

Mutual information for 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 with two communities. We aim to describe the mutual information between the observed network and the actual community structure in the sparse regime, where the total number of nodes diverges while the average degree of a given node remains bounded. Our main contributions are a conjecture for the limit of this quantity, which we express in terms of a Hamilton-Jacobi equation posed over a space of probability measures, and a proof that this conjectured limit provides a lower bound for the asymptotic mutual information. The well-posedness of the Hamilton-Jacobi equation is obtained in our companion paper. In the case when links across communities are more likely than links within communities, the asymptotic mutual information is known to be given by a variational formula. We also show that our conjectured limit coincides with this formula in this case.

1 Introduction and main results

The stochastic block model is the simplest generative model for networks with a community structure. It was first introduced in the machine learning and statistics literature [34, 37, 70, 71] but soon emerged independently in a variety of other scientific disciplines. In the theoretical computer science community it is often termed the planted partition model [16, 18, 32] while the mathematics literature often refers to it as the inhomogeneous random graph model [15]. Since its introduction, the stochastic block model has become a test bed for clustering and community detection algorithms used in social networks [59], protein-to-protein interaction networks [27], recommendation systems [44], medical prognosis [69], DNA folding [19], image segmentation [68] and natural language processing [8] among others. In this paper we focus on the sparse stochastic block model with two communities which we now describe.

Consider N individuals belonging to exactly one of two communities; it will be convenient to describe the communities using vectors of ± 1 with the agreement that people with the same label belong to the same group. In this way, a vector

$$\sigma^* = (\sigma_1^*, \dots, \sigma_N^*) \in \Sigma_N = \{-1, +1\}^N \quad (1.1)$$

can be used to encode the two communities. The labels $\sigma_i^* \sim P^*$ are taken to be i.i.d. Bernoulli random variables with probability of success $p \in (0, 1)$ and expectation \bar{m} ,

$$p = P^*(1) = \mathbb{P}\{\sigma_i^* = 1\} \quad \text{and} \quad \bar{m} = \mathbb{E} \sigma_1^* = 2p - 1. \quad (1.2)$$

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

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

The case $p = 1/2$ is termed the symmetric stochastic block model, and for reasons that will become apparent below has received the greatest attention. The assignment vector σ^* follows a product distribution,

$$\sigma^* \sim P_N^* = (P^*)^{\otimes N}, \quad (1.3)$$

so the expected sizes of the communities are Np and $N(1-p)$. Using the assignment vector σ^* , a random undirected graph $\mathbf{G}_N = (G_{ij})_{i,j \leq N}$ with vertex set $\{1, \dots, N\}$ is constructed by stipulating that an edge between node i and node j is present with conditional probability

$$\mathbb{P}\{G_{ij} = 1 | \sigma^*\} = \begin{cases} a_N & \text{if } \sigma_i^* = \sigma_j^* \\ b_N & \text{if } \sigma_i^* \neq \sigma_j^* \end{cases} \quad (1.4)$$

for some $a_N, b_N \in (0, 1)$ independently of all other edges. In other words, the probability that an edge is present between node i and node j depends only on whether or not the individuals i and j belong to the same community. To express (1.4) more succinctly, it is convenient to introduce the average and the gap of a_N and b_N ,

$$c_N = \frac{a_N + b_N}{2} \quad \text{and} \quad \Delta_N = \frac{a_N - b_N}{2} \in (-c_N, c_N), \quad (1.5)$$

in such a way that

$$\mathbb{P}\{G_{ij} = 1 | \sigma^*\} = c_N + \sigma_i^* \sigma_j^* \Delta_N. \quad (1.6)$$

The data $\mathbf{G}_N = (G_{ij})$ is said to be sampled from the stochastic block model, and the inference task is to recover the assignment vector σ^* as accurately as possible given the graph \mathbf{G}_N . In the case when $\Delta_N \leq 0$, it is more likely for an edge to be present between nodes in different communities and the model is called *disassortative*. When $\Delta_N > 0$ connections are more likely between individuals in the same community and the model is termed *assortative*.

Recently, the stochastic block model has attracted much renewed attention. On a practical level, we mention for instance extensions allowing for overlapping communities [7] that have proved to be a good fit for real data sets in massive networks [35]. On a theoretical level, the predictions put forth in [29] using deep but non-rigorous statistical physics arguments have been particularly stimulating. The theoretical study of the stochastic block model has seen significant progress in two main directions: exact recovery and detection. The exact recovery task aims to determine the regimes of a_N and b_N for which there exists an algorithm that completely recovers the two communities with high probability. Of course, a necessary condition for exact recovery is connectivity of the random graph \mathbf{G}_N ; this makes exact recovery impossible in the sparse regime. The sharp threshold for exact recovery was obtained in [2, 52], where it was shown that in the symmetric dense regime, $p = 1/2$, $a_N = a \log(N)/N$ and $b_N = b \log(N)/N$, exact recovery is possible, and efficiently so, if and only if $\sqrt{a} - \sqrt{b} \geq 2$. On the other hand, the detection task is to construct a partition of the graph \mathbf{G}_N that is positively correlated with the assignment vector σ^* with high probability. The sharp threshold for detection in the sparse regime was obtained in [47, 50, 53], where it was shown that in the symmetric sparse regime, $p = 1/2$, $a_N = a/N$ and $b_N = b/N$, detection is solvable, and efficiently so, if and only if $(a - b)^2 > 2(a + b)$. Notice that detection is much easier in the asymmetric case [20]. Indeed, the expected degree of node i conditional on its community membership is given by

$$\mathbb{E}[\deg(i) | \sigma_i^*] = (N - 1)(c_N + \bar{m} \Delta_N \sigma_i^*), \quad (1.7)$$

so meaningful information about the community structure is revealed from the degrees of nodes.

Despite this clear picture regarding the thresholds for exact recovery and detection in the setting of two communities, several questions remain open. In this paper, we focus on the problem of quantifying exactly how much information about the communities can be recovered by observing the graph \mathbf{G}_N . The mutual information between the assignment vector σ^* and the random graph \mathbf{G}_N is defined by

$$I(\mathbf{G}_N; \sigma^*) = \mathbb{E} \log \frac{\mathbb{P}(\mathbf{G}_N, \sigma^*)}{\mathbb{P}(\mathbf{G}_N)\mathbb{P}(\sigma^*)} = \mathbb{E} \log \frac{\mathbb{P}(\mathbf{G}_N | \sigma^*)}{\mathbb{P}(\mathbf{G}_N)}. \quad (1.8)$$

The asymptotic value of this mutual information has been computed in the dense regime [9, 30, 42] and the sparse disassortative regime [4, 28]. Its determination in the assortative sparse regime has proved more challenging. After we posted a first version of this paper to the arXiv, this problem was resolved in [72] in the case $p = 1/2$, building upon the earlier works [3, 39, 51, 54]. The approach developed there does not generalize well to more complex models such as when more than two communities are present [36]. In contrast, our aim here is to propose a new approach to the analysis of the community detection problem that would be robust to model modifications. We will discuss in much greater detail the relationship between these very recent works and our contribution near the end of the introduction; see also [1, 5, 6] for more on problems with more than two communities.

Henceforth, we focus exclusively on the sparse stochastic block model with

$$c_N = \frac{c}{N} \quad \text{and} \quad \Delta_N = \frac{\Delta}{N} \quad (1.9)$$

for some $c > 0$ and some non-zero $\Delta \in (-c, c)$. The case $\Delta = 0$ is trivial since it corresponds to the case where the graph \mathbf{G}_N and the assignment vector σ^* are independent. The probability (1.6) that an edge is present between node i and node j becomes

$$\mathbb{P}\{G_{ij} = 1 | \sigma^*\} = \frac{c + \Delta \sigma_i^* \sigma_j^*}{N} \quad (1.10)$$

for the family of conditionally independent Bernoulli random variables $\mathbf{G}_N = (G_{ij})$. The expected degree of any node i remains bounded with N ,

$$\mathbb{E} \deg(i) = \frac{N-1}{N} (c + \Delta \bar{m}^2), \quad (1.11)$$

so we are indeed in the sparse regime. The likelihood of the model is given by

$$\mathbb{P}\{\mathbf{G}_N = (G_{ij}) | \sigma^* = \sigma\} = \prod_{i < j} \left(\frac{c + \Delta \sigma_i \sigma_j}{N} \right)^{G_{ij}} \left(1 - \frac{c + \Delta \sigma_i \sigma_j}{N} \right)^{1-G_{ij}} = \frac{\exp(H_N^\circ(\sigma))}{N^{\sum_{i < j} G_{ij}}}, \quad (1.12)$$

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

$$\mathbb{P}\{\sigma^* = \sigma | \mathbf{G}_N = (G_{ij})\} = \frac{\exp(H_N^\circ(\sigma)) P_N^*(\sigma)}{\int_{\Sigma_N} \exp(H_N^\circ(\tau)) dP_N^*(\tau)} \quad (1.13)$$

associated with the Hamiltonian

$$H_N^\circ(\boldsymbol{\sigma}) = \sum_{i < j} \log \left[(c + \Delta \sigma_i \sigma_j)^{G_{ij}} \left(1 - \frac{c + \Delta \sigma_i \sigma_j}{N} \right)^{1-G_{ij}} \right]. \quad (1.14)$$

Moreover, up to an error vanishing with N and a simple additive constant, the normalized mutual information (1.8) coincides with the free energy

$$\bar{F}_N^\circ = \frac{1}{N} \mathbb{E} \log \int_{\Sigma_N} \exp H_N^\circ(\boldsymbol{\sigma}) dP_N^*(\boldsymbol{\sigma}). \quad (1.15)$$

Indeed, (1.12) and Bayes' formula imply that

$$I(\mathbf{G}_N; \boldsymbol{\sigma}^*) = \binom{N}{2} \mathbb{E} \log (c + \Delta \sigma_1^* \sigma_2^*)^{G_{12}} \left(1 - \frac{c + \Delta \sigma_1^* \sigma_2^*}{N} \right)^{1-G_{12}} - N \bar{F}_N^\circ. \quad (1.16)$$

Averaging with respect to the randomness of G_{12} and Taylor expanding the logarithm reveals that

$$\frac{1}{N} I(\mathbf{G}_N; \boldsymbol{\sigma}^*) = \frac{1}{2} \mathbb{E} (c + \Delta \sigma_1^* \sigma_2^*) \log (c + \Delta \sigma_1^* \sigma_2^*) - \frac{c}{2} - \frac{\Delta \bar{m}^2}{2} - \bar{F}_N^\circ + \mathcal{O}(N^{-1}). \quad (1.17)$$

To study the mutual information (1.8), we take the perspective of statistical physics and instead focus on the free energy (1.15). Notice that the mutual information between two independent random variables is zero, so for $\Delta = 0$, the free energy is found by setting the right-hand side of (1.17) equal to zero.

For technical reasons, it will be convenient to modify the free energy (1.15) without changing its limiting value. Conditionally on $\boldsymbol{\sigma}^*$, the modified Hamiltonian will be a sum of independent random variables, and the sum will be over a Poisson-distributed number of terms. The main advantage of this construction is that we can then conveniently vary the continuous parameter encoding the Poisson random variable, and in particular study derivatives with respect to this parameter. To be more precise, we introduce a random variable $\Pi_1 \sim \text{Poi} \left(\frac{N}{2} \right)$ as well as an independent family of i.i.d. random matrices $(G^k)_{k \in \mathbb{N}}$ each having conditionally independent entries $(G_{i,j}^k)_{i,j \leq N}$ taking values in $\{0, 1\}$ with conditional distribution

$$\mathbb{P}\{G_{i,j}^k = 1 | \boldsymbol{\sigma}^*\} = \frac{c + \Delta \sigma_i^* \sigma_j^*}{N}. \quad (1.18)$$

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

$$H_N(\boldsymbol{\sigma}) = \sum_{k \leq \Pi_1} \log \left[(c + \Delta \sigma_{i_k} \sigma_{j_k})^{G_{i_k, j_k}^k} \left(1 - \frac{c + \Delta \sigma_{i_k} \sigma_{j_k}}{N} \right)^{1-G_{i_k, j_k}^k} \right], \quad (1.19)$$

and write

$$\bar{F}_N = \frac{1}{N} \mathbb{E} \log \int_{\Sigma_N} \exp H_N(\boldsymbol{\sigma}) dP_N^*(\boldsymbol{\sigma}) \quad (1.20)$$

for its associated free energy. We show in Appendix A that the difference between the free energies in (1.15) and (1.20) tends to 0 as N tends to infinity. Together with (1.17), this implies that

$$\frac{1}{N} I(\mathbf{G}_N; \boldsymbol{\sigma}^*) = \frac{1}{2} \mathbb{E} (c + \Delta \sigma_1^* \sigma_2^*) \log (c + \Delta \sigma_1^* \sigma_2^*) - \frac{c}{2} - \frac{\Delta \bar{m}^2}{2} - \bar{F}_N + o_N(1). \quad (1.21)$$

The problem of finding the asymptotic value of the mutual information (1.8) has therefore been reduced to the task of determining the limit of the free energy (1.20). The main contributions of this work are the conjecture that

$$\lim_{N \rightarrow \infty} \bar{F}_N = f(1, 0), \quad (1.22)$$

where $f(t, \mu)$ is the solution to an infinite-dimensional Hamilton-Jacobi equation defined in (1.41), and a proof that $f(1, 0)$ provides an upper bound for the limit of the free energy; the matching lower bound will remain open.

To motivate and define the infinite-dimensional Hamilton-Jacobi equation (1.41), we will introduce an “enriched” free energy functional by transforming the free energy (1.20) into a function of a “time” variable $t \geq 0$ and a non-negative measure μ . The “time” variable will be used to vary the intensity of the Poisson point process Π_1 appearing in the Hamiltonian (1.19). We introduced the Hamiltonian H_N to replace H_N° in order to allow for convenient integration-by-parts-like calculations as we study derivatives with respect to this parameter t . The non-negative measure μ will be decomposed as $\mu = s\bar{\mu}$ for $s \geq 0$ and a probability measure $\bar{\mu}$. It will be used to consider a situation in which we also observe the graph of connections of a simpler setting in which each individual i can form connections with its own set of neighbour candidates. To be more specific, each individual i will have an independent number $\text{Poi}(sN)$ of neighbour candidates indexed by the pairs (i, k) for $k \leq \text{Poi}(sN)$. Each candidate neighbour (i, k) will be independently assigned a random “type” $x_{i,k}$ sampled from the distribution $\bar{\mu}$, and an edge will be present between individual i and its candidate neighbour (i, k) with probability $N^{-1}(c + \Delta\sigma_i^* x_{i,k})$. In the inference problem, the “types” $x_{i,k}$ are revealed to the statistician. The lack of interactions between individuals makes this piece of information much simpler to understand than the original community detection problem we aim to make progress upon. In total, this allows us to define an “enriched” free energy, function of t and μ , and the quantity in (1.20) can then be recovered by evaluating this enriched free energy at $t = 1$ and $\mu = 0$. We next aim to study whether variations in the t variable can be suitably compensated by variations in the μ variable, leaving the free energy roughly constant. More precisely, we hope to discover that the derivative in t of this functional can be expressed, up to a small error, as a function of its derivative in μ . On a heuristic level, one can see that this indeed seems to be possible, as will be clarified below. Combining ideas from the theory of viscosity solutions with the multi-overlap concentration result in [14], we will be able to prove one inequality between the limit free energy and the solution to the partial differential equation that arises. Although we expect the converse bound to also be valid, significant technical challenges prevent us from proving it at the moment. The difficulty is that the control we have on the “small error” appearing in the equation at finite N is relatively weak. In particular, we cannot show, and do not expect, that it becomes small as N tends to infinity for each individual choice of t and μ . On the other hand, controlling the error after we perform a small averaging over t and μ is possible, but does not suffice for the identification of the limit.

Let us now define the enriched free energy precisely. Denote by $\text{Pr}[-1, 1]$ the set of probability measures on $[-1, 1]$, and given $\mu \in \text{Pr}[-1, 1]$, consider a sequence $x = (x_{i,k})$ of i.i.d. random variables with law μ . For each $s > 0$ and $i \geq 1$, let $\Pi_{i,s} \sim \text{Poi}(sN)$ be independent over $i \geq 1$, and

introduce the Hamiltonian on Σ_N ,

$$\tilde{H}_N^{s,\mu}(\sigma) = \sum_{i \leq N} \sum_{k \leq \Pi_{i,s}} \log \left[(c + \Delta \sigma_i x_{i,k}) \tilde{G}_{i,k}^x \left(1 - \frac{c + \Delta \sigma_i x_{i,k}}{N} \right)^{1 - \tilde{G}_{i,k}^x} \right], \quad (1.23)$$

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

$$\mathbb{P}\{\tilde{G}_{i,k}^x = 1 | \sigma^*, x\} = \frac{c + \Delta \sigma_i^* x_{i,k}}{N}. \quad (1.24)$$

As alluded to above, this is the Hamiltonian associated with the task of inferring the signal σ^* from the data

$$\tilde{\mathcal{D}}_N^{s,\mu} = (\Pi_{i,s}, (x_{i,k})_{k \leq \Pi_{i,s}}, (\tilde{G}_{i,k}^x)_{k \leq \Pi_{i,s}})_{i \leq N}, \quad (1.25)$$

in the sense that the Gibbs measure associated with $\tilde{H}_N^{s,\mu}$ is the conditional law of σ^* given the data $\tilde{\mathcal{D}}_N^{s,\mu}$ (as in the identity (1.13) for the Hamiltonian H_N° and the data \mathbf{G}_N). For each $t \geq 0$, let $\Pi_t \sim \text{Poi}t\binom{N}{2}$, and consider a time-dependent version of the Hamiltonian (1.19) defined on Σ_N by

$$H_N^t(\sigma) = \sum_{k \leq \Pi_t} \log \left[(c + \Delta \sigma_{i_k} \sigma_{j_k}) G_{i_k, j_k}^k \left(1 - \frac{c + \Delta \sigma_{i_k} \sigma_{j_k}}{N} \right)^{1 - G_{i_k, j_k}^k} \right]. \quad (1.26)$$

Notice that this is the Hamiltonian associated with the task of inferring the signal σ^* from the data

$$\mathcal{D}_N^t = (\Pi_t, (i_k, j_k)_{k \leq \Pi_t}, (G_{i_k, j_k}^k)_{k \leq \Pi_t}). \quad (1.27)$$

We now introduce an enriched Hamiltonian on Σ_N ,

$$\tilde{H}_N^{t,s,\mu}(\sigma) = H_N^t(\sigma) + \tilde{H}_N^{s,\mu}(\sigma), \quad (1.28)$$

and denote by

$$\tilde{F}_N(t, s, \mu) = \frac{1}{N} \mathbb{E} \log \int_{\Sigma_N} \exp \tilde{H}_N^{t,s,\mu}(\sigma) dP_N^*(\sigma) \quad (1.29)$$

its associated free energy. Observe that $\tilde{F}_N(1, 0, \mu) = \bar{F}_N$ and that (1.28) is the Hamiltonian associated with inferring the signal σ^* from the data

$$\tilde{\mathcal{D}}_N^{t,s,\mu} = (\mathcal{D}_N^t, \tilde{\mathcal{D}}_N^{s,\mu}), \quad (1.30)$$

where the randomness in these two data sets is taken to be independent conditionally on σ^* . To obtain a Hamilton-Jacobi equation, it will be convenient to reinterpret the enriched free energy (1.29) as a function of the time-parameter $t > 0$ and a finite measure μ ; the parameter s will become the total mass of this finite measure. We denote by \mathcal{M}_s the space of signed measures on $[-1, 1]$,

$$\mathcal{M}_s = \{\mu \mid \mu \text{ is a signed measure on } [-1, 1]\}, \quad (1.31)$$

and by \mathcal{M}_+ the cone of non-negative measures on this interval,

$$\mathcal{M}_+ = \{\mu \in \mathcal{M}_s \mid \mu \text{ is a non-negative measure}\}. \quad (1.32)$$

We follow the convention that a signed measure can only take finite values, and in particular, every $\mu \in \mathcal{M}_+$ must have finite total mass. This implies that every non-zero measure $\mu \in \mathcal{M}_+$ induces a probability measure,

$$\bar{\mu} = \frac{\mu}{\mu[-1, 1]} \in \text{Pr}[-1, 1]. \quad (1.33)$$

Given a measure $\mu \in \mathcal{M}_+$, we define the Hamiltonian $H_N^{t, \mu}$ on Σ_N by

$$H_N^{t, \mu}(\sigma) = \tilde{H}_N^{t, \mu[-1, 1], \bar{\mu}}(\sigma), \quad (1.34)$$

where $\tilde{H}_N^{0, 0} = 0$ for the zero measure by continuity. The free energy associated with this Hamiltonian is given by

$$\bar{F}_N(t, \mu) = \tilde{F}_N(t, \mu[-1, 1], \bar{\mu}), \quad (1.35)$$

and once again $\bar{F}_N = \bar{F}_N(1, 0)$, where 0 denotes the zero measure. The free energy in (1.35) will be termed the enriched free energy, and in Section 2 we will show that, up to a ‘‘small error’’, it satisfies an infinite-dimensional Hamilton-Jacobi equation which we now describe.

Introduce the function $g : [-1, 1] \rightarrow \mathbb{R}$ defined by

$$g(z) = (c + \Delta z)(\log(c + \Delta z) - 1) = (c + \Delta z)\log(c) + c \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} z^n - c \quad (1.36)$$

as well as the cone of functions

$$\mathcal{C}_\infty = \left\{ G_\mu : [-1, 1] \rightarrow \mathbb{R} \mid G_\mu(x) = \int_{-1}^1 g(xy) d\mu(y) \text{ for some } \mu \in \mathcal{M}_+ \right\} \quad (1.37)$$

and the non-linearity $C_\infty : \mathcal{C}_\infty \rightarrow \mathbb{R}$ given by

$$C_\infty(G_\mu) = \frac{1}{2} \int_{-1}^1 G_\mu(x) d\mu(x). \quad (1.38)$$

This non-linearity is well-defined by the Fubini-Tonelli theorem (see equations (1.6)-(1.7) in [31]). Given a function $f : [0, \infty) \times \mathcal{M}_+ \rightarrow \mathbb{R}$ and measures $\mu, \nu \in \mathcal{M}_+$, 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 ,

$$D_\mu f(t, \mu; \nu) = \lim_{\varepsilon \rightarrow 0} \frac{f(t, \mu + \varepsilon \nu) - f(t, \mu)}{\varepsilon}. \quad (1.39)$$

We will say that the Gateaux derivative of f admits a density at the measure $\mu \in \mathcal{M}_+$ if there exists a bounded measurable function $x \mapsto D_\mu f(t, \mu, x)$ defined on the interval $[-1, 1]$ with

$$D_\mu f(t, \mu; \nu) = \int_{-1}^1 D_\mu f(t, \mu, x) d\nu(x) \quad (1.40)$$

for every measure $\nu \in \mathcal{M}_+$. We will often abuse notation and identify the density $D_\mu f(t, \mu, \cdot)$ with the Gateaux derivative $D_\mu f(t, \mu)$. In Section 2 we will show that, up to an error vanishing with N , the Gateaux derivative of the enriched free energy is indeed of the form G_{μ^*} for some $\mu^* \in \mathcal{M}_+$. In fact, the measure μ^* is the law of the Gibbs average of a uniformly sampled spin

coordinate for the Gibbs measure associated with the Hamiltonian (1.28). We will also argue that, if this Gibbs measure satisfies suitable overlap concentration properties, then the time derivative of the free energy is essentially given by $\int_{-1}^1 g(xy) d\mu^*(x) d\mu^*(y)$. In short, assuming the validity of these overlap concentration properties, we are led to believe that the large- N limit of the free energy (1.35) should satisfy the infinite-dimensional Hamilton-Jacobi equation

$$\begin{cases} \partial_t f(t, \mu) = C_\infty(D_\mu f(t, \mu)) & \text{on } \mathbb{R}_{>0} \times \mathcal{M}_+, \\ f(0, \mu) = \psi(\mu) & \text{on } \mathcal{M}_+, \end{cases} \quad (1.41)$$

where the initial condition $\psi : \mathcal{M}_+ \rightarrow \mathbb{R}$ is the limit of $F_N(0, \cdot)$ and can be readily computed, see Lemma 3.1. The well-posedness of this equation is established in [31], and leads to the conjecture that the enriched free energy converges to the solution to this equation. Remembering (1.21), this translates into a conjecture for the asymptotic mutual information.

Conjecture 1. *If f denotes the unique viscosity solution to the infinite-dimensional Hamilton-Jacobi equation (1.41), then the limit of the free energy (1.20) is given by*

$$\lim_{N \rightarrow \infty} \bar{F}_N = f(1, 0). \quad (1.42)$$

In particular, the asymptotic value of the mutual information (1.8) is

$$\lim_{N \rightarrow \infty} \frac{1}{N} I(\mathbf{G}_N; \boldsymbol{\sigma}^*) = \frac{1}{2} \mathbb{E} (c + \Delta \sigma_1^* \sigma_2^*) \log (c + \Delta \sigma_1^* \sigma_2^*) - \frac{c}{2} - \frac{\Delta \bar{m}^2}{2} - f(1, 0). \quad (1.43)$$

The main result of this paper is a proof of the upper bound in Conjecture 1.

Theorem 1.1. *Denote by f the unique viscosity solution to the infinite-dimensional Hamilton-Jacobi equation (1.41). For every $t \geq 0$ and $\mu \in \mathcal{M}_+$, the limit of the enriched free energy (1.35) satisfies the upper bound $\limsup_{N \rightarrow \infty} \bar{F}_N(t, \mu) \leq f(t, \mu)$. In particular, the free energy (1.20) satisfies the upper bound*

$$\limsup_{N \rightarrow \infty} \bar{F}_N \leq f(1, 0). \quad (1.44)$$

Although the matching lower bound still remains open, we give some support in favor of Conjecture 1 by proving that, in the disassortative regime, it matches the variational formula for the asymptotic free energy obtained in [28]. (The condition $p = 1/2$ was also assumed in [28].) We state this formula using the notation in [13], where a more direct proof is obtained using an interpolation argument and a cavity computation. Denote by

$$\mathcal{M}_p = \left\{ \mu \in \text{Pr}[-1, 1] \mid \int_{-1}^1 x d\mu = \bar{m} \right\} \quad (1.45)$$

the set of probability measures with mean $\bar{m} = 2p - 1$, and introduce the functional $\mathcal{P} : \mathcal{M}_p \rightarrow \mathbb{R}$ defined by

$$\mathcal{P}(\mu) = \psi(\mu) + \frac{c}{2} + \frac{\Delta \bar{m}^2}{2} - \frac{1}{2} \mathbb{E}(c + \Delta x_1 x_2) \log(c + \Delta x_1 x_2), \quad (1.46)$$

where x_1 and x_2 are independent samples from the probability measure μ .

Theorem 1.2. *In the disassortative sparse stochastic block model with $\Delta \leq 0$, the limit of the free energy (1.20) is given by*

$$\lim_{N \rightarrow \infty} \bar{F}_N = \sup_{\mu \in \mathcal{M}_p} \mathcal{P}(\mu) = f(1, 0). \quad (1.47)$$

Our results generalize immediately to the case in which the measure P^* is arbitrary with compact support, with the understanding that the link probabilities are given by (1.10). We also believe that they generalize without much change to settings with more than two communities, although we have not checked every technical detail.

Before closing this introduction, we discuss alternatives to the conjecture and the approach proposed in this paper. It will facilitate this discussion to point out that the proof of Theorem 1.2 also yields that when $\Delta \leq 0$, we can identify the limit of $\bar{F}_N(t, \mu)$ for every $t \geq 0$ and $\mu \in \mathcal{M}_+$ as

$$\lim_{N \rightarrow \infty} \bar{F}_N(t, \mu) = f(t, \mu) = \sup_{v \in \text{Pr}[-1, 1]} \left(\psi(\mu + tv) - \frac{t}{2} \int_{-1}^1 G_v(y) dv(y) \right), \quad (1.48)$$

where we recall that G_v is defined in (1.36)-(1.37). The identity (1.48) also remains valid if we take the supremum over all $v \in \mathcal{M}_+$, and it is at times convenient to operate over variables that can vary freely inside a cone.

Concerning the limit of the free energy, one may hope that the formulas given in (1.47) and (1.48) in the case $\Delta \leq 0$ remain valid in general. It seems difficult to identify the exact range of validity of these formulas. We would be surprised if they hold for arbitrary measures P^* , but we could not quickly find a counter-example. We are however confident that these formulas will not generalize to settings with more than two communities.

To see this, we rely on the fact that the problem of identifying the limit of the free energy becomes simpler in the dense regime. Indeed, if the average degree of a node diverges as N tends to infinity, then central-limit-theorem effects take place, and one can equivalently study a fully-connected model with Gaussian noise [30, 42]. Such models have been studied extensively [10, 11, 12, 21, 22, 23, 24, 38, 42, 43, 45, 46, 48, 49, 55, 56, 66, 67]. In this setting, a formula analogous to (1.47)-(1.48) is known to be valid as long as the relevant non-linearity is convex; but in general, one needs to modify this formula into a “sup-inf” formulation. Possibly the simplest setting in which this happens is for the problem in which we observe a rank-one matrix of the form XY^\top plus noise, where X and Y are two vectors with i.i.d. coordinates. In this setting, the non-linearity replacing C_∞ in (1.41) is the mapping $(x, y) \mapsto xy$, which is non-convex. The functional to be optimized over as in (1.48) would look like $\psi(x_0 + tx, y_0 + ty) - \frac{t}{2}xy$. Finding counter-examples to the formula is made relatively easy by considering candidates with, say, $x = 0$; in this case, the counter-term xy vanishes, so we can freely choose y as large as desired to maximize the ψ functional and obtain a contradiction. A similar phenomenon also occurs in the context of spin glasses, and a more precise discussion of this point can be found in Subsection 6.2 of [57].

Coming back to the sparse setting investigated in this paper, we can leverage this observation to demonstrate that the formulas (1.47) or (1.48) will also be invalid in general. To give a concrete example, consider the following scenario, which can be thought of as a problem with four communities, or as a bipartite version of the two-community problem. We first color the N nodes in red or blue, say with groups of sizes about $N/2$. We think of this coloring as fixed, e.g. the red

nodes are the first $\lfloor N/2 \rfloor$ indices in $\{1, \dots, N\}$, and it is perfectly known to the statistician. Next, we attribute ± 1 labels to each node independently, possibly with different biases according to the color of the node. Finally, we draw links between nodes i and j according to the formula in (1.10), with the additional constraint that only links between nodes of different colors are allowed. The task is to study the asymptotic behavior of the mutual information between the ± 1 labels and the observed graph. This problem is constructed in such a way that, in the limit of diverging average degree, it reduces to the problem of observing a noisy version of XY^\top , as discussed in the previous paragraph — the vectors X and Y contain the ± 1 labels of the red and blue nodes respectively. Using the results of [30, 42] to justify the large-degree approximation, or possibly even directly, we are confident that we can then produce counter-examples to the formulas (1.47) and (1.48).

For fully-connected models with possibly non-convex non-linearities such as the XY^\top example, the limit of the free energy was identified in the form of a “sup-inf” formula; see [21] for the most general results. Translating this result into our present context would suggest that the limit free energy might be given by

$$\sup_{\rho \in \mathcal{M}_+} \inf_{\nu \in \mathcal{M}_+} \left(\psi(\nu) + \int_{-1}^1 G_\rho(y) d(\mu - \nu)(y) + t \int_{-1}^1 G_\rho(y) d\rho(y) \right). \quad (1.49)$$

The key ingredient for showing the validity of the corresponding formula in the dense regime is that the enriched free energy is a convex function of its parameters in this case. In our setting, the question would translate into whether the mapping $(t, \mu) \mapsto \bar{F}_N(t, \mu)$ is convex. However, it was shown in [41] that this mapping is in fact *not* convex in the sparse regime, even in the limit of large N . This non-convexity property not only breaks down the proof strategy of [21]; in fact, we can leverage it to assert that the quantity (1.49) can therefore *not* be the limit of the free energy in this case. Indeed, the expression in (1.49) is a supremum over ρ of affine functions of (t, μ) ; it therefore follows that the whole expression is convex in (t, μ) . By [41], it is therefore not possible that the expression in (1.49) be the limit of the free energy.

To sum up, if we aim for a formula that is robust to model changes, then both (1.48) and (1.49) can be ruled out. We do not know of alternative candidate variational formulas for the limit of the free energy. This situation seems analogous to that encountered in the context of spin glasses with possibly non-convex interactions, as discussed in Section 6 of [57].

We now turn to a discussion of the recent works [3, 36, 39, 51, 54, 72]. Noticing that the graph \mathbf{G}_N locally looks like a tree, these works aim to leverage a connection between community detection and a process of broadcasting on trees. We briefly describe the latter problem on a regular tree for convenience. We first attribute a random ± 1 random variable σ^* to the root node. Then, recursively and independently along each edge, we “broadcast” it to each child node, by flipping the sign of the spin with some fixed probability $\delta \in (0, 1)$. One basic question is to determine the mutual information between the spin σ^* attributed to the root node and the spins on all the nodes at a given depth, in the limit of large depth. A fruitful variant of this question consists in adding a “survey” of all nodes, by randomly revealing the spins attached to each node independently with some fixed probability ε . If, in the limit of large depth, the knowledge of the spins on all the leaf vertices does not bring meaningful additional information on σ^* on top of surveying compared with surveying alone, then one can directly relate the mutual information between σ^* and the survey to the mutual information in the community detection problem; in this case, one may speak of “boundary

irrelevance”. To decide whether boundary irrelevance holds, one can study the evolution of the log-likelihood ratio between the two hypotheses $\sigma^* = \pm 1$ upon revealing the boundary information at a given depth. One can indeed calculate the law of this quantity recursively as the depth varies, by iterating a fixed map called the “BP operator”. In order to establish the property of boundary irrelevance, it then essentially suffices to show that this BP operator admits a unique non-trivial fixed point. Building upon earlier works, it was recently established in [72] that this uniqueness property holds for the setting corresponding to the detection of two balanced communities, $p = 1/2$. As a byproduct, this yields a full identification of the limit of the mutual information (1.8) in this case. The uniqueness of a non-trivial fixed point to the BP operator has subsequently been shown to be false in general for models with more than two communities [36].

We now point out some connections between the present paper and this series of works, and discuss how our approach might ultimately be able to circumvent the difficulties associated with the possible existence of multiple fixed points to the BP operator. To start with, recall that the function ψ is the limit of $\bar{F}_N(0, \cdot)$, which itself corresponds to a simple inference problem in which there is no interaction between the nodes $\{1, \dots, N\}$. We can therefore identify this object explicitly, see Lemma 3.1. From Remark 3.2, we can also identify a mapping $\Gamma : \mathcal{M}_+ \rightarrow \text{Pr}[-1, 1]$ such that for every $\mu \in \mathcal{M}_+$, we have $D_\mu \psi(\mu, \cdot) = G_{\Gamma(\mu)}$. This mapping is closely related to the BP operator discussed above, and is described as follows. Let σ^* be sampled according to P^* , and conditionally on σ^* , let $\Pi(\mu)$ denote a Poisson point process with intensity measure $(c + \Delta\sigma^*x) d\mu(x)$. Then the probability measure $\Gamma(\mu)$ is defined to be the law of the random variable

$$\frac{\int_{\Sigma_1} \sigma \exp(-\Delta\sigma \int_{-1}^1 x d\mu) \prod_{x \in \Pi(\mu)} (c + \Delta\sigma x) dP^*(\sigma)}{\int_{\Sigma_1} \exp(-\Delta\sigma \int_{-1}^1 x d\mu) \prod_{x \in \Pi(\mu)} (c + \Delta\sigma x) dP^*(\sigma)}. \quad (1.50)$$

Notice next that the condition for the measure ν to be a critical point in the variational problem on the right side of (1.48) can be written as

$$G_\nu = D_\mu \psi(\mu + t\nu, \cdot). \quad (1.51)$$

At least when $\Delta < 0$, the mapping $\nu \mapsto G_\nu$ is also injective, so the relation (1.51) can be equivalently written as

$$\nu = \Gamma(\mu + t\nu). \quad (1.52)$$

Restricting to the case of $(t, \mu) = (1, 0)$, this boils down to finding fixed points of the mapping Γ . That there is a connection between the variational formula in Theorem 1.2 and some BP fixed point equation has already been observed in [28, 29] and elsewhere. The less classical question is to relate this to the Hamilton-Jacobi equation (1.41) for arbitrary Δ . In finite dimensions, Hamilton-Jacobi equations can be solved for short times using the method of characteristics. Moreover, the slope of the characteristic line is computed by evaluating the gradient of the non-linearity at the gradient of the initial condition. In our context, the characteristic line emanating from a measure $\nu \in \mathcal{M}_+$ is the trajectory

$$t' \mapsto (t', \nu - t'\Gamma(\nu)), \quad (1.53)$$

for t' varying in $\mathbb{R}_{\geq 0}$. As long as characteristic lines emanating from different choices of ν do not intersect each other, we can then calculate the value of the solution along each characteristic line using the equation and the fact that the gradient of the solution remains constant along each

line [33]. The condition (1.52) turns out to be equivalent to asking that the characteristic line emanating from $\mu + t\nu$ passes through the point (t, μ) , since the latter condition can be written as $\mu = \mu + t\nu - t\Gamma(\mu + t\nu)$. In other words, for each fixed (t, μ) , there is a simple one-to-one correspondence between the fixed points to (1.52) and the characteristic lines that pass through (t, μ) . The formula for prescribing the value of the solution along a characteristic line starting from $\mu + t\nu$ is then as in the supremum in (1.48). As long as t is sufficiently small that the equation (1.52) has a unique solution for each μ , this gives us a clear procedure for computing the solution to (1.40). Once characteristic lines start to intersect each other, the viscosity solution to (1.41) aggregates these conflicting trajectories in a physically reasonable way, and our conjecture is that the free energy \bar{F}_N is tracking this in the limit of large N .

