Efficient Realization of Geometric Constraint Systems via Optimal Recursive Decomposition and Cayley Convexification

Meera Sitharam

University of Florida

July 29, 2016

Talk Based On.. I

Papers on Decomposition [1] [3] [4] Convex Cayley Spaces, [7] [9] [11] [10] Applications [5] [6] [12] [2] [8]
Supported in part by NSF/DMS grants 2007,2011,2016

T. Baker, M. Sitharam, M. Wang, and J. Willoughby.
Optimal decomposition and recombination of isostatic geometric constraint systems for designing layered materials.
Computer Aided Geometric Design, 40:1 – 25, 2015.

Miklós Bóna, Meera Sitharam, and Andrew Vince.
Enumeration of viral capsid assembly pathways: Treeorbits under permutation group action.
Bulletin of Mathematical Biology, 73(4):726–753, 2011.

Christoph M. Hoffman, Andrew Lomonosov, and Meera Sitharam.
Decomposition plans for geometric constraint systems, part I: Performance measures for CAD.
Journal of Symbolic Computation, 31(4):367–408, 2001.

Christoph M Hoffman, Andrew Lomonosov, and Meera Sitharam.
Decomposition plans for geometric constraint systems, part ii: Algorithms.
Journal of Symbolic Computation, 31(4), 2001.

Aysegul Ozkan and Meera Sitharam.
EASAL (efficient atlasing, analysis and search of molecular assembly landscapes).
In *Proceedings of the ISCA 3rd International Conference on Bioinformatics and Computational Biology, BICoB-2011*, pages 233–238, 2011.

Meera Sitharam and Mavis Agbandje-mckenna.
Modelling virus self-assembly pathways: Avoiding dynamics using geometric constraint decomposition.
J. Comp. Biol, page 65, 2006.

Talk Based On.. II

Meera Sitharam and Heping Gao.
Characterizing graphs with convex and connected cayley configuration spaces.
Discrete & Computational Geometry, 43(3):594–625, 2010.

Meera Sitharam, Andrew Vince, Menghan Wang, and Mikls Bna.
Symmetry in sphere-based assembly configuration spaces.
Symmetry, 8(1):5, 2016.

Meera Sitharam and Menghan Wang.
How the beast really moves: Cayley analysis of mechanism realization spaces using caymos.
Computer-Aided Design SIAM SPM 2013 issue, 46:205 – 210, 2014.
2013 {SIAM} Conference on Geometric and Physical Modeling.

Meera Sitharam and Joel Willoughby.
On Flattenability of Graphs, pages 129–148.
ADG Springer Lecture Notes, 2015.

Menghan Wang and Meera Sitharam.
Algorithm 951: Cayley analysis of mechanism configuration spaces using caymos: Software functionalities and architecture.
ACM Trans. Math. Softw., 41(4):27:1–27:8, October 2015.

Ruijin Wu, Aysegul Ozkan, Antonette Bennett, Mavis Agbandje-Mckenna, and Meera Sitharam.
Robustness measure for an adeno-associated viral shell self-assembly is accurately predicted by configuration space atlasing using easal.
In *Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Biomedicine*, BCB '12, pages 690–695, New York, NY, USA, 2012. ACM.

Outline

Realizing Linkages given dimension

EDM Completion given rank = PSD Completion given rank

Definition (Realizing a Linkage)

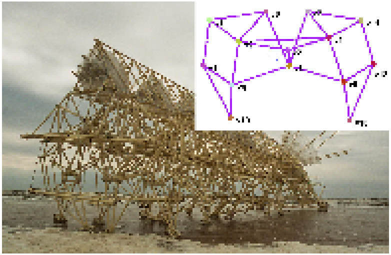
Given graph $G = (V, E, \delta)$ with $\delta : E \to \mathbf{Q}$,

- find/describe the **set** of all $p : V \to \mathbf{R}^d$ with $||p_u - p_v|| = \delta(u, v)$, modulo trivial transformations.
- equivalently, find/describe the **set** of all completions of $\delta(u, v) = ||p_u - p_v||$ from E to $V \times V$.

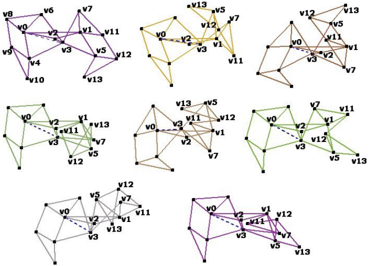
Realization = Solution of GCS

- Problem: Finding/Roadmapping the real solution set of the corresponding polynomial (typically quadratic) system.
- Extends to other Geometric Constraint Systems with underlying constraint (hyper)graphs (other distance metrics, types of constraints), with corresponding trivial transformation groups.
- Numerous applications: Computer Aided Mechanical/Structural design, Robotics, Graphics and Computer Vision, Molecular Configuration Spaces.

A linkage

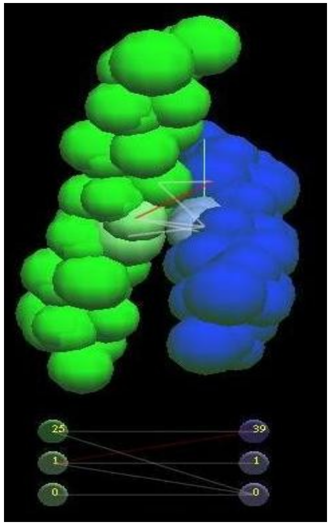


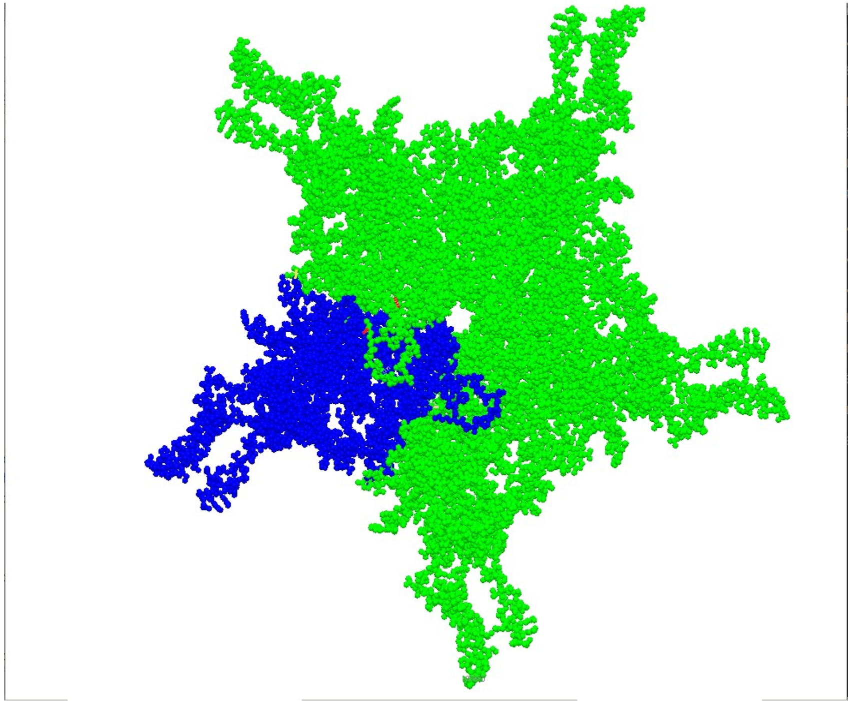
Connected Components



Problems

- Configuration Space Atlasing &
- Configurational Entropy Computation for
  - Assemblies of upto 5 rigid molecular motifs given
    - pair potentials and sterics
    - global constraints



