The Bethe Fartition Function and the SPA for Factor Graphs based on Homogeneous Real Stable Polynomials

Yuwen Huang and Pascal O. Vontobel

Department of Information Engineering
The Chinese University of Hong Kong
yuwen.huang@ieee.org, pascal.vontobel@ieee.org

ISIT 2024

Main results

Consider a **standard factor graph (S-FG)** N where **each local function** is defined based on a (possibly different) **multi-affine homogeneous real stable** (MAHRS) polynomial.

Then we prove that

1. The **projection** of the **local marginal polytope** (LMP) on the **edges** in N equals the **convex hull** of the set of **valid configurations** $\mathrm{conv}(\mathcal{C})$.

2. For the **typical** case where the S-FG has a **sum-product algorithm (SPA) fixed point** consisting of **positive-valued messages only**, the SPA finds the value of the **Bethe partition function** $Z_{\mathrm{B}}(\mathsf{N})$ **exponentially fast**.

3. The **Bethe free energy function** F_{B} has some **convexity properties**.

Outline

An introductory example

A setup based on binary matrices with prescribed row sums and column sums

Graphical-model-based approximation method

Main results

A more general setup

Main results for a more general setup

Numerical results

Future works and connection to other works

Outline

- **An introductory example**

  A setup based on binary matrices with prescribed row sums and column sums

  Graphical-model-based approximation method

  Main results

  A more general setup

  Main results for a more general setup

  Numerical results

  Future works and connections to other works

An introductory example

Consider the set of all **binary 3×3 matrices**.

We want to know the number of **binary** 3×3 matrices with **row sums** and **column sums** equaling **two**.

The following are **example binary** 3×3 matrices:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

An introductory example

Consider the set of all **binary** 3×3 **matrices**.

We want to know the number of **binary** 3×3 matrices with **row sums** and **column sums** equaling **two**.

The following are **example binary** 3×3 matrices:

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}}_{\times}, \quad \underbrace{\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}}_{\checkmark}, \quad \underbrace{\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}}_{\checkmark}.$$

The number of such matrices is **3!**.

An introductory example

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

- These binary matrices can be viewed as **binary contingency tables** of size 3×3 with **row sums** and **column sums** equaling **two**.
- The number of such **binary contingency tables** is $3!$.

Outline

Setup

Definition

1. $[n] \triangleq \{1, 2, \ldots, n\}$ for $n \in \mathbb{Z}_{\geq 1}$ and $[m] \triangleq \{1, 2, \ldots, m\}$ for $m \in \mathbb{Z}_{\geq 1}$.

2. $\boldsymbol{x} = \big(x(i,j)\big)_{i\in[n], j\in[m]}$: a **$\{0,1\}$-valued matrix** of size $n \times m$.

3. For the ***i*-th row** $\boldsymbol{x}(i,:)$, we introduce an integer r_i and impose a **constraint** on the **row sum**:

$$\mathcal{X}_{r_i} = \left\{ \boldsymbol{x}(i,:) \,\middle|\, \sum_{j\in[m]} x(i,j) = r_i \right\}.$$

4. For the ***j*-th column** $\boldsymbol{x}(:,j)$, we introduce an integer c_j and impose a **constraint** on the **column sum**:

$$\mathcal{X}_{c_j} = \left\{ \boldsymbol{x}(:,j) \,\middle|\, \sum_{i\in[n]} x(i,j) = c_j \right\}.$$

Setup

Definition

5. The set of **valid configurations** is defined to be

$$\mathcal{C} \triangleq \left\{ \boldsymbol{x} \in \{0,1\}^{n\times n} \;\middle|\; \begin{array}{l} \boldsymbol{x}(i,:) \in \mathcal{X}_{r_i},\ \forall i \in [n], \\ \boldsymbol{x}(:,j) \in \mathcal{X}_{c_j},\ \forall j \in [m] \end{array} \right\},$$

the set of **binary matrices** such that the **i-th row sum** is r_i and the **j-th column sum** is c_j.

