Hindawi
Journal of Mathematics
Volume 2018, Article ID 9417547, 13 pages
https://doi.org/10.1155/2018/9417547

Research Article

Multipopulation Spin Models: A View from Large Deviations Theoretic Window

Alex Akwasi Opoku and Godwin Osabutey

Mathematics and Statistics Department, University of Energy and Natural Resources, P. O. Box 214, Sunyani, Ghana

Correspondence should be addressed to Alex Akwasi Opoku; alex.opoku@uenr.edu.gh

Received 20 August 2018; Accepted 26 September 2018; Published 1 November 2018

Academic Editor: Niansheng Tang

This paper studies large deviations properties of vectors of empirical means and measures generated as follows. Consider a sequence $X_1, X_2, \dots, X_n$ of independent and identically distributed random variables partitioned into d-subgroups with sizes $n_1, \dots, n_d$. Further, consider a d-dimensional vector m_n whose coordinates are made up of the empirical means of the subgroups. We prove the following. (1) The sequence of vector of empirical means m_n satisfies large deviations principle with rate n and rate function I, when the sequence $X_1, X_2, \dots, X_n$ is $\mathbb{R}^l$ valued, with $l \geq 1$. (2) Similar large deviations results hold for the corresponding sequence of vector of empirical measures L_n if X_i's, $i = 1, 2, \dots, n$, take on finitely many values. (3) The rate functions for the above large deviations principles are convex combinations of the corresponding rate functions arising from the large deviations principles of the coordinates of m_n and L_n. The probability distributions used in the convex combinations are given by $\alpha = (\alpha^1, \dots, \alpha^d) = \lim_{n \to \infty}(1/n)(n_1, \dots, n_d)$. These results are consequently used to derive variational formula for the thermodynamic limit for the pressure of multipopulation Curie-Weiss (I. Gallo and P. Contucci (2008), and I. Gallo (2009)) and mean-field Pott's models, via a version of Varadhan's integral lemma for an equicontinuous family of functions. These multipopulation models serve as a paradigm for decision-making context where social interaction and other socioeconomic attributes of individuals play a crucial role.

1. Introduction

The early 1970s saw the utilization of two-population mean-field models in the study of the phase transitions and critical behaviour of antiferromagnetic systems. Such mean-field models were used as mean-field approximations of bipartite lattice systems for studying metamagnets [1, 2]. The two-population mean-field ideology was used in [3] to investigate the gibbs-non-gibbs transitions in gibbs measures for Curie-Weiss model subjected to Glauber dynamics. Here the analysis was based on complete analysis of the phase diagram of the evolving spins constrained to having a given magnetization. Phase transitions in such a constrained system is an indication of loss of the gibbs property for the evolving system and it is preserved otherwise.

Statistical mechanical models have seen applications in the socioeconomic literature. Here the focus is on how decisions of individuals are influenced by their socioeconomic environment. For instance, how one's choice of employment, residence, eduction, etc. is influenced by the social and economic environments. Spin models have appeared as natural models for such discrete choice context where social interactions play a crucial role [4–7]. More recently, the authors of [8–10] have introduced two-population mean-field models for a binary choice context where the reference population is partitioned into subgroups of individuals sharing the same socioeconomic attributes. The key assumption here is that individuals with the same attributes tend to behave the same way. In these papers it is assumed that the fractions of the individuals in the subgroups are independent of the size of the reference population. The thermodynamic limit for the pressure of the models was proved for many-body interaction version of the interacting Curie-Weiss model. But variational expression for the pressure and almost sure factorization of correlation function were proved for the case of one and two-body interactions case [10].

The aim of this paper is to set up large deviations machinery for assessing the variational formula for the pressure of the general model introduced in [10] and even extend the model there to the case where the fractions of the individuals in the subgroups are dependent on the size of the reference population. We also employ the tools developed here to derive variational formula for the pressure of a multibody multipopulation mean-field Potts' model.

We establish large deviation results for vectors of empirical means associated with a collection of random variables modelling the behaviour of interest of the subgroups that constitute the reference group. These empirical means are derived from uneven numbers of random variables. Thus the vector components are given by empirical mean of different numbers of independent and identically distributed random variables. Due to the variations in the sizes of the subgroups, the large populations asymptotics of the free energy results in proving a version of Varadhan's integral lemma for a sequence of functionals of the vector of empirical measures instead of the usual case where this functional is fixed throughout the asymptotics. We provide a necessary condition for such a sequence to admit the desired asymptotic result.

The rest of the paper is organized as follows: Section 2 discusses the generalities on large deviations theory and main results of the paper. In Section 3 we introduce the multipopulation Curie-Weiss and mean-field Potts' models to motivate the large deviations problem we address in this paper. The proofs of the results in Section 3 are given in Section 4.

2. Generalities on Large Deviations Theory and Main Results

Large deviation theory tells how, on an exponential scale, the probability for an atypical event decays to zero. More formerly, large deviations are defined as follows.

Definition 1. Let $\mathscr{X}$ be a complete separable metric space, $\mathscr{B}(\mathscr{X})$ the Borel σ-algebra of $\mathscr{X}$, and $\{\mu_n;\ n = 1, 2, \ldots\}$ a sequence of probability measures on $\mathscr{B}(\mathscr{X})$.

(1) $\{\mu_n\}$ is said to have a large deviation property if there exist a sequence of positive numbers $\{a_n;\ n = 1, 2, \ldots\}$ which tends to ∞ and a function I which maps $\mathscr{X}$ into $[0, \infty]$ such that the following hypothesis holds:

- (a) I is lower semicontinuous on $\mathscr{X}$.
- (b) I has compact level sets.
- (c) $\limsup_{n\longrightarrow\infty} a_n^{-1} \log \mu_n(K) \leq -\inf_{x\in K} I(x)$, for each closed set K in $\mathscr{X}$.
- (d) $\liminf_{n\longrightarrow\infty} a_n^{-1} \log \mu_n(G) \geq -\inf_{x\in G} I(x)$, for each open set G in $\mathscr{X}$.

(2) I is called the entropy/rate function of $\{\mu_n\}$.

In the above definition condition (b) implies that the rate function I is good.

2.1. The Set Up. Suppose $X_1, X_2, \ldots, X_n, \ldots$, is a sequence of independent and identically distributed $\mathbb{R}^l$-valued random variables, for a positive integer $l \geq 1$. Let $I_{n^1}, \ldots, I_{n^d}$ be a partition of the set $I_n = \{1, 2, \ldots, n\}$. The partition may be interpreted as the indexing set of the $d \geq 1$ subpopulations in a population of size n. Here we denote by n_i the size of the ith subpopulation and we assume that

$$\lim_{n\longrightarrow\infty} \alpha_n^i = \lim_{n\longrightarrow\infty} \frac{n_i}{n} = \alpha^i, \tag{1}$$

for any $i = 1, 2, \ldots, d$. For each subpopulation i we are interested in the empirical mean

$$S_n^i = \frac{1}{n_i} \sum_{k\in I_{n_i}} X_k \tag{2}$$

and the vector of empirical means for the multipopulation is given by

$$S_n = \left(S_n^1, \ldots, S_n^d\right). \tag{3}$$

For the case $l = 1$, it is clear that $S_n \in \mathbb{R}^d$, but it is different from the empirical mean of a sequence of $\mathbb{R}^d$-valued random variables. Note that in the latter each coordinate of the empirical mean is a sum of n random variables. In our case the coordinates of the empirical mean vector S_n are made up of sums of uneven numbers of random variables. In what follows we write μ_n for the distribution of S_n and the space $\mathscr{X} = (\mathbb{R}^l)^d$.

2.2. The $\mathbb{R}$-Valued Case. Suppose the sequence $X_1, X_2, \ldots$ of $\mathbb{R}$-valued random variables is independent and identically distributed with common distribution μ. Suppose the logarithmic moment generating function associated with μ is given by

$$\Lambda(\lambda) = \log M(\lambda) = \log E\left[e^{\lambda X_1}\right] \tag{4}$$

We assume $\Lambda(\lambda) < \infty$, for all $\lambda \in \mathbb{R}$. The Fenchel-Legendre transform of $\Lambda(\lambda)$ is defined as

$$\Lambda^*(x) = \sup_{\lambda\in\mathbb{R}} [\lambda x - \Lambda(\lambda)]. \tag{5}$$

We now state our first large deviations result for the vector of empirical means S_n. Recall that μ_n is the law of S_n.

Theorem 2. *The sequence of vector of empirical means S_n satisfies large deviations principle with rate n and rate function $\overline{\Lambda}^*(\cdot)$, given by*

$$\Lambda^*(x) = \sup_{\lambda\in\mathbb{R}^d} \left(\sum_{i=1}^{d} \alpha^i \left[\lambda^i x^i - \Lambda\left(\lambda^i\right)\right]\right). \tag{6}$$

2.3. The $\mathbb{R}^l$-Valued Case. In what follows we will consider the large deviations principle for the case of a sequence of independent and identically distributed (i.i.d.) $\mathbb{R}^l$-valued random variables $X_1, X_2, \ldots$. Here $l \geq 2$. As before, we denote by μ, a probability measure on $\mathbb{R}^l$, the common law of the i.i.d. sequence $X_1, X_2, \ldots$. Here again, we let

$$S_n^i = \frac{1}{n_i} \sum_{j\in I_{n_i}} X_j \in \mathbb{R}^l, \quad \text{for } i = 1, 2, \ldots, d, \tag{7}$$

and

$$S_n = \left(S_n^1, S_n^2, \ldots, S_n^d\right) \in \left(\mathbb{R}^l\right)^d. \tag{8}$$

Further, we assume that

$$\Lambda(\lambda) = \log E\left(e^{\langle \lambda, X_1\rangle}\right) < \infty, \quad \text{for every } \lambda \in \mathbb{R}^l. \tag{9}$$

In what follows we will write for every $x \in (\mathbb{R}^l)^d$, as

$$\begin{aligned} & x = \left(x^1, x^2, \ldots, x^d\right), \\ & \text{where } x^i = \left(x_1^i, x_2^i, \ldots, x_l^i\right) \in \mathbb{R}^l, \text{ for every } i = 1, 2, \ldots, d. \end{aligned} \tag{10}$$

In view of the above, we will write the inner product for any pair $x, y \in (\mathbb{R}^l)^d$ as follows:

$$\langle x, y\rangle = \sum_{i=1}^d \left\langle x^i, y^i\right\rangle = \sum_{i=1}^d \sum_{j=1}^l x_j^i y_j^i. \tag{11}$$

We define the $\mathbb{R}^d$-valued vectors $\overline{n} = (n_1, n_2, \ldots, n_d)$ and $\alpha = (\alpha^1, \alpha^2, \ldots, \alpha^d)$. The large deviations principle for the S_n vector for $l \geq 2$ case is summarized in the results below.