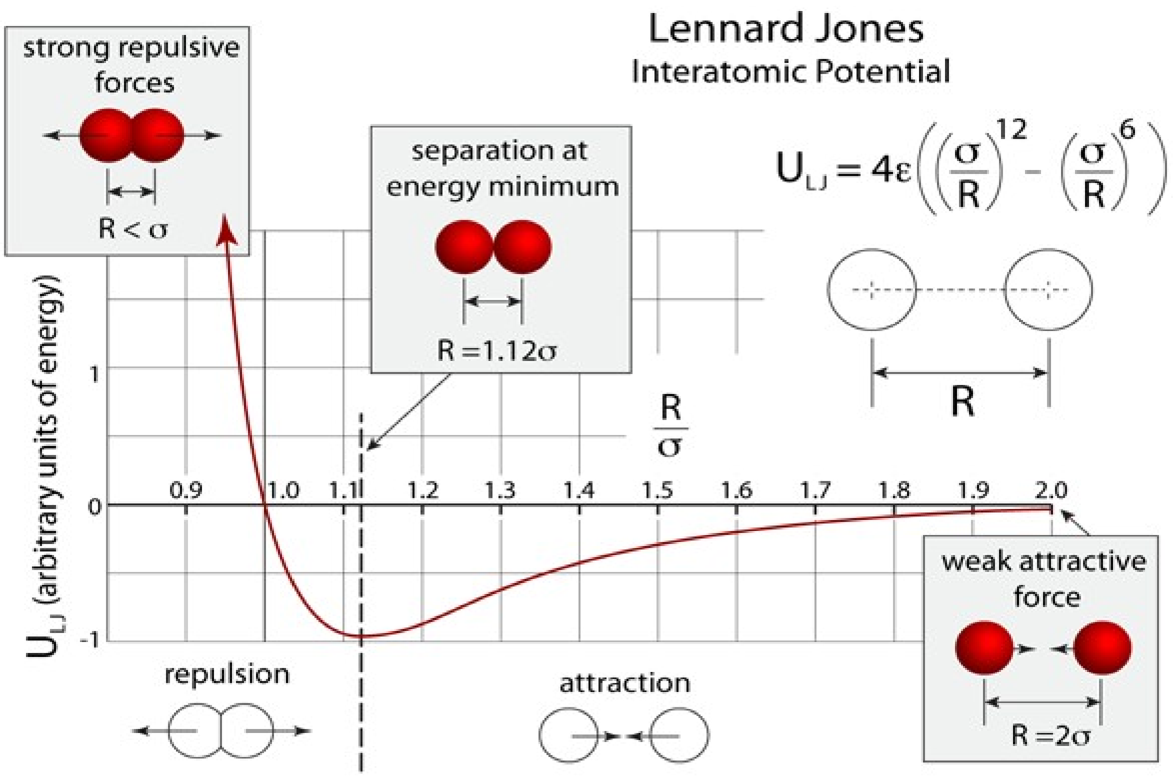


Problems

- Configuration Space Atlasing &
- Configurational Entropy Computation for
  - Assemblies of upto 5 rigid molecular motifs given
    - pair potentials and sterics
    - global constraints

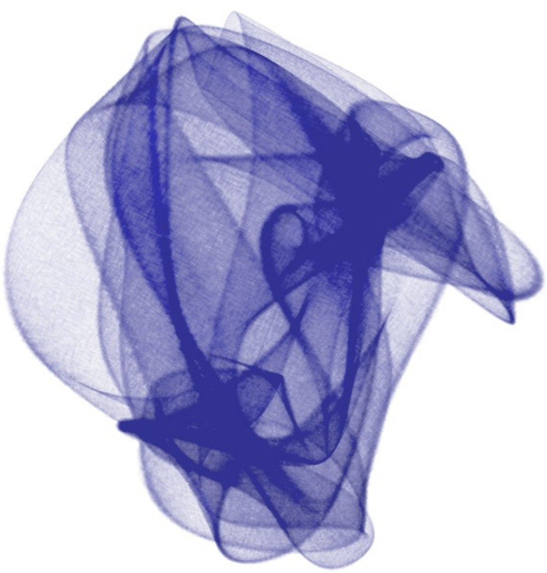
Courtesy:

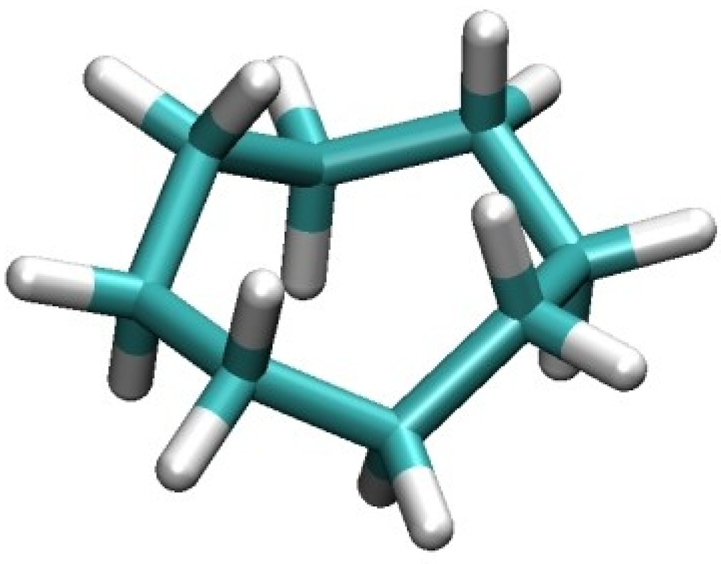
Atoms in Motion



Problems

- Configuration Space Atlasing &
- Configurational Entropy Computation for
  - Assemblies of upto 5 rigid molecular motifs
  - Small molecules with loop closure, pair potentials/sterics





Courtesy: CUIK project, Barcelona

Problems

- Configuration Space Atlasing &
- Configurational Entropy Computation for
  - Assemblies of upto 5 rigid molecular motifs
  - Small molecules with loop closure
  - Sticky sphere systems (sterics)

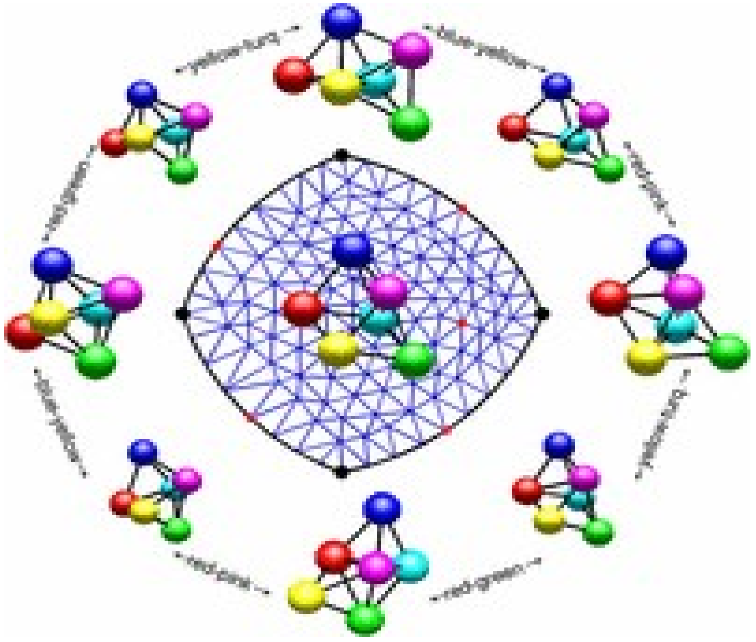
Courtesy:

“A geometrical approach to computing free-energy

landscapes from short-ranged potentials”

Miranda Holmes-Cerfon, Steven J. Gortler,

Michael P. Brenner PNAS v110(1)



