# When does the chaos in the Curie-Weiss model stop to propagate?[*]

Jonas Jalowy[†]     Zakhar Kabluchko[‡]     Matthias Löwe[§]

Alexander Marynych[¶]

### Abstract

We investigate increasing propagation of chaos for the mean-field Ising model of ferromagnetism (also known as the Curie-Weiss model) with $N$ spins at inverse temperature $\beta > 0$ and subject to an external magnetic field of strength $h \in \mathbb{R}$. Using a different proof technique than in Ben Arous and Zeitouni [Ann. Inst. H. Poincaré: Probab. Statist., 35(1): 85–102, 1999] we confirm the well-known propagation of chaos phenomenon: If $k = k(N) = o(N)$ as $N \to \infty$, then the $k$'th marginal distribution of the Gibbs measure converges to a product measure at $\beta < 1$ or $h \neq 0$ and to a mixture of two product measures, if $\beta > 1$ and $h = 0$. More importantly, we also show that if $k(N)/N \to \alpha \in (0, 1]$, this property is lost and we identify a non-zero limit of the total variation distance between the number of positive spins among any $k$-tuple and the corresponding binomial distribution.

## 1 Introduction

The Curie-Weiss model is a mean-field model for ferromagnetism in statistical mechanics. It is described by a sequence of probability measures, the Gibbs measures $\mu_N$, on the sets $\{-1, +1\}^N$. These measures are parametrized by a positive parameter $\beta > 0$ known as the inverse temperature and a real parameter $h \in \mathbb{R}$ interpreted as strength of

an external magnetic field. Given such $\beta > 0$ and $h \in \mathbb{R}$, the Gibbs measure takes the following form:

$$\mu_N(\sigma) := \mu_{N,\beta,h}(\sigma) := \frac{\exp\left(\frac{\beta}{2N}\sum_{i,j=1}^{N}\sigma_i\sigma_j + h\sum_{i=1}^{N}\sigma_i\right)}{Z_N} \qquad \sigma := (\sigma_i)_{i=1}^{N} \in \{-1, +1\}^N.$$
(1.1)

The normalizing constant

$$Z_N = Z_N(\beta, h) = \sum_{\sigma' \in \{-1,+1\}^N} \exp\left(\frac{\beta}{2N}\sum_{i,j=1}^{N}\sigma_i'\sigma_j' + h\sum_{i=1}^{N}\sigma_i'\right)$$
(1.2)

is called the partition function of the model.

There is a vast literature on the Curie-Weiss model with the main asymptotic results summarized in the textbooks [5], [11] and [14]; see also the papers [6, 10, 12]. The order parameter of the model is called the magnetization and is defined by

$$m_N := m_N(\sigma) := \frac{1}{N}\sum_{i=1}^{N}\sigma_i = \frac{2}{N}\mathcal{P}_N - 1,$$

where $\mathcal{P}_N := \mathcal{P}_N(\sigma) := |\{i = 1, \ldots, N : \sigma_i = +1\}|$ is the number of positive spins. Let $\mu_N \circ m_N^{-1}$ denote the distribution of the random variable $m_N$ under the probability measure $\mu_N$. The first-order limiting behavior of the magnetization is given by

$$\mu_N \circ m_N^{-1} \Rightarrow \begin{cases} \delta_{\mathfrak{m}(\beta,h)}, & \text{if } h \neq 0 \quad \text{or} \quad 0 < \beta \leq 1, \\ \frac{1}{2}\left(\delta_{\mathfrak{m}(\beta,0)} + \delta_{-\mathfrak{m}(\beta,0)}\right), & \text{if } h = 0 \quad \text{and} \quad \beta > 1, \end{cases}$$
(1.3)

which indicates a phase transition at $\beta = 1$ in the absence of the external magnetic field ($h = 0$). Here, $\Longrightarrow$ denotes weak convergence, $\delta_x$ is the Dirac measure at $x$ and $\mathfrak{m}(\beta, h)$ is the largest in absolute value solution to

$$z = \tanh(\beta z + h).$$
(1.4)

This solution is unique and positive if $h > 0$, unique and negative if $h < 0$, and it is equal to zero if $h = 0$ and $0 < \beta \leq 1$. If $h = 0$ and $\beta > 1$, equation (1.4) has two non-zero solutions $\mathfrak{m}(\beta, 0)$ and $-\mathfrak{m}(\beta, 0)$.

For the purpose of the present paper it is important that the weak convergence (1.3) is accompanied by the corresponding Gaussian approximations whenever $\beta \neq 1$ or $h \neq 0$. Let us briefly summarize these, postponing the rigorous statements to Section 2 below. By definition, $|\mathfrak{m}(\beta, h)| < 1$ and it is also true that $0 < \beta(1 - \mathfrak{m}^2(\beta, h)) < 1$ for all $\beta > 0$, $h \in \mathbb{R}$ excluding the "critical" case $(\beta, h) = (1, 0)$. Put

$$\mathrm{v}_{\beta,h}^2 := \frac{1 - \mathfrak{m}^2(\beta, h)}{1 - \beta(1 - \mathfrak{m}^2(\beta, h))}, \quad (\beta, h) \neq (1, 0).$$
(1.5)

If $0 < \beta < 1$ and $h = 0$, then $\sqrt{N}m_N$ has Gaussian fluctuations under $\mu_N$ with expectation $\mathfrak{m}(\beta, 0) = 0$ and variance $\mathrm{v}_{\beta,0}^2 = 1/(1 - \beta)$. Of course, this cannot be true for $\beta = 1$, $h = 0$. In this case, the distribution of $N^{1/4}m_N$ has a limiting density

$$f_1(x) = \frac{\exp(-\frac{1}{12}x^4)}{\int_{\mathbb{R}}\exp(-\frac{1}{12}y^4)dy}.$$
(1.6)

For $\beta > 1$, still assuming $h = 0$, there is a conditional central limit theorem for $\sqrt{N}(m_N - \mathfrak{m}(\beta, 0))$ (respectively $\sqrt{N}(m_N + \mathfrak{m}(\beta, 0))$, conditioned on $m_N > 0$ (respectively, $m_N < 0$). In this case the limiting expectation is 0 and the limiting variance is $\mathrm{v}_{\beta,0}^2$. Finally, if $h \neq 0$, there are no phase transitions as $\beta$ varies in $(0, \infty)$ and $\sqrt{N}(m_N - \mathfrak{m}(\beta, h))$ converges to the centred normal distribution with variance $\mathrm{v}_{\beta,h}^2$.

## 1.1 Propagation of chaos and the main results

For the time being, fix $k \in \mathbb{N}$ and pick any $k$-tuple among $N$ spins. The propagation of chaos paradigm for Gibbs measure states that for a mean-field model as the Curie-Weiss model the marginal distributions of the $k$ spins become asymptotically independent. This shall be investigated in the sequel. Since the family of random variables $(\sigma_i)_{i=1}^N$ is exchangeable under the Gibbs measure $\mu_N$, without loss of generality we may pick the first $k$ spins and consider their marginal distribution $\mu_{N,\beta,h}^{(k)}$. Let $\mathcal{P}_k := |\{j \in \{1, \ldots, k\} : \sigma_j = +1\}|$ be the number of positive spins among the picked ones. Note that $\mathcal{P}_k$ completely determines $\mu_{N,\beta,h}^{(k)}$, hence we might as well study the distribution of $\mathcal{P}_k$ under $\mu_N$, denoted by $\mu_N \circ \mathcal{P}_k^{-1}$. Intuitively, if $h = 0$ and $0 < \beta < 1$, as $N \to \infty$, the first $k$ spins should indeed be asymptotically independent and take values $\pm 1$ with the same probability $1/2$. This implies that the distribution of $\mathcal{P}_k$ should be close to the binomial distribution with parameters $k$ and $1/2$. This can formally be written as follows. Let $\text{Bin}(n, p)$ denote a binomial distribution with parameters $n \in \mathbb{N}_0$ and $p \in [0, 1]$:

$$\text{Bin}(n, p)(\{j\}) = \binom{n}{j} p^j (1-p)^{n-j}, \quad 0 \le j \le n. \tag{1.7}$$

Throughout the paper we shall slightly abuse the notation and write mixtures of distributions by simply writing a mixing distribution instead of a parameter which is being mixed. For example, for any distribution $\mathcal{L}(W)$ of a random variable $W$ on $[0, 1]$, the distribution $\text{Bin}(n, \mathcal{L}(W))$ should be understood as

$$\text{Bin}(n, \mathcal{L}(W))(\{j\}) = \mathbb{E}\big( \text{Bin}(n, W)(\{j\}) \big).$$

Recall that the total variation distance $d_{TV}$ between two probability measures $\mathfrak{M}_1$ and $\mathfrak{M}_2$ on $\mathbb{R}$ is defined by

$$d_{TV}(\mathfrak{M}_1, \mathfrak{M}_2) = \sup_{A \in \mathcal{B}(\mathbb{R})} |\mathfrak{M}_1(A) - \mathfrak{M}_2(A)|,$$

where $\mathcal{B}(\mathbb{R})$ is the Borel sigma-algebra. Recall also that $k \in \mathbb{N}$ is fixed for the time being. The 'propagation of chaos' phenomenon for the Curie-Weiss model tells us that, if $h = 0$ and $0 < \beta < 1$, then

$$\lim_{N \to \infty} d_{TV}(\mu_N \circ \mathcal{P}_k^{-1}, \text{Bin}(k, 1/2)) = 0. \tag{1.8}$$

On the other hand, for $h = 0$ and $\beta > 1$, it is known that

$$\lim_{N \to \infty} d_{TV}\left( \mu_N \circ \mathcal{P}_k^{-1}, \text{Bin}\left( k, \tfrac{1}{2}(\delta_{(1+\mathtt{m}(\beta,0))/2} + \delta_{(1-\mathtt{m}(\beta,0))/2}) \right) \right) = 0. \tag{1.9}$$

Note that the approximating distribution is a mixture of two binomial distributions, one with success probability $(1+\mathtt{m}(\beta,0))/2$ and the other one with $(1-\mathtt{m}(\beta,0))/2$. Finally, it is known that for $h \ne 0$ (and arbitrary $\beta > 0$), relation (1.8) holds with $1/2$ replaced by $(1+\mathtt{m}(\beta,h))/2$.

