A First Generation Nanofactory: An Architecture for Mechanosynthesis

Ralph C. Merkle
Senior Research Fellow
Institute for Molecular Manufacturing, IMM.org
President, Nanofactory Corporation

Cite as: Merkle RC. A First Generation Nanofactory: An Architecture for Mechanosynthesis. IMM Report No. 65, 2026; http://www.imm.org/Reports/rep065.pdf.

1 Foreword

The architecture of a first generation nanofactory is the focus of this paper. Progress towards that objective can be accelerated by creating a larger community to pursue it. The sections herein were written at various times between 2008 and 2026 as part of a privately funded internal research effort. They have been edited to unify the presentation. The various sections of this paper cover disparate aspects of a common theme: how to build a nanofactory starting from our existing technology base.

The author would like to express his strongest possible gratitude to Doug Arends, who had the vision and the courage to support this research when that support was needed most. His support has accelerated the development of the first nanofactory by at least a decade if not more.

The author owes particular thanks to his wife, Carol Shaw, for her patient support during the time this work required. He is also grateful to the many people who have commented on and contributed to this work over the years, but remains solely responsible for any flaws or errors. The greatest thanks are owed to Robert A. Freitas Jr. Many others have contributed, including (alphabetically by first name): Alex Vierich, Damian Allis, David Forrest, David Lister, Dusan Vobornik, James Lewis, James MacArthur, James "Jim" Robinson, James Ryley, Jeremy Barton, John Turriff, Khalil Sayed-Akhmad, Luis Sandoval Andrade, Mark Jobes, Mathieu Durand, Matt Moses, Michael Meakin, Sheldon Haird, Tad Hogg, Tait Takatani, Tom McKendree, Tom Moore, Vaibhav Thakore, and Xuran Zhu.

November 5, 2026 was the comment cutoff for this release. Further comments remain welcome and may inform corrections, revisions, separate publications, or other follow-up work.

2 Abstract

Manufactured products are made from atoms. The properties of those products depend on the arrangement of those atoms. The ideal manufacturing system would inexpensively arrange atoms in most of the ways permitted by physics and chemistry. Every atom and every bond between atoms would be there by design. This should enable atomically precise computers, computer memories, and quantum computers with capabilities orders of magnitude beyond those projected for conventional approaches, as well as atomically precise medical nanorobots able to heal and cure in cases well beyond the projected reach of conventional medicine.

For a development program, we define a nanofactory as a manufacturing system able to make most well-defined atomically precise products subject to certain limitations, including product size, rigidity, stability (both of the product and its intermediate build states during the manufacturing process), elemental composition, production rate (atoms per second or atoms per day), and power consumption (watts).

This paper proposes an architecture for such a first generation nanofactory and analyzes selected components. Importantly, even a first generation nanofactory could make more and better nanofactories. A critical question is: how do we build a first generation nanofactory? What relatively simple and limited nanofactory could serve as today's research goal yet also be sufficient to start a virtuous cycle of exponentially increasing capability and exponentially declining manufacturing cost in future generations of nanofactories?

The baseline nanofactory proposed herein targets $\sim 5 \times 10^{21}$ atoms of atomically precise product per day, ~100 milligrams of carbon equivalent product. It has an estimated power requirement (including cooling) of 325 kW. It would assemble products from atomically precise blocks of up to 1,000 atoms made from H, B, C, N, O, Si, Ge, and Pt.

Contents

1 Foreword **2**

2 Abstract **3**

3 Introduction **14**

3.1 The research program to make the baseline nanofactory . . . 17
3.2 How to build a nanofactory . . . 18
3.3 The baseline nanofactory: an overview . . . 19
3.3.1 Assumptions and the 2026 baseline . . . 20
3.3.2 Summary of Critical Parameters . . . 20
3.3.3 SRU Concept . . . 22
3.3.4 SRUs on a die . . . 22
3.4 Internal conceptual structure . . . 25
3.4.1 Mechanosynthesis . . . 25
3.4.2 SRU fabricators . . . 25
3.4.3 AP Blocks . . . 25
3.4.4 Bootstrapping, workstations and fabricators . . . 26
3.5 Why diamondoid? . . . 28
3.5.1 Diamond is the best for making almost anything . . . 28
3.5.2 Diamondoid is more than diamond . . . 28
3.5.3 Mechanosynthesis can make AP Blocks . . . 30
3.5.4 The 20 picometer requirement . . . 30
3.5.5 The reason for screw drives . . . 31
3.5.6 Advantages of mechanosynthesis . . . 31
3.6 Requirements for feedstock . . . 31
3.7 Molecular Tools . . . 32
3.7.1 Terminology for molecular tools . . . 33
3.7.2 A short reaction sequence . . . 37
3.7.3 Stability of molecular tools . . . 38
3.7.4 Some principles of mechanosynthesis . . . 39
3.7.5 Error rates in mechanosynthesis . . . 42
3.8 What we might manufacture . . . 45
3.9 Summary . . . 45

4 The Simple Replicating Unit or SRU **47**

4.1 Introduction . . . 47
4.1.1 SRU layout . . . 47

4.2 Shell . . . 48
4.3 Control . . . 50
4.3.1 Contact pillars . . . 50
4.3.2 Feet . . . 52
4.3.3 Wires . . . 52
4.3.4 Differential comb drives . . . 53
4.3.5 Bowden cables . . . 54
4.3.6 The demux . . . 54
4.4 Fabricator . . . 57
4.4.1 Stewart platform . . . 57
4.4.2 Programmable fabricator . . . 59
4.4.3 Assembly operations within the fabricator . . . 60
4.4.4 Molecular Tools . . . 60
4.4.5 Binding sites . . . 60
4.5 Balancing element intake and outflow . . . 61
4.6 Using platinum to generate H_2 gas . . . 62
4.7 Assembly . . . 64
4.7.1 Internal assembly arm . . . 64
4.8 Daughter Placement . . . 64
4.8.1 An SRU extension mechanism . . . 64
4.8.2 External Assembly Arm (EAA) . . . 65
4.8.3 Extruding the daughter . . . 66
4.8.4 Placing the daughter on the die . . . 68

5 Shell issues **69**
5.1 SRU Shell . . . 69
5.1.1 Blocks with dipolar links on top and bottom . . . 69
5.1.2 Blocks with dipolar links on four sides . . . 69
5.1.3 Hexagonal blocks with dipolar links on six sides . . . 69
5.1.4 Other linking groups . . . 70
5.1.5 Summary . . . 70

6 Control issues **71**
6.1 Broadcast Architecture . . . 71
6.1.1 Introduction . . . 71
6.1.2 The Instruction Set . . . 71
6.1.3 Stepper S (Select) . . . 72
6.1.4 Stepper D (Drive) . . . 72

6.1.5 RESET . . . 73
6.1.6 Turning the SRU ON or OFF . . . 73
6.1.7 Turn SRU ON . . . 73
6.1.8 IF <Test Unit> Turn SRU OFF . . . 74
6.1.9 Computing NOT . . . 76
6.1.10 Computing AND . . . 77
6.1.11 Universal Boolean Computing . . . 77
6.1.12 Enabling Simple Closed Loop Mechanosynthesis . . . 77
6.1.13 Summary of Instructions . . . 77
6.2 SRU Failure Handling . . . 78
6.2.1 Failures caused by radiation damage . . . 78
6.2.2 The target failure rate . . . 78
6.2.3 Fail-stop . . . 79
6.2.4 Errors during mechanosynthesis . . . 79
6.2.5 Simple failure model . . . 80
6.2.6 Letting fail-stopped SRUs accumulate: the baseline . . . 80
6.3 SRU Serial Numbers . . . 80
6.3.1 Scratchpad memory . . . 81
6.3.2 Communication between neighboring SRUs . . . 81
6.4 A Local Replication and Routing Algorithm . . . 82
6.4.1 Generations . . . 83
6.4.2 Synchronous Flag Update . . . 83
6.4.3 Parallel Daughter Actions . . . 84
6.4.4 Direct Placement . . . 84
6.4.5 Directed Transfer . . . 84
6.4.6 Persistent Daughter Motion . . . 85
6.4.7 Random Diffusion Transfer . . . 85
6.4.8 Local Conflict Resolution . . . 86
6.4.9 Time-Step Summary . . . 86
6.5 The Simulator . . . 87
6.6 10% fail-stopped SRUs on a die should be tolerable . . . 95
6.7 Convergent assembly and error mitigation . . . 97
6.8 Miscellaneous observations: . . . 98
6.9 Simple Closed Loop Mechanosynthesis . . . 100
6.9.1 Abstract . . . 100
6.9.2 Hysteresis and backlash . . . 101
6.9.3 An Example: Abstracting a Hydrogen . . . 101

6.9.4 Disabling the Fabricator . . . 103
6.9.5 Summary . . . 103
6.10 Build rate and replication time for an SRU . . . 103
6.11 Power for a daughter . . . 106
6.11.1 Umbilicals . . . 106

7 Fabricator issues **109**
7.1 Molecular composition of the feedstock . . . 109
7.1.1 Elemental and functional design . . . 109
7.1.2 Challenges *in* stabilizing the system . . . 109
7.1.3 Behavior of this mixture at 210 K . . . 112
7.1.4 Overall stability assessment . . . 112
7.2 Binding site leakage . . . 112
7.3 AP Blocks and Fabricators . . . 114
7.4 Decomposing large structures into small blocks . . . 117
7.4.1 Decomposing diamond into octahedrons and tetrahedrons . . . 118
7.4.2 Decomposing a flat diamond sheet into blocks . . . 119
7.4.3 A third decomposition of a flat diamond sheet into blocks . . . 120
7.4.4 Decomposition of hexagonal diamond into hexagonal blocks . . . 120
7.4.5 Decomposition into graphene strips . . . 123
7.4.6 Decomposition into coronene-like blocks . . . 124
7.4.7 Decomposition into parallelograms . . . 124
7.4.8 Decomposition into graphene parallelograms . . . 127
7.4.9 Decomposition into H-stabilized BN blocks . . . 127
7.4.10 Decomposition into w-BN blocks . . . 128
7.4.11 Decomposition into h-BN blocks . . . 132
7.4.12 A shell without interblock chemical bonds . . . 134
7.4.13 On the relative merits of B-N bonds versus C-C bonds . . . 137
7.5 Simple AP Blocks using B and N . . . 139
7.6 Molecular LEGO . . . 141
7.6.1 Introduction . . . 141
7.6.2 Terminology . . . 141
7.6.3 The various ways in which LEGO bricks can be connected . . . 141
7.6.4 LEGO walls . . . 142
7.6.5 LEGO floors and ceilings . . . 142
7.6.6 Crystal Structure . . . 146
7.6.7 Molecular LEGO blocks using interfacial Ge . . . 147
7.6.8 An example: an SRU shell from LEGO bricks . . . 147

7.6.9 SRU shell from Type A and Type B LEGO bricks . . . 149
7.6.10 Justification for the use of A and B blocks . . . 149
7.6.11 Extending the standard 2 by 4 LEGO brick . . . 150
7.6.12 Questions of interest . . . 150
7.7 Exemplar BN Molecular LEGO Blocks . . . 151
7.8 How many molecular BN LEGO Blocks in an SRU shell? . . . 153
7.9 BNO Molecular LEGO blocks that bond on four sides . . . 155
7.10 Molecular Tinkertoys . . . 157
7.10.1 Assembling space frames . . . 163
7.10.2 Other chiralities . . . 163
7.11 Fusing AP Blocks in the fabricator . . . 164
7.11.1 A question of terminology . . . 166
7.12 Graphene sheets from coronene . . . 166
7.13 Simple AP Blocks using Ge . . . 166
7.14 Stewart Platform Struts . . . 173
7.14.1 Buckling-Limited Force in the Stewart Platform Struts . . . 173
7.14.2 Idealized Strut Model . . . 174
7.14.3 Application to the Stewart Platform . . . 175
7.14.4 Two-Piece Threaded Strut Geometry . . . 176
7.14.5 Governing Buckling Equations . . . 176
7.14.6 Modeling of a strut . . . 176
7.14.7 Tube section properties (both sleeve and rod) . . . 179
7.14.8 Longitudinal stiffness k_s (N/m) . . . 179
7.14.9 Impact on axial stiffness of using AP Blocks for struts . . . 180
7.14.10 Number of carbon atoms in a block . . . 182
7.14.11 Helical structures from overlapping blocks . . . 182
7.15 Differential Electrostatic Comb Drives for All Actuation . . . 183
7.15.1 Overview . . . 183
7.15.2 Balancing forces: the differential comb drive . . . 183
7.15.3 Differential comb drives have constant capacitance . . . 184
7.15.4 Baseline design for a differential electrostatic comb drive . . . 184
7.15.5 The differential comb drive for the SRU . . . 184
7.15.6 Some System Level Observations . . . 196
7.16 H-trees for delivering power to the differential comb drives . . . 197
7.16.1 Capacitance of the H tree wiring on a die . . . 199
7.16.2 Power required to run the nanofactory . . . 200
7.16.3 LC delays in an H tree . . . 200

7.16.4 RC delays in an H tree . . . 201
7.17 Driving a demux with four binary Bowden cables . . . 203
7.18 Driving a screw drive with four-phase logic . . . 204
7.18.1 Rotary motion from linear motion . . . 204
7.18.2 The baseline: worm gears . . . 206
7.18.3 Summary . . . 206
7.19 Helical structures . . . 207
7.19.1 Artificial helical structures and screw drives . . . 207
7.19.2 How to design screw drives . . . 207
7.19.3 An example screw drive . . . 211

8 Die issues **216**
8.1 SRU-Die Interactions . . . 216
8.1.1 External electrical connections to the die . . . 216
8.1.2 Feet . . . 216
8.1.3 The first SRU . . . 217
8.1.4 Subsequent SRUs . . . 217
8.1.5 Questions . . . 217

9 Bigger than an SRU **219**
9.1 How to build 2 µm products . . . 219
9.1.1 Assembly in vacuum . . . 220
9.2 From 2D SRUs to 3D Space Frames . . . 221
9.2.1 Introduction to space frames . . . 221
9.2.2 Transitioning from surface SRUs to space frames . . . 226
9.2.3 Concepts for building larger structures . . . 229
9.3 Building more product . . . 230

10 Critical Parameters **232**
10.1 Productivity . . . 232
10.2 Build Rate . . . 232
10.3 SRU Count . . . 232
10.4 SRU site size . . . 232
10.5 Die Size and Count . . . 232
10.6 PC Board Size and Count . . . 233
10.7 SRU energy usage . . . 233
10.7.1 Comb drive throw and force: 2 nm and 1 nanonewton . . . 235
10.7.2 A global perspective . . . 237

10.7.3 Other energy losses . . . 237
10.7.4 Power requirements of the baseline nanofactory . . . 237
10.8 Cooling the nanofactory . . . 238

11 Baseline Implementation ("2026 Baseline") 240
11.1 Core Architecture . . . 240
11.2 Risks and Mitigation Strategies . . . 241
11.2.1 Baseline metabolism . . . 243
11.2.2 Positional compliance budget . . . 244
11.2.3 Strut design . . . 244
11.2.4 Comb drive force . . . 244
11.2.5 Feedstock mutual solubility . . . 244

12 Risk Mitigations 245
12.1 Self-alignment of the SRU in 6DOF . . . 245
12.2 Generating H_2 using AzAdam . . . 245
12.2.1 Hydrogen donation with AzAdam-H+ . . . 246
12.2.2 Ionizing AzAdam . . . 247
12.2.3 AzAdam+ on diamond . . . 247
12.2.4 Mechanosynthetic generation of H_2 . . . 247
12.2.5 Electrochemistry infrastructure . . . 251
12.2.6 Voltage boosting . . . 252
12.2.7 Sizing the voltage-boost capacitors . . . 254
12.2.8 Summary of generation of H_2 using AzAdam . . . 255
12.3 Disposal of excess gases other than venting . . . 256
12.3.1 Disposal of excess gas in containers . . . 256
12.3.2 Internal storage of waste gas . . . 257
12.3.3 More general balance . . . 257
12.4 Another Way of Generating Rotary Motion . . . 257
12.5 Split ring compound planetary gear . . . 259
12.6 Split ring planetary gear . . . 260
12.7 Reversed Harmonic Drive . . . 262
12.8 Using two indexers to rotate a gear . . . 263
12.9 Design for a vacuum dielectric differential electrostatic comb drive . . . 266
12.9.1 Classic comb drives . . . 266
12.9.2 Direct tunneling between vacuum-separated graphene electrodes . . . 268
12.9.3 Notation . . . 270
12.9.4 The design . . . 270

12.9.5 Parameter values . . . 271
12.9.6 Electrostatic equations and force . . . 271
12.9.7 Electrostatically active area of four differential comb drives . . . 273
12.9.8 Capacitance of differential comb drives . . . 273
12.9.9 Electronic time constant and heat . . . 273
12.9.10 Summary of vacuum electrostatic differential comb drive . . . 274
12.10 Power for a daughter . . . 274
12.10.1 Batteries . . . 274
12.11 Extrusion Mechanism . . . 275
12.11.1 Staggered pegs . . . 275
12.11.2 The MacArthur Mechanism . . . 275
12.12 Sliding Seal Extrusion Risk Mitigation . . . 276
12.12.1 Analysis of sliding seal risk when using hexagonal blocks . . . 276
12.12.2 Other candidate blocks . . . 278
12.12.3 First strategy . . . 279
12.12.4 Second strategy . . . 280
12.12.5 Third Strategy . . . 280
12.12.6 Fourth Strategy . . . 280
12.12.7 Fifth Strategy . . . 282
12.12.8 Conclusion . . . 285

13 Opportunities **286**
13.1 Future research in electromechanosynthetic (EMS) reactions . . . 286
13.1.1 “Donating” and “abstracting” electrons . . . 286
13.2 The positional control of electrons . . . 287
13.3 Additional concepts for building 2 µm products . . . 289
13.3.1 Abstract . . . 289
13.3.2 Introduction . . . 289
13.3.3 Vacuum Assembly in a Product Assembly Chamber (PAC) . . . 292
13.3.4 Concept 2 . . . 294
13.3.5 Concept 6: Unitary product fabrication with kiloSRUs . . . 296
13.4 Coordinates for each daughter . . . 296
13.5 Faster methods of populating dies . . . 297
13.5.1 Strategy 1 . . . 297
13.5.2 Strategy 2 . . . 299
13.5.3 Strategy 3 . . . 299
13.5.4 Summary . . . 300
13.6 Nested nanotubes . . . 300
13.7 Use of diamond as a dielectric in the comb drive . . . 302

14 Task list for SRU development **303**
14.1 Comb drives . . . 303
14.2 Wires to drive comb drives . . . 303
14.3 Build volume of SRU fabricator . . . 304
14.4 A demux driven by binary Bowden cables (or equivalent) . . . 304
14.5 Mechanical means of distributing power and control driven by a demux . . . 305
14.6 Conversion of linear to rotary motion . . . 305
14.7 Positional mechanosynthetic systems . . . 305
14.8 Assembly tools . . . 306
14.9 Enumerate a subset of assembly operations sufficient for assembling the SRU . . . 306
14.10 Describe feedback mechanisms to control SRU operations . . . 307
14.11 Determine the error rate of the SRU fabricator . . . 307
14.12 A shell . . . 307
14.13 Binding sites . . . 308
14.14 A programmable SRU fabricator/metabolizer . . . 308
14.15 Assembly arms . . . 308
14.16 Design electrical contacts between the SRU and the contact pillars . . . 309
14.17 Extrusion and placement of the daughter SRU by its parent . . . 309
14.18 Internal layout of SRU . . . 309
14.19 Removal of test items from workstation into liquid environment . . . 310
14.20 Removal of First SRU from workstation . . . 310
14.21 Design and Manufacture of Dies . . . 310
14.22 Design feedback systems from multiple SRUs on die to central computer . . . 311
14.23 Design of external support subsystems . . . 311
14.24 Various other things . . . 311
14.25 SRU replication and systems integration . . . 312

15 References **313**

16 Glossary **323**

17 Symbols, Abbreviations, and Units **334**
17.1 Principal symbols . . . 334
17.2 One-letter prefixes used before units . . . 338
17.3 Units . . . 338
17.4 Physical Constants . . . 339
17.5 Useful Constants . . . 339
17.6 Conversion factors . . . 340

18 Appendices **341**
18.1 Feedstock discussion . . . 341
18.2 Units used in this paper . . . 342
18.3 The force exerted by atmospheric pressure at the nanoscale . . . 342
18.4 Force on daughter from fluid flow . . . 342
18.4.1 Near-surface fluid velocity . . . 343
18.4.2 Viscous drag on the daughter . . . 343
18.4.3 Drag caused by deliberate movement of the daughter . . . 344
18.4.4 Conclusion . . . 344
18.5 Buckyballs as a storage method for hydrogen gas . . . 345
18.5.1 Summary . . . 345
18.5.2 Analysis . . . 345
18.5.3 Number of carbon atoms in the 15 nm cage . . . 346
18.5.4 Hydrogen capacity at 210 K . . . 346
18.5.5 Mechanical integrity . . . 346
18.5.6 Idealized size limit . . . 347

3 Introduction

> "If not us, who? If not now, when?"
> Variously attributed

The concept of programmable atomically precise manufacturing systems able to make copies of themselves was first advanced by Drexler [Dre92, Dre86], who combined the work of von Neumann [vN66] and Feynman [Fey60]. Von Neumann said self-replicating programmable general purpose manufacturing systems were possible, and Feynman said it was possible to arrange atoms in most of the ways permitted by physical law.

Today, we call such systems "nanofactories". The design space of possible nanofactories was reviewed by Freitas and Merkle [FM04b].

Previous proposals for nanofactories focused on either (a) feasibility, arguing that nanofactories were possible, or (b) relatively mature designs, arguing that nanofactories would be valuable. This paper proposes an architecture for a first-generation nanofactory and develops selected portions of that architecture to the level of quantitative engineering design. While still theoretical, it could serve as the starting point for a research program to develop a first generation nanofactory.

A research program to develop a nanofactory would ideally reach one of two conclusions. The first would be to demonstrate a nanofactory that could manufacture some very broad class[1] of atomically precise products. The second would be a proof that nanofactories are impossible. Today we have neither.

There is evidence (both experimental and theoretical) supporting the hypothesis that nanofactories are possible [Fey60, Dre92, FM04b, FM08a, FM04a, Fre09, CBN]. Recent experimental work has demonstrated mechanosynthetic reactions using inverted-mode STM (Scanning Tunneling Microscope) [B+25], including atomically precise C_2 donation on hydrogenated Si(100) [C+26], mechanosynthetic carbon donation and silicon abstraction on Si(100) [B+26], and the fabrication of non-planar molecular tools [H+25].

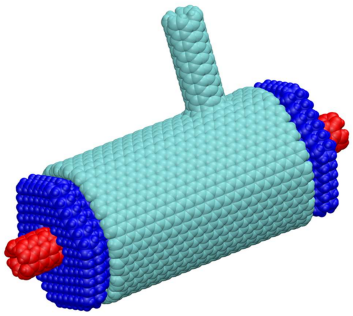
Further, there is reason to believe some of the products a nanofactory could make would be very valuable [MFH+18, HMA17, Dre92, Fre99, MFA25, MFR+19]. For example, theoretical analyses of atomically precise molecular mechanical switches suggest that their energy-delay product (EDP) might be eight orders of magnitude better than that of existing field-effect transistors (FETs), such as those in NVIDIA H100-class GPUs, and three orders of magnitude better than as-yet-theoretical atomically precise FETs [MFH+18, HMA17, MFA25]. Another attractive application of a nanofactory would be the fabrication of atomically precise quantum computers based on individually positioned atomically precise quantum circuit elements. In these architectures, the placement of each atom must be controlled with atomic precision because quantum interactions depend sensitively on position. Atomically precise manufacturing could therefore substantially simplify the construction of quantum computers [B+21, Kan98, E+25]. Atomically precise medical nanorobots able to treat cases that would be considered hopeless even by foreseeable extensions of conventional medicine have been proposed and analyzed theoretically [Fre99].

Moore's law began 60 years ago. Today it has slowed but remains exponential. Unfortunately, Moore's second law, also called Rock's law, is alive and well: semiconductor fab costs continue to

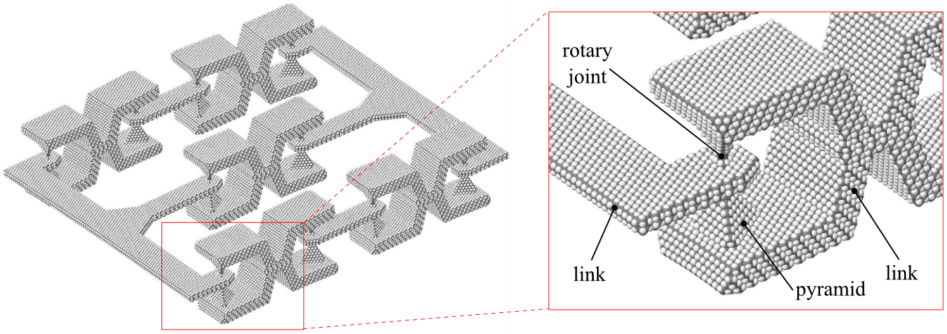
[1] We here introduce the concept that any given manufacturing system can manufacture a specific class of products. Just as a ribosome can manufacture any protein that satisfies certain constraints, such as its permitted amino-acid composition and maximum length before errors make manufacture prohibitive, any specific nanofactory might have limitations on the elements that it could incorporate, the size of object it could make, its minimum stiffness, the types of bonds it could make, and perhaps other limitations.

rise [Moo97]. Individual fabs already cost tens of billions of dollars. The number of semiconductor companies has dropped sharply in recent decades. "There are only three major semiconductor manufacturers of logic chips: TSMC, Samsung and Intel" [Cen24]. There is only one supplier in the world for extreme ultraviolet (EUV) lithography machines: ASML [ASM].

The logic elements shown in Figure 1 illustrate the potential of atomically precise manufacturing.



(a) A molecular field-effect transistor (mFET) (image from [MFA25]).



(b) A molecular mechanical logic element (lock) (image from [MFH+18]).

Figure 1: Molecular models of logic elements that we cannot make today but that are predicted to offer performance orders of magnitude better than is currently achievable.

Self-replicating manufacturing systems based on Atomically Precise Manufacturing (APM) have the potential to make computers with performance improved by multiple orders of magnitude relative to that projected for conventional computing technologies [MFH+18, HMA17, MFA25], while simultaneously reducing both capital cost and cost per logic element by multiple orders of magnitude.

Negligible capital cost is a consequence of a self-replicating manufacturing technology. Once the entire nanofactory can be manufactured by an existing nanofactory using only low cost (bulk chemical) inputs and energy, it follows that, in sufficiently high volume, manufacturing costs of such a nanofactory will decline until they approach the cost of the raw materials and energy required in the manufacturing process. The present baseline proposal does not meet this goal, as it has focused on being manufacturable from existing technology. It uses expensive inputs (such as semiconductors made with existing technology constrained by existing economics and manufacturing costs) and has poor energy efficiency. However, the baseline proposal would be sufficient to start the virtuous cycle that leads to future nanofactories able to meet this requirement. Moore's law would then be able

to continue using new logic technologies and new manufacturing technologies. These new self-replicating manufacturing technologies should eliminate the tens of billions of dollars in capital costs we now spend on fabs.

Many other applications could be improved by orders of magnitude following the introduction of low-cost atomically precise products. Improvements in medical capabilities following the introduction of medical nanorobots would be a particularly impactful area [Fre99].

If successful, the research program described here would result in nanofactories able to continue Moore's law for additional decades. A nanofactory could make very valuable products. Even a slow and expensive nanofactory could make more and better nanofactories. Given the extreme potential value of a nanofactory, such a research program would be worthwhile given even a modest probability of success.

Table 1: Nanofactory payoff matrix.

	We try	**We don't try**
We can do it	**Enormous payoff**	Nothing changes
We can't do it	Lose investment	Nothing changes

Correct use of the matrix in Table 1 requires an evaluation of the risk of failure. We make three claims about a funded program to develop a nanofactory:

1. The program has a low enough risk to justify it, given the high payoff.
2. The prospective program certainly justifies a lower cost effort to evaluate the risk.
3. The program could ramp up as the evaluated risk declines.

The reader is asked to make their own assessment of these claims.

There are already some high-level arguments that suggest the risk might not be as high as some might fear. For example, birds demonstrated that heavier-than-air flight was possible before we built airplanes. This provided evidence that the risk involved in a program to build airplanes was not sky high. Likewise, biology demonstrates that complex systems of molecular machines can inexpensively make more such systems. Potatoes are complex systems of biological molecular machines, yet they cost ~$1/kg because a potato can make more potatoes. This is a compact argument that at least some classes of complex atomically precise replicating manufacturing systems are both feasible and inexpensive.

Potatoes are based on ribosomes. A ribosome is a programmable molecular machine able to directly make a wide class of proteins and indirectly make complex biological systems such as cells and potatoes.

Similarly, the fabricator is a proposed programmable molecular machine which, if successfully developed, would be able to directly make a wide class of AP (Atomically Precise) Blocks and indirectly make complex molecular systems such as Simple Replicating Units (SRUs), powerful computers, and medical nanorobots. Biology has already demonstrated that a programmable capability able to directly make any selected member of a class of atomically precise objects can support a flexible,

low-cost manufacturing system. We seek to apply this concept to the class of diamondoid objects. Table 2 compares the corresponding components.

Table 2: A comparison of the core components of biology with the proposed components of a nanofactory.

	Biology	**Nanofactory**
Core synthetic device	Ribosome	Fabricator
Directly manufactures	Proteins	AP Blocks
Replicating unit	Cell	Simple Replicating Unit (SRU)
Indirectly manufactures	All living systems	Computers, medical nanorobots, and more

3.1 The research program to make the baseline nanofactory

The research program to make the baseline nanofactory proceeds in stages. Stage One is to master mechanosynthesis [H^+25, B^+25, C^+26, B^+26]. Once we can both (i) manufacture atomically precise blocks (AP Blocks) with perhaps hundreds of atoms in them, and (ii) assemble them (conceptually, we might imagine that this assembly process is similar to assembling LEGO bricks), we can proceed to Stage Two: developing and assembling a fabricator from AP Blocks. Once we complete Stage Two and have working fabricators, we can proceed to Stage Three: manufacturing enough AP Blocks to make the first Simple Replicating Unit (SRU). Once we complete Stage Three and have the first SRU, we can proceed to Stage Four: scale up. In Stage Four, the first SRU is placed on a site on a die, which provides power, instructions, and feedstock. That first SRU then replicates to occupy the billion or so pre-prepared sites on the surface of the die. Populating more dies by repeating this process completes the manufacture of the baseline nanofactory.

At the same time that we are developing our physical ability to make and assemble AP Blocks, we must also develop our design capabilities so that, once we have the physical ability, we will know which AP Blocks to make and have designs for the fabricator and the SRU based on those AP Blocks. Both the fabricator and the SRU are made from AP Blocks and are atomically precise. This should facilitate their design, analysis, and modeling.

Table 3 summarizes the four stages the research program might go through.

Table 3: Stages of the research program to develop the baseline nanofactory.

	Experimental work	**Advance preparation for next stage**
Stage One	Master Mechanosynthesis	Design the fabricator
Stage Two	Build a fabricator	Design the SRU
Stage Three	Build an SRU	Prepare SRU support environment
Stage Four	Scale up	Design products

Table 4 summarizes the atom count and build rate of the various systems involved in the manufacture of the baseline nanofactory. We start with the workstation, which builds the workstation

fabricator, which builds the first SRU, which then replicates to build more SRUs. In an ideal scenario, after 46 generations the first SRU and its descendants can manufacture the 5.8×10^{13} SRUs (mounted on 72,000 dies) in the baseline nanofactory.

Table 4: Atom counts and build rates of systems used to manufacture the baseline nanofactory.

	Workstation	**Workstation fabricator**	**First SRU**	**More SRUs**	**Baseline nanofactory**
Atoms	Many	10 million	100 million	100 million each	SRU atoms: 5.8×10^{21}
Makes	Workstation fabricator	First SRU	More SRUs	Baseline nanofactory	Products
Build rate	1 atom per second	10 to 100 atoms per second	1,000 atoms per second	1,000 atoms per second each	5.8×10^{16} atoms per second
Time to build target	116 days	12 to 116 days	28-hour replication time	46 generations (a few months)	Target dependent

This paper focuses primarily on the high level design of an early SRU. It touches on other issues.

3.2 How to build a nanofactory

The general algorithm for designing and building a nanofactory is remarkably simple.

1) Pick the most credible plan for building a nanofactory currently available.
2) Review the plan for risks.
3) Pick the risk you think is worst.
4) Resolve the risk. This resolution doesn't have to be perfect, but it should be good enough to make the selected risk of less concern than other risks in the plan.
5) Repeat.

Step (4) might involve analyzing a previously unexplored risk in greater detail to determine whether the existing plan works. If the existing plan assumes something that might not be true, investigate the assumption in depth. If the assumption appears to be true, explain that more clearly. If the assumption appears to be false, modify the plan by altering or eliminating the false assumption. If the assumption cannot be resolved to a satisfactory level of confidence, modify the plan by altering or eliminating the unresolvable assumption.

If you find the present plan to be of interest, start by looking for problems. Pick the worst problem and work on it. This work can either be experimental (if you've got access to the right equipment) or theoretical. In either case, publish a paper that shows how to solve or avoid the problem. If you can't solve or avoid the problem, seek help from others. If it remains unsolvable and unavoidable, publish it.

Continued iteration of this algorithm will eventually (a) produce a nanofactory, (b) show that a nanofactory is impossible, or (c) iterate indefinitely. If you have concluded that (a) is the most likely outcome, then the next question is: what course of action do you plan on taking?

3.3 The baseline nanofactory: an overview

Any nanofactory can be divided into an atomically precise eutactic core and a non-atomically precise support environment. The eutactic core is that part of the nanofactory which is "eutactic": i.e., each atom is in a designed, intended position. In our baseline nanofactory the atomically precise eutactic core consists of 58 trillion identical atomically precise Simple Replicating Units, or SRUs. Each SRU has a vacuum tight shell separating the external non-eutactic environment from the atomically precise eutactic interior. This shell is 500 nm by 100 nm by 60 nm. Each SRU is composed of ∼100M atoms and can manufacture and extrude atomically precise eutactic products similar in size to itself. An SRU is itself one such product that an SRU can manufacture and extrude.

The primary component of the non-atomically precise support environment for the baseline nanofactory is 125 boards immersed in liquid feedstock. Each board carries 576 integrated-circuit (IC) packages. Each package holds one die, nominally 1 cm^2, with approximately 8.05×10^8 SRU sites. Each site provides electrical power and instructions as well as mechanical support for one SRU. Both power and instructions are broadcast to the SRUs through contact pillars on the die surfaces. Identical instructions are broadcast to all SRUs.

SRUs are either ON or OFF. Those SRUs that are ON execute instructions in lockstep with the broadcast. SRUs that are OFF do not carry out manufacturing operations, but continue to respond to instructions that turn them ON. This allows subsets of SRUs to carry out simple conditional instructions, including error recovery routines in response to common error conditions.

The SRUs can be redirected to manufacture different products by changing the broadcast. The liquid feedstock provides the elements required by the SRU to manufacture atomically precise products. These elements are provided in the form of molecules in solution. Binding sites embedded in the shell of each SRU can bind to and transport individual feedstock molecules from the external liquid environment into the internal environment.

To manufacture the first baseline nanofactory we start with a single die with no SRUs on it. We attach the first SRU to a site on this first die. This first SRU is manufactured by a "molecular workstation" (or simply "workstation"), which, in concept, is an advanced variant of existing Scanning Probe Microscopes (SPMs) [Fre09, Fre10, H+25, B+25, M+23, MPS03, FM12a, FM12b, C+26, B+26, HBMR00, CBN].

That first SRU replicates and places its daughter on the die. Those two SRUs replicate and place their two daughters on the die. Those four SRUs replicate and place their four daughters on the die. At some point, daughters have to be passed from parent SRU to parent SRU until they can be placed on the die. This exponential growth in SRUs repeats until the first die is populated with SRUs.

In the next stage, all SRUs on the first die replicate. Each resulting daughter SRU is held aloft by its parent to allow it to be picked up from the first die and moved to an empty die. That is, the daughters are individually removed and placed individually on the 72,000 dies required for the nanofactory. All 72,000 dies are then incubated until all of them are fully populated. At some point during this process, these dies are placed in 72,000 IC packages, and those packages are placed on 125 PC boards, 576 to a board. The boards are then inserted into their support environment. This completes the manufacturing process.

Clearly, this plan is based on assumptions. There is a risk that one or more of these assumptions will not support the program in the manner expected. If this happens, we say that the risk has been realized. If a risk is realized, we must adopt an alternative plan. Before committing a large

budget to the program, we should systematically review the program risks and quantify them to the extent possible. For each significant risk we should also have one or more mitigation strategies that would allow the program to continue if the risk is realized. The actual risk of programmatic failure should be reduced as much as possible before major resources have been committed.

As we mitigate risks, so too we can exploit opportunities. For example, if mechanosynthesis at room temperature proved feasible, we could exploit this opportunity and simplify the system design by eliminating the refrigeration system.

3.3.1 Assumptions and the 2026 baseline

The formal name for the set of assumptions we make for this proposal is "the 2026 baseline for the nanofactory". We do not claim the 2026 baseline has a 100% chance of being correct. We do claim that major risks to the 2026 baseline have been considered and mitigation strategies proposed for risks that have raised concern.

Further derisking is both possible and worthwhile. While the threshold for starting a program to develop a nanofactory can reasonably be discussed and debated, continued derisking should eventually meet or exceed the threshold requirements. Research to derisk one or more first generation nanofactory designs should be pursued.

If the reader identifies a significant risk that has not been considered, please let us know. We are more than happy to add it to our list.

The baseline design is the particular first-generation nanofactory design analyzed in this paper. Alternative designs are introduced where they reduce identified risks, offer major opportunities, or illuminate important tradeoffs. In essence, the baseline is the author's best tradeoff of risk versus reward. More capable and rewarding systems are possible but would be riskier or more difficult to implement. Less capable systems are less risky but would limit further development of better nanofactories. It is unlikely that any given reader will agree with all the tradeoffs made by the author, but this initial baseline provides a framework for more extensive analysis and discussion.

The word "should", in context, usually indicates that the statement is assumed to be true for purposes of the baseline design. Each such statement corresponds to a risk. If something that "should" be true proves to be untrue, the risk has been realized. In this case, the baseline would need to be updated. We do not provide specific mitigation strategies for every statement that "should" be true. At the time of writing, we believe that statements that "should" be true likely are true. Many such statements can be risk-reduced by computational modeling. Some statements can be risk reduced by appropriate experiments. Early efforts to risk-reduce key assumptions are appropriate. It is expected that risks will continue to be realized and updates to the baseline will continue to be made until a nanofactory is actually built and demonstrated.

3.3.2 Summary of Critical Parameters

This section provides a concise summary of the parameters describing the baseline architecture, along with a brief description. This allows the reader to see the overall structure of the baseline nanofactory. A more thorough discussion is given in the section on Critical Parameters (Section 10).

These are nominal design estimates. Production assumes full occupancy by functioning SRUs, without losses from product rejection or reduced utilization. The build rate and replication time remain conditional on the detailed manufacturing and error-recovery schedule. Power estimates exclude buffer-network dissipation and other unbudgeted support loads.

Product atoms per nanofactory-day: 5×10^{21}**.** Atoms per nanofactory-second × 24 hours/day × 60 minutes/hour × 60 seconds/minute. Chosen to be profitable and achievable. Used as a guide for other parameters.

Carbon equivalent product mass per nanofactory-day: 100 milligrams. Product atoms per nanofactory-day × Mass of carbon atom.

SRUs per die: 805,306,368 = $\mathbf{2^{14} \times 2^{14} \times 3}$. Three SRUs have a site size of 0.6 × 0.6 µm. $2^{14} \times 0.6$ µm ≈ 1 cm. This number facilitates use of H trees.

Atoms per die-second: 805,306,368,000. SRUs per die × Atoms per fabricator-second

Atoms per nanofactory-second: 5.798×10^{16}**.** Atoms per die-second × dies per nanofactory.

Atoms per fabricator-second: 1,000. Chosen to be feasible with early designs.

Atoms per SRU-second: 1,000. One fabricator per SRU, so the same rate as a fabricator.

SRUs per nanofactory: 5.8×10^{13}**.** SRUs per die × dies per nanofactory.

Actuation power budget per SRU: 1 nW. See Estimating SRU energy usage (Section 10.7).

Actuation power budget per Nanofactory: 58 kW. Actuation power budget per SRU × SRUs per nanofactory

Footprint of an SRU: 500 nm × 100 nm. Driven by internal design.

SRU site dimensions: 600 nm × 200 nm. Chosen to hold an SRU with adequate margin between adjacent SRUs.

SRU die area: nominal 1 cm^2. A typical laptop or desktop CPU die area. Actual area is likely to be larger because of supporting structures on the die.

SRUs per IC package: 805,306,368. Same as SRUs per die.

Board dimensions: 450 mm × 450 mm. (18 inch × 18 inch) Common "large board" in industry.

Packages per board: 576. 24 × 24 packages per board. Enough board area per package.

Dies per board: 576. Same as packages per board.

SRUs per board: $\sim 4.64 \times 10^{11}$**.** (SRUs per package) × (packages per board).

Boards per nanofactory: 125. Chosen to manufacture $\sim 5 \times 10^{21}$ product atoms per day, ∼100 mg of carbon equivalent product. Number of boards can be adjusted to produce desired product mass per day.

Packages per nanofactory: 72,000. Packages/board × Boards/nanofactory.

Dies per nanofactory: 72,000. Same as Packages per nanofactory.

Actuation power budget per die: 0.805 W. (Actuation power budget per SRU) × SRUs per die. The power required to move the molecular machinery in the SRUs on the die.

Actuation power budget per Board: 464 W. (Actuation power budget per die) × dies per board

Voltage on differential comb drive: 0.19 V. Larger comb drives require less voltage but occupy more space. Lower voltage reduces waste heat. Comb-drive size was increased until the volume budget was exhausted.

Clock speed: 100 MHz. Preliminary analysis supports feasibility.

First-order nominal heat per nanofactory: ∼157 kW. Based on actuation power budget per nanofactory of 58 kW (2.5 aJ mechanical work per SRU per throw) plus non-recovering

capacitive switching losses; excludes buffer-network dissipation. See Power required to run the nanofactory (Section 7.16.2)

Heat per board: 1,256 W. Heat per nanofactory / Boards per nanofactory

Heat per die: 2.18 W. Heat per nanofactory / Dies per nanofactory

Power to cool a nanofactory to 210 K: 168 kW. See Cooling the nanofactory (Section 10.8)

Estimated heat + refrigeration electrical power: ~325 kW. Heat per nanofactory + Power to cool a nanofactory to 210 K.

Actuation energy budget per SRU-cycle: 10 aJ. Actuation power budget per SRU / Clock speed.

Differential comb drive throws per clock cycle: 4

Actuation energy budget per differential comb drive throw: 2.5×10^{-18} **J.** Actuation energy budget per SRU-cycle / 4

Actuation k_BT budget per differential comb drive throw at 210 K: 862. Actuation energy budget per differential comb drive throw / k_BT

Clock cycles per atom: 10^5**.** See Build rate and replication time for an SRU (Section 6.10)

Atoms per SRU: 100 M. An approximation derived from early design work.

SRU replication time: 10^5 seconds. ~28 hours. (Atoms per SRU) / (Atoms per SRU-sec).

Radiation hit rate per SRU per replication cycle: 3.0×10^{-6}**.**

Target failure rate per SRU per replication cycle: 10^{-3}**.** Or 0.1%.

3.3.3 SRU Concept

SRU: **S**imple **R**eplicating **U**nit

An SRU sits on the silicon surface astride six contact pillars. The proposed SRU is 500 nm × 100 nm × 60 nm. The footprint of the SRU on the silicon surface is 500 nm × 100 nm. The parent SRU extrudes a daughter SRU upwards through a fluid-tight seal on the 500 nm × 100 nm face. The extrusion port on the parent is 100 nm × 60 nm, allowing it to fit easily on the 500 nm × 100 nm face. The parent sits on three pairs of contact pillars, six contact pillars in all. These contact pillars provide power, ground, and four 100 MHz broadcast instruction signals. Beneath the contact pillars is the surface of the silicon die.

Figure 2 shows an icon for a fabricator with its inputs (feedstock, instructions, power) and outputs (heat, waste, AP Blocks). The icon for the fabricator is reminiscent of a macroscopic scanning probe microscope shrunk to ~100 nm. The SRU has an SRU fabricator as an internal subsystem. The SRU fabricator wields four positionally controlled molecular tools in vacuum, which it uses to mechanosynthesize atomically precise blocks called (very creatively) Atomically Precise Blocks, or AP Blocks. The SRU fabricator is both made from AP Blocks and manufactures AP Blocks. The SRU as a whole is likewise assembled from AP Blocks. The SRU can assemble the AP Blocks that the SRU fabricator makes. A simplified mental model of an AP Block is a very small atomically precise LEGO brick.

3.3.4 SRUs on a die

Figure 3 shows one SRU on a silicon die immersed in feedstock. The icon for an SRU is the icon for a fabricator inside a black rectangle that symbolizes the vacuum tight shell.

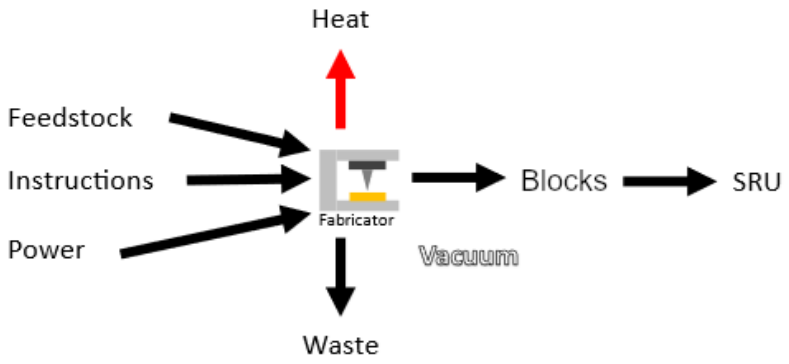


Figure 2: Conceptual fabricator inputs and outputs. Feedstock, instructions, and power support the manufacture of AP Blocks in vacuum, with heat and waste as outputs. The blocks are assembled into an SRU.

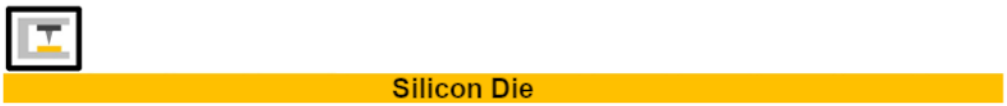


Figure 3: Initialization of a silicon die with a single SRU. The black rectangle represents the vacuum-tight shell surrounding its fabricator. The die provides the support environment for subsequent replication; the schematic is not to scale.

The silicon die in the cartoon is shown following its initialization with a single SRU.

Once the first SRU has been attached to the die surface, scale up proceeds by self-replication of the population of SRUs on the die surface. This population starts with one SRU and increases by doubling the number of SRUs at each replication cycle.

The first replication cycle starts when broadcast instructions to the first SRU instruct it to manufacture and extrude a daughter SRU, which it attaches to a site on the die surface adjacent to itself. At the end of the first replication cycle there are two SRUs attached to two sites on the die surface. In the second replication cycle these two SRUs produce two daughter SRUs that are then attached to sites on the die surface, for a total population at the end of the cycle of four SRUs. In the third replication cycle these four SRUs produce four daughters, bringing the total population to eight SRUs. Subsequent cycles continue this pattern, increasing the total population to 16, 32, 64, 128, and so forth. When necessary, daughter SRUs are passed hand-over-hand from one parent SRU to the next parent SRU on the die surface until the daughter SRUs can be attached to unoccupied (open) sites on the die surface. With a 10^5 second (~28 hour) SRU replication period and allowing for transportation overhead of daughter SRUs to the edge of the growth area, covering a die with SRUs might require a little more than one month. When the die is full (all sites are occupied), replication stops. Once the first die is full there will be more than enough SRUs to seed additional dies, which could then be populated in the same manner. Scale-up proceeds by populating the first die and then using that first die to manufacture additional SRUs that can be used to populate additional dies. Each die and its supporting environment are manufactured by existing conventional technology, while SRUs manufacture more SRUs using atomically precise mechanosynthesis.

Figure 4 schematically shows the growth of the SRU population on the silicon die toward full site occupancy.



Figure 4: Schematic population growth of SRUs on a silicon die. Replication expands the population from one SRU toward full site occupancy. The final row illustrates occupancy of the depicted region rather than an additional exact population doubling. Not to scale.

The rest of this paper addresses the major issues raised by this brief description and explores risks, mitigation strategies, and opportunities.

3.4 Internal conceptual structure

In this section we further define the internal conceptual structure of the SRU.

3.4.1 Mechanosynthesis

We define *mechanosynthesis*[2] as a method of manufacturing atomically precise products that uses positionally controlled molecular tools, particularly including radicals and other reactive species, in an inert environment. In this paper we use mechanosynthesis exclusively for the manufacture of AP Blocks from the elements H, B, C, N, O, Si, Ge and Pt. See Mechanosynthesis can make AP Blocks (Section 3.5.3).

AP Blocks and assemblies of AP Blocks are atomically precise products that are within the scope of this paper. Other methods of making atomically precise products are outside its scope, though obviously of great importance.

3.4.2 SRU fabricators

The fabricator inputs and outputs are shown in Figure 2.

We then define the baseline SRU fabricator as a subsystem in the SRU that is 100 nm × 100 nm × 60 nm in size, and which can carry out mechanosynthesis when embedded in a support environment that can:

i) Provide vacuum
ii) Provide instructions, primarily on how to position a set of molecular tools
iii) Provide power
iv) Provide feedstock
v) Remove waste materials
vi) Remove and assemble AP Blocks

The baseline SRU fabricator uses four Stewart platforms to provide six-degree-of-freedom control for each of four independently controlled molecular tools. The four Stewart platforms can pick up and put down individual molecular tools as they are needed while they build an atomically precise structure, much as two cooperating machinists might pick up and put down tools with their four hands while building a machine. These Stewart platforms also perform miscellaneous other mechanical functions.

3.4.3 AP Blocks

Fabricators make Atomically Precise Blocks, also known as **AP Blocks** or just "blocks" when context makes it clear what we are referring to.

Baseline AP Blocks:

[2] Drexler's definition [Dre92] of mechanosynthesis was "chemical synthesis controlled by mechanical systems operating with atomic-scale precision, enabling direct positional selection of reaction sites".

a) Are made from the elements H, B, C, N, O, Si, Ge, Pt.
b) Have 1,000 atoms or fewer
c) Are smaller than 15 nm in any dimension
d) Are rigid – no gases or liquids

The specific limits we have chosen are less significant than the fact that AP Blocks ***have*** limits. By imposing limits on the element palette, atom count, linear size, and type of structures an SRU fabricator can make, we can factor the problem of designing and building the first SRU into two largely independent problems: (1) designing the SRU from a specified class of AP Blocks and (2) physically bootstrapping the first SRU by building and assembling all the AP Blocks required for that first SRU.

If an SRU is assembled from AP Blocks, and AP Blocks are a relatively simple class of objects that can be manufactured by an SRU fabricator, then the problem of parts closure for an SRU has been converted into a design problem: design an SRU from AP Blocks. If an SRU is designed from AP Blocks, any AP Block can be made by the SRU, and the SRU can assemble AP Blocks, then an SRU can make another SRU of the same design.

3.4.4 Bootstrapping, workstations and fabricators

Bootstrapping consists of making and assembling all the AP Blocks for the first SRU. As the first SRU does not have the services of an already-existing SRU to manufacture and assemble the AP Blocks, we must have some other "bootstrap" method of making this first SRU. We do this by making all the AP Blocks for an SRU using a "workstation fabricator" and then assembling the resulting blocks with the workstation probe. See also AP Blocks and Fabricators (Section 7.3).

3.4.4.1 Definition of "Workstation"

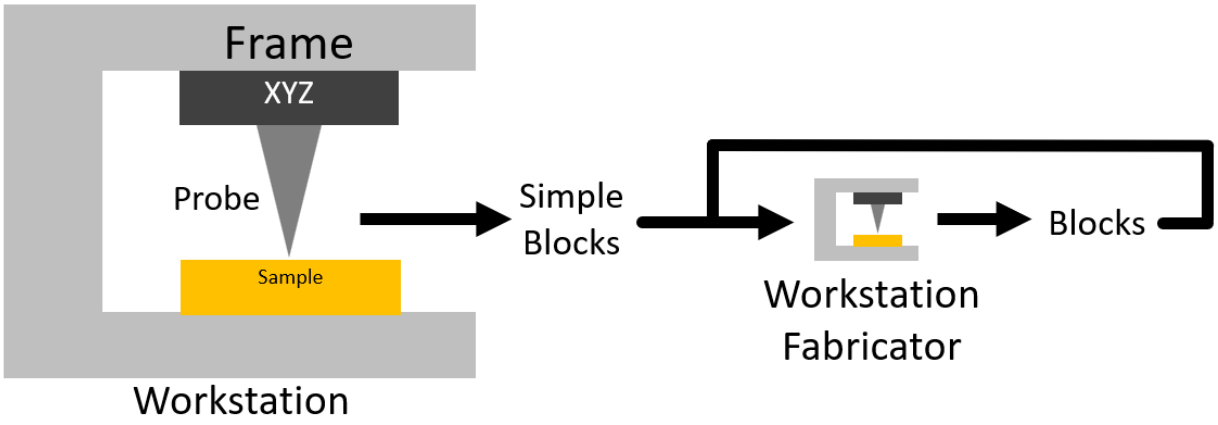


Figure 5: Bootstrapping a workstation fabricator from simple AP Blocks. The workstation makes and assembles the initial blocks into a workstation fabricator, which then manufactures further AP Blocks.

Figure 5 illustrates the relationship between a "workstation" and a "workstation fabricator".

A "workstation" is an advanced scanning probe microscope with enhancements to facilitate mechanosynthesis [Fre09, Fre10, H+25, CBN].

A "workstation fabricator" is a fabricator that resides within the workstation. The AP Blocks used for a workstation fabricator might be simpler than those used in the SRU fabricator. The AP Blocks from which the workstation fabricator is assembled must be made directly by the workstation. Assembly of the workstation fabricator is also done directly by the workstation.

The SRU fabricator, as a subsystem of the SRU, is made by assembling AP Blocks. The workstation fabricator is not a subsystem of the SRU. The workstation fabricator has different requirements than the SRU fabricator. The SRU fabricator must make all the AP Blocks for an SRU within the 10^5 second (∼28 hour) replication cycle of an SRU. The workstation fabricator must make all the AP Blocks for the first SRU within an "acceptable" bootstrap period, likely measured in months, and preferably not many months. That is, the workstation fabricator could be 100 times slower and could also be less reliable, as rejecting a higher percentage of AP Blocks because of imperfections during their mechanosynthesis is tolerable during the bootstrap of the first SRU.

The baseline SRU fabricator wields four molecular tools with six degrees of freedom each. This should be sufficient for the timely mechanosynthesis of all the AP Blocks for the SRU. One line of argument supporting this hypothesis is that wielding four tools with six degrees of freedom each was sufficient for all the reactions in *A Minimal Toolset for Positional Diamond Mechanosynthesis* [FM08a]. By contrast, the workstation fabricator can wield only one tool at a time. While this should be sufficient, more time would be spent positioning jigs, exchanging tools, and the like. It takes longer to build anything if you only have one arm.

The workstation fabricator itself is assembled from bootstrap AP Blocks (or simple AP Blocks or simple blocks). Bootstrap AP Blocks might be a subset of AP Blocks that are deliberately designed to be more limited in size, element types, and design complexity. We might, for example, limit bootstrap AP Blocks to 200 atoms, restrict them to a subset of our baseline palette (possibly just H, C and Ge) or impose design restrictions because of the limited capabilities of the workstation and the workstation fabricator. A future paper will focus on the design of the workstation fabricator and the nature of bootstrap AP Blocks. In this paper, we assume the workstation fabricator has been developed, that it can make blocks as specified, and that the blocks it makes can be assembled by the workstation.

The limitations on bootstrap AP Blocks arise because they must be manufactured directly by the workstation, which can wield only one molecular tool at a time in X, Y, and Z. While the workstation fabricator can more rapidly manufacture bootstrap AP Blocks, we must first bootstrap the workstation fabricator itself by building and assembling bootstrap AP Blocks using only the workstation. Once we have successfully bootstrapped the first workstation fabricator, we can use it to more rapidly build bootstrap AP Blocks for the second and succeeding workstation fabricators. However, all workstation fabricators must still be assembled using the workstation probe. That is, the workstation fabricator, while having achieved parts closure and many of the benefits of replication by virtue of being able to make the bootstrap AP Blocks from which it itself is made, is not yet able to achieve all the benefits of a fully self-replicating manufacturing system. For that, we must manufacture the first SRU and support its successful replication.

Any reasonably well defined and limited class of AP Blocks which is in fact compatible with the successful design of an SRU would (a) let us address the parts closure problem and (b) let us deal with the bootstrap problem for an SRU by making and assembling the AP Blocks for the first SRU by special means, and then having the first SRU and succeeding SRUs make and assemble the AP Blocks for all following SRUs.

The focus of this paper is on the SRU, its design and functioning. We only briefly mention the bootstrapping process for manufacturing the first SRU. For a 2008 perspective on what a molecular

workstation or a workstation fabricator might look like, the reader is referred to [FM08b, Fre09, Fre10]. More recent work is [H+25, B+25, C+26, B+26].

Most manufactured products would benefit from being assembled from AP Blocks. A few that would be of immense value are: (a) computers, (b) medical nanorobots, and (c) nanofactories.

See The Simple Replicating Unit or SRU (Section 4) for a more detailed discussion of the SRU.

In the following sections we examine the physical structure of the SRU and its support environment and examine the logical assumptions which result in that physical structure.

3.5 Why diamondoid?

3.5.1 Diamond is the best for making almost anything

Of particular interest in the work of Drexler is the focus on diamond and "diamondoid" materials.[3] Diamond is a materials scientist's dream come true, with properties that are the best or close to the best in many categories as shown in Table 5 [Uni, Appa]. Bending experiments on diamond nanoneedles, interpreted using finite-element modeling, indicate local tensile stresses at fracture up to 125 GPa [N+19]. The table lists a supplier-reported dielectric strength of 1 V/nm. The calculated intrinsic breakdown field of bulk diamond is approximately 2.4 V/nm [SBR12]. "Diamondoid" is a natural extension of "diamond" that includes more elements.

The ability to manufacture diamondoid products, to shape them in atomically precise detail, and to terminate their surfaces with the desired elements would be enormously valuable. Immensely powerful molecular computers [MFH+18, HMA17, Dre92, MFA25, Mer93b, MFR+19] and medical nanorobots able to heal and cure in cases that would today be considered hopeless [Fre99] would all become feasible.

3.5.2 Diamondoid is more than diamond

The term "diamondoid" includes a broader class of atomically precise structures than just diamond. It includes "strong, stiff structures containing dense, three dimensional networks of covalent bonds, formed chiefly from first and second row atoms with a valence of three or more" [Dre92]. In this paper, "diamondoid" includes strong, stiff, atomically precise structures containing dense networks of covalent bonds, including graphene structures and carbon nanotubes. This usage includes AP Blocks made from H, B, C, N, O, Si, Ge, and Pt. While Pt is not generally considered "diamondoid", we include it for convenience in terminology to avoid frequent references to "diamondoid and Pt". In the SRU, Pt is primarily used for its catalytic properties, rather than being used for structural purposes. In more general usage, "diamondoid" may also include other elements. It is worth emphasizing that mechanosynthesis is applicable to almost the entire periodic table, including metals such as Pt. Even noble-gas atoms can be positioned individually at cryogenic temperatures [ES90].

[3] As defined in Nanosystems [Dre92], diamondoid structures are "...strong, stiff structures containing dense, three dimensional networks of covalent bonds, formed chiefly from first and second row atoms with a valence of three or more. Many of the most useful diamondoid structures will in fact be rich in tetrahedrally coordinated carbon." Materials made from B, C and N have particularly high tensile strength; examples include diamond, lonsdaleite, graphene, SWNTs, and boron nitride variants thereof.

Table 5: Selected properties of diamond [Uni, Appa].

Property	**Value**	**Units**
Hardness	10,000	kgf/mm^2
Strength, tensile	> 1.2	GPa
Strength, compressive	> 110	GPa
Sound velocity	18,000	m/s
Density	3.52	g/cm^3
Young's modulus	1220	GPa
Poisson's ratio	0.2	Dimensionless
Thermal expansion coefficient	0.0000011	/K
Thermal conductivity	20.0	W/(cm·K)
Thermal shock parameter	30,000,000	W/m
Debye temperature	2,200	K
Dielectric constant	5.7	Dimensionless
Dielectric strength	10,000,000	V/cm
Electron mobility	2,200	$cm^2/(V·s)$
Hole mobility	1,600	$cm^2/(V·s)$
Electron saturated velocity	27,000,000	cm/s
Hole saturated velocity	10,000,000	cm/s
Electron affinity	small and negative on H-(111) surface	eV
Bandgap	5.45	eV
Resistivity	10^{13} - 10^{16}	Ohm-cm

3.5.3 Mechanosynthesis can make AP Blocks

Mechanosynthesis is a method of manufacturing atomically precise products that uses positionally controlled molecular tools, particularly including radicals and other reactive species in an inert environment (see Mechanosynthesis (Section 3.4.1)). We focus on the mechanosynthesis of atomically precise AP Blocks made from our baseline palette. We start by focusing on that subset of diamondoid structures known as diamond.

Chemical vapor deposition (CVD), which relies heavily on radical chemistry within a reduced-pressure reaction chamber, shows experimentally that radical chemistry can be used to synthesize diamond. In the CVD process, hydrocarbons like methane are introduced into a reduced-pressure reaction chamber alongside hydrogen gas. Energy is applied to this mix to form radicals, for example by creating a plasma.

While CVD demonstrates the ability to synthesize diamond, it does not demonstrate the atomically precise control required to assemble specific desired structures. CVD has a large random component that makes the synthesis of any particular specified diamond structure improbable.

Reactive molecular tools are at the heart of mechanosynthesis in a manner analogous to the highly reactive species that drive CVD. For example, methane dissociates into $\bullet CH_3$ and other radical species, which are the building blocks for diamond formation. These radicals facilitate the attachment of carbon atoms to the diamond lattice [BMC+09]. By positionally controlling the highly reactive species, we can reduce and possibly eliminate the randomness in the synthetic process.

Mechanosynthesis can in principle employ positionally controlled reactive species beyond those used in diamond CVD. Here, we focus on mechanosynthesis of a subset of the elements, and radicals thereof. Radicals of H, B, C, N, O, Si, Ge, and Pt can all come into play in the mechanosynthesis of AP Blocks.

While any inert environment could in principle be used (such as He or Ne), and possibly semi-inert environments (such as N_2), the baseline fabricator uses vacuum, which simplifies the analysis and clarifies the concepts. The baseline also uses only eight elements, rigid structures, and limited size. The mechanosynthetic approach should be extensible beyond the baseline palette and beyond diamond and diamondoid structures. How far it might be extended we leave to the future.

The use of positionally controlled reactive molecular tools in vacuum to synthesize diamond and its seemingly infinite and enormously valuable variants would seem to be an attractive route to "arrange the atoms one by one the way we want them" [Fey60]. Computational and experimental work support the hypothesis that mechanosynthesis is feasible and could be used to make a truly vast range of atomically precise diamondoid products of our choosing [Dre92, FM08a, Fre09, Fre10, B+25, M+23, FM12a, FM12b, C+26, HBMR00, Zho21, CBN].

3.5.4 The 20 picometer requirement

If positionally controlled molecular tools can be positioned to within 20 pm, it should be sufficient to reliably synthesize a very wide range of stable diamondoid structures. While many mechanosynthetic reactions could be reliably carried out by molecular tools positioned with accuracies of 50 or even 100 pm, it is expected that there will be reactions of interest that might be unreliable when carried out with these reduced accuracies. Reference [PFM04] reports positional tolerances of approximately 65.87 pm and 53.77 pm for placing a dimer at the global-minimum two-dimer configuration (GM2) while avoiding the nearest transition state. Figure 8 of [PFM04] also shows a smaller distance of 20.24 pm between LM2-2 and that transition state. We therefore adopt 20

pm as a conservative baseline intended to accommodate unusually demanding structures, including purposeful construction near closely spaced transition boundaries.

3.5.5 The reason for screw drives

Having adopted mechanosynthesis, which fundamentally depends on positionally controlled molecular tools, and a baseline positional accuracy for those tools of 20 pm, we use screw drives to achieve that accuracy. While screw drives are only one of many methods of achieving positional control, for purposes of simplicity and expositional clarity we have adopted screw drives as sufficient, relatively simple to design, and easy to explain. Other methods are left as either mitigations or opportunities.

3.5.6 Advantages of mechanosynthesis

Mechanosynthesis, the use of positionally controlled molecular tools (including radicals) in an inert environment such as vacuum, is an approach to synthesizing diamondoid products with the desired atomically precise control over the synthesis process while also addressing several system level issues. If we assume radicals are necessary to synthesize diamondoid structures, then it would seem difficult to use self-assembly or solution-based methods without creating uncontrolled radical-radical interactions. Radicals in solution would react with each other, with other chemicals in solution and with the solvent in an uncontrolled fashion.

There are alternatives to mechanosynthesis for the synthesis of diamond. For example, it is possible to dissolve carbon at high temperatures and pressures in metals or use detonation methods. If we use these methods, the necessary control of the synthetic process becomes problematic. If we do not use radicals, then the synthetic routes become more limited and complex, and the range of products appears to fall far short of what we desire. Diels–Alder and related reactions offer the possibility of synthesizing at least some range of hydrocarbon structures in solution, but the range of possible products appears to fall astronomically short of the goal.

We also impose the system constraint that a small atomically precise system be able to manufacture another small atomically precise system. We are excluding from our solution space those system proposals that require large and expensive systems to manufacture small and atomically precise systems. That is, we wish to avoid the current situation in the semiconductor industry where new fabs each cost tens of billions of dollars.

Given our current state of knowledge, the open path towards the goal is through positionally controlled molecular tools that use radicals: that is, through mechanosynthesis.

3.6 Requirements for feedstock

The system design proposed herein requires liquid feedstock, or at least non-solid feedstock. Earlier proposals (unpublished) using solid feedstock were both more complex and more difficult to bootstrap. Liquid feedstock systems can be much simpler. Having once decided on liquid feedstock, we are then left with the selection of specific feedstock molecules. We want the lowest temperature consistent with other requirements, as lower temperatures result in lower error rates from thermal noise. The melting points of individual candidate low temperature feedstock molecules, such as methane, diborane, nitrogen, oxygen, silane, and germane, are all at least 15 K below 125 K. However, we also require that the feedstock mixture remain stable as a mixture both at its operating temperature and during any earlier processing, possibly at higher temperatures. The inclusion of

oxygen is problematic in this regard, as it reacts with most of the other components. System designs that keep feedstock molecules separated from each other are again more complex and difficult to bootstrap. It would also be highly desirable if the feedstock did not impose significant safety issues for handling. Diborane, for example, is one of the most dangerous industrial gases. Exposure limits are measured in parts per billion. Handling exposes personnel to undesirable risks. Severe injury may not become apparent until hours after exposure.

It would also be desirable to avoid gases at the chosen operating temperature, as we both wish to avoid elevated operating pressures and also to maintain high volumetric concentrations to maintain intake rates and simplify binding site design.

Although the baseline temperature of 210 K, rather than the approximately 125 K discussed above, significantly increases error rates, low temperatures limit the pool of possible liquid feedstock molecules, while higher temperatures expand that pool, simplifying resolution of the issues outlined above. These considerations led to the current baseline feedstock (see Molecular composition of the feedstock (Section 7.1)) and operating temperature. A final determination that 210 K allows an acceptable compromise between competing issues awaits a more complete system design, including specification of all relevant reaction sequences and analysis of their error rates. The current baseline allows further design and analysis to move forward.

3.7 Molecular Tools

Freitas and Merkle [FM08a] proposed a set of nine positionally controlled radical molecular tools that could synthesize a subset of diamondoid structures and the self-same set of radical molecular tools from feedstock molecules in a vacuum environment, thus laying the groundwork for the present proposal.

It is important to understand, first, that the reactions proposed in [FM08a] (see the example in Section 3.7.2) do not take place in solution but instead take place in Ultra High Vacuum (UHV). Many of the highly reactive radicals used in the reaction sequences could immediately react with solvent molecules if these reactions took place in solution. Therefore, the proposal is to *not* use radicals in solution. If there is no solvent, the radicals cannot react with the (non-existent) solvent.

Second, these reactions use positionally controlled molecular tools. At a minimum, the molecular tools must always be controlled in their X, Y and Z positions, and at least held rigidly in roll, pitch and yaw. The highly reactive molecular tools cannot be allowed to contact each other or any surface except when this is explicitly called for by the reaction sequence. The HAbst tool, for example, was deliberately chosen to be one of the most highly reactive molecular species in existence. It must be held and positioned carefully, just as a hot soldering iron must be held and positioned carefully.

3.7.1 Terminology for molecular tools

Some clarifications in terminology are helpful. When names were proposed, molecular tools had specific functions and were to be "charged" and "discharged" [FM08a]. HDon donated hydrogen and was in either a charged state (with hydrogen) or discharged state (without hydrogen). Experience has generated new uses for old tools. Having a single name for two specific molecular tip structures and distinguishing between them by calling them "charged" and "discharged" has proven less flexible and more cumbersome in cases other than those for which they had originally been envisaged. This has led to a preference for naming tools based on their structure, rather than their function. For example, the term EAOGe-C2• has been used recently to refer to a specific molecule with **E**thene legs, an **A**damantane core, **O**xygen feet, and a **Ge**rmanium bridgehead bonded to a **C2** radical (Figure 6) [B$^+$25, C$^+$26]. This convention is appropriate for molecules with feet and legs but is less applicable when referring to molecular tools that lack feet and legs, as in both the SRU and the fabricator. It also sidesteps the fact that we need a name that refers to the molecular tip structure even if the supporting structure might change. That is, "a" molecular tool is actually a class of structures all of which have the same structure for the tip of the molecular tool, but need have little if anything in common for parts of the molecular structure more than a few bond lengths from the tip.

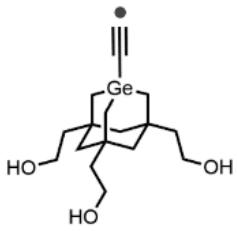


Figure 6: Molecular structure of EAOGe-C2•, with an adamantane core, ethene legs, oxygen feet, and a germanium bridgehead bonded to a C2 radical [B$^+$25, C$^+$26].

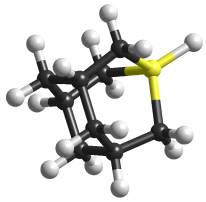
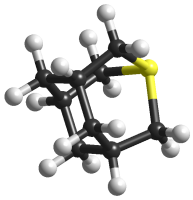
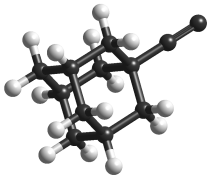
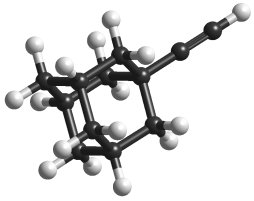
We update the earlier terminology by adopting the convention that the name of a molecular tool specifies one and only one molecular structure at the tooltip. One name, one tooltip structure. The charged and discharged states of a tool are subsumed under the more general concept of a tool having a payload, as in GeRad, GeRad-H, GeRad-CH2, GeRad-CC, or GeRad-CC-H. In this way, the state of a base tool plus payload (if any) can be specified in a way which is both more flexible and more easily understood. We define four base tools: AdamRad, SiRad, GeRad, and HAbst. These names are chosen to create the minimal disruption to existing naming conventions while still complying with the "one name, one tooltip structure" requirement. A purist might argue that HAbst should be named AdamRad-CC, but in deference to its widespread use we retain this name to specify the same structure as AdamRad-CC. The "discharged" HAbst is called HAbst-H or AdamRad-CC-H. Referring to the charged and discharged states is deprecated usage.

Even though each name refers to one tooltip structure, it's worth noting that each name refers to a class of molecular structures. The base of the tool is not specified by the name, only the tooltip. In the case of EAOGe-C2•, the name refers to the entire molecule including the base. EAOGe-C2• is included in the class of molecular tools called GeRad-CC. That is, the structure of the tool beyond the tip is not part of our naming system; what is included is the tooltip, which is an adamantane with a substitutional germanium at the bridgehead carbon and two carbons bonded to that bridgehead germanium.

The resulting names and usage are tabulated in Table 6. HDon now refers specifically to GeRad-

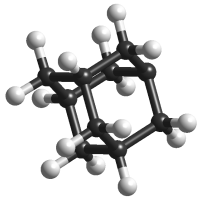
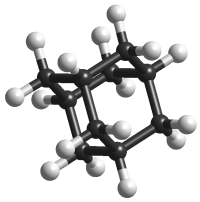
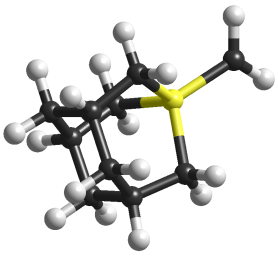
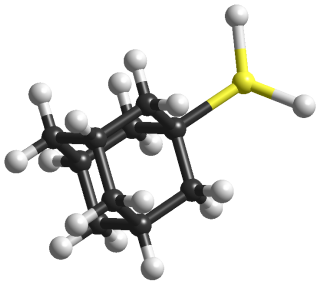
H, previously called the charged HDon tool. What was previously called the discharged HDon tool is called GeRad. HDon should not be used to refer to GeRad. HTrans refers specifically to HAbst-H(GeRad), previously called the charged HTrans tool, and not to anything else. What was previously called the discharged HTrans tool is now called HAbst-GeRad, more clearly showing its composition. The name HAbst now refers to exactly one tooltip molecular structure: AdamRad-CC.

Table 6: Molecular tool nomenclature and corresponding tooltip structures.

New name(s)	**Former name**	**Tooltip structure**
GeRad-H	GeRad-H HDon (charged) (deprecated)	1-germaadamantane; 
GeRad	HDon (discharged) (deprecated) HDon handle (deprecated)	1-germaadamantan-1-yl radical; 
HAbst **AdamRad-CC**	HAbst	2-(adamantan-1-yl)ethyn-1-yl; 1-adamantylethynyl radical; 
HAbst-H **AdamRad-CC-H**	HAbst (discharged) (deprecated)	1-ethynyladamantane; 

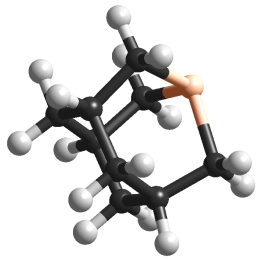
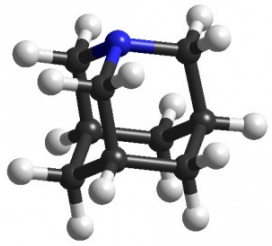
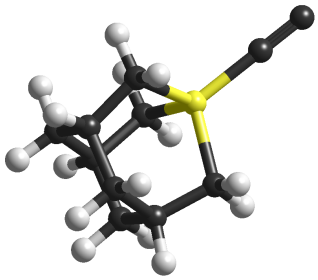
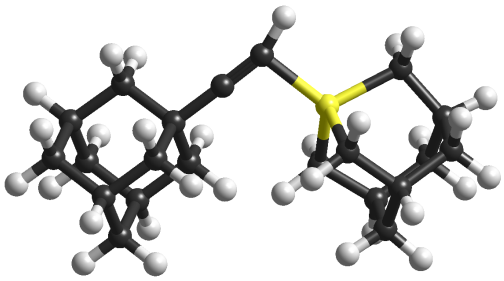
Continued on next page

Table 6 (continued)

New name(s)	Former name	Tooltip structure
AdamRad	AdamRad	adamantan-1-yl radical; 
AdamRad-H	AdamRad (discharged) (deprecated) Adam (deprecated)	adamantane; 
GeRad-CH2	Germylmethylene tool (deprecated)	
AdamRad-GeH2	Germylene tool (deprecated)	

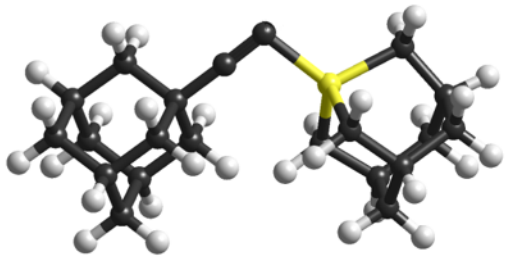
Continued on next page

Table 6 (continued)

New name(s)	Former name	Tooltip structure
SiRad	SiRad	1-silaadamantan-1-yl radical 
AzAdam		1-azaadamantane 
GeRad-CC	CC donation tool (deprecated)	2-(1-germaadamantan-1-yl)ethyn-1-yl radical 
HAbst-H(GeRad)	HTrans (charged) (deprecated)	

Continued on next page

Table 6 (continued)

New name(s)	Former name	Tooltip structure
HAbst-GeRad	HTrans (discharged) (deprecated)	

3.7.2 A short reaction sequence

Mechanosynthesis of diamond consists of sequences of tool applications, called reaction sequences. This example illustrates both a reaction sequence useful in fabricating atomically precise diamond structures and, by proxy, the concepts that can be used to mechanosynthetically make the wider class of diamondoid structures. An example reaction sequence (with updated names) dating from [FM08a] (or before) is shown in Figure 7.

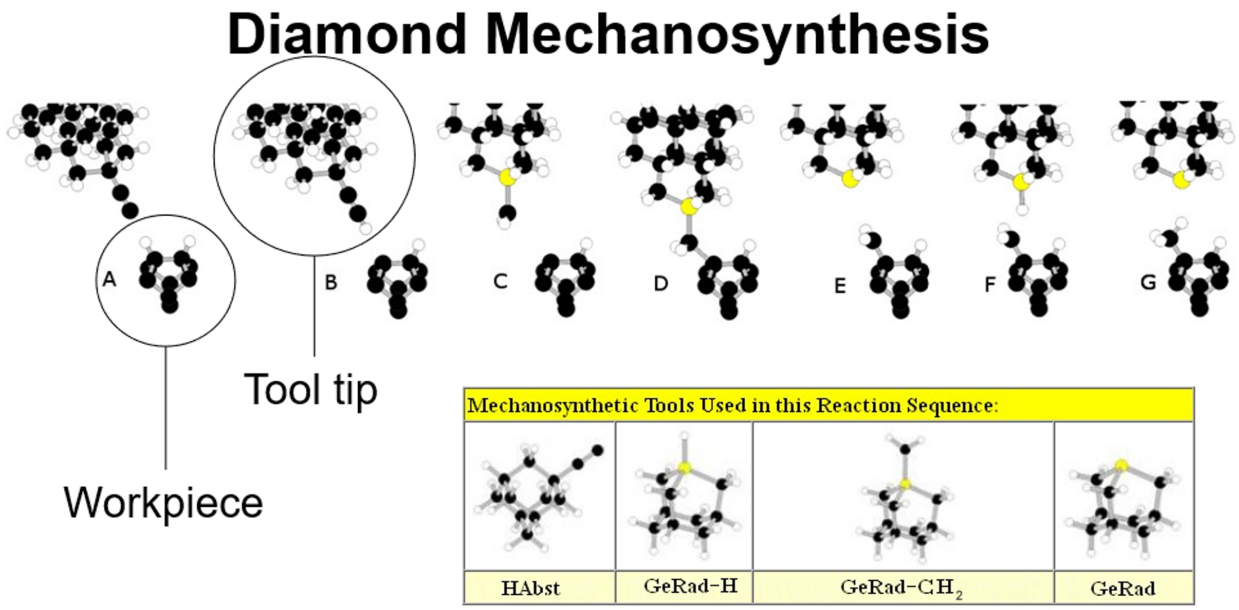


Figure 7: Example diamond mechanosynthesis sequence, with updated tool names [FM08a]. Panels A–G show the replacement of a surface hydrogen with a methyl group using HAbst, GeRad-CH2, and GeRad-H. GeRad is the tool remaining after payload transfer. The depicted workpiece represents part of a hydrogenated diamond (100) surface.

Note that the "Workpiece" shown in Figure 7 should be viewed as a small piece of the hydrogenated diamond (100) surface, not as an isolated (and quite reactive) molecular fragment. The different

tooltips would be brought up to the point on the surface in sequence to bring about the desired change to the chosen site on the surface. By successively applying a specific sequence of molecular tools to a sequence of specific surface sites, the diamond surface can be modified in a desired fashion to produce a desired atomically precise result. This is mechanosynthesis, and the specific sequence of tool applications is referred to as a "reaction sequence".

As can be seen in this example, replacing a single hydrogen on the hydrogenated diamond (100) surface with a CH_3 (methyl) group requires a reaction sequence that involves three tools and four reactions (counting C → D and D → E as two reactions).

In the first step, we remove a single selected hydrogen atom from the hydrogenated diamond surface. To do this, we position the distal carbon of HAbst adjacent to the hydrogen we desire to remove. Hydrogen forms a stronger bond to HAbst than it does to the surface of the diamond. To quote Temelso et al. [TSMF06]: "Modeling the diamond C(111) surface with isobutane and treating the ethynyl radical as a tooltip, hydrogen abstraction in this reaction is predicted to be barrierless". Therefore, the selected hydrogen is readily abstracted from the surface and bonds to the HAbst molecular tool, as shown in Figure 7, panels A–B.

Having abstracted the selected hydrogen atom, we remove the HAbst-H tool, leaving a carbon radical (a dangling bond) on the diamond surface. This dangling bond is highly reactive. As we carry out this reaction in UHV (Ultra High Vacuum), there is nothing for the carbon radical to react with, so it doesn't.

In the next step, we bring up the GeRad-CH2 tool, which has a $\bullet CH_2$ radical weakly bonded to germanium. The $\bullet CH_2$ radical will readily bond to the carbon radical on the diamond surface. This reaction, a radical-radical coupling, has zero barrier. This reaction takes place between steps C and D.

In the next step, we pull back on the GeRad-CH2 tool, stretching the Ge-C bond. As we pull the GeRad-CH2 tool away from the site, something is going to break. As the Ge-C bond is weaker than the C-C bond, the Ge-C bond breaks, leaving a $\bullet CH_2$ radical bonded to the diamond surface. This reaction is shown between steps D and E. Once the Ge-C bond is broken, the GeRad tool, with the $\bullet CH_2$ radical gone, is removed.

In the next step, we bring up GeRad-H to donate a hydrogen. GeRad-H has a hydrogen weakly bonded to a germanium. When the hydrogen is placed adjacent to the $\bullet CH_2$ radical bonded to the diamond surface, the hydrogen is bound less strongly to the germanium than it would be to the carbon. The reaction barrier is ∼43 zJ (0.27 eV) [TSMF07] when no force is applied. The hydrogen should transfer from the germanium to the carbon, as shown between steps F and G, if force is applied to lower the reaction barrier.

Finally, in step G we have arrived at the desired end state: the selective addition of a CH_3 to a specific site on the hydrogenated diamond surface.

Freitas and Merkle [FM08a] proposed 9 molecular tools using 65 reaction sequences and 328 reaction steps and analyzed 354 unique pathological side reactions in their proposed metabolism for H, C and Ge. For a production metabolism of eight elements, additional molecular tools, reaction steps and analysis will be required.

3.7.3 Stability of molecular tools

The molecular tools discussed in [FM08a] were analyzed for stability at room temperature. Barrier heights to reconstruction were all at or above (usually well above) 270 zJ (1.69 eV). By comparison,

thermal noise at room temperature is 4.14 zJ (0.0259 eV). For a barrier of 270 zJ and assuming an attempt rate of 10^{13}/second, the half life at room temperature (300 K) would be ~40 million years. The risk that these specific molecular tools might themselves be unstable is low.

3.7.4 Some principles of mechanosynthesis

1) If you pull on both ends of a bond hard enough, it will break, usually creating radicals.
    a. If you don't get what you want, pull on something different.
2) If two atoms are bonded, use two radicals to attack both atoms to break the bond.
3) When two radicals are brought together, they bond.
4) Hydrogen is a capping group.
    a. Add it using a weak radical.
    b. Remove it using a strong radical.
    c. Remove it from the strong radical by using two weak radicals (see Principle 2).
5) If two hydrogenated atoms are adjacent, you can bond them by removing the hydrogens.
6) If you want to use mechanosynthesis to make something:[4]
    a. Use these principles to create many potential mechanosynthetic reaction sequences for your target.
    b. Throw out the ones that don't work.

To expand on Principle 6 a bit, we might use the A* algorithm [HNR68, RN20] over the space of stable AP Blocks to search through reaction sequences starting from "the fabricator with an empty work space" and ending in "the fabricator with the desired AP Block in the work space". Transitions from state to state correspond to either (1) application of a specific molecular tool to some specific position on the workpiece to modify the workpiece or (2) spontaneous transitions in the workpiece caused by thermal noise, vacuum contaminants or radiation damage.

Application of a molecular tool ideally transitions to a single resultant state with 100% probability but might transition to two or more resultant states each with less than 100% probability. If a transition caused by a molecular tool to an undesired state has an unacceptably high probability, either (a) testing for the actual outcome can be used as part of a conditional sequence to bring the system back to an acceptable state with an acceptably high probability or (b) the proposed application of the molecular tool in that reaction sequence cannot be used.

Spontaneous transitions (caused by thermal noise, radiation, or contaminants) must either be such that it is possible to test for them and use a conditional reaction sequence to remediate them, or have a low enough probability that the SRU replication cycle still has a high enough probability of success that the system as a whole will continue to function. See SRU Failure Handling (Section 6.2).

[4] This general approach is called retrosynthetic analysis [Cor67, Cor88] when used with conventional chemical synthesis. It is similar to Drexler's "backward chaining" [Dre92], John Anderson's "Horizon Mission Methodology" [And93, And96], "meet in the middle" attacks in cryptography, and the classic A* algorithm [HNR68, RN20] used in computer science for finding the shortest path between two nodes in a graph.

3.7.4.1 Hydrogen abstraction and HAbst

The positionally controlled abstraction of a hydrogen atom from a diamondoid structure is among the oldest and best-studied elementary operations proposed for diamondoid mechanosynthesis (DMS). The canonical tool for this operation is a single ethynyl (acetylenyl) radical, $\cdot C \equiv CH$, mounted on a rigid molecular "handle." Proposed by Drexler in Nanosystems [Dre92], it was rapidly followed[5] by the first ab initio study by Musgrave et al. [MPMG91]. High-level coupled-cluster benchmarks were carried out by Temelso et al. [TSMF06, TSMF07].

Its use in a system context was part of the comprehensive minimal-toolset analysis by Freitas and Merkle, who called it HAbst [FM08a]. Tarasov et al. studied HAbst recharge in 2010 [TAI+10] and hydrogen donation in 2013 [TIA+13]. An independent high level kinetics determination was done by Bowman, Burke, Turney, and Schaefer [BBTS20]. Finally, CBN Nano Technologies used inverted-mode scanning tunneling microscopy for the first experimental demonstrations [H+25, B+25, C+26, B+26].

HAbst satisfies a critical set of design requirements—very high hydrogen affinity, steric accessibility, attachability to a positionable handle, mechanical and chemical stability, selectivity, and ease of regeneration (recharge). The choice of the ethynyl tip is a direct consequence of the first requirement: the terminal C–H bond of acetylene is among the strongest bonds to hydrogen known, so the radical formed by its removal has an exceptional affinity for hydrogen.

The abstraction tool is the prototype of subtractive mechanosynthesis generally. Recent experiments have not only removed hydrogen from a probe apex [B+25] but also removed silicon from Si(100) [B+26], using sub-angstrom positional control. HAbst can do more than just abstract hydrogen.

3.7.4.2 Hydrogen donation and GeRad-H

Hydrogen donation is the inverse of hydrogen abstraction, adding a single hydrogen atom to a chosen radical on a diamondoid workpiece. Site selective hydrogen donation was proposed abstractly by Drexler in *Nanosystems* [Dre92]. The use of germanium was validated quantitatively as a group-IV hydride donor by Temelso, Sherrill, Merkle and Freitas [TSMF07]. The tool most commonly associated with hydrogen donation is GeRad-H, proposed by Freitas and Merkle [FM08a] within their proposed process-complete minimal toolset. Tarasov et al. [TIA+13] analyzed the operating envelopes and optimal tool approach trajectories. This is a summary and tabulation of some donation energetics and the relevant bond dissociation energies (BDEs) in zeptojoules, with the original source value in parentheses (1 eV ≈ 160.22 zJ; 1 kcal mol^{-1} corresponds to approximately 6.948 zJ per molecule).

3.7.4.2.1 Energetics of Hydrogen Donation

Each donation is a radical H-transfer

$$\text{GeRad-H} + {}^{\bullet}\text{C (workpiece)} \rightarrow \text{GeRad} + \text{C–H(workpiece)} \tag{1}$$

Because the C–H bond formed is stronger than the Ge–H bond broken, every donation is exothermic. Table 7 gives the reaction energies for GeRad-H donating to representative depassivated sites, computed at B3LYP/6-311+G(2d,p) [FM08a].

[5] Although the publication of Nanosystems [Dre92] was nominally in 1992 and the publication by Musgrave et al. [MPMG91] was in 1991, drafts of Nanosystems were circulating prior to Musgrave's work.

Table 7: Donation reaction energies for GeRad-H [FM08a]. Negative values indicate exothermic reactions; zJ values are converted from the source eV.

Donation by GeRad-H to	**Reaction energy**
Dehydrogenated C(111) bridgehead	−98 zJ (−0.61 eV)
Adamantyl (AC) bridgehead radical	−96 zJ (−0.60 eV)
C(110) ridge site	−117 zJ (−0.73 eV)
C(100) dimer site	−130 zJ (−0.81 eV)
Phenyl (C_6H_5) radical	−194 zJ (−1.21 eV)

The underlying chemistry was benchmarked at CCSD(T) with correlation-consistent basis sets by Temelso et al. [TSMF07]. For the small models of Ge–H donating to a carbon radical (best converged basis): $\bullet CH_3 + GeH_4 \rightarrow CH_4 + \bullet GeH_3$ has a reaction energy of −163 zJ (−23.5 kcal mol^{-1}) over a barrier of 43 zJ (6.2 kcal mol^{-1}); $H\bullet + GeH_4 \rightarrow H_2 + \bullet GeH_3$ has a reaction energy of −142 zJ (−20.5 kcal mol^{-1}) over 23 zJ (3.3 kcal mol^{-1}). The closest small-molecule analogue of donation to a diamond C(111) site, $\bullet C(CH_3)_3 + HGe(CH_3)_3 \rightarrow HC(CH_3)_3 + \bullet Ge(CH_3)_3$, has an experimental enthalpy of −60 zJ (−8.7 kcal mol^{-1}), in general agreement with the −98 zJ B3LYP value for the full tool. Donation is thus modestly exothermic with a low, mechanically surmountable barrier.

3.7.4.2.2 Bond Dissociation Energies

Donation favorability is set by the difference between the C–H bond formed and the X–H bond broken; the relevant BDEs are collected in Table 8 (D_{298}). Values for the group-IV donors are anchored to the tertiary/bridgehead C–H bond (670 zJ (96.5 kcal mol^{-1})) and the reaction enthalpies of Temelso et al. [TSMF07].

Table 8: Representative bond dissociation energies. The high ≡C–H BDE makes HAbst a strong abstractor; the weaker Ge–H bond makes GeRad-H an effective donor that remains rechargeable.

Tool-H	**Bond (model compound)**	**BDE** D_{298}
HAbst-H	≡C–H (HC≡C–H)	924 zJ (133 kcal mol^{-1})
AdamRad-H	C–H (H–C(CH_3)$_3$)	670 zJ (96.5 kcal mol^{-1})
SiRad-H	Si–H (H–Si(CH_3)$_3$)	663 zJ (95.4 kcal mol^{-1})
GeRad-H	Ge–H (H–Ge(CH_3)$_3$)	610 zJ (87.8 kcal mol^{-1})
H-H	H-H	724 zJ (436.0 kJ mol^{-1}) [Dar70]
HAbst-H(GeRad)	C-H (HAbst-H(GeRad))	429 zJ (2.68 eV) [FM08a]

3.7.4.2.3 Status

Within the minimal toolset, GeRad-H and HAbst together manage all surface hydrogen, and the toolset achieves 100% process closure for the H, C, and Ge subset which it targets [FM08a], including recharge of GeRad back to GeRad-H (reaction sequence RS5, step 2) [FM08a]. GeRad-H also appears across the patent record — "Positional diamondoid mechanosynthesis" (US 8,171,568 B2;

US 8,276,211 B1) and "Build sequences for mechanosynthesis" (US 9,676,677 B2), Freitas and Merkle, assigned to CBN Nano Technologies [FM12a, FM12b, FM17]. The chemistry is favorable and the reaction network is closeable.

3.7.5 Error rates in mechanosynthesis

To quote from [FM08a]:

> Minimum acceptable reaction energies were guided by reaction reliability requirements. If the absolute value of an energy barrier or reaction exoergicity is E_{react}, then the Boltzmann population of the examined product structure may be approximated as $P_{\text{react}} \sim \exp(-E_{\text{react}}/(k_B T))$ with $k_B = 1.381 \times 10^{-23}$ J/K and T = temperature in K. At $T = 300$ K, a reaction exoergicity of $E_{\text{react}} = 0.40$ eV gives $P_{\text{react}} = 2 \times 10^{-7}$ for the failed reaction, reliable enough for products requiring $\lesssim 10^7$ reaction steps for fabrication (e.g., products containing $\sim 10^6$ heavy atoms, assuming $\sim$10 steps are required per heavy atom), at room temperature. At $T = 80$ K (LN_2 temperature), $E_{\text{react}} = 0.20$ eV gives $P_{\text{react}} = 3 \times 10^{-13}$ ($P_{\text{react}} = 5 \times 10^{-26}$ at 0.40 eV), indicating extreme reliability for fabricating products with very high atom counts at low temperature. Likewise, a barrier of $E_{\text{react}} = 0.40$ eV that blocks formation of an unwanted pathological structure along a second reaction pathway that competes with the desired more exoergic main reaction pathway gives $P_{\text{react}} = 2 \times 10^{-7}$ at $T = 300$ K for the unwanted structure, yielding similar reliability estimates as before. For these reasons, all proposed reactions have minimum energy barriers (blocking an undesired outcome) or relative exoergicities (favoring a desired outcome) of $\sim$0.40 eV for operation at 300 K and $\sim$0.20 eV for 80 K operation.

At our baseline temperature of 210 K, $k_B T$ is 2.9 zJ. The mechanosynthetic reactions in [FM08a] have a computed energy difference between the desired reaction and any erroneous reaction of $\geq$ 64 zJ (0.4 eV). This is 22 times thermal noise. Erroneous reactions would be expected to occur with a probability of $\sim e^{-22} \approx 2.8 \times 10^{-10}$ per reaction.

A typical mechanosynthetic reaction does not have an attempt frequency. Instead, it typically selects between competing potential reactions with different barrier heights. Consider two competing reactions. The probability of settling into the wrong minimum instead of the right minimum depends exponentially on the difference between the barrier heights. For example, an error that might occur when abstracting a particular hydrogen H_A would be to abstract the wrong hydrogen H_B. Once either hydrogen has been abstracted the tool + workpiece is in a sufficiently deep minimum that no more reactions would be expected in any relevant time frame.

Similarly, if we are attempting to bond a CH_2 to a radical site on a workpiece surface using the GeRad-CH2 tool, either the CH_2 bonds to the workpiece radical, or it reacts with something else, such as abstracting a hydrogen. In either case, the tool + workpiece is again in a sufficiently deep minimum that the possibility of further reactions can be neglected.

These considerations motivate the exponential estimate above, but do not establish a quantitative error probability for every mechanosynthetic step. More detailed analysis of the competing pathways along the imposed tool trajectory is needed to assess its accuracy.

For an SRU with $\sim 10^8$ atoms, and assuming 10 mechanosynthetic steps per atom, an error rate of 2.8×10^{-10} per reaction step implies we have a $\sim$76% chance of completing all blocks required for an SRU without any mechanosynthetic errors.

This implies the probability of an error during the synthesis of any given block is very low, while the probability of encountering a mechanosynthetic error in fabricating some block at some point during a replication cycle is significant. We need some mechanism to deal with infrequent errors during block mechanosynthesis. It's also the case that the estimated 76% success probability is highly uncertain: it might be substantially smaller or substantially larger. Our mechanism should be insensitive to even significant changes in this percentage.

Two obvious strategies exist for coping with this uncertainty.

(1) We test for errors during or after each mechanosynthetic step, allowing us to catch mechanosynthetic errors shortly after they occur. We then abort fabrication of the block if an error is detected. We can also do block level testing upon completion of each block. Upon detecting an error, we remove the block from the fabricator and store it, with associated discarded tools, in unused daughter space, then reset the fabricator to a known good state. This requires a block of code following completion of every block to remove the waste block and reset the fabricator. This code is required and must be broadcast regardless of whether any individual SRU does or does not execute this code. It also requires a strategy for providing the missing block (see discussion below). However, in exchange for this overhead we gain a robust ability to deal with substantially higher error rates.

(2) We (a) detect mechanosynthetic errors and block level errors and (b) provide enough error correction at the mechanosynthetic level so that the probability of an unrecoverable mechanosynthetic error is reduced to an acceptable level. An acceptable level is defined as a level which results in a failure rate per replication cycle of <0.1% (see SRU Failure Handling (Section 6.2)). If any block level error is detected or if any unrecoverable mechanosynthetic error is detected the SRU fail-stops. Block level errors should occur very infrequently if we are testing after every reaction, so fail-stop is an acceptable response to any remaining block level errors. This strategy entirely avoids the need for disposing of failed blocks and cleaning up the fabricator following a block abort but does force us to develop enough mechanosynthetic tests and sufficiently robust error recovery procedures to achieve the desired reduction in error rates.

In strategy (1), we require a method for dealing with a block whose fabrication was aborted. The simplest strategy is to reserve two fabrication attempts for every block. If fabrication succeeds on the first attempt, the second attempt is not executed. If fabrication fails on the first attempt, the second attempt is executed. Consider $M = 10^6$ blocks of 100 atoms each. Let p_{BL} be the probability that one block-fabrication attempt fails. Assuming independent block-fabrication attempts, the per-replication-cycle success probability, P_{success}, must satisfy:

$$P_{\mathrm{success}} = (1 - p_{\mathrm{BL}}^2)^M \geq 0.999. \tag{2}$$

Solving the boundary case $P_{\mathrm{success}} = 0.999$ in Eq. (2) for p_{BL}, this (remarkably simple) strategy requires a block level error rate of 1 in 31,600, or 3.16×10^{-5}. The error rate for each atom added to a block, from the previous error analysis, is 2.8×10^{-9}. If a block has 100 atoms this is an error rate per block of $\sim 2.8 \times 10^{-7}$. This is over two orders of magnitude less than the required block-level error rate of 3.16×10^{-5}.

Our estimated block level error rate is better than our requirement by somewhat over a factor of 100.

This simple "Build every block twice" strategy is robust against the kind of error rates we anticipate might occur and is remarkably simple. Because instructions are broadcast to all SRUs, time must be allocated for the second fabrication attempt even where a successful SRU idles. This doubles the time allocated to building blocks, effectively doubling the replication time. This strategy is therefore not used in our baseline.

A somewhat more sophisticated strategy would be to notice that we need many copies of most blocks – the number of unique blocks is very small. For most blocks we could build a few extra copies, keep a buffer of two or three blocks, and achieve similar reliability with very little increase in replication time. In somewhat greater detail, if our buffer had S slots, every time we needed a block we would draw it from the buffer and then manufacture a new block and place it into the buffer. We would build an extra block after every Kth block is withdrawn. If the buffer was already full, this extra block would not be built. If the buffer was not full, we would build the extra block. If we try to withdraw a block and the buffer is empty, we fail-stop. We adjust the parameters S and K to reduce the risk of fail-stop to less than 0.1% per replication cycle.

The baseline uses this spare-block buffer strategy; its buffer size S and replenishment interval K remain to be determined.

The spare-block policy is closely related to finite-buffer queueing and inventory-control models. The system can be modeled as a finite-state birth-death Markov process in which withdrawals decrease inventory, manufacturing replenishes inventory, and failure occurs only upon attempted withdrawal from an empty buffer. Standard queueing-inventory methods can be used to select buffer sizes that achieve an arbitrarily small stockout probability [STGH18, SAR23].

Strategy (1) requires some method of cleaning up the fabricator on block abort. Presumably, this involves removing the partially finished block from the fabricator, along with all tools that might have been involved in the first attempt. This aborted block, now called simply waste, must be put somewhere. In the baseline, this solid waste is stored in unused space in the daughter. Notice that we don't want to store waste in the parent SRU as it will accumulate over multiple generations until there is no space left. If waste is stored in the daughter, it does not accumulate. The estimated waste rate is low enough for this baseline storage strategy.

If solid-waste accumulation proved too high for daughter storage, venting failed blocks and discarded tools would be a risk mitigation. This is distinct from the baseline venting of molecular waste gases.

Strategy (2) avoids the need to deal with waste. It handles errors immediately after the reaction that caused them, when only a small, specific set of errors is possible. The error detection and correction must be sufficient to reduce the per generation error rate to 0.1% or better. We might be able to accomplish this by correcting the most likely one or two errors and fail-stopping on all others. Should this strategy suffice, the actual increase in replication time caused by adding mechanosynthetic error correcting code might result in a factor of two increase in replication time, which can be mitigated by Faster methods of populating dies (Section 13.5).

Error-rate estimates depend on the analysis of a specific set of molecular tools fabricating specific blocks. It seems unlikely that an early proposal would be optimal in terms of the mechanosynthetic error rate. As a consequence, we baseline the more robust strategy: we allow the fabrication of a block to be aborted, and the failed block to be removed from the fabricator for disposal by placement in the daughter. A mitigation strategy if the error rate is too high would be to vent the failed block from the SRU. Further research and analysis of the error rate is needed.

3.8 What we might manufacture

Once a silicon die has been populated, what might we manufacture? An individual SRU can extrude products that are about the size of an SRU, that is, rigid objects that are approximately 500 nm long (although because this dimension is extruded it can be longer) and 60 nm by 100 nm in cross section. So far, we have discussed only how to populate a die with SRUs, to the exclusion of all other products.

We don't have to fully populate the die with SRUs. We might populate every other site with an SRU in a checkerboard fashion. This would mean SRUs would occupy half the sites on the die, leaving the other half of the sites available for SRU-sized product objects. Of course, there's no requirement that half the sites be occupied by SRUs. We could reduce the percentage of sites occupied by SRUs by spacing them farther apart. Using the baseline proposal, we could reduce SRU occupancy to 10%, leaving 90% of the sites available for product objects that fit in the 500 nm × 100 nm × 60 nm form factor. This manufacturing process might eventually have manufacturing costs approaching the cost of silicon dies plus the materials and energy costs involved in processing. The products produced would be silicon dies with surfaces populated mostly with SRU-sized atomically precise products, and with some small percentage occupied by SRUs.

The most obvious thing we might want to manufacture would be computers. Each SRU-sized module could hold atomically precise logic elements, effectively packing the die with 90% SRU-sized logic modules. Direct atomically precise connections between logic modules could be put in place by the SRUs occupying the remaining 10% of sites. Each logic module could be placed on a customized site designed to support that particular module in that particular location. The modules, whether they used molecular mechanical logic [MFH+18, HMA17, Dre92, MFR+19] or electronic logic [MFA25], could interface with the outside world electronically through the wires in the die. Faster methods of populating a die might also be possible (see Faster methods of populating dies (Section 13.5)).

A second class of products would be medical nanorobots, which might be bounded in size by two µm in all three dimensions. This would require some additional manufacturing methods to increase the product size from SRU-sized components to 2 µm × 2 µm × 2 µm. Various methods of accomplishing this are discussed in Additional concepts for building 2 µm products (Section 13.3).

Unlike populated dies, which would themselves be products and could be handled as conventional macroscopic objects, medical nanorobots are microscopic in size and would have to be exported from the manufacturing system. As such bots would typically operate in a fluid environment, we might export them into the feedstock solution and then filter them out. We do not further address this issue in this paper.

An alternative manufacturing approach might be to build product objects "above" a fully populated die of SRUs. In this approach, each SRU on a fully populated die assembles SRU-sized product objects together in the liquid environment "above" the plane of the SRUs. Additional SRU-sized product objects ("product modules" or just "modules") could be added to the assembly every time the SRUs manufactured an additional die-full of modules. If additional modules were added "beneath" the growing assembly of modules, there is no hard limit on the thickness of the assembly of modules being pushed up from the die surface by the SRUs. See Bigger than an SRU (Section 9).

3.9 Summary

Nanofactories use vacuum mechanosynthesis of a specified set of elements to build products, including more nanofactories. Our baseline palette of elements is H, B, C, N, O, Si, Ge, and Pt.

Hydrogen and carbon provide structural elements (diamond, lonsdaleite, nanotubes), wires in the form of armchair (n, n) nanotubes, and insulators in the form of defect-free diamond or lonsdaleite, and possibly polyethylene. Germanium is included because it forms weaker bonds than B, C, N, and O. For example, as Ge–C bonds are weaker than C–C bonds, we can use germanium-tipped tools to deposit and release carbon fragments in a controlled fashion [FM08a, TSMF07, C^+26, B^+26].

We add nitrogen and boron because they can form dipolar bonds, which are very useful in bonding AP Blocks together using bonds that are reasonably strong but which allow disassembly (should that prove necessary). Dipolar bonds are also convenient in generating blocks from larger diamondoid structures ("blockification"). Starting with a solid block of a diamondoid workpiece, we can carve it into blocks by carefully replacing selected adjacent bonded atoms, such as C-C, Si-C, or the like, with adjacent B-N atoms that form a (weaker but structurally similar) dipolar bond. Boron and nitrogen also form boron nitride structures, including hexagonal boron nitride (h-BN). These materials have useful structural, physical, and electronic properties.

Si, along with C, forms silicon carbide. This material is strong, and its geometry can be usefully modified by substituting Si for C or C for Si. In addition, there is an entire industry based on silicon which we can benefit from. Starting a build from a silicon surface is particularly well studied and experimentally supported.

Oxygen is particularly useful in stabilizing certain diamond surfaces, such as the (100) surface. It can also form dipolar bonds with boron. When combined with silicon (as in quartz) it makes an effective insulator.

Platinum rounds out our list of eight elements. We use it for its catalytic abilities, particularly for the generation of H_2 gas. The inclusion of Pt emphasizes that mechanosynthesis is not restricted to diamondoid materials.

We adopt these eight elements, H, B, C, N, O, Si, Ge, and Pt, as our baseline. While still limiting our palette to only eight of the over 100 elements in the full periodic table, this still opens a very large range of options, some of which will be discussed later. Using these eight elements in the design of our baseline nanofactory makes the design easier than it would be with a more limited subset, while still being a small enough set to make experimental realization feasible.

4 The Simple Replicating Unit or SRU

4.1 Introduction

In this section we walk through the physical layout of the baseline SRU (Figure 8).
An **SRU**, or **S**imple **R**eplicating **U**nit, is

1. A **fabricator**
2. In a vacuum tight **shell**
3. On a silicon die that provides **power** and **control**
4. That can **assemble** a **daughter** SRU
5. Which is then **moved** to an open **site** on the **die**

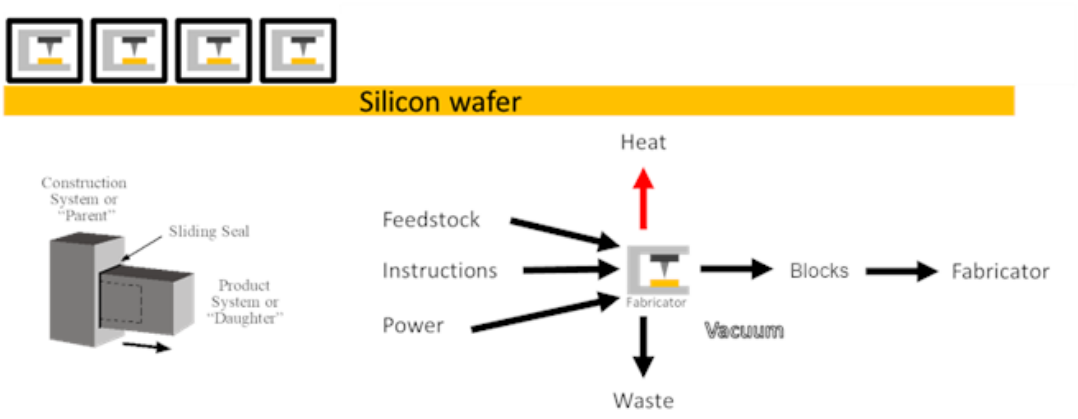


Figure 8: SRUs on a surface, daughter extrusion, and the fabricator.

4.1.1 SRU layout

The illustration in Figure 9 shows the SRU immediately after extrusion from its parent. The SRU is 500 nm × 100 nm × 60 nm. During extrusion, nothing extends beyond the 100 nm × 60 nm cross section. The Extrusion Sleeve (ES) of the SRU has not yet been extruded and the External Assembly Arm has not yet been attached to the surface of the shell.

This cartoon omits many details, including control lines (binary Bowden cables and wires), extrusion mechanisms, etc. The Stewart platforms are indicated by three struts instead of six to avoid clutter. The cartoon also omits the fact that a single SRU occupies an SRU site that is 600 nm × 200 nm. The site is larger than the actual SRU to provide clearance for operations involving the SRU, such as its initial placement on the site.

The SRU External Assembly Arm (EAA) is shown in Figure 10. It is a two-stage Stewart platform that provides 12DOF for manipulating assemblies external to the SRU. It is extruded as a separate product by the SRU. When a parent SRU extrudes a daughter SRU, it then immediately extrudes a daughter EAA, which it attaches to the daughter SRU (see External Assembly Arm (EAA)

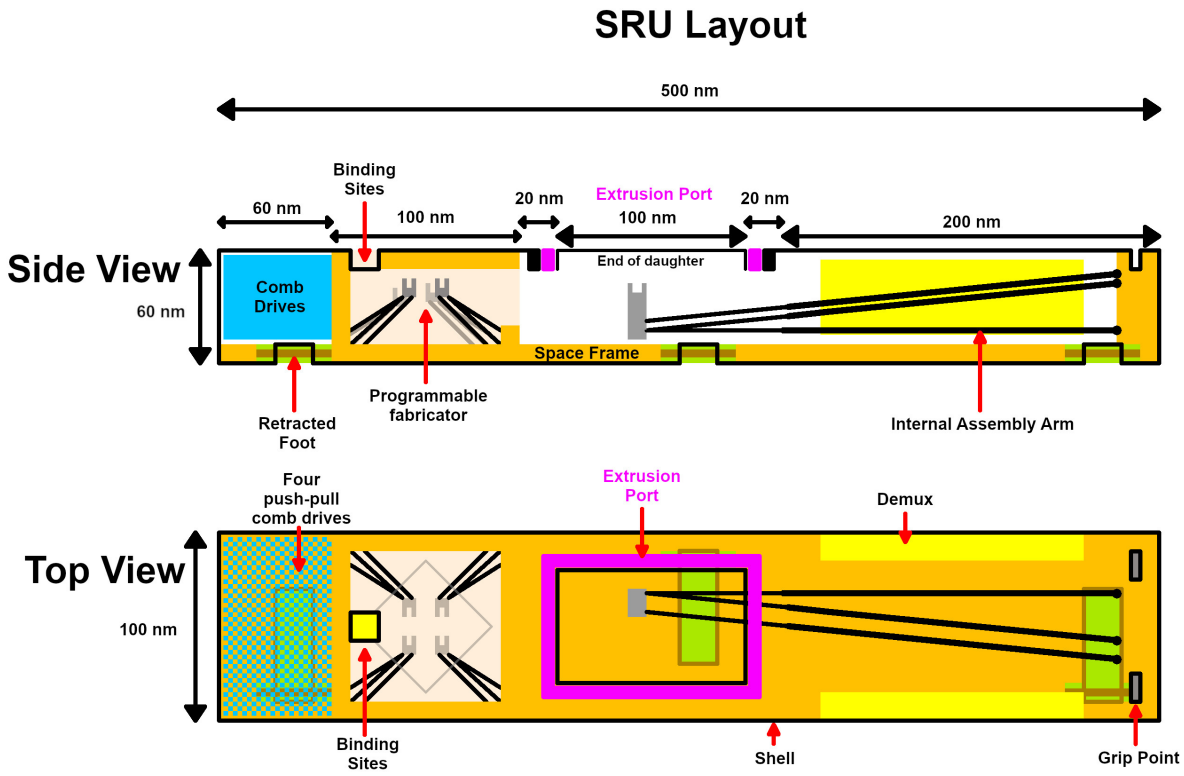


Figure 9: SRU layout, major components, and dimensions.

(Section 4.8.2)). The 60 nm by 100 nm base shown at the left end of the EAA creates a tight seal with the ES while the EAA is being extruded, preventing flooding of the SRU during this process, as shown in Figure 11(b).

The illustration in Figure 11(a) shows the SRU after it has extruded its ES (pink), which extends 225 nm above the shell of the parent SRU and surrounds ∼half of the just manufactured daughter SRU. About 255 nm of the daughter SRU extends above the ES, exposing the two grip points on the exterior of its shell at the top. The EAA is shown in its expanded state mounted on the right of the SRU.

The illustration in Figure 11(b) shows the just extruded EAA in its contracted state with its base just emerging from the ES. The parent's expanded EAA, if it has previously grasped and is holding the daughter SRU, can attach the daughter SRU to the base of the daughter's emerging EAA. The base, together with the ES, prevents flooding of the parent SRU.

4.2 Shell

The shell is 500 nm long, 100 nm wide, 60 nm high, and ∼1 nm thick. The shell is made of AP Blocks. A large part of the atom count (almost half) is in the shell.

The shell separates the inside of the SRU, which is vacuum, from the outside, which is a mixture of feedstock molecules.

Everything inside the shell is eutactic (atomically precise, "well ordered"). The shell itself is also atomically precise and eutactic. In a eutactic structure the position of every atom is the

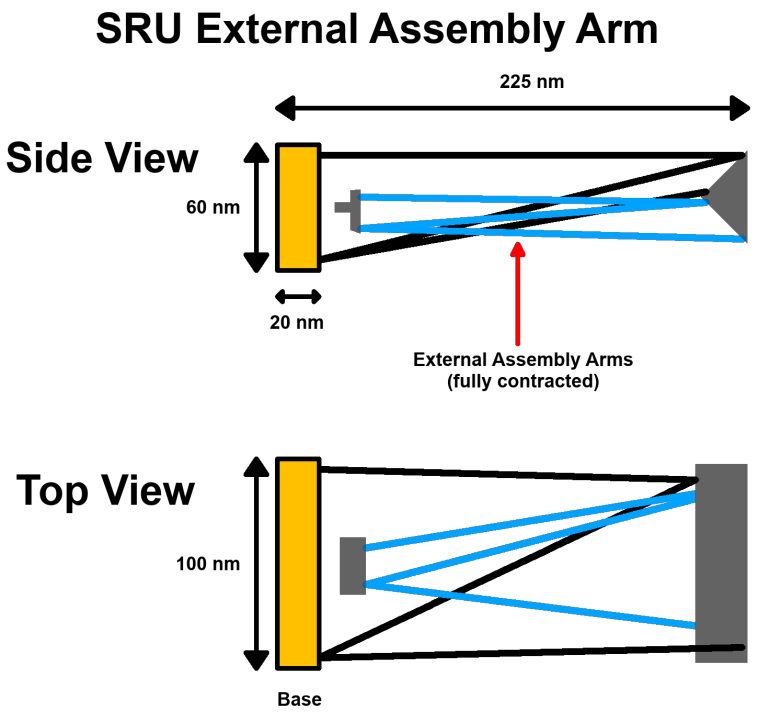
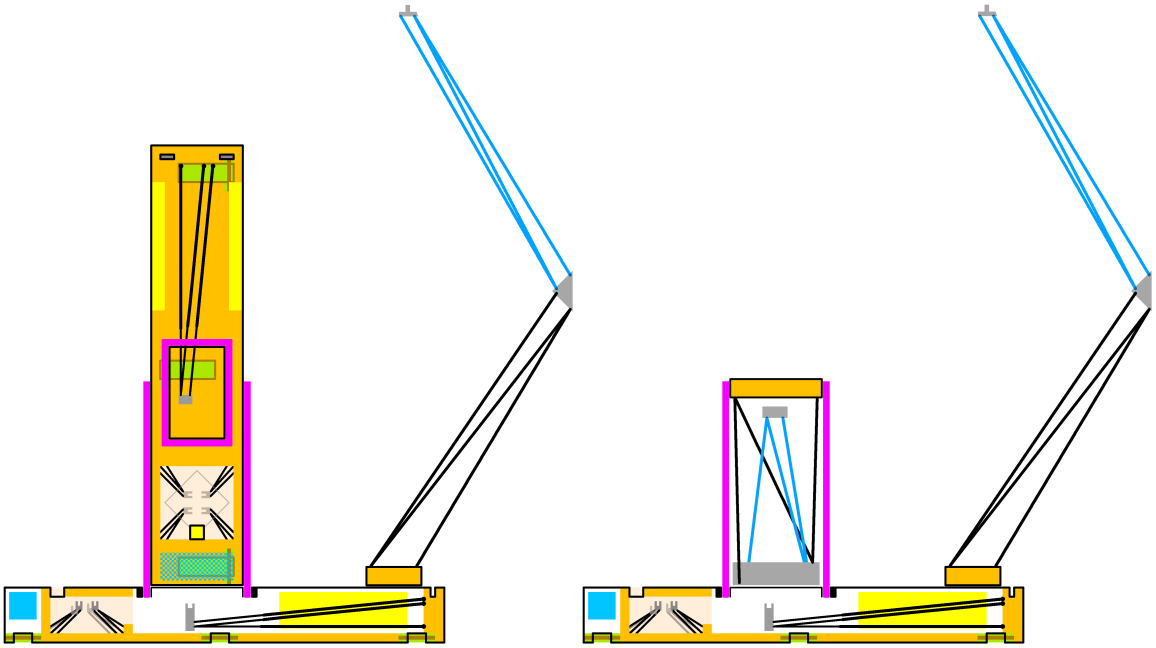


Figure 10: Side and top views of the External Assembly Arm.