Problems

- Configuration Space Atlasing &
- Computation of Free Energy & Formation Rate (kinetics) for
  - Assemblies of upto 5 rigid molecular motifs
  - Small molecules with loop closure
  - Sticky sphere systems
- Prediction of Crucial Interactions for Larger Assemblies e.g. Viral Shells

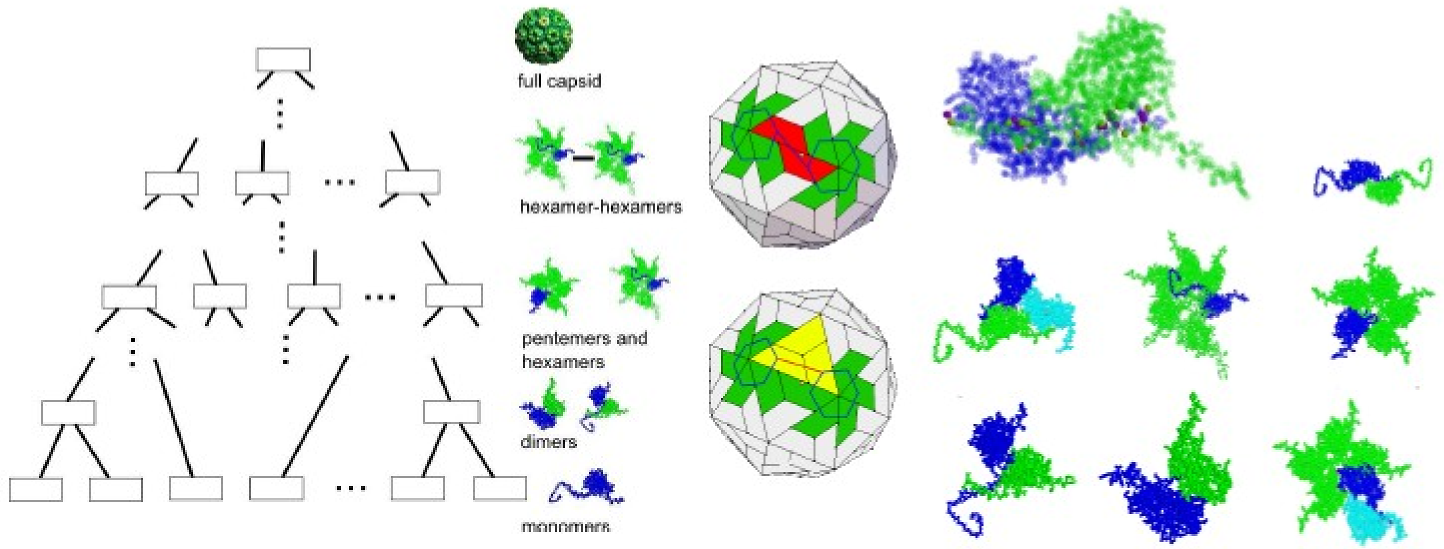


Figure 1: Left T=3 BMV assembly tree schematic; Mid: different bi-assembly interfaces (2-fold or 3-fold) for the same intermediate from the same pentamer subassemblies; Right top: 2-fold interface types, one magnified; mid: 3, 5 and 6 fold interfaces; bottom: bi-assemblies of monomers extracted from 5 and 6 fold and larger bi-assembly of small multimers from 5-fold at bottom right (the exhaustive list of such bi-assemblies of small multimers contains approx 24 for BMV)

Motivating Decomposition

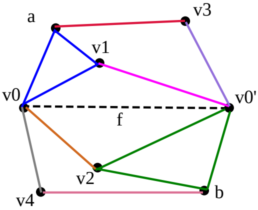
–Complexity of solving a quadratic system prohibitively high.
– Easy Case: Triangularizable System (maintaining degree 2) - *QRS (quadratically radically solvable,* or "ruler and compass" systems.
– A corresponding natural class of graphs:

Definition

For dimension 2, G is $\triangle$-decomposable if it is a single edge, or can be divided into 3 $\triangle$-decomposable subgraphs s.t. every two of them share a single vertex.

Note: $\triangle$-decomposable implies minimally rigid

- There is a base edge f with a graph construction from f: each step appends a new vertex shared by 2 $\triangle$-decomposable subgraphs (clusters)
- Corresponding linkages have a ruler and compass realization parallel to the graph theoretical construction
- Extends to arbitrary dimension d.



Outline

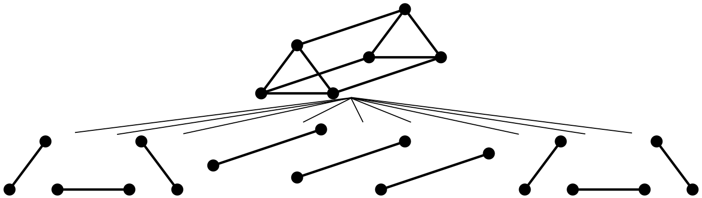
Decomposition for Recombination

Definition (Decomposition-recombination (DR-) plan)

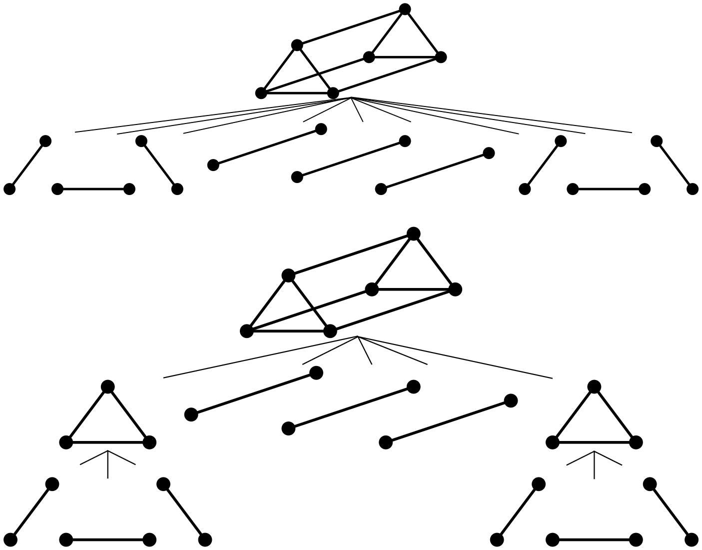
A *DR-plan* of constraint graph G is a forest where:

- Each node is a rigid subgraph of G.
- A root node is a vertex-maximal rigid subgraph.
- An internal node is the union of its children.
- A leaf node is a single edge

Example DR-Plans: $C_2 \times C_3$



Example DR-Plans: $C_2 \times C_3$



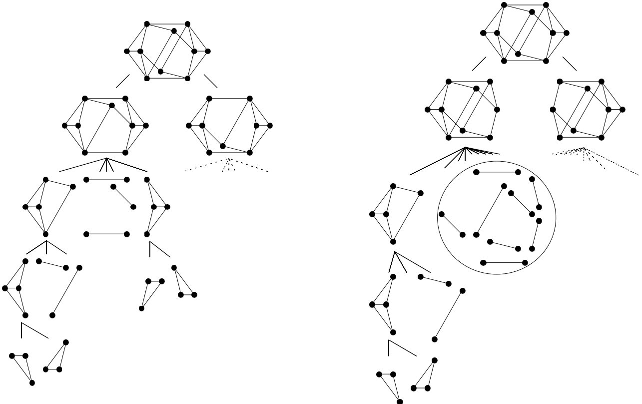
Optimal DR-Plan

An *optimal DR-plan* minimizes the maximum fan-in. Corresponds to the largest system that needs to be solved simultaneously.

Optimal DR-Plan

An *optimal DR-plan* minimizes the maximum fan-in. Corresponds to the largest system that needs to be solved simultaneously.

