

Mixing Times for the Mean-Field Blume-Capel Model via Aggregate Path Coupling

Yevgeniy Kovchegov · Peter T. Otto · Mathew Titus

Received: 10 June 2011 / Accepted: 25 July 2011 / Published online: 24 August 2011
© Springer Science+Business Media, LLC 2011

Abstract We investigate the relationship between the mixing times of the Glauber dynamics of a statistical mechanical system with its thermodynamic equilibrium structure. For this we consider the mean-field Blume-Capel model, one of the simplest statistical mechanical models that exhibits the following intricate phase transition structure: within a two-dimensional parameter space there exists a curve at which the model undergoes a second-order, continuous phase transition, a curve where the model undergoes a first-order, discontinuous phase transition, and a tricritical point which separates the two curves. We determine the interface between the regions of slow and rapid mixing. In order to completely determine the region of rapid mixing, we employ a novel extension of the path coupling method, successfully proving rapid mixing even in the absence of contraction between neighboring states.

Keywords Path coupling · Mixing times · Glauber dynamics · Large deviations · Blume-Capel model · Aggregate path coupling

1 Introduction

The notion of mixing times is fundamental in the study of stochastic processes and their applications. In addition to its purely theoretical value, mixing times are widely used in computational and physical sciences. In computer science, mixing times represent running

Y. Kovchegov · M. Titus
Department of Mathematics, Oregon State University, Corvallis, OR 97331, USA

Y. Kovchegov
e-mail: kovchegy@science.oregonstate.edu

M. Titus
e-mail: titusm@math.oregonstate.edu

P.T. Otto (✉)
Department of Mathematics, Willamette University, Salem, OR 97302, USA
e-mail: potto@willamette.edu

times for randomized algorithms such as Metropolis-Hastings that are responsible for providing high accuracy solutions to a large class of problems where classical algorithms have impractical running times. In physics, one important use of mixing times is to determine the feasibility of simulating statistical mechanical systems via so-called Glauber dynamics.

An important question of mixing times for statistical mechanical models is its relationship with the thermodynamic phase transition structure of the system. More specifically, as a system undergoes an equilibrium phase transition with respect to some parameter; e.g. temperature, how do the corresponding mixing times behave? This question has been well studied with recent rigorous results for one class of statistical mechanical models, namely those that undergo second-order, continuous phase transitions like the famous Ising model [9, 15, 16]. For these models, it has been shown that the mixing times undergo a transition at precisely the thermodynamic phase transition point. However, for models that exhibit the other type of phase transition: first-order, discontinuous; e.g. Potts model with $q > 2$ [8, 17] and the Blume-Capel model [1, 2, 5–7, 12] with weak interaction, as far as we know, this question has not been previously addressed. In this paper, we prove that the mixing time transition does not coincide with the thermodynamic equilibrium phase transition for a model that exhibits a first-order, discontinuous phase transition.

First-order, discontinuous phase transitions are more intricate than their counterparts, which makes rigorous analysis of these models traditionally more difficult. Furthermore, the more complex phase transition structure causes the models to fall outside the scope of standard mixing time techniques including the so-called ‘path coupling’ method [3]. The path coupling argument is a powerful tool used to prove rapid mixing for various Markov chains, including the Glauber dynamics for statistical mechanical models that exhibit a second-order, continuous phase transition. However, due to the intricacy of a first-order, discontinuous phase transition, the standard path coupling argument cannot be applied in all cases because contraction between couplings of all neighboring states does not exist.

In this paper we prove the mixing time rates for the mean-field Blume-Capel (BC) model [1, 2, 5–7], a statistical mechanical spin model ideally suited for the analysis of the relationship between the thermodynamic equilibrium behavior and mixing times due to its intricate phase transition structure. Specifically, the phase diagram of the BC model includes a curve at which the model undergoes a second-order, continuous phase transition, a curve where the model undergoes a first-order, discontinuous phase transition, and a tricritical point which separates the two curves.

As mentioned above, in a subset of the first-order, discontinuous phase transition region of the BC model, the standard path coupling method does not apply as it does throughout the second-order, continuous phase transition region. Therefore, we apply an innovative new method we call *aggregate path coupling* in order to prove rapid mixing for the BC model in the entire first-order, discontinuous phase transition region. While the standard path coupling method assumes contraction between *every* pair of states, our extended approach loosens this condition and requires contraction for only certain pairs of states. This new method is not restricted to only the BC model and can be applied to other statistical mechanical models that undergo a first-order, discontinuous phase transition and moreover is general enough to be a new tool in the theory of mixing times.

The paper is organized as follows. The mean-field Blume-Capel model is introduced in Sect. 2. There we describe the equilibrium phase transition structure of the model and state the large deviation principle of the magnetization. In Sect. 3, we define the Glauber dynamics for the BC model and the notion of mixing times of Markov chains. In Sect. 4, we introduce the path coupling for the BC model and prove the form of the mean coupling distance. In Sect. 5, we use the standard path coupling to prove rapid mixing in the single

phase region where the BC model undergoes a second-order, continuous phase transition. In Sect. 6, we modify the path coupling argument to prove rapid mixing in the region where the BC model undergoes a first-order, discontinuous phase transition. The slow mixing results are stated and proved in Sect. 7.

2 Mean-Field Blume-Capel Model

Statistical mechanical models are defined in terms of the Hamiltonian function. For the mean-field Blume-Capel model, the Hamiltonian function on the configuration space $\Omega^n = \{-1, 0, 1\}^n$ is defined by

$$H_{n,K}(\omega) = \sum_{j=1}^n \omega_j^2 - \frac{K}{n} \left(\sum_{j=1}^n \omega_j \right)^2$$

for configurations $\omega = (\omega_1, \dots, \omega_n)$. Here K represents the interaction strength of the model. Then for inverse temperature β , the mean-field Blume-Capel model is defined by the sequence of probability measures

$$P_{n,\beta,K}(\omega) = \frac{1}{Z_n(\beta, K)} \exp[-\beta H_{n,K}(\omega)]$$

where $Z_n(\beta, K) = \sum_{\omega \in \Omega^n} \exp[-\beta H_{n,K}(\omega)]$ is the normalizing constant called the *partition function*.

In [12], using large deviation theory [11], the authors proved the phase transition structure of the BC model. The analysis of $P_{n,\beta,K}$ was facilitated by expressing it in the form of a Curie-Weiss (mean-field Ising)-type model. This is done by absorbing the noninteracting component of the Hamiltonian into the product measure P_n that assigns the probability 3^{-n} to each $\omega \in \Omega^n$, obtaining

$$P_{n,\beta,K}(d\omega) = \frac{1}{\tilde{Z}_n(\beta, K)} \cdot \exp\left[n\beta K \left(\frac{S_n(\omega)}{n}\right)^2\right] P_{n,\beta}(d\omega). \tag{1}$$

In this formula $S_n(\omega)$ equals the total spin $\sum_{j=1}^n \omega_j$, $P_{n,\beta}$ is the product measure on Ω^n with identical one-dimensional marginals

$$\rho_\beta(d\omega_j) = \frac{1}{Z(\beta)} \cdot \exp(-\beta\omega_j^2) \rho(d\omega_j), \tag{2}$$

$Z(\beta)$ is the normalizing constant $\int_{\Lambda} \exp(-\beta\omega_j^2) \rho(d\omega_j) = 1 + 2e^{-\beta}$, and $\tilde{Z}_n(\beta, K)$ is the normalizing constant $[Z(\beta)]^n / Z_n(\beta, K)$.

Although $P_{n,\beta,K}$ has the form of a Curie-Weiss (mean-field Ising) model when rewritten as in (1), it is much more complicated because of the β -dependent product measure $P_{n,\beta}$ and the presence of the parameter K . These complications introduce new features to the BC model described above that are not present in the Curie-Weiss model [10].

