Poissonian Behavior of Ising Spin Systems in an External Field

A. Ganesh[1], B. M. Hambly[1,2], Neil O'Connell[1], Dudley Stark[1], and P. J. Upton[3,4]

[1] *BRIMS, Hewlett-Packard Labs*
Bristol, BS34 8QZ, UK

[2] *Department of Mathematics, University of Bristol*
University Walk, Bristol, BS8 1TW, UK

[3] *H. H. Wills Physics Laboratory, University of Bristol*
Tyndall Avenue, Bristol, BS8 1TL, UK

[4] *Fachbereich 7 Physik, Universität-GH Essen*
D-45117 Essen, Germany

Abstract

We apply the Stein-Chen method for Poisson approximation to spin-half Ising-type models, in positive external field, which satisfy the FKG inequality. In particular we show that, provided the density of minus spins is low and can be expanded as a convergent power series in the activity (fugacity) variable, the distribution of minus spins is approximately Poisson. The error of the approximation is inversely proportional to the number of lattice sites (we obtain upper and lower bounds on the total variation distance between the exact distribution and its Poisson approximation). We illustrate these results by application to specific models, including the mean-field and nearest-neighbour ferromagnetic Ising models.

KEY WORDS: Poisson approximation, Stein-Chen method, Ising models, nucleation

1 Introduction and Summary

A guiding principle in probability theory is the law of small numbers; that the distribution of rare events is approximately Poisson. In recent years there have been significant developments in quantifying this statement. In particular the 'Stein-Chen method' has been introduced which makes it possible to obtain good estimates on the accuracy of the Poisson approximation. A survey of the method and its application in a variety of mathematical problems can be found in [1].

The law of small numbers is often tacitly assumed in many areas of statistical physics. For instance, when one thermodynamic phase nucleates within another, it is often assumed that the distribution of nucleating droplets is Poisson, provided the density of nucleation centres is sufficiently low. In recent experimental work on dewetting of polymer films from solid surfaces [2, 3] the distribution of rupture sites was found to be Poisson, and the authors concluded that the dewetting process was dominated by the nucleation of defects which are intrinsic to the polymer — implicit here is the assumption that nucleating sites are Poisson distributed.

In this paper we apply the Stein-Chen method to classical lattice models in the low density regime. We restrict our attention to two-state Ising spin models, in a positive external field, which satisfy the FKG inequality. (The FKG inequality implies that the spin variables are positively correlated; for a precise definition see (1) below.) We give rigorous bounds for the error in the Poisson approximation which hold quite generally, and demonstrate that these bounds are, in some sense, the best possible. Our main result is that, roughly speaking, the Poisson approximation (for the distribution of minus spins) is justified, with error inversely proportional to the number of lattice sites, if the density is low and can be expanded as a convergent power series in the activity (fugacity) variable.

Barbour and Greenwood [4] have applied the Stein-Chen method to a class of Markov random fields, which includes the Ising model; the bounds they obtain for the Ising model are not very explicit. Using quite different methods, Fernández, Ferrari and Garcia [5] establish the asymptotic Poisson distribution of contours in the nearest-neighbour Ising model at low temperatures and zero external field.

The outline of the paper is as follows. In §2, we give a very brief introduction to the Stein-Chen method for Poisson approximation, and state the bounds which we shall apply. In §3, we introduce the class of models under consideration, and apply the bounds of the previous section to these models. Upper bounds are given in §§3.1. To demonstrate that these are optimal, in some sense, we establish lower bounds for the error in the approximation in §§3.2. Applications to specific models are presented in §§3.3. These include non-interacting models (high temperature limit), mean-field models, and the nearest-neighbour ferromagnetic Ising model. We finish with a conclusion in §4, an outline of the Stein-Chen method in Appendix A and some details which are needed for §§3.3 in Appendix B.

2 The law of small numbers

Suppose $W = \sum_{\alpha \in \Gamma} I_\alpha$, where Γ is some index set and I_α are Bernoulli random variables (that is, they take values 0 or 1). If $\rho_\alpha := \mathbb{E} I_\alpha$ is small enough and there is not much dependence between the I_α, then the 'law of small numbers' says that W should have approximately a Poisson distribution. This can be made precise. We shall restrict our attention to variables which satisfy the FKG inequality, which we now describe.