(a) (b)

Figure 11: Daughter and External Assembly Arm extrusion: (a) daughter extrusion with the arm expanded; (b) extrusion of the folded arm.

result of deliberate design, other than positional variations caused by thermal noise. In "typical" eutactic structures, every atom is in a single, known position (other than thermal noise). There are a few error-correction methods that allow for the eutactic system to be in one of two or three possible states which will, after appropriate conditional tests and conditional operations have been performed, be merged back into a single corrected state (see Simple Closed Loop Mechanosynthesis (Section 6.9)). Today, these multi-state configurations are expected to be of short duration and to converge rapidly back into a single corrected state. Complex multi-state configurations of long duration would be possible but remain for future research to explore.

Everything outside the shell (other than the external assembly arm and the extrusion sleeve) is non-eutactic (not atomically precise). It is part of the support environment and is manufactured by conventional methods.

Other inert environments inside the shell are possible. We might, for example, use gaseous helium. While this might seem to violate our definition of a eutactic environment, gaseous helium is inert and would not alter any fundamental assumptions about the operation of the SRU. The use of helium gas as an inert internal environment would result in a system where the location of the helium atoms was largely unknown, but the system could still be classed as "eutactic", giving some idea of the issues involved in a precise definition of "eutactic". The design of atomically precise hydraulic systems, where the working fluid is counted in and out molecule by molecule but the precise locations of individual molecules of working fluid are unknown, presents a similar issue. Such a hydraulic piston could achieve positional accuracy better than 20 pm [Mer23].

The shell might not be sufficiently rigid to support assembly operations (among other structural concerns), so a rigid space frame (see introduction to space frames (Section 9.2.1)) has been added inside the shell to provide a "shop floor" upon which "equipment" can be mounted. The space frame is expected to have an atom count that is less than the atom count of the shell but similar in magnitude.

The space frame consists of struts and nodes. If a strut is 10 times as long as it is wide, then a strut with a square cross section with sides of 1 nm (area 1 nm^2) would be 10 nm in length. This corresponds roughly to the pictured 10 nm thick space frame shown at the bottom of the side view. Taking a space frame with cubic unit cells, each unit cell being 10 nm on each edge, three axial struts per unit cell (one along each of X, Y, and Z), each 10 nm long, and three face-diagonal struts (in XY, YZ, and ZX), each $\sqrt{2} \times 10\,\text{nm}$ long, gives a total volume of ∼72 nm^3 per unit cell for the struts. This unit cell has a total volume of 1,000 nm^3, giving a density of ∼7% for such a space frame. As the space frame is 10 nm thick, while the shell is ∼1 nm thick, the atom count for a similar length of space frame should be ∼70% of the atom count of the shell.

We conclude that a 10 nm thick space frame can provide considerable additional stiffness in support of a ∼1 nm thick vacuum tight shell, and that the 10 nm thick space frame would add less than the mass of the ∼1 nm thick shell. A quantitative analysis of the increased stiffness is needed.

4.3 Control

4.3.1 Contact pillars

Each SRU site has six conductive contact pillars connected by vias to buried power, ground, and data conductors. The contact pillars are manufactured by conventional semiconductor manufacturing techniques. The contact pillars rise ∼60 nm above the surface. Each contact pillar is rectangular in outline, about 60 nm × 40 nm. They are arranged in three opposed pairs. The long

sides of each contact pillar face the other contact pillar in the pair. The facing surfaces are ∼40 nm apart. The SRU is positioned on top of the contact pillars. After placement, the SRU extends three "feet", one into the gap between the two pillars of each pair. It then drives two opposed contact spikes from each foot into the two contact pillars, making two separate electrical connections per foot. The opposed lateral forces cancel. The contact spikes are driven into the contact pillars at a slightly upward angle, pulling the SRU slightly closer to the surface after they are driven home. The contact spikes, assumed here to be nanotubes, provide both electrical contact and passive mechanical retention without thermal annealing.

A possible geometry is shown in Figure 12.

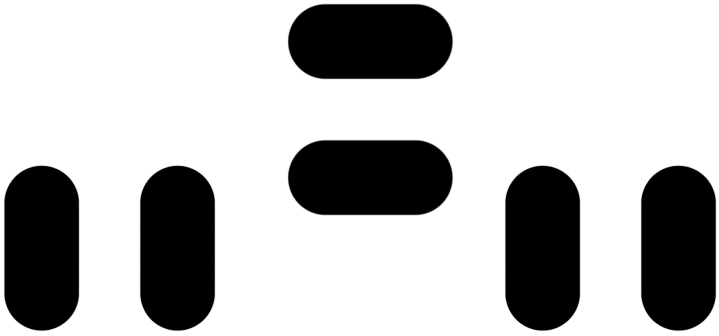


Figure 12: Arrangement of six contact pillars at an SRU site.

The contact spikes self-center when they are driven into the contact pillars. Even though the foot in the left panel is left-of-center, when the two contact spikes are driven into their respective contact pillars, the foot will be centered, as shown in the right panel (Figure 13).

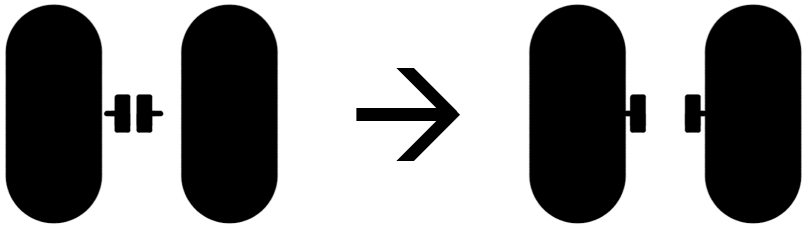


Figure 13: Self-centering of contact spikes on the contact pillars.

This self-centering is left-right for the left and right pairs of contact pillars but is up-down for the central pair of contact pillars, which are rotated 90° with respect to the left and right pairs. The feet are designed to allow movement in the direction orthogonal to the self-centering direction, allowing the feet to self-center the SRU in both X and Y. This is a natural outcome if the feet are lowered into position by rotation.

While this self-centering allows for control of two degrees of freedom, and possibly three, the SRU has six positional degrees of freedom. There are six contact pillars, so in principle if each contact pillar can control one degree of freedom, we could control all six degrees of freedom if there was one contact point between the SRU and each pillar. Whether we should use the contact pillars for such positional control while also using them for electrical contact is unclear. We could add additional pillars or other features specifically for positional control of the SRU. A great deal depends on the details of the design of the pillars and their interaction with the silicon surface.

The baseline assumption is that more complex positional control will not be needed, but control of the SRU in 6DOF with respect to surface alignment features is available as a risk mitigation; see Self-alignment of the SRU in 6DOF (Section 12.1).

Buried conductors carry Ground, V+, S1, S2, D1 and D2 across the die. See The demux (Section 4.3.6) for a brief overview of their significance. Vias carry the signals from lower layers to the surface. A single SRU with a 500 nm by 100 nm footprint would then sit on the die and connect electrically to the six contact pillars. The baseline assumes we adopt some mechanism for self-alignment of the SRU with respect to features on the silicon surface.

4.3.2 Feet

There are three feet, each 10 nm × 20 nm × 60 nm in size. They are held in three indented cavities in the shell of about the same dimensions. A shaft with an axis 5 nm from the end of the 60 nm length and anchored in the SRU enables each foot to be rotated outwards until it extends at right angles some 50 nm from the SRU. Once extended, one spike on each side of the foot can be thrust into the adjacent contact pillars. As there are three feet, this enables contact to be made with six contact pillars. Conveniently, this is the number of connections required to drive the six electrical inputs to the baseline SRU: Ground, V+, S1, S2, D1 and D2. See The demux (Section 4.3.6) for more information on these six electrical inputs.

More contact pillars could provide a higher data rate to the SRU, resulting in a faster build rate but also drawing more power and generating more heat.

Two feet are located at the ends of the SRU, and one foot in the middle. The central foot is rotated 90° with respect to the other two feet. As each foot must drive contacts into two contact pillars, we need to have an electrically separate connection on each side of each foot.

The dimensions are as proposed above.

With these dimensions, we would have to extend a spike more than 10 nm from each side of each 20 nm wide foot, and possibly as much as 20 nm or even somewhat more to make reliable contact with the contact pillar. The greater extension of the spike might be needed if the foot were off-center in the space between contact pillars, possibly pushing against one of the contact pillars on one side and leaving a wider gap on the other side.

Final foot design must be done in conjunction with the design of the die to which the SRU will be attached.

4.3.3 Wires

There are six wires from the feet to the Differential comb drives (Section 4.3.4).

The wires in the baseline design are (4,4) carbon nanotubes. The (n, n) armchair nanotubes, including the (4,4) nanotube, are metallic. Although the smaller (3,3) armchair nanotube has been observed experimentally, its stability without the protection of a surrounding nanotube is uncertain [GSI08]. Risk-reduction work is still needed to demonstrate that short, unpassivated (4,4) nanotubes are stable in the eutactic UHV environment and can be fused into longer conductive wires.

Three views of a 64-atom segment of a (4,4) carbon nanotube, minimized with HyperChem's implementation of AM1, are shown in Figure 14.

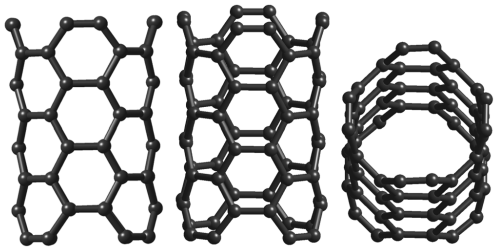
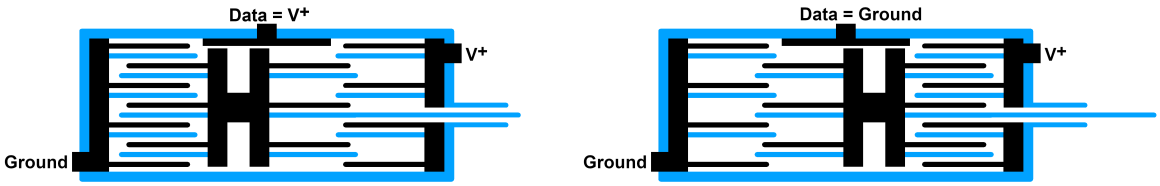


Figure 14: Molecular wire models calculated with AM1.

4.3.4 Differential comb drives

A differential comb drive converts a binary electrical signal (logical 0 or logical 1) into a binary mechanical signal in a binary Bowden cable (logical 0 or logical 1). The binary electrical signal can be either V+ or Ground. If a logical 0 is represented by Ground, then the illustration of a differential comb drive shows it with an input of Ground (logical 0) and the comb drive all the way to the right (mechanical logic 0). If the electrical signal switches to logical 1 (V+), then the differential comb drive would swing to its leftmost position and would drive the attached inner cable of the binary Bowden cable leftward as well. By this means, the comb drives in the SRU convert binary electrical inputs into binary mechanical outputs.

All mechanical operations in the SRU are driven by four differential comb drives. These four differential comb drives are in their turn driven by six wires: Ground, V+, S1, S2, D1 and D2 (Figure 15).



(a) (b)

Figure 15: Differential comb drive displacement states: (a) Data = V+; (b) Data = Ground.

For a more complete description of differential comb drives, see Differential Electrostatic Comb Drives for All Actuation (Section 7.15).

The total volume allocated to the four differential comb drives operating on 0.19 V, with a 24 nm shuttle throw reduced 12:1 to a 2 nm output throw, is 227,205 nm^3. The nominal output force is 1.25 nN before allowing for shuttle acceleration, leaving approximately 1 nN available for actuation. See Baseline design for a differential electrostatic comb drive (Section 7.15.4).

The space available for differential comb drives is 95 nm × 45 nm × 57 nm = 243,675 nm^3. We use all of it, mostly for the comb drives but a little for supporting infrastructure.

The six electrical signals are:

1. V+, the positive DC voltage supplied to the system, which is 0.19 V
2. Ground, 0 volts
3. S1, an abbreviation for “Select One”

4. S2, an abbreviation for “Select Two”
5. D1, an abbreviation for “Drive One”
6. D2, an abbreviation for “Drive Two”

The labels “S1”, “S2”, “D1” and “D2” are also applied to the corresponding contact pillars on the die surface, the internal wires that carry the corresponding electrical signal, the differential comb drives to which those wires are connected, and the binary Bowden cables driven by those comb drives.

The labels “V+” and “Ground” are applied only to the appropriate electrical contact pillars and the corresponding internal wires, not to any comb drives or binary Bowden cables.

The data lines, S1, S2, D1 and D2, are binary and can be either V+ or Ground. The binary Bowden cables S1, S2, D1 and D2 are the four inputs to the demux. For an explanation of what a demux is and what it does, as well as what the four data lines do, see “The demux (Section 4.3.6)”.

4.3.5 Bowden cables

Bowden cables, much like the derailleur cables or brake cables in bicycles, have an inner cable and an outer housing and can be used to transmit both compressive and tensile forces. They can be used to conveniently transmit force from the output of the differential comb drives to the inputs of the demux, and from the outputs of the demux to the inputs of the various devices around the SRU that need to be controlled. We anticipate using binary Bowden cables, which need only allow the inner cable to travel back and forth between two positions, separated by a small distance, perhaps 2 nm. It is possible to use analog Bowden cables for the transmission of analog data, such as an analog value between 0 and 1. We do not anticipate using analog transmission, although this would be possible. Multi-valued logic has generally lost against simple binary logic except in specialized cases, such as high-density memory.

4.3.6 The demux

At a high level, a demux, or demultiplexer, takes a small number of inputs and uses them to generate many outputs. This is most often used when the data lines are relatively expensive compared with the cost of the demultiplexer, and the lower data rate caused by using fewer data lines is less significant than the cost savings. In our case, we’re taking four mechanical inputs provided by four binary Bowden cables and generating ∼60 output pairs on ∼120 binary Bowden cables (as discussed below).

The demux takes as input four binary Bowden cables: S1, S2, D1 and D2. S1 and S2 (Select 1 and Select 2) select the pair of output binary Bowden cables to be driven, while D1 and D2 (Drive 1 and Drive 2) drive the selected output pair of binary Bowden cables. S1, S2, D1 and D2 begin their lives as electrical signals on wires, as discussed in the section on differential comb drives (Section 4.3.4). These signals then power four differential comb drives which power four binary Bowden cables. All differential comb drives also each have two more electrical inputs, V+ and Ground, which are also necessary in the electrical-to-mechanical conversion process. In total, therefore, there are six electrical inputs provided by six contact pillars on the die that are carried into the SRU through three feet by six internal wires (two wires per foot). These six electrical inputs are: V+, Ground, S1, S2, D1 and D2.

For a broader discussion of the demux and the meaning of S1, S2, D1 and D2, see Broadcast Architecture (Section 6.1).

There might be 60 outputs from the demux, each output consisting of two binary Bowden cables. This implies the demux might have ∼120 binary Bowden cables as outputs. The number of outputs can be estimated as follows. There are 7 Stewart platforms in the SRU. Four of these are in the programmable fabricator, one is used for the internal assembly arm, and two (one on top of the other) are used in the external assembly arm. Each of these Stewart platforms has six struts, giving 6 degrees of freedom for each Stewart platform. We assume the tool holder on each platform can grip a tool handle purely by movements of the platform and does not require an additional explicit degree of freedom to "grip" and "ungrip" the handle (see Stewart platform (Section 4.4.1)). This is $7 \times 6 = 42$ degrees of freedom in the Stewart platforms, and perhaps another 8 to 12 or so degrees of freedom in various other devices, for a rough total (with some additional margin) of 60 degrees of freedom, with each degree of freedom being driven by two binary Bowden cables. This totals roughly 120 binary Bowden cables, which is a good enough approximation for the moment.

4.3.6.1 The MacArthur Mechanism

James MacArthur designed and modeled the atomically precise core of a linear demultiplexer (see Figure 16), following exploration and modeling work on several alternative designs [MFR23]. The length of the demux is two demux segments of 140 nm each, for a total length of 280 nm in the SRU layout (Section 4.1.1), giving us $280/120 \approx 2.3$ nm per output, which should suffice. The two drive inputs to the demux, D1 and D2, operate in parallel, each driving one of the two 60-output demux segments. The two demuxes are physically positioned on opposite sides of the SRU.

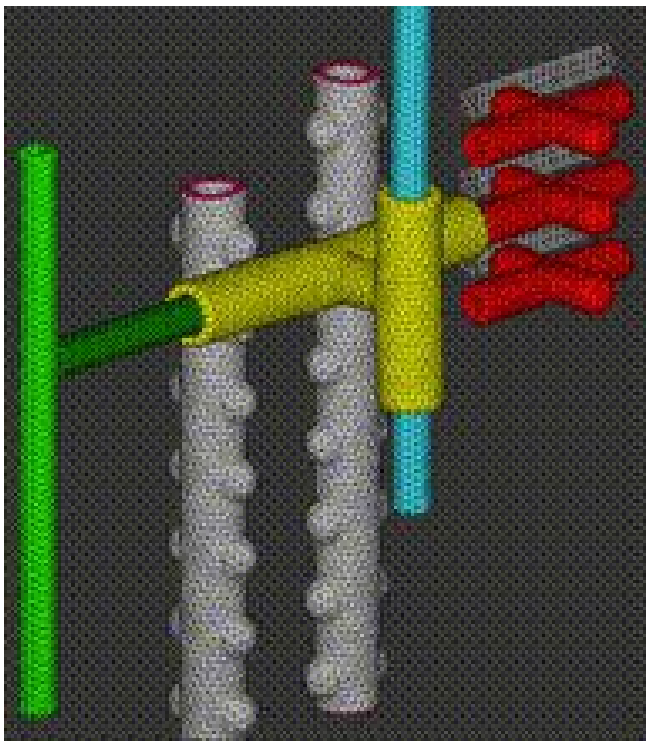


Figure 16: Atomically precise model of a demultiplexer segment by James MacArthur.

Three output nanotubes are shown in red. In the actual demux, there would be ∼120 red output

nanotubes, divided into two groups of 60, one group driven by D1, the other by D2. The yellow selector is moved from output to output by the two knobby nanotubes. The knobby nanotubes rotate beneath the yellow selector. When one or both knobs face upwards, the two knobby nanotubes move the yellow selector to the next output. Each knobby nanotube can individually slide past the yellow selector when the knobs are rotated sideways, while the other knobby nanotube holds the yellow selector lever in position. Each knobby nanotube moves back and forth between two axial positions and rotates between horizontal and vertical. That is, even though the yellow selector moves between sixty different output nanotubes, the two knobby nanotubes never move far at all: they just move back and forth between two positions and rotate between horizontal and vertical.

The two green nanotubes are driven by D1 and D2. The two green nanotubes are not physically joined. The horizontal green nanotube slides along the vertical green nanotube when the yellow selector is moved. D1 drives the first group of 60 outputs, and D2 drives the second group of sixty outputs. Assume the picture is showing the first demux, driven by D1. When a selected output is to be driven, a framework (not shown) drives the entire long green vertical bar against the horizontal green bar which slides through the yellow selector. When the long green bar pushes against the horizontal green bar, the horizontal green bar then pushes against the selected output.

There would be two yellow selectors, one selector for each group of sixty red output nanotubes. Regardless of whether the two physical demultiplexers are located apart from each other, side-by-side or one after the other, the two logical selectors must be kept in logical sync with each other. Keeping them in physical sync by having the yellow selectors physically joined would be one implementation method for keeping them in logical sync. In the present design, because the two sections of the demux are physically separated, the yellow selectors in the two separate sections will be driven in parallel by duplicated copies of S1 and S2. That is, the two yellow selectors are not physically joined together but must still be kept in logical sync. To do this, the same signals will be used to drive both yellow selectors, forcing both yellow selectors to move in sync even though they are physically separated. This implies that the two input binary Bowden cables carrying S1 and S2 must be split into four output cables carrying S1A, S1B, S2A and S2B, with S1A = S1B and S2A = S2B. This physical duplication of the signal carried by an input binary Bowden cable into two output binary Bowden cables is not difficult, but it does imply that S1 will have to deliver twice the force required to actuate S1A or S1B individually. Similar comments apply to S2. This might require an increase in the size of the comb drive driving S1 and S2. There is substantial margin in the relevant parameters should this adjustment be needed.

A four-phase clocking sequence for the two knobby nanotubes might be the following. If we assign the knobby nanotubes the names of L (for Left) and R (for Right), and treat their rotational positions as locked (the knob holds the yellow selector in place) and unlocked (the knob is rotated out of the way and does not hold the yellow selector in place), and start with L locked and R unlocked, the four-phase clocking sequence might be:

1) Pull L back and push R forward (simultaneously), moving the selector back one
2) Unlock L and lock R (simultaneously), leaving the selector unchanged
3) Push L forward and pull R back (simultaneously), moving the selector back one
4) Lock L and unlock R (simultaneously), leaving the selector unchanged

This sequence would advance the selector by two output positions per complete four-phase clock cycle. Reversing the sequence would move the selector in the opposite direction.

4.3.6.2 A screw drive

We might use a variant of the MacArthur mechanism [MFR23] by driving the yellow selector with a screw drive (Section 7.19.1) instead of with two knobby nanotubes. In this variant, the two knobby nanotubes are replaced with a long screw, and the yellow selector is attached to a nut that rides along the screw. The screw is rotated by S1 and S2 in either a clockwise or counterclockwise direction. The pitch of the threads moves the selector forwards or backwards one output after a full four-phase clock cycle, depending on which direction the screw is rotated.

As the screw's thermally induced angular uncertainty becomes worse the longer the screw, and as any angular error translates into axial error along the axis of the screw, this angular error must be carefully considered and controlled in any design. By contrast, the MacArthur mechanism is less sensitive to angular error. Until the angular error becomes so large that the knobby nanotubes allow the selector to slip from one position to the next, angular error does not cause axial error.

The axial positional uncertainty in a screw design, where the screw has radius r, is:

$$\sigma_z = \frac{p}{2\pi}\sqrt{\frac{k_B T L}{GJ}}, \tag{3}$$

where

$$J = \frac{\pi r^4}{2}, \tag{4}$$

k_B is Boltzmann's constant, J is the polar second moment of area, L is the length of the screw, r is the screw radius, p is the screw lead, T is the temperature, and G is the shear modulus. Reasonable approximations for our purpose might be: L is 100 nm, r is 1 nm, for a single start screw the screw lead is the same as the pitch p, which might be 0.7 nm, T is 210 K, and G for diamond is ∼500 GPa. This results in $\sigma_z \approx 2.14\,\text{pm}$.

The binary Bowden cables from the demux to the four Stewart platforms of the fabricator and the internal and external assembly Stewart platforms are not shown on the layout. The binary Bowden cables to the second external Stewart platform, which sits atop the first external Stewart platform, are also not shown. Had they been shown, they would have twined around the struts of the first external Stewart platform in a loose-fitting helical fashion, allowing the struts to extend and contract while not interfering with the function of the binary Bowden cables that control the second external Stewart platform.

4.4 Fabricator

4.4.1 Stewart platform

A Stewart platform consists of a base, a platform, and six struts. Each strut can be lengthened or shortened. The platform can be positioned in X, Y and Z and in roll, pitch and yaw with respect to the base. Therefore, the Stewart platform is a 6DOF or six degrees of freedom positional device. The geometry of Stewart platforms can vary considerably, depending on the length of the struts, their positioning on the platform and base, and the size and geometry of the platform and base (Figure 17).

We assume the platform has a tool holder that can "pick up" or "put down" a molecular tool. Each molecular tool has a handle which the tool holder can grasp. The SRU fabricator has four

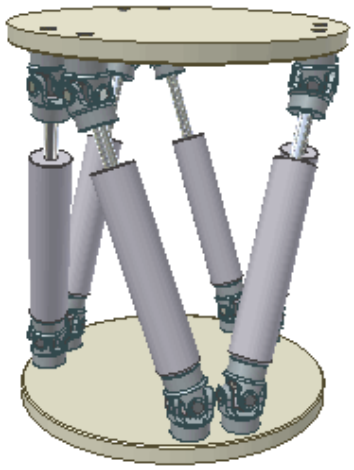


Figure 17: Stewart platform. Still frame from UtzOnBike's *Hexapod general Anim* (frame 10), licensed under CC BY-SA 3.0.

Stewart platforms and can wield four molecular tools at a time. The internal assembly arm wields one tool, one end of which has a block-specific "gripper" which grips a block, while the other end has a handle that the tool holder can grasp.

We leave for future research a tool holder implemented as an active gripper with a single driven degree of freedom, as a hand might close on the handle of a screwdriver. For the baseline, we assume the tool holder is implemented purely by the positional motion of the Stewart platform. Such passive robotic tool interface systems can have a wide range of possible solutions. Various methods of sliding the tool holder around or into the handle and possibly snapping it into place are possible. van der Waals forces rather than mechanical snaps could also be used. Passive tool attachment avoids the need to provide an additional active degree of freedom to attach and detach the handle, reducing the need for an additional control line to the tool holder [Rob13]. Similar comments apply to the external assembly arm, which has an EAA tool holder that grasps the grip point on the external shell of the SRU. See the discussion under Assembly tools (Section 14.8).

4.4.1.1 Stewart Platform Strut

Each strut (Figure 18) consists of a threaded rod and a long internally threaded sleeve (coupling nut), with both components contributing to the total adjustable length of the actuator. Rotation of the threaded rod with respect to the threaded sleeve causes the strut to lengthen or shorten, much as a nut on a threaded rod will move along the rod when the rod is rotated.

For our Stewart platform, each strut is driven by a pair of binary Bowden cables. The pair of binary Bowden cables causes a rotational drive to rotate the threaded rod in fixed rotational increments. This stepped rotation can be either clockwise or counterclockwise, causing the strut to extend or contract by a fixed increment. Each step causes the strut to extend or contract by approximately 20 pm. See Driving a screw drive with four-phase logic (Section 7.18) for more details.

The threaded rod and threaded sleeve are both helical structures (Section 7.19). Such structures can be made by repeating a single molecule, protein, or unit. Examples include the tobacco mosaic virus and microtubules in cells. For our design, the single repeating unit will be a block. See Impact on axial stiffness of using AP Blocks for struts (Section 7.14.9) for a design approach that

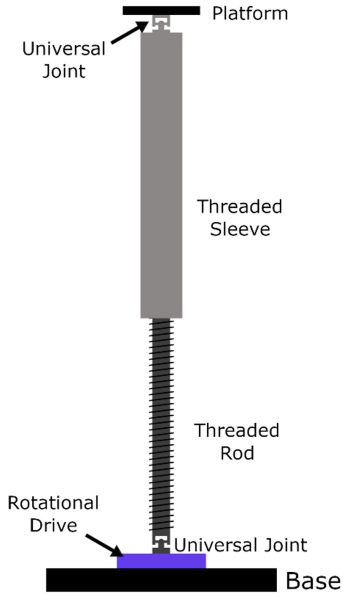


Figure 18: Threaded strut of a Stewart platform.

maintains stiffness in a blockified design.

4.4.2 Programmable fabricator

Rather than use a hard-coded (hard automation) metabolizer, the baseline SRU uses a programmable fabricator that doubles as a metabolizer to metabolize feedstock. It has four six-degree-of-freedom positional devices (Stewart platforms) that terminate in tool holders, allowing them to pick up and put down a wide range of tools held in tool racks. These programmable positional devices should also be able to handle all the metabolic reactions involved in processing feedstock and recharging molecular tools. This means the actual metabolic reactions can be modified late in the design cycle, the changes being made by alterations to the software driving the programmable fabricator.

The use of four 6DOF manipulators for the programmable fabricator was suggested by informal conversations suggesting that this should be sufficient for both metabolizing feedstock and recharging molecular tools. It's also consistent with the fact that four tool tips each with 6DOF were sufficient for all the mechanosynthetic reactions in *A Minimal Toolset for Positional Diamond Mechanosynthesis* [FM08a]. Given the computational and experimental experience with a single XYZ positional device, four 6DOF manipulators should be sufficient for this function.

We use a space frame (see introduction to space frames (Section 9.2.1)) to provide the metrological frame for the programmable fabricator. The space frame wraps around the fabricator, providing a five-sided box. The missing wall makes it possible for the fabricator to pass finished blocks from the fabricator through to the internal assembly arm. The programmable fabricator has a small hole (visible as a yellow region in Figure 9) so that liquid feedstock can gain access to the binding site from the external solution. It also has one or two holes for the charging probe(s) needed by the Electrochemistry infrastructure (Section 12.2.5).

In this layout, we use four Stewart platforms to provide the 6DOF manipulators required by the fabricator. Each Stewart platform has a range of motion that goes beyond the build volume,

allowing it to access tool holders and other supporting infrastructure. The gray square in Figure 9 outlines the build area on the surface of one wall upon which the four Stewart platforms can cooperate to metabolize a feedstock molecule, recharge a spent molecular tool, or build an AP Block. The expected dimensions of the square are ∼20 nm from edge to edge, a design target for building blocks up to 15 nm in size.

The struts in Figure 9 have a threaded rod estimated to be about 49 nm long, and a threaded sleeve of about the same length. Rough analysis suggests this should be sufficient to access the region within the square. A more detailed analysis of the accessible working volume and the actual reaction sequences and metabolic sequences will be needed to verify that the range of motion of the proposed Stewart platforms is, in fact, sufficient both to execute the necessary mechanosynthetic reactions and to access the supporting tool racks and binding sites.

Now that we have one layout for the programmable fabricator, it is easier to envision alternatives. We might, for example, alter the various dimensions involved or substitute different types of robotic arms for the Stewart platforms. More extensive modifications should be possible. The field of robotics offers a rich set of possibilities.

4.4.3 Assembly operations within the fabricator

Although assembly operations are normally carried out by the internal assembly arm, limited assembly operations can be carried out by the fabricator, provided the finished assembly and all intermediates can fit into the available space within the fabricator. Because the fabricator has four Stewart platforms, it can more easily position multiple blocks for complex assembly operations. In addition, the fabricator might have greater positional accuracy than the internal assembly arm.

4.4.4 Molecular Tools

For a brief introduction to molecular tools, see the section on Molecular Tools (Section 3.7) in the Introduction (Section 3). The fabricator wields molecular tools to manufacture AP Blocks (or simply blocks when the context makes clear what we are referring to). Blocks are atomically precise. Baseline blocks are made from the elements H, B, C, N, O, Si, Ge, and Pt by mechanosynthesis at 210 K. Each such block might typically have a few hundred atoms. In the baseline proposal, we assume that blocks are limited to a maximum of 1,000 atoms and 15 nm in their longest dimension.

4.4.5 Binding sites

One or more binding sites are used to bind to and bring into the SRU any required feedstock molecules. The baseline assumes drawer-style binding sites rather than sorting rotors. The binding sites are located within easy reach of the programmable fabricator so that they can be operated by the fabricator arms, avoiding the need for further binary Bowden cables to control their operation. Feedstock molecules can be directly metabolized by the programmable fabricator. The baseline feedstock molecules, tetraethylgermane, pentanenitrile, triethyl borate, tetraethyl orthosilicate (TEOS), and (ethylcyclopentadienyl)trimethylplatinum(IV), provide all eight elements in our baseline palette. In the baseline, excess H and N are vented as H_2 and N_2 directly into the feedstock solution. Excess carbon would also be vented directly into the feedstock solution as C_2H_2 (acetylene) gas. Excess oxygen is vented as CO_2. Excess silicon is vented as tetramethylsilane. As a consequence, removal of H_2, N_2, C_2H_2, CO_2, and tetramethylsilane from the feedstock is required

during intake. If necessary to achieve reliable rejection of vented waste (and other contaminants), a multistage cascade could be used (see Binding site leakage (Section 7.2)).

4.5 Balancing element intake and outflow

Ingestion of feedstock which has a different ratio of our eight elements than the product being built might result in excess amounts of any of the eight elements. We need a strategy for dealing with this imbalance.

Our baseline elemental palette is H, B, C, N, O, Si, Ge, and Pt. Our baseline feedstock molecules are shown in Figure 19.

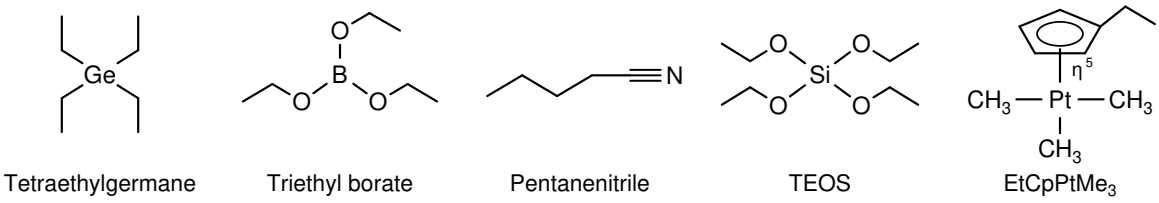


Figure 19: Baseline feedstock molecular structures.

For our baseline, we avoid an excess of Ge by ingesting only enough tetraethylgermane to provide for the actual Ge required in the product. If we need additional C or H after ingesting enough tetraethylgermane to satisfy our germanium requirement, we ingest some other feedstock molecule.

We avoid an excess of B by the same method. We ingest only enough triethyl borate to satisfy the actual requirement. If we need additional H, C, N, O, or Si, we ingest either pentanenitrile or TEOS (tetraethyl orthosilicate).

We avoid an excess of Pt by the same method. We ingest only enough $EtCpPtMe_3$ to satisfy the very limited actual requirement.

This means we might have an excess of or a deficiency of any of H, C, N, O or Si: the elements other than Ge, B, and Pt.

If we have a deficiency of any of these, we ingest additional pentanenitrile or TEOS until the deficiency has been remedied.

Having ingested an excess of pentanenitrile or TEOS to remedy a deficiency, we might have an excess of H, C, N, O, or Si.

If we have an excess of any of them, we vent waste molecules.

Excess H is dealt with by venting H_2 gas. Excess N is dealt with by venting N_2 gas.

We are now left with a possible excess of C, O, or Si.

Excess Si is dealt with by venting liquid tetramethylsilane. The carbon and hydrogen required come from pentanenitrile.

We now have a possible excess of C or O.

Excess O is dealt with by venting CO_2 gas. If we have a deficiency of carbon, we ingest pentanenitrile until the deficiency is dealt with. Excess nitrogen from this ingestion is dealt with by venting N_2 gas. Excess hydrogen from this ingestion is dealt with by venting H_2 gas. We ingest just enough pentanenitrile to cover the carbon needed for removing excess oxygen.

This leaves us with a possible excess of C.

Excess C is dealt with by venting C_2H_2 gas (acetylene, see Future research in electromechanosynthetic (EMS) reactions (Section 13.1)). If we need extra hydrogen to make the C_2H_2, we ingest pentanenitrile. Nitrogen from this ingestion is vented. Pentanenitrile is C_5H_9N. We dispose of the excess carbons by venting C_2H_2. Quantitatively, we vent five C_2H_2 for every two pentanenitriles ingested for carbon. Those five C_2H_2 require 10 hydrogens, while the two pentanenitriles provide 18. The excess 8 hydrogens are vented as H_2.

A detailed analysis of ingested and vented molecules would be needed to determine the amount of waste produced. Assuming a finite buffer of each element to reduce the immediate need to ingest feedstock molecules and vent waste would require a more detailed analysis. Optimizing the sequence of blocks manufactured to minimize the waste stream given the available buffer space would require a yet more detailed analysis. Nevertheless, the preceding analysis makes it clear that we can successfully synthesize product of whatever elemental composition might be required.

Qualitatively, the product is likely to have significantly less hydrogen than the feedstock. The feedstock molecules tend to be hydrogen rich, while structural elements tend to be larger with less surface available for hydrogen, and some of that surface will be occupied by B and N to facilitate dipolar bonds. Looking at proposed blocks for assembling the shell leads to the conclusion that we might also have a significant amount of excess carbon, as we would need to ingest significant amounts of carbon to satisfy the requirement for structural B and N in, for example, the hexagonal blocks in one proposal for the shell (see Decomposition of hexagonal diamond into hexagonal blocks (Section 7.4.4)).

The baseline design calls for the production of $\sim 5 \times 10^{21}$ atomically precise product atoms per day, equivalent to 100 mg/day if the average atomic mass is that of carbon. The total volume of feedstock to cover all 125 boards might be ∼one liter/board or 125 liters and possibly more. If we assume one kilogram/liter the total weight of feedstock might be 125 kg, or ∼one million times the weight of the product produced. Waste is likely to be roughly similar in magnitude to the product, ∼100 milligrams/day. This implies that after one day waste will constitute one millionth or less of the feedstock by weight, even with no attempt to filter it out. As outgassing will naturally remove H_2, N_2, C_2H_2, and CO_2 if we provide this opportunity in the system design, we are left with $C_4H_{12}Si$ (tetramethylsilane). If we periodically heat the feedstock to above the boiling point of tetramethylsilane, 299 K (near room temperature), outgassing can also be used to remove it. The lowest boiling point of any component is 391 K. Heating to ∼350 K should suffice to purge the feedstock of tetramethylsilane. Given that feedstock is filtered on intake by the SRU, relatively high levels of waste could be tolerated if that proved useful for the system design.

4.6 Using platinum to generate H_2 gas

A route for producing molecular hydrogen inside a nanofactory would be to use a small platinum collector surface as a passive recombination target for mechanosynthetically delivered hydrogen. A GeRad-H tool delivers individual hydrogen atoms to exposed sites on a platinum surface. The platinum surface then performs the only chemically nonpositional step required: two adsorbed hydrogen atoms migrate, encounter one another, form H_2, and desorb into the gas phase. The net reaction is therefore $2H^* \rightarrow H_2(g)$, where H^* denotes hydrogen adsorbed on the platinum surface.

The GeRad-H tool provides the atomically controlled transfer step; the platinum collector releases H_2 without positional control. Once hydrogen has been deposited on platinum, H_2 gas forms by conventional diffusive processes. Hydrogen adsorption, surface mobility, recombinative desorption,

and the catalytic competence of platinum for hydrogen reactions are all well-established features of platinum surface chemistry [CEP76, HTS14, C+16]. The SRU delivers hydrogen to a clean platinum target region, and H_2 gas is released.

The platinum is a catalyst rather than a stoichiometric reagent; the collector is not consumed during normal operation; it is merely cycled through hydrogen adsorption and desorption states. The output product, H_2, is volatile under the relevant vacuum conditions and so naturally leaves the surface, preventing permanent occupation of the catalytic sites. The process is therefore self-clearing in the intended direction: deposited hydrogen is converted into a stable gas molecule and removed from the active collector.

As the H_2 gas is released from an unspecified point on the Pt surface, the entire process must be confined to an isolated chamber on the SRU to prevent internal H_2 poisoning. The gas in this small chamber will have to be evacuated to the exterior of the SRU.

The method also has a useful failure mode. A missed GeRad-H stroke leaves the tool still charged, or fails to add hydrogen to the collector, rather than generating a highly reactive free radical in the work volume. An over-supplied collector merely increases local hydrogen coverage until recombination/desorption restores available surface sites. The principal engineering requirements are consequently: maintain a clean platinum surface, provide enough exposed area for the desired hydrogen throughput, avoid catalyst poisons, and vent the produced H_2. The latter step could be carried out by a positive displacement pump. It would not be necessary to pump down to a perfect vacuum over the platinum catalyst, only to provide a sufficient vacuum that transfer of hydrogen to the platinum surface was not blocked by the buildup of H_2 gas. The positive displacement pump would only need to expel excess H_2 gas.

Entry of GeRad-H into and removal of GeRad out of the platinum catalyst chamber would need to be done using a method that blocked exit of H_2 from the chamber into the eutactic environment (vacuum environment) of the SRU. Methods similar to those used during ingestion of feedstock to prevent entry of H_2 might be used.

The design is therefore modular, nonconsumptive with respect to the catalyst, and compatible with our baseline architecture.

The energetics of the known GeRad-H recharge sequence suggest a second and possibly more attractive route to mechanosynthetic hydrogen generation: replace the GeRad hydrogen acceptor in the recharge step with a low-coordinate platinum cluster. In the recharge reaction, the transferable hydrogen is bound to the HAbst-H(GeRad) tool by only ∼429 zJ (2.68 eV) [FM08a], while formation of GeRad-H is favored by ∼128 zJ (0.8 eV). Literature Pt–H energetics indicate that low-coordinate Pt atoms and small Pt clusters can bind H strongly enough to make the analogous transfer thermodynamically favorable in some cases. However, we do not have DFT verification of the reaction, including checks for Pt–C or Pt–Ge side bonding.

The gas-phase Pt–H bond is strong: a recent PtH-specific calculation gives $D_0(\text{Pt–H}) = 329.6 \pm 3.9$ kJ mol^{-1} (∼548 zJ (3.42 eV)) [Iri22]. This value places Pt–H in the same general range as the Ge–H bond in the GeRad-H tool (∼562 zJ (3.51 eV)), and stronger than the weak donor C–H bond in the HAbst-H(GeRad) tool used in the GeRad-H recharge sequence, ∼429 zJ (2.68 eV) [FM08a].

In principle, a single platinum atom might suffice to support the generation of H_2 gas.

We use the platinum collector method in the baseline for two primary reasons. First, the generation of H_2 gas from platinum is well established. Second, once we ingest a small amount of platinum-containing feedstock and build the collector, no continuous platinum supply is required. The disadvantages are as follows. First, the development of the platinum ingestion and metabolism

reactions requires more work. Second, we would need to develop a suitable platinum tool and develop and validate the reactions used to manufacture the small platinum collector surface using said tool. Third, adsorption by the platinum surface of H from a GeRad-H or HAbst-H(GeRad) tool needs to be validated.

4.7 Assembly

4.7.1 Internal assembly arm

The internal assembly arm assembles AP Blocks. We assume that such blocks must be positioned with six degrees of freedom to be properly assembled. We also assume that they must be assembled in vacuum. This means that blocks can be "fused". That is, blocks can have highly reactive surfaces that bond to other blocks when brought into contact. Because assembly is done inside the SRU shell, and the inside of the SRU shell is a vacuum, this condition is satisfied.

Assembly operations can also take place within the fabricator, provided the finished assembly and any intermediates can fit inside the fabricator. This allows more positional devices to be used in the assembly operation.

4.7.1.1 Range of motion of internal assembly arm

The range of motion for the internal assembly arm was estimated by assuming that it would be sufficient if each of its struts were able to move the full 100 nm range of the extrusion port plus an additional 20 nm past its far edge and another 20 nm preceding its near edge. This leads to a range of motion for each strut of 20 nm + 100 nm + 20 nm = 140 nm. As each strut consists of two telescoping tubes, each of which must be at least 140 nm long, then the length of each strut from the base to the platform, including a 40 nm margin, would have to be $2 \times 140 + 40 = 320$ nm. The margin was chosen rather arbitrarily. An additional 10 nm was allocated for the drive mechanism at the base-end of the strut. This gives the totals shown in the layout.

The tip of the platform of the internal assembly arm, if the platform was small, would not be able to move beyond the rectangular outline of the SRU shell into the extrusion port, which would be a significant limitation. To mitigate this, the platform, as shown, has been extended in length so that its "long nose" can extend into the extrusion port beyond the rectangular outline of the shell.

The range of motion of the internal assembly arm extends to the programmable fabricator, allowing the fabricator to hand off blocks to the internal assembly arm, which can then add these blocks to the daughter SRU being extruded through the extrusion port.

4.8 Daughter Placement

4.8.1 An SRU extension mechanism

At a high level, the extrusion mechanism consists of an extrusion sleeve (shown in pink in the illustration of the SRU layout (Section 4.1.1)) through which daughter SRUs are extruded. The extrusion sleeve itself has inner dimensions 100 nm × 60 nm, matching the two shortest exterior dimensions of the SRU itself. When an SRU is first extruded from its parent, the extrusion sleeve is not extended and, in fact, most of it remains to be built.

As a 60 nm × 100 nm area is only 0.6 aha, atmospheric pressure will only exert ∼0.6 nN of force on the daughter SRU as it emerges from the parent SRU (see The force exerted by atmospheric pressure at the nanoscale (Section 18.3)).

Overcoming this pressure force is a requirement that the baseline proposal should be able to meet.

After the SRU has been extruded and placed on the die surface, its first task is to manufacture and extrude the extrusion sleeve, through which daughter SRUs will be extruded. The extrusion sleeve extends at least 225 nm beyond the shell of the SRU.

The extrusion sleeve is $\geq$ 225 nm long because it must be at least as long as the $\sim$225 nm external length of the contracted external assembly arm, plus whatever allowance might be required to ensure that whatever is extruded next completely blocks the extrusion sleeve once the base of the EAA has cleared the end of the ES. If the extrusion sleeve is not long enough, the base of the EAA will be clear of the ES, allowing feedstock to pour into the interior of the SRU from the far end of the ES before its near end can be blocked. This is more easily understood by viewing the SRU layout (Section 4.1.1).

We will need a mechanism to extrude both the extrusion sleeve and also the daughter SRU. We consider such mechanisms below.

4.8.1.1 A screw extrusion mechanism

The baseline assumes a screw mechanism to extrude both the extrusion sleeve from the parent and the daughter from the extrusion sleeve.

If we are extending the daughter from the extrusion sleeve by this mechanism, we would conceptually press a screw into the shell of the SRU, imprinting it with the complementary shape of the screw pressed against it. In reality, of course, we would manufacture the shell with this shape already built into it, but conceptually it's easier to describe as forming the shell by pressing a screw into it and thereby shaping it. If we do this at several sites around the shell, we would then be able to apply a distributed force at many places around the shell of the daughter by rotating a series of screws pressed against the shell. Rotating the screws synchronously would move the daughter outward through the port and out of the parent.

Having once extended the daughter out of the port by some small distance, perhaps a few nanometers, the parent could then build an additional layer of shell, including embedded screw threads, into the newly built piece of the shell of the SRU. Additional components of the daughter could also be added as was appropriate.

This "build as you go" method would let us build an entire SRU, including the shell, in small incremental steps.

The same method could be used to extrude the sleeve from the parent.

While the imprint of the screws would mean the shell of the extruded daughter would not be perfectly smooth, it would still not leak. Imagine pressing a screw against a piece of clay, but only hard enough that the ridges of the screw leave marks upon the clay. The valleys in the clay are isolated, and if the hardened clay were pressed against a flat rubber surface, liquid in one valley would not be able to leak into another such valley.

4.8.2 External Assembly Arm (EAA)

The external assembly arm is a two-stage Stewart platform [BCK19] stored in the SRU in fully contracted (not extended) form. In the fully contracted form in which the two-stage Stewart platform is manufactured, all twelve struts of both Stewart platforms are shortened as much as possible to little more than half of their fully extended length, and the second stage Stewart platform is folded back to overlap with the first stage Stewart platform. The requirement for the

EAA is that it reach at least to the next site, lengthwise, from the parent SRU, and that it be able to pass daughter SRUs hand-over-hand. The baseline uses a 600 nm × 200 nm site grid. The EAA must reach the daughter's grip point at the intended neighboring sites, including diagonal neighbors, and support hand-over-hand transfer. Corresponding positions on adjacent sites are 600 nm apart lengthwise, 200 nm apart widthwise, and approximately 632 nm apart diagonally. The exact arm extension also depends on the mounting and grip-point offsets and the required motion envelope.

The baseline design target is the 14-site placement pattern in Figure 20, including adjacent sites lengthwise and four rows widthwise. Four widthwise site pitches correspond to 800 nm between corresponding site positions. This specifies the required placement capability; detailed kinematic validation remains necessary. The parent is shown in blue; the sites where the parent can place the daughter are shown in green.

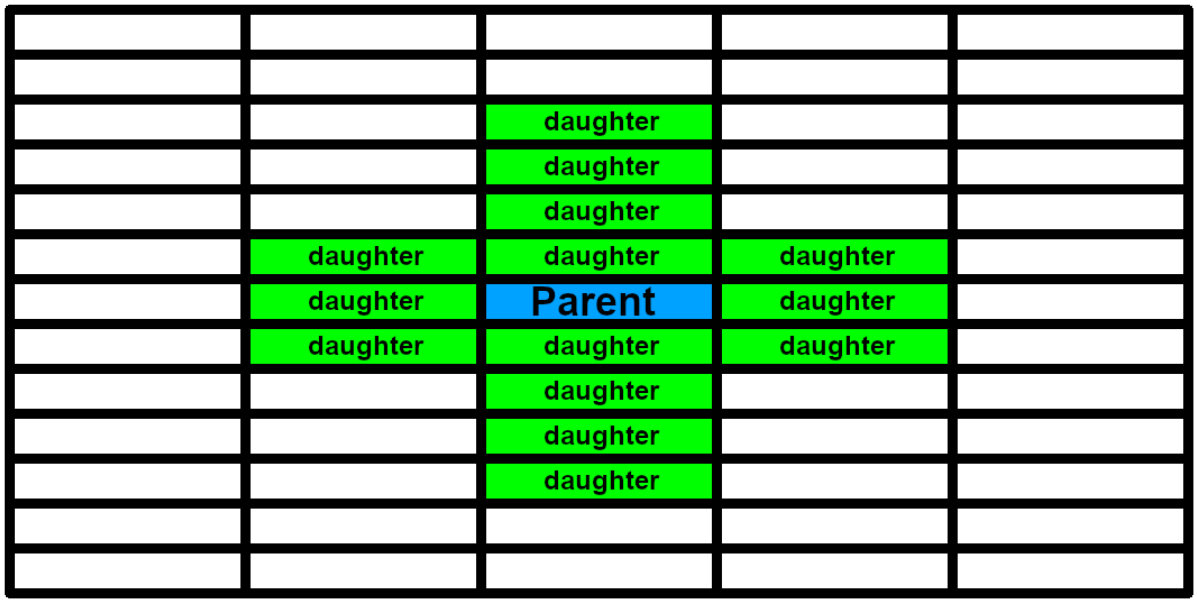


Figure 20: Baseline placement target for the External Assembly Arm of a parent SRU.

4.8.3 Extruding the daughter

After the parent SRU extrudes the daughter SRU (Figure 11(a)), the parent SRU initiates manufacture of the daughter's EAA.

Manufacture of the daughter's EAA causes continued extrusion of the daughter SRU until the daughter SRU fully emerges from the ES, which has, until this point, surrounded part of the daughter SRU and held it in position. As the daughter SRU finally emerges from the ES (pushed out by the daughter's EAA), it is gripped by the parent's EAA and held in position until extrusion of the daughter SRU is complete. At this point, the daughter SRU can be positioned by the parent's EAA in whatever position might be desired (Figure 21).

A possible concern is that the flow of liquid feedstock will create forces on the daughter SRU that prevent it from being successfully positioned. Approximating the daughter SRU as a cylinder 500 nm long and 50 nm in radius gives a lateral force of order 1 pN or less under the assumed feedstock-

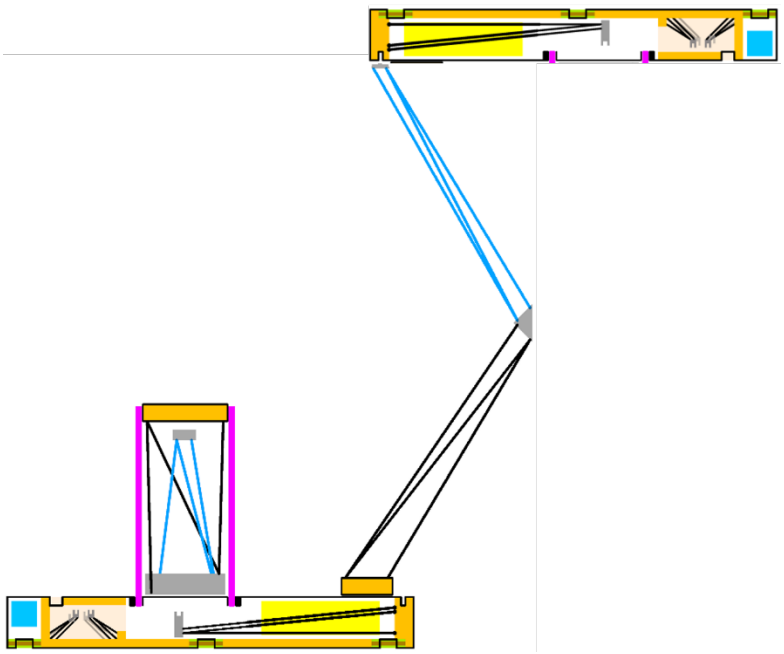


Figure 21: External Assembly Arm and extrusion sleeve during daughter placement.

flow conditions. This is small compared with the approximately 1 nN available actuation force (see Force on daughter from fluid flow (Section 18.4)).

The illustration in Figure 21 shows the SRU after it has extruded its ES (shown in pink), which extends 225 nm beyond the shell of the parent SRU and surrounds the just manufactured daughter's EAA. Part of the base of the daughter's EAA protrudes above the ES. The daughter SRU has been extruded and is held by the parent's EAA.

Once the daughter's EAA has been manufactured, but while it is still in the ES, the parent can position the daughter SRU over the extruded base of the daughter's EAA and attach the EAA base to the daughter SRU.

There is a risk that in an array of closely packed parent SRUs on the die surface, there might not be enough room for all the daughter SRUs to be aligned with and attached to their EAAs at the same time. Should this risk be realized, a mitigation would be for the parent SRUs to be divided, for example, into six groups, with each group taking its turn to attach its daughters' EAAs. In this way, each group would have plenty of room because when one group was swinging its daughter SRU around, any adjacent parent SRUs that might cause a conflict would be holding their daughter SRUs out of the way.

Admittedly, this would mean only one parent SRU in six would be attaching its daughter's EAA at a time, so this operation would take six times as long, but this is acceptable for a risk mitigation strategy because this operation is expected to use only an insignificant percentage of the total replication time.

In summary: after extruding the ES and manufacturing both the daughter SRU and the daughter EAA, the parent SRU attaches the daughter EAA to the daughter SRU. After attachment, the EAA is controlled by binary Bowden cables that run through the shell at the point of attachment (see Umbilicals (Section 6.11.1)).

Because there are two grip points, a daughter SRU can be passed from one parent to an adjacent

parent, with each parent gripping a different grip point on the daughter SRU during the handoff. This allows parents to pass daughters hand-over-hand, with alternating grip-points at each handoff.

When an SRU is to be placed onto the die, it is necessary to extend its feet and drive the electrical spikes in the feet into the contact pillars. This requires power, at a point in time when the daughter SRU is not yet connected to power. The adoptive parent is the SRU that places the daughter SRU into an open site; it might be remote from the manufacturing parent. The baseline assumes that the daughter supplies an umbilical, which the adoptive parent connects to itself to provide power.

Once the daughter EAA is attached to the daughter SRU, the assembled result is ready to either be (1) attached to the die surface or (2) passed hand-over-hand to an open site, there to be attached to the die surface.

4.8.4 Placing the daughter on the die

The simplest placement method would use an adjacent SRU, already aligned to features on the die, to position the daughter relative to its own known position. The accumulation of positional errors would be prevented by having the daughter, following its own attachment to the die surface, determine its own position with respect to alignment features on the die surface prior to assisting in the placement of any subsequent adjacent SRUs onto the die surface. See also A Local Replication and Routing Algorithm (Section 6.4) and Feet (Section 4.3.2).

5 Shell issues

Read Section 4.2 (Shell) first; it describes the SRU shell. This section examines specific shell issues in greater detail.

5.1 SRU Shell

See Decomposing large structures into small blocks (Section 7.4) for more details about possible molecular designs of blocks used to make the shell.

The SRU shell is a barrier between the feedstock solution outside and the eutactic environment (vacuum) inside. It is made from AP Blocks made by the SRU Fabricator that are assembled during the extrusion of the daughter SRU by the parent. It is extruded through the parent's extrusion sleeve, which uses a sliding seal to prevent entry of feedstock during the extrusion process. We assume the sliding seal between the shell and the extrusion sleeve prevents entry of undesired contaminants including H_2 gas.

We will consider three types of blocks from which to make the shell.

1) The first kind of block might be termed molecular LEGO blocks. They are modeled after LEGO bricks. See Molecular LEGO (Section 7.6). Like LEGO bricks, molecular LEGO blocks are rectangular and connect on their top and bottom faces, here via dipolar bonds.

2) The second kind of block we consider is also rectangular, but instead of having dipolar bonds on its top and bottom, it has dipolar bonds on the front, the back, the left, and the right faces. The top and bottom do not have linking groups. These blocks can form a sheet. See BNO Molecular LEGO that bond on four sides (Section 7.9).

3) The third kind of block is hexagonal. It has six sides, all of which can form dipolar bonds, and a top and bottom that do not form links to other blocks. See Decomposition of hexagonal diamond into hexagonal blocks (Section 7.4.4).

We discuss these three types of blocks in turn.

5.1.1 Blocks with dipolar links on top and bottom

An illustration of stacking LEGO bricks into a shell is shown in Figure 22 (the structure and the LEGO bricks are not to scale).

Each brick is held in place in the shell by many dipolar bonds on its top and bottom (see Exemplar BN Molecular LEGO Blocks (Section 7.7)).

5.1.2 Blocks with dipolar links on four sides

A wall can be made from blocks that have dipolar bonds on four sides; see BNO Molecular LEGO that bond on four sides (Section 7.9).

5.1.3 Hexagonal blocks with dipolar links on six sides

An attractive design alternative is the use of hexagonal blocks; see Decomposition of hexagonal diamond into hexagonal blocks (Section 7.4.4).

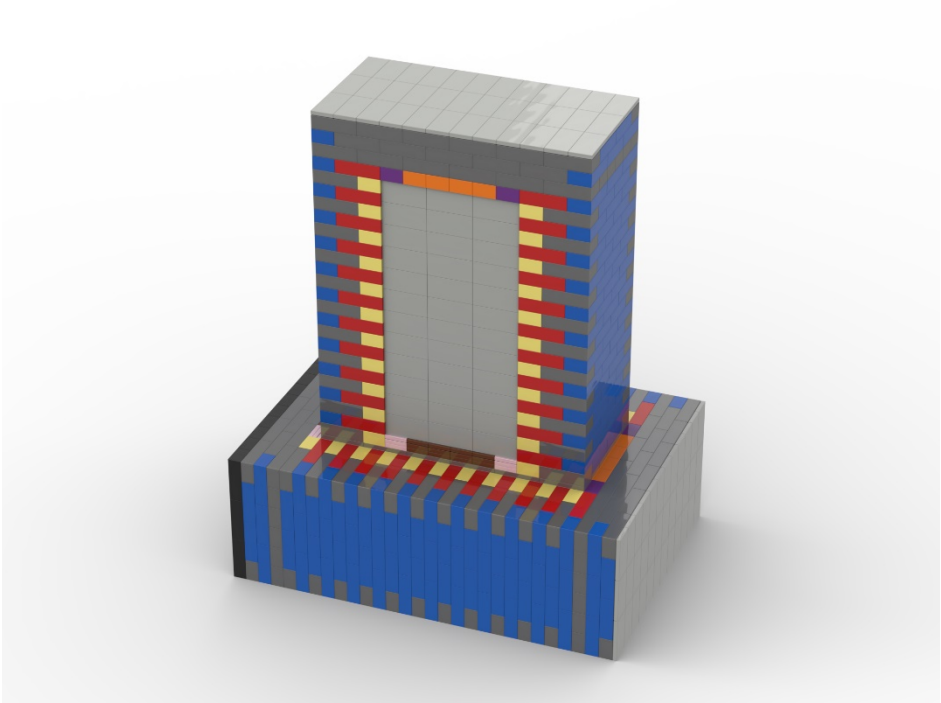


Figure 22: LEGO model of a shell assembled from LEGO bricks. This rendering shows the parent SRU extruding a daughter SRU, with the top surface of the granddaughter SRU occupying the extrusion port of the daughter SRU.

5.1.4 Other linking groups

We could also assemble blocks in the shell by linking them with van der Waals forces, dovetail joints, hooks, snaps, loops, nuts and bolts, and the like. Fusion of surfaces (such as the diamond (110) surface and possibly others) would also be possible, though the difficulty of reversing such fusion operations might introduce a requirement for a lower error rate in their assembly. In addition, germanium radicals form relatively weak bonds to each other and could be used in assembly operations that bond blocks together.

5.1.5 Summary

At the moment, the leading candidates for blocks from which to assemble the shell use dipolar bonds.

6 Control issues

This section should be read after reading the Control (Section 4.3) section under The Simple Replicating Unit or SRU (Section 4). This section covers specific issues surrounding control in greater detail.

6.1 Broadcast Architecture

6.1.1 Introduction

SRUs mounted on a die receive broadcast instructions through the contact pillars on the die. Broadcast instructions are, by definition, identical for all SRUs. There are variations on the broadcast architecture in which there are different blocks, with each block having its own computer broadcasting instructions to its own SRUs in its own block, with different computers broadcasting different instructions to different SRUs in different blocks. This approach is not in the baseline.

We outline a baseline set of instructions that an SRU should be able to decode and execute. As we want the SRU to be as simple as possible, the instructions that the SRU can execute are intended to be implementable on minimal hardware, while still providing adequate flexibility to enable the SRU to carry out its primary functions.

Each SRU is placed on a die. A large part of the die's surface is covered with a repeating pattern of SRU sites, each of which has six contact pillars, the six contact pillars being V+, Ground, S1, S2, D1, and D2. V+ is the positive DC voltage. Ground is 0 volts. S1, S2, D1 and D2 are data signals that vary between Ground and V+. Together, these six lines power and control four differential comb drives, which in turn drive a demux on the SRU that controls all aspects of the SRU's function. S1, S2, D1 and D2 are connected to the control inputs of the four differential comb drives that drive the mechanical demux. The S1 and S2 inputs cause the demux to select the Next or Previous output pair by using four phase clocking. The D1 and D2 inputs drive the selected output pair of binary Bowden cables either forward or backward, again by using four phase clocking.

That is, the six contact pillars provide power and control for four electrostatic differential comb drives. These four differential comb drives, along with their respective binary Bowden cable outputs, are called S1, S2, D1 and D2. S1 and S2 are a pair of binary Bowden cables that select an output, while D1 and D2 are a pair of binary Bowden cables that drive the selected output. The selected output is also a pair of binary Bowden cables.

S1 and S2 can be referred to collectively as S. D1 and D2 can be referred to collectively as D.

6.1.2 The Instruction Set

Our instruction set is oriented around the SRU. The SRU can be viewed as a box run by several Stewart platforms, and each Stewart platform is run by several struts. Our instruction set selects one strut from one Stewart platform, and then either extends it or retracts it. It then selects another strut (possibly from a different Stewart platform) and extends or retracts this second strut. It then selects a third strut and extends or retracts it. And so on. It keeps this up repeatedly, with all the Stewart platforms in all the SRUs running in lockstep, and one strut in each SRU being extended or retracted at a time.

Select a strut. Extend or retract. Select another strut. Extend or retract. That's pretty much it. Some struts are in Stewart platforms that carry out mechanosynthesis. Some struts are in Stewart platforms that carry out assembly of AP Blocks. Some struts are in Stewart platforms that take the finished daughter SRU and remove it from the parent. Some struts run the binding sites that bring in feedstock. It's all struts being extended and retracted.

Sometimes a subset of the SRUs will have to do something different. When this happens, all the other SRUs are turned OFF and only the selected SRUs continue to function. The selected SRUs do whatever they have to do, then they carefully position themselves back to exactly where they were when they diverged from their fellow SRUs. Then, everyone turns ON and continues forward, again with everyone in lockstep.

That's the general idea. We now look in more detail at our baseline realization. What are the instructions in the instruction set?

6.1.3 Stepper S (Select)

Stepper S can be stepped forward or backward in single-step increments. It is used to select among some number N of outputs. The outputs are numbered from 0 to $N-1$. Instructions available to Stepper S are:

Step S forward one output (A.K.A. Select Next Output).

Step S backward one output (A.K.A Select Previous Output).

If the output is currently set to 0, and the "Step S backward one output" instruction is given, the output will remain at 0 and the SRU will be set to the "ON" state.

If the output is currently set to $N-1$, and the "Step S forward one output" instruction is given, the output will remain at $N-1$. If a Test Unit has been loaded into the Test Station and the Test Unit is TRUE, then the SRU will enter the "OFF" state; otherwise, the SRU will remain in the state that it is in.

The baseline instructions map directly onto the implementation of the differential comb drives connected to the demux. Figure 23 shows a demux for a small example where $N = 8$. Each of the four input binary Bowden cables to the demux is driven by its own differential comb drive. The four input binary Bowden cables are called S1, S2, D1 and D2. The 16 output binary Bowden cables are divided into 8 pairs, Output Pair 0, Output Pair 1, … Output Pair 7. In the baseline system, N is ~60 (60 pairs of binary output Bowden cables, as discussed in Section 4.3.6.1).

6.1.4 Stepper D (Drive)

Stepper D can likewise be stepped forward or backward in single-step increments. It is used to drive the single selected output forward or backward. A single output might be used to drive a single degree of freedom of a robotic device, such as a fabricator or assembly arm. Instructions available to Stepper D are:

Step D forward one step (A.K.A. Drive Current Output Forward).

Step D backward one step (A.K.A. Drive Current Output Backward).

If the given output is already at the forward limit of its range of motion and the "Drive Current Output Forward" instruction is given, the driven output does not move forward.

If the given output is already at the backward limit of its range of motion and the "Drive Current Output Backward" instruction is given, the driven output does not move backward.

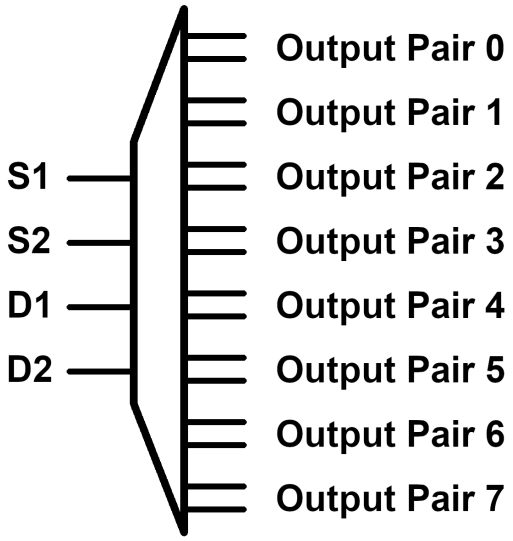


Figure 23: Demultiplexer with four binary Bowden cable inputs and eight output pairs.

6.1.5 RESET

To restore the system to a well-known state we need to have a RESET instruction. The baseline method of synthesizing a RESET instruction from the four instructions already described is to first Select Previous Output N times, guaranteeing that the Selected Output is 0. This would reset the Select Output stepper motor and turn the SRU "ON".

The method of resetting each individual driven output would depend on the driven output but would typically consist of repeatedly executing Drive Current Output Backward enough times to drive the selected degree of freedom to the end of its range of motion.

6.1.6 Turning the SRU ON or OFF

The SRU is in one of two modes: ON or OFF. In the ON mode, the SRU executes instructions as it normally does: both Stepper S and Stepper D are active and respond to instructions. In the OFF mode, Stepper S continues to respond to selection instructions, but Stepper D is disabled. Consequently, the fabricator, assembly arms, and other driven mechanisms do not move. The SRU returns to ON when a backward-selection instruction is issued with output 0 selected.

6.1.7 Turn SRU ON

If the output is currently set to 0, and the "Step S backward one output" instruction is given, the output will remain at 0 and the SRU will be set to the "ON" state.[6]

This instruction can be executed in either the ON or OFF mode. In the ON mode, this instruction is a no-op, as the SRU is already ON. In the OFF mode, this instruction turns the SRU ON.

Anyone familiar with the various supervisor states in current operating systems will recognize this as the simplest and most basic of possibilities for providing a mechanism for multiple operating states. We adopt this mechanism primarily because implementing it is remarkably simple compared

[6]With thanks to James MacArthur for this proposed implementation.

with possible alternatives for conditional control. In the computer industry the SIMD (single instruction stream, multiple data stream) architecture has been known since at least the **ILLIAC IV** (which the author programmed during a summer in the 1970s). It involves many processing elements which benefit from a single common instruction decode unit, but which are capable of either carrying out the decoded instruction or doing nothing depending on a binary condition. This architecture greatly simplifies the processing elements while still retaining most of the flexibility of a conventional computer in data handling. By using a similar architecture in the SRU, we can eliminate the need for an onboard computer and retain only enough local logic to enable or disable execution of broadcast drive instructions while continuing to execute selection instructions. A single bit in the SRU determines whether Stepper D responds to drive instructions; Stepper S remains operational in either mode. This might be viewed as a minimalist implementation of the SIMD concept. In this implementation the hardware requirements for the SIMD architecture boil down to (a) implementing a one-bit memory that can be turned "on" or "off" by testing a condition and (b) an ability to enable or disable the "D" driver output of the demux based on whether the one-bit memory is "on" or "off".

6.1.8 IF <Test Unit> Turn SRU OFF

We now need some instruction that will conditionally turn OFF an SRU.

At this point, we divide the process of testing a condition into two parts. In the first part, we test the condition using a binary Test Unit. The binary Test Unit has two stable states (TRUE and FALSE) separated by a potential barrier. It can transition from one stable state to the other by either being pulled on with sufficient force or being pushed on with sufficient force. "Sufficient force" is likely something in the range of a few nanonewtons, and the distance moved might be anywhere from 0.1 to 2 nm, depending on the specifics of the condition being tested for. A cartoon illustrating a simple concept for a Test Unit is shown in Figure 24. Note that van der Waals forces will keep the sliding block at one end or the other end of its throw once it has moved there.

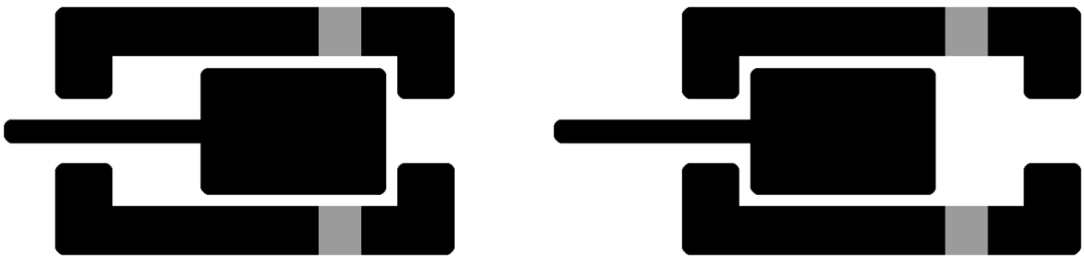


Figure 24: Conceptual bistable Test Unit.

Each Test Unit can be custom designed for the specific condition being tested for, so that the force and displacement required to overcome the barrier can be tuned for the specific condition being tested. However, all of them must have a standard interface to be used by the Test Station when interrogating whether the Test Unit is TRUE or FALSE, as seen in Figure 25, where a probe rod can be used to interrogate the state of the Test Unit.

The Test Unit can be picked up by an assembly arm or fabricator using a molecular gripper (Figure 26), positioned, and used to test something. It will then remain in one of its two stable states reliably in the face of thermal noise and other disturbances (such as might occur during normal handling) until it can be read out in the Test Station to conditionally turn the SRU OFF.

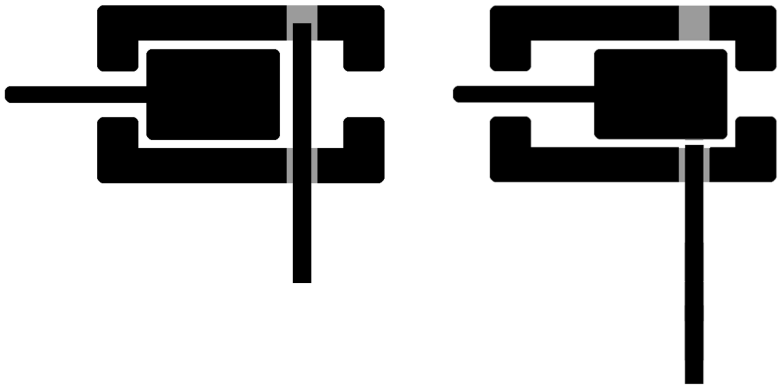


Figure 25: Probe readout of a Test Unit.

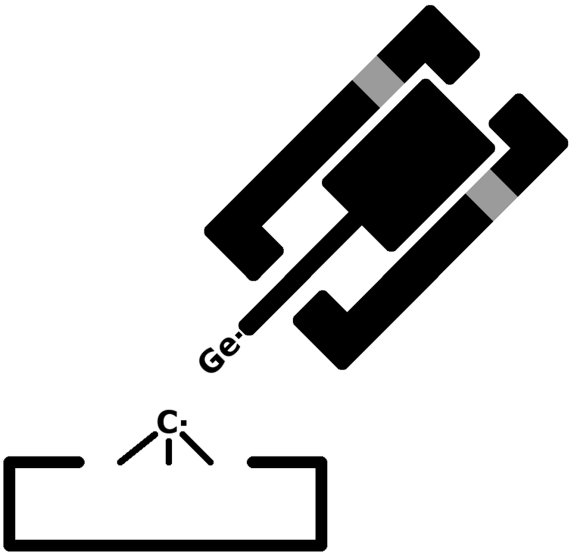


Figure 26: Conceptual Test Unit held by a molecular gripper.

A single SRU has a maximum power draw (P_{max}) of one nanowatt, or 10^{-17} J (10 aJ) per clock cycle at our chosen 100 MHz clock frequency (see SRU energy usage (Section 10.7)). If we want to set our Test Unit in one clock cycle, we must live within that 10 aJ energy budget. We could (if necessary) spend several cycles to set the Test Unit. Breaking a single C-C bond requires about 0.55 aJ, so the 10 aJ budget seems reasonable.

Once the Test Unit has been set (either TRUE or FALSE), we can use it in one of two instructions to conditionally turn off the SRU:

If <Test Unit> Turn SRU OFF

If NOT <Test Unit> Turn SRU OFF

Strictly speaking, we only need one instruction. For the sake of flexibility (and the current exposition) we assume the availability of two instructions.

Perhaps the simplest conceptual implementation would be to place the Test Unit into a "Test Station", a mechanism that conditionally turns the SRU OFF based on the condition of the Test Unit. We can imagine the Test Station as a jig that holds the Test Unit in place when we try to turn the SRU OFF. If the process of turning OFF the SRU involves a rod driven at right angles to the Test Unit, much as in rod logic, then the Test Station can test the state of the Test Unit using this rod. If the Test Unit blocks the movement of the rod that turns the SRU OFF, then the SRU won't turn OFF, and will instead remain ON. If the Test Unit fails to block the rod, then the SRU will be turned OFF. This might have an energy efficiency like rod logic. As we noted, we have enough energy to be wasteful in this testing process. Simplicity in this process seems (at least at the moment) more valuable than energy efficiency, so we choose simplicity.

Complex logical queries involving multiple binary conditions can be handled by using multiple Test Units, testing multiple binary conditions, and then reading out the multiple Test Units in the Test Station to determine what action to take next. In essence, the use of multiple Test Units gives us flexible access to Boolean logic.

While slow and inefficient, this kind of logic does not require that we build any kind of logic units, computers, computational devices or the like. The Test Units themselves can be quite simple. The Test Station directly turns the SRU OFF based on the condition of the Test Unit. As our need for logic operations is limited, this limited capability should be sufficient.

A variety of options for speeding up testing are possible if this should become a rate-limiting step in fabrication of AP Blocks. They are not part of the baseline.

6.1.9 Computing NOT

Suppose, for example, that Test Unit C has been set with some result, and we wish to set Test Unit D with the negation of the logical value in Test Unit C. We might proceed as follows:

Turn SRU ON.

Execute a series of instructions that would set Test Unit D to the value TRUE. This could be done by applying appropriate mechanical force to Test Unit D using either an assembly arm or a fabricator arm.

Then, execute the instruction "If NOT Test Unit C Turn SRU OFF".

If Test Unit C was TRUE, then the SRU remains ON and continues to execute instructions. If, however, Test Unit C was FALSE, the SRU is now OFF.

Now execute a series of instructions that sets Test Unit D to FALSE. This will only happen if Test Unit C was TRUE.

Turn SRU ON.

Test Unit D is now set to the negation of the logical value of Test Unit C. If Test Unit C was TRUE, Test Unit D is FALSE. If Test Unit C was FALSE, Test Unit D is TRUE.

6.1.10 Computing AND

Likewise, we could set Test Unit E to the logical AND of Test Unit C and Test Unit D by using the following series of instructions:

Turn SRU ON.

Execute a series of instructions that sets Test Unit E to FALSE.

If NOT Test Unit C Turn SRU OFF.

If NOT Test Unit D Turn SRU OFF.

Execute a series of instructions that sets Test Unit E to TRUE. These instructions will only be executed if both C and D are TRUE.

Turn SRU ON.

Test Unit E is now set to the logical AND of Test Unit C and Test Unit D.

6.1.11 Universal Boolean Computing

As the combination of AND and NOT is universal, we can compute any Boolean function.

6.1.12 Enabling Simple Closed Loop Mechanosynthesis

As this set of instructions can implement any Boolean function, it should be sufficient to implement Simple Closed Loop Mechanosynthesis (Section 6.9).

6.1.13 Summary of Instructions

Step S forward one output (A.K.A. Select Next Output).

Step S backward one output (A.K.A Select Previous Output).

Step D forward one step (A.K.A. Drive Current Output Forward).

Step D backward one step (A.K.A. Drive Current Output Backward).

Synthesized instructions:

RESET: To reset the Output Selector, “Select Previous Output” N times. For each output, repeat “Drive Current Output Backward” enough times to drive that degree of freedom to the end of its range of motion.

Turn SRU ON: Execute Select Previous Output when Output is 0.

IF <Test Unit> Turn SRU OFF: Place <Test Unit> in Test Station. Execute Select Next Output when Output is $N - 1$. If <Test Unit> is TRUE, then the state of the SRU is set to “OFF”; otherwise, it remains unchanged.

IF NOT <Test Unit> Turn SRU OFF: Place <Test Unit> in the Test Station and use the inverted test. If <Test Unit> is FALSE, set the SRU state to "OFF"; otherwise, leave it unchanged.

In Section 6.9 we describe how the baseline instruction set might be used to improve the reliability of a mechanosynthetic hydrogen abstraction reaction by using error checking and correction.

6.2 SRU Failure Handling

6.2.1 Failures caused by radiation damage

Drexler [Dre92] analyzed radiation failure in eutactic systems, and provided an approximate rate of damage of 10^{15} hits/(kg · rad), and an approximate damage rate, assuming background radiation of 0.5 rad/year, of 1.5×10^7 hits/(kg · s). An SRU with 10^8 carbon atoms has $10^8/(6.02214076 \times 10^{23}\,\mathrm{mol}^{-1}) \approx 1.66 \times 10^{-16}$ moles of atoms per SRU. There are 0.012 kg of carbon atoms per mole, implying one all-carbon SRU has a mass of $(1.66 \times 10^{-16}\,\mathrm{mol})(0.012\,\mathrm{kg/mol}) \approx 1.99 \times 10^{-18}\,\mathrm{kg}$ per SRU. Such an SRU would suffer $(1.99 \times 10^{-18}\,\mathrm{kg})(1.5 \times 10^7\,\mathrm{hits/(kg \cdot s)}) \approx 3.0 \times 10^{-11}$ hits per SRU per second. Multiplying by 86,400 seconds per day gives approximately 2.6×10^{-6} hits per SRU per day. Over one 10^5 second replication cycle, the corresponding rate is approximately 3.0×10^{-6} hits per SRU, about 330 times below the allowable failure probability of 0.001 per replication cycle. This hit rate implies that $2.6 \times 10^{-6} \times 5.8 \times 10^{13} \approx 1.5 \times 10^8$ SRUs in a baseline nanofactory would be hit every day. Under the deliberately pessimistic assumption that every hit causes failure, approximately 1.5×10^8 SRUs would therefore fail each day. This is a carbon-equivalent estimate; the result scales with the actual mass of the SRU and therefore with its elemental composition.

There are two reasons to believe this estimate overstates the failure rate. First, it assumes radiation errors occur randomly, when in fact radiation hits are distributed not at random, but along particle tracks. This clustering means that an SRU, if it is hit, is more likely to be hit a second time, but is also less likely to be hit at all.

Second, we assume any damage at all causes the SRU to fail. It should be clear that many types of damage will not interfere with SRU function, and the SRU will continue operating. We have not analyzed such situations because the analysis would be complex. Once we actually build SRUs, we can experimentally determine their actual failure rate, and this experimental result will include SRUs that have received some damage, but which continue to function correctly despite this damage.

It should be possible to drive unrecoverable thermal errors that occur during mechanosynthesis below this rate of 2.6×10^{-6} radiation hits per SRU per day (see Simple Closed Loop Mechanosynthesis (Section 6.9)), and the same should be true of failures caused by leakage of contaminants into the internal eutactic volume of the SRU (see Leakage of H_2 through a sliding seal (Section 7.4.4.1)). The likely result of future evolution of nanofactories is failure rates that approach the error rate dictated by radiation damage, primarily because our ability to conveniently and inexpensively reduce radiation damage appears to be limited. Shielding against all background radiation seems difficult. At the same time, such rates appear to be low enough that they do not appear to seriously impact the economics of a nanofactory, and a variety of approaches for fault tolerance are available.

6.2.2 The target failure rate

While the radiation-induced failure rate of $\sim 3.0 \times 10^{-11}$ hits per SRU per second is interesting, does it have any bearing on the failure rate we should target? For the baseline, we would accept

any failure rate that might be characterized as "tolerable". Spending resources on reducing the failure rate below what is necessary does not seem like the preferred initial strategy, even if, in the long term, we might seek failure rates similar to the radiation failure rates.

What failure rate is "tolerable"? It will take ∼46 generations, starting from a single SRU, to produce the 5.8×10^{13} SRUs in a nanofactory. To populate one die takes ∼30 generations. A one percent failure rate implies that 30 generations will leave a die with $0.99^{30} \approx 74\%$ functioning SRUs. After 46 generations, $0.99^{46} \approx 63\%$ of the expected SRUs will be functioning. While this would be tolerable if we could replace failed SRUs, in the baseline we're assuming fail-stop without replacement. A 1% failure rate seems too high under these conditions. If we target a 0.1% failure rate, or one in a thousand SRUs fails during a replication cycle, then ∼97% of the SRU sites on a die will be occupied by functioning SRUs. If we consider 46 generations, we get ∼95.5% (though we can reasonably assume that each die begins with a hand-picked functioning SRU).

A failure rate of 0.001 per SRU per replication cycle seems tolerable.

If one SRU in 1,000 can fail per replication cycle, this can also be expressed as a failure rate of 10^{-8} failures per SRU per second, or 0.0864% SRU failures per day. This can also be viewed as a success rate of 0.999136 per day. Over the course of a year, this results in $(0.999136)^{365} \approx 73\%$ survival of the SRUs in a nanofactory. This might be characterized as "barely tolerable", and confirms it as our target failure rate, on the theory that "barely tolerable" is a reasonable goal for the first nanofactory. Once we achieve this "barely tolerable" error rate, we can then work on reducing it into the "tolerable" and then "acceptable" range.

This "barely tolerable" error rate should become quite acceptable once an effective means of replacing failed SRUs is incorporated into the system design – which we have rejected for the baseline design to keep it simple.

6.2.3 Fail-stop

This leads to a second question: what strategy should we take in dealing with failures? For the baseline, we adopt the simplest strategy. A failed SRU stops operation and stays in place, not working. More formally, we adopt "fail-stop" as a strategy [SS83]. In this approach, each SRU is either functioning correctly, or it has failed and has stopped executing. This approach simplifies error handling by excluding the wide variety of possible error conditions in which the SRU has suffered from an error condition which allows it to continue functioning in an erroneous state. To minimize time spent in such an erroneous state, the SRU executes frequent self-tests so that an error condition is detected promptly and the fail-stop state is entered quickly following an error.

6.2.4 Errors during mechanosynthesis

One of the major functions of an SRU is mechanosynthesis of blocks. To ensure the SRU complies with the fail-stop criteria, we assume testing of the mechanosynthetic operations at a sufficient frequency to catch errors promptly. At its most extreme, we could test following each mechanosynthetic operation. Detection of an unrecoverable error would cause entry into the fail-stop state. Note that many errors are not unrecoverable. For example, errors that cause failure of the current reaction sequence and disposal of the current block would not, in the sense used here, be "unrecoverable". We might refer to the unrecoverable errors we are talking about here as "hard unrecoverable errors". That is, by definition, detection of a hard unrecoverable error causes the SRU to immediately enter the fail-stop state.

Another way of phrasing this is to say every SRU is either functioning correctly (including correct functioning during recovery from soft failures) or has fail-stopped. We take care in the design of the SRU to reduce the probability of any other possible state to a low enough level that it can be neglected.

6.2.5 Simple failure model

For our simple failure model, if a failure occurs during replication, the SRU fail-stops. Both the SRU and the daughter have failed. This is an all-or-none branching model in which, during each replication cycle, a functioning parent either (1) survives and produces a functioning daughter, contributing two functioning SRUs with probability q, or (2) both parent and daughter fail-stop with probability $1 - q$. Under that model the expected population after g generations is $(2q)^g$.

6.2.6 Letting fail-stopped SRUs accumulate: the baseline

We now have a choice to make: do we simply allow fail-stopped SRUs to accumulate, like so much dead wood, while the remaining SRUs continue functioning despite their presence? Or do we replace fail-stopped SRUs? The first approach is certainly the simplest. We adopt it for the baseline. Replacing failed SRUs is more difficult but has the advantage that it can tolerate a much higher error rate. It remains as a risk mitigation strategy.

If we allow failed SRUs to accumulate, then if we are populating a die starting with a single SRU, we must have an algorithm that allows a parent SRU to pass its daughter to some other parent (some other SRU attached to the surface of the die) that will eventually let it reach a parent that is next to an open slot on the surface of the die to which it can attach the daughter. That is, daughter SRUs must be passed hand-over-hand until they reach an open site on the silicon die to which they can be attached. This means daughter SRUs must be routed around dead parents. This means, at a minimum, that functioning SRUs must be able to detect failed SRUs so that they do not attempt to pass daughter SRUs to dead parents. We further require that this algorithm be very simple, preferably using only a few bits of information, because the SRUs have extremely limited information processing capacity.

6.3 SRU Serial Numbers

In the baseline, each SRU has a location identifier to support routing and coordination on the die. Here, an SRU's serial number identifies its location on the die, rather than its manufacturing history. As there are $805{,}306{,}368 = 2^{14} \times 2^{14} \times 3$ SRUs on a die, we can encode the location using a 14-bit X index and a 14-bit Y index to identify a $600\,\text{nm} \times 600\,\text{nm}$ square, plus two bits to identify one of its three SRU sites, for a total of 30 bits.

Where might an SRU acquire these 30 bits?

The first SRU that we place on the die might be told by a series of instructions to encode its position in its scratchpad memory.

After attachment, a daughter could obtain the coordinates of its adoptive parent or another neighbor and compute its own coordinates from their known relative site positions.

One example where knowing your own serial number would be useful would be coordination of nearby SRUs during manufacture of a larger product. When we divide our SRUs into 4×10 arrays of 40 SRUs to build products larger than an SRU, each SRU has to know its position in the array.

Another example would be diagnostics. Suppose we are concerned that one particular SRU is behaving abnormally, and we want to interrogate it. We broadcast instructions to all the SRUs and either include the serial number of the specific SRU (if we know it) or broadcast a series of test conditions so that the SRU we're interested in can figure out which one it is and then tell us its own serial number. We might then tell the neighboring SRUs to carry out tests on the SRU we're interested in. Of course, this gets into the issue of how to get data back from a single selected SRU to the conventional instruction computer – which is expected to be slow, perhaps painfully so.

6.3.1 Scratchpad memory

We assume the serial number is recorded on a scratchpad memory that might be described as a "binary abacus", where each bead on the abacus can occupy one of two positions, and the number of beads is some reasonable number: perhaps 100. Each bead can be pushed back and forth between its two positions by using the internal assembly arm of the SRU. Which position each bead is in could be tested by using a Test Unit. As we have assumed there are 30 bits in the serial number, and if we assume that it takes 1,000 instructions to interrogate each bit, then it might take about 30,000 instructions for an SRU to interrogate its own serial number while deciding on a course of action. At one instruction per clock cycle and a clock frequency of 100 MHz, this would take approximately 300 μs. As long as the serial number is not accessed too frequently, this works.

In essence, the abacus provides a local scratchpad whose contents can differ between SRUs. This scratchpad could be read using Test Units and written using the internal assembly arm. The instructions to do the reading and writing would come from the broadcast instructions broadcast to all SRUs. These broadcasts would come through the four data lines: S1, S2, D1 and D2. Using these scratchpads would have poor performance. However, because these scratchpads would be accessed infrequently, the overhead for using them in the operations envisioned for the baseline should be acceptable. Scratchpad consultation would be acceptable during initialization, during routing, and before deciding whether to manufacture or assemble a block. Once a block's mechanosynthetic sequence begins, it proceeds without further scratchpad accesses.

6.3.2 Communication between neighboring SRUs

One proposed arrangement would allow an SRU to communicate with its four neighbors, designated N, S, E, and W. Each SRU would have four communication rods, one for each neighboring site. Facing rods would connect to form a bidirectional mechanical link. For example, a southern SRU would extend a rod N, while its northern neighbor would extend a rod S.

To accommodate some lateral misalignment, each rod might end in a flat endplate. Facing endplates could couple through an attractive surface interaction or a mechanical connection analogous to a hook-and-loop fastener. The surface chemistry or mechanical construction remains to be specified. The coupling would need to transmit both pushing and pulling forces without slipping enough to confuse the two bit positions, and remain effective in the surrounding feedstock.

During each transmission phase, one SRU would drive the connected rods while the other allowed its rod to move and read its position. The transmitting SRU would extend or retract the connected rods to encode a 0 or a 1, and the receiving SRU would use a Test Unit to distinguish the two positions. Their roles would reverse in a subsequent phase, allowing communication in both directions over the same link.

Both SRUs would receive the same broadcast instruction stream. During northward transmission,

each SRU would position its north-facing rod according to its outgoing bit, while leaving its south-facing rod free to move. A subsequent broadcast sequence would cause each SRU to read the bit received through its south-facing rod. During southward transmission, the roles of the north-facing and south-facing rods would reverse. Eastward and westward communication would follow the same pattern.

To use baseline actuation and sensing capabilities, the rods would be moved back and forth by the internal assembly arm and read using Test Units. Each rod would pass through a sliding seal in the SRU shell, allowing the internal arm to move the externally connected rod while maintaining the internal vacuum.

A connection and liveness test would also be needed to distinguish a valid bit value from a disconnected link or a failed neighbor. A stationary rod alone would not make this distinction.

6.4 A Local Replication and Routing Algorithm

We consider the expansion of a population of simple replicating units, or SRUs, on a finite two-dimensional square lattice. The lattice used in the simulations in Figures 27–32 is (32×32), although the algorithm is defined for any rectangular grid. Each lattice site has four von Neumann neighbors, corresponding to the north, south, east, and west directions. Diagonal adjacency is not used.

Each site is in one of four states:

- EMPTY
- LIVE
- LIVE-WITH-DAUGHTER
- DEAD

A LIVE site contains a fixed parent SRU connected to the die surface wiring. A LIVE-WITH-DAUGHTER site contains both a fixed parent SRU and a mobile daughter SRU. An EMPTY site contains no SRU. A DEAD site represents a permanently blocked location occupied by a non-functioning parent SRU; no daughter may be placed on or transferred through a dead site.

Each live parent maintains four Boolean routing flags:

$$\text{openN},\ \text{openS},\ \text{openE},\ \text{openW}$$

which represent local information about possible routes toward an open site. These flags do not by themselves authorize movement in the designated directions. Rather, they are local routing signals used to bias daughter motion through the connected network of live parents. Actual movement is permitted only into a valid neighboring site of the appropriate type.

The simulation begins with a single live parent near the center of the grid. All other sites are initially empty. Time is divided into generations, and each generation is divided into discrete time steps.

6.4.1 Generations

Each baseline generation consists of a manufacturing phase followed by a routing phase. During the manufacturing phase, every live parent that does not already hold a daughter manufactures one daughter. Parents that already hold daughters skip this phase. Newly placed SRUs begin manufacturing in the next generation.

The routing phase lasts for a predetermined number of steps specified by the control computer. The generation ends when this scheduled interval expires, regardless of whether all daughters have been placed. No feedback from the SRUs is required to determine the generation boundary. Any unplaced daughters are retained for routing in the next generation. If a parent holding an unplaced daughter dies, both the parent and the daughter are treated as dead.

The original simulator (Section 6.5) uses condition-based generation boundaries: a generation ends when all daughters have been placed or routing stalls. It does not terminate routing after a fixed number of steps. In the simulations performed, routing always continued to completion at the mortality rates investigated, although it could slow down. No daughters carried over between generations. Mortality was applied independently to all live parents between generations, before the next generation's daughters were manufactured and routed. Each live parent becomes permanently dead with probability $P_{\text{mortality}}$, the mortality rate per generation. If all live parents die, the simulation ends in failure. If no empty sites remain, the grid is considered complete.

Cases in which a daughter is completely trapped and continues to pass from parent to parent without reaching an empty site were not simulated. Thus, the simulations did not exercise a stalled generation boundary or the death of a parent holding a carried-over daughter. These simulation conventions and limitations are distinct from the baseline's fixed routing schedule and carryover policy; they do not require the baseline controller to detect completion or failures across all SRUs.

6.4.2 Synchronous Flag Update

At the beginning of each time step, all routing flags are updated simultaneously using information from adjacent sites from the previous time step.

A live parent adjacent to at least one empty site acts as an open destination. Such a parent sets all four routing flags to true:

$$\text{openN} == \text{openS} == \text{openE} == \text{openW} == \text{TRUE}$$

A live boundary parent adjacent to an empty site is also a local source. Its routing flags propagate inward along its row or column under the same update rules, allowing parents in the interior to route daughters toward that source.

Dead sites and grid boundaries block actual movement. A daughter may never move into a dead site or outside the grid.

Routing information propagates through live parents directionally. If a live parent has a live neighbor in direction (D), and the live parent is not itself a local source adjacent to empty space, then its own flag in direction (D) is set equal to the neighbor's flag in the same direction. For example, if site (A) has a live neighbor (B) immediately east of it, then

A.openE $\leftarrow$ B.openE

The opposite flag is not used. Thus A.openE does not copy B.openW. The interpretation is that openD means "a daughter moving in direction (D) is moving along a locally indicated route toward open space."

Previous flag values are retained unless changed by these rules. This permits routing information to propagate through the live-parent network over multiple time steps. If a neighbor dies or the indicated direction has no neighbor at a grid boundary, no flag is copied from that direction. The previous flag value is retained unless the local-source rule sets all four flags to true. Retaining a flag does not permit movement into a dead site or outside the grid.

6.4.3 Parallel Daughter Actions

After the synchronous flag update, daughter motion occurs in three ordered subphases:

1. direct placement onto empty sites;
2. directed transfer along flag-supported directions;
3. fallback diffusion through the live-parent network.

Each subphase is parallel. Eligibility is computed from the lattice state at the beginning of that subphase. Candidate moves are then resolved locally by a proposal-and-acceptance rule, and all accepted moves are committed simultaneously.

A daughter may successfully move at most once during a time step.

6.4.4 Direct Placement

A daughter on a parent adjacent to one or more empty sites may be placed directly onto an empty neighboring site. The daughter then becomes a new fixed parent, and the source parent remains live but no longer has a daughter.

For each daughter with at least one adjacent empty site, a placement proposal is generated. If the daughter has a remembered direction of motion (see Section 6.4.6) and the empty site in that direction is available, the daughter proposes that forward placement. Otherwise, it selects randomly among adjacent empty sites.

If multiple daughters propose placement into the same empty site, the empty site accepts one proposal uniformly at random. All accepted placements occur simultaneously. Rejected daughters remain at their source sites and may participate in later subphases only if doing so does not violate the persistence rules described in Section 6.4.6.

Newly created parents start with all routing flags false. They do not manufacture daughters until the next generation.

6.4.5 Directed Transfer

If a daughter is not placed directly, it may be transferred from its current parent to an adjacent live parent that does not currently have a daughter. A directed transfer from source site S to target site T in direction D is eligible if

$$\text{S.openD} == \text{TRUE}$$

and T is the live parent adjacent to S in direction D, with no daughter currently present.

Each eligible daughter proposes at most one directed transfer. If multiple daughters propose to the same target parent, the target accepts one proposal uniformly at random. All accepted transfers occur simultaneously.

6.4.6 Persistent Daughter Motion

Daughters maintain a one-step memory of their direction of travel. This prevents daughters from executing an unbiased random walk in the interior of the live-parent region, where purely random back-and-forth motion can dominate late in a generation.

Each daughter's remembered motion state consists of two components:

$$d, s$$

where $d \in \{\mathrm{N}, \mathrm{S}, \mathrm{E}, \mathrm{W}\}$ is the remembered direction, and s is a Boolean indicating whether that direction was supported by a true routing flag at the time it was chosen.

There are therefore two persistence modes.

6.4.6.1 Unsupported persistence

A daughter is in unsupported persistence when it is moving in a randomly chosen direction that was not supported by a true routing flag. In this state, the daughter continues in the same direction if the forward target is available. However, if it locally encounters a different available direction whose corresponding routing flag is true, it switches to that flag-supported direction. Its persistence state then becomes supported.

This rule allows daughters that initially wander randomly through an apparently uninformative region to lock onto useful routing information when they encounter it.

6.4.6.2 Supported persistence

A daughter is in supported persistence when its remembered direction was supported by a true routing flag. In this state, the daughter continues in the remembered direction as long as the forward target is available. It does not switch merely because a different flag-supported side direction appears. It changes direction only if its current direction is blocked.

Thus, unsupported persistence is exploratory, while supported persistence is route-following.

6.4.7 Random Diffusion Transfer

If a daughter is not placed and does not perform a directed transfer, it may perform a fallback diffusion transfer. This transfer does not require the relevant routing flag to be true. It only requires an adjacent live parent without a daughter.

The diffusion rule is also persistent. If the daughter has a remembered direction and the forward target is available, it proposes the forward move. If the remembered direction is blocked, the daughter randomly selects among available neighboring live parents without daughters.

If the daughter is in unsupported persistence and observes a different available flag-supported direction, it does not continue random diffusion in its unsupported direction. Instead, it attempts

to switch into the flag-supported direction through the directed-transfer rule. If it loses a proposal conflict for that target, it waits rather than immediately turning sideways in the same time step.

As with other subphases, multiple proposals to the same target are resolved locally by random acceptance.

6.4.8 Local Conflict Resolution

All parallel movement subphases use the same conflict-resolution mechanism. Candidate moves are represented as directed edges

$$(S, T)$$

where S is the source parent containing a daughter and T is the proposed target site.

The resolver enforces two constraints:

1. each source may issue at most one proposal in a subphase;
2. each target may accept at most one incoming proposal.

When a source has multiple candidate targets, one proposal is selected randomly unless persistence uniquely determines the forward move. When a target receives multiple proposals, it accepts one uniformly at random. Accepted moves are then committed simultaneously.

This procedure is deliberately local. It does not compute a global maximum matching or perform global path planning. Consequently, the algorithm remains compatible with SRUs acting on local information.

6.4.9 Time-Step Summary

One time step consists of the following operations:

1. Update all routing flags synchronously.
2. Generate direct-placement proposals.
   (a) Resolve placement conflicts locally.
   (b) Commit all accepted placements simultaneously.
3. Generate flag-supported directed-transfer proposals.
   (a) Resolve transfer conflicts locally.
   (b) Commit all accepted directed transfers simultaneously.
4. Generate fallback diffusion-transfer proposals.
   (a) Resolve diffusion conflicts locally.
   (b) Commit all accepted diffusion transfers simultaneously.

For the baseline, the time steps above are repeated until the predetermined number of time steps has elapsed.

Once the predetermined number of time steps has elapsed, the next generation begins. In the next generation, parents that still hold daughters turn OFF during the manufacturing phase so that they do not manufacture another daughter. They return to ON when the manufacturing phase ends and participate in routing their retained daughters.

Interpretation

The algorithm is intended to model growth by simple local units that have no global map of the occupied region. Routing flags provide only local information, and the update algorithm for the routing flags uses only local information from adjacent SRUs and adjacent sites. Daughters moving from parent to parent looking for an open site use this information when available but can also diffuse in random directions through the live-parent network when no routing signal is present. The two-state persistence rule reduces unproductive backtracking while preserving locality: daughters do not know where open space is globally, but they can continue in a chosen direction and can switch from random exploration to flag-supported routing when a local signal becomes available.

The baseline uses a fixed routing schedule to avoid requiring feedback from the SRUs to determine generation boundaries. For each generation, the control computer specifies a routing interval based on an estimate of the required number of steps, with an added margin. An interval longer than necessary leaves some routing steps unused; an interval shorter than necessary leaves some daughters unplaced. Unplaced daughters carry over to the next generation, and their holding parents skip that generation's manufacturing phase before rejoining routing.

During early development, the presence or absence of a daughter at a site could be determined by examining the die under an optical microscope. This examination could be done either by leaving the die uncovered, or by using a transparent cover. This would allow empirical evaluation of any formula used to determine how many steps are required to find a site for all daughters.

The resulting dynamics combine parallel replication, local routing, stochastic conflict resolution, and mortality. This makes the model suitable for exploring whether a population of locally acting replicators can efficiently spread across a finite substrate despite obstacles, dead sites, and incomplete routing information.

6.5 The Simulator

A simulator for the algorithm described above was written for a 32×32 grid. The initial state followed by snapshots of subsequent states is shown in Figures 27 and 28.

Figure 28 shows an SRU population simulation with a 2 percent failure rate per generation.

This is generation 10 (Figure 28). Had there been no failed SRUs, there would be enough SRUs to fill every site. Because of the failed SRUs (shown as black squares on the grid), there are not enough SRUs to fill the grid. In this simulation, 2% of parents are assumed to fail at the end of each generation. In Figure 28, we see the last two daughters in generation 10 moving towards open sites.

The chosen 2% per generation error rate is much higher than our baseline error rate of 0.1% (0.001) per generation. This algorithm is robust to errors even at a much higher error rate than our baseline error rate.

As a generation takes 10^5 seconds for each parent to manufacture a daughter, and as time steps might take ∼one second, we see that the replication time of an SRU dominates the timestep time

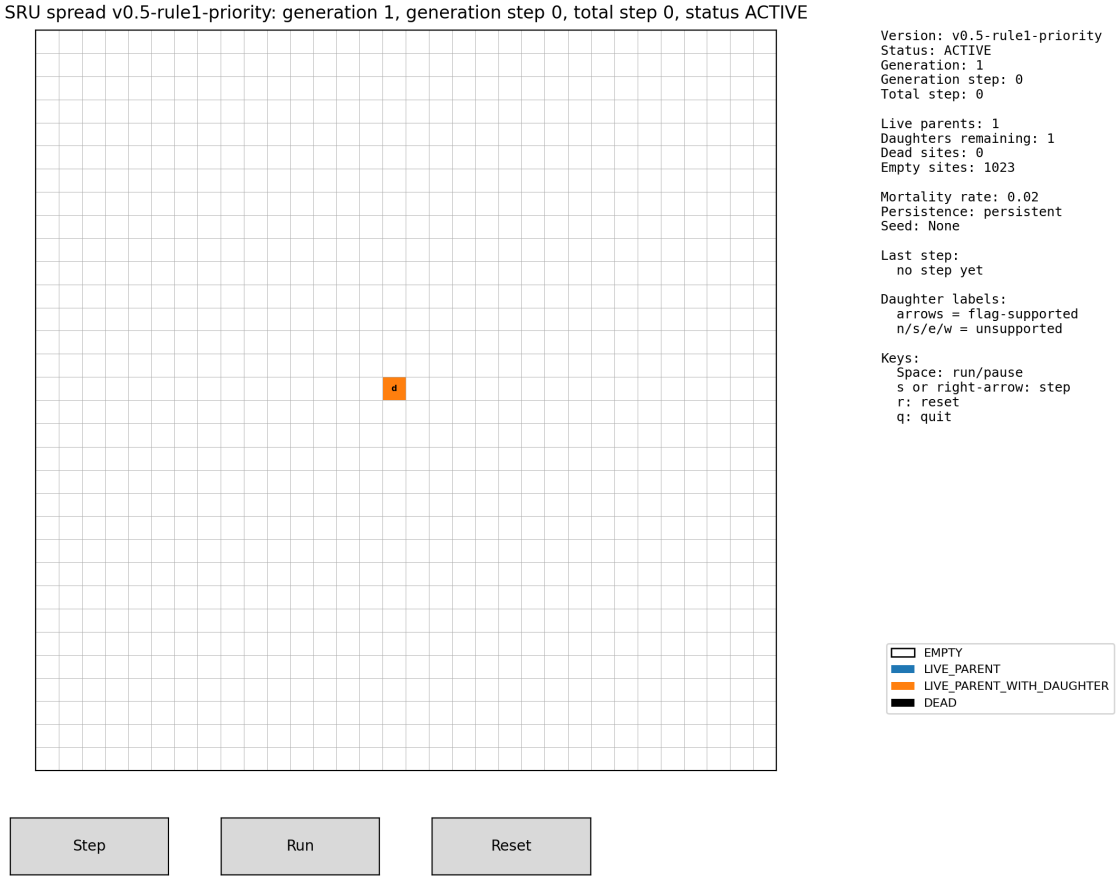


Figure 27: Initial state of the SRU population simulator. White sites are empty; blue sites contain live parents; orange sites contain live parents holding daughters; black sites contain dead SRUs.

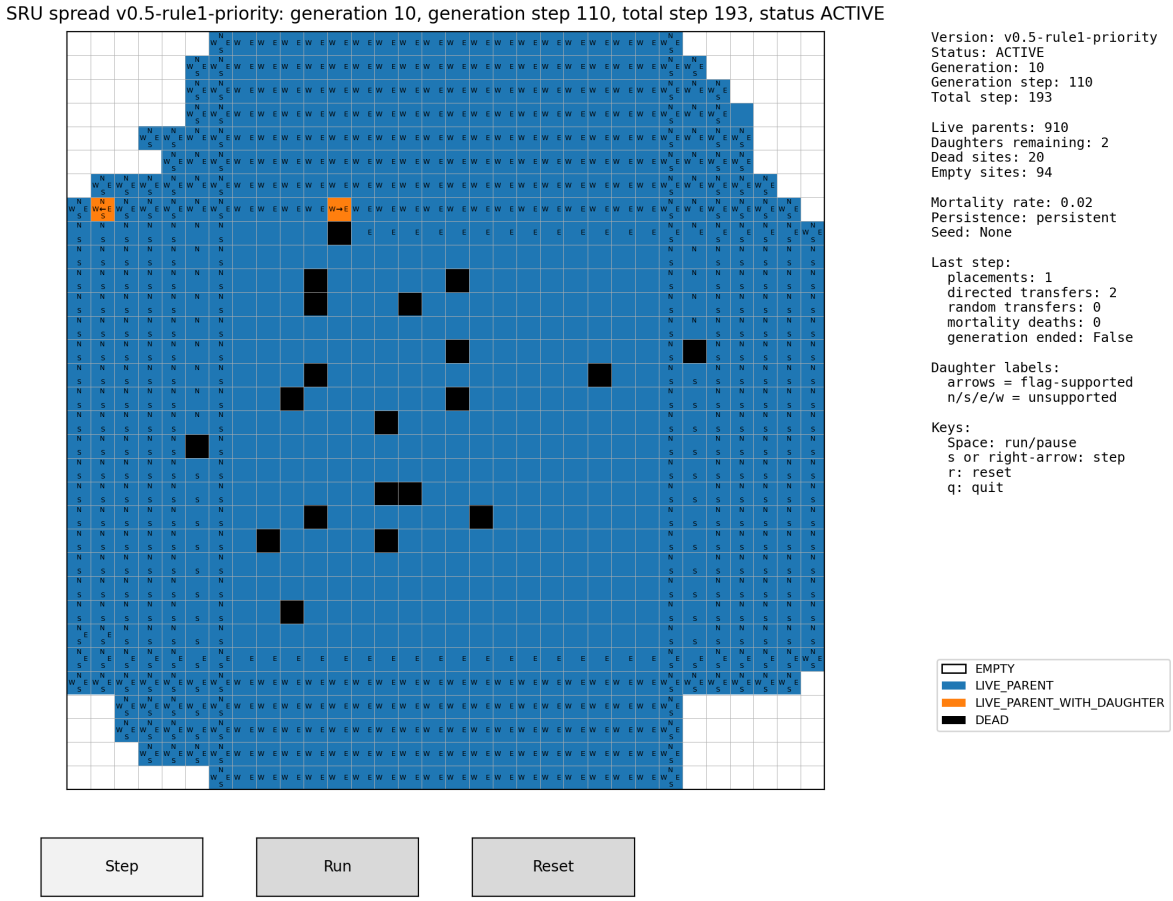


Figure 28: SRU population simulation with a 2 percent failure rate per generation. Site colors have the meanings given in Figure 27.

in this simulation. Even for a full die containing 8.05×10^8 SRUs, replication time should still dominate routing time. For a daughter to migrate from one end of the die to the other should take less than 30,000 steps. At a second per step, this would be about 8 hours. The final generation out of 30 generations to populate a die might take 10^5 seconds to manufacture daughters, and perhaps a day to route almost all of them to open sites. But this is the final generation. Daughters from the preceding generation might migrate half as far, and daughters from the generation preceding that would migrate half as far again. The sum total of all migration time over all ∼30 generations would be small compared with the total manufacturing time, making a high migration time for the last few generations acceptable. Eventually, for a large enough surface, the time required for SRUs to migrate on the surface will become larger than the manufacturing time spent by an SRU to fabricate a daughter SRU, but this should not be an issue for the baseline parameters.

Consider the following idealized case, which assumes that there are no errors in any step and assumes a simple growth pattern in which each new generation is numbered, and assumes each SRU occupies a square SRU site for convenience in illustrating the principles. We start with the small square in the lower left hand corner and assume it is already occupied by a single SRU. Call this generation 0. Generation 0 replicates itself, and the daughter SRU is moved one square to the right. This is generation 1. Generation 0 and 1 then replicate themselves and move their daughters one square up. This is generation 2. Generations 2 and before all replicate themselves and move their daughters two squares to the right. This is generation 3. Generations 3 and before all replicate themselves and move their daughters two squares up. Generations 4 and before all replicate themselves and move their daughters four squares to the right. Generations 5 and before all replicate themselves and move their daughters four squares up. Generations 6 and before all replicate themselves and move their daughters eight squares to the right. At this point, we have filled in the rectangle shown in the image. This pattern can continue until the entire die is filled in (Figure 29).

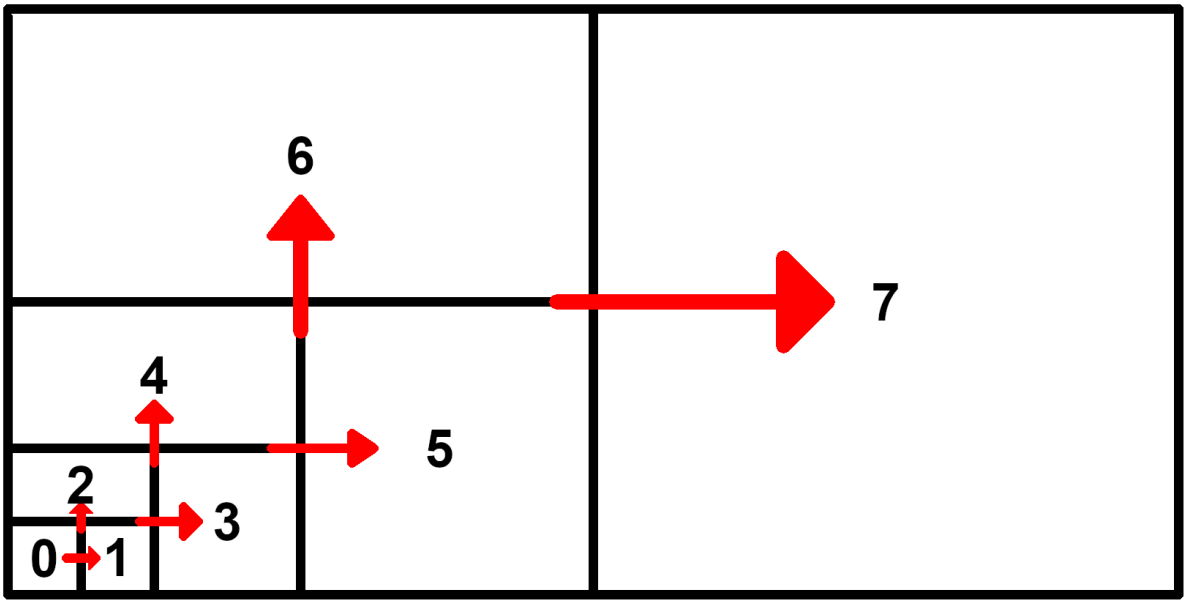


Figure 29: Deterministic SRU growth pattern on a die.

We adopt the convention that a single SRU occupying a single site on the die can simultaneously be a member of many generations. We do this because in most biological examples, parents die and are

gone, so earlier generations are no longer present. In replicating SRUs, earlier generations persist indefinitely and continue to have offspring. This leads to counterintuitive situations. While there seems to be no perfectly intuitive convention for mapping generations onto SRUs, allowing parents, grandparents, great grandparents, and so forth to be members of the current generation because they are actively manufacturing daughters leads to less confusion. This results in a situation where a single SRU can be a member of many different generations at the same time.

Generation 0 is the single parent occupying the first square SRU site in the lower left hand corner of the die. When a new generation of daughter SRUs is manufactured and extruded, both the parents doing the manufacturing and the manufactured daughter SRUs are in the new generation, generation $g + 1$. However, only the parents are in generation g.

Generation g occupies either (1) a square or (2) a rectangle made from two squares fused along one edge. If g is even, it is a square $2^{g/2}$ units on a side; if g is odd, it is a rectangle with sides $2^{(g-1)/2}$ and $2^{(g+1)/2}$.

The total population after replication cycle g, including the parents, is 2^g SRUs. For $g \geq 1$, the number of daughters newly manufactured during that cycle is 2^{g-1}.

The maximum number of steps taken by a daughter in generation g is $2^{\lceil g/2 \rceil} - 1$. Thus, for successive generations, the maximum movement distances are 1, 1, 3, 3, 7, 7, 15, 15, and so forth.

For even N, the cumulative number of parallel routing time steps required to complete generations 1 through N in this idealized no-failure model is S_N. Because daughters move in parallel, each generation contributes its maximum routing distance rather than the sum of all individual daughters' movements:

$$\begin{aligned} S_N = {} & 1 + 1 + 3 + 3 + 7 + 7 + 15 + 15 + 31 + 31 + \cdots \\ & + (2^{N/2} - 1) + (2^{N/2} - 1). \end{aligned} \tag{5}$$

After a little math, we get:

$$\begin{aligned} S_N &= 2^{N/2+2} - N - 4, \\ S_N &\approx 4 \cdot 2^{N/2} \qquad \text{for large even } N. \end{aligned} \tag{6}$$

Here S_N measures cumulative elapsed routing time in units of one routing step, for even N in the no-failure model.

We can double-check by computing the partial sums by hand, which gives the series we want:

2, 8, 22, 52, 114

Equation (6) gives the same numbers for $N = 2$, 4, 6, 8 and 10.

The approximation in Eq. (6) neglects the $N + 4$ correction in the exact no-failure result.

For $N = 30$, $S_{30} = 131{,}038$. If one step takes one second, this idealized model gives ∼36 hours of elapsed routing time on the die, plus the time spent in the 30 replication cycles actually manufacturing daughters. The total time migrating daughters is not unreasonable compared with the ∼35 days required to manufacture the 30 or so generations of daughters required to populate the die.

Alternatives that could reduce die-population time by transferring multiple SRUs onto a destination die are discussed in Faster methods of populating dies (Section 13.5).

This idealized routing-time estimate can inform the baseline's fixed routing schedule. The actual scheduled routing time also includes the margin allocated by the control computer for each generation.

When we introduce fail-stopped SRUs, the no-failure result in Eq. (6) becomes an approximation rather than an exact formula. Despite that, it indicates that the time spent manufacturing daughters dominates the time required to populate a die.

