

Lecture Note 8

Instructor: September 22

Disclaimer: *These notes have not been subjected to the usual scrutiny accorded to formal publications. They may be distributed outside this class only with the permission of the Instructor.*

Introduction

In the last lecture we introduced the concept of coupling of Markov chains and its use in bounding the mixing time in some simple examples. In this lecture we explore its use in more challenging scenarios.

8.1 Graph Colorings

One of the most spectacular successes of coupling has been in the analysis of MCMC algorithms for the benchmark problem of sampling colorings of a graph u.a.r. As we have seen, this leads to an fpras for counting colorings, or equivalently, for the partition function of the antiferromagnetic Potts model at zero temperature (where the only legal configurations are proper colorings). Throughout, our input will be an undirected graph $G = (V, E)$ and a set of colors $[q]$. We will write n for the number of vertices of G and Δ for its maximum degree. Recall that we assume $q \geq \Delta + 1$, in which case a q -coloring exists and can be found easily by a greedy algorithm. For $q = \Delta$ the decision problem, while still polynomial time solvable [Bro41], is non-trivial, and for $q \leq \Delta - 1$ even approximate counting is impossible unless $\text{NP}=\text{RP}$ [GŠV15].

A natural Markov chain for this problem is the Glauber dynamics, as discussed in an earlier lecture¹. Given any proper coloring of G , this process does the following:

- Pick a vertex $v \in V$ u.a.r. and a color $c \in [q]$ u.a.r.
- Recolor v with c if this yields a proper coloring, else do nothing.

This Markov chain is symmetric and aperiodic, but is not always irreducible: in particular, if $q \leq \Delta + 1$, we can have “frozen” colorings in which no move is possible (even though other proper colorings do exist). A simple example is when G is a clique on $\Delta + 1$ vertices.

However, it turns out that this problem cannot occur when the number of colors is at least $\Delta + 2$:

Important Exercise: show that the above Markov chain is irreducible if $q \geq \Delta + 2$.

Here are two famous conjectures regarding graph colorings:

1. Random sampling of proper colorings can be done in polynomial time whenever $q \geq \Delta + 1$.

¹Strictly speaking, the Markov chain given below differs very slightly from the (heat-bath) Glauber dynamics of the previous lecture: in that version, we would replace the color at v with a color chosen uniformly from the set of allowable colors at v . Analytically it is slightly more convenient to work with the more uniform version here, which picks from all possible colors. Effectively this just introduces a self-loop probability at each state.

2. The above Glauber dynamics has mixing time $O(n \ln n)$ whenever $q \geq \Delta + 2$.

Unfortunately, we cannot yet (quite!) prove either of these conjectures. However, much progress has been made on the second conjecture over the past 30 years, culminating in a very recent result that largely settles the second one (see the end of this note). We will begin by giving an elementary coupling argument that works when $q \geq 4\Delta + 1$. Note that intuitively we would expect that, the more colors we have, the easier it should be to prove that the Markov chain is rapidly mixing; this is because with more colors the vertices become “more independent.”

Claim 8.1. *Provided that $q \geq 4\Delta + 1$, the mixing time of the Glauber dynamics is $O(n \log n)$.*

This theorem (in the stronger version $q \geq 2\Delta + 1$) is due to Jerrum [Jer95], and independently to Salas and Sokal [SS97]. Note that $O(n \log n)$ mixing time is (at least intuitively) the best we could hope for, since by coupon collecting it takes that long before all the vertices have a chance to be recolored. (Actually the coupon collecting analogy is not strictly accurate; however, the $\Omega(n \log n)$ lower bound—for this, and indeed for any spin system on a bounded degree graph—can be proved by a more careful argument [HS05].)

Proof. We couple two copies X_t and Y_t of the chain, as follows: Let X_t and Y_t choose the same vertex v and the same color c at each step, and recolor v with c if possible. Let us now analyze this coupling.