There is a natural partial ordering on $\{0,1\}^\Gamma$ defined by $x \le y$ if $x_\alpha \le y_\alpha$ for all $\alpha \in \Gamma$. The random variables I_α are said to satisfy the Fortuin-Kasteleyn-Ginibre (FKG) inequality [6] if, for all non-decreasing f and g on $\{0,1\}^\Gamma$,

$$\mathbb{E}\{f(I_\alpha,\ \alpha \in \Gamma) g(I_\alpha,\ \alpha \in \Gamma)\} \ge \mathbb{E} f(I_\alpha,\ \alpha \in \Gamma) \mathbb{E} g(I_\alpha,\ \alpha \in \Gamma). \tag{1}$$

(We say that f is *non-decreasing* if $f(x) \le f(y)$ whenever $x < y$.) Random variables which satisfy the FKG inequality are said to be *positively associated*. Note that this property implies pairwise positive correlation, but not conversely.

Let $\mathcal{L}(X)$ denote the *law* of a random variable X. For random variables X and Y, the total variation distance between $\mathcal{L}(X)$ and $\mathcal{L}(Y)$ is defined as

$$d_{\mathrm{TV}}(\mathcal{L}(X), \mathcal{L}(Y)) := \sup_A |\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)|, \tag{2}$$

where the supremum is taken over measurable sets. Denote by $\mathrm{Po}(\lambda)$ the law of a Poisson random variable with mean λ. Set $\lambda := \sum_\alpha \rho_\alpha$.

Theorem 2.1 *[1, Corollary 2.C.4] In the above setting, if the FKG inequality is satisfied, we have*

$$d_{\mathrm{TV}}(\mathcal{L}(W), \mathrm{Po}(\lambda)) \leq \lambda^{-1}(1 - e^{-\lambda}) \left\{ \mathrm{var}\, W - \lambda + 2 \sum_{\alpha \in \Gamma} \rho_\alpha^2 \right\}. \tag{3}$$

We can also compare the whole process of indicators $(I_\alpha, \alpha \in \Gamma)$ with a Poisson process $(J_\alpha, \alpha \in \Gamma)$, with $\mathbb{E} J_\alpha = \rho_\alpha$. (The J's are independent Poisson random variables.) It is convenient to describe the process as an integer-valued measure on Γ: set $\Xi = \sum_\alpha I_\alpha \delta_\alpha$ and $\Pi = \sum_\alpha J_\alpha \delta_\alpha$, where δ_α denotes a point mass at α, i.e., for $A \subset \Gamma$,

$$\delta_\alpha\{A\} = \begin{cases} 1 & \text{if } A \supseteq \alpha; \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

We will denote the law of the Poisson random measure Π by $\mathrm{Po}(\boldsymbol{\rho})$, where $\boldsymbol{\rho} := \{\rho_\alpha\}_{\alpha \in \Gamma}$. The total variation distance may be defined in a way similar to (2), but now we take the supremum over subsets A of integer-valued measures.

Theorem 2.2 *[1, Remark 10.1.4] In the above setting, if the FKG inequality is satisfied, we have*

$$d_{\mathrm{TV}}(\mathcal{L}(\Xi), \mathrm{Po}(\boldsymbol{\rho})) \leq \mathrm{var}\, W - \lambda + 2 \sum_{\alpha \in \Gamma} \rho_\alpha^2. \tag{5}$$

These theorems are proved using the Stein-Chen method, a powerful technique in probability theory used to approximate one distribution by another. We refer the interested reader to Appendix A for a brief outline.

The Poisson distribution has the property that the variance is equal to the mean. Although this is not a characterisation, the above bounds indicate that whenever the FKG property holds, if the index set Γ is large and λ is moderate, then $\sum_{\alpha \in \Gamma} \rho_\alpha^2$ is small and *the distance to the Poisson approximation is essentially the extent to which the variance exceeds the mean.* This provides a useful rule of thumb. We remark here that for positively associated variables, the variance usually exceeds the mean (but not always).

Similar bounds can be obtained for the case of negatively associated indicator variables, where the opposite of the FKG inequality holds; in this case the spin variables are negatively correlated. For example, [1, Remark 10.1.4] there is an analogue of (5) which states that

$$d_{\mathrm{TV}}(\mathcal{L}(\Xi), \mathrm{Po}(\boldsymbol{\rho})) \leq \lambda - \mathrm{var}\, W.$$

Note that in this case the variance is always less than the mean.

