Some universal estimates on convergence to equilibrium in reversible Markov chains[*]

Mykhaylo Shkolnikov[†]

Abstract

We obtain universal estimates on the convergence to equilibrium and the times of coupling for continuous time irreducible reversible finite-state Markov chains, both in the total variation and in the L^2 norms. The estimates in total variation norm are obtained using a novel identity relating the convergence to equilibrium of a reversible Markov chain to the increase in the entropy of its one-dimensional distributions. In addition, for chains reversible with respect to the uniform measure, we show how the global convergence to equilibrium can be controlled using the entropy accumulated by the chain.

1 Introduction

Recently, the convergence to equilibrium of slowly mixing Markov chains appearing in statistical physics has attracted much attention. In this framework continuous time irreducible reversible Markov chains are defined by choosing the transition rates from a state (usually, a spin configuration) a to a state (spin configuration) b to be proportional to $e^{-\beta(E(b)-E(a))_+}$, where E is an energy functional and β stands for the inverse temperature, which in this context is chosen to be large: $\beta \gg 1$. In this low temperature regime, a recurring feature is that the energy landscape given by E divides the state space into sets of metastable (or, stable) states, which are separated by potential wells. The convergence to equilibrium of the corresponding Markov chain, started in a metastable state, is then governed by the time it takes to overcome the respective potential wells in order to reach the part of the state space with the lowest energy.

The potential theoretic approach to metastability developed in the articles [9], [10] and [11] (see also the excellent summaries [7] and [8]) has been used to obtain precise information on metastable transitions for reversible Markov chains associated with

---

[†]University of California, Berkeley, USA. E-mail: mshkolni@gmail.com

several models of statistical physics. These include certain disordered mean field models (see [10]) and, more specifically, the Curie-Weiss model with a random field taking finitely many values (see [10], [3] and [4]). Other examples of slowly mixing reversible Markov chains, in which the metastable behavior has been analyzed in detail, include the Glauber dynamics for the two-dimensional Ising model on a torus and its generalizations (see [20] and [21]), the three-dimensional Ising model on a torus (see [2]), the Ising model on rectangular subsets of $\mathbb{Z}^d$ (see [23]), the classical Curie-Weiss model (see [13] and [19]) and the Potts model on rectangular subsets of $\mathbb{Z}^d$ (see [6]). Moreover, the first exit problem from a domain for reversible chains with exponentially small transition probabilities was studied in the article [22].

Here, we take a different viewpoint. Instead of analyzing a specific Markov chain in detail, we try to understand some universal aspects of the ultrametric structure, that is, the presence of multiple time scales in a general irreducible reversible finite-state Markov chain. We obtain universal estimates on the convergence to equilibrium and the times of coupling in this abstract framework. We prove such results both in the total variation norm and in the L^2 norm. In the case of the total variation norm, we utilize a novel entropy identity relating the convergence to equilibrium of the chain to the increase of the entropy of its one-dimensional distributions. Finally, in the case that the chain is reversible with respect to the uniform measure, we show how the entropy of the one-dimensional distributions can be utilized to control the global convergence to equilibrium of the chain.

To give examples of the type of results we obtain, we introduce a set of notations. Let X be a continuous time irreducible reversible Markov chain on a set $I = \{1, 2, \ldots, n\}$ of n elements. Moreover, for an $a \in I$ and a $t \geq 0$ let P_t^a be the one-dimensional distribution of the chain at time t, when started in a. Finally, write $\|.\|_{TV}$ for the total variation norm, ν for the invariant distribution of X and let $H(\nu)$ be the entropy $-\sum_{a \in I} \nu(a) \log \nu(a)$ of the invariant distribution ν.
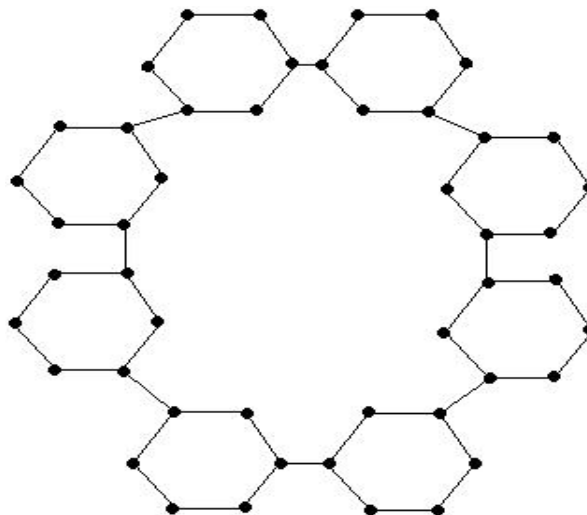

Figure 1: 6-cycles arranged in an 8-cycle

Before stating our first result rigorously, we would like to provide the reader with some intuition by giving an example. Fix natural numbers $M, N \geq 3$ and consider the

graph given by an arrangement of M N-cycles in a cycle of size M (see Figure 1). Now, let X be the continuous time Markov chain on this graph, which has a transition rate $\rho_1 > 0$ for neighboring vertices belonging to the same N-cycle and a transition rate $\rho_2 > 0$ for neighboring vertices belonging to different N-cycles. Since the generating matrix of this Markov chain is symmetric, it is reversible with respect to the uniform distribution on the set of vertices of the graph. Next, suppose that ρ_2 is much smaller than ρ_1. Then, it is intuitively clear that, for every vertex a of the described graph, the quantity $\|P_{3t}^a - P_t^a\|_{TV}$ can be only large on the two disjoint time intervals, during which the Markov chain mixes on the N-cycle containing a and on the M-cycle comprised by the M N-cycles, respectively. Under the scale-invariant measure, which has the density $\frac{1}{t}$ on the time axis $[0, \infty)$, the union of these two time intervals has a total measure of order $\log N + \log M = \log(MN)$. Thus, it is logarithmic in the size of the state space of X. The purpose of Theorems 1 and 3 below is to show that the latter property is universal for continuous time irreducible reversible Markov chains, in the sense that the order of magnitude in this example is an upper bound on the size of the corresponding quantity for a general reversible Markov chain. Due to the universal (worst-case) nature of the results, we do not expect the bounds to be sharp in the examples from statistical physics described at the beginning of the introduction.

Theorem 1. *Let X be a continuous time irreducible reversible Markov chain on the set $I = \{1, 2, \ldots, n\}$ and let ν be its invariant distribution. Then, the following is true.*

(a) *For every $\delta > 0$, there is a constant $C(\delta) > 0$ depending only on δ (and not on n or the particular Markov chain) such that*

$$\sum_{a \in I} \nu(a) \int_0^\infty \mathbf{1}_{\{t \geq 0: \, \|P_{3t}^a - P_t^a\|_{TV} \geq \delta\}} \frac{1}{t} \, \mathrm{d}t \leq C(\delta) \, H(\nu). \tag{1.1}$$

In particular, for every $\delta, \epsilon > 0$, there exists a constant $C_\epsilon(\delta) > 0$ depending only on δ and ϵ (and not on n or the particular Markov chain) such that

$$\nu\left(a \in I : \int_0^\infty \mathbf{1}_{\{t \geq 0: \, \|P_{3t}^a - P_t^a\|_{TV} \geq \delta\}} \frac{1}{t} \, \mathrm{d}t \geq C_\epsilon(\delta) \, H(\nu)\right) < \epsilon. \tag{1.2}$$

(b) *For every $\delta > 0$, there is a constant $\tilde{C}(\delta) > 0$ depending only on δ (and not on n or the particular Markov chain) such that*

$$\sum_{(a,b) \in I^2} \nu(a) \, \nu(b) \int_0^\infty \mathbf{1}_{\{t \geq 0: \, \|P_t^a - P_t^b\|_{TV} - \|P_{3t}^a - P_{3t}^b\|_{TV} \geq \delta\}} \frac{1}{t} \, \mathrm{d}t \leq \tilde{C}(\delta) \, H(\nu). \tag{1.3}$$

In particular, for every $\delta, \epsilon > 0$, there exists a constant $\tilde{C}_\epsilon(\delta) > 0$ depending only on δ and ϵ (and not on n or the particular Markov chain) such that

$$(\nu \times \nu)\left((a,b) \in I^2 : \int_0^\infty \mathbf{1}_{\{t \geq 0: \|P_t^a - P_t^b\|_{TV} - \|P_{3t}^a - P_{3t}^b\|_{TV} \geq \delta\}} \frac{1}{t} \mathrm{d}t \geq \tilde{C}_\epsilon(\delta) \, H(\nu)\right) < \epsilon. \tag{1.4}$$

We remark at this point that universal estimates as in Theorem 1 can only be obtained under the *scale-invariant measure* $\frac{1}{t} \, \mathrm{d}t$ on the time axis $[0, \infty)$, which has the property

$$\int_{t_1}^{t_2} \frac{1}{t} \, \mathrm{d}t = \int_{\eta t_1}^{\eta t_2} \frac{1}{t} \, \mathrm{d}t \tag{1.5}$$

for all $\eta > 0$ and $0 < t_1 < t_2 < \infty$. This can be easily seen by slowing down or speeding up the chain by a constant factor.

