The resource theory of tensor networks

Matthias Christandl, Vladimir Lysikov,
Vincent Steffan, Albert Werner, Freek Witteveen

Tensor network representations from the geometry of entangled states

Matthias Christandl[1], Angelo Lucia[1,2,3], Péter Vrana[1,4,5], Albert H. Werner[1,2,*]

The resource theory of tensor networks

Matthias Christandl,[1, *] Vladimir Lysikov,[2, †] Vincent Steffan,[1, ‡] Albert H. Werner,[1, §] and Freek Witteveen[1, ¶]

[1] *Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark*

[2] *Faculty of Computer Science, Ruhr University Bochum, Universitätsstraße 150, 44801 Bochum, Germany*

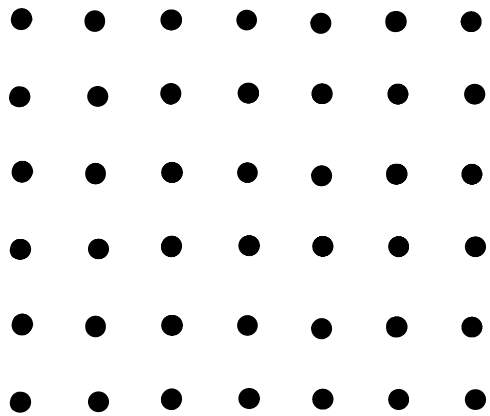
(Dated: July 17, 2023)

Plan

1. Motivation
2. What is a tensor network?
3. Resource theory of tensors
4. The resource theory of tensor networks

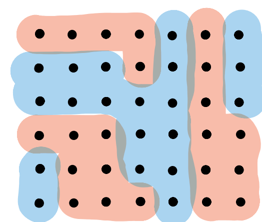
① Motivation

Paradigmatic example: lattice of spin particles



N sites, each can be $\uparrow$ or $\downarrow$

Hilbert space $\mathcal{H} = (\mathbb{C}^2)^{\otimes N}$

$|\Psi\rangle = \sum_{\text{conf } \sigma} \psi_\sigma \Big|$  $\Big\rangle$

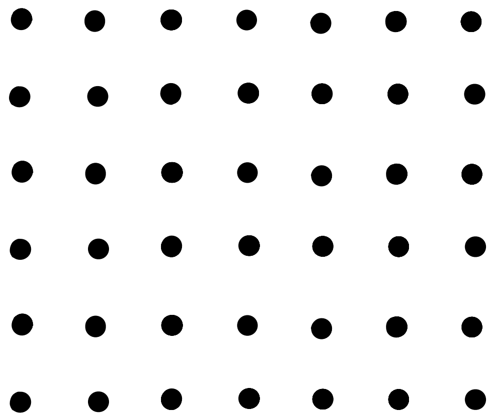
$\cap$

$(\mathbb{C}^2)^{\otimes N}$

2^N coefficients, inefficient description!

1 Motivation

Paradigmatic example: lattice of spin particles

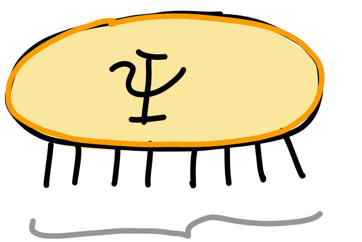


N sites, each can be ↑ or ↓

Hilbert space $\mathcal{H} = (\mathbb{C}^2)^{\otimes N}$

$$|\Psi\rangle = \sum_{i_1 \dots i_N \in \{0,1\}^N} \psi_{i_1 \dots i_N} \, |i_1\rangle \dots |i_N\rangle$$

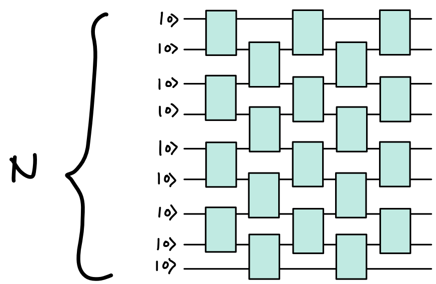
think of $\{\psi_{i_1 \dots i_N}\}_{i_1 \dots i_N}$ as a tensor



N legs

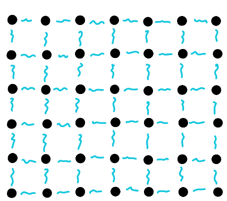
However, most quantum states/tensors we care about have additional structure & may have an 'efficient' description!

1. States prepared by a polynomial sized quantum circuit:



describing the gates requires only poly(N) parameters

2. Ground states of local Hamiltonians

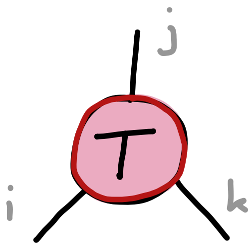


$$H = \sum_{\substack{ij \\ \text{neighbours}}} h_{ij}$$

Ground state $H|\psi_{gs}\rangle = E_{gs}|\psi_{gs}\rangle \Longrightarrow$ Area law entanglement

② What is a tensor network?

<u>tensor</u>

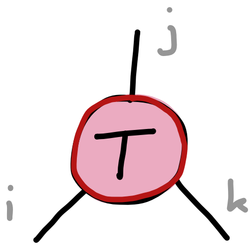


$$\left\{ T_{ijk} \in \mathbb{C} \right\}_{i,j,k=1}^{d_1, d_2, d_3}$$

$$T = \sum_{i,j,k} T_{ijk} |i\rangle |j\rangle |k\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} \otimes \mathbb{C}^{d_3}$$

② What is a tensor network?

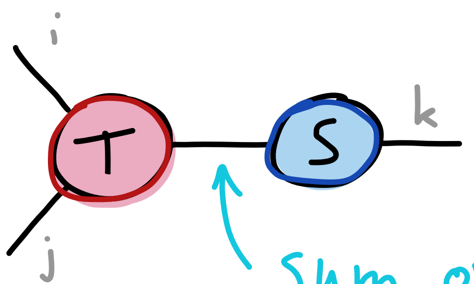
tensor



$$\{T_{ijk} \in \mathbb{C}\}_{i,j,k=1}^{d_1,d_2,d_3}$$

$$T = \sum_{i,j,k} T_{ijk} |i\rangle|j\rangle|k\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} \otimes \mathbb{C}^{d_3}$$

Tensor contraction: "connect legs"

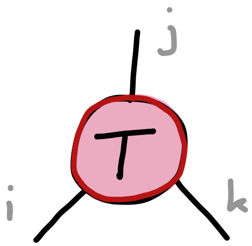


$$= \sum_{\ell} T_{ij\ell} S_{\ell k}$$

sum over internal index ℓ

② What is a tensor network?

tensor

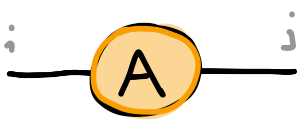


$$\{T_{ijk} \in \mathbb{C}\}_{i,j,k=1}^{d_1,d_2,d_3}$$

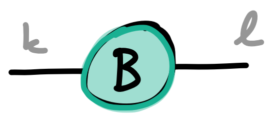
$$T = \sum_{i,j,k} T_{ijk} |i\rangle |j\rangle |k\rangle \in \mathbb{C}^{d_1} \otimes \mathbb{C}^{d_2} \otimes \mathbb{C}^{d_3}$$

Tensor contraction: "connect legs"

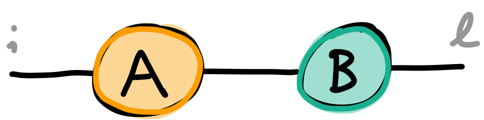
Example: matrix multiplication



A_{ij}

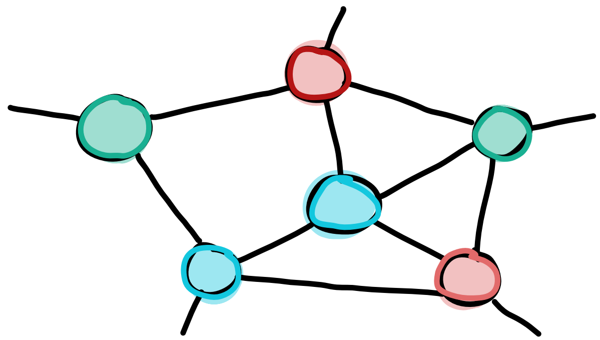


$B_{k\ell}$

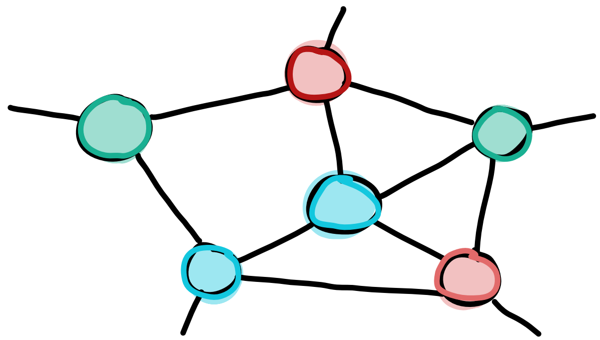


$\sum_j A_{ij} B_{j\ell}$

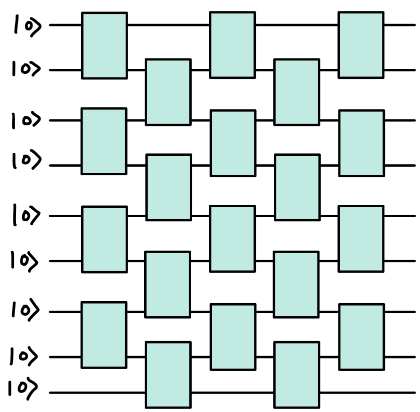
In general, have some **network** of tensors



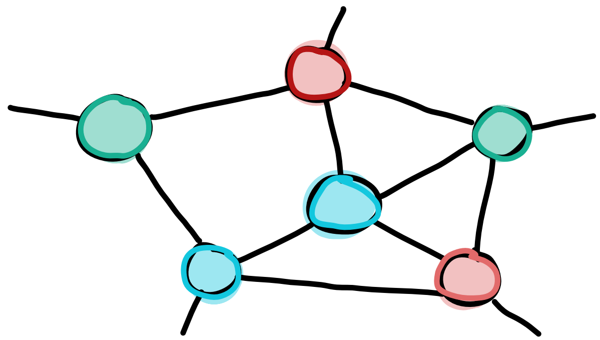
In general, have some **network** of tensors



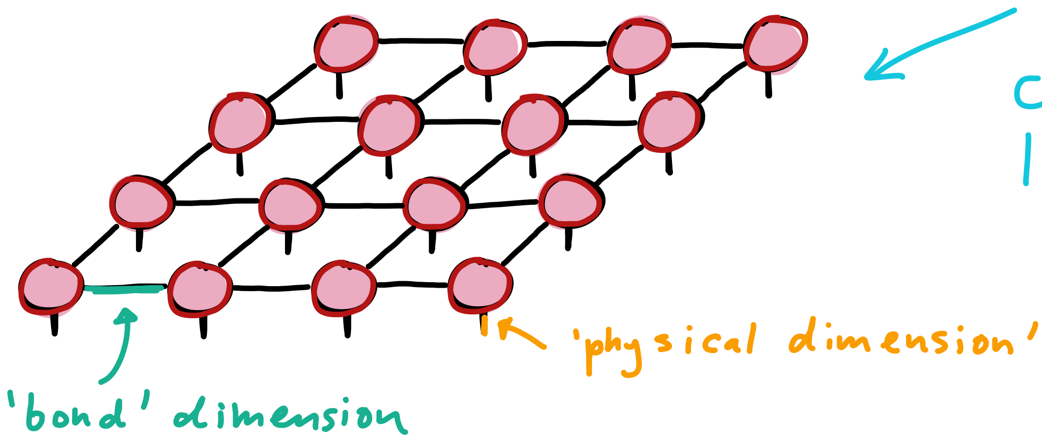
① Quantum circuits



In general, have some network of tensors



2 Lattice PEPS

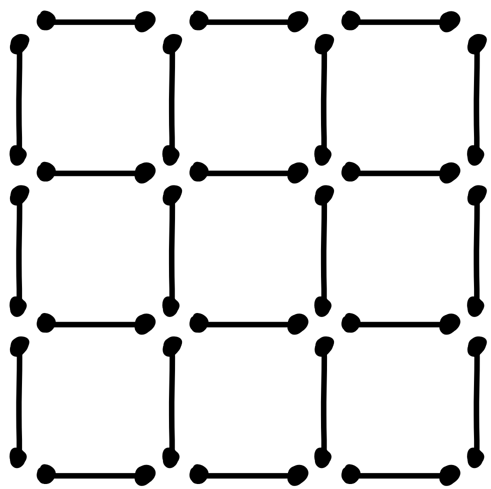


supposed to be a good class of states for lattice ground states for small D

A different perspective on the same thing

PEPS = 'projected entangled pair states'

① Start with a collection of level D entangled pairs on a graph G



 $= \sum_{i=1}^{D} |ii\rangle$

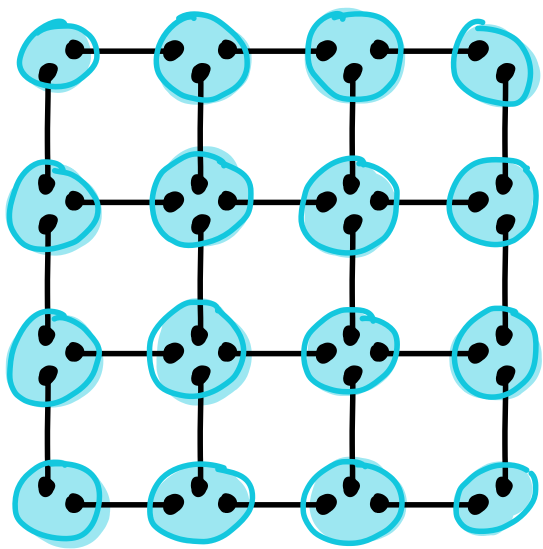
$$|\phi\rangle_G = \bigotimes_{e\in E} |\phi_e\rangle$$

$$|\phi_e\rangle = \sum_{i=1}^{D} |ii\rangle$$

A different perspective on the same thing

PEPS = 'projected entangled pair states'