3 Application to Ising Models

Ising spins $\sigma_i \in \{-1, 1\}$ are placed on sites i of a d-dimensional lattice $\Lambda \subset \mathbb{Z}^d$ with a total number of $|\Lambda| = N$ sites. The Hamiltonian, $\mathcal{H}_\Lambda(\sigma)$, acting on each spin configuration $\sigma \in \{-1, 1\}^\Lambda$, has the general form

$$\mathcal{H}_\Lambda(\sigma) = -\sum_{A \subset \Lambda} J_A \, \sigma^A - H \sum_{i \in \Lambda} \sigma_i \tag{6}$$

where $\sigma^A := \prod_{i \in A} \sigma_i$, H is the applied magnetic field and the interactions J_A are taken to be translationally invariant. The probability distribution of σ, $\mathbb{P}_\Lambda(\sigma)$, is given by the Gibbs measure

$$\mathbb{P}_\Lambda(\sigma) = \frac{1}{Z_\Lambda(z, \beta)} e^{-\beta \mathcal{H}_\Lambda(\sigma)} \tag{7}$$

where β is the inverse temperature, z is the 'activity' variable $z = e^{-2h}$ with $h = \beta H$ and $Z_\Lambda(z, \beta) = \sum_\sigma e^{-\beta \mathcal{H}_\Lambda(\sigma)}$ is the partition function. The indicator for $\sigma_i = -1$ is denoted by $I_i = (1 - \sigma_i)/2$. We consider values of H and β for which $\rho_i := \mathbb{E} I_i$ is small for all $i \in \Lambda$ (for example by having H sufficiently large and positive). We define $W := \sum_{i \in \Lambda} I_i$ (total number of negative spins in Λ) and the density of negative spins, ρ, is defined by $\rho := N^{-1} \sum_{i \in \Lambda} \rho_i$. We also define the integer-valued measure, Ξ, on Λ by $\Xi := \sum_{i \in \Lambda} I_i \delta_i$ where δ_i is a point mass at site i (defined as in (4)).

In what follows we shall be applying expansions in the activity, z, so we re-express the partition function as follows. On defining $\Phi_\Lambda(\sigma) := -\sum_{A \subset \Lambda} J_A \sigma^A$ we write

$$Z_\Lambda(z, \beta) = e^{-\beta \Phi_\Lambda(+)} z^{-N/2} \mathcal{Z}_\Lambda(z, \beta) \tag{8}$$

where $\sigma = +$ denotes the configuration with $\sigma_i = +1$ for all $i \in \Lambda$ and

$$\mathcal{Z}_\Lambda(z, \beta) = 1 + \sum_{n=1}^{N} z^n Q_n(\beta; \Lambda) \tag{9}$$

with

$$Q_n(\beta; \Lambda) := \sum_{\sigma : W = n} e^{-\beta[\Phi_\Lambda(\sigma) - \Phi_\Lambda(+)]}. \tag{10}$$

The density, ρ, can then be determined by

$$\rho(z) = \frac{z}{N} \frac{\partial}{\partial z} \ln \mathcal{Z}_\Lambda(z, \beta) \tag{11}$$

from which follows the activity expansion

$$\rho(z) = \sum_{n=1}^{\infty} a_n(\beta; \Lambda) z^n \tag{12}$$

provided convergence holds.

3.1 Upper Bounds

The aim here is to apply the bounds of Section 2 to the models defined by (6) and (7). We shall restrict our attention to Ising models which satisfy the FKG inequality so that bounds (3) and (5) are valid. This is certainly true if J_A are non-negative. We have

$$\operatorname{var} W = \sum_{i\in\Lambda}\sum_{j\in\Lambda} \operatorname{cov}(I_i, I_j) \tag{13}$$

and $\lambda := \mathbb{E}W = \sum_{i\in\Lambda}\rho_i$. The magnetisation on site $i \in \Lambda$ is defined as $m_i := \mathbb{E}\sigma_i = 1 - 2\rho_i$, the average magnetisation across the lattice, m, is $m := N^{-1}\sum_{i\in\Lambda} m_i$ and the susceptibility, χ, is $\chi := \partial m/\partial h$. It immediately follows from (13) that $\operatorname{var} W = \frac{1}{4}N\chi$ and therefore

$$\operatorname{var} W - \lambda = N\left(\tfrac{1}{4}\chi - \rho\right). \tag{14}$$

Alternatively, one notes that $\frac{1}{4}\chi = z\partial\rho/\partial z$. From hereon, for simplicity, we assume homogeneity in that $\rho_i = \rho$ for all $i \in \Lambda$ although generalisations to cases where this does not hold can easily be extended from what follows. Hence, we now have

$$\operatorname{var} W - \lambda + 2\sum_{i\in\Lambda}\rho_i^2 = N\left(z\frac{\partial\rho}{\partial z} - \rho + 2\rho^2\right), \tag{15}$$

with $\lambda = N\rho$. We now state the following proposition.

