Sampling Balanced Forests of Grids in Polynomial Time

Sarah Cannon[*], Wesley Pegden[†], and Jamie Tucker-Foltz[‡]

January 12, 2024

Abstract

We prove that a polynomial fraction of the set of k-component forests in the $m \times n$ grid graph have equal numbers of vertices in each component, for any constant k. This resolves a conjecture of Charikar, Liu, Liu, and Vuong, and establishes the first provably polynomial-time algorithm for (exactly or approximately) sampling balanced grid graph partitions according to the spanning tree distribution, which weights each k-partition according to the product, across its k pieces, of the number of spanning trees of each piece. Our result follows from a careful analysis of the probability a uniformly random spanning tree of the grid can be cut into balanced pieces.

Beyond grids, we show that for a broad family of lattice-like graphs, we achieve balance up to any multiplicative $(1 \pm \varepsilon)$ constant with constant probability, and up to an additive constant with polynomial probability. More generally, we show that, with constant probability, components derived from uniform spanning trees can approximate any given partition of a planar region specified by Jordan curves. These results imply polynomial time algorithms for sampling approximately balanced tree-weighted partitions for lattice-like graphs.

Our results have applications to understanding political districtings, where there is an underlying graph of indivisible geographic units that must be partitioned into k population-balanced connected subgraphs. In this setting, tree-weighted partitions have interesting geometric properties, and this has stimulated significant effort to develop methods to sample them.

1 Introduction

We consider the following question: given a graph G and an integer constant k, how can one randomly sample partitions of G into k connected pieces, each of equal size? We address this question in the context of the *spanning tree distribution* on partitions, under which the weight of a partition is proportional to the product of the numbers of spanning trees in each partition class. This distribution has been the subject of intense research in the context of mathematical approaches to the analysis of political districtings [8, 10, 17, 26, 13, 27, 29]. While efficient algorithms exist to sample from this distribution when there are no size constraints on the partition classes, there is no general recipe for converting such a sampler to an efficient sampler for the *balanced* spanning tree distribution, where we condition the spanning tree distribution on the event that the partition classes are equal in size. For the prototypical case of grid graphs, the following conjecture of Charikar, Liu, Liu, and Vuong asserted that rejection sampling would suffice:

Conjecture 1 (Charikar, Liu, Liu, and Vuong [10])**.** *For the $m \times n$ grid graph, the proportion of balanced k-partitions under the spanning tree distribution is at least $1/\mathrm{poly}(m,n)$, when $k = O(1)$.*

[*]Claremont McKenna College, scannon@cmc.edu

[†]Carnegie Mellon University, wes@math.cmu.edu

[‡]Harvard University, jtuckerfoltz@gmail.com

We confirm this conjecture as follows:

Theorem 2. *Let G be an $m \times n$ grid graph where $m \geq n$ and $k|m$. The probability that a k-partition from the spanning tree distribution is balanced is at least*

$$\frac{1}{\beta^{k^2} n^{5k-5} m^{3k-3}} \tag{1}$$

for a fixed constant β.

We note that the assumption that k divides the longer dimension is mostly for ease of exposition. With some more effort (and worse constant factors) one could require just $k|nm$, with essentially the same proof techniques. Theorem 2 will follow from Theorem 15, which will assert that, for a uniformly random spanning tree of the $m \times n$ grid graph ($m \geq n, k|m$), there is a $1/\text{poly}(mn)$ chance that there are $k-1$ edges whose removal divides the tree into equal-size components. Section 3 is devoted to proving Theorem 15, along with stronger bounds for the special case of $k = 2$.

The relative frequency of balanced partitions under the spanning tree distribution is particularly salient given the significant progress made in sampling algorithms for this distribution. For example, in 2020, leveraging recent breakthroughs in the polynomial-method approach to Markov chain mixing, Anari, Liu, Gharan, Vinzant, and Vuong gave an $O(N \log^2 N)$ approximate sampler based on the 'down-up' walk on the complement of k-component forests of an N-vertex graph [1]. In Section 2.5, we discuss the use of our results in the context of an additional rejection step for approximate samplers based on Markov chains, and also show how to exactly sample from the spanning tree distribution on balanced k-partitions in expected time $O(N^{3k-2} \log N)$ for a grid graph with N vertices. These are the first provably polynomial-time algorithms for (perfectly or approximately) sampling from the spanning tree distribution on balanced partitions.

In Section 4, we turn to analyze partitions in grid-like graphs under looser notions of balance. If we are interested in dividing a random spanning tree into components that are only approximately balanced (up to a $(1 \pm \varepsilon)$ multiplicative error), we show on lattice-like graphs (including grids) that this is possible with *constant* probability; Corollary 21 gives the precise statement for grids. In fact, we prove a more general result, which is that a uniform spanning tree on a sufficiently refined lattice-like graph will, with probability bounded below by a constant, be splittable into components that approximately match any partition of a region of the plane given by a collection of Jordan curves (Figure 1). In particular, suppose Λ_n is a sequence of infinite planar graphs of decreasing scale embedded in $\mathbb{R}^2$ which are lattice-like (see Definition 18). For example, our definition of "lattice-like" is broad enough to apply almost surely to the sequence where Λ_n is the Delauney triangulation of a Poisson point cloud in $\mathbb{R}^2$ of rate n. If D is a fixed plane graph, and Ω_{D,Λ_n} denotes a region of Λ_n whose boundary approximates the boundary of the outer face of D, we have that:

Theorem 3 (Informal version of Theorem 20)**.** *Given any plane graph D with $k+1$ faces, let $\phi_1, \ldots, \phi_k \subseteq \mathbb{R}^2$ denote its inner faces. For any $\varepsilon > 0$, as $n \to \infty$, there is a constant lower bound, depending only on the plane graph and ε, on the probability that a random spanning tree T of Ω_{D,Λ_n} contains $k-1$ edges whose removal disconnects T into components $C_1, \ldots, C_k$, where each C_i is at Hausdorff distance $< \varepsilon$ from a corresponding face ϕ_i of D.*

If one is interested in the even stronger condition of *additive* approximate balance, we show one can achieve this in lattice-like graphs as well, with an additional assumption regarding the uniformity of the density of vertices. This stronger assumption would fail, for example, for the previously mentioned case of Delauney triangulations of random point sets, but would still satisfied for any finite-degree, doubly-periodic connected plane graph.

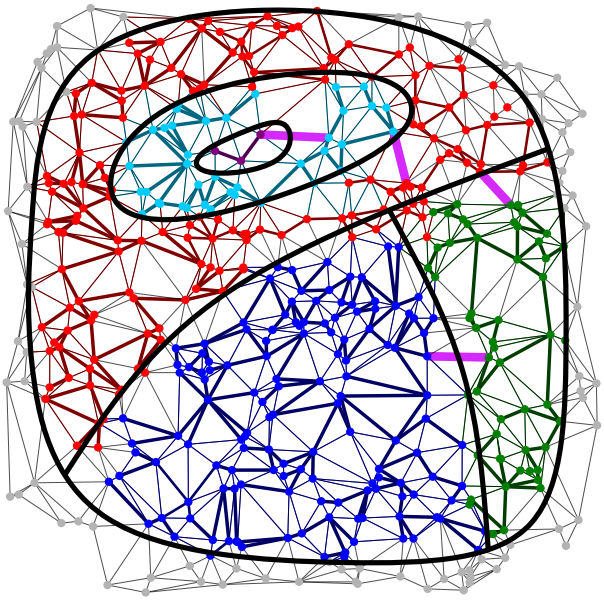


Figure 1: A partition of a region of a lattice-like graph approximating a division of the plane given by Jordan curves, and induced by the components remaining after deleting the four bright purple edges from a spanning tree of the region. Theorem 20 shows that given a division of the plane by curves, a random spanning tree of a sufficiently refined lattice-like graph can, with probability bounded below by a constant, be cut into components inducing a partition whose classes each has small Hausdorff distance from the corresponding face of the drawing.

Theorem 4 (Informal version of Theorem 22)**.** *If Λ_n is a sequence of infinite planar graphs of decreasing scale embedded in $\mathbb{R}^2$ which are* uniformly *lattice-like (see Definition 19) and Ω_{D,Λ_n} denotes a region of Λ_n whose boundary approximates the unit square, then there exists a constant $A > 0$ such that as $n \to \infty$, the probability a random spanning tree T of Ω_{D,Λ_n} contains $k-1$ edges whose removal produces k pieces whose sizes are equal up to an additive A vertices in each part is at least* $1/\mathrm{poly}(n)$.

Balance up to an additive constant is the best one could hope for in the framework of Theorem 4; exact balance may not be possible because of the structure of Λ_n or the way it's trimmed to approximate the unit square. However, we expect that even exact balance would be possible with mild additional assumptions on the local behavior of random walks in the lattice (as well as the necessary divisibility conditions).

Combining these results with known algorithms and rejection sampling gives corresponding polynomial time sampling algorithms in all of these settings.

Finally, in Section 5, we empirically evaluate the probabilities that random spanning trees of 10×10, 50×50, and 100×100 grids can be split at various locations into pieces of various sizes. These experiments visually confirm the analytical results in this paper.