To give an example of a result in the framework of L^2 convergence, a set of auxiliary notations is needed. For simplicity, we assume for the moment that X is irreducible and reversible with respect to the uniform distribution on I. In this case, writing $\mathcal{L}$ for the generating matrix of X, we can conclude that the matrix $-\mathcal{L}$ is symmetric and admits an orthonormal basis of eigenvectors $v_1, v_2, \ldots, v_n$ corresponding to eigenvalues $0 = \lambda_1 < \lambda_2 \leq \lambda_3 \leq \ldots \leq \lambda_n$. Fixing a pair of initial states (a, b) and letting e_a (resp. e_b) be the vector, whose only non-zero component is the a-th one (resp. the b-th one) and equals to 1, we have the decomposition

$$e_a - e_b = \sum_{l=2}^{n} \mu_l \, v_l. \tag{1.6}$$

Finally, we define the set

$$A(a,b) := \{0, \mu_2^2, \mu_2^2 + \mu_3^2, \ldots, \mu_2^2 + \mu_3^2 + \ldots + \mu_n^2\} \subset [0, 2] \tag{1.7}$$

and a family of its neighborhoods

$$A^\delta(a,b) := [0, \delta\mu_2^2] \cup [(1-\delta)\mu_2^2, \mu_2^2 + \delta\mu_3^2] \cup \ldots \cup [2 - \delta\mu_n^2, 2] \tag{1.8}$$

for $\delta \in \left(0, \frac{1}{2}\right)$, and write $\|.\|_2$ for the L^2 norm with respect to the counting measure on I.

Theorem 2. *In the setting just described the following is true. For every $\delta \in \left(0, \frac{1}{2}\right)$, there is a constant $K(\delta) > 0$ such that*

$$\int_0^\infty \mathbf{1}_{\left\{t \geq 0: \, \|P_t^a - P_t^b\|_2^2 \notin A^\delta(a,b)\right\}} \frac{1}{t} \, dt \leq K(\delta) \, n \tag{1.9}$$

for all pairs of initial states a, b. The constant $K(\delta)$ depends only on δ, but not on a, b, n or the particular Markov chain X.

The rest of the paper is organized as follows. In Section 2.1, we prove a stronger version of Theorem 1 in the case that the invariant distribution ν is the uniform distribution on I. In order to do this, we show a novel entropy identity (see Lemma 2.1) allowing us to relate the increase in the entropy of the one-dimensional distributions of the Markov chain to the convergence of the chain to its equilibrium. In Section 2.2, we prove Theorem 1 by suitably adapting the entropy identity of Lemma 2.1 to the general setting. In Section 3.1, we give a global (or, averaged) version of Theorem 2 and present the proof of Theorem 2. Subsequently, we explain in Section 3.2, how Theorem 2 extends to general continuous time irreducible reversible finite-state Markov chains. Finally, in Section 4, we show in the case that the chain is reversible with respect to the uniform distribution, how the entropy of the one-dimensional distributions of the chain can be used to obtain a control on the global convergence of the chain to its equilibrium.

2 Estimates in total variation norm

In this section we give a control on the convergence to equilibrium and the times of coupling with respect to the total variation norm by analyzing the change in the entropy of the Markov chain over time.

2.1 Markov chains reversible with respect to the uniform distribution

The following theorem is a stronger version of Theorem 1 for the special case of Markov chains, which are reversible with respect to the uniform distribution.

Theorem 3. *Consider the setting of Theorem* 1 *and assume, in addition, that the Markov chain X is reversible with respect to the uniform distribution on $I = \{1, 2, \dots, n\}$. Then:*

(a) There is a constant $C(\delta) > 0$ depending only on δ (and not on n or the particular Markov chain) such that for all initial states a of the Markov chain:

$$\int_0^\infty \mathbf{1}_{\{t \geq 0: \|P_{3t}^a - P_t^a\|_{TV} \geq \delta\}} \, \frac{1}{t} \, \mathrm{d}t \leq C(\delta) \log n. \tag{2.1}$$

(b) There is a constant $\tilde{C}(\delta) > 0$ depending only on δ (and not on n or the particular Markov chain) such that for all pairs (a, b) of initial states of the Markov chain:

$$\int_0^\infty \mathbf{1}_{\{t \geq 0: \|P_t^a - P_t^b\|_{TV} - \|P_{3t}^a - P_{3t}^b\|_{TV} \geq \delta\}} \, \frac{1}{t} \, \mathrm{d}t \leq \tilde{C}(\delta) \log n. \tag{2.2}$$

The proof relies on the following entropy identity.

Lemma 2.1. *Let $X(t)$, $t \geq 0$ be a Markov chain as in Theorem* 3 *started in an initial state $a \in I$. Then, for all $t \geq 0$:*

$$H(P_{2t}^a) - H(P_t^a) = H(P_{t,2t}^a | P_{3t,2t}^a). \tag{2.3}$$

Hereby, $H(.|.)$ stands for the relative entropy and $P_{u,s}^a$ stands for the law of the random vector $(X(u), X(s))$. In particular, by Pinsker's inequality (see e.g. Lemma 11.6.1 in [14]), the inequality

$$\|P_t^a - P_{3t}^a\|_{TV} \leq \sqrt{2(H(P_{2t}^a) - H(P_t^a))} \tag{2.4}$$

holds for all $t \geq 0$ and all initial states $a \in I$.

Proof of Lemma 2.1. We start the proof with the following elementary computation, which only relies on the Markov property of X:

$$
\begin{aligned}
H(P_{2t}^a) - H(P_t^a) &= -\sum_{i \in I} P_{2t}^a(i) \log P_{2t}^a(i) + \sum_{i \in I} P_t^a(i) \log P_t^a(i) \\
&= -\sum_{j \in I} P_t^a(j) \Big\{ \Big(\sum_{i \in I} P_t^j(i) \log P_{2t}^a(i) \Big) - \log P_t^a(j) \Big\} \\
&= -\sum_{j \in I} P_t^a(j) \sum_{i \in I} P_t^j(i) \Big\{ \log P_{2t}^a(i) - \log P_t^a(j) \Big\} \\
&= -\sum_{(i,j) \in I^2} P_t^a(j) \, P_t^j(i) \log \frac{P_{2t}^a(i) \, P_t^j(i)}{P_t^a(j) \, P_t^j(i)}.
\end{aligned}
$$

We now exploit the symmetry of the transition matrices of the Markov chain X (which is due to the reversibility of the uniform distribution and the detailed balance condition) to deduce

$$P_{2t}^a(i) \, P_t^j(i) = P_{2t}^a(i) \, P_t^i(j) \tag{2.5}$$

for all $(i, j) \in I^2$. In addition, the Markov property of X yields

$$P_{2t}^a(i) \, P_t^i(j) = \mathbb{P}^a(X(2t) = i, X(3t) = j), \tag{2.6}$$

$$P_t^a(j) \, P_t^j(i) = \mathbb{P}^a(X(t) = j, X(2t) = i) \tag{2.7}$$

for all $(i, j) \in I^2$. Putting the latter three observations together, we end up with the lemma. $\qquad\square$

In the proof of Theorem 3 we will need the following simple calculus lemma.

