Equi-energy sampling does not converge rapidly on the mean-field Potts model with three colors close to the critical temperature

Mirko Ebbers and Matthias Löwe

Fachbereich Mathematik und Informatik, Universität Münster, Einsteinstraße 62, 48149 Münster, Germany
E-mail: mirkoebbers@uni-muenster.de and maloewe@uni-muenster.de

Received 22 September 2019, revised 3 January 2020
Accepted for publication 7 February 2020
Published 25 February 2020

Abstract

Equi-energy sampling (EES, for short) is a method to speed up the convergence of the Metropolis chain, when the latter is slow. We show that there are still models like the mean-field Potts model, where EES does not converge rapidly in certain temperature regimes. Indeed we will show that EES is slowly mixing on the mean-field Potts model, in a regime below the critical temperature. Though we will concentrate on the Potts model with three colors, our arguments remain valid for any number of colors $q \geq 3$, if we adapt the temperature regime. For the situation of the mean-field Potts model this answers a question posed in Hua and Kou (2011 Stat. Sin. 21 1687–711).

Keywords: equi-energy algorithm, Potts model, Metropolis algorithm, Markov chain Monte Carlo method, Curie–Weiss model

1. Introduction

Sampling methods are of utmost importance in applied mathematics, e.g. in Bayesian statistics, computational physics, econometrics, or computational biology. In many cases one wants to sample a random element drawn from a finite set $\Omega$ according to a probability distribution $\pi$ on $(\Omega, \mathcal{P}(\Omega))$. But even this problem may be less trivial than it sounds. Sometimes $\Omega$ may be finite, yet very large. E.g. when modeling a ferromagnetic material on $N$ atoms, $\Omega$ is of the form $\{-1, +1\}^N$ and for real size systems, $N$ is of the order $10^{23}$, thus $|\Omega|$ is of the order $10^{46}$.
Hence a straightforward Monte-Carlo simulation would take exponentially long in the system size $N$ and thus would be much too expensive. In other situations, the size of $\Omega$ may not even be known, as e.g. in the so-called knapsack problem (Löwe and Meise 2001).

One potential solution of this problem lies in the use of Markov Chain Monte Carlo (MCMC, for short) algorithms. They rely on an aperiodic and irreducible Markov chain on $\Omega$ that has $\pi$ as its invariant (i.e. stationary) distribution. One runs this Markov chain, stops it after some long enough time, and takes the current state as a sample element. The ergodic Theorem for Markov chains ensures that this element is almost distributed according to $\pi$, given one has waited long enough. This method immediately raises two questions:

1. Can one find for each $\pi$ a Markov chain that converges in distribution to $\pi$? I.e. can one find for each $\pi$ an irreducible and aperiodic Markov chain that has $\pi$ as its invariant measure? This question is answered in the affirmative by the Metropolis–Hastings chain (see e.g. Häggström (2002)).

2. How long do we need to wait to get a sample with a distribution that is reasonably close to $\pi$? If this waiting time is polynomial in the problem instance we speak about fast or rapid convergence, otherwise, in particular, if the mixing time is exponential in the problem instance, we will say the algorithm converges torpidly or slowly.

It is, however, well known that like the Glauber dynamics the Metropolis–Hastings algorithm usually converges slowly, when the target distribution is multi-modal. Such situations occur e.g. in statistical physics in the presence of a phase transition. Hence slow convergence applies to a number of interesting situations, among them the low temperature phase of the Curie–Weiss model (see e.g. the discussion in Mossel and Sly (2013)). In the next section we will introduce a close relative of the Curie–Weiss model, the three state mean-field Potts model. This will be our test model for the EES to be introduced in section 3. In section 4 we will show that this sampler mixes slowly when applied to the (three state) mean-field Potts model in a certain temperature regime. Here, a key argument is based on a property of the mean-field Potts model that is closely related to its first order phase transition (see lemma 4.3 and the remark below it): The limit point of the order parameter $m_N$ (see (2.1)) at high temperatures remains a local maximum of its distribution also in a certain part of the low temperature regime. Hence a Metropolis–Hastings chain started in a neighborhood of this high temperature limit point will typically not escape from this neighborhood in polynomial time (in the part of the low temperature regime described above). But then also EES cannot improve the performance of the Metropolis–Hastings algorithm, because there are simply no observations of the global maxima of the distribution of $m_N$ the algorithm could jump at. The proof is completed by combining this observation with the very powerful technical argument of conductance, also known as Cheeger’s inequality (see theorem 4.4).

2. The mean-field Potts model

Let us now introduce the mean-field Potts model. Consider the space $\Omega = E^N$, where $E = \{1, 2, \ldots, q\}$, $q \in \mathbb{N}$, $q \geq 3$, and $N \in \mathbb{N}$ (to avoid some complications in the future, we can think of $N$ being a multiple of $q$). The elements of $E$ are sometimes referred to as colors.

For convenience, in this note, we will restrict to the case of $q = 3$, i.e. the mean-field Potts model with three states or colors taken from the set $E = \{1, 2, 3\}$. This from now on will be our standing assumption for the rest of the note. However, we remark that our argument remains valid for general $q \geq 3$, if one changes the regime of temperatures appropriately. We will come back to this remark later.
On $\Omega$ we construct an energy function given by

$$H_N(\sigma) := \frac{1}{2N} \sum_{i,j=1}^{N} \delta_{\sigma_i = \sigma_j}, \quad \sigma \in \Omega.$$  

Here $\delta_\sigma$ denotes the indicator function for an event $A$ (which is formally the Dirac measure for the event $A$, to stress that our notation is consistent with our later use of $\delta$). Note that $H_N$ can be written as a function of the vector

$$m_N(\sigma) := (m_N^1(\sigma), m_N^2(\sigma), m_N^3(\sigma))$$

$$:= \left( \frac{1}{N} \sum_{i=1}^{N} \delta_{\sigma_i = 1}, \frac{1}{N} \sum_{i=1}^{N} \delta_{\sigma_i = 2}, \frac{1}{N} \sum_{i=1}^{N} \delta_{\sigma_i = 3} \right).$$  

