

Large deviations, a phase transition, and logarithmic Sobolev inequalities in the block spin Potts model*

Holger Knöpfel[†] Matthias Löwe[‡] Holger Sambale[§]

Abstract

We introduce and analyze a generalization of the blocks spin Ising (Curie-Weiss) models that were discussed in a number of recent articles. In these block spin models each spin in one of s blocks can take one of a finite number of $q \geq 3$ values or colors, hence the name block spin Potts model. We prove a large deviation principle for the percentage of spins of a certain color in a certain block. These values are represented in an $s \times q$ matrix. We show that for uniform block sizes there is a phase transition. In some regime the only equilibrium is the uniform distribution of all colors in all blocks, while in other parameter regimes there is one predominant color, and this is the same color with the same frequency for all blocks. Finally, we establish log-Sobolev-type inequalities for the block spin Potts model.

Keywords: block spin Potts model; large deviation principle; logarithmic Sobolev inequality; phase transition.

MSC2020 subject classifications: 60F10; 82B20.

Submitted to ECP on January 12, 2021, final version accepted on April 25, 2021.

1 Introduction

Mean-field models as the Curie–Weiss model are approximations of lattice models. They often show qualitatively interesting results (see [13] for a survey). Mean-field block models have been proposed as an approximation of lattice models for meta-magnets, see e. g. [24]. To describe them, assume that we have N interacting particles that carry a spin. Also assume that we can group these particles into several groups. The interaction is such that particles within the same group interact with one interaction strength, while particles in different groups have another, usually smaller, interaction strength. In a sequence of papers the statistical mechanics of such models was studied from various points of view, see [17, 15, 9, 30, 25, 27, 26]. In particular they were discussed as models for social interactions between several groups, e. g. in [16, 2, 34, 32] (the latter paper studies a combination of Ising models on Erdős–Rényi graphs as in [7, 21, 22] and block

*Research of the second author was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy EXC 2044-390685587, Mathematics Münster: Dynamics-Geometry-Structure. Research of the third author was funded by the Deutsche Forschungsgemeinschaft via the CRC 1283 *Taming uncertainty and profiting from randomness and low regularity in analysis, stochastics and their applications*.

[†]University of Münster, Germany. E-mail: Holger.knoepfel@rub.de

[‡]University of Münster, Germany. E-mail: maloeve@math.uni-muenster.de

[§]University of Bielefeld, Germany. E-mail: hsambale@math.uni-bielefeld.de

models). Recently, block models have also been studied in a statistical context (see [3], [31]). Here the task is to exactly recover the block structure from a given number of realizations of the model. It turns out that this can be done surprisingly effectively.

However, all the literature cited above deals with Ising spins, i. e. the spins take two values (usually ± 1). Of course, the physics literature knows many more spin models than just the Ising model, in particular models with a continuous spin as Heisenberg models and XY-models.

On the discrete side Potts models (cf. e. g. [38, 23, 14, 10]) are the most natural generalization of Ising models. For them each particle carries a spin from a finite set (of cardinality 3 or larger). The aim of the present note is to investigate block spin Potts models as a natural generalization of block spin Ising models. We will basically concentrate on models where the blocks have approximately identical size and where the interaction is purely ferromagnetic, i. e. particles tend to have the same spin, no matter, whether they are in the same block or in different ones. Similar to [30] and [27] our main tool are large deviation techniques. Indeed, as we will see in Section 3, it is not too difficult to establish a large deviation principle for the “block magnetizations”. However, to derive a limit theorem with an explicit limit law from there turns out to be more complicated than in the case of Ising spins (which is quite a common feature in Potts models).

The rest of this note is organized in the following way. In the next section we will describe the block spin Potts model. Section 3 contains a large deviation analysis of this model. In Section 4, we will concentrate on a version with blocks of asymptotically equal size and compute the possible limit laws for such models. Finally, in Section 5 we prove and briefly discuss (modified) logarithmic Sobolev inequalities for the block spin Potts model.

Let us mention at this point that, while we were finishing the current manuscript we learned that in [29] the author studies a very similar model: Here the number of blocks is restricted to two, but they may be of different size. His techniques, however are different from ours. Moreover, extending his results, we prove a large deviation principle, are able to locate the minima of the rate functions and show logarithmic Sobolev inequalities.

2 The model

In the sequel we will consider the following model. Take the set $S = \{1, \dots, N\}$ and partition S into s sets S_1, \dots, S_s . These sets will, of course, depend on N and we assume that the limits $\gamma_k := \lim_{N \rightarrow \infty} \frac{|S_k|}{N} \in (0, 1)$ exist (and, of course, $\sum_{k=1}^s \gamma_k = 1$). Moreover, take an integer $q \geq 3$ and for $\omega \in \{1, \dots, q\}^S$ and $0 < \alpha < \beta$ introduce the Hamiltonian $H_{N,\alpha,\beta}(\omega) := H_N(\omega) := -\frac{\beta}{2N} \sum_{i \sim j} \mathbb{1}_{\omega_i = \omega_j} - \frac{\alpha}{2N} \sum_{i \not\sim j} \mathbb{1}_{\omega_i = \omega_j}$. Here $i \sim j$ means that the indices i and j belong to the same block S_k (where the case $i = j$ is included) for some $k \in \{1, \dots, s\}$, while we write $i \not\sim j$, if this is not the case. With $H_{N,\alpha,\beta}$ we will associate the Gibbs measure

$$\mu_{N,\alpha,\beta}(\omega) := \mu_N(\omega) := \frac{\exp(-H_N(\omega))}{Z_{N,\alpha,\beta}}$$

where, of course, $Z_{N,\alpha,\beta} := Z_N := \sum_{\omega'} \exp(-H_N(\omega'))$. For $k \in \{1, \dots, s\}$ and $c \in \{1, \dots, q\}$ denote by $m_{k,c}$ the relative number of spins of “color” c in the block S_k , i. e.

$$m_{k,c} := m_{k,c}(\omega) := \frac{1}{|S_k|} \sum_{i \in S_k} \mathbb{1}_{\omega_i = c},$$

and set $M_N := (m_{k,c}) \in M(s \times q)$. Note that M_N is an order parameter of the model in the sense that the Hamiltonian is a function of M_N rather than ω .

Indeed, since $\mathbb{1}_{\omega_i=\omega_j} = \sum_{c=1}^q \mathbb{1}_{\omega_i=c} \mathbb{1}_{\omega_j=c}$ we have

$$\begin{aligned}
 -2NH_N(\omega) &= \beta \sum_{c=1}^q \sum_{k=1}^s \sum_{i \in S_k} \sum_{j \in S_k} \mathbb{1}_{\omega_i=c} \mathbb{1}_{\omega_j=c} + \alpha \sum_{c=1}^q \sum_{k \neq k'} \sum_{i \in S_k} \sum_{j \in S_{k'}} \mathbb{1}_{\omega_i=c} \mathbb{1}_{\omega_j=c} \\
 &= \text{tr}(B^t AB)
 \end{aligned} \tag{2.1}$$

where $A_{\alpha,\beta} := A \in M(s \times s)$ is the symmetric matrix with entries β on and α off the diagonal (the *block interaction matrix*) and $B \in M(s \times q)$ has entries

$$b_{k,c} = \sum_{i \in S_k} \mathbb{1}_{\omega_i=c} = |S_k| m_{k,c}. \tag{2.2}$$

Now A is positive definite for $0 < \alpha < \beta$ due to $x^t Ax = (\beta - \alpha)x^2 + \alpha(\sum_k x_k)^2$ for $x \in \mathbb{R}^s$. So using the (unique) positive definite symmetric matrix \sqrt{A} we see that $\text{tr}(B^t AB) = \text{tr}((\sqrt{A}B)^t(\sqrt{A}B)) = [\sqrt{A}B, \sqrt{A}B]$ denoting by $[\cdot, \cdot]$ the Frobenius scalar product. Hence the Hamiltonian is a positive definite quadratic form of the matrix B and we will write $\text{tr}(B^t AB) = \langle B, B \rangle_A$. Now introducing the diagonal matrix $\Gamma_N \in M(s \times s)$ given by $(\Gamma_N)_{k,k} = |S_k|$ we finally rewrite (2.1) as

$$H_N(\omega) = -\frac{1}{2N} \langle \Gamma_N M_N, \Gamma_N M_N \rangle_A. \tag{2.3}$$

It is therefore natural to study the distribution of M_N under the Gibbs measure μ_N .