Theorem 3. *Let S_n be defined as above. Then S_n satisfies a large deviations principle with rate n and rate function $\overline{\Lambda}^*$ given by*

$$\Lambda^*(x) = \sup_{\lambda \in (\mathbb{R}^l)^d} \sum_{i=1}^d \alpha^i \left[\left\langle \lambda^i, x^i\right\rangle - \Lambda\left(\lambda^i\right)\right]. \tag{12}$$

2.4. Large Deviations Principle for a Vector of Empirical Measures. In this section we consider the large deviations principle for a vector of empirical measures of the subgroups when the sequence of i.i.d. random variables takes finitely many values. Specifically, let Σ be a finite set and $X_1, X_2, X_3, \ldots$ be an independent and identically distributed sequence of Σ-valued random variables with common law $\mu = (\mu_1, \ldots, \mu_{|\Sigma|}) \in \mathbb{R}^{|\Sigma|}$. Here $\mu_i = \mu(a_i)$, for any $a_i \in \Sigma$. The empirical measure for the random variables indexed by I_{n^i} is given by

$$\begin{aligned} & L_{n_i}[x] = \left(L_{n_i}[X](a_1), \ldots, L_{n_i}[X](a_{|\Sigma|})\right) \in \mathbb{R}^{|\Sigma|}, \\ & \qquad \text{where } L_{n_i}[X](a_k) = \frac{1}{n_i} \sum_{j \in I_{n_i}} 1_{a_k}\left(X_j\right). \end{aligned} \tag{13}$$

Here we have put $1_{a_k}(X_j) = 1_{\{a\}}(X_j)$, the indicator function for $\{a_k\} \subset \Sigma$. Note that $L_{n_i}[X](a_k)$ is the relative frequency of a_k among the random sample $(X_j)_{j \in I_{n_i}}$.

The vector of empirical measures is then given by

$$L_n[X] = \left(L_{n_1}[X], L_{n_2}[X], \ldots, L_{n_d}[X]\right) \in \left(\mathbb{R}^{|\Sigma|}\right)^d. \tag{14}$$

Note that $L_n[X]$ is a vector whose coordinates are probability measures on Σ. To see the connection between our vector of empirical measures and the vector of empirical means we considered earlier, we introduce the following sequence of random variables: for positive integer k, define

$$Y_k = \left(1_{a_1}\left(X_k\right), 1_{a_2}\left(X_k\right), \ldots, 1_{a_{|\Sigma|}}\left(X_k\right)\right). \tag{15}$$

Then the sequence $Y_1, Y_2, Y_3, \ldots$ is an i.i.d. $\mathbb{R}^{|\Sigma|}$–valued random variables are with the property that

$$\begin{aligned} S_n^i &= \frac{1}{n_i} \sum_{j \in I_{n_i}} X_j = L_{n_i}[X], \\ S_n &= L_n[X]. \end{aligned} \tag{16}$$

Thus the empirical mean of the Y-sequence is the same as the empirical measure of the X-sequence and the vector of empirical means associated with the Y-sequence coincides with the vector of empirical measures of the X-sequence.

For every $\lambda \in \mathbb{R}^{|\Sigma|}$, note that

$$\Lambda(\lambda) = \log E\left(e^{\langle \lambda, Y_1\rangle}\right) = \log \sum_{j=1}^{|\Sigma|} e^{\lambda^j} \mu_j. \tag{17}$$

Further, for any probability measure ν on Σ, the relative entropy of ν relative to μ is given by

$$H(\nu \mid \mu) = \sum_{j=1}^{|\Sigma|} \nu_j \log\left(\frac{\nu_j}{\mu_j}\right). \tag{18}$$

Suppose $\mathscr{M}_1(\Sigma)$ is the set of all probability measures on Σ. Then $\mathscr{M}_1(\Sigma) \subset \mathbb{R}^{|\Sigma|}$ and $\mathscr{M}_1(\Sigma)^d \subset (\mathbb{R}^{\Sigma})^d$. For any $\nu \in \mathscr{M}_1(\Sigma)$, it can be written in the form $\nu = (\nu^1, \nu^2, \ldots, \nu^d)$, where $\nu^i \in (\Sigma)$, for any $i = 1, 2, \ldots, d$. The following result is the large deviations principle for the vector of empirical measures $L_n[X]$.

Theorem 4. *The sequence of vectors of empirical measures $L_n[X]$ satisfies a large deviations principle with rate n and rate function*

$$\Lambda^*(\nu) = \sup_{\lambda \in (\mathbb{R}^{|\Sigma|})^d} \sum_{i=1}^d \alpha^i \left[\left\langle \lambda^i, \nu^i\right\rangle - \Lambda\left(\lambda^i\right), \right] \tag{19}$$

for every $\nu \in \mathscr{M}_1(\Sigma)^d$. In particular,

$$\Lambda^*(\nu) = \sum_{i=1}^d \alpha^i H\left(\nu^i \mid \mu\right). \tag{20}$$

2.5. Varadhan's Integral Lemma. In this section we consider Varadhan's integral lemma for a sequence of equicontinuous functions. Here we put $\mathscr{X} = (\mathbb{R}^l)^d$.

Theorem 5. *Suppose $\{\mu_n\}$ satisfies a large deviation principle with rate n and a good rate function $I : \mathscr{X} \longrightarrow \mathbb{R}$. Further, let the sequence of functions $\phi_n : \mathscr{X} \longrightarrow \mathbb{R}$ be equicontinuous converging point-wise to a function $\phi : \mathscr{X} \longrightarrow \mathbb{R}$. Assume either the tail condition*

$$\lim_{M \longrightarrow \infty} \limsup_{n \longrightarrow \infty} \frac{1}{n} \log E\left(e^{n\phi_n(S_n)} 1_{\{\phi_n(S_n \geq M)\}}\right) = -\infty \tag{21}$$

or the following moment generating condition for some $\gamma > 1$,

$$\limsup_{n\longrightarrow\infty}\frac{1}{n}\log E\left[e^{n\gamma\phi_n(S_n)}\right] < \infty. \quad (22)$$

Then

$$\lim_{n\longrightarrow\infty}\frac{1}{n}\log E\left(e^{n\phi_n(S_n)}\right) = \sup_{x\in\mathcal{X}}\left[\phi(x) - I(x)\right]. \quad (23)$$

3. Applications

Let us now introduce the models that motivated the large deviations questions addressed in this paper. We are interested in a model that will capture how individual decisions or choices are influenced by the choices of the rest of the people in their reference group. Additionally, individuals do have attributes, such as gender, place of residence, level of education, and ethnicity, that also influence their decisions. The Curie-Weiss case introduced in [10] is discussed first and the corresponding mean-field Potts' case will be discussed after that. The method discussed here could also apply to continuous spin models with compact support. In particular, it will apply to the mean-field versions of the following models: the $O(n)$ model [11], the spherical model [12, 13], the liquid crystal model [11, 14, 15], and the Kuramoto model [16]. The detail analysis for continuous spin models and their phase diagram will be carried out in a future paper. We present the results for the Curie-Weiss and Potts's models here because we have already studied the thermodynamic limits of these models in [17, 18].

3.1. Multipopulation Curie-Weiss Model. Suppose each individual in a population of N agents chooses a binary action, such as voting YES or NO on some issue at some common time. This binary action is coded by $\sigma_i \in \{1, -1\}$ with

$$\sigma_i = \begin{cases} +1, & \text{if individual } i \text{ votes YES} \\ -1, & \text{if individual } i \text{ votes NO,} \end{cases} \quad (24)$$
$$\text{for } i \in \{1, \ldots, N\}.$$

The choices made by all the N individuals are also coded by $\sigma \in \Omega_N = \{-1, +1\}^N$. The level of satisfaction of the population for deciding on $\sigma \in \Omega_N$ is given by the Curie-Weiss Hamiltonian

$$H_N(\sigma) = \frac{1}{2N}\sum_{i,l=1}^{N} J_{il}\sigma_i\sigma_l + \sum_{i=1}^{N} h_i\sigma_i. \quad (25)$$

The function H_N on the configurations σ represents the utility of individuals as a result of their choices and the influences on them while making those choices [19]. $H_N(\sigma)$ measures the level of satisfaction of the entire population for making the choice σ. The higher is the $H_N(\sigma)$, the higher is the level of satisfaction of the population. It has two parts; the first part models the social incentive of individuals in the population and the second part models the private incentive of individuals. Here J_{il} measures the influence of individual i on individual l. When J_{il} is positive means conformity or imitation is rewarded and conformity is not rewarded when J_{il} is negative. h_i controls the part of the utility that is specific to individual i.

