

# Subexponential mixing for partition chains on grid-like graphs

Alan Frieze\*

Wesley Pegden†

## Abstract

We consider the problem of generating uniformly random partitions of the vertex set of a graph such that every piece induces a connected subgraph. For the case where we want to have partitions with linearly many pieces of bounded size, we obtain approximate sampling algorithms based on Glauber dynamics which are fixed-parameter tractable with respect to the bandwidth of  $G$ , with simple-exponential dependence on the bandwidth. For example, for rectangles of constant or logarithmic width this gives polynomial-time sampling algorithms. More generally, this gives sub-exponential algorithms for bounded-degree graphs without large expander sub-graphs (for example, we obtain  $2^{O(\sqrt{n})}$  time algorithms for square grids).

In the case where we instead want partitions with a small number of pieces of linear size, we show that Glauber dynamics can have exponential mixing time, even just for the case of 2 pieces, and even for 2-connected sub-graphs of the grid with bounded bandwidth.

## 1 Introduction

In this paper we consider the mixing time of Glauber dynamics for uniform sampling of partitions of the vertex-set of a bounded-degree graph into  $q$  subsets which each induce connected subgraphs of  $G$ . Our positive results in this paper focus on the case where  $q$  is linear in the number of vertices of  $G$  and each subset in the partition is constrained to have bounded size  $\leq B$ ; in particular, this generalizes the setting of the well-studied monomer-dimer model [21, 20] where  $B = 2$ .

The problem of randomly partitioning a graph into connected pieces arises naturally in the context of evaluating political districtings, where Markov chain methods have been applied and developed over the past decade. (For this application, our positive results correspond to the case where there are many small districts; in the U.S. context, one might imagine the case of the state legislative districting of a state, rather than its districting into US Congressional districts.) While there are rigorous statistical approaches based on Markov chains which can avoid the sampling problem when the only goal is outlier detection [9, 8], the benefits of sampling have led to the development of a large range of Markov chains intended to generate good samples from partition spaces [11, 15, 17, 4, 7, 2, 3]. The mixing properties of these Markov chains, however, have only been supported by heuristic evidence, and there is still a dearth of results allowing rigorous approximate sampling from specified target distributions on such partitions.

The mixing times we prove will depend on the *bandwidth* of  $G$ , which is the minimum over labelings  $\sigma$  of the vertices of  $G$  by distinct numbers  $1, \dots, n$  of the maximum difference  $\sigma(u) - \sigma(v)$  of any adjacent pair  $u \sim v$  of vertices:

$$(1.1) \quad \text{band}(G) = \min_{\sigma} \max_{u \sim v} |\sigma(u) - \sigma(v)|.$$

(Throughout, we will use  $\sigma$  to denote a labeling realizing this minimum.) In particular, when  $G$  is a  $k \times \ell$  grid with  $n = k\ell$  vertices we have  $\text{band}(G) = \min(k, \ell)$  [10], and will obtain running times of order  $\text{poly}(n)2^{O(\min(k, \ell))}$ .

In this paper, a contiguous  $q$ -partition of  $G$  is a partition of the vertex set into  $q$  (not necessarily nonempty) pieces, each inducing a connected subgraph of  $G$ , and a  $B$ -bounded contiguous  $q$ -partition is a contiguous  $q$ -partition in which every piece has size  $\leq B$ .

In general, it can be NP-hard to determine whether a given graph admits a partition into connected pieces of prescribed sizes, even for planar graphs of bounded degree [14]. But things change if we allow some slack in the

\*Carnegie Mellon University. Research Supported in part by NSF grant DMS1952285. email: frieze@cmu.edu

†Carnegie Mellon University. Research Supported in part by NSF grant DMS1700365. email: wes@math.cmu.edu

size constraints. For example, we will see that one can efficiently partition any graph of maximum degree  $\Delta$  into connected pieces of sizes between  $B/(\Delta - 1)$  and  $B$ , for any  $B \geq 1$ . The Hamiltonicity of grid graphs means they can be partitioned into connected pieces of any sizes summing to the total number of vertices; i.e., essentially no slack is needed at all. Moreover, in both of these cases, the partitioning can also be done locally, without global modifications to the graph. The following definition captures these properties:

**DEFINITION 1.1.** *A family of graphs  $\mathcal{G}$  is  $(B, \alpha)$ -partitionable if for any  $\varepsilon > 0$  and  $q > (1 + \varepsilon)n/(\alpha B)$ , there's a  $K$  sufficiently large and a  $\delta > 0$  such for any sufficiently large  $G \in \mathcal{G}$ :*

- *There is a polynomial time algorithm to find a  $B$ -bounded contiguous  $q$ -partition of  $G$ , and*
- *Given any  $B$ -bounded contiguous- $q$ -partition  $\omega$  of  $G$ , there are  $\delta|V(G)|$  distinct  $B$ -bounded contiguous partitions  $\omega_i$  of  $G$ , using fewer than  $q$  partition classes, which each agree with  $\omega$  except on a connected sub-graph  $H_i$  of  $G$  with at most  $K$  vertices.*

Essentially, the second part of Definition 1.1 requires that any  $B$ -bounded contiguous partition using sufficiently many nonempty classes can be locally modified to use fewer nonempty color classes. Note that no infinite family of graphs can be  $(B, \alpha)$ -partitionable for any value  $\alpha > 1$ . On the other hand, we will prove the following result about bounded-degree graphs:

**PROPOSITION 1.1.** *For any  $B$ , connected graphs of maximum degree  $\Delta$  are  $(B, \frac{1}{\Delta-1} \frac{B-1}{B})$ -partitionable.*

Proposition 1.1 cannot be improved in general, but, for example, grids satisfy a strong version of Definition 1.1:

**PROPOSITION 1.2.** *Rectangular grids are  $(B, 1)$ -partitionable for all  $B$ .*

Our main result is the following:

**THEOREM 1.1.** *If  $\mathcal{G}$  is a  $(B, \alpha)$ -partitionable family of graphs of bounded maximum degree, then for any*

$$q \geq (1 + \varepsilon) \frac{n}{\alpha B},$$

*Glauber dynamics for the space of  $B$ -bounded contiguous  $q$ -partitions of  $G$  has mixing time at most*

$$\text{poly}(n) 2^{O(\text{band}(G))}.$$

We define the Glauber dynamics for this state space in Section 2.1.

The mixing time here depends exponentially just on the bandwidth of  $G$ ; they give algorithms for approximate sampling that are fixed-parameter-tractable for the bandwidth, and in particular we have subexponential (i.e.,  $2^{o(n)}$ ) algorithms for  $B$ -bounded contiguous  $q$ -partitions any time  $\mathcal{G}$  is a family of graphs with sublinear bandwidth. From a result of Böttcher, Pruessmann, Taraz, and Würfl [6], this is equivalent to requiring graphs in  $\mathcal{G}$  to have sublinear treewidth, or requiring that graphs in  $\mathcal{G}$  have no linear-size subgraphs with good expansion. Thus our results give subexponential time algorithms to approximately uniformly sample  $B$ -bounded contiguous  $q$ -partitions in all these cases (and for all planar graphs, by the planar separator theorem). Note, for example, that a square grid graph on  $n$  vertices has bandwidth  $\sqrt{n}$ , as can be realized by the row-by-row ordering if its vertices.

Najt, Deford, and Solomon analyzed the worst-case complexity of sampling connected  $q$ -partitions in the case where  $q$  is bounded and thus the typical piece-size is linear [22]. They show, for example that it is NP-hard to sample connected 2-partitions for planar graphs. They also show that sampling connected  $q$ -partitions (for bounded  $q$ ) is fixed-parameter tractable in the treewidth, although with a tower function dependence on the treewidth. On Markov Chains, they show that there are families of graphs (even among, say, 3-connected planar graphs of bounded degree) for which Glauber dynamics is torpidly mixing for the case of connected 2-partitions. In Section 3, we will show that Glauber dynamics can even be torpidly mixing for  $q = 2$  on bounded-bandwidth subgraphs of the grid. In particular, we will give an example of a family of graphs which are simply the union of two overlapping grid graphs of bounded bandwidth, with exponential mixing time.