2 Apply linear maps at the vertices of G



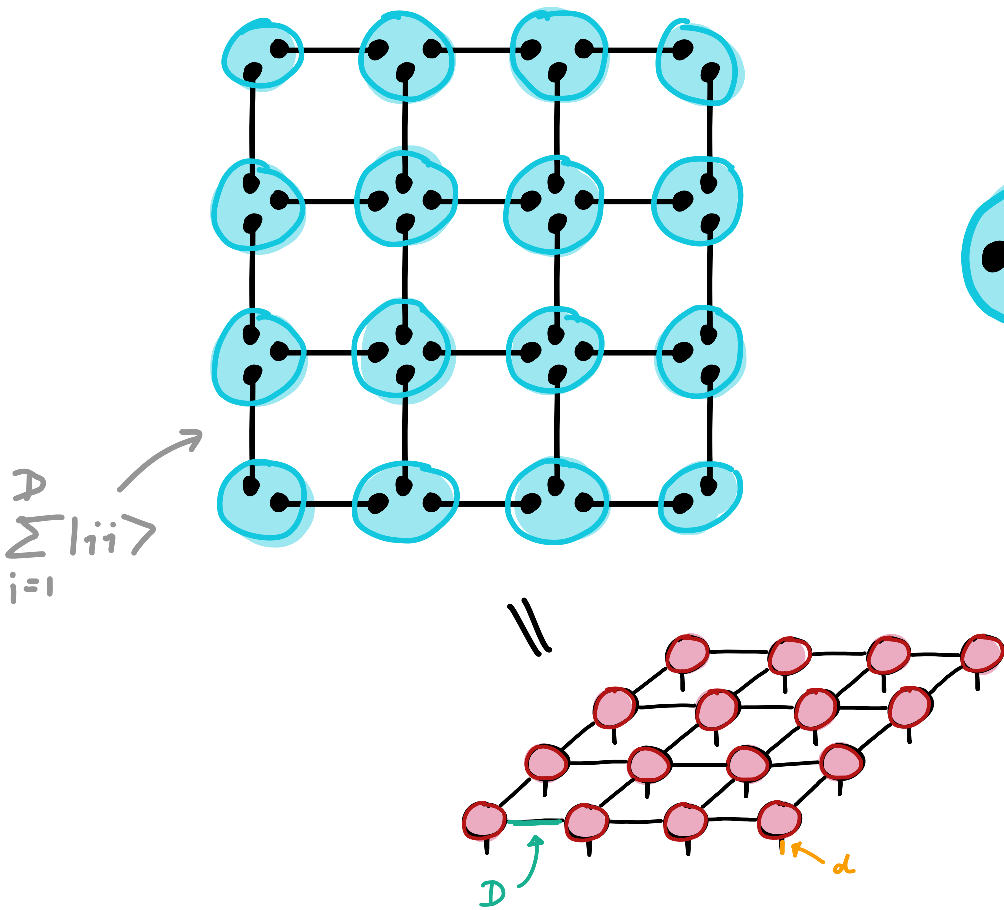
$$= M : (\mathbb{C}^D)^{\otimes 4} \to \mathbb{C}^d$$

$$|\Psi\rangle = \Big(\bigotimes_{v\in V} M_v\Big)\Big(\bigotimes_{e\in E} |\phi_e\rangle\Big)$$

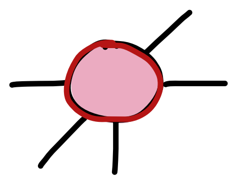
A different perspective on the same thing

PEPS = 'projected entangled pair states'

2 Apply linear maps at the vertices of G



$$= M : (\mathbb{C}^D)^{\otimes 4} \to \mathbb{C}^d$$

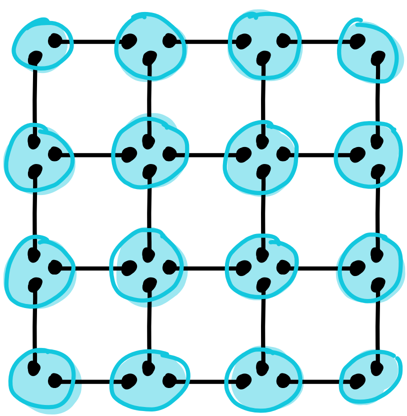


$$\in (\mathbb{C}^D)^{\otimes 4} \otimes \mathbb{C}^d$$

A different perspective on the same thing

PEPS = 'projected entangled pair states'

1. Start with a collection of level D entangled pairs on a graph G
2. Apply linear maps at the vertices of G



Can think of tensor network state as:

We start with local entanglement, and then apply local linear maps

"SLOCC"

tensor network states = states obtained by SLOCC from a network of max. entangled states

Entanglement structures

A natural generalization:

Start with different states than maximally entangled!

Entanglement structures

A natural generalization:

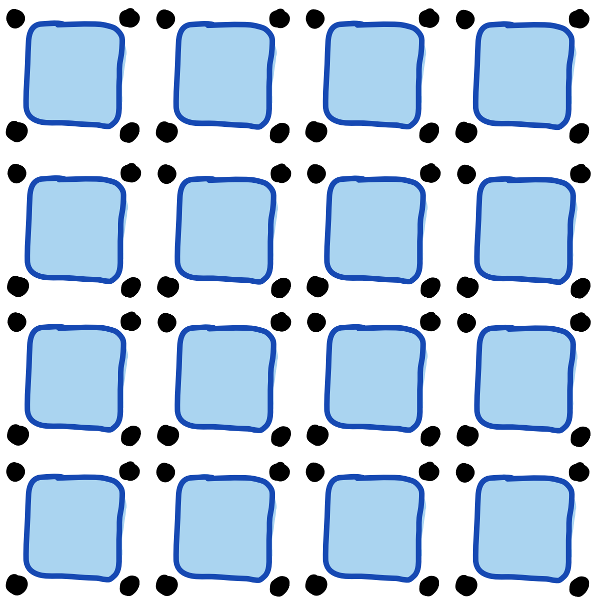
Start with different states than maximally entangled!

Note: for 2-party states this does not make a difference, given $|\phi_e\rangle = \sum_{i=1}^{D} \lambda_e |ii\rangle$ can apply SLOCC to transform to max entangled

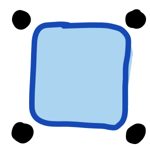
Entanglement structures

A natural generalization:

Start with different states than maximally entangled!



For example:

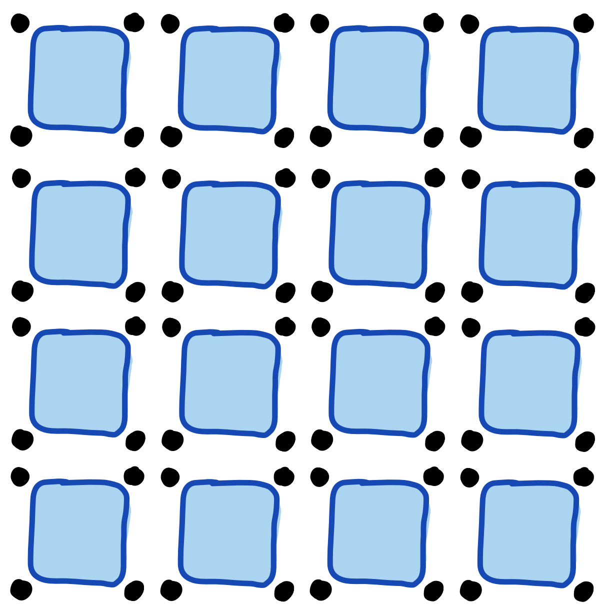
 $= \sum_{i=1}^{r} |iiii\rangle$

GHZ state

Entanglement structures

A natural generalization:

Start with different states than maximally entangled!



In general:

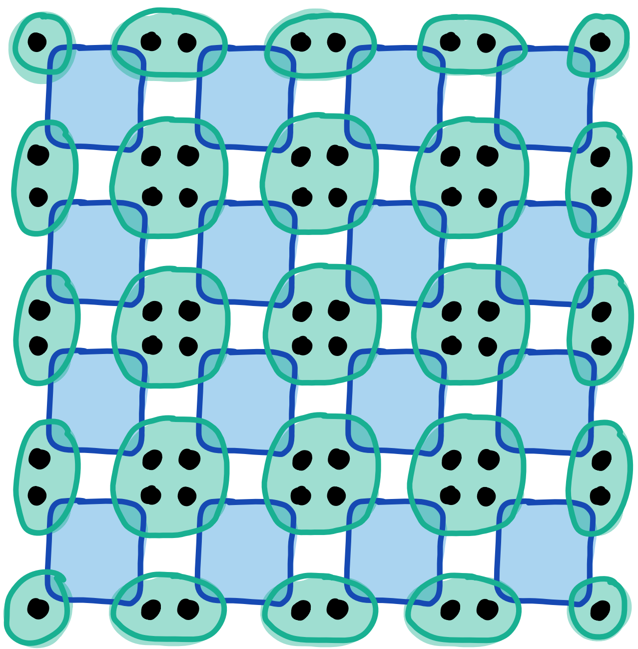
Hypergraph $G = (V, E)$

$$|\phi\rangle_G = \bigotimes_{e \in E} |\phi_e\rangle$$

Entanglement structures

A natural generalization:

Start with different states than maximally entangled!



In general:

Hypergraph $G = (V, E)$

$$|\phi\rangle_G = \bigotimes_{e \in E} |\phi_e\rangle$$

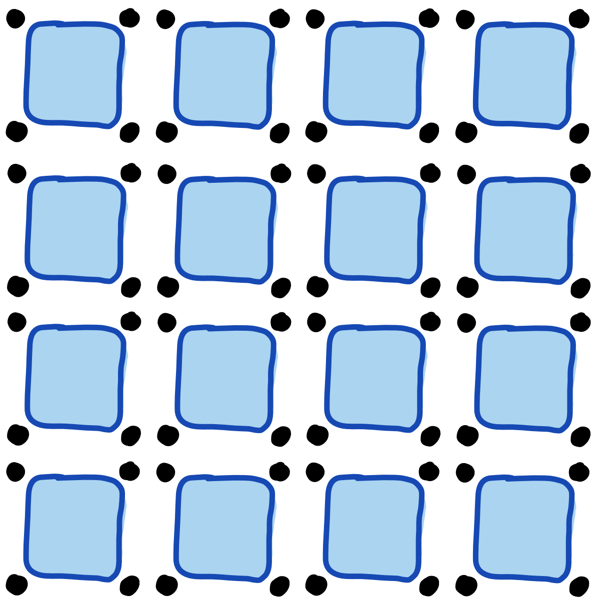
Apply linear maps at vertices to get tensor network state

$$|\Psi\rangle = \left(\bigotimes_{v \in V} M_v\right)\left(\bigotimes_{e \in E} |\phi_e\rangle\right)$$

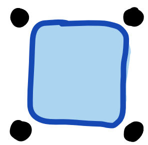
Entanglement structures

A natural generalization:

Start with different states than maximally entangled!



example:

 $= \sum_{i=1}^{r} |iiii\rangle$

GHZ state

$\Downarrow$

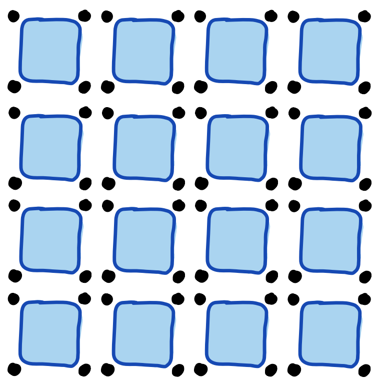
Contraction

$$\sum_{i=1}^{r} S_{i\, j_1 j_2 j_3}\, T_{i\, k_1 k_2 k_3}\, U_{i\, l_1 l_2 l_3}\, V_{i\, m_1 m_2 m_3}$$

Entanglement structures

A natural generalization: → Christandl-...

Start with different states than maximally entangled!

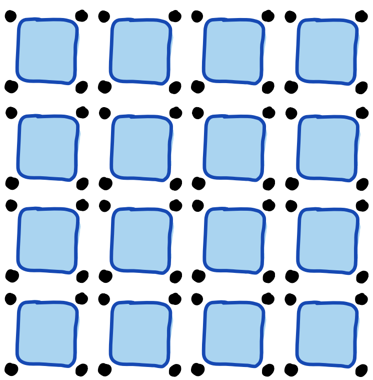


Why would you do this?

Entanglement structures

A natural generalization: Christandl-...

Start with different states than maximally entangled!



Why would you do this?

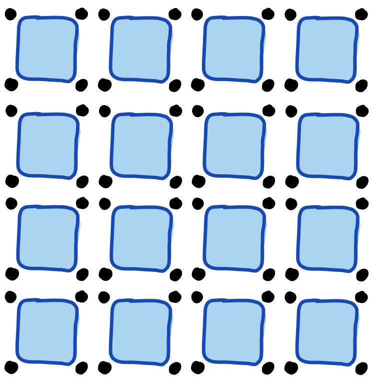
- Not strictly bigger class of states, can always rewrite to the usual notion for large enough D or extra vertices

Entanglement structures

Christandl -...

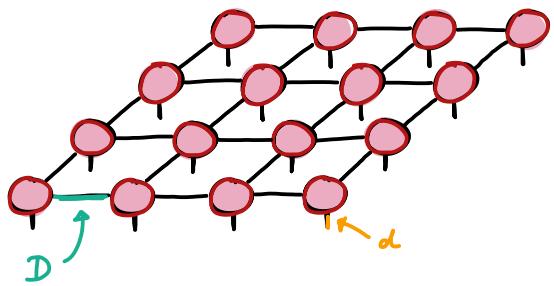
A natural generalization:

Start with different states than maximally entangled!



Why would you do this?

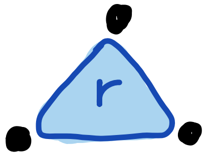
- Not strictly bigger class of states
- For certain models, max. ent. does not capture the correlations → e.g. spin liquids
- Different entanglement structure → better numerics
  PESS, Xie et al
- Representations of states with invertible Mv
  injective PEPS, Molnar et al



The larger the bond dimension D, the more expressive
(but also: algorithms scale (badly) with D)

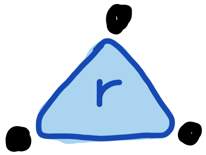
What is "bond dimension" for general entanglement structures?

An important example...

Let 

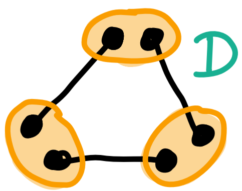
$= \sum_{i=1}^{r} |iii\rangle$ GHZ state

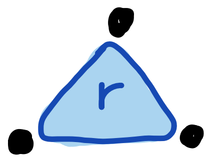
An important example...