Next we reparametrize the parameters J_{il} and h_i as follows: suppose that each individual i in the reference population has k attributes $a_i = (a_i^{(1)}, \ldots, a_i^{(k)}) \in \{0, 1\}^k$. For instance, suppose that attributes 1 and 2 are, respectively, employment $a_i^{(1)}$ and marital status $a_i^{(2)}$, then

$$a_i^{(1)} = \begin{cases} 1, & \text{if individual } i \text{ is employed} \\ 0, & \text{if individual } i \text{ is unemployed,} \end{cases} \quad (26)$$

and

$$a_i^{(2)} = \begin{cases} 1, & \text{if an individual } i \text{ married} \\ 0, & \text{if an individual } i \text{ is unmarried.} \end{cases} \quad (27)$$

Therefore, with respect to the attributes, the reference group can be partitioned into 2^k nonoverlapping subgroups. Members in a given subgroup share the same attributes and it is therefore reasonable to assume that they also behave the same way. In view of this, we shall assume in what follows that $J_{il} = J_{gg'}$, for all choices of i coming from subgroup g and all l taken from subgroup g'. Further, we assume that $h_i = h^g$ for all individuals in group g. In the sequel we will let I_{N_g} be the set of individuals in subgroup g, for $g = 1, \ldots, 2^k$, with $|I_{N_g}| = N_g$. Therefore, $N = N_1 + \cdots + N_{2^k}$. Note that $I_N = \{1, 2, \ldots, N\} = \bigcup_{g=1}^{2^k} I_{N_g}$ with $I_{N_g} \cap I_{N_{g'}} = \emptyset$, for $g \neq g'$. Define for every $N \in \mathbb{N}$

$$\gamma_N^g = \frac{N_g}{N}, \qquad m_N^g = \frac{1}{N_g}\sum_{i\in I_{N_g}}\sigma_i. \quad (28)$$

Note that γ_N^g is the proportion of the individuals found in subgroup g and m_N^g is the magnetization of the spins in subgroup g.

Therefore, it follows from (25) and the above parametrization of J_{il} and h_i that

$$\begin{aligned} H_N(\sigma) &= \frac{1}{2N}\sum_{g=1}^{2^k}\sum_{g'=1}^{2^k}\left(\sum_{i\in I_{N_g}}\sum_{l\in I_{N_{g'}}} J_{il}\sigma_i\sigma_l\right) + \sum_{g=1}^{2^k}\sum_{i\in I_{N_g}} h_i\sigma_i \\ &= N\left[\sum_{g=1}^{2^k}\sum_{g'=1}^{2^k}\frac{J_{gg'}}{2}\gamma_N^g\gamma_N^{g'}m_N^g m_N^{g'} + \sum_{g=1}^{2^k} h^g\gamma_N^g m_N^g\right]. \end{aligned} \quad (29)$$

Note that if $k = 0$, we get the original Curie-Weiss Hamiltonian. For the case $k \geq 1$, we end up with 2^k Curie-Weiss models on the subgroups $I_{N_1}, \ldots, I_{N_{2^k}}$, that are interacting with one another. Here we have 2^k subgroups because we have attributes that are binary. We could allow the attributes to have any finite number of alternatives and the alternatives for the attributes need not to come from the same set. Therefore, in what follows we will assume there are $d \geq 1$ subgroups and that the Hamiltonian takes the form

$$H_N(\sigma) = N\left[\sum_{g=1}^{d}\sum_{g'=1}^{d}\frac{J_{gg'}}{2}\gamma_N^g\gamma_N^{g'}m_N^g m_N^{g'} + \sum_{g=1}^{d}h^g\gamma_N^g m_N^g\right]. \tag{30}$$

The Hamiltonian in (30) consists of one-body and two-body interactions. In what follows we will extend the number of bodies in the interaction to range from 1 to d, where d is the number of subgroups [10]. We consider Hamiltonian of the form

$$\begin{aligned} H_N(\sigma) &= N\left[\sum_{r=1}^{d}\sum_{g_1,\ldots,g_r=1}^{d}\frac{J_{g_1\cdots g_r}}{r}\prod_{k=1}^{r}\gamma_N^{g_k}m_N^{g_k}.\right] \\ &= \sum_{r=1}^{d}\frac{1}{N^{r-1}}\sum_{g_1,\ldots,g_r=1}^{d}\frac{J_{g_1\cdots g_r}}{r}\sum_{i_1\in I_{N_{g_1}}}\cdots\sum_{i_r\in I_{N_{g_r}}}\sigma_{i_1}\ldots\sigma_{i_r}. \end{aligned} \tag{31}$$

Note from (30) and (31) that for $r = 1$, $J_{g_1} = h^g$ and for $r = 2$, $J_{g_1 g_2} = J_{gg'}$. The $J_{g_1,\ldots,g_r}$ are interaction coefficients associated with the r-body interaction among individuals coming from the subgroups $g_1, \ldots, g_r$, respectively. Thus the interaction is defined with the help of a tensor $J_{g_1,\ldots,g_r}$ of rank r for each of the r-body interactions [10].

Further, we assume that there is a probability measure γ on the set $\{1, 2, \ldots, d\}$, such that $\gamma = (\gamma^1, \ldots, \gamma^d)$ and

$$\lim_{N\longrightarrow\infty}\gamma_N^g = \gamma^g, \quad \text{for any } g = 1, 2, \ldots, d. \tag{32}$$

Note that the model we consider here is more general than the cases in [8, 10], in that for any finite N the fractions of the subgroups $\gamma_N^1, \ldots, \gamma_N^d$ are dependent on N. In [8, 10] these fractions are chosen to be independent of N. This simplifies the proofs, especially the existence of the thermodynamic limit.

In the sequel we will use the following notation: let $\Delta = [-1, 1] \subset \mathbb{R}$. Note that $m_N = (m_N^{g_1}, \ldots, m_N^{g_d}) \in \Delta^d$. Further, for every positive integer N, define a map $F_N : \Delta^d \longrightarrow \mathbb{R}$ as

$$F_N(m) = \sum_{r=1}^{d}\sum_{g_1,\ldots,g_r=1}^{d}\frac{J_{g_1\cdots g_r}}{r}\prod_{k=1}^{r}\gamma_N^{g_k}m^{g_k}. \tag{33}$$

Note that the F_N's are uniformly bounded by $\sum_{r=1}^{d}\sum_{g_1,\ldots,g_r=1}^{d}|J_{g_1\cdots g_r}/r|$. It is also clear that

$$\begin{aligned} F(m) &= \sum_{r=1}^{d}\sum_{g_1,\ldots,g_r=1}^{d}\frac{J_{g_1\cdots g_r}}{r}\prod_{k=1}^{r}\gamma^{g_k}m^{g_k} \\ &= \lim_{N\longrightarrow\infty}\sum_{r=1}^{d}\sum_{g_1,\ldots,g_r=1}^{d}\frac{J_{g_1\cdots g_r}}{r}\prod_{k=1}^{r}\gamma_N^{g_k}m^{g_k}. \end{aligned} \tag{34}$$

Since the maps $m \longmapsto F_N(m)$ are continuous for every N, and Δ^d is a compact subset of $\mathbb{R}^d$, it follows from Theorems 7.13 and 7.24 of [20] that the sequence $\{F_N\}$ is equicontinuous. Further, the Hamiltonian H_N in (31) become

$$H_N(\sigma) = NF_N(m_N). \tag{35}$$

Suppose the spins $\sigma_1, \sigma_2, \ldots$ are independent and identically distributed sequence of random variables with

$$P(\sigma_1 = 1) = P(\sigma_1 = -1) = \frac{1}{2}. \tag{36}$$

We denote by P_N the corresponding product measure on $\Omega_N = \{-1, 1\}^N$. The equilibrium state ω_N associated with the Hamiltonian H_N in (31) is given by

$$\omega_N(\sigma) = \frac{e^{H_N(\sigma)}P_N(\sigma)}{Z_N}, \quad \text{for } \sigma \in \Omega_N, \tag{37}$$

where

$$Z_N = \sum_{\tilde{\sigma}\in\Omega_N}e^{H_N(\tilde{\sigma})}P_N(\tilde{\sigma}) = \int_{\Delta^d}e^{NF_N(m_N)}Q_N(m) \tag{38}$$

is the partition function of the model and Q_N is the law of vector of empirical means $m_N = (m_N^1, \ldots, m_N^d)$ under P_N. In (38) we have used (35). In what follows μ, h, and J shall be as follows:

$$\begin{aligned} \gamma &= (\gamma^1, \ldots, \gamma^d), \\ h &= (h^1, \ldots, h^d) \in \mathbb{R}^d, \\ J &= (J_{g_1,\ldots,g_r})_{1\leq g_1,\ldots,g_r,r\leq d}. \end{aligned} \tag{39}$$

The pressure function of the model is then given by

$$p_N = \frac{1}{N}\log Z_N. \tag{40}$$

The large N behaviour of the model is governed by the pressure function. It is known from [18] that the thermodynamic limit

$$p(\gamma, h, J) = \lim_{N\longrightarrow\infty}\frac{1}{N}\log Z_N \tag{41}$$

exists. The proof of the case $\gamma_N^g = \gamma^g$ was earlier given in [8, 10].

Theorem 6. *For choice of the parameters γ, h, and J, the limiting pressure admits the following variational representation:*

$$p(\gamma, h, J) = \sup_{\mu\in[-1,1]^d}\left[F(\mu) - \sum_{g=1}^{d}\gamma^g I(\mu^g)\right], \tag{42}$$

where

$$I(\mu^g) = \frac{1-\mu^g}{2}\log(1-\mu^g) + \frac{1+\mu^g}{2}\log(1+\mu^g) \tag{43}$$

and F is given in (34).