1.1 Random sampling of political districting plans

In the context of the analysis of districting plans, sampling algorithms enable the generation of large *ensembles* of plans, which are useful for several purposes (detecting outliers, understanding the impacts of rules, evaluating the stated intentions of map-drawers, and more). Ensemble analysis has

been used in many academic studies, including [2, 3, 5, 6, 7, 8, 12, 14, 16, 17, 19, 20, 22, 23, 26, 32], as well as in mathematicians' expert reports in court cases [4, 11, 18, 25].

Randomly sampling political districting plans is equivalent to a sampling problem for suitable partitions of a graph, with vertices representing small geographic regions such as precincts or census blocks and edges representing adjacencies. Because they represent physical geography, these graphs are typically planar or nearly planar. While they are not usually perfect grids (except at times in cities), there is general consensus that grids are the logical simplified setting to first consider. By going beyond grids to lattice-like graphs, we move to a much more expressive graph class that can describe significant additional real-world geography.

A districting plan with k districts is a partition of this graph into k pieces, which are generally required to be connected. Throughout, we will call a partition of a graph into k connected pieces a *k-partition*, and we will refer to the k partition classes of a partition as *districts*.

In the context of redistricting, there are other constraints on partitions one must consider, including those related to population and shape. Our interest in balanced partitions stems from common requirements that districts have equal or near-equal populations. While our first main result resolves a conjecture about exactly balanced partitions, in practice most processes for sampling political districting plans do not aim for exact population balance but instead aim to keep the population to within a tolerance of 1-2%. This naturally corresponds to the setting of Theorem 3, where district sizes may vary by a multiplicative $1 \pm \varepsilon$ factor. Related to district shape, the spanning-tree distribution we analyze is targeted by several sampling algorithms designed for redistricting analysis [8, 17, 26], and has been shown to strongly correlate with geometric properties intended to capture legal requirements for 'compactness' of districts [13, 27, 29].

Unlike Markov chains such as the up-down walk, which operating in a context without balance constraints, we know that the approaches cited above such as *recombination Markov chains* can have exponential mixing time for some special families of graphs (including carefully chosen subgraphs of the grid) [10]. Even on rectangular grids, recombination chains with strict balance constraints can fail to be ergodic if there are many small districts [30]. Positive mixing time results for any reasonable class of graphs are not available. However, by giving the first polynomial lower bounds in grid and grid-like graphs on the probability of finding edges that cut random spanning trees in balanced ways, our approach also addresses a crucial factor for Markov chains like those in [8, 17, 26] that aim to achieve balance by preserving it at every step, by only using such balanced cuts in transitions.

Other Markov chains employed in the redistricting context include Glauber dynamics for contiguous partitions, which exchange individual vertices between districts. Here, without any additional constraints or weighting, stationary distributions are uniform on partitions with connected districts. Mixing time can again be exponential for some classes of graphs [21]. In fact, even in the absence of balance, it is not known whether the Glauber dynamics has polynomial mixing time for partitions of grid graphs into k connected pieces, or indeed whether any polynomial time algorithm to uniformly sample partitions of grid graphs uniformly randomly into k connected pieces exists, even for $k = 2$.

1.2 Approach

Rather than working with the tree distribution on partitions, we work with the uniform distribution on spanning trees. As we prove in Lemma 8, if there is a polynomial lower bound on the probability a random spanning tree can be split into k equal-sized components, there is a polynomial lower bound on the probability a random tree-weighted forest with k components is balanced. The majority of our work therefore focuses on uniformly random spanning trees and the probability

they can be split into components with desired properties.

Spanning trees of planar graphs are in bijective correspondence with the spanning trees of their dual graphs: If T is a spanning tree of G, its dual spanning tree T^* contains all edges in G^* whose corresponding edges are missing from T. The first key idea behind our approach is to study the structure of T^* rather than T. If T is a spanning tree of G with dual tree T^*, then the k connected components of $T \setminus e_1, \ldots, e_{k-1}$ are bounded by k cycles in $T^* \cup e_1 \cup \cdots \cup e_{k-1}$. In particular, to show that components with certain sizes or structure can be created by removing edges in T, it suffices to show that suitable boundary cycles almost already exist in T^*.

The second key idea is to study the probability of such suitable near-cycles occurring in T^* by analyzing the steps of Wilson's algorithm on the dual graph. For an arbitrary root vertex, Wilson's algorithm builds a uniformly random spanning tree by running a series of loop-erased random walks from arbitrary starting points to the component containing the root [31]. By choosing the root to be the dual graph vertex corresponding to the exterior face and carefully choosing the starting points of each random walk, we are able to show the algorithm is sufficiently likely to produce paths in T^* that have the properties we desire.

For some of our results on general lattice sequences, we will use a particular implementation of Wilson's algorithm described in Section 4.4 in which, having completed one loop erased random walk, we (sometimes) choose the next starting point for a new loop-erased random walk as the exit vertex of simple random walk within the induced subgraph of the already-built tree itself. This allows us to analyze the progress of the algorithm in long phases that may include many separate loop-erased random walks, but for which these separate loop-erased random walks can all be seen as being generated using a single random walk on the graph.

2 Preliminaries

2.1 Notation

For a positive integer n, we denote $[n] := \{1, 2, \ldots, n\}$. Unless otherwise specified, all graphs we consider are undirected with no self-loops, but multiple edges may be allowed between any pair of vertices. The $m \times n$ grid graph is the graph with vertex set $[m] \times [n]$, with an edge between (i, j) and (i', j') whenever $|i' - i| + |j' - j| = 1$. We always draw grid graphs in a Cartesian coordinate system, with m being the horizontal dimension and n being the vertical dimension. We denote by $\mathbb{Z}^2$ the infinite grid graph, where the vertex set is $\mathbb{Z} \times \mathbb{Z}$ and the edge relation is the same as in finite grids.

A forest is a graph with no cycles, and a tree is a connected forest. A *k-forest* is a forest with k connected components. A forest is *balanced* if every connected component has exactly the same number of vertices. If T is a tree and $S \subseteq E(T)$, we define $T \setminus S$ to be the forest F with vertex set $V(F) := V(T)$ and edge set $E(F) := E(T) \setminus S$. Thus, a tree T is *k-splittable* if there is some set $S \subseteq E(T)$ of size $k - 1$ such that $T \setminus S$ is a balanced k-forest.

For a graph H, we let $\mathrm{sp}(H)$ denote the number of spanning trees of H. For a k-partition P of G with districts P_1, …, P_k, we denote by π_{sp} the spanning tree distribution, given by

$$\pi_{\mathrm{sp}}(P) = \frac{\prod_{i=1}^{k} \mathrm{sp}(P_i)}{Z},$$

where Z is the normalizing constant, also called the partition function, given by

$$Z = \sum_{k\text{-partitions } P} \prod_{i=1}^{k} \mathrm{sp}(P_i).$$

Note that the uniform distribution over k-forests of G is equivalent to the spanning tree distribution over k-partitions of G when a forest is identified with its connected components.

2.2 Duality

Let G be a connected, planar graph, and fix an embedding of G in the plane with no edges crossing. The *dual graph* of G (with respect to the embedding) is the graph G^* whose vertices are faces of G, with an edge between two faces $a^* \in V(G^*)$ and $b^* \in V(G^*)$ whenever the two faces share a common boundary edge. Note that we count the outer face of G as a vertex of G^* as well.

For any edge $e \in E(G)$, let $e^* \in E(G^*)$ be the edge between the faces it bounds. For any set of edges $S \subseteq E(G)$, we analogously define $S^* := \{e^* \mid e \in S\} \subseteq E(G^*)$. The following lemma is a standard result.

Lemma 5. *Assume that G is connected and embedded in the plane such that no edge of G has the same face on both sides. Then $e \mapsto e^*$ is a bijection between edges of G and edges of G^*, and $T \mapsto T^* := (V(G^*), E(G^*) \setminus E(T)^*)$ is a bijection between spanning trees of G and spanning trees of G^*.*

Note that T^* does not contain the edges e^* for each $e \in T$, but rather those edges that are *not* in this set. The hypotheses of Lemma 5 hold for all $m \times n$ grid graphs with $m, n > 1$.

2.3 Wilson's algorithm

Wilson's algorithm [31] is important for us not just because it samples uniformly random trees efficiently, thus serving as a key subroutine in our perfect sampling algorithm (See Section 2.5.1), but also because our proofs rely on running Wilson's algorithm in a specific way.

For an input graph G, the steps of Wilson's algorithm are as follows:

1. Set $T \leftarrow \{r\}$ for an arbitrary "root" vertex $r \in V(G)$
2. While T does not connect all vertices of G:
   (a) Do a loop-erased random walk[1] starting at an arbitrary vertex $v \notin T$ until it reaches a vertex of T
   (b) Add all vertices and edges along this loop-erased random walk to T
3. Return T

Importantly, it does not matter which vertex is initially chosen as the root, and in each iteration of the while loop, it does not matter at which vertex not in T the next loop-erased random walk begins. Regardless of what arbitrary choices are made at these steps, one can prove the end result is a perfectly uniformly random spanning tree of G. We use this crucial fact in our proofs, analyzing the process of Wilson's algorithm (in the dual graph G^*) from carefully-chosen starting vertices.

Recall that the hitting time $\tau_u(v)$ of u from v is the expected time before a simple random walk reaches v from u, and the commute time between u and v is $\tau_u(v) + \tau_v(u)$. A *π-random* vertex of G is a vertex chosen according to the stationary distribution of the simple random walk on G, $\pi(v) = deg(v)/2m$. Wilson characterizes the expected running time of his algorithm (measured by the number of times we need to find a random neighbor of a vertex) in terms of the commute time as follows:

[1] That is, every time the random walk revisits a node u, erase the cycle and resume the random walk from u.

