(Non)-Degenerate Ground States in the H–P model

David Bremner

University of New Brunswick

Outline

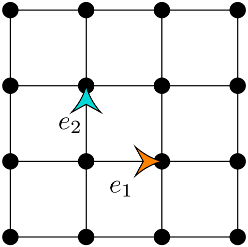
Outline

Lattices

Lattice

Given linearly independant vectors $B = \{ b_1 \dots b_d \}$ in $\mathbb{R}^d$, the *lattice*

$$L(B) := \{ \sum_{i=1}^{d} z_i b_i \mid z_i \in \mathbb{Z} \}$$



Lattice Graphs

Given l.i. $B \subset \mathbb{R}^d$, the *lattice graph*

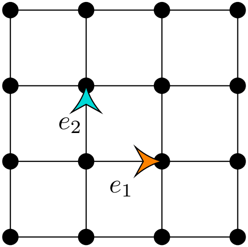
$$G(B) = (L(B), \{ (p, q) : \|p - q\| = 1 \})$$

Lattices

Lattice

Given linearly independant vectors $B = \{ b_1 \dots b_d \}$ in $\mathbb{R}^d$, the *lattice*

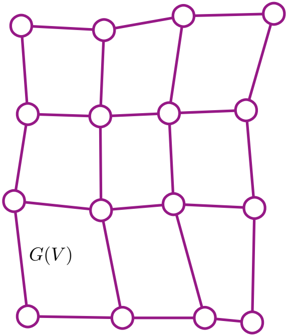
$$L(B) := \{ \sum_{i=1}^{d} z_i b_i \mid z_i \in \mathbb{Z} \}$$



Lattice Graphs

Given l.i. $B \subset \mathbb{R}^d$, the *lattice graph*

$$G(B) = (L(B), \{ (p, q) : \|p-q\| = 1 \})$$

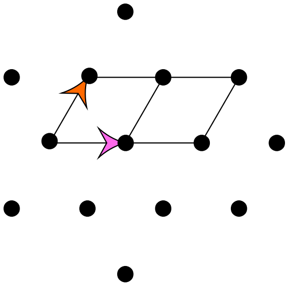


Lattices

Lattice

Given linearly independant vectors $B = \{ b_1 \ldots b_d \}$ in $\mathbb{R}^d$, the *lattice*

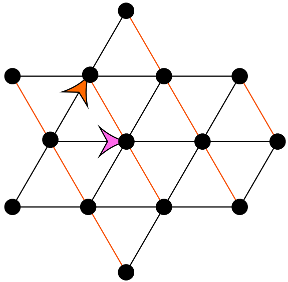
$$L(B) := \{ \sum_{i=1}^{d} z_i b_i \mid z_i \in \mathbb{Z} \}$$



Lattice Graphs

Given l.i. $B \subset \mathbb{R}^d$, the *lattice graph*

$$G(B) = (L(B), \{ (p,q) : \|p - q\| = 1 \})$$



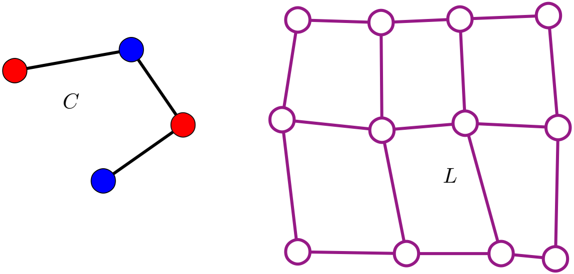
Lattice Models

Combinatorial Setting

Polymer A chain C (node sequence) with coloured (classified) nodes.

Lattice A vertex regular (sufficiently large) graph L.

Folding An embedding of C into L.



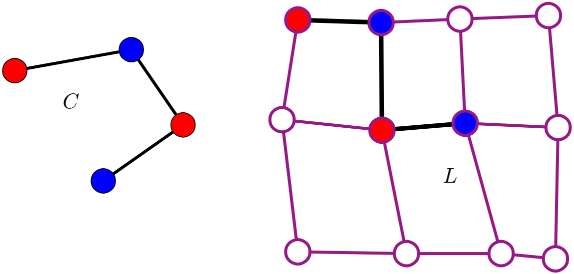
Lattice Models

Combinatorial Setting

Polymer A chain C (node sequence) with coloured (classified) nodes.

Lattice A vertex regular (sufficiently large) graph L.

Folding An embedding of C into L.



Lattice Models

Combinatorial Setting

Polymer A chain C (node sequence) with coloured (classified) nodes.

Lattice A vertex regular (sufficiently large) graph L.

Folding An embedding of C into L.

Energy Model

Energy ϕ : embedding $\rightarrow \mathbb{R}$

Locality ϕ is usually a function of a small neighbourhood in L.

Optimality Minimum energy embeddings (*ground states*) are considered optimal.

Lattice Models: pro and contra

Pro

Physics is hard Global optimization models have $\Omega(3^n)$ local optima.

Chemistry is lattice-like Close packed proteins are crystal-like.

Thought Experiment Can a small subset of forces explain folding?

Contra

Discrete optimization is hard Computing optimal embeddings is NP-hard.

Approximation is rough Close energy $\not\Leftrightarrow$ close shape?

Lattice artifacts Crude approximation of shape. Parity. Chirality.

Lattice Models: pro and contra

Pro

Physics is hard Global optimization models have $\Omega(3^n)$ local optima.

Chemistry is lattice-like Close packed proteins are crystal-like.

Thought Experiment Can a small subset of forces explain folding?

Contra

Discrete optimization is hard Computing optimal embeddings is NP-hard.

Approximation is rough Close energy $\not\Leftrightarrow$ close shape?

Lattice artifacts Crude approximation of shape. Parity. Chirality.

What Lattice, What Energy function?

Lattice

2D square
: Some interest for ≤ 30 monomers

3D cubic
: Basic local structures (helix) are 3D.

2D triangular
: Solve parity problems

Energy Function

- Hydrophobic/Hydrophilic forces by far strongest
- Helical structures can be designed by using only hydrophobicity,
- β-sheets have few local interactions

What Lattice, What Energy function?

Lattice

2D square: Some interest for ≤ 30 monomers

3D cubic: Basic local structures (helix) are 3D.

