Contents

Abstract

We prove comparison results for the Swendsen–Wang (SW) dynamics, the heat-bath (HB) dynamics for the Potts model and the single-bond (SB) dynamics for the random-cluster model on arbitrary graphs. In particular, we prove that rapid (i.e. polynomial) mixing of HB implies rapid mixing of SW on graphs with bounded maximum degree and that rapid mixing of SW and rapid mixing of SB are equivalent. Additionally, the spectral gap of SW and SB on planar graphs is bounded from above and from below by the spectral gap of these dynamics on the corresponding dual graph with suitably changed temperature.

As a consequence we obtain rapid mixing of the Swendsen–Wang dynamics for the Potts model on the two-dimensional square lattice at all non-critical temperatures, as well as rapid mixing for the two-dimensional Ising model at all temperatures. Furthermore, we obtain new results for general graphs at high or low enough temperatures.

Acknowledgements. First and foremost, I would like to thank the advisor of my PhD thesis (the present paper is based on it), Erich Novak, for his constant support and the many discussions we had, not only about mathematics. This has certainly led to the development of my present interest in and fun with mathematics.

I am also grateful to Aicke Hinrichs and Daniel Rudolf for helpful comments and suggestions during the work on my thesis, as well as for the valuable discussions we had about various topics.

Finally, I want to express my thanks to my colleagues and friends Erich, Aicke, Daniel, Henning Kempka, Lev Markhasin, Philipp Rudolph, Winfried Sickel, Markus Weimar and Heidi Weyhausen who were (more or less regular) members of our daily coffee break that served with its friendly and relaxed atmosphere as a source of extra energy every day.

2010 *Mathematics Subject Classification*: Primary 60K35; Secondary 60J10, 82C20.
Key words and phrases: Swendsen–Wang dynamics, rapid mixing, spectral gap, Ising model, Potts model, random-cluster model, single-bond dynamics.
Received 18.2.2013; revised version 18.9.2013.

1. Introduction and results

We study the mixing properties of Markov chains for the *q-state Potts model*. The Potts model with $q \in \mathbb{N}$ states at *inverse temperature* $\beta \geq 0$ consists of the set $\Omega_{\mathrm{P}} = \{1, \dots, q\}^V$ of (not necessarily proper) colorings of a graph $G = (V, E)$ together with the probability measure

$$\pi(\sigma) = \frac{1}{Z} e^{\beta |E(\sigma)|}, \qquad \sigma \in \Omega_{\mathrm{P}},$$

where $E(\sigma)$ denotes the set of edges of the graph G with equally colored endvertices and Z is the normalization constant. This model, especially in the case $q = 2$ where it is called the *Ising model*, attracted much interest over the last decades not only in statistical physics but also in several branches of mathematics and computer science. The goal is to sample from π, at least approximately, and since exact sampling is in general not feasible, Markov chains are typically used. There are a couple of (more or less efficient) Markov chains to attack this problem and, usually, it is an easy task to show that their distributions converge in the long-time limit towards the right distribution. However, it is much more delicate to prove bounds on the *mixing time*, i.e. the number of steps a Markov chain has to run in order that its distribution is "close" to its limit distribution. In the following, a chain is said to be *rapidly mixing* for a family of graphs if, for each graph of this family, the Markov chain can be defined analogously on it and its mixing time is bounded by a polynomial in the number of vertices of the graph.

We will consider two Markov chains for the Potts model, namely the *heat-bath* (*HB*) *dynamics* and the *Swendsen–Wang* (*SW*) *dynamics*. The heat-bath dynamics is the most common chain for this purpose. It is a local Markov chain that can be described as follows. Suppose that the current configuration is $\sigma = (\sigma(v))_{v \in V} \in \Omega_{\mathrm{P}}$. In each step, a vertex $v \in V$ of the underlying graph is chosen uniformly at random and a new color is assigned to v with respect to the conditional probability given that the color of all other vertices is fixed, so that the new configuration τ satisfies $\tau(u) = \sigma(u)$ for all $u \neq v$. The heat-bath dynamics is proven to be rapidly mixing in several instances (see Section 2.5), but is typically slowly (i.e. not rapidly) mixing if the inverse temperature β is large. For example, if the underlying graph is the two-dimensional square lattice $\mathbb{Z}_L^2$ of side length L, it is known that there exists a *critical inverse temperature* $\beta_c(q)$ such that the heat-bath dynamics is rapidly mixing if $\beta < \beta_c(q)$ and slowly mixing if $\beta > \beta_c(q)$ (the latter seems to be proven only for $q = 2$ but it is at least expected to be true for all $q \geq 2$). See Section 2.5 for more known results and the specific bounds.

The second Markov chain under consideration is the Swendsen–Wang dynamics that changes the color of a large portion of the vertices in each step. One step of this chain,

given the current state $\sigma \in \Omega_P$, can be described by the following two-step procedure. First, generate a subset $A \subset E(\sigma)$ of the edges of the graph with equally colored endvertices, such that every edge of $E(\sigma)$ is included in A with probability $1 - e^{-\beta}$. In the second step, assign independently and uniformly at random a color from $\{1, \dots, q\}$ to each connected component of the subgraph (V, A). This gives a new Potts configuration. It is widely believed that this Markov chain is rapidly mixing in much more cases than the heat-bath dynamics, e.g. at low temperatures (large β), and is therefore the preferred algorithm in practice since its invention around 1987. Nevertheless, results in this direction are rare. Besides some rapid mixing results concerning special classes of graphs like trees, cycles and the complete graph, or results for sufficiently high or low temperatures, there is no result that shows that it is generally advisable to use Swendsen–Wang instead of heat-bath dynamics.

Our first main result (see Chapter 3) shows that rapid mixing of heat-bath dynamics implies rapid mixing of Swendsen–Wang, if the underlying graph has bounded maximum degree, which partially confirms the above-mentioned intuition.

In order to give precise statements we need some notation. We define the spectral gap of a Markov chain with transition matrix P as 1 minus the second largest eigenvalue of P in absolute value and we denote it by $\lambda(P)$. It is well-known that polynomial (in $|V|$) bounds on the mixing time are equivalent to polynomial bounds on the inverse spectral gap (see Lemma 2.3), and since the spectral gap seems to be more convenient for comparison results, this is the quantity we are interested in. (Note that we classify Markov chains by their transition matrices, since the quantities of interest do not depend on anything else.)

For the following comparison results we consider the two above-mentioned Markov chains for the Potts model, as defined in Section 2.4, and we use this notation for their transition matrices:

- P_{HB} for the heat-bath dynamics for the Potts model (see (2.18)),
- P_{SW} for the Swendsen–Wang dynamics for the Potts model (see (2.21)).

The first result that we want to present here is a comparison between heat-bath and Swendsen–Wang dynamics for the Potts model.

THEOREM 3.1. *Suppose that P_{SW} (resp. P_{HB}) is the transition matrix of the Swendsen–Wang (resp. heat-bath) dynamics for the q-state Potts model at inverse temperature β on a graph G with maximum degree Δ. Then*

$$\lambda(P_{\mathrm{SW}}) \geq c_{\mathrm{sw}} \lambda(P_{\mathrm{HB}}),$$

where

$$c_{\mathrm{sw}} := c_{\mathrm{sw}}(\Delta, \beta, q) = q^{-1}(qe^{2\beta})^{-2\Delta}. \tag{3.1}$$

This result implies rapid mixing of the Swendsen–Wang dynamics in some new instances. But, since the heat-bath dynamics is typically slow at low temperatures, it is not helpful for large values of β. For this reason we switch to dynamics for a closely related model on the edges of the underlying graph, i.e. the *random-cluster (RC) model*, where it was possible to deduce some lower bounds on the spectral gap at low temperatures

from lower bounds at high temperatures. The random-cluster model consists of the set of subsets of the edges $\Omega_{\mathrm{RC}} = \{A : A \subset E\}$ and the probability measure

$$\mu(A) = \frac{1}{Z}\left(\frac{p}{1-p}\right)^{|A|} q^{c(A)}, \qquad A \subset E,$$

where $c(A)$ denotes the number of connected components of the subgraph (V, A), and $p := 1 - e^{-\beta}$. In fact, the Swendsen–Wang dynamics is based on the tight connection of Potts and random-cluster models. That is, there exists a clever coupling of π and μ, say ν, such that the conditional probabilities, given either a Potts or a random-cluster configuration, are equal to the probability distributions that are involved in the first and the second step of the Swendsen–Wang dynamics, respectively (see Section 2.3).

From this construction it is obvious that the Swendsen–Wang dynamics (with its two steps in reverse order) also defines a Markov chain for the random-cluster model. Additionally, we define the *single-bond* (*SB*) *dynamics*, which is a local Markov chain for the random-cluster model that chooses an edge $e \in E$ of the graph uniformly at random and includes or deletes e from the current random-cluster configuration with a certain probability (see (2.27)). The transition matrices of these two dynamics are denoted as follows:

- $\widetilde{P}_{\mathrm{SW}}$ for the Swendsen–Wang dynamics for the RC model (see (2.22)),
- $\widetilde{P}_{\mathrm{SB}}$ for the single-bond dynamics for the RC model (see (2.27)).

As an easy corollary of the construction of the Swendsen–Wang dynamics we obtain $\lambda(P_{\mathrm{SW}}) = \lambda(\widetilde{P}_{\mathrm{SW}})$ (see Lemma 2.6). Thus, every result on $\widetilde{P}_{\mathrm{SW}}$ immediately yields a result on P_{SW}.

It turns out that Swendsen–Wang and single-bond dynamics can be represented on the joint Potts/random-cluster model (i.e. the model corresponding to the coupling ν) using the same "building blocks", which leads to the second main result of this paper. We prove that rapid mixing of Swendsen–Wang dynamics is equivalent to rapid mixing of single-bond dynamics.

THEOREM 4.8. *Let $\widetilde{P}_{\mathrm{SW}}$ (resp. $\widetilde{P}_{\mathrm{SB}}$) be the transition matrix of the Swendsen–Wang (resp. single-bond) dynamics for the random-cluster model on a graph with $m \geq 3$ edges. Then*

$$\lambda(\widetilde{P}_{\mathrm{SB}}) \leq \lambda(\widetilde{P}_{\mathrm{SW}}) \leq 8m \log m \cdot \lambda(\widetilde{P}_{\mathrm{SB}}).$$

Finally, in Chapter 5 we restrict attention to a special class of graphs, namely *planar graphs*, and, using the notion of *dual graphs*, it will be possible to relate the mixing properties of the Swendsen–Wang dynamics on the original graph to the mixing properties on its dual with a suitable change of the temperature parameter. This is done by proving such a result for the single-bond dynamics and, using the comparison result above, translating it to the Swendsen–Wang dynamics.

For this, assume that the underlying graph is planar, i.e. can be drawn in the plane without intersecting edges. Furthermore, given a random-cluster model on a planar graph G with parameters p and q, we call the random-cluster model on the dual graph $G^\dagger$ with parameters $p^* = \frac{q(1-p)}{p+q(1-p)}$ and q the *dual model*.

THEOREM 5.6. *Let $\widetilde{P}_{\mathrm{SW}}$ (resp. $\widetilde{P}_{\mathrm{SB}}$) be the transition matrix of the Swendsen–Wang (resp. single-bond) dynamics for the random-cluster model on a planar graph G with m edges and let $\widetilde{P}^{\dagger}_{\mathrm{SW}}$ (resp. $\widetilde{P}^{\dagger}_{\mathrm{SB}}$) be the SW (resp. SB) dynamics for the dual model. Then*

$$\lambda(\widetilde{P}_{\mathrm{SB}}) \le q\lambda(\widetilde{P}^{\dagger}_{\mathrm{SB}}),$$
$$\lambda(\widetilde{P}_{\mathrm{SW}}) \le 8qm\log m \cdot \lambda(\widetilde{P}^{\dagger}_{\mathrm{SW}}).$$

Since the dual model of the dual model is the primal random-cluster model (if G is connected), we get the bound in the other direction by applying Theorem 5.6 twice. This theorem allows us to obtain some new results on rapid mixing at low temperatures directly from known results at high temperatures.

Now we turn to applications of the theorems above. But first, note that we do not give a direct analysis of the Markov chains under consideration. Thus, all results on rapid mixing in specific settings rely ultimately on already known mixing results. In fact, the three results given below are based on lower bounds on the spectral gap of the heat-bath dynamics for the Potts model. We refer to Section 2.5 for a collection of all previously known results that are necessary for the analysis in this paper.

The first application, which was the main reason for our study, deals with the two-dimensional square lattice $\mathbb{Z}^2_L$ of side length L. This is the graph $\mathbb{Z}^2_L = (V_{L,2}, E_{L,2})$ with vertex set $V_{L,2} = \{1,\dots,L\}^2 \subset \mathbb{Z}^2$ and edge set $E_{L,2} = \{\{u,v\} \subset V_{L,2} : |u-v| = 1\}$, where $|\cdot|$ denotes the Euclidean norm.

THEOREM 5.9. *Let P_{SW} be the transition matrix of the Swendsen–Wang dynamics for the q-state Potts model on $\mathbb{Z}^2_L$ at inverse temperature β. Let $n = L^2$. Then there exist constants $c_\beta = c_\beta(q), c' > 0$ and $C < \infty$ such that*

$$\lambda(P_{\mathrm{SW}}) \ge c_\beta/n \qquad \text{for } \beta < \beta_c(q),$$
$$\lambda(P_{\mathrm{SW}}) \ge \frac{c_\beta}{n^2\log n} \qquad \text{for } \beta > \beta_c(q),$$
$$\lambda(P_{\mathrm{SW}}) \ge c' n^{-C} \qquad \text{for } q = 2 \text{ and } \beta = \beta_c(2),$$

where $\beta_c(q) = \log(1+\sqrt{q})$.

This result shows rapid mixing of the Swendsen–Wang dynamics for the Potts model on the two-dimensional square lattice $\mathbb{Z}^2_L$ at *all* non-critical temperatures, i.e. at all $\beta \neq \beta_c(q)$, as well as rapid mixing at all temperatures in the case $q = 2$. Note that it was not even known that the SW dynamics mixes rapidly for $\beta < \beta_c$, as it is known for the heat-bath dynamics.

As a byproduct we obtain the following result for the single-bond dynamics for the random-cluster model.

THEOREM 5.10. *Let $\widetilde{P}_{\mathrm{SB}}$ be the transition matrix of the single-bond dynamics for the RC model on $\mathbb{Z}^2_L$ with parameters p and q. Let $m = 2L(L-1) = |E_{L,2}|$. Then there exist constants $c_p = c_p(q), c' > 0$ and $C < \infty$ such that*

$$\lambda(\widetilde{P}_{\mathrm{SB}}) \geq \frac{c_p}{m^2 \log m} \qquad \text{for } p \neq p_c(q),$$
$$\lambda(\widetilde{P}_{\mathrm{SB}}) \geq c' m^{-C} \qquad \text{for } q = 2 \text{ and } p = p_c(2),$$

where $p_c(q) = \frac{\sqrt{q}}{1+\sqrt{q}}$.

For the third application we consider the more general family of planar graphs of bounded maximum degree.

COROLLARY 5.7. *The Swendsen–Wang dynamics for the random-cluster model with parameters p and q on a planar, simple and connected graph G with m edges and maximum degree $\Delta \geq 6$ satisfies*

$$\lambda(\widetilde{P}_{\mathrm{SW}}) \geq \begin{cases} \dfrac{c(1-\varepsilon)}{m} & \text{if } p \leq \dfrac{\varepsilon}{3\sqrt{\Delta-3}}, \\ \dfrac{c(1-\varepsilon)}{m^2 \log m} & \text{if } p \geq 1 - \dfrac{\varepsilon}{q\Delta^\dagger}, \end{cases}$$

for some $c = c(\Delta, p, q) > 0$ and $\varepsilon > 0$, where $\Delta^\dagger$ is the maximum degree of a dual graph of G.

Additionally, we present a similar result for arbitrary graphs of bounded maximum degree. In this case the Swendsen–Wang dynamics is proven to be rapidly mixing if $p \leq \varepsilon/\Delta$ (see Corollary 3.3 and Lemma 2.6). These results enlarge the previously known set of temperatures where rapid mixing is known (see [34] or Theorem 2.18), but lead to a worse bound on the spectral gap.

2. Detailed introduction

In this chapter we provide the necessary definitions and notations. The experienced reader could skip this chapter and visit it when necessary.

First we give an elementary introduction to Markov chains on finite state spaces. See Levin, Peres and Wilmer [39] for more details. Then we define the measure of efficiency of Markov chains which we are concerned with, the spectral gap, and explain the relation of this quantity to the *rate of convergence* of a Markov chain to its stationary distribution, as well as the relation to another quantity, the mixing time. In the subsequent sections we focus on Markov chains for Potts and random-cluster models, and give precise definitions of the models and the Markov chains used. Finally, in Section 2.5, we state several known results on their mixing properties.

2.1. Markov chains. Let Ω be a finite set, P be an $\Omega \times \Omega$-matrix with $P(x,y) \geq 0$ and $\sum_{z\in\Omega} P(x,z) = 1$ for all $x, y \in \Omega$, and $X = (X_t)_{t\in\mathbb{N}}$, $X_t \in \Omega$, be a sequence of Ω-valued random variables defined on a probability space $(\mathfrak{X}, \mathcal{F}, \mathbb{P})$. We call X a *Markov chain on Ω with transition matrix P* if, for all $t \geq 1$ and all $x_0, \ldots, x_t \in \Omega$ such that $\mathbb{P}(X_0 = x_0, \ldots, X_{t-1} = x_{t-1}) > 0$, we have

$$\mathbb{P}(X_t = x_t \mid X_0 = x_0, \ldots, X_{t-1} = x_{t-1}) = \mathbb{P}(X_t = x_t \mid X_{t-1} = x_{t-1}) = P(x_{t-1}, x_t).$$

This is called the *Markov property*. Note that this implies that, for all $x, y \in \Omega$ and $s \in \mathbb{N}$ with $\mathbb{P}(X_s = x) > 0$,

$$P^t(x, y) = \sum_{z \in \Omega} P^{t-1}(x, z)P(z, y) = \mathbb{P}(X_{s+t} = y \mid X_s = x) \quad \text{for all } t \geq 0,$$

which defines the *t-step transition probabilities*. It is clear that the Markov property implies that the knowledge of P and of the distribution of $X = (X_t)_{t \in \mathbb{N}}$ at any time, say 0, determines the distribution of the Markov chain at all future times: Suppose that the distribution of the Markov chain at time 0 is given by some probability mass function η_0, i.e. $\eta_0(x) = \mathbb{P}(X_0 = x)$ for all $x \in \Omega$; then the distribution of the Markov chain at time $t \geq 1$ satisfies

$$\mathbb{P}(X_t = x) = \sum_{z \in \Omega} \eta_0(z) P^t(z, x). \tag{2.1}$$

Therefore, we can simulate one realization of the Markov chain X on Ω with transition matrix P and *initial distribution* η_0 by first generating a state x_0 with respect to the distribution η_0, and then generating successively the state x_t with respect to $P(x_{t-1}, \cdot)$, $t \geq 1$. In particular, if the initial distribution of the Markov chain is concentrated at a single state, i.e. $\eta_0(x) = \mathbb{1}(x = x_0)$ for some $x_0 \in \Omega$, the distributions of the Markov chain take the simple form $\mathbb{P}(X_t = x) = P^t(x_0, x)$. Here, the *indicator function* $\mathbb{1}$ is defined by $\mathbb{1}(x = x_0) = 1$ if $x = x_0$ and $\mathbb{1}(x = x_0) = 0$ otherwise. In general, $\mathbb{1}$ equals 1 if the statement in the parentheses is true and equals 0 otherwise.

From now on we will identify probability mass functions with row-vectors, so that equality (2.1) simplifies to $\mathbb{P}(X_t = x) = \eta_0 P^t(x)$, and for a probability mass function π we define, for $A \subset \Omega$, the probability measure π by

$$\pi(A) := \sum_{x \in A} \pi(x).$$

(By convention, we will use π interchangeably as a measure and a mass function.)

Markov chains are typically used to sample (approximately) from distributions for which direct simulation is not feasible. For this it is necessary that, at least in the limit, the distribution of the Markov chain reaches this distribution, say π, i.e.

$$\pi(x) = \lim_{t \to \infty} \mathbb{P}(X_t = x) \quad \text{for all } x \in \Omega. \tag{2.2}$$

It is in general not guaranteed that such limits exist, but we can present a sufficient condition for existence. We call a Markov chain *irreducible* if for all $x, y \in \Omega$, there exists t such that $P^t(x, y) > 0$, and *aperiodic* if for all $x \in \Omega$, $\gcd\{t \geq 1 : P^t(x, x) > 0\} = 1$, where $\gcd\{J\}$ denotes the greatest common divisor of all elements in $J \subset \mathbb{N}$. Clearly, for an irreducible Markov chain, $P(x, x) > 0$ for any $x \in \Omega$ is a sufficient condition for aperiodicity. All Markov chains in this paper will be irreducible and aperiodic, and we use *ergodic* as an abbreviation for irreducible and aperiodic. It is well-known that the above limit (2.2) exists for aperiodic Markov chains X and, if the Markov chain is also irreducible, the limit distribution π is independent of the initial distribution η_0 and satisfies $\pi(x) > 0$, $\forall x \in \Omega$. Thus, $\pi P = \pi$, which follows obviously from (2.2), is equivalent

to (2.2) whenever X is an ergodic Markov chain with transition matrix P. We call a distribution π with $\pi P = \pi$ a *stationary distribution* of the Markov chain with transition matrix P.

Another important concept is the reversibility of Markov chains. A Markov chain with transition matrix P is *reversible* (or satisfies *detailed balance*) *with respect to* π if for all $x, y \in \Omega$,

$$\pi(x)P(x,y) = \pi(y)P(y,x).$$

It is not hard to prove that, under this condition, π is a stationary distribution of P. Finally, we call a Markov chain *lazy*, if its transition matrix P satisfies $P(x,x) \geq \frac{1}{2}$, $\forall x \in \Omega$. Obviously, lazy Markov chains are aperiodic. If X is a Markov chain with transition matrix P, and $\widetilde{X}$ is a Markov chain with transition matrix Q, such that $Q(x,y) = \frac{1}{2}P(x,y)$ for all $x \neq y$, $x, y \in \Omega$, and $\mathbb{P}(\widetilde{X}_0 = x) = \mathbb{P}(X_0 = x)$, $\forall x \in \Omega$, then we say that $\widetilde{X}$ is the *lazy version of* X. Necessarily, $Q(x,x) = \frac{1}{2} + \frac{1}{2}P(x,x)$ for all $x \in \Omega$.

Throughout this paper we will refer to properties of a Markov chain with transition matrix P as properties of P, unless they depend on more than P. For example we say for an ergodic Markov chain with transition matrix P that is reversible with respect to π, that P is ergodic and reversible with respect to π.

2.2. Spectral gap and mixing time. In the following we want to estimate the efficiency of Markov chains for approximate sampling from their stationary distribution. For more details and aspects of the convergence of Markov chains to their stationary distribution see Levin, Peres and Wilmer [39].

To quantify this efficiency we first define the *total variation distance* of two distributions ν and π on Ω by

$$\|\nu - \pi\|_{\mathrm{TV}} := \frac{1}{2}\sum_{x\in\Omega} |\nu(x) - \pi(x)| = \max_{A\subset\Omega} |\nu(A) - \pi(A)|.$$

See [39, Prop. 4.2] for the second equality. Using this as a metric on the set of all probability measures on Ω it is natural to ask how fast the convergence in (2.2) takes place, i.e. how fast the distribution of a Markov chain for increasing time t converges to its stationary distribution. The next statement, which is called the *Convergence Theorem*, provides a more quantitative version of (2.2) (see e.g. [39, Theorem 4.9]).

THEOREM 2.1. *Let P be ergodic with stationary distribution π. Then there exist constants $\alpha \in [0,1)$ and $C > 0$ such that*

$$\max_{x\in\Omega} \|P^t(x,\cdot) - \pi\|_{\mathrm{TV}} \leq C\alpha^t. \tag{2.3}$$

This theorem shows that, for every $x \in \Omega$, the distribution of a Markov chain with transition matrix P and initial distribution concentrated at x converges exponentially fast (in total variation) to its stationary distribution. In fact, the worst initial distribution, i.e. the distribution η_0 that maximizes $\|\eta_0 P^t(\cdot) - \pi\|_{\mathrm{TV}}$, is concentrated at a single state. To see this, note that for a distribution η_0 on Ω,

$$\|\eta_0 P^t(\cdot) - \pi\|_{\mathrm{TV}} = \Big\| \sum_{x\in\Omega} \eta_0(x) P^t(x,\cdot) - \pi \Big\|_{\mathrm{TV}} = \Big\| \sum_{x\in\Omega} \eta_0(x)(P^t(x,\cdot) - \pi) \Big\|_{\mathrm{TV}}$$
$$\leq \sum_{x\in\Omega} \eta_0(x) \|P^t(x,\cdot) - \pi\|_{\mathrm{TV}} \leq \max_{x\in\Omega} \|P^t(x,\cdot) - \pi\|_{\mathrm{TV}}.$$

Thus, it is enough to consider the maximum as in Theorem 2.1 to get a statement for arbitrary initial distributions.

For given $\varepsilon > 0$, we are interested in the minimal time t such that the total variation distance between the distribution of our Markov chain at time t and its stationary distribution is at most ε, independent of the initial distribution. This t is called the *ε-mixing time* and is defined by

$$t_{\mathrm{mix}}(P,\varepsilon) := \min\Big\{ t \geq 0 : \max_{x\in\Omega} \|P^t(x,\cdot) - \pi\|_{\mathrm{TV}} \leq \varepsilon \Big\}.$$

Using the fact that $t_{\mathrm{mix}}(P,\varepsilon) \leq \lceil \log(\varepsilon^{-1}) \rceil t_{\mathrm{mix}}(P, 1/2e)$ (see e.g. [39, p. 55]), it is enough to consider the *mixing time* $t_{\mathrm{mix}}(P) := t_{\mathrm{mix}}(P, 1/2e)$. Unless otherwise stated, log denotes the natural logarithm. Obviously, bounds on t_{mix} imply (and follow from) bounds on the optimal, i.e. smallest, constants C and α in Theorem 2.1. To be precise, if we know some constants C and α for which (2.3) holds, we obtain $t_{\mathrm{mix}}(P) \leq \log(2eC) \cdot \log(1/\alpha)^{-1}$. For the reverse direction note that

$$\max_{x\in\Omega} \|P^t(x,\cdot) - \pi\|_{\mathrm{TV}} \leq e^{-\lfloor t/t_{\mathrm{mix}} \rfloor} \leq e^{1 - t/t_{\mathrm{mix}}}$$

with $t_{\mathrm{mix}} = t_{\mathrm{mix}}(P)$ (see [39, eq. (4.34)]).

In this paper we are concerned with bounding the optimal constant α in (2.3) and, in particular, comparing the optimal constants for different Markov chains. We will see (Lemma 2.3) that this also implies bounds on the mixing time.

For this we introduce the spectral gap of a Markov chain. Let P be the transition matrix of a Markov chain on Ω that is ergodic and reversible with respect to π. We regard P as an operator that maps functions $f : \Omega \to \mathbb{R}$ to functions by

$$Pf(x) := \sum_{y\in\Omega} P(x,y) f(y). \tag{2.4}$$

Such an operator is called the *Markov operator* that corresponds to P, and we will use the same notation for the Markov operator and its corresponding transition matrix. Note that $Pf(x)$ is the expectation of f with respect to the distribution $P(x,\cdot)$, i.e. the distribution of the Markov chain with transition matrix P and initial distribution concentrated at $x \in \Omega$ after one step. If we consider a function f on Ω as an element of $\mathbb{R}^\Omega$ (i.e. a column vector), then Pf is simply matrix multiplication. Additionally, we endow the function space $\mathbb{R}^\Omega$ with the inner product

$$\langle f, g \rangle_\pi := \sum_{x\in\Omega} f(x) g(x) \pi(x), \qquad f, g \in \mathbb{R}^\Omega, \tag{2.5}$$

and denote by $L_2(\pi)$ the inner product space (or Hilbert space) that consists of $\mathbb{R}^\Omega$ with the inner product $\langle \cdot, \cdot \rangle_\pi$. In particular, the norm in $L_2(\pi)$ is given by

$$\|f\|_\pi^2 := \langle f, f \rangle_\pi = \sum_{x\in\Omega} f(x)^2 \pi(x), \qquad f \in \mathbb{R}^\Omega.$$

Using the inner product (2.5) we define the *adjoint operator* P^* of P as the (unique) operator that satisfies $\langle f, Pg\rangle_\pi = \langle P^* f, g\rangle_\pi$ for all $f, g \in L_2(\pi)$. Then P^* is also a Markov operator and the corresponding transition matrix is

$$P^*(x,y) = \frac{\pi(y)}{\pi(x)} P(y,x), \qquad x, y \in \Omega. \tag{2.6}$$

Since P is reversible with respect to π, we obtain $P = P^*$. Hence, P defines a self-adjoint operator. This implies that P has only real eigenvalues, i.e. real numbers ξ with $Pf = \xi f$ for some $0 \neq f \in \mathbb{R}^\Omega$ (see e.g. [36, Thm. 9.1-1]). By the ergodicity of P, these eigenvalues $\{\xi_i\}$ satisfy $-1 < \xi_i \leq 1$; additionally, $Pf = f$ if and only if f is constant (see [39, Lemma 12.1]).