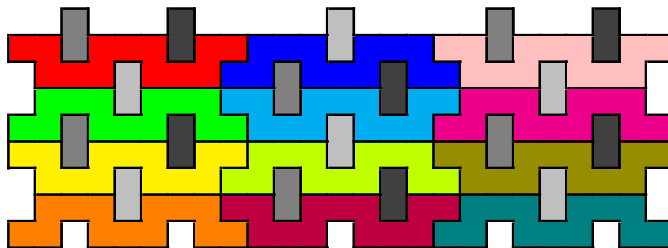


Figure 1: When tiled, this configuration has maximum partition class size 13, average class size  $32/5 = 6.4$ , but is rigid (no Glauber dynamics transitions are possible) if  $B = 13$  unless there are unused colors. Any change to existing pieces which preserves continuity either disconnects one of the large pieces, or increases the size of a large piece beyond the threshold  $B = 13$ .

## 2 Proofs of positive results

**2.1 Markov chains for partitions** In this section we define the Glauber dynamics for contiguous partitions. For a fixed connected graph  $G$  on  $n$  vertices, we let  $\Omega_{q,B,G}$  consist of all  $B$ -bounded contiguous- $q$ -partitions of  $G$ .

We consider a Markov Chain  $\mathcal{M}$  on  $\Omega_{q,B,G}$  defined by the following transition procedure from a state  $\omega \in \Omega_{q,B,G}$ :

- (a) Choose a uniformly random vertex  $v \in G$ ;
- (b) Choose a uniformly random set among all sets in the partition  $\omega$ ;
- (c) Remove  $v$  from its current set and make it a member of the randomly chosen set, if this move would result in a partition belonging to the state space  $\Omega_{q,B,G}$ .

When step (b) is carried out, there is a question of how one handles sets which occur multiple times in a fixed partition class  $\omega$ —do we consider  $\omega$  to be a set of partition classes or a multiset? The only set which can occur multiple times as a partition class is the empty set, which will appear  $q - \kappa$  times, where  $\kappa$  is the number of nonempty color classes in  $\omega$ . In the case where we include the empty-set with multiplicity in the uniform random selection in Step (b), the stationary distribution of  $\mathcal{M}$  is not uniform on partitions but on colorings; e.g., partitions with more nonempty classes are more likely. In the case where we include the empty-set without multiplicity, the uniform distribution on partitions is stationary. For simplicity, we work here just in the second setting, where the uniform distribution on partitions is stationary.

In fact,  $\mathcal{M}$  can fail to be irreducible on the state space  $\Omega_{q,B,G}$ , even when  $q = Cn$  and  $B$  is large and  $G$  is a grid graph, as can be seen from the example in Figure 1. But we will define  $\Omega_{q,B,G}^K$  to be those  $q$ -partitions  $\omega$  of  $G$  which agree with some  $\omega'$  in  $\Omega_{q,B,G}$  on all but at most  $K$  vertices. This is the chain we will show is rapidly mixing using the method of canonical paths. Note that that  $|\Omega_{q,B,G}^K| = O(n^K |\Omega_{q,B,G}|)$  and so we can use  $\Omega_{q,B,G}^K$  to obtain near uniform samples from  $\Omega_{q,B,G}$  in polynomial time.

**2.2 Three Lemmas** A crucial ingredient of our proof is the so-called *mountain-climbing problem* [18, 16]. Roughly speaking, this problem asserts that two mountain climbers, each beginning at sea level, can traverse a 2-dimensional mountain range in such a way that they meet at the summit, and at all times are at equal altitudes (Figure 2). Note that, in general, this requires the mountain climbers to change directions (and sometimes walk away from the summit), and that quadratically many such changes can be needed, in the number of local extrema of the mountain range.

To state the version we will use, define, for any integrable function  $f : S^1 \rightarrow \mathbb{R}$  whose integral on the whole circle is 0, the set of intervals  $\mathcal{I}^f$  to be those intervals of the circle on which the integral of  $f$  is 0. We include singletons and the whole circle in  $\mathcal{I}^f$  as degenerate cases. We then have the following:

**THEOREM 2.1.** *Let  $f : S^1 \rightarrow \{-1, +1\}$  be any function on the circle with finitely many discontinuities, and  $\int_{S^1} f(x)dx = 0$ . Then there is a continuous function  $I : [0, T] \rightarrow \mathcal{I}^f$  with  $I(0)$  a singleton and  $I(T) = S^1$ .*

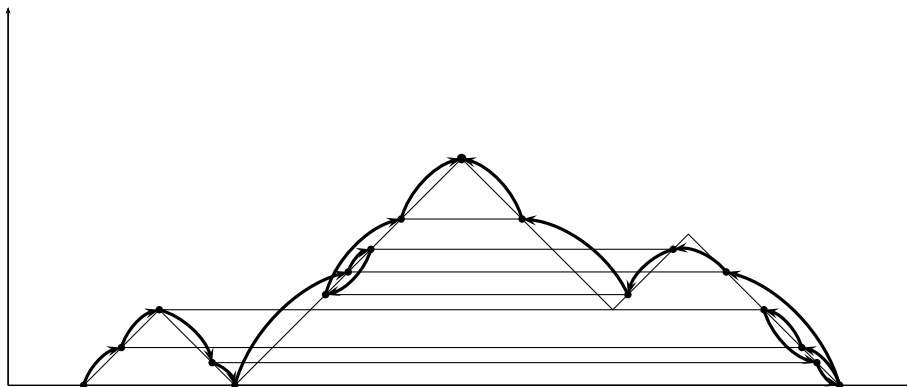


Figure 2: The mountain climbing lemma: two mountain climbers starting at the same height, which is a global minimum for the mountain range, can coordinate their movements so that they are at all times at equal heights, and at some point reach a common summit.

The connection between this and the version of the mountain climbing lemma depicted in Figure 2 is made by defining a function  $F$  on the circle by  $F(x) = \int_{t=0}^x f(x)$ , where integration is done, say, in the counterclockwise direction. Note that  $F$  is continuous and in our setup, is piecewise linear with finitely many linear pieces. Taking  $x_0$  to be the location of a minimum of this function, the mountain climbing lemma ensures that there are continuous paths  $\alpha : [0, 1] \rightarrow S^1$  and  $\beta : [0, 1] \rightarrow S^1$  from  $x_0$  which never cross  $x_0$ , which eventually meet, and which have  $F(\alpha(t)) = F(\beta(t))$  at all times. These functions give the endpoints of the intervals defined by the function  $I(t)$ .

Our analysis in this paper will make use of the following discrete version of Theorem 2.1:

LEMMA 2.1. *Let the vertices  $v_1, v_2, \dots, v_n$  of the  $n$  cycle  $C_n$  (indexed in cyclic order) be assigned labels  $\ell(v_i) \in [-K, K]$ , such that*

$$\sum_{i=1}^n \ell(v_i) = r.$$

*Then there is a sequence of subsets*

$$\emptyset = X_0, X_1, X_2, \dots, X_T = V(C_n)$$

*of  $V(C_n)$ , such that each  $X_i$ ,  $0 < i < T$  is a path, such the symmetric difference of  $X_i$  and  $X_{i+1}$  is single vertex for all  $0 \leq i < T$ , and such that*

$$\forall 0 \leq i \leq T, \quad \left| \sum_{v \in X_i} \ell(v) - \frac{|X_i|}{n} r \right| \leq 2K.$$

*Proof.* It suffices to prove the lemma in the case when  $r = 0$ ; the general case reduces to this case, by replacing each  $\ell(v_i)$  with  $\ell(v_i) - r/n$ .

