

# STEIN'S METHOD FOR DEPENDENT RANDOM VARIABLES OCCURRING IN STATISTICAL MECHANICS

PETER EICHELSBACHER<sup>1</sup> AND MATTHIAS LÖWE<sup>2</sup>

**Abstract:** We obtain rates of convergence in limit theorems of partial sums  $S_n$  for certain sequences of dependent, identically distributed random variables, which arise naturally in statistical mechanics, in particular, in the context of the Curie-Weiss models. Under appropriate assumptions there exists a real number  $\alpha$ , a positive real number  $\mu$ , and a positive integer  $k$  such that  $(S_n - n\alpha)/n^{1-1/2k}$  converges weakly to a random variable with density proportional to  $\exp(-\mu|x|^{2k}/(2k)!)$ . We develop Stein's method for exchangeable pairs for a rich class of distributional approximations including the Gaussian distributions as well as the non-Gaussian limit distributions with density proportional to  $\exp(-\mu|x|^{2k}/(2k)!)$ . Our results include the optimal Berry-Esseen rate in the Central Limit Theorem for the total magnetization in the classical Curie-Weiss model, for high temperatures as well as at the critical temperature  $\beta_c = 1$ , where the Central Limit Theorem fails. Moreover, we analyze Berry-Esseen bounds as the temperature  $1/\beta_n$  converges to one and obtain a threshold for the speed of this convergence. Single spin distributions satisfying the Griffiths-Hurst-Sherman (GHS) inequality like models of liquid helium or continuous Curie-Weiss models are considered.

*MSC 2000:* Primary 60F05, secondary 60G09, 60K35, 82B20, 82D40

*Keywords and phrases.* Berry-Esseen bound, Stein's method, exchangeable pairs, Curie-Weiss models, critical temperature, GHS-inequality

This research was done partly at the Mathematisches Forschungsinstitut Oberwolfach during a stay within the Research in Pairs Programme in September 2008.

---

<sup>1</sup>Ruhr-Universität Bochum, Fakultät für Mathematik, NA3/68, D-44780 Bochum, Germany, [Peich@math.ruhr-uni-bochum.de](mailto:Peich@math.ruhr-uni-bochum.de)

<sup>2</sup>Universität Münster, Fachbereich Mathematik und Informatik, Institut für Mathematische Statistik, Einsteinstr. 62, D-48149 Münster, Germany [maloewe@math.uni-muenster.de](mailto:maloewe@math.uni-muenster.de)

## 1. INTRODUCTION AND MAIN RESULT

There is a long tradition in considering mean–field models in statistical mechanics. The Curie–Weiss model is famous, since it exhibits a number of properties of real substances, such as multiple phases, metastable states and others, explicitly. The aim of this paper is to prove Berry–Esseen bounds for the sums of dependent random variables occurring in statistical mechanics under the name Curie–Weiss models. To this end, we will develop Stein’s method for exchangeable pairs (see [22]) for a rich class of distributional approximations. For an overview of results on the Curie–Weiss models and related models, see [10], [12], [14].

For a fixed positive integer  $d$  and a finite subset  $\Lambda$  of  $\mathbb{Z}^d$ , a ferromagnetic crystal is described by random variables  $X_i^\Lambda$  which represent the spins of the atom at sites  $i \in \Lambda$ , where  $\Lambda$  describes the macroscopic shape of the crystal. In *Curie–Weiss* models, the joint distribution at fixed temperature  $T > 0$  of the spin random variables is given by

$$P_{\Lambda,\beta}((x_i)) := P_{\Lambda,\beta}((X_i^\Lambda)_{i \in \Lambda} = (x_i)_{i \in \Lambda}) := \frac{1}{Z_\Lambda(\beta)} \exp\left(\frac{\beta}{2|\Lambda|} \left(\sum_{i \in \Lambda} x_i\right)^2\right) \prod_{i \in \Lambda} d\varrho(x_i). \quad (1.1)$$

Here  $\beta := T^{-1}$  is the inverse temperature and  $Z_\Lambda(\beta)$  is a normalizing constant known as the partition function and  $|\Lambda|$  denotes the cardinality of  $\Lambda$ . Moreover  $\varrho$  is the distribution of a single spin in the limit  $\beta \rightarrow 0$ . We define  $S_\Lambda = \sum_{i \in \Lambda} X_i^\Lambda$ , the *total magnetization* inside  $\Lambda$ . We take without loss of generality  $d = 1$  and  $\Lambda = \{1, \dots, n\}$ , where  $n$  is a positive integer. We write  $n$ ,  $X_i^{(n)}$ ,  $P_{n,\beta}$  and  $S_n$ , respectively, instead of  $|\Lambda|$ ,  $X_i^\Lambda$ ,  $P_{\Lambda,\beta}$ , and  $S_\Lambda$ , respectively. In the case where  $\beta$  is fixed we may even sometimes simply write  $P_n$ .

We assume that  $\varrho$  is in the class  $\mathcal{B}$  of non-degenerate symmetric Borel probability measures on  $\mathbb{R}$  which satisfy

$$\int \exp\left(\frac{bx^2}{2}\right) d\varrho(x) < \infty \quad \text{for all } b > 0. \quad (1.2)$$

In the classical Curie–Weiss model, spins are distributed in  $\{-1, +1\}$  according to  $\varrho = \frac{1}{2}(\delta_{-1} + \delta_1)$ . More generally, the Curie–Weiss model carries an additional parameter  $h > 0$  called *external magnetic field* which leads to the modified measure, given by

$$P_{n,\beta,h}(x) = \frac{1}{Z_{n,\beta,h}} \exp\left(\frac{\beta}{2n} S_n^2 + \beta h S_n\right) d\varrho^{\otimes n}(x), \quad x = (x_i).$$

The measures  $P_{n,\beta,h}$  is completely determined by the value of the total magnetization. It is therefore called an *order parameter* and its behaviour will be studied in this paper. The non-negative external magnetic field strength may even depend on the site:

$$P_{n,\beta,h_1,\dots,h_n}(x) = \frac{1}{Z_{n,\beta,h_1,\dots,h_n}} \exp\left(\frac{\beta}{2n} S_n^2 + \beta \sum_{i=1}^n h_i x_i\right) d\varrho^{\otimes n}(x), \quad x = (x_i). \quad (1.3)$$

In the general case (1.1), we will see (analogously to the treatment in [12, 14]) that the asymptotic behaviour of  $S_n$  depends crucially on the extremal points of a function  $G$  (which is a transform of the

rate function in a corresponding large deviation principle): define

$$\phi_\varrho(s) := \log \int \exp(sx) d\varrho(x)$$

and

$$G_\varrho(\beta, s) := \frac{\beta s^2}{2} - \phi_\varrho(\beta s). \quad (1.4)$$

We shall drop  $\beta$  in the notation for  $G$  whenever there is no danger of confusion, similarly we will suppress  $\varrho$  in the notation for  $\phi$  and  $G$ . For any measure  $\varrho \in \mathcal{B}$ ,  $G$  was proved to have global minima, which can be only finite in number, see [12, Lemma 3.1]. Define  $C = C_\varrho$  to be the discrete, non-empty set of minima (local or global) of  $G$ . If  $\alpha \in C$ , then there exists a positive integer  $k := k(\alpha)$  and a positive real number  $\mu := \mu(\alpha)$  such that

$$G(s) = G(\alpha) + \frac{\mu(\alpha)(s - \alpha)^{2k}}{(2k)!} + \mathcal{O}((s - \alpha)^{2k+1}) \quad \text{as } s \rightarrow \alpha. \quad (1.5)$$

The numbers  $k$  and  $\mu$  are called the *type* and *strength*, respectively, of the extremal point  $\alpha$ . Moreover, we define the maximal type  $k^*$  of  $G$  by the formula

$$k^* = \max\{k(\alpha); \alpha \text{ is a global minimum of } G\}.$$

Note that the  $\mu(\alpha)$  can be calculated explicitly: one gets

$$\mu(\alpha) = \beta - \beta^2 \phi''(\beta \alpha) \quad \text{if } k = 1 \quad (1.6)$$

while

$$\mu(\alpha) = -\beta^{2k} \phi^{(2k)}(\beta \alpha) \quad \text{if } k \geq 2 \quad (1.7)$$

(see [14]).

An interesting point is, that the global minima of  $G$  of maximal type correspond to stable states, meaning that multiple minima represent a mixed phase and a unique global minimum a pure phase. For details see the discussions in [14].

The following is known about the fluctuation behaviour of  $S_n$  under  $P_n$ . In the classical model ( $\varrho$  is the symmetric Bernoulli measure), for  $0 < \beta < 1$ , in [12] the Central Limit Theorem is proved:

$$\frac{\sum_{i=1}^n X_i}{\sqrt{n}} \rightarrow N(0, \sigma^2(\beta))$$

in distribution with respect to the Curie–Weiss finite volume Gibbs states with  $\sigma^2(\beta) = (1 - \beta)^{-1}$ . Since for  $\beta = 1$  the variance  $\sigma^2(\beta)$  diverges, the Central Limit Theorem fails at the critical point. In [12] it is proved that for  $\beta = 1$  there exists a random variable  $X$  with probability density proportional to  $\exp(-\frac{1}{12}x^4)$  such that as  $n \rightarrow \infty$

$$\frac{\sum_{i=1}^n X_i}{n^{3/4}} \rightarrow X$$

in distribution with respect to the finite-volume Gibbs states. Asymptotic independence properties and propagation of chaos for blocks of size  $o(n)$  have been investigated in [2].

In general, given  $\varrho \in \mathcal{B}$ , let  $\alpha$  be one of the global minima of maximal type  $k$  and strength  $\mu$  of  $G_\varrho$ . Then

$$\frac{S_n - n\alpha}{n^{1-1/2k}} \rightarrow X_{k,\mu,\beta}$$

in distribution, where  $X_{k,\mu,\beta}$  is a random variable with probability density  $f_{k,\mu,\beta}$ , defined by

$$f_{1,\mu,\beta}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-x^2/2\sigma^2) \quad (1.8)$$

and for  $k \geq 2$

$$f_{k,\mu,\beta}(x) = \frac{\exp(-\mu x^{2k}/(2k)!)}{\int \exp(-\mu x^{2k}/(2k)!) dx}. \quad (1.9)$$

Here,  $\sigma^2 = \frac{1}{\mu} - \frac{1}{\beta}$  so that for  $\mu = \mu(\alpha)$  as in (1.6),  $\sigma^2 = ([\phi''(\beta\alpha)]^{-1} - \beta)^{-1}$  (see [12], [14]). Moderate deviation principles have been investigated in [7].

In [11] and [14], a class of measures  $\varrho$  is described exhibiting a behaviour similar to that of the classical Curie–Weiss model. Assume that  $\varrho$  is any symmetric measure that satisfies the Griffiths–Hurst–Sherman (GHS) inequality,

$$\frac{d^3}{ds^3} \phi_\varrho(s) \leq 0 \quad \text{for all } s \geq 0, \quad (1.10)$$

(see also [13, 16]). One can show that in this case  $G$  has the following properties: There exists a value  $\beta_c$ , the inverse critical temperature, and  $G$  has a unique global minimum at the origin for  $0 < \beta \leq \beta_c$  and exactly two global minima, of equal type, for  $\beta > \beta_c$ . For  $\beta_c$  the unique global minimum is of type  $k \geq 2$  whereas for  $\beta \in (0, \beta_c)$  the unique global minimum is of type 1. At  $\beta_c$  the law of large numbers still holds, but the fluctuations of  $S_n$  live on a smaller scale than  $\sqrt{n}$ . This critical temperature can be explicitly computed as  $\beta_c = 1/\phi''(0) = 1/\text{Var}_\varrho(X_1)$ . By rescaling the  $X_i$  we may thus assume that  $\beta_c = 1$ .

Alternatively, the GHS-inequality can be formulated in the terms of  $Z_{n,\beta,h_1,\dots,h_n}$ , defined in (1.3):

$$\begin{aligned} 0 &\geq \frac{\partial^3}{\partial h_i \partial h_j \partial h_k} \log Z_{n,\beta,h_1,\dots,h_n} \\ &= E(X_i X_j X_k) - E(X_i)E(X_j X_k) - E(X_j)E(X_i X_k) \\ &\quad - E(X_k)E(X_i X_j) + 2E(X_i)E(X_j)E(X_k) \end{aligned} \quad (1.11)$$

for all (not necessarily distinct) sites  $i, j, k \in \{1, \dots, n\}$ . Here  $\mathbb{E}$  denotes the expectation with respect to  $P_{n,\beta,h_1,\dots,h_n}$ . The GHS inequality has a number of interesting implications, see [11].

With GHS, we will denote the set of measures  $\varrho \in \mathcal{B}$  such that the GHS-inequality (1.10) is valid (for  $P_{n,\beta,h_1,\dots,h_n}$  in the sense of (1.11)). We will give examples in Section 7.

**Remark 1.1.** In [12, Lemma 4.1], for  $\varrho \in \mathcal{B}$  it is proved that  $G$  has a unique global minimum if and only if

$$\int \exp(sx) d\varrho(x) < \exp(s^2/2), \quad \text{for } s \text{ real,}$$

where the right hand side of this strict inequality is the moment generating function of a standard normal random variable. Moreover, in the same Lemma it is proved that  $G$  has a local minimum at the origin of type  $k$  and strength  $\mu$  if and only if

$$\bar{\mu}_j - \mu_j(\varrho) = \begin{cases} 0 & \text{for } j = 0, 1, \dots, 2k - 1, \\ \mu > 0 & \text{for } j = 2k. \end{cases}$$

Here  $\mu_j(\varrho)$  and  $\bar{\mu}_j$  define the  $j$ 'th moment of  $\varrho$  and the  $j$ 'th moment of a standard normal random variables, respectively. Note that this in particular implies  $\mu_1(\varrho) = \mathbb{E}_\varrho(X_1) = 0$ .

The aim of this paper is to prove the following theorems:

### 1.1 Results for the classical Curie-Weiss model

**Theorem 1.2** (classical Curie-Weiss model, optimal Berry-Esseen bounds outside the critical temperature). *Let  $\varrho = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$  and  $0 < \beta < 1$ . We have*

$$\sup_{z \in \mathbb{R}} \left| P_n \left( S_n / \sqrt{n} \leq z \right) - \Phi_\beta(z) \right| \leq C n^{-1/2}, \quad (1.12)$$

where  $\Phi_\beta$  denotes the distribution function of the normal distribution with expectation zero and variance  $(1 - \beta)^{-1}$ , and  $C$  is an absolute constant, depending on  $\beta$ , only.

**Theorem 1.3** (classical Curie-Weiss model, optimal Berry-Esseen bounds at the critical temperature). *Let  $\varrho = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$  and  $\beta = 1$ . We have*

$$\sup_{z \in \mathbb{R}} \left| P_n \left( S_n / n^{3/4} \leq z \right) - F(z) \right| \leq C n^{-1/2}, \quad (1.13)$$

where

$$F(z) := \frac{1}{Z} \int_{-\infty}^z \exp(-x^4/12) dx, \quad (1.14)$$

$Z := \int_{\mathbb{R}} \exp(-x^4/12) dx$  and  $C$  is an absolute constant.

**Theorem 1.4** (Berry-Esseen bounds for size-dependent temperatures). *Let  $\varrho = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$  and  $0 < \beta_n < \infty$  depend on  $n$  in such a way that  $\beta_n \rightarrow 1$  monotonically as  $n \rightarrow \infty$ . Then the following assertions hold:*

- (1) *If  $\beta_n - 1 = \frac{\gamma}{\sqrt{n}}$  for some  $\gamma \neq 0$ , we have*

$$\sup_{z \in \mathbb{R}} \left| P_n \left( S_n / n^{3/4} \leq z \right) - F_\gamma(z) \right| \leq C n^{-1/2} \quad (1.15)$$

with

$$F_\gamma(z) := \frac{1}{Z} \int_{-\infty}^z \exp\left(-\frac{x^4}{12} + \frac{\gamma x^2}{2}\right) dx.$$

where  $Z := \int_{\mathbb{R}} \exp\left(-\frac{x^4}{12} + \frac{\gamma x^2}{2}\right) dx$  and  $C$  is an absolute constant.

- (2) *If  $|\beta_n - 1| \ll n^{-1/2}$ ,  $S_n / n^{3/4}$  converges in distribution to  $F$ , given in (1.14). Moreover, if  $|\beta_n - 1| = \mathcal{O}(n^{-1})$ , (1.13) holds true.*

(3) If  $|\beta_n - 1| \gg n^{-1/2}$ , the Kolmogorov distance of the distribution of  $\sqrt{\frac{1-\beta_n}{n}} \sum_{i=1}^n X_i$  and the normal distribution  $N(0, (1 - \beta_n)^{-1})$  converges to zero. Moreover, if  $|\beta_n - 1| \gg n^{-1/4}$ , we obtain

$$\sup_{z \in \mathbb{R}} \left| P_n \left( \frac{\sqrt{(1-\beta_n)} S_n}{\sqrt{n}} \leq z \right) - \Phi_{\beta_n}(z) \right| \leq C n^{-1/2}$$

with an absolute constant  $C$ .

**Remark 1.5.** In [1], Barbour obtained distributional limit theorems, together with rates of convergence, for the equilibrium distributions of a variety of one-dimensional Markov population processes. In section 3 he mentioned, that his results can be interpreted in the framework of [12]. As far as we understand, his result (3.9) can be interpreted as the statement (1.13), but with the rate  $n^{-1/4}$ .

**Remark 1.6.** In the first assertion of Theorem 1.4, our method of proof allows to compare the distribution of  $S_n/n^{3/4}$  alternatively with the distribution with Lebesgue-density proportional to

$$\exp\left(-\frac{\beta_n^3 x^4}{12} + \frac{\gamma x^2}{2}\right).$$

To be able to compare the distribution of interest with a distribution depending on  $n$  (on  $\beta_n$ ), is one of the advantages of Stein's method. The proof of this statement follows immediately from the proof of Theorem 1.4.