Lemma 2.2. *Let $g : \mathbb{R} \to [0, \infty)$ be a non-decreasing function, which satisfies*

$$\lim_{u \to -\infty} g(u) = p, \quad \lim_{u \to \infty} g(u) = q \tag{2.8}$$

for some non-negative real constants $p \leq q$. Then, for every $r > 0$ and $\epsilon > 0$, one has the inequality

$$\int_{-\infty}^{\infty} \mathbf{1}_{\{g(u+r)-g(u) \geq \epsilon\}} \, \mathrm{d}u \leq \frac{r(q-p)}{\epsilon} \leq \frac{rq}{\epsilon}. \tag{2.9}$$

Proof of Lemma 2.2. It suffices to observe the elementary inequality

$$\mathbf{1}_{\{g(u+r)-g(u) \geq \epsilon\}} \leq \frac{g(u+r) - g(u)}{\epsilon}, \tag{2.10}$$

which leads to

$$\int_{-\infty}^{\infty} \mathbf{1}_{\{g(u+r)-g(u) \geq \epsilon\}} \, \mathrm{d}u \leq \frac{1}{\epsilon} \lim_{K \to \infty} \left(\int_{-K+r}^{K+r} g(u) \, \mathrm{d}u - \int_{-K}^{K} g(u) \, \mathrm{d}u \right) = \frac{r(q-p)}{\epsilon}$$

and, hence, yields the lemma. $\qquad \square$

We are now ready for the proof of Theorem 3.

Proof of Theorem 3. First, we note that part (b) of the theorem is a consequence of part (a) due to the inequalities

$$\|P_t^a - P_t^b\|_{TV} - \|P_{3t}^a - P_{3t}^b\|_{TV} \leq \|P_t^a - P_{3t}^a\|_{TV} + \|P_t^b - P_{3t}^b\|_{TV} \tag{2.11}$$

and

$$\mathbf{1}_{\{\|P_t^a - P_{3t}^a\|_{TV} + \|P_t^b - P_{3t}^b\|_{TV} \geq \delta\}} \leq \mathbf{1}_{\{\|P_t^a - P_{3t}^a\|_{TV} \geq \delta/2\}} + \mathbf{1}_{\{\|P_t^a - P_{3t}^a\|_{TV} \geq \delta/2\}}. \tag{2.12}$$

We turn now to the proof of part (a). Due to the inequality (2.4), it suffices to prove that for every $\overline{\delta} > 0$ there is a constant $\overline{C}(\overline{\delta}) > 0$ depending only on $\overline{\delta}$ (and not on a, n or the Markov chain X) such that

$$\int_0^{\infty} \mathbf{1}_{\{t \geq 0: \, H(P_{2t}^a) - H(P_t^a) \geq \overline{\delta}\}} \frac{1}{t} \, \mathrm{d}t \leq \overline{C}(\overline{\delta}) \log n. \tag{2.13}$$

Introducing the function $g : \mathbb{R} \to [0, \infty)$, $g(u) = H(P_{e^u}^a)$, we can rewrite the latter inequality as

$$\int_{-\infty}^{\infty} \mathbf{1}_{\{u: \, g(u+\log 2) - g(u) \geq \overline{\delta}\}} \, \mathrm{d}u \leq \overline{C}(\overline{\delta}) \log n. \tag{2.14}$$

Noting that $\lim_{u \to \infty} g(u) = \lim_{t \to \infty} H(P_t^a) = \log n$ (since the uniform distribution is the unique stationary distribution of X), we see that the desired inequality holds with $\overline{C}(\overline{\delta}) = \frac{\log 2}{\overline{\delta}}$ as a consequence of Lemma 2.2. $\qquad \square$

2.2 General reversible Markov chains

In this subsection we consider a general continuous time irreducible reversible Markov chain X on I and will prove Theorem 1. To start with, we recall the detailed balance condition:

$$\nu(a) \, \mathbb{P}^a(X(t) = b) = \nu(b) \, \mathbb{P}^b(X(t) = a), \tag{2.15}$$

which holds for all times $t \geq 0$ and all pairs of states $(a, b) \in I^2$. We now give the proof of Theorem 1.

Proof of Theorem 1. The first assertion in part (b) of the theorem is a direct consequence of the inequality (2.12) (which clearly remains true in the more general setting of the present theorem) and the first assertion in part (a) of the theorem. Moreover, the second assertions in both parts of the theorem follow from the first assertions in the corresponding parts of the theorem and Markov's inequality. For these reasons, we only need to prove the first assertion in part (a) of the theorem.

To this end, we fix an initial state $a \in I$ and note that the same computation as in the proof of Lemma 2.1 above yields:

$$H(P_{2t}^a) - H(P_t^a) = \sum_{(i,j) \in I^2} P_t^a(j)\, P_t^j(i) \log \frac{P_t^a(j)\, P_t^j(i)}{P_{2t}^a(i)\, P_t^j(i)}.$$

As before, we have by the Markov property

$$P_t^a(j)\, P_t^j(i) = \mathbb{P}^a(X(t) = j, X(2t) = i). \tag{2.16}$$

Moreover, the detailed balance condition (2.15) gives

$$P_{2t}^a(i)\, P_t^j(i) = P_{2t}^a(i)\, P_t^i(j) \frac{\nu(i)}{\nu(j)} = \mathbb{P}^a(X(2t) = i,\ X(3t) = j) \frac{\nu(i)}{\nu(j)}. \tag{2.17}$$

Plugging this in, we get

$$H(P_{2t}^a) - H(P_t^a) = H(P_{t,2t}^a | P_{3t,2t}^a) + \sum_{(i,j) \in I^2} \mathbb{P}^a(X(t) = j, X(2t) = i) \log \frac{\nu(i)}{\nu(j)}, \tag{2.18}$$

where $P_{t,2t}^a$ and $P_{3t,2t}^a$ denote the laws of the vectors $(X(t), X(2t))$ and $(X(3t), X(2t))$, conditioned on $X(0) = a$. In addition, writing $\log \frac{\nu(i)}{\nu(j)} = \log \nu(i) - \log \nu(j)$ and summing, we obtain

$$H(P_{2t}^a) - H(P_t^a) - H(P_{t,2t}^a | P_{3t,2t}^a)$$
$$= \sum_{i \in I} P^a(X(2t) = i) \log \nu(i) - \sum_{j \in I} P^a(X(t) = j) \log \nu(j).$$

Finally, integrating both sides of the latter equation with respect to ν and using the fact that ν is the invariant distribution of the Markov chain X, we end up with the *averaged entropy identity*

$$\sum_{a \in I} \nu(a) \big(H(P_{2t}^a) - H(P_t^a) \big) = \sum_{a \in I} \nu(a)\, H(P_{t,2t}^a | P_{3t,2t}^a). \tag{2.19}$$

In particular, in view of Pinsker's inequality (see e.g. Lemma 11.6.1 in [14]), this implies the estimate

$$\sum_{a \in I} \nu(a) \big(H(P_{2t}^a) - H(P_t^a) \big) \geq \frac{1}{2} \sum_{a \in I} \nu(a) \, \| P_t^a - P_{3t}^a \|_{TV}^2. \tag{2.20}$$

On the other hand, the first inequality in part (a) of the theorem is equivalent to

$$\sum_{a \in I} \nu(a) \int_{-\infty}^{\infty} \mathbf{1}_{\{u \in \mathbb{R}:\ \| P_{3e^u}^a - P_{e^u}^a \|_{TV} \geq \delta\}}\, du \leq C(\delta)\, H(\nu). \tag{2.21}$$

This in turn would follow from $\mathbf{1}_{\{u \in \mathbb{R}:\ \| P_{3e^u}^a - P_{e^u}^a \|_{TV} \geq \delta\}} \leq \frac{\| P_{3e^u}^a - P_{e^u}^a \|_{TV}^2}{\delta^2}$, if we can prove

$$\sum_{a \in I} \nu(a) \int_{-\infty}^{\infty} \frac{\| P_{3e^u}^a - P_{e^u}^a \|_{TV}^2}{\delta^2}\, du \leq C(\delta)\, H(\nu). \tag{2.22}$$

However, due to the estimate (2.20), the left-hand side in the latter inequality is bounded above by

$$
\sum_{a \in I} \nu(a) \int_{-\infty}^{\infty} \frac{2 \big(H(P_{2e^u}^a) - H(P_{e^u}^a) \big)}{\delta^2} \, \mathrm{d}u
$$
$$
= \frac{2}{\delta^2} \sum_{a \in I} \nu(a) \lim_{K \to \infty} \left(\int_{-K+\log 2}^{K+\log 2} H(P_{e^u}^a) \, \mathrm{d}u - \int_{-K}^{K} H(P_{e^u}^a) \, \mathrm{d}u \right)
$$
$$
= \frac{2 \log 2 \, H(\nu)}{\delta^2}.
$$

This finishes the proof. $\qquad\square$

3 Estimates in L^2 norm

Throughout the first subsection of this section, we assume for the simplicity of notation that the continuous time Markov chain X is irreducible and reversible with respect to the uniform distribution on the set $I = \{1, 2, \ldots, n\}$. We first give a global version of Theorem 2 in Theorem 4 and then prove Theorem 2 at the end of the first subsection. Subsequently, in the second subsection, we give the analogues of these results for a general continuous time irreducible reversible Markov chain on I.

