DOI: 10.1214/EJP.v15-777

Electronic Journal of Probability

Vol. 15 (2010), Paper no. 30, pages 962–988.

Journal URL
http://www.math.washington.edu/~ejpecp/

# Stein's method for dependent random variables occurring in Statistical Mechanics*

Peter Eichelsbacher† Matthias Löwe‡

**Abstract**

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 $\exp(-\mu|x|^{2k}/(2k)!)$. As a consequence we obtain convergence rates 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. Our results include a 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.

**Key words:** Berry-Esseen bound, Stein's method, exchangeable pairs, Curie-Weiss models, critical temperature, GHS-inequality.

**AMS 2000 Subject Classification:** Primary 60F05; Secondary: 82B20, 82B26.

Submitted to EJP on March 7,, final version accepted May 20, 2010.

*Both authors have been supported by Deutsche Forschungsgemeinschaft via SFB/TR 12. This research was done partly at the Mathematisches Forschungsinstitut Oberwolfach during a stay within the Research in Pairs Programme.

†Ruhr-Universität Bochum, Fakultät für Mathematik, NA3/68, D-44780 Bochum, Germany, Peich@math.ruhr-uni-bochum.de

‡Universität Münster, Fachbereich Mathematik und Informatik, Institut für Mathematische Statistik, Einsteinstr. 62, D-48149 Münster, Germany maloewe@math.uni-muenster.de

## 1 Introduction

Stein's method is a powerful tool to prove distributional approximation. One of its advantages is that is often automatically provides also a rate of convergence and may be applied rather effectively also to classes of random variables that are stochastically dependent. Such classes of random variables are in natural way provided by spin system in statistical mechanics. The easiest of such models are mean-field models. Among them the Curie–Weiss model is well known for exhibiting a number of properties of real substances, such as spontaneous magnetization or metastability. The aim of this paper is to develop Stein's method for exchangeable pairs (see [20]) for a rich class of distributional approximations and thereby prove Berry-Esseen bounds for the sums of dependent random variables occurring in statistical mechanics under the name Curie-Weiss models. For an overview of results on the Curie–Weiss models and related models, see [9], [11], [13].

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

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

Here $\beta := T^{-1}$ is the inverse temperature and $Z_{\Lambda}(\beta)$ is a normalizing constant, that turns $P_{\Lambda,\beta}$ into a probability measure, known as the partition function and $|\Lambda|$ denotes the cardinality of $\Lambda$. Moreover $\varrho$ is the distribution of a single spin in the limit $\beta \to 0$. 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 $P_n$. In the *classical Curie–Weiss model*, spins are distributed in $\{-1,+1\}$ according to $\varrho = \frac{1}{2}(\delta_{-1} + \delta_1)$. The measures $P_{n,\beta}$ 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 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 [18] and [11] the Central Limit Theorem is proved: $\frac{\sum_{i=1}^n X_i}{\sqrt{n}} \to N(0, \sigma^2(\beta))$ 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 [18] and [11] 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 $\frac{\sum_{i=1}^n X_i}{n^{3/4}} \to X$ as $n \to \infty$ in distribution with respect to the finite-volume Gibbs states.

Stein introduced in [20] 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 ([20, 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 $0 < \lambda < 1$. Stein proved that for any uniformly Lipschitz function $h$, $|\mathbb{E}h(W) - \mathbb{E}h(Z)| \leq \delta\|h'\|$ with $Z$ denoting a standard normally distributed random variable and

$$\delta = 4\mathbb{E}\left|1 - \frac{1}{2\lambda}\mathbb{E}\big((W'-W)^2|W\big)\right| + \frac{1}{2\lambda}\mathbb{E}|W-W'|^3. \tag{1.2}$$

Stein's approach has been successfully applied in many models, see e.g. [20] or [21] and references therein. In [16], the range of application was extended by replacing the linear regression property by a weaker condition.

For a motivation of our paper we consider the construction of an exchangeable pair $(W,W')$ in the classical Curie-Weiss model for $W = (1/\sqrt{n})\sum_{i=1}^n X_i$, proving an approximate regression property. 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 $(X_j)_{j\neq i}$, independently form $X_i$. 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 $W - W' = \frac{X_I - X_I'}{\sqrt{n}}$. For $\mathscr{F} := \sigma(X_1,\dots,X_n)$ we obtain

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

The conditional distribution at site $i$ is given by

$$P_n\big(x_i|(x_j)_{j\neq i}\big) = \frac{\exp\big(x_i\,\beta\, m_i(x)\big)}{\exp\big(\beta m_i(x)\big) + \exp\big(-\beta m_i(x)\big)}, \quad \text{with} \quad m_i(x) := \frac{1}{n}\sum_{j\neq i} x_j,\ i = 1,\dots,n.$$

It follows that

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

The Taylor-expansion $\tanh(x) = x + \mathcal{O}(x^3)$ leads to

$$\mathbb{E}[W-W'|W] = \frac{1-\beta}{n}W + R = \frac{\lambda}{\sigma^2}W + R \tag{1.3}$$

with $\lambda := \frac{1}{n}$, $\sigma^2 := (1-\beta)^{-1}$ and $\mathbb{E}|R| = \mathcal{O}\big(\frac{1}{n^{3/2}}\big)$. With (1.3) we are able to apply Theorem 1.2 in [16]: for any uniformly Lipschitz function $h$, $|\mathbb{E}h(W) - \mathbb{E}h(Z)| \leq \delta'\|h'\|$ with

$$\delta' = 4\mathbb{E}\left|1 - \frac{1}{2\lambda}\mathbb{E}\big((W'-W)^2|W\big)\right| + \frac{1}{2\lambda}\mathbb{E}|W-W'|^3 + 19\frac{\sqrt{\mathbb{E}R^2}}{\lambda}. \tag{1.4}$$

Using (1.3) we will be able to prove a Berry-Esseen rate for $W$, see Section 3.

In the critical case $\beta = 1$ of the classical Curie-Weiss model the Taylor expansion $\tanh(x) = x - x^3/3 + \mathcal{O}(x^5)$, (1.3) would lead to

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

for some $\tilde{R}$. 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 $W := \frac{1}{n^{3/4}} \sum_{i=1}^n X_i$. Constructing the exchangeable pair $(W, W')$ in the same manner as before we will obtain

$$\mathbb{E}[W - W'|W] = \frac{1}{n^{3/2}} \frac{W^3}{3} + R(W) =: -\lambda\psi(W) + R(W) \tag{1.5}$$

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 $0 < \lambda < 1$.

In Section 2 we develop Stein's method for exchangeable pairs for a rich class of other distributional approximations than normal approximation. Moreover, we prove certain refinements of Stein's method for exchangeable pairs in the case of normal approximation. In Section 3 we apply our general approach to consider Berry-Esseen bounds for the rescaled total magnatization for general single spin distributions $\varrho$ in (1.1) satisfying the GHS-inequality. This inequality ensures the application of correlation-inequalities due to Lebowitz for bounding the variances and other low order moments, which appear in Stein's method. Rates of convergence in limit theorems for partial sums in the context of the Curie-Weiss model will be proven (Theorems 3.1, 3.2 and 3.3). Moreover, we prove a Berry-Essen rate in the Central Limit Theorem for the total magentization in the classical Curie-Weiss model, for high tempertaures (Theorem 3.7) as well as at the critical temperature $\beta = 1$ (Theorem 3.8). We analyze Berry-Esseen bounds as the temperature $1/\beta_n$ converges to one and obtain a threshold for the speed of this convergence. Section 4 contains a collection of examples including the Curie-Weiss model with three states, modeling liquid helium, and a continuous Curie-Weiss model, where the single spin distribution $\varrho$ is a uniform distribution.

During the preparation of our manusscript we became aware of a preprint of S. Chatterjee ans Q.-M. Shao about Stein's method with applications to the Curie-Weiss model [4]. As far as we understand, there the authors give an alternative proof of Theorem 3.7 and 3.8.

## 2 The exchangeable pair approach for distributional approximations

We will begin with modifying Stein's method by replacing the linear regression property by

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

where $\psi(x)$ depends on a continuous distribution under consideration. Let us 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 [21], following [20, Chapter 6]. In [21], 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 [3, Lemma 2.1].

