Characterizing the Bethe Partition Function of Double-Edge Factor Graphs via Graph Covers

Pascal O. Vontobel

Department of Information Engineering
The Chinese University of Hong Kong

Workshop on Inference Problems, August 31, 2020

Motivation

Consider the following setup:

- Let $\mathcal{A}$ be some finite, but large, set.
- Let g be a function over $\mathcal{A}$.

In this presentation we are interested in evaluating exactly or approximately expressions like

$$Z \triangleq \sum_{\mathbf{a} \in \mathcal{A}} g(\mathbf{a}) \, .$$

Motivation

Consider the following setup:

- Let $\mathcal{A}$ be some finite, but large, set.
- Let g be a function over $\mathcal{A}$.

In this presentation we are interested in evaluating exactly or approximately expressions like

$$Z \triangleq \sum_{\mathbf{a} \in \mathcal{A}} g(\mathbf{a}) \, .$$

- If $g : \mathcal{A} \to \mathbb{R}_{\geq 0}$ then evaluating Z is in general non-trivial. However, thanks to $g(\mathbf{a}) \geq 0$, the terms in the above summation

  "add up constructively"

  and so good approximations are often possible.

Motivation

Consider the following setup:

- Let $\mathcal{A}$ be some finite, but large, set.
- Let g be a function over $\mathcal{A}$.

In this presentation we are interested in evaluating exactly or approximately expressions like

$$Z \triangleq \sum_{\mathbf{a} \in \mathcal{A}} g(\mathbf{a}) \, .$$

- If $g : \mathcal{A} \to \mathbb{R}_{\geq 0}$ then evaluating Z is in general non-trivial.
  However, thanks to $g(\mathbf{a}) \geq 0$, the terms in the above summation

  "add up constructively"

  and so good approximations are often possible.

- If $g : \mathcal{A} \to \mathbb{C}$ then evaluating Z is even more challenging in general.
  In particular, because the real and the imaginary part of $g(\mathbf{a})$ can be both positive and negative, the terms in the above summation

  "add up constructively and destructively."

Motivation

Example: Let $n \in \mathbb{Z}_{>0}$, $\alpha \in \mathbb{C}$, and

$$Z_n \triangleq \sum_{\ell=0}^{n} \binom{n}{\ell} (1-\alpha)^{n-\ell} \alpha^{\ell} \, .$$

Motivation

Example: Let $n \in \mathbb{Z}_{>0}$, $\alpha \in \mathbb{C}$, and

$$Z_n \triangleq \sum_{\ell=0}^{n} \binom{n}{\ell} (1-\alpha)^{n-\ell} \alpha^{\ell} .$$

(Of course, in this particular case, we can easily evaluate Z_n exactly. Namely,

$$Z_n = ((1-\alpha)+\alpha)^n = 1^n = 1 .$$

The point of this example is to discuss bounding techniques that are more broadly applicable.)

Motivation

Example: Let $n \in \mathbb{Z}_{>0}$, $\alpha \in \mathbb{C}$, and

$$Z_n \triangleq \sum_{\ell=0}^{n} \binom{n}{\ell} (1-\alpha)^{n-\ell} \alpha^{\ell} .$$

(Of course, in this particular case, we can easily evaluate Z_n exactly. Namely,

$$Z_n = ((1-\alpha)+\alpha)^n = 1^n = 1 .$$

The point of this example is to discuss bounding techniques that are more broadly applicable.)

Let us define the following notation:

- $\mathcal{L}_n \triangleq \{0, 1, \ldots, n\}$.
- $c_{n,\ell} \triangleq \binom{n}{\ell} (1-\alpha)^{n-\ell} \alpha^{\ell}$, $\ell \in \mathcal{L}_n$.

With this,

$$Z_n \triangleq \sum_{\ell \in \mathcal{L}_n} c_{n,\ell} .$$

Motivation

Example (continued): Let us first consider the case $\alpha \in \mathbb{R}$ with $0 \leq \alpha \leq 1$.

Then all terms $c_{n,\ell}$ are non-negative real numbers, and so

$$\max_{\ell \in \mathcal{L}_n} c_{n,\ell} \;\leq\; {\color{red} Z_n} \;\leq\; |\mathcal{L}_n| \cdot \max_{\ell \in \mathcal{L}_n} c_{n,\ell} \,.$$

Motivation

Example (continued): Let us first consider the case $\alpha \in \mathbb{R}$ with $0 \leq \alpha \leq 1$.

Then all terms $c_{n,\ell}$ are non-negative real numbers, and so

$$\max_{\ell \in \mathcal{L}_n} c_{n,\ell} \;\leq\; Z_n \;\leq\; |\mathcal{L}_n| \cdot \max_{\ell \in \mathcal{L}_n} c_{n,\ell}\,.$$

Because $|\mathcal{L}_n| = n + 1$, this implies that

$$\max_{\ell \in \mathcal{L}_n} \frac{1}{n} \cdot \log(c_{n,\ell}) \;\leq\; \frac{1}{n} \cdot \log(Z_n) \;\leq\; \frac{\log(n+1)}{n} + \max_{\ell \in \mathcal{L}_n} \frac{1}{n} \cdot \log(c_{n,\ell})\,.$$

Motivation

Example (continued): Let us first consider the case $\alpha \in \mathbb{R}$ with $0 \leq \alpha \leq 1$.

Then all terms $c_{n,\ell}$ are non-negative real numbers, and so

$$\max_{\ell \in \mathcal{L}_n} c_{n,\ell} \;\leq\; Z_n \;\leq\; |\mathcal{L}_n| \cdot \max_{\ell \in \mathcal{L}_n} c_{n,\ell} \,.$$

Because $|\mathcal{L}_n| = n + 1$, this implies that

$$\max_{\ell \in \mathcal{L}_n} \frac{1}{n} \cdot \log(c_{n,\ell}) \;\leq\; \frac{1}{n} \cdot \log(Z_n) \;\leq\; \frac{\log(n+1)}{n} + \max_{\ell \in \mathcal{L}_n} \frac{1}{n} \cdot \log(c_{n,\ell}) \,.$$

Therefore, we can get a good approximation of Z_n

by finding the largest term in the summation.

Motivation

Example (continued): Let us first consider the case $\alpha \in \mathbb{R}$ with $0 \leq \alpha \leq 1$.

Then all terms $c_{n,\ell}$ are non-negative real numbers, and so

$$\max_{\ell \in \mathcal{L}_n} c_{n,\ell} \;\leq\; Z_n \;\leq\; |\mathcal{L}_n| \cdot \max_{\ell \in \mathcal{L}_n} c_{n,\ell}\,.$$

Because $|\mathcal{L}_n| = n + 1$, this implies that

$$\max_{\ell \in \mathcal{L}_n} \frac{1}{n} \cdot \log(c_{n,\ell}) \;\leq\; \frac{1}{n} \cdot \log(Z_n) \;\leq\; \frac{\log(n+1)}{n} + \max_{\ell \in \mathcal{L}_n} \frac{1}{n} \cdot \log(c_{n,\ell})\,.$$

Therefore, we can get a good approximation of Z_n

by finding the largest term in the summation.

In particular, in the limit $n \to \infty$, we obtain

$$\lim_{n \to \infty} \frac{1}{n} \cdot \log(Z_n) \;=\; \lim_{n \to \infty} \max_{\ell \in \mathcal{L}_n} \frac{1}{n} \cdot \log(c_{n,\ell})\,.$$

Motivation

Example (continued): Let us first consider the case $\alpha \in \mathbb{R}$ with $0 \leq \alpha \leq 1$.

Let $h(\alpha) \triangleq -\alpha \cdot \log(\alpha) - (1-\alpha) \cdot \log(1-\alpha)$ be the binary entropy function.

For simplicity of exposition, assume that $n\alpha \in \mathbb{Z}_{\geq 0}$. Because

$$\begin{aligned} \max_{\ell \in \mathcal{L}_n} c_{n,\ell} \;=\; c_{n,\ell}\big|_{\ell=\alpha n} &= \binom{n}{n\alpha}(1-\alpha)^{n(1-\alpha)}\alpha^{n\alpha} \\ &= \exp\big(nh(\alpha) + o(n)\big) \cdot \exp\big(-nh(\alpha)\big) \\ &= \exp(o(n)), \end{aligned}$$

we get

$$\frac{o(n)}{n} \leq \frac{1}{n} \cdot \log(Z_n) \leq \frac{\log(n+1)}{n} + \frac{o(n)}{n}.$$

Motivation

Example (continued): Let us first consider the case $\alpha \in \mathbb{R}$ with $0 \leq \alpha \leq 1$.

Let $h(\alpha) \triangleq -\alpha \cdot \log(\alpha) - (1-\alpha) \cdot \log(1-\alpha)$ be the binary entropy function.
For simplicity of exposition, assume that $n\alpha \in \mathbb{Z}_{\geq 0}$. Because

$$\begin{aligned} \max_{\ell \in \mathcal{L}_n} c_{n,\ell} \;=\; c_{n,\ell}\big|_{\ell = \alpha n} &= \binom{n}{n\alpha} (1-\alpha)^{n(1-\alpha)} \alpha^{n\alpha} \\ &= \exp\big(nh(\alpha) + o(n)\big) \cdot \exp\big(-nh(\alpha)\big) \\ &= \exp(o(n)) \,, \end{aligned}$$

we get

$$\frac{o(n)}{n} \;\leq\; \frac{1}{n} \cdot \log(Z_n) \;\leq\; \frac{\log(n+1)}{n} + \frac{o(n)}{n} \,.$$

In particular, in the limit $n \to \infty$, we obtain

$$\lim_{n \to \infty} \frac{1}{n} \cdot \log(Z_n) \;=\; 0 \,.$$

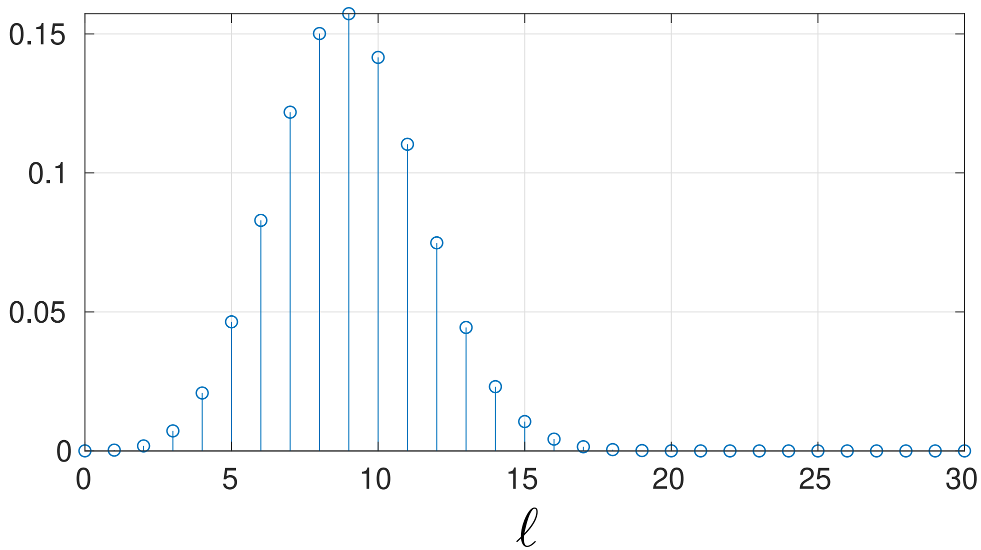
Motivation

Example (continued): Let us first consider the case $\alpha \in \mathbb{R}$ with $0 \leq \alpha \leq 1$.

Terms $c_{n,\ell}$ appearing in the sum

$$\color{red}{Z_n}\color{black} \;=\; \sum_{\ell=0}^{n} c_{n,\ell} \;=\; \sum_{\ell=0}^{n} \binom{n}{\ell} (1-\alpha)^{n-\ell} \alpha^{\ell}$$

for $n = 30,\ \alpha = 0.3$:



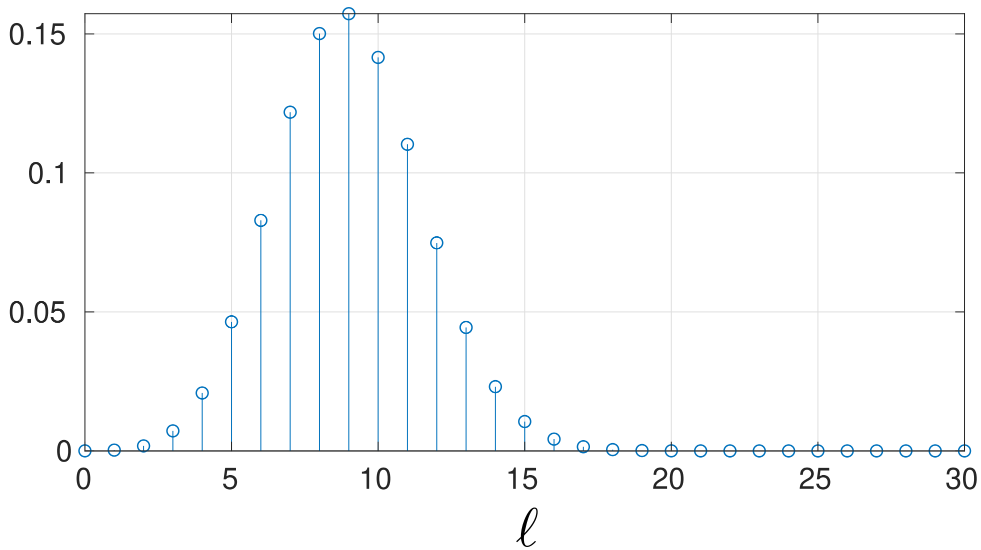
Motivation

Example (continued): Let us first consider the case $\alpha \in \mathbb{R}$ with $0 \leq \alpha \leq 1$.

Terms $c_{n,\ell}$ appearing in the sum

$$Z_n = \sum_{\ell=0}^{n} c_{n,\ell} = \sum_{\ell=0}^{n} \binom{n}{\ell} (1-\alpha)^{n-\ell} \alpha^{\ell}$$

for $n = 30, \ \alpha = 0.3$:



“Terms add up constructively.”

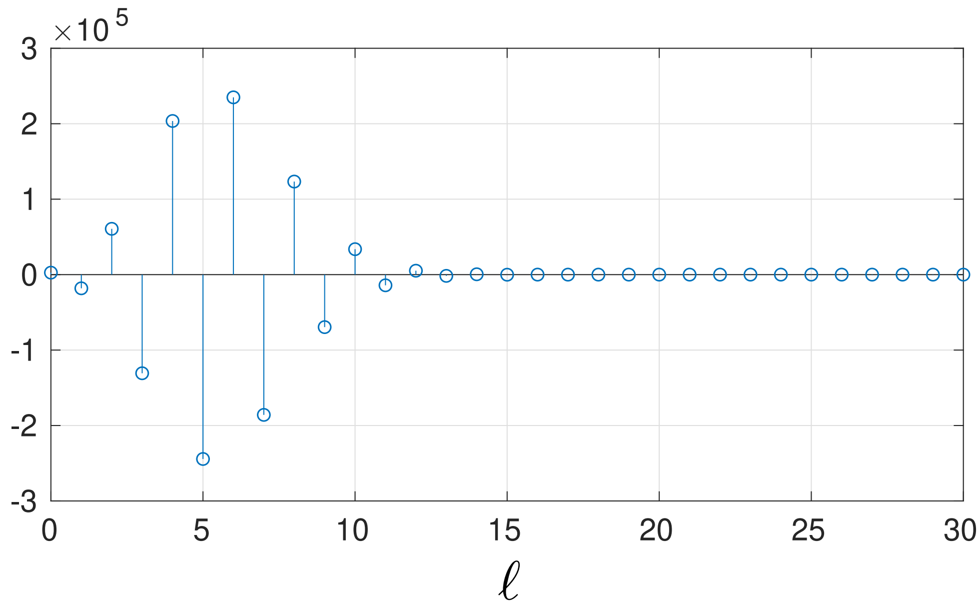
Motivation

Example (continued): Let us now consider the case $\alpha \in \mathbb{R}$ with $\alpha < 0$.

Terms $c_{n,\ell}$ appearing in the sum

$$\color{red}{Z_n} = \sum_{\ell=0}^{n} c_{n,\ell} = \sum_{\ell=0}^{n} \binom{n}{\ell} (1-\alpha)^{n-\ell} \alpha^{\ell}$$

for $n = 30,\ \alpha = -0.3$:



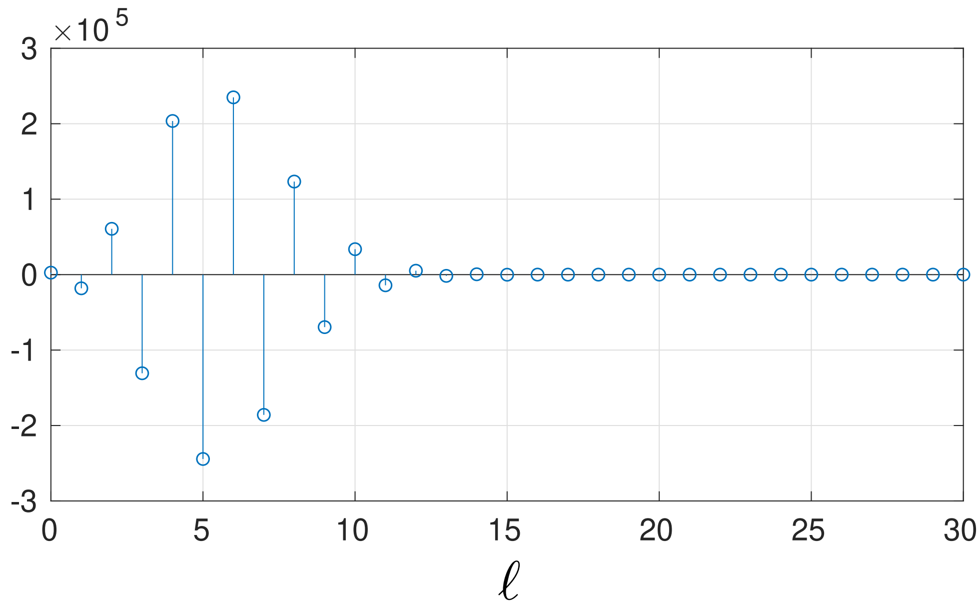
Motivation

Example (continued): Let us now consider the case $\alpha \in \mathbb{R}$ with $\alpha < 0$.

Terms $c_{n,\ell}$ appearing in the sum

$$\color{red}{Z_n} = \sum_{\ell=0}^{n} c_{n,\ell} = \sum_{\ell=0}^{n} \binom{n}{\ell} (1-\alpha)^{n-\ell} \alpha^{\ell}$$

for $n = 30,\ \alpha = -0.3$:



“Terms add up constructively and destructively.”

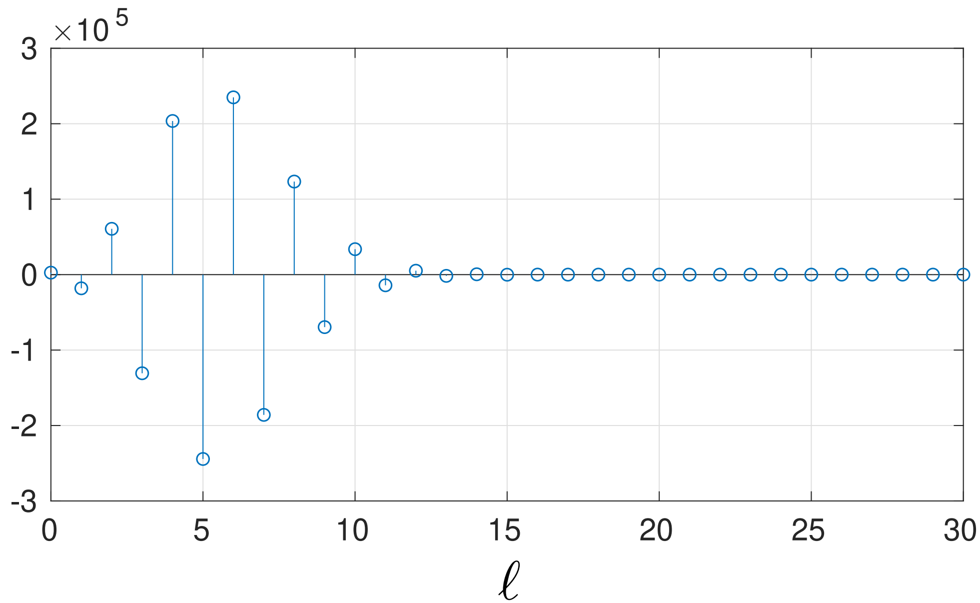
Motivation

Example (continued): Let us now consider the case $\alpha \in \mathbb{R}$ with $\alpha < 0$.

Terms $c_{n,\ell}$ appearing in the sum

$$\color{red}{Z_n} = \sum_{\ell=0}^{n} c_{n,\ell} = \sum_{\ell=0}^{n} \binom{n}{\ell} (1-\alpha)^{n-\ell} \alpha^{\ell}$$

for $n = 30, \ \alpha = -0.3$:



The term with largest magnitude gives a bad estimate of $\color{red}{Z_n}$.

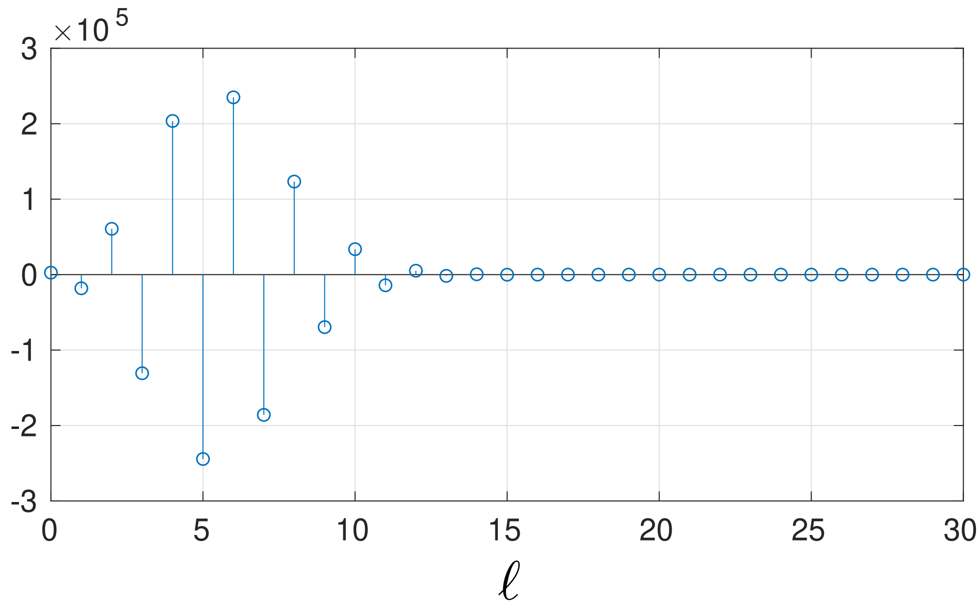
Motivation

Example (continued): Let us now consider the case $\alpha \in \mathbb{R}$ with $\alpha < 0$.

Terms $c_{n,\ell}$ appearing in the sum

$$Z_n = \sum_{\ell=0}^{n} c_{n,\ell} = \sum_{\ell=0}^{n} \binom{n}{\ell} (1-\alpha)^{n-\ell} \alpha^{\ell}$$

for $n = 30,\ \alpha = -0.3$:

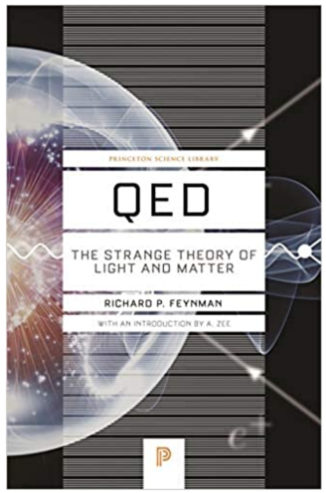


The term with largest magnitude does not even give the correct sign of Z_n.

Motivation

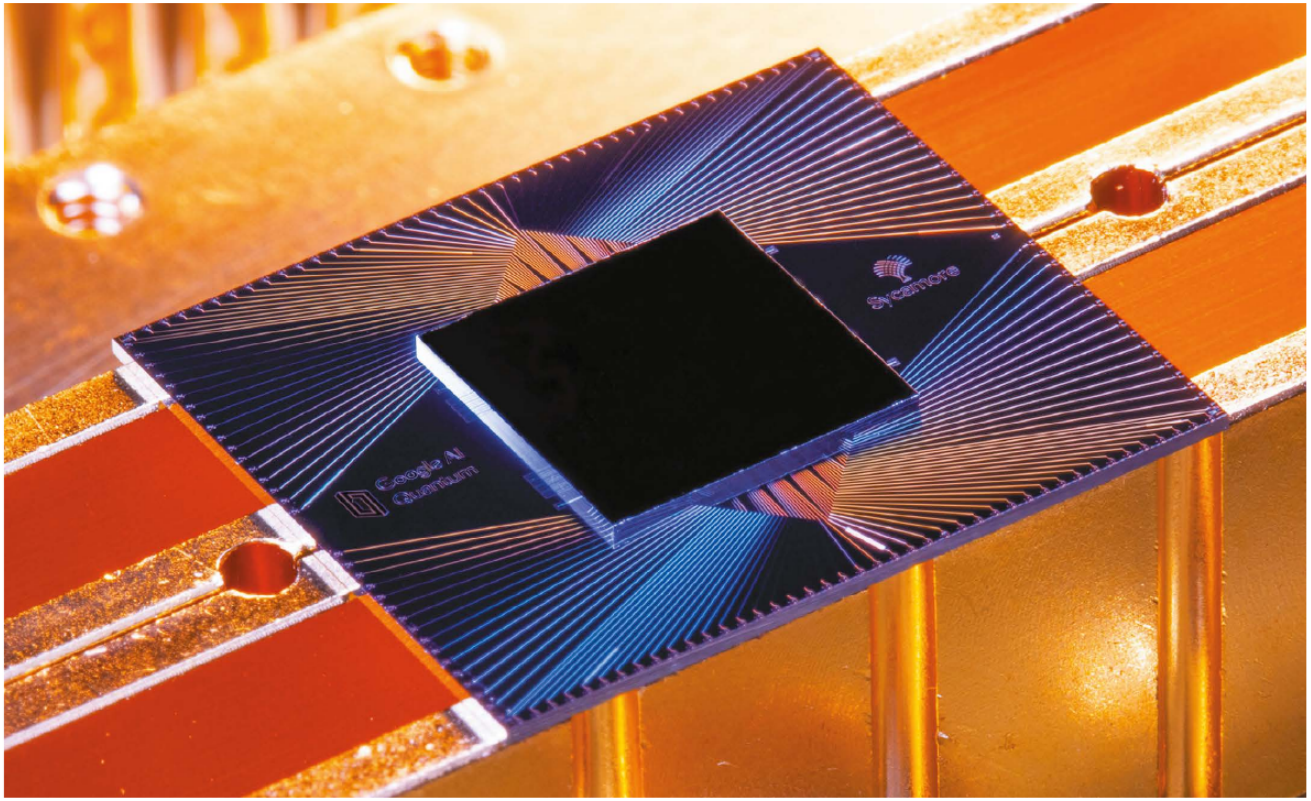
Motivation

R. P. Feynman
QED: The Strange Theory of Light and Matter
Princeton Science Library



Motivation

24 October 2019: Google publishes a paper **claiming quantum supremacy**



The Sycamore chip is composed of 54 qubits, each made of superconducting loops.

GOOGLE PUBLISHES LANDMARK QUANTUM SUPREMACY CLAIM

The company says that its quantum computer is the first to perform a calculation that would be practically impossible for a classical machine.

Overview

- Standard normal factor graphs (S-NFG):
  - Basics
  - A combinatorial interpretation of the Bethe partition sum, i.e., the Bethe approximation of the partition sum
- Double-edge normal factor graphs (DE-NFG):
  - Basics
  - A combinatorial interpretation of the Bethe partition sum, i.e., the Bethe approximation of the partition sum
- Conclusions / Outlook

Overview

- Standard normal factor graphs (S-NFG):
  - Basics
  - A combinatorial interpretation of the Bethe partition sum,
    i.e., the Bethe approximation of the partition sum **[known]**
- Double-edge normal factor graphs (DE-NFG):
  - Basics
  - A combinatorial interpretation of the Bethe partition sum,
    i.e., the Bethe approximation of the partition sum **[novel]**
- Conclusions / Outlook

Overview

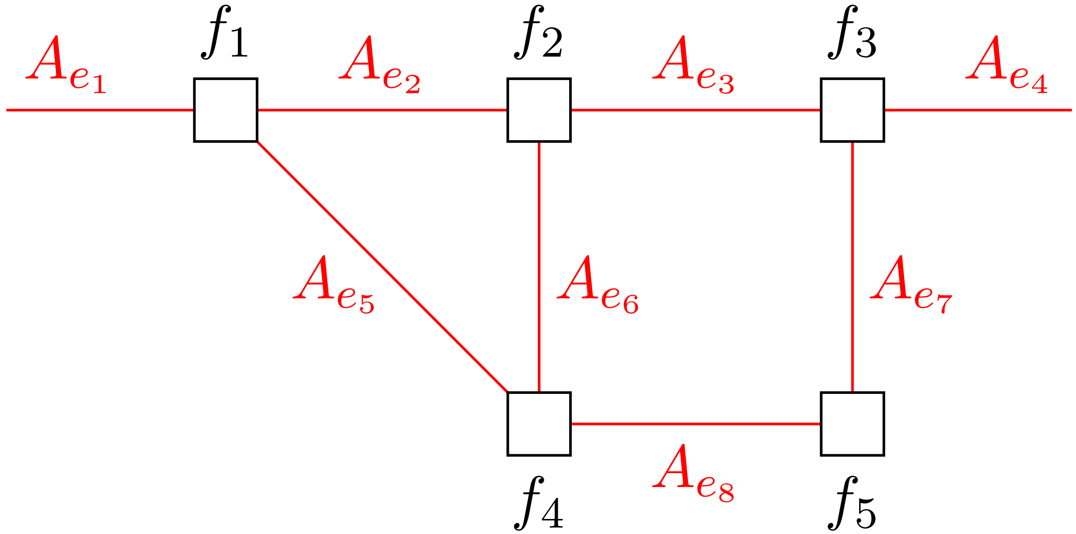
- Standard normal factor graphs (S-NFG):
  - Basics
  - A combinatorial interpretation of the Bethe partition sum, i.e., the Bethe approximation of the partition sum **[known]**
- Double-edge normal factor graphs (DE-NFG):
  - Basics
  - A combinatorial interpretation of the Bethe partition sum, i.e., the Bethe approximation of the partition sum **[novel]**
- Conclusions / Outlook

This presentation is based on joint work with Yuwen HUANG (CUHK).