6. We want to compute the number of the **valid configurations** $|\mathcal{C}|$.

Outline

Graphical-model-based approximation method

Main idea

1. Define a **standard factor graph (S-FG)** N whose partition function equals

$$Z(\mathsf{N}) = |\mathcal{C}|.$$

2. Run the **sum product algorithm (SPA)**, a.k.a. **belief propagation** (BP), on the S-FG N to compute the **Bethe approximation of** $|\mathcal{C}|$, denoted by $Z_{\mathrm{B}}(\mathsf{N})$.

Graphical-model-based approximation method

Example

Consider $n = m = 3$ **and** $r_i = c_j = 2$, *i.e.*, $\boldsymbol{x} \in \{0,1\}^{3\times 3}$.

The i**-th row** $\boldsymbol{x}(i,:) \in \mathcal{X}_{r_i}$ and the j**-th column** $\boldsymbol{x}(:,j) \in \mathcal{X}_{c_j}$, where

$$\mathcal{X}_{r_i} = \{(1,1,0),(0,1,1),(1,0,1)\}, \quad \mathcal{X}_{c_j} = \{(1,1,0)^\mathsf{T},(0,1,1)^\mathsf{T},(1,0,1)^\mathsf{T}\}.$$

1. The **local functions**:

$$f_{\mathrm{l},i}\big(\boldsymbol{x}(i,:)\big) \triangleq \begin{cases} 1 & \text{if } \boldsymbol{x}(i,:) \in \mathcal{X}_{r_i} \\ 0 & \text{otherwise} \end{cases}, \quad f_{\mathrm{r},j}\big(\boldsymbol{x}(:,j)\big) \triangleq \begin{cases} 1 & \text{if } \boldsymbol{x}(:,j) \in \mathcal{X}_{c_j} \\ 0 & \text{otherwise} \end{cases}.$$

2. The **support** of the local functions:

$$\mathcal{X}_{f_{\mathrm{l},i}} \triangleq \left\{\boldsymbol{x}(i,:) \in \{0,1\}^3 \;\middle|\; f_{\mathrm{l},i}\big(\boldsymbol{x}(i,:)\big) > 0\right\} = \mathcal{X}_{r_i},$$
$$\mathcal{X}_{f_{\mathrm{r},j}} \triangleq \left\{\boldsymbol{x}(:,j) \in \{0,1\}^3 \;\middle|\; f_{\mathrm{r},j}\big(\boldsymbol{x}(:,j)\big) > 0\right\} = \mathcal{X}_{c_j}.$$

Graphical-model-based approximation method

3. The $\{0,1\}$-valued **global function**:

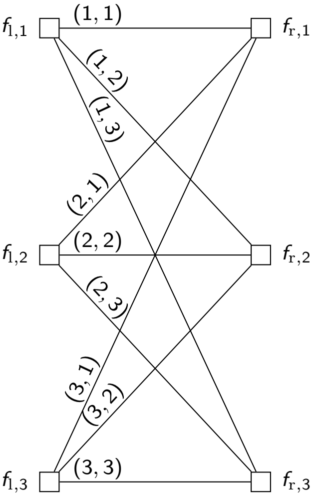
$$g(\boldsymbol{x}) \triangleq f_{\mathrm{l},1}\big(x(1,1), x(1,2), x(1,3)\big) \cdot f_{\mathrm{l},2}\big(x(2,1), x(2,2), x(2,3)\big) \cdots f_{\mathrm{r},2}\big(x(1,2), x(2,2), x(3,2)\big) \cdot f_{\mathrm{r},3}\big(x(1,3), x(2,3), x(3,3)\big).$$

The **previously defined** set of **valid configurations** is equal to the **support** of the global function:

$$\mathcal{C} = \left\{ \boldsymbol{x} \in \{0,1\}^{3\times 3} \;\middle|\; g(\boldsymbol{x}) > 0 \right\}.$$

4. The **partition function**:

$$Z(\mathsf{N}) \triangleq \sum_{\boldsymbol{x} \in \{0,1\}^{3\times 3}} g(\boldsymbol{x}) = |\mathcal{C}|.$$