Another alternative to the conjecture proposed here would be that the limit of the free energy is the maximal value one gets by plugging every possible solution of the fixed-point equation (1.52) into the functional inside the supremum in (1.48). But in view of the discussion in the previous paragraph, counter-examples to the variational formula in (1.48) seem to produce counter-examples to this possibility as well.

To conclude this introduction, we give a brief outline of the paper. In Section 2 we show that, up to an error vanishing with N , the enriched free energy (1.35) satisfies the infinite-dimensional Hamilton-Jacobi equation (1.41), provided that all multi-overlaps concentrate. The derivative computations that lead to the Hamilton-Jacobi equation are similar in spirit to those in Lemma 6 of [61], with some new ideas required to compute the Gateaux derivative. Section 3 is devoted to establishing the well-posedness of the infinite-dimensional Hamilton-Jacobi equation (1.41) using the results in [31], which in turn follows ideas from [25, 26, 58, 57]. In Section 4, a finitary version of the multi-overlap concentration result in [14] is combined with the strategy introduced in [58, 57] to prove Theorem 1.1. The final section is devoted to the proof of Theorem 1.2. Using the Hopf-Lax formula established in [31], the variational expression in (1.47) is shown to coincide with the right side of (1.47), and we can thus appeal to Theorem 1.1 to obtain an upper bound for the limit free energy. The matching lower bound is obtained through an interpolation argument taken from [13]. So as to not disrupt the flow of the paper, a number of technical arguments have been postponed to the appendices. In Appendix A, it is shown that the free energy functionals (1.15) and (1.20) are asymptotically equivalent. The proof relies on the binomial-Poisson approximation. Appendix B is devoted to proving that a perturbed version of the enriched free energy (1.35) is self-averaging, in the sense that the unaveraged free energy concentrates around its average value. This concentration result plays an important part in the proof of Theorem 1.1 and relies upon the generalized Efron-Stein inequality [17]. In Appendix C, a finitary version of the multi-overlap concentration result in [14] is established. In addition to being finitary, the most notable difference between our multi-overlap result and that in [14] is that we show multi-overlap concentration for any perturbation parameter satisfying a condition that may be verified in practice, as opposed to obtaining multi-overlap concentration on average over the set of perturbation parameters. This additional control on the choice of parameters is essential in the proof of Theorem 1.1.

Acknowledgements. *We would like to warmly thank Dmitry Panchenko and Jean Barbier for sharing their notes [13] on the free energy in the disassortative sparse stochastic block model with us, which provided us with a very useful starting point and helped us with many of the computations in Section 2.*

2 The Hamilton-Jacobi equation

In this section, we compute the derivative of the enriched free energy (1.35) with respect to $t \geq 0$ and $\mu \in \mathcal{M}_+$. This will allow us to see that, up to an error vanishing with N , the enriched free energy heuristically satisfies (1.41). It will be convenient to write $\langle \cdot \rangle$ for the average with respect to the Gibbs measure associated with the Hamiltonian (1.34). This means that for any bounded and measurable function $f = f(\sigma^1, \dots, \sigma^n)$ of finitely many replicas,

$$\langle f(\sigma^1, \dots, \sigma^n) \rangle = \langle f \rangle = \frac{\int_{\Sigma_N^n} f(\sigma^1, \dots, \sigma^n) \prod_{\ell \leq n} \exp H_N^{t, \mu}(\sigma^\ell) dP_N^*(\sigma^\ell)}{\left(\int_{\Sigma_N} \exp H_N^{t, \mu}(\sigma) dP_N^*(\sigma) \right)^n}. \quad (2.1)$$

In this notation, the replicas $\sigma^1, \dots, \sigma^n$ represent i.i.d. samples under the random measure $\langle \cdot \rangle$. By construction, we have that

$$\langle f(\sigma^1) \rangle = \mathbb{E} [f(\sigma^*) | \mathcal{D}_N^{t, \mu}], \quad (2.2)$$

where $\mathcal{D}_N^{t, \mu} = \tilde{\mathcal{D}}_N^{t, \mu[-1, 1], \bar{\mu}}$ is the data defined in (1.30).

Our computations will be considerably simplified by the Nishimori identity. This identity will allow 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. This identity states that, for every bounded and measurable function $f = f(\sigma^1, \dots, \sigma^n, \mathcal{D}_N^{t, \mu})$ of finitely many replicas and the data,

$$\mathbb{E} \langle f(\sigma^1, \sigma^2, \dots, \sigma^n, \mathcal{D}_N^{t, \mu}) \rangle = \mathbb{E} \langle f(\sigma^*, \sigma^2, \dots, \sigma^n, \mathcal{D}_N^{t, \mu}) \rangle. \quad (2.3)$$

This can be first verified for functions of product form using (2.2), and then extended to all bounded and measurable functions by a monotone class argument.

We now turn our attention to the computation of the time derivative of the enriched free energy (1.35). We fix a finite measure $\mu \in \mathcal{M}_+$ and proceed as in Lemma 6 of [61]. For each parameter $\lambda > 0$ and every integer $m \geq 0$, we denote by

$$\pi(\lambda, m) = \frac{\lambda^m}{m!} \exp(-\lambda) \quad (2.4)$$

the mass attributed to the atom m by a $\text{Poi}(\lambda)$ distribution. It will be convenient to set the convention that $\pi(\lambda, -1) = 0$. We write

$$H_{N, m}(\sigma) = \sum_{k \leq m} \log \left[(c + \Delta \sigma_{i_k} \sigma_{j_k})^{G_{i_k, j_k}^k} \left(1 - \frac{c + \Delta \sigma_{i_k} \sigma_{j_k}}{N} \right)^{1 - G_{i_k, j_k}^k} \right] \quad (2.5)$$

for the Hamiltonian (1.19) conditional on there being m terms in the sum, and introduce the partition function

$$Z_{N, m} = \int_{\Sigma_N} \exp \left(H_{N, m}(\sigma) + \tilde{H}_N^{\mu[-1, 1], \bar{\mu}}(\sigma) \right) dP_N^*(\sigma). \quad (2.6)$$

In this notation, the enriched free energy (1.35) may be expressed as

$$\bar{F}_N(t, \mu) = \frac{1}{N} \sum_{m \geq 0} \pi \left(t \binom{N}{2}, m \right) \mathbb{E} \log Z_{N, m}. \quad (2.7)$$

To take the time derivative of this expression, we will rely upon the simple fact that

$$\partial_\lambda \pi(\lambda, m) = \pi(\lambda, m-1) - \pi(\lambda, m). \quad (2.8)$$

Lemma 2.1. *For any $t > 0$ and $\mu \in \mathcal{M}_+$,*

$$\partial_t \bar{F}_N(t, \mu) = \frac{1}{2} \mathbb{E} (c + \Delta \langle \sigma_i \sigma_j \rangle) \log (c + \Delta \langle \sigma_i \sigma_j \rangle) - \frac{\Delta \bar{m}^2}{2} - \frac{c}{2} + \mathcal{O}(N^{-1}), \quad (2.9)$$

where the indices $i, j \in \{1, \dots, N\}$ are uniformly sampled independently of all other sources of randomness.

Proof. To simplify notation, let $\lambda(t) = t \binom{N}{2}$. Leveraging (2.8) to differentiate the right-hand side of (2.7) yields

$$\begin{aligned} \partial_t \bar{F}_N(t, \mu) &= \frac{1}{N} \binom{N}{2} \sum_{m \geq 0} (\pi(\lambda(t), m-1) - \pi(\lambda(t), m)) \mathbb{E} \log Z_{N,m} \\ &= \frac{1}{N} \binom{N}{2} \sum_{m \geq 0} \pi(\lambda(t), m) \mathbb{E} \log \frac{Z_{N,m+1}}{Z_{N,m}}. \end{aligned} \quad (2.10)$$

Denote by $i, j \in \{1, \dots, N\}$ uniformly sampled indices, and write $G_{i,j}$ for a random variable with conditional distribution (1.18). These random variables are taken to be independent of all other sources of randomness and of each other. Since

$$Z_{N,m+1} \stackrel{d}{=} \int_{\Sigma_N} (c + \Delta \sigma_i \sigma_j)^{G_{i,j}} \left(1 - \frac{c + \Delta \sigma_i \sigma_j}{N}\right)^{1-G_{i,j}} \exp\left(H_{N,m}(\sigma) + H_N^{\mu[-1,1], \bar{\mu}}(\sigma)\right) dP_N^*(\sigma),$$

it follows by (2.10) and the definition of the Gibbs average in (2.1) that

$$\partial_t \bar{F}_N(t, \mu) = \frac{1}{N} \binom{N}{2} \mathbb{E} \log \left\langle (c + \Delta \sigma_i \sigma_j)^{G_{i,j}} \left(1 - \frac{c + \Delta \sigma_i \sigma_j}{N}\right)^{1-G_{i,j}} \right\rangle.$$

Remembering the explicit form of the conditional distribution (1.18), and averaging with respect to the randomness of $G_{i,j}$ reveals that

$$\begin{aligned} \partial_t \bar{F}_N(t, s, \mu) &= \frac{1}{2} \mathbb{E} (c + \Delta \sigma_i^* \sigma_j^*) \log \langle c + \Delta \sigma_i \sigma_j \rangle \\ &\quad + \frac{N}{2} \mathbb{E} \left(1 - \frac{c + \Delta \sigma_i^* \sigma_j^*}{N}\right) \log \left\langle 1 - \frac{c + \Delta \sigma_i \sigma_j}{N} \right\rangle + \mathcal{O}(N^{-1}), \end{aligned}$$

Taylor expanding the logarithm and keeping only first order terms reduces this to

$$\begin{aligned} \partial_t \bar{F}_N(t, \mu) &= \frac{1}{2} \mathbb{E} (c + \Delta \sigma_i^* \sigma_j^*) \log \langle c + \Delta \sigma_i \sigma_j \rangle - \frac{\Delta}{2} \mathbb{E} \langle \sigma_i \sigma_j \rangle - \frac{c}{2} + \mathcal{O}(N^{-1}) \\ &= \frac{1}{2} \mathbb{E} (c + \Delta \sigma_i^* \sigma_j^*) \log \langle c + \Delta \sigma_i \sigma_j \rangle - \frac{\Delta}{2} \mathbb{E} \sigma_i^* \sigma_j^* - \frac{c}{2} + \mathcal{O}(N^{-1}) \\ &= \frac{1}{2} \mathbb{E} (c + \Delta \sigma_i^* \sigma_j^*) \log (c + \Delta \langle \sigma_i \sigma_j \rangle) - \frac{\Delta \bar{m}^2}{2} - \frac{c}{2} + \mathcal{O}(N^{-1}), \end{aligned}$$

where the second equality uses the Nishimori identity (2.3) and the third equality uses the fact that i and j are distinct with overwhelming probability. Noticing that the Gibbs average $\langle \sigma_i \sigma_j \rangle$ is a measurable function of the data by (2.2), and applying the Nishimori identity (2.3) completes the proof. \blacksquare

To compare (2.9) with the Gateaux derivative of the enriched free energy which we will compute below, it will be convenient to Taylor expand the logarithm. This will make the dependence of the time-derivative of the enriched free energy on the multi-overlaps

$$R_{\ell_1, \dots, \ell_n} = \frac{1}{N} \sum_{i \leq N} \sigma_i^{\ell_1} \dots \sigma_i^{\ell_n} \quad (2.11)$$

associated with the enriched Hamiltonian (1.34) explicit. Here (σ^ℓ) denotes a sequence of i.i.d. replicas sampled from the Gibbs measure (2.1). To simplify notation, we will write $R_{[n]} = R_{1, \dots, n}$.

Corollary 2.2. *For any $t > 0$ and $\mu \in \mathcal{M}_+$,*

$$\partial_t \bar{F}_N(t, \mu) = \frac{1}{2} (c + \Delta \bar{m}^2) \log(c) + \frac{c}{2} \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} \mathbb{E} \langle R_{[n]}^2 \rangle - \frac{c}{2} + \mathcal{O}(N^{-1}). \quad (2.12)$$

Proof. A Taylor expansion of the logarithm shows that

$$\log(c + \Delta \langle \sigma_i \sigma_j \rangle) = \log(c) + \log\left(1 + \frac{\Delta}{c} \langle \sigma_i \sigma_j \rangle\right) = \log(c) - \sum_{n \geq 1} \frac{(-\Delta/c)^n}{n} \langle \sigma_i \sigma_j \rangle^n.$$

Together with the Nishimori identity (2.3) this implies that

$$\begin{aligned} \mathbb{E} (c + \Delta \langle \sigma_i \sigma_j \rangle) \log(c + \Delta \langle \sigma_i \sigma_j \rangle) &= (c + \Delta \bar{m}^2) \log(c) \\ &\quad - \sum_{n \geq 1} \frac{(-\Delta/c)^n}{n} \mathbb{E} (c + \Delta \langle \sigma_i \sigma_j \rangle) \langle \sigma_i \sigma_j \rangle^n. \end{aligned} \quad (2.13)$$

Averaging with respect to the randomness of the uniformly sampled indices $i, j \in \{1, \dots, N\}$ reveals that

$$\mathbb{E} (c + \Delta \langle \sigma_i \sigma_j \rangle) \langle \sigma_i \sigma_j \rangle^n = c \mathbb{E} \langle R_{[n]}^2 \rangle + \Delta \mathbb{E} \langle R_{[n+1]}^2 \rangle.$$

Remembering that $|\Delta| < c$ and noticing that $\mathbb{E} \langle R_{[1]}^2 \rangle = \bar{m}^2 + \mathcal{O}(N^{-1})$ by the Nishimori identity, it follows that

$$\begin{aligned} \sum_{n \geq 1} \frac{(-\Delta/c)^n}{n} \mathbb{E} (c + \Delta \langle \sigma_i \sigma_j \rangle) \langle \sigma_i \sigma_j \rangle^n &= -\Delta \mathbb{E} \langle R_{[1]}^2 \rangle + c \sum_{n \geq 2} \left(\frac{(-\Delta/c)^n}{n} - \frac{(-\Delta/c)^n}{n-1} \right) \mathbb{E} \langle R_{[n]}^2 \rangle + \mathcal{O}(N^{-1}) \\ &= -\Delta \bar{m}^2 - c \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} \mathbb{E} \langle R_{[n]}^2 \rangle + \mathcal{O}(N^{-1}). \end{aligned}$$

Substituting this into (2.13) and invoking Lemma 2.1 completes the proof. \blacksquare

The computation of the Gateaux derivative of the enriched free energy (1.35) at a measure $\mu \in \mathcal{M}_+$ in the direction of a probability measure $\nu \in \text{Pr}[-1, 1]$,

$$D_\mu \bar{F}_N(t, \mu; \nu) = \lim_{\varepsilon \rightarrow 0} \frac{\bar{F}_N(t, \mu + \varepsilon \nu) - \bar{F}_N(t, \mu)}{\varepsilon}, \quad (2.14)$$

is slightly more involved. It will be useful to compute the derivative of the free energy (1.29) with respect to the parameter $s \geq 0$ first. Fix a probability measure $\mu \in \text{Pr}[-1, 1]$ and a time $t \geq 0$. For each $i \leq N$ write

$$\begin{aligned} \tilde{H}_{N,m}^{s,i}(\sigma) = & \sum_{j \neq i} \sum_{k \leq \Pi_{j,s}} \log \left[(c + \Delta \sigma_j x_{j,k})^{\tilde{G}_{j,k}^x} \left(1 - \frac{c + \Delta \sigma_j x_{j,k}}{N} \right)^{1 - \tilde{G}_{j,k}^x} \right] \\ & + \sum_{k \leq m} \log \left[(c + \Delta \sigma_i x_{i,k})^{\tilde{G}_{i,k}^x} \left(1 - \frac{c + \Delta \sigma_i x_{i,k}}{N} \right)^{1 - \tilde{G}_{i,k}^x} \right] \end{aligned} \quad (2.15)$$

for the Hamiltonian (1.23) conditional on the i 'th Poisson sum containing m terms, and denote by

$$Z_{N,m}^{s,i} = \int_{\Sigma_N} \exp(H_N^t(\sigma) + \tilde{H}_{N,m}^{s,i}(\sigma)) dP_N^*(\sigma) \quad (2.16)$$

its associated partition function. In this notation, the free energy (1.29) may be expressed as

$$\tilde{F}_N(t, s, \mu) = \frac{1}{N} \sum_{m \geq 0} \pi(sN, m) \mathbb{E} \log Z_{N,m}^{s,i}. \quad (2.17)$$

Lemma 2.3. *For any $t > 0$, $s > 0$ and $\mu \in \text{Pr}[-1, 1]$,*

$$\partial_s \tilde{F}_N(t, s, \mu) = \mathbb{E} (c + \Delta \langle \sigma_i \rangle x_i) \log (c + \Delta \langle \sigma_i \rangle x_i) - c - \Delta \bar{m} \mathbb{E} x_1 + \mathcal{O}(N^{-1}), \quad (2.18)$$

where the index $i \in \{1, \dots, N\}$ is uniformly sampled and the random variables (x_i) are sampled from the measure μ independently of all other sources of randomness.

Proof. Conditioning on the number of terms in each of the Poisson sums that appear in the definition of the free energy (1.29) and leveraging the product rule as well as equations (2.17) and (2.8), we see that

$$\partial_s \tilde{F}_N(t, s, \mu) = \frac{1}{N} \sum_{i \leq N} \sum_{m \geq 0} \partial_s \pi(sN, m) \mathbb{E} \log Z_{N,m}^{s,i} = \sum_{i \leq N} \sum_{m \geq 0} \pi(sN, m) \mathbb{E} \log \frac{Z_{N,m+1}^{s,i}}{Z_{N,m}^{s,i}}. \quad (2.19)$$

For each $i \leq N$, denote by x_i a sample from the measure μ , and write \tilde{G}_i^x for a random variable with conditional distribution (1.24). These random variables are taken to be independent for different values of $i \leq N$, and independent of all other sources of randomness. Since

$$Z_{N,m+1}^{s,i} \stackrel{d}{=} \int_{\Sigma_N} (c + \Delta \sigma_i x_i)^{\tilde{G}_i^x} \left(1 - \frac{c + \Delta \sigma_i x_i}{N} \right)^{1 - \tilde{G}_i^x} \exp(H_N^t(\sigma) + \tilde{H}_{N,m}^{s,i}(\sigma)) dP_N^*(\sigma), \quad (2.20)$$

it follows by (2.19) and the definition of the Gibbs average in (2.1) that

$$\partial_s \tilde{F}_N(t, s, \mu) = \sum_{i \leq N} \mathbb{E} \log \left\langle (c + \Delta \sigma_i x_i)^{\tilde{G}_i^x} \left(1 - \frac{c + \Delta \sigma_i x_i}{N} \right)^{1 - \tilde{G}_i^x} \right\rangle. \quad (2.21)$$

Remembering the explicit form of the conditional distribution (1.18) reveals that

$$\begin{aligned} \partial_s \tilde{F}_N(t, s, \mu) &= \frac{1}{N} \sum_{i \leq N} \mathbb{E} (c + \Delta \sigma_i^* x_i) \log \langle c + \Delta \sigma_i x_i \rangle \\ &\quad + \sum_{i \leq N} \mathbb{E} \left(1 - \frac{c + \Delta \sigma_i^* x_i}{N} \right) \log \left\langle 1 - \frac{c + \Delta \sigma_i x_i}{N} \right\rangle. \end{aligned}$$

Taylor expanding the logarithm and keeping only first order terms reduces this to

$$\begin{aligned} \partial_s \tilde{F}_N(t, s, \mu) &= \frac{1}{N} \sum_{i \leq N} \mathbb{E} (c + \Delta \sigma_i^* x_i) \log \langle c + \Delta \sigma_i x_i \rangle - c - \frac{\Delta}{N} \sum_{i \leq N} \mathbb{E} x_i \mathbb{E} \langle \sigma_i \rangle + \mathcal{O}(N^{-1}) \\ &= \frac{1}{N} \sum_{i \leq N} \mathbb{E} (c + \Delta \sigma_i^* x_i) \log (c + \Delta \langle \sigma_i \rangle x_i) - c - \Delta \bar{m} \mathbb{E} x_1 + \mathcal{O}(N^{-1}), \end{aligned}$$

where the second equality uses the Nishimori identity (2.3). Noticing that the Gibbs average $\langle \sigma_i \rangle$ is a measurable function of the data by (2.2) and applying the Nishimori identity (2.3) completes the proof. \blacksquare

Before leveraging this result to compute the Gateaux derivative (2.14), we begin with some distributional identities which will simplify the calculation. Fix a finite measure $\mu \in \mathcal{M}_+$ and a probability measure $\nu \in \text{Pr}[-1, 1]$. Let $s = \mu[-1, 1]$ and fix $\varepsilon > 0$. Introduce the measure $\lambda = \mu + \varepsilon \nu$ and observe that

$$\bar{\lambda} = \frac{\lambda}{s + \varepsilon} = \frac{s}{s + \varepsilon} \bar{\mu} + \frac{\varepsilon}{s + \varepsilon} \nu. \quad (2.22)$$

Denote by $(x_{i,k})$ i.i.d. random variables sampled from the measure $\bar{\mu}$, and write $(y_{i,k})$ for i.i.d. random variables sampled from the measure ν . Given i.i.d. random variables $(w_{i,k})$ with distribution $\text{Ber}(\frac{s}{s+\varepsilon})$, notice that by (2.22) the random variables

$$z_{i,k} = x_{i,k}^{w_{i,k}} y_{i,k}^{1-w_{i,k}} \quad (2.23)$$

are i.i.d. with distribution $\bar{\lambda}$. In particular, if $(\tilde{G}_{i,k}^z)$ are independent random variables with conditional distribution (1.24), the Hamiltonian (1.23) may be expressed as

$$\tilde{H}_N^{s+\varepsilon, \bar{\lambda}}(\sigma) \stackrel{d}{=} \sum_{i \leq N} \sum_{k \leq \Pi_{i,s+\varepsilon}} \log \left[(c + \Delta \sigma_i z_{i,k}) \tilde{G}_{i,k}^z \left(1 - \frac{c + \Delta \sigma_i z_{i,k}}{N} \right)^{1 - \tilde{G}_{i,k}^z} \right]. \quad (2.24)$$

This identity will allow us to linearize the enriched free energy (1.35) upon realizing that

$$\bar{F}_N(t, \mu + \varepsilon \nu) = \tilde{F}_N(t, s + \varepsilon, \bar{\lambda}). \quad (2.25)$$

To make the computation as clear as possible, it will be convenient to introduce additional notation. In the same spirit as (2.15), for each $i \leq N$, write

$$\tilde{H}_{N,m}^{s,i,+}(\sigma) = \tilde{H}_{N,m}^{s,i} + \log \left[(c + \Delta \sigma_i y_i) \tilde{G}_i^y \left(1 - \frac{c + \Delta \sigma_i y_i}{N} \right)^{1 - \tilde{G}_i^y} \right] \quad (2.26)$$

for the Hamiltonian (1.23) conditional on the i 'th Poisson sum containing $m+1$ terms one of which is sampled from the measure ν . Denote by

$$Z_{N,m}^{s,i,+} = \int_{\Sigma_N} \exp(H_N^t(\sigma) + \tilde{H}_{N,m}^{s,i,+}(\sigma)) dP_N^* \quad (2.27)$$

its associated partition function. It will be useful to record the following consequence of Taylor's theorem,

$$\left(\frac{s}{s+\varepsilon}\right)^{\sum_{i \leq N} m_i} = 1 - \frac{\varepsilon}{s+\varepsilon} \sum_{i \leq N} m_i + o(\varepsilon), \quad (2.28)$$

as well as the elementary identity,

$$m\pi(\lambda, m) = \lambda\pi(\lambda, m-1). \quad (2.29)$$

Lemma 2.4. *For any $t > 0$, $\mu \in \mathcal{M}_+$ and $\nu \in \text{Pr}[-1, 1]$,*

$$\begin{aligned} D_\mu \bar{F}_N(t, \mu; \nu) &= \mathbb{E}(c + \Delta \langle \sigma_i \rangle y_i) \log(c + \Delta \langle \sigma_i \rangle y_i) \\ &\quad + N \mathbb{E}\left(1 - \frac{c + \Delta \langle \sigma_i \rangle y_i}{N}\right) \log\left(1 - \frac{c + \Delta \langle \sigma_i \rangle y_i}{N}\right) \end{aligned} \quad (2.30)$$

where the index $i \in \{1, \dots, N\}$ is uniformly sampled and the random variables (y_i) are sampled from the measure ν independently of all other sources of randomness.

Proof. Leveraging (2.24), conditioning on the number of random variables $(w_{i,k})$ that are equal to one and using (2.28), we see that

$$\begin{aligned} \tilde{F}_N(t, s + \varepsilon, \bar{\lambda}) &= \tilde{F}_N(t, s + \varepsilon, \bar{\mu}) - \frac{\varepsilon}{s + \varepsilon} \sum_{i \leq N} \sum_{m \geq 0} m\pi((s + \varepsilon)N, m) \frac{1}{N} \mathbb{E} \log Z_{n,m}^{s+\varepsilon,i} \\ &\quad + \frac{\varepsilon}{s + \varepsilon} \sum_{i \leq N} \sum_{m \geq 0} (m+1)\pi((s + \varepsilon)N, m+1) \frac{1}{N} \mathbb{E} \log Z_{N,m}^{s+\varepsilon,i,+} + o(\varepsilon). \end{aligned}$$

Keeping in mind (2.29), this simplifies to

$$\begin{aligned} \tilde{F}_N(t, s + \varepsilon, \bar{\lambda}) &= \tilde{F}_N(t, s + \varepsilon, \bar{\mu}) \\ &\quad + \varepsilon \sum_{i \leq N} \sum_{m \geq 0} \pi((s + \varepsilon)N, m) \left(\mathbb{E} \log \frac{Z_{N,m}^{s+\varepsilon,i,+}}{Z_{N,m}^{s+\varepsilon,i}} - \mathbb{E} \log \frac{Z_{N,m+1}^{s+\varepsilon,i}}{Z_{N,m}^{s+\varepsilon,i}} \right) + o(\varepsilon). \end{aligned} \quad (2.31)$$

For each $i \leq N$, denote by x_i a sample from the measure $\bar{\mu}$ and by y_i a sample from the measure ν . Write \tilde{G}_i^x and \tilde{G}_i^y for random variables with conditional distribution (1.24). These random variables are taken to be independent for different values of $i \leq N$, and independent of all other sources of randomness. Combining (2.31) with (2.20) and the distributional identity

$$Z_{N,m}^{s+\varepsilon,i,+} \stackrel{d}{=} \int_{\Sigma_N} (c + \Delta \sigma_i y_i)^{\tilde{G}_i^y} \left(1 - \frac{c + \Delta \sigma_i y_i}{N}\right)^{1 - \tilde{G}_i^y} \exp(H_N^t(\sigma) + \tilde{H}_{N,m}^{s+\varepsilon,i}(\sigma)) dP_N^*(\sigma)$$

yields

$$\begin{aligned}\tilde{F}_N(t, s + \varepsilon, \bar{\lambda}) &= \tilde{F}_N(t, s + \varepsilon, \bar{\mu}) + \varepsilon \sum_{i \leq N} \mathbb{E} \log \left\langle (c + \Delta \sigma_i y_i)^{\tilde{G}_i^y} \left(1 - \frac{c + \Delta \sigma_i y_i}{N}\right)^{1 - \tilde{G}_i^y} \right\rangle \\ &\quad - \varepsilon \sum_{i \leq N} \mathbb{E} \log \left\langle (c + \Delta \sigma_i x_i)^{\tilde{G}_i^x} \left(1 - \frac{c + \Delta \sigma_i x_i}{N}\right)^{1 - \tilde{G}_i^x} \right\rangle + o(\varepsilon).\end{aligned}\quad (2.32)$$

Together with (2.21), this implies that

$$\begin{aligned}D_\mu \bar{F}_N(t, \mu; \nu) &= \lim_{\varepsilon \rightarrow 0} \frac{\tilde{F}_N(t, s + \varepsilon, \bar{\lambda}) - \tilde{F}_N(t, s + \varepsilon, \bar{\mu})}{\varepsilon} + \partial_s \bar{F}_N(t, s, \bar{\mu}) \\ &= \sum_{i \leq N} \mathbb{E} \log \left\langle (c + \Delta \sigma_i y_i)^{\tilde{G}_i^y} \left(1 - \frac{c + \Delta \sigma_i y_i}{N}\right)^{1 - \tilde{G}_i^y} \right\rangle.\end{aligned}$$

Notice that the Gibbs averages in (2.32) depend on ε , so in taking this limit we have implicitly used the fact that this dependence is continuous. Proceeding exactly as in the proof of Lemma 2.3 (after display (2.21)) completes the proof. \blacksquare

To compare (2.30) with the time-derivative of the enriched free energy in Corollary 2.2 it will be convenient to once again Taylor expand the logarithm. We will write

$$\begin{aligned}D_\mu \bar{F}_N(t, \mu, x) &= \mathbb{E} (c + \Delta \langle \sigma_i \rangle x) \log (c + \Delta \langle \sigma_i \rangle x) \\ &\quad + N \mathbb{E} \left(1 - \frac{c + \Delta \langle \sigma_i \rangle x}{N}\right) \log \left(1 - \frac{c + \Delta \langle \sigma_i \rangle x}{N}\right)\end{aligned}\quad (2.33)$$

for the density of the Gateaux derivative $D_\mu \bar{F}_N(t, \mu)$. Taylor expanding the logarithm shows that

$$D_\mu \bar{F}_N(t, \mu, x) = \mathbb{E} (c + \Delta \langle \sigma_i \rangle x) \log (c + \Delta \langle \sigma_i \rangle x) - c - \Delta \bar{m} x + \mathcal{O}(N^{-1}).\quad (2.34)$$

Corollary 2.5. *For every $t > 0$ and $\mu \in \mathcal{M}_+$,*

$$D_\mu \bar{F}_N(t, \mu, x) = (c + \Delta \bar{m} x) \log(c) + c \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} \mathbb{E} \langle R_{[n]} \rangle x^n - c + \mathcal{O}(N^{-1}).\quad (2.35)$$

Proof. Fix $\nu \in \text{Pr}[-1, 1]$. A Taylor expansion of the logarithm shows that

$$\begin{aligned}\mathbb{E} (c + \Delta \langle \sigma_i \rangle y_i) \log (c + \Delta \langle \sigma_i \rangle y_i) &= (c + \Delta \bar{m} \mathbb{E} y_1) \log(c) \\ &\quad - \sum_{n \geq 1} \frac{(-\Delta/c)^n}{n} \mathbb{E} (c + \Delta \langle \sigma_i \rangle y_i) y_i^n \langle \sigma_i \rangle^n,\end{aligned}\quad (2.36)$$

where we have used that y_i is sampled from $\nu \in \text{Pr}[-1, 1]$ and is independent of all other sources of randomness. Since $|\Delta| < c$ and $\mathbb{E} \langle \sigma_i \rangle = \bar{m}$ by the Nishimori identity,

$$\begin{aligned}\sum_{n \geq 1} \frac{(-\Delta/c)^n}{n} \mathbb{E} (c + \Delta \langle \sigma_i \rangle y_i) y_i^n \langle \sigma_i \rangle^n &= \sum_{n \geq 1} \frac{(-\Delta/c)^n}{n} (c \mathbb{E} y_i^n \mathbb{E} \langle \sigma_i \rangle^n + \Delta \mathbb{E} y_i^{n+1} \mathbb{E} \langle \sigma_i \rangle^{n+1}) \\ &= -\Delta \mathbb{E} y_1 \mathbb{E} \langle \sigma_i \rangle + c \sum_{n \geq 2} \left(\frac{(-\Delta/c)^n}{n} - \frac{(-\Delta/c)^n}{n-1} \right) \mathbb{E} y_i^n \mathbb{E} \langle \sigma_i \rangle^n \\ &= -\Delta \bar{m} \mathbb{E} y_1 - c \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} \mathbb{E} \langle R_{[n]} \rangle \mathbb{E} y_i^n.\end{aligned}$$

Substituting this into (2.36) and recalling (2.34) completes the proof. \blacksquare

This result implies that the Gateaux derivative density (2.33) is close to an element in the cone of functions (1.37). Indeed, if $\mu^* = \mathcal{L}(\langle \sigma_i \rangle)$ denotes the law of the Gibbs average of a uniformly sampled spin coordinate, then (2.34) may be formally rewritten as

$$D_\mu \bar{F}_N(t, \mu, x) \simeq \int_{-1}^1 g(xy) d\mu^*(y) = G_{\mu^*}(x), \quad (2.37)$$

using the Nishimori identity (2.3) to assert that $\mathbb{E}\langle \sigma_i \rangle = \bar{m}$. It follows by another application of the Nishimori identity that

$$C_\infty(D_\mu \bar{F}_N(t, \mu)) \simeq \frac{1}{2}(c + \Delta \bar{m}^2) \log(c) + \frac{c}{2} \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} (\mathbb{E}\langle R_{[n]} \rangle)^2 - \frac{c}{2}. \quad (2.38)$$

Comparing this with the expression in Corollary 2.2, and assuming the approximate concentration of all the multi-overlaps,

$$\mathbb{E}\langle R_{[n]}^2 \rangle \simeq (\mathbb{E}\langle R_{[n]} \rangle)^2, \quad (2.39)$$

reveals that, up to a small error, the enriched free energy (1.35) formally satisfies the infinite-dimensional Hamilton-Jacobi equation (1.41),

$$\partial_t f(t, \mu) = C_\infty(D_\mu f(t, \mu)) \quad \text{on } \mathbb{R}_{>0} \times \mathcal{M}_+. \quad (2.40)$$

As already mentioned in the introduction, the difficulty in making this informal derivation rigorous is that we do not expect the concentration of the multi-overlaps (2.39) to be valid for each choice of the parameters t and μ . On the positive side, the arguments in [14] reveal that the concentration of the multi-overlaps can be enforced through a small perturbation of the Hamiltonian which does not affect the limit of the free energy, for most values of the perturbation parameters. Yet, the solution theory to Hamilton-Jacobi equations is rather sensitive to details, and in particular, this control “for most values” or after a suitable local averaging is not sufficient to allow us to conclude. The next section will clarify the nature of this solution theory.

3 Well-posedness of the Hamilton-Jacobi equation

In this section, we leverage the main result in [31] to establish the well-posedness of the infinite-dimensional Hamilton-Jacobi equation (1.41). The first order of business will be to identify the initial condition in (1.41). For each integer $N \geq 1$, let

$$\psi_N(\mu) = \bar{F}_N(0, \mu) = \tilde{F}_N(0, \mu[-1, 1], \bar{\mu}), \quad (3.1)$$

and notice that the initial condition $\psi : \mathcal{M}_+ \rightarrow \mathbb{R}$ in (1.41) should be the limit of the initial conditions (ψ_N) . Following [13], we will first compute this limit for discrete measures $\mu \in \mathcal{M}_+$, and later show that the convergence may be extended to all measures in \mathcal{M}_+ through a density argument. Given a measure $\mu \in \mathcal{M}_+$, it will be convenient to write $\Pi_\pm(\mu)$ for the generalized Poisson point process with mean measure $(c \pm \Delta x) d\mu(x)$ on $[-1, 1]$. By generalized Poisson process we mean that an atom $a \in [-1, 1]$ with $\mu(\{a\}) > 0$ is counted with independent Poisson multiplicity having mean $(c \pm \Delta a)\mu(\{a\})$. It turns out that the limit of the initial conditions (3.1) is given by an appropriate average with respect to the randomness of the generalized Poisson point processes $\Pi_\pm(\mu)$, which we denote by \mathbb{E} as usual.

Lemma 3.1. For any discrete measure $\mu \in \mathcal{M}_+$, the sequence $(\psi_N(\mu))$ converges to

$$\begin{aligned} \psi(\mu) = & -\mu[-1, 1]c + p \mathbb{E} \log \int_{\Sigma_1} \exp(-\mu[-1, 1]\Delta\sigma \mathbb{E}x_1) \prod_{x \in \Pi_+(\mu)} (c + \Delta\sigma x) dP^*(\sigma) \\ & + (1-p) \mathbb{E} \log \int_{\Sigma_1} \exp(-\mu[-1, 1]\Delta\sigma \mathbb{E}x_1) \prod_{x \in \Pi_-(\mu)} (c + \Delta\sigma x) dP^*(\sigma), \end{aligned} \quad (3.2)$$

where x_1 has law $\bar{\mu}$.