2D triangular: Solve parity problems

Energy Function

- Hydrophobic/Hydrophilic forces by far strongest
- Helical structures can be designed by using only hydrophobicity,
- β-sheets have few local interactions

H–P model

Hydrophobic/Hydrophilic

- Hydrophobic (H) repels water
- Polar (Hydrophilic) (P) attracts water
- Model: H's attract each other and P's are neutral

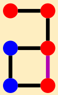
Amino Acid	Code	Classification
Leucine	L	H
Serine	S	P
Glycine	G	H
Threonine	T	P
⋮	⋮	⋮

H–P model

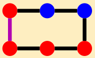
Hydrophobic/Hydrophilic

- Hydrophobic (H) repels water
- Polar (Hydrophilic) (P) attracts water
- Model: H's attract each other and P's are neutral

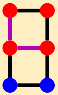
H–H contacts



1 contact



1 contact



2 contacts

optimal Maximum number of contacts

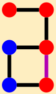
stable/degenerate optimal embedding

H–P model

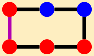
Hydrophobic/Hydrophilic

- Hydrophobic (H) repels water
- Polar (Hydrophilic) (P) attracts water
- Model: H's attract each other and P's are neutral

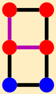
H–H contacts



1 contact



1 contact



2 contacts

optimal Maximum number of contacts

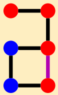
stable/nondegenerate Unique optimal embedding

H–P model

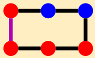
Hydrophobic/Hydrophilic

- Hydrophobic (H) repels water
- Polar (Hydrophilic) (P) attracts water
- Model: H's attract each other and P's are neutral

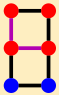
H–H contacts



1 contact



1 contact



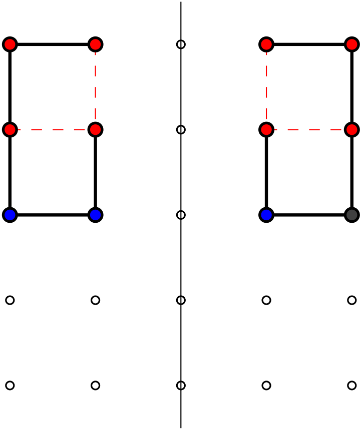
2 contacts

optimal Maximum number of contacts

unstable/degenerate Many optimal embeddings

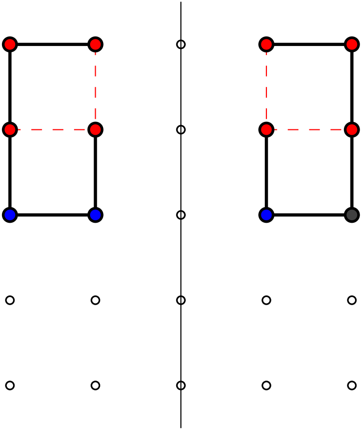
What counts as "unique"?

- Most lattices have *isometries*, i.e. distance preserving transformations.
- Isometries preserve contacts.
- Counting contacts is oblivious to *chirality*



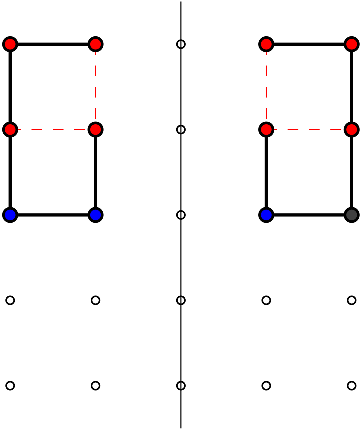
What counts as "unique"?

- Most lattices have *isometries*, i.e. distance preserving transformations.
- Isometries preserve contacts.
- Counting contacts is oblivious to *chirality*



What counts as "unique"?

- Most lattices have *isometries*, i.e. distance preserving transformations.
- Isometries preserve contacts.
- Counting contacts is oblivious to *chirality*



Computational Complexity of the H–P model

Algorithmic results

- NP-Complete for 3D (Berger & Leighton 1998)
- NP-Complete for 2D (Crescenzi et al., JCB 1998)
- 3/8-approximation for 3D and 1/4-approximation for 2D (Hart and Istrail, STOC 1995).

Fight hardness with more restricted problem?

- *H*-connected optimal embedding.
  - 3D NP-hard gadgets have this property
  - 2D gadgets do not
- Unique optimal embeddings
  - neither 2D nor 3D NP-hard gadgets are stable

Computational Complexity of the H–P model

Algorithmic results

- NP-Complete for 3D (Berger & Leighton 1998)
- NP-Complete for 2D (Crescenzi et al., JCB 1998)
- 3/8-approximation for 3D and 1/4-approximation for 2D (Hart and Istrail, STOC 1995).

Fight hardness with more restricted problem?

- *H*-connected optimal embedding.
  - 3D NP-hard gadgets have this property
  - 2D gadgets do not
- Unique optimal embeddings
  - neither 2D nor 3D NP-hard gadgets are stable

The Protein Folding "Paradoxes"

Protein Folding Paradox (Levinthal 1968)

There are an exponential number of foldings ("conformations"), but proteins fold quickly.

New Improved Protein Folding Paradox (1998)

Finding the optimal folding in the H–P model is NP-complete, but proteins still fold quickly.

The Protein Folding "Paradoxes"

Protein Folding Paradox (Levinthal 1968)

There are an exponential number of foldings ("conformations"), but proteins fold quickly.

New Improved Protein Folding Paradox (1998)

Finding the optimal folding in the H–P model is NP-complete, but proteins still fold quickly.

Why care about uniqueness?

Motivations

- An important property of real proteins
- Possible resolution to NP-hardness "paradox".
- "Sequence design: the hard part is uniqueness" (Dill et al., 1995)

Evidence

Experimental designed polymers have many optimal foldings

Algorithmic designing to fold to a shape is easy. (Kleinberg 1999)

Simulation machine designed H–P-polymers tend to collapse below design state (Yue et al. 1995)

Why care about uniqueness?