The starting point of the analysis of the phase-transition structure of the BC model is the large deviation principle (LDP) satisfied by the spin per site or *magnetization* S_n/n with

respect to $P_{n,\beta,K}$. In order to state the form of the rate function, we introduce the cumulant generating function c_β of the measure ρ_β defined in (2); for $t \in \mathbb{R}$ this function is defined by

$$c_\beta(t) = \log \int_{\Lambda} \exp(t\omega_1) \rho_\beta(d\omega_1) = \log \left[\frac{1 + e^{-\beta}(e^t + e^{-t})}{1 + 2e^{-\beta}} \right].$$

We also introduce the Legendre-Fenchel transform of c_β , which is defined for $z \in [-1, 1]$ by

$$J_\beta(z) = \sup_{t \in \mathbb{R}} \{tz - c_\beta(t)\}$$

and is finite for $z \in [-1, 1]$. J_β is the rate function in Cramér’s theorem, which is the LDP for S_n/n with respect to the product measures $P_{n,\beta}$ [10, Theorem II.4.1] and is one of the components of the proof of the LDP for S_n/n with respect to the BC model $P_{n,\beta,K}$. This LDP is stated in the next theorem and is proved in Theorem 3.3 in [12].

Theorem 2.1 *For all $\beta > 0$ and $K > 0$, with respect to $P_{n,\beta,K}$, S_n/n satisfies the large deviation principle on $[-1, 1]$ with exponential speed n and rate function*

$$I_{\beta,K}(z) = J_\beta(z) - \beta K z^2 - \inf_{y \in \mathbb{R}} \{J_\beta(y) - \beta K y^2\}.$$

In other words, for any closed subset F ,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log P_{n,\beta,K} \{S_n/n \in F\} \leq -I_{\beta,K}(F) \tag{3}$$

and for any open subset G ,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log P_{n,\beta,K} \{S_n/n \in G\} \geq -I_{\beta,K}(G) \tag{4}$$

where $I_{\beta,K}(A) = \inf_{z \in A} I_{\beta,K}(z)$.

The LDP in the above theorem implies that those $z \in [-1, 1]$ satisfying $I_{\beta,K}(z) > 0$ have an exponentially small probability of being observed as $n \rightarrow \infty$. Hence we define the set of equilibrium macrostates by

$$\tilde{\mathcal{E}}_{\beta,K} = \{z \in [-1, 1] : I_{\beta,K}(z) = 0\}.$$

For $z \in \mathbb{R}$ we define

$$G_{\beta,K}(z) = \beta K z^2 - c_\beta(2\beta K z) \tag{5}$$

and as in [13] and [14] refer to it as the *free energy functional* of the model. The calculation of the zeroes of $I_{\beta,K}$ —equivalently, the global minimum points of $J_{\beta,K}(z) - \beta K z^2$ —is greatly facilitated by the following observations made in Proposition 3.4 in [12]:

- (1) The global minimum points of $J_{\beta,K}(z) - \beta K z^2$ coincide with the global minimum points of $G_{\beta,K}$, which are much easier to calculate.
- (2) The minimum values $\min_{z \in \mathbb{R}} \{J_{\beta,K}(z) - \beta K z^2\}$ and $\min_{z \in \mathbb{R}} \{G_{\beta,K}(z)\}$ coincide.

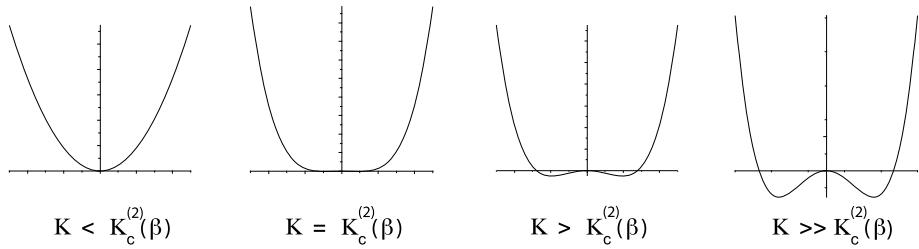


Fig. 1 The free-energy functional $G_{\beta,K}$ for $\beta \leq \beta_c$

Item (1) gives the alternate characterization that

$$\tilde{\mathcal{E}}_{\beta,K} = \{z \in [-1, 1] : z \text{ minimizes } G_{\beta,K}(z)\}. \tag{6}$$

The free energy functional $G_{\beta,K}$ exhibits two distinct behaviors depending on whether $\beta \leq \beta_c = \log 4$ or $\beta > \beta_c$. In the first case, the behavior is similar to the Curie-Weiss (mean-field Ising) model. Specifically, there exists a critical value $K_c^{(2)}(\beta)$ defined in (7) such that for $K < K_c^{(2)}(\beta)$, $G_{\beta,K}$ has a single minimum point at $z = 0$. At the critical value $K = K_c^{(2)}(\beta)$, $G_{\beta,K}$ develops symmetric non-zero minimum points and a local maximum point at $z = 0$. This behavior corresponds to a continuous, second-order phase transition and is illustrated in Fig. 1.

On the other hand, for $\beta > \beta_c$, $G_{\beta,K}$ undergoes two transitions at the values denoted by $K_1(\beta)$ and $K_c^{(1)}(\beta)$. For $K < K_1(\beta)$, $G_{\beta,K}$ again possesses a single minimum point at $z = 0$. At the first critical value $K_1(\beta)$, $G_{\beta,K}$ develops symmetric non-zero local minimum points in addition to the global minimum point at $z = 0$. These local minimum points are referred to as *metastable states* and we refer to $K_1(\beta)$ as the *metastable critical value*. This value is defined implicitly in Lemma 3.9 of [12] as the unique value of K for which there exists a unique $z > 0$ such that

$$G'_{\beta,K_1(\beta)}(z) = 0 \quad \text{and} \quad G''_{\beta,K_1(\beta)}(z) = 0.$$

As K increases from $K_1(\beta)$ to $K_c^{(1)}(\beta)$, the local minimum points decrease until at $K = K_c^{(1)}(\beta)$, the local minimum points reach zero and $G_{\beta,K}$ possesses three global minimum points. Therefore, for $\beta > \beta_c$, the BC model undergoes a phase transition at $K = K_c^{(1)}(\beta)$, which is defined implicitly in [12]. Lastly, for $K > K_c^{(1)}(\beta)$, the symmetric non-zero minimum points drop below zero and thus $G_{\beta,K}$ has two symmetric non-zero global minimum points. This behavior corresponds to a discontinuous, first-order phase transition and is illustrated in Fig. 2.

In the next two theorems, the structure of $\tilde{\mathcal{E}}_{\beta,K}$ corresponding to the behavior of $G_{\beta,K}$ just described is stated which depends on the relationship between β and the critical value $\beta_c = \log 4$. We first describe $\tilde{\mathcal{E}}_{\beta,K}$ for $0 < \beta \leq \beta_c$ and then for $\beta > \beta_c$. In the first case $\tilde{\mathcal{E}}_{\beta,K}$ undergoes a continuous bifurcation as K increases through the critical value $K_c^{(2)}(\beta)$ defined in (7); physically, this bifurcation corresponds to a second-order phase transition. The following theorem is proved in Theorem 3.6 in [12].