Remark 3.2. From this lemma and its extension to any $\mu \in \mathcal{M}_+$ proved in Proposition 3.6 below, one can also show that for every $\mu \in \mathcal{M}_+$, the density of the Gateaux derivative $D_\mu \psi(\mu)$ is

$$\begin{aligned} D_\mu \psi(\mu, x) = & p \mathbb{E} \langle c + \Delta\sigma x \rangle_+ \log \langle c + \Delta\sigma x \rangle_+ \\ & + (1-p) \mathbb{E} \langle c + \Delta\sigma x \rangle_- \log \langle c + \Delta\sigma x \rangle_- - c - \Delta\bar{m}x, \end{aligned} \quad (3.3)$$

where $\langle \cdot \rangle_\pm$ denote the Gibbs averages given by

$$\langle f(\sigma) \rangle_\pm := \frac{\int_{\Sigma_1} f(\sigma) \exp(-\mu[-1, 1]\Delta\sigma \mathbb{E}x_1) \prod_{x \in \Pi_\pm(\mu)} (c + \Delta\sigma x) dP^*(\sigma)}{\int_{\Sigma_1} \exp(-\mu[-1, 1]\Delta\sigma \mathbb{E}x_1) \prod_{x \in \Pi_\pm(\mu)} (c + \Delta\sigma x) dP^*(\sigma)}. \quad (3.4)$$

Proof of Lemma 3.1. Since $\mu \in \mathcal{M}_+$ is a discrete measure, it may be expressed as

$$\mu = \sum_{\ell \leq K} p_\ell \delta_{a_\ell}$$

for some integer $K \geq 1$, some atoms $a_\ell \in [-1, 1]$ and some weights $p_\ell \geq 0$. Let $s = \mu[-1, 1]$, and introduce independent Poisson random variables $\Pi_{i,s} \sim \text{Poi}(sN)$ in such a way that

$$\psi_N(\mu) = \frac{1}{N} \sum_{i \leq N} \mathbb{E} \log \int_{\Sigma_1} \exp \sum_{k \leq \Pi_{i,s}} \log \left[(c + \Delta\sigma x_{i,k}) \tilde{G}_{i,k}^x \left(1 - \frac{c + \Delta\sigma x_{i,k}}{N} \right)^{1 - \tilde{G}_{i,k}^x} \right] dP^*(\sigma),$$

where $(x_{i,k})$ are i.i.d. random variables with law $\bar{\mu}$. Since each of the expectations in this average is the same,

$$\psi_N(\mu) = \mathbb{E} \log \int_{\Sigma_1} \exp \sum_{k \leq \Pi_{1,s}} \log \left[(c + \Delta\sigma x_k) \tilde{G}_k^x \left(1 - \frac{c + \Delta\sigma x_k}{N} \right)^{1 - \tilde{G}_k^x} \right] dP^*(\sigma), \quad (3.5)$$

where (x_k) are i.i.d. random variables with law $\bar{\mu}$ and \tilde{G}_k^x has conditional distribution (1.24) for $i = 1$ and $x_{1,k}$ is replaced by x_k . To simplify this further, introduce the random index sets

$$\mathcal{I}_0 = \{k \leq \Pi_{1,s} \mid \tilde{G}_k^x = 0\} \quad \text{and} \quad \mathcal{I}_1 = \{k \leq \Pi_{1,s} \mid \tilde{G}_k^x = 1\}.$$

Decomposing the sum in (3.5) according to the partition $\{k \leq \Pi_{1,s}\} = \mathcal{I}_0 \sqcup \mathcal{I}_1$ and applying Taylor's theorem to the logarithm reveals that

$$\psi_N(\mu) = \mathbb{E} \log \int_{\Sigma_1} \prod_{k \in \mathcal{I}_1} (c + \Delta\sigma x_k) \exp \left(-\frac{\Delta\sigma}{N} \sum_{k \in \mathcal{I}_0} x_k \right) dP^*(\sigma) - \left(\frac{c}{N} + \mathcal{O}(N^{-2}) \right) \mathbb{E} |\mathcal{I}_0|.$$

Conditionally on σ^* , $\Pi_{1,s}$ and (x_k) , the random variable $|\mathcal{J}_0|$ is a sum of Bernoulli random variables with probability of success $1 - \frac{c + \Delta\sigma_1^* x_k}{N}$. It therefore has mean

$$\mathbb{E}|\mathcal{J}_0| = \mathbb{E}\Pi_{1,s} \mathbb{E} \left(1 - \frac{c + \Delta\sigma_1^* x_1}{N} \right) = sN \left(1 - \frac{c + \Delta\bar{m} \mathbb{E}x_1}{N} \right) \quad (3.6)$$

and variance bounded by

$$\text{Var}|\mathcal{J}_0| = \mathbb{E}\Pi_{1,s} \mathbb{E} \left(1 - \frac{c + \Delta\sigma_1^* x_1}{N} \right) \left(\frac{c + \Delta\sigma_1^* x_1}{N} \right) \leq s(c + |\Delta|). \quad (3.7)$$

Using (3.6) and introducing the random index sets $\mathcal{J}_1^\ell = \{k \in \mathcal{J}_1 \mid x_k = a_\ell\}$ reveals that

$$\psi_N(\mu) = \mathbb{E} \log \int_{\Sigma_1} \prod_{\ell \leq K} \prod_{k \in \mathcal{J}_1^\ell} (c + \Delta\sigma a_\ell) \exp \left(-\frac{\Delta\sigma}{N} \sum_{k \in \mathcal{J}_0} x_k \right) dP^*(\sigma) - cs + \mathcal{O}(N^{-1}).$$

Observe that for any $\sigma \in \Sigma_1$,

$$\begin{aligned} \mathbb{E} \left| -\frac{\Delta\sigma}{N} \sum_{k \in \mathcal{J}_0} x_k + \Delta\sigma s \mathbb{E}x_1 \right| &\leq |\Delta|s \mathbb{E} \left| \frac{1}{Ns} \sum_{k \in \mathcal{J}_0} x_k - \mathbb{E}x_1 \right| \\ &\leq |\Delta|s \mathbb{E} \left| \frac{1}{Ns} \sum_{k \leq Ns} x_k - \mathbb{E}x_1 \right| + \frac{\Delta}{N} (\text{Var}|\mathcal{J}_0| + |\mathbb{E}|\mathcal{J}_0| - Ns) \end{aligned}$$

where we have used the fact that $|x_k| \leq 1$ and Jensen's inequality in the second inequality. Recalling (3.7) and invoking the strong law of large numbers shows that

$$\psi_N(\mu) = \mathbb{E} \log \int_{\Sigma_1} \exp(-\Delta\sigma s \mathbb{E}x_1) \prod_{\ell \leq K} \prod_{k \in \mathcal{J}_1^\ell} (c + \Delta\sigma a_\ell) dP^*(\sigma) - cs + o_N(1). \quad (3.8)$$

The Poisson coloring theorem (see Chapter 5 in [40]) implies that $|\mathcal{J}_1^\ell|$ is a Poisson random variable with mean

$$\mathbb{E}\Pi_{1,s} \cdot \mathbb{P}\{\tilde{G}_1^x = 1, x_1 = a_\ell\} = sN \cdot \frac{c + \Delta\sigma_1^* a_\ell}{N} \cdot \bar{\mu}(a_\ell) = (c + \Delta\sigma_1^* a_\ell) \mu(a_\ell),$$

so averaging (3.8) over the randomness of σ^* yields

$$\begin{aligned} \psi_N(\mu) &= -cs + p \mathbb{E} \log \int_{\Sigma_1} \exp(-\Delta\sigma s \mathbb{E}x_1) \prod_{x \in \Pi_+(\mu)} (c + \Delta\sigma x) dP^*(\sigma) \\ &\quad + (1-p) \mathbb{E} \log \int_{\Sigma_1} \exp(-\Delta\sigma s \mathbb{E}x_1) \prod_{x \in \Pi_-(\mu)} (c + \Delta\sigma x) dP^*(\sigma) + o_N(1). \end{aligned}$$

This completes the proof. ■

To extend this convergence to all measures in \mathcal{M}_+ , we will rely upon the continuity of the functional (3.2) with respect to the Wasserstein distance on the space of probability measures,

$$W(\mathbb{P}, \mathbb{Q}) = \sup \left\{ \left| \int_{-1}^1 h(x) d\mathbb{P}(x) - \int_{-1}^1 h(x) d\mathbb{Q}(x) \right| \mid \|h\|_{\text{Lip}} \leq 1 \right\} \quad (3.9)$$

$$= \inf \left\{ \int_{[-1,1]^2} |x-y| d\mathbf{v}(x,y) \mid \mathbf{v} \in \text{Pr}([-1,1]^2) \text{ has marginals } \mathbb{P} \text{ and } \mathbb{Q} \right\}. \quad (3.10)$$

Here $\|\cdot\|_{\text{Lip}}$ denotes the Lipschitz semi-norm

$$\|h\|_{\text{Lip}} = \sup_{x \neq x' \in [-1, 1]} \frac{|h(x) - h(x')|}{|x - x'|} \quad (3.11)$$

defined on the space of functions $h : [-1, 1] \rightarrow \mathbb{R}$. This continuity will be obtained as a consequence of the following uniform bound on the spatial derivatives of the Gateaux derivative density (2.33).

Lemma 3.3. *For every N large enough (relative to c), $\mu \in \mathcal{M}_+$, $t \geq 0$ and $x \in [-1, 1]$,*

$$|D_\mu \bar{F}_N(t, \mu, x)| \leq 2c(2 + |\log(2c)| + |\log(c - |\Delta|)|), \quad (3.12)$$

$$|\partial_x D_\mu \bar{F}_N(t, \mu, x)| \leq c(1 + |\log(2c)| + |\log(c - |\Delta|)|). \quad (3.13)$$

Proof. Recall from (2.33) that

$$\begin{aligned} D_\mu \bar{F}_N(t, \mu, x) &= \mathbb{E} (c + \Delta \langle \sigma_i \rangle x) \log (c + \Delta \langle \sigma_i \rangle x) \\ &\quad + N \mathbb{E} \left(1 - \frac{c + \Delta \langle \sigma_i \rangle x}{N} \right) \log \left(1 - \frac{c + \Delta \langle \sigma_i \rangle x}{N} \right). \end{aligned}$$

It follows by a direct computation that

$$\partial_x D_\mu \bar{F}_N(t, \mu, x) = \Delta \mathbb{E} \langle \sigma_i \rangle \log (c + \Delta \langle \sigma_i \rangle x) - \Delta \mathbb{E} \langle \sigma_i \rangle \log \left(1 - \frac{c + \Delta \langle \sigma_i \rangle x}{N} \right).$$

Since all spin configuration coordinates are bounded by one and $|\Delta| < c$, Taylor's theorem implies that for N large enough,

$$\begin{aligned} |D_\mu \bar{F}_N(t, \mu, x)| &\leq 2c(2 + |\log(2c)| + |\log(c - |\Delta|)|), \\ |\partial_x D_\mu \bar{F}_N(t, \mu, x)| &\leq c(1 + |\log(2c)| + |\log(c - |\Delta|)|). \end{aligned}$$

Notice that the choice of N only depends on c as $x \in [-1, 1]$ and $|\Delta| < c$. This completes the proof. \blacksquare

Lemma 3.4. *The initial condition ψ_N satisfies the Lipschitz bound*

$$|\psi_N(\mathbb{P}) - \psi_N(\mathbb{Q})| \leq c(1 + |\log(2c)| + |\log(c - |\Delta|)|) W(\mathbb{P}, \mathbb{Q}) \quad (3.14)$$

for all probability measures $\mathbb{P}, \mathbb{Q} \in \text{Pr}[-1, 1]$.

Proof. The fundamental theorem of calculus and the definition of the Gateaux derivative in (1.39) imply that

$$\psi_N(\mathbb{P}) - \psi_N(\mathbb{Q}) = \int_0^1 \frac{d}{dt} \psi_N(\mathbb{Q} + t(\mathbb{P} - \mathbb{Q})) dt = \int_0^1 D_\mu \psi_N(\mathbb{Q} + t(\mathbb{P} - \mathbb{Q}); \mathbb{P} - \mathbb{Q}) dt.$$

Since the Gateaux derivative of the initial condition admits a continuously differentiable density,

$$|\psi_N(\mathbb{P}) - \psi_N(\mathbb{Q})| \leq \int_0^1 \left| \int_{-1}^1 f_t(x) d\mathbb{P}(x) - \int_{-1}^1 f_t(x) d\mathbb{Q}(x) \right| dt$$

for the continuously differentiable function $f_t(x) = D_\mu \psi_N(\mathbb{Q} + t(\mathbb{P} - \mathbb{Q}), x)$. The mean value theorem and (3.13) reveal that $\|f_t\|_{\text{Lip}} \leq c(1 + |\log(2c)| + |\log(c - |\Delta|)|)$. It follows by definition of the Wasserstein distance (3.9) that

$$|\psi_N(\mu) - \psi_N(\nu)| \leq c(1 + |\log(2c)| + |\log(c - |\Delta|)|)W(\mathbb{P}, \mathbb{Q}).$$

This completes the proof. \blacksquare

Lemma 3.5. *The functional $\psi : \mathcal{M}_+ \rightarrow \mathbb{R}$ defined by (3.2) is continuous with respect to the weak convergence of measures. This means that for any sequence of measures $(\mu_n) \subset \mathcal{M}_+$ converging weakly to a measure $\mu \in \mathcal{M}_+$, we have*

$$\lim_{N \rightarrow \infty} \psi(\mu_n) = \psi(\mu). \quad (3.15)$$

Proof. To alleviate the exposition, we will instead prove the continuity of the functional

$$\psi^1(\mu) = \mathbb{E} \log \int_{\Sigma_1} \exp(-\mu[-1, 1] \Delta \sigma \mathbb{E} x_1) \prod_{x \in \Pi_+(\mu)} (c + \Delta \sigma x) dP^*(\sigma)$$

with respect to the weak convergence of measures. Up to an additive constant, the asymptotic initial condition $\psi(\mu)$ is the weighted average of $\psi^1(\mu)$ and another functional of the same form whose continuity can be established using an identical argument, so this suffices. For each measure $\mu \in \mathcal{M}_+$ introduce the Hamiltonian

$$H(\sigma, \mu) = -\Delta \sigma \int_{-1}^1 x d\mu(x) + \sum_{x \in \Pi_+(\mu)} (c + \Delta \sigma x)$$

in such a way that the asymptotic initial condition is its associated free energy,

$$\psi^1(\mu) = \mathbb{E} \log \int_{\Sigma_1} \exp H(\sigma, \mu) dP^*(\sigma).$$

Consider a sequence of measures $(\mu_n) \subset \mathcal{M}_+$ converging weakly to a measure $\mu \in \mathcal{M}_+$, and let Π_n and Π be independent Poisson random variables with means $\mu_n[-1, 1]$ and $\mu[-1, 1]$, respectively. Introduce a collection $(X_k^n, X_k)_{k \in \mathbb{N}}$ of i.i.d. random vectors with joint law $\nu \in \text{Pr}([-1, 1]^2)$ having marginals $\bar{\mu}_n$ and $\bar{\mu}$. In this way, the coordinates $(X_k^n)_{k \in \mathbb{N}}$ are i.i.d. with law $\bar{\mu}_n$, the coordinates $(X_k)_{k \in \mathbb{N}}$ are i.i.d. with law $\bar{\mu}$, and we have the equalities in distribution

$$\sum_{x \in \Pi_+(\mu_n)} (c + \Delta \sigma x) \stackrel{d}{=} \sum_{k \leq \Pi_n} (c + \Delta \sigma X_k^n) \quad \text{and} \quad \sum_{x \in \Pi_+(\mu)} (c + \Delta \sigma x) \stackrel{d}{=} \sum_{k \leq \Pi} (c + \Delta \sigma X_k).$$

It follows that for any $\sigma \in \Sigma_1$,

$$|H(\sigma, \mu_n) - H(\sigma, \mu)| \leq |\Delta| \left| \int_{-1}^1 x d\mu_n(x) - \int_{-1}^1 x d\mu(x) \right| + c |\Pi_n - \Pi| + |\Delta| \left| \sum_{k \leq \Pi_n} X_k^n - \sum_{k \leq \Pi} X_k \right|,$$

and therefore,

$$|\psi^1(\mu_n) - \psi^1(\mu)| \leq |\Delta| \left| \int_{-1}^1 x d\mu_n(x) - \int_{-1}^1 x d\mu(x) \right| + c \mathbb{E} |\Pi_n - \Pi| + |\Delta| \mathbb{E} \left| \sum_{k \leq \Pi_n} X_k^n - \sum_{k \leq \Pi} X_k \right|.$$

To simplify this further, define the random variable $\Pi'_n = \min(\Pi_n, \Pi)$, introduce a Poisson random variable Π''_n independent of all other sources of randomness with mean $|\mu_n[-1, 1] - \mu[-1, 1]|$, and define the collection of random variables (Z_k^n) by

$$Z_k^n = \begin{cases} X_k^n & \text{if } \Pi'_n = \Pi \\ X_k & \text{otherwise} \end{cases}.$$

The basic properties of Poisson random variables and the fact that $|Z_k^n| \leq 1$ imply that

$$\begin{aligned} |\psi^1(\mu_n) - \psi^1(\mu)| &\leq c|\Delta| \left| \int_{-1}^1 x d(\mu_n - \mu)(x) \right| + c \mathbb{E} \Pi''_n + |\Delta| \mathbb{E} \sum_{k \leq \Pi''_n} |Z_k^n| + |\Delta| \mathbb{E} \sum_{k \leq \Pi'_n} |X_k^n - X_k| \\ &\leq c \left(\left| \int_{-1}^1 x d(\mu_n - \mu)(x) \right| + 2 \mathbb{E} \Pi''_n + \mathbb{E} \Pi'_n \int_{[-1, 1]^2} |x - y| d\nu(x, y) \right). \end{aligned}$$

Taking the infimum over all couplings $\nu \in \Pr([-1, 1]^2)$ with marginals $\bar{\mu}_n$ and $\bar{\mu}$ reveals that for n large enough,

$$|\psi^1(\mu_n) - \psi^1(\mu)| \leq c \left(\left| \int_{-1}^1 x d(\mu_n - \mu)(x) \right| + 2|\mu_n[-1, 1] - \mu[-1, 1]| + 3\mu[-1, 1]W(\bar{\mu}_n, \bar{\mu}) \right),$$

where we have used that $\mathbb{E} \Pi'_n \leq \mathbb{E} \Pi + \mathbb{E} \Pi_n \leq 3\mu[-1, 1]$ for n large enough as μ_n converges weakly to μ . Letting n tend to infinity and recalling that the Wasserstein distance (3.10) metrizes the weak convergence of probability measures completes the proof. \blacksquare

Proposition 3.6. *For any measure $\mu \in \mathcal{M}_+$, the sequence $(\psi_N(\mu))$ converges to (3.2).*

Proof. Consider a sequence $(\mu_n)_{n \geq 1}$ of discrete measures such that $\mu_n[-1, 1] = \mu[-1, 1]$ for all $n \geq 1$ and $\bar{\mu}_n \rightarrow \bar{\mu}$ with respect to the Wasserstein distance (3.9). By the triangle inequality and Lemma 3.4,

$$\begin{aligned} |\psi(\mu) - \psi_N(\mu)| &\leq |\psi(\mu) - \psi(\mu_n)| + |\psi(\mu_n) - \psi_N(\mu_n)| + |\psi_N(\mu_n) - \psi_N(\mu)| \\ &\leq |\psi(\mu) - \psi(\mu_n)| + |\psi(\mu_n) - \psi_N(\mu_n)| \\ &\quad + c(1 + |\log(2c)| + |\log(c - |\Delta|)|)W(\bar{\mu}_n, \bar{\mu}) \end{aligned}$$

where we have used the fact that $\mu_n[-1, 1] = \mu[-1, 1]$. Combining Lemma 3.5 with Lemma 3.1 to let $N \rightarrow \infty$ and then $n \rightarrow \infty$ completes the proof. \blacksquare

This result identifies the initial condition for the infinite-dimensional Hamilton-Jacobi equation (1.41). We may now leverage the main result in [31] to establish the well-posedness of this infinite-dimensional Hamilton-Jacobi equation. Before we do this, let us introduce the notation used in [31]; for more details and motivation regarding this notation, we encourage the reader to consult [31]. Given an integer $K \geq 1$, we write

$$\mathcal{D}_K = \left\{ k = \frac{i}{2^K} \mid -2^K \leq i < 2^K \right\} \quad (3.16)$$

for the set of dyadic rationals on $[-1, 1]$ at scale K . We index vectors using the set of dyadic rationals, writing $x = (x_k)_{k \in \mathcal{D}_K} \in \mathbb{R}_{\geq 0}^{\mathcal{D}_K}$ for a non-negative sequence with values in $\mathbb{R}_{\geq 0}$ indexed by the dyadic rationals \mathcal{D}_K . More generally, given two sets \mathbb{A} and \mathcal{B} , we write $\mathbb{A}^{\mathcal{B}}$ for the set of functions from \mathcal{B} to \mathbb{A} . We denote the set of discrete measures supported on the dyadic rationals at scale K in the interval $[-1, 1]$ by

$$\mathcal{M}_+^{(K)} = \left\{ \mu \in \mathcal{M}_+ \mid \mu = \frac{1}{|\mathcal{D}_K|} \sum_{k \in \mathcal{D}_K} x_k \delta_k \text{ for some } x = (x_k)_{k \in \mathcal{D}_K} \in \mathbb{R}_{\geq 0}^{\mathcal{D}_K} \right\}, \quad (3.17)$$

and we project a general measure $\mu \in \mathcal{M}_+$ onto $\mathcal{M}_+^{(K)}$ via the mapping

$$x^{(K)}(\mu) = (|\mathcal{D}_K| \mu[k, k + 2^{-K}])_{k \in \mathcal{D}_K} \in \mathbb{R}_{\geq 0}^{\mathcal{D}_K}, \quad (3.18)$$

whose inverse assigns to each $x \in \mathbb{R}_{\geq 0}^{\mathcal{D}_K}$ the measure

$$\mu_x^{(K)} = \frac{1}{|\mathcal{D}_K|} \sum_{k \in \mathcal{D}_K} x_k \delta_k \in \mathcal{M}_+^{(K)}. \quad (3.19)$$

We identify any real-valued function $f : \mathbb{R}_{\geq 0} \times \mathcal{M}_+^{(K)} \rightarrow \mathbb{R}$ with the function

$$f^{(K)}(t, x) = f(t, \mu_x^{(K)}) \quad (3.20)$$

defined on $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}^{\mathcal{D}_K}$, and we identify the projection of the initial condition with the function

$$\psi^{(K)}(x) = \psi(\mu_x^{(K)}) \quad (3.21)$$

defined on $\mathbb{R}_{\geq 0}^{\mathcal{D}_K}$. The Gateaux derivative of a real-valued function $f : \mathbb{R}_{\geq 0} \times \mathcal{M}_+^{(K)} \rightarrow \mathbb{R}$ at the measure $\mu \in \mathcal{M}_+^{(K)}$ becomes the gradient $|\mathcal{D}_K| \nabla f^{(K)}(t, x^{(K)}(\mu))$ by duality. Indeed,

$$D_\mu f(t, \mu; \nu) = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} f^{(K)}(t, x^{(K)}(\mu) + \varepsilon x^{(K)}(\nu)) = \nabla f^{(K)}(t, x^{(K)}(\mu)) \cdot x^{(K)}(\nu) \quad (3.22)$$

for any direction $\nu \in \mathcal{M}_+^{(K)}$. We fix $b > 0$ large enough so the modified kernel

$$\tilde{g}_b(z) = g(z) + b \quad (3.23)$$

is strictly positive, and we introduce the symmetric matrices

$$G^{(K)} = \frac{1}{|\mathcal{D}_K|^2} (g(kk'))_{k, k' \in \mathcal{D}_K} \quad \text{and} \quad \tilde{G}_b^{(K)} = \frac{1}{|\mathcal{D}_K|^2} (\tilde{g}_b(kk'))_{k, k' \in \mathcal{D}_K} \quad (3.24)$$

in $\mathbb{R}^{\mathcal{D}_K \times \mathcal{D}_K}$. We also define the projected cone

$$\tilde{\mathcal{C}}_{b, K} = \left\{ \tilde{G}_b^{(K)} x \in \mathbb{R}^{\mathcal{D}_K} \mid x \in \mathbb{R}_{\geq 0}^{\mathcal{D}_K} \right\}, \quad (3.25)$$

the projected non-linearity $\tilde{\mathcal{C}}_{b, K} : \tilde{\mathcal{C}}_{b, K} \rightarrow \mathbb{R}$ given by

$$\tilde{\mathcal{C}}_{b, K}(\tilde{G}_b^{(K)} x) = \frac{1}{2} \tilde{G}_b^{(K)} x \cdot x = \frac{1}{2|\mathcal{D}_K|^2} \sum_{k, k' \in \mathcal{D}_K} \tilde{g}_b(kk') x_k x_{k'}, \quad (3.26)$$

and the closed convex set

$$\mathcal{H}_{=1,K} = \left\{ G^{(K)}x \in \mathbb{R}^{\mathcal{D}_K} \mid x \in \mathbb{R}_{\geq 0}^{\mathcal{D}_K} \text{ and } |||x|||_1 = a \right\}. \quad (3.27)$$

We will measure quantities in the normalized ℓ^1 and $\ell^{1,*}$ norms,

$$|||x|||_1 = \frac{1}{|\mathcal{D}_K|} \sum_{k \in \mathcal{D}_K} |x_k| \quad \text{and} \quad |||y|||_{1,*} = \max_{k \in \mathcal{D}_K} |\mathcal{D}_K| |y_k|, \quad (3.28)$$

and it will be convenient to write

$$B_{K,R} = \{y \in \mathbb{R}^{\mathcal{D}_K} \mid |||y|||_{1,*} \leq R\} \quad (3.29)$$

for the ball of radius $R > 0$ centered at the origin with respect to the normalized- $\ell^{1,*}$ norm in $\mathbb{R}^{\mathcal{D}_K}$. Recall that a function $h : \mathbb{R}^d \rightarrow \mathbb{R}$ is said to be non-decreasing if $h(y) \leq h(y')$ for all $y, y' \in \mathbb{R}^d$ with $y' - y \in \mathbb{R}_{\geq 0}^d$. Proposition 2.3 in [31] will give the existence of a uniformly Lipschitz continuous and non-decreasing non-linearity $\tilde{H}_{b,K,R}$ which agrees with $\tilde{C}_{b,K}$ on the ball $\mathcal{C}_{b,K} \cap B_{K,R}$, and we will define the solution to the infinite-dimensional Hamilton-Jacobi equation (1.41) by

$$f(t, \mu) = \lim_{K \rightarrow \infty} \left(\tilde{f}_{b,R}^{(K)}(t, x^{(K)}(\mu)) - b |||x^{(K)}(\mu)|||_1 - \frac{bt}{2} \right), \quad (3.30)$$

where $\tilde{f}_{b,R}^{(K)} : [0, \infty) \times \mathbb{R}_{\geq 0}^{\mathcal{D}_K} \rightarrow \mathbb{R}$ is the unique solution to the Hamilton-Jacobi equation

$$\partial_t \tilde{f}^{(K)}(t, x) = \tilde{H}_{b,K,R}(\nabla \tilde{f}^{(K)}(t, x)) \quad \text{on} \quad \mathbb{R}_{>0} \times \mathbb{R}_{>0}^{\mathcal{D}_K} \quad (3.31)$$

subject to the projection of the initial condition $\tilde{\psi}_b : \mathcal{M}_+ \rightarrow \mathbb{R}$ defined by

$$\tilde{\psi}_b(\mu) = \psi(\mu) + b \int_{-1}^1 d\mu. \quad (3.32)$$

The fact that the function f defined by (3.30) does not depend on the choice of the constants $b \in \mathbb{R}$ and $R > 0$ sufficiently large is established in Proposition 3.14, while the existence and uniqueness of the appropriate notion of solution to the Hamilton-Jacobi equation (3.31) is established in Proposition 3.12. In order to prove these propositions, we will rely on some results from [31]. To appeal to these results, we must start by verifying that the initial condition ψ satisfies a certain number of hypotheses introduced there. To state these hypotheses, given a closed convex set $\mathcal{K} \subset \mathbb{R}^{\mathcal{D}_K}$, write

$$\mathcal{K}' = \mathcal{K} + B_{K, 2^{-K/2}} \quad (3.33)$$

for the neighbourhood of radius $2^{-K/2}$ around \mathcal{K} in the normalized- $\ell^{1,*}$ norm, and denote by

$$\text{TV}(\mu, \nu) = \sup \{ |\mu(A) - \nu(A)| \mid A \text{ is a measurable subset of } [-1, 1] \} \quad (3.34)$$

the total variation distance on \mathcal{M}_+ . A non-differential criterion for the gradient of a Lipschitz continuous function to lie in a closed convex set is given in Proposition B.2 of [31].

H1 The initial condition $\psi : \mathcal{M}_+ \rightarrow \mathbb{R}$ is Lipschitz continuous with respect to the total variation distance (3.34),

$$|\psi(\mu) - \psi(\nu)| \leq \|\psi\|_{\text{Lip,TV}} \text{TV}(\mu, \nu) \quad (3.35)$$

for all measures $\nu, \mu \in \mathcal{M}_+$.

H2 The initial condition $\psi : \mathcal{M}_+ \rightarrow \mathbb{R}$ has the property that each of the projected initial conditions has its gradient in the set $\mathcal{K}'_{=1,K}$,

$$\nabla \psi^{(K)} \in L^\infty(\mathbb{R}_{\geq 0}^d; \mathcal{K}'_{=1,K}). \quad (3.36)$$

H3 The initial condition $\psi : \text{Pr}[-1, 1] \rightarrow \mathbb{R}$ is Lipschitz continuous with respect to the Wasserstein distance (3.9),

$$|\psi(\mathbb{P}) - \psi(\mathbb{Q})| \leq \|\psi\|_{\text{Lip,W}} W(\mathbb{P}, \mathbb{Q}) \quad (3.37)$$

for all probability measures $\mathbb{P}, \mathbb{Q} \in \text{Pr}[-1, 1]$.

Lemma 3.7. *The initial condition ψ in (3.2) satisfies (H1)-(H3).*

Proof. Recall that Lemma 3.3 implies the existence of a constant $C > 0$ which depends only on c and Δ such that for every integer $N \geq 1$, $\mu \in \mathcal{M}_+$ and $t \geq 0$,

$$|D_\mu \bar{F}_N(t, \mu, x)| \leq C \quad \text{and} \quad |\partial_x D_\mu \bar{F}_N(t, \mu, x)| \leq C. \quad (3.38)$$

To establish (H1) notice that for every integer $N \geq 1$ and $\mu, \nu \in \mathcal{M}_+$,

$$\psi_N(\mu) - \psi_N(\nu) = \int_0^1 D_\mu \psi_N(\nu + t(\mu - \nu); \mu - \nu) dt = \int_0^1 \int_{-1}^1 f_t(x) d(\mu - \nu)(x) dt$$

for the continuously differentiable function $f_t(x) = D_\mu \psi_N(\nu + t(\mu - \nu), x)$. To bound this integral by the total variation distance, let $\eta = \mu - \nu \in \mathcal{M}_s$, and use the Hahn-Jordan decomposition to write $\eta = \eta^+ - \eta^-$ for measures $\eta^+, \eta^- \in \mathcal{M}_+$ with the property that for some measurable set $D \subset [-1, 1]$ and all measurable sets $E \subset [-1, 1]$,

$$\eta^+(E) = \eta(E \cap D) \geq 0 \quad \text{and} \quad \eta^-(E) = -\eta(E \cap D^c) \geq 0.$$

The triangle inequality and the first bound in (3.38) imply that

$$\begin{aligned} |\psi_N(\mu) - \psi_N(\nu)| &\leq \left| \int_0^1 \int_{-1}^1 f_t(x) d\eta^+(x) \right| + \left| \int_0^1 \int_{-1}^1 f_t(x) d\eta^-(x) \right| \\ &\leq C(\eta^+[-1, 1] + \eta^-[-1, 1]) \\ &\leq 2C \text{TV}(\mu, \nu). \end{aligned}$$

Using Proposition 3.6 to let N tend to infinity establishes (H1). To prove (H2) notice that by (3.22) and (2.34), for every $y \in \mathbb{R}_{\geq 0}^{\mathcal{D}_K}$, there exists some probability measure $\mu^* \in \text{Pr}[-1, 1]$ with

$$\partial_{x_k} \psi_N^{(K)}(y) = \frac{1}{|\mathcal{D}_K|} D_\mu \psi_N(\mu_y^{(K)}, k) = \frac{1}{|\mathcal{D}_K|} G_{\mu^*}(k) + \mathcal{O}(N^{-1}).$$

If $\mu_K^* = \mu_{x^{(K)}(\mu^*)}^{(K)} \in \mathcal{M}_+^{(K)}$ denotes the projection of μ^* onto $\mathcal{M}_+^{(K)}$, then the mean value theorem implies that

$$|G_{\mu^*}(k) - G_{\mu_K^*}(k)| \leq \sum_{k' \in \mathcal{D}_K} \int_{k'}^{k'+2^{-K}} |g(ky) - g(kk')| d\mu^*(y) \leq \frac{\|g'\|_\infty}{2^K},$$

where we have used that $\mu_K^*(k') = \mu^*[k', k' + 2^{-K}]$ for every dyadic $k' \in \mathcal{D}_K$. This means that

$$\partial_{x_k} \psi_N^{(K)}(y) = \frac{1}{|\mathcal{D}_K|} G_{\mu_K^*}(k) + \mathcal{O}(2^{-2K}) + \mathcal{O}(N^{-1}) = G^{(K)}_{x^{(K)}}(\mu_K^*)_k + \mathcal{O}(2^{-2K}) + \mathcal{O}(N^{-1})$$

so, for K large enough, we have $\nabla \psi_N^{(K)}(y) = w + \mathcal{O}(N^{-1})$ for some $w \in \mathcal{H}'_{=1,K}$. At this point fix $c \in \mathbb{R}$ and $x, x' \in \mathbb{R}^d$ with $(x' - x) \cdot z \geq c$ for every $z \in \mathcal{H}'_{=1,K}$. The fundamental theorem of calculus reveals that

$$\psi_N^{(K)}(x') - \psi_N^{(K)}(x) = \int_0^1 \nabla \psi_N^{(K)}(tx' + (1-t)x) \cdot (x' - x) dt \geq c + \mathcal{O}(N^{-1}).$$

Using Proposition 3.6 to let N tend to infinity shows that $\psi^{(K)}(x') - \psi^{(K)}(x) \geq c$ and Proposition B.2 in [31] gives (H2). Finally, (H3) is a consequence of Lemma 3.4 and Proposition 3.6. This completes the proof. \blacksquare

This result allows us to invoke Proposition 2.3 in [31] to extend the non-linearity $\tilde{C}_{b,K}$ in (3.26). It will be convenient to write $M_b = \max_{[-1,1]} \tilde{g}_b$ and $m_b = \min_{[-1,1]} \tilde{g}_b > 0$.

Proposition 3.8. *For every $R > 0$, there exists a non-decreasing non-linearity $\tilde{H}_{b,K,R} : \mathbb{R}^{\mathcal{D}_K} \rightarrow \mathbb{R}$ which agrees with $\tilde{C}_{b,K}$ on $\tilde{\mathcal{C}}_{b,K} \cap B_{K,R}$ and satisfies the Lipschitz continuity property*

$$|\tilde{H}_{b,K,R}(y) - \tilde{H}_{b,K,R}(y')| \leq \frac{8RM_b}{m_b^2} \| \|y - y'\| \|_{1,*} \quad (3.39)$$

for all $y, y' \in \mathbb{R}^{\mathcal{D}_K}$.

The well-posedness of the Hamilton-Jacobi equation (3.31) is the content of Theorem 1.1 in [31]. Before stating this result, let us remind the reader of the notion of a viscosity solution and introduce some more notation.

Definition 3.9. *An upper semi-continuous function $u : [0, \infty) \times \mathbb{R}_{\geq 0}^{\mathcal{D}_K} \rightarrow \mathbb{R}$ is said to be a viscosity subsolution to (3.31) if, given any $\phi \in C^\infty((0, \infty) \times \mathbb{R}_{> 0}^{\mathcal{D}_K})$ with the property that $u - \phi$ has a local maximum at $(t^*, x^*) \in (0, \infty) \times \mathbb{R}_{> 0}^{\mathcal{D}_K}$,*

$$(\partial_t \phi - \tilde{H}_{b,K,R}(\nabla \phi))(t^*, x^*) \leq 0. \quad (3.40)$$

Definition 3.10. *A lower semi-continuous function $v : [0, \infty) \times \mathbb{R}_{\geq 0}^{\mathcal{D}_K} \rightarrow \mathbb{R}$ is said to be a viscosity supersolution to (3.31) if, given any $\phi \in C^\infty((0, \infty) \times \mathbb{R}_{> 0}^{\mathcal{D}_K})$ with the property that $v - \phi$ has a local minimum at $(t^*, x^*) \in (0, \infty) \times \mathbb{R}_{> 0}^{\mathcal{D}_K}$,*

$$(\partial_t \phi - \tilde{H}_{b,K,R}(\nabla \phi))(t^*, x^*) \geq 0. \quad (3.41)$$

Definition 3.11. A continuous function $f \in C([0, \infty) \times \mathbb{R}_{\geq 0}^{\mathcal{D}_K})$ is said to be a viscosity solution to (3.31) if it is both a viscosity subsolution and a viscosity supersolution to (3.31).

Given functions $h : \mathbb{R}_{\geq 0}^{\mathcal{D}_K} \rightarrow \mathbb{R}$ and $u : [0, \infty) \times \mathbb{R}_{\geq 0}^{\mathcal{D}_K} \rightarrow \mathbb{R}$, define the semi-norms

$$\|h\|_{\text{Lip},1} = \sup_{x \neq x' \in \mathbb{R}_{\geq 0}^{\mathcal{D}_K}} \frac{|h(x) - h(x')|}{\|x - x'\|_1} \quad \text{and} \quad [u]_0 = \sup_{\substack{t > 0 \\ x \in \mathbb{R}_{\geq 0}^{\mathcal{D}_K}}} \frac{|u(t, x) - u(0, x)|}{t}. \quad (3.42)$$

Introduce the space of functions with Lipschitz initial condition that grow at most linearly in time,

$$\mathfrak{L} = \{u : [0, \infty) \times \mathbb{R}_{\geq 0}^{\mathcal{D}_K} \rightarrow \mathbb{R} \mid u(0, \cdot) \text{ is Lipschitz continuous and } [u]_0 < \infty\}, \quad (3.43)$$

and its subset of uniformly Lipschitz functions,

$$\mathfrak{L}_{\text{unif}} = \left\{ u \in \mathfrak{L} \mid \sup_{t \geq 0} \|u(t, \cdot)\|_{\text{Lip},1} < \infty \right\}. \quad (3.44)$$

Combining Theorem 1.1 with the arguments leading to the second conclusion of Lemma 6.1 in [31] gives the following well-posedness result for the Hamilton-Jacobi equation (3.31).

Proposition 3.12. For every $R > 0$, the Hamilton-Jacobi equation (3.31) admits a unique viscosity solution $\tilde{f}_{b,R}^{(K)} \in \mathfrak{L}_{\text{unif}}$ subject to the initial condition $\tilde{\psi}_b^{(K)}$. Moreover, the solution has its gradient in the closed convex set $\mathcal{K}'_{=1,K}$,

$$\nabla \tilde{f}_{b,R}^{(K)} \in L^\infty([0, \infty) \times \mathbb{R}_{\geq 0}^{\mathcal{D}_K}; \mathcal{K}'_{=1,K}), \quad (3.45)$$

and it satisfies the Lipschitz bound

$$\sup_{t > 0} \|\tilde{f}_{b,R}^{(K)}(t, \cdot)\|_{\text{Lip},1} = \|\tilde{\psi}_b^{(K)}\|_{\text{Lip},1} \leq \|\tilde{\psi}_b\|_{\text{Lip,TV}}. \quad (3.46)$$

In addition to this existence and uniqueness result, it will be important to record the following comparison principle for the projected Hamilton-Jacobi equation (3.31). This comparison principle is a consequence of Corollary A.12 in [31].