Let 

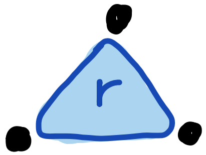
$= \sum_{i=1}^{r} |iii\rangle$ GHZ state

Minimal bond dimension D ?

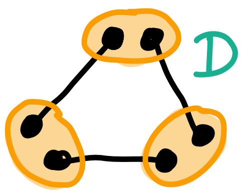


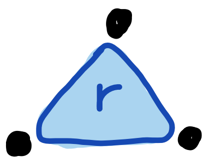
$=$ 

An important example...

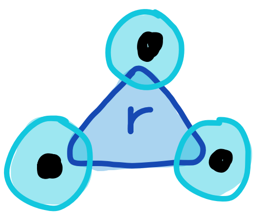
Let  $= \sum_{i=1}^{r} |iii\rangle$ GHZ state

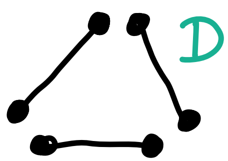
Minimal bond dimension D ?



$=$ 

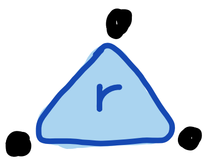
Conversely, minimal r such that



$=$ 

?

An important example...

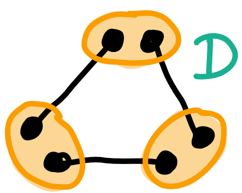
Let 

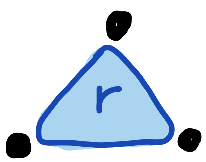
$= \sum_{i=1}^{r} |iii\rangle$ GHZ state

Minimal bond dimension D ?

We don't know!

(we know for a slight variation)

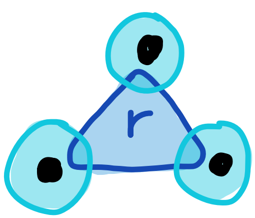


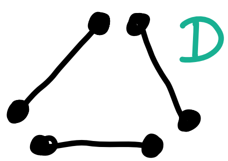
$=$ 

Conversely, minimal r such that

We don't know!

(and many mathematicians thought long and hard about it)

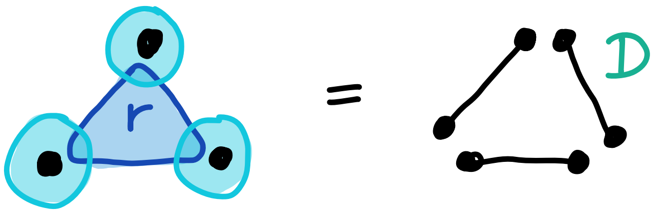


$=$ 

?

③ The resource theory of tensors

Why do mathematicians care about



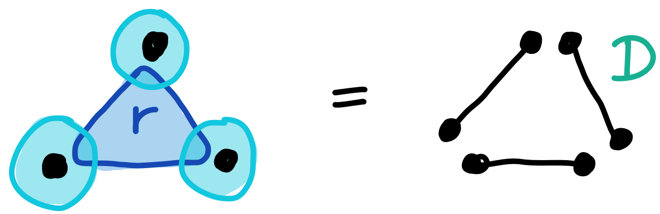
?

Related to the complexity of matrix multiplication

$$\begin{pmatrix} * & * \\ * & * \end{pmatrix}\begin{pmatrix} \bullet & \bullet \\ \bullet & \bullet \end{pmatrix} = \begin{pmatrix} \blacksquare & \blacksquare \\ \blacksquare & \blacksquare \end{pmatrix}$$

How many multiplications needed?

③ The resource theory of tensors

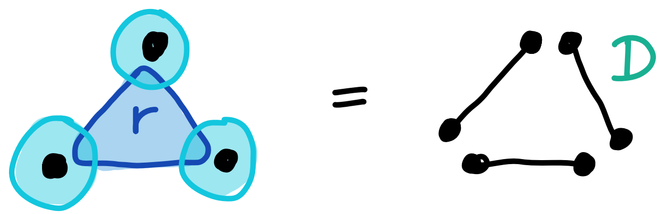


Related to the complexity of matrix multiplication

$$\begin{pmatrix} * & * \\ * & * \end{pmatrix}\begin{pmatrix} \bullet & \bullet \\ \bullet & \bullet \end{pmatrix} = \begin{pmatrix} \blacksquare & \blacksquare \\ \blacksquare & \blacksquare \end{pmatrix}$$

How many multiplications needed? $\cancel{8}$ 7 Strassen '69

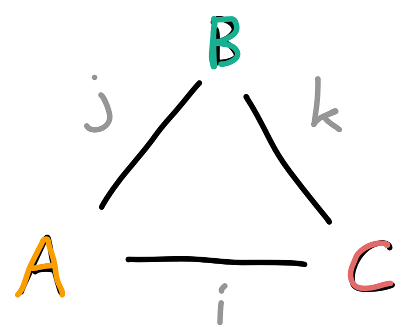
③ The resource theory of tensors



Related to the complexity of matrix multiplication

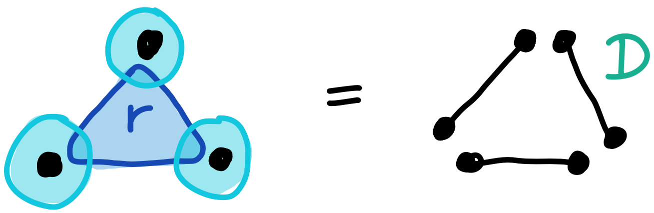
$$\begin{pmatrix} * & * \\ * & * \end{pmatrix}\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} = \begin{pmatrix} \blacksquare & \blacksquare \\ \blacksquare & \blacksquare \end{pmatrix}$$

$$\sum_j A_{ij} B_{jk} = C_{ik}$$



③ The resource theory of tensors

Related to the complexity of matrix multiplication



$\rightsquigarrow$ $D \times D$ matrix multiplication using r multiplications

Major open question: as $D \to \infty$, find best ω s.t.

$$r = \mathcal{O}(D^{\omega})\ ?$$

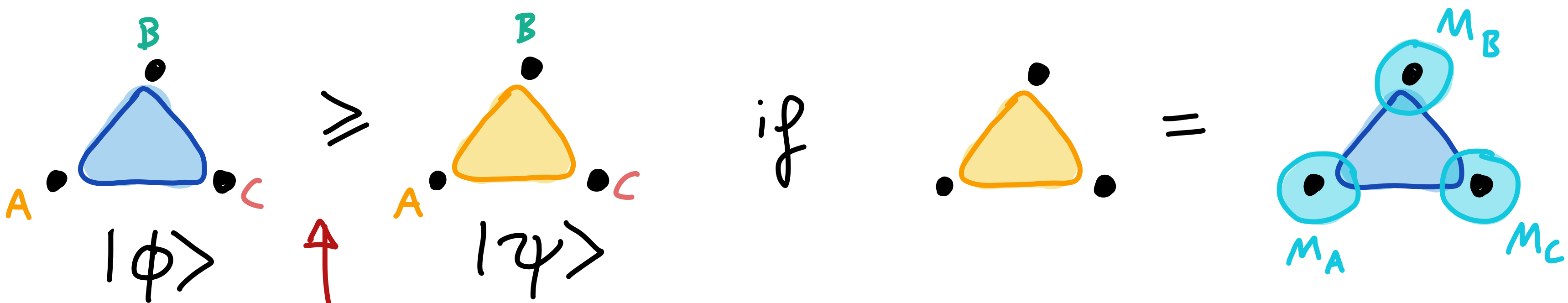
Known: $r \geq \Omega(D^2)$ $\quad (\omega \geq 2)$

$$r \leq \mathcal{O}(D^{2.371\ldots}) \quad (\omega \leq 2.371\ldots)$$

This motivated the **resource theory of tensors**:

$|\phi\rangle, |\psi\rangle$ 3-tensors on parties A, B, C

$$|\phi\rangle \geq |\psi\rangle \quad \text{if} \quad |\psi\rangle = (M_A \otimes M_B \otimes M_C)|\phi\rangle$$



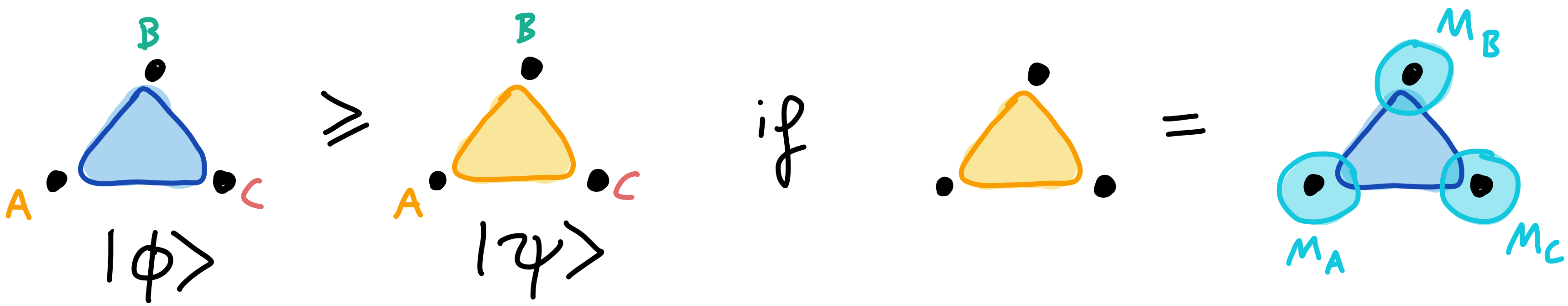
known as a "restriction"

i.e. we say that "$|\phi\rangle$ restricts to $|\psi\rangle$"

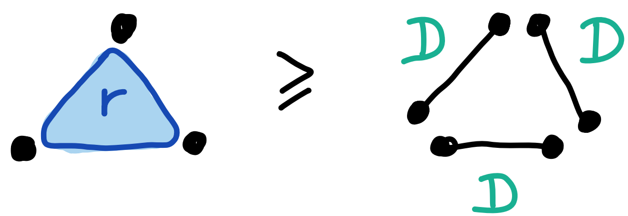
This motivated the resource theory of tensors:

$|\phi\rangle, |\psi\rangle$ 3-tensors on parties A, B, C

$$|\phi\rangle \geq |\psi\rangle \quad \text{if} \quad |\psi\rangle = (M_A \otimes M_B \otimes M_C)|\phi\rangle$$



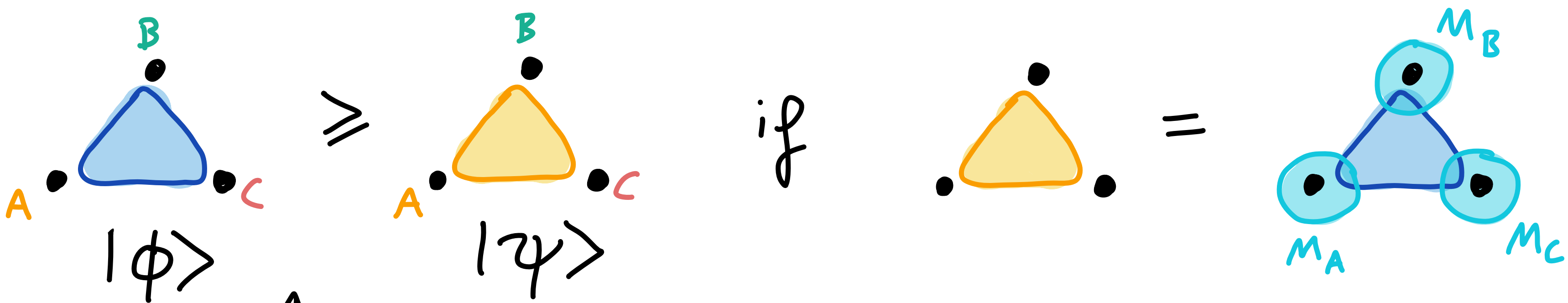
so the previous question becomes: for which r is



This motivated the resource theory of tensors:

$|\phi\rangle, |\psi\rangle$ 3-tensors on parties A, B, C

$$|\phi\rangle \geq |\psi\rangle \text{ if } |\psi\rangle = (M_A \otimes M_B \otimes M_C)|\phi\rangle$$



"$|\phi\rangle$ is a more powerful computational resource than $|\psi\rangle$"

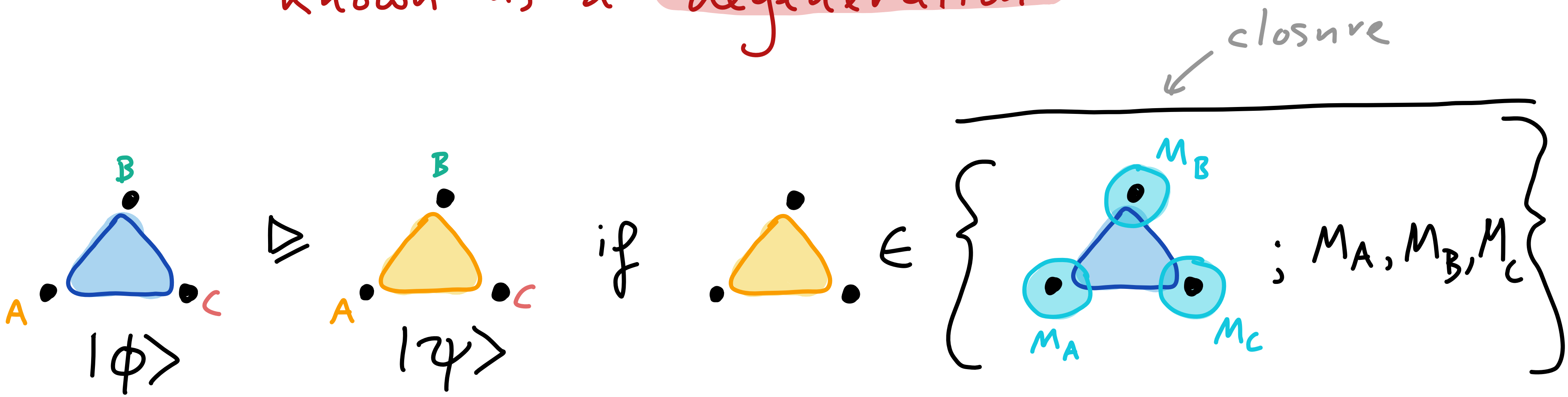
Can also see this as SLOCC entanglement theory

There is also an approximate notion:

$|\phi\rangle, |\psi\rangle$ 3-tensors on parties A, B, C

$$|\phi\rangle \trianglerighteq |\psi\rangle \quad \text{if} \quad |\psi\rangle = \lim_{\varepsilon \to 0} \left(M_A(\varepsilon) \otimes M_B(\varepsilon) \otimes M_C(\varepsilon) \right) |\phi\rangle$$

known as a "degeneration"



There is also an approximate notion:

$|\phi\rangle, |\psi\rangle$ 3-tensors on parties A, B, C

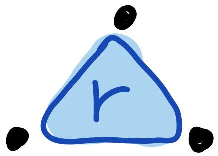
$$|\phi\rangle \trianglerighteq |\psi\rangle \text{ if } |\psi\rangle = \lim_{\varepsilon \to 0} \left(M_A(\varepsilon) \otimes M_B(\varepsilon) \otimes M_C(\varepsilon)\right) |\phi\rangle$$

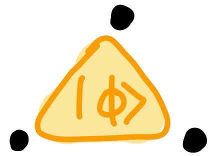
known as a "degeneration"

Finally there is also an asymptotic version

(clearly relevant for the asymptotics of the algorithm)

An aside on terminology

Best r s.t. 