Indeed, one easily checks that

$$H_N(\sigma) = \frac{N}{2} \sum_{i=1}^{3} (m_N^i(\sigma))^2.$$

$m_N$ is therefore called an order parameter of the model. With $H_N$ we associate a Gibbs measure $\pi_\beta$ at inverse temperature $\beta > 0$, i.e.

$$\pi_\beta(\sigma) = \frac{e^{\beta H_N(\sigma)}}{Z_\beta}, \quad \sigma \in \Omega.$$  

Here $Z_\beta = \sum_{\sigma} e^{\beta H(\sigma)}$ is the partition function. Note that conventionally in statistical physics the energy function $H$ would carry an additional minus sign, and the Gibbs measure (as well as the partition function) would be defined in terms of the exponential of minus $\beta$ times that energy. Since the two minus signs would cancel and lead to the same definition of the Gibbs measure our Gibbs measure does not carry these conventional minus signs.

The mean-field Potts model was studied in a variety of papers. We refer to Ellis and Wang (1990) and Kesten and Schonmann (1989), who showed that there is a critical inverse temperature $\beta_c$. This critical inverse temperature in the 3 states mean-field Potts model equals $\beta_c = 4 \log 2$ (see Cuff et al (2012), which discusses the very interesting phenomenon of a temperature-dependent cut-off effect for the Glauber dynamics of the model).

At $\beta_c$ the model undergoes a first order phase transition. More precisely, the order parameter $m_N$ of the model converges in distribution to $a_0 := (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, when $\beta < \beta_c$. At smaller temperatures one observes the following: For $\beta \geq \beta_c$ there is $1 > m^*(\beta) > \frac{1}{3}$ and vectors $a_1(\beta), a_2(\beta), a_3(\beta) \in \mathbb{R}^3$ such that the vector $a_i(\beta)$ has $m^*(\beta)$ in its $i$th component and all other components are equal such that they sum up to one. For $\beta > \beta_c$ the distribution of $m_N$ converges to $\frac{1}{3} \sum_{i=1}^{3} \delta_{a_i(\beta)}$. Here $\delta$ denotes the Dirac-measure. Finally, for $\beta = \beta_c$ there are, moreover, weights $\lambda_1, \lambda_2 > 0$ that sum up to 1, such that the distribution of $m_N$ converges to $\lambda_1 \delta_{a_1} + \lambda_2 \sum_{i=1}^{3} \delta_{a_i(\beta)}$.

The phase transition is of first order, since $m^*(\beta_c) > \frac{1}{3}$, i.e. the jump is discontinuous. Moreover, the vector $a_0$ remains a local maximum of the distribution of $m_N$ for some temperatures below the critical temperature. Such a behavior can also be observed for general values of $q \geq 3$ at other values for $\beta_c$. We will come back to this fact in section 4, lemmas 4.2 and 4.3, because it is of utmost importance for the proof of theorem 4.1 to be given in section 4.
3. Equi-energy sampling

Various modifications of the Metropolis–Hastings algorithm have been proposed to speed up its convergence. Among them the so-called swapping algorithm (see Geyer (1991)), the exchange Monte Carlo method (see Hukushima and Nemoto (1996)), parallel tempering (see Orlandini (1998)) and the simulated tempering algorithm (see Marinari and Parisi (1992), Geyer and Thompson (1995), and Madras (1998)) are very popular in applications. Another variant are Multicanonical Monte Carlo Simulations, introduced by Berg and Neuhaus (1992), also see Berg (2000). It is related to umbrella sampling (see Torrie and Valleau (1977)) and is close in spirit to the swapping algorithm, simulated tempering, as well as EES. A major difference, however, is, how an a priori estimate of the probability distribution of interest is produced. Therefore, we have not yet been able to analyze so far, whether Multicanonical Monte Carlo Simulations suffer from the same shortcomings as swapping, parallel tempering or EES (see next paragraph).

In many situations the algorithms described in the previous paragraph seem indeed to be able to improve the convergence of the Metropolis chain, however, there are also negative theoretical results about these algorithms. Madras and Zheng (2003) show that the swapping chain converges quickly for the Curie–Weiss model. On the other hand, Bhatnagar and Randall (2004) and Bhatnagar and Randall (2016) prove that both, the swapping algorithm and simulated tempering, are slowly mixing for the 3-state Potts model and conjecture that this is caused by the first order phase transition in the Potts model (also see our discussion in the remark following lemma 4.3). Qualitative properties of the swapping algorithm and parallel tempering were studied in Doll et al (2018). A first rapid convergence result for the Swapping Algorithm in an disordered situation was proved in Löwe and Vermet (2009). Ebbers and Löwe (2009) show that in disordered models the conjecture by Bhatnagar and Randall is not correct. They prove that the Swapping Algorithm mixes slowly on the Random Energy Model, even though this model has only a third order phase transition. In the Blume-Emery-Griffiths model both, rapid or torpid mixing may occur as was shown in Ebbers et al (2014).

Another idea to improve the performance of the Metropolis chain is the so called Equi-energy sampling algorithm (see e.g. Kou et al (2006)). This model was tested on the Ising model in Hua and Kou (2011) and the question, how fast it converges, was posed. For the Potts model, we will answer this question in the next section. In particular, we will show that EES may be slowly mixing on relevant models from statistical mechanics. Variants of EES were studied, among others, in Baragatti et al (2013).