Theorem 2.2 For $0 < \beta \leq \beta_c$, we define

$$K_c^{(2)}(\beta) = \frac{1}{2\beta c''_\beta(0)} = \frac{e^\beta + 2}{4\beta}. \tag{7}$$

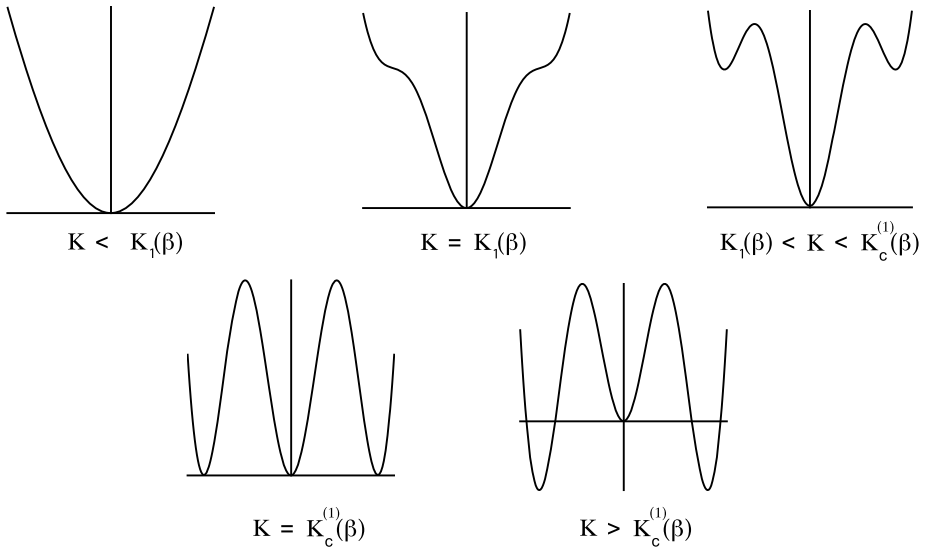


Fig. 2 The free-energy functional $G_{\beta,K}$ for $\beta > \beta_c$

For these values of β , $\tilde{\mathcal{E}}_{\beta,K}$ has the following structure.

- (a) For $0 < K \leq K_c^{(2)}(\beta)$, $\tilde{\mathcal{E}}_{\beta,K} = \{0\}$.
- (b) For $K > K_c^{(2)}(\beta)$, there exists $z(\beta, K) > 0$ such that $\tilde{\mathcal{E}}_{\beta,K} = \{\pm z(\beta, K)\}$.
- (c) $z(\beta, K)$ is a positive, increasing, continuous function for $K > K_c^{(2)}(\beta)$, and as $K \rightarrow (K_c^{(2)}(\beta))^+$, $z(\beta, K) \rightarrow 0$. Therefore, $\tilde{\mathcal{E}}_{\beta,K}$ exhibits a continuous bifurcation at $K_c^{(2)}(\beta)$.

For $\beta \in (0, \beta_c)$, the curve $(\beta, K_c^{(2)}(\beta))$ is the curve of second-order critical points. As we will see in a moment, for $\beta \in (\beta_c, \infty)$ the BC model also has a curve of first-order critical points, which we denote by $(\beta, K_c^{(1)}(\beta))$.

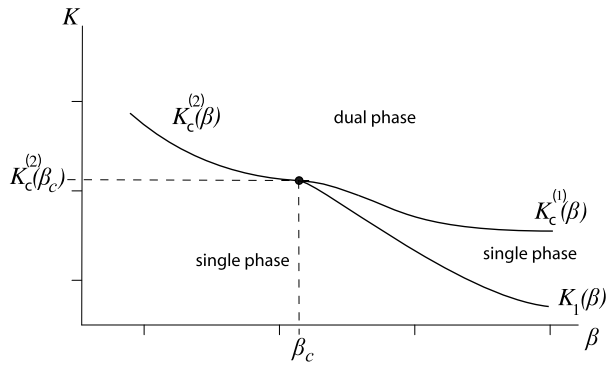
We now describe $\tilde{\mathcal{E}}_{\beta,K}$ for $\beta > \beta_c$. In this case $\tilde{\mathcal{E}}_{\beta,K}$ undergoes a discontinuous bifurcation as K increases through an implicitly defined critical value. Physically, this bifurcation corresponds to a first-order phase transition. The following theorem is proved in Theorem 3.8 in [12].

Theorem 2.3 For all $\beta > \beta_c$, $\tilde{\mathcal{E}}_{\beta,K}$ has the following structure in terms of the quantity $K_c^{(1)}(\beta)$ defined implicitly for $\beta > \beta_c$ on page 2231 of [12].

- (a) For $0 < K < K_c^{(1)}(\beta)$, $\tilde{\mathcal{E}}_{\beta,K} = \{0\}$.
- (b) There exists $z(\beta, K_c^{(1)}(\beta)) > 0$ such that $\tilde{\mathcal{E}}_{\beta,K_c^{(1)}(\beta)} = \{0, \pm z(\beta, K_c^{(1)}(\beta))\}$.
- (c) For $K > K_c^{(1)}(\beta)$ there exists $z(\beta, K) > 0$ such that $\tilde{\mathcal{E}}_{\beta,K} = \{\pm z(\beta, K)\}$.
- (d) $z(\beta, K)$ is a positive, increasing, continuous function for $K \geq K_c^{(1)}(\beta)$, and as $K \rightarrow K_c^{(1)}(\beta)^+$, $z(\beta, K) \rightarrow z(\beta, K_c^{(1)}(\beta)) > 0$. Therefore, $\tilde{\mathcal{E}}_{\beta,K}$ exhibits a discontinuous bifurcation at $K_c^{(1)}(\beta)$.

The phase diagram of the BC model is depicted in Fig. 3. The LDP stated in Theorem 2.1 implies the following weak convergence result used in the proof of rapid mixing in the first-order, discontinuous phase transition region. It is part (a) of Theorem 6.5 in [12].

Fig. 3 Equilibrium phase transition structure of the mean-field Blume-Capel model



Theorem 2.4 For β and K for which $\tilde{\mathcal{E}}_{\beta,K} = \{0\}$,

$$P_{n,\beta,K} \{S_n/n \in dx\} \implies \delta_0 \quad \text{as } n \rightarrow \infty.$$

We end this section with a final result that was not included in the original paper [12] but will be used in the proof of the slow mixing result for the BC model. The result states that not only do the global minimum point of $G_{\beta,K}$ and $I_{\beta,K}$ coincide, but so do the local minimum points.

Lemma 2.5 In the case where $G_{\beta,K}$ and $I_{\beta,K}$ are strictly convex at their minimum points, a point \tilde{z} is a local minimum point of $G_{\beta,K}$ if and only if it is a local minimum point of $I_{\beta,K}$.

Proof Assume that \tilde{z} is a local minimum point of $G_{\beta,K}$. Then \tilde{z} is a critical point of $G_{\beta,K}$ which implies that $\tilde{z} = c'_\beta(2\beta K \tilde{z})$. By the theory of Legendre-Fenchel transforms, $J'_\beta(z) = (c'_\beta)^{-1}(z)$ and thus

$$I'_{\beta,K}(\tilde{z}) = J'_\beta(\tilde{z}) - 2\beta K \tilde{z} = (c'_\beta)^{-1}(\tilde{z}) - 2\beta K \tilde{z} = 0.$$

Next, since \tilde{z} is a local minimum point of $G_{\beta,K}$,

$$G''_{\beta,K}(\tilde{z}) > 0 \iff c''_\beta(2\beta K \tilde{z}) < \frac{1}{2\beta K}.$$

Therefore,

$$I''_{\beta,K}(\tilde{z}) = J''_\beta(\tilde{z}) - 2\beta K = \frac{1}{c''_\beta(2\beta K \tilde{z})} - 2\beta K > 0$$

and we conclude that \tilde{z} is a local minimum point of $I_{\beta,K}$. The other direction is obtained by reversing the argument. \square

3 Glauber Dynamics and Mixing Times

The Markov chain for statistical mechanical models studied in this paper is called the Glauber dynamics, one of the most commonly used physical dynamics for these models. See [4] for more on Glauber dynamics. For the BC model, the Glauber dynamics evolve by

selecting a vertex i at random and updating the spin at i according to the distribution $P_{n,\beta,K}$, conditioned to agree with the spins at all vertices not equal to i . If the current configuration is ω and vertex i is selected, then the chance of the spin at i is updated to $+1$ is equal to

$$p_{+1}(\omega, i) = \frac{e^{2\beta K \tilde{S}(\omega, i)/n}}{e^{2\beta K \tilde{S}(\omega, i)/n} + e^{\beta - (\beta K)/n} + e^{-2\beta K \tilde{S}(\omega, i)/n}} \tag{8}$$

where $\tilde{S}(\omega, i) = \sum_{j \neq i} \omega_j$ is the total spin of the neighboring vertices of i . Similarly, the probabilities of i updating to 0 and -1 are

$$p_0(\omega, i) = \frac{e^{\beta - (\beta K)/n}}{e^{2\beta K \tilde{S}(\omega, i)/n} + e^{\beta - (\beta K)/n} + e^{-2\beta K \tilde{S}(\omega, i)/n}} \tag{9}$$

and

$$p_{-1}(\omega, i) = \frac{e^{-2\beta K \tilde{S}(\omega, i)/n}}{e^{2\beta K \tilde{S}(\omega, i)/n} + e^{\beta - (\beta K)/n} + e^{-2\beta K \tilde{S}(\omega, i)/n}}. \tag{10}$$

$p_{+1}(\omega, i)$ is increasing with respect to $\tilde{S}(\omega, i)$, $p_{-1}(\omega, i)$ is decreasing with respect to $\tilde{S}(\omega, i)$, and $p_0(\omega, i)$ is decreasing for $\tilde{S}(\omega, i) > 0$ and increasing for $\tilde{S}(\omega, i) < 0$.