We define the (absolute) *spectral gap* of P by

$$\lambda(P) := 1 - \max\{|\xi| : \xi \text{ is an eigenvalue of } P,\, \xi \neq 1\}.$$

It is well-known that the spectral gap can be written in terms of norms of the Markov operator P. For this we define the spectral norm (or simply *operator norm*) of an operator P by

$$\|P\|_\pi := \|P\|_{L_2(\pi)\to L_2(\pi)} = \max_{\|f\|_\pi = 1} \|Pf\|_\pi. \tag{2.7}$$

We use $\|\cdot\|_\pi$ interchangeably for functions and operators, because it will be clear from the context which norm is used. For a self-adjoint P the operator norm $\|P\|_\pi$ equals the largest eigenvalue of P in absolute value. To give a representation of the spectral gap we also need the operator S_π that is defined by $S_\pi f = \langle f, 1\rangle_\pi$. This (Markov) operator corresponds to the (transition) matrix $S_\pi(x,y) = \pi(y)$ for $x, y \in \Omega$. Obviously, S_π has only the eigenvalues 1 and 0, and the eigenspace to eigenvalue 0 is $\{f \in L_2(\pi) : \langle f, 1\rangle_\pi = 0\}$, which is also the union of all eigenspaces of P for eigenvalues different from 1. Thus,

$$\lambda(P) = 1 - \|P - S_\pi\|_\pi. \tag{2.8}$$

As stated above, the spectral gap and the speed of convergence in Theorem 2.1 are closely related. The next lemma (see [39, Cor. 12.6]) demonstrates this relation.

Lemma 2.2. *Let P be ergodic and reversible with respect to π. Then*

$$\lim_{t\to\infty} \Big(\max_{x\in\Omega} \|P^t(x,\cdot) - \pi\|_{\mathrm{TV}}\Big)^{1/t} = 1 - \lambda(P).$$

This shows that $1 - \lambda(P)$ is the optimal constant α in (2.3). One may hope that this asymptotic equality holds, at least approximately, also non-asymptotically, i.e. that there exist constants $c, C > 0$ such that

$$c(1-\lambda(P))^t \leq \max_{x\in\Omega} \|P^t(x,\cdot) - \pi\|_{\mathrm{TV}} \leq C(1-\lambda(P))^t$$

for all $t \in \mathbb{N}$, and in fact, this inequality holds with constants $c = \frac{1}{2}$ [39, (12.13)] and $C^{-1} = \min_{x\in\Omega} \pi(x)$ [39, (12.11)]. Plugging this into the definition of the mixing time we deduce that mixing time and spectral gap of a Markov chain (on finite state spaces) satisfy the following inequality (see e.g. [39, Theorems 12.3 & 12.4]).

LEMMA 2.3. *Let P be the transition matrix of a reversible, ergodic Markov chain with state space Ω and stationary distribution π. Then*

$$\lambda(P)^{-1} - 1 \leq t_{\mathrm{mix}}(P) \leq \log\left(\frac{2e}{\pi_{\min}}\right)\lambda(P)^{-1},$$

where $\pi_{\min} := \min_{x\in\Omega} \pi(x)$.

We are interested in the spectral gap of specific Markov chains and, in particular, in the dependence on the size of the state space if the Markov chain can be defined analogously on an unbounded family of state spaces. Thus, if we consider an *indexed family* of state spaces $\{\Omega_n\}_{n\in\mathbb{N}}$ with a corresponding family of transition matrices $\{P_n\}_{n\in\mathbb{N}}$, we say that the Markov chain is *rapidly mixing* for the given family if $\lambda(P_n)^{-1} \leq c\log(|\Omega_n|)^C$ for all $n \in \mathbb{N}$ and some $c, C < \infty$.

REMARK 2.4. The main reason for sampling from a given probability distribution π on Ω is the approximation of expectations $S_\pi f = \langle f, 1\rangle_\pi$ for certain functions $f \in \mathbb{R}^\Omega$. This is frequently done e.g. in statistical physics to deepen the understanding of the underlying model. If exact sampling from π is not feasible, then this often can be done by *Markov chain Monte Carlo methods* that can be described as follows. Choose an initial state $x_0 \in \Omega$ (deterministically or by some distribution), then simulate k steps of the Markov chain with transition matrix P to obtain a sequence $x_1, \dots, x_k \in \Omega$ and, finally, take the average $A_k(f) := \frac{1}{k+1}\sum_{i=0}^{k} f(x_i)$. Usually it is better to omit several states at the beginning of the sequence to improve the performance (see e.g. Rudolf [62]). The number of these omitted states is called *burn-in* in the literature. We know from the *Ergodic Theorem* [39, Thm. 4.16] that $\lim_{k\to\infty} A_k(f) = S_\pi f$ almost surely, whenever P is irreducible and has stationary distribution π. In this context the spectral gap plays an important role in bounding the number k of steps of the Markov chain that are necessary to achieve a prescribed error $\varepsilon > 0$. For a bound on k for the probabilistic error criterion that depends on the spectral gap and the ε-mixing time see Levin et al. [39, Sec. 12.6]. Another bound that deals with the mean square error and depends only on the spectral gap is given in [60, 61]; see also Novak and Woźniakowski [54] for some context and results in a more general setting. Overall, one can say that the existence of a rapidly mixing Markov chain for a family of state spaces $\{\Omega_n\}_{n\in\mathbb{N}}$ with corresponding measures $\{\pi_n\}_{n\in\mathbb{N}}$ leads to an efficient algorithm for the approximation of expectations of functions (with respect to π_n) defined on Ω_n, i.e. an algorithm that needs time proportional to ε^{-2} times a polynomial in $\log(|\Omega_n|)$ (times the variance of the considered function).

To finish this section we present a simple technique to compare the spectral gaps of two Markov chains on the same state space, but with possibly different stationary distributions. This result is well-known (see e.g. [15] or [39]), but since it is used several times in this paper we present its proof here. See also Dyer, Goldberg, Jerrum and Martin [21] for a survey on more general techniques for comparison of Markov chains.

LEMMA 2.5. *Suppose P_1 (resp. P_2) is an ergodic and reversible transition matrix with stationary distribution π_1 (resp. π_2) on Ω. If there exist constants $a, A > 0$ such that*

$$\frac{\pi_1(x)P_1(x,y)}{\pi_2(x)P_2(x,y)} \leq A \quad \text{and} \quad \frac{\pi_1(x)}{\pi_2(x)} \geq a$$

for all $x, y \in \Omega$, then

$$\lambda(P_1) \leq \frac{A}{a}\lambda(P_2).$$

If P_2 has only non-negative eigenvalues it is enough to verify the conditions for $x \neq y$.

Proof. For $f \in \mathbb{R}^\Omega$ define

$$\mathcal{E}_1(f) := \langle (I - P_1)f, f\rangle_{\pi_1} \quad \text{and} \quad \mathcal{F}_1(f) := \langle (I + P_1)f, f\rangle_{\pi_1}$$

with the identity $If := f$. Equivalently we define $\mathcal{E}_2$ and $\mathcal{F}_2$ for P_2. By reversibility and ergodicity, P_1 has only real eigenvalues

$$1 = \xi_1(P_1) > \xi_2(P_1) \geq \cdots \geq \xi_{|\Omega|}(P_1) > -1,$$

i.e. $\lambda(P_1) = \min\{1 - \xi_2(P_1), 1 + \xi_{|\Omega|}(P_1)\}$. Thus, using the "min-max characterization" of the eigenvalues [33, Thms. 4.2.2 & 4.2.11] and the fact that $\xi_1(P_1)$ corresponds to the constant eigenfunction, we obtain

$$1 - \xi_2(P_1) = 1 - \max_{\substack{f\neq 0:\\ \langle f,1\rangle_{\pi_1}=0}} \frac{\langle Pf, f\rangle_{\pi_1}}{\langle f, f\rangle_{\pi_1}} = \min_{\substack{f\neq 0:\\ \langle f,1\rangle_{\pi_1}=0}} \frac{\mathcal{E}_1(f)}{\langle f, f\rangle_{\pi_1}} = \min_{\substack{f\neq 0:\\ \langle f,1\rangle_{\pi_1}=0}} \frac{\mathcal{E}_1(f)}{\mathrm{Var}_{\pi_1}(f)}$$

with $\mathrm{Var}_{\pi_1}(g) = \langle g, g\rangle_{\pi_1} - \langle g, 1\rangle^2_{\pi_1}$ for $g \in \mathbb{R}^\Omega$. Noting that $\mathcal{E}_1(f) = \mathcal{E}_1(f + c)$ and $\mathrm{Var}_{\pi_1}(f) = \mathrm{Var}_{\pi_1}(f + c)$ for every $c \in \mathbb{R}$, we get

$$1 - \xi_2(P_1) = \min_{\substack{f\in\mathbb{R}^\Omega:\\ \mathrm{Var}_{\pi_1}(f)\neq 0}} \frac{\mathcal{E}_1(f)}{\mathrm{Var}_{\pi_1}(f)},$$

where $\mathrm{Var}_{\pi_1}(f) \neq 0$ iff f is not constant. Equivalently,

$$1 + \xi_{|\Omega|}(P_1) = 1 + \min_{f\neq 0} \frac{\langle Pf, f\rangle_{\pi_1}}{\langle f, f\rangle_{\pi_1}} = \min_{f\neq 0} \frac{\mathcal{F}_1(f)}{\langle f, f\rangle_{\pi_1}}.$$

It is easy to check (using reversibility) that

$$\mathcal{E}_1(f) = \frac{1}{2}\sum_{x,y\in\Omega} (f(x) - f(y))^2 \pi_1(x)P_1(x, y),$$

$$\mathcal{F}_1(f) = \frac{1}{2}\sum_{x,y\in\Omega} (f(x) + f(y))^2 \pi_1(x)P_1(x, y).$$

It follows from the first assumption of this lemma that $\mathcal{E}_1(f) \leq A\mathcal{E}_2(f)$ and $\mathcal{F}_1(f) \leq A\mathcal{F}_2(f)$ for every $f \in \mathbb{R}^\Omega$. The second assumption implies $\langle f, f\rangle_{\pi_1} \geq a\langle f, f\rangle_{\pi_2}$ and hence $1 + \xi_{|\Omega|}(P_1) \leq (A/a)(1 + \xi_{|\Omega|}(P_2))$. Additionally, for $f \in \mathbb{R}^\Omega$ with $\langle f, 1\rangle_{\pi_1} = 0$, we obtain

$$\mathrm{Var}_{\pi_2}(f) = \langle f, f\rangle_{\pi_2} - \langle f, 1\rangle^2_{\pi_2} \leq \langle f, f\rangle_{\pi_2} \leq \frac{1}{a}\langle f, f\rangle_{\pi_1} = \frac{1}{a}\mathrm{Var}_{\pi_1}(f),$$

which implies $\mathrm{Var}_{\pi_1}(f) \geq a\mathrm{Var}_{\pi_2}(f)$ for every $f \in \mathbb{R}^\Omega$ and therefore $1 - \xi_2(P_1) \leq (A/a)(1 - \xi_2(P_2))$. Note that if P_2 has only non-negative eigenvalues it is enough to compare $\mathcal{E}_1$ and $\mathcal{E}_2$, since $\lambda(P_2) = 1 - \xi_2(P_2)$. Both do not depend on the diagonal elements of P_1 and P_2. This proves the claim. ∎

2.3. The models. In this section we introduce the models that we study in this paper. Although we are mainly interested in sampling from the Potts model, we additionally need the closely related random-cluster and Fortuin–Kasteleyn–Edwards–Sokal (FKES) models. Since all these models are defined on an underlying graph, we begin with providing some general graph terminology; see e.g. Diestel [17] or Mohar and Thomassen [53] for a more comprehensive introduction to graph theory.

A *graph* G is a pair (V, E), where V is the finite set of *vertices* and E is the set of *edges*, together with a function φ that assigns to each edge $e \in E$ a set of at most two vertices, i.e. $\varphi(e) = \{u, v\}$ for some $u, v \in V$, which are called its *endvertices*. We denote the endvertices of an edge $e \in E$ by $e^{(1)}$ and $e^{(2)}$, i.e. $\varphi(e) = \{e^{(1)}, e^{(2)}\}$. Let $\varphi(E) = \{\varphi(e) : e \in E\}$. We say that u and v are *neighbors* in G if $\{u, v\} \in \varphi(E)$. Furthermore, u and v are called *connected*, if there exist vertices $v_0, \dots, v_n \in V$ such that $v_0 = u$, $v_n = v$ and $\{v_{i-1}, v_i\} \in \varphi(E)$, $i = 1, \dots, n$. We write $u \leftrightarrow v$ if u and v are connected (in G). Now suppose we have two graphs, $G = (V_G, E_G, \varphi_G)$ and $H = (V_H, E_H, \varphi_H)$. We say that H is a *subgraph* of G if $V_H \subset V_G$, $E_H \subset E_G$ and $\varphi_H(e) = \varphi_G(e)$ for all $e \in E_H$. If additionally $V_H = V_G$, we say H is a *spanning subgraph* of G. Let $A \subset E$ be a subset of the edges of the graph $G = (V, E, \varphi)$. Then $G_A := (V, A, \varphi)$ is a spanning subgraph of G and we write $u \overset{A}{\longleftrightarrow} v$ if $u, v \in V$ are connected in G_A, i.e. there exist vertices $v_0, \dots, v_n \in V$ such that $v_0 = u$, $v_n = v$ and $\{v_{i-1}, v_i\} \in \varphi(A)$, $i = 1, \dots, n$. Clearly, $\overset{A}{\longleftrightarrow}$ defines an equivalence relation on V for every $A \subset E$; its equivalence classes are called *connected components* of G_A. Two distinct edges $e_1, e_2 \in E$ that have the same endvertices, i.e. $\varphi(e_1) = \varphi(e_2)$, are said to be *parallel*, and edges with equal endvertices, i.e. $e \in E$ with $e^{(1)} = e^{(2)}$ (or $|\varphi(e)| = 1$), are called *loops*.

Most graphs of this paper are *simple*, i.e. contain no parallel edges. (Some authors use the term "graph" for simple graphs, and "multigraph" for graphs as defined above.) The advantage of simple graphs is that $|\varphi(E)| = |E|$ and thus φ is a bijection between E and $\varphi(E)$. In this case we identify $\varphi(E)$ with E and write $G = (V, E)$ for the graph (V, E, φ), as well as $e = \{u, v\}$ for $\varphi(e) = \{u, v\}$. With a slight abuse of notation we omit φ also if the graph is not simple, i.e. we write $G = (V, E)$ for (V, E, φ), and make the convention that $\{u, v\} = e$ (resp. $\{u, v\} \in E$) simply means $\{u, v\} = \varphi(e)$ (resp. $\{u, v\} \in \varphi(E)$). (Note that we will use this notation only in one direction, i.e. given $e \in E$ we obtain a unique $\{u, v\} \subset V$ with $\{u, v\} = e$. The other way cannot be done uniquely if the graph contains parallel edges.)

As example graphs one can have in mind the two-dimensional *square lattice* $\mathbb{Z}_L^2$ of side length L, i.e. the graph $\mathbb{Z}_L^2 = (V_{L,2}, E_{L,2})$ with vertex set $V_{L,2} = \{1, \dots, L\}^2 \subset \mathbb{Z}^2$ and edge set $E_{L,2} = \{\{u, v\} \subset V_{L,2} : |u - v| = 1\}$, where $|\cdot|$ denotes the ℓ_2 norm (see Figure 1, left), and a *tree*, which is a graph $T = (V, E)$ with $|E| = |V| - 1$ and exactly one connected component (see Figure 1, right). Obviously, deletion of any edge of T increases the number of connected components.

For the further analysis we need the following graph quantities. First of all, we will refer to the number of vertices of a graph, i.e. $|V|$ for $G = (V, E)$, as the *size* of the graph. The *degree* of a vertex v in $G = (V, E)$, i.e. $\big|\{e \in E : v \in e\}\big|$, is denoted by $\deg_G(v)$, and

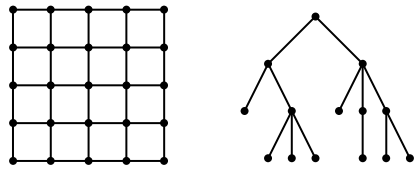

Fig. 1. The graph $\mathbb{Z}_5^2$ (left) and a tree (right)

we use $\Delta(G)$ for the *maximum degree* of G, i.e.

$$\Delta(G) = \max_{v \in V} \deg_G(v).$$

For example, $\Delta(\mathbb{Z}_L^2) = 4$ for every $L \geq 3$. Additionally, for a graph $G = (V, E)$ we denote by $c(G_A)$ the *number of connected components* of the graph $G_A = (V, A)$, $A \subset E$. Thus, $c(G_A)$ is the number of equivalence classes in V with respect to $\overset{A}{\longleftrightarrow}$. If the graph is fixed we simply write $c(A)$ for $c(G_A)$. Note that, if $T = (V, E)$ is a tree, $c(T_A) = |V| - |A|$ for all $A \subset E$.

Now we introduce the models under consideration. For this, fix a graph $G = (V, E)$, a natural number $q \geq 1$ and a real number $\beta \geq 0$. Typically, β is called the *inverse temperature*. The *q-state Potts model* on G is defined as the set of possible *configurations* $\Omega_{\mathrm{P}} := \Omega_{\mathrm{P}}(G) = [q]^V$, where $[q]{:=}\{1, \ldots, q\}$ is the set of *colors* (or spins), together with the probability measure

$$\pi(\sigma) := \pi_{\beta,q}^G(\sigma) = \frac{1}{Z(G, \beta, q)} e^{\beta |E(\sigma)|}, \quad \sigma \in \Omega_{\mathrm{P}}, \tag{2.9}$$

where

$$E(\sigma) := \{e \in E : \sigma(u) = \sigma(v) \text{ for } \{u, v\} = e\} \tag{2.10}$$

is the set of edges with equally colored endvertices. The *normalization constant* (also called *partition function*) Z is given by

$$Z(G, \beta, q) := \sum_{\sigma \in \Omega_{\mathrm{P}}(G)} e^{\beta |E(\sigma)|}.$$

This measure is called the *Potts measure* (or *Boltzmann distribution*) and if $q = 2$ we call the Potts model the *Ising model*. The inverse temperature β determines the interaction strength between neighboring vertices in the graph and since β is non-negative, configurations with more equally colored neighbors have a larger weight in π. In particular, at infinite temperature, i.e. $\beta = 0$, the measure π is the uniform distribution on Ω_{P} and the larger β the more weight goes to the (almost) constant configurations.

A closely related model is the *random-cluster model* (also known as the *FK model*), that was introduced by Fortuin and Kasteleyn [24]. It is defined on the graph $G = (V, E)$ by its state space $\Omega_{\mathrm{RC}} = \{A : A \subseteq E\}$ and the *random-cluster* (*RC*) *measure*

$$\mu(A) := \mu_{p,q}^G(A) = \frac{1}{Z\big(G, \log\big(\frac{1}{1-p}\big), q\big)} \left(\frac{p}{1-p}\right)^{|A|} q^{c(A)}, \quad A \subset E, \tag{2.11}$$

where $p \in [0, 1]$, $c(A)$ is the number of connected components in the graph (V, A) and

$Z(\cdot,\cdot,\cdot)$ is the same normalization constant as for the Potts model [29, Thm. 1.10] (see (2.13) and (2.14)). For a historical treatment and related topics see Grimmett [29].

To describe the connection of Potts and random-cluster models we state a coupling of the corresponding measures that is due to Edwards and Sokal [23]. This coupling leads to the third model that will be considered, namely the *Fortuin–Kasteleyn–Edwards–Sokal* (*FKES*) model (or simply the *joint model*). For this let $p = 1 - e^{-\beta}$, as we will assume henceforth. The joint model is defined on $\Omega_{\mathrm{J}} := \Omega_{\mathrm{P}} \times \Omega_{\mathrm{RC}}$ by the *FKES measure*

$$\nu(\sigma, A) := \nu_{p,q}^{G}(\sigma, A) = \frac{1}{Z\big(G, \log\big(\frac{1}{1-p}\big), q\big)} \left(\frac{p}{1-p}\right)^{|A|} \mathbb{1}(A \subset E(\sigma)) \tag{2.12}$$

for $(\sigma, A) \in \Omega_{\mathrm{J}}$. Again, the normalization constant $Z(\cdot,\cdot,\cdot)$ is the same as for the Potts measure [29, Thm. 1.10]. To see this, note that

$$A \subset E(\sigma) \;\Leftrightarrow\; \forall \{u, v\} \in A : \sigma(u) = \sigma(v),$$

which implies that σ is constant on each connected component of (V, A). Hence, for a given $A \subset E$ there are exactly $q^{c(A)}$ configurations $\sigma \in \Omega_{\mathrm{P}}$ with $\nu(\sigma, A) > 0$. Using this, for $\sigma \in \Omega_{\mathrm{P}}$ and $A \subset E$ we obtain

$$\begin{aligned} \sum_{\tau \in \Omega_{\mathrm{P}}} \left(\frac{p}{1-p}\right)^{|A|} \mathbb{1}(A \subset E(\tau)) &= \left(\frac{p}{1-p}\right)^{|A|} |\{\tau \in \Omega_{\mathrm{P}} : A \subset E(\tau)\}| \\ &= \left(\frac{p}{1-p}\right)^{|A|} q^{c(A)}, \end{aligned} \tag{2.13}$$

and since $p = 1 - e^{-\beta}$,

$$\begin{aligned} \sum_{B \in \Omega_{\mathrm{RC}}} \left(\frac{p}{1-p}\right)^{|B|} \mathbb{1}(B \subset E(\sigma)) &= \sum_{k=0}^{|E(\sigma)|} \binom{|E(\sigma)|}{k} \left(\frac{p}{1-p}\right)^{k} \\ &= \left(1 + \frac{p}{1-p}\right)^{|E(\sigma)|} = e^{\beta |E(\sigma)|}. \end{aligned} \tag{2.14}$$

Summing over $A \in \Omega_{\mathrm{RC}}$ in (2.13) and over $\sigma \in \Omega_{\mathrm{P}}$ in (2.14) proves that the normalization constants of π, μ and ν are equal. Another fact that can be deduced from (2.13) and (2.14) is that the marginal distributions of ν are exactly π and μ, respectively (see [23] or [29]). This means that for all $\sigma \in \Omega_{\mathrm{P}}$ and $A \subset E$,

$$\pi(\sigma) = \sum_{A \subset E} \nu(\sigma, A) = \nu(\sigma, \Omega_{\mathrm{RC}}) \quad \text{and} \quad \mu(A) = \sum_{\sigma \in \Omega_{\mathrm{P}}} \nu(\sigma, A) = \nu(\Omega_{\mathrm{P}}, A).$$

We define the conditional probability of $\sigma \in \Omega_{\mathrm{P}}$ with respect to ν given $A \subset E$ by

$$\nu(\sigma | A) := \begin{cases} \nu(\sigma, A)/\nu(\Omega_{\mathrm{P}}, A) & \text{if } (\sigma, A) \in \Omega_{\mathrm{J}}, \\ 0 & \text{otherwise.} \end{cases} \tag{2.15}$$

For $\nu(A \,|\, \sigma)$ replace $\nu(\Omega_{\mathrm{P}}, A)$ by $\nu(\sigma, \Omega_{\mathrm{RC}})$. We obtain

$$\sum_{A \in \Omega_{\mathrm{RC}}} \mu(A)\nu(\sigma \,|\, A) = \pi(\sigma) \quad \text{and} \quad \sum_{\sigma \in \Omega_{\mathrm{P}}} \pi(\sigma)\nu(A \,|\, \sigma) = \mu(A).$$

This can be interpreted in the following way. Assume that we can simulate a random variable $X \in \Omega_{\mathrm{RC}}$ that is distributed with respect to μ. The random variable $Y \in \Omega_{\mathrm{P}}$ obtained by sampling from $\nu(\cdot \,|\, X)$ is then distributed with respect to π. In fact, the conditional probabilities take a rather simple form, namely $\nu(\sigma|A) = q^{-c(A)}\mathbb{1}(A \subset E(\sigma))$ and $\nu(A \,|\, \sigma) = p^{|A|}(1-p)^{|E(\sigma)|-|A|}\mathbb{1}(A \subset E(\sigma))$. Thus, given a RC configuration $X \sim \mu$, i.e. one that is distributed with respect to μ, we can generate a Potts configuration $Y \sim \pi$ by assigning a random color from $\{1, \dots, q\}$ independently to each connected component of the graph (V, X). For the reverse way, given $Y \sim \pi$, include all edges $e = \{e^{(1)}, e^{(2)}\} \in E$ with $Y(e^{(1)}) = Y(e^{(2)})$ in X with probability p. Hence, if the Potts and the RC model are defined on the same graph G and $p = 1 - e^{-\beta}$, then the possibility of efficient sampling from either model enables also efficient sampling from the other.

We will see in the following section that the non-local Markov chain considered, namely the Swendsen–Wang dynamics, is based on this connection of the Potts and random-cluster models.

2.4. The dynamics. This section is devoted to the definition of the Markov chains that will be used to sample approximately from the Potts and the random-cluster model. In Chapters 3 and 4 we will compare the spectral gaps of these Markov chains and, as a consequence, obtain new results on the mixing properties of the Swendsen–Wang and the single-bond dynamics. Recall that we have fixed a graph $G = (V, E)$, some $q \in \mathbb{N}$, a real $\beta \geq 0$ and $p = 1 - e^{-\beta}$. We distinguish between dynamics for the Potts model and dynamics for the random-cluster model, and we will propose one *local* and one *non-local* Markov chain in both cases. "Local" means that the Markov chain changes the current state in one step only locally. In other words: let P (resp. $\widetilde{P}$) be the transition matrix of a Markov chain on Ω_{P} (resp. Ω_{RC}) with stationary distribution π (resp. μ); then we say that P (resp. $\widetilde{P}$), or its corresponding Markov chain, is *local* if $|\{v \in V : \sigma(v) \neq \tau(v)\}| \leq 1$ for all $\sigma, \tau \in \Omega_{\mathrm{P}}$ with $P(\sigma, \tau) > 0$ (resp. $|(A \setminus B) \cup (B \setminus A)| \leq 1$ for all $A, B \in \Omega_{\mathrm{RC}}$ with $\widetilde{P}(A, B) > 0$). Thus, local Markov chains update only one vertex (resp. edge) of the current configuration per step.

Another notion of locality is that the transition probabilities, e.g. $P(\sigma, \tau)$ for $\sigma, \tau \in \Omega_{\mathrm{P}}$, depend only locally on the states σ and τ. This is called the *finite range interaction property* of the transition probabilities (see e.g. [45]), and means (in the case of the Potts model) that for every $\sigma, \tau \in \Omega_{\mathrm{P}}$ with $P(\sigma, \tau) > 0$ there exists a subset of the vertices $V_0 \subset V$, with $|V_0|$ "not too large", such that $P(\sigma, \tau) = P(\sigma', \tau')$ for $\sigma', \tau' \in \Omega_{\mathrm{P}}$ with $\sigma(v) = \sigma'(v)$ and $\tau(v) = \tau'(v)$ for all $v \in V_0$. An analogous formulation can be found for Markov chains for the RC model. This property is especially interesting for computational reasons, because it enables efficient computation of the transition probabilities and therefore efficient simulation of the Markov chain. From the local dynamics of this section only the one for the Potts model satisfies the finite range interaction property.

2.4.1. Dynamics for the Potts model

The heat-bath dynamics. We begin with the definition of the local Markov chain for the Potts model. This Markov chain, namely the *heat-bath* (*HB*) *dynamics* (or *Glauber dynamics*), was introduced by Glauber [27] in 1963 for the Ising model and, since then,

has become the most studied Markov chain for the q-state Potts model (especially for $q = 2$). The one-step transitions can be described as follows. Suppose that the current state of the Markov chain is $\sigma \in \Omega_{\mathrm{P}}$. Then:

(HB1) Choose a vertex $v \in V$ uniformly at random.
(HB2) Let $\Omega_{\sigma,v} := \{\tau \in \Omega_{\mathrm{P}} : \tau(u) = \sigma(u), \forall u \neq v\}$ and choose the next state of the Markov chain, say τ, with respect to $\pi(\cdot \,|\, \Omega_{\sigma,v})$, that is, the conditional probability (with respect to π) given that τ differs from σ only at v.