The principle observation to motivate EES is that a main obstacle to fast mixing is the presence of a phase transition in the model. This, in turn, may be characterized by a multi-modal distribution of a macroscopic observable. Usually, then the (projected) Metropolis chain enters one of the modes rapidly and stays there for an exponentially long time. The EES tries to avoid this behavior by introducing shortcuts in the state space. These shortcuts are created by the observations of Metropolis chains at higher temperatures where the above mentioned modes are less pronounced or possibly not even present. More precisely, additionally to the Metropolis steps one allows also for jumps to points of the same or a similar energy as the present one, given one has observed these points already at higher temperatures (otherwise, the algorithm would require the exact structure of the energy function, in which case simulations would probably be pointless). The EES has been discussed in Kou et al (2006), its convergence was shown in the same article, and, using a different technique, in Andrieu et al (2008). We will now give an exact description of a version of this algorithm.
Let us first briefly recall the Metropolis–Hastings algorithm, which is the basis of the EES. To define the first let \( K_{\text{gen}} \) denote the following aperiodic, symmetric and irreducible Markov chain on \( \Omega \):

\[
K_{\text{gen}}(\sigma, \tau) = \begin{cases} 
\frac{1}{N} & \text{if } \sigma = \tau \\
\frac{1}{2N} & \text{if } d_H(\sigma, \tau) = 1 \\
0 & \text{otherwise.}
\end{cases}
\]

(3.1)

Here \( d_H \) is the Hamming distance and \( K_{\text{gen}} \) is a Markov chain, because every \( \sigma \in \Omega \) has \( 2N \) neighbors. Define the corresponding Metropolis–Hastings chain for the probability \( \pi_\beta \) as \( T_\beta(\cdot, \cdot) \):

\[
T_\beta(x, y) = \begin{cases} 
K_{\text{gen}}(x, y) & \text{if } x \neq y \text{ and } H(y) \geq H(x) \\
K_{\text{gen}}(x, y) \frac{\pi_\beta(x)}{\pi_\beta(y)} & \text{if } x \neq y \text{ and } H(y) < H(x) \\
1 - \sum_{z \neq x} T_\beta(x, z) & \text{otherwise.}
\end{cases}
\]

(3.2)

Note that \( T_\beta(\cdot, \cdot) \) sometimes is slow in natural situations, e.g. when sampling from the low temperature distribution of the Curie–Weiss model (see e.g. Madras and Piccioni (1999), of course, the \( T_\beta \) has to be adapted to the situation of the Curie–Weiss model). To speed up its convergence, we consider the EES. To define it, we first introduce a sequence of energy levels:

\[
\text{inf}_x H(x) := h_0 < h_1 < \ldots < h_M = \sup_x H(x).
\]

In our context, it is easily checked that

\[
\text{inf}_x H(x) = \frac{N}{6} \quad \text{and} \quad \sup_x H(x) = \frac{N}{2}
\]

(3.3)

which will be used later. Moreover, introduce a sequence of inverse temperature levels

\[
0 = \beta_0 < \beta_1 < \ldots < \beta_M = \beta
\]

where we assume that \( \beta \) is the temperature we want to sample from. It will often be convenient to take \( \beta_i = i \frac{6}{M} \). Note that \( M \) may and will depend on \( N \), which is not made explicit in Kou et al (2006), otherwise our construction, so far, agrees with the construction in Kou et al (2006). We will make an explicit choice for \( M \) and give reasons for this choice, after the description of the algorithm.

For this, we will also need a dummy state \( \tau \) and define \( \tilde{\Omega} := \Omega \cup \{\tau\} \). Let \( \mathcal{M} \) be an \((M + 1) \times |\Omega|\) matrix over \( \Omega \), which is initially filled with \( \tau \), only.

The EES consists of alternations between two steps. One is a usual Metropolis step at a (random) temperature level \( \beta_i \). The other one is an equi-energy jump at the same temperature, if \( i \geq 1 \). At inverse temperature 0 there are only Metropolis moves. We store the resulting energies of the states we see at temperature \( \beta_i \) by entering them into the \( i \)th row of the matrix \( \mathcal{M} \), if the state has not been seen before. In this case it replaces one of the \( \iota \)'s (in a pre-described order). To explain the equi-energy step assume that the chain is at temperature \( \beta_i \), \( i \geq 1 \) and in state \( \sigma \). We determine the energy level \( k \), such that \( h_{k-1} < H(\sigma) \leq h_k \) and choose (with equal probabilities) a state \( \tau \) from all states \( \tau' \) with \( h_{k-1} < H(\tau') \leq h_k \), which we have already seen at temperature level \( \beta_{i-1} \). This new state is accepted with probability \( \min\{1, \frac{\pi_\beta(\tau')\pi_\beta_{h_{k-1}(\sigma)}}{\pi_\beta(\sigma)\pi_\beta_{h_k(\tau')}}\} \).

Otherwise, especially, if we have not seen any state in the same energy band in the \( i - 1 \)st row of \( \mathcal{M} \), the chain stays where it is. We denote the corresponding transition matrix (on \( \tilde{\Omega} \)) by \( Q \). Note that \( Q \) in general depends on time. We will not make this explicit, because we will just analyze the algorithm in the ‘best case scenario’, where the matrix \( \mathcal{M} \) does not contain
any $i$’s anymore. However, under this assumption, we will still be able to show that EES is slowly converging on the three state mean-field Potts model in a certain temperature regime. Formally

$$Q_i(\sigma, \tau) := Q_{n,i}(\sigma, \tau) := \frac{1}{|B_{n,k}|} \min\{1, \frac{\pi_{\beta_i}(\tau)\pi_{\beta_{i-1}}(\sigma)}{\pi_{\beta_i}(\sigma)\pi_{\beta_{i-1}}(\tau)}\} 1_{\tau \in B_{n,k}}.$$ (3.4)

Here $n$ is the time variable and $B_{n,k}$ is the set of states $\tau'$ with $h_{k-1} < H(\tau') \leq h_k$, which we have already seen at temperature level $\beta_{i-1}$ by time $n$.