Proposition 6 (Wilson)**.** *The expected number of times we generate a random neighbor for a vertex in the course of running Wilson's algorithm on a graph G with root r is precisely the expected commute time between r and a π-random vertex v.* □

For general graphs with N vertices and M edges, it is well-known that the hitting time and thus the commute time between any pair of vertices is at most $O(NM)$ [24]; this implies that Wilson's algorithm runs in time $O(N^2)$ for any planar graph on N vertices. However, this can be improved for grid graphs by considering the dual graph and a carefully-chosen root:

Proposition 7. *Wilson's algorithm runs in expected time $O(N \log N)$ on the dual of any grid graph on N vertices, when the root is chosen to be the dual vertex corresponding to the outer face of the grid graph.*

This is easily proved using the characterization of the commute time in terms of effective resistance; we include a proof in the appendix.

2.4 Splittability and the spanning tree distribution

Here we explicitly connect the uniform distribution over spanning trees of a graph G with the uniform distribution over k-forests of G. This enables us to analyze the likelihood of obtaining a balanced partition when sampling from the spanning tree distribution over forests, as the up-down walk of [10] (approximately) does; see Section 2.5.2.

Lemma 8. *If the probability a uniformly random spanning tree of G with N vertices and M edges is k-splittable is at least α, then the probability a uniformly random k-forest of P is balanced is at least*

$$\frac{\alpha}{N^{k-1}(M-N+1)^{k-1}}.$$

The (short) proof of this lemma can be found in the appendix. We now use it to prove Theorem 2

Proof that Theorem 15 implies Theorem 2. Sampling a k-partition P of G according to the spanning tree distribution is the same as sampling a k-forest F of G uniformly at random and then considering its connected components. By Theorem 15, the probability that a uniformly random spanning tree T of G is k-splittable is a least $\frac{1}{\beta^{k^2} n^{3k-3} m^{k-1}}$ for some fixed constant β. By Lemma 8, as G has nm vertices and strictly less than $2nm$ edges, the probability a uniformly random k-forest is balanced is therefore at most $\frac{1}{\beta^{k^2} n^{5k-5} m^{3k-3}}$. □

There is hope this bound could be improved by studying random forests directly, rather than studying spanning trees and then considering cutting them to obtain forests.

2.5 Algorithms for sampling balanced tree-weighted partitions

Our theorems imply that known approaches sampling (not necessarily balanced) k-partitions according to the spanning tree distribution in polynomial time can be combined with a rejection sampling step to obtain an expected polynomial time algorithm for sampling balanced k-partitions. Here we present two methods by which this could be done, for the case of sampling exactly balanced k-partitions.

2.5.1 Perfectly sampling balanced k-partitions with Wilson's algorithm

Wilson's algorithm generates a perfectly uniform random spanning tree of a graph G. We can use it to randomly sample a balanced k-partition as follows.

1. Uniformly sample a random spanning tree T of G using Wilson's algorithm.
2. Check if T has $k-1$ edges whose removal disconnects T into k components of equal size. If no, reject and return to step 1.
3. If yes, create a k-partition P of G comprised of the connected components when these $k-1$ edges are removed from T.
4. Create a graph G/P which contracts each district of P into a single point and retains all edges between components with the appropriate multiplicity.
5. Compute the number s of spanning trees of G/P.
6. Return P with probability $1/s$. With the remaining probability $(s-1)/s$ reject and return to step 1.

Theorem 9. *For N-vertex grid graphs, this algorithm produces a balanced k-partition drawn perfectly from the spanning tree distribution in expected running time $O(N^{3k-2}\log N)$.*

See the appendix for a proof of this theorem. Briefly, the expected run time bounds are because it takes expected time $O(N\log N)$ steps to sample a random spanning tree and check if it is k-splittable, $O(N^{2k-2})$ attempts in expectation to see a k-splittable tree, and $O(N^{k-1})$ attempts to be successful in the final rejection of Step 5, by Theorem 15.

2.5.2 Approximately sampling balanced k-partitions with the up-down walk

An alternate method using the up-down Markov chain described in Charikar et al. can produce an approximately uniformly random k-forest [1, 10]. We briefly motivate and describe this approach here.

On any graph G, the spanning forests with at least k components form a matroid whose bases are exactly the k-component spanning forests of G. The well-known down-up chain on bases of a matroid mixes in time $O(r(\log r + \log\log n))$ when bases have r elements and the matroid has n total elements [15]; when run for longer than its mixing time, this chain produces an approximately uniformly random basis. For k-component forests, this down-up chain randomly removes an edge of the forest (to produce $k+1$ components), and then randomly adds back in an edge connecting two different components. It's mixing time is $O((N-k)(\log(N-k)+\log\log M))$ for graphs with N vertices and M edges; for constant k, this becomes $O(N\log N)$. However, naively implementing one down-up step requires $O(M)$ time, making the overall time for this chain to produce an approximate sample $O(NM\log N)$.

This was improved in [1] by considering the up-down walk instead. This walk which randomly adds an edge to the forest. If adding this edge creates a cycle, a random edge of the cycle is removed. If adding this edge did not create a cycle (e.g. the edge connected two components of the forest) then a random edge of the forest is removed. This chain has mixing time $O(M\log M)$ for graphs with M edges, as the up-down walk can also be viewed as the down-up walk on the dual matroid whose bases are the complements of k-component forests. Using a link-cut tree data structure, each up-down step can be implemented in amortized quasi-constant time, resulting in an overall

runtime of $O(M \log^2 M)$ to produce one random sample. For planar graphs where $M = O(N)$, this mixing time is $O(N \log^2 N)$. It is this up-down chain, rather than the usual down-up chain, that we will use.

The following is our algorithm for approximately randomly sampling a balanced k-forest of a N-vertex graph.

1. Run the up-down Markov chain on k-forests for some fixed amount of time longer than its mixing time.

2. If the current state of the chain is a balanced k-forest, return the partition P consisting of the connected components of the forest. Else, return to step 1 and repeat.

Theorem 10. *For N-vertex grid graphs, this algorithm produces a balanced k-partition drawn approximately from the spanning tree distribution in expected running time $O(N^{4k-3} \log^2 N)$.*

See the appendix for a proof of this theorem; briefly, the expected running time is because it takes $O(N \log^2 N)$ steps to approximately sample a random k-forest and $O(N^{4k-4})$ attempts in expectation to see a balanced one, by Theorem 2. Note that, with the relatively crude estimates we employ to deduce Theorem 2 from Theorem 15 (in Section 2.4), the runtime we prove for this approximate sampling approach is actually worse than for the exact sampler above. There is little reason to believe this to be the truth, however.

3 Exact balance on grid graphs

3.1 Exactly balanced bipartitions

In this section we prove the following:

Theorem 11. *Let G be a grid graph with N vertices, where N is even. The probability that a uniformly random spanning tree T of G is 2-splittable is at least $1/N^2$.*

In fact, we prove a stronger result, namely that specific edges near the center of the grid have a decent probability of being the edge that splits the tree. Formally, If G is an $m \times n$ grid graph, we define a *horizontal central edge* of G to be an edge of the form $\{(i, j), (i, j+1)\}$ that is as close to the center of G as possible. Note that there may be 1, 2, or 4 horizontal central edges depending on the parities of m and n.

Lemma 12. *Let G be an $m \times n$ grid graph where $m \geq n$, and mn is even, and let $e \in E(G)$ be any horizontal central edge of G. Then the probability that a uniformly random spanning tree T of G contains e, and $T \setminus \{e\}$ is a balanced 2-forest, is at least*

$$\begin{cases} \frac{1}{mn^3} & \text{if } m \text{ is even} \\ \frac{1}{4mn^3} & \text{if } m \text{ is odd} \end{cases}.$$

To prove Lemma 12, we will require two further lemmas.

Lemma 13. *Let G be the $m \times n$ grid graph induced by the subset $[m] \times [n]$ of the grid $\mathbb{Z}^2$, and $(i_0, j_0) \in V(G)$. The probability that a random walk from (i_0, j_0) in $\mathbb{Z}^2$ exits G for the first time to a vertex (i', j') with $j' > 0$ is at least $\frac{j_0}{n+1}$.*

Proof. Let $\bar{G}$ be the subset of of $\mathbb{Z}^2$ induced by vertices (x, y) with $0 \leq i \leq m+1$, $0 \leq j \leq n+1$. Consider the experiment in which we conduct a random walk from $(i_0, j_0) \in \bar{G}$, and then flip a biased coin whose heads probability is $\frac{j'}{n+1}$, where (i', j') is the first vertex the walk visits in $\bar{G} \setminus G$.