Assume now that $k = k(N)$ depends on $N$ and $\lim_{N \to \infty} k(N) = \infty$. The aim of the this note is to answer the question whether the analogues of (1.8) and (1.9) hold true in this case and if not, what is a threshold on the growth of $k(N)$ such that the limit in (1.8) or (1.9) becomes non-zero and what is the value of the limit. Note that for $k(N)/N \to 0$, that is $k = o(N)$, propagation of chaos has been explicitly shown to hold for $\beta \ne 1$ and $h = 0$ in [4] (note that the authors consider relative entropy rather than total variation distance and that their model also covers $h \ne 0$ implicitly). Our answer is provided by Theorem 1.1 below, which is our main result. To simplify its formulation, let us introduce additional notation.

For $\mathtt{m} \in \mathbb{R}$ and $\mathtt{v}^2 > 0$ put

$$\varphi(t; \mathtt{m}, \mathtt{v}^2) := \frac{1}{\mathtt{v}\sqrt{2\pi}} \exp\left(-\frac{(t-\mathtt{m})^2}{2\mathtt{v}^2}\right), \quad t \in \mathbb{R}. \tag{1.10}$$

That is to say, $t \mapsto \varphi(t; \mathtt{m}, \mathtt{v}^2)$ is the density of a Gaussian distribution with mean $\mathtt{m}$ and variance $\mathtt{v}^2$. Define

$$D(\mathtt{v}_1^2, \mathtt{v}_2^2) := \frac{1}{2} \int_{\mathbb{R}} \left|\varphi(t; 0, \mathtt{v}_1^2) - \varphi(t; 0, \mathtt{v}_2^2)\right| \mathrm{d}t \tag{1.11}$$

to be the total variation distance between two centred Gaussian distributions, which can easily be calculated in terms of the error function. Then our main result reads as follows.

**Theorem 1.1.** *Let $k = k(N)$ be any sequence of positive integers such that*

$$\lim_{N \to \infty} k(N) = +\infty, \quad \lim_{N \to \infty} \frac{k(N)}{N} = \alpha \in [0, 1]. \tag{1.12}$$

*If $h = 0$ and $\beta \in (0, 1)$, then*

$$\lim_{N \to \infty} d_{TV}(\mu_N \circ \mathcal{P}_{k(N)}^{-1}, \mathrm{Bin}(k(N), 1/2)) = D\left(\frac{1 - \beta + \alpha\beta}{4(1-\beta)}, \frac{1}{4}\right). \tag{1.13}$$

*If $h = 0$ and $\beta > 1$, then*

$$\lim_{N \to \infty} d_{TV}\left(\mu_N \circ \mathcal{P}_{k(N)}^{-1}, \mathrm{Bin}\left(k(N), \frac{1}{2}(\delta_{(1+\mathtt{m}(\beta,0))/2} + \delta_{(1-\mathtt{m}(\beta,0))/2})\right)\right)$$
$$= D\left(\frac{1 - \mathtt{m}^2(\beta, 0)}{4}\left(1 + \frac{\alpha\beta(1 - \mathtt{m}^2(\beta, 0))}{1 - \beta(1 - \mathtt{m}^2(\beta, 0))}\right), \frac{1 - \mathtt{m}^2(\beta, 0)}{4}\right). \tag{1.14}$$

*If $h \neq 0$ and $\beta > 0$, then*

$$\lim_{N \to \infty} d_{TV}\left(\mu_N \circ \mathcal{P}_{k(N)}^{-1}, \mathrm{Bin}(k(N), (1 + \mathtt{m}(\beta, h))/2)\right)$$
$$= D\left(\frac{1 - \mathtt{m}^2(\beta, h)}{4}\left(1 + \frac{\alpha\beta(1 - \mathtt{m}^2(\beta, h))}{1 - \beta(1 - \mathtt{m}^2(\beta, h))}\right), \frac{1 - \mathtt{m}^2(\beta, h)}{4}\right). \tag{1.15}$$

*In particular, if $k(N) = o(N)$, then $\alpha = 0$ and the right-hand sides in (1.13), (1.14) and (1.15) vanish, meaning that the 'propagation of chaos' persists for any $k(N)$ of a sublinear growth.*

Let us now informally discuss the case when $\alpha > 0$. For simplicity, we consider (1.13). The limit on the right-hand side is non-zero, which suggests that there is a residual dependence between the $k(N)$ spins under the Gibbs measure. The reason for the non-zero limit is the fact that the distribution of $\mathcal{P}_{k(N)}$ and the corresponding binomial distribution satisfy central limit theorems with different variances, the variance of $\mathcal{P}_{k(N)}$ being strictly larger, which comes from the fact that the spins are positively correlated under the Gibbs measure. The distance between these normal distributions appears on the right-hand side of (1.13). In Theorem 3.5, we shall determine a *mixed* binomial distribution which approximates the distribution of $\mathcal{P}_{k(N)}$ under $\mu_N$. In some sense, this describes the residual dependence between the spins under the Gibbs measure.

**Remark 1.2.** The exchangeability of the measure $\mu_N$ has been used to investigate the Curie-Weiss model for example, in [17, Section 5.2] and [2]. In particular, an explicit representation of $\mu_N$ as a mixture of Bernoulli measures (valid for each fixed $N$) can be found in [17, Theorem 5.6]. A general propagation of chaos principle stating that the distribution of $k$ entries in a finite exchangeable vector of length $n$ can be approximated by a mixture of i.i.d. distributions can be found in [7].

The paper is organized as follows. Our proof relies on local limit theorems for the magnetization $m_N$ and also for the total number of positive spins $\mathcal{P}_N$ under $\mu_N$. In some regimes those are known. We collect the corresponding results in Section 2 below. The proofs of these local limit theorems, which we have not been able to locate in the literature, are given in Section 4. The proof of Theorem 1.1 is given in Section 3, including the statement of residual dependence. Two auxiliary technical results related to calculations of the total variation distance are presented in Section 5.

## 2   Local limit theorem for the magnetization

Denote by $\mathcal{N}(\mathtt{m}, \mathtt{v}^2)$ a Gaussian distribution with mean $\mathtt{m}$ and variance $\mathtt{v}^2$, so

$$\mathcal{N}(\mathtt{m}, \mathtt{v}^2)(A) = \int_A \varphi(t; \mathtt{m}, \mathtt{v}^2)\mathrm{d}t, \quad A \in \mathcal{B}(\mathbb{R}).$$

Put $\delta_N := (1 - (-1)^N)/2$. This correction term appears below in the local limit theorems for $m_N$, since $Nm_N$ always has the same parity as $N$.

**Proposition 2.1.** *Assume that $h \neq 0$ or $0 < \beta < 1$. Then*

$$\mu_N \left( \sqrt{N}(m_N - \mathtt{m}(\beta, h)) \in \cdot \right) \implies \mathcal{N}\left(0, \mathtt{v}_{\beta,h}^2\right), \quad N \to \infty,$$

*and the following local limit theorem holds true:*

$$\lim_{N \to \infty} \sqrt{N} \sup_{\ell \in \mathbb{Z}} \left| \mu_N \left( \frac{Nm_N + \delta_N}{2} = \ell \right) - \varphi\left( \ell; \frac{N\mathtt{m}(\beta, h)}{2}, \frac{N\mathtt{v}_{\beta,h}^2}{4} \right) \right| = 0. \qquad (2.1)$$

**Proposition 2.2.** *Assume that $h = 0$ and $\beta > 1$. Then*

$$\mu_N \left( \sqrt{N}(m_N - (+)\mathtt{m}(\beta, 0)) \in \cdot \mid m_N > (<)0 \right) \implies \mathcal{N}\left(0, \mathtt{v}_{\beta,0}^2\right), \quad N \to \infty,$$

*and the following local limit theorem holds true:*

$$\lim_{N \to \infty} \sqrt{N} \sup_{\ell \in \mathbb{Z}} \left| \mu_N \left( \frac{Nm_N + \delta_N}{2} = \ell \right) \right.$$
$$\left. - \frac{1}{2} \left( \varphi\left( \ell; \frac{N\mathtt{m}(\beta, 0)}{2}, \frac{N\mathtt{v}_{\beta,0}^2}{4} \right) + \varphi\left( \ell; -\frac{N\mathtt{m}(\beta, 0)}{2}, \frac{N\mathtt{v}_{\beta,0}^2}{4} \right) \right) \right| = 0. \qquad (2.2)$$

For some values of $(\beta, h)$ the above local limit theorems can be extracted from the vast literature on the Curie-Weiss model. For example, in the high-temperature regime $\beta \in (0, 1)$ and for every $h \in \mathbb{R}$, (2.1) has been proved in [19, Theorem 4.5 and Eq. (4.4)]. If $h > 0$ and $\beta > 0$, then (2.1) can be found in [1, Theorem 2.14 and Lemma 1.1]. The missing case $h < 0$ and $\beta > 0$ in Proposition 2.1 can be derived by the same methods. Finally, if $h = 0$, a local limit theorem for a multi-group Curie-Weiss model in the high-temperature regime $\beta \in (0, 1)$ has been derived in [13]. Quite (un-)expectedly, we have not been able to locate Proposition 2.2 in the literature, because of the non-standard approximation by a mixture of normal distributions. We shall give an elementary proof based on the Stirling approximation in Section 4.