Proposition 3.1 *For values of H and β where ρ can be expressed as a convergent power series (activity expansion) in z for a sequence $\Lambda \to \mathbb{Z}^d$ and, if ρ is sufficiently small, then the following bounds hold*

$$d_{\mathrm{TV}}(\mathcal{L}(W), \mathrm{Po}(\lambda)) \;\leq\; \frac{1}{N}\lambda(1-e^{-\lambda})C(\beta)\left[1 + O\left(\frac{\lambda}{N}\right)\right], \tag{16}$$

$$d_{\mathrm{TV}}(\mathcal{L}(\Xi), \mathrm{Po}(\boldsymbol{\rho})) \;\leq\; \frac{1}{N}\lambda^2 C(\beta)\left[1 + O\left(\frac{\lambda}{N}\right)\right], \tag{17}$$

where $C(\beta)$ is a constant for fixed β.

Proof This follows easily from (15) and Theorems 2.1 and 2.2. We start from the activity expansion (12) which, for a sequence of lattices Λ tending to infinity, stays convergent in a region $D \subset \mathbb{C}$ and will be simple for sufficiently small $|\rho|$. So clearly, one has

$$z\frac{\partial\rho}{\partial z} - \rho = \sum_{n=2}^{\infty}(n-1)a_n(\beta;\Lambda)z^n, \tag{18}$$

which is also convergent and simple. For sufficiently small $|\rho|$ it is possible to invert (12) and write z as a convergent series of ρ which is also simple [7]. Substituting this into (18) gives

$$z\frac{\partial\rho}{\partial z} - \rho = \sum_{n=2}^{\infty} b_n(\beta;\Lambda)\rho^n \tag{19}$$

and since this stays convergent for all Λ tending to infinity one has

$$z\frac{\partial\rho}{\partial z} - \rho + 2\rho^2 \le C(\beta)\rho^2 + O(\rho^3). \tag{20}$$

Substituting this into (15) and applying Bounds (3) and (5) of Section 2 completes the proof.

We now make some remarks:

Remark 3.2 The general condition for convergence of activity expansions for models defined by (6) can be expressed in terms of β and J_A (see, e.g., Ruelle [8] for details). Certainly, if z is sufficiently small (H sufficiently large) convergence is assured.

Remark 3.3 In the complex z plane, singularities in $\rho = \rho(z)$ occur only at all those values of z for which $Z_\Lambda(z,\beta) = 0$ (though these cannot lie on the positive real axis, i.e., the physical segment, when Λ is finite) [9]. Thus, the activity expansion for $\rho(z)$ is convergent in the circle $|z| < d_\Lambda$ where d_Λ is the distance from the origin to the nearest zero of $Z_\Lambda(z,\beta)$ and therefore convergence is guaranteed for all Λ when $H > -(2\beta)^{-1}\ln d^*$ where $d^* = \inf_\Lambda d_\Lambda$.

Remark 3.4 For the case of models having pairwise ferromagnetic interactions only, i.e., $J_A = 0$ whenever $|A| \neq 2$ and $J_{\{i,j\}} \geq 0$ for site pairs $A = \{i,j\}$, we can apply the theorem of Lee and Yang [10]. This states that $Z_\Lambda(z,\beta) = 0$ implies $|z| = 1$ for all Λ. Thus, from Remark 3.3, the activity expansion for ρ is convergent for all positive H no matter how small. However, if H is small positive, β needs to be large in order to keep ρ small so as to satisfy the conditions of Proposition 3.1. Unfortunately, $C(\beta)$ may well diverge as $\beta \to \infty$ and this will prove to be the case in some of the examples considered below.

Remark 3.5 Further insight can be gained by expressing the bounds in Proposition 3.1 in terms of ρ as follows

$$d_{\mathrm{TV}}(\mathcal{L}(W), \mathrm{Po}(\lambda)) \;\le\; (1-e^{-\lambda})\, C(\beta)\,\rho\,[1+O(\rho)] \tag{21}$$
$$d_{\mathrm{TV}}(\mathcal{L}(\Xi), \mathrm{Po}(\boldsymbol{\rho})) \;\le\; \lambda\, C(\beta)\,\rho\,[1+O(\rho)] \tag{22}$$

and recall that λ is kept fixed whilst ρ must be small (so clearly $N = O(1/\rho)$). Two ways of making ρ small are listed below:

1. Taking $H \to \infty$ with β fixed and moderate. In this case $\rho \sim e^{-2\beta H}$ as $H \to \infty$ (with $N = O(e^{2\beta H})$).
2. Taking $\beta \to \infty$ (low-temperature limit) with H fixed *positive* and moderate. For the nearest-neighbour ferromagnetic Ising model (with coupling constant J and lattice coordination number q) it is known that $\rho \sim e^{-2q\beta J}$ as $\beta \to \infty$ [with $N = O(e^{2q\beta J})$] and the comments in Remark 3.4 fully apply here. The requirement that $C(\beta)\rho \to 0$ as $\beta \to \infty$ holds for this model provided $q > 2$ since $C(\beta) \sim e^{4\beta J}$ as $\beta \to \infty$, as can be seen in §§3.3.3 below.