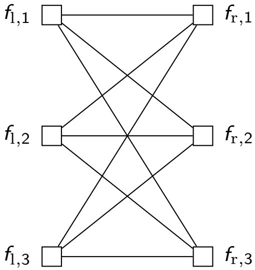
Graphical-model-based approximation method

5. The **Bethe approximation** of the partition function, *i.e.*, the **Bethe partition function**, is defined to be

$$Z_{\mathrm{B}}(\mathsf{N}) \triangleq \exp\left(-\min_{\boldsymbol{\beta}\in\mathcal{L}(\mathsf{N})} F_{\mathrm{B}}(\boldsymbol{\beta})\right),$$

where F_{B} is the **Bethe free energy (BFE)** function,

where $\mathcal{L}(\mathsf{N})$ is the **local marginal polytope** (LMP) (see, *e.g.*, [Wainwright and Jordan, 2008]).



6. Then we run the **sum-product algorithm (SPA)**, a.k.a. **belief propagation (BP)**, on the S-FG N to get $Z_{\mathrm{B}}(\mathsf{N})$.

Outline

An introductory example

A setup based on binary matrices with prescribed row sums and column sums

Graphical-model-based approximation method

- **Main results**

A more general setup

Main results for a more general setup

Numerical results

Future works and connection to other works

Main results

1. The **projection** of the **LMP** on the **edges** in N **equals** conv($\mathcal{C}$).
   (For general S-FGs, this projection is a **relaxation** of conv($\mathcal{C}$), *i.e.*, conv($\mathcal{C}$) is a **strict subset** of this **projection**.)

2. For the **typical case** where N has an **SPA fixed point** consisting of **positive-valued messages only**, the SPA finds the value of Z_{B}(N) **exponentially fast**.

3. The **BFE function** has some **convexity properties**.

Comments

- A **generalization** of parts of the results in [Vontobel, 2013].
- Even though the S-FG has a **non-trivial cyclic structure**, the SPA has **a good performance**.

Main results

Comments

For the setup where $n = m$, $r_i = 1$, and $c_j = 1$, it holds that

- $\mathcal{C} = \{\boldsymbol{x} \mid \boldsymbol{x}$ is a **permutation matrix** of size n-by-$n\}$
- The **projection** of the **LMP** on the **edges** equals the set of **doubly stochastic matrices** of size n-by-n.

Birkhoff–von Neumann theorem

The set of **doubly stochastic matrices** of size n-by-n is the **convex hull** of the set of the **permutation matrices** of size n-by-n.

The main result that $\text{conv}(\mathcal{C})$ equals the **projection** of the **LMP** on the **edges** for our considered S-FG, can be viewed as a **generalization**.

Outline

An introductory example

A setup based on binary matrices with prescribed row sums and column sums

Graphical-model-based approximation method

Main results

- **A more general setup**

Main results for a more general setup

Numerical results

Future works and connection to other works

A more general setup

An example S-FG

Consider $n = m = 3$ **and** $r_i = c_j = 2$. Then

$$f_{1,i}\big(\boldsymbol{x}(i,:)\big) = \begin{cases} 1 & \text{if } \boldsymbol{x}(i,:) \in \{(1,1,0),(0,1,1),(1,0,1)\} \\ 0 & \text{otherwise} \end{cases},$$

which corresponds to a **multi-affine homogeneous real stable (MAHRS) polynomial** w.r.t. the **indeterminates** in $\boldsymbol{L} \triangleq \big(L_1, L_2, L_3\big) \in \mathbb{C}^3$:

$$\begin{aligned} p_i(\boldsymbol{L}) &= \sum_{\boldsymbol{x}(i,:)\in\{0,1\}^3} f_{1,i}\big(\boldsymbol{x}(i,:)\big) \cdot \prod_{j\in[3]} \big(L_j\big)^{\boldsymbol{x}(i,j)} \\ &= L_1 \cdot L_2 + L_2 \cdot L_3 + L_1 \cdot L_3, \end{aligned}$$

Remark

- For details of **real stable polynomials**, see, *e.g.*, [Gharan, 2020]

Consider a **more general** setup where **each local function** is defined based on a **(possibly different) MAHRS polynomial**.