3 A Large Deviation Principle for M_N

In this section we prove a Large Deviation Principle (LDP) for the matrix M_N . The analysis of the corresponding rate function will help us to determine the limiting behavior of M_N and to prove the existence of a phase transition. Let us briefly recall the definition of a large deviation principle (cf. [12] and [11] for a rich survey): For a Polish space \mathcal{X} and an increasing sequence of non-negative real numbers $(a_n)_{n \in \mathbb{N}}$ a sequence of probability measures $(\nu_n)_n$ on \mathcal{X} is said to satisfy an LDP with speed a_n and rate function $I : \mathcal{X} \rightarrow \mathbb{R}$ (a lower semi-continuous function with compact level sets $\{x : I(x) \leq L\}$ for all $L > 0$), if for all Borel sets $B \in \mathcal{B}(\mathcal{X})$ we have

$$-\inf_{x \in \text{int}(B)} I(x) \leq \liminf_{n \rightarrow \infty} \frac{\log \nu_n(B)}{a_n} \leq \limsup_{n \rightarrow \infty} \frac{\log \nu_n(B)}{a_n} \leq -\inf_{x \in \text{cl}(B)} I(x).$$

Here $\text{int}(B)$ and $\text{cl}(B)$ denote the topological interior and closure of a set B , respectively. A sequence of random variables $X_n : \Omega \rightarrow \mathcal{X}$ satisfies an LDP with speed a_n and rate function $I : \mathcal{X} \rightarrow \mathbb{R}$ under a sequence of measures μ_n if the push-forward sequence $\nu_n := \mu_n \circ X_n$ does.

Now if $M_N(k)$ denotes the row k of M_N for a fixed k , then under the uniform measure ρ on $\{1, \dots, q\}$ the vector $M_N(k)$ is the empirical vector of a $|S_k|$ -fold drawing from the alphabet $\{1, \dots, q\}$. Thus, under the uniform measure $\rho^{|S_k|}$ the vector $M_N(k)$ obeys an LDP with speed $|S_k|$ and a rate function that is given by the relative entropy of a probability measure $\nu \in \mathcal{M}^1(\{1, \dots, q\})$ with respect to ρ , $H(\nu|\rho) := \sum_{c=1}^q \nu(c) \log \frac{\nu(c)}{\rho(c)}$ (see e.g. [11, Th. 2.1.10] for a reference). Note that

$$H(\nu|\rho) = \sum_{c=1}^q \nu(c) \log \nu(c) + \sum_{c=1}^q \nu(c) \log q =: H(\nu) + \log q,$$

where $H(\nu)$ is the entropy of ν and we adopt the convention that $0 \log 0 = 0$.

Now $M_N(k)$, $k = 1, \dots, s$ are independent random vectors. Consequently, for $\bar{\rho}_N := \otimes_{k=1}^s \rho^{|S_k|}$,

$$\frac{1}{N} \log \bar{\rho}_N(M_N \in B) = \frac{1}{N} \sum_{k=1}^s \log \rho^{|S_k|}(M_N(k) \in B_k).$$

for any set $B = \prod_{k=1}^s B_k$ with Borel sets $B_k \subseteq \mathbb{R}^q$ (here we associate probabilities ν on the set $\{1, \dots, q\}$ with vectors in \mathbb{R}^q and define $H(\nu|\rho) = H(\nu) = \infty$, if $\nu \in \mathbb{R}^q$ does not have non-negative components summing to 1). Together with the above mentioned LDP for the components $M_N(k)$ and the assumption that $|S_k|/N$ converges to γ_k as $N \rightarrow \infty$, this observation implies that the matrix M_N under $\bar{\rho}_N$ obeys an LDP with speed N and rate function

$$I(\nu) := \sum_{k=1}^s \gamma_k H(\nu_k|\rho) = \log q + \sum_{k=1}^s \gamma_k H(\nu_k) \tag{3.1}$$

Here $\nu := (\nu_k)_{1 \leq k \leq s} \in M(s \times q)$ and the ν_k are probabilities on $\{1, \dots, q\}$, otherwise $I(\nu)$ is defined to be ∞ . Thus we have seen

Proposition 3.1. *Under the measure $\bar{\rho}_N$ the matrix valued random variable M_N obeys an LDP with speed N and rate function I given by (3.1).*

Proposition 3.1 together with the representation of our Hamiltonian in terms of the matrix M_N (2.3) immediately yields an LDP for M_N under the Gibbs measure μ_N .

Theorem 3.2. *Under the Gibbs measure $\mu_{\alpha,\beta,N}$ the matrix valued random variable M_N obeys an LDP with speed N and rate function $J = J_{\alpha,\beta}$*

$$\begin{aligned} J(\nu) &:= - \left[\frac{1}{2} \langle \Gamma \nu, \Gamma \nu \rangle_A - I(\nu) \right] + \sup_{\mu} \left[\frac{1}{2} \langle \Gamma \mu, \Gamma \mu \rangle_A - I(\mu) \right] \\ &= - \left[\frac{\beta}{2} \sum_{c=1}^q \sum_{k=1}^s \gamma_k^2 \nu_{k,c}^2 + \frac{\alpha}{2} \sum_{c=1}^q \sum_{k \neq k'} \gamma_k \gamma_{k'} \nu_{k,c} \nu_{k',c} - I(\nu) \right] \\ &\quad + \sup_{\mu} \left[\frac{\beta}{2} \sum_{c=1}^q \sum_{k=1}^s \gamma_k^2 \mu_{k,c}^2 + \frac{\alpha}{2} \sum_{c=1}^q \sum_{k \neq k'} \gamma_k \gamma_{k'} \mu_{k,c} \mu_{k',c} - I(\mu) \right]. \end{aligned} \tag{3.2}$$

Here Γ is the $s \times s$ diagonal matrix with $(\Gamma)_{kk} = \gamma_k$, $\nu := (\nu_k)_k$ and $\mu := (\mu_k)_k$ are $s \times q$ -matrices, and the ν_k and μ_k are probabilities on $\{1, \dots, q\}$, otherwise $J(\nu)$ is defined to be ∞ .

Proof. Starting from (2.3) we write $H_N(\omega) = -\frac{N}{2} \langle (N^{-1} \Gamma_N) M, (N^{-1} \Gamma_N) M \rangle_A$ and so the assumption that $|S_k|/N \rightarrow \gamma_k \forall k$, Proposition 3.1, together with Varadhan’s Lemma ([12, Th. III.13]) and the tilted LDP ([12, Th. III.17]) in the version of [27, Lemma 2.1] show the result. \square

It may be more convenient to reformulate the LDP above in terms of an LDP for the matrix M'_N with entries $m'_{k,c} := m'_{k,c}(\omega) := \frac{1}{N} \sum_{i \in S_k} \mathbb{1}_{\omega_i=c}$, so that asymptotically, $M'_N \approx \Gamma M_N$. Then, of course, the relevant matrices are the matrices $\nu' := \Gamma \nu = (\gamma_k \nu_{kc})_{kc}$ where ν are the matrices appearing in Theorem 3.2. Using (3.1) and calculating $\sum_{k=1}^s \sum_{c=1}^q \nu'_{k,c} \log \nu'_{k,c} = I(\nu) - \log q + H(\gamma)$ where of course $\gamma = (\gamma_k)$, we can identify the new rate function. Indeed, as the term $H(\gamma) - \log q$ is independent of ν' , we can reformulate the LDP as follows.