If in Theorem 1.4 (2)  $|\beta_n - 1| \gg n^{-1}$  the speed of convergence reduces to  $\mathcal{O}(\sqrt{n}|\beta_n - 1|)$ . Likewise, if in Theorem 1.4 (3)  $|\beta_n - 1| \ll n^{-1/4}$ , the speed of convergence is  $\mathcal{O}(\frac{1}{n|1-\beta_n|})$ . This reduced speed of convergence reflects the influence of two potential limiting measures. Next to the "true" limit there is also the limit measure from part (1) of Theorem 1.4, which in these cases is relatively close to our measures of interest.

## 1.2 Results for a general class of Curie-Weiss models

More generally, we obtain Berry-Esseen bounds for sums of dependent random variables occurring in the general Curie-Weiss models. We will be able to obtain Berry-Esseen-type results for  $\varrho$ -a.s. bounded single-spin variables  $X_i$ :

**Theorem 1.7.** *Given  $\varrho \in \mathcal{B}$  in GHS, let  $\alpha$  be the global minimum of type  $k$  and strength  $\mu$  of  $G_\varrho$ . Assume that the single-spin random variables  $X_i$  are bounded  $\varrho$ -a.s. In the case  $k = 1$  we obtain*

$$\sup_{z \in \mathbb{R}} \left| P_n \left( \frac{S_n}{\sqrt{n}} \leq z \right) - \Phi_W(z) \right| \leq C n^{-1/2}, \quad (1.16)$$

where  $W := S_n/\sqrt{n}$  and  $\Phi_W$  denotes the distribution function of the normal distribution with mean zero and variance  $\mathbb{E}(W^2)$  and  $C$  is an absolute constant depending on  $0 < \beta < 1$ . For  $k \geq 2$  we obtain

$$\sup_{z \in \mathbb{R}} \left| P_n \left( \frac{S_n - n\alpha}{n^{1-1/2k}} \leq z \right) - \widehat{F}_{W,k}(z) \right| \leq C_k n^{-1/k} \quad (1.17)$$

where  $\widehat{F}_{W,k}(z) := \int_{-\infty}^z \widehat{f}_{W,k}(x) dx$  with  $\widehat{f}_{W,k}$  defined by

$$\widehat{f}_{W,k}(x) := \frac{\exp\left(-\frac{x^{2k}}{2k \mathbb{E}(W^{2k})}\right)}{\int \exp\left(-\frac{x^{2k}}{2k \mathbb{E}(W^{2k})}\right) dx}$$

with  $W := \frac{S_n - n\alpha}{n^{1-1/2k}}$  and  $C_k$  is an absolute constant.

**Theorem 1.8.** *Let  $\varrho \in \mathcal{B}$  satisfy the GHS-inequality and assume that  $\beta_c = 1$ . Let  $\alpha$  be the global minimum of type  $k$  with  $k \geq 2$  and strength  $\mu_k$  of  $G_\varrho$  and let the single-spin variable  $X_i$  be bounded. Let  $0 < \beta_n < \infty$  depend on  $n$  in such a way that  $\beta_n \rightarrow 1$  monotonically as  $n \rightarrow \infty$ . Then the following assertions hold true:*

(1) *If  $\beta_n - 1 = \frac{\gamma}{n^{1-k}}$  for some  $\gamma \neq 0$ , we have*

$$\sup_{z \in \mathbb{R}} \left| P_n \left( \frac{S_n - n\alpha}{n^{1-1/2k}} \leq z \right) - F_{W,k,\gamma}(z) \right| \leq C_k n^{-1/k} \quad (1.18)$$

with

$$F_{W,k,\gamma}(z) := \frac{1}{Z} \int_{-\infty}^z \exp\left(-c_W^{-1} \left( \frac{\mu_k}{(2k)!} x^{2k} - \frac{\gamma}{2} x^2 \right)\right) dx.$$

where  $Z := \int_{\mathbb{R}} \exp\left(-c_W^{-1} \left( \frac{\mu_k}{(2k)!} x^{2k} - \frac{\gamma}{2} x^2 \right)\right) dx$ , with  $W := \frac{S_n - n\alpha}{n^{1-1/2k}}$ ,

$$c_W := \frac{\mu_k}{(2k)!} \mathbb{E}(W^{2k}) - \gamma \mathbb{E}(W^2)$$

and  $C_k$  is an absolute constant.

(2) *If  $|\beta_n - 1| \ll n^{-(1-1/k)}$ ,  $\frac{S_n - n\alpha}{n^{1-1/2k}}$  converges in distribution to  $\widehat{F}_{W,k}$ , defined as in Theorem 1.7. Moreover, if  $|\beta_n - 1| = \mathcal{O}(n^{-1})$ , (1.17) holds true.*

(3) *If  $|\beta_n - 1| \gg n^{-(1-1/k)}$ , the Kolmogorov distance of the distribution of  $W := \sqrt{\frac{1-\beta_n}{n}} \sum_{i=1}^n X_i$  and the normal distribution  $N(0, \mathbb{E}(W^2))$  converges to zero. Moreover, if  $|\beta_n - 1| \gg n^{-(1/2-1/2k)}$ , we obtain*

$$\sup_{z \in \mathbb{R}} \left| P_n \left( \frac{\sqrt{(1-\beta_n)} S_n}{\sqrt{n}} \leq z \right) - \Phi_W(z) \right| \leq C n^{-1/2}$$

with an absolute constant  $C$ .

**Remark 1.9.** Since the symmetric Bernoulli law is GHS, Theorems 1.7 and 1.8 include Berry-Esseen type results for this case. But these results differ from the results in Theorem 1.2, 1.3 and 1.4 with respect to the limiting laws: the laws in 1.7 and 1.8 depend on moments of  $W$ . The bounds in Theorems 1.2-1.4 are easier to obtain; moreover their proofs apply Corollary 2.8 and part (2) of Theorem 4.6 which are less involved versions of Stein's method for exchangeable pairs.

For arbitrary  $\varrho \in \text{GHS}$  we are able to proof good bounds with respect to the Wasserstein-metric. For any class of test functions  $\mathcal{H}$ , a distance on probability measures on  $\mathbb{R}$  can be defined by

$$d_{\mathcal{H}}(P, Q) = \sup_{h \in \mathcal{H}} \left| \int h dP - \int h dQ \right|.$$

The class of test functions  $h$  for the Wasserstein distance  $d_w$  is just the Lipschitz functions  $\text{Lip}(1)$  with constant no greater than 1. The total variation distance is given by the set  $\mathcal{H}$  of indicators of Borel sets, the Kolmogorov distance  $d_K$  by the set of indicators of half lines.

Only for technical reasons, we consider now a modified model. Let

$$\widehat{P}_{n,\beta,h}(x) = \frac{1}{\widehat{Z}_{n,\beta,h}} \exp\left(\frac{\beta}{n} \sum_{1 \leq i < j \leq n} x_i x_j + \beta h \sum_{i=1}^n x_i\right) d\varrho^{\otimes n}(x), \quad x = (x_i).$$

**Theorem 1.10.** *Given the Curie-Weiss model  $\widehat{P}_{n,\beta}$  and  $\varrho \in \mathcal{B}$  in GHS, let  $\alpha$  be the global minimum of type  $k$  and strength  $\mu$  of  $G_\varrho$ . In the case  $k = 1$ , for any uniformly Lipschitz function  $h$  we obtain for  $W = S_n/\sqrt{n}$  that*

$$|\mathbb{E}(h(W)) - \Phi_W(h)| \leq \|h'\| C \frac{\max(\mathbb{E}|X_1|^3, \mathbb{E}|X'_1|^3)}{\sqrt{n}}.$$

Here  $C$  is a constant depending on  $0 < \beta < 1$  and  $\Phi_W(h) := \int_{\mathbb{R}} h(z) \Phi_W(dz)$ . The random variable  $X'_i$  is drawn from the conditional distribution of the  $i$ 'th coordinate  $X_i$  given  $(X_j)_{j \neq i}$  (this choice will be explained in Section 3). For  $k \geq 2$  we obtain for any uniformly Lipschitz function  $h$  and for  $W := \frac{S_n - n\alpha}{n^{1-1/2k}}$

$$|\mathbb{E}(h(W)) - \widehat{F}_{W,k}(h)| \leq \|h'\| \left( C_1 \frac{1}{\sqrt{n}} + \frac{C_2 \max(\mathbb{E}|X_1|^3, \mathbb{E}|X'_1|^3)}{n^{1-1/2k}} \right).$$

Here  $C_1, C_2$  are constants, and  $\widehat{F}_{W,k}(h) := \int_{\mathbb{R}} h(z) \widehat{F}_{W,k}(dz)$ .

**Remark 1.11.** Assume that there exists a  $\delta$  such that for any uniformly Lipschitz function  $h$ ,  $|\mathbb{E}h(W) - F(h)| \leq \delta \|h'\|$ , where  $W$  is a random variable,  $F(h) := \int_{\mathbb{R}} h(z) F(dz)$  for some distribution function  $F$ , then from the definition of the Wasserstein distance it follows immediately that  $\sup_{h \in \text{Lip}(1)} |\mathbb{E}h(W) - F(h)| \leq \delta$ . Moreover, the Kolmogorov distance  $\sup_z |P(W \leq z) - F(z)|$  can be bounded by  $c_F \delta^{1/2}$ , where  $c_F$  is some constant depending on  $F$  (the proof follows the lines of [6, Theorem 3.1]).

**Remark 1.12.** In [11], the distribution of the spins  $\varrho$  are allowed to depend on the site. They define a subclass  $\mathcal{G}$  of  $\mathcal{B}$  such that for  $\varrho_1, \dots, \varrho_n \in \mathcal{G}$  the GHS inequality holds. In Section 7 we present a large class of measures which belong to  $\mathcal{G}$  (see [11, Theorem 1.2]). The GHS inequality itself has a number of interesting implications like the concavity of the average magnetization as a function of the external field  $h$  or the monotonicity of correlation length in Ising models. These and other implications can be found in [11] and references therein. Note that for  $\varrho \in \text{GHS}$ ,  $\phi_\varrho(s) \leq \frac{1}{2} \sigma_\varrho^2 s^2$  for all real  $s$ , where  $\sigma_\varrho^2 = \int_{\mathbb{R}} x^2 \varrho(dx)$ . These measures are called *sub-Gaussian*. Very important for our proofs of Berry-Esseen bounds will be the following correlation-inequality due to Lebowitz [18]: If  $\mathbb{E}$  denotes the expectation with respect to the measure  $P_{n,\beta,h_1,\dots,h_n}$ , one observes easily that for any

$\varrho \in \mathcal{B}$  and sites  $i, j, k, l \in \{1, \dots, n\}$  the following identity holds:

$$\begin{aligned} & \frac{\partial^3}{\partial h_i \partial h_j \partial h_k} \mathbb{E}(X_l) \Big|_{\text{all } h_i=0} \\ &= \mathbb{E}(X_i X_j X_k X_l) - \mathbb{E}(X_i X_j) \mathbb{E}(X_k X_l) - \mathbb{E}(X_i X_k) \mathbb{E}(X_j X_l) - \mathbb{E}(X_i X_l) \mathbb{E}(X_j X_k). \end{aligned} \quad (1.19)$$

Lebowitz [18] proved that if  $\varrho \in \text{GHS}$ , then (1.19) is non-positive (see [10, V.13.7.(b)] and [17]). Stein's method reduces to the computation of, or bounds on, *low order* moments, perhaps even only on variances of certain quantities. Such variance computations can be very difficult. We will see in the proof of Theorem 1.7 and Theorem 1.8 the use of Lebowitz' inequality for bounding the variances successfully.

In the situation of Theorem 1.7 and Theorem 1.8 we can bound higher order moments as follows:

**Lemma 1.13.** *Given  $\varrho \in \mathcal{B}$ , let  $\alpha$  be one of the global minima of maximal type  $k$  for  $k \geq 1$  and strength  $\mu$  of  $G_\varrho$ . For*

$$W := \frac{S_n - n\alpha}{n^{1-1/2k}}$$

*we obtain for any  $l \in \mathbb{N}$*

$$\mathbb{E}|W|^l \leq \text{const.}(l).$$

We prepare for the proof of Lemma 1.13. It considers a well known transformation – sometimes called the *Hubbard-Stratonovich transformation* – of our measure of interest.

**Lemma 1.14.** *Let  $m \in \mathbb{R}$  and  $0 < \gamma < 1$  be real numbers. Consider the measure  $Q_{n,\beta} := (P_n \circ (\frac{S_n - nm}{n^\gamma})^{-1}) * \mathcal{N}(0, \frac{1}{\beta n^{2\gamma-1}})$  where  $\mathcal{N}(0, \frac{1}{\beta n^{2\gamma-1}})$  denotes a Gaussian random variable with mean zero and variance  $\frac{1}{\beta n^{2\gamma-1}}$ . Then for all  $n \geq 1$  the measure  $Q_{n,\beta}$  is absolutely continuous with density*

$$\frac{\exp(-nG(\frac{s}{n^{1-\gamma}} + m))}{\int_{\mathbb{R}} \exp(-nG(\frac{s}{n^{1-\gamma}} + m)) ds}, \quad (1.20)$$

*where  $G$  is defined in equation (1.4).*

**Remark 1.15.** As shown in [12], Lemma 3.1, our condition (1.2) ensures that

$$\int_{\mathbb{R}} \exp\left(-nG\left(\frac{s}{n^{1-\gamma}} + m\right)\right) ds$$

is finite, such that the above density is well defined.

*Proof of Lemma 1.14.* The proof of this lemma can be found at many places, e.g. in [12], Lemma 3.3. □

*Proof of Lemma 1.13.* We apply the Hubbard-Stratonovich transformation with  $\gamma = 1 - 1/2k$ . It is clear that this does not change the finiteness of any of the moments of  $W$ . Using the Taylor expansion (1.5) of  $G$ , we see that the density of  $Q_{n,\beta}$  with respect to Lebesgue measure is given by  $\text{Const.} \exp(-x^{2k})$  (up to negligible terms, see e.g. [12], [7]). A measure with this density, of course, has moments of any finite order. □

**Remark 1.16.** As we will see, we only have to bound  $\mathbb{E}(W^4)$  in the classical model, when  $0 < \beta < 1$ . This can be obtained directly using the definition of  $P_n$  and Taylor-expansion. But already for the classical model, for  $\beta = 1$ , it is quite cumbersome to bound higher order moments via direct calculations.

In Section 2, we develop in Theorem 2.5, Corollary 2.8 and Corollary 2.9 refinements of Stein's method for exchangeable pairs in the case of normal approximation. As a first application we prove Theorem 1.2 in Section 3. In Section 4 we develop Stein's method for exchangeable pairs for a rich class of other distributional approximations. Obtaining good bounds for the solutions of the corresponding Stein equations in the appendix, we prove Theorem 1.3 and Theorem 1.4 in Section 5, applying Theorem 4.6. In Section 6, we proof Theorems 1.7, 1.8 and 1.10, applying Corollary 2.9 and Theorem 4.7. Section 7 contains a collection of examples including the Curie-Weiss model with three states, studying liquid helium, and a continuous Curie-Weiss model, where the single spin distribution  $\varrho$  is a uniform distribution.

## 2. STEIN'S METHOD WITH EXCHANGEABLE PAIRS FOR NORMAL APPROXIMATION

Stein introduced in [22] the exchangeable pair approach. Given a random variable  $W$ , Stein's method is based on the construction of another variable  $W'$  (some coupling) such that the pair  $(W, W')$  is exchangeable, i.e. their joint distribution is symmetric. The approach essentially uses the elementary fact that if  $(W, W')$  is an exchangeable pair, then  $\mathbb{E}g(W, W') = 0$  for all antisymmetric measurable functions  $g(x, y)$  such that the expectation exists. A theorem of Stein ([22, Theorem 1, Lecture III]) shows that a measure of proximity of  $W$  to normality may be provided in terms of the exchangeable pair, requiring  $W' - W$  to be sufficiently small. He assumed the linear regression property

$$\mathbb{E}(W'|W) = (1 - \lambda)W$$

for some  $0 < \lambda < 1$ . This approach has been successfully applied in many models, see [22] and for example [23] and references therein. In [19], the range of application was extended by replacing the linear regression property by a weaker condition, allowing to hold the regression property only approximately. The exchangeable pair approach is also successful for other distributional approximations, as will be shown in Section 4. We develop Stein's method by replacing the linear regression property by

$$\mathbb{E}(W'|W) = W + \lambda \psi(W) + R(W),$$

where  $\psi(x)$  will be depend on a continuous distribution under consideration. Before we consider in this section the case of normal approximation, we mention that this is not the first paper to study other distributional approximations via Stein's method. For a rather large class of continuous distributions, the Stein characterization was introduced in [23], following [22, Chapter 6]. In [23], the method of exchangeable pairs was introduced for this class of distribution and used in a simulation context. Recently, the exchangeable pair approach was introduced for exponential approximation in [4, Lemma 2.1].

For measuring the distance of the distribution of  $W$  and the standard normal distribution (or any other distribution), we would like to bound

$$|\mathbb{E}h(W) - \Phi(h)|$$

for a class of test functions  $h \in \mathcal{H}$ , where  $\Phi(h) := \int_{-\infty}^{\infty} h(z)\Phi(dz)$  and  $\Phi$  is the standard normal distribution function. One advantage of Stein's method is that we are able to obtain bounds for different distances like the Wasserstein distance  $d_w$ , the total variation distance  $d_{TV}$  or the Kolmogorov distance  $d_K$ . In [19], the exchangeable pair approach of Stein was developed for a broad class of non smooth functions  $h$ , applying standard smoothing inequalities.