Define $d_t := d(X_t, Y_t)$ to be the number of vertices where the colors disagree. For every step in our chain a vertex v and a color c are chosen, which could result in:

- “Good moves” (d_t decreases by 1): the chosen vertex v is a disagreeing vertex, and the chosen color c is not present at any neighbor of v in either X_t or Y_t . In this case, the coupling will recolor v to the same color in both X_t and Y_t , thus eliminating one disagreeing vertex. Since the neighbors of v have at most 2Δ distinct colors in X_t and Y_t , there are at least $d_t(q - 2\Delta)$ good moves available.
- “Bad moves” (d_t increases by 1): the chosen vertex v is not a disagreeing vertex but is a neighbor of some disagreeing vertex v' , and the chosen color c is one of the colors of v' (in either X_t or Y_t). In this case, v will be recolored in one of the chains but not the other, resulting in a new disagreement at v . Since each disagreeing vertex v' has at most Δ neighbors, and there are only two corresponding “bad” colors c at v' , the number of bad moves available is at most $2d_t\Delta$.
- “Neutral moves” (d_t is unchanged): all choices of (v, c) that do not fall into one of the above two categories lead to no change in d_t .

Note that the difference between the numbers of good and bad moves is (at least) $d_t(q - 4\Delta)$, so we expect the distance to decrease when $q \geq 4\Delta + 1$. We now make this argument precise.

Since each move has the same probability, namely $\frac{1}{qn}$, we can compute the expected change in d_t under one step of the coupling:

$$\mathbb{E}[d_{t+1}|X_t, Y_t] \leq d_t - \frac{d_t(q - 2\Delta)}{qn} + \frac{2d_t\Delta}{qn} = d_t \left(1 - \frac{q - 4\Delta}{qn}\right).$$

Iterating this we get that

$$\mathbb{E}[d_t|X_0, Y_0] \leq \left(1 - \frac{q - 4\Delta}{qn}\right)^t d_0 \leq \left(1 - \frac{q - 4\Delta}{qn}\right)^t n \leq \left(1 - \frac{1}{qn}\right)^t n. \quad (8.1)$$

This implies that

$$\begin{aligned} \Pr[d_t > 0 | X_0, Y_0] &= \Pr[d_t \geq 1 | X_0, Y_0] \\ &\leq \mathbb{E}[d_t | X_0, Y_0] && \text{(Markov's Ineq.)} \\ &\leq \left(1 - \frac{1}{qn}\right)^t n. \end{aligned}$$

To ensure that $\Delta(t) \leq \varepsilon$, it is sufficient to let $t = qn(\ln n + \ln \varepsilon^{-1})$. Thus in particular the mixing time is $\tau_{\text{mix}} = O(qn \log n)$. \square

Remark: The fact that the above mixing time bound increases with q is an artefact of our crude approximation $q - 4\Delta \geq 1$ in the last step of line (8.1). Clearly (**exercise!**) the value q in the final bound on τ_{mix} can be replaced by $\frac{q}{q-4\Delta} \leq 4\Delta + 1$.

Exercise (harder): Devise an improved coupling that achieves a similar $O(n \log n)$ mixing time for $q \geq 2\Delta + 1$. [Hint: Instead of X_t, Y_t always choosing the same color, couple pairs of “bad” colors in the neighborhood of a vertex and make X_t, Y_t each choose different members of the pair. We’ll see a much easier way to get this improvement in the next section.]

8.2 Path coupling

“Path coupling,” an idea introduced by Bubley and Dyer [BD97], is a powerful engineering tool that makes it much easier to design couplings in complex examples. We’ll apply it to our colorings example to obtain a simple proof that the coupling still works for $q \geq 2\Delta + 1$.

Definition 8.2. A pre-metric on Ω is a connected undirected graph with positive edge weights with the property that every edge is a shortest path. We call two elements $x, y \in \Omega$ adjacent if (x, y) is an edge in the pre-metric.

Notice that a pre-metric extends to a metric in the obvious way (just take shortest path distances in the graph of the pre-metric).

Path coupling says that, when defining a coupling for a Markov chain on Ω , it is enough to specify the coupling *only for pairs of states that are adjacent in the pre-metric*. This will usually be much easier than specifying the coupling for arbitrary pairs of states.

This fact is expressed in the following theorem:

Theorem 8.3. Suppose there exists a coupling $(X, Y) \rightarrow (X', Y')$ defined only on pairs (X, Y) that are adjacent in the pre-metric such that

$$\mathbb{E}[d(X', Y') | X, Y] \leq (1 - \alpha)d(X, Y) \text{ for some } \alpha \in [0, 1], \quad (8.2)$$

where d is the metric extending the pre-metric. Then this coupling can be extended to a coupling which satisfies (8.2) on all pairs (X, Y) .