Lemma 3.13. If $u, v \in \mathfrak{L}_{\text{unif}}$ are respectively a continuous subsolution and a continuous supersolution to (3.31), then

$$\sup_{\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}^{\mathcal{D}_K}} (u(t, x) - v(t, x)) = \sup_{\mathbb{R}_{\geq 0}^{\mathcal{D}_K}} (u(0, x) - v(0, x)). \quad (3.47)$$

The existence of the limit (3.30) defining $f(t, \mu)$ is a consequence of Theorems 1.2 and 1.4 in [31].

Proposition 3.14. Let $b \in \mathbb{R}$ be such that the kernel \tilde{g}_b defined in (3.23) is positive on $[-1, 1]$, let $\tilde{\psi}_b$ be defined by (3.32), and for each integer $K \geq 1$ and $R > \|\tilde{\psi}_b\|_{\text{Lip,TV}}$, denote by $\tilde{f}_{b,R}^{(K)}$ the unique viscosity solution to the Hamilton-Jacobi equation (3.31) subject to the initial condition $\tilde{\psi}_b^{(K)}$. The limit (3.30) exists, is finite and is independent of R and b . This limit is defined to be the solution to the infinite-dimensional Hamilton-Jacobi equation (1.41).

To establish Theorem 1.1, the idea will be to show that the enriched free energy (1.35) is essentially a viscosity subsolution to the infinite-dimensional Hamilton-Jacobi equation (1.41).

4 The free energy upper bound

In this section, we combine the computations in Section 2 with the arguments in [57] to essentially show that the enriched free energy (1.35) is a viscosity subsolution to the infinite-dimensional Hamilton-Jacobi equation (1.41). This equation is given a precise meaning by Proposition 3.14. As alluded to in Section 2, we will perturb the enriched Hamiltonian (1.34) to enforce the concentration of all the multi-overlaps (2.11) without changing the limit of the associated free energy. The perturbation Hamiltonian we will now describe was introduced in [14] to prove a general multi-overlap concentration result whose finitary version we establish in Appendix C.

Fix an integer K_+ which will be chosen sufficiently large in the course of this section, and write $\lambda = (\lambda_0, \lambda_1, \dots, \lambda_{K_+})$ for a perturbation parameter with $\lambda_0 \in [1/2, 1]$ and $\lambda_k \in [2^{-k-1}, 2^{-k}]$ for $1 \leq k \leq K_+$. Given a sequence (ε_N) with $\varepsilon_N = N^\gamma$ for some $-1/8 < \gamma < 0$ and a standard Gaussian vector $Z_0 = (Z_{0,1}, \dots, Z_{0,N})$ in \mathbb{R}^N , introduce the Gaussian perturbation Hamiltonian

$$H_N^{\text{gauss}}(\sigma, \lambda_0) = \mathcal{H}_0 = \sum_{i \leq N} (\lambda_0 \varepsilon_N \sigma_i^* \sigma_i + \sqrt{\lambda_0 \varepsilon_N} Z_{0,i} \sigma_i) \quad (4.1)$$

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

$$Y^{\text{gauss}} = \sqrt{\lambda_0 \varepsilon_N} \sigma^* + Z_0. \quad (4.2)$$

Notice that

$$1 \geq \varepsilon_N \rightarrow 0 \quad \text{and} \quad N \varepsilon_N \rightarrow \infty. \quad (4.3)$$

Similarly, consider a sequence (s_N) with $s_N = N^\eta$ for $4/5 < \eta < 1$ in such a way that

$$\frac{s_N}{N} \rightarrow 0 \quad \text{and} \quad \frac{s_N}{\sqrt{N}} \rightarrow \infty. \quad (4.4)$$

Fix a sequence of i.i.d. random variables (π_k) with $\text{Poi}(s_N)$ distribution as well as a sequence $e = (e_{jk})$ of random variables with $\text{Exp}(1)$ distribution. For every $j \leq \pi_k$, sample i.i.d. random indices i_{jk} uniformly from the set $\{1, \dots, N\}$, and define the exponential perturbation Hamiltonian by

$$\mathcal{H}_k = \sum_{j \leq \pi_k} \left(\log(1 + \lambda_k \sigma_{i_{jk}}) - \frac{\lambda_k e_{jk} \sigma_{i_{jk}}}{1 + \lambda_k \sigma_{i_{jk}}^*} \right) \quad \text{and} \quad H_N^{\text{exp}}(\sigma) = \sum_{1 \leq k \leq K_+} \mathcal{H}_k. \quad (4.5)$$

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

$$Y_{jk}^{\text{exp}} = \frac{e_{jk}}{1 + \lambda_k \sigma_{i_{jk}}^*} \quad (4.6)$$

for $j \leq \pi_k$ and $k \geq 1$. Introduce the perturbed Hamiltonian

$$H_N(\sigma, \lambda) = H_N^{t, \mu}(\sigma) + H_N^{\text{gauss}}(\sigma, \lambda_0) + H_N^{\text{exp}}(\sigma, \lambda) \quad (4.7)$$

as well as its associated free energy

$$\bar{F}_N^{\text{pert}}(t, \mu, \lambda) = \frac{1}{N} \mathbb{E} \log \int_{\Sigma_N} \exp H_N(\sigma, \lambda) dP_N^*(\sigma). \quad (4.8)$$

Since the Gibbs measure associated with the Hamiltonian in (4.7) is still a conditional expectation as in (2.2), it will still satisfy the Nishimori identity (2.3). An essential property of the perturbation Hamiltonians (4.1) and (4.5) is that they do not affect the asymptotic behavior of the enriched free energy (1.35).

Lemma 4.1. *For every $t > 0$, $\mu \in \mathcal{M}_+$ and λ , the enriched free energy (1.35) and the perturbed free energy (4.8) are asymptotically equivalent,*

$$\lim_{N \rightarrow \infty} |\bar{F}_N^{\text{pert}}(t, \mu, \lambda) - \bar{F}_N(t, \mu)| = 0. \quad (4.9)$$

Proof. A direct computation reveals that

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

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

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

while

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

Since these bounds are uniform in σ , it follows that

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

The third term was obtained by taking the expectation with respect to the randomness of e first and then with respect to the randomness of (π_k) . Leveraging (4.3) and (4.4) to let N tend to infinity completes the proof. \blacksquare

With this result in mind, we abuse notation and redefine the perturbed free energy (4.8),

$$F_N(t, \mu, \lambda) = \frac{1}{N} \log \int_{\Sigma_N} \exp H_N(\sigma, \lambda) dP_N^*(\sigma) \quad \text{and} \quad \bar{F}_N(t, \mu, \lambda) = \mathbb{E} F_N(t, \mu, \lambda). \quad (4.10)$$

For every integer $K \geq 1$ and $x \in \mathbb{R}_{\geq 0}^K$, we denote by

$$F_N^{(K)}(t, x, \lambda) = F_N(t, \mu_x^{(K)}, \lambda) \quad \text{and} \quad \bar{F}_N^{(K)}(t, x, \lambda) = \mathbb{E} F_N^{(K)}(t, x, \lambda) \quad (4.11)$$

the finite-dimensional projections of these perturbed free energy functionals. In the same spirit as (3.30), given $b \in \mathbb{R}$ such that the kernel \tilde{g}_b defined in (3.23) is positive on $[-1, 1]$, introduce translated versions of these free energy functionals,

$$F'_N(t, \mu, \lambda) = F_N(t, \mu, \lambda) + b \int_{-1}^1 d\mu + \frac{bt}{2} \quad \text{and} \quad \bar{F}'_N(t, \mu, \lambda) = \mathbb{E} F'_N(t, \mu, \lambda). \quad (4.12)$$

For every integer $K \geq 1$ and $x \in \mathbb{R}_{\geq 0}^{\mathcal{Q}_K}$, we denote by

$$F_N^{(K)}(t, x, \lambda) = F_N'(t, \mu_x^{(K)}, \lambda) \quad \text{and} \quad \bar{F}_N^{(K)}(t, x, \lambda) = \mathbb{E} F_N^{(K)}(t, x, \lambda) \quad (4.13)$$

the finite-dimensional projections of these modified free energy functionals. Similarly, we write

$$\bar{F}_N^{(K)}(t, x) = \bar{F}_N(t, \mu_x^{(K)}) \quad \text{and} \quad \bar{F}_N^{\prime(K)}(t, x) = \bar{F}_N^{(K)}(t, x) + b\|x\|_1 + \frac{bt}{2} \quad (4.14)$$

for the finite-dimensional projections of the enriched free energy (4.12) and its translation according to (3.30). Combining Lemmas 2.1, 2.4 and 3.1 with the Arzela-Ascoli theorem, it is possible to extract a subsequential limit $\tilde{F}^{(K)}$ from the sequence defined by (4.14) for varying N . Passing to a further subsequence and using a diagonalization argument, it is also possible to ensure that

$$\tilde{F}^{(K)}(t, x) = \limsup_{N \rightarrow \infty} \bar{F}_N^{\prime(K)}(t, x) \quad (4.15)$$

for all $(t, x) \in (0, \infty) \times \mathbb{R}_{\geq 0}^{\mathcal{Q}_K}$. The key to establishing Theorem 1.1 will be to show that, in some sense, the subsequential limit $\tilde{F}^{(K)}$ is an approximate subsolution to the Hamilton-Jacobi equation (3.31) for some $R > \|\tilde{\psi}_b\|_{\text{Lip,TV}} + \|\tilde{g}_b\|_{\infty} + \|\tilde{g}'_b\|_{\infty} + 1$ which will remain fixed throughout this section.

We fix a smooth function $\phi \in C^{\infty}((0, \infty) \times \mathbb{R}_{> 0}^{\mathcal{Q}_K})$ with the property that the difference $\tilde{F}^{(K)} - \phi$ achieves a local maximum at some point $(t_{\infty}, x_{\infty}) \in (0, \infty) \times \mathbb{R}_{> 0}^{\mathcal{Q}_K}$. Recalling that the index K_+ controls the number of terms in the perturbation Hamiltonian (4.5), we introduce the parameter

$$\lambda_{\infty} = \frac{(1, 2^{-1}, 2^{-2}, \dots, 2^{-K_+}) + (2^{-1}, 2^{-2}, 2^{-3}, \dots, 2^{-K_+-1})}{2} \quad (4.16)$$

as well as the smooth function

$$\tilde{\phi}(t, x, \lambda) = \phi(t, x) + (t - t_{\infty})^2 + \|x - x_{\infty}\|_2^2 + \|\lambda - \lambda_{\infty}\|_2^2. \quad (4.17)$$

It is clear that $(t, x, \lambda) \mapsto \tilde{F}^{(K)}(t, x) - \tilde{\phi}(t, x, \lambda)$ has a strict local maximum at $(t_{\infty}, x_{\infty}, \lambda_{\infty})$. Arguing as in the proof of Lemma 4.1 shows that $(t, x, \lambda) \mapsto \bar{F}_N^{\prime(K)}(t, x, \lambda)$ converges to $(t, x, \lambda) \mapsto \tilde{F}^{(K)}(t, x)$ locally uniformly. It is therefore possible to find a sequence (t_N, x_N, λ_N) which converges to the point $(t_{\infty}, x_{\infty}, \lambda_{\infty})$ and has the property that $(t, x, \lambda) \mapsto \bar{F}_N^{\prime(K)}(t, x, \lambda) - \tilde{\phi}(t, x, \lambda)$ attains a local maximum at (t_N, x_N, λ_N) . More precisely, it is possible to find a constant $C < \infty$ which is allowed to depend on $K, K_+, t_{\infty}, x_{\infty}$ and the function ϕ such that

$$\begin{aligned} & \left(\bar{F}_N^{\prime(K)} - \tilde{\phi} \right)(t_N, x_N, \lambda_N) \\ &= \sup \left\{ \left(\bar{F}_N^{\prime(K)} - \tilde{\phi} \right)(t, x, \lambda) \mid |t - t_N| + \|x - x_N\|_2 + \|\lambda - \lambda_N\|_2 \leq C^{-1} \right\}. \end{aligned} \quad (4.18)$$

We will use such a constant $C < \infty$ at various places in this proof, and we understand that its value may need to be increased as we proceed; the important point is that it does not depend on N . The choices of λ_{∞} in (4.16) and $x_{\infty} \in \mathbb{R}_{> 0}^{\mathcal{Q}_K}$ ensure that when N is large enough $(\lambda_N)_k \in (2^{-k-1}, 2^{-k})$ for $0 \leq k \leq K_+$ and $x_N \in \mathbb{R}_{> 0}^{\mathcal{Q}_K}$. Increasing $C < \infty$ if necessary, it is therefore possible to guarantee

that for N large enough the supremum on the right-hand side of (4.18) is taken over triples (t, x, λ) with $t > 0$, $x \in \mathbb{R}_{>0}^{\mathcal{D}_K}$ and $\lambda_k \in [2^{-k-1}, 2^{-k}]$ for $1 \leq k \leq K_+$. It follows that

$$\partial_t \left(\bar{F}_N^{(K)} - \tilde{\phi} \right) (t_N, x_N, \lambda_N) = 0, \quad \nabla_x \left(\bar{F}_N^{(K)} - \tilde{\phi} \right) (t_N, x_N, \lambda_N) = 0 \quad (4.19)$$

and

$$\nabla_\lambda \left(\bar{F}_N^{(K)} - \tilde{\phi} \right) (t_N, x_N, \lambda_N) = \nabla_\lambda \left(\bar{F}_N^{(K)} - \tilde{\phi} \right) (t_N, x_N, \lambda_N) = 0. \quad (4.20)$$

The majority of this section will be devoted to using the second equality in (4.20) in conjunction with the main result in Appendix C to show the concentration of a finite but very large number of the multi-overlaps (2.11). We will then combine this finitary multi-overlap concentration result with the computations in Section 2 to establish the following crucial result.

Lemma 4.2. *Fix $R > \|\tilde{\psi}_b\|_{\text{Lip,TV}} + \|\tilde{g}_b\|_\infty + \|\tilde{g}'_b\|_\infty + 1$. For every $\varepsilon > 0$, there exists a choice of integer $K_+ \geq 1$ in the perturbed Hamiltonian (4.7) with the property that for any integer $K \geq 1$, it is possible to find a constant $\mathcal{E}_{\varepsilon,K}$ with*

$$\limsup_{N \rightarrow \infty} \left| \left(\partial_t \bar{F}_N^{(K)} - \tilde{\mathbb{H}}_{b,K,R} \left(\nabla_x \bar{F}_N^{(K)} \right) \right) (t_N, x_N, \lambda_N) \right| \leq \mathcal{E}_{\varepsilon,K} \quad (4.21)$$

and $\lim_{\varepsilon \rightarrow 0} \lim_{K \rightarrow \infty} \mathcal{E}_{\varepsilon,K} = 0$.

For the time being, let us prove Theorem 1.1 assuming Lemma 4.2.

Proof of Theorem 1.1 assuming Lemma 4.2. Given $\varepsilon > 0$, invoke Lemma 4.2 to find an integer $K_+ \geq 1$ in the perturbed Hamiltonian (4.7) with the property that for any integer $K \geq 1$, it is possible to find a constant $\mathcal{E}_{\varepsilon,K}$ with

$$\limsup_{N \rightarrow \infty} \left| \left(\partial_t \bar{F}_N^{(K)} - \tilde{\mathbb{H}}_{b,K,R} \left(\nabla_x \bar{F}_N^{(K)} \right) \right) (t_N, x_N, \lambda_N) \right| \leq \mathcal{E}_{\varepsilon,K} \quad (4.22)$$

and $\lim_{\varepsilon \rightarrow 0} \lim_{K \rightarrow \infty} \mathcal{E}_{\varepsilon,K} = 0$. Given an integer $K \geq 1$, the idea will be to show that the test function $\phi \in C^\infty((0, \infty) \times \mathbb{R}_{>0}^{\mathcal{D}_K})$ introduced above satisfies the subsolution condition in (3.40) for the non-linearity $\mathbb{H}_{b,K,R}$ at the point of contact (t_∞, x_∞) up to the small error $\mathcal{E}_{\varepsilon,K}$. This will mean that the subsequential limit $\tilde{F}^{(K)}$ of the modified free energy (4.14) is a viscosity subsolution to the Hamilton-Jacobi equation (3.31) up to a small error. More precisely, the function

$$\tilde{F}_\varepsilon^{(K)} = \tilde{F}^{(K)} - t \mathcal{E}_{\varepsilon,K} \quad (4.23)$$

will be a viscosity subsolution to (3.31). This observation will allow us to leverage the comparison principle in Lemma 3.13 to bound the limit superior of the enriched free energy (1.35) by the solution f to the infinite-dimensional Hamilton-Jacobi equation (1.41) constructed in Proposition 3.14. We proceed in two steps.

Step 1: $\tilde{F}_\varepsilon^{(K)}$ subsolution.

Since $x_N \rightarrow x_\infty$ assume without loss of generality that $(x_N) \subset \mathbb{R}_{>0}^{\mathcal{D}_K}$. It follows by (4.19) that

$$\left(\partial_t \tilde{\phi} - \tilde{\mathbb{H}}_{b,K,R} \left(\nabla_x \tilde{\phi} \right) \right) (t_N, x_N, \lambda_N) = \left(\partial_t \bar{F}_N^{(K)} - \tilde{\mathbb{H}}_{b,K,R} \left(\nabla_x \bar{F}_N^{(K)} \right) \right) (t_N, x_N, \lambda_N),$$

so letting N tend to infinity and combining the definition of $\tilde{\phi}$ with (4.22) yields

$$(\partial_t \tilde{\phi} - \tilde{H}_{b,K,R}(\nabla_x \tilde{\phi}))(t_\infty, x_\infty) \leq \mathcal{E}_{\varepsilon,K}.$$

This shows that the function (4.23) satisfies the subsolution condition in (3.40).

Step 2: comparison principle.

The comparison principle in Lemma 3.13 gives the upper bound

$$\tilde{F}^{(K)}(t, x) \leq \tilde{f}_{b,R}^{(K)}(t, x) + t \mathcal{E}_{\varepsilon,K}. \quad (4.24)$$

We have implicitly used that $\tilde{F}^{(K)}$ and $\tilde{f}_{b,R}^{(K)}$ are continuous and have the same initial condition by Proposition 3.6. We have also used that they both belong to the solution space $\mathfrak{L}_{\text{unif}}$ by Proposition 3.12, (2.1), (2.4) and a simple application of the mean value theorem. With (4.24) in mind, fix a finite measure $\mu \in \mathcal{M}_+$, and introduce the discrete measure

$$\mu^{(K)} = \mu_{x^{(K)}(\mu)}$$

defined in (3.19). It is readily verified that $\bar{\mu}^{(K)} \rightarrow \bar{\mu}$ with respect to the Wasserstein distance (3.9). Moreover, an identical argument to that in Lemma 3.4 leveraging the second bound in (3.13) reveals that

$$\bar{F}_N(t, \mu) \leq C' \mu[-1, 1] W(\bar{\mu}, \bar{\mu}^{(K)}) + \bar{F}_N^{(K)}(t, x^{(K)}(\mu))$$

for some constant C' that depends only on c . We use the letter C' instead of C to emphasize that the constant C' does not depend on K . Letting N tend to infinity, recalling (4.14), (4.15) and leveraging (4.24) yields

$$\limsup_{N \rightarrow \infty} \bar{F}_N(t, \mu) \leq C' \mu[-1, 1] W(\bar{\mu}, \bar{\mu}^{(K)}) + \tilde{f}_{b,R}^{(K)}(t, x^{(K)}(\mu)) - b \|x^{(K)}(\mu)\|_1 - \frac{bt}{2} + t \mathcal{E}_{\varepsilon,K}.$$

Invoking Lemma 4.2 and Proposition 3.14 to let K tend to infinity and then ε tend to zero completes the proof. \blacksquare

The rest of this section is devoted to the proof of Lemma 4.2 which will be obtained by combining the computations of Section 2 with the main result of Appendix C to show the concentration of a finite but very large number of the multi-overlaps (2.11). In the notation of [14], for any perturbation parameter λ , let

$$\lambda_{0,N} = \varepsilon_N \lambda_0, \quad (4.25)$$

and introduce the quantities

$$\mathcal{L}_0 = \frac{\mathcal{H}'_0}{N \varepsilon_N}, \quad \text{where} \quad \mathcal{H}'_0 = \partial_{\lambda_0} H_N^{\text{gauss}}(\sigma, \lambda_0) = \varepsilon_N \left(\sigma \cdot \sigma^* + \frac{\sigma \cdot Z_0}{2\sqrt{\lambda_{0,N}}} \right), \quad (4.26)$$

$$\mathcal{L}_k = \frac{\mathcal{H}'_k}{s_N}, \quad \text{where} \quad \mathcal{H}'_k = \partial_{\lambda_k} H_N^{\text{exp}}(\sigma, \lambda) = \sum_{j \leq \pi_k} \sigma_{ijk} \left(\frac{1}{1 + \lambda_k \sigma_{ijk}} - \frac{e^{jk}}{(1 + \lambda_k \sigma_{ijk}^*)^2} \right) \quad (4.27)$$

for $1 \leq k \leq K_+$. The importance of these quantities stems from the fact that

$$\partial_{\lambda_0} F_N^{(K)}(t, x, \lambda) = \frac{1}{N} \langle \mathcal{H}'_0 \rangle, \quad \partial_{\lambda_0}^2 F_N^{(K)}(t, x, \lambda) = \frac{1}{N} \left(\langle (\mathcal{H}'_0 - \langle \mathcal{H}'_0 \rangle)^2 \rangle - \frac{\varepsilon_N^2}{4\lambda_{0,N}^{3/2}} \langle \sigma \rangle \cdot Z_0 \right), \quad (4.28)$$

$$\partial_{\lambda_k} F_N^{(K)}(t, x, \lambda) = \frac{1}{N} \langle \mathcal{H}'_k \rangle, \quad \partial_{\lambda_k}^2 F_N^{(K)}(t, x, \lambda) = \frac{1}{N} \left(\langle (\mathcal{H}'_k - \langle \mathcal{H}'_k \rangle)^2 \rangle + \langle \mathcal{H}''_k \rangle \right) \quad (4.29)$$

for $1 \leq k \leq K_+$ while

$$\partial_{\lambda_k \lambda_j} F_N^{(K)}(t, x, \lambda) = \frac{1}{N} \left(\langle \mathcal{H}'_j \mathcal{H}'_k \rangle - \langle \mathcal{H}'_j \rangle \langle \mathcal{H}'_k \rangle \right) \quad (4.30)$$

for $0 \leq j \neq k \leq K_+$. Here, and for the remainder of this section, the Gibbs average $\langle \cdot \rangle$ will always be associated with the perturbed Hamiltonian (4.7) evaluated at a triple (t, x, λ) which will be clear from the context. It will also be convenient to record that for $1 \leq k \leq K_+$,

$$\mathcal{H}''_k = \partial_{\lambda_k}^2 H_N^{\text{exp}}(\sigma, \lambda) = \sum_{j \leq \pi_k} \left(-\frac{1}{(1 + \lambda_k \sigma_{ijk})^2} + 2 \frac{\sigma_{ijk} \sigma_{ijk}^* e_{jk}}{(1 + \lambda_k \sigma_{ijk}^*)^3} \right) \text{ and } |\mathbb{E} \langle \mathcal{H}''_k \rangle| \leq C s_N. \quad (4.31)$$

To obtain the concentration of the multi-overlaps (2.11) we will show the concentration (C.8) of the quantities \mathcal{L}_k for the Gibbs measure with parameters given by the contact point (t_N, x_N, λ_N) . This concentration will be deduced from the fact that the averaged free energy is being ‘‘touched from above’’ by a smooth function at the contact point, thereby constraining its Hessian at this point, together with the concentration of the free energy $F_N^{(K)}$ about its average $\bar{F}_N^{(K)}$. Due to the constraint on the Hessian at the contact point, we will be able to extend the concentration result on the free energy into an estimate on the concentration of its gradient. We decompose this argument into a series of four lemmas: the first two essentially bound the Hessian of the perturbed free energy (4.11) from above and from below; the third leverages the free energy concentration result in Appendix B to estimate the uniform L^p -distance between the quenched and averaged free energies (4.11); while the fourth extends this to a control on the gradient of the free energy.

Lemma 4.3. *For any perturbation parameter λ with $\|\lambda\|_2 \leq C^{-1}$,*

$$\bar{F}_N^{(K)}(t_N, x_N, \lambda_N + \lambda) - \bar{F}_N^{(K)}(t_N, x_N, \lambda_N) - \lambda \cdot \nabla_\lambda \bar{F}_N^{(K)}(t_N, x_N, \lambda_N) \leq C \|\lambda\|_2^2. \quad (4.32)$$

Proof. Fix a perturbation parameter λ with $\|\lambda\|_2 \leq C^{-1}$, and notice that (4.18) gives

$$\bar{F}_N^{(K)}(t_N, x_N, \lambda_N + \lambda) - \bar{F}_N^{(K)}(t_N, x_N, \lambda_N) \leq \tilde{\phi}(t_N, x_N, \lambda_N + \lambda) - \tilde{\phi}(t_N, x_N, \lambda_N).$$

On the other hand, Taylor’s formula with integral remainder implies that

$$\begin{aligned} & \bar{F}_N^{(K)}(t_N, x_N, \lambda_N + \lambda) - \bar{F}_N^{(K)}(t_N, x_N, \lambda_N) \\ &= \lambda \cdot \nabla_\lambda \bar{F}_N^{(K)}(t_N, x_N, \lambda_N) + \int_0^1 (1-s) \lambda \cdot \nabla_\lambda^2 \bar{F}_N^{(K)}(t_N, x_N, \lambda_N + s\lambda) \lambda \, ds, \end{aligned} \quad (4.33)$$

and similarly,

$$\begin{aligned} & \tilde{\phi}(t_N, x_N, \lambda_N + \lambda) - \tilde{\phi}(t_N, x_N, \lambda_N) \\ &= \lambda \cdot \nabla_\lambda \tilde{\phi}(t_N, x_N, \lambda_N) + \int_0^1 (1-s) \lambda \cdot \nabla_\lambda^2 \tilde{\phi}(t_N, x_N, \lambda_N + s\lambda) \lambda \, ds. \end{aligned}$$

Combining (4.19) with the chain rule shows that $\lambda \cdot \nabla_\lambda \tilde{\phi}(t_N, x_N, \lambda_N) = \lambda \cdot \nabla_\lambda \bar{F}_N^{(K)}(t_N, x_N, \lambda_N)$, and therefore

$$\begin{aligned} \int_0^1 (1-s) \lambda \cdot \nabla_\lambda^2 \bar{F}_N^{(K)}(t_N, x_N, \lambda_N + s\lambda) \lambda \, ds &\leq \int_0^1 (1-s) \lambda \cdot \nabla_\lambda^2 \tilde{\phi}(t_N, x_N, \lambda_N + s\lambda) \lambda \, ds \\ &\leq C \|\lambda\|_2^2. \end{aligned}$$

Substituting this into (4.33) gives

$$\begin{aligned} \bar{F}_N^{(K)}(t_N, x_N, \lambda_N + \lambda) - \bar{F}_N^{(K)}(t_N, x_N, \lambda_N) &= \bar{F}_N^{(K)'}(t_N, x_N, \lambda_N + \lambda) - \bar{F}_N^{(K)'}(t_N, x_N, \lambda_N) \\ &\leq \lambda \cdot \nabla_\lambda \bar{F}_N^{(K)'}(t_N, x_N, \lambda_N) + C \|\lambda\|_2^2 \\ &= \lambda \cdot \nabla_\lambda \bar{F}_N^{(K)}(t_N, x_N, \lambda_N) + C \|\lambda\|_2^2. \end{aligned}$$

This completes the proof. ■

Lemma 4.4. *There exists a random variable X with $\mathbb{E}X^2 \leq C$ such that, for all perturbation parameters λ with $\|\lambda\|_2 \leq C^{-1}$,*

$$F_N^{(K)}(t_N, x_N, \lambda_N + \lambda) - F_N^{(K)}(t_N, x_N, \lambda_N) - \lambda \cdot \nabla_\lambda F_N^{(K)}(t_N, x_N, \lambda_N) \geq -X \|\lambda\|_2^2. \quad (4.34)$$

Proof. Since t_N and x_N remain fixed throughout, write $F_N^{(K)}(\lambda)$ for $F_N^{(K)}(t_N, x_N, \lambda)$. Introduce the function

$$h(\lambda) = F_N^{(K)}(\lambda) - \frac{\sqrt{\lambda_{0,N}}}{N} \sum_{i \leq N} |Z_{0,i}| + \frac{1}{N} \sum_{1 \leq k \leq K_+} \sum_{j \leq \pi_k} (8\lambda_k^2 e_{jk} - \log(1 - \lambda_k)). \quad (4.35)$$

Leveraging (4.28) and Hölder's inequality, one can see that

$$\partial_{\lambda_0}^2 h(\lambda) = \frac{1}{N} \langle (\mathcal{H}'_0 - \langle \mathcal{H}'_0 \rangle)^2 \rangle - \frac{\varepsilon_N^2}{4N\lambda_{0,N}^{3/2}} \langle \sigma \rangle \cdot Z_0 + \frac{\varepsilon_N^2}{4N\lambda_{0,N}^{3/2}} \sum_{i \leq N} |Z_{0,i}| \geq \frac{1}{N} \langle (\mathcal{H}'_0 - \langle \mathcal{H}'_0 \rangle)^2 \rangle.$$

Using (4.29) and (4.31) reveals that for $1 \leq k \leq K_+$,

$$\begin{aligned} \partial_{\lambda_k}^2 h(\lambda) &= \frac{1}{N} \langle (\mathcal{H}'_k - \langle \mathcal{H}'_k \rangle)^2 \rangle \\ &\quad + \frac{1}{N} \left\langle \sum_{j \leq \pi_k} \left(-\frac{1}{(1 + \lambda_k \sigma_{i_{jk}})^2} + 2 \frac{\sigma_{i_{jk}} \sigma_{i_{jk}}^* e_{jk}}{(1 + \lambda_k \sigma_{i_{jk}}^*)^3} + 16e_{jk} + \frac{1}{(1 - \lambda_k)^2} \right) \right\rangle. \end{aligned}$$

Since $\lambda_k \leq 1/2$ and all spin configuration coordinates are bounded by one, it is actually the case that

$$\partial_{\lambda_k}^2 h(\lambda) \geq \frac{1}{N} \langle (\mathcal{H}'_k - \langle \mathcal{H}'_k \rangle)^2 \rangle.$$

Together with (4.30), this shows that $\nabla_{\lambda}^2 h$ is positive definite and therefore h is convex. It follows that for any perturbation parameter λ with $\|\lambda\|_2 \leq C^{-1}$,

$$h(\lambda_N + \lambda) \geq h(\lambda_N) + \lambda \cdot \nabla_{\lambda} h(\lambda_N).$$

Remembering the definition of h in (4.35), this shows that the left-hand side of (4.34) is bounded from below by

$$\begin{aligned} & \frac{1}{N} \left(\sqrt{(\lambda_N)_{0,N} + \lambda_{0,N}} - \sqrt{(\lambda_N)_{0,N}} - \frac{\lambda_{0,N}}{2\sqrt{(\lambda_N)_{0,N}}} \right) \sum_{i \leq N} |Z_{0,i}| \\ & + \frac{1}{N} \sum_{1 \leq k \leq K_+} \sum_{j \leq \pi_k} 8e_{jk} \left((\lambda_N)_k^2 + 2\lambda_k(\lambda_N)_k - ((\lambda_N)_k + \lambda_k)^2 \right) \\ & + \frac{1}{N} \sum_{1 \leq k \leq K_+} \sum_{j \leq \pi_k} \left(\log \left(\frac{1 - (\lambda_N)_k - \lambda_k}{1 - (\lambda_N)_k} \right) + \frac{\lambda_k}{1 - (\lambda_N)_k} \right). \end{aligned}$$

Increasing C if necessary, Taylor's theorem with differential remainder gives a perturbation parameter $\tilde{\lambda}$ with $\tilde{\lambda}_k \in [2^{-k-1}, 2^{-k}]$ for $0 \leq k \leq K_+$ whose value might not be the same at each occurrence such that

$$\begin{aligned} & \sqrt{(\lambda_N)_{0,N} + \lambda_{0,N}} - \sqrt{(\lambda_N)_{0,N}} - \frac{\lambda_{0,N}}{2\sqrt{(\lambda_N)_{0,N}}} = -\frac{(\lambda_{0,N})^2}{8\tilde{\lambda}_{0,N}^{3/2}} \geq -\sqrt{\varepsilon_N} \lambda_0^2 \geq -\lambda_0^2 \\ & (\lambda_N)_k^2 + 2\lambda_k(\lambda_N)_k - ((\lambda_N)_k + \lambda_k)^2 = -\lambda_k^2 \\ & \log \left(\frac{1 - (\lambda_N)_k - \lambda_k}{1 - (\lambda_N)_k} \right) + \frac{\lambda_k}{1 - (\lambda_N)_k} = -\frac{\lambda_k^2}{2(1 - \tilde{\lambda}_k)^2} \geq -2\lambda_k^2. \end{aligned}$$

It follows that the left-hand side of (4.34) is bounded from below by

$$-\frac{\lambda_0^2}{N} \sum_{i \leq N} |Z_{0,i}| - \frac{1}{N} \sum_{1 \leq k \leq K_+} \lambda_k^2 \sum_{j \leq \pi_k} (8e_{jk} + 2) \geq -X \|\lambda\|_2^2$$

for the random variable

$$X = \frac{1}{N} \sum_{i \leq N} |Z_{0,i}| + \frac{1}{N} \sum_{1 \leq k \leq K_+} \sum_{j \leq \pi_k} (8e_{jk} + 2).$$

Using the Cauchy-Schwarz inequality, taking the average with respect to the randomness of (e_{jk}) before the average with respect to the randomness of (π_k) and remembering (4.4) shows that

$$\begin{aligned} \mathbb{E} X^2 & \leq \frac{C}{N^2} \left(\mathbb{E} \left(\sum_{i \leq N} |Z_{0,i}| \right)^2 + \sum_{1 \leq k \leq K_+} \mathbb{E} \left(\sum_{j \leq \pi_k} (8e_{jk} + 2) \right)^2 \right) \\ & \leq \frac{C}{N^2} \left(N \mathbb{E} |Z_{0,1}| + (N^2 - N) \mathbb{E} |Z_{0,1} Z_{0,2}| + \sum_{1 \leq k \leq K_+} \mathbb{E} \pi_k \sum_{j \leq \pi_k} (8e_{jk} + 2)^2 \right) \\ & \leq \frac{C}{N^2} (N^2 + s_N^2 + s_N) \leq C. \end{aligned}$$

This completes the proof. ■

Lemma 4.5. For every $M > 0$ small enough, $p \in [1, \infty)$ and $\varepsilon > 0$, there exists $C < \infty$ not depending on N such that

$$\left(\mathbb{E} \sup_{\|\lambda\|_\infty \leq M} \left| \left(F_N^{(K)} - \bar{F}_N^{(K)} \right) (t_N, x_N, \lambda_N + \lambda) \right|^p \right)^{\frac{1}{p}} \leq CN^{-\frac{1}{2} + \varepsilon}. \quad (4.36)$$

Proof. Let $0 < M < 1/2$ be small enough so that $(\lambda_N + \lambda)_k \in [2^{-k-1}, 2^{-k}]$ for $0 \leq k \leq K_+$ whenever $\|\lambda\|_\infty \leq M$, and for each perturbation parameter λ introduce the random variable

$$Y(\lambda) = \frac{1}{N} \sum_{0 \leq k \leq K_+} |\langle \mathcal{H}'_k \rangle|,$$

where the Gibbs average is associated with the perturbed Hamiltonian (4.7) evaluated at the triple $(t_N, x_N, \lambda_N + \lambda)$. The relevance of these random variables stems from the fact that by the mean value theorem, (4.28) and (4.29), for every λ, λ' in the ℓ^∞ -ball of radius M ,

$$\left| F_N^{(K)}(t_N, x_N, \lambda_N + \lambda) - F_N^{(K)}(t_N, x_N, \lambda_N + \lambda') \right| \leq C \sup_{\|\eta\|_\infty \leq M} Y(\eta) \|\lambda - \lambda'\|_1.$$

Averaging this inequality also shows that for every λ, λ' in the ℓ^∞ -ball of radius M ,

$$\left| \bar{F}_N^{(K)}(t_N, x_N, \lambda_N + \lambda) - \bar{F}_N^{(K)}(t_N, x_N, \lambda_N + \lambda') \right| \leq \mathbb{E} \sup_{\|\eta\|_\infty \leq M} Y(\eta) \|\lambda - \lambda'\|_1$$

These two bounds imply that for any even integer $q \geq 2$,

$$\begin{aligned} \mathbb{E} \sup_{\|\lambda\|_\infty \leq M} \left| \left(F_N^{(K)} - \bar{F}_N^{(K)} \right) (t_N, x_N, \lambda_N + \lambda) \right|^q &\leq \mathbb{E} \sup_{\lambda \in A_\varepsilon} \left| \left(F_N^{(K)} - \bar{F}_N^{(K)} \right) (t_N, x_N, \lambda_N + \lambda) \right|^q \\ &\quad + C\varepsilon^q \mathbb{E} \sup_{\|\lambda\|_\infty \leq M} Y(\lambda)^q \end{aligned}$$

for $A_\varepsilon = \varepsilon\mathbb{Z}^{1+K_+} \cap \{\|\lambda\|_\infty \leq M\}$. Indeed, every λ is at most at distance $\varepsilon(K_+ + 1)$ from an element in A_ε with respect to the ℓ^1 -norm. Bounding the supremum over A_ε by the sum over A_ε and invoking the free energy concentration result in Proposition B.4 shows that

$$\mathbb{E} \sup_{\|\lambda\|_\infty \leq M} \left| \left(F_N^{(K)} - \bar{F}_N^{(K)} \right) (t_N, x_N, \lambda_N + \lambda) \right|^q \leq C|A_\varepsilon|N^{-\frac{q}{2}} + C\varepsilon^q \mathbb{E} \sup_{\|\lambda\|_\infty \leq M} |Y(\lambda)|^q. \quad (4.37)$$