Theorem 3.3. *Under the Gibbs measure $\mu_{\alpha,\beta,N}$ the matrix valued random variable M'_N obeys an LDP with speed N and rate function*

$$J'(\nu) = \begin{cases} - \left[\frac{1}{2} \langle \nu, \nu \rangle_A - \sum_{k=1}^s \sum_{c=1}^q \nu_{k,c} \log \nu_{k,c} \right] \\ \quad + \sup_{\mu \in C(\gamma)} \left[\frac{1}{2} \langle \mu, \mu \rangle_A - \sum_{k=1}^s \sum_{c=1}^q \mu_{k,c} \log \mu_{k,c} \right] & \nu \in C(\gamma) \\ \infty & \nu \notin C(\gamma) \end{cases}$$

where $C(\gamma) = \{\mu = (\mu_{kc}) \in M(s \times q) : \mu_{kc} \geq 0 \text{ and } \sum_{c=1}^q \mu_{k,c} = \gamma_k \text{ for all } k\}$.

Note that every matrix $\mu \in C(\gamma)$ can actually be considered as a probability distribution on $[sq]$ and the term $-\sum_{k \in [s]} \sum_{c \in [q]} \mu_{k,c} \log(\mu_{k,c})$ is its entropy. However, the set $C(\gamma)$ places restrictions on the mass that can be placed on every block.

4 Equilibria for uniform block sizes

An LDP as in Theorem 3.2 or Theorem 3.3 is, in principle, able to determine the limit distributions of the matrix valued random variables M_N and M'_N under the sequence of Gibbs measure $\mu_{\alpha,\beta,N}$. Indeed, they are given by the minima of the corresponding rate functions.

Corollary 4.1. *The weak limit points of the sequence (M_N) under the measures $\mu_{N,\alpha,\beta}$ are given by the minima of $J(\cdot)$ and the weak limit points of the sequence (M'_N) under the measures $\mu_{N,\alpha,\beta}$ are given by the minima of $J'(\cdot)$*

Proof. This is folklore in large deviation theory and not difficult to prove, when realizing that the upper bound in an LDP implies that any measurable subset of \mathbb{R} whose closure does not contain a minimum of the rate function has a probability that converges to 0. \square

We will, in the sequel, determine the minima of J' which will also give us the minima of J . We start with the observation that the minimum points of J' are the maximum points of

$$\begin{aligned} G(\mu) &:= \frac{1}{2} \langle \mu, \mu \rangle_A - \sum_{k=1}^s \sum_{c=1}^q \mu_{k,c} \log \mu_{k,c} \\ &= \sum_{c=1}^q \sum_{k=1}^s \frac{\beta}{2} \mu_{kc}^2 + \sum_{c=1}^q \sum_{k' \neq k} \frac{\alpha}{2} \mu_{kc} \mu_{k'c} - \sum_{c=1}^q \sum_{k=1}^s \mu_{kc} \log \mu_{kc}, \end{aligned} \tag{4.1}$$

where $\mu = (\mu_{kc}) \in C(\gamma)$ and we have set $0 \log 0 := 0$.

Lemma 4.2. *G attains its maximum on the set*

$$C^+(\gamma) := \{\mu \in C(\gamma) : 0 < \mu_{kc} < 1 \text{ for all } 1 \leq k \leq s, 1 \leq c \leq q\}.$$

Proof. Suppose one of μ 's entries equals 0, without loss of generality $\mu_{11} = 0$. Then, there is $2 \leq i \leq q$ such that $\mu_{1i} \geq \gamma_1/(q-1)$. Note that G is the sum of a polynomial of degree two in the μ_{kc} 's and $-\sum_{k=1}^s \sum_{c=1}^q \mu_{kc} \log \mu_{kc}$. Now $-t \log t$ has derivative infinity at 0. Hence, for $\varepsilon > 0$ small enough, we have $G(\mu) < G(\mu')$ where μ' is the matrix that we obtain from μ , if we replace μ_{11} by $\mu'_{11} = \varepsilon$, μ_{1i} by $\mu'_{1i} = \mu_{1i} - \varepsilon$ and leave the other entries unaltered. \square

Let us now apply the method of Lagrange multipliers to find the maximum points of G . Let $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_s)$. We then need to find the critical points of $L(\mu, \lambda) = G(\mu) - \sum_{k=1}^s \lambda_k (\sum_{c=1}^q \mu_{kc} - \gamma_k)$. Differentiating with respect to the μ_{kc} , $1 \leq k \leq s$, $1 \leq c \leq q$ gives the following set of equations

$$0 = \partial_{\mu_{kc}} G(\mu) - \lambda_k = \beta \mu_{kc} + \alpha \sum_{\substack{k'=1 \\ k' \neq k}}^s \mu_{k'c} - \log \mu_{kc} - 1 - \lambda_k. \tag{4.2}$$

Summing these equations over all c yields $q(1 + \lambda_k) = \beta \gamma_k + \alpha \sum_{\substack{k'=1 \\ k' \neq k}}^s \gamma_{k'}$ and plugging this into (4.2) we finally arrive at our system of critical equations

$$\beta \left(\mu_{kc} - \frac{\gamma_k}{q} \right) + \alpha \sum_{\substack{k'=1 \\ k' \neq k}}^s \left(\mu_{k'c} - \frac{\gamma_{k'}}{q} \right) = \log \frac{\mu_{kc}}{\sqrt[q]{\prod_d \mu_{kd}}}. \tag{4.3}$$

that any maximum point has to solve. Let us rephrase the value of G in critical points by multiplying (4.3) by μ_{kc} and summing over c :

$$\sum_c (\beta \mu_{kc}^2 + \alpha \sum_{\substack{k'=1 \\ k' \neq k}}^s \mu_{kc} \mu_{k'c} - 2 \mu_{kc} \log \mu_{kc}) = \frac{\gamma_k}{q} (\beta \gamma_k + \alpha (1 - \gamma_k)) - \sum_c \left(\frac{\gamma_k}{q} + \mu_{kc} \right) \log \mu_{kc}.$$

Hence, if μ^{crit} is a critical point of G , we can write $G(\mu^{\text{crit}})$ as

$$G(\mu^{\text{crit}}) = \frac{1}{2q} ((\beta - \alpha) \|\gamma\|^2 + \alpha) - \frac{1}{2q} \sum_{k=1}^s \sum_{c=1}^q (\gamma_k + q \mu_{kc}^{\text{crit}}) \log \mu_{kc}^{\text{crit}}. \tag{4.4}$$

Next we will see that of all critical points only those where all rows of μ have the same, e. g. an increasing order, are relevant.

Lemma 4.3. For $\mu \in C^+(\gamma)$ with increasing rows $\mu_{k1} \leq \dots \leq \mu_{kq}$, $1 \leq k \leq s$, we have

$$\sum_{k' \neq k} \sum_{c=1}^q \mu_{kc} \mu_{k'c} \geq \sum_{k' \neq k} \sum_{c=1}^q \mu_{k\sigma_k(c)} \mu_{k'\sigma_{k'}(c)} \tag{4.5}$$

for all s -tuples $(\sigma_k)_k$ of permutations $\sigma_k \in S_q$.

In particular, for $\alpha > 0$, the function G can only be maximal in a point μ' , if the rows of μ' are ordered in the same way, i. e. if there is a $\sigma \in S_q$ such that $\mu'_{k\sigma(1)} \leq \dots \leq \mu'_{k\sigma(q)}$ for all $1 \leq k \leq s$.