Given two random variables $X$ and $Y$ defined on a common probability space, we denote the Kolmogorov distance of the distributions of $X$ and $Y$ by

$$d_{\mathrm{K}}(X,Y) := \sup_{z\in\mathbb{R}} |P(X \le z) - P(Y \le z)|.$$

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. Let $I = (a,b)$ be a real interval, where $-\infty \le a < b \le \infty$. 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 $I$, has only countably many sign changes, and is continuous at the sign changes. Suppose moreover that $\int_I p(x)|\log(p(x))|\,dx < \infty$ and that $\psi(x) := \frac{p'(x)}{p(x)}$ is regular.

In [21] it is proved, that a random variable $Z$ is distributed according to the density $p$ if and only if $\mathbb{E}\big(f'(Z) + \psi(Z)f(Z)\big) = f(b-)p(b-) - f(a+)p(a+)$ for a suitably chosen class $\mathscr{F}$ of functions $f$. The corresponding Stein identity is

$$f'(x) + \psi(x)f(x) = h(x) - P(h), \tag{2.6}$$

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

$$f(x) = \frac{\int_a^x \big(h(y) - Ph\big)p(y)\,dy}{p(x)}. \tag{2.7}$$

For the function $h(x) := 1_{\{x\le z\}}(x)$ let $f_z$ be the corresponding solution of (2.6). We will make the following assumptions:

**Assumption (B1)** Let $p$ be a density fulfilling Assumption (D). We assume that for any absolutely continuous function $h$, the solution $f_h$ of (2.6) satisfies

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

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

$$f_z'(x) + \psi(x)f_z(x) = 1_{\{x\le z\}}(x) - P(z) \tag{2.8}$$

satisfies

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

and

$$|(\psi(x) f_z(x))'| = \Big|\Big(\frac{p'(x)}{p(x)} f_z(x)\Big)'\Big| \le d_4 \tag{2.9}$$

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

**Remark 2.1.** In the case of the normal approximation, $\psi(x) = -x$. Assumption (B2) includes a bound for $(x f_z(x))'$ for the solution $f_z$ of the classical Stein equation. It is easy to observe that $|(x f_z(x))'| \le 2$ by direct calculation (see [5, Proof of Lemma 6.5]). However, in the normal approximation case, using this bound leads to a worse Berry-Esseen constant. We will be able to improve the Berry-Esseen constant applying $d_2 = d_3 = 1$ and $d_1 = \sqrt{2\pi}/4$ (see Theorem 2.6 and Theorem 2.4).

We will see, that all limit laws in our class of Curie-Weiss models satisfy Condition (2.9):

**Lemma 2.2.** *The densities $f_{k,\mu,\beta}$ in* (3.35) *and* (3.36) *and the densities in Theorem 3.1 and Theorem 3.2 satisfy Assumptions (D), (B1) and (B2).*

*Proof.* We defer the proof to the appendix, since they only involve careful analysis. □

**Remark 2.3.** We restrict ourselves to solutions of the 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 2.2, one would also be able to derive good bounds (in the sense that Assumption (B1) and (B2) are fulfilled) even for measures with a probability density of the form

$$p(x) = b_k \exp\big(-a_k V(x)\big), \tag{2.10}$$

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 $x \ge d$ 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 3. A measure with density $p$ in (2.10) is usually called a Gibbs measure. Stein's method for discrete Gibbs measures is developed in [7]. Our remark might be of use for a potential application of Stein's method to Gibbs measures with continuous spins.

The following result is a refinement of Stein's result [20] for exchangeable pairs.

**Theorem 2.4.** *Let $p$ be a density fulfilling Assumption (D). Let $(W, W')$ be an exchangeable pair of real-valued random variables such that*

$$\mathbb{E}[W'|W] = W + \lambda \psi(W) - R(W) \tag{2.11}$$

*for some random variable $R = R(W)$, $0 < \lambda < 1$ and $\psi = p'/p$. Then*

$$\mathbb{E}(W - W')^2 = -2\lambda \mathbb{E}[W \psi(W)] + 2\mathbb{E}[W R(W)]. \tag{2.12}$$

*Moreover it holds for a random variable $Z$ distributed according to $p$:* **(1)***: Under Assumption (B1), for any uniformly Lipschitz function $h$, we obtain $|\mathbb{E}h(W) - \mathbb{E}h(Z)| \leq \delta \|h'\|$ with*

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

**(2)***: Under Assumption (B2), we obtain for any $A > 0$*

$$\begin{aligned} d_{\mathrm{K}}(W,Z) &\leq d_2 \sqrt{\mathbb{E}\left(1 - \frac{1}{2\lambda}\mathbb{E}[(W'-W)^2|W]\right)^2} + \big(d_1 + \frac{3}{2}A\big)\frac{\sqrt{\mathbb{E}(R^2)}}{\lambda} \\ &+ \frac{1}{\lambda}\big(\frac{d_4 A^3}{4}\big) + \frac{3A}{2}\mathbb{E}(|\psi(W)|) + \frac{d_3}{2\lambda}\mathbb{E}\big((W-W')^2 1_{\{|W-W'|\geq A\}}\big). \end{aligned} \quad (2.13)$$

With (2.12) 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 2.4 are only useful, if $-\mathbb{E}[W\psi(W)]$ is close to 1 and $\frac{\mathbb{E}(WR)}{\lambda}$ is small. Alternatively, bounds can be obtained by comparing with a modified distribution that 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 $\mathbb{E}\big(\mathbb{E}[W\psi(W)]f'(Z) + \psi(Z)f(Z)\big) = 0$ for a suitably chosen class of functions.

**Theorem 2.5.** *Let $p$ be a density fulfilling Assumption (D). Let $(W, W')$ be an exchangeable pair of random variables such that (2.11) 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*

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

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

$$\begin{aligned} d_{\mathrm{K}}(W,Z_W) &\leq \frac{d_2}{2\lambda}\big(\mathrm{Var}\big(\mathbb{E}[(W-W')^2|W]\big)^{1/2} + \big(d_1 + d_2\sqrt{\mathbb{E}(W^2)} + \frac{3}{2}A\big)\frac{\sqrt{\mathbb{E}(R^2)}}{\lambda} \\ &+ \frac{1}{\lambda}\big(\frac{d_4 A^3}{4}\big) + \frac{3A}{2}\mathbb{E}(|\psi(W)|) + \frac{d_3}{2\lambda}\mathbb{E}\big((W-W')^2 1_{\{|W-W'|\geq A\}}\big). \end{aligned} \quad (2.14)$$

*Proof of Theorem 2.4.* The proof is an adaption of the results in [20]. For any function $f$ such that $\mathbb{E}(\psi(W)f(W))$ exists we obtain

$$\begin{aligned} 0 &= \mathbb{E}\big((W-W')(f(W') + f(W))\big) \\ &= \mathbb{E}\big((W-W')(f(W') - f(W))\big) - 2\lambda\mathbb{E}(\psi(W)f(W)) + 2\mathbb{E}(f(W)R(W)), \end{aligned} \quad (2.15)$$

which is equivalent to

$$\mathbb{E}(\psi(W)f(W)) = -\frac{1}{2\lambda}\mathbb{E}\big((W-W')(f(W) - f(W'))\big) + \frac{1}{\lambda}\mathbb{E}(f(W)R(W)). \quad (2.16)$$

**Proof of (1):** Let $f = f_h$ be the solution of the Stein equation (2.6), and define

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

By (2.16), following the calculations on page 21 in [5], we obtain

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

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

$$|\mathbb{E}h(W)-\mathbb{E}h(Z)|\le\|h'\|\Big(c_2\mathbb{E}\Big|1-\frac{1}{2\lambda}\mathbb{E}\big((W-W')^2|W\big)\Big|+\frac{c_3}{4\lambda}\mathbb{E}|W-W'|^3+\frac{c_1}{\lambda}\sqrt{\mathbb{E}(R^2)}\Big).\quad (2.17)$$

**Proof of (2):** Now let $f=f_z$ be the solution of the Stein equation (2.8). Using (2.16), we obtain

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

Now the bounds in Assumption (B2) give

$$|T_1|\le d_2\sqrt{\mathbb{E}\Big(1-\frac{1}{\lambda}\mathbb{E}[(W'-W)^2|W]\Big)^2}\quad\text{and}\quad|T_2|\le\frac{d_1}{\lambda}\sqrt{\mathbb{E}(R^2)}.$$

