Critical point representation of the mutual information in the sparse stochastic block model

Tomas Dominguez[*]   Jean-Christophe Mourrat[†]

Abstract

We consider the problem of recovering the community structure in the stochastic block model. We aim to describe the mutual information between the observed network and the actual community structure as the number of nodes diverges while the average degree of a given node remains bounded. Our main contribution is a representation of the limit of this quantity, assuming it exists, as an explicit functional evaluated at a critical point of that functional. While we mostly focus on the two-community setting for clarity, we expect our method to be robust to model generalizations. We also present an example involving four communities where we show the invalidity of a plausible candidate variational formula for this limit.

1   Introduction

1.1   Informal summary of the main results

We consider a problem of community detection in the context of the stochastic block model. We focus on the case with two communities, where the community structure is encoded by a vector

$$\sigma^* := (\sigma_1^*, \ldots, \sigma_N^*) \in \Sigma_N := \{-1, +1\}^N. \tag{1.1}$$

We understand that individuals i and j belong to the same community if and only if $\sigma_i^* = \sigma_j^*$. The labels $(\sigma_i^*)_{i \leq N}$ are sampled independently from a Bernoulli distribution P^* with probability of success $p \in (0,1)$ and expectation $\overline{m}$,

$$p := P^*\{1\} = \mathbb{P}\{\sigma_i^* = 1\} \quad \text{and} \quad \overline{m} := \mathbb{E}\,\sigma_i^* = 2p - 1. \tag{1.2}$$

The assignment vector σ^* is thus distributed according to the product law

$$\sigma^* \sim P_N^* := (P^*)^{\otimes N}. \tag{1.3}$$

We fix parameters $c > 0$ and $\Delta \in (-c, c) \setminus \{0\}$. The case when $\Delta = 0$ would be trivial and we prefer to exclude it. Given the community structure σ^*, we sample a random undirected graph $\mathbf{G}_N := (G_{ij})_{i,j \leq N}$ with vertex set $\{1, \ldots, N\}$ by stipulating that an edge between node i and node j is present with conditional probability

$$\mathbb{P}\{G_{ij} = 1 \mid \sigma^*\} := \frac{c + \Delta\sigma_i^*\sigma_j^*}{N}, \tag{1.4}$$

---
[*]DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TORONTO, TOMAS.DOMINGUEZCHIOZZA@MAIL.UTORONTO.CA

[†]DEPARTMENT OF MATHEMATICS, ENS LYON AND CNRS, JEAN-CHRISTOPHE.MOURRAT@ENS-LYON.FR

and that this sampling is performed independently for each pair $i \neq j \in \{1, \ldots, N\}$. In words, we create a link between two vertices with probability $(c + \Delta)/N$ if those two vertices belong to the same community; and we create this link with probability $(c - \Delta)/N$ otherwise. In the case when $\Delta > 0$, it is more likely for an edge to be present between nodes belonging to the same community, and the model is called *assortative*; otherwise it is called *disassortative*. The inference task that we consider is that of recovering the community structure σ^* given the observation of the graph $\mathbf{G}_N$. We focus in particular on characterizing the mutual information between the graph $\mathbf{G}_N$ and the community structure σ^*, in the limit of large N.

We denote by $\mathscr{M}_+$ the space of non-negative measures over $[-1, 1]$ with finite total mass, and by $\mathscr{M}_p$ the set of probability measures with mean $\overline{m} = 2p - 1$,

$$\mathscr{M}_p := \left\{ \nu \in \Pr[-1, 1] \mid \int_{-1}^{1} x \, d\nu(x) = \overline{m} \right\}. \tag{1.5}$$

Our main results will be stated in terms of a fixed point of the functional $\Gamma_{t, \mu} : \mathscr{M}_p \to \mathscr{M}_p$, which is defined for each $t \geq 0$ and $\mu \in \mathscr{M}_+$ according to the following procedure. Fixing $\nu \in \mathscr{M}_p$, we sample ε^* according to P^*, and conditionally on ε^*, we let $\Pi_{t, \mu}(\nu)$ be a Poisson point process with intensity measure $(c + \Delta \varepsilon^* x) \, d(\mu + t\nu)(x)$; we then set

$$\Gamma_{t, \mu}(\nu) := \mathrm{Law} \left(\frac{\int_{\Sigma_1} \varepsilon e^{-\varepsilon \overline{m}} \prod_{x \in \Pi_{t, \mu}(\nu)} (c + \Delta \varepsilon x) \, dP^*(\varepsilon)}{\int_{\Sigma_1} e^{-\varepsilon \overline{m}} \prod_{x \in \Pi_{t, \mu}(\nu)} (c + \Delta \varepsilon x) \, dP^*(\varepsilon)} \right). \tag{1.6}$$

In other words, the quantity between the large parentheses in (1.6) is a random variable, as it depends on the realization of ε^* and $\Pi_{t, \mu}(\nu)$, and we define $\Gamma_{t, \mu}(\nu)$ to be the law of this random variable.

Let $(\sigma^\ell)_{\ell \geq 1} = ((\sigma_i^\ell)_{i \leq N})_{\ell \geq 1}$ be independent random variables that are sampled according to the conditional law of σ^* given the observation of $\mathbf{G}_N$. Our first main result states that, up to a perturbation of the inference problem that does not change the asymptotic behavior of the mutual information, and up to the extraction of a subsequence in N, the asymptotic law of $(\sigma_i^\ell)_{i \leq N, \ell \geq 1}$ can be sampled as follows. For some measure ν^* that is a fixed point of the operator $\Gamma_{1, 0}$:

S1 generate a sequence $(x_i)_{i \geq 1}$ of i.i.d. random variables with law ν^*;

S2 conditionally on $(x_i)_{i \geq 1}$, generate ± 1-valued random variables σ_i^ℓ with mean x_i, independently over $i, \ell \geq 1$.

Our second main result states that, assuming that the limit mutual information exists, it can be expressed in terms of an explicit functional $\mathscr{P}_{1, 0} : \mathscr{M}_+ \to \mathbb{R}$ that is evaluated at a fixed point of the map $\Gamma_{1, 0}$. The functional $\mathscr{P}_{1, 0}$ is closely related to the map $\Gamma_{1, 0}$ since, at least formally, a measure is a critical point of $\mathscr{P}_{1, 0}$ if and only if it is a fixed point of the map $\Gamma_{1, 0}$.

While we think of our results as being rather precise, since they substantially narrow down the space of possible asymptotic behaviors for the mutual information and for the conditional law of σ^* given the observation of $\mathbf{G}_N$, we also wish to stress that our results do not provide a complete characterization of the limits of these objects. Indeed, this is due to the fact that the map $\Gamma_{1, 0}$ may admit several fixed points in general. We will show that the existence of a unique fixed point can be

guaranteed in a regime of small signal-to-noise ratio, but we do not expect this uniqueness property to extend to arbitrary choices of parameters.

A complete characterization of the limit of the mutual information has already been achieved in a number of cases. The disassortative case was settled in [2, 15] when $p = 1/2$, and in [18] for arbitrary p. The technique in [18] is based on partial differential equations (see also [19] for a presentation of the approach), and the key point is that in the disassortative case, the non-linearity appearing in this equation is *convex*. We expect this complete characterization to extend to settings with more than two communities, as long as the corresponding non-linearity remains convex. This would be analogous to [11, Theorem 1.1] in the context of spin glasses, a result which itself builds upon [9, 22, 42, 43, 44, 45, 46, 47, 52]. Moreover, whenever this convexity property holds, the limit mutual information can be expressed as a variational formula. From the point of view of partial differential equations, this relates to the validity of a Hopf-Lax representation of the solution, see [20, Theorem 1.3].

The assortative case presents a greater challenge. It was successfully resolved when $p = 1/2$ in [53], building upon [1, 24, 34, 36]. The approach in [53] relies upon the study of the fixed points of an operator that relates to the setting of broadcasting on trees, and the key step of the proof is to establish the uniqueness of a non-trivial fixed point. This uniqueness property does not generalize to models with more than two communities [21]. Despite similarities, we could not identify a precise correspondence between the operator $\Gamma_{1,0}$ whose fixed points enter into our analysis and the operators appearing in [53]. Part of the difficulty is that the first step of the approach in these earlier works is to decompose the mutual information into an integral involving a family of new statistical inference problems with an additional survey mechanism. When one can guarantee that there is a unique non-trivial fixed point to each problem in this family, there is a tight relationship between the true community-detection problem and the problems involving broadcasting on trees. On the other hand, when uniqueness fails, this relationship seems less rigid to us, and we do not immediately see how one could derive results comparable to those obtained in the present paper in the context explored in these earlier works.

We expect the method presented here, primarily based on cavity calculations, to be robust to model variations such as situations involving more than two communities, and we also expect the lower bound on the mutual information obtained in [18] to generalize similarly. The results we present here and those in [18] are analogous to those obtained for spin glasses in [11] and in [39, 40] respectively. We have preferred to stick to a two-community model here to facilitate the presentation, but as an illustration of the robustness of the approach, we will at least indicate as we proceed the few small adaptations that need to be implemented in order to cover the case when then measure P^* is arbitrary with support in $[-1, 1]$ in place of $\{-1, 1\}$.

In order to make further progress and obtain a full characterization of the limit mutual information in general, one must first ask what a good candidate for the limit could be. One proposal based on partial differential equations was put forward in [18]. It would of course be convenient to identify a more explicit formulation of the candidate limit. For dense versions of the problem we consider, where the average degree of the graph diverges to infinity, a sort of central limit theorem takes place at the level of each node, and one can relate the problem to a simpler setup such as that of rank-one matrix estimation in the presence of additive Gaussian noise, as shown in [17, 26] (see also [19, Section 4.5]). For the latter class of problems, a saddle-point variational

formula was shown to be valid in a very wide range of settings in [14] (see also [19, Chapter 4]), building upon [4, 5, 6, 10, 13, 12, 23, 26, 27, 28, 29, 31, 32, 37, 38, 50, 51]. From the point of view of partial differential equations, this corresponds to the possibility to write a Hopf formula based on the convexity of the initial condition (as opposed to the convexity of the non-linearity for the Hopf-Lax formula, cf. [19, Theorems 3.8 and 3.13]). However, the convexity of the relevant "initial condition" in our setting is invalid [25], and as explained around [18, (1.49)], this implies that this tentative saddle-point Hopf formula is invalid in our setting with sparse graph connectivity. An alternative candidate limit for the mutual information could be considered by extending the Hopf-Lax variational formula known to be valid in the disassortative case. While we could not rule this out in the two-community case, we show here that this candidate variational formula is invalid in general. To do so, we focus on a version of the two-community problem in which the graph additionally possesses a bipartite structure. The model can be thought of as involving four communities, and it is inspired by the bipartite spin-glass model investigated in [39] (see in particular [39, Section 6]). In this context, we show that a naive generalization of the formula valid in the standard disassortative two-community case would lead to a prediction that is in contradiction with the main results of this paper.

As was already said, the main focus of this paper is to obtain precise characterizations of the mutual information and conditional law of the community structure given the observation, in the limit of large system size. In the two-community case with $p = 1/2$, it was shown in [30, 33, 35] that there exists an efficient algorithm that allows one to recover non-trivial information about the community structure whenever $\Delta^2 > c$, and that it is information-theoretically impossible to do so otherwise. To the best of our knowledge, whether or not there exists an efficient algorithm that recovers as much of the signal as is information-theoretically possible is not known. In certain settings involving more than two communities, one expects the existence of a gap between the regime for which there exists an efficient algorithm for recovering a non-trivial fraction of the community structure, and the regime for which it is information-theoretically possible to do so. The results of [3] vindicate a significant part of this picture. A precise characterization of the information-theoretic limit is not known however. These open questions are directly related to the main purposes of the present paper. Much of the recent work on the community detection problem has been triggered by [16], in which the results reviewed in this paragraph and much more had been anticipated using non-rigorous methods.

1.2 Precise statement of the main results

We now turn to a more precise description of our main results. As in [18], our starting point is to embed the community-detection problem under consideration into a richer family parametrized by a pair $(t, \mu) \in \mathbb{R}_{\geq 0} \times \mathcal{M}_+$. The conjecture in [18] is that, up to a simple affine transformation, the limit $f(t, \mu)$ of the re-scaled mutual information between the signal and the observation solves an explicit Hamilton-Jacobi equation. As is well-known in finite dimensions, first-order Hamilton-Jacobi equations can be solved for a short time using the method of characteristics (see for instance [19, Exercise 3.10 and solution]). This method allows us to uncover the value of the solution to the equation along each characteristic line separately. At larger times though, difficulties occur because characteristic lines may intersect, and will typically not agree on what the value of the

solution ought to be. Our main result can be interpreted as saying that, while this might indeed occur, there is always one characteristic line that prescribes the correct value for $f(t,\mu)$. In the language of [19, Section 3.5], one may say that the graph of f always belongs to the wavefront of the equation.

As already stated, the central object of interest to us is the mutual information between the community structure σ^* and the observed random graph $\mathbf{G}_N$, that is,

$$I(\mathbf{G}_N;\sigma^*) := \mathbb{E}\log\frac{\mathbb{P}(\mathbf{G}_N,\sigma^*)}{\mathbb{P}(\mathbf{G}_N)\mathbb{P}(\sigma^*)} = \mathbb{E}\int_{\mathbb{R}^N}\log\left(\frac{\mathrm{d}P_{\sigma^*|\mathbf{G}_N}}{\mathrm{d}P_N^*}(\sigma)\right)\mathrm{d}P_{\sigma^*|\mathbf{G}_N}(\sigma), \qquad (1.7)$$

where $P_{\sigma^*|\mathbf{G}_N}$ denotes the conditional law of σ^* given $\mathbf{G}_N$. The likelihood of the model is given by

$$\mathbb{P}\{\mathbf{G}_N = (G_{ij}) \mid \sigma^* = \sigma\} = \prod_{i<j}\left(\frac{c+\Delta\sigma_i\sigma_j}{N}\right)^{G_{ij}}\left(1-\frac{c+\Delta\sigma_i\sigma_j}{N}\right)^{1-G_{ij}}, \qquad (1.8)$$

and Bayes' formula implies that the posterior of the model is the Gibbs measure

$$\mathbb{P}\{\sigma^* = \sigma \mid \mathbf{G}_N = (G_{ij})\} = \frac{\exp\left(H_N^\circ(\sigma)\right)P_N^*(\sigma)}{\int_{\Sigma_N}\exp\left(H_N^\circ(\tau)\right)\mathrm{d}P_N^*(\tau)} \qquad (1.9)$$

associated with the Hamiltonian

$$H_N^\circ(\sigma) := \sum_{i<j}\log\left[(c+\Delta\sigma_i\sigma_j)^{G_{ij}}\left(1-\frac{c+\Delta\sigma_i\sigma_j}{N}\right)^{1-G_{ij}}\right]. \qquad (1.10)$$

Up to an error vanishing with N and a simple additive constant, the normalized mutual information (1.7) coincides with the free energy

$$\overline{F}_N^\circ := \frac{1}{N}\mathbb{E}\log\int_{\Sigma_N}\exp H_N^\circ(\sigma)\,\mathrm{d}P_N^*(\sigma). \qquad (1.11)$$

Indeed, one can observe that

$$I(\mathbf{G}_N;\sigma^*) = \binom{N}{2}\mathbb{E}\log(c+\Delta\sigma_1^*\sigma_2^*)^{G_{12}}\left(1-\frac{c+\Delta\sigma_1^*\sigma_2^*}{N}\right)^{1-G_{12}} - N\overline{F}_N^\circ; \qquad (1.12)$$

averaging with respect to the randomness of G_{12} and Taylor-expanding the logarithm yields that

$$\frac{1}{N}I(\mathbf{G}_N;\sigma^*) = \frac{1}{2}\mathbb{E}\left(c+\Delta\sigma_1^*\sigma_2^*\right)\log\left(c+\Delta\sigma_1^*\sigma_2^*\right) - \frac{c}{2} - \frac{\Delta\overline{m}^2}{2} - \overline{F}_N^\circ + \mathcal{O}\left(N^{-1}\right). \qquad (1.13)$$

In place of the mutual information, the main protagonist of this paper is in fact the free energy $\overline{F}_N^\circ$ and its extension, which we now proceed to introduce.

We begin by describing the enrichment of the free energy involving the parameter $t \geq 0$. For each $t \geq 0$, we introduce a random variable $\Pi_t \sim \mathrm{Poi}\,t\binom{N}{2}$ as well as an independent family of i.i.d. random matrices $(G^k)_{k\geq 1}$ each having conditionally independent entries $(G_{i,j}^k)_{i,j\leq N}$ taking values in $\{0,1\}$ with conditional distribution

$$\mathbb{P}\{G_{i,j}^k = 1 \mid \sigma^*\} := \frac{c+\Delta\sigma_i^*\sigma_j^*}{N}. \qquad (1.14)$$

Given a collection of random indices $(i_k, j_k)_{k \geq 1}$ sampled uniformly at random from $\{1, \ldots, N\}^2$, independently of the other random variables, we define the t-dependent Hamiltonian H_N^t on Σ_N by

$$H_N^t(\sigma) := \sum_{k \leq \Pi_t} \log \left[\left(c + \Delta \sigma_{i_k} \sigma_{j_k} \right)^{G_{i_k, j_k}^k} \left(1 - \frac{c + \Delta \sigma_{i_k} \sigma_{j_k}}{N} \right)^{1 - G_{i_k, j_k}^k} \right], \tag{1.15}$$

and denote by

$$\overline{F}_N(t) := \frac{1}{N} \mathbb{E} \log \int_{\Sigma_N} \exp H_N^t(\sigma) \, \mathrm{d} P_N^*(\sigma) \tag{1.16}$$

its corresponding free energy. Notice that this is the Hamiltonian associated with the task of inferring the signal σ^* from the data

$$\mathscr{D}_N^t := \left(\Pi_t, (i_k, j_k)_{k \leq \Pi_t}, (G_{i_k, j_k}^k)_{k \leq \Pi_t} \right). \tag{1.17}$$

By this we mean that the law of σ^* given the observation of $\mathscr{D}_N^t$ is described by a formula analogous to that in (1.9), with H_N° replaced by H_N^t there. By [18, Appendix A], the difference between the free energies $\overline{F}_N^\circ$ and $\overline{F}_N(1)$ converges to zero as N tends to infinity. The main advantage of this construction is that it allows one to study derivatives of the free energy with respect to the "time" parameter t. It is also convenient to study derivatives of the free energy with respect to an additional parameter denoted by $\mu \in \mathcal{M}_+$. Given $\mu \in \mathcal{M}_+$, we let $(\Lambda_i(N\mu))_{i \leq N}$ be independent Poisson point processes with intensity measure $N\mu$, and we let the μ-dependent Hamiltonian H_N^μ on Σ_N be given by

$$H_N^\mu(\sigma) := \sum_{i \leq N} \sum_{x \in \Lambda_i(N\mu)} \log \left[\left(c + \Delta \sigma_i x \right)^{G_i^x} \left(1 - \frac{c + \Delta \sigma_i x}{N} \right)^{1 - G_i^x} \right], \tag{1.18}$$

where $(G_i^x)_{i \leq N}$ are conditionally independent random variables taking values in $\{0, 1\}$ with conditional distribution

$$\mathbb{P}\{G_i^x = 1 \mid \sigma^*, x\} := \frac{c + \Delta \sigma_i^* x}{N}. \tag{1.19}$$

For a review of Poisson point processes, we refer the reader to [19, Chapter 5]. The function H_N^μ is the Hamiltonian associated with the task of inferring the signal σ^* from the data

$$\mathscr{D}_N^\mu := \left(\Lambda_i(N\mu), (G_i^x)_{x \in \Lambda_i(N\mu)} \right)_{i \leq N}. \tag{1.20}$$

Finally, for each $(t, \mu) \in \mathbb{R}_{\geq 0} \times \mathcal{M}_+$, we introduce the enriched Hamiltonian

$$H_N^{t, \mu}(\sigma) := H_N^t(\sigma) + H_N^\mu(\sigma) \tag{1.21}$$

as well as its associated free energy

$$\overline{F}_N(t, \mu) := \frac{1}{N} \mathbb{E} \log \int_{\Sigma_N} \exp H_N^{t, \mu}(\sigma) \, \mathrm{d} P_N^*(\sigma). \tag{1.22}$$

The function $H_N^{t, \mu}$ is the Hamiltonian associated with inferring the signal σ^* from the data

$$\mathscr{D}_N^{t, \mu} := (\mathscr{D}_N^t, \mathscr{D}_N^\mu), \tag{1.23}$$

6

where the randomness in these two datasets is taken to be independent conditionally on σ^*. We recall that by [18, Appendix A], the difference between the free energy $\overline{F}_N^\circ$ in (1.11) and $\overline{F}_N(1,0)$ tends to zero as N tends to infinity. The problem of finding the asymptotic value of the mutual information (1.7) is therefore a specific instance of the task of determining the limit of the enriched free energy (1.22).

To state the main results of this paper, we introduce additional notation. We define the function $g : [-1,1] \to \mathbb{R}$ by

$$g(z) := (c + \Delta z)\big(\log(c + \Delta z) - 1\big) = (c + \Delta z)\log(c) + c \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} z^n - c, \qquad (1.24)$$

and for each measure $\mu \in \mathscr{M}_+$, we let $G_\mu : [-1,1] \to \mathbb{R}$ be the function obtained by marginalizing the kernel $(x,y) \mapsto g(xy)$,

$$G_\mu(x) := \int_{-1}^{1} g(xy)\,\mathrm{d}\mu(y). \qquad (1.25)$$

Given $(t,\mu) \in \mathbb{R}_{\geq 0} \times \mathscr{M}_+$, we introduce the functional $\mathscr{P}_{t,\mu} : \mathscr{M}_+ \to \mathbb{R}$ defined by

$$\mathscr{P}_{t,\mu}(\nu) := \psi(\mu + t\nu) - \frac{t}{2}\int_{-1}^{1} G_\nu(y)\,\mathrm{d}\nu(y), \qquad (1.26)$$

where $\psi : \mathscr{M}_+ \to \mathbb{R}$ denotes the function

$$\psi(\mu) := -\mu[-1,1]c + p\,\mathbb{E}\log \int_{\Sigma_1} \exp\Big(-\Delta\sigma\int_{-1}^{1} x\,\mathrm{d}\mu(x)\Big) \prod_{x \in \Pi_+(\mu)} (c + \Delta\sigma x)\,\mathrm{d}P^*(\sigma)$$

$$+ (1-p)\,\mathbb{E}\log \int_{\Sigma_1} \exp\Big(-\Delta\sigma\int_{-1}^{1} x\,\mathrm{d}\mu(x)\Big) \prod_{x \in \Pi_-(\mu)} (c + \Delta\sigma x)\,\mathrm{d}P^*(\sigma) \quad (1.27)$$

for Poisson point processes $\Pi_\pm(\mu)$ with respective intensity measures $(c \pm \Delta x)\,\mathrm{d}\mu(x)$ on $[-1,1]$. The function ψ is the initial condition of the partial differential equation that was identified in [18], since it arises as the large-N limit of $\overline{F}_N(0,\mu)$. Our first main result states that the limit of the free energy, assuming it exists, is given by the functional (1.26) evaluated at one of its critical points. As mentioned in the previous subsection, it will be possible to recast the critical-point condition as a fixed-point equation for the operator $\Gamma_{t,\mu}$ defined in (1.6), so that our first main result reads as follows.

Theorem 1.1. *Suppose that the sequence of enriched free energies $(\overline{F}_N)_{N \geq 1}$ converges pointwise to some limit $f : \mathbb{R}_{\geq 0} \times \mathscr{M}_+ \to \mathbb{R}$. For every $(t,\mu) \in \mathbb{R}_{\geq 0} \times \mathscr{M}_+$, the map $\nu \mapsto \Gamma_{t,\mu}(\nu)$ admits a fixed point $\nu^* \in \mathscr{M}_p$ with the property that*

$$f(t,\mu) = \mathscr{P}_{t,\mu}(\nu^*). \qquad (1.28)$$

Our second result concerns the asymptotic behavior of the conditional law of the community structure σ^* given the observation. We use the bracket notation $\langle \cdot \rangle$ to denote the expectation with

respect to the conditional law of σ^* given the observation, with σ being the canonical random variable under $\langle \cdot \rangle$. Explicitly, this means that for every function $h : \Sigma_N \to \mathbb{R}$, we have

$$\langle h(\sigma) \rangle := \frac{\int_{\Sigma_N} h(\sigma) \exp H_N^{t,\mu}(\sigma) \, dP_N(\sigma)}{\int_{\Sigma_N} \exp H_N^{t,\mu}(\sigma) \, dP_N(\sigma)}. \tag{1.29}$$

We stress that the expectation $\langle \cdot \rangle$ is itself random, since it depends on the observation. We denote by $(\sigma^\ell)_{\ell \geq 1} = ((\sigma_i^\ell)_{i \leq N})_{\ell \geq 1}$ independent copies, often called replicas, of the random variable σ under $\langle \cdot \rangle$. We also introduce the set $\mathrm{Reg}(\mathscr{M}_+)$ of measures in $\mathscr{M}_+$ that admit a smooth and strictly positive density with respect to the Lebesgue measure on $[-1,1]$,

$$\mathrm{Reg}(\mathscr{M}_+) := \big\{ \mu \in \mathscr{M}_+ \mid \mu \text{ admits a smooth and strictly positive density on } [-1,1] \big\}. \tag{1.30}$$

Our next result states that for most choices of $(t,\mu) \in \mathbb{R}_{\geq 0} \times \mathscr{M}_+$, we can identify a fixed point of the map $\nu \mapsto \Gamma_{t,\mu}(\nu)$ for which the equality (1.28) holds, and for which we can also identify the limit law of $(\sigma^\ell)_{\ell \geq 1}$ after a small perturbation of the inference problem. More precisely, we will show that this can be done at any point $(t,\mu) \in \mathbb{R}_{\geq 0} \times \mathrm{Reg}(\mathscr{M}_+)$ at which the limit free energy is Gateaux differentiable; we refer to Section 2 for a definition of Gateaux differentiability and a precise discussion of why the limit free energy is Gateaux differentiable at most points. The convergence in law of $(\sigma_i^\ell)_{i \leq N, \ell \geq 1}$ is in the sense of finite-dimensional distributions.