The baseline algorithm is remarkably simple and appears to meet our requirements. Simplicity is important, as the computational capacity of an SRU is extremely limited. Simulation results support the conclusion that this algorithm works well with the baseline error rate of 0.1% per generation. Even with a failure rate of 10% (0.1) per generation, it still populates the die, though with a few trapped SRUs and some additional time for nearly trapped SRUs to find their way out of almost-cut-off regions (see Figures 30 and 31).

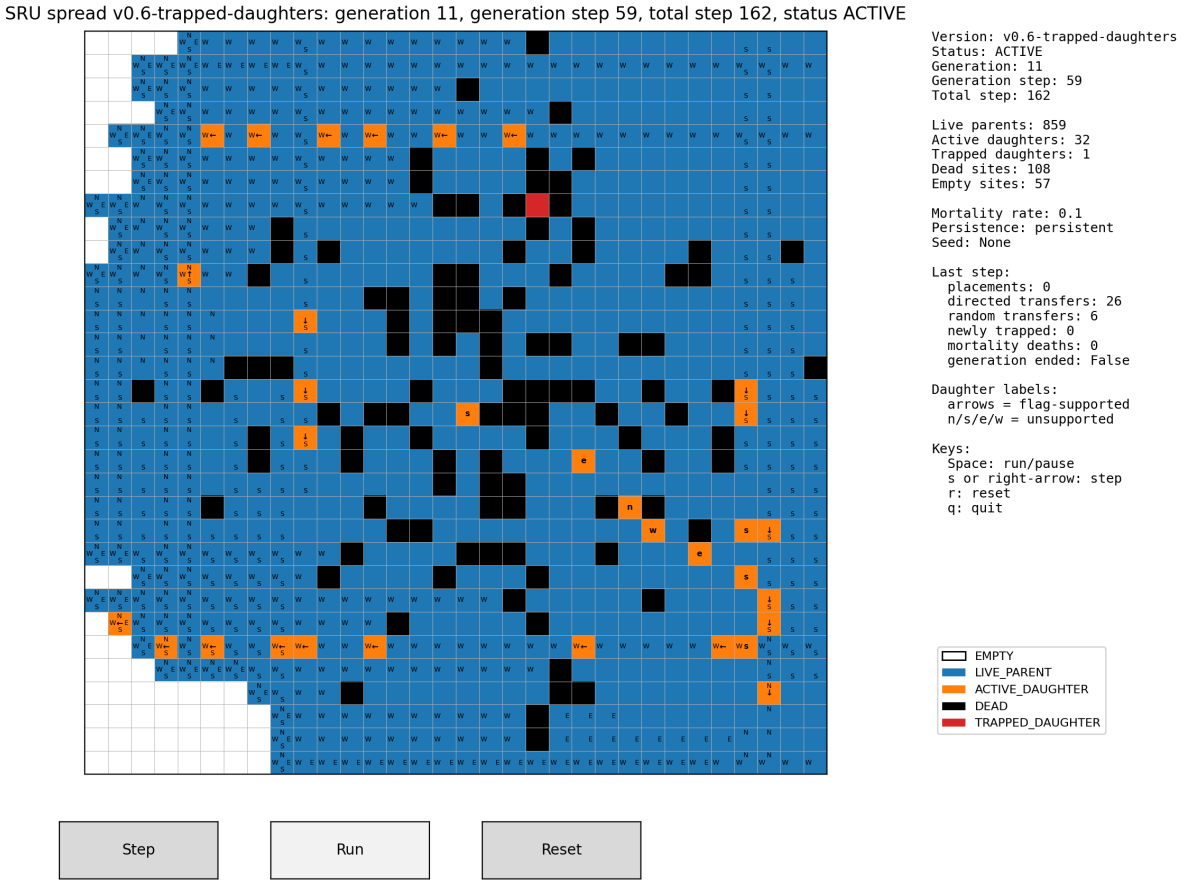


Figure 30: SRU population simulation with a 10 percent failure rate per generation. White sites are empty; blue sites contain live parents; orange marks active daughters; black sites contain dead SRUs; red sites contain trapped daughters.

Figure 31 shows an additional view of the simulation with a 10 percent failure rate per generation.

Even failure rates of 20% (0.2) per generation could be tolerated with small changes to the algorithm to abandon almost-trapped daughter SRUs that would spend a very long time escaping from confinement or, in some cases, never escape. Figure 32 shows a result of running the simulator with a 20% failure rate per generation.

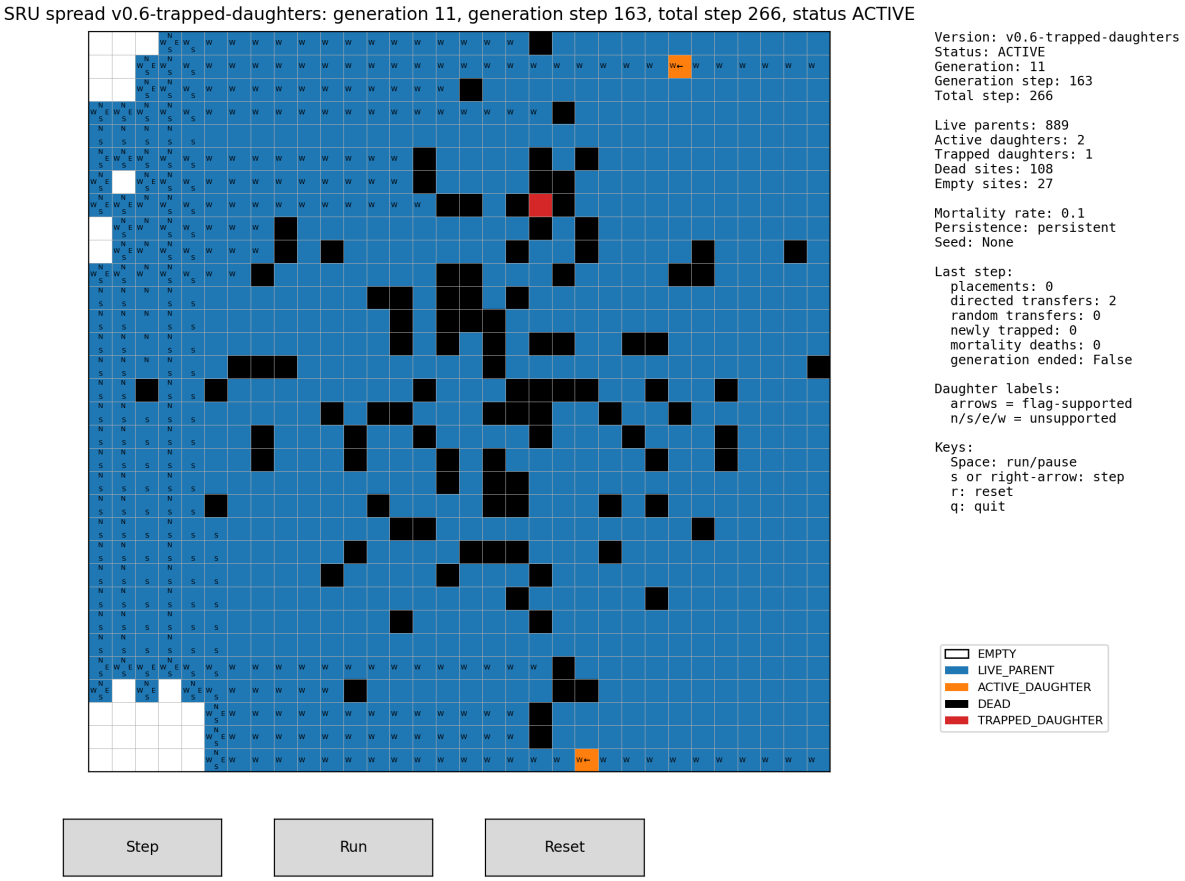


Figure 31: Additional view of the simulation with a 10 percent failure rate per generation. Site colors have the meanings given in Figure 30.

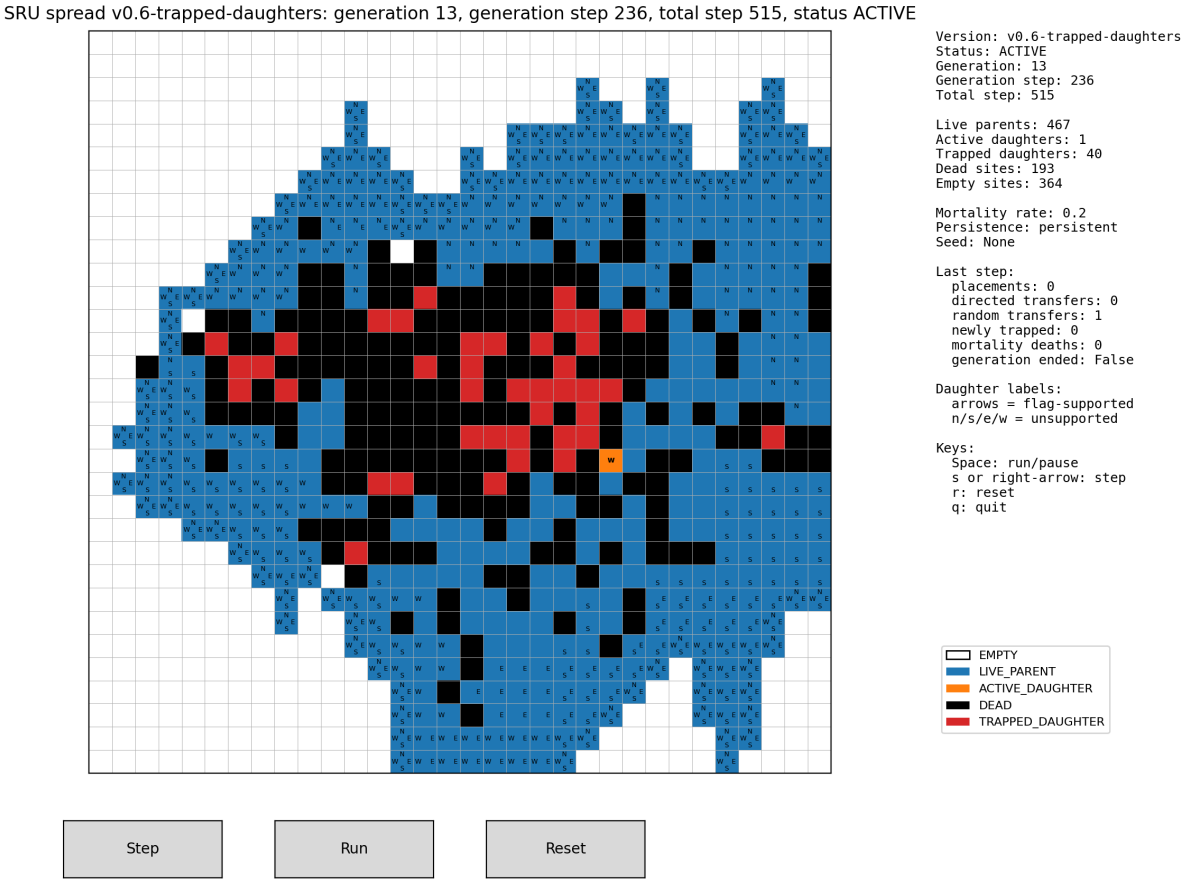


Figure 32: SRU population simulation with a 20 percent failure rate per generation. Site colors have the meanings given in Figure 30.

6.6 10% fail-stopped SRUs on a die should be tolerable

Populating a die starting with a single SRU is a fundamental task in manufacturing the baseline nanofactory. In each generation, the SRUs already on the die (the parents) manufacture new SRUs (the daughters), thereby doubling the population of SRUs. The daughter SRUs must then be attached to the die, thereby becoming, in their turn, parents. To do this, each daughter must be at an open site on the surface of the die and attached to it or transported by a series of parents to such an open site and attached to it. The percentage of fail-stopped SRUs populating the die is an important factor in this analysis.

The percentage of fail-stopped SRUs on a die is different than the failure rate per SRU per replication cycle. A 1% failure rate per SRU per generation would lead, after 20 generations, to a much higher percentage of failed SRUs on the die (Figure 33).

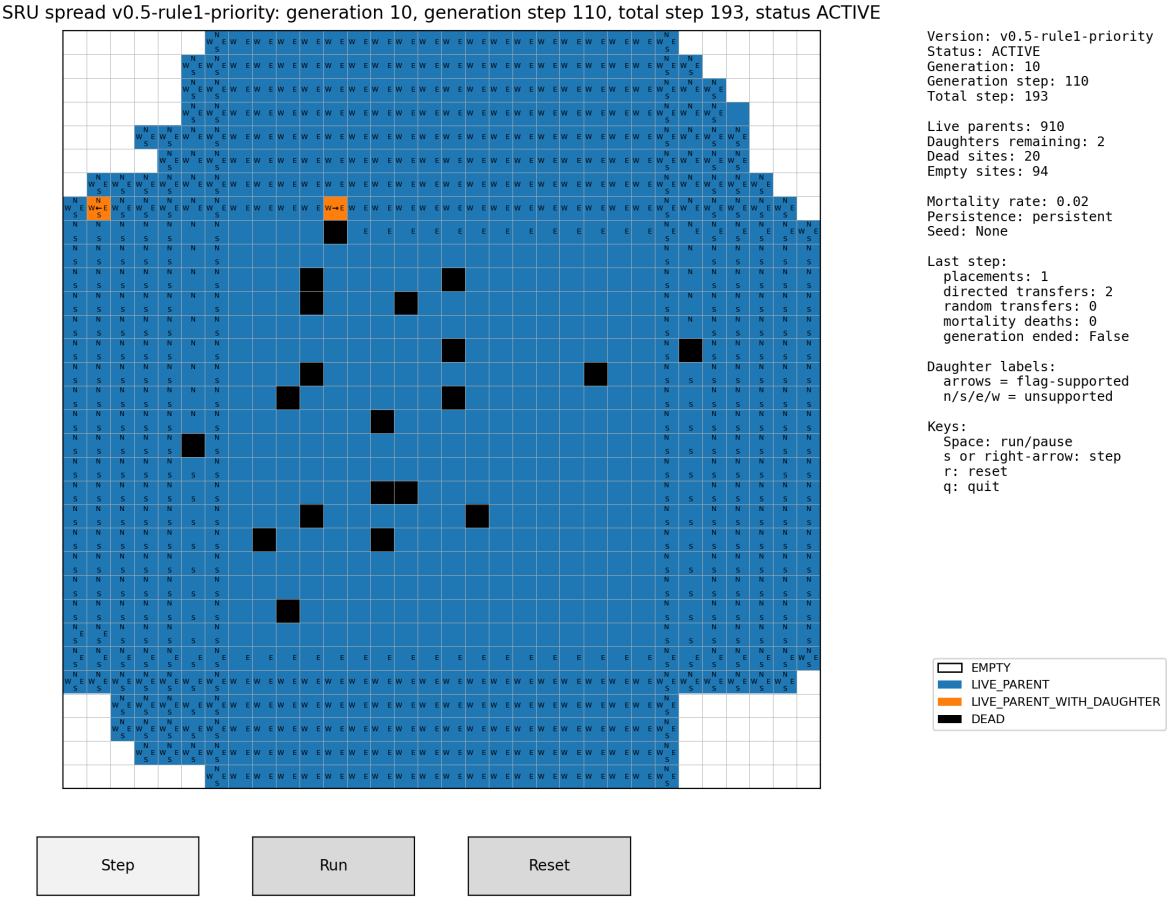


Figure 33: Simulated SRU population on a grid with failed sites. Screenshot repeated from Figure 28.

We start by considering a single parent on an otherwise unpopulated grid of sites and show a plausible map of the available sites to which it might directly attach a daughter (Figure 34).

The dimensions of this grid and the available sites are based on the SRU layout (Section 4.1.1). By increasing the length of the External Assembly Arm (EAA), the number of sites accessible to the parent for placing a daughter can be increased. The number of accessible sites shown is a reasonable minimum for a baseline SRU.

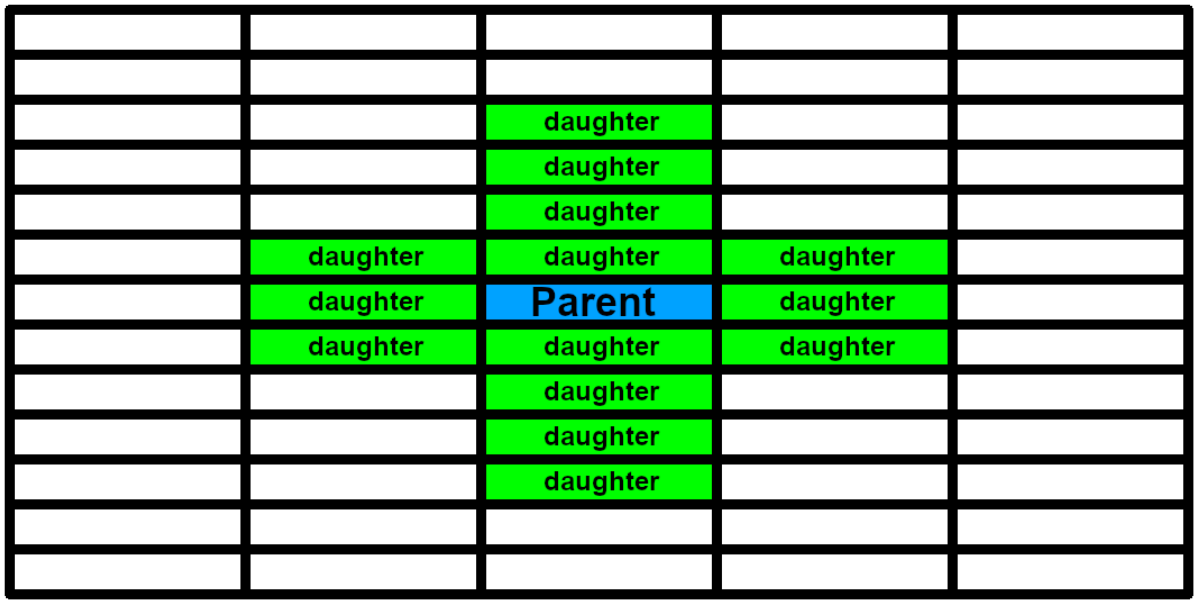


Figure 34: Sites reachable from a central parent SRU. Drawing repeated from Figure 20.

In Section 6.4, we analyzed the growth of SRUs on a die, assuming that each SRU could populate only the four sites adjacent to it in four directions on the surface: north, south, east and west. We concluded that growth was robust even with high error rates per SRU (5% per generation or possibly more).

The baseline design target is for each SRU to place daughter SRUs onto the 14 neighboring sites shown; detailed kinematic validation remains necessary. This substantially increases the ability of SRUs to grow reliably on a surface, even with a substantial percentage of fail-stopped SRUs already occupying that surface. By using a simulated annealing algorithm, in which the probability of selecting a move towards the goal is high but not unity, migrating daughter SRUs can avoid deterministic traps, that is, configurations of fail-stopped SRUs that prevent a deterministic algorithm from making progress because of an unwillingness to move away from the goal for a few steps. Biased random algorithms might be somewhat slower in cases where there are few fail-stopped SRUs, but will be more robust in cases where many fail-stopped SRUs can block progress. With the option of picking any of up to 14 direct next steps, and with a probability of 10% that any particular site is blocked, the probability that all 14 sites are blocked drops to 0.1^{14} = one in 100 trillion. Even allowing for the fact that blockages on successive sites along a path are somewhat correlated, it's clear that the probability of blocking an algorithm that appropriately incorporates some random movement will be very low.

Using an EAA with a longer reach would further improve the probability of success.

Making the useful life of a nanofactory acceptable when failure rates are so high requires that the nanofactory be able to replace fail-stopped SRUs. If fail-stopped SRUs could be replaced, failure rates of 10% per SRU per replication cycle might be tolerable.

While this exercise suggests a remarkably high failure rate could be tolerated, the baseline assumes a failure rate of 0.1% per SRU per generation. It is difficult to reduce the failure rate to below the radiation hit probability of $\sim 3.0 \times 10^{-6}$ per SRU per replication cycle. Appropriately designed SRUs should be able to reduce thermally caused errors to below the radiation rate.

6.7 Convergent assembly and error mitigation

In general, we need to keep errors under control during the manufacturing process. We can do this using the concept of convergent assembly, or hierarchy [Mer97]. Small parts make bigger parts make yet bigger parts. When you assemble a bigger part from smaller parts, test the bigger part thoroughly. If it fails the test, throw it out. This way, the tested parts at every level of the hierarchy can have negligible errors and yields of near perfect parts can be kept high at all levels of the hierarchy.

Consider a single error during, for example, a mechanosynthetic reaction. How might it be dealt with?

1. We can deal with it by an error handling routine at the reaction level and correct the error. The mechanosynthetic reaction is corrected and is completed successfully. The reaction sequence for the block continues.

2. If a mechanosynthetic error is detected but cannot be corrected at the reaction level, recovery proceeds at the block level by aborting synthesis and disposing of the partially fabricated block. This would be a "soft" unrecoverable error, not a "hard" unrecoverable error. Synthesis of the block is aborted. The fabricator goes into idle mode until the broadcast fabrication sequence for that block ends, at which point "block disposal" is executed and the partially fabricated block is disposed of.

3. The error occurs during a mechanosynthetic reaction but is not detected. If the error is detected during a subsequent mechanosynthetic reaction that depends on the correct result of the failed operation, then we have the previous case where synthesis of the block can be aborted. If no subsequent mechanosynthetic step reveals the error (for example, if the error occurred as a final reaction on the surface of a block that would not be used in a subsequent reaction), we will not detect the error during the normal course of continued mechanosynthesis of the block.

4. Assume we don't detect the error, but the block is still damaged. If we are testing after each mechanosynthetic reaction to verify the reaction completed as expected, then somehow an error occurred even though we tested each and every reaction step to verify its correctness. This implies we are now dealing with a low probability event. "Fail-stop" of the entire SRU might be a reasonable strategy. We might, for example, inspect the finished block to make sure it is correct, and, upon detecting an error in the block, place the SRU in a "fail-stop" state, on the grounds that a very low probability error has occurred and the state of the SRU is unknown and attempts to recover might result not in recovery but in a further degraded state. That is, we treat this as a "hard" error.

5. Alternatively, we might inspect the block and if it fails to pass inspection, we could initiate the block disposal routine. However, as there is a good chance we are dealing with a form of damage not caused by a simple mechanosynthetic error, we should now start thinking about inspecting the fabricator itself for errors. Any of the molecular tools the fabricator wields might have been damaged, so as part of the block disposal routine we might also dispose of all molecular tools. We might also carry out a self-test routine to verify the fabricator can position all Stewart platforms correctly. As the simplest thing to do if we detect an error in the fabricator itself is to execute fail-stop, we need to weigh the cost of carrying out a thorough set of tests versus the (hopefully low) probability that an error actually occurred.

That is, even if we might be able to throw out a part that is damaged following successful mechanosynthesis, we might find it simpler to treat this error as a "hard" error rather than a "soft" error. Which choice to make is likely to depend on idiosyncratic features of the error recovery method and its difficulty of implementation.

6. To the extent we can successfully detect hard failures and fail-stop the SRU, our problem then reduces to dealing with fail-stopped SRUs. The hope is that the probability of abnormally operating SRUs can be kept low by having enough tests that force entry into the fail-stopped state. If we wish to replace fail-stopped SRUs, then we need neighboring SRUs to have a method of detecting fail-stopped SRUs, of removing them, and of inserting a replacement SRU into the site vacated by the removed SRU. Replacing fail-stopped SRUs is not part of the baseline.

7. This suggests, as an opportunity, that the SRU manufacture a small and reliable "SRU mounting bracket" which attaches directly to the site, and to which the SRU then attaches. A fail-stopped SRU could then be removed from its mounting bracket by neighboring SRUs, leaving the mounting bracket intact for the replacement SRU.

8. If repeated attempts to replace an SRU at a given site fail, then there needs to be some method of simply marking the site (rather than the SRU at the site) as permanently failed and giving up. The permanently dead site would then just have to be worked around in any algorithm that dealt with SRUs on the surface as a group. For example, the algorithm that passes daughter SRUs to the edges of the block of parent SRUs that are already bonded to the die surface will have to flow daughter SRUs hand-over-hand around "dead" sites on the die surface, until the daughter SRUs reach the edge of the block and can be attached to the die.

The expectation is that in early systems extensive error detection will be used on mechanosynthetic reactions to ensure that any errors are either immediately detected and corrected using relatively simple correction methods or are instead detected and cause the SRU to fail-stop. Anything that might be called sophisticated error recovery is not part of the baseline. If this causes a significant number of SRUs to fail-stop, then we tolerate it. Only in later systems might more sophisticated error recovery be attempted should it prove necessary. As has been discussed, system level function can continue in the face of 10% or even more fail-stopped SRUs on a die.

Nanosystems [Dre92] mentions fail-stop on page 393. Fail-stop is a well-known strategy for high reliability computation [SS83].

6.8 Miscellaneous observations:

An error in a fabricator in one SRU would not normally be able to cause an error in a different SRU because SRUs are isolated from each other by (a) the vacuum tight shell surrounding each SRU and (b) the distance between SRUs. The only time an SRU would normally be able to damage another SRU is during the construction and placement of the daughter.

Recoverable block-abort errors during mechanosynthesis cause disposal of (a) the workpiece as well as disposal of (b) any molecular tools that might conceivably have been damaged. This deals with the risk that such an error might damage a molecular tool, which could then render the fabricator permanently inoperable. The baseline proposal calls for molecular tools to have a "handle" which can be grasped by a "tool holder" (also see the discussion in the Stewart platform (Section 4.4.1)

section). While the highly reactive "tool" end of a molecular tool might suffer damage during an aborted reaction, the probability that the handle, which is chemically inert, might suffer from a hard error should be no greater than the probability that any other inert component of the SRU might suffer a hard error.

The baseline proposal for an SRU does not include conveyer belts. Conveyer belts will likely be faster but are more specialized in their application, require greater design effort, and result in larger system size and greater difficulty in isolating errors. Cascade errors are of more concern as system size increases. Conveyer belts will likely be used in later generations of nanofactories, but in earlier generations generally, and in our baseline in particular, design simplicity and compact size of the replicating component are of greater importance. Stewart platforms or other programmable positional devices are slower but will also be more compact and move design complexity out of hardware and into more easily modified software.

Failure of a wire on the die will cause failure of all SRUs connected to that wire. Failures of wires on dies are well studied. Semiconductor technology is both well-known and mature. We will be using it well within known boundaries. This source of failures can be well controlled.

An SRU has almost no logic in it. Logic operations are extremely infrequent. On those occasions when a few logic operations are needed, we can afford to use reliable, albeit slow, methods. This should result in error rates in logic operations well within our desired operating parameters. The control computer, which is a conventional computer external to the SRUs and the die on which they reside, should have an error rate that is well understood and can be as low as error rates in well-designed high reliability computer systems can reasonably be. Errors during transmission of the ∼100 MHz control signals can likewise be kept to levels that are as low as a conventional mature technology can reasonably achieve.

One of the benefits of the SRU approach is that it prevents cascading errors from propagating beyond an SRU. Once an SRU has completed a replication cycle, the daughter SRU is almost completely insulated from errors in the parent and from other SRUs.

The SRU concept uses broadcast instructions. Therefore, errors in the broadcast can propagate to all SRUs and disrupt system operation. This risk is mitigated by the fact that the broadcast originates from and takes place over existing well-known and well-studied technology, so methods of controlling errors are well studied and mature. While the SRU itself is new technology, the SRU is a eutactic environment and, therefore, it should be possible to model the operations of the SRU both to validate its proper operation and to find and correct many, if not most, errors in its design early in the design process.

The only mechanosynthetic reaction examined in detail in this error-recovery example, the hydrogen abstraction reaction (see Simple Closed Loop Mechanosynthesis (Section 6.9)), exhibits little propensity for hard errors. Recovery from the soft errors to which it does seem prone does not seem to be unduly difficult.

The hydrogen abstraction reaction, which appears to result in a modest raw error rate and recoverable errors, seems more prone to errors (abstraction of the wrong hydrogen) than radical-radical coupling reactions, like that involving $\bullet CH_2$ donation, which involves an existing radical on the workpiece coupling to a $\bullet CH_2$ radical positionally controlled by GeRad. A GeRad-•CH2 approaching a hydrogenated diamond workpiece usually has a low probability of abstracting a hydrogen and a high probability of reacting with a radical site even if its positional accuracy is not perfect. An ethynyl radical will readily abstract the wrong hydrogen if it is not positioned with sufficient accuracy. If we can recover from errors caused by that highly reactive molecular tool, we should be able to deal with errors caused by most other molecular tools.

The fear that SRUs will fail at some very high frequency seems excessive. Biology operates with error rates of 1 in 1,000 to 1 in 10,000. It does this by building a module (a protein) which it then tests and throws out if it is faulty. We can do the same. Our module size (atom count of a block) is dictated by the error rate. Given the error rate, we select a block size that we can manufacture with some high probability of success. Once we know the block size, we design the whole system from blocks of that size. We can decompose large compound diamond parts into such blocks (see Decomposing large structures into small blocks (Section 7.4)).

We then manufacture the blocks, throwing out some small percentage of the blocks when they prove to be faulty. The resulting perfect blocks are then assembled. Assembly operations should be of higher reliability than mechanosynthetic operations, because blocks are bigger than atoms and the alignment requirements for assembly can be designed to be less stringent than the alignment requirements for mechanosynthesis. Remember that the alignment requirements for mechanosynthesis are dictated by nature. The alignment requirements for assembly operations are designed by us, and we can deliberately design assembly operations that favor reliable assembly.

If, for some reason, we must use assembly operations that have low reliability, we can add an additional level in the hierarchy of parts: blocks, second level parts made from (perhaps) a dozen or so blocks, and third level parts made from as many second level parts as we want. Second level parts are then inspected and faulty second level parts are thrown out. This simple hierarchical assembly approach lets us tolerate a relatively high error rate in the assembly of blocks into second level parts. This strategy might be used if we want to fuse blocks into second level parts. The fusion (bonding) of blocks into second level parts might be irreversible and might involve precise alignment requirements. Errors during this process might occur at undesirably high rates, forcing us to inspect and sometimes discard faulty second level parts. See Convergent assembly and error mitigation (Section 6.7).

If a fusion operation is irreversible, then the resulting fused part should be small enough so that it can be discarded. If such errors occur infrequently, a dedicated region (junk heap) can be devoted to holding the discarded parts. If such errors occur more frequently, then the discarded parts could be ejected from the SRU by means of a separate airlock.

We assume in the baseline proposal that assembly operations are either reversible, or that if irreversible assembly operations are used, errors occur at a low enough frequency that a dedicated region in the daughter being manufactured can be used for the disposal of discarded parts. Alternatively, if the error rate is very low indeed, the SRU can fail-stop upon detection of such an error.

It would be possible, though not desirable, to include a separate airlock for the ejection of discarded parts. This would be undesirable because it would imply a high error rate necessitating a need to discard a large volume of waste, because it would increase the design complexity to include the disposal airlock and associated parts handling in the design, and because it would result in a longer replication time. It should, however, allow the system design to tolerate a very high error rate. This high-error-tolerance is not in the baseline but is available for risk mitigation.

6.9 Simple Closed Loop Mechanosynthesis

6.9.1 Abstract

Open loop mechanosynthesis in the context of the broadcast architecture involves “blind” positional control of molecular tools to carry out sequences of mechanosynthetic reactions with a reliability

sufficient to ensure some predetermined probability that the entire sequence will result in the successful fabrication of some desired block. Open loop mechanosynthesis is driven by a fixed sequence of control instructions. Low overhead closed loop mechanosynthesis involves one or more binary (true/false) tests mediated by Test Units (see Test Unit conditional OFF operation (Section 6.1.8)) but is still driven by a fixed sequence of instructions and therefore remains compatible with the broadcast architecture. This is accomplished by effectively disabling the stream of instructions on one branch of the binary test. Even though the fixed instruction sequence continues whether the test is true or false, if the test is false the instruction sequence is effectively disregarded. The goal remains the same: to carry out a sequence of mechanosynthetic reactions with sufficient reliability to achieve some predetermined probability that the overall sequence will be successful.

6.9.2 Hysteresis and backlash

An important point in the design of a low overhead closed loop system is that the positional device (robotic arm) should be designed in such a way that it does not exhibit hysteresis or backlash. That is, when we move to a particular position, the actual position of the device does not depend on the history of operation of the device. As the signals throughout the control chain, up to the final rotations that operate the screw drive, are binary, the issue of backlash is pushed into the backlash in the screw drive mechanism controlling the extension and contraction of the Stewart platform struts. As these mechanisms are atomically precise, it is reasonable to require that when the screw mechanism is nominally held in one position by the binary signal controlling it, either (a) the barriers between any two minima on the potential energy surface of the screw are not significantly bigger than ambient thermal noise (no backlash) or (b) minima separated by significant barriers to rotation result in extension and retraction errors of the strut that are smaller than the 20 pm objective (backlash is small compared to the requirement).

The control chain starts with binary voltages on contact pillars, which then drive binary differential electrostatic comb drives, which then drive a binary demux, which then drives binary Bowden cables, which then drive a binary rotational stepper mechanism, which then drives an analogue screw mechanism. There will be no backlash in the binary chain. If we imagine the rotational stepper mechanism at the very end of this binary control chain stepping a toothed wheel, then a rotational error that is small compared with the angular separation of two adjacent teeth might be possible. Provided this angular error results in a positional error of a size smaller than our desired positional accuracy, then we can still achieve our desired positional accuracy.

The risk of accumulated positional error from hysteresis and backlash could be mitigated by evaluating a specific atomically precise design using molecular mechanics. This statement is a specific instance of the more general claim that the atomically precise design for an SRU could have its risks mitigated by the use of computational chemistry methods.

We illustrate low overhead closed loop broadcast mechanosynthesis by considering a simple hydrogen abstraction reaction, and how to achieve higher reliability even if there is a risk that the HAbst molecular tool abstracts the wrong hydrogen atom from the workpiece, therefore requiring a "patch and correct" sequence to achieve a desired target reliability.

6.9.3 An Example: Abstracting a Hydrogen

We apply the SIMD concept to mechanosynthesis, and specifically to a hydrogen abstraction reaction. See Turn SRU ON (Section 6.1.7) for a summary of the SIMD concept as we apply it to

the SRU. We assume we are applying HAbst to abstract hydrogen H_A, and that there is some risk that we might abstract the wrong hydrogen, H_B. We assume there is a 0.001 chance of abstracting H_B, and a 0.999 chance of abstracting H_A. We want to increase the probability of abstracting H_A and decrease the probability of abstracting H_B. The labels H_A, H_B, and H_C denote individual hydrogen atoms, not hydrogen molecules. We also assume, for purposes of this example, that the probability that one and only one of these two events has occurred is high enough that we can neglect other conceivable error states. That is, we have not abstracted some third hydrogen atom H_C nor have we caused any other erroneous configuration of atoms on the build surface.

The first instruction sequence is to apply HAbst to abstract H_A. With 0.999 probability, we will have abstracted H_A, and with only 0.001 probability we will have abstracted H_B.

Then, we bring GeRad to the carbon atom that carries H_B (or to its radical site if H_B was erroneously abstracted). Either that site is occupied by H_B, in which case the GeRad will not bond to the carbon atom (because it is not a radical, the carbon atom still being bonded to H_B) or H_B was erroneously abstracted and the GeRad bonds to the carbon atom.

We are now able to perform a test by pulling back on the GeRad. Let us assume that 1 nN of force applied to the GeRad will allow it to withdraw from H_B if H_B was not abstracted, and the same 1 nN will not be sufficient to withdraw it from the carbon atom if H_B was abstracted and the GeRad was bonded to the carbon atom. We now have a simple binary test that can be used to determine which of two states we are in, and which of two courses of action we must follow. Either we abstracted H_B (the wrong hydrogen atom) or we did not.

If we did not abstract H_B, then we disable the next steps. There are various ways, which we will consider later, of disabling the fabricator so that it does not carry out these steps. Here, we simply note that if we did not abstract H_B, processing is disabled for the next few steps.

If we DID abstract H_B, then we must remediate the incorrect abstraction. We do this by the following sequence:

We first remove the bonded testing GeRad using a withdrawal force stronger than the 1 nN test force. We then use GeRad-H to donate a hydrogen to the carbon atom, thereby replacing the abstracted H_B. We then repeat the application of HAbst to abstract H_A.

We then re-enable the fabricator regardless of what the test result was so that it can continue processing.

By this sequence of steps, we have enhanced the reliability of the hydrogen abstraction reaction from 0.999 to 0.999999. That is, we have one chance in 1,000 of failing the first time. If we succeed the first time then all is well, and we continue successfully. If we fail on the first attempt we try a second time. If we succeed on the second attempt then all is well, and we continue successfully. Only if we fail on the first attempt and fail on the second attempt do we fail overall, and this occurs only one time in one million.

If we want a higher reliability than one error in a million attempts, we can add an additional test. That is, in the abstract we would want to use a looping construct that continues testing and repairing until we are successful. However, because all fabricators must operate in lockstep when instructions are being broadcast, we have to "unroll" the loop some finite number of times and broadcast the instructions in the loop this fixed and finite number of times. If each attempt has a failure rate of 1 in 1,000, and we want a failure rate of 1 in 1,000,000,000, then we must unroll the loop into three sequential copies and broadcast the instructions for the body of the loop three times, testing after the first two copies of the instructions to see if we have achieved the desired result. Unrolling the loop three times makes execution of this operation three times longer for all

fabricators but makes the result exponentially more reliable. Because test-and-repeat improves the reliability of the operation exponentially, the number of repetitions will be a logarithmic function of the desired reliability. This implies that in a typical case a small integer number of repetitions will be sufficient to achieve the desired reliability.

This method does not guarantee that we will not fail, but it does allow us to decrease the risk of failure and increase the probability of success while remaining within the broadcast paradigm.

We have neglected the risk that the GeRad-H tool will fail, or that the testing process itself will fail, but hopefully these failure modes have a low probability. Donating a hydrogen should not have a high probability of failure, and bonding of GeRad to a carbon radical should also have a low probability of failure. In any event, before confident statements can be made about these probabilities, more detailed analysis will be required.

6.9.4 Disabling the Fabricator

To carry out this method, a one-bit memory enables or disables the demux's D drive response and hence the fabricator's motion. When the memory is OFF, Stepper D does not respond to drive instructions; Stepper S continues to execute selection instructions. The memory is set or cleared using the Test Unit mechanism (Section 6.1.8) to determine whether H_B has been abstracted.

See Broadcast Architecture (Section 6.1) for a discussion of how this might be done.

6.9.5 Summary

The broadcast architecture can simplify the system level design of a nanofactory. The low overhead closed loop method allows us to use the broadcast architecture even when we need to test and remediate some mechanosynthetic reactions to achieve acceptable reliability. Further analysis of specific reaction sequences will be needed to fully understand the range of blocks that can be fabricated using this approach.

6.10 Build rate and replication time for an SRU

The following estimate of total atoms processed per second is adapted from "A Nanofactory Roadmap" [FM08b], section 3.4.1.4, Manipulator Speed and NFW Productivity, page 288. Here NFW means Nanoscale Fabrication Workstation.

We make the following assumptions:

1) Each atom placed requires the sequential application of four tools
2) Each tool moves 15 nm from the tool rack and 15 nm back
3) Each strut extends or contracts in 20 pm steps
4) Each step takes one clock cycle
5) Worst case, all six struts each have to take one step to move the platform 20 pm
6) The clock is 100 MHz

15 nm is 15,000 pm. 15,000 pm / 20 pm is 750 steps per strut. To move the platform 20 pm in the worst case requires that all six struts have to extend or contract by 20 pm. This means we have to multiply our 750-step budget by six to determine the total number of steps (clock cycles) required to move the platform 15 nm. That is $6 \times 750 = 4{,}500$ steps. But we also have to select the struts as well as drive them. This implies using clock cycles to drive S1 and S2, as well as using clock cycles to drive D1 and D2. However, over long distances we can select a single strut and then drive the selected strut many times before selecting the next strut. Therefore, assume we need 10% of the steps used for driving D1 and D2 to drive S1 and S2. This yields 4,500 + 450 = 4,950 steps per 15 nm movement of the platform, with each step taking one clock cycle.

However, we do not take just one 15 nm trip; we have to move repeatedly from the workpiece to the tool rack and back again four times, for a total of eight trips of this length. This takes $8 \times 4{,}950 = 39{,}600$ steps.

The eight-trip budget above is conservative; actual tool-refresh trips may be fewer. As an aside, we note that we only need one tool to move back and forth between the workpiece and the tool rack, as only one tool is modified by a reaction and would need to be refreshed. If you want to abstract a hydrogen, only one tool needs to be refreshed. If you abstract one hydrogen, donate a CH_2, and then donate a hydrogen, that requires the application of three tools, and only three tools need to make the trip from the tool rack to the build site and back.

Additional steps will be used to move the internal assembly arm and the external assembly arm, but there are hundreds of steps used by the SRU fabricator to position a molecular tool for each step used to position a block made of hundreds of atoms. Positioning a molecular tool and positioning an entire block both take a similar number of clock cycles. Block assembly accounts for about 0.1% of the total steps in the conservative positioning estimate below. The total clock cycle budget is little changed by block assembly operations.

If blocks are 1,000 atoms, then there will be 10^5 blocks in a 100M atom SRU.

Even if we position each block with 20 pm accuracy and take 10,000 steps (200 nm) in X, Y and Z, and a like number of steps for roll, pitch, and yaw, and pad the resulting estimate to 10^5 steps, it should take no more than $10^5 \times 10^5 = 10^{10}$ clocks to position all 10^5 blocks. At 100 MHz this is just 100 seconds or 0.1% of a 10^5 second replication cycle.

We will have to metabolize feedstock to recharge our tools. A 15 nm trip to the binding site followed by a 15 nm return trip will be required for each feedstock molecule we metabolize (assuming the binding site is 15 nm from the tool rack), as once the molecule to be metabolized has been fetched it can be placed close to the tool rack. Further movements back and forth between the molecule being metabolized and the tool rack will be short. This means ∼10,000 steps to fetch the feedstock molecule and bring it close to the tool rack. If metabolizing feedstock to recharge a tool requires a similar number of tool applications per atom of feedstock processed, and if the distance between the tool rack and the feedstock molecule is as large as 3 nm, we will need 20% of the steps taken by the mechanosynthetic process, or 39,600/5 = 7,920 steps. This plus our 10,000 steps to fetch the feedstock molecule gives us a total number of steps to process each atom added to the workpiece of 40,000 + 10,000 + 7,920 <60,000 steps.

This metabolizes the feedstock, allows for mechanosynthesis of blocks, and covers the small number of steps needed for block assembly by the internal assembly arm and the even smaller number of steps needed to position daughter SRUs by the external assembly arm.

We assume we need $\sim 10^5$ steps per atom, leaving us with an extra 40,000 steps to cover any remaining miscellaneous tasks the SRU might need to carry out.

The budget needed for error testing and error recovery at the per-reaction level is not clear at the present time. We assume that we will bias SRU mechanosynthetic steps towards those that are reliable and which need little if any error testing. For planning purposes and for purposes of the 2026 baseline we assume that a total of 10^5 steps per atom will be sufficient. See Faster methods of populating dies (Section 13.5) and Error rates in mechanosynthesis (Section 3.7.5).

If it takes 10^5 steps to add an atom to a block, and the clock rate is 100 MHz, then our average SRU fabricator build rate, and hence the average SRU build rate, will be 1,000 atoms per second.

Rather than assuming that all blocks will be completed without error, we assume that each block must pass multiple tests after its completion. Failure of any test, either during or after completion, results in rejection of the failed block.

A test immediately following each mechanosynthetic step, followed by rejecting the block if the test failed, would add little overhead. Picking up the test unit, setting the test unit to TRUE or FALSE based on some condition, and then moving the test unit to the test station would require about the same amount of time as any other tool application. That is, instead of applying four tools we would apply five tools, the fifth tool being the test unit. This would increase the number of clock cycles required by only one quarter. We can, therefore, frequently test for failure conditions that are infrequent, but which would cause block failure if they occurred. If a failure occurred, we would set the "Block Abort" flag and turn the SRU OFF until completion of the current block fabrication cycle. Following the completion of each block we test the "Block Abort" flag. If TRUE, we would execute the block disposal sequence.

We can reduce the number of steps required by dedicating one of the four SRU fabricator Stewart platforms to holding a mobile tool rack. In concept, this is rather like a painter's palette. Painters can hold their palette in one hand while painting with the other hand, allowing them to recharge their brush with fresh paint by moving their brush a short distance from their picture to their palette quickly and easily, rather than painfully moving back and forth between their painting and a palette on a table ten feet away. If one Stewart platform is holding the tool rack, this implies there would only be three Stewart platforms available for mechanosynthesis. However, if any particular reaction required all four Stewart platforms, the one holding the tool rack could put it down and join the mechanosynthetic process.

Holding a tool rack close to the workpiece could substantially reduce the steps spent transporting tools between the rack and the workpiece.

In summary: we reserve 100,000 clock cycles for the placement of each atom. Of those 100,000, 40,000 are reserved for the actual mechanosynthetic steps required to fabricate the block, 20,000 are reserved for metabolism of feedstock molecules, and the final 40,000 include the <1% of the steps needed for block assembly and placement of daughter SRUs. Most of the final 40,000 steps are a reserve that can be deployed for unreliable reactions. The greatest uncertainty in this estimate comes from the number of steps that will actually be needed for checking and error correcting individual mechanosynthetic steps. If individual steps are reliable, the reserve will go mostly unused. If individual steps are frequently unreliable this reserve might be exceeded, increasing the estimated replication time.

A build rate of 1,000 atoms/second per SRU corresponds to a manufacturing rate of 60 × 60 × 1,000 = 3,600,000 atoms per hour, or 100 M atoms in 10^5 seconds.

Our 10^5 second (∼28 hour) replication time seems reasonable for a 100M atom SRU.

Assuming all atoms in the product being built are carbon atoms, and as the mass of a carbon atom is $\sim 2 \times 10^{-26}$ kg, the mass of a 100,000,000 atom SRU will be $\sim 2 \times 10^{-18}$ kg.

Assuming 4 contact pillars to transmit data gives 4×10^8 bits/second $\times$ 10^5 seconds $= 4 \times 10^{13}$ bits $= 5 \times 10^{12}$ bytes $= 5$ TB.

Commercial 16 TB hard drives have specified maximum sustained transfer rates exceeding 2×10^9 bits/s; for example, the Seagate Exos X16 specifies 261 MB/s at the outer diameter [Sea19].

This is feasible even without data compression.

6.11 Power for a daughter

After a parent SRU produces a daughter SRU, at some point the daughter will need to receive power. For the baseline we adopt the following approach.

We assume the parent SRU receives power and control electronically from the die on which it is mounted, and that the daughter SRU will eventually receive power and control electronically from the die to which it will likewise be mounted.

The daughter is manufactured with an attached electrical umbilical (6-conductor cable). The umbilical cable is spoolable and can pay out from the daughter, like the fiber optic tether from a drone. The daughter will be attached to the die by some adjacent already-connected SRU. The adjacent SRU attaches the daughter's umbilical to itself. The cable has sufficient length to allow the adjacent SRU to position the daughter on the die surface, whereupon the (now powered) daughter can deploy its feet to attach to the contact pillars on the die surface (or carry out any other pre-attachment procedure that might be needed).

The astute reader might notice that the adjacent SRU is already holding the daughter with its external assembly arm at the time that it must attach the daughter's umbilical to itself. One method of solving this puzzle is to have the relevant plug affixed to an appropriate surface of the daughter SRU so that the adjacent SRU can position both the daughter SRU and the plug simultaneously. This allows the adjacent SRU to plug in the daughter's cable by appropriately positioning the entire daughter.

6.11.1 Umbilicals

Power for the daughter SRU, if needed prior to its attachment to the contact pillars on the die surface, is provided by a six-conductor electrical umbilical from an adjacent SRU that is already electrically connected to the die. This approach takes advantage of the systems already present in the SRU to utilize electric power and instructions from six electrical signal lines.

While a daughter is receiving power through an adjacent SRU, the adjacent SRU may temporarily be required to supply both its own power and the daughter's power. If both operated simultaneously at the baseline power of 1 nW per SRU, the local load would be approximately 2 nW, or 10.5 nA at 0.19 V. The local contacts and power-distribution branch could be designed to carry twice the nominal SRU current for a short time. Alternatively, the baseline operating sequence might avoid this doubled load by not actuating the supplying SRU and the daughter simultaneously.

The simplest electrical umbilical would be six (4,4) nanotubes, insulated from each other and possibly contained in an insulating sheath around all six. Because it is insulated, such a cable need not rely on the insulating properties of the feedstock solution. It carries electrical signals from a previously attached SRU that is already connected to the die and adjacent to the daughter SRU being attached.

Such a six-conductor cable would carry Ground, V+, S1, S2, D1, and D2—the minimum electrical interface assumed by the baseline design.

While nanotube plugs exposed to the surrounding feedstock seem likely to work, we could use plugs designed to keep the nanotubes in vacuum during their entire cycle of operation by using appropriate sliding seals, as depicted in Figures 35 and 36.

Initially, the left and right connectors are separated (Figure 35, left panel).

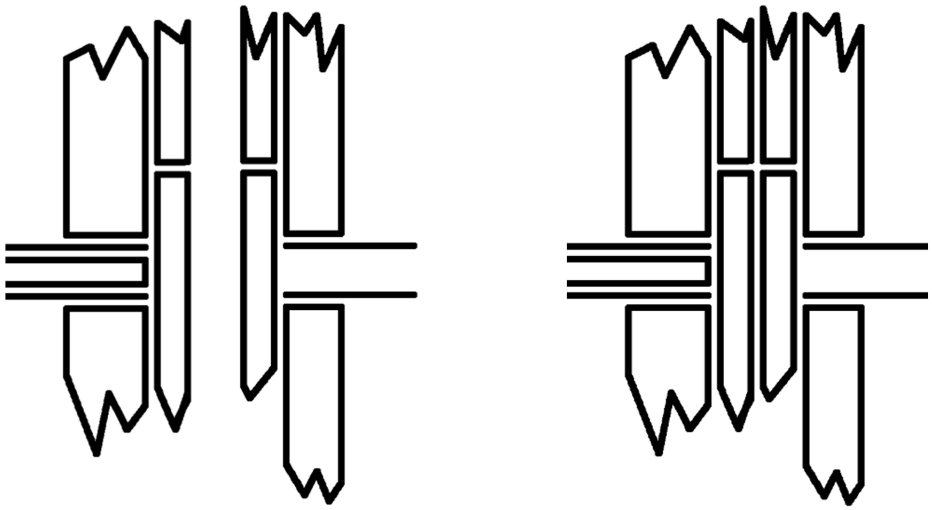


Figure 35: Umbilical connector approach and mating.

Then, the left and right connector faces are brought together, and the two atomically precise surfaces are sealed together (Figure 35, right panel).

We open the sliding seals that separate the internal (vacuum sealed) nanotubes (Figure 36, left panel).

We then slide one nanotube into the other inside the vacuum sealed space (Figure 36, right panel).

This approach can also be used to connect mechanical umbilicals if we wish to avoid contact with the feedstock and whatever impurities it might contain. This assumes that a vacuum seal is required and that sliding seals can be vacuum tight. The baseline assumption that sliding seals can be vacuum tight seems reasonable but carries some risk. De-risking by various theoretical methods would be helpful. Mitigation strategies, should this risk be realized, are reviewed in Sliding Seal Extrusion Risk Mitigation (Section 12.12). A risk mitigation strategy might be to use mechanical power transfer through wobble sticks. This would eliminate the need for sliding seals in this case.

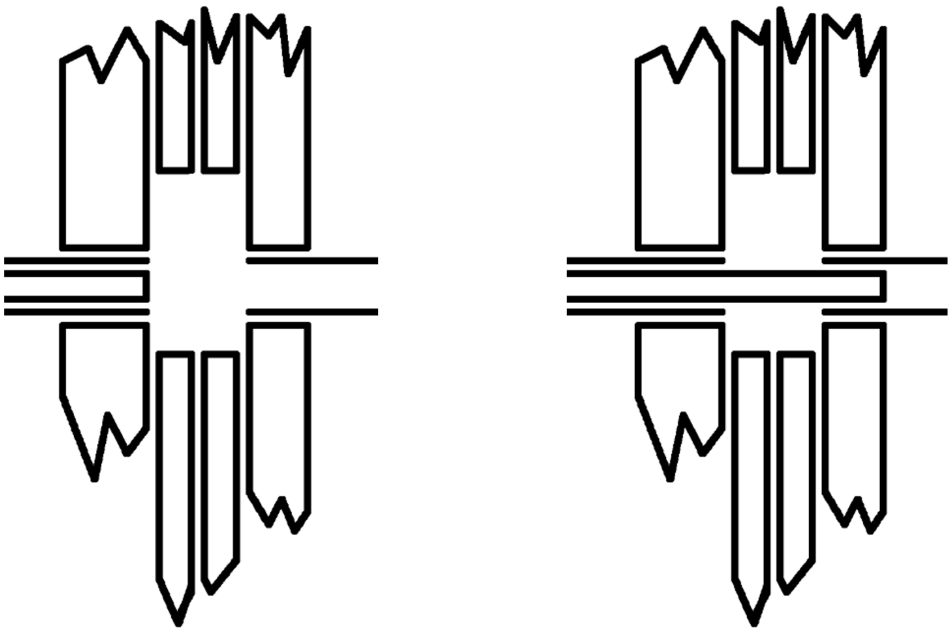


Figure 36: Opening the sliding seals and joining the umbilical nanotubes.

7 Fabricator issues

This section should be read after the Fabricator (Section 4.4) discussion under The Simple Replicating Unit or SRU (Section 4). This section covers specific issues surrounding the design and use of the fabricator in greater detail.

7.1 Molecular composition of the feedstock

If we are to fabricate products with H, B, C, N, O, Si, Ge, and Pt, we need feedstock molecules that include all eight elements. The mixture should be liquid at 210 K. We add the requirement that each pure feedstock component be liquid at 210 K, as this is easier to validate. All feedstock molecules should have an acceptable shelf life when mixed. The feedstock molecules must be mutually soluble. In addition, verification that the feedstock molecules and the waste molecules are all mutually soluble and stable is needed.

7.1.1 Elemental and functional design

- **C, H** → pentanenitrile, TEOS
- **Ge** → tetraethylgermane
- **N** → pentanenitrile
- **Si, O** → tetraethyl orthosilicate (TEOS)
- **B, O** → triethyl borate
- **Pt** → (ethylcyclopentadienyl)trimethylplatinum(IV) ($EtCpPtMe_3$)

From a chemistry standpoint, these candidate feedstock compounds (Figure 37) form a heterogeneous set:

- nonpolar vs. polar aprotic
- neutral organics vs. alkoxides
- light vs. heavier main-group elements

The expectation is that they coexist as largely independent species in a shared liquid environment.

7.1.2 Challenges *in* stabilizing the system

Stability here has three distinct dimensions that must all be satisfied simultaneously.

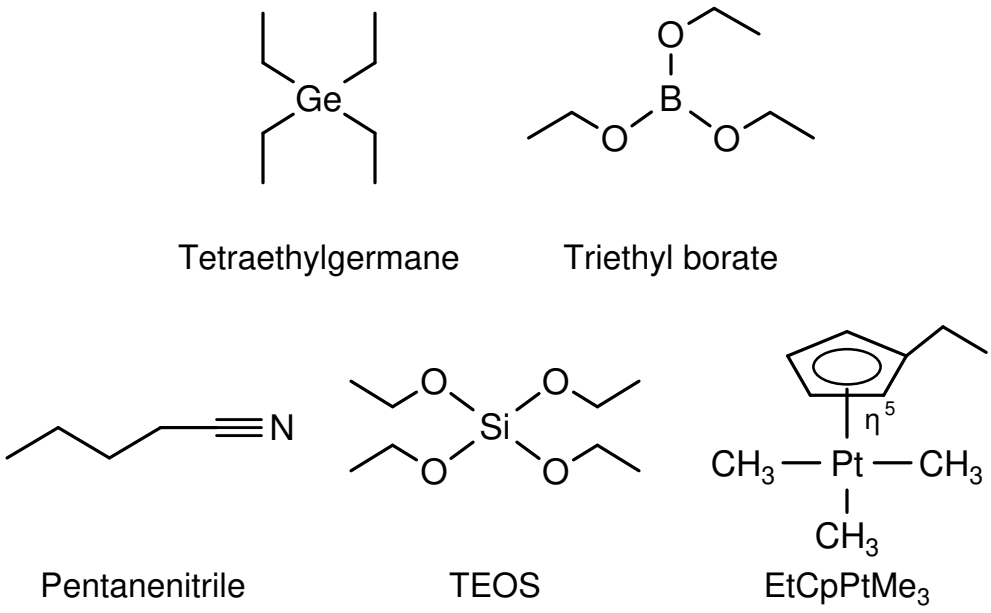


Figure 37: Molecular structures of candidate feedstock compounds.

7.1.2.1 Chemical stability

The most important reactivity channel in this system is hydrolysis:

- TEOS reacts with water to form silica (SiO_2-like networks) and ethanol
- Triethyl borate hydrolyzes to boric acid and ethanol

These reactions are:

- thermodynamically favorable
- triggered by trace water, not bulk conditions

The dominant requirement for chemical stability is therefore rigorous exclusion of water (and other protic contaminants).

Temperature then modulates the rate:

- lower temperature → slower kinetics
- higher temperature → faster kinetics

7.1.2.2 Phase stability

The mixture spans a wide polarity range:

- Pentanenitrile → polar aprotic, non-basic

- TEOS / triethyl borate → polar, reactive alkoxides
- Tetraethylgermane → weakly polar organometallic

This creates a classical risk: mutual solubility is not guaranteed across all conditions.
At low temperature this becomes more pronounced because:

- the stabilizing entropic contribution $T\Delta S_{\text{mix}}$ decreases with temperature (for ideal mixing)
- phase separation becomes more favorable
- crystallization thresholds are approached

Each component also has its own melting point; Table 9 summarizes melting and boiling points of candidate feedstock compounds.

Table 9: Melting and boiling points of candidate feedstock compounds.

Compound	Melting point	Boiling Point
TEOS / tetraethyl orthosilicate	**191.0 K** (NIST WebBook)	**440 K**
Triethyl borate	**188.4 K** (NIST WebBook)	**391 K**
Tetraethylgermane	**180.5 K** (NIST WebBook)	**435 K**
Pentanenitrile	**176.9 K** (NIST WebBook)	**414 K**
$EtCpPtMe_3$	$\leq$**195 K** (US5929267)	$\sim$**443 K**

Near their melting points, mixtures tend to show:

- partial crystallization
- slush formation
- local phase separation

7.1.2.3 Physical (transport) stability

At low temperatures:

- viscosity increases
- diffusion slows

- equilibration times lengthen

Consequences include:

- persistence of local composition gradients
- slower mixing
- increased sensitivity to small perturbations

7.1.3 Behavior of this mixture at 210 K

At 210 K, phase and transport behavior are more favorable than at the lower temperatures discussed above, although lower temperatures can improve chemical (kinetic) stability.

At 210 K, the temperature is at least 15 K above the pure-component melting point of $EtCpPtMe_3$, reported as $\leq$ 195 K; all other components have lower melting points.

7.1.4 Overall stability assessment

The pure-component melting points are encouraging: each component should be liquid at 210 K when considered individually. They do not, however, establish the phase behavior, mutual solubility, viscosity, diffusion coefficients, or long-term chemical stability of the five-component mixture. The mixture may exhibit phase separation, eutectic behavior, compound formation, or crystallization not predictable from the pure-component melting points alone. We therefore treat operation as a homogeneous liquid at 210 K as a baseline assumption requiring experimental validation.

7.1.4.1 Summary

The available data make 210 K a plausible starting temperature, but measurements of the actual mixture are required before its chemical, phase, and transport stability can be characterized with confidence.

7.2 Binding site leakage

Binding sites for feedstock molecules might allow passage of either feedstock contaminants or vented waste gases into the interior of the SRU. If the rate of such leakage is low enough, the resulting SRU failure rate will remain below our target of 0.1% per SRU per replication cycle. Above some threshold, it will exceed our target failure rate. How can leakage be kept below this threshold?

The most basic strategy is to adopt a staged cascade. The following figure, from [Dre92], illustrates the concept (Figure 38).

We could implement the system more or less as shown, with separate chambers holding liquid feedstock with varying concentrations of impurities and the desired feedstock. Alternatively, we might actually use only one feedstock molecule per "chamber". If a single stage allows the passage of at most one feedstock molecule, we can then interrogate the state of that (possibly absent) single feedstock molecule by the use of Test Units. Should the putative feedstock molecule fail the test, it could either (1) be returned to the previous stage or (2) be vented from the SRU. These two options are quite different.

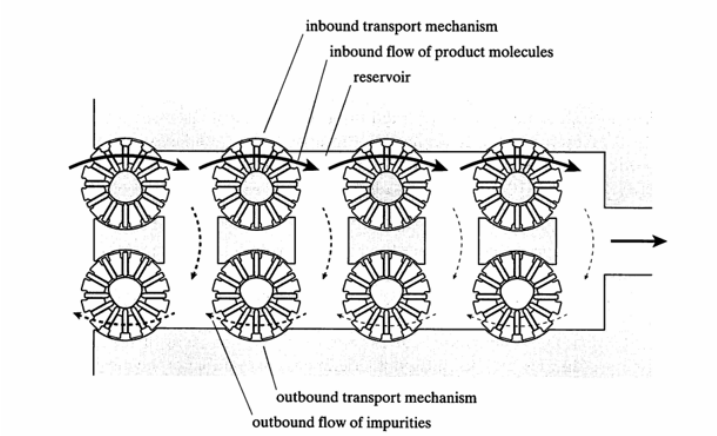


Figure 13.4. Schematic diagram of a staged-cascade process based on sorting rotors like those of Figure 13.1. (Note the inversion of the cams in the lower rotors.)

Figure 38: Multistage molecular sorting cascade, from [Dre92].

In the first case, illustrated in Figure 38, sorting of product towards the output of the filtration process and sorting of impurities backwards to eventually be ejected by the first stage back into the feedstock both take place at the same time. This is the more general case and, provided that at least some selective filtration takes place at each stage, would, in the idealized case, eventually produce any desired purity of output.

Each stage moving to the right would have a higher concentration of the desired product molecule and a lower concentration of impurities. The total volume being filtered forwards needs to match the total volume being filtered backwards, and the net filtration moving to the right, taking into account both filters between two adjacent chambers, needs to at least somewhat improve the concentration of the desired molecule and somewhat reduce the concentration of impurities. Conceptually, a separate such filtration process would be used for each type of feedstock molecule.

The simplest implementation of such a system would be a single stage moving only forwards. That is, one binding site would bind to the desired feedstock molecule directly from the feedstock solution and move it into the SRU. Such an implementation would require that the single stage achieve a level of selectivity sufficient to support further operations at a reliability that would meet the 0.1% failure rate per SRU per replication cycle.

A second simplification would be to have a series of forward-only stages, with rejected outputs from a stage being vented from the SRU. If each stage has an independent false-acceptance probability p, then n stages reduce the false-acceptance probability to p^n. If N contaminant molecules are presented during one replication cycle, the corresponding failure probability is bounded by approximately Np^n. The number and types of tests must therefore be chosen so that Np^n is below the target failure probability of 10^{-3} per SRU per replication cycle. Tests should examine different molecular properties to reduce common-mode failures. The probability of false rejection of the desired feedstock would have to be low enough that a sufficient quantity of the desired feedstock emerged from the final stage.

A sequence of such tests should be able to reliably pass only the desired feedstock molecules. One such test would be to clamp the putative feedstock molecule between two half-binding sites (half-sites). Should the two half-sites close completely, that would constitute passing one test. Should they fail to close completely, either because the wrong feedstock molecule was present or because

the desired feedstock molecule was present but there was some additional (undesired) molecule present, then the test would fail.

A second test would be to measure the force required to open the two half-sites after they had closed. If the opening force was incorrect, again, the test would fail.

A single test might be sufficient to achieve the desired reliability. If not, additional tests would be added until the desired reliability was achieved.

Finally, it would be possible to use the full system. This would use a sequence of chambers connected sequentially from left to right, with each pair of adjacent chambers having both a method of moving desired feedstock from the left chamber to the right chamber, and a method of moving undesired impurities from the right chamber to the left chamber. In such a system (1) higher concentrations of the desired feedstock and lower concentrations of impurities would be to the right, and (2) higher concentrations of impurities and lower concentrations of the desired feedstock would be to the left.

The actual method implemented would be the simplest that achieved the target failure rate per SRU per replication cycle.

The general method lets us reject waste gases and impurities in the feedstock with high reliability by increasing the number of stages of filtration that each input is required to go through.

7.3 AP Blocks and Fabricators

We make the following assumptions about AP Blocks.

1. **The SRU is assembled entirely from AP Blocks.**
2. **AP Blocks are atomically precise.**
3. **AP Blocks have a fixed maximum size.**
4. **AP Blocks are made by Fabricators.**
5. **Fabricators are assembled entirely from AP Blocks.**
6. **An SRU Fabricator is a subsystem within the SRU.**
7. **A workstation fabricator is a subsystem within a workstation.**

We want to manufacture atomically precise products; therefore, we are engaged in Atomically Precise Manufacturing, or APM. Further, we assume we are manufacturing everything from Atomically Precise Blocks, or AP Blocks. Think of LEGO bricks, only very small and very precise. When the context is clear, we might refer to AP Blocks simply as "blocks".

AP Blocks are ***assembled***. That is, we refer to the mechanical act of positioning two or more AP Blocks so that they remain connected (fused, bonded, connected, interlocked, or otherwise joined) as *assembling* the AP Blocks. How AP Blocks are assembled depends on the AP Blocks.

The workstation fabricator is assembled entirely from AP Blocks. The SRU Fabricator, as a subsystem of the SRU, is also assembled entirely from AP Blocks. The AP Blocks from which the workstation fabricator is assembled are initially manufactured by the workstation. The workstation fabricator can make the AP Blocks used to assemble further workstation fabricators and also the AP Blocks used to assemble the first SRU.

As it seems likely that driving down the mechanosynthetic error rate in early systems will both take time and be difficult, we should assume as high a mechanosynthetic error rate as can be tolerated. We adopt a simple model, assuming each atom placed has a fixed probability of being placed in error. With such a model, if we wish to build atomically precise blocks with 100 atoms with an error rate of 0.001 errors per atom placed, we will have a probability of success per block of 90%. However, the probability of success for a 1,000 atom block would be 37%. This would require rejecting almost two AP Blocks for every AP Block that was accepted. This might be tolerable during the manufacture of the first SRU but might be unacceptable in a production system. That is, the workstation fabricator might be able to reject two out of three blocks manufactured because it only has to manufacture enough blocks for the first SRU, but such a high error rate seems unacceptable for the SRU fabricator.

While the workstation fabricator can dispose of waste blocks in the workstation, which effectively has unlimited volume, the SRU fabricator would ideally dispose of a small number of waste blocks inside the daughter SRU being manufactured (or within the SRU-sized product being manufactured), which has limited volume available for disposal of waste. The baseline assumes the SRU fabricator has an error rate low enough that the limited waste disposal space would be sufficient.

A mitigation strategy for dealing with a higher error rate would be to dispose of a higher volume of rejected blocks. The SRU design could be augmented by the addition of an airlock through which rejected blocks could be ejected. This would allow an effectively unlimited number of blocks to be rejected. While a high error rate would lengthen the SRU replication time, it should not cause outright failure except in extreme cases.

We assume that AP Blocks are manufactured by one of the following:

(i) The workstation itself;

(ii) Molecular machinery within the workstation called the workstation fabricator; or

(iii) Molecular machinery within the SRU called the SRU Fabricator.

We assume a maximum size of 1,000 atoms for a block (see AP Blocks (Section 3.4.3)). A plausible target error rate for the workstation fabricator might be ∼0.001 errors per atom placed during the manufacture of the first SRU. Target error rates for the SRU Fabricator should be lower.

While the baseline assumes a maximum of 1,000 atoms per block, there is an opportunity to design a nanofactory from AP Blocks that are smaller in size: a few hundred atoms or even less. For the present theoretical analysis, the exact maximum size of AP Blocks is less critical than the fact that there is a maximum size, and that a complete atomically precise design of an SRU based on AP Blocks would necessarily provide us with a list of the blocks needed in its manufacture, and hence the maximum block size needed to assemble an SRU from that design.

We note here that the workstation fabricator is likely to be made from fewer AP Blocks, and smaller ones, than the SRU fabricator. The workstation fabricator will operate more slowly and with a higher error rate than the SRU Fabricator. There are reasons for this. The first workstation fabricator must be made from AP Blocks "hand crafted" by the workstation. The fewer and smaller they are, the better. The workstation fabricator can literally take months to fabricate the AP Blocks needed for the first SRU, while the SRU Fabricator needs to manufacture all the AP Blocks for another SRU (including the SRU Fabricator) within one SRU replication time of 10^5 seconds. Both the manufacturing constraints and the performance requirements are very different,

so the designs for the two fabricators are likely to differ. There are good reasons to design two different fabricators for the two different tasks.

To reiterate: the workstation makes the AP Blocks needed to assemble the first workstation fabricator. Designing the workstation fabricator from the fewest possible AP Blocks that are themselves possible for the workstation to make directly in a reasonable time is important. The workstation fabricator will presumably be able to make AP Blocks faster than the workstation, making it possible to manufacture all the AP Blocks needed for the first SRU within a matter of months.

The SRU Fabricator will be better yet, able to make all the AP Blocks for an SRU within the 10^5 second (∼28 hour) replication cycle of the SRU.

The baseline assumes the design of the first SRU that we build (the first generation) is the same as the design of subsequent generations of SRUs. We make this assumption because (a) it's easier to design a single SRU and (b) AP Blocks are largely the same, whether made by the workstation fabricator or the SRU Fabricator.

This type of logic does not apply to the workstation fabricator. The workstation fabricator will be different than the SRU Fabricator. The SRU Fabricator will be a significant redesign and improvement over the workstation fabricator. We expect that the workstation will have a harder time making the AP Blocks for the first workstation fabricator, and that we will diligently try to design the workstation fabricator from the simplest and fewest AP Blocks. As the workstation fabricator can take many months to make all the AP Blocks for the SRU, it can be ∼100 times slower than the SRU Fabricator. This should help simplify the design of the workstation fabricator.

The lower error rate of the SRU Fabricator creates an opportunity to redesign the SRU using blocks that have more than 1,000 atoms. The baseline does not do this because redesigning the SRU looks like significant additional design work that is not strictly needed to make the first SRU operational. Even if the mechanosynthetic error rate of an SRU fabricator is lower than the error rate of a workstation fabricator, the initial SRU design would still work, even if it did not fully benefit from the lower error rate. Therefore, we design the SRU from AP Blocks that are of a size that can be manufactured by the workstation fabricator.

We have to design an SRU, and that first design must necessarily be composed of blocks that can be made by the workstation fabricator. Those same blocks can be made by the SRU fabricator, so a second SRU design is not required.

For the experimentalists reading this discussion, the lesson is simple. The first experimental target for the workstation is building and assembling enough AP Blocks to make the workstation fabricator.

This deceptively simple statement assumes that the design of a workstation fabricator is a fixed target, when in fact there are many possible designs for workstation fabricators. The selection of a particular design for implementation involves choosing a design whose implementation is feasible, and which can also make a sufficiently useful range of AP Blocks to enable making the AP Blocks needed for the SRU (as well as the workstation fabricator itself).

Experimentalists want AP Blocks that at least have a plausible pathway for their experimental development. We must be able to make them in the workstation without the assistance of any fabricator. Among other things, this means the AP Blocks should have as few atoms and as few different elements as possible. AP Blocks made from only the elements H, C, and Ge might conceivably be enough to make a working workstation fabricator. It's hard to estimate the minimum possible size of block that would actually be sufficient.

Risk reduction for the assumption that the workstation fabricator could be made from smaller AP

Blocks should be possible by completing an atomically precise design and theoretical analysis (including relevant computational modeling) of a workstation fabricator. This analysis and modeling should be simplified by the fact that the design will be broken into AP Blocks, as well as by the fact that the design of the workstation fabricator should be deliberately structured so that it can be analyzed theoretically and modeled computationally.

The workstation fabricator must be able to fabricate the AP Blocks for an SRU. To specify the AP Blocks needed for an SRU in full would seem to require the design of an SRU. A complete atomically precise design and theoretical analysis (including relevant computational modeling) of an SRU should be feasible. Again, this analysis and modeling should be simplified by the fact that the design will be broken into blocks, as well as by the fact that the design of the SRU will be deliberately structured so that it can be analyzed theoretically and modeled computationally.

To summarize: while experimentalists will want the smallest AP Blocks with the fewest elements, the design team will want the largest possible AP Blocks with the largest palette of elements as this will simplify the design problem. The larger the blocks and the more elements, and the fewer restrictions on block structure, the easier it will be to design a workstation fabricator. Large complex blocks with many elements will unfortunately make the job of experimentalists more difficult.

It seems likely that some negotiations will be involved before the specifications of the AP Blocks are finalized. Some candidate AP Blocks are shown here to illustrate the concept (Figure 39).

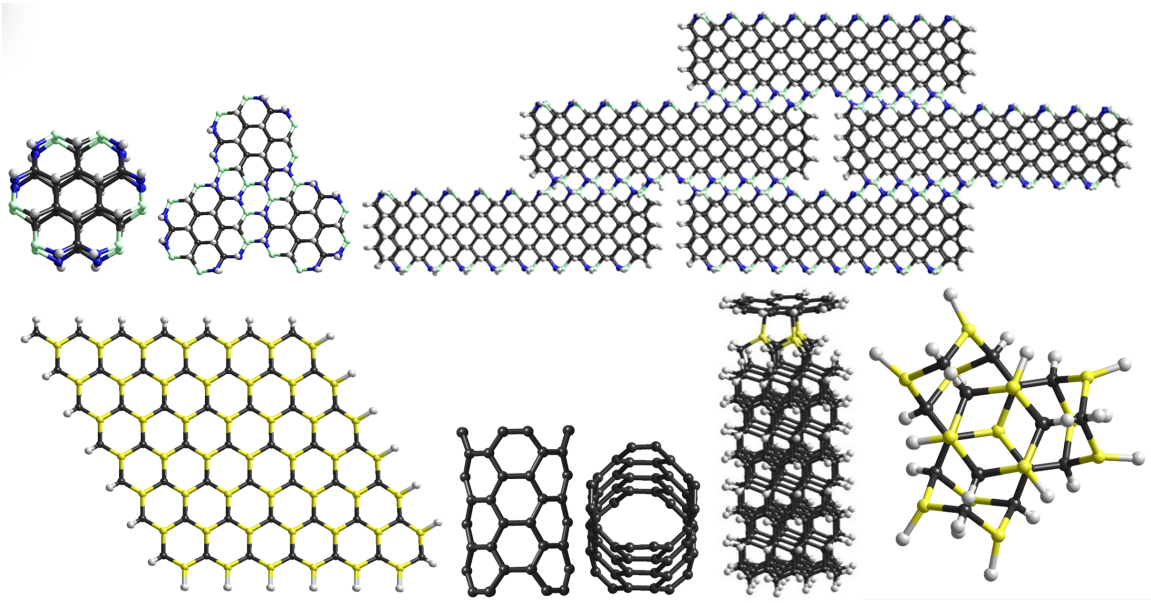


Figure 39: Candidate atomically precise building blocks. Some examples of candidate AP Blocks, including a segment of (4,4) nanotube and a proposed TriGe build site.

7.4 Decomposing large structures into small blocks

It seems very possible that early fabricators will have high error rates, limiting the size of blocks they can build. While it's unclear exactly how big such blocks might be, it's plausible to assume that early blocks might be limited to 100 to 1,000 atoms in size. Should this be true, it would be helpful to have some heuristic methods for decomposing larger diamond and graphene structures into similar structures made from small blocks.

We discuss one such heuristic method here.

7.4.1 Decomposing diamond into octahedrons and tetrahedrons

We start by assuming a large original structure is made from diamond.

We take the original solid diamond structure and conceptually slice it along {111} planes, where the {111} planes are defined by the usual Miller indices in the diamond crystal lattice. Slicing along these planes allows us to separate the structure into candidate small blocks, with each candidate small block connected to adjacent candidate small blocks by bonds that connect a single atom in one candidate small block to a single atom in the adjacent candidate small block. In one approach, we can slice any large diamond structure into octahedra and tetrahedra. We can then take each bond that connects a single atom in one candidate small block to a single atom in an adjacent candidate small block and make one of the atoms a nitrogen and the other a boron. This means the two candidate small blocks are now connected by dipolar bonds. We now clean up the messy regions where three or more candidate small blocks come together and cause problems (which is why this is a heuristic approach, and not an algorithm) to finish the decomposition.

The result of applying this approach to create a general set of octahedral and tetrahedral building blocks from which a structure like a diamond of arbitrary size can be built is shown in Figure 40.

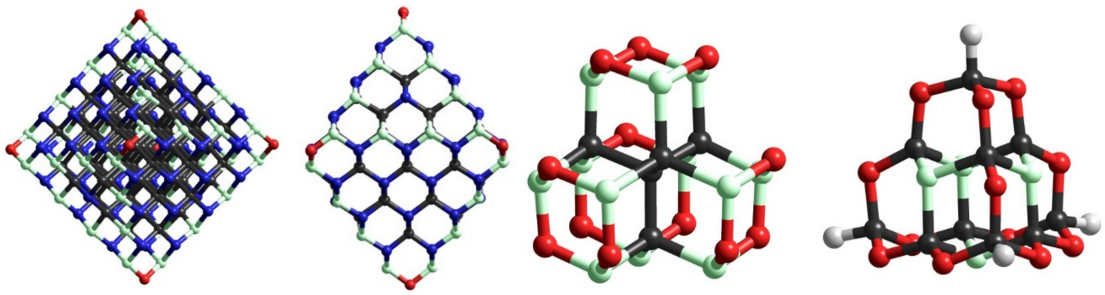


Figure 40: Octahedral and tetrahedral diamond building blocks. Two views of a 165 atom octahedron on the left, followed by two tetrahedra on the right, the first of 30 atoms and the second of 29 atoms.

One possible assembly is shown in Figure 41.

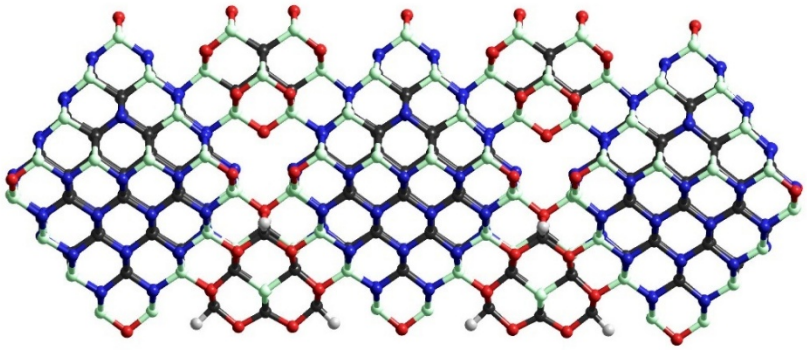


Figure 41: Assembly of octahedral and tetrahedral blocks with dipolar bonds.

In this example, we have one repeated octahedron and two slightly different repeated tetrahedra (Figure 41). It is possible to find examples where a single octahedral structure and a single tetrahedral structure are repeated to fill 3-space.

For a more general discussion of how octahedra and tetrahedra can be used to fill 3-space, see Introduction to space frames (Section 9.2.1).

Design cuts in the original diamond structure can be made along not only {111} planes, but also along {110} planes. Cuts along {110} planes can be dealt with in a manner analogous to that described for cuts along {111} planes. However, cuts along {100} planes require different treatment. They might be made most easily by deleting atoms and hydrogenating the resulting surfaces. This would not bond the resulting surfaces together, which could be a disadvantage. If we want to bond the two surfaces together with dipolar bonds, we might think about exploring a more speculative approach outside the baseline elemental palette: we might try cuts along {100} planes and then change atoms on one side of the cut to oxygen, and atoms on the other side of the cut to beryllium. Whether or not this approach works in vacuum and results in a stable beryllium-terminated surface that does not reconstruct in some undesired fashion, and which will also bond appropriately with the facing oxygen terminated surface, has not yet been evaluated. The hope is that each oxygen atom would form two dipolar bonds to two adjacent beryllium atoms in the neighboring block, and that each beryllium atom would form two dipolar bonds to two adjacent oxygen atoms in the neighboring block.

7.4.2 Decomposing a flat diamond sheet into blocks

A second example might involve cutting a diamond wall, perhaps intended for use in the shell of the SRU or the tines in a comb drive, into blocks (rectangular bricks) as shown in Figure 42.

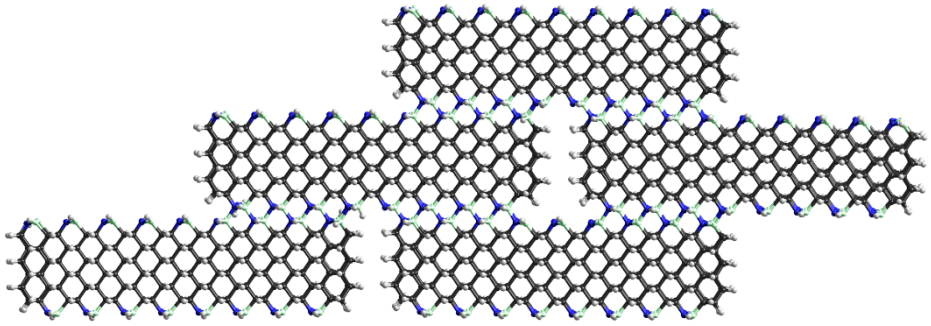


Figure 42: Rectangular molecular LEGO blocks assembled into a wall.

In this case, the resulting rectangular "bricks" (to use LEGO terminology) have three pairs of faces, with Miller indices of (110), ($\bar{1}$10), and (001). The (001) planes were not simply cut and replaced with B-N dipolar bonds, as were the upper and lower (110) surfaces, but in fact two layers of atoms adjacent to the cut plane were deleted, and the resulting dangling bonds were terminated with hydrogens. These (001) surfaces are not bonded together, weakening the structure even more than if dipolar bonds had replaced the carbon-carbon bonds in the original structure, but despite this, the resulting structure is stable because it is held together by the bonds on the upper and lower surfaces of the small blocks ("bricks") that resulted from the cutting process. For further information, see Molecular LEGO (Section 7.6).

7.4.3 A third decomposition of a flat diamond sheet into blocks

In Figure 43, we see a side view of a single AP Block. This AP Block has (110), ($\bar{1}$10), and (001) surfaces (like the previous AP Block). However, it differs in how the surfaces bond to each other. The surface facing us is the (110) surface and the surface hidden from us is the ($\bar{1}\bar{1}$0) surface. The surfaces facing up and down in this view are (001) and (00$\bar{1}$), and the surfaces facing left and right are (1$\bar{1}$0) and ($\bar{1}$10). All surfaces of the {110} family are terminated with boron and nitrogen (or oxygen). Note that boron can also form dipolar bonds with oxygen, as well as nitrogen. This AP Block can tile the (001) plane, bonding to four neighboring blocks through its four {110} faces in the [110], [1$\bar{1}$0], [$\bar{1}$10] and [$\bar{1}\bar{1}$0] directions, by dipolar bonds.

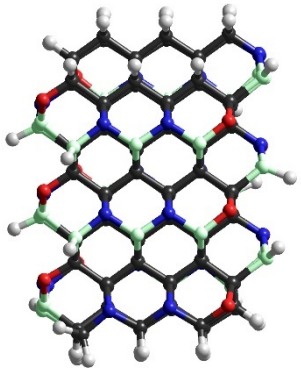


Figure 43: Diamond sheet building block with four bonding directions.

For further information on this AP Block, see Extending the standard 2 by 4 LEGO brick (Section 7.6.11).

7.4.4 Decomposition of hexagonal diamond into hexagonal blocks

Rather than decomposing a block of diamond into blocks, we might try to decompose a sheet of hexagonal diamond (also known as lonsdaleite) into hexagonal blocks.

The three proposed blocks each have 72 atoms: 12 boron, 12 nitrogen, 24 carbons and 24 hydrogens. Swapping borons and nitrogens in these three blocks creates two new blocks and one that remains the same.

As shown in Figure 44, these blocks have a hexagonal motif. They have a carbon interior which, because it is like unstrained hexagonal diamond, provides a stable core. The front and back surfaces are hydrogenated, which gives us a block that can be used to form sheets. The edges of the sheet use boron and nitrogen (Figure 44).

When the B and N atoms are properly arranged on the edges of a block, multiple identical copies of the same block can be assembled using dipolar bonds between the blocks (Figure 45). This lets the blocks tile the plane.

All five types of blocks can be fused using dipolar bonds and can tile the plane. We illustrate this by showing three copies of one type of block fused into a single assembly (Figure 45).

The three assembled blocks are ∼3,124 zJ (19.5 eV) more stable when assembled than they are unassembled, according to HyperChem's implementation of AM1 (Figure 45).

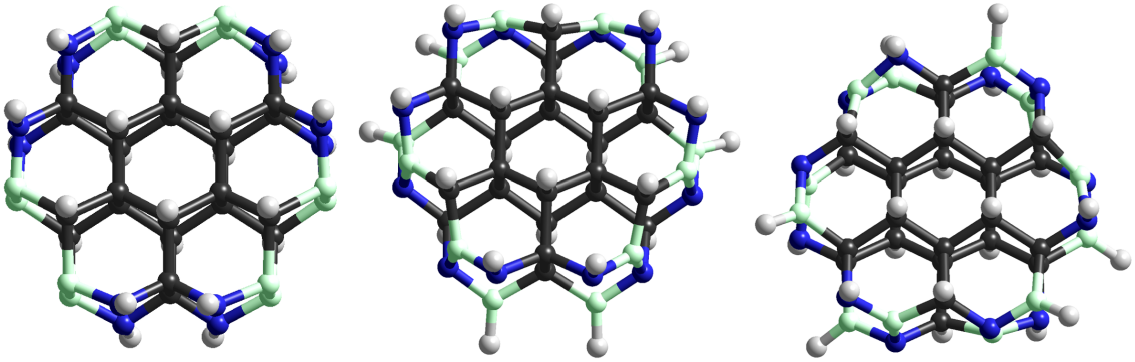


Figure 44: Hexagonal diamond building block.

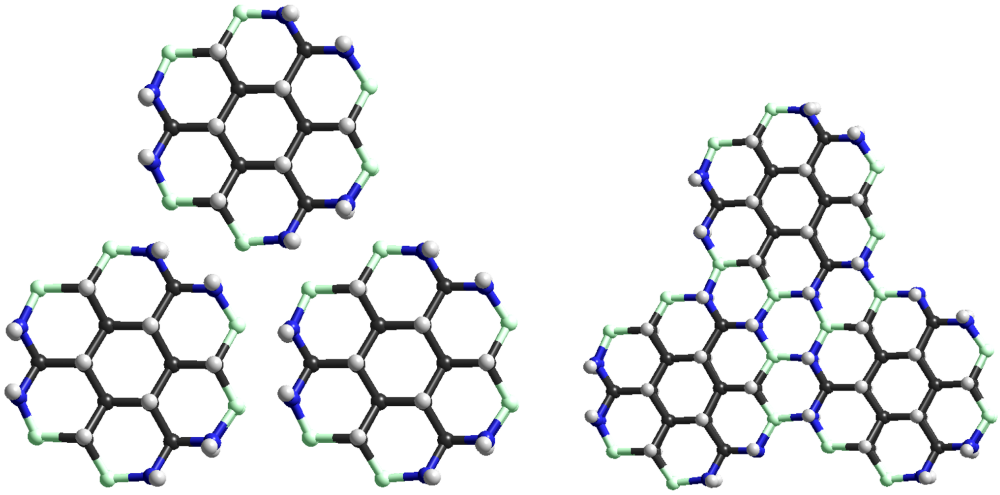


Figure 45: Three hexagonal diamond blocks fused into an assembly.

As there are four dipolar bonds between each pair of blocks and three interfaces in the three-block assembly, this implies an average stabilization of approximately 260 zJ (1.62 eV or 37.5 kcal/mol) per dipolar bond. In a wall, each block would be connected to its six neighbors by 24 such bonds, corresponding to an estimated total removal energy of approximately 6,240 zJ (39 eV or 900 kcal/mol).

The bonded pair of 1-azaadamantane and 1-boraadamantane would be a useful model system. Study of this system could be used to estimate the strength of the dipolar bonds between these proposed blocks.

The preliminary AM1 results suggest that the five proposed blocks may be sufficiently stable for assembly in UHV at cryogenic temperatures. Higher-level calculations and ultimately experiment are needed to test this conclusion.

These blocks should be able to tile the plane, as illustrated by the 19 identical hexagonal AP Blocks in Figure 46.

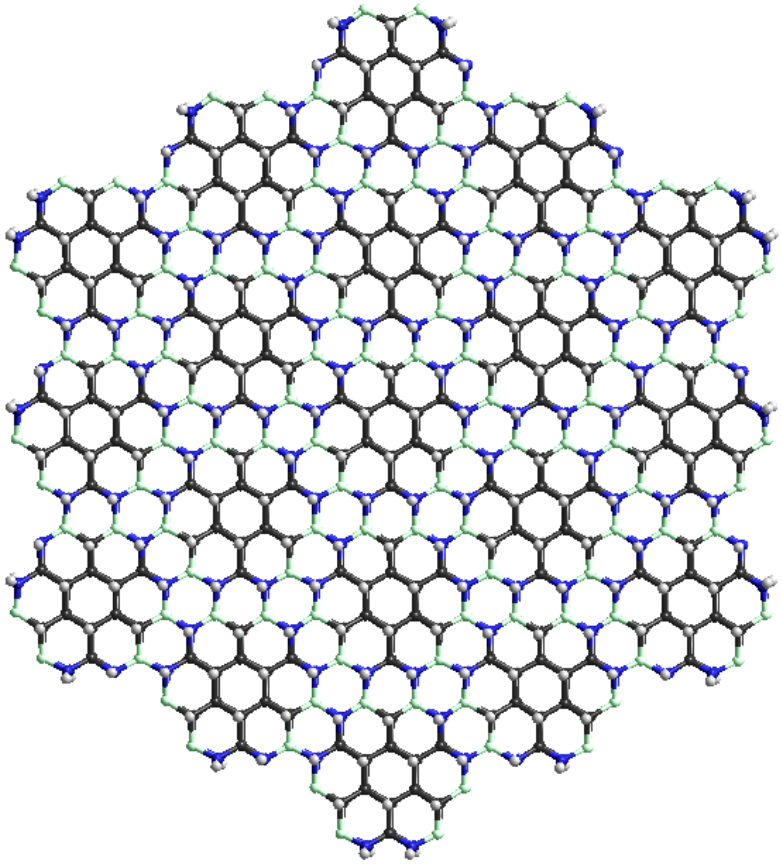


Figure 46: Nineteen hexagonal blocks tiling a segment of a wall.

7.4.4.1 Leakage of H_2 through a sliding seal

During extrusion of the daughter from the parent SRU, the exterior of the daughter's shell and the interior of the external extrusion sleeve (ES) slide against each other. Assuming both surfaces are made from an extended structure composed of hexagonal blocks, as depicted in Figure 46, we might ask whether an H_2 molecule could squeeze between the two surfaces. A simple approximation to

this problem is to calculate the energy required to insert an H_2 between two surfaces each composed of 19 hexagonal blocks (each like the 19-block assembly in Figure 46). Using HyperChem's MM+ molecular mechanics force field, we find this energy is ~243 zJ if the two 19-block assemblies are rotationally aligned with each other. However, if we initially misalign the two assemblies rotationally so that their facing crystal surfaces do not match up, the energy required to insert H_2 declines to 187 zJ. Thermal noise, k_BT at 210 K, is ~2.9 zJ. The insertion energy for misaligned blocks is therefore approximately $64.5\,k_BT$, corresponding to a Boltzmann factor of approximately 10^{-28}. If this factor bounds the leakage probability, as must be established by further research, the concern that such leakage might compromise the eutactic environment within the shell of an SRU would be effectively excluded.

If further calculations show that 187 zJ is a lower bound on the activation barrier along the minimum-energy leakage pathway, the following deliberately conservative attempt-rate calculation applies.

For a single SRU, H_2 has ~10^5 seconds to enter the crack between the external extrusion sleeve (ES) of the parent and the shell of the daughter. This crack is 2 × 100 nm + 2 × 60 nm = 320 nm in length. If we assume the feedstock is one percent H_2 (well beyond the worst case) and that the attempt rate is $10^{12}\,s^{-1}$ per nm of crack (higher than plausible) then there will be $320 \times 0.01 \times 10^{12}$ attempts per SRU per second, or $320 \times 10^{10} \times 10^5$ attempts per 10^5 seconds, or 3.2×10^{17} attempts per SRU per replication cycle. Even given these conservative assumptions, the failure rate per SRU per replication cycle is estimated to be $<3.2 \times 10^{17} \times 10^{-28} = 3.2 \times 10^{-11}$. This is 2.8×10^{-11} failures per SRU per day. As the nanofactory has 5.8×10^{13} SRUs the number of failed SRUs per day from this cause would be $2.8 \times 10^{-11} \times 5.8 \times 10^{13} = 1{,}624$. Among 5.8×10^{13} SRUs, approximately 1,624 such failures per day would be negligible relative to the baseline failure rate allowance of <0.0864% per SRU per day.

The estimated radiation caused failure rate is almost five orders of magnitude higher: 2.6×10^{-6} radiation hits per SRU per day (see Section 6.2.1). Once we reduce the failure rate from leakage to below the failure rate from radiation, we have effectively done as well as we need to do, although further investigation of the actual radiation caused failure rate is still needed.

The baseline assumes that it is possible to design sliding seals that interpose a high enough barrier to the relevant minimum-energy pathway between the SRU's external (non eutactic) environment and its internal (eutactic) environment to effectively reduce the probability of contaminant entry to an acceptable level, specifically, to a level low enough to meet our 0.1% failure rate per SRU per replication cycle. A preliminary calculation suggests this is possible. Further research to validate this assumption is needed.

We adopt sliding seals able to effectively block contaminant entry as our baseline.

7.4.5 Decomposition into graphene strips

We should be able to decompose a graphene sheet into strips that can be linked together with dipolar bonds (Figure 47).

This strip has 262 atoms and is 8.1 nm long. It should be conductive along its length. Multiple such strips might be linked through dipolar bonds to create a sheet, which should also be conductive along the strips (Figure 48).

These strips might be held in place on structural elements by simple van der Waals forces, making them useful as conducting surfaces for the tines of electrostatic comb drives.

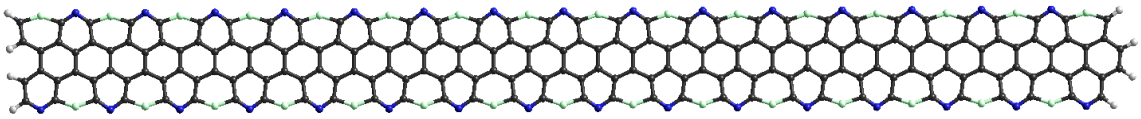


Figure 47: Graphene strip with dipolar bonding sites.

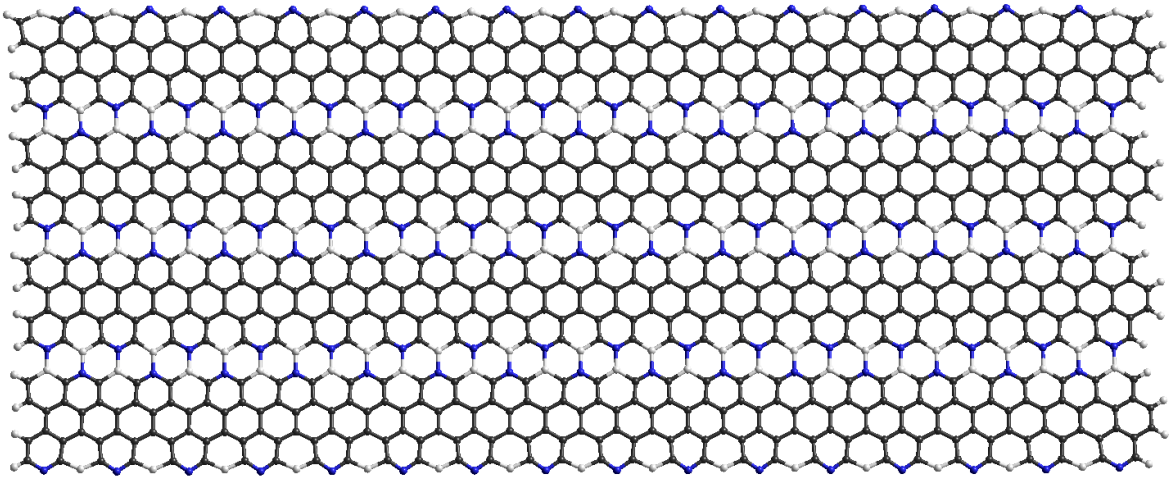


Figure 48: Graphene strips linked into a sheet.

7.4.6 Decomposition into coronene-like blocks

We can repeat analogous statements for graphene based structures based on coronene-like blocks (Figure 49).

Figure 50 shows an assembly of coronene-like graphene blocks.

Figure 51 shows an additional view of coronene-like graphene blocks.

7.4.7 Decomposition into parallelograms

Hexagonal blocks are not the smallest blocks linked by dipolar bonds that can tile the plane in a hexagonal pattern. The block in Figure 52 can tile the plane like parallelograms can tile the plane.

The core of this block is again carbon (Figure 52). AM1 predicts it will be stable in vacuum. Synthesis in solution might be challenging. Mechanosynthesis should be feasible. This block has 48 atoms: 12 C, 10 B, 10 N, and 16 H.

If we arrange four such blocks so that they can be assembled by dipolar bonds, we get the arrangement in Figure 53.

We can then assemble the four blocks using only dipolar bonds, producing the assembly in Figure 54.

Again, we anticipate the barrier to assembly will be negligible (Figure 54). Having shown how four such blocks can be assembled, it should be clear that this block can tile the plane, and that this tiling should be able to prevent passage of gas or liquid unless the tiling itself is physically disrupted by, for example, high temperature or very reactive chemicals.

Again, further research is required.

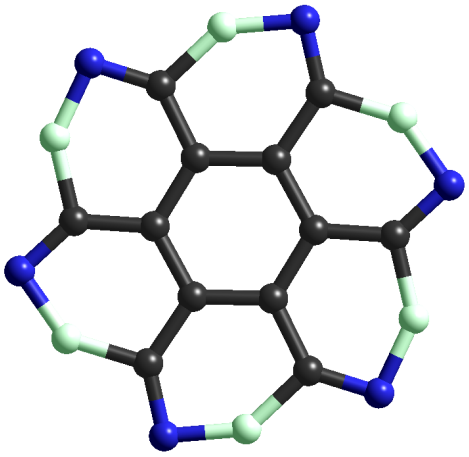


Figure 49: Coronene-like graphene building block.

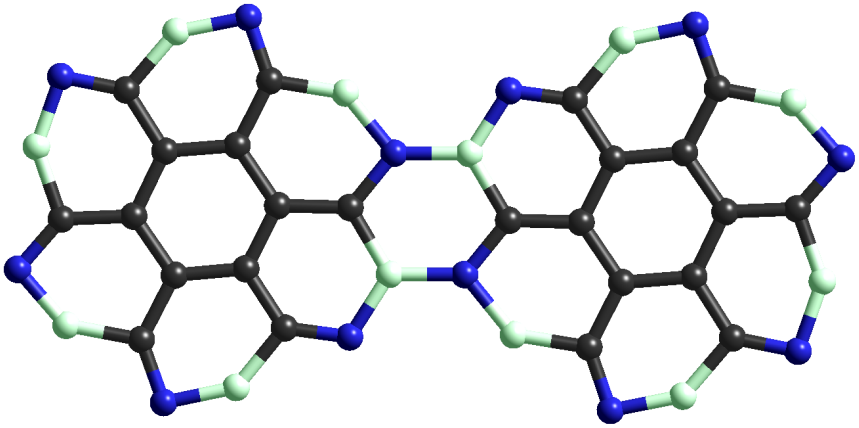


Figure 50: Assembly of coronene-like graphene blocks.

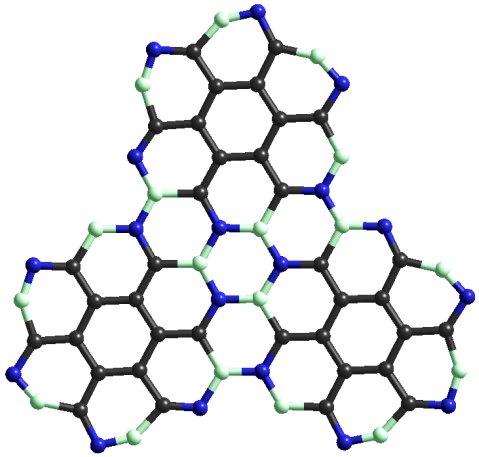


Figure 51: Three assembled coronene-like graphene blocks.

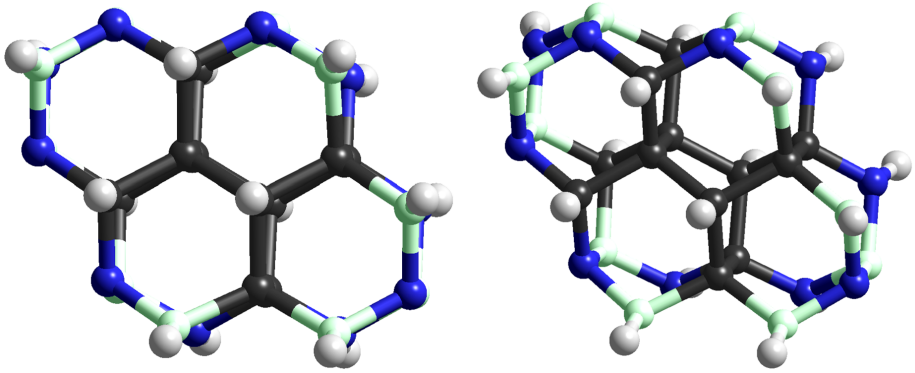


Figure 52: Parallelogram building block with a carbon core.

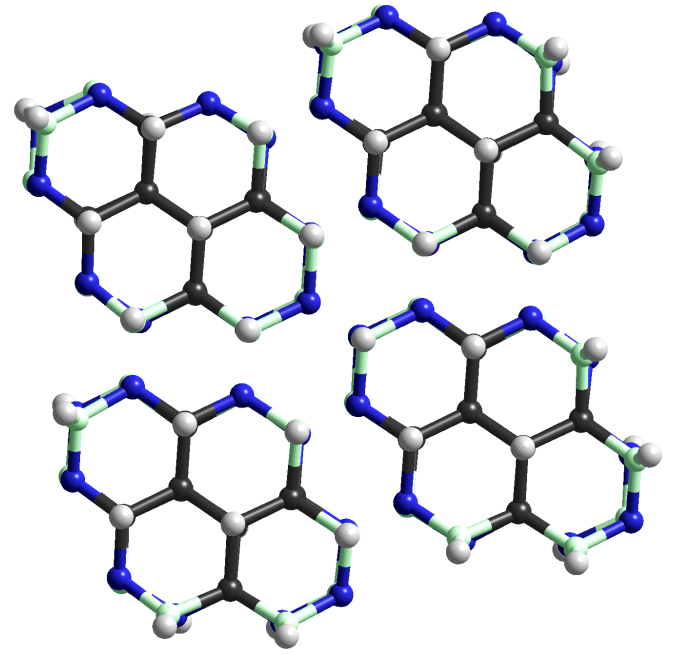


Figure 53: Four parallelogram blocks arranged for assembly.

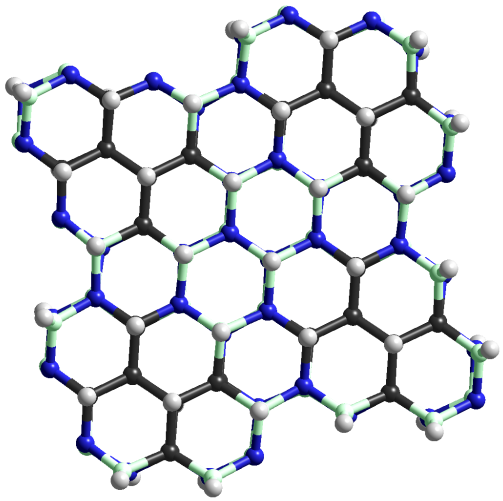


Figure 54: Four parallelogram blocks assembled with dipolar bonds.

7.4.8 Decomposition into graphene parallelograms

Similar arguments can be applied to the graphene analog of the preceding block (Figure 55).

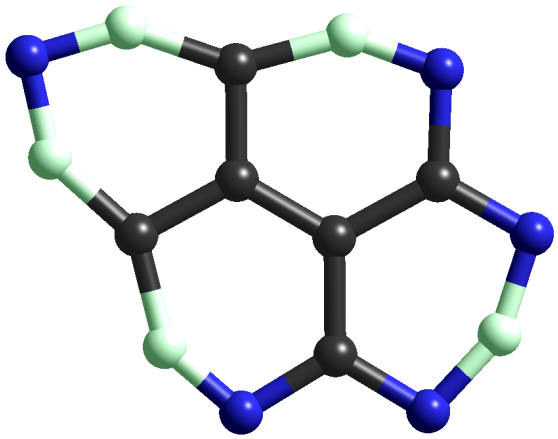


Figure 55: Graphene parallelogram building block.

Figure 56 shows an assembly of graphene parallelogram blocks.

7.4.9 Decomposition into H-stabilized BN blocks

The previous blocks used a carbon core to provide stability and surface borons and nitrogens to allow assembly with dipolar bonds. In a pure BN block we eliminate carbon and use only boron and nitrogen. The stability of the isolated blocks is less certain and will require more analysis. However, if blocks are to be manufactured by the fabricator using mechanosynthesis, which in this design takes place in vacuum, then we should review blocks that theoretically should be stable in vacuum and might lend themselves to assembly in vacuum. Wurtzite boron nitride (w-BN), the boron nitride analog of hexagonal diamond, has basal (0001) surfaces analogous to the (111)

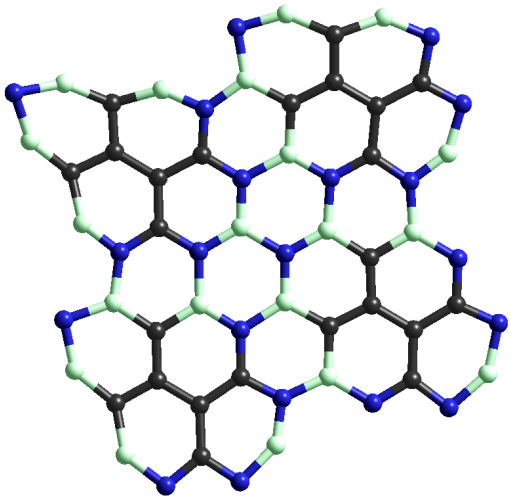


Figure 56: Assembly of graphene parallelogram blocks.

surfaces of cubic diamond, but with different stacking. Calculations support hydrogen termination as a means of stabilizing cubic BN (111) surfaces [RL07]. By analogy, hydrogen termination might stabilize the corresponding (0001) surfaces of w-BN blocks such as that in Figure 57. The stability of this particular block in vacuum requires further analysis.

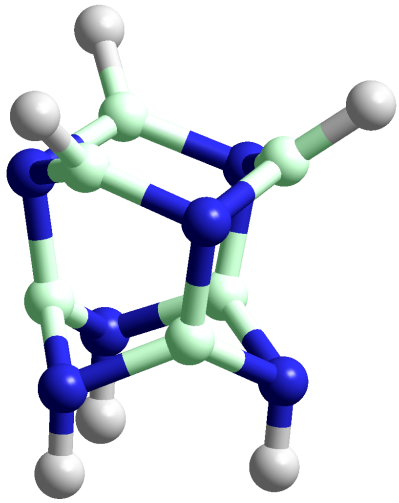


Figure 57: Hydrogen-stabilized boron nitride building block.

Individual blocks could be gripped by an adamantane gripper which had a boron and a nitrogen, allowing it to form two dipolar bonds with such a block, as shown in Figure 58.

7.4.10 Decomposition into w-BN blocks

HyperChem's implementation of AM1 predicts that the minimal w-BN "block" in Figure 59 is unstable (has an imaginary vibrational frequency).

Interestingly, HyperChem's AM1 predicts that the molecular structure in Figure 60, corresponding

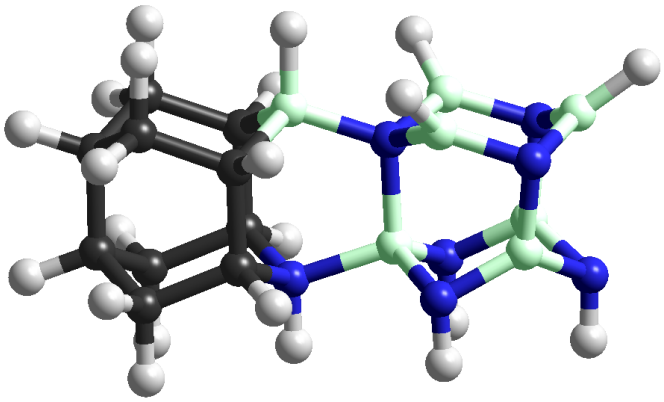


Figure 58: Adamantane gripper holding a boron nitride block. The adamantane gripper is on the left, the gripped block is on the right. The adamantane gripper on the left would have to be extended to include the handle, which is not shown in this model.

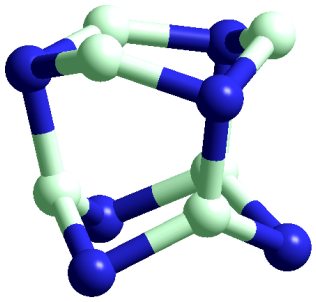


Figure 59: Minimal wurtzite boron nitride block.

to that same "block" held by a gripper, ***is*** stable (has all positive vibrational frequencies) in vacuum.

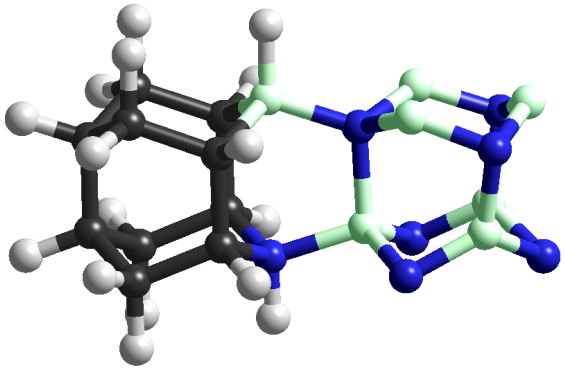


Figure 60: Minimal wurtzite boron nitride block held by a gripper.

Fusing two such minimal w-BN blocks produces the structure in Figure 61.

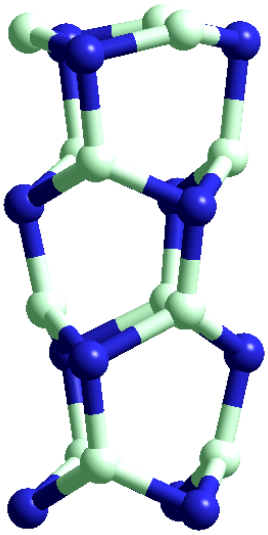


Figure 61: Two fused minimal wurtzite boron nitride blocks.

This not-quite-minimal piece of w-BN is, according to HyperChem's implementation of AM1, stable (has all positive vibrational frequencies) in vacuum (Figure 61).

This suggests a possible strategy for synthesizing w-BN in a wide variety of shapes using mechanosynthesis. Use minimal blocks of w-BN that are stable when held by a gripper and fuse them into larger w-BN assemblies that can maintain their stability once fused into the larger structure.

Possible instability of the w-BN (0001) surface might also be addressed by using blocks that are H-passivated on one surface, such as the block in Figure 62.

This block can then be held by a gripper, as shown in Figure 63.

The gripped block in Figure 63 illustrates how three grippers could hold and position blocks during assembly, as shown in Figure 64.

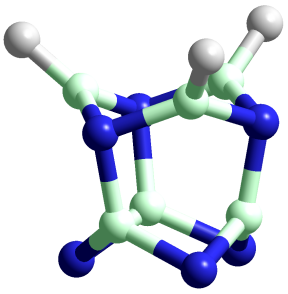


Figure 62: Wurtzite boron nitride block with a hydrogen-passivated surface. HyperChem's AM1 predicts this structure is stable.

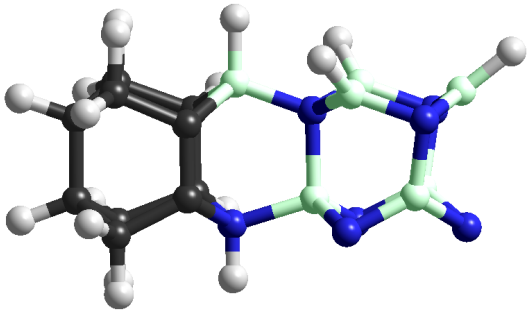


Figure 63: Gripper holding a hydrogen-passivated wurtzite boron nitride block. HyperChem's AM1 predicts this structure is stable.

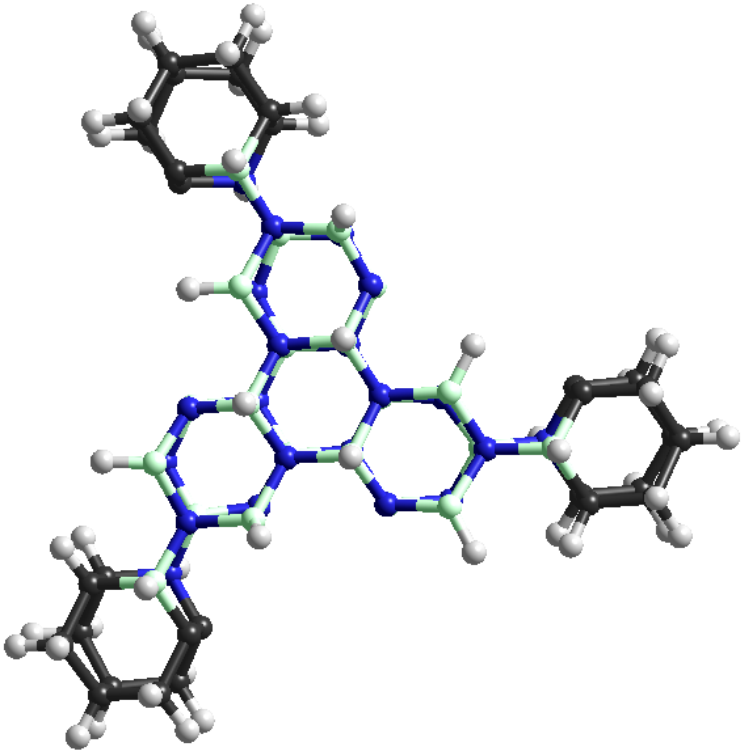


Figure 64: Assembly of wurtzite boron nitride blocks using three tools.

Removing the tools leaves the assembly shown in Figure 65.

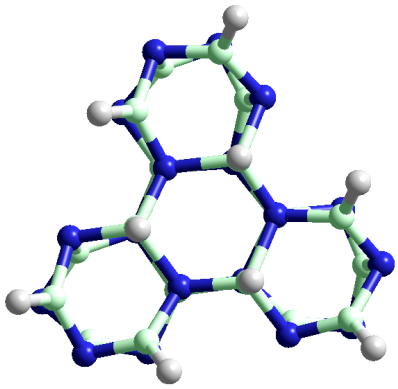


Figure 65: Wurtzite boron nitride assembly after removal of the tools. HyperChem's AM1 predicts this structure is stable.

Clearly, further research on what w-BN and cubic boron nitride (c-BN) structures can be built, and what tools can be used to build them, is needed. AM1 is used here to provide only a preliminary assessment of the possibilities.

7.4.11 Decomposition into h-BN blocks

Figure 66 shows the simplest h-BN block.

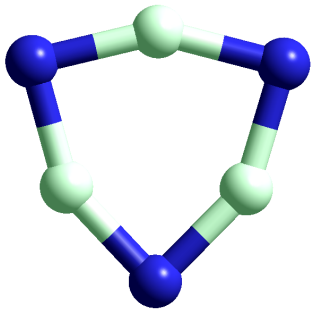


Figure 66: Minimal hexagonal boron nitride building block.

If we attach this block to a gripper (without the handle) we get the structure in Figure 67.

When three minimal h-BN blocks are fused together, we get the structure in Figure 68.

HyperChem's AM1 also predicts that the three-block assembly is stable (Figure 68).

HyperChem's AM1 predicts the coronene analog will also be stable (Figure 69).

Because sheets of h-BN can be important components in comb drives, it would be highly desirable to find stable h-BN blocks that could be assembled into larger sheets of h-BN suitable for use in such comb drives. See Differential comb drives (Section 4.3.4).