In general, finding optimal is NP-hard.



Uses of DR-Planning

- (Determining complexity of) Realization.
- Decomposition of the stress and flex matrices.
- Completion to Rigid.
- Interactive removal of dependent edges/constraints.

History

2001: Formalized in HoffmanLomonosovSitharamJSC2001

Late 1980's: Began with triangle-decomposable graphs. Corresponds to systems that can be triangularized and therefore have quadratic radical solutions (QRS).

1990's-2000's: Older algorithms were *bottom-up* and were based on maximum matching. E.g., Frontier. Polynomial time, ensuring some properties other than optimality.

2015: When graph is independent, our paper BakerSitharamWangWilloughbyCAGD2015 contains a top-down $O(|V|^3)$ algorithm with a formal guarantee to find an optimal DR-plan.

Outline

Summary of Results

If the geometric constraint system we are considering...

- Has an underlying abstract rigidity matroid $\rightarrow$ We can push the structure theorems through.

Summary of Results

If the geometric constraint system we are considering...

- Has an underlying abstract rigidity matroid $\rightarrow$ We can push the structure theorems through.
- Is independent $\rightarrow$ We achieve optimality of DR-plan.

Summary of Results

If the geometric constraint system we are considering...

- Has an underlying abstract rigidity matroid $\rightarrow$ We can push the structure theorems through.
- Is independent $\rightarrow$ We achieve optimality of DR-plan.
- Has an underlying sparsity matroid $\rightarrow$ We get a polynomial time algorithm.

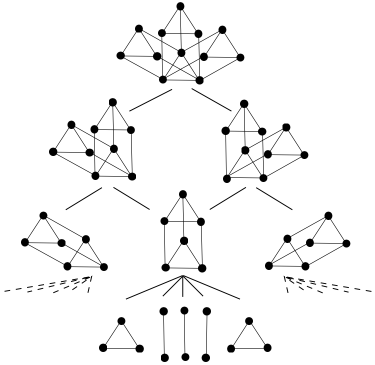
For 2D linkages we have $O(|V|^3)$ time algorithm.

Canonical DR-Plan

Definition (Canonical DR-plan)

A DR-plan that satisfies the additional two properties:

1. Children are rigid vertex-maximal proper subgraphs (rvmps) of the parent.
2. If all pairs of rvmps intersect trivially then all of them are children, otherwise exactly two that intersect non-trivially are children.



Importance of Canonical DR-Plan

We restrict the space of DR-plans for the input to this special class of canonical DR-plans.

Theorem

*A canonical DR-plan exists for a graph G and any canonical DR-plan is optimal if G is independent.**

*Applies when G has an underlying abstract rigidity matroid.

Proof is non-trivial.

Algorithmic Result

Theorem

*Computing an optimal DR-plan for an independent graph G has time complexity $O(|V|^3)$.**

*Provided there exists underlying sparsity matroid.

Algorithmic Result

Theorem

*Computing an optimal DR-plan for an independent graph G has time complexity $O(|V|^3)$.**

*Provided there exists underlying sparsity matroid.

Proof outline:

1. We define a new class of DR-plans.
2. We show it has fan-in no larger than a canonical DR-plan. (Non-trivial proof.)
3. We show how to build it in time $O(|V|^3)$.

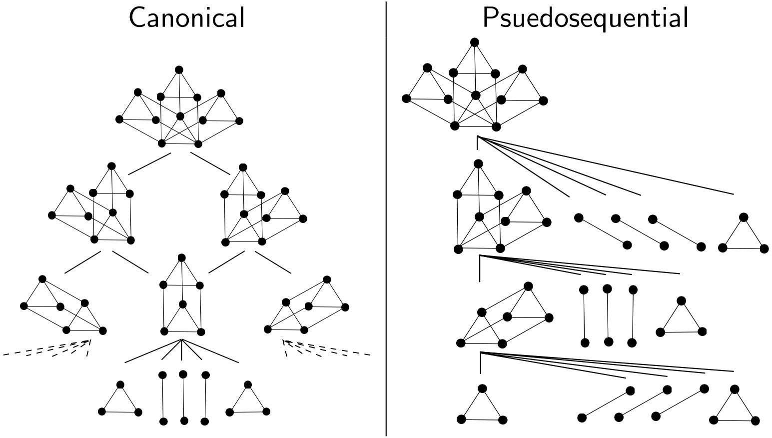
Pseudosequential DR-plan

Definition (Pseudosequential DR-plan)

A DR-plan where, if all pairs of rvmps of a node intersect

- Trivially: then all of them are children.
- Non-trivially: then exactly two that intersect non-trivially, C_1 and C_2, are used to find the children; they are C_1 and the pseudosequential DR-plan of $C_2 \setminus C_1$.

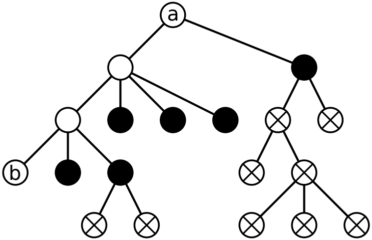
Example DR-Plans: Canonical vs. Pseudosequential



Branches

Definition (Branch)

Branch(T, a, b) of tree T is every node on the path from a to b and their children.



A Pseudosequential DR-Plan Branch from the Leaves

Given G and $e \in G$, there exists a pseudosequential DR-plan P_G where the leaves of branch(P_G, G, e) is exactly* the rvmps of $G \setminus e$.

A Pseudosequential DR-Plan Branch from the Leaves

Given G and $e \in G$, there exists a pseudosequential DR-plan P_G where the leaves of branch(P_G, G, e) is exactly* the rvmps of $G \setminus e$.

We can find the rvmps of $G \setminus e$ in $O(|V|^2)$.

A Pseudosequential DR-Plan Branch from the Leaves

Given G and $e \in G$, there exists a pseudosequential DR-plan P_G where the leaves of branch(P_G, G, e) is exactly* the rvmps of $G \setminus e$.

We can find the rvmps of $G \setminus e$ in $O(|V|^2)$.

Given a preprocessing step of finding the rvmps of $G \setminus f$ for all f, branch(P_G, G, e) can be built in time $O(|V|^2)$.

A Pseudosequential DR-Plan Branch from the Leaves

Given G and $e \in G$, there exists a pseudosequential DR-plan P_G where the leaves of branch(P_G, G, e) is exactly* the rvmps of $G \setminus e$.

We can find the rvmps of $G \setminus e$ in $O(|V|^2)$.

Given a preprocessing step of finding the rvmps of $G \setminus f$ for all f, branch(P_G, G, e) can be built in time $O(|V|^2)$.

Building the branch (from G to e):

1. Compute the rvmps of $G \setminus e$, $\{L_i\}$.
2. For each $L \in \{L_i\}$
   1. Choose an arbitrary edge $f \in L$ and compute the rvmps of $G \setminus f$, $\{M_i\}$.
   2. Compare the intersection of L with each M_i to get its position relative to the other leaves.
3. Compute nodes on the path from G to e.