Motivations

- An important property of real proteins
- Possible resolution to NP-hardness "paradox".
- "Sequence design: the hard part is uniqueness" (Dill et al., 1995)

Evidence

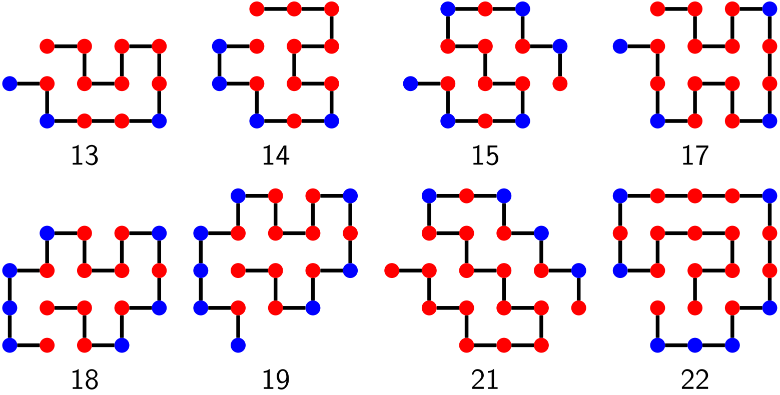
Experimental designed polymers have many optimal foldings

Algorithmic designing to fold to a shape is easy. (Kleinberg 1999)

Simulation machine designed H–P-polymers tend to collapse below design state (Yue et al. 1995)

Simulation Results

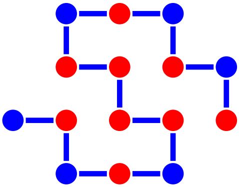
- About 2% of sequences up to length 18 length have unique optimal foldings



Outline

Terminology

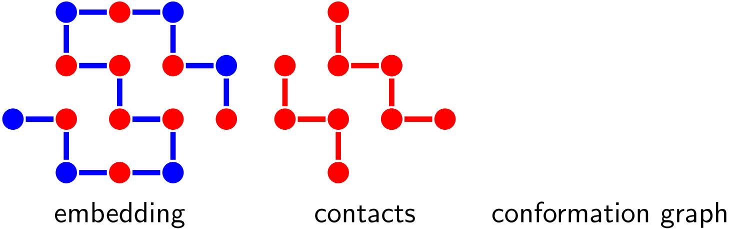
- A pair of H nodes adjacent in an embedding, but not on the chain P, is called a *contact*
- *contact graph* $V =$ H nodes; $E =$ contacts
- The *conformation graph* consists of the edges of polymer P, along with the contacts.



embedding contacts conformation graph

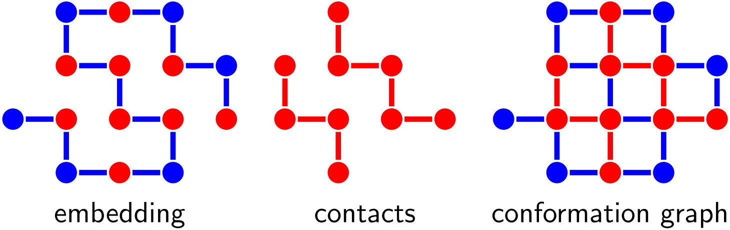
Terminology

- A pair of H nodes adjacent in an embedding, but not on the chain P, is called a *contact*
- *contact graph* $V =$ H nodes; $E =$ contacts
- The *conformation graph* consists of the edges of polymer P, along with the contacts.



Terminology

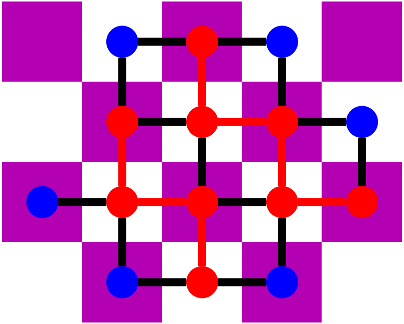
- A pair of H nodes adjacent in an embedding, but not on the chain P, is called a *contact*
- *contact graph* $V =$ H nodes; $E =$ contacts
- The *conformation graph* consists of the edges of polymer P, along with the contacts.



Parity and Lattice Graphs

Parity

Define the *parity* of lattice point $\sum z_i b_i$ as $\sum z_i \bmod 2$.



Bipartite Lattice Graphs

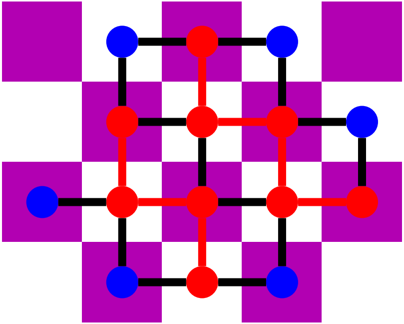
In the square and cubic lattice graphs:

- Every edge changes parity.
- Contacts exist only between H nodes of different parity.
- No odd cycles are possible.
- The maximal contact graph of a closed chain consists of disjoint even cycles.

Parity and Lattice Graphs

Parity

Define the *parity* of lattice point $\sum z_i b_i$ as $\sum z_i \bmod 2$.



Bipartite Lattice Graphs

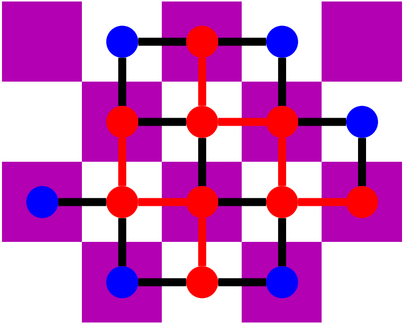
In the square and cubic lattice graphs:

- Every edge changes parity.
- Contacts exist only between H nodes of different parity.
- No odd cycles are possible.
- The maximal contact graph of a closed chain consists of disjoint even cycles.

Parity and Lattice Graphs

Parity

Define the *parity* of lattice point $\sum z_i b_i$ as $\sum z_i \bmod 2$.