When $(i_0, j_0) \in G$, the event that the coin is a heads is contained in the event that the random walk first exits G to a vertex (i', j') with $j' > 0$. And for this experiment, we have that the probability of heads is exactly $\frac{j_0}{n+1}$, since this holds on the boundary $\bar{G} \setminus G$, and since this linear function is the unique harmonic extension of this boundary condition; that is, the unique function such that the value at each vertex is the average of the values at the neighbors, as is the case for the probability of heads in the experiment defined above. □

Lemma 14. *Let X be a discrete probability distribution supported on a set of size k. Then*

$$\Pr_{x_1, x_2 \sim X \times X}(x_1 = x_2) \geq \frac{1}{k}.$$

Proof. Suppose the probabilities of each element in the support of X are $p_1, p_2, \ldots, p_k$, where these values sum to one. Then the probability that two independent samples are the same is given by $\sum_{i=1}^{k} p_i^2$. Let $\mathbf{p} = (p_1, p_2, \ldots, p_k)$ and let $\mathbf{v} = \left(\frac{1}{k}, \frac{1}{k}, \ldots, \frac{1}{k}\right)$ be length k vectors. We see $||\mathbf{p}||^2 = \sum_{i=1}^{k} p_i^2$, $||\mathbf{v}||^2 = 1/k$, and $|\langle \mathbf{p}, \mathbf{v} \rangle|^2 = 1/k^2$, so the lemma follows immediately from the Cauchy-Schwarz inequality. □

Proof of Lemma 12. Assume $n > 1$ (otherwise there is nothing to show; the probability is one). Then note that Lemma 5 applies to G. We first consider the case where n is odd (so m must be even). In this case there is a unique horizontal central edge e, connecting the vertices $(m/2, (n+1)/2)$ and $(m/2+1, (n+1)/2)$. Let G^* be the dual graph of G in the plane, and denote the outer face by $r^* \in V(G^*)$. Let $a^* \in V(G^*)$ be the face above e and let $b^* \in V(G^*)$ be the face below e, as in Figure 2.

Consider the following algorithm for generating a uniformly random spanning tree T of G. Run Wilson's algorithm on G^* with r^* as the root, starting the first loop-erased random walk from a^* and the second random walk from b^* (if it is not already added to the tree in the first random walk). The remaining random walks in Wilson's algorithm can be executed from arbitrary starting points. This gives us spanning tree T^* of G^*. We then output the primal tree T whose dual is T^*. Since Wilson's algorithm gives a uniformly random sample from the set of spanning trees of G^*, and those dual trees are in bijection with the primal spanning trees of G (Lemma 5), this algorithm gives us a uniformly random sample from the set of spanning trees of G.

We apply Lemma 13 to the $(m-1) \times \frac{n-1}{2}$ dual sub-grid outlined in the top (red) rectangle in Figure 2, with initial vertex a^*. Note that $j_0 = 1$ because the coordinate system is shifted so that a^* is in the bottom row. Lemma 13 says that a random walk from a^* will first exit the sub-grid above, to the left, or to the right (just not below) with probability at least

$$\frac{1}{\frac{n-1}{2} + 1} > \frac{1}{(n-1)+1} = \frac{1}{n}.$$

This clearly applies to our loop-erased random walk as well: The probability that the first walk in Wilson's algorithm, which starts from a^*, makes it to the outer face r^* without ever entering the bottom half of the grid is at least $\frac{1}{n}$. Assuming this happens, we may then apply the same argument to the bottom (blue) rectangle, for the next random walk starting from b^*. By independence, with probability at least $\frac{1}{n^2}$, both paths will have made it to r^* without crossing the horizontal midline.

Assume that this happens, as it does in Figure 2. Let P_a^* be the path from a^* to the boundary, and let P_b^* be the path from b^* to the boundary. Since both P_a^* and P_b^* will be included in T^*,

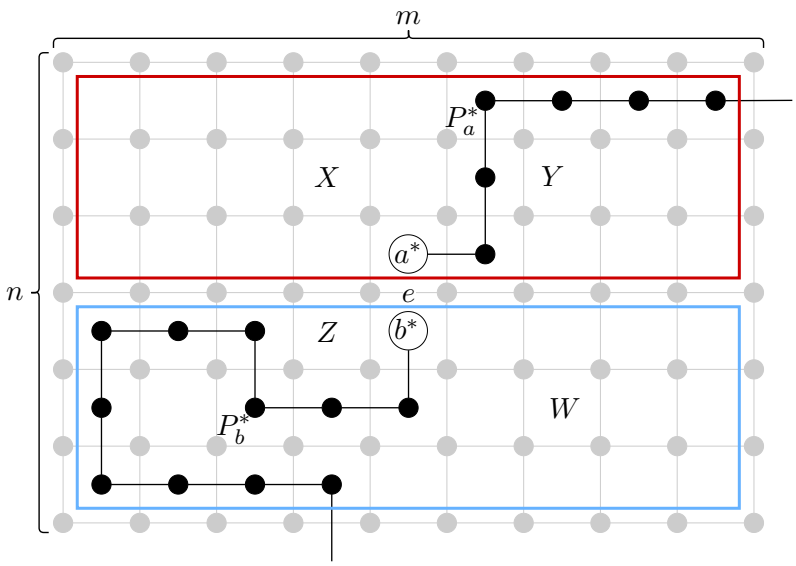


Figure 2: A possible run of the dual graph spanning tree sampling algorithm in the proof of Lemma 12 when m is odd. In this example, $m = 10$ and $n = 7$. The primal graph G is depicted in gray, and the first two random walks in the dual graph G^* are depicted in black.

we know that T cannot cross these paths. This means e must be included in T. Moreover, P_a^* and P_b^* completely determine the number of vertices on each side of e in T, as follows. Suppose there are X vertices in the top-half of the grid to the left of P_a^*, Y vertices in the top-half of the grid to the right of P_a^*, Z vertices in the bottom-half of the grid to the left of P_b^*, and W vertices in the bottom-half of the grid to the right of P_b^*. Then the subtree of T to the left of e will have $X + Z + \frac{m}{2}$ vertices, and the subtree to the right of e will have $Y + W + \frac{m}{2}$ vertices (the $\frac{m}{2}$ terms come from the vertices on the horizontal midline). Observe that the distribution, over the random path P_a^*, of the possible values of $X - Y$ is independent of and identical to the distribution, over the random path P_b^*, of the possible values of $W - Z$. Both distributions can take any integral value from $-\frac{m-1}{2}n$ to $\frac{m-1}{2}n$. Thus, applying Lemma 14, we know that, with probability at least

$$\frac{1}{(m\frac{n-1}{2}) - (-m\frac{n-1}{2}) + 1} = \frac{1}{mn - m + 1} > \frac{1}{mn},$$

we have $X - Y = W - Z$, which implies

$$X + Z + \frac{m}{2} = Y + W + \frac{m}{2},$$

i.e., the subtrees are balanced.

Thus, we have shown that the probability e is included in a uniformly random spanning tree T of G and splits it into a balanced 2-forest is at least

$$\frac{1}{n^2} \cdot \frac{1}{mn} = \frac{1}{mn^3}.$$

The remaining cases, where n is even, are almost the same. There are just a few minor additional assumptions we must impose about what happens to the random walks at the very beginning, as illustrated in Figure 3.

If m is even as well, there are two horizontal central edges bordering the unique central face of the grid. Without loss of generality, take e to be the top one, then define a^* and b^* as before. We

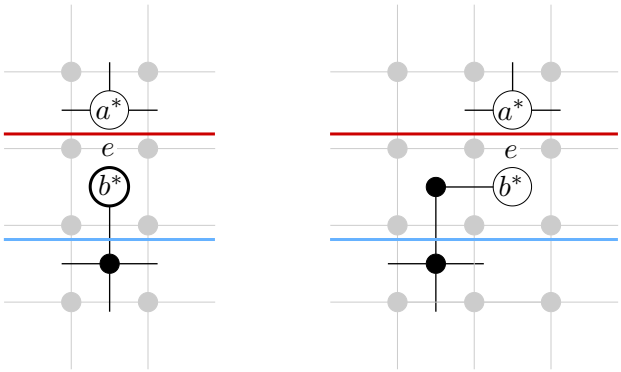


Figure 3: The cases in the proof of Lemma 12 when m is even, in which we must assume that the initial steps of the random walk from b^* takes a specific path into the blue rectangle, from which it never leaves until hitting the outer face.

suppose that the random walk from b^* first steps directly downward, as in Figure 3 (left). This happens with probability $\frac{1}{4}$. From there, by the same arguments as before, noting that the two subgrids are now each $m \times (n/2-1)$, the probabilities that the paths leave their respective red and blue rectangles at the boundary of the grid are both at least

$$\frac{1}{\left(\frac{n}{2}-1\right)+1} = \frac{2}{n}.$$

The probabilities that the number of vertices on each side are the same is at least

$$\frac{1}{(m\frac{n-2}{2}) - (-m\frac{n-2}{2}) + 1} = \frac{1}{mn - 2m + 1} > \frac{1}{mn}.$$

Thus, the probability that e splits a uniformly random spanning tree T into a balanced 2-forest is at least

$$\frac{1}{4} \cdot \left(\frac{2}{n}\right)^2 \cdot \frac{1}{mn} = \frac{1}{mn^3}.$$

Finally, consider the case where m is odd and n is even. Now there are four horizontal central edges, of which we pick the top-right one without loss of generality. With probability $\frac{1}{16}$, the random walk from b^* first steps to the left and then down into the blue rectangle, as in Figure 3 (middle). Now we can again apply the same arguments as above to the subgrids of dimensions $m \times (n/2-1)$ showing the probability the remaining paths leave their subgrids at the boundary of G are both at least $\frac{2}{m}$. While the random walk in the top grid no longer begins exactly in the center of the top grid (it can't, because this grid is now of even width), the top and bottom grids are rotationally symmetric, with the top walk beginning just one unit left of center and the bottom walk beginning one unit right of center. As before, the distributions of difference of the number of vertices on each side of the path are identical and are supported on sets of size at most mn, so by Lemma 14, the probability these differences are identical is at least $\frac{1}{mn}$. Thus, the probability that e splits a uniformly random spanning tree T into a balanced 2-forest is at least

$$\frac{1}{16} \cdot \left(\frac{2}{n}\right)^2 \cdot \frac{1}{mn} = \frac{1}{4mn^3}.$$ □

We now use this lemma to prove Theorem 11.