$\geq$ 

i.e. smallest

GHZ state

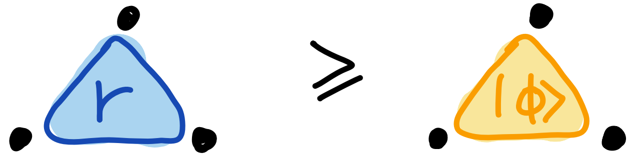
is the tensor rank of $|\phi\rangle$

Note: equivalent $|\phi\rangle = \sum_{i=1}^{r} |a_i\rangle |b_i\rangle |c_i\rangle$

for some vectors $|a_i\rangle, |b_i\rangle, |c_i\rangle$

An aside on terminology

Best r s.t.

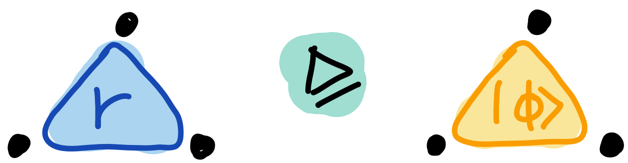


i.e. smallest

GHZ state

is the **tensor rank** of $|\phi\rangle$

Best r s.t.



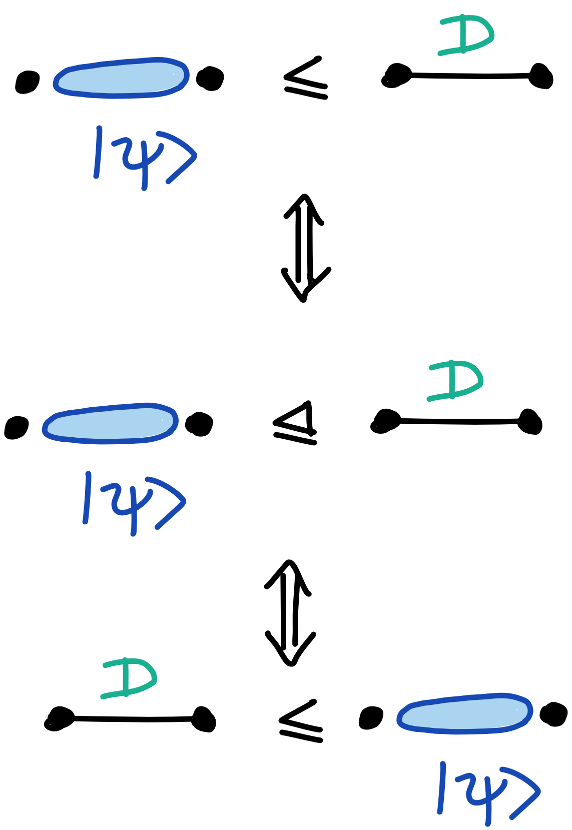
i.e. smallest

GHZ state

is the **border tensor rank** of $|\phi\rangle$

An aside on terminology

For 2 parties, both notions reduce to the usual rank (or Schmidt rank)



$$\Longleftrightarrow \quad |\psi\rangle = \sum_{i=1}^{D} \lambda_i \, |a_i\rangle |b_i\rangle$$

Lots of theory for this resource theory!

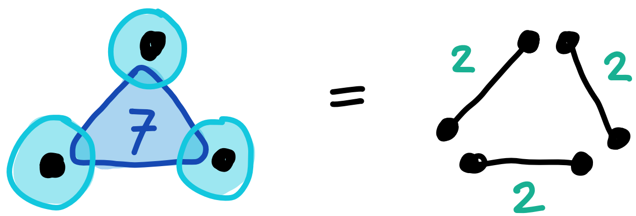
3 useful things

Lots of theory for this resource theory!

3 useful things

① Explicit constructions

Given some tensors, maps that show a restriction, e.g. (or degeneration)



This corresponds to faster algorithms (upper bounds)

Lots of theory for this resource theory!

3 useful things

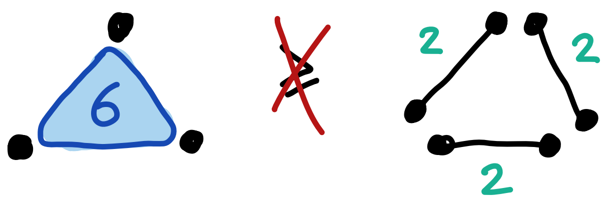
talk by JM Landsberg

(2) General obstructions

Methods to show that $|\phi\rangle \not\geq |\psi\rangle$

(or for degenerations)

For example, we know



This corresponds to obstructions to faster algorithms

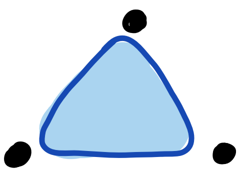
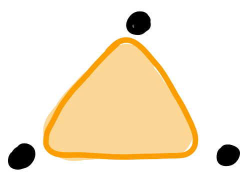
(lower bounds)

Lots of theory for this resource theory!

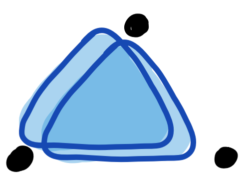
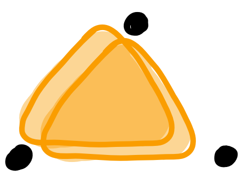
3 useful things

③ The resource theory of tensors is not multiplicative

There exist states/tensors s.t.

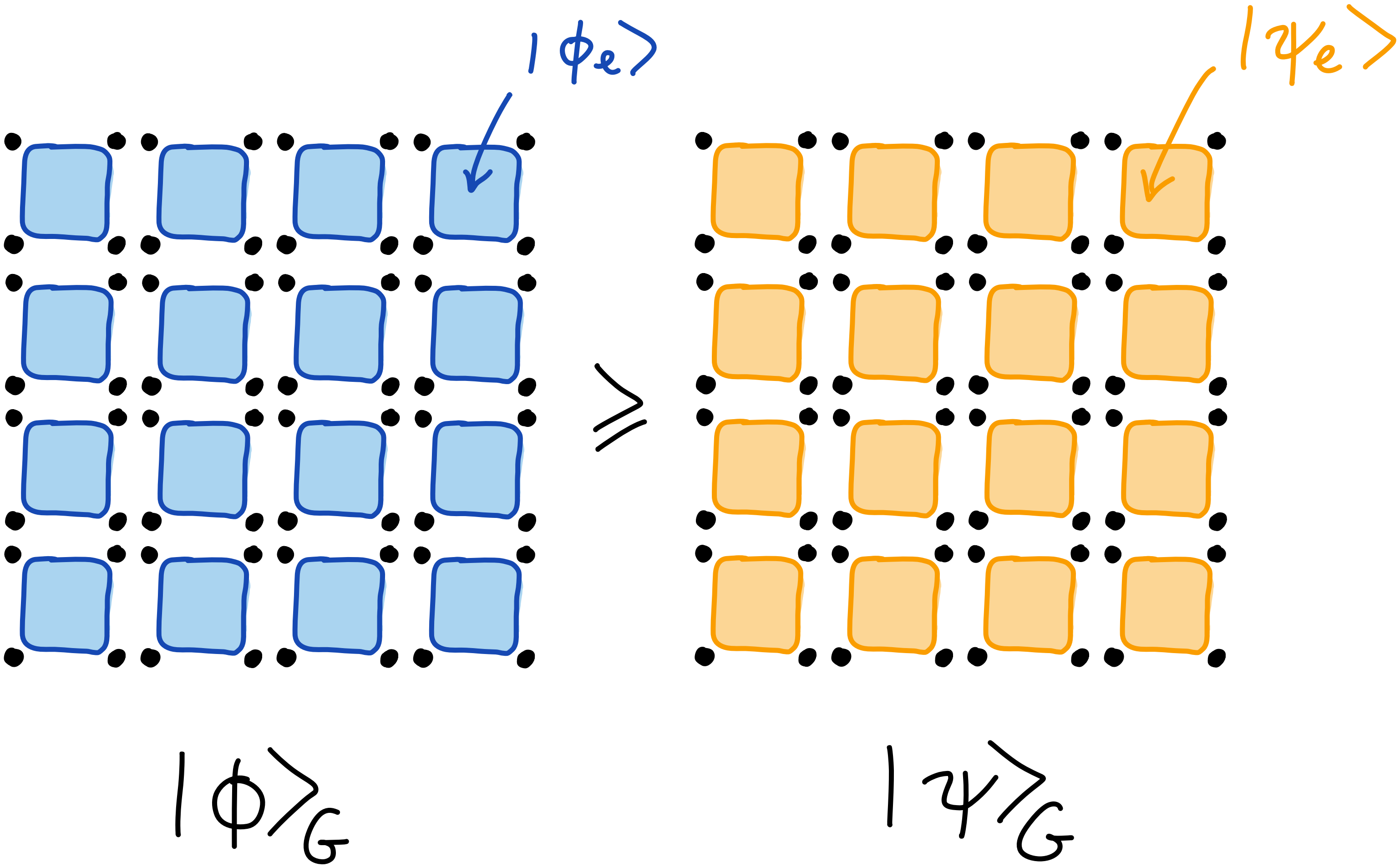
 $\not\geq$  $|\phi\rangle \not\geq |\psi\rangle$

but

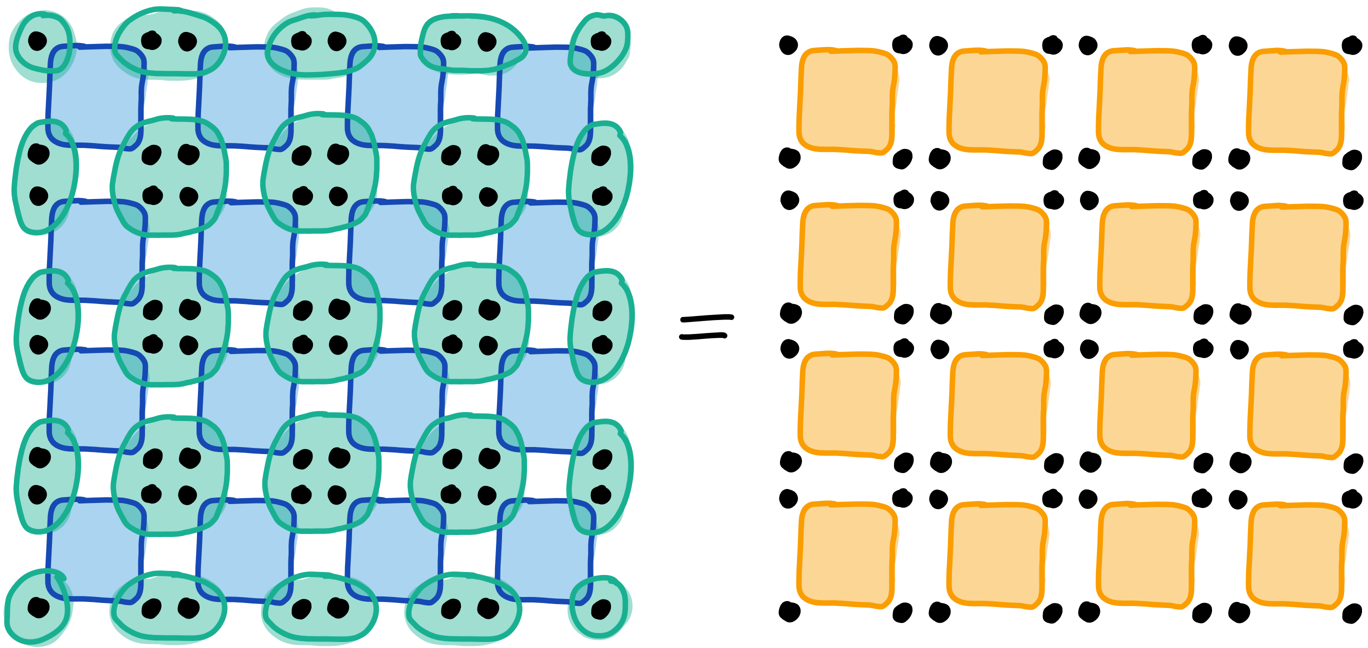
 $\geq$  $|\phi\rangle^{\otimes 2} \geq |\psi\rangle^{\otimes 2}$

This relates to the asymptotic performance of algorithms

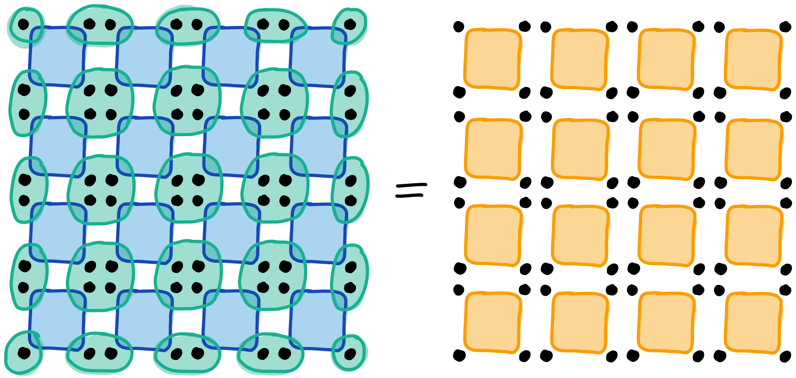
④ The resource theory of tensor networks



④ The resource theory of tensor networks



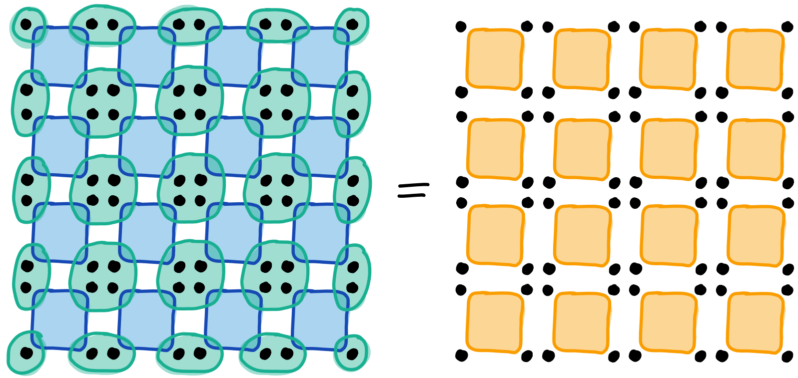
so $|\phi\rangle_G \geq |\psi\rangle_G$ if $|\psi\rangle_G$ has a tensor network representation using $|\phi\rangle_G$



$$|\phi\rangle_G \geq |\psi\rangle_G$$

Note: every $|\Psi\rangle$ which has a tensor network representation using $|\psi\rangle_G$ also has a representation using $|\phi\rangle_G$

$\Longrightarrow$ $|\phi\rangle_G$ is a more powerful resource than $|\psi\rangle_G$

$\lim_{\varepsilon \to 0}$ 

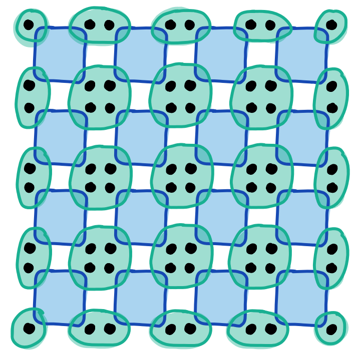
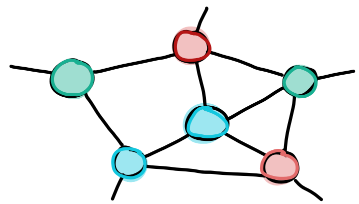
$|\phi\rangle_G \trianglerighteq |\psi\rangle_G$

Same for degenerations!