One might expect, that we indeed use all states we have seen previously, rather than the ones we explored with the chain at temperature $\beta_{i-1}$. However, there is hardly any difference between the two chains, because if temperatures are very different the chains will typically also see states of very different energies. Our choice has the advantage that it is easy to see that the global chain to be described below is reversible and moreover, it agrees with the choice in the literature, see Kou et al (2006).

Based on this, we build a matrix that describes the movement of all particles simultaneously. This operator $R$ will be a matrix on $\Omega^{M+1}$, of course. We lift the movement of the $i$th particle to $\Omega^{M+1}$ by building

$$Q_i := \bigotimes_{j=0}^{i-1} I \otimes Q_i \bigotimes_{j=i+1}^{M} I$$

where $I$ is the identity matrix. Similarly, we consider the matrix $T_i$ that lifts the Metropolis step $T_{\beta_i}$ to $\Omega^{M+1}$, i.e. we consider

$$T_i := \bigotimes_{j=0}^{i-1} I \otimes T_{\beta_i} \bigotimes_{j=i+1}^{M} I.$$ 

Combining these operators the EES is defined by

$$R = \frac{1}{(M + 1)^2} \sum_{j,k,l=0}^{M} Q_j T_k Q_l.$$ 

Note that the versions of the EES given in Kou et al (2006) and Andrieu et al (2008) differ from each other and also our version is slightly different from those. However, the spirit of the algorithms is the same.

In the sequel, we will only consider a number of energy levels $M$ that depends linearly on $N$, such that $M = dN$. We will furthermore assume that $h_i$ are equi-distant. Indeed, this choice of $M$ is somewhat arbitrary, allowing for a polynomial dependence between $M$ and $N$ would not alter the algorithm much. However, choosing $M$, e.g. exponentially large in $N$, would lead to empty, or almost empty energy bands which would make the equi-energy step obsolete. Moreover, it would obviously lead to exponential relaxation times (in $N$), because exponentially many temperatures have to be simulated. On the other hand, having $M$ too small, e.g. constant, leads to almost non-interactive components (i.e. an equi-energy jump is almost never accepted) and EES stands no chance of increasing the speed of convergence compared to the standard Metropolis algorithm.

Of course, eventually we will only be interested in the $M + 1$st coordinate of this Markov chain. However, studying it entirely, seems easier. First of all, let us note that indeed, the distribution of the $M + 1$st coordinate converges to $\pi_{\beta_i}$. 


Theorem 3.1. The distribution of the $M+1$’st coordinate converges to $\pi_\beta$ as time tends to infinity.

Proof. This is the content of Andrieu et al. (2008) for their version of the EES. For our version the assertion follows from the ergodic theorem for Markov chains. Indeed, denote by $\mathcal{S}$ the Markov chain on $\Omega^{M+1} \times \Omega$, where $\Omega$ is the space of all $(M + 1) \times 3^N$ matrices. $\mathcal{S}$ will behave in its first component like $\mathcal{R}$ while in the second component we keep record of the filling of $\mathcal{M}$. Observe that each $T_i := T_{\beta_i}$ is reversible with respect to $\pi_{\beta_i}$ and $\mathcal{M}$ does not play any role for it. On the other hand, once we reach a situation where $\mathcal{M}$ is entirely filled with states different from $\iota$ (we denote this state of $\mathcal{M}$ by $M_0$ in the second coordinate of $\mathcal{S}$), i.e. we have seen all states at all temperatures, also all the equi-energy steps $Q_i$ are reversible with respect to $\pi_{\beta_i}$. This is, because $Q_i(\sigma, \tau) > 0$, if and only if, $\sigma$ and $\tau$ lie in the same energy band and follows from the construction of the transition probabilities. Thus, once $M_0$ is reached—which happens almost surely in finite time—$\mathcal{S}$ is reversible with respect to

$$\pi := \prod_{i=0}^{M} \pi_{\beta_i} \times \delta_{M_0}.$$  

Then the convergence follows from the convergence theorem for Markov chains. This, in particular, yields the assertion of the theorem. \qed

The proof is somewhat misleading, as it seems to indicate, that for exponentially large state spaces there is no hope that EES may converge in polynomial time, since first the state $M_0$ has to be reached. However, if we consider the high temperature situation $\beta < \beta_c$ in the Potts model the Metropolis–Hastings chain converges to its invariant distribution in polynomial time, even without any equi-energy steps.

On the other hand, we will see in the next section that in part of the low temperature regime $\beta > \beta_c$ the situation is even worse. Even, when we start $\mathcal{S}$ in the optimal state $M_0$ in its second component, i.e. when we assume the second component is already in $M_0$, the mixing time may be exponential.

4. Torpid mixing of EES on the low temperature mean-field Potts model

We now come to the central result of the note.

Theorem 4.1. EES is slowly mixing on the 3-state mean-field Potts model, when $\beta_c < \beta < 3$, even when the second component of the Markov chain $\mathcal{S}$ introduced in the proof of theorem 3.1 above is in state $M_0$.

We will prepare the proof of the theorem by explaining the ideas and stating some lemmas. In the proof of the theorem we will exploit one of the main differences between the mean-field Potts model and the Curie–Weiss model (i.e. when $E = \{1, 2\}$) at low temperatures. This difference lies in the fact, that in the Curie–Weiss model the state where both colors occur equally often is a local minimum of the Gibbs measure at all low temperatures, while it is a local maximum of the Gibbs measure in the mean-field Potts model for some temperatures in the low temperature regime (also see lemma 4.3). In particular, in the Curie–Weiss model, the Gibbs measure is flat in this state at the critical temperature while it exposes a local maximum in this state at the critical temperature in the Potts model. Thus, in the latter, EES will be very
reluctant to move far away from a state $\sigma$ with $m_N(\sigma) \approx (\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$. This is the core idea, even though the technical steps are somewhat more involved.