Proof. Recall the rearrangement inequality [20, Th. 368]: When $x_1 \leq \dots \leq x_n$ and $y_1 \leq \dots \leq y_n$ are sequences of real numbers, then for every permutation $\pi \in S_n$ one has

$$\sum_{i=1}^n x_i y_{\pi(i)} \leq x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

and the inequality is strict if there are indices $j < j'$ with $x_j < x_{j'}$ and $y_{\pi(j)} > y_{\pi(j')}$. Applying this to row k and k' of μ gives $\sum_{c=1}^q \mu_{kc} \mu_{k'c} \geq \sum_{c=1}^q \mu_{k\sigma_k(c)} \mu_{k'\sigma_{k'}(c)}$ for every two permutations $\sigma_k, \sigma_{k'} \in S_q$. Summing over all $k' \neq k$ yields (4.5). In particular, μ' is not a maximum point, if we have two rows $k \neq k'$ and indices $j < j'$ with $\mu'_{kj} < \mu'_{k'j}$ and $\mu'_{k'j} > \mu'_{kj}$. \square

So we can and will assume in the following, that all rows of a critical point μ of G are increasing. The next lemma determines the structure of a critical μ .

Lemma 4.4. Let μ be a critical point of G .

1. If $\mu_{kc} = \mu_{kc'}$ for a k and $c \neq c'$ then $\mu_{k'c} = \mu_{k'c'}$ for all $1 \leq k' \leq s$.
2. Each row of μ has at most two different entries.

Proof. Subtracting (4.3) for c from the equation for c' yields $\alpha \sum_{k' \neq k} \mu_{k'c} = \alpha \sum_{k' \neq k} \mu_{k'c'}$ and thus, by increasing order, $\mu_{k'c} = \mu_{k'c'}$ for all $1 \leq k' \leq s$. This is the first claim.

For two different columns $c \neq c'$ we obtain from (4.3)

$$(\beta - \alpha)(\mu_{kc'} - \mu_{kc}) + \alpha \sum_{k'=1}^s (\mu_{k'c'} - \mu_{k'c}) = \log \mu_{kc'} - \log \mu_{kc}.$$

Now if we had three columns $c < c' < c''$ with $\mu_{kc} < \mu_{kc'} < \mu_{kc''}$ for one row k (and hence for all due to the first part of this lemma) we would have

$$\frac{1}{\alpha} = \sum_{k=1}^s \frac{\mu_{kc'} - \mu_{kc}}{\alpha \sum_{k'=1}^s (\mu_{k'c'} - \mu_{k'c})} = \sum_{k=1}^s \left(\frac{\log \mu_{kc'} - \log \mu_{kc}}{\mu_{kc'} - \mu_{kc}} - (\beta - \alpha) \right)^{-1} \tag{4.6}$$

and the same equation for the pair c, c'' . However, every summand on the right of that latter equation would be larger (and positive) than the corresponding summand in (4.6) by concavity of the logarithm. Hence we have a contradiction. \square

So far we have proved, that according to Lemma 4.2, Lemma 4.3 and Lemma 4.4 we can constrain our search for maximum points of G to matrices μ with positive entries, increasingly ordered rows and having at most two different columns. Taking into account that the entries in row k sum up to γ_k we see that the largest column μ^+ (component by component) of μ together with the number $1 \leq r \leq q$ of columns equal to μ^+ is all the information we need to build up μ . So either μ has q identical columns γ/q or (the increasingly ordered) μ reads

$$\mu = (\underbrace{\mu^- \dots \mu^-}_{q-r} \underbrace{\mu^+ \dots \mu^+}_r)$$

with $\mu^- = (\gamma - r\mu^+)/ (q - r)$ for some $1 \leq r \leq q - 1$ and $\mu_k^+ > \gamma_k/q > \mu_k^-$ for all k . At this stage we do not know how to proceed with the case of an arbitrary γ . However, a proof for the case of asymptotically equal block sizes $\gamma_k = 1/s$ for all k is readily accomplished.

Proposition 4.5. *Let $\gamma_k = 1/s$ for all $1 \leq k \leq s$ and $Q \in M(s \times q)$ with identical entries $1/sq$. If a critical μ with $G(Q) \leq G(\mu)$ does not have identical rows then μ is not a maximum point.*

Proof. Recall that the system of critical equations (4.3) reads

$$\alpha \sum_{k'=1}^s \left(\mu_{k'}^+ - \frac{1}{sq} \right) = \frac{q-r}{q} \log \frac{s(q-r)\mu_k^+}{1-sr\mu_k^+} - (\beta - \alpha) \left(\mu_k^+ - \frac{1}{sq} \right)$$

and regard the right hand side of this equation as a function of μ_k^+ , say $\psi(\mu_k^+)$. Now if the largest entry of the vector μ^+ occurs in line K (and perhaps somewhere else) but not in every line, then $\psi(\mu_K^+) - \alpha \left(s\mu_K^+ - \frac{1}{q} \right) < 0$. Since $\psi(t)$ diverges to $+\infty$ when t approaches $1/sr$ from below we can find a $t_0 > \mu_K^+$ with $\psi(t_0) = \alpha(st_0 - 1/q)$. Now building a matrix ν by taking instead of r columns equal to μ^+ just r columns with identical entries t_0 and completing the matrix with $q - r$ columns with identical entries $(1 - srt_0)/(s(q - r))$ then clearly ν is (well defined and) a critical point. So we just have to prove that $G(\mu) < G(\nu)$.

To that end observe that according to (4.4) the value of G in critical points p is given up to constants as $\sum_k w(p_k^+)$ with $w : (0, 1/(sr)) \rightarrow \mathbb{R}$

$$w(x) = -((q - r) + q(1 - sr)x) \log \frac{1 - sr x}{s(q - r)} - r(1 + sqx) \log x.$$

Calculating $w'(x) = srq \log \frac{1-srx}{s(q-r)x} + r \frac{sqx-1}{x(1-srx)}$ and $w''(x) = \frac{r(sqx-1)(2sr x-1)}{x^2(sr x-1)^2}$ we see that if $q > 2r$ then the graph of w , coming from $+\infty$ at the vertical asymptotic line $x = 0$, has a saddle point at $x = 1/sq$ changing from bending to the left to bending to the right. It decreases to the second inflection point at $x = 1/(2sr)$, passes, now bending to the left again, the unique minimum point at say ξ and disappears to $+\infty$ approaching the vertical line at $x = 1/(sr)$. In this case clearly $w(\mu_k^+) \leq \max(w(\mu_K^+), w(1/sq))$ for all k since $1/sq < \mu_k^+ \leq \mu_K^+$. Now $w(\mu_K^+) < w(1/sq)$ would imply the contradiction $G(\mu) = \sum_k w(\mu_k^+) < sw(1/sq) = G(Q)$, so we have $w(1/sq) \leq w(\mu_K^+)$ and therefore $\mu_K^+ > \xi$. This means $w(t_0) > w(\mu_K^+)$ and thus $G(\nu) = sw(t_0) > sw(\mu_K^+) \geq G(\mu)$. If $q \leq 2r$ then w is strictly increasing on $[1/sq, 1/sr)$ so that $G(\nu) = sw(t_0) > \sum_k w(\mu_k^+) = G(\mu)$. \square

Wrapping up what we have seen, we state

Proposition 4.6. *Let $\gamma \in \mathbb{R}^s$ have identical entries $1/s$. The function $G =: G^{\text{bP}}$ in the block spin Potts model is maximal on $C(\gamma)$ if its rows are identical and equal to a maximizer of the corresponding target function*

$$G^{\text{P}}(v) = \frac{1}{2s}(\beta + (s - 1)\alpha) \sum_{c=1}^q v_c^2 - \sum_{c=1}^q v_c \log v_c \tag{4.7}$$

on the set $V := \{v = (v_1, \dots, v_q) \mid \sum_c v_c = 1, v_c > 0\}$ in the Potts model.