Do the previous results **hold** in this **more general setup**?

Yes!

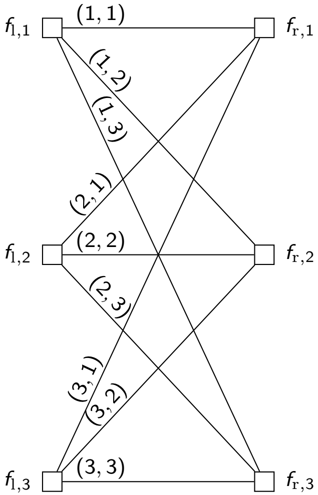
An MAHRS Polynomials-based S-FG

The standard factor graph (S-FG) N consists of

1. **edges**: $(1,1),(1,2),\ldots,(3,3)$;
2. **Binary** matrix

$$\boldsymbol{x} \triangleq \begin{pmatrix} x(1,1) & x(1,2) & x(1,3) \\ x(2,1) & x(2,2) & x(2,3) \\ x(3,1) & x(3,2) & x(3,3) \end{pmatrix}.$$

3. **Nonnegative-valued** local functions $f_{\mathrm{l},1},\ldots,f_{\mathrm{r},3}$;



An MAHRS Polynomials-based S-FG

6. The **local function** $f_{\mathrm{l},i}$ on the **LHS** is defined to be the mapping:

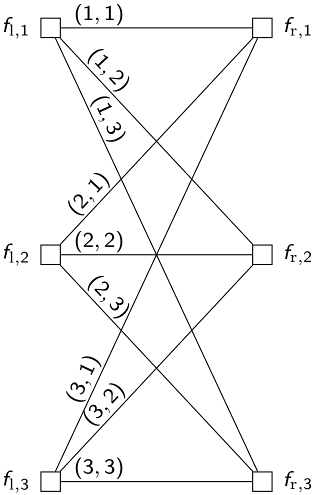
$$\{0,1\}^3 \to \mathbb{R}_{\geq 0}, \quad \boldsymbol{x}(i,:) \mapsto f_{\mathrm{l},i}\big(\boldsymbol{x}(i,:)\big)$$

such that it corresponds to an **MAHRS polynomial**.

7. The **support** of $f_{\mathrm{l},i}$:

$$\mathcal{X}_{f_{\mathrm{l},i}} \triangleq \left\{ \boldsymbol{x}(i,:) \in \{0,1\}^3 \;\middle|\; f_{\mathrm{l},i}\big(\boldsymbol{x}(i,:)\big) > 0 \right\}.$$

8. A **similar idea** in the definitions of $f_{\mathrm{r},j}$ and $\mathcal{X}_{f_{\mathrm{r},j}}$ on the **RHS**.



An MAHRS Polynomials-based S-FG

9. The **nonnegative-valued global function**:

$$g(\boldsymbol{x}) \triangleq f_{\mathrm{l},1}\big(\boldsymbol{x}(1,:)\big) \cdot f_{\mathrm{l},2}\big(\boldsymbol{x}(2,:)\big) \cdot f_{\mathrm{l},3}\big(\boldsymbol{x}(3,:)\big) \cdot f_{\mathrm{r},1}\big(\boldsymbol{x}(:,1)\big) \cdot f_{\mathrm{r},2}\big(\boldsymbol{x}(:,2)\big) \cdot f_{\mathrm{r},3}\big(\boldsymbol{x}(:,3)\big).$$

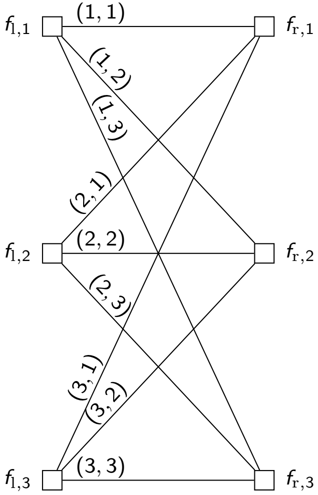
10. The set of **valid configurations**:

$$\mathcal{C} \triangleq \left\{\boldsymbol{x} \in \{0,1\}^{3\times 3} \;\middle|\; g(\boldsymbol{x}) > 0\right\},$$

which is also the **support** of the **global function**.