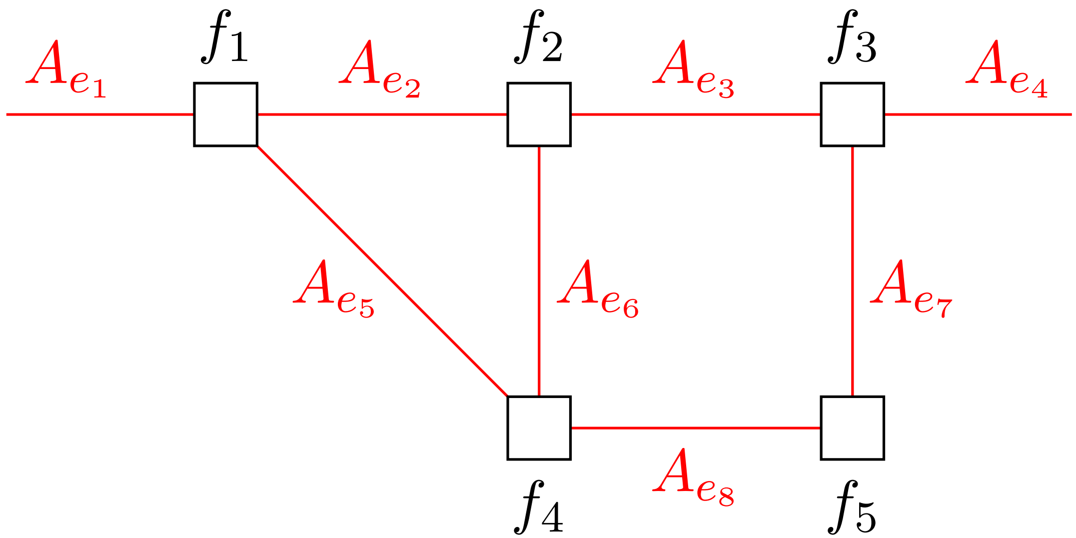
Y. Huang and P. O. Vontobel", “Characterizing the Bethe partition function of double-edge factor graphs via graph covers,” ISIT 2020. [Longer version in preparation.]

Standard normal factor graphs (S-NFGs)

Standard Normal Factor Graph (S-NFG)



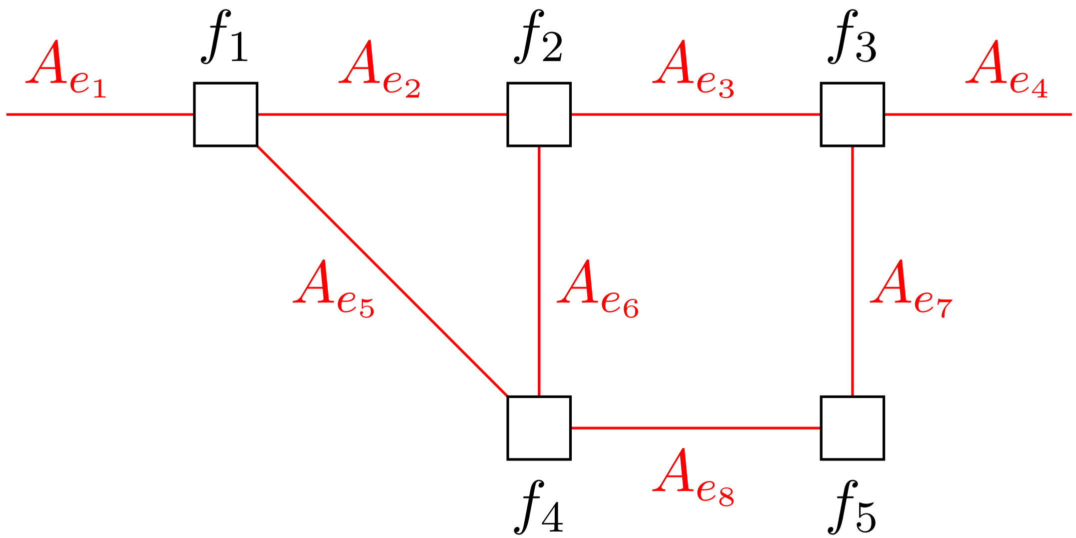
Standard Normal Factor Graph (S-NFG)



Global function:

$$g(a_{e_1}, \ldots, a_{e_8}) \triangleq f_1(a_{e_1}, a_{e_2}, a_{e_5}) \cdot f_2(a_{e_2}, a_{e_3}, a_{e_6}) \cdot f_3(a_{e_3}, a_{e_4}, a_{e_7}) \cdot f_4(a_{e_5}, a_{e_6}, a_{e_8}) \cdot f_5(a_{e_7}, a_{e_8})$$

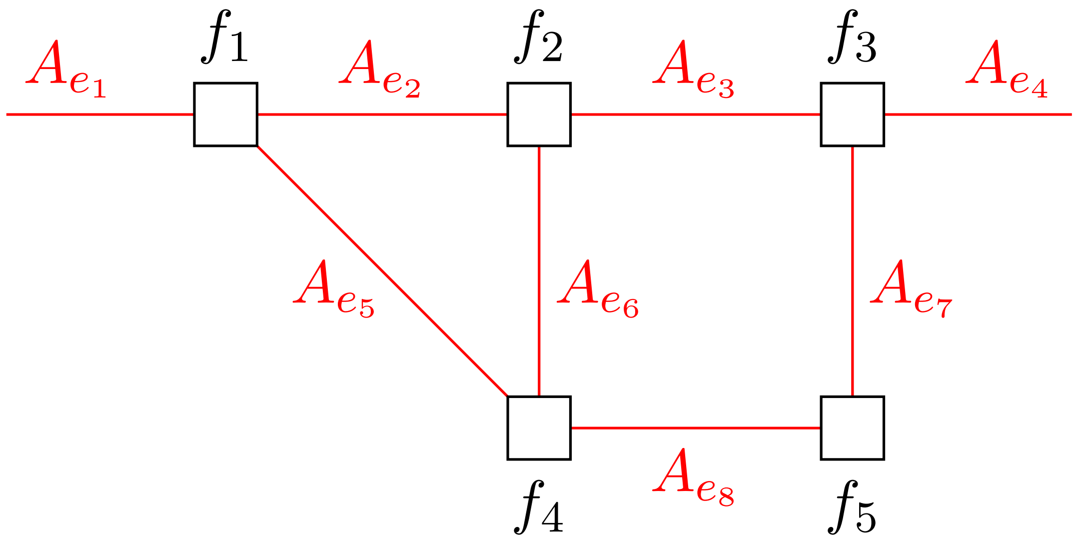
Standard Normal Factor Graph (S-NFG)



Global function:

$$g(\mathbf{a}) \triangleq \prod_{f} f(\mathbf{a}_f)$$

Standard Normal Factor Graph (S-NFG)



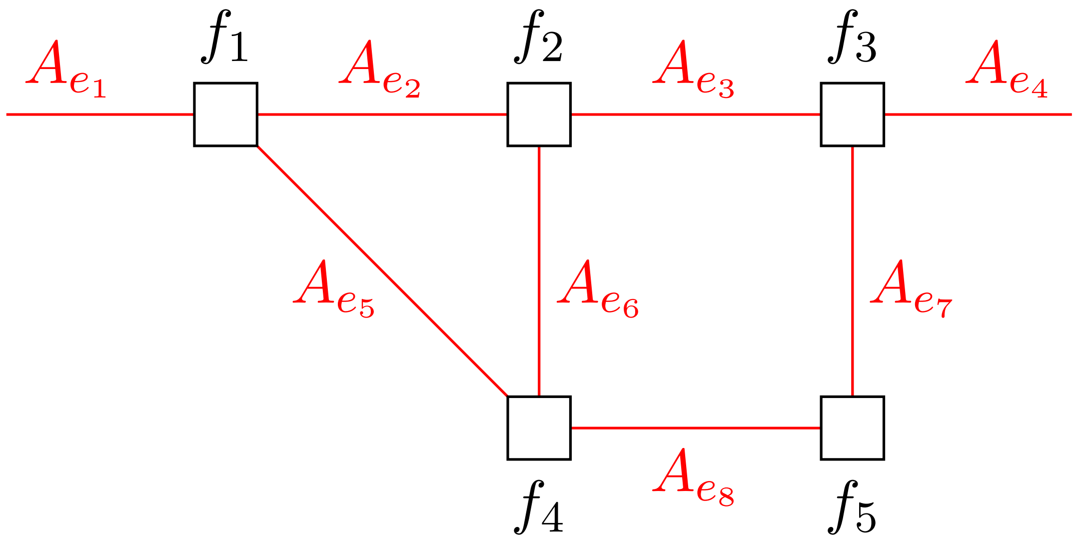
Global function:

$$g(\mathbf{a}) \triangleq \prod_{f} f(\mathbf{a}_f)$$

Partition sum:

$$Z \triangleq \sum_{\mathbf{a}} g(\mathbf{a})$$

Standard Normal Factor Graph (S-NFG)



Global function:

$$g(\mathbf{a}) \triangleq \prod_{f} f(\mathbf{a}_f)$$

Assumption from here on:

$$f(\mathbf{a}_f) \geq 0 \quad \forall f, \forall \mathbf{a}_f$$

Partition sum:

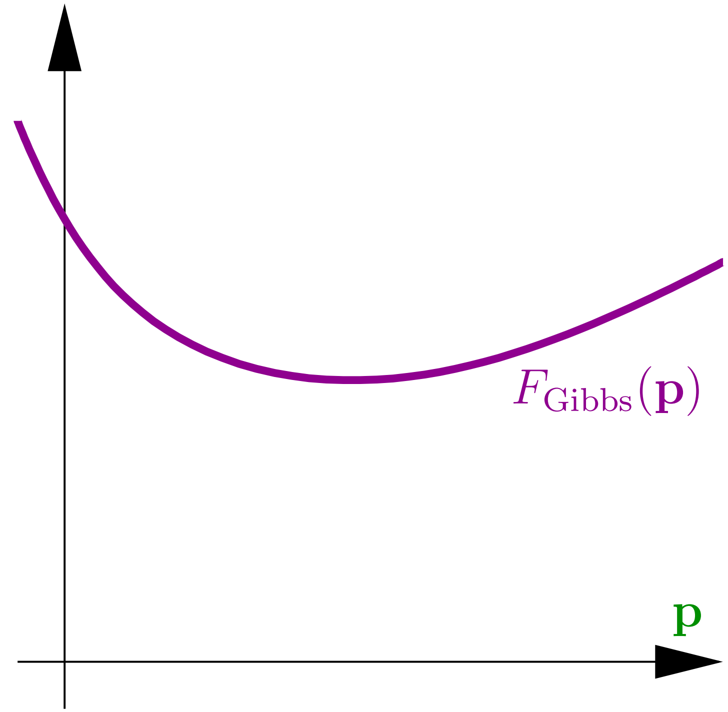
$$Z \triangleq \sum_{\mathbf{a}} g(\mathbf{a})$$

The Gibbs free energy function

Gibbs Free Energy Function

The Gibbs free energy function

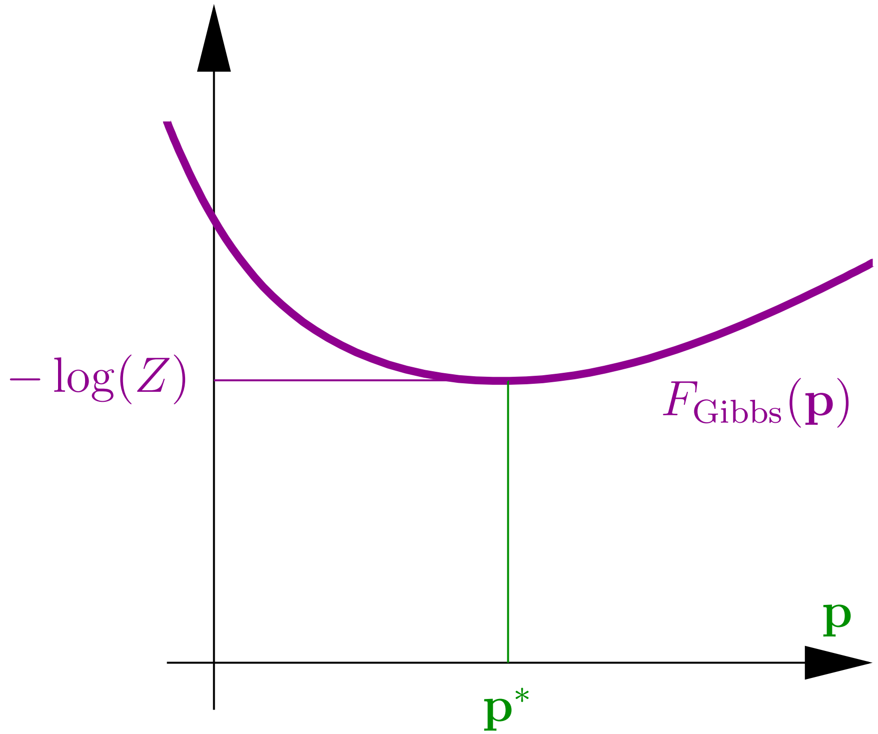
$$F_{\text{Gibbs}}(\mathbf{p}) \triangleq -\sum_{\mathbf{a}} p_{\mathbf{a}} \cdot \log\big(g(\mathbf{a})\big) + \sum_{\mathbf{a}} p_{\mathbf{a}} \cdot \log(p_{\mathbf{a}}).$$



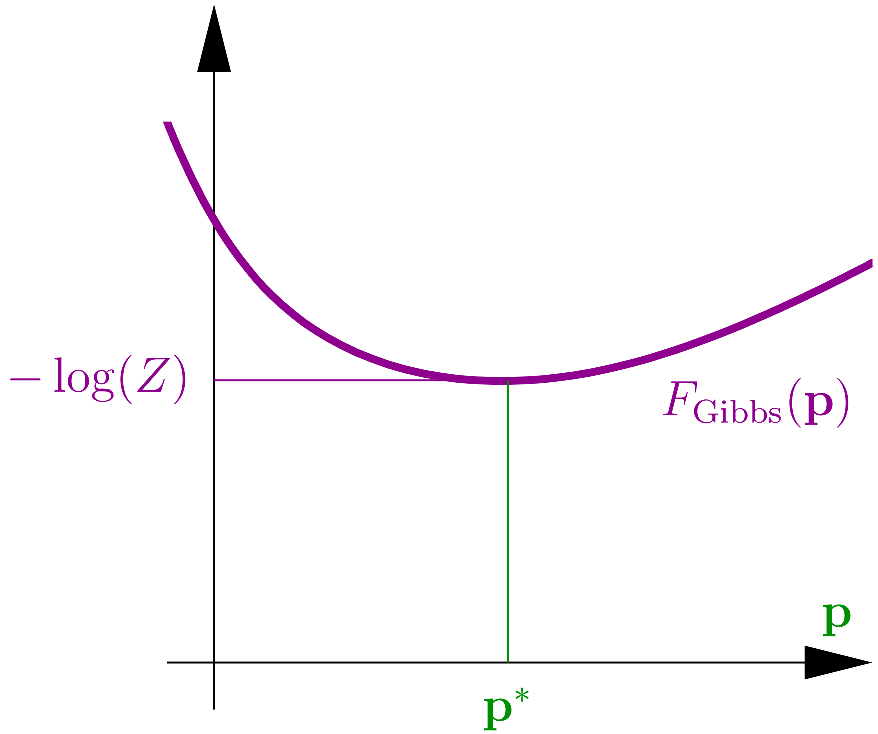
Gibbs Free Energy Function

The Gibbs free energy function

$$F_{\mathrm{Gibbs}}(\mathbf{p}) \triangleq -\sum_{\mathbf{a}} p_{\mathbf{a}} \cdot \log\big(g(\mathbf{a})\big) + \sum_{\mathbf{a}} p_{\mathbf{a}} \cdot \log(p_{\mathbf{a}}).$$



Gibbs Free Energy Function



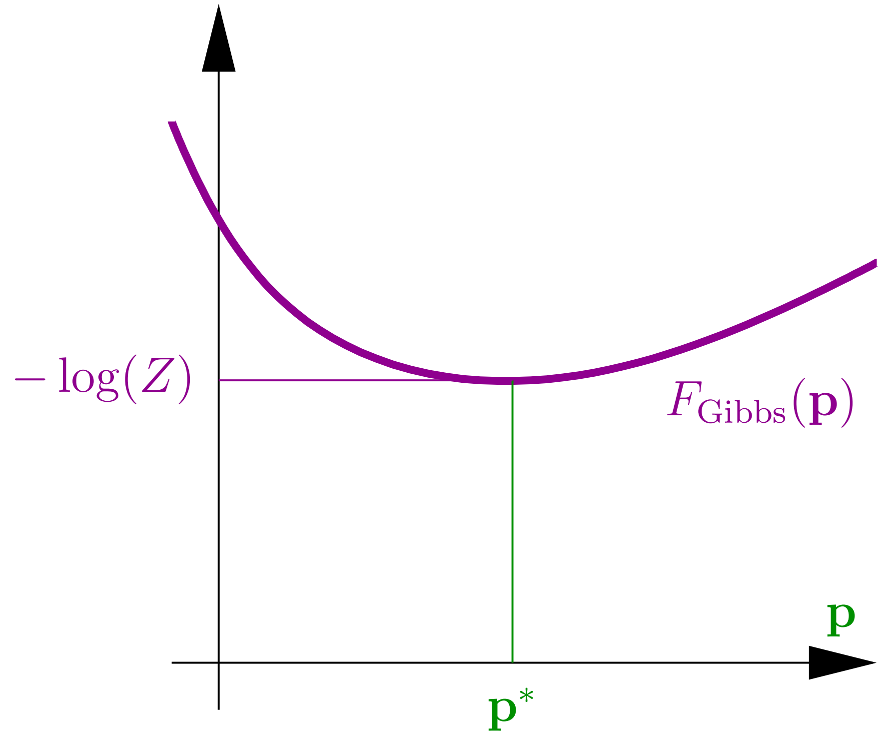
The Gibbs free energy function

$$F_{\text{Gibbs}}(\mathbf{p}) \triangleq -\sum_{\mathbf{a}} p_{\mathbf{a}} \cdot \log\left(g(\mathbf{a})\right) + \sum_{\mathbf{a}} p_{\mathbf{a}} \cdot \log(p_{\mathbf{a}}).$$

is defined such that its minimal value is related to the partition function:

$$Z = \exp\left(-\min_{\mathbf{p}} F_{\text{Gibbs}}(\mathbf{p})\right).$$

Gibbs Free Energy Function



The Gibbs free energy function

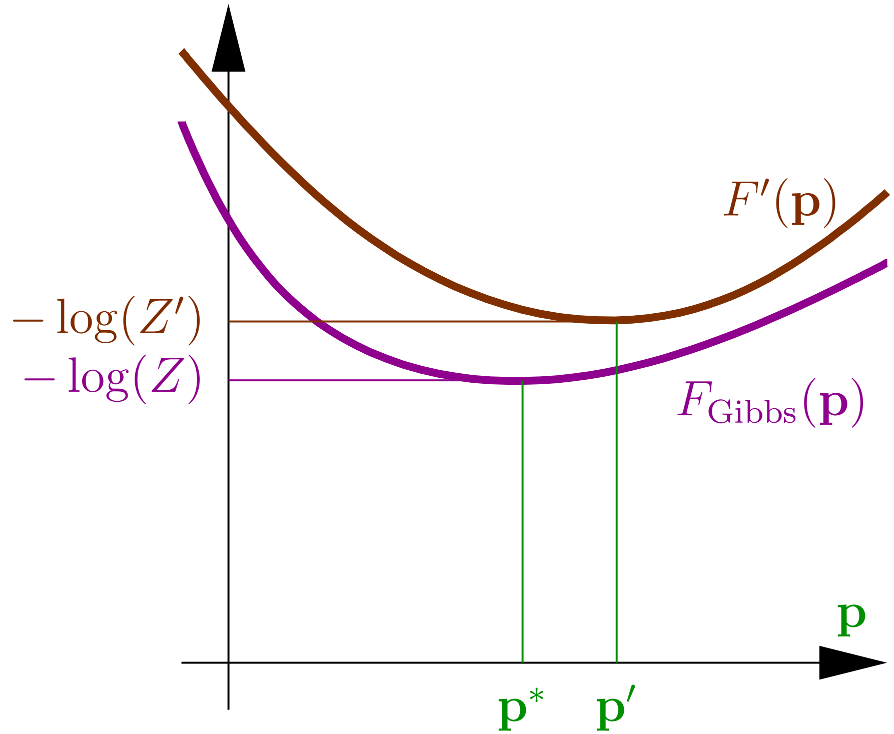
$$F_{\text{Gibbs}}(\mathbf{p}) \triangleq -\sum_{\mathbf{a}} p_{\mathbf{a}} \cdot \log\left(g(\mathbf{a})\right) + \sum_{\mathbf{a}} p_{\mathbf{a}} \cdot \log(p_{\mathbf{a}}).$$

is defined such that its minimal value is related to the partition function:

$$Z = \exp\left(-\min_{\mathbf{p}} F_{\text{Gibbs}}(\mathbf{p})\right).$$

Nice, but it does not yield any computational savings by itself.

Gibbs Free Energy Function



The Gibbs free energy function

$$F_{\text{Gibbs}}(\mathbf{p}) \triangleq -\sum_{\mathbf{a}} p_{\mathbf{a}} \cdot \log\left(g(\mathbf{a})\right) + \sum_{\mathbf{a}} p_{\mathbf{a}} \cdot \log(p_{\mathbf{a}}).$$

is defined such that its minimal value is related to the partition function:

$$Z = \exp\left(-\min_{\mathbf{p}} F_{\text{Gibbs}}(\mathbf{p})\right).$$

But it suggests other optimization schemes.

The Bethe approximation

Bethe Approximation

The Bethe approximation to the Gibbs free energy function yields such an alternative optimization scheme.

Bethe Approximation

The Bethe approximation to the Gibbs free energy function yields such an alternative optimization scheme.

This approximation is interesting because of the following theorem:

Theorem (Yedidia/Freeman/Weiss, 2000):

Fixed points of the sum-product algorithm (SPA)

correspond to

stationary points of the Bethe free energy function.

Bethe Approximation

The Bethe approximation to the Gibbs free energy function yields such an alternative optimization scheme.

This approximation is interesting because of the following theorem:

Theorem (Yedidia/Freeman/Weiss, 2000):

Fixed points of the sum-product algorithm (SPA)

correspond to

stationary points of the Bethe free energy function.

Definition:

We define the Bethe approximation Z_{Bethe} of the partition sum Z to be

$$Z_{\mathrm{Bethe}} \triangleq \exp\left(-\min_{\boldsymbol{\beta}} F_{\mathrm{Bethe}}(\boldsymbol{\beta})\right).$$

Bethe Approximation

Primal formulation:

$$Z_{\mathrm{Bethe}} \triangleq \exp\left(-\min_{\boldsymbol{\beta}} F_{\mathrm{Bethe}}(\boldsymbol{\beta})\right).$$

Bethe Approximation

Primal formulation:

$$Z_{\text{Bethe}} \triangleq \exp\left(- \min_{\boldsymbol{\beta}} F_{\text{Bethe}}(\boldsymbol{\beta}) \right) .$$

Pseudo-dual formulation:

$$Z_{\text{Bethe}} = \max_{\substack{\text{SPA LLR messages} \\ \text{fixed point } \boldsymbol{\lambda}}} \frac{\prod_{f \in \mathcal{F}} Z_f(\boldsymbol{\lambda})}{\prod_{e \in \mathcal{E}_{\text{full}}} Z_e(\boldsymbol{\lambda})} ,$$

where

$$Z_f(\boldsymbol{\lambda}) \triangleq \sum_{\mathbf{a}_f} f(\mathbf{a}_f) \cdot \prod_{e \in \partial f} \mathrm{e}^{-\lambda_{e \to f}(a_{f,e})} , \quad f \in \mathcal{F} ,$$

$$Z_e(\boldsymbol{\lambda}) \triangleq \sum_{a_e} \mathrm{e}^{-\lambda_{e \to f}(a_e) - \lambda_{e \to f'}(a_e)} , \qquad e = (f, f') \in \mathcal{E}_{\text{full}} .$$

Bethe Approximation

Primal formulation:

$$Z_{\text{Bethe}} = \exp\left(- \min_{\boldsymbol{\beta}} \; F_{\text{Bethe}}(\boldsymbol{\beta}) \right) .$$

Pseudo-dual formulation:

$$Z_{\text{Bethe}} = \max_{\substack{\text{SPA messages} \\ \text{fixed point } \boldsymbol{\mu}}} \; \frac{\prod_{f \in \mathcal{F}} Z_f(\boldsymbol{\mu})}{\prod_{e \in \mathcal{E}_{\text{full}}} Z_e(\boldsymbol{\mu})} ,$$

where

$$Z_f(\boldsymbol{\mu}) \triangleq \sum_{\mathbf{a}_f} f(\mathbf{a}_f) \cdot \prod_{e \in \partial f} \mu_{e \to f}(a_{f,e}) , \quad f \in \mathcal{F} ,$$

$$Z_e(\boldsymbol{\mu}) \triangleq \sum_{a_e} \mu_{e \to f}(a_e) \cdot \mu_{e \to f'}(a_e) , \qquad e = (f, f') \in \mathcal{E}_{\text{full}} .$$

Bethe Approximation

This talk it about better understanding approximations given by the Bethe approximation / SPA for factor graphs.

Bethe Approximation

This talk it about better understanding approximations given by the Bethe approximation / SPA for factor graphs.

Some areas where factor graphs and the Bethe approximation / SPA have turned out to be useful:

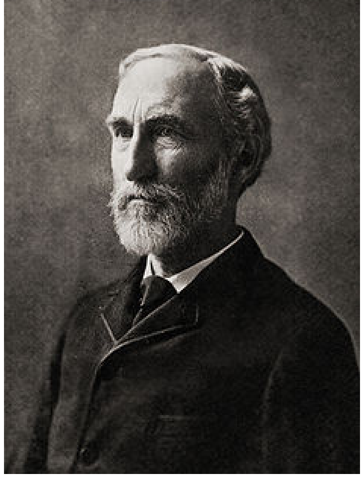
- Low-density parity-check (LDPC) and turbo codes.
- Counting patterns in constrained coding.
- Some image processing tasks.
  (E.g., early vision problems such as stereo, optical flow, and image restoration.)
- Estimating the permanent of a non-negative matrix.
- Pattern maximum likelihood (PML) estimate.
  (PML estimate: estimating sorted p.m.f.s based on relatively few samples.)
- Etc.