Let  $L = \sum_i |\ell(v_i)|$ , and, with the aim of applying Theorem 2.1, recursively define intervals  $I_0, I_1, \dots, I_n$  corresponding to each vertex  $v_i$  of  $C_n$ , each of the form  $[x_i, x_{i+1})$ , where the length  $x_{i+1} - x_i$  of  $I_i$  is given by

$$x_{i+1} - x_i = \frac{1}{L} |\ell(v_i)|.$$

We then define  $f$  on  $S^1$  to be 1 on intervals  $I_i$  whose corresponding vertex  $v_i$  has a positive label  $\ell(v_i)$ , and  $-1$  otherwise. Theorem 2.1 gives us the continuous interval function  $I : [0, T] \rightarrow \mathcal{I}^f$ ; let the corresponding left- and right-endpoint functions be  $a(t)$  and  $b(t)$ .

Now define  $\mathcal{X}(t)$  to be the cyclically ordered list of intervals of  $I_i$  intersecting  $I(t)$ . We define a graph whose vertex set  $Y \subseteq V(C_n) \times V(C_n)$  consisting of all the pairs  $(v_i, v_j)$  for which there exists a  $t$  such that  $\mathcal{X}(t) = (I_i, I_{i+1}, I_{i+2}, \dots, I_j)$ . Define two vertices  $(v_i, v_j)$  and  $(v_s, v_t)$  in this graph to be adjacent if there is a

$t_0$  such that for all  $\delta > 0$ ,  $\mathcal{X}((t_0 - \delta, t_0 + \delta))$  includes both  $(I_i, \dots, I_j)$  and  $(I_s, \dots, I_t)$ . By continuity of  $I(t)$ , we have that  $s \in \{i - 1, i, i + 1\}$  and  $t \in \{j - 1, j, j + 1\}$ . Viewing the points on the edges and vertices of this graph as a topological space,  $I(t)$  induces a continuous path in this topological space from a pair  $(v_i, v_i)$  whose corresponding list of intervals is a singleton, to a pair  $(v_i, v_{i-1})$  whose corresponding list includes every interval. In particular, these two vertices lie in the same connected component of this graph and there is a (discrete) path in the graph from one to the other. To obtain the statement of the theorem, we replace any edges of the path  $(v_i, v_j) \sim (v_s, v_t)$  for which  $s = i \pm 1$  and  $t = j \pm 1$  with a path of length 2 of the form  $(v_i, v_j) \sim (v_i, v_t) \sim (v_s, v_t)$ .

Since the integral of  $\int_{I(t)} f(x)dx = 0$  for any  $t$ , and the sum of the lengths of the intervals in  $\mathcal{X}(t)$  differs from this integral by at most  $2K$  for any  $t$ , this path gives the sequence of subsets claimed by the theorem.  $\square$

Interestingly, this paper is not the first time the mountain climbing has been applied to Markov chains, having seen earlier use in [1, 5, 13]. In some sense the mountain-climbing lemma can be viewed as a fundamental result about balanced route-planning, a perspective that can explain its utility in the construction of canonical paths, where it is essential to define paths that stay within the constraints defining the boundary of the state space of one's Markov chain, while avoiding veering so far from the boundary that one is inevitably over-utilizing a small volume of high-dimensional space and thus introducing unacceptable levels of congestion to the collection of paths.

The next lemma will be used to control the congestion near the endpoints of our canonical paths.

**LEMMA 2.2.** *For any bipartite graph on the bipartition  $A \dot{\cup} B$ , where the maximum degree of vertices in  $B$  is  $D$  and the minimum degree of vertices in  $A$  is  $\varepsilon D$ , there is a mapping  $\theta : A \rightarrow B$  such  $(a, \theta(a))$  is an edge for all  $a \in A$ , and every preimage  $\theta^{-1}(b)$  has size at most  $\lceil 1/\varepsilon \rceil$ .*

*Proof.* We can modify the bipartite graph by replacing each vertex  $b \in B$  with  $k = \lceil 1/\varepsilon \rceil$  copies, each joined to the same neighbors in  $A$ . The resulting graph has the property that the minimum degree on the left is at least the maximum degree on the right; by Hall's theorem, there is a complete matching of  $A$ , which induces the desired function  $\theta$ .  $\square$

The final lemma in this section simply captures the balance we can ensure when partitioning arbitrary connected graphs of bounded degree into connected pieces.

**LEMMA 2.3.** *If  $T$  is a tree on  $n \geq 2$  vertices and of maximum degree at most  $\Delta$  with root  $v_0$  of degree  $\leq \Delta - 1$ , then for any  $B < n$ , we can partition  $T$  into subtrees  $T''$  and  $T'$  so that  $\frac{B-1}{\Delta-1} < |V(T'')| \leq B$  and  $v_0 \in T'$ .*

*Proof.* Consider the trees  $T_1, T_2, \dots, T_d$  ( $d \leq \Delta - 1$ ) that are the connected components of  $T \setminus v_0$ , and let  $v_i$  be the neighbor of  $v_0$  in  $T_i$ . Since  $\frac{n-1}{\Delta-1} > \frac{B-1}{\Delta-1}$ , there is some  $i$  such that  $|V(T_i)| > \frac{B-1}{\Delta-1}$ . If  $|V(T_i)| \leq B$  as well, we're done; otherwise, we apply induction to the tree  $T_i$  rooted at  $v_i$ .  $\square$

**2.3 Partitionable graphs** In this section we revisit Definition 1.1. We begin by proving Proposition 1.1, showing that graphs of maximum degree  $\Delta$  are  $(B, \frac{1}{\Delta-1} \frac{B-1}{B})$ -partitionable.

*Proof.* [Proof of Proposition 1.1] Applying Lemma 2.3 recursively shows that we can partition  $G$  into pieces of size greater than  $\frac{B-1}{\Delta-1}$  and at most  $B$ . Moreover this can be done in polynomial time since at each application, there are only linearly many choices for the division claimed to exist by the Lemma. This proves the first part of the definition holds as claimed.

We will show the second part of the definition is satisfied for any

$$(2.2) \quad K = \lceil 2/\varepsilon \rceil B\Delta + 1.$$

Let  $G$  be any connected graph and  $c$  any  $B$ -bounded contiguous- $\kappa$ -coloring of  $G$  with

$$\kappa \geq \frac{(1 + \varepsilon)n(\Delta - 1)}{B - 1}.$$

Define the graph  $H$  whose vertex-set is the color classes of  $c$ , with two color classes being adjacent in  $H$  if they are joined by any edge in  $G$ . Note that the maximum degree of  $H$  is at most  $B \cdot \Delta$ . Letting  $k = \lceil \frac{2}{\varepsilon} \rceil$ ,

we see from Lemma 2.3 that we can partition  $H$  into connected pieces  $H_1, \dots, H_s$  each of whose size satisfies  $k < |V(H_i)| \leq k(B\Delta - 1) + 1$ , except for one piece that may have smaller size. In particular, by merging the small piece with another, we may assume that each piece satisfies

$$k < |V(H_i)| \leq kB\Delta + 1.$$

Each  $H_i$  corresponds to a subgraph  $G(H_i)$  of  $G$  whose vertex set is the union of some color classes of  $c$ . Since the average size of a color class of  $c$  is less than  $\frac{1}{1+\varepsilon} \frac{B-1}{\Delta-1}$ , there is a  $\delta > 0$  (depending on  $B, \Delta, \varepsilon$ ) and  $\delta|V(G)|$  indices  $i_1, i_2, \dots$  such that each  $G(H_{i_j})$  is the union of color classes of  $c$  which have average size less than  $\frac{1}{1+\varepsilon/2} \frac{B-1}{\Delta-1}$ , for all  $j$ .