11. The **partition function**:

$$Z(\mathsf{N}) \triangleq \sum_{\boldsymbol{x}\in\mathcal{C}} g(\boldsymbol{x}).$$



Outline

Known results

Consider an **S-FG** N where **each local function** is defined based on a (possibly different) **MAHRS** polynomial.

Remarks

- Exactly computing $Z(\mathsf{N})$ is a **#P-complete problem** in general.
- **Run the SPA** to find the value of the **Bethe partition function** $Z_{\mathrm{B}}(\mathsf{N})$ that **approximates** $Z(\mathsf{N})$.
- [Straszak and Vishnoi, 2019, Theorem 3.2]: $Z_{\mathrm{B}}(\mathsf{N}) \leq Z(\mathsf{N})$.
- **Other real-stable-polynomial-based approximation of** $Z(\mathsf{N})$ [Gurvits, 2015, Brändén et al., 2023].

Main results

Consider an **S-FG** N where **each local function** is defined based on a (possibly different) **MAHRS** polynomial.

- The **support** $\mathcal{X}_{f_{\mathrm{l},i}}$ on the LHS corresponds to **a set of bases of a matroid** [Brändén, 2007].
- The support of the **product** of the **local functions** on the **LHS** is $\{\mathcal{X}_{f_{\mathrm{l},1}} \times \mathcal{X}_{f_{\mathrm{l},2}} \times \cdots \times \mathcal{X}_{f_{\mathrm{l},n}}\}$.
- Similarly for the local functions and the **support** on the **LHS**.
- The support of the **global function** equals the **intersection** of the bases of **matroids**:

$$\mathcal{C} = \{\mathcal{X}_{f_{\mathrm{l},1}} \times \mathcal{X}_{f_{\mathrm{l},2}} \times \cdots \times \mathcal{X}_{f_{\mathrm{l},n}}\} \bigcap \{\mathcal{X}_{f_{\mathrm{r},1}} \times \mathcal{X}_{f_{\mathrm{r},2}} \times \cdots \times \mathcal{X}_{f_{\mathrm{r},m}}\}$$

Main results

1. The **convex hull** $\mathrm{conv}(\mathcal{C})$ is the **projection of the LMP** on the **edges**.
   (Based on results on intersection of matroids [Oxley, 2011].)

2. For the typical case where the S-FG has an SPA **fixed point** consisting of **positive-valued messages only**, the SPA finds the value of $Z_{\mathrm{B}}(\mathsf{N})$ **exponentially fast**.
   (Based on the properties of **real stable polynomials** in [Brändén, 2007].)

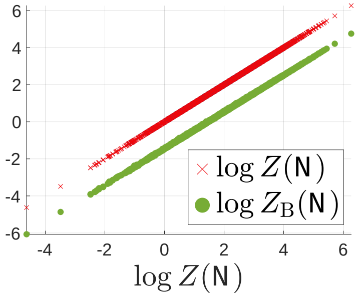
3. The **Bethe free energy function** F_{B} has some **convexity properties**.
   The proof of the convexity is **new**.
   (Based on the **dual** form of $Z_{\mathrm{B}}(\mathsf{N})$ in [Straszak and Vishnoi, 2019, Anari and Gharan, 2021].)

Outline

Numerical results

Setup

- We first consider the case $n = m = 6$ **and** $r_i = c_j = 2$.
- We independently randomly generate 3000 **instances** of N.