Previous work Christandl et al.

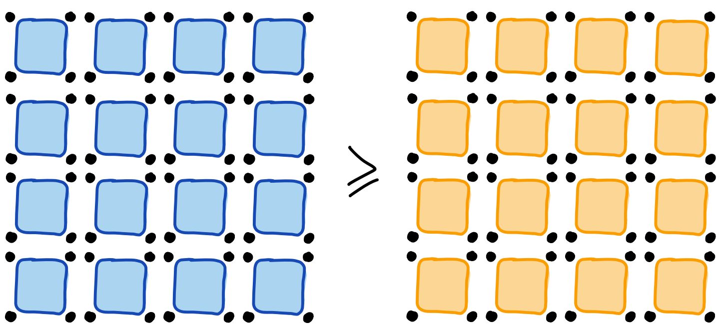
tensor network degenerations can be converted to tensor network restriction at small overhead in important cases

Intermezzo: algebraic complexity of contraction

Compute coefficient of  or 

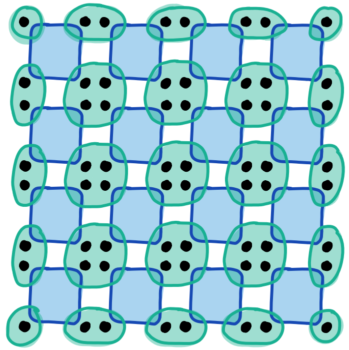
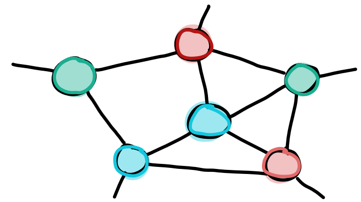
$\Longrightarrow$ Polynomial in the tensor entries.

What is the minimal #multiplications needed to evaluate this polynomial?

① 

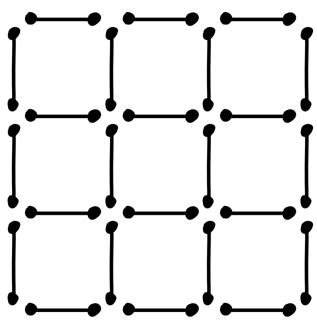
implies contracting  is easier than 

Intermezzo: algebraic complexity of contraction

Compute coefficient of  or 

$\Longrightarrow$ Polynomial in the tensor entries.

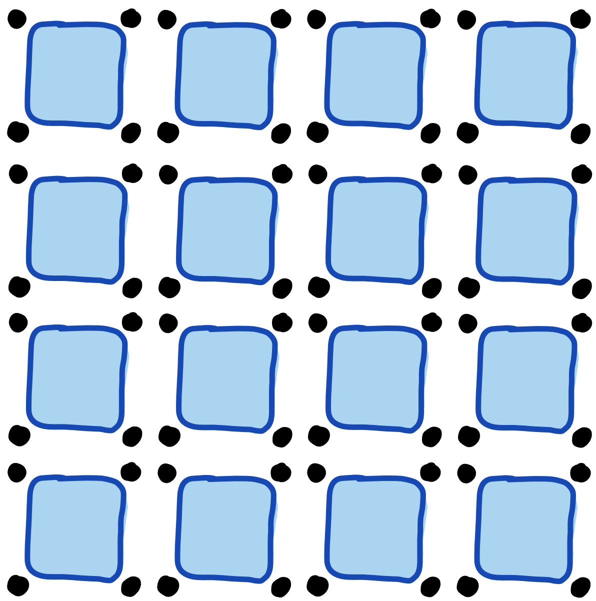
What is the minimal #multiplications needed to evaluate this polynomial?

2 $D=2$  contraction is VNP-complete

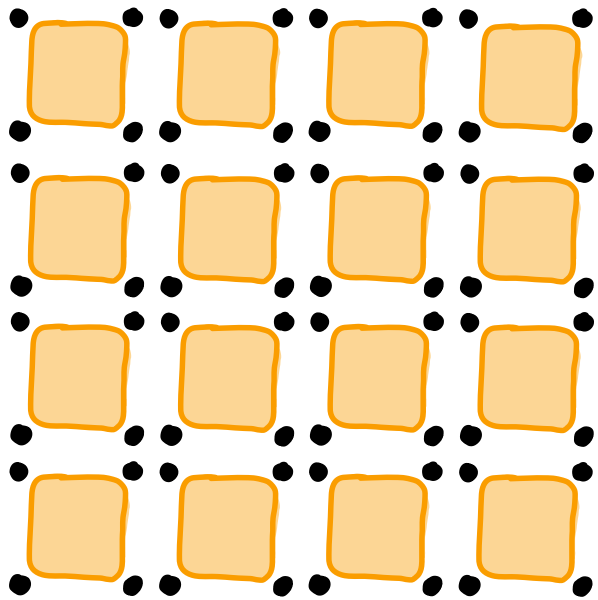
as hard as permanent

algebraic version of #P

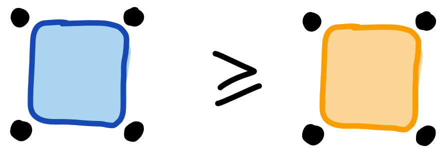
This leaves us with the (difficult) question:

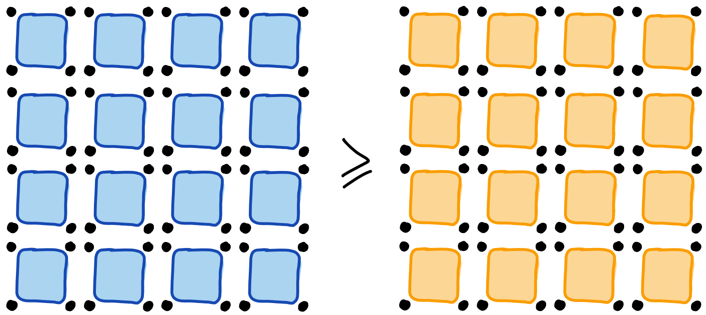


$\overset{?}{\geqslant}$



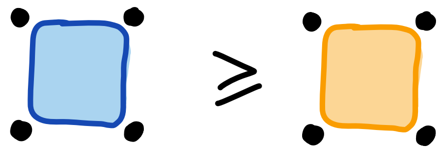
$$|\phi\rangle_G \qquad |\psi\rangle_G$$

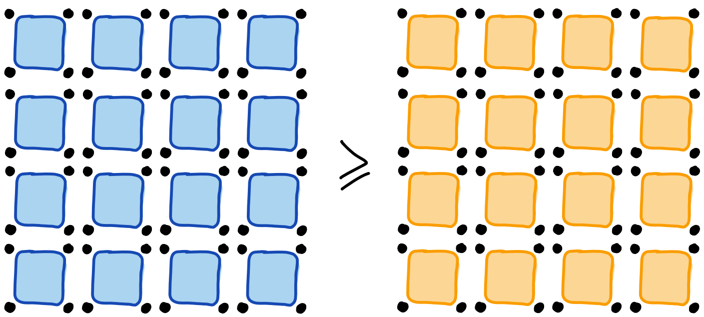
Special case: if 



single-plaquette transformation

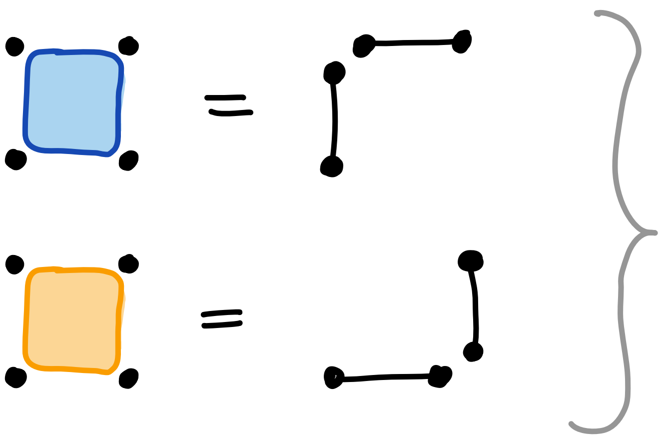
Is this all there is?

Special case: if 

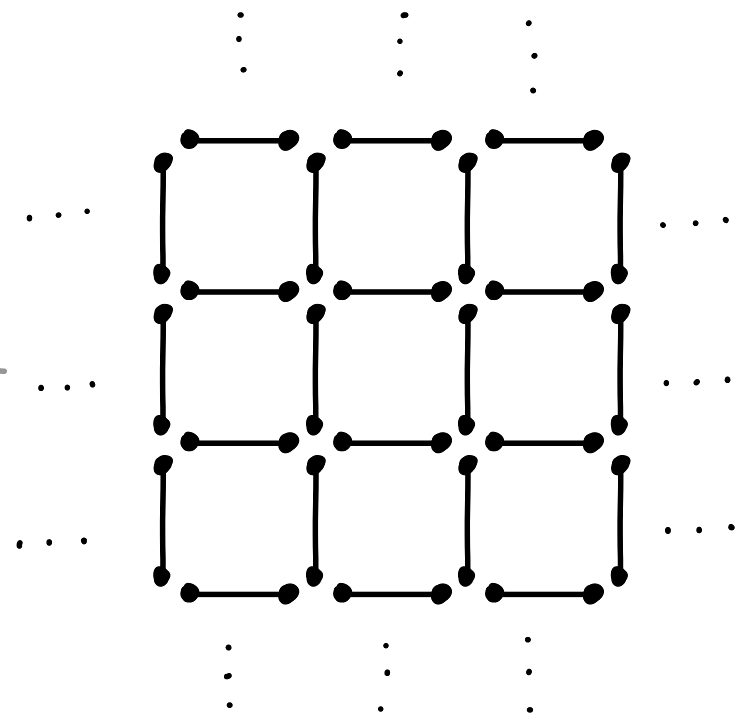


single-plaquette transformation

Is this all there is? NO!



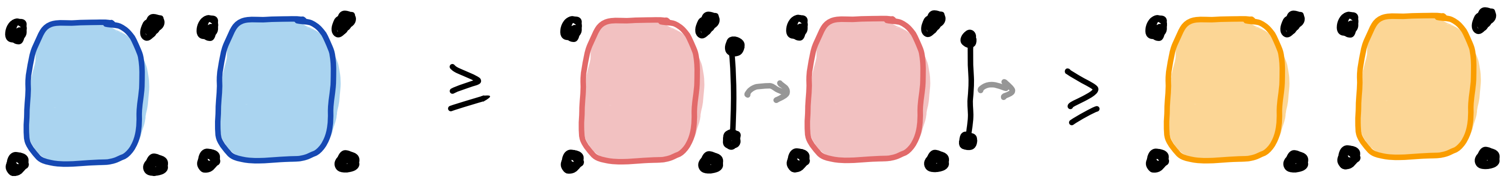
on the lattice both give



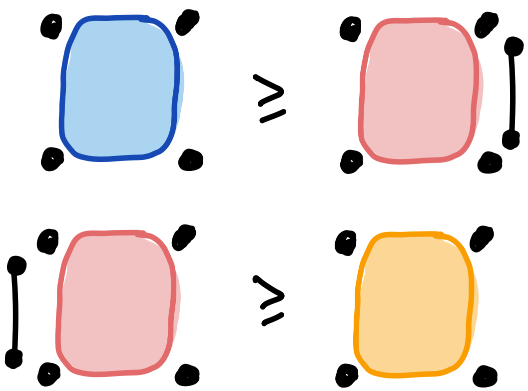
Ok, sure ... but that's kind of an uninteresting example!

Ok, sure ... but that's kind of an uninteresting example!

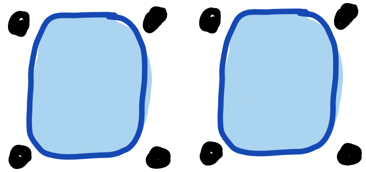
More generally, can have constructions where

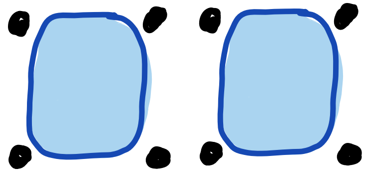


i.e.