Let $c_1, c_2, c_3$ be numbers in $[0, 1]$ that add up to 1 and such that $c_i N$ is an integer for each $i = 1, 2, 3$. Then for $\sigma$, such that $m_N(\sigma_c) = (c_1, c_2, c_3) =: c$ we have that

$$\pi_\beta(m_N(\sigma) = c) = \frac{N}{\pi_{Nc_1, Nc_2, Nc_3}} \frac{\exp(\beta H_N(\sigma_c))}{Z_\beta}$$

$$= \frac{1}{N} \exp(-N(\sum_{i=1}^{N} c_i \log c_i + \Delta(c))) \frac{\exp(\beta H_N(\sigma_c))}{Z_\beta}$$

$$= \frac{\exp(N(f(c) + \Delta(c)))}{NZ_\beta}$$

where

$$f(c) = \frac{3}{2} \sum_{i=1}^{3} \left( \frac{\beta}{2} c_i^2 - c_i \log c_i \right)$$

and $\Delta(c)$ is $o(1)$. Note that we used Stirling’s formula to derive the second equality in (4.1) and the fact that we can rewrite $H_N(\sigma)$ as $H_N(\sigma) = \frac{1}{2} \sum_{i=1}^{3} c_i^2$, if $m_N(\sigma) = c$.

Letting $C := \{(c_1, c_2, c_3) \in [0, 1]^3 : \sum_{i=1}^{3} c_i = 1\}$ to be the domain of $f$ (and the set of all probabilities on the space $E$), Gore and Jerrum show:

**Lemma 4.2 (See (Gore and Jerrum 1999, proposition 1)).** Let $c$ be a local maximum of $f$. Then $c$ satisfies:

1. $c$ lies in the interior of $C$.
2. Either $a_i = \frac{1}{3}$ for all $i = 1, 2, 3$, or there are $0 < \alpha < \frac{1}{3} < \alpha' < 1$, such that $a_i \in \{\alpha, \alpha'\}$ for all $i = 1, 2, 3$. In the latter case there is exactly one $a_i$ equal to $\alpha'$, while all the other $a_j, j \neq i$ are equal to $\alpha$.

Analyzing the function $f$ around the point $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ we find that (in accordance with lemma 4.2) it might be a local but not a global maximum of $f$, if $\beta > \beta_c = 4 \log 2$ is not too large (a similar observation was already made in Kesten and Schonmann (1989)):

**Lemma 4.3.** If $4 \log 2 = \beta_c < \beta < 3$ then $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ is a local mode of $\pi_\beta$, if $N$ is large enough, i.e. $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ is a local maximum point of $\pi_\beta$.

**Proof.** In view of (4.1) it suffices to analyze $f$. For $x > 0$, and $a \in [0, 1]$ consider

$$h(x) := f(\frac{1}{3} + x, \frac{1}{3} - ax, \frac{1}{3} - (1 - a)x).$$

It is easy matter to check that $h'(0) = 0$ and $h''(0) = -(6 - 2\beta)(a^2 + a + 1)$. The assertion follows.

**Remark.** Lemma 4.3 is a main reason why theorem 4.1 is true. It is not difficult to check that the same behavior is true for general $q \geq 3$ in an appropriately chosen temperature regime (depending on $q$). Therefore, also theorem 4.1 could be proven for general $q \geq 3$. Indeed, the property shown in lemma 4.3 is intrinsically related to the first order phase transition of the mean-field Potts model. Such a phase transition can be characterized by the discontinuous transition of the accumulation point(s) of an order parameter of the model at the critical inverse temperature $\beta_c$. In the Potts model this order parameter is the variable $m_N$. However, in most natural models, these new accumulation point(s) are already local maxima of the dis-
tribution of the order parameter for some smaller values of $\beta$. Similarly, the old accumulation point(s) remain local maxima of the distribution of the order parameter for some larger values of $\beta$. This is exactly the statement of lemma 4.3.

Another key ingredient of the proof is a conductance argument (also known as Cheeger’s inequality in Diaconis and Stroock (1991))

**Theorem 4.4 (Sinclair and Jerrum (1989)).** Let $P$ be a Markov chain on a finite set $S$. Assume it is reversible with respect to $\pi$. For all $S' \subseteq S$, define

$$\Phi_{S'} = \sum_{x \in S', y \not\in S'} \frac{\pi(x)P(x,y)}{\pi(S')}$$

The conductance $\Phi$ given by

$$\Phi = \min_{S' : \pi(S') \leq \frac{1}{2}} \Phi_{S'}.$$

Then the following holds true for the spectral gap $\Gamma(P)$ of $P$:

$$\frac{\Phi^2}{2} \leq \Gamma(P) \leq 2\Phi.$$

As follows e.g. from Diaconis and Stroock (1991) a spectral gap that is the inverse of a polynomial in the problem instance results in fast mixing of the Markov chain. On the other hand, if the spectral gap is the inverse of an exponential in the problem instance, the Markov chain mixes slowly. An immediate consequence of theorem 4.4 is that the Metropolis algorithm alone is slowly mixing on the low temperature Potts model.