The mixing time is a measure of the convergence rate of a Markov chain to its stationary distribution and is defined in terms of the *total variation distance* between two distributions μ and ν defined by

$$\|\mu - \nu\|_{TV} = \sup_{A \subset \Omega} |\mu(A) - \nu(A)| = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)|.$$

Given the convergence of the Markov chain, we define the maximal distance to stationary to be

$$d(t) = \max_{x \in \Omega} \|P^t(x, \cdot) - \pi\|_{TV}$$

where $P^t(x, \cdot)$ is the transition probability of the Markov chain starting in configuration x and π is its stationary distribution. Then, given $\varepsilon > 0$, the *mixing time* of the Markov chain is defined by

$$t_{\text{mix}}(\varepsilon) = \min\{t : d(t) \leq \varepsilon\}.$$

In general, rigorous analysis of mixing times is very difficult and the proof of exact mixing time asymptotics of even some basic chains remain elusive. See [15] for a detailed survey on the theory of mixing times.

Rates of mixing times are generally categorized into two groups: *rapid mixing* which implies the mixing time exhibits polynomial growth with respect to the system size, and *slow mixing* which implies that the mixing time grows exponentially with the system size. Determining where a model undergoes rapid mixing is of major importance, as it is in this region that the application of the dynamics is physically feasible.

A classical tool in proving rapid mixing for Markov chains defined on graphs, including the Glauber dynamics of statistical mechanical models, is the path coupling technique [3]. It will be shown that this technique can be directly applied to the BC model in the second-order, continuous phase transition region but fails in a subset of the first-order, discontinuous phase transition region. For the latter region, we develop an extension of the path coupling method we call *aggregate path coupling* to prove rapid mixing. The path coupling method for the BC model is introduced in the next section.

4 Path Coupling

The basis to the path coupling argument used to prove rapid mixing is the following variation of the coupling inequality.

Lemma 4.1 *Let μ and ν be two probability distributions on Ω . Then*

$$\|\mu - \nu\|_{TV} = \inf\{P\{X \neq Y\} : (X, Y) \text{ is a coupling of } \mu \text{ and } \nu\}.$$

This lemma implies that the total variation distance to stationarity of the Glauber dynamics Markov chain of the BC model can be bounded above by the probability $P(X_t \neq Y_t)$ for a coupling chain (X_t, Y_t) starting at some $(X_0, Y_0) = (\sigma, \tau)$. The path coupling method is used to prove rapid mixing in the situations when probability $P(X_t \neq Y_t)$ contracts with each time step regardless of (σ, τ) .

The idea of the path coupling method is to view a coupling that starts in configurations σ and τ as a sequence of couplings that start in *neighboring* configurations (x_i, x_{i+1}) such that $(\sigma = x_0, x_1, x_2, \dots, x_r = \tau)$. Then the contraction of the original coupling distance can be obtained by proving contraction between neighboring configurations which is often easier to show. See [15] for more on path coupling.

We begin by setting up the coupling rules for the Glauber dynamics of the mean-field Blume-Capel model. Define the path metric ρ on $\Omega^n = \{-1, 0, 1\}^n$ by

$$\rho(\sigma, \tau) = \sum_{j=1}^n \mathbf{1}\{\sigma_j \neq \tau_j\}, \tag{11}$$

the number of sites at which the configurations σ and τ differ.

Let σ and τ be two configurations with $\rho(\sigma, \tau) = 1$; i.e. σ and τ are neighboring configurations. The spins of σ and τ agree everywhere except at a single vertex i . Assume that $\sigma_i < \tau_i$. We next describe the path coupling (X, Y) of one step of the Glauber dynamics starting in configuration σ with one starting in configuration τ . Pick a vertex k uniformly at random. We use a single random variable as the common source of noise to update both chains, so the two chains agree as often as possible. In particular, let U be a uniform random variable on $[0, 1]$ and set

$$X(k) = \begin{cases} -1 & \text{if } 0 \leq U \leq p_{-1}(\sigma, k), \\ 0 & \text{if } p_{-1}(\sigma, k) < U \leq p_{-1}(\sigma, k) + p_0(\sigma, k), \\ +1 & \text{if } p_{-1}(\sigma, k) + p_0(\sigma, k) < U \leq 1 \end{cases}$$

and

$$Y(k) = \begin{cases} -1 & \text{if } 0 \leq U \leq p_{-1}(\tau, k), \\ 0 & \text{if } p_{-1}(\tau, k) < U \leq p_{-1}(\tau, k) + p_0(\tau, k), \\ +1 & \text{if } p_{-1}(\tau, k) + p_0(\tau, k) < U \leq 1. \end{cases}$$

Set $X(j) = \sigma_j$ and $Y(j) = \tau_j$ for $j \neq k$.

Since $\sigma_i < \tau_i$, for all $j \neq i$, $\tilde{S}(\sigma, j) < \tilde{S}(\tau, j)$ and thus

$$p_{+1}(\tau, k) > p_{+1}(\sigma, k) \quad \text{and} \quad p_{-1}(\tau, k) < p_{-1}(\sigma, k).$$

The path metric ρ on the coupling above takes on the following possible values,

$$\rho(X, Y) = \begin{cases} 0 & \text{if } k = i, \\ 1 & \text{if } k \neq i \text{ and both chains updates the same,} \\ 2 & \text{if } k \neq i \text{ and the chains update differently.} \end{cases}$$

The application of the path coupling technique to prove rapid mixing is dependent on whether the mean coupling distance with respect to the path metric ρ , denoted by $\mathbb{E}_{\sigma, \tau}[\rho(X, Y)]$, contracts over *all* pairs of neighboring configurations.

In the lemma below and following corollary, we derive a working form for the mean coupling distance.

Lemma 4.2 *Let ρ be the path metric defined in (11) and (X, Y) be the path coupling of one step of the Glauber dynamics of the mean-field Blume-Capel model where X and Y start in neighboring configurations σ and τ . Define*

$$\varphi_{\beta, K}(x) = \frac{2 \sinh(\frac{2\beta K}{n}x)}{2 \cosh(\frac{2\beta K}{n}x) + e^{\beta - \frac{\beta K}{n}}}. \tag{12}$$

Then

$$\mathbb{E}_{\sigma, \tau}[\rho(X, Y)] = \frac{n-1}{n} + \frac{(n-1)}{n} [\varphi_{\beta, K}(S_n(\tau)) - \varphi_{\beta, K}(S_n(\sigma))] + O\left(\frac{1}{n^2}\right).$$

Proof Let n_{-1}, n_0 and n_{+1} denote the number of $-1, 0$ and $+1$ spins, respectively, in configuration σ , not including the spin at vertex i , where the configurations differ. Note that $n_{-1} + n_0 + n_{+1} = n - 1$.