For $\sigma \in \Omega_{\mathrm{P}}$, $k \in [q]$ and $u, v \in V$, let

$$\sigma^{v,k}(u) := \begin{cases} \sigma(u) & \text{if } u \neq v, \\ k & \text{if } u = v, \end{cases} \tag{2.16}$$

be the configurations that differ from σ at most at $v \in V$ and let

$$d_{v,k}(\sigma) := |\{e \in E : \sigma(u) = k \text{ for } \{u, v\} = e\}|$$

be the number of edges in G connecting v to vertices with color k in σ. It is easy to see that the set $\Omega_{\sigma,v}$ from (HB2) can be written as $\Omega_{\sigma,v} = \bigcup_{l=1}^{q} \{\sigma^{v,l}\}$, and that the conditional probabilities satisfy

$$\pi(\sigma^{v,k} \,|\, \Omega_{\sigma,v}) = \frac{\pi(\sigma^{v,k})}{\sum_{l=1}^{q} \pi(\sigma^{v,l})} = \frac{e^{\beta d_{v,k}(\sigma)}}{\sum_{l=1}^{q} e^{\beta d_{v,l}(\sigma)}}, \tag{2.17}$$

and $\pi(\tau \,|\, \Omega_{\sigma,v}) = 0$ whenever $\tau \neq \sigma^{v,k}$ for all k. Therefore we can write the transition matrix of the heat-bath dynamics on Ω_{P} as

$$P_{\mathrm{HB}}(\sigma, \tau) := P^{G}_{\mathrm{HB},\beta,q}(\sigma, \tau) = \frac{1}{|V|} \sum_{v \in V} \frac{\pi(\tau)}{\sum_{l=1}^{q} \pi(\sigma^{v,l})} \mathbb{1}(\tau \in \Omega_{\sigma,v}). \tag{2.18}$$

Clearly, P_{HB} is reversible with respect to π and, as long as $\beta < \infty$, also ergodic.

An interesting feature of this simple construction of a Markov chain is that the (temporal) mixing properties of the heat-bath dynamics are proven to be equivalent to some *spatial mixing* properties of the associated Potts measure if the underlying graph is a rectangular subset of the (physically most relevant) d-dimensional integer lattice $\mathbb{Z}^d$; see e.g. [1, 22, 32, 46, 47, 48, 49, 65] or the survey [45] by Martinelli. Similar results are known for trees [52]. It turns out that certain spatial properties of the model imply tight bounds on the spectral gap (or mixing time) of the heat-bath dynamics. In Section 2.5 we will state some of these results, in particular for the two-dimensional square lattice where we have an almost complete characterization of the mixing properties.

One drawback of the heat-bath dynamics is that it is typically slowly (i.e. not rapidly) mixing at low temperatures (large β); see e.g. [6, 7, 39, 64]. We refer again to Section 2.5. In fact, at zero temperature ($\beta = \infty$) the heat-bath dynamics is not even irreducible, although this is the simplest case: uniform distribution on q (constant) configurations.

In the next paragraph we introduce the Swendsen–Wang dynamics, which overcomes the slow mixing behavior at low temperatures. This Markov chain is the primarily studied Markov chain of this paper.

The Swendsen–Wang dynamics. Inspired by the representation of the Potts models of Fortuin and Kasteleyn [24], Swendsen and Wang [67] invented around 1987 a Markov chain that makes large steps through the state space and is surprisingly simple to implement. We call this Markov chain, which is presently the best candidate for the most efficient dynamics to sample from the Potts model on general families of graphs, the Swendsen–Wang (SW) dynamics.

As stated several times, this Markov chain is based on the connection of the random-cluster and Potts models that is given by the coupling of the corresponding measures of Edwards and Sokal [23].

First, recall that this coupling is given by the FKES measure (2.12), i.e.

$$\nu(\sigma, A) = \frac{1}{Z}\left(\frac{p}{1-p}\right)^{|A|} \mathbb{1}(A \subset E(\sigma)), \qquad \sigma \in \Omega_{\mathrm{P}},\ A \in \Omega_{\mathrm{RC}},$$

with $E(\sigma)$ from (2.10) and some $p \in [0,1]$, and that the conditional probabilities (with respect to ν) of σ given A (or of A given σ) satisfy

$$\nu(\sigma \,|\, A) = q^{-c(A)} \mathbb{1}(A \subset E(\sigma)), \tag{2.19}$$

$$\nu(A \,|\, \sigma) = p^{|A|}(1-p)^{|E(\sigma)|-|A|} \mathbb{1}(A \subset E(\sigma)) \tag{2.20}$$

(see (2.15) and the discussion thereafter).

The *Swendsen–Wang dynamics* (for the Potts model) makes use of these conditional probabilities in such a way that, given the current state $\sigma \in \Omega_{\mathrm{P}}$, the next state $\tau \in \Omega_{\mathrm{P}}$ is generated by first sampling an $A \subset E$ from $\nu(\cdot \,|\, \sigma)$, and then sampling τ from $\nu(\cdot \,|\, A)$. Thus, the transition matrix of the Swendsen–Wang dynamics for the q-state Potts model on G at inverse temperature $\beta = -\ln(1-p)$ is given by

$$\begin{aligned} P_{\mathrm{SW}}(\sigma,\tau) := P^G_{\mathrm{SW},\beta,q}(\sigma,\tau) &= \sum_{A \subset E} \nu(A \,|\, \sigma)\nu(\tau \,|\, A) \\ &= (1-p)^{|E(\sigma)|} \sum_{A \subset E} \left(\frac{p}{1-p}\right)^{|A|} q^{-c(A)} \mathbb{1}(A \subset E(\sigma) \cap E(\tau)). \end{aligned} \tag{2.21}$$

Using the same interpretation of this sampling with respect to the conditional probabilities that was given in Section 2.3, we can describe one step of the Swendsen–Wang dynamics as the following two-step procedure:

(SW1) Given a Potts configuration $\sigma \in \Omega_{\mathrm{P}}$ on G, delete each edge of $E(\sigma)$ independently with probability $1-p = e^{-\beta}$. This gives $A \in \Omega_{\mathrm{RC}}$.

(SW2) Assign a random color from $[q]$ independently to each connected component of (V, A). Vertices of the same component get the same color. This gives $\tau \in \Omega_{\mathrm{P}}$.

Ergodicity of this Markov chain is easy to check. For this, imagine that the generated edge set A of the first step (SW1) is the empty set (this happens with probability $(1-p)^{|E(\sigma)|} > 0$ for $p < 1$), and then, in step (SW2), every possible configuration can be generated with probability $q^{-|V|}$. Hence, $P_{\mathrm{SW}}(\sigma,\tau) > 0$ for all $\sigma, \tau \in \Omega_{\mathrm{P}}$, which implies ergodicity. If $p = 1$ ($\beta = \infty$), then the only elements of Ω_{P} with positive measure are the q constant configurations, for which we have $P_{\mathrm{SW}}(\sigma,\tau) = q^{-1} > 0$. For the reversibility of P_{SW} with respect to π note that $\pi(\sigma) = (1/Z)(1-p)^{-|E(\sigma)|}$, since $p = 1 - e^{-\beta}$. We

deduce from (2.21) that

$$\pi(\sigma)P_{\mathrm{SW}}(\sigma,\tau) = \frac{1}{Z}\sum_{A\subset E} p^{|A|}(1-p)^{-|A|}q^{-c(A)}\mathbb{1}(A\subset E(\sigma)\cap E(\tau)),$$

which is symmetric in σ,τ. This shows that the SW dynamics is reversible and ergodic for any G, q and β, but, in contrast with the heat-bath dynamics, this Markov chain seems to be efficient also for large β. For results showing that this is true for large enough β see e.g. [34] and [44]. Additionally, there are plenty of numerical results suggesting that the Swendsen–Wang dynamics is rapidly mixing at all temperatures in various cases. These include the Ising model and the 3-state Potts model on the two-dimensional square lattice [19, 63] and the Ising model on the three-dimensional cubic lattice [56].

However, rigorous proofs of such statements are rare. There are, as far as we know, presently only four cases where rapid mixing at all temperatures is proven. These are on trees, cycles [11, 34, 41] and narrow grids [10] for all $q\in\mathbb{N}$ and on the complete graph for $q=2$ [41, 42]. One of our goals is to add the Ising model ($q=2$) on the two-dimensional square lattice to this list.

There are also some rigorous results regarding slow mixing. We state three of them: Li and Sokal [40] proved that the inverse spectral gap is larger than some constant times the *specific heat* (that is proportional to the second derivative with respect to β of the Potts normalization constant Z, see (2.9)). This shows (at least) that the inverse spectral gap cannot be bounded independently of the size of the graph in many cases of interest. Furthermore, Gore and Jerrum [28] showed slow mixing of the SW dynamics on the complete graph for $q\geq 3$ for some value of the inverse temperature. Recently, Borgs, Chayes and Tetali [5] gave tight upper and lower bounds on the spectral gap (or mixing time) for SW on rectangular subsets of $\mathbb{Z}^d$, $d\geq 2$, with periodic boundary conditions at some (critical) temperature for q large enough, which show that the spectral gap is exponentially small in the size of the graph. Since we need some of the above results in the following, we state them in more detail in Section 2.5.

Apart from the results given above, the Swendsen–Wang dynamics resisted a precise analysis in (physically) relevant cases, where the heat-bath dynamics is well-understood. For instance, it was not even proven that SW is rapidly mixing for the Ising model on the two-dimensional square lattice for all temperatures above the critical one, which has been known for the heat-bath dynamics for 20 years [47, 48].

We will prove this statement in Chapter 3 by comparison with the heat-bath dynamics. Furthermore we will obtain rapid mixing at all temperatures below the critical one.

The main challenge in proving rapid mixing of the Swendsen–Wang dynamics for the two-dimensional square lattice is the low temperature (large β) regime, where no Markov chain is proven so far to be rapid down to the critical temperature. Therefore, we need to consider dynamics for the corresponding random-cluster model that seems to be (and we will see that it indeed is) more convenient for relating spectral gaps at high and low temperatures.

2.4.2. Dynamics for the random-cluster model. We begin this subsection with the definition of a non-local Markov chain for the random-cluster model. This is again the

Swendsen–Wang dynamics. Since it is defined by means of conditional probabilities with respect to the FKES measure ν (see (2.21)), this Markov chain can be naturally defined also for the random-cluster model.

The Swendsen–Wang dynamics. The SW dynamics for the random-cluster model is based on the same connection that was given in the last subsection, but here we make the two-step procedure in reverse order:

($\widetilde{\text{SW}}1$) Given a random-cluster configuration $A \in \Omega_{\text{RC}}$ on G, assign a random color from $[q]$ independently to each connected component of (V, A). Vertices of the same component get the same color. This gives $\sigma \in \Omega_{\text{P}}$.

($\widetilde{\text{SW}}2$) Take $E(\sigma)$ and delete each edge independently with probability $1-p$. This gives the new RC configuration $B \in \Omega_{\text{RC}}$.

Thus, the transition matrix of the Swendsen–Wang dynamics for the random-cluster model on G with parameters p and q is given by

$$\begin{aligned}\widetilde{P}_{\text{SW}}(A,B) &:= \widetilde{P}^G_{\text{SW},p,q}(A,B) = \sum_{\sigma\in\Omega_{\text{P}}} \nu(\sigma \,|\, A)\nu(B \,|\, \sigma) \\ &= q^{-c(A)}\left(\frac{p}{1-p}\right)^{|B|} \sum_{\sigma\in\Omega_{\text{P}}} (1-p)^{|E(\sigma)|}\mathbb{1}(A\cup B \subset E(\sigma)). \qquad (2.22)\end{aligned}$$

Ergodicity and reversibility of this Markov chain can be shown by similar ideas to those above. Although the Swendsen–Wang dynamics for the RC model appears quite often in the literature (especially its generalization to non-integer values of q; see e.g. [8, 9, 13, 14]) we are not aware of any attempt to prove mixing properties of this Markov chain. The following lemma shows however that every result on the spectral gap of the SW dynamics for the Potts model is also valid for SW for the corresponding random-cluster model.

LEMMA 2.6. *Let $\widetilde{P}_{\text{SW}}$ (resp. P_{SW}) be the Swendsen–Wang dynamics for the random-cluster (resp. q-state Potts) model with parameters p and q (resp. at inverse temperature β with $p = 1 - e^{-\beta}$). Then*

$$\lambda(P_{\text{SW}}) = \lambda(\widetilde{P}_{\text{SW}}).$$

Proof. Define the $\Omega_{\text{P}} \times \Omega_{\text{RC}}$-matrix $\mathcal{C}$ by

$$\mathcal{C}(\sigma, A) := \nu(A \,|\, \sigma) = \frac{\nu(\sigma, A)}{\pi(\sigma)}, \qquad \sigma \in \Omega_{\text{P}},\ A \in \Omega_{\text{RC}}.$$

Regard $\mathcal{C}$ as an operator that maps from $L_2(\mu)$ to $L_2(\pi)$ (cf. (2.4)) by

$$\mathcal{C}g(\sigma) = \sum_{A\in\Omega_{\text{RC}}} \mathcal{C}(\sigma, A)g(A)$$

for $g \in \mathbb{R}^{\Omega_{\text{RC}}}$. By the usual definition of the adjoint of an operator between different Hilbert spaces (see e.g. [36, Def. 3.9-1]), the adjoint operator $\mathcal{C}^*$ of $\mathcal{C}$ satisfies

$$\langle \mathcal{C}^* f, g\rangle_\mu = \langle f, \mathcal{C}g\rangle_\pi$$

for all $f \in \mathbb{R}^{\Omega_{\text{P}}}$ and $g \in \mathbb{R}^{\Omega_{\text{RC}}}$. Thus, the matrix corresponding to $\mathcal{C}^*$ is given by

$$\mathcal{C}^*(A, \sigma) = \frac{\pi(\sigma)}{\mu(A)}\mathcal{C}(\sigma, A) = \frac{\nu(\sigma, A)}{\mu(A)} = \nu(\sigma \,|\, A).$$

The definitions of the Swendsen–Wang dynamics in (2.21) and (2.22) imply $P_{\mathrm{SW}} = \mathcal{C}\mathcal{C}^*$ and $\widetilde{P}_{\mathrm{SW}} = \mathcal{C}^*\mathcal{C}$. Additionally, we define $S_\pi(\sigma,\tau) := \pi(\tau)$, $S_\mu(A,B) := \mu(B)$ and $S_{(\pi,\mu)}(\sigma,A) = \mu(A)$ for all $\sigma,\tau \in \Omega_{\mathrm{P}}$ and $A,B \in \Omega_{\mathrm{RC}}$. It is straightforward to verify that $S^*_{(\pi,\mu)}(A,\sigma) = \pi(\sigma)$, $\sigma \in \Omega_{\mathrm{P}}$, $A \in \Omega_{\mathrm{RC}}$. Thus, $S_{(\pi,\mu)}S^*_{(\pi,\mu)} = S_\pi$ and $S^*_{(\pi,\mu)}S_{(\pi,\mu)} = S_\mu$. We obtain $\mathcal{C}\mathcal{C}^* - S_\pi = (\mathcal{C} - S_{(\pi,\mu)})(\mathcal{C} - S_{(\pi,\mu)})^*$ since by definition $S_{(\pi,\mu)}\mathcal{C}^* = S_\pi = S^*_\pi = \mathcal{C}S^*_{(\pi,\mu)}$ (see (2.15) and below). Analogously we get $\mathcal{C}^*\mathcal{C} - S_\mu = (\mathcal{C} - S_{(\pi,\mu)})^*(\mathcal{C} - S_{(\pi,\mu)})$. Recall the definition of the spectral gap from (2.8). It follows from simple properties of the norm of (adjoint) operators between Hilbert spaces (see e.g. [36, Thm. 3.9-4]) that

$$\begin{aligned} 1 - \lambda(P_{\mathrm{SW}}) &= \|P_{\mathrm{SW}} - S_\pi\|_\pi = \|(\mathcal{C} - S_{(\pi,\mu)})(\mathcal{C} - S_{(\pi,\mu)})^*\|_\pi = \|\mathcal{C} - S_{(\pi,\mu)}\|^2_{L_2(\mu)\to L_2(\pi)} \\ &= \|(\mathcal{C} - S_{(\pi,\mu)})^*(\mathcal{C} - S_{(\pi,\mu)})\|_\mu = \|\widetilde{P}_{\mathrm{SW}} - S_\mu\|_\mu = 1 - \lambda(\widetilde{P}_{\mathrm{SW}}), \end{aligned}$$

which proves the statement. ∎

The single-bond dynamics. The last kind of Markov chains we want to consider are the local Markov chains for the random-cluster model. In fact, we need more than one construction in this case. This comes from the fact that the two dynamics, namely heat-bath and single-bond dynamics, both have properties that are needed for the further analysis.

First of all we introduce the *heat-bath* (HB) *dynamics* for the random-cluster model. Similarly to the heat-bath dynamics for the Potts model, suppose that $A \in \Omega_{\mathrm{RC}}$ is the current state of the Markov chain. Then:

($\widetilde{\mathrm{HB}}$1) Choose an edge $e \in E$ uniformly at random.
($\widetilde{\mathrm{HB}}$2) Choose the next state, say $B \in \Omega_{\mathrm{RC}}$, with respect to $\mu(\cdot \mid \{A \cup e, A \setminus e\})$, i.e. the conditional probability (with respect to μ) given that B differs from A only by e.

We use $A \cup e$ instead of $A \cup \{e\}$ (similarly for $\cap$, $\setminus$). Recall that, for $A \in \Omega_{\mathrm{RC}}$, we write $u \overset{A}{\longleftrightarrow} v$ if $u,v \in V$ are connected in the subgraph (V,A) and that we denote the endpoints of $e \in E$ by $e^{(1)}$ and $e^{(2)}$. Write $u \overset{A}{\not\leftrightarrow} v$ if u and v are not connected in (V,A). If we additionally define $A^e := A \ominus e$ with the symmetric difference $\ominus$, i.e.

$$A^e = \begin{cases} A \cup e & \text{if } e \notin A, \\ A \setminus e & \text{if } e \in A, \end{cases} \tag{2.23}$$

we can write the transition probability matrix $\widetilde{P}_{\mathrm{HB}} := \widetilde{P}^G_{\mathrm{HB},p,q}$ of the heat-bath dynamics for the RC model as

$$\widetilde{P}_{\mathrm{HB}}(A,B) = \frac{1}{|E|}\sum_{e\in E} \frac{\mu(B)}{\mu(A)+\mu(A^e)}\mathbb{1}(B \setminus e = A \setminus e). \tag{2.24}$$

Note that unless $A = B$, at most one summand in the above sum is not zero. Using the definition of μ from (2.11) we obtain

$$\frac{\mu(A)}{\mu(A^e)} = \left(\frac{p}{1-p}\right)^{|A|-|A^e|} q^{c(A)-c(A^e)}, \tag{2.25}$$

where $c(A) = c(A^e)$ if and only if $e^{(1)} \overset{A\setminus e}{\longleftrightarrow} e^{(2)}$. Otherwise, $c(A) - c(A^e) = 2\mathbb{1}(e \notin A) - 1$.

Hence, $\widetilde{P}_{\mathrm{HB}}$ satisfies

$$\widetilde{P}_{\mathrm{HB}}(A,B) = \frac{1}{|E|}\sum_{e\in E}\begin{cases} p & \text{if } B = A\cup e \text{ and } e^{(1)} \overset{A\setminus e}{\longleftrightarrow} e^{(2)},\\ 1-p & \text{if } B = A\setminus e \text{ and } e^{(1)} \overset{A\setminus e}{\longleftrightarrow} e^{(2)},\\ \frac{p}{p+q(1-p)} & \text{if } B = A\cup e \text{ and } e^{(1)} \overset{A\setminus e}{\not\longleftrightarrow} e^{(2)},\\ \frac{q(1-p)}{p+q(1-p)} & \text{if } B = A\setminus e \text{ and } e^{(1)} \overset{A\setminus e}{\not\longleftrightarrow} e^{(2)}, \end{cases} \tag{2.26}$$

for $A, B \in \Omega_{\mathrm{RC}}$. We immediately obtain from this equation that $\widetilde{P}_{\mathrm{HB}}$ does not have the finite range interaction property, since we have to check connectivity of two vertices in each step, which is in general a non-local property.

In addition to the need of this Markov chain in an intermediate step of our proof of rapid mixing for the Swendsen–Wang dynamics, the heat-bath dynamics is interesting in its own right. One reason is that it also provides a Markov chain for the random-cluster model with non-integer values of $q > 0$. It would be interesting to prove a similar relation between temporal mixing properties of $\widetilde{P}_{\mathrm{HB}}$ and some "spatial mixing" property in the RC model, as it is known for single-spin dynamics for the Potts model. Another reason is that the heat-bath dynamics was proposed by Propp and Wilson [57] to produce exact samples from μ (and so also from the Potts model) using their famous *coupling from the past* procedure. Therefore, every mixing time bound on $\widetilde{P}_{\mathrm{HB}}$ leads to a bound on the (expected) cost of their procedure.

To the best of our knowledge, there is presently only one result on mixing of local dynamics for the RC model. This is the upper bound on the mixing time of Ge and Štefankovič [25] that shows rapid mixing of a Metropolis-type version of $\widetilde{P}_{\mathrm{HB}}$ for every $p \in (0,1)$ and $q > 0$ if the underlying graph has bounded *tree-width*. We will state this result later (see Theorem 2.19).

The advantage of the heat-bath dynamics for our purposes is that its spectral gap, if the underlying graph G is *planar*, can be shown to be the same as the spectral gap of the heat-bath dynamics for the RC model on the *dual graph* of G with suitable values of the parameters p and q (see Chapter 5). This provides us with a relation of the mixing properties of the Swendsen–Wang dynamics at high and low temperatures (see Theorem 5.6).

The second local Markov chain that we consider is inspired by the local behavior of the Swendsen–Wang dynamics and we will see (Theorem 4.8) that it enables us to give lower and upper bounds on the spectral gap of the SW dynamics in terms of the spectral gap of this Markov chain. The aforementioned local behavior can be demonstrated by the following example. Fix $p \in (0,1)$, $q \in \mathbb{N}$ and let the graph be given by $G_1 = (\{u,v\}, \{\{u,v\}\})$, i.e. the graph with two vertices and one edge between them. Let $e = \{u,v\}$. Then it is easy to verify that the transition probabilities of the SW dynamics satisfy $\widetilde{P}^{G_1}_{\mathrm{SW}}(\emptyset, e) = p/q$ and $\widetilde{P}^{G_1}_{\mathrm{SW}}(e, \emptyset) = 1-p$. For this note that, following the steps ($\widetilde{\mathrm{SW}}$1) and ($\widetilde{\mathrm{SW}}$2), if the current configuration is $\emptyset \subset \{e\}$ we assign independently a random color from $[q]$ to each of the vertices, and if their colors agree (which happens with probability $1/q$), we put an edge between them with probability p. Hence, $\widetilde{P}^{G_1}_{\mathrm{SW}}(\emptyset, e) = p/q$. The second equality can be shown analogously. We use these transition probabilities to construct a local Markov

chain for the RC model on arbitrary graphs, which we call the *single-bond dynamics*, as follows. Let the current state be $A \subset E$, choose an edge $e \in E$ uniformly at random and decide if e shall be included in the new configuration or not depending on the connectivity of the endvertices of e in A. If the endvertices are not connected in A include e with probability $\widetilde{P}_{\mathrm{SW}}^{G_1}(\emptyset, e)$, otherwise include e with probability $\widetilde{P}_{\mathrm{SW}}^{G_1}(e, e)$. Note that in the heat-bath dynamics, we check connectivity in $A \setminus e$ instead of connectivity in A.

The transition matrix of the *single-bond* (SB) *dynamics* is given by

$$\widetilde{P}_{\mathrm{SB}}(A, B) = \frac{1}{|E|} \sum_{e \in E} \begin{cases} p & \text{if } B = A \cup e \text{ and } e^{(1)} \overset{A}{\longleftrightarrow} e^{(2)}, \\ 1 - p & \text{if } B = A \setminus e \text{ and } e^{(1)} \overset{A}{\longleftrightarrow} e^{(2)}, \\ \frac{p}{q} & \text{if } B = A \cup e \text{ and } e^{(1)} \overset{A}{\not\longleftrightarrow} e^{(2)}, \\ 1 - \frac{p}{q} & \text{if } B = A \setminus e \text{ and } e^{(1)} \overset{A}{\not\longleftrightarrow} e^{(2)}. \end{cases} \tag{2.27}$$

Ergodicity of $\widetilde{P}_{\mathrm{SB}}$ is clear. For the reversibility with respect to μ it is enough to prove

$$\frac{\widetilde{P}_{\mathrm{SB}}(A \setminus e, A \cup e)}{\widetilde{P}_{\mathrm{SB}}(A \cup e, A \setminus e)} = \frac{\mu(A \cup e)}{\mu(A \setminus e)} \overset{(2.25)}{=} \begin{cases} \frac{p}{1-p} & \text{if } e^{(1)} \overset{A \setminus e}{\longleftrightarrow} e^{(2)}, \\ \frac{p}{q(1-p)} & \text{if } e^{(1)} \overset{A \setminus e}{\not\longleftrightarrow} e^{(2)}, \end{cases}$$

for every $A \in \Omega_{\mathrm{RC}}$ and $e \in E$. First note that $e^{(1)}$ and $e^{(2)}$ are always connected in $A \cup e$, hence we divide by $\widetilde{P}_{\mathrm{SB}}(A \cup e, A \setminus e) = 1 - p$ independently of A and e in the above equation. Additionally, the numerator of the above fraction is p or p/q depending on connectivity in $A \setminus e$, as desired.

As for the heat-bath dynamics for the RC model, the single-bond dynamics is also a valid local Markov chain for non-integer values of q, but here we have to assume $q \geq 1$ in order to ensure that $\widetilde{P}_{\mathrm{SB}}$ is well-defined.

In Section 4.1 we will present the usefulness of this dynamics by providing a common representation of this and the Swendsen–Wang dynamics on the joint (FKES) model, using the same "building blocks". We finish this section with an inequality between the spectral gaps of $\widetilde{P}_{\mathrm{SB}}$ and $\widetilde{P}_{\mathrm{HB}}$.

LEMMA 2.7. *For $\widetilde{P}_{\mathrm{HB}}$ and $\widetilde{P}_{\mathrm{SB}}$ for the random-cluster model with parameters p and q we have*

$$\left(1 - p\left(1 - \frac{1}{q}\right)\right) \lambda(\widetilde{P}_{\mathrm{HB}}) \leq \lambda(\widetilde{P}_{\mathrm{SB}}) \leq \lambda(\widetilde{P}_{\mathrm{HB}}).$$

Proof. First we show that $\widetilde{P}_{\mathrm{HB}}$ and $\widetilde{P}_{\mathrm{SB}}$ have only non-negative eigenvalues. For this write $\widetilde{P}_{\mathrm{HB}} = \frac{1}{|E|} \sum_{e \in E} \widetilde{P}_e$ with

$$\widetilde{P}_e(A, B) := \frac{\mu(B)}{\mu(A) + \mu(A^e)} \mathbb{1}(B \setminus e = A \setminus e), \qquad A, B \in \Omega_{\mathrm{RC}}$$

(see (2.24)). Obviously, $\widetilde{P}_e$, $e \in E$, is reversible with respect to μ and satisfies $\widetilde{P}_e^2 = \widetilde{P}_e$, since the distributions $\widetilde{P}_e(A \setminus e, \cdot)$ and $\widetilde{P}_e(A \cup e, \cdot)$ are equal. This shows that all $\widetilde{P}_e$, $e \in E$, are projections (see [36, Thm. 9.5-1]), and thus positive, i.e. $\langle \widetilde{P}_e g, g \rangle_\mu \geq 0$ for all $g \in L_2(\mu)$ (see [36, Thm. 9.5-2]). Using the fact that the sum of positive operators is positive we obtain positivity of $\widetilde{P}_{\mathrm{HB}}$. It follows that $\widetilde{P}_{\mathrm{HB}}$ has only non-negative eigenvalues (see [33,

Obs. 7.1.4]). Similar arguments lead to the same statement for $\widetilde{P}_{\mathrm{SB}}$ (see Remark 4.4). By Lemma 2.5, and since $\widetilde{P}_{\mathrm{HB}}$ and $\widetilde{P}_{\mathrm{SB}}$ have the same stationary distribution, it is sufficient to show

$$\left(1 - p\left(1 - \frac{1}{q}\right)\right)\widetilde{P}_{\mathrm{HB}}(A,B) \le \widetilde{P}_{\mathrm{SB}}(A,B) \le \widetilde{P}_{\mathrm{HB}}(A,B)$$

for $A \neq B \in \Omega_{\mathrm{RC}}$, i.e. for $B = A^e$ for some $e \in E$ (otherwise $\widetilde{P}_{\mathrm{HB}}(A,B) = \widetilde{P}_{\mathrm{SB}}(A,B) = 0$). By reversibility we obtain

$$\frac{\widetilde{P}_{\mathrm{HB}}(A \setminus e, A \cup e)}{\widetilde{P}_{\mathrm{SB}}(A \setminus e, A \cup e)} = \frac{\widetilde{P}_{\mathrm{HB}}(A \cup e, A \setminus e)}{\widetilde{P}_{\mathrm{SB}}(A \cup e, A \setminus e)} \overset{(2.27)}{=} \frac{|E|}{1-p}\widetilde{P}_{\mathrm{HB}}(A \cup e, A \setminus e).$$

Since

$$\frac{1-p}{|E|} \le \widetilde{P}_{\mathrm{HB}}(A \cup e, A \setminus e) \le \frac{1}{|E|}\,\frac{q(1-p)}{p+q(1-p)}$$

for $q \ge 1$, it follows that

$$1 \le \frac{\widetilde{P}_{\mathrm{HB}}(A, A^e)}{\widetilde{P}_{\mathrm{SB}}(A, A^e)} \le \frac{q}{p+q(1-p)} = \frac{1}{1-p(1-q^{-1})} \tag{2.28}$$

for all $A \in \Omega_{\mathrm{RC}}$ and $e \in E$. ∎

2.5. Known results. In this section we present a selection of known results on the mixing properties for the above introduced algorithms. In fact, for the heat-bath dynamics for the Potts model, we will state only results that are needed for the further analysis. For the Swendsen–Wang dynamics and the local Markov chains for the random-cluster model we try to give a complete overview of the known results.