3.1 Markov chains reversible with respect to the uniform distribution

In the following theorem we show that, for most of the time on the scale-invariant clock, the square of the L^2 distance between the one-dimensional distributions of the Markov chain started in a and the one-dimensional distributions of the Markov chain started in b, averaged over all pairs $(a, b) \in I^2$, stays close to the lattice

$$
A_L := \left\{ 0, \frac{2}{n}, \frac{4}{n}, \ldots, \frac{2(n-1)}{n} \right\}. \tag{3.1}
$$

This statement can be viewed as a global (or, averaged) version of Theorem 2. To make this statement precise, we write A_L^δ for the $\frac{2\delta}{n}$-neighborhood of A_L in $\left[0, \frac{2(n-1)}{n} \right]$, where δ is a number in $\left(0, \frac{1}{2} \right)$, and can formulate the following result.

Theorem 4. *In the setting of Theorem 2, for all $0 < \delta < \frac{1}{2}$, there exists a constant $K(\delta) > 0$ such that the estimate*

$$
\int_0^\infty \mathbf{1}_{\left\{ t \geq 0: \, \frac{1}{n^2} \sum_{(a,b) \in I^2} \| P_t^a - P_t^b \|_2^2 \notin A_L^\delta \right\}} \frac{1}{t} \, \mathrm{d}t \leq K(\delta) \, n \tag{3.2}
$$

holds. Hereby, the constant $K(\delta)$ depends only on δ, and not on n or the particular Markov chain X.

Proof. To start with, we recall the notation $\mathcal{L}$ for the generating matrix of the Markov chain X, so that, in particular, the transition matrix P_t corresponding to a time $t \geq 0$ is given by $e^{t\mathcal{L}}$. Since X is irreducible and reversible with respect to the uniform distribution, the matrix $-\mathcal{L}$ is symmetric and non-negatively definite and has the eigenvalues $0 = \lambda_1 < \lambda_2 \leq \lambda_3 \leq \ldots \leq \lambda_n$. In particular, each of the matrices P_t, $t \geq 0$ is symmetric, positively definite and has the eigenvalues

$$
1, \, e^{-\lambda_2 t}, \, e^{-\lambda_3 t}, \, \ldots, \, e^{-\lambda_n t}. \tag{3.3}
$$

Writing $\|.\|_2$ for the L^2 norm with respect to the counting measure on I and $\langle .,.\rangle_2$ for the corresponding scalar product, we can make the following computation:

$$
\begin{aligned}
\sum_{(a,b)\in I^2} \|P_t^a - P_t^b\|_2^2 &= 2n \sum_{a\in I} \|P_t^a\|_2^2 - 2 \sum_{(a,b)\in I^2} \langle P_t^a, P_t^b\rangle_2 \\
&= 2n \sum_{(a,c)\in I^2} \mathbb{P}^a(X(t)=c)^2 - 2\Big\langle \sum_{a\in I} P_t^a, \sum_{b\in I} P_t^b \Big\rangle_2 \\
&= 2n\big(e^{-2\lambda_2 t} + e^{-2\lambda_3 t} + \ldots + e^{-2\lambda_n t}\big),
\end{aligned}
$$

which is valid for all $t \geq 0$.

Next, we set

$$f(t) := e^{-2\lambda_2 t} + e^{-2\lambda_3 t} + \ldots + e^{-2\lambda_n t}, \tag{3.4}$$

$$t_k := \inf\{t \geq 0: \ f(t) \leq k-\delta\}, \quad k = 1, 2, \ldots, n-1. \tag{3.5}$$

The continuity of the function f implies $f(t_k) = k-\delta$. In particular, it follows that

$$f(2t_k) \leq \max_{x_1+x_2+\ldots+x_{n-1}=k-\delta,\, 0\leq x_i\leq 1} (x_1^2 + x_2^2 + \ldots + x_{n-1}^2). \tag{3.6}$$

Moreover, since the maximum of the convex function

$$(x_1, \ldots, x_{n-1}) \mapsto (x_1^2 + x_2^2 + \ldots + x_{n-1}^2)$$

is taken over a convex set, it must be attained at a boundary point of that set. In other words, at the optimizing point it must hold $x_l \in \{0,1\}$ for at least one $1 \leq l \leq n-1$. Eliminating the corresponding variable, we obtain a maximization problem of the same type and conclude that at least one another coordinate $x_{l'}$ has to belong to the set $\{0,1\}$. Proceeding with the same argument we conclude that for each optimizing point $(x_1, x_2, \ldots, x_{n-1})$ there must be $(k-1)$ coordinates, which are equal to 1, $(n-k-1)$ coordinates, which are equal to 0, and one coordinate, which is equal to $1-\delta$. Thus, we have:

$$f(2t_k) \leq f(t_k) - (1-\delta) + (1-\delta)^2 = f(t_k) - \delta(1-\delta). \tag{3.7}$$

Now, either $f(2t_k) \leq (k-1)+\delta$, or we can proceed with the same argument to conclude

$$f(4t_k) \leq f(2t_k) - \delta(1-\delta) \leq f(t_k) - 2\delta(1-\delta). \tag{3.8}$$

Proceeding further with the same argument, we end up with

$$f(2^R t_k) \leq f(t_k) - (1-2\delta) = (k-1)+\delta \tag{3.9}$$

for $R = \left\lceil \frac{1-2\delta}{\delta(1-\delta)} \right\rceil$, where $\lceil . \rceil$ denotes the closest integer from above.

Hence, setting

$$\tilde{t}_k = \inf\{t \geq 0: \ f(t) \leq (k-1)+\delta\}, \quad k = 1, 2, \ldots, n-1, \tag{3.10}$$

we have the estimate

$$\int_0^\infty \mathbf{1}_{[t_k, \tilde{t}_k]} \frac{1}{t}\, dt = \log \frac{\tilde{t}_k}{t_k} \leq \log 2 \left\lceil \frac{1-2\delta}{\delta(1-\delta)} \right\rceil =: K(\delta). \tag{3.11}$$

Finally, using this and the identity

$$\int_0^\infty \mathbf{1}_{\left\{t\geq 0:\ \frac{1}{n^2}\sum_{(a,b)\in I^2} \|P_t^a - P_t^b\|_2^2 \notin A_L^\delta\right\}} \frac{1}{t}\, dt = \sum_{k=1}^n \int_0^\infty \mathbf{1}_{[t_k, \tilde{t}_k]} \frac{1}{t}\, dt, \tag{3.12}$$

we readily obtain the theorem. $\qquad\qquad\square$

Remark 1. We note at this point that the main result of the article [17] implies that, for any vector $(1, \tilde{\lambda}_2, \tilde{\lambda}_3, \ldots, \tilde{\lambda}_n)$ with

$$1 > \tilde{\lambda}_2 \geq \tilde{\lambda}_3 \geq \ldots \geq \tilde{\lambda}_n > 0, \tag{3.13}$$

there is a symmetric doubly stochastic matrix S with eigenvalues $1, \tilde{\lambda}_2, \tilde{\lambda}_3, \ldots, \tilde{\lambda}_n$.