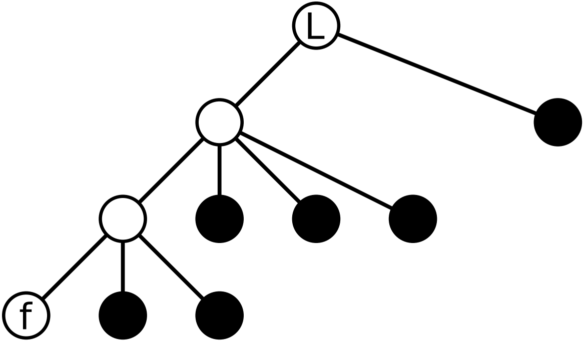
Finding an Entire Pseudosequential DR-Plan

Building the DR-plan (of G):

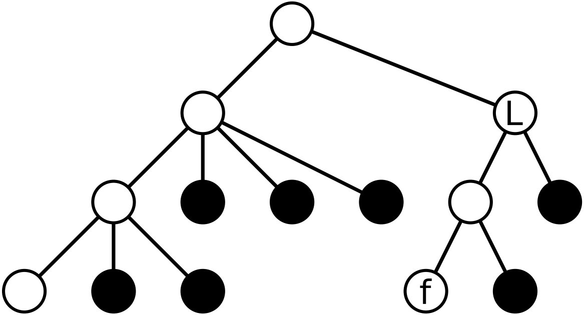
1. Preprocessing: Compute the rvmps of $G \setminus e$, for all e.
2. Start with G as the single node in the DR-plan.
3. Recursively compute a branch for each leaf in the DR-plan.

Algorithm Demonstration

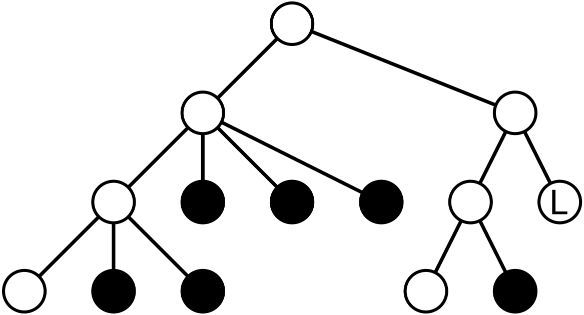
Algorithm Demonstration



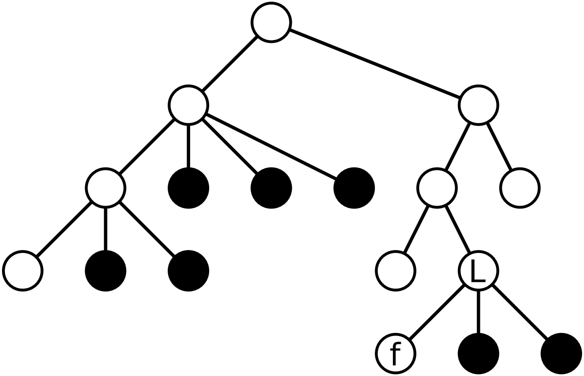
Algorithm Demonstration



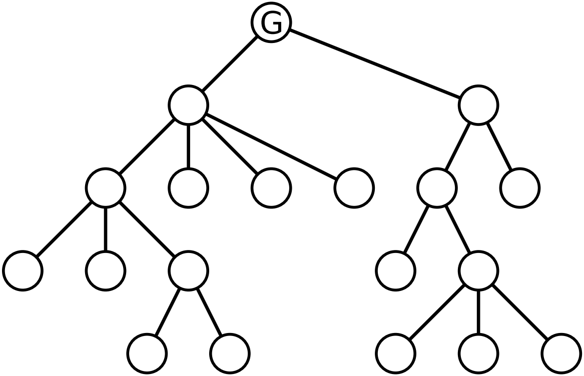
Algorithm Demonstration



Algorithm Demonstration



Algorithm Demonstration



Finding an Entire Pseudosequential DR-Plan

Building the DR-plan (of G):

1. Preprocessing: Compute the rvmps of $G \setminus e$, for all e.
2. Start with G as the single node in the DR-plan.
3. Recursively compute a branch for each leaf in the DR-plan.

Outline

What Next?

What comes after optimal DR-planning? We've decomposed to the extent possible.

Naïvely, we would, bottom-up, recombine the solved children into parents.
Recombining is equivalent to solving an indecomposable system.

OMD

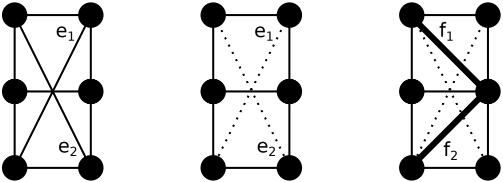
Definition (Optimal modification for decomposition (OMD))

Informally, drop some edges and add some others to make
–an easily realizable system (i.e. small max fan-in DR-plan),
–easy to search for lengths of added edges (*Cayley parameters*) that meet desired lengths of dropped edges.

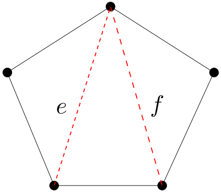
Dropped edges: Chosen so that the realization space has a *convex Cayley parameterization*.

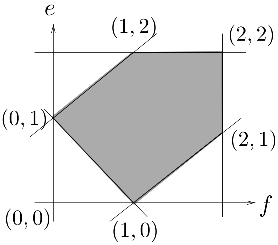
Added edges: Cayley parameters that convexify the realization space. Additionally, realization of modified linkage can be efficiently computed (e.g., triangle-decomposable.)

Convexity permits efficient search the realization space of modified linkage for realizations that satisfy the dropped constraints.



Convex Cayley Characterization





[SitharamGaoDCG2010], [SitharamWilloughbyADG2015]:
–Characterizes graphs that have (Strong/Weak) Convex Cayley Parameterization in dimension d
–Strong: Directly equivalent to *d-flattenability*, i.e., gram dimension $\leq d$.
–Results extend to linkages in other norms

d-flattenability

Definition

A graph G is d-flattenable under norm $||.||$ if for any m and any realization $r : V(G) \rightarrow \mathbf{R}^m$ there is a realization $r' : V(G) \rightarrow \mathbf{R}^d$ with $||r(u) - r(v)|| = ||r'(u) - r'(v)||$ for every $(u, v) \in E(G)$.

Analogous definition for flattenability of frameworks (G, r)

d-flattenability

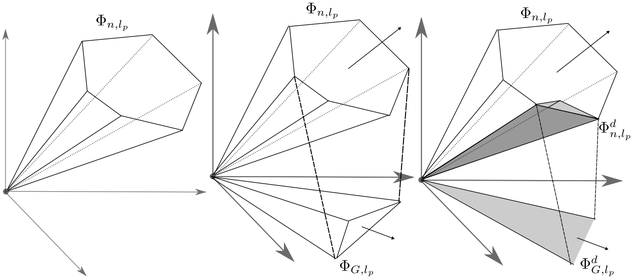
Observation

Let $\Phi_{l_p}(n)$ be the cone of vectors of $\binom{n}{2}$ pairwise l_p^p distances on n points. Let $\Phi_{l_p}^d(n)$ be the stratum of the cone consisting of those vectors when the points are in $\mathbf{R}^d$. Then G is d-flattenable if and only if both objects have the same projection on on the edge set G.

d-flattenability

Observation

Let $\Phi_{l_p}(n)$ be the cone of vectors of $\binom{n}{2}$ pairwise l_p^p distances on n points. Let $\Phi^d_{l_p}(n)$ be the stratum of the cone consisting of those vectors when the points are in $\mathbf{R}^d$. Then G is d-flattenable if and only if both objects have the same projection on on the edge set G.



Strong (Inherently) Convex Cayley Spaces

Definition

SitharamGaoDCG2010 A graph H has an *inherently convex Cayley space* in d-dimensions for a given norm l_q, $1 \leq q \leq \infty$, if the projection of $\Phi^d_{l_q}(n)$ on the edge set of H is convex. I.e., the space of realizable edge-length-vectors for H in d-dimensions and given norm, is convex.

Strong (Inherently) Convex Cayley Spaces

Definition

SitharamGaoDCG2010 A graph H has an *inherently convex Cayley space* in d-dimensions for a given norm l_q, $1 \leq q \leq \infty$, if the projection of $\Phi^d_{l_q}(n)$ on the edge set of H is convex. I.e., the space of realizable edge-length-vectors for H in d-dimensions and given norm, is convex.