Note that (8.2) says that the distance between X and Y (as measured in the metric d) decreases in expectation by a factor $\alpha > 0$. Just as in our analysis of the Glauber dynamics for graph coloring in the previous section, assuming d takes non-negative integer values this immediately leads to a bound on the mixing time of $\tau_{\text{mix}} = O\left(\frac{1}{\alpha} \log D\right)$, where D is the maximum distance between any two states. (In that application, we had $\alpha = \frac{1}{qn}$ and $D = n$.)

Proof. Let (X, Y) be arbitrary, not necessarily adjacent. Consider any shortest path in the pre-metric

$$X = Z_0 \rightarrow Z_1 \rightarrow Z_2 \rightarrow \dots \rightarrow Z_{k-1} \rightarrow Z_k = Y.$$

We construct a coupling of one move of (X, Y) by composing couplings for each adjacent pair (Z_i, Z_{i+1}) along this path, as follows:

- Map (Z_0, Z_1) to (Z'_0, Z'_1) according to the coupling.
- For each $i \geq 1$ in sequence, map (Z_i, Z_{i+1}) to (Z'_i, Z'_{i+1}) according to the coupling, but conditional on Z'_i already being chosen.

This process constructs a coupling $(X, Y) \rightarrow (X', Y') = (Z'_0, Z'_k)$. [**Exercise:** Convince yourself that this is indeed a valid coupling!]

Now the expected change in distance under the coupling can be analyzed as follows:

$$\begin{aligned} \mathbb{E}[d(X', Y') | X, Y] &\leq \mathbb{E} \left[\sum_{i=0}^{k-1} d(Z'_i, Z'_{i+1}) \middle| X, Y \right] \\ &\leq (1 - \alpha) \sum_{i=0}^{k-1} d(Z_i, Z_{i+1}) \\ &= (1 - \alpha)d(X, Y), \end{aligned}$$

which establishes (8.2) as required. \square

8.2.1 Application to graph coloring

We now apply path coupling to the Glauber dynamics for graph coloring, and see how it leads to a simpler, more elegant and tighter analysis.

Theorem 8.4. *Provided $q \geq 2\Delta + 1$, $\tau_{\text{mix}} = O(n \log n)$.*

Proof. We will use the following pre-metric: two colorings X, Y are adjacent iff they differ at exactly one vertex. In this case, we set $d(X, Y) = 1$. We will extend this pre-metric to the *Hamming metric*: $d(X, Y)$ is the number of vertices at which X, Y differ.

Note, however, that in order to have the Hamming distance indeed be the extension of the pre-metric described above, we need to let our state space include non-valid colorings as well. This is because the shortest path from one coloring to another might involve temporarily assigning illegal colors to vertices. We keep the transitions as before, i.e., we do not allow the MC to make a transition to an invalid coloring. Thus the state space is not irreducible, but rather it consists of a single irreducible component (namely, all proper colorings) plus some transient components consisting of invalid colorings. It is easy to see that this Markov chain converges to the uniform distribution on the proper colorings, as before; and, moreover, that a bound on the mixing time derived using coupling extends to a bound on the mixing time of the original Markov chain (without the invalid colorings). [**Exercise:** Verify this by going back to the basic coupling bound in the previous note.]

Now let X, Y be any two (not necessarily proper) colorings that are *adjacent*; this means that they differ at only one vertex, say v_0 . We define our coupling for (X, Y) as follows:

- Pick the same vertex v in both chains.

- If v is not in the neighborhood $N(v_0)$ of the unique disagreeing vertex v_0 , then pick the same color c in both chains.
- If $v \in N(v_0)$, match up the choice of colors as follows:

$$\begin{aligned} c_X &\longleftrightarrow c_Y \\ c_Y &\longleftrightarrow c_X \\ c &\longleftrightarrow c, \end{aligned}$$

where c_X, c_Y are the colors of v_0 in X, Y respectively, and $c \notin \{c_X, c_Y\}$.

- Having chosen a vertex v and a color c in both chains, recolor v with c in each chain if possible.

This is clearly a valid coupling because each of X, Y , viewed in isolation, makes a move according to the Markov chain. Note that we don't need to explicitly define the coupling on other pairs (X, Y) as this is taken care of by path coupling.

To analyze the above coupling, we again count the numbers of “good” and “bad” moves. A “good” move corresponds to choosing the disagreeing vertex v_0 and some color not present among its neighbors; hence there are at least $q - \Delta$ good moves. A “bad” move corresponds to picking a neighbor $v \in N(v_0)$ together with one choice of color (namely, the combination c_Y in X and c_X in Y , for then v is recolored to different colors in the two chains and becomes a new disagreeing vertex; note that the complementary choice c_X in X and c_Y in Y is not a bad move because neither chain will recolor v in this case). Hence the number of bad moves is at most Δ (i.e., the number of neighbors of v_0). All other moves are neutral (i.e., do not change the distance).