Theorem 1.2. *Suppose that the sequence of enriched free energies $(\overline{F}_N)_{N \geq 1}$ converges pointwise to some limit $f : \mathbb{R}_{\geq 0} \times \mathscr{M}_+ \to \mathbb{R}$. If $(t,\mu) \in \mathbb{R}_{\geq 0} \times \mathrm{Reg}(\mathscr{M}_+)$ is such that $f(t,\cdot)$ is Gateaux differentiable at μ, then we can identify a probability measure $\nu^* \in \mathscr{M}_p$ such that the Gateaux derivative density of f at (t,μ) is G_{ν^*}. Moreover, the probability measure ν^* is a fixed point of the map $\Gamma_{t,\mu}$, and we have*

$$f(t,\mu) = \mathscr{P}_{t,\mu}(\nu^*). \tag{1.31}$$

Finally, there exists a sequence $(N_k, \lambda^{N_k})_{k \geq 1}$ such that $(N_k)_{k \geq 1}$ increases to infinity, the spin array $(\sigma_i^\ell)_{i \leq N_k, \ell \geq 1}$ converges in law under $\mathbb{E}\langle \cdot \rangle$ with the modified Hamiltonian (4.6) with perturbation parameters $(\lambda^{N_k})_{k \geq 1}$, and the limit law is generated by ν^* according to (S1)-(S2). The perturbation parameters λ^{N_k} introduced in the modified Hamiltonian (4.6) are sufficiently small that the difference between the free energy (1.22) and the free energy with the perturbed Hamiltonian (4.6) tends to zero as N tends to infinity.

One limitation of Theorems 1.1 and 1.2 is that they assume the existence of the limit free energy, which we do not know a priori. To a large extent, this is a side effect of the fact that we do not know how to single out the correct critical point in general, as was discussed in the previous subsection. Correspondingly, this assumption can be removed whenever we can guarantee the uniqueness of a fixed point to $\Gamma_{t,\mu}$. As shown in Lemma 5.3, the second part of Theorem 1.2 in fact does not require that we assume that the free energy converges. Indeed, to show this result we only need to ensure that the free energy converges along a subsequence, and this can always be obtained using the Arzelà-Ascoli theorem.

The next proposition certifies the uniqueness of a fixed point in a regime of small signal-to-noise ratio.

Proposition 1.3. *There exists $C < +\infty$ such that for all $t < C^{-1}$ and $\mu \in \mathscr{M}_+$, the map $\nu \mapsto \Gamma_{t,\mu}(\nu)$ admits a unique fixed point ν^* in $\mathscr{M}_p$. Moreover, the sequence of enriched free energies $(\overline{F}_N(t,\mu))_{N \geq 1}$ converges to $\mathscr{P}_{t,\mu}(\nu^*)$,*

$$\lim_{N \to +\infty} \overline{F}_N(t,\mu) = \mathscr{P}_{t,\mu}(\nu^*). \tag{1.32}$$

For completeness, we also explain how to recover the variational formula already obtained in [2, 15] (for $p = 1/2$) and in [18] for the limit free energy in the disassortative case.

Proposition 1.4. *In the disassortative case, that is when $\Delta \leq 0$, we have for every $(t,\mu) \in \mathbb{R}_{\geq 0} \times \mathscr{M}_+$ that*

$$\lim_{N \to +\infty} \overline{F}_N(t,\mu) = \sup_{\nu \in \mathscr{M}_p} \mathscr{P}_{t,\mu}(\nu). \tag{1.33}$$

As explained in the previous subsection, we will also show the invalidity of a candidate variational formula for the limit free energy. Since this involves a variant of the model with an additional bipartite structure, we prefer to postpone a precise description of the setting and result to Section 7.

1.3 Organization of the paper

In Section 2, we recall the expressions for the derivatives of the free energy obtained in [18]. We then leverage these to establish the uniform Lipschitz continuity of the free energy, the almost everywhere differentiability of the limit free energy, and the asymptotic behavior of the derivative of the free energy. Section 3 is a brief overview of the main multioverlap concentration result in [8] which implies that the spin array sampled from the asymptotic Gibbs measure is generated according to (**S1**)-(**S2**) for some measure ν^* in $\mathscr{M}_p$. In Section 4, the cavity computations leading to the functional $\mathscr{P}_{t,\mu}$ in (1.26) are performed. In Section 5, the cavity representation is combined with the multioverlap concentration result in Section 3 and the regularity properties of the free energy obtained in Section 2 to prove Theorems 1.1 and 1.2. Section 6 is devoted to the proofs of Propositions 1.3 and 1.4. Finally, in Section 7, it is argued that, unlike in the disassortative setting, one should not in general expect the limit free energy to be given by evaluating the functional $\mathscr{P}_{t,\mu}$ at the measure $\nu \in \mathscr{M}_p$ which maximizes its value.

2 Regularity properties of the free energy

In this section we discuss three regularity properties of the enriched free energy (1.22). In Section 2.1, we establish a Lipschitz continuity bound for the enriched free energy which implies in particular that the free energy restricted to the space of probability measures is Lipschitz continuous with respect to the Wasserstein distance. In Section 2.2, we prove that a Lipschitz continuous function on the space of cumulative distribution functions on $[-1,1]$, viewed as a subset of $L^2[-1,1]$,

is Gateaux differentiable on a dense subset of its domain. Finally, in Section 2.3, we leverage a convexity property of the free energy to prove that the sequence of derivatives of the free energy converges to the derivative of the limit free energy at any point of differentiability of said limit. We now state the three main results of this section. Since the ideas developed in these proofs will not reappear later, the reader may consider skipping these proofs on first reading.

The Lipschitz continuity result in Section 2.1 involves the Wasserstein distance. The *Wasserstein distance* between two probability measures $\mu, \nu \in \Pr[-1,1]$ is defined by

$$W(\mu, \nu) := \sup \left\{ \left| \int_{-1}^{1} h(x) \, d\mu(x) - \int_{-1}^{1} h(x) \, d\nu(x) \right| \, \Big| \, \|h\|_{\mathrm{Lip}} \leq 1 \right\}, \tag{2.1}$$

where $\|\cdot\|_{\mathrm{Lip}}$ denotes the Lipschitz semi-norm

$$\|h\|_{\mathrm{Lip}} := \sup_{x \neq x' \in [-1,1]} \frac{|h(x) - h(x')|}{|x - x'|} \tag{2.2}$$

on the space of Lipschitz continuous functions $h : [-1,1] \to \mathbb{R}$. The Kantorovich-Rubinstein theorem [48, Theorem 4.15] ensures that the Wasserstein distance admits the dual representation

$$W(\mu, \nu) = \inf \left\{ \mathbb{E}|X - Y| \mid X \sim \mu \text{ and } Y \sim \nu \right\}. \tag{2.3}$$

In the one-dimensional setting, as shown in [41, Theorem 1.5.1 and Corollary 1.5.3], the Wasserstein distance admits an explicit representation in terms of the cumulative distribution function or the quantile function. Indeed, if $F_\mu : [-1,1] \to [0,1]$ and $F_\mu^{-1} : [0,1] \to [-1,1]$ denote the cumulative distribution function and the quantile transform of a probability measure $\mu \in \Pr[-1,1]$,

$$F_\mu(x) := \mu[-1, x] \quad \text{and} \quad F_\mu^{-1}(u) := \inf\{x \in \mathbb{R} \mid F_\mu(x) \geq u\}, \tag{2.4}$$

then for all probability measures $\mu, \nu \in \Pr[-1,1]$,

$$W(\mu, \nu) = \int_0^1 \left| F_\mu^{-1}(u) - F_\nu^{-1}(u) \right| du = \int_{-1}^1 \left| F_\mu(x) - F_\nu(x) \right| dx. \tag{2.5}$$

In this notation, the main Lipschitz continuity result for the enriched free energy reads as follows.

Proposition 2.1. *There exists $C < +\infty$ such that for any $(t, \mu), (t, \mu') \in \mathbb{R}_{\geq 0} \times \mathcal{M}_+$,*

$$\left| \overline{F}_N(t, \mu) - \overline{F}_N(t', \mu') \right| \leq C \left(|t - t'| + \left| \mu[-1,1] - \mu'[-1,1] \right| + \mu[-1,1] |F_{\overline{\mu}} - F_{\overline{\mu}'}|_{L^2} \right). \tag{2.6}$$

In Sections 2.2-2.3 and throughout the paper, the notion of Gateaux differentiability that we take may differ from the most classical one, so we proceed to define it precisely. As in [11], we fix a topological vector space X and a subset U of X. We denote by X^* the continuous dual of X, and we write $\langle \cdot, \cdot \rangle$ for the duality pairing. We will not necessarily take U to be an open set, so, given a point $q \in U$, the set

$$\mathrm{Adm}(U, q) := \left\{ x \in X \mid \text{there exists } r > 0 \text{ such that for all } t \in [0, r] \text{ we have } q + tx \in U \right\} \tag{2.7}$$

of admissible directions along which a small line segment starting at q is contained in U will play an important part in the definition. Indeed, we say that a function $h : U \to \mathbb{R}$ is *Gateaux differentiable* at $q \in U$ if the following two conditions hold.

GD1 For every $x \in \mathrm{Adm}(U,q)$, the limit

$$Dh(q;x) := \lim_{\varepsilon \searrow 0} \frac{h(q + \varepsilon x) - h(q)}{\varepsilon} \tag{2.8}$$

exists.

GD2 There is a unique $y \in X^*$ such that, for every $x \in \mathrm{Adm}(U,q)$, we have $Dh(q;x) = \langle y, x \rangle$.

In this case, the dual vector y is called the *Gateaux derivative* of h at q.

We will mainly be concerned with the case $X = \mathbb{R}_{\geq 0} \times \mathscr{M}_+$ endowed with the product of the Euclidean topology and the topology of weak convergence, which is induced by the Wasserstein distance (2.1). In this setting, given a function $f : \mathbb{R}_{\geq 0} \times \mathscr{M}_+ \to \mathbb{R}$, a time $t \geq 0$, and measures $\mu \in \mathscr{M}_+$ and $\nu \in \mathrm{Adm}(\mathscr{M}_+, \mu)$, we denote by $D_\mu f(t, \mu; \nu)$ the Gateaux derivative of the function $f(t, \cdot)$ at the measure μ in the direction ν,

$$D_\mu f(t, \mu; \nu) := \lim_{\varepsilon \to 0} \frac{f(t, \mu + \varepsilon \nu) - f(t, \mu)}{\varepsilon}. \tag{2.9}$$

We say that the Gateaux derivative of $f(t, \cdot)$ admits a density at the measure $\mu \in \mathscr{M}_+$ if there exists a continuous function $x \mapsto D_\mu f(t, \mu, x)$ defined on the interval $[-1, 1]$ such that for all $\nu \in \mathrm{Adm}(\mathscr{M}_+, \mu)$,

$$D_\mu f(t, \mu; \nu) = \int_{-1}^{1} D_\mu f(t, \mu, x) \, d\nu(x). \tag{2.10}$$

We often abuse notation and identify the density $D_\mu f(t, \mu, \cdot)$ with the Gateaux derivative $D_\mu f(t, \mu)$. Implicitly identifying the continuous dual of $\mathscr{M}_+$ with the space of continuous functions on $[-1, 1]$, this definition of Gateaux differentiability coincides with that in (**GD1**)-(**GD2**) but differs slightly from that in [18] where the density is only required to be measurable and bounded as opposed to continuous. However, all Gateaux derivative densities appearing in [18] are of the finite-volume free energy and are continuous, so this difference is insignificant.

For technical reasons, it will be convenient for the dense subset of $\mathbb{R}_{\geq 0} \times \mathscr{M}_+$ on which we establish the Gateaux differentiability of the free energy to be composed of well-behaved measures. Recall the definition in (1.30) of the set $\mathrm{Reg}(\mathscr{M}_+)$ of measures in $\mathscr{M}_+$ that admit a smooth and strictly positive density with respect to Lebesgue measure on $[-1, 1]$. A key property of this set is that for any $\mu, \nu \in \mathrm{Reg}(\mathscr{M}_+)$ and $t \in \mathbb{R}$ with $|t|$ small enough, we have $\mu + t\nu \in \mathscr{M}_+$. In particular, the set of admissible directions at $\mu \in \mathrm{Reg}(\mathscr{M}_+)$ contains $\mathrm{Reg}(\mathscr{M}_+)$. In this notation, the main Gateaux differentiability results for the enriched free energy proved in Sections 2.2 and 2.3 read as follows.

Proposition 2.2. *Suppose that the sequence of enriched free energies $(\overline{F}_N)_{N \geq 1}$ converges pointwise to some limit $f : \mathbb{R}_{\geq 0} \times \mathscr{M}_+ \to \mathbb{R}$ along a subsequence $(N_k)_{k \geq 1}$. The subsequential limit f is Gateaux differentiable jointly in its two variables on a subset of $\mathbb{R}_{>0} \times \mathrm{Reg}(\mathscr{M}_+)$ that is dense in $\mathbb{R}_{\geq 0} \times \mathscr{M}_+$.*

Proposition 2.3. *Suppose that the sequence of enriched free energies $(\overline{F}_N)_{N \geq 1}$ converges pointwise to some limit $f : \mathbb{R}_{\geq 0} \times \mathscr{M}_+ \to \mathbb{R}$ along a subsequence $(N_k)_{k \geq 1}$. For each $t \geq 0$, if $f(t, \cdot)$ is Gateaux differentiable at $\mu \in \mathrm{Reg}(\mathscr{M}_+)$, then $D_\mu \overline{F}_{N_k}(t, \mu, \cdot)$ converges weakly to $D_\mu f(t, \mu, \cdot)$.*

11

In proving Propositions 2.1 and 2.2, it will be convenient to remember from [18, Lemmas 2.1 and 2.3] that the enriched free energy (1.22) is Gateaux differentiable at every $(t,\mu) \in \mathbb{R}_{\geq 0} \times \mathscr{M}_+$ with

$$\partial_t \overline{F}_N(t,\mu) = \frac{1}{2}\mathbb{E}\left(c + \Delta\langle\sigma_1\sigma_2\rangle\right)\log\left(c + \Delta\langle\sigma_1\sigma_2\rangle\right) - \frac{\Delta\overline{m}^2}{2} - \frac{c}{2} + \mathscr{O}(N^{-1}), \tag{2.11}$$

$$D_\mu \overline{F}_N(t,\mu,x) = \mathbb{E}\left(c + \Delta\langle\sigma_1\rangle x\right)\log\left(c + \Delta\langle\sigma_1\rangle x\right) - c - \Delta\overline{m}x + \mathscr{O}(N^{-1}). \tag{2.12}$$

We have used $\langle\cdot\rangle$ to denote the Gibbs average associated with the enriched Hamiltonian (1.21). This means that for any bounded and measurable function $f = f(\sigma^1,\ldots,\sigma^n)$ of finitely many replicas,

$$\langle f(\sigma^1,\ldots,\sigma^n)\rangle := \langle f \rangle := \frac{\int_{\Sigma_N^n} f(\sigma^1,\ldots,\sigma^n)\prod_{\ell\leq n}\exp H_N^{t,\mu}(\sigma^\ell)\,dP_N^*(\sigma^\ell)}{\left(\int_{\Sigma_N}\exp H_N^{t,\mu}(\sigma)\,dP_N^*(\sigma)\right)^n}. \tag{2.13}$$

The computations leading to (2.11)-(2.12) are considerably simplified by the *Nishimori identity*. This identity allows us to freely interchange one replica σ^ℓ by the signal σ^* when taking an average with respect to all sources of randomness, thus avoiding a cascade of new replicas as we differentiate the free energy. More precisely, it states that, for every bounded and measurable function $f = f(\sigma^1,\ldots,\sigma^n,\mathscr{D}_N^{t,\mu})$ of finitely many replicas and the data,

$$\mathbb{E}\left\langle f\left(\sigma^1,\sigma^2,\ldots,\sigma^n,\mathscr{D}_N^{t,\mu}\right)\right\rangle = \mathbb{E}\left\langle f\left(\sigma^*,\sigma^2,\ldots,\sigma^n,\mathscr{D}_N^{t,\mu}\right)\right\rangle. \tag{2.14}$$

This can be first verified for functions of product form using that the Gibbs average is the conditional law of the signal σ^* given the data $\mathscr{D}_N^{t,\mu}$, and then extended to all bounded and measurable functions by a monotone class argument as in [19, Proposition 4.1].

The proof of Proposition 2.3 will rely on a convexity property of the free energy that will be obtained through information-theoretic arguments taken from [25].

2.1 Uniform Lipschitz continuity of the free energy

The Lipschitz continuity of the free energy stated in Proposition 2.1 will be deduced from the Lipschitz continuity of an extension of the free energy. Given a probability measure $\mu \in \mathrm{Pr}[-1,1]$, consider a sequence $x = (x_{i,k})_{i,k\geq 1}$ of i.i.d. random variables with law μ. For each $s > 0$ and $i \geq 1$, let $\Pi_{i,s} \sim \mathrm{Poi}(sN)$ be independent over $i \geq 1$, and introduce the Hamiltonian on Σ_N defined by

$$\widetilde{H}_N^{s,\mu}(\sigma) := \sum_{i\leq N}\sum_{k\leq\Pi_{i,s}}\log\left[\left(c + \Delta\sigma_i x_{i,k}\right)^{\widetilde{G}_{i,k}^x}\left(1 - \frac{c + \Delta\sigma_i x_{i,k}}{N}\right)^{1-\widetilde{G}_{i,k}^x}\right], \tag{2.15}$$

where the random variables $(\widetilde{G}_{i,k}^x)_{i,k\geq 1}$ are independent with conditional distribution

$$\mathbb{P}\left\{\widetilde{G}_{i,k}^x = 1 \mid \sigma^*,x\right\} := \frac{c + \Delta\sigma_i^* x_{i,k}}{N}. \tag{2.16}$$

Recalling the definition of the time-dependent Hamiltonian in (1.15), we define the extended free energy functional

$$\widetilde{F}_N(t,s,\mu) := \frac{1}{N}\mathbb{E}\log\int_{\Sigma_N}\exp\left(H_N^t(\sigma) + \widetilde{H}_N^{s,\mu}(\sigma)\right)dP_N^*(\sigma). \tag{2.17}$$

12

For any $\mu \in \mathscr{M}_+$, we have $\overline{F}_N(t,\mu) = \widetilde{F}_N(t,\mu[-1,1],\overline{\mu})$. Indeed, this is how the enriched free energy (1.22) was defined in [18]. The Lipschitz continuity of the extended free energy (2.17) will be obtained by combining the mean value theorem with the derivative expressions (2.11)-(2.12). It will also be convenient to remember from [18, Lemma 2.3] that

$$\partial_s \widetilde{F}_N(t,s,\mu) = \mathbb{E}\left(c + \Delta \langle \sigma_1 \rangle x_1\right) \log\left(c + \Delta \langle \sigma_1 \rangle x_1\right) - c - \Delta \overline{m} \mathbb{E} x_1 + \mathcal{O}(N^{-1}), \qquad (2.18)$$

where the random variable x_1 has law μ.

Lemma 2.4. *There exists $C < +\infty$ such that for any $(t,s,\mu), (t',s',\nu) \in \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathrm{Pr}[-1,1]$,*

$$|\widetilde{F}_N(t,s,\mu) - \widetilde{F}_N(t',s',\nu)| \leq C\left(|t - t'| + |s - s'| + s' W(\mu,\nu)\right). \qquad (2.19)$$

Proof. The derivative expressions (2.11) and (2.18) reveal that $\partial_t \widetilde{F}_N(t,s,\mu)$ and $\partial_s \widetilde{F}_N(t,s,\mu)$ are bounded by some constant independent of the triple (t,s,μ). By the mean value theorem and the triangle inequality it therefore suffices to show that there exists $C < +\infty$ such that for all $t,s \in \mathbb{R}_{>0}$ and $\mu,\nu \in \mathrm{Pr}[-1,1]$,

$$|\widetilde{F}_N(t,s,\mu) - \widetilde{F}_N(t,s,\nu)| \leq CsW(\mu,\nu). \qquad (2.20)$$

The fundamental theorem of calculus and the definition of the Gateaux derivative imply that

$$\widetilde{F}_N(t,s,\mu) - \widetilde{F}_N(t,s,\nu) = \overline{F}_N(t,s\mu) - \overline{F}_N(t,s\nu) = s\int_0^1 D_\mu \overline{F}_N\left(t,s\nu + us(\mu - \nu);\mu - \nu\right) \mathrm{d}u.$$

For each $u \in [0,1]$, let $f_{t,s,u}(x) := D_\mu \overline{F}_N\left(t,s\nu + us(\mu - \nu);x\right)$ in such a way that

$$\left|\widetilde{F}_N(t,s,\mu) - \widetilde{F}_N(t,s,\nu)\right| \leq s\int_0^1 \left|\int_{-1}^1 f_{t,s,u}(x)\,\mathrm{d}(\mu - \nu)(x)\right| \mathrm{d}u. \qquad (2.21)$$

The mean value theorem and the bound

$$\left|\partial_x D_\mu \overline{F}_N(t,\mu,x)\right| \leq c\left(1 + |\log(2c)| + |\log(c - |\Delta|)|\right) \qquad (2.22)$$

established for any $(t,\mu) \in \mathbb{R}_{\geq 0} \times \mathscr{M}_+$ in [18, Lemma 3.3] imply that

$$\|f_{t,s,u}\|_{\mathrm{Lip}} \leq c\left(1 + |\log(2c)| + |\log(c - |\Delta|)|\right).$$

Together with the definition of the Wasserstein distance in (2.1) and the bound (2.21), this establishes (2.20) and completes the proof. ∎

Proof of Proposition 2.1. The Lipschitz bound in Lemma 2.4, the representation (2.5) for the one-dimensional Wasserstein distance, and the Cauchy-Schwarz inequality give a constant $C < +\infty$ such that, for all $(t,s,\mu), (t',s',\nu) \in \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathrm{Pr}[-1,1]$,

$$|\widetilde{F}_N(t,s,\mu) - \widetilde{F}_N(t',s',\nu)| \leq C\left(|t - t'| + |s - s'| + s'|F_\mu - F_\nu|_{L^2}\right).$$

Remembering that for any $\mu \in \mathscr{M}_+$, we have $\overline{F}_N(t,\mu) = \widetilde{F}_N(t,\mu[-1,1],\overline{\mu})$ completes the proof. ∎

2.2 Almost everywhere differentiability of the limit free energy

The Gateaux differentiability of the free energy stated in Proposition 2.2 will be deduced from the Gateaux differentiability of the free energy thought of as a function on the space of paths

$$\mathscr{D}[0,1] := \big\{ q : [-1,1] \to [0,1] \mid q \text{ is right-continuous}$$

$$\text{and non-decreasing with } q(1) = 1 \big\} \subset L^2[-1,1]. \quad (2.23)$$

This space of paths is in one-to-one correspondence with the set of cumulative distribution functions via the mapping $\mathscr{F} : \mathrm{Pr}[-1,1] \to \mathscr{D}[0,1]$ defined by $\mathscr{F}(\mu) := F_\mu$. Through this mapping, any function $h : \mathbb{R}_{\geq 0} \times \mathscr{M}_+ \to \mathbb{R}$ may be identified with the function $\widetilde{h}' : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathscr{D}[0,1] \to \mathbb{R}$ defined by

$$\widetilde{h}'(t,s,q) := h\big(t, s\mathscr{F}^{-1}(q)\big). \quad (2.24)$$

In particular, the free energy (1.22) may be identified with the function $\widetilde{F}'_N : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathscr{D}[0,1] \to \mathbb{R}$ defined by

$$\widetilde{F}'_N(t,s,q) := \overline{F}_N\big(t, s\mathscr{F}^{-1}(q)\big). \quad (2.25)$$

Notice that the notions of Gateaux differentiability for a function $h : \mathbb{R}_{\geq 0} \times \mathscr{M}_+ \to \mathbb{R}$ and its corresponding function $\widetilde{h}' : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathscr{D}[0,1] \to \mathbb{R}$ differ. On the one hand, for each fixed $t \geq 0$, the Gateaux derivative of the function $h(t,\cdot)$ at the measure $\mu \in \mathscr{M}_+$ may be identified with the continuous density $D_\mu h(t,\mu,\cdot)$ defined according to (2.10). On the other, for each fixed $t,s \geq 0$, the Gateaux derivative of the function $\widetilde{h}'(t,s,\cdot)$ at the path $q \in \mathscr{D}[0,1]$ may be identified with the square-integrable function $\partial_q \widetilde{h}'(t,s,q,\cdot) \in L^2[-1,1]$ having the property that for all $q' \in \mathrm{Adm}(\mathscr{D}[0,1],q)$,

$$\widetilde{h}'(t,s,q+\varepsilon q') = \widetilde{h}'(t,s,q) + \varepsilon \int_{-1}^1 \partial_q \widetilde{h}'(t,s,q,u) q'(u) \, \mathrm{d}u + o(\varepsilon). \quad (2.26)$$

However, it will be important to observe that whenever $\widetilde{h}'$ is Gateaux differentiable at a point $(t,s,q) \in \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathscr{D}[0,1]$, then h must be Gateaux differentiable at the corresponding point $(t, s\mathscr{F}^{-1}(q)) \in \mathbb{R}_{\geq 0} \times \mathscr{M}_+$.