In particular, one can find a matrix $\mathcal{L} = S - \mathrm{Id}$ with the following two properties:

(a) $\mathcal{L}$ generates a continuous time Markov chain, which is irreducible and reversible with respect to the uniform measure.

(b) The matrix $-\mathcal{L}$ has the eigenvalues

$$(0, \lambda_2, \lambda_3, \ldots, \lambda_n) = (0, 1 - \tilde{\lambda}_2, 1 - \tilde{\lambda}_3, \ldots, 1 - \tilde{\lambda}_n) \tag{3.14}$$

for a given sequence $\tilde{\lambda}_2, \tilde{\lambda}_3, \ldots, \tilde{\lambda}_n$ as in (3.13).

This together with the proof of Theorem 4 shows that the order n of the upper bound in Theorem 4 is optimal. As will become clear from the proofs below, the same is true for the upper bound of Theorem 2, and the counterparts of these results for general continuous time irreducible reversible Markov chains treated in Section 3.2.

We proceed with the proof of Theorem 2.

Proof of Theorem 2. To start with, we introduce the following notations:

$$t_k(a, b) := \inf \left\{ t \geq 0 : \|P_t^a - P_t^b\|_2^2 \leq \sum_{l=2}^{k-1} \mu_l^2 + (1 - \delta)\mu_k^2 \right\}, \ k = 2, \ldots, n$$

$$\tilde{t}_k(a, b) := \inf \left\{ t \geq 0 : \|P_t^a - P_t^b\|_2^2 \leq \sum_{l=2}^{k-1} \mu_l^2 + \delta\mu_k^2 \right\}, \ k = 2, \ldots, n$$

and note that

$$\int_0^\infty \mathbf{1}_{\left\{ t \geq 0 : \ \|P_t^a - P_t^b\|_2^2 \notin A^\delta(a,b) \right\}} \frac{1}{t} \, dt = \sum_{k=2}^n \log \frac{\tilde{t}_k(a, b)}{t_k(a, b)}. \tag{3.15}$$

From now on, we fix a $k = 2, 3, \ldots, n$ and will show that

$$\log \frac{\tilde{t}_k(a, b)}{t_k(a, b)} \leq K(\delta) \tag{3.16}$$

for a suitable constant $K(\delta) > 0$, which depends only on δ (but not on a, b, k or n). To this end, we note that the identity

$$\|P_t^a - P_t^b\|_2^2 = \sum_{l=2}^n \mu_l^2 e^{-2t\lambda_l}, \quad t \geq 0 \tag{3.17}$$

and the inequality $0 < \lambda_2 \leq \lambda_3 \leq \ldots \leq \lambda_n$ imply the estimate

$$\|P_{2t_k(a,b)}^a - P_{2t_k(a,b)}^b\|_2^2 \leq \max_{\sum_{l=2}^n \mu_l^2 x_l = \|P_{t_k(a,b)}^a - P_{t_k(a,b)}^b\|_2^2, \ 1 \geq x_2 \geq x_3 \geq \ldots \geq x_n \geq 0} \sum_{l=2}^n \mu_l^2 x_l^2. \tag{3.18}$$

Moreover, if we have $\mu_l^2 > 0$ for all $l = 2, 3, \ldots, n$, then the function

$$(x_2, x_3, \ldots, x_n) \mapsto \sum_{l=2}^n \mu_l^2 x_l^2$$

is stricly convex and must attain its maximum at a vertex point of the convex polyhedron

$$\left\{ (x_2, x_3, \ldots, x_n) : \sum_{l=2}^{n} \mu_l^2 x_l = \|P_{t_k(a,b)}^a - P_{t_k(a,b)}^b\|_2^2, \ 1 \ge x_2 \ge x_3 \ge \ldots \ge x_n \ge 0 \right\}.$$

If we have $\mu_l^2 = 0$ for some $l \in \{2, 3, \ldots, n\}$, then we can elimininate the corresponding coordinate in the maximization problem and make the same conclusion for the reduced maximization problem. For this reason, we may assume without loss of generality that $\mu_l^2 > 0$ for all $l = 2, 3, \ldots, n$. Moreover, since the hyperplane $\sum_{l=2}^{n} \mu_l^2 x_l = \|P_{t_k(a,b)}^a - P_{t_k(a,b)}^b\|_2^2$ is $(n-2)$-dimensional, the vertices of the polyhedron above are given by points $1 \ge x_2 \ge x_3 \ge \ldots \ge x_n \ge 0$, for which $(n-2)$ of the inequalities

$$1 \ge x_2, \ x_2 \ge x_3, \ \ldots, \ x_n \ge 0$$

are in fact equalities.

Thus, each optimizing point of the maximization problem above can be described as follows: There is a partition of $\{2, 3, \ldots, n\}$ into three sets I_1, I_2, I_3 of the form $\{2, 3, \ldots, l_1\}$, $\{l_1 + 1, l_1 + 2, \ldots, l_2\}$, $\{l_2 + 1, l_2 + 2, \ldots, n\}$, respectively, such that, for all $l \in I_1$ it holds $x_l = 1$, for all $l \in I_2$ we have $x_l = \zeta$ for a suitable $\zeta \in [0, 1]$, and for all $l \in I_3$ it holds $x_l = 0$. Moreover, the identity

$$\|P_{t_k(a,b)}^a - P_{t_k(a,b)}^b\|_2^2 = \sum_{l=2}^{k-1} \mu_l^2 + (1 - \delta)\mu_k^2 \tag{3.19}$$

shows that the value of ζ is given by

$$\zeta = \frac{\left(\sum_{l=2}^{k-1} \mu_l^2 + (1 - \delta)\mu_k^2 \right) - \left(\sum_{l \in I_1} \mu_l^2 \right)}{\sum_{l \in I_2} \mu_l^2}, \tag{3.20}$$

and that $I_1 \subset \{2, 3, \ldots, k-1\}$. To proceed, we introduce the set

$$\tilde{I}_2 := \left(\{2, 3, \ldots, k-1\} \backslash I_1 \right) \subset I_2 \tag{3.21}$$

and conclude

$$\zeta = \frac{\left(\sum_{l \in \tilde{I}_2} \mu_l^2 \right) + (1 - \delta)\mu_k^2}{\sum_{l \in I_2} \mu_l^2}, \quad \tilde{I}_2 \subsetneq I_2. \tag{3.22}$$

This allows us to make the following computation:

$$\|P_{2t_k(a,b)}^a - P_{2t_k(a,b)}^b\|_2^2 = \|P_{t_k(a,b)}^a - P_{t_k(a,b)}^b\|_2^2 - \left(\sum_{l \in I_2} \mu_l^2 \zeta \right) + \left(\sum_{l \in I_2} \mu_l^2 \zeta^2 \right)$$

$$= \|P_{t_k(a,b)}^a - P_{t_k(a,b)}^b\|_2^2 - \frac{\left(\sum_{l \in \tilde{I}_2} \mu_l^2 + (1 - \delta)\mu_k^2 \right) \left(\sum_{l \in I_2 \backslash \tilde{I}_2} \mu_l^2 - (1 - \delta)\mu_k^2 \right)}{\sum_{l \in I_2} \mu_l^2}.$$

Next, we note that the latter fraction is of the form $\frac{A \cdot B}{A + B} = \frac{1}{\frac{1}{A} + \frac{1}{B}}$, whereby: $A \ge (1-\delta)\mu_k^2$ and $B \ge \delta\mu_k^2$. Thus,

$$\|P_{2t_k(a,b)}^a - P_{2t_k(a,b)}^b\|_2^2 \ \le \ \|P_{t_k(a,b)}^a - P_{t_k(a,b)}^b\|_2^2 - \frac{1}{\frac{1}{(1-\delta)\mu_k^2} + \frac{1}{\delta\mu_k^2}}$$

$$= \ \|P_{t_k(a,b)}^a - P_{t_k(a,b)}^b\|_2^2 - \frac{\mu_k^2}{\frac{1}{1-\delta} + \frac{1}{\delta}}.$$