Since each move occurs with the same probability $\frac{1}{qn}$, and since $d(X_t, Y_t) = 1$ for all adjacent pairs X_t, Y_t , we can conclude that

$$\mathbb{E}[d(X_{t+1}, Y_{t+1}) | X_t, Y_t] \leq \left(1 - \frac{q - 2\Delta}{qn}\right) d(X_t, Y_t).$$

Thus, assuming $q \geq 2\Delta + 1$, (8.2) holds with $\alpha = \frac{q - 2\Delta}{qn}$ and $D = n$, so we get $\tau_{\text{mix}} = O(n \log n)$ as claimed. \square

Theorem 8.4 was first proved by Jerrum [Jer95] (see also Salas and Sokal [SS97]) using a direct coupling rather than path coupling (see the exercise at the end of the previous section). The resulting analysis, though still elementary, is quite a bit more delicate than that above.

8.3 Going beyond the 2Δ bound

The ultimate goal of research on MCMC for graph coloring is to resolve the following:

Conjecture 8.5. τ_{mix} is $O(n \log n)$ for all $q \geq \Delta + 2$.

Note that this is precisely the range of q for which the Markov chain is guaranteed to be connected. The Conjecture is known to hold for the very special case of Δ -regular trees [MSW06].

There is a further interesting related conjectured connection between colorings on trees and on general graphs. The range $q \geq \Delta + 1$ is precisely the region in which the Gibbs measure on the infinite Δ -regular tree is *unique*, i.e., if we set any boundary condition (fixed coloring) on the leaves of the tree, the asymptotic

influence of this boundary condition on the distribution of the color at the root of the tree goes to zero with distance if and only if $q \geq \Delta + 1$. It is believed that the same holds for general graphs (i.e., among all graphs of maximum degree Δ , the decay of influence with distance is slowest for the tree). Resolving either this conjecture, or the related one above for Markov chain mixing, would be of great interest in combinatorics, computer science and statistical physics.

We now briefly summarize the large body of work over the past 30 years aimed at resolving Conjecture 8.5; this work contains many important ideas that have other applications in the analysis of Markov chains.

8.3.1 Ideas for improving $2\Delta \rightarrow 1.76\Delta$

Recall our path coupling analysis in the proof of Theorem 8.4 above. Our bound of $q - \Delta$ on the number of good moves there was pessimistic, because in a typical coloring we would expect significantly fewer than Δ colors (the maximum possible number) to be represented in the neighborhood $N(v_0)$.

Let $A(X, v)$ denote the number of *available* colors at v in coloring X , i.e., the number of colors not represented in $N(v)$. Since the number of bad moves is (at most) Δ , the previous analysis will still go through provided we have $A(X, v) > \Delta$.

To get a handle on $A(X, v)$, suppose each vertex is colored independently and u.a.r. in X . Then by linearity of expectation we have

$$\mathbb{E}[A(X, v)] = q \left(1 - \frac{1}{q}\right)^\Delta \approx qe^{-\frac{\Delta}{q}}$$

is the expected value. Thus we will have $\mathbb{E}[A(X, v)] > \Delta$ provided $qe^{-\frac{\Delta}{q}} > \Delta$, which is true whenever $q > \alpha\Delta$ where α is the solution to $x = e^{\frac{1}{x}}$. In particular, $\alpha \approx 1.76$. So we might hope that we get $O(n \log n)$ mixing time for $q > 1.76\Delta$.

This is the crux of the proof presented in [DF06]. However, quite a bit of work remains to be done: we need to justify why it is OK to work with a random coloring; in a random coloring, the neighbors of v are not colored independently; and we cannot work just with the *expected* value $\mathbb{E}[A(X, v)]$.

To sketch how to turn the above intuition into a real proof, we follow the development of [HV05].

Definition 8.6. *Say that a coloring X of G satisfies the local uniformity property if, for all v , $A(X, v)$ is at least $q(e^{-\Delta/q} - \delta)$.*

Here δ is an arbitrarily small positive constant. Thus local uniformity says that the number of available colors at all vertices is not much less than the expected value we calculated above.