They proved the following:

**Theorem 2.1** (Rinott, Rotar: 1997). *Consider a random variable  $W$  with  $\mathbb{E}(W) = 0$  and  $\mathbb{E}(W^2) = 1$ . Let  $(W, W')$  be an exchangeable pair (i.e., their joint distribution is symmetric). Define a random variable  $R = R(W)$  by*

$$E(W'|W) = (1 - \lambda)W + R,$$

where  $\lambda$  is a number satisfying  $0 < \lambda < 1$ . If moreover

$$|W' - W| \leq A$$

for a constant  $A$ . Then one obtains

$$\sup_{z \in \mathbb{R}} |P(W \leq z) - \Phi(z)| \leq \frac{12}{\lambda} \sqrt{\text{var}\{\mathbb{E}[(W' - W)^2|W]\}} + 37 \frac{\sqrt{\mathbb{E}(R^2)}}{\lambda} + 48\sqrt{2/\pi} \frac{A^3}{\lambda} + \sqrt{2/\pi} \frac{A^2}{\sqrt{\lambda}}. \quad (2.1)$$

**Remark 2.2.** Rinott and Rotar also proved a bound in the case, where  $|W' - W|$  is not assumed to be bounded. In this case, the last two summands on the right hand side of (2.1) have to be replaced by

$$\sqrt{\frac{a}{\lambda} \mathbb{E}|W' - W|^3}.$$

This estimation is crude, since even for a normalized sum of  $n$  independent variables  $W$ , it leads to a bound of the order  $n^{-1/4}$ . The advantage of the results in [19] is, that these bounds do not only apply to indicators on half lines, but also to a broad class of non smooth test functions, see [19, Section 1.2].

Chen and Shao introduced a concentration inequality approach. Here a concentration inequality is proved using the Stein identity (see [5] and [6]). In the context of the construction of an exchangeable pair, in [20] Shao and Su proved the following theorem:

**Theorem 2.3** (Shao, Su: 2005). *Let  $W$  be a random variable with  $\mathbb{E}(W) = 0$  and  $\mathbb{E}(W^2) \leq 1$  and  $(W, W')$  be an exchangeable pair such that*

$$E(W'|W) = (1 - \lambda)W$$

with  $0 < \lambda < 1$ , then for any  $a > 0$

$$\sup_{z \in \mathbb{R}} |P(W \leq z) - \Phi(z)| \leq \sqrt{\mathbb{E}\left(1 - \frac{1}{2\lambda} E((W - W')^2 | W)\right)^2} + \frac{0.41a^3}{\lambda} + 1.5a + \frac{1}{2\lambda} \mathbb{E}((W - W')^2 1_{\{|W - W'| \geq a\}}). \quad (2.2)$$

If  $|W - W'| \leq A$ , then the bound reduces to

$$\sup_{z \in \mathbb{R}} |P(W \leq z) - \Phi(z)| \leq \sqrt{\mathbb{E}\left(1 - \frac{1}{2\lambda} E((W - W')^2 | W)\right)^2} + \frac{0.41A^3}{\lambda} + 1.5A. \quad (2.3)$$

**Remark 2.4.** When  $|W - W'|$  is bounded, (2.3) improves (2.1) with respect to the constants.

Following the lines of the proofs in [19] and [20], we obtain the following refinement: Given two random variables  $X$  and  $Y$  defined on a common probability space, we denote by

$$d_K(X, Y) := \sup_{z \in \mathbb{R}} |P(X \leq z) - P(Y \leq z)|$$

the Kolmogorov distance of the distributions of  $X$  and  $Y$ .

**Theorem 2.5.** Let  $(W, W')$  be an exchangeable pair of real-valued random variables such that

$$E(W' | W) = (1 - \lambda)W + R$$

for some random variable  $R = R(W)$  and with  $0 < \lambda < 1$ . Assume that  $\mathbb{E}(W^2) \leq 1$ . Let  $Z$  be a random variable with standard normal distribution. Then for any  $A > 0$ ,

$$\begin{aligned} d_K(W, Z) &\leq \sqrt{\mathbb{E}\left(1 - \frac{1}{2\lambda} \mathbb{E}[(W' - W)^2 | W]\right)^2} + \left(\frac{\sqrt{2\pi}}{4} + 1.5A\right) \frac{\sqrt{\mathbb{E}(R^2)}}{\lambda} \\ &\quad + \frac{0.41A^3}{\lambda} + 1.5A + \frac{1}{2\lambda} \mathbb{E}((W - W')^2 1_{\{|W - W'| \geq A\}}). \end{aligned}$$

If  $|W - W'| \leq A$  for a constant  $A$ , we obtain the bound

$$d_K(W, Z) \leq \sqrt{\mathbb{E}\left(1 - \frac{1}{2\lambda} \mathbb{E}[(W' - W)^2 | W]\right)^2} + \left(\frac{\sqrt{2\pi}}{4} + 1.5A\right) \frac{\sqrt{\mathbb{E}(R^2)}}{\lambda} + \frac{0.41A^3}{\lambda} + 1.5A. \quad (2.4)$$

**Remark 2.6.** When  $|W - W'|$  is bounded, (2.4) improves (2.1) with respect to the Berry-Esseen constants.

*Proof.* We sketch the proof: For a function  $f$  with  $|f(x)| \leq C(1 + |x|)$  we obtain

$$\begin{aligned} 0 &= \mathbb{E}((W - W')(f(W') + f(W))) \\ &= \mathbb{E}((W - W')(f(W') - f(W))) + 2\lambda \mathbb{E}(W f(W)) - 2\mathbb{E}(f(W) R). \end{aligned} \quad (2.5)$$

Let  $f = f_z$  denote the solution of the Stein equation

$$f'_z(x) - x f_z(x) = 1_{\{x \leq z\}}(x) - \Phi(z). \quad (2.6)$$

We obtain

$$\begin{aligned}
P(W \leq z) - \Phi(z) &= \mathbb{E}(f'(W) - Wf(W)) \\
&= \mathbb{E}(f'(W)) - \frac{1}{2\lambda} \mathbb{E}((W - W')(f(W) - f(W'))) - \frac{1}{\lambda} \mathbb{E}(f(W)R) \\
&= \mathbb{E}\left(f'(W)\left(1 - \frac{1}{2\lambda}(W - W')^2\right)\right) + \mathbb{E}\left(f'(W)\frac{1}{2\lambda}(W - W')^2\right) \\
&\quad - \frac{1}{2\lambda} \mathbb{E}((W - W')(f(W) - f(W'))) - \frac{1}{\lambda} \mathbb{E}(f(W)R) \\
&= \mathbb{E}\left(f'(W)\left(1 - \frac{1}{2\lambda}(W - W')^2\right)\right) - \frac{1}{2\lambda} \mathbb{E}(2f(W)R) \\
&\quad - \frac{1}{2\lambda} \mathbb{E}\left[(W - W')(f(W) - f(W') - (W - W')f'(W))\right] \\
&=: T_1 + T_2 + T_3.
\end{aligned} \tag{2.7}$$

Using  $|f'(x)| \leq 1$  for all real  $x$  (see [6, Lemma 2.2]), we obtain the bound

$$|T_1| \leq \sqrt{\mathbb{E}\left(1 - \frac{1}{2\lambda} \mathbb{E}[(W' - W)^2|W]\right)^2}.$$

Using  $0 < f(x) \leq \sqrt{2\pi}/4$  (see [6, Lemma 2.2]), we have

$$|T_2| \leq \frac{\sqrt{2\pi}}{4\lambda} \mathbb{E}(|R|) \leq \frac{\sqrt{2\pi}}{4\lambda} \sqrt{\mathbb{E}(R^2)}.$$

Bounding  $T_3$  we apply the concentration technique, see [20]:

$$\begin{aligned}
(-2\lambda)T_3 &= \mathbb{E}\left((W - W')1_{\{|W - W'| > A\}} \int_{-(W - W')}^0 (f'(W + t) - f'(W))dt\right) \\
&\quad + \mathbb{E}\left((W - W')1_{\{|W - W'| \leq A\}} \int_{-(W - W')}^0 (f'(W + t) - f'(W))dt\right).
\end{aligned} \tag{2.8}$$

The modulus of the first term can be bounded by  $\mathbb{E}((W - W')^2 1_{\{|W - W'| > A\}})$  using  $|f'(x) - f'(y)| \leq 1$  for all real  $x$  and  $y$  (see [6, Lemma 2.2]). Using the Stein identity (2.6), the second summand can be represented as

$$\begin{aligned}
&\mathbb{E}\left((W - W')1_{\{|W - W'| \leq A\}} \int_{-(W - W')}^0 ((W + t)f(W + t) - Wf(W))dt\right) \\
&\quad + \mathbb{E}\left((W - W')1_{\{|W - W'| \leq A\}} \int_{-(W - W')}^0 (1_{\{W + t \leq z\}} - 1_{\{W \leq z\}})dt\right) =: U_1 + U_2.
\end{aligned}$$

Next observe that  $|U_1| \leq 0.82A^3$ , see [20]: by the mean value theorem one gets

$$(W + t)f(W + t) - Wf(W) = W(f(W + t) - f(W)) + tf(W + t) = W\left(\int_0^1 f'(W + ut)tdu\right) + tf(W + t).$$

Hence

$$|(W + t)f(W + t) - Wf(W)| \leq |W||t| + |t|\sqrt{2\pi}/4 = |t|(\sqrt{2\pi}/4 + |W|).$$

Using  $\mathbb{E}|W| \leq \sqrt{\mathbb{E}(W^2)} \leq 1$  gives the bound. The term  $U_2$  can be bounded by

$$\mathbb{E}((W - W')^2 I_{\{0 \leq (W - W') \leq A\}} 1_{\{z \leq W \leq z + A\}}).$$

Under the assumptions of our Theorem we proceed as in [20] and obtain the following concentration inequality:

$$\mathbb{E}((W - W')^2 I_{\{0 \leq (W - W') \leq a\}} 1_{\{z \leq W \leq z + A\}}) \leq 3A(\lambda + \mathbb{E}(R)). \quad (2.9)$$

To see this, we apply the estimate

$$\mathbb{E}((W - W')^2 I_{\{0 \leq (W - W') \leq A\}} 1_{\{z \leq W \leq z + A\}}) \leq \mathbb{E}((W - W')(f(W) - f(W'))),$$

see [20]; here  $f$  is defined by  $f(x) := -1.5A$  for  $x \leq z - A$ ,  $f(x) := 1.5A$  for  $x \geq z + 2A$  and  $f(x) := x - z - A/2$  in between. Now we apply (2.5) and get

$$\mathbb{E}((W - W')^2 I_{\{0 \leq (W - W') \leq A\}} 1_{\{z \leq W \leq z + A\}}) \leq 2\lambda \mathbb{E}(Wf(W)) + 2\mathbb{E}(f(W)R) \leq 3A(\lambda + \mathbb{E}(|R|)),$$

where we used  $\mathbb{E}(|W|) \leq \sqrt{\mathbb{E}(W^2)} \leq 1$ . Similarly, we obtain

$$U_2 \geq -3A(\lambda + \mathbb{E}(R)).$$

□

**Remark 2.7.** In Theorem 2.5, we assumed  $\mathbb{E}(W^2) \leq 1$ . Alternatively, let us assume that  $\mathbb{E}(W^2)$  is finite. Then the proof of Theorem 2.5 shows, that the third and the fourth summand of the bound (2.4) change to

$$\frac{A^3}{\lambda} \left( \frac{\sqrt{2\pi}}{16} + \frac{\sqrt{\mathbb{E}(W^2)}}{4} \right) + 1.5A \mathbb{E}(|W|).$$

In the following corollary, we discuss the Kolmogorov-distance of the distribution of a random variable  $W$  to a random variable distributed according to  $N(0, \sigma^2)$ , the normal distribution with mean zero and variance  $\sigma^2$ .

**Corollary 2.8.** *Let  $\sigma^2 > 0$  and  $(W, W')$  be an exchangeable pair of real-valued random variables such that*

$$E(W'|W) = \left(1 - \frac{\lambda}{\sigma^2}\right)W + R \quad (2.10)$$

for some random variable  $R = R(W)$  and with  $0 < \lambda < 1$ . Assume that  $\mathbb{E}(W^2)$  is finite. Let  $Z_\sigma$  be a random variable distributed according to  $N(0, \sigma^2)$ . If  $|W - W'| \leq A$  for a constant  $A$ , we obtain the bound

$$\begin{aligned} d_K(W, Z_\sigma) &\leq \sqrt{\mathbb{E}\left(1 - \frac{1}{2\lambda} \mathbb{E}[(W' - W)^2 | W]\right)^2} + \left(\frac{\sigma\sqrt{2\pi}}{4} + 1.5A\right) \frac{\sqrt{\mathbb{E}(R^2)}}{\lambda} \\ &+ \frac{A^3}{\lambda} \left(\frac{\sqrt{2\pi}\sigma^2}{16} + \frac{\sqrt{\mathbb{E}(W^2)}}{4}\right) + 1.5A\sqrt{\mathbb{E}(W^2)}. \end{aligned} \quad (2.11)$$

*Proof.* Let us denote by  $f_\sigma := f_{\sigma,z}$  the solution of the Stein equation

$$f'_{\sigma,z}(x) - \frac{x}{\sigma^2} f_{\sigma,z}(x) = 1_{\{x \leq z\}}(x) - F_\sigma(z) \quad (2.12)$$

with  $F_\sigma(z) := \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^z \exp(-\frac{y^2}{2\sigma^2}) dy$ . It is easy to see that the identity  $f_{\sigma,z}(x) = \sigma f_z(\frac{x}{\sigma})$ , where  $f_z$  is the solution of the corresponding Stein equation of the standard normal distribution, holds true. Using [6, Lemma 2.2] we obtain  $0 < f_\sigma(x) < \sigma \frac{\sqrt{2\pi}}{4}$ ,  $|f'_\sigma(x)| \leq 1$ , and  $|f'_\sigma(x) - f'_\sigma(y)| \leq 1$ . With (2.10) we arrive at

$$P(W \leq z) - F_\sigma(z) = T_1 + T_2 + T_3$$

with  $T_i$ 's defined in (2.7). Using the bounds of  $f_\sigma$  and  $f'_\sigma$ , the bound of  $T_1$  is the same as in the proof of Theorem 2.5, whereas the bound of  $T_2$  changes to

$$|T_2| \leq \sigma \frac{\sqrt{2\pi}}{4\lambda} \sqrt{\mathbb{E}(R^2)}.$$

Since we consider the case  $|W - W'| \leq A$ , we have to bound

$$T_3 = -\frac{1}{2\lambda} \mathbb{E} \left( (W - W') 1_{\{|W - W'| \leq A\}} \int_{-(W - W')}^0 (f'(W + t) - f'(W)) dt \right).$$

Using the Stein identity (2.12), the mean value theorem as well as the concentration inequality-argument along the lines of the proof of Theorem 2.5, we obtain

$$|T_3| \leq \frac{A^3}{\lambda} \left( \frac{\sqrt{\mathbb{E}(W^2)}}{4} + \frac{\sigma \sqrt{2\pi}}{16} \right) + 1.5A \left( \sqrt{\mathbb{E}(W^2)} + \frac{\sqrt{\mathbb{E}(R^2)}}{\lambda} \right).$$

Hence the corollary is proved.  $\square$

With (2.10) we obtain  $\mathbb{E}(W - W')^2 = \frac{2\lambda}{\sigma^2} \mathbb{E}(W^2) - 2\mathbb{E}(W R)$ . Therefore

$$\mathbb{E} \left( 1 - \frac{1}{2\lambda} \mathbb{E}[(W' - W)^2 | W] \right) = 1 - \frac{\mathbb{E}(W^2)}{\sigma^2} + \frac{\mathbb{E}(W R)}{\lambda}, \quad (2.13)$$

so that the bound in Corollary 2.8 is only useful when  $\mathbb{E}(W^2)$  is close to  $\sigma^2$  (and  $\mathbb{E}(W R)/\lambda$  is small). An alternative bound can be obtained comparing with a  $N(0, \mathbb{E}(W^2))$ -distribution.