Lemma 2.5. *If the function $\widetilde{h}' : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathscr{D}[0,1] \to \mathbb{R}$ is Gateaux differentiable jointly in its three variables at the point $(t,s,q) \in \mathbb{R}_{>0} \times \mathbb{R}_{>0} \times \mathscr{D}[0,1]$, then the function $h : \mathbb{R}_{\geq 0} \times \mathscr{M}_+ \to \mathbb{R}$ used to define it according to (2.24) is Gateaux differentiable jointly in its two variables at the point $(t, s\mathscr{F}^{-1}(q)) \in \mathbb{R}_{>0} \times \mathscr{M}_+$.*

Proof. For simplicity of notation, we will drop the time variable and prove that whenever the function $\widetilde{h}' : \mathbb{R}_{\geq 0} \times \mathscr{D}[0,1] \to \mathbb{R}$ is Gateaux differentiable jointly in its two variables at the point $(s,q) \in \mathbb{R}_{>0} \times \mathscr{D}[0,1]$, then the function $h : \mathscr{M}_+ \to \mathbb{R}$ used to define it according to (2.24) without the time variable is Gateaux differentiable at the point $s\mathscr{F}^{-1}(q)$. Generalizing the proof to incorporate the time variable requires only minor changes in notation. Let $\mu := s\mathscr{F}^{-1}(q)$, and fix $v \in \mathrm{Adm}(\mathscr{M}_+,\mu)$ as well as $\varepsilon > 0$ small enough so that $\mu + \varepsilon v \in \mathscr{M}_+$. We define $s' := v[-1,1]$, and observe that by Gateaux differentiability of $\widetilde{h}'$ at $(s,q) \in \mathbb{R}_{>0} \times \mathscr{D}[0,1]$,

$$h(\mu + \varepsilon v) - h(\mu) = \widetilde{h}'\bigg(s + \varepsilon s', q + \varepsilon \frac{s' F_{\overline{v} - \overline{\mu}}}{s + \varepsilon s'} \bigg) - \widetilde{h}'(s,q)$$

$$= \varepsilon \bigg(s' \partial_s \widetilde{h}'(s,q) + \frac{s'}{s} \int_{-1}^1 \partial_q \widetilde{h}'(s,q,u) F_{\overline{v} - \overline{\mu}}(u) \, \mathrm{d}u \bigg) + o(\varepsilon).$$

14

The Fubini-Tonelli theorem implies that the second term between parentheses is given by

$$\frac{1}{s}\int_{-1}^{1}\int_{-1}^{1}\partial_q\widetilde{h}'(s,q,u)\mathbb{1}_{[-1,u]}(w)\,d\nu(w)\,du - \frac{1}{s}\int_{-1}^{1}\int_{-1}^{1}\partial_q\widetilde{h}'(s,q,u)F_{\overline{\mu}}(u)\,du\,d\nu(w)$$
$$= \frac{1}{s}\int_{-1}^{1}\left(\int_{w}^{1}\partial_q\widetilde{h}'(s,q,u)\,du - \int_{-1}^{1}\partial_q\widetilde{h}'(s,q,u)F_{\overline{\mu}}(u)\,du\right)d\nu(w).$$

It follows that

$$h(\mu+\varepsilon\nu)-h(\mu) = \varepsilon\int_{-1}^{1}D_\mu(w)\,d\nu(w)+o(\varepsilon)$$

for the continuous function $D_\mu : [-1,1] \to \mathbb{R}$ defined by

$$D_\mu(w) := \partial_s\widetilde{h}'(s,q) + \frac{1}{s}\left(\int_{w}^{1}\partial_q\widetilde{h}'(s,q,u)\,du - \int_{-1}^{1}\partial_q\widetilde{h}'(s,q,u)F_{\overline{\mu}}(u)\,du\right).$$

We have used that $\partial_q\widetilde{h}'(s,q,\cdot) \in L^2[-1,1]$ to obtain the continuity of D_μ. This establishes the Gateaux differentiability of h at $\mu \in \mathscr{M}_+$ and completes the proof. ∎

To account for the fact that we would like to establish Gateaux differentiability on a subset of $\mathbb{R}_{>0} \times \mathrm{Reg}(\mathscr{M}_+)$ that is dense in $\mathbb{R}_{\geq 0} \times \mathscr{M}_+$, we also introduce the space of paths

$$\mathscr{R}[0,1] := \left\{q \in \mathscr{D}[0,1] \mid q \text{ is smooth with strictly positive derivative}\right\}. \qquad (2.27)$$

This space of paths is in one-to-one correspondence with the probability measures in $\mathrm{Reg}(\mathscr{M}_+)$. To establish the Gateaux differentiability of a function defined on an infinite-dimensional Banach space, we will follow the arguments in [11] and rely on the notion of a Gaussian null set. A Borel subset B of a separable Banach space X is said to be a *Gaussian null set* if for every non-degenerate Gaussian measure μ on X, we have $\mu(B) = 0$. Recall that a probability measure μ on X is said to be a *non-degenerate Gaussian measure* if for every non-zero $y \in X^*$, the measure $\mu \circ y^{-1}$ is a Gaussian measure with non-zero variance. Gaussian null sets appear when invoking the following two lemmas.

Lemma 2.6. *Let $(w_n)_{n\geq 1}$ be a sequence in a separable Banach space X that has dense linear span and satisfies $\lim_{n\to+\infty}|w_n| = 0$, and let K be the closed convex hull of $\{0\}\cup\{w_n \mid n \geq 1\}$. For every $x \in X$, the set $x + K$ is not a Gaussian null set.*

Proof. This is [11, Lemma 2.5] which itself is adapted from [49, Lemma 3]. ∎

Lemma 2.7. *Let X be a separable Banach space. If a function $h : X \to \mathbb{R}$ is Lipschitz continuous, then it is Gateaux differentiable outside of a Gaussian null subset of X.*

Proof. This is [11, Theorem 2.6] which itself is a special case of [49, Theorem 6]. ∎

Lemma 2.8. *If $h : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathscr{D}[0,1] \to \mathbb{R}$ is locally Lipschitz continuous in its first two variables and uniformly Lipschitz continuous with respect to the L^2-norm in its third variable, then it is Gateaux differentiable jointly in its three variables (in the sense of (2.26)) on a subset of $\mathbb{R}_{>0} \times \mathbb{R}_{>0} \times \mathscr{R}[0,1]$ that is dense in $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathscr{D}[0,1]$.*

15

Proof. For simplicity of notation, we will instead prove that a function $h : \mathscr{D}[0,1] \to \mathbb{R}$ that is locally Lipschitz continuous with respect to the L^2-norm is Gateaux differentiable on a subset of $\mathscr{R}[0,1]$ that is dense in $\mathscr{D}[0,1]$. Our proof generalizes with only modifications in notation to the setting of a function defined on $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathscr{D}[0,1]$. Up to multiplying the function h with a cutoff function that vanishes when the total mass of the measure becomes larger than some predetermined threshold, we will also assume without loss of generality that $h : \mathscr{D}[0,1] \to \mathbb{R}$ is in fact uniformly Lipschitz continuous with respect to the L^2-norm. The proof proceeds in three steps. First, we construct a Lipschitz extension $\overline{h} : L^2[-1,1] \to \mathbb{R}$ of h, then we show that there exists a Gaussian null set $\mathscr{N}$ such that h is Gateaux differentiable on $\mathscr{D}[0,1] \setminus \mathscr{N}$, and finally we prove that $\mathscr{R}[0,1] \setminus \mathscr{N}$ is dense in $\mathscr{D}[0,1]$.

Step 1: constructing the Lipschitz extension. We denote by $P : L^2[-1,1] \to \mathscr{D}[0,1]$ the orthogonal projection onto the closed convex set $\mathscr{D}[0,1]$. We refer the reader to [19, Exercise 2.16] and its solution for a justification that this projection P is well-defined based on the fact that $\mathscr{D}[0,1]$ is a closed convex subset of $L^2[-1,1]$. It is also shown there using the convexity of $\mathscr{D}[0,1]$ that, for every $q \in L^2[-1,1]$ and $p \in \mathscr{D}[0,1]$,

$$\langle q - P(q), p - P(q) \rangle_{L^2} \leq 0.$$

It follows that for all $q, q' \in L^2[-1,1]$,

$$\langle q - P(q), P(q') - P(q) \rangle_{L^2} \leq 0 \leq \langle q' - P(q'), P(q') - P(q) \rangle_{L^2}.$$

Rearranging and using the Cauchy-Schwarz inequality yields

$$|P(q) - P(q')|_{L^2}^2 \leq \langle q' - q, P(q') - P(q) \rangle_{L^2} \leq |q' - q|_{L^2} |P(q) - P(q')|_{L^2}$$

from which it follows that

$$|P(q) - P(q')|_{L^2} \leq |q' - q|_{L^2}.$$

This means that the extension $\overline{h} : L^2[-1,1] \to \mathbb{R}$ of h defined by

$$\overline{h}(q) := h(P(q))$$

is Lipschitz continuous.

Step 2: Gateaux differentiability outside a Gaussian null set. By Lemma 2.7, there exists a Gaussian null set $\mathscr{N} \subset L^2[-1,1]$ such that $\overline{h}$ is Gateaux differentiable on $L^2[-1,1] \setminus \mathscr{N}$. We now fix $q \in \mathscr{D}[0,1] \setminus \mathscr{N}$, and prove that h is Gateaux differentiable at q. Let $y \in (L^2[-1,1])^* \simeq L^2[-1,1]$ denote the Gateaux derivative of $\overline{h}$ at q. The limit $Dh(q;x)$ exists for every $x \in \mathrm{Adm}(\mathscr{D}[0,1],q)$, and it is equal to $\langle y, x \rangle_{L^2}$. Let us now assume that there exists $y' \in L^2[-1,1]$ with $\langle y', x \rangle_{L^2} = \langle y, x \rangle_{L^2}$ for all $x \in \mathrm{Adm}(\mathscr{D}[0,1],q)$, and prove that we must have $y' = y$. Observe that $\mathscr{D}[0,1] - q \subset \mathrm{Adm}(\mathscr{D}[0,1],q)$ by convexity of $\mathscr{D}[0,1]$. Moreover, the orthogonal complement of $\mathscr{D}[0,1] - q$ in $L^2[-1,1]$ is trivial. Indeed, suppose that $z \in L^2[-1,1]$ is such that $\langle z, q' \rangle_{L^2} = \langle z, q \rangle_{L^2}$ for all $q' \in \mathscr{D}[0,1]$. If $\mu' \in \mathrm{Pr}[-1,1]$ is such that $q' = F_{\mu'}$, then the Fubini-Tonelli theorem implies that

$$\langle z, q' \rangle_{L^2} = \int_{-1}^{1} z(u)\mu'[-1,u]\,\mathrm{d}u = \int_{-1}^{1} \int_{-1}^{u} z(u)\,\mathrm{d}\mu'(v)\,\mathrm{d}u = \int_{-1}^{1} \int_{v}^{1} z(u)\,\mathrm{d}u\,\mathrm{d}\mu'(v) = 0.$$

16

Since $\mathscr{D}[0,1]$ is in one-to-one correspondence with $\mathrm{Pr}[-1,1]$, this means that for all $\mu' \in \mathrm{Pr}[-1,1]$,

$$\int_{-1}^{1} \int_{v}^{1} z(u) \, du \, d\mu'(v) = \langle z, q \rangle_{L^2}.$$

Choosing $\mu' = \delta_v$ for each $v \in [-1,1]$, and differentiating the resulting expression with respect to v shows that $z = 0$. This means that the closed linear span of $\mathscr{D}[0,1] - q$ is $L^2[-1,1]$ and that $\langle y', x \rangle_{L^2} = \langle y, x \rangle_{L^2}$ for all $x \in L^2[-1,1]$. It follows that $y' = y$, and therefore that h is Gateaux differentiable on $\mathscr{D}[0,1] \setminus \mathscr{N}$.

Step 3: density of $\mathscr{R}[0,1] \setminus \mathscr{N}$ in $\mathscr{D}[0,1]$. The set $\mathscr{R}[0,1]$ has dense linear span in $L^2[-1,1]$, so the separability of $L^2[-1,1]$ allows us to find a sequence $(w_n)_{n \geq 1} \subset \mathscr{R}[0,1]$ whose linear span is dense in $L^2[-1,1]$. Re-scaling each w_n, we can assume without loss of generality that $|w_n|_{L^2} \leq 1$ and that $\lim_{n \to +\infty} |w_n|_{L^2} = 0$. Let K be as in Lemma 2.6, and observe that $K \subset \mathscr{R}[0,1]$ by convexity of $\mathscr{R}[0,1]$. At this point, fix $q \in \mathscr{R}[0,1]$ as well as $\varepsilon > 0$. Since $q + \varepsilon K$ is not a Gaussian null set, we have that $q + \varepsilon K \not\subset \mathscr{N}$. In particular, there exists $y \in q + \varepsilon K \setminus \mathscr{N}$. Notice that for any $\kappa \in K$, we have $(q + \varepsilon \kappa)' > 0$ so in fact $y \in \mathscr{R} \setminus \mathscr{N}$ with $|y - q|_{L^2} \leq \varepsilon$. This establishes the density of $\mathscr{R}[0,1] \setminus \mathscr{N}$ in $\mathscr{R}[0,1]$. Together with the density of $\mathscr{R}[0,1]$ in $\mathscr{D}[0,1]$ which can be obtained through mollifiers, this completes the proof. ∎

Proof of Proposition 2.2. We denote by $\widetilde{f}' : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathscr{D}[0,1] \to \mathbb{R}$ the limit of the sequence of extended free energies $(\widetilde{F}'_N)_{N \geq 1}$ defined in (2.25) along the subsequence $(N_k)_{k \geq 0}$. Invoking Lemma 2.8, we can find a dense subset $\mathscr{A}$ of $\mathbb{R}_{>0} \times \mathbb{R}_{>0} \times \mathscr{R}[0,1]$ on which $\widetilde{f}'$ is Gateaux differentiable jointly in its three variables. By Lemma 2.5, the limit free energy $f : \mathbb{R}_{\geq 0} \times \mathscr{M}_+ \to \mathbb{R}$ is Gateaux differentiable jointly in its two variables on the set

$$\mathscr{B} := \left\{ \left(t, s\mathscr{F}^{-1}(q) \right) \mid (t,s,q) \in \mathscr{A} \right\} \subset \mathbb{R}_{\geq 0} \times \mathrm{Reg}(\mathscr{M}_+). \tag{2.28}$$

We now show that this set is dense in $\mathbb{R}_{\geq 0} \times \mathscr{M}_+$. We fix $(t,\mu) \in \mathbb{R}_{\geq 0} \times \mathscr{M}_+$, and aim to show that there is a sequence $(t_n, \mu_n)_{n \geq 1} \subset \mathscr{B}$ such that $t_n \to t$ and $\mu_n \to \mu$ weakly. Let $s := \mu[-1,1]$, and consider the point $(t, s, F_{\overline{\mu}}) \in \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathscr{D}[0,1]$. There exists a sequence $(t_n, s_n, q_n)_{n \geq 1} \subset \mathscr{A}$ converging to $(t, s, F_{\overline{\mu}})$, and we define $\mu_n := s_n \mathscr{F}^{-1}(q_n)$. By compactness, the sequence μ_n admits a subsequential limit. Since $F_{\mu_n} = s_n q_n \to s F_{\overline{\mu}} = F_{\mu}$ in L^2, the only possible subsequential limit is μ so indeed $\mu_n \to \mu$ weakly. ∎

2.3 Convergence of the derivative free energy

The convergence of the derivative free energy stated in Proposition 2.3 will be deduced from a convexity property of the free energy. Although the enriched free energy (1.22) is known not to be convex [25], we now use the information-theoretic arguments in [25, Proposition 4.1] to show that it is convex in certain directions. Together with a computation similar to that in [11, Proposition 5.4] we then establish Proposition 2.3.

Proposition 2.9. *For all $(t,\mu) \in \mathbb{R}_{\geq 0} \times \mathscr{M}_+$ and $v \in \mathrm{Pr}[-1,1]$, the map $r \mapsto \overline{F}_N(t, \mu + rv)$ is convex.*

Proof. The superposition principle for Poisson point processes implies that, jointly in $\sigma \in \Sigma_N$,

$$H_N^{\mu+r\nu}(\sigma) \overset{d}{=} H_N^\mu(\sigma) + H_N^{r\nu}(\sigma),$$

where we understand that the Hamiltonians $H_N^\mu(\sigma)$ and $H_N^{r\nu}(\sigma)$ are independent conditionally on the ground truth σ^*. Conditionally on the ground truth σ^*, the Poisson point processes $(\Lambda_i(N\mu))_{i \leq N}$, and the Poisson random variable Π_t, we define the "prior" measure on Σ_N by

$$P_{t,\mu}^*(\sigma) := \exp\left(H_N^t(\sigma) + H_N^\mu(\sigma)\right) P_N^*(\sigma)$$

in such a way that

$$\overline{F}_N(t,\mu+r\nu) = \frac{1}{N}\mathbb{E}\log \int_{\Sigma_N} \exp H_N^{r\nu}(\sigma)\,\mathrm{d}P_{t,\mu}^*(\sigma).$$

The measure $P_{t,\mu}^*$ is not necessarily a probability measure, but this will not have any significant effect. The proof now proceeds in two steps. First, we relate the free energy $\overline{F}_N(t,\mu+r\nu)$ to a mutual information $I_N(r)$, and argue that $r \mapsto \overline{F}_N(t,\mu+r\nu)$ is convex in r if and only if this mutual information is concave. Then, we compute the second derivative of the mutual information, and we show that it is non-positive.

Step 1: introducing a mutual information. Conditionally on the ground truth σ^*, the Poisson point processes $(\Lambda_i(N\mu))_{i \leq N}$, and the Poisson random variable Π_t, we sample a signal $\sigma^\diamond$ from the normalized prior distribution $\overline{P}_{t,\mu}^*$ as well as a family $\Lambda^{r\nu} := (\Lambda_i(Nr\nu))_{i \leq N}$ of Poisson point processes. Conditionally on the family of Poisson point processes $\Lambda^{r\nu}$, we sample the Bernoulli random variables $\widetilde{\mathbf{G}}_N := \{\widetilde{G}_i^x \mid x \in \Lambda_i(Nr\nu), i \leq N\}$ with conditional law

$$\mathbb{P}\left\{\widetilde{G}_i^x = 1 \mid \sigma^\diamond, x\right\} := \frac{c + \Delta\sigma_i^\diamond x}{N}.$$

The likelihood of the model is given by

$$\mathbb{P}\left\{\widetilde{\mathbf{G}}_N = (\widetilde{G}_i^x) \mid \sigma^\diamond = \sigma\right\} = \prod_{i < N}\prod_{x \in \Lambda_i(Nr\nu)} \left(\frac{c+\Delta\sigma_i x}{N}\right)^{\widetilde{G}_i^x}\left(1 - \frac{c+\Delta\sigma_i x}{N}\right)^{1-\widetilde{G}_i^x},$$

and Bayes' formula implies that the posterior of the model is the Gibbs measure

$$\mathbb{P}\left\{\sigma^\diamond = \sigma \mid \widetilde{\mathbf{G}}_N = (\widetilde{G}_i^x)\right\} = \frac{\exp\left(H_N^\diamond(\sigma)\right) P_{t,\mu}^*(\sigma)}{\int_{\Sigma_N}\exp\left(H_N^\diamond(\tau)\right)\mathrm{d}P_{t,\mu}^*(\tau)}$$

associated with the Hamiltonian

$$H_N^\diamond(\sigma) := \sum_{i \leq N}\sum_{x \in \Lambda_i(Nr\nu)}\log\left[\left(c+\Delta\sigma_i x\right)^{\widetilde{G}_i^x}\left(1 - \frac{c+\Delta\sigma_i x}{N}\right)^{1-\widetilde{G}_i^x}\right]. \tag{2.29}$$

It follows that

$$I_N(r) := \frac{1}{N}I\left(\widetilde{\mathbf{G}}_N;\sigma^\diamond\right) = \mathbb{E}\sum_{x \in \Lambda_1(Nr\nu)}\log\left[\left(c+\Delta\sigma_1^\diamond x\right)^{\widetilde{G}_1^x}\left(1 - \frac{c+\Delta\sigma_1^\diamond x}{N}\right)^{1-\widetilde{G}_1^x}\right] - \overline{F}_N^\diamond(r) \tag{2.30}$$

18

for the free energy defined on Σ_N by

$$\overline{F}_N^\diamond(r) := \frac{1}{N}\mathbb{E}\log\int_{\Sigma_N}\exp H_N^\diamond(\sigma)\,dP_{t,\mu}^*(\sigma).$$

Together with [19, Proposition 5.12] on Poisson averages, this implies that

$$I_N(r) = Nr\,\mathbb{E}\int_{-1}^1\log\left[(c+\Delta\sigma_1^\diamond x)^{\widetilde{G}_1^x}\left(1-\frac{c+\Delta\sigma_1^\diamond x}{N}\right)^{1-\widetilde{G}_1^x}\right]d\nu(x) - \overline{F}_N^\diamond(r),$$

so the free energy $\overline{F}_N^\diamond(r)$ is convex if and only if the mutual information $I_N(r)$ is concave. Averaging the free energy $\overline{F}_N^\diamond(r)$ over the randomness of the ground truth σ^*, the Poisson point processes $(\Lambda_i(N\mu))_{i\leq N}$, and the Poisson random variable Π_t yields the free energy $\overline{F}_N(t,\mu+r\nu)$. It therefore suffices to prove that $r\mapsto I_N(r)$ is concave.

Step 2: showing the mutual information is concave. The properties of Poisson point processes imply that the Hamiltonian (2.29) may be written as

$$H_N^\diamond(\sigma) = \sum_{i\leq N}\sum_{k\leq\Pi_{i,r}}\log\left[(c+\Delta\sigma_i x_{i,k})^{\widetilde{G}_{i,k}}\left(1-\frac{c+\Delta\sigma_i x_{i,k}}{N}\right)^{1-\widetilde{G}_{i,k}}\right],$$

where the random variables $(x_{i,k})_{i,k\geq 1}$ are i.i.d. with law ν, the random variables $(\Pi_{i,r})_{i\leq N}$ are independent Poisson random variables with mean Nr, and $(\widetilde{G}_{i,k})_{i,k\geq 1}$ are Bernoulli random variables with conditional distribution

$$\mathbb{P}\{\widetilde{G}_{i,k}=1\mid\sigma^\diamond,x_{i,k}\} := \frac{c+\Delta\sigma_i^\diamond x_{i,k}}{N}.$$

This means that the mutual information (2.30) depends on r only through the Poisson random variable $\Pi_{i,r}$. With this in mind, given $i,j\leq N$ and $L_1,L_2\geq 0$, introduce the conditional mutual information functionals

$$I_N^i(L_1) := \frac{1}{N}I\big(\widetilde{\mathbf{G}}_N;\sigma^\diamond\mid\Pi_{i,r}=L_1\big)\quad\text{and}\quad I_N^{i,j}(L_1,L_2) := I\big(\widetilde{\mathbf{G}}_N;\sigma^\diamond\mid\Pi_{i,r}=L_1,\Pi_{j,r}=L_2\big).$$

A direct computation leveraging the product rule shows that

$$\partial_r I_N(r) = \sum_{i\leq N}\sum_{L\geq 0}\pi(rN,L)\big(I_N^i(L+1)-I_N^i(L)\big).$$

Another direct computation leveraging the product rule reveals that

$$\begin{aligned}
\partial_r^2 I_N(r) = {}&N\sum_{i\leq N}\sum_{L_1\geq 0}\pi(rN,L_1)\big(I_N^i(L_1+2)-2I_N^i(L_1+1)+I_N^i(L_1)\big)\\
&+N\sum_{i\leq N}\sum_{j\neq i}\sum_{L_1,L_2\geq 0}\pi(rN,L_1)\pi(rN,L_2)\big(I_N^{i,j}(L_1+1,L_2+1)-I_N^{i,j}(L_1+1,L_2)\\
&\hspace{6cm}-I_N^{i,j}(L_1,L_2+1)+I_N^{i,j}(L_1,L_2)\big).
\end{aligned}$$

It therefore suffices to show that for fixed $i \neq j \leq N$ and $L_1, L_2 \geq 0$,

$$I_N^i(L_1 + 2) - 2I_N^i(L_1 + 1) + I_N^i(L_1) \leq 0, \qquad (2.31)$$

$$I_N^{i,j}(L_1 + 1, L_2 + 1) - I_N^{i,j}(L_1 + 1, L_2) - I_N^{i,j}(L_1, L_2 + 1) + I_N^{i,j}(L_1, L_2) \leq 0. \qquad (2.32)$$

We will only prove (2.32) as the proof of (2.31) is similar. To alleviate notation, conditionally on the events $\Pi_{i,r} = L_1$ and $\Pi_{j,r} = L_2$ introduce the random vectors

$$\widetilde{\mathbf{G}}_N' := \left\{ \widetilde{G}_{m,k} \mid m \neq i,\ m \neq j \text{ and } k \leq \Pi_{m,r} \right\} \quad \text{and} \quad Z := \left((\widetilde{G}_{i,\ell})_{\ell \leq L_1}, (\widetilde{G}_{j,\ell})_{\ell \leq L_2}, \widetilde{\mathbf{G}}_N' \right).$$

By the chain rule for mutual information,

$$I_N^{i,j}(L_1 + 1, L_2 + 1) = I((\widetilde{G}_{i,L_1+1}, \widetilde{G}_{j,L_2+1}), Z; \sigma^\diamond) = I((\widetilde{G}_{i,L_1+1}, \widetilde{G}_{j,L_2+1}); \sigma^\diamond \mid Z) + I(Z; \sigma^\diamond),$$

$$I_N^{i,j}(L_1 + 1, L_2) = I(\widetilde{G}_{i,L_1+1}, Z; \sigma^\diamond) = I(\widetilde{G}_{i,L_1+1}; \sigma^\diamond \mid Z) + I(Z; \sigma^\diamond),$$

$$I_N^{i,j}(L_1, L_2 + 1) = I(\widetilde{G}_{j,L_2+1}, Z; \sigma^\diamond) = I(\widetilde{G}_{j,L_2+1}; \sigma^\diamond \mid Z) + I(Z; \sigma^\diamond).$$