The partition sum

and its Bethe approximation

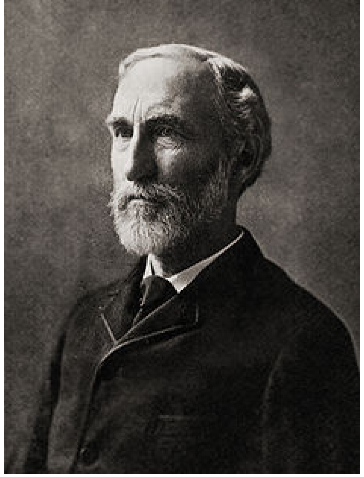
combinatorial	$Z \triangleq \sum_{\mathbf{a}} g(\mathbf{a})$	$Z_{\text{Bethe}} = ???$
analytical	$Z = \exp\left(-\min_{\mathbf{p}} F_{\text{Gibbs}}(\mathbf{p})\right)$	$Z_{\text{Bethe}} \triangleq \exp\left(-\min_{\boldsymbol{\beta}} F_{\text{Bethe}}(\boldsymbol{\beta})\right)$

(Temperature $T = 1$)



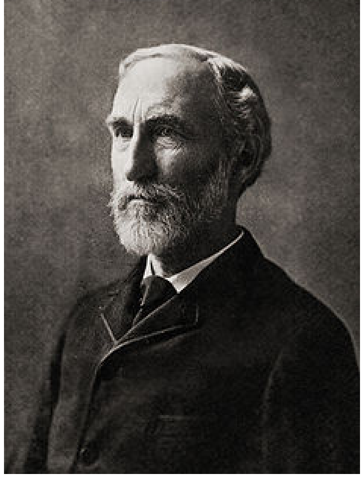
combinatorial	$Z \triangleq \sum_{\mathbf{a}} g(\mathbf{a})$	$Z_{\text{Bethe}} = ???$
analytical	$Z = \exp\left(-\min_{\mathbf{p}} F_{\text{Gibbs}}(\mathbf{p})\right)$	$Z_{\text{Bethe}} \triangleq \exp\left(-\min_{\boldsymbol{\beta}} F_{\text{Bethe}}(\boldsymbol{\beta})\right)$

(Temperature $T = 1$)



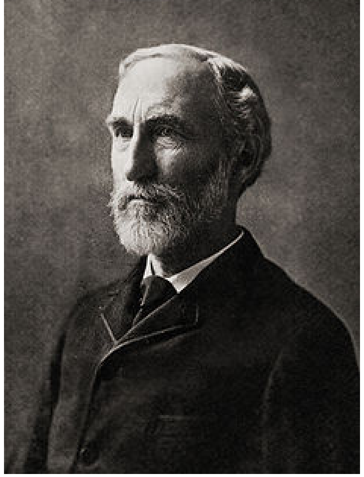
combinatorial	$Z \triangleq \sum_{\mathbf{a}} g(\mathbf{a})$	$Z_{\text{Bethe}} = ???$
analytical	$Z = \exp\left(-\min_{\mathbf{p}} F_{\text{Gibbs}}(\mathbf{p})\right)$	$Z_{\text{Bethe}} \triangleq \exp\left(-\min_{\boldsymbol{\beta}} F_{\text{Bethe}}(\boldsymbol{\beta})\right)$

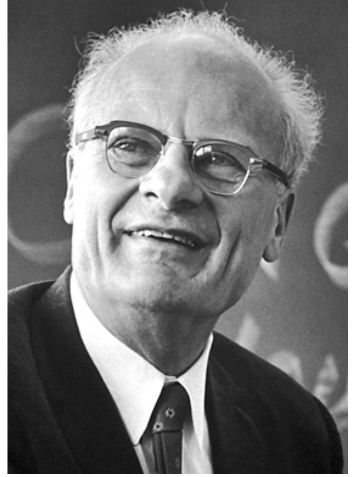
(Temperature $T = 1$)



combinatorial	$Z \triangleq \sum_{\mathbf{a}} g(\mathbf{a})$	$Z_{\text{Bethe}} = ???$
analytical	$Z = \exp\left(-\min_{\mathbf{p}} F_{\text{Gibbs}}(\mathbf{p})\right)$	$Z_{\text{Bethe}} \triangleq \exp\left(-\min_{\boldsymbol{\beta}} F_{\text{Bethe}}(\boldsymbol{\beta})\right)$

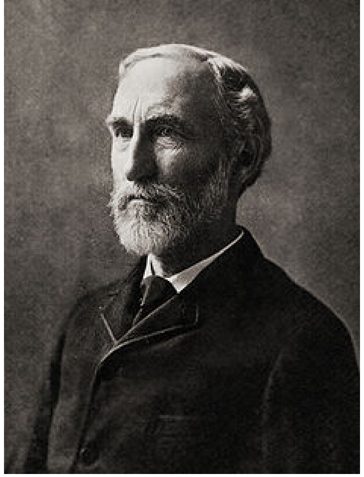
(Temperature $T = 1$)

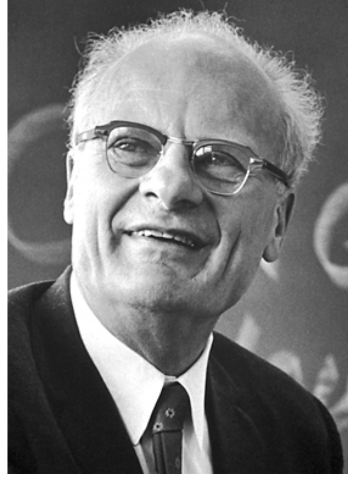




combinatorial	$Z \triangleq \sum_{\mathbf{a}} g(\mathbf{a})$	$Z_{\text{Bethe}} = ???$
analytical	$Z = \exp\left(-\min_{\mathbf{p}} F_{\text{Gibbs}}(\mathbf{p})\right)$	$Z_{\text{Bethe}} \triangleq \exp\left(-\min_{\boldsymbol{\beta}} F_{\text{Bethe}}(\boldsymbol{\beta})\right)$

(Temperature $T = 1$)



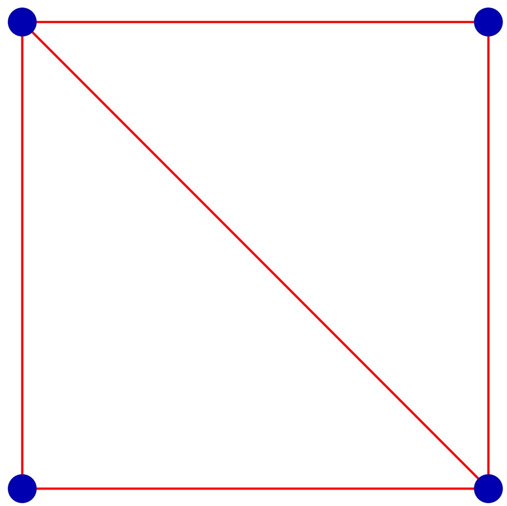


combinatorial	$Z \triangleq \sum_{\mathbf{a}} g(\mathbf{a})$	$Z_{\text{Bethe}} = ???$
analytical	$Z = \exp\left(-\min_{\mathbf{p}} F_{\text{Gibbs}}(\mathbf{p})\right)$	$Z_{\text{Bethe}} \triangleq \exp\left(-\min_{\boldsymbol{\beta}} F_{\text{Bethe}}(\boldsymbol{\beta})\right)$

(Temperature $T = 1$)

Finite graph covers

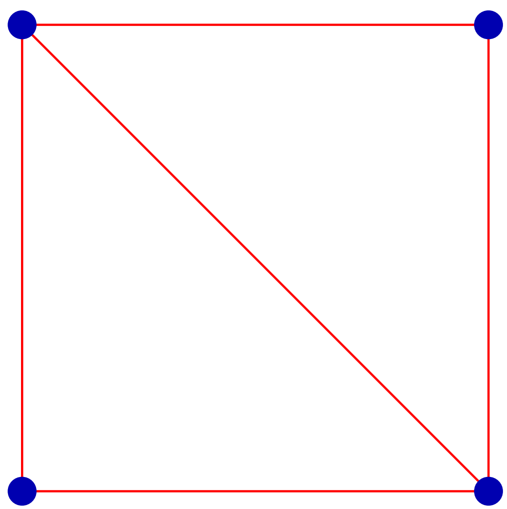
Finite Graph Covers



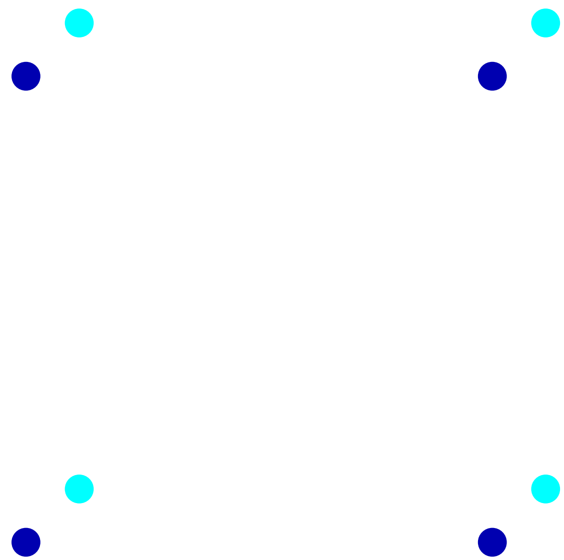
original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



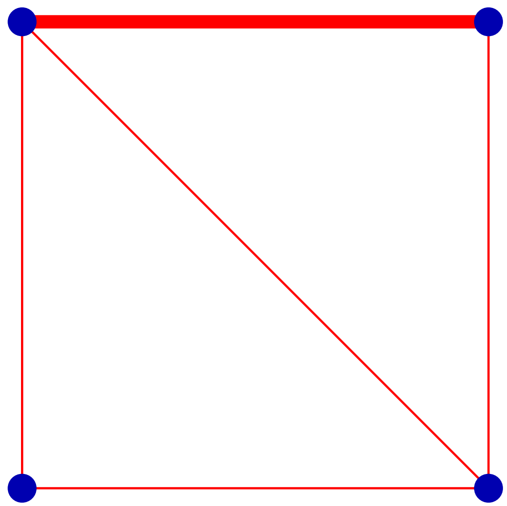
original graph



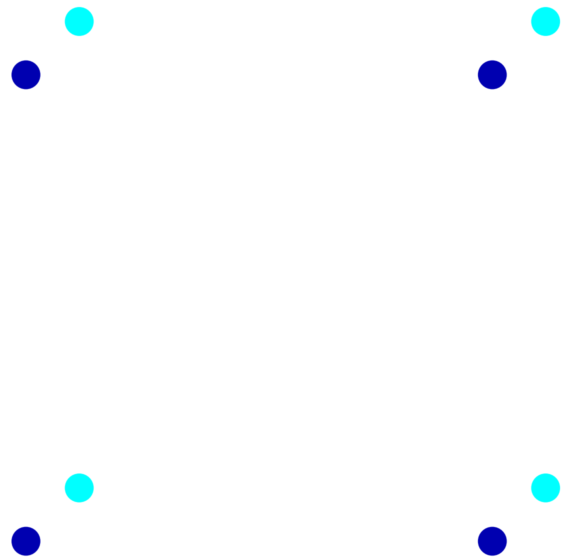
2-fold cover of
original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



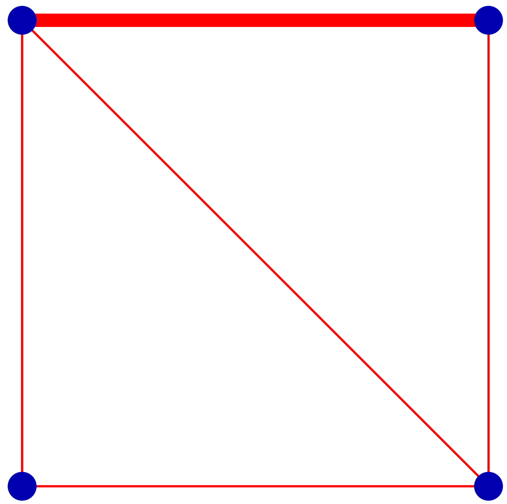
original graph



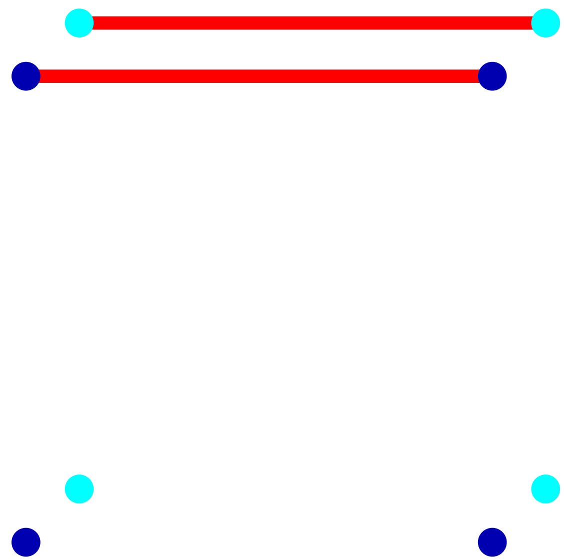
2-fold cover of
original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



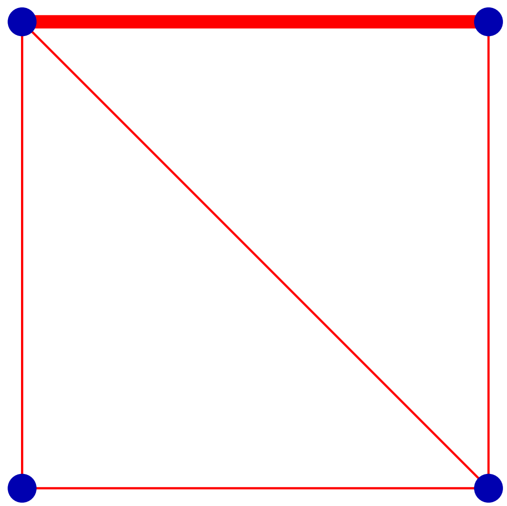
original graph



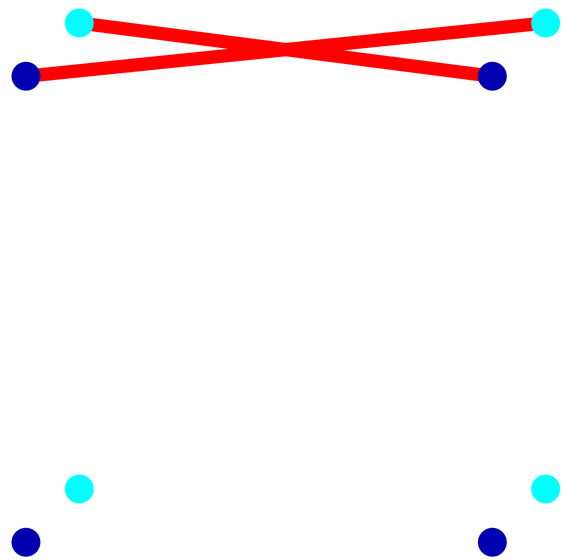
2-fold cover of original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



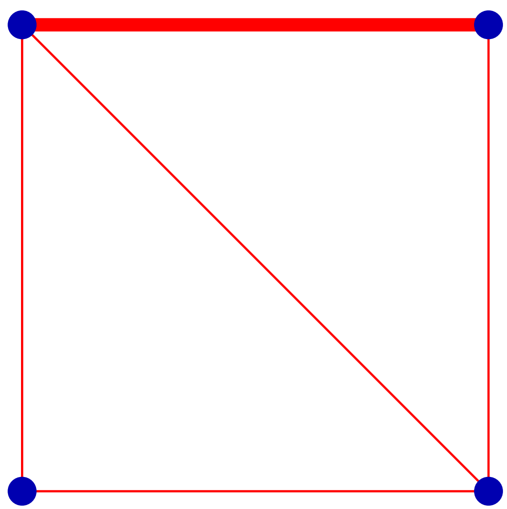
original graph



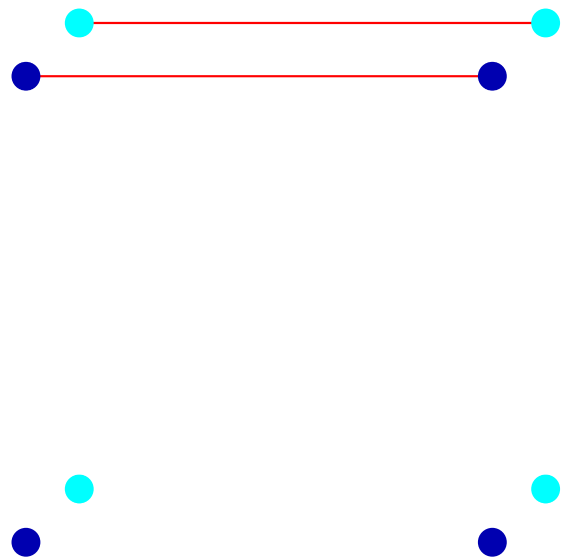
2-fold cover of original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



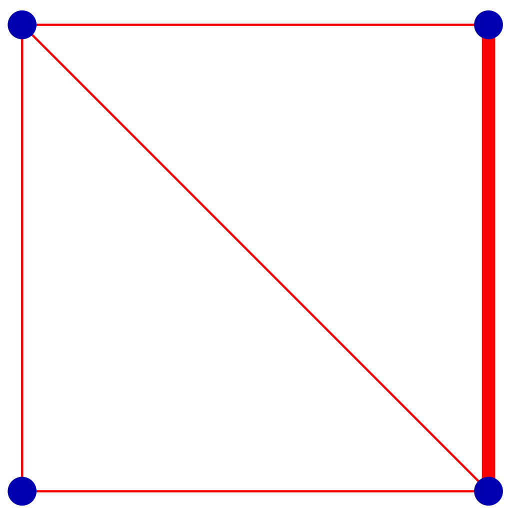
original graph



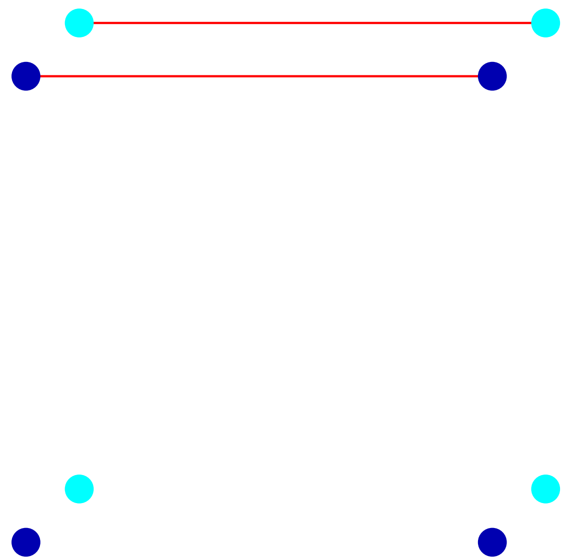
2-fold cover of original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



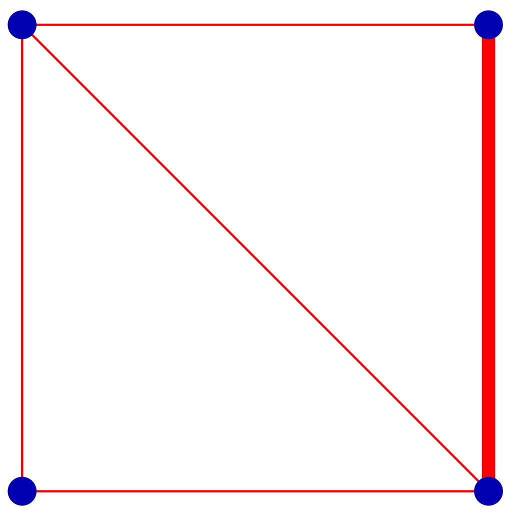
original graph



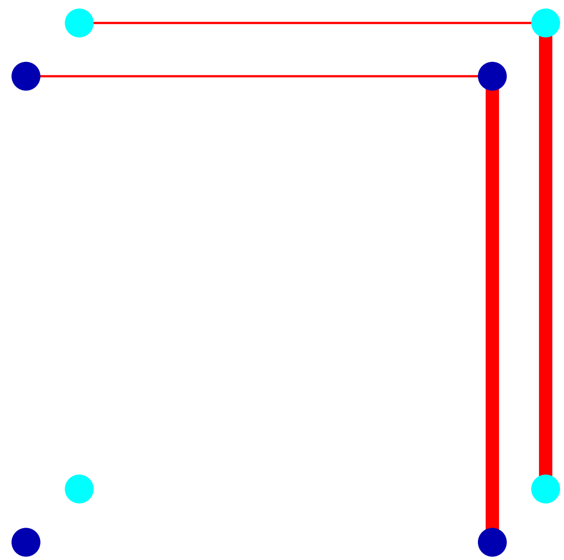
2-fold cover of
original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



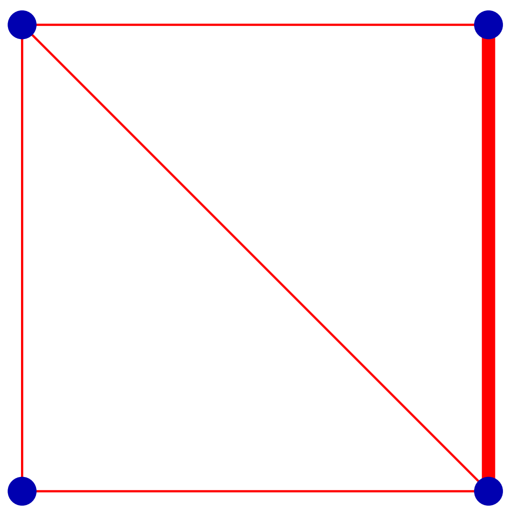
original graph



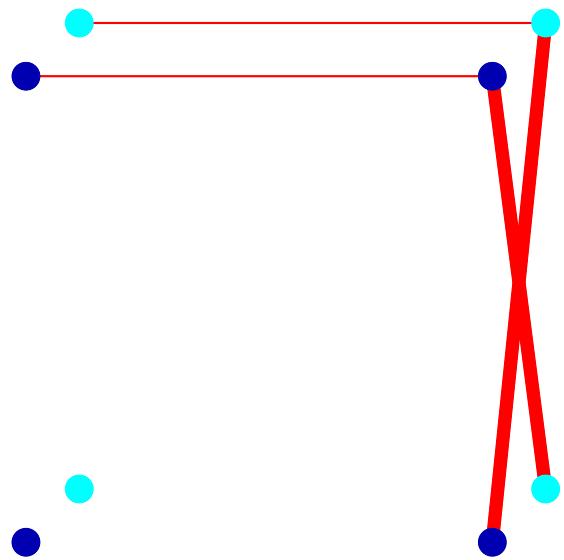
2-fold cover of
original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



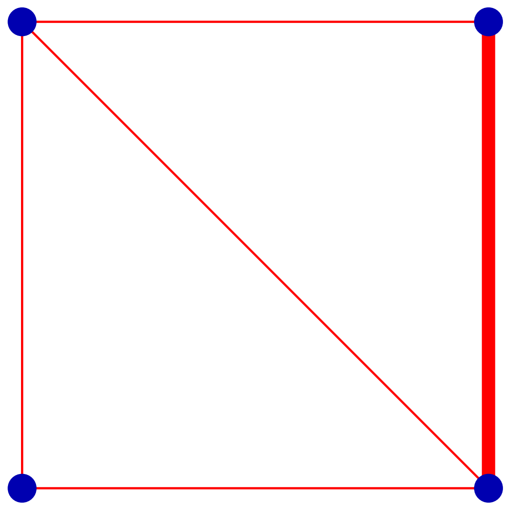
original graph



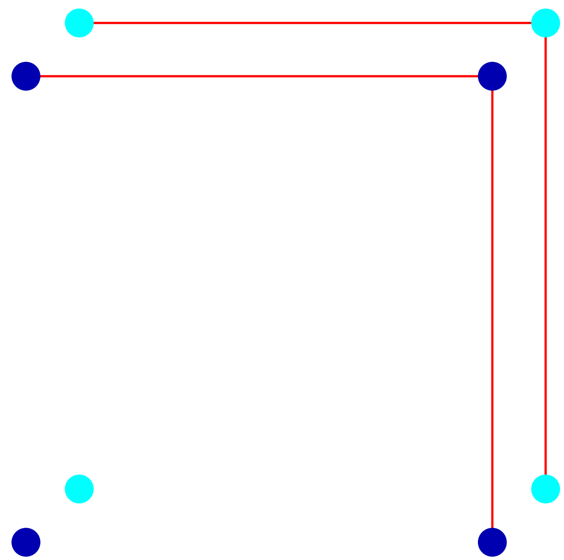
2-fold cover of original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



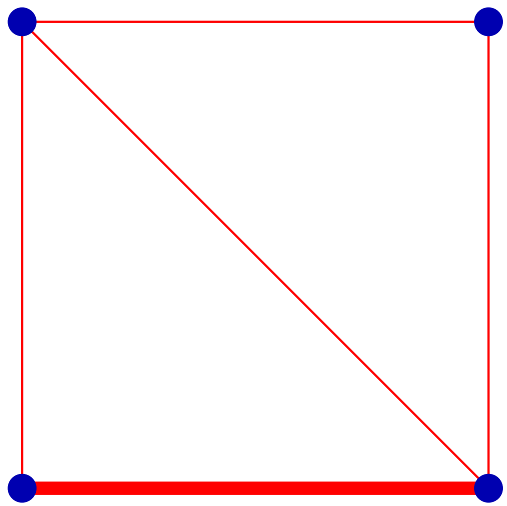
original graph



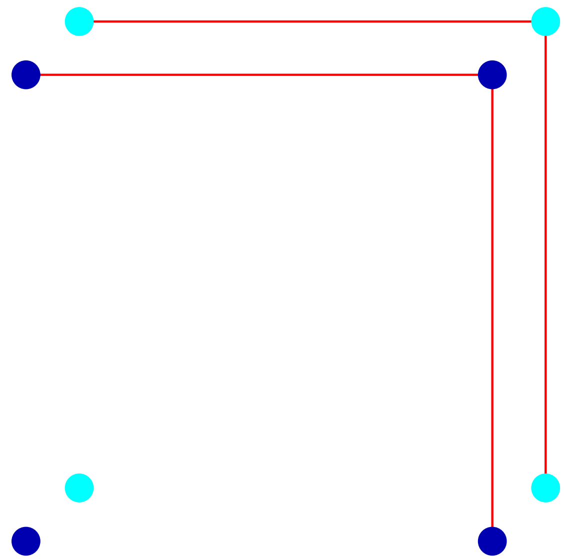
2-fold cover of
original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



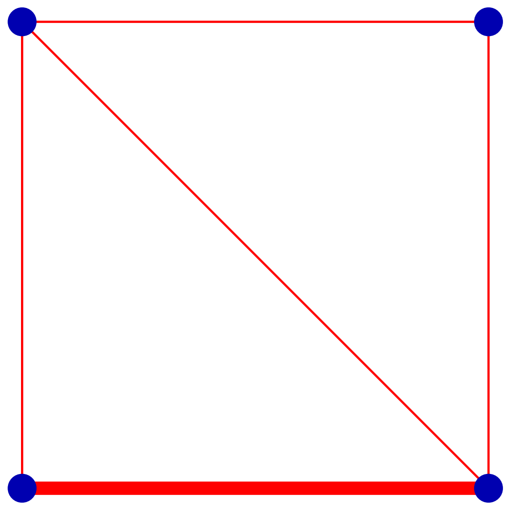
original graph



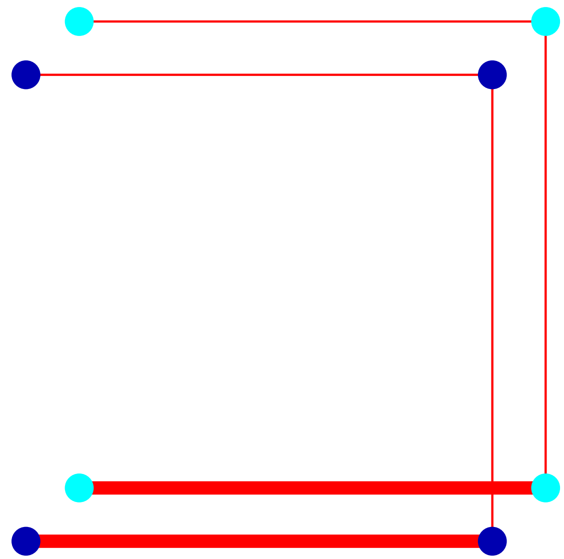
2-fold cover of original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



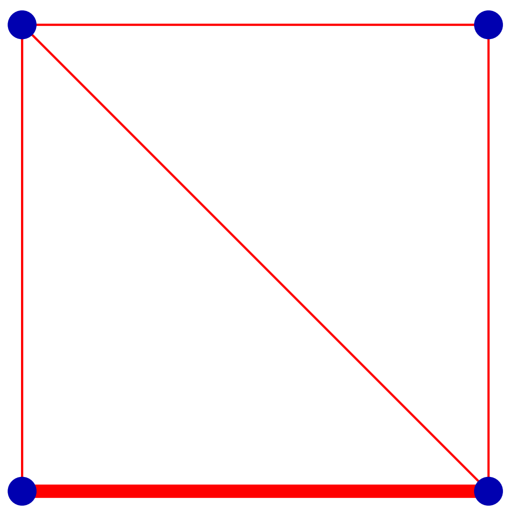
original graph



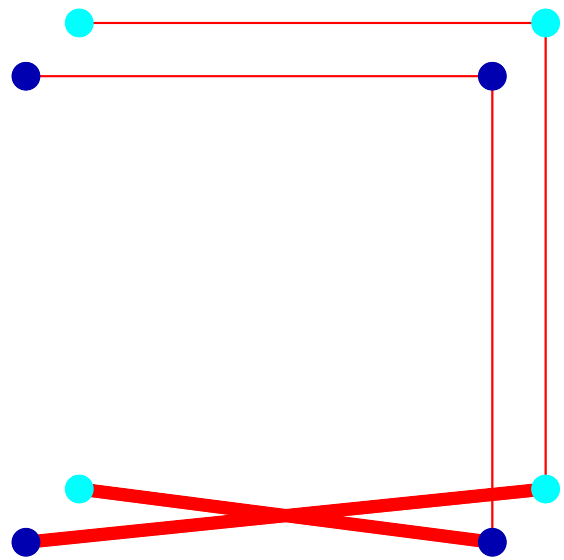
2-fold cover of original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



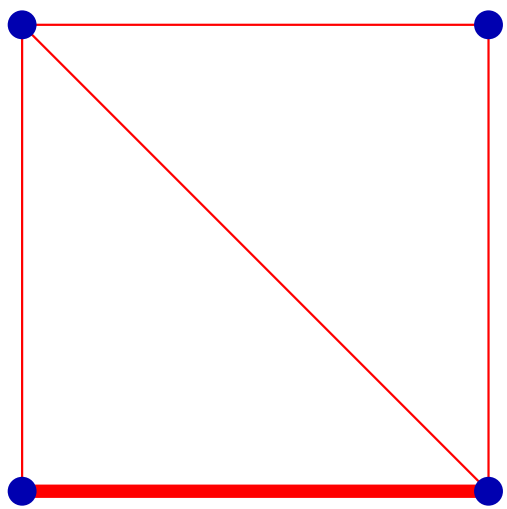
original graph



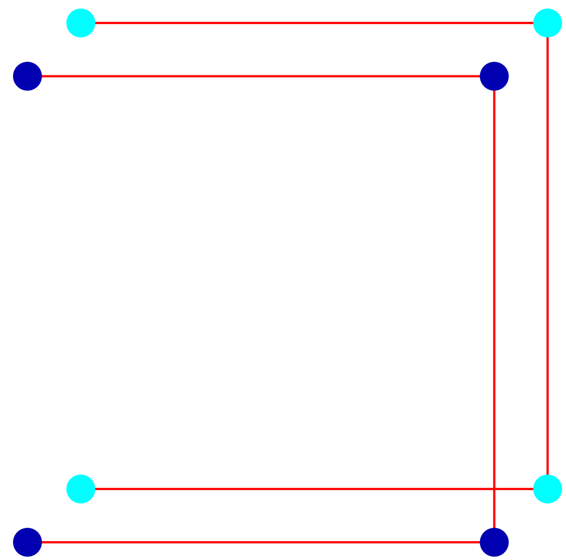
2-fold cover of
original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