Define $\varepsilon(-1)$ to be the probability that X and Y update differently when the chosen vertex $k \neq i$ is a -1 spin. Similarly, define $\varepsilon(0)$ and $\varepsilon(+1)$. Then the mean coupling distance can be expressed as

$$\begin{aligned} \mathbb{E}_{\sigma, \tau}[\rho(X, Y)] &= \frac{n_{-1}}{n}(1 - \varepsilon(-1)) + \frac{n_0}{n}(1 - \varepsilon(0)) + \frac{n_{+1}}{n}(1 - \varepsilon(+1)) \\ &\quad + 2 \left[\frac{n_{-1}}{n}\varepsilon(-1) + \frac{n_0}{n}\varepsilon(0) + \frac{n_{+1}}{n}\varepsilon(+1) \right] \\ &= \frac{n-1}{n} + \frac{n_{-1}}{n}\varepsilon(-1) + \frac{n_0}{n}\varepsilon(0) + \frac{n_{+1}}{n}\varepsilon(+1). \end{aligned}$$

The probability that X and Y update differently when the chosen vertex $k \neq i$ is a -1 spin is given by

$$\begin{aligned} \varepsilon(-1) &= [p_{-1}(\sigma, k) - p_{-1}(\tau, k)] + [(p_{-1}(\sigma, k) + p_0(\sigma, k)) - (p_{-1}(\tau, k) + p_0(\tau, k))] \\ &= [p_{+1}(\tau, k) - p_{+1}(\sigma, k)] + [p_{-1}(\sigma, k) - p_{-1}(\tau, k)] \\ &= [p_{+1}(\tau, k) - p_{-1}(\tau, k)] + [p_{-1}(\sigma, k) - p_{+1}(\sigma, k)] \\ &= \frac{2 \sinh(\frac{2\beta K}{n}(S_n(\tau) + 1))}{2 \cosh(\frac{2\beta K}{n}(S_n(\tau) + 1)) + e^{\beta - \frac{\beta K}{n}}} - \frac{2 \sinh(\frac{2\beta K}{n}(S_n(\sigma) + 1))}{2 \cosh(\frac{2\beta K}{n}(S_n(\sigma) + 1)) + e^{\beta - \frac{\beta K}{n}}} \end{aligned}$$

$$\begin{aligned} &= \varphi_{\beta,K}((S_n(\tau) + 1)) - \varphi_{\beta,K}((S_n(\sigma) + 1)) \\ &= \varphi_{\beta,K}(S_n(\tau)) - \varphi_{\beta,K}(S_n(\sigma)) + O\left(\frac{1}{n^2}\right). \end{aligned}$$

Similarly, we have

$$\varepsilon(0) = \varphi_{\beta,K}(S_n(\tau)) - \varphi_{\beta,K}(S_n(\sigma))$$

and

$$\varepsilon(+1) = \varphi_{\beta,K}((S_n(\tau) - 1)) - \varphi_{\beta,K}((S_n(\sigma) - 1)) = \varphi_{\beta,K}(S_n(\tau)) - \varphi_{\beta,K}(S_n(\sigma)) + O\left(\frac{1}{n^2}\right)$$

and the proof is complete. □

For c_β defined in (3), we have

$$\varphi_{\beta,K}(x) = c'_\beta \left(\frac{2\beta K}{n} x \right) (1 + O(1/n))$$

which yields the following corollary.

Corollary 4.3 *Let ρ be the path metric defined in (11) and (X, Y) be the path coupling where X and Y start in neighboring configurations σ and τ . Then*

$$\mathbb{E}_{\sigma,\tau}[\rho(X, Y)] = \frac{n-1}{n} + \frac{(n-1)}{n} \left[c'_\beta \left(2\beta K \frac{S_n(\tau)}{n} \right) - c'_\beta \left(2\beta K \frac{S_n(\sigma)}{n} \right) \right] + O\left(\frac{1}{n^2}\right).$$

By the above corollary, we conclude that the mean coupling distance of a coupling starting in neighboring configurations contracts; i.e. $\mathbb{E}_{\sigma,\tau}[\rho(X, Y)] < \rho(\sigma, \tau) = 1$, if

$$\begin{aligned} &\left[c'_\beta \left(2\beta K \frac{S_n(\tau)}{n} \right) - c'_\beta \left(2\beta K \frac{S_n(\sigma)}{n} \right) \right] \\ &\approx 2\beta K \left[\frac{S_n(\tau)}{n} - \frac{S_n(\sigma)}{n} \right] c''_\beta \left(2\beta K \frac{S_n(\sigma)}{n} \right) < \frac{1}{n-1}. \end{aligned}$$

Since σ and τ are neighboring configurations and $S_n(\tau) > S_n(\sigma)$, this is equivalent to

$$c''_\beta \left(2\beta K \frac{S_n(\sigma)}{n} \right) < \frac{1}{2\beta K}. \tag{13}$$

Therefore, contraction of the mean coupling distance, and thus rapid mixing, depends on the concavity behavior of the function c'_β . This is also precisely what determines the type of thermodynamic equilibrium phase transition (continuous, second-order versus discontinuous, first-order) that is exhibited by the mean-field Blume-Capel model. We state the concavity behavior of c'_β in the next theorem which is proved in Theorem 3.5 in [12]. The results of the theorem are depicted in Fig. 4

Theorem 4.4 *For $\beta > \beta_c = \log 4$ define*

$$w_c(\beta) = \cosh^{-1} \left(\frac{1}{2} e^\beta - 4e^{-\beta} \right) \geq 0. \tag{14}$$

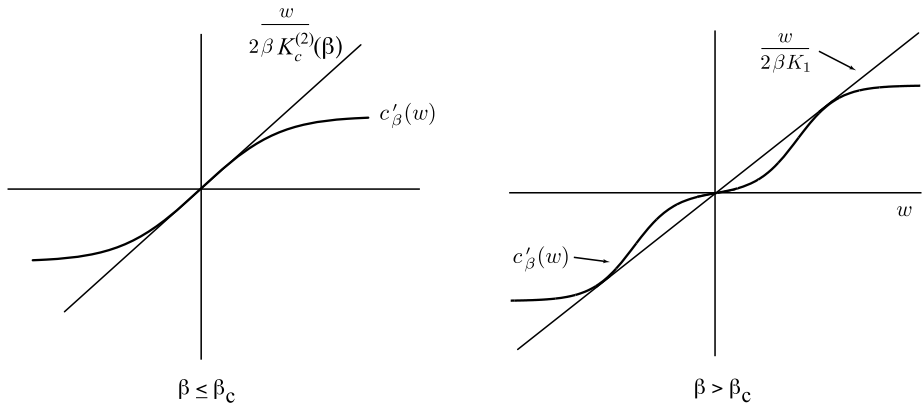


Fig. 4 Behavior of $c'_\beta(w)$ for large and small β

The following conclusions hold.

- (a) For $0 < \beta \leq \beta_c$, $c'_\beta(w)$ is strictly concave for $w > 0$.
- (b) For $\beta > \beta_c$, $c'_\beta(w)$ is strictly convex for $0 < w < w_c(\beta)$ and $c'_\beta(w)$ is strictly concave for $w > w_c(\beta)$.

By part (a) of the above theorem, for $\beta \leq \beta_c$, $c''_\beta(x) \leq c''_\beta(0) = 1/(2\beta K_c^{(2)}(\beta))$. Therefore, by (13), the mean coupling distance contracts between all pairs of neighboring states whenever $K < K_c^{(2)}(\beta)$.

By contrast, for $\beta > \beta_c$, we will show that rapid mixing occurs whenever $K < K_1(\beta)$ where $K_1(\beta)$ is the metastable critical value introduced in Sect. 2 and depicted in Fig. 2. However, since the supremum $\sup_{[-1,1]} c''_\beta(x) > \frac{1}{2\beta K_1(\beta)}$, the condition $K < K_1(\beta)$ is not sufficient for (13) to hold. That is, $K < K_1(\beta)$ does not imply the contraction of the mean coupling distance between all pairs of neighboring states. However, we prove rapid mixing for all $K < K_1(\beta)$ in Sect. 6 by using an extension to the path coupling method that we refer to as *aggregate path coupling*.