The following Fact captures formally the intuition from our informal calculation above:

Fact 8.7. *Let G be triangle-free and have maximum degree Δ . Assume that the number of colors satisfies $q \geq \max\{\Delta + 2/\delta, C \log n\}$, where C is a constant that depends on δ . Then a random q -coloring of G satisfies the local uniformity property (with constant δ) with probability at least $1 - O(n^{-4})$.*

The Fact is really only interesting in the case where $\Delta \geq C \log n$. In that case the condition on q is only $q \geq \Delta + \text{const}$, which will certainly hold in our application.

We will not prove this Fact here; the proof follows from a straightforward application of large deviation bounds on the colors of the neighbors of a given vertex v , conditional on an arbitrary fixed coloring of the rest of the graph. (Note that, since G is triangle-free, these colors are conditionally independent. And since the number of neighbors is $\geq C \log n$, large deviation bounds hold.) For the details, see [HV05].

To use this fact, we assume $\Delta \geq C \log n$ and we return to our original coupling idea (without path coupling). Now we note that, rather than coupling two arbitrary initial states (X_0, Y_0) , it is enough to couple an arbitrary state X_0 with a *stationary* Y_0 , i.e., in the coupling we may assume that Y_0 is a uniformly random coloring. **[Exercise:** Check this by going back to our original justification of coupling in the previous note.] So at all times t , Y_t is a uniformly random coloring, and thus by Fact 8.7 it is locally uniform with high probability. In fact, if we let $T = cn \log n$ (for some constant c), then applying a union bound over times t we see that Y_t satisfies local uniformity for $t = 0 \dots T$ with probability $\geq 1 - O(n^{-2})$.

Now, exploiting local uniformity to bound the number of good moves as sketched earlier, we get an expected contraction $1 - O(\frac{1}{n})$ in distance at each step, from which we can conclude that if $q \geq 1.76\Delta$ then $\Pr[X_T \neq Y_T] \leq O(n^{-2})$. (Here we are choosing the constant c large enough that the probability of not having coupled is this small. Note that some work still needs to be done here; in particular, we need to check that it is enough that just one of the two colorings, Y_t , satisfies local uniformity.)

Hence, we obtain that

$$\mathbb{E}[d(X_T, Y_T)|X_0] \leq n(O(n^{-2}) + O(n^{-2})) = O(n^{-1}),$$

where the first $O(n^{-2})$ term bounds the probability that local uniformity fails, and the second bounds the probability that $\Pr[X_T \neq Y_T]$, and the factor n comes from the fact that we can always bound $d(X_T, Y_T)$ by its maximum value n . Thus, using Markov's inequality we get

$$\Pr[d(X_T, Y_T) > 0|X_0] = \Pr[d(X_T, Y_T) \geq 1|X_0] \leq \mathbb{E}[d(X_T, Y_T)|X_0] \leq O(n^{-1}).$$

This establishes the result with $q \geq 1.76\Delta$ for triangle-free graphs of degree $\Delta \geq C \log n$.

See [HV05] for the missing details of the above argument.

8.3.2 Further Reading

There was a huge amount of activity on this problem in the early 2000's, starting several years after the initial $q > 2\Delta$ result; see [FV07] for a comprehensive survey.

While a graduate student at Berkeley, Eric Vigoda [Vig99] was the first to show (by comparing it to a slightly more complicated Markov chain) that the Glauber dynamics has mixing time $O(n \log n)$ for q below 2Δ ; specifically, his result holds for $q \geq \frac{11}{6}\Delta$. This remains essentially the best known result for a local Markov chain without any additional restrictions on the graph; the constant was recently improved to $\frac{11}{6} - \varepsilon$ for a very small $\varepsilon > 0$ [CDM⁺19].