We shall actually need local limit theorems for $\mathcal{P}_N$ rather than $m_N$. They follow immediately from Propositions 2.1 and 2.2 using the obvious relation $\mathcal{P}_N = \frac{N}{2}(1 + m_N)$ together with the bound

$$|\varphi(t_1, \mathtt{m}, \mathtt{v}^2) - \varphi(t_2, \mathtt{m}, \mathtt{v}^2)| \leq C \frac{|t_2 - t_2|}{\mathtt{v}^2}, \quad t_1, t_2 \in \mathbb{R},$$

for some absolute constant $C > 0$, which is a consequence of the mean value theorem for differentiable functions. The above bound allows us to neglect the correction term $\delta_N$ appearing in local limit theorems for $m_N$.

**Corollary 2.3.** *Assume that $h \neq 0$ or $0 < \beta < 1$. Then*

$$\mu_N \left( N^{-1/2}(\mathcal{P}_N - \tfrac{N}{2}(1 + \mathtt{m}(\beta, h))) \in \cdot \right) \implies \mathcal{N}\left(0, \tfrac{1}{4}\mathtt{v}_{\beta,h}^2\right), \quad N \to \infty,$$

*and the following local limit theorem holds true*

$$\lim_{N \to \infty} \sqrt{N} \sup_{\ell \in \mathbb{Z}} \left| \mu_N\left(\mathcal{P}_N = \ell\right) - \varphi(\ell; \tfrac{N}{2}(1 + \mathtt{m}(\beta, h)), \tfrac{N}{4}\mathtt{v}_{\beta,h}^2) \right| = 0.$$

**Corollary 2.4.** *Assume that $h = 0$ and $\beta > 1$. Then*

$$\mu_N \left( N^{-1/2}(\mathcal{P}_N - \tfrac{N}{2}(1 + (-)\mathtt{m}(\beta, 0))) \in \cdot \mid \mathcal{P}_N > (<)\tfrac{N}{2} \right) \implies \mathcal{N}\left(0, \tfrac{1}{4}\mathtt{v}_{\beta,0}^2\right), \quad N \to \infty,$$

*and the following local limit theorem holds true*

$$\lim_{N \to \infty} \sqrt{N} \sup_{\ell \in \mathbb{Z}} \Big| \mu_N\left(\mathcal{P}_N = \ell\right)$$
$$- \tfrac{1}{2} \left( \varphi\left(\ell; \tfrac{N}{2}(1 + \mathtt{m}(\beta, 0)), \tfrac{N}{4}\mathtt{v}_{\beta,0}^2\right) + \varphi\left(\ell; \tfrac{N}{2}(1 - \mathtt{m}(\beta, 0)), \tfrac{N}{4}\mathtt{v}_{\beta,0}^2\right) \right) \Big| = 0.$$

## 3  Proof of Theorem 1.1

We embark on a simple observation which is a consequence of exchangeability of the spins under the Gibbs measure $\mu_N$. Given that $\mathcal{P}_N = i \in \mathbb{N}_0$, the conditional distribution of $\mathcal{P}_k$ is hypergeometric with parameters $N$, $i$ and $k$ denoted hereafter $\mathrm{HyperG}(N, i, k)$. Recall that

$$\mathrm{HyperG}(n, i, k)(\{j\}) = \frac{\binom{i}{j}\binom{n-i}{k-j}}{\binom{n}{k}}, \quad 0 \leq j \leq \min(i, k), \quad i \leq n.$$

In other words, the distribution of $\mathcal{P}_{k(N)}$ can be represented as the following mixture of hypergeometric distributions:

$$\mu_N \circ \mathcal{P}_{k(N)}^{-1} = \mathrm{HyperG}(N, \mathcal{P}_N, k(N)). \tag{3.1}$$

The family of hypergeometric distributions possesses the following property which is of major importance for us.

**Lemma 3.1.** *For $n \in \mathbb{N}_0$, $k \leq n$ and $p \in [0, 1]$, it holds $\mathrm{HyperG}(n, \mathrm{Bin}(n, p), k) = \mathrm{Bin}(k, p)$.*

*Proof.* For $0 \leq j \leq k$,

$$\mathrm{HyperG}(n, \mathrm{Bin}(n, p), k)(\{j\})$$
$$= \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} \mathrm{HyperG}(n, i, k)(\{j\}) = \sum_{i=j}^{n} \binom{n}{i} p^i (1-p)^{n-i} \frac{\binom{i}{j}\binom{n-i}{k-j}}{\binom{n}{k}}$$
$$= \sum_{i=j}^{n} \binom{k}{j}\binom{n-k}{i-j} p^i (1-p)^{n-i} = \binom{k}{j} p^j (1-p)^{k-j} \sum_{i=j}^{n} \binom{n-k}{i-j} p^{i-j}(1-p)^{n-k-i+j}$$
$$= \binom{k}{j} p^j (1-p)^{k-j} \sum_{i=0}^{n-j} \binom{n-k}{i} p^i (1-p)^{n-k-i} = \binom{k}{j} p^j (1-p)^{k-j} = \mathrm{Bin}(k, p)(\{j\}).$$

Alternatively, we can argue probabilistically: If each of $n$ balls is colored black or white with probability $p$ and $1 - p$, respectively, and then a sample of $k$ balls is drawn at random from $n$ balls, then the number of black balls in the sample has binomial distribution with parameters $(k, p)$. $\square$

The subsequent proof of Theorem 1.1 proceeds according to the following scheme:

*Step 1.* We approximate $\mathcal{L}(\mathcal{P}_N)$ by an appropriate *mixed* binomial distribution $\mathrm{Bin}(N, \mathcal{L}(W))$, where $W$ is a random variable taking values in $[0,1]$ and $\mathcal{L}(X)$ denotes the distribution of a random variable $X$. The approximation is understood in the sense of the $d_{TV}$-distance which, as we shall show, converges to 0. To accomplish this step we employ the local limit theorems for $\mathcal{P}_N$ provided by Corollaries 2.3 and 2.4. It turns out that as the mixing distribution $W$ we can take a beta distribution (or a mixture of two beta-distributions) with properly adjusted parameters.

*Step 2.* Lemma 3.1 implies that the $\mu_N$-distribution of $\mathcal{P}_k$ is close (in a sense of the $d_{TV}$-distance) to the mixed binomial distribution $\mathrm{Bin}(k, \mathcal{L}(W))$. Formal verification of this employs the well-known characterization of the $d_{TV}$-distance

$$d_{TV}(\mathfrak{M}_1, \mathfrak{M}_2) = \inf \ \mathbb{P}\{X \neq Y\}, \tag{3.2}$$

where the infimum is taken over all pairs $(X, Y)$ of random variables such that $X$ is distributed according to $\mathfrak{M}_1$ and $Y$ is distributed according to $\mathfrak{M}_2$.

*Step 3.* We derive a local limit theorem for $\mathrm{Bin}(k, \mathcal{L}(W))$.

*Step 4.* We calculate the total variation distance between $\mathrm{Bin}(k, \mathcal{L}(W))$ and the three binomial distributions appearing in Theorem 1.1 by using local limit theorems for binomial distributions in conjunction with another well-known formula for $d_{TV}$: If measures $\mathfrak{M}_1$ and $\mathfrak{M}_2$ are supported on $\mathbb{Z}$, then

$$d_{TV}(\mathfrak{M}_1, \mathfrak{M}_2) = \frac{1}{2} \sum_{k \in \mathbb{Z}} |\mathfrak{M}_1(\{k\}) - \mathfrak{M}_2(\{k\})|, \tag{3.3}$$

see Propositions 5.1 and Proposition 5.2 below.

