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

Yuwen Huang and Pascal O. Vontobel
Department of Information Engineering
The Chinese University of Hong Kong
hy018@ie.cuhk.edu.hk, pascal.vontobel@ieee.org

**Abstract**

For standard factor graphs (S-FGs), i.e., factor graphs with local functions taking on non-negative real values, Vontobel gave a characterization of the Bethe approximation to the partition function in terms of the partition function of finite graph covers. The proof of that statement heavily relied on the method of types.

In this paper we give a similar characterization for so-called double-edge factor graphs (DE-FGs), which are a class of factor graphs where local functions take on complex values and have to satisfy some positive semi-definiteness constraints. Such factor graphs are of interest in quantum information processing.

In general, approximating the partition function of DE-FGs is more challenging than for S-FGs because the partition function is a sum of complex values and not just a sum of non-negative real values. In particular, for proving the above-mentioned characterization of the Bethe approximation in terms of finite graph covers, one cannot use the method of types anymore. We overcome this challenge by applying the loop calculus transform by Chertkov and Chernyak, along with using the symmetric-subspace transform, a novel technique for factor graphs that should be of interest beyond proving the main result of this paper. Currently, the characterization of the Bethe approximation of the partition function of DE-FGs is for DE-FGs satisfying an (easily checkable) condition. However, based on numerical results, we suspect that the characterization holds more broadly.

## I. Introduction

Standard factor graphs (S-FGs) [1]–[3], i.e., factor graphs with local functions taking on non-negative real values, are used for representing statistical models. They are widely applied in different areas including statistical mechanics (see, e.g., [4]), coding theory (see, e.g., [5]), and communications (see, e.g., [6]). Many important inference problems in these areas can be transformed into computing the marginals and the partition function of S-FGs. In order to approximate these quantities efficiently, a popular choice is the sum-product algorithm (SPA), also known as loopy belief propagation (LBP), which is a heuristic algorithm and has been successfully used for approximating the marginal functions and the partition function for many classes of S-FGs.

There are various characterizations of the pseudo-marginal functions obtained via a fixed point of the SPA. When an S-FG is cycle free, the pseudo-marginal functions at the fixed-point of the SPA are the exact marginal functions of the global function represented by the associated S-FG [4]. In terms of S-FGs with cycles, the authors in [7] showed that the fixed points of the SPA correspond to the stationary points of the constrained Bethe free energy function. Because of this connection of the SPA to the Bethe free energy function, the approximation of the partition function obtained with the help of the SPA is often called the Bethe approximation of the partition function, or simply the Bethe partition function. In [8], [9], Chertkov and Chernyak presented a technique they called loop calculus that yields an expression relating the partition function and the Bethe partition function. Notable extensions to these results were presented by Mori [10].

Finite graph covers of S-FGs are a useful theoretical tool [11]–[15] for understanding the Bethe partition function and the SPA. A combinatorial characterization of the Bethe partition function of an S-FG in terms of the partition function of finite graph covers of the S-FG was given in [12]. Leveraging this result, Ruozzi *et al.* [13], [14] proved statements relating the partition function and the Bethe partition function for certain classes of S-FGs, in particular solving a conjecture by Sudderth *et al.* about log-supermodular graphical models [16]. Moreover, graph covers were used in [15] to characterize the behavior of the max-product algorithm for Gaussian graphical models. (For Gaussian graphical models, the max-product algorithm is essentially equal to the sum-product algorithm.)

All the above results are for factor graphs that here are called S-FGs, i.e., for factor graphs with local functions taking on non-negative real values. Recently, some papers have considered more general factor graphs, namely factor graphs where the local functions take on complex values, in particular toward representing quantities of interest in quantum information processing [17]–[19]. (For connections of these factor graphs to other graphical notations in physics, see [18, Appendix A].) The structure of these factor graphs is not completely arbitrary, and so, toward formalizing such factor graphs, Cao and Vontobel [20] introduced double-edge factor graphs (DE-FGs), where local functions take on complex values and have to satisfy some positive semi-definiteness constraints.

DE-FGs have the property [20] that the partition function is a non-negative real number. However, because the partition function is the sum of complex numbers (notably, the real and the imaginary parts of these complex numbers can have positive or negative sign), approximating the partition function is in general much more challenging for DE-FGs than for S-FGs. This problem is known as the numerical sign problem in applied mathematics and theoretical physics (see, e.g., [21]).

It is straightforward to define the SPA for DE-FGs, along with defining the Bethe partition function based on fixed points of the SPA [20, Def. 8].[1] One can prove that the Bethe partition function is a non-negative real number. Given this, one wonders if there is a characterization of the Bethe partition function of a DE-FG in terms of the partition functions of its graph covers analogous to the result for S-FGs in [12]. Numerical calculations that we performed for small DE-FGs suggested that such a result should also hold for DE-FGs, however, proving such a result turned out to be much more challenging than for S-FGs. The reason for this is that the method of types, which was the main ingredient for the proof in the case of S-FGs, works essentially only for approximating sums of non-negative real values. Therefore, a novel proof approach needed to be developed for the case of DE-FGs; this novel proof approach is the main result of this paper besides the statement itself. (Currently, the characterization of the Bethe partition function of DE-FGs in terms of its finite graph covers is for DE-FGs satisfying an (easily checkable) condition. However, based on numerical results, we suspect that the characterization holds more broadly.)

Our proof approach is based on the following three main ingredients:

- In a first step, we apply the loop calculus transform[2] to the DE-FG based on a stable fixed point of the SPA. (In case the SPA has multiple stable fixed points, we assume that it is the stable fixed point with the largest Bethe partition function value.) The benefit of the loop calculus transform is that the transformed SPA fixed point messages have a very simple form independent of the edge and the direction of the message.
- In a second step, we apply the symmetric-subspace transform, a novel technique for factor graphs that should be of interest beyond the proof of the main result of this paper. Namely, for any positive integer $M$, the symmetric-subspace transform allows one to express the average partition function of $M$-covers of the DE-NFG in terms of some integral.
- In a third step, we let $M \to \infty$ and evaluate the integral with the help of Laplace's method. In order to apply Laplace's method, we study the Hessian matrix of the relevant function by relating it to the Jacobian matrix of the SPA. The negative

[1]Note that in [20] the Bethe partition function is formulated based on the so-called pseudo-dual Bethe free energy function for S-FGs [20, Def. 8]). Although generalizing the primal Bethe free energy function from S-FGs to DE-FGs is formally straightforward, it poses challenges because of the multi-valuedness of the complex logarithm, and is left for future research.

[2]The loop calculus transform can be seen as a particular instance of a holographic transform [22].

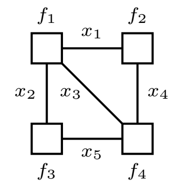



Fig. 1: NFG $\mathsf{N}$ in Example 1

semi-definiteness of the Hessian matrix at the point of interest follows from the stability of the SPA fixed point.[3]

We leave it as an open problem to broadly characterize when the relevant function has a global maximum at the point of interest. We give an (easily checkable) condition when this is the case. However, we suspect that it holds more generally.

In the following, we will use, without essential loss of generality, normal factor graphs (NFGs), i.e., factor graphs where variables are associated with edges [2], [3].

The rest of this paper is structured as follows. Section II reviews the basics of S-NFGs and the Bethe partition function. Section III discusses finite graph covers and the main result of [12], i.e., the combinatorial characterization of the Bethe partition function of an S-NFG in terms of the average partition functions of its graph covers. Section IV reviews DE-NFGs. Section V defines the loop calculus transform. Section VI introduces the symmetric-subspace transform. Finally, Section VII combines the techniques from the previous sections[4].

## II. Standard NFGs (S-NFGs)

In this section, we review some basic concepts and properties of an S-NFG, along with the Bethe approximation of the partition function of the S-NFG. We use an example to introduce the key concepts of an S-NFG.

**Example 1.** *Consider the multivariate function* $g(x_1,\ldots,x_5) := f_1(x_1,x_2,x_3)\cdot f_2(x_1,x_4)\cdot f_3(x_2,x_5)\cdot f_4(x_3,x_4,x_5)$*, where* $g$*, the so-called global function, is defined to be the product of the so-called local functions* $f_1$*,* $f_2$*,* $f_3$ *and* $f_4$*. We can visualize the factorization of* $g$ *with the help of the S-NFG* $\mathsf{N}$ *in Fig. 1. Note that the S-NFG* $\mathsf{N}$ *shown in Fig. 1 consists of four function nodes* $f_1,\ldots,f_4$ *and five (full) edges with associated variables* $x_1,\ldots,x_5$*.*

For simplicity, we consider only S-NFGs with full edges, because S-NFGs with half edges can be turned into S-NFGs with only full edges by adding dummy 1-valued function nodes to half edges without changing any marginals or the partition function.

**Definition 2.** *The S-NFG* $\mathsf{N}(\mathcal{F},\mathcal{E}_{\mathrm{full}},\mathcal{X})$ *consists of:*

1) *The graph* $(\mathcal{F},\mathcal{E}_{\mathrm{full}})$ *with vertex set* $\mathcal{F}$ *and edge set* $\mathcal{E}_{\mathrm{full}}$*, where* $\mathcal{E}_{\mathrm{full}}$ *consists of all full edges in* $\mathsf{N}$*. With some slight abuse of notation,* $f\in\mathcal{F}$ *will denote a function node and the corresponding local function.*
2) *The alphabet* $\mathcal{X} := \prod_{e\in\mathcal{E}_{\mathrm{full}}}\mathcal{X}_e$*, where* $\mathcal{X}_e$ *is the alphabet associated with edge* $e\in\mathcal{E}_{\mathrm{full}}$*.*

**Definition 3.** *For a given S-NFG* $\mathsf{N}(\mathcal{F},\mathcal{E}_{\mathrm{full}},\mathcal{X})$*, we define:*

1) *For every function node* $f\in\mathcal{F}$*, the set* $\partial f$ *is the set of edges incident on* $f$*.*
2) *The alphabet for each local function* $f\in\mathcal{F}$ *is defined to be* $\mathcal{X}_f := \prod_{e\in\partial f}\mathcal{X}_e$*.*

[3]At a high level, this last result is similar in spirit to some results by Heskes [23] and Watanabe [24]. However, in contrast to [24], we make a statement about all the eigenvalues of the Hessian matrix and not just the product of the eigenvalues, i.e., the determinant of the Hessian.

[4]Should there be any difference between the conference version and this version. This version would prevail.

Fig. 2: Left: NFG $\mathsf{N}$. Right: samples of possible 2-covers $\hat{\mathsf{N}}$ of $\mathsf{N}$.

3) *An assignment* $\boldsymbol{x} := (x_e)_{e\in\mathcal{E}_{\mathrm{full}}} \in \mathcal{X}$ *is called a configuration of the S-NFG. For each* $f \in \mathcal{F}$*, a configuration* $\boldsymbol{x} \in \mathcal{X}$ *induces the vector* $\boldsymbol{x}_{\partial f}$ *with components* $\boldsymbol{x}_{\partial f} := (x_e)_{e\in\partial f} \in \mathcal{X}_f$*.*
4) *The local function* $f$ *associated with function node* $f \in \mathcal{F}$ *denotes an arbitrary mapping* $f : \mathcal{X}_f \to \mathbb{R}_{\geq 0}$*.*
5) *The global function* $g$ *is defined to be the mapping* $g : \mathcal{X} \to \mathbb{R}_{\geq 0},\ \boldsymbol{x} \mapsto \prod_{f\in\mathcal{F}} f(\boldsymbol{x}_{\partial f})$*.*
6) *The partition function is* $Z(\mathsf{N}) := \sum_{\boldsymbol{x}\in\mathcal{X}} g(\boldsymbol{x})$*.*

If there is no ambiguity, in the following we will use the short-hands $\sum_{\boldsymbol{x}}$, $\sum_{\boldsymbol{x}_{\partial f}}$, $\sum_{x_e}$, $\sum_e$, $\sum_f$ for $\sum_{\boldsymbol{x}\in\mathcal{X}}$, $\sum_{\boldsymbol{x}_{\partial f}\in\mathcal{X}_f}$, $\sum_{x_e\in\mathcal{X}_e}$, $\sum_{e\in\mathcal{E}_{\mathrm{full}}}$, $\sum_{f\in\mathcal{F}}$, respectively.

The following definition of the Bethe partition function is based on the so-called pseudo-dual Bethe free energy function, not based on the (primal) Bethe free energy function.

**Definition 4.** *Consider the collection of messages* $\boldsymbol{\mu} := \{\boldsymbol{\mu}_{e\to f}\}_{f\in\partial e, e\in\mathcal{E}_{\mathrm{full}}}$*, with* $\boldsymbol{\mu}_{e\to f}$ *having entries* $\mu_{e\to f}(x_e) \in \mathbb{R}_{\geq 0}$*,* $x_e \in \mathcal{X}_e$*. Let* $Z_{\mathrm{B}}(\mathsf{N},\boldsymbol{\mu}) := \prod_f Z_f(\mathsf{N},\boldsymbol{\mu}) \big/ \prod_e Z_e(\mathsf{N},\boldsymbol{\mu})$*, where*

$$Z_f(\mathsf{N},\boldsymbol{\mu}) := \sum_{\boldsymbol{x}_{\partial f}} f(\boldsymbol{x}_{\partial f}) \prod_{e\in\partial f} \mu_{e\to f}(x_e),\ f \in \mathcal{F},$$

$$Z_e(\mathsf{N},\boldsymbol{\mu}) := \sum_{x_e} \mu_{e\to f}(x_e)\mu_{e\to f'}(x_e), \quad e = (f,f') \in \mathcal{E}_{\mathrm{full}}.$$

*The Bethe partition function* $Z_{\mathrm{B}}(\mathsf{N})$ *is defined to be maximum of* $Z_{\mathrm{B}}(\mathsf{N},\boldsymbol{\mu})$ *over all SPA fixed-point messages* $\boldsymbol{\mu}$ *of* $\mathsf{N}$*.*

## III. Finite Graph Covers

This section reviews the concept of finite graph covers of a graph in general, and of NFGs in particular [11], [12], [25].

**Definition 5.** *A graph* $\hat{\mathsf{G}} = (\hat{\mathcal{V}},\hat{\mathcal{E}})$ *with vertex set* $\hat{\mathcal{V}}$ *and edge set* $\hat{\mathcal{E}}$ *is the cover of a graph* $\mathsf{G} = (\mathcal{V},\mathcal{E})$ *if there exists a graph homomorphism* $\pi : \hat{\mathcal{V}} \to \mathcal{V}$ *such that for each* $v \in \mathcal{V}$ *and* $\hat{v} \in \hat{\mathcal{V}}$*, the set* $\partial\hat{v}$*, i.e., the neighborhood of* $\hat{v}$*, is mapped to* $\partial v$ *bijectively. Given a cover* $\hat{\mathsf{G}} = (\hat{\mathcal{V}},\hat{\mathcal{E}})$*, if* $|\pi^{-1}(v)| = M \in \mathbb{Z}_{>0}$ *for all* $v \in \mathcal{V}$*, where* $\pi^{-1}$ *is the pre-image of* $v$ *under mapping* $\pi$*, then* $\hat{\mathsf{G}}$ *is called an* $M$*-cover.*

Consequently, we can denote the vertex set of an $M$-cover as $\hat{\mathcal{V}} := \mathcal{V} \times [M]$. It means that for $(v,m) \in \hat{\mathcal{V}}$, $\pi((v,m)) = v$ and for edge $\big((v,m),(v',m')\big) \in \hat{\mathcal{E}}$, $\pi\big(((v,m),(v',m'))\big) = (v,v')$. Every $M$-cover, $\hat{\mathsf{G}} = (\hat{\mathcal{V}},\hat{\mathcal{E}})$ of a graph $\mathsf{G} = (\mathcal{V},\mathcal{E})$ can be obtained in the following way. For each edge $(v,v') \in \mathcal{E}$, select a permutation $\sigma_{v,v'} \in \mathcal{S}_M$, where $\mathcal{S}_M$ is the group of all permutation of $M$ elements, and then add the edge $\big((v,m),\ (v',\sigma_{v,v'}(m))\big)$ to $\hat{\mathsf{G}}$ for every $m \in [M]$.

**Example 6.** *An S-NFG* $\mathsf{N}$ *with* $4$ *vertices and* $5$ *edges is illustrated in Fig. 2(left). Possible* $2$*-covers of* $\mathsf{N}$ *are shown in Fig. 2(right). In particular, one can verify that every M-cover of* $\mathsf{N}$ *has* $M \cdot 4$ *function nodes and* $M \cdot 5$ *edges.*

In the following, we review the characterization of the Bethe partition function in terms of finite graph covers [12].

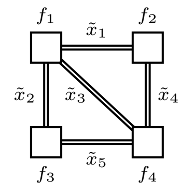



Fig. 3: DE-NFG. in Example 9.

**Definition 7.** *Consider some S-NFG* $\mathsf{N}$. *For any positive integer* $M$, *we define the degree-*$M$ *Bethe partition function to be* $Z_{\mathrm{B},M}(\mathsf{N}) := \sqrt[M]{\langle Z(\hat{\mathsf{N}})\rangle_{\hat{\mathsf{N}}\in\hat{\mathcal{N}}_M}}$, *where* $\langle Z(\hat{\mathsf{N}})\rangle_{\hat{\mathsf{N}}\in\hat{\mathcal{N}}_M}$ *represents the average of* $Z(\hat{\mathsf{N}})$ *over all* $\hat{\mathsf{N}} \in \mathcal{N}_M$, *where* $\mathcal{N}_M$ *is the set of all* $M$*-covers of* $\mathsf{N}$.

**Theorem 8** ([12])**.** *Consider some S-NFG* $\mathsf{N}$. *It holds that*[5]

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

## IV. Double-Edge NFGs (DE-NFGs)

In this section, we review DE-NFGs [20]. Similar to Section II, we only discuss the DE-NFGs where the edges are all full edges.

**Example 9.** *The DE-NFG in Fig. 3 depicts the following factorization*

$$g(\boldsymbol{x},\boldsymbol{x}') := f_1(x_1,x_2,x_3,x_1',x_2',x_3') \cdot f_2(x_1,x_4,x_1',x_4') \cdot f_3(x_2,x_5,x_2',x_5') \cdot f_4(x_3,x_4,x_5,x_3',x_4',x_5').$$

*Note that with a double edge* $e$ *we associate the variable* $\tilde{x}_e = (x_e, x_e')$, *where* $x_e$ *and* $x_e'$ *take value in the same alphabet* $\mathcal{X}_e$. *Moreover, if* $e$ *is incident on a function node* $f$, *then* $\tilde{x}_e$ *is an argument of the local function* $f$, *i.e., both* $x_e$ *and* $x_e'$ *are arguments of* $f$.

Note that DE-NFGs can also contain single edges (see [20] for details). However, without loss of generality, we can consider only DE-NFGs with double edges, because DE-NFGs with single edges can be turned into DE-NFGs with only double edges by changing single edges into double edges and suitably reformulating function nodes.

The definition of DE-NFGs is given as follows:

**Definition 10.** *A DE-NFG* $\mathsf{N}(\mathcal{F},\mathcal{E}_{\mathrm{full}},\tilde{\mathcal{X}})$ *consists of the following objects:*

1) *A graph* $(\mathcal{F},\mathcal{E}_{\mathrm{full}})$ *with vertex set* $\mathcal{F}$ *and edge set* $\mathcal{E}_{\mathrm{full}}$, *where* $\mathcal{F}$ *is also known as the set of function nodes.*
2) *An alphabet* $\tilde{\mathcal{X}} := \prod_{e\in\mathcal{E}_{\mathrm{full}}} \tilde{\mathcal{X}}_e$, *where* $\tilde{\mathcal{X}}_e := \mathcal{X}_e \times \mathcal{X}_e$ *is the alphabet of variables* $(x_e, x_e')$ *associated with edge* $e \in \mathcal{E}_{\mathrm{full}}$.
3) *A set of local functions* $\{f\}_{f\in\mathcal{F}}$. *Note that each* $f$ *corresponds to function node* $f \in \mathcal{F}$.