Proceeding with the same argument, we conclude that

$$\|P^a_{2^R t_k(a,b)} - P^b_{2^R t_k(a,b)}\|^2_2 \le \|P^a_{t_k(a,b)} - P^b_{t_k(a,b)}\|^2_2 - \mu^2_k(1 - 2\delta) \tag{3.23}$$

for all natural numbers $R \ge \frac{1-2\delta}{\frac{1}{1-\delta} + \frac{1}{\delta}}$. In particular, we conclude that

$$\tilde{t}_k(a,b) \le 2^{\left\lceil \frac{1-2\delta}{\frac{1}{1-\delta} + \frac{1}{\delta}} \right\rceil} t_k(a,b), \tag{3.24}$$

where $\lceil . \rceil$ denotes the closest integer from above. This shows the claim (3.16) with

$$K(\delta) = \log 2 \cdot \left\lceil \frac{1 - 2\delta}{\frac{1}{1-\delta} + \frac{1}{\delta}} \right\rceil \tag{3.25}$$

and finishes the proof. □

3.2 General reversible Markov chains

We proceed with the analogues of Theorems 2 and 4 for a general continuous time irreducible reversible Markov chain X. To state the results, we introduce the following set of notations. We write ν for the invariant measure of X as before, and let D be the diagonal matrix, whose diagonal entries are given by $\nu(i)$, $i \in I$. Then, by the detailed balance condition (2.15), the matrix $D^{1/2}P_t D^{-1/2}$ is symmetric for all $t \ge 0$. Moreover, since the matrices $D^{1/2}P_t D^{-1/2}$, $t \ge 0$ commute, they have a joint orthonormal basis of eigenvectors $v_1, v_2, \ldots, v_n$ corresponding to sets eigenvalues

$$1 > e^{-\lambda_2 t} \ge e^{-\lambda_3 t} \ge \ldots \ge e^{-\lambda_n t}, \quad t \ge 0, \tag{3.26}$$

respectively (see chapter 3 of the book [1] for more details). In addition, for any fixed pair (a, b) of initial states, we let

$$D^{-1/2}(e_a - e_b) = \sum_{l=2}^{n} \tilde{\mu}_l v_l \tag{3.27}$$

be the expansion of the vector $D^{-1/2}(e_a - e_b)$ in terms of the basis $v_1, v_2, \ldots, v_n$ (note that the vector $D^{-1/2}(e_a - e_b)$ is orthogonal to the eigenvector corresponding to the eigenvalue 1 of the matrices $D^{1/2}P_t D^{-1/2}$, $t \ge 0$). Finally, define the sets

$$\tilde{A}_L := \{0, 1, 2, \ldots, n-1\},$$
$$\tilde{A}(a,b) := \{0, \tilde{\mu}^2_2, \tilde{\mu}^2_2 + \tilde{\mu}^2_3, \ldots, \tilde{\mu}^2_2 + \tilde{\mu}^2_3 + \ldots + \tilde{\mu}^2_n\},$$

and their neighborhoods

$$\tilde{A}^\delta_L := [0, \delta] \cup [1 - \delta, 1 + \delta] \cup [2 - \delta, 2 + \delta] \cup \ldots \cup [n - 1 - \delta, n - 1],$$
$$\tilde{A}^\delta(a,b) := [0, \delta\tilde{\mu}^2_2] \cup [(1-\delta)\tilde{\mu}^2_2, \tilde{\mu}^2_2 + \delta\tilde{\mu}^2_3] \cup \ldots \cup \left[\sum_{l=2}^{n-1} \tilde{\mu}^2_l + (1-\delta)\tilde{\mu}^2_n, \sum_{l=2}^{n} \tilde{\mu}^2_l \right],$$

$0 < \delta < \frac{1}{2}$. With these notations, the analogues of Theorems 2 and 4 read as follows.

Theorem 5. *Let* $\|.\|_{L^2(\nu^{-1})}$ *and* $\langle ., . \rangle_{L^2(\nu^{-1})}$ *be the* L^2 *norm and scalar product with respect to the weights* $\nu(i)^{-1}$, $i \in I$. *Then, for all* $0 < \delta < \frac{1}{2}$, *there is a constant* $\tilde{K}(\delta) > 0$ *such that the inequalities*

$$\int_0^\infty \mathbf{1}_{\left\{ t \ge 0: \sum_{(a,b) \in I^2} \nu(a)\,\nu(b)\,\|P^a_t - P^b_t\|^2_{L^2(\nu^{-1})} \notin \tilde{A}^\delta_L \right\}} \frac{1}{t}\, dt \le \tilde{K}(\delta)\, n \tag{3.28}$$

and

$$\int_0^\infty \mathbf{1}_{\left\{ t \geq 0: \, \|P_t^a - P_t^b\|_{L^2(\nu^{-1})}^2 \notin \tilde{A}^\delta(a,b) \right\}} \frac{1}{t} \, \mathrm{d}t \leq \tilde{K}(\delta) \, n, \quad (a,b) \in I^2, \tag{3.29}$$

hold true. The constant $\tilde{K}(\delta)$ *depends only on* δ, *and not on* n *or the particular Markov chain* X.

Proof. In order to prove (3.28), we use the fact that ν is the invariant distribution of the Markov chain X to deduce the identities

$$\sum_{(a,b) \in I^2} \nu(a) \, \nu(b) \, \|P_t^a - P_t^b\|_{L^2(\nu^{-1})}^2$$

$$= 2 \sum_{a \in I} \nu(a) \, \|P_t^a\|_{L^2(\nu^{-1})}^2 - 2 \Big\langle \sum_{a \in I} \nu(a) \, P_t^a, \sum_{b \in I} \nu(b) \, P_t^b \Big\rangle$$

$$= 2 \sum_{a \in I} \nu(a) \sum_{c \in I} P_t^a(c)^2 \, \nu(c)^{-1} - 2 \langle \nu, \nu \rangle_{L^2(\nu^{-1})}$$

$$= 2 \sum_{(a,c) \in I^2} \nu(a) \, P_t^a(c)^2 \, \nu(c)^{-1} - 2,$$

which hold for all $t \geq 0$. Moreover, the latter sum is given by the sum of squares of the entries of the matrix $D^{1/2} P_t D^{-1/2}$ and is, hence, equal to $1 + \sum_{l=2}^n e^{-2\lambda_l t}$. Thus,

$$\sum_{(a,b) \in I^2} \nu(a) \, \nu(b) \, \|P_t^a - P_t^b\|_{L^2(\nu^{-1})}^2 = \sum_{l=2}^n e^{-2\lambda_l t} \quad t \geq 0. \tag{3.30}$$

From this point on, one can proceed as in the proof of Theorem 4 to show (3.28).

Now, we turn to the proof of (3.29). To this end, we note that the detailed balance condition (2.15) implies $D P_t = P_t^T D$, $t \geq 0$, where the superscript T stands for the transpose of a matrix. This allows us to make the computation

$$P_t^a - P_t^b = ((e_a - e_b)^T P_t)^T = P_t^T (e_a - e_b) = D^{1/2} \big(D^{1/2} P_t D^{-1/2} \big) D^{-1/2} (e_a - e_b)$$

$$= D^{1/2} \sum_{l=2}^n \tilde{\mu}_l \, e^{-\lambda_l t} \, v_l = \sum_{l=2}^n \tilde{\mu}_l \, e^{-\lambda_l t} \, D^{1/2} v_l$$

for all $t \geq 0$. Next, we observe that the vectors $D^{1/2} v_1, D^{1/2} v_2, \ldots, D^{1/2} v_n$ form an orthonormal basis with respect to the scalar product $\langle .,. \rangle_{L^2(\nu^{-1})}$, since the vectors $v_1, v_2, \ldots, v_n$ form an orthonormal basis with respect to the standard Euclidean scalar product. Hence,

$$\|P_t^a - P_t^b\|_{L^2(\nu^{-1})}^2 = \sum_{l=2}^n \tilde{\mu}_l^2 \, e^{-2\lambda_l t}, \quad t \geq 0. \tag{3.31}$$

From this point on, one only needs to follow the arguments in the proof of Theorem 2 to end up with (3.29). $\qquad \square$

Remark 2. It is worth noting that the estimates of Theorems 2, 4 and 5 hold for

$$\|P_t^a - \nu\|_2^2, \quad \frac{2}{n} \sum_{a \in I} \|P_t^a - \nu\|_2^2, \quad 2 \sum_{a \in I} \nu(a) \|P_t^a - \nu\|_{L^2(\nu^{-1})}^2 \quad \text{and} \quad \|P_t^a - \nu\|_{L^2(\nu^{-1})}^2$$

in place of

$$\|P_t^a - P_t^b\|_2^2, \quad \frac{1}{n^2} \sum_{(a,b) \in I^2} \|P_t^a - P_t^b\|_2^2, \quad \sum_{(a,b) \in I^2} \nu(a) \nu(b) \|P_t^a - P_t^b\|_{L^2(\nu^{-1})}^2$$

$$\text{and} \quad \|P_t^a - P_t^b\|_{L^2(\nu^{-1})}^2,$$

respectively. The same proofs apply, with the only difference being that one needs to expand the vectors $(e_a - \nu)$ and $D^{-1/2}(e_a - \nu)$ in terms of an orthonormal basis of eigenvectors of the matrices P_t, $t \geq 0$ and $D^{1/2} P_t D^{-1/2}$, $t \geq 0$, respectively.