**Proposition 4.5.** The Metropolis algorithm mixes slowly on the Potts model, if $\beta > \beta_c$.

**Proof.** Take the macro-state $a_1 := a_1(\beta)$, i.e. the maximum point $a = (a_1, a_2, a_3)$ of $f$, where $a_1 > a_2 = a_3$. This point exists according to lemma 4.2 and because we are in the low temperature region. Since $a_1$ is a maximum of $f$, there is $\epsilon > 0$ such that $f$ is decreasing on the ball of radius $2\epsilon$ centered in $a_1$, $B_{2\epsilon}(a_1)$, when we walk from the center to the boundary on a straight line. $a_1$ is one of the three points in which the distribution of $m_N$ concentrates for large $N$ and that are equally likely. Thus for

$$\mathcal{B}_{2\epsilon} := \{ \sigma : m_N(\sigma) \in B_{2\epsilon}(a_1) \}$$

we obtain that

$$\frac{1}{4} \leq \pi_\beta(m_N(\sigma) \in \mathcal{B}_{2\epsilon}) \leq \frac{1}{3},$$

when $N$ is large enough and $\epsilon > 0$ is fixed and small enough. Moreover, due to the exponential structure of $\pi_\beta$, i.e.

$$\pi_\beta(m_N(\sigma) = c) = \frac{\exp(N(f(c) + \Delta(c)))}{NZ_\beta},$$

and the behavior of $f$ on $B_{2\epsilon}(a_1)$ (on $B_{2\epsilon}(a_1)$, the function $f$ decreases like a multiple of the square of the two norms) we obtain that
\[
\pi_\beta(m_N(\sigma) \in B_{2\varepsilon} \setminus B_{\varepsilon}) \leq e^{-c'N}
\]
for a suitably chosen constant \(c' > 0\). But this implies that the set \(B_{2\varepsilon}\) constitutes a ‘bad cut’. Indeed with the notation of the previous theorem we see that

\[
\Phi \leq \sum_{x \in B_{2\varepsilon}, y \not\in B_{\varepsilon}} \pi_\beta(m_N(\sigma) \in B_{2\varepsilon} \setminus B_{\varepsilon}) \pi_\beta(m_N(\sigma) \in B_{2\varepsilon}) \leq e^{-c'N}.
\]

Thus \(T_\beta\) mixes slowly, when \(\beta > \beta_c\).

As a consequence, if EES is fast on the low temperature Potts model, this will have to be caused by the equi-energy steps. However, the following important observation is that we will not be able to switch between two states that are at very different distances from the center mode \(a_0 := (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})\) by an equi-energy step. More precisely:

**Lemma 4.6.** For each \(\varepsilon > 0\) and each \(\varepsilon > \delta > 0\) there is a number of spins \(N_0\) such that for all \(N > N_0\) and whenever \(\sigma\) and \(\tau\) satisfy \(||m_N(\sigma) - a_0||_1 < \delta\) and \(||m_N(\tau) - a_0||_1 > \varepsilon\) (where \(|| \cdot ||_1\) denotes the 1-norm on \(\mathbb{C}\)) then

\[
Q_M(\sigma, \tau) = 0.
\]

Here \(Q_M\) is defined in (3.4).

**Proof.** The proof mainly shows that under the given conditions the energies of \(\sigma\) and \(\tau\) are too far apart from each other. Indeed, observe that \(Q_M(\sigma, \tau) > 0\) requires \(\sigma\) and \(\tau\) to be in the same energy band. Thus there is \(i \in \{0, \ldots, M - 1\}\) such that \(h_i < H_N(\sigma) = H_N(\tau) \leq h_{i+1}\). Now each \(\sigma_{a_0}\) with \(m_N(\sigma_{a_0}) = a_0\) minimizes the energy to \(H_N(\sigma_{a_0}) = \frac{N}{2} \times 3 \times \frac{1}{3} = \frac{N}{2}\). On the other hand, the states where all spins point into the same direction have maximal energy \(\frac{N}{2}\) see (3.3).

Thus, recalling that \(M = dN\), the width of the energy bands is

\[
h_{i+1} - h_i = \frac{1}{M}(h_M - h_0) = \frac{\frac{N}{2} - \frac{N}{6}}{M} = \frac{1}{3d}.
\]

Therefore, \(\sigma\) and \(\tau\) are only in the same energy band, if

\[
|H_N(\sigma) - H_N(\tau)| = \frac{N}{2} \left(||m_N(\sigma)||_2^2 - ||m_N(\tau)||_2^2\right) \leq \frac{1}{3d},
\]

i.e. if the two norms \(||m_N(\sigma)||_2^2\) and \(||m_N(\tau)||_2^2\) satisfy

\[
||m_N(\sigma)||_2^2 - ||m_N(\tau)||_2^2 \leq \frac{1}{6dN}.
\]

Since 1-norm and 2-norm are equivalent on \(\mathbb{C}\) this proves the assertion.
We will again use a conductance argument to prove theorem 4.1. In order to prepare it we lift the balls \( B_{\epsilon} \) to \( \Omega^{M+1} \): For \( \epsilon > 0 \) let
\[
\mathcal{B}_{\epsilon}^{M+1} := \{ x \in \Omega^{M+1} : m_N(x_M) \in B_{\epsilon}(a_0) \}.
\]
From now on we will assume that \( \beta_c < \beta < 3 \). Recall that then still \( a_0 \) is a local (but not a global) maximum of the function \( f \). Let us fix \( \epsilon > 0 \) so small, that still \( f \) is decreasing on \( B_{\epsilon}(a_0) \) when we move away from the center (in particular, \( a_0 \) is the only mode of \( \pi_\beta \) on \( B_{\epsilon}(a_0) \)). Moreover, let us fix \( \delta > \epsilon \) and \( N_0 \) so large that even with two equi-energy steps and a Metropolis step in between, a \( \sigma \) with \( ||m_N(\sigma) - a_0|| \geq \epsilon \) cannot be reached from a \( \tau \) with \( ||m_N(\tau) - a_0|| < \delta \).

This can be constructed as in lemma 4.6. Indeed, we will need the following: For \( \delta > 0 \) given with \( \delta < \delta' < \epsilon \) there is \( N_1 \), such that if \( N \geq N_1 \) an equi-energy jump started in \( m_N \in B_{\delta'}(a_0) \) will not leave \( B_{\delta'}(a_0) \). The subsequent Metropolis step can only increase the 1-distance of \( m_N \) to \( a_0 \) by at most \( 1/N \), hence \( m_N \) is still in, say, \( B_{\delta''}(a_0) \), for some \( \delta' < \delta'' < \epsilon \).