The details of a DE-NFG are listed as follows.

**Definition 11.** *Given a DE-NFG* $\mathsf{N}(\mathcal{F},\mathcal{E}_{\mathrm{full}},\tilde{\mathcal{X}})$, *we make the following definitions.*

1) *For each* $f \in \mathcal{F}$, *the set* $\partial f$ *is defined to be the set of double edges connected to* $f$. *Similarly, the set* $\partial e$ *is the set of function nodes that are connected to edge* $e$.

[5]Actually, in [12] the right-hand side was based on $Z_{\mathrm{B}}(\mathsf{N})$ computed via the minimum of the primal Bethe free energy function [7]. Therefore, the reformation here is for NFGs $\mathsf{N}$ for which the minimum of the primal Bethe free energy function can be characterized by an SPA fixed point, which is typically the case [7].

Fig. 4: Left: DE-NFG $\tilde{\mathsf{N}}$. Right: samples of possible 2-covers $\hat{\mathsf{N}}$ of $\mathsf{N}$.

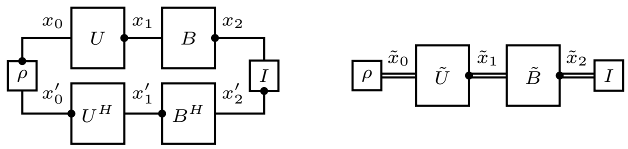



Fig. 5: Left: NFG in Example 12. Right: DE-NFG in Example 12.

2) *The alphabet for the set of variables* $((x_e, x_e'))_{e \in \partial f}$ *is* $\tilde{\mathcal{X}}_f := \prod_{e \in \partial f} \tilde{\mathcal{X}}_e$.
3) *For simplicity, we define* $\tilde{x}_e := (x_e, x_e') \in \tilde{\mathcal{X}}_e$ *to be a configuration for each edge* $e \in \mathcal{E}_{\text{full}}$. *For* $f \in \mathcal{F}$, *the associated configuration is* $\tilde{\boldsymbol{x}}_{\partial f} := (\tilde{x}_e)_{e \in \partial f} \in \tilde{\mathcal{X}}_f$. *A configuration of* $\mathsf{N}$ *is* $\tilde{\boldsymbol{x}} := (\tilde{x}_e)_{e \in \mathcal{E}_{\text{full}}} \in \tilde{\mathcal{X}}$. *Note that* $\tilde{\boldsymbol{x}}_{\partial f}$ *is induced by* $\tilde{\boldsymbol{x}}$.
4) *For each* $f \in \mathcal{F}$, $f$ *is defined to be an arbitrary mapping*
$$f : \tilde{\mathcal{X}}_f \to \mathbb{C}, \quad \tilde{\boldsymbol{x}}_{\partial f} \mapsto f(\tilde{\boldsymbol{x}}_{\partial f}).$$
*In particular, we require that for each local function* $f$, *the corresponding square matrix* $\big(f(\boldsymbol{x}_{\partial f}, \boldsymbol{x}'_{\partial f})\big)_{\boldsymbol{x}_{\partial f}, \boldsymbol{x}'_{\partial f}}$ *with row indices* $\boldsymbol{x}_{\partial f}$ *and column indices* $\boldsymbol{x}'_{\partial f}$, *which is known as Choi-matrix representation [26], is a complex-valued, Hermitian, positive semi-definite (PSD) matrix.*
5) *The global function* $g$ *represents a mapping*
$$g : \tilde{\mathcal{X}} \to \mathbb{C}, \quad \tilde{\boldsymbol{x}} \mapsto \prod_{f \in \mathcal{F}} f(\tilde{\boldsymbol{x}}_{\partial f}).$$
6) *The partition function is defined to be*
$$Z(\mathsf{N}) := \sum_{\tilde{\boldsymbol{x}} \in \tilde{\mathcal{X}}} g(\tilde{\boldsymbol{x}}).$$
*One can show [20] that* $Z(\mathsf{N}) \in \mathbb{R}_{\geq 0}$.
7) *It is straightforward to extend the SPA from S-NFGs to DE-NFGs. Note that the Choi-matrix representation of the SPA message* $\boldsymbol{\mu}_{f \to e}$ *can be represented by the square matrix* $\big(\mu_{f \to e}(x_e, x_e')\big)_{x_e, x_e'}$, *with row indices* $x_e$ *and column indices* $x_e'$. *In particular, if the initial messages represent complex-valued, Hermitian, PSD matrices, then all subsequent messages will also represent complex-valued, Hermitian, PSD matrices. (We refer to [20] for details.)*
8) *The expressions for* $Z_{\mathrm{B}}(\mathsf{N}, \boldsymbol{\mu})$ *and* $Z_{\mathrm{B}}(\mathsf{N})$ *in Definition 4 can be extended straightforwardly. However, now the messages* $\boldsymbol{\mu}_{e \to f}$ *represent complex-valued, Hermitian, PSD matrices.*
9) *It is straightforward to define finite graph covers for DE-NFGs. For example, Fig. 4(right) shows possible double covers of the DE-NFG in Fig. 4(left).*
10) *It is straightforward to extend the definition of* $Z_{\mathrm{B},M}(\mathsf{N})$ *in Definition 7 to DE-NFGs.*

One of the motivations for considering DE-NFGs is that many NFGs with complex-valued local functions in quantum information theory can be transformed into DE-NFGs.

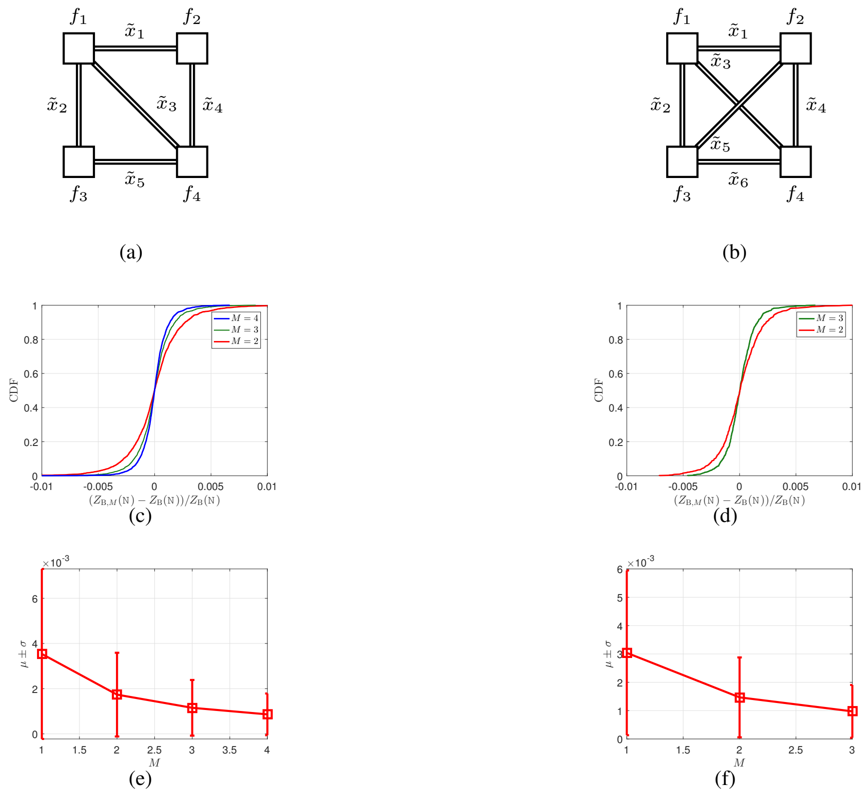



Fig. 6: DE-NFGs and the numerical results in Example 14.

**Example 12** ([17], [18])**.** *Let us consider an NFG in Fig. 5 (left). Note that at the end of each edge there is a dot used for denoting the first index of the matrix with respect to the function node. In particular, it describes a quantum system with two consecutive unitary evolutions:*

- *At the first stage, a quantum mechanical system is prepared in a mixed state represented by a PSD density matrix $\rho$.*
- *Then the system experiences two consecutive unitary evolutions, which are represented by the pair of factor nodes $U$ and $B$, respectively.*

*By suitably merging the edges, rearranging the order of the variables for each local function and defining new local functions, we can obtain the DE-NFG in Fig. 5 (right). For example, the entries of matrix $\tilde{U}$ are defined to be $\tilde{U}(x_0, x_1, x_0', x_1') := U(x_0, x_1) \cdot \overline{U(x_0', x_1')}$. One can verify that $\tilde{U}$ is a PSD matrix with row indices $(x_0, x_1)$ and column indices $(x_0', x_1')$.*

Based on the many connections between S-NFGs and DE-NFGs, and based on some numerical experiments, it is tempting to make the following conjecture.

**Conjecture 13.** *Consider some DE-NFG $\mathsf{N}$. It holds that*

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

Let us provide some promising numerical results for Conjecture 13.

## *A. Numerical Results for Conjecture 13*

**Example 14.** *Consider a DE-NFG in Fig. 6(a), where the alphabet size of each edge is fixed to be $4$, i.e., $|\mathring{\mathcal{X}}_{\mathrm{e}}| = 4$ and $\mathcal{X}_{\mathrm{e}} = \{0, 1\}$ for all $e \in \mathcal{E}_{\mathrm{full}}$. The number of realization is $3000$. In each realization, the local functions are randomly and independently generated.*

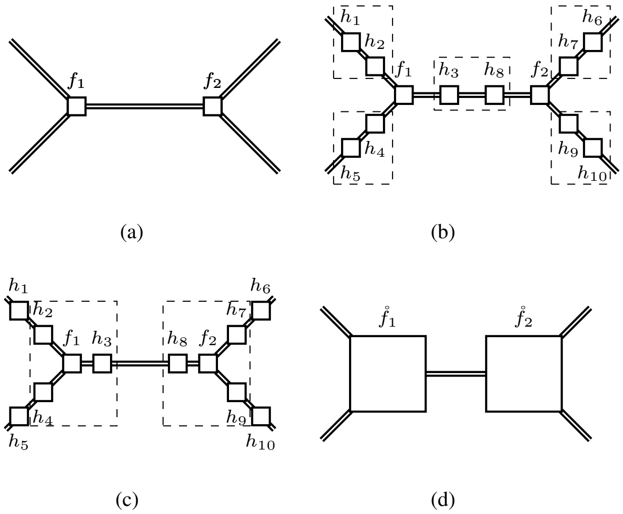



Fig. 7: Exemplifying the LCT for a part of an example DE-NFG.

*Fig. 6(c) shows the relationship between* $(Z_{\mathrm{B},M}(\mathsf{N}) - Z_{\mathrm{B}}(\mathsf{N}))/Z_{\mathrm{B}}(\mathsf{N})$ *and the degree of the covers* $M$ *in terms of the cumulative distribution functions, while Fig. 6(e) illustrates the metric* $\mu \pm \sigma$ *with respect to* $|Z_{\mathrm{B},M}(\mathsf{N}) - Z_{\mathrm{B}}(\mathsf{N})|/Z_{\mathrm{B}}(\mathsf{N})$*, where* $\mu$ *and* $\sigma$ *respectively represent the mean and the standard deviation of the corresponding numerical results. We can see that as* $M$ *increases, the degree-*$M$ *Bethe partition function* $Z_{\mathrm{B},M}(\mathsf{N})$ *approaches* $Z_{\mathrm{B}}(\mathsf{N})$*.*

**Example 15.** *We also investigate another DE-NFG shown in Fig. 6(b) with* $|\mathring{\mathcal{X}}_e| = 4$ *and* $\mathcal{X}_e = \{0, 1\}$ *for all* $e \in \mathcal{E}_{\mathrm{full}}$*. The total number of realizations is 3000. For every realization, all local functions are generated randomly and independently. The cumulative distribution functions of* $(Z_{\mathrm{B},M}(\mathsf{N}) - Z_{\mathrm{B}}(\mathsf{N}))/Z_{\mathrm{B}}(\mathsf{N})$ *for* $M = 2$ *and* $M = 3$*, are respectively plotted in Fig. 6(d). In addition, the impact of degree of covers* $M$ *on* $\mu \pm \sigma$ *is shown in Fig. 6(f), where* $\mu$ *and* $\sigma$ *are respectively the mean and the standard deviation of the metric* $|Z_{\mathrm{B},M}(\mathsf{N}) - Z_{\mathrm{B}}(\mathsf{N})|/Z_{\mathrm{B}}(\mathsf{N})$*. We notice that when the degree of covers* $M$ *increases, the gap between* $Z_{\mathrm{B}}(\mathsf{N})$ *and* $Z_{\mathrm{B},M}(\mathsf{N})$*is narrowed.*

We can prove Conjecture 13 for DE-NFGs satisfying the constraint mentioned at the end of Section VII. However, we suspect that Conjecture 13 holds for a much larger class of DE-NFGs.

## V. Loop-Calculus Transform (LCT)

Loop calculus as defined by Chertkov and Chernyak [8], [9] and further developed by Mori [10], can be formulated as a certain holographic transform [22] applied to a factor graph, and will henceforth be called loop-calculus transform (LCT). In the following, we give a high-level introduction to the LCT as applied to a DE-NFG with the help of Fig. 7.

- Consider some DE-NFG $\mathsf{N}$. (Fig. 7(a) shows parts of such a DE-NFG.)
- Consider some fixed-point messages $\boldsymbol{\mu}$ for this $\mathsf{N}$.
- The DE-NFG in Fig. 7(b) is obtained from the DE-NFG in Fig. 7(a) by applying suitable opening-the-box operations [3]. Importantly, the new function nodes $h_i$ are based on $\boldsymbol{\mu}$ and are such that the partition function is unchanged.
  For example, for the edge $e = (f_1, f_2)$ in Fig. 7(a), we introduce function nodes $h_3$ and $h_8$ with local functions $h_3 : \tilde{\mathcal{X}}_e \times \mathring{\mathcal{X}}_e \to \mathbb{C}$ and $h_8 : \tilde{\mathcal{X}}_e \times \mathring{\mathcal{X}}_e \to \mathbb{C}$, where $\mathring{\mathcal{X}}_e := \{0, 1, \ldots, |\mathring{\mathcal{X}}_e| - 1\}^2$. The partition function of the DE-NFG is unchanged because $h_3$ and $h_8$ satisfy $\sum_{\mathring{\tilde{x}}_e} h_3(\tilde{x}_{e,f_1}, \mathring{\tilde{x}}_e) \cdot h_8(\tilde{x}_{e,f_2}, \mathring{\tilde{x}}_e) = [\tilde{x}_{e,f_1} = \tilde{x}_{e,f_2}]$.
- The DE-NFG in Fig. 7(c) is essentially the same as the DE-NFG in Fig. 7(b).

- The DE-NFG in Fig. 7(d) is obtained from the DE-NFG in Fig. 7(c) by applying suitable closing-the-box operations [3] to the DE-NFG in Fig. 7(c). For example, the local function $\mathring{f}_1$ with set of adjacent edges $\partial f_1 = (e, e', e'')$ is obtained via $\mathring{f}_1(\tilde{\mathring{x}}_e, \tilde{\mathring{x}}_{e'}, \tilde{\mathring{x}}_{e''}) := \sum_{\tilde{x}_{e,f}, \tilde{x}_{e',f}, \tilde{x}_{e'',f}} f_1(\tilde{x}_{e,f}, \tilde{x}_{e',f}, \tilde{x}_{e'',f}) \cdot h_3(\tilde{x}_{e,f}, \tilde{\mathring{x}}_e) \cdot h_2(\tilde{x}_{e',f}, \tilde{\mathring{x}}_{e'}) \cdot h_4(\tilde{x}_{e'',f}, \tilde{\mathring{x}}_{e''})$. Again, the partition function of the DE-NFG is unchanged.

### A. Definition of LCT for DE-NFGs

In the following, we use variables $\tilde{\mathring{\boldsymbol{x}}}_{\partial f} := (\mathring{\boldsymbol{x}}_{\partial f}, \mathring{\boldsymbol{x}}'_{\partial f})$ for denoting the variable associated with function $\mathring{f}$ and $\tilde{\mathring{x}}_e := (\mathring{x}_e, \mathring{x}'_e)$ for denoting the variable on each edge $e \in \mathcal{E}_{\text{full}}$ for the DE-NFG after LCT and variables $\tilde{\boldsymbol{x}}_{\partial f} := (\boldsymbol{x}_{\partial f}, \boldsymbol{x}'_{\partial f})$ and $\tilde{x}_e$ for denoting the corresponding variables for the DE-NFG before LCT. By definition, we have $\tilde{\mathring{\boldsymbol{x}}}_{\partial f} = (\tilde{\mathring{x}}_e)_{e \in \partial f}$, $\mathring{x}_e$, $\mathring{x}'_e \in \{0, \ldots, |\mathcal{X}_e| - 1\}$ and $\mathring{\mathcal{X}}_e := \mathcal{X}_e \times \mathcal{X}_e$ for each $e \in \partial f$, $f \in \mathcal{F}$.

Given a solution of $Z_{\mathrm{B}}(\mathsf{N})$ for a DE-NFG $\mathsf{N}$, which is denoted as the set of messages $\boldsymbol{\mu}^* := (\boldsymbol{\mu}^*_{e \to f})_{e \in \partial f, f \in \mathcal{F}}$, we define two matrices $M_{e \to f} \in \mathbb{C}^{|\tilde{\mathcal{X}}_e| \times |\mathring{\mathcal{X}}_e|}$ and $M_{e \to f'} \in \mathbb{C}^{|\tilde{\mathcal{X}}_e| \times |\mathring{\mathcal{X}}_e|}$ for each $e = (f, f') \in \mathcal{E}_{\text{full}}$ with entries

$$
\begin{aligned}
M_{e \to f}(\tilde{x}_e, \tilde{\mathring{0}}) &:= c_{e,1} \mu^*_{e \to f}(\tilde{x}_e), \\
M_{e \to f}(\tilde{x}_e, \tilde{\mathring{x}}_e) &:= \frac{c_{e,1}}{c_{e,3}} \mu^*_{e \to f}(\tilde{x}_e) \Big( \big[ \tilde{x}_e = \tilde{\mathring{x}}_e \big] - b_e(\tilde{\mathring{x}}_e) \Big), \ \tilde{\mathring{x}}_e \in \mathring{\mathcal{X}}_e \backslash \{\tilde{\mathring{0}}\}, \\
M_{e \to f'}(\tilde{x}_e, \tilde{\mathring{0}}) &:= c_{e,2} \mu^*_{e \to f'}(\tilde{x}_e), \\
M_{e \to f'}(\tilde{x}_e, \tilde{\mathring{x}}_e) &:= c_{e,2} c_{e,3} Z_e(\mathsf{N}, \boldsymbol{\mu}^*) \phi_e(\tilde{x}_e, \tilde{\mathring{x}}_e), \ \tilde{\mathring{x}}_e \in \mathring{\mathcal{X}}_e \backslash \{\tilde{\mathring{0}}\},
\end{aligned}
$$

where $\tilde{\mathring{0}} := (0, 0)$, $c_{e,1} \in \mathbb{R}$ and $c_{e,2} \in \mathbb{R}$ are arbitrary constants satisfying $c_{e,1} c_{e,2} = Z_e(\mathsf{N}, \boldsymbol{\mu}^*)^{-1}$, $c_{e,3} \in \mathbb{R}$ is an arbitrary non-zero constant and

$$
\phi_e(\tilde{x}_e, \tilde{\mathring{x}}_e) := \begin{cases}
-\frac{1}{\mu^*_{e \to f}(\tilde{\mathring{0}})} & \tilde{\mathring{x}}_e = \tilde{\mathring{0}}, \ \mu^*_{e \to f}(\tilde{\mathring{0}}) \neq 0, \\
-1 & \tilde{\mathring{x}}_e = \tilde{\mathring{0}}, \ \mu^*_{e \to f}(\tilde{\mathring{0}}) = 0, \\
\frac{[\tilde{\mathring{x}}_e = \tilde{x}_e]}{\mu^*_{e \to f}(\tilde{\mathring{x}}_e)} & \tilde{\mathring{x}}_e \neq \tilde{\mathring{0}}, \ \mu^*_{e \to f}(\tilde{\mathring{x}}_e) \neq 0, \\
\big[ \tilde{\mathring{x}}_e = \tilde{x}_e \big] & \tilde{\mathring{x}}_e \neq \tilde{\mathring{0}}, \ \mu^*_{e \to f}(\tilde{\mathring{x}}_e) = 0.
\end{cases}
$$

One can verify that

$$
\begin{aligned}
\sum_{\tilde{x}_e} M_{e \to f}(\tilde{x}_e, \tilde{\mathring{x}}_{e,f}) M_{e \to f'}(\tilde{x}_e, \tilde{\mathring{x}}_{e,f'}) &= \Big[ \tilde{\mathring{x}}_{e,f} = \tilde{\mathring{x}}_{e,f'} \Big], \\
\sum_{\tilde{\mathring{x}}_e} M_{e \to f}(\tilde{x}_{e,f}, \tilde{\mathring{x}}_e) M_{e \to f'}(\tilde{x}_{e,f}, \tilde{\mathring{x}}_e) &= \big[ \tilde{x}_{e,f} = \tilde{x}_{e,f'} \big],
\end{aligned}
$$

for all $\tilde{\mathring{x}}_e$, $\tilde{\mathring{x}}_{e,f}$, $\tilde{\mathring{x}}_{e,f'} \in \mathring{\mathcal{X}}_e$ and $\tilde{x}_e$, $\tilde{x}_{e,f}$, $\tilde{x}_{e,f'} \in \tilde{\mathcal{X}}$.