4 Bounds on the global convergence to equilibrium through the entropy

We have seen in Section 2 that one can obtain a control on the convergence to equilibrium and the times of coupling by analyzing the entropy that is accumulated by the Markov chain over time. In this section, we pursue this idea further and give estimates on the approach to equilibrium on subsets of macroscopic size (that is, with size proportional to the size of the state space) for continuous time irreducible Markov chains which are *reversible with respect to the uniform distribution*. To this end, for each $0 < \kappa < 1$ and $t \geq 0$, we introduce the set

$$EQ_\kappa(t) = \left\{ a \in I : \ P_t(a) \in \left(\frac{1-\kappa}{n}, \frac{1+\kappa}{n} \right) \right\} \subset \{1, 2, \ldots, n\} = I, \qquad (4.1)$$

where, with a slight abuse of notation, we wrote P_t for the law of the random variable $X(t)$. For each $t \geq 0$, the set $EQ_\kappa(t)$ should be viewed as the part of the state space on which the probability measure P_t is close to the equilibrium distribution of the Markov chain X. We are interested in lower bounds on the size $|EQ_\kappa(t)|$ of such sets.

The following result differs from existing results employing entropy to study the convergence to equilibrium in total variation distance and L^2 norm (such as the ones in [5], [15] and the references there) in that it does not rely on functional inequalities satisfied by the invariant distribution of the chain, but rather uses information on the entropy of the one-dimensional marginals of the chain.

Theorem 6. *Fix real numbers $0 < \kappa < 1$ and $0 < \alpha < \frac{1}{2}$, set $\tilde{\alpha} = \frac{1}{2} - \alpha$, and, on the interval $[0, \tilde{\alpha}]$, define the function*

$$
\begin{aligned}
F_\kappa(\alpha_1) \ = \ & \alpha_1(1-\kappa)\log(1-\kappa) + (\tilde{\alpha} - \alpha_1)(1+\kappa)\log(1+\kappa) \\
& + (1 - (1-\kappa)\alpha_1 - (1+\kappa)(\tilde{\alpha} - \alpha_1)) \log\left(\frac{1 - \tilde{\alpha} - \tilde{\alpha}\kappa + 2\kappa\alpha_1}{1 - \tilde{\alpha}} \right)
\end{aligned}
$$

taking non-negative values. Then, the entropy estimate

$$\log n - H(P_t) < \min_{0 \leq \alpha_1 \leq \tilde{\alpha}} F_\kappa(\alpha_1) \qquad (4.2)$$

implies the lower bound

$$|EQ_\kappa(t)| \geq \alpha n. \qquad (4.3)$$

Hereby, depending on the values of κ and α, the minimum in (4.2) is attained at 0, $\tilde{\alpha}$ or α_1^* (defined in (4.9) below).

Proof. We fix numbers κ and α as in the statement of the theorem and suppose that the inequality (4.3) does not hold. We will show that this implies that the entropy bound (4.2) cannot hold. To start with, we introduce the notation $p_a := P_t(a)$, $a \in I$, and make the decomposition

$$H(P_t) = - \sum_{a \in EQ_\kappa(t)} p_a \log p_a - \sum_{b \notin EQ_\kappa(t)} p_b \log p_b. \qquad (4.4)$$

For any given value of $\rho := \sum_{a \in EQ_\kappa(t)} p_a \in [0, 1]$, the maximum of the function $-\sum_{b \notin EQ_\kappa(t)} p_b \log p_b$ is attained on the interior boundary of the set

$$\left\{ \sum_{b \notin EQ_\kappa(t)} p_b = 1 - \rho : \ p_b \notin \left(\frac{1-\kappa}{n}, \frac{1+\kappa}{n} \right) \right\}. \tag{4.5}$$

Indeed, this is a consequence of the fact that the function

$$(p_b : \ b \notin EQ_\kappa(t)) \mapsto - \sum_{b \notin EQ_\kappa(t)} p_b \log p_b \tag{4.6}$$

is concave and attains its maximum over the convex set $\{\sum_{b \notin EQ_\kappa(t)} p_b = 1 - \rho\}$ at the point $\left(\frac{1-\rho}{n - |EQ_\kappa(t)|}, \frac{1-\rho}{n - |EQ_\kappa(t)|}, \cdots, \frac{1-\rho}{n - |EQ_\kappa(t)|} \right)$, which is not an element of the set in (4.5). The latter statement follows from the inequalities

$$\frac{1-\rho}{n - |EQ_\kappa(t)|} \geq \frac{1 - \frac{1+\kappa}{n}|EQ_\kappa(t)|}{n - |EQ_\kappa(t)|} > \frac{1-\kappa}{n} \tag{4.7}$$

and

$$\frac{1-\rho}{n - |EQ_\kappa(t)|} \leq \frac{1 - \frac{1-\kappa}{n}|EQ_\kappa(t)|}{n - |EQ_\kappa(t)|} < \frac{1+\kappa}{n} \tag{4.8}$$

with the respective second inequalities in the latter two chains of estimates being consequences of $|EQ_\kappa(t)| < \frac{n}{2}$.

From the preceeding argument we conclude that at least one of the coordinates of a point in the set (4.5), which maximizes the function in (4.6), has be equal to $\frac{1-\kappa}{n}$ or $\frac{1+\kappa}{n}$. Eliminating this coordinate and proceeding with the same argument, we deduce that at least $\frac{n}{2} - |EQ_\kappa(t)|$ coordinates of an optimizing point have to be equal to $\frac{1-\kappa}{n}$ or $\frac{1+\kappa}{n}$. Now, eliminating all coordinates, which belong to the set $\left\{ \frac{1-\kappa}{n}, \frac{1+\kappa}{n} \right\}$, we deduce the following: If the inequality (4.3) fails, then the entropy $H(P_t)$ cannot exceed the entropy of a probability measure on a set of n elements, for which at least $\frac{n}{2} - |EQ_\kappa(t)|$ of its weights belong to the set $\left\{ \frac{1-\kappa}{n}, \frac{1+\kappa}{n} \right\}$ and the rest of its weights is equal. In other words, denoting the proportion of weights, which are equal to $\frac{1-\kappa}{n}$, by α_1 and the proportion of weights, which are equal to $\frac{1+\kappa}{n}$, by α_2, we have: $H(P_t) \leq \max_{\alpha_1, \alpha_2} \overline{F}_\kappa(\alpha_1, \alpha_2)$ with

$$\begin{aligned} \overline{F}_\kappa(\alpha_1, \alpha_2) &= \Big[-\alpha_1(1-\kappa) \log \frac{1-\kappa}{n} - \alpha_2(1+\kappa) \log \frac{1+\kappa}{n} \\ &\quad -(1 - (1-\kappa)\alpha_1 - (1+\kappa)\alpha_2) \log \frac{1 - (1-\kappa)\alpha_1 - (1+\kappa)\alpha_2}{(1 - \alpha_1 - \alpha_2)n} \Big]. \end{aligned}$$

Hereby, the maximum is taken under the constraints $\frac{1}{2} - \alpha \leq \alpha_1 + \alpha_2 \leq 1$, $\alpha_1 \geq 0$, $\alpha_2 \geq 0$, $1 - (1-\kappa)\alpha_1 - (1+\kappa)\alpha_2 \geq 0$.