Ideally, such a set of h-BN blocks would (1) be stable at the relevant operating temperature in UHV and (2) be assemblable into larger h-BN sheets. While some purely B and N blocks would be

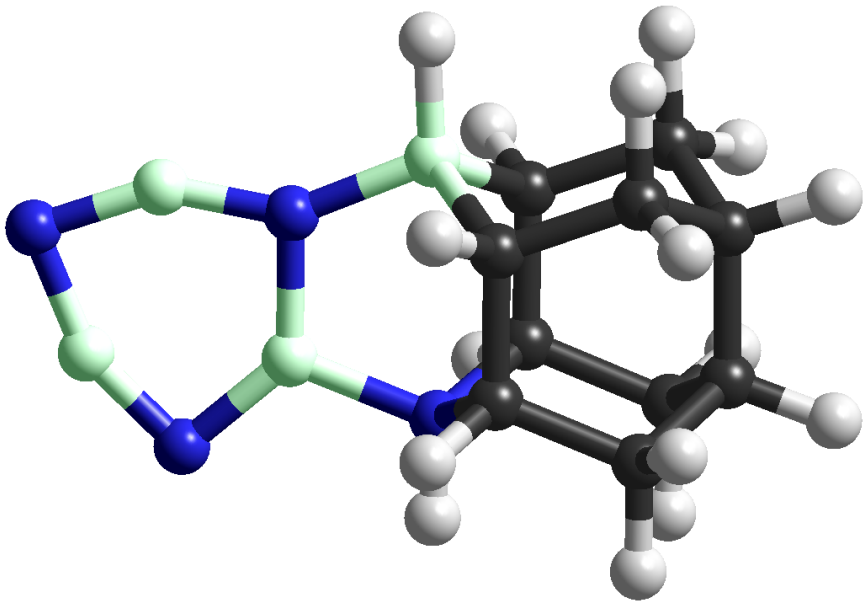


Figure 67: Gripper holding a hexagonal boron nitride block.

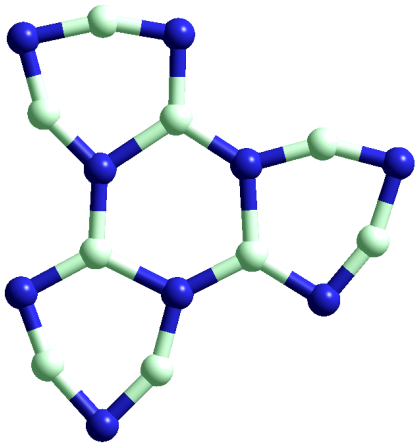


Figure 68: Three fused hexagonal boron nitride blocks.

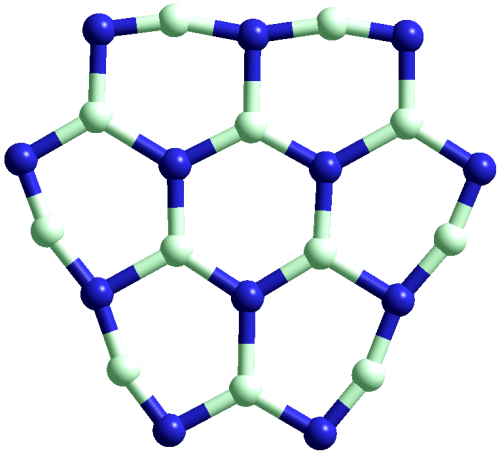


Figure 69: Hexagonal boron nitride analog of coronene.

needed to assemble sheets of h-BN, blocks that were H passivated on at least some edges would be needed to assemble an H-passivated h-BN sheet.

Should the simplest h-BN block in Figure 66 prove stable at some suitable temperature in UHV and be amenable to handling by a gripper, and should certain basic assembly operations prove feasible, then large sheets of h-BN could be made by mechanosynthesis. Further research on this possibility seems appropriate.

7.4.12 A shell without interblock chemical bonds

A possible design for a set of blocks without interblock chemical bonds might be based on the conceptual LEGO models in Figures 70–73.

The first building block, shown in Figure 70, left panel, is a flat plate with four posts. These plates can tile a plane, shown in Figure 70, right panel. Note that four posts from four adjacent plates form a 2×2 cluster on the surface of the tiling.

We can now create a second building block, shown in yellow in Figure 71, which we might call a "ring plate" because it has a hole in the center which forms a ring around the 2×2 cluster created by the four posts from four adjacent flat plates. Again, the ring plate has four posts. The image in Figure 71 shows, on the left, a ring plate in isolation and, on the right, a single ring plate ringing a 2×2 cluster of posts, holding the four flat plates together. Note that AP Blocks based on this motif would be held together by van der Waals forces: the ring plate would be held in place by its attraction to the four flat plates beneath it, and the four flat plates would be held together both by van der Waals forces, and also by the ring plate holding the four posts from the four flat plates together.

Add more ring plates to the surface of the plane of flat plates, holding the flat plates together (Figure 72).

We can hold the first layer of ring plates together by adding a second layer of ring plates (Figure 73). The yellow and blue ring plates are identical in design; the color difference is used simply to clarify the fact that there is now an additional layer of ring plates added on top of the first layer. Additional layers can be added, as might prove useful in stabilizing the structure. The entire structure relies only on van der Waals forces.

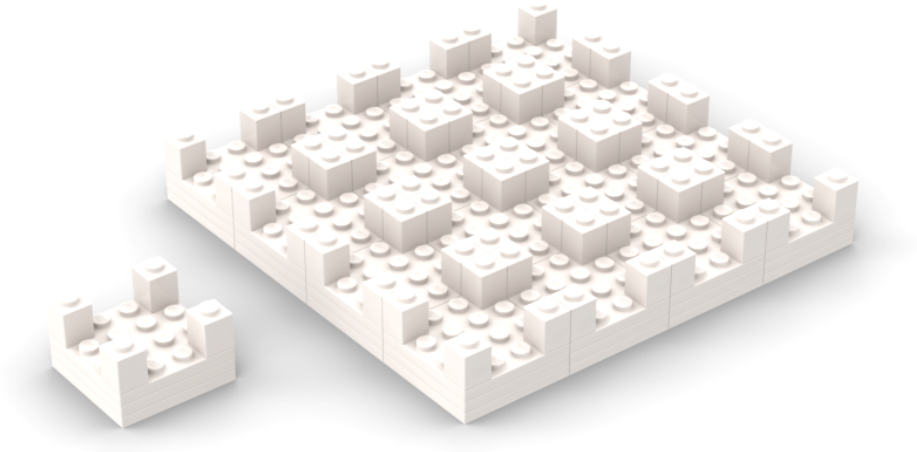


Figure 70: LEGO model of flat plates and their planar tiling.

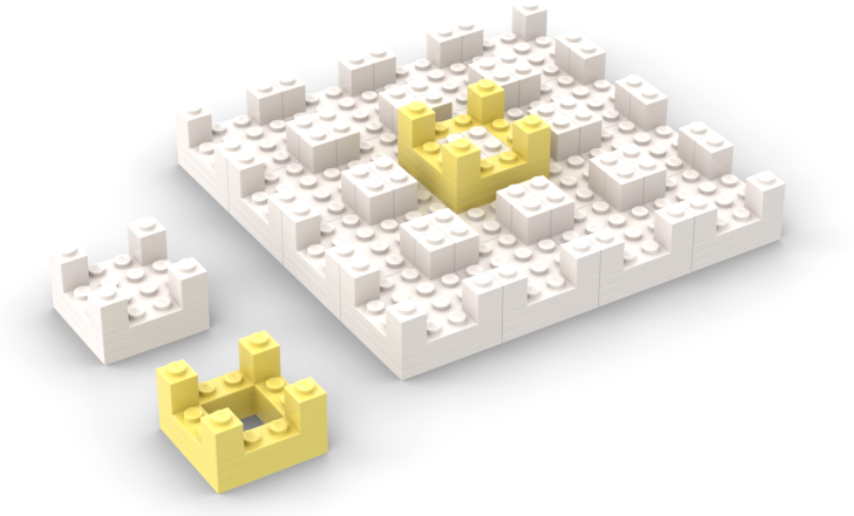


Figure 71: Ring plate joining four adjacent flat plates.

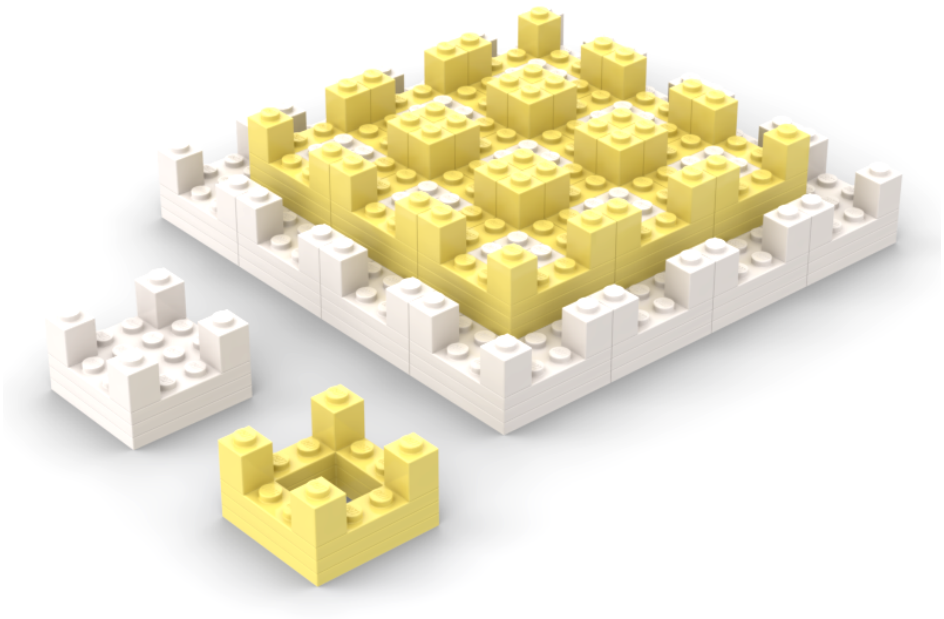


Figure 72: Ring plates securing a plane of flat plates.

Note that the bumps on the LEGO bricks would not be present or used on the blocks derived from this concept, and are present only because the software used to generate these images adds them (Figure 73).

This planar structure can be extended in X and Y as far as might be desired (Figure 73). Joining two such planar structures at a 90-degree corner will require the design of a few additional custom blocks.

The flat plates are characterized by a height H (not including the post), an X and Y size, and a post X and Y size: PX and PY. The height of the post above the flat plate is also H, making the total height of the plate plus post 2H. That is, the flat plates are characterized by five numbers: H, X, Y, PX, and PY.

The ring plates also have a ring, characterized by an X and Y dimension RX and RY, as well as the dimensions of the flat plate. Each ring plate is characterized by seven numbers: H, X, Y, PX, PY, RX, and RY.

By adjusting the different dimensions and using the three types of diamondoid blocks that have regular rectangular edges (with surface orientations (100), (010), (001), (110), $(\bar{1}10)$, (001), and (111), $(\bar{1}10)$, $(11\bar{2})$) as the basis for our design, we obtain a large set of flat plates and ring plates that we might explore. We could extend the design space, but this range should suffice for now. The impact of these parameters on the van der Waals forces, and the consequences for the structural integrity and permeability of the resulting wall, can reasonably be investigated. It seems likely that walls that are sufficiently leakproof to reasonable feedstock solutions can be found without unreasonable effort.

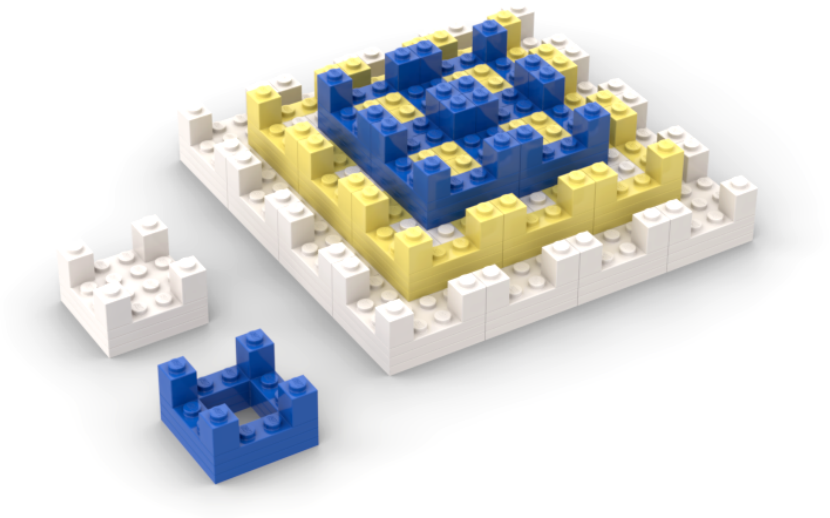


Figure 73: Two layers of ring plates stabilizing the underlying plates.

7.4.13 On the relative merits of B-N bonds versus C-C bonds

The diamond (110) surface is stable with respect to reconstruction. Fusing two such stable surfaces on two AP Blocks might be a more desirable alternative to the use of B-N dipolar bonds. Two such fused AP Blocks would form a stronger and more rigid connection between the two AP Blocks, as fusion allows the diamond crystal structure to continue, without interruption, across the assembly interface. On the other hand, there are some arguments against this. We consider the pros and cons of the two approaches. Note that the comments here could be extended to include oxygen and possibly a few other elements.

Favoring the B-N bonds would be the following:

Firstly, correctly fusing two (110) diamond surfaces requires greater positional accuracy. Misalignment along the trough by over 125 pm (half of the ∼250 pm distance between carbon atoms) would place each carbon radical more than halfway towards the next carbon radical, thereby likely causing it to bond to the wrong carbon atom. In a simplified registration argument, alternating B and N atoms along the trough doubles the complementary repeat distance from approximately 250 pm to 500 pm. A shift of 250 pm aligns N with N and B with B, preventing the intended dipolar bonds. The halfway threshold for sliding into the next complementary potential well is thus about 250 pm, rather than the approximately 125 pm for carbon. This is a registration argument, not a calculated potential energy surface or a validated positional tolerance. By using a random placement of B and N atoms on a surface, it should be possible to effectively prevent two AP Blocks from bonding at all unless they are in perfect alignment, using an approach that more closely resembles self-assembly in solution than positional assembly in vacuum. Random assignment does have the drawback that N-N and B-B bonds would be present on the resulting surfaces, reducing stability.

Secondly, disassembling two AP Blocks held together with dipolar B-N bonds should be feasible,

allowing the correction of errors. The value of this depends, of course, on both the error rate and the cost of errors. If errors are sufficiently infrequent, then this is of little value. Alternatively, if errors occur with some frequency, but are of little consequence, then this would also be of little value. An example of the latter might be if we used fusion of AP Blocks to build small assemblies that could be tested for correctness. If the small assembly failed the test, the whole assembly could be discarded at little cost. On the other hand, if a large structure that could not be discarded was being assembled, such as a feedstock transport system inside an ultra-high vacuum, then a single leak might be catastrophic and the inability to remove a fused piece placed in error might have a very high cost. In such a case, the use of dipolar B-N bonds that allowed easy removal of a piece placed in error would have high value.

Thirdly, the use of B and N allows the stabilization of the (111) surface, which reconstructs if the surface is purely unpassivated carbon. If B and N atoms are appropriately placed on this surface, it can be quite stable and not reconstruct. It would likely be easier to design stable surfaces and small AP Blocks (which are mostly surface) if we allow the inclusion of B and N in their design. If we do not allow either B or N, the (111) surface would have to be stabilized by hydrogen passivation.

Favoring the use of fused surfaces would be the following:

Firstly, C-C bonds are much stronger than dipolar B-N bonds. The force required to break a dipolar B-N bond is smaller than that required to break a C-C bond.

Secondly, the surface area required to bond two AP Blocks together with similar strength should be smaller, leading, possibly, to smaller AP Blocks.

Third, a C–C bond is stiffer than a dipolar B-N bond, though this argument is not as compelling as the first. The tensile strength of a chain is limited by the strength of its weakest link. On the other hand, the stiffness of that chain in tension is moderately dependent on a weaker link. For example, in an eleven-link chain, making one link 20% as stiff as each of the other ten reduces the chain stiffness by about 27% relative to eleven equally stiff links.

Fourthly, the inclusion of B and N complicates many system issues, including the need to develop appropriate molecular tools, reaction sequences for B and N, feedstock molecules, metabolic sequences, and the like.

Summary

The general heuristic applied to cuts along the {111} and {110} planes relies on replacing carbon atoms with boron, nitrogen or oxygen atoms in such a fashion that small blocks are cut out from a larger block of diamond, these small blocks having facing surfaces that have complementary boron and nitrogen (or possibly oxygen) atoms that create dipolar bonds that hold them together. When a cut along a (001) plane is viewed as desirable, strength is sacrificed by simply removing one or two layers of atoms adjacent to the cut and hydrogenating the exposed dangling bonds.

The resulting structure is weaker than the original all-carbon structure but has the same general shape. The electrical properties might have been altered. If general shape is the primary property of concern, this approach seems useful.

This approach should also be useful when applied to graphene structures. A nanotube, for example, might be separated into separate distinct segments, with each segment terminated by rings of boron and nitrogen atoms that would be complementary to the rings of nitrogen and boron atoms that terminate adjacent segments. Again, the electrical properties might be altered, but the mechanical properties should be similar, if somewhat weaker.

This approach requires, of course, the ability to make blocks that include at least boron and nitrogen, and possibly oxygen. As a consequence, implementing this approach will require the

design and synthesis of appropriate molecular tools and reaction sequences for using them. It should be possible to make relatively small blocks using this approach, under a hundred atoms, as exposed surfaces of boron and nitrogen will be able to stick directly to complementary surfaces of nitrogen and boron on other blocks.

It should be noted that while the specific examples in this discussion illustrate the principles involved, they might exhibit unknown errors or pathologies. Until appropriate computational modeling and review have been conducted, or experimental validation confirms their expected functioning, there is a risk that these particular examples might not work even if the general principles are correct.

7.5 Simple AP Blocks using B and N

As one of the premises of this design approach to a nanofactory is that we can design it from AP Blocks, it might be worth taking some time to consider why there are reasons for believing that we can build a truly wide range of possible products, including simple nanofactories, from even simple (small atom count) AP Blocks. This section gives examples of some simple AP Blocks that might be used to build a reasonably wide range of structures.

Dipolar bonds, especially between boron and nitrogen, can be used to form a wide range of structures [Mer00, AS02, GAMJ10, GJMA09]. Two molecules that might be considered members of one of the simplest classes of AP Blocks are tetra-aza-adamantane and its boron analog, tetra-bora-adamantane [GAMJ10, GJMA09] (Figure 74). Tetra-aza-adamantane has been synthesized and is chemically stable. "This white crystalline compound is highly soluble in water and polar organic solvents." [Ott].

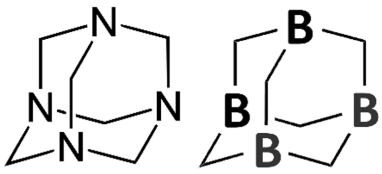


Figure 74: Chemical structures of tetra-aza-adamantane and tetra-bora-adamantane.

There is no readily available record of the synthesis of the boron analog. Assuming it is stable in vacuum, the two could be used to build structures that would be analogous to silicon carbide, with the nitrogen substituted adamantanes replacing silicon and the boron substituted adamantanes replacing carbon in the molecular structure. The Si-C bonds in the silicon carbide would be replaced by dipolar bonds between B and N atoms in neighboring adamantane molecules. Theoretical calculations support the stability of the resulting crystal [GAMJ10, GJMA09]. The 2009 study reports:

> "Additionally, a hypothetical molecular crystal in a zincblende structure, involving the pair tetra-bora-adamantane and tetra-aza-adamantane, was investigated. This molecular crystal presented a direct and large electronic bandgap and a bulk modulus of 20 GPa. The viability of using those functionalized molecules as fundamental building blocks for nanostructure self-assembly is discussed." [GJMA09]

It's not clear that this pair of molecules provides sufficient design flexibility to support the functionality of a workstation fabricator or an SRU, but they would seem to provide a very large design

space with a remarkably small class of molecules. Even a modest increase in the size of the allowed AP Blocks from the size of this pair would cause a combinatorial explosion in the design space, making it very difficult to search and therefore very difficult to make definitive statements about what is not possible. The image in Figure 75 is from [Mer00] and shows "a central 1,3,5,7 tetrabora-adamantane MBB linked to four surrounding 1,3,5,7-tetraaza-adamantane MBBs."

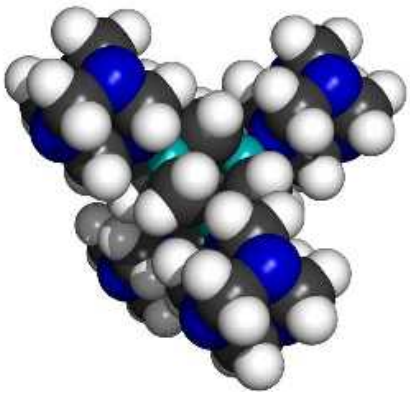


Figure 75: Tetrabora-adamantane linked to four tetraaza-adamantanes, from [Mer00].

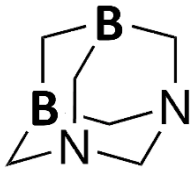


Figure 76: A diaza-dibora-adamantane molecular building block for extended assemblies.

A single building block (Figure 76) might suffice to create a very large space of possible structures. Whether or not its actual properties match the idealized properties one might hope for requires further research.

In essence, this block can make structures that resemble sp^3 carbon-based polycyclic structures, with this block substituted for carbon atoms in the structure. The orientation of this diaza-dibora-adamantane must be correct, or instead of a dipolar (B-N) bond, there will be a non-bond between either two nitrogens (N-N) or two borons (B-B). This molecule could form five- and seven-membered rings of itself, unlike the two molecules shown earlier, which can only form rings with an even number of members. Its synthesis might be a worthwhile challenge, should theoretical modeling show that it is stable in vacuum and could be used in the manner hoped for. However, even if it could be so used, designing all the components of a fabricator from this minimal AP Block appears quite challenging. This 22-atom minimal AP Block does inspire at least some hope that AP Blocks with fewer than 50 atoms and using only C, H, Ge, N, and B (see also Section 7.13) might be sufficient, if only we were clever enough, to make a workstation fabricator. Giving ourselves some additional leeway by increasing the atom count to a few hundred arguably gives us a plausible chance to design AP Blocks sufficient for a workstation fabricator.



Figure 77: Standard 2-by-4 LEGO brick used to explain molecular construction rules.

7.6 Molecular LEGO

7.6.1 Introduction

LEGO bricks are well-known general-purpose construction blocks with well-defined properties. Technically, they might be described as rectangular cuboids or right rectangular prisms. The standard 2×4 LEGO brick (Figure 77) has the following properties:

(a) Bonds on top and bottom (studs on top, holes on bottom)

(b) Does not bond on its front, back or sides (no studs or holes)

Such LEGO bricks can form a wide range of three-dimensional objects, particularly including walls, ceilings, and floors. The walls can have openings (windows or ports). Sealed enclosures (rooms, closets, boxes) can readily be made from LEGO bricks.

If we could design molecular versions of LEGO bricks, we might be able to use them for the SRU shell or other subsystems of the SRU.

7.6.2 Terminology

LEGO bricks are usually referred to as "bricks". As our molecular equivalent is a class of AP Blocks, we will refer to the molecular version as blocks, molecular LEGO AP Blocks, AP Blocks, molecular LEGO Blocks, or some such term. We beg indulgence from the LEGO community for this deviation from LEGO orthodoxy in our terminology in the molecular realm, as there is a long-standing tradition in the molecular realm of talking about molecular building "blocks", rather than "bricks".

7.6.3 The various ways in which LEGO bricks can be connected

Two standard 2×4 LEGO bricks can be stacked directly on top of each other, in which case 8 studs on one brick are inserted into 8 holes on another brick. The two bricks are then connected or bonded together. More commonly, two 2×4 LEGO bricks can be bonded together by inserting four studs on the end of one brick into 4 holes on the end of another brick. We will not consider other types of bonding, such as those involving only one or two studs. When two LEGO bricks are bonded using four studs, either they form a straight bond, or they form a right-angle bond. A standard LEGO brick can bond both ways, either a straight bond or a right-angle bond.

It will prove convenient to have two types of molecular LEGO Blocks, which we arbitrarily call A and B. A blocks can form straight bonds with other A blocks but cannot form right-angle bonds

with other A blocks. Likewise, B blocks can form straight bonds with other B blocks but cannot form right-angle bonds with other B blocks. A blocks can form right-angle bonds with B blocks, and B blocks can form right-angle bonds with A blocks.

These rules are illustrated pictorially in Figures 78 and 79. Type A (gray) and Type B (blue) blocks ***cannot*** form right-angle bonds with blocks of their own type but ***can*** form right-angle bonds with each other (Figure 78).

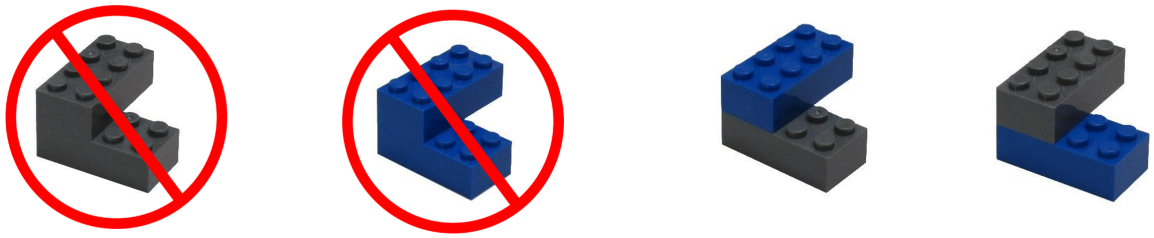


Figure 78: Allowed and prohibited right-angle connections of Type A and Type B blocks.

A (gray) or B (blue) blocks ***can*** form straight bonds with themselves. This allows us to form molecular struts (long runs of straight bonds of molecular LEGO blocks of one type), which can be made from either A (gray) or B (blue) molecular LEGO blocks (Figure 79). Each strut is a beam two bricks high, with the length needed to span the face being constructed.



Figure 79: Straight connections forming struts of Type A or Type B blocks.

We adopt these rules because they will prove convenient when we start to design our molecular LEGO Blocks. As these rules are fairly easy to visualize with LEGO bricks, we continue to describe them in terms of LEGO bricks before we turn to describing a molecular implementation.

7.6.4 LEGO walls

Struts of one type can be stacked on top of each other to form a wall (Figure 80).

Two walls can be connected at right angles (Figure 81).

These walls can also form an enclosure (Figure 82).

7.6.5 LEGO floors and ceilings

Each top or bottom face consists of two crossed layers of struts. Each strut is two bricks high, so the completed face is four bricks thick. Crossing two struts provides strength in both X and Y directions (Figure 83).

Crossing many struts in both directions creates an assembly that is four bricks high (Figure 84).

A solid assembly four bricks high, made of many crossed struts, will be relatively strong. Looking down any single 2×2 column in such an assembly will show that the four 2×4 bricks in that column



Figure 80: LEGO wall made from stacked struts.

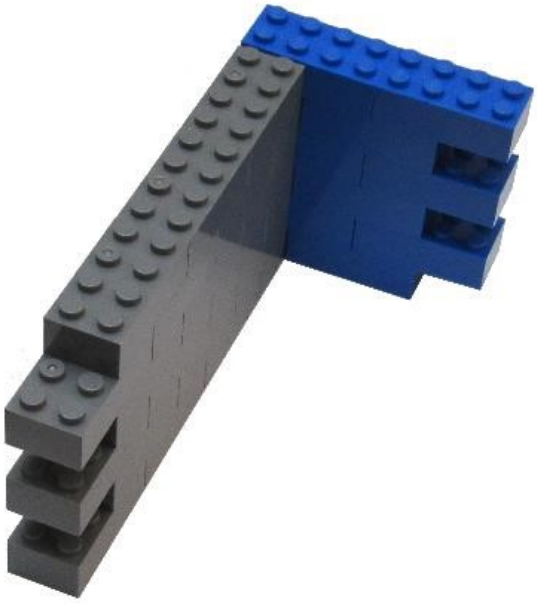


Figure 81: Two LEGO walls joined at a right angle.

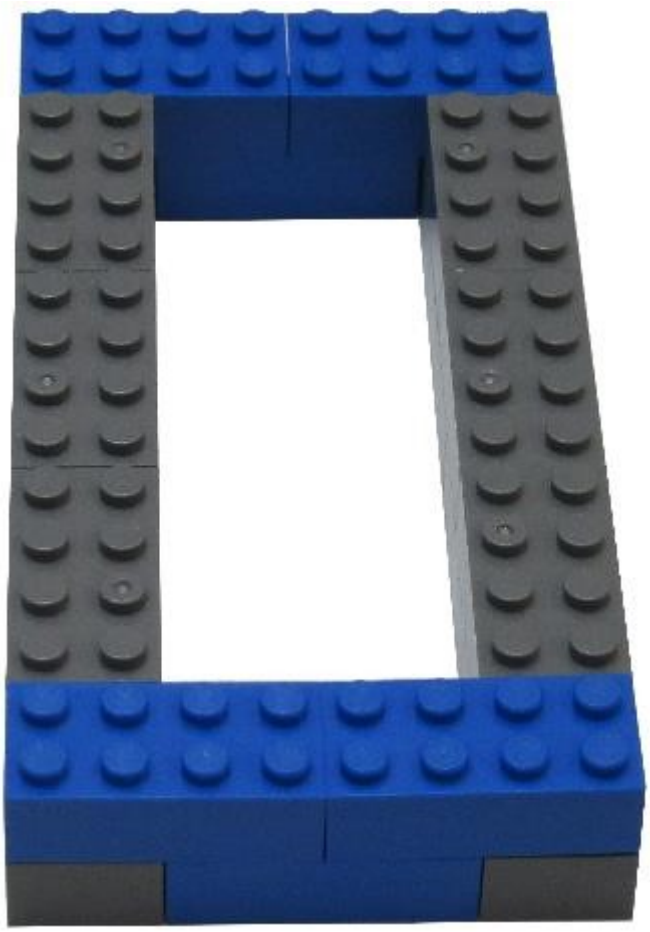


Figure 82: LEGO walls forming an enclosure.

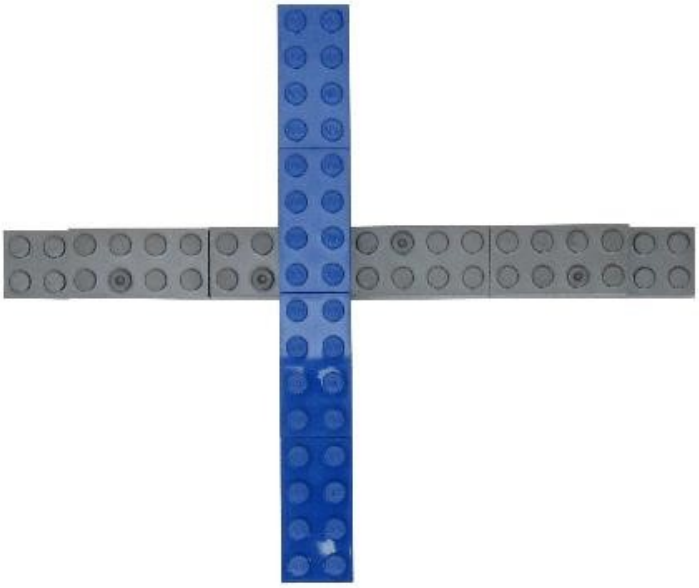


Figure 83: LEGO struts arranged for a floor or ceiling.

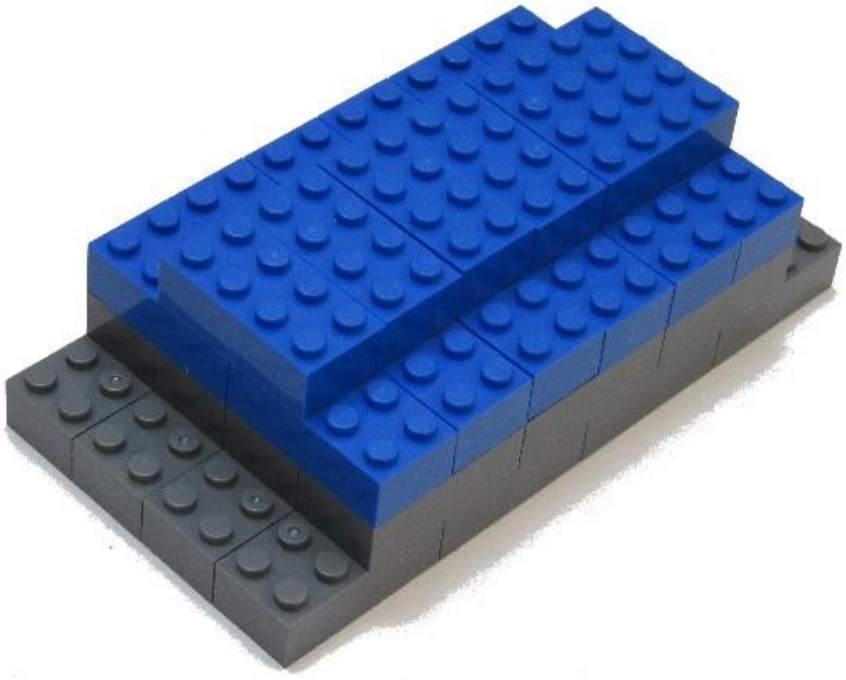


Figure 84: LEGO model of a floor or ceiling assembly.

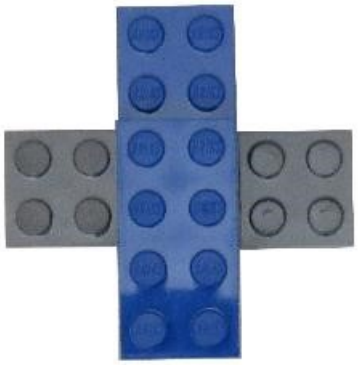


Figure 85: Additional view of a LEGO floor or ceiling assembly.

connect that column to four adjacent columns. This gives good strength to such an assembly (Figure 85).

We have shown that 2×4 LEGO bricks are sufficient to let us build walls, enclosures, and even closed rooms with both walls and ceilings, and further that we can restrict ourselves to using A and B blocks and still achieve the same result.

7.6.6 Crystal Structure

The crystal surfaces of diamond blocks can be defined by Miller indices. The families {100}, {110}, and {111} describe the three principal surface families of diamond.[7] The Miller indices $(11\bar{2})$ describe a well-known surface. For cubic diamond, two planes are perpendicular if the dot product of their normal vectors is zero.[8] Using the Miller indices as the components of these normals, this condition is $h_1h_2 + k_1k_2 + l_1l_2 = 0$.

Here are three sets of three mutually perpendicular planes in cubic diamond, specified by their Miller indices. You can verify that the planes in each set are mutually perpendicular by applying this test to every pair:

(100), (010), (001)

$(\bar{1}10)$, (001)

(111), $(\bar{1}10)$, $(11\bar{2})$

We can pick one of these three sets of orthogonal surfaces and use it as the basis for defining a molecular LEGO block by assigning the Top/Bottom face of the block to one Miller index in the set, the Front/Back face of the block to a second Miller index in the set, and the Left/Right face of the block to a third Miller index in the set.

Then we can pick a specific surface structure for each face.

For the Top/Bottom, we might assign B/N atom types to the Top, and complementary B/N atom types to the Bottom, thus ensuring that we can bond the bottom of the block to the top of an identical block below it. While this would not be in our baseline, it might be used as a risk mitigation or opportunity.

The Front/Back and Left/Right faces need only be terminated in some stable fashion, although we might add the additional constraint that the crack between two adjacent blocks should prevent or slow the diffusion of individual molecules of a gas or liquid through the crack.

A variety of surface terminations are possible, including the alternatives below. Each requires a suitable block orientation and compatible assignment of the other faces; these are not simultaneous face assignments for a single rectangular block:

1) Terminate Top/Bottom with Si or Ge radicals to form weak bonds with facing blocks.

2) Terminate Front/Back (111) surfaces with nitrogen to facilitate sliding over facing (111) surfaces, or to minimize leakage.

[7] We use parentheses for a specific plane, such as (100), braces for a family of equivalent planes, such as {100}, and square brackets for a direction, such as [100]. An overbar negates an individual index. The {100} family includes (100), (010), (001), $(\bar{1}00)$, $(0\bar{1}0)$, and $(00\bar{1})$. For the cubic diamond lattice, the plane (hkl) is perpendicular to the direction $[hkl]$, so its normal has Cartesian components proportional to (h, k, l) when the coordinate axes follow the cubic crystal axes. For example, $(\bar{1}10)$ has a normal vector $(-1.0, 1.0, 0.0)$. This correspondence does not hold in general for noncubic crystals such as lonsdaleite: plane normals are defined by reciprocal-lattice vectors, whereas crystallographic directions are defined by direct-lattice vectors.

[8] For two vectors $\mathbf{a} = (a_1, a_2, a_3)$ and $\mathbf{b} = (b_1, b_2, b_3)$, their dot product is $\mathbf{a} \cdot \mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3$.

3) Terminate Left/Right (100) surfaces with properly oriented oxygen to minimize leakage.

4) Terminate Left/Right (110) surfaces with B/N to assist in bonding and strengthening a wall.

Note that we are considering possibilities outside our baseline in this discussion.
Molecular LEGO blocks might prove useful for, among other things:

1) The shell of an SRU

2) The lining of an extrusion sleeve

3) Curved molecular LEGO blocks for a 40-tooth gear

4) Curved molecular LEGO blocks for a threaded rod

5) Curved molecular LEGO blocks for a threaded sleeve

7.6.7 Molecular LEGO blocks using interfacial Ge

The molecular LEGO block in Figure 86 uses only carbon, hydrogen, and germanium. See also Simple AP Blocks using Ge (Section 7.13). Assembly takes advantage of the fact that opposing Ge radicals will bond to each other with no barrier.

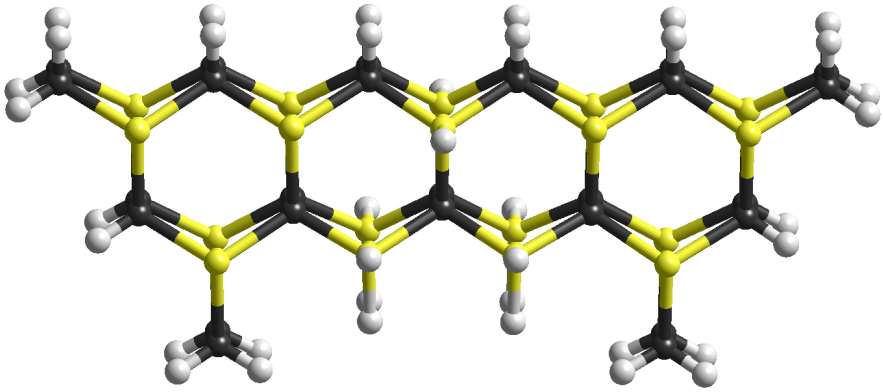


Figure 86: Top view of a molecular LEGO block with germanium radical bonding sites.

Figure 86 shows a top view of a molecular LEGO block. Note the three Ge radicals on the left and the three Ge radicals on the right. An additional six radicals are on the opposite side and are partially obscured in this view. It has 82 atoms.

Figure 87 shows a side view of the same molecular LEGO block, showing all four assembly points on the block, each assembly point consisting of three Ge radicals.

Figure 88 shows a side view of four blocks bonded together using the Ge radicals.

We show only the "A" blocks and leave the design of the "B" blocks as an exercise for the reader.

7.6.8 An example: an SRU shell from LEGO bricks

An SRU shell could be made from four walls (front, back, left side, and right side) and two assemblies for the top and bottom. This would be sufficient to form a rectangular cuboid or right rectangular prism. Given properly designed AP Blocks, such a 3D enclosure should be vacuum tight.

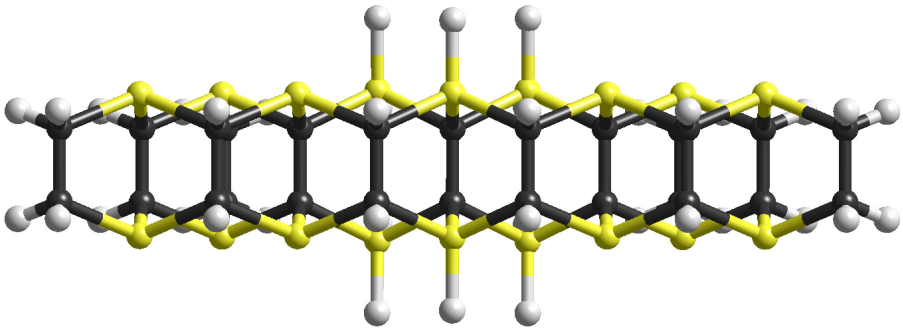


Figure 87: Side view of a molecular LEGO block with germanium radical bonding sites.

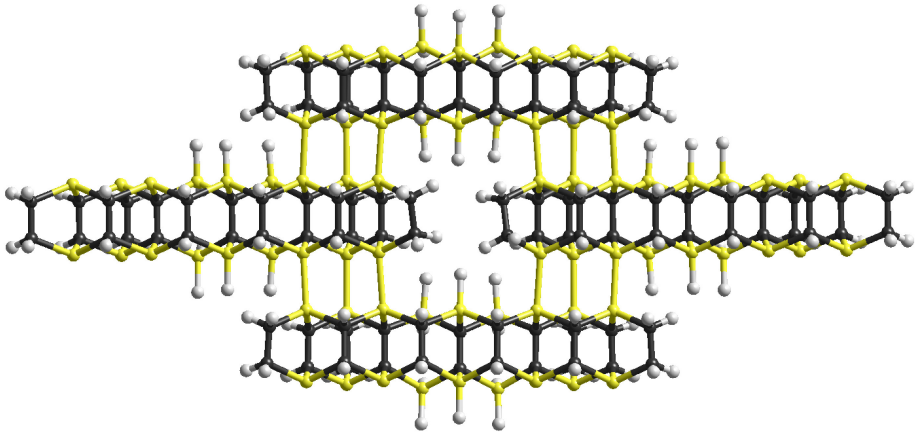


Figure 88: Side view of four molecular LEGO blocks bonded through germanium radicals.

7.6.9 SRU shell from Type A and Type B LEGO bricks

In this section, we show a design approach for an SRU shell illustrated by a literal LEGO brick design. This design is not to scale, in the sense that it has many fewer LEGO bricks than a corresponding design for an actual SRU shell would have molecular LEGO blocks. However, this toy design shows all the needed types of LEGO bricks that would be required, even if there are orders of magnitude too few of them (Figure 89).

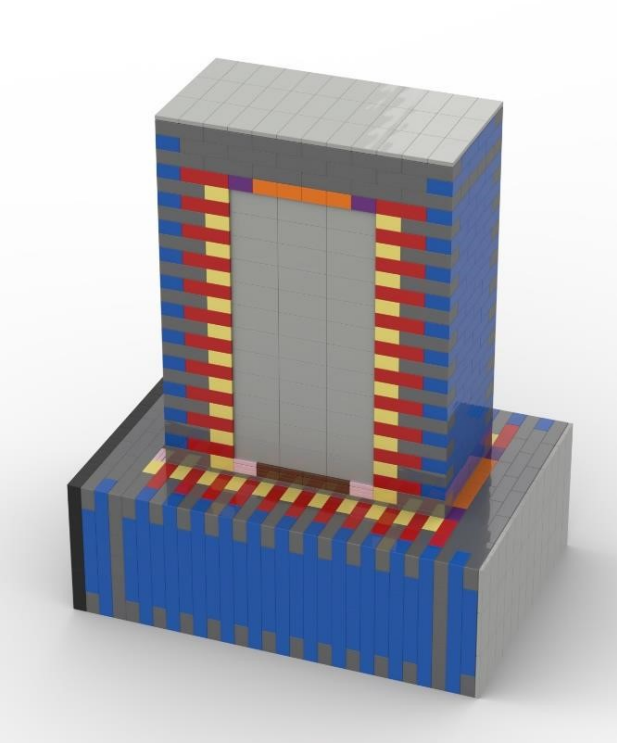


Figure 89: LEGO model of an SRU shell using Type A and Type B blocks.

Type A: Dark gray: front face (the extrusion-port side) and back face; red: extrusion-port lining; green: extrusion-port lining, not visible.

Type B: Blue: left and right sides; light gray: top, visible in the extrusion port; black: bottom; yellow: extrusion-port lining; orange: lining near the top; brown: lining near the bottom; purple: lining at the top edges; pink: lining at the bottom edges.

7.6.10 Justification for the use of A and B blocks

Now that we see the result, a small version of an SRU shell made from LEGO bricks, we can also see why we adopted A and B blocks in the actual molecular LEGO blocks. The walls made from molecular LEGO blocks must be vacuum tight. At multiple points in the shell, blocks that meet at right angles must seal tightly against each other along their lengths. To do this, their width must be half their length, or they will not seal. Designing atomically precise blocks that obey this constraint is both difficult and unnecessary. The A and B blocks can more easily be designed to seal along their lengths by separately designing the details of each block.

7.6.11 Extending the standard 2 by 4 LEGO brick

In a build where every LEGO brick has studs on the top and holes on the bottom, every brick in the entire structure will be oriented with the studs pointing "up" (in the same direction) and the holes pointing "down". While in-plane rotations of 90°, 180°, and 270° are allowed, no other rotations (with the simple rules we've adopted) are allowed.

Another point of interest is that we adopted the constraint that our molecular LEGO blocks bond only on the top and bottom, like the usual LEGO bricks. However, the top, bottom, front, and back of the A blocks are {110} diamond surfaces, so there is no reason we can't make these four sides all bond, making a more flexible block that can more easily form a wider range of more stable structures.[9] Such a brick could be stacked both vertically and horizontally. In LEGO parlance, this is called a SNOT (Studs Not On Top)[10] brick. An example of a SNOT 2×4 brick with side studs is shown in Figure 90.

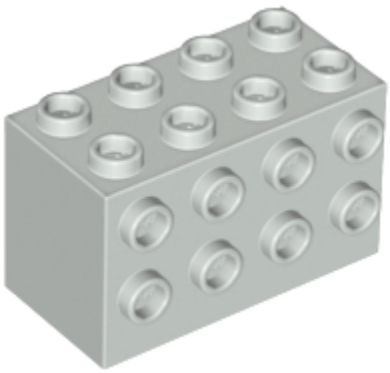


Figure 90: LEGO brick with side studs for connections in multiple orientations.

Once we allow SNOT bricks, a LEGO structure can have bricks that are oriented upside down and sideways. That is, multiple rotations of 90° are allowed around X, Y and Z axes for individual bricks in the build.

In our molecular LEGO blocks, however, rotations continue to be disallowed. Our A and B blocks can only be translated and cannot be rotated. Despite this, we can create builds using molecular LEGO blocks like LEGO using SNOT bricks. Beyond our A and B blocks, we might wish to introduce C, D, or even more block types, given our constraint that blocks can only be translated.

7.6.12 Questions of interest

How stiff is an SRU made of the proposed molecular LEGO Blocks? Is this enough?

If not, how might the SRU be made stiffer?

Can we gain increased stiffness without increasing atom count?

How might we evaluate the sliding seals?

If a design for an SRU shell has some leakage during the replication cycle, what measures could be taken to

[9] The B block is somewhat different, as the front and back belong to {100} and the left and right sides to {110}. However, the same general principles apply. Four sides can be used to bond to other blocks, not just two.

[10] Although, technically, some SNOT bricks have holes (rather than studs) not on top.

(a) Reduce the amount of leakage or, alternatively,

(b) Deal with any leakage that does occur?

Could the proposed molecular LEGO Blocks be made smaller? What advantages might this offer?

What other LEGO-based Blocks might be feasible?

What other four-high blocks have the property that every column is connected to every adjacent column?

What advantages might be gained by going outside the LEGO-inspired design framework? Adding exotically shaped bricks? Bonding on the side(s) as well as top and bottom? Other significant components of the SRU need to be broken down into building blocks. How far can these methods be generalized, and what are their limitations?

7.7 Exemplar BN Molecular LEGO Blocks

Figure 91 shows an example of a Type A molecular LEGO block (see The various ways in which LEGO bricks can be connected (Section 7.6.3)). It is made of carbon, hydrogen, boron and nitrogen. The element count is C: 432, B: 68, N: 68, H: 168, for a total of 736 atoms.

It is boron–nitrogen-terminated on the top (110) and bottom $(\bar{1}\bar{1}0)$.

It is H-terminated on the front $(\bar{1}10)$ and back $(1\bar{1}0)$.

It is H-terminated on the left (001) and right $(00\bar{1})$ sides.

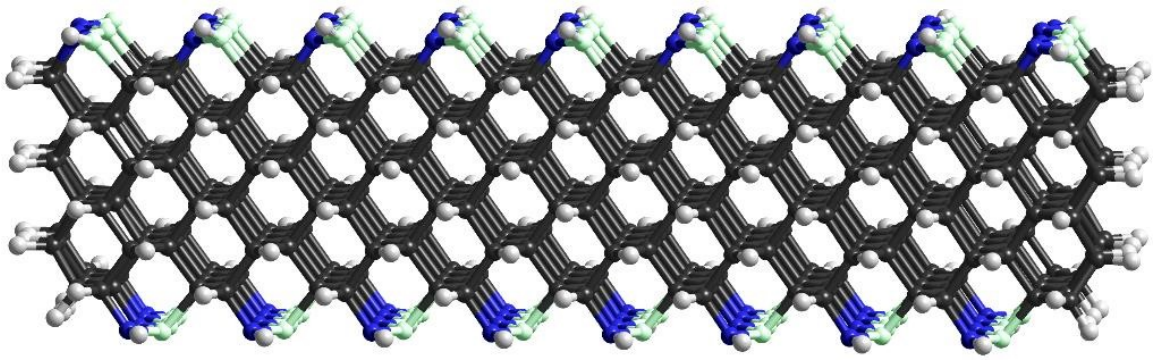


Figure 91: Exemplar Type A molecular BN LEGO block.

Five Type A blocks can be used to form part of a wall (Figure 92). Such a wall should substantially slow if not stop diffusion of most if not all molecular species through the cracks between blocks. The horizontal center-to-center block pitch is 3.62 nm. Successive staggered rows are separated vertically by 1.02 nm; corresponding block centers in alternate rows are therefore separated vertically by 2.04 nm.

The area per Type A block in the plane of the wall is 3.69 nm^2 (horizontal pitch times row spacing), giving an areal atom count of approximately 200 atoms/nm^2.

Figure 93 shows an example of a Type B molecular LEGO block. It is made of carbon, hydrogen, boron and nitrogen. The element count is C: 432, B: 63, N: 63, H: 134, for a total of 692 atoms.

It is boron–nitrogen-terminated on the top (110) and bottom $(\bar{1}\bar{1}0)$.

It is H-terminated on the front (001) and back $(00\bar{1})$.

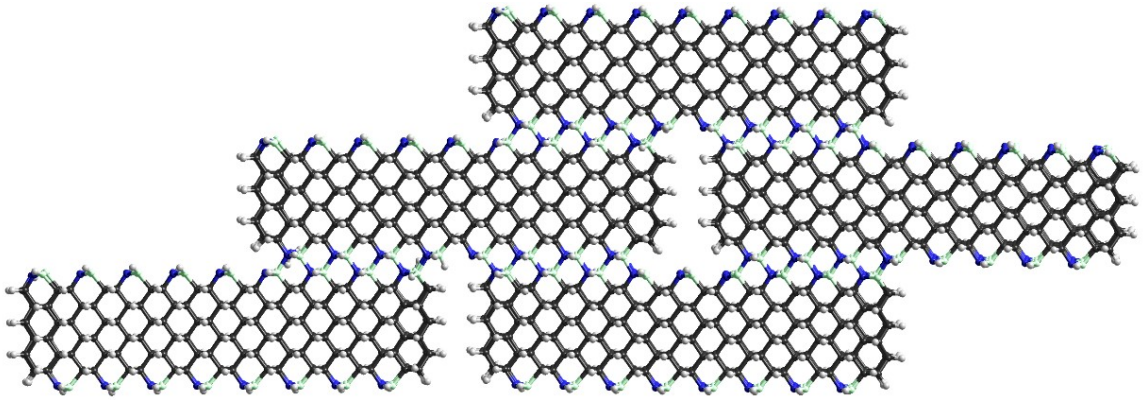


Figure 92: Wall segment assembled from Type A molecular BN LEGO blocks.

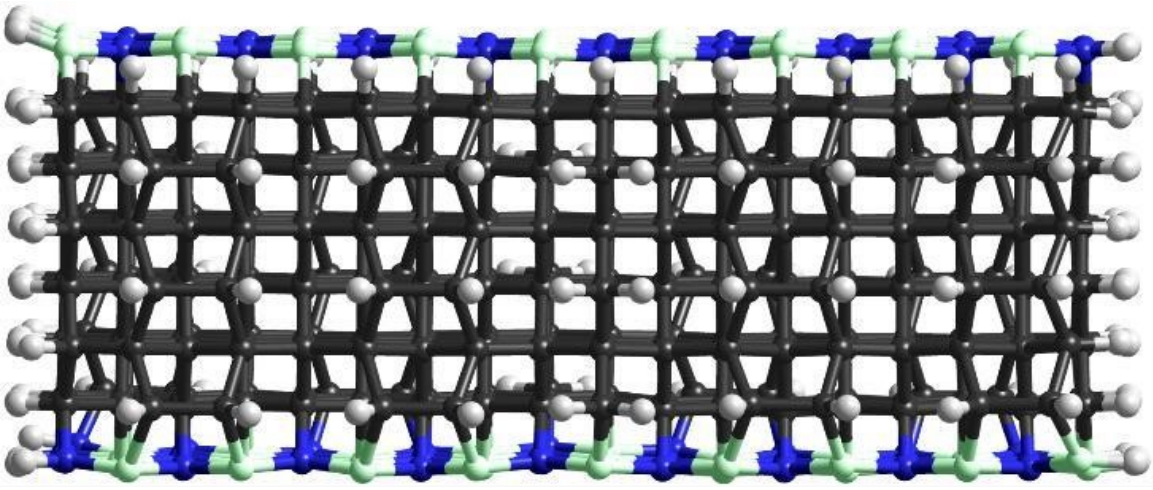


Figure 93: Exemplar Type B molecular BN LEGO block.

It is H-terminated on the left $(1\bar{1}0)$ and right $(\bar{1}10)$ sides.

The Type B block has three pairs of opposite faces, with mutually perpendicular orientations represented by (110), $(\bar{1}10)$, and (001). Four such blocks can form part of a wall (Figure 94). Again, such a wall should effectively slow or stop diffusion of molecules and even noble-gas atoms through the cracks between blocks.

The vertical block pitch is 1.02 nm and the horizontal block pitch is 2.54 nm. The area per Type B block in the plane of the wall is 2.59 nm^2, giving an areal atom count of approximately 267 atoms/nm^2 of wall face.

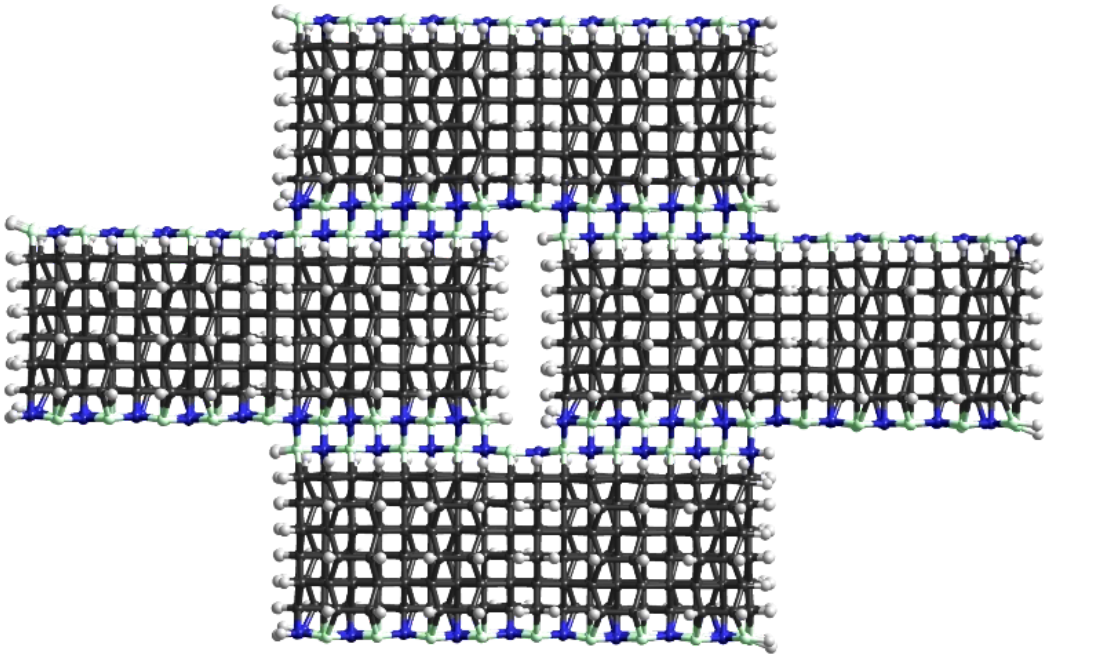


Figure 94: Wall segment assembled from Type B molecular BN LEGO blocks.

Type A (bottom) and Type B (left) blocks can be connected at right angles, as illustrated in Figure 95.

If we tiled the plane with either Type A or Type B blocks in the orientation shown in Figure 95(a), a rectangular repeat region measuring 3.62 nm × 2.54 nm (or the reverse) would contain two blocks. The area per block in this tiling would therefore be approximately 4.60 nm^2, for either Type A or Type B.

7.8 How many molecular BN LEGO Blocks in an SRU shell?

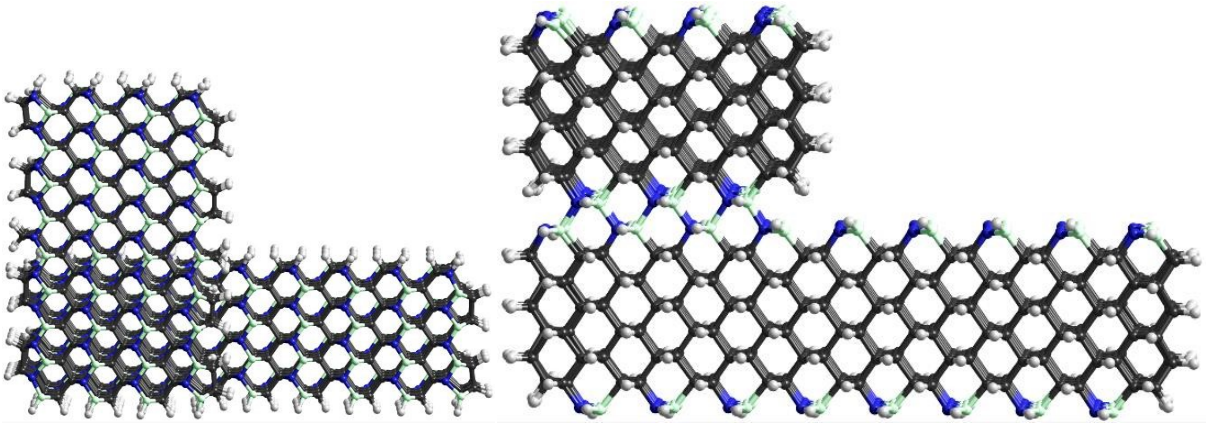
SRU shell dimensions: 500 nm × 100 nm × 60 nm.

A face front area: 3.69 nm^2 (dimensions measured from HyperChem model)

B face front area: 2.59 nm^2

A or B top area: 4.60 nm^2

The four side faces are each one block thick. The top and bottom faces are each four blocks thick, with two layers of Type A blocks and two layers of Type B blocks. Thus, each block type accounts



(a) (b)

Figure 95: Right-angle connection of Type A and Type B molecular BN LEGO blocks: (a) right-angle connection of Type A and Type B molecular BN LEGO blocks. (b) additional view of the Type A and Type B right-angle connection.

for half the combined top-and-bottom block count. Let $N_{A,\text{top+bottom}}$ and $N_{B,\text{top+bottom}}$ denote these counts for each type.

$$\text{SRU A faces:} \quad \frac{2 \times 500 \times 100\,\text{nm}^2}{3.69\,\text{nm}^2} \approx 27{,}100\,\text{A} \tag{7}$$

$$\text{SRU B faces:} \quad \frac{2 \times 500 \times 60\,\text{nm}^2}{2.59\,\text{nm}^2} \approx 23{,}166\,\text{B} \tag{8}$$

$$N_{A,\text{top+bottom}} = N_{B,\text{top+bottom}} = \frac{2 \times 2 \times 60 \times 100\,\text{nm}^2}{4.60\,\text{nm}^2} \approx 5{,}217.4 \tag{9}$$

$$\text{Subtract extrusion port:} \quad -\frac{60 \times 100\,\text{nm}^2}{3.69\,\text{nm}^2} \approx -1{,}626\,\text{A}. \tag{10}$$

In the top-and-bottom calculation, the first factor of two counts the two faces; the second counts the two layers of each block type per face. Rounding each type's count upward gives 5,218 Type A blocks and 5,218 Type B blocks.

Total: 30,692 Type A blocks and 28,384 Type B blocks.

(The top and bottom faces are each four blocks thick; see LEGO floors and ceilings (Section 7.6.5)).

This gives a total of 59,076 molecular LEGO blocks for an SRU shell that is a single layer of blocks thick except on the top and bottom, which are four layers thick. Table 10 summarizes atom counts for Type A and Type B molecular BN LEGO blocks in an SRU shell.

Table 10: Atom counts for Type A and Type B molecular BN LEGO blocks in an SRU shell.

Element	Atoms per A Block	Atoms in all A Blocks	Atoms in all A and B Blocks
Carbon A	432	13,258,944	
Boron A	68	2,087,056	
Nitrogen A	68	2,087,056	
Hydrogen A	168	5,156,256	
All atoms (A)	736	22,589,312	
	Per B Block	**All B Blocks**	
Carbon B	432	12,261,888	
Boron B	63	1,788,192	
Nitrogen B	63	1,788,192	
Hydrogen B	134	3,803,456	
All atoms (B)	692	19,641,728	
All atoms (A + B)			42,231,040

This proposal uses ∼42M atoms in the shell, somewhat less than half of our 100M atom budget.

7.9 BNO Molecular LEGO blocks that bond on four sides

We illustrate a simple AP Block in Figure 96 that can bond on four sides, functioning like a molecular LEGO SNOT block. Notice that the bumps on the left side correspond to hollows on the right side. Borons line up with nitrogens or oxygens. The four sides are (110), $(\bar{1}\bar{1}0)$, $(\bar{1}10)$, and $(1\bar{1}0)$. The top and bottom are (001) and $(00\bar{1})$. In Figure 96, we see a side image of a single AP Block. The surface facing us is the (110) surface. The surfaces facing up and down in this view are (001) and $(00\bar{1})$, and the surfaces facing left and right are $(1\bar{1}0)$ and $(\bar{1}10)$. All surfaces of the {110} family are terminated with boron and nitrogen (or oxygen). Note that boron can also form dipolar bonds with oxygen, as well as nitrogen.

Figure 97 shows two side views of this AP Block. Note that the ridges and hollows on the right side of the AP Block fit into the hollows and ridges on the left side of the AP Block.

The two AP Blocks can be assembled (fused) and will be held together by dipolar bonds between boron atoms and either nitrogen or oxygen atoms, as shown in Figure 98.

It is expected there will be little, if any, barrier to this assembly (Figure 98).

Figure 99 shows a top view of four such molecular LEGO SNOT blocks bonding to each other. This view has been rotated by 90° from the previous view. The surface facing us is the (001) surface.

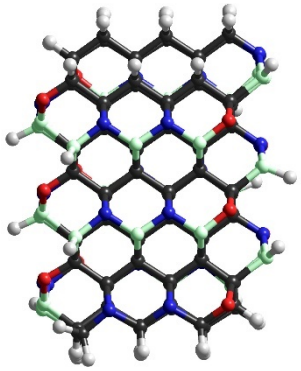


Figure 96: BNO molecular LEGO block with bonding sites on four sides. Drawing repeated from Figure 43. This AP Block has 243 atoms/block and 229 atoms/nm^2 of wall face.

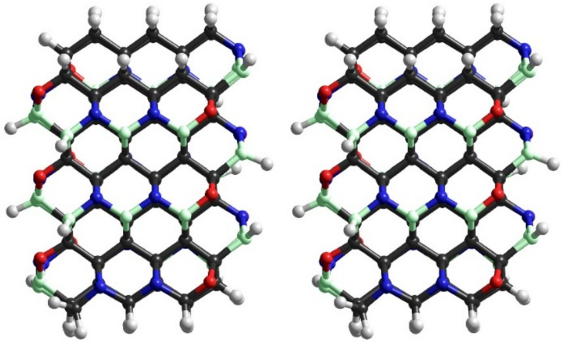


Figure 97: Side views of complementary surfaces of BNO molecular LEGO blocks.

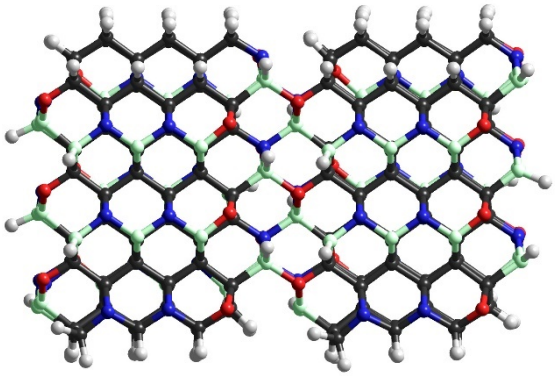


Figure 98: Two BNO molecular LEGO blocks joined by dipolar bonds.

This block can tile the plane of the page both to the left and right and also up and down in this orientation. It should also be gas and liquid tight, as passage of individual molecules through the central passage between the four blocks is energetically implausible at room temperature. For example, the barrier to passage of even H_2 is at least 497 zJ (3.1 eV) according to HyperChem's MM+. Each block in such a tiling occupies a nearly square 1.018 nm × 1.043 nm cell, again according to HyperChem. As there are 243 atoms/block, this means there are 229 atoms/nm^2 of wall face.

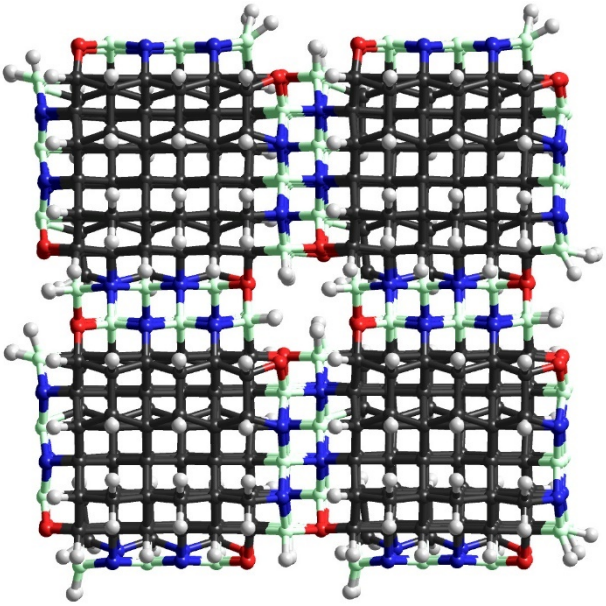


Figure 99: Top view of four joined BNO molecular LEGO blocks.

Had we not used oxygen in this AP Block, the nitrogen atoms at the edges would each have required a hydrogen atom for proper termination. As can be seen in Figure 99, showing four AP Blocks, the region in the center where the four AP Blocks meet would have been crowded with hydrogen atoms, resulting in "hard" contacts. To avoid this, all instances of N-H were replaced with O, freeing up space at the intersections of the AP Blocks and avoiding hard contacts between hydrogen atoms.

7.10 Molecular Tinkertoys

We have drawn inspiration from LEGO bricks, which assemble in a certain fashion into largely solid structures, such as walls. We can also draw separate inspiration from Tinkertoys. Unlike LEGO bricks, Tinkertoys are divided into two types: rods and spools. Spools have holes that can accept rods. If we ask what might be analogous to rods and spools at the molecular scale, we think naturally of nanotubes.

Perhaps the simplest assembly method would be to insert one nanotube into another. In particular, (4,4) nanotubes fit in (9,9) nanotubes.

First, we show three views of a short segment of a 72 atom (4,4) nanotube (Figure 100).

Second, we show three views of a 288 atom (9,9) nanotube (Figure 101).

Finally, we show the (4,4) nanotube nested in the (9,9) nanotube (Figure 102).

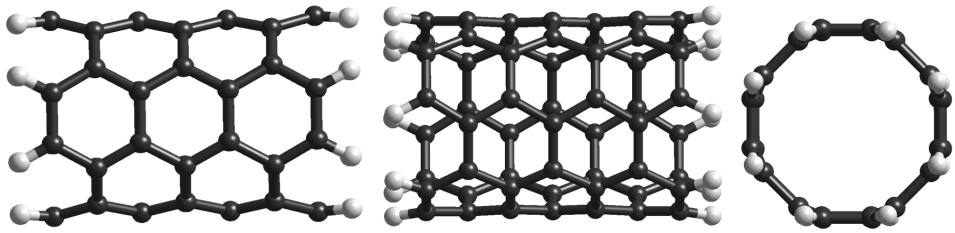


Figure 100: Three views of a short (4,4) nanotube segment.

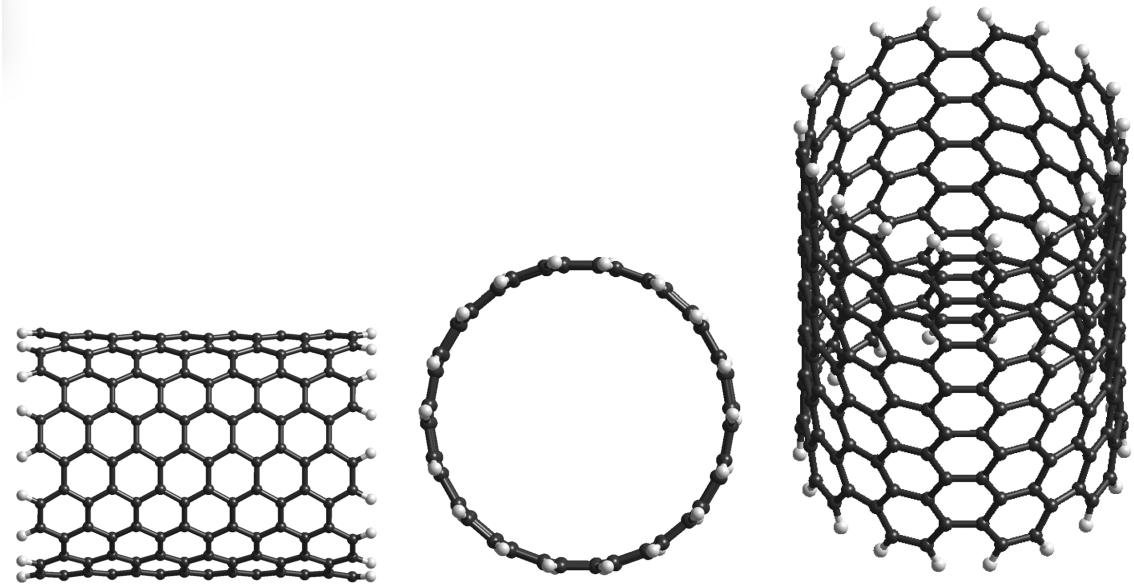


Figure 101: Three views of a (9,9) nanotube segment.

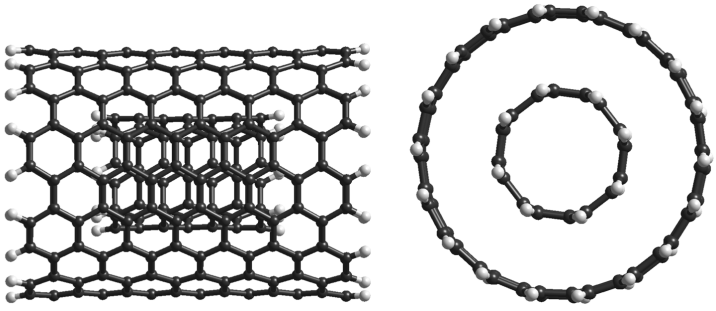


Figure 102: (4,4) nanotube nested inside a (9,9) nanotube.

The energy of the (4,4) nanotube inside the (9,9) nanotube is 660 zJ (95 kcal/mol) lower than when they are separated (Figure 102). If we put the (4,4) nanotube across the lip of the (9,9) nanotube (separate analysis, not shown), the force pulling the (4,4) nanotube into the (9,9) nanotube is 0.73 nN (estimate using HyperChem's MM+).

A 416 atom XYZ six-way block with six (4,4) nanotubes that can link to adjacent blocks in X, Y, and Z is shown in Figure 103. It is composed only of five-, six-, and seven-membered rings. There are no four- or eight-membered rings.

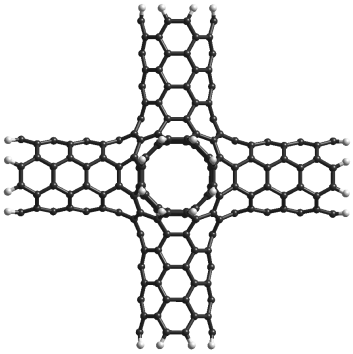


Figure 103: Six-way XYZ nanotube connector.

Figure 104 shows two adjacent XYZ blocks linked by a (9,9) linker.

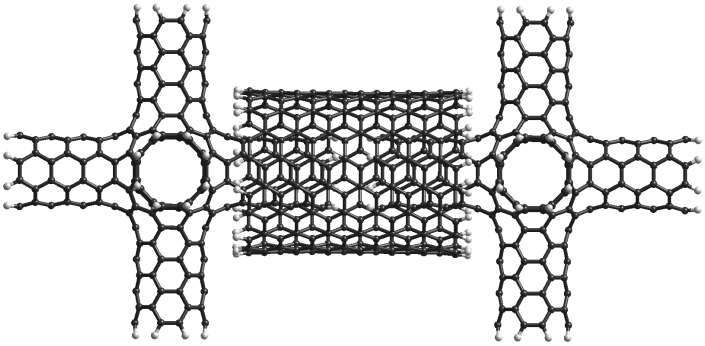


Figure 104: Two XYZ blocks joined by a (9,9) nanotube linker.

Two (4,4) nanotubes can be linked by placing both inside a single segment of (9,9) nanotube. As $W = Fd$, where W is separation work, F is the approximate retaining force, and d is the pullout distance associated with the nanotube overlap, we can increase the binding energy of a single block-block linkage by increasing the length of the (4,4) nanotube segments and the length of the (9,9) linker. If the approximate force is 0.73 nN, then 0.73 nN $\times$ 1 nm = 0.73×10^{-18} J. Boltzmann's constant times temperature is $k_B T = (1.380649 \times 10^{-23}\,\mathrm{J\,K^{-1}})(210\,\mathrm{K}) \approx 2.9 \times 10^{-21}\,\mathrm{J}$. Dividing 0.73×10^{-18} J by 2.9×10^{-21} J gives 252 times thermal noise. That is, a one nanometer overlap between a (4,4) and a (9,9) nanotube takes about 252 times $k_B T$ to break, implying this event will essentially never occur from thermal noise. Even a 0.25 nanometer overlap would be 63 times thermal noise, again implying the link would effectively never be broken by thermal noise.

We conclude that a modest overlap between one (4,4) and one (9,9) nanotube would result in a link effectively immune to thermal noise. Multiple such links, as would occur in a three-dimensional "jungle gym" of 6-way blocks, would be even more stable.

Figure 105 shows a 224 atom planar three-way block with three (4,4) nanotubes. Again, it is composed of only five-, six-, and seven-membered rings.

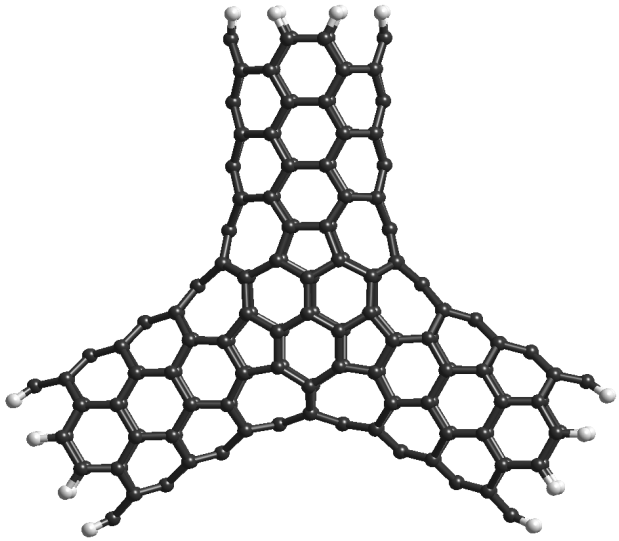


Figure 105: Planar three-way nanotube connector.

Figure 106 shows a 336 atom 4-way block with four (4,4) nanotubes, also composed of only five-, six-, and seven-membered rings.

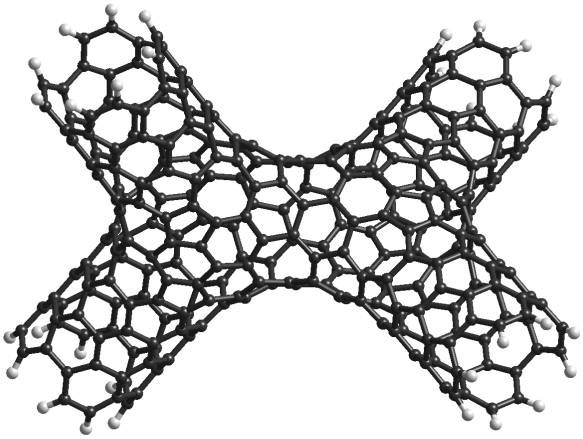


Figure 106: Four-way nanotube connector.

Figure 107 shows an additional view of a six-way XYZ nanotube connector.

We combine six reduced size (164 atom) 3-way blocks to create a 984 atom 6-way planar block (Figure 108).

If we look at the 6-way XYZ block from a different perspective, we can see the three-fold symmetry inherent in a cube (Figure 109).

If we trap the 6-way XYZ block in the planar 6-way block, we get a 984 + 416 = 1,400 atom 12-way assembly which can be used as a node in a space frame (Figure 110).

Figure 111 shows an additional view of the twelve-way space-frame node.

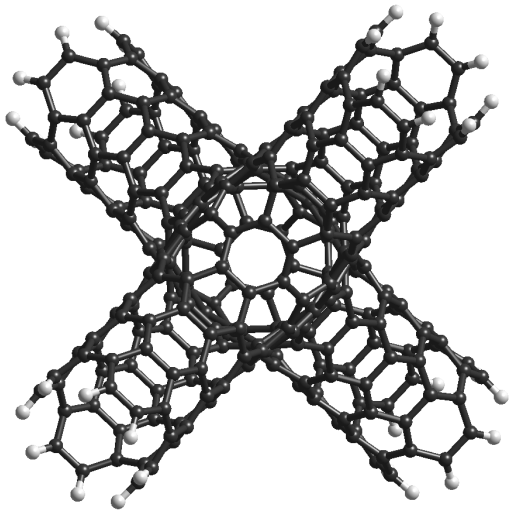


Figure 107: Additional view of a six-way XYZ nanotube connector.

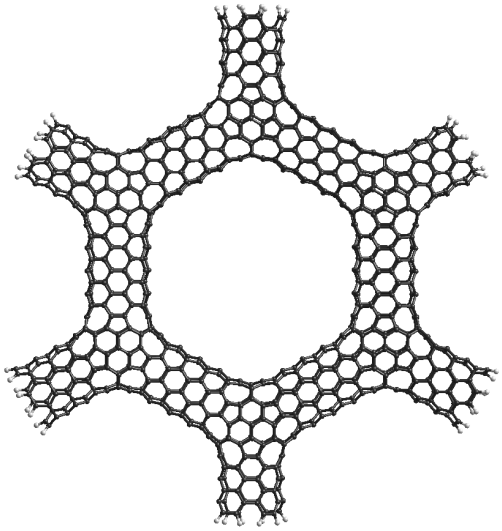


Figure 108: Planar six-way connector assembled from six three-way blocks.

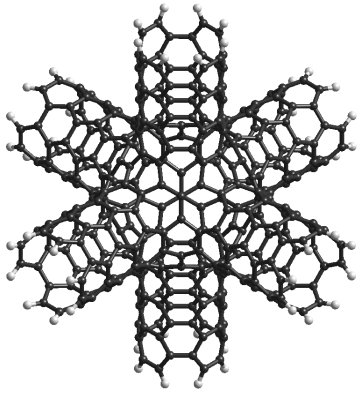


Figure 109: View showing the three-fold symmetry of a six-way XYZ connector.

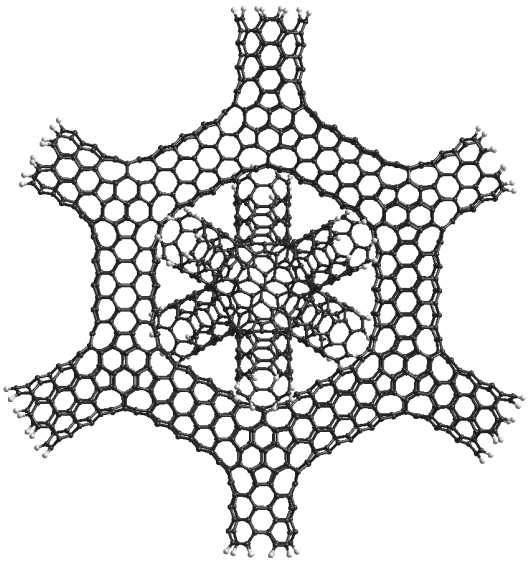


Figure 110: Twelve-way space-frame node combining planar and XYZ connectors.

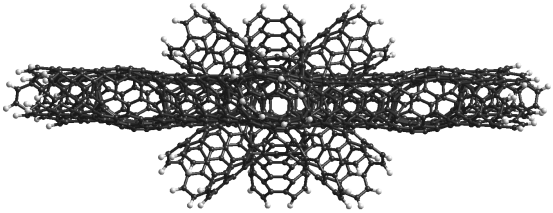


Figure 111: Additional view of the twelve-way space-frame node.

7.10.1 Assembling space frames

A specific assembly sequence for a space frame depends on the order in which the nodes and struts are assembled. While it would be most pleasant if all assembly sequences were valid, this is not always the case. If we wish to add the final node (the final 12-way strut) to the middle of a space frame, there are issues of access. Even if we have access, if linking two nodes requires that we first pull them apart before being able to properly connect them, then we require enough flexibility in what has been built so far to allow us to pull the relevant nodes further apart than they will be in the final structure. And if we wish to repair a space frame by removing a node, routing it out of the space frame through a series of channels, and replacing it with a node brought in through a series of channels, our difficulties increase.

How can we insert a new 12-way block into the middle of an existing space frame? If we assume that a single node in a space frame is missing, even assuming we had access to put a new node into place, we'd then have the problem of wrapping the (9,9) linker around the abutting (4,4) nanotube ends to be linked. This would require the ability to pull the ends apart far enough to slide the (9,9) linker onto both ends. This might or might not be possible, depending on the length of the struts in the space frame and our ability to push them aside during the joining process.

Consider, however, the problem of building an arbitrary space frame de novo.

We could divide the final assembled space frame into two halves along a plane perpendicular to one of the twelve struts coming from a node (a node is just the 12-way block). This would allow us to separate the two halves far enough apart to slip the (9,9) linkers onto the exposed (4,4) nanotubes on one half, and then bring the two halves together in a single assembly operation.

Alternatively, we might design a different linker which could be wrapped around a (4,4) nanotube without having to move it. Such a linker might be U shaped, and could simply be slipped into place. This would sacrifice some of the van der Waals surface area used to keep the butt ends of the two (4,4) nanotubes together, but would simplify assembly.

By combining the 3-way, 4-way, XYZ 6-way, planar 6-way, and 12-way blocks we can build more complex structures. The 12-way block lets us build rigid space-frame lattices.

7.10.2 Other chiralities

Although we constructed the 3-way and other structures from (4,4) nanotubes, the same thing can be done with other chiralities, for example a (6,0) nanotube (Figure 112).

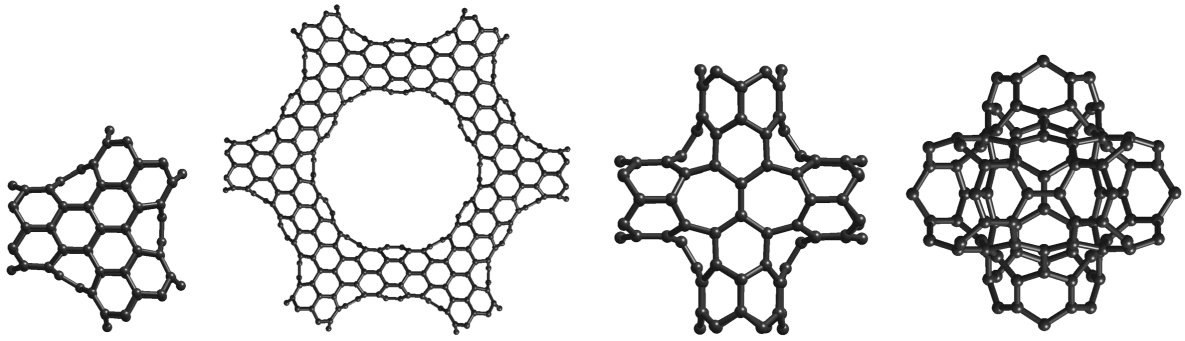


Figure 112: Nanotube building block with (6,0) chirality.

7.11 Fusing AP Blocks in the fabricator

We generally assume that AP Blocks are manufactured in the fabricator and then assembled either by the workstation or the internal assembly arm of the SRU. However, in some cases it might be more convenient to assemble some AP Blocks using the fabricator. In particular, the fusion[11] of two AP Blocks might better be done by the fabricator. While extensive fusion operations of AP Blocks within the fabricator run into size restrictions, a few AP Blocks could be fused into a single assembly. This might be particularly advantageous in the SRU fabricator, as it can hold and position four molecular tools simultaneously during the fusion process, facilitating simultaneous operations.

This raises an interesting point: in convergent assembly (see Convergent assembly and error mitigation (Section 6.7)), an important component of the process is to inspect and reject defective sub-assemblies before proceeding with further assembly operations. By this means, it is possible to manufacture large assemblies with high reliability even when the manufacture of the sub-assemblies is error prone.

In our case, we want to inspect and reject defective AP Blocks within the SRU fabricator prior to fusing them into larger structures. This assumes there is an error rate involved in manufacturing AP Blocks which is large enough that it must be mitigated. We consider a simple example.

We want to make a 6-way planar connector from six 3-way connectors (Figure 113).

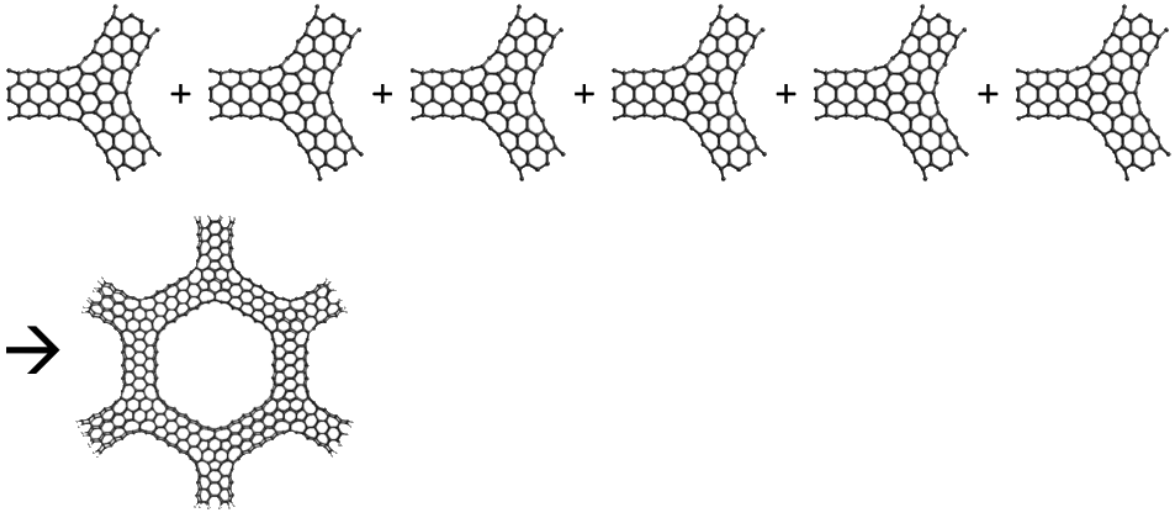


Figure 113: Assembly of a planar six-way connector from six three-way connectors.

However, we want to mitigate the risk that one of the 3-way connectors is defective.

We assume that one arm of a 3-way connector can be fused with one arm of another 3-way connector.

To mitigate the risk that one 3-way connector is defective, and therefore that fabrication of the planar six-way connector fails, the fabricator makes seven 3-way connectors. We assume one is defective for purposes of this example (Figure 114).

[11] Fusion operations involve the formation of covalent bonds, typically by pressing two reactive surfaces together. An example might be the diamond (110) surface. It is stable against reconstruction, and when two such surfaces are pressed together with sufficiently accurate alignment, they should fuse to form a single unbroken diamond crystal structure. The (100) surface has a known reconstruction, and when two such surfaces are pressed together they may form covalent bonds, possibly as four-membered rings or double bonds, without continuing the diamond crystal structure.

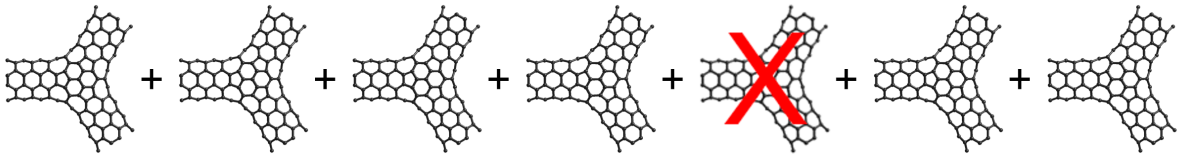


Figure 114: Redundant three-way connectors for assembly despite a defective block.

We are, however, using the broadcast architecture. How do we efficiently deal with this case without a combinatorial explosion of instructions?

First, the SRU fabricator manufactures seven 3-way connectors and places them on seven storage sites, numbered from one through seven. After completing this task, the defective unit is identified. This might be done by testing during manufacture to identify a failed mechanosynthetic step, or by testing after the fact to determine that one of the 3-way connectors is defective. Whenever the testing is conducted, we assume the defective unit is identified and disposed of.

At this point, we have seven storage sites holding six successfully manufactured 3-way connectors and one empty storage site. These seven sites would be physical locations within the SRU fabricator.

Now we broadcast the following sequence of instructions:

If siteNumber 1 is empty, Move 3-way connector at siteNumber 2 to siteNumber 1

If siteNumber 2 is empty, Move 3-way connector at siteNumber 3 to siteNumber 2

If siteNumber 3 is empty, Move 3-way connector at siteNumber 4 to siteNumber 3

If siteNumber 4 is empty, Move 3-way connector at siteNumber 5 to siteNumber 4

If siteNumber 5 is empty, Move 3-way connector at siteNumber 6 to siteNumber 5

If siteNumber 6 is empty, Move 3-way connector at siteNumber 7 to siteNumber 6

(Do loops must be unrolled before they are broadcast. The actual broadcast instructions must be fully decoded before they are transmitted, as each SRU can only execute very simple instructions.)

The number of instructions broadcast to carry out these operations is negligible compared with the much larger number of fully decoded mechanosynthetic instructions that were broadcast to actually manufacture the seven 3-way connectors.

Following completion of this instruction sequence, we now have six 3-way connectors in sites one through six, waiting to be assembled into the 6-way planar connector.

In this example, if two connectors are defective, we cannot make the desired six-way planar connector. This might result in a fail-stopped SRU unless some other redundancy mechanism can deal with this error. Fail-stop is acceptable if the probability of its occurrence is low enough.

This concept generalizes to larger numbers of parts that have more failures. The essential concept is that we manufacture enough more of any particular type of AP Block to reduce the probability of running out of that part to an acceptably low level. Further, this risk mitigation can take place within the SRU fabricator provided the number of buffered parts does not exceed the limits of the storage capacity of the SRU fabricator.

This approach effectively allows the fabricator to make larger assemblies with more atoms than might first appear to be the case. It does require that the assemblies be decomposable into directly manufacturable AP Blocks that can be fused (or otherwise linked) together into a single larger structure.

7.11.1 A question of terminology

Assume we use the method outlined above to make a 4,000 atom block from five 800 atom blocks. The 800 atom blocks should clearly be called "blocks", because (a) they were manufactured in the fabricator, which normally makes blocks; (b) they were manufactured using a single reaction sequence; and (c) they form a single covalently bonded network of atoms. These five blocks satisfy our 1,000 atom limit on block size. However, the 4,000 atom "block" is larger than our 1,000 atom limit, even though it was made in the fabricator. It was not made using a single reaction sequence. Instead, it was assembled from five bona fide blocks. Should we call the assembly of five blocks an "assembly" because that's how it was made, or should we call it a "block" because a more reliable fabricator, able to make it in a single reaction sequence, would legitimately call it a "block", or should we call it a "block" because the fabricator made it, and we're going to call anything the fabricator makes a "block"?

To add to our confusion, the SRU's internal assembly arm can carry out assembly operations, and some assembly operations fuse blocks together into single covalently bonded networks of atoms. If the internal assembly arm of the SRU manufactured our 4,000 atom assembly by assembling five 800 atom blocks, is it a "block" or an "assembly"? Normally, the SRU's internal assembly arm only makes assemblies and never blocks. Do we call this an "assembly" because it was made by the SRU? Or an "assembly" because the operations that made it were assembly operations? Or a "block" because it's a single covalently bonded network of atoms?

As there does not appear to be a single answer that satisfies all our goals, for purposes of this paper a "block" is anything made by a single reaction sequence. If two blocks, each consisting of many atoms, are fused together, the result is called an "assembly" whether the fusion was done by the SRU or the fabricator.

From this perspective, only the fabricator makes blocks, but it also has a limited capability to assemble blocks into assemblies. The SRU, using the internal assembly arm, can only assemble blocks and assemblies of blocks into assemblies.

This answer is not ideal, because we don't know, just by looking at a structure, whether it is a block or an assembly. We only find out if we ask about its history. But it does mean that only fabricators make blocks, even though they can dabble in assembly, and the SRU only assembles, even if some assemblies are single networks of covalently bonded atoms.

7.12 Graphene sheets from coronene

We might use coronene as an AP Block and use it to make larger graphene sheets in the fabricator. By dehydrogenating the edges of two coronene molecules, we can fuse them together into a larger graphene sheet. By fusing this larger sheet with another coronene molecule, we can make a yet larger graphene sheet (Figures 115 and 116).

7.13 Simple AP Blocks using Ge

Figure 117 shows the simple AP Block 1,3,5,7-tetragermatricyclo[3.3.1.1^{3,7}]decane.

By removing the hydrogen atoms from the germanium atoms, we would be left with four germanium radicals that would readily bond to other germanium radicals on other molecules. If the adamantane cage remains stable, the resulting minimal AP Blocks might be used to build a wide range of structures. Whether this approach, in vacuum, would allow physical construction of the obvious

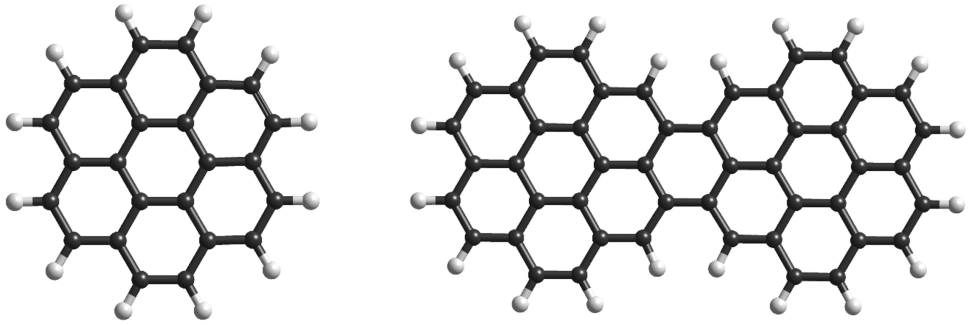


Figure 115: Growth of a graphene sheet by fusing coronene blocks.

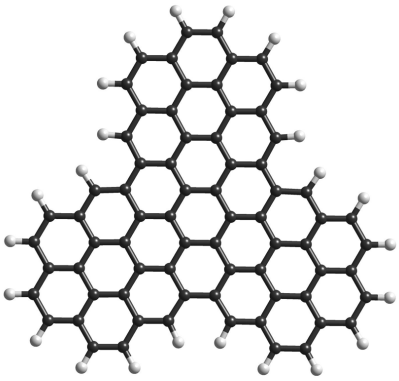


Figure 116: Additional view of a graphene sheet assembled from coronene blocks.

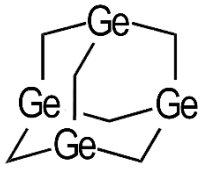


Figure 117: Chemical structure of a germanium-substituted adamantane building block.

set of design possibilities, and whether this set of design possibilities would encompass a workstation fabricator, are both open questions. The primary observation is that remarkably simple AP Blocks open a very wide design space, making it reasonable that modest increases in the AP Block atom count should open a design space large enough to encompass a workstation fabricator, and a somewhat larger increase in the atom count should make it less difficult to actually find a design in that somewhat larger design space. That is, search is easier when there are fewer constraints. A sufficiently small search space will have no solutions. Increasing the search space might admit one or two solutions, but they might require examining most of the search space to find. Further increasing the search space can make it easier to find solutions because there are fewer arbitrary constraints to deal with.

Another candidate set of larger tetrahedral AP Blocks might be assembled by bonding two blocks together by using Ge radicals on their surfaces. These two blocks are both based on a substituted GeC adamantane core. They fit together like the two elements in a zinc blende crystal lattice structure. They are illustrated in Figures 118 and 119.

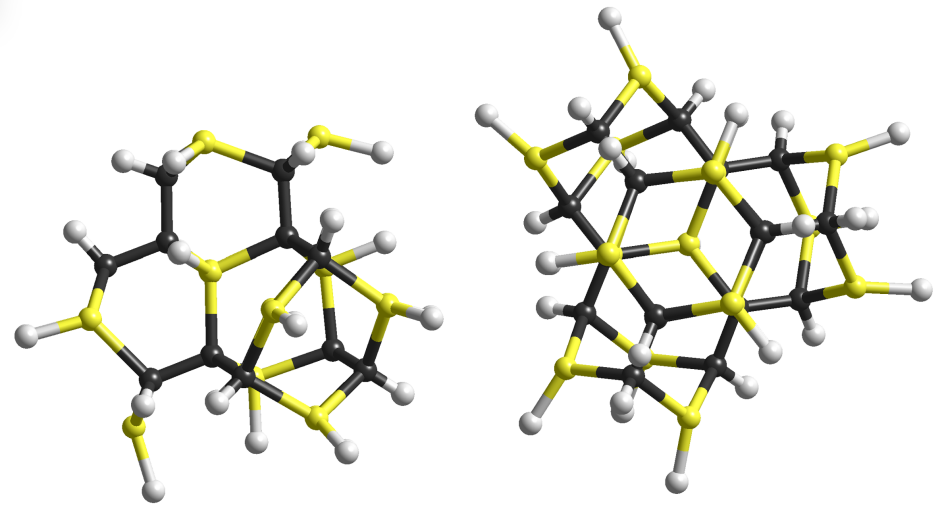


Figure 118: Complementary germanium-carbon adamantane building blocks.

Figure 119 shows an assembly of complementary germanium-carbon adamantane blocks.

Notice that we assemble these blocks by placing three Ge radicals on one block adjacent to three Ge radicals on the other block, at which point radical-radical bonding takes place. As each block has four linkage sites, and each linkage site has three Ge radicals, each block in the bulk is held in place by 12 Ge-Ge bonds. Blocks on the surface would be held in place by three, six, or nine Ge-Ge bonds. This crystal lattice resembles an aerogel, in that there are large holes in the structure with nodes connected by long struts. If a denser structure is desired, one method of increasing the density might be to build several similar interpenetrating crystal frameworks, each displaced from its neighbors so that their atoms do not overlap. This kind of interpenetrating crystal structure has been seen experimentally [Erm88].

It might also be interesting to see what we can build from 1,000 atoms of C, H, and Ge if we are limited to flat (0001) lonsdaleite sheets only four layers thick. These basal planes are analogous to cubic diamond (111) planes, but with different stacking. If the (0001) Ge-up faces are stable against reconstruction (plausible), then two such faces might reversibly bond to each other. While there is no reason to believe that design examples using this class to carry out basic functions required by

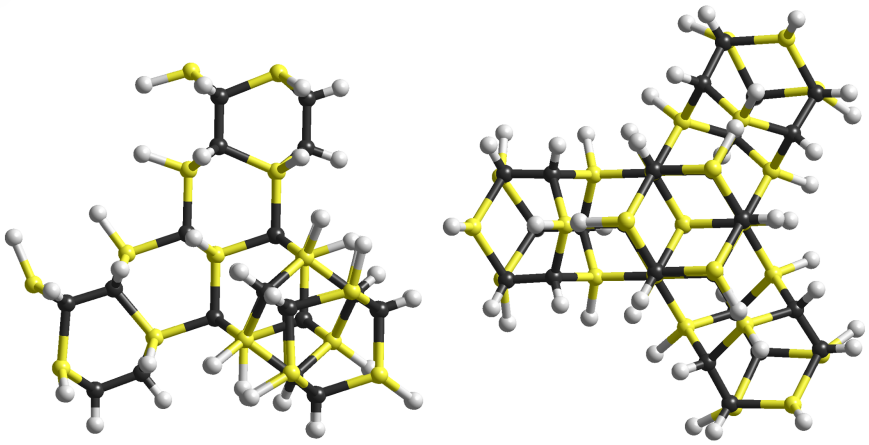


Figure 119: Assembly of complementary germanium-carbon adamantane blocks.

a workstation fabricator would be difficult to find, this has not yet been confirmed.

We show an example (Figure 120).

Figure 121 shows an additional view of a germanium-carbon thin-sheet building block.

Figure 122 shows the bonding surfaces of a germanium-carbon thin-sheet building block.

The germanium atoms on the top and bottom are radicals, each having only three bonds. This allows this AP Block to bond strongly to other blocks on the top and bottom, as illustrated in Figure 123 with two blocks which are in a "quarter overlap" structure. Technically, there is less than one quarter overlap, as the central line of germaniums is left unbonded to provide clearance for another block to fit adjacent to the first, but the name seems appropriate.

Figure 124 shows an additional view of overlapping germanium-carbon blocks.

We could add further blocks, continuing this wall of AP Blocks for as long as we might like and as high as we might like. Figure 125 shows 8 blocks with four in each layer.

Figure 126 shows an additional view of the two-layer germanium-carbon assembly.

We can overlap these AP Blocks in various ways. We might, for example, add four such blocks in a "half overlap" pattern (Figure 127).

It seems clear that this construction pattern of multiple layers of what might technically be called all Ge-up twinned hexagonal germanium-carbide thin-sheet AP Blocks might be useful in a variety of atomically precise construction projects.

Further analysis and research on the illustrated AP Blocks are needed to confirm that they are stable, that the Ge radicals on one block will bond to corresponding Ge radicals on a second block, that the resulting assemblage of blocks is adequately stable, that they can be separated without damage, that they are not prone to unexpected reconstructions, and that we can use them to build useful molecular structures. In particular, whether this particular block can be an effective barrier to hydrogen, as required by the baseline design, has not yet been analyzed.

This class of AP Blocks has the potential to build a wide range of two-dimensional shapes, is thin, can form layers that adhere to each other, and uses only the elements C, H and Ge. Stacking layers should be able to form 3D structures. Variants created by placing germanium adatoms on the (100) and (110) surfaces will provide a variety of design alternatives to investigate, as will modifications to the edge structure to enhance the ability to block the passage of H_2.

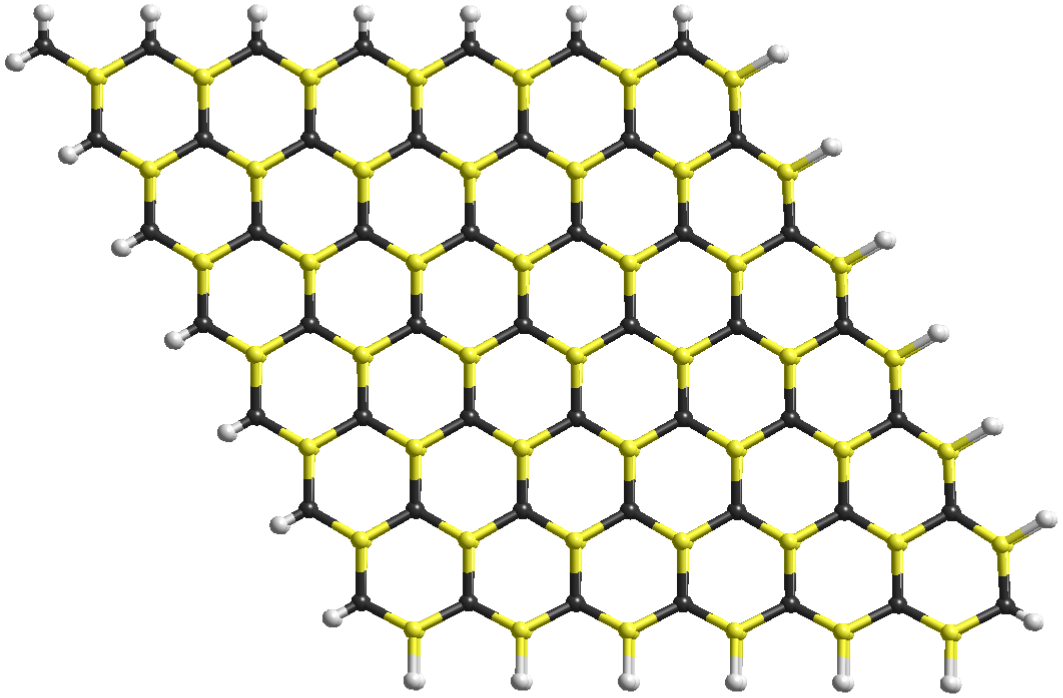


Figure 120: Germanium-carbon thin-sheet building block.

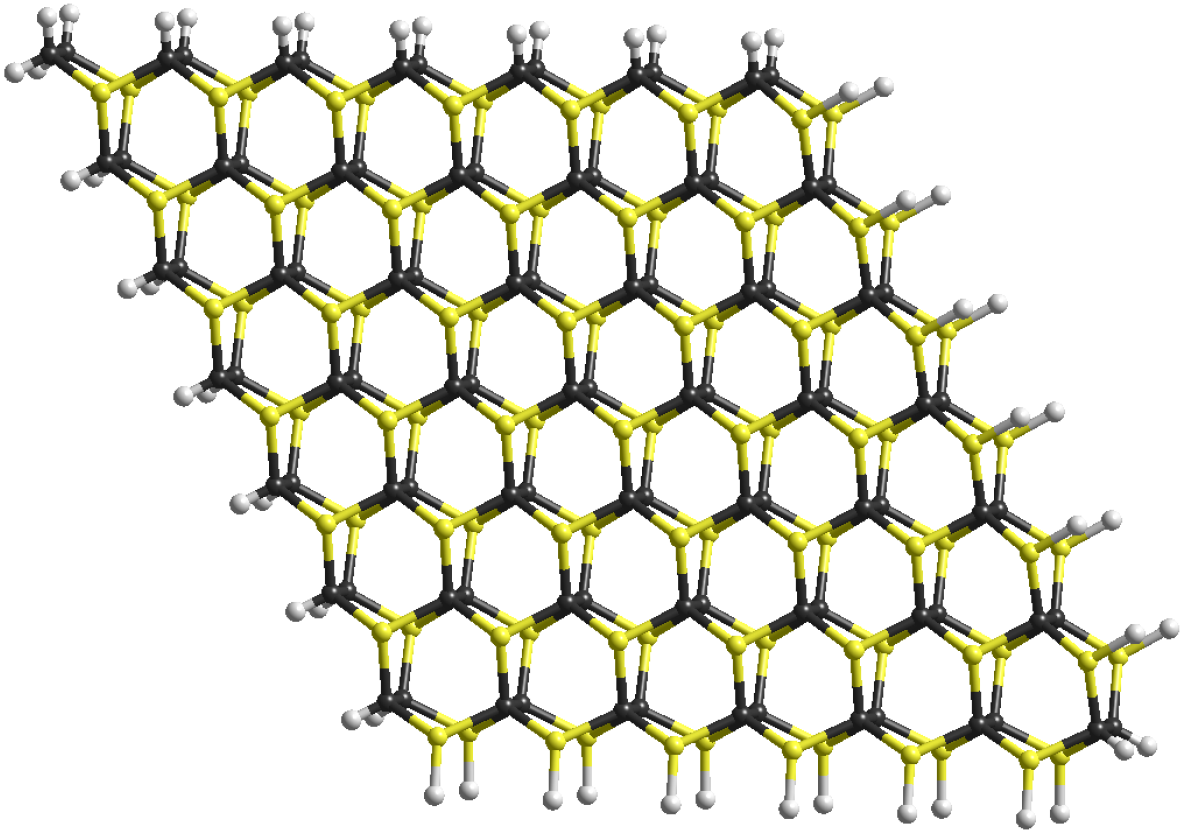


Figure 121: Additional view of a germanium-carbon thin-sheet building block.

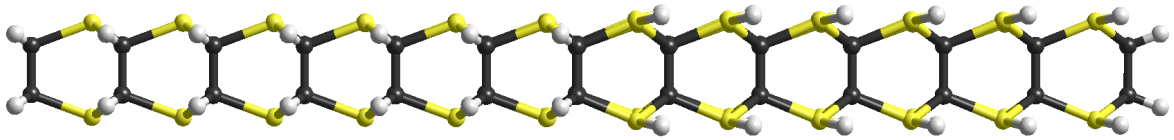


Figure 122: Bonding surfaces of a germanium-carbon thin-sheet building block. This AP Block has a total of 248 atoms: 54 H, 96 Ge, and 98 C.

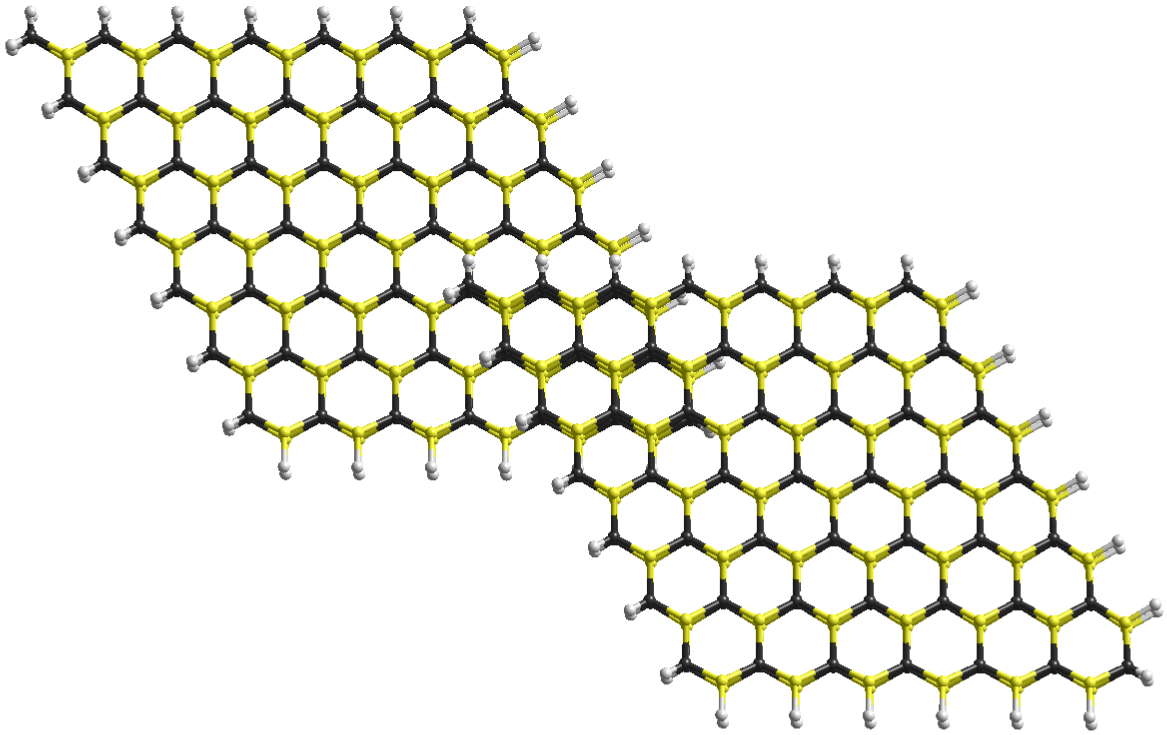


Figure 123: Overlapping germanium-carbon blocks forming a wall.

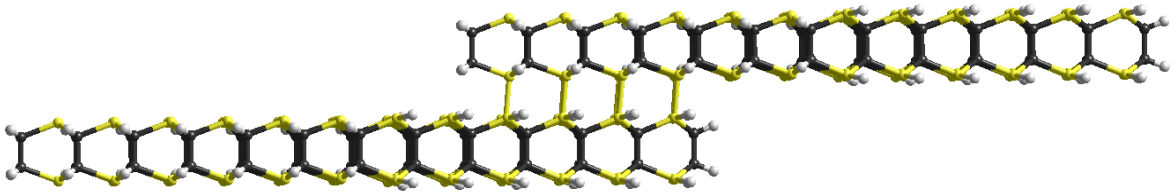


Figure 124: Additional view of overlapping germanium-carbon blocks.

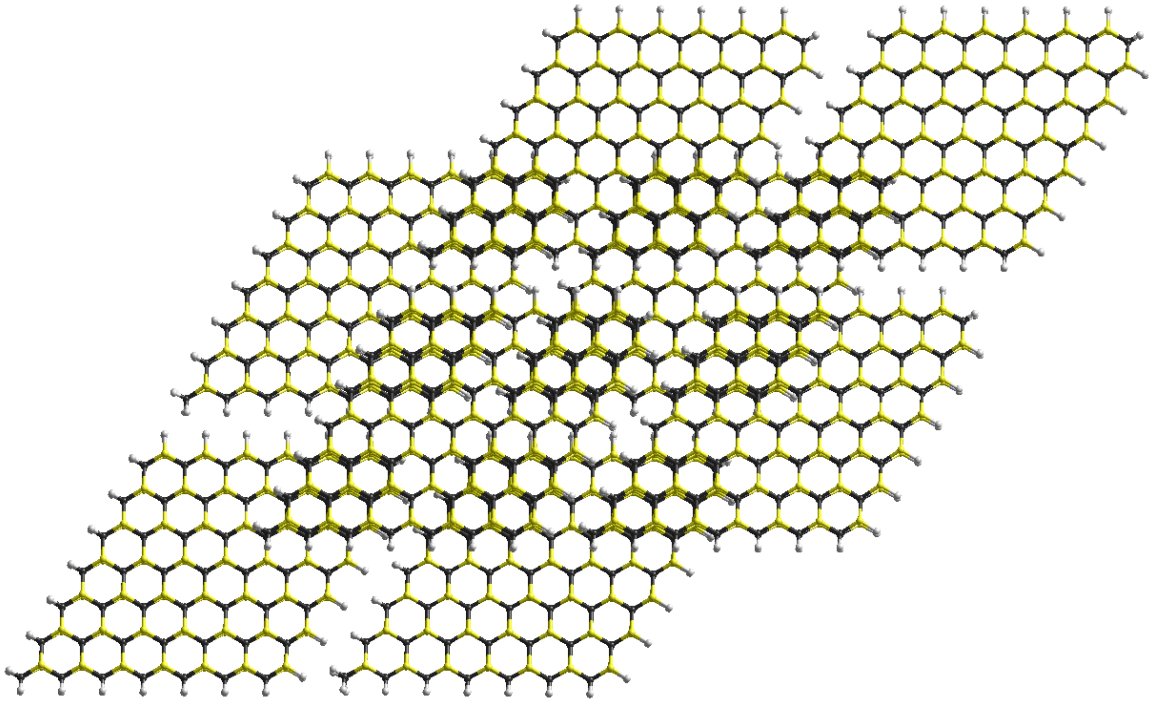


Figure 125: Eight germanium-carbon blocks arranged in two layers.

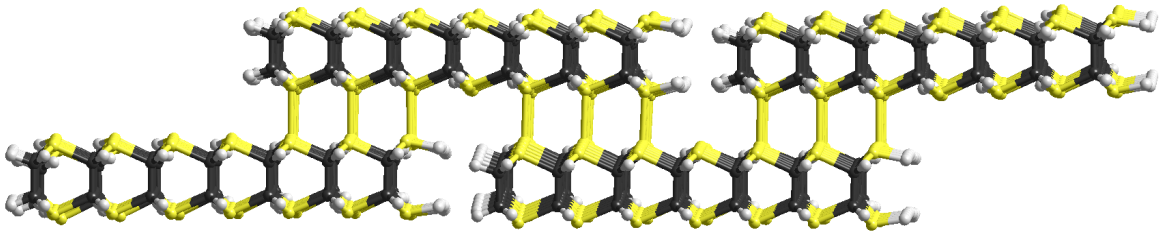


Figure 126: Additional view of the two-layer germanium-carbon assembly.