Proof. The maximum of G^{bP} on $C(\gamma)$ is attained on the subset of matrices with identical rows taken from $\{v/s \mid v \in V\}$ and the value is equal to $\frac{\beta+(s-1)\alpha}{2s} \sum_{c=1}^q v_c^2 - \sum_{c=1}^q v_c \log \frac{v_c}{s}$. However, up to a minus sign and ignoring the summand $\log s$ this is the free energy functional in a Potts model at inverse temperature $(\beta + (s - 1)\alpha)/s$. \square

The following theorem hence follows from the results in [23, 14], where the critical temperature and the behaviour of the Potts model is computed.

Theorem 4.7. *Consider the block spin Potts model in the asymptotically uniform case $\gamma_k = s^{-1}$. Denote by $\zeta_q := 2\frac{q-1}{q-2} \log(q - 1)$ the critical inverse temperature in the q -color Potts model, and let $g := \frac{\beta+(s-1)\alpha}{s}$. Then the q -color block spin Potts model has a phase transition. More precisely, if $g < \zeta_q$, then the distribution of M'_N under the Gibbs measure $\mu_{N,\alpha,\beta}$ concentrates in a unique point, the matrix with all entries identical to $1/(sq)$.*

To describe the “low temperature” behavior define the function $\varphi : [0, 1] \rightarrow \mathbb{R}^q$:

$$\varphi(t) := \left(\frac{1 + (q - 1)t}{sq}, \frac{1 - t}{sq}, \dots, \frac{1 - t}{sq} \right)$$

and let $u(g)$ be the largest solution of the equation $u = \frac{1 - e^{-gu}}{1 + (q-1)e^{-gu}}$. Let $n^1(g) := \varphi(u(g))$ and $n^i(g)$ be $n^1(g)$ with the i 'th and the first coordinate interchanged, $i = 2, \dots, q$. Let $\nu^i(g)$ be the matrix with all rows identical to $n^i(g)$ and Q be the matrix with all entries identical to $1/(qs)$.

Then, if $g > \zeta_q$ the distribution of M'_N under the Gibbs measure $\mu_{N,\alpha,\beta}$ concentrates in a (uniform) mixture of the Dirac measures in $\nu^1(g), \dots, \nu^q(g)$.

At $g = \zeta_q$ the limit points of M'_N under the Gibbs measure $\mu_{N,\alpha,\beta}$ are Q and $\nu^1(g), \dots, \nu^q(g)$.

5 Logarithmic Sobolev inequalities

In this section, we present logarithmic Sobolev inequalities (LSIs) for block spin Potts models. LSIs are frequently used e. g. in concentration of measure theory, where they form the core of the well-known entropy method (cf. the monographs [28, 6]). Recently, LSIs for various type of finite spin systems have been established ([18, 35]), a line of research we continue by considering block spin Potts models. For $\omega = (\omega_i)_{i \in S} \in \{1, \dots, q\}^S$ and $i \in S$, let $\omega_{i^c} := (\omega_j)_{j \neq i}$. Moreover, for any function $f : \{1, \dots, q\}^S \rightarrow \mathbb{R}$, we define a certain “difference operator” by

$$|\mathfrak{d}f|(\omega) = \left(\sum_{i \in S} \int (f(\omega) - f(\omega_{i^c}, \omega'_i))^2 d\mu_N(\omega'_i \mid \omega_{i^c}) \right)^{1/2},$$

where $\mu_N(\cdot \mid \omega_{i^c})$ denotes the regular conditional probability. The integrals may be regarded as a kind of “local variance” in the i -th coordinate. The difference operator \mathfrak{d} is a well-known object, and $\int |\mathfrak{d}f|^2 d\mu_N$ can be regarded as a Dirichlet form (cf. [18, Rem. 2.2]). Finally, for any non-negative function f , $\text{Ent}_{\mu_N}(f) := \int f \log(f) d\mu_N - \int f d\mu_N \log(\int f d\mu_N)$ denotes the entropy.

Theorem 5.1. Assume that $2q\beta e^\beta < 1$.

1. For N large enough, μ_N satisfies an LSI with respect to \mathfrak{d} , i. e. for any $f: \{1, \dots, q\}^S \rightarrow \mathbb{R}$,

$$\text{Ent}_{\mu_N}(f^2) \leq 2\sigma_1^2 \int |\mathfrak{d}f|^2 d\mu_N. \tag{5.1}$$

2. For N large enough and any function $f: \{1, \dots, q\}^S \rightarrow \mathbb{R}$ we have

$$\text{Ent}_{\mu_N}(e^f) \leq \sigma_2^2 \sum_{i \in S} \int \text{Cov}_{\mu_{(\cdot|\omega_{i^c})}}(f(\omega_{i^c}, \cdot), e^{f(\omega_{i^c}, \cdot)}) d\mu(\omega). \tag{5.2}$$

3. For N large enough and any function $f: \{1, \dots, q\}^S \rightarrow \mathbb{R}$ we have

$$\text{Ent}_{\mu_N}(e^f) \leq \frac{\sigma_3^2}{2} \int |\mathfrak{d}f|^2 e^f d\mu_N. \tag{5.3}$$

Here, $\sigma_1, \sigma_2, \sigma_3 > 0$ are constants which depend on β and q only.

See Remark 5.3 for explicit values for $\sigma_1, \sigma_2, \sigma_3$ and some notes about the condition on N . Both (5.2) and (5.3) are also known as modified logarithmic Sobolev inequalities.

To see how Theorem 5.1 implies concentration results, note that based on (5.1) we obtain L^p norm inequalities and concentration bounds for any $f \in L^\infty(\mu_N)$, cf. [18, Prop. 2.4, Th. 1.5]. Moreover, (5.3) implies subgaussian tails for Lipschitz functions f (i. e. $|\mathfrak{d}f| \leq 1$) via the Herbst argument, see e. g. [36] (including some more advanced situations, cf. Section 2.4). Recently, in [1] it has been established that (5.3) gives rise to L^p bounds (via Beckner inequalities) as well. Especially for spin systems and Glauber dynamics, cf. the discussion in Section 4.3 therein.

As a simple example, let $T_{k,c}(\omega) := \sum_{i \in S_k} \mathbb{1}_{\omega_i=c}$ denote the number of vertices in the block S_k which have colour c . Using $|\mathfrak{d}T_{k,c}|^2 \leq |S_k|$ and [36, Eq. (1.2)], we immediately obtain that $\mu_N(|T_{k,c} - \mu_N(T_{k,c})| \geq t) \leq 2 \exp\left(-\frac{t^2}{2|S_k|\sigma_3^2}\right)$, where $\mu_N(T_{k,c}) := \int T_{k,c} d\mu_N$. Note that for N large, this probability approaches $2 \exp(-t^2/(2N\gamma_k\sigma_3^2))$. A class of non-Lipschitz functions for which concentration bounds based on inequalities of type (5.1) and (5.3) have been established are multilinear (i. e. affine in each variable) polynomials, cf. e. g. [19, Th. 5], [1, Cor. 5.4].

Finally, note that (5.2) is frequently used in the context of Markov processes, and it is equivalent to exponential decay of the relative entropy along the Glauber semigroup (cf. e. g. [4, 8]). It moreover implies that the associated Glauber dynamics is rapidly mixing, i. e. its mixing time is $O(N \log N)$, see [35, Th. 2.2]. This complements [5], where a different situation was considered (the usual Potts model without blocks but on graphs with fixed maximal degree).