Since, in the original papers, some results are given in terms of spectral gap and some with mixing times, we first state a corollary to Lemma 2.3 that we need for translation.

Corollary 2.8. *Let P (resp. $\widetilde{P}$) be the transition matrix of a reversible, ergodic Markov chain for the q-state Potts (resp. random-cluster) model on a graph $G = (V, E)$ at inverse temperature β (resp. with parameters p and q). Then*

$$\lambda(P)^{-1} - 1 \le t_{\mathrm{mix}}(P) \le (2 + \beta|E| + |V|\log q)\lambda(P)^{-1},$$
$$\lambda(\widetilde{P})^{-1} - 1 \le t_{\mathrm{mix}}(\widetilde{P}) \le \left(2 + |E|\log\frac{1}{p(1-p)} + |V|\log q\right)\lambda(\widetilde{P})^{-1}.$$

In particular, this shows that every result of this paper can also be written in terms of the mixing time, loosing the same factor as in Corollary 2.8.

We begin with the probably most studied instance: the heat-bath dynamics for the q-state Potts model on the two-dimensional square lattice. For $d \ge 1$, define the d-dimensional *hypercubic lattice* $\mathbb{Z}_L^d$ of side length L as the graph $\mathbb{Z}_L^d = (V_{L,d}, E_{L,d})$ with vertex set $V_{L,d} = \{1, \dots, L\}^d \subset \mathbb{Z}^d$ and edge set $E_{L,d} = \{\{u, v\} \subset V_{L,d} : |u - v| = 1\}$, where $|\cdot|$ denotes the Euclidean norm. For $d = 2$ we call $\mathbb{Z}_L^2$ the two-dimensional square lattice of side length L. In this case ($d = 2$) there is an almost complete characterization of the spectral gap, that was established over the last decades, in particular for the Ising model ($q = 2$). Beginning with the work of Holley [32], Aizenman and Holley [1] and Stroock and Zegarliński [66], who showed rapid mixing of the heat-bath dynamics given some spatial mixing property, which is called *complete analyticity* (or *Dobrushin–*

Shlosman mixing condition) (see [18]), it was finally proven by Martinelli and Olivieri [47, 48] that the heat-bath dynamics for the Ising model is rapidly mixing up to the *critical temperature*, i.e. if the inverse temperature β satisfies $\beta < \beta_c(2) = \log(1+\sqrt{2})$. Using results of Cesi, Chayes, Chayes, Guadagni, Martinelli and Schonmann [7, 6, 64] it is known that this is (almost) best possible in the sense that the spectral gap of the HB dynamics on $\mathbb{Z}_L^2$ at inverse temperature β is smaller than $\exp(-cL)$, for some $c > 0$, if $\beta > \beta_c(2)$. Only recently has rapid mixing at the critical temperature β_c been proven by Lubetzky and Sly [43]. For the proof of rapid mixing of the heat-bath dynamics for the Potts model at all $\beta < \log(1+\sqrt{q})$ we need *exponential decay of connectivities* in the RC model (see Beffara and Duminil-Copin [3]). This implies *weak mixing* in the Potts model (see Alexander [2]), and thus rapid mixing of the heat-bath dynamics (see Martinelli, Olivieri and Schonmann [49]).

Before we summarize these results in Theorem 2.10, we introduce a variant of the Potts measure (2.9) and the heat-bath dynamics (2.18) with (constant) boundary condition. For this define the *boundary* of $V_{L,d}$ by

$$\partial V_{L,d} := \{v = (v_1, \dots, v_d) \in V_{L,d} : v_i \in \{1, L\} \text{ for some } i \in [d]\}$$

and let $d^+_{L,d}(v)$ be the number of neighbors of v in $\mathbb{Z}^d \setminus V_{L,d}$, i.e.,

$$d^+_{L,d}(v) := |\{u \in \mathbb{Z}^d \setminus V_{L,d} : \ |v-u| = 1\}|.$$

The Potts measure on $\mathbb{Z}_L^d$ with 1-*boundary condition* is defined by

$$\pi^{\mathbb{Z}_L^d,1}_{\beta,q}(\sigma) := \bar{Z}^{-1}\pi^{\mathbb{Z}_L^d}_{\beta,q}(\sigma) \prod_{v\in\partial V_{L,d}} \exp\big(\beta d^+_{L,d}(v)\mathbb{1}(\sigma(v)=1)\big), \qquad \sigma \in \Omega_{\mathrm{P}}(\mathbb{Z}_L^d), \tag{2.29}$$

where $\bar{Z}$ is the proper normalization constant and $\pi^{\mathbb{Z}_L^d}_{\beta,q}$ is defined as in (2.9). This measure can be interpreted as the conditional distribution of the configurations on $V_{L,d}$ given that all vertices of $\mathbb{Z}^d \setminus V_{L,d}$ have color 1.

Remark 2.9. The critical inverse temperature $\beta_c(d,q)$ for the q-state Potts model on $\mathbb{Z}^d$, that was cited above for $d = 2$, is generally defined by

$$\beta_c(d,q) := \inf\{\beta : M_{d,q}(\beta) > 0\}, \tag{2.30}$$

where

$$M_{d,q}(\beta) := \lim_{L\to\infty} \frac{1}{|V_{L,d}|} \sum_{v\in V_{L,d}} \left(\pi^{\mathbb{Z}_L^d,1}_{\beta,q}(\{\sigma : \sigma(v)=1\}) - \frac{1}{q}\right).$$

(It is well-known that these limits exist; see e.g. Grimmett [29].) We write $\beta_c(q)$ for $\beta_c(2,q)$. A closed formula for $\beta_c(q)$ was first established by Onsager [55] in the case $q = 2$ by giving an explicit formula for $M_{2,2}(\beta)$. A proof of the equality $\beta_c(q) = \log(1+\sqrt{q})$ for all $q \geq 2$, which was expected to be true, has been given only recently by Beffara and Duminil-Copin [3]. For $d \geq 3$ it is still a challenging open problem to give an explicit formula for the critical inverse temperature. However, it is known (see Laanait et al. [37]) that

$$\beta_c(d,q) = \frac{1}{d}\log q + \mathcal{O}(q^{-1/d})$$

for q large enough.

Let $\pi^1 := \pi_{\beta,q}^{\mathbb{Z}_L^2,1}$. Similarly to (2.18) we define the transition matrix of the heat-bath dynamics for the Potts model on $\mathbb{Z}_L^2$ with 1-boundary condition by

$$P_{\mathrm{HB},1}(\sigma,\tau) := P_{\mathrm{HB},1,\beta,q}^{\mathbb{Z}_L^2}(\sigma,\tau) = \frac{1}{|V_{L,2}|}\sum_{v\in V_{L,2}} \frac{\pi^1(\tau)}{\sum_{l=1}^q \pi^1(\sigma^{v,l})}\mathbb{1}(\tau\in\Omega_{\sigma,v}) \tag{2.31}$$

for all $\sigma,\tau\in\Omega_{\mathrm{P}}(\mathbb{Z}_L^2)$. We summarize the rapid mixing results stated above in the following theorem.

THEOREM 2.10. *Let P_{HB} be the transition matrix of the heat-bath dynamics for the q-state Potts model on $\mathbb{Z}_L^2$ at inverse temperature β. Let $n = L^2 = |V_{L,2}|$. Then there exist constants $c_\beta = c_\beta(q) > 0$ and $C < \infty$ such that*

$$\lambda(P_{\mathrm{HB}}) \geq \frac{c_\beta}{n} \qquad \text{for } \beta < \beta_c(q)$$

and

$$\lambda(P_{\mathrm{HB}}) \geq n^{-C} \qquad \text{for } q = 2 \text{ and } \beta = \beta_c(2),$$

where $\beta_c(q) = \log(1+\sqrt{q})$. These bounds hold also if we replace P_{HB} by $P_{\mathrm{HB},1}$.

Proof. The results, as originally given in [43, 47], refer to a continuous-time Markov process for the Potts model. See e.g. [45] for an introduction to the "graphical construction" of the continuous-time heat-bath dynamics for the Ising model. In fact, these papers present lower bounds on $\mathrm{gap}(\mathbb{Z}_L^2)$, which is defined by

$$\mathrm{gap}(\mathbb{Z}_L^2) := \inf_{\substack{f\in L_2(\pi):\\ \mathrm{Var}_\pi(f)=1}} \frac{1}{2}\sum_{\sigma\in\Omega_{\mathrm{P}}}\sum_{v\in V_{L,2}}\sum_{k=1}^q \pi(\sigma)\frac{\pi(\sigma^{v,k})}{\sum_{l=1}^q \pi(\sigma^{v,l})}(f(\sigma^{v,k})-f(\sigma))^2$$

(cf. [49, Sec. 3]). By the variational characterization of the eigenvalues of reversible transition matrices (see e.g. [16] or the proof of Lemma 2.5) we can write

$$1-\xi_2 = \inf_{\substack{f\in L_2(\pi):\\ \mathrm{Var}_\pi(f)=1}} \frac{1}{2}\sum_{\sigma\in\Omega_{\mathrm{P}}}\sum_{\tau\in\Omega_{\mathrm{P}}} \pi(\sigma)P_{\mathrm{HB}}(\sigma,\tau)(f(\tau)-f(\sigma))^2,$$

where ξ_2 is the second largest eigenvalue of P_{HB}. By definition, $\mathrm{gap}(\mathbb{Z}_L^2) = n(1-\xi_2)$ (see (2.18)). But P_{HB} has in general only non-negative eigenvalues, and thus $\lambda(P_{\mathrm{HB}}) = 1-\xi_2$. To see this, write $P_{\mathrm{HB}} = \frac{1}{|V|}\sum_{v\in V} P_v$ with $P_v(\sigma,\tau) = \pi(\tau\,|\,\Omega_{\sigma,v})$ (cf. (2.17)), and note that $P_v^2 = P_v$ since $\Omega_{\sigma,v} = \Omega_{\tau,v}$ for all σ,τ with $P_v(\sigma,\tau) > 0$. That is, P_v is a projection (see [36, Thm. 9.5-1]). It follows that all P_v, $v\in V$, and thus P_{HB}, have only non-negative eigenvalues by [36, 9.5-2]. Similar ideas were used in the proof of Lemma 2.7.

Hence, it is enough to show that there exist constants $\widetilde{c}_\beta > 0$ and $\widetilde{C} < \infty$ such that

$$\mathrm{gap}(\mathbb{Z}_L^2) \geq \begin{cases} \widetilde{c}_\beta & \text{for } \beta < \beta_c(q), \\ n^{-\widetilde{C}} & \text{for } q = 2 \text{ and } \beta = \beta_c(2). \end{cases}$$

The second inequality is proven in [43, Thm. 4.2]. By [49, Thm. 3.2] the first inequality is equivalent to a weak mixing property of the Potts measure (see [49, eq. (1.11)]). This weak mixing property is shown [2, Thm. 3.6] to hold whenever correlations decay exponentially

or, equivalently, we have exponential decay of connectivities in the corresponding infinite-volume random-cluster model, i.e., for all $u, v \in \mathbb{Z}^2$ we have

$$\lim_{L\to\infty} \mu_{p,q}^{\mathbb{Z}_L^2}(\{A \subset E_{L,2} : u \overset{A}{\longleftrightarrow} v\}) \leq \alpha_1 e^{-\alpha_2|u-v|}$$

with some $0 < \alpha_1(p,q), \alpha_2(p,q) < \infty$ and Euclidean norm $|\cdot|$. This was proven by Beffara and Duminil-Copin [3, Thm. 2] for all $q \geq 1$ and $p < p_c(q) := 1 - e^{-\beta_c(q)}$. Furthermore, the statements of this theorem hold true if we consider the case of 1-boundary condition. For this see [43, Thm. 1] and note that the result of [49, Thm. 3.2] holds for arbitrary boundary conditions. This proves the theorem. ∎

Remark 2.11. As stated above it is known that the heat-bath dynamics for the Ising model (without boundary conditions) is slow mixing if $\beta > \beta_c(2)$. Additionally, we are only aware of a result that shows an exponentially small upper bound on the spectral gap if $\beta > \beta_c(q)$ and q is large enough for some specific (periodic) boundary condition; see [5, Thm. 1.2] or Theorem 2.12 below. However, it is reasonable to believe that the HB dynamics is slowly mixing for all $q \geq 2$ and $\beta > \beta_c(q)$ on $\mathbb{Z}_L^2$ without (or with periodic) boundary condition.

We now turn to another class of underlying graphs, namely to rectangular subsets of the hypercubic lattice $\mathbb{Z}^d$. In fact, we consider only the case of *periodic* boundary condition. For this consider the *cycle* C_L of length L, that is the graph $C_L = (\{1, \dots, L\}, \widetilde{E}_L)$ with $\widetilde{E}_L := \{\{v, v+1\} : v \in \{1, \dots, L-1\}\} \cup \{1, L\}$, and define, for two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, the *graph product* of G_1 and G_2, written $G_1 \times G_2$, as the graph with vertex set $V_1 \times V_2$, where $(u_1, u_2), (v_1, v_2) \in V_1 \times V_2$ are neighbors in $G_1 \times G_2$ iff either u_1 and v_1 are neighbors in G_1 and $u_2 = v_2$, or u_2 and v_2 are neighbors in G_2 and $u_1 = v_1$. We then define the *d-dimensional torus* $\widetilde{\mathbb{Z}}_L^d$ of side length L by the d-fold graph product

$$\widetilde{\mathbb{Z}}_L^d := C_L^d = C_L \times \cdots \times C_L. \tag{2.32}$$

From Borgs, Chayes and Tetali [5] we obtain the following theorem.

Theorem 2.12. *Let P_{HB} (resp. P_{SW}) be the transition matrix of the heat-bath (resp. Swendsen–Wang) dynamics for the q-state Potts model on $\widetilde{\mathbb{Z}}_L^d$, $d \geq 2$, at inverse temperature β. Then there exist constants $k_1, k_2 < \infty$ and a constant $k_3 > 0$ (all depending on d, β and q) such that, for q and L large enough,*

$$e^{-(k_1+k_2\beta)L^{d-1}} \leq \lambda(P_{\mathrm{HB}}) \leq e^{-k_3\beta L^{d-1}} \quad \text{for all } \beta \geq \beta_c(d,q)$$

and

$$e^{-(k_1+k_2\beta)L^{d-1}} \leq \lambda(P_{\mathrm{SW}}) \leq e^{-k_3\beta L^{d-1}} \quad \text{for } \beta = \beta_c(d,q)$$

with $\beta_c(d,q)$ from (2.30)*. In fact, the lower bounds hold for all β, q and L.*

This theorem shows (at least for large q) that the heat-bath dynamics is slowly mixing at and below the critical temperature and, additionally, that also the Swendsen–Wang dynamics has an exponentially small spectral gap at the critical temperature if q is large enough. We will see in Chapter 4 that an analogous result holds for the single-bond dynamics for the random-cluster model.

Before we discuss other results for the Swendsen–Wang dynamics, we state a result for the HB dynamics on a more general class of graphs. For this, fix a graph $G = (V, E)$ and define the *adjacency matrix* A_G of G by $A_G(u, v) := \mathbb{1}(\{u, v\} \in E)$, $u, v \in V$. Additionally, write $\|A_G\|$ for its (unweighted) operator norm, i.e. $\|A_G\| = \max_{|x|=1} |A_G x|$, where the maximum is taken over $x \in \mathbb{R}^{|V|}$ and $|\cdot|$ is the Euclidean norm. In the literature $\|A_G\|$ is called the *principal eigenvalue* of the graph G. The following theorem is based on a result of Hayes [31].

THEOREM 2.13. *The heat-bath dynamics for the q-state Potts model at inverse temperature β on a graph G with n vertices satisfies*

$$\lambda(P_{\mathrm{HB}}) \geq \frac{1-\varepsilon}{n}$$

whenever $\beta \leq 2\varepsilon \|A_G\|^{-1}$.

Proof. Note first that one key ingredient for the result is Observation 11 of [31]. The result is stated only for the Ising model ($q = 2$), but it can be generalized quite easily by induction on q. We state the proof here for completeness. For this, we have to show, for all $u, v \in V$, that

$$\rho_{u,v} \leq \tanh(\beta/2) A_G(u, v),$$

where $\rho_{u,v}$, i.e. the *influence of v on u*, is defined by

$$\begin{aligned}\rho_{u,v} &:= \max_{\substack{\sigma \in \Omega_{\mathrm{P}} \\ \tau \in \Omega_{\sigma,v}}} \frac{1}{2} \sum_{k=1}^{q} \left| \frac{\pi(\sigma^{u,k})}{\sum_{l=1}^{q} \pi(\sigma^{u,l})} - \frac{\pi(\tau^{u,k})}{\sum_{l=1}^{q} \pi(\tau^{u,l})} \right| \\ &= \max_{\substack{\sigma \in \Omega_{\mathrm{P}} \\ \tau \in \Omega_{\sigma,v}}} \frac{1}{2} \sum_{k=1}^{q} \left| \frac{e^{\beta d_{u,k}(\sigma)}}{\sum_{l=1}^{q} e^{\beta d_{u,l}(\sigma)}} - \frac{e^{\beta d_{u,k}(\tau)}}{\sum_{l=1}^{q} e^{\beta d_{u,l}(\tau)}} \right|\end{aligned}$$

(see [31, Def. 4]). As before, $\Omega_{\sigma,v}$ is the set of configurations that differ from σ only at v (cf. (HB2) on page 20), and $d_{u,k}(\sigma)$ is the number of neighbors of u in G with color k in σ. Note that in [31] the above bound is stated with β in place of $\beta/2$. This comes from the difference in the normalization of the measure.

Obviously, $\rho_{u,v} = 0$ if $\{u, v\} \notin E$ since the term inside the absolute value depends only on the colors of the neighbors of u, which are equal if v is neither of them. Now fix some neighbors $u, v \in V$ and configurations $\sigma, \tau \in \Omega_{\mathrm{P}}$ with $\tau \in \Omega_{\sigma,v}$. Let $r_k := d_{u,k}(\sigma)$. Since σ and τ differ only at one neighbor of u, there exist $i, j \in [q]$ such that $d_{u,i}(\tau) = r_i + 1$, $d_{u,j}(\tau) = r_j - 1$ and $d_{u,k}(\tau) = r_k$ for all $k \neq i, j$. Assume without loss of generality that $i = 1$ and $j = 2$. Then some simple calculations show that

$$\frac{1}{2} \sum_{k=1}^{q} \left| \frac{e^{\beta d_{u,k}(\sigma)}}{\sum_{l=1}^{q} e^{\beta d_{u,l}(\sigma)}} - \frac{e^{\beta d_{u,k}(\tau)}}{\sum_{l=1}^{q} e^{\beta d_{u,l}(\tau)}} \right| = \frac{N_q(r)}{D_q(r)}, \qquad r = (r_1, \ldots, r_q),$$

where

$$N_q(r) := (e^{\beta} - e^{-\beta}) e^{\beta(r_1 + r_2)} + \max\{(e^{\beta} - 1) e^{\beta r_1}, (1 - e^{-\beta}) e^{\beta r_2}\} \sum_{k=3}^{q} e^{\beta r_k}$$

and

$$D_q(r) := \Big(\sum_{k=1}^{q} e^{\beta r_k}\Big)\Big(e^{\beta(r_1+1)} + e^{\beta(r_2-1)} + \sum_{l=3}^{q} e^{\beta r_l}\Big).$$

We will prove that $N_q(r)/D_q(r) \le \tanh(\beta/2)$ for all $r \in \mathbb{R}^q$ and $q \in \mathbb{N}$ by induction.

Let us first recall the $q = 2$ case from Observation 11 of [31]. In this case the last sum in the definition of N_2 and D_2 disappears. Thus, for $r \in \mathbb{R}^2$,

$$\begin{aligned}\frac{N_2(r)}{D_2(r)} &= \frac{(e^{\beta} - e^{-\beta})e^{\beta(r_1+r_2)}}{(e^{\beta r_1} + e^{\beta r_2})(e^{\beta(r_1+1)} + e^{\beta(r_2-1)})} = \frac{e^{\beta} - e^{-\beta}}{e^{\beta} + e^{-\beta} + e^{\beta(r_1-r_2+1)} + e^{-\beta(r_1-r_2+1)}} \\ &\le \frac{e^{\beta} - e^{-\beta}}{e^{\beta} + e^{-\beta} + 2} = \tanh(\beta/2),\end{aligned}$$

where the last inequality comes from $e^x + e^{-x} \ge 2$, $x \in \mathbb{R}$. Now assume that the statement holds for $q-1$ and all $s = (r_1, \dots, r_{q-1}) \in \mathbb{R}^{q-1}$ and let $r = (r_1, \dots, r_{q-1}, r_q)$ for some $r_q \in \mathbb{R}$. We obtain

$$\begin{aligned}\frac{N_q(r)}{D_q(r)} &= \frac{N_{q-1}(s) + e^{\beta r_q}\max\{(e^{\beta}-1)e^{\beta r_1}, (1-e^{-\beta})e^{\beta r_2}\}}{D_{q-1}(s) + (e^{\beta}+1)e^{\beta(r_1+r_q)} + (1+e^{-\beta})e^{\beta(r_2+r_q)} + e^{2\beta r_q} + 2e^{\beta r_q}\sum_{k=3}^{q-1} e^{\beta r_k}} \\ &\le \frac{N_{q-1}(s) + e^{\beta r_q}\max\{(e^{\beta}-1)e^{\beta r_1}, (1-e^{-\beta})e^{\beta r_2}\}}{D_{q-1}(s) + e^{\beta r_q}[(e^{\beta}+1)e^{\beta r_1} + (1+e^{-\beta})e^{\beta r_2}]} \le \tanh\Big(\frac{\beta}{2}\Big),\end{aligned}$$

since $a/b \le t$ and $c/d \le t$ imply $(a+c)/(b+d) \le t$ for $a, b, c, d, t \ge 0$. This proves Observation 11 of [31] for all $q \in \mathbb{N}$ and thus, under the assumptions of this theorem, we find by [31, Thm. 6] that

$$\max_{\sigma \in \Omega_{\mathrm{P}}(G)} \|P_{\mathrm{HB}}^t(\sigma, \cdot) - \pi\|_{\mathrm{TV}} \le n\Big(1 - \frac{1-\varepsilon}{n}\Big)^t.$$

Using Lemma 2.2 we obtain the result. ∎

In particular we have the following corollary (see [31]).

Corollary 2.14. *The heat-bath dynamics for the q-state Potts model at inverse temperature β on a graph G with n vertices and maximum degree Δ satisfies*

$$\lambda(P_{\mathrm{HB}}) \ge \frac{1-\varepsilon}{n}$$

if $\beta \le 2\varepsilon/\Delta$.

Proof. Using [33, Thm. 8.1.22] we obtain $\|A_G\| \le \max_{u \in V} \sum_{v \in V} A_G(u, v) = \Delta(G)$. Thus, the result follows from Theorem 2.13. ∎

In [31] one can find also an improvement of this corollary if we restrict to the class of *planar* graphs. We will use this in Chapter 5 to obtain a result for the Swendsen–Wang dynamics on planar graphs (see Corollary 5.7).

Remark 2.15. There are, of course, a lot of other results concerning mixing properties of the heat-bath dynamics for the Potts model. These include, e.g., rapid mixing for heat-bath dynamics on trees at all temperatures [4]. Additionally, only recently was the complete picture of rapid mixing or lack thereof established for the HB dynamics on the *complete graph* [12, 38].

Now we want to state some known results on the mixing properties of the Swendsen–Wang dynamics. We try to give an overview of all known results, but we omit results that involve boundary conditions (or external magnetic field) (see e.g. [44, 50, 51]), and results where the underlying graph is random (see [11]).

We start with the known results on the complete graph, where at least for $q = 2$ the mixing behavior is completely known. Denote by K_n the *complete graph* on n vertices, i.e. $K_n = ([n], \binom{[n]}{2})$, where $\binom{[n]}{2}$ is the set of all two-element subsets of $[n]$. For two non-negative functions $f, g : \mathbb{N} \to \mathbb{R}$, we write $f(n) = \Theta(g(n))$ iff $0 < \lim_{n\to\infty} \frac{f(n)}{g(n)} < \infty$. The result below is due to Long, Nachmias and Peres [42, 41] in the case $q = 2$ (see also [10]). For $q \geq 3$ the results are adopted from Gore and Jerrum [28] and Huber [34].

THEOREM 2.16. *Let P_{SW} be the transition matrix of the Swendsen–Wang dynamics for the q-state Potts model on the complete graph K_n at inverse temperature $\beta = \log\left(\frac{n}{n-c}\right)$, $c \geq 0$ (or $p = 1 - e^{-\beta} = c/n$). Then, with $t_{\mathrm{mix}} := t_{\mathrm{mix}}(P_{\mathrm{SW}})$,*

(i) *for $q = 2$ and*

- $c < 2$: $t_{\mathrm{mix}} = \Theta(1)$,
- $c = 2$: $t_{\mathrm{mix}} = \Theta(n^{1/4})$,
- $c > 2$: $t_{\mathrm{mix}} = \Theta(\log n)$;

(ii) *for $q \geq 3$ and*

- $c = \dfrac{2(q-1)\log(q-1)}{q-2}$: $t_{\mathrm{mix}} \geq e^{\varepsilon\sqrt{n}}$ *for some $\varepsilon > 0$,*
- $c < \dfrac{1}{3}$: $t_{\mathrm{mix}} \leq C \log n$ *for some $C < \infty$,*
- $c > 2q\log(3qn)$: $t_{\mathrm{mix}} \leq Cqn$ *for some $C < \infty$.*

As far as we know, Theorems 2.12 and 2.16(ii) contain the only presently known slow mixing results for the Swendsen–Wang dynamics (except for results on random graphs). Other classes of graphs where rapid mixing of SW is known at all temperatures, but now for every $q \in \mathbb{N}$, are trees and cycles (see Cooper and Frieze [11] and Long [41]).

THEOREM 2.17. *Let T be a tree on n vertices and C_n be a cycle of length n. Then*

- $t_{\mathrm{mix}}(P^T_{\mathrm{SW},\beta,q}) = \Theta(\log n)$, *and*
- $t_{\mathrm{mix}}(P^{C_n}_{\mathrm{SW},\beta,q}) \leq cn \log n$ *for some $c < \infty$.*

In fact, we have $\lambda(P^T_{\mathrm{SW},\beta,q}) = 1 - p(1 - 1/q)$.

For the statement on the spectral gap consider the construction of [41, Chap. 7] and Exercise 12.7 of [39]. The last result that we want to present here for the SW dynamics is, similarly to Corollary 2.14, a result on graphs of bounded maximum degree from [34] (see also [11]).

THEOREM 2.18. *The Swendsen–Wang dynamics for the q-state Potts model at inverse temperature β on a graph G with n vertices and maximum degree Δ satisfies*

$$t_{\mathrm{mix}}(P_{\mathrm{SW}}) \leq C \log n \quad \text{for some } C < \infty$$

whenever $\beta \leq \frac{1}{3(\Delta-1)}$.

Now we turn to a result for the single-bond (or heat-bath) dynamics for the random-cluster model. In fact, this is the only result on the mixing properties of this Markov chain that we are aware of.

For this define the *linear width* of a graph $G = (V, E)$ as the smallest number ℓ such that there exists an ordering $e_1, \dots, e_{|E|}$ of the edges with the property that for every $i \in [|E|]$ there are at most ℓ vertices that are an endvertex of both an edge in $\{e_1, \dots e_i\}$ and an edge in $\{e_{i+1}, \dots e_{|E|}\}$. See [25] for bounds on the linear width of paths, cycles, trees and a bound in terms of a related quantity, the tree width.

The following result is due to Ge and Štefankovič [25] (we only state the $q \geq 1$ case).