To bound the moments of $\sup_{\|\lambda\|_\infty \leq M} |Y(\lambda)|$ fix $1 \leq k \leq K_+$. Hölder's inequality and (4.27) reveal that

$$\begin{aligned} \mathbb{E} \sup_{\|\lambda\|_\infty \leq M} |\langle \mathcal{H}'_k \rangle|^q &\leq \mathbb{E} \sup_{\|\lambda\|_\infty \leq M} \left| \sum_{j \leq \pi_k} \frac{1}{1 - (\lambda_N + \lambda)_k} + \frac{e_{jk}}{(1 - (\lambda_N + \lambda)_k)^2} \right|^q \\ &\leq \mathbb{E} \left| \sum_{j \leq \pi_k} \frac{4(1 + e_{jk})}{(1 - 2M)^2} \right|^q \leq \mathbb{E} \pi_k^{q-1} \sum_{j \leq \pi_k} \sum_{j \leq \pi_k} \left| \frac{4(1 + e_{jk})}{(1 - 2M)^2} \right|^q \\ &\leq C \mathbb{E} \pi_k^q, \end{aligned}$$

where the last inequality is found by averaging over the randomness of (e_{jk}) . Similarly (4.26) and Hölder's inequality give

$$\begin{aligned} \mathbb{E} \sup_{\|\lambda\|_\infty \leq M} |\langle \mathcal{H}_0' \rangle|^q &\leq \varepsilon_N^q \mathbb{E} \sup_{\|\lambda\|_\infty \leq M} \left| |\sigma \cdot \sigma^*| + \frac{|\sigma \cdot Z_0|}{2\varepsilon_N^{q/2} ((\lambda_N)_0 + \lambda_0)^{q/2}} \right|^q \\ &\leq C\varepsilon_N^q \left(N^q + \frac{2^q \mathbb{E} |\sigma \cdot Z_0|^q}{2\varepsilon_N^{q/2} (1-2M)^q} \right) \leq C\varepsilon_N^{\frac{q}{2}} N^q. \end{aligned}$$

Combining these two inequalities with Lemma B.2 and recalling the choices (4.3) and (4.4) of ε_N and s_N shows that

$$\mathbb{E} \sup_{\|\lambda\|_\infty \leq M} |Y(\lambda)|^q \leq C. \quad (4.38)$$

Substituting this into (4.37) and noticing that $|A_\varepsilon|$ is of order $\varepsilon^{-(K_++1)}$ yields

$$\mathbb{E} \sup_{\|\lambda\|_\infty \leq M} \left| \left(F_N^{(K)} - \bar{F}_N^{(K)} \right) (t_N, x_N, \lambda_N + \lambda) \right|^q \leq C(\varepsilon^{-(K_++1)} N^{-\frac{q}{2}} + \varepsilon^q).$$

Taking $1/q$ 'th powers and choosing $\varepsilon = N^{-\frac{q}{2(q+K_++1)}}$ gives

$$\left(\mathbb{E} \sup_{\|\lambda\|_\infty \leq M} \left| \left(F_N^{(K)} - \bar{F}_N^{(K)} \right) (t_N, x_N, \lambda_N + \lambda) \right|^q \right)^{\frac{1}{q}} \leq CN^{-\frac{q}{2(q+K_++1)}}.$$

Notice that the power on the right-hand side can be made arbitrarily close to $-\frac{1}{2}$ by taking q large enough. Invoking Jensen's inequality completes the proof. \blacksquare

Lemma 4.6. *For every $\varepsilon > 0$, there exists a constant $C < \infty$ not depending on N such that*

$$\mathbb{E} \left\| \nabla_\lambda \left(F_N^{(K)} - \bar{F}_N^{(K)} \right) (t_N, x_N, \lambda_N) \right\|_2^2 \leq CN^{-\frac{1}{2} + \varepsilon}. \quad (4.39)$$

Proof. Given $\mu \in [0, C^{-1}]$, consider the random perturbation parameter

$$\lambda = \mu \cdot \frac{\nabla_\lambda \left(F_N^{(K)} - \bar{F}_N^{(K)} \right) (t_N, x_N, \lambda_N)}{\left\| \nabla_\lambda \left(F_N^{(K)} - \bar{F}_N^{(K)} \right) (t_N, x_N, \lambda_N) \right\|_2}.$$

Combining Lemma 4.3 with Lemma 4.4 shows that

$$\begin{aligned} \left(F_N^{(K)} - \bar{F}_N^{(K)} \right) (t_N, x_N, \lambda_N + \lambda) - \left(F_N^{(K)} - \bar{F}_N^{(K)} \right) (t_N, x_N, \lambda_N) \\ \geq \mu \left\| \nabla_\lambda \left(F_N^{(K)} - \bar{F}_N^{(K)} \right) (t_N, x_N, \lambda_N) \right\|_2 - (C + X) \|\lambda\|_2^2. \end{aligned}$$

Rearranging, squaring and taking expectations yields

$$\mu^2 \mathbb{E} \left\| \nabla_\lambda \left(F_N^{(K)} - \bar{F}_N^{(K)} \right) (t_N, x_N, \lambda_N) \right\|_2^2 \leq C \left(\mathbb{E} \sup_{\|\lambda\|_\infty \leq C^{-1}} \left| \left(F_N^{(K)} - \bar{F}_N^{(K)} \right) (t_N, x_N, \lambda_N + \lambda) \right|^2 + \mu^4 \right),$$

where we have used the fact that $\mathbb{E}X^2 \leq C$ and $\|\lambda\|_2 = \mu$. Invoking Lemma 4.5 gives

$$\mathbb{E} \left\| \nabla_{\lambda} \left(F_N^{(K)} - \bar{F}_N^{(K)} \right) (t_N, x_N, \lambda_N) \right\|_2^2 \leq C \left(\frac{1}{N^{1-2\varepsilon} \mu^2} + \mu^2 \right).$$

Optimizing over μ leads to the choice $\mu = N^{-\frac{1}{4} + \frac{\varepsilon}{2}}$ and completes the proof. \blacksquare

Lemma 4.7. *For any $1 \leq k \leq K_+$, there exists a constant $C < \infty$ not depending on N such that*

$$\mathbb{E} \langle (\mathcal{L}_0 - \mathbb{E} \langle \mathcal{L}_0 \rangle)^2 \rangle \leq CN^{-\frac{1}{4}} \quad \text{and} \quad \mathbb{E} \langle (\mathcal{L}_k - \mathbb{E} \langle \mathcal{L}_k \rangle)^2 \rangle \leq CN^{-\frac{1}{20}}. \quad (4.40)$$

Here, the Gibbs average $\langle \cdot \rangle$ is associated with the perturbed Hamiltonian (4.7) evaluated at the contact point (t_N, x_N, λ_N) .

Proof. A direct computation using (4.28) shows that

$$\begin{aligned} N^2 \varepsilon_N^2 \mathbb{E} \langle (\mathcal{L}_0 - \mathbb{E} \langle \mathcal{L}_0 \rangle)^2 \rangle &= \mathbb{E} \langle (\mathcal{H}'_0 - \langle \mathcal{H}'_0 \rangle)^2 \rangle + \mathbb{E} \left(\langle \mathcal{H}'_0 \rangle - \mathbb{E} \langle \mathcal{H}'_0 \rangle \right)^2 \\ &= N \partial_{\lambda_0}^2 \bar{F}_N^{(K)} (t_N, x_N, \lambda_N) + \frac{\varepsilon_N^2}{4 \lambda_{0,N}^{3/2}} \mathbb{E} \langle \sigma \rangle \cdot Z_0 \\ &\quad + N^2 \mathbb{E} \left(\partial_{\lambda_0} \left(F_N^{(K)} - \bar{F}_N^{(K)} \right) (t_N, x_N, \lambda_N) \right)^2. \end{aligned}$$

It follows by Lemma 4.3 and Lemma 4.6 that for any $\varepsilon > 0$,

$$\mathbb{E} \langle (\mathcal{L}_0 - \mathbb{E} \langle \mathcal{L}_0 \rangle)^2 \rangle \leq \frac{C}{N^2 \varepsilon_N^2} \left(N + N \varepsilon_N^{2-\frac{3}{2}} + N^{2-\frac{1}{2}+\varepsilon} \right) = C \left(N^{2|\gamma|-1} + N^{\frac{3}{2}|\gamma|-1} + N^{2|\gamma|+\varepsilon-\frac{1}{2}} \right).$$

Remembering that $-1/8 < \gamma < 0$ gives the first bound in (4.40). To establish the second bound, fix $1 \leq k \leq K_+$. A direct computation using (4.29) yields

$$\begin{aligned} s_N^2 \mathbb{E} \langle (\mathcal{L}_k - \mathbb{E} \langle \mathcal{L}_k \rangle)^2 \rangle &= \mathbb{E} \langle (\mathcal{H}'_k - \langle \mathcal{H}'_k \rangle)^2 \rangle + \mathbb{E} \left(\langle \mathcal{H}'_k \rangle - \mathbb{E} \langle \mathcal{H}'_k \rangle \right)^2 \\ &= N \partial_{\lambda_k}^2 \bar{F}_N^{(K)} (t_N, x_N, \lambda_N) - \mathbb{E} \langle \mathcal{H}''_k \rangle \\ &\quad + N^2 \mathbb{E} \left(\partial_{\lambda_k} \left(F_N^{(K)} - \bar{F}_N^{(K)} \right) (t_N, x_N, \lambda_N) \right)^2. \end{aligned}$$

Invoking (4.31), Lemma 4.3 and Lemma 4.6 reveals that for any $\varepsilon > 0$,

$$\mathbb{E} \langle (\mathcal{L}_k - \mathbb{E} \langle \mathcal{L}_k \rangle)^2 \rangle \leq \frac{C}{s_N^2} \left(N + s_N + N^{\frac{3}{2}+\varepsilon} \right) = C \left(N^{1-2\eta} + N^{-\eta} + N^{\frac{3}{2}+\varepsilon-2\eta} \right).$$

Choosing $\varepsilon = 1/20$, and recalling that $-1/8 < \gamma < 0$ and $4/5 < \eta < 1$ completes the proof. \blacksquare

This result implies the fundamental assumption (C.8) in Appendix C. Combining this with Lemma C.1, Lemma C.2 and Proposition C.3 and fixing $\varepsilon > 0$, it is possible to find $\delta > 0$ so the statement of Proposition C.6 holds. In particular, setting $K_+ = \lfloor \delta^{-1} \rfloor$ in the perturbed Hamiltonian (4.7) ensures that

$$\mathbb{E} \langle (\mathcal{R}_{[m]} - \mathbb{E} \langle \mathcal{R}_{[m]} \rangle)^2 \rangle \leq \varepsilon \quad (4.41)$$

for $1 \leq m \leq \lfloor \varepsilon^{-1} \rfloor$. Together with the computations in Section 2, this multi-overlap concentration allows us to finally give a proof of Lemma 4.2.

Proof of Lemma 4.2. To alleviate notation, we always implicitly assume that $\bar{F}_N^{(K)}$ and its derivatives are evaluated at the contact point (t_N, x_N, λ_N) . The definition of the modified free energy in (4.12) and Corollary 2.2 reveal that

$$\partial_t \bar{F}_N^{(K)} = \frac{1}{2}(c + \Delta \bar{m}^2) \log(c) + \frac{c}{2} \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} \mathbb{E} \langle R_{[n]}^2 \rangle - \frac{c}{2} + \frac{b}{2} + \mathcal{O}(N^{-1}). \quad (4.42)$$

On the other hand, the duality relation (3.22), the definition of the modified free energy in (4.12) and Corollary 2.5 imply that for any $k \in \mathcal{D}_K$,

$$\partial_{x_k} \bar{F}_N^{(K)} = \frac{1}{|\mathcal{D}_K|} \left((c + \Delta \bar{m}k) \log(c) + c \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} \mathbb{E} \langle R_{[n]} \rangle k^n - c + b \right) + \mathcal{O}(N^{-1}).$$

If we denote by $\mu^* = \mathcal{L}(\langle \sigma_i \rangle)$ the law of the Gibbs average of a uniformly sampled spin coordinate, then the Nishimori identity (2.3) and the definition of \tilde{g}_b in (3.23) allow us to rewrite this as

$$\partial_{x_k} \bar{F}_N^{(K)} = \frac{1}{|\mathcal{D}_K|} \int_{-1}^1 \tilde{g}_b(ky) d\mu^*(y) + \mathcal{O}(N^{-1}).$$

The mean value theorem shows that

$$\begin{aligned} |\mathcal{D}_K| \left| \frac{1}{|\mathcal{D}_K|} \int_{-1}^1 \tilde{g}_b(ky) d\mu^*(y) - \tilde{G}_b^{(K)} x^{(K)}(\mu^*)_k \right| &\leq \sum_{k' \in \mathcal{D}_K} \int_{k'}^{k'+2^{-K}} |\tilde{g}_b(ky) - \tilde{g}_b(kk')| d\mu^*(y) \\ &\leq \frac{\|\tilde{g}'_b\|_\infty}{2^K}, \end{aligned}$$

which means that

$$\|\|\nabla_x \bar{F}_N^{(K)} - \tilde{G}_b^{(K)} x^{(K)}(\mu^*)\|\|_{1,*} \leq \frac{\|\tilde{g}'_b\|_\infty}{2^K} + \mathcal{O}(N^{-1}).$$

In particular,

$$\|\|\tilde{G}_b^{(K)} x^{(K)}(\mu^*)\|\|_{1,*} \leq \|\|\nabla_x \bar{F}_N^{(K)}\|\|_{1,*} + \frac{\|\tilde{g}'_b\|_\infty}{2^K} + \mathcal{O}(N^{-1}) \leq \|\tilde{g}_b\|_\infty + \|\tilde{g}'_b\|_\infty + \mathcal{O}(N^{-1}) \leq R$$

for N large enough. Remembering that $\tilde{H}_{b,K,R}$ coincides with $\tilde{C}_{b,K}$ on $\tilde{\mathcal{C}}_{b,K} \cap B_{K,R}$ and leveraging the Lipschitz continuity of $\tilde{H}_{b,K,R}$ in Proposition 3.8 gives

$$\left| \tilde{H}_{b,K,R} \left(\nabla_x \bar{F}_N^{(K)} \right) - \tilde{C}_{b,K} \left(\tilde{G}_b^{(K)} x^{(K)}(\mu^*) \right) \right| \leq \frac{8RM_b \|\tilde{g}'_b\|_\infty}{2^K m_b^2} + \mathcal{O}(N^{-1}).$$

Another application of the mean value theorem shows that

$$\left| \tilde{C}_{b,K} \left(\tilde{G}_b^{(K)} x^{(K)}(\mu^*) \right) - \frac{1}{2} \int_{-1}^1 \int_{-1}^1 \tilde{g}_b(xy) d\mu^*(y) d\mu^*(x) \right| \leq \frac{\|\tilde{g}'_b\|_\infty}{2^K},$$

while a direct computation using the Nishimori identity (2.3) reveals that

$$\frac{1}{2} \int_{-1}^1 \int_{-1}^1 \tilde{g}_b(xy) d\mu^*(y) d\mu^*(x) = \frac{1}{2}(c + \Delta \bar{m}^2) \log(c) + \frac{c}{2} \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} (\mathbb{E} \langle R_{[n]} \rangle)^2 - \frac{c}{2} + \frac{b}{2}.$$

It follows by (4.42) that up to an error vanishing with N ,

$$\left| \partial_t \bar{F}_N^{(K)} - \tilde{H}_{K,R} \left(\nabla_x \bar{F}_N^{(K)} \right) \right| \leq \left| \frac{c}{2} \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} \mathbb{E} \langle (R_{[n]} - \mathbb{E} \langle R_{[n]} \rangle)^2 \rangle \right| + \frac{8RM_b \|\tilde{g}'_b\|_\infty}{2^K m_b^2} + \frac{\|\tilde{g}'\|_\infty}{2^K}.$$

Invoking the multi-overlap concentration (4.41), noticing that the multi-overlaps are bounded by one and using the formula for the sum of a geometric series implies that, up to an error vanishing with N ,

$$\left| \partial_t \bar{F}_N^{(K)} - \tilde{H}_{K,R} \left(\nabla_x \bar{F}_N^{(K)} \right) \right| \leq \frac{\varepsilon c^2}{2(c-|\Delta|)} + \frac{c}{2} \sum_{n \geq \lceil \varepsilon^{-1} \rceil} (|\Delta|/c)^n + \frac{8RM_b \|\tilde{g}'_b\|_\infty}{2^K m_b^2} + \frac{\|\tilde{g}'\|_\infty}{2^K}.$$

Defining $\mathcal{E}_{\varepsilon,K}$ to be the right-hand side of this expression completes the proof. \blacksquare

5 The disassortative sparse stochastic block model

In this section, we leverage Theorem 1.1 to recover the known variational formula for the sparse stochastic block model in the disassortative regime, $\Delta \leq 0$. The idea will be to apply Theorem 1.5 in [31] to the infinite-dimensional Hamilton-Jacobi equation (1.41) and obtain an infinite-dimensional Hopf-Lax formula. This Hopf-Lax formula will coincide with the standard variational formula obtained in [13, 28] and stated in Theorem 1.1. The key insight will be that in the disassortative setting, for b large enough, the function \tilde{g}_b in (3.23) is such that

$$\int_{-1}^1 \int_{-1}^1 \tilde{g}_b(xy) d\mu(x) d\mu(y) \geq 0 \quad (5.1)$$

for every signed measures $\mu \in \mathcal{M}_s$. This assumption is equivalent to the non-negative definiteness of each of the matrices $\tilde{G}_b^{(K)}$ and to the convexity of each of the projected non-linearities (3.26).

Lemma 5.1. *If $\Delta \leq 0$ and b is large enough, then the function $\tilde{g}_b : [-1, 1] \rightarrow \mathbb{R}$ defined in (3.23) satisfies (5.1).*

Proof. By a simple approximation argument, it suffices to establish (5.1) for a discrete signed measure of the form

$$\mu = \frac{1}{|\mathcal{D}_K|} \sum_{k \in \mathcal{D}_K} x_k \delta_k$$

for some $x \in \mathbb{R}^{\mathcal{D}_K}$. For such a measure,

$$\int_{-1}^1 \int_{-1}^1 \tilde{g}_b(xy) d\mu(x) d\mu(y) = \frac{1}{|\mathcal{D}_K|^2} \sum_{k,k' \in \mathcal{D}_K} \tilde{g}_b(kk') x_k x_{k'} = x^T \tilde{G}_b^{(K)} x,$$

so (5.1) is equivalent to the non-negative definiteness of each of the matrices $\tilde{G}_b^{(K)}$. Observe that for any $k, k' \in \mathcal{D}_K$,

$$\left(\tilde{G}_b^{(K)} \right)_{kk'} = \frac{1}{|\mathcal{D}_K|^2} \tilde{g}_b(kk') = \frac{1}{|\mathcal{D}_K|^2} \left(b + c \log(c) - c + \Delta k k' \log(c) + c \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} (kk')^n \right).$$

If we introduce the vectors $\mathbf{k} = (k)_{k \in \mathcal{D}_K}$ and $\mathbf{1} = (1)_{k \in \mathcal{D}_K}$, and write $\odot n$ for the n -fold Hadamard product on the space of $\mathcal{D}_K \times \mathcal{D}_K$ matrices, this implies that

$$\tilde{G}_b^{(K)} = \frac{1}{|\mathcal{D}_K|^2} \left((b + c \log(c) - c) \mathbf{1} \mathbf{1}^T + \Delta \log(c) \mathbf{k} \mathbf{k}^T + \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} (\mathbf{k} \mathbf{k}^T)^{\odot n} \right) = \lim_{M \rightarrow \infty} \tilde{G}_{b,M}^{(K)}$$

for the matrix

$$\tilde{G}_{b,M}^{(K)} = \frac{1}{|\mathcal{D}_K|^2} \left((b + c \log(c) - c) \mathbf{1} \mathbf{1}^T + \Delta \log(c) \mathbf{k} \mathbf{k}^T + \sum_{2 \leq n \leq M} \frac{(-\Delta/c)^n}{n(n-1)} (\mathbf{k} \mathbf{k}^T)^{\odot n} \right).$$

Choosing $b > 2c|\log(c)| + c$ ensures that the first two terms in this sum define a non-negative definite matrix. Using that $\Delta \leq 0$ and the Schur product theorem, we see that the matrix $\tilde{G}_{b,M}^{(K)}$ is a positive linear combination of non-negative definite matrices, and is therefore non-negative definite. Noticing that the limit of non-negative definite matrices is again non-negative definite completes the proof. \blacksquare

This result allows us to apply Theorem 1.5 in [31] to the infinite-dimensional Hamilton-Jacobi equation (1.41) and obtain a variational formula for its solution. To state this formula concisely, through a slight abuse of notation, introduce the functional $\mathcal{P} : \mathbb{R}_{>0} \times \mathcal{M}_+ \times \mathcal{M}_+ \rightarrow \mathbb{R}$ defined by

$$\mathcal{P}(t, \mu, \nu) = \psi(\mu + t\nu) - \frac{t}{2} \int_{-1}^1 G_\nu(y) \, d\nu(y). \quad (5.2)$$

Theorem 1.5 in [31] and (H2) imply that the unique solution to the infinite-dimensional Hamilton-Jacobi equation (1.41) is given by the Hopf-Lax formula

$$f(t, \mu) = \sup_{\nu \in \text{Pr}[-1,1]} \mathcal{P}(t, \mu, \nu) \quad (5.3)$$

for every $t > 0$ and $\mu \in \mathcal{M}_+$. Combining this representation formula with Theorem 1.1 and a simple interpolation argument taken from [13], we now prove Theorem 1.2.

Proof of Theorem 1.2. By Theorem 1.1 and the Hopf-Lax formula (5.3), the limit of the free energy (1.20) satisfies the upper bound

$$\limsup_{N \rightarrow \infty} \bar{F}_N \leq \sup_{\nu \in \text{Pr}[-1,1]} \mathcal{P}(1, 0, \nu), \quad (5.4)$$

where 0 denotes the zero measure. To show that the right-hand side of this expression may be bounded by the supremum of the functional (1.46) over measures $\mu \in \mathcal{M}_p$, for $b > 1$ large enough so the function \tilde{g}_b defined in (3.23) is strictly positive on $[-1, 1]$ and each integer $N \geq 1$, introduce the functional $\mathcal{P}_{b,N} : \mathbb{R}_{\geq 0} \times \mathcal{M}_+ \times \mathcal{M}_+ \rightarrow \mathbb{R}$ defined by

$$\tilde{\mathcal{P}}_{b,N}(t, \mu, \nu) = \tilde{\psi}_{b,N}(\mu + t\nu) - \frac{t}{2} \int_{-1}^1 \tilde{G}_{b,\nu}(y) \, d\nu(y).$$

Here $\tilde{\psi}_{b,N} : \mathcal{M}_+ \rightarrow \mathbb{R}$ denotes the initial condition

$$\tilde{\psi}_{b,N}(\mu) = \psi_N(\mu) + b \int_{-1}^1 d\mu$$

and $\tilde{G}_{b,v} : [-1, 1] \rightarrow \mathbb{R}$ denotes the function

$$\tilde{G}_{b,v}(x) = \int_{-1}^1 \tilde{g}_b(xy) d\nu(y).$$

By Theorem 1.5 in [31], there exists a probability measure $\tilde{\nu} \in \text{Pr}[-1, 1]$ which maximizes the right-hand side of (5.4). It follows by Proposition 3.6 that

$$\sup_{\nu \in \text{Pr}[-1, 1]} \mathcal{P}(1, 0, \nu) = \lim_{N \rightarrow \infty} \tilde{\mathcal{P}}_{b,N}(1, 0, \tilde{\nu}) - \frac{b}{2} \leq \limsup_{N \rightarrow \infty} \sup_{\nu \in \mathcal{M}_+} \tilde{\mathcal{P}}_{b,N}(1, 0, \nu) - \frac{b}{2}. \quad (5.5)$$

An identical argument to that in Lemma 5.1 of [31] gives a sequence of maximizing measures $(\nu_N) \subset \mathcal{M}_+$ with

$$\sup_{\nu \in \mathcal{M}_+} \tilde{\mathcal{P}}_{b,N}(1, 0, \nu) = \tilde{\mathcal{P}}_{b,N}(1, 0, \nu_N). \quad (5.6)$$

By Corollary 2.5 and (2.37), the Gateaux derivative density of the initial condition $\tilde{\psi}_{b,N}$ at the measure ν_N is given by $D_\mu \tilde{\psi}_{b,N}(\nu_N, \cdot) = \tilde{G}_{b,\nu_N^*} + \mathcal{O}(N^{-1})$ for some measure $\nu_N^* \in \mathcal{M}_p$. Up to adding errors of $\mathcal{O}(N^{-1})$ throughout, the proof of Theorem 1.3 in [31] applies and reveals that each maximizer $\nu_N \in \mathcal{M}_+$ satisfies the approximate first order condition

$$D_\mu \tilde{\psi}_{b,N}(\nu_N, \cdot) = \tilde{G}_{b,\nu_N} + \mathcal{O}(N^{-1}).$$

This means that

$$\tilde{G}_{b,\nu_N^*} = D_\mu \tilde{\psi}_{b,N}(\nu_N, \cdot) + \mathcal{O}(N^{-1}) = \tilde{G}_{b,\nu_N} + \mathcal{O}(N^{-1}).$$

Together with the definition of \tilde{g}_b in (3.23), this implies that

$$\begin{aligned} (b + c \log(c) - c) \int_{-1}^1 d\nu_N(y) + \Delta \log(c) \int_{-1}^1 y d\nu_N(y)x + c \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} \int_{-1}^1 y^n d\nu_N(y)x^n \\ = (b + c \log(c) - c) + \Delta \log(c) \bar{m}x + c \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} \int_{-1}^1 y^n d\nu_N^*(y)x^n + \mathcal{O}(N^{-1}). \end{aligned}$$

Since $b > 1$ and $c \log(c) - c \geq -1$, we must have

$$\int_{-1}^1 d\nu_N(y) = 1 + \mathcal{O}(N^{-1}) \quad \text{and} \quad \int_{-1}^1 y^n d\nu_N(y) = \int_{-1}^1 y^n d\nu_N^*(y) + \mathcal{O}(N^{-1})$$

for all $n \geq 2$. We have used the fact that $\Delta \neq 0$ and accounted for the fact that c could be equal to one. Applying the Prokhorov theorem and passing to a subsequence if necessary, it is therefore possible to ensure that the sequences (ν_N) and (ν_N^*) converge weakly to probability measures $\nu \in \text{Pr}[-1, 1]$ and $\nu^* \in \mathcal{M}_p$ with

$$\int_{-1}^1 y^n d\nu(y) = \int_{-1}^1 y^n d\nu^*(y)$$

for all $n \neq 1$. Since the set of polynomials with degree one coefficient equal to zero form of a sub-algebra of the space of continuous functions on the compact set $[-1, 1]$, the Stone-Weierstrass

theorem implies that $\mathbf{v} = \mathbf{v}^* \in \mathcal{M}_p$. Recalling that we denote $\bar{\mathbf{v}}_N$ for the probability measure induced by \mathbf{v}_N , and arguing as in the proof of Lemma 3.7, we have

$$\begin{aligned} |\tilde{\Psi}_{b,N}(\mathbf{v}_N) - \tilde{\Psi}_b(\mathbf{v}^*)| &\leq |\tilde{\Psi}_{b,N}(\mathbf{v}_N) - \tilde{\Psi}_{b,N}(\bar{\mathbf{v}}_N)| + |\tilde{\Psi}_{b,N}(\bar{\mathbf{v}}_N) - \tilde{\Psi}_{b,N}(\mathbf{v}^*)| + |\tilde{\Psi}_{b,N}(\mathbf{v}^*) - \tilde{\Psi}_b(\mathbf{v}^*)| \\ &\leq C\text{TV}(\mathbf{v}_N, \bar{\mathbf{v}}_N) + CW(\bar{\mathbf{v}}_N, \mathbf{v}^*) + |\tilde{\Psi}_{b,N}(\mathbf{v}^*) - \tilde{\Psi}_b(\mathbf{v}^*)|, \end{aligned}$$

for some constant $C > 0$ that depends only on c . Recalling that the Wasserstein distance (3.9) metrizes the weak convergence of probability measures on $[-1, 1]$, observing that

$$\text{TV}(\mathbf{v}_N, \bar{\mathbf{v}}_N) = |1 - \mathbf{v}_N[-1, 1]|,$$

and using Proposition 3.6 and (5.6) to let N tend to infinity in (5.5) shows that

$$\begin{aligned} \sup_{\mathbf{v} \in \text{Pr}[-1, 1]} \mathcal{P}(1, 0, \mathbf{v}) &\leq \limsup_{N \rightarrow \infty} \tilde{\mathcal{P}}_{b,N}(1, 0, \mathbf{v}_N) - \frac{b}{2} = \tilde{\Psi}_b(\mathbf{v}^*) - \frac{1}{2} \int_{-1}^1 \tilde{G}_{b, \mathbf{v}^*}(y) d\mathbf{v}^*(y) - \frac{b}{2} \\ &\leq \sup_{\mathbf{v} \in \mathcal{M}_p} \mathcal{P}(1, 0, \mathbf{v}). \end{aligned}$$

Substituting this upper bound into (5.4) yields

$$\limsup_{N \rightarrow \infty} \bar{F}_N \leq \sup_{\mathbf{v} \in \mathcal{M}_p} \mathcal{P}(1, 0, \mathbf{v}).$$

To express this upper bound in terms of the functional (1.46), fix $\mathbf{v} \in \mathcal{M}_p$ and denote by x_1 and x_2 two independent samples from the probability measure \mathbf{v} . The definition of g in (1.36) implies that

$$\mathcal{P}(1, 0, \mathbf{v}) = \psi(\mathbf{v}) + \frac{c}{2} - \frac{1}{2}(c + \Delta\bar{m}^2) \log(c) - \frac{c}{2} \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} (\mathbb{E}x_1^n)^2.$$

A Taylor expansion of the logarithm shows that

$$c \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} (\mathbb{E}x^n)^2 = \mathbb{E}(c + \Delta x_1 x_2) \log(c + \Delta x_1 x_2) - (c + \Delta\bar{m}^2) \log(c) - \Delta\bar{m}^2 \quad (5.7)$$

from which it follows that

$$\mathcal{P}(1, 0, \mathbf{v}) = \psi(\mathbf{v}) + \frac{c}{2} - \frac{1}{2} \mathbb{E}(c + \Delta x_1 x_2) \log(c + \Delta x_1 x_2) + \frac{\Delta\bar{m}^2}{2} = \mathcal{P}(\mathbf{v}),$$

and therefore

$$\limsup_{N \rightarrow \infty} \bar{F}_N \leq \sup_{\mathbf{v} \in \mathcal{M}_p} \mathcal{P}(\mathbf{v}).$$

This establishes the upper bound in (1.47). To prove the corresponding lower bound, we follow [13] and proceed by interpolation. Given a measure $\mathbf{v} \in \mathcal{M}_p$, introduce the interpolating free energy

$$\varphi(t) = \tilde{F}_N(t, 1-t, \mathbf{v}) \quad (5.8)$$

for the free energy \tilde{F}_N defined in (1.29). The derivative computations in Corollary 2.2 and Lemma 2.3 together with a computation identical to that in Corollary 2.5 imply that

$$\begin{aligned}\varphi'(t) &= \partial_t \tilde{F}_N(t, 1-t, \mathbf{v}) - \partial_s \tilde{F}_N(t, 1-t, \mathbf{v}) \\ &= \frac{c}{2} - \frac{1}{2}(c + \Delta \bar{m}^2) \log(c) + \frac{c}{2} \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} \mathbb{E} \langle R_{[n]}^2 \rangle - c \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} \mathbb{E} \langle R_{[n]} \rangle \mathbb{E} x_1^n \\ &= \frac{c}{2} - \frac{1}{2}(c + \Delta \bar{m}^2) \log(c) - \frac{c}{2} \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} (\mathbb{E} x_1^n)^2 + \frac{c}{2} \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} \mathbb{E} \langle (R_{[n]} - \mathbb{E} x_1^n)^2 \rangle.\end{aligned}$$

It follows by (5.7) that

$$\varphi'(t) = \frac{c}{2} + \frac{\Delta \bar{m}^2}{2} - \frac{1}{2} \mathbb{E}(c + \Delta x_1 x_2) \log(c + \Delta x_1 x_2) + \frac{c}{2} \sum_{n \geq 2} \frac{(-\Delta/c)^n}{n(n-1)} \mathbb{E} \langle (R_{[n]} - \mathbb{E} x_1^n)^2 \rangle.$$

Since the final term in this equality is non-negative, the fundamental theorem of calculus reveals that

$$\bar{F}_N \geq \psi_N(\mathbf{v}) + \frac{c}{2} + \frac{\Delta \bar{m}^2}{2} - \frac{1}{2} \mathbb{E}(c + \Delta x_1 x_2) \log(c + \Delta x_1 x_2),$$

where we have used that $\varphi(1) = \bar{F}_N$ and $\varphi(0) = \psi_N(\mathbf{v})$. Using Proposition 3.6 to let N tend to infinity gives the lower bound

$$\liminf_{N \rightarrow \infty} \bar{F}_N \geq \psi(\mathbf{v}) + \frac{c}{2} + \frac{\Delta \bar{m}^2}{2} - \frac{1}{2} \mathbb{E}(c + \Delta x_1 x_2) \log(c + \Delta x_1 x_2) = \mathcal{P}(\mathbf{v}).$$

Taking the supremum over all measures $\mathbf{v} \in \mathcal{M}_p$ completes the proof. \blacksquare

A Asymptotic equivalence of free energy functionals

In this appendix we show that the free energy functionals \bar{F}_N° and \bar{F}_N defined in (1.15) and (1.20), respectively, are asymptotically equivalent. This will be a consequence of the binomial-Poisson approximation. To state this result concisely, given a separable metric space S , recall the definition of the total variation distance

$$\text{TV}(\mathbb{P}, \mathbb{Q}) = \sup \{ |\mathbb{P}(A) - \mathbb{Q}(A)| \mid A \text{ is a measurable subset of } S \} \quad (\text{A.1})$$

between probability measures $\mathbb{P}, \mathbb{Q} \in \text{Pr}(S)$. Approximating any measurable function with values in S by a sequence of simple functions, one can check that the total variation distance admits the dual representation

$$\text{TV}(\mathbb{P}, \mathbb{Q}) = \sup \left\{ \left| \int_{-1}^1 f(x) d\mathbb{P}(x) - \int_{-1}^1 f(x) d\mathbb{Q}(x) \right| \mid f : S \rightarrow [0, 1] \text{ measurable} \right\}. \quad (\text{A.2})$$

Using the Hahn-Jordan decomposition, it is also possible to show that

$$\text{TV}(\mathbb{P}, \mathbb{Q}) = \inf \{ \mathbb{P}\{X \neq Y\} \mid X \sim \mathbb{P} \text{ and } Y \sim \mathbb{Q} \}. \quad (\text{A.3})$$

We will use this result for discrete probability measures supported on the set of natural numbers in which case this representation follows from the Kantorovich-Rubinstein theorem (see Theorem 4.15 in [65]). The binomial-Poisson approximation is an upper bound on the total variation distance between the convolution of Bernoulli distributions and an appropriate Poisson distribution.