By Lemma 2.3, this time applied to a  $G(H_{i_j})$ , there is a partition of  $G(H_{i_j})$  into pieces of size at most  $B$  and size at least  $\frac{B-1}{\Delta-1}$ , except for one piece that may have smaller size; if the partition of  $G(H_{i_j})$  has  $\leq k$  classes then we're done, as this partition can be used to give an alternative coloring  $c'$  which agrees with  $c$  except on  $G(H_{i_j})$ , which satisfies  $|G(H_{i_j})| \leq K$  by (2.2). But the partition cannot have more than  $k$  classes, since then the average size of a piece in our partition would be at least

$$\frac{k}{k+1} \frac{B-1}{\Delta-1} \geq \frac{1}{1+\varepsilon/2} \frac{B-1}{\Delta-1}.$$

Thus for each of the  $\delta|V(G)|$  indices  $i_j$ , we obtain a distinct contiguous coloring using fewer than  $\kappa$  colors, and differing from  $c$  on a connected subgraph of most  $K$  vertices.  $\square$

Next we prove Proposition 1.2, asserting that grids are  $(B, 1)$ -partitionable for all  $B$ .

*Proof.* [Proof of Proposition 1.2] The first part of the definition follows from the fact that grid graphs have Hamilton paths, and paths can be partitioned as claimed.

Next, let  $G$  be a  $k \times \ell$  grid on  $n = k\ell$  vertices and assume  $k \leq \ell$ . We define parameters in two cases:

**Case 1:** If  $k \leq 400B^2/\varepsilon^2$ , we set  $L = k$ , and  $M = \lceil \frac{20B}{\varepsilon} \rceil$ .

**Case 2:** If  $k, \ell > 400B^2/\varepsilon^2$ , we set  $L = M = \lceil \frac{20B}{\varepsilon} \rceil$ .

By the choice of  $L$ , there is a  $k' \times \ell'$  subgrid  $G'$  of  $G$  with  $k'$  and  $\ell'$  multiples of  $L$  and  $M$ , respectively, and for which  $n' := k'\ell' \geq n - 40B/\varepsilon$ . In particular, for  $n$  sufficiently large, given any  $B$ -bounded contiguous- $\kappa$ -coloring  $c$  of  $G$  for  $\kappa \geq (1 + \varepsilon)n/B$ , at least  $(1 + \varepsilon/2)n/B$  of its color classes intersect  $G'$ . We can then partition  $G'$  into  $L \times M$  grid subgraphs  $H_i$ . In Case 1 the area of each such grid subgraph is  $\geq 20kB/\varepsilon$  and for each  $H_i$ , the set  $B(H_i)$  of boundary vertices—i.e., vertices that have neighbors in other  $H_j$ 's—has size at most  $2k$  for all  $i$ . In Case 2, the area of each  $H_i$  is  $\geq 400B^2/\varepsilon^2$  and each boundary  $B(H_i)$  has size at most  $80B/\varepsilon + 4$ . In both cases, we have

$$(2.3) \quad \frac{|B(H_i)|}{|V(H_i)|} \leq \frac{\varepsilon}{4B}.$$

Thus, fixing a  $B$ -bounded contiguous- $\kappa$ -coloring, at most  $\varepsilon n/4B$  color classes can intersect a boundary vertex of an  $H_i$ . Subtracting from the  $\geq (1 + \varepsilon/2)n/B$  color classes that intersect  $G'$ , we are left with at least

$$(1 + \varepsilon/4)n/B \geq (1 + \varepsilon/4)n'/B$$

color classes that are subsets of  $V(G')$  and intersect exactly one  $H_i$ . Since no  $H_i$  can contain more than  $L \cdot M$  color classes, we have for

$$\delta = \frac{\varepsilon}{8BLM}$$

that there are  $\delta n$  indices  $i$  such that there are at least  $(1 + \varepsilon/8)LM/B > LM/B$  color classes which are subsets of  $H_i$ . Each of these can be used to provide a contiguous- $\kappa'$ -coloring for  $\kappa' < \kappa$  satisfying the definition. Indeed, as each such  $H_i$  is Hamiltonian, each can be partitioned into contiguous pieces of arbitrary sizes summing to  $|V(H_i)|$ ; in particular, each can be partitioned into color classes of size exactly  $B$ , which in each case requires just  $LM/B$  color classes per  $H_i$ .  $\square$

The first part of Definition 1.1 ensures that we can efficiently find an element of  $\Omega_{q,B,G}$  as a starting point for our algorithm. Theorem 1.1 will then follow from bounds on the mixing times of the Markov chain, which we achieve using the method of canonical paths.

**2.4 Outline of the proof strategy** We bound the mixing time of  $\mathcal{M}$  using the method of canonical paths [12, 19, 23], which requires us to define a path between any pair of states so that no transition is used by too many paths.

The basic idea of our approach to defining canonical paths is simple. Given the graph  $G$ , we use the bandwidth ordering  $\sigma$  and divide the vertex set  $G$  into intervals of length  $2B\text{band}(G)$  in the ordering  $\sigma$ ; this ensures that any piece of a  $B$ -bounded contiguous  $q$ -partition intersects at most two of these intervals.

Next, given two valid partitions  $\omega$  and  $\omega'$  of  $G$ , the basic idea is that we would like to, one interval at a time, replace the classes of  $\omega$  in one interval with the classes of  $\omega'$ . First (in Phase 0 below), we must reduce the number of nonempty partition classes so that we have enough room, e.g., to break classes into singletons.

But when carrying out these copy operations, the crucial thing we need to maintain is that all intermediate states use few enough partition classes that they are valid members of the state space, but also sufficiently many partition classes that they are not a high-congestion node (note that we should expect there to be few partitions which use few classes; thus it is not surprising that canonical paths cannot make heavy use of such partitions as intermediate states). This is where the mountain climbing problem enters. Lemma 2.1 ensures that there is some order in which we can copy (and uncopy!) the partition  $\omega'$  into the partition  $\omega$  so that we eventually have entirely replaced  $\omega$  with  $\omega'$ , but at all intermediate states have a “just right” number of nonempty partition classes.

**2.5 Defining Valid Canonical Paths** For  $G$  an  $(\alpha, B)$ -partitionable graph, and any  $\varepsilon > 0$ , let  $K$  be as given in Definition 1.1. We now define canonical paths between any states  $\omega, \omega'$  of  $\Omega = \Omega_{q, B, G}^K$ , for  $q > (1 + \varepsilon)n/(\alpha B)$ . Define  $\kappa(\omega)$  to be the number of non-null color classes  $\omega$ . We let

$$t = 2B\text{band}(G).$$

**Phase 0** From the definition of  $(B, \alpha)$ -partitionable and our choice of  $q > (1 + \varepsilon)n/(\alpha B)$ , we can define paths

$$(2.4) \quad \omega = \omega_0, \omega_1, \dots, \omega_{s_0}$$

and

$$(2.5) \quad \omega' = \omega'_0, \omega'_1, \dots, \omega'_{t_0}$$

in  $\Omega$  such that  $\omega_{s_0}$  and  $\omega'_{t_0}$  both have fewer than  $q - 10t$  nonempty color classes, and where  $s_0, t_0 \leq 10Kt$ . For example, applying the definition of  $(B, \alpha)$ -partitionable to  $\omega$  gives us a sequence  $\omega = \omega_0, \dots, \omega_k$ , for  $k < K$ , where  $\omega_k \in \Omega_{q-1, B, G}$  and for all  $0 < i < k$  we have  $\omega_i \in \Omega_{q-1, B, G}^K$ . Applying this repeatedly, we eventually obtain a sequence ending in the state  $\omega_{s_0} \in \Omega_{q-10t, B, G}$ , with all intermediate states in  $\Omega_{q-1, B, G}^K$ .