**Theorem 16.** *Given $\boldsymbol{\mu}^* \in \mathrm{FI}(\mathsf{N})$ which is one of the solutions of $Z_{\mathrm{B}}(\mathsf{N})$, we define $\mathring{\mathsf{N}}$ to be the LCT of the DE-NFG $\mathsf{N}$ and $\mathring{f}$ be its local function for function node $f$. In particular,*

$$
\mathring{f}(\tilde{\mathring{\boldsymbol{x}}}_{\partial f}) := \sum_{\tilde{\boldsymbol{x}}_{\partial f}} f(\tilde{\boldsymbol{x}}_{\partial f}) \prod_{e \in \partial f} M_{e \to f}(\tilde{x}_e, \tilde{\mathring{x}}_e), \tag{1}
$$

*for any $\tilde{\mathring{\boldsymbol{x}}}_{\partial f} \in \mathring{\mathcal{X}}_f$ and $f \in \mathcal{F}$.*

*The DE-NFG $\mathring{\mathsf{N}}$ has the following properties:*

- *$\mathring{\mathsf{N}}$ has the same underlying graph topology as DE-NFG $\mathsf{N}$.*
- *The messages $\mathring{\boldsymbol{\mu}}_0 := (\mathring{\boldsymbol{\mu}}_{e \to f,0})_{e \in \partial f, f \in \mathcal{F}}$ with entries $\mathring{\mu}_{e \to f,0}(\tilde{\mathring{x}}_e) = 1$ if $\tilde{\mathring{x}}_e = \tilde{\mathring{0}}$ and $\mathring{\mu}_{e \to f,0}(\tilde{\mathring{x}}_e) = 0$ otherwise for all $e \in \partial f$, $f \in \mathcal{F}$ are fixed-point messages of the SPA for $\mathring{\mathsf{N}}$.*

- $Z_{\mathrm{B}}(\mathring{\mathsf{N}}, \mathring{\boldsymbol{\mu}}_0) = Z_{\mathrm{B}}(\mathsf{N})$.
- $Z(\mathring{\mathsf{N}}) = Z(\mathsf{N}) = Z_{\mathrm{B}}(\mathsf{N}) + \sum_{\mathring{\boldsymbol{x}} \in \mathring{\mathcal{X}} \setminus \{\mathring{\mathbf{0}}\}} \prod_{f \in \mathcal{F}} \mathring{f}(\mathring{\boldsymbol{x}}_{\partial f})$, *where* $\mathring{\mathcal{X}} := \prod_{e \in \mathcal{E}_{\mathrm{full}}} \mathring{\mathcal{X}}_e$, $\mathring{\mathbf{0}} := (\mathring{0})_{e \in \mathcal{E}_{\mathrm{full}}}$.
- $Z_{\mathrm{B}}\left(\mathring{\hat{\mathsf{N}}}, \bigotimes_{m=1}^{M} \mathring{\boldsymbol{\mu}}_0\right) = (Z_{\mathrm{B}}(\mathsf{N}))^M$ *for* $\mathring{\hat{\mathsf{N}}} \in \mathring{\mathcal{N}}_M$, *where* $\mathring{\mathcal{N}}_M$ *is the set of all degree-*$M$ *covers of* $\mathring{\mathsf{N}}$.

*Proof.* Omitted. ■

**Proposition 17.** *For each local function* $f \in \mathcal{F}$*, the corresponding Choi-matrix representation* $(\mathring{f}(\mathring{\boldsymbol{x}}_{\partial f}, \mathring{\boldsymbol{x}}'_{\partial f}))_{\mathring{\boldsymbol{x}}_{\partial f}, \mathring{\boldsymbol{x}}'_{\partial f}}$ *with row indices* $\mathring{\boldsymbol{x}}_{\partial f}$ *and column indices* $\mathring{\boldsymbol{x}}'_{\partial f}$ *is a Hermitian matrix.*

*Proof.* See Appendix A. ■

### B. Definition of LCT for NFGs

Similarly, we use variables $\mathring{\boldsymbol{x}}_{\partial f} \in \mathcal{X}_f$ for denoting the variable associated with function node $f$ and $\mathring{x}_e$ for denoting the variable on each edge $e \in \mathcal{E}_{\mathrm{full}}$ for the NFG after LCT and variables $\boldsymbol{x}_{\partial f} \in \mathcal{X}_f$ and $x_e$ for denoting the corresponding variables for the NFG before LCT. By definition, we have $\mathring{\boldsymbol{x}}_{\partial f} = (\mathring{x}_e)_{e \in \partial f}$ and $\mathring{x}_e \in \{0, \ldots, |\mathcal{X}_e| - 1\}$ for each $e \in \partial f$, $f \in \mathcal{F}$.

Given a solution of $Z_{\mathrm{B}}(\mathsf{N})$ for an NFG $\mathsf{N}$, which is denoted as the set of messages $\boldsymbol{\mu}^* := (\boldsymbol{\mu}^*_{e \to f})_{e \in \partial f, f \in \mathcal{F}}$, we define two matrices $M_{e \to f} \in \mathbb{C}^{|\mathcal{X}_e| \times |\mathcal{X}_e|}$ and $M_{e \to f'} \in \mathbb{C}^{|\mathcal{X}_e| \times |\mathcal{X}_e|}$ for each $e = (f, f') \in \mathcal{E}_{\mathrm{full}}$ with entries

$$
\begin{aligned}
M_{e \to f}(x_e, 0) &:= c_{e,1} \mu^*_{e \to f}(x_e), \\
M_{e \to f}(x_e, \mathring{x}_e) &:= \frac{c_{e,1}}{c_{e,3}} \mu^*_{e \to f}(x_e) \Big( \Big[ x_e = \mathring{x}_e \Big] - b_e(\mathring{x}_e) \Big), \ \mathring{x}_e \in \mathcal{X}_e \setminus \{0\}, \\
M_{e \to f'}(x_e, 0) &:= c_{e,2} \mu^*_{e \to f'}(x_e), \\
M_{e \to f'}(x_e, \mathring{x}_e) &:= c_{e,2} c_{e,3} Z_e(\mathsf{N}, \boldsymbol{\mu}^*) \phi_e(x_e, \mathring{x}_e), \ \mathring{x}_e \in \mathcal{X}_e \setminus \{0\},
\end{aligned}
$$

where $x_e \in \mathcal{X}_e$, $c_{e,1}$, $c_{e,2} \in \mathbb{R}$ are two arbitrary constants satisfying $c_{e,1} c_{e,2} = Z_e(\mathsf{N}, \boldsymbol{\mu}^*)^{-1}$, $c_{e,3} \in \mathbb{R}$ is an arbitrary non-zero constant and

$$
\phi_e(x_e, \mathring{x}_e) := \begin{cases}
-\frac{1}{\mu^*_{e \to f}(0)} & \mathring{x}_e = 0, \ \mu^*_{e \to f}(0) \neq 0, \\
-1 & \mathring{x}_e = 0, \ \mu^*_{e \to f}(0) = 0, \\
\frac{\left[ \mathring{x}_e = x_e \right]}{\mu^*_{e \to f}(\mathring{x}_e)} & \mathring{x}_e \neq 0, \ \mu^*_{e \to f}(\mathring{x}_e) \neq 0, \\
\Big[ \mathring{x}_e = x_e \Big] & \mathring{x}_e \neq 0, \ \mu^*_{e \to f}(\mathring{x}_e) = 0.
\end{cases}
$$

## VI. Symmetric-Subspace Transform (SST)

The symmetric-subspace transform (SST) can be used to analyze $M$-covers of S-NFGs or DE-NFGs. The developments in this section were motivated by [27], where, in terms of the language of the present paper, the authors transformed a certain integral into the average partition function of some NFGs, where the average is over double covers. (Although [27] needs to consider only double covers, it is clear that their results can be extended to general $M$-covers.) However, compared with [27], we go in the other direction, i.e., we express the average partition function of a DE-NFGs, where the average is over $M$-covers of some DE-NFG, in terms of some integral.

The developments in this section were partially also motivated by the results in [28].

For simplicity, in the following we consider only the case $M = 2$.

- Consider some DE-NFG $\mathsf{N}$. (Fig. 8(a) shows parts of this DE-NFG.).[6]

[6]Actually, for proving Conjecture 13, we consider a DE-NFG that is the LCT of some DE-NFG.

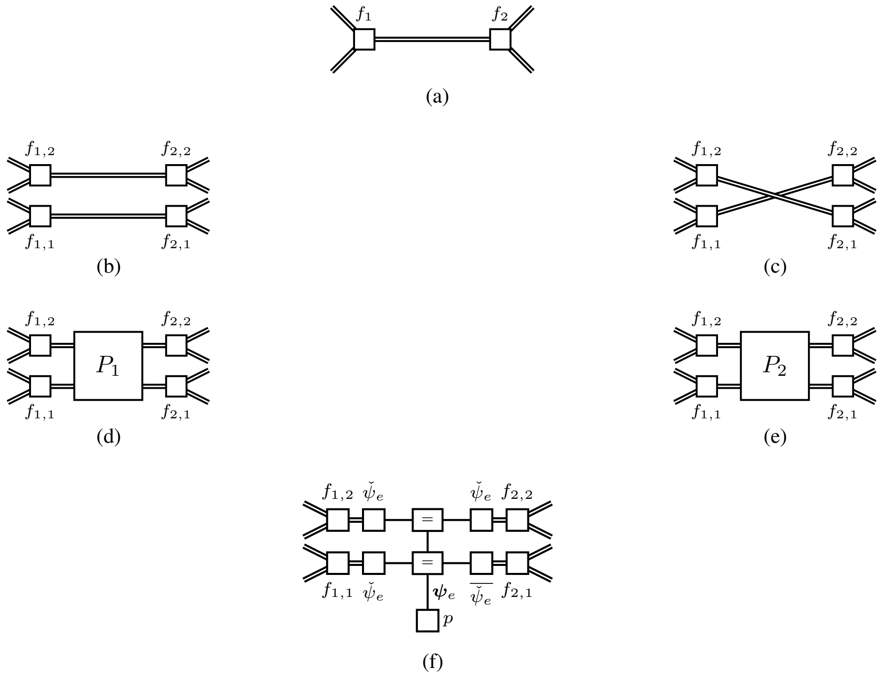



Fig. 8: Partial DE-NFGs illustrating the SST.

- Let $e = (f_1, f_2)$ be the edge connecting $f_1$ and $f_2$ in Fig. 8(a). When considering double covers of $\mathsf{N}$, the two copies of $e$ can connect the two copies of $f_1$ and $f_2$ either as shown in Fig. 8(b) or as shown in Fig. 8(c). In fact, in half the double covers the connections will be as in Fig. 8(b), and in half the double covers the connections will be as in Fig. 8(c).
- The connection pattern in Fig. 8(b) can be encoded with the help of a $\{0,1\}$-valued function node $P_1$, see Fig. 8(d). Similarly, the connection pattern in Fig. 8(c) can be encoded with the help of a $\{0,1\}$-valued function node $P_2$, see Fig. 8(e).
- We can represent the function values of $\frac{1}{2}P_1 + \frac{1}{2}P_2$ in terms of a square matrix. Crucially, this matrix can be obtained by evaluating the following integral (for $M = 2$)

$$(M+1) \int_{\boldsymbol{\psi}_e \in \mathbb{CP}^{d-1}} (\check{\boldsymbol{\psi}}_e \check{\boldsymbol{\psi}}_e^{\mathsf{H}})^{\otimes M} \, \mathrm{d}\mu_{\mathrm{FS}}(\boldsymbol{\psi}_e), \tag{2}$$

  where $\mathbb{CP}^{d-1}$ is the complex projective space with $d = |\tilde{\mathcal{X}}_e|^M$; $\check{\boldsymbol{\psi}}_e := \boldsymbol{\psi}_e / \|\boldsymbol{\psi}_e\|$; $\check{\boldsymbol{\psi}}_e^{\mathsf{H}}$ is the Hermitian transpose of $\check{\boldsymbol{\psi}}_e$; $(\check{\boldsymbol{\psi}}_e \check{\boldsymbol{\psi}}_e^{\mathsf{H}})^{\otimes M}$ is the $M$-fold tensor product of the matrix $\check{\boldsymbol{\psi}}_e \check{\boldsymbol{\psi}}_e^{\mathsf{H}}$; and $\mu_{\mathrm{FS}}$ denotes the Fubini-Study (FS) measure.

  The above integral can be expressed in terms of an NFG as shown in Fig. 8(f), where (2) is implemented by parameterized functions $\check{\boldsymbol{\psi}}_e$ that are parameterized by $\boldsymbol{\psi}_e$, which is distributed according to $p(\boldsymbol{\psi}_e)$.

### A. An Alternative Characterization of $\langle Z(\hat{\mathsf{N}}) \rangle_{\hat{\mathsf{N}} \in \hat{\mathcal{N}}_M}$ using SST

Here we give an alternative characterization of $\langle Z(\hat{\mathsf{N}}) \rangle_{\hat{\mathsf{N}} \in \hat{\mathcal{N}}_M}$.

**Lemma 18.** *Given an edge $e \in \mathcal{E}_{\mathrm{full}}$, we suppose that the entries in $\boldsymbol{w}_e = [w_e(0), \ldots, w_e(2|\mathring{\mathcal{X}}_e| - 1)]^T \in \mathbb{R}^{2|\mathring{\mathcal{X}}_e|}$ are independent and identically distributed (i.i.d.) real-valued random variables following standard normal distribution. For any $(K_i)_{i=0}^{|\mathring{\mathcal{X}}_e|-1} \in \mathbb{Z}_{\geq 0}^{|\mathring{\mathcal{X}}_e|}$, we have*

$$\mathbb{E}\left( \frac{\prod_{i=0}^{|\mathring{\mathcal{X}}_e|-1} \big( (w_e(i))^2 + (w_e(i+1))^2 \big)^{K_i}}{\big( \sum_{j=0}^{2|\mathring{\mathcal{X}}_e|-1} (w_e(j))^2 \big)^{\sum_{k=0}^{|\mathring{\mathcal{X}}_e|-1} K_k}} \right) = \frac{\prod_{i=0}^{|\mathring{\mathcal{X}}_e|-1} K_i!}{\big( \sum_{j=0}^{|\mathring{\mathcal{X}}_e|-1} K_j + 1 \big)!}.$$

*Proof.* See Appendix B. ∎

To apply SST, we define

$$\begin{aligned}
\boldsymbol{\psi} &:= (\boldsymbol{\psi}_e)_{e \in \mathcal{E}_{\mathrm{full}}}, \\
\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}} &:= (\mathring{\tilde{x}}_{e,f})_{e \in \partial f, f \in \mathcal{F}}, \\
\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}) &:= \prod_f \mathring{f}(\mathring{\tilde{\boldsymbol{x}}}_{\partial f,f}), \\
\mathring{g}_2(\boldsymbol{\psi}) &:= \sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}} \mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}) \prod_e \check{\psi}_e(\mathring{\tilde{x}}_{e,f}) \overline{\check{\psi}_e(\mathring{\tilde{x}}_{e,f'})}, \\
\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi}) &:= \big(\mathring{g}_2(\boldsymbol{\psi})\big)^M,
\end{aligned} \tag{3}$$

where $\mathring{\tilde{\boldsymbol{x}}}_{\partial f,f} \in \mathring{\mathcal{X}}_f$, $\mathring{\tilde{x}}_{e,f} \in \mathring{\mathcal{X}}_e$, $\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}} := \sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}} \in (\mathring{\mathcal{X}})^2}$,

$$\boldsymbol{\psi}_e := \Big[ w_e(0) + i w_e(1) \quad \ldots \quad w_e(2|\mathring{\mathcal{X}}_e| - 2) + i w_e(2|\mathring{\mathcal{X}}_e| - 1) \Big]^T \in \mathbb{C}^{|\mathring{\mathcal{X}}_e|},$$

with index $\mathring{\tilde{x}}_e$ and the entries in $\boldsymbol{w}_e = [w_e(0), \ldots, w_e(2|\mathring{\mathcal{X}}_e| - 1)]^T \in \mathbb{R}^{2|\mathring{\mathcal{X}}_e|}$ are i.i.d. real-valued random variables and follow standard normal distribution. By definition of LCT, we have $\mathring{g}(\mathring{\tilde{\mathbf{0}}}) = Z_{\mathrm{B}}(\mathsf{N}) \in \mathbb{R}_{\geq 0}$.