Proof of Thoerem 11. Recall that $N = nm$ is the total number of vertices. Assuming $m \geq n$, we know that $mn^3 \leq m^2n^2 = N^2$. Thus, in the case where m is even, we simply choose one of the horizontal central edges, which, by Lemma 12, splits a random tree into a balanced 2-forest with probability at least $\frac{1}{mn^3} \geq \frac{1}{N^2}$. In the case where m is odd (and so n must be even) there are 4 horizontal central edges, each of which will split a random tree into a balanced 2-forest with probability at least $\frac{1}{4mn^3}$. Since these 4 events are mutually exclusive, one of these four will give a balanced split with probability at least $\frac{1}{mn^3} \geq \frac{1}{N^2}$. □

3.2 Exactly balanced k-partitions

In this section we prove the following:

Theorem 15. *For $m \geq n$, let G be an $m \times n$ grid graph, and let k be a positive integer dividing m. There exists a set $S \subseteq E(G)$ of size $k-1$ such that the probability a uniformly random spanning tree T of G contains each edge in S, and $T \setminus S$ is a balanced k-forest, at least*

$$\frac{1}{\beta^{k^2} n^{3k-3} m^{k-1}} \tag{2}$$

for a fixed constant β.

The proof proceeds along similar lines as the proof of Lemma 12. We require the following stronger lemmas about random walks on grids. This is similar to Lemma 13, except that now we are also not allowed to hit the left or right sides, which makes the proof significantly more involved. We begin by proving a result about square grids, then extend this to tall thin rectangles.

Lemma 16. *For any odd integer ℓ, let G be the $(2\ell+1) \times (2\ell+1)$ grid graph whose coordinates are given by $\{-\ell, \ldots, \ell\} \times \{0, \ldots 2\ell\}$. For a random walk in $\mathbb{Z}^2$ beginning at vertex $(0,0)$, the probability this random walk first exits G to a vertex with coordinate $(i, 2\ell+1)$ for some integer $-\ell \leq i \leq \ell$, is at least $1/B\ell$ for some constant B.*

Proof. Let $\mathcal{L}$ denote the event we are interested in, that a random walk in $\mathbb{Z}^2$ beginning at vertex $(0,0)$ first exits G to a vertex with coordinate $(i, 2\ell+1)$. Our proof giving a lower bound on $\Pr(\mathcal{L})$ will use the following three claims:

Claim 1. *A simple random walk on $\mathbb{Z}$ from 0 reaches vertex n within $2n^2$ steps with probability at least $\frac{1}{4}$.*

Claim 2. *For any j, conditioned on the event that a simple random walk on $\mathbb{Z}$ from 0 of fixed length $j \geq x$ ends at the vertex $x > 0$, the probability that the walk never revisits 0 is $\frac{x}{j}$.*

Claim 3. *There is a constant λ such that for any D, a simple random walk on $\mathbb{Z}$ from 0 of length $\leq D\ell^2$ has probability $\geq e^{-\lambda D}$ of never leaving the interval $[-\ell, \ell]$.*

We give elementary proofs of these claims below, but first let us use them to prove that for an absolute constant B,

$$\Pr(\mathcal{L}) \geq \frac{1}{B\ell}. \tag{3}$$

In particular, we will consider a walk of fixed length $40\ell^2$, and show that with probability at least $\frac{1}{B\ell}$, such a walk from $(0,0)$ will exit G from the top side and do so before exiting any other side.

The random walk of length $40\ell^2$ is associated to an i.i.d. sequence of length $40\ell^2$ over the four directions in the lattice, which we denote (in cyclic order) by N, E, S, W. We generate the random walk in two steps:

1. We generate a uniform binary sequence of length $40\ell^2$ over the symbols $\{H, V\}$.
2. We independently replace each symbol H with a symbol E or W with probability $\frac{1}{2}$, and each symbol V with a symbol N or S with probability $\frac{1}{2}$.

Observe first that, with high probability, there are between $18\ell^2$ and $22\ell^2$ symbols of each type H and V after the first step. In particular, we will simply condition on the event that the number of symbols H and V fall in this particular range. By associating steps S and W to -1 and N and E to $+1$, we can associate to the sequence of replacements for the H's a random walk W_H on $\mathbb{Z}$, and to the sequence of replacements for the V's a 1D random walk W_V on $\mathbb{Z}$. We need to put a lower bound on the probability that W_H never leaves the interval $[-\ell, \ell]$ AND W_V reaches $2\ell+1$ without visiting -1. Conditioned on the lengths of the walks W_V, W_H, these are independent events; thus it suffice to prove lower bounds on the probabilities of these two events conditioned on any fixed lengths for W_V and W_H between $18\ell^2$ and $22\ell^2$.

We first consider the case of W_V. We define the events $\mathcal{E}', \mathcal{E}_j$ by

$$\mathcal{E}' = \{W_V \text{ reaches } 2\ell+1 \text{ within } 9\ell^2 \text{ steps}\}$$
$$\mathcal{E}_j = \{j \text{ is the index of last visit of } W_V \text{ to } 2\ell+1 \text{ among first } 18\ell^2 \text{ steps}\}.$$

Observe that

$$\bigcup_{j \le 18\ell^2} \mathcal{E}_j = \mathcal{E}'. \tag{4}$$

Since $\ell \ge 1$ implies $2(2\ell+1)^2 \le 18\ell^2$, we have from Claim 1 that

$$\Pr(\mathcal{E}') \ge \tfrac{1}{4}. \tag{5}$$

Now fixing some $j \le 18\ell^2$ steps and conditioning on the event that W_V is at $2\ell+1$ on the jth step, we have from Claim 2 that

$$\begin{aligned}&\Pr\left(W_V \text{ never visits } -1 \text{ before step } j \mid W_V \text{ at } 2\ell+1 \text{ on step } j\right)\\ &\quad\ge \Pr\left(W_V \text{ never revisits 0 before step } j \mid W_V \text{ at } 2\ell+ \text{ on step } j\right) \ge \frac{2\ell+1}{18\ell^2} \ge \frac{1}{9\ell}.\end{aligned} \tag{6}$$

Now (4), (5), and (6) imply that

$$\begin{aligned}\Pr\left(W_V \text{ visits } 2\ell+1 \text{ before } -1\right) &= \Pr(\mathcal{E}')\Pr(W_V \text{ visits } 2\ell+1 \text{ before } -1 \mid \mathcal{E}')\\ &= \Pr(\mathcal{E}') \sum_{j \le 9\ell^2} \Pr(\mathcal{E}_j \mid \mathcal{E}') \Pr\left(W_V \text{ visits } 2\ell+1 \text{ before } -1 \mid \mathcal{E}', \mathcal{E}_j\right)\\ &\ge \Pr(\mathcal{E}') \sum_{j \le 9\ell^2} \Pr(\mathcal{E}_j \mid \mathcal{E}') \Pr\left(W_V \text{ avoids } -1 \text{ through step } j \mid \mathcal{E}', \mathcal{E}_j\right)\\ &\ge \frac{1}{4} \cdot \frac{1}{9\ell} = \frac{1}{36\ell}.\end{aligned}$$

In the final inequality, we have used the fact that for $j \le 18\ell^2$, when conditioning on both events $\mathcal{E}'$ and $\mathcal{E}_j$, the initial segment of W_V of length j is still a uniformly random walk among all walks of length j from 0 to $2\ell+1$, as replacing such an initial segment with any other does not change the outcome of the events $\mathcal{E}', \mathcal{E}_j$.

Finally, since W_H is of length $\leq 22\ell^2$, we have from Claim 3 (with $D = 22$) that with probability at least $\exp(-22\lambda)$, W_H never leaves the interval $[-\ell, \ell]$. As we have conditioned ahead of time on the lengths of these two walks, the events corresponding to the two probabilities are independent, and so we have

$$\Pr(W_V \text{ visits } 2\ell+1 \text{ before } -1 \text{ AND } W_H \text{ stays in } [-\ell,\ell]) \geq \frac{1}{B\ell},$$

for an absolute constant B. This completes the proof of (3); it only remains to prove the three claims.

Proof of Claim 1. Recall that simple random walk begun from s on $\{0, 1, \ldots, M\}$ reaches an endpoint of this interval in expected time $s(M-s)$ (see e.g., [24, Proposition 2.1]). Thus a simple random walk on $\mathbb{Z}$ from 0 reaches either $-n$ or n in expected time n^2. This means that a simple random walk from 0 of length $\geq 2n^2$ reaches $-n$ or n with probability at least $\frac{1}{2}$, as Markov's inequality implies that the probability that the hitting time exceeds $2n^2$ is $\leq \frac{n^2}{2n^2}$. By symmetry the probability it reaches n is thus at least $\frac{1}{4}$. This proves Claim 1.