THEOREM 2.19. *Let $\widetilde{P}_{\mathrm{SB}}$ be the transition matrix of the single-bond dynamics for the random-cluster model with parameters p and $q \geq 1$ on a graph $G = (V, E)$ with linear width bounded by ℓ. Let $m := |E|$. Then*

$$\lambda(\widetilde{P}_{\mathrm{SB}}) \geq \frac{1}{2q^{\ell+1}} \frac{1}{m^2}. \tag{2.33}$$

Proof. In [25] the authors consider the (lazy) Metropolis version of the single-bond dynamics. This Markov chain has transition probabilities

$$\widetilde{P}_{\mathrm{M}}(A, A^e) = \frac{1}{2|E|} \min\left\{1, q^{c(A^e)-c(A)} \left(\frac{p}{1-p}\right)^{|A^e|-|A|}\right\}, \quad A \subset E,$$

with $\widetilde{P}_{\mathrm{M}}(A, A)$ such that $\widetilde{P}_{\mathrm{M}}$ is a stochastic matrix and with A^e from (2.23). For this Markov chain they prove a lower bound on the congestion, which is defined as follows. Let $\Gamma := \{\gamma_{AB} : A, B \subset E\}$, where γ_{AB} are paths from A to B in the graph $\mathcal{H} = (\Omega_{\mathrm{RC}}, \mathcal{E})$ with $\mathcal{E} = \{(A, B) : \widetilde{P}_{\mathrm{M}}(A, B) > 0\}$. Then we define the *congestion* of $\widetilde{P}_{\mathrm{M}}$ (with respect to Γ) by

$$\varrho(\widetilde{P}_{\mathrm{M}}, \Gamma) := \max_{(B_1,B_2)\in\mathcal{E}} \frac{1}{\mu(B_1)\widetilde{P}_{\mathrm{M}}(B_1, B_2)} \sum_{A,C:(B_1,B_2)\in\gamma_{AC}} |\gamma_{AC}|\mu(A)\mu(C),$$

where $|\gamma_{AC}|$ denotes the length of the path. The bound of [25, Lemma 16] is $\varrho(\widetilde{P}_{\mathrm{M}}, \Gamma) \leq 2|E|^2 q^\ell$ for a suitable choice of Γ, and so, by [16, Prop. 1′] (note that $\widetilde{P}_{\mathrm{M}}$ is lazy),

$$\lambda(\widetilde{P}_{\mathrm{M}})^{-1} \leq 2|E|^2 q^\ell.$$

But since it is easy to show that $\widetilde{P}_{\mathrm{M}}(A, B) \leq q\widetilde{P}_{\mathrm{SB}}(A, B)$ for all $A \neq B \subset E$ and that $\widetilde{P}_{\mathrm{SB}}$ has only non-negative eigenvalues (see Remark 4.4), we can conclude from Lemma 2.5 that

$$\lambda(P_{\mathrm{SB}})^{-1} \leq q\lambda(P_{\mathrm{M}})^{-1} \leq 2|E|^2 q^{\ell+1}. \quad \blacksquare$$

Finally, we want to mention an algorithm that allows approximate sampling in polynomial time from the Ising model on arbitrary graphs and at all temperatures. This algorithm is due to Randall and Wilson [58] and is based on the seminal work of Jerrum and Sinclair [35], which shows the first (and presently the only) polynomial-time algorithm to approximate the partition function of an arbitrary (ferromagnetic) Ising system. We are not aware of an explicit bound on the expected running time of this sampling procedure, but there is a bound for the algorithm for the partition function in [35]. Maybe

due to its simplicity, the Swendsen–Wang dynamics is, however, still the preferred algorithm in practice and, needless to say, it would be amazing to obtain rapid mixing of this Markov chain in the same regime.

3. Comparison with single-spin dynamics

This chapter is based on [68]. We prove by comparison that the spectral gap of the Swendsen–Wang dynamics (SW) is bounded from below by some constant times the spectral gap of the heat-bath chain (HB). This result leads to rapid mixing of SW on graphs of bounded degree whenever HB mixes rapidly.

We will prove the following theorem, which is a minor improvement of [68, Thm. 1].

THEOREM 3.1. *Suppose that P_{SW} (resp. P_{HB}) is the transition matrix of the Swendsen–Wang (resp. heat-bath) dynamics for the q-state Potts model at inverse temperature β on a graph G with maximum degree Δ. Then*

$$\lambda(P_{\mathrm{SW}}) \geq c_{\mathrm{sw}} \lambda(P_{\mathrm{HB}}),$$

where

$$c_{\mathrm{sw}} = c_{\mathrm{sw}}(\Delta, \beta, q) := q^{-1}(qe^{2\beta})^{-2\Delta}. \tag{3.1}$$

REMARK 3.2. The inequality of Theorem 3.1 is probably off by a factor of $|V(G)|$, because we compare the SW dynamics with a Markov chain that changes only the color of one vertex of the graph per step. We conjecture that a bound of the form $\lambda(P_{\mathrm{SW}}) \geq c|V(G)|\lambda(P_{\mathrm{HB}})$, for some constant $c > 0$, holds and that, in particular, this constant c has a "better" dependence on the parameters involved. Unfortunately, this does not seem to be possible to show with our techniques.

Before we prove Theorem 3.1 we state some corollaries that can be deduced directly from the known results for the heat-bath dynamics (see Section 2.5).

The first corollary deals with the class of graphs with bounded maximum degree. It relies on a slight generalization of a result of Hayes [31], who gives a simple condition on β, depending on the maximum degree, for rapid mixing of the heat-bath dynamics for the Ising model (see Theorem 2.13 and Corollary 2.14).

COROLLARY 3.3. *The Swendsen–Wang dynamics for the q-state Potts model at inverse temperature β on a graph G with n vertices and maximum degree Δ satisfies*

$$\lambda(P_{\mathrm{SW}}) \geq c_{\mathrm{sw}}(1-\varepsilon)/n$$

with $c_{\mathrm{sw}} = c_{\mathrm{sw}}(\Delta, \beta, q)$ from (3.1) and $\varepsilon > 0$, if

$$\beta \leq 2\varepsilon/\Delta.$$

This result improves that of Huber [34] (see Theorem 2.18) in the range of applicability, which was $\beta \leq 1/(3(\Delta-1))$ before, but while the result of [34] is a logarithmic (in n) upper bound on the mixing time, Corollary 3.3 together with Corollary 2.8 leads only to

a quadratic bound. We will see a further improvement of the above result in Chapter 5 if we consider only *planar* graphs (see Corollary 5.7).

The second corollary, which we call a theorem because of its importance for the rest of this work, gives a bound on the spectral gap for the Swendsen–Wang dynamics on the square lattice. For this recall that the two-dimensional square lattice of side length L is the graph $\mathbb{Z}_L^2 = (V_{L,2}, E_{L,2})$ with vertex set $V_{L,2} = \{1, \dots, L\}^2 \subset \mathbb{Z}^2$ and edge set $E_{L,2} = \{\{u, v\} \subset V_{L,2} : |u - v| = 1\}$ (see Figure 1).

THEOREM 3.4. *Let P_{SW} be the transition matrix of the Swendsen–Wang dynamics for the q-state Potts model on $\mathbb{Z}_L^2$ at inverse temperature β. Let $n = L^2 = |V_{L,2}|$. Then there exist constants $c_\beta = c_\beta(q) > 0$ and $C < \infty$ such that*

$$\lambda(P_{\mathrm{SW}}) \geq c_\beta / n \qquad \text{for } \beta < \beta_c(q),$$

and

$$\lambda(P_{\mathrm{SW}}) \geq c_{\mathrm{sw}} n^{-C} \quad \text{for } q = 2 \text{ and } \beta = \beta_c(2),$$

with $c_{\mathrm{sw}} = c_{\mathrm{sw}}(4, \beta_c(2), 2)$ from (3.1) and $\beta_c(q) = \log(1 + \sqrt{q})$.

This result is a consequence of Theorem 3.1 and the corresponding result for the heat-bath dynamics (see Theorem 2.10 as well as the references in its proof).

REMARK 3.5. Note that only the presence of q under the exponent in the definition of c_{sw} from (3.1) prevents us from the application of Theorem 3.1 to the complete graph K_n. This comes from $\Delta(K_n) = n - 1$ and the usual normalization of the inverse temperature to $\beta = c/n$ for some $c > 0$ (cf. Theorem 2.16), which would lead (without the q) to a lower bound on $c_{\mathrm{sw}}(K_n, c/n, q)$ independent of n.

3.1. Proof of Theorem 3.1. The proof is based on standard techniques for the comparison of Markov chains (see Lemma 2.5), together with an appropriate choice of an auxiliary Markov chain that can be compared to both Swendsen–Wang and heat-bath dynamics. For the remainder of this section fix a graph $G = (V, E)$, some $\beta \geq 0$ and $q \in \mathbb{N}$, and recall that we denote by π the measure for the q-state Potts model on G at inverse temperature β (see (2.9)). We will analyze the auxiliary Markov chain with transition probability matrix

$$Q = P_{\mathrm{HB}} P_{\mathrm{SW}} P_{\mathrm{HB}} \tag{3.2}$$

where P_{HB} is from (2.18) and P_{SW} from (2.21). Since P_{HB} and P_{SW} are reversible with respect to π, we see that Q is also reversible.

The first lemma shows that a Markov chain with transition matrix Q has a larger spectral gap than the heat-bath dynamics.

LEMMA 3.6. *With the definitions above we get*

$$\lambda(Q) \geq \lambda(P_{\mathrm{HB}}).$$

Proof. With $S_\pi(\sigma,\tau) = \pi(\tau)$ for $\sigma,\tau \in \Omega_{\mathrm{P}}$, we have $Q - S_\pi = (P_{\mathrm{HB}} - S_\pi)P_{\mathrm{SW}}(P_{\mathrm{HB}} - S_\pi)$, which is self-adjoint. Hence we can write the spectral gap (see (2.8)) as

$$\begin{aligned} 1 - \lambda(Q) &= \|Q - S_\pi\|_\pi = \|(P_{\mathrm{HB}} - S_\pi)P_{\mathrm{SW}}(P_{\mathrm{HB}} - S_\pi)\|_\pi \\ &\le \|P_{\mathrm{HB}} - S_\pi\|_\pi^2 \|P_{\mathrm{SW}}\|_\pi \le \|P_{\mathrm{HB}} - S_\pi\|_\pi = 1 - \lambda(P_{\mathrm{HB}}), \end{aligned}$$

where we use submultiplicativity of the spectral norm as well as $\|P_{\mathrm{SW}}\|_\pi \le 1$ and $\|P_{\mathrm{HB}} - S_\pi\|_\pi \le 1$. ∎

To prove a lower bound on the spectral gap of P_{SW} it remains to prove $\lambda(P_{\mathrm{SW}}) \ge c\lambda(Q)$ for some $c > 0$. For this we need an estimate of the transition probabilities of the Swendsen–Wang dynamics on G with respect to some subgraph of G. Therefore we denote the transition matrix of the Swendsen–Wang dynamics for the q-state Potts model on a graph G at inverse temperature β throughout this section by P_G, i.e.

$$P_G := P^G_{\mathrm{SW},\beta,q} \tag{3.3}$$

(cf. (2.21)). We prove the following lemma.

Lemma 3.7. *Let $G = (V,E)$ be a graph and $G_0 = (V,E_0)$ be a spanning subgraph of G with $E_0 \subset E$. Then*

$$a_1^{|E\setminus E_0|} P_{G_0}(\sigma,\tau) \le P_G(\sigma,\tau) \le a_2^{|E\setminus E_0|} P_{G_0}(\sigma,\tau)$$

for all $\sigma,\tau \in \Omega_{\mathrm{P}}$, where

$$a_1 = a_1(\beta) := e^{-\beta} \quad \text{and} \quad a_2 = a_2(\beta,q) := 1 + q(e^\beta - 1).$$

Proof. The first inequality is already known from the proof of Lemma 3.3 in [5], but we state it here for completeness. Let $p = 1 - e^{-\beta}$ and note that $E_0(\sigma) \subset E(\sigma)$ for all $\sigma \in \Omega_{\mathrm{P}}$. We deduce by (2.21) that

$$\begin{aligned} P_G(\sigma,\tau) &= \sum_{A\subset E} p^{|A|}(1-p)^{|E(\sigma)|-|A|} q^{-c(A)} \mathbb{1}(A \subset E(\sigma) \cap E(\tau)) \\ &\ge \sum_{A\subset E_0} p^{|A|}(1-p)^{|E(\sigma)|-|A|} q^{-c(A)} \mathbb{1}(A \subset E_0(\sigma) \cap E_0(\tau)) \\ &= (1-p)^{|E(\sigma)|-|E_0(\sigma)|} P_{G_0}(\sigma,\tau) \ge (1-p)^{|E\setminus E_0|} P_{G_0}(\sigma,\tau). \end{aligned}$$

For the second inequality suppose for now $E_0 = E \setminus e$ with some $e \in E$ and note that $c(A \cup e) \ge c(A) - 1$. We get

$$\begin{aligned} P_G(\sigma,\tau) &= \sum_{A\subset E(\sigma)\cap E(\tau)} p^{|A|}(1-p)^{|E(\sigma)|-|A|} q^{-c(A)} \\ &= \sum_{\substack{A\subset E(\sigma)\cap E(\tau):\\ e\in A}} p^{|A|}(1-p)^{|E(\sigma)|-|A|} q^{-c(A)} \\ &\quad + \sum_{\substack{A\subset E(\sigma)\cap E(\tau):\\ e\notin A}} p^{|A|}(1-p)^{|E(\sigma)|-|A|} q^{-c(A)} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{A' \subset E_0(\sigma) \cap E_0(\tau)} p^{|A' \cup e|} (1-p)^{|E(\sigma)| - |A' \cup e|} q^{-c(A' \cup e)} \\
&\quad + \sum_{A' \subset E_0(\sigma) \cap E_0(\tau)} p^{|A'|} (1-p)^{|E(\sigma)| - |A'|} q^{-c(A')} \\
&\le \frac{qp}{1-p} \sum_{A' \subset E_0(\sigma) \cap E_0(\tau)} p^{|A'|} (1-p)^{|E_0(\sigma)| - |A'|} q^{-c(A')} + P_{G_0}(\sigma, \tau) \\
&= \left(\frac{qp}{1-p} + 1 \right) P_{G_0}(\sigma, \tau) = (1 + q(e^\beta - 1)) P_{G_0}(\sigma, \tau).
\end{aligned}
$$

For $|E \setminus E_0| > 1$ one can iterate this procedure $|E \setminus E_0|$ times. ∎

We use this lemma to prove that the transition probability from σ to τ is similar to the probability of going from a neighbor of σ to a neighbor of τ. Recall that $\sigma^{v,k}$ is defined by $\sigma^{v,k}(v) = k \in [q]$ and $\sigma^{v,k}(u) = \sigma(u)$, $u \neq v$.

LEMMA 3.8. *Let $\sigma, \tau \in \Omega_{\mathrm{P}}$, $v \in V$ and $k, l \in [q]$. Then*

$$
P_G(\sigma^{v,k}, \tau^{v,l}) \le a_3^{\deg_G(v)} P_G(\sigma, \tau) \quad \textit{with} \quad a_3 = a_3(\beta, q) := q e^{2\beta} - (q-1) e^\beta,
$$

where $\deg_G(v)$ denotes the degree of the vertex v in G.

Proof. Define $E_v := \{e \in E : v \in e\}$ and $G_v := (V, E \setminus E_v)$. Then $v \in V$ is an isolated vertex in G_v. By the definition of the Swendsen–Wang dynamics we get

$$
P_{G_v}(\sigma^{v,k}, \tau^{v,l}) = P_{G_v}(\sigma, \tau).
$$

If we set $E_0 = E \setminus E_v$ we get $|E \setminus E_0| = \deg_G(v)$. We deduce by Lemma 3.7 that

$$
P_G(\sigma^{v,k}, \tau^{v,l}) \le a_2^{\deg_G(v)} P_{G_v}(\sigma^{v,k}, \tau^{v,l}) = a_2^{\deg_G(v)} P_{G_v}(\sigma, \tau) \le \left(\frac{a_2}{a_1} \right)^{\deg_G(v)} P_G(\sigma, \tau)
$$

with a_1 and a_2 from Lemma 3.7. ∎

Now we are able to prove the main result of this chapter.

Proof of Theorem 3.1. Because of Lemma 3.6 we only have to prove $\lambda(P_{\mathrm{SW}}) \ge c_{\mathrm{SW}} \lambda(Q)$. Let

$$
c := \max_{\substack{\sigma_1, \sigma_2, \tau_1, \tau_2 \in \Omega_{\mathrm{P}}: \\ \sigma_1 \sim \sigma_2,\, \tau_1 \sim \tau_2}} \frac{P_{\mathrm{SW}}(\sigma_1, \tau_1)}{P_{\mathrm{SW}}(\sigma_2, \tau_2)}, \tag{3.4}
$$

where $\sigma \sim \tau :\Leftrightarrow \sum_{v \in V} \mathbb{1}(\sigma(v) \neq \tau(v)) \le 1$, i.e. σ and τ differ in at most one vertex. Note that $P_{\mathrm{HB}}(\sigma, \tau) \neq 0$ if and only if $\sigma \sim \tau$. We find for $\sigma_1, \tau_1 \in \Omega_{\mathrm{P}}$ that

$$
\begin{aligned}
Q(\sigma_1, \tau_1) &= \sum_{\sigma_2, \tau_2 \in \Omega_{\mathrm{P}}} P_{\mathrm{HB}}(\sigma_1, \sigma_2) P_{\mathrm{SW}}(\sigma_2, \tau_2) P_{\mathrm{HB}}(\tau_2, \tau_1) \\
&\le c P_{\mathrm{SW}}(\sigma_1, \tau_1) \sum_{\sigma_2 \sim \sigma_1} P_{\mathrm{HB}}(\sigma_1, \sigma_2) \sum_{\tau_2 \sim \tau_1} P_{\mathrm{HB}}(\tau_2, \tau_1) \le qc P_{\mathrm{SW}}(\sigma_1, \tau_1).
\end{aligned}
$$

Using Lemma 2.5 we obtain

$$
\lambda(Q) \le qc \lambda(P_{\mathrm{SW}}).
$$

It remains to bound c. Recall that $\deg_G(v) \le \Delta$ for all $v \in V(G)$. With a_3 from Lemma 3.8 we get, for $\sigma_1, \sigma_2, \tau_1, \tau_2 \in \Omega_{\mathrm{P}}$ with $\sigma_1 \sim \sigma_2$ and $\tau_1 \sim \tau_2$,

$$\frac{P_{\mathrm{SW}}(\sigma_1,\tau_1)}{P_{\mathrm{SW}}(\sigma_2,\tau_2)} \le a_3^{\Delta} \frac{P_{\mathrm{SW}}(\sigma_2,\tau_1)}{P_{\mathrm{SW}}(\sigma_2,\tau_2)} \le a_3^{2\Delta} \frac{P_{\mathrm{SW}}(\sigma_2,\tau_2)}{P_{\mathrm{SW}}(\sigma_2,\tau_2)} = a_3^{2\Delta}.$$

Finally,

$$\lambda(Q) \le qc\lambda(P_{\mathrm{SW}}) \le qa_3^{2\Delta}\lambda(P_{\mathrm{SW}}) \le q(qe^{2\beta})^{2\Delta}\lambda(P_{\mathrm{SW}}). \ \blacksquare$$

3.2. A slight generalization. In this section we present a generalization of Theorem 3.1 that is necessary to handle also graphs with a single vertex of maximal degree. The idea behind this modification is that the transition probabilities of the Swendsen–Wang dynamics are invariant under global flips of the color of all vertices, i.e.

$$P_{\mathrm{SW}}(\sigma,\tau) = P_{\mathrm{SW}}(\sigma + k, \tau + l), \quad \sigma,\tau \in \Omega_{\mathrm{P}},\ k,l \in [q], \tag{3.5}$$

where $(\sigma + k)(v) := (\sigma(v) + k - 1 \bmod q) + 1$ for all $v \in V$. To see this, note that $P_{\mathrm{SW}}(\sigma,\tau)$ (cf. (2.21)) does not depend on the precise colors in σ and τ, but on the "edges of agreement" $E(\sigma)$ and $E(\tau)$ (see (2.10)), which are invariant under global flips. If we now consider the heat-bath dynamics with additional global flips, i.e. in each step make one step using P_{HB} and then change the color of all vertices at once by a random increment, it is reasonable to conjecture that SW has also an (almost) larger spectral gap than this Markov chain. In the following theorem we prove this statement, but in a different form, namely, we consider the original Markov chains under the condition that the color of the configurations is fixed at a single vertex. Because of symmetry we can assume that the fixed color equals 1. This can be interpreted as "boundary conditions" for the neighbors of this single vertex.

For some fixed vertex $w \in V$ in the graph $G = (V, E)$ we denote by

$$\Lambda_w := \{1, \dots, q\}^{V \setminus w}$$

the set of all colorings of the vertices $V \setminus w$ and define a probability measure on Λ_w by

$$\pi^w(\bar\sigma) := q\pi(\bar\sigma^1), \quad \bar\sigma \in \Lambda_w, \tag{3.6}$$

with π from (2.9), where $\bar\sigma^1 \in \Omega_{\mathrm{P}}$ is a coloring of V such that $\bar\sigma^1(u) = \bar\sigma(u)$, $u \ne w$, and $\bar\sigma^1(w) = 1$. One may think of π^w as the conditional probability measure on the colorings of the vertices $V \setminus w$ with respect to π given that w is colored 1.

It is not difficult to show (using ideas similar to those of the proof of Theorem 3.9 below) that the Markov chain on Λ_w with transition matrix $\bar P(\bar\sigma,\bar\tau) := qP_{\mathrm{SW}}(\sigma,\tau)$, where $\sigma,\tau \in \Omega_{\mathrm{P}}$ and $\bar\sigma,\bar\tau \in \Lambda_w$ are such that

$$\bar\sigma(v) := (\sigma(v) - \sigma(w) \bmod q) + 1, \quad v \in V \setminus w,$$

is reversible with respect to π^w and has the same spectral gap as P_{SW}. Additionally, we define the heat-bath dynamics that is reversible with respect to π^w similarly to (2.18) by

$$P^w_{\mathrm{HB}}(\bar\sigma,\bar\tau) := P^{G,w,1}_{\mathrm{HB},\beta,q}(\bar\sigma,\bar\tau) = \frac{1}{|V|-1} \sum_{v \in V\setminus w} \frac{\pi^w(\bar\tau)}{\sum_{l=1}^q \pi^w(\bar\sigma^{v,l})} \mathbb{1}(\bar\tau \in \Omega^w_{\sigma,v}), \tag{3.7}$$

where $\Omega^w_{\bar\sigma,v} := \{\bar\tau \in \Lambda_w : \bar\tau(u) = \bar\sigma(u),\ \forall u \ne v\}$. The following theorem gives a comparison inequality between P_{SW} and P^w_{HB} that will be essential in utilizing results on the heat-

bath dynamics for the Potts model with boundary conditions (cf. (2.29)) for the analysis of the SW dynamics.

THEOREM 3.9. *Suppose that P_{SW} is the transition matrix of the Swendsen–Wang dynamics for the q-state Potts model at inverse temperature β on a graph $G = (V, E)$. Furthermore, let $w \in V$ be any vertex and P^w_{HB} be as in* (3.7)*. Then*

$$\lambda(P_{\mathrm{SW}}) \geq \widetilde{c}_{\mathrm{sw}} \lambda(P^w_{\mathrm{HB}}) \quad \textit{with} \quad \widetilde{c}_{\mathrm{sw}} = \widetilde{c}_{\mathrm{sw}}(G, \beta, q) := q^{-1}(qe^{2\beta})^{-2\widetilde{\Delta}},$$

where

$$\widetilde{\Delta} := \max_{u \in V \setminus w} \deg_G(u)$$

is the maximum degree over $V \setminus w$.

Before we prove Theorem 3.9, we present an application of it. For this recall the definition of the two-dimensional square lattice $\mathbb{Z}^2_L = (V_{L,2}, E_{L,2})$ of side length L from Section 2.5, and that the boundary of $V_{L,2}$ is defined by $\partial V_{L,2} := \{v = (v_1, v_2) \in V_{L,2} : v_i \in \{1, L\} \text{ for some } i \in [2]\}$. Now we introduce a new auxiliary vertex v^* and let $\mathbb{Z}^{2\dagger}_L$ be the graph $\mathbb{Z}^2_{L-1}$ with additional vertex v^* and edges between v^* and all boundary vertices $u \in \partial V_{L-1,2}$. That is, $\mathbb{Z}^{2\dagger}_L := (V^\dagger_{L,2}, E^\dagger_{L,2}, \varphi)$ is the (multi-)graph (cf. Section 2.3) with vertex set $V^\dagger_{L,2} = V_{L-1,2} \cup v^*$ and edge set $E^\dagger_{L,2}$ such that the set of endpoints of the edges in $E^\dagger_{L,2}$ satisfies $\varphi(E^\dagger_{L,2}) = E_{L-1,2} \cup \{\{v^*, u\} : u \in \partial V_{L-1,2}\}$ and, for each of the vertices $(1,1), (1, L-1), (L-1, 1), (L-1, L-1) \in V^\dagger_{L,2}$, there are two parallel edges to v^* (see Figure 2). Now we can deduce the following directly from Theorem 2.10.

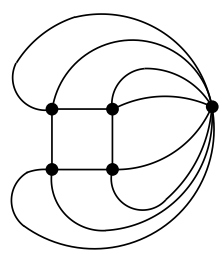

Fig. 2. The graph $\mathbb{Z}^{2\dagger}_3$

COROLLARY 3.10. *Let P_{SW} be the transition matrix of the Swendsen–Wang dynamics for the q-state Potts model on $\mathbb{Z}^{2\dagger}_L$ at inverse temperature β. Let $n = |V^\dagger_{L,2}|$. Then there exists a constant $c_\beta = c_\beta(q) > 0$ such that*

$$\lambda(P_{\mathrm{SW}}) \geq c_\beta / n \qquad \textit{for } \beta < \beta_c(q)$$

with $\beta_c(q) = \log(1 + \sqrt{q})$.

Proof. Let the auxiliary vertex v^* be the vertex with fixed color from Theorem 3.9. Then we see that $P^{v^*}_{\mathrm{HB}}$ from (3.7) equals $P_{\mathrm{HB},1}$ from (2.31). Thus, the result follows from Theorem 2.10. ∎

We finish this chapter with the proof of Theorem 3.9.

Proof of Theorem 3.9. The proof is very similar to the proof of Theorem 3.1. First, recall that for $\sigma \in \Omega_{\mathrm{P}}$, $\bar{\sigma}(v) = (\sigma(v) - \sigma(w) \bmod q) + 1$, $v \in V \setminus w$, with $\bar{\sigma} \in \Lambda_w$. Additionally, for $\bar{\sigma} \in \Lambda_w$, we denote by $\bar{\sigma}^1 \in \Omega_{\mathrm{P}}$ the configuration with $\bar{\sigma}^1(v) = \bar{\sigma}(v)$, $v \neq w$, and $\bar{\sigma}^1(w) = 1$.