Bounding $T_3$ we apply the concentration technique, see [17]:

$$\begin{aligned}(-2\lambda)\,T_3 &= \mathbb{E}\Big((W-W')1_{\{|W-W'|>A\}}\int_{-(W-W')}^{0}(f'(W+t)-f'(W))dt\Big) \\ &+ \mathbb{E}\Big((W-W')1_{\{|W-W'|\le A\}}\int_{-(W-W')}^{0}(f'(W+t)-f'(W))dt\Big).\end{aligned}\quad (2.18)$$

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

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

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

The term $U_2$ can be bounded by $\mathbb{E}\big((W-W')^2 I_{\{0\le(W-W')\le A\}}\, 1_{\{z\le W\le z+A\}}\big)$. Under the assumptions of our Theorem we proceed as in [17] and obtain the following concentration inequality:

$$\mathbb{E}\big((W-W')^2 I_{\{0\le(W-W')\le a\}}\, 1_{\{z\le W\le z+A\}}\big) \le 3A(\lambda\mathbb{E}(|\psi(W)|)+\mathbb{E}(|R|)). \tag{2.19}$$

To see this, we apply the estimate $\mathbb{E}\big((W-W')^2 I_{0\le(W-W')\le A} 1_{z\le W\le z+A}\big) \le \mathbb{E}\big((W-W')(f(W)-f(W'))\big)$, hence $U_2$ can be bounded by $\mathbb{E}\big((W-W')(f(W)-f(W'))\big) = 2\mathbb{E}\big(f(W)R(W)\big) - 2\lambda\,\mathbb{E}\big(\psi(W)f(W)\big)$, where we applied (2.16), and where $f$ is defined by $f(x):=-1.5A$ for $x\le z-A$, $f(x):=1.5A$ for $x\ge z+2A$ and $f(x):=x-z-A/2$ in between. Thus $U_2\le 3A\big(\mathbb{E}(|R|)+\lambda\mathbb{E}(|\psi(W)|)\big)$. Similarly we obtain $U_2\ge -3A\big(\mathbb{E}(|R|)+\lambda\mathbb{E}(|\psi(W)|)\big)$. □

*Proof of Theorem 2.5.* The main observation is the following identity:

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

with $T_3$ defined as in the proof of Theorem 2.4. We apply the Cauchy-Schwarz inequality to get $\mathbb{E}\Big|\mathbb{E}[(W-W')^2]-\mathbb{E}[(W-W')^2|W]\Big| \le \big(\mathrm{Var}\big(\mathbb{E}[(W-W')^2|W]\big)\big)^{1/2}$. The proof follows the lines of the proof of Theorem 2.4. □

In the special case of normal approximation, the following Theorem improves Theorem 1.2 in [16]:

**Theorem 2.6.** *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 $\mathbb{E}(W^2)\le 1$. Let $Z$ be a random variable with standard normal distribution. Then for any $A>0$,*

$$d_K(W,Z) \le \sqrt{\mathbb{E}\left(1-\frac{1}{2\lambda}\mathbb{E}[(W'-W)^2|W]\right)^2}+\left(\frac{\sqrt{2\pi}}{4}+1.5A\right)\frac{\sqrt{\mathbb{E}(R^2)}}{\lambda} \tag{2.20}$$

$$+\frac{0.41A^3}{\lambda}+1.5A+\frac{1}{2\lambda}\mathbb{E}\big((W-W')^2 1_{\{|W-W'|\ge A\}}\big). \tag{2.21}$$

**Remark 2.7.** When $|W-W'|$ is bounded, our estimate improves (1.10) in [16, Theorem 1.2] with respect to the Berry-Esseen constants.

*Proof.* Let $f=f_z$ denote the solution of the Stein equation

$$f_z'(x)-xf_z(x)=1_{\{x\le z\}}(x)-\Phi(z). \tag{2.22}$$

We obtain

$$\begin{aligned}P(W\le z)-\Phi(z) &= \mathbb{E}\left(f'(W)\big(1-\frac{1}{2\lambda}(W-W')^2\big)\right)-\frac{1}{2\lambda}\mathbb{E}(2f(W)R)\\ &\quad -\frac{1}{2\lambda}\mathbb{E}\Big[(W-W')\big(f(W)-f(W')-(W-W')f'(W)\big)\Big]\\ &=: T_1+T_2+T_3.\end{aligned} \tag{2.23}$$

Using $|f'(x)| \le 1$ for all real $x$ (see [5, Lemma 2.2]), we obtain the bound $|T_1| \le \big(\mathbb{E}\big(1-\frac{1}{2\lambda}\mathbb{E}[(W'-W)^2|W]\big)^2\big)^{1/2}$. Using $0 < f(x) \le \sqrt{2\pi}/4$ (see [5, Lemma 2.2]), we have $|T_2| \le \frac{\sqrt{2\pi}}{4\lambda}\mathbb{E}(|R|) \le \frac{\sqrt{2\pi}}{4\lambda}\sqrt{\mathbb{E}(R^2)}$. Bounding $T_3$ we apply (2.18) for $(-2\lambda)T_3$. The modulus of the first term can be bounded by $\mathbb{E}\big((W-W')^2 1_{\{|W-W'|>A\}}\big)$ using $|f'(x)-f'(y)| \le 1$ for all real $x$ and $y$ (see [5, Lemma 2.2]). Using the Stein identity (2.22), the second summand can be represented as

$$\mathbb{E}\left((W-W')1_{\{|W-W'|\le A\}}\int_{-(W-W')}^{0}\big((W+t)f(W+t)-Wf(W)\big)dt\right)$$
$$+\mathbb{E}\left((W-W')1_{\{|W-W'|\le A\}}\int_{-(W-W')}^{0}(1_{\{W+t\le z\}}-1_{\{W\le z\}})dt\right) =: U_1+U_2.$$

Next observe that $|U_1| \le 0.82A^3$, see [17]: By the mean value theorem one gets

$$(W+t)f(W+t)-Wf(W) = W(f(W+t)-f(W))+tf(W+t) = W\Big(\int_0^1 f'(W+ut)t\,du\Big)+tf(W+t).$$

Hence $|(W+t)f(W+t)-Wf(W)| \le |W||t|+|t|\sqrt{2\pi}/4 = |t|(\sqrt{2\pi}/4+|W|)$. Using $\mathbb{E}|W| \le \sqrt{\mathbb{E}(W^2)} \le 1$ gives the bound. The term $U_2$ can be bounded by $\mathbb{E}\big((W-W')^2 I_{\{0\le(W-W')\le A\}}1_{\{z\le W\le z+A\}}\big)$ and $\mathbb{E}\big((W-W')^2 I_{\{0\le(W-W')\le a\}}1_{\{z\le W\le z+A\}}\big) \le 3A(\lambda+\mathbb{E}(R))$, see [17]. Similarly, we obtain $U_2 \ge -3A(\lambda+\mathbb{E}(R))$. □

**Remark 2.8.** In Theorem 2.6, we assumed $\mathbb{E}(W^2) \le 1$. Alternatively, let us assume that $\mathbb{E}(W^2)$ is finite. Then the proof of Theorem 2.6 shows, that the third and the fourth summand of the bound (2.20) change to $\frac{A^3}{\lambda}\big(\frac{\sqrt{2\pi}}{16}+\frac{\sqrt{\mathbb{E}(W^2)}}{4}\big)+1.5A\mathbb{E}(|W|)$.