For the proof of Theorem 5.1, recall that for product measures $\mu = \otimes_{i=1}^n \mu_i$, the entropy functional tensorizes, i. e. $\text{Ent}_\mu(f) \leq \sum_{i=1}^n \int \text{Ent}_{\mu_i}(f) d\mu$, and therefore, proving LSIs reduces to controlling each coordinate separately, i. e. a “one-dimensional” case. For non-product measures, if the dependencies are sufficiently weak, an analogue can be shown which is called an *approximate tensorization property*. A criterion for approximate tensorization in probability spaces with finitely many atoms was introduced in [33], which we will exploit in the sequel.

Proposition 5.2. Assume that $2q\beta e^\beta < 1$. For N large enough, the approximate tensorization property of entropy holds, i. e.

$$\text{Ent}_{\mu_N}(f) \leq C \sum_{i \in S} \int \text{Ent}_{\mu_N(\cdot|\omega_{i^c})}(f(\omega_{i^c}, \cdot)) d\mu_N(\omega)$$

with C depending on β and q only.

Proof. We shall apply Marton’s approximate tensorization result [33] in the slightly rewritten and corrected form stated in [35, Th. 4.1]. Essentially, we need to control the conditional probabilities $\mu_N(\cdot|\omega_{i^c})$, which we rewrite in the sequel, generalizing the case of the usual Potts model without blocks as in [37, Prop. 2.16]. Recalling the matrix $B = (b_{k,c})$ from (2.2), we fix two sites $i, j \in S$ and define $b_{k,ij,c} := \sum_{\substack{\nu \in S_k \\ \nu \notin \{i,j\}}} \mathbb{1}_{\omega_\nu=c}$. Then, we may decompose

$$\begin{aligned}
 -H_N(\omega) &= \frac{\beta}{2N} \sum_{c=1}^q \sum_{k=1}^s (b_{k,ij,c} + \mathbb{1}_{i \in S_k, \omega_i=c} + \mathbb{1}_{j \in S_k, \omega_j=c})^2 \\
 &\quad + \frac{\alpha}{2N} \sum_{c=1}^q \sum_{k=1}^s \sum_{k':k' \neq k} (b_{k,ij,c} + \mathbb{1}_{i \in S_k, \omega_i=c} + \mathbb{1}_{j \in S_k, \omega_j=c}) \\
 &\quad \quad \quad (b_{k',ij,c} + \mathbb{1}_{i \in S_{k'}, \omega_i=c} + \mathbb{1}_{j \in S_{k'}, \omega_j=c}).
 \end{aligned}$$

Now, writing

$$\begin{aligned}
 (b_{k,ij,c} + \mathbb{1}_{i \in S_k, \omega_i=c} + \mathbb{1}_{j \in S_k, \omega_j=c})^2 &= b_{k,ij,c}^2 + 2b_{k,ij,c}(\mathbb{1}_{i \in S_k, \omega_i=c} + \mathbb{1}_{j \in S_k, \omega_j=c}) \\
 &\quad + \mathbb{1}_{i \in S_k, \omega_i=c} + \mathbb{1}_{j \in S_k, \omega_j=c} + 2 \cdot \mathbb{1}_{i \in S_k, \omega_i=c} \mathbb{1}_{j \in S_k, \omega_j=c}
 \end{aligned}$$

and summing over c and k , we obtain

$$\begin{aligned}
 &\sum_{c=1}^q \sum_{k=1}^s (b_{k,ij,c} + \mathbb{1}_{i \in S_k, \omega_i=c} + \mathbb{1}_{j \in S_k, \omega_j=c})^2 \\
 &= \sum_{c=1}^q \sum_{k=1}^s b_{k,ij,c}^2 + 2 \sum_{k=1}^s b_{k,ij,\omega_i} \mathbb{1}_{i \in S_k} + 2 \sum_{k=1}^s b_{k,ij,\omega_j} \mathbb{1}_{j \in S_k} + 2 + 2 \cdot \mathbb{1}_{i \sim j, \omega_i=\omega_j}.
 \end{aligned}$$

Similarly,

$$\begin{aligned}
 &(b_{k,ij,c} + \mathbb{1}_{i \in S_k, \omega_i=c} + \mathbb{1}_{j \in S_k, \omega_j=c})(b_{k',ij,c} + \mathbb{1}_{i \in S_{k'}, \omega_i=c} + \mathbb{1}_{j \in S_{k'}, \omega_j=c}) \\
 &= b_{k,ij,c} b_{k',ij,c} + b_{k,ij,c}(\mathbb{1}_{i \in S_{k'}, \omega_i=c} + \mathbb{1}_{j \in S_{k'}, \omega_j=c}) + b_{k',ij,c}(\mathbb{1}_{i \in S_k, \omega_i=c} + \mathbb{1}_{j \in S_k, \omega_j=c}) \\
 &\quad + (\mathbb{1}_{i \in S_k, \omega_i=c} + \mathbb{1}_{j \in S_k, \omega_j=c})(\mathbb{1}_{i \in S_{k'}, \omega_i=c} + \mathbb{1}_{j \in S_{k'}, \omega_j=c}),
 \end{aligned}$$

leading to

$$\begin{aligned}
 &\sum_{c=1}^q \sum_{k=1}^s \sum_{k':k' \neq k} (b_{k,ij,c} + \mathbb{1}_{i \in S_k, \omega_i=c} + \mathbb{1}_{j \in S_k, \omega_j=c})(b_{k',ij,c} + \mathbb{1}_{i \in S_{k'}, \omega_i=c} + \mathbb{1}_{j \in S_{k'}, \omega_j=c}) \\
 &= \sum_{c=1}^q \sum_{k=1}^s \sum_{k':k' \neq k} b_{k,ij,c} b_{k',ij,c} + 2 \sum_{k=1}^s b_{k,ij,\omega_i} \mathbb{1}_{i \notin S_k} + 2 \sum_{k=1}^s b_{k,ij,\omega_j} \mathbb{1}_{j \notin S_k} + 2 \cdot \mathbb{1}_{i \not\sim j, \omega_i=\omega_j}.
 \end{aligned}$$

Altogether, we arrive at the representation

$$\begin{aligned}
 -H_N(\omega) &= \frac{\beta}{2N} \sum_{c=1}^q \sum_{k=1}^s b_{k,ij,c}^2 + \frac{\beta}{N} \sum_{k=1}^s b_{k,ij,\omega_i} \mathbb{1}_{i \in S_k} + \frac{\beta}{N} \sum_{k=1}^s b_{k,ij,\omega_j} \mathbb{1}_{j \in S_k} + \frac{\beta}{N} \\
 &\quad + \frac{\beta}{N} \mathbb{1}_{i \sim j, \omega_i=\omega_j} + \frac{\alpha}{2N} \sum_{c=1}^q \sum_{k=1}^s \sum_{k':k' \neq k} b_{k,ij,c} b_{k',ij,c} \\
 &\quad + \frac{\alpha}{N} \sum_{k=1}^s b_{k,ij,\omega_i} \mathbb{1}_{i \notin S_k} + \frac{\alpha}{N} \sum_{k=1}^s b_{k,ij,\omega_j} \mathbb{1}_{j \notin S_k} + \frac{\alpha}{N} \mathbb{1}_{i \not\sim j, \omega_i=\omega_j}.
 \end{aligned}$$