Our implementation of Steps 1-3 relies on the next proposition. For $\alpha, \beta > 0$, $\mathrm{Beta}(\alpha, \beta)$ denotes a beta distribution with the density

$$\mathrm{Beta}(\alpha, \beta)(\mathrm{d}x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}\mathrm{d}x}{B(\alpha, \beta)} \mathbb{1}_{\{x \in (0,1)\}},$$

where $B$ is the Euler beta-function. In what follows we assume that all auxiliary random variables are defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

**Proposition 3.2.** *Assume that $(\Theta_N)_{N \in \mathbb{N}}$ is a sequence of $\mathbb{N}_0$-valued random variables that satisfy a local limit theorem of the form: for a fixed integer $K \in \mathbb{N}$, a collection of positive weights $p_1, \ldots, p_K$ satisfying $\sum_{j=1}^{K} p_j = 1$, $a_j \in (0,1)$ and $\sigma_j^2 \in (a_j(1-a_j), \infty)$, $j = 1, \ldots, K$, it holds*

$$\lim_{N \to \infty} \sqrt{N} \sup_{\ell \in \mathbb{Z}} \left| \mathbb{P}(\Theta_N = \ell) - \sum_{j=1}^{K} p_j \varphi(\ell; Na_j, N\sigma_j^2) \right| = 0. \tag{3.4}$$

*Suppose that $k = k(N)$ is a sequence of positive integers such that (1.12) holds for some $\alpha \in [0,1]$ and put*

$$\gamma_{j,1} := \frac{a_j^2(1-a_j)}{\sigma_j^2 - a_j(1-a_j)}, \quad \gamma_{j,2} := \frac{a_j(1-a_j)^2}{\sigma_j^2 - a_j(1-a_j)}, \quad j = 1, \ldots, K. \tag{3.5}$$

*For $N \in \mathbb{N}$, let $X_N$ be a random variable with a mixed hypergeometric distribution $\mathrm{HyperG}(N, \mathcal{L}(\Theta_N), k(N))$ and $Y_{N,k(N)}$ be a random variable with the mixed binomial*

*distribution* $\mathrm{Bin}(k(N), \sum_{j=1}^{K} p_j \, \mathrm{Beta}(\gamma_{j,1} N, \gamma_{j,2} N))$, *that is,*

$$\mathbb{P}(Y_{N,k(N)} = \ell) = \int_0^1 \binom{k(N)}{\ell} x^\ell (1-x)^{k(N)-\ell} \left( \sum_{j=1}^{K} p_j \, \mathrm{Beta}(\gamma_{j,1} N, \gamma_{j,2} N)(\mathrm{d}x) \right),$$

*for $\ell = 0, \ldots, k(N)$. Then $\lim_{N\to\infty} d_{TV}(\mathcal{L}(X_N), \mathcal{L}(Y_{N,k(N)})) = 0$.*

**Remark 3.3.** We shall use this proposition with $\Theta_N = \mathcal{P}_N$ and $K = 1$, $p_1 = 1$ (in conjunction with Corollary 2.3) or $K = 2$, $p_1 = p_2 = 1/2$ (in conjunction with Corollary 2.4).

**Remark 3.4.** Let us provide an informal explanation of Proposition 3.2. Consider for simplicity the case $K = 1$, $p = 1$. Then, (3.4) states that $\Theta_N$ satisfies a local limit theorem with asymptotic centering $N a_1$. One is therefore tempted to approximate $\Theta_N$ by the binomial distribution $\mathrm{Bin}(N, a_1)$, however the variance of this distribution, which equals $N a_1 (1 - a_1)$, is strictly smaller than the asymptotic variance $N \sigma_1^2$ appearing in (3.4) due to the assumption $\sigma_1^2 > a_1(1 - a_1)$. Instead, we approximate $\Theta_N$ by a *mixed binomial* distribution that artificially blows up the variance until the variances match. The informal explanation of this is that both distributions satisfy a local limit theorem with the same centering and normalization. In the case $k(N) = N$, we have $\Theta_N \overset{d}{=} X_N$ and Proposition 3.2 states that the total variation distance between the distribution of $X_N$ and $\mathrm{Bin}(N, \mathrm{Beta}(\gamma_{1,1} N, \gamma_{1,2} N))$ converges to $0$. Moreover, in the formal limit $\sigma_1^2 \downarrow a_1(1 - a_1)$, we retrieve Lemma 3.1.

*Proof of Proposition 3.2.* We start by noting that, for fixed $j = 1, \ldots, K$, the pair $(\gamma_{j,1}, \gamma_{j,2})$ is a unique solution to the following system of equations

$$\begin{cases} N \frac{\gamma_{j,1}}{\gamma_{j,1}+\gamma_{j,2}} = N a_j, \\ N \left( \frac{\gamma_{j,1} \gamma_{j,2}}{(\gamma_{j,1}+\gamma_{j,2})^2} + \frac{\gamma_{j,1} \gamma_{j,2}}{(\gamma_{j,1}+\gamma_{j,2})^3} \right) = N \sigma_j^2. \end{cases} \tag{3.6}$$

On the left-hand side of the first equation we recognize the mean of $Y_{N,N}$, whereas the left-hand side of the second equation is equal to the variance of $Y_{N,N}$ up to a term $O(1)$. This suggests that the distribution of $\Theta_N$ is close to the distribution of $Y_{N,N}$. In fact, assume that we have proved

$$\lim_{N\to\infty} d_{TV}(\mathcal{L}(\Theta_N), \mathcal{L}(Y_{N,N})) = 0. \tag{3.7}$$

According to (3.2) there exists a sequence of pairs $(\Theta'_N, Y'_{N,N})$ such that

$$\lim_{N\to\infty} \mathbb{P}(\Theta'_N \neq Y'_{N,N}) = 0$$

and $\Theta'_N$ (respectively, $Y'_{N,N}$) has the same distribution as $\Theta_N$ (respectively, $Y_{N,N}$), for every $N \in \mathbb{N}$. Therefore,

$$\begin{aligned} & d_{TV}(\mathcal{L}(X_N), \mathrm{HyperG}(N, \mathcal{L}(Y_{N,N}), k(N))) \\ =\, & d_{TV}(\mathrm{HyperG}(N, \mathcal{L}(\Theta_N), k(N)), \mathrm{HyperG}(N, \mathcal{L}(Y_{N,N}), k(N))) \\ =\, & d_{TV}(\mathrm{HyperG}(N, \mathcal{L}(\Theta'_N), k(N)), \mathrm{HyperG}(N, \mathcal{L}(Y'_{N,N}), k(N))) \to 0. \end{aligned}$$

But this immediately yields the claim since $\mathrm{HyperG}(N, \mathcal{L}(Y_{N,N}), k(N))$ has the same distribution as $Y_{N,k(N)}$ by Lemma 3.1. Thus, it remains to prove (3.7).

To this end, we shall check that $Y_{N,N}$ satisfies exactly the same local limit theorem as $\Theta_N$, that is,

$$\lim_{N\to\infty} \sqrt{N} \sup_{\ell \in \mathbb{Z}} \left| \mathbb{P}(Y_{N,N} = \ell) - \sum_{j=1}^{K} p_j \varphi(\ell; N a_j, N \sigma_j^2) \right| = 0. \tag{3.8}$$

We shall actually prove a stronger result for later considerations, namely

$$\lim_{N\to\infty} \sup_{\ell\in\mathbb{Z}} \left| \sqrt{k(N)}\mathbb{P}(Y_{N,k(N)} = \ell) - \sum_{j=1}^{K} p_j \varphi(\ell; k(N)a_j, k(N)\sigma_{\alpha,j}^2) \right| = 0, \qquad (3.9)$$

where

$$\sigma_{\alpha,j}^2 := \frac{\gamma_{j,1}\gamma_{j,2}}{(\gamma_{j,1} + \gamma_{j,2})^2} + \frac{\alpha\gamma_{j,1}\gamma_{j,2}}{(\gamma_{j,1} + \gamma_{j,2})^3}. \qquad (3.10)$$

Relation (3.8) follows from (3.9) upon setting $k(N) = N$. The intuitive fact that (3.8) and (3.4) imply (3.7) follows from Proposition 5.2 below.

By the very definition of $Y_{N,k(N)}$ it suffices to prove (3.9) only in the case $K = 1$, $p_1 = 1$ and $j = 1$. We find it instructive to first prove a central limit theorem for $Y_{N,k(N)}$ of the form

$$\frac{Y_{N,k(N)} - a_1 k(N)}{\sqrt{k(N)}} \implies \mathcal{N}(0, \sigma_{\alpha,1}^2), \qquad (3.11)$$

which explains the formula for the variance $\sigma_{\alpha,1}^2$ in (3.9). Let $B_N$ be a random variable with the beta-distribution $\mathrm{Beta}(\gamma_{1,1}N, \gamma_{1,2}N)$ and $C_N^{(1)}$, $C_N^{(2)}$ be independent random variables with gamma-distributions with parameters $(\gamma_{1,1}N, 1)$ and $(\gamma_{1,2}N, 1)$, respectively. Using a representation

$$B_N \stackrel{d}{=} \frac{C_N^{(1)}}{C_N^{(1)} + C_N^{(2)}},$$

for the beta-distribution, see [15, Theorem 3], and the central limit theorem for $C_N^{(1)}$ and $C_N^{(2)}$, it is easy to check that

$$\sqrt{N}\left(B_N - \frac{\gamma_{1,1}}{\gamma_{1,1} + \gamma_{1,2}}\right) \implies \mathcal{N}(0, \mathbf{s}^2), \quad N \to \infty, \qquad (3.12)$$

where $\mathbf{s}^2 := \frac{\gamma_{1,1}\gamma_{1,2}}{(\gamma_{1,1} + \gamma_{1,2})^3}$. Moreover, let $(U_k)_{k\in\mathbb{N}}$ be a sequence of independent copies of a random variable with the uniform distribution on $[0,1]$. It is well-known that

$$\left(\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N} \mathbb{1}_{\{U_i \le t\}} - t\right)\right)_{t\in[0,1]} \implies (W(t))_{t\in[0,1]}, \quad N \to \infty, \qquad (3.13)$$

in the Skorokhod $J_1$-topology on $D[0,1]$, where $(W(t))_{t\in[0,1]}$ is a standard Brownian bridge. By independence and our assumption $k(N)/N \to \alpha$, we have a joint convergence