**Theorem 19.** *The function* $\big\langle Z(\hat{\mathsf{N}}) \big\rangle_{\hat{\mathsf{N}} \in \hat{\mathcal{N}}_M}$ *equals*

$$\big\langle Z(\hat{\mathsf{N}}) \big\rangle_{\hat{\mathsf{N}} \in \hat{\mathcal{N}}_M} = (M+1)^{|\mathcal{E}_{\mathrm{full}}|} \int \mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi}) \mathrm{d}\mu_{\mathrm{FS}}(\boldsymbol{\psi}).$$

*Proof.* See Appendix C. ■

## VII. Combining Everything

We could apply the SST to all the edges of the DE-NFG $\mathsf{N}$ under consideration. However, the SST turns out to be much more useful when applied to the LCT $\mathring{\mathsf{N}}$ of $\mathsf{N}$ based on some SPA-fixed point messages $\boldsymbol{\mu}$. This is because of the properties of $\mathring{\mathsf{N}}$ mentioned in Theorem 16. Namely, after applying the SST to all edges of $\mathring{\mathsf{N}}$, we obtain an NFG $\mathring{\mathsf{N}}_{\mathrm{SST}}$ with the following properties:

- On the one hand, the partition function satisfies

$$Z(\mathring{\mathsf{N}}_{\mathrm{SST}}) = \big(Z_{\mathrm{B},M}(\mathsf{N})\big)^M. \tag{4}$$

- On the other hand, the partition function satisfies

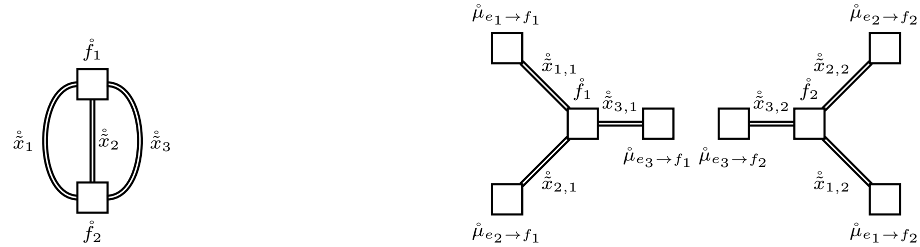
$$Z(\mathring{\mathsf{N}}_{\mathrm{SST}}) = (M+1)^{|\mathcal{E}_{\mathrm{full}}|} \int \mathrm{Re}\big(\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})\big) \, \mathrm{d}\mu(\boldsymbol{\psi}), \tag{5}$$

  where $\boldsymbol{\psi} := \{\boldsymbol{\psi}_e\}_{e \in \mathcal{E}_{\mathrm{full}}}$ and where the integral is over some suitable set and measure as implied by (2). The function $\mathring{g}_{\mathrm{SST}}$ appearing in this integral satisfies $\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi}_0) = \big(Z_{\mathrm{B}}(\mathsf{N}, \boldsymbol{\mu})\big)^M$, where $\boldsymbol{\psi}_0 := (\boldsymbol{\psi}_{e,0})_{e \in \mathcal{E}_{\mathrm{full}}}$ and $\psi_{e,0}(\mathring{\tilde{x}}_e) = 1$ if $\mathring{\tilde{x}}_e = \mathring{\tilde{0}}$ and $\psi_e(\mathring{\tilde{x}}_e) = 0$ otherwise for all $e \in \mathcal{E}_{\mathrm{full}}$.

**Theorem 20.** *Function* $\mathrm{Re}\big(\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})\big)$ *has a local maximum at* $\boldsymbol{\psi} = \boldsymbol{\psi}_0$

*Proof.* See Appendix D. ■

The proof of Conjecture 13 can be obtained by combining (4)–(5), taking the limit $M \to \infty$, and using Laplace's method. In order to show that $\mathrm{Re}\big(\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})\big)$ has a global maximum at $\boldsymbol{\psi} = \boldsymbol{\psi}_0$, one can for example limit the consideration to DE-NFGs $\mathsf{N}$ whose LCT $\mathring{\mathsf{N}}$ satisfies the following condition: for every function node $f$, it holds that $\mathring{f}(\mathring{\tilde{\mathbf{0}}}) > \sum_{\mathring{\tilde{\boldsymbol{x}}}_f \neq \mathring{\tilde{\mathbf{0}}}} |\mathring{f}(\mathring{\tilde{\boldsymbol{x}}}_f)|$. However, we suspect that (much) weaker conditions are required.




(a) (b)

Fig. 9: DE-NFGs in Example 21

## VIII. Comparison between $\mathring{g}_2(\boldsymbol{\psi})$ and the Bethe Partition Function

Recall that the function $Z_{\mathrm{B}}(\mathring{\mathsf{N}},\mathring{\boldsymbol{\mu}})$ for the DE-NFG after LCT equals

$$Z_{\mathrm{B}}(\mathring{\mathsf{N}},\mathring{\boldsymbol{\mu}}) = \frac{\prod_f Z_f(\mathring{\mathsf{N}},\mathring{\boldsymbol{\mu}})}{\prod_e Z_e(\mathring{\mathsf{N}},\mathring{\boldsymbol{\mu}})},$$

where $\mathring{\boldsymbol{\mu}} := \{\mathring{\boldsymbol{\mu}}_{e\to f}\}_{e\in\partial f, f\in\mathcal{F}}$ and $\mathring{\boldsymbol{\mu}}_{e\to f} \in \mathbb{C}^{|\mathring{\mathcal{X}}_e|}$.

**Example 21.** *Consider the DE-NFG $\mathring{\mathsf{N}}$ in Fig. 9(a). The corresponding function $Z_{\mathrm{B}}(\mathring{\mathsf{N}},\mathring{\boldsymbol{\mu}})$ is the partition function of the DE-NFG in Fig. 9(b), which equals to*

$$Z_{\mathrm{B}}(\mathring{\mathsf{N}},\mathring{\boldsymbol{\mu}}) = \frac{\left(\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\partial f_1}}, \mathring{f}_1(\mathring{\tilde{\boldsymbol{x}}}_{\partial f_1}) \prod_{e\in(1,2,3)} \mathring{\mu}_{e\to f_1}(\mathring{\tilde{x}}_{e,1})\right)\left(\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\partial f_2}}, \mathring{f}_2(\mathring{\tilde{\boldsymbol{x}}}_{\partial f_2}) \prod_{e\in(1,2,3)} \mathring{\mu}_{e\to f_1}(\mathring{\tilde{x}}_{e,2})\right)}{\prod_{e\in(1,2,3)}\left(\sum_{\mathring{\tilde{x}}_{e,f}} \mathring{\mu}_{e\to f_1}(\mathring{\tilde{x}}_{e,1})\mathring{\mu}_{e\to f_2}(\mathring{\tilde{x}}_{e,2})\right)},$$

*where $\mathring{\tilde{\boldsymbol{x}}}_{\partial f_1} := (\mathring{\tilde{x}}_{1,1}, \mathring{\tilde{x}}_{2,1}, \mathring{\tilde{x}}_{3,1})$ and $\mathring{\tilde{\boldsymbol{x}}}_{\partial f_2} := (\mathring{\tilde{x}}_{1,2}, \mathring{\tilde{x}}_{2,2}, \mathring{\tilde{x}}_{3,2})$.*

The Wirtinger derivatives of $Z_{\mathrm{B}}(\mathsf{N},\boldsymbol{\mu})$ with respect to $\mathring{\mu}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_1})$ and $\overline{\mathring{\mu}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_1})}$ for $e_1 = (f_1, f_1')$ equal

$$\begin{aligned}
\frac{\partial Z_{\mathrm{B}}(\mathsf{N},\mathring{\boldsymbol{\mu}})}{\partial \mathring{\mu}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_1})} &= \frac{\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}\in S((e_1,f_1),\mathring{\tilde{x}}_{e_1})} \mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\mathring{\mu}_{e_1\to f_1'}(\mathring{\tilde{x}}_{e,f'}) \prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}} \mathring{\mu}_{e\to f}(\mathring{\tilde{x}}_{e,f})\mathring{\mu}_{e\to f'}(\mathring{\tilde{x}}_{e,f'})}{\prod_{e\in\mathcal{E}_{\mathrm{full}}}\left(\sum_{\mathring{\tilde{x}}_{e,f}} \mathring{\mu}_{e\to f}(\mathring{\tilde{x}}_{e,f})\mathring{\mu}_{e\to f'}(\mathring{\tilde{x}}_{e,f})\right)} \\
&\quad - \frac{\mathring{\mu}_{e_1\to f_1'}(\mathring{\tilde{x}}_{e_1}) \sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}} \mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}) \prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}} \mathring{\mu}_{e\to f}(\mathring{\tilde{x}}_{e,f})\mathring{\mu}_{e\to f'}(\mathring{\tilde{x}}_{e,f'})}{\left(\sum_{\mathring{\tilde{x}}_{e_1,f_1}} \mathring{\mu}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_1,f_1})\mathring{\mu}_{e_1\to f_1'}(\mathring{\tilde{x}}_{e_1,f_1})\right)^2 \prod_{e\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}}\left(\sum_{\mathring{\tilde{x}}_{e,f}} \mathring{\mu}_{e\to f}(\mathring{\tilde{x}}_{e,f})\mathring{\mu}_{e\to f'}(\mathring{\tilde{x}}_{e,f})\right)}, \\
\frac{\partial Z_{\mathrm{B}}(\mathsf{N},\mathring{\boldsymbol{\mu}})}{\partial \overline{\mathring{\mu}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_1})}} &= 0,
\end{aligned}$$

where

$$S((e_1,f_1),\mathring{\tilde{x}}_{e_1}) := \{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}} \in (\mathring{\mathcal{X}})^2 :\ \mathring{\tilde{x}}_{e_1,f_1} = \mathring{\tilde{x}}_{e_1}\},$$

for all $\mathring{\tilde{x}}_{e_1,f_1} \in \mathring{\mathcal{X}}_{e_1}$, $e_1 = (f_1, f_1') \in \mathcal{E}_{\mathrm{full}}$. At the stationary point of $Z_{\mathrm{B}}(\mathsf{N},\mathring{\boldsymbol{\mu}})$ i.e., $\frac{\partial Z_{\mathrm{B}}(\mathsf{N},\mathring{\boldsymbol{\mu}})}{\partial \mathring{\mu}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_1})} = 0$, for every $e_1 \in \partial f$ and $f \in \mathcal{F}$, we have

$$\begin{aligned}
\mathring{\mu}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_1}) &\propto \sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}\in S((e_1,f_1'),\mathring{\tilde{x}}_{e_1})} \mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\mathring{\mu}_{e_1\to f_1}(\mathring{\tilde{x}}_{e,f'}) \prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}} \mathring{\mu}_{e\to f}(\mathring{\tilde{x}}_{e,f})\mathring{\mu}_{e\to f'}(\mathring{\tilde{x}}_{e,f'}), \\
\mathring{\mu}_{e_1\to f_1'}(\mathring{\tilde{x}}_{e_1}) &\propto \sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}\in S((e_1,f_1),\mathring{\tilde{x}}_{e_1})} \mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\mathring{\mu}_{e_1\to f_1'}(\mathring{\tilde{x}}_{e,f'}) \prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}} \mathring{\mu}_{e\to f}(\mathring{\tilde{x}}_{e,f})\mathring{\mu}_{e\to f'}(\mathring{\tilde{x}}_{e,f'}),
\end{aligned}$$

which is the expression for the fixed-point of the SPA.

**Proposition 22.** *Recall that $\mathring{g}_2(\boldsymbol{\psi})$ is defined in* (3)*. At the stationary point of $\mathring{g}_2(\boldsymbol{\psi})$, i.e., $\frac{\partial \mathring{g}_2(\boldsymbol{\psi})}{\partial \psi_{e_1}(\tilde{\mathring{x}}_{e_1})} = \frac{\partial \mathring{g}_2(\boldsymbol{\psi})}{\partial \overline{\psi_{e_1}(\tilde{\mathring{x}}_{e_1})}} = 0$, for every $e_1 \in \partial f$ and $f \in \mathcal{F}$, the vector $\boldsymbol{\psi}$ corresponds to the SPA fixed-point.*

*Proof.* From (16) and (17), at the stationary point of $\mathring{g}_2(\boldsymbol{\psi})$, we have

$$\psi_{e_1}(\tilde{\mathring{x}}_{e_1}) \propto \sum_{\tilde{\boldsymbol{x}}_{\mathcal{E}_{\text{full}},\mathcal{F}} \in S((e_1,f_1'),\tilde{\mathring{x}}_{e_1})} \mathring{g}(\tilde{\boldsymbol{x}}_{\mathcal{E},\mathcal{F}}) \psi_{e_1}(\tilde{\mathring{x}}_{e_1,f_1}) \prod_{e=(f,f') \in \mathcal{E}_{\text{full}} \setminus \{e_1\}} \psi_e(\tilde{\mathring{x}}_{e,f}) \overline{\psi_e(\tilde{\mathring{x}}_{e,f'})},$$

which is the expression for the fixed-point of SPA such that $\mathring{\boldsymbol{\mu}}_{e\to f} = \boldsymbol{\psi}_e$ and $\mathring{\boldsymbol{\mu}}_{e\to f'} = \overline{\boldsymbol{\psi}_e}$ for all $e = (f,f') \in \mathcal{E}_{\text{full}}$.

■

## Appendix A
## Proof of Proposition 17

The following operations on a Hermitian matrix $M_e \in \mathbb{C}^{|\mathcal{X}_e|^2 \times |\mathcal{X}_e|^2}$ with row index $x_e$, $\mathring{x}_e$ and column index $x_e'$, $\mathring{x}_e'$ are Hermitian-preserving operations:

1) $cM_e$ for some $c \in \mathbb{R}$.
2) The Hadamard product between $M_e$ and $M_{e'}$, where $M_{e'}$ is also a Hermitian matrix with the same size as $M_e$.
3) The partial trace of $M_e$. The matrix

$$\sum_{x_e, x_e'} M_e\big((x_e, \mathring{x}_e), (x_e', \mathring{x}_e')\big),$$

with row index $\mathring{x}_e$ and column index $\mathring{x}_e'$ is a Hermitian matrix. Also the matrix

$$\sum_{\mathring{x}_e, \mathring{x}_e'} M_e\big((x_e, \mathring{x}_e), (x_e', \mathring{x}_e')\big),$$

with row index $x_e$ and column index $x_e'$ is a Hermitian matrix.

Given a function $f$'s corresponding Choi-matrix representation [26] $(f(\boldsymbol{x}_{\partial f,f}, \boldsymbol{x}'_{\partial f,f}))_{\boldsymbol{x}_{\partial f,f}, \boldsymbol{x}'_{\partial f,f}}$ and its associated matrices for loop-calculus transform, i.e., $M_{e\to f}((x_{e,f}, \mathring{x}_e), (x'_{e,f'}, \mathring{x}'_e))_{\tilde{x}_{e,f}, \tilde{\mathring{x}}_e \in \tilde{\mathcal{X}}}$ in Choi-matrix representation with row indices $(x_{e,f}, \mathring{x}_e)$ and column indices $(x'_{e,f'}, \mathring{x}'_e)$ such that $M_{e\to f}((x_{e,f}, \mathring{x}_e), (x'_{e,f'}, \mathring{x}'_e)) = M_{e\to f}(\tilde{x}_{e,f}, \tilde{\mathring{x}}_e)$ for $e \in \partial f$, one can verify that they are all Hermitian matrices. The loop-calculus transform in (1) can be decomposed into two consecutive operations:

1) The Hadamard product between matrices $(f(\boldsymbol{x}_{\partial f,f}, \boldsymbol{x}'_{\partial f,f}))_{\boldsymbol{x}_{\partial f,f}, \boldsymbol{x}'_{\partial f,f}} \bigotimes I$ and $\bigotimes_{e\in\partial f} M_{e\to f}((x_{e,f}, \mathring{x}_e), (x'_{e,f'}, \mathring{x}'_e))_{\tilde{x}_{e,f}, \tilde{\mathring{x}}_e \in \tilde{\mathcal{X}}}$, where $I$ is the identity matrix with size $\prod_{e\in\partial f} |\mathcal{X}_e| \times \prod_{e\in\partial f} |\mathcal{X}_e|$.
2) The partial trace over the indices of the edges between local function and its neighboring loop-calculus transform matrices, i.e.,

$$f(\mathring{\boldsymbol{x}}_{\partial f}, \mathring{\boldsymbol{x}}'_{\partial f}) = \sum_{\boldsymbol{x}_{\partial f,f}, \boldsymbol{x}'_{\partial f,f}} f(\boldsymbol{x}_{\partial f,f}, \boldsymbol{x}'_{\partial f,f}) \prod_{e\in\partial f} M_{e\to f}((x_{e,f}, \mathring{x}_e), (x'_{e,f}, \mathring{x}'_e)).$$

Both these two operations are Hermitian-preserving operations. Thus the resulting matrix $(f(\mathring{\boldsymbol{x}}_{\partial f}, \mathring{\boldsymbol{x}}'_{\partial f}))_{\mathring{\boldsymbol{x}}_{\partial f}, \mathring{\boldsymbol{x}}'_{\partial f}}$ is Hermitian.

## Appendix B
## Proof of Lemma 18

The main idea of this proof comes from [29]. In the following, we prove Lemma 18 when $|\mathring{\mathcal{X}}_e| = 2$ and this proof can be generalized to cases with alphabet size $|\mathring{\mathcal{X}}_e|$ greater than 2.

Let us define $\boldsymbol{w}_e := (w_e(0), w_e(1), w_e(2), w_e(3))$. Since $w_e(0)$, $w_e(1)$, $w_e(2)$, $w_e(3)$ are uniformly distributed when $\|\boldsymbol{w}_e\| = r$, we have

$$\mathbb{E}\big(f_1(\boldsymbol{w}_e)\big) = \int_{\mathbb{CP}^3} f_1(\boldsymbol{w}_e) \, \mathrm{d}\mu_{\text{FS}}(\boldsymbol{w}_e),$$

where $f_1(\boldsymbol{w}_e) := ((w_e(0))^2 + (w_e(1))^2)^{K_1}((w_e(2))^2 + (w_e(3))^2)^{K_2}/\|\boldsymbol{w}_e\|^{2(K_1+K_2)}$. After some calculations, one obtains

$$\int_{\mathbb{R}^4} \frac{1}{4\pi^2} f_1(\boldsymbol{w}_e) e^{-\frac{\|\boldsymbol{w}_e\|^2}{2}} \, \mathrm{d}x = 2^{K_1+K_2} K_1! K_2!. \tag{6}$$

Let us define a sphere in $\mathbb{R}^4$ with radius $r$ to be $S_4(r) := \{\boldsymbol{w}_e \in \mathbb{R}^4 | \|z\| = r\}$. Evaluating (6) in polar coordinates gives

$$\begin{aligned}
&\frac{1}{4\pi^2} \int_{S_4} \int_0^\infty r^{2(K_1+K_2)+3} e^{-\frac{r^2}{2}} f_1(\boldsymbol{w}_e) \, \mathrm{d}r \, \mathrm{d}\sigma(\boldsymbol{w}_e) \\
&= \frac{1}{4\pi^2} \int_0^\infty r^{2(K_1+K_2)+3} e^{-\frac{r^2}{2}} \, \mathrm{d}r \int_{S_4} f_1(\boldsymbol{w}_e) \, \mathrm{d}\sigma(\boldsymbol{w}_e) \\
&= \frac{2^{K_1+K_2+1}(K_1+K_2+1)!}{4\pi^2} \int_{S_4} f_1(\boldsymbol{w}_e) \, \mathrm{d}\sigma(\boldsymbol{w}_e).
\end{aligned} \tag{7}$$

Since both $\frac{2}{4\pi^2}\sigma(\boldsymbol{w}_e)$ and $\mu_{\mathrm{FS}}(\boldsymbol{w}_e)$ are measures induced from the Haar measure [30], comparing (6) with (7), we have

$$\begin{aligned}
\int_{S_4} \frac{2 f_1(\boldsymbol{w}_e)}{4\pi^2} \, \mathrm{d}\sigma(\boldsymbol{w}_e) &= \int_{\mathbb{CP}^3} f_1(\boldsymbol{w}_e) \, \mathrm{d}\mu_{\mathrm{FS}}(\boldsymbol{w}_e) \\
&= \frac{K_1! K_2!}{(K_1+K_2+1)!}.
\end{aligned}$$