In particular, the conditional probabilities given ω_{i^c} can be written as

$$\begin{aligned} \mu_N(c|\omega_{i^c}) &= \frac{\exp(-H_N(\omega_{i^c}, c))}{\sum_{c'=1}^q \exp(-H_N(\omega_{i^c}, c'))} = \frac{1}{1 + \sum_{c':c' \neq c} \exp(H_N(\omega_{i^c}, c) - H_N(\omega_{i^c}, c'))} \\ &= h\left(\sum_{c':c' \neq c} \exp(H_N(\omega_{i^c}, c) - H_N(\omega_{i^c}, c'))\right), \end{aligned}$$

where $h(x) = 1/(1+x)$, which is 1-Lipschitz. Using the representation derived above, we have

$$\begin{aligned} H_N(\omega_{i^c}, c) - H_N(\omega_{i^c}, c') &= \frac{\beta}{N} \sum_{k=1}^s (b_{k,ij,c'} - b_{k,ij,c}) \mathbb{1}_{i \in S_k} + \frac{\beta}{N} (\mathbb{1}_{\omega_j=c'} - \mathbb{1}_{\omega_j=c}) \mathbb{1}_{i \sim j} \\ &\quad + \frac{\alpha}{N} \sum_{k=1}^s (b_{k,ij,c'} - b_{k,ij,c}) \mathbb{1}_{i \notin S_k} + \frac{\alpha}{N} (\mathbb{1}_{\omega_j=c'} - \mathbb{1}_{\omega_j=c}) \mathbb{1}_{i \not\sim j} \\ &= \frac{1}{N} \sum_{k=1}^s (b_{k,ij,c'} - b_{k,ij,c}) (\beta \mathbb{1}_{i \in S_k} + \alpha \mathbb{1}_{i \notin S_k}) + \frac{1}{N} (\mathbb{1}_{\omega_j=c'} - \mathbb{1}_{\omega_j=c}) (\beta \mathbb{1}_{i \sim j} + \alpha \mathbb{1}_{i \not\sim j}). \end{aligned}$$

From this representation, we now derive two facts: first,

$$\min_{i \in S} \min_{\omega \in \{1, \dots, q\}^S} \mu_N(\omega_i | \omega_{i^c}) \geq \gamma_1 \tag{5.4}$$

for some $\gamma_1 > 0$ which only depends on β but not on N (here we have used that $\alpha < \beta$).

Moreover, we need to control the operator norm $\|J\|_{2 \rightarrow 2}$ of the $N \times N$ matrix J whose entries J_{ij} are given by $J_{ij} = \sup d_{TV}(\mu_N(\cdot | \omega_{i^c}), \mu_N(\cdot | \sigma_{i^c}))$, where the sup is taken over all configurations ω, σ which differ at site j only. Obviously, this can be upper bounded by the $\ell^\infty \rightarrow \ell^\infty$ norm $\|J\|_{\infty \rightarrow \infty}$. Fixing two such configurations, i. e. $\omega_j \neq \sigma_j, \omega_{j^c} = \sigma_{j^c}$, it follows that for any $i \neq j$,

$$\begin{aligned} d_{TV}(\mu_N(\cdot | \omega_{i^c}), \mu_N(\cdot | \sigma_{i^c})) &= \frac{1}{2} \sum_{c=1}^q |\mu_N(c | \omega_{i^c}) - \mu_N(c | \sigma_{i^c})| \\ &= \frac{1}{2} \sum_{c=1}^q \left| h\left(\sum_{c':c' \neq c} \exp(H_N(\omega_{i^c}, c) - H_N(\omega_{i^c}, c'))\right) \right. \\ &\quad \left. - h\left(\sum_{c':c' \neq c} \exp(H_N(\sigma_{i^c}, c) - H_N(\sigma_{i^c}, c'))\right) \right| \\ &\leq \frac{1}{2} \sum_{c=1}^q \left| \sum_{c':c' \neq c} \left(\exp(H_N(\omega_{i^c}, c) - H_N(\omega_{i^c}, c')) - \exp(H_N(\sigma_{i^c}, c) - H_N(\sigma_{i^c}, c')) \right) \right| \\ &\leq \frac{1}{2} \sum_{c=1}^q \sum_{c':c' \neq c} \exp\left(\frac{1}{N} \sum_{k=1}^s (b_{k,ij,c'} - b_{k,ij,c}) (\beta \mathbb{1}_{\{i \in S_k\}} + \alpha \mathbb{1}_{\{i \notin S_k\}})\right) \\ &\quad \left| \exp\left(\frac{1}{N} (\mathbb{1}_{\omega_j=c'} - \mathbb{1}_{\omega_j=c}) (\beta \mathbb{1}_{i \sim j} + \alpha \mathbb{1}_{i \not\sim j})\right) \right. \\ &\quad \left. - \exp\left(\frac{1}{N} (\mathbb{1}_{\sigma_j=c'} - \mathbb{1}_{\sigma_j=c}) (\beta \mathbb{1}_{i \sim j} + \alpha \mathbb{1}_{i \not\sim j})\right) \right| \\ &\leq \frac{1}{2} e^\beta \sum_{c=1}^q \sum_{c':c' \neq c} \left| \exp\left(\frac{1}{N} (\mathbb{1}_{\omega_j=c'} - \mathbb{1}_{\omega_j=c}) (\beta \mathbb{1}_{i \sim j} + \alpha \mathbb{1}_{i \not\sim j})\right) \right. \\ &\quad \left. - \exp\left(\frac{1}{N} (\mathbb{1}_{\sigma_j=c'} - \mathbb{1}_{\sigma_j=c}) (\beta \mathbb{1}_{i \sim j} + \alpha \mathbb{1}_{i \not\sim j})\right) \right| \\ &=: e^\beta \sum_{c=1}^q \sum_{c':c' \neq c} I(c, c'). \end{aligned}$$

Here, in the first inequality we have used the Lipschitz property, in the second one that $b_{k,ij,c}(\omega) = b_{k,ij,c}(\sigma)$ for all c since $\omega_{j^c} = \sigma_{j^c}$, and in the last one $\alpha < \beta$ once again. Note that if $\{c, c'\} \cap \{\omega_j, \sigma_j\} = \emptyset$, we have $I(c, c') = 0$. It follows that the right-hand side can be written as

$$e^\beta \left(\sum_{c' \neq \omega_j} I(\omega_j, c') + \sum_{c' \neq \sigma_j} I(\sigma_j, c') + \sum_{c \neq \omega_j} I(c, \omega_j) + \sum_{c \neq \sigma_j} I(c, \sigma_j) \right).$$

By Taylor expansion, assuming $i \sim j$ it is not hard to see that

$$I(\omega_j, \sigma_j) = I(\sigma_j, \omega_j) \leq \frac{2\beta}{N} + o((\beta/N)^2), \text{ and}$$

$$I(\omega_j, c') = I(\sigma_j, c') = I(c, \omega_j) = I(c, \sigma_j) \leq \frac{\beta}{N} + O((\beta/N)^2),$$

where $c, c' \notin \{\omega_j, \sigma_j\}$. The same bounds with β replaced by α hold if $i \approx j$. Altogether,

$$\begin{aligned} \|J\|_{\infty \rightarrow \infty} &= \max_{i \in S} \sum_{j \in S} |J_{ij}| \leq \frac{4(N-1)}{2} e^\beta \left((q-2) \frac{\beta}{N} + 2 \frac{\beta}{N} + O_\beta(N^{-2}) \right) \\ &\leq 2q\beta e^\beta + O_\beta(N^{-1}). \end{aligned}$$

In particular, if N is sufficiently large, then for any $\gamma_2 < 1 - 2q\beta e^\beta$,

$$\|J\|_{2 \rightarrow 2} \leq 1 - \gamma_2. \tag{5.5}$$

Finally, combine (5.4) and (5.5) and apply [35, Th. 4.1] (which requires $\gamma_2 \in (0, 1)$). \square

Proof of Theorem 5.1. Use Proposition 5.2 in [35, Proof of Th. 4.1] and [36, (2.10)–(2.12)]. \square

Remark 5.3. By a closer look at the proofs, one can give explicit values for the constants in Theorem 5.1 and Proposition 5.2 depending on γ_1 and γ_2 as in (5.4) and (5.5), namely $C = (\gamma_1 \gamma_2^2)^{-1}$, $\sigma_2^2 = \sigma_3^2 = C$, $\sigma_1^2 = \log(\gamma_1^{-1})C / \log(4)$. Moreover, requiring N to be large enough means that N must be so large that $2q\beta e^\beta + O_\beta(N^{-1}) < 1$ in the asymptotics leading to (5.5).