It follows that (2.32) is equivalent to the inequality

$$I((\widetilde{G}_{i,L_1+1}, \widetilde{G}_{j,L_2+1}); \sigma^\diamond \mid Z) - I(\widetilde{G}_{i,L_1+1}; \sigma^\diamond \mid Z) - I(\widetilde{G}_{j,L_2+1}; \sigma^\diamond \mid Z) \leq 0. \qquad (2.33)$$

Another application of the chain rule for mutual information reveals that

$$I((\widetilde{G}_{i,L_1+1}, \widetilde{G}_{j,L_2+1}); \sigma^\diamond \mid Z) = I(\widetilde{G}_{i,L_1+1}; \sigma^\diamond \mid Z) + I(\widetilde{G}_{j,L_2+1}; \sigma^\diamond \mid Z, \widetilde{G}_{i,L_1+1})$$

and that the second term on the right side of this expression is given by

$$I(\widetilde{G}_{j,L_2+1}; (\sigma^\diamond, \widetilde{G}_{i,L_1+1}) \mid Z) - I(\widetilde{G}_{j,L_2+1}; \widetilde{G}_{i,L_1+1} \mid Z)$$
$$= I(\widetilde{G}_{j,L_2+1}; \sigma^\diamond \mid Z) + I(\widetilde{G}_{j,L_2+1}; \widetilde{G}_{i,L_1+1} \mid Z, \sigma^\diamond) - I(\widetilde{G}_{j,L_2+1}; \widetilde{G}_{i,L_1+1} \mid Z).$$

It follows that the left side of (2.33) is

$$I(\widetilde{G}_{j,L_2+1}; \widetilde{G}_{i,L_1+1} \mid Z, \sigma^\diamond) - I(\widetilde{G}_{j,L_2+1}; \widetilde{G}_{i,L_1+1} \mid Z) = -I(\widetilde{G}_{j,L_2+1}; \widetilde{G}_{i,L_1+1} \mid Z) \leq 0,$$

where we have used that, conditionally on $\sigma^\diamond$, the random variables $\widetilde{G}_{j,L_2+1}$, $\widetilde{G}_{i,L_1+1}$, and Z are independent to assert that $I(\widetilde{G}_{j,L_2+1}; \widetilde{G}_{i,L_1+1} \mid Z, \sigma^\diamond) = 0$. This establishes (2.33) and completes the proof. ∎

Proof of Proposition 2.3. Fix $r, \eta > 0$ and a probability measure $\nu \in \Pr[-1, 1] \cap \mathrm{Reg}(\mathscr{M}_+)$. We define the measure $\nu_\eta \in \mathrm{Reg}(\mathscr{M}_+)$ by $\nu_\eta := \eta \nu$. The convexity property of the enriched free energy in Proposition 2.9 gives the upper bound

$$\overline{F}_N(t, \mu + r\nu_\eta) = \overline{F}_N(t, \mu + (r \cdot 1 + (1-r) \cdot 0)\nu_\eta) \leq r\overline{F}_N(t, \mu + \nu_\eta) + (1-r)\overline{F}_N(t, \mu).$$

Subtracting $\overline{F}_N(t, \mu)$ from both sides of this inequality, dividing by r, and letting r tend to zero yields

$$\eta D_\mu \overline{F}_N(t, \mu; \nu) \leq \overline{F}_N(t, \mu + \nu_\eta) - \overline{F}_N(t, \mu).$$

If $D_{t,\mu}$ denotes a weak subsequential limit of $D_\mu \overline{F}_{N_K}(t, \mu, \cdot)$, then letting K tend to infinity, dividing by η, and letting η tend to zero gives

$$D_{t,\mu}(\nu) \leq D_\mu f(t, \mu; \nu).$$

Repeating the argument with η replaced by $-\eta$, which is possible as $\mu, \nu \in \mathrm{Reg}(\mathscr{M}_+)$, shows that for any $\nu \in \mathrm{Reg}(\mathscr{M}_+)$, we have $D_{t,\mu}(\nu) = D_\mu f(t, \mu; \nu)$. By density of $\mathrm{Reg}(\mathscr{M}_+)$ in $\mathscr{M}_+$, this completes the proof. ∎

3 Multioverlap concentration through perturbation

In the next section we will perform the cavity computation that gives rise to the functional (1.26) and leads to the critical point representation in Theorem 1.1. This cavity computation will require an understanding of the asymptotic behavior of the array $(\sigma_i^\ell)_{i,\ell \geq 1}$ of i.i.d. replicas sampled from the Gibbs measure (2.13). In this section, we leverage the main result in [8] to argue that, up to a small perturbation of the enriched Hamiltonian (1.21) that does not affect the limit of the enriched free energy (1.22), the asymptotic spin array $(\sigma_i^\ell)_{i,\ell \geq 1}$ is generated by some probability measure $v \in \mathscr{M}_p$ according to (S1)-(S2). In particular, as we will see in Lemma 3.4, this means that the average of a product of finitely many replicas is asymptotically equal to a product of moments of v. More precisely, given any finite set of n replicas and collection $\{\mathscr{C}_\ell\}_{\ell \leq n}$ of finite indices,

$$\lim_{N \to +\infty} \mathbb{E} \prod_{\ell \leq n} \left\langle \prod_{i \in \mathscr{C}_\ell} \sigma_i^\ell \right\rangle = \prod_{\ell \leq n} \mathbb{E} x_\ell^{|\mathscr{C}_\ell|}, \tag{3.1}$$

where the $(x_\ell)_{\ell \geq 1}$ are i.i.d. with law v. This observation will be crucial in Lemma 4.2, where we study the asymptotic behavior of the cavity representation and first arrive at the functional (1.26).

The perturbation Hamiltonians that will enforce the rigid asymptotic structure described by (S1)-(S2) depend on a perturbation parameter $\lambda := (\lambda_k)_{k \geq 0}$ with $\lambda_k \in [2^{-k-1}, 2^{-k}]$ for $k \geq 0$, and they enforce the concentration of the multioverlaps

$$R_{1,\ldots,n} := \frac{1}{N} \sum_{i \leq N} \sigma_i^1 \cdots \sigma_i^n. \tag{3.2}$$

As will be explained below, here $(\sigma_i^\ell)_{i,\ell \geq 1}$ denote i.i.d. replicas from the Gibbs measure associated with the sum of the Hamiltonian (1.21) and the two perturbation Hamiltonians that we now introduce. The first perturbation Hamiltonian will ensure the concentration of the overlap $R_{1,2}$. Given a sequence $(\varepsilon_N)_{N \geq 1}$ with $\varepsilon_N := N^\gamma$ for some $-1/8 < \gamma < 0$ and a standard Gaussian vector $Z_0 := (Z_{0,1}, \ldots, Z_{0,N})$ in $\mathbb{R}^N$, we introduce the Gaussian perturbation Hamiltonian

$$H_N^{\text{gauss}}(\sigma, \lambda_0) := \mathscr{H}_0 := \sum_{i \leq N} \left(\lambda_0 \varepsilon_N \sigma_i^* \sigma_i + \sqrt{\lambda_0 \varepsilon_N} Z_{0,i} \sigma_i \right) \tag{3.3}$$

associated with the task of recovering the signal σ^* from the data

$$Y^{\text{gauss}} := \sqrt{\lambda_0 \varepsilon_N} \sigma^* + Z_0. \tag{3.4}$$

Notice that

$$1 \geq \varepsilon_N \to 0 \quad \text{and} \quad N \varepsilon_N \to +\infty. \tag{3.5}$$

The second perturbation Hamiltonian will ensure the validity of the Franz-de Sanctis identities [8, 18] which together with the concentration of the overlap $R_{1,2}$ give the concentration of all the multioverlaps (3.2). We consider a sequence $(s_N)_{N \geq 1}$ with $s_N := N^\eta$ for some $4/5 < \eta < 1$, so that in particular

$$\frac{s_N}{N} \to 0 \quad \text{and} \quad \frac{s_N}{\sqrt{N}} \to +\infty. \tag{3.6}$$

We fix a sequence of i.i.d. random variables $(\pi_k)_{k\geq 1}$ with $\mathrm{Poi}(s_N)$ distribution as well as a sequence $e := (e_{jk})_{j,k\geq 1}$ of random variables with $\mathrm{Exp}(1)$ distribution and a family $(i_{jk})_{j,k\geq 1}$ of random indices uniformly sampled from the set $\{1,\ldots,N\}$. We define the exponential perturbation Hamiltonian by

$$H_N^{\mathrm{exp}}(\sigma) := \sum_{k\geq 1} \mathcal{H}_k, \quad \text{where} \quad \mathcal{H}_k := \sum_{j\leq \pi_k} \left(\log(1+\lambda_k \sigma_{i_{jk}}) - \frac{\lambda_k e_{jk} \sigma_{i_{jk}}}{1+\lambda_k \sigma_{i_{jk}}^*} \right). \tag{3.7}$$

This is the Hamiltonian associated with the task of recovering the signal σ^* from the independently generated data

$$Y^{\mathrm{exp}} := \left\{ \frac{e_{jk}}{1+\lambda_k \sigma_{i_{jk}}^*} \mid j \leq \pi_k \text{ and } k \geq 1 \right\}. \tag{3.8}$$

For each $(t,\mu) \in \mathbb{R}_{\geq 0} \times \mathcal{M}_+$, the perturbed enriched Hamiltonian is the sum of the enriched Hamiltonian (1.21) and the perturbation Hamiltonians just defined,

$$H_N^{t,\mu}(\sigma,\lambda) := H_N^{t,\mu}(\sigma) + H_N^{\mathrm{gauss}}(\sigma,\lambda_0) + H_N^{\mathrm{exp}}(\sigma,\lambda). \tag{3.9}$$

As usual, we denote by

$$F_N(t,\mu,\lambda) := \frac{1}{N} \log \int_{\Sigma_N} \exp H_N^{t,\mu}(\sigma,\lambda) \, \mathrm{d}P_N^*(\sigma) \quad \text{and} \quad \overline{F}_N(t,\mu,\lambda) := \mathbb{E} F_N(t,\mu,\lambda) \tag{3.10}$$

its associated free energy functionals, and by $\langle \cdot \rangle$ its corresponding Gibbs average. This means that for any bounded and measurable function $f = f(\sigma^1,\ldots,\sigma^n)$ of finitely many replicas,

$$\langle f(\sigma^1,\ldots,\sigma^n) \rangle := \langle f \rangle := \frac{\int_{\Sigma_N^n} f(\sigma^1,\ldots,\sigma^n) \prod_{\ell \leq n} \exp H_N^{t,\mu}(\sigma^\ell,\lambda) \, \mathrm{d}P_N^*(\sigma^\ell)}{\left(\int_{\Sigma_N} \exp H_N^{t,\mu}(\sigma,\lambda) \, \mathrm{d}P_N^*(\sigma) \right)^n}. \tag{3.11}$$

Since the Gibbs measure associated with the perturbed Hamiltonian (3.9) is a conditional expectation, it satisfies the Nishimori identity (2.14).

Lemma 3.1. *For all $(t,\mu) \in \mathbb{R}_{\geq 0} \times \mathcal{M}_+$ and perturbation parameter λ, the enriched free energy* (1.22) *and the perturbed free energy* (3.10) *are asymptotically equivalent,*

$$\lim_{N\to+\infty} \left| \overline{F}_N(t,\mu,\lambda) - \overline{F}_N(t,\mu) \right| = 0. \tag{3.12}$$

Proof. This corresponds to [18, Lemma 4.1] but we provide a proof for completeness. A direct computation reveals that

$$\left| \overline{F}_N(t,\mu,\lambda) - \overline{F}_N(t,\mu) \right| \leq \frac{1}{N} \mathbb{E} \max_{\sigma \in \Sigma_N} \left| H_N^{\mathrm{gauss}}(\sigma,\lambda_0) \right| + \frac{1}{N} \mathbb{E} \max_{\sigma \in \Sigma_N} \left| H_N^{\mathrm{exp}}(\sigma,\lambda) \right|.$$

For any spin configuration $\sigma \in \Sigma_N$,

$$\left| H_N^{\mathrm{gauss}}(\sigma,\lambda_0) \right| \leq N\varepsilon_N + \sqrt{\varepsilon_N} \sum_{i\leq N} |Z_{0,i}|$$

22

and

$$\left|H_N^{\exp}(\sigma,\lambda)\right| \leq \sum_{1\leq k\leq K'}\sum_{j\leq \pi_k}\left(\log(1+\lambda_k)+\frac{\lambda_k e_{jk}}{1-\lambda_k}\right).$$

Since these bounds are uniform in σ, it follows that

$$\left|\overline{F}_N(t,\mu,\lambda)-\overline{F}_N(t,\mu)\right| \leq \varepsilon_N + \sqrt{\varepsilon_N}\,\mathbb{E}Z_{0,1} + \frac{s_N}{N}\sum_{k\geq 1}\left(\log(1+\lambda_k)+\frac{\lambda_k}{1-\lambda_k}\right).$$

The third term was obtained by taking the expectation with respect to the randomness of e, and then with respect to the randomness of $(\pi_k)_{k\geq 1}$. Leveraging (3.5)-(3.6) to let N tend to infinity completes the proof. ∎

Lemma 3.2. *For every $n\geq 1$, the multioverlap $R_{1,\dots,n}$ in (3.2) associated with the perturbed Gibbs measure (3.11) concentrates on average over the perturbation parameters,*

$$\lim_{N\to+\infty}\mathbb{E}_\lambda\,\mathbb{E}\left\langle (R_{1,\dots,n}-\mathbb{E}\langle R_{1,\dots,n}\rangle)^2\right\rangle = 0, \tag{3.13}$$

where $\mathbb{E}_\lambda$ denotes the average with respect to the uniform random variables λ_k on $[2^{-k-1},2^{-k}]$.

Proof. We introduce the concentration function

$$v_N := \frac{1}{N}\sup\left\{\mathbb{E}\left|F_N(t,\mu,\lambda)-\overline{F}_N(t,\mu,\lambda)\right|^2 \mid \lambda_k \in [2^{-k-1},2^{-k}] \text{ for all } k\geq 0\right\}, \tag{3.14}$$

and think of the perturbation parameters $\lambda := (\lambda_k)_{k\geq 0}$ as being sampled independently with λ_k uniform on $[2^{-k-1},2^{-k}]$ for $k\geq 0$. The concentration function v_N is bounded from above by a constant independent of N by [18, Proposition B.4], so (3.13) follows from [8, Theorem 2.2]. This completes the proof. ∎

Arguing as in [43, Lemma 3.3], it is now possible to find a deterministic sequence $(\lambda^N)_{N\geq 1}$ along which, for all $n\geq 1$,

$$\lim_{N\to+\infty}\mathbb{E}\left\langle (R_{1,\dots,n}-\mathbb{E}\langle R_{1,\dots,n}\rangle_N)^2\right\rangle_N = 0. \tag{3.15}$$

Here, the Gibbs average $\langle\cdot\rangle_N$ for a system of size N is associated with the perturbation parameter λ^N, and the multioverlap $R_{1,\dots,n}$ is for i.i.d. replicas $(\sigma_i^\ell)_{i,\ell\geq 1}$ sampled from this perturbed Gibbs measure. The multioverlap concentration (3.15) ensures that the asymptotic spin array is generated by some probability measure $\nu \in \mathscr{M}_p$ according to (S1)-(S2).

Proposition 3.3. *Let $(\lambda^N)_{N\geq 1}$ be a deterministic sequence along which (3.15) holds for all $n\geq 1$, and denote by $(\sigma_i^\ell)_{i,\ell\geq 1}$ i.i.d. replicas sampled from the Gibbs measure $\langle\cdot\rangle_N$ in (3.11) associated with the perturbation parameter λ^N. For any subsequence $(N_k)_{k\geq 1}$ along which the distribution of the spin array $(\sigma_i^\ell)_{i,\ell\geq 1}$ converges in the sense of finite-dimensional distributions, there exists a probability measure $\nu \in \mathscr{M}_p$ (which may depend on the subsequence) such that the asymptotic spin array is generated by ν according to (S1)-(S2).*

Proof. This is immediate from [8, Corollary 2.3]. We have implicitly used that $\mathbb{E}\langle R_1 \rangle = \mathbb{E}\,\sigma_1^* = \overline{m}$ by the Nishimori identity to assert that $v \in \mathscr{M}_p$. ∎

The two-step process (**S1**)-(**S2**) to generate the asymptotic spin array from the measure $v \in \mathscr{M}_p$ allows us to write the asymptotic spin array $(\sigma_i^\ell)_{i,\ell \geq 1}$ explicitly in terms of the i.i.d. random variables $(x_i)_{i \geq 1}$ sampled from v and a family of i.i.d. uniform random variables $(u_{i,\ell})_{i,\ell \geq 1}$ on $[0,1]$ also independent of $(x_i)_{i \geq 1}$,

$$\sigma_i^\ell := 2\mathbb{1}\left\{ u_{i,\ell} \leq \frac{1+x_i}{2} \right\} - 1. \tag{3.16}$$

From this representation of the asymptotic spins we can readily establish (3.1).

Lemma 3.4. *Let $(\lambda^N)_{N \geq 1}$ be a deterministic sequence along which (3.15) holds for all $n \geq 1$, denote by $(\sigma_i^\ell)_{i,\ell \geq 1}$ i.i.d. replicas sampled from the Gibbs measure $\langle \cdot \rangle_N$ in (3.11) associated with the perturbation parameter λ^N, and suppose that the distributional limit of the spin array $(\sigma_i^\ell)_{i,\ell \geq 1}$ is generated by a probability measure $v \in \mathscr{M}_p$ according to (**S1**)-(**S2**). If $(x_\ell)_{\ell \geq 1}$ denote i.i.d. samples from v, then for any finite set of n replicas and collection $\{\mathscr{C}_\ell\}_{\ell \leq n}$ of finite indices,*

$$\lim_{N \to +\infty} \mathbb{E} \prod_{\ell \leq n} \left\langle \prod_{i \in \mathscr{C}_\ell} \sigma_i^\ell \right\rangle_N = \prod_{\ell \leq n} \mathbb{E} x_\ell^{|\mathscr{C}_\ell|}. \tag{3.17}$$

Proof. Let $\mathscr{C} := \{(i,\ell) \mid \ell \leq n \text{ and } i \in \mathscr{C}_\ell\}$, and observe that

$$\mathbb{E} \prod_{\ell \leq n} \left\langle \prod_{i \in \mathscr{C}_\ell} \sigma_i^\ell \right\rangle_N = \mathbb{E} \left\langle \prod_{(i,\ell) \in \mathscr{C}} \sigma_i^\ell \right\rangle_N.$$

Recalling the representation (3.16) of an asymptotic spin in terms of a family $(u_{i,\ell})_{i,\ell \geq 1}$ of i.i.d. uniform random variables on $[0,1]$, and letting N tend to infinity in the previous display reveals that

$$\lim_{N \to +\infty} \mathbb{E} \prod_{\ell \leq n} \left\langle \prod_{i \in \mathscr{C}_\ell} \sigma_i^\ell \right\rangle_N = \mathbb{E} \prod_{(i,\ell) \in \mathscr{C}} \left(2\mathbb{1}\left\{ u_{i,\ell} \leq \frac{1+x_i}{2} \right\} - 1 \right) = \mathbb{E} \prod_{(i,\ell) \in \mathscr{C}} x_i,$$

where the independence of the $(u_{i,\ell})_{i,\ell \geq 1}$ has played its part in the second equality. Leveraging the independence of the $(x_i)_{i \geq 1}$ completes the proof. ∎

This formula to compute the asymptotic moments of the spin array will be applied to a slight modification of the Hamiltonian (1.21) in Section 4. The arguments to justify the formula in this modified setting are identical to those just given.

Remark 3.5. We discuss briefly how to extend the results of the present paper and cover the more general case when the measure P^* is arbitrary with support in $[-1,1]$ in place of $\{-1,1\}$, with the understanding that we stick with the formula (1.4) to compute the link probabilities. In this case, the description of the limit law of $(\sigma_i^\ell)_{i,\ell \geq 1}$ must naturally be modified from (**S1**)-(**S2**), since these explicitly refer to ± 1 random variables. Instead, we would first need to sample i.i.d. random probability measures $(\mu_i)_{i \geq 1}$, each with law $\Xi \in \mathrm{Pr}(\mathrm{Pr}[-1,1])$, and then conditionally on $(\mu_i)_{i \geq 1}$, to sample σ_i^ℓ according to the law μ_i independently over $i, \ell \geq 1$. One key observation though is

that Lemma 3.4 remains valid as stated, provided that we understand that ν is the law of the mean value of μ_1, that is,

$$\nu = \mathrm{Law}\left(\int x \, d\mu_1(x) \right).$$

Moreover, the cavity computations ultimately only require us to be able to evaluate moments of the form described by Lemma 3.4. So even in this more general setting, we do not really need to keep track of the more complex object $\Xi \in \mathrm{Pr}(\mathrm{Pr}[-1,1])$, and can proceed as in the ± 1 case by focusing on this simpler measure $\nu \in \mathcal{M}_p$.

4   Cavity computation

In this section we perform the cavity computation that gives rise to the functional (1.26) and leads to the critical point representation in Theorem 1.1. The idea behind the cavity computation is to determine the effect of perturbing the Hamiltonian (3.9) for the system of size N by a single additional spin. To emphasize this additional spin, we write $\rho := (\varepsilon, \sigma) \in \Sigma_{N+1}$ for $\sigma \in \Sigma_N$ and $\varepsilon \in \{-1, +1\}$. Similarly, we write $\rho^* := (\varepsilon^*, \sigma^*) \in \Sigma_{N+1}$ for $\sigma^* \in \Sigma_N$ and $\varepsilon^* \in \Sigma_1$ to denote the ground truth sampled from P_{N+1}^*. Given a pair $(t, \mu) \in \mathbb{R}_{\geq 0} \times \mathcal{M}_+$ and a perturbation parameter $\lambda = (\lambda_k)_{k \geq 0}$, we will compare the systems with Hamiltonians $H_{N+1}^{t,\mu}(\varepsilon, \sigma, \lambda)$ and $H_N^{t,\mu}(\sigma, \lambda)$. This comparison will be done at the level of the free energies, and the re-scaled difference between $\overline{F}_{N+1}(t, \mu, \lambda)$ and $\overline{F}_N(t, \mu, \lambda)$ will be expressed as the difference between the modified Gibbs averages of two quantities that we may call the "cavity fields". We now describe the modified Hamiltonian associated with this modified Gibbs average as well as the two cavity fields.

For each $t \geq 0$, we introduce a random variable $\Pi_t \sim \mathrm{Poi}\, t \binom{N}{2}$ as well as an independent family of i.i.d. random matrices $(\widetilde{G}^k)_{k \geq 1}$ each having conditionally independent entries $(\widetilde{G}_{i,j}^k)_{i,j \leq N}$ taking values in $\{0, 1\}$ with conditional distribution

$$\mathbb{P}\{\widetilde{G}_{i,j}^k = 1 \mid \sigma^*\} := \frac{c + \Delta \sigma_i^* \sigma_j^*}{N+1}. \tag{4.1}$$

Given a collection of random indices $(i_k, j_k)_{k \geq 1}$ sampled uniformly at random from $\{1, \ldots, N\}^2$, independently of the other random variables, we define the modified time-dependent Hamiltonian $\mathcal{H}_N^t$ on Σ_N by

$$\mathcal{H}_N^t(\sigma) := \sum_{k \leq \Pi_t} \log \left[\left(c + \Delta \sigma_{i_k} \sigma_{j_k} \right)^{\widetilde{G}_{i_k, j_k}^k} \left(1 - \frac{c + \Delta \sigma_{i_k} \sigma_{j_k}}{N+1} \right)^{1 - \widetilde{G}_{i_k, j_k}^k} \right]. \tag{4.2}$$

Similarly, for each $\mu \in \mathcal{M}_+$, consider a sequence $(\Lambda_i((N+1)\mu))_{i \leq N}$ of independent Poisson point processes with mean measure $(N+1)\mu$, and define the modified measure-dependent Hamiltonian $\mathcal{H}_N^\mu$ on Σ_N by

$$\mathcal{H}_N^\mu(\sigma) := \sum_{i \leq N} \sum_{x \in \Lambda_i((N+1)\mu)} \log \left[\left(c + \Delta \sigma_i x \right)^{\widetilde{G}_i^x} \left(1 - \frac{c + \Delta \sigma_i x}{N+1} \right)^{1 - \widetilde{G}_i^x} \right], \tag{4.3}$$

25

where $(\widetilde{G}_i^x)_{i \le N}$ are conditionally independent random variables taking values in $\{0,1\}$ with conditional distribution

$$\mathbb{P}\big\{\widetilde{G}_i^x = 1 \mid \sigma^*, x\big\} := \frac{c + \Delta\sigma_i^* x}{N+1}. \tag{4.4}$$

For each pair $(t,\mu) \in \mathbb{R}_{\ge 0} \times \mathscr{M}_+$, the modified enriched Hamiltonian is the sum of the modified time-dependent and measure-dependent Hamiltonians,

$$\mathscr{H}_N^{t,\mu}(\sigma) := \mathscr{H}_N^t(\sigma) + \mathscr{H}_N^\mu(\sigma), \tag{4.5}$$

and, given a perturbation parameter $\lambda := (\lambda_k)_{k \ge 0}$, its perturbed version is

$$\mathscr{H}_N^{t,\mu}(\sigma, \lambda) := \mathscr{H}_N^{t,\mu}(\sigma) + H_N^{\text{gauss}}(\sigma, \lambda_0) + H_N^{\text{exp}}(\sigma, \lambda) \tag{4.6}$$

for the perturbation Hamiltonians (3.3) and (3.7). As usual, we write

$$\overline{F}_N'(t,\mu) := \frac{1}{N} \mathbb{E} \log \int_{\Sigma_N} \exp \mathscr{H}_N^{t,\mu}(\sigma) \, dP_N^*(\sigma) \tag{4.7}$$

for the modified free energy, and for each perturbation parameter $\lambda = (\lambda_k)_{k \ge 0}$, we denote by

$$\overline{F}_N'(t,\mu,\lambda) := \frac{1}{N} \mathbb{E} \log \int_{\Sigma_N} \exp \mathscr{H}_N^{t,\mu}(\sigma, \lambda) \, dP_N^*(\sigma) \tag{4.8}$$

its perturbed version. As in (3.11), we write $\langle \cdot \rangle'$ for the Gibbs average with respect to the perturbed and modified Hamiltonian (4.6). This means that for any bounded and measurable function $f = f(\sigma^1, \ldots, \sigma^n)$ of finitely many replicas,

$$\langle f(\sigma^1, \ldots, \sigma^n) \rangle' := \langle f \rangle' := \frac{\int_{\Sigma_N^n} f(\sigma^1, \ldots, \sigma^n) \prod_{\ell \le n} \exp \mathscr{H}_N^{t,\mu}(\sigma^\ell, \lambda) \, dP_N^*(\sigma^\ell)}{\big(\int_{\Sigma_N} \exp \mathscr{H}_N^{t,\mu}(\sigma, \lambda) \, dP_N^*(\sigma) \big)^n}. \tag{4.9}$$

This is the modified Gibbs measure relative to which the cavity fields will be averaged. It is readily verified that this Gibbs measure is a conditional expectation, and therefore satisfies the Nishimori identity (2.14).