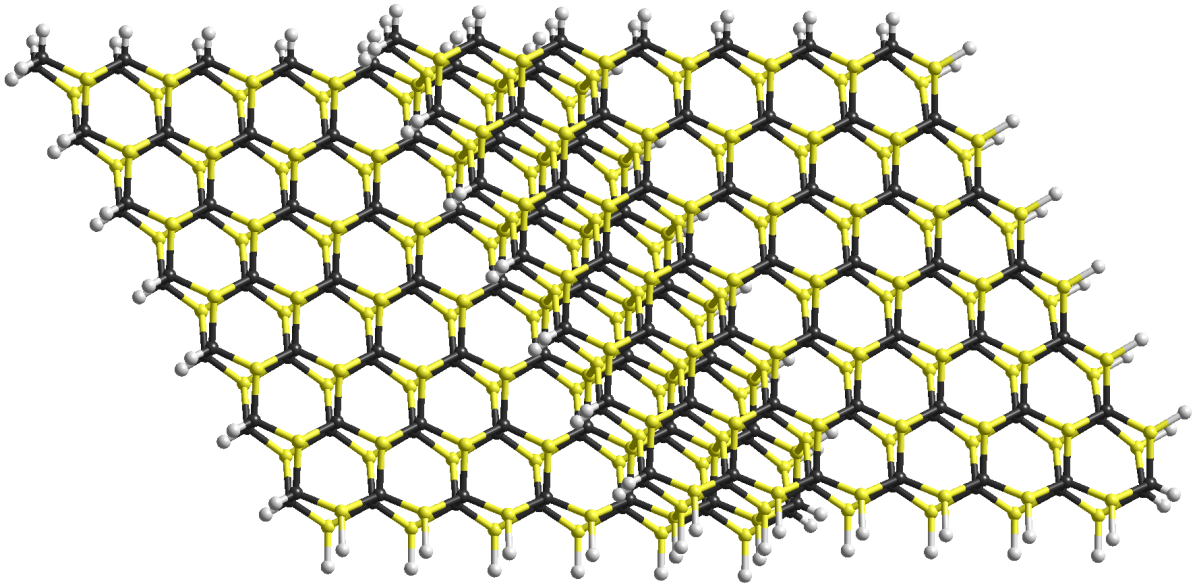


Figure 127: Half-overlap arrangement of germanium-carbon blocks.

7.14 Stewart Platform Struts

In this section we analyze the strength and stiffness of the struts of the Stewart platforms (Figure 128).

We use continuum approximations to analyze (1) the buckling force of a strut (nN), (2) the longitudinal stiffness (N/m), (3) the tensile strength (nN), and (4) the atom count. We compute these from the geometry (length, radius) and material properties (the Young's modulus) of the strut. Continuum mechanics is used as an order-of-magnitude guide; atomistic modeling/MD would be used to validate at the smallest diameters.

Struts in a Stewart platform are in pure tension or compression. The universal joints at both ends of the strut ensure that bending moments are small. Residual bending moments are neglected in this analysis. There is little torsional force. The slight torsional force is a consequence of the screw drive converting a small component of the compressive and tensile forces into torsion of the strut. We neglect this torsional force.

While the following designs assume the Young's modulus of diamond (∼one TPa), the actual designs will have to be converted into multiple blocks, that is, "blockified". This blockification process might cause a factor of four reduction in the Young's modulus (see Impact on axial stiffness of using AP Blocks for struts (Section 7.14.9)).

7.14.1 Buckling-Limited Force in the Stewart Platform Struts

The fundamental mechanical limit on the force that can be exerted by a Stewart platform is, in practice, often set not by material strength but by elastic buckling of its struts. The present analysis therefore begins with a deliberately simple abstraction of each strut as a slender tubular member loaded primarily in axial compression.

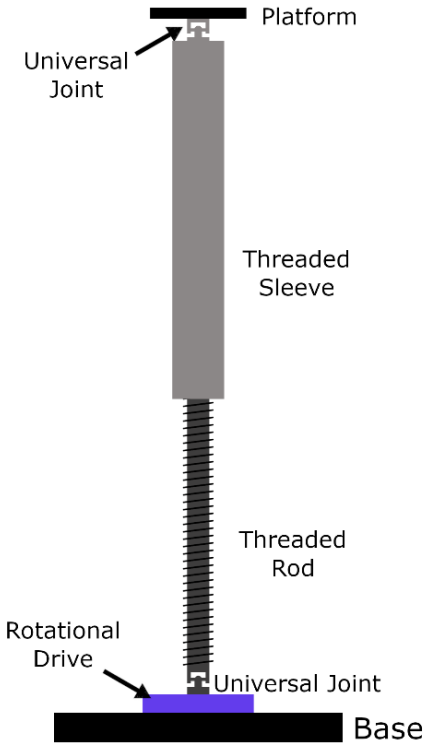


Figure 128: Stewart platform strut. Drawing repeated from Figure 18.

7.14.2 Idealized Strut Model

At the most basic level, each Stewart-platform strut may be modeled as a **tube**: a thin-walled, high-stiffness tubular column of length L, elastic modulus E, wall thickness w_t, and second moment of area I. For such a member subjected to axial compression, the governing failure mode for a sufficiently large slenderness ratio is buckling rather than yielding.

For each tube, D_o is the outer diameter, D_i is the inner diameter, and $D_m = (D_o + D_i)/2 = D_o - w_t$ is the mean diameter. The corresponding radii are $r_o = D_o/2$, $r_i = D_i/2$, and $r_m = D_m/2$. Exact section-property formulas use D_o, D_i; the thin-wall model uses D_m or r_m.

Under these assumptions, the maximum sustainable compressive load is given by the Euler critical buckling force

$$P_{cr} = \frac{\pi^2 EI}{(KL)^2}. \tag{11}$$

Here K is the effective-length factor.

Because the struts are terminated with universal joints that permit rotation and do not intentionally transmit bending moments, each strut is modeled as a pinned–pinned column. Under this assumption, the buckling load is reduced relative to a fixed-ended strut of the same geometry. All subsequent buckling calculations use $K = 1$ in Eq. (11), corresponding to pinned–pinned boundary conditions.

The following relations describe the buckling force and second moment of area:

The critical buckling force is **inversely proportional to the square of the strut length**:

$$P_{cr} \propto 1/L^2.$$

For a tubular strut the second moment of area for bending is

$$I = \frac{\pi}{4}\left(r_o^4 - r_i^4\right). \tag{12}$$

For a tubular strut of fixed wall thickness and material, with mean radius $r_m = (r_o + r_i)/2$ and wall thickness w_t:

$$I = \pi r_m^3 w_t \left(1 + \frac{w_t^2}{4r_m^2}\right). \tag{13}$$

If $w_t \ll r_m$ then $w_t^2/(4r_m^2)$ can be neglected, and we get:

$$I \approx \pi r_m^3 w_t. \tag{14}$$

If $w_t/r_m = 0.2$, this approximation is accurate to 1%. Even if $w_t/r_m = 0.4$, it is still accurate to 4%. As a consequence, we assume:

$$P_{cr} \approx \frac{\pi^2 E(\pi r_m^3 w_t)}{L^2} = \pi^3 E w_t \frac{r_m^3}{L^2}. \tag{15}$$

For $E = 10^{12}\,\text{Pa}$ and wall thickness $w_t = 0.5\,\text{nm}$, we have:

$$P_{cr} \approx (15{,}503\,\text{N/m})\,\frac{r_m^3}{L^2}. \tag{16}$$

If r_m and L are in nanometers then P_{cr} will be in nanonewtons.

That is, $P_{cr} \propto r_m^3/L^2$.

The thin-wall scaling in Eq. (16) dominates the mechanical design space of the platform: modest increases in strut diameter yield large gains in load capacity, while increases in strut length rapidly degrade achievable force.

7.14.3 Application to the Stewart Platform

In a Stewart platform, the platform-level force capability is obtained by resolving the individual strut forces along the platform geometry. Since each actuator must operate below its own buckling limit, the **minimum buckling load among the six struts** effectively bounds the maximum force of the entire mechanism.

Thus, even when material stresses remain far below yield, the platform becomes buckling-limited well before material failure is approached. Only when the platform is exerting forces in specific directions which minimize or eliminate compressive forces on individual struts are we not buckling-limited. This allows much larger forces to be exerted in specific directions, which can be useful in some particular cases.

7.14.4 Two-Piece Threaded Strut Geometry

In this section, we carry out an approximate analysis to support the feasibility of the design approach we have adopted. In particular, this approach should be able to achieve the desired force and positional accuracy within a volume that is not prohibitive.

In the design proposed here, each strut is not a single continuous tube but is instead composed of two coaxial elements:

1. A **threaded outer sleeve**, and
2. A **threaded inner rod**.

This composite geometry introduces local variations in stiffness, diameter, and moment of inertia along the strut length. Strictly speaking, the Euler buckling load must therefore be computed using a piecewise elastic column model with spatially varying EI.

However, because the global buckling mode remains dominated by the full strut length rather than by local instability in any of the sections, the composite strut behaves like a uniform diamond tube with comparable effective bending stiffness and pinned ends.

As a result, the critical buckling force of the assembled strut can be approximated by that of an equivalent uniform tubular column.

7.14.5 Governing Buckling Equations

The analysis that follows therefore uses the Euler buckling law in Eq. (11).

This approach preserves the essential $1/L^2$ and r_m^3 scaling and keeps the analysis simple.

These observations lead to fairly simple design heuristics. First, determine the geometry and force you wish the platform to exert. If possible, arrange for the direction of force application to be geometrically favored so that the forces on the struts are primarily tensile and not compressive. Then increase the radius r_m of the tubes comprising the struts until they can carry the remaining compressive loads without buckling.

7.14.6 Modeling of a strut

The parameters that describe the strut are listed below; Table 21 provides a consolidated symbol and units reference.

- T: temperature (K)
- E: Young's modulus (Pa)[12]
- D_1: outer diameter of threaded sleeve (m)
- D_2: outer diameter of threaded rod, treated as a tube (m)
- D_m: mean diameter of the inner rod, $D_m = D_2 - w_t$ (m), used for the conservative equivalent uniform tube.

[12]See Impact on axial stiffness of using AP Blocks for struts (Section 7.14.9), where it is concluded that blockifying the strut design might reasonably degrade the Young's modulus by a factor of four.

- L_1: length of both sleeve and rod (m)
- L_2: length of rod extending beyond sleeve (m)
- w_t: wall thickness of both sleeve and rod (m)

Lengths in SI formulas are in meters; where nanometer inputs give nanonewton outputs, this is stated explicitly. We use the inner rod dimensions for the equivalent uniform tube as a conservative assumption.

We assume the temperature T is 210 K.

We assume E = the Young's modulus of diamond $\approx 10^{12}$ N/m^2 (or Pa). The actual value is somewhat higher. This approximate value is adequate for the present analysis, where a conservative bound is sufficient. For a block redesign of this proposal it is likely that the Young's modulus will be reduced by a factor of ~four (see Impact on axial stiffness of using AP Blocks for struts (Section 7.14.9)).

We assume the wall thickness w_t is 0.5 nm.

$$\text{We assume} \quad D_2 = D_1 - 2w_t = D_1 - 1\,\text{nm} \tag{17}$$

$$\text{Total end-to-end strut length} \quad L = L_1 + L_2. \tag{18}$$

For modeling the buckling force, axial tensile strength, and positional uncertainty in the length of the strut caused by thermal noise, we treat the strut as a single tube with mean diameter D_m and wall thickness w_t. This model is least accurate when the threaded rod is fully retracted into the sleeve ($L_2 = 0$). In that case, the actual stiffness and buckling force will be higher than predicted by the model. The model will tend to understate the actual stiffness and buckling strength, a conservative error.

This model will be most accurate when the strut is fully extended and the strut is most likely to buckle. As this is the regime where buckling is of most concern, increased accuracy in this regime is an attractive property of this approximate model.

Axial stiffness follows Eq. (27), with tube area $A \approx \pi D_m w_t$. Thermal-noise axial uncertainty follows the equipartition relation in Eq. (23). Buckling follows Eq. (11) with $K = 1$ and $I \approx \pi D_m^3 w_t/8$. The thin-wall specialization, using $r_m = D_m/2$, is:

$$\begin{aligned} P_{cr} &\approx \frac{\pi^2 E \pi (D_m/2)^3 w_t}{L^2} = \frac{\pi^3 E w_t}{8}\frac{D_m^3}{L^2} \\ &\approx (1{,}938\,\text{N/m})\,\frac{D_m^3}{L^2}. \end{aligned} \tag{19}$$

If D_m and L are entered in nanometers the result is in nanonewtons.

The actual strut consists of a threaded sleeve and a threaded rod (treated as a tube). The exact axial compliance depends on the series and parallel combination of the sleeve, rod, and threaded overlap. For this approximate analysis, however, we conservatively treat the strut as a uniform tube having mean diameter D_m, wall thickness w_t, and total length L.

Carbon atom counts treat the sleeve and rod separately, and are estimated from total diamond volume and an atom density of 176 C atoms/nm^3 (see Useful Constants (Section 17.5)).

For a tubular cylindrical strut segment of mean diameter D_m and wall thickness w_t and length L, the volume is:

$$V_S = \pi D_m w_t L. \tag{20}$$

If the diameter D_m, the wall thickness w_t, and the length L are all measured in nm, then the result will be in nm^3. If we multiply this by 176, the number of carbon atoms in a cubic nanometer of diamond, we'll have the number of carbon atoms in the strut segment. For a fixed thickness of 0.5 nm, and if D_m and L are in nm, we have:

$$\begin{aligned} \text{Carbon atoms} &\approx (176\,\mathrm{nm}^{-3})\pi w_t D_m L \\ &\approx (276\,\mathrm{nm}^{-2}) D_m L. \end{aligned} \tag{21}$$

Applying Eq. (27), the axial stiffness of a tubular strut made of material with Young's modulus E, of mean diameter D_m, length L, and wall thickness w_t is approximately:

$$k_s \approx \frac{E\pi w_t D_m}{L} \approx (1{,}571\,\mathrm{N/m})\frac{D_m}{L}. \tag{22}$$

We also have the fundamental equation [Dre92]:

$$\sigma_x = \sqrt{\frac{k_B T}{k_s}}. \tag{23}$$

Here, σ_x is the RMS axial positional uncertainty.

For the ideal continuous-diamond thin-wall stiffness above ($E = 1\,\mathrm{TPa}$),

$$\sigma_x \approx \sqrt{\frac{k_B T L}{(1{,}571\,\mathrm{N/m}) D_m}}. \tag{24}$$

At $T = 210$ K, for this ideal continuous-diamond case, we get:

$$\sigma_x \approx (1.36 \times 10^{-12}\,\mathrm{m})\sqrt{\frac{L}{D_m}}. \tag{25}$$

For the baseline blockified strut, we assume an effective axial modulus $E_{\mathrm{eff}} = E/4 = 0.25\,\mathrm{TPa}$, following Section 7.14.9. This divides axial stiffness by four and doubles RMS uncertainty relative to the ideal continuous-diamond examples. This derating is a design assumption requiring validation.

Table 11 gives both cases for $L = 50\,\mathrm{nm}$ and $D_m = 10\,\mathrm{nm}$. Within this thin-wall estimate, increasing w_t from 0.5 to 2.0 nm multiplies k_s by four and halves σ_x. The 12 pm comparison threshold is an illustrative estimate, not a specified tolerance or an allocated error budget; the baseline positional accuracy remains 20 pm. Tail probabilities are two-sided Gaussian probabilities $P(|x| > 12\,\mathrm{pm})$, calculated from the unrounded thermal uncertainties.

Table 11: Thin-wall axial uncertainty estimates at 210 K for a 50 nm strut of mean diameter 10 nm. Ideal continuous-diamond and factor-four derated baseline cases are shown separately.

Case	Wall (nm)	RMS (pm)	12 pm / RMS	Two-sided tail
Ideal, 1 TPa	0.5	3.04	3.95	7.81×10^{-5}
Ideal, 1 TPa	2.0	1.52	7.90	2.79×10^{-15}
Baseline, 0.25 TPa	0.5	6.08	1.98	4.83×10^{-2}
Baseline, 0.25 TPa	2.0	3.04	3.95	7.81×10^{-5}

7.14.7 Tube section properties (both sleeve and rod)

For either the sleeve or rod, modeled as a tube with outer diameter D_o and wall thickness w_t:
The inner diameter, cross-sectional area, and second moment of area about any transverse axis are:

$$D_i = D_o - 2w_t \tag{26a}$$

$$A(D_o, w_t) = \frac{\pi}{4}(D_o^2 - D_i^2) \tag{26b}$$

$$I(D_o, w_t) = \frac{\pi}{64}(D_o^4 - D_i^4). \tag{26c}$$

(For circular tubes, buckling is isotropic in lateral direction.)

7.14.8 Longitudinal stiffness k_s (N/m)

7.14.8.1 Stiffness of each tube segment

Axial stiffness of a uniform member:

$$k_s = \frac{EA}{L}. \tag{27}$$

For the strut, whose total length is $L_1 + L_2$, this gives the lower bound on axial stiffness:

$$k_s \geq \frac{EA(D_2, w_t)}{L_1 + L_2}. \tag{28}$$

Here $A(D_2, w_t)$ is the rod area from Eq. (26b), since D_2 is its outer diameter. The non-strict inequality includes equality in the uniform-member limiting case. This lower bound on stiffness is arrived at by the conservative assumption that the stiffness of the most compliant segment of the strut is the stiffness of the entire strut.

7.14.8.2 Longitudinal positional uncertainty from thermal noise

Model the strut in the axial direction as a spring of stiffness k_s. At temperature T, equipartition gives:

$$\frac{1}{2}k_s\langle x^2\rangle = \frac{1}{2}k_B T. \tag{29}$$

The RMS axial positional uncertainty is $\sigma_x = \sqrt{\langle x^2\rangle}$, given by Eq. (23), where k_B is Boltzmann's constant.

7.14.8.3 Tensile strength check (verify it exceeds buckling)

A simple "material limit" tensile-force estimate is:

$$F_{\text{tensile,max}} = \sigma_{\text{tensile}}\, A_{\text{min}}, \tag{30}$$

where A_{min} is the smaller of the sleeve and rod cross-sectional areas, and σ_{tensile} is the allowable tensile stress.

Bending experiments on diamond nanoneedles, interpreted using finite-element modeling, indicate local tensile stresses at fracture up to 125 GPa [N+19]. This experimental result is distinct from the theoretical, direction-dependent strength range used below.

Whether tensile strength is "much more" than buckling depends on slenderness. The exact section properties are given by Eqs. (26b) and (26c).

Combining these section properties with Eqs. (30) and (11), the thin-wall force ratio scales roughly like:

$$\frac{F_{\text{tensile,max}}}{P_{cr}} \sim \frac{2}{\pi^2}\frac{\sigma_{\text{tensile}}}{E}\left(\frac{L}{r_m}\right)^2. \tag{31}$$

Here r_m is the mean tube radius, and $K = 1$ as in Eq. (11).

The theoretical tensile strength of atomically precise diamond nanoneedles is on the order of **∼90–225 GPa**, depending on direction [JZ11]. The ratio of the tensile strength to Young's modulus, $\sigma_{\text{tensile}}/E$, is therefore ∼0.1 to 0.2. If we take $\sigma_{\text{tensile}}/E = 0.1$, then the buckling force will equal the maximum tensile force when the L/r_m ratio is ∼7. An approximate example might be a strut which was 5 nm in mean radius and 35 nm long. The buckling force will fall below the maximum tensile force as the strut grows longer than 35 nm. In compression, longer struts buckle more easily.

For **longer, slender struts**, tensile strength rapidly becomes far larger than buckling, making buckling the governing mode; for **short, thick** struts, the two can be comparable.

7.14.8.4 Volumes

The sleeve and rod volumes and their total carbon count are:

$$\begin{aligned} V_{\text{sleeve}} = A(D_1, w_t)L_1, \qquad V_{\text{rod}} = A(D_2, w_t)L_1, \\ N_C \approx (V_{\text{sleeve}} + V_{\text{rod}})(1.76 \times 10^{29}\,\text{atoms/m}^3). \end{aligned} \tag{32}$$

(Where all lengths/diameters are in meters in this formula.)

7.14.8.5 Buckling force equations (pinned–pinned, $K = 1$).

The buckling load is given by Eq. (11), with $K = 1$ for the pinned–pinned strut.

7.14.9 Impact on axial stiffness of using AP Blocks for struts

The plan for blockifying the tubular structures in a Stewart platform strut is to use relatively long, thin blocks that overlap approximately half their length. Imagine the staves in a barrel, and now extend the barrel length by sliding every other stave half its length along the axis of the barrel. Then, stack barrels on top of each other, each barrel with half its staves interlocked with the barrel

above, and the other half interlocked with the barrel below. This overlapped tubular structure provides good axial stiffness while breaking the structure into blocks of roughly 500 atoms each (see Section 7.14.10).

For an order-of-magnitude estimate of the axial stiffness of an overlapped tubular structure, assume each block is 11 nm long, 0.5 nm wide, 0.5 nm thick, and made of diamond. Adjacent blocks overlap by approximately 5 nm, leaving a 1 nm axial gap between the butt ends of collinear blocks. Load transfer between overlapping blocks is assumed to occur through Ge–Ge bonds formed between germanium-terminated surface sites.

The relevant contact area per overlap is therefore approximately:

$$A_{\text{overlap}} = 5\,\text{nm} \times 0.5\,\text{nm} = 2.5\,\text{nm}^2. \tag{33}$$

Using the diamond (111) up-atom surface density as the nominal density of available termination sites gives roughly 18 sites/nm^2. Because the block face is only about 0.5 nm wide, edge effects and the requirement for three back-bonds into the block are expected to reduce the usable Ge termination density. Applying a derating factor of 2 gives an effective Ge–Ge bond density of approximately:

$$\rho_{\text{Ge–Ge}} \approx 9\,\text{bonds/nm}^2. \tag{34}$$

Thus the number of Ge–Ge load-transfer bonds per overlap is estimated as:

$$N_{\text{Ge–Ge}} \approx \rho_{\text{Ge–Ge}} A_{\text{overlap}} \approx (9\,\text{nm}^{-2})(2.5\,\text{nm}^2) \approx 23. \tag{35}$$

Taking the effective tangential (shear-transfer) stiffness of a single Ge–Ge linkage to be of order 50–100 N/m, the overlap joint stiffness is estimated to be 1.2×10^3–2.3×10^3 N/m.

$$k_{\text{joint}} \approx N_{\text{Ge–Ge}} k_{\text{Ge–Ge}} \approx 23(50\,\text{N/m}) \approx 1.2 \times 10^3\,\text{N/m}. \tag{36}$$

The axial stiffness of one diamond-like block segment is estimated from Eq. (27), using the block cross section and effective length. With $E \approx 1$ TPa, $A = 0.25$ nm^2, and an effective axial length of 11 nm, this gives:

$$k_{\text{block}} \approx \frac{(10^{12}\,\text{Pa})(0.25 \times 10^{-18}\,\text{m}^2)}{11 \times 10^{-9}\,\text{m}} \approx 23\,\text{N/m}. \tag{37}$$

Thus, under these assumptions, the block compliance dominates the axial compliance of a single load path:

$$\frac{1}{k_{\text{path}}} \approx \frac{1}{k_{\text{block}}} + \frac{1}{k_{\text{joint}}}. \tag{38}$$

Since $k_{\text{joint}} \gg k_{\text{block}}$, the effective path stiffness is close to the block stiffness, approximately 23 N/m. This suggests that, for the assumed geometry and Ge–Ge bond density, the overlapped Ge–Ge joint is not the limiting compliance; rather, axial compliance is dominated by stretching of the narrow diamond-like block itself. Additional Ge–Ge bonds would increase strength and redundancy but would provide little additional axial stiffness unless the block cross section were also increased.

The overlapping design results in an effective cross-sectional area of the blocks which is about half the cross-sectional area of the corresponding tubular design. Because each block is dominated by surface area, we will likely see a further reduction in effective axial modulus, possibly by as much as a factor of two. The total derating from the modulus of diamond might therefore be as much as a factor of four; this assumption is used for positional-uncertainty calculations and should be validated experimentally and/or theoretically when a blockified design becomes available.

We therefore assume the effective axial strut modulus, after blockification, will be degraded from the modulus of diamond by a factor of four.

This analysis neglects the further improvement in stiffness that might be achieved by compressing the butt-ends of longitudinally adjacent blocks into each other. This improvement could be significant.

While this general approach should be applicable to the axial strength of the design, the shear load is not shared equally among the Ge-Ge pairs because blocks stretch axially under load. The most heavily loaded bonds might rupture first, limiting total strength of block-block connections.

7.14.10 Number of carbon atoms in a block

The number of carbon atoms in a single block can be estimated from its volume and the atomic density of diamond. The block volume is:

$$V_{\text{block}} = 11\,\text{nm} \times 0.5\,\text{nm} \times 0.5\,\text{nm} = 2.75\,\text{nm}^3. \tag{39}$$

Diamond contains approximately 176 carbon atoms/nm^3. Therefore, the number of carbon atoms in a single block is:

$$N_C \approx (176\,\text{nm}^{-3})(2.75\,\text{nm}^3) \approx 484. \tag{40}$$

Thus, a single 11 nm × 0.5 nm × 0.5 nm block contains approximately 500 carbon atoms. This estimate is consistent with the small cross-sectional dimensions assumed above and indicates that edge and surface effects are likely to be significant. A substantial fraction of the atoms reside at or near the surface, supporting the expectation that the density of germanium termination sites will be lower than the ideal density predicted from an infinite diamond (111) surface.

7.14.11 Helical structures from overlapping blocks

We assume that the struts are made from multiple repeating units, like a polymer made by connecting monomers. In our example 11 nm block, we can reasonably assume that its 11 nm length is composed of a repeating unit that might have a repeat length of 0.5 to 1.0 nm. This produces 10 to 20 repeat units in our 11 nm example block. To create a helical structure, we consider a single slice down the full length of the tubular structure (the tubular structure being composed of many individual overlapping blocks). Recall that this tubular structure might be 50 to 200 nm in length, with the slice being just as long. We then slide all blocks on one side of the slice down by one repeat unit. That is, we slide all blocks on one side of the slice down by 0.5 to 1.0 nm. The individual repeat units remain aligned around the circumference of the tube. Most blocks remain aligned with their neighbors as before. As we go around the circumference once, one pair of blocks is misaligned with its neighbor by one monomeric unit. This will not have a significant effect on

the axial stiffness of the tube but will create a helical structure with the blocks now almost but not quite aligned with the axis of the helix. The axial stiffness remains close to what it was when the blocks were all perfectly aligned.

This is very close to the design method for helical structures described in Artificial helical structures and screw drives (Section 7.19.1).

A plane that cuts the tube at right angles to its axis will usually cut through all the blocks around the circumference. However, such a plane which cuts through the one nm axial gap between blocks will only cut through half as many blocks. This is where we lost a factor of about two in the axial stiffness of the tube.

If, instead of keeping all the axial gaps aligned axially around the circumference, we spread them out along the axis of the tube, then each plane cutting the tube would cut through almost all the blocks around the circumference, restoring most of the factor of two lost from aligning the gaps. Because there is a discrete integer number of blocks going around the circumference, the axial displacement of successive gaps must meet certain numerical conditions if we want the final gap in a single turn around the circumference to line up with the first gap in the next turn around the circumference.

We do not analyze this design approach here, other than to note that it should increase both the axial stiffness and the axial strength. For purposes of our baseline analysis, we assume that the axial stiffness of our tube is derated by a factor of four from the axial stiffness of diamond.

7.15 Differential Electrostatic Comb Drives for All Actuation

7.15.1 Overview

Key Takeaways

A small fixed-size linear electrostatic module (comb drive) could be used as the core for a family of linear electrostatic actuators.

This family could be used to meet all electrostatic actuation requirements in the workstation fabricator and the SRU. Direct mechanical actuation by motion of the SPM probe is a separate option.

This fixed-size module could be manufactured by a fabricator with a limited build volume.

Multiple identical modules could be assembled into larger linear electrostatic actuators with whatever force and throw might be required.

Continuous rotary motion can be derived from linear actuators.

7.15.2 Balancing forces: the differential comb drive

Differential comb drives were described by Tang et al. (1989) [TNH89] who said "the structure can be driven differentially (push-pull) using the two combs" (see Figure 129).

When a single comb actuator is used, van der Waals forces tend to pull the two combs together. If we used a single comb actuator as our linear actuator there would be a force pulling the two combs together at all times and pulling more forcefully when voltage is applied. If we only want to apply force when voltage is applied to our electrostatic actuator, then we must somehow cancel out the "resting" force. We prefer a design where the van der Waals forces are almost balanced, and only when the device is charged is the balance broken and net force created.

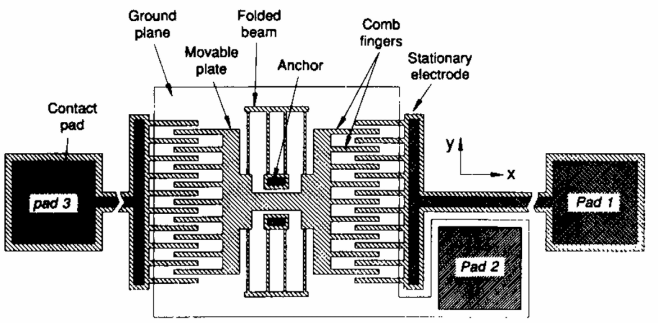
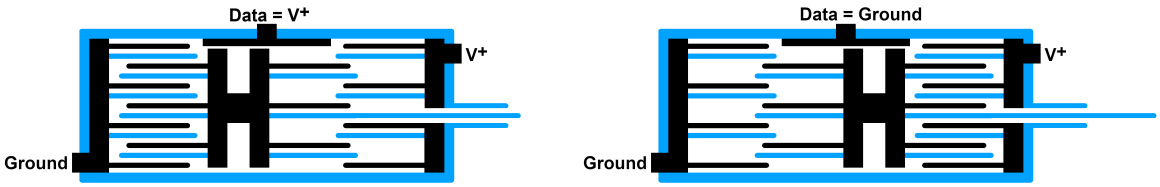


Figure 129: Published sketch of a differential comb drive, from [TNH89].

Our baseline solution is to use a differential comb drive, as illustrated (Figure 130):



(a) (b)

Figure 130: Differential comb drive displacement states: (a) Data = V+; (b) Data = Ground. Voltage can be applied to the central "Data" input to make the central black double comb move either left or right. Drawing repeated from Figure 15(a) and Figure 15(b).

7.15.3 Differential comb drives have constant capacitance

One differential comb drive consists of two opposing comb actuators acting on one shared shuttle. Each comb actuator has interdigitated stationary and moving tines. Four differential comb drives are installed in an SRU; one pair is enabled at a time, either S1 and S2 or D1 and D2 (see Section 4.3.4). Only one drive in the selected pair changes state during each clock phase. Here, an active pair means two complete differential comb drives, not the two comb actuators within one drive.

From a system perspective, an important property of a differential comb drive is that its capacitance remains approximately constant. As the capacitance of one opposing comb actuator increases because the overlap of the tines increases, the capacitance of the other opposing comb actuator decreases by the same amount because the overlap of its tines decreases. The capacitive load presented to the power supply remains the same.

7.15.4 Baseline design for a differential electrostatic comb drive

The cartoon in Figure 130 is for an atomically precise design which is described in the following paragraphs.

7.15.5 The differential comb drive for the SRU

This section covers the design of the differential comb drives required by the SRU.

7.15.5.1 Notation

The notation used in the equations for the various dimensions and materials parameters is taken from the diagram in Figure 131, which illustrates a MEMS scale comb drive. With that notational issue out of the way, we proceed with the discussion.

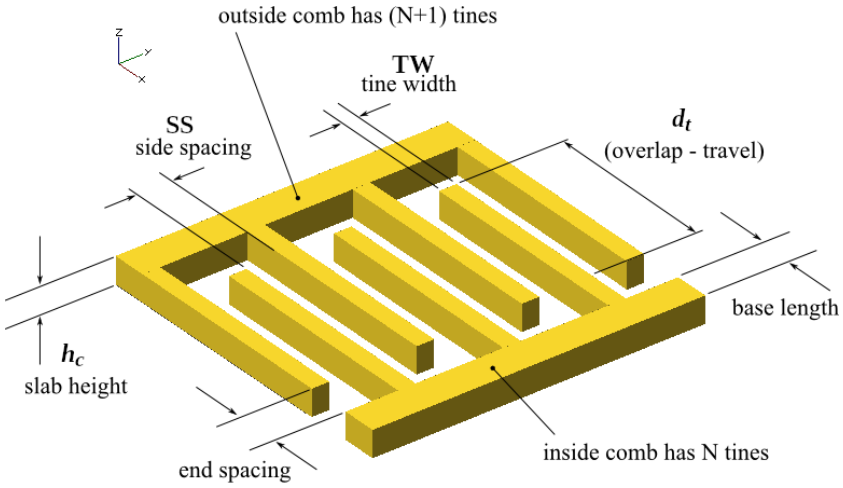


Figure 131: Comb drive geometry and notation.

7.15.5.2 h-BN design

The baseline h-BN differential comb drive operates from a 0.19 V supply and has a 24 nm shuttle throw. A 12:1 transmission provides a 2 nm output throw, with approximately 1 nN available to drive the SRU machinery. These satisfy the requirements of the SRU.

In this design, we alternate five sheets of insulating hexagonal boron nitride (h-BN) with one sheet of graphene. That is, we start with one conductive graphene sheet connected to V+, then we add five h-BN sheets, then we add one graphene sheet connected on the other side to ground, then we add five h-BN sheets, then we add one graphene sheet again connected to V+, then we add five h-BN sheets, then we add one graphene sheet again connected to ground, and so forth.

For baseline power budgeting, we assume that five atomically precise layers of h-BN (∼1.65 nm thick) limit leakage current to no more than 1 nA/µm^2 (1 fA/nm^2) at an applied voltage of 0.19 V. Experimental measurements show that tunneling conductance decreases exponentially with h-BN layer count [B$^+$12] (Figure 132). Inspection of the reported finite-bias current-voltage curves, together with the additional reduction expected from a fifth layer, suggests that 1 nA/µm^2 is a conservative design value. Because the dielectric is manufactured atomically precisely within the eutactic environment of the SRU, the baseline assumes that it contains no unintended defect-assisted conduction paths.

The Side spacing, or S_S, is the thickness of five sheets of h-BN. One sheet of h-BN is ∼330 pm thick, so the insulator is five times that. The tine width is one graphene sheet, which is also ∼330 pm thick. The inside comb has N tines (N graphene sheets) connected to ground and the outside comb has $N + 1$ tines ($N + 1$ graphene sheets) connected to V+. There are $2N$ insulating layers, each insulating layer being 5×330 pm in thickness. These connections describe a single energized

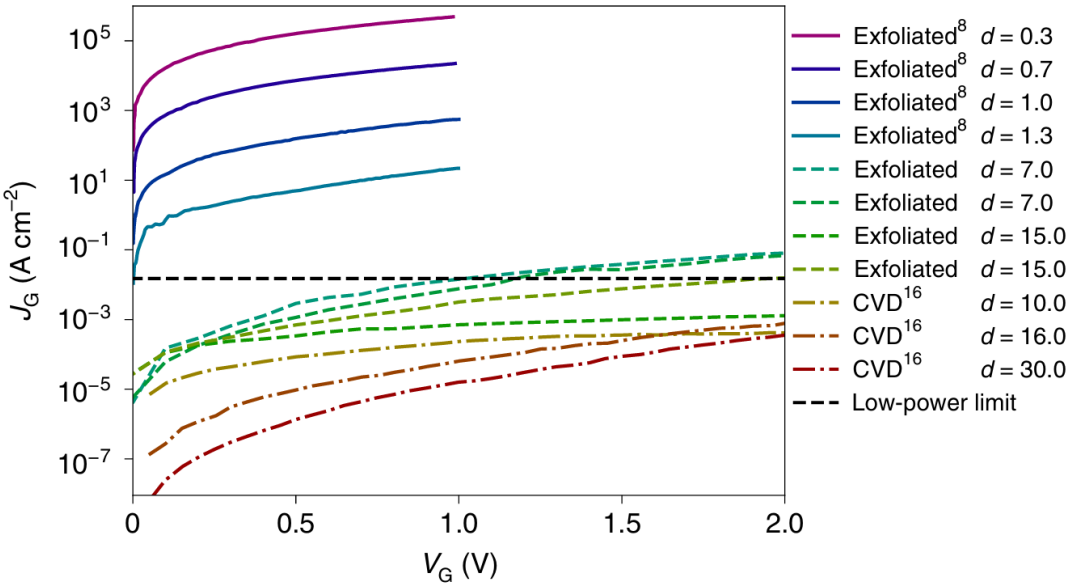


Figure 132: Measured leakage current density through h-BN samples of several thicknesses, reproduced from Figure 2 of [K+21], including measurements from [B+12]. The horizontal axis is applied gate bias V_G (V); the vertical axis is current density J_G (A/cm^2). The legend gives physical h-BN thickness d in nm and distinguishes exfoliated and chemical-vapor-deposited (CVD) samples.

comb actuator; the complete differential comb drive uses the Data-shuttle switching arrangement described below. This makes the comb drive total thickness:

$$\begin{aligned} \text{Total thickness} &= 2N(5 \times 330\,\text{pm}) + N(330\,\text{pm}) \\ &\quad + (N+1)(330\,\text{pm}) \\ &= (12N+1)(330\,\text{pm}). \end{aligned} \tag{41}$$

We can set N to whatever value is convenient for our design. This means the total thickness can be adjusted in increments of 3.96 nm. For convenience in subsequent calculations, we approximate the total thickness as $N \times 3.96\,\text{nm}$.

The slab height h_c (labeled h in Figure 131) is in a direction at right angles to the thickness. We can set the height of the comb drive to whatever value is convenient for our design.

The length is at right angles to both the height h_c and the thickness. The total length changes as the comb drive extends and contracts.

The shuttle throw is denoted by d_t (labeled "Throw" in Figure 131). This is the travel before transmission, not the output throw at the binary Bowden cable. The length consists of the following components:

- The overlap is $d_t + 1\,\text{nm}$ when the comb drive is in its contracted position: the throw d_t plus 1 nm of margin.
- The overlap is 1 nm when the comb drive is in its extended position because of the 1 nm of margin.

- The end spacing is 2 nm. There are 2 ends, which doubles the contribution of the end spacing to the total length.
- The base length is 2 nm. The base length contributes twice to the total length, once at each end.

The total length in the contracted position (going sequentially through the lengths): $d_t + 1\,\text{nm}$ (overlap margin) $+ 2 \times 2\,\text{nm}$ (twice the end spacing) $+ 2 \times 2\,\text{nm}$ (twice the base length) $= d_t + 9\,\text{nm}$.

The total length in the extended position is the contracted length plus the shuttle throw: $2d_t + 9\,\text{nm}$.

A single comb actuator therefore has a length of:

- Length: $d_t + 9\,\text{nm}$ when contracted.
- Length: $2d_t + 9\,\text{nm}$ when extended.

When used in a differential comb drive, sharing the central 2 nm base between the two opposing comb actuators reduces the combined length by 2 nm. The length is therefore fixed at:

$$(d_t + 9\,\text{nm}) + (2\,d_t + 7\,\text{nm}) = 3\,d_t + 16\,\text{nm}. \tag{42}$$

The three orthogonal directions for one differential comb drive are therefore:

- Length: $3d_t + 16\,\text{nm}$
- Thickness: $N \times 3.96\,\text{nm}$ for any desired value of N.
- Height: any desired value.

The physical volume V_{drive} of a single differential comb drive is:

$$V_{\text{drive}} = (3d_t + 16\,\text{nm})(N \times 3.96\,\text{nm})h_c$$

By adjusting d_t, N, and h_c we can fit a comb drive into a rectangular space. However, the force the comb drive will generate at a given voltage might not be the force we desire. We could achieve a desired force by adjusting the voltage or alternatively by using pulleys or levers on the comb drive output.

The energy a comb drive delivers per throw is:

$$\text{Energy} = F\,d_t \propto d_t\,N\,(3.96\,\text{nm})\,h_c\,V^2. \tag{43}$$

7.15.5.3 Mass of the shuttle

The shuttle is the two moving combs back to back, one from each opposing comb actuator. It has a central plate with a total volume of:

$$\begin{aligned} V_{\text{Center}} &= \text{base}\,h_c\,N\,(3.96\,\text{nm}) \\ &= (2\,\text{nm})\,h_c\,N\,(3.96\,\text{nm}). \end{aligned} \tag{44}$$

Multiplying the number of shuttle tines, $2N$, by the area of each tine gives the total shuttle tine surface area:

$$\text{Shuttle tine surface area} = 2\,N\,(d_t + 3\,\text{nm})\,h_c. \tag{45}$$

Diamond has ∼176 carbon atoms per nm^3.

Graphene has ∼38.2 carbon atoms per nm^2.

The number of carbon atoms in the central plate is:

$$C_{\text{plate}} \approx (176\,\text{nm}^{-3})\,[2\,\text{nm}\,h_c\,N\,(3.96\,\text{nm})]. \tag{46}$$

While the calculation here assumes the plate is simple diamond, in fact it would have to be made conductive, either by heavy doping with boron or by the judicious additions of conductive graphene, so that the tines to which it is connected will all carry the same voltage. The assumption is that this (now explicitly required to be conductive) plate is in sliding contact with an appropriate conductive strip which communicates the data signal to the plate.

The number of carbon atoms in the shuttle's tines is:

$$C_{\text{tines}} \approx (38.2\,\text{nm}^{-2}) \times 2\,N\,(d_t + 3\,\text{nm})\,h_c. \tag{47}$$

7.15.5.4 Shuttle travel within one half-cycle.

The two shuttle designs are compared in Table 12. The baseline mass used below is 1.08×10^{-20} kg.

Table 12: Shuttle geometry and mass for the baseline h-BN design and a larger alternative without a 12:1 transmission. Atom counts use Eqs. (46) and (47); mass uses 1.993×10^{-26} kg per carbon atom.

Quantity	**Baseline h-BN**	**Larger alternative**
Tine count N	7	22
Height h_c	22.5 nm	87 nm
Shuttle throw d_t	24 nm	2 nm
Plate carbon atoms	219,542	2,667,963
Tine carbon atoms	324,891	731,148
Total carbon atoms	544,433	3,399,111
Shuttle mass	1.08×10^{-20} kg	6.77×10^{-20} kg

From $F = ma$ and $d = \frac{1}{2}at^2$ we get:

$$d = \frac{1}{2}\frac{F}{m}\,t^2. \tag{48}$$

The available time is one half-cycle of the 100 MHz clock, 5 ns (the cycle time is 10 ns). The force is 0.1036 nN, so the distance is:

$$d = \frac{1}{2}\frac{0.1036 \times 10^{-9}\,\text{N}}{1.08 \times 10^{-20}\,\text{kg}} \times (5 \times 10^{-9}\,\text{s})^2 \approx 120\,\text{nm}. \tag{49}$$

As the desired travel is 24 nm, this design meets the baseline requirement. More to the point, it delivers enough force to move the shuttle those 24 nm and then still has sufficient force to provide ∼one nN to the molecular machinery of the SRU. At a 12:1 gear down, it provides $12 \times 0.1036 \approx 1.25$ nN of force, with ∼0.25 nN to simply move the shuttle within 5 ns and the remaining ∼1 nN available to power the molecular machinery of the SRU.

7.15.5.5 A design tradeoff

To reduce the energy spent accelerating the shuttle, we could increase the force and shorten the throw. Evaluating Eq. (56) for the h-BN geometry gives:

$$F \approx N\, V^2\, h_c \left(0.0182\, \frac{\text{nN}}{\text{V}^2\,\text{nm}}\right). \tag{50}$$

We can increase the force and shorten the throw while still delivering the same energy per throw by increasing N and h_c. If we increase N and h_c significantly, the force increases and the required travel decreases, which means we can move the shuttle more easily over a shorter distance using a larger force. Ultimately, we could deliver approximately 1 nN of force directly from the differential comb drive without any need to use a pulley system or a lever. We leave this possibility for a future redesign. A clean-slate redesign might place the comb drives with N and h_c aligning with the 500 nm and 100 nm axes of the overall SRU, with the throw along its 60 nm axis. This is a layout alternative to the baseline. For example, selecting h_c∼87 nm and $N = 22$, with $N \times 3.96\,\text{nm} \approx 87\,\text{nm}$, and lowering the voltage to 0.17 V, gives a force of ∼1.0 nN and requires a throw of 2 nm while avoiding the need for a block and tackle or a lever.

For the larger alternative in Table 12, the shuttle mass is 6.77×10^{-20} kg. Moving it 2 nm in 5 ns requires, from Eq. (48):

$$F = \frac{2dm}{t^2} = \frac{2(2\,\text{nm})(6.77 \times 10^{-20}\,\text{kg})}{(5\,\text{ns})^2} \approx 0.011\,\text{nN}. \tag{51}$$

This is only 1.1% of the nominal 1 nN drive force. Shuttle acceleration therefore consumes only a small fraction of the force available to drive the SRU's molecular machinery and can be neglected in this approximation.

7.15.5.6 Electrostatic equations and force

For the comb-drive symbols and their units, see Table 21. The electrostatically active area of one comb actuator in the contracted position is:

$$A_{\text{contracted}} = h_c\,(d_t + 1\,\text{nm})\,2\,N. \tag{52}$$

The electrostatically active area of one comb actuator in the extended position is:

$$A_{\text{extended}} = h_c\,(1\,\text{nm})\,2\,N. \tag{53}$$

The total change in area of the plates between the contracted and extended positions is the height times the throw times twice the number of tines on the inside comb, or

$$\Delta A_{\text{comb}} = h_c\, d_t\, 2\, N. \tag{54}$$

The total electrostatically active area of one differential comb drive is $A_{\text{drive}} = A_{\text{contracted}} + A_{\text{extended}}$:

$$A_{\text{drive}} = A_{\text{contracted}} + A_{\text{extended}} = 2Nh_c(d_t + 2\,\text{nm}). \tag{55}$$

Here h_c and d_t are lengths, V is the applied voltage, and N is the dimensionless tine count. The numerical coefficients below carry units, so these variables represent physical quantities rather than unitless numerical values. Here ϵ_r is the relative permittivity of h-BN, and ϵ_0 is the vacuum permittivity. The force delivered by one differential comb drive is:

$$\begin{aligned} F &= \frac{NV^2\epsilon_r\epsilon_0 h_c}{S_S} \\ &\approx NV^2 h_c \left(0.0182\,\frac{\text{nN}}{\text{V}^2\,\text{nm}}\right). \end{aligned} \tag{56}$$

The maximum energy one differential comb drive can deliver per throw is:

$$W = F\,d_t \approx N\,V^2\,h_c\,d_t \left(0.0182\,\frac{\text{aJ}}{\text{V}^2\,\text{nm}^2}\right). \tag{57}$$

7.15.5.7 Electronic time constant and heat

Applying the parallel-plate capacitance law in Eq. (95) to one differential comb drive gives:

$$\begin{aligned} C &= \frac{\epsilon_0\epsilon_r\,A_{\text{drive}}}{S_S} = \frac{(8.85\times10^{-12}\,\text{F/m})\times 3.4\,h_c\,(d_t + 2\,\text{nm})\,2N}{1.65\times10^{-9}\,\text{m}} \\ &\approx (0.0182\,\text{aF/nm}^2)\,h_c\,(d_t + 2\,\text{nm})\,2N. \end{aligned} \tag{58}$$

The coefficient is expressed in aF/nm^2; multiplying by the area gives the capacitance.

A useful resistance scale is the resistance quantum:

$$R_K = \frac{h}{e^2} \approx 25{,}812\,\Omega, \tag{59}$$

where R_K is the resistance quantum, h is Planck's constant and e is the elementary charge.

For an illustrative estimate, assume a molecular circuit has about four junctions, each with resistance of order R_K, giving approximately 100 kΩ (10^5 Ω). Actual junction resistances depend on their electronic structure and contacts. The assumed resistance gives an RC time constant below, where as denotes attoseconds:

$$\begin{aligned} \tau = RC &\approx (100\,\text{k}\Omega)(0.0182\,\text{aF/nm}^2)h_c(d_t + 2\,\text{nm})2N \\ &= (1{,}820\,\text{as/nm}^2)h_c(d_t + 2\,\text{nm})2N. \end{aligned} \tag{60}$$

Each differential comb drive maintains the same electrostatically active area A_{drive} during operation: as one opposing comb actuator contracts, the other extends by the same amount. The total electrostatically active area and hence the leakage current per drive remain constant. For the two drives in the selected clock pair, $A_{\text{active}} = 2A_{\text{drive}}$; for all four installed drives, $A_{\text{installed}} = 4A_{\text{drive}}$. These areas distinguish the operating pair from the full installed set.

Five layers of h-BN are ∼1.65 nm thick and should reduce leakage current to about 1 nA/μm^2 or 1 fA/nm^2 for an applied voltage of 0.19 V (see above).

This gives a leakage current of:

$$I_{\text{leak}} = j_{\text{leak}} A_{\text{drive}}, \qquad j_{\text{leak}} \approx 1\,\text{fA/nm}^2. \tag{61}$$

Computing power loss from leakage current:

$$P_{\text{leak}} = V I_{\text{leak}} = V j_{\text{leak}} A_{\text{drive}}. \tag{62}$$

7.15.5.8 Space occupied by four differential comb drives

The space available for four differential comb drives is 95 nm × 45 nm × 57 nm = 243,675 nm^3.

The formula for the space occupied by one differential comb drive is:

$$V_{\text{drive}} = (3d_t + 16\,\text{nm})(N \times 3.96\,\text{nm} + 0.330\,\text{nm})h_c. \tag{63}$$

Using Eq. (63), we might choose 88 nm × 22.5 nm × 28.05 nm as the dimensions for one of the four differential comb drives. We allocate 90 nm in the length dimension, providing a 2 nm allowance. This fits in the allotted space with a small amount of additional space for either a block and tackle or a lever to convert the low force long throw output into a high force short throw output.

Setting the three corresponding dimensions to $3d_t + 16\,\text{nm}$, h_c, and $N \times 3.96\,\text{nm} + 0.330\,\text{nm}$ gives us:

$$d_t = 24\,\text{nm}, \qquad h_c = 22.5\,\text{nm}, \qquad N = 7. \tag{64}$$

7.15.5.9 Force generated

Evaluating Eq. (56) for $N = 7$ and $h_c = 22.5\,\text{nm}$ gives:

$$\begin{aligned} F &\approx N\,V^2\,h_c \left(0.0182\,\frac{\text{nN}}{\text{V}^2\,\text{nm}}\right) \\ &= 7V^2(22.5\,\text{nm}) \left(0.0182\,\frac{\text{nN}}{\text{V}^2\,\text{nm}}\right) \approx (2.87\,\text{nN/V}^2)\,V^2. \end{aligned} \tag{65}$$

And the energy generated by the differential comb drive is therefore:

$$W = F\,d_t \approx (2.87\,\text{nN/V}^2)\,V^2\,(24\,\text{nm}) \approx (68.9\,\text{aJ/V}^2)\,V^2. \tag{66}$$

We need to deliver 2 nm × 1.25 nN = 2.5 aJ of energy. Therefore:

$$V = \sqrt{\frac{2.5\,\text{aJ}}{68.9\,\text{aJ/V}^2}} \approx 0.19\,\text{V}. \tag{67}$$

The voltage, V, is set to 0.19 V, and therefore $F \approx V^2 \times 2.87\,\text{nN/V}^2 \approx 0.1036\,\text{nN}$

$$W = Fd_t = (0.1036\,\text{nN})(24\,\text{nm}) \approx 2.49\,\text{aJ}.$$

The force delivered directly from this differential comb drive is too low. We use a block and tackle to increase the force. If the two blocks both have 6 pulleys, this gives us a 12:1 mechanical advantage, converting a force of 0.1036 nN over 24 nm to a force of ∼1.25 nN over 2 nm.

7.15.5.10 RC time constant

Using Eq. (60), the RC time constant is:

$$\tau = RC \approx (1{,}820\,\mathrm{as/nm^2})(22.5\,\mathrm{nm})(26\,\mathrm{nm})(14) \approx 15\,\mathrm{ps}. \tag{68}$$

One cycle is 10 ns or 10,000 ps. $15\,\mathrm{ps} \ll 10{,}000\,\mathrm{ps}$.

7.15.5.11 Leakage current

The leakage current of one differential comb drive will be:

$$I_{\mathrm{leak}} \approx 8.2\,\mathrm{pA}, \qquad P_{\mathrm{leak}} = V I_{\mathrm{leak}} \approx 1.56\,\mathrm{pW}. \tag{69}$$

For four differential comb drives this is 6.2 pW, or 0.6% of our 1 nW power budget per SRU. This is an acceptable leakage current.

7.15.5.12 Sliding Drag

"Sliding drag" is analogous to viscous drag and has power dissipation governed by a similar equation [Mer16, N+14]. The sliding drag equation that governs power dissipation when two atomically precise surfaces slide over each other (shown assuming MKS units) is:

$$P_{\mathrm{drag}} = K_{\mathrm{drag}}\, S\, v^2 \qquad \text{(sliding drag power equation)}, \tag{70}$$

where:

P_{drag} is the power dissipated (in W or J/s) by the sliding process.

S is the contact area in square meters ($\mathrm{m^2}$) between the two sliding surfaces,

v is the interfacial speed in meters/second (m/s) between the two sliding surfaces,

K_{drag} is the applicable drag coefficient, which depends on the temperature, the two surfaces, and the compressive force applied between the two surfaces. It has units of $\mathrm{J \cdot s/m^4}$, also equivalent to $\mathrm{kg/(m^2 \cdot s)}$.

For our baseline differential comb drive, the relevant surface area is $S = A_{\mathrm{drive}}$ from Eq. (55):

$$S = A_{\mathrm{drive}} = 2Nh_c(d_t + 2\,\mathrm{nm}) \approx 8.190 \times 10^{-15}\,\mathrm{m^2}. \tag{71}$$

While the actual speed profile of the moving shuttle is not constant, we approximate the interfacial speed as a constant for simplicity of analysis, which introduces a tolerable quantitative error in an essentially qualitative argument:

$$v = 24\,\mathrm{nm}/5\,\mathrm{ns} = 4.8\,\mathrm{nm/ns} = 4.8\,\mathrm{m/s}. \tag{72}$$

Using the sliding-drag law in Eq. (70):

$$P_{\mathrm{drag}} = K_{\mathrm{drag}} S v^2 \approx K_{\mathrm{drag}}(1.89 \times 10^{-13}\,\mathrm{m^4/s^2}). \tag{73}$$

We need to keep the heat loss from sliding friction in the differential comb drives under control. Quantitatively, we take this to mean that $P_{\mathrm{drag}} \leq 10\%$ of 1 nanowatt, or $P_{\mathrm{drag}} \leq 10^{-10}\,\mathrm{W}$. Combining this bound with Eq. (73) gives:

$$K_{\text{drag}} \leq \frac{10^{-10}\,\text{W}}{1.89 \times 10^{-13}\,\text{m}^4/\text{s}^2} \approx 529\,\text{kg}\,\text{m}^{-2}\,\text{s}^{-1}. \tag{74}$$

We need information on the coefficient of friction for h-BN on graphene, which is addressed in "Sliding friction of graphene/hexagonal-boron nitride heterojunctions: a route to robust superlubricity" [MLHU17]. The abstract says "The origin of ultra-low friction exhibited by heterogeneous junctions of graphene and hexagonal boron nitride (h-BN) is revealed. For aligned interfaces, we identify a characteristic contact size, below which the junction behaves like its homogeneous counterparts with friction forces that grow linearly with the contact area. Superlubricity sets in due to the progressive appearance of Moiré patterns resulting in a collective stick-slip motion of the elevated super-structure ridges that turns into smooth soliton-like gliding with increasing contact size. Incommensurability effects are enhanced in misaligned contacts, where the friction coefficients further drop by orders of magnitude."

If the two sliding surfaces are misaligned, the drag coefficient becomes very small. "Let us now turn to discuss the rotated interfaces with a misfit angle of $\theta = 30°$." In this misaligned case "The kinetic friction coefficient of both the homogeneous and heterogeneous misaligned junctions are found to be extremely small (in fact, below our numerical accuracy)". While there should be some drag, it seems as though our baseline constraint, that $K_{\text{drag}} \leq 529\,\text{kg}\,\text{m}^{-2}\,\text{s}^{-1}$, can be met by an appropriate atomically precise design. Further research on this point to determine the drag of a specific differential comb drive design should be pursued.

In our differential comb drive, it would seem sufficient to have long strips of graphene and h-BN that slide past each other, with one aligned so that its edges are zigzag, and the other aligned so its edges are armchair. These two strips would both be hexagonal in structure, but with their lattices rotated 30° with respect to each other. The repeat distances (assuming both had the same lattice constant) would be approximately 0.246 nm and 0.246 nm $\times\sqrt{3} \approx 0.426$ nm, which are incommensurate. The two lattice constants differ by ∼1.8%, so the purely mathematical result does not hold exactly, but two rotationally misaligned strips a few tens of nm long and wide should have remarkably low sliding friction.

We now discuss two ideas for further reducing sliding drag.

First idea: Five h-BN sheets between two graphene sheets provide six sliding interfaces: two graphene/h-BN and four h-BN/h-BN interfaces. If the five intervening h-BN sheets are not held in place but are instead free to slide, the relative motion between the sheets could be distributed across those six interfaces. With equal velocity sharing and comparable drag coefficients and contact areas, the velocity-squared drag model gives $6K_{\text{drag}}S(v/6)^2 = K_{\text{drag}}Sv^2/6$, approximately one-sixth of the dissipation of a single sliding interface.

Second idea: we observe that drag is a function of contact area. If we turn some of the h-BN sheets into Swiss cheese it would nominally reduce the contact area, possibly by a significant factor, which should reduce drag. Two sheets with the same pattern of holes might well increase drag, so which sheets should be modified and exactly how they might be modified requires further study. An initial attempt might modify only the h-BN sheets adjacent to graphene sheets. Putting holes in some of the h-BN sheets would also cause a drop in the dielectric constant of the insulating layers; the resulting dielectric constant would need to be evaluated.

7.15.5.13 Capacitance of one differential comb drive

Evaluating the capacitance estimate in Eq. (58) gives

$$C \approx (0.0182\,\text{aF}/\text{nm}^2)(22.5\,\text{nm})(24\,\text{nm} + 2\,\text{nm})(2 \times 7) \approx 149\,\text{aF}.$$

This design allows operation at 0.19 V instead of the 2 V used by the vacuum-dielectric alternative, while maintaining the required performance in a smaller volume. Low-voltage operation is critical to reducing power consumption.

7.15.5.14 Friction and wear

Properly designed atomically precise surfaces can slide over each other with remarkably little friction and without the asperity-driven wear characteristic of conventional surfaces [HMA17, Dre92, Mer93a], allowing the atomically precise comb-drive design to meet the friction requirements (see Sliding Drag (Section 7.15.5.12)). Any residual wear would require discrete events such as bond breaking, atom transfer, or surface reconstruction. For thermally activated events, the probability is governed approximately by the Arrhenius equation. The design objective is to maintain sufficiently large activation barriers, including the effects of applied mechanical stress, that the aggregate probability of such an event remains below the baseline failure threshold of 0.1% per SRU per replication cycle. Nonthermal damage mechanisms, including radiation damage, are treated separately; Drexler estimates them to remain below this threshold under the assumed operating conditions [Dre92].

Using the deliberately conservative assumption that every one of the 100,000,000 atoms in an SRU is an independently susceptible site, we can estimate the required activation barrier. Let the attempt frequency per site be $\nu_0 = 10^{13}\,\mathrm{s}^{-1}$, the replication time be $t = 10^5\,\mathrm{s}$, the number of potentially susceptible sites be $N_s = 10^8$, and the acceptable probability of at least one thermally activated failure during a replication cycle be $p = 0.001$. For rare independent events:

$$p \approx N_s \nu_0 t \exp\left(-\frac{E_a}{k_B T}\right). \tag{75}$$

Solving for the required activation energy gives:

$$E_a > k_B T \ln\left(\frac{N_s \nu_0 t}{p}\right). \tag{76}$$

At $T = 210$ K, $k_B T \approx 2.90$ zJ. Therefore:

$$\begin{aligned} E_a &> 2.90\,\mathrm{zJ}\ \ln\left[\frac{10^8 \times 10^{13}\,\mathrm{s}^{-1} \times 10^5\,\mathrm{s}}{10^{-3}}\right] \\ E_a &> 2.90\,\mathrm{zJ}\ \ln(10^{29}) \approx 194\,\mathrm{zJ}. \end{aligned} \tag{77}$$

Thus, even under the extremely conservative assumption that every atom in the SRU is a potential independent failure site, an activation barrier greater than approximately 194 zJ would reduce the predicted probability of a thermally activated failure to below 0.1% per 10^5 second replication cycle. This estimate assumes independent events, a conventional attempt frequency, and no lowering of the activation barrier by mechanical stress. A detailed design analysis must therefore use the lowest stress-modified activation barrier along the relevant failure pathways.

7.15.5.15 h-BN parameter values

Table 13 separates the inputs and calculated outputs for one baseline differential comb drive. The 12:1 transmission uses two blocks with six pulleys in each block. The matching vacuum-design values appear in Table 19.

Table 13: Baseline h-BN differential comb-drive inputs and outputs. Values are per complete differential drive, before transmission unless stated otherwise.

Parameter	**Symbol**	**Value**	**Role / assumption**
Side spacing	S_S	1.65 nm	Input; five 330 pm layers
End spacing	—	2 nm	Input
Base length	—	2 nm	Input
Slab height	h_c	22.5 nm	Input
Relative permittivity	ϵ_r	~ 3.4	h-BN assumption
Vacuum permittivity	ϵ_0	8.85×10^{-12} F/m	Physical constant
Voltage	V	0.19 V	Selected operating input
Tine count	N	7	Input
Shuttle throw	d_t	24 nm	Input
Raw shuttle force	F	0.1036 nN	Calculated; before transmission
Transmitted force	—	~ 1.25 nN	12:1 transmission
Transmitted throw	—	2 nm	12:1 transmission
Capacitance per drive	C	149 aF	Calculated
RC time constant	τ	15 ps	Calculated
Leakage per drive	I_{leak}	8.2 pA	Estimated
Leakage power per drive	P_{leak}	1.56 pW	Estimated

With $N = 7$, the exact thickness in Eq. (41) is 28.05 nm.

Calculated dimensions of each differential comb drive: 88 nm × 22.5 nm × 28.05 nm.

Allocated dimensions of each differential comb drive, including the 2 nm length allowance: 90 nm × 22.5 nm × 28.05 nm.

Dimensions of space allocated to the four differential comb drives: 90 nm × 45 nm × 56.1 nm. This gives an allocated volume of 227,205 nm^3.

7.15.5.16 Summary of h-BN electrostatic differential comb drive

The h-BN electrostatic differential comb drive, four of which are used in the baseline design to drive the SRU, has a 24 nm shuttle throw. A 12:1 transmission provides a 2 nm output throw with approximately 1 nN available to drive the SRU machinery. It operates at 0.19 V. Leakage current uses ∼0.6% of our power budget. The RC time constant (electronic switching time) is 15 picoseconds, 0.15% of the 10,000 ps cycle time. Four such comb drives have a combined allocated volume of 227,205 nm^3.

The volume allocated to four differential comb drives, each of which is able to deliver 1 nN of force from 0.19 V inputs, is 227,205 nm^3. The capacitance of one differential comb drive is 149 aF.

7.15.6 Some System Level Observations

One of the purposes of the electrostatic actuator is to provide a building block (basic module, actuator module, electrostatic actuator module) which can be used throughout all stages of development.

The baseline module is intended to have the following properties.

1) It is conceptually simple. It consists of two end plates, graphene tines attached to the right end plate, graphene tines attached to the left end plate, and a shuttle that moves back and forth with graphene shuttle tines separated from the left and right end plate tines by h-BN dielectric attached to the left and right end plates. Variations on this core design are possible.

2) The force that draws the two end plates together consists of (a) a "fixed" van der Waals component that does not vary with the voltage applied to the left and right tines, and (b) a "varying" electrostatic component that varies depending on the voltage. Various methods of canceling out the fixed force can be used, allowing the varying electrostatic actuation force to dominate. In our design, the fixed force is directly canceled because the shuttle is pulled in opposite directions by the two opposing fixed forces. The fixed force is larger in designs that use sliding contact between the conducting tines and the insulating dielectric.

3) Multiple modules can be assembled into multi-module assemblies (compound actuators) that have any desired force or throw. If the fabricator can manufacture a small module whose force or throw is insufficient, that module can be incorporated into a larger compound actuator.

4) It should be possible to carry out a static analysis of small variants of the basic module to determine the force it can apply when various voltages are applied, the leakage current that will flow at those voltages, and its capacitance. The dynamic response of the module can be approximated by a series of closely spaced static analyses, giving us a semi-static analysis of the dynamic response.

5) Knowing the I/V and Force/V curves of an individual module, along with its capacitance C, should let us adequately characterize the module for use in a fabricator or SRU.

6) We can adequately characterize the I/V, Force/V and C characteristics of a compound actuator given the I/V, Force/V and C characteristics of the modules from which it is composed.

7) All electrostatic actuation requirements during development and operation of a nanofactory could be met by modular electrostatic actuators. Some actuation requirements will be met by using the SPM probe (e.g., by mechanical actuation using physical motion of the SPM probe).

8) We could, in principle, design a single basic electrostatic actuator module and use it for all actuation (other than simple mechanical actuation mediated by the SPM probe) throughout the development and operation of the baseline nanofactory.

9) The parameters describing a single electrostatic actuator or differential comb drive could be systematically varied to evaluate and optimize its performance based on any of a variety of criteria, such as volume, throw, force, heat, internal drag, and the like.

Derisking these assumptions both individually as to the extent of their individual validity, and collectively as to the value of them from the perspective of the overall system, could initially be done computationally using theoretical models, before experimental versions are available.

7.16 H-trees for delivering power to the differential comb drives

An H tree (Figure 133) is a fractal pattern commonly used in semiconductor electronics as a wiring pattern for distributing clock and other signals to a large area. We use H trees to broadcast control signals to all the SRUs on a die. The recursive branching is also illustrated by the turtle analogy in Figure 134. The three dimensional turtle image reflects the fact that H trees in semiconductors often cross multiple layers. From this perspective, a turtle is a single H while the turtle's legs are vias connecting a single H on one layer to the center of an H on an adjacent layer. Such a three dimensional H tree provides increased flexibility for (1) wires of greater width than would be allowed if the entire fractal pattern were confined to a single layer, and (2) breaks or deviations from the regular pattern.

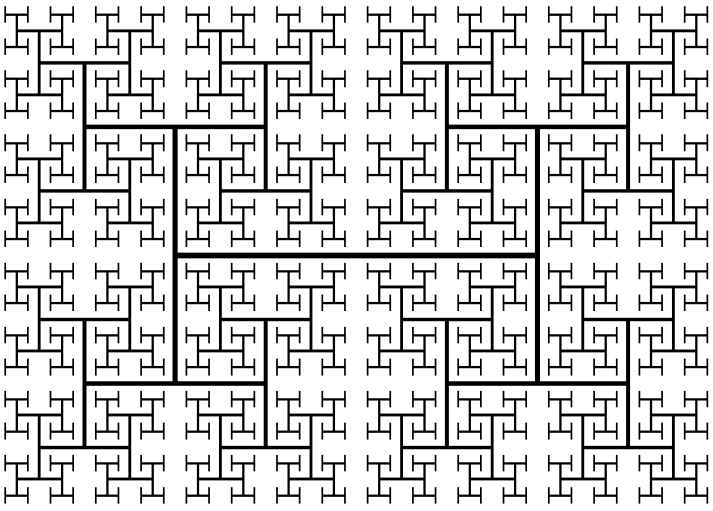


Figure 133: H-tree power distribution pattern. Static frame from Loopake's H-tree animation, licensed under CC BY-SA 4.0.



Figure 134: Turtle illustration of recursive branching. Picture by ChatGPT 5.6. Infinite turtles, all the way down.

We first determine the total wire length of one H tree distributing one of S1, S2, D1 or D2 to all the SRUs on a die.

Each die contains $2^{14} \times 2^{14} \times 3 = 805{,}306{,}368$ SRUs. There are H trees at levels 0 through 13 plus a 14^{th} half-level which has only horizontal crossbars and no vertical crossbars. The 14^{th} level has three SRUs in a 600 nm × 600 nm square with a single horizontal crossbar connecting all three. This single crossbar covers all three SRUs at this level. This crossbar runs 400 nm from the center of the first SRU across the center of the 2nd SRU and terminates in the center of the third SRU. Four such 0.6 µm × 0.6 µm squares form the next level: a square that is 1.2 µm × 1.2 µm. Call this level 13. Call the length of its sides "S_{13}". Level 13 has an H with a horizontal crossbar that is 1/2 S_{13} in length, and two vertical sidebars that are both 1/2 S_{13} in length. That is, this level has wires that are 3/2 S_{13} in length. Therefore, each level 13 square has $3/2 \times 1.2\ \mu\text{m} = 1.8\ \mu\text{m}$ of wiring. As there are $2^{13} \times 2^{13}$ such level 13 squares, the total length of level 13 wiring is 121 m.

In general, at level k, the square side length is $S(k) = (0.6 \times 10^{-6}\,\text{m})2^{14-k}$.

Each successive coarser level has one-quarter as many squares and twice the wire length per square, so its total wire length is halved.

Let $L_{\text{wire},0:13}$ denote the total length of wiring from level 13 to level 0 inclusive. Then:

$$\begin{aligned} L_{\text{wire},0:13} &= \left(2^{13} \times 2^{13} \times \frac{3}{2} S_{13}\right)\left(1 + \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^{13}}\right) \\ &\approx (121\,\text{m}) \times 2 = 242\,\text{m}. \end{aligned} \tag{78}$$

The terminal crossbars at level 14 add $(400\,\text{nm})2^{28} \approx 107\,\text{m}$, giving **approximately 349 meters of wire per H tree**.

7.16.1 Capacitance of the H tree wiring on a die

On-chip interconnect capacitance depends on metal layer, linewidth, spacing, and surrounding conductors, but 0.2 fF/μm (0.2 nF/m) is a commonly used first-order estimate for metal wiring [WH11].

At 0.2 nF/m, 349 meters of wire is 69.8 nF for the wiring broadcasting each of S1, S2, D1, and D2. This is the capacitance external to the SRUs on a die and does not count the internal capacitance of the comb drives.

Each SRU also contributes an additional 149 aF from the differential electrostatic comb drives internal to the SRU to which S1, S2, D1, and D2 are connected. This is $2^{14} \times 2^{14} \times 3$ SRUs × 149 aF/SRU = 120 nF for the whole die, bringing our total to 69.8 nF + 120 nF ≈ **190 nF for each data line on the die.**

The electronic loss from charging this capacitive load with a square wave will be CV^2 per cycle. This loss occurs even if the comb drives don't move. The two active lines dissipate $2 \times (190\,\text{nF}) \times (0.19\,\text{V})^2 \approx 13.7\,\text{nJ}$ per cycle; at 10^8 cycles/s this is **1.37 W per die**.

In addition to this electronic loss, there is also a mechanical loss. While the core requirement is that each differential comb drive and its transmission deliver an available output force of 1 nN, we design them to deliver a nominal output force of 1.25 nN. While this appears to be greater than our requirement, we found upon analysis that we needed to design the drive assemblies to deliver 1.25 nN at the output so that they will effectively be able to deliver 1 nN. There is a loss mechanism that degrades the effective force; see Shuttle travel within one half-cycle (Section 7.15.5.4).

At the transmission output, the nominal force is approximately 1.25 nN over a 2 nm throw. With $W = F_{\text{out}} d_{\text{out}}$, where d_{out} is the output throw, this corresponds to approximately 2.5 aJ per throw. The mechanical actuation losses are therefore 2.5 aJ per active line per throw. There are four throws per cycle, so this dissipates 10 aJ per cycle, or at 100 MHz we lose 1 nW. Multiplying by the number of SRUs on a die gives an additional loss of 0.805 W per die. Adding the mechanical loss to the electrical loss gives **2.18 W per die** of electrical and mechanical heat.

As there are 72,000 dies in the nanofactory, this is 157 kW. This is ∼2.7 times the 58,000 W "Actuation power budget per Nanofactory". The actuation power budget per Nanofactory is the maximum power actually needed to actuate all the SRUs. It does not include the electrical energy wasted in delivering that 58,000 watts to the SRUs, nor the energy wasted by the SRUs themselves in converting square-wave electrical power into mechanical power.

We have also not yet included the cooling costs, which are in addition to what we have just computed. See Cooling the nanofactory (Section 10.8).

7.16.2 Power required to run the nanofactory

The present first-order estimate for H-tree, load, and SRU switching dissipation is approximately 157 kW, excluding the additional dissipation that might be introduced by the buffer network once it is incorporated into the design.

With that qualification, the electrical power required to run the baseline nanofactory, using a simple square wave non-recovering power distribution method, should be ∼157 kW. Cooling this load will require additional power.

7.16.3 LC delays in an H tree

When a signal propagates from the root of an H tree to a leaf, there will be an LC delay. This delay is not expected to be significant, as the RC delay dominates, but we compute it anyway.

There is a useful first order approximation that says a wire has an inductance of ∼1 nH for every mm of its length.

Combining this inductance estimate with a capacitance of 0.2 pF/mm gives the signal propagation speed:

$$v = \frac{1}{\sqrt{L'C'}} \approx 7.07 \times 10^7 \,\text{m/s},$$

where L' and C' are inductance and capacitance per unit length. This is approximately 0.24 times the speed of light.

The die side length is $S_0 = (0.6\,\mu\text{m})2^{14} \approx 9.83\,\text{mm}$. The root-to-leaf path contributes $S_0/2$ at the first stage, with successive stage lengths halving. Including the final connection of at most 200 nm to an SRU, the path length is approximately

$$\ell \approx \frac{S_0}{2}\left(1 + \frac{1}{2} + \cdots + \frac{1}{2^{13}}\right) \approx S_0 \approx 9.83\,\text{mm}.$$

The propagation delay is therefore:

$$\text{Delay} = \frac{\ell}{v} \approx 0.139\,\text{ns}. \tag{79}$$

At 100 MHz, this is 0.139 ns out of a 10 ns cycle. This is ∼1.4% of a cycle time, which is acceptable.

We conclude that the LC propagation delay between the root and a leaf should not be a significant concern. Even variations in this delay (jitter) should be small enough not to be a concern.

7.16.4 RC delays in an H tree

RC delays are a significant concern in the design of a signal distribution system. We expect to use an H tree to directly distribute signals, but with signal buffering at one or more levels. Our analysis here is intended to establish the need for and the feasibility of a buffered H tree for distribution of S1, S2, D1 and D2, as well as to provide a general understanding of the approach. Further analysis will be needed to decide on the design of the buffered H tree, including what level(s) need to be buffered and the specifics of the buffering system.

One method of scaling an H tree is to assume constant resistance for each stage between successive levels of the H tree. The linear size doubles, and the area quadruples, going from level k to level $k-1$. Thus, $S(k-1) = 2S(k)$, and the corresponding wire lengths at level $k-1$ are twice those at level k. To maintain constant resistance, the cross-sectional area of the wires at level $k-1$ must therefore also double. For a rectangular wire of cross-sectional area height × width, we could scale both height and width by $\sqrt{2}$. We choose conductor cross sections at each level so that each stage contributes approximately 2.1 Ω of resistance. There are 14 such stages from the root to an SRU, giving a total root-to-SRU resistance of $14 \times 2.1\,\Omega = 29.4\,\Omega$.

The final stage leading to level 14 is a special case because it breaks the four-fold pattern and distributes the signal among three SRUs rather than four half-size H subtrees. For this first-order analysis, we nevertheless assign approximately the same 2.1 Ω resistance to this stage. The resulting approximation does not affect the qualitative conclusions of the analysis.

Table 14 summarizes H-tree geometry and SRU population by level.

Table 14: H-tree geometry and SRU population by level.

Level	Square side (µm)	Squares / die	SRUs / subtree
0	9830.4000	1	805,306,368
1	4915.2000	4	201,326,592
2	2457.6000	16	50,331,648
3	1228.8000	64	12,582,912
4	614.4000	256	3,145,728
5	307.2000	1,024	786,432
6	153.6000	4,096	196,608
7	76.8000	16,384	49,152
8	38.4000	65,536	12,288
9	19.2000	262,144	3,072
10	9.6000	1,048,576	768
11	4.8000	4,194,304	192

Table 14 (continued)

Level	Square side (µm)	Squares / die	SRUs / subtree
12	2.4000	16,777,216	48
13	1.2000	67,108,864	12
14	0.6000	268,435,456	3

The capacitance at the root is shown in Table 15 as 1.898×10^5 pF (189.8 nF). The corresponding resistance is shown as 29.4 Ω, allowing us to compute the crude RC delay: approximately 5,580 ns. This is inaccurately long for an H tree. An Elmore delay would be a better approximation. The simpler RC delay (shown in Table 15) conveys the qualitative message: we cannot simply drive signals from the root; we need some sort of buffering at one or more intermediate levels.

Buffered H trees are commonly used to distribute signals in integrated circuits which are about 1 cm^2 in size. The design of a die able to deliver ground, V+, S1, S2, D1, and D2 to the roughly 800 M SRU sites while satisfying the requirements of the baseline should be within the current state of the art.

Table 15 summarizes H-tree capacitance, resistance, and RC delay by level.

Table 15: Estimated H-tree capacitance, resistance, and RC time constant by level for one signal tree. RC values are rounded to three significant figures.

Level	Subtree wire C (pF)	Subtree SRU load C (pF)	Subtree total C (pF)	Remaining R to center leaf (Ω)	Crude RC time constant (ps)
0	6.979E+04	1.200E+05	1.898E+05	29.400	5.58×10^6
1	1.745E+04	3.000E+04	4.744E+04	27.300	1.30×10^6
2	4.361E+03	7.499E+03	1.186E+04	25.200	2.99×10^5
3	1.090E+03	1.875E+03	2.965E+03	23.100	6.85×10^4
4	2.724E+02	4.687E+02	7.412E+02	21.000	1.56×10^4
5	6.807E+01	1.172E+02	1.852E+02	18.900	3.50×10^3
6	1.699E+01	2.929E+01	4.629E+01	16.800	778
7	4.237E+00	7.324E+00	1.156E+01	14.700	170
8	1.053E+00	1.831E+00	2.884E+00	12.600	36.3
9	2.605E-01	4.577E-01	7.182E-01	10.500	7.54
10	6.368E-02	1.144E-01	1.781E-01	8.400	1.50
11	1.520E-02	2.861E-02	4.381E-02	6.300	0.276
12	3.440E-03	7.152E-03	1.059E-02	4.200	0.0445
13	6.800E-04	1.788E-03	2.468E-03	2.100	0.00518

Table 15 (continued)

Level	Subtree wire C (pF)	Subtree SRU load C (pF)	Subtree total C (pF)	Remaining R to center leaf (Ω)	Crude RC time constant (ps)
14	8.000E-05	4.470E-04	5.270E-04	0.000	0

Table 15 assumes the capacitance of a wire $C = 0.2$ fF per µm [WH11], and that the total resistance between levels is a constant 2.1 Ω.

7.17 Driving a demux with four binary Bowden cables

Sub-angstrom helical or screw drive (Section 7.19.1) mechanisms are an attractive approach for achieving the 20 picometer positional accuracy required for mechanosynthesis. At the same time, many decades of experience in the manufacture of logic devices suggests that binary device designs are advantageous when compared with analog or multi-valued digital device designs. This supports the use of binary device designs as much as possible, switching to rotary stepped motion only at the last moment when it is required to drive the screw drive mechanism used by the Stewart platforms. Rotational stepped motion might conveniently be driven by binary mechanical signals from a pair of binary Bowden cables. A pair of binary Bowden cables would be two cables, with each individual inner cable being able to assume one of two fixed positions, one position representing "0" and the other position representing "1". The two positions of each inner cable might be separated by, for example, 2 nm. Various mechanical implementations can be used to convert a sequence of 1s and 0s on a pair of binary Bowden cables into stepped clockwise and counterclockwise motions of a wheel, thus driving a screw drive and achieving the desired 20 pm steps required for mechanosynthesis in the fabricator. See Driving a screw drive with four-phase logic (Section 7.18).

Given this, the output of the demux (Section 4.3.6) might reasonably be tens of pairs of binary Bowden cables, driving various stepped wheels, stepped sliders (which could likewise be stepped left or right to actuate various mechanical devices) or other devices as might be required in the SRU.

The input to the demux might, in such a case, be four binary Bowden cables (two pairs of binary Bowden cables). One input pair would select which output pair would be driven and would allow the demux to select the previous output pair or the next output pair, and the other input pair would actually drive the selected output pair. That is, the demux would have four input binary Bowden cables (two input pairs) and 2 × N output binary Bowden cables (N output pairs).

While we might well choose some other approach, this design is both plausible and allows us to more clearly imagine "how it might work".

Four differential comb drives convert the electrical inputs S1, S2, D1 and D2 into the four binary Bowden-cable inputs to the demux. Each drive has two opposing comb actuators sharing a shuttle attached to one inner cable (Figure 130); see Baseline design for a differential electrostatic comb drive (Section 7.15.4).

The six electrical inputs—V+, Ground, S1, S2, D1 and D2—reach the SRU through six contact pillars and six spikes on three feet; see Feet (Section 4.3.2) for the attachment mechanism. The resulting control chain runs from the contact pillars through the four comb drives and their Bowden cables to the demux, which drives approximately 60 output pairs to actuate the SRU's mechanical devices. Figure 135 shows a symbol for the demux and its cable connections.

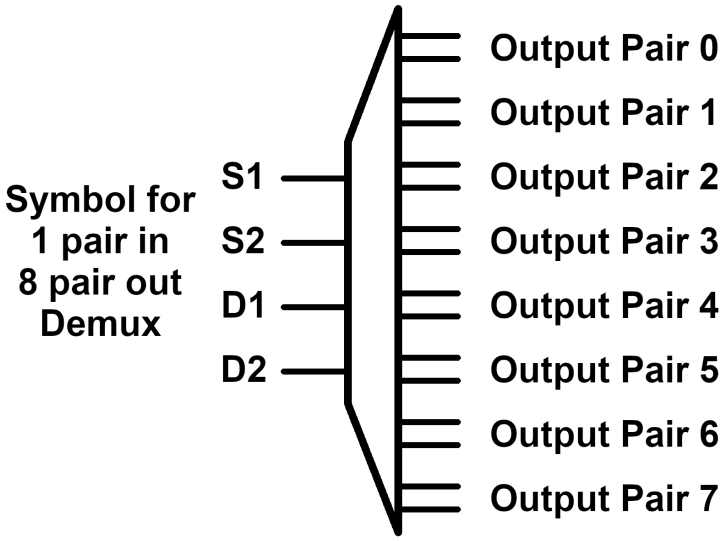


Figure 135: Demultiplexer symbol and binary Bowden cable connections. Each of the black lines going into or out of the demux is one binary Bowden cable. On the left, we see four binary Bowden cables going into the demux: S1, S2, D1 and D2.

The two selectors, S1 and S2, operate together to select which pair of output binary Bowden cables is to be driven by the two drive inputs, D1 and D2. S1 and S2 use four-phase clocking to move sequentially forwards and backwards through the output pairs, from zero to seven in the illustrative eight-output demux. After a pair of output binary Bowden cables has been selected, that pair is driven by the input pair of binary Bowden cables D1 and D2. For more, see Broadcast Architecture (Section 6.1).

The contact pillars on the die are connected to an external computer. The contact pillars are driven by the external computer at 100 MHz.

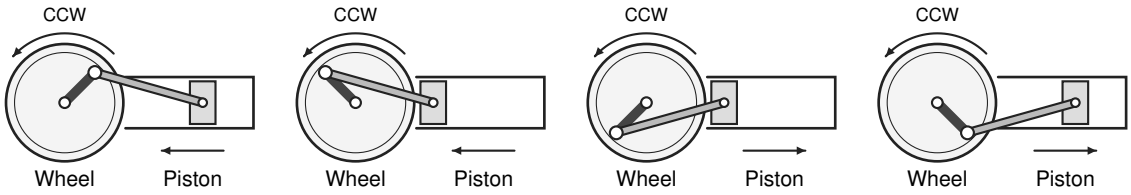
7.18 Driving a screw drive with four-phase logic

The screw in a screw drive must be stepped through some fixed rotary motion to advance the screw by a fixed distance. We have a positional accuracy requirement of 20 pm (see The 20 picometer requirement (Section 3.5.4)). A screw with threads with a pitch of ∼0.45 nm (as in the example screw drive in Section 7.19.3) would need to rotate 20/450 × 360° =16° to deliver a distance step of 20 pm. A full 360° rotation would take 22.5 such 16° steps. This leads to a requirement to rotate a screw by roughly ≤16°. This requirement is equivalent to a requirement to rotate a shaft or a gear by that same 16°. This requirement depends on the pitch of the screw.

7.18.1 Rotary motion from linear motion

The comb drives provide linear motion, which must be converted into rotary motion to turn the screw drives. A piston driving a wheel illustrates a simple conversion mechanism (Figure 136):

We could imagine replacing the piston with an electrostatic comb drive (or, indirectly, by a binary Bowden cable driven by a comb drive), which would allow us to scale this method down to a very small size. Unfortunately, the macroscopic piston relies on the momentum of the wheel, which is



(a) (b) (c) (d)

Figure 136: Four successive states of a piston driving a wheel counterclockwise (CCW). Arrows indicate motion. From a to b and from c to d, piston force makes the wheel spin faster. From b to c, and from d to a, inertia keeps the wheel spinning.

essentially absent at the molecular scale. In the absence of momentum, a molecular-scale wheel would randomly rotate forward and backward under the influence of a linear driving force, such as that provided by a single comb actuator.

We could, however, drive the wheel with two comb drives at right angles to each other (Figure 137):

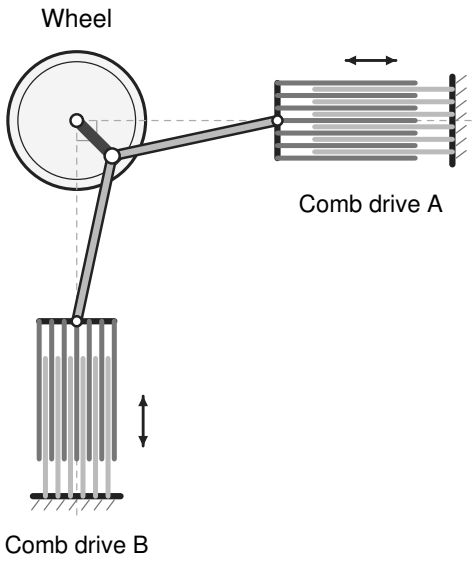


Figure 137: Two perpendicular comb drives acting on a shared crank pin. The crank position matches Figure 136(d). The rotating wheel never coasts from inertia, but is always driven by one or the other comb.

By using this method, the rotating wheel is never allowed to rotate freely. It is always held in position and is never at "top dead center" for both comb drives simultaneously, as would happen with a single comb drive. This method also allows us to rotate the wheel either forward or backward, by controlling the phase of the comb drives.

For the two binary inputs A and B, one four-phase cycle follows $00 \rightarrow 10 \rightarrow 11 \rightarrow 01 \rightarrow 00$. Reversing this order reverses the wheel's motion. The perpendicular connections shown in Figure 137 let one drive maintain control as the other passes through dead center.

The main problem with these designs is that they rotate the output shaft once per four-phase clock cycle, which is too rapid in many of our applications. This can be addressed by gearing down the rotational motion.

7.18.2 The baseline: worm gears



Figure 138: Wooden worm gear transmission. Photograph: Rama, *Engrenage à vis creuse CNAM-IMG 0602*, licensed under CC BY-SA 2.0 France; reproduced without modification.

A worm gear (Figure 138) is a right-angle mechanical transmission consisting of a screw-like shaft (the "worm") meshing with a toothed wheel. Rotation of the worm advances the wheel by one tooth per turn, enabling very large speed reduction and corresponding torque amplification in a compact form. The sliding contact between worm and wheel often produces a self-locking effect, allowing motion in one direction while resisting back-driving.

If we had a gear with 24 teeth, then rotating that gear by one tooth would step it 360° / 24 = 15°, which would in turn advance a helical screw drive with a thread pitch of 450 pm by 450/24 = 18.75 pm.

7.18.3 Summary

If we are to use a screw drive to achieve our 20 pm positional accuracy requirement, and the finest pitch of our screw is ∼450 pm, this translates into a requirement for ≤16° rotational steps of the screw. As we are using four-phase clocking, we must translate a four-phase clock signal carried on two binary Bowden cables into rotational steps of ≤16°.

Our baseline proposal is a worm gear. If the worm were directly driven by two binary Bowden cables at one turn per clock cycle, a 24 tooth gear would deliver 15° rotational increments per cycle, which meets our requirements.

Implicit in this analysis and our use of Stewart platforms is that the use of screw drives to extend and contract the struts of a Stewart platform by increments of 20 pm will roughly translate into movement of the platform by something like 20 pm. When the Stewart platform is close to a singular configuration, this assumption might be significantly in error. The layout of the Stewart platforms and the expected operating envelopes are intended to avoid singular configurations, making our crude analysis tolerable as a first pass. A more accurate analysis of the kinematics of Stewart platforms would be helpful.

7.19 Helical structures

7.19.1 Artificial helical structures and screw drives

The design approach for making the large helical structures required for the struts of the Stewart platform (Section 4.4.1) (the threaded sleeve and the threaded rod) from AP Blocks will be to find a simple repeat unit from which they can be built. Consider, for example, the simple hydrocarbon bearing (Figure 139):

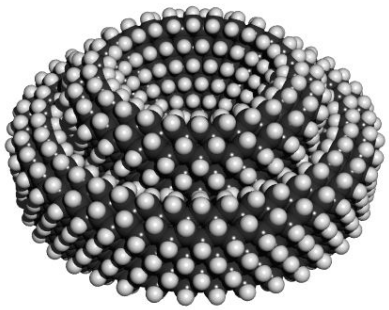


Figure 139: Hydrocarbon bearing as a starting point for helical structures.

Now extend both the central shaft and the sleeve. Slice lengthwise and slide one half of the slice one unit cell along the other half of the slice. This creates two helices that slide smoothly against each other, one being the threaded rod and the other being the threaded sleeve. Both the threaded rod and the threaded sleeve are two-dimensional surfaces. Cut two patches out of these two surfaces, each patch having fewer than 1,000 atoms. Adjust the patches so that, when repeated in two dimensions, they will form two surfaces that approximate the original two surfaces from which they were taken so that the resulting two surfaces can slide over each other with a low barrier to helical rotation. Experience in the design of atomically precise structures is helpful in this endeavor.

The result will be two helical struts that can be screwed together, each helical strut being formed from a single repeating small (<1,000 atom) atomically precise AP Block. This forms a screw drive.

For a more detailed description of how this might work with a blockified design, see Helical structures from overlapping blocks (Section 7.14.11).

7.19.2 How to design screw drives

The screw drive in this section was designed by James MacArthur. This section was adapted from his unpublished report [Mac]. The screw drive has a threaded rod and a threaded sleeve. In Figure 140, r_1 is the diameter of the threaded rod and r_2 is the diameter of the threaded sleeve. In this design, the threads of the threaded rod and of the threaded sleeve are hydrogen atoms on the surface of lonsdaleite. If $r_2 - r_1$ is too big, the threads won't engage and the screw drive will exert little if any axial force. On the other hand, if $r_2 - r_1$ is too small, the torque required to rotate the screw drive will be too high.

While the design in this section cannot be manufactured by the SRU Fabricator because it exceeds the allowed atom counts and needs to be converted to AP Blocks, it illustrates an approach for designing an atomically precise helical screw drive which is expected to require a relatively small torque to rotate.

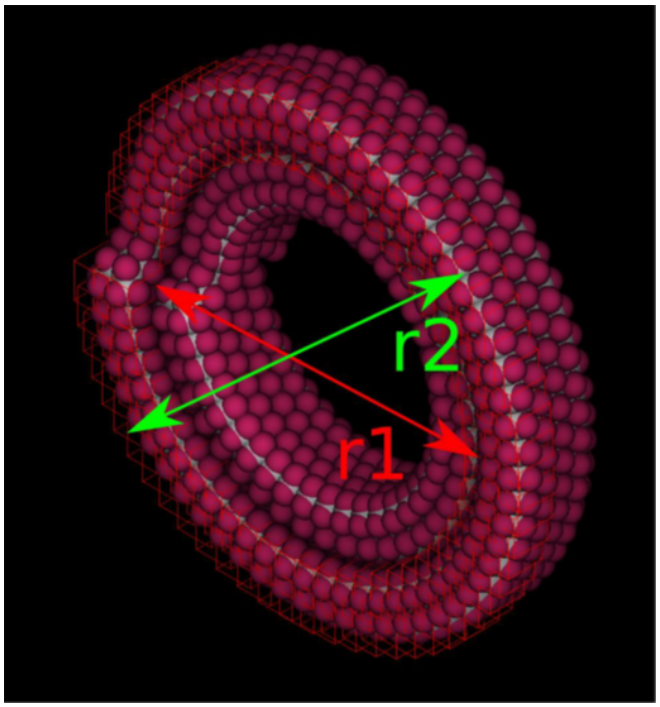


Figure 140: MacArthur molecular screw drive. From James MacArthur's unpublished report [Mac].

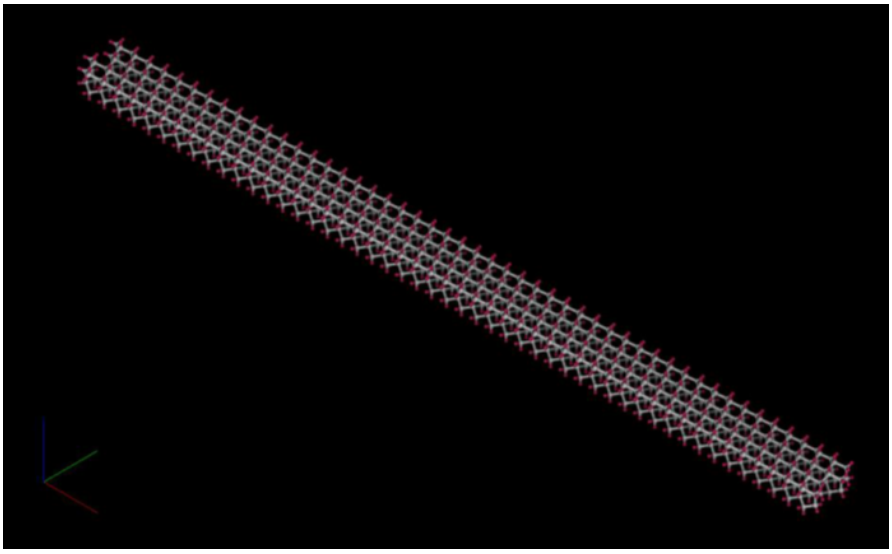


Figure 141: Lonsdaleite strip used to construct a helical screw drive. From James MacArthur's unpublished report [Mac].

To make the screw drive, start with a strip of lonsdaleite (Figure 141).

Wrap it into a ring and offset the ends (Figure 142).

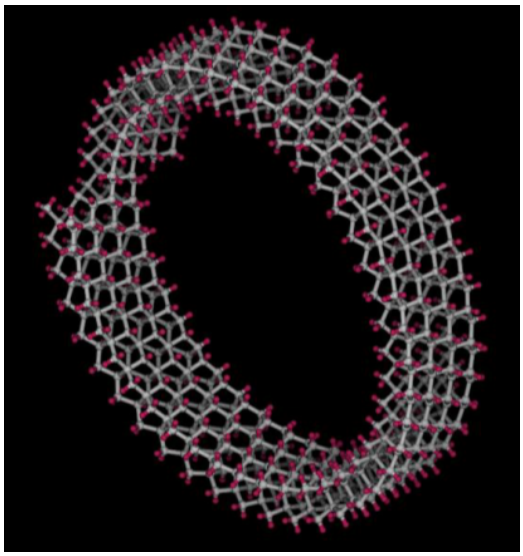


Figure 142: Lonsdaleite strip wrapped into an offset ring. From James MacArthur's unpublished report [Mac].

MacArthur provides an illustrative analysis of the tradeoff between torque and axial force as a function of r_1 and r_2 (see Figure 143).

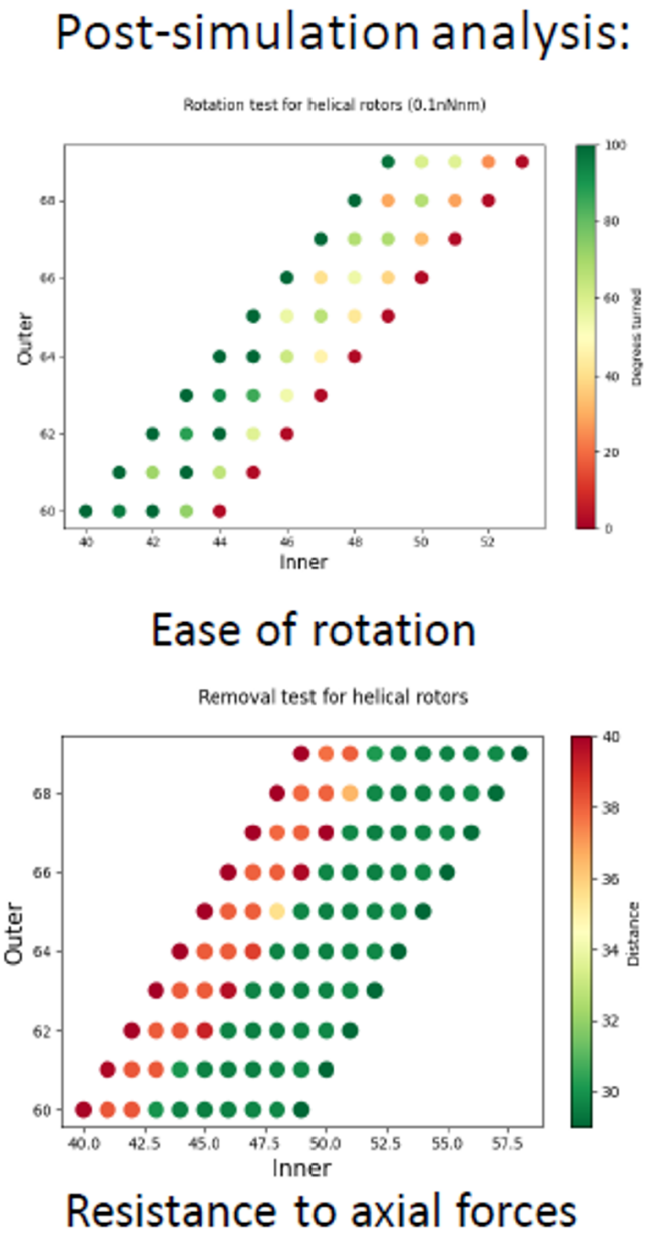


Figure 143: Calculated screw-drive tradeoff between ease of rotation and resistance to axial force. From James MacArthur's unpublished report [Mac].

7.19.3 An example screw drive

The screw drive in this section was designed by James Ryley. This section was adapted from his unpublished report [Ryl].

As with the preceding design, this design exceeds the allowed atom counts and would need conversion to AP Blocks. It illustrates an atomically precise helical screw drive expected to require relatively little torque to rotate.

7.19.3.1 Inner rod design

A rectangular sheet of lonsdaleite was wrapped into a hoop, and the edges fused. But, before edge fusing, one edge was shifted down a single unit cell. A lonsdaleite unit cell is ∼4.4 Å along the axis of the hoop. This resulted in, if we apply CNT nomenclature to a diamond hoop, something analogous to a 39-1 but instead of being made of graphene, the hoop was made of a sheet of lonsdaleite 2 unit cells thick.

Because the hoop is a 39-1, with what in cubic diamond would be the (111) face (in lonsdaleite sometimes called the (0001) face) pointing outward and inward, a very gradual spiral (a double spiral or helix actually, because each unit cell is 2 hydrogens wide) is formed from what would normally be straight rows of hydrogens on the surface. Every other hydrogen was removed from the surface and replaced with -CN. This allows the -CN to form threads that are raised above a hydrogenated trough, allowing greater axial force (Figures 144 and 145).

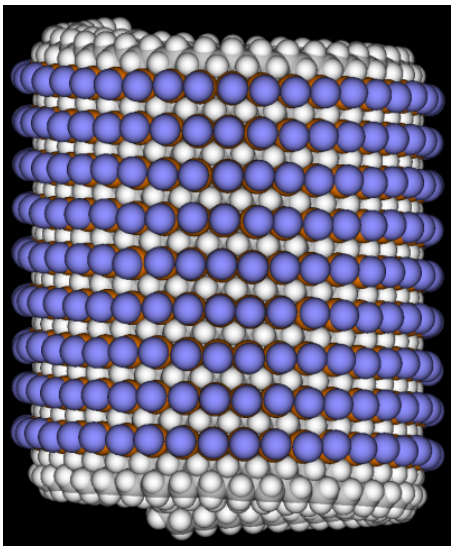


Figure 144: Side view of a lonsdaleite threaded rod with CN threads. From James Ryley's unpublished report [Ryl].

The inner rod's minimum radius is limited by how tightly diamond can be bent without causing too much strain (Figure 145). The chemical stability of this structure has not yet been verified.

7.19.3.2 Outer threaded sleeve

The outer threaded sleeve into which the threaded rod fits has to be designed a bit differently. The -CN projections, when pointed inward, are too crowded. So, every other one along the spiral path was removed. The size of the threaded sleeve was determined empirically. It ended up being a 63-1 (Figures 146 and 147).

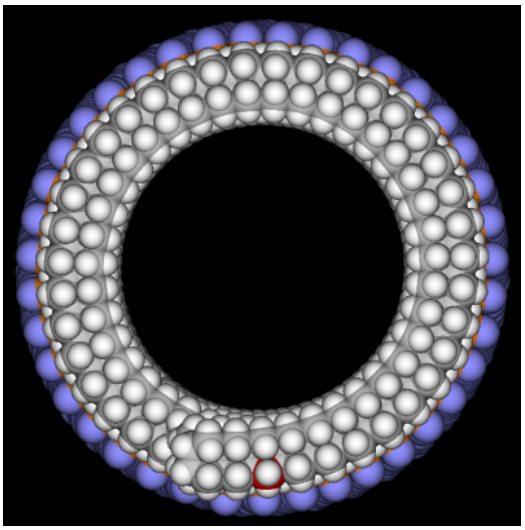


Figure 145: Top view of a lonsdaleite threaded rod with CN threads. From James Ryley's unpublished report [Ryl].

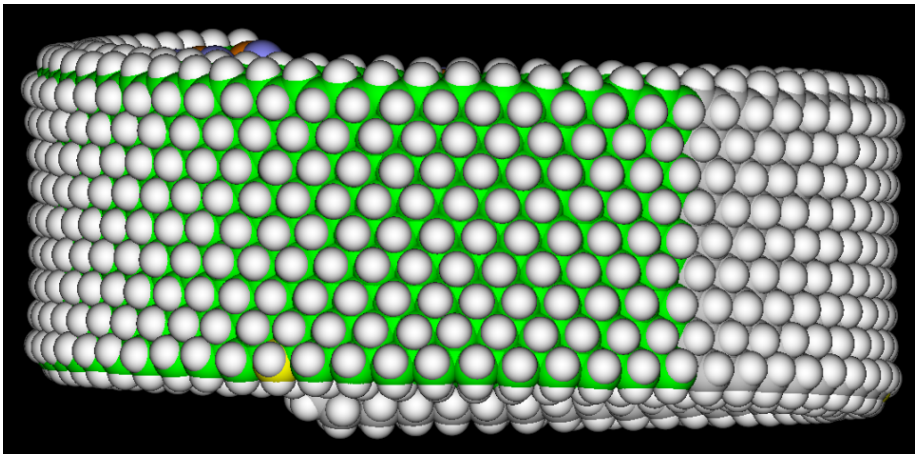


Figure 146: Side view of an internally threaded outer sleeve. From James Ryley's unpublished report [Ryl].

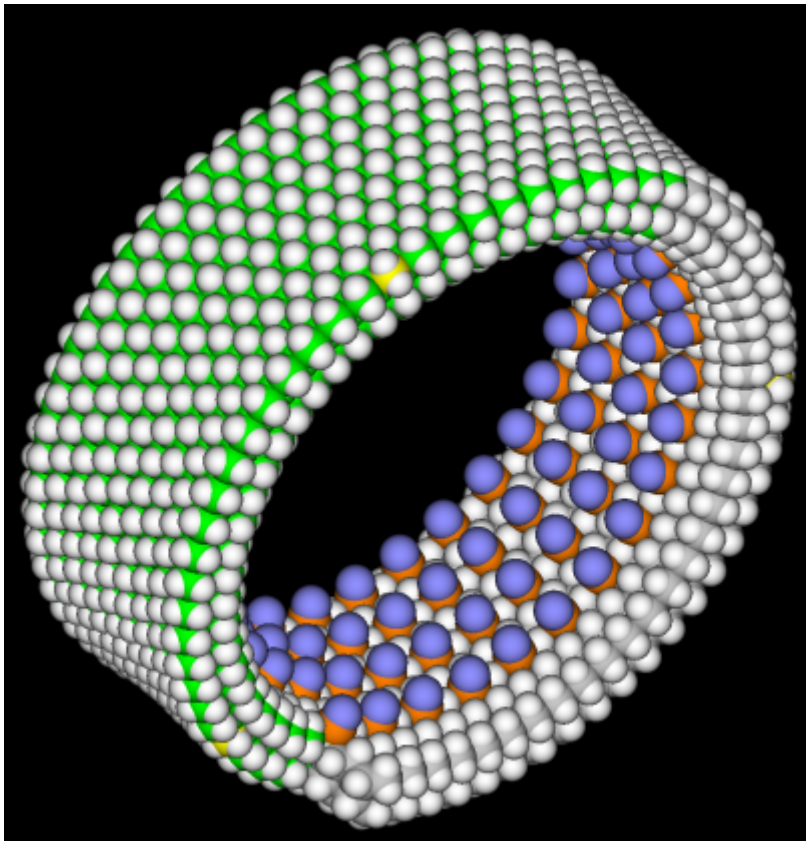


Figure 147: Oblique view of an internally threaded outer sleeve. From James Ryley's unpublished report [Ryl].

7.19.3.3 Assembled screw drive

The assembled mechanism is shown in side and top views (Figures 148 and 149).

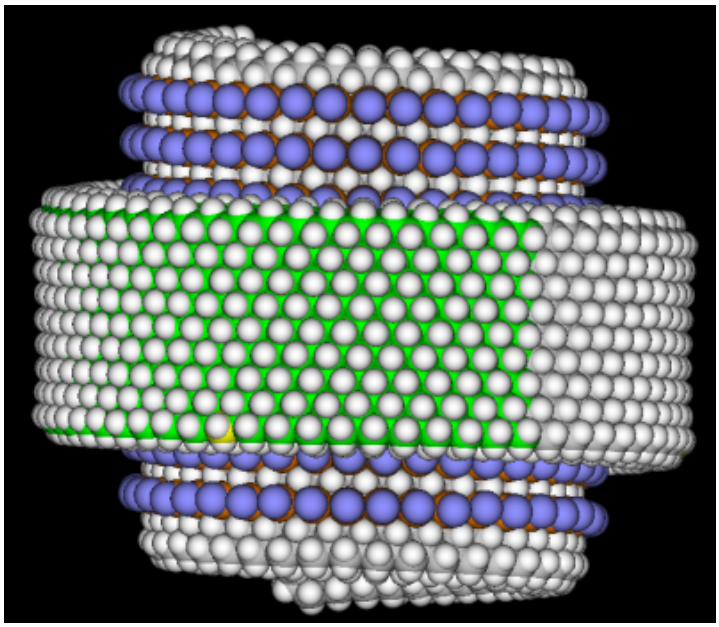


Figure 148: Side view of the assembled Ryley screw drive. From James Ryley's unpublished report [Ryl].

The diameter at the midpoint of the thread interface is about 42 Å, making the circumference about 132 Å (Figure 149). When looked at as an inclined plane, given a pitch of about 4.54 Å, the rise/run is about 1/29.

7.19.3.4 Incommensurate surfaces

Before the axial offset that forms the helix is introduced, the inner and outer hoops have 39 and 63 periods around their circumferences, respectively. These counts share a factor of three, so their least common multiple (LCM) is 819, rather than their product of 2457. In the modified outer sleeve, every other inward-facing projection was removed along the continuous helix. The resulting thread pattern can be described as a 31-1; the alternation follows the helix rather than restarting on each turn. The counts 39 and 31 share no common factor, so their LCM is their product, 1209. However, these repeat counts alone do not establish how easily the rod and sleeve slide relative to each other. The lengths of the threads and the extent of their overlap must also be considered.

Too short and you do not get the nice averaged out PES (Potential Energy Surface) of the surfaces over the LCM. Too long and you start repeating (and since they would exactly add up, partial repeats might make the surfaces more commensurate). Since the threaded sleeve is only about 4 threads wide, this raises the question of how 4 overlaps of a 39-1 and a 31-1 behave.

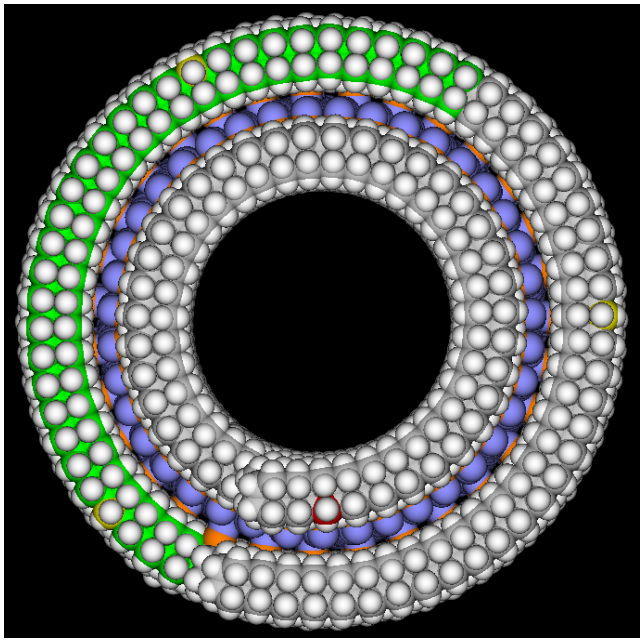


Figure 149: Top view of the assembled Ryley screw drive. From James Ryley's unpublished report [Ryl].

8 Die issues

8.1 SRU-Die Interactions

Introduction

The SRU (Simple Replicating Unit) is supported on a die. The baseline die is a conventional lithographically produced semiconductor die. E-beam lithography on smaller wafers at a custom fabrication facility might be an attractive option. The SRU is an atomically precise box that draws power and control signals from the die. System designs might involve anywhere from one die in early research to the 72,000 dies in the baseline system. Each die would be immersed in the baseline feedstock solution at a temperature of 210 K.

Size of an SRU

The baseline SRU is 500 nm long and 100 nm wide. The height is 60 nm. The gap separating two SRUs is 100 nm. An SRU site is 600 nm × 200 nm. The baseline SRU uses six spikes driven into six contact pillars. See Contact pillars (Section 4.3.1) for more information about the spikes and contact pillars. Later generation SRUs might have as many more electrical connections as might be accommodated given the limitations for a device the size of an SRU and might use different contact technologies. The baseline design uses contact pillars for ground, V+, and four 100 MHz data lines. The baseline V+ voltage is 0.19 V. Operating frequencies well into the gigahertz range might be used in later generations. We target a maximum actuation power draw in the range of one nanowatt (10^{-9} watts) per SRU. SRUs are arranged in a monolayer on the surface of the die and would not extend into 3D in early generations.

8.1.1 External electrical connections to the die

We assume that ground is carried across the die by a ground plane, and that V+ is carried across the die by its own plane.

We assume the four control signals, S1, S2, D1 and D2, are carried by four fractal H trees (see H-trees for delivering power to the differential comb drives (Section 7.16)).

The six contact pillars at each SRU site above the surface of the die connect to the fine 40 nm wide wires, the ground plane, and the V+ plane by vias.

We provide connections through the die using through-silicon vias (TSVs). These form a 16 × 16 grid with 0.614 mm spacing. This would produce 16 × 16 = 256 TSVs. This reduces resistance to facilitate rapid charging and discharging of the wires connected to the SRU differential comb drives.

8.1.2 Feet

A single SRU has 3 “feet”. Each foot is narrower than the gap between two contact pillars on the surface of the die. Each foot could therefore fit between two contact pillars. Each foot would be atomically precise. Atomically precise machinery in each foot extends a spike from each side of the foot into the two contact pillars between which the foot was placed. This allows the spikes to make good electrical connections with the two contact pillars. As the baseline SRU requires six connections, a baseline SRU has three feet, each foot making two connections.

The gaps between paired contact pillars, the pillar heights, the relative positions of the pillars, and the variation in these dimensions are critical if we are to ensure that the SRUs, which are

atomically precise, can make proper contact with the non-atomically precise surface of the die at the SRU sites.

8.1.3 The first SRU

After the first SRU is placed on the die, it then follows instructions transmitted through the contact pillars, draws on feedstock, and manufactures and extrudes a daughter SRU. The External Assembly Arm on the parent SRU then positions the daughter SRU on the die adjacent to the parent SRU. This requires that the parent be able to position the daughter with respect to the SRU site chosen for the daughter. The parent is positioned on its own contact pillars, so the parent knows its own position. It positions the daughter properly on its adjacent contact pillars. When the daughter drives the spikes home into the contact pillars, each foot self-centers between them. This allows the daughter to correct any slight drift that might have occurred in the parent's position.

The feet of the daughter, and hence the daughter, must be positioned with sufficient accuracy by the parent that the daughter's feet enter between their contact pillars on the die surface.

8.1.4 Subsequent SRUs

After the first few replication cycles, daughter SRUs might find that no site adjacent to their manufacturing parent is open. In this case, the daughter SRU would be passed hand-over-hand, from parent to parent, until it reached the edge of the occupied SRU sites. At this point, the parent SRU holding the daughter SRU would place the daughter SRU into the unoccupied SRU site.

When a "daughter" is placed into a site on the die surface and electrically connected, it becomes a "parent".

By repeatedly passing daughter SRUs hand-over-hand until they encounter an unoccupied SRU site on the die surface and then placing the daughter SRU into the unoccupied SRU site, the die surface becomes covered with SRUs. Once completely covered, broadcast instructions would change. Instead of manufacturing more SRUs, the instructions would cause the SRUs to manufacture some useful atomically precise product.

8.1.5 Questions

The general picture is of a large number of small atomically precise devices (boxes) sitting on a conventional silicon die. A device has a footprint of 100 nm by 500 nm and is 60 nm high. It might draw in the range of 10^{-9} W (one nW), at 0.19 V, giving a current draw per device of as much as 5.3 nA. Each die is immersed in the baseline feedstock solution at a temperature of 210 K. The device will be in electrical contact with the six contact pillars on its site. The contact pillars will provide ground, V+, and four 100 MHz control signals. The baseline method of connecting to the contact pillars is through three "feet", with each foot resting between two contact pillars and having two separate connections, one to each of the two contact pillars that it's between.

This architecture raises several issues with respect to the manufacture of the system.

1) E-beam lithography would seem better for prototyping and might be acceptable in cost for production. What facilities for manufacturing wafers are available, at what cost, and in what volume? What about converting wafers into die? Early volumes in the tens of thousands of die might be plausible.

2) The design of the dies. The baseline is to use a standard H tree layout (see H-trees for delivering power to the differential comb drives (Section 7.16)) with buffer drivers as might be needed. In essence, we're copying standard clock distribution methods to distribute the broadcast control signals. Backside delivery of ground, V+, S1, S2, D1 and D2 through TSVs seems likely.

See Contact pillars (Section 4.3.1) for further discussion of this layout.

SRUs are mounted on pre-defined sites built using standard semiconductor manufacturing methods. Each site has six "contact pillars" arranged in three pairs. One contact pillar is ground and one is V+. The other four contact pillars carry the four data signals. We assume ground is carried by a ground plane and V+ is carried by its own plane. The first high level question is: does the baseline proposal work? The second high level question is: might there be a better way?

3) Quantitatively, what sort of bandwidth might we support to each device? We're targeting 100 MHz. Further research is needed to determine how high a frequency can be achieved and what frequency would be a desirable target in future generations.

4) What about ringing and noise? Crosstalk? We want very low error rates. We would like to operate with a good probability of zero errors for 24 hours (that is, we would like a single device to operate for a 24-hour period with no errors). As our overall allowable SRU failure rate is ∼0.0864% per day, the chance of a single error caused by electronic transmission of data during that 24-hour period must consume only part of that overall failure budget. This should be achievable.

5) How do we connect from an external computer to the die(s) at ∼210 K reliably? TSVs from the backside?

6) At what temperature do the electrons in silicon freeze out and semiconductor operations become either unreliable or flatly infeasible? This should not be a problem at our baseline temperature of 210 K, but during early research work or if we adopt lower temperatures for risk mitigation this might become an issue.

7) Maintenance operations might involve thermal cycling, which is well known to cause problems. Is there any experience with these issues that we should be aware of? Can we mitigate this risk? Is thermal cycling from room temperature to 210 K problematic? Lower temperatures?

8) Who might provide the kind of facility described here? Desiderata include rapid prototyping, low volume, and reasonable cost.