Bipartite Lattice Graphs

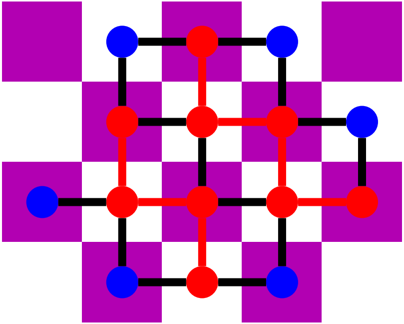
In the square and cubic lattice graphs:

- Every edge changes parity.
- Contacts exist only between H nodes of different parity.
- No odd cycles are possible.
- The maximal contact graph of a closed chain consists of disjoint even cycles.

Parity and Lattice Graphs

Parity

Define the *parity* of lattice point $\sum z_i b_i$ as $\sum z_i \bmod 2$.



Bipartite Lattice Graphs

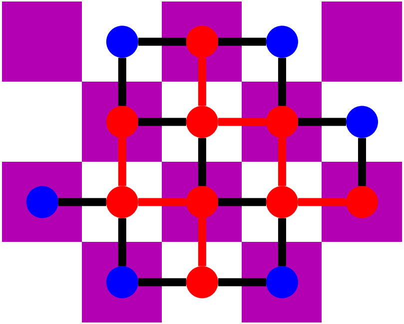
In the square and cubic lattice graphs:

- Every edge changes parity.
- Contacts exist only between H nodes of different parity.
- No odd cycles are possible.
- The maximal contact graph of a closed chain consists of disjoint even cycles.

Parity and Lattice Graphs

Parity

Define the *parity* of lattice point $\sum z_i b_i$ as $\sum z_i$ mod 2.



Bipartite Lattice Graphs

In the square and cubic lattice graphs:

- Every edge changes parity.
- Contacts exist only between H nodes of different parity.
- No odd cycles are possible.
- The maximal contact graph of a closed chain consists of disjoint even cycles.

Outline

Degenerate Ground States

Setting

- 2D square lattice
- open or closed chains
- Degenerate ground state ≡ many optimal embeddings

Fact

Any embedding of P^k is optimal

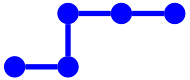
Degenerate Ground States

Setting

- 2D square lattice
- open or closed chains
- Degenerate ground state ≡ many optimal embeddings

Fact

Any embedding of P^k *is optimal*



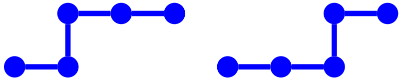
Degenerate Ground States

Setting

- 2D square lattice
- open or closed chains
- Degenerate ground state $\equiv$ many optimal embeddings

Fact

Any embedding of P^k *is optimal*



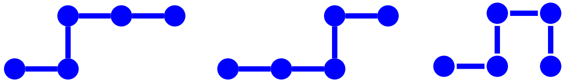
Degenerate Ground States

Setting

- 2D square lattice
- open or closed chains
- Degenerate ground state ≡ many optimal embeddings

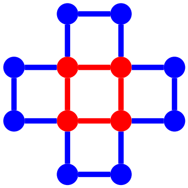
Fact

Any embedding of P^k *is optimal*



Fact

Any optimal embedding of the closed chain $(\text{PHP})^{4k}$ has a contact graph consisting of k four cycles.



Skip proof

Proof.

Consider a big contact graph cycle...

Fact

Any optimal embedding of the closed chain $(\mathrm{PHP})^{4k}$ *has a contact graph consisting of* k *four cycles.*

Proof.

Consider a big contact graph cycle...

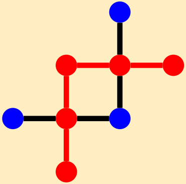
□

Fact

Any optimal embedding of the closed chain $(\mathrm{PHP})^{4k}$ *has a contact graph consisting of k four cycles.*

Proof.

Consider a big contact graph cycle...



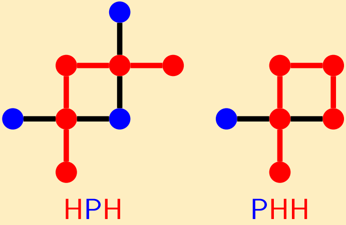
□

Fact

Any optimal embedding of the closed chain $(\mathrm{PHP})^{4k}$ *has a contact graph consisting of* k *four cycles.*

Proof.

Consider a big contact graph cycle. . .



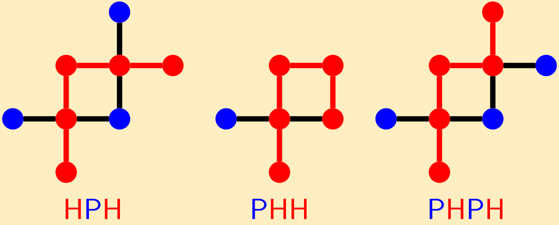
□

Fact

Any optimal embedding of the closed chain $(\mathrm{PHP})^{4k}$ *has a contact graph consisting of* k *four cycles.*

Proof.

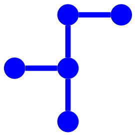
Consider a big contact graph cycle...



□

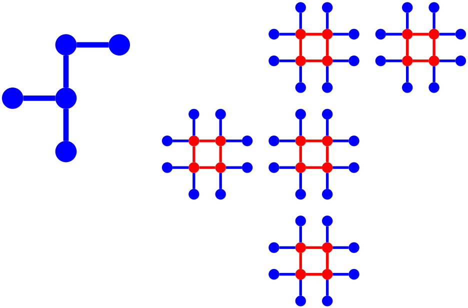
Fact

There are as many optimal embeddings of $(\mathrm{PHP})^{4k}$ *as there are (embeddings of) k-node lattice trees.*



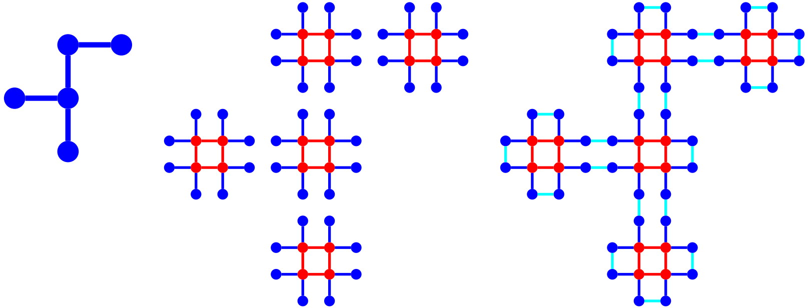
Fact

There are as many optimal embeddings of $(\text{PHP})^{4k}$ *as there are (embeddings of) k-node lattice trees.*



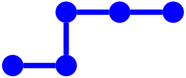
Fact

There are as many optimal embeddings of (PHP)^{4k} *as there are (embeddings of) k-node lattice trees.*



Fact (Stanley)

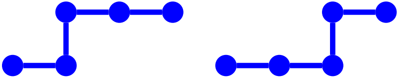
There are $\Omega((1+\sqrt{2})^k)$ embeddings of k-node lattice trees.