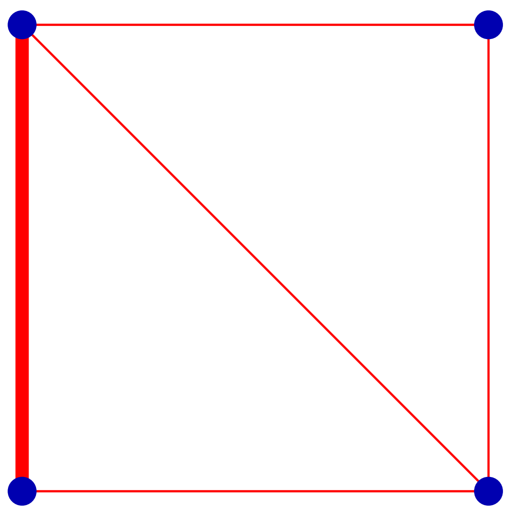
original graph



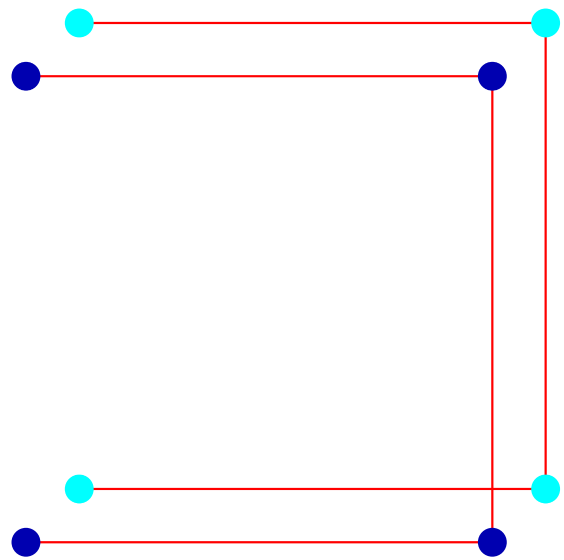
2-fold cover of original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



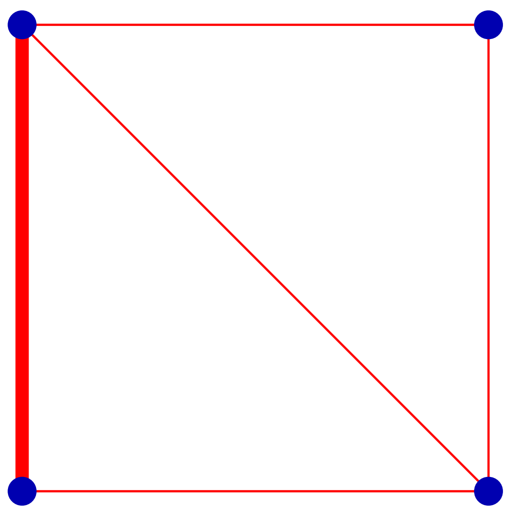
original graph



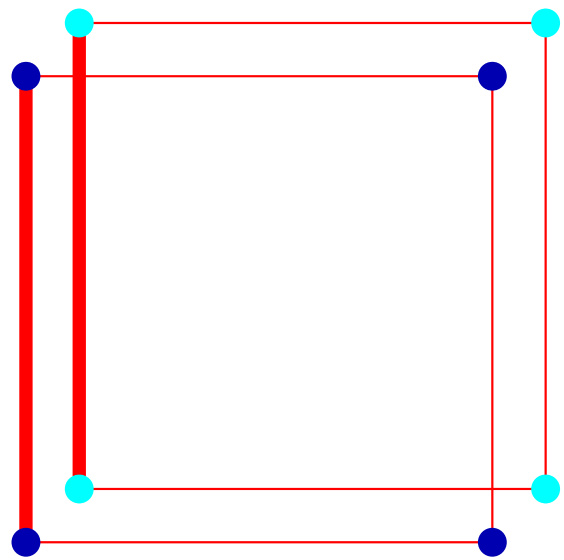
2-fold cover of original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



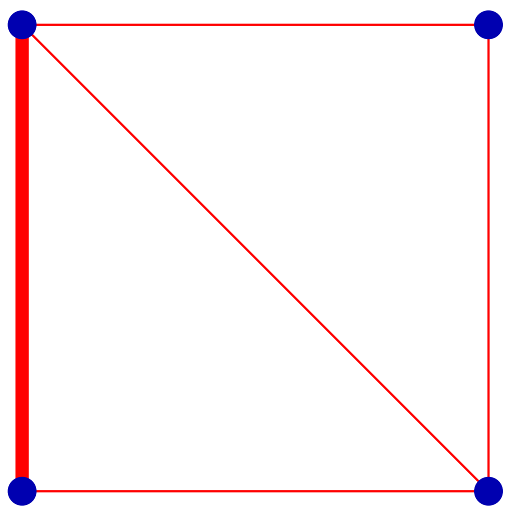
original graph



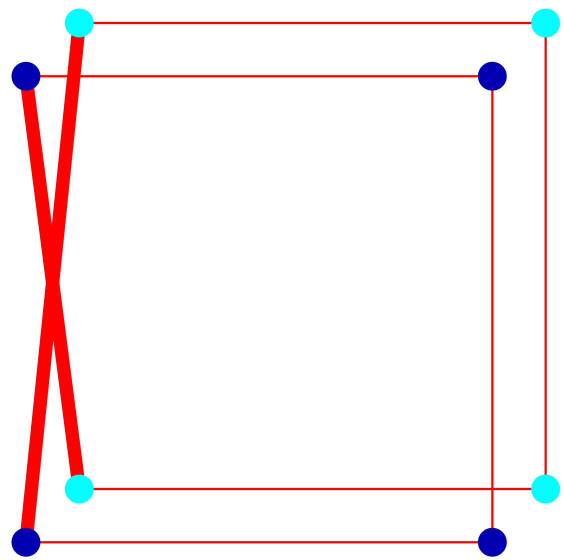
2-fold cover of original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



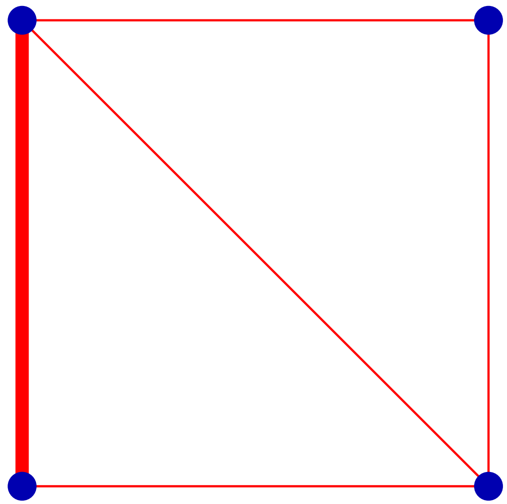
original graph



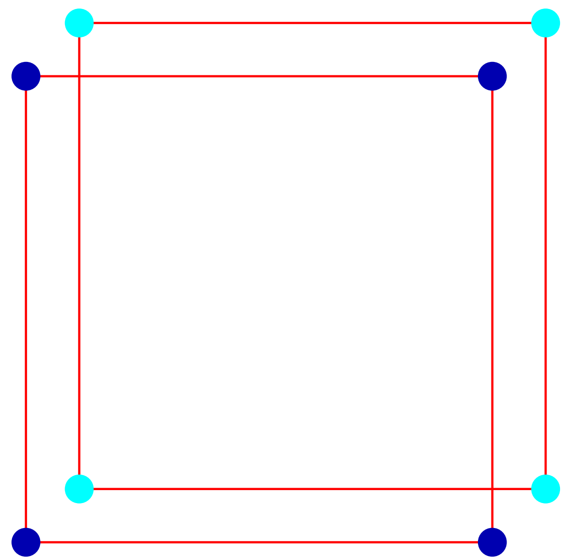
2-fold cover of original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



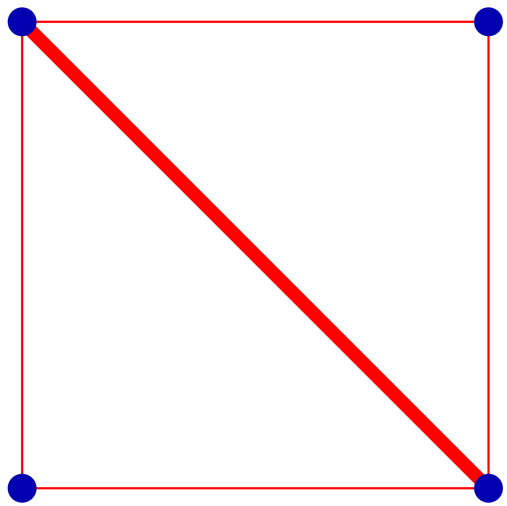
original graph



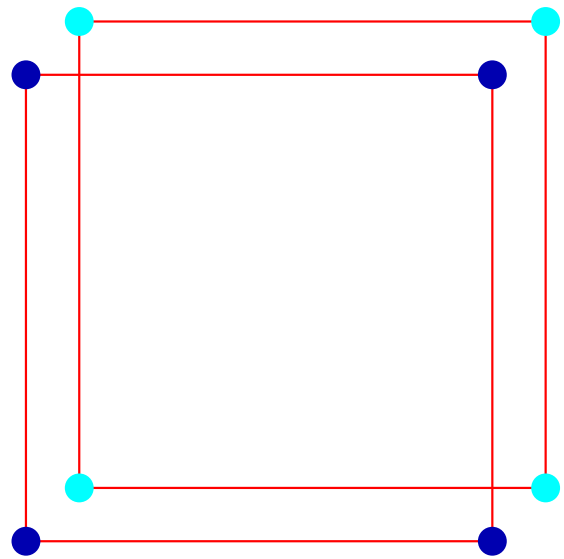
2-fold cover of original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



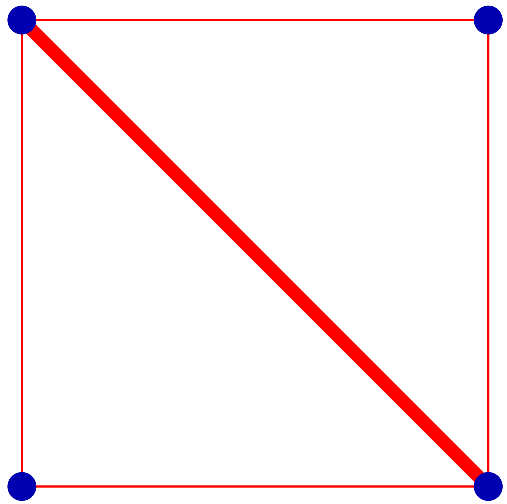
original graph



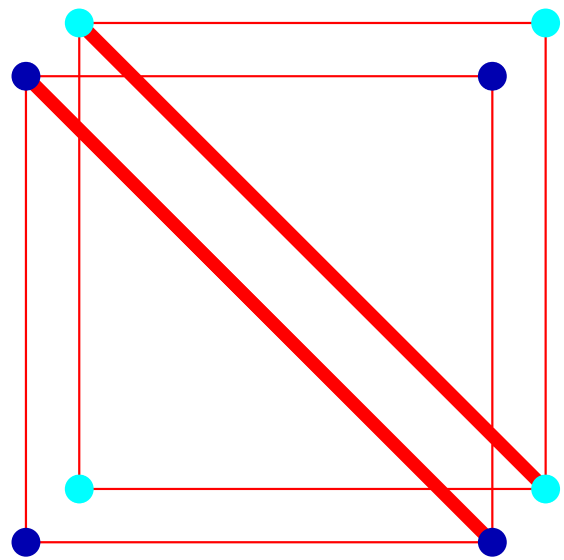
2-fold cover of
original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



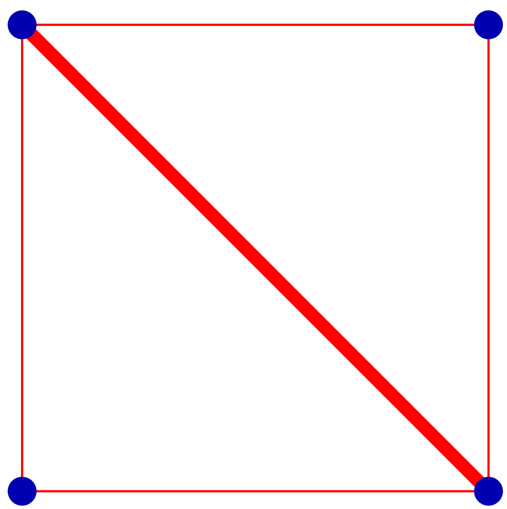
original graph



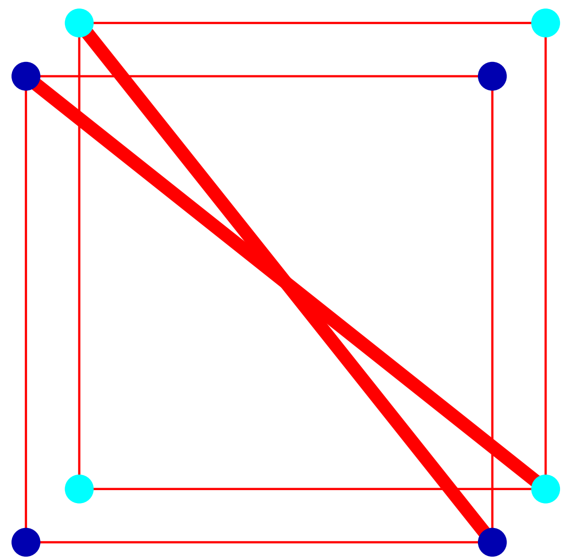
2-fold cover of original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



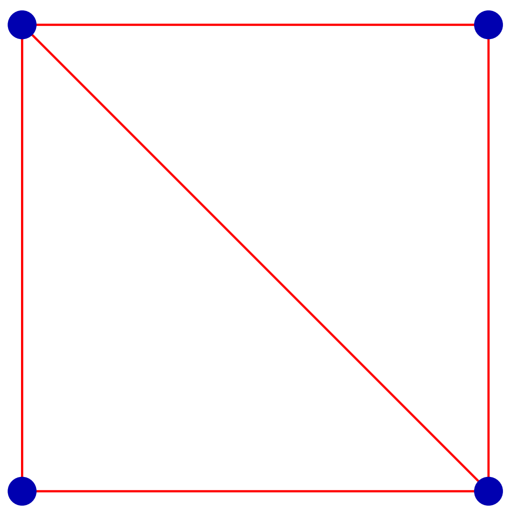
original graph



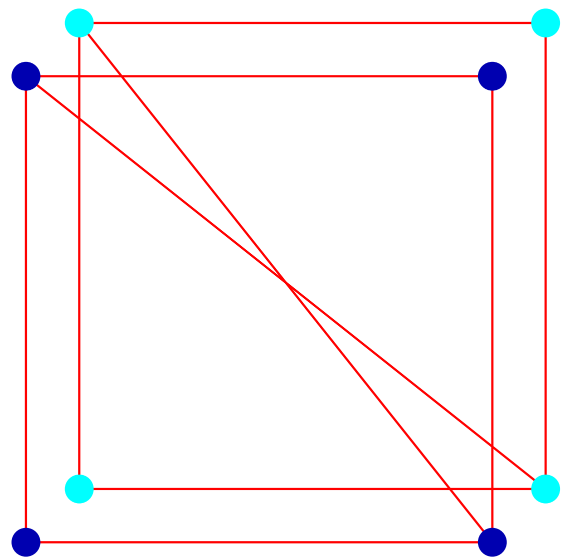
2-fold cover of original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



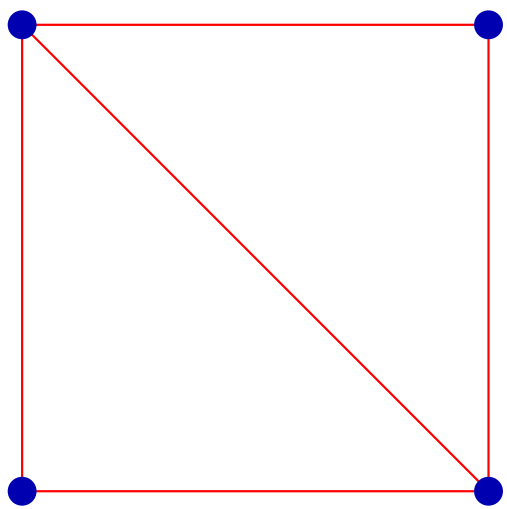
original graph



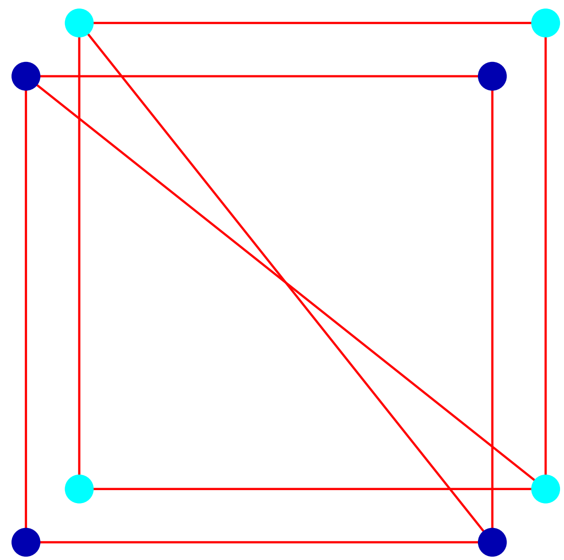
2-fold cover of original graph

Definition: A double cover of a graph is . . .

Finite Graph Covers



original graph

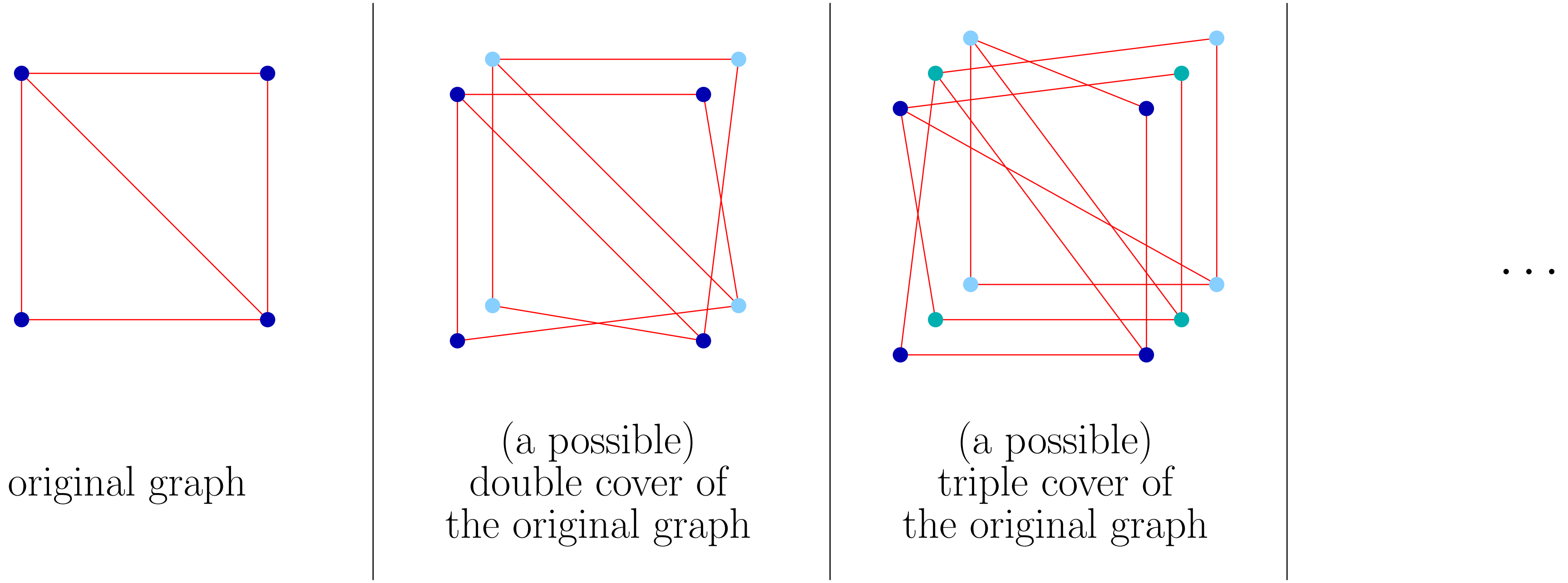


2-fold cover of original graph

Definition: A double cover of a graph is . . .

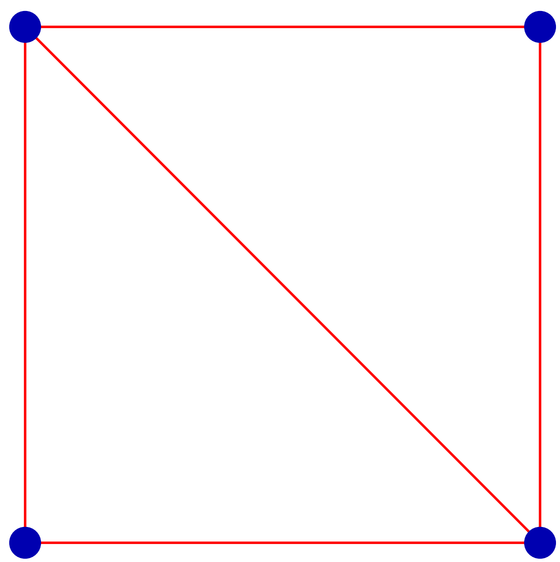
Note: the above graph has $2! \cdot 2! \cdot 2! \cdot 2! \cdot 2! = (2!)^5$ double covers.

Graph Covers

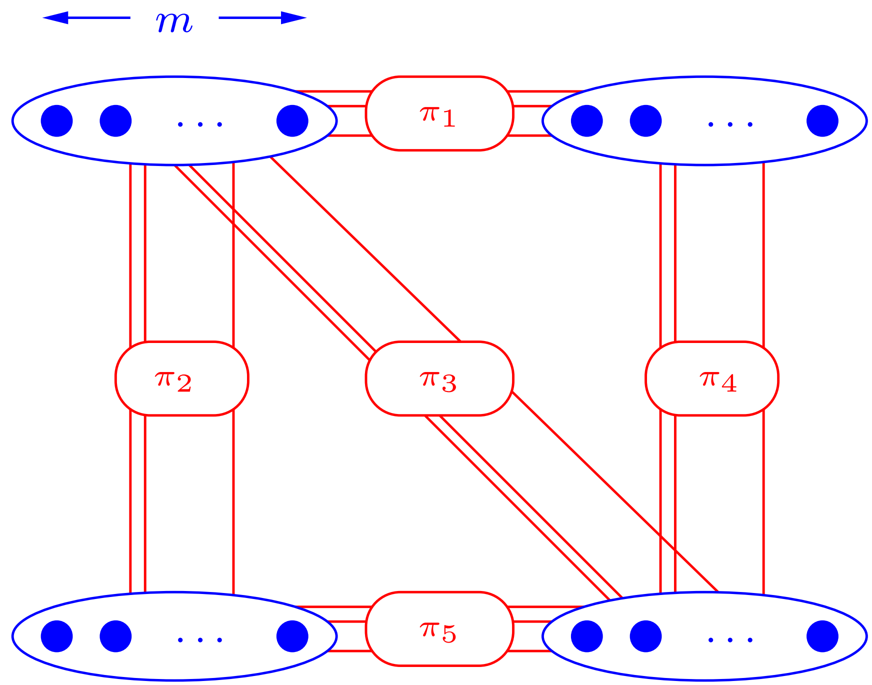


Besides double covers, a graph also has many triple covers, quadruple covers, quintuple covers, etc.

Graph Covers



original graph



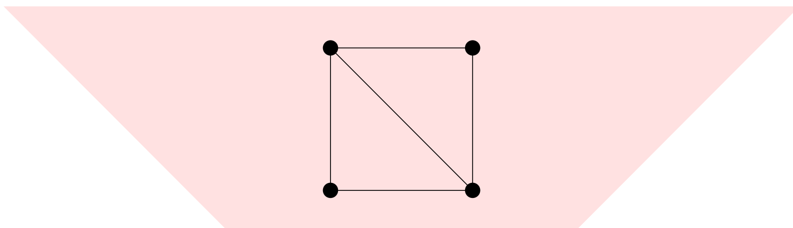
(possible)
m-fold cover of
original graph

An m-fold cover is also called a cover of degree m.

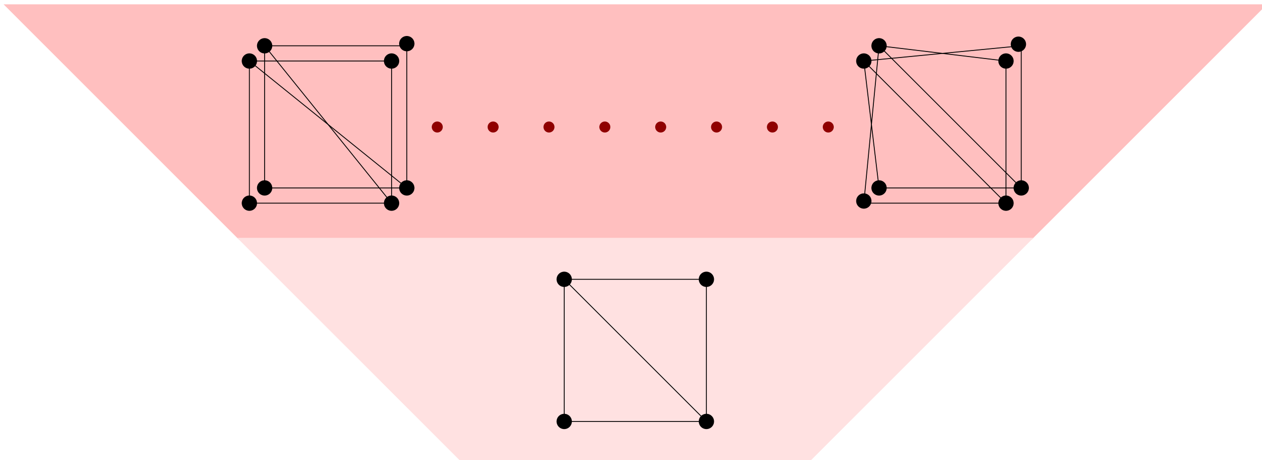
Do not confuse this degree with the degree of a vertex!