Theorem

SitharamWilloughby2015 For any norm, a graph H has an inherently convex Cayley realization space in d dimensions if and only if H is d-flattenable.

Observation

- *For any norm and any dimension d, d-flattenability and Strong Cayley convexity in dimension d are minor-closed properties, with finite forbidden minor characterizations.*
- *Graphs of tree-width d (among others) have inherently convex Cayley configuration spaces.*

Applications

- Application (l_2, Inherently Convex, multiple Cayley parameters): EASAL for molecular/sticky-sphere assembly OzkanSitharamBiCoB2011; WuEtAlACMBCB2013,16; SitharamEtAl2015,16; OzkanEtAl2016A,B,C

Applications

- Application (l_2, Inherently Convex, multiple Cayley parameters): EASAL for molecular/sticky-sphere assembly OzkanSitharamBiCoB2011; WuEtAlACMBCB2013,16; SitharamEtAl2015,16; OzkanEtAl2016A,B,C
- Application (l_2, Non-Convex, single Cayley parameter): CayMos for CAD Mechanisms SitharamWangSPM2014, WangSitharamTOMS2015, SitharamWangGao2013a,b

EASAL – Virus Assembly

- Approximate computation of volume of potential energy basins → free energy change for <u>each node</u> of assembly tree
- Topology of configuration space → formation rates for <u>each node</u> of assembly tree
- <u>Likelihood</u> of each assembly tree

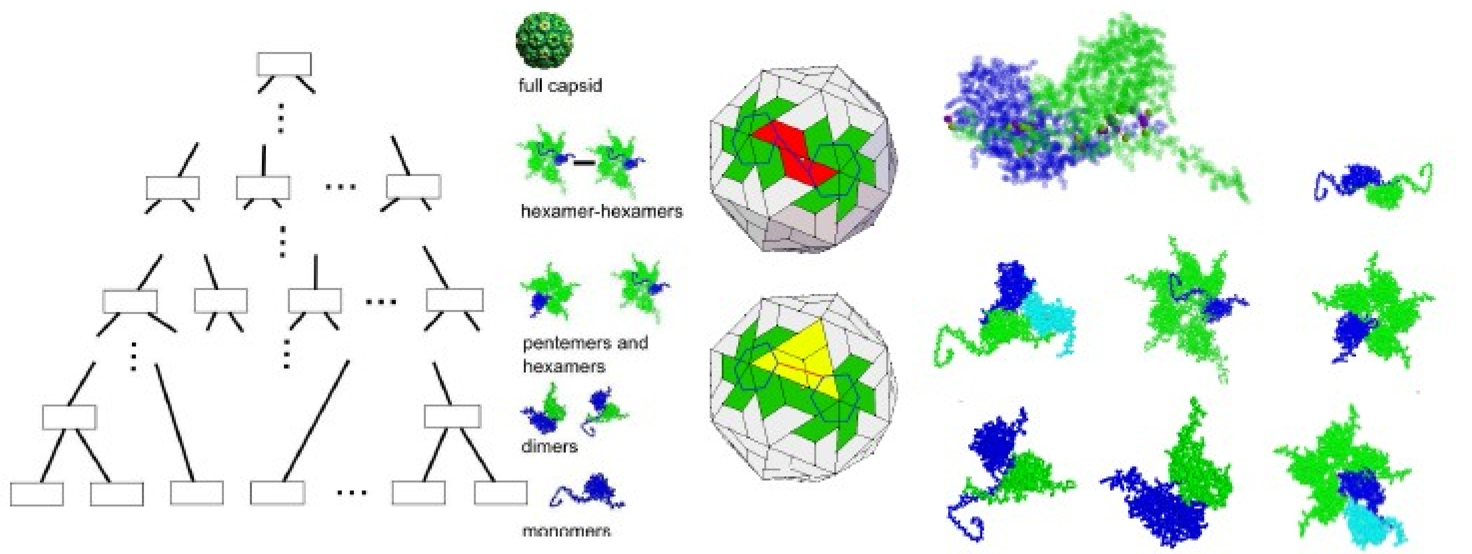
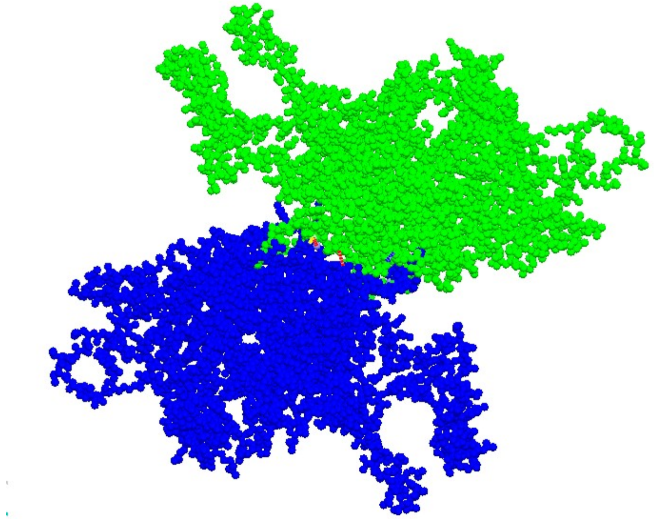
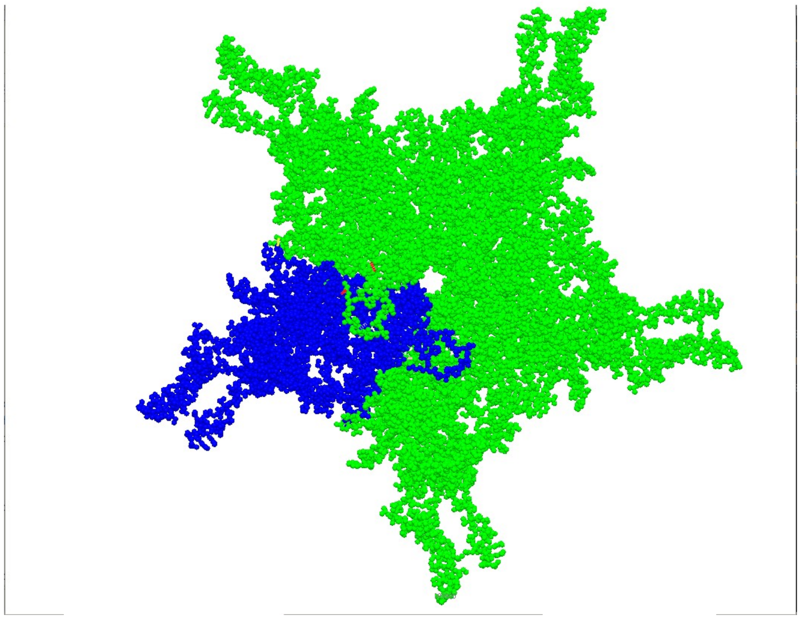


Figure 1: Left T=3 BMV assembly tree schematic; Mid: different bi-assembly interfaces (2-fold or 3-fold) for the same intermediate from the same pentamer subassemblies; Right top: 2-fold interface types, one magnified; mid: 3, 5 and 6 fold interfaces; bottom: bi-assemblies of monomers extracted from 5 and 6 fold and larger bi-assembly of small multimers from 5-fold at bottom right (the exhaustive list of such bi-assemblies of small multimers contains approx 24 for BMV)