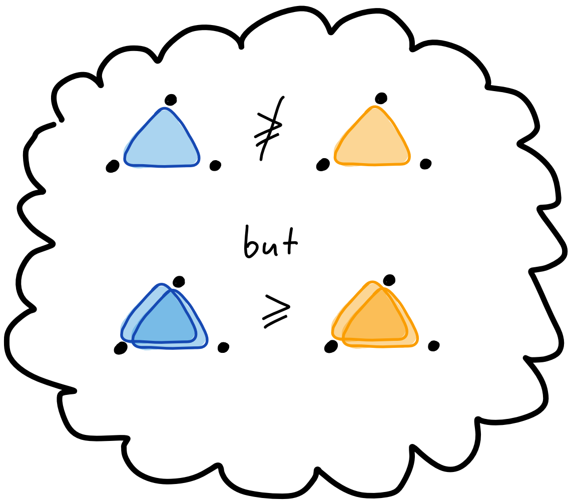


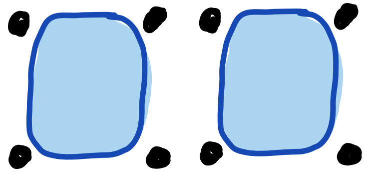
& exchange of 2-party entanglement

Since  only adjacent in two vertices, maybe all transformations consist of some exchange of max. ent. states?

Since  only adjacent in two vertices, maybe all transformations consist of some exchange of max. ent. states?

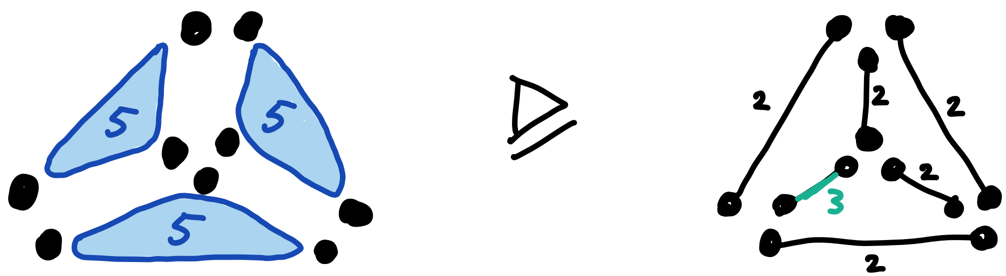
Recall ∘ ○ O



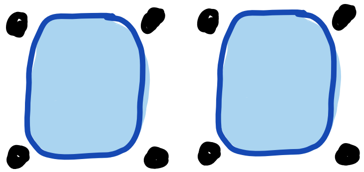
Since  only adjacent in two vertices, maybe all transformations consist of some exchange of max. ent. states?

NO!

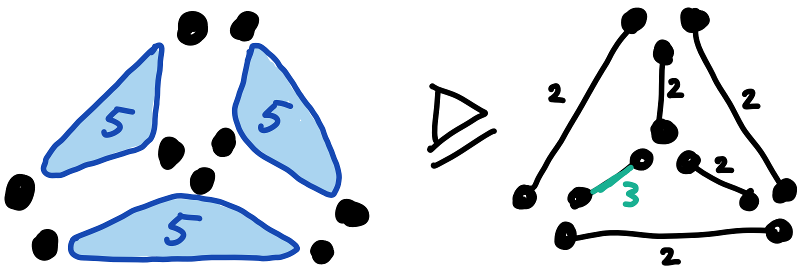
We give an explicit example:

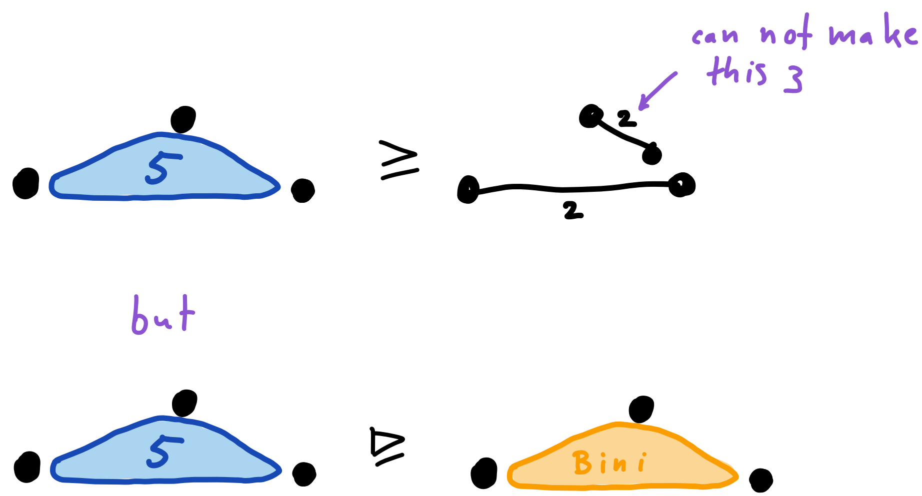


based on a method for faster matrix multiplication

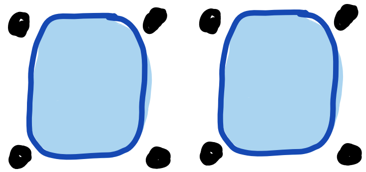
Since  only adjacent in two vertices, maybe all transformations consist of some exchange of max. ent. states?

NO!





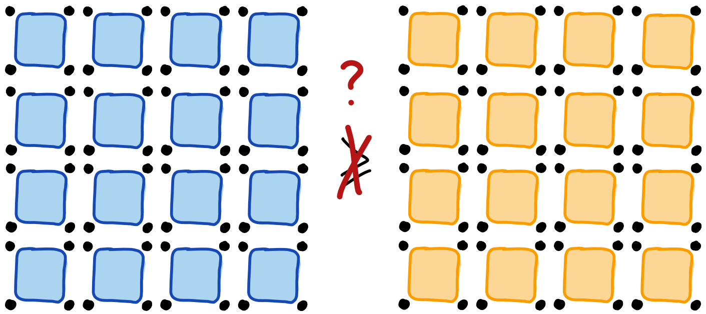
based on a method for faster matrix multiplication

Since  only adjacent in two vertices, maybe all transformations consist of some exchange of max. ent. states?

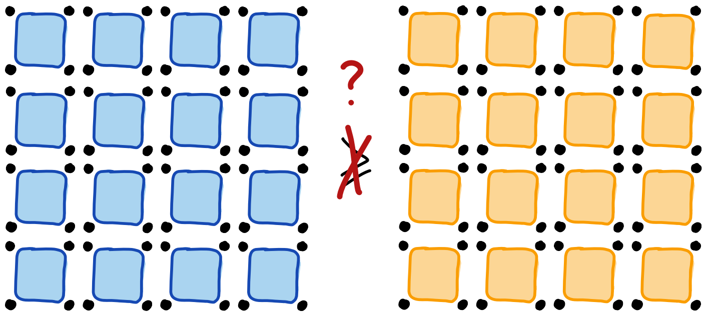
NO!

This shows that the resource theory of tensor networks is nontrivial (and one cannot just look at single plaquettes)

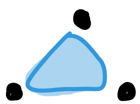
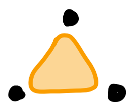
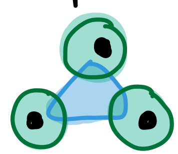
This also means it is in general complicated to show obstructions...

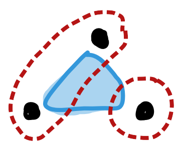
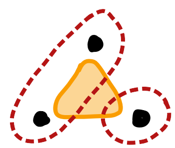


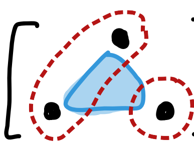
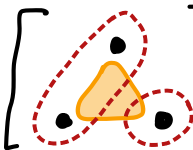
This also means it is in general complicated to show obstructions...



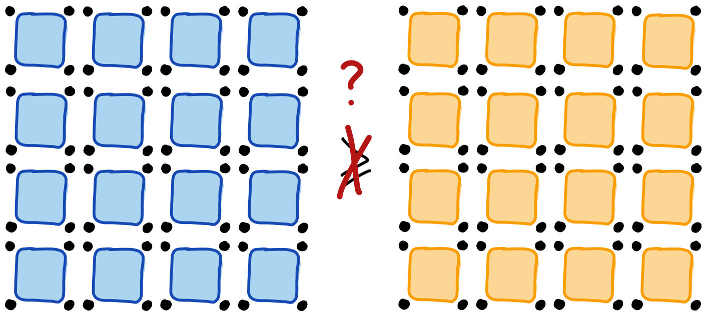
Basic method: based on bipartite rank

If  $\geq$ , so  $=$ 

then in particular  $\geq$  as 2-tensors

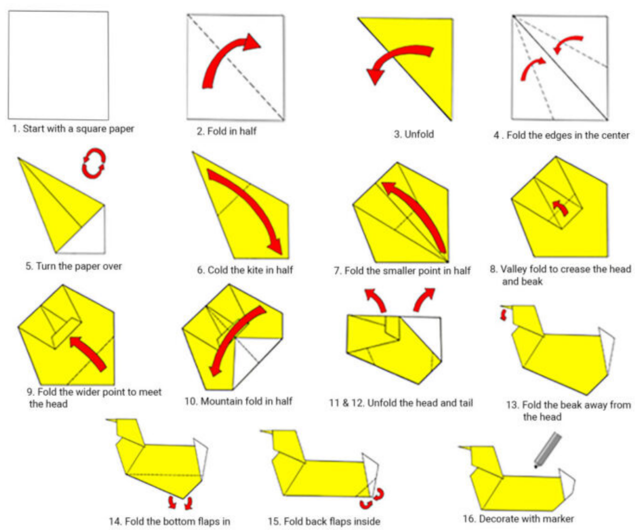
so rank[] $\geq$ rank[]

This also means it is in general complicated to show obstructions...

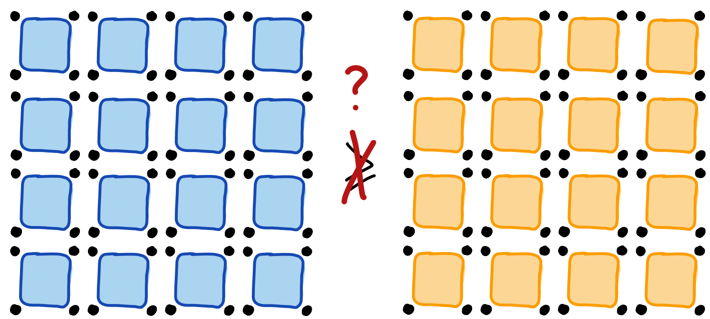


We give a general strategy:

① "Fold" the tensor network



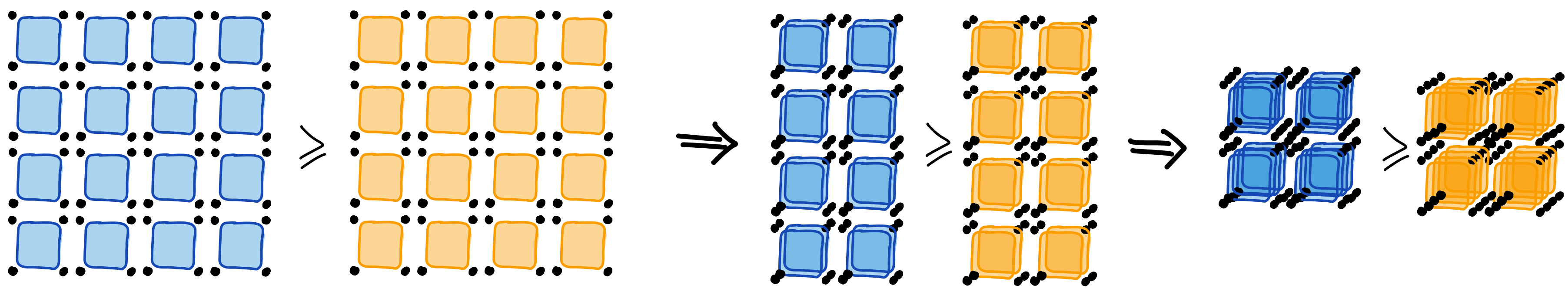
This also means it is in general complicated to show obstructions...



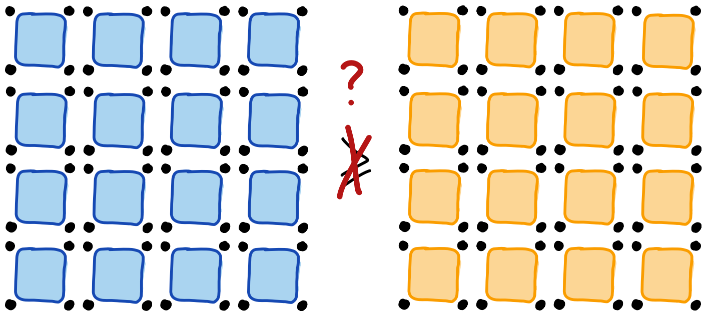
We give a general strategy:

1. "Fold" the tensor network

i.e. group together parties



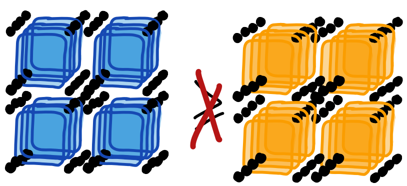
This also means it is in general complicated to show obstructions...