## APPENDIX C
## PROOF OF THEOREM 19

Let us define a matrix $D_{\sigma_e,M}$ for each edge such that given $\sigma_e \in \mathcal{S}_M$, $\mathring{\tilde{\boldsymbol{x}}}_{1,M}$, $\mathring{\tilde{\boldsymbol{x}}}_{2,M} \in \mathring{\mathcal{X}}_e^M$, the entry of matrix $D_{\sigma_e,M}$ with row index $\mathring{\tilde{\boldsymbol{x}}}_{1,M}$ and column index $\mathring{\tilde{\boldsymbol{x}}}_{2,M}$ equals

$$D_{\sigma_e,M}(\mathring{\tilde{\boldsymbol{x}}}_{1,M}, \mathring{\tilde{\boldsymbol{x}}}_{2,M}) := \prod_{m=1}^{M} \Big[ \mathring{\tilde{x}}_{1,m} = \sigma_e(\mathring{\tilde{x}}_{2,m}) \Big],$$

where $I \in \mathbb{R}^{|\mathring{\mathcal{X}}_e| \times |\mathring{\mathcal{X}}_e|}$ is the identity matrix. Notice that $D_{\sigma_e,M}$ is a permutation matrix. If $\sigma_e \in \mathcal{S}_M$ denotes the identity permutation, then $D_{\sigma_e,M} = I^{\otimes M}$, where $I^{\otimes M}$ is the standard $M$-fold Kronecker product of the identity matrix. Define

$$T_{e,M} := \frac{1}{M!} \sum_{\sigma \in \mathcal{S}_M} D_{\sigma_e,M},$$

$$\beta_{e,\mathring{\tilde{\boldsymbol{x}}}_{e,M}}(\mathring{\tilde{x}}_e) := \sum_{m=1}^{M} \Big[ \mathring{\tilde{x}}_{e,m} = \mathring{\tilde{x}}_e \Big], \tag{8}$$

for all $\mathring{\tilde{\boldsymbol{x}}}_{e,M} \in \mathring{\mathcal{X}}_e^M, \mathring{\tilde{x}}_e \in \mathring{\mathcal{X}}, e \in \mathcal{E}_{\mathrm{full}}$. One can verify that for $\mathring{\tilde{\boldsymbol{x}}}_{1,M}, \mathring{\tilde{\boldsymbol{x}}}_{2,M} \in \mathring{\mathcal{X}}^M$,

$$T_{e,M}(\mathring{\tilde{\boldsymbol{x}}}_{1,M}, \mathring{\tilde{\boldsymbol{x}}}_{2,M}) = \frac{\prod_{\mathring{\tilde{x}}_e} \beta_{e,\mathring{\tilde{\boldsymbol{x}}}_{1,M}}(\mathring{\tilde{x}}_e)!}{M!} \cdot \Big[ \boldsymbol{\beta}_{e,\mathring{\tilde{\boldsymbol{x}}}_{1,M}} = \boldsymbol{\beta}_{e,\mathring{\tilde{\boldsymbol{x}}}_{2,M}} \Big], \tag{9}$$

where $\boldsymbol{\beta}_{e,\mathring{\tilde{\boldsymbol{x}}}_{1,M}} = \boldsymbol{\beta}_{e,\mathring{\tilde{\boldsymbol{x}}}_{2,M}}$ means the vector $\boldsymbol{\beta}_{e,\mathring{\tilde{\boldsymbol{x}}}_{1,M}}$ equals the vector $\boldsymbol{\beta}_{e,\mathring{\tilde{\boldsymbol{x}}}_{2,M}}$ entrywise. With these definitions, we have

$$\big\langle Z(\hat{\mathsf{N}}) \big\rangle_{\hat{\mathsf{N}} \in \hat{\mathcal{N}}_M} = \sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F},M}} \prod_{m=1}^{M} \prod_{f} \mathring{f}(\mathring{\tilde{\boldsymbol{x}}}_{\partial f,f,m}) \cdot \prod_{e} T_{e,M}(\mathring{\tilde{\boldsymbol{x}}}_{e,f,M}, \mathring{\tilde{\boldsymbol{x}}}_{e,f',M}), \tag{10}$$

where $e = (f, f') \in \mathcal{E}_{\mathrm{full}}$, $\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F},M} := (\mathring{\tilde{x}}_{e,f,m})_{e \in \partial f, f \in \mathcal{F}, m \in [M]}$ and $\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F},M}} = \sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F},M} \in \mathring{\mathcal{X}}^{2M}}$. As a result, we can reduce the problem of computing $\big\langle Z(\hat{\mathsf{N}}) \big\rangle_{\hat{\mathsf{N}} \in \hat{\mathcal{N}}_M}$ to a problem of computing the partition function over the original DE-NFG $\mathsf{N}$ with a matrix $T_{e,M}$ on each edge and a function $\prod_{m=1}^{M} f$ for each function node.

For $\mathring{\tilde{\boldsymbol{x}}}_{1,M}, \mathring{\tilde{\boldsymbol{x}}}_{2,M} \in \mathring{\mathcal{X}}^M$, the entry in $\mathbb{E}\big((\check{\boldsymbol{\psi}}_e \check{\boldsymbol{\psi}}_e^{\mathsf{H}})^{\otimes M}\big)$ with row index $\mathring{\tilde{\boldsymbol{x}}}_{1,M}$ and column index $\mathring{\tilde{\boldsymbol{x}}}_{2,M}$ equals

$$\mathbb{E}\Bigg( \prod_{\mathring{\tilde{x}}_e} \big| \check{\psi}_e(\mathring{\tilde{x}}_e) \big|^{2\beta_{e,\mathring{\tilde{\boldsymbol{x}}}_{1,M}}(\mathring{\tilde{x}}_e)} \Bigg) \cdot \Big[ \boldsymbol{\beta}_{e,\mathring{\tilde{\boldsymbol{x}}}_{1,M}} = \boldsymbol{\beta}_{e,\mathring{\tilde{\boldsymbol{x}}}_{2,M}} \Big]. \tag{11}$$

Recall that $\boldsymbol{\beta}_{e,\mathring{\tilde{\boldsymbol{x}}}_{1,M}}$ and $\boldsymbol{\beta}_{e,\mathring{\tilde{\boldsymbol{x}}}_{2,M}}$ are defined in (8). With Proposition 18, the entry of matrix $\mathbb{E}\big((\check{\boldsymbol{\psi}}_e\check{\boldsymbol{\psi}}_e^{\mathsf{H}})^{\otimes M}\big)$ with row index $\mathring{\tilde{\boldsymbol{x}}}_{1,M}$ and column index $\mathring{\tilde{\boldsymbol{x}}}_{2,M}$ equals

$$\frac{\prod_{\mathring{\tilde{x}}_e}\beta_{e,\mathring{\tilde{\boldsymbol{x}}}_{1,M}}(\mathring{\tilde{x}}_e)!}{(M+1)!}\cdot\Big[\boldsymbol{\beta}_{e,\mathring{\tilde{\boldsymbol{x}}}_{1,M}}=\boldsymbol{\beta}_{e,\mathring{\tilde{\boldsymbol{x}}}_{2,M}}\Big], \tag{12}$$

where $\prod_{\mathring{\tilde{x}}_e}$ denotes $\prod_{\mathring{\tilde{x}}_e\in\mathring{\tilde{\mathcal{X}}}_e}$. Comparing the above term with (9), we have $(M+1)\mathbb{E}\big((\check{\boldsymbol{\psi}}_e\check{\boldsymbol{\psi}}_e^{\mathsf{H}})^{\otimes M}\big)=T_{e,M}$.

## Appendix D
## Proof of Theorem 20

Here we prove the local maximum at $\boldsymbol{\psi}_0$ by proving that the Hessian matrix of $-\operatorname{Re}(\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi}))$ with respect to $\boldsymbol{\psi}$ at $\boldsymbol{\psi}=\boldsymbol{\psi}_0$ is PSD. For simplicity, we define

$$g_{(e_1,f_1)}(\mathring{\tilde{x}}_{e_1}):=g_1\left(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}:\substack{\mathring{\tilde{x}}_{e_1,f_1}=\mathring{\tilde{x}}_{e_1},\\ \mathring{\tilde{x}}_{e,f}=\mathring{\tilde{0}}\ \forall(e,f)\in(\mathcal{E}_{\mathrm{full}}\times\mathcal{F})\\ \setminus\{(e_1,f_1)\}:\ e\in\partial f}\right), \tag{13}$$

$$g_{(e_1,f_1),(e_2,f_2)}(\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2}):=g_1\left(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}:\substack{\mathring{\tilde{x}}_{e_1,f_1}=\mathring{\tilde{x}}_{e_1},\ \mathring{\tilde{x}}_{e_2,f_2}=\mathring{\tilde{x}}_{e_2},\\ \mathring{\tilde{x}}_{e,f}=\mathring{\tilde{0}}\ \forall(e,f)\in(\mathcal{E}_{\mathrm{full}}\times\mathcal{F})\\ \setminus\{(e_1,f_1),(e_2,f_2)\}:\ e\in\partial f}\right), \tag{14}$$

$$S((e_1,f_1),(e_2,f_2),\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2}):=\{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}\in(\mathring{\tilde{\mathcal{X}}})^2:\ \mathring{\tilde{x}}_{e_1,f_1}=\mathring{\tilde{x}}_{e_1},\ \mathring{\tilde{x}}_{e_2,f_2}=\mathring{\tilde{x}}_{e_2}\}, \tag{15}$$

for all $\mathring{\tilde{x}}_{e_1}\in\mathring{\tilde{\mathcal{X}}}_{e_1}$, $\mathring{\tilde{x}}_{e_2}\in\mathring{\tilde{\mathcal{X}}}_{e_2}$, $e_1=(f_1,f_1')$, $e_2=(f_2,f_2')\in\mathcal{E}_{\mathrm{full}}$.

### A. First Derivatives of $\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})$ with respect to $\boldsymbol{\psi}$ at $\boldsymbol{\psi}=\boldsymbol{\psi}_0$

**Lemma 23.** *When $\boldsymbol{\psi}=\boldsymbol{\psi}_0$, we have*

$$\left.\frac{\partial\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})}{\partial\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}\right|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0}=0,$$
$$\left.\frac{\partial\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})}{\partial\overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}}\right|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0}=0,$$

*for all $e_1=(f_1,f_1')\in\mathcal{E}_{\mathrm{full}}$, $\mathring{\tilde{x}}_{e_1}\in\mathring{\tilde{\mathcal{X}}}_{e_1}$.*

*Proof.* The Wirtinger derivatives of $\mathring{g}_2(\boldsymbol{\psi})$ with respect to $\psi_{e_1}(\mathring{\tilde{x}}_{e_1})$ and $\overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}$ equal, respectively,

$$\begin{aligned}\frac{\partial\mathring{g}_2(\boldsymbol{\psi})}{\partial\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}&=\frac{\partial\operatorname{Re}(\mathring{g}_2(\boldsymbol{\psi}))}{\partial\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}+i\frac{\partial\operatorname{Im}(\mathring{g}_2(\boldsymbol{\psi}))}{\partial\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}\\&=\frac{\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}\in S((e_1,f_1),\mathring{\tilde{x}}_{e_1})}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1,f_1'})}\prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}}\psi_e(\mathring{\tilde{x}}_{e,f})\overline{\psi_e(\mathring{\tilde{x}}_{e,f'})}}{\prod_{e\in\mathcal{E}_{\mathrm{full}}}\|\boldsymbol{\psi}_e\|^2}\\&\quad-\frac{\overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}}\psi_e(\mathring{\tilde{x}}_{e,f})\overline{\psi_e(\mathring{\tilde{x}}_{e,f'})}}{\|\boldsymbol{\psi}_{e_1}\|^4\prod_{e\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}}\|\boldsymbol{\psi}_e\|^2},\end{aligned} \tag{16}$$

$$\begin{aligned}\frac{\partial\mathring{g}_2(\boldsymbol{\psi})}{\partial\overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}}&=\frac{\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}\in S((e_1,f_1'),\mathring{\tilde{x}}_{e_1})}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\psi_{e_1}(\mathring{\tilde{x}}_{e_1,f_1})\prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}}\psi_e(\mathring{\tilde{x}}_{e,f})\overline{\psi_e(\mathring{\tilde{x}}_{e,f'})}}{\prod_{e\in\mathcal{E}_{\mathrm{full}}}\|\boldsymbol{\psi}_e\|^2}\\&\quad-\frac{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}}\psi_e(\mathring{\tilde{x}}_{e,f})\overline{\psi_e(\mathring{\tilde{x}}_{e,f'})}}{\|\boldsymbol{\psi}_{e_1}\|^4\prod_{e\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}}\|\boldsymbol{\psi}_e\|^2}.\end{aligned} \tag{17}$$

At $\boldsymbol{\psi}=\boldsymbol{\psi}_0$, we have

$$\left.\frac{\partial\mathring{g}_2(\boldsymbol{\psi})}{\partial\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}\right|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0}=g_{(e_1,f_1)}(\mathring{\tilde{x}}_{e_1})-\Big[\mathring{\tilde{x}}_{e_1}=\mathring{\tilde{0}}\Big]\cdot\mathring{g}(\mathring{\tilde{\mathbf{0}}}),$$
$$\left.\frac{\partial\mathring{g}_2(\boldsymbol{\psi})}{\partial\overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}}\right|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0}=g_{(e_1,f_1')}(\mathring{\tilde{x}}_{e_1})-\Big[\mathring{\tilde{x}}_{e_1}=\mathring{\tilde{0}}\Big]\cdot\mathring{g}(\mathring{\tilde{\mathbf{0}}}).$$

When $\mathring{\tilde{x}}_{e_1} = \mathring{\tilde{0}}$, the above expressions equal zero. When $\mathring{\tilde{x}}_{e_1} \neq \mathring{\tilde{0}}$, by definition of LCT, we also have $g_{(e_1,f_1)}(\mathring{\tilde{x}}_{e_1}) = 0$ and $g_{(e_1,f_1')}(\mathring{\tilde{x}}_{e_1}) = 0$, which implies

$$\left.\frac{\partial \mathring{g}_2(\boldsymbol{\psi})}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})}\right|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} = 0,$$

$$\left.\frac{\partial \mathring{g}_2(\boldsymbol{\psi})}{\partial \overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}}\right|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} = 0.$$

Note that

$$\frac{\partial \mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})} = M(\mathring{g}_2(\boldsymbol{\psi}))^{M-1} \frac{\partial \mathring{g}_2(\boldsymbol{\psi})}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})},$$

$$\frac{\partial \mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})}{\partial \overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}} = M(\mathring{g}_2(\boldsymbol{\psi}))^{M-1} \frac{\partial \mathring{g}_2(\boldsymbol{\psi})}{\partial \overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}}.$$

■

**Lemma 24.** *At $\boldsymbol{\psi} = \boldsymbol{\psi}_0$, it holds*

$$\left.\frac{\partial \overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})}\right|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} = 0,$$

$$\left.\frac{\partial \overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial \overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}}\right|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} = 0.$$

*Proof.* We have

$$\begin{aligned}\frac{\partial \overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})} = &\frac{\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}} \in S((e_1,f_1'),\mathring{\tilde{x}}_{e_1})} \overline{\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}) \psi_{e_1}(\mathring{\tilde{x}}_{e_1,f_1})} \prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}} \overline{\psi_e(\mathring{\tilde{x}}_{e,f})} \psi_e(\mathring{\tilde{x}}_{e,f'})}{\prod_{e\in\mathcal{E}_{\mathrm{full}}} \|\boldsymbol{\psi}_e\|^2} \\ &- \frac{\overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})} \sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}} \overline{\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})} \prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}} \overline{\psi_e(\mathring{\tilde{x}}_{e,f})} \psi_e(\mathring{\tilde{x}}_{e,f'})}{\|\boldsymbol{\psi}_{e_1}\|^4 \prod_{e\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}} \|\boldsymbol{\psi}_e\|^2},\end{aligned}$$

$$\begin{aligned}\frac{\partial \overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial \overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}} = &\frac{\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}} \in S((e_1,f_1),\mathring{\tilde{x}}_{e_1})} \overline{\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})} \psi_{e_1}(\mathring{\tilde{x}}_{e_1,f_1'}) \prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}} \overline{\psi_e(\mathring{\tilde{x}}_{e,f})} \psi_e(\mathring{\tilde{x}}_{e,f'})}{\prod_{e\in\mathcal{E}_{\mathrm{full}}} \|\boldsymbol{\psi}_e\|^2} \\ &- \frac{\psi_{e_1}(\mathring{\tilde{x}}_{e_1}) \sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}} \overline{\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})} \prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}} \overline{\psi_e(\mathring{\tilde{x}}_{e,f})} \psi_e(\mathring{\tilde{x}}_{e,f'})}{\|\boldsymbol{\psi}_{e_1}\|^4 \prod_{e\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}} \|\boldsymbol{\psi}_e\|^2}.\end{aligned}$$

When $\boldsymbol{\psi} = \boldsymbol{\psi}_0$, we have

$$\begin{aligned}\left.\frac{\partial \overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})}\right|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} &= \overline{g_{(e_1,f_1')}(\mathring{\tilde{x}}_{e_1})} - \left[\mathring{\tilde{x}}_{e_1} = \mathring{\tilde{0}}\right] \cdot \mathring{g}(\mathring{\tilde{\mathbf{0}}}) \\ &= 0,\end{aligned}$$

$$\begin{aligned}\left.\frac{\partial \overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial \overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}}\right|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} &= \overline{g_{(e_1,f_1)}(\mathring{\tilde{x}}_{e_1})} - \left[\mathring{\tilde{x}}_{e_1} = \mathring{\tilde{0}}\right] \cdot \mathring{g}(\mathring{\tilde{\mathbf{0}}}) \\ &= 0,\end{aligned}$$

for all $e_1 \in \mathcal{E}_{\mathrm{full}}$, $\mathring{\tilde{x}}_{e_1} \in \mathring{\mathcal{X}}_{e_1}$. ■

### *B. Hessian Matrix of $\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})$ with respect to $\boldsymbol{\psi}$ at $\boldsymbol{\psi} = \boldsymbol{\psi}_0$*

**Lemma 25.** *When $\boldsymbol{\psi} = \boldsymbol{\psi}_0$, we have*

$$\left.\frac{\partial}{\partial \psi_{e_2}(\mathring{\tilde{x}}_{e_2})} \overline{\left(\frac{\partial \overline{\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})}}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})}\right)}\right|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} = M(\mathring{g}(\mathring{\tilde{\mathbf{0}}}))^{M-1} \frac{\partial}{\partial \psi_{e_2}(\mathring{\tilde{x}}_{e_2})} \left.\overline{\left(\frac{\partial \overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})}\right)}\right|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0}, \tag{18}$$

$$\frac{\partial}{\partial \overline{\psi_{e_2}(\mathring{\tilde{x}}_{e_2})}} \overline{\left( \frac{\partial \overline{\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})}}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})} \right)} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} = M(\mathring{g}(\mathring{\tilde{\mathbf{0}}}))^{M-1} \frac{\partial}{\partial \overline{\psi_{e_2}(\mathring{\tilde{x}}_{e_2})}} \overline{\left( \frac{\partial \overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})} \right)} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0}, \tag{19}$$

$$\frac{\partial}{\partial \psi_{e_2}(\mathring{\tilde{x}}_{e_2})} \overline{\left( \frac{\partial \overline{\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})}}{\partial \overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}} \right)} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} = M(\mathring{g}(\mathring{\tilde{\mathbf{0}}}))^{M-1} \frac{\partial}{\partial \psi_{e_2}(\mathring{\tilde{x}}_{e_2})} \overline{\left( \frac{\partial \overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial \overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}} \right)} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0}, \tag{20}$$

$$\frac{\partial}{\partial \overline{\psi_{e_2}(\mathring{\tilde{x}}_{e_2})}} \overline{\left( \frac{\partial \overline{\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})}}{\partial \overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}} \right)} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} = M(\mathring{g}(\mathring{\tilde{\mathbf{0}}}))^{M-1} \frac{\partial}{\partial \overline{\psi_{e_2}(\mathring{\tilde{x}}_{e_2})}} \overline{\left( \frac{\partial \overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial \overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}} \right)} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0}, \tag{21}$$

*for all* $e_1 = (f_1, f_1')$, $e_2 = (f_2, f_2') \in \mathcal{E}_{\mathrm{full}}$, $\mathring{\tilde{x}}_{e_1} \in \mathring{\mathcal{X}}_{e_1}$, $\mathring{\tilde{x}}_{e_2} \in \mathring{\mathcal{X}}_{e_2}$.