Lemma A.1. *Consider independent Bernoulli random variables $X_i \sim \text{Ber}(p_i)$ for $i \leq n$, and let $\lambda_n = \sum_{i \leq n} p_i$. If $S_n = \sum_{i \leq n} X_i$ and $\Pi_n \sim \text{Poi}(\lambda_n)$, then*

$$\text{TV}(S_n, \Pi_n) \leq \sum_{i \leq n} p_i^2. \quad (\text{A.4})$$

Proof. See Theorem 2.4 in [65]. ■

Proposition A.2. *The free energies (1.15) and (1.20) are asymptotically equivalent,*

$$\lim_{N \rightarrow \infty} |\bar{F}_N - \bar{F}_N^\circ| = 0. \quad (\text{A.5})$$

Proof. Introduce the Hamiltonians

$$\begin{aligned} \tilde{H}_N^\circ(\sigma) &= \sum_{i < j} \left(G_{ij} \log(c + \Delta \sigma_i \sigma_j) - \frac{c + \Delta \sigma_i \sigma_j}{N} \right) \\ \tilde{H}_N(\sigma) &= \sum_{k \leq \Pi_1} \left(G_{i_k, j_k}^k \log(c + \Delta \sigma_{i_k} \sigma_{j_k}) - \frac{c + \Delta \sigma_{i_k} \sigma_{j_k}}{N} \right) \end{aligned}$$

on Σ_N , and denote by

$$\tilde{F}_N^\circ = \frac{1}{N} \mathbb{E} \log \int_{\Sigma_N} \exp \tilde{H}_N^\circ(\sigma) dP_N^*(\sigma) \quad \text{and} \quad \tilde{F}_N = \frac{1}{N} \mathbb{E} \log \int_{\Sigma_N} \exp \tilde{H}_N(\sigma) dP_N^*(\sigma)$$

their associated free energy functionals. A Taylor expansion of the logarithm shows that for any $\sigma \in \Sigma_N$,

$$\begin{aligned} |H_N^\circ(\sigma) - \tilde{H}_N^\circ(\sigma)| &\leq \sum_{i < j} \left| \frac{c + \Delta \sigma_i \sigma_j}{N} - (1 - G_{ij}) \left(\frac{c + \Delta \sigma_i \sigma_j}{N} \right) \right| + \mathcal{O}(1) \\ &\leq \frac{c + |\Delta|}{N} \sum_{i < j} G_{ij} + \mathcal{O}(1) \end{aligned}$$

and

$$\begin{aligned} |H_N(\sigma) - \tilde{H}_N(\sigma)| &\leq \sum_{k \leq \Pi_1} \left| \frac{c + \Delta \sigma_{i_k} \sigma_{j_k}}{N} - (1 - G_{i_k, j_k}^k) \left(\frac{c + \Delta \sigma_{i_k} \sigma_{j_k}}{N} \right) \right| + \mathcal{O}(\Pi_1/N^2) \\ &\leq \frac{c + |\Delta|}{N} \sum_{k \leq \Pi_1} G_{i_k, j_k}^k + \mathcal{O}(\Pi_1/N^2). \end{aligned}$$

Since these bounds are uniform in $\sigma \in \Sigma_N$ and $\mathbb{E} \Pi_1 = \binom{N}{2}$,

$$|\bar{F}_N^\circ - \tilde{F}_N^\circ| \leq \frac{c + |\Delta|}{N^2} \sum_{i < j} \mathbb{E} G_{ij} + \mathcal{O}(N^{-1}) \leq \frac{(c + |\Delta|)^2}{N} + \mathcal{O}(N^{-1}) = \mathcal{O}(N^{-1}),$$

$$|\bar{F}_N - \tilde{F}_N| \leq \frac{c + |\Delta|}{N^2} \mathbb{E} \sum_{k \leq \Pi_1} G_{i_k, j_k}^k + \mathcal{O}(\mathbb{E} \Pi_1/N^3) \leq \frac{(c + |\Delta|)^2}{N^3} \mathbb{E} \Pi_1 + \mathcal{O}(N^{-1}) = \mathcal{O}(N^{-1}).$$

By the triangle inequality, it therefore suffices to show that

$$\lim_{N \rightarrow \infty} |\tilde{F}_N - \tilde{F}_N^\circ| = 0. \quad (\text{A.6})$$

We now rewrite \tilde{F}_N in a way that more closely resembles \tilde{F}_N° . For each pair $i \leq j$ introduce the random index set

$$\mathcal{I}_{i,j} = \{k \leq \Pi_1 \mid (i_k, j_k) = (i, j) \text{ or } (i_k, j_k) = (j, i)\},$$

and observe that

$$\begin{aligned} \tilde{H}_N(\sigma) &= \sum_{i \leq j} \sum_{k \in \mathcal{I}_{i,j}} \left(G_{i_k, j_k}^k \log(c + \Delta \sigma_{i_k} \sigma_{j_k}) - \frac{c + \Delta \sigma_{i_k} \sigma_{j_k}}{N} \right) \\ &= \sum_{i < j} \left(\tilde{G}_{i,j} \log(c + \Delta \sigma_i \sigma_j) - \frac{c + \Delta \sigma_i \sigma_j}{N} \right) - \sum_{i \leq N} \left(\tilde{G}_{i,i} \log(c + \Delta) - \frac{c + \Delta}{N} \right) \end{aligned}$$

for the random variables

$$\tilde{G}_{i,j} = \sum_{k \in \mathcal{I}_{i,j}} G_{i,j}^k.$$

The Poisson coloring theorem (see Chapter 5 in [40]) implies that $\tilde{G}_{i,j}$ is a Poisson random variable with mean

$$\tilde{\lambda}_{i,j} = \mathbb{E} \Pi_1 \cdot \mathbb{P}\{(i_1, j_1) = (i, j) \text{ or } (i_1, j_1) = (j, i)\} \cdot \mathbb{P}\{G_{i,j}^1 = 1\} = \begin{cases} \frac{N-1}{N} \cdot \frac{c + \Delta \sigma_i^* \sigma_j^*}{N} & \text{if } i < j, \\ \frac{N-1}{2N} \cdot \frac{c + \Delta}{N} & \text{if } i = j. \end{cases}$$

If we introduce the Hamiltonian

$$\tilde{H}'_N(\sigma) = \sum_{i < j} \left(\tilde{G}_{i,j} \log(c + \Delta \sigma_i \sigma_j) - \frac{c + \Delta \sigma_i \sigma_j}{N} \right)$$

and its associated free energy

$$\tilde{F}'_N = \frac{1}{N} \mathbb{E} \log \int_{\Sigma_N} \exp \tilde{H}'_N(\sigma) dP_N^*(\sigma),$$

then

$$|\tilde{F}'_N - \tilde{F}_N| \leq \frac{|\log(c + \Delta)|}{N} \sum_{i \leq N} \mathbb{E} \tilde{G}_{i,i} + \frac{c + |\Delta|}{N} \leq \frac{|\log(c + \Delta)|}{N} \sum_{i \leq N} \tilde{\lambda}_{i,i} + \frac{2c}{N} = \mathcal{O}(N^{-1}).$$

Together with (A.6) and the triangle inequality, this means that it suffices to show that

$$\lim_{N \rightarrow \infty} |\tilde{F}'_N - \tilde{F}_N^\circ| = 0. \quad (\text{A.7})$$

At this point, for any random vector $Y = (Y_{i,j})_{i < j}$ introduce the Hamiltonian

$$\tilde{H}_N(\sigma, Y) = \sum_{i < j} Y_{i,j} \log(c + \Delta \sigma_i \sigma_j)$$

and the measure

$$\tilde{P}_N^*(\sigma) = \exp\left(-\sum_{i<j} \frac{c + \Delta\sigma_i\sigma_j}{N}\right) P_N^*(\sigma).$$

Write

$$F_N(Y) = \frac{1}{N} \log \int_{\Sigma_N} \exp \tilde{H}_N(\sigma, Y) d\tilde{P}_N^*(\sigma) \quad \text{and} \quad \tilde{F}_N(Y) = \mathbb{E} F_N(Y)$$

for the associated free energy functionals, and denote by $\Pi_{i,j}$ a Poisson random variable with distribution

$$\Pi_{i,j} \sim \text{Poi}\left(\frac{c + \Delta\sigma_i^* \sigma_j^*}{N}\right)$$

conditionally on σ^* . Observe that

$$\partial_{Y_{i,j}} F_N(Y) = \frac{1}{N} \langle \partial_{Y_{i,j}} \tilde{H}_N(\sigma, Y) \rangle = \frac{1}{N} \langle \log(c + \Delta\sigma_i\sigma_j) \rangle \leq \frac{C}{N}$$

for some deterministic constant $C > 0$ that depends only on c . Here, the bracket $\langle \cdot \rangle$ denotes the Gibbs measure associated with the Hamiltonian $\tilde{H}_N(\sigma, Y)$. It follows by the mean value theorem that

$$\begin{aligned} |\tilde{F}'_N - \tilde{F}_N(\Pi)| &= |\tilde{F}_N(\tilde{G}) - \tilde{F}_N(\Pi)| \leq \frac{C}{N} \sum_{i<j} \mathbb{E} |\tilde{G}_{i,j} - \Pi_{i,j}| \\ |\tilde{F}_N(\Pi) - \tilde{F}_N^\circ| &= |\tilde{F}_N(\Pi) - \tilde{F}_N(G)| \leq \frac{C}{N} \sum_{i<j} \mathbb{E} |\Pi_{i,j} - G_{ij}|. \end{aligned}$$

To bound the first of the sums observe that $\Pi_{i,j} \stackrel{d}{=} \tilde{G}_{i,j} + \Pi'_{i,j}$ for a Poisson random variable with distribution

$$\Pi'_{i,j} \sim \text{Poi}\left(\frac{c + \Delta\sigma_i^* \sigma_j^*}{N^2}\right)$$

conditionally on σ^* . This means that

$$|\tilde{F}'_N - \tilde{F}_N(\Pi)| \leq \frac{C}{N} \sum_{i<j} \mathbb{E} \Pi'_{i,j} \leq \frac{Cc}{N} = \mathcal{O}(N^{-1}),$$

and therefore,

$$|\tilde{F}'_N - \tilde{F}_N^\circ| \leq \frac{C}{N} \sum_{i<j} \mathbb{E} |\Pi_{i,j} - G_{ij}| + \mathcal{O}(N^{-1}). \quad (\text{A.8})$$

If we write $\lambda_{i,j} = \frac{c + \Delta\sigma_i^* \sigma_j^*}{N}$, then

$$\begin{aligned} \mathbb{E} |\Pi_{i,j} - G_{ij}| &\leq \mathbb{E} |\Pi_{i,j} - G_{ij}| \mathbb{1}\{\Pi_{i,j} \geq 2\} + \mathbb{E} |\Pi_{i,j} - G_{ij}| \mathbb{1}\{\Pi_{i,j} \leq 2\} \mathbb{1}\{\Pi_{i,j} \neq G_{ij}\} \\ &\leq \mathbb{E} \Pi_{i,j} \mathbb{1}\{\Pi_{i,j} \geq 2\} + \mathbb{P}\{\Pi_{i,j} \geq 2\} + 3\mathbb{P}\{\Pi_{i,j} \neq G_{ij}\} \\ &\leq 3\mathbb{P}\{\Pi_{i,j} \neq G_{ij}\} + \mathcal{O}(\lambda_{i,j}^2), \end{aligned}$$

where we have used the fact that $|G_{ij}| \leq 1$. Taking the infimum over all couplings of $\Pi_{i,j}$ and G_{ij} , recalling the definition of the total variation distance in (A.3) and invoking the binomial-Poisson approximation in Lemma A.1 shows that

$$\mathbb{E} |\Pi_{i,j} - G_{ij}| \leq 3\text{TV}(\Pi_{i,j}, G_{ij}) + \mathcal{O}(\lambda_{i,j}^2) = \mathcal{O}(\lambda_{i,j}^2) = \mathcal{O}(N^{-2}).$$

Substituting this into (A.8) and letting N tend to infinity establishes (A.7). This completes the proof. \blacksquare

B Concentration of the free energy

In this appendix we discuss the concentration of the free energy associated with the perturbed Hamiltonian (4.7). For simplicity of notation, we will ignore the Hamiltonian (1.34) and focus instead on the perturbed Hamiltonian

$$H'_N(\sigma) = H_N(\sigma) + H_N^{\text{gauss}}(\sigma) + H_N^{\text{exp}}(\sigma), \quad (\text{B.1})$$

where the randomness of each of the Hamiltonians (1.26), (4.1) and (4.5) is independent of the randomness of the other Hamiltonians. The more general case is treated in an identical fashion but the notation becomes too cumbersome for comfort. We will often need to make the dependence of the perturbed Hamiltonian (B.1) on one of its sources of randomness σ^* , Π_t , $\mathcal{I}_1 = (i_k, j_k)_{k \leq \Pi_t}$, $\mathcal{G} = (G_{i_k, j_k}^k)_{k \leq \Pi_t}$, $e = (e_{jk})$, $\Pi' = (\pi_k)_{k \geq 0}$, $\mathcal{I}_2 = (i_{jk})_{j \leq \Pi'}$ and $Z = (Z_{0,i})_{i \leq N}$ explicit. To do so, we will abuse notation and write $H'_N(X)$ when we want to study the dependence on the source of randomness X . The main objects of study will be the free energy,

$$F'_N = \frac{1}{N} \log \int \exp H'_N(\sigma) dP_N^*(\sigma), \quad (\text{B.2})$$

and its average

$$\bar{F}'_N = \frac{1}{N} \mathbb{E} \log \int \exp H'_N(\sigma) dP_N^*(\sigma). \quad (\text{B.3})$$

For each even $p \geq 2$, we will bound the concentration function,

$$v_{N,p} = \sup \left\{ \mathbb{E} |F'_N - \bar{F}'_N|^p \mid \lambda_k \in [2^{-k-1}, 2^{-k}] \text{ for all } k \geq 0 \right\}, \quad (\text{B.4})$$

by means of the generalized Efron-Stein inequality (see Theorem 15.5 of [17]).

Lemma B.1 (Generalized Efron-Stein inequality). *Let $X = (X_1, \dots, X_n)$ and $X' = (X'_1, \dots, X'_n)$ be two independent copies of a vector of independent random variables, and let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a measurable function. Introduce the random variable $Z = f(X)$, and for each $1 \leq i \leq n$ let $Z'_i = f(X_1, \dots, X_{i-1}, X'_i, X_{i+1}, \dots, X_n)$. If $q \geq 2$, then*

$$\mathbb{E} |Z - \mathbb{E} Z|^q \leq C \mathbb{E} \left| \sum_{i \leq n} \mathbb{E}_{X'} (Z - Z'_i)^2 \right|^{\frac{q}{2}}, \quad (\text{B.5})$$

where $C > 0$ is a constant that depends only on q .

A key observation that will be used repeatedly without further explanation is the following: given two sources of randomness X and X' , a configuration-independent bound on the difference of the Hamiltonians $H'_N(X)$ and $H'_N(X')$,

$$\max_{\sigma \in \Sigma_N} |H'_N(X) - H'_N(X')| \leq Y, \quad (\text{B.6})$$

gives a control by the possibly random Y on the difference of the free energy functionals $F'_N(X)$ and $F'_N(X')$,

$$|F'_N(X) - F'_N(X')| \leq \frac{Y}{N}. \quad (\text{B.7})$$

The following bounds on the moments of Poisson and binomial random variables will also play their part.

Lemma B.2. *If Π is a $\text{Poi}(\lambda)$ random variable for some $\lambda \geq 1$ and $k \geq 2$ is an integer, then*

$$\mathbb{E}\Pi^k \leq C\lambda^k \quad \text{and} \quad \mathbb{E}(\Pi - \mathbb{E}\Pi)^k \leq C\lambda^{\lfloor k/2 \rfloor} \quad (\text{B.8})$$

for some constant $C > 0$ that depends only on k .

Proof. Denote by $\left\{ \begin{smallmatrix} k \\ j \end{smallmatrix} \right\}$ the number of ways to partition a k element set into j non-empty subsets. In combinatorics, such numbers are known as Stirling numbers of the second kind, and they have the property that for any integer $m \geq 0$,

$$m^k = \sum_{j \leq k} \left\{ \begin{smallmatrix} k \\ j \end{smallmatrix} \right\} (m)_j, \quad (\text{B.9})$$

where $(m)_j = m(m-1)\cdots(m-j+1)$ is the falling factorial. The basic properties of the Poisson distribution imply that

$$\begin{aligned} \mathbb{E}\Pi^k &= \sum_{m \geq 0} \sum_{j \leq k} \left\{ \begin{smallmatrix} k \\ j \end{smallmatrix} \right\} (m)_j \frac{\lambda^m}{m!} \exp(-\lambda) = \sum_{j \leq k} \left\{ \begin{smallmatrix} k \\ j \end{smallmatrix} \right\} \lambda^j \sum_{m \geq j} \frac{\lambda^{m-j}}{(m-j)!} \exp(-\lambda) = \sum_{j \leq k} \left\{ \begin{smallmatrix} k \\ j \end{smallmatrix} \right\} \lambda^j \\ &\leq \max(1, \lambda^k) B_k, \end{aligned}$$

where B_k denotes the k 'th Bell number. This establishes the first bound in (B.8). We now prove by induction that for each $k \geq 2$, the function $M_k(\lambda) = \mathbb{E}(\Pi - \mathbb{E}\Pi)^k$ is a polynomial of degree $\lfloor k/2 \rfloor$. The base case holds since $M_2(\lambda) = \text{Var}\Pi = \lambda$, so assume the result holds for all $2 \leq i \leq k$. By the product rule

$$\begin{aligned} M'_k(\lambda) &= - \sum_{m \geq 0} k(m-\lambda)^{k-1} \frac{\lambda^m}{m!} \exp(-\lambda) + \sum_{m \geq 0} (m-\lambda)^k m \frac{\lambda^{m-1}}{m!} \exp(-\lambda) - M_k(\lambda) \\ &= -kM_{k-1}(\lambda) + \sum_{m \geq 0} (m-\lambda)^k (m-\lambda + \lambda) \frac{\lambda^{m-1}}{m!} \exp(-\lambda) - M_k(\lambda) \\ &= -kM_{k-1}(\lambda) + \frac{1}{\lambda} (M_{k+1}(\lambda) + \lambda M_k(\lambda)) - M_k(\lambda) \\ &= -kM_{k-1}(\lambda) + \frac{1}{\lambda} M_{k+1}(\lambda). \end{aligned}$$

Invoking the induction hypothesis shows that $M_{k+1}(\lambda)$ has degree $\max(\lfloor k/2 \rfloor, 1 + \lfloor (k-1)/2 \rfloor)$. This completes the proof. \blacksquare

Lemma B.3. *If X is a $\text{Bin}(n, p)$ random variable with $np \geq 1$ and $k \geq 1$ is an integer, then*

$$\mathbb{E}X^k \leq C(np)^k \quad (\text{B.10})$$

for some constant $C > 0$ that depends only on k .

Proof. Using (B.9) and identifying the probability density function of a $\text{Bin}(n-j, p)$ shows that

$$\begin{aligned}\mathbb{E}X^k &= \sum_{m \leq n} \sum_{j \leq k} \binom{k}{j} (m)_j \frac{n!}{(n-m)!m!} p^m (1-p)^{n-m} \\ &\leq \sum_{j \leq k} \binom{k}{j} (np)^j \sum_{j \leq m \leq n} \frac{(n-j)!}{(n-m)!(m-j)!} p^{m-j} (1-p)^{n-m} \\ &\leq \max(1, (np)^k) B_k,\end{aligned}$$

where B_k denotes the k 'th Bell number. This completes the proof. \blacksquare

Proposition B.4. *For any two sequences (ε_N) and (s_N) satisfying (4.3) and (4.4) respectively and every even $p \geq 2$,*

$$v_{N,p} \leq \frac{C(1+t^p)}{N^{p/2}} \quad (\text{B.11})$$

for some constant $C > 0$ that depends only on p, c and Δ .

Proof. To alleviate notation, write $C > 0$ for a constant that depends only on p, c and Δ whose value might not be the same at each occurrence. Given a source of randomness X , write \mathbb{E}_X for the average with respect to the randomness of X . The proof will rely upon the generalized Efron-Stein inequality in Lemma B.1 and the fact that

$$\mathbb{E} = \mathbb{E}_{\sigma^*} \mathbb{E}_Z \mathbb{E}_{\Pi'} \mathbb{E}_{\mathcal{J}_2} \mathbb{E}_e \mathbb{E}_{\Pi_t} \mathbb{E}_{\mathcal{J}_1} \mathbb{E}_{\mathcal{G}|\sigma^*}.$$

Introduce the averaged free energy functionals

$$\widehat{F}'_N = \mathbb{E}_{\Pi_t} \mathbb{E}_{\mathcal{J}_1} \mathbb{E}_{\mathcal{G}|\sigma^*} F'_N \quad \text{and} \quad \widetilde{F}'_N = \mathbb{E}_Z \mathbb{E}_{\Pi'} \mathbb{E}_{\mathcal{J}_2} \mathbb{E}_e \widehat{F}'_N$$

in such a way that

$$\mathbb{E} (F'_N - \overline{F}'_N)^p \leq C \left(\mathbb{E} (F'_N - \widehat{F}'_N)^p + \mathbb{E} (\widehat{F}'_N - \widetilde{F}'_N)^p + \mathbb{E} (\widetilde{F}'_N - \overline{F}'_N)^p \right). \quad (\text{B.12})$$

We will now bound each of these terms separately.

Step 1: proving $\mathbb{E} (F'_N - \widehat{F}'_N)^p = \mathcal{O}((t/N)^{p/2})$.

We decompose this further into

$$\begin{aligned}\mathbb{E} (F'_N - \widehat{F}'_N)^p &\leq C \left(\mathbb{E} (F'_N - \mathbb{E}_{\mathcal{G}|\sigma^*} F'_N)^p + \mathbb{E} (\mathbb{E}_{\mathcal{G}|\sigma^*} F'_N - \mathbb{E}_{\mathcal{J}_1} \mathbb{E}_{\mathcal{G}|\sigma^*} F'_N)^p \right. \\ &\quad \left. + \mathbb{E} (\mathbb{E}_{\mathcal{J}_1} \mathbb{E}_{\mathcal{G}|\sigma^*} F'_N - \widehat{F}'_N)^p \right) = C(I + II + III),\end{aligned} \quad (\text{B.13})$$

and proceed to bound I, II and III individually. By the generalized Efron-Stein inequality

$$I \leq C \mathbb{E} \left| \sum_{\ell \leq \Pi_t} \mathbb{E}_{\mathcal{G}^{(\ell)}|\sigma^*} \left(F'_N(\mathcal{G}) - F'_N(\mathcal{G}^{(\ell)}) \right)^2 \right|^{p/2},$$

where $(\tilde{G}_{i,j}^{(\ell)})_{i,j \in \mathbb{N}}$ is an independent copy of $(G_{i,j}^{(\ell)})_{i,j \in \mathbb{N}}$. Since $|\Delta| < c$ and all spin configuration coordinates are bounded by one,

$$\begin{aligned} |H'_N(\mathcal{G}) - H'_N(\mathcal{G}^{(\ell)})| &\leq |\log(c + \Delta \sigma_{i_\ell} \sigma_{j_\ell})| |G_{i_\ell, j_\ell}^\ell - \tilde{G}_{i_\ell, j_\ell}^\ell| \\ &\quad + \left| \log \left(1 - \frac{c + \Delta \sigma_{i_\ell} \sigma_{j_\ell}}{N} \right) \right| |G_{i_\ell, j_\ell}^\ell - \tilde{G}_{i_\ell, j_\ell}^\ell| \leq C |G_{i_\ell, j_\ell}^\ell - \tilde{G}_{i_\ell, j_\ell}^\ell|. \end{aligned}$$

It follows that

$$\begin{aligned} I &\leq \frac{C}{N^p} \mathbb{E} \left| \sum_{\ell \leq \Pi_t} \mathbb{E}_{\mathcal{G}^{(\ell)} | \sigma^*} |G_{i_\ell, j_\ell}^\ell - \tilde{G}_{i_\ell, j_\ell}^\ell|^2 \right|^{p/2} \leq \frac{C}{N^p} \mathbb{E} \left| \sum_{\ell \leq \Pi_t} \left(\left(1 - \frac{2}{N} \right) G_{i_\ell, j_\ell}^\ell + \frac{1}{N} \right) \right|^{p/2} \\ &\leq \frac{C}{N^p} \left(\mathbb{E} \left| \sum_{\ell \leq \Pi_t} G_{i_\ell, j_\ell}^\ell \right|^{p/2} + \frac{1}{N^{p/2}} \mathbb{E} \Pi_t^{p/2} \right). \end{aligned}$$

Notice that $\sum_{\ell \leq \Pi_t} G_{i_\ell, j_\ell}^\ell$ follows a Bin $\left(\Pi_t, \frac{c + \Delta \sigma_{i_\ell}^* \sigma_{j_\ell}^*}{N} \right)$ distribution conditionally on Π_t . Invoking Lemma B.2 and Lemma B.3 yields

$$I \leq \frac{C}{N^{p+\frac{p}{2}}} \mathbb{E} \Pi_t^{p/2} \leq \frac{C t^{p/2}}{N^{p/2}}. \quad (\text{B.14})$$

Another application of the generalized Efron-Stein inequality gives

$$II \leq \mathbb{E} \left| \sum_{\ell \leq \Pi_t} \mathbb{E}_{\mathcal{I}_1^{(\ell)}} \left(\mathbb{E}_{\mathcal{G} | \sigma^*} F'_N(\mathcal{I}_1) - \mathbb{E}_{\mathcal{G} | \sigma^*} F'_N(\mathcal{I}_1^{(\ell)}) \right) \right|^{p/2},$$

where $\mathcal{I}_1^{(\ell)}$ has an independent copy (i'_ℓ, j'_ℓ) of (i_ℓ, j_ℓ) at the ℓ 'th coordinate but otherwise coincides with \mathcal{I}_1 . Taylor expanding the logarithm and remembering that $G_{i_\ell, j_\ell}^\ell \in \{0, 1\}$, it is readily verified that

$$|H'_N(\mathcal{I}_1)| \leq |G_{i_\ell, j_\ell}^\ell| |\log(c + \Delta \sigma_{i_\ell} \sigma_{j_\ell})| + |1 - G_{i_\ell, j_\ell}^\ell| \left| \log \left(1 - \frac{c + \Delta \sigma_{i_\ell} \sigma_{j_\ell}}{N} \right) \right| \leq C \left(|G_{i_\ell, j_\ell}^\ell| + \frac{1}{N} \right).$$

This means that $|H'_N(\mathcal{I}_1) - H'_N(\mathcal{I}_1^{(\ell)})| \leq C \left(|G_{i_\ell, j_\ell}^\ell| + |G_{i'_\ell, j'_\ell}^\ell| + \frac{1}{N} \right)$, and therefore

$$\left(\mathbb{E}_{\mathcal{G} | \sigma^*} F'_N(\mathcal{I}_1) - \mathbb{E}_{\mathcal{G} | \sigma^*} F'_N(\mathcal{I}_1^{(\ell)}) \right)^2 \leq \frac{C}{N^2} \left(\mathbb{E}_{\mathcal{G} | \sigma^*} |G_{i_\ell, j_\ell}^\ell| + \mathbb{E}_{\mathcal{G} | \sigma^*} |G_{i'_\ell, j'_\ell}^\ell| + \frac{1}{N} \right)^2 \leq \frac{C}{N^4}.$$

It follows that

$$II \leq \frac{C}{N^{2p}} \mathbb{E} \Pi_t^{p/2} \leq \frac{C t^{p/2}}{N^p} \leq \frac{C t^{p/2}}{N^{p/2}}. \quad (\text{B.15})$$

A final application of the generalized Efron-Stein inequality reveals that

$$III \leq \mathbb{E} \left| \mathbb{E}_{\Pi'_t} \left(\mathbb{E}_{\mathcal{I}_1} \mathbb{E}_{\mathcal{G} | \sigma^*} F'_N(\Pi_t) - \mathbb{E}_{\mathcal{I}_1} \mathbb{E}_{\mathcal{G} | \sigma^*} F'_N(\Pi'_t) \right) \right|^{p/2},$$

where Π'_t is an independent copy of Π_t . Slightly abusing notation and redefining Π'_t to be the maximum between Π_t and Π'_t , we see that

$$\begin{aligned} |H'_N(\Pi'_t) - H'_N(\Pi_t)| &\leq \sum_{\Pi_t \leq k \leq \Pi'_t} \left(|G_{i_k, j_k}^k| |\log(c + \Delta \sigma_{i_k} \sigma_{j_k})| + |1 - G_{i_k, j_k}^k| \left| \log \left(1 - \frac{c + \Delta \sigma_{i_k} \sigma_{j_k}}{N} \right) \right| \right) \\ &\leq C \sum_{\Pi_t \leq k \leq \Pi'_t} \left(|G_{i_k, j_k}^k| + \frac{1}{N} \right). \end{aligned}$$

It follows that

$$\left| \mathbb{E}_{\mathcal{J}_1} \mathbb{E}_{\mathcal{G}} |_{\sigma^*} F'_N(\Pi_t) - \mathbb{E}_{\mathcal{J}_1} \mathbb{E}_{\mathcal{G}} |_{\sigma^*} F'_N(\Pi'_t) \right| \leq \frac{C}{N} \mathbb{E}_{\mathcal{J}_1} \mathbb{E}_{\mathcal{G}} |_{\sigma^*} \sum_{\Pi_t \leq k \leq \Pi'_t} \left(|G_{i_k, j_k}^k| + \frac{1}{N} \right) \leq \frac{C}{N^2} |\Pi'_t - \Pi_t|,$$

and by Jensen's inequality and Lemma B.2,

$$III \leq \frac{C}{N^{2p}} \mathbb{E} \left| \mathbb{E}_{\Pi'_t} |\Pi'_t - \Pi_t|^2 \right|^{p/2} \leq \frac{C}{N^{2p}} \mathbb{E} |\Pi_t - \mathbb{E} \Pi_t|^p \leq \frac{C t^{p/2}}{N^p} \leq \frac{C t^{p/2}}{N^{p/2}}. \quad (\text{B.16})$$

Combining (B.13), (B.14), (B.15) and (B.16) reveals that $\mathbb{E} (F'_N - \widehat{F}'_N)^p = \mathcal{O}((t/N)^{p/2})$.

Step 2: proving $\mathbb{E} (\widehat{F}'_N - \widetilde{F}'_N)^p = \mathcal{O}(N^{-p/2})$.

We decompose this further into

$$\begin{aligned} \mathbb{E} (\widehat{F}'_N - \widetilde{F}'_N)^p &\leq C \left(\mathbb{E} (\widehat{F}'_N - \mathbb{E}_e \widehat{F}'_N)^p + \mathbb{E} (\mathbb{E}_e \widehat{F}'_N - \mathbb{E}_{\mathcal{J}_2} \mathbb{E}_e \widehat{F}'_N)^p + \mathbb{E} (\mathbb{E}_{\mathcal{J}_2} \mathbb{E}_e \widehat{F}'_N - \mathbb{E}_{\Pi'} \mathbb{E}_{\mathcal{J}_2} \mathbb{E}_e \widehat{F}'_N)^p \right. \\ &\quad \left. + \mathbb{E} (\mathbb{E}_{\Pi'} \mathbb{E}_{\mathcal{J}_2} \mathbb{E}_e \widehat{F}'_N - \widetilde{F}'_N)^p \right) = C(I + II + III + IV), \end{aligned} \quad (\text{B.17})$$

and proceed to bound *I*, *II*, *III* and *IV* individually. By the generalized Efron-Stein inequality

$$I \leq \mathbb{E} \left| \sum_{k \geq 0} \sum_{j \leq \pi_k} \mathbb{E}_{e^{(jk)}} \left(\widehat{F}'_N(e) - \widehat{F}'_N(e^{(jk)}) \right) \right|^{p/2},$$

where $e^{(jk)}$ has an independent copy e'_{jk} of e_{jk} at the jk 'th coordinate but otherwise coincides with e . Since

$$\left| H'_N(e) - H'_N(e^{(jk)}) \right| \leq \frac{|\lambda_k \sigma_{i_{jk}}|}{|1 + \lambda_k \sigma_{i_{jk}}^*|} |e_{jk} - e'_{jk}| \leq \frac{\lambda_k}{1 - \lambda_k} |e_{jk} - e'_{jk}|,$$

and $\lambda_k \in [2^{-k-1}, 2^{-k}]$, we have

$$I \leq \frac{C}{N^p} \mathbb{E} \left| \sum_{k \geq 0} \frac{1}{2^k} \cdot \frac{1}{2^k} \sum_{j \leq \pi_k} \mathbb{E}_{e'_{jk}} |e_{jk} - e'_{jk}|^2 \right|^{p/2}.$$

It follows by two applications of Hölder's inequality and Jensen's inequality that

$$I \leq \frac{C}{N^p} \mathbb{E} \sum_{k \geq 0} \left(\frac{1}{2^k} \sum_{j \leq \pi_k} \mathbb{E}_{e'_{jk}} |e_{jk} - e'_{jk}|^2 \right)^{p/2} \leq \frac{C}{N^p} \mathbb{E} \sum_{k \geq 0} \frac{1}{2^{\frac{kp}{2}}} \pi_k^{\frac{p}{2}-1} \sum_{j \leq \pi_k} \mathbb{E}_{e'_{jk}} |e_{jk} - e'_{jk}|^p.$$

Recalling that $e_{jk} \sim \text{Exp}(1)$ while $\pi_{jk} \sim \text{Poi}(s_N)$ and invoking Lemma B.2 gives

$$I \leq \frac{C}{N^p} \sum_{k \geq 0} \frac{1}{2^{\frac{kp}{2}}} \mathbb{E} \pi_k^{\frac{p}{2}} \leq C \left(\frac{s_N}{N^2} \right)^{p/2} \leq \frac{C}{N^{p/2}}. \quad (\text{B.18})$$

Similarly, by the generalized Efron-Stein inequality,

$$II \leq C \mathbb{E} \left| \sum_{k \geq 0} \sum_{j \leq \pi_k} \mathbb{E}_{\mathcal{J}_2^{(jk)}} \left(\mathbb{E}_e \widehat{F}'_N(\mathcal{J}_2) - \mathbb{E}_e \widehat{F}'_N(\mathcal{J}_2^{(jk)}) \right) \right|^{p/2},$$

where $\mathcal{S}_2^{(jk)}$ has an independent copy i'_{jk} of i_{jk} at the jk 'th coordinate but otherwise coincides with \mathcal{S}_2 . By the mean value theorem,

$$\begin{aligned} |H'_N(\mathcal{S}_2) - H'_N(\mathcal{S}_2^{(jk)})| &\leq |\log(1 + \lambda_k \sigma_{i_{jk}}) - \log(1 + \lambda_k \sigma_{i'_{jk}})| + \lambda_k e_{jk} \left| \frac{\sigma_{i_{jk}}}{1 + \lambda_k \sigma_{i_{jk}}^*} - \frac{\sigma_{i'_{jk}}}{1 + \lambda_k \sigma_{i'_{jk}}^*} \right| \\ &\leq C\lambda_k(1 + e_{jk}). \end{aligned}$$

It follows once again by two applications of Hölder's inequality and Lemma B.2 that

$$II \leq \frac{C}{N^p} \mathbb{E} \sum_{k \geq 0} \frac{1}{2^{\frac{kp}{2}}} \pi_k^{\frac{p}{2}-1} \sum_{j \leq \pi_k} (1 + e_{jk})^p \leq C \left(\frac{s_N}{N^2} \right)^{p/2} \leq \frac{C}{N^{p/2}}. \quad (\text{B.19})$$

Another application of the generalized Efron-Stein inequality yields

$$III \leq C \mathbb{E} \left| \sum_{k \geq 0} \mathbb{E}_{\Pi^{(k)}} \left(\mathbb{E}_{\mathcal{S}_2} \mathbb{E}_e \widehat{F}'_N(\Pi') - \mathbb{E}_{\mathcal{S}_2} \mathbb{E}_e \widehat{F}'_N(\Pi^{(k)}) \right) \right|^{p/2},$$

where $\Pi^{(k)}$ has an independent copy π'_k of π_k at the k 'th coordinate but otherwise coincides with Π' . Slightly abusing notation and redefining $\Pi^{(k)}$ to be the process with the larger k 'th coordinate, we see that

$$\begin{aligned} |H'_N(\Pi') - H'_N(\Pi^{(k)})| &\leq \sum_{\pi_k \leq j \leq \pi'_k} \left| \log(1 + \lambda_k \sigma_{i_{jk}}) - \frac{\lambda_k e_{jk} \sigma_{i_{jk}}}{1 + \lambda_k \sigma_{i_{jk}}^*} \right| \\ &\leq \sum_{\pi_k \leq j \leq \pi'_k} \left(\left| \lambda_k \sigma_{i_{jk}} - \frac{\lambda_k e_{jk} \sigma_{i_{jk}}}{1 + \lambda_k \sigma_{i_{jk}}^*} \right| + C\lambda_k^2 \right) \\ &\leq \sum_{\pi_k \leq j \leq \pi'_k} \left(\lambda_k \left| 1 - \frac{e_{jk}}{1 + \lambda_k \sigma_{i_{jk}}^*} \right| + C\lambda_k^2 \right) \\ &\leq \lambda_k \sum_{\pi_k \leq j \leq \pi'_k} \left(\frac{|1 - e_{jk}| + \lambda_k}{1 - \lambda_k} + C\lambda_k \right). \end{aligned}$$

It follows by two applications of the Cauchy-Schwarz inequality that

$$\begin{aligned} III &\leq \frac{C}{N^p} \mathbb{E} \sum_{k \geq 0} \frac{1}{2^{\frac{kp}{2}}} \left(\sum_{\pi_k \leq j \leq \pi'_k} \left(\frac{|1 - e_{jk}| + \lambda_k}{1 - \lambda_k} + C\lambda_k \right) \right)^p \\ &\leq \frac{C}{N^p} \mathbb{E} \sum_{k \geq 0} \frac{1}{2^{\frac{kp}{2}}} |\pi_k - \pi'_k|^{p-1} \sum_{\pi_k \leq j \leq \pi'_k} \left(\frac{|1 - e_{jk}| + \lambda_k}{1 - \lambda_k} + C\lambda_k \right)^p \\ &\leq \frac{C}{N^p} \sum_{k \geq 0} \frac{1}{2^{\frac{kp}{2}}} \mathbb{E} |\pi_k - \pi'_k|^p. \end{aligned}$$

Since $\mathbb{E} |\pi_k - \pi'_k|^p \leq C \mathbb{E} |\pi_k - \mathbb{E} \pi_k|^p \leq s_N^{p/2}$ by Lemma B.2, this implies that

$$III \leq C \left(\frac{s_N}{N^2} \right)^{p/2} \leq \frac{C}{N^{p/2}}. \quad (\text{B.20})$$

A final application of the generalized Efron-Stein inequality gives

$$IV \leq C \mathbb{E} \left| \sum_{i \leq N} \mathbb{E}_{Z^{(i)}} \left(\mathbb{E}_{\Pi'} \mathbb{E}_{\mathcal{J}_2} \mathbb{E}_e \widehat{F}'_N(Z) - \mathbb{E}_{\Pi'} \mathbb{E}_{\mathcal{J}_2} \mathbb{E}_e \widehat{F}'_N(Z^{(i)}) \right) \right|^{p/2},$$

where $Z^{(i)}$ has an independent copy $Z'_{i,0}$ of $Z_{i,0}$ at the i 'th coordinate but otherwise coincides with Z . Combining Hölder's inequality with the bound

$$|H'_N(Z) - H'_N(Z^{(i)})| \leq \sqrt{\lambda_0 \varepsilon_N} |Z_{i,0} - Z'_{i,0}|$$

reveals that

$$IV \leq C \left(\frac{\lambda_0 \varepsilon_N}{N} \right)^p \mathbb{E} \left| \sum_{i \leq N} \mathbb{E}_{Z^{(i)}} |Z_{i,0} - Z'_{i,0}|^2 \right|^{p/2} \leq \frac{C}{N^p} N^{\frac{p}{2}-1} \sum_{i \leq N} \mathbb{E} |Z_{i,0} - Z'_{i,0}|^p \leq \frac{C}{N^{p/2}}. \quad (\text{B.21})$$

Together with (B.17), (B.18), (B.19) and (B.20), this shows that $\mathbb{E} (\widehat{F}'_N - \widetilde{F}'_N)^p = \mathcal{O}(N^{-p/2})$.