Finally, there is \( N_2 \), such that if \( N \geq N_2 \) an equi-energy jump started in \( m_N \in B_{\delta''}(a_0) \) will not leave \( B_{\delta''}(a_0) \). We will from now on always take \( N \geq N_0 := \max\{N_1, N_2\} \).

All this is necessary because the chains \( \mathcal{R} \) and \( \mathcal{S} \) possibly comprise two such jumps. Next we prove

**Lemma 4.7.** Let \( \beta_c < \beta < 3 \) and \( \epsilon > \delta > 0 \) and \( N_0 \) be chosen as above. Then, there exists \( c'' > 0 \) such that for \( \hat{\pi}(x) := \prod_{i=0}^M \pi_\beta(x_i), x \in \Omega^{M+1}, \) we have
\[
\frac{\pi(\mathcal{B}_{\epsilon}^{M+1} \setminus \mathcal{B}_{\delta}^{M+1})}{\pi(\mathcal{B}_{\epsilon}^{M+1})} \leq e^{-c''N}.
\]

**Proof.** According to our above analysis \( a_0 \) is a local (but not a global) maximum point of the distribution of \( m_N \) under \( \pi_\beta = \pi_\beta := \pi_M \), if \( \beta_c < \beta < 3 \). Therefore
\[
\frac{\pi_M(\{ \sigma : m_N(\sigma) \in \mathcal{B}_{\epsilon}(a_0) \} \setminus \mathcal{B}_{\delta}(a_0) \})}{\pi(\{ \sigma : m_N(\sigma) \in \mathcal{B}_{\epsilon}(a_1) \})} \leq e^{-c''N}
\]
for \( c'' > 0 \) chosen appropriately. The proof of this statement follows the concepts of the proof of proposition 4.5. This fact is easily transferred to the measure \( \hat{\pi} \) due to its product structure.

With the help of this lemma we will be able to establish that the set \( \mathcal{B}_{\epsilon}^{M+1} \) constitutes a ‘bad cut’ for the Markov chain \( \mathcal{S} \).

**Proposition 4.8.** Consider the Markov chain \( \mathcal{S} \) on the state space \( \Omega_{\mathbb{M}} := \Omega^{M+1} \times \mathbb{M} \) where again \( \mathbb{M} \) is the space of all \( (M + 1) \times 3^N \) matrices. Here the first coordinate keeps record of the current state of \( M + 1 \) chains, while the second coordinate tracks the filling of the matrix \( \mathcal{M} \).

If \( \beta_c < \beta < 3 \) and the second coordinate of \( \mathcal{S} \) is equal to \( M_0 \) its conductance \( \Phi(\mathcal{S}) \) satisfies
\[
\Phi(\mathcal{S}) \leq e^{-c''N}
\]
for some \( c'' > 0 \), if \( N \) is large enough.

**Proof.** Since \( \beta_c < \beta \) clearly \( \pi_M(\{ \sigma : m_N(x_M) \in B_{\epsilon}(a_0) \}) < \frac{1}{2} \), when \( N \) is large enough. Thus also
\[ \pi(\Omega^M \times \{ \sigma : m_N(x_M) \in B_{\delta}(a_0) \} \times M_0) = \pi(\mathfrak{B}_{\varepsilon}^{M+1} \times M_0) < \frac{1}{2} \]

for \( N \) large enough. Thus

\[ \Phi(S) \leq \Phi(\mathfrak{B}_{\varepsilon}^{M+1} \times M_0)(S) \]

\[ \leq \sum_{\tilde{\sigma} \in (\mathfrak{B}_{\varepsilon}^{M+1} \times M_0) \setminus (\mathfrak{B}_{\delta}^{M+1} \times M_0)} \pi(\tilde{\sigma}) \]

\[ \leq \frac{\pi(\mathfrak{B}_{\varepsilon}^{M+1} \times M_0)}{\pi(\mathfrak{B}_{\delta}^{M+1} \times M_0)} \]

\[ \leq \frac{\pi((\mathfrak{B}_{\varepsilon}^{M+1} \times M_0) \setminus (\mathfrak{B}_{\delta}^{M+1} \times M_0))}{\pi(\mathfrak{B}_{\varepsilon}^{M+1} \times M_0)} \]

\[ = \frac{\tilde{\pi}(\mathfrak{B}_{\varepsilon}^{M+1} \setminus \mathfrak{B}_{\delta}^{M+1})}{\tilde{\pi}(\mathfrak{B}_{\varepsilon}^{M+1})} \leq e^{-c'N}. \]

Here we used, of course, our previous estimates together with our construction of \( \delta \). Starting in \( B_{\delta}(a_0) \) the combination of an equi-energy jump, a Metropolis move, and another equi-energy jump will not leave \( B_{\varepsilon}(a_0) \) according to lemma 4.6 and the construction of \( \delta \) and \( N_0 \). 

Now we have prepared everything to prove theorem 4.1.

**Proof of theorem 4.1.** Just note, that conditioned on the event that the second coordinate of \( S \) is in \( M_0 \) (which it cannot leave anymore), \( S \) is reversible with respect to \( \pi \). Hence we can apply theorem 4.4 together with the conductance estimate of proposition 4.8 to obtain the desired result.

**Remark.** Note that a similar proof would not work in the Curie–Weiss model, because there the ‘center point’ \((1/2, 1/2)\), i.e. the \( \sigma \)-s where both directions for the spins occur equally often, is always a local minimum of the Gibbs measure at low temperatures.

Moreover, note that we could adapt the proof to different values of \( q \geq 3 \) as mentioned above.

Finally, a similar argument should work for ‘more disordered’ models, as Potts models on sufficiently dense Erdős–Rényi graphs, as e.g. analyzed in Kabluchko et al (2019) for \( q = 2 \).