**Corollary 2.9.** *In the situation of Corollary 2.8, let  $Z_W$  denote the  $N(0, \mathbb{E}(W^2))$  distribution. We obtain*

$$\begin{aligned} d_K(W, Z_W) &\leq \frac{\sigma^2}{2\lambda} (\text{Var}(\mathbb{E}[(W' - W)^2 | W]))^{1/2} + \sigma^2 \left( \frac{\sqrt{\mathbb{E}(W^2)} \sqrt{2\pi}}{4} + 1.5A \right) \frac{\sqrt{\mathbb{E}(R^2)}}{\lambda} \\ &+ \sigma^2 \frac{A^3}{\lambda} \left( \frac{\sqrt{\mathbb{E}(W^2)} \sqrt{2\pi}}{16} + \frac{\sqrt{\mathbb{E}(W^2)}}{4} \right) + \sigma^2 1.5A \sqrt{\mathbb{E}(W^2)} + \sigma^2 \frac{\sqrt{\mathbb{E}(W^2)} \sqrt{\mathbb{E}(R^2)}}{\lambda}. \end{aligned} \quad (2.14)$$

*Proof.* With (2.13) we get  $\mathbb{E}(W^2) = \sigma^2 \left( \frac{1}{2\lambda} (\mathbb{E}(W - W')^2 + 2\mathbb{E}(W R)) \right)$ . With the definition of  $T_2$  and  $T_3$  as in (2.7) we obtain

$$\begin{aligned} \mathbb{E}(\mathbb{E}(W^2) f'(W) - W f(W)) &= \sigma^2 \mathbb{E} \left( \frac{\mathbb{E}(W - W')^2 + 2\mathbb{E}(W R)}{2\lambda} f'(W) \right) - \mathbb{E}(W f(W)) \\ &= \sigma^2 \mathbb{E} \left( f'(W) \left( \frac{\mathbb{E}(W - W')^2 - \mathbb{E}[(W - W')^2 | W]}{2\lambda} \right) \right) + \sigma^2 (T_2 + T_3) + \sigma^2 \frac{\mathbb{E}(W R)}{\lambda}. \end{aligned} \quad (2.15)$$

Remark that now  $\sigma^2$  in (2.10) is a parameter of the exchangeable-pair identity and no longer the parameter of the limiting distribution. We apply (2.12) and exchange every  $\sigma^2$  in (2.12) with  $\mathbb{E}(W^2)$ . Applying Cauchy-Schwarz to the first summand and bounding the other terms as in the proof of Corollary 2.8 leads to the result.  $\square$

### 3. BERRY-ESSEEN BOUNDS FOR THE CLASSICAL CURIE-WEISS MODEL

Let  $\rho$  be the symmetric Bernoulli measure and  $0 < \beta < 1$ . Then

$$W := W_n := \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i.$$

converges in distribution to a  $N(0, \sigma^2)$  with  $\sigma^2 = (1 - \beta)^{-1}$ :

*Proof of Theorem 1.2.* We consider the usual construction of an exchangeable pair. We produce a spin collection  $X' = (X'_i)_{i \geq 1}$  via a *Gibbs sampling* procedure: select a coordinate, say  $i$ , at random and replace  $X_i$  by  $X'_i$  drawn from the conditional distribution of the  $i$ 'th coordinate given  $(X_j)_{j \neq i}$ . Let  $I$  be a random variable taking values  $1, 2, \dots, n$  with equal probability, and independent of all other random variables. Consider

$$W' := W - \frac{X_I}{\sqrt{n}} + \frac{X'_I}{\sqrt{n}} = \frac{1}{\sqrt{n}} \sum_{j \neq I} X_j + \frac{X'_I}{\sqrt{n}}.$$

Hence  $(W, W')$  is an exchangeable pair and

$$W - W' = \frac{X_I - X'_I}{\sqrt{n}}.$$

Let  $\mathcal{F} := \sigma(X_1, \dots, X_n)$ . Now we obtain

$$\mathbb{E}[W - W' | \mathcal{F}] = \frac{1}{\sqrt{n}} \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i - X'_i | \mathcal{F}] = \frac{1}{n} W - \frac{1}{\sqrt{n}} \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X'_i | \mathcal{F}].$$

The conditional distribution at site  $i$  is given by

$$P_n(x_i | (x_j)_{j \neq i}) = \frac{\exp(x_i \beta m_i(x))}{\exp(\beta m_i(x)) + \exp(-\beta m_i(x))},$$

with

$$m_i(x) := \frac{1}{n} \sum_{j \neq i} x_j, \quad i = 1, \dots, n.$$

It follows that

$$\mathbb{E}[X'_i | \mathcal{F}] = \mathbb{E}[X_i | (X_j)_{j \neq i}] = \tanh(\beta m_i(X)).$$

Now  $\frac{1}{\sqrt{n}} \frac{1}{n} \sum_{i=1}^n \tanh(\beta m_i(X)) = \frac{1}{\sqrt{n}} \frac{1}{n} \sum_{i=1}^n (\tanh(\beta m_i(X)) - \tanh(\beta m(X))) + \frac{1}{\sqrt{n}} \tanh(\beta m(X)) =: R_1 + R_2$  with  $m(X) := \frac{1}{n} \sum_{i=1}^n X_i$ . Taylor-expansion  $\tanh(x) = x + \mathcal{O}(x^3)$  leads to

$$R_2 = \frac{1}{\sqrt{n}} \beta m(X) + \frac{1}{\sqrt{n}} \mathcal{O}(m(X)^3) = \frac{\beta}{n} W + \mathcal{O}\left(\frac{W^3}{n^2}\right).$$

Hence

$$\mathbb{E}[W - W'|W] = \frac{1-\beta}{n}W + R = \frac{\lambda}{\sigma^2}W + R \quad (3.1)$$

with  $\lambda := \frac{1}{n}$ ,  $\sigma^2 := (1-\beta)^{-1}$  and  $R := \mathcal{O}\left(\frac{W^3}{n^2}\right) - R_1$ . Since  $|W - W'| = \left|\frac{X_I - X'_I}{\sqrt{n}}\right| \leq \frac{1}{\sqrt{n}} =: A$ , we are able to apply Corollary 2.8. From Lemma 1.13 we know that for  $\varrho$  being the symmetric Bernoulli distribution and for  $0 < \beta < 1$  we have  $\mathbb{E}(W^4) \leq \text{const.}$ . Applying this it follows that the fourth term in (2.11) can be bounded by  $1.5A \frac{\sqrt{\mathbb{E}(W^2)}}{\sigma^2} \leq \frac{(1-\beta)\text{const.}}{\sqrt{n}}$ , and the third summand in (2.11) can be estimated as follows:

$$\frac{A^3}{\lambda} \left( \frac{\sqrt{2\pi}}{16} \sqrt{(1-\beta)} + \frac{\text{const.}}{4}(1-\beta) \right) \leq \frac{1}{\sqrt{n}} \sqrt{(1-\beta)} \text{const.}$$

Moreover we obtain  $\mathbb{E}|R| \leq \mathbb{E}|R_1| + \mathcal{O}\left(\frac{\mathbb{E}|W^3|}{n^2}\right)$ . Since  $\tanh(x)$  is 1-Lipschitz we obtain  $|R_1| \leq \frac{1}{\sqrt{n}}|m_i(X) - m(X)| \leq \frac{1}{n^{3/2}}$ . Therefore, with Lemma 1.13, we get  $\mathbb{E}|R| = \mathcal{O}\left(\frac{1}{n^{3/2}}\right)$  and thus, the second summand in (2.11) can be bounded by

$$\text{const.} \left( \frac{\sqrt{2\pi}}{4\sqrt{(1-\beta)}} + 1.5\frac{1}{\sqrt{n}} \right) \frac{1}{\sqrt{n}} = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right).$$

To bound the first summand in (2.11), we obtain  $(W - W')^2 = \frac{X_I^2}{n} - \frac{2X_I X'_I}{n} + \frac{X'^2_I}{n}$ . Hence

$$\mathbb{E}[(W - W')^2|\mathcal{F}] = \frac{2}{n} - \frac{2}{n^2} \sum_{i=1}^n X_i \tanh(\beta m_i(X)),$$

and therefore

$$\begin{aligned} 1 - \frac{1}{2\lambda} \mathbb{E}[(W - W')^2|\mathcal{F}] &= \frac{1}{n} \sum_{i=1}^n X_i \tanh(\beta m_i(X)) \\ &= \frac{1}{n} \sum_{i=1}^n X_i (\tanh(\beta m_i(X)) - \tanh(\beta m(X))) + m(X) \tanh(\beta m(X)) \\ &=: R_1 + R_2. \end{aligned}$$

By Taylor expansion we get  $R_2 = \frac{\beta}{n}W^2 + \mathcal{O}\left(\frac{W^4}{n^2}\right)$  and using Lemma 1.13 we obtain  $\mathbb{E}|R_2| = \mathcal{O}(n^{-1})$ . Since  $\tanh(x)$  is 1-Lipschitz we obtain  $|R_1| \leq \frac{1}{n}$ . Hence  $\mathbb{E}|R_1 + R_2| = \mathcal{O}(n^{-1})$  and Theorem 1.2 is proved.  $\square$

Now we discuss the critical case  $\beta = 1$ , when  $\varrho$  is the symmetric Bernoulli distribution. For  $\beta = 1$ , using the Taylor expansion  $\tanh(x) = x - x^3/3 + \mathcal{O}(x^5)$ , (3.1) would lead to

$$\mathbb{E}[W - W'|W] = \frac{W^3}{3} \frac{1}{n^2} + \tilde{R}$$

for some  $\tilde{R}$ . Hence it is no longer possible to apply Corollary 2.8. Moreover the prefactor  $\lambda := \frac{1}{n^2}$  would give growing bounds. In other words, the criticality of the temperature value  $1/\beta_c = 1$  can

also be recognized by Stein's method. We already know that at the critical value, the sum of the spin-variables has to be rescaled. Let us now define

$$W := \frac{1}{n^{3/4}} \sum_{i=1}^n X_i. \quad (3.2)$$

Constructing the exchangeable pair  $(W, W')$  in the same manner as before we will obtain

$$\mathbb{E}[W - W'|W] = \frac{1}{n^{3/2}} \frac{W^3}{3} + R(W) =: -\lambda\psi(W) + R(W). \quad (3.3)$$

with  $\lambda = \frac{1}{n^{3/2}}$  and a remainder  $R(W)$  presented later. Considering the density  $p(x) = C \exp(-x^4/12)$ , we have

$$\frac{p'(x)}{p(x)} = \psi(x).$$

This is the starting point for developing Stein's method for limiting distributions with a regular Lebesgue-density  $p(\cdot)$  and an exchangeable pair  $(W, W')$  which satisfies the condition

$$\mathbb{E}[W - W'|W] = -\lambda\psi(W) + R(W) = -\lambda \frac{p'(W)}{p(W)} + R(W)$$

with  $0 < \lambda < 1$ . To prove (3.3), observe that

$$\mathbb{E}[W - W'|W] = \frac{1}{n}W - \frac{1}{n^{3/4}} \frac{1}{n} \sum_{i=1}^n \tanh(m_i(X)).$$

By Taylor expansion and the identity  $m_i(X) = m(X) - \frac{X_i}{n}$  we obtain

$$\frac{1}{n^{3/4}} \frac{1}{n} \sum_{i=1}^n \tanh(m_i(X)) = \frac{1}{n}W - \frac{1}{n^{3/2}} \frac{W^3}{3} - R(W)$$

with  $R(W)$  such that  $\mathbb{E}|R(W)| = \mathcal{O}(n^{-2})$ . The exact form of  $R(W)$  will be presented in Section 5.

#### 4. THE EXCHANGEABLE PAIR APPROACH FOR DISTRIBUTIONAL APPROXIMATIONS

Motivated by the classical Curie-Weiss model at the critical temperature, we will develop Stein's method with the help of exchangeable pairs as follows. For a rather large class of continuous distributions, the Stein characterization was introduced in [23], following the lines of [22, Chapter 6]. The densities occurring as limit laws in models of statistical mechanics belong to this class. Let  $I$  be a real interval, where  $-\infty \leq a < b \leq \infty$ . A function is called *regular* if  $f$  is finite on  $I$  and, at any interior point of  $I$ ,  $f$  possesses a right-hand limit and a left-hand limit. Further,  $f$  possesses a right-hand limit  $f(a+)$  at the point  $a$  and a left-hand limit  $f(b-)$  at the point  $b$ .

Let us assume, that the regular density  $p$  satisfies the following condition:

**Assumption (D)** Let  $p$  be a regular, strictly positive density on an interval  $I = [a, b]$ . Suppose  $p$  has a derivative  $p'$  that is regular on  $I$  and has only countably many sign changes and being continuous at the sign changes. Suppose moreover that  $\int_I p(x) |\log(p(x))| dx < \infty$  and assume that

$$\psi(x) := \frac{p'(x)}{p(x)} \quad (4.1)$$

is regular.

In [23, Proposition] it is proved, that a random variable  $Z$  is distributed according to the density  $p$  if and only if

$$\mathbb{E}(f'(Z) + \psi(Z) f(Z)) = f(b-)p(b-) - f(a+)p(a+)$$

for a suitably chosen class  $\mathcal{F}$  of functions  $f$ . The proof is integration by parts. The corresponding Stein identity is

$$f'(x) + \psi(x) f(x) = h(x) - P(h), \quad (4.2)$$

where  $h$  is a measurable function for which  $\int_I |h(x)| p(x) dx < \infty$ ,  $P(x) := \int_{-\infty}^x p(y) dy$  and  $P(h) := \int_I h(y) p(y) dy$ . The solution  $f := f_h$  of this differential equation is given by

$$f(x) = \frac{\int_a^x (h(y) - Ph) p(y) dy}{p(x)}. \quad (4.3)$$

For the function  $h(x) := 1_{\{x \leq z\}}(x)$  let  $f_z$  be the corresponding solution of (4.2). We will make the following assumptions:

**Assumption (B1)** Let  $p$  be a density fulfilling Assumption (D). We assume that for any absolute continuous function  $h$ , the solution  $f_h$  of (4.2) satisfies

$$\|f_h\| \leq c_1 \|h'\|, \quad \|f'_h\| \leq c_2 \|h'\| \quad \text{and} \quad \|f''_h(x)\| \leq c_3 \|h'\|,$$

where  $c_1, c_2$  and  $c_3$  are constants.

**Assumption (B2)** Let  $p$  be a density fulfilling Assumption (D) We assume that the solution  $f_z$  of

$$f'_z(x) + \psi(x) f_z(x) = 1_{\{x \leq z\}}(x) - P(z) \quad (4.4)$$

satisfies

$$|f_z(x)| \leq d_1, \quad |f'_z(x)| \leq d_2 \quad \text{and} \quad |f'_z(x) - f'_z(y)| \leq d_3$$

and

$$|(\psi(x) f_z(x))'| = \left| \left( \frac{p'(x)}{p(x)} f_z(x) \right)' \right| \leq d_4 \quad (4.5)$$

for all real  $x$  and  $y$ , where  $d_1, d_2, d_3$  and  $d_4$  are constants.

At first glance, Condition (4.5) seem to be a rather strong or at least a rather technical condition.

**Remark 4.1.** In the case of the normal approximation,  $\psi(x) = -x$ , we have to bound  $(x f_z(x))'$  for the solution  $f_z$  of the classical Stein equation. But it is easy to observe that  $|(x f'_z(x))'| \leq 2$  by direct calculation (see [6, Proof of Lemma 6.5]). However, in the normal approximation case, this bound would lead to a worse Berry-Esseen constant (compare Theorem 2.5 with Theorem 4.6). Hence in this case we only use  $d_2 = d_3 = 1$  and  $d_1 = \sqrt{2\pi}/4$ .

We will see, that for all distributions appearing as limit laws in our class of Curie-Weiss models, Condition (4.5) can be proved:

**Lemma 4.2.** *The densities  $f_{k,\mu,\beta}$  in (1.8) and (1.9) and the densities in Theorem 1.4, Theorem 1.7 and Theorem 1.8 satisfy Assumptions (D), (B1) and (B2).*

*Proof.* We defer the proofs to the appendix, since they only involve careful analysis.  $\square$

**Remark 4.3.** With respect to all densities which appear as limiting distributions in our theorems, we restrict ourselves to bound solutions (and its derivatives) of the corresponding Stein equation characterizing distributions with probability densities  $p$  of the form  $b_k \exp(-a_k x^{2k})$ . Along the lines of the proof of Lemma 4.2, one would be able to present good bounds (in the sense that Assumption (B1) and (B2) are fulfilled) even for measures with a probability density of the form

$$p(x) = b_k \exp(-a_k V(x)), \quad (4.6)$$

where  $V$  is even, twice continuously differentiable, unbounded above at infinity,  $V' \neq 0$  and  $V'$  and  $1/V'$  are increasing on  $[0, \infty)$ . Moreover one has to assume that  $\frac{V''(x)}{|V'(x)|}$  can be bounded by a constant for  $x \geq d$  with some  $d \in \mathbb{R}_+$ . We sketch the proof in the appendix. It is remarkable, that this class of measures is a subclass of measures which are GHS, see Section 7. A measure with density  $p$  in (4.6) is usually called a Gibbs measure. Stein's method for discrete Gibbs measures is developed in [8]. Our remark might be of use applying Stein's method for some continuous Gibbs measure approximation.

**Remark 4.4.** In the case of comparing with an *exponential distribution* with parameter  $\mu$ , it is easy to see, that Assumption (D) and (B2) is fulfilled, see [23, Example 1.6] for (D) and [4, Lemma 2.1] for (B2). We have  $\psi(x) = -\mu$  and  $\|f_z\| \leq 1$ ,  $\|f'_z\| \leq 1$  and  $\sup_{x,y \geq 0} |f'_z(x) - f'_z(y)| \leq 1$ . Thus  $|(\psi(x)f_z(x))'| = \mu|f'_z(x)| \leq \mu$ .

**Remark 4.5.** From (4.3) we obtain

$$f_z(x) = \frac{(1 - P(z))P(x)}{p(x)} \quad \text{for } x \leq z$$

and

$$f_z(x) = \frac{P(z)(1 - P(x))}{p(x)} \quad \text{for } x \geq z.$$

Hence

$$\psi(x)f_z(x) = \frac{P(x)p'(x)}{p^2(x)}(1 - P(z)) \quad \text{for } x \leq z$$

and

$$\psi(x)f_z(x) = \frac{(1 - P(x))p'(x)}{p^2(x)}(P(z)) \quad \text{for } x \geq z.$$

Therefore one has to bound the derivative of

$$\frac{P(x)p'(x)}{p^2(x)} \quad \text{and} \quad \frac{(1 - P(x))p'(x)}{p^2(x)},$$

respectively, to check Condition (4.5).

The following result is a refinement of Stein's result [22] for exchangeable pairs.