We give a general strategy:

2 Adapt & apply known obstruction methods

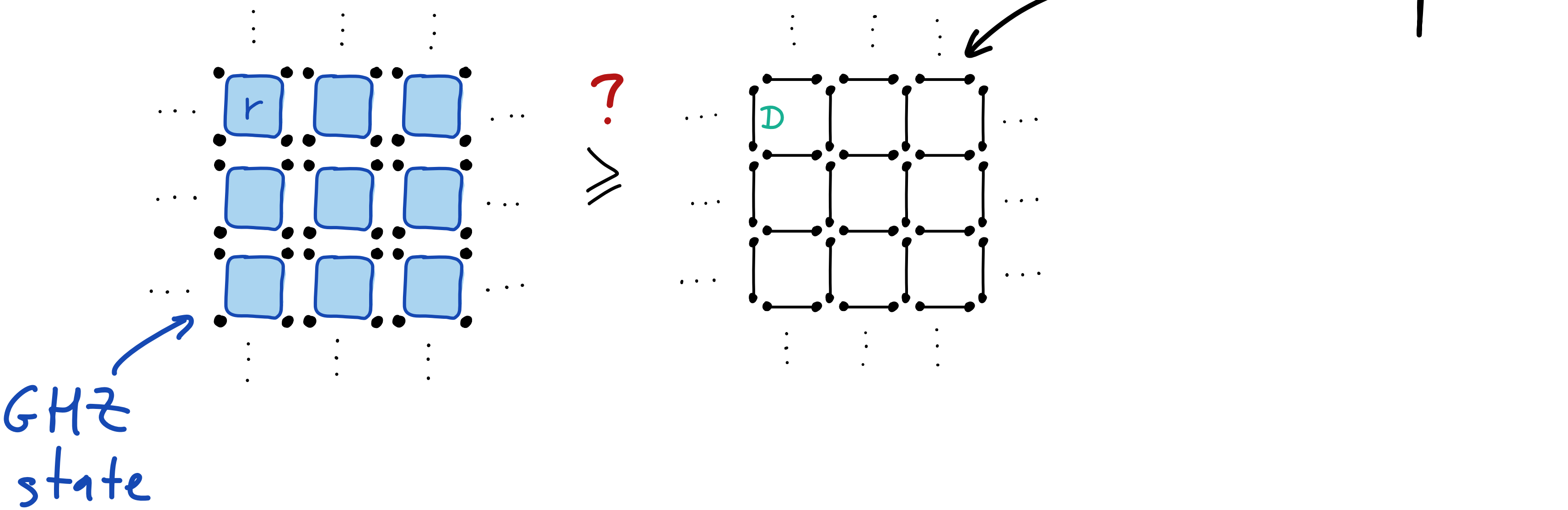
show



now a problem with only few parties

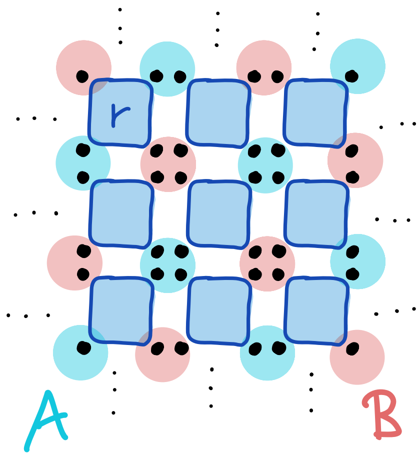
Simple example

When is

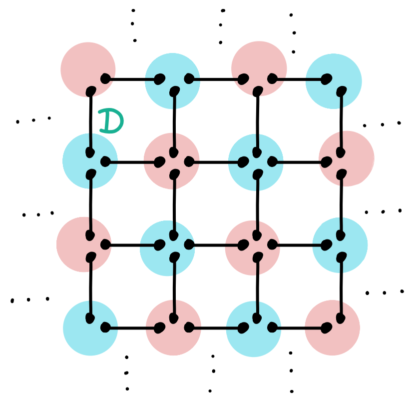


Simple example

When is

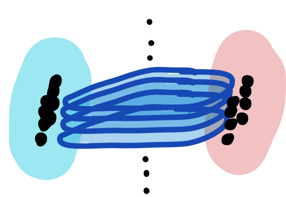


$\overset{?}{\geq}$



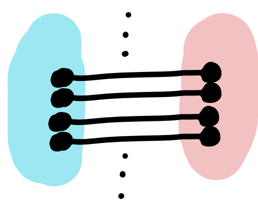
group into 2 parties

$\Downarrow$



rank $r^{|E|}$ between A, B

$\Downarrow$

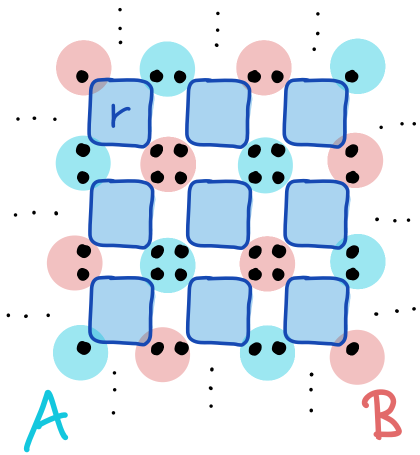


rank $D^{2|E|}$ between A, B

Simple example

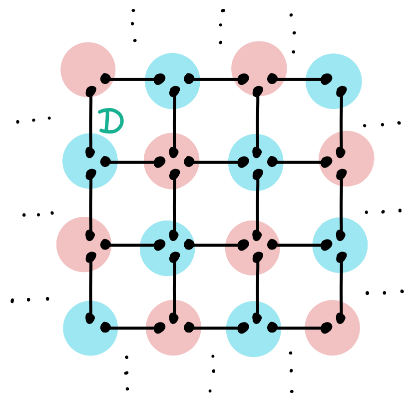
lower bound technique is just rank here!

When is

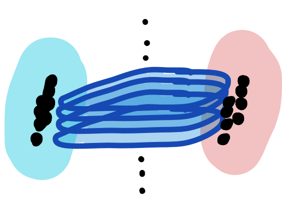


$\overset{?}{\geq}$

group into 2 parties



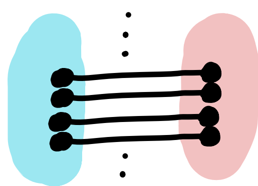
$\Downarrow$



rank $r^{|E|}$

between A, B

$\Downarrow$

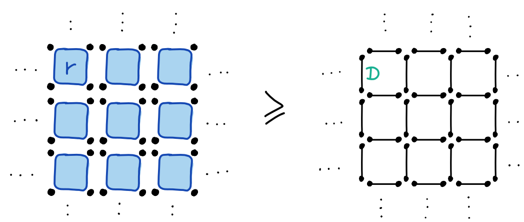


rank $D^{2|E|}$

between A, B

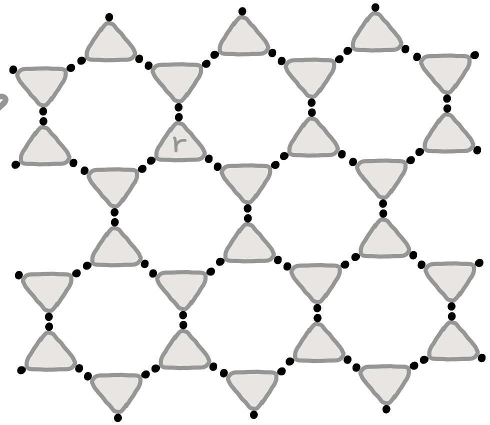
$\Longrightarrow \quad r \geq D^2$

and in that case

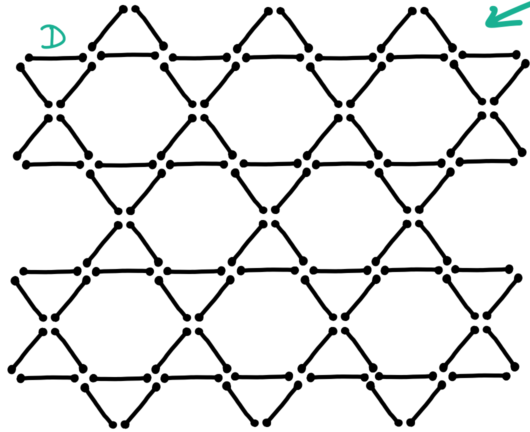


More complicated examples

GHZ
level r

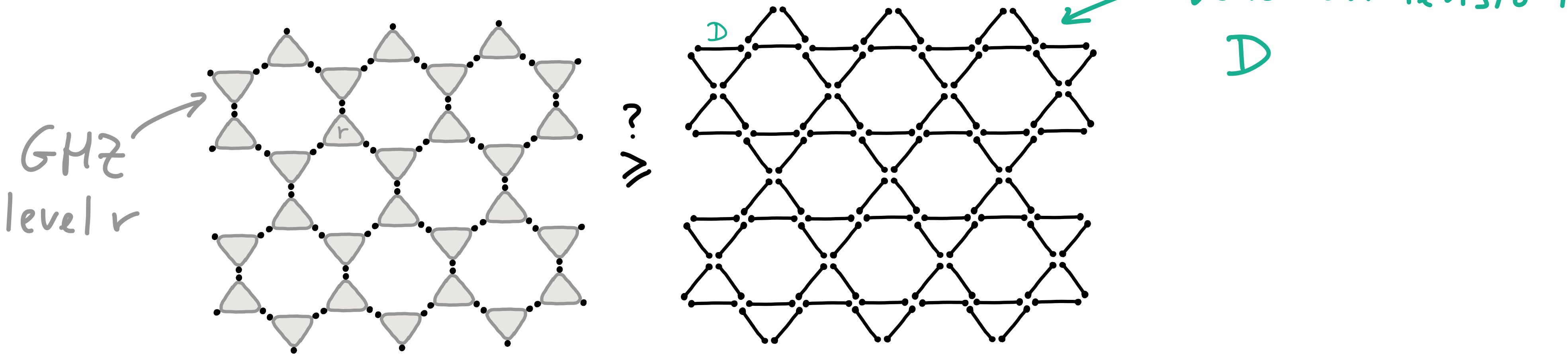


$\overset{?}{\geq}$

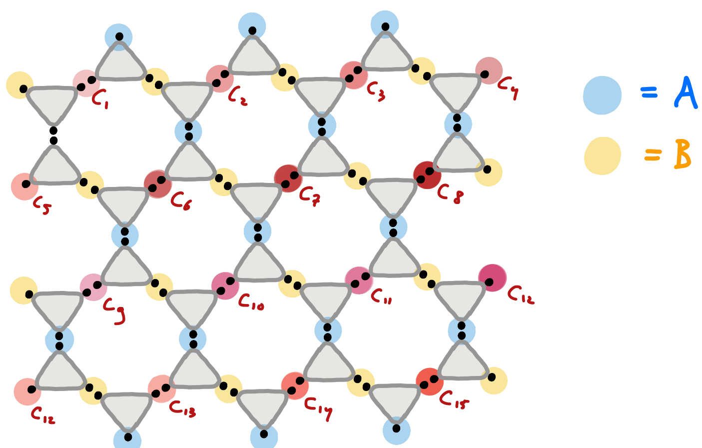


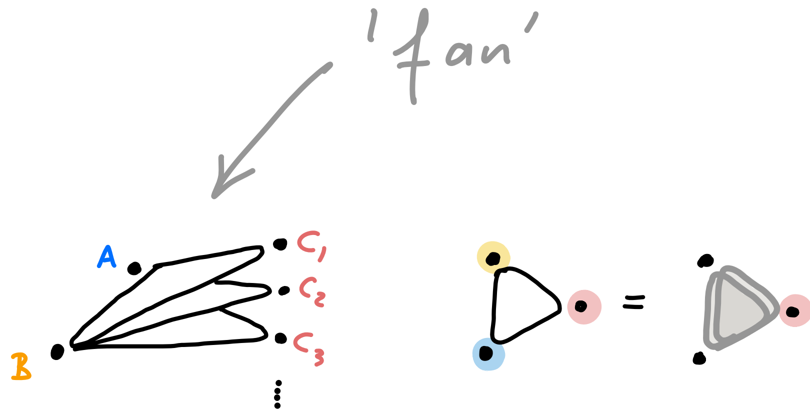
bond dimension
D

More complicated examples

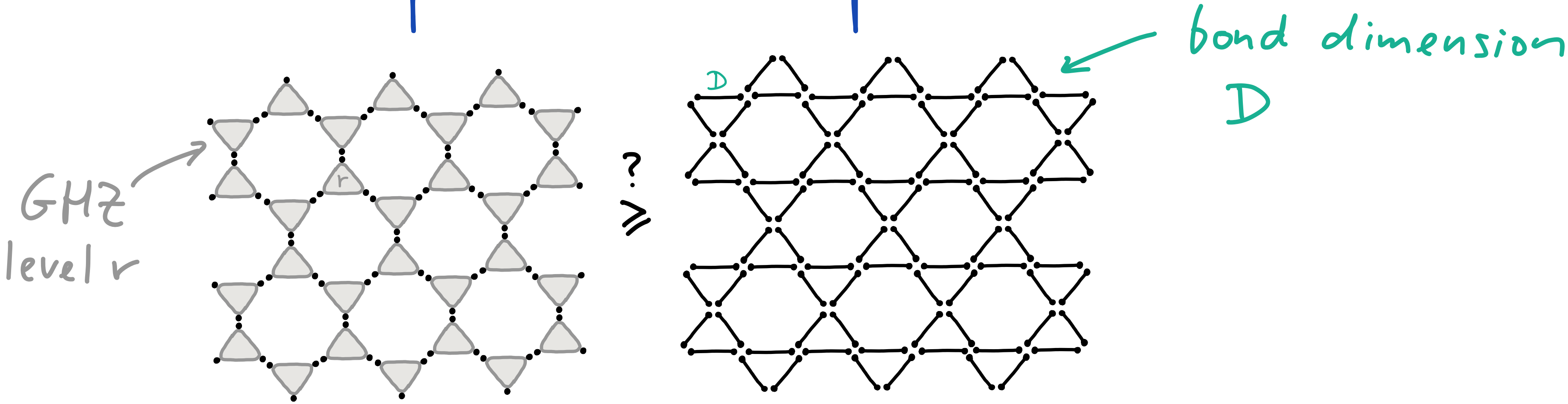


① Fold to a 'fan'

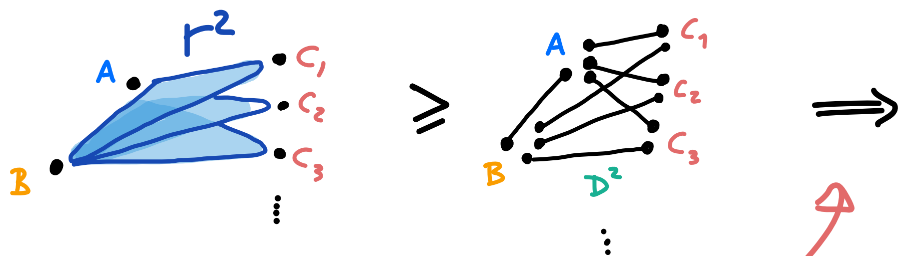




More complicated examples



② Extend known obstruction to fan

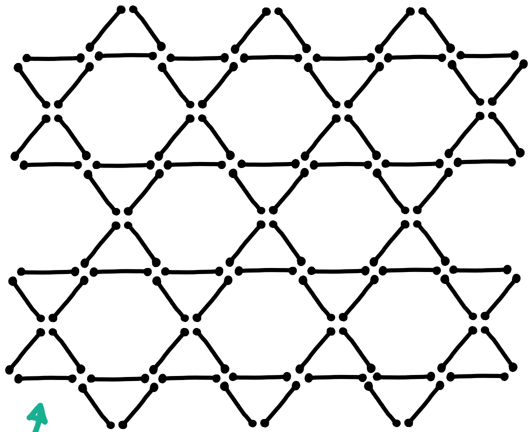


$$r^2 \geq 2D^4 - D^2$$

derives from Koszul flattening

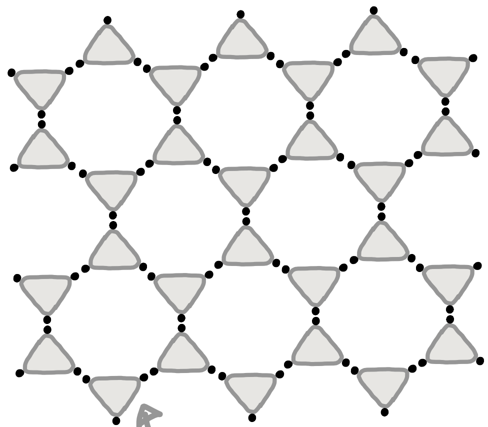
nontrivial bound! (talk JM Landsberg)