Proof. The proof of this theorem follows from (38), (40), and (41) and Theorems 2 and 5 upon setting $\sigma_i = X_i$, $N_g = n_g$, $N = n$, $\gamma^g = \alpha^g$, $\gamma = \alpha$, and $F_N = \phi_n$, and noting that the F_N's form an equicontinuous family and they are uniformly bounded. $\square$

3.2. Multipopulation Mean-Field Potts' Model. Suppose this time round that the individuals in the population of N agents choose from finite number of alternatives, say various alternatives of employment, at some common time. This discrete choice action of individual i is coded by $\sigma_i \in E = \{1, \ldots, q\}$, where $q \geq 2$, with

$$\sigma_i = \begin{cases} 1, & \text{if individual } i \text{ chooses employment alternative 1} \\ \vdots & \\ q, & \text{if individual } i \text{ chooses employment alternative q} \end{cases} \quad \text{for } i \in \{1, \ldots, N\}. \tag{44}$$

The choices made by all the N individuals are also coded by $\sigma \in \Omega_N = \{1, \ldots, q\}^N$. The level of satisfaction of the population for deciding on $\sigma \in \Omega_N$ is given by the mean-field Potts' Hamiltonian

$$H_N(\sigma) = \frac{1}{2N} \sum_{i,l=1}^{N} J_{il} \delta_{\sigma_i,\sigma_l} + \sum_{i=1}^{N} h_i \delta_{\sigma_i,1}. \tag{45}$$

Here $\delta_{\sigma_i,\sigma_j}$ is the Dirac-delta measure. The function H_N on the configurations σ represents the utility of individuals as a result of their choices and the influences on them while making a decision [19]. H_N and its parameters have the usual interpretation given for the Curie-Weiss model.

Using the partition $I_{N_1}, \ldots, I_{N_d}$ of I_N, we define for every $N \in \mathbb{N}$

$$L_N^g[\sigma](s) = \frac{1}{N_g} \sum_{i \in I_{N_g}} \delta_{\sigma_i,s}, \quad \text{for any } s = 1, 2, \ldots, q. \tag{46}$$

Therefore

$$L_N^g[\sigma] = \left(L_N^g[\sigma](1), \ldots, L_N^g[\sigma](q)\right) \tag{47}$$

is the empirical measure for the choices of the individuals in subgroup g. It is an element in $\mathscr{P}(E)$, the set of probability measures on $E = \{1, \ldots, q\}$. Further,

$$L_N[\sigma] = \left(L_N^1[\sigma], \ldots, L_N^{2^k}[\sigma]\right) \tag{48}$$

is the vector of empirical measures. Note that $L_N[\sigma] \in \mathscr{P}(E)^{2^k}$ and γ_N^g is the proportion of the individuals found in subgroup g.

Therefore, it follows from (45) and the above parametrization of J_{il} and h_i that

$$\begin{aligned} H_N(\sigma) &= \frac{1}{2N} \sum_{g=1}^{2^k} \sum_{g'=1}^{2^k} \left(\sum_{i \in I_{N_g}} \sum_{l \in I_{N_{g'}}} J_{il} \delta_{\sigma_i,\sigma_l} \right) \\ &\quad + \sum_{g=1}^{2^k} \sum_{i \in I_{N_g}} h_i \delta_{\sigma_i,1} \\ &= N \left[\sum_{g=1}^{2^k} \sum_{g'=1}^{2^k} \frac{J_{gg'}}{2} \gamma_N^g \gamma_N^{g'} \left\langle L_N^g[\sigma], L_N^{g'}[\sigma] \right\rangle \right. \\ &\quad \left. + \sum_{g=1}^{2^k} h^g \gamma_N^g \left\langle L_N^g[\sigma], e_1 \right\rangle \right]. \end{aligned} \tag{49}$$

In the second equation above we have used that

$$\begin{gathered} \sum_{s=1}^{q} \left(\sum_{i \in I_{N_g}} \delta_{\sigma_i,s} \right) \left(\sum_{l \in I_{N_{g'}}} \delta_{\sigma_l,s} \right) = \sum_{i \in I_{N_g}} \sum_{l \in I_{N_{g'}}} \delta_{\sigma_i,\sigma_l}, \\ \langle x, y \rangle = \sum_{i=1}^{2^k} x_i y_i, \\ \text{for any } x, y \in \mathbb{R}^{2^k}, \\ e_1 = (1, 0, 0, \ldots, 0) \in \mathbb{R}^{2^k}. \end{gathered} \tag{50}$$

Note that if $k = 0$, we get the original mean-field Potts' Hamiltonian. For the case $k \geq 1$, we end up with 2^k mean-field Potts' models on the subgroups $I_{N_1}, \ldots, I_{N_{2^k}}$, that are interacting with one another. Here we have 2^k subgroups because we have attributes that are binary. We could allow the attributes to have any finite number of alternatives and the alternatives for the attributes need not come from the same set. Therefore, in what follows we will assume there are $d \geq 1$ subgroups and this gives rise to the Hamiltonian

$$\begin{aligned} H_N(\sigma) = N &\left[\sum_{g=1}^{d} \sum_{g'=1}^{d} \frac{J_{gg'}}{2} \gamma_N^g \gamma_N^{g'} \left\langle L_N^g[\sigma], L_N^{g'}[\sigma] \right\rangle \right. \\ &\quad \left. + \sum_{g=1}^{d} h^g \gamma_N^g \left\langle L_N^g[\sigma], e_1 \right\rangle \right]. \end{aligned} \tag{51}$$

The Hamiltonian in (51) consists of one-body and two-body interactions. In what follows we will extend the number of bodies in the interaction to range from 1 to d, where d is the number of subgroups. We consider Hamiltonian of the form

$$H_N(\sigma)=\sum_{r=1}^{d}\frac{1}{N^{r-1}}\sum_{g_1,\dots,g_r=1}^{d}\frac{J_{g_1\cdots g_r}}{r}\sum_{i_1\in I_{N_{g_1}}}\cdots\cdot\sum_{i_r\in I_{N_{g_r}}}\delta_{\sigma_{i_1},\dots,\sigma_{i_r}}=N\left[\sum_{r=1}^{d}\sum_{g_1,\dots,g_r=1}^{d}\frac{J_{g_1\cdots g_r}}{r}\cdot\prod_{k=1}^{r}\gamma_N^{g_k}\left\langle L_N^{g_1}[\sigma],\dots,L_N^{g_r}[\sigma]\right\rangle.\right]. \tag{52}$$

In the above we have used that

$$\sum_{s=1}^{q}\left(\sum_{i_1\in I_{N_{g_1}}}\delta_{\sigma_{i_1},s}\right)\cdots\left(\sum_{i_r\in I_{N_{g_r}}}\delta_{\sigma_{i_r},s}\right)=\sum_{i_1\in I_{N_{g_1}}}\cdots\sum_{i_r\in I_{N_{g_r}}}\delta_{\sigma_{i_1},\dots,\sigma_{i_r}},\tag{53}$$
$$\left\langle x^1,\dots,x^r\right\rangle=\sum_{i=1}^{d}\prod_{k=1}^{r}x_i^k,\quad\text{for any } x^1,\dots,x^r\in\mathbb{R}^d.$$

Note from (51) and (52) that for $r=1$, $J_{g_1}=h^g$ and $\langle x^1\rangle=\langle x_1,e_1\rangle$. The $J_{g_1,\dots,g_r}$ are interaction coefficients associated with the r-body interaction among individuals coming from subgroups $g_1,\dots,g_r$, respectively. Thus the interaction is defined with the help of a tensor $J_{g_1,\dots,g_r}$ of rank r for each of the r-body interactions. The model considered here is mean-field Pott's version of the Curie-Weiss model considered in [8, 10, 17]. Further, for any finite N the fractions of the subgroups $\gamma_N^1,\dots,\gamma_N^d$ are dependent on N. In [8, 10] these fractions were chosen to be independent of N, which simplified the proofs, especially the proof of the existence of the thermodynamic limit.

Note that $L_N[\sigma]=(L_N^1[\sigma],\dots,L_N^d[\sigma])\in\mathcal{P}(E)^d$. Further, for every positive integer N, define a map $F_N:\mathcal{P}(E)^d\longrightarrow\mathbb{R}$ as

$$F_N(\nu)=\sum_{r=1}^{d}\sum_{g_1,\dots,g_r=1}^{d}\frac{J_{g_1\cdots g_r}}{r}\prod_{k=1}^{r}\gamma_N^{g_k}\sum_{i=1}^{d}\prod_{k=1}^{r}\nu_i^{g_k}.\tag{54}$$

Note that the F_N's are uniformly bounded by $\sum_{r=1}^{d}\sum_{g_1,\dots,g_r=1}^{d}|J_{g_1\cdots g_r}/r|$. It is also clear that

$$F(\nu)=\sum_{r=1}^{d}\sum_{g_1,\dots,g_r=1}^{d}\frac{J_{g_1\cdots g_r}}{r}\prod_{k=1}^{r}\gamma^{g_k}\sum_{i=1}^{d}\prod_{k=1}^{r}\nu_i^{g_k}=\lim_{N\longrightarrow\infty}\sum_{r=1}^{d}\sum_{g_1,\dots,g_r=1}^{d}\frac{J_{g_1\cdots g_r}}{r}\prod_{k=1}^{r}\gamma_N^{g_k}\sum_{i=1}^{d}\prod_{k=1}^{r}\nu_i^{g_k}.\tag{55}$$

Since the maps $\nu\longmapsto F_N(\nu)$ are continuous for every N, and Δ^d is a compact subset of $(\mathbb{R}^q)^d$, it follows from Theorems 7.13 and 7.24 of [20] that the sequence $\{F_N\}$ is equicontinuous. Further, the Hamiltonian H_N in (52) become

$$H_N(\sigma)=NF_N(L_N[\sigma]).\tag{56}$$

Suppose the spins $\sigma_1,\sigma_2,\dots$ are independent and identically distributed sequence of E-valued random variables with

$$P(\sigma_1=s)=\rho(s)=\frac{1}{q},\quad\text{for every } s\in E.\tag{57}$$

We denote by P_N the corresponding product measure ρ^N on $\Omega_N=E^N$. The equilibrium state ω_N associated with the Hamiltonian H_N in (52) is given by

$$\omega_N(\sigma)=\frac{e^{H_N(\sigma)}P_N(\sigma)}{Z_N},\quad\text{for }\sigma\in\Omega_N,\tag{58}$$

where

$$Z_N=\sum_{\widetilde{\sigma}\in\Omega_N}e^{H_N(\widetilde{\sigma})}P_N(\widetilde{\sigma})=\int_{\mathcal{P}(E)^d}e^{NF_N(\nu)}Q_N(d\nu)\tag{59}$$

is the partition function of the model. Here Q_N is the law of the vector of empirical measures $L_N(\sigma)$. The pressure function of the model is then given by

$$p_N=\frac{1}{N}\log Z_N.\tag{60}$$

It follows from [17] that the thermodynamic limit

$$p(\gamma,h,J)=\lim_{N\longrightarrow\infty}\frac{1}{N}\log Z_N\tag{61}$$

exists.