To reduce congestion, we choose the sequences (2.4), (2.5) carefully, as follows. For each  $\kappa$  with  $q - 10t \leq \kappa \leq q$ , we define  $\Omega_\kappa \subseteq \Omega$  to be those colorings with exactly  $\kappa$  nonempty color classes. We define a bipartite graph on  $A \dot{\cup} B$  for  $A = \Omega_\kappa, B = \Omega_{\kappa-1}$  by letting  $\omega$  be adjacent to  $\omega'$  in this graph for  $\omega \in \Omega_\kappa, \omega' \in \Omega_{\kappa-1}$  if  $\omega'$  differs from  $\omega$  only on a connected subgraph of at most  $K$  vertices. From Definition 1.1, there is  $\delta > 0$  such that the degree of a vertex  $\omega \in A$  is least  $\delta n$ . On other hand, very crudely, the number of subsets of  $V(G)$  of size  $K$  inducing connected subgraphs of  $G$  is at most  $n\Delta^K$ , and each such set can be partitioned in  $\leq 2^{\Delta K}$  ways, by choosing which edges lie within versus between partition classes. This gives an upper bound on the maximum degree of the bipartite graph. Thus by Lemma 2.2, there is a mapping  $\theta_\kappa : \Omega_\kappa \rightarrow \Omega_{\kappa-1}$  such that  $\theta_\kappa(\omega)$  differs from  $\omega$  only on a connected subgraph of at most  $K$  vertices, and such that for all  $\omega' \in \Omega_{\kappa-1}$ , the preimage  $\theta_\kappa^{-1}(\omega')$  has size at most, say,

$$(2.6) \quad \frac{3^{\Delta K}}{\delta}.$$

We fix choices of the maps  $\theta_\kappa$  for each  $\kappa$ , and use these maps when choosing sequences in (2.4) and (2.5). For notational convenience, we write  $\bar{\omega} = \omega_{s_0}$  and  $\bar{\omega}' = \omega'_{t_0}$ .

**Setting up the next step** It remains to define a path from  $\bar{\omega}$  to  $\bar{\omega}'$ .

We let  $\Phi(\omega)$  be the set of nonempty color classes of  $\omega$  and let  $\Phi_i(\omega)$  denote the set of color classes  $C \in \Phi(\omega)$  for which

$$\min_{v \in C} \left\lfloor \frac{\sigma(v)}{t} \right\rfloor = i.$$

Observe that for  $i = 0, \dots, m - 1$  ( $m = \lfloor n/t \rfloor + 1$ ) the  $\Phi_i(\omega)$  form a partition of  $\Phi(\omega)$ . We let  $V(\Phi_i(\omega))$  denote the underlying vertex set of  $\Phi_i(\omega)$ ; that is, the union of its elements. Observe that by our choice of  $t$ , we have

$$(2.7) \quad (\forall \omega \in \Omega) \quad V(\Phi_i(\omega)) \subseteq V(\Phi_i(\omega)) \cup V(\Phi_{i+1}(\omega)), \quad \text{and}$$

$$(2.8) \quad (\forall \omega, \omega' \in \Omega) \quad V(\Phi_i(\omega')) \subseteq V(\Phi_{i-1}(\omega)) \cup V(\Phi_i(\omega)) \cup V(\Phi_{i+1}(\omega)),$$

We write  $\phi_i(\omega)$  for the size of  $\Phi_i(\omega)$ ; i.e., the number of color classes it contains, and similarly write  $\phi(\omega)$  for the size of  $\Phi(\omega)$ .

Now we label the vertices  $v_0, v_1, \dots, v_m = v_0$  of the cycle  $C_m$  (indexed in cyclic order) with the labels

$$\ell(v_i) = \phi_i(\bar{\omega}') - \phi_i(\bar{\omega}),$$

which satisfy

$$|\ell(v_i)| \leq t$$

and let

$$(2.9) \quad X_0, X_1, \dots, X_T$$

be the sequence of subsets of the vertices of the cycle guaranteed to exist by Lemma 2.1. Note that we can assume  $T \leq m^2/2$ , since in the worst case all pairs of endpoints are used by endpoints of the paths  $X_i$ .

Recall that each  $X_i$  in  $X_1, \dots, X_{T-1}$  induces a subgraph of  $C_m$  that is a path, whose vertex set we call  $I_i$ . Its complement is also a path, and we denote the interior of this complementary path by  $O_i$ . Finally, we let  $B_i = C_m \setminus (I_i \cup O_i)$ ; so, for all  $i = 1, \dots, T - 1$ ,  $|B_i| = 2$ , and in this case we write  $b_k^+$  and  $b_k^-$  for the vertices in  $B_k$  which are adjacent to the greatest and least vertices of  $I_k$ , respectively, in cyclic order.

Now with Phase 1 beginning from the state  $\bar{\omega}_0 = \bar{\omega}$ , we will define Phases  $1, \dots, T$  so each Phase  $k$  ends at a state  $\bar{\omega}_k$  such that:

- (a) For each  $v_i \in I_k$ , every color class  $C' \in \Phi_i(\bar{\omega}')$  is a color class of  $\bar{\omega}_k$ .
- (b) For each  $v_i \in O_k$ , every color class  $C \in \Phi_i(\bar{\omega})$  is a color class of  $\bar{\omega}_k$ .
- (c) For  $v_i = b_k^+$ , every vertex in  $(V(\Phi_i(\bar{\omega})) \cup V(\Phi_{i-1}(\bar{\omega}))) \setminus V(\Phi_{i-1}(\bar{\omega}'))$  is a singleton color class in  $\bar{\omega}_k$ , while for  $v_i = b_k^-$ , every vertex in  $(V(\Phi_i(\bar{\omega})) \cup V(\Phi_{i+1}(\bar{\omega}))) \setminus V(\Phi_{i+1}(\bar{\omega}'))$  is a singleton color class in  $\bar{\omega}_k$ .

In particular, these conditions imply that

$$(2.10) \quad \phi(\bar{\omega}_k) = \sum_{v_i \in O_k} \phi_i(\bar{\omega}) + \sum_{v_i \in I_k} \phi_i(\bar{\omega}') + E_k = \sum_{v_i \in I_k \cup O_k} \phi_i(\bar{\omega}) + \sum_{v_i \in I_k} \ell(v_i) + E_k,$$

for

$$|E_k| \leq 2t.$$

In particular, we have with  $r = \phi(\bar{\omega}') - \phi(\bar{\omega})$  in the statement of Lemma 2.1

$$(2.11) \quad \phi(\bar{\omega}_k) \leq \phi(\bar{\omega}) + \frac{|I_k|}{m} (\phi(\bar{\omega}') - \phi(\bar{\omega})) + 4t \leq \max(\phi(\bar{\omega}), \phi(\bar{\omega}')) + 4t \leq q - 6t.$$



**Phase  $k$**  Inductively, having completed Phase  $k - 1$  (or, for  $k = 1$ , beginning from  $\bar{\omega}_0 = \bar{\omega}$ , we proceed as follows.

Recall the  $X_i$  from (2.9) which we use to define our phases, and the associated sets  $I_i, O_i, B_i$ .  $X_k$  differs from  $X_{k-1}$  in a single vertex, in the sense that there is  $v_{i_k} \in B_{k-1}$  and an  $i'_k = i_k \pm 1$ , so that either  $i'_k \in O_{k-1}$  and

$$(2.12) \quad B_k = (B_k \setminus \{v_{i_k}\}) \cup \{v_{i'_k}\} \quad I_k = I_{k-1} \cup \{v_{i_k}\} \quad O_k = O_{k-1} \setminus \{v_{i'_k}\}$$

or else  $i'_k \in I_{k-1}$  and

$$(2.13) \quad B_k = (B_k \setminus \{v_{i_k}\}) \cup \{v_{i'_k}\} \quad I_k = I_{k-1} \setminus \{v_{i_k}\} \quad O_k = O_{k-1} \cup \{v_{i'_k}\}.$$

Consider the first case. Here, we proceed as follows.