9) Whether the first SRU is attached in liquid or vacuum remains open. If attachment is in vacuum, positioning might involve macroscopic nano-positioning and high-resolution imaging (possibly SEM). What equipment might satisfy this requirement? The stability of the SRU under SEM imaging has not been evaluated.

10) Subsequent devices can be positioned with respect to earlier devices or positioned with respect to positional markers placed on the die surface. What lithographic fiducial markers might be appropriate? What fiducial markers have been used previously?

11) Can the baseline nanotube contact spikes make reliable electrical contact with the metal pillars? What modeling software is available for these nanotube-to-metal contacts? What can existing nanotube-to-silicon or nanotube-to-metal contact models contribute?

12) There has been some thought of providing feedstock through microfluidics in the die. This might involve a thin film (tens of nanometers) over a channel or reservoir in the silicon. The thin film might possibly be made of silicon nitride (SiN), silicon dioxide (SiO_2), or various metal thin films. An atomically precise needle would then punch a hole in the thin film to access the contents of the reservoir. Who might manufacture such thin films on silicon? What software is available for modeling the behavior of such thin films? Who has expertise in fluidics in silicon dies?

9 Bigger than an SRU

Once we can populate multiple dies with SRUs, we will want to build products that are bigger than an SRU. Perhaps the first product we'll want to address will be medical nanorobots, both because they should be able to save lives and because they should be very profitable. "In 2021 alone, global spending on health reached a staggering $9.8 trillion, accounting for 10.3% of global GDP. Yet, over the last decade, life expectancy has stagnated in many countries, including the United States (US), which alone spends over $4 trillion annually on health." [Wor24]

The manufacture of medical nanorobots focuses on products that are ∼2 µm in linear dimensions.

Once we can manufacture such products, we'll want to reduce their manufacturing costs. In essence, this means more SRUs that are more energy efficient in a smaller volume, hence the discussion of moving from 2D SRUs on a die to 3D SRUs in a space frame.

Finally, we'll want to build larger products. We might start this exploration with products that are 10 mm in linear dimensions, perhaps by extruding 10 mm product cubes from each 1 cm × 1 cm die on a board. To do this in a practical time frame would require a higher density of SRUs on the die surface.

This section provides only a taste of the many ways that might be used to build atomically precise products that are larger than an SRU. Once we establish a beachhead in this new domain, we'll rapidly expand it.

First, we focus on how we might manufacture products that are 2 µm in X, Y and Z.

9.1 How to build 2 µm products

On the assumption that medical nanorobots are about 2 µm in size, we can approximate the problem of building them by asking how we might build 2 µm × 2 µm × 2 µm products in a rectangular region on the die surface that is 2.0 µm by 2.4 µm in area. A rectangular array of 40 SRUs is laid out in the rectangular region as shown (Figure 150).

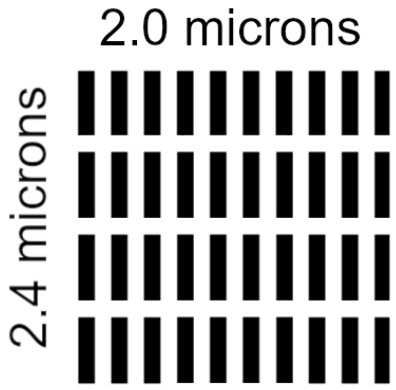


Figure 150: Four-by-ten SRU array for assembly of larger products.

Each SRU is 500 nm × 100 nm, on a surface site that is 600 nm × 200 nm in area. The total site size for 40 SRUs is 2.4 µm × 2.0 µm.

Each SRU exports SRU-sized components which are then assembled in the liquid environment by the external assembly arms on the 40 SRUs into products, such as 2 µm medical nanorobots.

We use a 2 µm atomically precise cube as a volume proxy for a medical nanorobot, assuming the same average atomic density as an SRU. Table 16 summarizes the inputs and results. The component count and manufacturing latency are:

$$N_{\text{components}} = \frac{V_{\text{product}}}{V_{\text{component}}}, \qquad t_{\text{product}} = \frac{N_{\text{components}} t_{\text{component}}}{40},$$

where the volumes are those of the product cube and an SRU-sized component, and $t_{\text{component}}$ is the time for one SRU to manufacture one component. The product atom count follows from 10^8 atoms per component; dividing the nanofactory's atom throughput by this count gives its product throughput.

Table 16: Inputs and estimated results for manufacturing a 2 µm product cube.

Input	**Value**
SRU-sized component dimensions	$500 \times 100 \times 60$ nm
Component volume	3×10^6 nm^3
Product cube volume	8×10^9 nm^3
Atoms per component	10^8
SRUs per manufacturing site	40
Manufacturing time per component	10^5 s
Nanofactory atom throughput	5×10^{21} atoms/day
Result	**Estimate**
Components per product	∼2,667
Atoms per product	2.7×10^{11}
Manufacturing latency	6.67×10^6 s, roughly 80 days
Nanofactory product throughput	1.9×10^{10} products/day

The baseline nanofactory should produce $\sim 1.5 \times 10^{12}$ (1.5 trillion) medical nanorobots every 80 days, each medical nanorobot being manufactured during the full 80 day period on its own 2.4 µm × 2.0 µm manufacturing site by the coordinated efforts of 40 SRUs. Depending on the mission, these 1.5 trillion medical nanorobots might constitute anywhere from one dose to 100 doses, where a single dose might be injected into a person for a single specific purpose. A mapping mission might use fewer medical nanorobots, possibly allowing 100 doses. A more active mission might be an injection of respirocytes. This might use the entire 1.5 trillion medical nanorobots or even more for a single dose.

The baseline method of exporting finished 2-µm products is to release them into the surrounding fluid. Fluid flow would then carry them out of the product manufacturing area and to some suitable product collection and processing area. It would also be possible to take advantage of the propensity of product objects to sink or float: products that sink could be allowed to settle to the bottom of a collection area, while products that float could be skimmed from the top.

9.1.1 Assembly in vacuum

If assembly requires vacuum rather than liquid feedstock, for example to use reactive surfaces, the 40 SRUs could first build a product-sized enclosure sealed around their upper surfaces. After

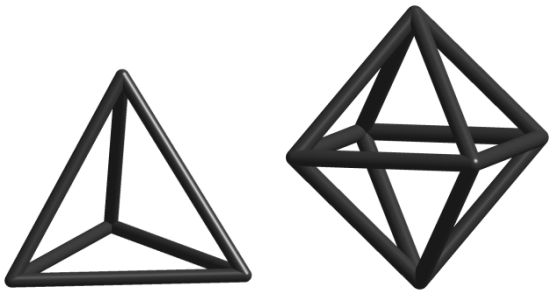
pump-down, they would manufacture components and assemble the product inside the enclosure; see Vacuum Assembly in a Product Assembly Chamber (PAC) (Section 13.3.3).

9.2 From 2D SRUs to 3D Space Frames

9.2.1 Introduction to space frames

This section is adapted from [Mer14].

A space frame is an extremely lightweight, strong, rigid, three dimensional structure that can have an arbitrary shape. It relies heavily on the use of triangles, a structural element which offers great stiffness at low weight. The simplest three dimensional structures that can be made of triangles are the tetrahedron and octahedron (Figure 151).



(a) (b)

Figure 151: Tetrahedron and octahedron: (a) tetrahedron. (b) octahedron.

If we combine these two geometrical forms in a uniform fashion we get a space frame, common in architecture and construction (Figure 152). Space frames were independently invented by Alexander Graham Bell, Buckminster Fuller, and others interested in high strength, light weight, and simplicity [Wik].

9.2.1.1 Struts and Nodes

Space frames are built from struts and nodes. Struts are in pure compression or pure tension and connect nodes. In a uniform space frame, all struts are the same length. Nodes are attached to as many as 12 struts. In a uniform space frame, a node is at the center of a rhombic dodecahedron and the struts are orthogonal to its faces.

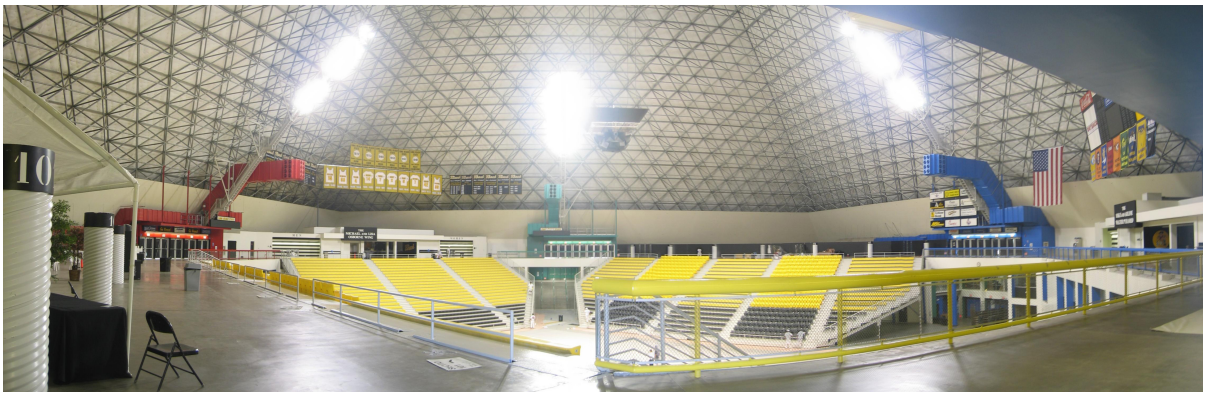
In non-uniform space frames, the length of struts can vary, as can the angles between struts at a node. This results in multiple different types of struts and nodes and can, in the worst case, result in every strut and every node in the finished structure being unique. In a uniform space frame, all struts and all nodes are identical – which greatly simplifies manufacturing, supply chain logistics, and assembly operations.

A node can be thought of as a sphere in the center of a cube with twelve short struts connecting to the centers of the 12 edges of the cube.

In Figure 153, there are 12 connections, with the 6 colored red in a plane, and the 6 colored green forming two tripods, one inverted. In this view, a space frame can be thought of as a stacked set of planes, each plane supported from below by tripods and supporting another plane above on tripods. There are 4 sets of stacked planes, each set stacked in parallel to one of the 4 faces of a



(a)



(b)

Figure 152: Exterior and interior views of the Pyramid space frame: (a) exterior of the Pyramid at California State University, Long Beach. (b) interior panorama showing the roof space frame. The Pyramid is the largest space frame in North America. Photos by Buchanan-Hermit: *Csulb-pyr1* and *Csulb12-pano*, reproduced unmodified under the attribution permissions on their source pages.

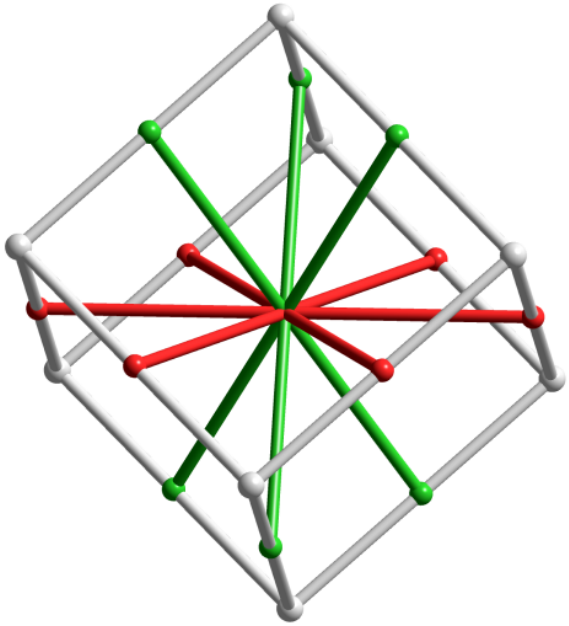


Figure 153: Tripodal view of a space-frame node.

tetrahedron, the 4 sets of stacked planes interlocking to form the space frame. By inspection and symmetry, we can see that any two red half-struts in the cube are at 60°, 120° or 180° angles to each other.

Another perspective on a space frame is illustrated in Figure 154, in which the 12 connections are divided into 4 colored red in a plane, and 8 colored green that form two four-legged tetrapods, one inverted. In this case, the space frame can be thought of as a stacked set of planes, each plane supported from below by tetrapods and supporting another plane above on inverted tetrapods. There are 3 sets of stacked planes, each set of planes stacked in parallel to one of the faces of the cube, the 3 sets of stacked planes interlocking to form the space frame. By inspection and symmetry, we can see that any two red half-struts in the cube are at 90° or 180° angles to each other. We conclude that the angle between any two struts or half-struts connected to a node must be 60°, 90°, 120° or 180°.

Note that the cubes in this space frame occupy 3 dimensions much as the red squares occupy a red-and-black checkerboard in 2 dimensions – if you move left, right, forward, backward, up or down from a cube in the space frame, you find an empty cube-sized hole. There are also eight body-diagonal directions, which we might designate (1,1,1), (1,1,-1), (1,-1,1), (1,-1,-1), (-1,1,1), (-1,1,-1), (-1,-1,1) and (-1,-1,-1). We arbitrarily designate two members of this set of eight directions as D1 and D2. If we are at any cube, we will find that by moving in direction D1 and then moving in direction D2, we will again be at some cube in this space frame, for any directions D1 and D2. If we just move in direction D1, we will be in a cube-sized hole.

Note also that the frame of the cube illustrated in Figures 153 and 154, outlined in gray, is not part of the physical space frame. It is included in the illustration solely to help the mind understand the nature and symmetries of the system. The 12 red and green half-struts shown connecting the node at the center of the gray reference cube to its edges constitute a single "monomer"; when many copies of it "crystallize," they form a space frame. Each half-strut would bond linearly to another half-strut on another monomer to form a whole strut that would join the two nodes from the two

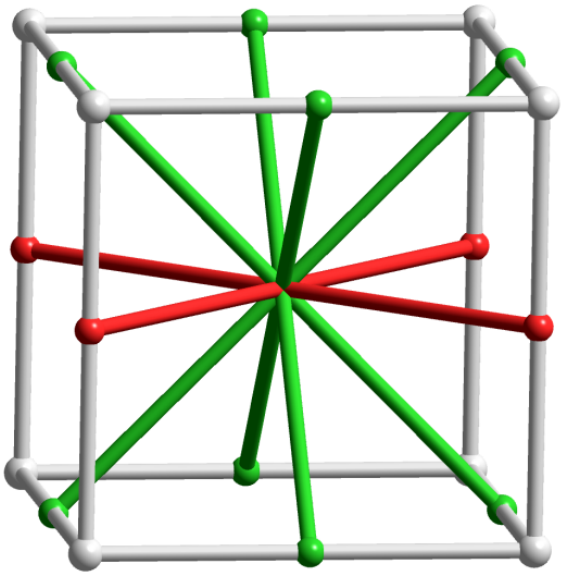


Figure 154: Tetrapodal view of a space-frame node.

monomers.

A unit cell of a space frame is illustrated in Figure 155. It is composed of 4 space frame nodes alternately occupying 4 of the 8 cubes into which the unit cell can be divided.

Consistent with the above discussion, Figure 156 illustrates a segment of space frame showing the tripods and inverted tripods between two planes with triangular facets, while Figure 157 illustrates a segment of space frame showing the tetrapods and inverted tetrapods between two planes with square facets. Note that the nodes are identical in both illustrations. The difference is solely in the lattice plane we are exposing on the surface. The square two-plane space frame motif is typical in many open roof applications in architecture, for example, the sheltering cover between buildings in a shopping center or in front of a theater.

In a sufficiently large structure (that is, if we neglect the surface of the structure and consider only the bulk) the ratio of struts to nodes is 6 to 1, because every strut connects 2 nodes, and every node is surrounded by 12 struts, each of which is shared with another node. To put it another way, our finished structure could be made from a single "monomer": a node pre-attached to 12 half-struts. A space frame is formed when these "monomers" form a three-dimensional "polymer" in which the half-struts bond to each other.

As it requires only 3 coordinates to specify the location of a point (neglecting singularities), the space frame, which "specifies" the location of a node using 12 incident struts, has considerable redundancy in constraining each node. The shared inventory count is six struts per node because each strut joins two nodes. Put another way, the space frame architecture can distribute a load applied to any given node among multiple struts and nodes, and can degrade gracefully in the presence of weakness or even failure of individual struts or nodes. On the other hand, by eliminating superfluous struts and nodes, a space frame can be tailored to provide a three dimensional scaffolding of the desired shape with just the strength and stiffness that are required.

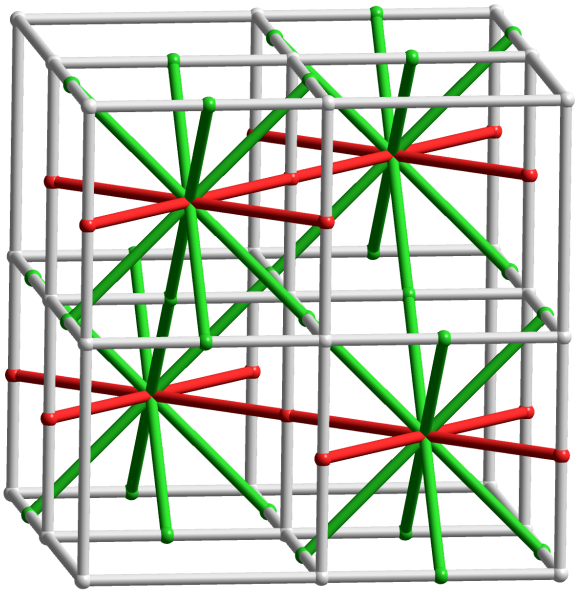


Figure 155: Space-frame unit cell containing four nodes and a tetrahedron.

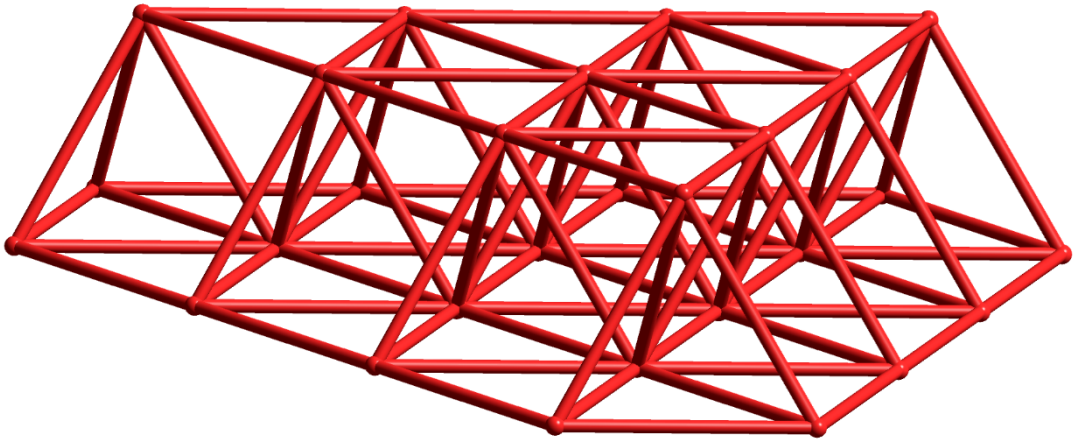


Figure 156: Tripodal view of a space frame with multiple nodes.

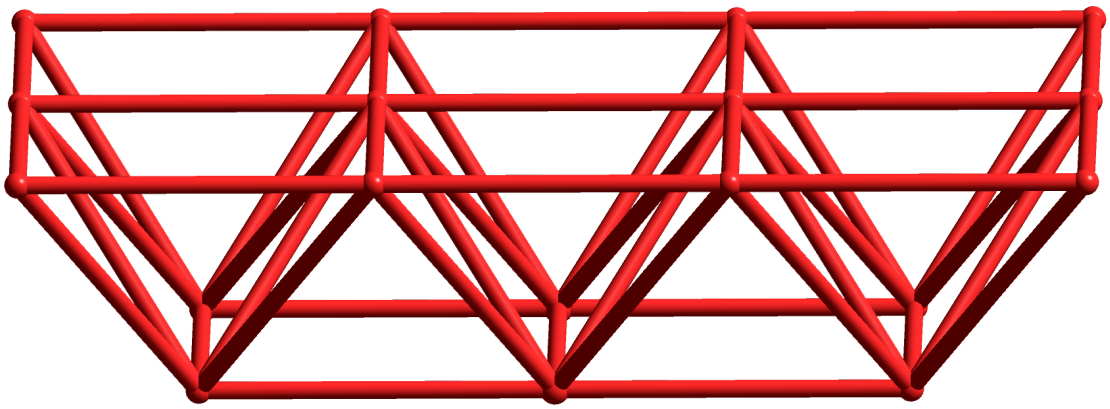


Figure 157: Tetrapodal view of a space frame with multiple nodes.

9.2.2 Transitioning from surface SRUs to space frames

"SRUs on a surface" has many advantages as a design motif for a nanofactory that we want to build from existing technology. On the other hand, if we want to increase the manufacturing rate by increasing the density of SRUs per unit volume, then space frames become advantageous as a design motif for future nanofactories. This raises the question of how to transition from earlier surface-based nanofactories to later space frame based nanofactories, which we address in this section.

Figures 158–163 show the construction of a space frame from an orientation showing it as layers of rectangular arrays of nodes connected by struts, each layer offset from the layer above and below, and the nodes in each layer connected to the nodes in the layer above and below by diagonal struts. This perspective emphasizes the rectangular (rather than triangular) nature of a space frame, which is helpful if we wish to build a space frame starting with a rectangular array of SRUs on a surface.

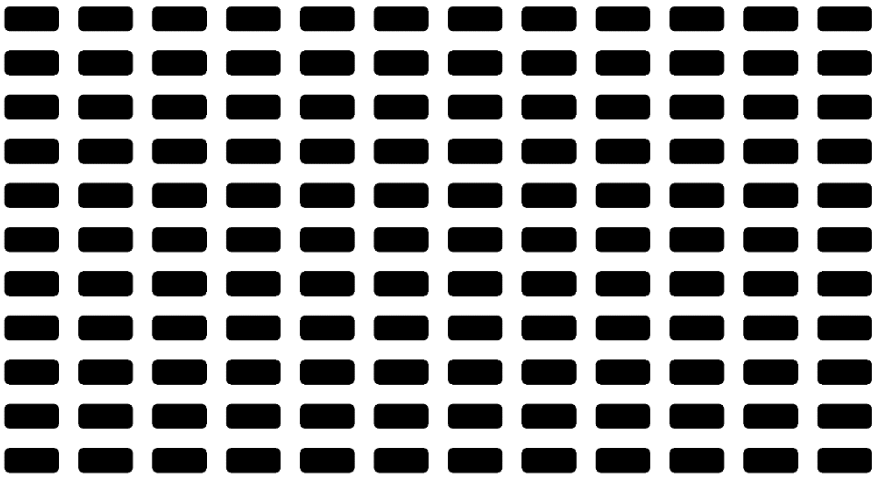
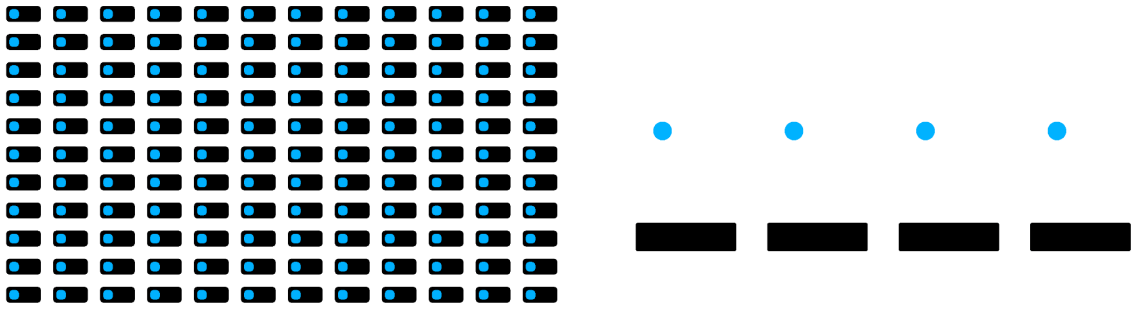


Figure 158: Rectangular SRU array used as the starting surface for space-frame assembly.

Figure 158 shows an array of SRUs viewed from above. Notice that this is a rectangular array,

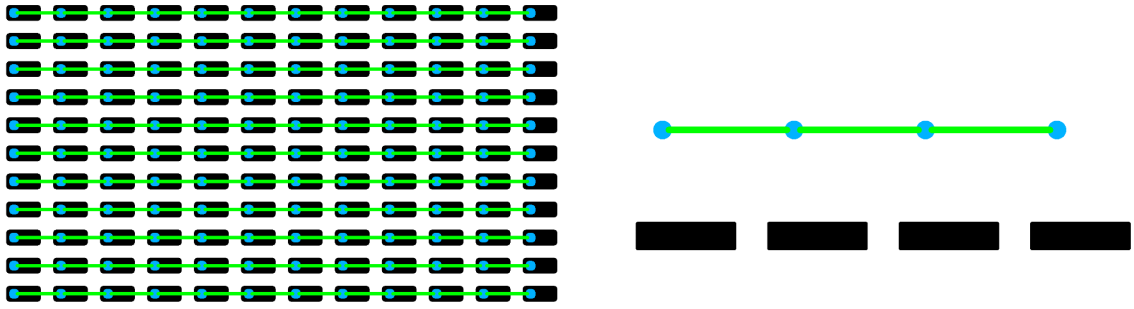
hence building a rectangular face of a space frame is relatively easy.



(a) (b)

Figure 159: Space-frame nodes positioned above an SRU array: (a) plan view; (b) side view.

Figure 159 shows an array of SRUs with nodes positioned above them. On the left, the view from above. On the right, a view from the side. The nodes would be held in position by assembly arms attached to the surface of the SRUs. The assembly arms are not shown in this simple illustration.



(a) (b)

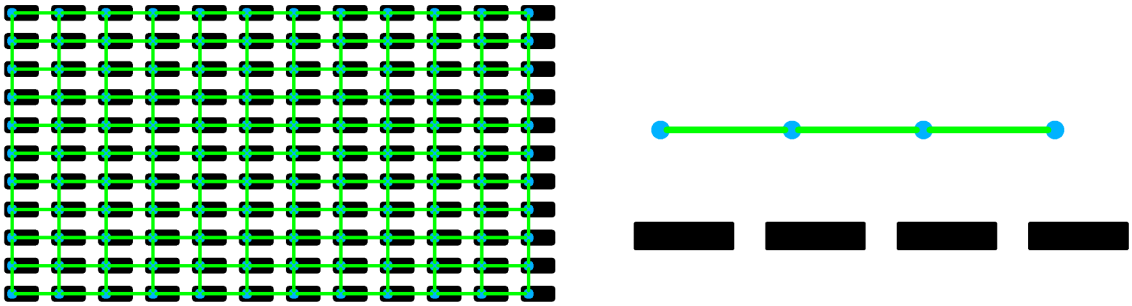
Figure 160: Space-frame nodes joined by horizontal struts: (a) plan view; (b) side view.

Figure 160 shows an array of SRUs with nodes connected by horizontal struts positioned above them. Both a view from above and a view from the side are shown. The green struts would be positioned by a second set of assembly arms and aligned with and attached to the blue nodes. A variation on a dovetail joint should be sufficient for this attachment.

Figure 161 shows an array of SRUs with nodes connected by horizontal and vertical struts positioned above them. On the left, the view from above. On the right, a view from the side. Again, the vertical struts would be positioned by a set of assembly arms and aligned with and attached to the blue nodes, and again a variation on a dovetail joint should be sufficient for this attachment.

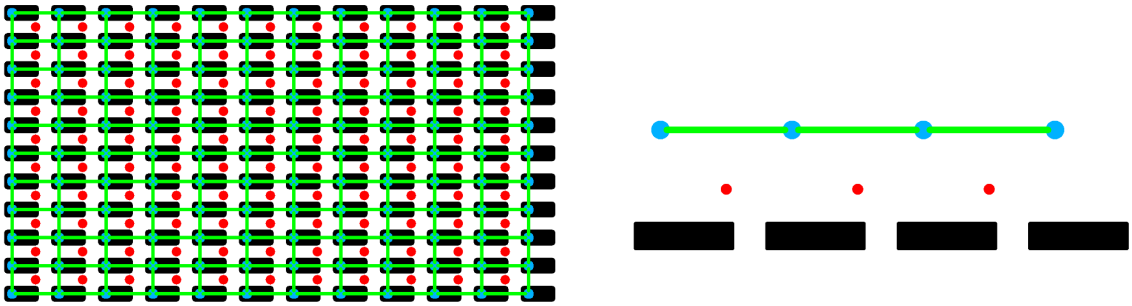
Figure 162 shows an array of SRUs with nodes connected by horizontal and vertical struts positioned above them, and additional nodes (red) centered between the blue nodes. On the left, the view from above. On the right, a view from the side. Again, the red nodes would be held in place during the assembly process by assembly arms attached to the surfaces of the SRUs.

Figure 163 shows an array of SRUs with nodes connected by horizontal and vertical struts positioned above them, and additional nodes (red) centered between the blue nodes, and diagonal struts (yellow) between the red nodes and the adjacent blue nodes. On the left, the view from above.



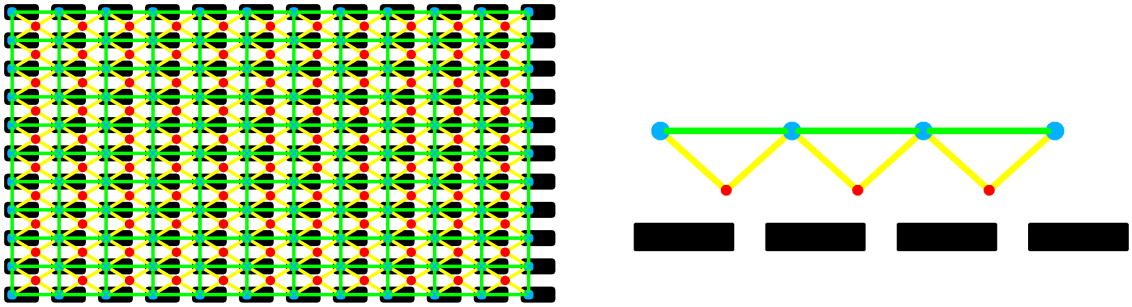
(a) (b)

Figure 161: Space-frame nodes joined by horizontal and vertical struts: (a) plan view; (b) side view, unchanged from Figure 160(b).



(a) (b)

Figure 162: Additional nodes (red) positioned between the original nodes (blue): (a) plan view; (b) side view.



(a) (b)

Figure 163: Diagonal struts (yellow) joining the additional nodes to the original nodes: (a) plan view; (b) side view.

On the right, a view from the side. The diagonal struts (yellow) would be positioned and attached between the blue and red nodes by assembly arms mounted on the SRUs (black).

An additional set of horizontal and vertical green struts would then be added between the red nodes, and the entire space frame layer would then be pushed upwards and the entire process repeated as many times as desired to build as many layers of space frame as might be needed.

While we have depicted a space frame in which the rectangular dimensions of the space frame match the 600 nm × 200 nm site dimensions, we might prefer to have a space frame whose rectangular dimensions are 600 nm × 600 nm, thereby covering three SRUs with one unit cell of the space frame. This has the advantage that three external assembly arms would be available during the construction of a unit cell, simplifying the handling of struts and nodes during the construction of the space frame.

9.2.3 Concepts for building larger structures

Products larger than 2 µm in size are not the focus of this paper. However, once we have built the baseline nanofactory, there is a wide range of possible methods for increasing product size. We might start by using some of the methods already discussed in the published literature [Dre92, FM04b, Mer97].

Once we have dies populated with SRUs, methods that start from this existing capability might be used. For example, we might extrude a 3D solid product as large in cross section as the 1 cm^2 die from which it is extruded by building it up layer by layer from just above the die surface. In this approach, the 8.05×10^8 SRUs on the die surface all contribute by adding SRU-sized parts to the whole being built.

The major problem with this approach is that each SRU site adds only one SRU-sized product object to the growing product layer every 10^5 seconds. If we adopt the baseline packing density of SRUs on the surface, every 600 nm × 200 nm surface site will contribute only one SRU-sized product object every replication cycle. On average, one SRU-sized product object contributes 500 nm × 100 nm × 60 nm of product volume to a surface area of 600 nm × 200 nm, or about 25 nm per replication cycle or ∼22 nm/day. By way of comparison, fingernails grow almost 5,000 times faster at ∼100,000 nm per day (∼3 mm/month).

To boost this growth rate, we might stack SRUs in 3D above the die surface. That is, the product that we extrude would be a space frame with embedded SRUs. If we extrude 1,024 layers with embedded SRUs, the product could grow at ∼22 µm/day, about one fifth that of a fingernail. If we design our growth properly it should take us ∼12 days to grow 1,024 layers, as outlined in the next few paragraphs.

Each layer, floating in feedstock and connected by electrical umbilical cords and some easily changed mechanical links to the layers above and below, would be designed so there is enough space between layers to build a daughter layer, also connected by umbilical cords. To do this, the umbilical cords must be twice as long as the distance between the layers so that there is room to extrude the daughter SRUs from which the daughter layer is composed. Once each daughter has been extruded and the new umbilical connected between the parent and the daughter, as well as the second new umbilical between the daughter and the parent SRU in the next layer above, the old umbilical between the two parents can be disconnected. The new layers are now parent-daughter-parent-daughter-parent-daughter… and the new umbilicals can be extended to their full length by adjusting the mechanical links that connect the layers, giving enough room for another layer between every existing layer.

In this way, 1 layer becomes 2 layers becomes 4 layers becomes 8 layers ... becomes 512 layers becomes 1,024 layers. As each doubling of the number of layers takes 10^5 seconds, in 10^6 seconds (∼12 days) we have 1,024 layers.

If we have 1,024 layers of SRUs, and if each layer passes SRU-sized products up the stack of layers to the top layer, and the SRUs on the top layer then add these SRU-sized products to the growing end of the one cm^2 product object, then it will grow at not 22 nm per day, but 22 µm per day, or 660 µm per month.

At this point we might ask: what limits our ability to carry out the exponential growth of these one cm^2 layers? There are three immediate answers: broadcast signal degradation, feedstock, and heating.

Broadcast signal degradation can be solved by using repeaters. These receive the broadcast, restore it to its original form, and send it on after some processing delay.

Feedstock provided in solution must flow at some sufficiently fast rate to provide sufficient feedstock for manufacturing purposes. This rate is typically low enough that it will not be the limiting factor.

Heating and the associated raw cost of energy are likely to be the first and most significant limiting problems in scaling of this approach. Actuation alone would dissipate approximately 824 W/cm^2. Applying the present baseline ratio between total switching heat and actuation power suggests approximately 2.2 kW/cm^2, before accounting for additional repeaters and interlayer wiring. The precise result depends on the three-dimensional power-distribution architecture.

It's useful to remember that each "layer" has as many SRUs as an entire die as that term is used in the earlier sections of this paper. One of the reasons for adopting a design in which a 1 cm^2 die supports only a single layer of SRUs is that this approach keeps heating to an actuation load of approximately 0.8 W per die (a present first-order total switching load of approximately 2.18 W per die), which is quite manageable.

The actual energy efficiencies of the design proposals advanced herein remain to be seen. Once we know the actual average power requirements of an SRU, we can consider how many layers might be reasonable. We can reasonably expect rapid improvements in later systems. This new technology is likely to see exponentially improving energy efficiency for some time, just as computing systems have seen decades of exponentially improving energy and manufacturing efficiency, with a doubling time ranging from every 1.57 years to, more recently, every 2.29 years [PPEL25].

As average power consumption goes down, the cost of manufacturing also goes down. The theoretical energy requirements for molecular mechanical computation are remarkably small [MFH+18, MFR+19]. This suggests that nanofactories can be designed that dissipate many orders of magnitude less energy than the crude 1 nW actuation power budget assumed by our baseline proposal for an SRU.

Potatoes illustrate low-cost biological production using atomically precise cellular machinery, with product costs well under $10 per kilogram. To scale up nanofactory production to the volumes we see in agricultural production, we need energy efficiencies that at least approach the energy efficiency seen in plants. We should be able to do as well as plants, or better, once the technology has matured.

9.3 Building more product

At least a few paths forward for building more and bigger products once the single-layer die can be made to work seem clear.

1) Use contact printing of dies to rapidly populate additional dies.

2) Stack SRUs on a die in 3D on top of each other as deeply as we can to increase productivity per die, limited primarily by energy efficiency (heat).

3) Incrementally (and hopefully rapidly) improve the energy efficiency of SRUs.

4) Extrude product from die surfaces.

In other words, the immediate paths forward involve designing nanofactories which have more dies, more productivity per die, and less heat produced per gram of product.

10 Critical Parameters

10.1 Productivity

Productivity of the nanofactory: 100 milligrams of carbon-equivalent product per day. This can also be stated in atoms/day, which can be computed by noting that one mole of carbon atoms is 12 grams or $\sim 6 \times 10^{23}$ atoms (approximately Avogadro's constant, the exact value of which is $6.022\,140\,76 \times 10^{23}$ mol^{-1}). Dividing 12 grams by 120 gives 100 milligrams, while dividing 6×10^{23} atoms by 120 gives 5×10^{21} atoms. Therefore, 100 milligrams/day is also $\sim 5 \times 10^{21}$ atoms/day. This target was chosen because it's enough to be profitable and small enough to be achievable.

10.2 Build Rate

The build rate per SRU fabricator is 1,000 atoms/second. This build rate was chosen because it should be achievable by an early design and because it should be sufficient to achieve the 100 milligrams/day productivity target.

There is one fabricator per SRU, so the build rate per SRU is also 1,000 atoms/second.

10.3 SRU Count

We need to know the total number of SRUs required to achieve the 100 milligrams/day manufacturing rate: 5×10^{21} atoms/day divided by $24 \times 60 \times 60$ seconds in a day gives $\sim 5.8 \times 10^{16}$ atoms/second, which implies (at 1,000 atoms/second) that we need 5.8×10^{13} SRUs (or fabricators) in the baseline nanofactory.

10.4 SRU site size

The site size is 600 nm × 200 nm. We arrive at this figure by starting with the footprint of an SRU on a die and putting 100 nm in X and 100 nm in Y between SRUs. An SRU has a 500 nm × 100 nm footprint on the die surface. Adding our assumed spacing between SRUs gives us an SRU site size of 600 nm × 200 nm.

10.5 Die Size and Count

Three SRU sites with a site size of 600 nm × 200 nm each will occupy a square 0.6 µm on a side. Using H trees starting from this size (see H-trees for delivering power to the differential comb drives (Section 7.16)) gives us squares that have side lengths of $2^{14-k} \times 0.6$ µm for $k = 0, 1, 2, 3, \ldots 14$. For k=0, we have sides for our square that are ~9.8 mm. The area of this square is close to 1 cm^2. A die size of 1 cm^2 is fairly typical in the semiconductor industry for a laptop or desktop CPU. The actual die size will depend on detailed decisions by the group doing the semiconductor die design. Regardless of its actual size, we will continue to refer to it as a 1 cm^2 die. Given the site size of 600 nm × 200 nm, and quantizing to $2^{14} \times 2^{14} \times 3 = 805{,}306{,}368$ SRUs on a die, a nominally 1 cm^2 die will have $\sim 8.05 \times 10^8$ SRUs on it. If our nanofactory requires 5.8×10^{13} SRUs, and each die has $\sim 8.05 \times 10^8$ SRUs, then we need 72,000 dies to accommodate all 5.8×10^{13} SRUs, where each die is nominally 1 cm^2. We therefore assume our nanofactory has 72,000 silicon dies, each die with $\sim 8.05 \times 10^8$ SRUs. Scaling the baseline nanofactory is easy. If you want more product you add more dies to the nanofactory.

Nominal die size: 1 cm^2

SRUs per die: 805,306,368

Number of dies: 72,000

10.6 PC Board Size and Count

Each die is mounted in an IC package, and each package is mounted on a printed circuit board (PCB). A typical reasonably large PCB might be 450 mm × 450 mm. If each PCB holds a 24 × 24 array of ICs, we get 352 mm^2 of PCB area per package, which is compatible with industry practice for a nominal one cm^2 die. If we have 805,306,368 SRUs per IC, and 576 ICs per board, we have $\sim 4.64 \times 10^{11}$ SRUs per board. If we have 125 PC boards, this gives us $125 \times 4.64 \times 10^{11} \approx 5.8 \times 10^{13}$ SRUs, as required for the nanofactory.

PC Board size: 450 mm × 450 mm

Packages per board (single sided): 576

SRUs per board: 4.64×10^{11}

Boards per nanofactory: 125

10.7 SRU energy usage

The simplest way to estimate the energy used by an SRU is to first observe that all actuation power for an SRU comes from four differential comb drives. Therefore, if we know the force and throw of these four differential comb drives we can determine the energy they deliver per throw from the classic equation $W = Fd_t$, where d_t is the actuator throw.

This estimate neglects the power lost in delivering electrical power to the four differential comb drives. Electrical losses when square waves are used to charge the wiring and the comb drives are significant and are not accounted for in this "actuation power per SRU" computation. They are accounted for separately. There are three types of power:

1) The actuation power required by the mechanical systems of the SRU (the "Actuation power budget per SRU")

2) The power lost delivering that actuation power and

3) The power required by the cooling system.

We start by estimating the "Actuation power budget per SRU".

The four differential comb drives are operated in two pairs: two "Select" differential comb drives and two "Drive" differential comb drives. When a pair of differential comb drives is active (selected for clocking), only one of them changes state and is being driven at a time, while the other differential comb drive in the pair is not being driven and hence not drawing power. S1 and S2 are 90° out of phase and 2.5 ns apart in a 10 ns cycle. When S1 is driven, S2 is not driven and vice versa. This means that only one of these four differential comb drives is drawing power during each phase of the four-phase clocking cycle to drive mechanical actuators either forwards or backwards, while the others are not.

This can also be seen by looking at the four-phase clock cycle. If we have two clock signals, A and B, driving two differential comb drives, and both A and B start low, then the clocking sequence

is: A goes high, B goes high, A goes low, B goes low. Power is drawn, and the differential comb drives are actuated, only on a transition. As can be seen, there are four transitions during one full clock cycle of the four-phase clock signal. One full clock cycle at 100 MHz is 10 ns. Therefore, one of the differential comb drives transitions every 2.5 ns.

We therefore multiply the energy required per throw by four times the clock frequency to get the power used by the SRU.

If the force generated by the comb drives is sufficient to actuate the molecular machinery of the SRU, that is, if the differential comb drives do not actually stall because they cannot generate sufficient force, we can neglect any inefficiencies in the molecular machinery of the SRU itself. Once we have an upper bound on the force, and we know the distance traveled, we can compute an upper bound on the energy required per throw of a comb drive.

We also need to estimate the electrical losses in providing the power to the comb drives, such as the resistance of the wires on the die through which the current to the comb drives flows, the energy lost charging and discharging those wires, the resistance through the nanotubes from the wires to the comb drives, and the like. We also need to estimate the power required for cooling. These are estimated in Power required to run the nanofactory (Section 7.16.2).

A 2 nm throw with a force of 1 nN is more than adequate to provide reliable mechanical actuation for a wide range of molecular machines, as shown by proposals for molecular mechanical logic [MFH+18, HMA17, Dre92, MFR+19]. This gives a maximum energy requirement (E_{max}) per actuation of 2×10^{-18} J. We provide a 25% larger energy budget because this particular design requires additional energy to accelerate the shuttle within the available 5 ns between transitions of the same Drive signal (there are 2.5 ns between successive transitions across the two Drive signals). This increases the energy budget per actuation from 2.0×10^{-18} J to 2.5×10^{-18} J. We emphasize that this additional 0.5×10^{-18} J requirement is caused by the specific nature of the differential comb drive design, and is not a general property of comb drives. See Mass of the shuttle (Section 7.15.5.3).

$$E_{\text{max}} = 2.5 \times 10^{-18}\,\text{J}. \tag{80}$$

Here E_{max} is **the maximum energy that one of the four differential comb drives powering an SRU will expend during a single throw**.

This can also be expressed in terms of k_BT at our expected operating temperature of 210 K, where k_B is Boltzmann's constant. As $k_B = 1.380649 \times 10^{-23}$ J K^{-1}, k_BT at 210 K is $\sim 2.9 \times 10^{-21}$ J. Therefore $E_{\text{max}}/(k_BT) \approx 2.5 \times 10^{-18}$ J / 2.9×10^{-21} J ≈ 862. That is, $E_{\text{max}} \approx 862 \times k_BT$. This is a more than adequate margin to ensure reliable operation near the expected operating temperature of 210 K.

At our planned clock frequency of 100 MHz the actuation power will be 100 MHz $\times$ 4 $\times$ E_{max} or $(10^8\,\text{s}^{-1}) \times 4 \times (2.5 \times 10^{-18}\,\text{J}) = 1 \times 10^{-9}\,\text{J/s}$ = one nanowatt per SRU. Again, the factor of 4 is because four phase logic has four clock transitions per cycle, and the differential comb drives draw power during every clock transition.

$$P_{\text{max}} \approx 10^{-9}\,\text{W}. \tag{81}$$

Here P_{max} is **the maximum actuation power for one SRU**.

We expect that the energy required per throw will usually be less than E_{max} (less than 2.5×10^{-18} J) because less force will be required than the full 1 nN. That is, the rising and falling edges of the driving signal need not be vertical but could be sloped. The maximum power draw would only

occur if full force was required for the full stroke. This occurs when the driving signal is a square wave.

The baseline assumes we use square waves and makes no attempt to use resonant energy transfers. It should be possible to substantially improve the efficiency of the baseline proposal.

The **actuation power budget per SRU should be one nanowatt when operating at 100 MHz**. This is sufficient to achieve 100 milligrams/day (see Summary of Critical Parameters (Section 3.3.2)).

10.7.1 Comb drive throw and force: 2 nm and 1 nanonewton

We briefly justify the claim that a 2 nm throw and a 1 nanonewton force will be sufficient for the four differential comb drives that power the SRU in the following discussion. Computational studies of molecular mechanical computer designs suggest that a 2 nm output travel should be sufficient to achieve highly reliable operation, and that substantially less than 1 nN of force should also be sufficient. Applying the thermal uncertainty relation in Eq. (23), in variance form, gives

$$\sigma_x^2 = \frac{k_B T}{k_s}. \tag{82}$$

Here σ_x is the RMS positional uncertainty in the direction of actuation, in meters, k_B is Boltzmann's constant, T is temperature in kelvin, and k_s is stiffness in N/m [Dre92]. Assuming a low k_s of 1 N/m for a molecular machine and T ~210 K, we derive σ_x = 54 pm. A 10 σ_x positional error of 540 pm or 0.54 nm would have a probability of about one chance in 10^{23}. This astronomically low probability suggests that positional errors caused purely by thermal noise in properly designed molecular mechanical systems can be neglected when an actuation distance of 0.54 nm is used. Our use of 2 nm is more than enough to deal with errors caused by thermal noise. Other sources of error can be dealt with (see Hysteresis and backlash (Section 6.9.2) and Friction and wear (Section 7.15.5.14)). We conclude a 2 nm throw should be more than sufficient.

While other links in the actuation chain can be approximated by a classical spring with a linear restoring force, the differential comb drive itself is a constant force spring. That is, once the voltage on the differential comb drive has settled to either ground or V+, the force pushing or pulling the differential comb drive will be constant until it has reached the end of its throw. The energy required to cause the differential comb drive itself to be 0.2 nm out of position is Fd, with displacement $d = 0.2\,\text{nm}$: 1 nN × 0.2 nm = 0.2×10^{-18} J. The probability of such an error from thermal noise would be $e^{-\text{energy}/(k_B T)} = e^{-((0.2\times10^{-18}\,\text{J})/(1.38\times10^{-23}\,\text{J/K}\times210\,\text{K}))}$ ~1 in 10^{30}. Again, 2 nm is more than sufficient.

The 2 nm throw comes more from issues of reliable mechanical actuation than from issues of thermal noise. Reliably engaging and disengaging with a tooth on a gear is easier if the tooth is not the smallest possible and the clearances are not the minimum conceivable. Larger clearances can also reduce unwanted van der Waals interactions with nearby surfaces.

We also need sufficient force to effectuate substantial mechanical movements. Six nN is sufficient to break a carbon-carbon single bond. It is reasonable to assume that significantly less than 6 nN will be sufficient force to actuate a wide range of molecular machines, including those in our SRU.

Very low forces are sufficient to actuate well designed atomically precise rotary joints and molecular machines based on such actuation. The rotary drag equation [Mer16] that governs the required force and power dissipation when molecular rotary joints are used at ordinary temperatures is:

$$P = K\,\omega^2 \qquad \text{(rotary drag power equation)}, \tag{83}$$

where:

P is the power dissipated (in W or J/s) by the rotational process.

ω is the rotational speed (in radians/second or rad/s) between the housing and the rotor.

K is the applicable rotary drag coefficient (in J · s or kg m^2/s), with radians treated as dimensionless in the power equation.

Equation (83) is applicable at "ordinary" temperatures. As temperatures become sufficiently low, this equation might significantly overstate energy dissipation, even taking refrigeration into account [Mer93b] – which is a conservative error, as dissipating less energy than we estimate would still result in acceptable performance.

Computational results show that forces in the pN range are sufficient to actuate well designed molecular mechanical logic systems [Mer16]. Computationally, the rotary drag for an atomically precise bearing based on a single bond, specifically the carbon-carbon acetylenic bond, has been estimated. "The acetylenic joint has a rotational frictional drag coefficient of 2×10^{-35} kg m^2 s^{-1}" [HMA17]. The power loss when operating such a rotary joint at 10^8 Hz would therefore be:

$$\begin{aligned} P &= K\omega^2 \\ &\approx (2 \times 10^{-35}\,\text{kg}\,\text{m}^2/\text{s})(2\pi \times 10^8\,\text{s}^{-1})^2 \\ &\approx 7.9 \times 10^{-18}\,\text{W} \qquad \text{(per acetylenic rotary joint)}. \end{aligned} \tag{84}$$

As we are operating at 100 MHz our switching time is 10^{-8} seconds, giving roughly $(10^{-8}\,\text{s})(7.9 \times 10^{-18}\,\text{W}) = 7.9 \times 10^{-26}$ J per single-rotor switching operation. As $W = Fd$ (work equals force times distance d) and we know W and our distance is assumed to be 2 nm, we can compute $F = W/d \approx 7.9 \times 10^{-26}$ J / 2 nm $\approx 4 \times 10^{-17}$ N. Actual logic operations have 10 to 100 rotary joints, so the force required to overcome drag should be increased by a corresponding factor to $\sim 10^{-15}$ N. The potential energy surface is unlikely to be perfectly flat as a function of the rotational position of the rotary joints. If the barriers to rotation are a significant fraction of k_BT (thermal noise at the operating temperature of the system), then some additional force might be required to overcome them. The maximum computed torque in the acetylenic rotary joint is 1.2 pN·nm [HMA17]. For a 2 nm lever arm, this torque corresponds to a force of 0.6 pN. While this is a simple atomically precise rotor and actual forces in molecular machines using acetylenic rotary joints are likely to be higher, it gives a feeling for the kinds of magnitudes we are talking about.

It seems fair to state, given actuation forces that are ~1 pN in the well-studied ideal molecular machine system given above, that a margin of three orders of magnitude to 1 nN should be sufficient to move the molecular machinery of the SRU.

The primary purpose of our 1 nN force is to allow, over several clock cycles, the application of significant force, including bond-breaking forces when that is needed, by the Stewart platform. Each complete clock cycle consists of four throws of D1 and D2, producing one full rotation of the input shaft. Gearing converts this rotation into approximately 15° of strut rotation, extending or contracting the strut by approximately 20 pm, depending on the thread pitch. At an available input force of 1 nN over 2 nm per throw, the four throws supply a total mechanical work of approximately 8 aJ per clock cycle. In a lossless transmission, this would correspond to an average axial strut force of approximately 400 nN over the 20 pm displacement. Gearing and other mechanical losses reduce

the available force. Strut force provides a useful proxy for the force available at the Stewart-platform tip, although the actual tip force depends on the platform geometry.

10.7.2 A global perspective

Almost all clock cycles are devoted to moving the SRU fabricator Stewart platforms by actuating their struts. An axial strut force of 400 nN acting through 20 pm requires 8 aJ per complete clock cycle, matching the nominal mechanical work supplied by four 1 nN, 2 nm input throws. The actuation-energy budget allows 2.5 aJ per throw, or 10 aJ per complete clock cycle, including the allowance for shuttle acceleration (see SRU energy usage (Section 10.7)). These figures describe the input budget and ideal strut output; actual platform-tip force and displacement depend on transmission losses and platform geometry.

10.7.3 Other energy losses

We enumerate here the primary sources of energy loss which impact SRU energy efficiency. These are the links in the chain from the electrical power source to the robotic actuators.

1. Losses in electrical signals traveling through the wires on a die and atomically precise molecular wires from the contact pillar on the die to the differential comb drives.
2. The conversion losses between the electrostatic actuator and mechanical motion. The baseline uses graphene/h-BN differential comb drives; their mechanical losses are estimated in Section 7.15.5.12. Vacuum-separated tines are an alternative, discussed in Section 12.9, with different loss mechanisms.
3. The losses in the demux (Section 4.3.6).
4. Any losses in transporting the signal from the demux to the actuators (losses in the binary Bowden cables).
5. Any losses in the actuators.
6. The losses in the four-phase screw drive that gives us sub-angstrom positioning accuracy (Section 7.18).

- The losses in the demux, Bowden cables, actuators, and screw-drive mechanisms are part of the losses involved in moving the molecular machinery. See Comb drive throw and force: 2 nm and 1 nanonewton (Section 10.7.1).

10.7.4 Power requirements of the baseline nanofactory

If our four differential comb drives have a 2 nm stroke with a maximum of 1 nN force, this provides 2×10^{-18} J per stroke or 8×10^{-18} J per clock cycle. This powers the SRU. We have added a 25% margin to this number, increasing our energy dissipation to 10^{-17} J per clock cycle. This results in a mechanical actuation power at 100 MHz of one nW per SRU (10^{-9} W). As there are 5.8×10^{13} SRUs, this leads to a mechanical actuation power of 58 kW.

The expected actuation power required by the baseline nanofactory is 58 kW.

10.8 Cooling the nanofactory

As the actuation power used by a single SRU is 1 nW (see SRU energy usage (Section 10.7)) and there are 5.8×10^{13} SRUs in a nanofactory, the actuation power usage of a nanofactory $\approx 10^{-9}$ W/SRU $\times$ 5.8×10^{13} SRUs = 58 kW.

However, that actuation power is the power used directly by the SRU and does not include the power lost in charging and discharging the wires leading to the SRU. When that is taken into account (see Power required to run the nanofactory (Section 7.16.2)) the actual heat dissipated by the nanofactory is 157 kW.

We have structured the nanofactory as 125 boards immersed in liquid feedstock. Each board produces ~157 kW / 125 = 1,256 W of heat. As the boards are already immersed in feedstock, the simplest method of removing this heat is to flow feedstock over the boards. The cooling analysis here uses tetraethylgermane's properties for illustration. For the full five-component feedstock (tetraethylgermane, pentanenitrile, triethyl borate, TEOS, $EtCpPtMe_3$), a more detailed heat-capacity and density calculation of the actual mixture composition at 210 K is required to establish final cooling requirements.

Tetraethylgermane has a heat capacity of ~300 J/(mol·K) at 298 K. We might plausibly estimate the heat capacity at 210 K as 250 J/(mol·K). Each board will have to dissipate 1,256 W of heat. If we assume a temperature rise of 10 K, this will require a flow of tetraethylgermane of ~0.50 moles/second. Tetraethylgermane has a room temperature density of $\sim 5.2 \times 10^{-3}$ moles per cc [Gel15]. Using this room temperature density as an approximation at 210 K, a flow of 0.50 moles/second is the same as a flow of 96 cc/second. Therefore, a flow of tetraethylgermane of 96 cc/second over each of our 125 boards should be sufficient to cool boards that each generate 1,256 W, assuming a relatively modest 10 K temperature increase in the flow of tetraethylgermane cooling the board. The total flow over all 125 boards would be 125 $\times$ 96 cc/sec $\approx$ 12 liters per second. The density and heat capacity of the final feedstock mixture at its operating temperature are needed to refine this estimate.

We must then remove the 157 kW absorbed by the circulating feedstock and cool the returning feedstock by the assumed 10 K temperature rise before recirculation.

The COP (Coefficient of Performance) of the cooling system could be at best:

$$\mathrm{COP}_{\mathrm{Carnot}} = \frac{T_c}{T_h - T_c} = \frac{210\,\mathrm{K}}{300\,\mathrm{K} - 210\,\mathrm{K}} \approx 2.33. \tag{85}$$

If we had this theoretically best possible cooling system, the actual power required to cool a 157 kW load would be:

$157\,\mathrm{kW}/2.33 \approx 67\,\mathrm{kW}$.

However, we don't have such a system. A more realistic COP might be 40% of the theoretical best, slightly less than half as good. In this case:

$157\,\mathrm{kW}/(2.33 \times 0.4) \approx 168\,\mathrm{kW}$.

Adding this to the 157 kW used to power the system, we can sum the actuation, electrical, and cooling power together to get 325 kW.

In Ottawa, commercial power might run C$0.17 per kWh. At this rate, the 325 kW load would cost ~C$55/hour to run the nanofactory. This would be C$1,320 per day. Electric power might cost ~C$1,320 per 100 milligrams of product, varying with the elemental mix of the product.

If we assume a single die has a cost of C$50, then 72,000 dies would cost C$3.6M. Doubling this to roughly account for other infrastructure required by the nanofactory might result in a cost of C$7M to C$8M. Amortized over three years, this might be ∼C$7,300 per day. Adding this to the electric power gives costs approximating C$8,620 per day.

Over the course of a year, it might cost C$3.1 M to operate the nanofactory. A production rate of approximately 100 mg/day over 365 days gives 36.5 grams/year, or C$3.1 M / 36.5 grams ≈ **C$85 K / gram**. The baseline nanofactory might manufacture atomically precise products for < C$100,000 per gram.

These estimates are subject to significant uncertainty. For example, SRU failure rates are difficult to reliably estimate today. As SRU reliability improves, more SRUs will remain functional for longer, increasing productivity. If they remain high, a more complex system able to replace failed SRUs would be needed, which would increase productivity, though at the cost of schedule delay (required to design the more complex system) and increased development costs. Completion of the baseline nanofactory should be followed by subsequent improvements and an exponential decline in manufacturing cost over the following years. An example of this kind of decline is the semiconductor industry, which has seen a factor of ∼two decline in cost per transistor every 18 months for decades. We can reasonably expect that a manufacturing technology that manufactures itself would see declines in manufacturing cost at least as great, if not greater, particularly with the expected enhancement in design capabilities provided by AI.

11 Baseline Implementation ("2026 Baseline")

*The baseline is a nanofactory composed of **58 trillion identical Simple Replicating Units (SRUs)**, each 500 nm × 100 nm × 60 nm and containing ∼100M atoms, each mounted on a 600 nm × 200 nm site on one of 72,000 nominally 1 cm^2 silicon dies.*

The performance figures below are nominal design estimates, assuming full occupancy by functioning SRUs without product-rejection or utilization losses. Replication time depends on the detailed manufacturing and error-recovery schedule. The first-order power estimate excludes buffer-network dissipation and other unbudgeted support loads.

11.1 Core Architecture

- ***Local control and routing:*** *Each SRU has a 30-bit location identifier stored in local scratchpad memory and can communicate with neighboring SRUs. A proposed mechanical implementation uses coupled rods passing through sliding seals, with actuation by the internal assembly arm and sensing by Test Units. See Section 6.3, including Sections 6.3.1 and 6.3.2; coupling details and connection/liveness testing remain to be developed.*
- ***Elements (Baseline Palette):*** *H, B, C, N, O, Si, Ge, and Pt.*
- ***AP Blocks:*** *Diamondoid atomically precise blocks made of atoms drawn from the baseline palette (≤1,000 atoms, ≤15 nm, rigid).*
- ***Feedstock:*** *A solution of (1) tetraethylgermane (C, H, Ge), (2) pentanenitrile (C, H, N), (3) triethyl borate (C, H, B), (4) tetraethyl orthosilicate (TEOS) (C, H, O, Si), and (5) $EtCpPtMe_3$ ((ethylcyclopentadienyl)trimethylplatinum(IV)) (C, H, Pt).*
- ***Waste disposal:*** *Waste molecules (H_2, C_2H_2, N_2, CO_2, $C_4H_{12}Si$ (tetramethylsilane)) vented directly into the feedstock. Tetramethylsilane is a liquid at 210 K; the other waste species are treated as gases under the assumed operating pressures.*
- ***SRU failure rate:*** *<0.1% per replication cycle. Alternatively, <0.0864% per day.*
- ***SRU position on the silicon surface:*** *self-aligned. The method of self-alignment is unspecified; it awaits specifics of the fabrication process used to make surface alignment features.*
- ***SRU structure:*** *Several subsystems inside a vacuum-tight **shell**, all made by assembling AP Blocks.*
- ***Failure handling strategy:*** *fail-stop. Failed SRUs are allowed to accumulate in place.*
- ***SRU Fabricator:*** *a programmable **fabricator** 100 nm × 100 nm × 60 nm within the SRU able to make AP Blocks. It provides **mechanosynthetic capability** at **210 K** using 4 Stewart platforms, each with 6DOF, positionally controlling four molecular tools in vacuum, with **20 pm positional accuracy**. Struts extend and contract using screw drives.*
- ***Internal Assembly Arm:*** *a Stewart Platform able to assemble AP Blocks made by the SRU fabricator. It has access to the base of the daughter SRU as it is being manufactured and extruded and also access to the fabricator output.*
- ***External Assembly Arm:*** *two stacked Stewart Platforms external to the SRU able to pass daughter SRUs to adjacent SRUs till the daughters can be placed on the die surface.*

- ***Control:*** *Broadcast architecture—6 contact pillars per SRU (Ground, V+, S1, S2, D1, D2) at* ***100 MHz*** *for the data lines. All ON SRUs execute identical instructions in lockstep (SIMD).*
- ***SIMD architecture:*** *Single Instruction stream, Multiple Data stream. Each SRU is either ON or OFF. SRUs that are ON decode the instruction broadcast. SRUs that are OFF continue to execute selection instructions but do not execute drive instructions. SRUs can be turned ON or OFF in response to various conditions.*
- ***Actuation:*** *4 differential electrostatic comb drives (0.19 V, h-BN dielectric, 24 nm shuttle throw) → 12:1 transmissions (2 nm output throw, approximately 1 nN available force) → binary Bowden cables → mechanical demux → ∼60 output pairs (120 cables) driving all SRU degrees of freedom.*
- ***Wires:*** *(4,4) carbon nanotubes.*
- ***H_2 generation:*** *Platinum catalysis to generate H_2 from HAbst-H(GeRad).* See Section 4.6.
- ***Replication:*** *∼28 hour cycle, 1,000 atoms/sec build rate. Daughter SRU extruded through a sliding seal, then placed on die via External Assembly Arm (EAA).*
- ***Scale:*** 125 *PC boards* × *576 ICs* × 805,306,368 *SRU sites/die = 57,982,058,496,000* ≈ *58T*
- ***Actuation power budget:*** *58 kW.*
- ***First-order electrical dissipation, including non-recovering capacitive switching losses:*** *approximately 157 kW.*
- ***Estimated total electrical input including refrigeration:*** *approximately 325 kW.*

11.2 Risks and Mitigation Strategies

Table 17 summarizes baseline design risks and corresponding mitigations.

Table 17: Baseline design risks and corresponding mitigations.

	Risk	***Mitigation(s)***
1	***Sliding seal leakage*** *during daughter extrusion — feedstock enters vacuum interior*	1. *Cycle die between feedstock & vacuum;* 2. *Multi-stage seals with inter-stage pumps;* 3. *Pressurized inert gas (He) inside SRU;* 4. *"Flooded SRU" with only the fabricator in vacuum behind an airlock;* 5. *Nanofluidic channels in die, die surface in vacuum.*

Table 17 (continued)

	Risk	***Mitigation(s)***
2	***Waste disposal*** *— binding sites may fail to reject vented H_2, C_2H_2, N_2, CO_2, or $C_4H_{12}Si$ during feedstock ingestion*	1. *Use a multi-stage cascade, or test each putative feedstock molecule individually using, for example, half-site closure or half-site opening tests, to achieve the desired reliability* (see Binding site leakage (Section 7.2)); 2. *Package waste gas in buckyballs (e.g., C_{26460} holding ∼30,000 H_2; see Buckyballs as a storage method for hydrogen gas (Section 18.5));* 3. *store gas internally and vent in batch after daughter is complete.*
3	***Platinum catalyst*** *— Generation of H_2 gas using a platinum catalyst might not work.*	*Use AzAdam electrochemistry (see Generating H_2 using AzAdam (Section 12.2)).*
4	***(4,4) nanotube stability*** *— unpassivated segments might not be stable or fusible*	*Use other nanotube chiralities or an entirely different design for nanowires.*
5	***Rotary motion gear down*** *— worm gear baseline may not meet requirements*	1. *Split-ring compound planetary gear;* 2. *reversed harmonic drive;* 3. *dual-indexer gear drive.*
6	***Operating temperature too high*** *(210 K) — positional uncertainty or marginal reaction reliability blocks mechanosynthesis*	1. *Thermal cycling: enlarge SRU to hold all feedstock internally at 210 K, then (1) ingest all the feedstock required to build one SRU at 210 K, (2) cool to temperature required for mechanosynthesis, (3) build the daughter, and finally (4) rewarm to 210 K before ingesting feedstock for the next SRU replication cycle.* 2. *Use feedstock substances that are liquids or gases at lower temperatures. Use ethylene, germane, diborane, silane, N_2, and drop oxygen. A compatible platinum source would also be required; a suitable precursor and its solubility and stability at the selected operating temperature remain to be established.*

Table 17 (continued)

	Risk	***Mitigation(s)***
7	***Daughter power before die attachment*** — *a 6-conductor umbilical from an adjacent already-connected SRU doesn't work.*	1. *Use a chemical or mechanical battery, ejectable after use.* 2. *Attach unpowered daughter to die.*
8	***Extrusion mechanism*** — *screw-based extrusion may be insufficient*	1. *Staggered-peg mechanism (pegs in shell holes, spring-driven);* 2. *MacArthur mechanism (knobbed rods with locking/unlocking cycle—fully reversible, no "snaps").*
9	***Mechanosynthetic error rate*** *too high for 1,000-atom blocks*	*Simple Closed Loop Mechanosynthesis: test-and-repeat within the broadcast paradigm (exponential reliability improvement with logarithmic overhead). Airlock for ejecting rejected blocks if waste volume is insufficient. Temperature reduction to lower error rates.*
10	***Build rate too low*** — *the actual operations at a 100 MHz clock rate take much longer than expected.*	*Allow replication time to take significantly longer than 10^5 seconds. Increase clock speed to more than 100 MHz. Decrease number of clock cycles to add an atom to a build site. Increase the number of data lines from four to six or more. See Build rate and replication time of an SRU (Section 6.10).*

Several items have been assumed but should be further analyzed.

11.2.1 Baseline metabolism

While Freitas and Merkle [FM08a] provide a metabolism that achieves quantitative closure for H, C, and Ge, including how to regenerate all the tools they proposed starting from those selfsame tools, and while there is a body of both theoretical and experimental evidence supporting the feasibility of mechanosynthesis [BMC+09, TSMF06, TSMF07, PFM04, H+25, B+25, M+23, FM12a, FM12b, C+26, B+26, HBMR00, O+03, MPMG91, TAI+10, TIA+13, BBTS20, FM17, HR03, Zho21, CBN], there has been no analysis of a complete set of mechanosynthetic reactions for the eight element baseline palette. What's needed is

1. A complete description of the baseline toolset.
2. Reaction sequences that recharge all tools in the baseline toolset starting from the baseline feedstock molecules.
3. Reaction sequences that build every tool starting with the baseline tools.
4. A description of baseline AP Blocks.

5. Reaction sequences that build a reasonable subset of baseline AP Blocks. This provides evidence supporting the claim that this toolset can make "most" baseline AP Blocks.

6. Verification of all reaction sequences using a credible level of theory (DFT with a good basis set, or better).

Items 2 and 3 must jointly achieve quantitative closure. A set of reaction sequences that gradually exhausts any critical resource (other than the feedstock molecules) does not exhibit quantitative closure.

Developing and verifying the baseline metabolism computationally will require a significant amount of work. This effort could begin now, as it does not require experimental capabilities.

11.2.2 Positional compliance budget

Computational work on a compliance budget can begin now. It should combine all sources of compliance in tool-tip positioning to estimate positional uncertainty at 210 K. A similar compliance budget for the internal assembly arm and the external assembly arm is also needed.

11.2.3 Strut design

A significant component of the compliance budgets depends on the atomically precise design of the Stewart platforms. The ideal strut-stiffness estimate and the derated design assumption (Section 7.14.9) indicate the magnitude but do not provide the actual stiffness of a specific atomically precise design. Again, work on the design and analysis of an atomically precise strut can begin now.

11.2.4 Comb drive force

The baseline assumes that a one nN force delivered by each of the four differential comb drives will be adequate. The actual force depends on the details of the driven molecular machinery. In the absence of a detailed design, a "best estimate" was used based on experience designing molecular mechanical computers. Further analysis is needed, preferably based on a full atomically precise design of the relevant molecular machinery.

11.2.5 Feedstock mutual solubility

A more detailed analysis of the mutual solubility of the proposed feedstock components is needed. A best guess about the elemental composition of the SRU would allow an estimate of the ratios of the feedstock molecules in the feedstock solution. The mutual solubility of the components could then be more carefully evaluated.

12 Risk Mitigations

12.1 Self-alignment of the SRU in 6DOF

If we want to self-align the SRU with features on the silicon surface, that is, we want to adjust its position in six degrees of freedom, then six contact points should suffice. If the six contact points are on some form of pillar, whether electrically active or not, the semiconductor fabrication process can provide some degree of control over the pillar location and the pillar height. Ignoring, for the moment, the issue of electrical conductivity and focusing purely on 6DOF positional control, there are many ways of distributing the six contact points. We illustrate one example (Figure 164).

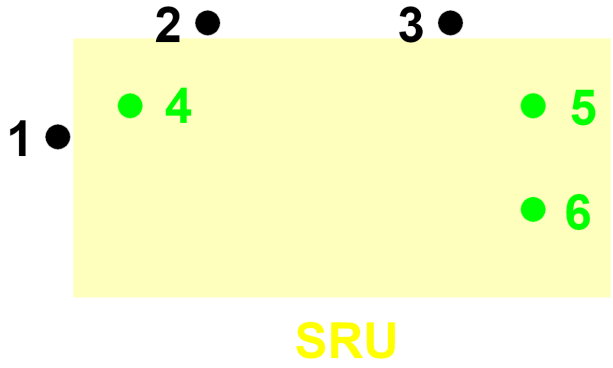


Figure 164: Six-contact arrangement for self-alignment of an SRU on the die.

The SRU, shown in yellow, sits atop three short green pillars, 4, 5 and 6, and is in contact with three tall pillars, 1, 2, and 3. If the SRU is pushed into the page (-Z), as well as left (-X) and up (Y), then it will contact all six pillars, the black pillars on their sides and the green pillars on their tops. Adjusting the height of pillar 4 controls the roll. Adjusting the height of pillar 6 adjusts the pitch. Adjusting the height of all three green pillars adjusts the Z position. Adjusting the position of pillar 2 adjusts the yaw. Adjusting both pillars 2 and 3 adjusts the Y position. Adjusting the position of pillar 1 adjusts the X position.

The fabrication process will define the position and variance of these six contact points. If the positional control provided by this approach is insufficient, it would be possible for the SRU to use six screws to control the six contact points with greater precision. That is, six small pads would be pushed out of the SRU surface at the six contact points, these pads being mounted on the end of a screw that would allow their position to be adjusted with respect to the SRU surface. Such fine adjustment would require reference points more accurate than those provided by fabrication on the silicon surface; an adjacent SRU might provide them.

Such fine control over SRU position may not be needed, but could be investigated for risk mitigation should the need arise.

12.2 Generating H_2 using AzAdam

Our feedstock is generally richer in H than the products being manufactured, so we must dispose of the excess hydrogen. The risk mitigation strategy described here is to generate H_2 using an electrochemical extension to mechanosynthesis that is based on AzAdam tools and is compatible with our baseline architecture. These tools incorporate a nitrogen atom at the adamantane bridgehead. Excess H_2 would then be directly vented from the SRU.

Electrochemistry offers a conceptually simple proposal for generating H_2. It requires that we pursue one of the following options:

(1) Select a voltage for V+ compatible with both electrochemical generation of H_2 and operation of the electrostatic differential comb drives, or

(2) Introduce a seventh and possibly an eighth contact pillar for the positive and negative voltages required for the electrochemical production of H_2 or

(3) Boost V+ to internal voltages sufficient for electrochemical production of H_2 or

(4) Something else.

For this mitigation strategy, we assume we use conventional voltage boosting methods coupled with an elementary electromechanosynthetic reaction utilizing AzAdam. The core reaction has been validated computationally (see discussion below).

12.2.1 Hydrogen donation with AzAdam-H+

While GeRad-H can effectively donate a hydrogen in many cases, there are cases where it is not effective. In particular, in this section we discuss the generation of H_2 gas via mechanosynthesis. Unfortunately, the reaction

$$\text{GeRad-H} + \text{H-GeRad} \rightarrow \text{GeRad} + \text{H}_2 + \text{GeRad} \tag{86}$$

is both endoergic and has a barrier. However, the reaction

$$\begin{aligned} &\text{e}^- + \text{AzAdam-H}^+ + \text{H-GeRad} \\ &\rightarrow \text{AzAdam-H} + \text{H-GeRad} \\ &\rightarrow \text{AzAdam} + \text{H}_2 + \text{GeRad} \end{aligned} \tag{87}$$

is exoergic.

Even more exoergic would be

$$\begin{aligned} &2\,\text{e}^- + \text{AzAdam-H}^+ + \text{H-AzAdam}^+ \\ &\rightarrow \text{AzAdam-H} + \text{H-AzAdam} \\ &\rightarrow 2\,\text{AzAdam} + \text{H}_2 \end{aligned} \tag{88}$$

This reaction should not only be strongly exoergic but should also have no barrier. Figure 165 shows the AzAdam tool tip used for hydrogen donation.

Calculations by Damian Allis (2013) show that the electromechanosynthetic (EMS) reaction

$$\text{GeRad-H} + \text{AzAdam}^+ \rightarrow \text{GeRad} + \text{AzAdam-H}^+ \tag{89}$$

favors the production of AzAdam-H+ by 44 zJ (0.27393 eV) with a barrier of ~16 zJ (0.1 eV).

Donation of H from HAbst-H(GeRad) to AzAdam+ should be more favorable energetically than donation from GeRad-H by 181 zJ (see Table 7). Therefore, the reaction

$$\text{HAbst-H(GeRad)} + \text{AzAdam}^+ \rightarrow \text{HAbst-GeRad} + \text{AzAdam-H}^+ \tag{90}$$

should be energetically favorable by 44 zJ + 181 zJ = 225 zJ.

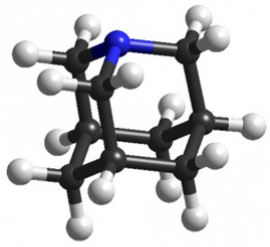


Figure 165: 1-Azaadamantane active tip of the AzAdam tool used for hydrogen donation.

12.2.2 Ionizing AzAdam

To donate H from either GeRad-H or HAbst-H(GeRad), we must first ionize AzAdam to produce AzAdam+.

The ionization potential of 1-Azaadamantane is 1,272 zJ (7.94 eV) [BS19]. The baseline operating voltage of the SRU is 0.19 V. By boosting the operating voltage using standard methods [Her22], we can generate the voltage needed to ionize AzAdam and implement the electromechanosynthetic reaction

$$\text{AzAdam} \rightarrow \text{AzAdam}^{+} + \text{e}^{-} \tag{91}$$

thereby producing the required charge state on the nitrogen.

12.2.3 AzAdam+ on diamond

AzAdam+ bonded to an otherwise perfect diamond support is expected to behave as a localized, deep, nitrogen-centered surface redox state rather than as a shallow donor whose charge delocalizes into the diamond bulk. Migration of the positive charge into the diamond would require formation of a bulk or surface hole state, or transfer involving some other defect/dopant level, and is therefore not expected in ideal insulating diamond. Quantitative confirmation is needed using, for example, a charged slab or embedded-cluster DFT calculation of the AzAdam/AzAdam+ 0/+ transition level relative to the diamond band edges.

12.2.4 Mechanosynthetic generation of H_2

A theoretically supported path for generating H_2 within the context of the current SRU architecture is to use electrochemistry. What does it mean to carry out an electrochemical reaction in the context of mechanosynthesis? To answer this question, we consider the AzAdam tool. See Hydrogen donation with AzAdam-H+ (Section 12.2.1).

The AzAdam tool is a molecular tool much like the other molecular tools that have been proposed: GeRad, HAbst, and the like. Unlike the other tools, its active tip involves a nitrogen atom. The active tip of the AzAdam tool is 1-Azaadamantane, shown in Figure 166.

Nitrogen participates in many reactions in both its neutral state and in its ionized +1 state. Removing an electron from ammonia produces a radical cation, which can combine with a hydrogen atom to form ammonium. AzAdam follows an analogous sequence. This reaction is written:

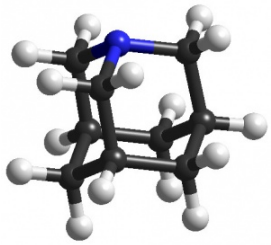


Figure 166: 1-Azaadamantane active tip of the AzAdam tool, illustrating its molecular structure.

$$NH_3^{\bullet+} + H^{\bullet} \rightleftharpoons NH_4^+ \tag{92}$$

For purposes of mechanosynthesis we're going to consider the following states of the AzAdam tool:

AzAdam

AzAdam+

AzAdam-H+ (stable)

AzAdam-H (unstable, releases H)

We can convert from AzAdam to AzAdam+ by removing an electron as in Eq. (91). We can then donate a hydrogen from HAbst-H(GeRad) to AzAdam+ using Eq. (90).

The stable AzAdam-H+ can then be converted to AzAdam + H• by adding an electron in a second electromechanosynthetic reaction:

$$\begin{aligned} \text{AzAdam-H}^+ + e^- &\rightarrow \text{AzAdam-H} \\ &\rightarrow \text{AzAdam} + H^{\bullet} \end{aligned} \tag{93}$$

It is essential that AzAdam+ and AzAdam-H+ be stable in vacuum. If a stray electron tunneled to the tip and neutralized the positive charge, the tools would fail to work. Fortunately, the AzAdam tip is mounted on a larger diamond pyramid. Diamond is an insulator, so by enlarging the diamond pyramid on which the AzAdam tip is mounted we can provide whatever insulation might be required to prevent a stray electron from tunneling to the tip. Further, there are no "stray electrons" in a eutactic structure, only electrons that are present by design. It will still be necessary to prevent tunneling from any electron that, by design, is present within tunneling distance of the tip.

We should be able to design a molecular tool that will keep both AzAdam+ and AzAdam-H+ stable in the context of an SRU once they have been created.

To create AzAdam+ we must remove an electron from AzAdam: we must ionize it. Before we do this, we must know its ionization potential.

12.2.4.1 The required voltage

The ionization potential of NH_3 is 1,631 zJ (10.18 eV). The ionization potential of 1-Azaadamantane is 1,272 zJ (7.94 eV) [BS19]. The work function of a carbon SWNT is ∼801 zJ (5 eV) [SA01]. This suggests that charging a hemispherically capped carbon (4,4) nanotube to +4 V might be sufficient to ionize an AzAdam to AzAdam$^+$. That is, an electron would prefer to move to the +4 V nanotube by 641 zJ (4 eV) because of its +4 V electric potential, plus an additional ∼801 zJ (5 eV) from the work function, giving a 1442 zJ (9 eV) total bias. This should overcome the 1,272 zJ (7.94 eV) ionization potential. On this simplified energy-level estimate, the electron transfer is energetically favorable by approximately 170 zJ (1.1 eV). A detailed calculation must include the finite probe geometry, probe charging energy, image-charge effects, local electric field, molecular orientation, separation, and electron-transfer rate. Regardless of the specific value, the use of voltage boosting (Section 12.2.6) should be able to provide sufficient voltage to ionize an AzAdam molecular tool.

12.2.4.2 Donating H to AzAdam+

Assuming we can create an AzAdam+, we must now add a hydrogen atom to it.

Calculations by Damian Allis (Mar 13 2013) show that the mechanosynthetic reaction in Eq. (89) favors the production of AzAdam-H+ by 44 zJ (0.27393 eV) with a barrier of ∼16 zJ (0.1 eV) (Figure 167). At 210 K (the baseline operating temperature for the SRU), k_BT is 2.9 zJ (where k_B is Boltzmann's constant, $1.380\,649 \times 10^{-23}\,\mathrm{J\,K^{-1}}$). The exoergicity is approximately $44/2.9 \approx 15$ times k_BT, implying the probability that the hydrogen would be found on the GeRad rather than the AzAdam+ would be less than one in three million. The molecular geometry is shown in Figure 168.

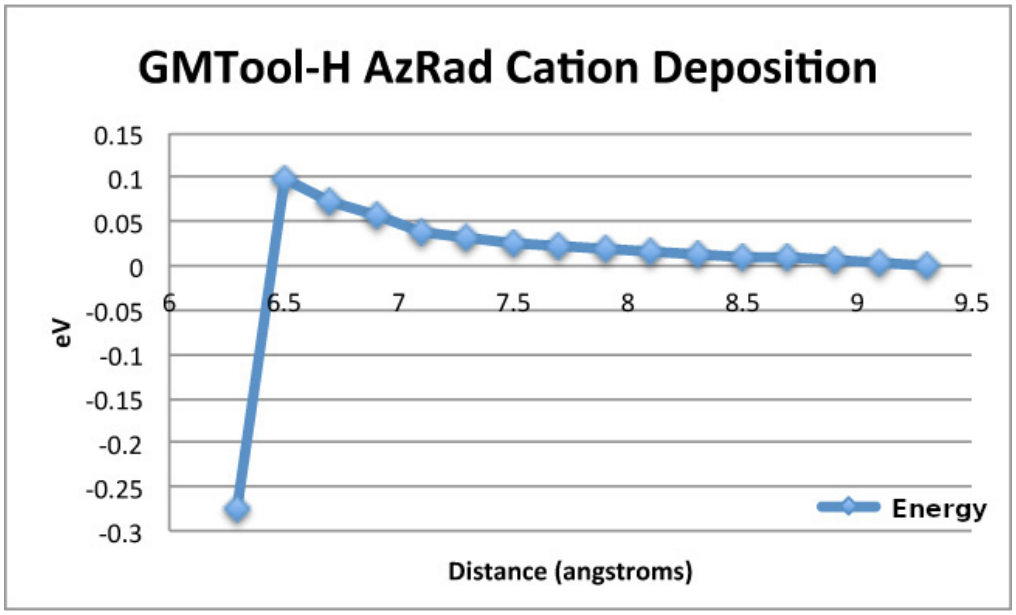


Figure 167: Calculated energy profile for hydrogen donation to AzAdam+. Energy (vertical axis, eV) is plotted against the approach distance of GeRad-H toward AzAdam+ (horizontal axis, angstroms). Calculation by Damian Allis, March 13, 2013.

The BDE of GeRad-H is ∼610 zJ and that of HAbst-H(GeRad) is ∼429 zJ (see Hydrogen donation and GeRad-H (Section 3.7.4.2)). If we donate an H from HAbst-H(GeRad) that should be 181 zJ

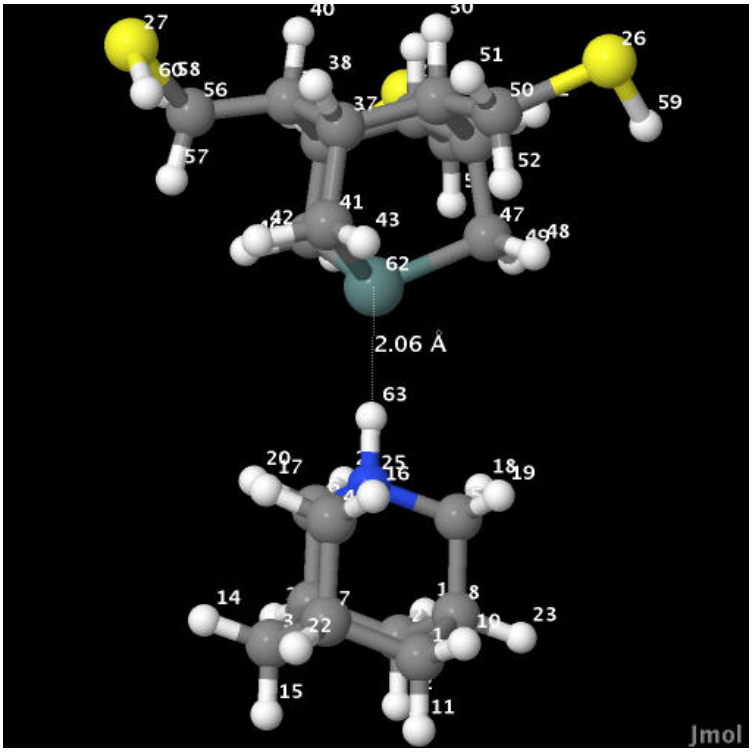


Figure 168: Molecular geometry for hydrogen donation to AzAdam+. Calculation by Damian Allis, March 13, 2013.

more favorable than a donation from GeRad-H.

Therefore, future research should consider the reaction in Eq. (90).

12.2.4.3 Releasing H from AzAdam-H+

AzAdam-H^+ should be stable in vacuum and can readily be converted to AzAdam-H by the addition of an electron. AzAdam-H is the neutral species and does not form a bond to the hydrogen, despite the reactivity of atomic hydrogen. This has obvious utility in any situation where the selective addition of a single hydrogen might be useful, but where there might otherwise be a barrier to the addition. This would be useful when donating H to Ge, as in one reaction required in [FM08a].

AzAdam-H^+ could be positioned opposite a second GeRad-H and, with the addition of an electron via a suitable negative probe, the following electrochemical reaction should take place:

$$\text{AzAdam-H}^+ + \text{H-GeRad} + \text{e}^- \rightarrow \text{AzAdam} + \text{H}_2 + \text{GeRad} \tag{94}$$

As the formation of H-H yields 724 zJ, while breaking the GeRad-H bond costs only 610 zJ, there is 114 zJ favoring the reaction (assuming the neutral AzAdam-H bond has negligible strength).

The two-AzAdam hydrogen-release reaction in Eq. (88) would have the entire 724 zJ of the H-H bond driving it forwards. This requires the donation of two electrons.

Summary: By extending our mechanosynthetic framework to include electrochemistry, we should be able to generate one H_2 gas molecule from two GeRad-H molecular tools. The particular mechanosynthetic reactions proposed require the use of nitrogen. This implies a requirement for the ingestion and metabolism of a nitrogen-bearing feedstock in our baseline proposal.

12.2.5 Electrochemistry infrastructure

These three mechanosynthetic reactions, when used in sequence, should be able to synthesize H_2:

1) $\text{AzAdam} \rightarrow \text{AzAdam}^+ + e^-$ (electrochemical with positive probe)

2) $\text{AzAdam}^+ + \text{H-GeRad} \rightarrow \text{AzAdam-H}^+ + \text{GeRad}$ (no probe)

3) $\text{AzAdam-H}^+ + \text{H-GeRad} + e^- \rightarrow \text{AzAdam} + H_2 + \text{GeRad}$ (electrochemical with negative probe)

    a. Alternative reaction:
    $2\,\text{AzAdam-H}^+ + 2\,e^- \rightarrow 2\,\text{AzAdam} + H_2$ (electrochemical with negative probe)

To use these electrochemical reactions we must first provide an infrastructure capable of supporting them. That is, we must be able to add and remove electrons as required.

One method would be to approach a mechanosynthetic tip with a charging probe. For example, to remove an electron from an AzAdam tip, we might approach it with a positively charged small diameter single-walled nanotube (SWNT) with a hemispherical fullerene cap, such as a (4,4) nanotube.

The voltage at the end of the charging probe must be measured with respect to some reference voltage, such as the voltage in "free space" surrounding the end of the probe. In our context, "free space" is the volume within the SRU fabricator. To provide a reference voltage within the SRU fabricator we might put it in a conductive enclosure made of graphene and conductive nanotubes connected to ground. Ground is one of the six contact pillars to which the SRU is connected. Openings in this enclosure would be required in a few places but should not substantially alter its ability to provide a "free space" reference voltage for the mechanosynthetic reactions taking place within the volume of the SRU fabricator.

A typical Van de Graaff generator is a conductive metallic sphere with an opening to allow entry of charge on a belt. The free space voltage within the sphere is very close to the voltage of the metallic sphere itself, even though there is a hole in the metallic sphere (Figure 169).

Another example of an enclosure that provides a voltage in free space within its volume might be the (20,20) carbon SWNT used as the gate in the molecular Field Effect Transistor (mFET) in [MFA25]. This enclosure is a cylinder that is 4.7 nm long and 2.8 nm in diameter. Despite the open ends of this cylindrical "enclosure", the spatial region near the center of the cylinder should have a free space voltage that is a close enough approximation to the gate voltage to allow this mFET to function.

In a similar fashion, even though our conductive enclosure embedded in the walls of the SRU fabricator has openings to allow access to the charging probes that provide positive and negative voltages for electrochemical reactions, it should be sufficient to provide a reference voltage when we remove a single electron from our AzAdam molecular tool tip and when we add a single electron to our AzAdam-H+ molecular tool tip.

Our electrochemical infrastructure therefore consists primarily of the following:

1) A conductive enclosure surrounding the SRU fabricator;

2) Two charged probes, one positive and one negative, that enter through openings in the conductive enclosure, and which donate and abstract individual electrons, as needed, to specific

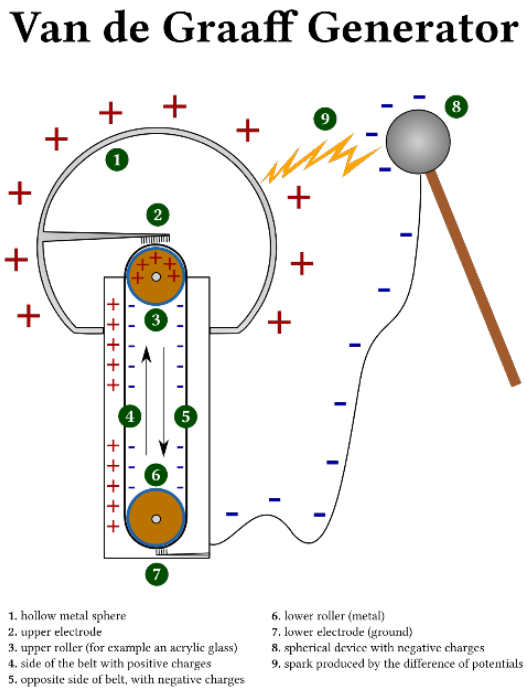


Figure 169: Van de Graaff generator illustrating an electric potential inside a conducting enclosure. Raster reproduction of Omphalosskeptic's diagram, after Dake and Gonfer00, licensed under CC BY-SA 3.0.

mechanosynthetic reactions. As carbon SWNTs have a work function of ~801 zJ (5 eV) [SA01], it would be advantageous to use them for the positive probe which removes electrons. For the same reason, it would be disadvantageous to use them for the negative probe. The use of an ohmic contact to a diamond probe which had nitrogen-terminated (111) faces and one hydrogenated face through which electrons could be donated is one possibility that might be considered. The hydrogenated diamond (111) surface has a negative electron affinity, which would be helpful in donating electrons, while the nitrogen terminated (111) surface is expected to have a positive electron affinity, which would be helpful in preventing electrons from departing the diamond probe at random undesired locations.

3) A voltage-boost subsystem (described below) providing the positive and negative voltages needed for the two probes.

12.2.6 Voltage boosting

Having once placed our SRU fabricator in a conductive enclosure tied to ground, we must then boost the voltage of V+ to a suitable value to allow the desired electrochemical reactions to proceed.

Voltage boosting is the art of taking a direct-current (DC) source—say a 3 V DC source—and stepping it up to a higher, more useful level, such as the 12 V needed to drive a string of white LEDs.

Compact voltage boosters can be made from a handful of capacitors in a switched-capacitor charge pump. By rapidly re-wiring between charging one (or more) capacitors and discharging them in series, a charge pump can double, triple, or even invert a supply, making it useful for on-chip

tasks like biasing AMOLED displays or driving GaN gate drivers [Her22, Sch01]. The trade-off is that the output resistance rises with each added stage, so efficiency fades as load current climbs (Figure 170).

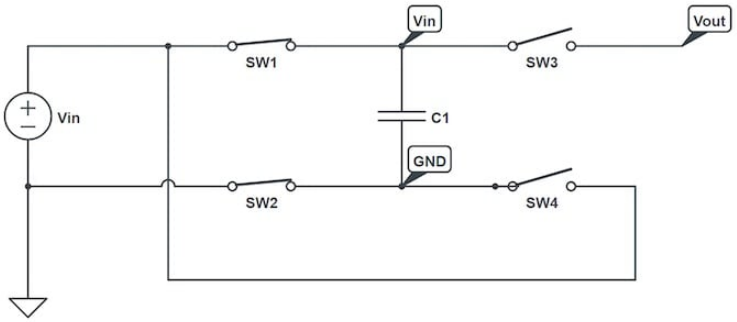


Figure 170: Charging phase of a switched-capacitor voltage booster. From [Her22].

Figure 171 shows the common phase of a switched-capacitor voltage booster.

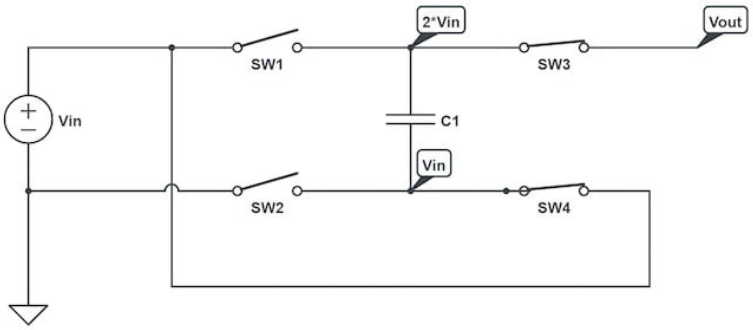


Figure 171: Series discharging phase of a switched-capacitor voltage booster. From [Her22].

Voltage boosting is conceptually simple. A capacitor C1 has two terminals. Referring to Figures 170 and 171, we start with SW1 and SW2 closed, while SW3 and SW4 are open. Capacitor C1 is charged by Vin. Then, SW1 and SW2 are opened and SW3 and SW4 are closed. Vin and C1 are now connected in series, and the voltage across them is summed, effectively doubling the voltage of Vin.

This is the essence of voltage boosting. Charge a capacitor from an input voltage, disconnect the capacitor, and then reconnect the capacitor in series with the input voltage to provide a boosted output voltage.

This process can be repeated in multiple stages, each stage providing an additional voltage boost.

This process requires switches. Although a detailed design for a P-channel mFET (molecular Field Effect Transistor) is available [MFA25], this approach, at least for now, has some drawbacks. Among them are the lack of a similar design for an N-channel mFET, the need to design circuits able to handle high voltages and minimize leakage currents, and the greater design effort as compared with mechanical switches. Unless the electronic circuit was driven by an electronic controller of some

sort, it could still only operate at the mechanical speed of the mechanical drivers (the outputs of the demux) already used in the SRU.

A mechanical switch might be made from carbon nanotube wires, possibly by sliding a length of wire between two parallel wires, opening and closing the switch. In Figure 172, the black lines are conductors, possibly (4,4) carbon SWNT, while the green lines are insulators, possibly chirally matching h-BN. Sliding the horizontal connecting wire back and forth opens and closes the switch. While the geometry shown minimizes contact area between the sliding wire and the switched wires, this is primarily for convenience of this illustration. In 3D, other options that increase contact area are available. One can imagine the vertical conductors are nanotubes terminated with a 4-way planar connector large enough to surround the horizontal conductor/insulator. Leakage current in an appropriately designed switch falls exponentially with increasing separation between the poles. A separation of a few nanometers is proposed.

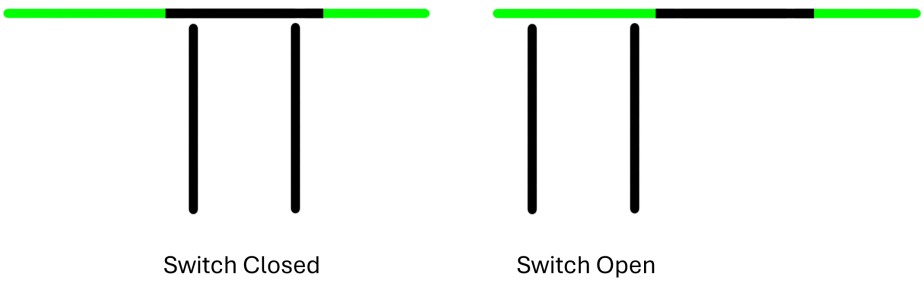


Figure 172: Closed and open states of a nanoscale switch for the voltage booster. Black is conductive, green is insulating.

12.2.7 Sizing the voltage-boost capacitors

To estimate the volume required for the voltage-boost capacitors, we consider an idealized capacitor consisting of two square graphene plates, each measuring 10 nm × 10 nm, separated by 2 nm of vacuum. We use its calculated capacitance and active volume as a reference for scaling to the approximately 20 aF capacitance required by the voltage-boost circuit.

12.2.7.1 Capacitance of a parallel-plate capacitor

$$C = \frac{\epsilon_r \, \epsilon_0 \, A}{d}. \tag{95}$$

Here $\epsilon_r = 1$ for vacuum, $\epsilon_0 \approx 8.85 \times 10^{-12}\,\mathrm{F/m}$ (see Physical Constants (Section 17.4)), $A = (10\,\mathrm{nm})(10\,\mathrm{nm}) = 1.0 \times 10^{-16}\,\mathrm{m}^2$, and $d = 2\,\mathrm{nm}$. Thus,

$$C = \frac{(8.854 \times 10^{-12}\,\mathrm{F/m})(1.0 \times 10^{-16}\,\mathrm{m}^2)}{2.0 \times 10^{-9}\,\mathrm{m}} \approx 4.4 \times 10^{-19}\,\mathrm{F} = 0.44\,\mathrm{aF}. \tag{96}$$

A 10 nm × 10 nm parallel-plate geometry with a 2 nm vacuum gap has a calculated capacitance of approximately 0.44 aF. Capacitance and active volume scale linearly with plate area when the plate separation is held constant.

12.2.7.2 Capacitance of a wire

On-chip interconnect capacitance depends on metal layer, linewidth, spacing, and surrounding conductors, but 0.2 fF/μm (0.2 aF/nm) is a commonly used first-order estimate for dense, minimum-pitch metal wiring [WH11].

12.2.7.3 Inductance of a wire

We similarly use 1 nH/mm (≈1 pH/µm) as an order-of-magnitude placeholder. The actual inductance and propagation velocity must be calculated from the complete signal-and-return geometry [Bog14].

12.2.7.4 Volume of a small capacitor

A 10 nm × 10 nm vacuum-gap parallel-plate capacitor with a 2 nm plate separation, charged to 2 V, holds a classical charge of 8.9×10^{-19} C.

The active volume between the plates is 10 nm × 10 nm × 2 nm = 200 nm^3.

12.2.7.5 Volume of the voltage boosting capacitors

If we deliver the electrons from the voltage boosting circuitry to the charging probe through a molecular wire we must charge the wire, which has a capacitance of ∼0.2 aF/nm (Section 12.2.7.2). A 100 nm wire (somewhat longer than we require) should be ∼20 aF. If we want to charge the wire's capacitance in a few boost cycles, we'd need a capacitor of roughly similar capacitance to that of the wire: 20 aF. As a 0.44 aF capacitor has a volume of 200 nm^3, our 20 aF capacitor will need about 200 nm^3 × 20/0.44 ≈ 9,100 nm^3 of active volume. Properly switched, we can boost voltage from V+ to 6 V using approximately (6 V) / V+ such capacitors, rounded up to an integer. For this arrangement, the total active capacitor volume therefore depends on the selected input voltage V+, with a higher input voltage requiring fewer capacitors. There will be a drop in voltage when we first charge the molecular wire, but a few cycles can charge the molecular wire to our required 4 V and more.

12.2.8 Summary of generation of H_2 using AzAdam

The nanofactory uses small feedstock molecules which have excess hydrogen. We could dispose of excess hydrogen by generating H_2 using an electrochemical extension of mechanosynthesis. For this approach, the SRU fabricator is placed in a (not entirely enclosing) conductive enclosure made from graphene and nanotubes connected to the ground contact pillar, one of the six contact pillars to which the SRU is electrically connected. The "free space" voltage within the fabricator is therefore ground (or close to it). The V+ voltage provided by a second contact pillar connected to the SRU is then boosted to 6 V using a standard capacitive voltage boosting circuit which uses mechanical switches. This voltage boosting circuitry then charges the wire connecting this circuitry to a charging probe in the fabricator. The wire and the charging probe must charge to 4 V. We could boost to a higher voltage if needed. The capacitor volume depends on the selected input voltage, target voltage, and capacitances. If this mitigation is adopted, the supply voltage would need to be revisited along with the circuit design and its volume and power requirements; it need not remain at the baseline value of 0.19 V. The positive charging probe, a carbon SWNT operated at the boosted voltage, ionizes two AzAdam tools. Each receives a hydrogen from GeRad-H, producing two AzAdam-H^+ tools. Donation of two electrons from the negative probe then allows the two tools to release H_2 and regenerate two neutral AzAdam tools (Eq. (88)). The H_2 is then ejected

from the SRU. The feedstock intake must filter out the H_2 gas to preserve the internal SRU eutactic environment (see Binding sites (Section 4.4.5)). H_2 gas entering the SRU is assumed to create a hard failure that must be prevented by blocking its entry.

12.3 Disposal of excess gases other than venting

12.3.1 Disposal of excess gas in containers

The baseline waste products are H_2, C_2H_2, N_2, CO_2, and tetramethylsilane. A uniform method for disposing of excess gases would be to package them into large spherical containers, such as large buckyballs. Volume increases as the cube, area as the square. As the radius increases, the material needed to build the sphere becomes an ever smaller fraction of the waste it contains. For simplicity, we would prefer to avoid pressures that solidify the waste gases. At 210 K, H_2 and N_2 do not solidify at 1,000 atmospheres and reach densities comparable to their normal-boiling-point liquids, as quantified below.

CO_2 solidifies well below 1,000 atmospheres. Tetramethylsilane and acetylene might solidify at 1,000 atmospheres. If we wish to dispose of a waste molecule which goes through a phase transition to a solid below 1,000 atmospheres, we could add waste to our waste-containing sphere until it had reached an equilibrium where it was largely but not completely solid. At this equilibrium content, where the solid phase and either the liquid or gas phase co-existed, the shape of the solid phase would be able to adjust to the spherical shape of the waste container. This equilibrium point would also provide space efficient storage, as the waste would mostly be in the solid phase. As this equilibrium pressure would be different for different waste molecules, it might be necessary to use separate waste containers for separate types of waste.

At 210 K and 1,000 atmospheres, hydrogen reaches approximately 86% of the density of liquid hydrogen at its normal boiling point. Further compression is possible, but with diminishing returns: nearly doubling the pressure increases density by only about 37% [LHL08, Nee]. At 210 K and 1,000 atmospheres, nitrogen likewise reaches approximately 86% of the density of liquid nitrogen at its normal boiling point. Nearly doubling the pressure increases its density by only about 18% [SLJ+00, Nee]. These comparisons are density benchmarks; neither fluid is liquid at 210 K, and ordinary liquid density is not an ultimate compression limit.

We carry out the calculations for H_2. Packaging gases under pressure into buckyballs has the advantage of using relatively few carbon atoms, as the surface area (hence carbon atom count) of a buckyball increases as the square of the diameter, d^2, while its volume increases as d^3. In addition, large buckyballs will not clog up input binding sites. A disadvantage is that we might have to assemble them from AP Blocks using the internal assembly arm. Two possible AP Blocks would be coronene and corannulene ($C_{20}H_{10}$), a curved molecule with a five-membered ring at its center instead of a six-membered ring. Corannulene has 20 carbon atoms, as opposed to coronene's 24. As buckyballs have exactly twelve five-membered rings, we could assemble a 240 atom buckyball from 12 corannulene blocks. The assembly operation would presumably require selective dehydrogenation of pairs of carbon atoms during the assembly of the curved corannulene blocks, as the assembly operation in this case would involve bonding of carbon atoms. Larger buckyballs could be produced by adding coronenes as well as the 12 corannulene molecules into the buckyball. The assembly operation in this case would seem to require selective dehydrogenation of specific carbon atoms from the two coronene molecules to be bonded together, unless the unhydrogenated perimeters of coronene and corannulene were sufficiently stable before assembly. As the diameter of C_{240} is 1.42 nm, and the atom count is smaller than the 1,000 atom size limit we have adopted for an AP Block,

we could manufacture C_{240} in the fabricator and export it after it had been filled with, for example, H_2 molecules. Unfortunately, such a small buckyball would not contain many H_2 molecules.

$C_{26\,460}$ has a diameter of ∼15 nm and should be able to contain ∼30,000 molecules of H_2 at 1,000 atmospheres of pressure at 210 K. Bond strain caused by the internal pressure would be ∼50 pN per bond, <1% of the C-C bond rupture force of ∼6 nN (6,000 pN). See Buckyballs as a storage method for hydrogen gas (Section 18.5). It might be possible to manufacture the entire sphere and load it with H_2 while still in the fabricator. It could then be exported as a sealed unit for disposal. For manufacturing purposes, the sphere could be broken down into blocks, each block consisting of a graphene sheet surrounded by an edge, the edges being designed to assemble/bond to other edges. Principles of convergent assembly (see Convergent assembly and error mitigation (Section 6.7)) could be used to mitigate the impact of assembly failures.

The complexity of this design would have to be compared with the complexity of designing binding sites that blocked entry of waste gas. It would also be necessary to analyze what happens to a large number of 15 nm buckyballs expelled from the SRU.

12.3.2 Internal storage of waste gas

Another method of dealing with waste gases would be to store them internally until manufacture of the daughter SRU was completed, after which they would be expelled at once. Following this, the feedstock would be cleared of the expelled waste gases prior to initiating further ingestion.

This would create a cycle: build a daughter while storing waste gas internally. After manufacture of the daughter had been completed, vent the waste gas. Then pump down the feedstock to remove the expelled waste gas. This last step could be extended over a period of some hours given that the time to manufacture a daughter is over a day. Removal of the gas-infused feedstock and pump down to a vacuum would be feasible if necessary. Finally, reintroduction of feedstock that did not have waste gas would establish the preconditions for repeating the cycle.

12.3.3 More general balance

Alternatively, we could have multiple feedstock molecules (more than two), each one rich in a particular element. For illustration, consider just the hydrogen–carbon balance with two feedstocks: we could have one feedstock molecule that was carbon rich and one that was hydrogen rich. If the carbon rich feedstock had enough carbon, we could simply process more of the carbon rich feedstock if we needed more carbon, and more of the hydrogen rich feedstock when we needed more hydrogen.

Numerically, if each carbon atom in the carbon rich feedstock has Hc hydrogen atoms, and each carbon atom in the hydrogen rich feedstock has Hh hydrogen atoms, and each carbon atom in the product has Hp hydrogen atoms, then we must have Hh > Hp > Hc. If our hydrogen rich feedstock were methane (CH_4), then the other feedstock might be carbon nanotubes.

Most small feedstock molecules have more hydrogen than can be used in the manufacture of products. Notable exceptions are all-carbon feedstocks, including fullerenes such as C_{60} and C_{70}, and carbon nanotubes. Plausible metabolic pathways for buckyballs have so far not been developed. Nanotubes are an attractive option. These remain opportunities at this time.

12.4 Another Way of Generating Rotary Motion

The ratchet and pawl can be used to drive a rotary shaft, as illustrated here (Figure 173):

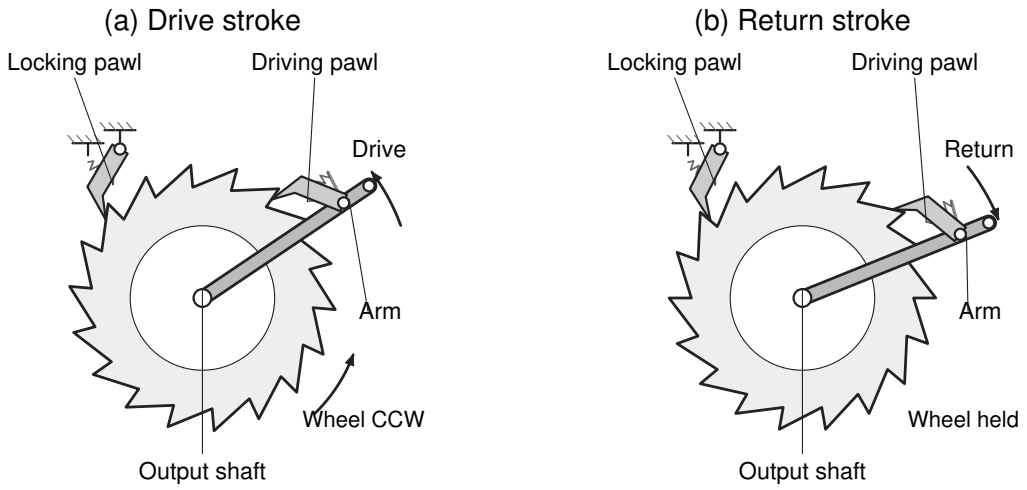


Figure 173: Ratchet-and-pawl mechanism, shown schematically. (a) The driving arm advances the wheel counterclockwise while the locking pawl rides over a sloped tooth face. (b) On the return stroke, the locking pawl holds the wheel while the driving pawl rides over a sloped tooth face.

An actuator attached to the end of the driving-pawl arm should be sufficient (Figure 173). When actuated, it would cause the driving-pawl arm to move counterclockwise and advance the ratchet wheel by one tooth. The actuator would then return the arm to its original position. During this return stroke, the locking pawl holds the wheel against reverse motion while the driving pawl rides over the next tooth, ready for another advance.

The major disadvantage of this approach is that it involves two springs and two "snapping" motions that are inherently dissipative. To make sure it operates reliably, each snapping motion must dissipate energy much greater than k_BT. If the additional energy budget can be tolerated, the resulting concept appears otherwise to be attractive.

If our clock is 100 MHz, then we might drive the pawl once per clock cycle. The actual power stroke might be one of the four clock transitions.

Total energy dissipation might be on the order of 100 k_BT for the locking pawl, and another 100 k_BT for the driving pawl, plus whatever energy was required to apply actual force to the rotating wheel. That is, there would be an overhead of about 200 k_BT above and beyond the actual energy cost of the useful force applied to rotate the shaft and produce some useful mechanical motion. The useful mechanical energy budget per clock transition is approximately 690 k_BT at 210 K; the full 2.5 aJ actuator budget, including the 25% margin, is approximately 862 k_BT (see Summary of Critical Parameters (Section 3.3.2)). Recall that driving an output uses four clock transitions per clock cycle (two transitions for D1 and two for D2). Our 200 k_BT is well below this limit, even assuming we drive the mechanism through the "snap" using only the power from a single clock transition.

If we assume a system operating frequency of 100 MHz and a system operating temperature of

~210 K, then each such device will have a power dissipation in snapping per SRU of about

$$(10^8\,\mathrm{s}^{-1})(200)(1.38\times10^{-23}\,\mathrm{J/K})(210\,\mathrm{K})\approx 5.8\times10^{-11}\,\mathrm{W}.$$

If the power requirements of one SRU are 10^{-9} W, then we are using 5.8×10^{-11} / 10^{-9} or 5.8% of our power on "snaps". Another way of viewing this: 58 trillion SRUs operating at 100 MHz in our nanofactory implies we would waste about 3,364 watts in heat out of our 58 kW budget because we used a ratchet and pawl mechanism. This is again 5.8% of our heat budget. Additional heat would be dissipated actually driving the output load on the output shafts.

The astute reader will notice that the mechanism illustrated will drive rotary motion in one direction only. The reader might enjoy thinking about how to drive motion in two directions before googling "reversible ratchet and pawl".

12.5 Split ring compound planetary gear

Among the many possible gearbox designs which might be used to reduce the rotational speed of a shaft, of particular interest is the split ring compound planetary gear, illustrated here (Figure 174):

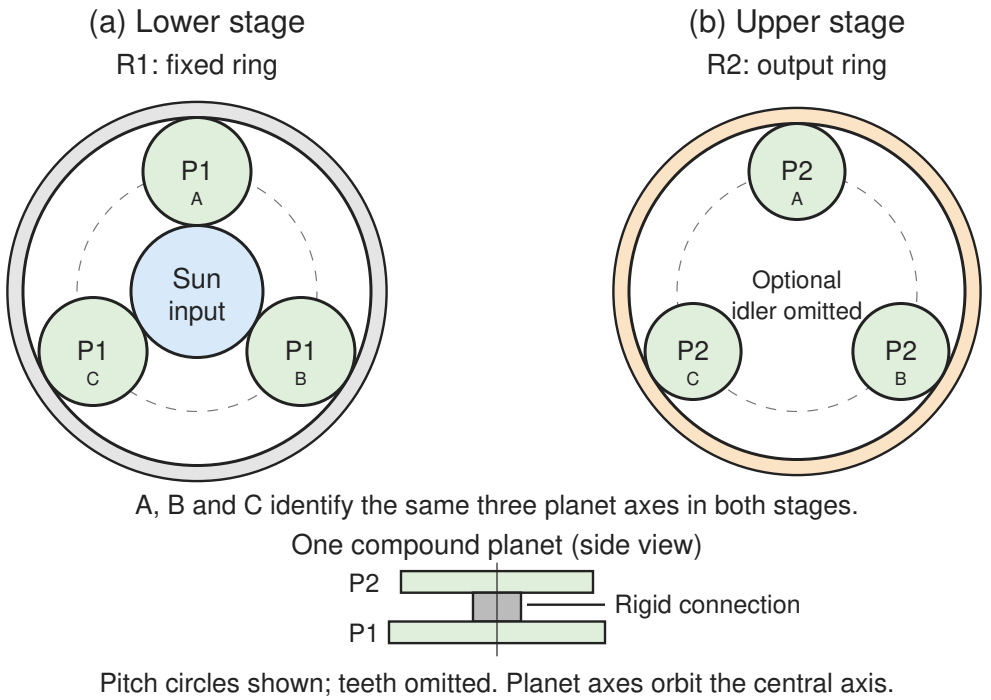


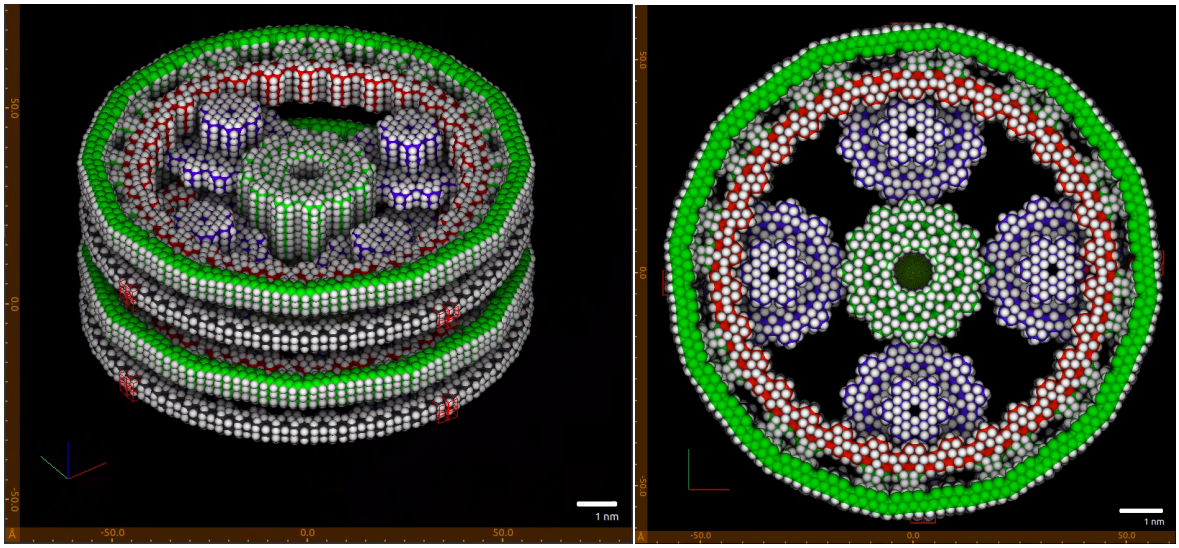
Figure 174: Split-ring compound planetary gear mechanism [Jua18]: (a) lower stage with input sun and fixed ring; (b) upper stage with output ring. Corresponding planet sections P1 and P2 are rigidly connected on each of the three planet axes. Pitch-circle schematic; optional upper idler omitted. For a related implementation, see the MPRT project.

The split-ring compound planetary gear is a specialized gear mechanism designed to achieve exceptionally high reduction ratios within a compact and rigid structure. Unlike harmonic drives, which rely on flexible components, this gear system utilizes entirely rigid parts, making it suitable for applications where flexibility is undesirable or impractical.