The limiting pressure admits the following variational formula representation.

Theorem 7. *For any choice of the parameters γ, h, and J, the limiting pressure admits the following variational representation:*

$$p(\gamma,h,J)=\sup_{\nu\in\Delta^d}\left[F(\nu)-\sum_{g=1}^{d}\gamma^g H(\nu^g\mid\rho)\right],\tag{62}$$

where

$$H(\nu^g\mid\rho)=\sum_{s=1}^{q}\nu_s^g\log\left(\frac{\nu_s^g}{\rho(s)}\right)\tag{63}$$

is the relative entropy of ν^g with respect to ρ and F is given in (55).

Proof. The proof of this theorem follows from (59) to (61) and Theorems 4 and 5 upon setting $\sigma_i=X_i$, $N_g=n_g$, $N=n$, $\gamma^g=\alpha^g$, $\gamma=\alpha$, and $F_N=\phi_n$, and noting that the F_N's form an equicontinuous family and they are uniformly bounded. □

4. Proofs

The proofs of the results of this paper are given in this section. In the proof below we will use the following properties of the functions Λ and Λ^*. For proof of these properties we refer the reader to the proof of Lemma 2.2.5 of [21].

Lemma 8.

(1) *Λ is a convex function and Λ^* is a convex rate function.*

(2) *If Λ is only finite at $\lambda = 0$, then Λ^* is identically zero. If $\Lambda(\lambda) < \infty$ for some $\lambda > 0$, then $\overline{x} = E(X_1) < \infty$ (possibly $\overline{x} = -\infty$), and for all $x \geq \overline{x}$,*

$$\Lambda^*(x) = \sup_{\lambda\geq 0}[\lambda x - \Lambda(\lambda)] \tag{64}$$

is for $x > \overline{x}$, a nondecreasing function. Similarly, if $\Lambda(\lambda) < \infty$ for some $\lambda < 0$, then $\overline{x} = E(X_1) > -\infty$ (possibly $\overline{x} = \infty$), and for all $x \leq \overline{x}$,

$$\Lambda^*(x) = \sup_{\lambda\leq 0}[\lambda x - \Lambda(\lambda)] \tag{65}$$

is for $x < \overline{x}$, a nonincreasing function.

When $\overline{x}$ is finite, $\Lambda^*(\overline{x}) = 0$, and always

$$\inf_{x\in\mathbb{R}}\Lambda^*(x) = 0. \tag{66}$$

4.1. Proof of Theorem 2

Proof. The proof comes in two steps. In step one we will establish a large deviations upper bound and the corresponding lower bound is proved in step two. The proof is an adaptation of arguments used to prove the large deviations principle for the empirical mean of i.i.d sequence of $\mathbb{R}^d$-valued random vectors.

Step 1. Let $I(x) = \Lambda^*(x)$ be the intended rate function of the problem and define the δ-rate function as follows:

$$I^\delta(x) = \min\left\{I(x) - \delta, \frac{1}{\delta}\right\}. \tag{67}$$

The proof of the upper bound will follow if we can show for every $\delta > 0$ and every closed subset F of $\mathbb{R}^d$ that

$$\lim_{n\longrightarrow\infty}\sup\frac{1}{n}\log\mu_n(F) \leq \delta - \inf_{x\in F}I^\delta(x). \tag{68}$$

We will first of all prove this inequality for compact subsets and extend it to closed subsets with the help of exponential tightness argument. Suppose K is a compact subset of $\mathbb{R}^d$. Then for every $\delta > 0$ and $q \in K$, we can find a $\lambda_q \in \mathbb{R}^d$ such that

$$\sum_{i=1}^{q}\alpha^i\left[q^i\lambda_q^i - \Lambda\left(\lambda_q^i\right)\right] \geq I^\delta(q). \tag{69}$$

It follows from the definition of Λ^* that such a λ_q exists. For any $\lambda_q,\ u \in \mathbb{R}^d$, define

$$\lambda_{q,u} = \left(u_1\lambda_q^1, \ldots, u_d\lambda_q^d\right). \tag{70}$$

Therefore if $\overline{n} = (n_1, \ldots, n_d)$, then $(1/n)\lambda_{q,\overline{n}} = (\alpha_n^1\lambda_q^1, \ldots, \alpha_n^d\lambda_q^d)$. For each $q \in \mathbb{R}^d$, choose $\rho_q > 0$ such that $\rho_q|\lambda_q| \leq \delta$ and define

$$B_{q,\rho_q} = \left\{x \in \mathbb{R}^d : |x - q| < \rho_q\right\}, \tag{71}$$

where

$$\begin{aligned}|x| &= \sqrt{\langle x, x\rangle},\\ \langle \lambda, q\rangle &= \sum_{i=1}^{d}\lambda^i q^i.\end{aligned} \tag{72}$$

Note that for any $q \in K$

$$\begin{aligned}-\inf_{x\in B_{q,\rho_q}}\left\langle\lambda_{q,\alpha}, x\right\rangle &\leq \rho_q\left|\lambda_{q,\alpha}\right| - \left\langle\lambda_{q,\alpha}, q\right\rangle\\ &\leq \delta - \left\langle\lambda_{q,\alpha}, q\right\rangle\end{aligned} \tag{73}$$

since $0 \leq \alpha^i \leq 1$ for each $i = 1, 2, \ldots, d$. Therefore it follows from the exponential Chebycheff inequality that

$$\begin{aligned}&\frac{1}{n}\log\mu_n\left(B_{q,\rho_q}\right)\\ &\quad\leq \frac{1}{n}\log E\left[\exp\left(\left\langle\lambda_{q,\overline{n}}, S_n\right\rangle - \inf_{x\in B_{q,\rho_q}}\left\langle\lambda_{q,\overline{n}}, x\right\rangle\right)\right]\\ &\quad= -\inf_{x\in B_{q,\rho_q}}\frac{1}{n}\left\langle\lambda_{q,\overline{n}}, x\right\rangle\\ &\qquad+ \frac{1}{n}\log E\left[\exp\left(\sum_{i=1}^{d}\lambda_q^i\sum_{j\in I_{n_i}}X_j\right)\right]\\ &\quad= -\inf_{x\in B_{q,\rho_q}}\frac{1}{n}\left\langle\lambda_{q,\overline{n}}, x\right\rangle + \frac{1}{n}\sum_{i=1}^{d}n_i\Lambda\left(\lambda_q^i\right)\\ &\quad\leq \delta - \left(\sum_{i=1}^{d}\alpha_n^i\lambda_q^i q^i - \sum_{i=1}^{d}\alpha_n^i\Lambda\left(\lambda_q^i\right)\right).\end{aligned} \tag{74}$$

The second equality uses that the sequence $X_1, X_2, \ldots$ is i.i.d. and the second inequality uses that $n_i/n = \alpha_n^i$. Since K is compact, it has a finite covering consisting of $N = N(K, \delta)$ open balls $B_{q_l,\rho_{q_l}}$ centred at $q_1, \ldots q_N$. It follows from the subadditivity property of probability measures and the choice of λ_{q_l}'s that

$$\begin{aligned}\log\mu_n(K) &\leq \frac{1}{n}\log N + \delta\\ &\quad- \min_{l=1,\ldots,N}\left(\sum_{i=1}^{d}\alpha_n^i\lambda_{q_l}^i q_l^i - \sum_{i=1}^{d}\alpha_n^i\Lambda\left(\lambda_{q_l}^i\right)\right).\end{aligned} \tag{75}$$

Therefore

$$\limsup_{n\to\infty}\log\mu_n(K)$$
$$\begin{aligned}&\le \delta-\limsup_{n\to\infty}\min_{l=1,\dots,N}\left(\sum_{i=1}^{d}\alpha_n^i\left[\lambda_{q_l}^i q_l^i-\Lambda\left(\lambda_{q_l}^i\right)\right]\right)\\&=\delta-\min_{l=1,\dots,N}\left(\sum_{i=1}^{d}\alpha^i\left[\lambda_{q_l}^i q_l^i-\Lambda\left(\lambda_{q_l}^i\right)\right]\right)\\&\le\delta-\min_{l=1,\dots,N}I^\delta\left(q_l\right)\le\delta-\inf_{q\in K}I^\delta(q)\end{aligned}\tag{76}$$

The equality above uses that the quantity $\min_{l=1,\dots,N}(\sum_{i=1}^{d}\alpha_n^i[\lambda_{q_l}^i q_l^i-\Lambda(\lambda_{q_l}^i)])$ is finite, by the finiteness assumption on Λ, which is a concave function of α_n and consequently it is continuous of α_n.