The first cavity field will have a time-dependent and a measure-dependent component. For each $t \ge 0$, we introduce a random variable $\Pi_t' \sim \text{Poi}(tN)$ as well as an independent family of i.i.d. random vectors $(G'^k)_{k \ge 1}$ each having conditionally independent entries $(G'^k_i)_{i \le N}$ taking values in $\{0,1\}$ with conditional distribution

$$\mathbb{P}\big\{G'^k_i = 1 \mid \varepsilon^*, \sigma^*\big\} := \frac{c + \Delta\varepsilon^* \sigma_i^*}{N+1}. \tag{4.10}$$

Given a collection of random indices $(\ell_k)_{k \ge 1}$ sampled uniformly at random from $\{1, \ldots, N\}$, independently of the other random variables, we define the time-dependent cavity field $z_N^t(\varepsilon, \sigma)$ on Σ_{N+1} by

$$z_N^t(\varepsilon, \sigma) := \sum_{k \le \Pi_t'} \log \left[(c + \Delta\varepsilon\sigma_{\ell_k})^{G'^k_{\ell_k}} \left(1 - \frac{c + \Delta\varepsilon\sigma_{\ell_k}}{N+1} \right)^{1 - G'^k_{\ell_k}} \right]. \tag{4.11}$$

Similarly, for each $\mu \in \mathcal{M}_+$, we define the measure-dependent cavity field $z_N^\mu(\varepsilon)$ on Σ_1 by

$$z_N^\mu(\varepsilon) := \sum_{x \in \Lambda_i(N\mu)} \log\left[(c + \Delta\varepsilon x)^{G'^x} \left(1 - \frac{c + \Delta\varepsilon x}{N+1}\right)^{1 - G'^x} \right], \tag{4.12}$$

where G'^x are conditionally independent random variables taking values in $\{0, 1\}$ with conditional distribution

$$\mathbb{P}\{G'^x = 1 \mid \varepsilon^*, x\} := \frac{c + \Delta\varepsilon^* x}{N+1}. \tag{4.13}$$

For each pair $(t, \mu) \in \mathbb{R}_{\geq 0} \times \mathcal{M}_+$, the cavity field $z_N^{t,\mu}(\varepsilon, \sigma)$ is the sum of the time-dependent and measure-dependent cavity fields,

$$z_N^{t,\mu}(\varepsilon, \sigma) := z_N^t(\varepsilon, \sigma) + z_N^\mu(\varepsilon). \tag{4.14}$$

The second cavity field will only have a time-dependent component. For each $t \geq 0$, introduce an independent family of i.i.d. random matrices $(G''^k)_{k \geq 1}$ each having conditionally independent entries $(G''^k_{i,j})_{i,j \leq N}$ taking values in $\{0, 1\}$ with conditional distribution

$$\mathbb{P}\{G''^k_{i,j} = 1 \mid \sigma^*\} := \frac{c + \Delta\sigma_i^* \sigma_j^*}{N(N+1)}. \tag{4.15}$$

Given a collection of random indices $(\ell_k, m_k)_{k \geq 1}$ sampled uniformly at random from $\{1, \ldots, N\}^2$, independently of the other random variables, we define the cavity field $y_N(\sigma)$ on Σ_N by

$$y_N^t(\sigma) := \sum_{k \leq \Pi_t} \log\left[(c + \Delta\sigma_{\ell_k}\sigma_{m_k})^{G''^k_{\ell_k, m_k}} \left(1 - \frac{c + \Delta\sigma_{\ell_k}\sigma_{m_k}}{N(N+1)}\right)^{1 - G''^k_{\ell_k, m_k}} \right]. \tag{4.16}$$

The cavity computation reveals that the re-scaled difference between $\overline{F}_{N+1}(t, \mu, \lambda)$ and $\overline{F}_N(t, \mu, \lambda)$ is the difference between the cavity fields (4.14) and (4.16) averaged with respect to the Gibbs measure (4.9). More precisely, for each pair $(t, \mu) \in \mathbb{R}_{\geq 0} \times \mathcal{M}_+$, the asymptotic difference between these re-scaled free energies is given by

$$A_N(t, \mu, \lambda) := \mathbb{E} \log \int_{\Sigma_1} \langle \exp z_N^{t,\mu}(\varepsilon, \sigma) \rangle' \, dP^*(\varepsilon) - \mathbb{E} \log \langle \exp y_N^t(\sigma) \rangle'. \tag{4.17}$$

To prove this, it will be convenient to introduce notation for approximate distributional identities. Given two Hamiltonians H_N^1 and H_N^2 on Σ_N, we write

$$H_N^1(\sigma) \overset{d}{\approx} H_N^2(\sigma) \iff N\overline{F}_N^1 = N\overline{F}_N^2 + o_N(1), \tag{4.18}$$

where $\overline{F}_N^1$ and $\overline{F}_N^2$ denote the free energies associated with the Hamiltonians H_N^1 and H_N^2, respectively. Oftentimes we will not prove rigorously that two Hamiltonians are approximately equal in distribution, contenting ourselves with showing that a given realization of them differs by a quantity that is stochastically of negative degree in N. However, all approximate distributional identities that we introduce may be justified rigorously through an interpolation argument similar to that in Lemmas 3.1 or 4.4.

Lemma 4.1. *For all $(t, \mu) \in \mathbb{R}_{\geq 0} \times \mathcal{M}_+$, perturbation parameter λ, and $N \geq 1$, the re-scaled difference between the perturbed free energy* (3.10) *for a system of size $N+1$ and a system of size N is given by*

$$(N+1)\overline{F}_{N+1}(t, \mu, \lambda) - N\overline{F}_N(t, \mu, \lambda) = A_N(t, \mu, \lambda) + o_N(1), \qquad (4.19)$$

where the error term is uniform over the perturbation parameter $\lambda = (\lambda_k)_{k \geq 1}$ with $\lambda_k \in [2^{-k-1}, 2^{-k}]$ for $k \geq 0$.

Proof. For simplicity, we will ignore the perturbation Hamiltonians (3.3) and (3.7); however, the choices (3.5)-(3.6) of the sequences $(\varepsilon_N)_{N \geq 1}$ and $(s_N)_{N \geq 1}$ can be used to show that the error incurred by replacing the perturbation Hamiltonian $H_{N+1}^{\text{gauss}}(\sigma, \lambda_0) + H_{N+1}^{\text{exp}}(\sigma, \lambda)$ for the system of size $N+1$ by that for the system of size N can be absorbed into the error term on the right side of (4.19). The proof proceeds in two steps. First we relate the time-dependent Hamiltonian (1.15) for the system of size $N+1$ to that for the system of size N, and then we do the same for the measure-dependent Hamiltonian (1.18).

Step 1: time-dependent cavity computation. The probability that the tuple $\{i_k, j_k\}$ in the time-dependent Hamiltonian (1.15) contains the index $N+1$ is

$$1 - \left(\frac{N}{N+1}\right)^2 = \frac{2N+1}{(N+1)^2} = \frac{2}{N+1} - \frac{1}{(N+1)^2}.$$

It follows by the Poisson coloring theorem that

$$H_{N+1}^t(\rho) \stackrel{d}{\approx} \mathcal{H}_N^t(\sigma) + z_N^t(\varepsilon, \sigma) \qquad (4.20)$$

for the modified time-dependent Hamiltonian (4.2) and the time-dependent cavity field (4.11). We have implicitly used that

$$t\binom{N+1}{2}\left(1 - \frac{2}{N+1}\right) = t\binom{N}{2} \quad \text{and} \quad t\binom{N+1}{2} \cdot \frac{2}{N+1} = tN.$$

Another application of the Poisson coloring theorem reveals that

$$H_N^t(\sigma) \stackrel{d}{\approx} \mathcal{H}_N^t(\sigma) + y_N^t(\sigma), \qquad (4.21)$$

for the modified time-dependent Hamiltonian (4.2) and the time-dependent cavity field (4.16). We have implicitly used that

$$\frac{1}{N} - \frac{1}{N+1} = \frac{1}{N(N+1)}.$$

Step 2: measure-dependent cavity computation. The measure-dependent Hamiltonian (1.18) admits the approximate distributional decomposition

$$H_{N+1}^\mu(\rho) \stackrel{d}{\approx} \mathcal{H}_N^\mu(\sigma) + z_N^\mu(\varepsilon, \sigma) \qquad (4.22)$$

28

for the modified measure-dependent Hamiltonian (4.3) and the measure-dependent cavity field (4.12). Similarly, the Poisson coloring theorem implies that

$$H_N^\mu(\sigma) \overset{d}{\approx} \mathcal{H}_N^\mu(\sigma). \tag{4.23}$$

Combining the cavity computations (4.20)-(4.23) with the definitions of the modified enriched Hamiltonian (4.5), the Gibbs average (4.9), and the cavity fields (4.14) and (4.16) completes the proof. ∎

The cavity representation in Lemma 4.1 suggests that to understand the limit of the free energy, it is helpful to understand the asymptotic behavior of the quantity A_N in (4.17). Indeed, if both limits existed, they would coincide. Unfortunately, it will not be possible to show that the limit of A_N exists in general; however, we will now argue that any subsequential limit of A_N along which the multioverlaps (3.2) concentrate in the sense of (3.15) is given by the functional (1.26) evaluated at a measure $\nu \in \mathcal{M}_p$ which generates the asymptotic spin array according to (S1)-(S2).

Lemma 4.2. *Fix $(t,\mu) \in \mathbb{R}_{\geq 0} \times \mathcal{M}_+$, and suppose there is a sequence $(N_k, \lambda^{N_k})_{k \geq 1}$ such that $(N_k)_{k \geq 1}$ increases to infinity and the multioverlap concentration (3.15) for the Hamiltonian (4.6) with perturbation parameters $(\lambda^{N_k})_{k \geq 1}$ holds for every $n \geq 1$. Then there are a subsequence $(N'_k, \lambda^{N'_k})_{k \geq 1}$ and a measure $\nu \in \mathcal{M}_p$ such that the asymptotic spin array associated with the perturbation parameters $(\lambda^{N'_k})_{k \geq 1}$ is generated by ν according to (S1)-(S2) and*

$$\lim_{k \to +\infty} A_{N'_k}\big(t, \mu, \lambda^{N'_k}\big) = \mathscr{P}_{t,\mu}(\nu). \tag{4.24}$$

Proof. Using the Prokhorov theorem, we can find a subsequence $(N'_k)_{k \geq 1}$ along which the distribution of the spin array $(\sigma_i^\ell)_{i,\ell \geq 1}$ under the Gibbs measure $\mathbb{E}\langle \cdot \rangle'$ converges in the sense of finite-dimensional distributions. For simplicity of notation, we will denote this subsequence N'_k simply by N. By Proposition 3.3, there exists a probability measure $\nu \in \mathcal{M}_p$ such that the asymptotic spin array is generated by ν according to (S1)-(S2). The rest of the proof is devoted to establishing (4.24) and proceeds in two steps. Each step determines the asymptotic behavior of one of the quantities

$$A_N^1(t, \mu, \lambda^N) := \mathbb{E}\log\int_{\Sigma_1} \langle \exp z_N^{t,\mu}(\varepsilon, \sigma)\rangle' \, dP^*(\varepsilon) \quad \text{and} \quad A_N^2(t, \mu, \lambda^N) := \mathbb{E}\log\langle \exp y_N^t(\sigma)\rangle'$$

which make up $A_N(t, \mu, \lambda^N)$. The main difficulty will be that there are two limits to account for: the convergence of the spin array $(\sigma_i^\ell)_{i,\ell \geq 1}$ and the convergence of the Poisson sums. Nonetheless, it will be possible to show that

$$\lim_{N \to +\infty} A_N^1(t, \mu, \lambda^N) = \psi(\mu + t\nu) \quad \text{and} \quad \lim_{N \to +\infty} A_N^2(t, \mu, \lambda^N) = \frac{t}{2}\int_{-1}^1 G_\nu(y)\, d\nu(y) \tag{4.25}$$

for the initial condition $\psi : \mathcal{M}_+ \to \mathbb{R}$ in (1.27) and the function $G_\nu : [-1,1] \to \mathbb{R}$ in (1.25). To alleviate notation, the dependence of A_N^1, A_N^2, z_N and y_N on t, μ and λ^N will be kept implicit, and we will write $s := \mu[-1,1]$.

Step 1: limit of A_N^1. The modified Hamiltonian (4.6) corresponds to a problem in statistical inference, so the Nishimori identity (2.14) implies that

$$A_N^1 = \mathbb{E}\log \int_{\Sigma_1} \langle \exp z_N(\varepsilon,\sigma^1)\rangle' \, dP^*(\varepsilon)$$

for the cavity field

$$z_N(\varepsilon,\sigma^1) := \sum_{k \leq \Pi_t'} \log\left[\left(c + \Delta\varepsilon\sigma_{\ell_k}^1\right)^{G_{\ell_k}'^k}\left(1 - \frac{c + \Delta\varepsilon\sigma_{\ell_k}^1}{N+1}\right)^{1-G_{\ell_k}'^k}\right] + z_N^\mu(\varepsilon),$$

where, through a slight abuse of notation, the family of i.i.d. random matrices $(G'^k)_{k \geq 1}$ each has conditionally independent entries $(G_i'^k)_{i \leq N}$ taking values in $\{0,1\}$ with conditional distribution

$$\mathbb{P}\{G_i'^k = 1 \mid \varepsilon^*, \sigma^2\} := \frac{c + \Delta\varepsilon^*\sigma_i^2}{N+1}.$$

The Poisson coloring theorem implies that the number of terms in the first sum defining $z_N(\varepsilon,\sigma^1)$ that have $G_{\ell_k}'^k$ equal to zero is Poisson with mean

$$tN \cdot \left(1 - \frac{c + \Delta\varepsilon^*\sigma_{\ell_k}^2}{N+1}\right) = tN + \mathcal{O}(1).$$

Together with a Taylor expansion of the logarithm and the law of large numbers, this implies that the contribution of the terms in the first term defining $z_N(\varepsilon,\sigma^1)$ that have $G_{\ell_k}'^k$ equal to zero is

$$-\sum_{k \leq tN} \frac{c + \Delta\varepsilon\sigma_{\ell_k}^1}{N} + \mathcal{O}\left(N^{-1}\right) = -ct - \Delta t\varepsilon\overline{m} + \mathcal{O}\left(N^{-1}\right).$$

A similar argument shows that the contribution of the terms in $z_N^\mu(\varepsilon)$ that have G'^x equal to zero is $-cs - \Delta s\varepsilon\,\mathbb{E}x_1 + \mathcal{O}(N^{-1})$. If we define

$$z_N'(\varepsilon,\sigma^1) := \sum_{k \leq \Pi_t'} \log\left(c + \Delta\varepsilon\sigma_{\ell_k}^1\right)^{G_{\ell_k}'^k} + \sum_{k \leq \Pi_s'} \log\left(c + \Delta\varepsilon x_k\right)^{G_k'^x},$$

where $(x_k)_{k \geq 1}$ are i.i.d. random variables with law μ and, through a slight abuse of notation, the random variables $(G_k'^x)_{k \geq 1}$ are conditionally independent with conditional distribution

$$\mathbb{P}\{G_k'^x = 1 \mid \varepsilon^*, x_k\} := \frac{c + \Delta\varepsilon^* x_k}{N+1},$$

then this implies that

$$A_N^1 = \mathbb{E}\log \int_{\Sigma_1} \exp(-\Delta\varepsilon(t\overline{m} + s\,\mathbb{E}x_1))\langle \exp z_N'(\varepsilon,\sigma^1)\rangle' \, dP^*(\varepsilon) - c(t+s) + \mathcal{O}\left(N^{-1}\right). \quad (4.26)$$

We now fix $\delta > 0$ and focus on the first term in this expression which we denote by $A_N^{1,1}$. Averaging with respect to the randomness of $(G'^k_{\ell_k})_{k\geq 1}$ and $(G'^x_k)_{k\geq 1}$, and using symmetry between sites reveals that $A_N^{1,1}$ is given by averaging the sum over $r \leq \Pi'_t$ and $q \leq \Pi'_s$ of

$$\binom{\Pi'_t}{r}\binom{\Pi'_s}{q}\prod_{i=1}^{r}\frac{c+\Delta\varepsilon^*\sigma_i^2}{N+1}\prod_{i=r+1}^{\Pi'_t}\left(1-\frac{c+\Delta\varepsilon^*\sigma_i^2}{N+1}\right)\prod_{j=1}^{q}\frac{c+\Delta\varepsilon^*x_j}{N+1}\prod_{j=q+1}^{\Pi'_s}\left(1-\frac{c+\Delta\varepsilon^*x_j}{N+1}\right)$$

$$\log\left\langle\int_{\Sigma_1}\exp(-\Delta\varepsilon(t\overline{m}+s\,\mathbb{E}x_1))\prod_{i=1}^{r}(c+\Delta\varepsilon\sigma_i^1)\prod_{j=1}^{q}(c+\Delta\varepsilon x_j)\,\mathrm{d}P^*(\varepsilon)\right\rangle'$$

with respect to the randomness of σ^2, $(x_k)_{k\geq 1}$, Π'_t, Π'_s, and ε^*. Since the expression in the second line of this display grows at most linearly in r and q, the concentration of Π'_t and Π'_s about tN and sN, respectively, and the rapid decay of the binomial coefficients allow us to truncate both sums at a large enough integer M by incurring an error of at most δ uniformly over N. Together with the approximations

$$\prod_{i=r+1}^{\Pi'_t}\left(1-\frac{c+\Delta\varepsilon^*\sigma_i^1}{N+1}\right)=\exp(-tc-\Delta t\varepsilon^*\overline{m})+o_N(1),$$

$$\prod_{j=q+1}^{\Pi'_s}\left(1-\frac{c+\Delta\varepsilon^*x_j}{N+1}\right)=\exp(-sc-\Delta s\varepsilon^*\,\mathbb{E}x_1)+o_N(1),$$

this implies that, up to an error vanishing in N and uniform in δ, the quantity $A_N^{1,1}$ is given by

$$\mathbb{E}\sum_{r\leq\Pi'_t\wedge M}\sum_{q\leq\Pi'_s\wedge M}\binom{\Pi'_t}{r}\binom{\Pi'_s}{q}\frac{e^{-c(t+s)-\Delta\varepsilon^*(t\overline{m}+s\,\mathbb{E}x_1)}}{(N+1)^{r+q}}\left\langle\prod_{i=1}^{r}(c+\Delta\varepsilon^*\sigma_i^2)\prod_{j=1}^{q}(c+\Delta\varepsilon^*x_j)\right.$$

$$\left.\log\left\langle\int_{\Sigma_1}e^{-\Delta\varepsilon(t\overline{m}+s\,\mathbb{E}x_1)}\prod_{i=1}^{r}(c+\Delta\varepsilon\sigma_i^1)\prod_{j=1}^{q}(c+\Delta\varepsilon x_j)\,\mathrm{d}P^*(\varepsilon)\right\rangle'\right\rangle'.$$

Up to incurring a further error vanishing in N and uniform in δ, the Weierstrass approximation theorem can be used to replace the logarithm by a polynomial. This results in a linear combination of finitely many terms that are functions of finitely many spins at a time. Invoking Lemma 3.4 to replace each of the averaged spins in this expression by a sample from the measure ν, tracing back the manipulations we have done to $A_N^{1,1}$, and letting δ tend to zero reveals that

$$A_N^{1,1}=\mathbb{E}\log\int_{\Sigma_1}\exp(-\Delta\varepsilon(t\overline{m}+s\,\mathbb{E}x_1))\exp z_N'(\varepsilon)\,\mathrm{d}P^*(\varepsilon)+o_N(1),$$

for the cavity field

$$z_N'(\varepsilon):=\sum_{k\leq\Pi'_t}\log(c+\Delta\varepsilon y_k)^{G'^y_k}+\sum_{k\leq\Pi'_s}\log(c+\Delta\varepsilon x_k)^{G'^x_k},\tag{4.27}$$

where $(y_k)_{k\geq 1}$ are independent samples from the measure $\nu\in\mathcal{M}_p$ which generates the asymptotic spin array according to (**S1**)-(**S2**). In other words, it is possible to take the limit of the spin array while leaving the finite-volume Poisson sums unaffected. To simplify this expression further,

31

suppose that the probability measures $\overline{\mu}$ and ν are discrete,

$$\overline{\mu} := \sum_{\ell \leq K} p_\ell \delta_{a_\ell} \quad \text{and} \quad \nu := \sum_{\ell \leq K} q_\ell \delta_{a_\ell}$$

for some integer $K \geq 1$, some weights $p_\ell, q_\ell \in [0,1]$, and some masses $a_\ell \in [-1,1]$. For each $\ell \leq K$, introduce the index sets

$$\mathscr{I}_t^\ell := \{k \leq \Pi_t' \mid G_k'^y = 1 \text{ and } y_k = a_\ell\} \quad \text{and} \quad \mathscr{I}_s^\ell := \{k \leq \Pi_s' \mid G_k'^x = 1 \text{ and } x_k = a_\ell\}$$

in such a way that

$$A_N^{1;1} = \mathbb{E}\log \int_{\Sigma_1} \exp(-\Delta\varepsilon(t\overline{m} + s\,\mathbb{E}x_1)) \prod_{k \leq \Pi_t'} (c + \Delta\varepsilon y_k)^{G_k'^y} \prod_{k \leq \Pi_s'} (c + \Delta\varepsilon x_k)^{G_k'^x} \mathrm{d}P^*(\varepsilon) + o_N(1)$$

$$= \mathbb{E}\log \int_{\Sigma_1} \exp(-\Delta\varepsilon(t\overline{m} + s\,\mathbb{E}x_1)) \prod_{\ell \leq K} \prod_{k \in I_t^\ell \cup I_s^\ell} (c + \Delta\varepsilon a_\ell)\,\mathrm{d}P^*(\varepsilon) + o_N(1).$$

The Poisson coloring theorem implies that $|I_t^\ell \cup I_s^\ell|$ is Poisson with mean

$$\mathbb{E}\Pi_t'\mathbb{P}\{G_1'^y = 1, y_1 = a_\ell\} + \mathbb{E}\Pi_s'\mathbb{P}\{G_1'^x = 1, x_1 = a_\ell\} = sp_\ell(c + \Delta\varepsilon^* a_\ell) + tq_\ell(c + \Delta\varepsilon^* a_\ell) + o_N(1)$$
$$= (c + \Delta\varepsilon^* a_\ell)(\mu + t\nu)(a_\ell) + o_N(1).$$

It follows that

$$A_N^{1;1} = \mathbb{E}\log \int_{\Sigma_1} \exp(-\Delta\varepsilon(t\overline{m} + s\,\mathbb{E}x_1)) \prod_{x \in \Pi_*(\mu+t\nu)} (c + \Delta\varepsilon x)\,\mathrm{d}P^*(\varepsilon) + o_N(1)$$

$$= \psi(\mu + t\nu) + c(t+s) + o_N(1),$$

where $\Pi_*(\mu + t\nu)$ is a Poisson point process with mean measure $(c + \Delta\varepsilon^* x)\,\mathrm{d}(\mu + t\nu)(x)$. Extending this result to the case when μ and ν are arbitrary can be done through a continuity argument leveraging Proposition 2.1. Combining this with (4.26) establishes the first equality in (4.25).