Step 3: proving $\mathbb{E} (\widetilde{F}'_N - \overline{F}'_N)^p = \mathcal{O}((t^2/N)^{p/2})$

Controlling the final term in (B.12) requires more care since \widetilde{F}'_N depends on σ^* both through F'_N and through the conditional expectation $\mathbb{E}_{\mathcal{G}|\sigma^*}$. To simplify notation, write $\mathbb{E}' = \mathbb{E}_Z \mathbb{E}_{\Pi'} \mathbb{E}_{\mathcal{J}_2} \mathbb{E}_e \mathbb{E}_{\Pi_t} \mathbb{E}_{\mathcal{J}_1}$ in such a way that by the generalized Efron-Stein inequality

$$\begin{aligned} \mathbb{E} (\widetilde{F}'_N - \overline{F}'_N)^p &\leq C \mathbb{E} \left| \sum_{\ell \leq N} \mathbb{E}_{\sigma^{*,(\ell)}} \left(\mathbb{E}' \mathbb{E}_{\mathcal{G}|\sigma^*} F'_N(\sigma^*) - \mathbb{E}' \mathbb{E}_{\mathcal{G}|\sigma^{*,(\ell)}} F'_N(\sigma^{*,(\ell)}) \right) \right|^{p/2} \\ &\leq C \mathbb{E} \left| \sum_{\ell \leq N} \mathbb{E}_{\sigma^{*,(\ell)}} \left(\mathbb{E}' \mathbb{E}_{\mathcal{G}|\sigma^*} F'_N(\sigma^*) - \mathbb{E}' \mathbb{E}_{\mathcal{G}|\sigma^*} F'_N(\sigma^{*,(\ell)}) \right) \right|^{p/2} \\ &\quad + C \mathbb{E} \left| \sum_{\ell \leq N} \mathbb{E}_{\sigma^{*,(\ell)}} \left(\mathbb{E}' \mathbb{E}_{\mathcal{G}|\sigma^*} F'_N(\sigma^{*,(\ell)}) - \mathbb{E}' \mathbb{E}_{\mathcal{G}|\sigma^{*,(\ell)}} F'_N(\sigma^{*,(\ell)}) \right) \right|^{p/2} \\ &= C(I + II), \end{aligned} \quad (\text{B.22})$$

where $\sigma^{*,(\ell)}$ has an independent copy $\widetilde{\sigma}_\ell^*$ of σ_ℓ^* at the ℓ 'th coordinate but otherwise coincides with σ^* . Since

$$\begin{aligned} |H'_N(\sigma^*) - H'_N(\sigma^{*,(\ell)})| &\leq \lambda_0 \varepsilon_N |\sigma_i| |\sigma_\ell^* - \widetilde{\sigma}_\ell^*| + \sum_{k \geq 1} \sum_{j: i_{jk}=\ell} \left| \frac{\lambda_k e_{jk} \sigma_\ell}{1 + \lambda_k \sigma_\ell^*} - \frac{\lambda_k e_{jk} \sigma_\ell}{1 + \lambda_k \widetilde{\sigma}_\ell^*} \right| \\ &\leq 2\varepsilon_N + \sum_{k \geq 1} \sum_{j: i_{jk}=\ell} \frac{2\lambda_k^2 e_{jk}}{(1 - \lambda_k)^2}, \end{aligned}$$

and $\mathbb{E} e_{jk} = 1$, the Fubini-Tonelli theorem and the basic properties of the multinomial distribution imply that

$$\begin{aligned} |\mathbb{E}' \mathbb{E}_{\mathcal{G}|\sigma^*} F'_N(\sigma^*) - \mathbb{E}' \mathbb{E}_{\mathcal{G}|\sigma^*} F'_N(\sigma^{*,(\ell)})| &\leq \frac{2\varepsilon_N}{N} + \frac{1}{N} \mathbb{E}_{\Pi'} \sum_{k \geq 1} \frac{2\lambda_k^2}{(1 - \lambda_k)^2} \mathbb{E}_{\mathcal{J}_2} |\{j : i_k = \ell\}| \\ &\leq \frac{2}{N} + \frac{1}{N^2} \mathbb{E}_{\Pi'} \sum_{k \geq 1} \frac{2\lambda_k^2}{(1 - \lambda_k)^2} \pi_k \leq \frac{2}{N} + \frac{s_N}{N^2} \leq \frac{3}{N}. \end{aligned}$$

It follows that

$$I \leq C \left(\frac{N}{N^2} \right)^{p/2} = \frac{C}{N^{p/2}}. \quad (\text{B.23})$$

To bound II we will use an interpolation argument. Fix $1 \leq \ell \leq N$ and condition on all sources of randomness other than \mathcal{G} . For each $u \in [0, 1]$, $G \in \{0, 1\}^{\Pi_t}$ and $k \leq \Pi_t$, let

$$\begin{aligned} P_u^{1,k}(G) &= G_k \left(\frac{c + \Delta \sigma_\ell^{*,u} \sigma_{j_k}^*}{N} \right) + (1 - G_k) \left(1 - \frac{c + \Delta \sigma_\ell^{*,u} \sigma_{j_k}^*}{N} \right), \\ P_u^{2,k}(G) &= G_k \left(\frac{c + \Delta \sigma_\ell^{*,u} \sigma_{i_k}^*}{N} \right) + (1 - G_k) \left(1 - \frac{c + \Delta \sigma_\ell^{*,u} \sigma_{i_k}^*}{N} \right), \\ P_u^3(G) &= G_k \left(\frac{c + \Delta (\sigma_\ell^{*,u})^2}{N} \right) + (1 - G_k) \left(1 - \frac{c + \Delta (\sigma_\ell^{*,u})^2}{N} \right), \end{aligned}$$

where $\sigma_\ell^{*,u} = (1 - u)\sigma_\ell^* + u\tilde{\sigma}_\ell^*$. Write $\sigma^{*,u}$ for the configuration with ℓ 'th coordinate $\sigma_\ell^{*,u}$ which otherwise coincides with σ^* , and introduce the sets

$$\begin{aligned} \mathcal{I}_1^1 &= \{k \mid i_k = l \text{ and } j_k \neq l\} & \mathcal{I}_1^3 &= \{k \mid i_k = j_k = l\} \\ \mathcal{I}_1^2 &= \{k \mid i_k \neq l \text{ and } j_k = l\} & \mathcal{I}_1^4 &= \{k \mid i_k \neq l \neq j_k\}. \end{aligned}$$

Denote $\tilde{\mathcal{G}} = (\mathcal{G}_k)_{k \in \mathcal{I}_1^4}$, $\mathcal{G}^{(k)} = \mathcal{G} \setminus \mathcal{G}_k$, $\tilde{G} = (G_k)_{k \in \mathcal{I}_1^4}$ and $G^{(k)} = G \setminus G_k$. Define the interpolating free energy

$$\varphi(u) = \sum_{G \in \{0,1\}^{\Pi_t}} F'_N(\sigma^{*,(\ell)}, G) \mathbb{P}\{\tilde{\mathcal{G}} = \tilde{G} \mid \sigma^*\} \cdot \prod_{k \in \mathcal{I}_1^1} P_u^{1,k}(G) \prod_{k \in \mathcal{I}_1^2} P_u^{2,k}(G) \prod_{k \in \mathcal{I}_1^3} P_u^{3,k}(G)$$

in such a way that $\varphi(1) = \mathbb{E}_{\mathcal{G} \mid \sigma^*} F'_N(\sigma^{*,(\ell)})$ and $\varphi(0) = \mathbb{E}_{\mathcal{G} \mid \sigma^{*,(\ell)}} F'_N(\sigma^{*,(\ell)})$. By the product rule,

$$\begin{aligned} \varphi'(u) &= \sum_{k \in \cup_{i \leq 3} \mathcal{I}_1^i} \sum_{G \in \{0,1\}^{\Pi_t}} F'_N(\sigma^{*,(\ell)}, G) \mathbb{P}\{\mathcal{G}^{(k)} = G^{(k)} \mid \sigma^{*,u}\} (2G_k - 1) \\ &\quad \left(\frac{\Delta(\tilde{\sigma}_\ell^* - \sigma_\ell^*)(\sigma_{j_k}^* \mathbb{1}_{\{k \in \mathcal{I}_1^1\}} + \sigma_{i_k}^* \mathbb{1}_{\{k \in \mathcal{I}_1^2\}} + 2\sigma_\ell^{*,u} \mathbb{1}_{\{k \in \mathcal{I}_1^3\}})}{N} \right) \\ &= \sum_{k \in \mathcal{I}_1^1} D_k \frac{\Delta(\tilde{\sigma}_\ell^* - \sigma_\ell^*) \sigma_{j_k}^*}{N} + \sum_{k \in \mathcal{I}_1^2} D_k \frac{\Delta(\tilde{\sigma}_\ell^* - \sigma_\ell^*) \sigma_{i_k}^*}{N} \\ &\quad + \sum_{k \in \mathcal{I}_1^3} D_k \frac{2\Delta((1-u)\sigma_\ell^* + u\tilde{\sigma}_\ell^*)(\tilde{\sigma}_\ell^* - \sigma_\ell^*)}{N} \end{aligned}$$

for $D_k = \mathbb{E}_{\mathcal{G}^{(k)} \mid \sigma^{*,u}} F'_N(\sigma^{*,(\ell)}, \mathcal{G}^{(k)}, \mathcal{G}_k = 1) - \mathbb{E}_{\mathcal{G}^{(k)} \mid \sigma^{*,u}} F'_N(\sigma^{*,(\ell)}, \mathcal{G}^{(k)}, \mathcal{G}_k = 0)$. Since

$$\left| H'_N(\mathcal{G}^{(k)}, \mathcal{G}_k = 1) - H'_N(\mathcal{G}^{(k)}, \mathcal{G}_k = 0) \right| \leq \left| \log(c + \Delta \sigma_{i_k} \sigma_{j_k}) - \log \left(1 - \frac{c + \Delta \sigma_{i_k} \sigma_{j_k}}{N} \right) \right| \leq C,$$

we have $|D_k| \leq \frac{C}{N}$, so the fundamental theorem of calculus yields

$$\left| \mathbb{E}_{\mathcal{G} \mid \sigma^*} F'_N(\sigma^{*,(\ell)}) - \mathbb{E}_{\mathcal{G} \mid \sigma^{*,(\ell)}} F'_N(\sigma^{*,(\ell)}) \right| \leq \sup_{u \in (0,1)} |\varphi'(u)| \leq \frac{C}{N^2} (|\mathcal{I}_1^1| + |\mathcal{I}_1^2| + |\mathcal{I}_1^3|).$$

It follows by the basic properties of the binomial distribution that

$$\begin{aligned} \left(\mathbb{E}' \mathbb{E}_{\mathcal{G}|\sigma^*} F'_N(\sigma^{*,(\ell)}) - \mathbb{E}' \mathbb{E}_{\mathcal{G}|\sigma^{*,(\ell)}} F'_N(\sigma^{*,(\ell)}) \right)^2 &\leq \frac{C}{N^4} \sum_{1 \leq i \leq 3} (\mathbb{E}_{\Pi_i} |\mathcal{I}_1^i|)^2 = \frac{C}{N^4} \sum_{1 \leq i \leq 3} \left(\mathbb{E}_{\Pi_i} \frac{\Pi_i}{N} \right)^2 \\ &\leq \frac{Ct^2}{N^2}, \end{aligned}$$

and thus

$$H \leq C \left(\frac{Nt^2}{N^2} \right)^{p/2} = \frac{Ct^p}{N^{p/2}}. \quad (\text{B.24})$$

Combining (B.22), (B.23) and (B.24) reveals that $\mathbb{E}(\tilde{F}'_N - \bar{F}'_N)^p = \mathcal{O}((t^2/N)^{p/2})$. Together with (B.12), step 1 and step 2 this completes the proof. \blacksquare

C Multi-overlap concentration

In this appendix we prove a finitary version of the main result in [14] regarding the concentration of the multi-overlaps (2.11). Instead of focusing on the stochastic block model (1.19) or (1.34), we will work in the general setting of optimal Bayesian inference; this presents no additional difficulty, and we suspect that our restatement of the multi-overlap concentration result in [14] will be useful for the analysis of other statistical inference models.

Let us describe a general optimal Bayesian inference model following [14]. We consider a ground-truth signal $\sigma^* \in \Sigma_N = \{-1, +1\}^N$ with independent coordinates generated from a prior distribution P^* ,

$$\sigma^* \sim P^* = \prod_{i \leq N} P_i^*. \quad (\text{C.1})$$

We suppose that the data $\mathcal{D} = \mathcal{D}(\sigma^*)$ is sampled conditionally on the unknown signal σ^* from a probability distribution P_{out} ,

$$\mathcal{D} \sim P_{\text{out}}\{\cdot | \sigma^*\}. \quad (\text{C.2})$$

The inference task is to recover the signal σ^* as accurately as possible given the data \mathcal{D} under the assumption that the likelihood P_{out} and the prior P^* are known to the statistician. In this setting the posterior of the model can be written explicitly. Indeed, if

$$H_N(\sigma) = \log P_{\text{out}}\{\mathcal{D} | \sigma^* = \sigma\} \quad (\text{C.3})$$

denotes the Hamiltonian or log-likelihood of the model, then the Gibbs measure or posterior distribution of the model is given by Bayes' formula,

$$G_N(d\sigma) = \mathbb{P}\{\sigma^* \in d\sigma | \mathcal{D}\} = \frac{\exp H_N(\sigma) P^*(d\sigma)}{\int \exp H_N(\tau) P^*(d\tau)}. \quad (\text{C.4})$$

In addition to working in the context of optimal Bayesian inference, we will assume that the Hamiltonian (C.3) satisfies symmetry between sites. This means that for any permutation ρ of the spin indices,

$$\mathbb{P}\{\sigma^* \in d\sigma | \mathcal{D}\} \stackrel{d}{=} \mathbb{P}\{\rho(\sigma^*) \in d\sigma | \mathcal{D}\}. \quad (\text{C.5})$$

Notice that both the stochastic block model (1.19) and its enriched version (1.34) fall into the setting we have just described. For instance, in the enriched stochastic block model we have $P_i^* = \text{Ber}(p)$ and $\mathcal{D} = \tilde{\mathcal{D}}^{t,\mu[-1,1],\bar{\mu}}$, where the data $\tilde{\mathcal{D}}^{t,s,\mu}$ was defined in (1.30).

The concentration of the multi-overlaps (2.11) associated with the Hamiltonian (C.3) will be enforced through a small perturbation which will not affect the limit of the associated free energy (1.20). Recall the choice of the sequences (ε_N) and (s_N) satisfying (4.3) and (4.4), respectively. Fix an integer $K_+ \geq 1$, and for each perturbation parameter $\lambda \in \mathbb{R}^{1+K_+}$ with $\lambda_k \in [2^{-k-1}, 2^{-k}]$ for $0 \leq k \leq K_+$ recall the definition of the perturbation Hamiltonians (4.1) and (4.5). Introduce the perturbed Hamiltonian

$$H_N(\sigma, \lambda) = H_N(\sigma) + H_N^{\text{gauss}}(\sigma, \lambda_0) + H_N^{\text{exp}}(\sigma, \lambda), \quad (\text{C.6})$$

where the randomness of each Hamiltonian is independent of the randomness of the other Hamiltonians. The multi-overlaps associated with this perturbed Hamiltonian are defined as in (2.11),

$$R_{\ell_1, \dots, \ell_n} = \frac{1}{N} \sum_{i \leq N} \sigma_i^{\ell_1} \cdots \sigma_i^{\ell_n}, \quad (\text{C.7})$$

where (σ^ℓ) denotes a sequence of i.i.d. replicas sampled from the Gibbs measure $\langle \cdot \rangle$ associated with the perturbed Hamiltonian (C.6). It is actually these multi-overlaps that will be shown to concentrate. The proof of Lemma 4.1 shows that the free energy functionals associated with the Hamiltonians (C.3) and (C.6) are asymptotically equivalent; for our purposes these two Hamiltonians can therefore be thought to describe the same model.

To establish the concentration of the multi-overlaps (C.7) we will closely follow the arguments in [14]. The authors in [14] obtain the concentration of the multi-overlaps (C.7) for some perturbation parameter λ by showing that it holds on average over the set of admissible perturbation parameters. In the proof of Theorem 1.1 we will need to be able to obtain multi-overlap concentration for a specific perturbation parameter. Following the strategy in [14], we will propose a verifiable condition on a perturbation parameter λ which ensures the concentration of its associated multi-overlaps. To be more precise, we will obtain the concentration of the multi-overlaps (C.7) up to a small error for any sequence of perturbation parameters (λ^N) with

$$\lim_{N \rightarrow \infty} \mathbb{E} \langle (\mathcal{L}_k - \mathbb{E} \langle \mathcal{L}_k \rangle)^2 \rangle = 0. \quad (\text{C.8})$$

The quantities \mathcal{L}_k are defined in (4.26) and (4.27) for $0 \leq k \leq K_+$, and, through a slight abuse of notation, we have written $\langle \cdot \rangle$ for the Gibbs average with respect to the perturbed Hamiltonian (C.6) associated with the perturbation parameters (λ^N) . If necessary, we will write $\langle \cdot \rangle_N$ to emphasize the dependence of this Gibbs measure on N .

We begin by showing that (C.8) implies the concentration of the magnetization R_1 and of the overlap $R_{1,2}$; the former will be immediate from the Nishimori identity (2.3) while the latter will follow from a standard application of the Gaussian integration by parts formula (see Lemma 1.4 in [60]).

Lemma C.1. *For any integer $N \geq 1$,*

$$\mathbb{E} \langle (R_1 - \mathbb{E} \langle R_1 \rangle)^2 \rangle \leq \frac{1}{N}. \quad (\text{C.9})$$

Proof. Applying the Nishimori identity reveals that

$$\mathbb{E} \langle (R_1 - \mathbb{E} \langle R_1 \rangle)^2 \rangle = \frac{1}{N^2} \sum_{i \leq N} \mathbb{E} (\sigma_i^* - \mathbb{E} \sigma_i^*)^2 \leq \frac{1}{N}.$$

This completes the proof. ■

Lemma C.2. *For any integer $N \geq 1$,*

$$\mathbb{E} \langle (R_{1,2} - \mathbb{E} \langle R_{1,2} \rangle)^2 \rangle \leq 4 \mathbb{E} \langle (\mathcal{L}_0 - \mathbb{E} \langle \mathcal{L}_0 \rangle)^2 \rangle. \quad (\text{C.10})$$

Proof. The proof is taken from the Appendix of [14], and it consists in testing the concentration of the overlap $R_{1,*} = \frac{\sigma \cdot \sigma^*}{N}$ against the Hamiltonian \mathcal{L}_0 by means of the Gaussian integration by parts formula. Using (4.26) shows that

$$\begin{aligned} \mathbb{E} \langle (R_{1,*} - \mathbb{E} \langle R_{1,*} \rangle) (\mathcal{L}_0 - \mathbb{E} \langle \mathcal{L}_0 \rangle) \rangle &= \mathbb{E} \langle R_{1,*} (R_{1,*} - \mathbb{E} \langle R_{1,*} \rangle) \rangle \\ &\quad + \frac{1}{2N \sqrt{\lambda_{0,N}}} \mathbb{E} \langle R_{1,*} (\sigma \cdot Z_0 - \mathbb{E} \langle \sigma \cdot Z_0 \rangle) \rangle. \end{aligned} \quad (\text{C.11})$$

The Gaussian integration by parts formula and the Nishimori identity imply that

$$\mathbb{E} \langle \sigma \cdot Z \rangle = N \sqrt{\lambda_{0,N}} (1 - \mathbb{E} \langle R_{1,*} \rangle) \quad \text{and} \quad \mathbb{E} \langle R_{1,*} \sigma \cdot Z \rangle = N \sqrt{\lambda_{0,N}} (\mathbb{E} \langle R_{1,*} \rangle - \mathbb{E} \langle R_{1,*} \rangle^2).$$

Substituting these two equalities into (C.11) reveals that

$$\mathbb{E} \langle (R_{1,*} - \mathbb{E} \langle R_{1,*} \rangle) (\mathcal{L}_0 - \mathbb{E} \langle \mathcal{L}_0 \rangle) \rangle = \frac{1}{2} \mathbb{E} \langle (R_{1,*} - \mathbb{E} \langle R_{1,*} \rangle)^2 \rangle + \frac{1}{2} \mathbb{E} \langle (R_{1,*} - \mathbb{E} \langle R_{1,*} \rangle)^2 \rangle.$$

(There seems to be a sign error in equation (5.2) of [14].) It follows by the Nishimori identity that

$$\mathbb{E} \langle (R_{1,*} - \mathbb{E} \langle R_{1,*} \rangle) (\mathcal{L}_0 - \mathbb{E} \langle \mathcal{L}_0 \rangle) \rangle \geq \frac{1}{2} \mathbb{E} \langle (R_{1,2} - \mathbb{E} \langle R_{1,2} \rangle)^2 \rangle.$$

Invoking the Cauchy-Schwarz inequality and the Nishimori identity completes the proof. ■

The concentration of the multi-overlaps (C.7) is considerably more complicated to obtain, and follows from the Franz-de Sanctis identities described in [14]. The first section of this appendix will be devoted to establishing this implication. In the second section we will prove a finitary version of the multi-overlap concentration result in [14] which will be uniform over an appropriate class of random probability measures. This uniformity plays its part in the proof Lemma 4.2.

C.1 Franz-de Sanctis identities

The Franz-de Sanctis identities may be thought of as the Ghirlanda-Guerra identities of optimal Bayesian inference. A random probability measure which satisfies the Ghirlanda-Guerra identities must have an ultrametric support; a deep insight which leads to the appearance of the intricate Poisson-Dirichlet probability cascades in many spin glass models [60]. The Franz-de Sanctis identities enforce a much simpler and more rigid structure on a random probability measure which we

will describe in due course. To state these identities, it will be convenient to fix a uniform index $i \in \{1, \dots, N\}$ and an exponential random variable $e \sim \text{Exp}(1)$ independent of all other sources of randomness, and introduce the random variables

$$y_{ik} = \frac{e}{1 + \lambda_k \sigma_i^*}, \quad \theta_{ik}^\ell = \log(1 + \lambda_k \sigma_i^\ell) - \lambda_k y_{ik} \sigma_i^\ell, \quad \text{and} \quad d_{ik}^\ell = \frac{y_{ik} \sigma_i^\ell}{1 + \lambda_k \sigma_i^*} \quad (\text{C.12})$$

for $1 \leq k \leq K_+$.

Proposition C.3 (Franz-de Sanctis identities in inference). *For any $1 \leq k \leq K_+$ and any function f_n of finitely many spins on n replicas and of the signal σ^* with $\|f_n\|_{L^\infty} \leq 1$,*

$$\left| \mathbb{E} \frac{\langle f_n d_{ik}^1 \exp(\sum_{\ell \leq n} \theta_{ik}^\ell) \rangle}{\langle \exp(\theta_{ik}) \rangle} - \mathbb{E} \langle f_n \rangle \mathbb{E} \frac{\langle d_{ik} \exp(\theta_{ik}) \rangle}{\langle \exp(\theta_{ik}) \rangle} \right| \leq \left(2 \mathbb{E} \langle (\mathcal{L}_k - \mathbb{E} \langle \mathcal{L}_k \rangle)^2 \rangle + \frac{16}{s_N} \right)^{1/2}. \quad (\text{C.13})$$

The Franz-de Sanctis identities are proved in a similar fashion to the Ghirlanda-Guerra identities by testing the concentration of the quantities

$$\widetilde{\mathcal{L}}_k = \frac{1}{s_N} \sum_{j \leq \pi_k} \frac{\sigma_{ijk} e^{jk}}{(1 + \lambda_k \sigma_{ijk}^*)^2} \quad (\text{C.14})$$

defined for $1 \leq k \leq K_+$ against an arbitrary function of finitely many spins and of the signal σ^* . Notice that $\widetilde{\mathcal{L}}_k$ is none other than the second term in the sum defining each \mathcal{L}_k in (4.27). The reason for focusing only on the second term is that the first term concentrates automatically by the Nishimori identity. This is the content of Proposition 3.4 in [14] which we reproduce here for completeness. We present a slightly simpler proof than that in [14] which was kindly shared with us by Dmitry Panchenko.

Lemma C.4. *For any $1 \leq k \leq K_+$ and every large enough $N \geq 1$,*

$$\mathbb{E} \langle (\widetilde{\mathcal{L}}_k - \mathbb{E} \langle \widetilde{\mathcal{L}}_k \rangle)^2 \rangle \leq 2 \mathbb{E} \langle (\mathcal{L}_k - \mathbb{E} \langle \mathcal{L}_k \rangle)^2 \rangle + \frac{16}{s_N}. \quad (\text{C.15})$$

Proof. Introduce the quantity

$$g(\sigma, \pi_k) = \sum_{j \leq \pi_k} \frac{\sigma_{ijk}}{1 + \lambda_k \sigma_{ijk}}$$

in such a way that $\widetilde{\mathcal{L}}_k = s_N^{-1} g(\sigma, \pi_k) - \mathcal{L}_k$. Write Var for the variance with respect to the measure $\mathbb{E} \langle \cdot \rangle$. Since the variance of a sum of two random variables is bounded by twice the sum of the variance of each of the random variables,

$$\text{Var}(\widetilde{\mathcal{L}}_k) \leq 2 \left(\text{Var}(\mathcal{L}_k) + \frac{1}{s_N^2} \text{Var}(g) \right). \quad (\text{C.16})$$

By the Nishimori identity and a direct computation,

$$\text{Var}(g(\sigma, \pi_k)) = \mathbb{E} \left(\sum_{j \leq \pi_k} \frac{\sigma_{ijk}^*}{1 + \lambda_k \sigma_{ijk}^*} \right)^2 - \left(\mathbb{E} \sum_{j \leq \pi_k} \frac{\sigma_{ijk}^*}{1 + \lambda_k \sigma_{ijk}^*} \right)^2. \quad (\text{C.17})$$

Recalling that the coordinates of the signal σ^* are i.i.d. and averaging with respect to the randomness of the indices (i_{jk}) reveals that

$$\begin{aligned}\mathbb{E}\left(\sum_{j \leq \pi_k} \frac{\sigma_{i_{jk}}^*}{1 + \lambda_k \sigma_{i_{jk}}^*}\right)^2 &= \frac{1}{N} \mathbb{E} \sum_{j, j' \leq \pi_k} \frac{1}{(1 + \lambda_k \sigma_1^*)^2} + \frac{N^2 - N}{N} \mathbb{E} \sum_{j, j' \leq \pi_k} \frac{\sigma_1^* \sigma_2^*}{(1 + \lambda_k \sigma_1^*)(1 + \lambda_k \sigma_2^*)} \\ &\leq \frac{4 \mathbb{E} \pi_k^2}{N} + \mathbb{E} \pi_k^2 \left(\mathbb{E} \frac{\sigma_1^*}{1 + \lambda_k \sigma_1^*}\right)^2,\end{aligned}$$

where we have used that $\sigma_1^2 = 1$ and $\lambda_k \leq 1/2$. Similarly,

$$\left(\mathbb{E} \sum_{j \leq \pi_k} \frac{\sigma_{i_{jk}}^*}{1 + \lambda_k \sigma_{i_{jk}}^*}\right)^2 = (\mathbb{E} \pi_k)^2 \left(\mathbb{E} \frac{\sigma_1^*}{1 + \lambda_k \sigma_1^*}\right)^2.$$

Substituting these two bounds into (C.17), recalling (4.4) and choosing N large enough yields

$$\text{Var}(g(\sigma, \pi_k)) \leq \frac{4 \mathbb{E} \pi_k^2}{N} + \text{Var}(\pi_k) \left(\mathbb{E} \frac{\sigma_1^*}{1 + \lambda_k \sigma_1^*}\right)^2 \leq 8s_N.$$

Plugging this into (C.16) completes the proof. ■

Proof of Proposition C.3. We follow the proof of Theorem 3.3 in [14]; we will not give full details, and instead encourage the interested reader to consult [14]. The Cauchy-Schwarz inequality and the fact that $\|f_n\|_{L^\infty} \leq 1$ imply that

$$\begin{aligned}|\mathbb{E}\langle f_n \widetilde{\mathcal{L}}_k(\sigma^1) \rangle - \mathbb{E}\langle f_n \rangle \mathbb{E}\langle \widetilde{\mathcal{L}}_k(\sigma) \rangle| &\leq \mathbb{E}\langle (f_n - \langle f_n \rangle)^2 \rangle^{1/2} \mathbb{E}\langle (\widetilde{\mathcal{L}}_k - \mathbb{E}\langle \widetilde{\mathcal{L}}_k \rangle)^2 \rangle^{1/2} \\ &\leq \mathbb{E}\langle (\widetilde{\mathcal{L}}_k - \mathbb{E}\langle \widetilde{\mathcal{L}}_k \rangle)^2 \rangle^{1/2}.\end{aligned}$$

By Lemma C.4 it is therefore sufficient to prove that

$$\mathbb{E}\langle f_n \widetilde{\mathcal{L}}_k(\sigma^1) \rangle = \mathbb{E} \frac{\langle f_n d_{ik}^1 \exp(\sum_{\ell \leq n} \theta_{ik}^\ell) \rangle}{\langle \exp(\theta_{ik}) \rangle^n} \quad \text{and} \quad \mathbb{E}\langle \widetilde{\mathcal{L}}_k(\sigma) \rangle = \mathbb{E} \frac{\langle d_{ik} \exp(\theta_{ik}) \rangle}{\langle \exp(\theta_{ik}) \rangle}. \quad (\text{C.18})$$

Since π_k is independent of all other sources of randomness, taking the expectation with respect to this random variable first shows that

$$\mathbb{E}\langle f_n \widetilde{\mathcal{L}}_k(\sigma^1) \rangle = \sum_{r \geq 1} \frac{s_N^{r-1}}{(r-1)!} \exp(-s_N) \mathbb{E}\langle f_n D_{1k}^1 \rangle_{\pi_k=r}, \quad (\text{C.19})$$

where $D_{1k}^1 = \sigma_{i_{1k}}^1 e_{1k} / (1 + \lambda_k \sigma_{i_{1k}}^*)^2$. To simplify this expression, we will isolate the first replica σ^1 appearing in each of the averages. It will be convenient to introduce the quantities

$$\Theta_{jk}^\ell = \log(1 + \lambda_k \sigma_{i_{jk}}^\ell) - \frac{\lambda_k e_{jk} \sigma_{i_{jk}}^\ell}{1 + \lambda_k \sigma_{i_{jk}}^*} \quad \text{and} \quad \mathcal{H}_k^{r-1}(\sigma^\ell) = \sum_{2 \leq j \leq r} \Theta_{jk}^\ell$$

as well as the partially perturbed Hamiltonian

$$H'_N(\boldsymbol{\sigma}) = H_N(\boldsymbol{\sigma}) + H_N^{\text{gauss}}(\boldsymbol{\sigma}) + \sum_{\substack{1 \leq k' \leq K_+ \\ k' \neq k}} \mathcal{H}_{k'}, \quad (\text{C.20})$$

where \mathcal{H}_k is defined in (4.5). Denoting by $\langle \cdot \rangle'_{\pi_k=r}$ the Gibbs measure corresponding to the Hamiltonian $H'_N(\boldsymbol{\sigma}) + \mathcal{H}_k^{r-1}(\boldsymbol{\sigma}^\ell)$, we have

$$\mathbb{E} \langle f_n D_{1k}^1 \rangle_{\pi_k=r} = \mathbb{E} \mathbb{E}_{i_{1k}} \mathbb{E}_{e_{1k}} \frac{\langle f_n D_{1k}^1 \exp(\sum_{\ell \leq n} \Theta_{1k}^\ell) \rangle'_{\pi_k=r}}{(\langle \exp(\Theta_{1k}) \rangle'_{\pi_k=r})^n} \quad (\text{C.21})$$

for each $r \geq 1$. Since the uniform random variable i_{1k} and the exponential random variable e_{1k} no longer appear in the Gibbs average $\langle \cdot \rangle'_{\pi_k=r}$, we may replace them by a uniform random variable $i \in \{1, \dots, N\}$ and an exponential random variable $e \sim \text{Exp}(1)$ independent of all other sources of randomness as in the statement of the theorem. To emphasize this change, we also replace D_{1k}^1 and Θ_{1k}^ℓ by $d_{ik}^1 = \sigma_i^1 e / (1 + \lambda_k \sigma_i^*)^2$ and $\theta_{ik}^\ell = \log(1 + \lambda_k \sigma_i^\ell) - \lambda_k \sigma_i^\ell e / (1 + \lambda_k \sigma_i^*)$, respectively. Notice that this matches the definitions in (C.12). In this new notation (C.21) reads

$$\mathbb{E} \langle f_n D_{1k}^1 \rangle_{\pi_k=r} = \mathbb{E} \frac{\langle f_n d_{ik}^1 \exp(\sum_{\ell \leq n} \theta_{ik}^\ell) \rangle'_{\pi_k=r}}{(\langle \exp(\theta_{ik}) \rangle'_{\pi_k=r})^n}.$$

Substituting this into (C.19) and making the change of variables $m = r - 1$ reveals that

$$\mathbb{E} \langle f_n \widetilde{\mathcal{L}}_k(\boldsymbol{\sigma}^1) \rangle = \sum_{m \geq 0} \frac{s_N^m}{m!} \exp(-s_N) \mathbb{E} \frac{\langle f_n d_{ik}^1 \exp(\sum_{\ell \leq n} \theta_{ik}^\ell) \rangle'_{\pi_k=m+1}}{(\langle \exp(\theta_{ik}) \rangle'_{\pi_k=m+1})^n}.$$

Notice that whenever $\pi_k = m + 1$, the Hamiltonian defining the Gibbs average $\langle \cdot \rangle'_{\pi_k=m+1}$ is given by $H'_N(\boldsymbol{\sigma}) + \mathcal{H}_k^m(\boldsymbol{\sigma})$. This Hamiltonian has the same distribution as the Hamiltonian in (B.1) defining the original Gibbs average $\langle \cdot \rangle_{\pi_k=m}$. It follows that

$$\mathbb{E} \langle f_n \widetilde{\mathcal{L}}_k(\boldsymbol{\sigma}^1) \rangle = \sum_{m \geq 0} \frac{s_N^m}{m!} \exp(-s_N) \mathbb{E} \frac{\langle f_n d_{ik}^1 \exp(\sum_{\ell \leq n} \theta_{ik}^\ell) \rangle_{\pi_k=m}}{(\langle \exp(\theta_{ik}) \rangle_{\pi_k=m})^n} = \mathbb{E} \frac{\langle f_n d_{ik}^1 \exp(\sum_{\ell \leq n} \theta_{ik}^\ell) \rangle}{\langle \exp(\theta_{ik}) \rangle^n}.$$

This is the first equality in (C.18). The second equality in (C.18) is obtained by taking $n = 1$ and $f_1 = 1$ in the first equality. This completes the proof. \blacksquare

Applying this result along a sequence of perturbation parameters (λ^N) satisfying (C.8) reveals that

$$\lim_{N \rightarrow \infty} \left| \mathbb{E} \frac{\langle f_n d_{ik}^1 \exp(\sum_{\ell \leq n} \theta_{ik}^\ell) \rangle_N}{\langle \exp(\theta_{ik}) \rangle_N^n} - \mathbb{E} \langle f_n \rangle_N \mathbb{E} \frac{\langle d_{ik} \exp(\theta_{ik}) \rangle_N}{\langle \exp(\theta_{ik}) \rangle_N} \right| = 0 \quad (\text{C.22})$$

for any $1 \leq k \leq K_+$ and any function f_n of finitely many spins on n replicas and of the signal $\boldsymbol{\sigma}^*$ with $\|f_n\|_{L^\infty} \leq 1$. Observing that the denominators $\langle \exp(\theta_{ik}) \rangle_N$ do not depend on the signal $\boldsymbol{\sigma}^*$, it is possible to use the Nishimori identity (2.3) to replace all occurrences of the signal $\boldsymbol{\sigma}^*$ in (C.22) by another replica. For convenience of notation we will denote this new replica by $\boldsymbol{\sigma}^\diamond$ to distinguish

it from the signal σ^* and at the same time not occupy any specific index. The equations in (C.22) now read

$$\lim_{N \rightarrow \infty} \left| \mathbb{E} \mathbb{E}_\diamond \frac{\langle f_n d_{ik}^1 \exp(\sum_{\ell \leq n} \theta_{ik}^\ell) \rangle_N}{\langle \exp(\theta_{ik}) \rangle_N^n} - \mathbb{E} \mathbb{E}_\diamond \langle f_n \rangle_N \mathbb{E} \mathbb{E}_\diamond \frac{\langle d_{ik} \exp(\theta_{ik}) \rangle_N}{\langle \exp(\theta_{ik}) \rangle_N} \right| = 0, \quad (\text{C.23})$$

where \mathbb{E}_\diamond denotes the Gibbs average with respect to the replica σ^\diamond only, the bracket $\langle \cdot \rangle_N$ denotes the Gibbs average with respect to all other standard replicas, the function f_n depends on finitely many spins on the n standard replicas and on σ^\diamond , and, with some abuse of notation,

$$y_{ik} = \frac{e}{1 + \lambda_k \sigma_i^\diamond}, \quad \theta_{ik}^\ell = \log(1 + \lambda_k \sigma_i^\ell) - \lambda_k y_{ik} \sigma_i^\ell, \quad \text{and} \quad d_{ik}^\ell = \frac{y_{ik} \sigma_i^\ell}{1 + \lambda_k \sigma_i^\diamond} \quad (\text{C.24})$$

with $\lambda_k = \lambda_k^N$ for $1 \leq k \leq K_+$. We now simplify (C.23) for functions f_n that do not depend on the spin coordinate indexed by 1. Introduce the collection of functions

$$F_n = \left\{ \text{functions } f_n \text{ of finitely many spins } \sigma_i^\ell, \sigma_i^\diamond \text{ with } 2 \leq i \leq N \right. \\ \left. \text{of the } n \text{ standard replicas } (\sigma^\ell)_{\ell \leq n} \text{ and the special replica } \sigma^\diamond \text{ with } \|f_n\|_{L^\infty} \leq 1 \right\} \quad (\text{C.25})$$

and the quantities

$$y_k = \frac{e}{1 + \lambda_k \sigma_1^\diamond}, \quad \theta_k^\ell = \log(1 + \lambda_k \sigma_1^\ell) - \lambda_k y_k \sigma_1^\ell, \quad \text{and} \quad d_k^\ell = \frac{y_k \sigma_1^\ell}{1 + \lambda_k \sigma_1^\diamond}. \quad (\text{C.26})$$

For functions $f_n \in F_n$, the symmetry between sites (C.5) and the fact that $i \in \{2, \dots, N\}$ with overwhelming probability in the limit, allow us to replace the uniform random index $i \in \{1, \dots, N\}$ by the index 1. The Franz de-Sanctis identities together with assumption (C.8) therefore have the following important implication.