- (i) By (b),  $\bar{\omega}_{k-1}$  agrees with  $\bar{\omega}$  on  $\Phi_{i'_k}(\bar{\omega})$ . One by one, for each color class in  $\Phi_{i'_k}(\bar{\omega})$ , color each of the vertices of the color class with a distinct unused color (this can be done recursively, beginning with the leaves of a spanning tree of the color class, as to not disconnect it) so that all vertices in these color classes become singleton color classes. This increases the number of colors used by less than  $2t$ , so we will never use more than  $q$  colors. Call the result  $\bar{\omega}'_{k-1}$ .
- (ii) By (c) and (2.8), after completing Step (i), every vertex in  $V(\Phi_{i_k}(\bar{\omega}'))$  is a singleton color class in  $\bar{\omega}'_{k-1}$ . Now, one by one, we merge these singletons (by assigning them identical colors) so that the resulting coloring agrees with  $\bar{\omega}'$  on all color classes in  $\Phi_{i_k}(\bar{\omega})$ . This step only decreases the number of colors used, so we will not exceed  $q$  colors.

The resulting coloring is  $\bar{\omega}_k$ , and by construction satisfies (a), (b), and (c). In the second case, the analogous steps are:

- (i) By (a),  $\bar{\omega}_{k-1}$  agrees with  $\bar{\omega}$  on  $\Phi_{i'_k}(\bar{\omega}')$ . One by one, for each color class in  $\Phi_{i'_k}(\bar{\omega})$ , color each of the vertices of the color class with a distinct unused color, increasing the number of colors used by less than  $2t$ . The result  $\bar{\omega}'_{k-1}$ .
- (ii) By (c), after completing Step (i), every vertex in  $V(\Phi_{i_k}(\bar{\omega}'))$  is a singleton color class in  $\bar{\omega}'_{k-1}$ . Now, one by one, we merge these singletons (by assigning them identical colors) so that the resulting coloring agrees with  $\bar{\omega}'$  on all color classes in  $\Phi_{i_k}(\bar{\omega})$ .

The steps between phases never increase the number of non-null color classes by more than  $2t$ , and at the boundaries  $\bar{\omega}_k$  we have  $\phi(\bar{\omega}_k) \leq q - 6t$  as given in (2.11), which followed from our use of Lemma 2.1.

**2.6 Bounding the congestion** In the previous section we defined paths  $\Gamma_{\omega, \omega'}$  between any states  $\omega, \omega' \in \Omega$ , with the length  $|\Gamma_{\omega, \omega'}|$  satisfying

$$|\Gamma_{\omega, \omega'}| \leq 2tm^2 + 20t \leq 2n^2,$$

as Phase 0 consumes at most  $20t$  steps, and each of the  $T = \binom{m}{2}$  Phases  $1, \dots, T$  consumes at most  $4t$  steps.

Now, letting  $E$  denote the graph of transitions of the Markov Chain  $\mathcal{M}$ , we aim to bound the congestion

$$(2.14) \quad B = \max_{e \in E} \frac{2|E|}{|\Omega|^2} \sum_{\substack{\omega, \omega' \in \Omega \\ \Gamma_{\omega, \omega'} \ni e}} |\Gamma_{\omega, \omega'}| \leq \frac{4n^2|E|}{|\Omega|^2} \cdot \max_{e \in E} \#\{\{\omega, \omega'\} \mid \Gamma_{\omega, \omega'} \ni e\}.$$

Since  $|E|/|\Omega|$  is polynomial in  $n$ , we are particularly interested in bounding the ratio

$$(2.15) \quad \frac{\max_{e \in E} \#\{\{\omega, \omega'\} \mid \Gamma_{\omega, \omega'} \ni e\}}{|\Omega|},$$

which we bound by the vertex-congestion

$$(2.16) \quad \frac{\max_{\omega_1 \in \Omega} \#\{\{\omega, \omega'\} \mid \Gamma_{\omega, \omega'} \text{ visits } \omega_1\}}{|\Omega|}.$$

To this end, fix an  $\omega_1 \in \Omega$ . By the definition of the canonical paths, if  $\omega_1$  is visited by  $\Gamma_{\omega, \omega'}$  then

- (A) There are  $\bar{\omega}$  and  $\bar{\omega}'$  which can be reached from  $\omega$  and  $\omega'$ , respectively, from a sequence of at most  $10t$  applications of the maps  $\theta_\kappa$  defined in Phase 0, such that  $\phi(\bar{\omega}), \phi(\bar{\omega}') \leq q - 10t$ , and such that:
- (B) There are sets  $I, O \subseteq C_m$  inducing connected subgraphs of  $C_m$  with  $|I| + |O| = m - 3$ , such  $I$  and  $O$  are separated by at least one vertex on each side, and such that  $\omega_1$  agrees with  $\bar{\omega}'$  on all color classes in  $\Phi_i(\bar{\omega}')$  for all  $v_i \in I$ , and with  $\bar{\omega}$  on all color classes in  $\Phi_i(\bar{\omega})$  for  $v_i \in O$ .  
(Note that we have  $|I| + |O| = m - 3$  instead of  $|I| + |O| = m - 2$  because  $\omega_1$  can come from the middle of a phase rather than an endpoint.)
- (C) From our application of Lemma 2.1 in our construction of the canonical paths, these sets satisfy

$$\sum_{v_i \in I} (\phi_i(\bar{\omega}') - \phi_i(\bar{\omega})) = \frac{|I|}{m} (\phi(\bar{\omega}') - \phi(\bar{\omega})) \pm E \quad \text{for } |E| \leq 4t,$$

and thus, by subtracting both sides from  $\phi(\bar{\omega}') - \phi(\bar{\omega})$ , also satisfy

$$\sum_{v_i \in O} (\phi_i(\bar{\omega}') - \phi_i(\bar{\omega})) = \frac{|O|}{m} (\phi(\bar{\omega}') - \phi(\bar{\omega})) \pm F \quad \text{for } |F| \leq 8t.$$

- (D) We thus have that

$$(2.17) \quad \sum_{v_i \in I} \phi_i(\omega_1) - \phi_i(\bar{\omega}) = \frac{|I|}{m} (\phi(\omega_1) - \phi(\bar{\omega})) \pm E,$$

$$(2.18) \quad \sum_{v_i \in O} \phi_i(\bar{\omega}') - \phi_i(\omega_1) = \frac{|O|}{m} (\phi(\bar{\omega}') - \phi(\omega_1)) \pm F.$$

Our immediate task is to bound the number of choices for  $\bar{\omega}$  and  $\bar{\omega}'$  given  $\omega_1$ . There are less than  $m^2$  choices for the sets  $I$  and  $O$  as in (B). Then, given  $I$  and  $O$ , all color classes of  $\bar{\omega}'$  in sets  $\Phi_i(\bar{\omega}')$  for  $v_i \in I$  appear as color classes of  $\omega_1$ ; all that is unknown is the color classes in  $\Phi_i(\bar{\omega}')$  for  $v_i \in C_m \setminus I$ . Similarly, all color classes of  $\bar{\omega}$  in sets  $\Phi_i(\bar{\omega})$  for  $v_i \in O$  appear as color classes of  $\omega_1$ ; all that is unknown is the color classes in  $\Phi_i(\bar{\omega})$  for  $v_i \in C_m \setminus O$ .