**Theorem 4.6.** *Let  $p$  be a density fulfilling Assumption (D). Let  $(W, W')$  be an exchangeable pair of real-valued random variables such that*

$$\mathbb{E}[W'|W] = W + \lambda\psi(W) - R(W) \quad (4.7)$$

for some random variable  $R = R(W)$ ,  $0 < \lambda < 1$  and  $\psi$  defined in (4.1). Then

$$\mathbb{E}(W - W')^2 = -2\lambda\mathbb{E}[W\psi(W)] + 2\mathbb{E}[WR(W)]. \quad (4.8)$$

We obtain the following assertions:

- (1) *Let  $Z$  be a random variable distributed according to  $p$ . Under Assumption (B1), for any uniformly Lipschitz function  $h$ , we obtain*

$$|\mathbb{E}h(W) - \mathbb{E}h(Z)| \leq \delta\|h'\|$$

with

$$\delta := c_2\mathbb{E}\left|1 - \frac{1}{2\lambda}\mathbb{E}[(W - W')^2|W]\right| + \frac{c_3}{4\lambda}\mathbb{E}|W - W'|^3 + \frac{c_1}{\lambda}\sqrt{\mathbb{E}(R^2)}.$$

- (2) *Let  $Z$  be a random variable distributed according to  $p$ . Under Assumption (B2), we obtain for any  $A > 0$*

$$\begin{aligned} d_K(W, Z) &\leq d_2\sqrt{\mathbb{E}\left(1 - \frac{1}{2\lambda}\mathbb{E}[(W' - W)^2|W]\right)^2} + (d_1 + \frac{3}{2}A)\frac{\sqrt{\mathbb{E}(R^2)}}{\lambda} \\ &+ \frac{1}{\lambda}\left(\frac{d_4A^3}{4}\right) + \frac{3A}{2}\mathbb{E}(|\psi(W)|) + \frac{d_3}{2\lambda}\mathbb{E}((W - W')^2\mathbf{1}_{\{|W - W'| \geq A\}}). \end{aligned} \quad (4.9)$$

With (4.8) we obtain

$$\mathbb{E}\left(1 - \frac{1}{2\lambda}\mathbb{E}[(W - W')^2|W]\right) = 1 + \mathbb{E}[W\psi(W)] - \frac{\mathbb{E}(WR)}{\lambda}.$$

Therefore the bounds in Theorem 4.6 are unlikely to be useful unless  $-\mathbb{E}[W\psi(W)]$  is close to 1 and  $\frac{\mathbb{E}(WR)}{\lambda}$  is small. Alternatively bounds can be obtained comparing not with a distribution given by  $p$  but with a modification which involves  $\mathbb{E}[W\psi(W)]$ . Let  $p_W$  be a probability density such that a random variable  $Z$  is distributed according to  $p_W$  if and only if

$$\mathbb{E}(\mathbb{E}[W\psi(W)]f'(Z) + \psi(Z)f(Z)) = 0$$

for a suitably chosen class of functions.

**Theorem 4.7.** *Let  $p$  be a density fulfilling Assumption (D). Let  $(W, W')$  be an exchangeable pair of real-valued random variables such that (4.7) holds. If  $Z_W$  is a random variable distributed according to  $p_W$ , we obtain under (B1), for any uniformly Lipschitz function  $h$  that  $|\mathbb{E}h(W) - \mathbb{E}h(Z_W)| \leq \delta'\|h'\|$  with*

$$\delta' := \frac{c_2}{2\lambda}(\text{Var}(\mathbb{E}[(W - W')^2|W]))^{1/2} + \frac{c_3}{4\lambda}\mathbb{E}|W - W'|^3 + \frac{c_1 + c_2\sqrt{\mathbb{E}(W^2)}}{\lambda}\sqrt{\mathbb{E}(R^2)}.$$

Under Assumption (B2) we obtain for any  $A > 0$

$$\begin{aligned} d_K(W, Z_W) &\leq \frac{d_2}{2\lambda} (\text{Var}(\mathbb{E}[(W - W')^2|W]))^{1/2} + (d_1 + d_2\sqrt{\mathbb{E}(W^2)} + \frac{3}{2}A) \frac{\sqrt{\mathbb{E}(R^2)}}{\lambda} \\ &\quad + \frac{1}{\lambda} \left( \frac{d_4 A^3}{4} \right) + \frac{3A}{2} \mathbb{E}(|\psi(W)|) + \frac{d_3}{2\lambda} \mathbb{E}((W - W')^2 1_{\{|W - W'| \geq A\}}). \end{aligned} \quad (4.10)$$

*Proof of Theorem 4.6.* Interestingly enough, the proof is a quite simple adaption of the results in [22] and follows the lines of the proof of Theorem 2.5. For a function  $f$  with  $|f(x)| \leq C(1 + |x|)$  we obtain

$$\begin{aligned} 0 &= \mathbb{E}((W - W')(f(W') + f(W))) \\ &= \mathbb{E}((W - W')(f(W') - f(W))) - 2\lambda \mathbb{E}(\psi(W) f(W)) + 2\mathbb{E}(f(W) R(W)), \end{aligned} \quad (4.11)$$

which is equivalent to

$$\mathbb{E}(\psi(W) f(W)) = -\frac{1}{2\lambda} \mathbb{E}((W - W')(f(W) - f(W'))) + \frac{1}{\lambda} \mathbb{E}(f(W) R(W)) \quad (4.12)$$

**Proof of (1):** Now let  $f = f_h$  be the solution of the Stein equation (4.2), and define

$$\widehat{K}(t) := (W - W')(1_{\{-(W - W') \leq t \leq 0\}} - 1_{\{0 < t \leq -(W - W')\}}) \geq 0.$$

By (4.12), following the calculations on page 21 in [6], we simply obtain

$$\begin{aligned} |\mathbb{E}h(W) - \mathbb{E}h(Z)| &= |\mathbb{E}(f'(W) + \psi(W) f(W))| \\ &= |\mathbb{E}(f'(W) \left(1 - \frac{1}{2\lambda} (W - W')^2\right) + \frac{1}{2\lambda} \mathbb{E} \left( \int_{\mathbb{R}} (f'(W) - f'(W + t)) \widehat{K}(t) dt \right) \\ &\quad + \frac{1}{\lambda} \mathbb{E}(f(W) R(W))|. \end{aligned}$$

Using  $\int_{\mathbb{R}} |t| \widehat{K}(t) dt = \frac{1}{2} \mathbb{E}|W - W'|^3$ , the bounds in Assumption (B1) give:

$$|\mathbb{E}h(W) - \mathbb{E}h(Z)| \leq \|h'\| \left( c_2 \mathbb{E} \left| 1 - \frac{1}{2\lambda} \mathbb{E}((W - W')^2|W) \right| + \frac{c_3}{4\lambda} \mathbb{E}|W - W'|^3 + \frac{c_1}{\lambda} \sqrt{\mathbb{E}(R^2)} \right). \quad (4.13)$$

**Proof of (2):** Now let  $f = f_z$  be the solution of the Stein equation (4.4). As in (2.7), using (4.12), we obtain

$$\begin{aligned} P(W \leq z) - P(z) &= \mathbb{E}(f'(W) + \psi(W) f(W)) \\ &= \mathbb{E} \left( f'(W) \left(1 - \frac{1}{2\lambda} (W - W')^2\right) + \frac{1}{2\lambda} \mathbb{E}(2f(W) R) \right. \\ &\quad \left. - \frac{1}{2\lambda} \mathbb{E} \left[ (W - W')(f(W) - f(W') - (W - W') f'(W)) \right] \right) \\ &= T_1 + T_2 + T_3. \end{aligned}$$

Now the bounds in Assumption (B2) give

$$|T_1| \leq d_2 \sqrt{\mathbb{E} \left( 1 - \frac{1}{\lambda} \mathbb{E}[(W' - W)^2|W] \right)^2}$$

and

$$|T_2| \leq \frac{d_1}{\lambda} \sqrt{\mathbb{E}(R^2)}.$$

Using the decomposition (2.8) of  $(-2\lambda)T_3$ , the modulus of the first term can be bounded by  $d_3 \mathbb{E}((W - W')^2 1_{\{|W - W'| > A\}})$ . Using the Stein identity (4.4), the second summand can be represented as

$$\begin{aligned} & \mathbb{E}\left((W - W') 1_{\{|W - W'| \leq A\}} \int_{-(W - W')}^0 (-\psi(W + t) f(W + t) + \psi(W) f(W)) dt\right) \\ & + \mathbb{E}\left((W - W') 1_{\{|W - W'| \leq A\}} \int_{-(W - W')}^0 (1_{\{W + t \leq z\}} - 1_{\{W \leq z\}}) dt\right) =: U_1 + U_2. \end{aligned}$$

With  $g(x) := (\psi(x)f(x))'$  we obtain

$$-\psi(W + t) f(W + t) + \psi(W) f(W) = - \int_0^t g(W + s) ds.$$

Since  $|g(x)| \leq d_4$  we obtain  $|U_1| \leq \frac{A^3}{2} d_4$ .

Analogously to the steps in the proof of Theorem 2.5,  $U_2$  can be bounded by

$$\mathbb{E}((W - W')(f(W) - f(W'))) = 2\mathbb{E}(f(W) R(W)) - 2\lambda \mathbb{E}(\psi(W) f(W)),$$

where we applied (4.12), and where  $f$  is defined by  $f(x) := -1.5A$  for  $x \leq z - A$ ,  $f(x) := 1.5A$  for  $x \geq z + 2A$  and  $f(x) := x - z - A/2$  in between. Thus  $U_2 \leq 3A(\mathbb{E}(|R|) + \lambda \mathbb{E}(|\psi(W)|))$ . Similarly we obtain  $U_2 \geq -3A(\mathbb{E}(|R|) + \lambda \mathbb{E}(|\psi(W)|))$ .  $\square$

*Proof of Theorem 4.7.* The main observation is the following identity:

$$\begin{aligned} \mathbb{E}(-\mathbb{E}[W\psi(W)] f'(W) + \psi(W)f(W)) &= \mathbb{E}\left(f'(W) \left(\frac{\mathbb{E}[(W - W')^2] - 2\mathbb{E}[WR]}{2\lambda}\right)\right) + \mathbb{E}(\psi(W) f(W)) \\ &= \mathbb{E}\left(f'(W) \left(\frac{\mathbb{E}[(W - W')^2] - \mathbb{E}[(W - W')^2|W]}{2\lambda}\right)\right) + \frac{1}{\lambda} \left(\mathbb{E}[f(W) R] - \mathbb{E}[\mathbb{E}(WR) f'(W)]\right) + T_3 \end{aligned}$$

with  $T_3$  defined as in the proof of Theorem 4.6. Now we can apply the Cauchy-Schwarz inequality to get

$$\mathbb{E}|\mathbb{E}[(W - W')^2] - \mathbb{E}[(W - W')^2|W]| \leq (\text{Var}(\mathbb{E}[(W - W')^2|W]))^{1/2}.$$

Now the proof follows the lines of the proof of Theorem 4.6.  $\square$

**Remark 4.8.** We discuss an alternative bound in Theorem 4.6 in the case that  $(\psi(x)f_z(x))'$  cannot be bounded uniformly. By the mean value theorem we obtain in general

$$-\psi(W + t)f(W + t) + \psi(W)f(W) = \psi(W) \left(-\int_0^1 f'(W + st) t ds\right) + f(W + t) \left(-\int_0^1 \psi'(W + st) t ds\right).$$

This gives

$$|-\psi(W + t)f(W + t) + \psi(W)f(W)| \leq d_2 |\psi(W)| |t| + d_1 \int_0^1 |\psi'(W + st)| |t| ds.$$

Now we get the bound

$$\frac{1}{2\lambda} |U_1| \leq \frac{d_2 A^3}{4\lambda} \mathbb{E}(|\psi(W)|) + \frac{d_1}{2\lambda} \mathbb{E}(V)$$

with

$$V := \left(|W - W'| 1_{\{|W - W'| \leq A\}} \int_{-(W - W')}^0 \int_0^1 |\psi'(W + st)| |t| ds dt\right).$$

Let us consider the example  $\psi(x) = -x^3/3$ . Now

$$\psi'(W + st) = |(W + st)^2| = |W^2 + 2stW + s^2t^2|,$$

hence

$$|(W + st)^2||t| \leq |t||W^2| + |t^2|2|W||s| + |t^3||s^2|$$

and integration over  $s$  gives

$$\int_0^1 |\psi'(W + st)||t|ds \leq |t||W^2| + |t^2||W| + |t^3|/3.$$

Integration over  $t$  leads to

$$\int_{-(W-W')}^0 \left( \int_0^1 |\psi'(W + st)||t|ds \right) \leq \frac{|\Delta|^2}{2}|W^2| + \frac{|\Delta|^3}{3}|W| + \frac{|\Delta|^4}{12}.$$

with  $\Delta := (W - W')$ . Hence we get

$$\mathbb{E}(V) \leq \left( \frac{A^3}{2}\mathbb{E}|W^2| + \frac{A^4}{3}\mathbb{E}|W| + \frac{A^5}{12} \right).$$

We will see in Section 5, that this bound is good enough for an alternative proof of Theorem 1.3.

## 5. BERRY-ESSEEN BOUND AT THE CRITICAL TEMPERATURE

*Proof of Theorem 1.3.* We start with (3.3), where  $W$  is given by (3.2). We will calculate the remainder term  $R(W)$  more carefully: By Taylor expansion and the identities  $m_i(X) = m(X) - X_i/n$  and  $m(X) = \frac{1}{n^{1/4}}W$  we obtain

$$\frac{1}{n^{3/4}} \frac{1}{n} \sum_{i=1}^n \tanh(m_i(X)) = \frac{1}{n}W - \frac{1}{n^{3/2}} \frac{W^3}{3} - \mathcal{O}\left(\frac{W}{n^2}\right) + \mathcal{O}\left(\frac{W^3}{n^{5/2}}\right) + \mathcal{O}(S(W))$$

with

$$S(W) = \frac{1}{n^{3/4}} \frac{1}{n} \sum_{i=1}^n m_i(X)^5 = \mathcal{O}\left(\frac{W^5}{n^2}\right) + \mathcal{O}\left(\frac{W^3}{n^{7/2}}\right) + \mathcal{O}\left(\frac{W^2}{n^{21/4}}\right) + \mathcal{O}\left(\frac{W}{n^6}\right).$$

From Lemma 1.13 we know that for  $\varrho$  being the symmetric Bernoulli distribution and  $\beta = 1$  we get  $\mathbb{E}|W|^6 \leq \text{const.}$ . Using this we get the exchangeable pair identity (3.3) with  $R(W) = \mathcal{O}\left(\frac{1}{n^2}\right)$ . With Lemma 4.2, we can now apply Theorem 4.6, using  $|W - W'| \leq \frac{1}{n^{3/4}} =: A$ . We obtain  $1.5A \mathbb{E}(|\psi(W)|) \leq \text{const.} \frac{1}{n^{3/4}}$  and  $\frac{d_4 A^3}{4\lambda} = \frac{d_4}{4} \frac{1}{n^{3/4}}$ . Using  $\mathbb{E}|R(W)| \leq \text{const.} \frac{1}{n^2}$  we get

$$\left(d_1 + \frac{3}{2}A\right) \frac{\mathbb{E}|R(W)|}{\lambda} \leq \text{const.} \frac{1}{\sqrt{n}}.$$

Moreover we obtain

$$\mathbb{E}[(W - W')^2 | \mathcal{F}] = \frac{2}{n^{3/2}} - \frac{2}{n^{5/2}} \sum_{i=1}^n X_i \tanh(m_i(X)).$$

Hence applying Theorem 4.6 we have to bound the expectation of

$$T := \left| \frac{1}{n} \sum_{i=1}^n X_i \tanh(m_i(X)) \right|.$$

Again using Taylor and  $m_i(X) = m(X) - \frac{X_i}{n}$  and Lemma 1.13, the leading term of  $T$  is  $\frac{W^2}{n^{1/2}}$ . Hence  $\mathbb{E}(T) = \mathcal{O}(n^{-1/2})$  and Theorem 1.3 is proved.  $\square$

**Remark 5.1.** In Remark 4.8, we presented an alternative bound via Stein's method without proving a uniform bound for  $(\psi'(x)f_z(x))'$ . As we can see, the additional terms in this bound are of smaller order than  $\mathcal{O}(n^{-1/2})$ , using  $A = n^{-3/4}$ .

*Proof of Theorem 1.4.* (1) Let  $\beta_n - 1 = \frac{\gamma}{\sqrt{n}}$  and  $W = S_n/n^{3/4}$ . For the distribution function  $F_\gamma$  in Theorem 1.4 we obtain  $\psi(x) = \gamma x - \frac{1}{3}x^3$ . Moreover we have

$$\mathbb{E}[W - W'|W] = \frac{1 - \beta_n}{n}W + \frac{\beta_n^3}{n^{3/2}}\frac{W^3}{3} + R(\beta_n, W) \quad (5.1)$$

with  $R(\beta_n, W) = \mathcal{O}(n^{-2})$ . With  $\beta_n - 1 = \frac{\gamma}{\sqrt{n}}$  we obtain

$$\mathbb{E}[W - W'|W] = -\frac{\gamma}{n^{3/2}}W + \frac{\beta_n^3}{n^{3/2}}\frac{W^3}{3} + R(\beta_n, W) = -\frac{1}{n^{3/2}}\psi(W) + \tilde{R}(\beta_n, W)$$

with  $\tilde{R}(\beta_n, W) = \mathcal{O}(n^{-2})$ . Now we only have to adapt the proof of Theorem 1.3 step by step, using, that the sixth moment of  $W$  is bounded for varying  $\beta_n$ , see Lemma 1.13. Hence by Lemma 4.2 and Theorem 4.6, part (1) is proved.

(2): we consider the case  $|\beta_n - 1| = \mathcal{O}(n^{-1})$  and  $W = S_n/n^{3/4}$ . Now in (5.1), the term  $\frac{1 - \beta_n}{n}W$  will be a part of the remainder:

$$\mathbb{E}[W - W'|W] = \frac{\beta_n^3}{n^{3/2}}\frac{W^3}{3} + R(\beta_n, W) + \frac{1 - \beta_n}{n}W =: -\frac{\beta_n^3}{n^{3/2}}\psi(W) + \hat{R}(\beta_n, W)$$

with  $\psi(x) := x^3/3$ . Along the lines of the proof of Theorem 1.3, we have to bound  $\frac{\mathbb{E}|\hat{R}|}{\lambda}$  with  $\lambda = \frac{\beta_n^3}{n^{3/2}}$ . But since by assumption

$$\lim_{n \rightarrow \infty} \frac{1}{\lambda} \frac{1 - \beta_n}{n} = \frac{\sqrt{n}(1 - \beta_n)}{\beta_n^3} = 0,$$

applying Theorem 4.6, we obtain the convergence in distribution for any  $\beta_n$  with  $|\beta_n - 1| \ll n^{-1/2}$ , and we obtain the Berry-Esseen bound of order  $\mathcal{O}(1/\sqrt{n})$  for any  $|\beta_n - 1| = \mathcal{O}(n^{-1})$ .