$$\left(\left(\sqrt{k(N)}\left(\frac{1}{k(N)}\sum_{i=1}^{k(N)} \mathbb{1}_{\{U_i \le t\}} - t\right)\right)_{t\in[0,1]}, B_N, \sqrt{k(N)}\left(B_N - a_1\right)\right)$$
$$\implies ((W(t))_{t\in[0,1]}, a_1, \mathcal{N}(0, \alpha\mathbf{s}^2)), \quad n \to \infty, \quad (3.14)$$

on $D[0,1] \times \mathbb{R} \times \mathbb{R}$ endowed with the product topology. Applying the map

$$D[0,1] \times \mathbb{R} \times \mathbb{R} \ni (f(\cdot), x, y) \longmapsto f(x) + y \in \mathbb{R},$$

which is a.s. continuous at the limiting point in (3.14) we arrive at

$$\frac{Y_{N,k(N)} - a_1 k(N)}{\sqrt{k(N)}} \implies W(a_1) + \mathcal{N}(0, \alpha\mathbf{s}^2), \quad N \to \infty,$$

which is equivalent to (3.11), since $\mathrm{Var}\,(W(a_1)) = a_1(1 - a_1)$. Fix $\varepsilon \in (0, \min(a_1, 1 - a_1))$ and note that

$$
\begin{aligned}
\mathbb{P}(|Y_{N,k(N)} - a_1 k(N)| \geq \varepsilon k(N)) &= \mathbb{P}(|Y_{N,k(N)} - \mathbb{E}Y_{N,k(N)}| \geq \varepsilon k(N)) \\
&\leq \frac{\mathrm{Var}(Y_{N,k(N)})}{\varepsilon^2 (k(N))^2} \\
&= \frac{k(N)\mathbb{E}B_N(1 - B_N) + N^{-1}k^2(N)\mathrm{Var}(\sqrt{N}B_N)}{\varepsilon^2 (k(N))^2} \\
&= O((k(N)^{-1}),
\end{aligned}
$$

where the penultimate equality follows from the law of total variance, and the last equality is a consequence of $\lim_{N \to \infty} \mathrm{Var}(\sqrt{N}B_N) = \mathbf{s}^2$; see (3.12). The above estimate in conjunction with a standard tail estimate for the normal law implies that (3.9) is equivalent to

$$
\sqrt{k(N)} \sup_{\ell:\, |\ell - a_1 k(N)| \leq \varepsilon k(N)} \left| \mathbb{P}(Y_{N,k(N)} = \ell) - \varphi(\ell; a_1 k(N), \sigma_{\alpha,1}^2 k(N)) \right| \longrightarrow 0, \quad N \to \infty.
\tag{3.15}
$$

The latter can be deduced from the following explicit formula

$$
\mathbb{P}(Y_{N,k(N)} = \ell) = \binom{k(N)}{\ell} \frac{B(\ell + \gamma_{1,1}N, k(N) - \ell + \gamma_{1,2}N)}{B(\gamma_{1,1}N, \gamma_{1,2}N)}
\tag{3.16}
$$

$$
= \frac{k(N) + 1}{B(\ell + 1, k(N) - \ell + 1)} \frac{B(\ell + \gamma_{1,1}N, k(N) - \ell + \gamma_{1,2}N)}{B(\gamma_{1,1}N, \gamma_{1,2}N)}, \quad \ell = 0, \ldots, k(N),
\tag{3.17}
$$

together with the asymptotic relation

$$
B(k(N)x, k(N)y) = (1 + o(1)) \sqrt{\frac{2\pi}{k(N)}} \sqrt{\frac{x + y}{xy}} \left( \frac{x^x y^y}{(x + y)^{x+y}} \right)^{k(N)}, \quad N \to \infty, \tag{3.18}
$$

for the beta-function, which is uniform in $x, y \in [\delta, \delta^{-1}]$, for every fixed $\delta \in (0, 1)$. Our choice of $\varepsilon$ ensures that all the arguments of the beta-functions in (3.16) lie in $[\delta k(N), \delta^{-1}k(N)]$ for a sufficiently small $\delta > 0$. The uniform asymptotic relation (3.18) is a consequence of Stirling's formula with a uniform estimate of the remainder; see, for example, Eq. (5.11.10) and (5.11.11) in [18]. The proof of Proposition 3.2 is complete. $\square$

By combining (3.1) and Corollary 2.3 with Proposition 3.2 applied for $\Theta_N = \mathcal{P}_N$, $K = 1$, $a_1 = (1 + \mathbf{m}(\beta, h))/2$ and $\sigma_1^2 := \mathbf{v}_{\beta,h}^2/4$, we obtain the next theorem. Note that $4a_1(1 - a_1) = 1 - \mathbf{m}_{\beta,h}^2 < \mathbf{v}_{\beta,h}^2 = 4\sigma_1^2$ in view of (1.5). Also some simple algebra yields

$$
\gamma_1(\beta, h) := \gamma_{1,1} = \frac{1 - \beta(1 - \mathbf{m}^2(\beta, h))}{2\beta(1 - \mathbf{m}(\beta, h))}, \quad \gamma_2(\beta, h) := \gamma_{1,2} = \frac{1 - \beta(1 - \mathbf{m}^2(\beta, h))}{2\beta(1 + \mathbf{m}(\beta, h))}.
\tag{3.19}
$$

**Theorem 3.5.** *Assume that $h \neq 0$ or $0 < \beta < 1$. Suppose that $k = k(N)$ satisfies (1.12). Then*

$$
\lim_{N \to \infty} d_{TV} \left( \mu_N \circ \mathcal{P}_{k(N)}^{-1}, \mathrm{Bin}\left(k(N), \mathrm{Beta}\left(\gamma_1(\beta, h)N, \gamma_2(\beta, h)N\right)\right) \right) = 0,
$$

*where $\gamma_1(\beta, h)$ and $\gamma_2(\beta, h)$ are given in (3.19).*

**Remark 3.6.** The theorem above describes the residual dependence in the propagation of chaos phenomenon via a beta-binomial distribution, which in turn has a simple interpretation as follows. Consider a Pólya urn which initially contains $\gamma_1(\beta, h)N$ positive spins (white balls) and $\gamma_2(\beta, h)N$ negative spins (black balls). Balls are drawn one at a

time and immediately returned to the urn together with a new ball of the same color. The number of white balls drawn after $k(N)$ trials has the beta-binomial distribution $\mathrm{Bin}\,(k(N), \mathrm{Beta}\,(\gamma_1(\beta, h)N, \gamma_2(\beta, h)N))$. Thus, Theorem 3.5 tells us that the number of positive spins under the Gibbs measure is close in distribution to the number of white balls drawn from the above Pólya urn after $k(N)$ trials.

Similarly, Corollary 2.4 and Proposition 3.2 applied with $K = 2$, $p_1 = p_2 = 1/2$ yield the next result.

**Theorem 3.7.** *Assume that $h = 0$ and $\beta > 1$. Suppose that $k = k(N)$ satisfies (1.12). Then*

$$
\lim_{N \to \infty} d_{TV} \Bigg( \mu_N \circ \mathcal{P}_{k(N)}^{-1},
$$
$$
\mathrm{Bin}\left(k(N), \frac{1}{2}\left(\mathrm{Beta}(\gamma_1(\beta, h)N, \gamma_2(\beta, h)N) + \mathrm{Beta}(\gamma_2(\beta, h)N, \gamma_1(\beta, h)N)\right)\right) \Bigg) = 0.
$$

*Proof of Theorem 1.1.* We shall treat three cases separately.

*Case $h = 0$ and $\beta \in (0, 1)$.* In this case $\mathtt{m}(\beta, h) = \mathtt{m}(\beta, 0) = 0$ and $\gamma_1(\beta, 0) = \gamma_2(\beta, 0) = (1 - \beta)/(2\beta)$. In view of Theorem 3.5 it suffices to check that

$$
\lim_{N \to \infty} d_{TV}\left(\mathrm{Bin}\,(k(N), \mathrm{Beta}\,(\gamma_1(\beta, h)N, \gamma_2(\beta, h)N)), \mathrm{Bin}(k(N), 1/2)\right)
$$
$$
= D\left(\frac{1 - \beta + \alpha\beta}{4(1 - \beta)}, \frac{1}{4}\right). \tag{3.20}
$$

From (3.9) we infer

$$
\lim_{N \to \infty} \sqrt{k(N)} \sup_{\ell \in \mathbb{N}} \Bigg| \mathrm{Bin}\,(k(N), \mathrm{Beta}\,(\gamma_1(\beta, h)N, \gamma_2(\beta, h)N))\,(\{\ell\})
$$
$$
- \varphi\left(\ell; k(N)/2, \frac{1 - \beta + \alpha\beta}{4(1 - \beta)} k(N)\right) \Bigg| = 0,
$$

where the formula for the limiting variance follows from (3.10). By the classical de Moivre-Laplace local limit theorem

$$
\lim_{N \to \infty} \sqrt{k(N)} \sup_{\ell \in \mathbb{N}} |\mathrm{Bin}\,(k(N), 1/2)\,(\{\ell\}) - \varphi\,(\ell; k(N)/2, k(N)/4)| = 0.
$$

Equality (3.20) now follows from Proposition 5.1 below.