Step 2: limit of A_N^2. The modified Hamiltonian (4.6) corresponds to a problem in statistical inference, so the Nishimori identity (2.14) implies that

$$A_N^2 = \mathbb{E}\log\langle \exp y_N(\sigma^1)\rangle'$$

for the cavity field

$$y_N(\sigma^1) := \sum_{k \leq \Pi_t} \log\left[(c + \Delta\sigma_{\ell_k}^1 \sigma_{m_k}^1)^{G''_{\ell_k,m_k}^k} \left(1 - \frac{c + \Delta\sigma_{\ell_k}^1 \sigma_{m_k}^1}{N(N+1)}\right)^{1 - G''_{\ell_k,m_k}^k} \right],$$

where, through a slight abuse of notation, the family of i.i.d. random matrices $(G''^k)_{k \geq 1}$ each has conditionally independent entries $(G''^k_{i,j})_{i,j \leq N}$ taking values in $\{0,1\}$ with conditional distribution

$$\mathbb{P}\{G''^k_{i,j} = 1 \mid \sigma^2\} := \frac{c + \sigma_{\ell_k}^2 \sigma_{m_k}^2}{N(N+1)}.$$

The Poisson coloring theorem implies that the number of terms in the sum defining $y_N(\sigma^1)$ that have $G''^k_{\ell_k,m_k}$ equal to zero is Poisson with mean

$$t\binom{N}{2} \cdot \left(1 - \frac{c + \Delta\sigma^2_{\ell_k}\sigma^2_{m_k}}{N(N+1)}\right) = \frac{t}{2}N^2 + \mathscr{O}(1).$$

Together with a Taylor expansion of the logarithm and the law of large numbers, this implies that the contribution of the terms in the sum defining $y_N(\sigma^1)$ that have $G'^k_{\ell_k}$ equal to zero is

$$-\sum_{k \leq tN^2/2} \frac{c + \Delta\sigma^1_{\ell_k}\sigma^1_{m_k}}{N(N+1)} + \mathscr{O}\left(N^{-1}\right) = -\frac{t}{2}\left(c + \Delta\overline{m}^2\right) + \mathscr{O}\left(N^{-1}\right).$$

If we define

$$y'_N(\sigma^1) := \sum_{k \leq \Pi_t} \log(c + \Delta\sigma^1_{\ell_k}\sigma^1_{m_k})^{G''^k_{\ell_k,m_k}},$$

then this implies that

$$A^2_N = \mathbb{E}\log\langle\exp y'_N(\sigma^1)\rangle' - \frac{t}{2}\left(c - \Delta\overline{m}^2\right) + \mathscr{O}\left(N^{-1}\right). \tag{4.28}$$

We now focus on the first term in this expression which we denote by $A^{2;1}_N$. A similar argument as that in Step 1 shows that it is possible to take the limit of the spin array while leaving the finite-volume Poisson sum unaffected. More precisely, it is possible to show that

$$A^{2;1}_N = \mathbb{E}\sum_{k \leq \Pi_t} \log(c + \Delta y_{1,k}y_{2,k})^{G''^y_k} + o_N(1),$$

where $(y_{1,k})_{k \geq 1}$ and $(y_{2,k})_{k \geq 1}$ are independent samples from the measure $v \in \mathscr{M}_p$ which generates the asymptotic spin array according to (**S1**)-(**S2**), and $(G''^y_k)_{k \geq 1}$ has conditional distribution

$$\mathbb{P}\big\{G''^y_k = 1 \mid y_{1,k}, y_{2,k}\big\} := \frac{c + \Delta y_{1,k}y_{2,k}}{N(N+1)}.$$

Averaging with respect to the randomness of $(G''^y_k)_{k \geq 1}$, and using the independence of $(y_{1,k})_{k \geq 1}$ and $(y_{2,k})_{k \geq 1}$ reveals that

$$A^{2;1}_N = \frac{1}{N(N+1)}\mathbb{E}\sum_{k \leq \Pi_t}(c + \Delta y_{1,k}y_{2,k})\log(c + \Delta y_{1,k}y_{2,k}) + o_N(1)$$

$$= \frac{t}{2}\mathbb{E}(c + \Delta y_1 y_2)\log(c + \Delta y_1 y_2) + o_N(1).$$

Combining this with (4.28) and recalling the definition of the function $G_v : [-1,1] \to \mathbb{R}$ in (1.25) gives

$$\lim_{N \to +\infty} A^2_N = \frac{t}{2}\left(\mathbb{E}(c + \Delta y_1 y_2)\log(c + \Delta y_1 y_2) - \Delta\overline{m}^2 - c\right) = \frac{t}{2}\int_{-1}^{1} G_v(y)\,\mathrm{d}v(y).$$

This completes the proof. ∎

33

This result can now be combined with the cavity representation in Lemma 4.1 to bound the asymptotic free energy from above and from below by the functional (1.26) evaluated at two possibly different measures $\nu^{\pm} \in \mathcal{M}_p$ which generate the asymptotic spin array associated with two sets of perturbation parameters according to (S1)-(S2). This result will be improved in the next section.

Proposition 4.3. *For all $(t, \mu) \in \mathbb{R}_{\geq 0} \times \mathcal{M}_+$ there are sequences $\left(N_k^{\pm}, \lambda^{N_k^{\pm}}\right)_{k \geq 1}$ and measures $\nu^{\pm} \in \mathcal{M}_p$ such that the following holds. The sequences $(N_k^{\pm})_{k \geq 1}$ increase to infinity, the multi-overlap concentration (3.15) for the Hamiltonian (4.6) with perturbation parameters $(\lambda^{N_k^{\pm}})_{k \geq 1}$ holds for $n \geq 1$, the asymptotic spin array associated with the perturbation parameters $(\lambda^{N_k^{\pm}})_{k \geq 1}$ is generated by $\nu^{\pm}$ according to (S1)-(S2), and*

$$\mathscr{P}_{t,\mu}(\nu^-) \leq \liminf_{N \to +\infty} \overline{F}_N(t,\mu) \leq \limsup_{N \to +\infty} \overline{F}_N(t,\mu) \leq \mathscr{P}_{t,\mu}(\nu^+). \tag{4.29}$$

Proof. Combining Lemma 3.2 with the arguments in [43, Lemma 3.3], it is possible to find a sequence of perturbation parameters $(\lambda^N)_{N \geq 1}$ such that the multioverlap concentration (3.15) for the perturbed Hamiltonian (4.6) holds for all $n \geq 1$. Remembering that the perturbation parameters do not affect the asymptotic behavior of the free energy by Lemma 3.1 and that the $\liminf$ of a Cesàro sum is bounded from below by the $\liminf$ of its general term, we see that

$$\liminf_{N \to +\infty} \overline{F}_N(t,\mu) \geq \liminf_{N \to +\infty} A_N(t,\mu,\lambda^N).$$

Invoking Lemma 4.2 and passing to a further subsequence along which $(A_N(t,\mu,\lambda^N))_{N \geq 1}$ converges to its $\liminf$ if necessary, gives a sequence $(N_k^-, \lambda^{N_k^-})_{k \geq 1}$ and a measure $\nu^- \in \mathcal{M}_p$ such that the asymptotic spin array associated with the perturbation parameters $(\lambda^{N_k^-})_{k \geq 1}$ is generated by ν^- according to (S1)-(S2) and

$$\liminf_{N \to +\infty} A_N\left(t,\mu,\lambda^N\right) = \lim_{k \to +\infty} A_{N_k^-}\left(t,\mu,\lambda^{N_k^-}\right) = \mathscr{P}_{t,\mu}\left(\nu^-\right).$$

Together with the previous display this gives the lower bound in (4.29). The upper bound in this inequality is proved in an identical manner by instead using that the $\limsup$ of a Cesàro sum is bounded from above by the $\limsup$ of its general term. This completes the proof. ∎

We close this section by showing that the asymptotic behavior of the enriched free energy (1.22) and the modified enriched free energy (4.7) are equivalent. This will play its part in the next section when we show that whenever the free energy converges, we can turn (4.29) into an equality.

Lemma 4.4. *For all $(t, \mu) \in \mathbb{R}_{\geq 0} \times \mathcal{M}_+$, the enriched free energy (1.22) and the modified enriched free energy (4.7) are asymptotically equivalent,*

$$\lim_{N \to +\infty} \left|\overline{F}_N(t,\mu) - \overline{F}'_N(t,\mu)\right| = 0. \tag{4.30}$$

Proof. Recall from (4.21) and (4.23) that the enriched Hamiltonian (1.21) and the modified enriched Hamiltonian (4.5) differ by the cavity field y_N^t in (4.16) in the approximate sense defined in (4.18),

$$H_N^{t,\mu}(\sigma) \stackrel{d}{\approx} \mathscr{H}_N^{t,\mu}(\sigma) + y_N^t(\sigma). \tag{4.31}$$

34

We now leverage an interpolation argument to establish the asymptotic equivalence (4.30). For each $u \in [0,1]$, we define the interpolating Hamiltonian

$$\mathscr{H}_N^{t,\mu}(u,\sigma) := \mathscr{H}_N^{t,\mu}(\sigma) + u y_N^t(\sigma),$$

and denote by $\overline{F}_N(u)$ and $\langle \cdot \rangle_u$ its associated free energy and Gibbs average, respectively. The fundamental theorem of calculus implies that

$$\left| \overline{F}_N(1) - \overline{F}_N(0) \right| \leq \sup_{u \in [0,1]} \left| \frac{\mathrm{d}}{\mathrm{d}u} \overline{F}_N(u) \right| \leq \frac{1}{N} \sup_{u \in [0,1]} \mathbb{E} \left\langle |y_N^t(\sigma)| \right\rangle_u. \tag{4.32}$$

Averaging with respect to the randomness of $(G''^k)_{k \geq 1}$ shows that for each $u \in [0,1]$,

$$\mathbb{E} \left\langle |y_N^t(\sigma)| \right\rangle_u \leq \mathbb{E} \Pi_t \left(\frac{(c + |\Delta|)(\log(c + |\Delta|) + 1)}{N(N+1)} + \mathscr{O}\left(N^{-4} \right) \right) = \mathscr{O}(1).$$

Substituting this into (4.32), recalling that $\overline{F}_N(1) = \overline{F}_N(t,\mu) + o_N(1)$ by the approximate distributional identity (4.31), and observing that $\overline{F}_N(0) = \overline{F}'_N(t,\mu)$ completes the proof. $\blacksquare$

5 Critical point representation

In this section we finally prove Theorems 1.1 and 1.2. We begin by showing that whenever the limit free energy exists and is Gateaux differentiable, it may be identified with a measure $\nu \in \mathscr{M}_p$ through the function G_ν in (1.25). We then show that the map $\eta \mapsto G_\eta$ is injective on $\mathscr{M}_p$, meaning that the identified measure $\nu \in \mathscr{M}_p$ is unique. We also prove that this measure generates some subsequential limit of the spin array associated with the modified and perturbed Hamiltonian (4.6). We then argue that under these same assumptions, the measure ν is a valid choice for both the measure ν^+ and the measure ν^- in Proposition 4.3. Through arguments similar to those in Lemma 4.2, at any point $(t,\mu) \in \mathbb{R}_{\geq 0} \times \mathscr{M}_+$ of differentiability of the limit free energy, we relate the Gateaux derivative of the limit free energy at (t,μ) to the Gateaux derivative of the asymptotic initial condition (1.27) at $\mu + t\nu$. This leads to the equality $G_\nu = D_\mu \psi(\mu + t\nu, \cdot)$ which we show to be equivalent to the fixed point equation $\nu = \Gamma_{t,\mu}(\nu)$ associated with the operator (1.6). Finally, we combine all these insights with Proposition 2.2 on the almost everywhere Gateaux differentiability of the limit free energy to establish Theorems 1.1 and 1.2.

To relate the Gateaux derivative of the limit free energy to that of the initial condition, it will be convenient to recall from [18, Remark 3.2] that for any $\mu \in \mathscr{M}_+$, the density of the Gateaux derivative $D_\mu \psi(\mu)$ is

$$D_\mu \psi(\mu, x) = \mathbb{E}\langle c + \Delta \varepsilon x \rangle_* \log \langle c + \Delta \varepsilon x \rangle_* - c - \Delta \overline{m} x, \tag{5.1}$$

where, conditionally on the ground truth σ^*, the bracket $\langle \cdot \rangle_*$ denotes the Gibbs average associated with the Hamiltonian defined on Σ_1 by

$$H^\mu(\varepsilon) := - \int_{-1}^1 (c + \Delta \varepsilon x) \, \mathrm{d}\mu(x) + \sum_{x \in \Pi_*(\mu)} \log(c + \Delta \varepsilon x). \tag{5.2}$$

Here $\Pi_*(\mu)$ denotes a Poisson point process on $[-1,1]$ with mean measure $(c+\Delta\sigma^*x)\,\mathrm{d}\mu(x)$. This means that for any bounded and measurable function $f = f(\varepsilon^1,\dots,\varepsilon^n)$ of finitely many replicas,

$$\langle f(\varepsilon^1,\dots,\varepsilon^n)\rangle_* := \langle f\rangle_* := \frac{\int_{\Sigma_1^n} f(\varepsilon^1,\dots,\varepsilon^n)\prod_{\ell\le n}\exp H^\mu(\varepsilon^\ell)\,\mathrm{d}P^*(\varepsilon^\ell)}{\left(\int_{\Sigma_1}\exp H^\mu(\varepsilon)\,\mathrm{d}P^*(\varepsilon)\right)^n}. \tag{5.3}$$

It will also be convenient to remember from the proof of [18, Lemma 2.4] that the Gateaux derivative of the enriched free energy (1.22) at $(t,\mu)\in\mathbb{R}_{\ge 0}\times\mathcal{M}_+$ in the direction of a probability measure $\eta\in\mathrm{Pr}[-1,1]$ is given by

$$D_\mu\overline{F}_N(t,\mu;\eta) = \sum_{i\le N}\mathbb{E}\log\left\langle (c+\Delta\sigma_i w_i)^{\widetilde{G}_i^w}\left(1-\frac{c+\Delta\sigma_i w_i}{N}\right)^{1-\widetilde{G}_i^w}\right\rangle, \tag{5.4}$$

where the bracket $\langle\cdot\rangle$ denotes the Gibbs average (2.13), the $(w_i)_{i\ge 1}$ are i.i.d. random variables sampled from η, and $(\widetilde{G}_i^w)_{i\le N}$ are conditionally independent random variables taking values in $\{0,1\}$ with conditional distribution

$$\mathbb{P}\{\widetilde{G}_i^w = 1\mid\sigma^*,w_i\} := \frac{c+\Delta\sigma_i^* w_i}{N}. \tag{5.5}$$

Lemma 5.1. *Suppose that the sequence of enriched free energies $(\overline{F}_N)_{N\ge 1}$ converges pointwise to some limit $f:\mathbb{R}_{\ge 0}\times\mathcal{M}_+\to\mathbb{R}$ along a subsequence $(N_k)_{k\ge 1}$, and fix $(t,\mu)\in\mathbb{R}_{\ge 0}\times\mathrm{Reg}(\mathcal{M}_+)$. If $f(t,\cdot)$ is Gateaux differentiable at $\mu\in\mathrm{Reg}(\mathcal{M}_+)$, then there exists $\nu\in\mathcal{M}_p$ with*

$$G_\nu = D_\mu f(t,\mu,\cdot). \tag{5.6}$$

Proof. For each integer $k\ge 1$, let $\nu_{N_k} := \mathscr{L}(\langle\sigma_1\rangle)$ denote the law (under $\mathbb{P}$) of the Gibbs average of a spin variable, where the Gibbs measure is as in (2.13). The Nishimori identity (2.14), the derivative expression (2.12), and the definition of the kernel g in (1.24) imply that

$$D_\mu\overline{F}_{N_k}(t,\mu,\cdot) = G_{\nu_{N_k}}(\cdot) + \mathcal{O}(N_k^{-1}). \tag{5.7}$$

By the Prokhorov theorem, the sequence $(\nu_{N_k})_{k\ge 1}\subset\mathcal{M}_p$ admits a subsequential limit $\nu\in\mathcal{M}_p$. Invoking Proposition 2.3 to let k tend to infinity in (5.7) completes the proof. ∎

Lemma 5.2. *If $\nu,\nu'\in\mathcal{M}_p$ are such that $G_\nu = G_{\nu'}$, then $\nu = \nu'$.*

Proof. The equality $G_\nu = G_{\nu'}$ means that for all $x\in[-1,1]$,

$$c\log(c) - c + \Delta\overline{m}x + c\sum_{n\ge 2}\frac{(-\Delta/c)^n}{n(n-1)}\int_{-1}^1 y^n\,\mathrm{d}\nu(y)x^n$$

$$= c\log(c) - c + \Delta\overline{m}x + c\sum_{n\ge 2}\frac{(-\Delta/c)^n}{n(n-1)}\int_{-1}^1 y^n\,\mathrm{d}\nu'(y)x^n.$$

It follows that ν and ν' have all the same moments of order $n\ge 2$, and are therefore equal by the Stone-Weierstrass theorem. ∎

Lemma 5.3. *Suppose that the sequence of enriched free energies $(\overline{F}_N)_{N \geq 1}$ converges pointwise to some limit $f : \mathbb{R}_{\geq 0} \times \mathscr{M}_+ \to \mathbb{R}$ along a subsequence $(N_k)_{k \geq 1}$. If $(t, \mu) \in \mathbb{R}_{\geq 0} \times \mathrm{Reg}(\mathscr{M}_+)$ is such that $f(t, \cdot)$ is Gateaux differentiable at μ and $\nu \in \mathscr{M}_p$ is such that $G_\nu = D_\mu f(t, \mu, \cdot)$ (which exists by Lemma 5.1), then there is a sequence $\left(N'_k, \lambda^{N'_k}\right)_{k \geq 1}$ such that the following holds. The sequence $(N'_k)_{k \geq 1} \subset (N_k)_{k \geq 1}$ increases to infinity, the multioverlap concentration (3.15) for the Hamiltonian (4.6) with perturbation parameters $(\lambda^{N'_k})_{k \geq 1}$ holds for every $n \geq 1$, and the asymptotic spin array associated with the perturbation parameters $(\lambda^{N'_k})_{k \geq 1}$ is generated by ν according to (S1)-(S2).*

Proof. Combining Lemma 3.2 with the arguments in [43, Lemma 3.3], it is possible to find a sequence of perturbation parameters $(\lambda^N)_{N \geq 1}$ such that the multioverlap concentration (3.15) for the perturbed Hamiltonian (4.6) holds for all $n \geq 1$. Using the Prokhorov theorem, we can find a subsequence $(N'_k)_{k \geq 1}$ of $(N_k)_{k \geq 1}$ along which the distribution of the spin array $(\sigma^\ell_i)_{i, \ell \geq 1}$ under the Gibbs measure $\mathbb{E}\langle \cdot \rangle'_{N_k}$ with perturbation parameter λ^{N_k} converges in the sense of finite-dimensional distributions. By Proposition 3.3, there exists a probability measure $\nu^* \in \mathscr{M}_p$ such that the asymptotic spin array is generated by ν^* according to (S1)-(S2). We now show that $\nu^* = \nu$. An identical computation to that leading to the derivative expression (2.12) reveals that the Gateaux derivative of the perturbed modified free energy (4.8) is given by

$$D_\mu \overline{F}'_{N'_k}(t, \mu, \lambda^{N'_k}, x) = \left(c + \Delta \overline{m} x\right) \log(c) + c \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} \mathbb{E}\langle R_{1,\dots,n} \rangle'_{N'_k} x^n - c + o_{N'_k}(1).$$

Remembering Lemmas 4.4 and 3.1, and arguing as in the proof of Proposition 2.3 to let k tend to infinity reveals that $D_\mu f(t, \mu, \cdot) = G_{\nu^*}$. Leveraging the assumption that $G_\nu = D_\mu f(t, \mu, \cdot)$ implies that $G_\nu = G_{\nu^*}$. Invoking Lemma 5.2 completes the proof. ∎

Lemma 5.4. *Suppose that the sequence of enriched free energies $(\overline{F}_N)_{N \geq 1}$ converges pointwise to some limit $f : \mathbb{R}_{\geq 0} \times \mathscr{M}_+ \to \mathbb{R}$, and fix $(t, \mu) \in \mathbb{R}_{\geq 0} \times \mathrm{Reg}(\mathscr{M}_+)$. If $f(t, \cdot)$ is Gateaux differentiable at $\mu \in \mathrm{Reg}(\mathscr{M}_+)$ and $\nu \in \mathscr{M}_p$ is such that $G_\nu = D_\mu f(t, \mu, \cdot)$ (which exists by Lemma 5.1), then*

$$f(t, \mu) = \mathscr{P}_{t,\mu}(\nu). \tag{5.8}$$

Proof. Let $(N_k, \lambda^{N^\pm_k})_{k \geq 1}$ and $\nu^\pm \in \mathscr{M}_p$ be as in the proof of Proposition 4.3. By inequality (4.29), it suffices to show that $\nu^- = \nu^+ = \nu$. Since $\nu^\pm$ generate the asymptotic spin array associated with the perturbation parameters $(\lambda^{N^\pm_k})_{k \geq 1}$ according to (S1)-(S2), arguing as in the proof of Lemma 5.3 with ν^* replaced by $\nu^\pm$, it is possible to show that $G_{\nu^\pm} = D_\mu f(t, \mu, \cdot) = G_\nu$. Invoking Lemma 5.2 gives $\nu = \nu^\pm$ and completes the proof. ∎

Lemma 5.5. *Suppose that the sequence of enriched free energies $(\overline{F}_N)_{N \geq 1}$ converges pointwise to some limit $f : \mathbb{R}_{\geq 0} \times \mathscr{M}_+ \to \mathbb{R}$ along a subsequence $(N_k)_{k \geq 1}$, and fix $(t, \mu) \in \mathbb{R}_{\geq 0} \times \mathrm{Reg}(\mathscr{M}_+)$. If $f(t, \cdot)$ is Gateaux differentiable at $\mu \in \mathrm{Reg}(\mathscr{M}_+)$ and $\nu \in \mathscr{M}_p$ is such that $G_\nu = D_\mu f(t, \mu, \cdot)$ (which exists by Lemma 5.1), then*

$$D_\mu f(t, \mu, \cdot) = D_\mu \psi(\mu + t\nu, \cdot). \tag{5.9}$$

Proof. Let $(N'_k, \lambda^{N'_k})_{k \geq 1}$ be as in the statement of Lemma 5.3. To alleviate the exposition, we abuse notation and denote this sequence by $(N, \lambda^N)_{N \geq 1}$. An identical computation to that leading to the

derivative expression (5.4) reveals that the Gateaux derivative of the perturbed free energy (3.10) in the direction of a probability measure $\eta \in \mathrm{Pr}[-1,1]$ is given by

$$D_\mu \overline{F}_{N+1}(t,\mu,\lambda^N;\eta) = \sum_{i \leq N+1} \mathbb{E} \log \left\langle (c + \Delta\sigma_i w_i)^{\widetilde{G}_i^w} \left(1 - \frac{c + \Delta\sigma_i w_i}{N+1}\right)^{1-\widetilde{G}_i^w} \right\rangle_{N+1},$$

where the bracket $\langle \cdot \rangle_{N+1}$ denotes the Gibbs average (3.11) with perturbation parameter λ^N, the $(w_i)_{i \geq 1}$ are i.i.d. random variables sampled from η, and $(\widetilde{G}_i^w)_{i \leq N+1}$ are conditionally independent random variables taking values in $\{0,1\}$ with conditional distribution

$$\mathbb{P}\{\widetilde{G}_i^w = 1 \mid \sigma^*, w_i\} := \frac{c + \Delta\sigma_i^* w_i}{N+1}.$$

Taylor-expanding the logarithm and using the law of large numbers shows that the contribution of the terms in this sum that have $\widetilde{G}_i^w$ equal to zero is

$$-c - \frac{\Delta \mathbb{E} w_1}{N+1} \sum_{i \leq N+1} \sigma_i + \mathcal{O}(N^{-1}) = -c - \Delta\overline{m} \mathbb{E} w_1 + \mathcal{O}(N^{-1}).$$

Together with the symmetry between sites, this implies that

$$D_\mu \overline{F}_{N+1}(t,\mu,\lambda^N;\eta) = (N+1)\mathbb{E}\log\langle (c+\Delta\varepsilon w_1)^{\widetilde{G}_1^w}\rangle_{N+1} - c - \Delta\overline{m}\mathbb{E}w_1. \tag{5.10}$$

The distributional identities (4.20) and (4.22) reveal that the first term in this expression is given by

$$D^1 := \mathbb{E}\log\int_{\Sigma_1} \langle (c+\Delta\varepsilon w_1)^{\widetilde{G}_1^w}\exp z_N^{t,\mu}(\varepsilon,\sigma)\rangle' - \mathbb{E}\log\int_{\Sigma_1}\langle \exp z_N^{t,\mu}(\varepsilon,\sigma)\rangle'\,\mathrm{d}P^*(\varepsilon)$$

for the cavity field $z_N^{t,\mu}$ in (4.14) and the modified Gibbs average (4.9). By Lemma 5.3, the choice of sequence $(N,\lambda^N)_{N \geq 1}$ ensures that the asymptotic spin array $(\sigma_i^\ell)_{i,\ell \geq 1}$ is generated by ν according to (S1)-(S2), so arguing as in Step 1 of the proof of Lemma 4.2 it is possible to show that

$$D^1 = \mathbb{E}\log\int_{\Sigma_1} e^{-c(t+s)-\Delta\varepsilon(t\overline{m}+s\mathbb{E}x_1)}(c+\Delta\varepsilon w_1)^{G'^w}\exp z_N'(\varepsilon)\,\mathrm{d}P^*(\varepsilon)$$

$$- \mathbb{E}\log\int_{\Sigma_1} e^{-c(t+s)-\Delta\varepsilon(t\overline{m}+s\mathbb{E}x_1)}\exp z_N'(\varepsilon)\,\mathrm{d}P^*(\varepsilon) + o_N(1),$$

where $s := \mu[-1,1]$, the cavity field z_N' is defined in (4.27), and G'^w has conditional distribution

$$\mathbb{P}\{G'^w = 1 \mid \varepsilon^*, w_1\} := \frac{c + \Delta\varepsilon^* w_1}{N+1}.$$

Continuing to argue as in Step 1 of the proof of Lemma 4.2 gives

$$D^1 = \mathbb{E}\log\langle (c+\Delta\varepsilon w_1)^{G'^w}\rangle_* + o_N(1) = \frac{1}{N+1}\mathbb{E}\langle c+\Delta\varepsilon w_1\rangle_*\log\langle c+\Delta\varepsilon w_1\rangle_* + o_N(1).$$