*Proof.* We have

$$\begin{aligned} \frac{\partial}{\partial \psi_{e_2}(\mathring{\tilde{x}}_{e_2})} \overline{\left( \frac{\partial \overline{\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})}}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})} \right)} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} &= M(M-1)(\mathring{g}(\mathring{\tilde{\mathbf{0}}}))^{M-2} \cdot \frac{\partial \mathring{g}_2(\boldsymbol{\psi})}{\partial \psi_{e_2}(\mathring{\tilde{x}}_{e_2})} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} \cdot \overline{\left( \frac{\partial \overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})} \right)} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} \\ &\quad + M(\mathring{g}(\mathring{\tilde{\mathbf{0}}}))^{M-1} \frac{\partial}{\partial \psi_{e_2}(\mathring{\tilde{x}}_{e_2})} \overline{\left( \frac{\partial \overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})} \right)} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} \\ &= M(\mathring{g}(\mathring{\tilde{\mathbf{0}}}))^{M-1} \frac{\partial}{\partial \psi_{e_2}(\mathring{\tilde{x}}_{e_2})} \overline{\left( \frac{\partial \overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})} \right)} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0}. \end{aligned}$$

The proof of expressions (19)–(21) is similar and thus it is omitted here. ■

Let $\phi_e : \mathring{\mathcal{X}} \to \left[ \sum_{e' \in \mathcal{E}_{\mathrm{full}}} |\mathring{\mathcal{X}}_{e'}| \right]$ be the bijection that sends an element $\mathring{\tilde{x}}_e \in \mathring{\mathcal{X}}_e$ to the position of $(\ldots, \mathring{\tilde{0}}, \mathring{\tilde{x}}_e, \mathring{\tilde{0}}, \ldots) \in \prod_{e' \in \mathcal{E}_{\mathrm{full}}} \mathring{\mathcal{X}}_{e'}$ in lexicographical order among all vectors in $\prod_{e' \in \mathcal{E}_{\mathrm{full}}} \mathring{\mathcal{X}}_{e'}$.

The matrices $H_{\boldsymbol{\psi},1}$, $H_{\boldsymbol{\psi},2}$, $H_{\boldsymbol{\psi},3}$ and $H_{\boldsymbol{\psi},4}$ with column index $\phi_{e_1}(\mathring{\tilde{x}}_{e_1})$ and row index $\phi_{e_2}(\mathring{\tilde{x}}_{e_2})$ is defined to be

$$H_{\boldsymbol{\psi},1}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}), \phi_{e_1}(\mathring{\tilde{x}}_{e_1})) := H_{\boldsymbol{\psi},1,\mathrm{re}}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}), \phi_{e_1}(\mathring{\tilde{x}}_{e_1})) + i H_{\boldsymbol{\psi},1,\mathrm{im}}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}), \phi_{e_1}(\mathring{\tilde{x}}_{e_1})),$$

$$H_{\boldsymbol{\psi},2}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}), \phi_{e_1}(\mathring{\tilde{x}}_{e_1})) := H_{\boldsymbol{\psi},2,\mathrm{re}}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}), \phi_{e_1}(\mathring{\tilde{x}}_{e_1})) + i H_{\boldsymbol{\psi},2,\mathrm{im}}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}), \phi_{e_1}(\mathring{\tilde{x}}_{e_1})),$$

$$H_{\boldsymbol{\psi},3}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}), \phi_{e_1}(\mathring{\tilde{x}}_{e_1})) := H_{\boldsymbol{\psi},3,\mathrm{re}}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}), \phi_{e_1}(\mathring{\tilde{x}}_{e_1})) + i H_{\boldsymbol{\psi},3,\mathrm{im}}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}), \phi_{e_1}(\mathring{\tilde{x}}_{e_1})),$$

$$H_{\boldsymbol{\psi},4}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}), \phi_{e_1}(\mathring{\tilde{x}}_{e_1})) := H_{\boldsymbol{\psi},4,\mathrm{re}}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}), \phi_{e_1}(\mathring{\tilde{x}}_{e_1})) + i H_{\boldsymbol{\psi},4,\mathrm{im}}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}), \phi_{e_1}(\mathring{\tilde{x}}_{e_1})),$$

for all $\mathring{\tilde{x}}_{e_1} \in \mathring{\mathcal{X}}_{e_1}$, $\mathring{\tilde{x}}_{e_2} \in \mathring{\mathcal{X}}_{e_2}$, $e_1 = (f_1, f_1')$, $e_2 = (f_2, f_2') \in \mathcal{E}_{\mathrm{full}}$, where

$$H_{\boldsymbol{\psi},1,\mathrm{re}}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}), \phi_{e_1}(\mathring{\tilde{x}}_{e_1})) := \frac{\partial}{\partial \psi_{e_2}(\mathring{\tilde{x}}_{e_2})} \overline{\left( \frac{\partial \, \mathrm{Re}(\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi}))}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})} \right)} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} \in \mathbb{C},$$

$$H_{\boldsymbol{\psi},1,\mathrm{im}}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}), \phi_{e_1}(\mathring{\tilde{x}}_{e_1})) := \frac{\partial}{\partial \psi_{e_2}(\mathring{\tilde{x}}_{e_2})} \overline{\left( \frac{\partial \, \mathrm{Im}(\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi}))}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})} \right)} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} \in \mathbb{C},$$

$$H_{\boldsymbol{\psi},2,\mathrm{re}}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}), \phi_{e_1}(\mathring{\tilde{x}}_{e_1})) := \frac{\partial}{\partial \psi_{e_2}(\mathring{\tilde{x}}_{e_2})} \overline{\left( \frac{\partial \, \mathrm{Re}(\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi}))}{\partial \overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}} \right)} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} \in \mathbb{C},$$

$$H_{\boldsymbol{\psi},2,\mathrm{im}}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}), \phi_{e_1}(\mathring{\tilde{x}}_{e_1})) := \frac{\partial}{\partial \psi_{e_2}(\mathring{\tilde{x}}_{e_2})} \overline{\left( \frac{\partial \, \mathrm{Im}(\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi}))}{\partial \overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}} \right)} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} \in \mathbb{C},$$

$$H_{\boldsymbol{\psi},3,\mathrm{re}}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}), \phi_{e_1}(\mathring{\tilde{x}}_{e_1})) := \frac{\partial}{\partial \overline{\psi_{e_2}(\mathring{\tilde{x}}_{e_2})}} \overline{\left( \frac{\partial \, \mathrm{Re}(\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi}))}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})} \right)} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} \in \mathbb{C},$$

$$H_{\boldsymbol{\psi},3,\mathrm{im}}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}), \phi_{e_1}(\mathring{\tilde{x}}_{e_1})) := \frac{\partial}{\partial \overline{\psi_{e_2}(\mathring{\tilde{x}}_{e_2})}} \overline{\left( \frac{\partial \, \mathrm{Im}(\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi}))}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})} \right)} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} \in \mathbb{C},$$

$$H_{\boldsymbol{\psi},4,\mathrm{re}}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}), \phi_{e_1}(\mathring{\tilde{x}}_{e_1})) := \frac{\partial}{\partial \overline{\psi_{e_2}(\mathring{\tilde{x}}_{e_2})}} \overline{\left( \frac{\partial \, \mathrm{Re}(\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi}))}{\partial \overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}} \right)} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} \in \mathbb{C},$$

$$H_{\boldsymbol{\psi},4,\mathrm{im}}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}),\phi_{e_1}(\mathring{\tilde{x}}_{e_1})) := \frac{\partial}{\partial\overline{\psi_{e_2}(\mathring{\tilde{x}}_{e_2})}}\overline{\left(\frac{\partial\,\mathrm{Im}(\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi}))}{\partial\overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}}\right)}\Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} \in \mathbb{C}.$$

**Lemma 26.** *We have*

$$H_{\boldsymbol{\psi},1}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}),\phi_{e_1}(\mathring{\tilde{x}}_{e_1})) = M(\mathring{g}(\mathring{\tilde{\mathbf{0}}}))^{M-1}\cdot\left[\mathring{\tilde{x}}_{e_1}\neq\mathring{\tilde{0}},\mathring{\tilde{x}}_{e_2}\neq\mathring{\tilde{0}}\right]\cdot\left(g_{(e_1,f_1'),(e_2,f_2)}(\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2}) - [e_1=e_2]\cdot\left[\mathring{\tilde{x}}_{e_1}=\mathring{\tilde{x}}_{e_2}\right]\cdot\mathring{g}(\mathring{\tilde{\mathbf{0}}})\right), \quad (22)$$

$$H_{\boldsymbol{\psi},2}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}),\phi_{e_1}(\mathring{\tilde{x}}_{e_1})) = M(\mathring{g}(\tilde{\mathbf{0}}))^{M-1}\cdot\left[\mathring{\tilde{x}}_{e_1}\neq\mathring{\tilde{0}},\mathring{\tilde{x}}_{e_2}\neq\mathring{\tilde{0}}\right]\cdot[e_1\neq e_2]\cdot g_{(e_1,f_1),(e_2,f_2)}(\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2}), \quad (23)$$

$$H_{\boldsymbol{\psi},3}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}),\phi_{e_1}(\mathring{\tilde{x}}_{e_1})) = M(\mathring{g}(\mathring{\tilde{\mathbf{0}}}))^{M-1}\cdot\left[\mathring{\tilde{x}}_{e_1}\neq\mathring{\tilde{0}},\mathring{\tilde{x}}_{e_2}\neq\mathring{\tilde{0}}\right]\cdot[e_1\neq e_2]\cdot g_{(e_1,f_1'),(e_2,f_2')}(\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2}), \quad (24)$$

$$H_{\boldsymbol{\psi},4}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}),\phi_{e_1}(\mathring{\tilde{x}}_{e_1})) = M(\mathring{g}(\mathring{\tilde{\mathbf{0}}}))^{M-1}\cdot\left[\mathring{\tilde{x}}_{e_1}\neq\mathring{\tilde{0}},\mathring{\tilde{x}}_{e_2}\neq\mathring{\tilde{0}}\right]\cdot\left(g_{(e_1,f_1),(e_2,f_2')}(\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2}) - [e_1=e_2]\cdot\left[\mathring{\tilde{x}}_{e_1}=\mathring{\tilde{x}}_{e_2}\right]\cdot\mathring{g}(\mathring{\tilde{\mathbf{0}}})\right), \quad (25)$$

*for all* $\mathring{\tilde{x}}_{e_1}\in\mathring{\tilde{\mathcal{X}}}_{e_1}$, $\mathring{\tilde{x}}_{e_2}\in\mathring{\tilde{\mathcal{X}}}_{e_2}$, $e_1=(f_1,f_1')$, $e_2=(f_2,f_2')\in\mathcal{E}_{\mathrm{full}}$.

*Proof.* Note that

$$\begin{aligned}
H_{\boldsymbol{\psi},1}(\phi_{e_2}(\mathring{\tilde{x}}_{e_2}),\phi_{e_1}(\mathring{\tilde{x}}_{e_1})) &= \frac{\partial}{\partial\psi_{e_2}(\mathring{\tilde{x}}_{e_2})}\overline{\left(\frac{\partial\,\mathrm{Re}(\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi}))}{\partial\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}\right)}\Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} + i\,\frac{\partial}{\partial\psi_{e_2}(\mathring{\tilde{x}}_{e_2})}\overline{\left(\frac{\partial\,\mathrm{Im}(\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi}))}{\partial\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}\right)}\Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} \\
&= \frac{\partial}{\partial\psi_{e_2}(\mathring{\tilde{x}}_{e_2})}\overline{\left(\frac{\partial\frac{1}{2}\left(\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})+\overline{\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})}\right)}{\partial\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}\right)}\Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} + i\,\frac{\partial}{\partial\psi_{e_2}(\mathring{\tilde{x}}_{e_2})}\overline{\left(\frac{\partial\frac{1}{2i}\left(\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})-\overline{\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})}\right)}{\partial\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}\right)}\Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} \\
&= \frac{\partial}{\partial\psi_{e_2}(\mathring{\tilde{x}}_{e_2})}\overline{\left(\frac{\partial\overline{\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})}}{\partial\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}\right)}\Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0}. \quad (26)
\end{aligned}$$

With Lemma 25, it is sufficient to consider the Hessian of $\overline{\mathring{g}_2(\boldsymbol{\psi})}$. We have

$$\begin{aligned}
\overline{\left(\frac{\partial\overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}\right)} &= \frac{\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}\in S((e_1,f_1'),\mathring{\tilde{x}}_{e_1})}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\psi_{e_1}(\mathring{\tilde{x}}_{e_1,f_1})\prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}}\psi_e(\mathring{\tilde{x}}_{e,f})\overline{\psi_e(\mathring{\tilde{x}}_{e,f'})}}{\prod_{e\in\mathcal{E}_{\mathrm{full}}}\|\boldsymbol{\psi}_e\|^2} \\
&\quad - \frac{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}}\psi_e(\mathring{\tilde{x}}_{e,f})\overline{\psi_e(\mathring{\tilde{x}}_{e,f'})}}{\|\boldsymbol{\psi}_{e_1}\|^4\prod_{e\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}}\|\boldsymbol{\psi}_e\|^2}. \quad (27)
\end{aligned}$$

Suppose $e_1\neq e_2$, we have

$$\begin{aligned}
\frac{\partial}{\partial\psi_{e_2}(\mathring{\tilde{x}}_{e_2})}\overline{\left(\frac{\partial\overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}\right)} &= \frac{\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}\in S((e_1,f_1'),(e_2,f_2),\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2})}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\psi_{e_1}(\mathring{\tilde{x}}_{e_1,f_1})\overline{\psi_{e_2}(\mathring{\tilde{x}}_{e_2,f_2'})}\prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1,e_2\}}\psi_e(\mathring{\tilde{x}}_{e,f})\overline{\psi_e(\mathring{\tilde{x}}_{e,f'})}}{\prod_{e\in\mathcal{E}_{\mathrm{full}}}\|\boldsymbol{\psi}_e\|^2} \\
&\quad - \frac{\overline{\psi_{e_2}(\mathring{\tilde{x}}_{e_2})}\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}\in S((e_1,f_1'),\mathring{\tilde{x}}_{e_1})}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\psi_{e_1}(\mathring{\tilde{x}}_{e_1,f_1})\prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}}\psi_e(\mathring{\tilde{x}}_{e,f})\overline{\psi_e(\mathring{\tilde{x}}_{e,f'})}}{\|\boldsymbol{\psi}_{e_2}\|^4\prod_{e\in\mathcal{E}_{\mathrm{full}}\setminus\{e_2\}}\|\boldsymbol{\psi}_e\|^2} \\
&\quad - \frac{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}\in S((e_2,f_2),\mathring{\tilde{x}}_{e_2})}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\overline{\psi_{e_2}(\mathring{\tilde{x}}_{e_2,f_2'})}\prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}\setminus\{e_2\}}\psi_e(\mathring{\tilde{x}}_{e,f})\overline{\psi_e(\mathring{\tilde{x}}_{e,f'})}}{\|\boldsymbol{\psi}_{e_1}\|^4\prod_{e\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}}\|\boldsymbol{\psi}_e\|^2} \\
&\quad + \frac{\psi_{e_1}(\mathring{\tilde{x}}_{e_1})\overline{\psi_{e_2}(\mathring{\tilde{x}}_{e_2})}\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}}\psi_e(\mathring{\tilde{x}}_{e,f})\overline{\psi_e(\mathring{\tilde{x}}_{e,f'})}}{\|\boldsymbol{\psi}_{e_1}\|^4\|\boldsymbol{\psi}_{e_2}\|^4\prod_{e\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1,e_2\}}\|\boldsymbol{\psi}_e\|^2}.
\end{aligned}$$

If $e_1=e_2$, the derivative of $\overline{\left(\frac{\partial\overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}\right)}$ with respect to $\psi_{e_1}(\mathring{\tilde{x}}_{e_2})$ is

$$\begin{aligned}
\frac{\partial}{\partial\psi_{e_2}(\mathring{\tilde{x}}_{e_2})}\overline{\left(\frac{\partial\overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial\psi_{e_1}(\mathring{\tilde{x}}_{e_1})}\right)} &= \frac{\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}\in S((e_1,f_1'),(e_1,f_1),\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2})}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}}\psi_e(\mathring{\tilde{x}}_{e,f})\overline{\psi_e(\mathring{\tilde{x}}_{e,f'})}}{\prod_{e\in\mathcal{E}_{\mathrm{full}}}\|\boldsymbol{\psi}_e\|^2} \\
&\quad - \frac{\overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_2})}\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}\in S((e_1,f_1'),\mathring{\tilde{x}}_{e_1})}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\psi_{e_1}(\mathring{\tilde{x}}_{e_1,f_1})\prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}}\psi_e(\mathring{\tilde{x}}_{e,f})\overline{\psi_e(\mathring{\tilde{x}}_{e,f'})}}{\|\boldsymbol{\psi}_{e_1}\|^4\prod_{e\in\mathcal{E}\setminus\{e_2\}}\|\boldsymbol{\psi}_e\|^2} \\
&\quad - \frac{\left[\mathring{\tilde{x}}_{e_1}=\mathring{\tilde{x}}_{e_2}\right]\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}}\psi_e(\mathring{\tilde{x}}_{e,f})\overline{\psi_e(\mathring{\tilde{x}}_{e,f'})}}{\|\boldsymbol{\psi}_{e_1}\|^4\prod_{e\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}}\|\boldsymbol{\psi}_e\|^2}
\end{aligned}$$

$$- \frac{\psi_{e_1}(\mathring{\tilde{x}}_{e_1}) \sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}} \in S((e_1,f_1),\mathring{\tilde{x}}_{e_2})} \mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}) \overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_1,f_1'})} \prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}} \psi_e(\mathring{\tilde{x}}_{e,f}) \overline{\psi_e(\mathring{\tilde{x}}_{e,f'})}}{\|\boldsymbol{\psi}_{e_1}\|^4 \prod_{e\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}} \|\boldsymbol{\psi}_e\|^2}$$

$$+ \frac{2\psi_{e_1}(\mathring{\tilde{x}}_{e_1}) \overline{\psi_{e_1}(\mathring{\tilde{x}}_{e_2})} \sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}} \mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}) \prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}} \psi_e(\mathring{\tilde{x}}_{e,f}) \overline{\psi_e(\mathring{\tilde{x}}_{e,f'})}}{\|\boldsymbol{\psi}_{e_1}\|^6 \prod_{e\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}} \|\boldsymbol{\psi}_e\|^2}.$$