We now prove the mixing times for the mean-field Blume-Capel model, which varies depending on the parameter values (β, K) and their position with respect to the thermodynamic phase transition curves. We begin with the case $\beta \leq \beta_c$ where the model undergoes a continuous, second-order phase transition and $K \leq K_c^{(2)}(\beta)$ which corresponds to the single phase region.

5 Rapid Mixing for Continuous Phase Transition Region

We begin by stating the standard path coupling argument used to prove rapid mixing for the mean-field Blume-Capel model in the continuous, second-order phase transition region. The result is proved in Corollary 14.7 of [15].

Theorem 5.1 *Suppose the state space Ω of a Markov chain is the vertex set of a graph with path metric ρ . Suppose that for each edge $\{\sigma, \tau\}$ there exists a coupling (X, Y) of the distributions $P(\sigma, \cdot)$ and $P(\tau, \cdot)$ such that*

$$\mathbb{E}_{\sigma, \tau}[\rho(X, Y)] \leq \rho(\sigma, \tau)e^{-\alpha} \quad \text{for some } \alpha > 0.$$

Then

$$t_{\text{mix}}(\varepsilon) \leq \left\lceil \frac{-\log(\varepsilon) + \log(\text{diam}(\Omega))}{\alpha} \right\rceil.$$

In this section, we assume $\beta \leq \beta_c$ which implies that the BC model undergoes a continuous, second-order phase transition at $K = K_c^{(2)}(\beta)$ defined in (7). By Theorem 4.4, for $\beta \leq \beta_c$, $c'_\beta(x)$ is concave for $x > 0$. See the first graph of Fig. 4 as reference. We next state and prove the rapid mixing result for the mean-field Blume-Capel model in the second-order, continuous phase transition regime.

Theorem 5.2 *Let $t_{\text{mix}}(\varepsilon)$ be the mixing time for the Glauber dynamics of the mean-field Blume-Capel model on n vertices and $K_c^{(2)}(\beta)$ the continuous phase transition curve defined in (7). Then for $\beta \leq \beta_c = \log 4$ and $K < K_c^{(2)}(\beta)$,*

$$t_{\text{mix}}(\varepsilon) \leq \frac{n}{\alpha} (\log n + \log(1/\varepsilon))$$

for any $\alpha \in (0, \frac{K_c^{(2)}(\beta) - K}{K_c^{(2)}(\beta)})$ and n sufficiently large.

Proof Let (X, Y) be a coupling of the Glauber dynamics of the BC model that begin in neighboring configurations σ and τ with respect to the path metric ρ defined in (11). By Corollary 4.3 of Lemma 4.2,

$$\mathbb{E}_{\sigma, \tau}[\rho(X, Y)] = 1 - \left(\frac{1}{n} - \frac{(n-1)}{n} \left[c'_\beta \left(2\beta K \frac{S_n(\tau)}{n} \right) - c'_\beta \left(2\beta K \frac{S_n(\sigma)}{n} \right) \right] \right) + O\left(\frac{1}{n^2}\right).$$

Observe that c''_β is an even function and that for $\beta \leq \beta_c$, $\sup_x c''_\beta(x) = c''_\beta(0)$. Therefore, by the mean value theorem and Theorem 2.2,

$$\begin{aligned} \mathbb{E}_{\sigma, \tau}[\rho(X, Y)] &\leq 1 - \frac{[1 - (n-1)(2\beta K/n)c''_\beta(0)]}{n} + O\left(\frac{1}{n^2}\right) \\ &\leq \exp \left\{ -\frac{1 - 2\beta K c''_\beta(0)}{n} + O\left(\frac{1}{n^2}\right) \right\} \\ &= \exp \left\{ \frac{1}{n} \left(\frac{K_c^{(2)}(\beta) - K}{K_c^{(2)}(\beta)} \right) + O\left(\frac{1}{n^2}\right) \right\} \\ &< e^{-\alpha/n} \end{aligned}$$

for any $\alpha \in (0, \frac{K_c^{(2)}(\beta) - K}{K_c^{(2)}(\beta)})$ and n sufficiently large. Thus, for $K < K_c^{(2)}(\beta)$, we can apply Theorem 5.1, where the diameter of the configuration space of the BC model Ω^n is n , to complete the proof. \square

6 Rapid Mixing for Discontinuous Phase Transition Region

Here we consider the region $\beta > \beta_c$, where the mean-field Blume-Capel model undergoes a first-order discontinuous phase transition. In this region, the function $c'_\beta(x)$ which determines whether the mean coupling distance contracts (Corollary 4.3) is no longer strictly

concave for $x > 0$ (Theorem 4.4). See the second graph in Fig. 4 for reference. We will show that rapid mixing occurs whenever $K < K_1(\beta)$ where $K_1(\beta)$ is the metastable critical value defined in Sect. 2 and depicted in Fig. 2.

As shown in Sect. 4, in order to apply the standard path coupling technique of Theorem 5.1, we need the inequality (13) to hold for all values of $S_n(\sigma)$ and thus $\sup_{[-1,1]} c''_\beta(x) < \frac{1}{2\beta K}$. However since $\sup_{[-1,1]} c''_\beta(x) > \frac{1}{2\beta K_1(\beta)}$, the condition $K < K_1(\beta)$ is not sufficient for the contraction of the mean coupling distance between all pairs of neighboring states which is required to prove rapid mixing using the standard path coupling technique stated in Theorem 5.1.

In order to prove rapid mixing in the region where $\beta > \beta_c$ and $K < K_1(\beta)$, we take advantage of the result in Theorem 2.4 which states the weak convergence of the magnetization S_n/n to a point-mass at the origin. Thus, in the coupling of the dynamics, the magnetization of the process that starts at equilibrium will stay mainly near the origin. As a result, for two starting configurations σ and τ , one of which has near-zero magnetization ($S_n(\sigma)/n \approx 0$), the mean coupling distance of a coupling starting in these configurations will be the aggregate of the mean coupling distances between neighboring states along a minimal path connecting the two configurations. Although not all pairs of neighbors in the path will contract, we show that in the aggregate, contraction between the two configurations still holds.

In the next lemma we prove contraction of the mean coupling distance in the aggregate and then the rapid mixing result for the mean-field Blume-Capel model is proved in the theorem following the lemma by applying the new aggregate path coupling method.

Lemma 6.1 *Let (X, Y) be a coupling of one step of the Glauber dynamics of the BC model that begin in configurations σ and τ , not necessarily neighbors with respect to the path metric ρ defined in (11). Suppose $\beta > \beta_c$ and $K < K_1(\beta)$. Then for any $\alpha \in (0, \frac{K_1(\beta) - K}{K_1(\beta)})$ there exists an $\varepsilon > 0$ such that, asymptotically as $n \rightarrow \infty$,*

$$\mathbb{E}_{\sigma, \tau}[\rho(X, Y)] \leq e^{-\alpha/n} \rho(\sigma, \tau) \tag{15}$$

whenever $|S_n(\sigma)| < \varepsilon n$.

Proof Observe that for $\beta > \beta_c$ and $K < K_1(\beta)$,

$$|c'_\beta(x)| \leq \frac{|x|}{2\beta K_1(\beta)} \quad \text{for all } x.$$

We will show that for a given $\alpha' \in (\frac{1}{2\beta K_1(\beta)}, \frac{1-\alpha}{2\beta K})$, there exists $\varepsilon > 0$ such that

$$c'_\beta(x) - c'_\beta(x_0) \leq \alpha'(x - x_0) \quad \text{whenever } |x_0| < \varepsilon \tag{16}$$

as $c'_\beta(x)$ is a continuously differentiable increasing odd function and $c'_\beta(0) = 0$.