**Remark.** We have just seen that EES mixes slowly on the 3-state Potts model at \( \beta_c < \beta < 3 \), even when we know the energies of the entire set of states. We also argued that at high temperatures these temperature steps are not necessary, because already the Metropolis chain itself converges rapidly. However, one may doubt that there are reasonable models, at all, in which EES converges rapidly while the Metropolis algorithm does not. The point is, that, if we have not filled \( \mathcal{M} \) almost entirely, a temperature jump may provide the desired tunneling effect, but to a rather unfavorable point of the target distribution.
Acknowledgments

Research of the second author was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy EXC 2044 390685587, Mathematics Münster: Dynamics Geometry-Structure.

ORCID IDs

Matthias Löwe ⚖ https://orcid.org/0000-0001-9893-3770

References

Andrieu C, Jasra A, Doucet A and Del Moral P 2008 A note on convergence of the equi-energy sampler Stoch. Anal. Appl. 26 298–312
Baragatti M, Grimaud A and Pommeret D 2013 Parallel tempering with equi-energy moves Stat. Comput. 23 323–39
Berg B A 2000 Introduction to multicanonical Monte Carlo simulations Monte Carlo methods (Toronto, ON, 1998) (Fields Institute Communications vol 26) (Providence, RI: American Mathematical Society) pp 1–24
Berg B A and Neuhaus T 1992 Multicanonical ensemble: a new approach to simulate first-order phase transitions Phys. Rev. Lett. 68 9–12
Bhatnagar N and Randall D 2004 Torpid mixing of simulated tempering on the Potts model Proc. 15th Annual ACM-SIAM Symp. on Discrete Algorithms (New York: ACM) pp 478–87 (electronic)
Bhatnagar N and Randall D 2016 Simulated tempering and swapping on mean-field models J. Stat. Phys. 164 495–530
Cuff P, Ding J, Louidor O, Lubetzky E, Peres Y and Sly A 2012 Glauber dynamics for the mean-field Potts model J. Stat. Phys. 149 432–77
Diaconis P and Stroock D 1991 Geometric bounds for eigenvalues of Markov chains Ann. Appl. Probab. 1 36–61
Doll J, Dupuis P and Nyquist P 2018 A large deviations analysis of certain qualitative properties of parallel tempering and infinite swapping algorithms Appl. Math. Optim. 78 103–44
Ebbers M, Knöpfel H, Löwe M and Vermel F 2014 Mixing times for the swapping algorithm on the Blume–Emery–Griffiths model Random Struct. Algorithms 45 38–77
Ebbers M and Löwe M 2009 Torpid mixing of the swapping chain on some simple spin glass models Markov Process. Relat. Fields 15 59–80
Ellis R S and Wang K 1990 Limit theorems for the empirical vector of the Curie–Weiss–Potts model Stoch. Process. Appl. 35 59–79
Geyer C J 1991 Markov chain Monte Carlo maximum likelihood Computing Science and Statistics: Proc. 23rd Symp. on the Interface Interface Foundation pp 156–63 (Fairfax Station)
Geyer C J and Thompson E A 1995 Annealing Markov chain Monte Carlo with applications to ancestral inference J. Am. Stat. Assoc. 90 909–20
Gore V K and Jerrum M R 1999 The Swendsen–Wang process does not always mix rapidly J. Stat. Phys. 97 67–86
Häggström O 2002 Finite Markov Chains and Algorithmic Applications (London Mathematical Society Student Texts vol 52) (Cambridge: Cambridge University Press)
Hua X and Kou S C 2011 Convergence of the equi-energy sampler and its application to the Ising model Stat. Sin. 21 1687–711
Hukushima K and Nemoto K 1996 Exchange Monte Carlo method and application to spin glass simulations J. Phys. Soc. Japan 65 1604–8
Kabluchko Z, Löwe M and Schubert K 2019 Fluctuations of the magnetization for Ising models on dense Erdős–Rényi random graphs J. Stat. Phys. 177 78–94
Kesten H and Schonmann R H 1989 Behavior in large dimensions of the Potts and Heisenberg models Rev. Math. Phys. 1 147–82
Kou S C, Zhou Q and Wong W H 2006 Equi-energy sampler with applications in statistical inference and statistical mechanics Ann. Stat. 34 1581–652 (With discussions and a rejoinder by the authors)

Löwe M and Meise C 2001 Note on the knapsack Markov chain Stoch. Process. Appl. 94 155–70

Löwe M and Vermet F 2009 The swapping algorithm for the Hopfield model with two patterns Stoch. Process. Appl. 119 3471–93

Madras N 1998 Umbrella sampling and simulated tempering Numerical Methods for Polymeric Systems (IMA Volume in Mathematics and its Applications vol 102) ed S G Whittington (New York: Springer) pp 19–32

Madras N and Piccioni M 1999 Importance sampling for families of distributions Ann. Appl. Probab. 9 1202–25

Madras N and Zheng Z 2003 On the swapping algorithm Random Struct. Algorithms 22 66–97

Marinari E and Parisi G 1992 Simulated tempering: a new Monte Carlo scheme Europhys Lett. 19 451–8

Mossel E and Sly A 2013 Exact thresholds for Ising–Gibbs samplers on general graphs Ann. Probab. 41 294–328

Orlandini E 1998 Monte Carlo study of polymer systems by multiple Markov chain method Numerical Methods for Polymeric Systems (IMA Volume in Mathematics and Its Applications vol 102) ed S G Whittington (New York: Springer) pp 33–57

Sinclair A and Jerrum M 1989 Approximate counting, uniform generation and rapidly mixing Markov chains Inform. Comput. 82 93–133

Torrie G and Valleau J 1977 Nonphysical sampling distributions in Monte Carlo free-energy estimation: umbrella sampling J. Comput. Phys. 23 187–99