(3) Finally we consider  $|\beta_n - 1| \gg n^{-1/2}$  and  $W = \sqrt{\frac{1 - \beta_n}{n}}S_n$ . Now we obtain

$$\mathbb{E}[W - W'|W] = \frac{1 - \beta_n}{n}W + \frac{\beta_n W}{n^2} + \frac{\beta_n^3}{n^2(1 - \beta_n)}\frac{W^3}{3} + R(\beta_n, W) =: -\lambda\psi(W) + \tilde{R}(\beta_n, W)$$

with  $\lambda = \frac{1 - \beta_n}{n}$  and  $\psi(x) = -x$ . We apply Corollary 2.8: with  $A = \frac{1}{\sqrt{n}}(1 - \beta_n)^{1/2}$ , one obtains  $\lambda^{-1}A^3 = n^{-1/2}(1 - \beta_n)^{1/2}$  and

$$\frac{\mathbb{E}|\tilde{R}(\beta_n, W)|}{\lambda} \leq \frac{\text{const}}{n(1 - \beta_n)^2}.$$

Moreover

$$\mathbb{E}[(W - W')^2 | W] = \frac{2(1 - \beta_n)}{n} - \frac{2(1 - \beta_n)}{n} \frac{1}{n} \sum_{i=1}^n X_i \tanh(\beta_n m_i(X)).$$

Hence

$$\left| 1 - \frac{1}{2\lambda} \mathbb{E}[(W - W')^2 | W] \right| = \left| \frac{\beta_n}{n(1 - \beta_n)} W^2 - \frac{\beta_n}{n} - \frac{\beta_n^3}{n^2(1 - \beta_n)^2} \frac{W^4}{3} + R(\beta_n, W) \right| = \mathcal{O}\left(\frac{\beta_n}{n(1 - \beta_n)}\right).$$

Hence with  $|\beta_n - 1| \gg n^{-1/2}$  we obtain convergence in distribution. Under the additional assumption  $|\beta_n - 1| \gg n^{-1/4}$  we obtain the Berry-Esseen result.  $\square$

## 6. PROOF OF THE GENERAL CASE

*Proof of Theorem 1.7.* Given  $\varrho$  which satisfies the GHS-inequality and let  $\alpha$  be the global minimum of type  $k$  and strength  $\mu(\alpha)$  of  $G_\varrho$ . In case  $k = 1$  it is known that the random variable  $\frac{S_n}{\sqrt{n}}$  converges in distribution to a normal distribution  $N(0, \sigma^2)$  with  $\sigma^2 = \mu(\alpha)^{-1} - \beta^{-1} = (\sigma_\varrho^{-2} - \beta)^{-1}$ , see for example [10, V.13.15]. Hence in this case we will apply Corollary 2.9 (to obtain better constants for our Berry-Esseen bound in comparison to Theorem 4.7).

Consider  $k \geq 1$ . We just treat the case  $\alpha = 0$  and denote  $\mu = \mu(0)$ . The more general case can be done analogously. For  $k = 1$ , we consider  $\psi(x) = -\frac{x}{\sigma^2}$  with  $\sigma^2 = \mu^{-1} - \beta^{-1}$ . For any  $k \geq 2$  we consider

$$\psi(x) = -\frac{\mu}{(2k - 1)!} x^{2k-1}.$$

We define

$$W := W_{k,n} := \frac{1}{n^{1-1/(2k)}} \sum_{i=1}^n X_i$$

and  $W'$ , constructed as in Section 3, such that

$$W - W' = \frac{X_I - X'_I}{n^{1-1/(2k)}}.$$

We obtain

$$\mathbb{E}[W - W' | \mathcal{F}] = \frac{1}{n} W - \frac{1}{n^{1-1/(2k)}} \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X'_i | \mathcal{F}).$$

Now we have to calculate the conditional distribution at site  $i$  in the general case:

**Lemma 6.1.** *In the situation of Theorem 1.7, if  $X_1$  is  $\varrho$ -a.s. bounded, we obtain*

$$\mathbb{E}(X'_i | \mathcal{F}) = (m_i(X) - \frac{1}{\beta} G'_\varrho(\beta, m_i(X))) (1 + \mathcal{O}(1/n))$$

with  $m_i(X) := \frac{1}{n} \sum_{j \neq i} X_j = m(X) - \frac{X_i}{n}$ .

*Proof.* We compute the conditional density  $g_\beta(x_1|(X_i)_{i \geq 2})$  of  $X_1 = x_1$  given  $(X_i)_{i \geq 2}$  under the Curie-Weiss measure:

$$\begin{aligned} g_\beta(x_1|(X_i)_{i \geq 2}) &= \frac{e^{\beta/2n(\sum_{i \geq 2} x_1 X_i + \sum_{i \neq j \geq 2} X_i X_j + x_1^2)}}{\int e^{\beta/2n(\sum_{i \geq 2} x_1 X_i + \sum_{i \neq j \geq 2} X_i X_j + x_1^2)} \varrho(dx_1)} \\ &= \frac{e^{\beta/2n(\sum_{i \geq 2} x_1 X_i + x_1^2)}}{\int e^{\beta/2n(\sum_{i \geq 2} x_1 X_i + x_1^2)} \varrho(dx_1)}. \end{aligned}$$

Hence we can compute  $\mathbb{E}[X_1'|\mathcal{F}]$  as

$$\mathbb{E}[X_1'|\mathcal{F}] = \frac{\int x_1 e^{\beta/2n(\sum_{i \geq 2} x_1 X_i + x_1^2)} \varrho(dx_1)}{\int e^{\beta/2n(\sum_{i \geq 2} x_1 X_i + x_1^2)} \varrho(dx_1)}.$$

Now, if  $|X_1| \leq c$   $\varrho$ -a.s

$$\mathbb{E}[X_1'|\mathcal{F}] \leq \frac{\int x_1 e^{\beta/2n(\sum_{i \geq 2} x_1 X_i)} \varrho(dx_1) e^{\beta c^2/2n}}{\int e^{\beta/2n(\sum_{i \geq 2} x_1 X_i + x_1^2)} \varrho(dx_1) e^{-\beta c^2/2n}}$$

and

$$\mathbb{E}[X_1'|\mathcal{F}] \geq \frac{\int x_1 e^{\beta/2n(\sum_{i \geq 2} x_1 X_i)} \varrho(dx_1) e^{-\beta c^2/2n}}{\int e^{\beta/2n(\sum_{i \geq 2} x_1 X_i + x_1^2)} \varrho(dx_1) e^{\beta c^2/2n}}.$$

By computation of the derivative of  $G_\varrho$  we see that

$$\frac{\int x_1 e^{\beta/2n(\sum_{i \geq 2} x_1 X_i)} \varrho(dx_1)}{\int e^{\beta/2n(\sum_{i \geq 2} x_1 X_i + x_1^2)} \varrho(dx_1)} e^{\pm \beta c^2/n} = (m_1(X) - \frac{1}{\beta} G'_\varrho(\beta, m_1(X))) (1 \pm \beta c^2/n).$$

□

**Remark 6.2.** If we consider the Curie-Weiss model with respect to  $\widehat{P}_{n,\beta}$ , the conditional density  $g_\beta(x_1|(X_i)_{i \geq 2})$  under this measure becomes

$$g_\beta(x_1|(X_i)_{i \geq 2}) = \frac{e^{\beta/2n(\sum_{i \geq 2} x_1 X_i)}}{\int e^{\beta/2n(\sum_{i \geq 2} x_1 X_i)} \varrho(dx_1)}.$$

Thus we obtain  $\mathbb{E}(X_i'|\mathcal{F}) = (m_i(X) - \frac{1}{\beta} G'_\varrho(\beta, m_i(X)))$  without the boundedness assumption for the  $X_1$ .

Applying Lemma 6.1 and the presentation (1.5) of  $G_\varrho$ , it follows that

$$\mathbb{E}[W - W'|W] = \frac{1}{n}W - \frac{1}{n^{1-1/(2k)}} \left( \frac{1}{n} \sum_{i=1}^n \left( m_i(X) - \frac{\mu}{\beta(2k-1)!} m_i(X)^{2k-1} + \mathcal{O}(m_i(X)^{2k}) \right) \right).$$

With  $m_i(X) = m(X) - \frac{X_i}{n}$  and  $m(X) = \frac{1}{n^{1/(2k)}}W$  we obtain

$$\frac{1}{n^{1-1/(2k)}} \frac{1}{n} \sum_{i=1}^n m_i(X) = \frac{1}{n}W - \frac{1}{n^2}W$$

and

$$\frac{1}{n^{1-1/(2k)}} \frac{1}{n} \sum_{i=1}^n \frac{\mu}{\beta(2k-1)!} m_i(X)^{2k-1} = \frac{1}{n^{1-1/(2k)}} \frac{\mu}{\beta(2k-1)!} \sum_{l=0}^{2k-1} \binom{2k-1}{l} m(X)^{2k-1-l} \frac{(-1)^l}{n^l} \frac{1}{n} \sum_{i=1}^n X_i^l.$$

For any  $k \geq 1$  the first summand ( $l = 0$ ) is

$$\frac{1}{n^{2-\frac{1}{k}}} \frac{\mu}{\beta(2k-1)!} W^{2k-1} = -\frac{1}{n^{2-\frac{1}{k}}} \psi(W). \quad (6.1)$$

To see this, let  $k = 1$ . Since we set  $\phi''(0) = 1$ , we obtain  $\mu(0) = \beta - \beta^2$  and therefore  $\frac{1}{\beta}\mu(0)W = (1 - \beta)W$ . In the case  $k \geq 2$  we know that  $\beta = 1$ . Hence in both cases, (6.1) is checked. Summarizing we obtain for any  $k \geq 1$

$$\mathbb{E}[W - W'|W] = -\frac{1}{n^{2-\frac{1}{k}}} \psi(W) + R(W) =: -\lambda\psi(W) + R(W)$$

with

$$\begin{aligned} R(W) &= \frac{1}{n^2} W + \frac{\mu}{\beta(2k-1)!} \frac{1}{n^{1-1/(2k)}} \sum_{l=1}^{2k-1} \binom{2k-1}{l} m(X)^{2k-1-l} \frac{(-1)^l}{n^l} \frac{1}{n} \sum_{i=1}^n X_i^l + \mathcal{O}(m(X)^{2k}) \\ &= \frac{1}{n^2} W + \frac{\mu}{\beta(2k-1)!} \sum_{l=1}^{2k-1} \binom{2k-1}{l} \frac{1}{n^{2-\frac{1}{k}-\frac{l}{2k}}} W^{2k-1-l} \frac{(-1)^l}{n^l} \frac{1}{n} \sum_{i=1}^n X_i^l + \mathcal{O}\left(\frac{W^{2k}}{n}\right). \end{aligned}$$

With Lemma 1.13 we know that  $\mathbb{E}|W|^{2k} \leq \text{const.}$  We will apply Corollary 2.9, if  $k = 1$  and Theorem 4.7 for  $k \geq 2$ . In both cases we apply Lemma 4.2. Since the spin variables are assumed to be bounded  $\varrho$ -a.s, we have

$$|W - W'| \leq \frac{\text{const.}}{n^{1-\frac{1}{2k}}} =: A.$$

Let  $k = 1$ . Now  $\lambda = \frac{1}{n}$ ,  $A = \text{const.}/n^{-1/2}$ ,  $\mathbb{E}(W^4) \leq \text{const.}$  The leading term of  $R$  is  $W/n^2$ . Hence the last four summands in (2.14) of Corollary 2.9 are  $\mathcal{O}(n^{-1/2})$ .

For  $k \geq 2$  we obtain  $\frac{3A}{2}\mathbb{E}(|\psi(W)|) = \mathcal{O}(n^{\frac{1}{2k}-1})$  and  $\frac{1}{\lambda}\left(\frac{d_4 A^3}{4}\right) = \mathcal{O}(n^{\frac{1}{2k}-1})$ . The leading term in the second term of  $R(W)$  is the first summand ( $l = 1$ ), which is of order  $\mathcal{O}(n^{-3+\frac{1}{k}+\frac{1}{2k}})$ . With  $\lambda = n^{\frac{1}{k}-2}$  we obtain

$$\frac{\mathbb{E}(|R|)}{\lambda} \leq \frac{\mathbb{E}(|W|)}{\lambda n^2} + \mathcal{O}(n^{\frac{1}{2k}-1}) \quad \text{and} \quad \frac{\mathbb{E}(|W|)}{\lambda n^2} = \mathcal{O}(n^{1/k}).$$

Hence the last four summands in (4.10) of Theorem 4.7 are  $\mathcal{O}(n^{-1/k})$ .

Finally we have to consider the variance of  $\frac{1}{2\lambda}\mathbb{E}[(W - W')^2|W]$ . Hence we have to bound the variance of

$$\frac{1}{2n} \sum_{i=1}^m X_i^2 + \frac{1}{2n} \sum_{i=1}^n E[(X_i^2)|\mathcal{F}] + \frac{1}{n} \sum_{i=1}^n X_i \left( m_i(X) - \frac{1}{\beta} G'_\varrho(\beta, m_i(X)) \right) (1 + \mathcal{O}(1/n)). \quad (6.2)$$

Since we assume that  $\varrho \in \text{GHS}$ , we can apply the correlation-inequality due to Lebowitz (see Remark 1.12)

$$\mathbb{E}(X_i X_j X_k X_l) - \mathbb{E}(X_i X_j) \mathbb{E}(X_k X_l) - \mathbb{E}(X_i X_k) \mathbb{E}(X_j X_l) - \mathbb{E}(X_i X_l) \mathbb{E}(X_j X_k) \leq 0.$$

The choice  $i = k$  and  $j = l$  leads to the bound

$$\text{Cov}(X_i^2, X_j^2) = \mathbb{E}(X_i^2 X_j^2) - \mathbb{E}(X_i^2) \mathbb{E}(X_j^2) \leq 2(\mathbb{E}(X_i X_j))^2.$$

With Lemma 1.13 we know that  $(\mathbb{E}(X_i X_j))^2 \leq \text{const.} n^{-2/k}$ . This gives

$$\text{Var}\left(\frac{1}{2n} \sum_{i=1}^n X_i^2\right) = \frac{1}{4n^2} \sum_{i=1}^n \text{Var}(X_i^2) + \frac{1}{4n^2} \sum_{1 \leq i < j \leq n} \text{Cov}(X_i^2, X_j^2) = \mathcal{O}(n^{-1}) + \mathcal{O}(n^{-2/k}).$$

Using a conditional version of Jensen's inequality we have

$$\text{Var}\left(\mathbb{E}\left(\frac{1}{2n} \sum_{i=1}^n X_i^2 \middle| \mathcal{F}\right)\right) \leq \text{Var}\left(\frac{1}{2n} \sum_{i=1}^n X_i^2\right).$$

Hence the variance of the second term in (6.2) is of the same order as the variance of the first term. Applying (1.5) for  $G_\rho$ , the variance of the third term in (6.2) is of the order of the variance of  $W^2/n^{1/k}$ . Summarizing the variance of (6.2) can be bounded by 9 times the maximum of the variances of the three terms in (6.2), which is a constant times  $n^{-2/k}$ , and therefore for  $k \geq 1$  we obtain

$$\left(\text{Var}\left(\frac{1}{2\lambda} \mathbb{E}[(W - W')^2 | W]\right)\right)^{1/2} = \mathcal{O}(n^{-1/k}).$$

Note that for  $k \geq 2$

$$\frac{\psi(x)}{-\mathbb{E}[W\psi(W)]} = -\frac{x^{2k-1}}{\mathbb{E}(W^{2k})}.$$

Hence we compare the distribution of  $W$  with a distribution with Lebesgue-probability density proportional to  $\exp\left(-\frac{x^{2k}}{2k\mathbb{E}(W^{2k})}\right)$ .  $\square$

*Proof of Theorem 1.8.* Since  $\alpha = 0$  and  $k = 1$  for  $\beta \neq 1$  while  $\alpha = 0$  and  $k \geq 2$  for  $\beta = 1$ ,  $G_\rho(\cdot)$  can now be expanded as

$$G(s) = G(0) + \frac{\mu_1}{2}s^2 + \frac{\mu_k}{(2k)!}s^{2k} + \mathcal{O}(s^{2k+1}) \quad \text{as } s \rightarrow 0.$$

Hence  $\frac{1}{\beta_n} G'_\rho(s) = \frac{\mu_1}{\beta_n}s + \frac{\mu_k}{\beta_n(2k-1)!}s^{2k-1} + \mathcal{O}(s^{2k})$ . With Lemma 6.1 and  $\mu_1 = (1 - \beta_n)\beta_n$  we obtain

$$\mathbb{E}[X_i | \mathcal{F}] = \beta_n m_i(X) - \frac{\mu_k}{\beta_n(2k-1)!} m_i(X)^{2k-1} (1 + \mathcal{O}(1/n)).$$

We get

$$\mathbb{E}[W - W' | W] = \frac{1 - \beta_n}{n} W + \frac{\beta_n}{n^2} W + \frac{1}{n^{2-1/k}} \frac{\mu_k}{\beta_n(2k-1)!} W^{2k+1} + R(\beta_n, W).$$

The remainder  $R(\beta_n, W)$  is the remainder in the proof of Theorem 1.7 with  $\mu$  exchanged by  $\mu_k$  and  $\beta$  exchanged by  $\beta_n$ .

Let  $\beta_n - 1 = \frac{\gamma}{n^{1-1/k}}$  and  $W = n^{1/(2k)-1} \sum_{i=1}^n X_i$ . We obtain

$$\mathbb{E}[W - W' | W] = -\frac{1}{n^{2-1/k}} \psi(W) + \frac{\beta_n}{n^2} W + R(\beta_n, W), \quad (6.3)$$

where  $\psi(x) = \gamma x - \frac{\mu_k}{\beta_n(2k-1)!} x^{2k-1}$ . As in the proof of Theorem 1.7 we obtain that  $R(\beta_n, W) = \mathcal{O}(n^{-2})$ . Now we only have to adapt the proof of Theorem 1.7 step by step, applying Lemma 1.13, Lemma 4.2 and Theorem 4.7.