At its core, the split-ring compound planetary gear comprises two distinct ring gears—commonly referred to as the top and bottom rings—with slightly different tooth counts. For instance, one configuration might feature 49 teeth on the top ring and 50 on the bottom. These rings interact with compound planet gears, each consisting of two gears mounted on the same axis which rotate together but which differ in size. The bottom sun gear provides the input and drives the planet gears, which simultaneously engage both ring gears. The bottom ring gear is held stationary, while the top ring gear is driven by the planets. The top ring gear provides the output. The slight difference in tooth count between the ring gears causes a gradual phase shift as the planets rotate, resulting in a significant reduction in output speed. The top sun gear, if present, is just an idler to keep the planets axially aligned and engaged with the top ring gear.

This design enables the gear system to achieve a very large combined reduction ratio. For example, a bottom stage reduction of 3.5:1 combined with a top stage reduction of 49:1 yields an overall reduction of approximately 171.5:1. Such high reduction ratios are typically challenging to attain with conventional gear systems without increasing size or complexity.

In summary, the split-ring compound planetary gear offers a compact and efficient solution for achieving high reduction ratios with entirely rigid components. Its unique configuration makes it an attractive option for precision applications where space is limited, and flexibility is not desired [Jua18, Sau, Jan21] (Figure 175).



(a) (b)

Figure 175: Two-stage split ring compound planetary gear, 100:1 Gear Ratio. Two views of the same gear. Atomically precise design by Mark Jobes, 2024-09-27, from an unpublished internal presentation [Job].

12.6 Split ring planetary gear

Figures 176 and 177 illustrate possible split-ring gearing arrangements generated with Pål Saugstad's calculator [Sau]. These use single planet gears rather than the paired, coaxial planet gears of the preceding compound design. They illustrate gearing options rather than specific atomically precise designs.

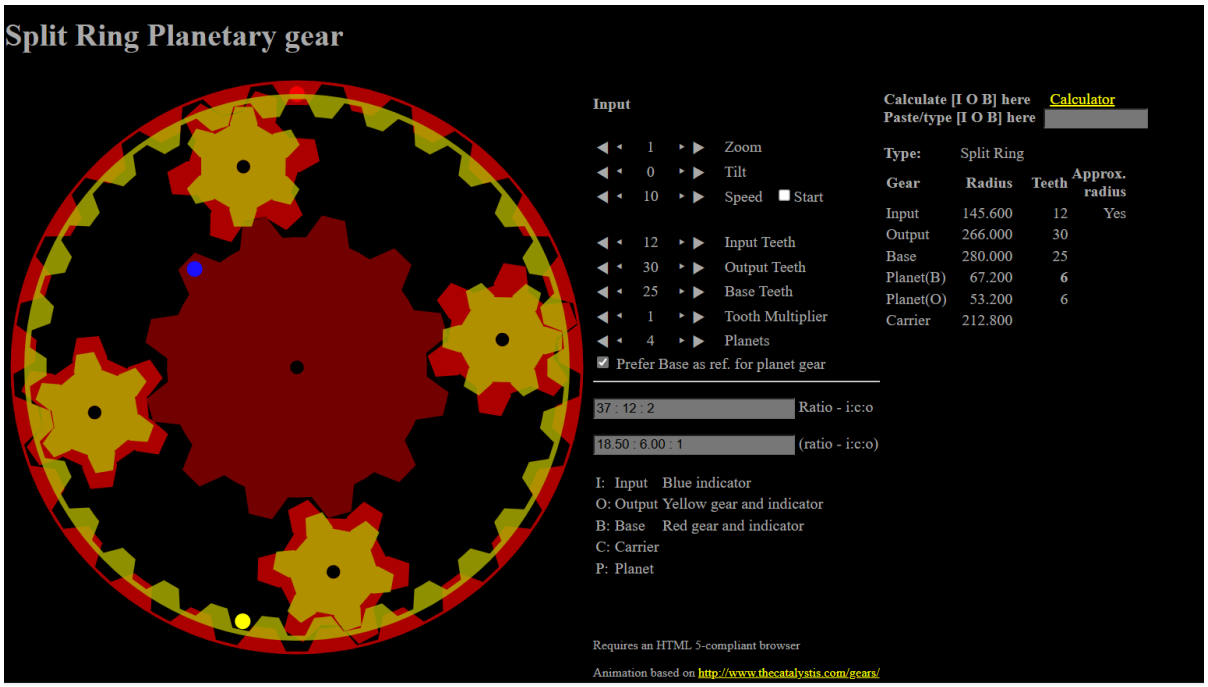


Figure 176: Split-ring planetary gear arrangement with 25 base-ring teeth and an indicated input/output ratio of 18.5:1, from Saugstad's calculator [Sau].

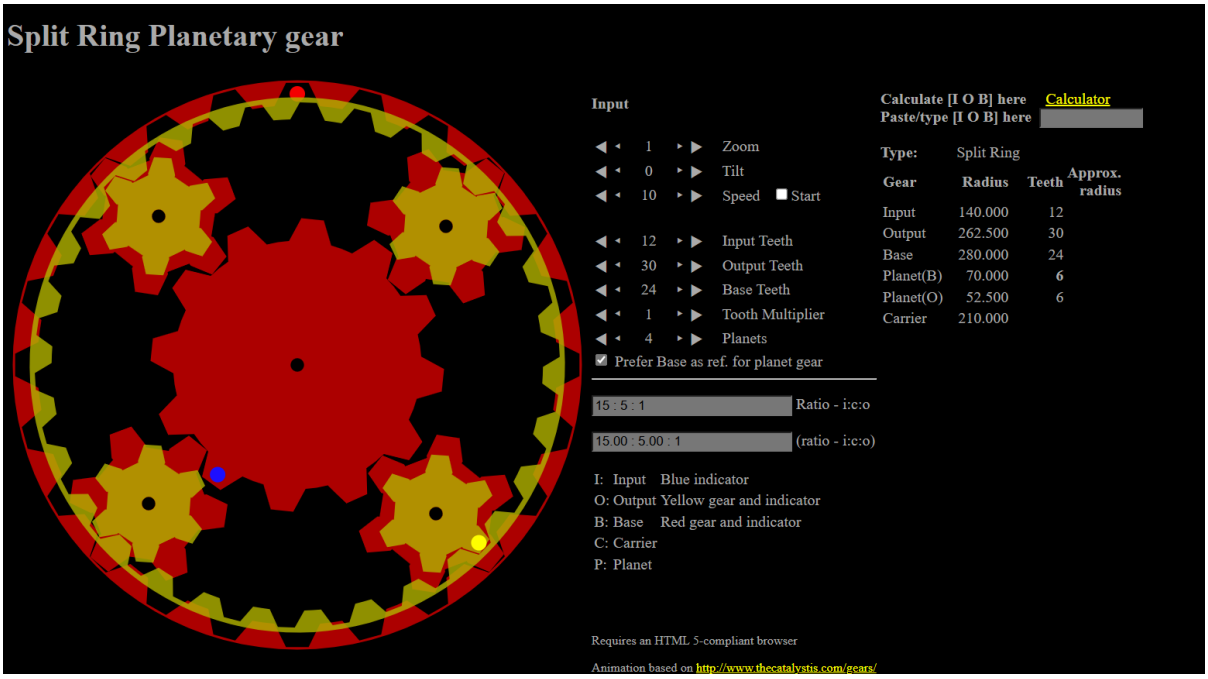


Figure 177: Split-ring planetary gear arrangement with 24 base-ring teeth and an indicated input/output ratio of 15:1, from Saugstad's calculator [Sau].

12.7 Reversed Harmonic Drive

Matt Moses suggested the use of a reversed harmonic drive (or “direct drive strain wave motor”) [Mos] (Figure 178).

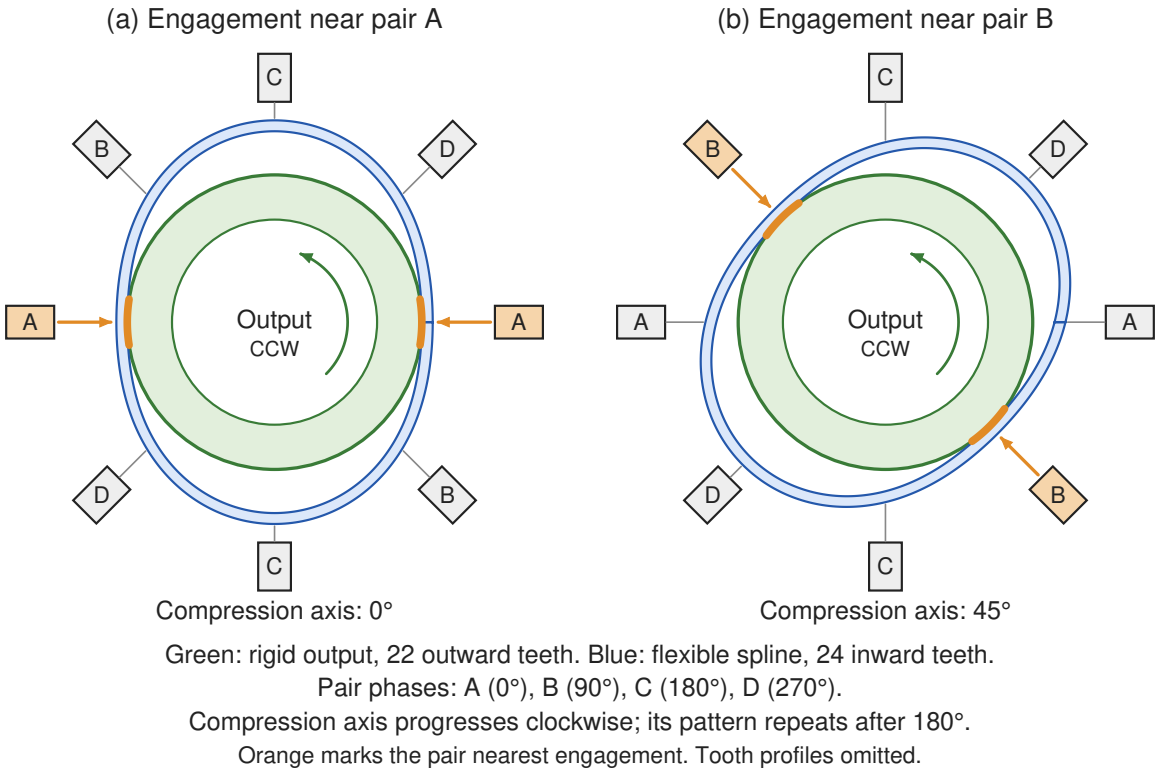


Figure 178: Planar schematic of the reversed harmonic drive in two successive compression positions. Equal letters identify opposing actuator pairs. Orange highlights the actuator pair nearest the engagement regions. The blue flexible spline drives the green rigid output counterclockwise as the compression axis advances clockwise. Tooth profiles are omitted.

In the reversed harmonic drive, the teeth on the rigid output gear (shown in green) are driven by teeth on a flexible spline (shown in blue). There are also eight linear actuators, which in our case would be driven by two binary Bowden cables. This is because each of the four pairs is driven by one phase, and two of the four phases in four phase clocking can be derived from the other two phases by inverting them. We can provide two phases, 0° and 90°, with two binary Bowden cables. We can then generate the other two phases by inverting the first two. The inverse of 0° is 180° and the inverse of 90° is 270°. Once we have all four phases, we drive each of the four pairs of actuators with one of the four phases, thus driving all eight actuators from two binary Bowden cables. The two actuators in each pair are directly opposite each other across the center of the circle. The two actuators in a pair compress the blue spline teeth into the green gear teeth at the same time, and then both paired actuators pull the blue spline teeth away from the green gear teeth at the same time. The opposing engagement regions advance around the circle, creating a rotating wave of compression during the four-phase clock cycle. Because each pair acts at opposite points, the compression pattern repeats after a half-turn. Because there are 22 outward facing green teeth and 24 inward facing blue teeth, one complete clock cycle causes the inner green teeth to rotate one tooth counterclockwise.

We next show an atomically precise design by Sheldon Haird (slides courtesy of S. Haird) (Figure 179).

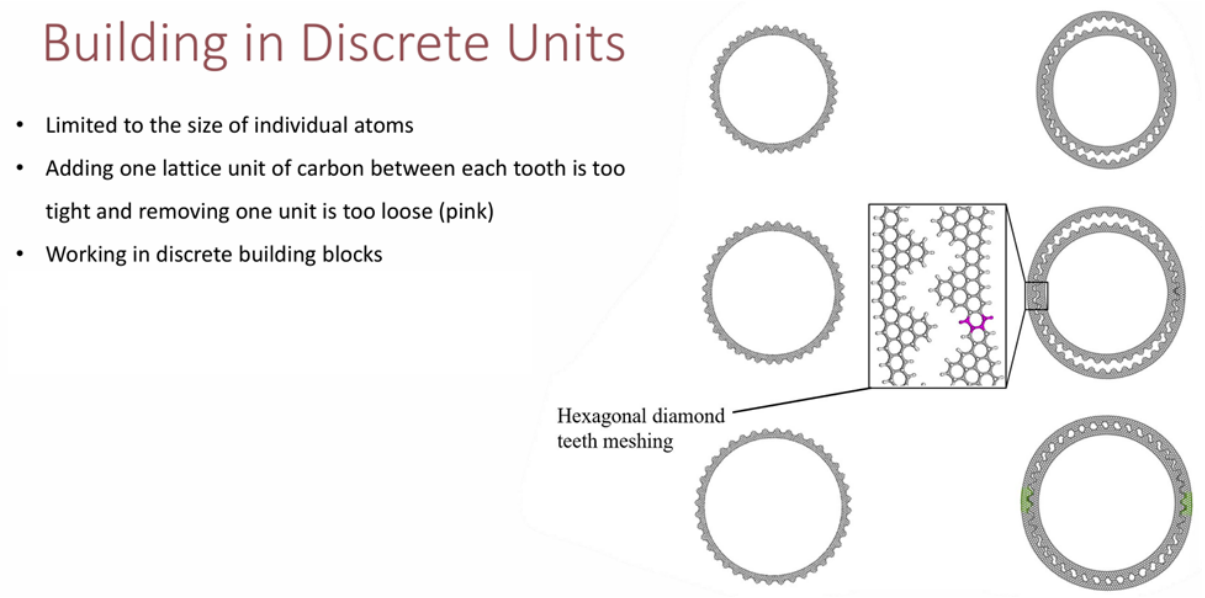


Figure 179: Building discrete units for Sheldon Haird's atomically precise reversed harmonic drive (slide courtesy of S. Haird).

Figure 180 shows the discrete building blocks for the reversed harmonic drive.

Figure 181 shows the assembled reversed harmonic drive and its tooth-meshing geometry.

Figure 182 shows a simulated actuation of the reversed harmonic drive.

12.8 Using two indexers to rotate a gear

If we have two binary Bowden cables driven by four phase clocking, and want to drive our gear forwards one tooth per clock cycle, we might:

1) Insert an indexer between two teeth.
2) Drive the indexer forwards by the distance between two teeth.
3) Withdraw the indexer from between the two teeth.
4) Restore the indexer by moving it back to its position just prior to Step 1.

The insert/withdraw motions would be controlled by the first binary Bowden cable, while the drive/restore motion would be controlled by the second binary Bowden cable.

This works if we assume the gear doesn't move after the indexer has been withdrawn. On the molecular scale, this is a bad assumption. We will need a second indexer that is (a) inserted when the first indexer is withdrawn and (b) withdrawn when the first indexer is inserted. The gear will remain locked in place during the interval in which one indexer is being withdrawn while the other is being inserted, if their movements, sizes, and the sizes of the gear teeth are chosen properly.

With this, the movements of the second indexer, synced to the timesteps of the first indexer, are:



Figure 180: Discrete building blocks for the reversed harmonic drive (slide courtesy of S. Haird).

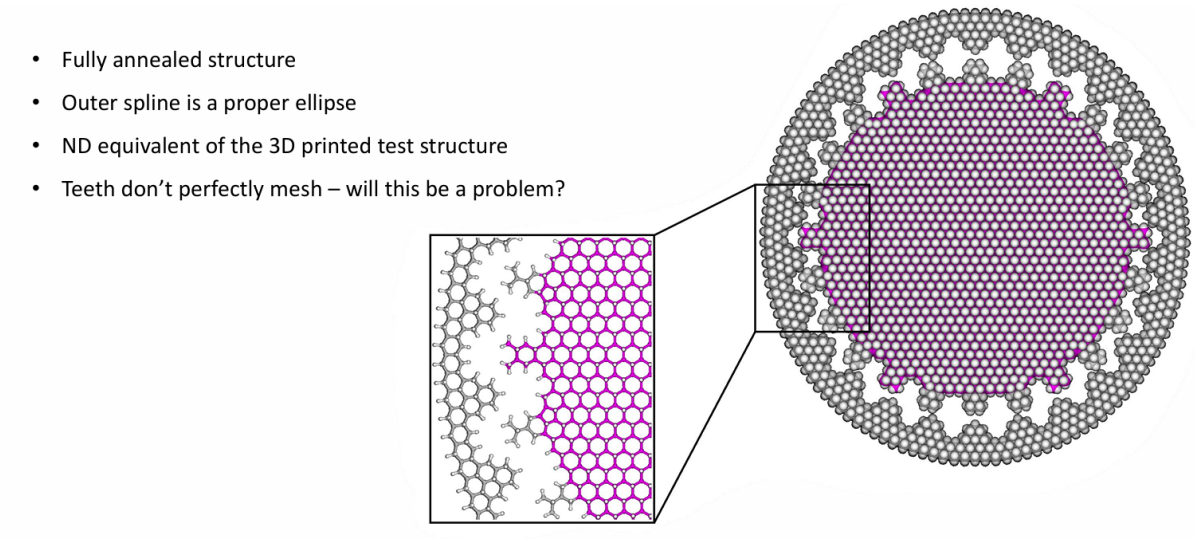


Figure 181: Assembled reversed harmonic drive and tooth-meshing geometry (slide courtesy of S. Haird).

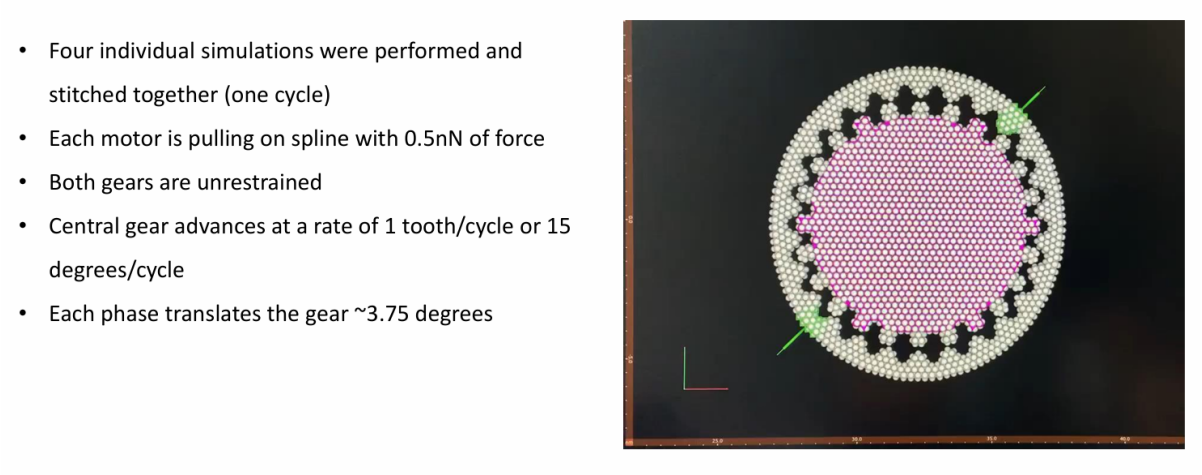


Figure 182: Simulated actuation of the reversed harmonic drive (slide courtesy of S. Haird; owned by Nanofactory CBN and reproduced with permission). In the reported four-phase simulation, the central gear advances approximately 3.75° per phase and 15° (one tooth) per complete cycle.

1) Withdraw the second indexer from between two teeth.
2) Do nothing.
3) Insert the second indexer between two teeth.
4) Do nothing.

The insert/withdraw motions of the two indexers are both driven by the first cable, but 180° out of phase. As a result, the gear is always locked in position. When the first indexer is inserted and the second indexer withdrawn, it is locked into position by the first indexer. When the second indexer is inserted and the first indexer withdrawn, it is locked into position by the second indexer.

During the cycle when the first indexer is being withdrawn and the second indexer is being inserted, the gear is locked into position provided we assume that it is locked into position if either indexer is more than half inserted. To provide some margin for error in the timing of the insertion and withdrawal of the two indexers, they should be designed so that the gear is locked into position as long as either indexer is more than, say, 30% inserted. Only when one indexer is more than 70% withdrawn is it clear of the gear. This way, small errors in timing of the insertion and withdrawal of the two indexers will never leave the gear entirely unlocked.

This approach can be used to drive a toothed wheel forwards by one tooth every clock cycle. When the four-phase clock cycle is put into reverse, the gear will also move in reverse, as each step in this cycle is itself reversible. A similar mechanism was described by Tom Moore (personal communication).

If we use the two “Do nothing” intervals to “Restore the second indexer to its position just prior to step 3” and “Drive the second indexer forwards by the distance between two teeth”, then we can move forwards (or backwards) by two teeth every clock cycle.

12.9 Design for a vacuum dielectric differential electrostatic comb drive

12.9.1 Classic comb drives

The classic comb drive involves tines separated by vacuum or air [LGE96, TNH89]. This can be scaled to an atomically precise design using either boron-doped diamond (per the available B from triethyl borate feedstock; the design shown is courtesy of Sheldon Haird) or an alternative design in which the tines are covered with graphene to provide electrical conductivity, while mechanical strength to keep the tines straight in the presence of van der Waals forces is provided by the diamond supporting structure (Figure 183).

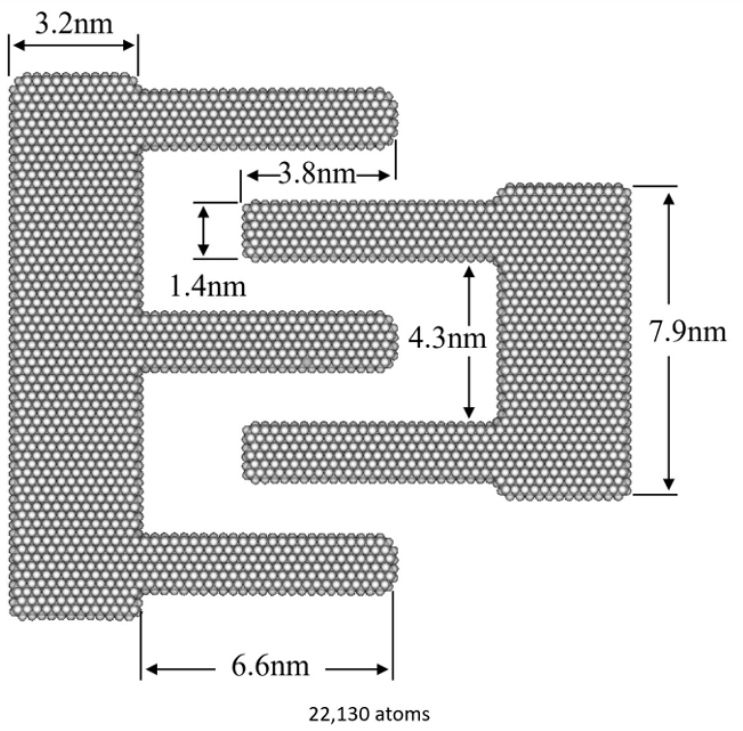


Figure 183: Classic electrostatic comb drive. Image owned by Nanofactory CBN; reproduced with permission.

Figure 184 shows the classic design with a specific set of parameters.

This design meets the baseline force, travel, and leakage requirements for the SRU; its atom count and volume remain major design constraints. When we assume a dielectric constant of 5.7 for diamond [Appb] and consider that the force generated by a comb drive is proportional to the dielectric constant, an idealized comparison at unchanged geometry predicts approximately 5.7 times the force with diamond rather than vacuum as the dielectric. This suggests size savings might be possible.

If we used diamond between the tines, the analysis of (a) the leakage current and breakdown voltage of diamond and (b) the drag when the tines slide over the dielectric during comb drive actuation would both be more complex. However, if we adopt vacuum as our dielectric, both of these issues become simpler to analyze, and the analysis itself is both easier to understand and less prone to error.

In the classic design, current through the 2 nm vacuum gap assumed below is small enough that

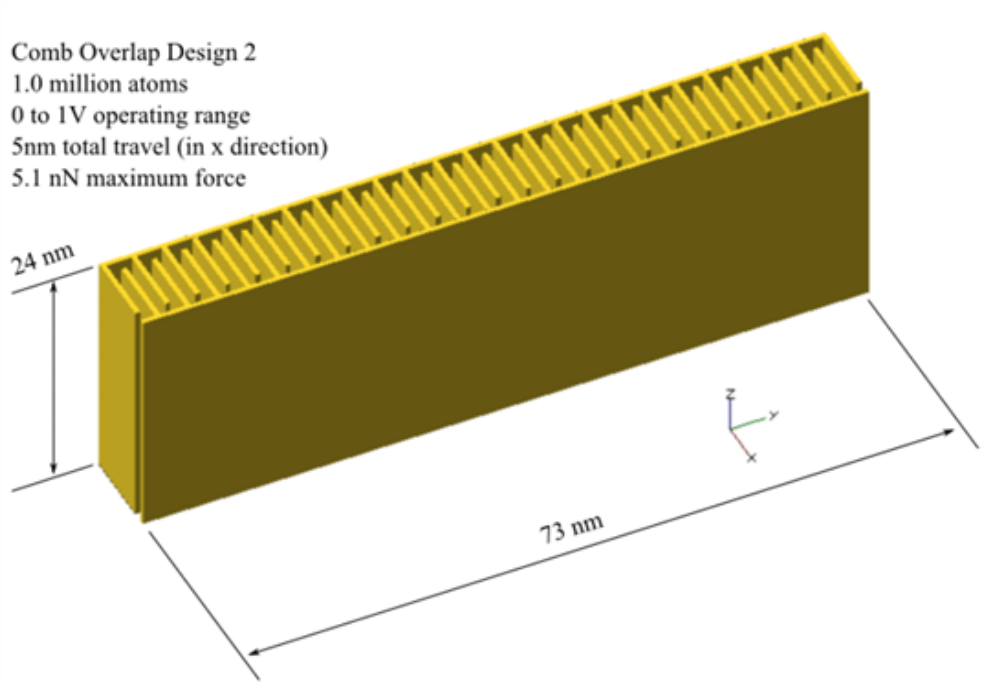


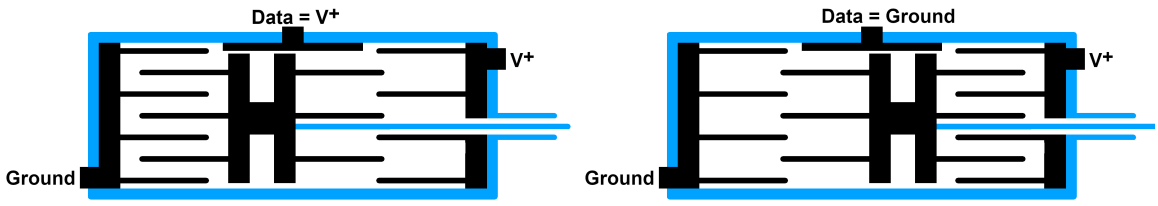
Figure 184: Classic comb drive with a specific set of design parameters. Image owned by Nanofactory CBN; reproduced with permission.

it can be neglected. As we assume atomically precise surfaces with no asperities, irregularities or particles to facilitate a conductive path from one plate to another, the theoretical tunneling current (which is substantially smaller than our leakage limit) should be applicable. To put it another way, if we design a classic eutactic comb drive with vacuum as the dielectric, the leakage current should be well within our requirements.

The final issue is how to convert this solid diamond design into AP Blocks. One general approach is to build the structural elements such as tines and the supports for the tines from AP Blocks whose surfaces are terminated with Ge radicals, like those illustrated in Molecular LEGO blocks using interfacial Ge (Section 7.6.7) and Simple AP Blocks using Ge (Section 7.13). The electrical components could be made from graphene and conductive nanotubes, such as the (4,4) nanotube (see Wires (Section 4.3.3)). Graphene sheets composed of several hundred carbon atoms could be attached to the structural elements by bonding radical sites on the graphene to the Ge radical sites on the supports.

Another approach would be to use dipolar bonds to link diamond blocks for structural elements (see Decomposing large structures into small blocks (Section 7.4)). Conductive elements might be composed of nanotubes and various graphene structures. For example, graphene strips with B- and N-terminated edges could be used (see Decomposition into graphene strips (Section 7.4.5)). An atomically precise design for a comb drive should be feasible. Further research to design and analyze specific atomically precise designs is needed.

The cartoon in Figure 185 is for an atomically precise design which is described in the following paragraphs.



(a) (b)

Figure 185: Atomically precise differential comb drive displacement states: (a) Data = V+; (b) Data = Ground.

12.9.2 Direct tunneling between vacuum-separated graphene electrodes

For two clean graphene sheets separated by a vacuum gap in UHV, the leakage mechanism at sub-nanometer to nanometer separations is expected to be direct quantum tunneling. The dominant material parameter is the graphene work function. Measurements and analyses of graphene commonly place the work function near that of graphite, approximately 721 zJ (4.5 eV) – 737 zJ (4.6 eV). Yu et al. [YZR+09] note a graphite-like value near 737 zJ (4.6 eV) as a useful reference point.

The current density was estimated using a Simmons/WKB direct-tunneling model. Simmons [Sim63] derived the standard thin-barrier tunneling expression for electrodes separated by an insulating film, including the case of an arbitrary barrier profile; a vacuum gap can be treated as the limiting case in which the insulating barrier height is set primarily by the electrode work function. For a first estimate, the barrier is modeled at zero bias as a rectangular barrier of height

$$\Phi \simeq W_G \simeq 721.0\,\mathrm{zJ} \quad (4.5\,\mathrm{eV}). \tag{97}$$

The barrier becomes trapezoidal when tilted by the applied voltage V across a gap d. Neglecting image-force lowering, atomic-scale roughness, and momentum-selection effects, the intermediate-bias Simmons form may be written approximately as

$$\begin{aligned} J \approx \frac{e}{2\pi h d^2}\Bigg[&\left(\Phi - \frac{eV}{2}\right)\exp\left(-\frac{4\pi d}{h}\sqrt{2m\left(\Phi - \frac{eV}{2}\right)}\right) \\ &- \left(\Phi + \frac{eV}{2}\right)\exp\left(-\frac{4\pi d}{h}\sqrt{2m\left(\Phi + \frac{eV}{2}\right)}\right)\Bigg], \end{aligned} \tag{98}$$

where e is the elementary charge, h is the Planck constant, d is the vacuum separation, Φ is the barrier energy, m is the free-electron mass, and V is the applied bias.

Equation (98) should be regarded as an order-of-magnitude tunneling estimate, not as a substitute for an atomistic nonequilibrium Green's-function calculation.

Using $\Phi = 721.0$ zJ (4.5 eV), Table 18 summarizes estimated leakage current densities for the listed gaps and applied biases.

Table 18: Estimated tunneling current across vacuum gaps at selected applied biases.

Vacuum gap d	**Applied bias** V	**Electric field**	**Estimated** J	**Current per nm^2**
0.5 nm	1 V	2 V/nm	$\sim 2 \times 10^5$ A/cm^2	$\sim$2 nA
1.0 nm	1 V	1 V/nm	$\sim$3 A/cm^2	$\sim$30 fA
2.0 nm	2 V	1 V/nm	$\sim 10^{-8}$ A/cm^2	$\sim 10^{-22}$ A

These values are best interpreted logarithmically. The exponential sensitivity to distance is the main result. For a 721.0 zJ (4.5 eV) vacuum barrier, the WKB decay constant is approximately

$$\kappa = \frac{\sqrt{2\, m\, \Phi}}{\hbar} \simeq 10.9\,\mathrm{nm}^{-1}, \tag{99}$$

so the transmission factor scales roughly as

$$T \sim \exp(-2\,\kappa\, d). \tag{100}$$

Consequently, reducing the vacuum gap from 1.0 nm to 0.5 nm increases the tunneling probability by approximately

$$\exp\left(2\kappa(0.5\,\mathrm{nm})\right) \sim 5 \times 10^4. \tag{101}$$

This exponential dependence also implies that an uncertainty of only 0.1 nm in the effective separation changes the predicted current by roughly a factor of

$$\exp(2\kappa\,(0.1\,\mathrm{nm})) \sim 9. \tag{102}$$

Thus, the estimate is not precise at the atomic scale unless the separation and electrode geometry are specified with comparable precision.

The calculated values omit image-potential lowering of the vacuum barrier, which would tend to increase the current. They should therefore not be interpreted as conservative upper bounds on leakage. Monolayer graphene near charge neutrality has finite quantum capacitance, so some of the applied voltage can appear as a shift of the graphene chemical potentials rather than as an electrostatic drop across the vacuum gap. For gaps of 0.5–2.0 nm, this correction is typically a factor-of-order-unity effect and does not change the qualitative conclusion. The unmodified-barrier Simmons model predicts leakage in the tens-of-femtoamps-per-square-nanometer range for a 1 nm vacuum separation at 1 V and approximately 10^{-22} A/nm^2 for a 2 nm separation at 2 V. The robust conclusion is the extreme sensitivity of leakage to the effective gap and barrier profile; a more complete calculation including image-potential effects would be required to establish the actual leakage current.

As the total area of all four differential comb drives (see Electrostatically active area of four differential comb drives (Section 12.9.7)) is 1,920 nm^2, and our estimated leakage current is 10^{-22}

A/nm^2, our estimated total leakage current per SRU will be 1.9×10^{-19} A. $P = VI$, giving us about 3.8×10^{-19} watts of lost power from leakage current per SRU. As our power budget per SRU is 10^{-9} watts, this is less than a billionth of our power budget.

In summary, vacuum-separated graphene electrodes provide a comparatively well-defined limiting case for leakage estimation. For 2 V across a 2 nm gap, leakage should be negligible.

12.9.3 Notation

Table 21 summarizes the symbols and units used below.

One differential comb drive consists of two opposing comb actuators acting on one shared shuttle. Four differential comb drives are installed; one pair is enabled at a time, either S1 and S2 or D1 and D2. Only one drive in the selected pair changes state during each clock phase. Here, active refers to membership in the selected pair, not simultaneous switching. An active pair therefore means two complete differential comb drives.

The notation used in the equations for the various dimensions and materials parameters is shown in the diagram in Figure 186, which illustrates a MEMS scale comb drive using vacuum as its dielectric. With that notational issue out of the way, we proceed with the discussion.

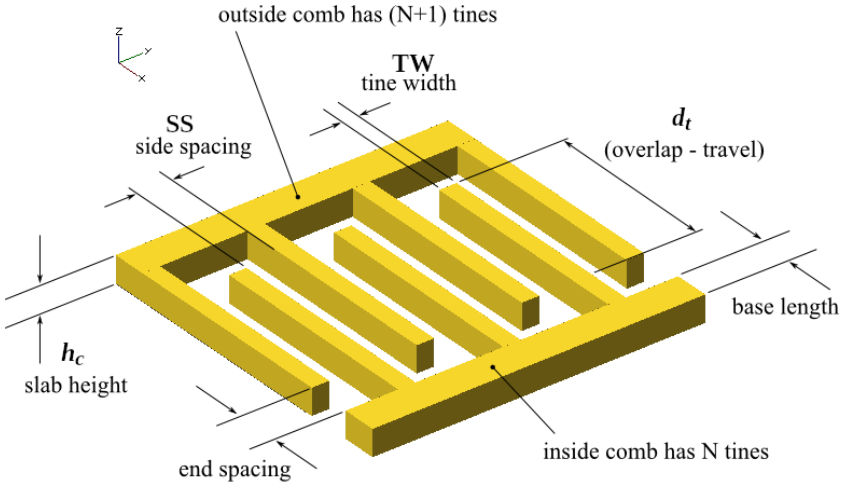


Figure 186: Comb drive dimensions and their names. Drawing repeated from Figure 131.

12.9.4 The design

This alternative design is for a differential comb drive driven by 2 volts, with a 2 nm shuttle throw and a force of 1 nN.

The Side Spacing (SS) is the distance between tines, that is, the distance between the facing conductive coverings of the diamond tines. A continuous carbon nanotube around each tine is a conceptually simple example; the choice and construction of the conductive covering remain open. The SS is 2.0 nm. The Tine Width (TW) of each tine is 2.5 nm, including the conductive covering. As $N = 6$, the inside comb has 6 tines and the outside comb has 7 tines. There are $2 \times N = 12$

insulating vacuum layers between tines that are SS = 2.0 nm in thickness. This makes the comb drive total thickness:

$$\begin{aligned}\text{Total thickness} &= 12\,\text{SS} + (6+7)\,\text{TW} \\ &= 12(2\,\text{nm}) + 13(2.5\,\text{nm}) = 56.5\,\text{nm}.\end{aligned} \tag{103}$$

The slab height h_c is in a direction at right angles to the thickness. We can set the height of the comb drive to whatever value is convenient for our design. We set the height h_c of the comb drive to 10 nm.

The length is at right angles to both the height h_c and the thickness. The total length changes as the comb drive extends and contracts.

The actuator throw is denoted by d_t (labeled "Throw" in Figure 186). The length consists of the following components (with $d_t = 2\,\text{nm}$):

> The overlap is 3 nm when the comb drive is in its contracted position, 2 nm of throw plus 1 nm of margin.
>
> The overlap is 1 nm when the comb drive is in its extended position because of the 1 nm of margin.
>
> The end spacing is 1.5 nm (when the comb drive is in its contracted position). There are 2 ends, which doubles the contribution of the end spacing to the total length.
>
> The base length is 3.5 nm. The base length contributes twice to the total length, once at each end.

The total length in the contracted position (going sequentially through the lengths): 3 nm (overlap) + 2 × 1.5 nm (twice the end spacing) + 2 × 3.5 nm (twice the base length) = 13 nm.

The total length in the extended position is the length in the contracted position + the throw. This is 13 nm + 2 nm = 15 nm.

The dimensions, throw margin, and tine count are collected in Table 19.

In summary, the length is 13 nm (when contracted), the height is 10 nm, and the thickness is 56.5 nm. This gives a total volume of 7,345 nm^3 for a single comb actuator when contracted. When extended it would occupy 8,475 nm^3. A single differential comb drive would occupy 7,345 nm^3 + 8,475 nm^3, as one opposing comb actuator will be contracted and the other extended, giving 15,820 nm^3. **Four such differential comb drives would occupy 63,280 nm^3.**

12.9.5 Parameter values

The following are the parameter values used to define these four electrostatic differential comb drives. Here V_{comb} is the volume of a single comb actuator, and V_{drive} is the volume of one differential comb drive.

12.9.6 Electrostatic equations and force

For one opposing comb actuator, the contracted and extended overlaps are 3 nm and 1 nm, respectively. The corresponding active areas and area change are:

Table 19: Vacuum differential comb-drive inputs and outputs, using the same columns as the baseline comparison in Table 13.

Parameter	**Symbol**	**Value**	**Role / assumption**
Side spacing	SS	2.0 nm	Input; vacuum gap
End spacing	—	1.5 nm	Input
Base length	—	3.5 nm	Input
Tine width	TW	2.5 nm	Input; includes sheath
Slab height	h_c	10 nm	Input
Relative permittivity	ϵ_r	1	Vacuum assumption
Vacuum permittivity	ϵ_0	8.85×10^{-12} F/m	Physical constant
Voltage	V	2 V	Input
Tine count	N	6	Input; 7 outer tines
Actuator thickness	—	56.5 nm	Calculated
Throw margin	—	1.0 nm	Input
Shuttle throw	d_t	2 nm	Input; no transmission
Raw shuttle force	F	~ 1.06 nN	Calculated
Actuator length	—	13–15 nm	Contracted–extended
Volume per actuator	V_{comb}	7,345–8,475 nm^3	Contracted–extended
Volume per drive	V_{drive}	$\sim$ 15,820 nm^3	Two opposing actuators
Capacitance per drive	C	2.1 aF	Calculated

$$A_{\text{contracted}} = h_c(3\,\text{nm})2N = 360\,\text{nm}^2 \tag{104a}$$
$$A_{\text{extended}} = h_c(1\,\text{nm})2N = 120\,\text{nm}^2 \tag{104b}$$
$$\Delta A_{\text{comb}} = h_c d_t 2N = 240\,\text{nm}^2. \tag{104c}$$

Evaluating Eq. (56) for the vacuum-design inputs gives:

$$\begin{aligned} F &= \frac{N\,V^2\,\epsilon_r\,\epsilon_0\,h_c}{\text{SS}} \\ &= \frac{6 \times (2\,\text{V})^2 \times 1 \times (8.85 \times 10^{-12}\,\text{F/m}) \times (10 \times 10^{-9}\,\text{m})}{2.0 \times 10^{-9}\,\text{m}} \approx 1.06 \times 10^{-9}\,\text{N}. \end{aligned} \tag{105}$$

12.9.7 Electrostatically active area of four differential comb drives

Using the contracted and extended areas in Eqs. (104a) and (104b), the electrostatically active area of one differential comb drive is $A_{\text{drive}} = A_{\text{contracted}} + A_{\text{extended}} = 360\,\text{nm}^2 + 120\,\text{nm}^2 = 480\,\text{nm}^2$. The two drives in the selected clock pair have $A_{\text{active}} = 2A_{\text{drive}} = 960\,\text{nm}^2$. **All four installed differential comb drives have** $A_{\text{installed}} = 4A_{\text{drive}} = 1{,}920\,\text{nm}^2$. This last value is the total area used in the leakage estimate for the installed set.

12.9.8 Capacitance of differential comb drives

The capacitance of a single differential comb drive, using Eq. (95) with $A = A_{\text{drive}}$, is:

$$C = \frac{\epsilon_0 \epsilon_r\,A_{\text{drive}}}{\text{SS}} = \frac{(8.85 \times 10^{-12}\,\text{F/m})\,(480 \times 10^{-18}\,\text{m}^2)}{2 \times 10^{-9}\,\text{m}} \approx 2.1 \times 10^{-18}\,\text{F}. \tag{106}$$

Here $\epsilon_r = 1$ is the relative permittivity of vacuum, $\epsilon_0 \approx 8.85 \times 10^{-12}\,\text{F/m}$ is the vacuum permittivity, the area A_{drive} of the two opposing comb actuators within one differential comb drive was derived in the previous section as 480 nm^2, and the spacing between plates is the side spacing SS specified as 2.0 nm for this vacuum-dielectric alternative.

The combined capacitance of the two differential comb drives in the selected clock pair is 4.2 aF.

There are four differential comb drives, but only two of them are active at any given time.

12.9.9 Electronic time constant and heat

The resistance quantum $R_K \approx 25{,}812\,\Omega$, introduced in Eq. (59), gives a characteristic scale for resistances through atomically precise nanojunctions.

A molecular circuit might involve a few such junctions, making 100 kΩ ($10^5\ \Omega$) a plausible resistance. This gives an RC time constant of:

$$\tau = RC \approx (4 \times 25{,}812\,\Omega)(2.1 \times 10^{-18}\,\text{F}) \approx 0.22 \times 10^{-12}\,\text{s}, \tag{107}$$

or less than one picosecond. This is three orders of magnitude margin compared with our requirements of ∼one nanosecond, as we require a bandwidth comfortably above our clock frequency of 100 MHz.

Leakage current across a 2.0 nm vacuum gap at two volts is negligible.

12.9.10 Summary of vacuum electrostatic differential comb drive

This electrostatic differential comb drive, four of which could replace the main h-BN differential comb drives that drive the SRU in the baseline design, has a 2 nm throw and can apply a 1 nN force. Its 2 V operation, in place of the baseline 0.19 V, would require a revised supply or voltage boosting (Section 12.2.6). Two differential comb drives are active at a time: either (a) S1 and S2 or (b) D1 and D2. All four are never active at the same time. Leakage current at two volts between tines over a 2.0 nm gap is negligible. Capacitance of a single differential comb drive is 2.1×10^{-18} F. A typical resistance of four atomically precise electrical connections in series might be 100,000 Ω. The resulting RC time constant (electronic switching time) is less than one picosecond, much faster than the approximately 1,000 ps electrical-response requirement for operation at 100 MHz. The two active differential comb drives have a combined total capacitance of 4.2×10^{-18} F. This capacitance does not change during actuation of the comb drive but instead remains constant. All four differential comb drives have a combined total volume of ∼63,280 nm^3. Adding an allowance to account for supporting structures, the total volume for four vacuum-dielectric comb drives proposed as an alternative to run an SRU should be less than 80,000 nm^3.

12.10 Power for a daughter

12.10.1 Batteries

Batteries add additional design and implementation overhead but might offer advantages in some scenarios. They are not in the baseline, but might be useful both as an opportunity and a risk mitigation strategy. Battery operation would likely be required for less than a second, just enough time for the daughter SRU to attach its three feet, each making contact with two pillars, to the six contact pillars, after which the contact pillars would provide power to the daughter.

A mechanical battery has low energy density, but would work reliably and be relatively simple to design and manufacture. It could be manufactured in the fully wound state and ejected from the daughter SRU after use, allowing recovery of the space that it occupied.

A candidate one-shot bootstrap source might be a hydrogen–fluorine fuel cell. At 210 K both H_2 and F_2 are gases, while the HF reaction product is liquid rather than solid. The fluorine and HF chamber could be lined with a fluorine-terminated Si-up SiC(111) surface, which is remarkably chemically inert. An atomically precise membrane containing negatively charged proton-conducting pores would transport protons from the hydrogen compartment while excluding fluoride ions. The principal remaining uncertainties are the design and stability of the membrane's active sites, electron transfer at the fluorine electrode, and prevention of liquid HF from obstructing the reaction interface.

12.11 Extrusion Mechanism

12.11.1 Staggered pegs

An alternative approach, available for risk mitigation, would be to adopt the method proposed for the Merkle–Freitas Hydrocarbon Molecular Assembler [FM04b, Sec. 4.11.3.4, item 10]. In this method, we start with a series of pegs inserted into small holes around the shell of the daughter. By exerting force on the pegs, the shell can be moved incrementally out of the extrusion port. Each peg can then be moved to another hole in the shell that is closer to the parent. That is, a ring of holes around the shell is filled with pegs, and then the subsequent rings of holes, initially empty, are available for the pegs to be moved to as the daughter is gradually extruded from the parent.

In more detail, each peg is pulled outward through the extrusion port by a spring attached to the inside shell of the parent. Each spring is weak enough that the peg to which it is attached can be pulled back by the internal assembly arm of the parent, but is strong enough that, taken collectively, the sum of the spring forces will reliably move the daughter out of the parent through the extrusion port. Initially, each peg rests against a stop on the parent that prevents the daughter from sliding further out the extrusion port, even though it is being pulled by a spring attached to the parent. Then, the internal assembly arm of the parent successively moves each peg to the next hole, where it is no longer against the stop. At some point, the force pulling the daughter out is large enough to exceed the force pushing it in, at which point it wants to slide out, but is prevented by the remaining pegs that are up against the stops. The remaining pegs, which are now preventing the daughter from sliding out, continue to be pulled out and put in their new holes. Eventually, all the pegs have been placed in their new holes, and the daughter slides out one "step" into its new position.

When only the last peg is holding the daughter in place, the daughter might be "tipped" in the extrusion port. If this is an issue, a more complex design that locks the daughter in place until all the pegs have been moved and then abruptly unlocks the daughter at multiple points (rather than a single point), allowing the daughter to slide smoothly and without tilting, could be used.

It is worth noting that the small holes in the shell of the daughter will not cause leaks during extrusion. This is because the extrusion sleeve extends long sidewalls after the parent SRU has been extruded from *its* parent. As long as the holes in the shell are shorter than the length of the sidewalls of the extrusion sleeve, liquid entering the holes will not be able to enter the parent but will remain confined in the holes. We refer to these sidewalls as the "Extrusion Sleeve" or ES for short.

A major purpose of the extrusion mechanism is to maintain a vacuum tight seal during the extrusion process. The entry of any molecules of gas or liquid into the interior of the SRU is highly undesirable, and the probability of such entry during a replication cycle should be kept very low. For more on this topic, see Sliding Seal Extrusion Risk Mitigation (Section 12.12).

12.11.2 The MacArthur Mechanism

A second alternative mechanism, also available for risk mitigation, might adapt MacArthur's atomically precise proposal for a demultiplexer (Section 4.3.6). In MacArthur's proposal, two rods labeled A and B with knobs on them are rotated, extended and retracted in a specific pattern that moves an orthogonal rod lying across them either forwards or backwards along their length.

To adapt this approach, take 2N rods, divided into A rods and B rods, and distribute them around the base of the daughter to be extruded. Assume the A and B rods only have knobs at one location

along their lengths (although it would be possible to have a few knobs, if this was thought desirable) unlike the A and B rods in the demux, which had many knobs along the entire long length of their A and B rods. That is, to extrude the shell we use short rods rather than the long rods used in the demultiplexer application. When the A or B rods rotate, their knobs can rotate into transverse slots in the shell of the daughter, locking the rod and the shell together. There are many slots along the length of the shell. When the knobs are rotated out of the slots in the shell the rod is not locked to the shell and is free to move independently of it.

Start with the A-rod knobs locked into their slots on the shell, and the B-rod knobs free of their slots. The A-rods are in their "back" position, while the B-rods are in their "forward" position.

1. The A rods move forward and the B rods move back. The A rods are engaged, so the daughter slides out one step, while the B rods are disengaged, so their movement back does not move the daughter.
2. The A and B rods both rotate 90°. This causes the A rods to disengage, and the B rods to engage.
3. The A rods move back and the B rods move forward. The B rods are engaged, so the daughter slides out one step, while the A rods are disengaged, so their movement back does not move the daughter.
4. The A and B rods both rotate -90°. This causes the A rods to engage, and the B rods to disengage.

Repeat steps 1-4 as needed to extrude the shell.

It is worth noting that the MacArthur mechanism has wide utility in a variety of applications. It is fully reversible conceptually, and with proper design and reduced speed can approach thermodynamic reversibility as closely as might be desired. That is, there are no fundamentally irreversible steps in its operation, nothing that "snaps" or "clicks", no springs or latches, just continuously and smoothly moving mechanical operations. We can have many knobs and a short moving element that comes into contact with only one or a few knobs (as in the demux), or we can have one or a few knobs that come into contact with a long moving element (as in the example of the daughter being extruded from the parent). Either works, and the number of knobs in contact with the moving element can be adjusted as might be desired.

12.12 Sliding Seal Extrusion Risk Mitigation

In the SRU concept the daughter is extruded from the parent during a 10^5 second (∼28 hour) replication cycle. This extrusion process ideally assumes that a sliding seal exists between the shell of the extruded daughter and the extrusion sleeve of the parent (Figure 187).

12.12.1 Analysis of sliding seal risk when using hexagonal blocks

Hexagonal blocks incorporating B and N should be able to tile the plane, as illustrated by the 19 identical hexagonal AP Blocks shown assembled in Figure 188. See Decomposition of hexagonal diamond into hexagonal blocks (Section 7.4.4). A single block is linked by 24 dipolar bonds to its neighboring blocks. Assuming each dipolar bond is 260 zJ (1.6 eV), 24 such bonds would result in each block being bound into the wall by 6,240 zJ (39 eV).

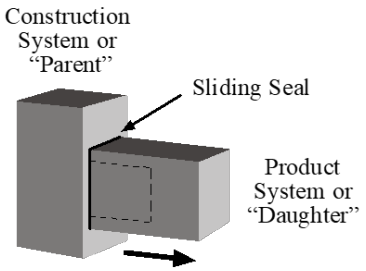


Figure 187: Sliding seal between the daughter shell and the parent extrusion sleeve.

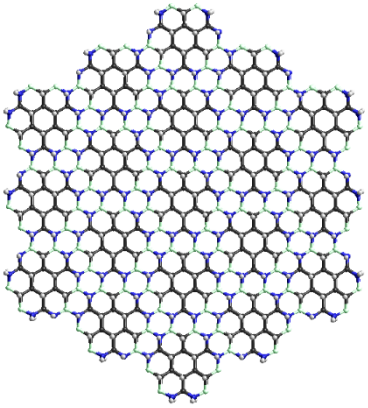


Figure 188: Nineteen hexagonal blocks assembled into a wall segment. Drawing repeated from Figure 46.

See Leakage of H_2 through a sliding seal (Section 7.4.4.1) for the analysis of sliding seal leakage when using shells composed of such blocks. The conclusion is that the estimated failure rate per SRU per replication cycle from this cause should be less than 3.2×10^{-11}. This more than meets our failure rate requirement of <0.1% per replication cycle (see SRU Failure Handling (Section 6.2)). This cause of SRU failures should not be a significant risk.

Larger contaminant molecules should have a significantly lower probability of entering through the sliding seals.

12.12.2 Other candidate blocks

Analysis of sliding seal leakage for other candidate blocks proposed for the shell is a topic for future research.

One such block would use dipolar bonds on four sides of a rectangular block. A view of four such blocks, assembled into a four block unit by dipolar bonds, is shown in Figure 99. See BNO Molecular LEGO that bond on four sides (Section 7.9) for further details.

Five LEGO like AP Blocks (see Molecular LEGO (Section 7.6)) are shown in Figure 189 and could also be used for the shell.

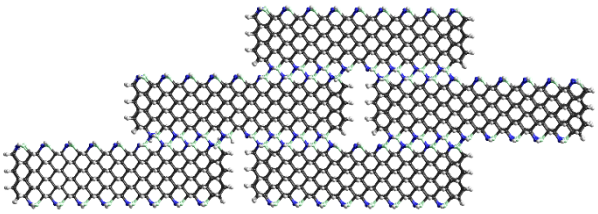


Figure 189: Five molecular LEGO blocks assembled into a wall. Drawing repeated from Figure 42.

LEGO like AP Blocks can be assembled much like LEGO bricks. The example in Figure 190 shows LEGO bricks assembled into a wall, while the LEGO like AP Blocks in Figure 189 have likewise been assembled into a wall.

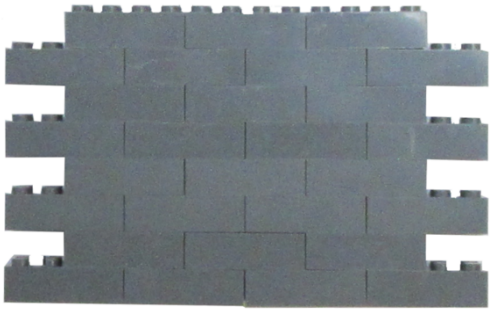


Figure 190: LEGO bricks illustrating wall assembly. Image repeated from Figure 80.

LEGO bricks can be used to illustrate a design for a shell and extrusion port (the extrusion sleeve is not shown in this design exercise), and to show the daughter being extruded from the extrusion port (Figure 22).

The following risk mitigation strategies were developed prior to the sliding seal analysis shown above, which suggests these mitigation strategies might not be required. They are presented here

because (1) they deal with this risk should it be realized, (2) they show design alternatives that could prove useful for dealing with other risks, and (3) they will be useful in providing higher performance alternatives in future designs.

The following strategies are presented in no particular order.

12.12.3 First strategy

The replication cycle of an SRU, 10^5 seconds, gives sufficient time to cycle the die between feedstock solution when the SRUs are absorbing feedstock, and vacuum when the daughters are being extruded. This immerses each SRU in feedstock part of the time, and puts it in vacuum part of the time.

We keep the existing SRU architecture but add an internal tank to the SRU that can hold enough feedstock to enable the construction of a daughter SRU. This is possible because the internal volume of an SRU can be made as large as might be desired by sufficiently extending its longest dimension.

The outside cross-section of an SRU is 100 nm × 60 nm. Assuming walls approximately 1 nm thick, the corresponding internal cross-section is approximately 98 nm × 58 nm = 5,684 nm^2. Extending the SRU by one nanometer therefore adds approximately 5,684 nm^3 of tank volume. Tetraethylgermane has a density of approximately 3.16 molecules/nm^3 and contains eight carbon atoms per molecule, giving approximately 25.28 carbon atoms/nm^3. Each nanometer of extension therefore adds approximately 5,684 × 25.28 ≈ 143,692 carbon atoms to the stored feedstock. Extending the two 100 nm walls and two 60 nm walls requires approximately 320 nm^3 of diamond, or 320 × 176 = 56,320 carbon atoms. The net carbon inventory added per nanometer of extension is therefore approximately 87,372 atoms. On this simplified carbon-equivalent calculation, storing enough feedstock to construct a 100-million-atom SRU requires approximately 100,000,000/87,372 ≈ 1,145 nm of additional length, giving a total SRU length of approximately 1,645 nm.

Note that this 1,645 nm length already takes into account the additional carbon atoms needed to build the tank itself.

The resulting system operates in two phases, which keep alternating until all sites on the die are occupied by SRUs.

1. **Feedstock ingestion phase:** We first flood the dies with feedstock. During this phase, each SRU pumps feedstock into its internal tank. The tank is big enough to hold all the feedstock required to build an SRU.

2. **SRU construction and extrusion phase:** We pump down the chamber that holds the dies, producing a vacuum. Each SRU then manufactures and extrudes a daughter SRU, using the feedstock from its internal tank. These daughter SRUs are then attached to the die surface.

This architecture reduces the difficulty of the task of the sliding seal. During extrusion, the sliding seal must prevent anything from entering the SRU from the external environment. If the external environment is a good quality vacuum, the rate at which contaminants enter the system should be reducible to a tolerable rate.

If the rate at which contaminants can enter the SRU is sufficiently reduced, we can operate in a regime where the probability of contaminants entering the internal SRU environment during

a replication cycle falls below the threshold that allows operation for the duration of multiple replication cycles without the entry of any contaminants.

This approach requires a chamber over the dies that can both hold feedstock solution and be pumped down to a good quality vacuum, alternating between the two states as required.

If sliding seals leak, this approach should effectively mitigate the leakage. However, it requires a larger SRU (1,645 nm in length, as compared with 500 nm), takes more time for a replication cycle, reduces the number of SRUs that can fit on the die surface, and requires more support equipment (a chamber that can alternate between vacuum and liquid feedstock).

12.12.4 Second strategy

This strategy uses multistage sliding seals combined with pumps between stages. Each stage would pump into the preceding stage, thereby maintaining a higher quality vacuum than the preceding stage. Additional stages would be added until the desired quality of vacuum was achieved. This strategy is employed successfully on the macroscale and should be adaptable to our case at the molecular scale. The major disadvantage is the requirement for active pumping (Figure 191).

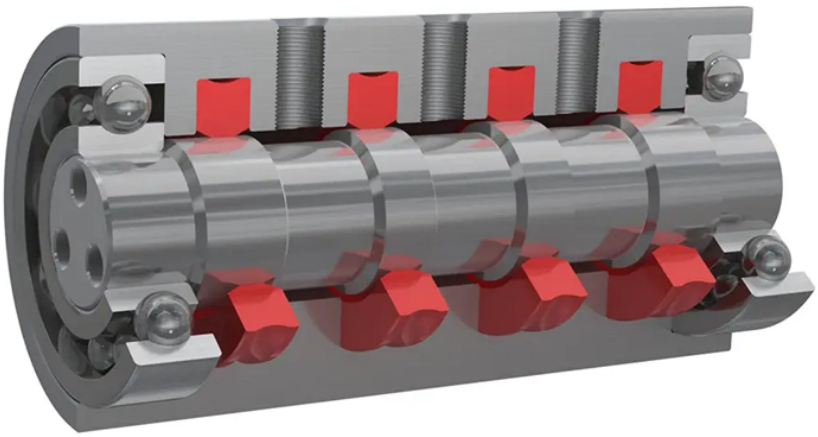


Figure 191: Commercial multistage rotary seal. Image source: AHP Seals, “Swivel Seals” [AHP].

Figure 192 shows a multistage seal concept for daughter extrusion.

12.12.5 Third Strategy

The interior of the SRU would be filled with a pressurized inert gas, such as helium. During extrusion of the daughter, any tendency to leak into the interior would be countered by the flow of the inert gas in the opposite direction. This strategy requires a pump in the SRU to pump the inert gas into the SRU, and a filter to pass the inert gas and reject everything else.

12.12.6 Fourth Strategy

In this strategy, we keep the interior of the fabricator in vacuum and isolate it by an airlock. Blocks are exported from the fabricator through the airlock. The rest of the SRU is flooded. Only the parts of the shell required for structural support are retained; the rest are eliminated.

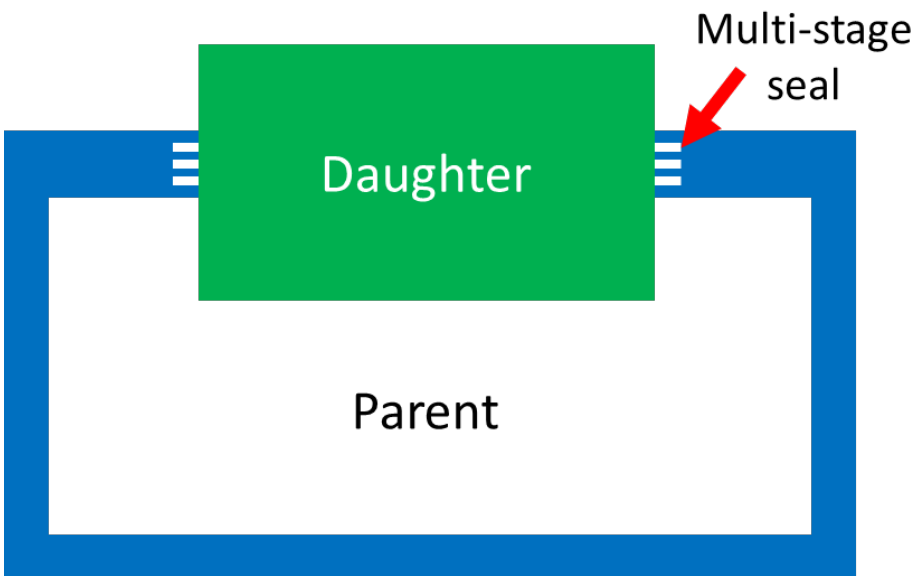


Figure 192: Multistage seal concept for daughter extrusion.

As the build volume, and therefore the size of the airlock(s), will be limited (the build volume might be 15 × 15 × 15 nm), this strategy should not significantly increase the volume of the SRU, even if we need a series of airlocks, each with better quality vacuum, to export blocks from the interior of the fabricator and make them accessible to the interior assembly arm. Exposing blocks to feedstock might exclude the design option of using AP Blocks that had surfaces that reacted with feedstock, which would be a drawback of this approach.

An optimization that this strategy might enable would be the elimination of either the internal or the external assembly arm (but not both), as the barrier between the two could be removed, allowing one of them to carry out the functions of both.

A potential advantage of this strategy would be to open a wider range of possible architectures for a nanofactory. Only enough of the shell would need to be retained to seal off the fabricator and for mechanical support of the major components in the interior of the SRU, such as the base of the internal and external assembly arms, the feet, the demultiplexer, the daughter extrusion port, and the like. As most of this mechanical support function has already been taken over by the space frame, most of the shell and a significant part of the mass of the SRU could be eliminated.

The fabricator itself, if manufactured from AP Blocks in solution, would have to be pumped down once it had been made vacuum tight.

However, a flooded SRU might rule out fusion of blocks as an assembly option. The reactive block surfaces required for block-block fusion would be directly exposed to the feedstock. Unless the feedstock was designed so that it did not react with the reactive surfaces of the blocks, this feedstock-block reaction might cause problems. If dipolar bonds between nitrogen and boron were sufficiently unreactive with the feedstock solution, one could imagine bonding blocks together with dipolar bonds when they were immersed in feedstock solution. However, the use of radical coupling reactions, using, for example, carbon, silicon or germanium radicals, would be more problematic. Of concern would be the ability to make wires and graphene sheets. Wires, by definition, must be conductive. Long, thin conductive structures are often continuously bonded, implying the use of non-bonded assembly methods might be inconvenient for electrically conductive structures such

as wires. It would be convenient to make wires from segments of (4,4) carbon SWNT. Several options are available, including synthesis of short segments ∼1 nm in length (64 C atoms, see Wires (Section 4.3.3)) or even 0.5 nm (32 C atoms). Such short segments might be exported from the fabricator as usual if the unterminated ends do not react with the feedstock solution. Alternatively, the C atoms at the ends of the short segments could be replaced with B and N. Complementary B and N atoms at the ends of the short segments would allow them to be assembled if they do not react with the feedstock.

The flooded SRU proposal differs significantly from the Freitas biphase assembler [FM04a]. In that proposal, the fabricator floats freely in solution, manufactures parts (corresponding to AP Blocks in the SRU) and extrudes them into solution. The parts are then assembled in solution by an assembler, also floating in solution, into product. The parts must be soluble, imposing a constraint on the space of allowable parts, and must be designed so they cannot self-assemble while floating in solution, imposing a second constraint on the parts. Finally, it must be possible to identify and bind to each part uniquely in solution, distinguishing it from all other parts regardless of orientation when it encounters the binding site, imposing a third constraint on their design space. None of these constraints are present in the flooded SRU approach, as AP Blocks are made by mechanosynthesis in vacuum (as in the Freitas biphase assembler) but are then not released freely into solution but remain positionally controlled when transferred from the fabricator through the airlock to the interior assembly arm of the SRU. They are bound by a gripper, but the same gripper might be used on multiple AP Blocks, as the blocks are distinguished from each other by their position and orientation – which is controlled, rather than random. AP Blocks are then positionally controlled during assembly and never have an opportunity to clump together at random in an uncontrolled fashion.

Although the flooded SRU is a possible alternative, we do not pursue it in the present paper. We simply point out that this possibility exists and might significantly reduce the atom count of the SRU. However, an "airlock" mechanism to allow blocks to be moved out of the vacuum inside the fabricator and into the feedstock now internal to the SRU would need to be added, and strategies for dealing with block assembly in the presence of feedstock would be required. The design of certain mechanical components, such as the differential comb drives, would also have to be reviewed either to ensure correct operation in liquid or to provide localized vacuum.

12.12.7 Fifth Strategy

In this strategy, we maintain the surface of the die in vacuum at all times, and pipe feedstock to the SRUs through nanofluidic channels etched into the surface of the die. We outline the general issues involved in this approach in the following section.

12.12.7.1 E-beam lithography

E-beam lithography (EBL) is a high-resolution and precise lithographic technique used to create nanoscale patterns, such as nanofluidic channels, on various substrates, including silicon dies. Etching and bonding, and sacrificial-layer methods, offer candidate fabrication approaches [WNC+17]. The following is an example of an etched-channel process; the materials and sealing steps must be chosen for the intended substrate and feedstock:

1. Substrate preparation: The substrate, typically a silicon die, is cleaned to remove contaminants, and a thin layer of an electron-sensitive resist material (either positive or negative resist) is spin-coated onto the surface.

2. Exposure: The coated substrate is placed in an e-beam lithography system, where a focused beam of electrons is scanned across the resist layer, following a predetermined pattern corresponding to the desired nanofluidic channel design. The electron beam causes chemical changes in the resist material, either breaking down the polymer chains (positive resist) or crosslinking them (negative resist).

3. Development: After exposure, the substrate is immersed in a developer solution. Development removes exposed regions for positive resist or unexposed regions for negative resist, revealing the patterned mask. Openings in the mask expose the substrate where nanofluidic channels are to be etched.

4. Etching: The patterned resist layer serves as a mask for etching the nanofluidic channels into the substrate. This can be done using either wet or dry etching techniques, such as reactive ion etching (RIE) or deep reactive ion etching (DRIE), which selectively remove material from the exposed areas of the substrate, creating the nanofluidic channels.

5. Resist removal: Once the etching process is complete, the remaining resist material is removed, leaving behind the nanofluidic channels etched into the substrate.

6. Sealing the channel: To cover and seal the nanofluidic channels, a bonding process is employed. This typically involves attaching a transparent material, such as glass or a polymer, onto the etched substrate using techniques like anodic bonding, fusion bonding, or adhesive bonding. This step encloses the nanofluidic channels and creates a sealed micro- or nanofluidic device.

The length of the nanofluidic channels can vary depending on the application requirements, ranging from a few µm to several mm or even cm. However, as the channel length increases, the potential for issues like increased fluidic resistance and fabrication complexity may also rise.

E-beam lithography enables the fabrication of nanofluidic channels with widths and heights down to tens of nanometers, offering high resolution and precision for various applications in fields such as molecular biology, nanotechnology, and analytical chemistry.

Compatibility with the actual feedstock and suitability of a penetrable covering film remain to be established.

12.12.7.2 Covering material

The choice of covering material for sealing micro- or nanofluidic channels depends on factors such as transparency, chemical compatibility, biocompatibility, bonding technique compatibility, and mechanical properties. Of the common materials, we require a thin film that can be penetrated by an atomically precise system made by the SRU. This suggests the use of thin films:

- Candidate membrane materials include silicon nitride, silicon dioxide (SiO_2), and various metal thin films
- Thickness range: Tens of nanometers to several µm
- Advantages: Can be deposited directly onto the substrate using techniques like chemical vapor deposition (CVD) or physical vapor deposition (PVD), and can be used as a barrier layer or as part of a multilayered covering

12.12.7.3 Sacrificial layer

When using a thin film to cover and seal nanofluidic channels, a sacrificial material is typically employed to fill the channels temporarily during the deposition process. This prevents the thin film from closing off or clogging the channels during deposition.

The process typically involves the following steps:

1. Filling the channels: After the nanofluidic channels are etched into the substrate, they are filled with a sacrificial material, such as a polymer, photoresist, or even a metal that can be later removed selectively without damaging the substrate or the deposited thin film.
2. Thin film deposition: With the channels filled, the thin film can be deposited on the substrate using techniques like chemical vapor deposition (CVD) or physical vapor deposition (PVD), depending on the chosen material. The thin film covers the entire substrate, including the filled channels.
3. Sacrificial material removal: Once the thin film deposition is complete, the sacrificial material is removed from the channels using a selective removal process, such as wet etching, dry etching, or thermal decomposition. This process uncovers the nanofluidic channels, which are now sealed by the thin film layer.

The use of sacrificial materials is a common practice in micro- and nanofabrication processes where sealing of delicate features is required. It is important to choose a sacrificial material that is compatible with the substrate, thin film material, and the intended application to ensure successful fabrication and functionality of the final device.

12.12.7.4 Filling channels with sacrificial materials

Filling nanofluidic channels with sacrificial material without spilling excess material onto the surface of the die can be challenging. However, there are several methods to achieve this with minimal residue or excess material on the die surface:

1. Controlled deposition: By using techniques such as spin coating, spray coating, or inkjet printing, it is possible to control the amount of sacrificial material deposited on the die surface. By optimizing the deposition parameters, such as spin speed, deposition time, or printing pattern, the material can be selectively applied to fill the channels without excessive spillover.
2. Surface treatment: Applying a surface treatment or modifying the surface properties of the die, such as by using a hydrophobic or hydrophilic coating, can help confine the sacrificial material to the channels, minimizing its spread on the surface.
3. Viscosity control: Adjusting the viscosity of the sacrificial material can help control its flow and spread on the die surface. A higher viscosity material is less likely to spread or spill over onto the surface, making it easier to fill the channels with minimal residue.
4. Removal of excess material: After filling the channels, any excess sacrificial material on the die surface can be removed through a cleaning process, such as a rinse with a solvent, or by using a mechanical removal technique like a doctor blade or a soft squeegee. It is crucial to ensure that the cleaning process does not damage the substrate or dislodge the sacrificial material from the channels.

5. Sequential processing: In some cases, the sacrificial material can be deposited in a two-step process. First, the material is deposited to fill the channels, followed by a selective removal or planarization step to remove excess material from the die surface. This can be done using techniques like chemical-mechanical polishing (CMP) or selective etching.

The choice of method depends on the specific requirements of the fabrication process, the properties of the sacrificial material, and the dimensions of the nanofluidic channels. Careful process optimization and control are critical to successfully filling the channels without leaving excessive residue on the die surface.

12.12.7.5 Open questions

1. How to develop a die fabrication process that creates leakproof nanofluidic channels of the correct dimensions to carry feedstock to all SRUs on a die.
2. How the die fab process simultaneously allows for the creation of appropriate contact pillars on the surface in parallel with the nanofluidic channels.
3. How to flow bulk feedstock into the nanofluidic channels in the die (presumably from the edge of the die).
4. How each SRU can "drill" through to tap into the feedstock in a nanofluidic channel while maintaining a seal during and after the drilling process.
5. How product is removed from the vacuum chamber.

These problems are likely solvable, but it's unclear whether the cost and delays are justified as compared with the alternatives. This awaits further analysis.

12.12.8 Conclusion

A preliminary review of possible strategies for mitigating the risk of leakage during extrusion of the daughter from the parent shows that there are several options available should this become an issue.

13 Opportunities

13.1 Future research in electromechanosynthetic (EMS) reactions

Recall that electromechanosynthesis (EMS) denotes mechanosynthesis involving controlled electron transfer or the use of charged molecular groups in the reaction. EMS reactions are a fertile area for future research. While we have so far considered EMS primarily for generating H_2, several reactions involving AzAdam come to mind:

$$\text{HAbst-H(GeRad)} + \text{AzAdam}^+ \rightarrow \text{HAbst-GeRad} + [\text{AzAdam-H}]^+ \tag{108}$$

This repeats Eq. (90), already mentioned as appropriate for future research. Another might be:

$$\text{GeRad-CH}_2 + \text{AzAdam}^+ \rightarrow \text{GeRad} + \left[\text{AzAdam-CH}_2\right]^+ \tag{109}$$

This could potentially provide a CH_2 deposition tool that would not apply force to the workpiece to release the CH_2. The donation of the CH_2 would proceed when an electron, provided either from a nearby wire with an appropriate voltage or from a nearby positionally controlled electron donor group, was donated to the AzAdam-CH2+.

An alternative reaction that might be energetically more favorable for making AzAdam-CH2+ is:

$$\text{HAbst-CH}_2\text{(GeRad)} + \text{AzAdam}^+ \rightarrow \text{HAbst-GeRad} + \left[\text{AzAdam-CH}_2\right]^+ \tag{110}$$

Similar reactions to make AzAdam-CC+ might be:

$$\text{GeRad-CC} + \text{AzAdam}^+ \rightarrow \text{GeRad} + [\text{AzAdam-CC}]^+ \tag{111}$$

$$\text{HAbst-CC(GeRad)} + \text{AzAdam}^+ \rightarrow \text{HAbst-GeRad} + [\text{AzAdam-CC}]^+ \tag{112}$$

13.1.1 "Donating" and "abstracting" electrons

Further research is needed on the more general process of "donating" and "abstracting" electrons within the framework of positionally controlled molecular tools, much as we might donate or abstract H, to facilitate desired EMS reactions.

Related precedents exist in STM-induced single-molecule synthesis, where tunneling electrons and voltage pulses have been used to break bonds, form bonds, and change redox states after positional manipulation of adsorbates; however, the systematic design of reusable molecular tools whose charge state is deliberately set and then used for positionally controlled electron donation or abstraction appears to remain largely unexplored in the mechanosynthesis literature [FM08a, B+25, B+26, HBMR00, SA01, HR03, dBE16, DK11, WH10, Z+08].

Two specific reactions for generating charged groups which might fruitfully be investigated are the following electron-attachment reaction and the AzAdam ionization recalled from Eq. (91):

$$\text{GeRad-CC} + e^- \rightarrow \text{GeRad-CC}^- \tag{113}$$

$$\text{AzAdam} \rightarrow \text{AzAdam}^{+} + \text{e}^{-} \tag{114}$$

The resulting charged group could then be positioned, much like any other molecular tool, either to participate directly in a reaction or to donate or abstract an electron from a reaction. Groups with high electron affinity would be useful as electron acceptors. Positionally controlled electron carriers require a balance: the electron must be bound strongly enough to remain attached during transport but weakly enough to transfer to the intended acceptor. Electron affinities in the approximate range of 160–480 zJ (1–3 eV) might be useful, depending on the acceptor, temperature, applied electric field, and required retention time.

Additional EMS reactions that come to mind which might fruitfully be investigated are:

$$\text{AdamRad-CC}^{-} + \text{AdamRad} \rightarrow \text{AdamRad-CC-AdamRad} + \text{e}^{-} \tag{115}$$

$$\left[\text{AzAdam-CH}_2\right]^{+} + \left[\text{AzAdam-CH}_2\right]^{+} + 2\,\text{e}^{-} \rightarrow 2\,\text{AzAdam} + \text{C}_2\text{H}_4 \tag{116}$$

$$\left[\text{AzAdam-H}\right]^{+} + \left[\text{AzAdam-CH}_3\right]^{+} + 2\,\text{e}^{-} \rightarrow 2\,\text{AzAdam} + \text{CH}_4 \tag{117}$$

Generation of C_2H_2 (acetylene) might proceed by

$$\left[\text{AzAdam-CH}_2\right]^{+} + \left[\text{AzAdam-CH}_2\right]^{+} \rightarrow \left[\text{AzAdam-C}_2\text{H}_4\text{-AzAdam}\right]^{2+} \tag{118a}$$

$$\left[\text{AzAdam-C}_2\text{H}_4\text{-AzAdam}\right]^{2+} + 2\,\text{HAbst} \rightarrow \left[\text{AzAdam-C}_2\text{H}_2\text{-AzAdam}\right]^{2+} + 2\,\text{HAbst-H} \tag{118b}$$

$$\left[\text{AzAdam-C}_2\text{H}_2\text{-AzAdam}\right]^{2+} + 2\,\text{e}^{-} \rightarrow 2\,\text{AzAdam} + \text{C}_2\text{H}_2. \tag{118c}$$

Brackets enclose each complete charged tool-bound species; the superscript outside the brackets denotes its total charge, without specifying precise charge localization. In the pathway in Eq. (118), $[\text{AzAdam} - \text{C}_2\text{H}_4 - \text{AzAdam}]^{2+}$ is the complex formed by two AzAdam^{+} tools bonded to the two carbons in C_2H_4. Once this complex is formed, it might be possible to abstract two hydrogens from it, leaving the complex of two AzAdam^{+} tools each bonded to one of the carbons in C_2H_2. Further examination of these proposed electromechanosynthetic pathways to generate C_2H_4 and C_2H_2 is called for.

Future research into a much broader range of possibilities should be of interest.

13.2 The positional control of electrons

Mechanosynthetic reactions that use charged reactants require some method of charging those reactants. This might be done by transferring an electron to or from the tool from a charged probe. It might also be done by transferring charge from one tool to another.

The concept of positionally controlling a single electron in the context of electromechanosynthesis has not previously been fully explored. We consider one tool that might both accept and donate an electron: HAbst. We then consider a few of the possibilities.

HAbst has an electron affinity of ∼476 zJ (2.9689 eV) [ZGN07]. The transition from HAbst to HAbst^{-} might be mediated by proximity to a negatively charged probe. Once HAbst^{-} is available,

it might transfer its electron to a reaction, such as $[\text{AzAdam-H}]^+ + \text{H-GeRad} + e^- \rightarrow \text{AzAdam} + H_2 + \text{GeRad}$. This transfer might be spontaneous, or it might be assisted. An assist might be to bond HAbst^- to a second radical, as in

$$\text{HAbst}^- + \text{GeRad} \rightarrow \text{HAbst-GeRad} + e^- \tag{119}$$

On this simplified energy balance, electron detachment from HAbst^- requires approximately 476 zJ (2.97 eV), while formation of the HAbst–GeRad bond releases approximately 769 zJ (4.8 eV) [TAI+10]. The net reaction is therefore exothermic by approximately 293 zJ (1.83 eV), before accounting for local electrostatic potentials and structural relaxation.

Positionally controlled charged tools such as HAbst^- would also enable charge transfer across an electric field for various purposes. The Van de Graaff generator transfers charge from outside a charged conductive (typically metallic) sphere to inside through a hole in the sphere. This transfer is often mediated mechanically by charge on a moving belt going through the hole. We might build a molecular Van de Graaff generator in which charge transfer into the conductive sphere (possibly a spherical fullerene) is mediated by moving HAbst^- tools into the sphere, discharging them by bringing them to within tunneling range ($< \sim$1 nm) of a nanotube protrusion from the inside of the spherical fullerene (but not within bonding range of ~0.5 nm), and moving them back out of the fullerene. This would enable build-up of high voltages (>10 V) at the molecular scale by moving charged tools into and discharged tools out of (or vice versa) an appropriately designed molecular-scale conductive sphere. The energy required to add an electron to the negatively charged sphere is supplied as mechanical work done against electrostatic repulsion while moving the HAbst^- anion through the hole from just outside to just inside the sphere. This converts mechanical energy into electrical energy. Conversion of electrical to mechanical energy occurs when charged particles are moved mechanically through electric fields (Figure 193).

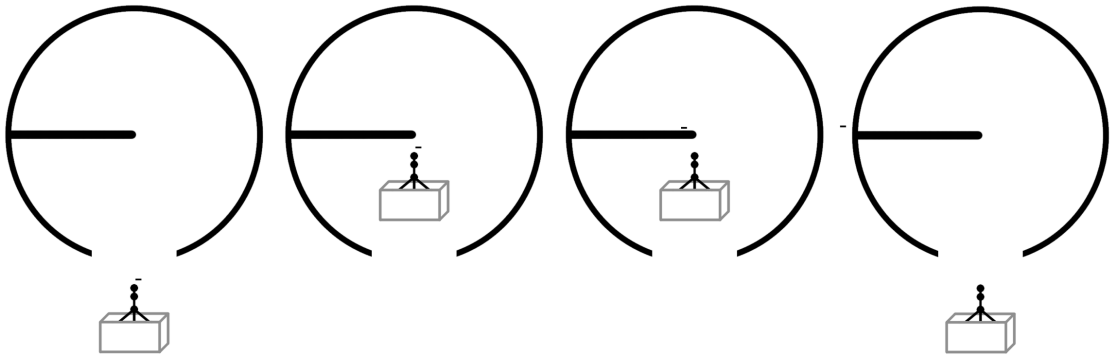


Figure 193: Conversion between mechanical and electrical energy using a charged tool and a conducting sphere. Left: a negatively charged HAbst^- outside a conductive fullerene sphere. Center left: the negatively charged HAbst^- has been moved by mechanical means (not shown) into the center of the conductive fullerene sphere. Center right: the electron on the HAbst^- tunnels to the nanotube extending from the left side of the conductive fullerene sphere. HAbst^- must be within tunneling range (~1 nm) but outside of bonding range (~0.5 nm) of the SWNT. Right: the electron has moved from the center of the sphere to its surface, and the HAbst has moved back outside the sphere.

The electron affinity of HAbst is ~476 zJ (2.9689 eV) [ZGN07], while the work function of a carbon SWNT is ~801 zJ (5 eV) [SA01]. Repeated transfers of an electron to the nanotube will increase

the voltage on the conductive fullerene sphere. In the ideal closed-conductor approximation, charge accumulated on the outer surface produces no electric field in the cavity. This shielding could allow transfer from an internal $HAbst^-$ even when accumulated charge inhibits transfer from outside. The charged tool itself produces a field and induces charge on the inner surface; the aperture and molecular scale of the proposed shell also limit this approximation. Mechanical work would be required to move the charged tool into the charged sphere, potentially converting mechanical energy into electrical energy.

Summary: energy can be converted between electrical and mechanical forms by using molecular scale apparatus that uses charged tools to move charge with or against local electric fields. High voltages can be generated by the mechanical motion of ionized molecular tools. These high voltages can be used to ionize molecular tools. Alternatively, such voltages can be used to apply force to charged molecular tools, or to any charged molecular structure.

13.3 Additional concepts for building 2 µm products

13.3.1 Abstract

There are many concepts for building 2 µm product objects, given that we start with SRUs on a surface. Two of these concepts were included in the baseline and described in Section 9.1. The remaining concepts are described here as either risk mitigations or possible opportunities. Some of these concepts were developed from conversations of the form "Why couldn't we do something like <fill in the blank>?" Sometimes the development process took unexpected turns when specific implementation issues were considered in greater detail.

13.3.2 Introduction

Once enough SRUs have been built on a 1 cm^2 die, it will be necessary to manufacture products of some size. We here assume that we want to build products up to 2 µm in linear size, a canonical product being a cube with dimensions 2 µm × 2 µm × 2 µm, although any product that fits within such a cube is assumed to be manufacturable if the cube is manufacturable.

The manufacturing rate of the following concepts is not tracked, meaning some concepts will manufacture product more slowly and others will manufacture more rapidly if the die surface area devoted to manufacturing is held constant. Alternatively, if we wish to manufacture products at the same rate, the die's surface area devoted to manufacturing will have to be increased or decreased appropriately.

We assume small products (smaller than the size of an SRU) are manufactured inside SRUs. They are exported into the surrounding liquid feedstock by means similar to SRU export and assembled there. The assembly of these small products primarily uses such non-bonding methods as dovetail joints, screw joints, bolts, pegs in holes, lap joints, van der Waals forces, and the like. Some bonded interactions (dipolar bonds, some limited forms of bonded interactions) might also be used. Products of an irregular shape could, if necessary, be exported in a disposable shell that conformed to the cross-sectionally extruded dimensions of an SRU.

In non-vacuum (or fluid) assembly, if vacuum regions within the final product are necessary, then the vacuum will have to be created by some post-manufacture method, such as pumping down the regions using vacuum pumps, also manufactured for the purpose. Elevating the temperature while pumping down the interior region to be evacuated ("baking out") is a proven method for improving

the quality of the vacuum, as more species become gases and therefore easier for the vacuum pumps to remove. If this process is used, any temperature labile fragments would need to be stabilized or disposed of until temperatures had been reduced.

We also assume that finished products can be released into the surrounding fluid. Fluid flow carries them out of the product manufacturing area and to some suitable product collection and processing area. It would also be possible to take advantage of the propensity of product objects to sink or float: products that sink could be collected from the bottom of the tank, while products that float could be skimmed from the top.

If assembly in vacuum is needed, see Section 13.3.3.

In some concepts it might be useful to bolt objects to the exterior of the shell (Figure 194).

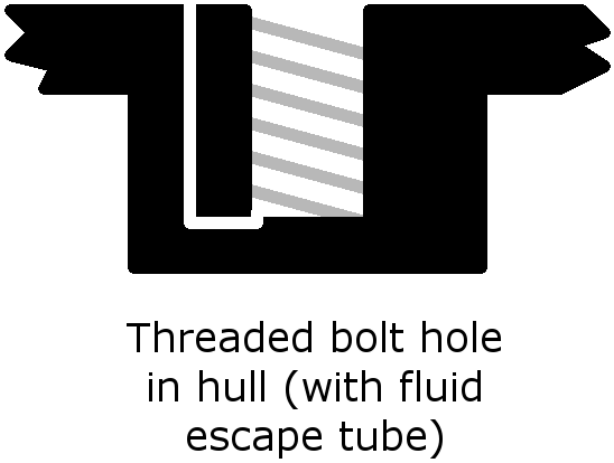


Figure 194: Attachment of an object to the exterior of an SRU shell.

A potential problem with atomically precise bolts in atomically precise bolt holes is that, when sufficient upward force is applied to the bolt, it might rotate and tend to unscrew, eventually falling out. This might be prevented by using a locking pin that is inserted into the head of the bolt and slides into a slot in the surface and prevents rotation of the bolt after it has been screwed in place. If designed appropriately, the locking pin would be held in place by van der Waals and other intermolecular forces, much like a molecule held in a binding site.

A locking pin may not be needed. VDW (van der Waals) will make the bolt want to screw in to maximize surface contact area. It will also want to pull the flange of the bolt head down against the surface. Van der Waals forces should tend to keep the bolt tightened, although a sufficiently large applied force could loosen it. If we can use VDW to facilitate self-assembly and adhesion, it could simplify the design.

Another possibility would be an atomically precise self-locking bolt with two facing surfaces that allow rotation of the bolt in one direction with low torque as it is screwed in place, but which prevent it from unscrewing due to the ratchet effect. Each time the bolt is turned past one of the ratchet surfaces, there would be a dissipative "snap" that might dissipate as much as a few hundred k_BT, effectively preventing back rotation of the bolt from thermal noise. Unscrewing the bolt would require the application of a large torque.

Self-locking bolts have the disadvantage that they are fundamentally dissipative. The use of a

locking pin, illustrated in Figure 195, is one alternative which could avoid the dissipative "snap" as a bolt is screwed in and locked in place. To unscrew the bolt, remove the locking pin and apply a small torque to unscrew it. The locking pin would be held in place by van der Waals forces, and if the energy required to remove it is tens to a few hundred k_BT, it will be stable in the face of thermal noise.

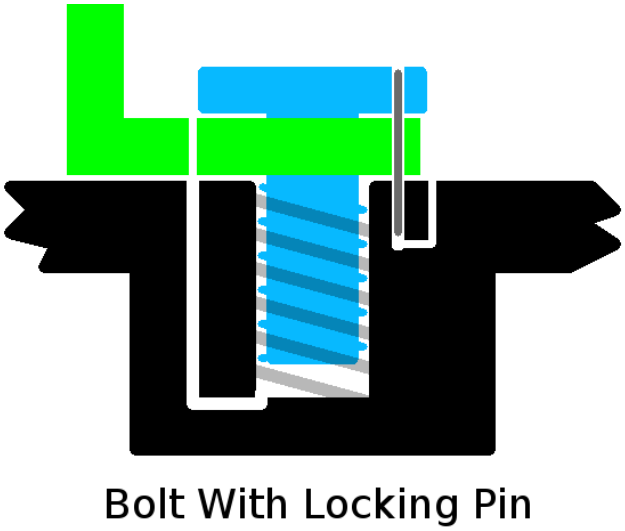


Figure 195: Bolt secured with a removable locking pin.

In all concepts for building 2 µm products, we assume individual SRUs are addressable, so that each SRU can be instructed to build specific SRU-sized components and can also be instructed to coordinate assembly of SRU-sized components with other SRUs to build larger products. SRUs within a 2.4 µm × 2.0 µm surface region can coordinate their activities to manufacture a product.

Recall that all active SRUs must execute exactly the same instructions. We might think about two common modes of operation. In the first mode, all SRUs are manufacturing an AP Block that will be used many times in the manufacture of the product. This mode of operation efficiently utilizes the productive capacity of the SRUs. In the second mode, a single SRU in a 2.4 µm × 2.0 µm region is carrying out a unique operation by executing the broadcast instructions while the other SRUs in that region are OFF. This mode of operation should be limited in the frequency with which it happens, particularly as the number of SRUs in a region increases, as it does not efficiently utilize the capacity of the SRUs. Unique assembly steps used in the manufacture of a product should be limited in number.

The good news is that the manufacture of repeatedly used blocks can fully utilize the productive capacity of the SRUs, while unique assembly operations used to assemble SRU-sized components will necessarily be few relative to mechanosynthetic operations required to fabricate blocks because there are many more total mechanosynthetic operations in fabrication of blocks than there are assembly operations in the assembly of SRU-sized components. Therefore, it should be feasible to manufacture products in a way that keeps most of the SRUs doing something useful most of the time (that is, that limits the average idle time across all SRUs to an acceptable percentage). This assumes that most blocks are repeated many times.

In the following we identify for each concept whether we anticipate assembly of SRU-sized components in vacuum or liquid, and the initial density of SRUs on the 1 cm^2 die before we begin product manufacturing operations, typically specified as the number of SRUs per 2.4 µm × 2.0 µm

region on the 1 cm^2 die. The replicative process whereby we achieved the necessary initial density of SRUs will be left unspecified here.

Table 20 summarizes a comparison of product assembly Concepts 2 and 6.

Table 20: Comparison of product assembly Concepts 2 and 6.

	Concept 2	**Concept 6**
Number of distinct SRU designs	1	3
SRU-sized component assembly environment	Liquid or Vacuum	Vacuum
Internal assembly from multiple fabricators	No	Yes
Required seed size	SRU	SRU

13.3.3 Vacuum Assembly in a Product Assembly Chamber (PAC)

If assembly in vacuum is needed, and if we wish to bond SRU-sized components together in vacuum, and if the assembly method is initially in liquid, then we first create a vacuum environment in which to make the final product. That is, we first create a non-vacuum (liquid) assembly environment which enables the assembly of SRU-sized components in liquid and then build a Product Assembly Chamber (PAC) out of SRU-sized components in liquid. We then pump down this PAC, creating an internal vacuum in the PAC, thus enabling further assembly of SRU-sized components in vacuum. Bakeout of the PAC prior to pump down might be used, requiring cycling of the die chamber between liquid and vacuum, as well as temperature control and the ability to "safe" the SRU against elevated temperatures. The SRUs involved are directly connected to the PAC, which allows said SRUs to export SRU-sized components directly from the internal vacuum inside each SRU into the vacuum of the PAC, thus eliminating the need for an airlock or extrusion sleeve between each SRU and the PAC.

There are several design choices for the PAC. First, the PAC can either (a) expand in size as the product gets larger or (b) be of fixed size, and the product must fit inside the fixed dimensions. Second, to get the product out of the PAC we can i) break the seal on the PAC and flood it, after which we can remove the product but must now somehow create a new vacuum for the next product, or ii) use an airlock, while maintaining the internal vacuum within the PAC at all times.

We illustrate the general concept by a series of cartoons illustrating one method of creating a vacuum assembly environment and exporting products. We start by manufacturing our PAC in the liquid environment and then pumping down a vacuum. A cartoon of a possible vacuum assembly environment, and the process of manufacturing and exporting a product, might look like the sequence in Figure 196. The vacuum pump is not depicted in these cartoons. The gray cross is the product to be exported.

In this particular approach, the sidewalls are built once, then a new "floor" is built for each product. The "floor" slides up the sidewalls using a sliding seal to maintain the vacuum, and the "floor" becomes the "ceiling" providing a build volume for the new product. Once the new product has been manufactured, a new "floor" is manufactured, the "ceiling" is jettisoned, and the product is

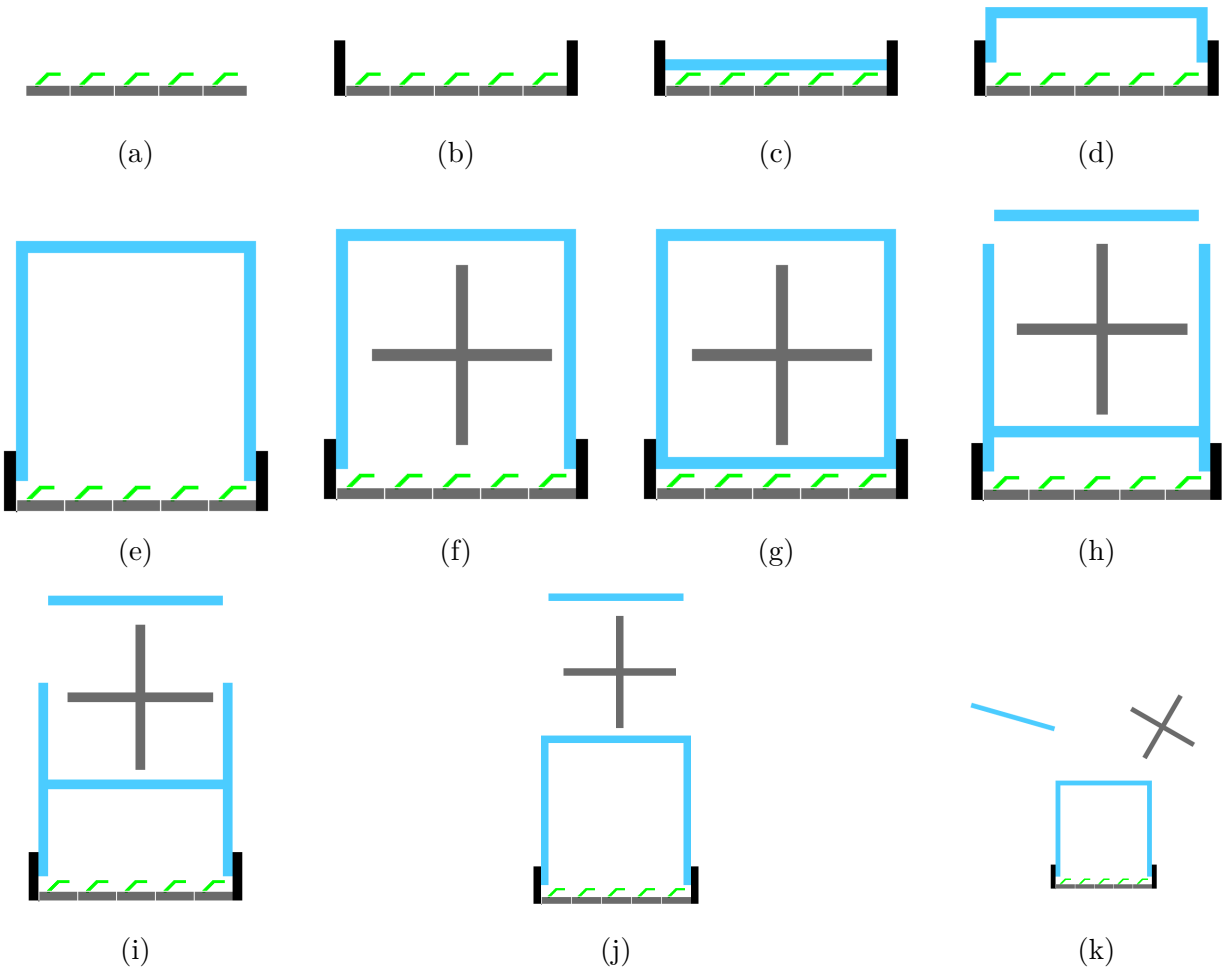


Figure 196: Vacuum assembly and product export in a Product Assembly Chamber. Panels (a)–(k), in reading order, show construction of the chamber, product assembly, opening, and export. The gray cross represents the product; the vacuum pump is not depicted.

expelled into the surrounding liquid. The process is then repeated as the new "floor" slides up the sidewalls and takes its place as the new "ceiling".

This concept discards the "ceiling" after each cycle. It would also be possible to integrate the "floor" into the product, reducing waste. Concepts that reuse the "ceiling" and turn it, once again, into a "floor" are also possible.

We assume the original SRUs are arrayed in a liquid environment. In Figure 196(a), their external assembly arms are shown in green. These arms use exported SRU-sized components to assemble the black external seal around the SRU bases in panel (b), then the blue vacuum-tight enclosure in panel (c). The enclosed volume must be pumped down to a vacuum suitable for assembly; the pump is not depicted. Panels (d) and (e) show expansion by raising the roof and extending the sides.

Panel (f) shows a product assembled in vacuum inside the PAC. A new floor seals the chamber below the product in panel (g), and the ceiling is opened in panel (h). Raising the floor exports the product while maintaining vacuum beneath it, as shown in panels (i) and (j). Panel (k) shows the product and discarded ceiling clear of the chamber, with the former floor now serving as the new ceiling. If the product can itself maintain a vacuum-tight sliding seal during extrusion, the ceiling need not be discarded. Alternatively, the ceiling could be reclaimed and recycled.

Other design concepts for a PAC are possible.

13.3.4 Concept 2

This concept has already been briefly described as part of the baseline. In this concept, there are 40 SRUs per 2.4 µm × 2.0 µm region on the 1 cm^2 die, and assembly of SRU-sized product objects into the final product takes place in the liquid feedstock.

We assume the 40 SRUs form a 4 × 10 array, more or less evenly laid out in the 2.4 µm × 2.0 µm region. Each SRU occupies a site that is 200 nm × 600 nm (Figure 150).

Assigning each SRU a pitch of 600 nm along its 500 nm dimension and 200 nm along its 100 nm dimension gives a 4 × 10 array occupying a periodic region 2.4 µm × 2.0 µm in size.

Each SRU exports SRU-sized components which are then assembled, in the liquid environment, by the external assembly arms into products.

To facilitate cooperation among the external assembly arms, the SRUs to which they are attached could first be aligned by using alignment struts. The alignment struts can be adjusted in length (see the discussion of turnbuckles and Figure 198). By using six alignment struts between adjacent pairs of SRUs, the adjacent pairs can be brought into alignment much as the platform of a Stewart platform can be brought into alignment with its base. First, the six struts are adjusted in length so that they match the existing lengths between their attachment points. This allows the struts to be attached without the need to move the SRUs to which they are attached. Finally, the six struts are adjusted in length until the adjacent SRUs are brought into their proper (aligned) positions. This sequence should be able to align the SRUs to within a few tens of picometers, more than sufficient so that the external assembly arms can assemble SRU-sized components produced by the SRUs into products in fluid using blind "pick and place" assembly operations.

Figure 197 shows a two-dimensional analog in which adjacent SRUs are connected by three alignment struts, rather than the six that would actually be used in three dimensions.

It's also worth noting that in this arrangement, the alignment struts overdetermine the position of the four SRUs in the two- dimensional plane, so we should either adjust the struts in small



Figure 197: Two-dimensional analog of SRUs connected by alignment struts.

increments to minimize excess tension on the struts during alignment or eliminate struts that result in an overdetermined system.

One possible implementation strategy for alignment struts would be to use turnbuckles (with thanks to David Lister, see https://en.wikipedia.org/wiki/Turnbuckle) (Figure 198):

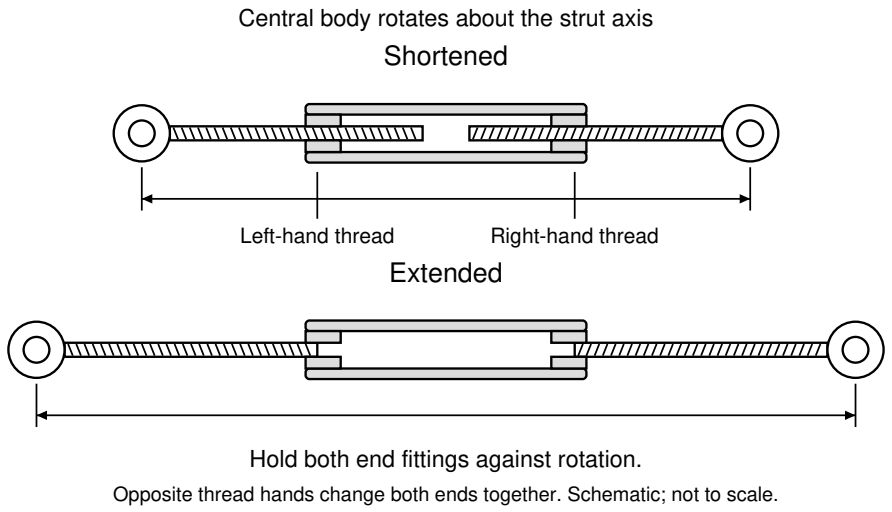


Figure 198: Turnbuckle principle for adjusting alignment struts: shortened (top) and extended (bottom). With the end fittings held against rotation, rotating the central body draws the oppositely threaded ends together or apart. Schematic, not to scale.

The alignment process could be facilitated if SRUs were attached to the surface through some sort of flexible mounting system that allowed the SRUs to be easily moved some distance with respect to the surface. As long as all 40 SRUs were mounted within some allowable distance from their correct position, aligning them should be feasible.

An atomically precise turnbuckle might spontaneously lengthen or shorten unless designed so that the rotating element would either "snap" to a series of stable positions or have a locking mechanism that prevented it from rotating once the turnbuckle had been set to the correct length. We assume one of these approaches is incorporated into the turnbuckle design.

The baseline nanofactory should produce $\sim 1.5 \times 10^{12}$ (1.5 trillion) medical nanorobots every 80 days, each medical nanorobot being manufactured during the full 80 day period on its own 2.4 µm × 2.0 µm manufacturing site by the coordinated efforts of 40 SRUs. See How to build 2 µm

products (Section 9.1) for further discussion.

If assembly in vacuum is required, a PAC would first be manufactured and pumped down in the liquid feedstock environment.

13.3.5 Concept 6: Unitary product fabrication with kiloSRUs

Proposal by Robert Freitas, Feb 2021.

13.3.5.1 Highlights

1. Builds products entirely in vacuum.
2. All products are unitary builds with no assembly from smaller parts required, up to 2 μm product sizes.
3. Starts from a seed that is similar in size to the baseline SRU.

13.3.5.2 Operating Summary

Define "SRU1" as a 500 nm × 200 nm × 100 nm SRU.

Start with a populated die of SRU1s.

1. Select those SRU1s on a 2D sparse grid spaced 4,400 nm × 8,800 nm apart. Ignore the other SRU1s on the silicon die. Have the selected SRU1s extrude an 8000 nm × 200 nm × 100 nm "SRU2" onto an opposing die in ∼20 days, filling ∼4% of the 1 cm^2 die surface. Conceptually, each SRU2 is made from a linear sequence of SRU1s, with a linear sequence of fabricators internally, and with a single 4,000 nm × 100 nm extrusion port.
2. Each SRU2 extrudes a 4000 nm × 4000 nm × 100 nm "SRU3" (a.k.a. a "kiloSRU") onto an opposing die in ∼10 days, filling 50% of the sites on the die surface. This procedure is repeated during the next 10 days, completely filling 100% of the sites on the 1 cm^2 die surface in a total of 20 days. Conceptually, each SRU3 has a rectangular grid of ∼320 densely packed fabricators mounted internally on its "floor". The densely packed fabricators share an internal assembly arm. As each SRU3 occupies a region that is 4.4 µm × 4.4 µm and also has ∼320 fabricators internally, this produces ∼twice the density of 40 fabricators per 2.4 µm × 2.0 µm region, which gives this proposal a raw manufacturing rate ∼twice that of the baseline per unit area.
3. Each SRU3 can extrude one 0.2-trillion-atom, 2 μm × 2 μm × 2 μm medical nanorobot particle along with its shipping container every ∼10 days, a total-product manufacturing rate of ∼twice that of the baseline (assuming the same number of boards).

13.4 Coordinates for each daughter

Another class of algorithms for populating a die (not included in the baseline but left as an opportunity) starts by giving each daughter SRU the target coordinates of its assigned destination. At each time step, the parent SRU compares its own coordinates with the target coordinates of the daughter SRU that it is holding in that time step. It then moves the daughter SRU in whatever

direction gets it closer to its target. One possible implementation might be to encode the coordinates of the target, in binary, on the shell of the daughter. The parent holding the daughter in any given timestep would then read the coordinates using a Test Unit on the end of the EAA.

Permanently blocking progress only occurs when all possible movements along a path to the target are blocked by fail-stopped SRUs, which is unlikely for "reasonable" failure rates if the EAA can reach even a relatively modest number of sites.

If a daughter SRU has eight candidate direct next steps, and each is independently occupied by a fail-stopped SRU with probability 0.1, the probability that all eight candidate next steps are blocked is $0.1^8 = 10^{-8}$, or one in 100,000,000.

13.5 Faster methods of populating dies

Our baseline method for populating a die is to (1) attach a single SRU to a single site on the die, and then (2) allow that single SRU to replicate, covering the die (see Daughter placement (Section 4.8.4) and A Local Replication and Routing Algorithm (Section 6.4)). If there are 2^N sites on the die, it would take N replication cycles to populate the die, along with whatever additional time was required to pass daughter SRUs, hand-over-hand, from parent to parent before they could be placed on an open site on the die.

A somewhat faster method would be to place two SRUs on two sites on the die. Instead of taking N replication cycles, this would take $N-1$ replication cycles to generate enough SRUs to populate all the sites on the die. If we placed the two SRUs side-by-side on the die surface, delays caused by hand-over-hand passage of daughter SRUs would be unchanged. If the two SRUs were separated from each other on the die by an appropriate distance, delays caused by hand-over-hand passage of daughter SRUs could also be reduced.

If both SRUs could be placed on the die at the same time, in parallel, there would be no loss of time from placing two SRUs on the surface instead of one.

If two is good, would four be better? That is, if we could place four SRUs on the surface in one parallel operation, we could generate enough SRUs to populate the 2^N sites on the die in $N-2$ replication cycles, saving us two replication cycles instead of just one.

This suggests three strategies:

> **Strategy 1:** 2^M SRUs (with $1 < M \leq N$) are placed on the destination die in a single densely packed block. The die is then fully populated by the replication of the newly placed SRUs and their descendants on the surface.
>
> **Strategy 2:** Batches of 2^M SRUs (with $1 < M \leq N$) are repeatedly transferred from a fully populated source die to a destination die until the source die is depleted of daughter SRUs and all sites on the destination die are occupied.
>
> **Strategy 3:** A source die with 2^N SRUs on it builds a flat sheet with appropriate wiring to support 2^N SRUs. The source die then transfers 2^N SRUs to the newly constructed flat sheet, which is then detached. The flat sheet then replaces a silicon die in future operations.

13.5.1 Strategy 1

Strategy 1 might transfer a square array of SRUs. We adopt the method of allocating SRUs to squares discussed elsewhere in this paper (see RC delays in an H tree (Section 7.16.4)). The

parameter k takes on values between 0 and 14, inclusive. A square at level k has size $2^{14-k} \times 0.6$ µm. The square at Level 0 has sides that are $2^{14-0} \times 0.6\,\mu m = 9{,}830.4\,\mu m$ long. The square at Level 14 has sides that are $2^{14-14} \times 0.6\,\mu m = 0.6\,\mu m$ long. For the sake of a concrete illustration, assume we have a Level 0 source die that is 9.8304 mm × 9.8304 mm and a Level 0 destination die that is also 9.8304 mm × 9.8304 mm. Assume we have a Level 3 "transfer die" that is 1.2288 mm × 1.2288 mm in size. That is, there are three dies: a source die and a destination die that are both at Level 0, and a transfer die that is at Level 3. The transfer die holds $2^{14-3} \times 2^{14-3} \times 3 = 12{,}582{,}912$ SRUs. (See RC delays in an H tree (Section 7.16.4)). The source die and the destination die each hold an 8 × 8 grid of transferable squares, where each cell in the grid is a Level 3 square that is 1.2288 mm × 1.2288 mm in size.

At the start of the process

(1) the source die is fully populated,

(2) the transfer die is fully populated and

(3) the destination die is unpopulated.

The reason the transfer die is fully populated is to facilitate the transfers to and from the transfer die. These transfers involve the active participation of the parent SRUs permanently mounted to the transfer die. The parent SRUs on the transfer die receive active control signals during the transfer process to attach to and release from the daughter SRUs being transferred.

The source die carries out one replication cycle, so that each parent SRU on the source die has manufactured one daughter SRU.

All sites on the transfer die are permanently occupied by one parent SRU. This "transfer SRU" starts with no daughter SRUs.

The transfer die is positioned over a Level 3 square on the source die. All daughter SRUs on the source die that are positioned under an SRU on the transfer die are transferred to the transfer die. This process involves the active participation of the transfer SRUs. In this step, a Level 3, 1.2288 mm × 1.2288 mm patch of daughter SRUs on the source die is transferred to the transfer die. As the source die is a Level 0 square which is 9.8304 mm × 9.8304 mm in size, most of the daughter SRUs on the source die remain on the source die.

Then the transfer die is positioned over the (unpopulated) destination die. A 1.2288 mm × 1.2288 mm Level 3 patch of daughter SRUs on the transfer die is then transferred to the destination die by the transfer SRUs on the transfer die. As the destination die is a Level 0 square measuring 9.8304 mm × 9.8304 mm, this means most of the destination die remains unpopulated. Only a small Level 3 patch that is 1.2288 mm × 1.2288 mm in size on the destination die becomes populated.

At the end of this process, 12,582,912 SRUs, the number of SRUs needed to saturate the Level 3 array of sites on the transfer die, have been moved from the initially fully populated source die to the initially completely unpopulated destination die.

After completing this first transfer, the transfer die now has no daughter SRUs and is free to position itself over a fully populated Level 3 patch on the source die and repeat the entire process with a new Level 3 patch transferred to a new destination die.

At the same time that the new transfer to a new die from the original die is taking place, the original destination die is allowed to manufacture further generations of SRUs until it is fully populated.

This is strategy 1.

13.5.2 Strategy 2

Strategy 2 is the same as strategy 1, except that, instead of replacing the destination die after completing one transfer, we keep transferring Level 3 batches of daughter SRUs from the source die to the destination die until the source die is depleted of daughter SRUs and all sites on the destination die have been occupied. In our example, each transfer handles a Level 3 area. As there are 64 Level 3 squares on the source and destination dies, it will take 64 transfers to fully deplete the population of daughter SRUs on the source die and fully populate all sites on the destination die.

We note that alignment becomes less difficult the larger the value of k, and more difficult the smaller the value of k. In the extreme, $k = 0$, we would be pressing a populated 9.8304 mm × 9.8304 mm source die against a populated transfer die whose SRUs hold no daughters, and a fully populated transfer die against an unpopulated destination die. This would pose significant alignment issues. Given the remarkable alignment achieved for EUV lithography systems, it is difficult to conclude that such a system would prove to be either impossible or uneconomic. However, as k becomes larger, the difficulties are reduced. In our baseline, only a single SRU is transferred, the least difficult option. Larger transfers could be explored using decreasing values of k, starting from $k = 14$ (three SRUs on a 0.6 × 0.6 µm patch) and working our way up to $k = 0$ (transferring the entire nominally 1 cm^2 die).

Transferring a Level 3 patch of SRUs 64 times from the source die to the destination die would allow us to mitigate the risk that replication times might be significantly longer than 10^5 seconds. If the replication time was one week, we could adopt this strategy. As long as the time required to transfer a Level 3 patch was reasonable, the total time to populate a single die would remain largely unchanged. If we assume the transfer of one patch takes ten minutes, then the transfer of 64 patches would take only 640 minutes, or less than half a day. This could be further reduced by transferring multiple patches in parallel.

If our replication time were to slow down to one generation per week, fully populating the first die would take 30 weeks, ∼seven months. Once we have one die, and assuming we can populate a new die using an existing populated die in less than one day, enough SRUs to populate the full 5.8×10^{13} SRUs in a nanofactory would take only an additional ∼17 weeks. This would give a total population time of approximately 47 weeks, less than a year.

Thus, even with one-week replication times, the stated population time would remain less than a year, provided we developed the ability to transfer SRUs from a populated die to an unpopulated die in Level 3 or Level 4 batches.

13.5.3 Strategy 3

In strategy 3 we replace the functionality of the silicon die with an atomically precise alternative. The silicon die is a flat sheet that provides electrical connections and physical support. A fully populated die could manufacture an APM sheet that provided both the electrical connections and the physical support. As the just-manufactured APM sheet would be perfectly aligned with the parent SRU sites on the source die, the source die would then manufacture a new generation of daughter SRUs and, after a single SRU replication time delay, attach all of them to the APM sheet. The APM sheet would then be detached as a unit. It would then be attached to a suitable support system, taking the place of the silicon dies in future manufacturing operations.

13.5.4 Summary

In summary, there are opportunities for significantly shortening the time required to populate a full 1 cm^2 die compared with the baseline strategy of placing a single SRU on the die and allowing it to replicate. Strategies 2 and 3 try to take advantage of the fact that a single 1 cm^2 die can produce enough SRUs to populate a second 1 cm^2 die in one replication cycle.

13.6 Nested nanotubes

Another approach to an atomically precise electrostatic actuator would use nested nanotubes. A cross-section through one possible design for a nested nanotube electrostatic actuator is illustrated as follows (Figure 199):



(a) (b)

Figure 199: Nested-nanotube electrostatic actuator: (a) contracted; (b) extended. Black denotes conducting material and blue denotes insulation.

In this illustration, black is conductive and blue is insulating (Figure 199). A possible implementation would be to use carbon SWNTs for the conductive black nanotubes, hexagonal boron nitride (h-BN) SWNTs for the blue insulating nanotubes, and boron-doped hexagonal diamond for the conductive end-caps. A stack of five layers of h-BN is ∼1.65 nm thick and should reduce leakage current to about 1 nA/µm^2 or 1 fA/nm^2 for an applied voltage of 0.19 V, as extrapolated from the measurements in Figure 132 [K+21, B+12].

Many such actuators could be connected in parallel. If all the inner SWNTs are connected at one end to a conducting plate, and all of the outer SWNTs are connected at the other end to a second conducting plate, the result would be a voltage activated linear actuator that could be made to have as large a force as might be required by adding more actuators in parallel between the two conducting plates.

Figure 200 shows how many such SWNT actuators might look in a schematic longitudinal cutaway, both contracted (a) and extended (b).

Having once connected enough individual actuators in parallel to achieve the required force, we can then repeat this compound structure in series to give as long an actuation length as might be desired. The contracted and extended compound actuators are shown in Figure 201(a) and (b), respectively:

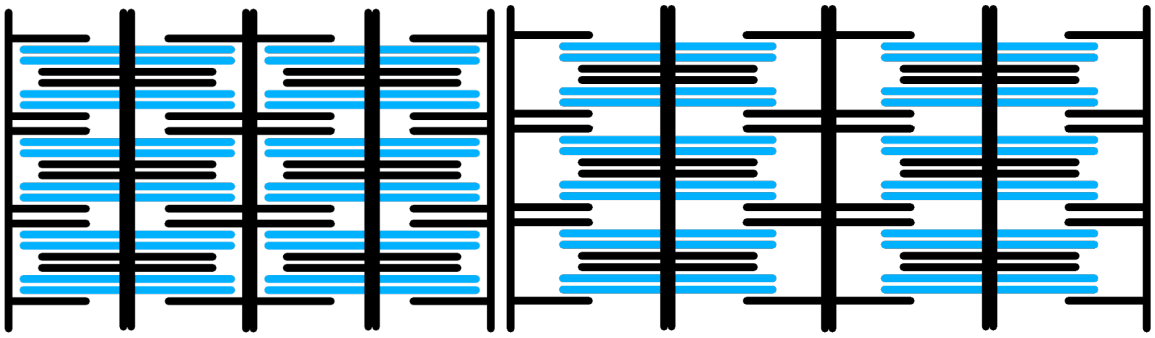
This structure suggests a method of creating a linear actuator which can be built from relatively small fixed-size modular units, perhaps on the order of five nm in length. This fits well with the concept of using a fabricator with a small build volume.