To extend the proof of the above to all closed subsets of $\mathbb{R}^d$, we need to establish exponential tightness of the measures μ_n. Let $\rho>0$ and define $H_\rho=[-\rho,\rho]^d$. Then $H_\rho^c\subset\cup_{i=1}^d\{x\in\mathbb{R}^d:\ |x^i|>\rho\}$ and

$$\mu_n\left(H_\rho^c\right)\le\sum_{i=1}^{d}\mu_n^i\left([\rho,\infty)\right)+\sum_{i=1}^{d}\mu_n^i\left((-\infty,-\rho]\right),\tag{77}$$

where μ_n^i is the law of the ith coordinate of S_n, i.e., the law of $S_n^i=(1/n_i)\sum_{k\in I_{n^i}}X_k$. Thus $\mu_n^i=\mu_{n_i}$. Therefore for any $\lambda^i\ge 0$

$$\begin{aligned}\mu_{n_i}\left([\rho,\infty)\right)&\le E\left[1_{S_n^i-\rho\ge0}\right]\le E\left[e^{n_i\lambda^i(S_n^i-\rho)}\right]\\&=e^{-n_i\lambda^i\rho}\prod_{k\in I_{n_i}}E\left[e^{\lambda^i X_k}\right]=e^{-n_i[\lambda^i\rho-\Lambda(\lambda^i)]}.\end{aligned}\tag{78}$$

Therefore, for every $\rho\ge\overline{x}$, it follows from (64) that

$$\mu_{n_i}\left([\rho,\infty)\right)\le e^{-n_i\Lambda^*(\rho)}.\tag{79}$$

Similarly, using (65) we get that

$$\mu_{n_i}\left((-\infty,-\rho]\right)\le e^{-n_i\Lambda^*(\rho)}.\tag{80}$$

It follows from the fact that $\Lambda(\lambda)<\infty$, for every $\lambda\in\mathbb{R}$, and Lemma 2.2.20 of [21] that $\Lambda^*(x)\longrightarrow\infty$ and $|x|\longrightarrow\infty$. Therefore it follows from (78), (79), and (80) that

$$\lim_{\rho\to\infty}\limsup_{n\to\infty}\frac{1}{n}\log\mu_n\left(H_\rho^c\right)=-\infty.\tag{81}$$

This then implies that the sequence of measures μ_n is exponentially tight. Hence the upper bound (76) holds for every closed set F of $\mathbb{R}^d$.

Step 2. Suppose $y\in\mathscr{D}_{\overline{\Lambda}^*}=\{x\in\mathbb{R}^d:\Lambda^*(x)<\infty\}$ and for some $\eta\in\mathbb{R}^d$, $y^i=\Lambda'(\eta^i)=(d\Lambda/d\lambda)(\eta^i)$, for any $i=1,2,\dots,d$. Define, for any $i=1,\dots,d$, a probability measure $\widetilde{\mu}^i$ on $\mathbb{R}$, as

$$\frac{d\widetilde{\mu}^i}{d\mu}\left(z^i\right)=e^{\eta^i z^i-\Lambda'(\eta^i)}\tag{82}$$

and suppose $X_1,X_2,\dots,X_n$ be an independent sequence of random variables with the law of X_j being $\widetilde{\mu}^i$ if $j\in I_{n_i}$. Let $\widetilde{\mu}_n$ be the corresponding law for S_n. Then

$$d\widetilde{\mu}_n(dz)=\prod_{i=1}^{d}e^{n_i(\eta^i z^i-\Lambda(\eta^i))}d\mu_{n_i}\left(dz^i\right).\tag{83}$$

This then implies that

$$\begin{aligned}d\mu_n(z)&=\prod_{i=1}^{d}e^{n_i(\Lambda(\eta^i)-\eta^i z^i)}d\widetilde{\mu}_n(z)\\&=\prod_{i=1}^{d}e^{n_i(\Lambda(\eta^i)-y^i\eta^i+y^i\eta^i-\eta^i z^i)}d\widetilde{\mu}_n(z).\end{aligned}\tag{84}$$

Therefore

$$\frac{1}{n}\log\mu\left(B_{y,\delta}\right)$$
$$\begin{aligned}&=\frac{1}{n}\log\int_{B_{y,\delta}}e^{\sum_{i=1}^{d}n_i(\Lambda(\eta^i)-y^i\eta^i+y^i\eta^i-\eta^i z^i)}d\widetilde{\mu}_n(z)\\&=\frac{1}{n}\log e^{\sum_{i=1}^{d}n_i(\Lambda(\eta^i)-y^i\eta^i)}\\&\quad+\frac{1}{n}\log\int_{z\in B_{y,\delta}}e^{\sum_{i=1}^{d}n_i(\eta^i(y^i-z^i))}d\widetilde{\mu}_n(z)\\&=\sum_{i=1}^{d}\alpha_n^i\left[\Lambda\left(\eta^i\right)-y^i\eta^i\right]\\&\quad+\frac{1}{n}\log\int_{z\in B_{y,\delta}}e^{\langle\eta_{\overline{n}},y-z\rangle}d\widetilde{\mu}_n(z)\\&\ge\sum_{i=1}^{d}\alpha_n^i\left[\Lambda\left(\eta^i\right)-y^i\eta^i\right]-\frac{1}{n}\left|\eta_{\overline{n}}\right|\delta\\&\quad+\frac{1}{n}\log\widetilde{\mu}_n\left(B_{y,\delta}\right)\\&=\sum_{i=1}^{d}\alpha_n^i\left[\Lambda\left(\eta^i\right)-y^i\eta^i\right]-\sqrt{\sum_{i=1}^{d}\left(\alpha_n^i\right)^2\left(\eta^i\right)^2}\delta\\&\quad+\frac{1}{n}\log\widetilde{\mu}_n\left(B_{y,\delta}\right)\\&\ge\sum_{i=1}^{d}\alpha_n^i\left[\Lambda\left(\eta^i\right)-y^i\eta^i\right]-|\eta|\delta+\frac{1}{n}\log\widetilde{\mu}_n\left(B_{y,\delta}\right),\end{aligned}\tag{85}$$

where $\eta_{\overline{n}}=(n_1\eta^1,\dots,n_d\eta^d)$. For any $i=1,\dots,d$, let X_j be such that $j\in I_{n_i}$. Then

$$\mathbb{E}_{\widetilde{\mu}^i}\left(X_j\right)=\frac{1}{\mathscr{M}\left(\eta^i\right)}\int xe^{\eta^i x}d\mu(x)\tag{86}$$

and by the dominated convergence theorem

$$\mathbb{E}_{\widetilde{\mu}^i}\left(X_j\right)=\Lambda'\left(\eta^i\right)=y^i.\tag{87}$$

Therefore by the weak law of large numbers, $\lim_{n\longrightarrow\infty}\widetilde{\mu}_n(B_{y,\delta}) = 1$ for all $\delta > 0$, since the ith component of $\widehat{S}_n$ converges in probability to y^i. Since

$$\liminf_{n\longrightarrow\infty}\sum\alpha_n^i\left[\Lambda\left(\eta^i\right)-y^i\eta^i\right]=\sum\alpha^i\left[\Lambda\left(\eta^i\right)-y^i\eta^i\right] \quad (88)$$

we have that

$$\liminf_{n\longrightarrow\infty}\frac{1}{n}\log\mu_n\left(B_{y,\delta}\right)\geq\sum_{i=1}^{d}\alpha^i\left[\Lambda\left(\eta^i\right)-y^i\eta^i\right]-|\eta|\,\delta, \quad (89)$$

and by the definition of Λ^*

$$\Lambda^*\left(y\right)\geq\sum_{i=1}^{d}\alpha^i\left[-\Lambda\left(\eta^i\right)+y^i\eta^i\right]. \quad (90)$$

Hence

$$\liminf_{n\longrightarrow\infty}\frac{1}{n}\log\mu_n\left(B_{y,\delta}\right)\geq\overline{\Lambda}^*\left(y\right)-|\eta|\,\delta. \quad (91)$$

Therefore

$$\begin{aligned}\liminf_{n\longrightarrow\infty}\frac{1}{n}\log\mu_n\left(B_{y,\delta}\right)&\geq\lim_{\delta\longrightarrow0}\liminf_{n\longrightarrow\infty}\frac{1}{n}\log\mu_n\left(B_{y,\delta}\right)\\&\geq-\overline{\Lambda}^*\left(y\right).\end{aligned} \quad (92)$$

Next, suppose that $y\in\mathscr{D}_{\overline{\Lambda}^*}$, but there is no $\eta\in\mathbb{R}^d$ such that $y^i = \Lambda'(\eta^i)$ for all $i = 1,\ldots,d$. Here we change the common law μ of the $X_1, X_2, \ldots$ sequence by adding a small normal random variable so that the perturbed μ will admit logarithmic moment generating function Λ_M, for which we can find an $\eta\in\mathbb{R}^d$ that satisfies $\Lambda'_M(\eta^i) = y^i$, for every $i = 1,\ldots,d$. With this we get a lower bound for the large deviations principle which is later used to deduce the lower for the unperturbed case. Formerly, let $Y_1, Y_2, \ldots$ be an i.i.d sequence given by

$$Y_i = X_i + \frac{V_i}{\sqrt{M}}, \quad (93)$$

where V_i is a sequence of independent and identically distributed standard normal random variables that are independent of the $X_1, X_2, \ldots$ sequence. The quantity $M < \infty$. Let ν be the common law of the Y_i's and $S_n^{(M)} = (S_{n,M}^i, \ldots, S_{n,M}^d)$, where

$$S_{n,M}^i = \frac{1}{n_i}\sum_{j\in I_{n_i}}Y_j, \quad \text{for } i = 1,2,\ldots,d. \quad (94)$$