Dyer and Frieze [DF06] were the first to obtain $\tau_{\text{mix}} = O(n \log n)$ for $q \geq \alpha\Delta$, with $\alpha \approx 1.76$, but using a more complicated argument than the one sketched above. (Rather than assuming Y_0 is stationary, they instead argued that, after sufficiently many steps T , Y_T satisfies a similar local uniformity condition.) Molloy [Mol04] improved the bound to $q \geq \beta\Delta$, where $\beta \approx 1.49$, for graphs meeting the above requirements (i.e., triangle-free and maximum degree $\Delta = \Omega(\log n)$). In [HV03], Hayes and Vigoda used a very delicate *non-Markovian* coupling to demonstrate $O(n \log n)$ mixing time for $q \geq (1 + \epsilon)\Delta$, but the $\Delta = \Omega(\log n)$ requirement remains, and moreover the associated constant in the O -expression depends on ϵ . Dyer et al. [DFHV04] show that the $q \geq \beta\Delta$ result of [Mol04] holds for graphs of girth ≥ 6 and sufficiently large constant degree. It is still open whether the stronger $(1 + \epsilon)\Delta$ result of [HV03] holds for constant-degree graphs. Random graphs with constant (average) degree are discussed in [DFV06], where it is shown that, with high probability over the choice of the graph, $O(n \log n)$ mixing time holds for q down to about $O(\log \log n)$, which is much smaller than the maximum degree (which is of order $\frac{\log n}{\log \log n}$).

Very recently, using more sophisticated techniques based on so-called *spectral independence* (a property that we will touch on later in the course), Chen *et al.* [CLMM23] made a significant breakthrough by almost

proving Conjecture 8.5. Their result requires the very slightly stronger lower bound $q \geq \Delta + 3$, and, more significantly, a lower bound of $\Omega_{\Delta}(1)$ on the girth of the graph (i.e., a constant that depends on Δ).

References

- [BD97] R. Bubley and M.E. Dyer. Path coupling: a technique for proving rapid mixing in Markov chains. *Proceedings of the 38th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 223–231, 1997.
- [Bro41] R.L. Brooks. On colouring the nodes of a network. *Mathematical Proceedings of the Cambridge Philosophical Society*, 37:194–197, 1941.
- [CDM⁺19] S. Chen, M. Delcourt, A. Moitra, G. Perarnau, and L. Postle. Improved bounds for randomly sampling colorings via linear programming. *Proceedings of ACM-SIAM SODA*, pages 2216–2234, 2019.
- [CLMM23] Z. Chen, K. Liu, N. Mani, and A. Moitra. Strong spatial mixing for colorings on trees and its algorithmic applications. ArXiv preprint 2304.01954. 2023.
- [DF06] M. Dyer and A. Frieze. Randomly coloring graphs with lower bounds on girth and maximum degree. *Random Structures and Algorithms*, 2006.
- [DFV06] M. Dyer, A. Flaxman, A. Frieze, and E. Vigoda. Randomly coloring sparse random graphs with fewer colors than the maximum degree. *Random Structures and Algorithms*, 29:450–465, 2006.
- [DFHV04] M. Dyer, A. Frieze, T.P. Hayes, and E. Vigoda. Randomly coloring constant degree graphs. *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 582–589, 2004.
- [FV07] A.M. Frieze and E. Vigoda. A survey on the use of Markov chains to randomly sample colourings. In *Combinatorics, Complexity and Chance*, pages 53–71. Oxford Lecture Series in Applied Math., Vol. 34, Oxford University Press, 2007.
- [GŠV15] A. Galanis, D. Štefankovič, and E. Vigoda. Inapproximability for antiferromagnetic spin systems in the tree nonuniqueness region. *Journal of the ACM*, 62, 2015.
- [HS05] T. Hayes and A. Sinclair. A general lower bound for mixing of single-site dynamics on graphs. *Proceedings of the 46th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 511–520, 2005.
- [HV03] T. Hayes and E. Vigoda. A non-Markovian coupling for randomly sampling colorings. *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, pages 618–627, 2003.
- [HV05] T. Hayes and E. Vigoda. Coupling with the stationary distribution and improved sampling for colorings and independent sets. *Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 971–979, 2005.
- [Jer95] M. Jerrum. A very simple algorithm for estimating the number of k-colourings of a low-degree graph. *Random Structures and Algorithms*, 7:157–165, 1995.
- [Mol04] M. Molloy. The Glauber dynamics on colorings of a graph with high girth and maximum degree. *SIAM Journal on Computing*, 33:721–737, 2004.
- [MSW06] F. Martinelli, A. Sinclair, and D. Weitz. Fast mixing for independent sets, colorings and other models on trees. *Random Structures & Algorithms*, 31:134–172, 2006.

- [SS97] J. Salas and A.D. Sokal. Absence of phase transition for antiferromagnetic Potts models via the Dobrushin uniqueness theorem. *Journal of Statistical Physics*, 86:551–579, 1997.
- [Vig99] E. Vigoda. Improved bounds for sampling colorings. *Proceedings of the 40th IEEE Symposium on Foundations of Computer Science*, pages 51–59, 1999.