Next, we note that $\overline{F}_\kappa$ can be written as $(\log n) - \tilde{F}_\kappa$, where $\tilde{F}_\kappa$ is given by

$$\begin{aligned} \tilde{F}_\kappa(\alpha_1, \alpha_2) &= \Big[\alpha_1(1-\kappa) \log(1-\kappa) + \alpha_2(1+\kappa) \log(1+\kappa) \\ &\quad + (1 - (1-\kappa)\alpha_1 - (1+\kappa)\alpha_2) \log \frac{1 - (1-\kappa)\alpha_1 - (1+\kappa)\alpha_2}{1 - \alpha_1 - \alpha_2} \Big]. \end{aligned}$$

Hence, $H(P_t) \leq (\log n) - \min_{\alpha_1, \alpha_2} \tilde{F}_\kappa(\alpha_1, \alpha_2)$, where the minimum is taken over the region described at the end of the previous paragraph. Now, a straightforward computation of the Hessian of $\tilde{F}_\kappa$ together with the constraint $1 - (1-\kappa)\alpha_1 - (1+\kappa)\alpha_2 \geq 0$

show that the function $\tilde{F}_\kappa$ is convex throughout the region over which its minimum is taken. In addition, the minimum of $\tilde{F}_\kappa$ over the region determined by the constraints $\alpha_1 \geq 0$, $\alpha_2 \geq 0$, $\alpha_1 + \alpha_2 \leq 1$, $1 - (1 - \kappa)\alpha_1 - (1 + \kappa)\alpha_2 \geq 0$ is attained at the point $(0, 0)$ and is equal to 0, since it corresponds to the highest value of the entropy that a probability measure on a set of n elements can take (namely, $\log n$). Thus, the minimum of $\tilde{F}_\kappa$ over the region of interest is attained on the segment given by the constraints $\alpha_1 \geq 0$, $\alpha_2 \geq 0$, $\alpha_1 + \alpha_2 = \frac{1}{2} - \alpha$. Plugging in $\frac{1}{2} - \alpha - \alpha_1$ instead of α_2, and recalling the notation $\tilde{\alpha} = \frac{1}{2} - \alpha$, we end up with $H(P_t) \leq (\log n) - \min_{0 \leq \alpha_1 \leq \tilde{\alpha}} F_\kappa(\alpha_1)$. This is the desired contradiction to (4.2).

We also observe that the function F_κ must be non-negative throughout $[0, \tilde{\alpha}]$, since the entropy of a probability measure on a set of n elements cannot exceed the value $\log n$. Moreover, since the function $\tilde{F}_\kappa$ is convex, the function F_κ is also convex. Furthermore, a straightforward computation shows that, depending on the values of κ and α, either the derivative of the function F_κ has no zeros on the interval $[0, \tilde{\alpha}]$, in which case F_κ attains its minimum at one of the boundary points, or the only zero of the derivative of F_κ on the interval $[0, \tilde{\alpha}]$ is given by

$$\alpha_1^* := \frac{(1-\kappa)^{-\frac{1}{2\kappa}}\left((1-\tilde{\alpha})\sqrt{1-\kappa}(1+\kappa)^{\frac{1+\kappa}{2\kappa}} - e(1-\kappa)^{\frac{1}{2\kappa}}\left(1 - \tilde{\alpha} - \tilde{\alpha}\kappa\right)\right)}{2e\kappa}, \tag{4.9}$$

in which case F_κ attains its minimum at α_1^*. This finishes the proof. $\quad\square$

Acknowledgments. The author would like to thank David J. Aldous for his comments throughout the preparation of this work. He is also grateful to Anton Bovier for his remarks on an early version of this manuscript.

References

[1] Aldous, D. J. and Fill, J. A. (2002). *Reversible Markov Chains and Random Walks on Graphs*. Available at `http://www.stat.berkeley.edu/~aldous/RWG/book.html`.

[2] Ben Arous, G. and Cerf, R. (1996). Metastability of the three-dimensional Ising model on a torus at very low temperatures. *Electron. J. Probab.* **10**. MR-1423463

[3] Bianchi, A., Bovier, A. and Ioffe, D. (2008). Sharp asymptotics for metastability in the Random Field Curie-Weiss model. *Electron. J. Probab.* **14**, 1541-1603. MR-2525104

[4] Bianchi, A., Bovier, A. and Ioffe, D. (2009). Pointwise estimates and exponential laws in metastable systems via coupling methods. To appear in *Ann. Probab.* MR-2917775

[5] Bobkov, S. and Tetali, P. (2006). Modified logarithmic Sobolev inequalities in discrete settings. *J. Theoret. Probab.* **19**, 28-36. MR-2283379

[6] Borgs, C., Chayes, J. and Tetali, P. (2012). Tight bounds for mixing of the Swendsen-Wang algorithm at the Potts transition point. **152**, 509-557. MR-2892955

[7] Bovier, A. (2004). Metastability and ageing in stochastic dynamics. In *Dynamics and Randomness II*, Nonlinear Phenomena and Complex Systems **10**. Kluwer Acad. Publishers. MR-2153326

[8] Bovier, A. (2006). Metastability: a potential theoretic approach. In *Proceedings of the ICM 2006*, 499-518. European Mathematical Society. MR-2275693

[9] Bovier, A., Eckhoff, M., Gayrard, V. and Klein, M. (2000). Metastability and small eigenvalues in Markov chains. J. Phys. **A 33**, L447-L451. MR-1804034

[10] Bovier, A., Eckhoff, M., Gayrard, V. and Klein, M. (2001). Metastability in stochastic dynamics of disordered mean-field models. *Probab. Theor. Rel. Fields* **119**, 99-161. MR-1813041

[11] Bovier, A., Eckhoff, M., Gayrard, V. and Klein, M. (2002). Metastability and low-lying spectra in reversible Markov chains. *Comm. Math. Phys.* **228**, 219-255. MR-1911735

[12] Bovier, A. and Manzo, F. (2002). Metastability in Glauber dynamics in the low temperature limit: Beyond exponential asymptotics. *J. Statist. Phys.* **107**, 757-779. MR-1898856

[13] Cassandro, M., Galvez, A., Olivieri, E. and Vares, M. E. Metastable behavior of stochastic dynamics: a pathwise approach. *J. Statist. Phys.* **35**, 603-634. MR-0749840

[14] Cover, T. M. and Thomas, J. A. (2006). *Elements of Information Theory*. 2nd ed. John Wiley & Sons, Inc. MR-2239987

[15] Diaconis, P. and Saloff-Coste, L. (1996). Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.* **6**, 695-750. MR-1410112

[16] Dembo A. and Zetouni O. (1998). *Large deviation techniques and applications.* 2nd ed. Springer, New York. MR-1619036

[17] Hwang, S.-G. and Pyo S.-S. (2004). The inverse eigenvalue problem for symmetric doubly stochastic matrices. *Linear Algebra and Its Applications* **379** 77-83. MR-2039298

[18] Kramers, H. A. (1940). Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica.* **7**, 284-304. MR-0002962

[19] Levin, D. A., Luczak, M. J. and Peres, Y. (2010). Glauber dynamics for the mean-field Ising model: cut-off, critical power law, and metastability. *Probab. Theory Related Fields.* **146**, 223-265. MR-2550363

[20] Neves, E. J. and Schonmann, R. H. (1991). Critical droplets and metastability for a Glauber dynamics at very low temperatures. *Comm. Math. Phys.* **137**, 209-230. MR-1101685

[21] Neves, E. J. and Schonmann, R. H. (1992). Behavior of droplets for a class of Glauber dynamics at very low temperature. *Probab. Theory Related Fields.* **91**, 331-354. MR-1151800

[22] Olivieri, E. and Scoppola. (1995). Markov chains with exponentially small transition probabilities: first exit problem from a general domain. I. The reversible case. *J. Statist. Phys.* **79**, 613-647. MR-1327899

[23] Thomas, L. E. (1989). Bound on the mass gap for finite volume stochastic Ising models at low temperature. *Comm. Math. Phys.* **126**, 1-11. MR-1027910