We define, for $\sigma \in \Omega_{\mathrm{P}}$ and $\bar{\tau} \in \Lambda_w$, the "flip" transition matrices

$$F_1(\sigma, \bar{\tau}) := \mathbb{1}(\bar{\tau} = \bar{\sigma}) \quad \text{and} \quad F_2(\bar{\tau}, \sigma) := \frac{1}{q} \sum_{l=0}^{q-1} \mathbb{1}(\sigma = \bar{\tau}^1 + l).$$

(If we consider F_1 as an operator mapping from $L_2(\pi^w)$ to $L_2(\pi)$, then $F_2 = F_1^*$.) With π from (2.9) and π^w from (3.6), it is easy to check that $\pi F_1 = \pi^w$ and $\pi^w F_2 = \pi$. Following the same ideas as in Section 3.1 with

$$Q = F_1 P_{\mathrm{HB}}^w F_2 P_{\mathrm{SW}} F_1 P_{\mathrm{HB}}^w F_2,$$

which is reversible with respect to π, we get

$$\|Q - S_\pi\|_\pi \leq \|F_1 P_{\mathrm{HB}}^w F_2 - S_\pi\|_\pi = \|F_1 (P_{\mathrm{HB}}^w - S_{\pi^w}) F_2\|_\pi \leq \|P_{\mathrm{HB}}^w - S_{\pi^w}\|_{\pi^w}.$$

This proves $\lambda(Q) \geq \lambda(P_{\mathrm{HB}}^w)$. It remains to show that $\lambda(P_{\mathrm{SW}}) \geq \widetilde{c}_{\mathrm{SW}} \lambda(Q)$. By the construction of the Swendsen–Wang dynamics we know that

$$P_{\mathrm{SW}}(\sigma, \tau) = P_{\mathrm{SW}}(\bar{\sigma}^1, \bar{\tau}^1), \quad \forall \sigma, \tau \in \Omega_{\mathrm{P}}.$$

Hence we deduce with

$$\widetilde{c} := \max_{\substack{\sigma_1, \sigma_2, \tau_1, \tau_2 \in \Omega_{\mathrm{P}}: \\ \bar{\sigma}_1^1 \sim \bar{\sigma}_2^1, \, \bar{\tau}_1^1 \sim \bar{\tau}_2^1}} \frac{P_{\mathrm{SW}}(\sigma_1, \tau_1)}{P_{\mathrm{SW}}(\sigma_2, \tau_2)}$$

(cf. (3.4)) that, for all $\sigma, \tau \in \Omega_{\mathrm{P}}$,

$$\begin{aligned}
Q(\sigma, \tau) &= \sum_{\substack{\bar{\sigma}_1, \bar{\tau}_1 \in \Lambda_w \\ \sigma_2, \tau_2 \in \Omega_{\mathrm{P}}}} P_{\mathrm{HB}}^w(\bar{\sigma}, \bar{\sigma}_1) F_2(\bar{\sigma}_1, \sigma_2) P_{\mathrm{SW}}(\sigma_2, \tau_2) P_{\mathrm{HB}}^w(\bar{\tau}_2, \bar{\tau}_1) F_2(\bar{\tau}_1, \tau) \\
&= \sum_{\substack{\bar{\sigma}_1 \in \Lambda_w, \\ \tau_2 \in \Omega_{\mathrm{P}}}} P_{\mathrm{HB}}^w(\bar{\sigma}, \bar{\sigma}_1) P_{\mathrm{SW}}(\bar{\sigma}_1^1, \bar{\tau}_2^1) P_{\mathrm{HB}}^w(\bar{\tau}_2, \bar{\tau}) F_2(\bar{\tau}, \tau) \\
&\leq \widetilde{c} P_{\mathrm{SW}}(\bar{\sigma}^1, \bar{\tau}^1) \sum_{\bar{\sigma}_1 \sim \bar{\sigma}} P_{\mathrm{HB}}^w(\bar{\sigma}, \bar{\sigma}_1) \cdot \frac{1}{q} \sum_{\tau_2 : \bar{\tau}_2 \sim \bar{\tau}} P_{\mathrm{HB}}^w(\bar{\tau}_2, \bar{\tau}) \\
&= \widetilde{c} P_{\mathrm{SW}}(\sigma, \tau) \sum_{\substack{\bar{\tau}_3 \in \Lambda_w : \\ \bar{\tau}_3 \sim \bar{\tau}}} P_{\mathrm{HB}}^w(\bar{\tau}_3, \bar{\tau}) \leq q \widetilde{c} P_{\mathrm{SW}}(\sigma, \tau).
\end{aligned}$$

With a_3 from Lemma 3.8 we see for $\sigma_1, \sigma_2, \tau_1, \tau_2 \in \Omega_{\mathrm{P}}$ with $\bar{\sigma}_1^1 \sim \bar{\sigma}_2^1$ and $\bar{\tau}_1^1 \sim \bar{\tau}_2^1$ that

$$\frac{P_{\mathrm{SW}}(\sigma_1, \tau_1)}{P_{\mathrm{SW}}(\sigma_2, \tau_2)} \leq a_3^{\widetilde{\Delta}} \frac{P_{\mathrm{SW}}(\sigma_2, \tau_1)}{P_{\mathrm{SW}}(\sigma_2, \tau_2)} \leq a_3^{2\widetilde{\Delta}} \frac{P_{\mathrm{SW}}(\sigma_2, \tau_2)}{P_{\mathrm{SW}}(\sigma_2, \tau_2)} = a_3^{2\widetilde{\Delta}},$$

because $\deg_G(v) \leq \widetilde{\Delta}$ for all $v \in V \setminus w$. By the same ideas as in the proof of Theorem 3.1 (and Lemma 2.5) we conclude that

$$\lambda(Q) \leq q \widetilde{c} \lambda(P_{\mathrm{SW}}) \leq q a_3^{2\widetilde{\Delta}} \lambda(P_{\mathrm{SW}}) \leq q (q e^{2\beta})^{2\widetilde{\Delta}} \lambda(P_{\mathrm{SW}}). \quad \blacksquare$$

4. Comparison with single-bond dynamics

The goal of this chapter is to present a comparison of Swendsen–Wang and single-bond dynamics for the random-cluster model. In fact, we prove that the spectral gaps of both Markov chains are equal up to a small polynomial in the number of edges of the underlying graph $G = (V, E)$, i.e. up to $c|E| \log |E|$ for some $c < \infty$. In particular, this proves that rapid mixing of both Markov chains is equivalent.

In this chapter, we first give a representation of both Markov chains (or their transition matrices) on the joint (FKES) model (see (2.12)) using the same "building blocks". Then we provide some technical lemmas that yield estimates on the norm of products of operators between (not necessarily equal) Hilbert spaces.

Using these ingredients we obtain several new rapid mixing results for the single-bond dynamics, but we can also use Theorem 2.19 to get a result on Swendsen–Wang dynamics on graphs of bounded *linear width*. Additionally we adopt the result of Theorem 2.12 to prove matching upper and lower bounds for the spectral gap of the single-bond dynamics on the d-dimensional torus (see (2.32)) at the critical temperature that show slow mixing if the number of colors is large enough.

4.1. Common representation. We want to represent the Swendsen–Wang and the single-bond dynamics on the joint model, which consists of the product state space $\Omega_{\mathrm{J}} := \Omega_{\mathrm{P}} \times \Omega_{\mathrm{RC}}$ and the FKES measure. This was done first in [70] and we follow the steps from that article. First recall that, if $G = (V, E)$, $p = 1 - e^{-\beta}$ and $q \in \mathbb{N}$ are fixed, we denote by μ (resp. ν) the random-cluster (resp. FKES) measure on G with parameters p and q (see (2.11) and (2.12)) and that $L_2(\mu)$ (resp. $L_2(\nu)$) is the Hilbert space that consists of $\mathbb{R}^{\Omega_{\mathrm{RC}}}$ (resp. $\mathbb{R}^{\Omega_{\mathrm{J}}}$) with the inner product $\langle \cdot, \cdot \rangle_\mu$ (resp. $\langle \cdot, \cdot \rangle_\nu$).

We introduce the stochastic matrix M that defines the mapping (by matrix multiplication) from the RC to the joint model by

$$M(B, (\sigma, A)) := q^{-c(B)} \mathbb{1}(A = B) \mathbb{1}(B \subset E(\sigma)) \tag{4.1}$$

for $B \in \Omega_{\mathrm{RC}}$ and $(\sigma, A) \in \Omega_{\mathrm{J}}$. Obviously, we find by definition of μ and ν that M satisfies $\mu M(\sigma, A) = \sum_{B \in \Omega_{\mathrm{RC}}} \mu(B) M(B, (\sigma, A)) = \nu(\sigma, A)$. (See the discussion after (2.15) for a possible interpretation of that equality.) As in (2.4), M defines an operator that maps from $L_2(\nu)$ to $L_2(\mu)$ by

$$Mf(B) := \sum_{(\sigma, A) \in \Omega_{\mathrm{J}}} M(B, (\sigma, A)) f(\sigma, A), \qquad f \in L_2(\nu). \tag{4.2}$$

Its adjoint operator M^*, i.e. the operator that satisfies

$$\langle M^* g, f \rangle_\nu = \langle g, Mf \rangle_\mu \qquad \text{for all } f \in L_2(\nu), g \in L_2(\mu),$$

can be given by the (stochastic) matrix

$$M^*((\sigma, A), B) = \mathbb{1}(A = B).$$

One simple and useful property of the matrices M and M^* is that their product satisfies $MM^*(A, B) = \mathbb{1}(A = B)$ for all $A, B \in \Omega_{\mathrm{RC}}$, or equivalently, the corresponding operators satisfy $MM^* = I$ with $If := f$ for all $f \in L_2(\mu)$.

REMARK 4.1. The matrices above will be used with the following interpretation in mind. We want to construct a Markov chain for the RC model that is reversible with respect to μ. This could be done, starting in a RC configuration, by making a "step" with M first, which leads to a configuration in the joint model (with an unchanged RC coordinate). Then make updates of the RC coordinate in the joint configuration that are reversible with respect to ν. Finally, forget about the coloring to obtain a new RC configuration. At first sight this is no advantage, but we will see that it is possible in some cases to find a representation of Markov chains for the RC model by rather simple update rules on the joint model, while the original updates are difficult to handle.

The following matrices provide the updates of the "RC coordinate" in the joint model. For $(\sigma, A), (\tau, B) \in \Omega_{\mathrm{J}}$ and $e \in E$ with $e = \{e^{(1)}, e^{(2)}\}$ let

$$T_e((\sigma, A), (\tau, B)) := \mathbb{1}(\sigma = \tau) \begin{cases} p, & B = A \cup e \text{ and } \sigma(e^{(1)}) = \sigma(e^{(2)}), \\ 1-p, & B = A \setminus e \text{ and } \sigma(e^{(1)}) = \sigma(e^{(2)}), \\ 1, & B = A \setminus e \text{ and } \sigma(e^{(1)}) \neq \sigma(e^{(2)}). \end{cases} \tag{4.3}$$

The lemma below shows some interesting properties of the matrices from (4.1) and (4.3). For example $\{T_e\}_{e \in E}$ is a family of commuting projections in $L_2(\nu)$. This will be important in the proof of the main result.

LEMMA 4.2. *Let M, M^* and T_e be the matrices above. Then*

(i) *M^*M and T_e are self-adjoint in $L_2(\nu)$.*
(ii) *$T_eT_e = T_e$ and $T_eT_{e'} = T_{e'}T_e$ for all $e, e' \in E$.*
(iii) *$\|T_e\|_\nu = 1$ and $\|M^*M\|_\nu = 1$.*

Proof. The first part of (i) is obvious since $(M^*M)^* = M^*(M^*)^* = M^*M$ (see e.g. [36, Thm. 3.9-4]). The second part can be checked easily using (2.6). Part (ii) comes from the fact that the transition probabilities depend only on the "coordinate" that will not be changed. For (iii) note that $\|M^*M\|_\nu = \|MM^*\|_\mu$ (see [36, Thm. 3.9-4]), and $MM^* = I$ with $If := f$, $f \in L_2(\mu)$. Hence, $\|M^*M\|_\nu = \|I\|_\mu = 1$. It remains to prove $\|T_e\|_\nu = 1$. By $T_eT_e = T_e$ and the self-adjointness of T_e we obtain $\|T_e\|_\nu \overset{\text{(ii)}}{=} \|T_e^2\|_\nu \overset{\text{(i)}}{=} \|T_e\|_\nu^2$. This implies $\|T_e\|_\nu \in \{0, 1\}$, but since $T_e g = g$ if g is constant, $\|T_e\|_\nu$ cannot be zero. ∎

We finish this section with a lemma that demonstrates the relation of M and T_e to the Swendsen–Wang and single-bond dynamics. Recall that $\widetilde{P}_{\mathrm{SW}}$ and $\widetilde{P}_{\mathrm{SB}}$ are their transition matrices (see (2.22) and (2.27)).

LEMMA 4.3. *Let M, M^* and T_e be the matrices above. Then*

(i) $\widetilde{P}_{\mathrm{SW}} = M\Big(\prod_{e \in E} T_e\Big)M^*$.

(ii) $\widetilde{P}_{\mathrm{SB}} = \dfrac{1}{|E|}\sum_{e \in E} MT_eM^* = M\Big(\dfrac{1}{|E|}\sum_{e \in E} T_e\Big)M^*$.

From Lemma 4.2(ii) we know that the order of multiplication in (i) is not important.

Proof of Lemma 4.3. For (i) note that

$$\Big(\prod_{e\in E} T_e\Big)((\sigma,A),(\tau,B)) = \mathbb{1}(\sigma=\tau)\mathbb{1}(B\subset E(\sigma))p^{|B|}(1-p)^{|E(\sigma)|-|B|}.$$

Hence,

$$\begin{aligned} M\Big(\prod_{e\in E} T_e\Big)M^*(A,B) &= \sum_{\sigma\in\Omega_{\mathrm{P}}} M(A,(\sigma,A))\Big(\prod_{e\in E} T_e\Big)((\sigma,A),(\sigma,B)) \\ &= \sum_{\sigma\in\Omega_{\mathrm{P}}} q^{-c(A)} \underbrace{\mathbb{1}(A\subset E(\sigma))\mathbb{1}(B\subset E(\sigma))}_{=\mathbb{1}(A\cup B\subset E(\sigma))} p^{|B|}(1-p)^{|E(\sigma)|-|B|} \\ &= P_{\mathrm{SW}}(A,B). \end{aligned}$$

For part (ii) we define $\mathbb{1}_e(\sigma) := \mathbb{1}(\sigma(e^{(1)}) = \sigma(e^{(2)}))$ and $\mathbb{1}_e(A) := \mathbb{1}(e^{(1)} \overset{A}{\longleftrightarrow} e^{(2)})$ for $\sigma\in\Omega_{\mathrm{P}}$, $A\in\Omega_{\mathrm{RC}}$ and $e\in E$ with $e=\{e^{(1)},e^{(2)}\}$. Now write

$$T_e((\sigma,A),(\sigma,B)) = \mathbb{1}(B=A\setminus e) + p\mathbb{1}_e(\sigma)[\mathbb{1}(B=A\cup e) - \mathbb{1}(B=A\setminus e)]$$

and note that $|\{\sigma\in\Omega_{\mathrm{P}} : A\subset E(\sigma)\}| = q^{c(A)}$ (because the colorings have to be constant on each of the $c(A)$ components) and

$$q^{-c(A)} \sum_{\sigma:\, A\subset E(\sigma)} \mathbb{1}_e(\sigma) = \frac{1}{q} + \mathbb{1}_e(A)\Big(1-\frac{1}{q}\Big).$$

Hence,

$$\begin{aligned} MT_eM^*(A,B) &= \sum_{\sigma} q^{-c(A)}\mathbb{1}(A\subset E(\sigma))T_e((\sigma,A),(\sigma,B)) \\ &= \mathbb{1}(B=A\setminus e) \\ &\quad + p[\mathbb{1}(B=A\cup e) - \mathbb{1}(B=A\setminus e)]q^{-c(A)} \sum_{\sigma:\, A\subset E(\sigma)} \mathbb{1}_e(\sigma) \\ &= \begin{cases} p, & B=A\cup e \text{ and } e^{(1)} \overset{A}{\longleftrightarrow} e^{(2)}, \\ 1-p, & B=A\setminus e \text{ and } e^{(1)} \overset{A}{\longleftrightarrow} e^{(2)}, \\ p/q, & B=A\cup e \text{ and } e^{(1)} \overset{A}{\nleftrightarrow} e^{(2)}, \\ 1-p/q, & B=A\setminus e \text{ and } e^{(1)} \overset{A}{\nleftrightarrow} e^{(2)}. \end{cases} \end{aligned}$$

Summing over E and dividing by $|E|$ shows equality to (2.27). ∎

Remark 4.4. We know from Lemma 4.2 that T_e, $e\in E$, is self-adjoint and idempotent. Hence, T_e is positive, i.e. $\langle T_ef,f\rangle_\nu \geq 0$ for all $f\in L_2(\nu)$ (see [36, Thms. 9.5-1 & 9.5-2]). In fact, all eigenvalues of T_e are either 0 and 1. Additionally, $\prod_{e\in E} T_e$ is positive, because it is the product of commuting and self-adjoint linear operators (see [36, Thm. 9.3-1]). Consequently,

$$\langle \widetilde{P}_{\mathrm{SW}}g,g\rangle_\mu = \Big\langle M\Big(\prod_{e\in E} T_e\Big)M^*g,g\Big\rangle_\mu = \Big\langle \Big(\prod_{e\in E} T_e\Big)M^*g, M^*g\Big\rangle_\nu \geq 0,$$

$$\langle \widetilde{P}_{\mathrm{SB}}g,g\rangle_\mu = \frac{1}{|E|}\sum_{e\in E}\langle MT_eM^*g,g\rangle_\mu = \frac{1}{|E|}\sum_{e\in E}\langle T_eM^*g,M^*g\rangle_\nu \geq 0$$

for all $g \in L_2(\mu)$. This shows that the transition matrices of Swendsen–Wang and single-bond dynamics are positive, which implies that they have only non-negative eigenvalues (see [33, Obs. 7.1.4]).

4.2. Technical lemmas. In this section we provide some technical lemmas that will be necessary for the analysis. We state them in a general form, because we guess that they could be useful also in other settings. First let us introduce the notation. Consider two Hilbert spaces H_1 and H_2 with the corresponding inner products $\langle\cdot,\cdot\rangle_{H_1}$ and $\langle\cdot,\cdot\rangle_{H_2}$. The norms in H_1 and H_2 are defined as usual as the square root of the inner product of a function with itself. Throughout this section we consider two bounded, linear operators, $R : H_2 \to H_1$ and $T : H_2 \to H_2$, such that

- T is self-adjoint, i.e. $T = T^*$, and
- T is positive, i.e. $\langle Tg, g\rangle_{H_2} \geq 0$ for all $g \in H_2$.

We denote by $\|\cdot\|_{H_1}$ (resp. $\|\cdot\|_{H_2\to H_1}$) the operator norms of operators mapping from H_1 to H_1 (resp. H_2 to H_1), i.e.

$$\|R\|_{H_2\to H_1} := \max_{\|g\|_{H_2}=1} \|Rg\|_{H_1}$$

and $\|\cdot\|_{H_1}$ is as in (2.7). As is well-known, the adjoint operator of R, i.e. $R^* : H_1 \to H_2$, with $\langle R^* f, g\rangle_{H_2} = \langle f, Rg\rangle_{H_1}$ for all $f \in H_1$ and $g \in H_2$, satisfies $\|R^*\|_{H_1\to H_2} = \|R\|_{H_2\to H_1}$ (see e.g. Kreyszig [36, Thm. 3.9-2]). Additionally note that self-adjointness of T implies that RT^kR^*, $k \in \mathbb{N}$, are self-adjoint operators on H_1.

LEMMA 4.5. *Let T and R be as above. Then*

$$\|RT^{k+1}R^*\|_{H_1} \leq \|T\|_{H_2}\|RT^kR^*\|_{H_1} \quad \text{for all } k \in \mathbb{N}.$$

In particular, if $\|T\|_{H_2} \leq 1$ this proves monotonicity in k.

Proof. By the assumptions, T has a unique positive square root $\widetilde{T}$, i.e. $T = \widetilde{T}\widetilde{T}$, which is again self-adjoint (see e.g. [36, Th. 9.4-2]). Using the fact that $\|A\|^2_{H_2\to H_1} = \|AA^*\|_{H_1}$ for every bounded linear operator $A : H_2 \to H_1$, we obtain

$$\begin{aligned}\|RT^{k+1}R^*\|_{H_1} &= \|R\widetilde{T}^{2k+2}R^*\|_{H_1} = \|R\widetilde{T}^{k+1}\|^2_{H_2\to H_1}\\ &\leq \|R\widetilde{T}^k\|^2_{H_2\to H_1}\|\widetilde{T}\|^2_{H_2} = \|R\widetilde{T}^{2k}R^*\|_{H_1}\|T\|_{H_2} = \|T\|_{H_2}\|RT^kR^*\|_{H_1}. \blacksquare\end{aligned}$$

LEMMA 4.6. *In the above setting let additionally $\|R\|^2_{H_2\to H_1} = \|RR^*\|_{H_1} \leq 1$. Then*

$$\|RTR^*\|^{2^k}_{H_1} \leq \|RT^{2^k}R^*\|_{H_1} \quad \text{for all } k \in \mathbb{N}.$$

Proof. The case $k = 0$ is obvious. Now suppose the statement is correct for $k-1$; then

$$\begin{aligned}\|RTR^*\|^{2^k}_{H_1} &= \|RTR^*\|^{2^{k-1}2}_{H_1} \leq \|RT^{2^{k-1}}R^*\|^2_{H_1}\\ &\leq \|RT^{2^{k-1}}\|^2_{H_2\to H_1}\|R^*\|^2_{H_1\to H_2}\\ &= \|RT^{2^{k-1}}T^{2^{k-1}}R^*\|_{H_1}\|RR^*\|_{H_2} \leq \|RT^{2^k}R^*\|_{H_1},\end{aligned}$$

which proves the statement for k. ∎

The next corollary combines the statements of the last two lemmas to give a result similar to Lemma 4.6 for arbitrary exponents.

COROLLARY 4.7. *Additionally to the general assumptions of this section let* $\|T\|_{H_2} \leq 1$ *and* $\|RR^*\|_{H_1} \leq 1$. *Then*

$$\|RTR^*\|_{H_1}^{2k} \leq \|RT^kR^*\|_{H_1} \quad \text{for all } k \in \mathbb{N}.$$

Proof. Let $\ell = \lceil \log_2 k \rceil$, so that $k \leq 2^\ell \leq 2k$. Since $\|RTR^*\|_{H_1} \leq 1$ by assumption, we obtain

$$\|RTR^*\|_{H_1}^{2k} \leq \|RTR^*\|_{H_1}^{2^\ell} \overset{4.6}{\leq} \|RT^{2^\ell}R^*\|_{H_1} \overset{4.5}{\leq} \|RT^kR^*\|_{H_1}. \ \blacksquare$$

In the following section we will apply these bounds for a specific choice of R and T.

4.3. The result. We prove the following theorem (see [69, 70]).

THEOREM 4.8. *Let* $\widetilde{P}_{\mathrm{SW}}$ (*resp.* $\widetilde{P}_{\mathrm{SB}}$) *be the transition matrix of the Swendsen–Wang* (*resp. single-bond*) *dynamics for the random-cluster model on a graph with* $m \geq 3$ *edges. Then*

$$\lambda(\widetilde{P}_{\mathrm{SB}}) \leq \lambda(\widetilde{P}_{\mathrm{SW}}) \leq 8m \log m \cdot \lambda(\widetilde{P}_{\mathrm{SB}}).$$

This theorem shows that the Swendsen–Wang dynamics is rapidly mixing if and only if the single-bond dynamics is rapidly mixing, since the spectral gaps can differ only by a polynomial in the number of edges of the graph. In Section 4.4 we will state some of the results that follow from the (already discussed) results for the Swendsen–Wang dynamics.

The proof of Theorem 4.8 is based on the bounds of the last section and a suitable choice of the involved operators. Recall that we consider both dynamics on a graph $G = (V, E)$ with m edges, i.e. $m = |E|$. Fix an arbitrary ordering $e_1, \ldots, e_m$ of the edges $e \in E$. We set the Hilbert spaces from Section 4.2 to

$$H_1 := L_2(\mu) \quad \text{and} \quad H_2 := L_2(\nu)$$

and define the operators

$$T := \frac{1}{m} \sum_{i=1}^{m} T_{e_i} \quad \text{and} \quad \mathcal{T} := \prod_{i=1}^{m} T_{e_i},$$

where the T_e, $e \in E$, are from the common representation in (4.3). Recall that $\widetilde{P}_{\mathrm{SB}} = MTM^*$ and $\widetilde{P}_{\mathrm{SW}} = M\mathcal{T}M^*$ by Lemma 4.3 with M from (4.1). Additionally we define

$$\mathcal{T}_\alpha := \prod_{i=1}^{m} T_{e_i}^{\alpha_i}$$

for $\alpha \in \mathbb{N}^m$. By Lemma 4.2(ii) we deduce for $\alpha, \gamma \in \mathbb{N}^m$ that $\mathcal{T}_\alpha = \mathcal{T}_\gamma$ if and only if $\{i : \alpha_i = 0\} = \{i : \gamma_i = 0\}$. Furthermore, $\mathcal{T}_\alpha = \mathcal{T}$ iff $\alpha_i > 0$ for all $i = 1, \ldots, m$.

To conclude the proof of Theorem 4.8 we need two lemmas that are stated in the sense of Section 4.2. Afterwards we will see how this implies the result. The first lemma follows from Lemma 4.5.

LEMMA 4.9. *Let* $R : H_2 \to H_1$ *be a bounded linear operator. Then*

$$\|R\mathcal{T}R^*\|_{H_1} \leq \|RT^kR^*\|_{H_1} \quad \text{for all } k \in \mathbb{N}.$$

Proof. See Remark 4.4 for arguments that show positivity of T. Additionally, Lemma 4.2 leads to $\|T\|_{H_2} \leq \frac{1}{m} \sum_{i=1}^{m} \|T_{e_i}\|_{H_2} = 1$. Hence T satisfies the assumptions of Lemma 4.5.

We obtain

$$\|RT^{\ell}R^*\|_{H_1} \le \|RT^kR^*\|_{H_1} \quad \text{for every } k \le \ell.$$

Note that $\lim_{\ell\to\infty} T^{\ell} = \mathcal{T}$ (in norm topology) implies $\lim_{\ell\to\infty} \|RT^{\ell}R^*\|_{H_1} = \|R\mathcal{T}R^*\|_{H_1}$, which yields the result. ∎

Lemma 4.9 (or 4.5) shows that, while k approaches infinity, $\|RT^kR^*\|_{H_1}$ monotonically approaches $\|R\mathcal{T}R^*\|_{H_1}$. This suggests that, for k large enough, these norms are close to each other. The next lemma yields such a reverse inequality.

Lemma 4.10. *Let $k = \lceil m\log(m/\varepsilon)\rceil$ and $R : H_2 \to H_1$ be a bounded linear operator with $\|RR^*\|_{H_1} \le 1$. Then*

$$\|RT^kR^*\|_{H_1} \le (1-\varepsilon)\big\|R\mathcal{T}R^*\big\|_{H_1} + \varepsilon.$$

Proof. Define the index sets $I_{m,k} := \{\alpha \in \mathbb{N}^m : \sum_{i=1}^m \alpha_i = k\}$ and $I^1_{m,k} := \{\alpha \in I_{m,k} : \alpha_i > 0, \forall i = 1,\dots,m\}$. Let $I^0_{m,k} := I_{m,k} \setminus I^1_{m,k}$ and denote by $\binom{k}{\alpha}$, for $\alpha \in I_{m,k}$, the multinomial coefficient, i.e.

$$\binom{k}{\alpha} = \frac{k!}{\alpha_1!\cdots\alpha_m!}.$$

Obviously (by the multinomial theorem [30, eq. (2.21)]),

$$\sum_{\alpha\in I_{m,k}} \binom{k}{\alpha} = m^k$$

and

$$Z_{m,k} := \sum_{\alpha\in I^0_{m,k}} \binom{k}{\alpha} \le \sum_{i=1}^m \sum_{\alpha\in I^0_{m,k}:\alpha_i=0} \binom{k}{\alpha} = m \sum_{\gamma\in I_{m-1,k}} \binom{k}{\gamma} = m(m-1)^k.$$

We write

$$T^k = \left(\frac{1}{m}\sum_{i=1}^m T_{e_i}\right)^k = \frac{1}{m^k}\sum_{\alpha\in I_{m,k}} \binom{k}{\alpha}\mathcal{T}_\alpha = \frac{1}{m^k}\sum_{\alpha\in I^1_{m,k}} \binom{k}{\alpha}\mathcal{T}_\alpha + \frac{1}{m^k}\sum_{\alpha\in I^0_{m,k}} \binom{k}{\alpha}\mathcal{T}_\alpha.$$

Note that we use for the second equality the fact that the T_e's are commuting by Lemma 4.2(ii). Since we know that $\mathcal{T}_\alpha = \mathcal{T}$ for every $\alpha \in I^1_{m,k}$ (note that $I^1_{m,k} = \emptyset$ for $k < m$) and $\|R\mathcal{T}_\alpha R^*\|_{H_1} \le 1$ for every $\alpha \in I_{m,k}$, we obtain

$$\begin{aligned}\|RT^kR^*\|_{H_1} &\le \frac{1}{m^k}\sum_{\alpha\in I^1_{m,k}} \binom{k}{\alpha}\big\|R\mathcal{T}_\alpha R^*\big\|_{H_1} + \frac{1}{m^k}\sum_{\alpha\in I^0_{m,k}} \binom{k}{\alpha}\|R\mathcal{T}_\alpha R^*\|_{H_1}\\ &\le \left(1 - \frac{Z_{m,k}}{m^k}\right)\|R\mathcal{T}R^*\|_{H_1} + \frac{Z_{m,k}}{m^k}.\end{aligned}$$

Using $(1-a)c + a \le (1-b)c + b$ for $c \le 1$ and $a \le b$, we find that

$$\|RT^kR^*\|_{H_1} \le \left(1 - m\left(1-\frac{1}{m}\right)^k\right)\big\|R\mathcal{T}R^*\big\|_{H_1} + m\left(1-\frac{1}{m}\right)^k.$$

Setting $k = \lceil m\log(m/\varepsilon)\rceil$ yields the result. ∎

Now we are able to prove the comparison result for SW and SB dynamics. For this let $S_1(B,(\sigma,A)) := \nu(\sigma,A)$ for all $B \in \Omega_{\mathrm{RC}}$ and $(\sigma,A) \in \Omega_{\mathrm{J}}$, which defines an operator (by (4.2)) that maps from H_2 to H_1. The adjoint operator S_1^* is then given by $S_1^*((\sigma,A),B) := \mu(B)$, and thus $S_1 S_1^*(A,B) = S_\mu(A,B) = \mu(B)$ for all $A,B \in \Omega_{\mathrm{RC}}$. For the proof we choose the operator

$$R := M - S_1$$

with M from (4.1). This operator has some useful properties which can be deduced directly from the properties of M and S_1. First of all, note that the matrix corresponding to S_1^* is constant in the first (FKES) coordinate. This readily implies $MS_1^* = S_\mu$. But S_μ defines a self-adjoint operator on $H_1 := L_2(\mu)$, and hence $S_\mu = S_\mu^* = (MS_1^*)^* = S_1 M^*$. It follows that

$$RR^* = (M - S_1)(M^* - S_1^*) = MM^* - S_\mu = I - S_\mu,$$

with $If := f$ for all $f \in L_2(\mu)$, because $MM^*(A,B) = \mathbb{1}(A = B)$. Since $(RR^*)^2 = (I - S_\mu)^2 = I - S_\mu = RR^*$, we deduce that R satisfies $\|RR^*\|_{H_1} \in \{0,1\}$ and thus the assumptions of Lemmas 4.9 and 4.10. Note that $\|RR^*\|_{H_1} = 1$ whenever $m > 1$, since then there always exists a non-constant function $f \in H_1$ with $\langle f, 1\rangle_\mu = 0$, i.e. $(I - S_\mu)f = f$.