In the design shown, a possible failure mode would be for some of the inner or insulating nanotubes to fall out of the larger nanotube. This can be mitigated by appropriate limitations on the range of motion of either the actuator itself or the actuated device.



(a) (b)

Figure 200: Parallel nanotube actuator array: (a) contracted; (b) extended.



(a) (b)

Figure 201: Compound nanotube actuator: (a) contracted; (b) extended.

13.7 Use of diamond as a dielectric in the comb drive

If we assume diamond between the tines instead of h-BN, we cannot exceed the breakdown field of diamond. The calculated intrinsic breakdown field of bulk diamond is approximately 2.4 V/nm [SBR12]. For this preliminary analysis, we assume an operating field of 1 V/nm, less than half the calculated intrinsic breakdown field. A design-specific analysis would still be required to determine leakage from direct tunneling and electrode injection at the selected diamond thickness.

For atomically precise, defect-free diamond, leakage paths associated with traps, grain boundaries, and extended defects are absent. The relevant residual conduction mechanisms are intrinsic field-induced processes: Zener/Keldysh band-to-band tunneling, Fowler–Nordheim injection at interfaces, and direct tunneling for ultrathin barriers.

A graphene/diamond design would have graphene sliding over diamond. For the same output force and operating voltage, the assumed dielectric constant of diamond (5.7) could allow a smaller comb-drive actuator than the h-BN design, subject to the required dielectric thickness and other design constraints. Properly designed atomically precise surfaces can slide over each other with remarkably little friction (see Friction and wear (Section 7.15.5.14)) [HMA17, Dre92, Mer93a], allowing this design to meet the friction requirements.

This option remains an attractive opportunity, requiring some additional analysis to properly derisk.

14 Task list for SRU development

If we are to build the baseline nanofactory, we must break down that process into a series of identifiable tasks and sub-tasks that, when completed, will lead to the desired result. What follows is a preliminary attempt to identify the major tasks that must be completed. This list is more heavily focused on design tasks than experimental tasks, as experimental tasks depend heavily on the availability and capabilities of the workstation. In this paper, we simply assume that workstation capabilities are available and that we need to proceed to the next step in the process: the design, construction and testing of molecular machines using the capabilities of the workstation.

Tasks marked "**" have highest priority; tasks marked "*" have high priority. Unmarked tasks have no assigned priority. Tasks are not listed in priority order.

14.1 Comb drives

1. ** Design one or more differential comb drives coupled, directly or through a transmission, to a binary Bowden cable or similar output system. The output must provide a 2 nm throw and 1 nN of available force. The design should seek to minimize the build volume required. The actuator design should be decomposed into AP Blocks that the SRU fabricator can manufacture.

2. * Analyze the performance of the electrostatic actuators in terms of maximum force, maximum throw, leakage current/voltage, breakdown voltage, simplicity of use, maximum force/volume, force/voltage, and any other relevant parameters.

3. * Evaluate the ability of the modules to produce output directly into binary Bowden cables and the ability to combine the output of binary Bowden cables in series and parallel to achieve the desired performance.

14.2 Wires to drive comb drives

1. ** Carry out a high level theoretical study of the stability of short lengths of (4,4) nanotubes.

2. ** Carry out a high level theoretical study of the ability to fuse those short lengths into longer lengths of (4,4) nanotubes.

3. * Design an atomically precise gripper for a (4,4) nanotube (possibly using GeRads).

4. * Model the ability of the atomically precise gripper to bind to a (4,4) nanotube.

5. * Design electrically conductive sharp angle turns (e.g., 90°) for (4,4) nanotubes. Propose at least one purely graphene design. (See Molecular Tinkertoys (Section 7.10)).

6. Propose other candidate molecular wires as seems appropriate.

7. * Analyze the candidate wires proposed above. Analysis should include electrical and mechanical properties. Radius of curvature is of particular interest, both in terms of mechanical and electrical properties. Evaluate contact resistance, maximum current carrying capacity, and requirements for either standoff insulators that hold the wires or external insulating jackets (possibly multiwall h-BN nanotubes as external insulating jackets).

14.3 Build volume of SRU fabricator

Carry out a trade-off study of different fixed build volumes for the SRU fabricator, ranging in size from 4 nm to 20 nm (and possibly smaller or larger, as might seem appropriate). Estimate the size of the AP Blocks required to build various SRU subsystems to provide guidance for the trade-off study.

1. Assess the block size for differential comb drives.
2. Assess the block size for threaded rods/sleeves needed for Stewart platforms.
3. Assess the block size for threaded rods/sleeves needed for positional accuracy.
4. Assess the block size for the shortest segment of linear rod, rotating shaft, or hydraulic tube needed to carry power or signal.
5. Assess the block sizes for the components of a binary Bowden cable.
6. Assess the block size for the minimal segment of electrical wire.
7. Assess the block size for components of a demux.
8. Assess the block size for components for rotary motion.
9. Assess the block size for a minimal component of a load-bearing element.
10. Assess the block size for a shell block.
11. Assess the block size of grippers for all the above.

14.4 A demux driven by binary Bowden cables (or equivalent)

1. ** Design a demux which has one pair of binary Bowden cables (S1 and S2) as inputs to **S**elect one pair of binary Bowden cables as outputs, and another pair of binary Bowden cables (D1 and D2) as inputs to **D**rive the selected pair of output binary Bowden cables. This makes a total of four binary Bowden cables as inputs to the demux (S1, S2, D1 and D2). The design should have at least eight pairs of binary Bowden cables as outputs. The design will eventually have to scale to ∼60 or more pairs of binary Bowden cables as outputs.
2. Consider the "control chain" concept as a possible demux for use in the SRU. What are the speed and energy disadvantages of this approach? Are these disadvantages significant in the context of the SRU? From [FM08b, p. 40]:

   > "Control Chain. The control chain consists of a single flexible molecular ribbon with a number of graphitic knobs integrally affixed to known positions along its length. A latchlike mechanism at the end of the piston rod engages the chain, converting piston-driven axial rod motion to transverse chain motion and forcing the chain to move. All normal operations use two pressure bands – an upper band from 2-3 atm and a lower band from 1-2 atm. The chain engages latching mechanisms at the control capstans of all elevator worm screws and Stewart platform struts. Cycling across the upper band steps the control chain forward, applying power to successive control capstans. Cycling across the bottom band steps the control

chain backwards, but without applying power to any control capstans. Cycling across both bands (from 1-3 atm) causes the control chain to step back and forth over the same control capstan, repeatedly applying power only to that one control capstan."

3. * Design other demuxes as deemed desirable. Evaluate the number of pairs of drive inputs that might be desirable. More drive inputs allow a higher data rate.

14.5 Mechanical means of distributing power and control driven by a demux

1. ** Design at least one binary Bowden cable.

2. ** Design a coupling so that two binary Bowden cables can be easily connected or disconnected. Provide at least the ability to couple (a) straight through and (b) at 90°. Consider bell cranks.

3. Consider how to fabricate and assemble pre-built straight binary Bowden cable segments with couplers at both ends that can be rapidly snapped together into longer runs of binary Bowden cables with couplers at both ends. Analyze whether this would simplify the manufacture of the many runs of binary Bowden cables used in the SRU.

4. Design other mechanical means of distributing power and control as seems appropriate. Possible means might include linearly driven rods, rotationally driven shafts, cables in tension, pairs of cables moving in opposition, hydraulic systems, and many other possibilities. Note that different inner and outer cables for binary Bowden cables, including differing chiralities of nanotubes, or the use of inner cables made of solid diamond rods, and many other possibilities for both inner and outer cables, provide a wide range of design options within the general space of "binary Bowden cables", all of which can be adapted to multi-valued or analog Bowden cables, which have utility in a variety of system contexts.

5. * Model and analyze selected systems designed here.

14.6 Conversion of linear to rotary motion

1. ** Design systems that convert the input from two binary Bowden cables to stepped rotary motion with steps of roughly 15° or less. See Driving a screw drive with four-phase logic (Section 7.18).

2. Design other approaches as might seem appropriate.

14.7 Positional mechanosynthetic systems

1. ** Design a number of different mechanosynthetic systems assembled from blocks. The minimum requirement is a single probe XYZ system (3DOF). A Stewart platform, although it provides 6DOF, can be simple to design with few parts.

2. ** Design a four probe system (a programmable fabricator) with each probe having 6DOF. Probes should be able to grasp tools from a tool rack. Workspace should be between 4 nm × 4 nm × 4 nm and 20 nm × 20 nm × 20 nm. Larger or smaller can be pursued with

appropriate justification. See the programmable fabricator in the SRU layout (Section 4.1.1) for a possible layout.

3. * Analyze the behavior of items 1 and 2.
4. Design a set of tools assembled from blocks.
5. Design a tool rack assembled from blocks.

14.8 Assembly tools

1. ** Design a set of interchangeable assembly tools able to grip all blocks used in the SRU fabricator. One end of each assembly tool should have a common handle. The other end should have a gripper, able to grip one or more blocks. A single type of tool holder can hold the handle of any tool. A tool holder affixed to the internal assembly arm allows it to "pick up" and "put down" any selected tool from a tool rack.
2. Note that there is a requirement that the SRU fabricator be able to pass an AP Block from the SRU fabricator to the internal assembly arm for assembly. One approach would be to pass a tool along with its gripped block from the SRU fabricator to the internal assembly arm. One method of doing this would be for the tool to have two handles, with the tool holder on the SRU fabricator able to hold the first handle and pass the tool to the internal assembly arm, which would attach its tool holder to the second handle before the tool holder on the SRU fabricator released the first handle. Alternatively, the AP Block might be gripped by two tools, one tool held by the SRU fabricator and the other tool held by the internal assembly arm. In this scenario, the gripper of the first tool (held by the SRU fabricator) would have to release the block so that the second tool, held by the internal assembly arm, could take possession of the block.
3. Review concepts that allow the tool holder to engage and disengage the handle. Concepts that rely only on the motion of the tool holder would be advantageous.
4. Review concepts that allow the gripper to grip/release a block.
5. What are the precision requirements for assembly operations?

14.9 Enumerate a subset of assembly operations sufficient for assembling the SRU

1. ** Enumerate possible assembly operations.
2. ** Propose one or more operational subsets of assembly operations such that each subset would be sufficient, by itself, to fully assemble the SRU.
3. * Develop ranking criteria by which to evaluate the subsets.
4. The subsets should be as small as possible.
5. Simple and reliable assembly operations should be given preference for inclusion in the subsets.
6. Irreversible assembly operations should be evaluated carefully before inclusion.

7. Dovetail joints, fusion of reactive surfaces, dipolar bonds, radical-radical coupling, and simple van der Waals forces are among the possibilities to be considered.

8. ** Design and model blocks that allow evaluation of all proposed assembly operations and then evaluate all assembly operations.

14.10 Describe feedback mechanisms to control SRU operations

1. ** Are simple binary test units sufficient to provide all needed feedback to control SRU operations?

2. * A molecular-scale analog of the qPlus force sensor should be feasible. The conventional qPlus is a quartz cantilever sensor derived from a tuning fork with one prong fixed [Gie98]. The SRU proposal as currently envisioned does not include this capability. Is it necessary? How might it be implemented? Will it require additional infrastructure in the SRU to implement? What benefits does it provide?

14.11 Determine the error rate of the SRU fabricator

1. ** Evaluate the positional uncertainty of the positional devices.

2. ** Evaluate the positional uncertainty of the molecular tools.

3. ** Evaluate the positional uncertainty of the workpiece.

4. ** What is the total positional uncertainty?

5. * For a given reaction sequence, evaluate the barrier heights to undesired side reactions. Evaluate the probability of those side reactions occurring.

6. * The SRU should be, with high probability, a eutactic environment, free of contaminants. Quantify the probability that the SRU internal environment might cease to be eutactic. What are the dominant failure mechanisms?

7. * Determine for a given reaction sequence if the Low Overhead Closed Loop Error Handling approach (Section 6.9) can successfully reduce the error rate to an acceptable level.

8. * Evaluate similar results for a selected set of representative reaction sequences.

9. ** Will small (100 to 1,000 atoms) rigid blocks be sufficient for all subsystems of the SRU?

14.12 A shell

1. Design as many tiling sets for the shell as seems appropriate. Tile size cannot exceed the build volume.

2. Design internal bracing (if needed) for the tiling sets of item 1. Would a space frame (Section 9.2.1) be useful as internal bracing?

3. Analyze the tiling sets and bracing designed in items 1 and 2. What is the stiffness of the bracing? Is it sufficient to provide an adequate metrological frame for the internal assembly arm? For the fabricator?

4. The Molecular LEGO (Section 7.6) approach can be used to generate tiling sets. There are also other approaches; see Decomposing large structures into small blocks (Section 7.4). Alternatives should be defined and analyzed.

14.13 Binding sites

1. * Design "tray" (or "drawer") style binding sites for tetraethylgermane.
2. Design tray-style binding sites for all baseline feedstock molecules. See Molecular composition of the feedstock (Section 7.1).
3. How effective are these designs at rejecting waste molecules and plausible contaminants? Is this "good enough"? (Analyze with respect to the baseline failure rate of 0.1% per generation.)
4. Design mechanisms to bind to baseline feedstock molecules while in the binding site trays using tools designed for this purpose.

14.14 A programmable SRU fabricator/metabolizer

1. ** Design a general-purpose programmable fabricator/metabolizer.
2. Design tools that bond to all the baseline feedstock molecules.
3. Design metabolic pathways that recharge GeRads and produce GeRad-H, GeRad-C2, GeRad-CH2, and other tools as needed from the baseline feedstock.
4. Propose alternative metabolic sequences for other non-baseline feedstocks, such as ethylene or acetylene.
5. Design and validate the baseline platinum-catalyzed H_2-generation subsystem, including hydrogen transfer from the molecular tool to the platinum collector and subsequent H_2 release. See Using platinum to generate H_2 gas (Section 4.6).
6. Evaluate the proposed AzAdam$^+$ reaction pathway as an alternative H_2-generation subsystem, including the complete tool-recharge cycle. See Generating H_2 using AzAdam (Section 12.2).
7. Carry out a trade-off study between different methods of disposing of excess H_2, including by the production of 1) H_2 gas, 2) H_2 gas in a container (such as a fullerene), 3) methane, 4) ethane, 5) polyethylene, or 6) some other hydrogen-rich structure.

14.15 Assembly arms

1. ** Design a Stewart platform (a.k.a. hexapod) made of blocks. This should be useable in both internal and external assembly arms, as well as in the programmable fabricator/metabolizer.
2. * Design internal assembly arm(s) made of blocks.
3. * Analyze the performance of the internal assembly arm(s).
4. * Design external assembly arm(s) made of blocks.

5. * Analyze the performance of the external assembly arm(s).

6. * Does a two-level Stewart platform provide adequate range of motion for the external assembly arm? How does the second stage attach to the first stage? Would this be helpful for the internal assembly arm?

14.16 Design electrical contacts between the SRU and the contact pillars

1. * Two spikes or screws could be driven into two adjacent contact pillars from a foot lowered from the SRU between two contact pillars. As a minimum of six contacts is required, three feet making two connections each would suffice.

2. Carry out a trade-off study of as many methods of electrically connecting an SRU to the contact pillars on a die surface as seems appropriate.

14.17 Extrusion and placement of the daughter SRU by its parent

1. * Design an extrusion mechanism.

2. Analyze the performance of the extrusion mechanism.

3. * How accurately must SRUs be positioned with respect to the die? With respect to each other?

4. * Design alignment features for the surface of the die.

5. * Design the process by which an extruded SRU is positioned on the die surface in proper alignment next to its parent by the external assembly arm(s).

6. * Review Sliding Seal Extrusion Risk Mitigation (Section 12.12) strategies and determine which extrusion strategies to pursue.

14.18 Internal layout of SRU

1. Review the SRU layout (Section 4.1.1).

2. Design a layout assuming only the SRU fabricator is in vacuum (the "flooded SRU"). How do you assemble a differential comb drive from blocks? How well does the differential comb drive function when not in vacuum? Can it be a sealed unit?

3. Design a layout that operates in vacuum and has feedstock input and waste discharge pipes. Develop scalable system concepts that provide input and discharge pipes.

4. Design a layout assuming positional devices other than Stewart platforms.

5. Design other layouts as seems appropriate.

6. Which layout is most easily built in a workstation?

14.19 Removal of test items from workstation into liquid environment

1. Design an external temperature controlled test environment. Our baseline temperature is 210 K. Review this assumption. Determine what range of temperatures might be appropriate for this system. Is 77 K to 210 K appropriate? "It's interesting to note that 63% propane, 37% ethane (molar concentration) remains liquid at 77 K" [TBJ08].

2. Design binding site/tray mechanism. Design test strategy.

3. Design sliding seal mechanism. Design test strategy.

4. Design electrical contact mechanism (hook and slot, grapplers, spikes, screws, other). Design test strategy.

5. Design extrusion mechanism. Design test strategy.

14.20 Removal of First SRU from workstation

1. Design a system to transfer the first SRU on a workstation die (∼1 mm × 1 mm) from the workstation to the Early Growth Environment.

2. Design the Early Growth Environment and the process for growth from First SRU to a million or so SRUs (an array of 1,000 × 1,000, less than one mm^2 in area) in the Early Growth Environment.

3. The Early Growth Environment might be one mm^3 of liquid feedstock maintained at 210 K with the workstation die immersed in it (after removal from the workstation) with the First SRU on the workstation die. Wiring would connect the workstation die to a control computer.

4. Design the transfer of early SRU(s) (possibly taken from an initial 1,000 × 1,000 array of SRUs) from Early Growth Environment onto production die(s).

14.21 Design and Manufacture of Dies

1. Instead of 1 cm^2 dies, explore the possibility of using dies of other sizes made from wafers patterned by electron-beam lithography, as these can be made relatively inexpensively at many facilities doing research and development.

2. Explore the design of nanofluidic channels on the die surface as part of evaluating the Fifth Strategy (Section 12.12.7).

3. Describe how alignment features are made on the surface of the die.

4. Analyze the impact of placing a die into the baseline feedstock solution at the proposed baseline operating temperature (or other proposed operating temperatures). Could we operate at a lower temperature, such as 77 K (consider the mix of propane and ethane at 77 K that remains liquid)?

5. * What is the conductivity of tetraethylgermane? How might it be changed by dissolved ionic solutes (as dissolved salt changes the conductivity of water)? Of other proposed feedstock solutions? Can the risk of a conductive solution be mitigated by adding an insulating surface barrier to the die during manufacture?

6. * What is the dielectric constant of tetraethylgermane? Address the same questions as in item 5.

7. See also Die issues (Section 8).

14.22 Design feedback systems from multiple SRUs on die to central computer

1. * Design electronic feedback system(s).

2. Design optical feedback system(s) from individual SRUs on a die.

3. Design optical output device(s) from individual SRUs (possibly using modulated nitrogen-vacancy (NV) centers, LEDs, or other devices).

4. Determine power requirements of LEDs and NV centers.

5. Evaluate bandwidth of optical and electronic feedback devices.

6. Design optical sensing system to select and demodulate optical signals from individual SRUs on a die.

7. Do a trade-off study of various feedback concepts.

8. * Can a dead SRU be removed from a die? Can it be replaced?

9. Can we use a scanning electron microscope (SEM) to image a functioning SRU? Will it interfere with SRU function?

14.23 Design of external support subsystems

1. Die tank and pumps.

2. Die interface (mechanical and electrical).

3. Cooling subsystem.

4. Feedstock filtration and product extraction.

5. Power subsystem.

6. Hardware computational subsystem.

7. Software.

14.24 Various other things

1. * Enumerate all other things needed for the SRU concept.

2. Develop layout plans for the subsystems described in item 1.

14.25 SRU replication and systems integration

1. Develop and analyze an end-to-end SRU replication sequence, building on the subsystem descriptions above. Identify the mechanical, electrical, and control conditions required at each transition, including daughter handoff, placement, and activation. Distinguish completion of daughter assembly from readiness to begin replication.

2. Verify that the parent can carry out each step with the available tools, workspace, and access paths. Check support and alignment of partially assembled structures, and preservation of the seal as the daughter's exterior features pass through it.

3. Reconcile the subsystem budgets for time, force, positional uncertainty, volume, power, and error recovery across the complete replication sequence. Identify shared constraints and quantify margins where possible. Distinguish actuation power from total heat dissipation, including electrical switching losses.

4. Develop and verify a collision-free schedule for daughter EAA attachment, handoff, and placement in a densely populated array. Include any required grouping of SRUs and account for its effect on replication and die-population time.

5. Evaluate whether the proposed error-detection and shutdown mechanisms adequately support the fail-stop model. Consider undetected errors, incorrect continued operation, and failures that obstruct daughter handoff or placement. Analyze common-cause failures separately, including those arising from shared instructions, design defects, or feedstock contamination.

6. Test components and progressively integrated mechanisms under their intended loads and operating environments. Distinguish successful assembly, functional performance, and demonstrated reliability, and identify which conclusions rest on simulation rather than experiment. Distinguish measurements available during laboratory testing from checks an operating SRU can perform.

15 References

[AHP] AHP Seals. Swivel seals. URL: https://ahpseals.com/swivel-seals/ [cited 2026-10-15].

[And93] J. L. Anderson. Horizon Mission Methodology—A Tool for the Study of Technology Innovation and New Paradigms. In *AIAA Paper 93-1134*, 1993. URL: https://ntrs.nasa.gov/citations/19930047019.

[And96] J. L. Anderson. Leaps of the Imagination: Interstellar Flight and the Horizon Mission Methodology. *Journal of the British Interplanetary Society*, 49:15–20, 1996. URL: https://bis-space.com/shop/product/leaps-of-the-imagination-interstellar-flight-and-the-horizon-mission-methodology/.

[Appa] Applied Diamond, Inc. Diamond Properties. Property chart: https://usappliedddiamond.com/wp-content/uploads/2016/02/chart3.pdf; background: https://usappliedddiamond.com/why-diamond/. URL: https://usappliedddiamond.com/products/ [cited 2026-10-06].

[Appb] Applied Diamond, Inc. Why Diamond? URL: https://usappliedddiamond.com/why-diamond/ [cited 2026-10-06].

[AS02] D. G. Allis and J. T. Spencer. The Design of Carbon Nanotube-Based Dative Structures from Supramolecular Principles. In *Foresight Institute Tenth Conference on Nanotechnology*, Bethesda, MD, Oct. 2002. Mirror: https://web.archive.org/web/*/somewhereville.com. URL: https://www.somewhereville.com/the-design-of-carbon-nanotube-based-dative-structures-from-supramolecular-principles/.

[ASM] ASML. EUV lithography systems. Accessed October 16, 2026. URL: https://www.asml.com/en/products/euv-lithography-systems.

[B+12] L. Britnell et al. Atomically thin boron nitride: a tunnelling barrier for graphene devices. *Nano Letters*, 12(3):1707–1710, 2012. URL: https://arxiv.org/abs/1202.0735, doi:10.1021/nl3002205.

[B+21] E. Bussmann et al. Atomic-precision advanced manufacturing for Si quantum computing. *MRS Bulletin*, 46:607–615, 2021. URL: https://link.springer.com/article/10.1557/s43577-021-00139-8, doi:10.1557/s43577-021-00139-8.

[B+25] E. Barrera et al. Inverted-Mode Scanning Tunneling Microscopy for Atomically Precise Fabrication. arXiv preprint, Dec. 2025. arXiv:2512.24431, doi:10.48550/arXiv.2512.24431.

[B+26] B. Blue et al. Towards Atom-by-Atom Fabrication: Mechanosynthetic donation and abstraction. arXiv preprint, June 2026. arXiv:2606.13876, doi:10.48550/arXiv.2606.13876.

[BBTS20] M. C. Bowman, A. D. Burke, J. M. Turney, and H. F. Schaefer, III. Conclusive determination of ethynyl radical hydrogen abstraction energetics and kinetics. *Molecular Physics*, 118(19–20):e1769214, 2020. doi:10.1080/00268976.2020.1769214.

[BCK19] D. Balaban, J. Cooper, and E. Komendera. Inverse Kinematics and Sensitivity Minimization of an n-Stack Stewart Platform. In *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019)*, pages 1793–1795, Montréal, Canada, May 2019. International Foundation for Autonomous Agents and Multiagent Systems. Extended abstract. URL: https://www.ifaamas.org/Proceedings/aamas2019/pdfs/p1793.pdf.

[BMC+09] J. E. Butler, Y. A. Mankelevich, A. Cheesman, J. Ma, and M. N. R. Ashfold. Understanding the chemical vapor deposition of diamond: recent progress. *Journal of Physics: Condensed Matter*, 21(36):364201, 2009. doi:10.1088/0953-8984/21/36/364201.

[Bog14] E. Bogatin. Estimating wire & loop inductance: Rule of Thumb #15. EDN, 2014. URL: https://www.edn.com/estimating-wire-loop-inductance-rule-of-thumb-15/.

[BS19] C. Beshara and G. Shustov. Inside-protonated 1-azaadamantane: computational studies on the structure, stability, and generation. *Canadian Journal of Chemistry*, 97(3):169–177, Mar. 2019. URL: https://utoronto.scholaris.ca/server/api/core/bitstreams/68112999-1e41-436b-9a0f-815880f3970c/content, doi:10.1139/cjc-2018-0407.

[C+16] N. Cheng et al. Platinum single-atom and cluster catalysis of the hydrogen evolution reaction. *Nature Communications*, 7:13638, 2016. doi:10.1038/ncomms13638.

[C+26] M. Cowie et al. Atomically precise mechanosynthesis of carbon structures on hydrogenated Si(100) by inverted-mode STM. arXiv preprint, May 2026. arXiv:2605.27250, doi:10.48550/arXiv.2605.27250.

[CBN] CBN Nano Technologies. Building the Atomic Manufacturing Platform. URL: https://cbnnanotechnologies.com/ [cited 2026-10-06].

[Cen24] Center for Strategic and International Studies. A Strategy for The United States to Regain its Position in Semiconductor Manufacturing, 2024. Archived: https://web.archive.org/web/2024*/csis.org/analysis/strategy-united-states-regain-its-position-semiconductor-manufacturing. URL: https://www.csis.org/analysis/strategy-united-states-regain-its-position-semiconductor-manufacturing.

[CEP76] K. Christmann, G. Ertl, and T. Pignet. Adsorption of hydrogen on a Pt(111) surface. *Surface Science*, 54(2):365–392, 1976. doi:10.1016/0039-6028(76)90232-6.

[Cor67] E. J. Corey. General Methods for the Construction of Complex Molecules. *Pure and Applied Chemistry*, 14:19–38, 1967. URL: https://publications.iupac.org/pac/pdf/1967/pdf/1401x0019.pdf.

[Cor88] E. J. Corey. Retrosynthetic Thinking—Essentials and Examples. *Chemical Society Reviews*, 17:111–133, 1988.

[Dar70] B. deB. Darwent. Bond dissociation energies in simple molecules. National Standard Reference Data Series 31, National Bureau of Standards, 1970. URL: https://nvlpubs.nist.gov/nistpubs/Legacy/NSRDS/nbsnsrds31.pdf, doi:10.6028/NBS.NSRDS.31.

[dBE16] M. L. den Boer and J. A. A. W. Elemans. Triggering chemical reactions by Scanning Tunneling Microscopy: From atoms to polymers. *European Polymer Journal*, 83:390–406, 2016. doi:10.1016/j.eurpolymj.2016.03.002.

[DK11] A. A. Dzhioev and D. S. Kosov. Kramers problem for nonequilibrium current-induced chemical reactions. *Journal of Chemical Physics*, 135:074701, 2011. doi:10.1063/1.3626521.

[Dre86] K. E. Drexler. *Engines of Creation*. Doubleday, 1986.

[Dre92] K. E. Drexler. *Nanosystems: Molecular Machinery, Manufacturing, and Computation*. John Wiley & Sons, New York, 1992. URL: https://web.archive.org/web/20250328091942/https://nanosyste.ms/.

[E+25] H. Edlbauer et al. An 11-qubit atom processor in silicon. *Nature*, 648:569–575, 2025. URL: https://www.nature.com/articles/s41586-025-09827-w, doi:10.1038/s41586-025-09827-w.

[Erm88] O. Ermer. Five-fold diamond structure of adamantane-1,3,5,7-tetracarboxylic acid. *Journal of the American Chemical Society*, 110(12):3747–3754, 1988. URL: https://pubs.acs.org/doi/pdf/10.1021/ja00220a005, doi:10.1021/ja00220a005.

[ES90] D. M. Eigler and E. K. Schweizer. Positioning single atoms with a scanning tunnelling microscope. *Nature*, 344:524–526, 1990. URL: https://www.nature.com/articles/344524a0, doi:10.1038/344524a0.

[Fey60] R. P. Feynman. There's plenty of room at the bottom: An Invitation to Enter a New Field of Physics. *Engineering and Science*, 23(5):22–36, Feb. 1960. URL: https://resolver.caltech.edu/CaltechES:23.5.1960Bottom.

[FM04a] R. A. Freitas, Jr. and R. C. Merkle. Freitas Biphase Assembler. In *Kinematic Self-Replicating Machines*, chapter 4.18. Landes Bioscience, Georgetown, TX, 2004. Concept proposed in 2000. URL: https://www.molecularassembler.com/KSRM/4.18.htm.

[FM04b] R. A. Freitas, Jr. and R. C. Merkle. *Kinematic Self-Replicating Machines*. Landes Bioscience, Georgetown, TX, 2004. URL: https://www.molecularassembler.com/KSRM.htm.

[FM08a] R. A. Freitas, Jr. and R. C. Merkle. A Minimal Toolset for Positional Diamond Mechanosynthesis. *Journal of Computational and Theoretical Nanoscience*, 5(5):760–861, May 2008. URL: http://www.MolecularAssembler.com/Papers/MinToolset.pdf, doi:10.1166/JCTN.2008.2531.

[FM08b] R. A. Freitas, Jr. and R. C. Merkle. A Nanofactory Roadmap: Research Proposal for a Comprehensive Diamond Nanofactory Development Program. Technical Report 58, Institute of Molecular Manufacturing, 2008. Updated 2025. URL: http://www.imm.org/Reports/rep058.pdf.

[FM12a] R. A. Freitas, Jr. and R. C. Merkle. Positional diamondoid mechanosynthesis, May 2012. U.S. Patent US8171568B2. URL: https://patents.google.com/patent/US8171568B2/en.

[FM12b] R. A. Freitas, Jr. and R. C. Merkle. Positional diamondoid mechanosynthesis, September 2012. U.S. Patent US8276211B1. URL: https://patents.google.com/patent/US8276211B1/en.

[FM17] R. A. Freitas, Jr. and R. C. Merkle. Build sequences for mechanosynthesis, June 2017. U.S. Patent US9676677B2. URL: https://patents.google.com/patent/US9676677B2/en.

[Fre99] R. A. Freitas, Jr. *Nanomedicine, Volume I: Basic Capabilities*. Landes Bioscience, Georgetown, TX, 1999. URL: https://www.nanomedicine.com/NMI.htm.

[Fre09] R. A. Freitas, Jr. Molecular Workstation Roadmap I: Survey of Key Technologies. Technical Report 59, Institute of Molecular Manufacturing, 2009. Updated 2025. URL: http://www.imm.org/Reports/rep059.pdf.

[Fre10] R. A. Freitas, Jr. Molecular Workstation Roadmap II: Survey of Single-Atom Tips. Technical Report 60, Institute of Molecular Manufacturing, 2010. Updated 2025. URL: http://www.imm.org/Reports/rep060.pdf.

[GAMJ10] J. C. Garcia, L. V. C. Assali, W. V. M. Machado, and J. F. Justo. Crystal engineering using functionalized adamantane. *Journal of Physics: Condensed Matter*, 22(31):315303, 2010. URL: https://arxiv.org/abs/1204.2863, doi:10.1088/0953-8984/22/31/315303.

[Gel15] Gelest, Inc. Tetraethylgermane Safety Data Sheet. Technical Report GET7150, Gelest, Inc., Jan. 2015. CAS 597-63-7. URL: https://www.gelest.com/wp-content/uploads/product_msds/GET7150-msds.pdf.

[Gie98] Franz J. Giessibl. High-speed force sensor for force microscopy and profilometry utilizing a quartz tuning fork. *Applied Physics Letters*, 73(26):3956–3958, 1998. URL: https://epub.uni-regensburg.de/25327/, doi:10.1063/1.122948.

[GJMA09] J. C. Garcia, J. F. Justo, W. V. M. Machado, and L. V. C. Assali. Functionalized adamantane: building blocks for nanostructure self-assembly. *Physical Review B*, 80:125421, 2009. doi:10.1103/PhysRevB.80.125421.

[GSI08] L. Guan, K. Suenaga, and S. Iijima. Smallest Carbon Nanotube Assigned with Atomic Resolution Accuracy. *Nano Letters*, 8(2):459–462, 2008. doi:10.1021/nl072396j.

[H+25] T. Huff et al. Molecular Tools for Non-Planar Surface Chemistry. arXiv preprint, Aug. 2025. arXiv:2508.16798, doi:10.48550/arXiv.2508.16798.

[HBMR00] S. W. Hla, L. Bartels, G. Meyer, and K. H. Rieder. Inducing All Steps of a Chemical Reaction with the Scanning Tunneling Microscope Tip: Towards Single Molecule Engineering. *Physical Review Letters*, 85(13):2777–2780, Sept. 2000. URL: https://courses.cs.duke.edu/spring04/cps296.4/papers/HBMR00.pdf.

[Her22] J. Hertz. The Fundamentals of a Charge Pump Circuit. All About Circuits, May 2022. URL: https://www.allaboutcircuits.com/technical-articles/switched-capacitor-circuits-charge-pump-circuit-basics/.

[HMA17] T. Hogg, M. S. Moses, and D. G. Allis. Evaluating the friction of rotary joints in molecular machines. *Molecular Systems Design & Engineering*, 2:235–252, 2017. URL: https://arxiv.org/abs/1701.08202, doi:10.1039/C7ME00021A.

[HNR68] P. E. Hart, N. J. Nilsson, and B. Raphael. A Formal Basis for the Heuristic Determination of Minimum Cost Paths. *IEEE Transactions on Systems Science and Cybernetics*, 4(2):100–107, 1968. URL: https://ai.stanford.edu/~nilsson/OnlinePubs-Nils/PublishedPapers/astar.pdf, doi:10.1109/TSSC.1968.300136.

[HR03] S. W. Hla and K. H. Rieder. STM Control of Chemical Reactions: Single-Molecule Synthesis. *Annual Review of Physical Chemistry*, 54:307–330, 2003. URL: https://www.annualreviews.org/content/journals/10.1146/annurev.physchem.54.011002.103852, doi:10.1146/annurev.physchem.54.011002.103852.

[HTS14] T. T. T. Hanh, Y. Takimoto, and O. Sugino. First-principles thermodynamic description of hydrogen electroadsorption on the Pt(111) surface. *Surface Science*, 625:104–111, 2014. doi:10.1016/j.susc.2014.03.006.

[Iri22] K. K. Irikura. Theoretical dissociation, ionization, and spin-orbit energetics of the diatomic platinum hydrides PtH, PtH^+, and PtH^-. *Journal of Chemical Physics*, 157:104304, 2022. doi:10.1063/5.0107501.

[Jan21] L. Janssen. The Split-Ring Compound Planetary - My Highest Reduction Yet. YouTube video, 2021. URL: https://www.youtube.com/watch?v=4gkbb2CBzQI.

[Job] Mark Jobes. [Two-stage split-ring compound planetary gear]. Unpublished internal presentation; descriptive title.

[Jua18] Juan Gg. Split ring compound planetary gear, Feb. 2018. URL: https://juangg-projects.blogspot.com/2018/02/split-ring-compound-epicyclicplanetary.html.

[JZ11] X. Jiang and J. Zhao. Mechanical and electronic properties of diamond nanowires under tensile strain from first principles. *Nanotechnology*, 22(40):405705, 2011. doi:10.1088/0957-4484/22/40/405705.

[K+21] T. Knobloch et al. The performance limits of hexagonal boron nitride as an insulator for scaled CMOS devices based on two-dimensional materials. *Nature Electronics*, 4:98–108, 2021. doi:10.1038/s41928-020-00529-x.

[Kan98] B. E. Kane. A silicon-based nuclear spin quantum computer. *Nature*, 393:133–137, 1998.

[Kel81] B. T. Kelly. *Physics of Graphite*. Applied Science Publishers, London, 1981.

[LGE96] R. Legtenberg, A. W. Groeneveld, and M. Elwenspoek. Comb-drive actuators for large displacements. *Journal of Micromechanics and Microengineering*, 6:320–329, 1996. URL: https://ris.utwente.nl/ws/files/6508177/comb-drive.pdf.

[LHL08] Eric W. Lemmon, Marcia L. Huber, and Jacob W. Leachman. Revised standardized equation for hydrogen gas densities for fuel consumption applications. *Journal of Research of the National Institute of Standards and Technology*, 113(6):341–350, 2008. URL: https://nvlpubs.nist.gov/nistpubs/jres/113/6/V113.N06.A05.pdf.

[LJPL09] J. W. Leachman, R. T. Jacobsen, S. G. Penoncello, and E. W. Lemmon. Fundamental Equations of State for Parahydrogen, Normal Hydrogen, and Orthohydrogen. *Journal of Physical and Chemical Reference Data*, 38:721–748, 2009. Tabulated density used here derived from the NIST normal-hydrogen equation of state. doi:10.1063/1.3160306.

[M+23] R. C. Merkle et al. Systems and methods for mechanosynthesis, July 2023. U.S. Patent US11708384B2. URL: https://patents.google.com/patent/US11708384B2/en.

[Mac] James MacArthur. [molecular screw drive]. Unpublished report; descriptive title.

[Mer93a] R. C. Merkle. A proof about molecular bearings. *Nanotechnology*, 4(2):86–90, 1993. doi:10.1088/0957-4484/4/2/004.

[Mer93b] R. C. Merkle. Two Types of Mechanical Reversible Logic. *Nanotechnology*, 4(2):114–131, 1993. URL: https://www.zyvex.com/nanotech/mechano.html, doi:10.1088/0957-4484/4/2/007.

[Mer97] R. C. Merkle. Convergent assembly. *Nanotechnology*, 8(1):18–22, 1997. URL: https://www.zyvex.com/nanotech/convergent.html, doi:10.1088/0957-4484/8/1/005.

[Mer00] R. C. Merkle. Molecular building blocks and development strategies for molecular nanotechnology. *Nanotechnology*, 11(2):89–99, 2000. URL: https://www.zyvex.com/nanotech/mbb/mbb.html, doi:10.1088/0957-4484/11/2/309.

[Mer14] R. C. Merkle. Molecular Space Frames: An Atomically Precise Aerogel. Technical Report 44, Institute of Molecular Manufacturing, 2014. URL: http://www.imm.org/Reports/rep044.pdf.

[Mer16] R. C. Merkle. Zero energy computing, 2016. Version 1.21, unpublished manuscript.

[Mer23] R. C. Merkle. Picometer Hydraulic Pistons, 2023. Version 1.0. URL: https://www.zyvex.com/nanotech/Picometer%20Hydraulic%20Pistons%20V1.0.pdf.

[MFA25] R. C. Merkle, R. A. Freitas, Jr., and D. G. Allis. Design of a molecular Field Effect Transistor (mFET). Technical Report 56, Institute of Molecular Manufacturing, Mar. 2025. URL: http://www.imm.org/Reports/rep056.pdf.

[MFH+18] R. C. Merkle, R. A. Freitas, Jr., T. Hogg, T. E. Moore, M. S. Moses, and J. Ryley. Mechanical computing systems using only links and rotary joints. *Journal of Mechanisms and Robotics*, 10(6):061006, 2018. URL: https://arxiv.org/abs/1801.03534, doi:10.1115/1.4041209.

[MFR+19] R. C. Merkle, R. A. Freitas, Jr., J. Ryley, M. S. Moses, and T. Hogg. Mechanical Computing Systems, November 2019. U.S. Patent US10481866B2. URL: https://patents.google.com/patent/US10481866B2/en.

[MFR23] J. MacArthur, R. A. Freitas, Jr., and J. F. Ryley, III. Selective Motion Transfer Mechanism, September 2023. U.S. Patent Application Publication US20230296163A1. URL: https://patents.google.com/patent/US20230296163A1/.

[MHM+74] Philip H. Mogul, Marc C. Hochberg, Robert Michiel, Glenn K. Nestel, Brian L. Wamsley, and Seth D. Coren. Physical properties of tetra-n-alkylgermanes (C1–C6). *Journal of Chemical & Engineering Data*, 19(1):4–6, 1974. doi:10.1021/je60060a029.

[MLHU17] D. Mandelli, I. Leven, O. Hod, and M. Urbakh. Sliding friction of graphene/hexagonal-boron nitride heterojunctions: a route to robust superlubricity. *Scientific Reports*, 7:10851, Sept. 2017. URL: https://www.nature.com/articles/s41598-017-10522-8.pdf, doi:10.1038/s41598-017-10522-8.

[Moo97] Gordon E. Moore. An update on Moore's law. Intel Developer Forum, Fall 1997, San Francisco, 1997. Keynote transcript, September 30. URL: https://www.intel.com/pressroom/archive/speeches/GEM93097.HTM.

[Mos] Matt Moses. Personal communication. Suggestion to use a reversed harmonic drive.

[MPMG91] C. B. Musgrave, J. K. Perry, R. C. Merkle, and W. A. Goddard, III. Theoretical studies of a hydrogen abstraction tool for nanotechnology. *Nanotechnology*, 2(4):187–195, 1991. doi:10.1088/0957-4484/2/4/004.

[MPS03] R. C. Merkle, E. G. Parker, and G. D. Skidmore. Method and system for self-replicating manufacturing stations, January 2003. U.S. Patent US6510359B1. URL: https://patents.google.com/patent/US6510359B1/en.

[N+14] A. Niguès et al. Ultrahigh interlayer friction in multiwalled boron nitride nanotubes. *Nature Materials*, 13:688–693, June 2014. URL: https://ilm.univ-lyon1.fr/images/ILM/Equipe_17/articles/articlePNEC/Nature-Mat-friction-Nigues-2014.pdf, doi:10.1038/NMAT3985.

[N+19] A. Nie et al. Approaching diamond's theoretical elasticity and strength limits. *Nature Communications*, 10:5533, 2019. doi:10.1038/s41467-019-13378-w.

[Nata] National Institute of Standards and Technology. Fundamental Physical Constants. See also: https://physics.nist.gov/cuu/pdf/wall_2022.pdf. URL: https://physics.nist.gov/cuu/Constants/ [cited 2026-10-06].

[Natb] National Institute of Standards and Technology. NIST Guide to the SI, Chapter 5: Units Outside the SI. See also: https://www.nist.gov/glossary/osac-lexicon. URL: https://www.nist.gov/pml/special-publication-811/nist-guide-si-chapter-5-units-outside-si [cited 2026-10-06].

[Natc] National Oceanic and Atmospheric Administration. Germane: CAMEO chemicals. NFPA 704 health rating: 4. URL: https://cameochemicals.noaa.gov/chemical/809 [cited 2026-10-15].

[Nee] Ian Neeser. Liquid hydrogen bulk storage introduction. Chart Industries presentation, hosted by the U.S. Department of Energy. Fluid-properties table, slide 2. URL: https://www.energy.gov/sites/default/files/2022-03/Liquid%20H2%20Workshop-Chart.pdf [cited 2026-10-15].

[O+03] N. Oyabu et al. Mechanical Vertical Manipulation of Selected Single Atoms by Soft Nanoindentation Using Near Contact Atomic Force Microscopy. *Physical Review Letters*, 90:176102, 2003. URL: https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.90.176102.

[Ott] Otto Chemie Pvt. Ltd. Hexamine, 99%: Specification sheet. Product H 1355; CAS 100-97-0. URL: https://www.ottokemi.com/specsheet/hexamine.aspx [cited 2026-10-15].

[PFM04] J. Peng, R. A. Freitas, Jr., and R. C. Merkle. Theoretical analysis of diamond mechanosynthesis. Part I. Stability of C2-mediated growth of nanocrystalline diamond C(110) surface. *Journal of Computational and Theoretical Nanoscience*, 1(1):62–70, 2004. URL: http://www.MolecularAssembler.com/Papers/JCTNPengMar04.pdf, doi:10.1166/jctn.2004.007.

[PPEL25] A. Prieto, B. Prieto, J. J. Escobar, and T. Lampert. Evolution of computing energy efficiency: Koomey's Law revisited. *Cluster Computing*, 28:42, 2025. doi:10.1007/s10586-024-04767-y.

[RL07] Henna Ruuska and Karin Larsson. Surface reactivities of (111), (100), and (110) planes of c-BN: A quantum mechanical approach. *Diamond and Related Materials*, 16(1):118–123, 2007. doi:10.1016/j.diamond.2006.04.016.

[RN20] S. J. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Pearson, 4th edition, 2020. URL: https://aima.cs.berkeley.edu/~russell/aima/translations.html.

[Rob13] P. Roberts. A Novel Passive Robotic Tool Interface. In *15th European Space Mechanisms & Tribology Symposium (ESMATS 2013)*, Noordwijk, The Netherlands, Sept. 2013. URL: https://www.esmats.eu/esmatspapers/pastpapers/pdfs/2013/roberts.pdf.

[Ryl] James Ryley. [Molecular screw drive]. Unpublished report; descriptive title.

[SA01] M. Shiraishi and M. Ata. Work function of carbon nanotubes. *Carbon*, 39(12):1913–1917, Oct. 2001. doi:10.1016/S0008-6223(00)00322-5.

[SAR23] K. Salini, P. S. Arya, and M. Rangaswamy. Queueing-inventory Systems: A Survey. arXiv preprint, Aug. 2023. arXiv:2308.06518, doi:10.48550/arXiv.2308.06518.

[Sau] P. Saugstad. Split Ring Planetary Gear Calculator. URL: https://saugstad.net/gear-animation/calculate.html [cited 2026-10-06].

[SBR12] Y. Sun, S. A. Boggs, and R. Ramprasad. The intrinsic electrical breakdown strength of insulators from first principles. *Applied Physics Letters*, 101:132906, 2012. doi:10.1063/1.4755841.

[Sch01] C. Schell. Voltage Doubler Design and Analysis. Technical Report AN-1119, National Semiconductor, June 2001. Reissued by Texas Instruments as SNAA095. URL: https://www.ti.com/lit/an/snaa095/snaa095.pdf.

[Sea19] Seagate Technology. Exos X16 data sheet, 2019. DS2011.1-1904US, April 2019. Maximum sustained outer-diameter transfer rate: 261 MB/s. URL: https://www.seagate.com/content/dam/seagate/migrated-assets/www-content/datasheets/pdfs/exos-x16-DS2011-1-1904US-en_US.pdf [cited 2026-10-15].

[Sim63] J. G. Simmons. Generalized Formula for the Electric Tunnel Effect between Similar Electrodes Separated by a Thin Insulating Film. *Journal of Applied Physics*, 34(6):1793–1803, 1963. doi:10.1063/1.1702682.

[SLJ+00] Roland Span, Eric W. Lemmon, Richard T. Jacobsen, Wolfgang Wagner, and Akimichi Yokozeki. A reference equation of state for the thermodynamic properties of nitrogen for temperatures from 63.151 to 1000 K and pressures to 2200 MPa. *Journal of Physical and Chemical Reference Data*, 29(6):1361–1433, 2000. URL: https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=907386, doi:10.1063/1.1349047.

[SS83] R. D. Schlichting and F. B. Schneider. Fail-stop processors: an approach to designing fault-tolerant computing systems. *ACM Transactions on Computer Systems*, 1(3):222–238, Aug. 1983. URL: https://www.cs.cornell.edu/fbs/publications/Fail_Stop.pdf, doi:10.1145/357369.357371.

[STGH18] J. F. Shortle, J. M. Thompson, D. Gross, and C. M. Harris. *Fundamentals of Queueing Theory*. Wiley Series in Probability and Statistics. Wiley, 5th edition, 2018.

[Sze81] S. M. Sze. *Physics of Semiconductor Devices*. Wiley, New York, 2nd edition, 1981.

[TAI+10] D. Tarasov, N. Akberova, E. Izotova, D. Alisheva, M. Astafiev, and R. A. Freitas, Jr. Optimal Tooltip Trajectories in a Hydrogen Abstraction Tool Recharge Reaction Sequence. *Journal of Computational and Theoretical Nanoscience*, 7(2):325–353, 2010. URL: https://www.molecularassembler.com/Papers/TarasovFeb2010.pdf.

[TBJ08] W. F. Tivol, A. Briegel, and G. J. Jensen. An Improved Cryogen for Plunge Freezing. *Microscopy and Microanalysis*, 14(5):375–379, Oct. 2008. doi:10.1017/S1431927608080781.

[TIA+13] D. Tarasov, E. Izotova, D. Alisheva, N. Akberova, and R. A. Freitas, Jr. Optimal Approach Trajectories for a Hydrogen Donation Tool in Positionally Controlled Diamond Mechanosynthesis. *Journal of Computational and Theoretical Nanoscience*, 10(9):1899–1907, 2013. URL: https://www.molecularassembler.com/Papers/TarasovSep2013.pdf.

[TNH89] W. C. Tang, T.-C. H. Nguyen, and R. T. Howe. Laterally Driven Polysilicon Resonant Microstructures. *Sensors and Actuators*, 20:25–32, 1989. URL: https://wctgroup.eng.uci.edu/publications/journal/J001-1989-S%26A-Tang.pdf.

[TSA21] M. Tolladay, F. Scarpa, and N. L. Allan. Interatomic forces breaking carbon-carbon bonds. *Carbon*, 175:420–428, 2021. URL: https://research-information.bris.ac.uk/files/263734146/bonbreaking2019_2.pdf, doi:10.1016/j.carbon.2020.12.088.

[TSMF06] B. Temelso, C. D. Sherrill, R. C. Merkle, and R. A. Freitas, Jr. High-level ab initio studies of hydrogen abstraction from prototype hydrocarbon systems. *Journal of Physical Chemistry A*, 110(38):11160–11173, 2006. URL: http://www.MolecularAssembler.com/Papers/TemelsoHAbst.pdf, doi:10.1021/jp061821e.

[TSMF07] B. Temelso, C. D. Sherrill, R. C. Merkle, and R. A. Freitas, Jr. Ab initio thermochemistry of the hydrogenation of hydrocarbon radicals using silicon-, germanium-, tin-, and lead-substituted methane and isobutane. *Journal of Physical Chemistry A*, 111(35):8677–8688, 2007. URL: http://www.MolecularAssembler.com/Papers/TemelsoHDon.pdf, doi:10.1021/jp071797k.

[Uni] University of Bristol, School of Chemistry. Principal Properties of Diamond. The page credits Spear and Dismukes, Synthetic Diamond—Emerging CVD Science and Technology, Wiley, New York, 1994; the book was not independently consulted. URL: https://www.chm.bris.ac.uk/motm/diamond/diamprop.htm [cited 2026-11-02].

[vN66] J. von Neumann. *Theory of Self-Reproducing Automata*. University of Illinois Press, Urbana, IL, 1966. Edited by A. W. Burks. URL: https://archive.org/details/theoryofselfrepr00vonn_0/page/n5/mode/2up.

[WH10] S. W. Wu and W. Ho. Two-photon-induced hot-electron transfer to a single molecule in a scanning tunneling microscope. *Physical Review B*, 82:085444, 2010. URL: https://doi.org/10.48550/arXiv.1008.2414, doi:10.1103/PhysRevB.82.085444.

[WH11] N. H. E. Weste and D. M. Harris. *CMOS VLSI Design: A Circuits and Systems Perspective*. Addison-Wesley, 4th edition, 2011. URL: https://pages.hmc.edu/harris/cmosvlsi/4e/lect/lect14.pdf.

[Wik] Wikipedia. Space frame. URL: https://en.wikipedia.org/wiki/Space_frame [cited 2026-10-06].

[WNC+17] Chao Wang, Sung-Wook Nam, John M. Cotte, Christopher V. Jahnes, Evan G. Colgan, Robert L. Bruce, Markus Brink, Michael F. Lofaro, Jyotica V. Patel, Lynne M. Gignac, Eric A. Joseph, Satyavolu Papa Rao, Gustavo Stolovitzky, Stanislav Polonsky, and Qinghuang Lin. Wafer-scale integration of sacrificial nanofluidic chips for detecting and manipulating single DNA molecules. *Nature Communications*, 8:14243, 2017. doi:10.1038/ncomms14243.

[Wor24] World Economic Forum. Health spending takes up 10% of the global economy: How can tech help reduce costs and improve lives?, Aug. 2024. URL: https://www.weforum.org/stories/2024/08/healthcare-costs-digital-tech/.

[YZR+09] Y.-J. Yu, Y. Zhao, S. Ryu, L. E. Brus, K. S. Kim, and P. Kim. Tuning the Graphene Work Function by Electric Field Effect. *Nano Letters*, 9(10):3430–3434, 2009. URL: https://arxiv.org/abs/0909.0020, doi:10.1021/nl901572a.

[Z+08] J. Zhang et al. Single-Molecule Electron Transfer in Electrochemical Environments. *Chemical Reviews*, 108(7):2737–2791, 2008. URL: https://pubs.acs.org/doi/full/10.1021/cr068073+, doi:10.1021/cr068073+.

[ZGN07] J. Zhou, E. Garand, and D. M. Neumark. Vibronic structure in C2H and C2D from anion slow electron velocity-map imaging spectroscopy. *Journal of Chemical Physics*, 127(11):114313, 2007. URL: https://webbook.nist.gov/cgi/cbook.cgi?Source=2007ZHO%2FGAR114313, doi:10.1063/1.2768932.

[Zho21] Q. Zhong. Constructing covalent organic nanoarchitectures molecule by molecule via scanning probe manipulation. *Nature Chemistry*, 13:1133–1139, 2021. URL: https://www.nature.com/articles/s41557-021-00773-4.pdf, doi:10.1038/s41557-021-00773-4.

16 Glossary

6DOF: Six Degrees Of Freedom. The ability to move in three translational (X, Y, Z) and three rotational (roll, pitch, yaw) axes.

Activation barrier; transition state: An activation barrier is the energy rise along a reaction pathway from the reactant configuration to the intervening transition state. The transition state is the unstable configuration at the barrier separating reactants from products. Barrier height affects reaction rates and thermal reliability; it is distinct from the energy difference between the reaction endpoints.

Adamantane: A diamondoid hydrocarbon molecule, used as a structural motif in molecular design.

Adsorption; desorption: Adsorption is the binding or accumulation of atoms or molecules at a surface; desorption is their release from it. These processes occur in the proposed platinum-mediated hydrogen chemistry. Adsorption differs from absorption into the bulk of a material.

Airlock: A transfer chamber or mechanism that moves material between environments while limiting their mixing. Proposed SRU alternatives use airlocks to move blocks or waste across the vacuum/solution boundary. An airlock is distinct from the baseline sliding seal around an extruding product.

AM1 (Austin Model 1): A semiempirical quantum-chemical method that uses fitted parameters to approximate molecular electronic structure. It is used for several molecular models in this paper. Its relatively low computational cost is useful for exploratory modeling, but its accuracy depends on the system and property being calculated.

AP Block: Atomically Precise Block. A small, rigid, atomically precise diamondoid building block made by mechanosynthesis. AP Blocks can be assembled by either the fabricator or the internal assembly arm. Assemblies made within the fabricator are limited in size. Used as the fundamental construction unit for SRUs, fabricators, and other products. The SRU is an assembly of AP Blocks. See also Bootstrap AP Blocks.

Assembly; assembly operation: An assembly is a structure made by joining two or more blocks or assemblies. An assembly operation joins or positions these parts rather than synthesizing a block through a single reaction sequence. In this paper, fusing blocks into one covalently bonded network still produces an assembly; the term depends on manufacturing history, not just final structure.

Atomic precision; atomically precise manufacturing (APM): Atomic precision specifies the identities and arrangement of individual atoms in a product. APM is manufacturing with this control. Mechanosynthesis is the particular synthesis method proposed here; atomic precision does not by itself specify a manufacturing method.

Attohectare (aha): A unit of area equal to that of a square 100 nm on a side, used for expressing nanoscale surface areas.

AzAdam: A molecular-tool family whose active tip is based on 1-azaadamantane, with nitrogen at an adamantane bridgehead. AzAdam denotes the neutral tip, AzAdam+ its positively charged form, and AzAdam-H+ a hydrogen-bearing positively charged form. Electron addition produces AzAdam-H, which can release hydrogen in the proposed reaction sequences. AzAdam chemistry is discussed as a hydrogen-generation risk mitigation.

B3LYP; basis set: B3LYP is a hybrid density functional combining Becke's three-parameter formulation with Lee–Yang–Parr correlation. A basis set is the set of mathematical functions

used to represent the electronic orbitals in a calculation. Notation such as B3LYP/6-311+G(2d,p) specifies both a functional and a basis set; neither is an experimental measurement.

Baseline: The set of design assumptions and choices for the 2026 nanofactory proposed in this paper. This is an informal design description; see Section 11.

Baseline palette: the elements H, B, C, N, O, Si, Ge, Pt.

Binary Bowden cable: A Bowden cable that transmits a binary (two-position) mechanical signal.

Binding site: A molecular structure that selectively binds an intended molecule or part. Feedstock binding sites capture molecules from the surrounding solution for transfer into the SRU. Their selectivity, release behavior, and resistance to leakage are design requirements. See also Half-site closure test and Half-site opening test.

Blockify: Decompose a design into AP Blocks: convert a continuum design or an atomically precise design that is not divided into AP Blocks into a similar design which maintains core functionality, but which is composed entirely of AP Blocks. This splits the conceptual design space into designs which entirely ignore manufacturing limitations and designs which can be made by assembling blocks from a fabricator. The hope is that such a two-phase design process, which first creates APM designs without the AP Block restriction and then blockifies those designs, will be easier than creating blockified designs de novo.

Bond dissociation energy (BDE): The energy required to break a specified bond by homolytic cleavage, giving each fragment one of the bonding electrons. Bond strengths help compare hydrogen abstraction and donation reactions. The quoted source determines the temperature and whether the tabulated quantity includes thermal enthalpy corrections.

Bootstrap AP Blocks: Blocks from which the workstation fabricator is assembled. Manufacturable directly by the workstation without using the workstation fabricator. Necessary to bootstrap the workstation fabricator. Elements in bootstrap AP Blocks are expected to be a proper subset of the baseline palette. The exact set of elements in bootstrap AP Blocks is not specified by this paper.

Bootstrapping: Building the first working system without relying on an already-existing copy. Here, a workstation first makes and assembles bootstrap AP Blocks into a workstation fabricator; that fabricator makes the AP Blocks for the first SRU, which the workstation assembles. Subsequent SRUs are made by replication.

Bowden cable: A mechanical cable system consisting of an inner cable that slides within an outer housing, used to transmit force or motion; in the SRU, used for transmitting binary mechanical signals.

Bridgehead: An atom shared by rings at a junction in a bridged ring structure. Adamantane has bridgehead sites used in the tool designs discussed here. Replacing a carbon at such a site with germanium or nitrogen helps define the active-tip structures of GeRad or AzAdam tools.

Broadcast: Instructions transmitted to all SRUs by the central computer, typically through Select (S1 and S2) and Drive (D1 and D2). All connected SRUs receive the transmission; execution depends on their ON/OFF state and the instruction.

Buckling; Euler buckling: Buckling is a loss of structural stability under compression, often causing a slender member to bend laterally. Euler buckling is the ideal elastic model used here to estimate a strut's critical compressive load from its stiffness, cross-section, length, and end constraints. Buckling is distinct from fracture or tensile failure.

Buckyball: A spherical molecule composed entirely of carbon atoms (fullerene), used for hydrogen storage or as a structural motif. The most common buckyball is C_{60}, which resembles a soccer ball.

Charge pump: An electronic circuit that uses capacitors to increase (boost) or invert voltage.

Closed-loop mechanosynthesis; feedback: Mechanosynthesis in which measurements or tests affect subsequent tool motion or execution. Feedback is the use of an observed result to adjust an action or control decision. This contrasts with carrying out prescribed motions without checking their outcome; the paper proposes simple binary feedback rather than requiring elaborate local computation.

Coefficient of performance (COP): For a refrigerator, the heat removed from the cold region divided by the work supplied to the cooling system. COP can exceed one because the system moves heat rather than converting all of it from work. The Carnot COP is an ideal upper limit for operation between specified temperatures.

Commensurate; incommensurate surfaces: Commensurate contacting lattices have matching periodicities with a common repeat pattern. Incommensurate lattices lack such a common repeat. Finite contacts can also behave as effectively incommensurate when no common repeat fits within the contact region. Mismatched periodicities or orientations can reduce coherent resistance to sliding, but finite size, deformation, and defects also affect friction.

Contact pillar: A raised feature on the die that provides an electrical connection to an SRU. In the baseline, spikes extended from three SRU feet contact six pillars carrying Ground, V+, S1, S2, D1, and D2. Pillar geometry is also relevant to locating and attaching the SRU.

Convergent assembly: A hierarchical manufacturing strategy where small parts are assembled into larger parts, which are then assembled into even larger structures, with testing at each stage to ensure reliability.

Coronene: $C_{24}H_{12}$, a flat molecule that looks like seven fused benzene rings, one ring in the center surrounded by six others.

Covalent bond: A chemical bond involving shared electron density between atoms. Covalent networks provide the structure of diamondoid blocks. Covalently fusing blocks creates chemical bonds across their interface, unlike attaching them through noncovalent interactions such as van der Waals forces.

CVD (Chemical Vapor Deposition): A process for depositing thin films of material from a vapor phase, used in diamond and nanomaterial synthesis.

Daughter (SRU): An SRU that is manufactured and extruded by a parent SRU during replication.

Demux (Demultiplexer): A device that takes a small number of inputs and generates many outputs. In the SRU, the demux might have four inputs and generate over 100 outputs.

Density functional theory (DFT): An electronic-structure method based on electron density, used to calculate molecular energies, structures, and related properties. Practical calculations use approximate functionals. Their predictions depend on the chosen approximation and numerical representation, and are computational results rather than experimental verification.

Diamondoid: As used in this paper: strong, stiff, atomically precise structures containing dense networks of covalent bonds, including graphene structures and carbon nanotubes. Baseline AP Blocks are diamondoid and are made from the baseline palette, with Pt included by the terminology convention described in Section 3.5.2. Additional elements are included in other contexts.

Die: A flat, typically silicon, substrate that provides sites, power, and control signals for SRUs.

Dielectric; dielectric constant; dielectric strength: A dielectric is an electrically insulating medium that can be polarized by an electric field; vacuum is the reference medium in the alternative comb-drive design. Dielectric constant is relative permittivity, which affects capacitance. Dielectric

strength is the electric field a material can withstand before electrical breakdown. A large dielectric constant does not itself imply a large dielectric strength.

Differential comb drive: An electrostatic actuator consisting of two opposing comb actuators acting on one shared shuttle. It converts binary electrical signals into binary mechanical motion and is used to drive binary Bowden cables. Four differential comb drives are installed in the baseline SRU; an active pair consists of two complete differential comb drives selected for clocking, with one drive changing state during each clock phase.

DOF: Degrees Of Freedom. A term often used to describe positional devices. A 3DOF device might provide X, Y, and Z positioning. Alternatively, a 3DOF device might provide roll, pitch and yaw positioning. A 6DOF device, such as a Stewart platform, positions a physical object in all six spatial degrees of freedom. A 1DOF device might position a slider along a straight line.

EAA: External Assembly Arm. An external assembly arm is a compound Stewart platform mounted on an SRU. It removes a daughter SRU, a newly manufactured daughter EAA, or another product from the ES after manufacture. It transfers SRUs or other products across the die surface using ciliary transport. It positions the daughter SRU when it is to be attached to the die surface.

EBL (E-beam lithography): Electron beam lithography, a high-resolution technique for patterning nanoscale features on wafers.

Electromechanosynthesis: Mechanosynthesis involving controlled electron transfer or the use of charged molecular groups in the reaction. See EMS, the corresponding abbreviation used in this paper.

EMS: Electromechanosynthesis; mechanosynthesis involving controlled electron transfer or the use of charged molecular groups in the reaction.

Energy-delay product (EDP): The energy dissipated by an operation multiplied by the time required for it. It combines energy use and speed in one performance metric. Comparisons require compatible definitions of the operation, energy, and delay; EDP is not the total power consumption of a computer.

Equipartition: The classical thermal-equilibrium principle that each independent quadratic contribution to energy has a mean energy of $k_BT/2$. Applied to a harmonic spring, it gives $k_s\langle x^2\rangle/2 = k_BT/2$, which underlies the thermal-displacement estimates here. Quantum effects can limit this classical approximation.

ES: Extrusion sleeve. After an SRU has been manufactured and extruded by its parent, it manufactures and extends its extrusion sleeve. The sleeve has a 60 nm × 100 nm cross section which seals against the daughter SRU as it is being extruded and prevents flooding of the parent SRU even if long irregularities or gaps are present in the shell of the daughter SRU. The External Assembly Arm (EAA) has a 20 nm long base with a 60 nm × 100 nm cross section that prevents flooding of the parent SRU during its manufacture.

EtCpPtMe$_3$: (Trimethyl)ethylcyclopentadienylplatinum(IV), the proposed platinum-containing baseline feedstock molecule. EtCp denotes an ethylcyclopentadienyl group and Me denotes a methyl group. Platinum recovered from this feedstock is intended for the hydrogen-generation catalyst.

Eutactic: "well arranged," from the Greek eus, "good," and taktikos, "of order or arrangement". "Characterized by precise molecular order, like that of a perfect crystal, the interior of a protein molecule, or a machine-phase system; contrasted to the disorder of bulk materials, solution environments, or biological structures on a cellular scale. Borderline cases can be identified, but perfection is not necessary. As a crystal with sparse defects is best described as a crystal (rather than as

amorphous), so a eutactic structure with sparse defects is best described as (imperfectly) eutactic, rather than as disordered." [Dre92]

Exoergic; endoergic; exoergicity: As used in the reaction-energy discussions here, an exoergic reaction has products lower in energy than the reactants; an endoergic reaction has products higher in energy. Exoergicity describes the energy decrease. The quoted calculation determines whether the quantity is an electronic energy, enthalpy, or free-energy difference; these are not interchangeable. An energetically favorable reaction can still have an activation barrier.

Fabricator: A subsystem within either the SRU or workstation that carries out mechanosynthesis to build AP Blocks. See Workstation fabricator and SRU Fabricator. This term can be used generically to refer to either, or, depending on context, can refer specifically to one or the other.

Fail-stop: An error-handling strategy where a device ceases operation upon detecting an unrecoverable error, preventing propagation of faults [SS83].

Feedstock: The supply of molecules (often in solution) that provides the raw elements required for mechanosynthesis.

Four-phase clocking: A repeating sequence of four drive phases that coordinates movement, engagement, and release in a mechanical mechanism. Properly phased motions of its elements produce successive steps of translation or rotation. The specific phase sequence depends on the mechanism; four phases do not imply four independently addressed control wires. Two Bowden cables can carry the four phases used in four-phase clocking.

Fullerene: A closed-cage carbon molecule, including roughly spherical cages such as C_{60} and elongated cages. Buckyballs are the roughly spherical members of this family. Fullerene-derived hemispherical structures are also used as nanotube caps in the proposed designs.

GeRad: A molecular tool based on adamantane with a bridgehead germanium radical, used in mechanosynthesis for specific chemical reactions.

GeRad-CH2: A molecular tool used in specific mechanosynthetic reaction sequences, often to deposit a CH_2 group onto a radical site on the workpiece.

GeRad-H: A molecular tool used to donate hydrogen atoms during mechanosynthesis. It is the GeRad tool with a hydrogen bonded to its germanium atom at an adamantane bridgehead.

GM tool: older name for GeRad-CH2

Graphene: A single layer of carbon atoms arranged in a hexagonal network. Graphene is used here in proposed structural blocks, conductive surfaces, and sliding interfaces. A carbon nanotube can be described geometrically as a rolled graphene sheet. Graphene and h-BN have similar planar lattice structures but different electronic properties.

Gripper; assembly tool: A gripper is the part of an assembly tool that holds an intended block or assembly. An assembly tool has a handle for attachment to the general tool holder and a working end suited to the part being manipulated. The gripper, tool holder, and handle perform different roles.

Grip point: There are two grip points on the exterior of the SRU. Each grip point can be grasped by an EAA tool holder. The grip point has the same general design as a handle, in that an EAA tool holder grasps the grip point by using 6DOF motions. As with a tool handle, the EAA tool holder can release the grip point by using 6DOF motions.

Ground (0 volts): The reference voltage in the SRU electrical system.

HAbst (Hydrogen Abstraction Tool): A highly reactive molecular tool used to remove hydrogen atoms from surfaces during mechanosynthesis. The AdamRad-CC hydrogen-abstraction tool

described under molecular-tool terminology is also called AdamRad-C2.

Half-site closure test: A method of testing an individual molecule or atomically precise structure by sandwiching it between two "half-sites", where each half-site is half of a binding site. If the two half-sites fail to close, the molecule fails the test.

Half-site opening test: A method of testing an individual molecule or atomically precise structure by sandwiching it between two "half-sites", where each half-site is half of a binding site. If the two half-sites close, the force required to open the two half-sites can be measured. If the opening force is outside the acceptable range, the half-site fails the opening test.

Handle: The handle is the tool's attachment feature, through which the tool holder positions, orients, and retains the tool. All molecular tools share a common handle design, allowing any molecular tool to be picked up or put down by any tool holder. In the baseline proposal, all positional devices are 6DOF and position a tool holder that can both attach to and release a handle purely by using motions of the positional device.

Hand-over-hand transfer; ciliary transport: Cooperative transport in which successive external assembly arms grasp, move, and release a daughter SRU or other product across the die. "Ciliary transport" is used here for this coordinated arm-based transport. It allows a daughter to reach an available site beyond its parent's immediate reach.

Harmonic drive: A gear mechanism that uses a flexible spline to achieve high reduction ratios.

h-BN (hexagonal boron nitride): A two-dimensional material similar to graphene, used as a structural element and as the dielectric in the baseline differential electrostatic comb drives.

HDon: older name for GeRad-H.

H-tree: A recursively branching wiring layout that distributes signals to many destinations along geometrically balanced paths. H-trees distribute the baseline SRU control signals across a die. Wiring resistance, capacitance, propagation delay, and the connected loads determine their electrical performance.

Hysteresis; backlash: Hysteresis is dependence of a system's response on its previous history. Backlash is lost motion due to clearance or slack, especially when reversing a transmission. Both can make actual position differ from the position inferred from commands; they are distinct from random thermal fluctuations.

Internal assembly arm: A Stewart-platform subsystem inside the SRU that uses assembly tools to manipulate and join AP Blocks supplied by the fabricator. It can reach the fabricator output and the base of the daughter or other product under construction. Unlike the external assembly arm, it works within the SRU's vacuum environment.

LEGO bricks: a standard LEGO brick is an interlocking plastic construction element featuring top studs and hollow bottom tubes (anti-studs) that lock securely together. The most common is the 2×4 LEGO brick. LEGO blocks are not LEGO bricks. LEGO blocks are atomically precise with sizes measured in nanometers.

LM2-2: A specific diamond surface structure on a (110) surface requiring high positional accuracy during mechanosynthesis. See Section 3.5.4.

Lonsdaleite: Hexagonal diamond, a crystalline form of carbon.

Mechanosynthesis: The synthesis of complex atomically precise products using positionally controlled molecular tools, particularly tools with radical tips, in an inert environment such as vacuum. This paper focuses primarily on diamondoid products.

MEMS (Microelectromechanical systems): Small manufactured devices combining mechanical structures with electrical sensing or actuation, commonly made using semiconductor fabrication methods. Conventional MEMS comb drives provide a geometric and operating comparison for the atomically precise comb-drive proposals here.

Metabolism (manufacturing): The network of reactions by which the manufacturing system obtains useful atoms from feedstock, builds products, reloads molecular tools, and disposes of waste. This use borrows the biological term but refers to engineered manufacturing chemistry, not biological metabolism.

mFET (molecular Field Effect Transistor): An atomically precise transistor [MFA25].

Molecular LEGO block; LEGO block: An AP Block inspired by LEGO bricks. For example, AP Blocks that, like standard 2×4 LEGO bricks, attach to the top and bottom of other similar AP Blocks, enabling the formation of walls and other common LEGO structures. LEGO bricks are different from LEGO blocks. LEGO bricks are macroscopic and made from plastic. LEGO blocks are atomically precise with sizes measured in nanometers.

Molecular Tinkertoy: AP Blocks inspired by Tinkertoys, in which "rods" and "spools" are used to assemble (rather porous) frameworks. The rods insert into holes in the spools. In the constructions used here, molecular Tinkertoys use rods made from nanotubes and spools made from fullerenes.

Molecular tool; tooltip: A molecular tool is a structure positioned by a tool holder to perform a reaction or manipulation. Its tooltip is the local molecular structure at the working end, particularly the reactive tip of a mechanosynthetic tool. Tool names in this paper specify the tooltip structure; more distant support and handle structures can vary.

Nanofactory: A programmable manufacturing system that makes atomically precise products. The baseline nanofactory combines many SRUs with a conventional support environment that supplies feedstock, power, control, and cooling. The SRUs manufacture parts and assemble products, including further SRUs.

Nanotube chirality; armchair and zigzag: The way a hexagonal sheet is wrapped to form a nanotube, specified by two integers (n, m) describing its wrapping vector. Armchair nanotubes have $n = m$; zigzag nanotubes have one index zero in the usual convention. The indices affect tube diameter and electronic properties. Armchair and zigzag also describe edge orientations in planar graphene or h-BN sheets.

Narrowcast: In contrast to the more usual broadcast, in which instructions are intended for most, if not all, SRUs, narrowcast instructions are instructions transmitted by the central computer which are intended for a small number of SRUs, possibly even a single SRU. Only the intended recipient SRUs are turned ON; all other SRUs are turned OFF. The small number of SRUs turned ON might all be localized to a small region on the surface of the die, or might otherwise be a small number of SRUs that share some common and unusual property. Narrowcast is expected to be essential for debugging, system setup, and handling of edge cases.

nN/aha (nanonewton per attohectare): A unit of pressure equal to 10^5 Pa. It is approximately equal to one atmosphere.

PAC (Product Assembly Chamber): A vacuum-tight chamber used for assembling products larger than an SRU.

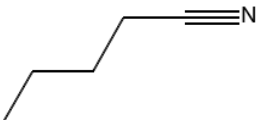
Parent (SRU): An SRU that manufactures and extrudes a daughter SRU.

Parts closure: The ability of a manufacturing system to make all the parts needed to construct another instance of itself. For the proposed SRU, this requires manufacturing its AP Blocks and

assembling them into another SRU. It does not imply that the SRU also manufactures the conventional die, power supply, or other external support equipment.

Passivation; depassivation: Passivation stabilizes a reactive surface site, for example by terminating a dangling bond with hydrogen. Depassivation removes that termination to expose a site for a subsequent reaction. Hydrogen abstraction and donation can therefore create and then stabilize reaction sites during mechanosynthesis.

PDB (Protein Data Bank format): A standard file format for representing three-dimensional structures of molecules.

Pentanenitrile: C_5H_9N. 

One component of the feedstock.

Potential energy surface (PES): The potential energy of a molecular system as a function of its atomic configuration. Minima correspond to locally stable configurations, while barriers separate them. Reaction paths, screw rotation, and sliding interfaces can be analyzed by considering their paths across a PES.

Programmable Fabricator: see SRU Fabricator.

Quantum capacitance: The capacitance associated with the change in a conductor's electronic chemical potential as charge is added or removed. It depends on the available electronic states and can alter how an applied voltage divides across graphene electrodes and a vacuum gap, in addition to the ordinary geometric capacitance.

Radical; dangling bond: A radical is an atom, molecule, or ion with an unpaired electron. A dangling bond is an unsatisfied bonding site, often represented as a reactive radical site in the molecular structures discussed here. Radicals can be neutral or charged: radical character and electrical charge are different properties.

Ratchet and pawl: A mechanical device that allows motion in one direction only. In a possible SRU application, a reciprocating pawl advances a ratchet wheel to rotate a shaft.

Reaction sequence: An ordered set of chemical operations, including positioning, donation, abstraction, and tool reloading, used to synthesize a structure. The paper calls a structure made through a single reaction sequence a block; joining previously made blocks produces an assembly.

RET: Resonant Energy Transfer. A method of charging and discharging wires repeatedly without dissipating the usual CV^2 energy per cycle.

RMS positional uncertainty; thermal noise: RMS (root-mean-square) positional uncertainty measures the typical displacement from a mean or intended position through the square root of the mean squared displacement. Thermal noise is fluctuation caused by thermal motion. For a classical harmonic positional mode with stiffness k_s at temperature T, the thermal RMS displacement is $\sqrt{k_B T/k_s}$. It is not a worst-case bound on positioning error.

S1, S2, D1, D2 (Control/data lines): Electrical or mechanical signals used to drive the demux. These signals are used to select (S) and drive (D) every degree of freedom in the SRU. A single selected degree of freedom might be stepping a rotary stepper backwards and forwards to control the extension or contraction of one strut of a Stewart platform.

Screw drive: A mechanism that converts relative rotation of a threaded rod and sleeve or nut into axial displacement. Its thread pitch relates rotation to linear travel. The baseline uses screw

drives to adjust Stewart-platform strut lengths with fine positional control.

Self-assembly: Formation of an organized structure through interactions among its components, without separately positioning each component through the assembly steps. This differs from the positionally controlled assembly performed by the proposed SRU assembly arms.

Shell: The barrier that separates the internal (eutactic) environment of the SRU from the external (non-eutactic, conventionally manufactured) environment. The shell itself is considered eutactic.

SIMD (Single Instruction stream Multiple Data stream): A parallel computing architecture where multiple processing elements execute the same instruction on different data. The general architectural approach used for the broadcast architecture used in the baseline.

Sliding seal: A seal between surfaces that move relative to one another while restricting fluid leakage. During product extrusion, the seal between the extrusion sleeve and the emerging product prevents the external feedstock solution from flooding the parent SRU. Its required leakage performance follows from the SRU reliability target.

SNOT (Studs Not On Top): A LEGO term for bricks that add studs on the sides rather than just the top. Used to describe molecular LEGO blocks that can be assembled in a manner similar to SNOT bricks.

Space frame: A lightweight, rigid, three-dimensional structure composed of struts and nodes, used for internal support in SRUs or larger assemblies.

Split ring compound planetary gear: A gear mechanism for achieving high reduction ratios in a compact, rigid structure.

SPM (Scanning Probe Microscope): A device capable of imaging surfaces at the atomic scale and manipulating individual atoms or molecules.

SRU: Simple Replicating Unit. An SRU might be 500 nm × 100 nm × 60 nm, mounted on a die surface. It manufactures and extrudes product objects through a fluid-tight Extrusion sleeve (ES). It receives instructions electronically from the contact pillars on the die surface. The die is immersed in the feedstock solution, which supplies the SRU with the raw materials for manufacturing product objects. Product objects include SRUs, EAAs, and other atomically precise structures that measure 500 nm × 100 nm × 60 nm. Because SRUs extrude product objects, the 500 nm dimension could be longer, limited primarily by the reliability of the mechanosynthetic manufacturing and extrusion process.

SRU Fabricator: The fabricator internal to the SRU. It has four Stewart platforms in a volume of 100 nm × 100 nm × 60 nm.

Stewart platform: A six-degree-of-freedom (6DOF) positional device consisting of a base, a platform, and six adjustable struts, used for precise positioning in SRU fabricators and assembly arms.

Stiffness; compliance; Young's modulus: Stiffness relates force to displacement; axial stiffness k_s is F divided by displacement in a linear spring model. Compliance is displacement per unit force, the reciprocal of stiffness for that degree of freedom. Young's modulus E is the material's tensile stress divided by tensile strain in the linear elastic regime. A structure's stiffness depends on geometry and interfaces as well as material modulus.

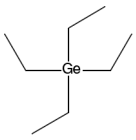
STM (Scanning Tunneling Microscope): A specific type of SPM which measures the tunneling current between the atom at the end of a sharp probe and a surface being scanned. The object being imaged and the probe must both be conductive. An STM is capable of atomic resolution and of manipulating individual atoms or molecules.

Superlubricity: A regime of exceptionally low sliding friction, often associated with a mismatch between the periodic structures of contacting surfaces. The paper considers rotationally misaligned graphene and h-BN interfaces as candidates for reducing sliding losses. The term does not imply exactly zero friction or guarantee negligible wear.

SWNT (single-walled nanotube): A cylindrical nanostructure made of a single layer of carbon atoms, used for electrical conduction.

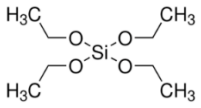
TEOS: Tetraethyl orthosilicate, one of the baseline feedstock molecules. See Tetraethyl orthosilicate.

Test Unit: A binary device with two stable states (TRUE/FALSE), used for conditional logic and error checking in SRUs. The test unit is positioned next to the condition to be tested in such a manner that it switches an internal mechanical bit in response to the condition. The test unit is then loaded into the test station (see Section 6.1.8), where the result of the test is used to conditionally turn the SRU OFF.



Tetraethylgermane: $Ge(C_2H_5)_4$.

One component of the feedstock.



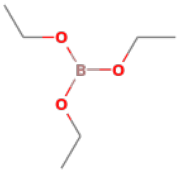
Tetraethyl orthosilicate: $Si(OC_2H_5)_4$.

One component of the feedstock solution.

Throw: The specified travel of an actuator between its operating positions. In the comb-drive calculations, d_t denotes shuttle throw. The baseline differential comb drive has a 24 nm shuttle throw. A 12:1 transmission reduces this to a 2 nm output throw at the binary Bowden cable. Throw is a displacement, distinct from the actuator's overall length, generated force, or number of operations in a clock cycle.

Tinkertoy: a commercially available construction set made of "rods" and "spools". The rods insert into holes in the spools. This construction style encompasses a space frame, which is made from struts and nodes.

Tool holder: The tool holder can pick up or put down any molecular tool by grasping its handle, which it can do purely by 6DOF motions. The tool holder in the baseline proposal is passive, not active. It does not require a separate control line to operate it. The handle snaps into place by overcoming a barrier that prevents it from being released by thermal noise or small mechanical perturbations. Deliberate application of sufficient force along a particular degree of freedom is used to unsnap (release, put down) the handle and the molecular tool to which it is attached. Many variations are possible.



Triethyl borate: $B(OC_2H_5)_3$.

One component of the feedstock.

TSV: TSV stands for Through-Silicon Via. A TSV is essentially a vertical electrical connection drilled (or etched) through a silicon die and then filled with a conductor, usually copper. Instead

of signals traveling horizontally across a chip, a TSV allows them to travel vertically from one die to another.

Turnbuckle: A mechanical device used to adjust the length or tension of cables or rods.

Ultra-high vacuum (UHV): A vacuum regime with extremely low residual gas pressure, used in surface science and atomic-scale manipulation. Low pressure reduces collisions and contamination from unwanted molecules. UHV describes an environmental pressure regime; "vacuum tight" in this paper instead describes acceptable leakage performance.

Umbilical: A multi-conductor cable (possibly made from nanotubes) that provides power and control signals to an SRU before it is attached to the die.

Vacuum tight: The leakage rate is low enough that leakage-related failures consume only part of the total 0.1% failure budget per SRU per replication cycle.

Van der Waals forces (VDW): Noncovalent interactions between atoms or molecules arising from permanent or induced electric multipoles, including dispersion. They provide adhesion between many of the proposed surfaces and parts without forming a continuous covalent network. At short separations, repulsion also limits how closely surfaces can approach.

V+ (Positive voltage): The positive DC voltage supplied to the SRU.

Waste: Waste molecules are vented from the SRU into the surrounding feedstock. Waste molecules are H_2, N_2, C_2H_2, CO_2, and tetramethylsilane.

Wire: a lithographically patterned metal interconnect in an integrated circuit, or an atomically precise interconnect such as a metallic nanotube.

WKB approximation: The Wentzel–Kramers–Brillouin semiclassical approximation for a quantum wavefunction in a varying potential. It is used here to estimate transmission through an electrical tunneling barrier and the strong dependence of leakage on barrier height and width [Sim63].

Workstation: An advanced scanning probe microscope (SPM) with enhancements to facilitate mechanosynthesis and assembly of AP Blocks.

Workstation fabricator: A fabricator subsystem within the workstation, assembled from bootstrap AP Blocks, and used to manufacture the blocks needed by the first SRU.

Worm gear: A transmission in which a screw-like shaft, the worm, meshes with a toothed wheel. It provides speed reduction and torque amplification, and can resist back-driving depending on its geometry and friction. In the baseline it supplies rotation to a screw drive; the screw drive then produces linear motion.

Wurtzite boron nitride (w-BN): A crystalline form of boron nitride with a hexagonal, tetrahedrally bonded wurtzite structure, distinct from layered hexagonal boron nitride (h-BN).

17 Symbols, Abbreviations, and Units

Abbreviations are defined in the Glossary (Section 16).

17.1 Principal symbols

Table 21 collects the principal symbols used in the calculations. Symbols are grouped by application; a symbol used in more than one context has separate entries where its meaning or units differ. Units are given in SI form. Calculations often use prefixed units such as nm, pm, nN, zJ, aJ, and aF; these prefixes are listed in Section 17.2. A unit of 1 denotes a dimensionless quantity, including a count or probability.

The table defines quantities rather than prescribing numerical values. Baseline h-BN and vacuum comb-drive parameter values are given in Tables 13 and 19; physical and material constants are listed in Sections 17.4 and 17.5. Local definitions in the accompanying calculation specify the applicable geometry, model, and parameter choices.

Table 21: Principal symbols, units, and context.

Symbol	Meaning	Units	Scope or qualification
Mechanical geometry, stiffness, and motion			
D_o, D_i	Outer and inner tube diameters	m	Tube cross-section geometry.
D_m, r_m	Mean tube diameter and radius	m	$D_m = (D_o + D_i)/2$; $r_m = D_m/2$.
D_1, D_2	Sleeve and rod outer diameters	m	Two-piece threaded strut; neither is a mean diameter.
w_t	Tube wall thickness	m	Sleeve and rod model.
L, L_1, L_2	Strut lengths	m	L_1: length of each sleeve and rod; L_2: exposed rod length; $L = L_1 + L_2$.
A	Load-bearing cross-sectional area	m^2	Axial stiffness and tensile strength; distinct from comb electrode area.
I	Second moment of area	m^4	Bending and Euler buckling, Eq. (11).
E, G	Young's and shear moduli	Pa	Axial/bending and torsional elasticity, respectively.
k_s	Spring stiffness	N/m	Axial strut: $k_s = EA/L$, Eq. (27); also used for molecular springs.
σ_x, σ_z	RMS positional uncertainty	m	Thermal axial uncertainty; screw-induced axial uncertainty, respectively.
σ_{tensile}	Allowable tensile stress	Pa	Material strength; not positional uncertainty.
P_{cr}	Critical buckling force	N	Euler column model; not power or pressure.

Continued on next page

Table 21 (continued)

Symbol	Meaning	Units	Scope or qualification
K	Effective-length factor	1	Euler buckling; $K = 1$ for pinned ends.
F	Force	N	Subscripts identify applied, drag, or bond forces.
m	Moving mass	kg	Shuttle motion; distinct from electron mass in tunneling.
d, t	Travel distance and elapsed time	m; s	Constant-force shuttle motion; tunneling uses a different meaning of d.
v, ω	Sliding and angular speeds	m/s; rad/s	Linear sliding and rotary drag, respectively.
J	Polar second moment of area	m^4	Screw/rod torsional elasticity; distinct from current density.
p	Screw lead	m	Axial advance per revolution; distinct from probability.
Comb drives and electrical quantities			
N	Inside-comb tine count	1	Outside comb has $N + 1$ tines; one comb actuator has $2N$ side gaps.
h_c, d_t	Comb height and actuator throw	m	Lengths in the comb geometry.
S_S, SS	Side spacing between tines	m	h-BN and vacuum design notation, respectively.
TW	Tine width	m	Vacuum design; includes the nanotube sheath.
$A_{\mathrm{contracted}}$, A_{extended}	Electrode overlap areas	m^2	Area of one comb actuator in each end position.
ΔA_{comb}	Change in overlap area	m^2	One comb actuator; $\Delta A_{\mathrm{comb}} = 2Nh_cd_t$.
A_{drive}	Electrode area per differential drive	m^2	Sum of contracted and extended areas of its two opposing actuators.
A_{active}, $A_{\mathrm{installed}}$	Areas of selected pair and installed set	m^2	$2A_{\mathrm{drive}}$ and $4A_{\mathrm{drive}}$; active does not mean simultaneous switching.
V	Applied voltage	V	Electrical calculations; not physical volume.
V_{comb}, V_{drive}	Physical volumes	m^3	One comb actuator and one differential drive, respectively.
C, R, τ	Capacitance, resistance, RC time constant	F; Ω; s	$\tau = RC$.
R_K	von Klitzing resistance	Ω	$R_K = h/e^2$; molecular-wire resistance estimate.
ϵ_r	Relative permittivity	1	Material dielectric constant; vacuum has $\epsilon_r = 1$.

Continued on next page

Table 21 (continued)

Symbol	Meaning	Units	Scope or qualification
P, P_{drag}	Power and drag power	W	Electrical losses and mechanical dissipation; not pressure.
W	Work or energy	J	Energy delivered per actuation; not the unit symbol W for watt.
$E_{\max}$	Maximum energy per driver actuation	J	Energy budget for one of the four mechanical drivers.
S	Sliding contact area	m^2	Sliding drag, Eq. (70).
K_{drag}	Sliding drag coefficient	$\text{kg}/(\text{m}^2\,\text{s})$	$P_{\text{drag}} = K_{\text{drag}} S v^2$.
K	Rotary drag coefficient	$\text{kg}\,\text{m}^2/\text{s}$	$P = K\omega^2$, Eq. (83); radians are dimensionless in SI.
$S(k), S_k$	H-tree square side length	m	k is the H-tree level; not sliding area.
$L_{\text{wire},0:13}$	Total H-tree wire length	m	Sum over levels 0 through 13.
Tunneling and thermal reliability			
J, j_{leak}	Current density	A/m^2	Vacuum tunneling and h-BN leakage estimates, respectively.
I, I_{leak}	Current and leakage current	A	Electrical use of I; not second moment of area.
d	Electrode separation	m	Simmons vacuum tunneling model; not shuttle travel.
m	Free-electron mass	kg	Simmons model; not shuttle mass.
Φ, W_G	Barrier energy and graphene work function	J	Vacuum leakage estimate; eV values are also supplied.
κ	Evanescent decay coefficient	m^{-1}	$\kappa = \sqrt{2m\Phi}/\hbar$.
T	Tunneling transmission probability	1	$T \sim \exp(-2\kappa d)$; not temperature.
T	Absolute temperature	K	Thermal fluctuations, activated events, and cooling.
E_a	Activation energy barrier	J	Wear estimate; distinct from Young's modulus E.
ν_0	Attempt frequency	s^{-1}	Thermally activated wear events; not angular speed.
N_s, p	Susceptible site count and failure probability	1	Wear model: $p \approx N_s \nu_0 t \exp[-E_a/(k_B T)]$.
Replication and filtering			
p_{BL}, M	Block attempt failure probability and block count	1	Retry model; M counts blocks required per replication.
P_{success}	Replication success probability	1	Block retry model, Eq. (2).

Continued on next page

Table 21 (continued)

Symbol	Meaning	Units	Scope or qualification
q, g	Branching success probability and generation index	1	All-or-none model: expected population $(2q)^g$.
N, S_N	Generation count and total step count	1	Die-filling model; not tine count or sliding area.
N, n, p	Contaminant count, filter-stage count, per-stage false acceptance probability	1	Independent-stage filter model; failure bound approximately Np^n.
Fluid flow, cooling, and fullerene storage			
Q, A_{channel}	Volumetric flow and channel area	m^3/s; m^2	$U_{\text{mean}} = Q/A_{\text{channel}}$.
h, y	Channel height and height above the die	m	Plane Poiseuille flow; this h is not Planck's constant.
$U_{\text{mean}}, u(y)$	Mean and local fluid velocities	m/s	Channel flow, Eq. (122).
μ	Dynamic viscosity	Pa s	Fluid drag; distinct from the prefix micro.
T_c, T_h	Cold and hot reservoir temperatures	K	Carnot cooling coefficient of performance.
$\text{COP}_{\text{Carnot}}$	Ideal cooling coefficient of performance	1	$T_c/(T_h - T_c)$.
P	Pressure	Pa	Fullerene hydrogen storage and atmospheric loads; not power.
$r_{\text{full}}, r_{\text{in}}$	Outer and interior fullerene radii	m	$r_{\text{in}} = r_{\text{full}} - 0.17\,\text{nm}$.
V	Interior fullerene volume	m^3	$V = (4/3)\pi r_{\text{in}}^3$; not voltage.
N_C	Carbon atom count	1	Fullerene cage inventory.
$n_{\text{H}_2}, N_{\text{H}_2}$	Hydrogen number density and molecule count	m^{-3}; 1	$N_{\text{H}_2} \approx n_{\text{H}_2} V$.
Physical constants and atomic densities			
$h, \hbar$	Planck and reduced Planck constants	J s	$\hbar = h/(2\pi)$; not channel or comb height.
k_B	Boltzmann constant	J/K	Thermal energy scale $k_B T$.
e	Elementary charge magnitude	C	Electrical and tunneling calculations.
ϵ_0	Vacuum permittivity	F/m	Electrostatic force and capacitance.
N_A	Avogadro constant	mol^{-1}	Conversion between particle and molar quantities.

Continued on next page

Table 21 (continued)

Symbol	Meaning	Units	Scope or qualification
n_C, n_{2D}	Diamond volume and graphene area atom densities	m^{-3}; m^{-2}	Material inventories; distinct from atom counts.
b	Graphitic C–C bond length	m	Graphene areal-density calculation.

17.2 One-letter prefixes used before units

m milli 10^{-3}

µ micro 10^{-6}

n nano 10^{-9}

p pico 10^{-12}

f femto 10^{-15}

a atto 10^{-18}

z zepto 10^{-21}

y yocto 10^{-24}

17.3 Units

Distance m meter

Time s second

Force N newton

Mass kg kilogram

Mass g gram 10^{-3} kg

Energy J joule

Capacitance F farad

Inductance H henry

Pressure Pa pascal

Power W watt

Voltage V volt

Current A ampere

Charge C coulomb

Temperature K kelvin

Resistance Ω ohm

Amount of substance mol mole

Frequency Hz hertz

The area-unit abbreviation aha denotes an attohectare ($10^{-14}\,m^2$).

17.4 Physical Constants

N_A Avogadro's constant $6.02214076 \times 10^{23}\,\mathrm{mol}^{-1}$ (exact) [Nata]

u, Da AMU, dalton 1.660 539 068 92(52) $\times$ 10^{-27} kg [Nata].

h Planck's constant $6.62607015 \times 10^{-34}\,\mathrm{J\,s}$ (exact) [Nata]

k_B Boltzmann's constant 1.380649×10^{-23} J/K [Nata]

e Elementary-charge magnitude $1.602176634 \times 10^{-19}\,\mathrm{C}$ [Nata]

ϵ_0 vacuum permittivity 8.854 187 8188(14) $\times$ 10^{-12} F/m [Nata]

17.5 Useful Constants

Density of diamond 3515 kg/m^3. Its carbon atom density, n_C, is approximately 8 atoms per lattice-spacing cubed:

$$n_C \approx 1.76 \times 10^{23}\,\mathrm{cm}^{-3} = 1.76 \times 10^{29}\,\mathrm{m}^{-3} = 176\,\mathrm{nm}^{-3}. \tag{120}$$

lattice spacing of diamond (300 K): 0.356 683 nm [Sze81]

lattice spacing of silicon (300 K): 0.543 095 nm [Sze81]

lattice spacing of germanium (300 K): 0.564 613 nm [Sze81]

lattice spacing of grey tin (300 K): 0.648 920 nm [Sze81]

C-C bond length in diamond: 0.154 448 nm (lattice spacing $\times\frac{1}{4}\sqrt{3}$)

C-C bond length in graphite: 0.142 nm [Kel81]

1-3 distance in graphite: 0.246 nm; C-C bond length $\times\sqrt{3}$

The areal carbon atom density of a monoatomic sheet of graphite is n_{2D}, with $b = 0.142\,\mathrm{nm}$ the C-C bond length:

$$n_{2D} = \frac{4}{3\sqrt{3}\,b^2} \approx 38.2\,\mathrm{nm}^{-2}. \tag{121}$$

Number of carbon atoms/nm^3 of graphite: 114

Distance between adjacent sheets in graphite: 0.335 nm [Kel81]

Basal surface energy of graphite: 0.234 J/m^2 [Kel81]

bandgap of diamond: 876 zJ (5.47 eV) [Sze81]

assumed dielectric constant of diamond: 5.7 [Appb]

bandgap of silicon: 179 zJ (1.12 eV) [Sze81]

mobility of electron in diamond: 1800 cm^2/(V $\times$ s) [Sze81]

mobility of electron in silicon: 1500 cm^2/(V $\times$ s) [Sze81]

Stiffness of carbon-carbon single bond: 440 N/m [Dre92]

Energy required to break carbon-carbon single bond: 0.556 aJ [Dre92]

Young's modulus of diamond: 1.05×10^{12} Pa

Tensile strength of diamond: 5×10^{10} Pa [Dre92]

17.6 Conversion factors

one calorie (cal) is 4.184 J [Natb]

one joule (J) is 0.0002390057361 kcal (derived from above)

one joule (J) per molecule is $1.439326185 \times 10^{20}$ kcal/mol (derived from above)

An energy of 1 eV per particle corresponds to approximately 96.485 kJ/mol (derived from the constants above).

one kcal/mol corresponds to $6.947695457 \times 10^{-21}$ J per molecule (derived from above)

one Hartree is 4.359 744 722 2060(48) $\times$ 10^{-18} J [Nata]

one electron volt (eV) is $1.602\,176\,634 \times 10^{-19}$ J (exact) [Nata]

18 Appendices

18.1 Feedstock discussion

The baseline elements are H, B, C, N, O, Si, Ge, and Pt. It's easy to design feedstock that includes C and H. There is more of an issue with Ge combined with our desire for low temperature. We need a germanium containing feedstock with a low melting point. The obvious choice is GeH_4, germane, with a melting point of 108 K. Germane raises serious health concerns: NOAA's CAMEO Chemicals database lists an NFPA 704 health rating of 4 ("Can be lethal") [Natc]. Safer germanium compounds include tetraethylgermane and tetraethoxygermanium. As tetraethylgermane is somewhat safer from a health perspective (NFPA 704 health rating of 2 rather than 3) and has a lower melting point (180.5 K versus 201 K) than tetraethoxygermanium, we baseline tetraethylgermane.

Other Ge sources considered raised greater safety concerns. Safer Ge compounds with a lower melting point are of interest.

Baseline feedstock:

Tetraethylgermane, pentanenitrile, tetraethyl orthosilicate (TEOS), triethyl borate, $EtCpPtMe_3$.

Tetraethylgermane has a melting point of 180.5 K. Pentanenitrile has a melting point of 176.9 K. It is expected to be miscible with tetraethylgermane. Enough tetraethylgermane is ingested to provide exactly the amount of Ge that is required by the products being manufactured. Likewise, enough triethyl borate is ingested to provide exactly the amount of B that is required by the products being manufactured. Enough pentanenitrile is ingested to provide the C, H, and N required by the products being manufactured. Enough TEOS is ingested to provide the required Si and O, along with its accompanying C and H. Enough $EtCpPtMe_3$ is ingested to supply the Pt required by the products being manufactured. Excess H, C, N, O, and Si are vented as H_2, C_2H_2, N_2, CO_2, and $C_4H_{12}Si$ (tetramethylsilane).

Toluene was originally included as an inert solvent/diluent. In the present feedstock formulation, however, all five feedstock compounds are liquids near 210 K (see the pure-compound melting points in Table 9), and toluene supplies only carbon and hydrogen while diluting the heteroatom sources. It will be omitted but can be used for risk mitigation if needed to suppress crystallization, phase separation, or excessive viscosity at the operating temperature.

Baseline: the SRU will generate and vent H_2, C_2H_2, N_2, CO_2, and $C_4H_{12}Si$ (tetramethylsilane) as needed to address any excess of H, C, N, O, or Si provided by the feedstock.

Baseline operating temperature: 210 K

The feedstock dictates the SRU temperature of operation. We pick a temperature of operation higher than the nominal melting point of any component by a comfortable margin.

Risk: There is a risk that the chosen baseline operating temperature will be too high, either because the positional uncertainty caused by the increased operating temperature causes an unacceptable increase in error rates, or because the reliability of marginal reactions becomes unacceptable due to thermal errors, which scale as $e^{-E_a/(k_B T)}$ with temperature, where E_a is the relevant activation barrier, T is temperature in kelvin and k_B is Boltzmann's constant.

Mitigation: Should this risk be realized, one mitigation is to adopt thermal cycling. The size of the SRU is increased to allow it to store all of the feedstock required to manufacture an SRU. The SRU then ingests the required feedstock at 210 K. The temperature of the system is then lowered to the temperature required for reliable mechanosynthesis. Manufacture and extrusion of the daughter SRU then take place at the lower temperature, using the stored feedstock. The temperature is then

increased to allow ingestion and storage of the feedstock required to manufacture the next daughter, and the cycle repeats.

18.2 Units used in this paper

We favor SI units. There are several reasons for this. First, we cross disciplines. Each discipline has its own favorite units. Agreement on a common set of units across disciplines is therefore useful. For example, joules are already widely used. Hartrees, kcal/mol, eV, and other units can all be converted to joules (or aJ, where $1\,\text{aJ} = 10^{-18}\,\text{J}$, or zJ, where $1\,\text{zJ} = 10^{-21}\,\text{J}$, when considering chemical energies among atoms), creating a common language.

Second, we often have recourse to Newtonian mechanics, involving such equations as $W = Fd$, where d is distance. If F is in newtons and d is in meters, it's most convenient if W is in joules. Other equations are also simplified if all units are standard SI. Where research in a relevant area uses non-SI units, we use SI first and non-SI in parentheses afterwards, e.g., 801 zJ (5 eV).

Third, many important fields already use SI units.

18.3 The force exerted by atmospheric pressure at the nanoscale

One standard atmosphere is exactly 101,325 Pa. A useful exact unit conversion is

$$10^5\,\text{Pa} = 1\,\text{nN/aha},$$

where nN is a nanonewton and aha is an attohectare. One aha is a square 100 nm on a side. One atmosphere is 101,325 N/m^2, or 1.01325 nN per attohectare (aha).

18.4 Force on daughter from fluid flow

As a new daughter is extruded by its parent, and as the new daughter is moved by a parent both to complete attachment of the external assembly arm and to pass the daughter hand-over-hand to a vacant SRU site, the daughter will experience hydrodynamic forces caused by the relative motion of the feedstock with respect to the daughter. In this section we estimate the magnitude of these forces.

For purposes of this analysis, we treat the SRU as a cylinder 500 nm long with a radius of 50 nm. Our first question is what lateral forces might be exerted on the daughter by the flow of feedstock over the surface of the die.

The flow rate of tetraethylgermane coolant over a board is approximately 96 cm^3/s (see Cooling the nanofactory (Section 10.8)). Dividing this among 576 dies and assuming 2 mm of clearance across the 10 mm die width gives:

$$\begin{aligned} Q &= \frac{96\,\text{cm}^3/\text{s}}{576} \approx 0.167\,\text{cm}^3/\text{s} = 1.67 \times 10^{-7}\,\text{m}^3/\text{s}, \\ A_{\text{channel}} &= (2\,\text{mm})(10\,\text{mm}) = 2.0 \times 10^{-5}\,\text{m}^2, \\ U_{\text{mean}} &= \frac{Q}{A_{\text{channel}}} \approx 8.33 \times 10^{-3}\,\text{m/s}. \end{aligned}$$

Thus, the average fluid velocity through the channel is about 8.3 mm/s. The daughter, however, remains very close to the stationary die surface, where the no-slip condition makes the local fluid velocity much smaller than the channel average.

18.4.1 Near-surface fluid velocity

We approximate the flow as fully developed laminar flow between two parallel plates separated by a distance $h = 2$ mm. For plane Poiseuille flow, the velocity at height y above the die is:

$$u(y) = 6U_{\text{mean}} \frac{y}{h} \left(1 - \frac{y}{h}\right). \tag{122}$$

Because the daughter remains within 1 µm of the die, y/h is at most 0.0005. Near the wall, the term $(1 - y/h)$ is therefore essentially one, giving:

$$u(y) \approx 6U_{\text{mean}} \frac{y}{h}. \tag{123}$$

At $y = 1$ µm, this gives:

$$u(1\,\mu\text{m}) \approx 6 \times 8.33 \times 10^{-3}\,\text{m/s} \times \frac{1 \times 10^{-6}\,\text{m}}{2 \times 10^{-3}\,\text{m}}$$
$$u(1\,\mu\text{m}) \approx 2.5 \times 10^{-5}\,\text{m/s} = 25\,\mu\text{m/s}$$

This is a conservative velocity for estimating the force on the daughter, because it assigns the velocity at the highest allowed position to the entire daughter. If a 500 nm-long daughter extends vertically from the die surface, its midpoint is only 250 nm above the surface and its height-averaged ambient flow velocity is approximately 6.3 µm/s.

18.4.2 Viscous drag on the daughter

The Reynolds number on the 100 nm daughter diameter is far below one, so inertial drag is negligible and the force is dominated by viscous Stokes flow. For transverse flow past a slender cylinder, a suitable back-of-the-envelope expression is:

$$F_{\text{drag}} \approx \frac{4\pi\mu L u}{\ln(L/r)}, \tag{124}$$

where μ is the dynamic viscosity, L is the cylinder length, r is its radius, and u is the relative fluid velocity. Mogul et al. report viscosity measurements for pure tetraethylgermane over approximately 278–337 K [MHM+74]. An Arrhenius fit to their tabulated data, $\ln[\mu/(\text{mPa}\cdot\text{s})] \approx -3.167 + (842.4\,\text{K})/T$, extrapolates to approximately 2.3 mPa·s at 210 K. This is an extrapolated pure-fluid estimate, not a measurement at 210 K; the viscosity of the operating feedstock mixture remains to be established. For this preliminary drag calculation, we use $\mu \approx 2.3 \times 10^{-3}\,\text{Pa}\cdot\text{s}$, $L = 500$ nm, $r = 50$ nm, and the conservative velocity $u = 25$ µm/s. Since $\ln(L/r) = \ln(10) \approx 2.30$:

$$F_{\text{drag}} \approx \frac{4\pi \times (2.3 \times 10^{-3}\,\text{Pa}\cdot\text{s}) \times (500 \times 10^{-9}\,\text{m}) \times (25 \times 10^{-6}\,\text{m/s})}{2.30}$$
$$F_{\text{drag}} \approx 1.6 \times 10^{-13}\,\text{N} = 0.16\,\text{pN}$$

Using the height-averaged flow velocity of approximately 6.3 µm/s for a daughter extending from the die surface instead gives approximately 0.040 pN. The cylinder has only a 10:1 aspect ratio, its orientation can vary, and the nearby wall can increase viscous resistance. These effects make the numerical coefficient uncertain by a factor of order unity. It is therefore reasonable to summarize the lateral force caused by the specified bulk coolant flow as being of order 1 pN or less.

Even a 1 pN load would be only one-thousandth of the approximately 1 nN available output force after the baseline comb drive's 12:1 transmission. The background flow of feedstock over the die should therefore not interfere significantly with extrusion, attachment of the external assembly arm, or hand-over-hand transfer of a daughter SRU.

18.4.3 Drag caused by deliberate movement of the daughter

The same approximation can be used to estimate drag when a parent moves a daughter through otherwise nearly stationary feedstock. For the assumed cylinder and viscosity, the drag is approximately proportional to the daughter's speed relative to the fluid:

$$F_{\text{move}} \approx \left(6.3 \times 10^{-9}\,\frac{\text{N s}}{\text{m}}\right) v_{\text{relative}}, \tag{125}$$

where v_{relative} is the speed relative to the fluid, measured in meters per second. Representative values are:

Table 22: Daughter-motion drag estimates from Eq. (125). Near-wall and geometry corrections are not included.

Relative speed	Estimated drag
1 mm/s	6.3 pN
1 cm/s	0.063 nN
5 cm/s	0.31 nN

Consequently, background coolant flow is not a significant load, but the speed chosen for deliberate daughter motion can be. Keeping sustained daughter speeds well below approximately 16 cm/s should keep the uncorrected viscous drag below 1 nN. Because proximity to the die and other SRUs can increase drag, a more conservative operating speed may be appropriate. All these drag estimates scale directly with the assumed viscosity. Actual transfer trajectories and speeds should be included in the eventual mechanical load budget.

18.4.4 Conclusion

The no-slip condition reduces the background fluid velocity within 1 µm of the die to no more than approximately 25 µm/s. The resulting lateral force on the daughter is approximately 0.04–0.16 pN for the extrapolated viscosity used here and, allowing for geometrical and near-wall corrections, is reasonably characterized as being of order 1 pN or less. This is negligible compared with the approximately 1 nN actuation-force scale of the SRU. Drag caused by deliberate daughter motion scales with speed and can become comparable to 1 nN at relative speeds of order 16 cm/s in this uncorrected model.

18.5 Buckyballs as a storage method for hydrogen gas

18.5.1 Summary

Table 23 summarizes geometric and mechanical estimates for a fullerene cage.

Table 23: Geometric and mechanical estimates for a fullerene cage.

Quantity	Estimate
Number of carbon atoms in a buckyball	$N_C \approx 4\pi r_{\text{full}}^2\,(38.2\,\text{nm}^{-2})$
Interior volume	$V_{\text{interior}} = (4/3)\pi(r_{\text{full}} - 0.17\,\text{nm})^3$
H_2 number density at 210 K and 1,000 atm [LJPL09]	approximately 18.2 molecules/nm^3
H_2 molecules in a buckyball	$N_{H_2} \approx V_{\text{interior}}\,(18.2\,\text{nm}^{-3})$
Idealized bond tension at 1,000 atm	$F_{\text{bond}} = (\sqrt{3}/2)Pb\,r_{\text{in}}^2/r_{\text{full}}$
Force per bond for $r_{\text{full}} = 7.5$ nm	approximately 0.089 nN
Rupture force of a C-C bond	approximately 6 nN [TSA21]
Idealized bond-force extrapolation radius	approximately 480 nm
H_2 capacity at the extrapolated radius	approximately 8.5×10^9 H_2 molecules

18.5.2 Analysis

Fullerenes (C_n, "buckyballs") offer a sealed, all-carbon nanocontainer whose size can, in principle, be tuned from a few hundred picometers to many nanometers. We combine simple shell geometry with real-gas thermodynamics to estimate the number of H_2 molecules that can be confined in a theoretical 15 nm-diameter megabuckyball and the mechanical load that the internal gas exerts on the carbon-carbon bonds of the cage.

At the baseline SRU operating temperature of 210 K and an internal pressure of 1,000 atm, a 15 nm-diameter cage stores approximately 30,000 H_2 molecules. In the idealized isotropic bond-network model below, the pressure-induced tension per C-C bond is approximately 0.089 nN, about 1.5% of the assumed 6 nN single-bond rupture criterion. This model therefore suggests a substantial bond-force margin at this size; it does not establish atomistic cage stability.

For nearly icosahedral fullerenes, the fullerene radius is approximately:

$$r_{\text{full}} \approx (0.046\,\text{nm})\,\sqrt{N_C}, \tag{126}$$

where N_C is the number of carbon atoms. The wall thickness is taken as one van der Waals radius of carbon, 0.17 nm. Therefore:

$$r_{\text{in}} = r_{\text{full}} - 0.17\,\text{nm}$$
$$V = \frac{4}{3}\pi r_{\text{in}}^3. \tag{127}$$

18.5.3 Number of carbon atoms in the 15 nm cage

For an outside radius of 7.50 nm, $A = 4\pi r_{\text{full}}^2 \approx 707\,\text{nm}^2$. Graphene contains approximately 38.2 carbon atoms/nm^2, so $N_C \approx 707\,\text{nm}^2 \times 38.2\,\text{nm}^{-2} \approx 2.7 \times 10^4$. These and the hydrogen-capacity results are collected in Table 24.

An illustrative icosahedral magic-number fullerene is $C_{26,460}$, corresponding to $(h, k) = (21, 21)$.

18.5.4 Hydrogen capacity at 210 K

The 0.17 nm wall allowance gives $r_{\text{in}} = 7.33\,\text{nm}$ and $V \approx 1.65 \times 10^3\,\text{nm}^3$ for the 15 nm cage.

NIST normal-hydrogen equation-of-state data [LJPL09] give a hydrogen density of approximately 60.5 kg/m^3 at 210 K and 100 MPa. Because 1,000 atm is 101.325 MPa, we use approximately 61 kg/m^3, corresponding to approximately 18.2 H_2 molecules per cubic nanometer. Equivalently, the compressibility factor is approximately 1.92. Here n_{H_2} denotes the hydrogen number density.

$$N_{H_2} \approx n_{H_2} V \approx (18.2\,\text{nm}^{-3})(1.65 \times 10^3\,\text{nm}^3) \approx 3.0 \times 10^4. \tag{128}$$

Thus, a 15 nm-diameter buckyball should hold approximately 30,000 H_2 molecules at 210 K and 1,000 atm.

18.5.5 Mechanical integrity

If the sphere is divided along a great circle, the total force tending to separate the two hemispheres is:

$$F = P\pi r_{\text{in}}^2 \approx 1.7 \times 10^{-8}\,\text{N}. \tag{129}$$

To estimate bond tension, model the shell as a locally regular honeycomb under uniform isotropic membrane tension $\mathcal{T}$ (force per length). Let the C-C bond length be $b = 0.142\,\text{nm}$ and the common axial bond tension be F_{bond}. A two-atom cell has area $3\sqrt{3}b^2/2$ and three distinct bonds, giving bond areal density $n_b = 2/(\sqrt{3}b^2)$. Under a small equal strain increment $d\epsilon$ in both in-plane directions, virtual work per area gives $2\mathcal{T}\,d\epsilon = n_b F_{\text{bond}} b\,d\epsilon$, hence $\mathcal{T} = F_{\text{bond}}/(\sqrt{3}b)$. This includes the projections of all three bond orientations.

Hemisphere force balance gives $2\pi r_{\text{full}}\mathcal{T} = P\pi r_{\text{in}}^2$, where P is the pressure difference across the shell. Therefore:

$$F_{\text{bond}} = \frac{\sqrt{3}}{2} P b \frac{r_{\text{in}}^2}{r_{\text{full}}} \approx 0.089\,\text{nN} \tag{130}$$

for the 15 nm cage at $P = 101.325\,\text{MPa}$. The corresponding membrane tension is approximately 0.363 N/m. The bond load is about 1.5% of the assumed 6 nN rupture force [TSA21].

Table 24: Calculated 15 nm fullerene example at 210 K and 1,000 atm. The idealized bond-force extrapolation to about 480 nm radius is a separate calculation.

Quantity	**Value**
Outside radius	7.50 nm
Interior radius	7.33 nm
Surface area	$707\,\mathrm{nm}^2$
Carbon atoms	$\sim 2.7 \times 10^4$
Interior volume	$1.65 \times 10^3\,\mathrm{nm}^3$
Hydrogen molecules	$\sim$ 30,000
Hemisphere separating force	1.7×10^{-8} N
Membrane tension	0.363 N/m
Idealized bond tension	0.089 nN

18.5.6 Idealized size limit

For fixed pressure and shell structure, the idealized bond tension approaches linear scaling with radius. With radii entered in nanometers, the derived relation is:

$$F_{\mathrm{bond}}[\mathrm{nN}] \approx 0.0124605\,\frac{(r_{\mathrm{full}}[\mathrm{nm}] - 0.17)^2}{r_{\mathrm{full}}[\mathrm{nm}]}. \tag{131}$$

Setting this force to the assumed 6 nN criterion gives:

$$r_{\mathrm{full}} \approx 481.86\,\mathrm{nm} \approx 480\,\mathrm{nm}. \tag{132}$$

The corresponding diameter is approximately 960 nm. This is an idealized bond-force extrapolation, not an established practical rupture size or a rigorous upper bound. It assumes uniform isotropic tension in a locally regular network and neglects pressure-driven changes in shape, bond length, and gas volume. It does not resolve pentagon-induced stresses, curvature-specific forces, thermal rupture, defects, or nonlinear deformation near rupture.

At 210 K and 1,000 atm, using the unrounded extrapolated radius, the corresponding hydrogen inventory is:

$$\begin{aligned} N_{\mathrm{H_2}} &\approx \frac{4}{3}\pi(481.86\,\mathrm{nm} - 0.17\,\mathrm{nm})^3(18.2\,\mathrm{nm}^{-3}) \\ &\approx 8.5 \times 10^9. \end{aligned} \tag{133}$$