In order to show (16), observe that $c''_\beta(0) = \frac{1}{2\beta K_c^{(2)}(\beta)} < \frac{1}{2\beta K_1(\beta)}$, and since c''_β is continuous, there exists a $\delta > 0$ such that

$$c''_\beta(x) < \alpha' \quad \text{whenever } |x| < \delta.$$

The mean value theorem implies that

$$c'_\beta(x) - c'_\beta(x_0) < \alpha'(x - x_0) \quad \text{for all } x_0, x \in (-\delta, \delta).$$

Now, let $\varepsilon = \frac{\alpha' - 1 / (2\beta K_1(\beta))}{\alpha' + 1 / (2\beta K_1(\beta))} \delta < \delta$. Then for any $|x_0| < \varepsilon$ and $|x| \geq \delta$,

$$\begin{aligned} |c'_\beta(x) - c'_\beta(x_0)| &\leq \frac{|x| + |x_0|}{2\beta K_1(\beta)} \\ &\leq \frac{(1 + \varepsilon/\delta)|x|}{2\beta K_1(\beta)} \\ &= \frac{|x - x_0|}{2\beta K_1(\beta)} \cdot \frac{1 + \varepsilon/\delta}{|1 - x_0/x|} \\ &\leq \frac{|x - x_0|}{2\beta K_1(\beta)} \cdot \frac{1 + \varepsilon/\delta}{1 - \varepsilon/\delta} \\ &= \alpha'|x - x_0|. \end{aligned}$$

Without loss of generality suppose that $S_n(\sigma) < S_n(\tau)$. Let $(\sigma = x_0, x_1, \dots, x_r = \tau)$ be a path connecting σ to τ and monotone increasing in ρ such that (x_{i-1}, x_i) are neighboring configurations. Here $r = \rho(\sigma, \tau)$. Then by Corollary 4.3 of Lemma 4.2 and (16), we have for $|S_n(\sigma)| < \varepsilon n$ and asymptotically as $n \rightarrow \infty$,

$$\begin{aligned} \mathbb{E}_{\sigma, \tau}[\rho(X, Y)] &\leq \sum_{i=1}^r \mathbb{E}_{x_{i-1}, x_i}[\rho(X_{i-1}, X_i)] \\ &= \frac{(n-1)}{n} \rho(\sigma, \tau) + \frac{(n-1)}{n} \left[c'_\beta\left(\frac{2\beta K}{n} S_n(\tau)\right) - c'_\beta\left(\frac{2\beta K}{n} S_n(\sigma)\right) \right] \\ &\quad + \rho(\sigma, \tau) \cdot O\left(\frac{1}{n^2}\right) \\ &\leq \frac{(n-1)}{n} \rho(\sigma, \tau) + \frac{(n-1)}{n} (S_n(\tau) - S_n(\sigma)) \frac{2\beta K \alpha'}{n} + \rho(\sigma, \tau) \cdot O\left(\frac{1}{n^2}\right) \\ &\leq \rho(\sigma, \tau) \left[1 - \left(\frac{1 - 2\beta K \alpha'}{n}\right) + O\left(\frac{1}{n^2}\right) \right] \\ &\leq e^{-\alpha'/n} \rho(\sigma, \tau). \end{aligned}$$

This completes the proof. □

Theorem 6.2 *Let $t_{\text{mix}}(\varepsilon)$ be the mixing time for the Glauber dynamics of the mean-field Blume-Capel model on n vertices and $K_1(\beta)$ be the metastable critical point. Then, for $\beta > \beta_c$ and $K < K_1(\beta)$,*

$$t_{\text{mix}}(\varepsilon) \leq \frac{n}{\alpha} (\log n + \log(2/\varepsilon))$$

for any $\alpha \in (0, \frac{K_1(\beta) - K}{K_1(\beta)})$ and n sufficiently large.

Proof Let (X_t, Y_t) be a coupling of the Glauber dynamics of the BC model such that $Y_0 \stackrel{\text{dist}}{=} P_{n, \beta, K}$, the stationary distribution. For a given $\alpha \in (0, \frac{K_1(\beta) - K}{K_1(\beta)})$, let ε be as in Lemma 6.1. For sufficiently large n ,

$$\begin{aligned}
 \|P^t(X_0, \cdot) - P_{n,\beta,K}\|_{TV} &\leq P\{X_t \neq Y_t\} \\
 &= P\{\rho(X_t, Y_t) \geq 1\} \\
 &\leq \mathbb{E}[\rho(X_t, Y_t)] \\
 &= \mathbb{E}[\mathbb{E}[\rho(X_t, Y_t) | X_{t-1}, Y_{t-1}]] \\
 &\leq \mathbb{E}[\mathbb{E}[\rho(X_t, Y_t) | X_{t-1}, Y_{t-1}] | |S_n(Y_{t-1})| < \varepsilon n] \\
 &\quad \cdot P\{|S_n(Y_{t-1})| < \varepsilon n\} + nP\{|S_n(Y_{t-1})| \geq \varepsilon n\}.
 \end{aligned}$$

By iterating (15), it follows that

$$\begin{aligned}
 \|P^t(X_0, \cdot) - P_{n,\beta,K}\|_{TV} &\leq e^{-\alpha/n} \mathbb{E}[\rho(X_{t-1}, Y_{t-1}) | |S_n(Y_{t-1})| < \varepsilon n] \\
 &\quad \cdot P\{|S_n(Y_{t-1})| < \varepsilon n\} + nP\{|S_n(Y_{t-1})| \geq \varepsilon n\} \\
 &\leq e^{-\alpha/n} \mathbb{E}[\rho(X_{t-1}, Y_{t-1})] + nP\{|S_n(Y_{t-1})| \geq \varepsilon n\} \\
 &\quad \vdots \\
 &\leq e^{-\alpha t/n} \mathbb{E}[\rho(X_0, Y_0)] + n \sum_{s=0}^{t-1} P\{|S_n(Y_s)| \geq \varepsilon n\} \\
 &= e^{-\alpha t/n} \mathbb{E}[\rho(X_0, Y_0)] + nt P_{n,\beta,K}\{|S_n/n| \geq \varepsilon\} \\
 &\leq ne^{-\alpha t/n} + nt P_{n,\beta,K}\{|S_n/n| \geq \varepsilon\}.
 \end{aligned}$$

We recall the result in Theorem 2.4 that for $\beta > \beta_c$ and $K < K_1(\beta)$

$$P_{n,\beta,K}\{S_n/n \in dx\} \implies \delta_0 \quad \text{as } n \rightarrow \infty.$$

Moreover, for any $\gamma > 1$ and n sufficiently large, the LDP stated in Theorem 2.1 implies that

$$\begin{aligned}
 \|P^t(X_0, \cdot) - P_{n,\beta,K}\|_{TV} &\leq ne^{-\alpha t/n} + nt P_{n,\beta,K}\{|S_n/n| \geq \varepsilon\} \\
 &< ne^{-\alpha t/n} + tne^{-\frac{n}{\gamma} I_{\beta,K}(\varepsilon)}.
 \end{aligned}$$

For $t = \frac{n}{\alpha}(\log n + \log(2/\varepsilon))$, the above right-hand side converges to $\varepsilon/2$ as $n \rightarrow \infty$. □

7 Slow Mixing

We complete our mixing time analysis of the mean-field Blume-Capel model by determining the slow mixing region of the (β, K) parameter space for both the continuous, second-order and discontinuous, first-order phase transition regions. To prove slow mixing for the BC model, we use the bottleneck or Cheeger constant argument which we state in the theorem below. The theorem is an application of Theorem 7.3 in [15] to the BC model.

Theorem 7.1 *For two configurations ω and τ , define the edge measure Q as follows:*

$$Q(\omega, \tau) = P_{n,\beta,K}(\omega)P(\omega, \tau) \quad \text{and} \quad Q(A, B) = \sum_{\omega \in A, \tau \in B} Q(\omega, \tau).$$

Here $P(\omega, \tau)$ is the transition probability corresponding to the Glauber dynamics of the mean-field Blume-Capel model. The bottleneck ratio of the set S is defined by

$$\Phi(S) = \frac{Q(S, S^c)}{P_{n,\beta,K}(S)} \quad \text{and} \quad \Phi_* = \min_{S: P_{n,\beta,K}(S) \leq \frac{1}{2}} \Phi(S).$$

Under these assumptions, we have

$$t_{\text{mix}} = t_{\text{mix}}(1/4) \geq \frac{1}{4\Phi_*}.$$

The next lemma states that bottlenecks, and thus slow mixing, occurs for the BC model whenever the free energy functional $G_{\beta,K}$ has a positive minimum point.