Substituting this into (5.10) and leveraging Lemma 3.1 and Proposition 2.3 to let N tend to infinity reveals that

$$D_\mu f(t,\mu;\eta) = \mathbb{E}\langle c+\Delta\varepsilon w_1\rangle_*\log\langle c+\Delta\varepsilon w_1\rangle_* - c - \Delta\overline{m}\mathbb{E}w_1.$$

Remembering the expression (5.1) for the Gateaux derivative of the initial condition completes the proof. ∎

Lemma 5.6. *For any $(t, \mu) \in \mathbb{R}_{\geq 0} \times \mathcal{M}_+$, a measure $\nu \in \mathcal{M}_p$ satisfies $G_\nu = D_\mu \psi(\mu + t\nu, \cdot)$ if and only if it is a fixed point of the map $\nu \mapsto \Gamma_{t,\mu}(\nu)$.*

Proof. We denote by y_1 a random variable sampled from ν. The definition of the function g in (1.24) implies that for every $x \in [-1, 1]$,

$$G_\nu(x) = \int_{-1}^{1} g(xy) \, d\nu^*(y) = (c + \Delta \mathbb{E} y_1 x) \log(c) + c \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} \mathbb{E} y_1^n x^n - c. \qquad (5.11)$$

On the other hand, the derivative expression (5.1) implies that

$$D_\mu \psi(\mu + t\nu, x) = \mathbb{E} g(\langle \varepsilon \rangle_* x) = (c + \Delta \mathbb{E} \langle \varepsilon \rangle_* x) \log(c) + c \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} \mathbb{E} \langle \varepsilon \rangle_*^n x^n - c. \qquad (5.12)$$

The equation $G_\nu = D_\mu \psi(\mu + t\nu, \cdot)$ is therefore equivalent to the equality between all moments of y_1 and $\langle \varepsilon \rangle_*$ of order 2 and higher. By the Stone-Weierstrass theorem, this is in turn equivalent to the equality between the laws of y_1 and $\langle \varepsilon \rangle_*$. These laws are equal if and only if the fixed point equation $\nu = \Gamma_{t,\mu}(\nu)$ is satisfied. This completes the proof. ∎

Proof of Theorem 1.1. Proposition 2.2 gives a sequence $(\mu_n)_{n \geq 1} \subset \mathrm{Reg}(\mathcal{M}_+)$ converging weakly to μ such that $f(t, \cdot)$ is Gateaux differentiable at every $\mu_n \in \mathrm{Reg}(\mathcal{M}_+)$. By Lemmas 5.4 and 5.5, for every $n \geq 1$, there is $\nu_n \in \mathcal{M}_p$ with

$$G_{\nu_n} = D_\mu f(t, \mu_n, \cdot) = D_\mu \psi(\mu + t\nu_n, \cdot) \quad \text{and} \quad f(t, \mu_n) = \mathscr{P}_{t, \mu_n}(\nu_n).$$

By the Prokhorov theorem, up to passing to a subsequence, we can assume that $(\nu_n)_{n \geq 1}$ converges weakly to some limit $\nu^* \in \mathcal{M}_p$. The continuity of the map $\nu \mapsto G_\nu$ and of the Gateaux derivative map $\nu \mapsto D_\mu \psi(\nu, \cdot)$ in (5.1) imply that ν^* satisfies the equation

$$G_{\nu^*} = D_\mu \psi(\mu + t\nu^*, \cdot). \qquad (5.13)$$

Similarly, the Lipschitz continuity of the limit free energy and of the asymptotic initial condition established in Proposition 2.1 imply that

$$f(t, \mu) = \lim_{n \to +\infty} f(t, \mu_n) = \lim_{n \to +\infty} \mathscr{P}_{t, \mu_n}(\nu_n) = \mathscr{P}_{t, \mu}(\nu^*).$$

Remembering that equation (5.13) is equivalent to the fixed point equation $\nu^* = \Gamma_{t,\mu}(\nu^*)$ by Lemma 5.6 completes the proof. ∎

Proof of Theorem 1.2. Lemmas 5.5 and 5.6 imply that ν^* is a fixed point of the map $\nu \mapsto \Gamma_{t,\mu}(\nu)$, and Lemma 5.4 ensures that $f(t, \mu) = \mathscr{P}_{t,\mu}(\nu^*)$. The second part of the statement on the asymptotic spin array follows from Lemma 5.3. This completes the proof. ∎

6 The disassortative and small-time settings

Two limitations of Theorems 1.1 and 1.2 are that they assume the existence of the limit free energy, and that they do not provide a simple characterization of the fixed point for which (1.28) holds. In this section we focus on two settings in which these limitations can be addressed: the disassortative setting when $\Delta \leq 0$ and the small-time setting when $t \in \mathbb{R}_{\geq 0}$ is close to zero.

In the disassortative stochastic block model with two communities, it can be shown that the limit free energy exists and is given by evaluating the functional $\mathscr{P}_{t,\mu}$ defined in (1.26) at the fixed point of the operator $\Gamma_{t,\mu}$ which maximizes $\mathscr{P}_{t,\mu}$. This is the content of Proposition 1.4 which we now prove.

Proof of Proposition 1.4. We fix $v \in \mathscr{M}_p$, and for each $u \in [0,1]$, we define the interpolating free energy

$$\varphi(u) := \overline{F}_N(tu, \mu + t(1-u)v)$$

in such a way that $\varphi(1) = \overline{F}_N(t,\mu)$ and $\varphi(0) = \psi(\mu + tv)$. We denote by x_1 a random variable sampled from v. Recall from the derivative computations in [18, Corollaries 2.2 and 2.5] that

$$\partial_t \overline{F}_N(t,\mu) = \frac{1}{2}\left(c + \Delta \overline{m}^2\right)\log(c) + \frac{c}{2}\sum_{n\geq 2}\frac{(-\Delta/c)^n}{n(n-1)}\mathbb{E}\left\langle R_{[n]}^2\right\rangle - \frac{c}{2} + \mathscr{O}(N^{-1}), \qquad (6.1)$$

$$D_\mu \overline{F}_N(t,\mu,x) = \left(c + \Delta \overline{m}x\right)\log(c) + c\sum_{n\geq 2}\frac{(-\Delta/c)^n}{n(n-1)}\mathbb{E}\langle R_{[n]}\rangle x^n - c + \mathscr{O}(N^{-1}). \qquad (6.2)$$

Together with the assumption that $\Delta \leq 0$ these imply that

$$\varphi'(u) = t\partial_t \overline{F}_N(tu, \mu + t(1-u)v) - t\int_{-1}^1 D_\mu \overline{F}_N(tu, \mu + t(1-u)v)\,dv$$

$$= t\left(\frac{c}{2} - \frac{1}{2}\left(c + \Delta \overline{m}^2\right)\log(c) + \frac{c}{2}\sum_{n\geq 2}\frac{(-\Delta/c)^n}{n(n-1)}\mathbb{E}\left\langle R_{[n]}^2\right\rangle - c\sum_{n\geq 2}\frac{(-\Delta/c)^n}{n(n-1)}\mathbb{E}\langle R_{[n]}\rangle \mathbb{E}x_1^n\right)$$

$$= -\frac{t}{2}\left(\left(c + \Delta \overline{m}^2\right)\log(c) + c\sum_{n\geq 2}\frac{(-\Delta/c)^n}{n(n-1)}\left(\mathbb{E}x_1^n\right)^2 - c\right) + \frac{tc}{2}\sum_{n\geq 2}\frac{(-\Delta/c)^n}{n(n-1)}\mathbb{E}\left\langle (R_{[n]} - \mathbb{E}x_1^n)^2\right\rangle$$

$$\geq -\frac{t}{2}\int_{-1}^1 G_v(y)\,dv(y).$$

Leveraging the fundamental theorem of calculus and letting N tend to infinity gives the lower bound

$$\liminf_{N\to+\infty}\overline{F}_N(t,\mu) \geq \sup_{v\in\mathscr{M}_p}\mathscr{P}_{t,\mu}(v).$$

The matching upper bound is immediate from Proposition 4.3. ∎

In the small-time setting, a Lipschitz continuity bound on the fixed point operator $\Gamma_{t,\mu}$ defined in (1.6) can be combined with the Banach fixed point theorem to show that $\Gamma_{t,\mu}$ admits a unique fixed point $v_{t,\mu} \in \mathscr{M}_p$. Together with a Lipschitz continuity bound on the map $(t,\mu) \mapsto v_{t,\mu}$ and an argument similar to that in Lemma 5.4, this uniqueness can be used to establish Proposition 1.3.

Lemma 6.1. *There exists a constant $C < +\infty$ such that the fixed point operator $\Gamma_{t,\mu} : \mathcal{M}_p \to \mathcal{M}_p$ defined in (1.6) satisfies the following Lipschitz continuity bounds.*

1. *For all $(t,\mu) \in \mathbb{R}_{\geq 0} \times \mathcal{M}_+$, and $\nu, \eta \in \mathcal{M}_p$,*

$$W\big(\Gamma_{t,\mu}(\nu), \Gamma_{t,\mu}(\eta)\big) \leq CtW(\nu, \eta). \tag{6.3}$$

2. *For all $(t,\mu), (t',\mu') \in \mathbb{R}_{\geq 0} \times \mathcal{M}_+$, and $\nu \in \mathcal{M}_p$,*

$$W\big(\Gamma_{t,\mu}(\nu), \Gamma_{t',\mu'}(\nu)\big) \leq C\big(|t' - t| + \mu[-1,1]W\big(\overline{\mu}, \overline{\mu}'\big) + |\mu[-1,1] - \mu'[-1,1]|\big). \tag{6.4}$$

Proof. It will be convenient to remember the dual representation (2.3) of the Wasserstein distance. With this representation in mind, we prove each bound separately.

1. Fix $(t,\mu) \in \mathbb{R}_{\geq 0} \times \mathcal{M}_+$ as well as $\nu, \eta \in \mathcal{M}_p$. Conditionally on the ground truth signal ε^* sampled from P^*, introduce the measures

$$\mathrm{d}\tilde{\nu}(x) := (c + \Delta \varepsilon^* x)\,\mathrm{d}\nu(x) \quad \text{and} \quad \mathrm{d}\tilde{\eta}(x) := (c + \Delta \varepsilon^* x)\,\mathrm{d}\eta(x). \tag{6.5}$$

We denote by $(X_k)_{k\geq 1}$ and $(Y_k)_{k\geq 1}$ i.i.d. samples from the probability measures ν' and η' induced by $\tilde{\nu}$ and $\tilde{\eta}$, and let Π_t be a Poisson random variable with mean $t(c + \Delta \varepsilon^* \overline{m})$. For each $u \in [0,1]$, we define the interpolating Hamiltonian H_u on Σ_1 by

$$H_u(\varepsilon) := -\varepsilon \overline{m} + \sum_{x \in \Pi_{0,\mu}(0)} \log(c + \Delta \varepsilon x) + \sum_{k \leq \Pi_t} \log\big(c + \Delta \varepsilon (X_k + u(Y_k - X_k))\big)$$

as well as the random variable

$$Z_u := \frac{\int_{\Sigma_1} \varepsilon \exp H_u(\varepsilon)\,\mathrm{d}P^*(\varepsilon)}{\int_{\Sigma_1} \exp H_u(\varepsilon)\,\mathrm{d}P^*(\varepsilon)}.$$

The superposition principle ensures that $\Gamma_{t,\mu}(\nu) \overset{d}{=} Z_0$ and $\Gamma_{t,\mu}(\eta) \overset{d}{=} Z_1$. If we write $\langle \cdot \rangle_u$ for the Gibbs measure associated with the Hamiltonian H_u, then

$$\frac{\mathrm{d}}{\mathrm{d}u} Z_u = \sum_{k \leq \Pi_t} (Y_k - X_k)\left(\left\langle \frac{\varepsilon}{c + \Delta\varepsilon(X_k + u(Y_k - X_k))}\right\rangle_u - \langle\varepsilon\rangle_u \left\langle \frac{1}{c + \Delta\varepsilon(X_k + u(Y_k - X_k))}\right\rangle_u\right).$$

It follows by the mean value theorem that for some constant $C < +\infty$ that depends only on c and Δ whose value may not be the same at each occurrence,

$$\mathbb{E}|Z_1 - Z_0| \leq C\mathbb{E}\sum_{k \leq \Pi_t} |Y_k - X_k| \leq tC\mathbb{E}|X_1 - Y_1|.$$

Taking the infimum over all couplings between $(X_k)_{k\geq 1}$ and $(Y_k)_{k\geq 1}$ yields

$$W\big(\Gamma_{t,0}(\nu), \Gamma_{t,0}(\eta)\big) = W(Z_0, Z_1) \leq tCW(\nu', \eta'), \tag{6.6}$$

where we recall that v' and η' denote the probability measures induced by $\tilde{v}$ and $\tilde{\eta}$. We now control the distance between the measures v' and η' by that between the measures v and η. The explicit representation (2.5) of the one-dimensional Wasserstein distance reveals that

$$W(v',\eta') = \int_{-1}^{1} |F_{v'}(x) - F_{\eta'}(x)|\,\mathrm{d}x,$$

and integrating by parts shows that for $x \in [-1,1]$,

$$F_{v'}(x) = \frac{1}{c + \Delta\varepsilon^* \overline{m}}\left(cF_v(x) + \Delta\varepsilon^*\left(xF_v(x) - \int_{-1}^{x} F_v(z)\,\mathrm{d}z\right)\right).$$

It follows that for $x \in [-1,1]$,

$$|F_{v'}(x) - F_{\eta'}(x)|\,\mathrm{d}x \leq C\big(|F_v(x) - F_\eta(x)| + W(v,\eta)\big),$$

and therefore

$$W(v',\eta') \leq CW(v,\eta).$$

Together with (6.6) this establishes the Lipschitz bound (6.3).

2. We fix $(t,\mu),(t',\mu') \in \mathbb{R}_{\geq 0} \times \mathscr{M}_+$ as well as $v \in \mathscr{M}_p$. To simplify notation, we denote by $\tilde{\mu}$, $\tilde{\mu}'$ and $\tilde{v}$ the measures defined as in (6.5), and let $\tilde{s} := \tilde{\mu}[-1,1]$ and $\tilde{s}' := \tilde{\mu}'[-1,1]$. Up to interchanging the roles of t and t' or μ and μ', assume without loss of generality that $t \leq t'$ and $\tilde{s} \leq \tilde{s}'$. For any measure $\eta \in \mathscr{M}_+$ write $\Lambda(\eta)$ for the Poisson point process with intensity measure η. Through a slight abuse of notation, we denote by $(X_k)_{k\geq 1}$ and $(Y_k)_{k\geq 1}$ i.i.d. samples from the probability measures induced by $\tilde{\mu}$ and $\tilde{\mu}'$, and let $\Pi_{\tilde{s}}$ and $\Pi_{\tilde{s}'-\tilde{s}}$ be Poisson random variables with means $\tilde{s}$ and $\tilde{s}'$, respectively. Through another abuse of notation, for each $u \in [0,1]$, we define the interpolating Hamiltonian H_u on Σ_1 by

$$H_u(\varepsilon) := -\varepsilon\overline{m} + \sum_{x\in\Lambda(t\tilde{v})}\log(c + \Delta\varepsilon x) + u\sum_{x\in\Lambda((t'-t)\tilde{v})}\log(c + \Delta\varepsilon x)$$
$$+ \sum_{k\leq\Pi_{\tilde{s}}}\log\big(c + \Delta\varepsilon(X_k + u(Y_k - X_k))\big) + u\sum_{k\leq\Pi_{\tilde{s}'-\tilde{s}}}\log(c + \Delta\varepsilon Y_k)$$

as well as the random variable

$$Z_u := \frac{\int_{\Sigma_1}\varepsilon\exp H_u(\varepsilon)\,\mathrm{d}P^*(\varepsilon)}{\int_{\Sigma_1}\exp H_u(\varepsilon)\,\mathrm{d}P^*(\varepsilon)}.$$

We have implicitly used that $t \leq t'$ and $\tilde{s} \leq \tilde{s}'$ to define H_u. The superposition principle ensures that $\Gamma_{t,\mu}(v) \overset{d}{=} Z_0$ and $\Gamma_{t',\mu'}(v) \overset{d}{=} Z_1$. If we write $\langle\cdot\rangle_u$ for the Gibbs measure associated with the Hamiltonian H_u, then

$$\frac{\mathrm{d}}{\mathrm{d}u}Z_u = \sum_{k\leq\Pi_{\tilde{s}}}(Y_k - X_k)\left(\left\langle\frac{\varepsilon}{c + \Delta\varepsilon(X_k + u(Y_k - X_k))}\right\rangle_u - \langle\varepsilon\rangle_u\left\langle\frac{1}{c + \Delta\varepsilon(X_k + u(Y_k - X_k))}\right\rangle_u\right)$$
$$+ \sum_{x\in\Lambda((t'-t)\tilde{v})}\big(\langle\varepsilon\log(c + \Delta\varepsilon x)\rangle_u - \langle\varepsilon\rangle_u\langle\log(c + \Delta\varepsilon x)\rangle_u\big)$$
$$+ \sum_{k\leq\Pi_{\tilde{s}'-\tilde{s}}}\big(\langle\varepsilon\log(c + \Delta\varepsilon Y_k)\rangle_u - \langle\varepsilon\rangle_u\langle\log(c + \Delta\varepsilon Y_k)\rangle_u\big).$$

42

It follows by the mean value theorem that for some constant $C < +\infty$ that depends only on c and Δ whose value may not be the same at each occurrence,

$$\mathbb{E}|Z_1 - Z_0| \leq C\left(\mathbb{E}\sum_{k \leq \Pi_{\bar{s}}}|Y_k - X_k| + \mathbb{E}|\Lambda((t'-t)\tilde{v})| + \mathbb{E}\Pi_{\bar{s}'-\bar{s}}\right)$$

$$\leq C\left(\mu[-1,1]\mathbb{E}|Y_1 - X_1| + |t'-t| + |\mu'[-1,1] - \mu[-1,1]|\right).$$

Taking the infimum over all couplings between $(X_k)_{k \geq 1}$ and $(Y_k)_{k \geq 1}$ completes the proof. ∎

Lemma 6.2. *There exists a constant $C < +\infty$ such that for all $t < C^{-1}$ and $\mu \in \mathscr{M}_+$, the fixed point operator $\Gamma_{t,\mu} : \mathscr{M}_p \to \mathscr{M}_p$ defined in (1.6) admits a unique fixed point $v_{t,\mu} \in \mathscr{M}_p$. Moreover, for all $(t,\mu),(t',\mu') \in \mathbb{R}_{\geq 0} \times \mathscr{M}_+$ with $t,t' < C^{-1}$, the fixed point map satisfies the Lipschitz bound*

$$W(v_{t,\mu}, v_{t',\mu'}) \leq C\left(|t'-t| + \mu[-1,1]W(\overline{\mu}, \overline{\mu}') + |\mu[-1,1] - \mu'[-1,1]|\right). \tag{6.7}$$

Proof. Let $C < +\infty$ be as in the statement of Lemma 6.1, and fix $L \in (0,1)$. For $t < LC^{-1}$, the Lipschitz bound (6.3) and the Banach fixed point theorem ensure that $\Gamma_{t,\mu}$ admits a unique fixed point $v_{t,\mu} \in \mathscr{M}_p$. Moreover, for any initial measure $v^0 \in \mathscr{M}_p$, the Banach fixed point theorem implies that the sequence $(v_{t,\mu}^n)_{n \geq 1}$ defined iteratively by

$$v_{t,\mu}^0 := v^0 \qquad \text{and} \qquad v_{t,\mu}^{n+1} := \Gamma_{t,\mu}(v_{t,\mu}^n) \text{ for } n \geq 0$$

converges weakly to $v_{t,\mu} \in \mathscr{M}_p$, and satisfies the rate bound

$$W(v_{t,\mu}, v^0) \leq \frac{1}{1-L}W(v_{t,\mu}^1, v^0) = \frac{1}{1-L}W(\Gamma_{t,\mu}(v^0), v^0).$$

Applying this with $v^0 := v_{t',\mu'}$ and recalling that this is the fixed point of $\Gamma_{t',\mu'}$ reveals that

$$W(v_{t,\mu}, v_{t',\mu'}) \leq \frac{1}{1-L}W(\Gamma_{t,\mu}(v_{t',\mu'}), v_{t',\mu'}) = \frac{1}{1-L}W(\Gamma_{t,\mu}(v_{t',\mu'}), \Gamma_{t',\mu'}(v_{t',\mu'}))$$

Invoking the Lipschitz bound (6.4) and redefining the constant $C < +\infty$ completes the proof. ∎

Proof of Proposition 1.3. The Arzelà-Ascoli theorem gives a subsequence $(N_k)_{k \geq 1}$ along which the free energy converges to some limit $f : \mathbb{R}_{\geq 0} \times \mathscr{M}_+ \to \mathbb{R}$. Arguing as in Proposition 4.3, find sequences $(N_k^{\pm}, \lambda^{N_k^{\pm}})_{k \geq 1}$ and measures $v^{\pm} \in \mathscr{M}_p$ such that the following holds. The sequences $(N_k^{\pm})_{k \geq 1} \subset (N_k)_{k \geq 1}$ increase to infinity, the multioverlap concentration (3.15) for the Hamiltonian (4.6) with perturbation parameters $(\lambda^{N_k^{\pm}})_{k \geq 1}$ holds for all $n \geq 1$, the asymptotic spin array associated with the perturbation parameters $(\lambda^{N_k^{\pm}})_{k \geq 1}$ is generated by $v^{\pm}$ according to (S1)-(S2), and

$$\mathscr{P}_{t,\mu}(v^-) \leq f(t,\mu) \leq \mathscr{P}_{t,\mu}(v^+). \tag{6.8}$$

The idea is to now fix $C < +\infty$ as in the statement of Lemma 6.2, and show that for $t < C^{-1}$, the measure $v^{\pm}$ is equal to the unique fixed point $v_{t,\mu} \in \mathscr{M}_p$ of the operator $\Gamma_{t,\mu}$. This shows that the sequence of enriched free energies admits the unique subsequential limit $\mathscr{P}_{t,\mu}(v_{t,\mu})$ for $t < C^{-1}$, and therefore converges to said limit in this small-time regime. It suffices to prove this at

43

points $(t,\mu) \in \mathbb{R}_{\geq 0} \times \mathrm{Reg}(\mathcal{M}_+)$ such that $f(t,\cdot)$ is Gateaux differentiable at μ. Indeed, the equality $f(t,\mu) = \mathscr{P}_{t,\mu}(v_{t,\mu})$ can then be extended to all measures $\mu \in \mathcal{M}_+$ using a density argument relying on Proposition 2.2 and the Lipschitz continuity results in Proposition 2.1 and Lemma 6.2. If $(t,\mu) \in \mathbb{R}_{\geq 0} \times \mathrm{Reg}(\mathcal{M}_+)$ is such that $f(t,\cdot)$ is Gateaux differentiable at μ and $v \in \mathcal{M}_p$ is such that $G_v = D_\mu f(t,\mu,\cdot)$, then arguing as in the proof of Lemma 5.3 with v^* replaced by $v^\pm$ and invoking Lemma 5.5, it is possible to show that

$$G_{v^\pm} = G_v = D_\mu f(t,\mu,\cdot) = D_\mu \psi(\mu + tv, \cdot).$$

By Lemma 5.2, the first of these equalities implies that $v^\pm = v$ which means that $v^\pm$ satisfies the fixed point equation

$$G_{v^\pm} = D_\mu \psi(\mu + tv^\pm, \cdot).$$

Invoking Lemma 5.6 reveals that $v^\pm$ is a fixed point of the operator $\Gamma_{t,\mu}$, and is therefore equal to the unique fixed point $v_{t,\mu} \in \mathcal{M}_p$ whenever $t < C^{-1}$. Together with (6.8) this completes the proof. ∎

We close this section by showing that the limit free energy is differentiable about the origin. The generalization of this result to the bipartite version of the stochastic block model will be used in Proposition 7.1 to argue that the generalization of the variational formula (1.33) does not hold in this setting.

Corollary 6.3. *If $C < +\infty$ is as in the statement of Proposition 1.3, and $f : [0, C^{-1}] \times \mathcal{M}_+ \to \mathbb{R}$ denotes the limit of the sequence of enriched free energies $(\overline{F}_N)_{N \geq 1}$ on $[0, C^{-1}] \times \mathcal{M}_+$, then the map $t \mapsto f(t,0)$ is differentiable at $t = 0$.*

Proof. We fix $t < C^{-1}$, and denote by $v_t \in \mathcal{M}_p$ the unique fixed point of the operator $\Gamma_{t,0}$. Observe that

$$f(t,0) - f(0,0) = \mathscr{P}_{t,0}(v_0) - \mathscr{P}_{0,0}(v_0) + \mathscr{P}_{t,0}(v_t) - \mathscr{P}_{t,0}(v_0). \tag{6.9}$$

On the one hand,

$$\mathscr{P}_{t,0}(v_0) - \mathscr{P}_{0,0}(v_0) = \psi(tv_0) - \psi(0) - \frac{t}{2}\int_{-1}^{1} G_{v_0}(y)\, dv_0(y). \tag{6.10}$$

On the other hand, if $v_{u,t} := (1-u)v_0 + uv_t$, then the fundamental theorem of calculus implies that

$$\mathscr{P}_{t,0}(v_t) - \mathscr{P}_{t,0}(v_0) = \int_0^1 D_\mu \mathscr{P}_{t,0}(v_{u,t}; v_t - v_0)\, du = \int_0^1 \int_{-1}^1 D_\mu \mathscr{P}_{t,0}(v_{u,t},x)\, d(v_t - v_0)(x)\, du.$$

A direct computation reveals that for all $u \in [0,1]$ and $x \in [-1,1]$,

$$D_\mu \mathscr{P}_{t,0}(v_{u,t},x) = t\big(D_\mu \psi(tv_{u,t},x) - G_{v_{u,t}}(x)\big).$$

Together with the mean value theorem and the derivative bound on the density of the initial condition's Gateaux derivative implied by (2.22), this means that there exists a constant $L > 0$ such that

44

the map $f_{u,t}(x) := D_\mu \mathscr{P}_{t,0}(v_{u,t},x)$ is Lipschitz continuous with $\|f_{u,t}\|_{\text{Lip}} \leq Lt$. It follows by definition of the Wasserstein distance (2.1) and the Lipschitz continuity of the map $t \mapsto v_t$ established in Lemma 6.2 that, up to redefining the constant L,

$$|\mathscr{P}_{t,0}(v_t) - \mathscr{P}_{t,0}(v_0)| \leq Lt^2.$$

Together with (6.9)-(6.10), this implies that the map $t \mapsto f(t,0)$ is differentiable at the origin with

$$\partial_t f(0,0) = D_\mu \psi(0; v_0) - \frac{1}{2} \int_{-1}^{1} G_{v_0}(y) \, dv_0(y).$$

This completes the proof. ∎

7 Critical point selection

In this section, we argue that we do not expect the variational formula

$$\lim_{N \to +\infty} \overline{F}_N(t,\mu) = \sup_{v \in \mathscr{M}_p} \mathscr{P}_{t,\mu}(v) \tag{7.1}$$

to hold for the general sparse stochastic block model. More precisely, we will argue that it does not generalize to a bipartite version of the stochastic block model which we now describe.