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

**Corollary 2.9.** *Let $\sigma^2 > 0$ and $(W,W')$ be an exchangeable pair of real-valued random variables such that*

$$E(W'|W) = \big(1-\frac{\lambda}{\sigma^2}\big)W+R \tag{2.24}$$

*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 $|W-W'| \le A$ for a constant $A$, we obtain the bound*

$$\begin{aligned} d_K(W,Z_\sigma) \;\le\; & \sqrt{\mathbb{E}\left(1-\frac{1}{2\lambda}\mathbb{E}[(W'-W)^2|W]\right)^2}+\left(\frac{\sigma\sqrt{2\pi}}{4}+1.5A\right)\frac{\sqrt{\mathbb{E}(R^2)}}{\lambda} \\ & + \frac{A^3}{\lambda}\left(\frac{\sqrt{2\pi\sigma^2}}{16}+\frac{\sqrt{\mathbb{E}(W^2)}}{4}\right)+1.5A\sqrt{\mathbb{E}(W^2)}. \end{aligned} \tag{2.25}$$

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

$$f'_{\sigma,z}(x)-\frac{x}{\sigma^2}f_{\sigma,z}(x) = 1_{\{x\le z\}}(x)-F_\sigma(z) \tag{2.26}$$

with $F_\sigma(z) := \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{z}\exp\left(-\frac{y^2}{2\sigma^2}\right)dy$. It is easy to see that the identity $f_{\sigma,z}(x) = \sigma f_z\left(\frac{x}{\sigma}\right)$, where $f_z$ is the solution of the corresponding Stein equation of the standard normal distribution, holds true. Using [5, Lemma 2.2] we obtain $0 < f_\sigma(x) < \sigma\frac{\sqrt{2\pi}}{4}$, $|f'_\sigma(x)| \le 1$, and $|f'_\sigma(x) - f'_\sigma(y)| \le 1$. With (2.24) we arrive at $P(W \le z) - F_\sigma(z) = T_1 + T_2 + T_3$ with $T_i$'s defined in (2.23). Using the bounds of $f_\sigma$ and $f'_\sigma$, the bound of $T_1$ is the same as in the proof of Theorem 2.6, whereas the bound of $T_2$ changes to $|T_2| \le \sigma\frac{\sqrt{2\pi}}{4\lambda}\sqrt{\mathbb{E}(R^2)}$. Since we consider the case $|W - W'| \le A$, we have to bound

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

Along the lines of the proof of Theorem 2.6, we obtain $|T_3| \le \frac{A^3}{\lambda}\left(\frac{\sqrt{\mathbb{E}(W^2)}}{4} + \frac{\sigma\sqrt{2\pi}}{16}\right) + 1.5A\left(\sqrt{\mathbb{E}(W^2)} + \frac{\sqrt{\mathbb{E}(R^2)}}{\lambda}\right)$. Hence the corollary is proved. □

With (2.24) we obtain $\mathbb{E}(W - W')^2 = \frac{2\lambda}{\sigma^2}\mathbb{E}(W^2) - 2\mathbb{E}(WR)$. Therefore

$$\mathbb{E}\left(1 - \frac{1}{2\lambda}\mathbb{E}[(W' - W)^2|W]\right) = 1 - \frac{\mathbb{E}(W^2)}{\sigma^2} + \frac{\mathbb{E}(WR)}{\lambda}, \tag{2.27}$$

so that the bound in Corollary 2.9 is only useful when $\mathbb{E}(W^2)$ is close to $\sigma^2$ (and $\mathbb{E}(WR)/\lambda$ is small). An alternative bound can be obtained comparing with a $N(0, \mathbb{E}(W^2))$-distribution.

**Corollary 2.10.** *In the situation of Corollary 2.9, let $Z_W$ denote the $N(0, \mathbb{E}(W^2))$ distribution. We obtain*

$$\begin{aligned} d_{\mathrm{K}}(W, Z_W) \;\le\; & \frac{\sigma^2}{2\lambda}\left(\mathrm{Var}\left(\mathbb{E}[(W' - W)^2|W]\right)\right)^{1/2} + \sigma^2\left(\frac{\sqrt{\mathbb{E}(W^2)}\sqrt{2\pi}}{4} + 1.5A\right)\frac{\sqrt{\mathbb{E}(R^2)}}{\lambda} \\ & + \sigma^2\frac{A^3}{\lambda}\left(\frac{\sqrt{\mathbb{E}(W^2)}\sqrt{2\pi}}{16} + \frac{\sqrt{\mathbb{E}(W^2)}}{4}\right) + \sigma^2\, 1.5A\sqrt{\mathbb{E}(W^2)} + \sigma^2\frac{\sqrt{\mathbb{E}(W^2)}\sqrt{\mathbb{E}(R^2)}}{\lambda}. \end{aligned} \tag{2.28}$$

*Proof.* With (2.27) we get $\mathbb{E}(W^2) = \sigma^2\left(\frac{1}{2\lambda}(E(W - W')^2 + 2\mathbb{E}(WR))\right)$. With the definition of $T_2$ and $T_3$ as in (2.23) we obtain

$$\begin{aligned} \mathbb{E}\left(\mathbb{E}(W^2)f'(W) - Wf(W)\right) &= \sigma^2\mathbb{E}\left(\frac{\mathbb{E}(W - W')^2 + 2\mathbb{E}(WR)}{2\lambda}f'(W)\right) - \mathbb{E}(Wf(W)) \\ = \sigma^2\mathbb{E}\left(f'(W)\left(\frac{\mathbb{E}(W - W')^2 - \mathbb{E}[(W - W')^2|W]}{2\lambda}\right)\right) &+ \sigma^2(T_2 + T_3) + \sigma^2\frac{\mathbb{E}(WR)}{\lambda}. \end{aligned} \tag{2.29}$$

Remark that now $\sigma^2$ in (2.24) is a parameter of the exchangeable-pair identity and no longer the parameter of the limiting distribution. We apply (2.26) and exchange every $\sigma^2$ in (2.26) with $\mathbb{E}(W^2)$. Applying Cauchy-Schwarz to the first summand and bounding the other terms as in the proof of Corollary 2.9 leads to the result. □

## 3 Berry-Esseen bounds for Curie-Weiss models

We assume that $\varrho$ in (1.1) is in the class $\mathscr{B}$ of non-degenerate symmetric Borel probability measures on $\mathbb{R}$ which satisfy

$$\int \exp\left(\frac{b\,x^2}{2}\right) d\varrho(x) < \infty \quad \text{for all} \quad b > 0. \tag{3.30}$$

For technical reasons we introduce a model with non-negative external magnetic field, where the strength may even depend on the site:

$$P_{n,\beta,h_1,\dots,h_n}(x) = \frac{1}{Z_{n,\beta,h_1,\dots,h_n}} \exp\big(\frac{\beta}{2n} S_n^2 + \beta \sum_{i=1}^n h_i\, x_i\big)\, d\varrho^{\otimes n}(x), \quad x = (x_i). \tag{3.31}$$

In the general case (1.1), we will see (analogously to the results in [11; 13]) 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 $\phi_\varrho(s) := \log \int \exp(s\,x)\, d\varrho(x)$ and

$$G_\varrho(\beta, s) := \frac{\beta\, s^2}{2} - \phi_\varrho(\beta\, s). \tag{3.32}$$

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 $\varrho \in \mathscr{B}$, $G$ was proved to have global minima, which can be only finite in number, see [11, 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

$$G(s) = G(\alpha) + \frac{\mu(\alpha)(s-\alpha)^{2k}}{(2k)!} + \mathscr{O}((s-\alpha)^{2k+1}) \quad \text{as} \quad s \to \alpha. \tag{3.33}$$

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

$$\mu(\alpha) = \beta - \beta^2 \phi''(\beta\,\alpha) \quad \text{if } k = 1 \quad \text{while} \qquad \mu(\alpha) = -\beta^{2k} \phi^{(2k)}(\beta\,\alpha) \quad \text{if } k \geq 2 \tag{3.34}$$

(see [13]). 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 [13]. In general, given $\varrho \in \mathscr{B}$, let $\alpha$ be one of the global minima of maximal type $k$ and strength $\mu$ of $G_\varrho$. Then $\frac{S_n - n\alpha}{n^{1-1/2k}} \to X_{k,\mu,\beta}$ in distribution, where $X_{k,\mu,\beta}$ is a random variable with probability density $f_{k,\mu,\beta}$, defined by

$$f_{1,\mu,\beta}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\big(-x^2/2\sigma^2\big) \tag{3.35}$$

and for $k \geq 2$

$$f_{k,\mu,\beta}(x) = \frac{\exp\big(-\mu x^{2k}/(2k)!\big)}{\int \exp\big(-\mu x^{2k}/(2k)!\big)\, dx}. \tag{3.36}$$

Here, $\sigma^2 = \frac{1}{\mu} - \frac{1}{\beta}$ so that for $\mu = \mu(\alpha)$ as in (3.34), $\sigma^2 = ([\phi''(\beta\alpha)]^{-1} - \beta)^{-1}$ (see [11], [13]). Moderate deviation principles have been investigated in [6].