*Case $h = 0$ and $\beta > 1$.* By the same reasoning as in the previous case, but applying Theorem 3.7 instead of Theorem 3.5, we see that it suffices to prove

$$
\lim_{N \to \infty} d_{TV}\Bigg( \mathrm{Bin}\left(k(N), \frac{1}{2}\left(\mathrm{Beta}(\gamma_1(\beta, 0)N, \gamma_2(\beta, 0)N) + \mathrm{Beta}(\gamma_2(\beta, 0)N, \gamma_1(\beta, 0)N)\right)\right),
$$
$$
\mathrm{Bin}\left(k(N), \frac{1}{2}(\delta_{(1+\mathtt{m}(\beta,0))/2} + \delta_{(1-\mathtt{m}(\beta,0))/2})\right) \Bigg)
$$
$$
= D\left(\frac{1 - \mathtt{m}^2(\beta, 0)}{4}\left(1 + \frac{\alpha\beta(1 - \mathtt{m}^2(\beta, 0))}{1 - \beta(1 - \mathtt{m}^2(\beta, 0))}\right), \frac{1 - \mathtt{m}^2(\beta, 0)}{4}\right). \tag{3.21}
$$

This again follows from (3.9) but now with $K = 2$, $p_1 = p_2 = 1/2$ and the de Moivre-Laplace local limit theorem in conjunction with Proposition 5.2 below. Put

$$
\mathfrak{s}^2(\beta, h) := \frac{1 - \mathtt{m}^2(\beta, h)}{4}\left(1 + \frac{\alpha\beta(1 - \mathtt{m}^2(\beta, h))}{1 - \beta(1 - \mathtt{m}^2(\beta, h))}\right),
$$

and note that in (3.9) the variances coincide:

$$\sigma_{\alpha,1}^2 = \sigma_{\alpha,2}^2 = \mathfrak{s}^2(\beta,0).$$

The second equality here follows upon substituting (3.19) into (3.10) and using that $\gamma_{2,1} = \gamma_{1,2} = \gamma_1(\beta,0)$, $\gamma_{1,1} = \gamma_{2,2} = \gamma_2(\beta,0)$ in this case.

*Case $h \neq 0$ and $\beta > 0$.* In this case we again apply Theorem 3.5. From (3.9) it follows that

$$\lim_{N \to \infty} \sqrt{k(N)} \sup_{\ell \in \mathbb{N}} \left| \text{Bin}\left(k(N), \text{Beta}\left(\gamma_1(\beta,h)N, \gamma_2(\beta,h)N\right)\right)(\{\ell\}) \right.$$
$$\left. - \varphi\left(\ell; \frac{1 + \mathfrak{m}(\beta,h)}{2}k(N), \mathfrak{s}^2(\beta,h)k(N)\right) \right| = 0.$$

Combining this with the de Moivre-Laplace local limit theorem and using Proposition 5.1 we arrive at (1.15). $\qquad\square$

**Remark 3.8.** Here we give a comparison of our findings with some known results on the uniform distribution on a high-dimensional sphere, which bears some similarities with the Curie-Weiss model in the context of the propagation of chaos phenomenon. For every $N \in \mathbb{N}$, consider a random vector $\xi_N = (\xi_{1;N}, \ldots, \xi_{N;N})$ which is uniformly distributed on the sphere $\sqrt{N}\mathbb{S}^{N-1}$ of radius $\sqrt{N}$ in $\mathbb{R}^N$. It is a classical result of Maxwell-Poincaré-Borel that, for every $k \in \mathbb{N}$, the distribution of any $k$ components of $\xi_N$ converges weakly to a $k$-dimensional standard normal distribution, as $N \to \infty$. Moreover, if $k = k(N)$ depends on $N$ such that $k(N)/N \to 0$, then it has been shown in [8, Section 2] that the total variation distance between the distribution of $(\xi_{1;n}, \ldots, \xi_{k;N})$ and the standard normal distribution on $\mathbb{R}^k$ converges to 0. On the other hand, for $k(N)/N \to \alpha$ with $\alpha \neq 0$, the total variation distance between these distributions converges to a non-zero limit which has been identified in [9, Theorem 1.6 (b)]; see also Eq. (2.12) on p. 403 in [8]. All these results are similar to what we know about the Curie-Weiss model. There is, however, one important difference: An approximation by a variance mixture of normal distributions is not possible if $\alpha > 0$ in the setting of $\xi_N$. A much more general result has been shown in [9, Theorem 2.3 (b)]. Let us give a short informal argument. Let $(\zeta_1, \ldots, \zeta_k)$ be a random vector with the standard normal distribution on $\mathbb{R}^k$ and let $R > 0$ be a mixing variable independent of the $\zeta_i$'s. We ask whether it is possible to approximate $(\xi_{1;N}, \ldots, \xi_{k;N})$ by $R \cdot (\zeta_1, \ldots, \zeta_k)$ in the total variation distance. By rotation invariance, it suffices to consider the distance between the squared radial parts; see Eq. (2.4) on p. 402 in [8]. Since $\frac{1}{N}(\xi_{1;N}^2 + \cdots + \xi_{k;N}^2)$ is beta-distributed with parameters $(k/2, (N-k)/2)$, it follows that

$$\mathbb{E}(\xi_{1;N}^2 + \cdots + \xi_{k;N}^2) = k, \qquad \text{Var}(\xi_{1;N}^2 + \cdots + \xi_{k;N}^2) = \frac{k(N-k)/4}{(N/2)^2(N/2+1)}N^2 \sim 2\alpha(1-\alpha)N.$$

On the other hand, the squared radial part of $R(\zeta_1, \ldots, \zeta_k)$ is $R^2\chi_k^2$, where $\chi_k^2 := \zeta_1^2 + \ldots + \zeta_k^2$ has a chi-square distribution with $k$ degrees of freedom, and we have

$$\mathbb{E}(R^2(\zeta_1^2 + \cdots + \zeta_k^2)) = \mathbb{E}(R^2)k, \qquad \text{Var}(R^2(\zeta_1^2 + \cdots + \zeta_k^2)) = \mathbb{E}(R^4)(2k + k^2) - (\mathbb{E}(R^2)k)^2.$$

where we used that $\mathbb{E}(\chi_k^4) = 2k + k^2$. If we want to match expectations, we need $\mathbb{E}(R^2) = 1$, but then $\mathbb{E}(R^4)(2k + k^2) - (\mathbb{E}(R^2)k)^2 \geq 2k + k^2 - k^2 = 2k \sim 2\alpha N$, and we cannot match the variances since $2\alpha > 2\alpha(1-\alpha)$. This is in sharp contrast to the case of the Curie-Weiss model, for which Theorem 3.5 shows that the larger variance of $\mathcal{P}_{k(N)}$ can be artificially matched by a mixed binomial distribution.

## 4 Proof of Proposition 2.2

Recall that we work under the assumptions $h = 0$ and $\beta > 1$, which imply $\mathtt{m} := \mathtt{m}(\beta, 0) \in (0, 1)$. For simplicity we also assume throughout this proof that $N = 2n$ is even. The case of odd $N$ can be treated similarly. For every $-n \le \ell \le n$, we have

$$\mu_N(\{\sigma : Nm_N(\sigma) = 2\ell\}) = \frac{|\{\sigma : Nm_N(\sigma) = 2\ell\}|}{Z_N(\beta)} \exp\left(\frac{\beta}{2N}(2\ell)^2\right) \tag{4.1}$$

$$= \frac{1}{Z_{2n}(\beta)}\binom{2n}{n+\ell}\exp(\beta\ell^2/n). \tag{4.2}$$

For further use we record the asymptotic formula for the partition function: if $\beta > 1$, then

$$Z_N(\beta) = (1 + o(1))\frac{2\mathtt{v}_{\beta,0}}{\sqrt{1 - \mathtt{m}^2}}2^N e^{-N(\mathcal{I}(\mathtt{m}) - \beta\mathtt{m}^2/2)}, \quad N \to \infty, \tag{4.3}$$

where $\mathcal{I}(x) := \frac{1}{2}((1+x)\log(1+x) + (1-x)\log(1-x))$ for $x \in [-1, 1]$, which can be derived from Théorème B in [3] by specializing to the Curie-Weiss model or from (3.19) in [16] by choosing $p = 1$; see also [20, Theorem 1.3]. For every fixed $n \in \mathbb{N}$, the function

$$\{-n, \ldots, n\} \ni \ell \longmapsto \binom{2n}{n+\ell}\exp(\beta\ell^2/n) \tag{4.4}$$

is even and attains two local maxima at $\ell \approx n\mathtt{m}$ and $\ell \approx -n\mathtt{m}$. This can be checked by calculating the ratio of its two consecutive values. Using the aforementioned symmetry and the standard tail estimate for the normal density, we see that (2.2) is equivalent to

$$\lim_{n\to\infty}\sqrt{N}\sup_{\ell\in\mathbb{Z}_{\ge 0}}\left|\mu_N(\{\sigma : Nm_N(\sigma) = 2\ell\}) - \frac{1}{2}\varphi\left(\ell; \frac{N\mathtt{m}}{2}, \frac{N\mathtt{v}_{\beta,0}^2}{4}\right)\right| = 0,$$

or, by (4.1), to

$$\lim_{n\to\infty}\sqrt{n}\sup_{\ell\in\mathbb{Z}_{\ge 0}}\left|\frac{1}{Z_{2n}(\beta)}\binom{2n}{n+\ell}\exp(\beta\ell^2/n) - \frac{1}{2}\varphi\left(\ell; n\mathtt{m}, \frac{n\mathtt{v}_{\beta,0}^2}{2}\right)\right| = 0. \tag{4.5}$$