When $\boldsymbol{\psi} = \boldsymbol{\psi}_0$ and $e_1 \neq e_2$, we have

$$\frac{\partial}{\partial \psi_{e_2}(\mathring{\tilde{x}}_{e_2})} \overline{\left( \frac{\partial \overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})} \right)} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} = g_{(e_1,f_1'),(e_2,f_2)}(\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2}) - \left[\mathring{\tilde{x}}_{e_2} = \mathring{\tilde{0}}\right] \cdot g_{(e_1,f_1')}(\mathring{\tilde{x}}_{e_1}) - \left[\mathring{\tilde{x}}_{e_1} = \mathring{\tilde{0}}\right] \cdot g_{(e_2,f_2)}(\mathring{\tilde{x}}_{e_2})$$
$$+ \left[\mathring{\tilde{x}}_{e_1} = \mathring{\tilde{0}}, \mathring{\tilde{x}}_{e_2} = \mathring{\tilde{0}}\right] \cdot \mathring{g}(\mathring{\tilde{\boldsymbol{0}}}).$$

When $\boldsymbol{\psi} = \boldsymbol{\psi}_0$ and $e_1 = e_2$, we have

$$\frac{\partial}{\partial \psi_{e_2}(\mathring{\tilde{x}}_{e_2})} \overline{\left( \frac{\partial \overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})} \right)} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} = g_{(e_1,f_1'),(e_1,f_1)}(\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2}) - \left[\mathring{\tilde{x}}_{e_2} = \mathring{\tilde{0}}\right] \cdot g_{(e_1,f_1')}(\mathring{\tilde{x}}_{e_1}) - \left[\mathring{\tilde{x}}_{e_1} = \mathring{\tilde{0}}\right] \cdot g_{(e_1,f_1)}(\mathring{\tilde{x}}_{e_2})$$
$$+ 2\left[\mathring{\tilde{x}}_{e_1} = \mathring{\tilde{0}}, \mathring{\tilde{x}}_{e_2} = \mathring{\tilde{0}}\right] \cdot \mathring{g}(\mathring{\tilde{\boldsymbol{0}}}) - \left[\mathring{\tilde{x}}_{e_1} = \mathring{\tilde{x}}_{e_2}\right] \cdot \mathring{g}(\mathring{\tilde{\boldsymbol{0}}}).$$

We notice that when at least one of $\mathring{\tilde{x}}_{e_1}$ and $\mathring{\tilde{x}}_{e_2}$ equals $\mathring{\tilde{0}}$, we have $\frac{\partial}{\partial \psi_{e_2}(\mathring{\tilde{x}}_{e_2})} \overline{\left( \frac{\partial \overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})} \right)} \Big|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} = 0$. To sum up, for any $e_1, e_2 \in \mathcal{E}_{\mathrm{full}}$, we have

$$\frac{\partial}{\partial \psi_{e_2}(\mathring{\tilde{x}}_{e_2})} \overline{\left( \frac{\partial \overline{\mathring{g}_2(\boldsymbol{\psi})}}{\partial \psi_{e_1}(\mathring{\tilde{x}}_{e_1})} \right)} \Bigg|_{\boldsymbol{\psi}=\boldsymbol{\psi}_0} = \left[\mathring{\tilde{x}}_{e_1} \neq \mathring{\tilde{0}}, \mathring{\tilde{x}}_{e_2} \neq \mathring{\tilde{0}}\right] \cdot \Big( g_{(e_1,f_1'),(e_2,f_2)}(\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2}) - [e_1 = e_2] \cdot \left[\mathring{\tilde{x}}_{e_1} = \mathring{\tilde{x}}_{e_2}\right] \cdot \mathring{g}(\mathring{\tilde{\boldsymbol{0}}}) \Big).$$

With (18), we can obtain (22). The proof for expressions (23)–(25) is similar to the proof of (22). Thus it is omitted here. ■

By definition, the Hessian of $\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})$ at $\boldsymbol{\psi} = \boldsymbol{\psi}_0$, denoted as $H_{\boldsymbol{\psi}}$, equals

$$H_{\boldsymbol{\psi}} = \begin{bmatrix} H_{\boldsymbol{\psi},1} & H_{\boldsymbol{\psi},2} \\ H_{\boldsymbol{\psi},2} & H_{\boldsymbol{\psi},4} \end{bmatrix}.$$

### C. Relating $H_{\boldsymbol{\psi}}$ to the Hessian of $\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})$ when $\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})$ is a function of real arguments $\boldsymbol{w}$

Let $J_{\boldsymbol{w}}$ be the Jacobian matrix of $\boldsymbol{\psi}$ with respect to $\boldsymbol{w}$ and $H_{\boldsymbol{w}}$ be the Hessian of $\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi})$ with respect to $\boldsymbol{w}$. From Section 6 in [31], the matrix $J_{\boldsymbol{w}}$ is invertible and we have

$$J_{\boldsymbol{w}} \cdot J_{\boldsymbol{w}}^{\mathsf{H}} = c \cdot I, \ c \in \mathbb{R}_{>0},$$
$$\boldsymbol{\psi} = J_{\boldsymbol{w}} \cdot \boldsymbol{w},$$
$$H_{\boldsymbol{w}} = J_{\boldsymbol{w}}^{\mathsf{H}} \cdot H_{\boldsymbol{\psi}} \cdot J_{\boldsymbol{w}}.$$

### D. Jacobian Matrix for SPA

Here we compute the Jacobian matrix of the SPA on the DE-NFG after LCT. For every the fixed-point messages of the SPA on the original DE-NFG, there is a corresponding fixed-point messages of the SPA on the DE-NFG after LCT. Given a set of fixed-point messages $\boldsymbol{\mu}$ of the SPA on the original DE-NFG, the corresponding fixed-point messages of the SPA on the DE-NFG after LCT, denoted as $\mathring{\boldsymbol{\mu}} = \{\mathring{\boldsymbol{\mu}}_{e\to f}\}_{e\in\mathcal{E}_{\mathrm{full}}, f\in\partial e}$ with element $\mathring{\boldsymbol{\mu}}_{e\to f} \in \mathbb{C}^{|\mathring{\tilde{\mathcal{X}}}|}$ equaling

$$\mathring{\boldsymbol{\mu}}_{e\to f} = \big(M_{e\to f}\big)^{-1} \cdot \boldsymbol{\mu}_{e\to f},$$
$$\mathring{\boldsymbol{\mu}}_{e\to f'} = \big(M_{e\to f'}\big)^{-1} \cdot \boldsymbol{\mu}_{e\to f'},$$

for all $f \in \partial e$, $e \in \mathcal{E}_{\mathrm{full}}$.

In the SPA, the terms $\mathring{\mu}_{e\to f}(\mathring{\tilde{0}})$ and $\mathring{\mu}_{e\to f'}(\mathring{\tilde{0}})$ are normalized to be 1 for all $e=(f,f')\in\mathcal{E}_{\text{full}}$ at each iteration. Suppose that we are at the $n$-th iteration and we want to obtain the $(n+1)$-th message $\mathring{\boldsymbol{\mu}}_{e_1\to f_1}^{(n+1)}\in\mathbb{C}^{|\mathring{\mathcal{X}}_{e_1}|}$ with index $\mathring{\tilde{x}}_{e_1}\in\mathring{\mathcal{X}}_{e_1}$ using SPA, we have

$$\mathring{\mu}_{e_1\to f_1}^{(n+1)}(\mathring{\tilde{x}}_{e_1})=\frac{\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\text{full}}},\mathcal{F}\in S((e_1,f_1'),\mathring{\tilde{x}}_{e_1})}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\text{full}},\mathcal{F}})\mathring{\mu}_{e_1\to f_1}^{(n)}(\mathring{\tilde{x}}_{e_1,f_1})\prod_{e=(f,f')\in\mathcal{E}_{\text{full}}\setminus\{e_1\}}\mathring{\mu}_{e\to f}^{(n)}(\mathring{\tilde{x}}_{e,f})\mathring{\mu}_{e\to f'}^{(n)}(\mathring{\tilde{x}}_{e,f'})}{\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\text{full}}},\mathcal{F}\in S((e_1,f_1'),\mathring{\tilde{0}})}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\text{full}},\mathcal{F}})\mathring{\mu}_{e_1\to f_1}^{(n)}(\mathring{\tilde{x}}_{e_1,f_1})\prod_{e=(f,f')\in\mathcal{E}_{\text{full}}\setminus\{e_1\}}\mathring{\mu}_{e\to f}^{(n)}(\mathring{\tilde{x}}_{e,f})\mathring{\mu}_{e\to f'}^{(n)}(\mathring{\tilde{x}}_{e,f'})}.$$

Let us define $\mathring{\boldsymbol{\mu}}^{(n)}:=\{\mathring{\boldsymbol{\mu}}_{e\to f}^{(n)}\}_{e\in\mathcal{E}_{\text{full}}\cap\partial f,f\in\mathcal{F}}$. Recall that $\mathring{\boldsymbol{\mu}}_0$ is defined in Theorem 16, t he Jacobian matrix $J_{\boldsymbol{\mu}}$ of the SPA at $\mathring{\boldsymbol{\mu}}^{(n)}=\mathring{\boldsymbol{\mu}}_0$ equals

$$J_{\boldsymbol{\mu}}=\begin{bmatrix}J_{\boldsymbol{\mu},1} & J_{\boldsymbol{\mu},2}\\ J_{\boldsymbol{\mu},3} & J_{\boldsymbol{\mu},4}\end{bmatrix},$$

where

$$J_{\boldsymbol{\mu},1}(\phi_{e_1}(\mathring{\tilde{x}}_{e_1}),\phi_{e_2}(\mathring{\tilde{x}}_{e_2})):=\left.\frac{\partial\mathring{\mu}_{e_1\to f_1}^{(n+1)}(\mathring{\tilde{x}}_{e_1})}{\partial\mathring{\mu}_{e_2\to f_2}^{(n)}(\mathring{\tilde{x}}_{e_2})}\right|_{\mathring{\boldsymbol{\mu}}^{(n)}=\boldsymbol{\mu}_0},$$

$$J_{\boldsymbol{\mu},2}(\phi_{e_1}(\mathring{\tilde{x}}_{e_1}),\phi_{e_2}(\mathring{\tilde{x}}_{e_2})):=\left.\frac{\partial\mathring{\mu}_{e_1\to f_1}^{(n+1)}(\mathring{\tilde{x}}_{e_1})}{\partial\mathring{\mu}_{e_2\to f_2'}^{(n)}(\mathring{\tilde{x}}_{e_2})}\right|_{\mathring{\boldsymbol{\mu}}^{(n)}=\boldsymbol{\mu}_0},$$

$$J_{\boldsymbol{\mu},3}(\phi_{e_1}(\mathring{\tilde{x}}_{e_1}),\phi_{e_2}(\mathring{\tilde{x}}_{e_2})):=\left.\frac{\partial\mathring{\mu}_{e_1\to f_1'}^{(n+1)}(\mathring{\tilde{x}}_{e_1})}{\partial\mathring{\mu}_{e_2\to f_2}^{(n)}(\mathring{\tilde{x}}_{e_2})}\right|_{\mathring{\boldsymbol{\mu}}^{(n)}=\boldsymbol{\mu}_0},$$

$$J_{\boldsymbol{\mu},4}(\phi_{e_1}(\mathring{\tilde{x}}_{e_1}),\phi_{e_2}(\mathring{\tilde{x}}_{e_2})):=\left.\frac{\partial\mathring{\mu}_{e_1\to f_1'}^{(n+1)}(\mathring{\tilde{x}}_{e_1})}{\partial\mathring{\mu}_{e_2\to f_2'}^{(n)}(\mathring{\tilde{x}}_{e_2})}\right|_{\mathring{\boldsymbol{\mu}}^{(n)}=\boldsymbol{\mu}_0},$$

for all $\mathring{\tilde{x}}_{e_1}\in\mathring{\mathcal{X}}_{e_1}$, $\mathring{\tilde{x}}_{e_2}\in\mathring{\mathcal{X}}_{e_2}$, $e_1=(f_1,f_1')$, $e_2=(f_2,f_2')\in\mathcal{E}_{\text{full}}$.

**Lemma 27.** *We have*

$$J_{\boldsymbol{\mu},1}(\phi_{e_1}(\mathring{\tilde{x}}_{e_1}),\phi_{e_2}(\mathring{\tilde{x}}_{e_2}))=\left[\mathring{\tilde{x}}_{e_1}\neq\mathring{\tilde{0}},\mathring{\tilde{x}}_{e_2}\neq\mathring{\tilde{0}}\right]\cdot\frac{g_{(e_1,f_1'),(e_2,f_2)}(\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2})}{\mathring{g}(\mathring{\tilde{\mathbf{0}}})},\tag{28}$$

$$J_{\boldsymbol{\mu},2}(\phi_{e_1}(\mathring{\tilde{x}}_{e_1}),\phi_{e_2}(\mathring{\tilde{x}}_{e_2}))=[e_1\neq e_2]\cdot\left[\mathring{\tilde{x}}_{e_1}\neq\mathring{\tilde{0}},\mathring{\tilde{x}}_{e_2}\neq\mathring{\tilde{0}}\right]\cdot\frac{g_{(e_1,f_1'),(e_2,f_2')}(\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2})}{\mathring{g}(\mathring{\tilde{\mathbf{0}}})},\tag{29}$$

$$J_{\boldsymbol{\mu},3}(\phi_{e_1}(\mathring{\tilde{x}}_{e_1}),\phi_{e_2}(\mathring{\tilde{x}}_{e_2}))=[e_1\neq e_2]\cdot\left[\mathring{\tilde{x}}_{e_1}\neq\mathring{\tilde{0}},\mathring{\tilde{x}}_{e_2}\neq\mathring{\tilde{0}}\right]\cdot\frac{g_{(e_1,f_1),(e_2,f_2)}(\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2})}{\mathring{g}(\mathring{\tilde{\mathbf{0}}})},\tag{30}$$

$$J_{\boldsymbol{\mu},4}(\phi_{e_1}(\mathring{\tilde{x}}_{e_1}),\phi_{e_2}(\mathring{\tilde{x}}_{e_2}))=\left[\mathring{\tilde{x}}_{e_1}\neq\mathring{\tilde{0}},\mathring{\tilde{x}}_{e_2}\neq\mathring{\tilde{0}}\right]\cdot\frac{g_{(e_1,f_1),(e_2,f_2')}(\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2})}{\mathring{g}(\mathring{\tilde{\mathbf{0}}})},\tag{31}$$

*for all* $\mathring{\tilde{x}}_{e_1}\in\mathring{\mathcal{X}}_{e_1}$, $\mathring{\tilde{x}}_{e_2}\in\mathring{\mathcal{X}}_{e_2}$, $e_1=(f_1,f_1')$, $e_2=(f_2,f_2')\in\mathcal{E}_{\text{full}}$.

*Proof.* Suppose $e_1\neq e_2$, the Wirtinger derivatives of $\mathring{\mu}_{e_1\to f_1}^{(n+1)}(\mathring{\tilde{x}}_{e_1})$ with respect to $\mathring{\mu}_{e_2\to f_2}^{(n)}(\mathring{\tilde{x}}_{e_2})$ are

$$\begin{aligned}\frac{\partial\mathring{\mu}_{e_1\to f_1}^{(n+1)}(\mathring{\tilde{x}}_{e_1})}{\partial\mathring{\mu}_{e_2\to f_2}^{(n)}(\mathring{\tilde{x}}_{e_2})}=&\Bigg(\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\text{full}}},\mathcal{F}\in S((e_1,f_1'),(e_2,f_2),\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2})}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\text{full}},\mathcal{F}})\mathring{\mu}_{e_1\to f_1}^{(n)}(\mathring{\tilde{x}}_{e_1,f_1})\mathring{\mu}_{e_2\to f_2'}^{(n)}(\mathring{\tilde{x}}_{e_2,f_2'})\\&\cdot\prod_{e=(f,f')\in\mathcal{E}_{\text{full}}\setminus\{e_1,e_2\}}\mathring{\mu}_{e\to f}^{(n)}(\mathring{\tilde{x}}_{e,f})\mathring{\mu}_{e\to f'}^{(n)}(\mathring{\tilde{x}}_{e,f'})\Bigg)\\&\cdot\Bigg(\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\text{full}}},\mathcal{F}\in S((e_1,f_1'),\mathring{\tilde{0}})}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\text{full}},\mathcal{F}})\mathring{\mu}_{e_1\to f_1}^{(n)}(\mathring{\tilde{x}}_{e_1,f_1})\prod_{e=(f,f')\in\mathcal{E}_{\text{full}}\setminus\{e_1\}}\mathring{\mu}_{e\to f}^{(n)}(\mathring{\tilde{x}}_{e,f})\mathring{\mu}_{e\to f'}^{(n)}(\mathring{\tilde{x}}_{e,f'})\Bigg)^{-1}\\&-\frac{\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\text{full}}},\mathcal{F}\in S((e_1,f_1'),\mathring{\tilde{x}}_{e_1})}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\text{full}},\mathcal{F}})\mathring{\mu}_{e_1\to f_1}^{(n)}(\mathring{\tilde{x}}_{e_1,f_1})\prod_{e=(f,f')\in\mathcal{E}_{\text{full}}\setminus\{e_1\}}\mathring{\mu}_{e\to f}^{(n)}(\mathring{\tilde{x}}_{e,f})\mathring{\mu}_{e\to f'}^{(n)}(\mathring{\tilde{x}}_{e,f'})}{\left(\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\text{full}}},\mathcal{F}\in S((e_1,f_1'),\mathring{\tilde{0}})}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\text{full}},\mathcal{F}})\mathring{\mu}_{e_1\to f_1}^{(n)}(\mathring{\tilde{x}}_{e_1,f_1})\prod_{e=(f,f')\in\mathcal{E}_{\text{full}}\setminus\{e_1\}}\mathring{\mu}_{e\to f}^{(n)}(\mathring{\tilde{x}}_{e,f})\mathring{\mu}_{e\to f'}^{(n)}(\mathring{\tilde{x}}_{e,f'})\right)^2}\end{aligned}$$

$$\cdot\left(\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}\in S((e_1,f_1'),(e_2,f_2),\mathring{\tilde{0}},\mathring{\tilde{x}}_{e_2})}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\mathring{\mu}^{(n)}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_1,f_1})\mathring{\mu}^{(n)}_{e_2\to f_2'}(\mathring{\tilde{x}}_{e_2,f_2'})\right.$$
$$\left.\cdot\prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1,\ e_2\}}\mathring{\mu}^{(n)}_{e\to f}(\mathring{\tilde{x}}_{e,f})\mathring{\mu}^{(n)}_{e\to f'}(\mathring{\tilde{x}}_{e,f'})\right), \tag{32}$$

$$\frac{\partial\mathring{\mu}^{(n+1)}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_1})}{\partial\overline{\mathring{\mu}^{(n)}_{e_2\to f_2}(\mathring{\tilde{x}}_{e_2})}}=0, \tag{33}$$

for all $\mathring{\tilde{x}}_{e_1}\in\mathring{\mathcal{X}}_{e_1},\ \mathring{\tilde{x}}_{e_2}\in\mathring{\mathcal{X}}_{e_2}$. When $e_1=e_2$, we have