Thus we will bound the number of choices for  $\bar{\omega}, \bar{\omega}'$  by bounding the choices for the sets

$$(2.19) \quad \Phi_i(\bar{\omega}') \text{ for } v_i \in C_m \setminus I, \quad \Phi_i(\bar{\omega}) \text{ for } v_i \in C_m \setminus O.$$

We will do this by first bounding the choices for the slightly smaller collection of sets

$$(2.20) \quad \Phi_i(\bar{\omega}') \text{ for } v_i \in O, \quad \Phi_i(\bar{\omega}) \text{ for } v_i \in I.$$

By (2.8), these sets satisfy the property that no  $C \in \Phi_i(\bar{\omega})$  for  $v_i \in I$  can intersect any  $C' \in \Phi_j(\bar{\omega}')$  for  $v_j \in O$ , and by (2.7) that the only vertices not assigned a color class in either collection of sets belong to  $V_{b^+} \cup V_{b^+ + 1} \cup V_{b^-} \cup V_{b^- + 1}$  where  $\{b^+, b^-\} = C_m \setminus I \setminus O$ . Thus, after fixing a choice for the color classes in the sets in (2.20), the result can be extended to a partition of the whole graph  $G$  by letting the unassigned vertices of  $V_{b^+} \cup V_{b^+ + 1} \cup V_{b^-} \cup V_{b^- + 1}$  be singleton classes. Doing so actually gives a valid  $q$ -partition of  $G$ , since this requires at most  $4t$  singleton classes, and by (D), the total number of color classes in the sets in (2.20) must satisfy

$$(2.21) \quad \sum_{v_i \in I} \phi_i(\bar{\omega}) + \sum_{v_i \in O} \phi_i(\bar{\omega}') \leq \frac{|I|}{m} \phi(\bar{\omega}) + \frac{|O|}{m} \phi(\bar{\omega}') \leq \max(\phi(\bar{\omega}), \phi(\bar{\omega}')) \leq q - 10t.$$

This implies that the number of choices for the sets in (2.20) is at most  $|\Omega|$ .



Figure 3: Partitions of  $G = G_{3 \times \ell}$  (for  $\ell = 21$ ) into two connected color classes. There are many more colorings which agree on the boundary vertices of  $G$  with the first and third shown coloring than with the middle shown coloring, which forms a bottleneck in the chain. In our discussion of this example, blue is color 1, and red is color 2.

$\Delta^n$  gives a crude upper bound on the number of forests in a graph on  $n$  vertices with maximum degree  $\leq \Delta$ . In turn this lets us bound the number of choices for the four sets

$$(2.22) \quad \Phi_i(\bar{\omega}') \text{ for } v_i \in C_m \setminus I \setminus O, \quad \Phi_i(\bar{\omega}) \text{ for } v_i \in C_m \setminus O \setminus I.$$

by  $\Delta^{2t}$  each, for a total of at most  $\Delta^{8t}$  choices. Thus in total we've bounded the number of choices for  $\bar{\omega}$  and  $\bar{\omega}'$  by

$$m^2 \Delta^{8t} |\Omega|.$$

Finally, accounting for Phase 0, the bound in (2.6) implies that there at most

$$(3^{\Delta K} / \delta)^{20t} m^2 \Delta^{8t} |\Omega|$$

choices for the states  $\omega, \omega'$ .

### 3 Torpid mixing for $q = 2$

In this section we give an example of a graph  $G$  which is a subgraph of the grid with bounded bandwidth for which the Markov chain  $\mathcal{M}$  is ergodic, but for which the mixing time of  $\mathcal{M}$  is exponential in the number of vertices of  $G$ .

It is simpler to give an example of torpid mixing for the chain operating on colors rather than partitions, so we do this first. (In particular, for this example, we will give a bottleneck which separates the set of colorings into two classes, which are equivalent under interchange of colors; thus this is not a bottleneck in the partition chain). For this we let  $G = G_{3 \times n}$  be a  $3 \times n$  grid graph, which has bandwidth 3.

When drawn in the plane in the natural way, the outer face of the  $k \times \ell$  grid graph  $G_{k \times \ell}$  is a  $2k + 2\ell - 4$  cycle which we denote by  $B_{k \times \ell}$ . The connectivity of the color classes implies that the intersection of each color class  $\omega^{-1}(i)$  with  $B_{k \times \ell}$  is a path, which we denote by  $P_i = P_i(\omega)$ . In our examples there will be only polynomially many colorings for which one of these paths is empty, thus we focus on the colorings where both paths are nonempty.

We fix a cyclic orientation of the cycle  $B_{k \times \ell}$  (when referring to Figure 3, we will use counterclockwise) so that we can refer to the *first* and *last* vertex of each path  $P_i$  in a consistent way.

Focusing now on the case  $k = 3$ , let define  $S$  to be those states  $\omega \in \Omega_{3 \times \ell}$  for which coordinate of the first vertex of  $P_1(\omega)$  is less than the second coordinate of the last vertex of  $P_1(\omega)$ . A representative element of  $S$  is illustrated at left in Figure 3 (note that we show the subgraph of the dual of the grid corresponding to  $G$ , and color its faces).

By symmetry we have that  $\pi(S) = \frac{1}{2}(1 - \rho)$  where  $\rho$  is the probability that the second coordinates of the first and last vertex of  $P_i(\omega)$  are actually equal. In particular, we have that  $\pi(S) \approx \frac{1}{2}$ .

The vertex-boundary  $\partial(S)$  of  $S$  consists of states for which the second coordinates of the first and last vertex of  $P_1$  differ by at most 1 (or: where one of the paths is empty: a linear-sized set). For these states, the intersection of each color class with the 'middle set'  $\{2\} \times [\ell]$  must also be a path; see Figure 3, middle. There are only  $O(\ell^2)$  such states, while the total number of states in  $\Omega$  is greater than  $2^\ell$  (when the top and bottom rows are constant with different colors, each of the middle vertices can be colored arbitrarily). Thus  $\pi(\partial S) = O(\frac{\ell^2}{2^\ell})$ , and the chain has exponential mixing time. (It is not hard to extend this argument to show that  $G_{4 \times \ell}$  has exponential mixing time, although understanding the general  $G_{c \times \ell}$  for constant  $c$  already seems like a challenge.)

The bottleneck we have identified for the color chain separates the chain into two regions, which are inverted by changing the labels of the two color classes; thus it does not show a bottleneck for the partition chain.

To show a bottleneck for the partition chain, we define  $G_{3 \times \ell}^4$ , for  $\ell$  odd, to be the subgraph of the  $\ell \times \ell$  grid graph induced by the set of vertices at graph distance  $\leq 1$  from the middle horizontal or middle vertical. Some

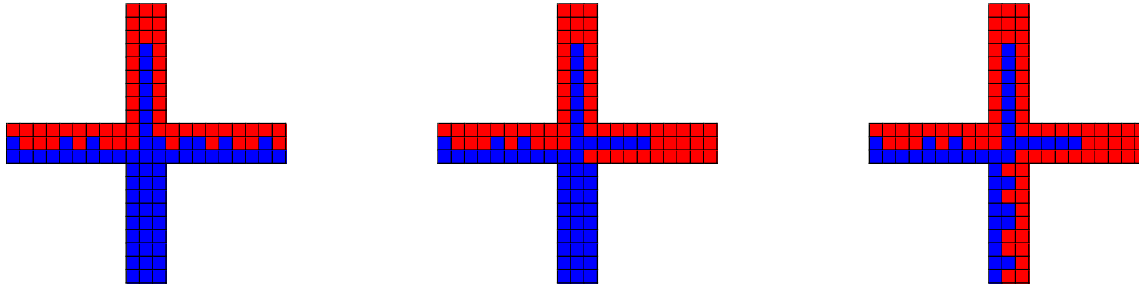


Figure 4: Partitions of  $G = G_{3 \times \ell}^4$  (for  $\ell = 21$ ) into two connected color classes. There are many more colorings which agree on the boundary vertices of  $G$  with the first and third shown coloring than with the middle shown coloring, which forms part of a bottleneck in the chain.

divisions of this graph (for  $\ell = 21$ ) into two color classes are shown in Figure 4. Note that the bandwidth of  $G_{3 \times \ell}^4$  is at most 12, which can be realized by a  $\sigma$  which orders the vertices in, say, increasing distance from the central vertex of  $G_{3 \times \ell}^4$  (the unique vertex of the graph which is fixed by all of its symmetries).