Corollary C.5 (Asymptotic Franz-de Sanctis identities in inference). *If (C.8) holds, then for every $1 \leq k \leq K_+$ and all functions $f_n \in F_n$,*

$$\lim_{N \rightarrow \infty} \left| \mathbb{E} \mathbb{E}_\diamond \frac{\langle f_n d_k^1 \exp(\sum_{\ell \leq n} \theta_k^\ell) \rangle_N}{\langle \exp(\theta_k) \rangle_N^n} - \mathbb{E} \mathbb{E}_\diamond \langle f_n \rangle_N \mathbb{E} \mathbb{E}_\diamond \frac{\langle d_k \exp(\theta_k) \rangle_N}{\langle \exp(\theta_k) \rangle_N} \right| = 0. \quad (\text{C.27})$$

C.2 Finitary multi-overlap concentration

The finitary version of the multi-overlap concentration result in [14] will be uniform over an appropriate class of random probability measures which we now describe. For each integer $N \geq 1$ consider the set of random probability measures on Σ_N thought of as a subset of $\{-1, 0, 1\}^N$,

$$\mathcal{G}_N = \left\{ G \mid G \text{ is a random probability measure on } \Sigma_N \times \{0\}^N \right\}, \quad (\text{C.28})$$

and say that a measure $G \in \mathcal{G}_N$ satisfies symmetry between sites if, for any sequence of i.i.d. replicas $(\sigma^\ell)_{\ell \geq 1}$ sampled from G ,

$$(\sigma_i^\ell)_{i, \ell \geq 1} \stackrel{d}{=} (\sigma_{\rho_1(i)}^{\rho_2(\ell)})_{i, \ell \geq 1} \quad (\text{C.29})$$

for any permutation ρ_1 on the finite set $\{1, \dots, N\}$ and any permutation ρ_2 of finitely many indices. Denote by \mathcal{G}_N^s the subset of \mathcal{G}_N which satisfies symmetry between sites. Notice that each Gibbs measure G_N defined by (C.4) can be thought of as an element of \mathcal{G}_N by setting $\sigma_i = 0$ when $i > N$ for any replica $\sigma \in \Sigma_N$ sampled from G_N . In this way, the symmetry between sites in (C.29) and (C.5) coincide, so in fact $G_N \in \mathcal{G}_N^s$. This identification also suggests that the appropriate notion of the multi-overlap (C.7) for a random probability measure $G \in \mathcal{G}_N$ should be

$$R_{\ell_1, \dots, \ell_n} = \frac{1}{N} \sum_{i \leq N} \sigma_i^{\ell_1} \dots \sigma_i^{\ell_n}, \quad (\text{C.30})$$

where (σ^ℓ) denotes a sequence of i.i.d. replicas sampled from the Gibbs measure G . Denoting by $\langle \cdot \rangle_G$ the average with respect to the random probability measure G we may now state the main result of this appendix.

Proposition C.6. *For every $\varepsilon > 0$ there exists $\delta > 0$ such that the following holds. Let $N \geq \lfloor \delta^{-1} \rfloor$ and $G \in \mathcal{G}_N$ be a random probability measure such that for all $1 \leq k \leq K_+ = \lfloor \delta^{-1} \rfloor$ and any function $f_n \in \mathbb{F}_n$,*

$$\mathbb{E} \langle (R_1 - \mathbb{E} \langle R_1 \rangle_G)^2 \rangle_G \leq \delta, \quad \mathbb{E} \langle (R_{1,2} - \mathbb{E} \langle R_{1,2} \rangle_G)^2 \rangle_G \leq \delta, \quad (\text{C.31})$$

$$\left| \mathbb{E} \mathbb{E}_\diamond \frac{\langle f_n d_k^1 \exp(\sum_{\ell \leq n} \theta_k^\ell) \rangle_G}{\langle \exp(\theta_k) \rangle_G^n} - \mathbb{E} \mathbb{E}_\diamond \langle f_n \rangle_G \mathbb{E} \mathbb{E}_\diamond \frac{\langle d_k \exp(\theta_k) \rangle_G}{\langle \exp(\theta_k) \rangle_G} \right| \leq \delta. \quad (\text{C.32})$$

Then for any $1 \leq m \leq \lfloor \varepsilon^{-1} \rfloor$, we have

$$\mathbb{E} \langle (R_{1, \dots, m} - \mathbb{E} \langle R_{1, \dots, m} \rangle_G)^2 \rangle_G \leq \varepsilon. \quad (\text{C.33})$$

The proof proceeds by contradiction and closely follows Section 3.5 and Section 3.7 of [14]. Suppose there exists some $\varepsilon > 0$ such that no matter the choice of $\delta > 0$, it is always possible to find some integer $N = N(\delta) \geq \lfloor \delta^{-1} \rfloor$ and some random probability measure $G = G(\delta) \in \mathcal{G}_N$ with

$$\mathbb{E} \langle (R_1 - \mathbb{E} \langle R_1 \rangle_G)^2 \rangle_G \leq \delta, \quad \mathbb{E} \langle (R_{1,2} - \mathbb{E} \langle R_{1,2} \rangle_G)^2 \rangle_G \leq \delta, \quad (\text{C.34})$$

$$\left| \mathbb{E} \mathbb{E}_\diamond \frac{\langle f_n d_k^1 \exp(\sum_{\ell \leq n} \theta_k^\ell) \rangle_G}{\langle \exp(\theta_k) \rangle_G^n} - \mathbb{E} \mathbb{E}_\diamond \langle f_n \rangle_G \mathbb{E} \mathbb{E}_\diamond \frac{\langle d_k \exp(\theta_k) \rangle_G}{\langle \exp(\theta_k) \rangle_G} \right| \leq \delta \quad (\text{C.35})$$

for any $1 \leq k \leq K_+ = \lfloor \delta^{-1} \rfloor$ and any $f_n \in \mathbb{F}_n$ for which there exists some $1 \leq m = m(\delta) \leq \lfloor \varepsilon^{-1} \rfloor$ with

$$\mathbb{E} \langle (R_{1, \dots, m} - \mathbb{E} \langle R_{1, \dots, m} \rangle_G)^2 \rangle_G > \varepsilon. \quad (\text{C.36})$$

Applying the Prokhorov theorem on the compact metric space $\{-1, 0, +1\}^{\mathbb{N}^2}$ and noticing that there are only finitely many choices for $m = m(\delta)$, it is possible to find a subsequence with $\delta \rightarrow 0$ along which the distribution of the array $(\sigma_i^\ell)_{i, \ell \geq 1}$ under the averaged Gibbs measure $\mathbb{E} \langle \cdot \rangle_{G(\delta)}$ converges in the sense of finite dimensional distributions and along which (C.34), (C.35) and (C.36) hold for every $k \geq 1$ and a fixed $1 \leq m \leq \lfloor \varepsilon^{-1} \rfloor$. Since $N(\delta) \rightarrow \infty$ and $G(\delta) \in \mathcal{G}_N$, in the limit, the distribution of spins will be a measure on $\{-1, +1\}^{\mathbb{N}^2}$ which will inherit the symmetry between sites (C.29). By the Aldous-Hoover representation (see Theorem 1.4 in [60]), this symmetry implies the existence of some function $\sigma : [0, 1]^4 \rightarrow \{-1, +1\}$ with

$$(\sigma_i^\ell)_{i, \ell \geq 1} \stackrel{d}{=} (\sigma(w, u_\ell, v_i, x_{i, \ell}))_{i, \ell \geq 1}, \quad (\text{C.37})$$

where w , (u_ℓ) , (v_i) and $(x_{i,\ell})$ are i.i.d. uniform random variables on $[0, 1]$. Since σ takes values in $\{-1, +1\}$, the distribution of the array (σ_i^ℓ) is encoded by the function

$$\bar{\sigma}(w, u, v) = \mathbb{E}_{x_{i,\ell}} \sigma(w, u, v, x_{i,\ell}) = \int_0^1 \sigma(w, u, v, x) dx. \quad (\text{C.38})$$

Indeed, the last coordinate $x_{i,\ell}$ is a dummy variable that corresponds to flipping a biased coin to generate a Bernoulli random variable with expected value $\bar{\sigma}(w, u, v)$. To clarify this further, let du and dv denote Lebesgue measure on $[0, 1]$, and define the random probability measure

$$G = G_w = du \circ (u \mapsto \bar{\sigma}(w, u, \cdot))^{-1} \quad (\text{C.39})$$

on the space of functions of $v \in [0, 1]$,

$$H = L^2([0, 1], dv) \cap \{\|\bar{\sigma}\|_{L^\infty} \leq 1\}, \quad (\text{C.40})$$

equipped with the topology of $L^2([0, 1], dv)$. As described in Section 2 of [64], the whole process of generating spins can be broken into the following steps:

- (i) generate the asymptotic Gibbs measure $G = G_w$ using the uniform random variable w ;
- (ii) consider an i.i.d. sequence $\bar{\sigma}^\ell = \bar{\sigma}(w, u_\ell, \cdot)$ of replicas from G , which are functions in H ;
- (iii) plug in i.i.d. uniform random variables $(v_i)_{i \geq 1}$ to obtain the array $\bar{\sigma}^\ell(v_i) = \bar{\sigma}(w, u_\ell, v_i)$;
- (iv) finally, use the random variables $(x_{i,\ell})$ to generate (σ_i^ℓ) by flipping a coin with expected value $\bar{\sigma}^\ell(v_i)$,

$$\sigma_i^\ell = 2\mathbb{1}\left\{x_{i,\ell} \leq \frac{1 + \bar{\sigma}^\ell(v_i)}{2}\right\} - 1. \quad (\text{C.41})$$

This suggests that the asymptotic Gibbs average should be the average with respect to the random variables (u_ℓ) and $(x_{i,\ell})$ that depend on the replica indices, which we will denote by

$$\langle \cdot \rangle = \mathbb{E}_{(u_\ell), (x_{i,\ell})}. \quad (\text{C.42})$$

We can also expect the asymptotic multi-overlap to be given by

$$R_{\ell_1, \dots, \ell_n}^\infty(w, (u_{\ell_j})_{j \leq n}) = \mathbb{E}_v \prod_{j \leq n} \bar{\sigma}(w, u_{\ell_j}, v) = \int_0^1 \prod_{j \leq n} \bar{\sigma}(w, u_{\ell_j}, v) dv. \quad (\text{C.43})$$

This intuition is confirmed by the two following results adapted from Section 3.5 and the Appendix in [14].

Lemma C.7. *For any finite set of n replicas and every collection $\{\mathcal{C}_\ell\}_{\ell \leq n}$ of finite indices,*

$$\lim_{\delta \rightarrow 0} \mathbb{E} \prod_{\ell \leq n} \left\langle \prod_{i \in \mathcal{C}_\ell} \sigma_i^\ell \right\rangle_{G(\delta)} = \mathbb{E}_{w, (v_i)} \prod_{\ell \leq n} \left\langle \prod_{i \in \mathcal{C}_\ell} \sigma_i^\ell \right\rangle. \quad (\text{C.44})$$

Proof. Let $\mathcal{C} = \{(i, \ell) \mid \ell \leq n \text{ and } i \in \mathcal{C}_\ell\}$. By definition of weak convergence in the sense of finite-dimensional marginal distributions,

$$\lim_{\delta \rightarrow 0} \mathbb{E} \prod_{\ell \leq n} \left\langle \prod_{i \in \mathcal{C}_\ell} \sigma_i^\ell \right\rangle_{G(\delta)} = \lim_{\delta \rightarrow 0} \mathbb{E} \left\langle \prod_{(i, \ell) \in \mathcal{C}} \sigma_i^\ell \right\rangle_{G(\delta)} = \mathbb{E} \prod_{(i, \ell) \in \mathcal{C}} \sigma(w, u_\ell, v_i, x_{i, \ell}).$$

Recalling the notation (C.42),

$$\mathbb{E} \prod_{(i, \ell) \in \mathcal{C}} \sigma(w, u_\ell, v_i, x_{i, \ell}) = \mathbb{E}_{w, (v_i)} \prod_{\ell \leq n} \mathbb{E}_{(u_\ell)} \prod_{i \in \mathcal{C}_\ell} \mathbb{E}_{(x_{i, \ell})} \sigma(w, u_\ell, v_i, x_{i, \ell}) = \mathbb{E}_{w, (v_i)} \prod_{\ell \leq n} \left\langle \prod_{i \in \mathcal{C}_\ell} \sigma_i^\ell \right\rangle.$$

as required. ■

Lemma C.8. *For any collection of sets $\{\mathcal{L}_i\}_{i \geq 1}$ only finitely many of which are non-empty,*

$$\lim_{\delta \rightarrow 0} \mathbb{E} \left\langle \prod_{i \geq 1} R_{\mathcal{L}_i} \right\rangle_{G(\delta)} = \mathbb{E}_w \left\langle \prod_{i \geq 1} R_{\mathcal{L}_i}^\infty \right\rangle. \quad (\text{C.45})$$

Proof. Write $N \geq \lfloor \delta^{-1} \rfloor$ for the unique integer with $G(\delta) \in \mathcal{G}_N$ and suppose without loss of generality that the sets \mathcal{L}_i for $i \leq j$ are non-empty while the sets \mathcal{L}_i for $i > j$ are empty. From (C.30),

$$\mathbb{E} \left\langle \prod_{i \geq 1} R_{\mathcal{L}_i} \right\rangle_{G(\delta)} = \frac{1}{N^j} \sum_{i_1, \dots, i_j} \mathbb{E} \left\langle \prod_{\ell_1 \in \mathcal{L}_1} \dots \prod_{\ell_j \in \mathcal{L}_j} \sigma_{i_1}^{\ell_1} \dots \sigma_{i_j}^{\ell_j} \right\rangle_{G(\delta)}.$$

The number of terms in this sum for which at least two of the indices i_1, \dots, i_j are equal is of order N^{j-1} and is therefore negligible in the limit. Moreover, the symmetry between sites (C.29) implies that whenever i_1, \dots, i_j are all distinct

$$\mathbb{E} \left\langle \prod_{\ell_1 \in \mathcal{L}_1} \dots \prod_{\ell_j \in \mathcal{L}_j} \sigma_{i_1}^{\ell_1} \dots \sigma_{i_j}^{\ell_j} \right\rangle_{G(\delta)} = \mathbb{E} \left\langle \prod_{\ell_1 \in \mathcal{L}_1} \dots \prod_{\ell_j \in \mathcal{L}_j} \sigma_1^{\ell_1} \dots \sigma_j^{\ell_j} \right\rangle_{G(\delta)} = \mathbb{E} \left\langle \prod_{i \geq 1} \prod_{\ell \in \mathcal{L}_i} \sigma_i^\ell \right\rangle_{G(\delta)}.$$

(This seems to fix a small typo in the second-to-last display of the Appendix in [14]). Combining these two observations shows that

$$\lim_{\delta \rightarrow 0} \mathbb{E} \left\langle \prod_{i \geq 1} R_{\mathcal{L}_i} \right\rangle_{G(\delta)} = \mathbb{E}_{w, (u_\ell)} \prod_{i \geq 1} \mathbb{E}_{v_i} \prod_{\ell \in \mathcal{L}_i} \mathbb{E}_{x_{i, \ell}} \sigma(w, u_\ell, v_i, x_{i, \ell}) = \mathbb{E}_{w, (u_\ell)} \prod_{i \geq 1} R_{\mathcal{L}_i}^\infty.$$

This completes the proof. ■

In the notation of (C.42) and (C.43), the asymptotic version of (C.34) and (C.35) therefore reads

$$\mathbb{E} \langle (R_1^\infty)^2 \rangle = (\mathbb{E} \langle R_1^\infty \rangle)^2, \quad \mathbb{E} \langle (R_{1,2}^\infty)^2 \rangle = (\mathbb{E} \langle R_{1,2}^\infty \rangle)^2 \quad (\text{C.46})$$

$$\mathbb{E} \mathbb{E}_\diamond \frac{\langle f_n d_k^1 \exp(\sum_{\ell \leq n} \theta_k^\ell) \rangle}{\langle \exp(\theta_k) \rangle^n} = \mathbb{E} \mathbb{E}_\diamond \langle f_n \rangle \mathbb{E} \mathbb{E}_\diamond \frac{\langle d_k \exp(\theta_k) \rangle}{\langle \exp(\theta_k) \rangle} \quad (\text{C.47})$$

for any $k \geq 1$ and any function $f_n \in \mathcal{F}_n$, while the asymptotic version of (C.36) becomes

$$\mathbb{E} \langle (R_{1, \dots, m}^\infty - \mathbb{E} \langle R_{1, \dots, m}^\infty \rangle)^2 \rangle > \varepsilon \quad (\text{C.48})$$

for some $1 \leq m \leq \lfloor \varepsilon^{-1} \rfloor$. We now derive the two most important consequences of the identities in (C.46) and (C.47) that will allow us to establish multi-overlap concentration. On the one hand, the concentration of the overlap $\mathbb{E}\langle (R_{1,2}^\infty)^2 \rangle = (\mathbb{E}\langle R_{1,2}^\infty \rangle)^2$ implies that the system lies in a ‘‘thermal pure state’’ and that the function $\bar{\sigma}(w, u, v)$ is therefore almost surely independent of u . The proof of this fact is taken from Theorem 3.1 in [62].

Lemma C.9. *If $\mathbb{E}\langle (R_{1,2}^\infty)^2 \rangle = (\mathbb{E}\langle R_{1,2}^\infty \rangle)^2$, then for almost all $u, v, w \in [0, 1]$,*

$$\bar{\sigma}(w, u, v) = \mathbb{E}_u \bar{\sigma}(w, u, v). \quad (\text{C.49})$$

Proof. Denote by \cdot the inner product on the Hilbert space (C.40),

$$\bar{\sigma}^1 \cdot \bar{\sigma}^2 = \mathbb{E}_v \bar{\sigma}^1(w, u_1, v) \bar{\sigma}^2(w, u_2, v) = R_{1,2}^\infty,$$

and observe that

$$0 = \mathbb{E}_{w, (u_\ell)} (R_{1,2}^\infty)^2 - \mathbb{E}_{w, (u_\ell)} R_{1,2}^\infty R_{3,4}^\infty = \mathbb{E}_w \text{Var}_{(u_\ell)} \bar{\sigma}^1 \cdot \bar{\sigma}^2.$$

It follows that for almost all $w \in [0, 1]$, the inner product $\bar{\sigma}^1 \cdot \bar{\sigma}^2$ of any two replicas sampled from the Gibbs measure G_w is constant almost surely. In other words, the measure G_w is concentrated on a single function which may depend on w . This completes the proof. \blacksquare

The second identity in (C.46) therefore implies that, instead of the equality in distribution (C.37), we actually have

$$(\sigma_i^\ell)_{i, \ell \geq 1} \stackrel{d}{=} (\sigma(w, v_i, x_{i, \ell}))_{i, \ell \geq 1} \quad (\text{C.50})$$

for any function $\sigma : [0, 1]^3 \rightarrow \{-1, +1\}$ such that $\int_0^1 \sigma(w, v, x) dx = \bar{\sigma}(w, v)$. In particular, the Gibbs average (C.42) simplifies to

$$\langle \cdot \rangle = \mathbb{E}_{(x_i, \ell)} \quad (\text{C.51})$$

while the multi-overlap (C.43) becomes

$$R_{\ell_1, \dots, \ell_n}^\infty(w) = \mathbb{E}_v \prod_{j \leq n} \bar{\sigma}(w, v) = \mathbb{E}_v (\bar{\sigma}(w, v)^n) = \int_0^1 \bar{\sigma}(w, v)^n dv. \quad (\text{C.52})$$

On the other hand, the asymptotic Franz-de Sanctis identities in (C.47) imply the following decoupling property of the asymptotic Gibbs measure. This is lemma 3.5 in [14].

Lemma C.10 (A decoupling lemma). *Fix $\lambda \in \{\lambda_k \mid k \geq 1\}$. If e_1, e_2 are independent $\text{Exp}(1)$ random variables and, for $j = 1, 2$,*

$$y_j = \frac{e_j}{1 + \lambda \sigma_j^\diamond}, \quad \theta_j = \log(1 + \lambda \sigma_j) - \lambda y_j \sigma_j, \quad \text{and} \quad d_j = \frac{y_j \sigma_j}{1 + \lambda \sigma_j^\diamond}, \quad (\text{C.53})$$

then

$$\mathbb{E} \mathbb{E}_\diamond \frac{\langle d_1 \exp(\theta_1) d_2 \exp(\theta_2) \rangle}{\langle \exp(\theta_1) \exp(\theta_2) \rangle} = \mathbb{E} \mathbb{E}_\diamond \frac{\langle d_1 \exp(\theta_1) \rangle}{\langle \exp(\theta_1) \rangle} \mathbb{E} \mathbb{E}_\diamond \frac{\langle d_2 \exp(\theta_2) \rangle}{\langle \exp(\theta_2) \rangle}. \quad (\text{C.54})$$

Proof. The idea is to combine a technical truncation argument with the Weierstrass approximation theorem and an application of (C.47) to the function $f_{n+1} = d_2^1 \exp \sum_{\ell \leq n+1} \theta_2^\ell$; we encourage the interested reader to consult [14]. \blacksquare

We are finally in a position to contradict (C.48) and prove Proposition C.6. The calculations are very similar in spirit to those in [63, 64], and are taken from Theorem 2.2 in [14].

Proof of Proposition C.6. We follow the proof of Theorem 2.2 in [14]; we will not give full details, and instead encourage the interested reader to consult [14]. Recall from (C.50) that $\sigma_j = \sigma(w, v_j, x_j)$ and $\sigma_j^\diamond = \sigma(w, v_j, x_j^\diamond)$. Since all random variables indexed by $j = 1, 2$ are independent, if we denote by $\mathbb{E}_{|w} = \mathbb{E}_{(e_j), (v_j), x_j, x_j^\diamond}$ the conditional expectation given w and we introduce the random variable

$$Y(w) = \mathbb{E}_{|w} \frac{\langle d_1 \exp(\theta_1) \rangle}{\langle \exp(\theta_1) \rangle} = \mathbb{E}_{|w} \frac{y_1}{1 + \lambda \sigma_1^\diamond} \frac{\langle \sigma_1 \exp(\theta_1) \rangle}{\langle \exp(\theta_1) \rangle},$$

which depends implicitly on λ through y_1 and θ_1 , then (C.54) reads

$$\mathbb{E} \left(\mathbb{E}_{|w} \frac{\langle d_1 \exp(\theta_1) \rangle}{\langle \exp(\theta_1) \rangle} \right) \left(\mathbb{E}_{|w} \frac{\langle d_2 \exp(\theta_2) \rangle}{\langle \exp(\theta_2) \rangle} \right) - \left(\mathbb{E} \frac{\langle d_1 \exp(\theta_1) \rangle}{\langle \exp(\theta_1) \rangle} \right) \left(\mathbb{E} \frac{\langle d_2 \exp(\theta_2) \rangle}{\langle \exp(\theta_2) \rangle} \right) = \mathbb{E} \text{Var}_{|w} Y(w) = 0.$$

This means that $Y = \mathbb{E}Y$ almost surely. To exploit this fact, through a slight abuse of notation, write σ for σ_1 and observe that conditionally on σ_1^\diamond ,

$$Y(w) = \mathbb{E}_{|w} \int_0^\infty \langle \exp(-\lambda \sigma y) \rangle \frac{\langle \sigma(1 + \lambda \sigma) \exp(-\lambda y \sigma) \rangle}{\langle (1 + \lambda \sigma) \exp(-\lambda y \sigma) \rangle} y \exp(-y) dy.$$

Using the analyticity of both

$$g_w : \gamma \mapsto g_w(\gamma) = \mathbb{E}_{|w} \int_0^\infty \langle \exp(-\gamma \sigma y) \rangle \frac{\langle \sigma(1 + \gamma \sigma) \exp(-\gamma y \sigma) \rangle}{\langle (1 + \gamma \sigma) \exp(-\gamma y \sigma) \rangle} y \exp(-y) dy$$

for a fixed w as well as its w -expectation $\mathbb{E} g_w(\gamma)$, it is possible to deduce that $Y(w) = \mathbb{E}Y$ for all λ in a small neighbourhood of the origin. With this in mind, introduce the random variable

$$Z(w) = \mathbb{E}_{|w} \int_0^\infty \langle \sigma(1 + \lambda \sigma) \exp(-\lambda \sigma) \rangle y \exp(-y) dy$$

which is deterministic by the first identity in (C.46). This implies that the random variable

$$X(w) = \frac{Z(w) - Y(w)}{\lambda} = \mathbb{E}_{|w} \int_0^\infty \langle \sigma \exp(-\lambda y \sigma) \rangle \frac{\langle \sigma(1 + \lambda \sigma) \exp(-\lambda y \sigma) \rangle}{\langle (1 + \lambda \sigma) \exp(-\lambda y \sigma) \rangle} y \exp(-y) dy$$

is deterministic for all λ in a small neighbourhood of the origin. In particular, all its λ -derivatives are also independent of w . We will now deduce from this observation that all multi-overlaps concentrate. Given $n \geq 1$, applying $\frac{\partial^n}{\partial \lambda^n}$ to the denominator in the expression inside the integral defining X and evaluating at $\lambda = 0$ yields the term

$$n! R_{1, \dots, n+2} \mathbb{E} e(e-1)^n,$$

where e is an $\text{Exp}(1)$ random variable. Since $\mathbb{E}e(e-1)^n = \mathbb{E}(e-1)^{n+1} + \mathbb{E}(e-1)^n > 0$ for all $n \geq 1$, the term obtained by applying all derivatives to the denominator in the expression inside the integral defining X produces the multi-overlap $R_{1,\dots,n+2}^\infty$. If along the way we apply a derivative of λ to any factor other than the denominator, this will not create a new replica, so all those terms will produce a linear combination of multi-overlaps on strictly less than $n+2$ replicas which by induction we assume to be independent of w . This establishes the concentration of all multi-overlaps and contradicts (C.48) thus completing the proof. ■

References

- [1] E. Abbe. *Community Detection and Stochastic Block Models*. Foundations and trends in communications and information theory. Now Publishers, 2018.
- [2] E. Abbe, A. S. Bandeira, and G. Hall. Exact recovery in the stochastic block model. *IEEE Transactions on Information Theory*, 62(1):471–487, 2016.
- [3] E. Abbe, E. Cornacchia, Y. Gu, and Y. 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.
- [4] E. Abbe and A. Montanari. Conditional random fields, planted constraint satisfaction, and entropy concentration. *Theory of Computing*, 11:413–443, 12 2015.
- [5] E. Abbe and C. Sandon. Detection in the stochastic block model with multiple clusters: proof of the achievability conjectures, acyclic BP, and the information-computation gap. *arXiv e-prints*, 2015.
- [6] E. Abbe and C. Sandon. Recovering communities in the general stochastic block model without knowing the parameters. In *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc., 2015.
- [7] E. M. Airoldi, D. M. Blei, S. E. Fienberg, and E. P. Xing. Mixed membership stochastic blockmodels. *Journal of Machine Learning Research*, 9(65):1981–2014, 2008.
- [8] B. Ball, B. Karrer, and M. E. J. Newman. Efficient and principled method for detecting communities in networks. *Physical Review E*, 84(3), 2011.
- [9] J. Barbier, C. L. Chan, and N. Macris. Mutual information for the stochastic block model by the adaptive interpolation method. In *2019 IEEE International Symposium on Information Theory*, page 405–409. IEEE Press, 2019.
- [10] J. Barbier, M. Dia, N. Macris, F. Krzakala, T. Lesieur, and L. 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.

- [11] J. Barbier and N. Macris. The adaptive interpolation method: a simple scheme to prove replica formulas in bayesian inference. *Probability Theory and Related Fields*, 174(3-4):1133–1185, 2019.
- [12] J. Barbier, N. Macris, and L. 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.
- [13] J. Barbier and D. Panchenko. Free energy in the stochastic block model in sparse regime: Disassortative case. *Personal communication*.
- [14] J. Barbier and D. Panchenko. Strong replica symmetry in high-dimensional optimal Bayesian inference. *Communications in Mathematical Physics*, 393:1199–1239, 2022.
- [15] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Struct. Algorithms*, 31(1):3–122, 2007.
- [16] R. B. Boppana. Eigenvalues and graph bisection: An average-case analysis. In *28th Annual Symposium on Foundations of Computer Science*, pages 280–285, 1987.
- [17] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities*. Oxford University Press, Oxford, 2013.
- [18] T. N. Bui, S. Chaudhuri, F. T. Leighton, and M. Sipser. Graph bisection algorithms with good average case behavior. *Combinatorica*, 7(2):171–191, 1987.
- [19] I. Cabreros, E. Abbe, and A. Tsirigos. Detecting community structures in Hi-C genomic data. In *Conference on Information Science and Systems*, pages 584–589, 2016.
- [20] F. Caltagirone, M. Lelarge, and L. Miolane. Recovering asymmetric communities in the stochastic block model. *IEEE Transactions on Network Science and Engineering*, 5(3):237–246, 2018.
- [21] H. Chen, J.-C. Mourrat, and J. Xia. Statistical inference of finite-rank tensors. *Annales Henri Lebesgue*, 5:1161–1189, 2022.
- [22] H.-B. Chen. Hamilton-Jacobi equations for nonsymmetric matrix inference. *Ann. Appl. Probab.*, 32(4):2540–2567, 2022.
- [23] H.-B. Chen and J. Xia. Limiting free energy of multi-layer generalized linear models. *Preprint arXiv:2108.12615*, 2021.
- [24] H.-B. Chen and J. Xia. Hamilton-Jacobi equations for inference of matrix tensor products. *Ann. Inst. Henri Poincaré Probab. Stat.*, 58(2):755–793, 2022.
- [25] H.-B. Chen and J. Xia. Hamilton-Jacobi equations from mean-field spin glasses. *arXiv e-prints*, 2022.
- [26] H.-B. Chen and J. Xia. Hamilton-Jacobi equations with monotone nonlinearities on convex cones. *arXiv e-prints*, 2022.

- [27] J. Chen and B. Yuan. Detecting functional modules in the yeast protein-protein interaction network. *Bioinformatics*, 22(18):2283–2290, 2006.
- [28] A. Coja-Oghlan, F. Krzakala, W. Perkins, and L. Zdeborová. Information-theoretic thresholds from the cavity method. *Advances in Mathematics*, 333:694–795, 2018.
- [29] A. Decelle, F. Krzakala, C. Moore, and L. Zdeborová. Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications. *Physical Review E*, 84:066106, 2011.
- [30] Y. Deshpande, E. Abbe, and A. Montanari. Asymptotic mutual information for the balanced binary stochastic block model. *Information and Inference: A Journal of the IMA*, 6(2):125–170, 2016.
- [31] T. Dominguez and J.-C. Mourrat. Infinite-dimensional Hamilton-Jacobi equations for statistical inference on sparse graphs. *arXiv e-prints*, 2022.
- [32] M. Dyer and A. Frieze. The solution of some random np-hard problems in polynomial expected time. *Journal of Algorithms*, 10(4):451–489, 1989.
- [33] L. Evans. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 2010.
- [34] S. E. Fienberg, M. M. Meyer, and S. S. Wasserman. Statistical analysis of multiple sociometric relations. *Journal of the American Statistical Association*, 80(389):51–67, 1985.
- [35] P. K. Gopalan and D. M. Blei. Efficient discovery of overlapping communities in massive networks. *Proc. Natl. Acad. Sci. USA*, 110(36):14534–14539, 2013.
- [36] Y. Gu and Y. Polyanskiy. Uniqueness of BP fixed point for the Potts model and applications to community detection. Preprint, arXiv:2303.14688.
- [37] P. W. Holland, K. B. Laskey, and S. Leinhardt. Stochastic blockmodels: First steps. *Social Networks*, 5(2):109–137, 1983.
- [38] J. Kadmon and S. Ganguli. Statistical mechanics of low-rank tensor decomposition. In *Advances in Neural Information Processing Systems*, pages 8201–8212, 2018.
- [39] V. Kanade, E. Mossel, and T. Schramm. Global and local information in clustering labeled block models. *IEEE Transactions on Information Theory*, 62(10):5906–5917, 2016.
- [40] J. Kingman. *Poisson Processes*. Oxford Studies in Probability. Clarendon Press, 1992.
- [41] A. Kireeva and J.-C. Mourrat. Breakdown of a concavity property of mutual information for non-gaussian channels. Preprint, arXiv:2304.05129.
- [42] M. Lelarge and L. Miolane. Fundamental limits of symmetric low-rank matrix estimation. *Probability Theory and Related Fields*, 173(3):859–929, 2019.

- [43] T. Lesieur, L. Miolane, M. Lelarge, F. Krzakala, and L. Zdeborová. Statistical and computational phase transitions in spiked tensor estimation. In *2017 IEEE International Symposium on Information Theory (ISIT)*, pages 511–515. IEEE, 2017.
- [44] G. Linden, B. Smith, and J. York. Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet Comput.*, 7:76–80, 2003.
- [45] C. Luneau, J. Barbier, and N. Macris. Mutual information for low-rank even-order symmetric tensor estimation. *Information and Inference: A Journal of the IMA*, 10(4):1167–1207, 2021.
- [46] C. Luneau, N. Macris, and J. 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.
- [47] L. 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.
- [48] V. Mayya and G. 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.
- [49] L. Miolane. Fundamental limits of low-rank matrix estimation: the non-symmetric case. *Preprint, arXiv:1702.00473*, 2017.
- [50] E. Mossel, J. Neeman, and A. Sly. Reconstruction and estimation in the planted partition model. *Probab. Theory Related Fields*, 162(3-4):431–461, 2015.
- [51] E. Mossel, J. Neeman, and A. Sly. Belief propagation, robust reconstruction and optimal recovery of block models. *The Annals of Applied Probability*, 26(4):2211–2256, 2016.
- [52] E. Mossel, J. Neeman, and A. Sly. Consistency thresholds for the planted bisection model. *Electronic Journal of Probability*, 21:1–24, 2016.
- [53] E. Mossel, J. Neeman, and A. Sly. A proof of the block model threshold conjecture. *Combinatorica*, 38(3):665–708, 2018.
- [54] E. Mossel and J. 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.
- [55] J.-C. Mourrat. Hamilton–Jacobi equations for finite-rank matrix inference. *The Annals of Applied Probability*, 30(5):2234–2260, 2020.
- [56] J.-C. Mourrat. Hamilton–Jacobi equations for mean-field disordered systems. *Annales Henri Lebesgue*, 4:453–484, 2021.
- [57] J.-C. Mourrat. Nonconvex interactions in mean-field spin glasses. *Probability and Mathematical Physics*, 2(2):61–119, 2021.

- [58] J.-C. Mourrat. Free energy upper bound for mean-field vector spin glasses. *Ann. Inst. Henri Poincaré Probab. Stat.*, to appear.
- [59] M. E. J. Newman, D. J. Watts, and S. H. Strogatz. Random graph models of social networks. *Proc. Natl. Acad. Sci. USA*, 99:2566–2572, 2002.
- [60] D. Panchenko. *The Sherrington-Kirkpatrick model*. Springer Monographs in Mathematics. Springer, New York, 2013.
- [61] D. Panchenko. Spin glass models from the point of view of spin distributions. *The Annals of Probability*, 41(3A):1315–1361, 2013.
- [62] D. Panchenko. On the replica symmetric solution of the K-sat model. *Electronic Journal of Probability*, 19(67):1–17, 2014.
- [63] D. Panchenko. Structure of 1-RSB asymptotic Gibbs measures in the diluted p -spin models. *Journal of Statistical Physics*, 155(1):1–22, 2014.
- [64] D. Panchenko. Structure of finite-RSB asymptotic Gibbs measures in the diluted spin glass models. *Journal of Statistical Physics*, 162(1):1–42, 2015.
- [65] D. Panchenko. *Lecture notes on probability theory*. 2019.
- [66] G. Reeves. Information-theoretic limits for the matrix tensor product. *IEEE Journal on Selected Areas in Information Theory*, 1(3):777–798, 2020.
- [67] G. Reeves, V. Mayya, and A. Volfovsky. The geometry of community detection via the mmse matrix. In *2019 IEEE International Symposium on Information Theory (ISIT)*, pages 400–404. IEEE, 2019.
- [68] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(8):888–905, 2000.
- [69] T. Sørlie, C. M. Perou, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. USA*, 98(19):10869–10874, 2001.
- [70] Y. J. Wang and G. Y. Wong. Stochastic blockmodels for directed graphs. *Journal of the American Statistical Association*, 82(397):8–19, 1987.
- [71] H. C. White, S. A. Boorman, and R. L. Breiger. Social structure from multiple networks. I. Blockmodels of roles and positions. *American Journal of Sociology*, 81(4):730–780, 1976.
- [72] Q. Yu and Y. Polyanskiy. Ising model on locally tree-like graphs: Uniqueness of solutions to cavity equations. Preprint, arXiv:2211.15242.