References

- [1] R. Adamczak, B. Polaczyk, and M. Strzelecki. *Modified log-Sobolev inequalities, Beckner inequalities and moment estimates*, arXiv preprint (2020). arXiv:2007.10209. MR-1164616
- [2] E. Agliari, R. Burioni, and P. Contucci. *A diffusive strategic dynamics for social systems*, J. Stat. Phys. **139** (2010), no. 3, 478–491. MR-2609450
- [3] Q. Berthet, P. Rigollet, and P. Srivastava. *Exact recovery in the Ising blockmodel*, Ann. Statist. **47** (2019), no. 4, 1805–1834. MR-3953436
- [4] S. G. Bobkov and P. Tetali. *Modified logarithmic Sobolev inequalities in discrete settings*, J. Theoret. Probab. **19** (2006), no. 2, 289–336. MR-2283379
- [5] M. Bordewich, C. Greenhill, and V. Patel. *Mixing of the Glauber dynamics for the ferromagnetic Potts model*, Random Structures Algorithms **48** (2016), no. 1, 21–52. MR-3432570
- [6] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities*, Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence, With a foreword by Michel Ledoux. MR-3185193
- [7] A. Bovier and V. Gayrard. *The thermodynamics of the Curie-Weiss model with random couplings*, J. Statist. Phys. **72** (1993), no. 3-4, 643–664. MR-1239568
- [8] P. Caputo, G. Menz, and P. Tetali. *Approximate tensorization of entropy at high temperature*, Ann. Fac. Sci. Toulouse Math. (6) **24** (2015), no. 4, 691–716. MR-3434252

- [9] F. Collet. *Macroscopic limit of a bipartite Curie-Weiss model: a dynamical approach*, J. Stat. Phys. **157** (2014), no. 6, 1301–1319. MR-3277768
- [10] M. Costeniuc, R. S. Ellis, and H. Touchette. *Complete analysis of phase transitions and ensemble equivalence for the Curie-Weiss-Potts model*, J. Math. Phys. **46** (2005), no. 6, 063301, 25. MR-2149837
- [11] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition. MR-1619036
- [12] F. den Hollander. *Large deviations*, volume 14 of *Fields Institute Monographs*. American Mathematical Society, Providence, RI, 2000. MR-1739680
- [13] R. S. Ellis. *Entropy, large deviations, and statistical mechanics*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1985 original. MR-2189669
- [14] R. S. Ellis and K. Wang. *Limit theorems for the empirical vector of the Curie-Weiss-Potts model*, Stochastic Process. Appl. **35** (1990), no. 1, 59–79. MR-1062583
- [15] M. Fedele and P. Contucci. *Scaling limits for multi-species statistical mechanics mean-field models*, J. Stat. Phys. **144** (2011), no. 6, 1186–1205. MR-2841921
- [16] I. Gallo, A. Barra, and P. Contucci. *Parameter evaluation of a simple mean-field model of social interaction*, Math. Models Methods Appl. Sci. **19** (2009), suppl., 1427–1439. MR-2554157
- [17] I. Gallo and P. Contucci. *Bipartite mean field spin systems. Existence and solution*, Math. Phys. Electron. J. **14** (2008), Paper 1, 21. MR-2407199
- [18] F. Götze, H. Sambale, and A. Sinulis. *Higher order concentration for functions of weakly dependent random variables*, Electron. J. Probab. **24** (2019), Paper No. 85, 19. MR-4003138
- [19] F. Götze, H. Sambale, and A. Sinulis. *Concentration inequalities for bounded functionals via generalized log-Sobolev inequalities*, J. Theoret. Probab., online first, 2020.
- [20] G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1988. Reprint of the 1952 edition. MR-0944909
- [21] Z. Kabluchko, M. Löwe, and K. Schubert. *Fluctuations of the magnetization for Ising models on dense Erdős-Rényi random graphs*, J. Stat. Phys. **177** (2019), no. 1, 78–94. MR-4003721
- [22] Z. Kabluchko, M. Löwe, and K. Schubert. *Fluctuations of the magnetization for Ising models on Erdős-Rényi random graphs – the regimes of small p and the critical temperature*, Journal of Physics A: Mathematical and Theoretical **53** (2020), no. 35, 355004, 37. MR-4137544
- [23] H. Kesten and R. H. Schonmann. *Behavior in large dimensions of the Potts and Heisenberg models*, Rev. Math. Phys. **1** (1989), no. 2-3, 147–182. MR-1070088
- [24] J. M. Kincaid and E. G. D. Cohen. *Phase diagrams of liquid helium mixtures and metamagnets: Experiment and mean field theory*, Physics Reports **22** (1975), no. 2, 57–143.
- [25] W. Kirsch and G. Toth. *Two groups in a Curie-Weiss model with heterogeneous coupling*, J. Theoret. Probab. **33** (2020), no. 4, 2001–2026. MR-4166190
- [26] W. Kirsch and G. Toth. *Two groups in a Curie-Weiss model*, Math. Phys. Anal. Geom. **23** (2020), no. 2, Paper No. 17, 14. MR-4098132
- [27] H. Knöpfel, M. Löwe, K. Schubert, and A. Sinulis. *Fluctuation results for general block spin Ising models*, J. Stat. Phys. **178** (2020), no. 5, 1175–1200. MR-4081224
- [28] M. Ledoux. *The concentration of measure phenomenon*, volume 89 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2001. MR-1849347
- [29] Q. Liu. *Limit theorems for the bipartite potts model*, J. Statist. Phys. **181** (2020), 2071–2093. MR-4179798
- [30] M. Löwe and K. Schubert. *Fluctuations for block spin Ising models*, Electron. Commun. Probab. **23** (2018), Paper No. 53, 12. MR-3852267
- [31] M. Löwe and K. Schubert. *Exact recovery in block spin Ising models at the critical line*, Electron. J. Stat. **14** (2020), no. 1, 1796–1815. MR-4090785
- [32] M. Löwe, K. Schubert, and F. Vermet. *Multi-group binary choice with social interaction and a random communication structure – a random graph approach*, Physica A: Statistical Mechanics and its Applications **556** (2020), 124735, 17. MR-4106748

- [33] K. Marton. *Logarithmic Sobolev inequalities in discrete product spaces*, *Combin. Probab. Comput.* **28** (2019), no. 6, 919–935. MR-4015662
- [34] A. A. Opoku, K. Owusu Edusei, and R. K. Ansah. *A conditional Curie-Weiss model for stylized multi-group binary choice with social interaction*, *J. Stat. Phys.* **171** (2018), no. 1, 106–126. MR-3773853
- [35] H. Sambale and A. Sinulis. *Concentration of measure for finite spin systems*, *Bernoulli* **26** (2020), no. 3, 1863–1890. MR-4091094
- [36] H. Sambale and A. Sinulis. *Modified log-Sobolev inequalities and two-level concentration*, *ALEA, Lat. Am. J. Probab. Math. Stat.* **18** (2021), 855–885.
- [37] A. Sinulis. *Higher order concentration of measure and applications*, Ph.D. thesis, Bielefeld Univ., 2019.
- [38] F. Y. Wu. *The potts model*, *Rev. Mod. Phys.* **54** (1982), no. 1, 235–268. MR-0641370

Acknowledgments. The authors would like to thank Arthur Sinulis for fruitful discussions.