Lemma 7.2 *Suppose the free energy functional $G_{\beta,K}$ defined in (5) has a positive minimum (either local or global) point. Then there exists a positive constant b and a strictly positive function $r(\beta, K)$ such that*

$$t_{\text{mix}} \geq b e^{r(\beta,K)n}.$$

Proof Suppose $G_{\beta,K}$ has a minimum (either local or global) point at $\tilde{z} > 0$. Let z' be the corresponding local maximum point of $G_{\beta,K}$ such that $0 \leq z' < \tilde{z}$. Define the bottleneck set

$$A = \left\{ \omega : z' < \frac{S_n(\omega)}{n} \leq 1 \right\}.$$

Since $z' \geq 0$, by the symmetry of $G_{\beta,K}$, $P_{n,\beta,K}(A) \leq \frac{1}{2}$. With respect to the Glauber dynamics of the BC model, in order to leave the set A in one step of the chain, the current configuration must lie in one of the two boundary sets

$$A_1 = \left\{ \omega : \frac{S_n(\omega)}{n} = z' + \frac{1}{n} \right\} \quad \text{or} \quad A_2 = \left\{ \omega : \frac{S_n(\omega)}{n} = z' + \frac{2}{n} \right\}.$$

We take n sufficiently large so that $z' + 2/n < \tilde{z}$.

For any $\gamma > 1$, the large deviations upper bound (3) implies that the edge measure Q for the bottleneck set A satisfies

$$Q(A, A^c) \leq P_{n,\beta,K}(A_1) + P_{n,\beta,K}(A_2) < e^{-\frac{n}{\gamma} I_{\beta,K}(z' + \frac{2}{n})}$$

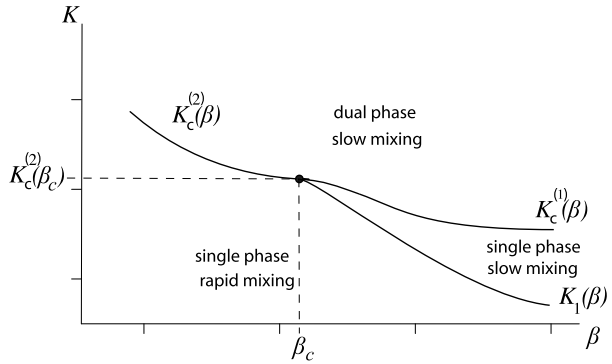
for n large enough. Moreover, since the LDP implies that $P_{n,\beta,K}(A) > e^{-n\gamma I_{\beta,K}(\tilde{z})}$, the bottleneck ratio satisfies

$$\Phi(A) < 2e^{-n[\gamma^{-1} I_{\beta,K}(z' + \frac{2}{n}) - \gamma I_{\beta,K}(\tilde{z})]},$$

all for a given $\gamma > 1$ and n large enough. Lastly, by Lemma 2.5 \tilde{z} is a local minimum point of $I_{\beta,K}$ since it is a local minimum point of $G_{\beta,K}$, and thus $I_{\beta,K}(z') > I_{\beta,K}(\tilde{z})$ and the proof is complete by Theorem 7.1. □

Next we state the slow mixing result for the mean-field Blume-Capel model. The result follows from Lemma 7.2 and the region in the (β, K) parameter space at which the free-energy functional $G_{\beta,K}$ possesses a positive minimum point. These regions were determined in [12] and depicted in Figs. 1 and 2.

Fig. 5 Mixing times and equilibrium phase transition structure of the mean-field Blume-Capel model



Corollary 7.3 Let $t_{\text{mix}} = t_{\text{mix}}(1/4)$ be the mixing time for the Glauber dynamics of the mean-field Blume-Capel model on n vertices. For (a) $\beta \leq \beta_c$ and $K > K_c^{(2)}(\beta)$, and (b) $\beta > \beta_c$ and $K > K_1(\beta)$, there exists a positive constant b and a strictly positive function $r(\beta, K)$ such that

$$t_{\text{mix}} \geq b e^{r(\beta, K)n}.$$

We summarize the mixing time results for the mean-field Blume-Capel model and its relationship to the model's thermodynamic phase transition structure in Fig. 5. As shown in the figure, in the second-order, continuous phase transition region ($\beta \leq \beta_c$) for the BC model, the mixing time transition coincides with the equilibrium phase transition. This is consistent with other models that exhibit this type of phase transition. However, in the first-order, discontinuous phase transition region ($\beta > \beta_c$) the mixing time transition occurs below the equilibrium phase transition at the metastable critical value.

References

1. Blume, M.: Theory of the first-order magnetic phase change in UO_2 . *Phys. Rev.* **141**, 517–524 (1966)
2. Blume, M., Emery, V.J., Griffiths, R.B.: Ising model for the λ transition and phase separation in He^3 – He^4 mixtures. *Phys. Rev. A* **4**, 1071–1077 (1971)
3. Bubley, R., Dyer, M.: Path coupling: a technique for proving rapid mixing in Markov chains. In: Proceedings of the 38th Annual Symposium on Foundations of Computer Science, pp. 223–231 (1997)
4. Brémaud, P.: Markov Chains, Gibbs Fields, Monte Carlo Simulation, and Queues. Texts in Applied Mathematics, vol. 31. Springer, New York (1999)
5. Capel, H.W.: On the possibility of first-order phase transitions in Ising systems of triplet ions with zero-field splitting. *Physica* **32**, 966–988 (1966)
6. Capel, H.W.: On the possibility of first-order phase transitions in Ising systems of triplet ions with zero-field splitting II. *Physica* **33**, 295–331 (1967)
7. Capel, H.W.: On the possibility of first-order phase transitions in Ising systems of triplet ions with zero-field splitting III. *Physica* **37**, 423–441 (1967)
8. Costeniuc, M., Ellis, R.S., Touchette, H.: Complete analysis of phase transitions and ensemble equivalence for the Curie-Weiss-Potts model. *J. Math. Phys.* **46**, 063301 (2005)
9. Ding, J., Lubetzky, E., Peres, Y.: The mixing time evolution of Glauber dynamics for the mean-field Ising model. *Commun. Math. Phys.* **289**, 725–764 (2009)
10. Ellis, R.S.: Entropy, Large Deviations and Statistical Mechanics. Springer, New York (1985). Reprinted in 2006 in *Classics in Mathematics*
11. Ellis, R.S., Haven, K., Turkington, B.: Large deviation principles and complete equivalence and nonequivalence results for pure and mixed ensembles. *J. Stat. Phys.* **101**, 999–1064 (2000)
12. Ellis, R.S., Otto, P.T., Touchette, H.: Analysis of phase transitions in the mean-field Blume-Emery-Griffiths model. *Ann. Appl. Probab.* **15**, 2203–2254 (2005)

13. Ellis, R.S., Machta, J., Otto, P.T.: Asymptotic behavior of the magnetization near critical and tricritical points via Ginzburg-Landau polynomials. *J. Stat. Phys.* **133**, 101–129 (2008)
14. Ellis, R.S., Machta, J., Otto, P.T.: Asymptotic behavior of the finite-size magnetization as a function of the speed of approach to criticality. *Ann. Appl. Probab.* **20**, 2118–2161 (2010)
15. Levin, D., Peres, Y., Wilmer, E.: *Markov Chains and Mixing Times*. Am. Math. Soc., Providence (2009)
16. Levin, D., Luczak, M., Peres, Y.: Glauber dynamics for the mean-field Ising model: cut-off, critical power law, and metastability. *Probab. Theory Relat. Fields* **146** (2010)
17. Wu, F.Y.: The Potts model. *Rev. Mod. Phys.* **54** 235–268 (1982)