- and probably lots more ($\Omega\left(\frac{3.79^k}{k}\right)$)

Fact (Stanley)

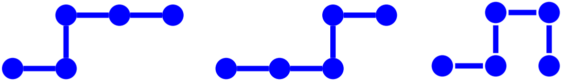
There are $\Omega((1+\sqrt{2})^k)$ *embeddings of* k*-node lattice trees.*



- and probably lots more ($\Omega\left(\frac{3.79^k}{k}\right)$)

Fact (Stanley)

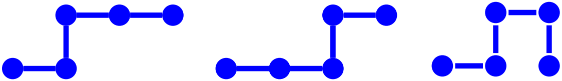
There are $\Omega((1+\sqrt{2})^k)$ *embeddings of* k*-node lattice trees.*



- and probably lots more ($\Omega\left(\frac{3.79^k}{k}\right)$)

Fact (Stanley)

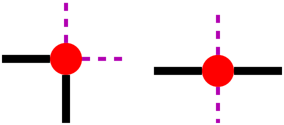
There are $\Omega((1+\sqrt{2})^k)$ embeddings of k-node lattice trees.



- and probably lots more $(\Omega\left(\frac{3.79^k}{k}\right))$

Outline

Missing Contacts and Perimeter



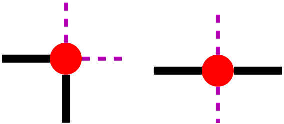
Missing Contact Neighbouring lattice point that is neither

- An H node
- nor a P node adjacent on chain.

Pseudocontacts H-H neighbours (on chain or otherwise). For a given chain, maximizing contacts is equivalent to

- maximizing pseudocontacts, or
- minimizing missing contacts.

Missing Contacts and Perimeter



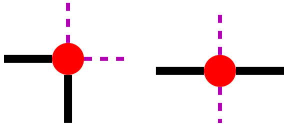
Missing Contact Neighbouring lattice point that is neither

- An H node
- nor a P node adjacent on chain.

Pseudocontacts H-H neighbours (on chain or otherwise). For a given chain, maximizing contacts is equivalent to

- maximizing pseudocontacts, or
- minimizing missing contacts.

Missing Contacts and Perimeter



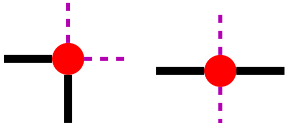
Missing Contact Neighbouring lattice point that is neither

- An H node
- nor a P node adjacent on chain.

Pseudocontacts H-H neighbours (on chain or otherwise). For a given chain, maximizing contacts is equivalent to

- maximizing pseudocontacts, or
- minimizing missing contacts.

Missing Contacts and Perimeter



Missing Contact Neighbouring lattice point that is neither

- An H node
- nor a P node adjacent on chain.

Pseudocontacts H-H neighbours (on chain or otherwise). For a given chain, maximizing contacts is equivalent to

- maximizing pseudocontacts, or
- minimizing missing contacts.

Square Contact Graphs

Fact

A chain with s^2 H nodes embedded in the square lattice has at most $2s^2 - 2s$ pseudocontacts, and this is achieved exactly when H nodes are embedded in a $s \times s$ square grid.

$$\min 2x + 2y$$

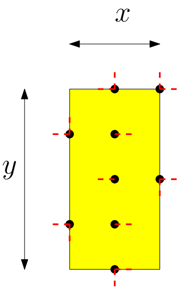
subject to

$$x \cdot y \geq s^2$$
$$x \geq 1$$
$$y \geq 1$$

Square Contact Graphs

Fact

A chain with s^2 H nodes embedded in the square lattice has at most $2s^2 - 2s$ pseudocontacts, and this is achieved exactly when H nodes are embedded in a $s \times s$ square grid.



$$\min 2x + 2y$$

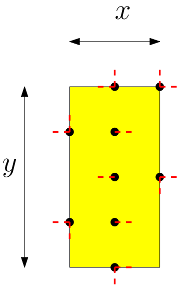
subject to

$$x \cdot y \geq s^2$$
$$x \geq 1$$
$$y \geq 1$$

Square Contact Graphs

Fact

A chain with s^2 H nodes embedded in the square lattice has at most $2s^2 - 2s$ pseudocontacts, and this is achieved exactly when H nodes are embedded in a $s \times s$ square grid.

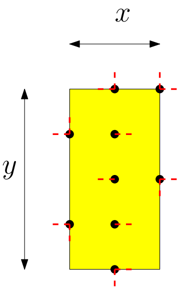


$$
\begin{aligned}
\min\ & 2x + 2y \\
\text{subject to}\quad & \\
& x \cdot y \geq s^2 \\
& x \geq 1 \\
& y \geq 1
\end{aligned}
$$

Square Contact Graphs

Fact

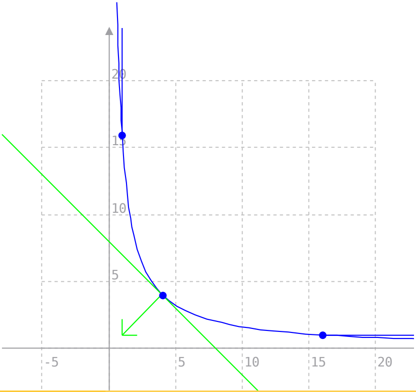
A chain with s^2 H nodes embedded in the square lattice has at most $2s^2 - 2s$ pseudocontacts, and this is achieved exactly when H nodes are embedded in a $s \times s$ square grid.