In [10] and [13], a class of measures $\varrho$ is described with a behaviour similar to that of the classical Curie–Weiss model. Assume that $\varrho$ is any symmetric measure that satisfies the GHS-inequality,

$$\frac{d^3}{ds^3}\phi_\varrho(s) \le 0 \quad \text{for all } s \ge 0, \tag{3.37}$$

(see also [12; 14]). 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 $0 < \beta \le \beta_c$ and exactly two global minima, of equal type, for $\beta > \beta_c$. For $\beta_c$ the unique global minimum is of type $k \ge 2$ whereas for $\beta \in (0, \beta_c)$ the unique global minimum is of type 1. At $\beta_c$ the fluctuations of $S_n$ live on a larger scale than $\sqrt{n}$. This critical temperature can be explicitly computed as $\beta_c = 1/\phi''(0) = 1/\operatorname{Var}_\varrho(X_1)$. By rescaling the $X_i$ we may thus assume that $\beta_c = 1$.

By GHS we denote the set of measures $\varrho \in \mathscr{B}$ such that the GHS-inequality (3.37) is valid. We will be able to obtain Berry-Esséen-type results for $\varrho$-a.s. bounded single-spin variables $X_i$:

**Theorem 3.1.** *Given $\varrho \in \mathscr{B}$ 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*

$$d_K\Big(\frac{S_n}{\sqrt{n}}, Z_W\Big) \le C n^{-1/2}, \tag{3.38}$$

*where $Z_W$ denotes a random variable distributed according to he normal distribution with mean zero and variance $\mathbb{E}(W^2)$ and $C$ is an absolute constant depending on $0 < \beta < 1$. For $k \ge 2$ we obtain*

$$d_K\Big(\frac{S_n - n\alpha}{n^{1-1/2k}}, Z_{W,k}\Big) \le C_k\, n^{-1/k}, \tag{3.39}$$

*where $Z_{W,k}$ denotes a random variable distributed according to the density $\widehat{f}_{W,k}$ defined by $\widehat{f}_{W,k}(x) := C \exp\big(-\frac{x^{2k}}{2k\,\mathbb{E}(W^{2k})}\big)$ with $C^{-1} = \int \exp\big(-\frac{x^{2k}}{2k\,\mathbb{E}(W^{2k})}\big)\,dx$ and $W := \frac{S_n - n\alpha}{n^{1-1/2k}}$ and $C_k$ is an absolute constant.*

**Theorem 3.2.** *Let $\varrho \in \mathscr{B}$ satisfy the GHS-inequality and assume that $\beta_c = 1$. Let $\alpha$ be the global minimum of type $k$ with $k \ge 2$ 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 $\beta_n \to 1$ monotonically as $n \to \infty$. Then the following assertions hold:* **(1)***: If $\beta_n - 1 = \frac{\gamma}{n^{1-\frac{1}{k}}}$ for some $\gamma \ne 0$, we have*

$$\sup_{z \in \mathbb{R}} \left| P_n\left(\frac{S_n - n\alpha}{n^{1-1/2k}} \le z\right) - F_{W,k,\gamma}(z) \right| \le C_k\, n^{-1/k} \tag{3.40}$$

*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 $Z := \int_{\mathbb{R}} \exp\big(-c_W^{-1}\big(\frac{\mu_k}{(2k)!}x^{2k} - \frac{\gamma}{2}x^2\big)\big)\,dx$, 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 3.1. Moreover, if $|\beta_n - 1| = \mathscr{O}(n^{-1})$, (3.39) 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*

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

*with an absolute constant $C$.*

For arbitrary $\varrho \in \mathrm{GHS}$ we are able to prove good bounds with respect to smooth test functions $h$. For technical reasons, we consider a modified model. Let

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

**Theorem 3.3.** *Given the Curie-Weiss model $\widehat{P}_{n,\beta}$ and $\varrho \in \mathscr{B}$ 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*

$$\left|\mathbb{E}\big(h(W)\big) - \Phi_W(h)\right| \le \|h'\|\, C\, \frac{\max\big(\mathbb{E}|X_1|^3, \mathbb{E}|X_1'|^3\big)}{\sqrt{n}}.$$

*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 $k \ge 2$ we obtain for any uniformly Lipschitz function $h$ and for $W := \frac{S_n - n\alpha}{n^{1-1/2k}}$*

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

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

**Remark 3.4.** Note that for $\varrho \in \mathrm{GHS}$, $\phi_\varrho(s) \le \frac{1}{2}\sigma_\varrho^2 s^2$ 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 [15]: If $\mathbb{E}$ denotes the expectation with respect to the measure $P_{n,\beta,h_1,\dots,h_n}$ in (3.31), one observes easily that for any $\varrho \in \mathscr{B}$ and sites $i, j, k, l \in \{1, \dots, n\}$ the following identity holds:

$$\begin{aligned} &\frac{\partial^3}{\partial h_i\, \partial h_j\, \partial h_k} \mathbb{E}(X_l)\bigg|_{\text{all } h_i = 0} \qquad (3.41) \\ = \;& \mathbb{E}(X_i X_j X_k X_l) - \mathbb{E}(X_i X_j)\mathbb{E}(X_k X_l) - \mathbb{E}(X_i X_k)\mathbb{E}(X_j X_l) - \mathbb{E}(X_i X_l)\mathbb{E}(X_j X_k). \end{aligned}$$

Lebowitz [15] proved that if $\varrho \in \mathrm{GHS}$, (3.41) is non-positive (see [9, V.13.7.(b)]). We will see in the proofs of Theorem 3.1 and Theorem 3.2 Lebowitz' inequality helps to bound the variances.

In the situation of Theorem 3.1 and Theorem 3.2 we can bound higher order moments as follows:

**Lemma 3.5.** *Given $\varrho \in \mathscr{B}$, let $\alpha$ be one of the global minima of maximal type $k$ for $k \ge 1$ and strength $\mu$ of $G_\varrho$. For $W := \frac{S_n - n\alpha}{n^{1-1/2k}}$ we obtain that $\mathbb{E}|W|^l \le \mathrm{const.}(l)$ for any $l \in \mathbb{N}$ .*

We prepare for the proof of Lemma 3.5. It considers a well known transformation – sometimes called the *Hubbard–Stratonovich transformation* – of our measure of interest.

**Lemma 3.6.** *Let $m \in \mathbb{R}$ and $0 < \gamma < 1$ be real numbers. Consider the measure $Q_{n,\beta} := \left(P_n \circ \left(\frac{S_n - nm}{n^\gamma}\right)^{-1}\right) * \mathcal{N}(0, \frac{1}{\beta n^{2\gamma-1}})$ 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*

$$\frac{\exp\left(-nG(\frac{s}{n^{1-\gamma}} + m)\right)}{\int_{\mathbb{R}} \exp\left(-nG(\frac{s}{n^{1-\gamma}} + m)\right) ds}, \tag{3.42}$$

*where $G$ is defined in equation* (3.32).

As shown in [11], Lemma 3.1, our condition (3.30) 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. The proof of the lemma can be found at many places, e.g. in [11], Lemma 3.3.

*Proof of Lemma 3.5.* 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 (3.33) of $G$, we see that the density of $Q_{n,\beta}$ with respect to Lebesgue measure is given by Const. $\exp(-x^{2k})$ (up to negligible terms, see e.g. [11], [6]). A measure with this density, of course, has moments of any finite order. □

Since the symmetric Bernoulli law is GHS, Theorems 3.1 and 3.2 include Berry-Esseen type results for the classical model. However the limiting laws depend on moments of $W$. Approximations with fixed limiting laws can be obtained in the classical case, since we can apply Corollary 2.9 and part (2) of Theorem 2.4:

**Theorem 3.7** (classical Curie-Weiss model, 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*

$$\sup_{z \in \mathbb{R}} \left| P_n\left(S_n/\sqrt{n} \leq z\right) - \Phi_\beta(z) \right| \leq C\, n^{-1/2}, \tag{3.43}$$

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

$$\sup_{z \in \mathbb{R}} \left| P_n\left(S_n/n^{3/4} \leq z\right) - F(z) \right| \leq C\, n^{-1/2}, \tag{3.44}$$

*where $F(z) := \frac{1}{Z}\int_{-\infty}^{z} \exp(-x^4/12)\, dx$, $Z := \int_{\mathbb{R}} \exp(-x^4/12)\, dx$ and $C$ is an absolute constant.*

Berry-Esséen bounds for size-dependent temperatures for $\varrho = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$ and $0 < \beta_n < \infty$ depending on $n$ in such a way that $\beta_n \to 1$ monotonically as $n \to \infty$ can be formulated similarly to the results in Theorem 3.2.