3.2 Lower Bound

It is possible to place a lower bound on the total variation distances by considering a specific event. In particular, for the Ising models, the simplest would be to compare $\mathbb{P}(W=0)=\mathbb{P}_\Lambda(+)$ with $\mathbb{P}(Z=0)=e^{-\lambda}$. We make the following proposition.

Proposition 3.6 *For values of H and β where ρ can be expressed as a convergent power series (activity expansion) in z for a sequence $\Lambda \to \mathbb{Z}^d$ and if ρ is sufficiently small then*

$$\left|\mathbb{P}_\Lambda(+) - e^{-\lambda}\right| = \frac{1}{N} e^{-\lambda}\lambda^2 \left|1-\tfrac{1}{2}C(\beta)\right| \left[1+O(N^{-1})\right] \tag{23}$$

where the constant $C(\beta)$ is the same as that in Proposition 3.1.

Proof Start by noting that

$$\begin{aligned}\mathbb{P}_\Lambda(+) &= \frac{1}{\mathcal{Z}_\Lambda(z,\beta)} \\ &= \exp\left[-N\int_0^z \frac{\rho(z')}{z'}\,dz'\right]\end{aligned} \tag{24}$$

where the second line follows directly from (11). Using the activity expansion (12) we have

$$\int_0^z \frac{\rho(z')}{z'}\,dz' = \sum_{n=1}^{\infty} \frac{1}{n} a_n(\beta;\Lambda) z^n \tag{25}$$

$$= \rho + \sum_{n=2}^{\infty} c_n(\beta;\Lambda)\rho^n \tag{26}$$

where (26) followed from (25) by inverting (12) which is possible for sufficiently small $|\rho|$. Therefore, putting $\lambda = N\rho$ we have

$$\left|\mathbb{P}_\Lambda(+) - e^{-\lambda}\right| = e^{-\lambda}\left|1-\exp\left[-\sum_{n=2}^{\infty} c_n(\beta;\Lambda)\lambda^n N^{1-n}\right]\right| \tag{27}$$

which, given that $c_2(\beta;\Lambda) = -\frac{1}{2}b_2(\beta;\Lambda)$ [where $b_n(\beta;\Lambda)$ are the same expansion coefficients used in (19)], completes the proof.

Remark 3.7 Clearly, $|\mathbb{P}_\Lambda(+) - e^{-\lambda}|$ provides a lower bound to $d_{\mathrm{TV}}(\mathcal{L}(\cdot), \mathrm{Po}(\cdot))$ for both $\mathcal{L}(W)$ and $\mathcal{L}(\Xi)$ and therefore their upper bounds cannot be tighter than $O(N^{-1})$. Thus, for fixed λ and β, we have

$$d_{\mathrm{TV}}(\mathcal{L}(\cdot), \mathrm{Po}(\cdot)) \sim \frac{1}{N} \tag{28}$$

as N gets large. [Here, we use the convention that $f(N) \sim g(N)$ as $N \to \infty$ denotes that $\ln f(N)/\ln g(N) \to 1$ as $N \to \infty$.]

3.3 Examples

Here we consider some specific choices for J_A in (6) and compute the constant $C(\beta)$ for these models.

3.3.1 Non-Interacting Model and the High-Temperature Limit

This model is defined by putting $J_A = 0$ for all $A \subset \Lambda$ (non-interacting paramagnet). The results so obtained are, of course, equivalent to taking, in (6) and (7), the simultaneous limit $\beta \to 0$, $H \to \infty$ such that $h = \beta H$ is kept finite.

For this case, it is straightforward to show that $\rho = z/(1+z)$ from which it follows that $b_2(0;\Lambda) = -1$ and $b_n(0;\Lambda) = 0$ for all $n \geq 3$. Also, $c_n(0;\Lambda) = 1/n$ in the expansion used for Proposition 3.6. Hence, it immediately follows that $C(0) = 1$.