$$\min 2x + 2y$$

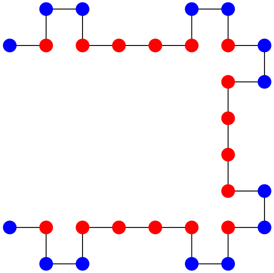
subject to

$$x \cdot y \geq s^2$$
$$x \geq 1$$
$$y \geq 1$$



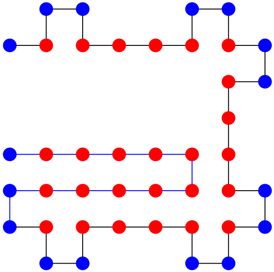
Uniquely achieving the $s \times s$ square

- Start by fixing the corners
- Make short loops of H nodes
- Repeat



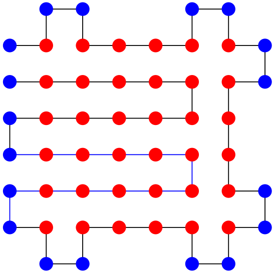
Uniquely achieving the $s \times s$ square

- Start by fixing the corners
- Make short loops of H nodes
- Repeat



Uniquely achieving the $s \times s$ square

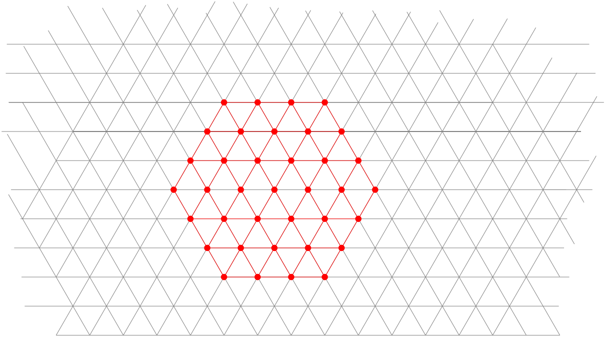
- Start by fixing the corners
- Make short loops of H nodes
- Repeat



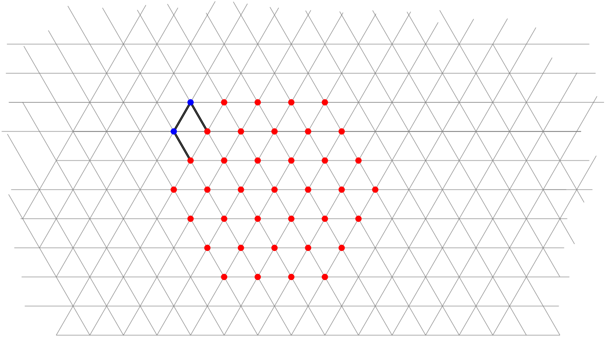
Hexagonal Contact Graphs

Fact

A chain $3s^2 - 3s + 1$ H nodes embedded in the triangular lattice has at most $9s^2 - 15s + 6$ pseudocontacts, and this is achieved exactly when the H nodes are embedded in a side-length s hexagonal grid.

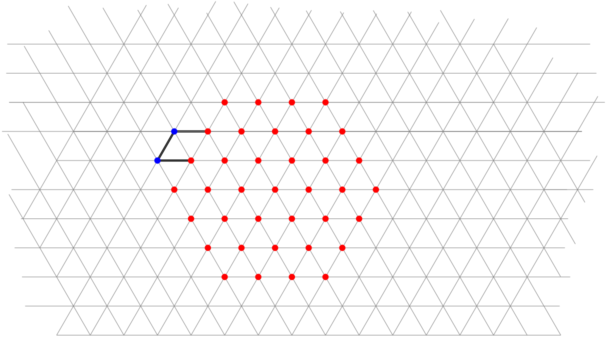


Uniquely realizing the hexagon



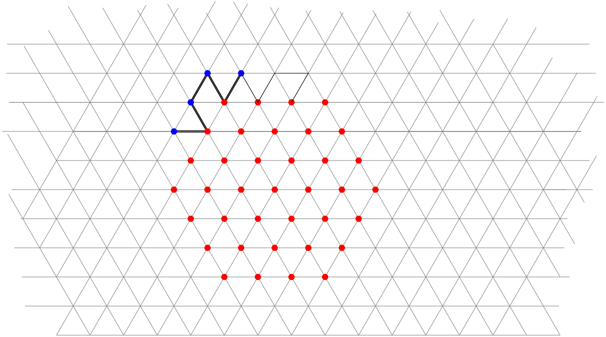
- Sides can be fixed
- but with wiggle
- Wiggle can be fixed
- And the interior filled

Uniquely realizing the hexagon



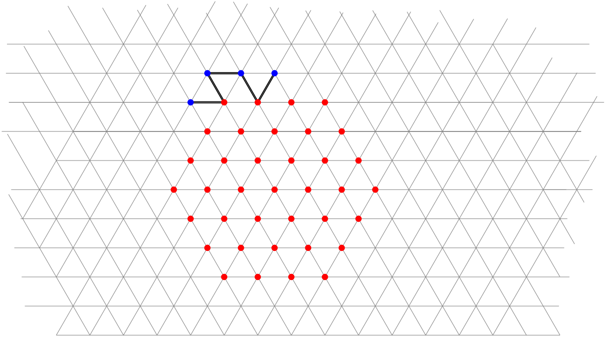
- Sides can be fixed
- but with wiggle
- Wiggle can be fixed
- And the interior filled

Uniquely realizing the hexagon



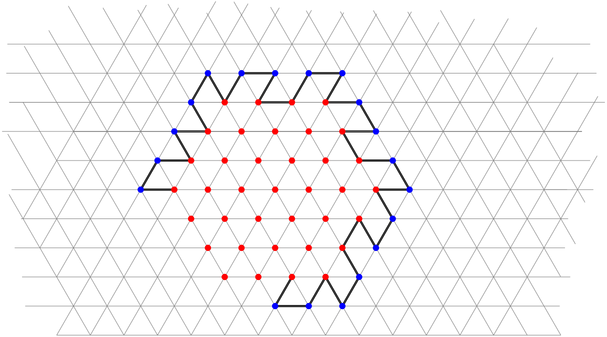
- Corners can be fixed
- but with wiggle
- Wiggle can be fixed
- And the interior filled