**Remark 3.9.** 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 [11]. As far as we understand, his result (3.9) can be interpreted as the statement (3.44), but with the rate $n^{-1/4}$.

*Proof of Theorem 3.1.* 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 [9, V.13.15]. Hence in this case we will apply Corollary 2.10 (to obtain better constants for our Berry-Esséen bound in comparison to Theorem 2.5). 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 the Introduction, such that $W - W' = \frac{X_I - X_I'}{n^{1-1/(2k)}}$. We obtain

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

**Lemma 3.10.** *In the situation of Theorem 3.1, if $X_1$ is $\varrho$-a.s. bounded, we obtain*

$$\mathbb{E}(X_i' | \mathscr{F}) = \left( m_i(X) - \frac{1}{\beta} G_\varrho'(\beta, m_i(X)) \right) \left( 1 + \mathscr{O}(1/n) \right)$$

*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:

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

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

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

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

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

and

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

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

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

□

**Remark 3.11.** 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 $\mathbb{E}(X_i'|\mathscr{F}) = \left(m_i(X) - \frac{1}{\beta} G_\varrho'(\beta, m_i(X))\right)$ without the boundedness assumption for the $X_1$.

Applying Lemma 3.10 and the presentation (3.33) of $G_\varrho$, it follows that

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

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

$$\frac{1}{n^{2-\frac{1}{k}}} \frac{\mu}{\beta(2k-1)!} W^{2k-1} = -\frac{1}{n^{2-\frac{1}{k}}} \psi(W). \tag{3.45}$$

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, (3.45) 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

$$R(W) = \frac{1}{n^2} W + \frac{\mu}{\beta(2k-1)!} \sum_{l=1}^{2k-1} \binom{2k-1}{l} \frac{1}{n^{2-\frac{1}{k}-\frac{l}{2k}}} W^{2k-1-l} \frac{(-1)^l}{n^l} \frac{1}{n} \sum_{i=1}^n X_i^l + \mathcal{O}(\frac{W^{2k}}{n}).$$

With Lemma 3.5 we know that $\mathbb{E}|W|^{2k} \leq \text{const.}$ 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 (2.28) of Corollary 2.10 are $\mathcal{O}(n^{-1/2})$. For $k \geq 2$ we obtain $\frac{3A}{2}\mathbb{E}(|\psi(W)|) = \mathcal{O}\left(n^{\frac{1}{2k}-1}\right)$ and $\frac{1}{\lambda}\left(\frac{d_4 A^3}{4}\right) = \mathcal{O}\left(n^{\frac{1}{2k}-1}\right)$. 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

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

Hence the last four summands in (2.14) of Theorem 2.5 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

$$\frac{1}{2n}\sum_{i=1}^{m} X_i^2 + \frac{1}{2n}\sum_{i=1}^{n} E[(X_i')^2|\mathscr{F}] + \frac{1}{n}\sum_{i=1}^{n} X_i\left(m_i(X) - \frac{1}{\beta}G_\varrho'(\beta, m_i(X))\right)(1+O(1/n)). \tag{3.46}$$

Since we assume that $\varrho \in$ GHS, we can apply the correlation-inequality due to Lebowitz (see Remark 3.4) $\mathbb{E}(X_iX_jX_kX_l) - \mathbb{E}(X_iX_j)\mathbb{E}(X_kX_l) - \mathbb{E}(X_iX_k)\mathbb{E}(X_jX_l) - \mathbb{E}(X_iX_l)\mathbb{E}(X_jX_k) \le 0$. The choice $i=k$ and $j=l$ leads to the bound $\mathrm{cov}(X_i^2, X_j^2) = \mathbb{E}(X_i^2X_j^2) - \mathbb{E}(X_i^2)\mathbb{E}(X_j^2) \le 2(\mathbb{E}(X_i X_j))^2$. With Lemma 3.5 we know that $(\mathbb{E}(X_iX_j))^2 \le \mathrm{const}.n^{-2/k}$. This gives

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

Using a conditional version of Jensen's inequality we have

$$\mathrm{Var}\big(\mathbb{E}\big(\frac{1}{2n}\sum_{i=1}^{n}X_i^2\big|\mathscr{F}\big)\big) \le \mathrm{Var}\big(\frac{1}{2n}\sum_{i=1}^{n}X_i^2\big).$$

Hence the variance of the second term in (3.46) is of the same order as the variance of the first term. Applying (3.33) for $G_\varrho$, the variance of the third term in (3.46) is of the order of the variance of $W^2/n^{1/k}$. Summarizing the variance of (3.46) can be bounded by 9 times the maximum of the variances of the three terms in (3.46), which is a constant times $n^{-2/k}$, and therefore for $k \ge 1$ we obtain

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

Note that for $k \ge 2$ $\frac{\psi(x)}{-\mathbb{E}[W\psi(W)]} = -\frac{x^{2k-1}}{\mathbb{E}(W^{2k})}$. Hence we compare the distribution of $W$ with a distribution with Lebesgue-probability density proportional to $\exp\big(-\frac{x^{2k}}{2k\mathbb{E}(W^{2k})}\big)$. □

*Proof of Theorem 3.2.* Since $\alpha = 0$ and $k=1$ for $\beta \ne 1$ while $\alpha = 0$ and $k \ge 2$ for $\beta = 1$, $G_\varrho(\cdot)$ can now be expanded as

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

Hence $\frac{1}{\beta_n}G_\varrho'(s) = \frac{\mu_1}{\beta_n}s + \frac{\mu_k}{\beta_n(2k-1)!}s^{2k-1} + \mathcal{O}(s^{2k})$. With Lemma 3.10 and $\mu_1 = (1-\beta_n)\beta_n$ we obtain

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

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 3.1 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

$$\mathbb{E}[W - W'|W] = -\frac{1}{n^{2-1/k}}\psi(W) + \frac{\beta_n}{n^2}W + R(\beta_n, W), \tag{3.47}$$

where $\psi(x) = \gamma x - \frac{\mu_k}{\beta_n (2k-1)!}x^{2k-1}$. As in the proof of Theorem 3.1 we obtain that $R(\beta_n, W) = \mathcal{O}(n^{-2})$. Now we only have to adapt the proof of Theorem 3.1 step by step, applying Lemma 3.5, Lemma 2.2 and Theorem 2.5.