Graph Cover Hierarchy

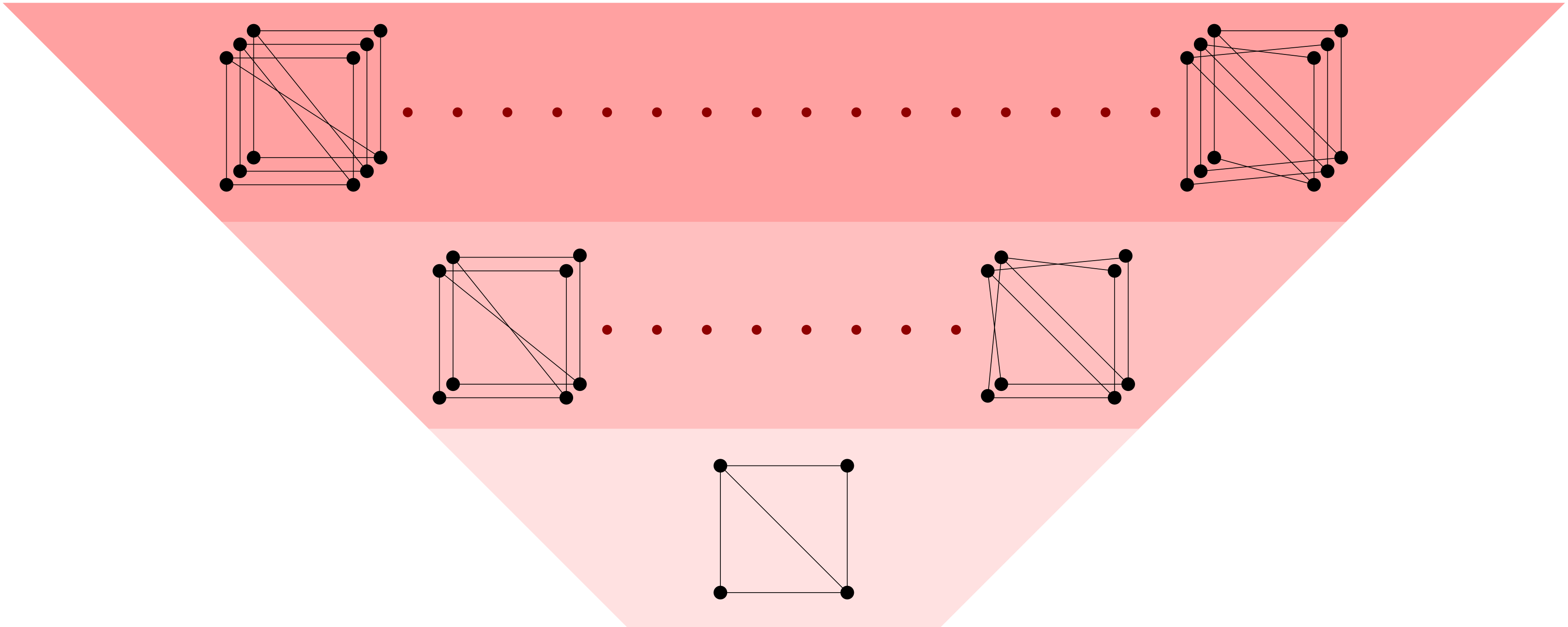
Graph Cover Hierarchy



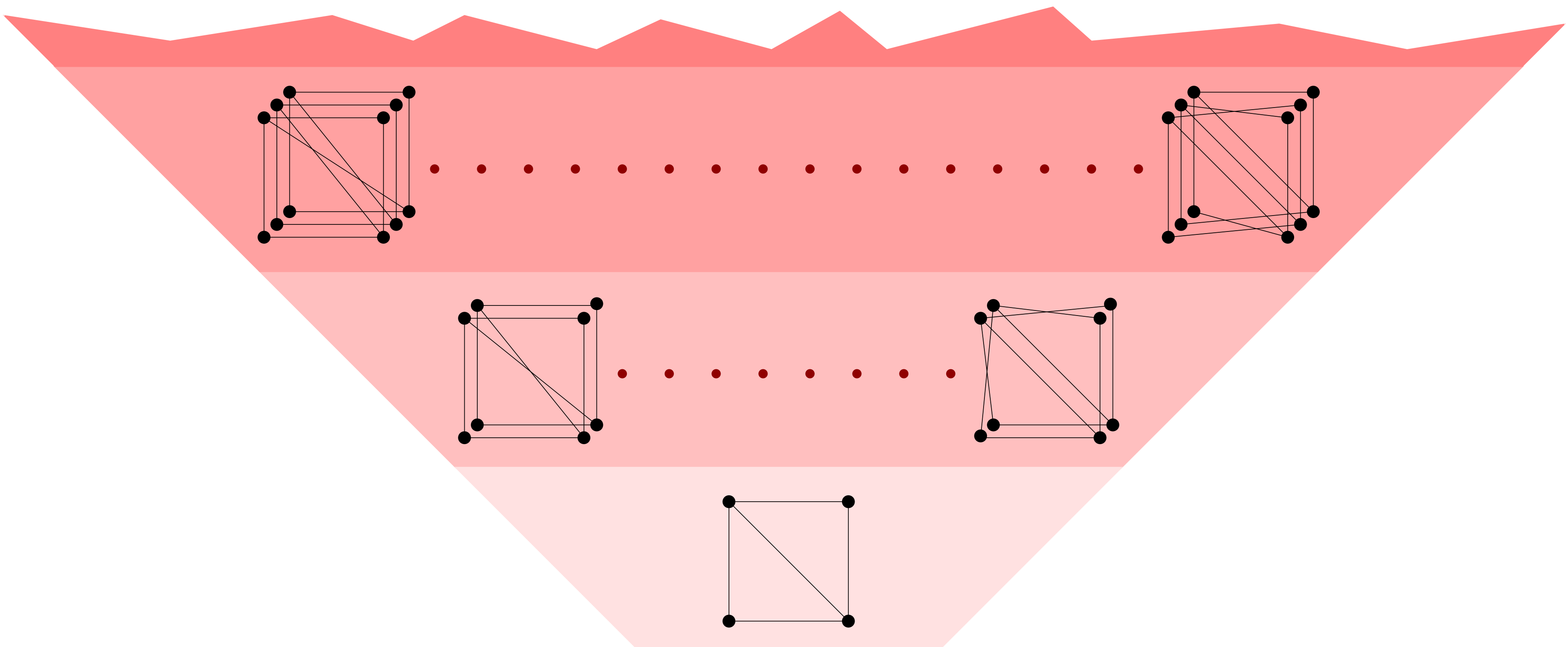
Graph Cover Hierarchy



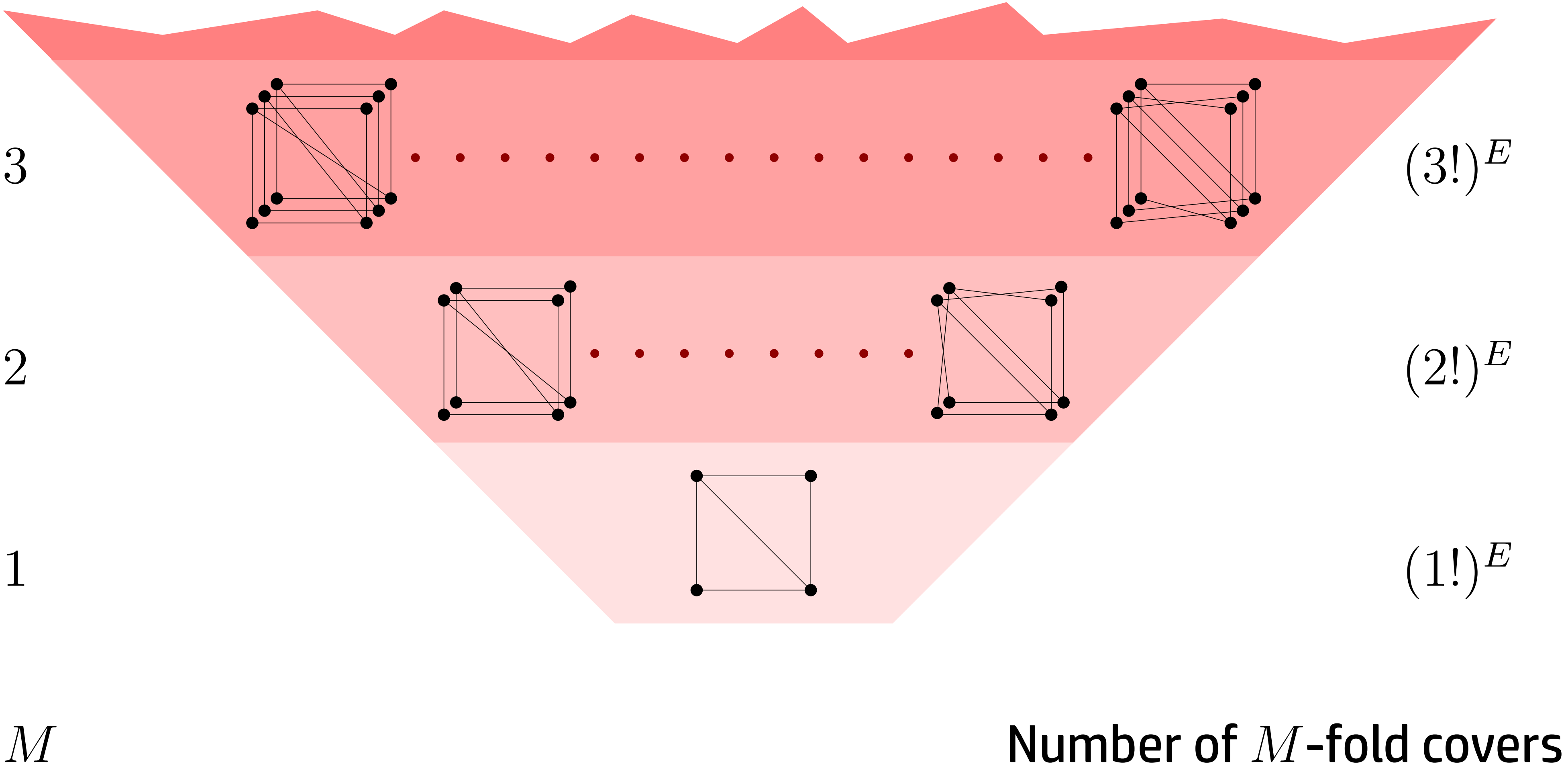
Graph Cover Hierarchy



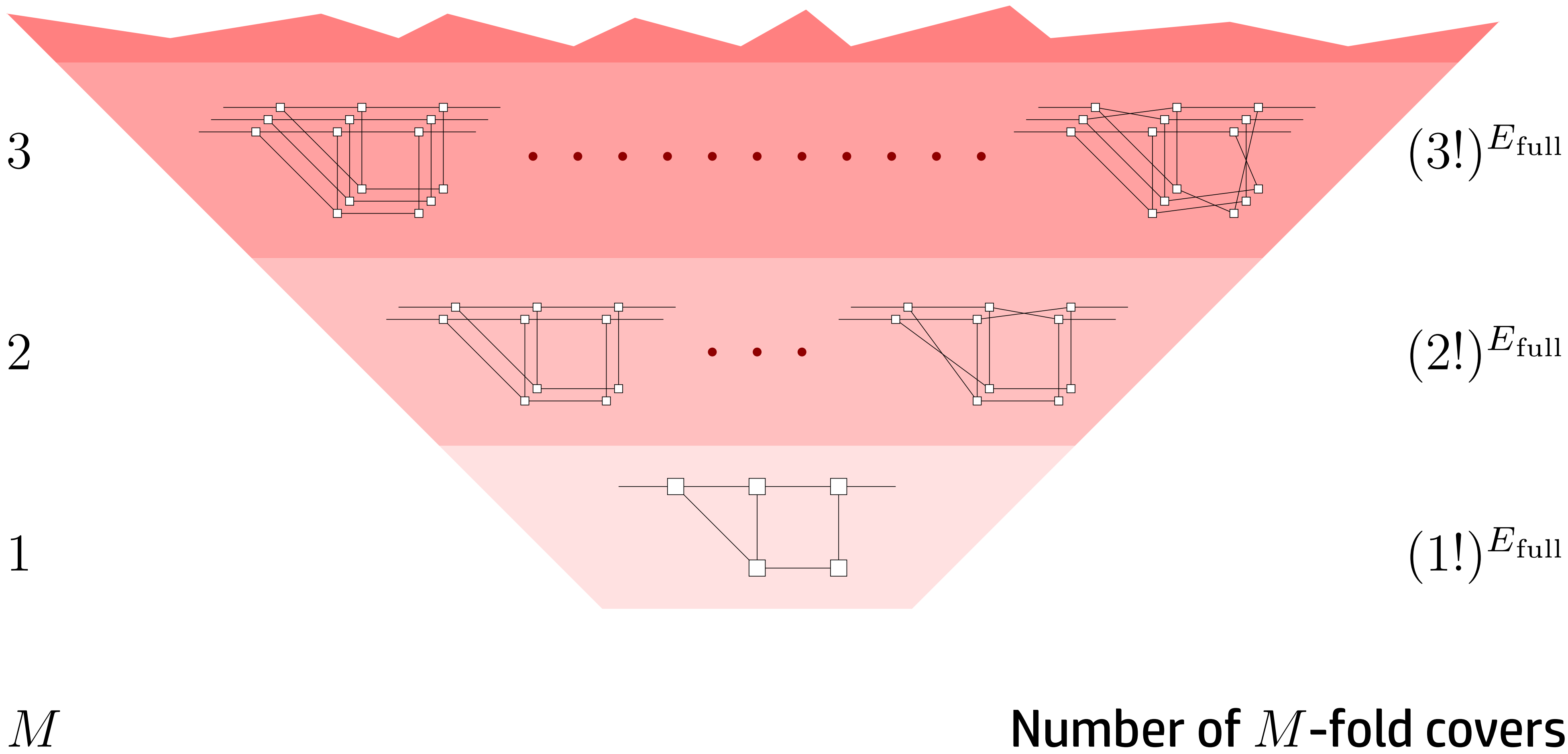
Graph Cover Hierarchy



Graph Cover Hierarchy



Graph Cover Hierarchy



Graph Covers

Graph covers (a.k.a. graph lifts) have appeared in various contexts in the literature:

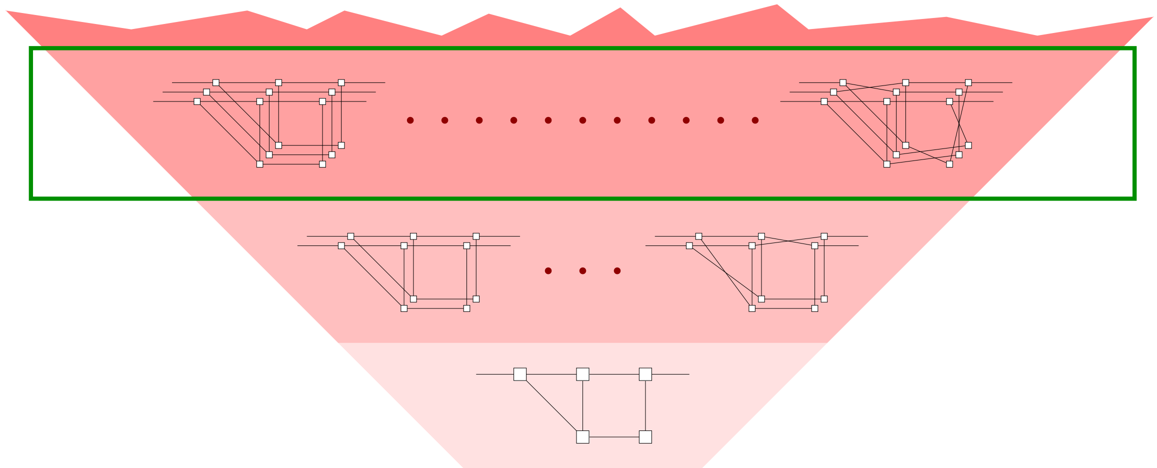
- D. Angluin (STOC 1980):
  Local and global properties in networks of processors.
- N. Linial et al.:
  Various papers on characterizing properties of graph covers.
- A. Marcus, D. A. Spielman, and N. Srivastava (FOCS 2013):
  have shown the existence of infinite families of regular bipartite Ramanujan graphs of every degree bigger than 2.

Graph covers in coding theory:

- Koetter and Vontobel (ISTC 2003):
  analysis of message-passing iterative decoders via graph covers.

A combinatorial interpretation of the Bethe partition sum

A Combinatorial Interpretation of the Bethe Partition Sum



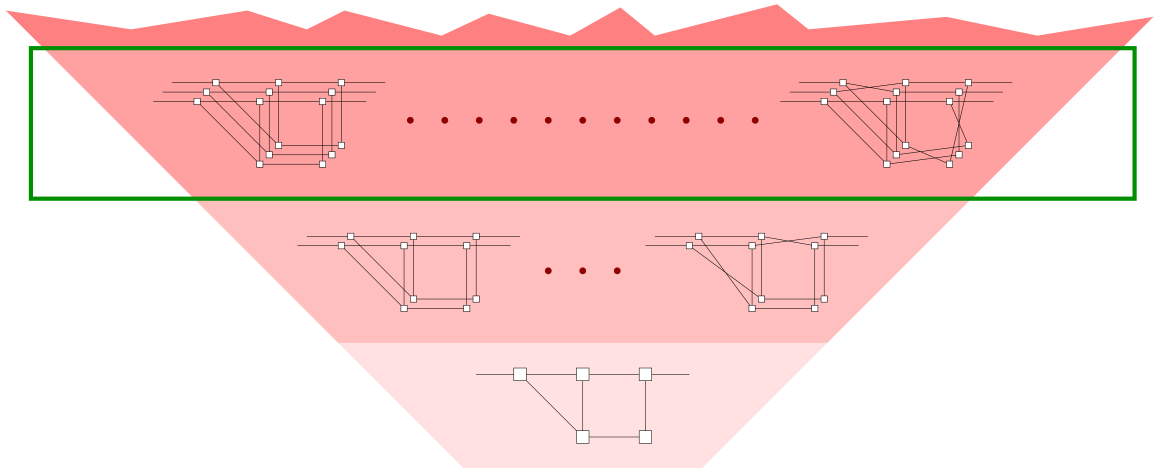
Definition:

- Let N be a factor graph.
- Let $M \in \mathbb{Z}_{>0}$.

We define the degree-M Bethe partition sum to be

$$Z_{\mathrm{Bethe},M}(\mathsf{N}) \triangleq \sqrt[M]{\left\langle Z(\widetilde{\mathsf{N}}) \right\rangle_{\widetilde{\mathsf{N}} \in \widetilde{\mathcal{N}}_M}}.$$

A Combinatorial Interpretation of the Bethe Partition Sum



Definition:

- Let N be a factor graph.
- Let $M \in \mathbb{Z}_{>0}$.

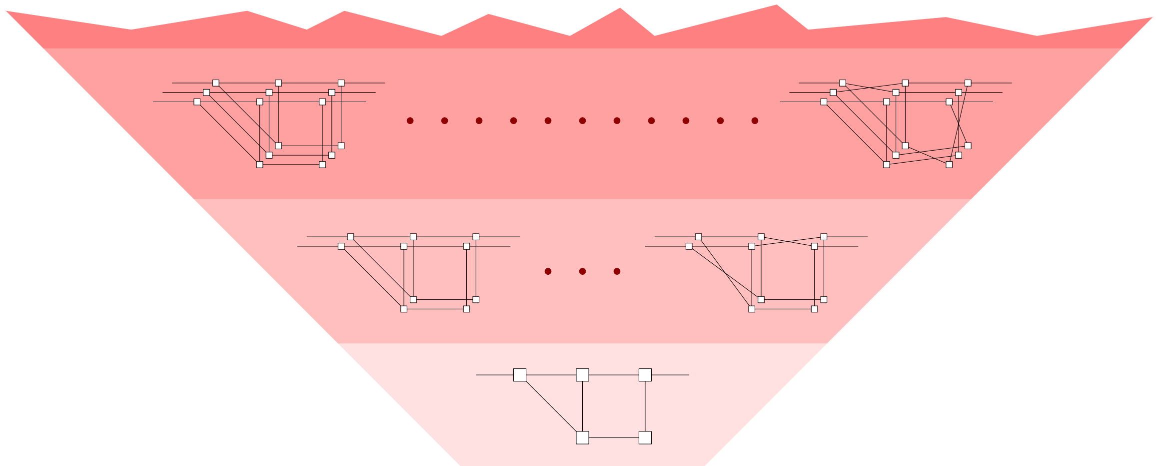
We define the degree-M Bethe partition sum to be

$$Z_{\mathrm{Bethe},M}(\mathsf{N}) \triangleq \sqrt[M]{\left\langle Z(\widetilde{\mathsf{N}}) \right\rangle_{\widetilde{\mathsf{N}} \in \widetilde{\mathcal{N}}_M}}.$$

Note that the RHS of the above expression is based on the partition sum, and not on the Bethe partition sum.

Degree-M Bethe Partition Sum

$Z_{\mathrm{Bethe},M}(\mathsf{N})$

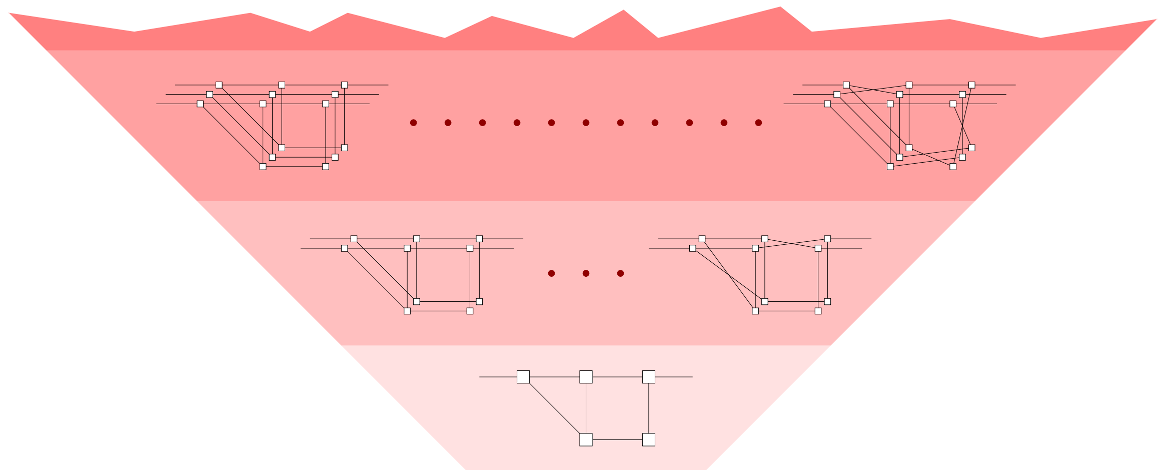


$$Z_{\mathrm{Bethe},M}(\mathsf{N}) \triangleq \sqrt[M]{\left\langle Z(\widetilde{\mathsf{N}}) \right\rangle_{\widetilde{\mathsf{N}} \in \widetilde{\mathcal{N}}_M}}$$

Degree-M Bethe Partition Sum

$Z_{\text{Bethe},M}(\mathsf{N})$

$Z_{\text{Bethe},M}(\mathsf{N})\Big|_{M=1} = Z(\mathsf{N})$



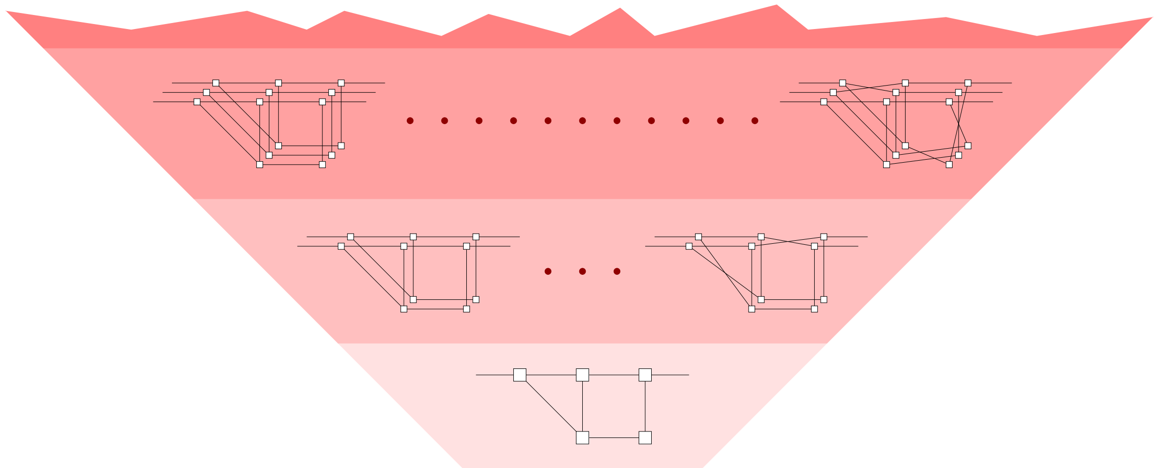
$$Z_{\text{Bethe},M}(\mathsf{N}) \triangleq \sqrt[M]{\left\langle Z(\widetilde{\mathsf{N}}) \right\rangle_{\widetilde{\mathsf{N}} \in \widetilde{\mathcal{N}}_M}}$$

Degree-M Bethe Partition Sum

$$Z_{\mathrm{Bethe},M}(\mathsf{N})\Big|_{M\to\infty} = Z_{\mathrm{Bethe}}(\mathsf{N})$$

$$Z_{\mathrm{Bethe},M}(\mathsf{N})$$

$$Z_{\mathrm{Bethe},M}(\mathsf{N})\Big|_{M=1} = Z(\mathsf{N})$$



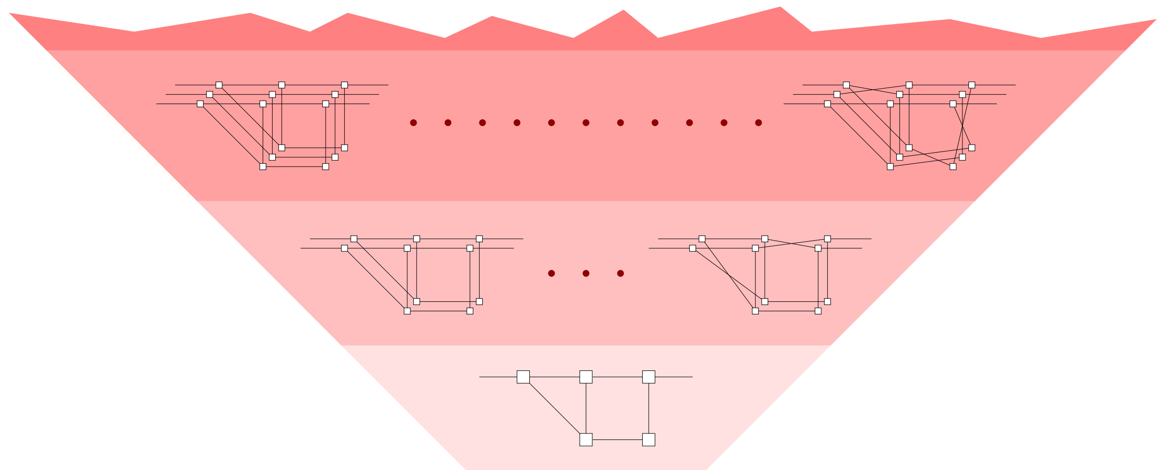
$$Z_{\mathrm{Bethe},M}(\mathsf{N}) \triangleq \sqrt[M]{\Big\langle Z(\widetilde{\mathsf{N}}) \Big\rangle_{\widetilde{\mathsf{N}}\in\widetilde{\mathcal{N}}_M}}$$

Degree-M Bethe Partition Sum

$$Z_{\mathrm{Bethe},M}(\mathsf{N})\Big|_{M\to\infty} = Z_{\mathrm{Bethe}}(\mathsf{N})$$ **(Theorem [V., 2013])**

$$Z_{\mathrm{Bethe},M}(\mathsf{N})$$

$$Z_{\mathrm{Bethe},M}(\mathsf{N})\Big|_{M=1} = Z(\mathsf{N})$$



$$Z_{\mathrm{Bethe},M}(\mathsf{N}) \triangleq \sqrt[M]{\left\langle Z(\widetilde{\mathsf{N}}) \right\rangle_{\widetilde{\mathsf{N}} \in \widetilde{\mathcal{N}}_M}}$$

Examples

Example 1: Factor graphs without cycles

If N does *not* contain any cycles, then

any M-cover $\widetilde{\mathrm{N}}$ of N consists of M disconnected copies of N

Example 1: Factor graphs without cycles

If N does *not* contain any cycles, then

any M-cover $\widetilde{\mathsf{N}}$ of N consists of M disconnected copies of N

and so

$$Z(\widetilde{\mathsf{N}}) = \big(Z(\mathsf{N})\big)^{M}.$$

Example 1: Factor graphs without cycles

If N does *not* contain any cycles, then

any M-cover $\widetilde{\mathsf{N}}$ of N consists of M disconnected copies of N

and so

$$Z(\widetilde{\mathsf{N}}) = \big(Z(\mathsf{N})\big)^M .$$

Consequently,

$$Z_{\mathrm{Bethe},M}(\mathsf{N}) = \sqrt[M]{\Big\langle Z(\widetilde{\mathsf{N}}) \Big\rangle_{\widetilde{\mathsf{N}} \in \widetilde{\mathcal{N}}_M}} = Z(\mathsf{N}).$$

Example 1: Factor graphs without cycles

If N does *not* contain any cycles, then

any M-cover $\widetilde{\mathsf{N}}$ of N consists of M disconnected copies of N

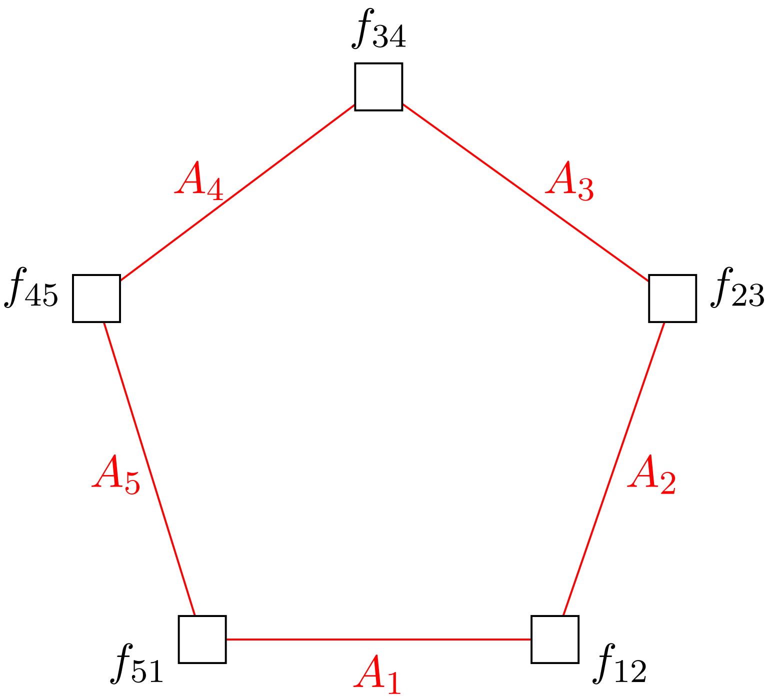
and so

$$Z(\widetilde{\mathsf{N}}) = \big(Z(\mathsf{N})\big)^{M}.$$

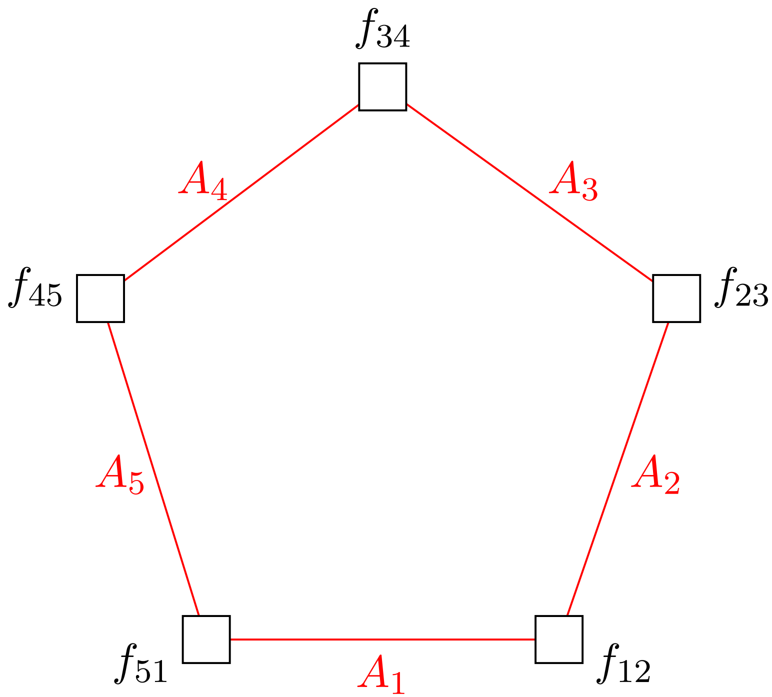
Consequently,

$$Z_{\mathrm{Bethe}}(\mathsf{N}) = Z(\mathsf{N}).$$

Example 2: 5-Cycle

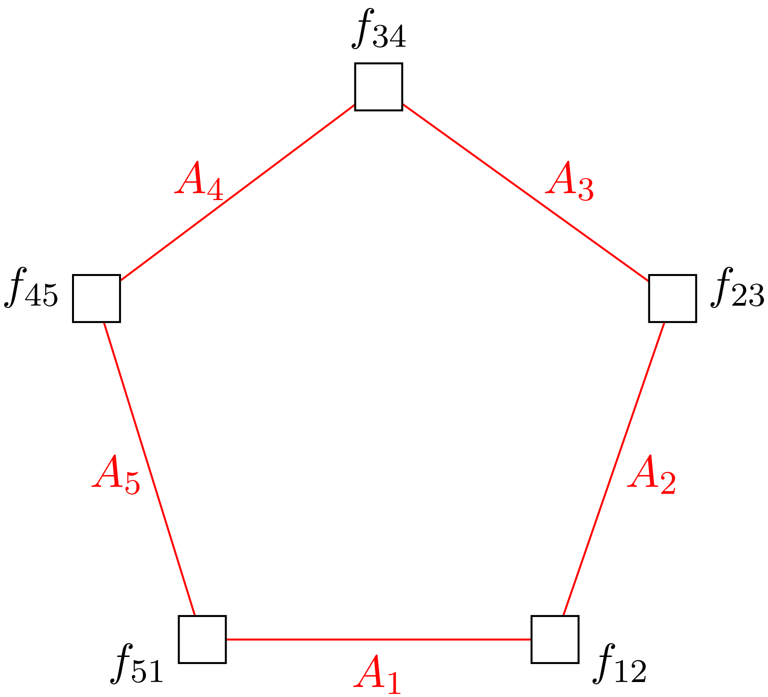


Example 2: 5-Cycle



- Let $\mathcal{A}_e \triangleq \{0, 1\}$ for all $e \in \mathcal{E}$.