Uniquely realizing the hexagon



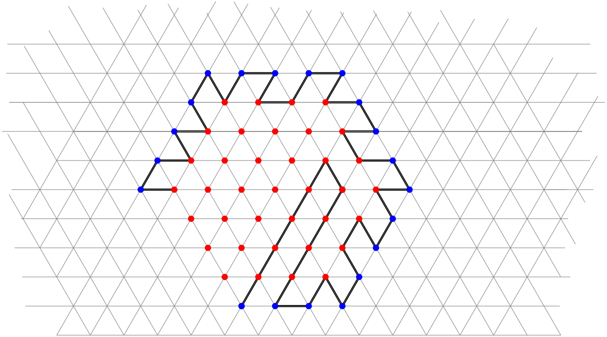
- Corners can be fixed
- but with wiggle
- Wiggle can be fixed
- And the interior filled

Uniquely realizing the hexagon



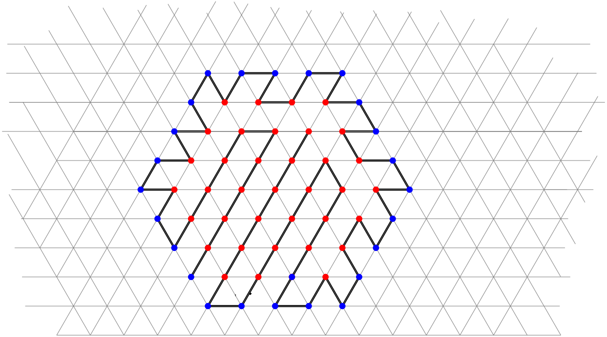
- can be fixed
- but with wiggle
- Wiggle can be fixed
- And the interior filled

Uniquely realizing the hexagon



- can be fixed
- but with wiggle
- Wiggle can be fixed
- And the interior filled

Uniquely realizing the hexagon



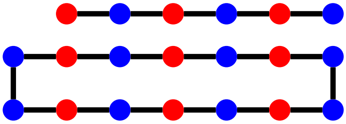
- can be fixed
- but with wiggle
- Wiggle can be fixed
- And the interior filled

Outline

Closed Chain Examples

$$A_m = (\text{HP})^m$$

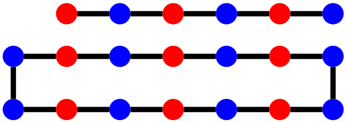
$$S_k = \text{P}\ A_{\lceil k/2 \rceil}\ \text{P}\ A_{\lfloor k/2 \rfloor}$$

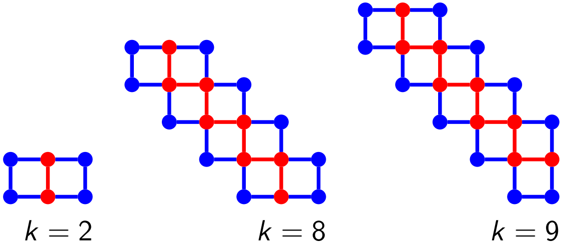


Closed Chain Examples

$A_m = (\text{HP})^m$

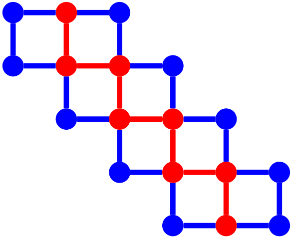
$S_k = \text{P}\ A_{\lceil k/2 \rceil}\ \text{P}\ A_{\lfloor k/2 \rfloor}$





Observation

There exists an embedding of S_k with 2 missing contacts.



Corollary

In any optimal embedding of S_k, both monochrome edges are on the bounding box.

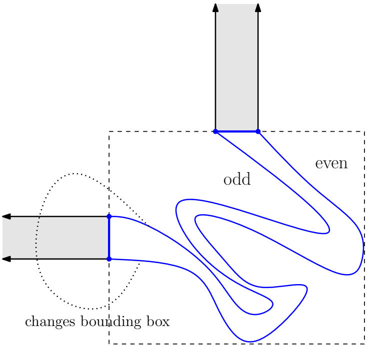
Internal and External Contacts

Definition

An exterior contact in an embedding of a closed chain C is one that does not subdivide the interior of C.

Lemma

There are no exterior contacts in an optimal embedding of S_k.



The conformation graph of S_k

Lemma

Over all optimal embeddings of S_k, the conformation graph is unique.

Theorem

There is a the unique optimal embedding (up to isometries) of S_k.

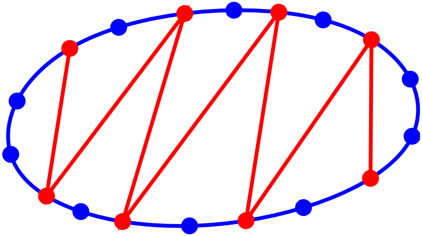
Proof.

Start with one of the four cycles, the embedding is forced. □

The conformation graph of S_k

Lemma

Over all optimal embeddings of S_k, the conformation graph is unique.



Theorem

There is a the unique optimal embedding (up to isometries) of S_k.

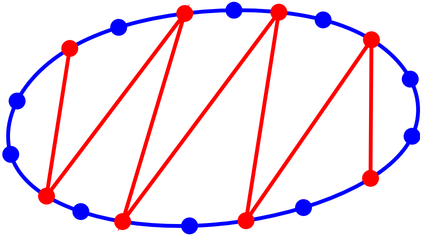
Proof.

Start with one of the four cycles, the embedding is forced. □

The conformation graph of S_k

Lemma

Over all optimal embeddings of S_k, the conformation graph is unique.



Theorem

There is a the unique optimal embedding (up to isometries) of S_k.

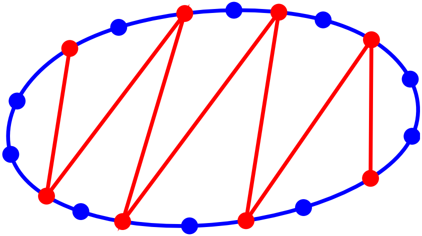
Proof.

Start with one of the four cycles, the embedding is forced. □

The conformation graph of S_k

Lemma

Over all optimal embeddings of S_k, the conformation graph is unique.



Theorem

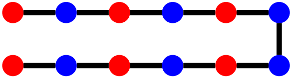
There is a the unique optimal embedding (up to isometries) of S_k.