3.3.2 Mean-Field Model

Here, the interactions are infinite-ranged pairwise ferromagnetic and defined by having $J_A = 0$ whenever $|A| \neq 2$ and $J_{\{i,j\}} = \tilde{J}/N$ for *all* pairs $A = \{i,j\}$ where $\tilde{J} \geq 0$. The location of the zeros of $Z_\Lambda(z,\beta)$ are known for this model [11] and lie along the circle $|z| = 1$ for all Λ — the same as those models where the theorem of Lee and Yang, mentioned in Remark 3.4, can be applied. Therefore, the activity expansion for $\rho(z)$, Eq. (12), is convergent for $|z| < 1$.

In order to generate an expansion in z one needs to determine $Q_n(\beta;\Lambda)$. For the mean-field model this can be expressed simply as

$$Q_n(\beta;\Lambda) = \binom{N}{n} \exp\left[2\beta\tilde{J}n\left(\frac{n}{N} - 1\right)\right] \tag{29}$$

and from this follows the expansion coefficients $\{b_n(\beta;\Lambda)\}_{n\geq 2}$. In particular, the first two [obtained from Eqs. (38) to (42) of Appendix B] are given by

$$\begin{aligned} b_2(\beta;\Lambda) &= 4\beta\tilde{J} - 1 + O(N^{-1}) && (30) \\ b_3(\beta;\Lambda) &= 8\beta\tilde{J}(2\beta\tilde{J} - 1) + O(N^{-1}) && (31) \end{aligned}$$

and from (30) one obtains $C(\beta)$ as follows

$$C(\beta) = 4\beta\tilde{J} + 1. \tag{32}$$

3.3.3 Nearest-Neighbour Ferromagnetic Ising Model

We consider a lattice, Λ, with coordination number q (e.g., $q = 2d$ for a d-cubic lattice) and take $J_A = J \geq 0$ for all A consisting of nearest-neighbour pairs whereas $J_A = 0$ otherwise.[5]

[5] In order to avoid more complicated lattice boundary considerations we assume, for simplicity, periodic boundary conditions (i.e., Λ wrapped on a d torus). Of course, the resulting value for $C(\beta)$ applies to more general lattice boundaries.

Again, the Lee-Yang theorem of Remark 3.4 applies here so that the expansion (12) is convergent for $|z| < 1$. The expansion coefficients $\{b_n(\beta;\Lambda)\}_{n\geq 2}$ can be systematically computed, the first two of which (see Appendix B) are given below

$$
\begin{aligned}
b_2(\beta;\Lambda) &= qe^{4\beta J} - q - 1 && (33)\\
b_3(\beta;\Lambda) &= q(q-3)e^{8\beta J} + 2q(2-q)e^{4\beta J} + q(q-1) && (34)
\end{aligned}
$$

and from (33) follows $C(\beta)$ which is given by

$$C(\beta) = qe^{4\beta J} - q + 1. \tag{35}$$

Note that, as required, if we put $J = \tilde{J}/q$ and take $q \to \infty$ we recover the results of the mean-field model given above with $\lim_{q\to\infty} C(\beta)$ given by (32).

4 Conclusion

When an uniaxial ferromagnet is deep enough within the positively magnetized region of the phase diagram, so that the probability of finding a negative spin is sufficiently small, it is usually assumed that the distribution of negative spins is Poisson. Here, under quite general conditions for a range of interacting spin systems, we have been able to rigorously quantify this picture and show that the distribution of minus spins is approximately Poisson with an error inversely proportional to the number of lattice sites. The two main conditions under which our proof holds are: (a) that the Gibbs measure for the system satisfies the FKG property (this is certainly true if all interactions are ferromagnetic) and (b) that the density of minus spins is sufficiently small and expandable as a convergent power series in the activity (a property that generally holds if one is sufficiently far from phase boundaries or if the temperature is high enough).

This may have implications in the study of nucleation. For, if at some subcritical temperature one suddenly reverses the magnetic field from a value deep within the positively magnetized region of the phase diagram to a negative value, the system becomes metastable and will eventually decay into the stable negatively magnetized phase. The minus spins present before the field reversal would act as 'seed' sites for the nucleating negatively-magnetized droplets. Thus, the distribution of droplets reflects the distribution of these negative spins and would therefore be approximately Poisson — in experimental situations this is often taken as a signature that the decay process of the metastable phase was due to nucleation rather than, say, through the formation of hydrodynamic-like instabilities typical of spinodal decomposition.

Acknowledgments

BMH acknowledges the support of a Royal Society Industry Fellowship held at BRIMS, Hewlett-Packard, where this research was undertaken. PJU acknowledges financial support from the EPSRC (UK) under Grant No. B/94/AF/1769.

A The Stein-Chen Method

In this appendix we give a brief outline of the Stein-Chen method for Poisson approximation.