Let $|\beta_n - 1| = \mathcal{O}(1/n)$ and $W = n^{1/(2k)-1}\sum_{i=1}^n X_i$. Now in (3.47), 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 $\psi(x) = -\frac{\mu_k}{(2k-1)!}x^{2k-1}$. Along the lines of the proof of Theorem 3.1, 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 $\lim_{n\to\infty}\frac{1}{\lambda}\frac{(1-\beta_n)}{n} = \beta_n(1-\beta_n)n^{1-1/k} = 0$. Thus with Theorem 2.5 we obtain convergence in distribution for any $\beta_n$ with $|\beta_n - 1| \ll n^{-(1-1/k)}$. Moreover we obtain the Berry-Esséen 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 2.10. With $A := \frac{\text{const.}(1-\beta_\text{n})^{1/2}}{\sqrt{n}}$ we obtain

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

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 distributuion. Under the additional assumption that $|\beta_n - 1| \gg n^{-(1/2-1/(2k))}$ we obtain the Berry-Esséen bound. □

*Proof of Theorem 3.3.* We apply Theorem 2.5. For unbounded spin variables $X_i$ we consider $\widehat{P}_{n,\beta}$ and apply Lemma 3.10 to bound $\frac{1}{\lambda}\sqrt{\text{Var}(\mathbb{E}[(W - W')^2|W])}$ exactly as in the proof of Theorem 3.1. By Theorem 2.5 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 $\mathbb{E}|X_1^2 X_1'| \leq \left(\mathbb{E}|X_1|^3\right)^{2/3} \left(\mathbb{E}|X_1'|^3\right)^{1/3} = \mathbb{E}|X_1|^3$. Hence we have $\frac{1}{\lambda}\mathbb{E}|W - W'|^3 \leq \frac{8}{n^{1-1/2k}}\mathbb{E}|X_i|^3$. □

*Proof of Theorem 3.7.* With $\frac{1}{\sqrt{n}}\frac{1}{n}\sum_{i=1}^n \tanh(\beta m_i(X)) = \frac{1}{\sqrt{n}}\frac{1}{n}\sum_{i=1}^n \big(\tanh(\beta m_i(X)) - \tanh(\beta m(X))\big) + \frac{1}{\sqrt{n}}\tanh(\beta m(X)) =: R_1 + R_2$, where $m(X) := \frac{1}{n}\sum_{i=1}^n X_i$, Taylor-expansion $\tanh(x) = x + \mathcal{O}(x^3)$ leads to

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

Hence

$$\mathbb{E}[W - W'|W] = \frac{1-\beta}{n}W + R = \frac{\lambda}{\sigma^2}W + R \tag{3.48}$$

with $\lambda := \frac{1}{n}$, $\sigma^2 := (1-\beta)^{-1}$ and $R := \mathcal{O}\big(\frac{W^3}{n^2}\big) - R_1$. Since $|W - W'| = \left|\frac{X_I - X_I'}{\sqrt{n}}\right| \leq \frac{1}{\sqrt{n}} =: A$, we are able to apply Corollary 2.9. From Lemma 3.5 we know that $\mathbb{E}(W^4) \leq$ const.. Hence the fourth term in (2.25) 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 (2.25) can be estimated as follows: $\frac{A^3}{\lambda}\left(\frac{\sqrt{2\pi}}{16}\sqrt{(1-\beta)} + \frac{\text{const.}}{4}(1-\beta)\right) \leq \frac{1}{\sqrt{n}}\sqrt{(1-\beta)}$const.. Moreover we obtain $\mathbb{E}|R| \leq \mathbb{E}|R_1| + \mathcal{O}\big(\frac{\mathbb{E}|W^3|}{n^2}\big)$. 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 3.5, we get $\mathbb{E}|R| = \mathcal{O}\big(\frac{1}{n^{3/2}}\big)$ and thus, the second summand in (2.25) can be bounded by const.$\left(\frac{\sqrt{2\pi}}{4\sqrt{(1-\beta)}} + 1.5\frac{1}{\sqrt{\mathrm{n}}}\right)\frac{1}{\sqrt{\mathrm{n}}} = \mathcal{O}\big(\frac{1}{\sqrt{\mathrm{n}}}\big)$. To bound the first summand in (2.25), we obtain $(W - W')^2 = \frac{X_I^2}{n} - \frac{2X_I X_I'}{n} + \frac{X_I'}{n}$. Hence $\mathbb{E}\big[(W - W')^2|\mathscr{F}\big] = \frac{2}{n} - \frac{2}{n^2}\sum_{i=1}^n X_i \tanh(\beta m_i(X))$, and therefore

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

By Taylor expansion we get $R_2 = \frac{\beta}{n}W^2 + \mathcal{O}\big(\frac{W^4}{n^2}\big)$ and using Lemma 3.5 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 3.7 is proved. □

*Proof of Theorem 3.8.* We obtain

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

with an $S(W)$ such that $\mathbb{E}(S(W)) = O(1/n^2)$. Using this we get the exchangeable pair identity (1.5) with $R(W) = \mathcal{O}\big(\frac{1}{n^2}\big)$. With Lemma 2.2, we can now apply Theorem 2.4, using $|W - W'| \leq \frac{1}{n^{3/4}} =: A$. We obtain $1.5A\mathbb{E}(|\psi(W)|) \leq \text{const.}\frac{1}{\mathrm{n}^{3/4}}$ and $\frac{d_4 A^3}{4\lambda} = \frac{d_4}{4}\frac{1}{n^{3/4}}$. Using $\mathbb{E}|R(W)| \leq \text{const.}\frac{1}{\mathrm{n}^2}$ we get $\big(d_1 + \frac{3}{2}A\big)\frac{\mathbb{E}|R(W)|}{\lambda} \leq \text{const.}\frac{1}{\sqrt{\mathrm{n}}}$. Moreover we obtain

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

Hence applying Theorem 2.4 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 3.5, the leading term of $T$ is $\frac{W^2}{n^{1/2}}$. Hence $\mathbb{E}(T) = \mathcal{O}(n^{-1/2})$ and Theorem 3.8 is proved. □

# 4 Examples

It is known that the following distributions $\varrho$ are GHS (see [10, Theorem 1.2]). The symmetric Bernoulli measure is GHS, as first noted in [8]. The family of measures $\varrho_a(dx) = a\,\delta_x + \big((1-a)/2\big)\big(\delta_{x-1} + \delta_{x+1}\big)$ for $0 \le a \le 2/3$ is GHS, whereas the GHS-inequality fails for $2/3 < a < 1$, see [19, p.153]. GHS contains all measures of the form $\varrho_V(dx) := \big(\int_{\mathbb{R}} \exp\big(-V(x)\big)\,dx\big)^{-1} \exp\big(-V(x)\big)\,dx$, where $V$ is even, continuously differentiable, and unbounded above at infinity, and $V'$ is convex on $[0,\infty)$. GHS contains all absolutely continuous measures $\varrho \in \mathscr{B}$ with support on $[-a,a]$ for some $0 < a < \infty$ provided $g(x) = d\varrho/dx$ is continuously differentiable and strictly postitive on $(-a,a)$ and $g'(x)/g(x)$ is concave on $[0,a)$. Measures like $\varrho(dx) = \text{const.} \exp\big(-ax^4 - bx^2\big)\,dx$ or $\varrho(dx) = \text{const.} \exp\big(-a\cosh x - bx^2\big)\,dx$ with $a > 0$ and $b$ real are GHS. Both are of physical interest, see [10] and references therein).

**Example 4.1** (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. 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 [22]. In [2] it was used to analyze the tri-critical point of liquid helium. A little computation shows that $\frac{d^3}{ds^3}\phi_\varrho(s) \le 0$. for all $s \ge 0$. Hence the GHS-inequality (3.37) is fulfilled (see also [10, Theorem 1.2]), which implies that there is one critical temperature $\beta_c$ such that there is one minimum of $G$ for $\beta \le \beta_c$ 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 $\beta \le \beta_c$ 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 3.1 and 3.2 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 $\sigma^2 = \frac{1}{1-\beta}$ and the Berry-Esseen rate is $\frac{C}{\sqrt{n}}$. 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. The rate is $\frac{C}{n^{1/3}}$. 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 $\exp\big(-c_W^{-1}\left(\frac{x^6}{120} - \gamma\frac{x^2}{2}\right)\big)$ with $c_W = \frac{1}{120}\mathbb{E}(W^6) - \gamma\mathbb{E}(W^2)$. The rate is $\frac{C}{n^{1/3}}$.

**Example 4.2** (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 $\frac{d\varrho(x_i)}{d\,x_i} = \frac{1}{2a}\mathbb{I}_{[-a,a]}(x_i)$ with $a = \sqrt{3}$. Then from a general result in [12, Theorem 2.4] (see also [10, Theorem 1.2]) it follows that $\varrho(x_i)$ obeys the GHS-inequality (3.37). 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 3.1 and 3.2 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 $-\mathbb{E}_\varrho(X_1^4) + 3\mathbb{E}_\varrho(X_1^2) = -\frac{9}{5} + 3 = \frac{6}{5}$. The rate is $\frac{C}{n^{1/4}}$. 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 $c_W = \frac{5}{6(8!)}\mathbb{E}(W^8) - \gamma\mathbb{E}(W^2)$. Here the Berry-Esseen rate is $\frac{C}{n^{1/4}}$.