Proof of Claim 2. This is a consequence of Bertrand's ballot theorem. If the walk of length $j \geq x$ ends at the vertex $x > 0$, it takes $R = \frac{j+x}{2}$ steps to the right and $L = \frac{j-x}{2}$ steps to the left. Bertrand's ballot theorem then implies that the probability that among any initial nonempty segment, there is never as many left steps as right steps, is precisely

$$\frac{R-L}{R+L} = \frac{x}{j}.$$

Proof of Claim 3. Choose $K = \frac{1}{20}\ell^2$. Consider K-step simple random walk on $\mathbb{Z}$ starting from 0. Observe that by Chernoff bounds, if S_K is the location after K steps,

$$\Pr(S_K \geq \ell/2) \leq e^{-(\ell/2)^2/2K} = e^{-5/2} < \frac{1}{10},$$

It follows that with probability $> \frac{9}{10}$, a simple random walk of length $K = \frac{1}{20}\ell^2$ from 0 in $\mathbb{Z}$ will not end at a vertex $S_K \geq \ell$. Recall the standard reflection principle for simple random walk:

$$\Pr(\max_{j \leq n} S_j \geq T) = \Pr(S_n \geq T) + \Pr(S_n \geq T+1). \tag{7}$$

It follows that with probability $\geq \frac{8}{10}$, the walk $S_0, \ldots, S_K$ will never exceed $\ell/2$. By symmetry, with probability $\geq \frac{6}{10} \geq \frac{1}{2}$, it will never leave the interval $\mathcal{I} = [-\ell/2, \ell/2]$. Thus, with probability at least $\frac{1}{4}$, the random walk will both not leave the interval $\mathcal{I}$ *and* end in the nonnegative (respectively, nonpositive) side of the interval. Repeated applications of this simple fact then give Claim 3 as follows:

We break the random walk $S_0, S_1, \ldots, S_{D\ell^2}$ into $D\ell^2/K = 20D$ walks, each of length K. From the observation of the previous paragraph, each such walk has probability at least $\frac{1}{4}$ of never deviating more than $\ell/2$ from its starting point *and* ending to the right or left of its starting point, according to whichever direction the origin 0 is from the starting point of the walk. In this way, with probability $\geq \left(\frac{1}{4}\right)^{20D}$, every walk begins and ends within $\ell/2$ of the origin, and never ventures more than $\ell/2$ from its starting point. In this way the whole concatenated walk remains within ℓ of the origin. This proves Claim 3 with $\lambda = 20 \ln 4$, and thus completes the proof of the Lemma. □

The following lemma extends these results beyond a square region to a narrow rectangle, provided its width is at least a constant fraction of its height.

Lemma 17. *Suppose there is a constant $\varepsilon > 0$ such that $m > \varepsilon n$. Let G be the $(m+1) \times (n+1)$ grid graph induced by the subset $\{0, \ldots, m\} \times \{0, \ldots, n\}$ of the grid $\mathbb{Z}^2$, where $m+1$ is odd, and let $(i_0, j_0) = (\frac{m}{2}, 0)$. The probability that a random walk from (i_0, j_0) in $\mathbb{Z}^2$ exits G for the first time to a vertex (i', j') with $j' = n+1$ is at least $\frac{1}{Ane^{A/\varepsilon}}$, for a fixed constant A.*

Proof. First, we consider the case where $m \geq n$. In this case, set $\ell = \lfloor (n+1)/2 \rfloor$, so that $2\ell + 1$ is equal to $n+1$ or $n+2$, whichever is odd. By Lemma 16, there is a constant B such that the probability the a random walk starting at $(m/2, 0)$ first exits the rectangle at some $(i', n+1)$ for some $i' \in [m/2 - \ell, m/2 + \ell]$ without ever leaving this interval of x-coordinates is at least $1/Bn$ for some constant B. When $m \geq n$, this interval is a subset of $[0, m]$, and so the probability a random walk from $(m/2)$ first exits G to some $(i', n+1)$ is at least $1/Bn$ as well; appropriate choice of A such that $Ae^A > B$ proves the lemma.

Now, suppose $m < n$. For $\ell = \left\lfloor \frac{m-2}{4} \right\rfloor$, define $\mathcal{S}_0$ to be the $(2\ell+1) \times (2\ell+1)$ square subset of G induced by the vertices (x, y) with $\left\lfloor \frac{m}{2} \right\rfloor - \ell \leq x \leq \left\lfloor \frac{m}{2} \right\rfloor + \ell$ and $0 \leq y < (2\ell+1)$. Given a random walk in $\mathbb{Z}^2$ beginning at $(\frac{m}{2}, 0)$, we define $\mathcal{L}_0$ to be the event that the random walk exits the square $\mathcal{S}_0$ for the first time along its top side. By Lemma 16, $\Pr(\mathcal{L}_0) \geq \frac{1}{B\ell} \geq \frac{1}{Bm}$ as $\ell \leq m$. We now show how one can extend this path from the top of $\mathcal{S}_0$ to the top boundary of G in a reasonably likely way.

For any vertex $v \in \mathbb{Z}^2$ and positive integer ℓ, for a random walk in $\mathbb{Z}^2$ starting from v, define events $\mathcal{L}_{v,\ell}^{\mathrm{NNW}}$ and $\mathcal{L}_{v,\ell}^{\mathrm{NNE}}$, as the events that random walk from v first exits the $(2\ell+1) \times (2\ell+1)$ square centered at v along the left half or right half, respectively, of the top side of the square. In the case where the random walk exits at the middle of the top side of the square, we say both events occur. In this way, by the symmetries of the square, we have that

$$\Pr(\mathcal{L}_{v,\ell}^{\mathrm{NNE}}) = \Pr(\mathcal{L}_{v,\ell}^{\mathrm{NNW}}) \geq \frac{1}{8}. \tag{8}$$

Now, given v and ℓ, we can define $\mathcal{L}_{v,\ell}^{N0}$ to be the event $\mathcal{L}_{v,\ell}^{NNE}$ whenever the horizontal coordinate of v is $< \frac{m}{2}$, and to be $\mathcal{L}_{v,\ell}^{NNW}$ whenever the horizontal coordinate of v is $\geq \frac{m}{2}$.

Let $v_0 = (\frac{m}{2}, 0)$, and let v_1 be the first vertex outside $\mathcal{S}_0$ encountered in a random walk in $\mathbb{Z}^2$ starting at v_0. For each $i \geq 1$, define $\mathcal{S}_i$ to be the $(2\ell+1) \times (2\ell+1)$ square centered at v_i, and let v_{i+1} be the first vertex outside $\mathcal{S}_i$ in a random walk beginning at v_i; see Figure 4. To prove the lemma, note that if $\mathcal{L}_0$ occurs and, for each subsequent v_i for $i = 1, \ldots, L$ where $L = \left\lceil \frac{2(n-2\ell)}{2\ell+2} \right\rceil$, event $\mathcal{L}_{v,\ell}^{N0}$ occurs, then the random walk from $(\frac{m}{2}, 0)$ first exits G to a vertex $(i', n+1)$: the initial event $\mathcal{L}_0$ ensures it never exits the bottom of the rectangle; the choice of $N0 = NNW$ or $N0 = NNE$ ensures it never exits the side of the rectangle; and L was chosen large enough to ensure exiting $\mathcal{S}_L$ along its top implies exiting G along its top. Therefore, to prove the lemma, it suffices to prove a lower bound, for any fixed sequence $v_1, \ldots, v_L$ on the product

$$\Pr(\mathcal{L}_0) \prod_{i=1}^{L} \Pr(\mathcal{L}_{v_i,\ell}^{N0} \mid v_i) = \Pr(\mathcal{L}_0) \prod_{i=1}^{L} \Pr(\mathcal{L}_{v_i,\ell}^{N0}) \geq \frac{1}{Bm8^L} \geq \frac{1}{Bn8^L}. \tag{9}$$

Note that the side length of the square is $2\ell + 1 \geq \frac{m}{2} - 1$, and so

$$L \leq \frac{2n}{2\ell+2} \leq \frac{2n}{m/2} = \frac{4n}{m} \leq \frac{4n}{\varepsilon n} = \frac{4}{\varepsilon}. \tag{10}$$

Combining (9) and (10), we conclude that the probability the random walk $S_0, S_1, S_2 \ldots$, exits at the very top is at least

$$\frac{1}{Bn8^L} \geq \frac{1}{Bn8^{4/\varepsilon}} \tag{11}$$

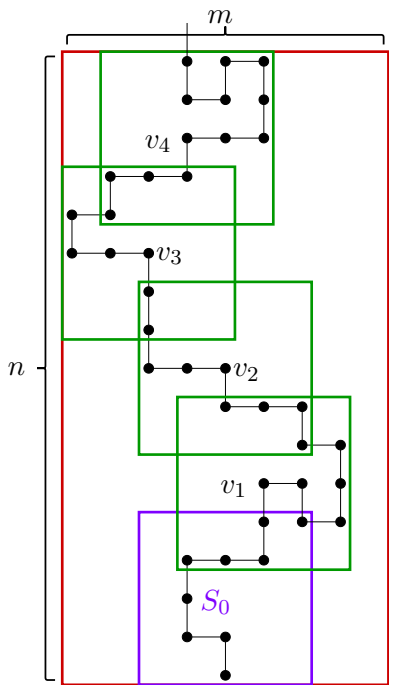


Figure 4: A random walk from the bottom of the red rectangle that first exits at the very top because each of the events $\mathcal{L}_0, \mathcal{L}^{N0}_{v_1,\ell}, \mathcal{L}^{N0}_{v_2,\ell}, \ldots, \mathcal{L}^{N0}_{v_L,\ell}$ occur. Here $m = 10$, $n = 16$, $\ell = 2$, and $L = 4$. The event $\mathcal{L}_0$ says that the walk from the bottom first exits the bottom purple outlined square to some vertex v_1 above the top. From there, each subsequent event $\mathcal{L}_{v_i,\ell}$ says that the walk exits the next green square along the top, on the side of the top boundary that is closer to the center.