Suppose $W = \sum_{\alpha\in\Gamma} I_\alpha$, where Γ is some index set and I_α are Bernoulli random variables. Suppose we can construct random variables U_α and V_α *on the same probability space* such that

$$\mathcal{L}(U_\alpha) = \mathcal{L}(W), \qquad \mathcal{L}(1+V_\alpha) = \mathcal{L}(W|I_\alpha = 1).$$

Such a construction is called a *coupling*. Then (see, for example, Theorem 2.A of [1]), for $\lambda := \sum_\alpha \rho_\alpha$, we have

$$d_{\mathrm{TV}}(\mathcal{L}(W), \mathrm{Po}(\lambda)) \leq \lambda^{-1}(1-e^{-\lambda}) \sum_{\alpha\in\Gamma} \rho_\alpha \mathbb{E}|U_\alpha - V_\alpha|. \tag{36}$$

The total variation distance is a measure of how close the distribution of W is to the Poisson distribution with mean λ; assuming that the ρ_α's are small, the notion of weak dependence can now be expressed as "for each α, $U_\alpha = V_\alpha$ with high probability."

We will sketch how (36) is obtained using the Stein-Chen method. The basic idea of the Stein-Chen method is to characterize the distribution of interest in some way, usually by a *Stein operator* on some function space. This operator is often related to a Markov process which has the approximating distribution as its stationary distribution. The Stein operator, $\mathcal{A}_\lambda$, for $\mathrm{Po}(\lambda)$, the Poisson distribution with mean λ, acts on bounded functions f on the non-negative integers and is defined by

$$\mathcal{A}_\lambda f(i) = \lambda f(i+1) - i f(i).$$

The operator $\mathcal{A}_\lambda$ is the generator of the immigration-death process with immigration rate λ and unit death rate, whose stationary distribution is Poisson with mean λ. The key fact is that $\mathbb{E}\mathcal{A}_\lambda f(Z) = 0$ *for all bounded f defined on the non-negative integers if, and only if, Z is Poisson with mean λ.* This suggests that if $\mathbb{E}\mathcal{A}_\lambda f(W)$ is small, for a sufficiently rich class of functions f, then the distribution of W is close to $\mathrm{Po}(\lambda)$.

This is indeed the case, and can be formulated as follows. The class of functions to consider are solutions, g, to the equation

$$\mathcal{A}_\lambda g = I[i \in A] - \mathbb{P}(Z \in A),$$

with $g(0) = 0$, for each subset A of the non-negative integers. Here I is the indicator function and $\mathcal{L}(Z) = \mathrm{Po}(\lambda)$. This is the *Stein equation.* It can be shown that if g is a solution, for any particular A, we have

$$\sup_i |g(i+1) - g(i)| \leq \lambda^{-1}(1-e^{-\lambda}).$$

Using this fact, and the above coupling, we have, for each A,

$$\begin{aligned}
|\mathbb{P}(W\in A) - \mathbb{P}(Z \in A)| &= |\mathbb{E}[\lambda g(W+1) - W g(W)]| \\
&= \left| \sum_\alpha \rho_\alpha \mathbb{E}[g(U_\alpha+1) - g(V_\alpha+1)] \right| \\
&\leq \lambda^{-1}(1-e^{-\lambda}) \sum_\alpha \rho_\alpha \mathbb{E}|U_\alpha - V_\alpha|
\end{aligned}$$

establishing the bound (36).

The method for obtaining Poisson *process* approximations is similar. In the notation of Section 2, suppose that for each $\alpha \in \Gamma$, random measures Ψ_α and Φ_α can be realized on the same probability space, in such a way that

$$\mathcal{L}(\Psi_\alpha) = \mathcal{L}(\Xi), \quad \mathcal{L}(\Phi_\alpha + \delta_\alpha) = \mathcal{L}(\Xi | I_\alpha = 1).$$

Then, using arguments similar to the above, it is possible to show (see Theorem 10.B of [1]) that

$$d_{\mathrm{TV}}(\mathcal{L}(\Xi), \mathrm{Po}(\boldsymbol{\rho})) \leq \sum_{\alpha \in \Gamma} \rho_\alpha \mathbb{E} d(\Psi_\alpha, \Phi_\alpha), \tag{37}$$

where

$$d(\Psi_\alpha, \Phi_\alpha) = \sum_{\beta \in \Gamma} |\Psi_\alpha\{\beta\} - \Phi_\alpha\{\beta\}|.$$

The bounds in the statements of Theorems 2.1 and 2.2 are obtained through a suitable choice of U_α and V_α (respectively Ψ_α and Φ_α).