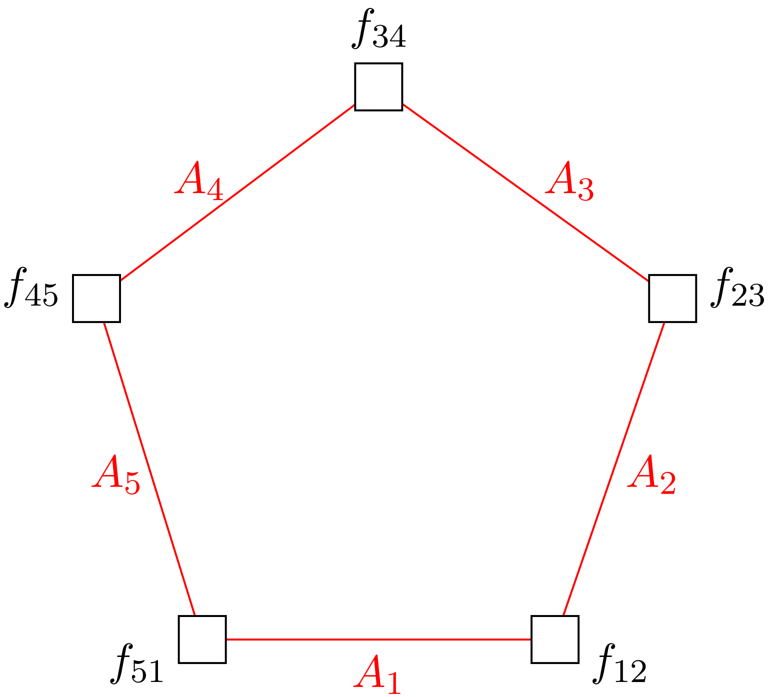
Example 2: 5-Cycle



- Let $\mathcal{A}_e \triangleq \{0, 1\}$ for all $e \in \mathcal{E}$.
- With the local function $f(a_e, a_{e+1})$, we can associate the matrix

$$\mathbf{M}_f = \begin{pmatrix} f(0,0) & f(0,1) \\ f(1,0) & f(1,1) \end{pmatrix}.$$

Example 2: 5-Cycle

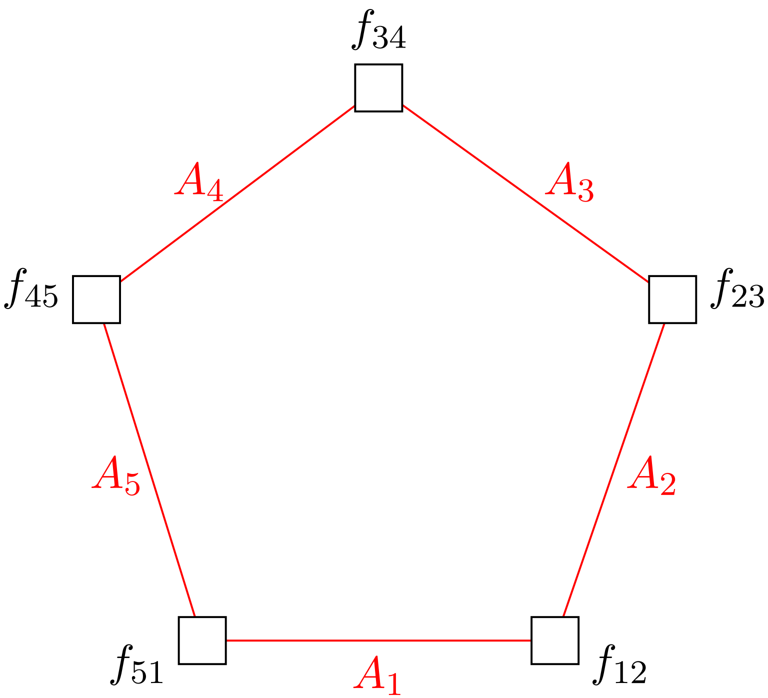


- Let $\mathcal{A}_e \triangleq \{0, 1\}$ for all $e \in \mathcal{E}$.
- With the local function $f(a_e, a_{e+1})$, we can associate the matrix

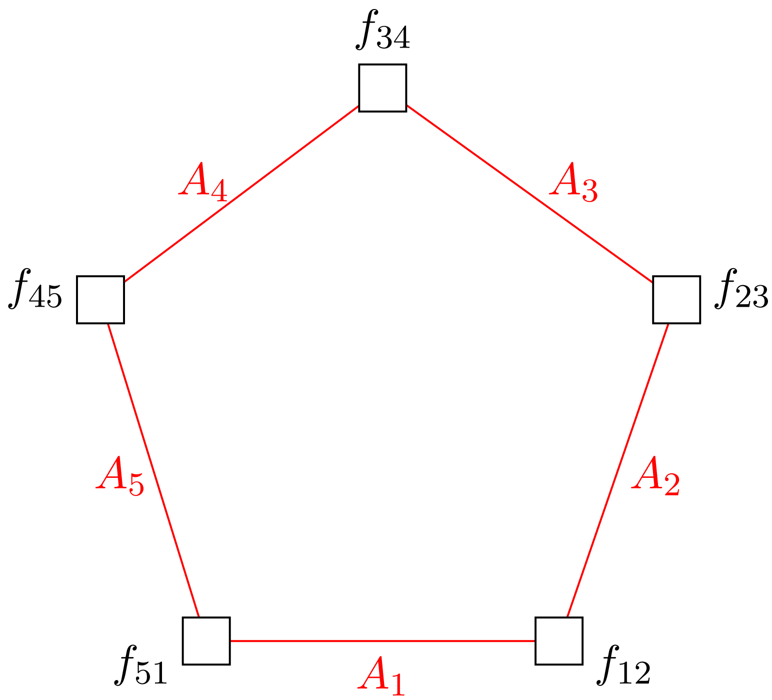
$$\mathbf{M}_f = \begin{pmatrix} f(0,0) & f(0,1) \\ f(1,0) & f(1,1) \end{pmatrix}.$$

- We assume that $\mathbf{M}_f = \mathbf{M}$ for all f, where $\mathbf{M}$ has
  - non-negative entries,
  - real eigenvalues λ_1 and λ_2 such that $\lambda_1 \geq |\lambda_2| \geq 0$.

Example 2: 5-Cycle



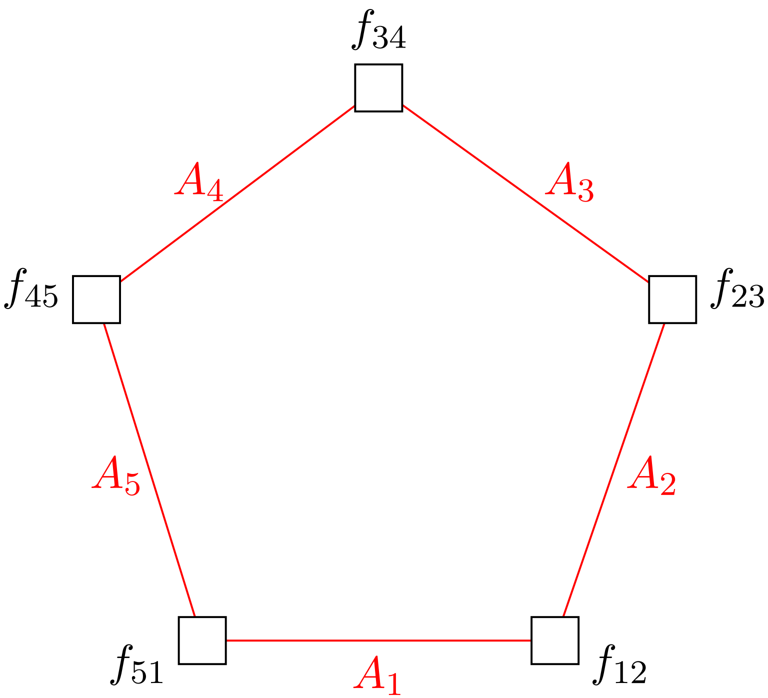
Example 2: 5-Cycle



- Partition sum:

$$Z(\mathsf{N}) = \mathrm{trace}(\mathbf{M}^5) = \lambda_1^5 + \lambda_2^5.$$

Example 2: 5-Cycle



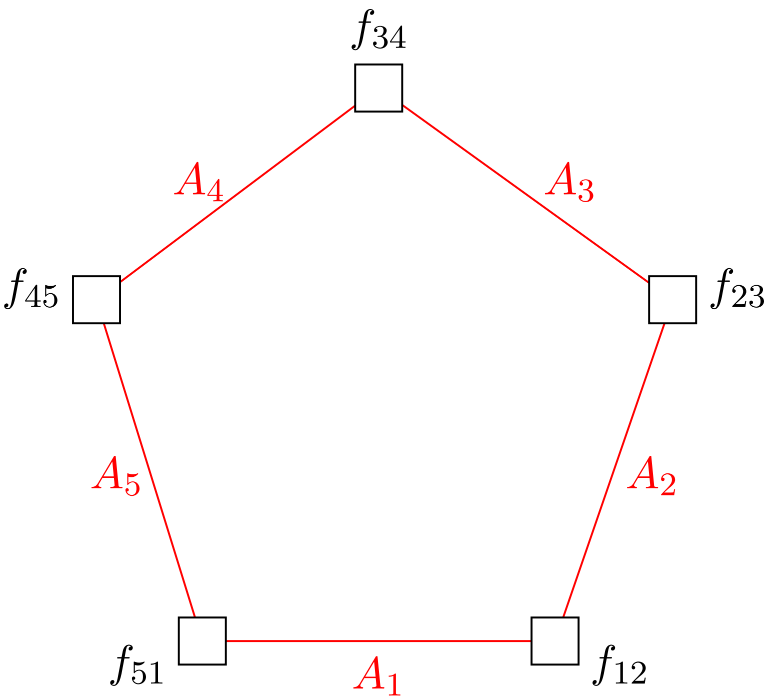
- Partition sum:

$$Z(\mathsf{N}) = \mathrm{trace}(\mathbf{M}^5) = \lambda_1^5 + \lambda_2^5.$$

- Degree-2 Bethe partition sum:

$$Z_{\mathrm{Bethe},2}(\mathsf{N}) = \sqrt[2]{\lambda_1^{10} + \lambda_1^5\lambda_2^5 + \lambda_2^{10}}.$$

Example 2: 5-Cycle



- Partition sum:

$$Z(\mathsf{N}) = \mathrm{trace}(\mathbf{M}^5) = \lambda_1^5 + \lambda_2^5.$$

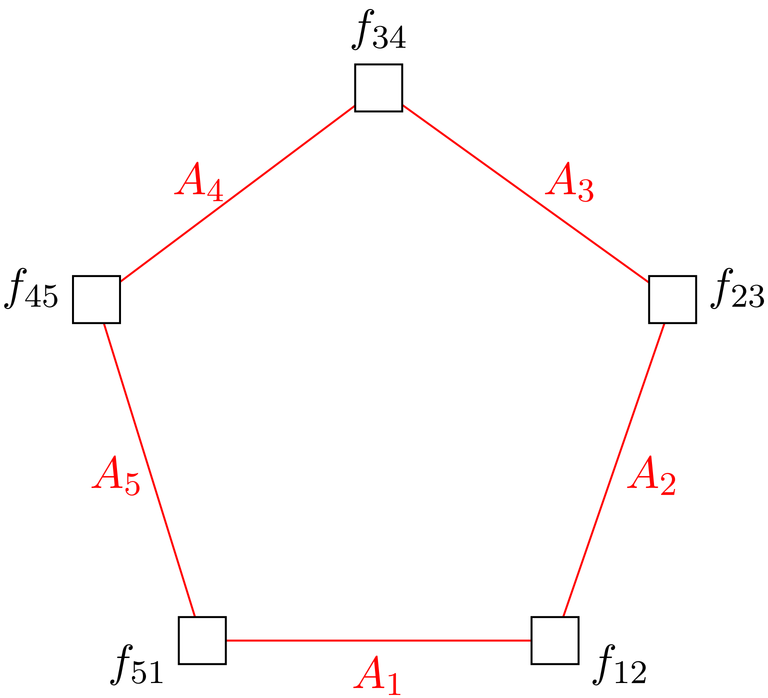
- Degree-2 Bethe partition sum:

$$Z_{\mathrm{Bethe},2}(\mathsf{N}) = \sqrt[2]{\lambda_1^{10} + \lambda_1^5 \lambda_2^5 + \lambda_2^{10}}.$$

- Degree-M Bethe partition sum:

$$Z_{\mathrm{Bethe},M}(\mathsf{N}) = \sqrt[M]{\sum_{m=0}^{M} \lambda_1^{5(M-m)} \lambda_2^{5m}}.$$

Example 2: 5-Cycle



- Partition sum:

$$Z(\mathsf{N}) = \mathrm{trace}(\mathbf{M}^5) = \lambda_1^5 + \lambda_2^5.$$

- Degree-2 Bethe partition sum:

$$Z_{\mathrm{Bethe},2}(\mathsf{N}) = \sqrt[2]{\lambda_1^{10} + \lambda_1^5 \lambda_2^5 + \lambda_2^{10}}.$$

- Degree-M Bethe partition sum:

$$Z_{\mathrm{Bethe},M}(\mathsf{N}) = \sqrt[M]{\sum_{m=0}^{M} \lambda_1^{5(M-m)} \lambda_2^{5m}}.$$

- Bethe partition sum:

$$Z_{\mathrm{Bethe}}(\mathsf{N}) = \lambda_1^5.$$

Log-Supermodular NFGs

Theorem 1 [Ruozzi 2012]

Let N be a binary log-supermodular NFG. Let $M \geq 1$. Then for any M-cover $\widetilde{\mathsf{N}}$ of N it holds that

$$Z(\widetilde{\mathsf{N}}) \leq Z(\mathsf{N})^M .$$

Log-Supermodular NFGs

Theorem 1 [Ruozzi 2012]

Let N be a binary log-supermodular NFG. Let $M \geq 1$. Then for any M-cover $\widetilde{\mathsf{N}}$ of N it holds that

$$Z(\widetilde{\mathsf{N}}) \leq Z(\mathsf{N})^M .$$

Theorem 2 [Ruozzi 2012]

Let N be a binary log-supermodular factor graph. Then

$$Z_{\mathrm{Bethe}}(\mathsf{N}) \leq Z(\mathsf{N}).$$

Log-Supermodular NFGs

Theorem 1 [Ruozzi 2012]

Let N be a binary log-supermodular NFG. Let $M \geq 1$. Then for any M-cover $\widetilde{\mathsf{N}}$ of N it holds that

$$Z(\widetilde{\mathsf{N}}) \leq Z(\mathsf{N})^M .$$

Theorem 2 [Ruozzi 2012]

Let N be a binary log-supermodular factor graph. Then

$$Z_{\text{Bethe}}(\mathsf{N}) \leq Z(\mathsf{N}).$$

Proof of Theorem 2:

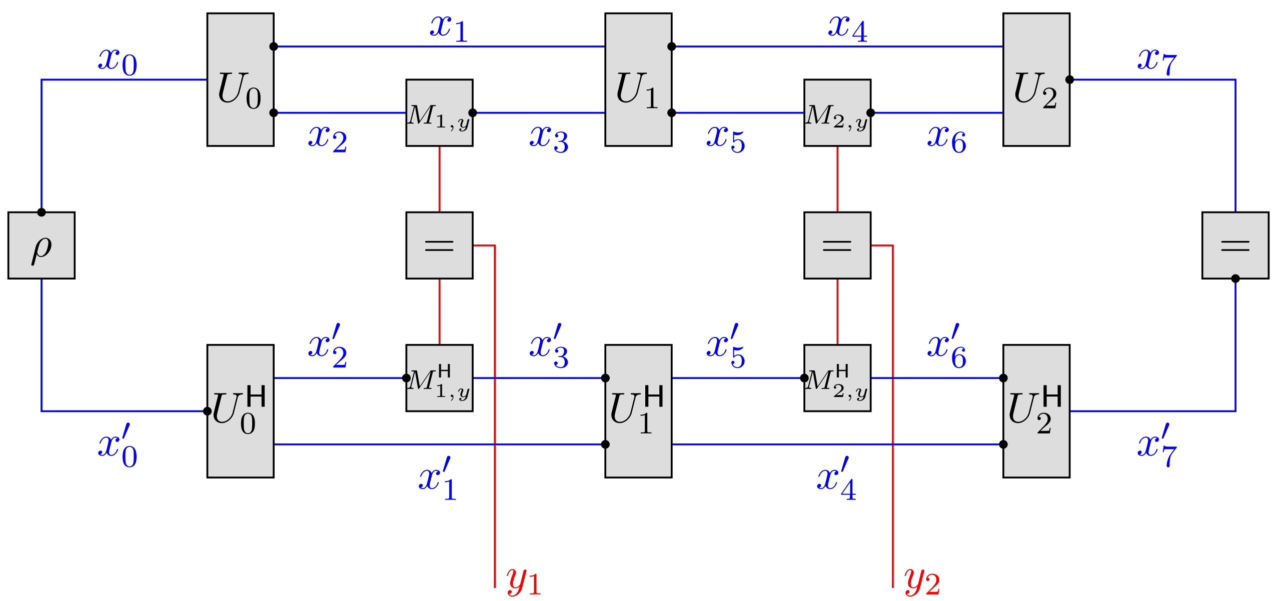
$$\begin{aligned} Z_{\text{Bethe}}(\mathsf{N}) &= \limsup_{M \to \infty} Z_{\text{Bethe},M}(\mathsf{N}) \\ &= \limsup_{M \to \infty} \sqrt[M]{\left\langle Z(\widetilde{\mathsf{N}}) \right\rangle_{\widetilde{\mathsf{N}} \in \widetilde{\mathcal{N}}_M}} \\ &\leq \limsup_{M \to \infty} \sqrt[M]{\left\langle Z(\mathsf{N})^M \right\rangle_{\widetilde{\mathsf{N}} \in \widetilde{\mathcal{N}}_M}} \\ &= Z(\mathsf{N}). \end{aligned}$$

Double-edge normal factor graphs (DE-NFGs)

Motivation for DE-NFGs: Part 1

(unitary evolutions and measurements)

Motivation for DE-NFGs



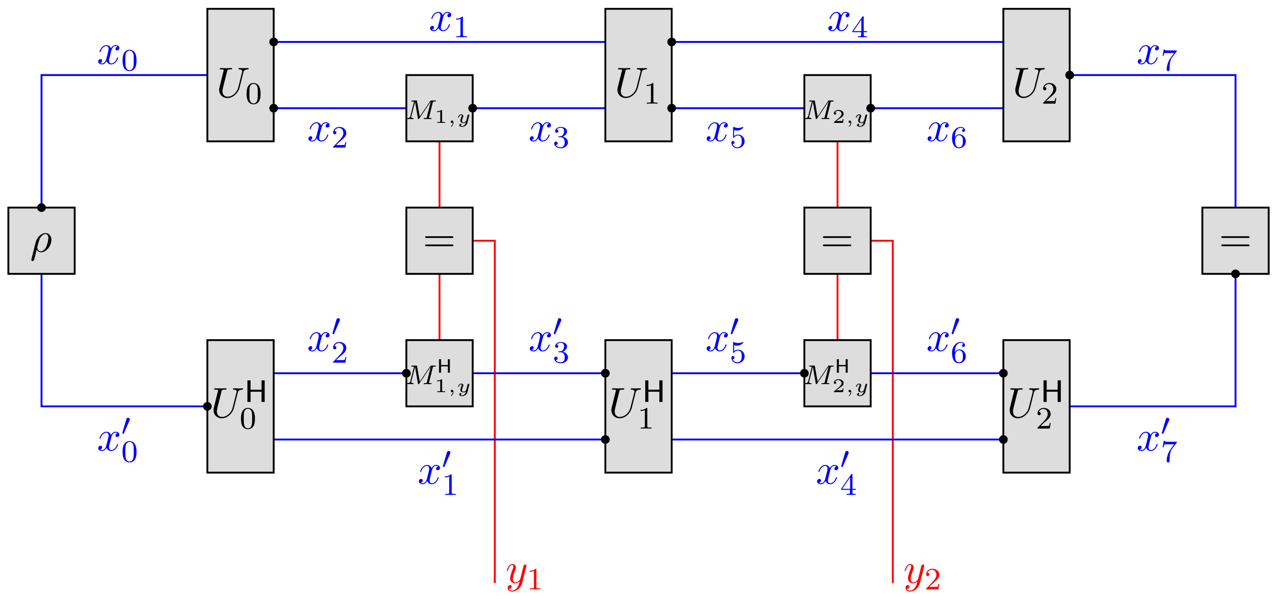
The above graphical model is an **NFG** that can be used to represent probabilities of interest in **quantum information processing** [Loeliger and Vontobel, ISIT 2012 and T-IT 2017].

Here:

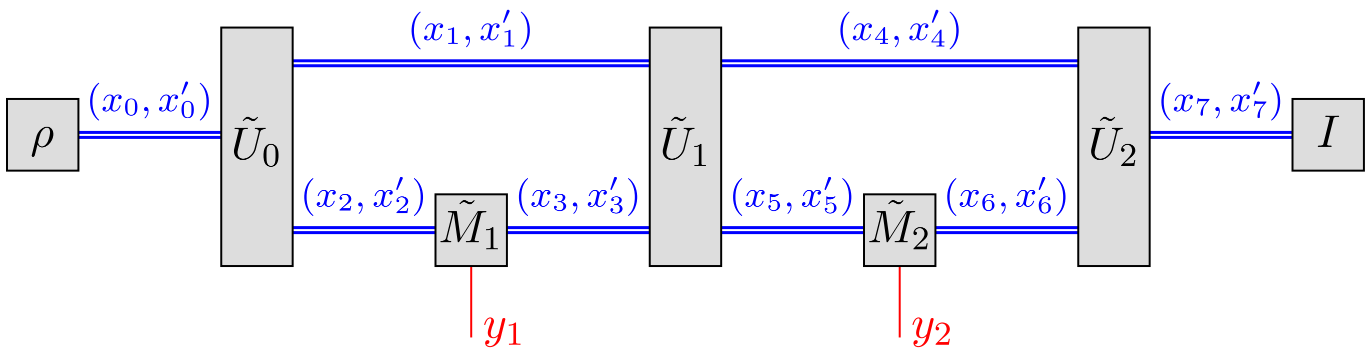
1. A system is prepared in some **state**.
2. The system **evolves unitarily**.
3. Part of the system is **measured**. $\rightarrow$ **Outcome** y_1.
4. The system **evolves unitarily**.
5. Part of the system is **measured**. $\rightarrow$ **Outcome** y_2.
6. The system **evolves unitarily**.

$$\Pr(Y_1 = y_1, Y_2 = y_2) = e(y_1, y_2)$$

From an NFG to a DE-NFG



After grouping pairs of blue variables and closing-the-box around suitable collections of function nodes, we obtain a graphical model that we call a **double-edge normal factor graph (DE-NFG)**.



Motivation for DE-NFGs: Part 2

(quantum teleportation)

Quantum Teleportation

Setup:

- Assume that Alice has a qubit **(Qubit 1)** in state ρ.
- Assume that Alice and Bob share an EPR pair **(Alice: Qubit 2; Bob: Qubit 3)**.
- Assume that Alice wants to transmit the state of **Qubit 1** (i.e., ρ) to Bob. However she is only allowed to use a classical channel, i.e., the **Qubit 1** cannot be sent via some quantum channel to Bob.

Quantum Teleportation

Setup:

- Assume that Alice has a qubit **(Qubit 1)** in state ρ.
- Assume that Alice and Bob share an EPR pair **(Alice: Qubit 2; Bob: Qubit 3)**.
- Assume that Alice wants to transmit the state of **Qubit 1** (i.e., ρ) to Bob. However she is only allowed to use a classical channel, i.e., the **Qubit 1** cannot be sent via some quantum channel to Bob.

Approach:

- Alice does some suitable operations and measurements on **Qubits 1 and 2**. Let the measurement results be $m_1, m_2 \in \{0, 1\}$.
- Alice transmits the measurement results m_1 and m_2 to Bob.
- Based on m_1 and m_2, Bob performs some operations on **Qubit 3**.
- In the end, **Qubit 3** will be in state ρ.

Quantum Teleportation

Setup:

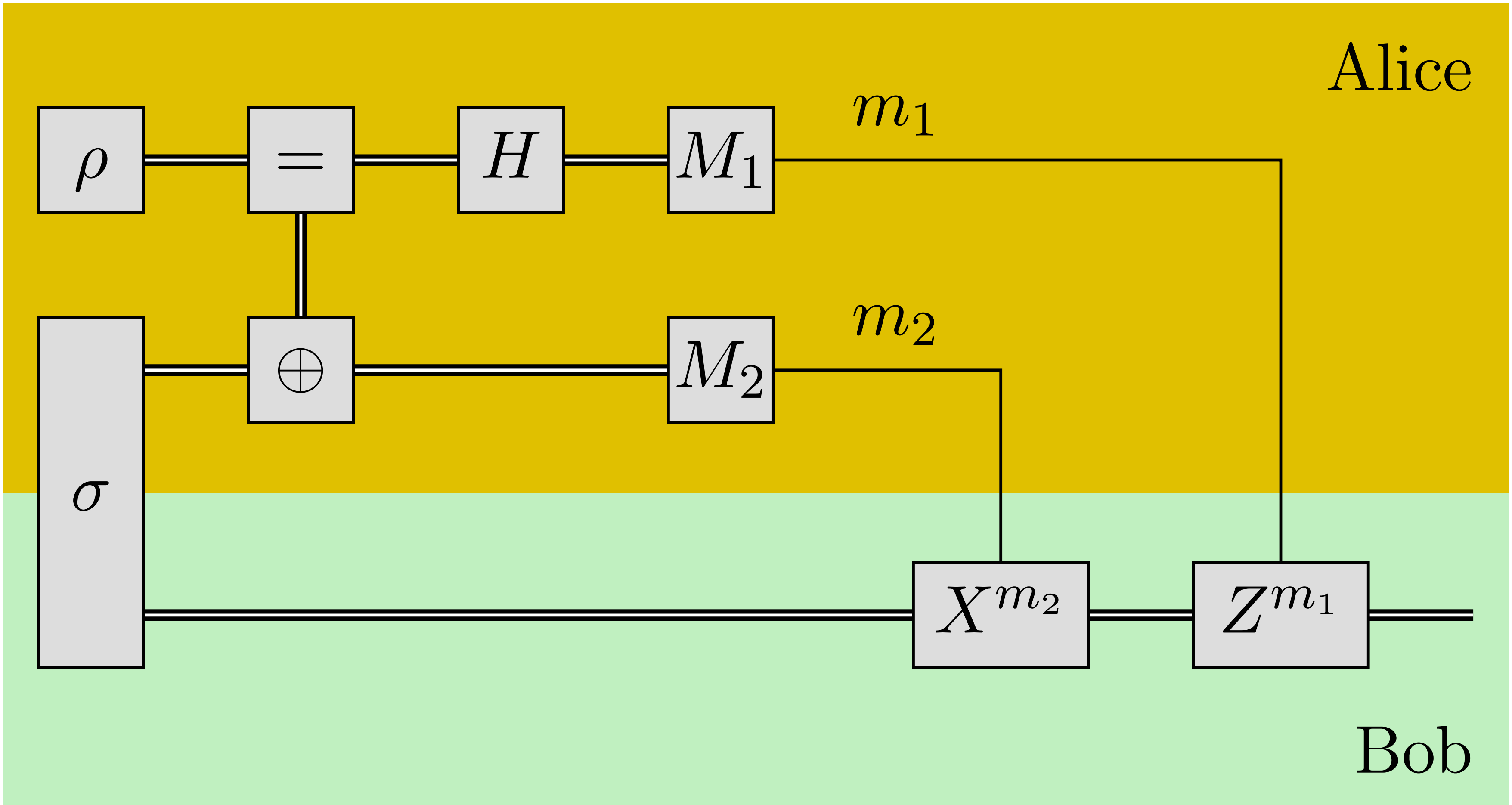
- Assume that Alice has a qubit **(Qubit 1)** in state ρ.
- Assume that Alice and Bob share an EPR pair **(Alice: Qubit 2; Bob: Qubit 3)**.
- Assume that Alice wants to transmit the state of **Qubit 1** (i.e., ρ) to Bob. However she is only allowed to use a classical channel, i.e., the **Qubit 1** cannot be sent via some quantum channel to Bob.

Approach:

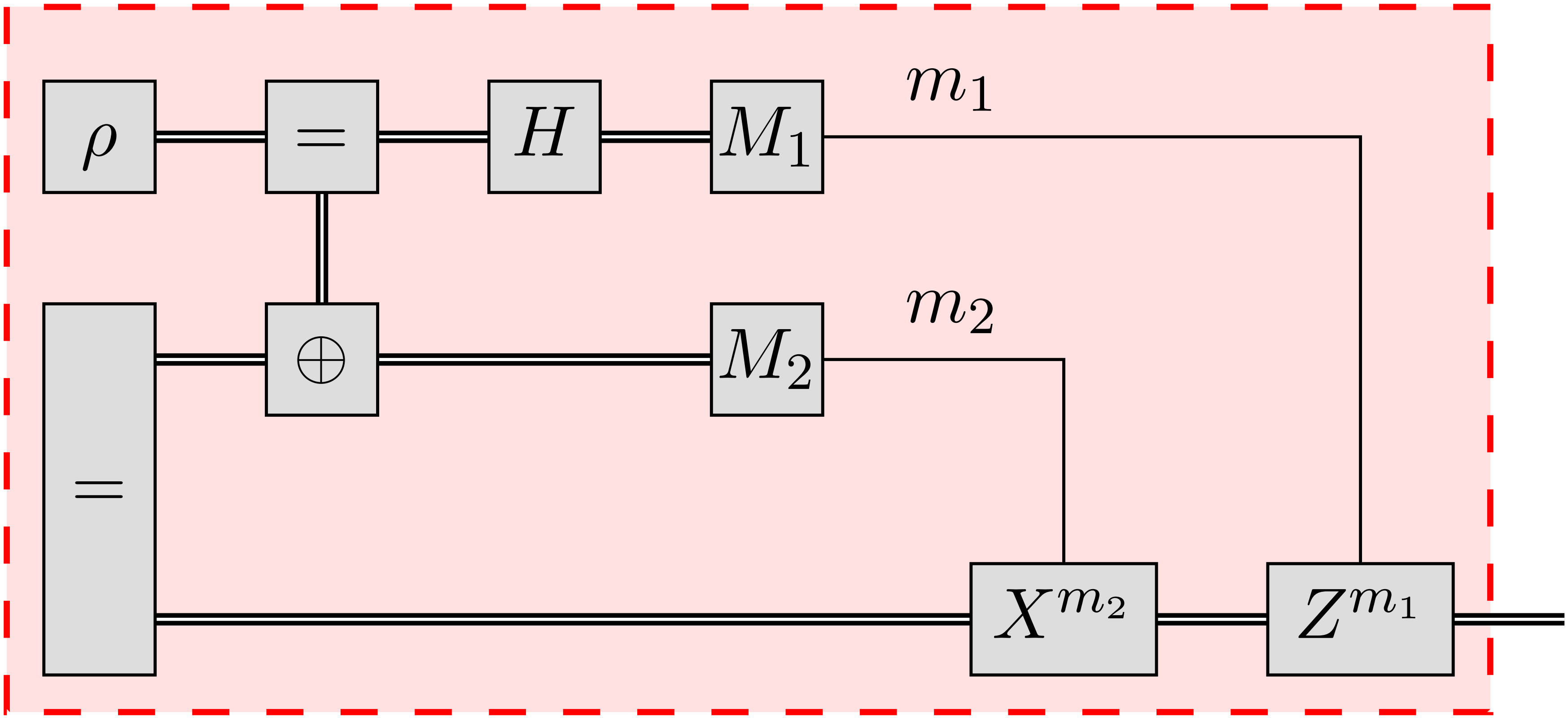
- Alice does some suitable operations and measurements on **Qubits 1 and 2**. Let the measurement results be $m_1, m_2 \in \{0, 1\}$.
- Alice transmits the measurement results m_1 and m_2 to Bob.
- Based on m_1 and m_2, Bob performs some operations on **Qubit 3**.
- In the end, **Qubit 3** will be in state ρ.