Observation

- $Z_{\mathrm{B}}(\mathrm{N}) \leq Z(\mathrm{N})$ ([Straszak and Vishnoi, 2019, Theroem 3.2]).
- $Z_{\mathrm{B}}(\mathrm{N})$ provides a **good estimate** of $Z(\mathrm{N})$ in this case.

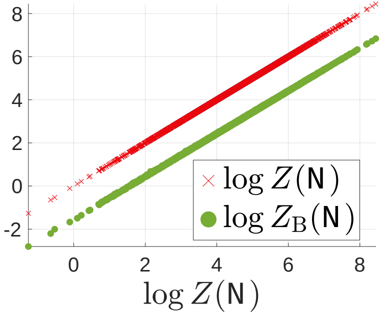
Numerical results

Setup

Consider **the same setup** as the previous case, but with $n = m = 6$ **replaced** by $n = m = 7$.

Observation

We can make **similar observations**.



Outline

An introductory example

A setup based on binary matrices with prescribed row sums and column sums

Graphical-model-based approximation method

Main results

A more general setup

Main results for a more general setup

Numerical results

- **Future works and connection to other works**

Future work

- Consider a **more general** S-FG, where each local function corresponds to a **more general** polynomial.
- Prove the **convergence** of the SPA for a **more general** S-FG.

Connection to other works

- **Polynomial approaches** to approximate **partition functions**.
  [Gurvits, 2011, Straszak and Vishnoi, 2017, Anari and Gharan, 2021]

- The properties of **real stable** polynomials and the **partition functions**.
  [Brändén, 2014, Borcea and Brändén, 2009, Borcea et al., 2009]

Selected References I

Anari, N. and Gharan, S. O. (2021).
A generalization of permanent inequalities and applications in counting and optimization.
Adv. Math., 383:107657.

Borcea, J. and Brändén, P. (2009).
The Lee-Yang and Pólya-Schur programs. II. Theory of stable polynomials and applications.
Commun. Pure Appl. Math., 62(12):1595–1631.

Borcea, J., Brändén, P., and Liggett, T. M. (2009).
Negative dependence and the geometry of polynomials.
J. Amer. Math. Soc., 22(2):521–567.

Selected References II

Brändén, P. (2007).
Polynomials with the half-plane property and matroid theory.
Adv. Math., 216(1):302–320.

Brändén, P. (2014).
The Lee-Yang and Pólya-Schur programs. I. Linear operators preserving stability.
Amer. J. Math., 136(1):241–253.

Brändén, P., Leake, J., and Pak, I. (2023).
Lower bounds for contingency tables via Lorentzian polynomials.
Israel J. Math., 253(1):43–90.

Selected References III

Gharan, S. O. (2020).

Course notes of polynomial paradigm in algorithm design.

Lecture 3.

Gurvits, L. (2011).

Unleashing the power of Schrijver's permanental inequality with the help of the Bethe approximation.

Elec. Coll. Comp. Compl.

Gurvits, L. (2015).

Boolean matrices with prescribed row/column sums and stable homogeneous polynomials: Combinatorial and algorithmic applications.

Inform. and Comput., 240:42–55.

Selected References IV

Oxley, J. (2011).
Matroid Theory.
Oxford University Press, New York, NY, USA, 2nd edition.

Straszak, D. and Vishnoi, N. K. (2017).
Real stable polynomials and matroids: optimization and counting.
In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2017, page 370–383, New York, NY, USA. Association for Computing Machinery.

Straszak, D. and Vishnoi, N. K. (2019).
Belief propagation, Bethe approximation, and polynomials.
IEEE Trans. Inf. Theory, 65(7):4353–4363.

Selected References V

Vontobel, P. O. (2013).
The Bethe permanent of a nonnegative matrix.
IEEE Trans. Inf. Theory, 59(3):1866–1901.

Wainwright, M. J. and Jordan, M. I. (2008).
Graphical models, exponential families, and variational inference.
Foundation and Trends in Machine Learning, 1(1–2):1–305.

Thank you!

Y. H. gratefully acknowledges useful discussions with Prof. Jonathan Leake.