For an appropriate choice of constant $A \geq \max\{B, 4\ln 8\}$, this is of the form $1/(Ane^{A/\varepsilon})$ claimed in the statement of the Lemma; note using the same constant A twice is just a simplification for convenient bookkeeping. □

We now use these lemmas to prove our main theorem.

Proof of Theorem 15. As in the proof of Lemma 12, we assume $n > 1$ so that Lemma 5 applies, and we first consider the case where n is odd. For each $1 \leq i \leq k-1$, we define $e_i := \{(\frac{m}{k}i, \frac{n+1}{2}), (\frac{m}{k}i + 1, \frac{n+1}{2})\}$. As shown in Figure 5, these edges lie on the horizontal midline of the grid and divide it vertically into k equal pieces. For each index i, let a_i^* and b_i^* be the faces respectively above and below e_i.

We generate a random spanning tree T of G by running Wilson's algorithm on the dual graph G^* with the outer face r^* as the root. We start the first $2(k-1)$ random walks from $a_1^*, b_1^*, a_2^*, b_2^*, \ldots, a_{k-1}^*, b_{k-1}^*$, assuming each of these vertices of G^* has not yet been added to the tree. The remaining starting points can be determined arbitrarily, and we output the tree T that is dual to the uniformly random spanning tree T^* of G^* sampled by Wilson's algorithm. We will show that $S := \{e_1, e_2, \ldots, e_k\}$ splits G into a balanced k-forest with sufficiently large probability.

We apply Lemma 17 to each of the red and blue rectangles depicted in Figure 5, which all have grid dimensions $(\frac{m}{k} - 1) \times \frac{n-1}{2}$ if $\frac{m}{k}$ is even and $\frac{m}{k} \times \frac{n-1}{2}$ if $\frac{m}{k}$ is odd. Independently, by applying Lemma 17 to each of these grids with $\varepsilon = 2/k$ and recalling $m \geq n$, each random walk makes it to

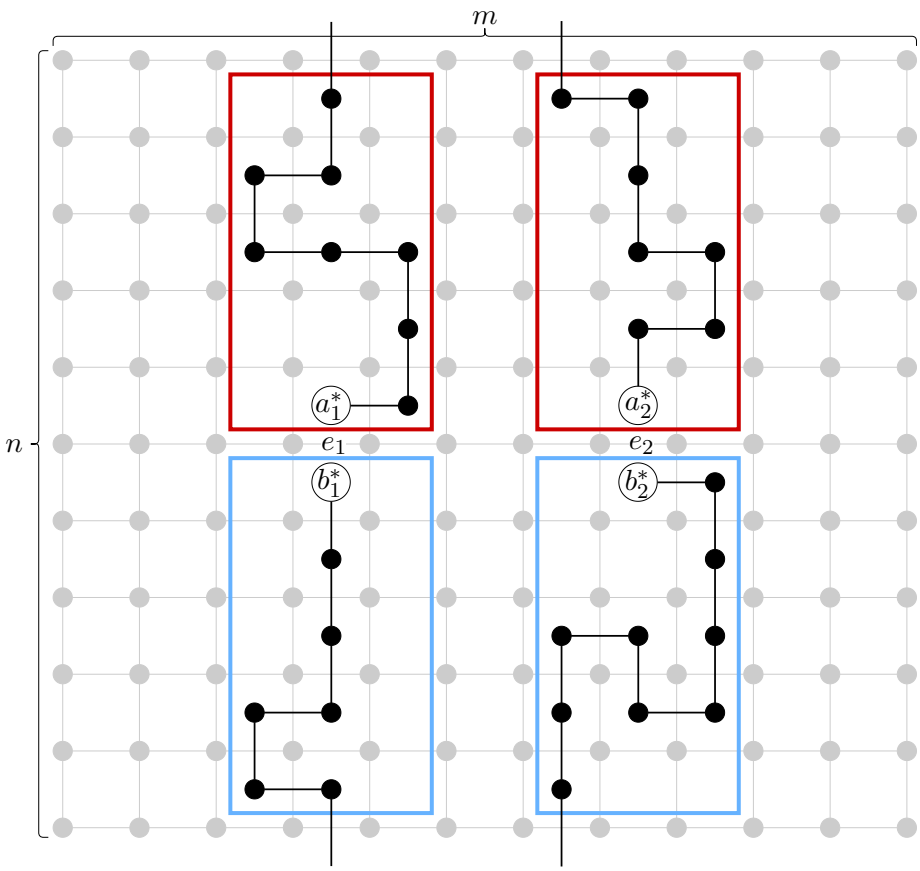


Figure 5: A possible run of the dual graph spanning tree sampling algorithm in the proof of Theorem 15 when m is odd. In this example, $m = 12$, $n = 11$, and $k = 3$. The primal graph G is depicted in gray, and the first four random walks in the dual graph G^* are depicted in black.

the far side of its rectangle with probability at least

$$\frac{1}{Ane^{-Ak/2}}, \tag{12}$$

for a fixed constant A.

When this happens, we then apply the same symmetry argument from the proof of Lemma 12 to each pair of rectangles to conclude that, within each red-blue pair of rectangles, the the number of vertices of G to the left of the paths is equal to the number of vertices in to the right of the paths with probability at least $\frac{1}{2p+1}$, where

$$p := \left(\frac{m}{k} - 1\right) \cdot \left(\frac{n-1}{2} - 1\right)$$

is (an upper bound on) the number of vertices of G in each rectangle. Simplifying this expression, a lower bound on this probability is $\frac{k}{mn}$. By independence, all pairs of rectangles will be balanced (as they are in Figure 5) with probability at least

$$\left(\frac{k}{mn}\right)^{k-1},$$

in which case $T \setminus S$ will be a balanced k-forest. Altogether, this will happen with probability at least

$$\left(\frac{1}{Ane^{Ak/2}}\right)^{2(k-1)} \cdot \left(\frac{k}{mn}\right)^{k-1} \geq \frac{1}{\beta^{k^2} m^{k-1} n^{3k-3}}$$

for β a fixed constant.

In the case where n is even, we may apply the same trick as in the proof of Lemma 12, assuming that each path goes downward one step first. This simply tacks on another factor of $\frac{1}{4^{k-1}}$ to the final probability bound, which can be incorporated into the β^{k^2} term. □

4 Approximating Partitions of Lattice Structures

In this section we move beyond grid graphs to a more general class of lattice-like graphs. In Section 4.1, we begin by defining what we mean by lattice-like graphs, including the more general definition which we use in our multiplicative approximate balance results and the stronger version (*uniform* lattice sequences) we use for our additive approximate balance results. In Section 4.2 we state our results for both lattice classes, including an interesting corollary for grids. In Section 4.3, we give lower bounds on the probability that a random walk in a lattice sequence follows a given curve in the plane. These bounds are used in the proofs of both our multiplicative and additive approximate balance results. In Section 4.4, we apply them to analyze a particular implementation of Wilson's algorithm that uses a random walk to choose the next starting place for a loop-erased random walk, which suffices to prove our multiplicative approximate balance result. Finally, in Section 4.5, we consider a shrinking sequence of rectangles to correct multiplicative approximate balanced into additive approximate balance.

4.1 Lattice sequences

We use $\text{dist}(x, y)$ to denote graph (e.g. lattice) distances, and $d(x, y)$ and $d(X, Y)$ to denote Euclidean distances between points and Hausdorff distance between sets, respectively. Recall that a *curve* γ is a continuous function $\gamma : [a, b] \to \mathbb{R}^2$ for $a < b$. Except where specified otherwise, we take $a = 0, b = 1$. A plane graph $D = (V, \Gamma)$ is a drawing of a (planar) graph in the plane without intersections. In particular, $V = V(D)$ is a finite set of points in the plane $\mathbb{R}^2$. Γ is a finite collection of curves given by continuous functions $\gamma_i : [0, 1] \to \mathbb{R}^2$ such that no two such curves intersect except possibly at their endpoints, and such that if E is the set of pairs of endpoints:

$$E = \{\{\gamma_i(0), \gamma_i(1)\} \mid i = 1, \ldots, |\Gamma|\},$$

then (V, E) is a graph. The faces of D are the connected components of $\mathbb{R}^2 \setminus \bigcup_{i=1}^{|\Gamma|} \text{im}(\gamma_i)$ (here $\text{im}(\gamma)$ is the image of the function γ), and we refer to the unique unbounded face as the *outer* face.

We state our results in terms of sequences of infinite lattice-like plane graphs that get finer and finer with properties defined as follows. We consider two slightly different notions: essentially, the second adds a requirement that fluctuations are bounded at some constant scale, which will be necessary for our result on partitions that are balanced up to additive error.