We again consider the outer face of the natural plane drawing of  $G$ , which is a  $4\ell - 2$  cycle, which we denote by  $B_{3 \times \ell}^4$ . This cycle has 4 vertices which have degree 4 in  $G$  (the corner vertices of the central  $3 \times 3$  square of the graph); we denote these vertices by  $b_1, b_2, b_3, b_4$ . They divide  $B_{3 \times \ell}^4$  into four paths;  $Q_1$  from  $b_4$  to  $b_1$ ,  $Q_2$  from  $b_1$  to  $b_2$ ,  $Q_3$  from  $b_2$  to  $b_3$ , and  $Q_4$  from  $b_3$  to  $b_4$ .

Now for a coloring  $\omega$ , we let  $\beta(\omega)$  be the set of  $i \in \{1, 2, 3, 4\}$ , for which the interior of the path  $Q_i$  includes vertices of both color. In particular, letting  $P_i = P_i(\omega)$  ( $i = 1, 2$ ) denote the intersection of color class  $i$  with  $B_{3 \times \ell}^4$  (which is necessarily a path),

$$\beta(\omega) = \{i \mid \overset{\circ}{Q}_i \cap P_1 \neq \emptyset \text{ AND } \overset{\circ}{Q}_i \cap P_2 \neq \emptyset\}.$$

For example, if in Figure 3,  $Q_1$  is the “North” boundary segment,  $Q_2$  the “East” boundary segment, and  $Q_3$  and  $Q_4$  “South” and “West”, respectively, then the coloring shown at left belongs has  $\beta(\omega) = \{2, 4\}$ , the middle coloring has  $\beta(\omega) = \{2\}$ , and the coloring at right has  $\beta(\omega) = \{2, 3\}$ .

LEMMA 3.1. *For all  $\omega$ , we have  $|\beta(\omega)| \leq 2$ .*

*Proof.* If  $i \in \beta(\omega)$ , then the interior of  $Q_i$  intersects both  $P_1$  and  $P_2$ . Since each of  $P_1$  and  $P_2$  has just two endpoints, and the interiors of the  $Q_i$  are pairwise disjoint, there can be at most two such  $i$ .  $\square$

The key point now is that for any distinct  $i, j \in \{1, 2, 3, 4\}$ , there are  $\Omega(2^\ell)$  colorings in  $\beta^{-1}(\{i, j\})$ ; indeed, for the case where half of  $P_1$  belongs to one color class and half to the other, we can freely choose the colors of the  $\approx \ell/2$  “middle vertices”—those that are adjacent to two vertices of this path. Thus the cases where this is true for both  $P_1$  and  $P_2$  already gives  $\Omega(2^\ell)$  members  $\omega \in \beta^{-1}(\{i, j\})$ . On the other hand, letting  $\beta_0$  be the set of colorings for which one of the paths  $P_i$  is empty, the set

$$\beta^{-1}(i) \cup \beta^{-1}(j) \cup \beta_0$$

forms a cut set separating  $\beta^{-1}(\{i, j\})$  from the rest of the Markov chain. But this set has size  $O(\text{poly}(\ell)2^{\ell/2})$ , showing that the mixing time of the chain is exponential in  $\ell$ .

## References

- [1] Georgios Amanatidis and Pieter Kleer. Rapid mixing of the switch Markov chain for 2-class joint degree matrices. *SIAM Journal on Discrete Mathematics*, 36(1):118–146, 2022.
- [2] Eric A Autry, Daniel Carter, Gregory Herschlag, Zach Hunter, and Jonathan C Mattingly. Multi-scale merge-split markov chain monte carlo for redistricting. *arXiv preprint arXiv:2008.08054*, 2020.
- [3] Eric A Autry, Daniel Carter, Gregory J Herschlag, Zach Hunter, and Jonathan C Mattingly. Metropolized multiscale forest recombination for redistricting. *Multiscale Modeling & Simulation*, 19(4):1885–1914, 2021.

- [4] Sachet Bangia, Christy Vaughn Graves, Gregory Herschlag, Han Sung Kang, Justin Luo, Jonathan C Mattingly, and Robert Ravier. Redistricting: Drawing the line. *arXiv preprint arXiv:1704.03360*, 2017.
- [5] Nayantara Bhatnagar, Dana Randall, Vijay V Vazirani, and Eric Vigoda. Random bichromatic matchings. *Algorithmica*, 50(4):418–445, 2008.
- [6] Julia Böttcher, Klaas P Pruessmann, Anusch Taraz, and Andreas Würfl. Bandwidth, expansion, treewidth, separators and universality for bounded-degree graphs. *European Journal of Combinatorics*, 31(5):1217–1227, 2010.
- [7] Daniel Carter, Zach Hunter, Gregory Herschlag, and Jonathan Mattingly. A merge-split proposal for reversible Monte Carlo Markov chain sampling of redistricting plans. *arXiv preprint arXiv:1911.01503*, 2019.
- [8] Maria Chikina, Alan Frieze, Jonathan C Mattingly, and Wesley Pegden. Separating effect from significance in markov chain tests. *Statistics and Public Policy*, 7(1):101–114, 2020.
- [9] Maria Chikina, Alan Frieze, and Wesley Pegden. Assessing significance in a markov chain without mixing. *Proceedings of the National Academy of Sciences*, 114(11):2860–2864, 2017.
- [10] Jarmila Chvátalová. Optimal labelling of a product of two paths. *Discrete Mathematics*, 11(3):249–253, 1975.
- [11] Daryl DeFord, Moon Duchin, and Justin Solomon. Recombination: A Family of Markov Chains for Redistricting. *Harvard Data Science Review*, 3(1), mar 31 2021. <https://hdr.mitpress.mit.edu/pub/lds8ptxu>.
- [12] Persi Diaconis and Daniel Stroock. Geometric bounds for eigenvalues of markov chains. *The annals of applied probability*, pages 36–61, 1991.
- [13] Martin Dyer, Mark Jerrum, and Haiko Müller. On the switch markov chain for perfect matchings. *Journal of the ACM (JACM)*, 64(2):1–33, 2017.
- [14] Martin E Dyer and Alan M Frieze. On the complexity of partitioning graphs into connected subgraphs. *Discrete Applied Mathematics*, 10(2):139–153, 1985.
- [15] Benjamin Fifield, Michael Higgins, Kosuke Imai, and Alexander Tarr. Automated redistricting simulation using markov chain monte carlo. *Journal of Computational and Graphical Statistics*, 29(4):715–728, 2020.
- [16] Jacob E Goodman, János Pach, and Chee K Yap. Mountain climbing, ladder moving, and the ring-width of a polygon. *The American Mathematical Monthly*, 96(6):494–510, 1989.
- [17] Gregory Herschlag, Han Sung Kang, Justin Luo, Christy Vaughn Graves, Sachet Bangia, Robert Ravier, and Jonathan C Mattingly. Quantifying gerrymandering in north carolina. *Statistics and Public Policy*, 7(1):30–38, 2020.
- [18] Tatsuo Homma. A theorem on continuous functions. In *Kodai Mathematical Seminar Reports*, volume 4, pages 13–16. Department of Mathematics, Tokyo Institute of Technology, 1952.
- [19] Mark Jerrum and Alistair Sinclair. Approximating the permanent. *SIAM journal on computing*, 18(6):1149–1178, 1989.
- [20] Mark Jerrum, Alistair Sinclair, and Eric Vigoda. A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. *Journal of the ACM (JACM)*, 51(4):671–697, 2004.
- [21] Claire Kenyon, Dana Randall, and Alistair Sinclair. Approximating the number of monomer-dimer coverings of a lattice. *Journal of Statistical Physics*, 83(3):637–659, 1996.
- [22] Lorenzo Najt, Daryl DeFord, and Justin Solomon. Complexity and geometry of sampling connected graph partitions. *arXiv preprint arXiv:1908.08881*, 2019.
- [23] Alistair Sinclair. Improved bounds for mixing rates of markov chains and multicommodity flow. *Combinatorics, probability and Computing*, 1(4):351–370, 1992.