Let  $|\beta_n - 1| = \mathcal{O}(1/n)$  and  $W = n^{1/(2k)-1} \sum_{i=1}^n X_i$ . Now in (6.3), the term  $\frac{1-\beta_n}{n}W$  will be a part of the remainder:

$$\begin{aligned} \mathbb{E}[W - W'|W] &= \frac{1}{n^{2-1/k}} \frac{\mu_k}{\beta_n(2k-1)!} W^{2k+1} + R(\beta_n, W) + \frac{\beta_n}{n^2} W + \frac{1-\beta_n}{n} W \\ &=: -\frac{1}{\beta_n n^{2-1/k}} \psi(W) + \hat{R}(\beta, W) \end{aligned}$$

with  $\psi(x) = -\frac{\mu_k}{(2k-1)!} x^{2k-1}$ . Following the lines of the proof of Theorem 1.7, we have to bound  $\frac{\mathbb{E}|\hat{R}(\beta_n, W)|}{\lambda}$  with  $\lambda := \frac{1}{\beta_n n^{2-1/k}}$ . Since by our assumption for  $(\beta_n)_n$  we have

$$\lim_{n \rightarrow \infty} \frac{1}{\lambda} \frac{(1-\beta_n)}{n} = \beta_n(1-\beta_n)n^{1-1/k} = 0.$$

Thus with Theorem 4.7 we obtain convergence in distribution for any  $\beta_n$  with  $|\beta_n - 1| \ll n^{-(1-1/k)}$ . Moreover we obtain the Berry-Esseen bound of order  $\mathcal{O}(n^{-1/k})$  for any  $|\beta_n - 1| = \mathcal{O}(n^{-1})$ .

Finally we consider  $|\beta_n - 1| \gg n^{-(1-1/2)}$  and  $W = \sqrt{\frac{1-\beta_n}{n}} S_n$ . A little calculation gives

$$\mathbb{E}[W - W'|W] = \frac{1-\beta_n}{n} W + \frac{\beta_n W}{n^2} + \frac{\mu_k}{(2k-1)! n^k (1-\beta_n)^{k-1} \beta_n} W^{2k-1} + R(\beta_n, W) =: -\lambda \psi(W) + \hat{R}(\beta_n, W)$$

with  $\psi(x) = -x$  and  $\lambda = \frac{1-\beta_n}{n}$ . Now we apply Corollary 2.9. With  $A := \frac{\text{const.}(1-\beta_n)^{1/2}}{\sqrt{n}}$  we obtain

$$\frac{A^3}{\lambda} \leq \frac{\text{const.}(1-\beta_n)^{1/2}}{\sqrt{n}} \quad \text{and} \quad \frac{\mathbb{E}|\hat{R}(\beta_n, W)|}{\lambda} \leq \frac{\text{const}}{n^{k-1}(1-\beta_n)^k}.$$

Remark that the bound on the right hand side is good for any  $|\beta_n - 1| \gg n^{-(1-1/k)}$ . Finally we have to bound the variance of  $\frac{1}{2\lambda} \mathbb{E}[(W - W')^2|W]$ . The leading term is the variance of

$$\frac{1}{n} \sum_{i=1}^n X_i \left( m_i(X) - \frac{1}{\beta} G'_\varrho(\beta, m_i(X)) \right),$$

which is of order  $\mathcal{O}\left(\frac{\beta_n}{n(1-\beta_n)}\right)$ . Hence with  $|\beta_n - 1| \gg n^{-(1-1/k)}$  we get convergence in distribution. Under the additional assumption that  $|\beta_n - 1| \gg n^{-(1/2-1/(2k))}$  we obtain the Berry-Esseen bound.  $\square$

*Proof of Theorem 1.10.* We apply Theorem 4.7. For unbounded spin variables  $X_i$  we consider  $\hat{P}_{n,\beta}$  and apply Lemma 6.1 to bound  $\frac{1}{\lambda} \sqrt{\text{Var}(\mathbb{E}[(W - W')^2|W])}$  exactly as in the proof of Theorem 1.7. By Theorem 4.7 it remains to bound  $\frac{1}{\lambda} \mathbb{E}|W - W'|^3$ . With  $\lambda = n^{-2+1/k}$  we have

$$\frac{1}{\lambda} \mathbb{E}|W - W'|^3 = \frac{1}{n^{1-1/2k}} \mathbb{E}|X_I - X'_I|^3 = \frac{1}{n^{1-1/2k}} \mathbb{E}|X_1 - X'_1|^3.$$

Now  $\mathbb{E}|X_1 - X'_1|^3 \leq \mathbb{E}|X_1|^3 + 3\mathbb{E}|X_1^2 X'_1| + 3\mathbb{E}|X_1(X'_1)^2| + \mathbb{E}|X'_1|^3$ . Using Hölder's inequality we obtain

$$\mathbb{E}|X_1^2 X'_1| \leq (\mathbb{E}|X_1|^3)^{2/3} (\mathbb{E}|X'_1|^3)^{1/3} \leq \max(\mathbb{E}|X_1|^3, \mathbb{E}|X'_1|^3).$$

Hence we have

$$\frac{1}{\lambda} \mathbb{E}|W - W'|^3 \leq \frac{8}{n^{1-1/2k}} \max(\mathbb{E}|X_1|^3, \mathbb{E}|X'_1|^3).$$

Thus the Theorem is proved.  $\square$

## 7. EXAMPLES

It is known that the following distributions  $\varrho$  are GHS (see [11, Theorem 1.2]). The symmetric Bernoulli measure is GHS, first noted in [9]. The family of measures

$$\varrho_a(dx) = a \delta_x + ((1-a)/2)(\delta_{x-1} + \delta_{x+1})$$

for  $0 \leq a \leq 2/3$  is GHS, whereas the GHS-inequality fails for  $2/3 < a < 1$ , see [21, p.153]. GHS contains all measures of the form

$$\varrho_V(dx) := \left( \int_{\mathbb{R}} \exp(-V(x)) dx \right)^{-1} \exp(-V(x)) dx,$$

where  $V$  is even, continuously differentiable, and unbounded above at infinity, and  $V'$  is convex on  $[0, \infty)$ . GHS contains all absolutely continuous measures  $\varrho \in \mathcal{B}$  with support on  $[-a, a]$  for some  $0 < a < \infty$  provided  $g(x) = d\varrho/dx$  is continuously differentiable and strictly positive on  $(-a, a)$  and  $g'(x)/g(x)$  is concave on  $[0, a)$ . Measures like  $\varrho(dx) = \text{const.} \exp(-ax^4 - bx^2) dx$  or  $\varrho(dx) = \text{const.} \exp(-a \cosh x - bx^2) dx$  with  $a > 0$  and  $b$  real are GHS. Both are of physical interest, see [11] and references therein).

**Example 7.1** (A Curie–Weiss model with three states). We will now consider the next simplest example of the classical Curie–Weiss model: a model with three states. Observe, that this is not the Curie–Weiss–Potts model [15], since the latter has a different Hamiltonian. Indeed the Hamiltonian considered in [15] is of the form  $\frac{1}{n} \sum_{i,j} \delta_{x_i, x_j}$ . It favours states with many equal spins, whereas in our case the spins also need to have large values. We choose  $\varrho$  to be

$$\varrho = \frac{2}{3} \delta_0 + \frac{1}{6} \delta_{-\sqrt{3}} + \frac{1}{6} \delta_{\sqrt{3}}.$$

This model seems to be of physical relevance. It is studied in [24]. In [3] it was used to analyze the tri-critical point of liquid helium. A little computation shows that

$$\frac{d^3}{ds^3} \phi_{\varrho}(s) = -6 \frac{\sinh(x \sqrt{3}) \sqrt{3} (\cosh(x \sqrt{3}) - 1)}{12 \cosh(x \sqrt{3}) + 6 \cosh(x \sqrt{3})^2 + \cosh(x \sqrt{3})^3 + 8} \leq 0$$

for all  $s \geq 0$ . Hence the GHS-inequality (1.10) is fulfilled (see also [11, Theorem 1.2]), which implies that there is one critical temperature  $\beta_c$  such that there is one minimum of  $G$  for  $\beta \leq \beta_c$  and two minima above  $\beta_c$ . Since  $\text{Var}_{\varrho}(X_1) = 2 \frac{1}{6} \cdot 3 = 1$  we see that  $\beta_c = 1$ . For  $\beta \leq \beta_c$  the minimum of  $G$  is located in zero while for  $\beta > 1$  the two minima are symmetric and satisfy

$$s = \frac{\sqrt{3} \sinh(\sqrt{3} \beta s)}{2 + \cosh(\sqrt{3} \beta s)}.$$

Now Theorem 1.7 and 1.8 tell that

- For  $\beta < 1$  the rescaled magnetization  $S_n/\sqrt{n}$  satisfies a Central Limit Theorem and the limiting variance is  $(1 - \beta)^{-1}$ . Indeed,  $\frac{d^2}{ds^2} \phi_{\varrho}(0) = \text{Var}_{\varrho}(X_1) = 1$ . Hence  $\mu_1 = \beta - \beta^2$  and  $\sigma^2 = \frac{1}{1-\beta}$ . Moreover we obtain

$$\sup_{z \in \mathbb{R}} \left| P_n \left( \frac{S_n}{\sqrt{n}} \leq z \right) - \Phi_W(z) \right| \leq \frac{C}{\sqrt{n}}.$$

- For  $\beta = \beta_c = 1$  the rescaled magnetization  $S_n/n^{5/6}$  converges in distribution to  $X$  which has the density  $f_{3,6,1}$ . Indeed  $\mu_2$  is computed to be 6. Moreover we obtain

$$\sup_{z \in \mathbb{R}} \left| P_n \left( \frac{S_n}{n^{5/6}} \leq z \right) - \widehat{F}_3(z) \right| \leq \frac{C}{n^{1/3}}$$

where the derivative of  $\widehat{F}_3$  is the rescaled density  $\exp\left(-\frac{x^6}{6\mathbb{E}(W^6)}\right)$ .

- If  $\beta_n$  converges monotonically to 1 faster than  $n^{-2/3}$  then  $\frac{S_n}{n^{5/6}}$  converges in distribution to  $\widehat{F}_3$ , whereas if  $\beta_n$  converges monotonically to 1 slower than  $n^{-2/3}$  then  $\frac{\sqrt{1-\beta_n} S_n}{\sqrt{n}}$  satisfies a Central Limit Theorem. Eventually, if  $|1 - \beta_n| = \gamma n^{-2/3}$ ,  $\frac{S_n}{n^{5/6}}$  converges in distribution to a random variable which probability distribution has the mixed Lebesgue-density

$$\exp\left(-c_W^{-1} \left( \frac{x^6}{120} - \gamma \frac{x^2}{2} \right)\right)$$

with  $c_W = \frac{1}{120}\mathbb{E}(W^6) - \gamma\mathbb{E}(W^2)$ . Moreover we have

$$\sup_{z \in \mathbb{R}} \left| P_n \left( \frac{S_n}{n^{5/6}} \leq z \right) - \frac{1}{Z} \int_{-\infty}^z \exp\left(-c_W^{-1} \left( \frac{x^6}{120} - \gamma \frac{x^2}{2} \right)\right) dx \right| \leq \frac{C}{n^{1/3}}.$$

**Example 7.2** (A continuous Curie–Weiss model). Last but not least we will treat an example of a continuous Curie–Weiss model. We choose as underlying distribution the uniform distribution on an interval in  $\mathbb{R}$ . To keep the critical temperature one we define

$$\frac{d\varrho(x_i)}{dx_i} = \frac{1}{2a} \mathbb{I}_{[-a,a]}(x_i)$$

with  $a = \sqrt{3}$ . Then from a general result in [13, Theorem 2.4] (see also [11, Theorem 1.2]) it follows that  $\varrho(x_i)$  obeys the GHS-inequality (1.10). Therefore there exists a critical temperature  $\beta_c$ , such that for  $\beta < \beta_c$  zero is the unique global minimum of  $G$  and is of type 1, while at  $\beta_c$  this minimum is of type  $k \geq 2$ . This  $\beta_c$  is easily computed to be one. Indeed,  $\mu_1 = \beta - \beta^2 \phi''(0) = \beta - \beta^2 \mathbb{E}_\varrho(X_1^2) = \beta(1 - \beta)$ , since  $\varrho$  is centered and has variance one. Thus  $\mu_1$  vanishes at  $\beta = \beta_c = 1$ . Eventually for  $\beta > 1$  there are again two minima which are solutions of

$$\frac{\sqrt{3}\beta}{\tanh(\sqrt{3}\beta x)} = \beta x + \frac{1}{x}.$$

Now again by Theorems 1.7 and 1.8

- For  $\beta < 1$  the rescaled magnetization  $S_n/\sqrt{n}$  obeys a Central Limit Theorem and the limiting variance is  $(1 - \beta)^{-1}$ . Indeed, since  $\mathbb{E}_\varrho(X_1^2) = 1$ ,  $\mu_1 = \beta - \beta^2$  and  $\sigma^2 = \frac{1}{1-\beta}$ .
- For  $\beta = \beta_c = 1$  the rescaled magnetization  $S_n/n^{7/8}$  converges in distribution to  $X$  which has the density  $f_{4,6/5,1}$ . Indeed  $\mu_2$  is computed to be

$$-\mathbb{E}_\varrho(X_1^4) + 3\mathbb{E}_\varrho(X_1^2) = -\frac{9}{5} + 3 = \frac{6}{5}.$$

Moreover we obtain

$$\sup_{z \in \mathbb{R}} \left| P_n \left( \frac{S_n}{n^{7/8}} \leq z \right) - \widehat{F}_4(z) \right| \leq \frac{C}{n^{1/4}}$$

where the derivative of  $\widehat{F}_4$  is the rescaled density  $\exp\left(-\frac{x^8}{8\mathbb{E}(W^8)}\right)$ .

- If  $\beta_n$  converges monotonically to 1 faster than  $n^{-3/4}$  then  $\frac{S_n}{n^{7/8}}$  converges in distribution to  $\widehat{F}_4$ , whereas if  $\beta_n$  converges monotonically to 1 slower than  $n^{-3/4}$  then  $\frac{\sqrt{1-\beta_n}S_n}{\sqrt{n}}$  satisfies a Central Limit Theorem. Eventually, if  $|1 - \beta_n| = \gamma n^{-3/4}$ ,  $\frac{S_n}{n^{7/8}}$  converges in distribution to the mixed density

$$\exp\left(-c_W^{-1}\left(\frac{6}{5}\frac{x^8}{8!} - \gamma\frac{x^2}{2}\right)\right)$$

with  $c_W = \frac{5}{6(8!)}\mathbb{E}(W^8) - \gamma\mathbb{E}(W^2)$ . Moreover we have

$$\sup_{z \in \mathbb{R}} \left| P_n\left(\frac{S_n}{n^{7/8}} \leq z\right) - \frac{1}{Z} \int_{-\infty}^z \exp\left(-c_W^{-1}\left(\frac{6}{5}\frac{x^8}{8!} - \gamma\frac{x^2}{2}\right)\right) dx \right| \leq \frac{C}{n^{1/4}}.$$

Note that there is some interesting change in limiting behaviour of all of these models at criticality. While for  $\beta < 1$  all of the models have the same rate of convergence for the Central Limit Theorem behaviour, in the limit at criticality the limiting distribution function as well as the distributions which depend on some moments of  $W$  becomes characteristic of the underlying distribution  $\varrho$ . Moreover the rate of convergence differs at criticality (for  $k \geq 3$ ).