Then

$$\begin{aligned}\Lambda_M\left(\lambda^1\right)&=\log\mathbb{E}\left(e^{\lambda^1Y_1}\right)=\Lambda\left(\lambda^1\right)+\frac{1}{2M}\left(\lambda^1\right)^2\\&\geq\Lambda\left(\lambda^1\right).\end{aligned} \quad (95)$$

Therefore

$$\begin{aligned}\overline{\Lambda}^*\left(y\right)&=\sup_{\lambda\in\mathbb{R}^d}\sum\alpha^i\left[y^i\lambda^i-\Lambda\left(\lambda^i\right)\right]\\&\geq\sup_{\lambda\in\mathbb{R}^d}\sum\alpha^i\left[y^i\lambda^i-\Lambda_M\left(\lambda^i\right)\right].\end{aligned} \quad (96)$$

Recall that $\overline{x} = \mathbb{E}(X_1) < \infty$. Hence by the Jensen's inequality we have that $\Lambda(\lambda)\geq\lambda\overline{x}$, for all $\lambda\in\mathbb{R}$. Therefore the function

$$g\left(\lambda\right)=\sum\alpha^i\left[\lambda^iy^i-\Lambda_M\left(\lambda^i\right)\right] \quad (97)$$

is finite, differentiable and it satisfies

$$\lim_{\rho\longrightarrow\infty}\sup_{|\lambda|>\rho}g\left(\lambda\right)=-\infty. \quad (98)$$

Therefore g(·) attains its supremum over $\mathbb{R}^d$ at some $\eta\in\mathbb{R}^d$, for which

$$0=\frac{\partial g}{\partial\lambda^i}\left(\eta^i\right)=\alpha^i\left(y^i-\Lambda'_M\left(\eta^i\right)\right),\quad i=1,\ldots,d. \quad (99)$$

This implies that

$$y^i=\Lambda'_M\left(\eta^i\right). \quad (100)$$

Therefore by the proceeding proof we have large deviations lower bound for the sequence of probability measures $\{\nu_n\}$, i.e., for every $\delta > 0$,

$$\liminf_{n\longrightarrow\infty}\log\nu_n\left(B_y,\delta\right)\geq-\overline{\Lambda}_M^*\left(y\right)\geq-\overline{\Lambda}^*\left(y\right)>-\infty. \quad (101)$$

Note that the process $S_n^{(M)}$ has the same distribution as $S_n + V/\sqrt{Mn}$, where V is the standard multivariate normal random vector. Therefore, in distributional sense

$$S_n=S_n^{(M)}-\frac{V}{\sqrt{Mn}}, \quad (102)$$

and using that

$$\begin{aligned}&\left\{S_n\in B_{y,2\delta}\right\}\\&\quad\supset\left\{S_n^{(M)}\in B_{y,\delta}\right\}\setminus\left\{S_n^{(M)}\in B_{y,\delta},\frac{V}{\sqrt{Mn}}\in B_{y,\delta}^c\right\}\end{aligned} \quad (103)$$

we have

$$\begin{aligned}\mathbb{P}\left(S_n\in B_{y,2\delta}\right)&\geq\mathbb{P}\left(S_n^{(M)}\in B_{y,\delta}\right)\\&\quad-\mathbb{P}\left(S_n^{(M)}\in B_{y,\delta},\frac{V}{\sqrt{Mn}}\in B_{y,\delta}^c\right).\end{aligned} \quad (104)$$

Therefore

$$\begin{aligned}\mu_n\left(B_{y,2\delta}\right)&\geq\nu_n\left(B_{y,\delta}\right)-\mathbb{P}\left(\frac{V}{\sqrt{Mn}}\in B_{y,\delta}^c\right)\\&=\nu_n\left(B_{y,\delta}\right)-\mathbb{P}\left(|V|\geq\sqrt{Mn}\delta\right)\\&\geq\nu_n\left(B_{y,\delta}\right)+\frac{M\delta^2n}{2}\geq\nu_n\left(B_{y,\delta}\right).\end{aligned} \quad (105)$$

Thus

$$\begin{aligned}\liminf_{n\longrightarrow\infty}\mu_n\left(B_{y,2\delta}\right)&\geq\liminf_{n\longrightarrow\infty}\nu_n\left(B_{y,\delta}\right)\geq-\overline{\Lambda}^*\left(y\right)\\&=-\inf_{x\in B_{y,2\delta}}\overline{\Lambda}^*\left(x\right).\end{aligned} \quad (106)$$

□

4.2. Proof of Theorem 3

Proof. The proof of this theorem follows from that of the $\mathbb{R}$-valued case upon making appropriate substitutions. For instance, every $\mathbb{R}$, $\mathbb{R}^d$, $\lambda^i y^i$, and $\Lambda'(\eta^i)$ in the proof of the $\mathbb{R}$−valued case should be replaced with $\mathbb{R}^l$, $(\mathbb{R}^l)^d$, $\langle\lambda^i, y^i\rangle$, and $\nabla\Lambda(\eta^i)$, respectively. In particular, in establishing the exponential tightness of the μ_n's we use that for any $\rho > 0$, we define $H_\rho = [-\rho, \rho]^{ld}$. Then $H_\rho^c \subset \cup_{i=1}^d \cup_{j=1}^l \{x \in \mathbb{R}^d : |x_j^i| > \rho\}$ and

$$\begin{aligned}\mu_n\left(H_\rho^c\right) \leq \sum_{i=1}^d \sum_{j=1}^l \mu_{n_i}^j\left([\rho, \infty)\right) \\ + \sum_{i=1}^d \sum_{j=1}^l \mu_{n_i}^j\left((-\infty, -\rho]\right),\end{aligned} \tag{107}$$

where x_j^i is the jth coordinate of the ith-component of $x = (x^1, \ldots, x^d) \in (\mathbb{R}^l)^d$ and $\mu_{n_i}^j$ is the law of the empirical mean

$$S_n^{i,j} = \frac{1}{n_i} \sum_{k \in I_{n_i}} X_k^j. \tag{108}$$

Here X_k^j is the jth coordinate of X_k. □

4.3. Proof of Theorem 4

Proof. The proof follows from the large deviations principle for vector of empirical means of $\mathbb{R}^{|\Sigma|}$-valued random vectors restricted to probability vectors in $\mathbb{R}^{|\Sigma|}$. Recall from (17) that $\Lambda(\cdot)$ is finite and differentiable convex function. Therefore the map $\lambda \longmapsto \sum_{i=1}^d \alpha^i[\langle\lambda^i, \nu^i\rangle - \Lambda(\lambda^i)]$ is concave. Therefore, the supremum of this map is attained at the value of λ such that

$$\nu^i - \nabla\Lambda\left(\lambda^i\right) = 0, \quad \text{for every } i = 1, 2, \ldots, d. \tag{109}$$

Let $c^i = \sum_{j=1}^{|\Sigma|} e^{\lambda_j^i}\mu_j$. Here λ_j^i is the jth component of the vector $\lambda^i \in \mathbb{R}^{|\Sigma|}$. Then for every $i = 1, \ldots, d$ and $j = 1, \ldots |\Sigma|$ it follows from (109) that

$$\lambda_i^j = \log\left(\frac{\nu_j^i}{\mu_j}\right) + \log c^i. \tag{110}$$

Substituting this form of λ into $\sum_{i=1}^d \alpha^i[\langle\lambda^i, \nu^i\rangle - \Lambda(\lambda^i)]$ leads to

$$\overline{\Lambda}^*(\nu) = \sum_{i=1}^d \alpha^i \sum_{j=1}^{|\Sigma|} \nu_j^i \log\left(\frac{\nu_j^i}{\mu_j}\right) = \sum_{i=1}^d \alpha^i H\left(\nu^i \mid \mu\right). \tag{111}$$

□

5. Proof of Theorem 5

Proof. The proof comes in three steps. In step one we proof a lower for (23). Using condition (21), we proof an upper bound for (23) in step two. Step three shows that condition (22) implies condition (21) and that completes the proof.

Step One. Fix $x \in \mathcal{X}$ and $\delta > 0$. Due to the equi-continuity of the sequence ϕ_n we have that the functions ϕ_n are lower semicontinuous. Thus, there exists a neighbourhood G of x such that

$$\inf_{y \in G} \phi_n(y) \geq \phi_n(x) - \delta, \quad \text{for all } n = 1, 2, 3, \ldots. \tag{112}$$

We then have that

$$E\left(e^{n\phi_n(S_n)}\right) \geq E\left(e^{n\phi_n(S_n)} 1_{S_n \in G}\right) \geq e^{n[\phi_n(x) - \delta]} \mu_n(G). \tag{113}$$

We get from here that

$$\begin{aligned}&\liminf_{n\longrightarrow\infty} \frac{1}{n} \log E\left(e^{n\phi_n(S_n)}\right) \\ &\quad \geq \liminf_{n\longrightarrow\infty} \phi_n(x) - \delta + \liminf_{n\longrightarrow\infty} \frac{1}{n} \log \mu_n(G) \\ &\quad \geq \phi(x) - \delta - \inf_{y \in G} I(y) \geq \phi(x) - I(x) - \delta.\end{aligned} \tag{114}$$

Since $x \in \mathcal{X}$ and $\delta > 0$ were arbitrary chosen, we have that

$$\liminf_{n\longrightarrow\infty} \frac{1}{n} \log E\left(e^{n\phi_n(S_n)}\right) \geq \sup_{x \in \mathcal{X}} [\phi(x) - I(x)]. \tag{115}$$