Put $\mathcal{C}_n := \{j \in \mathbb{Z}_{\ge 0} : j \in [\mathtt{m}n - n^{7/12}, \mathtt{m}n + n^{7/12}]\}$. By Stirling's formula, uniformly in $\ell \in \mathcal{C}_n$, it holds

$$2^{-2n}\binom{2n}{n+\ell} = (1 + o(1))\sqrt{\frac{1}{\pi n}}\frac{1}{\sqrt{1 - \mathtt{m}^2}}e^{-2n\mathcal{I}(\ell/n)}.$$

Combining this with (4.3) we deduce that, uniformly in $\ell \in \mathcal{C}_n$,

$$\frac{1}{Z_{2n}(\beta)}\binom{2n}{n+\ell}\exp(\beta\ell^2/n) = \frac{1 + o(1)}{2\mathtt{v}_{\beta,0}\sqrt{\pi n}}\exp\left(-2n\left(\mathcal{I}(\ell/n) - \mathcal{I}(\mathtt{m}) + \frac{\beta\mathtt{m}^2}{2} - \frac{\beta\ell^2}{2n^2}\right)\right).$$

By the definition of $\mathtt{m}$ the first derivative of $t \mapsto \mathcal{I}(t) - \beta t^2/2$ vanishes at $t = \mathtt{m}$ and the second derivative at $t = \mathtt{m}$ is equal to $\mathtt{v}_{\beta,0}^{-2}$. Thus, plugging the Taylor expansion

$$\mathcal{I}(\ell/n) = \mathcal{I}(\mathtt{m}) + \mathcal{I}'(\mathtt{m})\frac{\ell - \mathtt{m}n}{n} + \frac{\mathcal{I}''(\mathtt{m})}{2}\left(\frac{\ell - \mathtt{m}n}{n}\right)^2 + O\left(\left(\frac{\ell - \mathtt{m}n}{n}\right)^3\right),$$

we obtain

$$\frac{1}{Z_{2n}(\beta)}\binom{2n}{n+\ell}\exp(\beta\ell^2/n) = \frac{(1 + o(1))}{2\mathtt{v}_{\beta,0}\sqrt{\pi n}}\exp\left(-\frac{(\ell - \mathtt{m}n)^2}{\mathtt{v}_{\beta,0}^2 n}\right) = (1 + o(1))\frac{1}{2}\varphi\left(\ell; n\mathtt{m}, \frac{n\mathtt{v}_{\beta,0}^2}{2}\right),$$

which is again uniform over $\ell \in \mathcal{C}_n$. This yields that (4.5) holds with $\sup_{\ell \in \mathbb{Z}_{\geq 0}}$ replaced by $\sup_{\ell \in \mathbb{Z}_{\geq 0}, \ell \in \mathcal{C}_n}$.

To check that the indices outside $\mathcal{C}_n$ are negligible we use the same reasoning as in [16]; see pp. 548-549 therein. According to the estimate (3.8) in [16]

$$2^{-2n} \binom{2n}{2n+\ell} \leq \exp(-2n\mathcal{I}(\ell/n)), \quad -n \leq \ell \leq n.$$

Therefore, for some absolute constant $C_1 > 0$ all $n \in \mathbb{N}$ and $|\ell| \leq n$,

$$\frac{1}{Z_{2n}(\beta)} \binom{2n}{n+\ell} \exp(\beta\ell^2/n) \leq C_1 \exp\left(-2n\left(\mathcal{I}(\ell/n) - \mathcal{I}(\mathtt{m}) + \frac{\beta\mathtt{m}^2}{2} - \frac{\beta\ell^2}{2n^2}\right)\right).$$

The function $t \mapsto \mathcal{I}(t) - \beta t^2/2$ attains a local minimum on the positive half-line at $\mathtt{m}$. Therefore, for $|\ell - \mathtt{m}n| \geq n^{7/12}$, $|\ell| \leq n$, and some $C_2 > 0$

$$\mathcal{I}(\ell/n) - \mathcal{I}(\mathtt{m}) + \frac{\beta\mathtt{m}^2}{2} - \frac{\beta\ell^2}{2n^2} \geq C_2 \left(\frac{(\ell - \mathtt{m}n)^2}{n^2}\right) \geq C_2 n^{-5/6}.$$

Thus, for $|\ell - \mathtt{m}n| \geq n^{7/12}$ and $|\ell| \leq n$,

$$\frac{1}{Z_{2n}(\beta)} \binom{2n}{n+\ell} \exp(\beta\ell^2/n) \leq C_1 \mathrm{e}^{-2C_2 n^{1/6}}.$$

Combing this with a standard tail estimate for the normal distribution shows that (4.5) holds with $\sup_{\ell \in \mathbb{Z}_{\geq 0}}$ replaced by $\sup_{\ell \in \mathbb{Z}_{\geq 0}, \ell \notin \mathcal{C}_n}$. The proof of Proposition 2.2 is complete.

## 5 Auxiliary results

**Proposition 5.1.** *Assume that two sequences of integer-valued random variables* $(Q_N^{(1)})_{N \in \mathbb{N}}$ *and* $(Q_N^{(2)})_{N \in \mathbb{N}}$ *satisfy the local limit theorems*

$$\lim_{N \to \infty} \sqrt{N} \sup_{\ell \in \mathbb{Z}} \left|\mathbb{P}(Q_N^{(i)} = \ell) - \varphi(\ell; N\mathtt{m}, N\mathtt{v}_i^2)\right| = 0, \quad i = 1, 2, \tag{5.1}$$

*with the same mean* $\mathtt{m}$ *and arbitrary variances* $\mathtt{v}_1^2$ *and* $\mathtt{v}_2^2$*. Then*

$$\lim_{N \to \infty} d_{TV}(\mathcal{L}(Q_N^{(1)}), \mathcal{L}(Q_N^{(2)})) = D(\mathtt{v}_1^2, \mathtt{v}_2^2),$$

*where the right-hand side was defined in (1.11). In particular, if* $\mathtt{v}_1^2 = \mathtt{v}_2^2$*, then*

$$\lim_{N \to \infty} d_{TV}(\mathcal{L}(Q_N^{(1)}), \mathcal{L}(Q_N^{(2)})) = 0.$$

*Proof.* Without loss of generality assume that $\mathtt{m} = 0$. Fix $\varepsilon > 0$ and $M > 0$. Then, there exists $N_0(\varepsilon)$ such that

$$\left|\mathbb{P}(Q_N^{(i)} = \ell) - \varphi(\ell; 0, N\mathtt{v}_i^2)\right| \leq \frac{\varepsilon}{\sqrt{N}}, \quad i = 1, 2, \quad N \geq N_0(\varepsilon), \quad -M\sqrt{N} \leq \ell \leq M\sqrt{N}.$$

Using (3.3) we infer

$$\left|2d_{TV}(\mathcal{L}(Q_N^{(1)}), \mathcal{L}(Q_N^{(2)})) - \sum_{\ell: |\ell| \leq M\sqrt{N}} \left|\varphi(\ell; 0, N\mathtt{v}_1^2) - \varphi(\ell; 0, N\mathtt{v}_2^2)\right|\right|$$

$$\leq \mathbb{P}(|Q_N^{(1)}| > M\sqrt{N}) + \mathbb{P}(|Q_N^{(2)}| > M\sqrt{N}) + 4M\varepsilon$$

$$+ \sum_{\ell: |\ell| > M\sqrt{N}} \varphi(\ell; 0, N\mathtt{v}_1^2) + \sum_{\ell: |\ell| > M\sqrt{N}} \varphi(\ell; 0, N\mathtt{v}_2^2).$$

Note that

$$\lim_{N \to \infty} \sum_{\ell:\, |\ell| \le M\sqrt{N}} \left| \varphi(\ell; 0, N\mathrm{v}_1^2) - \varphi(\ell; 0, N\mathrm{v}_2^2) \right| \tag{5.2}$$

$$= \lim_{N \to \infty} \frac{1}{\sqrt{N}} \sum_{\ell:\, |\ell| \le M\sqrt{N}} \left| \varphi(N^{-1/2}\ell; 0, \mathrm{v}_1^2) - \varphi(N^{-1/2}\ell; 0, \mathrm{v}_2^2) \right|$$

$$= \int_{-M}^{M} \left| \varphi(t; 0, \mathrm{v}_1^2) - \varphi(t; 0, \mathrm{v}_2^2) \right| \mathrm{d}t, \tag{5.3}$$

since the sum on the left-hand side is a Riemann sum of a Riemann integrable function on $[-M, M]$. Thus,

$$\limsup_{N \to \infty} \left| 2d_{TV}(\mathcal{L}(Q_n^{(1)}), \mathcal{L}(Q_n^{(2)})) - \int_{-M}^{M} \left| \varphi(t; 0, \mathrm{v}_1^2) - \varphi(t; 0, \mathrm{v}_2^2) \right| \mathrm{d}t \right|$$

$$\le \int_{|t|>M} \varphi(t; 0, \mathrm{v}_1^2)\mathrm{d}t + \int_{|t|>M} \varphi(t; 0, \mathrm{v}_2^2)\mathrm{d}t + 4M\varepsilon$$

$$+ \limsup_{N \to \infty} N^{-1/2} \sum_{\ell:\, |\ell|>M\sqrt{N}} \varphi(N^{-1/2}\ell; 0, N\mathrm{v}_1^2)$$

$$+ \limsup_{N \to \infty} N^{-1/2} \sum_{\ell:\, |\ell|>M\sqrt{N}} \varphi(N^{-1/2}\ell; 0, N\mathrm{v}_2^2)$$

$$= 2\int_{|t|>M} \varphi(t; 0, \mathrm{v}_1^2)\mathrm{d}t + 2\int_{|t|>M} \varphi(t; 0, \mathrm{v}_2^2)\mathrm{d}t + 4M\varepsilon,$$

where for the last equality we used that $t \mapsto \varphi(t; 0, v_i^2)$, $i = 1, 2$, are directly Riemann integrable on $\mathbb{R}$. Sending $\varepsilon \to 0$ and then $M \to \infty$ completes the proof. $\qquad\square$

For the mixtures of normal densities we have the following generalization of Proposition 5.1.