# 5 Appendix

*Proof of Lemma 2.2.* Consider a probability density of the form

$$p(x) := p_k(x) := b_k \exp\left(-a_k x^{2k}\right) \tag{5.49}$$

with $b_k = \int_{\mathbb{R}} \exp\left(-a_k x^{2k}\right) dx$. 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 (5.49), 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

$$f_z(x) = \begin{cases} (1-P(z))\, P(x) \exp(a_k x^{2k}) b_k^{-1} & \text{for } x \leq z, \\ P(z)\,(1-P(x)) \exp(a_k x^{2k}) b_k^{-1} & \text{for } x \geq z \end{cases} \tag{5.50}$$

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

$$1 - P(x) \leq \frac{b_k}{2k\, a_k x^{2k-1}} \exp\left(-a_k x^{2k}\right), \tag{5.51}$$

whereas for $x < 0$ we have

$$P(x) \leq \frac{b_k}{2k\, a_k |x|^{2k-1}} \exp\left(-a_k x^{2k}\right). \tag{5.52}$$

By partial integration we have

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

Hence for any $x > 0$

$$b_k \left(\frac{x}{2k\, a_k x^{2k} + 2k - 1}\right) \exp\left(-a_k x^{2k}\right) \leq 1 - P(x). \tag{5.53}$$

With (5.51) we get for $x > 0$

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

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

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

Summarizing we obtain for $x > 0$

$$1 - P(x) \le \min\left(\frac{1}{2}, \frac{b_k}{2k\,a_k\,x^{2k-1}}\right)\exp\big(-a_k x^{2k}\big). \tag{5.54}$$

With (5.52) we get for $x < 0$

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

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

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

Summarizing we obtain for $x < 0$

$$P(x) \le \min\left(\frac{1}{2}, \frac{b_k}{2k\,a_k\,|x|^{2k-1}}\right)\exp\big(-a_k x^{2k}\big). \tag{5.55}$$

Applying (5.54) and (5.55) gives $0 < f_z(x) \le \frac{1}{2\,b_k}$ for all $x$. Note that for $x < 0$ we only have to consider the first case of (5.50), since $z \ge 0$. The constant $\frac{1}{2\,b_k}$ is not optimal. Following the proof of Lemma 2.2 in [5] or alternatively of Lemma 2 in [20, Lecture II] would lead to optimal constants. We omit this. It follows from (5.50) that

$$f_z'(x) = \begin{cases} (1 - P(z))\Big[1 + x^{2k-1}\,2k\,a_k\,P(x)\exp(a_k x^{2k})b_k^{-1}\Big] & \text{for } x \le z, \\ P(z)\Big[(1 - P(x))\,2k\,a_k\,x^{2k-1}\,\exp(a_k x^{2k})b_k^{-1} - 1\Big] & \text{for } x \ge z. \end{cases} \tag{5.56}$$

With (5.51) we obtain for $0 < x \le z$ that

$$f_z'(x) \le (1 - P(z))\Big[z^{2k-1}\,2k\,a_k\,P(z)\exp(a_k z^{2k})b_k^{-1}\Big] + 1 \le 2.$$

The same argument for $x \ge z$ leads to $|f_z'(x)| \le 2$. For $x < 0$ we use the first half of (5.50) and apply (5.52) to obtain $|f_z'(x)| \le 2$. Actually this bound will be improved later. Next we calculate the derivative of $-\psi(x)\,f_z(x)$:

$$(-\psi(x)f_z(x))' = \begin{cases} \frac{(1-P(z))}{b_k}\Big[P(x)e^{a_k x^{2k}}\Big(2k(2k-1)a_k x^{2k-2} + (2k)^2 a_k^2 x^{4k-2}\Big) + 2k a_k x^{2k-1} b_k\Big], & x \le z, \\ \frac{P(z)}{b_k}\Big[(1 - P(x))e^{a_k x^{2k}}\Big(2k(2k-1)a_k x^{2k-2} + (2k)^2 a_k^2 x^{4k-2}\Big) - 2k a_k x^{2k-1} b_k\Big], & x \ge z. \end{cases} \tag{5.57}$$

With (5.53) 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 (5.50)). Moreover with (5.51), (5.52) and (5.53) we obtain that

$$\lim_{x\to-\infty} 2k\, a_k\, x^{2k-1} f_z(x) = P(z) - 1 \quad \text{and} \quad \lim_{x\to\infty} 2k\, a_k\, x^{2k-1} f_z(x) = P(z). \tag{5.58}$$

Hence we have $|2k\, a_k\, x^{2k-1} f_z(x)| \leq 1$ and $|2k\, a_k\big(x^{2k-1} f_z(x) - u^{2k-1} f_z(u)\big)| \leq 1$ for any $x$ and $u$. From (5.51) it follows that $f_z'(x) > 0$ for all $x < z$ and $f_z'(x) < 0$ for $x > z$. With Stein's identity $f_z'(x) = -\psi(x)f_z(x) + 1_{\{x \leq x\}} - P(z)$ and (5.58) 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

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

Next we bound $(-\psi(x)f_z(x))'$. We already know that $(-\psi(x)f_z(x))' > 0$. Again we apply (5.51) and (5.52) 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 $h : \mathbb{R} \to \mathbb{R}$. Let $f_h$ be the solution of the Stein equation (2.6), 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 [5, 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 (2.6) and (5.53), 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 would mean to Stein's identity (2.6) 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 omitted. To bound the second derivative $f_h''$, we differentiate (2.6) and have

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

Similarly to [5, (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 (5.57) is non-negative to obtain

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

Moreover we know, that the quantity in (5.57) 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

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

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). □

*Sketch of the proof of Remark 2.3.* Now let $p(x) = b_k \exp\big(-a_k V(x)\big)$ and $V$ satisfies the assumptions listed in Remark 2.3. To proof that $f_z$ (with respect to $p$) satisfies Assumption (B2), we adapt (5.53) as well as (5.54) and (5.55), using the assumptions on $V$. We obtain for $x > 0$

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

and for $x > 0$

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

and for $x < 0$

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

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. □

## References


[1] A. D. Barbour, *Equilibrium distributions for Markov population processes*, Adv. in Appl. Probab. **12** (1980), no. 3, 591–614. MRMR578839

[2] 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.

[3] S. Chatterjee, J. Fulman, and A. Röllin, *Exponential approximation by Stein's method and spectral graph theory*, preprint, 2009.

[4] S. Chatterjee and Q.-M. Shao, *Stein's method of exchangeable pairs with application to the Curie-Weiss model*, preprint, 2009.

[5] L. H. Y. Chen and Q.-M. Shao, *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. MR2235448

[6] 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.

[7] P. Eichelsbacher and G. Reinert, *Stein's method for discrete Gibbs measures*, Ann. Appl. Probab. **18** (2008), no. 4, 1588–1618. MR2434182

[8] R. S. Ellis, *Concavity of magnetization for a class of even ferromagnets*, Bull. Amer. Math. Soc. **81** (1975), no. 5, 925–929. MR0376052

[9] R. S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*, Springer-Verlag, New York, 1985.

[10] 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. MR0395659

[11] 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.

[12] ———, *Necessary and sufficient conditions for the GHS inequality with applications to analysis and probability*, Trans. Amer. Math. Soc. **237** (1978), 83–99.

[13] 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.

[14] 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.

[15] J.L. Lebowitz, *GHS and other inequalities*, Comm. Math. Phys. **35** (1974), 87–92. MR0339738

[16] 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. MR1484798

[17] 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). MR2215787

[18] B. Simon and R. B. Griffiths, *The* $(\phi^4)_2$ *field theory as a classical Ising model*, Comm. Math. Phys. **33** (1973), 145–164. MR0428998

[19] ———, *The* $(\phi^4)_2$ *field theory as a classical Ising model*, Comm. Math. Phys. **33** (1973), 145–164. MR0428998

[20] C. Stein, *Approximate computation of expectations*, Institute of Mathematical Statistics Lecture Notes—Monograph Series, 7, Institute of Mathematical Statistics, Hayward, CA, 1986. MR882007

[21] 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. MR2118600

[22] C. J. Thompson, *Mathematical statistical mechanics*, The Macmillan Co., New York, 1972, A Series of Books in Applied Mathematics.