## 8. APPENDIX

*Proof of Lemma 4.2.* Consider a probability density of the form

$$p(x) := p_k(x) := b_k \exp(-a_k x^{2k}) \quad (8.1)$$

with  $b_k = \int_{\mathbb{R}} \exp(-a_k x^{2k}) dx$ . Clearly  $p$  satisfies Assumption (D). First we prove that the solutions  $f_z$  of the Stein equation, which characterizes the distribution with respect to the density (8.1), satisfies Assumption (B2). Let  $f_z$  be the solution of

$$f'_z(x) + \psi(x)f_z(x) = 1_{\{x \leq z\}}(x) - P(z).$$

Here  $\psi(x) = -2k a_k x^{2k-1}$ . We have

$$f_z(x) = \begin{cases} (1 - P(z)) P(x) \exp(a_k x^{2k}) b_k^{-1} & \text{for } x \leq z, \\ P(z) (1 - P(x)) \exp(a_k x^{2k}) b_k^{-1} & \text{for } x \geq z \end{cases} \quad (8.2)$$

with  $P(z) := \int_{-\infty}^z p(x) dx$ . Note that  $f_z(x) = f_{-z}(-x)$ , so we need only to consider the case  $z \geq 0$ . For  $x > 0$  we obtain

$$1 - P(x) \leq \frac{b_k}{2k a_k x^{2k-1}} \exp(-a_k x^{2k}), \quad (8.3)$$

whereas for  $x < 0$  we have

$$P(x) \leq \frac{b_k}{2k a_k |x|^{2k-1}} \exp(-a_k x^{2k}). \quad (8.4)$$

By partial integration we have

$$\int_x^\infty \frac{(2k-1)}{2k a_k} t^{-2k} \exp(-a_k t^{2k}) dt = -\frac{1}{2k a_k t^{2k-1}} \exp(-a_k t^{2k}) \Big|_x^\infty - \int_x^\infty \exp(-a_k t^{2k}) dt.$$

Hence for any  $x > 0$

$$b_k \left( \frac{x}{2k a_k x^{2k} + 2k - 1} \right) \exp(-a_k x^{2k}) \leq 1 - P(x). \quad (8.5)$$

With (8.3) we get for  $x > 0$

$$\frac{d}{dx} \left( \exp(a_k x^{2k}) \int_x^\infty \exp(-a_k t^{2k}) dt \right) = -1 + 2k a_k x^{2k-1} \exp(a_k x^{2k}) \int_x^\infty \exp(-a_k t^{2k}) dt < 0.$$

So  $\exp(a_k x^{2k}) \int_x^\infty \exp(-a_k t^{2k}) dt$  attains its maximum at  $x = 0$  and therefore

$$\exp(a_k x^{2k}) b_k \int_x^\infty \exp(-a_k t^{2k}) dt \leq \frac{1}{2}.$$

Summarizing we obtain for  $x > 0$

$$1 - P(x) \leq \min \left( \frac{1}{2}, \frac{b_k}{2k a_k x^{2k-1}} \right) \exp(-a_k x^{2k}). \quad (8.6)$$

With (8.4) we get for  $x < 0$

$$\frac{d}{dx} \left( \exp(a_k x^{2k}) \int_{-\infty}^x \exp(-a_k t^{2k}) dt \right) = 1 + 2k a_k x^{2k-1} \exp(a_k x^{2k}) \int_{-\infty}^x \exp(-a_k t^{2k}) dt > 0.$$

So  $\exp(a_k x^{2k}) \int_{-\infty}^x \exp(-a_k t^{2k}) dt$  attains its maximum at  $x = 0$  and therefore

$$\exp(a_k x^{2k}) b_k \int_{-\infty}^x \exp(-a_k t^{2k}) dt \leq \frac{1}{2}.$$

Summarizing we obtain for  $x < 0$

$$P(x) \leq \min \left( \frac{1}{2}, \frac{b_k}{2k a_k |x|^{2k-1}} \right) \exp(-a_k x^{2k}). \quad (8.7)$$

Applying (8.6) and (8.7) gives  $0 < f_z(x) \leq \frac{1}{2b_k}$  for all  $x$ . Note that for  $x < 0$  we only have to consider the first case of (8.2), since  $z \geq 0$ . The constant  $\frac{1}{2b_k}$  is not optimal. Following the proof of Lemma 2.2 in [6] or alternatively of Lemma 2 in [22, Lecture II] would lead to optimal constants. We omit this. It follows from (8.2) that

$$f'_z(x) = \begin{cases} (1 - P(z)) \left[ 1 + x^{2k-1} 2k a_k P(x) \exp(a_k x^{2k}) b_k^{-1} \right] & \text{for } x \leq z, \\ P(z) \left[ (1 - P(x)) 2k a_k x^{2k-1} \exp(a_k x^{2k}) b_k^{-1} - 1 \right] & \text{for } x \geq z. \end{cases} \quad (8.8)$$

With (8.3) we obtain for  $0 < x \leq z$  that

$$f'_z(x) \leq (1 - P(z)) \left[ z^{2k-1} 2k a_k P(z) \exp(a_k z^{2k}) b_k^{-1} \right] + 1 \leq 2.$$

The same argument for  $x \geq z$  leads to  $|f'_z(x)| \leq 2$ . For  $x < 0$  we use the first half of (8.2) and apply (8.4) to obtain  $|f'_z(x)| \leq 2$ . Actually this bound will be improved later. Next we calculate the derivative of  $-\psi(x) f_z(x)$ :

$$(-\psi(x) f_z(x))' = \begin{cases} \frac{(1-P(z))}{b_k} \left[ P(x) e^{a_k x^{2k}} \left( 2k(2k-1) a_k x^{2k-2} + (2k)^2 a_k^2 x^{4k-2} \right) + 2k a_k x^{2k-1} b_k \right], & x \leq z, \\ \frac{P(z)}{b_k} \left[ (1 - P(x)) e^{a_k x^{2k}} \left( 2k(2k-1) a_k x^{2k-2} + (2k)^2 a_k^2 x^{4k-2} \right) - 2k a_k x^{2k-1} b_k \right], & x \geq z. \end{cases} \quad (8.9)$$

With (8.5) we obtain  $(-\psi(x)f_z(x))' \geq 0$ , so  $-\psi(x)f_z(x)$  is an increasing function of  $x$  (remark that for  $x < 0$  we only have to consider the first half of (8.2)). Moreover with (8.3), (8.4) and (8.5) we obtain that

$$\lim_{x \rightarrow -\infty} 2k a_k x^{2k-1} f_z(x) = P(z) - 1 \quad \text{and} \quad \lim_{x \rightarrow \infty} 2k a_k x^{2k-1} f_z(x) = P(z). \quad (8.10)$$

Hence we have  $|2k a_k x^{2k-1} f_z(x)| \leq 1$  and  $|2k a_k (x^{2k-1} f_z(x) - u^{2k-1} f_z(u))| \leq 1$  for any  $x$  and  $u$ . From (8.3) it follows that  $f'_z(x) > 0$  for all  $x < z$  and  $f'_z(x) < 0$  for  $x > z$ . With Stein's identity  $f'_z(x) = -\psi(x)f_z(x) + 1_{\{x \leq z\}} - P(z)$  and (8.10) we have

$$0 < f'_z(x) \leq -\psi(z)f_z(z) + 1 - P(z) < 1 \quad \text{for} \quad x < z$$

and

$$-1 < -\psi(z)f_z(z) - P(z) \leq f'_z(x) < 0 \quad \text{for} \quad x > z.$$

Hence, for any  $x$  and  $y$ , we obtain

$$|f'_z(x)| \leq 1 \quad \text{and} \quad |f'_z(x) - f'_z(y)| \leq \max(1, -\psi(z)f_z(z) + 1 - P(z) - (-\psi(z)f_z(z) - P(z))) = 1.$$

Next we bound  $(-\psi(x)f_z(x))'$ . We already know that  $(-\psi(x)f_z(x))' > 0$ . Again we apply (8.3) and (8.4) to see that

$$(-\psi(x)f_z(x))' \leq \frac{2k-1}{|x|}$$

for  $x \geq z > 0$  and all  $x \leq 0$ . For  $0 < x \leq z$  this latter bound holds, as can be seen by applying this bound (more precisely the bound for  $(-\psi(x)f_z(x))' \frac{b_k}{P(z)}$  for  $x \geq z$ ) with  $-x$  for  $x$  to the formula for  $(\psi(x)f_z(x))'$  in  $x \leq z$ . For some constant  $c$  we can bound  $(\psi(x)f_z(x))'$  by  $c$  for all  $|x| \geq \frac{2k-1}{c}$ . Moreover, on  $[-\frac{2k-1}{c}, \frac{2k-1}{c}]$  the continuous function  $(-\psi(x)f_z(x))'$  is bounded by some constant  $d$ , hence we have proved

$$|-(\psi(x)f_z(x))'| \leq \max(c, d).$$

The problem of finding the optimal constant, depending on  $k$ , is omitted. Summarizing, Assumption (B2) is fulfilled for  $p$  with  $d_2 = d_3 = 1$  and some constants  $d_1$  and  $d_4$ .

Next we consider an absolutely continuous function  $h : \mathbb{R} \rightarrow \mathbb{R}$ . Let  $f_h$  be the solution of the Stein equation (4.2), that is

$$f_h(x) = \frac{1}{p(x)} \int_{-\infty}^x (h(t) - Ph) p(t) dt = -\frac{1}{p(x)} \int_x^{\infty} (h(t) - Ph) p(t) dt.$$

We adapt the proof of [6, Lemma 2.3]: without loss of generality we assume that  $h(0) = 0$  and put  $e_0 := \sup_x |h(x) - Ph|$  and  $e_1 := \sup_x |h'(x)|$ . From the definition of  $f_h$  it follows that  $|f_h(x)| \leq e_0 \frac{1}{2b_k}$ . An alternative bound is  $c_1 e_1$  with some constant  $c_1$  depending on  $\mathbb{E}|Z|$ , where  $Z$  denotes a random variable distributed according to  $p$ . With (4.2) and (8.5), for  $x \geq 0$ ,

$$|f'_h(x)| \leq |h(x) - Ph| - \psi(x) e^{a_k x^{2k}} \int_x^{\infty} |h(t) - Ph| e^{-a_k t^{2k}} dt \leq 2e_0.$$

An alternative bound is  $c_2 e_1$  with some constant  $c_2$  depending on the  $(2k - 2)$ 'th moment of  $p$ . This is using Stein's identity (4.2) to obtain

$$f'_h(x) = -e^{a_k x^{2k}} \int_x^\infty (h'(t) - \psi'(t) f(t)) e^{-a_k t^{2k}} dt.$$

The details are omit. To bound the second derivative  $f''_h$ , we differentiate (4.2) and have

$$f''_h(x) = (\psi^2(x) - \psi'(x)) f_h(x) - \psi(x)(h(x) - Ph) + h'(x).$$

Similarly to [6, (8.8), (8.9)] we obtain

$$h(x) - Ph = \int_{-\infty}^x h'(t) P(t) dt - \int_x^\infty h'(t)(1 - P(t)) dt.$$

It follows that

$$f_h(x) = -\frac{1}{b_k} e^{a_k x^{2k}} (1 - P(x)) \int_{-\infty}^x h'(t) P(t) dt - \frac{1}{b_k} e^{a_k x^{2k}} P(x) \int_x^\infty h'(t)(1 - P(t)) dt.$$

Now we apply the fact that the quantity in (8.9) is non-negative to obtain

$$\begin{aligned} |f''_h(x)| &\leq |h'(x)| + |(\psi^2(x) - \psi'(x)) f_h(x) - \psi(x)(h(x) - Ph)| \\ &\leq |h'(x)| + \left| \left( -\psi(x) - \frac{1}{b_k} (\psi^2(x) - \psi'(x)) e^{a_k x^{2k}} (1 - P(x)) \right) \int_{-\infty}^x h'(t) P(t) dt \right| \\ &\quad + \left| \left( \psi(x) - \frac{1}{b_k} (\psi^2(x) - \psi'(x)) e^{a_k x^{2k}} P(x) \right) \int_x^\infty h'(t)(1 - P(t)) dt \right| \\ &\leq |h'(x)| + e_1 \left( \psi(x) + \frac{1}{b_k} (\psi^2(x) - \psi'(x)) e^{a_k x^{2k}} (1 - P(x)) \right) \int_{-\infty}^x P(t) dt \\ &\quad + e_1 \left( -\psi(x) + \frac{1}{b_k} (\psi^2(x) - \psi'(x)) e^{a_k x^{2k}} P(x) \right) \int_x^\infty (1 - P(t)) dt. \end{aligned}$$

Moreover we know, that the quantity in (8.9) can be bounded by  $\frac{2k-1}{|x|}$ , hence

$$|f''_h(x)| \leq e_1 + e_1 \frac{2b_k(2k-1)}{|x|} \left( \int_{-\infty}^x P(t) dt + \int_x^\infty (1 - P(t)) dt \right).$$

Now we bound

$$\left| \int_{-\infty}^x P(t) dt + \int_x^\infty (1 - P(t)) dt \right| = |xP(x) - x(1 - P(x)) + 2 \int_x^\infty tp(t) dt| \leq 2|x| + 2\mathbb{E}|Z|,$$

where  $Z$  is distributed according to  $p$ . Summarizing we have  $|f''_h(x)| \leq c_3 \sup_x |h'(x)|$  for some constant  $c_3$ , using the fact that  $f_h$  and therefore  $f'_h$  and  $f''_h$  are continuous. Hence  $f_h$  satisfies Assumption (B1).  $\square$

*Sketch of the proof of Remark 4.3.* Now let  $p(x) = b_k \exp(-a_k V(x))$  and  $V$  satisfies the assumptions listed in Remark 4.3. To proof that  $f_z$  (with respect to  $p$ ) satisfies Assumption (B2), we adapt (8.5) as well as (8.6) and (8.7), using the assumptions on  $V$ . We obtain for  $x > 0$

$$b_k \left( \frac{V'(x)}{V''(x) + a_k V'(x)^2} \right) \exp(-a_k V(x)) \leq 1 - P(x).$$

and for  $x > 0$

$$1 - P(x) \leq \min\left(\frac{1}{2}, \frac{b_k}{a_k V'(x)}\right) \exp(-a_k V(x))$$

and for  $x < 0$

$$P(x) \leq \min\left(\frac{1}{2}, \frac{b_k}{a_k |V'(x)|}\right) \exp(-a_k V(x)).$$

Estimating  $(-\psi(x)f_z(x))'$  gives

$$(-\psi(x)f_z(x))' \leq \text{const.} \frac{V''(x)}{|V'(x)|}.$$

By our assumptions on  $V$ , the right hand side can be bounded for  $x \geq d$  with  $d \in \mathbb{R}_+$  and since  $\psi(x)f_z(x)$  is continuous, it is bounded everywhere.  $\square$

*Acknowledgement.* During the preparation of our manuscript we became aware of a preprint of S. Chatterjee and Q.-M. Shao about Stein's method with applications to the Curie-Weiss model. As far as we understand, there the authors give an alternative proof of Theorem 1.2 and 1.3.

## REFERENCES

1. A. D. Barbour, *Equilibrium distributions for Markov population processes*, Adv. in Appl. Probab. **12** (1980), no. 3, 591–614. MR **MR578839** (81h:60107)
2. G. Ben Arous and O. Zeitouni, *Increasing propagation of chaos for mean field models*, Ann. Inst. H. Poincaré Probab. Statist. **35** (1999), no. 1, 85–102. MR **2000a:60041**
3. M. Blume, V. J. Emery, and R. B. Griffiths, *Ising model for the  $\lambda$  transition and phase separation in  $He^3$ - $He^4$  mixtures*, Phys. Rev. A **4** (1971), 1071–1077.
4. S. Chatterjee, J. Fulman, and A. Röllin, *Exponential approximation by Stein's method and spectral graph theory*, preprint, 2009.
5. L. H. Y. Chen and Q.-M. Shao, *A non-uniform Berry-Esseen bound via Stein's method*, Probab. Theory Related Fields **120** (2001), no. 2, 236–254. MR **MR1841329** (2002h:60037)
6. ———, *Stein's method for normal approximation*, An introduction to Stein's method, Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap., vol. 4, Singapore Univ. Press, Singapore, 2005, pp. 1–59. MR **MR2235448**
7. P. Eichelsbacher and M. Löwe, *Moderate deviations for the overlap parameter in the Hopfield model*, Probab. Theory and Related Fields **130** (2004), no. 4, 441–472.
8. P. Eichelsbacher and G. Reinert, *Stein's method for discrete Gibbs measures*, Ann. Appl. Probab. **18** (2008), no. 4, 1588–1618. MR **MR2434182**
9. R. S. Ellis, *Concavity of magnetization for a class of even ferromagnets*, Bull. Amer. Math. Soc. **81** (1975), no. 5, 925–929. MR **MR0376052** (51 #12238)
10. R. S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*, Springer-Verlag, New York, 1985.
11. R. S. Ellis, J. L. Monroe, and C. M. Newman, *The ghs and other correlation inequalities for a class of even ferromagnets*, Comm. Math. Phys. **46** (1976), no. 2, 167–182. MR **MR0395659** (52 #16453)
12. R. S. Ellis and C. M. Newman, *Limit theorems for sums of dependent random variables occurring in statistical mechanics*, Z. Wahrsch. Verw. Gebiete **44** (1978), no. 2, 117–139.
13. ———, *Necessary and sufficient conditions for the GHS inequality with applications to analysis and probability*, Trans. Amer. Math. Soc. **237** (1978), 83–99. MR **58** #11282
14. R. S. Ellis, C.M. Newman, and J. S. Rosen, *Limit theorems for sums of dependent random variables occurring in statistical mechanics, II.*, Z. Wahrsch. Verw. Gebiete **51** (1980), no. 2, 153–169.

15. R. S. Ellis and K. Wang, *Limit theorems for the empirical vector of the Curie-Weiss-Potts model*, Stochastic Process. Appl. **35** (1990), no. 1, 59–79. MR **91m**:82023
16. R. B. Griffiths, C. A. Hurst, and S. Sherman, *Concavity of magnetization of an Ising ferromagnet in a positive external field*, J. Mathematical Phys. **11** (1970), 790–795. MR 42 #1412
17. K. Kondo, T. Otofujii, and Y. Sugiyama, *Correlation inequalities for a class of even ferromagnets*, J. Statist. Phys. **40** (1985), no. 3-4, 563–575. MR **MR806714** (**87a**:82011)
18. J.L. Lebowitz, *GHS and other inequalities*, Comm. Math. Phys. **35** (1974), 87–92. MR MR0339738 (49 #4495)
19. Y. Rinott and V. Rotar, *On coupling constructions and rates in the CLT for dependent summands with applications to the antivoter model and weighted  $U$ -statistics*, Ann. Appl. Probab. **7** (1997), no. 4, 1080–1105. MR **MR1484798** (**99g**:60050)
20. Q.-M. Shao and Z.-G. Su, *The Berry-Esseen bound for character ratios*, Proc. Amer. Math. Soc. **134** (2006), no. 7, 2153–2159 (electronic). MR **MR2215787** (**2008j**:60064)
21. B. Simon and R. B. Griffiths, *The  $(\phi^4)_2$  field theory as a classical Ising model*, Comm. Math. Phys. **33** (1973), 145–164. MR MR0428998 (55 #2018)
22. C. Stein, *Approximate computation of expectations*, Institute of Mathematical Statistics Lecture Notes—Monograph Series, 7, Institute of Mathematical Statistics, Hayward, CA, 1986. MR **MR882007** (**88j**:60055)
23. C. Stein, P. Diaconis, S. Holmes, and G. Reinert, *Use of exchangeable pairs in the analysis of simulations*, Stein’s method: expository lectures and applications, IMS Lecture Notes Monogr. Ser., vol. 46, Inst. Math. Statist., Beachwood, OH, 2004, pp. 1–26. MR **MR2118600** (**2005j**:65005)
24. C. J. Thompson, *Mathematical statistical mechanics*, The Macmillan Co., New York, 1972, A Series of Books in Applied Mathematics. MR 57 #8821