Proof.

Start with one of the four cycles, the embedding is forced. □

Open Chains

$$Z_k = (HP)^{\lceil k/2 \rceil}(PH)^{\lfloor k/2 \rfloor}$$



Theorem

Z_{2j} has a unique optimal embedding for all $j \geq 1$.

Skip proof

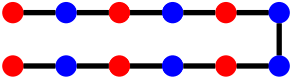
Proof.

(Sketch)

1. How can H nodes appear on the bounding box?
2. Both endpoints on the bounding box, and in contact.
3. The monochrome edge is on the bounding box.
4. The open case reduces to the closed case

Open Chains

$$Z_k = (\mathrm{HP})^{\lceil k/2 \rceil}(\mathrm{PH})^{\lfloor k/2 \rfloor}$$



Theorem

Z_{2j} has a unique optimal embedding for all $j \geq 1$.

Proof.

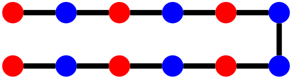
(Sketch)

1. How can H nodes appear on the bounding box?
2. Both endpoints on the bounding box, and in contact.
3. The monochrome edge is on the bounding box.
4. The open case reduces to the closed case

□

Open Chains

$$Z_k = (HP)^{\lceil k/2 \rceil}(PH)^{\lfloor k/2 \rfloor}$$



Theorem

Z_{2j} has a unique optimal embedding for all $j \geq 1$.

Proof.

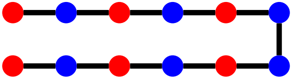
(Sketch)

1. How can H nodes appear on the bounding box?
2. Both endpoints on the bounding box, and in contact.
3. The monochrome edge is on the bounding box.
4. The open case reduces to the closed case

□

Open Chains

$$Z_k = (\text{HP})^{\lceil k/2 \rceil} (\text{PH})^{\lfloor k/2 \rfloor}$$



Theorem

Z_{2j} has a unique optimal embedding for all $j \geq 1$.

Proof.

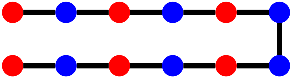
(Sketch)

1. How can H nodes appear on the bounding box?
2. Both endpoints on the bounding box, and in contact.
3. The monochrome edge is on the bounding box.
4. The open case reduces to the closed case

□

Open Chains

$$Z_k = (HP)^{\lceil k/2 \rceil}(PH)^{\lfloor k/2 \rfloor}$$



Theorem

Z_{2j} has a unique optimal embedding for all $j \geq 1$.

Proof.

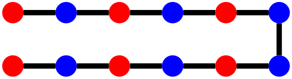
(Sketch)

1. How can H nodes appear on the bounding box?
2. Both endpoints on the bounding box, and in contact.
3. The monochrome edge is on the bounding box.
4. The open case reduces to the closed case

□

Open Chains

$$Z_k = (\text{HP})^{\lceil k/2 \rceil}(\text{PH})^{\lfloor k/2 \rfloor}$$



Theorem

Z_{2j} has a unique optimal embedding for all $j \geq 1$.

Proof.

(Sketch)

1. How can H nodes appear on the bounding box?
2. Both endpoints on the bounding box, and in contact.
3. The monochrome edge is on the bounding box.
4. The open case reduces to the closed case

□

Outline

Open and Closed Questions

Open Questions

1. Do real proteins fold uniquely in the H–P model?
2. Asymptotically, what fraction of n-node H–P-sequences fold uniquely?
3. Is H–P sequence folding still NP-complete when restricted to “nice” sequences?

Not so open questions

- There exist stable H–P trees in $3D$.
- There are stable chains in the 2D H-anything model.
- Minimal area and maximum contacts are not always simultaneously achievable.

Open and Closed Questions

Open Questions

1. Do real proteins fold uniquely in the H–P model?
2. Asymptotically, what fraction of n-node H–P-sequences fold uniquely?
3. Is H–P sequence folding still NP-complete when restricted to “nice” sequences?

Not so open questions

- There exist stable H–P trees in $3D$.
- There are stable chains in the 2D H-anything model.
- Minimal area and maximum contacts are not always simultaneously achievable.

Open and Closed Questions

Open Questions

1. Do real proteins fold uniquely in the H–P model?
2. Asymptotically, what fraction of n-node H–P-sequences fold uniquely?
3. Is H–P sequence folding still NP-complete when restricted to “nice” sequences?

Not so open questions

- There exist stable H–P trees in 3D.
- There are stable chains in the 2D H-anything model.
- Minimal area and maximum contacts are not always simultaneously achievable.

Open and Closed Questions

Open Questions

1. Do real proteins fold uniquely in the H–P model?
2. Asymptotically, what fraction of *n*-node H–P-sequences fold uniquely?
3. Is H–P sequence folding still NP-complete when restricted to "nice" sequences?

Not so open questions

- There exist stable H–P trees in $3D$.
- There are stable chains in the 2D H-anything model.
- Minimal area and maximum contacts are not always simultaneously achievable.

Open and Closed Questions

Open Questions

1. Do real proteins fold uniquely in the H–P model?
2. Asymptotically, what fraction of n-node H–P-sequences fold uniquely?
3. Is H–P sequence folding still NP-complete when restricted to "nice" sequences?

Not so open questions

- There exist stable H–P trees in $3D$.
- There are stable chains in the 2D H-anything model.
- Minimal area and maximum contacts are not always simultaneously achievable.

Open and Closed Questions

Open Questions

1. Do real proteins fold uniquely in the H–P model?
2. Asymptotically, what fraction of n-node H–P-sequences fold uniquely?
3. Is H–P sequence folding still NP-complete when restricted to “nice” sequences?

Not so open questions

- There exist stable H–P trees in $3D$.
- There are stable chains in the 2D H-anything model.
- Minimal area and maximum contacts are not always simultaneously achievable.

Credits

- Inspired by an article of Brian Hayes in American Scientist
- Initiated at a workshop on Molecular Reconfiguration organized by Godfried Toussaint.
- Non-globular examples with Oswin Aichholzer, Erik Demaine, Vera Sacristan and Mike Soss.
- Globular examples with Henk Meijer and Jit Bose