The following slides use DE-NFGs to show that the state of **Qubit 3** is indeed ρ.

Quantum Teleportation

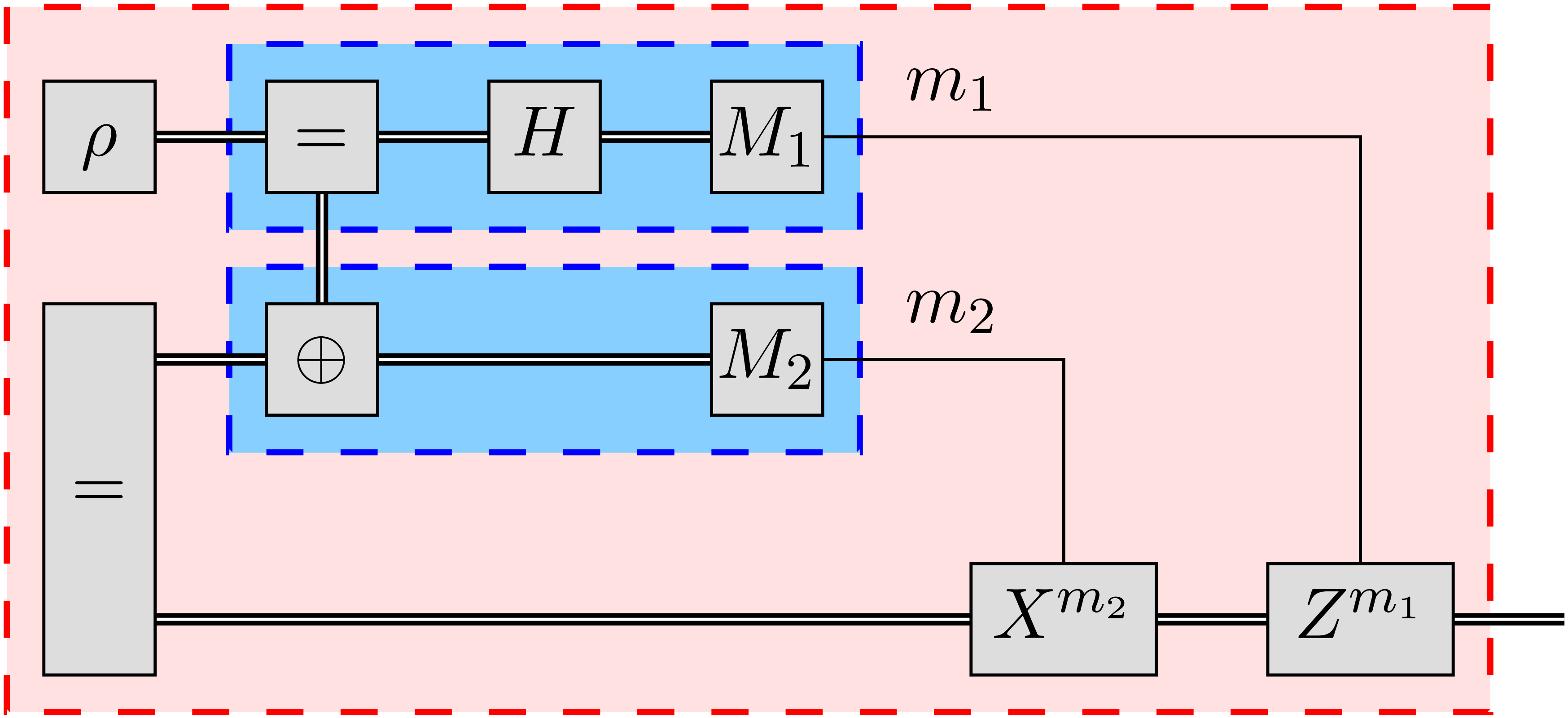


Quantum Teleportation



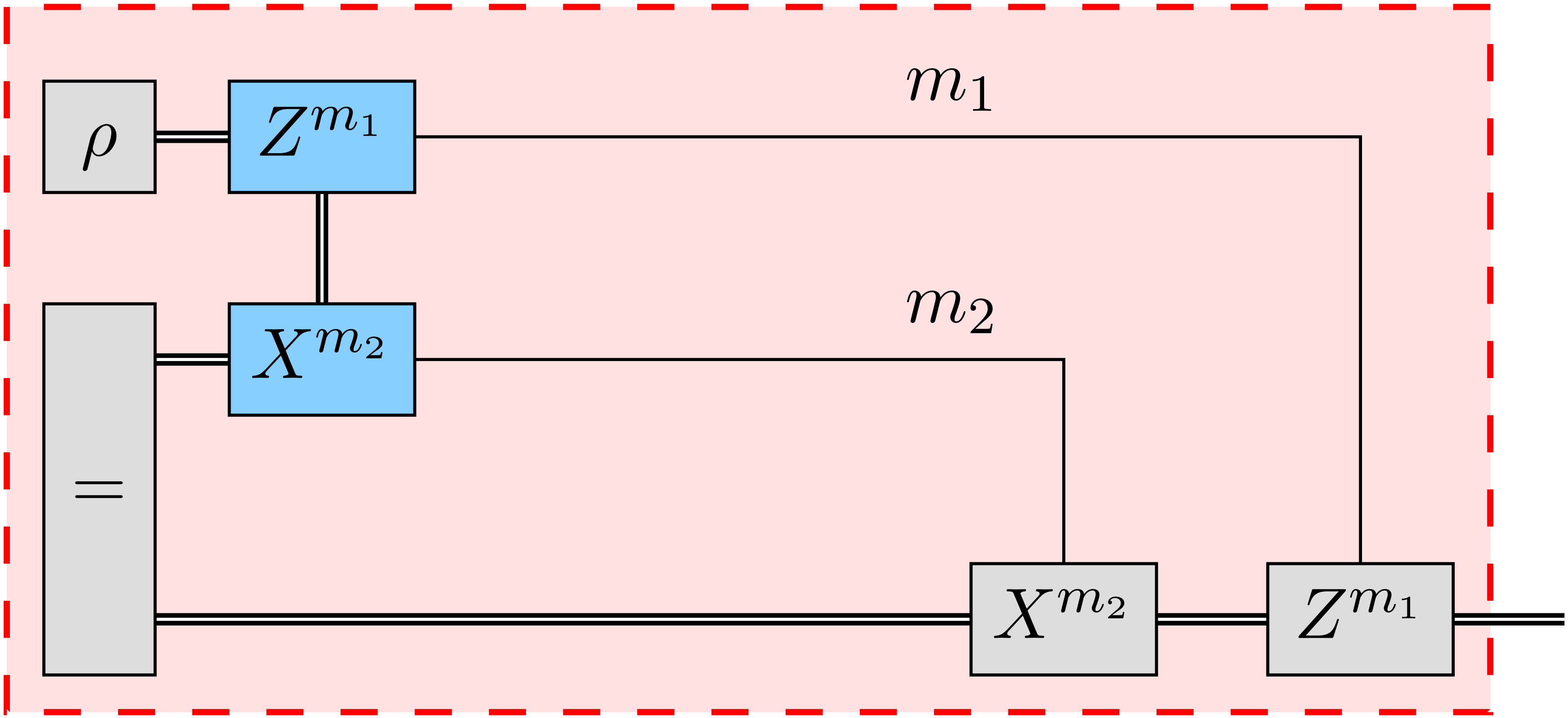
(Proportionality constants have been omitted.)

Quantum Teleportation



(Proportionality constants have been omitted.)

Quantum Teleportation



(Proportionality constants have been omitted.)

Definition of DE-NFGs

Definition of a DE-NFG

Definition: Consider the factorization

$$g(\mathbf{x}, \mathbf{x}'; \mathbf{y}) = \prod_{f \in \mathcal{F}} f(\mathbf{x}_{\partial f}, \mathbf{x}'_{\partial f}; \mathbf{y}_{\delta f})$$

represented by some DE-NFG. We will use the following conventions:

- We call g the **global function.**
- We call $f \in \mathcal{F}$ the **local functions**.
- For every function node $f \in \mathcal{F}$, the variables associated with the incident **double-edges** are collected in $\mathbf{x}_{\partial f}, \mathbf{x}'_{\partial f}$.
- For every function node $f \in \mathcal{F}$, the variables associated with the incident **single-edges** are collected in $\mathbf{y}_{\delta f}$.

(continued on next slide)

Definition of a DE-NFG

Definition (continued):

Most importantly, we require every **local function** $f \in \mathcal{F}$ to have the following property:

for every $\mathbf{y}_{\delta f}$, the **square matrix** $\big(f(\mathbf{x}_{\partial f}, \mathbf{x}'_{\partial f}; \mathbf{y}_{\delta f})\big)_{\mathbf{x}_{\partial f}, \mathbf{x}'_{\partial f}}$
with row indices $\mathbf{x}_{\partial f}$ and column indices $\mathbf{x}'_{\partial f}$ is a
complex-valued, hermitian, positive semi-definite (PSD) matrix.

Equivalently,

for every $\mathbf{y}_{\delta f}$, the function $f(\mathbf{x}_{\partial f}, \mathbf{x}'_{\partial f}; \mathbf{y}_{\delta f})$ is a
complex-valued, hermitian, positive semi-definite kernel function.

When a function node f has no incident double edges, then the above condition has to be understood as requiring the local function f to take on only non-negative real values.

Properties of DE-NFGs

The Partition Sum of a DE-NFG

Definition: Consider some DE-NFG. The **partition sum** associated with this DE-NFG is defined to be

$$Z \triangleq \sum_{\mathbf{x}, \mathbf{x}', \mathbf{y}} g(\mathbf{x}, \mathbf{x}'; \mathbf{y}) \,.$$

The Partition Sum of a DE-NFG

Definition: Consider some DE-NFG. The **partition sum** associated with this DE-NFG is defined to be

$$Z \triangleq \sum_{\mathbf{x}, \mathbf{x}', \mathbf{y}} g(\mathbf{x}, \mathbf{x}'; \mathbf{y}) \, .$$

Proposition:

The partition sum of a DE-NFG is a **non-negative real number**.

Sum-Product Algorithm for DE-NFGs

Assumptions: We make the following assumptions about the **initial messages**, i.e., about the messages at time $t = 0$:

- **Messages along single-edges** are

  positive real-valued functions.

- **Messages along double-edges** are

  complex-valued, hermitian, positive definite kernel functions.

Sum-Product Algorithm for DE-NFGs

Assumptions: We make the following assumptions about the **initial messages**, i.e., about the messages at time $t = 0$:

- **Messages along single-edges** are

  positive real-valued functions.

- **Messages along double-edges** are

  complex-valued, hermitian, positive definite kernel functions.

Proposition: Let the messages be initialized as specified above.
Then **for every iteration** $t \geq 1$ it holds that:

- **Messages along single-edges** are

  non-negative real-valued functions.

- **Messages along double-edges** are

  complex-valued, hermitian, positive semi-definite kernel functions.

Reminder: Bethe Approx. for S-NFGs

Primal formulation:

$$Z_{\text{Bethe}} \triangleq \exp\left(- \min_{\boldsymbol{\beta}} F_{\text{Bethe}}(\boldsymbol{\beta}) \right).$$

Pseudo-dual formulation:

$$Z_{\text{Bethe}} = \max_{\substack{\text{SPA messages} \\ \text{fixed point } \boldsymbol{\mu}}} \frac{\prod_{f \in \mathcal{F}} Z_f(\boldsymbol{\mu})}{\prod_{e \in \mathcal{E}_{\text{full}}} Z_e(\boldsymbol{\mu})},$$

where

$$Z_f(\boldsymbol{\mu}) \triangleq \sum_{\mathbf{a}_f} f(\mathbf{a}_f) \cdot \prod_{e \in \partial f} \mu_{e \to f}(a_{f,e}), \quad f \in \mathcal{F},$$

$$Z_e(\boldsymbol{\mu}) \triangleq \sum_{a_e} \mu_{e \to f}(a_e) \cdot \mu_{e \to f'}(a_e), \qquad e = (f, f') \in \mathcal{E}_{\text{full}}.$$

Reminder: Bethe Approx. for S-NFGs

Primal formulation:

$$Z_{\text{Bethe}} \triangleq \exp\left(- \min_{\boldsymbol{\beta}} F_{\text{Bethe}}(\boldsymbol{\beta}) \right) .$$

Pseudo-dual formulation:

$$Z_{\text{Bethe}} = \max_{\substack{\text{SPA messages} \\ \text{fixed point } \boldsymbol{\mu}}} \frac{\prod_{f \in \mathcal{F}} Z_f(\boldsymbol{\mu})}{\prod_{e \in \mathcal{E}_{\text{full}}} Z_e(\boldsymbol{\mu})} ,$$

(generalization to DE-NFGs straightforward)

where

$$Z_f(\boldsymbol{\mu}) \triangleq \sum_{\mathbf{a}_f} f(\mathbf{a}_f) \cdot \prod_{e \in \partial f} \mu_{e \to f}(a_{f,e}) , \quad f \in \mathcal{F} ,$$

$$Z_e(\boldsymbol{\mu}) \triangleq \sum_{a_e} \mu_{e \to f}(a_e) \cdot \mu_{e \to f'}(a_e) , \qquad e = (f, f') \in \mathcal{E}_{\text{full}} .$$

Reminder: Bethe Approx. for S-NFGs

Primal formulation:

$$Z_{\text{Bethe}} \triangleq \exp\left(- \min_{\boldsymbol{\beta}} F_{\text{Bethe}}(\boldsymbol{\beta}) \right) .$$

(generalization to DE-NFGs unclear)

Pseudo-dual formulation:

$$Z_{\text{Bethe}} = \max_{\substack{\text{SPA messages} \\ \text{fixed point } \boldsymbol{\mu}}} \frac{\prod_{f \in \mathcal{F}} Z_f(\boldsymbol{\mu})}{\prod_{e \in \mathcal{E}_{\text{full}}} Z_e(\boldsymbol{\mu})} ,$$

(generalization to DE-NFGs straightforward)

where

$$Z_f(\boldsymbol{\mu}) \triangleq \sum_{\mathbf{a}_f} f(\mathbf{a}_f) \cdot \prod_{e \in \partial f} \mu_{e \to f}(a_{f,e}) , \quad f \in \mathcal{F} ,$$

$$Z_e(\boldsymbol{\mu}) \triangleq \sum_{a_e} \mu_{e \to f}(a_e) \cdot \mu_{e \to f'}(a_e) , \qquad e = (f, f') \in \mathcal{E}_{\text{full}} .$$

The Bethe Partition Sum of a DE-NFG

Definition: Consider a collection of SPA messages $\boldsymbol{\mu} = \{\mu_{e \to f}\}_{f \in \mathcal{F}, e \in \mathcal{N}(f)}$, i.e., one message for every edge in both directions. Let

$$Z_{\text{Bethe}}(\boldsymbol{\mu}) \triangleq \frac{\prod_{f \in \mathcal{F}} Z_f(\boldsymbol{\mu})}{\prod_{e \in \mathcal{E}_{\text{full}}} Z_e(\boldsymbol{\mu})} ,$$

where

- for every $f \in \mathcal{F}$ we define

$$Z_f(\boldsymbol{\mu}) \triangleq \sum_{\mathbf{x}_{\partial f}, \mathbf{x}'_{\partial f}, \mathbf{y}_{\delta f}} f(\mathbf{x}_{\partial f}, \mathbf{x}'_{\partial f}; \mathbf{y}_{\delta f}) \cdot \left(\prod_{e \in \partial f} \mu_{e \to f}(x_e, x'_e) \right) \cdot \left(\prod_{e \in \delta f} \mu_{e \to f}(y_e) \right) ,$$

- for every single-edge $e = (f, f') \in \mathcal{E}$ we define

$$Z_e(\boldsymbol{\mu}) \triangleq \sum_{y_e} \mu_{e \to f}(y_e) \cdot \mu_{e \to f'}(y_e) ,$$

- for every double-edge $e = (f, f') \in \mathcal{E}$ we define

$$Z_e(\boldsymbol{\mu}) \triangleq \sum_{x_e, x'_e} \mu_{e \to f}(x_e, x'_e) \cdot \mu_{e \to f'}(x_e, x'_e) .$$

The Bethe Partition Sum of a DE-NFG

Proposition:

The function $Z_{\mathrm{Bethe}}(\boldsymbol{\mu})$ in previous definition has the following properties:

The Bethe Partition Sum of a DE-NFG

Proposition:

The function $Z_{\mathrm{Bethe}}(\boldsymbol{\mu})$ in previous definition has the following properties:

- Assume
  - that messages have the properties listed in the previous proposition;
  - that $Z_{\mathrm{Bethe}}(\boldsymbol{\mu})$ is well-defined, *i.e.*, $Z_e(\boldsymbol{\mu}) \neq 0$ for all $e \in \mathcal{E}$.

  Then

  $$Z_{\mathrm{Bethe}}(\boldsymbol{\mu}) \text{ is a non-negative real number.}$$

The Bethe Partition Sum of a DE-NFG

Proposition:

The function $Z_{\mathrm{Bethe}}(\boldsymbol{\mu})$ in previous definition has the following properties:

- Assume
  - that messages have the properties listed in the previous proposition;
  - that $Z_{\mathrm{Bethe}}(\boldsymbol{\mu})$ is well-defined, *i.e.*, $Z_e(\boldsymbol{\mu}) \neq 0$ for all $e \in \mathcal{E}$.

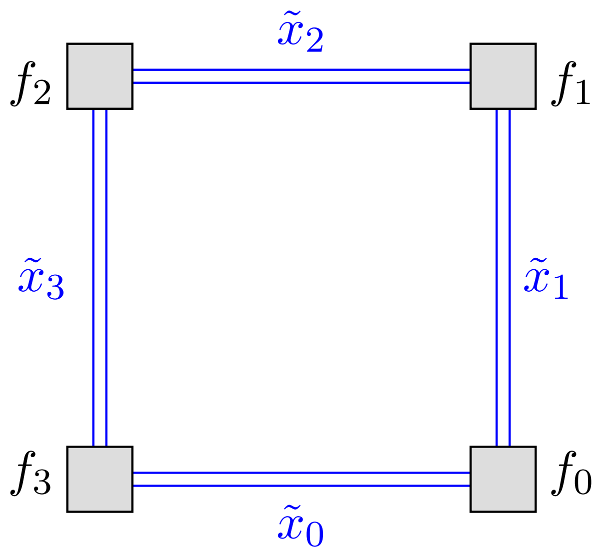
  Then

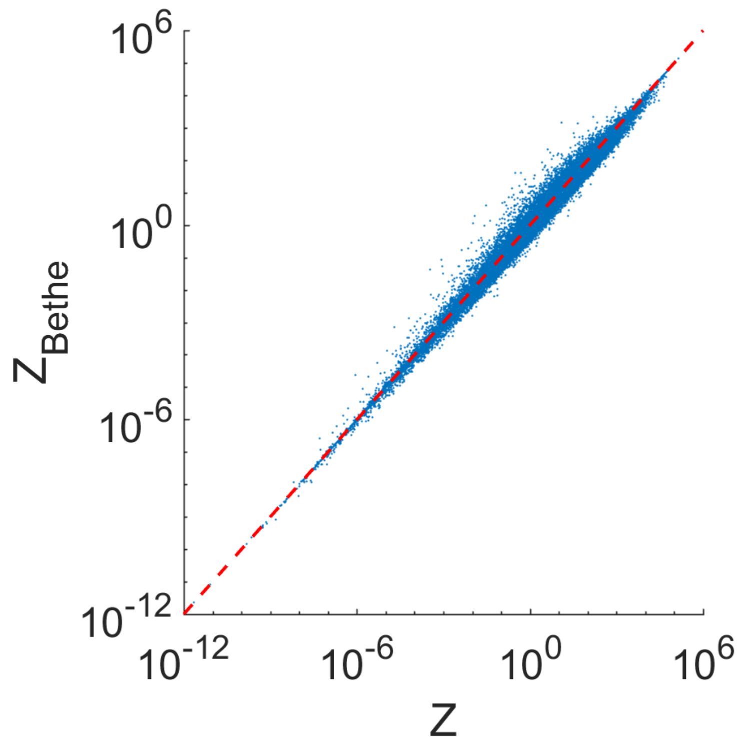
  $Z_{\mathrm{Bethe}}(\boldsymbol{\mu})$ **is a non-negative real number.**

- **Fixed points of the SPA $\hat{=}$ stationary points of the function $Z_{\mathrm{Bethe}}(\boldsymbol{\mu})$.** (This generalizes a theorem by Yedidia, Freeman, and Weiss.)

Examples of DE-NFGs

DE-NFG Example 1

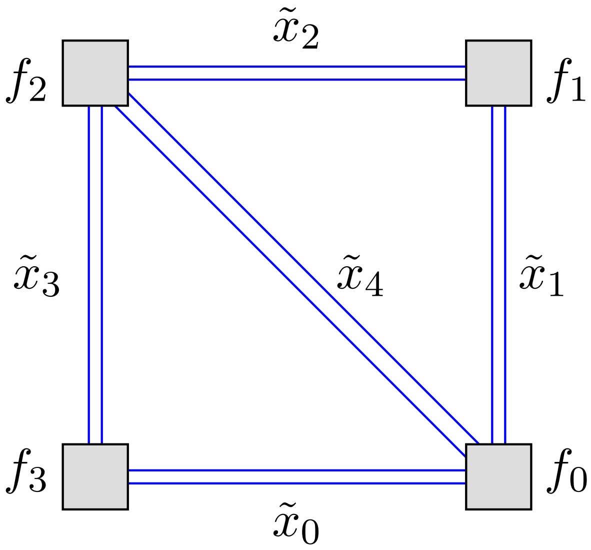


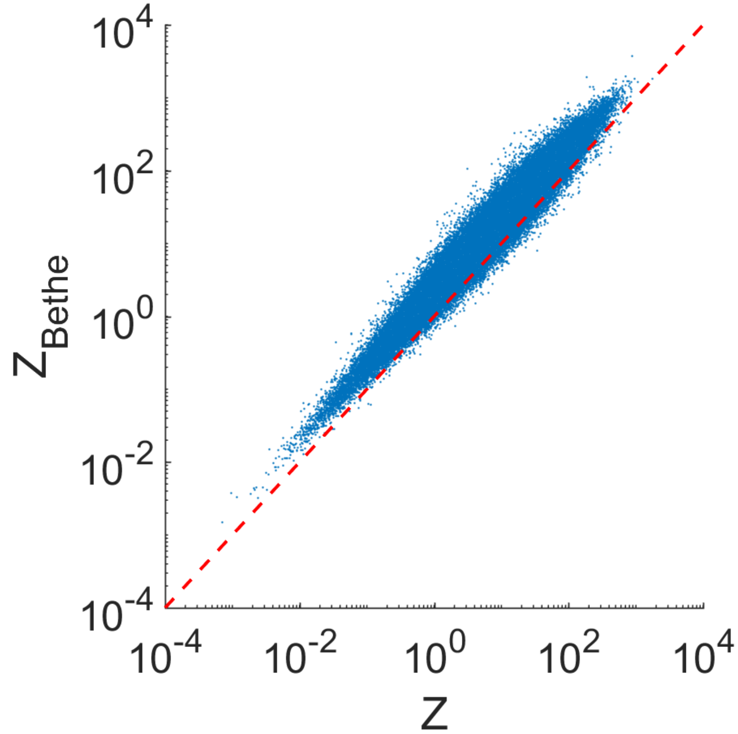


Setup for simulation results:

- $n = 4\,;\ \ |\mathcal{X}| = 2\,;\ \ 10^6$ experiments.
- $\mathbf{F} \triangleq \mathbf{U} \cdot \mathbf{D} \cdot \mathbf{U}^{\mathsf{H}}$ is **randomly generated** according to the following: procedure:
  - where $\mathbf{U}$ is a randomly generated unitary matrix (Haar measure),
  - where $\mathbf{D}$ is a diagonal matrix with i.i.d. diagonal entries sampled from a standard χ^2 distribution with one degree of freedom.

DE-NFG Example 2





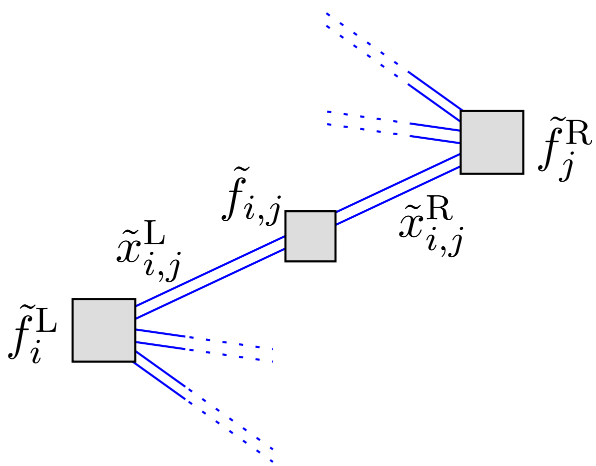
Setup for simulation results:

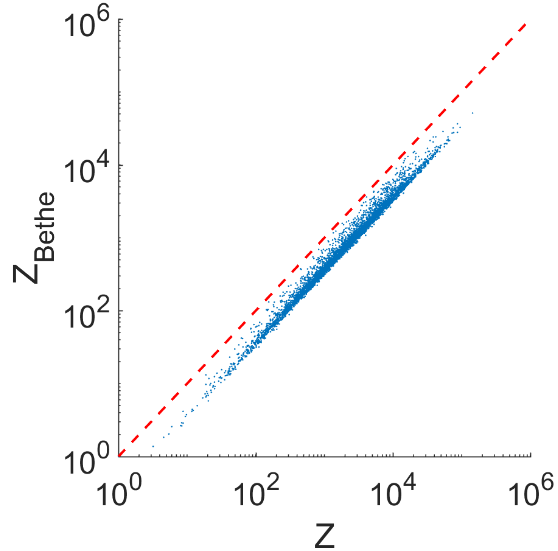
- $|\mathcal{X}| = 2\,;\quad 10^6$ experiments.
- For every instantiation, all local functions are generated independently.

  (This is In contrast to Example 1, where for every instantiation all local function were the same.)

DE-NFG Example 3

- Consider a certain type of **quantum computer based on linear optics** proposed by Aaronson and Arkhipov (2013).
- **Probabilities** that appear in that paper can be written as the partition sum of suitable DE-NFGs.
- Here we consider DE-NFGS that are generalizations of these DE-NFGs.
- These DE-NFGs are also generalizations of NFGs that appear when (approximately) computing **permanents of matrices**.