**Proposition 5.2.** *If we replace* (5.1) *by*

$$\lim_{N \to \infty} \sqrt{N} \sup_{\ell \in \mathbb{Z}} \left| \mathbb{P}(Q_N^{(i)} = \ell) - \sum_{j=1}^{K} p_j \varphi(\ell; N\mathrm{m}_j, N\mathrm{v}_{i,j}^2) \right| = 0, \quad i = 1, 2, \tag{5.4}$$

*for some $K \in \mathbb{N}$, a collection of positive weights $p_1, \ldots, p_K$ satisfying $\sum_{j=1}^{K} p_j = 1$, $\mathrm{m}_j \in \mathbb{R}$ such that $\mathrm{m}_i \ne \mathrm{m}_j$, $i \ne j$, and $\mathrm{v}_{i,j}^2 > 0$, $j = 1, \ldots, K$, then*

$$\lim_{N \to \infty} d_{TV}(\mathcal{L}(Q_N^{(1)}), \mathcal{L}(Q_N^{(2)})) = \sum_{j=1}^{K} p_j D(\mathrm{v}_{1,j}^2, \mathrm{v}_{2,j}^2).$$

*Proof.* We shall sketch the proof omitting details which are similar to those used in the proof of Proposition 5.1. Fix large positive constants $M_1, \ldots, M_K$. Split the range of summation on the right-hand side of equality

$$d_{TV}(\mathcal{L}(Q_N^{(1)}), \mathcal{L}(Q_N^{(2)})) = \frac{1}{2} \sum_{\ell \in \mathbb{Z}} \left| \mathbb{P}(Q_N^{(1)} = \ell) - \mathbb{P}(Q_N^{(2)} = \ell) \right|$$

into $K + 1$ sets

$$A_j(N) := \{\ell \in \mathbb{Z} : |\ell - \mathrm{m}_j N| \le M_j \sqrt{N}\}, \quad j = 1, \ldots, K, \quad A_{K+1}(N) := \mathbb{Z} \setminus \left( \cup_{j=1}^{K} A_j(N) \right).$$

These sets are pairwise disjoint for all large enough $N$, since $\mathtt{m}_j$'s are pairwise different. In view of (5.4) and by the standard tail estimate for the normal law, for every $j = 1, \ldots, K$,

$$\frac{1}{2} \sum_{\ell \in A_j(N)} \left| \mathbb{P}(Q_N^{(1)} = \ell) - \mathbb{P}(Q_N^{(2)} = \ell) \right|$$

$$= \frac{p_j}{2} \sum_{\ell \in A_j(N)} \left| \varphi(\ell; N\mathtt{m}_j, N\mathtt{v}_{1,j}^2) - \varphi(\ell; N\mathtt{m}_j, N\mathtt{v}_{2,j}^2) \right| + r_j(N, M_j),$$

where the remainders satisfy $\lim_{M_j \to \infty} \limsup_{N \to \infty} |r_j(N, M_j)| = 0$, $j = 1, \ldots, K$. Furthermore,

$$\lim_{\min_j M_j \to \infty} \limsup_{N \to \infty} \sum_{\ell \in A_{K+1}(N)} \left| \mathbb{P}(Q_N^{(1)} = \ell) - \mathbb{P}(Q_N^{(2)} = \ell) \right|$$

$$\leq \lim_{\min_j M_j \to \infty} \limsup_{N \to \infty} \left( \mathbb{P}(Q_N^{(1)} \in A_{K+1}(N)) + \mathbb{P}(Q_N^{(2)} \in A_{K+1}(N)) \right) = 0.$$

It remains to note that (5.2) yields

$$\lim_{M_j \to \infty} \lim_{N \to \infty} \frac{1}{2} \sum_{\ell \in A_j(N)} \left| \varphi(\ell; N\mathtt{m}_j, N\mathtt{v}_{1,j}^2) - \varphi(\ell; N\mathtt{m}_j, N\mathtt{v}_{2,j}^2) \right| = D(\mathtt{v}_{1,j}^2, \mathtt{v}_{2,j}^2), \quad j = 1, \ldots, K.$$

$\square$

## References

[1] A. D. Barbour, A. Röllin, and N. Ross. Error bounds in local limit theorems using Stein's method. *Bernoulli*, 25(2):1076–1104, 2019. MR3920366

[2] Y. Barhoumi-Andréani, M. Butzek, and P. Eichelsbacher. A surrogate by exchangeability approach to the Curie-Weiss model. *arXiv preprint arXiv:2305.06872*, 2023.

[3] G. Ben Arous and M. Brunaud. Méthode de Laplace: étude variationnelle des fluctuations de diffusions de type "champ moyen". *Stoch. Stoch. Rep.*, 31(1-4):79–144, 1990. MR1080535

[4] G. Ben Arous and O. Zeitouni. Increasing propagation of chaos for mean field models. *Ann. Inst. H. Poincaré Probab. Statist.*, 35(1):85–102, 1999. MR1669916

[5] A. Bovier. *Statistical Mechanics of Disordered Systems – A Mathematical Perspective*. Cambridge Series in Statistical and Probabilistic Mathematics, 2006. MR2252929

[6] S. Chatterjee and Q.-M. Shao. Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie-Weiss model. *Ann. Appl. Probab.*, 21(2):464–483, 2011. MR2807964

[7] P. Diaconis and D. Freedman. Finite exchangeable sequences. *Ann. Probab.*, 8:745–764, 1980. MR0577313

[8] P. Diaconis and D. Freedman. A dozen de Finetti-style results in search of a theory. *Ann. Inst. Henri Poincaré, Probab. Stat.*, 23:397–423, 1987. MR0898502

[9] P. Diaconis and D. A. Freedman. Conditional limit theorems for exponential families and finite versions of de Finetti's theorem. *J. Theor. Probab.*, 1(4):381–410, 1988. MR0958245

[10] P. Eichelsbacher and M. Löwe. Stein's method for dependent random variables occurring in statistical mechanics. *Electron. J. Probab.*, 15:no. 30, 962–988, 2010. MR2659754

[11] R. S. Ellis. *Entropy, Large Deviations, and Statistical Mechanics*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1985 original. MR2189669

[12] R. S. Ellis, C. M. Newman, and J. S. Rosen. Limit theorems for sums of dependent random variables occurring in statistical mechanics. II. Conditioning, multiple phases, and metastability. *Z. Wahrsch. Verw. Gebiete*, 51(2):153–169, 1980. MR0566313

[13] M. Fleermann, W. Kirsch, and G. Toth. Local central limit theorem for multi-group Curie-Weiss models. *J. Theoret. Probab.*, 35(3):2009–2019, 2022. MR4488568

[14] S. Friedli and Y. Velenik. *Statistical Mechanics of Lattice Systems*. Cambridge University Press, 2017. MR3752129

[15] M. V. Jambunathan. Some properties of Beta and Gamma distributions. *Ann. Math. Statistics*, 25:401–405, 1954. MR0061317

[16] Z. Kabluchko, M. Löwe, and K. Schubert. Fluctuations of the magnetization for Ising models on Erdos-Rényi random graphs—the regimes of low temperature and external magnetic field. *ALEA Lat. Am. J. Probab. Math. Stat.*, 19(1):537–563, 2022. MR4394308

[17] W. Kirsch. A survey on the method of moments, 2015. Available at: https://www.fernuni-hagen.de/mi/fakultaet/emeriti/docs/kirsch/momente.pdf.

[18] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark. *NIST Handbook of Mathematical Functions*. Cambridge University Press, 2010. MR2723248

[19] A. Röllin and N. Ross. Local limit theorems via Landau-Kolmogorov inequalities. *Bernoulli*, 21(2):851–880, 2015. MR3338649

[20] M. Shamis and O. Zeitouni. The Curie-Weiss model with complex temperature: phase transitions. *J. Stat. Phys.*, 172(2):569–591, 2018. MR3824953

# Electronic Journal of Probability
# Electronic Communications in Probability

## Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

## Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4], ProjectEuclid[5]

- Purely electronic

## Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1] LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2] EJMS: Electronic Journal Management System: https://vtex.lt/services/ejms-peer-review/
[3] IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4] BS: Bernoulli Society http://www.bernoulli-society.org/
[5] Project Euclid: https://projecteuclid.org/
[6] IMS Open Access Fund: https://imstat.org/shop/donation/