Additionally, we find from Lemma 4.3 that

$$\widetilde{P}_{\mathrm{SB}} - S_\mu = RTR^* \quad \text{and} \quad \widetilde{P}_{\mathrm{SW}} - S_\mu = R\mathcal{T}R^*.$$

Proof of Theorem 4.8. Recall that by definition

$$\lambda(\widetilde{P}_{\mathrm{SB}}) = 1 - \|\widetilde{P}_{\mathrm{SB}} - S_\mu\|_{H_1} = 1 - \|RTR^*\|_{H_1},$$
$$\lambda(\widetilde{P}_{\mathrm{SW}}) = 1 - \|\widetilde{P}_{\mathrm{SW}} - S_\mu\|_{H_1} = 1 - \|R\mathcal{T}R^*\|_{H_1}.$$

Obviously, Lemma 4.9 (with $k = 1$) implies the first inequality of Theorem 4.8.

Now let $k = \lceil m \log(m/\varepsilon) \rceil$. Since we know from Lemma 4.10 that

$$\|R\mathcal{T}R^*\|_{H_1} \geq \frac{1}{1-\varepsilon}(\|RT^kR^*\|_{H_1} - \varepsilon)$$

we obtain

$$\begin{aligned}\lambda(\widetilde{P}_{\mathrm{SW}}) = 1 - \|R\mathcal{T}R^*\|_{H_1} &\overset{4.10}{\leq} 1 - \frac{1}{1-\varepsilon}(\|RT^kR^*\|_{H_1} - \varepsilon) \\ &= \frac{1}{1-\varepsilon}(1 - \|RT^kR^*\|_{H_1}) \overset{4.7}{\leq} \frac{1}{1-\varepsilon}(1 - \|RTR^*\|_{H_1}^{2k}) \\ &\leq \frac{2k}{1-\varepsilon}(1 - \|RTR^*\|_{H_1}) = \frac{2k}{1-\varepsilon}\lambda(\widetilde{P}_{\mathrm{SB}}),\end{aligned}$$

where the last inequality comes from $1 - x^k \leq k(1-x)$ for $x \in [0,1]$. Setting $\varepsilon = \frac{1}{2}$, we obtain $\frac{2k}{1-\varepsilon} = 4k \leq 8m \log m$. ∎

4.4. Applications. We present three applications of Theorem 4.8. The first one concerns the Swendsen–Wang dynamics on graphs with bounded linear width and is based on a result of Ge and Štefankovič [25] (see Theorem 2.19 above). For this recall the definition of the *linear width* of a graph $G = (V,E)$ as the smallest number ℓ such that there exists an ordering $e_1, \ldots, e_{|E|}$ of the edges with the property that for every $i \in [|E|]$ there are at most ℓ vertices that have an adjacent edge in $\{e_1, \ldots, e_i\}$ and in $\{e_{i+1}, \ldots, e_{|E|}\}$.

COROLLARY 4.11. *Let $\widetilde{P}_{\mathrm{SW}}$ be the transition matrix of the Swendsen–Wang dynamics for the random-cluster model with parameters p and $q \in \mathbb{N}$ on a graph $G = (V, E)$ with linear width bounded by ℓ. Let $m := |E|$. Then*

$$\lambda(\widetilde{P}_{\mathrm{SW}}) \geq \frac{1}{2q^{\ell+1}} \frac{1}{m^2}. \tag{4.4}$$

Proof. The result follows from Theorem 4.8 together with Theorem 2.19. ∎

Note that Lemma 2.6 shows that the result holds true if we replace the random-cluster model by the corresponding Potts model at inverse temperature $\beta = -\log(1-p)$ and $\widetilde{P}_{\mathrm{SW}}$ by P_{SW}.

Now we turn to the single-bond dynamics. The second result of this section, analogously to Theorems 2.10 and 3.4, describes rapid mixing on $\mathbb{Z}_L^2 = (V_{L,2}, E_{L,2})$, i.e. on the two-dimensional square lattice of side length L, at high temperatures.

COROLLARY 4.12. *Let $\widetilde{P}_{\mathrm{SB}}$ be the transition matrix of the single-bond dynamics for the RC model on $\mathbb{Z}_L^2$ with parameters p and $q \geq 1$. Let $m = 2L(L-1) = |E_{L,2}|$. Then there exist constants $c_p = c_p(q), c' > 0$ and $C < \infty$ such that*

$$\lambda(\widetilde{P}_{\mathrm{SB}}) \geq \frac{c_p}{m^2 \log m} \qquad \text{for } p < p_c(q),$$

and

$$\lambda(\widetilde{P}_{\mathrm{SB}}) \geq c' m^{-C} \qquad \text{for } q = 2 \text{ and } p = p_c(2),$$

where $p_c(q) = \frac{\sqrt{q}}{1+\sqrt{q}}$.

Proof. The bounds for the Swendsen–Wang (SW) dynamics for the Potts model from Theorem 3.4 and Lemma 2.6 show that SW is rapidly mixing in the desired range of p. Theorem 4.8 leads to the specific bounds (since $|V_{L,2}| \leq |E_{L,2}| = 2L(L-1) \leq 2|V_{L,2}|$). ∎

In the next chapter we present techniques that allow us to relate the spectral gap of local Markov chains for the random-cluster model at high and low temperatures. In particular, we will extend the first bound of Corollary 4.12 to all non-critical temperatures (see Theorem 5.10).

The third and last result that we want to present here yields tight bounds on the spectral gap of the single-bond dynamics on the d-dimensional torus $\widetilde{\mathbb{Z}}_L^d$ of side length L (see (2.32)) at the critical temperature. These bounds show slow mixing if q is large enough. This complements (and also uses) a result of Borgs, Chayes and Tetali [5] (see Theorem 2.12).

THEOREM 4.13. *Let $\widetilde{P}_{\mathrm{SB}}$ (resp. $\widetilde{P}_{\mathrm{SW}}$) be the transition matrix of the single-bond (resp. Swendsen–Wang) dynamics for the random-cluster model on $\widetilde{\mathbb{Z}}_L^d$, $d \geq 2$, with parameters p and q. Then there exist constants $k_1, k_2 < \infty$ and a constant $k_3 > 0$ (depending all on d, β and q) such that, for q and L large enough,*

$$e^{-(k_1+k_2\beta)L^{d-1}} \leq \lambda(\widetilde{P}_{\mathrm{SB}}) \leq e^{-k_3\beta L^{d-1}}$$

and

$$e^{-(k_1+k_2\beta)L^{d-1}} \le \lambda(\widetilde{P}_{\mathrm{SW}}) \le e^{-k_3\beta L^{d-1}}$$

if $p = 1 - e^{-\beta_c(d,q)}$ *with* $\beta_c(d,q)$ *from* (2.30). *The lower bounds hold for all* p, q *and* L.

Proof. We obtain the second inequality immediately from Theorem 2.12 and Lemma 2.6. The result on the single-bond dynamics thus follows from Theorem 4.8. ∎

The original proof of the lower bound on $\lambda(\widetilde{P}_{\mathrm{SB}})$, as given in [70], uses the bound of Theorem 2.19 together with a bound on the linear width of $\widetilde{\mathbb{Z}}_L^d$.

REMARK 4.14. There are some more results on rapid mixing of the single-bond dynamics that follow directly from Theorem 4.8. These include rapid mixing on trees and cycles (cf. Theorem 2.17), on the complete graph (cf. Theorem 2.16) and on graphs with bounded degree if p is small enough (cf. Theorem 3.3). We do not state them here explicitly, because our main purpose is to prove rapid mixing of the Swendsen–Wang dynamics.

5. Rapid mixing in two dimensions

This chapter is devoted to the study of the spectral gaps of Swendsen–Wang and single-bond dynamics if the underlying graph for the random-cluster model has a special structure. Namely, we consider *planar* graphs, i.e. graphs that can be drawn in the plane without intersecting edges. For each such graph it is possible to define a corresponding *dual* graph and we will see that it is possible to prove bounds on the spectral gap (of both dynamics) on the original graph in terms of the spectral gap on the dual one, if the temperature is suitably changed. In particular, we prove results on rapid mixing of the dynamics at low temperatures, where the previous techniques, which rely on rapid mixing of single-spin dynamics for the Potts model at the same temperature, do not apply.

The plan of this chapter is as follows. In Section 5.1 we introduce the notion of representations of graphs in the plane and, consequently, the notion of planar graphs. This leads to the construction of dual graphs and we will see that there is a tight connection between the random-cluster model on a graph and its dual. In Section 5.2 we use this connection to prove the main results of this chapter, i.e. the comparison of the spectral gaps on a graph and on its dual graph. This ends in a proof of rapid mixing of the considered dynamics on the two-dimensional square lattice at all non-critical temperatures, see Section 5.3. Finally, in Section 5.4, we present a slight generalization of the given results to graphs that can be drawn on a surface with bounded genus.

5.1. Planar and dual graphs. The following introduction to embeddings of graphs in the plane and to the construction of dual graphs is presented according to Mohar and Thomassen [53] and we refer to this monograph for a more comprehensive study of the topic.

Throughout this chapter we consider only connected graphs. Let a connected graph $G = (V, E, \varphi)$ be given and recall that φ is the function that assigns to each edge $e \in E$ the set of its endvertices. For convenience, we will mostly omit φ from the notation. We

say that G is a *planar graph* if there is a *representation of G in the plane* $\mathbb{R}^2$. That is, there exists a finite set $\widetilde{V} \subset \mathbb{R}^2$, a bijection κ from V to $\widetilde{V}$ and a set of *simple curves* $\widetilde{E} = \{\gamma_e : [0,1] \to \mathbb{R}^2 : e \in E\}$ in $\mathbb{R}^2$ such that

- $\gamma_e(0) \in \widetilde{V}$ and $\gamma_e(1) \in \widetilde{V}$ for all $e \in E$,
- for all $u, v \in V$ and $e \in E$ we have
$$\{\gamma_e(0), \gamma_e(1)\} = \{\kappa(u), \kappa(v)\} \;\Leftrightarrow\; \varphi(e) = \{u, v\},$$
- $\gamma_e((0,1)) \cap \gamma_f([0,1]) = \emptyset$ for all $e \neq f \in E$.

Note that a curve $\gamma : [0,1] \to \mathbb{R}^2$ is always assumed to be continuous, and is called *simple* if it does not cross itself except for its ends, i.e. $\gamma(x) = \gamma(y)$ for $x \neq y$ implies $x, y \in \{0,1\}$. The above assumptions mean respectively that the ends of all curves are elements of $\widetilde{V}$, that γ_e connects the elements of $\widetilde{V}$ that correspond (by κ) to the endvertices of $e \in E$, and that the interior of each curve is disjoint from all other curves. The tuple $\widetilde{G} = (\widetilde{V}, \widetilde{E})$ is called a *representation of G in the plane* and we assume henceforth that we fix for every planar graph G a representation $\widetilde{G}$ in the plane. In the literature $\widetilde{G}$ is also called a *plane graph*, *planar embedding* or *drawing* of G. See e.g. Figure 3 for two examples of planar graphs (or, more precisely, their representations in the plane).

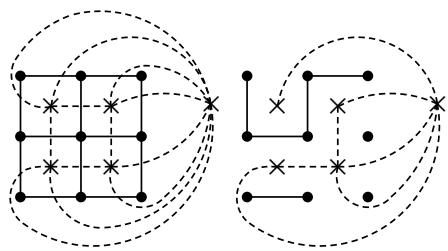

Fig. 3. Left: The graph $\mathbb{Z}_3^2$ (dots and solid lines) and its dual graph $\mathbb{Z}_3^{2\dagger}$ (crosses and dashed lines). Right: An RC configuration on $\mathbb{Z}_3^2$ and its dual configuration.

Clearly, being connected by a simple curve in a set $C \subset \mathbb{R}^2$ defines an equivalence relation on C, whose equivalence classes are called *regions of C*. A *face of C* is a region of $\mathbb{R}^2 \setminus C$ and we call the faces of $\bigcup_{e \in E} \gamma_e([0,1])$, i.e. the faces of the union of the images of all curves of $\widetilde{G}$, the *faces of $\widetilde{G}$*. Note that since we consider finite graphs, there is exactly one unbounded face. We call it the *outer face of $\widetilde{G}$*. In the following we need the number of faces of a representation $\widetilde{G}$ of a graph G, which we denote by $F(\widetilde{G})$. One may think that this number depends on the choice of the representation, but the following lemma, which is known as *Euler's* (*polyhedral*) *formula*, shows that this is not the case (see e.g. [53, Prop. 2.2.3]).

Lemma 5.1 (Euler's formula). *Let $G = (V, E)$ be a connected planar graph and $\widetilde{G}$ be a representation of G in the plane. Then*
$$2 = |V| - |E| + F(\widetilde{G}).$$

This proves that the number of faces does not depend on the particular representation and thus we write $F(G)$ for the *number of faces* of any representation of G.

In addition to the statement of Lemma 5.1 for the whole graph G, we need a result on its spanning subgraphs. For this recall that we denote by $G_A = (V, A)$, $A \subset E$, the graph G without the edges of $E \setminus A$ and that $c(G_A)$ (or simply $c(A)$) denotes the number of connected components of G_A. The representation of G_A in the plane is fixed to be the representation of G without the curves corresponding to the edges of $E \setminus A$ and we write $F(A)$ for the number of faces of the (fixed) representation of G_A. The next equation is a simple corollary of Lemma 5.1.

COROLLARY 5.2. *Let G be a planar graph and G_A, $A \subset E$, be a spanning subgraph of G. Then*

$$c(A) + 1 = |V| - |A| + F(A).$$

Proof. We know from Lemma 5.1 that for each connected component $G_i = (V_i, E_i)$ of G_A, $i = 1, \dots, c(A)$, we have $2 = |V_i| - |E_i| + F(G_i)$. Summing over i leads to $2c(A) = |V| - |A| + \sum_{i=1}^{c(A)} F(G_i)$, but in the last sum we count the outer face $c(A)$ times. Subtracting $c(A) - 1$ from both sides yields the result. ∎

Now we turn to the definition of the *dual graph* of a graph G. For this consider the (fixed) representation in the plane $\widetilde{G} = (\widetilde{V}, \widetilde{E})$ of the connected graph $G = (V, E)$. We define the dual graph as the (unique) graph $G^\dagger = (V^\dagger, E^\dagger)$ such that its representation in the plane $\widetilde{G}^\dagger = (\widetilde{V}^\dagger, \widetilde{E}^\dagger)$ satisfies:

- there is exactly one $\widetilde{v}^\dagger \in \widetilde{V}^\dagger$ in every face of $\widetilde{G}$, and
- for every $e \in E$ there is exactly one simple curve $\gamma_e^\dagger \in E^\dagger$ with endpoints in $\widetilde{V}^\dagger$ that intersects $\gamma_e \in \widetilde{E}$, but not γ_f, $f \neq e \in E$.

Note that, for every $\gamma_e \in \widetilde{E}$, the endpoints of the corresponding "dual" curve $\gamma_e^\dagger$ are unique, since there is only one vertex in every face of $\widetilde{G}$ and a curve can separate at most two faces from each other (by Jordan's Curve Theorem [53]). We call the edge $e_\dagger \in E^\dagger$ in the dual graph that belongs to the dual curve $\gamma_e^\dagger$ the *dual edge of* $e \in E$.

In general, the definition of the dual graph of G depends on the representation of G in the plane. Therefore recall that we fix such a representation whenever we fix the graph. This implies that the dual graph is well-defined, unique and that the dual of the dual graph is the original graph. (Here we presume that the representation in the plane of the dual graph is fixed to be the one constructed above.)

REMARK 5.3. Note that the dual graph of any planar graph is connected. Therefore the connectivity of G is necessary to ensure that the dual of the dual graph of G is G. Additionally, it is known (see [53, Thm. 2.6.7]) that if we restrict to 3-connected planar graphs, i.e. planar graphs with at least four vertices such that every subgraph obtained by deleting two vertices is still connected, then the dual graph is unique and thus independent of the representation.

In view of the application to Markov chains for the random-cluster model we also define a *dual configuration* $A^\dagger \subset E^\dagger$ on the dual graph $G^\dagger$ for every random-cluster configuration $A \subset E$ on G by the property that for all $e \in E$ (and corresponding $e_\dagger \in E^\dagger$)

$$e_\dagger \in A^\dagger \;\Leftrightarrow\; e \notin A. \tag{5.1}$$

Thus, a dual edge is present in $A^\dagger$ whenever the original one is absent from A. In particular, $|E| = |E^\dagger| = |A| + |A^\dagger|$ for every $A \subset E$. See Figure 3 for an example of a graph and its dual together with a proper pair of RC configurations.

Recall that $\widetilde{G}_A$, $A \subset E$, is the representation of the spanning subgraph $G_A = (V, A)$ in the plane that is adopted from $\widetilde{G}$ by forgetting the curves corresponding to $E \setminus A$, and $c(A)$ is the number of connected components of G_A. With a slight abuse of notation we write $c(A^\dagger)$, $A^\dagger \subset E^\dagger$, for the number of connected components of $G^\dagger_{A^\dagger}$.

It is easy to verify (cf. Figure 3) that every face of $\widetilde{G}_A$ contains a unique connected component of $G^\dagger_{A^\dagger}$ and vice versa. Thus, we have $c(A^\dagger) = F(A)$. Using Euler's formula (see Corollary 5.2) we obtain

$$c(A) + 1 = c(A^\dagger) + |V| - |A| = c(A^\dagger) + |V| - |E| + |A^\dagger|. \tag{5.2}$$

This equality leads immediately to the following lemma that shows the tight connection between the random-cluster measure on G and $G^\dagger$ (see e.g. Grimmett [29, eq. (6.4)]).

LEMMA 5.4. *Let $\mu^G_{p,q}$ be the random-cluster measure on a planar graph $G = (V, E)$ with parameters p and q. Furthermore, let $G^\dagger$ be the dual graph of G. Then*

$$\mu^G_{p,q}(A) = \mu^{G^\dagger}_{p^*,q}(A^\dagger) \quad \text{for all } A \subset E,$$

where p^* satisfies

$$\frac{p^*}{1-p^*} = \frac{q(1-p)}{p}. \tag{5.3}$$

Proof. Recall from (2.11) that

$$\mu^G_{p,q}(A) = \frac{1}{\widetilde{Z}(G,p,q)} \left(\frac{p}{1-p}\right)^{|A|} q^{c(A)},$$

where $\widetilde{Z}(G,p,q) = \sum_{A\subset E} \left(\frac{p}{1-p}\right)^{|A|} q^{c(A)}$ is the normalization constant. With (5.2) and (5.3) we obtain

$$\begin{aligned}
\left(\frac{p}{1-p}\right)^{|A|} q^{c(A)} &\overset{(5.2)}{=} \left(\frac{p}{1-p}\right)^{|A|} q^{c(A^\dagger)+|V|-|E|+|A^\dagger|-1} \\
&\overset{(5.3)}{=} \left(\frac{q(1-p^*)}{p^*}\right)^{|E|-|A^\dagger|} q^{c(A^\dagger)+|V|-|E|+|A^\dagger|-1} \\
&= q^{|V|-1} \left(\frac{p^*}{1-p^*}\right)^{|A^\dagger|-|E|} q^{c(A^\dagger)}.
\end{aligned}$$

Thus, with $\widetilde{Z}(G^\dagger, p^*, q) = q^{1-|V|}(p^*/(1-p^*))^{|E|}\widetilde{Z}(G,p,q)$ we get $\mu^G_{p,q}(A) = \mu^{G^\dagger}_{p^*,q}(A^\dagger)$ as desired. ∎

The relation from (5.3) has a unique *self-dual point*, i.e. a value of p such that $p = p^*$. This value is given by

$$p_{\mathrm{sd}}(q) := \frac{\sqrt{q}}{1+\sqrt{q}}. \tag{5.4}$$

Note that $p_{sd}(q)$ equals $1 - e^{\beta_c(q)}$, where $\beta_c(q)$ is the critical inverse temperature for the q-state Potts model on the two-dimensional square lattice (see Remark 2.9).

5.2. Dynamics on planar graphs. We prove that the spectral gap of the Swendsen–Wang and the single-bond dynamics for the random-cluster model on a planar graph G is bounded from above and from below by the spectral gap of the corresponding dynamics on the dual graph $G^\dagger$ if we change the temperature parameter p to p^* from the last section. Furthermore, we state a second corollary of Theorem 2.13 at the end of this section, that shows rapid mixing of the Swendsen–Wang dynamics for the Potts model on planar graphs if the inverse temperature is small or large enough (depending on the maximum degree). This is a modification of Corollary 3.3 to planar graphs (cf. Theorem 2.13).

For this fix a connected planar graph $G = (V, E)$ (together with its representation in the plane), some $p \in (0, 1)$ and a natural number q. Let $G^\dagger$ be the dual graph of G, and p^* be the unique value satisfying (5.3), i.e. $p^* = \frac{q(1-p)}{p+q(1-p)}$. We call the random-cluster model on $G^\dagger$ with parameters p^* and q the *dual model* and abbreviate $\mu_{p^*,q}^{G^\dagger}$ to $\mu^\dagger$. Additionally, given a Markov chain for the random-cluster model with transition matrix $\widetilde{P}$, we write $\widetilde{P}^\dagger$ for the transition matrix of the Markov chain for the corresponding dual model.

The first result demonstrates the usefulness of the heat-bath dynamics for the random-cluster model, which is one of the two previously defined local Markov chains for this model (see (2.24) for the definition of its transition matrix $\widetilde{P}_{\mathrm{HB}}$). It turns out that the spectral gap of $\widetilde{P}_{\mathrm{HB}}$ equals the spectral gap of $\widetilde{P}_{\mathrm{HB}}^\dagger$. This is a simple consequence of Lemma 5.4 and is probably known, but we could not find a reference for it.

LEMMA 5.5. *Let $\widetilde{P}_{\mathrm{HB}}$ (resp. $\widetilde{P}_{\mathrm{HB}}^\dagger$) be the transition matrix of the heat-bath dynamics for the random-cluster model on a planar graph G (resp. for the dual model). Then*

$$\lambda(\widetilde{P}_{\mathrm{HB}}) = \lambda(\widetilde{P}_{\mathrm{HB}}^\dagger).$$

Proof. By definition we have, for all $A, B \subset E$,

$$\widetilde{P}_{\mathrm{HB}}(A, B) = \frac{1}{|E|} \sum_{e \in E} \frac{\mu(B)}{\mu(A \cup e) + \mu(A \setminus e)} \mathbb{1}(B \setminus e = A \setminus e)$$

(see (2.24)). Using Lemma 5.4 and since $(A \cup e)^\dagger = A^\dagger \setminus e_\dagger$ (resp. $(A \setminus e)^\dagger = A^\dagger \cup e_\dagger$), by (5.1), we obtain

$$\begin{aligned}
\widetilde{P}_{\mathrm{HB}}(A, B) &= \frac{1}{|E|} \sum_{e \in E} \frac{\mu(B)}{\mu(A \cup e) + \mu(A \setminus e)} \mathbb{1}(B \setminus e = A \setminus e) \\
&= \frac{1}{|E|} \sum_{e \in E} \frac{\mu^\dagger(B^\dagger)}{\mu^\dagger((A \cup e)^\dagger) + \mu^\dagger((A \setminus e)^\dagger)} \mathbb{1}((B \setminus e)^\dagger = (A \setminus e)^\dagger) \\
&= \frac{1}{|E^\dagger|} \sum_{e_\dagger \in E^\dagger} \frac{\mu^\dagger(B^\dagger)}{\mu^\dagger(A^\dagger \setminus e_\dagger) + \mu^\dagger(A^\dagger \cup e_\dagger)} \mathbb{1}(B^\dagger \cup e_\dagger = A^\dagger \cup e_\dagger) = \widetilde{P}_{\mathrm{HB}}^\dagger(A^\dagger, B^\dagger).
\end{aligned}$$

Note for the last equality that $B^\dagger \cup e_\dagger = A^\dagger \cup e_\dagger$ if and only if $B^\dagger \setminus e_\dagger = A^\dagger \setminus e_\dagger$. This obviously implies that the matrices $\widetilde{P}_{\mathrm{HB}}$ and $\widetilde{P}_{\mathrm{HB}}^\dagger$ have the same eigenvalues and thus the same spectral gap. ∎

We immediately obtain the last ingredient for the proof of rapid mixing of Swendsen–Wang and single-bond dynamics on the two-dimensional square lattice at all non-critical temperatures (see Section 5.3).

THEOREM 5.6. *Let $\widetilde{P}_{\mathrm{SW}}$ (resp. $\widetilde{P}_{\mathrm{SB}}$) be the transition matrix of the Swendsen–Wang (resp. single-bond) dynamics for the random-cluster model on a planar graph G with m edges, and let $\widetilde{P}^{\dagger}_{\mathrm{SW}}$ (resp. $\widetilde{P}^{\dagger}_{\mathrm{SB}}$) be the SW (resp. SB) dynamics for the dual model. Then*

$$\lambda(\widetilde{P}_{\mathrm{SB}}) \le q\lambda(\widetilde{P}^{\dagger}_{\mathrm{SB}}) \quad \textit{and} \quad \lambda(\widetilde{P}_{\mathrm{SW}}) \le 8qm\log m \cdot \lambda(\widetilde{P}^{\dagger}_{\mathrm{SW}}).$$

Proof. With Lemmas 2.7 and 5.5 we get

$$\lambda(\widetilde{P}_{\mathrm{SB}}) \le \lambda(\widetilde{P}_{\mathrm{HB}}) = \lambda(\widetilde{P}^{\dagger}_{\mathrm{HB}}) \le \left(1 - p^*\left(1 - \frac{1}{q}\right)\right)^{-1} \lambda(\widetilde{P}^{\dagger}_{\mathrm{SB}}).$$

The constant on the right hand side is maximal for $p^* = 1$. This proves the first statement of the theorem. For the second we additionally use Theorem 4.8. Thus,

$$\lambda(\widetilde{P}_{\mathrm{SW}}) \le 8m\log m \cdot \lambda(\widetilde{P}_{\mathrm{SB}}) \le 8qm\log m \cdot \lambda(\widetilde{P}^{\dagger}_{\mathrm{SB}}) \le 8qm\log m \cdot \lambda(\widetilde{P}^{\dagger}_{\mathrm{SW}}). \ \blacksquare$$

Prior to the application of this result to the two-dimensional square lattice in the next section, we present the above-mentioned result on rapid mixing of the Swendsen–Wang dynamics at sufficiently high and low temperatures if the underlying graph is planar and has bounded maximum degree. This corollary relies on a result of Hayes [31].

COROLLARY 5.7. *The Swendsen–Wang dynamics for the random-cluster model with parameters p and q on a planar, simple and connected graph G with m edges and maximum degree $\Delta \ge 6$ satisfies*

$$\lambda(\widetilde{P}_{\mathrm{SW}}) \ge \begin{cases} \dfrac{c(1-\varepsilon)}{m} & \textit{if } p \le \dfrac{\varepsilon}{3\sqrt{\Delta - 3}}, \\ \dfrac{c(1-\varepsilon)}{m^2 \log m} & \textit{if } p \ge 1 - \dfrac{\varepsilon}{q\Delta^{\dagger}}, \end{cases}$$

for some $c = c(\Delta, p, q) > 0$ and $\varepsilon > 0$, where $\Delta^{\dagger}$ is the maximum degree of a dual graph of G.

Proof. From Theorem 2.13 together with the bound on the operator norm of the adjacency matrix of planar simple graphs from [31, Cor. 17] we obtain $\lambda(P_{\mathrm{HB}}) \ge (1-\varepsilon)/n$, where P_{HB} is the transition matrix of the heat-bath dynamics for the Potts model on G, and n is the number of vertices, if $\beta \le \varepsilon/\sqrt{3(\Delta-3)}$. Recall that $p = 1 - e^{-\beta}$, which implies $\beta \le p/(1-p)$; we deduce that the assumption on p yields the desired bound on β. By connectedness of G we have $m \ge n - 1$. Thus, Theorem 3.1 completes the proof of the first inequality.

For the second we use $\lambda(\widetilde{P}_{\mathrm{SW}}) \ge \lambda(\widetilde{P}^{\dagger}_{\mathrm{SW}})/(8qm\log m)$ from Theorem 5.6, where $\widetilde{P}^{\dagger}_{\mathrm{SW}}$ is the Swendsen–Wang dynamics for the dual model, i.e. the RC model on the dual graph $G^{\dagger}$ with parameters $p^* = \frac{q(1-p)}{p+q(1-p)}$ and $q \ge 1$. Since $p^* \le q(1-p)$, it is enough to prove $\lambda(\widetilde{P}^{\dagger}_{\mathrm{SW}}) \ge \widetilde{c}(1-\varepsilon)/m$ if $p^* \le \varepsilon/\Delta^{\dagger}$. But, because of $\beta \le p/(1-p)$ and $m \ge n-1$, this follows from Corollary 3.3 and Lemma 2.6. $\blacksquare$

Note that the maximum degree of the dual graph corresponds to the maximum number of edges that are needed to surround a single face in a planar representation of the original graph. This quantity is sometimes called the *maximum face-degree* of G in the literature.