Predicting crucial interactions for T=1,3 viral capsid shell assembly

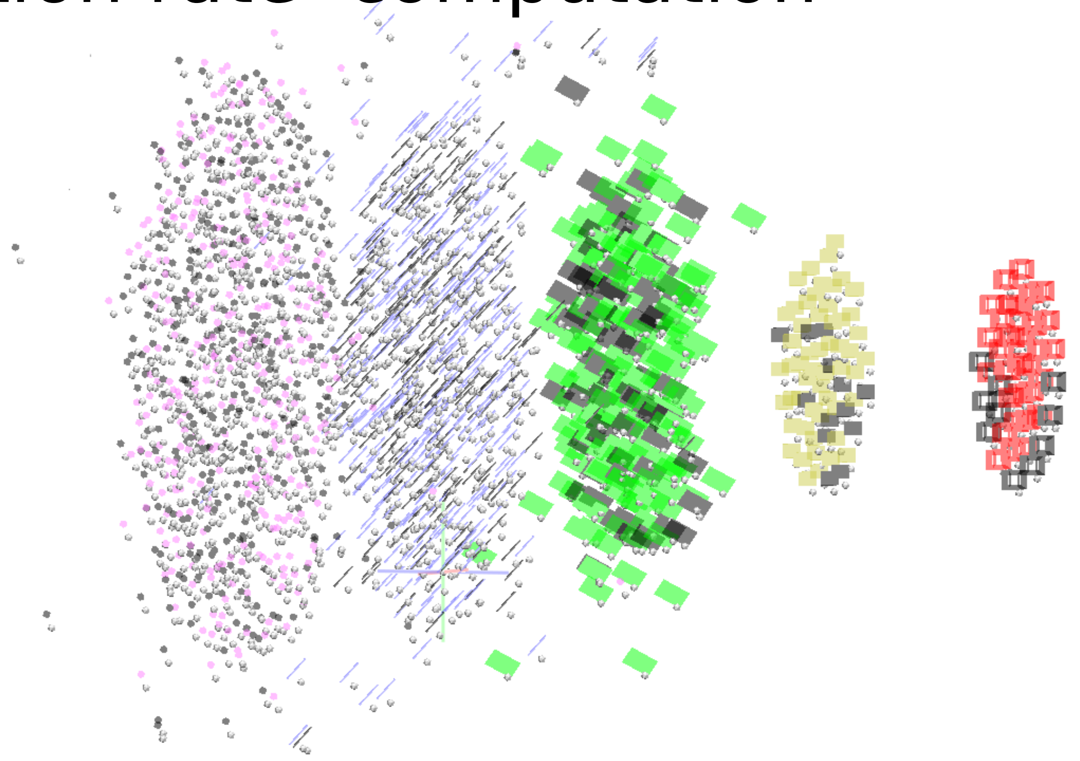


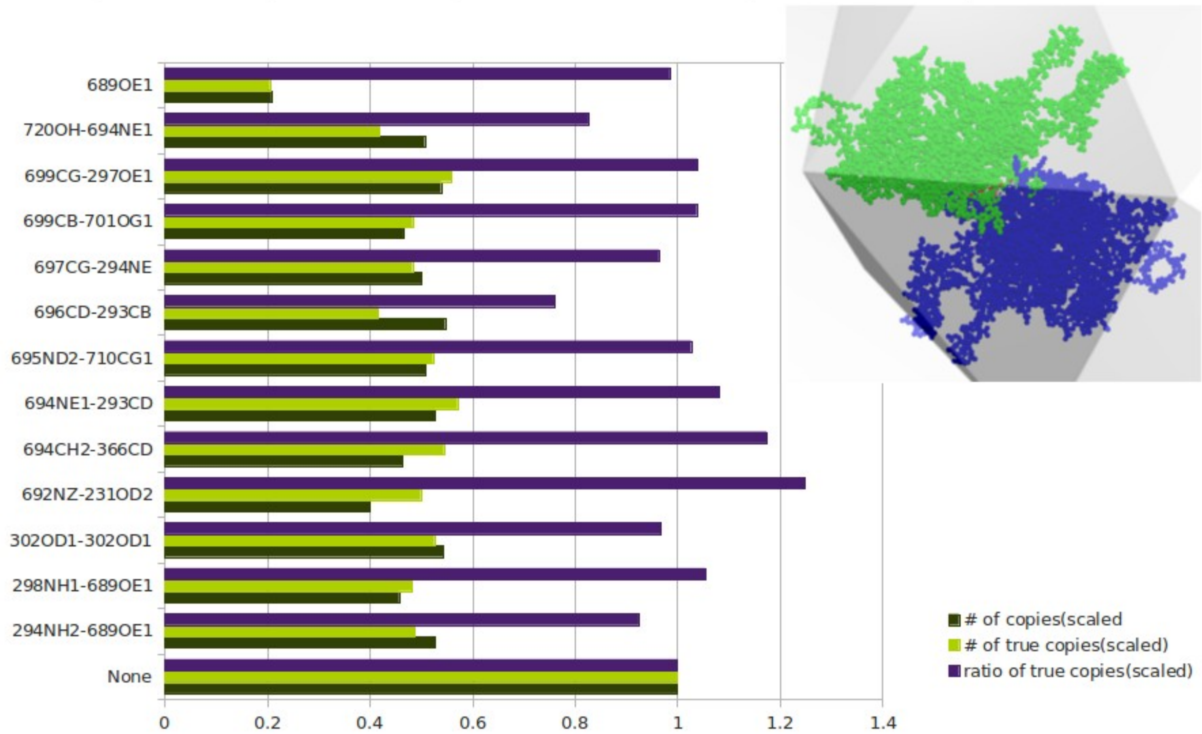


- predict minimal sets of geometric constraints whose removal disrupts assembly of viral shell.

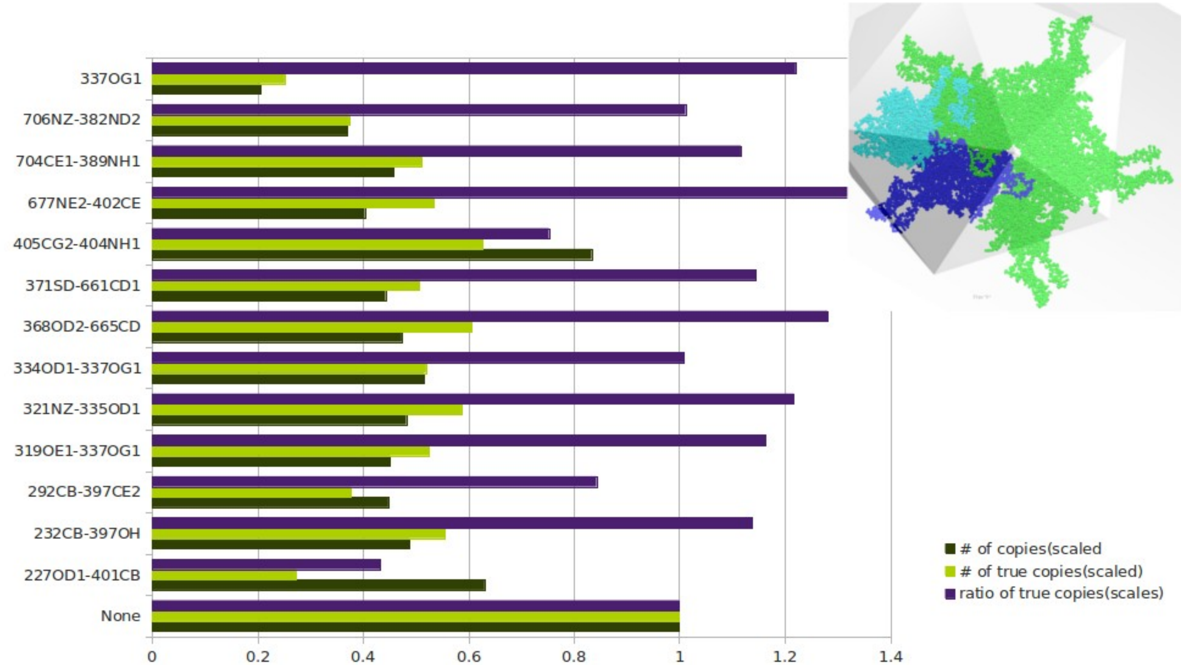
Predicting crucial interactions for T=1 viral shell assembly