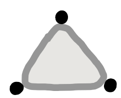
More complicated examples



D=2

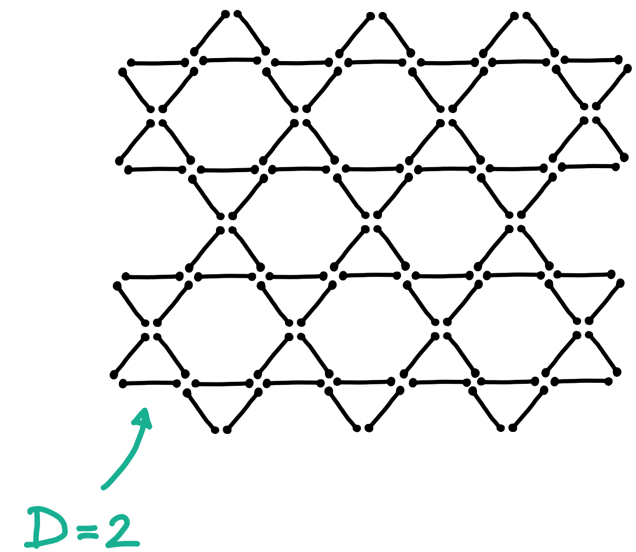
$\overset{?}{\geq}$



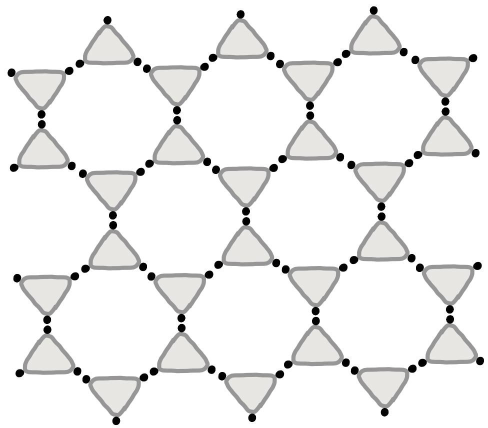
 $= |\lambda\rangle = \sum_{i,j,k=0}^{2} \varepsilon_{ijk} |ijk\rangle + |222\rangle$

related to RVB state,
example of spin liquid

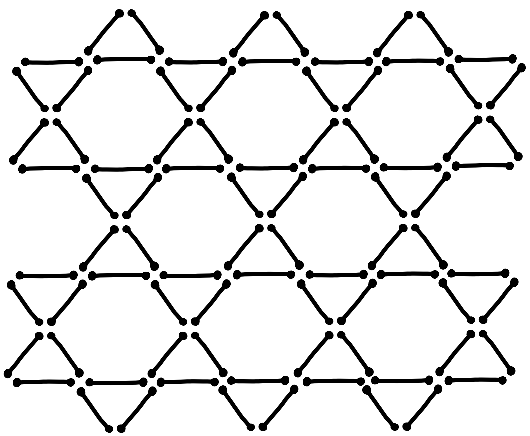
More complicated examples



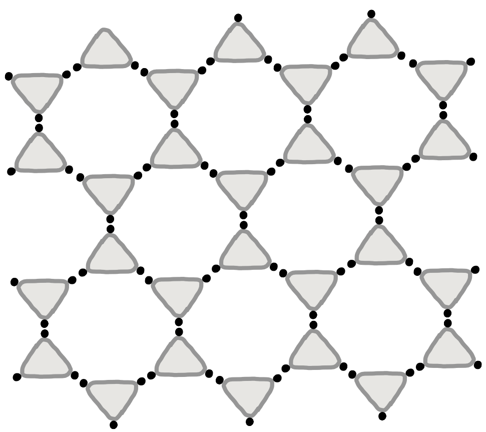
$\overset{?}{\geq}$

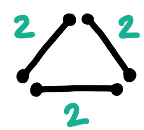


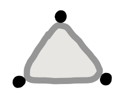
We know — Christandl, ...



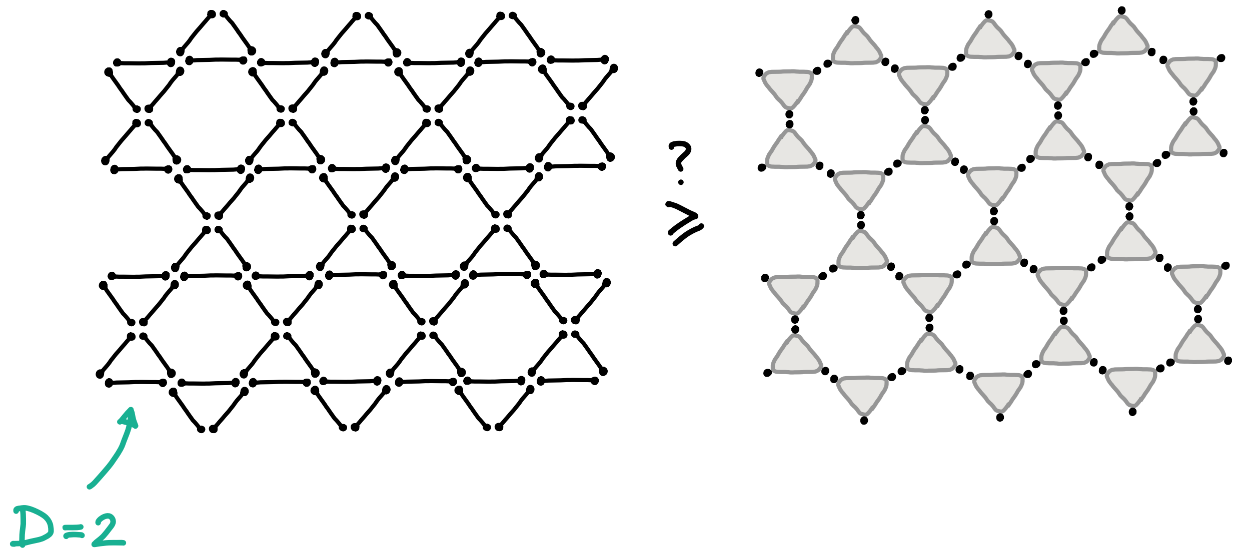
$\trianglerighteq$



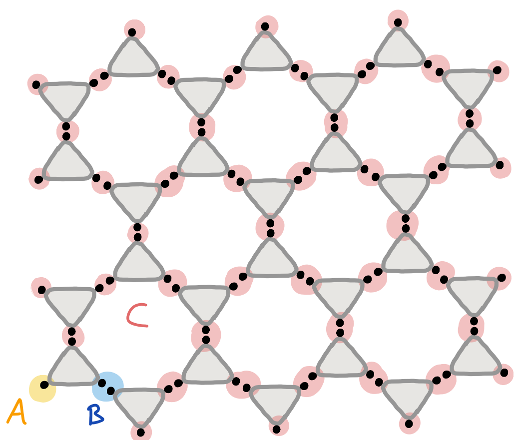
since 

$\trianglerighteq$ 

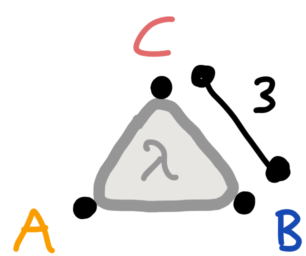
More complicated examples



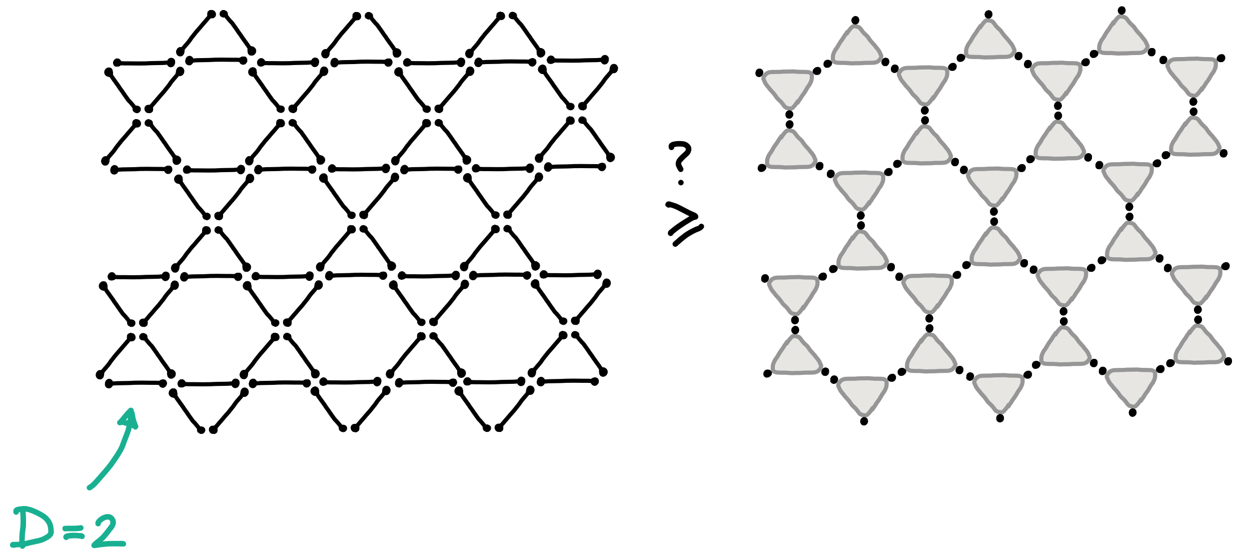
① Fold to a single plaquette



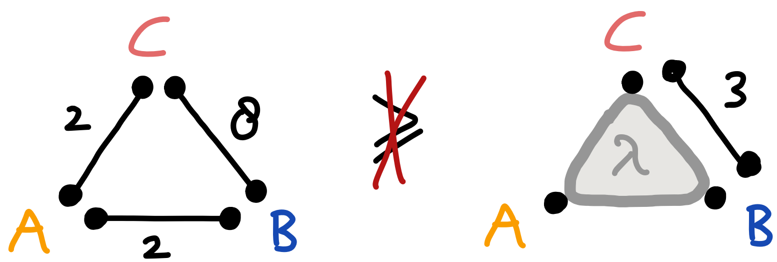
$\Longrightarrow$



More complicated examples



② Apply a version of the substitution method



Future directions

1 Connections to physics

Tensor networks are useful to study symmetries (in particular symmetry protected topological phases)

Does this resource theory have interesting consequences for SPT phases?

Future directions

2 Algorithms for (exact) TN contraction

If your algorithm is:

1. Choose an ordering of the edges
2. Contract edges in that order

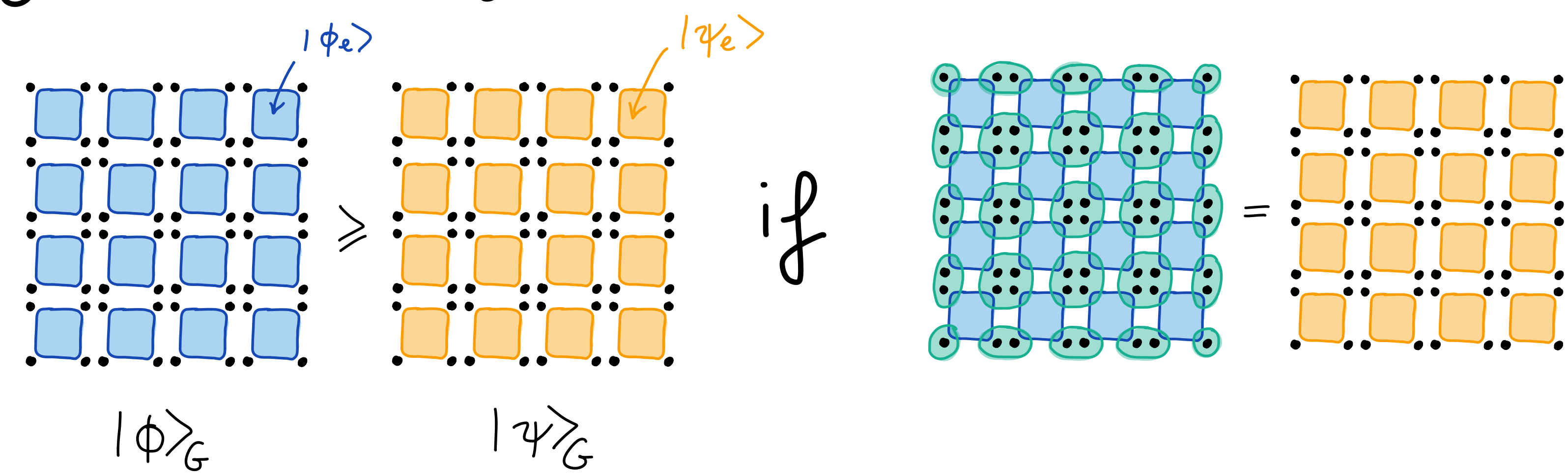
Complexity ~ matrix mult. of size $D^{\text{treewidth}}$

Can one find other algorithms, depending on tensor & hypergraph parameters?

Note: the problem is hard, so algos will be exponential

Conclusions

Resource theory for tensor networks with general entanglement structures



Conclusions

Resource theory for tensor networks with general entanglement structures

We develop:

1. $|\varphi\rangle_G \geq |\psi\rangle_G$ beyond single-plaquette transformations

2. Methods for showing $|\varphi\rangle_G \not\geq |\psi\rangle_G$

Conclusions

Resource theory for tensor networks with general entanglement structures

We develop:

1. $|\varphi\rangle_G \geq |\psi\rangle_G$ beyond single-plaquette transformations

2. Methods for showing $|\varphi\rangle_G \not\geq |\psi\rangle_G$

Key message elaborate methods from the study of matrix multiplication algorithms can be used to study tensor networks!