B Activity Expansions

Here we give some details on how the expansion coefficients $\{b_n(\beta; \Lambda)\}_{n \geq 2}$ are computed. We start by using Eq. (11) to generate the expansion for $\rho(z)$ [Eq. (12)] where $a_n(\beta; \Lambda)$ (for $n \geq 1$) can be expressed in terms of $\{Q_m(\beta; \Lambda)\}_{1 \leq m \leq n}$. The first three coefficients are given by

$$\begin{aligned}
N\, a_1(\beta; \Lambda) &= Q_1(\beta; \Lambda) & (38)\\
N\, a_2(\beta; \Lambda) &= 2Q_2(\beta; \Lambda) - Q_1^2(\beta; \Lambda) & (39)\\
N\, a_3(\beta; \Lambda) &= 3Q_3(\beta; \Lambda) - 3Q_1(\beta; \Lambda)Q_2(\beta; \Lambda) + Q_1^3(\beta; \Lambda). & (40)
\end{aligned}$$

If $\rho = \rho(z)$ is a simple function (as it will be for sufficiently small $|\rho|$), one can invert (12) and write $z = z(\rho)$ as a power series in ρ which when substituted into (18) gives the expansion coefficients $\{b_n(\beta; \Lambda)\}_{n \geq 2}$. The first two of these are given below

$$\begin{aligned}
b_2(\beta; \Lambda) &= a_2(\beta; \Lambda)/a_1^2(\beta; \Lambda) & (41)\\
b_3(\beta; \Lambda) &= 2[a_1(\beta; \Lambda)a_3(\beta; \Lambda) - a_2^2(\beta; \Lambda)]/a_1^4(\beta; \Lambda). & (42)
\end{aligned}$$

For the nearest-neighbour Ising model of Subsection 3.3.3 (with periodic boundary conditions) we find that

$$\begin{aligned}
Q_1(\beta; \Lambda)/N &= u^q & (43)\\
Q_2(\beta; \Lambda)/N &= \tfrac{1}{2}(N - q - 1)u^{2q} + \tfrac{1}{2}qu^{2q-2} & (44)\\
Q_3(\beta; \Lambda)/N &= \tfrac{1}{6}[N^2 - 3(q+1)N + 3q^2 + 3q + 2]u^{3q} & \\
&+ \tfrac{1}{2}q(N - 2q)u^{3q-2} + \tfrac{1}{2}q(q-1)u^{3q-4} & (45)
\end{aligned}$$

where $u = e^{-2\beta J}$. Equations (33) and (34) then follow after substituting (43) to (45) into (38) to (42).

References

[1] A. D. Barbour, L. Holst, and S. Janson, *Poisson Approximation*, (Clarendon Press, Oxford, 1992).

[2] K. Jacobs, S. Herminghaus and K. R. Mecke, Thin Liquid Polymer Films Rupture via Defects, *Langmuir* **14**, 965 (1998).

[3] S. Herminghaus, K. Jacobs, K. R. Mecke, J. Bischof, A. Fery, M. Ibn-Elhaj and S. Schlagowski, Spinodal Dewetting in Liquid and Metal Films, *Science* **282**, 916 (1998).

[4] A. D. Barbour and P. E. Greenwood, Rates of Poisson Approximation to Finite Range Random Fields, *Ann. Appl. Prob.* **3**, 91 (1993).

[5] R. Fernández, P. A. Ferrari and N. L. Garcia, Loss network representation of Peierls contours. Preprint. Los Alamos Archive math.PR/9806131, 1998.

[6] C. M. Fortuin, P. W. Kasteleyn and J. Ginibre, Correlation Inequalities in some Partially Ordered Sets, *Comm. Math. Phys.* **22**, 89 (1971).

[7] E. C. Titchmarsh, *The Theory of Functions,* 2nd ed., (Oxford University Press, London, 1939).

[8] D. Ruelle, *Statistical Mechanics: Rigorous Results,* (W. A. Benjamin, Inc., New York, 1969).

[9] C. N. Yang and T. D. Lee, Statistical Theory of Equations of State and Phase Transitions, I. Theory of Condensation, *Phys. Rev.* **87**, 404 (1952).

[10] T. D. Lee and C. N. Yang, Statistical Theory of Equations of State and Phase Transitions, II. Lattice Gas and Ising Model, *Phys. Rev.* **87**, 410 (1952).

[11] S. Katsura, Phase Transition of Husimi-Temperley Model of Imperfect Gas, *Prog. Theo. Phys.* **13**, 571 (1955).