- see how the atlas differs (in black) if a constraint is dropped.
- crucial constraints result in big changes in (*approximate*) configurational entropy + formation rate computation





- EASAL's prediction is confirmed by mutagenesis data from the *Agbandge-Mckenna lab at UF*



aav, dimer

Bond	Confirmed
D231 and K692	
P293 and W694	Yes
P293 and P696	Yes
R294 and E689	Yes
R294 and E697	Yes
R298 and E689	Yes
P366 and W694	Yes
Y720 and W694	Yes

mvm, dimer

Bond	Confirmed
F55 and L537	Yes
D127 and N540	
Q129 and V546	Yes
N133 and Q548	
D302 and T567	
D302 and T569	
D302 and N571	
N540 and S126	
P545 and W564	Yes
Y547 and S550	

bmv, dimer

Bond	Confirmed
I51and F184	Yes
Q172 and F184	Yes
D181 and D181	Yes

aav, pentamer

Bond	Confirmed
T337 and Q319	
T337 and N334	
N382 and K706	
R389 and Y704	
Y397 and S292	Yes
Q401 and N227	Yes
M402 and Q677	Yes

mvm, pentamer

Bond	Confirmed
Y47 and E251	
Y47 and I256	
Y47 and L258	
N149 and R260	
K153 and D171	Yes
K153 and N504	Yes
Y168 and T506	
Y168 and D507	
N170 and T173	

bmv, pentamer

Bond	Confirmed
P98 and T66	
P98 and E116	
P98 and F119	
P98 and Y157	Yes
S99 and E116	
E131 and K130	Yes
E131 and E131	

mvm, trimer

Bond	Confirmed
W283 and C216	Yes
W283 and N244	Yes
W283 and Q246	Yes
R287 and D101	Yes
Q291 and S194	Yes
Q 291 and S209	Yes
Q 291 and R212	Yes
R314 and D102	Yes
L453 and L481	Yes
N544 and F247	Yes
R584 and D474	Yes
R584 and E476	Yes

bmv, trimer3

Bond	Confirmed
E80 and E110	Yes
E80 and D148	
E80 and N151	
K81 and D139	Yes
E84 and T145	Yes
E84 and D148	Yes
Y188 and K130	Yes

bmv, hexamer

Bond	Confirmed
A25 and R26	
Q28 and Q28	
P29 and I31	
P98 and K105	

bmv, hex-pen

Bond	Confirmed
Q172 and F180	Yes
Q172 and T185	
V187 and K41	Yes
V187 and V132	Yes

bmv, trimer2

Bond	Confirmed
E110 and E80	
K130 and R189	Yes
D139 and F183	Yes
T145 and E84	Yes
D148 and E 84	Yes

EASAL – Virus Assembly

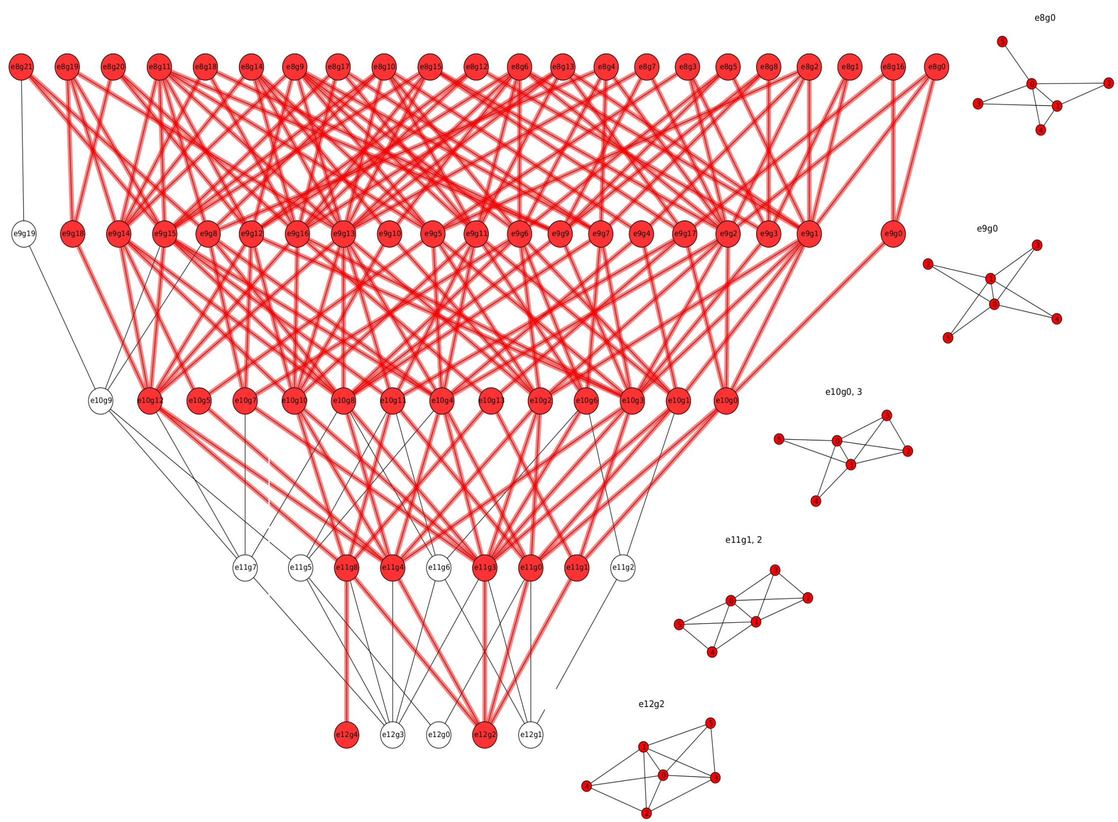
- Mysterious "Missing" factor: Combinatorial entropy – counting pathway symmetry equivalence classes (*)

- Sparse mutagenesis data – need to use kinetics, differential calorimetry and other combination of experimental data, including fine-grained MC/MD for cross-validation

(*) Bona, Sitharam, Vince "Tree orbits under permutation groups and application to virus assembly" Bulletin of Math Bio, 2011

EASAL – Sticky Spheres

- Complete computation of free energy and formation rates for 6,7,8 sticky sphere system



Opensource Software

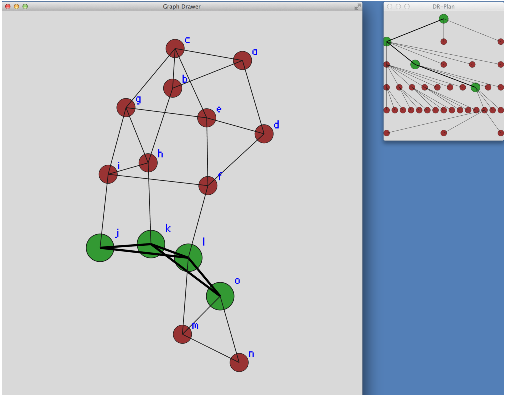
Available on my webpage.

Decomposition:

FRONTIER (GPL, bitbucket),

New version Under development Available at:

`cise.ufl.edu/~tbaker/drp`



More Opensource Software

Cayley Configuration spaces:
CayMos (for 2D mechanisms) (GPL, bitbucket)
EASAL (for molecular and sticky sphere assembly)