Step Two. Suppose that the functions are ϕ_n and uniformly bounded; i.e., there is a constant M, such that $\sup_{x\in\mathcal{M}} \phi_n(x) \leq M$. Then condition (21) holds. Let $\alpha < \infty$ and $\delta > 0$ and define $\Psi_I(\alpha) = \{x \in \mathcal{X} : I(x) \leq \alpha\}$. Then Ψ_I is a compact level set since I is a good rate function. It follows from the lower semicontinuity property of I, the upper semicontinuity property of ϕ_n, and the fact that $\mathcal{X}$ is a regular topological space that for every $x \in \Psi_I(\alpha)$, there is a neighbourhood A_x of x such that

$$\begin{aligned}&\inf_{y \in \overline{A_x}} I(y) \geq I(x) - \delta, \\ &\sup_{y \in \overline{A_x}} \phi_n(y) \leq \phi_n(x) + \delta, \quad \text{for all } n = 1, 2, 3, \ldots.\end{aligned} \tag{116}$$

Since $\Psi_I(\alpha)$ is compact, there are finite number of points $x_1, x_2, \ldots, x_N \in \Psi_I(\alpha)$ such that the neighbourhoods $A_{x_1}, A_{x_2}, \ldots, A_{x_N}$ cover $\Psi_I(\alpha)$. Therefore

$$\begin{aligned}E\left[e^{n\phi_n(S_n)}\right] &\leq \sum_{i=1}^N E\left[e^{n\phi_n(S_n)} 1_{\{S_n \in A_{x_i}\}}\right] \\ &\quad + e^{nM} \mu_n\left(\left(\bigcup_{i=1}^N A_{x_i}\right)^c\right) \\ &\leq \sum_{i=1}^N e^{n[\phi_n(x_i) + \delta]} \mu_n\left(\overline{A_{x_i}}\right) \\ &\quad + e^{nM} \mu_n\left(\left(\bigcup_{i=1}^N A_{x_i}\right)^c\right).\end{aligned} \tag{117}$$

The last inequality follows from (116). Note that $(\bigcup_{i=1}^{N} A_{x_i})^c \subset \Psi_I(\alpha)^c$ and

$$\limsup_{n\longrightarrow\infty}\frac{1}{n}\log E\left[e^{n\phi_n(S_n)}\right] \leq \max\left\{\max_{i=1}^{N}\left\{\limsup_{n\longrightarrow\infty}\phi_n(x_i)+\delta-\inf_{y\in\overline{A}_{x_i}} I(y)\right\},\right.$$
$$\left. M-\inf_{y\in(\bigcup_{i=1}^{N}A_{x_i})^c} I(y)\right\} \leq \max\left\{\max_{i=1}^{N}\{\phi(x_i)-I(x_i)+2\delta\}, M-\alpha\right\} \leq \max\left\{\sup_{x_i\in\mathcal{X}}\{\phi(x_i)-I(x_i)\}, M-\alpha\right\}+2\delta. \tag{118}$$

Since the ϕ_n's are bounded above, so is the limiting ϕ. The upper bound of interest follows upon taking the limits $\delta \longrightarrow 0$ and $\alpha \longrightarrow \infty$.

If the ϕ_n's are not uniformly bounded above, we set $\phi_n^M(x) = \phi_n(x) \wedge M \leq \phi_n(x)$. Then ϕ_n^M's are uniformly bounded above and using the arguments for the above proof we get that

$$\limsup_{n\longrightarrow\infty}\frac{1}{n}\log E\left[e^{n\phi_n(S_n)}\right] \leq \sup_{x\in\mathcal{X}}\left[\phi(x)-I(x)\right] \vee \limsup_{n\longrightarrow\infty}\frac{1}{n}\log E\left[e^{n\phi_n(S_n)}1_{\{\phi_n(S_n)\geq M\}}\right]. \tag{119}$$

The result for this case follows from the tail condition (21) as we let $M \longrightarrow \infty$.

Step Three. In this step we show that condition (22) implies the tail condition (21). In this regard we let $\epsilon > 0$, $\gamma > 1$, and define $X_n = \exp(n[\phi_n(S_n) - M])$. Then

$$e^{-nM}E\left[e^{n\phi_n(S_n)}1_{\{\phi_n(S_n)\geq M\}}\right] = E\left[X_n 1_{\{X_n\geq 1\}}\right] \leq E\left[X_n^{\gamma}\right] \leq e^{-n\gamma M}E\left[e^{n\gamma\phi_n(S_n)}\right]. \tag{120}$$

It follows from here that

$$\limsup_{n\longrightarrow\infty}\frac{1}{n}\log E\left[e^{n\phi_n(S_n)}1_{\{\phi_n(S_n)\geq M\}}\right] \leq M(1-\gamma)+\limsup_{n\longrightarrow\infty}\frac{1}{n}\log E\left[e^{n\gamma\phi_n(S_n)}\right]. \tag{121}$$

Therefore condition (22) implies condition (21) upon taking $M \longrightarrow \infty$ limit. □

6. Conclusion

This paper has developed large deviations machinery for the empirical means and measures for partitions of independent and identically distributed sequence of random variables. The large deviations result is further applied to derive the limiting free energy for multipopulation Curie-Weiss and Potts' models. The method proposed here can be applied to multipopulation versions of spin models with continuous spins such as the $O(n)$ model [11], the spherical model [12, 13], the liquid crystal model [11, 14, 15], and the Kuramoto model [16]. The multipopulation Potts' model may have applications in discrete choice context with more than two alternatives to choose from. This serves as a natural extension to the Ising cases considered in [8–10, 22, 23].

The knowledge gained from the study of the minimizers of the associated minimization problem that leads to the limiting pressure will offer insight to the scaling limit behaviour of the empirical measures associated with the multipopulation Potts' model. This will be a natural extension of the work in [24].

Data Availability

This paper uses no data.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors thank the staff at the Mathematics Department of University of Energy and Natural Resources for their support and kindness during the period in which this paper was written.

References

[1] E. G. D. Cohen, "Tricritical points in metamagnets and helium mixtures," in *Proceedings of the Fundamental Problems in Statistical Mechanics, Proceedings of the 1974 Wageningen Summer School*, 1974.

[2] J. M. Kincaid and E. G. D. Cohen, "Phase diagrams of liquid helium mixtures and metamagnets: Experiment and mean field theory," *Physics Reports*, vol. 22, no. 2, pp. 57–143, 1975.

[3] C. Külske and A. Le Ny, "Spin-flip dynamics of the Curie-Weiss model: loss of Gibbsianness with possibly broken symmetry," *Communications in Mathematical Physics*, vol. 271, no. 2, pp. 431–454, 2007.

[4] W. A. Brock and S. N. Durlauf, "Discrete choice with social interactions," *Review of Economic Studies*, vol. 68, no. 2, pp. 235–260, 2001.

[5] S. N. Durlauf, "How can statistical mechanics contribute to social science?" *Proceedings of the National Acadamy of Sciences of the United States of America*, vol. 96, no. 19, pp. 10582–10584, 1999.

[6] W. A Brock and S. N. Durlauf, *Discrete choice with social interaction I: Theory, NBER Working Paper Series*, 1995.

[7] D. McFadden, "Economic choices," *American Economic Review*, vol. 91, no. 3, pp. 351–378, 2001.

[8] I. Gallo and P. Contucci, "Bipartite mean field spin systems. Existence and solution," *Mathematical Physics Electronic Journal*, vol. 14, Paper 1, 21 pages, 2008.

[9] I. Gallo, A. Barra, and P. Contucci, "Parameter evaluation of a simple mean-field model of social interaction," *Mathematical Models and Methods in Applied Sciences*, vol. 19, no. suppl., pp. 1427–1439, 2009.

[10] I. Gallo, *An equilibrium approach to modelling social interaction [Ph.D. thesis]*, University of Bologna, Bologna, 2009.

[11] K. Kirkpatrick and T. Nawaz, "Asymptotics of mean-field $O(N)$ models," *Journal of Statistical Physics*, vol. 165, no. 6, pp. 1114–1140, 2016.

[12] M. Kac and C. J. Thompson, "Spherical model and the infinite spin dimensionality limit," *Physica Norvegica*, vol. 5, no. 3-4, pp. 163–168, 1971.

[13] M. Kastner and O. Schnetz, "On the mean-field spherical model," *Journal of Statistical Physics*, vol. 122, no. 6, pp. 1195–1213, 2006.

[14] J. L. Ericksen, "Liquid crystals with variable degree of orientation," *Archive for Rational Mechanics and Analysis*, vol. 113, no. 2, pp. 97–120, 1990.

[15] M. Fannes, H. Spohn, and A. Verbeure, "Equilibrium states for mean field models," *Journal of Mathematical Physics*, vol. 21, no. 2, pp. 355–358, 1980.

[16] M. Antoni and S. Ruffo, "Clustering and relaxation in Hamiltonian long-range dynamics," *Physical Review E*, 1995.

[17] A. A. Opoku and G. Osabutey, "A multi-population mean-field Potts model for discrete choice with social interaction".

[18] A. A. Opoku and G. Osabutey, "A multi-population Curie-Weiss model for binary choice with social interaction".

[19] A. A. Opoku, K. Owusu Edusei, and R. . Ansah, "A conditional Curie-Weiss model for stylized multi-group binary choice with social interaction," *Journal of Statistical Physics*, vol. 171, no. 1, pp. 106–126, 2018.

[20] W. Rudin, *Principles of Mathematical Analysis*, McGraw-Hill Book Co., New York, NY, USA, 3rd edition, 1964.

[21] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, vol. 38 of *Applications of Mathematics (New York)*, Springer, New York, NY, USA, 2nd edition, 1998.

[22] P. Contucci, I. Gallo, and S. Ghirlanda, *Equilibria of culture contact derived from ingroup and outgroup attitudes*, vol. 5, 2008.

[23] A. Bianchi, P. Contucci, and C. Giardina, "Thermodynamic limit for mean-field spin models," *Mathematical Physics Electronic Journal*, vol. 9, p. 15, 2003.

[24] M. Fedele and P. Contucci, "Scaling limits for multi-species statistical mechanics mean-field models," *Journal of Statistical Physics*, vol. 144, no. 6, pp. 1186–1205, 2011.