To motivate the definition of the bipartite version of the stochastic block model, consider a group of N students indexed by the elements $\{1,\ldots,N\}$ and N tutors also indexed by the elements $\{1,\ldots,N\}$. Each individual is assigned to an in-person group or a virtual group, depending on whether they prefer in-person or virtual classes. The idea is to match students preferring in-person instruction with tutors willing to deliver it, and similarly for students preferring virtual instruction. An allocation of the $2N$ students and tutors into the two groups can be identified with a vector

$$(x,y) := ((x_1,x_2,\ldots,x_N),(y_1,y_2,\ldots,y_N)) \in \Sigma_N^2, \tag{7.2}$$

where $x \in \Sigma_N$ encodes the allocation of students, $y \in \Sigma_N$ denotes the allocation of tutors, the label $+1$ represents the group preferring in-person instruction, and the label -1 represents the group preferring virtual instruction. We will assume that the labels $X_i \sim P^X$ and $Y_i \sim P^Y$ are taken to be i.i.d. from two probability measures P^X and P^Y on Σ_1. This means that the vector (X,Y) is sampled from the product distribution

$$P_N^X \otimes P_N^Y := (P^X)^{\otimes N} \otimes (P^Y)^{\otimes N}. \tag{7.3}$$

Using the assignment vector (X,Y), a random undirected bipartite graph $\mathbf{G}_N := (G_{ij})_{i,j \leq N}$ with vertex sets $\{1,\ldots,N\}$ and $\{1,\ldots,N\}$ is constructed by stipulating that an edge between nodes $i,j \in \{1,\ldots,N\}$ in different groups is present with conditional probability

$$\mathbb{P}\{G_{i,j} = 1 \mid X,Y\} := \frac{c + \Delta X_i Y_j}{N}, \tag{7.4}$$

45

where $c > 0$ and $\Delta \in (-c, c) \setminus \{0\}$ are fixed parameters. The likelihood of the model is given by

$$\mathbb{P}\{\mathbf{G}_N = (G_{ij})_{i,j \leq N} \mid X = x, Y = y\} = \prod_{i,j=1}^{N} \left(\frac{c + \Delta X_i Y_j}{N}\right)^{G_{ij}} \left(1 - \frac{c + \Delta X_i Y_j}{N}\right)^{1 - G_{ij}}, \qquad (7.5)$$

so Bayes' formula implies that the posterior of the model is the Gibbs measure

$$\mathbb{P}\{X = x, Y = y \mid \mathbf{G}_N = (G_{ij})_{i,j \leq N}\} = \frac{\exp H_N(x,y)}{\int_{\Sigma_N^2} \exp H_N(x', y') \, dP_N^X(x') \, dP_N^Y(y')} \qquad (7.6)$$

associated with the Hamiltonian on Σ_N^2 defined by

$$H_N^\circ(x,y) := \sum_{i,j=1}^{N} \log\left[(c + \Delta x_i y_j)^{G_{ij}} \left(1 - \frac{c + \Delta x_i y_j}{N}\right)^{1 - G_{ij}}\right]. \qquad (7.7)$$

The free energy corresponding to this Hamiltonian is

$$\overline{F}_N^\circ := \frac{1}{N} \mathbb{E} \log \int_{\Sigma_N^2} \exp H_N^\circ(x,y) \, dP_N^X(x) \, dP_N^Y(y). \qquad (7.8)$$

As in Section 1, it will be convenient to modify the free energy (7.8) without changing its limiting value, and then enrich it so that it depends on a time parameter $t \geq 0$ and measure parameters $\mu_1, \mu_2 \in \mathscr{M}_+$. For each $t \geq 0$, we introduce a random variable $\Pi_t \sim \text{Poi}(tN^2)$ as well as an independent family of i.i.d. random matrices $(G^k)_{k \geq 1}$ each having conditionally independent entries $(G_{i,j}^k)_{i,j \leq N}$ taking values in $\{0, 1\}$ with conditional distribution

$$\mathbb{P}\{G_{i,j}^k = 1 \mid X, Y\} := \frac{c + \Delta X_i Y_j}{N}. \qquad (7.9)$$

Given a collection of random indices $(i_k, j_k)_{k \geq 1}$ sampled uniformly at random from $\{1, \ldots, N\}^2$, independently of the other random variables, we define the time-dependent Hamiltonian H_N^t on Σ_N^2 by

$$H_N^t(x,y) := \sum_{k \leq \Pi_t} \log\left[(c + \Delta x_{i_k} y_{j_k})^{G_{i_k, j_k}^k} \left(1 - \frac{c + \Delta x_{i_k} y_{j_k}}{N}\right)^{1 - G_{i_k, j_k}^k}\right], \qquad (7.10)$$

and denote by

$$\overline{F}_N(t) := \frac{1}{N} \mathbb{E} \log \int_{\Sigma_N} \exp H_N^t(x,y) \, dP_N^X(x) \, dP_N^Y(y) \qquad (7.11)$$

its corresponding free energy. This is the Hamiltonian associated with the task of inferring the signal (X, Y) from the data

$$\mathscr{D}_N^t := \left(\Pi_t, (i_k, j_k)_{k \leq \Pi_t}, (G_{i_k, j_k}^k)_{k \leq \Pi_t}\right). \qquad (7.12)$$

Given $\mu_1, \mu_2 \in \mathscr{M}_+$, we consider sequences $(\Lambda_i(N\mu_1))_{i \leq N}$ and $(\Lambda_i(N\mu_2))_{i \leq N}$ of independent Poisson point processes with intensity measures $N\mu_1$ and $N\mu_2$, and we define the measure-dependent

Hamiltonian $H_N^{\mu_1,\mu_2}$ on Σ_N^2 by

$$H_N^{\mu_1,\mu_2}(x,y) := \sum_{i \leq N} \sum_{z \in \Lambda_i(N\mu_1)} \log\left[(c+\Delta x_i z)^{G_i^{z,X}} \left(1 - \frac{c+\Delta x_i z}{N}\right)^{1-G_i^{z,X}}\right]$$
$$+ \sum_{i \leq N} \sum_{z \in \Lambda_i(N\mu_2)} \log\left[(c+\Delta y_i z)^{G_i^{z,Y}} \left(1 - \frac{c+\Delta y_i z}{N}\right)^{1-G_i^{z,Y}}\right], \quad (7.13)$$

where $(G_i^{z,X})_{i \leq N}$ and $(G_i^{z,Y})_{i \leq N}$ are conditionally independent random variables taking values in $\{0,1\}$ with conditional distributions

$$\mathbb{P}\{G_i^{z,X} = 1 \mid X, z\} := \frac{c + \Delta X_i z}{N} \quad \text{and} \quad \mathbb{P}\{G_i^{z,Y} = 1 \mid Y, z\} := \frac{c + \Delta Y_i z}{N}. \quad (7.14)$$

This measure-dependent Hamiltonian is the Hamiltonian associated with the task of inferring the signal (X,Y) from the data

$$\mathscr{D}_N^{\mu_1,\mu_2} := \left(\Lambda_i(N\mu_1), (G_i^{z,X})_{z \in \Lambda_i(N\mu_1)}, \Lambda_i(N\mu_2), (G_i^{z,Y})_{z \in \Lambda_i(N\mu_2)}\right)_{i \leq N}. \quad (7.15)$$

Finally, for each $(t,\mu_1,\mu_2) \in \mathbb{R}_{\geq 0} \times \mathscr{M}_+^2$, we introduce the enriched Hamiltonian

$$H_N^{t,\mu_1,\mu_2}(x,y) := H_N^t(x,y) + H_N^{\mu_1,\mu_2}(x,y) \quad (7.16)$$

as well as its associated free energy

$$\overline{F}_N(t,\mu_1,\mu_2) := \frac{1}{N} \mathbb{E} \log \int_{\Sigma_N^2} \exp H_N^{t,\mu_1,\mu_2}(x,y) \, dP_N^X(x) \, dP_N^Y(y). \quad (7.17)$$

The Hamiltonian (7.16) is associated with the inference of the signal (X,Y) from the data

$$\mathscr{D}_N^{t,\mu_1,\mu_2} := (\mathscr{D}_N^t, \mathscr{D}_N^{\mu_1,\mu_2}), \quad (7.18)$$

where the randomness in these two datasets is taken to be independent conditionally on (X,Y). As usual, we write $\langle \cdot \rangle$ for the Gibbs average associated with the Hamiltonian (7.16). This means that for any bounded and measurable function $f = f(x^1,y^1 \ldots, x^n, y^n)$ of finitely many replicas,

$$\langle f \rangle := \frac{\int_{\Sigma_N^{2n}} f(x^1,y^1,\ldots,x^n,y^n) \prod_{\ell \leq n} \exp H_N^{t,\mu_1,\mu_2}(x^\ell, y^\ell) \, dP_N^X(x^\ell) \, dP_N^Y(y^\ell)}{\left(\int_{\Sigma_N^2} \exp H_N^{t,\mu_1,\mu_2}(x,y) \, dP_N^X(x) \, dP_N^Y(y)\right)^n}. \quad (7.19)$$

The enriched free energy in (7.17) is the main object of study in the bipartite stochastic block model much in the same way as the enriched free energy (1.22) was the main object of study in the two-community stochastic block model.

To show that the variational formula (7.1) does not generalize to the setting of the bipartite stochastic block model, we focus on the case with

$$P^X := \text{Ber}(1/2) \quad \text{and} \quad P^Y := \text{Ber}(p) \quad (7.20)$$

for some $p \neq 1/2$. This means that the labels X_i are centered while the labels Y_i each have non-zero mean

$$\overline{m} := 2p - 1 \neq 0. \tag{7.21}$$

Arguing as in Theorem 1.1, it can be shown that, if the sequence of enriched free energies $(\overline{F}_N)_{N \geq 1}$ converges pointwise to some limit $f : \mathbb{R}_{\geq 0} \times \mathcal{M}_+^2 \to \mathbb{R}$, then this limit free energy is given by evaluating the functional $\mathscr{P}_{t,\mu_1,\mu_2} : \mathcal{M}_+^2 \to \mathbb{R}$ defined by

$$\mathscr{P}_{t,\mu_1,\mu_2}(\nu_1, \nu_2) := \psi_X(\mu_1 + t\nu_1) + \psi_Y(\mu_2 + t\nu_2) - t \int_{-1}^1 \int_{-1}^1 g(xy)\, d\nu_1\, d\nu_2 \tag{7.22}$$

at some critical point $(\nu_1, \nu_2) \in \mathcal{M}_{1/2} \times \mathcal{M}_p$ of this functional. The initial conditions $\psi_X : \mathcal{M}_+ \to \mathbb{R}$ and $\psi_Y : \mathcal{M}_+ \to \mathbb{R}$ are given by (1.27) for the priors P^X and P^Y, respectively. The generalization of the variational formula (7.1) would then be that for all $(t, \mu_1, \mu_2) \in \mathbb{R}_{\geq 0} \times \mathcal{M}_+^2$, we have

$$\lim_{N \to +\infty} \overline{F}_N(t, \mu_1, \mu_2) = \sup_{\nu_1 \in \mathcal{M}_{1/2}} \sup_{\nu_2 \in \mathcal{M}_p} \mathscr{P}_{t,\mu_1,\mu_2}(\nu_1, \nu_2). \tag{7.23}$$

We now argue that this formula cannot hold for small $t > 0$.

Proposition 7.1. *If $\Delta > 0$, $c > e$ and $\overline{m}$ is sufficiently small, then the variational formula* (7.23) *for $\mu_1 = \mu_2 = 0$ does not hold in a neighborhood of $t = 0$.*

Proof. The strategy will be to show that the time derivative of the limit free energy at the origin, which can be shown to exist through similar arguments to those in the proof of Corollary 6.3, does not agree with the time derivative implied by the variational formula (7.23). Determining the time derivative of the enriched free energy (7.17) can be done through a similar computation to that in [18, Lemma 2.1], and reveals that for all $(t, \mu_1, \mu_2) \in \mathbb{R}_{\geq 0} \times \mathcal{M}_+^2$,

$$\partial_t \overline{F}_N(t, \mu_1, \mu_2) = \mathbb{E}\left(c + \Delta \langle x_i y_j \rangle\right) \log\left(c + \Delta \langle x_i y_j \rangle\right) - c + \mathcal{O}(N^{-1}).$$

The assumption that X is centered and independent of Y has played its part. Together with the fact that the Gibbs measure associated with the parameters $(t, \mu_1, \mu_2) = (0, 0, 0)$ is simply the average with respect to the prior $P^X \otimes P^Y$ and an analogue of Proposition 2.3, this implies that the time derivative of the limit free energy $f : \mathbb{R}_{\geq 0} \times \mathcal{M}_+^2 \to \mathbb{R}$ at the origin is given by

$$\partial_t f(0,0,0) = \mathbb{E}(c + \Delta \mathbb{E}(X_1 Y_1)) \log(c + \Delta \mathbb{E}(X_2 Y_2)) - c = c \log(c) - c. \tag{7.24}$$

On the other hand, the validity of the variational formula (7.23) in a neighborhood of $t = 0$ implies that for any $t > 0$ in such a neighborhood,

$$\frac{f(t,0,0) - f(0,0,0)}{t} \geq \frac{\mathscr{P}_{t,0,0}(\delta_0, \delta_{\overline{m}})}{t} = \frac{\psi_X(t\delta_0) + \psi_Y(t\delta_{\overline{m}}) - tg(0)}{t} = \frac{\psi_Y(t\delta_{\overline{m}})}{t},$$

where we have used that $f(0,0,0) = 0$ and $\psi_X(t\delta_0) = tg(0) = t(c\log(c) - c)$. Letting t tend to zero and recalling the expression (5.1) for the directional derivative of the asymptotic initial condition gives the lower bound

$$\partial_t f(0,0,0) \geq D_\mu \psi_Y(0; \delta_{\overline{m}}) = \mathbb{E}\langle c + \Delta \varepsilon \overline{m} \rangle_* \log\langle c + \Delta \varepsilon \overline{m} \rangle_* - c - \Delta \overline{m}^2.$$

48

Noticing that the Gibbs average $\langle \cdot \rangle_*$ is simply the average with respect to the prior P^Y and Taylor-expanding the logarithm implies that

$$\partial_t f(0,0,0) \geq (c + \Delta \overline{m}^2) \log(c + \Delta \overline{m}^2) - c - \Delta \overline{m}^2 = c \log(c) - c + \Delta \overline{m}^2 \log(c) + \mathscr{O}(\overline{m}^4).$$

Whenever $\overline{m}$ is small enough, $c > e$ and $\Delta > 0$, we therefore have

$$\partial_t f(0,0,0) > c \log(c) - c.$$

This contradicts (7.24) and completes the proof. ∎

References

[1] Emmanuel Abbe, Elisabetta Cornacchia, Yuzhou Gu, and Yury Polyanskiy. Stochastic block model entropy and broadcasting on trees with survey. In *Proceedings of Thirty Fourth Conference on Learning Theory*, volume 134, pages 1–25. PMLR, 2021.

[2] Emmanuel Abbe and Andrea Montanari. Conditional random fields, planted constraint satisfaction, and entropy concentration. *Theory of Computing*, 11:413–443, 12 2015.

[3] Emmanuel Abbe and Colin Sandon. Detection in the stochastic block model with multiple clusters: proof of the achievability conjectures, acyclic bp, and the information-computation gap. Preprint, arXiv:1512.09080.

[4] Jean Barbier, Mohamad Dia, Nicolas Macris, Florent Krzakala, Thibault Lesieur, and Lenka Zdeborová. Mutual information for symmetric rank-one matrix estimation: A proof of the replica formula. In *Advances in Neural Information Processing Systems (NIPS)*, volume 29, pages 424–432, 2016.

[5] Jean Barbier and Nicolas Macris. The adaptive interpolation method: a simple scheme to prove replica formulas in Bayesian inference. *Probab. Theory Related Fields*, 174(3-4):1133–1185, 2019.

[6] Jean Barbier, Nicolas Macris, and Léo Miolane. The layered structure of tensor estimation and its mutual information. In *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1056–1063. IEEE, 2017.

[7] Jean Barbier and Dmitry Panchenko. Free energy in the stochastic block model in sparse regime: Disassortative case. *Personal communication*.

[8] Jean Barbier and Dmitry Panchenko. Strong replica symmetry in high-dimensional optimal Bayesian inference. *Comm. Math. Phys.*, 393(3):1199–1239, 2022.

[9] Adriano Barra, Pierluigi Contucci, Emanuele Mingione, and Daniele Tantari. Multi-species mean field spin glasses. Rigorous results. *Ann. Henri Poincaré*, 16(3):691–708, 2015.

[10] Hong-Bin Chen. Hamilton-Jacobi equations for nonsymmetric matrix inference. *Ann. Appl. Probab.*, 32(4):2540–2567, 2022.

[11] Hong-Bin Chen and Jean-Christophe Mourrat. On the free energy of vector spin glasses with non-convex interactions. Preprint, arXiv:2311.08980.

[12] Hong-Bin Chen and Jiaming Xia. Hamilton-Jacobi equations for inference of matrix tensor products. *Ann. Inst. Henri Poincaré Probab. Stat.*, 58(2):755–793, 2022.

[13] Hong-Bin Chen and Jiaming Xia. Limiting free energy of multi-layer generalized linear models. Preprint, arXiv:2108.12615.

[14] Hongbin Chen, Jean-Christophe Mourrat, and Jiaming Xia. Statistical inference of finite-rank tensors. *Ann. H. Lebesgue*, 5:1161–1189, 2022.

[15] Amin Coja-Oghlan, Florent Krzakala, Will Perkins, and Lenka Zdeborová. Information-theoretic thresholds from the cavity method. *Adv. Math.*, 333:694–795, 2018.

[16] Aurelien Decelle, Florent Krzakala, Cristopher Moore, and Lenka Zdeborová. Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications. *Physical Review E*, 84:066106, 2011.

[17] Yash Deshpande, Emmanuel Abbe, and Andrea Montanari. Asymptotic mutual information for the balanced binary stochastic block model. *Inf. Inference*, 6(2):125–170, 2017.

[18] Tomas Dominguez and Jean-Christophe Mourrat. Mutual information for the sparse stochastic block model. *Ann. Probab.*, 52(2):434–501, 2024.

[19] Tomas Dominguez and Jean-Christophe Mourrat. *Statistical mechanics of mean-field disordered systems: a Hamilton-Jacobi approach*. Zurich Lectures in Advanced Mathematics. European Mathematical Society, Zürich, 2024.

[20] Tomas Dominguez and Jean-Christophe Mourrat. Infinite-dimensional Hamilton-Jacobi equations for statistical inference on sparse graphs. *SIAM J. Math. Anal.*, to appear.

[21] Yuzhou Gu and Yury Polyanskiy. Uniqueness of BP fixed point for the Potts model and applications to community detection. In *The Thirty Sixth Annual Conference on Learning Theory*, pages 837–884. PMLR, 2023.

[22] Francesco Guerra. Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.*, 233(1):1–12, 2003.

[23] Jonathan Kadmon and Surya Ganguli. Statistical mechanics of low-rank tensor decomposition. In *Advances in Neural Information Processing Systems*, pages 8201–8212, 2018.

[24] Varun Kanade, Elchanan Mossel, and Tselil Schramm. Global and local information in clustering labeled block models. *IEEE Transactions on Information Theory*, 62(10):5906–5917, 2016.

[25] Anastasia Kireeva and Jean-Christophe Mourrat. Breakdown of a concavity property of mutual information for non-gaussian channels. Preprint, arXiv:2304.05129.

[26] Marc Lelarge and Léo Miolane. Fundamental limits of symmetric low-rank matrix estimation. *Probab. Theory Related Fields*, 173(3-4):859–929, 2019.

[27] Thibault Lesieur, Léo Miolane, Marc Lelarge, Florent Krzakala, and Lenka Zdeborová. Statistical and computational phase transitions in spiked tensor estimation. In *2017 IEEE International Symposium on Information Theory (ISIT)*, pages 511–515. IEEE, 2017.

[28] Clément Luneau, Jean Barbier, and Nicolas Macris. Mutual information for low-rank even-order symmetric tensor estimation. *Inf. Inference*, 10(4):1167–1207, 2021.

[29] Clément Luneau, Nicolas Macris, and Jean Barbier. High-dimensional rank-one nonsymmetric matrix decomposition: the spherical case. In *2020 IEEE International Symposium on Information Theory (ISIT)*, pages 2646–2651. IEEE, 2020.

[30] Laurent Massoulié. Community detection thresholds and the weak Ramanujan property. In *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, page 694–703. Association for Computing Machinery, 2014.

[31] Vaishakhi Mayya and Galen Reeves. Mutual information in community detection with covariate information and correlated networks. In *2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 602–607. IEEE, 2019.

[32] Léo Miolane. Fundamental limits of low-rank matrix estimation: the non-symmetric case. Preprint, arXiv:1702.00473.

[33] Elchanan Mossel, Joe Neeman, and Allan Sly. Reconstruction and estimation in the planted partition model. *Probab. Theory Related Fields*, 162(3-4):431–461, 2015.

[34] Elchanan Mossel, Joe Neeman, and Allan Sly. Belief propagation, robust reconstruction and optimal recovery of block models. *Ann. Appl. Probab.*, 26(4):2211–2256, 2016.

[35] Elchanan Mossel, Joe Neeman, and Allan Sly. A proof of the block model threshold conjecture. *Combinatorica*, 38(3):665–708, 2018.

[36] Elchanan Mossel and Jiaming Xu. Local algorithms for block models with side information. In *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science*, pages 71–80, 2016.

[37] Jean-Christophe Mourrat. Hamilton-Jacobi equations for finite-rank matrix inference. *Ann. Appl. Probab.*, 30(5):2234–2260, 2020.

[38] Jean-Christophe Mourrat. Hamilton-Jacobi equations for mean-field disordered systems. *Ann. H. Lebesgue*, 4:453–484, 2021.

[39] Jean-Christophe Mourrat. Nonconvex interactions in mean-field spin glasses. *Probab. Math. Phys.*, 2(2):281–339, 2021.

[40] Jean-Christophe Mourrat. Free energy upper bound for mean-field vector spin glasses. *Ann. Inst. Henri Poincaré Probab. Stat.*, 59(3):1143–1182, 2023.

[41] Victor Panaretos and Yoav Zemel. *An Invitation to Statistics in Wasserstein Space*. Springer-Briefs in Probability and Mathematical Statistics. Springer International Publishing, 2020.

[42] Dmitry Panchenko. The Parisi ultrametricity conjecture. *Ann. of Math. (2)*, 177(1):383–393, 2013.

[43] Dmitry Panchenko. *The Sherrington-Kirkpatrick model*. Springer Monographs in Mathematics. Springer, New York, 2013.

[44] Dmitry Panchenko. The Parisi formula for mixed p-spin models. *Ann. Probab.*, 42(3):946–958, 2014.

[45] Dmitry Panchenko. The free energy in a multi-species Sherrington-Kirkpatrick model. *Ann. Probab.*, 43(6):3494–3513, 2015.

[46] Dmitry Panchenko. Free energy in the Potts spin glass. *Ann. Probab.*, 46(2):829–864, 2018.

[47] Dmitry Panchenko. Free energy in the mixed p-spin models with vector spins. *Ann. Probab.*, 46(2):865–896, 2018.

[48] Dmitry Panchenko. *Lecture notes on probability theory*. 2019.

[49] R. R. Phelps. Gaussian null sets and differentiability of Lipschitz map on Banach spaces. *Pacific J. Math.*, 77(2):523–531, 1978.

[50] Galen Reeves. Information-theoretic limits for the matrix tensor product. *IEEE Journal on Selected Areas in Information Theory*, 1(3):777–798, 2020.

[51] Galen Reeves, Vaishakhi Mayya, and Alexander Volfovsky. The geometry of community detection via the mmse matrix. In *2019 IEEE International Symposium on Information Theory (ISIT)*, pages 400–404. IEEE, 2019.

[52] Michel Talagrand. The Parisi formula. *Ann. of Math. (2)*, 163(1):221–263, 2006.

[53] Qian Yu and Yury Polyanskiy. Ising model on locally tree-like graphs: uniqueness of solutions to cavity equations. *IEEE Transactions on Information Theory*, 70(3):1913–1938, 2024.