Unfortunately, in the second inequality of Corollary 5.7 the lower bound on p contains $\Delta^{\dagger}$ and not its square root. The reason for this is that the proof of the bound on the

principal eigenvalue of a planar graph requires the graph to be simple, see [31, Cor. 17]. (Note that although Euler's formula is valid also for non-simple graphs, one cannot deduce from it that there is a vertex of small degree.) But since the proof of Corollary 5.7 uses dual graphs, and we cannot guarantee in general that the dual graph is simple, we have to use the stronger assumption. If we considered only 3-connected graphs (see Remark 5.3), then it is known that the dual graph is simple. In this case one could improve the lower bound on p to $1 - \varepsilon/(3q\sqrt{\Delta^\dagger - 3})$ under the additional assumption $\Delta^\dagger \geq 6$.

5.3. Rapid mixing on the square lattice. The goal of this section is to present an application of Theorem 5.6 to the random-cluster model on the two-dimensional square lattice. We prove that the Swendsen–Wang and single-bond dynamics are rapidly mixing for each $q \in \mathbb{N}$ if the parameter p satisfies $p \neq \frac{\sqrt{q}}{1+\sqrt{q}}$. Translated to the q-state Potts model this shows that the Swendsen–Wang dynamics is rapidly mixing at all non-critical temperatures, i.e. at all $\beta \neq \beta_c(q)$ (see Remark 2.9).

This is done by a successive application of the results of the previous chapters. Especially, we need Theorem 3.4, Corollary 3.10, Theorem 4.8 and Theorem 5.6. Note that, since we proved only comparison results between different Markov chains, the results on rapid mixing rely ultimately on the known results on the mixing properties of the heat-bath dynamics for the Potts model (see Section 2.5).

First of all, recall the definition of $\mathbb{Z}_L^2 = (V_{L,2}, E_{L,2})$, i.e. the two-dimensional square lattice of side length L, from Section 2.5 and of $\mathbb{Z}_L^{2\dagger}$ from Section 3.2. It is easy to see that $\mathbb{Z}_L^{2\dagger}$ is indeed the dual graph of $\mathbb{Z}_L^2$ (see Figure 3). Hence, we obtain the following result.

THEOREM 5.8. *Let $\widetilde{P}_{\mathrm{SW}}$ be the transition matrix of the Swendsen–Wang dynamics for the random-cluster model on $\mathbb{Z}_L^2$ with parameters p and q. Let $m = 2L(L-1) = |E_{L,2}|$. Then there exist constants $c_p = c_p(q)$, $c' > 0$ and $C < \infty$ such that*

$$\begin{aligned} \lambda(\widetilde{P}_{\mathrm{SW}}) &\geq c_p/m && \text{for } p < p_c(q), \\ \lambda(\widetilde{P}_{\mathrm{SW}}) &\geq \frac{c_p}{m^2 \log m} && \text{for } p > p_c(q), \\ \lambda(\widetilde{P}_{\mathrm{SW}}) &\geq c' m^{-C} && \text{for } q = 2 \text{ and } p = p_c(2), \end{aligned}$$

where $p_c(q) = \frac{\sqrt{q}}{1+\sqrt{q}}$.

Proof. The first and the last inequality follow from Theorem 3.4, Lemma 2.6 and the fact that $|V_{L,2}| \leq m = 2L(L-1) \leq 2|V_{L,2}|$. For the second let $\widetilde{P}_{\mathrm{SW}}^\dagger$ be the Swendsen–Wang dynamics on $\mathbb{Z}_L^{2\dagger}$ with parameters $p^* = \frac{q(1-p)}{p+q(1-p)}$ and q. Corollary 3.10 (together with Lemma 2.6) implies that $\lambda(\widetilde{P}_{\mathrm{SW}}^\dagger) \geq c_{p^*}/m$ for some $c_{p^*} < \infty$, if $p^* < p_c(q)$. (Recall that $p_c(q) = 1 - e^{-\beta_c(q)}$.) But since $p^* < p_c(q)$ if and only if $p > p_c(q)$, we can deduce the desired bound from Theorem 5.6. ∎

An immediate consequence is the following.

THEOREM 5.9. *Let P_{SW} be the transition matrix of the Swendsen–Wang dynamics for the q-state Potts model on $\mathbb{Z}_L^2$ at inverse temperature β. Let $n = L^2$. Then there exist*

constants $c_\beta = c_\beta(q)$, $c' > 0$ *and* $C < \infty$ *such that*

$$\begin{aligned} \lambda(P_{\mathrm{SW}}) &\geq c_\beta/n && \text{for } \beta < \beta_c(q), \\ \lambda(P_{\mathrm{SW}}) &\geq \frac{c_\beta}{n^2 \log n} && \text{for } \beta > \beta_c(q), \\ \lambda(P_{\mathrm{SW}}) &\geq c' n^{-C} && \text{for } q = 2 \text{ and } \beta = \beta_c(2), \end{aligned}$$

where $\beta_c(q) = \log(1 + \sqrt{q})$.

Proof. Apply Lemma 2.6 to Theorem 5.8 and note, again, that $n \leq |E_{L,2}| \leq 2n$. ∎

The bounds as given in Theorem 5.9 are certainly not optimal. We conjecture $\lambda(P_{\mathrm{SW}})$ to be bounded below by a constant for $\beta \neq \beta_c(q)$. However, this seems to be the first polynomial bound for the Swendsen–Wang dynamics in this regime.

We also obtain the following for the single-bond dynamics.

THEOREM 5.10. *Let* $\widetilde{P}_{\mathrm{SB}}$ *be the transition matrix of the single-bond dynamics for the RC model on* $\mathbb{Z}_L^2$ *with parameters* p *and* q. *Let* $m = 2L(L-1) = |E_{L,2}|$. *Then there exist constants* $c_p = c_p(q)$, $c' > 0$ *and* $C < \infty$ *such that*

$$\begin{aligned} \lambda(\widetilde{P}_{\mathrm{SB}}) &\geq \frac{c_p}{m^2 \log m} && \text{for } p \neq p_c(q), \\ \lambda(\widetilde{P}_{\mathrm{SB}}) &\geq c' m^{-C} && \text{for } q = 2 \text{ and } p = p_c(2), \end{aligned}$$

where $p_c(q) = \frac{\sqrt{q}}{1+\sqrt{q}}$.

Proof. The result for $p \leq p_c(q)$ was already given in Corollary 4.12. For the result for $p > p_c(q)$ take the bound of Corollary 3.10 and apply Lemma 2.6 and Theorem 4.8. ∎

5.4. Graphs of higher genus. In this final section we give a brief description of embeddings of graphs into surfaces, i.e. into two-dimensional topological manifolds, and show that it is possible to extend the results of Section 5.2 to this case. For convenience, we only consider connected, orientable and closed surfaces, which we simply call *surfaces* in what follows. The plan of this section is to define the *genus* of a surface (or a graph) and then, analogously to Section 5.1, to define representations of graphs into surfaces. Then we give a formulation of Euler's formula, define dual graphs and show the consequence for the random-cluster measure.

For this, let Γ_1 and Γ_2 be two surfaces and define their *connected sum* $\Gamma_1 \sqcup \Gamma_2$ as the surface that is obtained by cutting out a small disc from each of the surfaces Γ_1 and Γ_2 and gluing them together along the boundaries of the resulting holes. (See e.g. Giblin [26] for a more formal definition.) It is known that the surface $\Gamma_1 \sqcup \Gamma_2$ does not depend (up to homeomorphisms) on the choice of the discs that are cut out from the primal surfaces (see [26, Prop. 2.17]). We consider the surfaces $\mathbb{S}_h$, $h \geq 0$, which are given by $\mathbb{S}_{k+1} := \mathbb{S}_k \sqcup \mathbb{S}_1$, where $\mathbb{S}_0$ is the two-dimensional sphere and $\mathbb{S}_1$ is the torus. That is, $\mathbb{S}_h$ is the surface that can be obtained from a sphere by adding h handles to it. We call $\mathbb{S}_h$ the (orientable) *surface of genus* h. For example, the sphere is a surface of genus 0 and the torus a surface of genus 1. Note that each connected, closed and orientable surface is homeomorphic to precisely one of the surfaces $\mathbb{S}_h$, $h \geq 0$ (see [53, Thm. 3.1.3]).

Let $G = (V, E)$ be a connected graph. In the light of Section 5.1 we say that G has a *representation in* $\mathbb{S}_h$, namely $\widetilde{G} = (\widetilde{V}, \widetilde{E})$, if there exists a finite set $\widetilde{V} \subset \mathbb{S}_h$, a bijection κ from V to $\widetilde{V}$ and a set of simple, continuous curves $\widetilde{E} = \{\gamma_e : [0,1] \to \mathbb{S}_h : e \in E\}$ in $\mathbb{S}_h$ with the same three properties as given at the beginning of Section 5.1. Here, we consider planar graphs as graphs that admit a representation in the sphere $\mathbb{S}_0$.

The *genus* $\mathbf{g}(G)$ of the graph G is defined to be the smallest integer h such that G has a representation in $\mathbb{S}_h$, and we call a representation of G in $\mathbb{S}_{\mathbf{g}(G)}$ a *minimum genus representation*. For example, the graph $\widetilde{\mathbb{Z}}^2_L$, $L \geq 3$, i.e. the two-dimensional torus (see (2.32)), has genus 1.

Remark 5.11. Every finite graph can be represented in $\mathbb{S}_h$ if h is sufficiently large. For this consider the complete graph K_n, whose genus satisfies $\mathbf{g}(K_n) = \lceil (n-3)(n-4)/12 \rceil$ (see [59]), and observe that every graph on n vertices can be represented in the same surface than K_n (since they are subgraphs). Below we will see that there are constants involved in the results that depend exponentially on the genus. Therefore, the forthcoming results are only useful in cases where the genus is at most logarithmic in the size of the graph.

Recall that $F(\widetilde{G})$ denotes the number of faces of the representation $\widetilde{G}$. It turns out that there is a version of Euler's formula also in this case (see [53, eq. (3.7)]).

Lemma 5.12 (Euler's formula II). *Let $G = (V, E)$ be a connected graph and $\widetilde{G}$ be a minimum genus representation of G. Then*

$$2 - 2\,\mathbf{g}(G) = |V| - |E| + F(\widetilde{G}).$$

Following the ideas of Section 5.1 we fix a minimum genus representation $\widetilde{G}$ of G and fix the representation of the spanning subgraphs $G_A = (V, A)$, $A \subset E$, to be $\widetilde{G}$ without the curves corresponding to $E \setminus A$. We write $F(A)$ for the number of faces in this representation of G_A. Note that a subgraph could be of smaller genus, e.g. $G_\emptyset = (V, \emptyset)$ is obviously planar. However, we equip subgraphs of G with a representation in $\mathbb{S}_{\mathbf{g}(G)}$. We obtain the following corollary, which is the basis of the remaining results.

Corollary 5.13. *Let G be a connected graph and G_A, $A \subset E$, be a spanning subgraph of G. Then*

$$2 - 2\,\mathbf{g}(G) \leq |V| - |A| + F(A) - c(A) + 1 \leq 2.$$

Proof. We prove that the term in the middle cannot increase if we replace A by $A \cup e$, i.e. $|V| - |A \cup e| + F(A \cup e) - c(A \cup e) + 1 \leq |V| - |A| + F(A) - c(A) + 1$. This proves the claim since the first and second inequalities of the statement become equalities for $A = E$ and $A = \emptyset$, respectively. Thus, it is enough to prove $F(A \cup e) - F(A) \leq c(A \cup e) - c(A) + 1$ for all $A \subset E$. The right hand side of this inequality equals 1 if the endpoints of e are connected in (V, A), i.e. $e^{(1)} \overset{A}{\longleftrightarrow} e^{(2)}$, and 0 otherwise. Consequently, we want to verify $F(A \cup e) - F(A) \leq \mathbb{1}(e^{(1)} \overset{A}{\longleftrightarrow} e^{(2)})$, that is, if adding e to A induces a new face, then the endpoints of e are connected in G_A. But this is clearly true, because if a new face is induced by e, then e is contained in a "cycle" in $G_{A \cup e}$, which implies that the endpoints are already connected in G_A. ∎

Given a graph $G = (V, E)$ and a corresponding minimum genus representation $\widetilde{G} = (\widetilde{V}, \widetilde{E})$ we define the *dual graph* $G^\dagger = (V^\dagger, E^\dagger)$ just as in Section 5.1. That is, $G^\dagger$ is the unique graph that has a representation in $\mathbb{S}_{\mathbf{g}(G)}$ with a single vertex in every face of $\widetilde{G}$ and, for every curve $\gamma_e \in \widetilde{E}$, a curve that intersects only γ_e and connects the vertices (or the vertex) on both sides of γ_e. We write $e_\dagger$ for the edge that corresponds to this "dual" curve.

A *dual configuration* $A^\dagger \subset E^\dagger$ of $A \subset E$ is given by

$$e_\dagger \in A^\dagger \;\Leftrightarrow\; e \notin A$$

(cf. (5.1)). It is again easy to convince oneself that every face of the representation of $\widetilde{G}_A$ (adopted from $\widetilde{G}$) contains a connected component of $G^\dagger_{A^\dagger}$, i.e. $F(A) = c(A^\dagger)$. Thus,

$$c(A^\dagger) + |A^\dagger| - 1 \le c(A) - |V| + |E| \le c(A^\dagger) + |A^\dagger| - 1 + 2\,\mathbf{g}(G). \tag{5.5}$$

This implies an analogous result to Lemma 5.4 and, eventually, the following theorem. Recall that the *dual model* to the random-cluster model on G with parameters p and q is the RC model on $G^\dagger$ with parameters p^* and q, where $p^* = \frac{q(1-p)}{p+q(1-p)}$.

THEOREM 5.14. *Let $\widetilde{P}_{\mathrm{HB}}$, $\widetilde{P}_{\mathrm{SB}}$ and $\widetilde{P}_{\mathrm{SW}}$ (resp. $\widetilde{P}^\dagger_{\mathrm{HB}}$, $\widetilde{P}^\dagger_{\mathrm{SB}}$ and $\widetilde{P}^\dagger_{\mathrm{SW}}$) be the transition matrices of the heat-bath, single-bond and Swendsen–Wang dynamics for the random-cluster model on a graph G (resp. for the dual model). Then*

$$\lambda(\widetilde{P}_{\mathrm{HB}}) \le q^{8\,\mathbf{g}(G)}\lambda(\widetilde{P}^\dagger_{\mathrm{HB}})$$

and hence

$$\lambda(\widetilde{P}_{\mathrm{SB}}) \le q^{8\,\mathbf{g}(G)+1}\lambda(\widetilde{P}^\dagger_{\mathrm{SB}}),$$
$$\lambda(\widetilde{P}_{\mathrm{SW}}) \le 8q^{8\,\mathbf{g}(G)+1} m \log m \cdot \lambda(\widetilde{P}^\dagger_{\mathrm{SW}}).$$

Proof. We only prove the first inequality. The other two follow in the same way as Theorem 5.6 follows from Lemma 5.5. An easy computation (similar to the proof of Lemma 5.5) shows

$$q^{-2\,\mathbf{g}(G)}\mu^G_{p,q}(A) \le \mu^{G^\dagger}_{p^*,q}(A^\dagger) \le q^{2\,\mathbf{g}(G)}\mu^G_{p,q}(A)$$

for all $A \subset E$. By the definition of $\widetilde{P}_{\mathrm{HB}}$ (see (2.24)), this implies

$$\widetilde{P}_{\mathrm{HB}}(A, B) \le q^{4\,\mathbf{g}(G)}\widetilde{P}^\dagger_{\mathrm{HB}}(A^\dagger, B^\dagger).$$

But these two inequalities yield $\lambda(\widetilde{P}_{\mathrm{HB}}) \le q^{8\,\mathbf{g}(G)}\lambda(\widetilde{P}^\dagger_{\mathrm{HB}})$ (see Lemma 2.5). ∎

Theorem 5.14 can be used in the same way as Theorem 5.6 to prove results on the mixing properties of the Markov chains involved in specific cases. For example, a result, similar to Corollary 5.7, on graphs of bounded degree can be proven by using Theorem 2.13 and a bound on the principal eigenvalue of graphs with bounded genus (see e.g. [20]). Another application is that lower bounds on the spectral gap of the Markov chains for the Potts model on the two-dimensional square lattice with periodic boundary condition (i.e. on the two-dimensional torus $\widetilde{\mathbb{Z}}^2_L$) at high temperatures can be translated to lower bounds at low temperatures (cf. Section 5.3). We omit the details.

References

[1] M. Aizenman and R. Holley, *Rapid convergence to equilibrium of stochastic Ising models in the Dobrushin Shlosman regime*, in: Percolation Theory and Ergodic Theory of Infinite Particle Systems, IMA Vol. Math. Appl. 8, Springer, New York, 1987, 1–11.

[2] K. S. Alexander, *On weak mixing in lattice models*, Probab. Theory Related Fields 110 (1998), 441–471.

[3] V. Beffara and H. Duminil-Copin, *The self-dual point of the two-dimensional random-cluster model is critical for $q \geq 1$*, Probab. Theory Related Fields 153 (2012), 511–542.

[4] N. Berger, C. Kenyon, E. Mossel, and Y. Peres, *Glauber dynamics on trees and hyperbolic graphs*, Probab. Theory Related Fields 131 (2005), 311–340.

[5] C. Borgs, J. Chayes, and P. Tetali, *Tight bounds for mixing of the Swendsen–Wang algorithm at the Potts transition point*, Probab. Theory Related Fields 152 (2012), 509–557.

[6] F. Cesi, G. Guadagni, F. Martinelli, and R. H. Schonmann. *On the two-dimensional stochastic Ising model in the phase coexistence region near the critical point*, J. Statist. Phys. 85 (1996), 55–102.

[7] J. T. Chayes, L. Chayes, and R. H. Schonmann, *Exponential decay of connectivities in the two-dimensional Ising model*, J. Statist. Phys. 49 (1987), 433–445.

[8] L. Chayes and J. Machta, *Graphical representations and cluster algorithms I. Discrete spin systems*, Phys. A 239 (1997), 542–601.

[9] L. Chayes and J. Machta, *Graphical representations and cluster algorithms II*, Phys. A 254 (1998), 477–516.

[10] C. Cooper, M. E. Dyer, A. M. Frieze, and R. Rue, *Mixing properties of the Swendsen–Wang process on the complete graph and narrow grids*, J. Math. Phys. 41 (2000), 1499–1527.

[11] C. Cooper and A. M. Frieze, *Mixing properties of the Swendsen–Wang process on classes of graphs*, Random Structures Algorithms 15 (1999), 242–261.

[12] P. Cuff, J. Ding, O. Louidor, E. Lubetzky, Y. Peres, and A. Sly, *Glauber dynamics for the mean-field Potts model*, J. Statist. Phys. 149 (2012), 432–477.

[13] Y. Deng, T. M. Garoni, J. Machta, G. Ossola, M. Polin, and A. D. Sokal, *Critical behavior of the Chayes–Machta–Swendsen–Wang dynamics*, Phys. Rev. Lett. 99 (2007), 055701.

[14] Y. Deng, X. Qian, and H. W. J. Blöte, *Single-cluster dynamics for the random-cluster model*, Phys. Rev. E 80 (2009), 036707.

[15] P. Diaconis and L. Saloff-Coste, *Comparison theorems for reversible Markov chains*, Ann. Appl. Probab. 3 (1993), 696–730.

[16] P. Diaconis and D. Stroock, *Geometric bounds for eigenvalues of Markov chains*, Ann. Appl. Probab. 1 (1991), 36–61.

[17] R. Diestel, *Graph Theory*, 2nd ed., Grad. Texts in Math. 173, Springer, New York, 2000.

[18] R. L. Dobrushin and S. B. Shlosman, *Completely analytical interactions: constructive description*, J. Statist. Phys. 46 (1987), 983–1014.

[19] J. Du, B. Zheng, and J.-S. Wang, *Dynamic critical exponents for Swendsen–Wang and Wolff algorithms obtained by a nonequilibrium relaxation method*, J. Statist. Mech. 2006, P05004.

[20] Z. Dvořák and B. Mohar, *Spectral radius of finite and infinite planar graphs and of graphs of bounded genus*, J. Combin. Theory Ser. B 100 (2010), 729–739.

[21] M. Dyer, L. A. Goldberg, M. Jerrum, and R. Martin, *Markov chain comparison*, Probab. Surv. 3 (2006), 89–111.

[22] M. Dyer, A. Sinclair, E. Vigoda, and D. Weitz, *Mixing in time and space for lattice spin systems: a combinatorial view*, Random Structures Algorithms 24 (2004), 461–479.

[23] S. F. Edwards and A. D. Sokal, *Generalization of the Fortuin–Kasteleyn–Swendsen–Wang representation and Monte Carlo algorithm*, Phys. Rev. D 38 (1988), 2009–2012.

[24] C. M. Fortuin and P. W. Kasteleyn, *On the random-cluster model. I. Introduction and relation to other models*, Physica 57 (1972), 536–564.

[25] Q. Ge and D. Štefankovič, *A graph polynomial for independent sets of bipartite graphs*, in: Proc. 30th Internat. IARCS Annual Conf. on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2010), LIPIcs Leibniz Int. Proc. Inform. 8, Dagstuhl, Leibniz-Zent. Inform., Wadern, 2010, 240–250.

[26] P. Giblin, *Graphs, Surfaces and Homology*, 3rd ed., Cambridge Univ. Press, Cambridge, 2010.

[27] R. J. Glauber, *Time-dependent statistics of the Ising model*, J. Math. Phys. 4 (1963), 294–307.

[28] V. K. Gore and M. R. Jerrum, *The Swendsen–Wang process does not always mix rapidly*, in: Proc. 29th Annual ACM Sympos. on Theory of Computing STOC '97, ACM, New York, 1997, 674–681.

[29] G. Grimmett, *The Random-Cluster Model*, Grundlehren Math. Wiss. 333, Springer, Berlin, 2006.

[30] J. M. Harris, J. L. Hirst, and M. J. Mossinghoff, *Combinatorics and Graph Theory*, 2nd ed., Undergrad. Texts in Math., Springer, New York, 2008.

[31] T. P. Hayes, *A simple condition implying rapid mixing of single-site dynamics on spin systems*, in: Proc. 47th Annual IEEE Sympos. on Foundations of Computer Science (Berkeley, CA, 2006), IEEE Comput. Soc., Los Alamitos, CA, 2006, 39–46.

[32] R. Holley, *Possible rates of convergence in finite range, attractive spin systems*, in: Particle Systems, Random Media and Large Deviations (Brunswick, ME, 1984), Contemp. Math. 41, Amer. Math. Soc., Providence, RI, 1985, 215–234.

[33] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, Cambridge, 1990.

[34] M. Huber, *A bounding chain for Swendsen–Wang*, Random Structures Algorithms 22 (2003), 43–59.

[35] M. Jerrum and A. Sinclair, *Polynomial-time approximation algorithms for the Ising model*, SIAM J. Comput. 22 (1993), 1087–1116.

[36] E. Kreyszig, *Introductory Functional Analysis with Applications*, Wiley, New York, 1978.

[37] L. Laanait, A. Messager, S. Miracle-Solé, J. Ruiz, and S. Shlosman, *Interfaces in the Potts model. I. Pirogov–Sinai theory of the Fortuin–Kasteleyn representation*, Comm. Math. Phys. 140 (1991), 81–91.

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

[39] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov Chains and Mixing Times*, Amer. Math. Soc., Providence, RI, 2009.

[40] X. Li and A. D. Sokal, *Rigorous lower bound on the dynamic critical exponents of the Swendsen–Wang algorithm*, Phys. Rev. Lett. 63 (1989), 827–830.

[41] Y. Long, *Mixing time of the Swendsen–Wang dynamics on the complete graph and trees*, PhD thesis, Univ. of California, Berkeley, 2009.

[42] Y. Long, A. Nachmias, and Y. Peres, *Mixing time power laws at criticality*, in: Proc. 48th Annual IEEE Sympos. on Foundations of Computer Science (Providence, RI, 2007), IEEE Comput. Soc., Los Alamitos, CA, 205–214.

[43] E. Lubetzky and A. Sly, *Critical Ising on the square lattice mixes in polynomial time*, Comm. Math. Phys. 313 (2012), 815–836.

[44] F. Martinelli, *Dynamical analysis of low-temperature Monte Carlo cluster algorithms*, J. Statist. Phys. 66 (1992), 1245–1276.

[45] F. Martinelli, *Lectures on Glauber dynamics for discrete spin models*, in: Lectures on Probability Theory and Statistics (Saint-Flour, 1997), Lecture Notes in Math. 1717, Springer, Berlin, 1999, 93–191.

[46] F. Martinelli and E. Olivieri, *Finite volume mixing conditions for lattice spin systems and exponential approach to equilibrium of Glauber dynamics*, in: Cellular Automata and Cooperative Systems (Les Houches, 1992), NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. 396, Kluwer, Dordrecht, 1993, 473–490.

[47] F. Martinelli and E. Olivieri, *Approach to equilibrium of Glauber dynamics in the one phase region. I. The attractive case*, Comm. Math. Phys. 161 (1994), 447–486.

[48] F. Martinelli and E. Olivieri, *Approach to equilibrium of Glauber dynamics in the one phase region. II. The general case*, Comm. Math. Phys. 161 (1994), 487–514.

[49] F. Martinelli, E. Olivieri, and R. H. Schonmann, *For 2-D lattice spin systems weak mixing implies strong mixing*, Comm. Math. Phys. 165 (1994), 33–47.

[50] F. Martinelli, E. Olivieri, and E. Scoppola, *On the Swendsen–Wang dynamics. I. Exponential convergence to equilibrium*, J. Statist. Phys. 62 (1991), 117–133.

[51] F. Martinelli, E. Olivieri, and E. Scoppola, *On the Swendsen–Wang dynamics. II. Critical droplets and homogeneous nucleation at low temperature for the two-dimensional Ising model*, J. Statist. Phys. 62 (1991), 135–159.

[52] F. Martinelli, A. Sinclair, and D. Weitz, *Glauber dynamics on trees, boundary conditions and mixing time*, Comm. Math. Phys. 250 (2004), 301–334.

[53] B. Mohar and C. Thomassen, *Graphs on Surfaces*, Johns Hopkins Stud. Math. Sci., Johns Hopkins Univ. Press, Baltimore, MD, 2001.

[54] E. Novak and H. Woźniakowski, *Tractability of Multivariate Problems. Volume II, Standard Information for Functionals*, EMS Tracts Math. 12, Eur. Math. Soc., Zürich, 2010.

[55] L. Onsager, *Crystal statistics. I. A two-dimensional model with an order-disorder transition*, Phys. Rev. 65 (1944), 117–149.

[56] G. Ossola and A. D. Sokal, *Dynamic critical behavior of the Swendsen–Wang algorithm for the three-dimensional Ising model*, Nuclear Phys. B 691 (2004), 259–291.

[57] J. G. Propp and D. B. Wilson, *Exact sampling with coupled Markov chains and applications to statistical mechanics*, Random Structures Algorithms 9 (1996), 223–252.

[58] D. Randall and D. Wilson, *Sampling spin configurations of an Ising system*, in: Proc. Tenth Annual ACM-SIAM Sympos. on Discrete Algorithms SODA '99, SIAM, Philadelphia, 1999, 959–960.

[59] G. Ringel and J. W. T. Youngs, *Solution of the Heawood map-coloring problem*, Proc. Nat. Acad. Sci. U.S.A. 60 (1968), 438–445.

[60] D. Rudolf, *Explicit error bounds for lazy reversible Markov chain Monte Carlo*, J. Complexity 25 (2009), 11–24.

[61] D. Rudolf, *Error bounds for computing the expectation by Markov chain Monte Carlo*, Monte Carlo Methods Appl. 16 (2010), 323–342.

[62] D. Rudolf, *Explicit error bounds for Markov chain Monte Carlo*, Dissertationes Math. 485 (2012), 93 pp.

[63] J. Salas and A. Sokal, *Dynamic critical behavior of the Swendsen–Wang algorithm: The two-dimensional three-state Potts model revisited*, J. Statist. Phys. 87 (1997), 1–36.

[64] R. H. Schonmann, *Second order large deviation estimates for ferromagnetic systems in the phase coexistence region*, Comm. Math. Phys. 112 (1987), 409–422.

[65] D. W. Stroock and B. Zegarliński, *The equivalence of the logarithmic Sobolev inequality and the Dobrushin–Shlosman mixing condition*, Comm. Math. Phys. 144 (1992), 303–323.

[66] D. W. Stroock and B. Zegarliński, *The logarithmic Sobolev inequality for discrete spin systems on a lattice*, Comm. Math. Phys. 149 (1992), 175–193.

[67] R. H. Swendsen and J.-S. Wang, *Nonuniversal critical dynamics in Monte Carlo simulations*, Phys. Rev. Lett. 58 (1987), 86–88.

[68] M. Ullrich, *Comparison of Swendsen–Wang and heat-bath dynamics*, Random Structures Algorithms 42 (2013), 520–535.

[69] M. Ullrich, *Rapid mixing of Swendsen–Wang and single-bond dynamics in two dimensions*, arXiv:1202.6321 (2012).

[70] M. Ullrich, *Swendsen–Wang is faster than single-bond dynamics*, SIAM J. Discrete Math. 28 (2014), 37–48.

Index