$$\frac{\partial\mathring{\mu}^{(n+1)}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_1})}{\partial\mathring{\mu}^{(n)}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_2})}=\frac{\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}\in S((e_1,f_1'),(e_1,f_1),\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2})}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}}\mathring{\mu}^{(n)}_{e\to f}(\mathring{\tilde{x}}_{e,f})\mathring{\mu}^{(n)}_{e\to f'}(\mathring{\tilde{x}}_{e,f'})}{\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}\in S((e_1,f_1'),\mathring{\tilde{0}})}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\mathring{\mu}^{(n)}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_1,f_1})\prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}}\mathring{\mu}^{(n)}_{e\to f}(\mathring{\tilde{x}}_{e,f})\mathring{\mu}^{(n)}_{e\to f'}(\mathring{\tilde{x}}_{e,f'})}$$
$$-\frac{\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}\in S((e_1,f_1'),\mathring{\tilde{x}}_{e_1})}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\mathring{\mu}^{(n)}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_1,f_1})\prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}}\mathring{\mu}^{(n)}_{e\to f}(\mathring{\tilde{x}}_{e,f})\mathring{\mu}^{(n)}_{e\to f'}(\mathring{\tilde{x}}_{e,f'})}{\left(\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}\in S((e_1,f_1'),\mathring{\tilde{0}})}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\mathring{\mu}^{(n)}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_1,f_1})\prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}}\mathring{\mu}^{(n)}_{e\to f}(\mathring{\tilde{x}}_{e,f})\mathring{\mu}^{(n)}_{e\to f'}(\mathring{\tilde{x}}_{e,f'})\right)^2}$$
$$\cdot\left(\sum_{\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}}\in S((e_1,f_1'),(e_1,f_1),\mathring{\tilde{0}},\mathring{\tilde{x}}_{e_2})}\mathring{g}(\mathring{\tilde{\boldsymbol{x}}}_{\mathcal{E}_{\mathrm{full}},\mathcal{F}})\prod_{e=(f,f')\in\mathcal{E}_{\mathrm{full}}\setminus\{e_1\}}\mathring{\mu}^{(n)}_{e\to f}(\mathring{\tilde{x}}_{e,f})\mathring{\mu}^{(n)}_{e\to f'}(\mathring{\tilde{x}}_{e,f'})\right),$$

$$\frac{\partial\mathring{\mu}^{(n+1)}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_1})}{\partial\overline{\mathring{\mu}^{(n)}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_2})}}=0,$$

for all $\mathring{\tilde{x}}_{e_1}\in\mathring{\mathcal{X}}_{e_1},\ \mathring{\tilde{x}}_{e_2}\in\mathring{\mathcal{X}}_{e_2}$. Evaluating (32) at $\mathring{\boldsymbol{\mu}}^{(n)}=\boldsymbol{\mu}_0$ gives

$$\left.\frac{\partial\mathring{\mu}^{(n+1)}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_1})}{\partial\mathring{\mu}^{(n)}_{e_2\to f_2}(\mathring{\tilde{x}}_{e_2})}\right|_{\mathring{\boldsymbol{\mu}}^{(n)}=\boldsymbol{\mu}_0}=\frac{g_{(e_1,f_1'),(e_2,f_2)}(\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2})}{\mathring{g}(\mathring{\tilde{\mathbf{0}}})}-\frac{g_{(e_1,f_1')}(\mathring{\tilde{x}}_{e_1})g_{(e_2,f_2)}(\mathring{\tilde{x}}_{e_2})}{(\mathring{g}(\mathring{\tilde{\mathbf{0}}}))^2}, \tag{34}$$

for all $\mathring{\tilde{x}}_{e_1}\in\mathring{\mathcal{X}}_{e_1},\ \mathring{\tilde{x}}_{e_2}\in\mathring{\mathcal{X}}_{e_2}$. When any one of $\mathring{\tilde{x}}_{e_1},\ \mathring{\tilde{x}}_{e_2}$ equals $\mathring{\tilde{0}}$, expression (34) equals zero. As a result, it holds

$$\left.\frac{\partial\mathring{\mu}^{(n+1)}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_1})}{\partial\mathring{\mu}^{(n)}_{e_2\to f_2}(\mathring{\tilde{x}}_{e_2})}\right|_{\mathring{\boldsymbol{\mu}}^{(n)}=\boldsymbol{\mu}_0}=\left[\mathring{\tilde{x}}_{e_1}\neq\mathring{\tilde{0}},\mathring{\tilde{x}}_{e_2}\neq\mathring{\tilde{0}}\right]\cdot\frac{g_{(e_1,f_1'),(e_2,f_2)}(\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2})}{\mathring{g}(\mathring{\tilde{\mathbf{0}}})},$$

for all $\mathring{\tilde{x}}_{e_1}\in\mathring{\mathcal{X}}_{e_1},\ \mathring{\tilde{x}}_{e_2}\in\mathring{\mathcal{X}}_{e_2}$. For $e_1=e_2$, we have

$$\left.\frac{\partial\mathring{\mu}^{(n+1)}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_1})}{\partial\mathring{\mu}^{(n)}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_2})}\right|_{\mathring{\boldsymbol{\mu}}^{(n)}=\boldsymbol{\mu}_0}=\frac{g_{(e_1,f_1'),(e_1,f_1)}(\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2})}{\mathring{g}(\mathring{\tilde{\mathbf{0}}})}-\frac{g_{(e_1,f_1')}(\mathring{\tilde{x}}_{e_1})g_{(e_1,f_1)}(\mathring{\tilde{x}}_{e_2})}{(\mathring{g}(\mathring{\tilde{\mathbf{0}}}))^2}, \tag{35}$$

for all $\mathring{\tilde{x}}_{e_1}\in\mathring{\mathcal{X}}_{e_1},\ \mathring{\tilde{x}}_{e_2}\in\mathring{\mathcal{X}}_{e_2}$. Similarly, when any one of $\mathring{\tilde{x}}_{e_1},\ \mathring{\tilde{x}}_{e_2}$ equals $\mathring{\tilde{0}}$, expression (35) equals zero. Thus

$$\left.\frac{\partial\mathring{\mu}^{(n+1)}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_1})}{\partial\mathring{\mu}^{(n)}_{e_1\to f_1}(\mathring{\tilde{x}}_{e_2})}\right|_{\mathring{\boldsymbol{\mu}}^{(n)}=\boldsymbol{\mu}_0}=\left[\mathring{\tilde{x}}_{e_1}\neq\mathring{\tilde{0}},\mathring{\tilde{x}}_{e_2}\neq\mathring{\tilde{0}}\right]\cdot\frac{g_{(e_1,f_1'),(e_1,f_1)}(\mathring{\tilde{x}}_{e_1},\mathring{\tilde{x}}_{e_2})}{\mathring{g}(\mathring{\tilde{\mathbf{0}}})},$$

for all $\mathring{\tilde{x}}_{e_1}\in\mathring{\mathcal{X}}_{e_1},\ \mathring{\tilde{x}}_{e_2}\in\mathring{\mathcal{X}}_{e_2}$. The proof of expressions (30)–(31) is similar to the proof for expression (30) and thus it is omitted here. ∎

### *E. Relating $J_{\boldsymbol{\mu}}$ to $H_{\boldsymbol{\psi}}$*

**Lemma 28.** *Suppose that $\boldsymbol{\mu}=\boldsymbol{\mu}_0$ is a stable fixed-point of the SPA, then the eigenvalues of the matrix $-H_{\boldsymbol{\psi}}$ have non-negative real part.*

*Proof.* Since $\boldsymbol{\mu}=\boldsymbol{\mu}_0$ is a stable fixed-point of the SPA, by stability theory, the eigenvalues of the matrix $J_{\boldsymbol{\mu}}$, denoted as $\{\lambda_{\mathrm{J},i}\}_{i=0}^{2\sum_{e\in\mathcal{E}_{\mathrm{full}}}|\mathring{\mathcal{X}}_e|-1}$, satisfy $|\lambda_{\mathrm{J},i}|<1$ and $\mathrm{Re}(1-\lambda_{\mathrm{J},i})>0$ for all $i$. Let us define $I'$ to be a matrix with size $2\sum_{e\in\mathcal{E}_{\mathrm{full}}}|\mathring{\mathcal{X}}_e|\times2\sum_{e\in\mathcal{E}_{\mathrm{full}}}|\mathring{\mathcal{X}}_e|$ such that

$$I':=\begin{bmatrix}I_1 & \mathbf{0}\\ \mathbf{0} & I_1\end{bmatrix},$$

where $I_1 \in \mathbb{R}_{\geq 0}^{\sum_{e\in\mathcal{E}_{\mathrm{full}}} |\mathring{\mathcal{X}}_e| \times \sum_{e\in\mathcal{E}_{\mathrm{full}}} |\mathring{\mathcal{X}}_e|}$ has entry

$$I_1(\phi_{e_1}(\mathring{\tilde{x}}_{e_1}), \phi_{e_2}(\mathring{\tilde{x}}_{e_2})) := \left[\mathring{\tilde{x}}_{e_1} \neq \mathring{\tilde{0}}, \mathring{\tilde{x}}_{e_2} \neq \mathring{\tilde{0}}\right] \cdot [e_1 = e_2] \cdot \left[\mathring{\tilde{x}}_{e_1} = \mathring{\tilde{x}}_{e_2}\right],$$

for all $\mathring{\tilde{x}}_{e_1} \in \mathring{\mathcal{X}}_{e_1}$, $\mathring{\tilde{x}}_{e_2} \in \mathring{\mathcal{X}}_{e_2}$, $e_1 = (f_1, f_1')$, $e_2 = (f_2, f_2') \in \mathcal{E}_{\mathrm{full}}$. Comparing the entries in the matrix $I' - J_{\boldsymbol{\mu}}$ with the entries in matrix $-H_{\boldsymbol{\psi}}^T$, one obtains that $I' - J_{\boldsymbol{\mu}} = -H_{\boldsymbol{\psi}}^T$. With the previous analysis, the eigenvalues of $I' - J_{\boldsymbol{\mu}}$ also have non-negative real part. Thus we can derive that both $-H_{\boldsymbol{\psi}}^T$ and $-H_{\boldsymbol{\psi}}$ have eigenvalues with non-negative real part. ■

Let us define matrices $H_{\boldsymbol{\psi},\mathrm{re}}$ and $H_{\boldsymbol{\psi},\mathrm{im}}$ to be

$$H_{\boldsymbol{\psi},\mathrm{re}} := \begin{bmatrix} H_{\boldsymbol{\psi},1,\mathrm{re}} & H_{\boldsymbol{\psi},2,\mathrm{re}} \\ H_{\boldsymbol{\psi},2,\mathrm{re}} & H_{\boldsymbol{\psi},4,\mathrm{re}} \end{bmatrix},$$

$$H_{\boldsymbol{\psi},\mathrm{im}} := \begin{bmatrix} H_{\boldsymbol{\psi},1,\mathrm{im}} & H_{\boldsymbol{\psi},2,\mathrm{im}} \\ H_{\boldsymbol{\psi},2,\mathrm{im}} & H_{\boldsymbol{\psi},4,\mathrm{im}} \end{bmatrix}.$$

By definition, $H_{\boldsymbol{\psi},\mathrm{re}}$ and $H_{\boldsymbol{\psi},\mathrm{im}}$ are the Hessian matrices of $\mathrm{Re}(\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi}))$ and $\mathrm{Im}(\mathring{g}_{\mathrm{SST}}(\boldsymbol{\psi}))$, respectively.

**Theorem 29.** *Matrix $-H_{\boldsymbol{\psi},\mathrm{re}}$ is positive-semidefinite.*

*Proof.* By definition and [31], we know that $-H_{\boldsymbol{\psi},\mathrm{im}}$ and $-H_{\boldsymbol{\psi},\mathrm{re}}$ are Hermitian matrices. Thus matrix $-iH_{\boldsymbol{\psi},\mathrm{im}}$ is a skew-Hermitian matrix. Let $\boldsymbol{v}_i \in \mathbb{C}^{2\sum_{e\in\mathcal{E}_{\mathrm{full}}} |\mathring{\mathcal{X}}_e|}$ be a right eigenvector of $-H_{\boldsymbol{\psi}} = -(H_{\boldsymbol{\psi},\mathrm{re}} + iH_{\boldsymbol{\psi},\mathrm{im}})$ with eigenvalue $\lambda_i \in \mathbb{C}$ for all $i = 0, \ldots, 2\sum_{e\in\mathcal{E}_{\mathrm{full}}} |\mathring{\mathcal{X}}_e| - 1$. We have

$$\begin{aligned} -H_{\boldsymbol{\psi}}\boldsymbol{v}_i &= \lambda_i \boldsymbol{v}_i, \\ -\boldsymbol{v}_i^{\mathsf{H}} H_{\boldsymbol{\psi}}^{\mathsf{H}} &= \overline{\lambda_i} \boldsymbol{v}_i^{\mathsf{H}}, \\ \lambda_i \|\boldsymbol{v}_i\|^2 &= \boldsymbol{v}_i^{\mathsf{H}} \cdot (-H_{\boldsymbol{\psi}}) \cdot \boldsymbol{v}_i \\ &= \overline{\boldsymbol{v}_i^{\mathsf{H}} \cdot (-H_{\boldsymbol{\psi}})^{\mathsf{H}} \cdot \boldsymbol{v}_i}. \end{aligned}$$

Therefore,

$$\begin{aligned} \mathrm{Re}(\lambda_i)\|\boldsymbol{v}_i\|^2 &= \frac{1}{2}\boldsymbol{v}_i^{\mathsf{H}} \cdot \Big(-H_{\boldsymbol{\psi}} - (H_{\boldsymbol{\psi}})^{\mathsf{H}}\Big) \cdot \boldsymbol{v}_i \\ &= \frac{1}{2}\boldsymbol{v}_i^{\mathsf{H}} \cdot \Big(-H_{\boldsymbol{\psi},\mathrm{re}} - iH_{\boldsymbol{\psi},\mathrm{im}} - (H_{\boldsymbol{\psi},\mathrm{re}})^{\mathsf{H}} - (iH_{\boldsymbol{\psi},\mathrm{im}})^{\mathsf{H}}\Big) \cdot \boldsymbol{v}_i \\ &\overset{(a)}{=} \boldsymbol{v}_i^{\mathsf{H}} \cdot (-H_{\boldsymbol{\psi},\mathrm{re}}) \cdot \boldsymbol{v}_i \geq 0, \end{aligned} \tag{36}$$

where at step $(a)$ we have used the condition that matrix $iH_{\boldsymbol{\psi},\mathrm{im}}$ is a skew-Hermitian matrix. Note that the inequality $\boldsymbol{v}_i^{\mathsf{H}} \cdot \big(-H_{\boldsymbol{\psi},\mathrm{re}}\big) \cdot \boldsymbol{v}_i \geq 0$ in (36) holds for all $i$. Thus $-H_{\boldsymbol{\psi},\mathrm{re}}$ is positive-semidefinite. ■

## References


[1] F. R. Kschischang, B. J. Frey, and H.-. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

[2] G. D. Forney, "Codes on graphs: normal realizations," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 520–548, Feb. 2001.

[3] H.-A. Loeliger, "An introduction to factor graphs," *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 28–41, Jan. 2004.

[4] M. Mézard and A. Montanari, *Information, Physics and Computation*. Oxford, U.K.: Oxford Univ. Press, 2009.

[5] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2008.

[6] H. Wymeersch, *Iterative Receiver Design*. Cambridge, U.K.: Cambridge Univ. Press, 2007.

[7] J. S. Yedidia, W. T. Freeman, and Y. Weiss, "Constructing free-energy approximations and generalized belief propagation algorithms," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2282–2312, Jul. 2005.

[8] M. Chertkov and V. Y. Chernyak, "Loop series for discrete statistical models on graphs," *J. Stat. Mech: Theory Exp.*, vol. 2006, no. 06, pp. 6009–6009, Jun. 2006.

[9] V. Y. Chernyak and M. Chertkov, "Loop calculus and belief propagation for q-ary alphabet: Loop tower," in *Proc. IEEE Int. Symp. Information Theory*, Nice, France, Jun. 2007, pp. 316–320.

[10] R. Mori, "Loop calculus for non-binary alphabets using concepts from information geometry," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1887–1904, Apr. 2015.

[11] R. Koetter, W.-C. W. Li, P. O. Vontobel, and J. L. Walker, "Characterizations of pseudo-codewords of (low-density) parity-check codes," *Adv. Math.*, vol. 213, no. 1, pp. 205–229, 2007.

[12] P. O. Vontobel, "Counting in graph covers: A combinatorial characterization of the Bethe entropy function," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 6018–6048, Sep. 2013.

[13] N. Ruozzi, "The Bethe partition function of log-supermodular graphical models," in *Proc. Neural Information Processing Systems (NIPS)*, Lake Tahoe, NV, Dec. 2012, pp. 117–125.

[14] S. Shams, N. Ruozzi, and P. Csikvári, "Counting homomorphisms in bipartite graphs," in *Proc. IEEE Int. Symp. Information Theory*, Jul. 2019, pp. 1487–1491.

[15] N. Ruozzi, J. Thaler, and S. Tatikonda, "Graph covers and quadratic minimization," in *Proc. and Computing (Allerton) 2009 47th Annual Allerton Conf. Communication, Control*, Sep. 2009, pp. 1590–1596.

[16] E. B. Sudderth, M. J. Wainwright, and A. S. Willsky, "Loop series and Bethe variational bounds in attractive graphical models," in *Proc. Neural Inf. Proc. Sys. Conf.*, Vancouver, Canada, Dec. 3–8 2007.

[17] H.-A. Loeliger and P. O. Vontobel, "A factor-graph representation of probabilities in quantum mechanics," in *Proc. IEEE Int Symp. Information Theory*, Cambridge, MA, USA, Jul. 2012, pp. 656–660.

[18] ——, "Factor graphs for quantum probabilities," *IEEE Trans. Inf. Theory*, vol. 63, no. 9, pp. 5642–5665, Sep. 2017.

[19] R. Mori, "Holographic transformation, belief propagation and loop calculus for generalized probabilistic theories," in *Proc. IEEE Int Symp. Information Theory*, Hong Kong, China, Jun. 2015, pp. 1099–1103.

[20] M. X. Cao and P. O. Vontobel, "Double-edge factor graphs: Definition, properties, and examples," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Kaohsiung, Taiwan, Nov. 2017, pp. 136–140.

[21] E. Loh Jr., J. Gubernatis, R. Scalettar, S. White, D. Scalapino, and R. Sugar, "Sign problem in the numerical simulation of many-electron systems," *Phys. Rev. B, Condens. Matter*, vol. 41, no. 13, p. 9301, May 1990.

[22] A. Al-Bashabsheh and Y. Mao, "Normal factor graphs and holographic transformations," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 752–763, Feb. 2011.

[23] T. Heskes, "Stable fixed points of loopy belief propagation are local minima of the Bethe free energy," in *Proc. Neural Information Processing Systems (NIPS)*, Vancouver, Canada, Dec. 2003, pp. 359–366.

[24] Y. Watanabe, "Discrete geometric analysis of message passing algorithm on graphs," Ph.D. dissertation, Dept. of Statistical Sciences, School of Multidisciplinary Science, Graduate University for Advanced Studies, Mar. 2010. [Online]. Available: https://arxiv.org/pdf/1004.4942.pdf

[25] H. M. Stark and A. A. Terras, "Zeta functions of finite graphs and coverings," *Adv. in Math.*, vol. 121, no. 1, pp. 124–165, Jul. 1996.

[26] M.-D. Choi, "Completely positive linear maps on complex matrices," *Linear Algebra Appl.*, vol. 10, no. 3, pp. 285 – 290, 1975.

[27] C. J. Wood, J. D. Biamonte, and D. G. Cory, "Tensor networks and graphical calculus for open quantum systems," *Quantum Info. Comput.*, vol. 15, no. 9-10, pp. 759–811, Jul. 2015.

[28] P. O. Vontobel, "Analysis of double covers of factor graphs," in *Proc. Int. Conf. Sig. Proc. and Comm.*, Bangalore, India, June 12–15 2016, pp. 1–5.

[29] G. B. Folland, "How to integrate a polynomial over a sphere," *Amer. Math. Monthly*, vol. 108, no. 5, pp. 446–448, 2001.

[30] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2017.

[31] K. Kreutz-Delgado, "The complex gradient operator and the CR-calculus." [Online]. Available: https://arxiv.org/abs/0906.4835