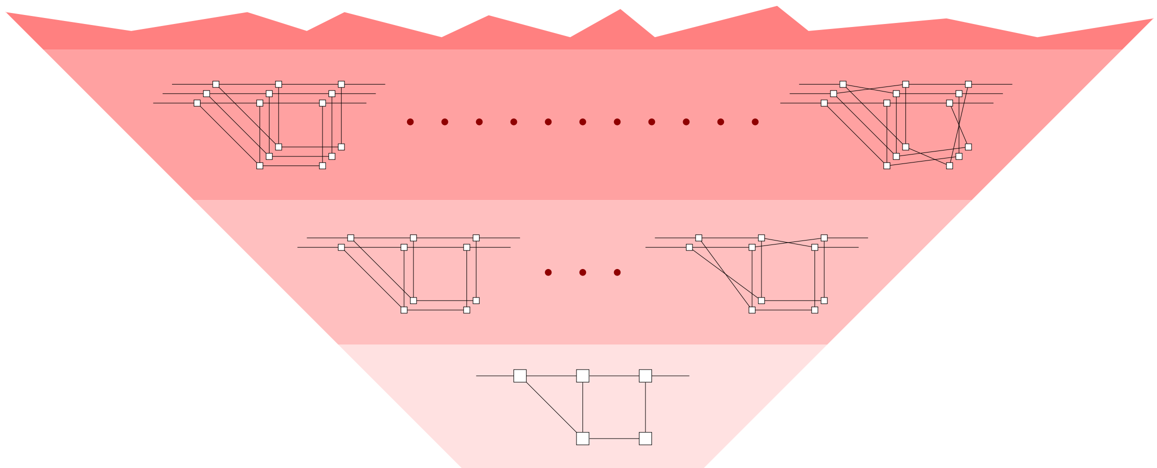
A combinatorial interpretation of the Bethe partition sum

Reminder: Z_{Bethe} for S-NFGs

$$Z_{\mathrm{Bethe},M}(\mathsf{N})\Big|_{M\to\infty} = Z_{\mathrm{Bethe}}(\mathsf{N})$$ **(Theorem [V., 2013])**

$$Z_{\mathrm{Bethe},M}(\mathsf{N})$$

$$Z_{\mathrm{Bethe},M}(\mathsf{N})\Big|_{M=1} = Z(\mathsf{N})$$



$$Z_{\mathrm{Bethe},M}(\mathsf{N}) \triangleq \sqrt[M]{\Big\langle Z(\widetilde{\mathsf{N}}) \Big\rangle_{\widetilde{\mathsf{N}} \in \widetilde{\mathcal{N}}_M}}$$

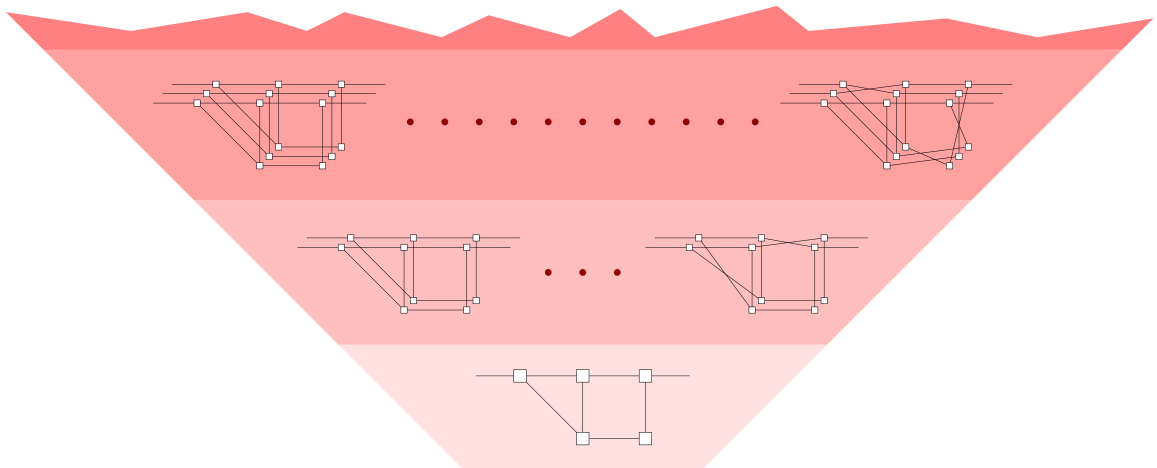
Reminder: Z_{Bethe} for S-NFGs

Does a similar theorem hold for DE-NFGs?

$$Z_{\mathrm{Bethe},M}(\mathsf{N})\Big|_{M\to\infty} = Z_{\mathrm{Bethe}}(\mathsf{N})$$ **(Theorem [V., 2013])**

$$Z_{\mathrm{Bethe},M}(\mathsf{N})$$

$$Z_{\mathrm{Bethe},M}(\mathsf{N})\Big|_{M=1} = Z(\mathsf{N})$$



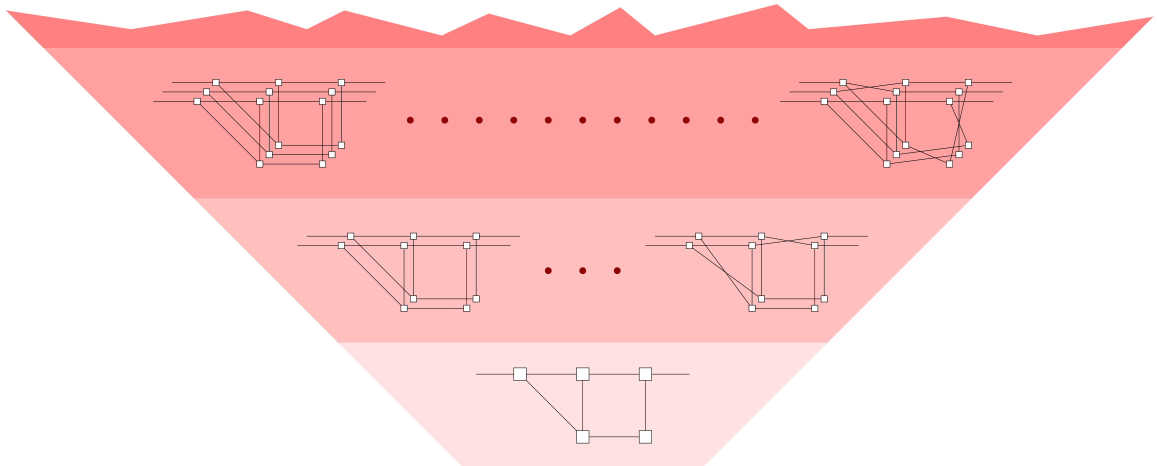
$$Z_{\mathrm{Bethe},M}(\mathsf{N}) \triangleq \sqrt[M]{\Big\langle Z(\widetilde{\mathsf{N}}) \Big\rangle_{\widetilde{\mathsf{N}} \in \widetilde{\mathcal{N}}_M}}$$

Reminder: Z_{Bethe} for S-NFGs

Problem: the proof for S-NFGs (based on the method of types) does not generalize to DE-NFGs.

$$Z_{\mathrm{Bethe},M}(\mathsf{N})\Big|_{M\to\infty} = Z_{\mathrm{Bethe}}(\mathsf{N})$$

(Theorem [V., 2013])



$$Z_{\mathrm{Bethe},M}(\mathsf{N})$$

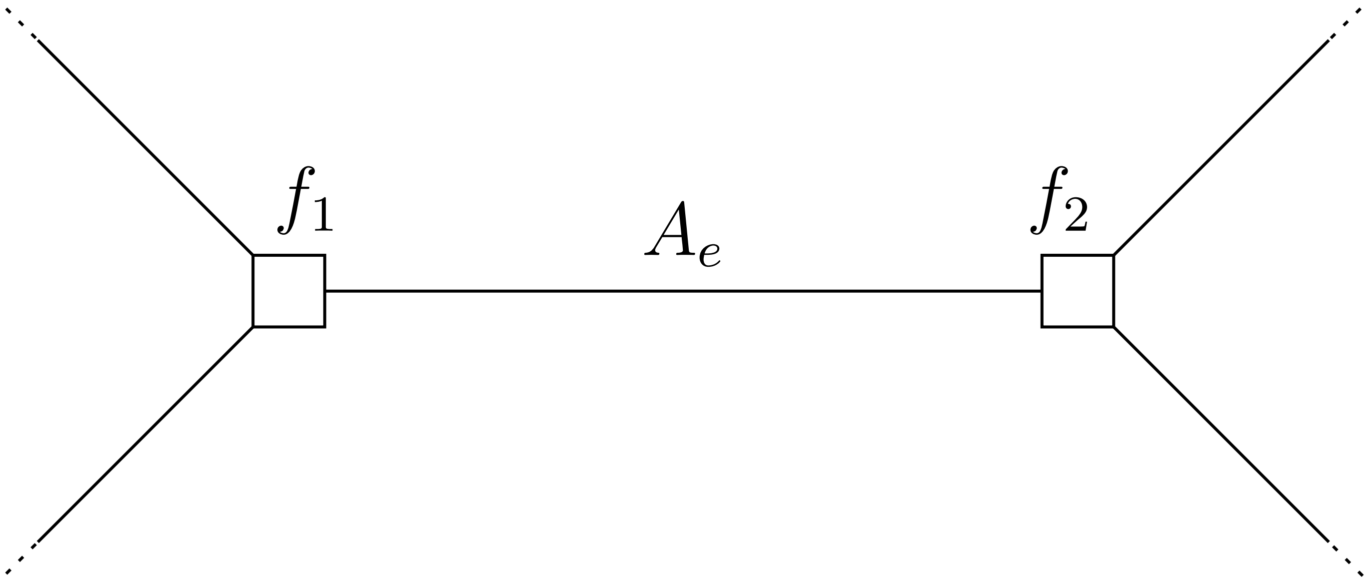
$$Z_{\mathrm{Bethe},M}(\mathsf{N})\Big|_{M=1} = Z(\mathsf{N})$$

$$Z_{\mathrm{Bethe},M}(\mathsf{N}) \triangleq \sqrt[M]{\left\langle Z(\widetilde{\mathsf{N}}) \right\rangle_{\widetilde{\mathsf{N}} \in \widetilde{\mathcal{N}}_M}}$$

Symmetric-subspace transform (SST)

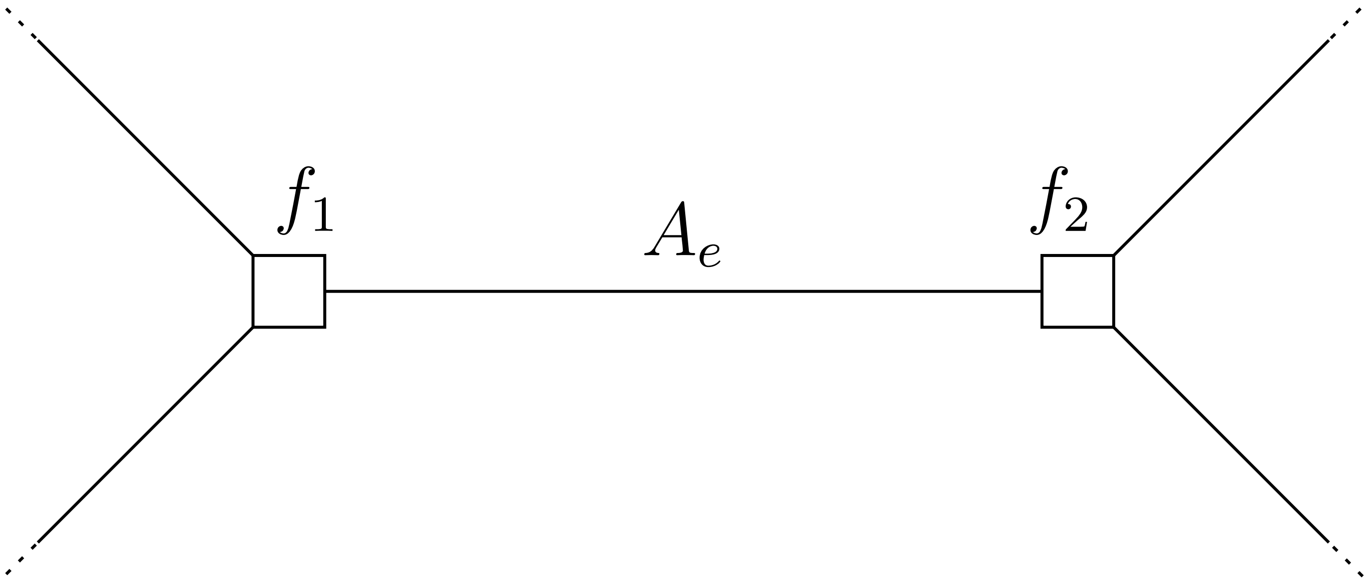
Symmetric-Subspace Transform

Assume that some part of our S-NFG N looks like this:



Symmetric-Subspace Transform

Assume that some part of our S-NFG N looks like this:

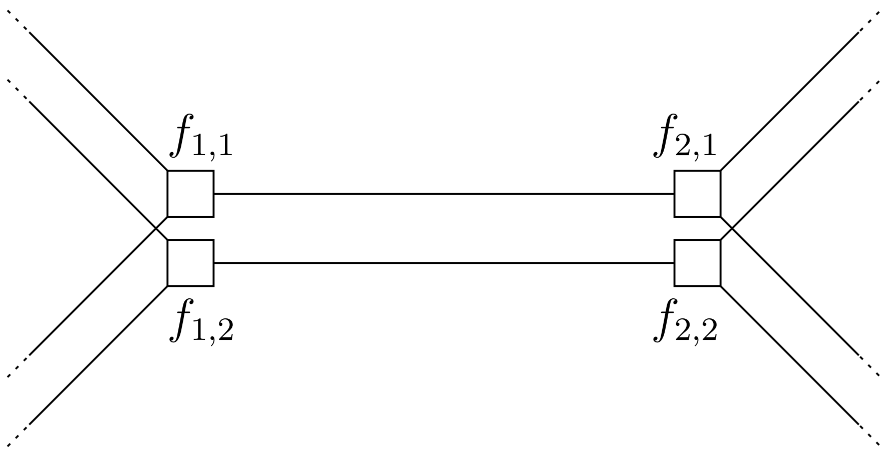


In the following, for simplicity, we assume that all variable alphabets are $\{0, 1\}$.

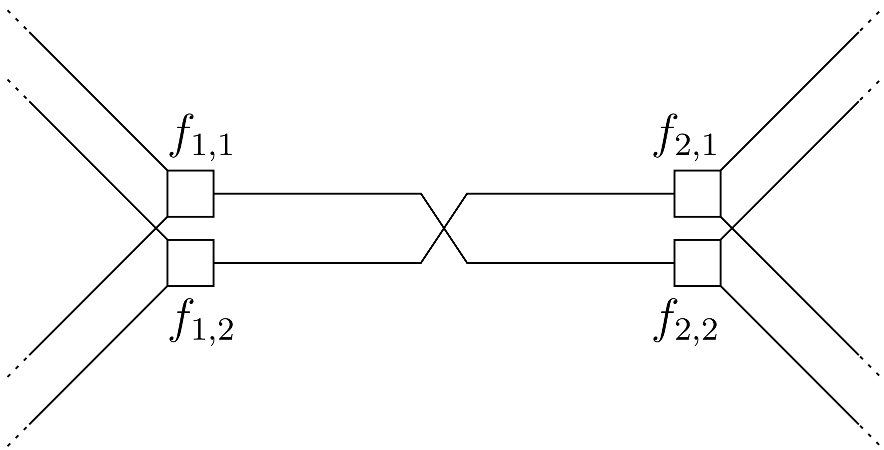
Symmetric-Subspace Transform

Let $\widetilde{N}$ be arbitrary double cover of N.

The corresponding part of $\widetilde{N}$ will either look like this

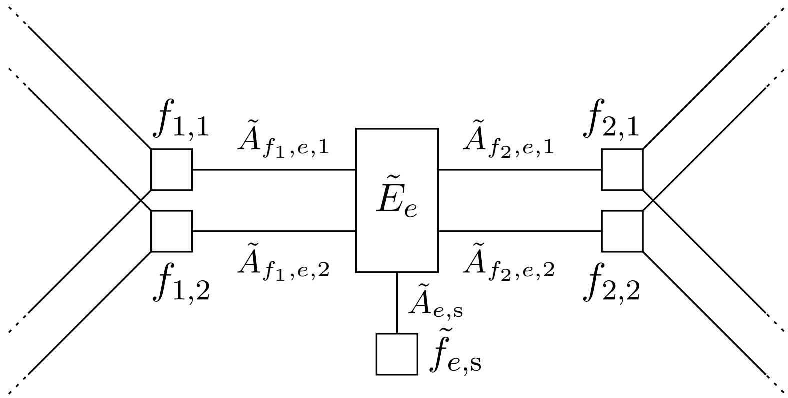


or like this



Symmetric-Subspace Transform

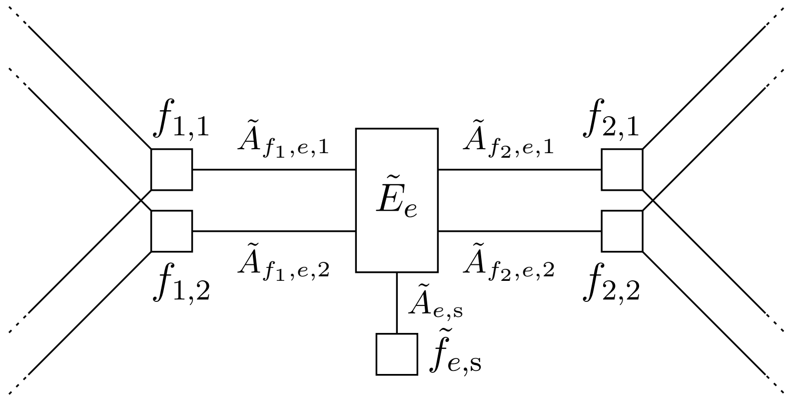
Independently of what the double cover looks like, its partition sum is equal to the partition sum of the following NFG



with suitably chosen function nodes.

Symmetric-Subspace Transform

Independently of what the double cover looks like, its partition sum is equal to the partition sum of the following NFG



with suitably chosen function nodes.

In particular, the matrices associated with

$$\tilde{E}_e\Big((\tilde{a}_{f_1,e,1}, \tilde{a}_{f_1,e,2}), (\tilde{a}_{f_2,e,1}, \tilde{a}_{f_2,e,2}), \tilde{a}_{e,\mathrm{s}} = 0\Big) ,$$

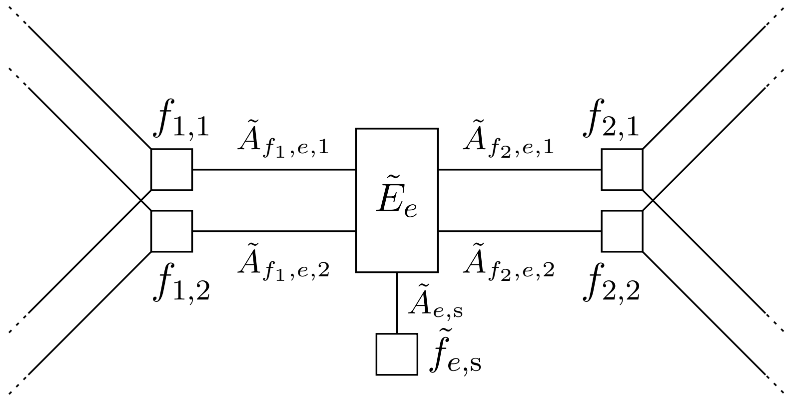
$$\tilde{E}_e\Big((\tilde{a}_{f_1,e,1}, \tilde{a}_{f_1,e,2}), (\tilde{a}_{f_2,e,1}, \tilde{a}_{f_2,e,2}), \tilde{a}_{e,\mathrm{s}} = 1\Big)$$

are, respectively,

$$\widetilde{\mathbf{E}}_{\mathrm{nocross}} \triangleq \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \widetilde{\mathbf{E}}_{\mathrm{cross}} \triangleq \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Symmetric-Subspace Transform

Independently of what the double cover looks like, its partition sum is equal to the partition sum of the following NFG



with suitably chosen function nodes.

Moreover, we defined

$$\tilde{f}_{e,\mathrm{s}}(0) \triangleq 1, \quad \tilde{f}_{e,\mathrm{s}}(1) \triangleq 0 \quad \text{(no crossing)},$$

$$\tilde{f}_{e,\mathrm{s}}(0) \triangleq 0, \quad \tilde{f}_{e,\mathrm{s}}(1) \triangleq 1 \quad \text{(crossing)}.$$

Symmetric-Subspace Transform

Note that

$$\widetilde{\mathbf{E}}_{\text{avg}} \triangleq \frac{1}{2} \cdot \widetilde{\mathbf{E}}_{\text{nocross}} + \frac{1}{2} \cdot \widetilde{\mathbf{E}}_{\text{cross}}$$

$$= \frac{1}{2} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \frac{1}{2} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Symmetric-Subspace Transform

Let

$$\boldsymbol{\psi} \triangleq \begin{pmatrix} \psi_0 \\ \psi_1 \end{pmatrix} \in \mathbb{C}^2 .$$

It follows that

$$\boldsymbol{\psi}^{\otimes 2} = \boldsymbol{\psi} \otimes \boldsymbol{\psi} = \begin{pmatrix} \psi_0 \cdot \psi_0 \\ \psi_0 \cdot \psi_1 \\ \psi_1 \cdot \psi_0 \\ \psi_1 \cdot \psi_1 \end{pmatrix} .$$

Symmetric-Subspace Transform

Assume that ψ is uniformly distributed among all vectors in $\mathbb{C}^2$ of length one. Then consider the matrix

$$\mathbf{M} \triangleq \mathsf{E}\left[\psi^{\otimes 2} \cdot \left(\psi^{\otimes 2}\right)^{\mathsf{H}}\right]$$

Claim:

$$\mathbf{M} \propto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Symmetric-Subspace Transform

$$\mathbf{M} = \begin{pmatrix} \mathsf{E}\left[\psi_0 \cdot \psi_0 \cdot \overline{\psi_0 \cdot \psi_0}\right] & \mathsf{E}\left[\psi_0 \cdot \psi_0 \cdot \overline{\psi_0 \cdot \psi_1}\right] & \mathsf{E}\left[\psi_0 \cdot \psi_0 \cdot \overline{\psi_1 \cdot \psi_0}\right] & \mathsf{E}\left[\psi_0 \cdot \psi_0 \cdot \overline{\psi_1 \cdot \psi_1}\right] \\ \mathsf{E}\left[\psi_0 \cdot \psi_0 \cdot \overline{\psi_0 \cdot \psi_0}\right] & \mathsf{E}\left[\psi_0 \cdot \psi_1 \cdot \overline{\psi_0 \cdot \psi_1}\right] & \mathsf{E}\left[\psi_0 \cdot \psi_1 \cdot \overline{\psi_1 \cdot \psi_0}\right] & \mathsf{E}\left[\psi_0 \cdot \psi_1 \cdot \overline{\psi_1 \cdot \psi_1}\right] \\ \mathsf{E}\left[\psi_1 \cdot \psi_0 \cdot \overline{\psi_0 \cdot \psi_0}\right] & \mathsf{E}\left[\psi_1 \cdot \psi_0 \cdot \overline{\psi_0 \cdot \psi_1}\right] & \mathsf{E}\left[\psi_1 \cdot \psi_0 \cdot \overline{\psi_1 \cdot \psi_0}\right] & \mathsf{E}\left[\psi_1 \cdot \psi_0 \cdot \overline{\psi_1 \cdot \psi_1}\right] \\ \mathsf{E}\left[\psi_1 \cdot \psi_1 \cdot \overline{\psi_0 \cdot \psi_0}\right] & \mathsf{E}\left[\psi_1 \cdot \psi_1 \cdot \overline{\psi_0 \cdot \psi_1}\right] & \mathsf{E}\left[\psi_1 \cdot \psi_1 \cdot \overline{\psi_1 \cdot \psi_0}\right] & \mathsf{E}\left[\psi_1 \cdot \psi_1 \cdot \overline{\psi_1 \cdot \psi_1}\right] \end{pmatrix}$$

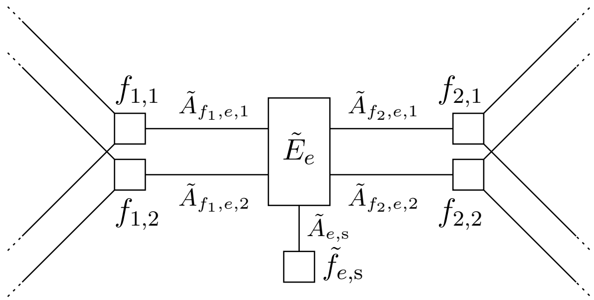
Symmetric-Subspace Transform

$$\mathbf{M} = \begin{pmatrix} \mathsf{E}\left[|\psi_0|^4\right] & \mathsf{E}\left[|\psi_0|^2 \cdot \psi_0 \cdot \overline{\psi_1}\right] & \mathsf{E}\left[|\psi_0|^2 \cdot \psi_0 \cdot \overline{\psi_1}\right] & \mathsf{E}\left[\psi_0^2 \cdot \overline{\psi_1}^2\right] \\ \mathsf{E}\left[|\psi_0|^2 \cdot \overline{\psi_0} \cdot \psi_1\right] & \mathsf{E}\left[|\psi_0|^2 \cdot |\psi_1|^2\right] & \mathsf{E}\left[|\psi_0|^2 \cdot |\psi_1|^2\right] & \mathsf{E}\left[\psi_0 \cdot |\psi_1|^2 \cdot \overline{\psi_1}\right] \\ \mathsf{E}\left[|\psi_0|^2 \cdot \overline{\psi_0} \cdot \psi_1\right] & \mathsf{E}\left[|\psi_0|^2 \cdot |\psi_1|^2\right] & \mathsf{E}\left[|\psi_0|^2 \cdot |\psi_1|^2\right] & \mathsf{E}\left[\psi_0 \cdot |\psi_1|^2 \cdot \overline{\psi_1}\right] \\ \mathsf{E}\left[\overline{\psi_0}^2 \cdot \psi_1^2\right] & \mathsf{E}\left[\overline{\psi_0} \cdot |\psi_1|^2 \cdot \psi_1\right] & \mathsf{E}\left[\overline{\psi_0} \cdot |\psi_1|^2 \cdot \psi_1\right] & \mathsf{E}\left[|\psi_1|^4\right] \end{pmatrix}$$

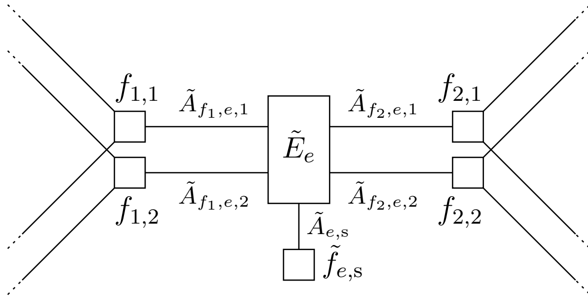
Symmetric-Subspace Transform

$$\mathbf{M} \propto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

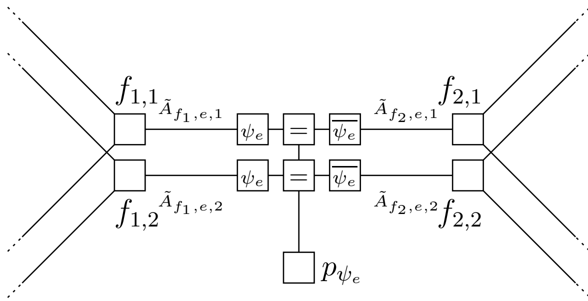
Symmetric-Subspace Transform



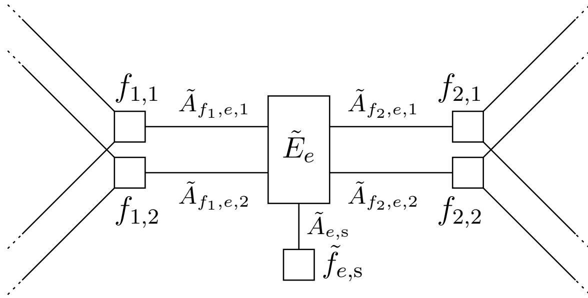
Symmetric-Subspace Transform



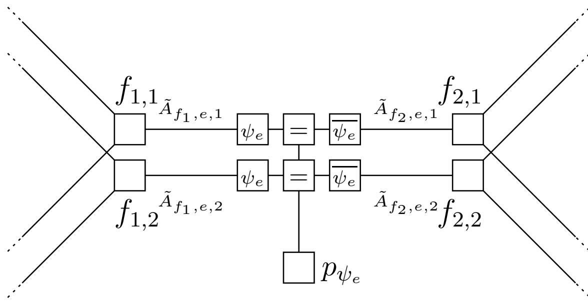
Using the above observation (omitting some proportionality constant):



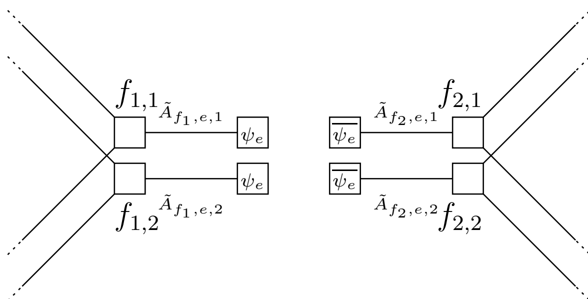
Symmetric-Subspace Transform



Using the above observation (omitting some proportionality constant):



After conditioning on ψ (omitting some proportionality constant):



Symmetric-Subspace Transform

The above considerations show that

$$Z_{\mathrm{Bethe},M}(\mathsf{N}) \;=\; \sqrt[M]{\alpha_{\mathsf{N},M} \cdot \int \mathrm{Re}\left(\left(g_{\mathrm{SST}}(\boldsymbol{\psi})\right)^{M}\right) \mathrm{d}\mu_{\mathrm{FS}}(\boldsymbol{\psi})}\,.$$

Symmetric-Subspace Transform

The above considerations show that

$$Z_{\mathrm{Bethe},M}(\mathsf{N}) \;=\; \sqrt[M]{\alpha_{\mathsf{N},M} \cdot \int \mathrm{Re}\Big(\big(g_{\mathrm{SST}}(\boldsymbol{\psi}) \big)^{M} \Big) \, \mathrm{d}\mu_{\mathrm{FS}}(\boldsymbol{\psi})} \, .$$

For DE-NFGs satisfying an (easily checkable) condition, we can use the Laplace method to analyze the above expression and show that

$$\limsup_{M \to \infty} Z_{\mathrm{Bethe},M}(\mathsf{N}) \;=\; Z_{\mathrm{Bethe}}(\mathsf{N}) \, .$$

Symmetric-Subspace Transform

The above considerations show that

$$Z_{\mathrm{Bethe},M}(\mathsf{N}) = \sqrt[M]{\alpha_{\mathsf{N},M} \cdot \int \mathrm{Re}\left(\left(g_{\mathrm{SST}}(\boldsymbol{\psi}) \right)^{M} \right) \mathrm{d}\mu_{\mathrm{FS}}(\boldsymbol{\psi})} \,.$$

For DE-NFGs satisfying an (easily checkable) condition, we can use the Laplace method to analyze the above expression and show that

$$\limsup_{M \to \infty} Z_{\mathrm{Bethe},M}(\mathsf{N}) = Z_{\mathrm{Bethe}}(\mathsf{N}) \,.$$

Actually, in order to obtain the above result, we also need to apply the so-called loop-calculus transform by Chertkov and Charnyak before applying the SST.

Conclusions / Outlook

Conclusions / Outlook

- Standard normal factor graphs (S-NFG):
  - Basics
  - A combinatorial interpretation of the Bethe partition sum, i.e., the Bethe approximation of the partition sum **[known]**
- Double-edge normal factor graphs (DE-NFG):
  - Basics
  - A combinatorial interpretation of the Bethe partition sum, i.e., the Bethe approximation of the partition sum **[novel]**

Y. Huang and P. O. Vontobel", "Characterizing the Bethe partition function of double-edge factor graphs via graph covers," ISIT 2020. [Longer version in preparation.]


Thank you!