Definition 18. A *lattice sequence* is a pair $(\{\Lambda_n\}, \rho)$, where $\{\Lambda_n\}$ is sequence of plane graphs with vertex sets $V(\Lambda_n) \subseteq \mathbb{R}^2$ for which there are corresponding dual plane graphs Λ_n^*, with each vertex $v \in V(\Lambda_n^*)$ a point in the corresponding face of Λ_n, such that for all $\varepsilon > 0$, there exists N such that for all $n > N$,

(a) For any adjacent pair $x, y \in V(\Lambda_n^*)$, $d(x, y) < \varepsilon$,

(b) For any $p \in \mathbb{R}^2$, the ball $B_\varepsilon(p)$ contains a vertex of Λ_n^*,

(c) For any dual vertex v and $\varepsilon > 0$ there is a division of the circle $C_{\varepsilon,v}$ of radius ε into arcs $A_1, \ldots, A_s$ each of length at most $\frac{1}{8} 2\pi\varepsilon$, such that the following holds for each $i \in [s]$: In a simple random walk $v = v_0, v_1, v_2, \ldots$, with probability at least $\rho > 0$, letting j the first index where $d(v_0, v_j) > \varepsilon$, the straight line segment joining v_{j-1} with v_j passes through $C_{\varepsilon,v}$ in A_i.

This definition generalizes sequences of finer and finer grids. Specifically, we claim that the family of plane graphs $\{\mathbb{Z}^2, \frac{1}{2}\mathbb{Z}^2, \frac{1}{3}\mathbb{Z}^2, \ldots\}$ is a lattice sequence with $\rho = \frac{1}{8}$. Property (a) and (b) say that, as we increase n, neighboring vertices in $\frac{1}{n}\mathbb{Z}^2$ become arbitrarily close, and the vertex set becomes arbitrarily dense. Property (c) holds using the partition of $C_{\varepsilon,v}$ obtained by drawing horizontal, vertical, and both 45 degree diagonal lines through v. By symmetry, a random walk from v is equally likely to exit through each of these 8 arcs.

The definition also applies to scalings of lattices like the triangular or hexagonal lattice, since, for example, property (c) holds for any sequence of lattices for which the scaling limit of random walk on the dual is Brownian motion. For the same reason, it applies a.s., for example, to the sequence where Λ_n denotes the Delauney triangulation of a Poisson cloud in $\mathbb{R}^2$ of rate n, see [28]. The choice of $\frac{1}{8}$ in the definition is made just for convenience; replacing it with any constant $< \frac{1}{2}$ would give exactly the same family of lattice sequences.

For our result about partitions that are balanced up to an additive error (Theorem 22), we require the following strengthening of this notion.

Definition 19. A *uniform lattice sequence* is a tuple $(\{\Lambda_n\}, \rho, R, C_1, C_2)$, where $\{\Lambda_n\}$ is sequence of plane graphs with vertex sets $V(\Lambda_n) \subseteq \mathbb{R}^2$ for which there are corresponding dual plane graphs Λ_n^*, with each vertex $v \in V(\Lambda_n^*)$ a point in the corresponding face of Λ_n, and positive constants ρ, R, C_1, C_2 so that, for all sufficiently large n:

(a) For any $p \in \mathbb{R}^2$, the ball $B_{R/n}(p)$ contains between C_1 and C_2 vertices of Λ_n^*,

(b) Any adjacent pair $x, y \in V(\Lambda_n^*)$ satisfies $d(x, y) < \frac{R}{n}$,

(c) For any dual vertex v there is a division of the circle $C_{r,v}$ of radius $r \geq \frac{R}{n}$ into arcs $A_1, \ldots, A_s$ each of length at most $\frac{1}{8} 2\pi R/n$, such that the following holds for each $i \in [s]$: In a simple random walk $v = v_0, v_1, v_2, \ldots$, with probability at least ρ, letting j the first index where $d(v_0, v_j) > \varepsilon$, the straight line segment joining v_{j-1} with v_j passes through $C_{r,v}$ in A_i.

This definition is also satisfied for grids, the triangular lattice, the hexagonal lattice, or indeed any finite-degree, doubly-periodic connected plane graph. However, it is not satisfied for triangulations of random point clouds, since as $n \to \infty$, as the fluctuations in density will be too large.

4.2 Statement of results

In addition to lattice sequences, we will also consider a fixed bounded plane graph D that gives the partition structure we are looking to approximate. In doing this, we will need to restrict the infinite graphs in the lattice sequence to a reasonable bounded subgraph that falls inside D, and we do this as follows. Let D be a bounded plane graph, and fix $\delta > 0$ that will be chosen later in terms of D (in Lemma 25). Given plane graph Λ_n with dual Λ_n^* from a lattice sequence, and a cycle C^* in Λ^* at Hausdorff distance $< \delta$ from the outer face boundary of D, we let Ω_{D,Λ_n} be the subgraph of of Λ_n lying inside C^*. We let Ω^*_{D,Λ_n} be the subgraph of Λ_n^* induced by all vertices of C^* along with the vertices of Λ_n^* lying inside C^*. In this way, we can consider the planar dual of Ω_{D,Λ_n} to be Ω^*_{D,Λ_n} with *wired boundary condition*, where the entire cycle C^* (rather than a single dual vertex) corresponds to the outer face of Ω_{D,Λ_n}. (In particular, for our proofs, we will run Wilson's algorithm on Ω^*_{D,Λ_n} with the cycle C^* identified as a single root.)

Given this plane graph D describing the partition structure we are looking to approximate, let $k+1$ be the number of faces and denote the k bounded faces by $\phi_1, \ldots, \phi_k$. We say a partition of the graph $\Omega_{D,\Lambda}$ into connected components $C_1, \ldots, C_k$ is *ε-compatible* with D if for all i and vertices $v \in \Omega_{D,\Lambda}$, the implication

$$v \in C_i \implies d(v, \phi_i) \leq \varepsilon \tag{13}$$

holds. By subdividing edges if necessary, we will assume that D has no loops, so that $\gamma(0) \neq \gamma(1)$ for all $\gamma \in \Gamma(D)$.

For a lattice sequence $(\{\Lambda_n\}, \rho)$ and a probability space on the set of spanning trees of Ω_{D,Λ_n} and given $\varepsilon > 0$, we define the event $\mathcal{E}_{D,\Lambda_n,\varepsilon}$, which holds whenever there are $k-1$ edges whose removal from T results in a forest with components $C_1, \ldots, C_k$ that is ε-compatible with D. The following is our main result for multiplicative balance (the formal version of Theorem 3).

Theorem 20. *Let $(\{\Lambda_n\}, \rho)$ be a lattice sequence, let D be a plane graph with $k+1$ faces, and let Ω_{D,Λ_n} be as above. For the uniform probability space on the set of spanning trees of a graph Ω_{D,Λ_n}, we have that as $n \to \infty$, $\Pr(\mathcal{E}_{D,\Lambda_n,\varepsilon})$ is bounded below by a constant depending only on D and ε.*

As a consequence, if we draw the partition so that the parts contain approximately equal numbers of vertices, we can conclude that random trees are splittable into approximately balanced pieces with constant probability. This is possible so long as $\{\Lambda_n\}$ has the property that for any $\delta > 0$ and R, there is an $\varepsilon > 0$ so that every ε ball $B_\varepsilon(p)$ satisfies $|B_\varepsilon(p) \cap V(\Lambda_n)| \leq \delta |B_R(0) \cap V(\Lambda_n)|$. In the case of grid graphs, for instance, we obtain the following corollary.

Corollary 21. *Fix $\varepsilon \geq 0$ and k a positive integer. Let m, n be positive integers such that $n \leq m$, $k|m$, and $20/n \leq \varepsilon \leq 1/(3k)$. Let G be an $m \times n$ grid graph. There is a constant $C(k, \varepsilon)$ such that the probability a uniformly random spanning tree of G is (k, ε)-approximately splittable is at least $C(k, \varepsilon)$.*

As our final main result, we give 1/poly lower bound on the probability that a random spanning tree in a square region of a uniform lattice sequence can be cut into pieces that differ in size by an *additive constant* on uniform lattice sequences (the formal version of Theorem 4).

Theorem 22. *Suppose $(\{\Lambda_n\}, \rho, R, C_1, C_2)$ is a uniform lattice sequence, and for all $n > N$, C_n^* in Λ_n^* is a cycle in the dual that is contained within the unit square $[0,1]^2$ and at Hausdorff distance $< R/n$ from its boundary, and $\Omega^S_{\Lambda_n}$ denotes the subgraph of Λ_n enclosd by the cycle C^*. Then there is a constant A such that with probability at least $\frac{1}{\mathrm{poly}(n)}$, a uniformly random spanning tree of $\Omega^S_{\Lambda_n}$ can be disconnected by the removal of $k-1$ edges into k components that that differ in size by at most A.*

Note that balance up to an additive constant is the best we can hope for in the generality we are working in here; for example, an $n \times n$ grid graph with n odd and an additional leaf appended to every vertex is a case to which Theorem 22 applies, but that does not admit an exactly balanced partition despite the fact that the total number of vertices is even.

4.3 Probability of staying near a curve

To prove Theorem 20, we first need the following lemma.

Lemma 23. *Let $(\{\Lambda_n\}, \rho)$ be a lattice sequence, let $\varepsilon > 0$, and let γ be a curve in the plane of length T. For $v_0 \in \Lambda_n$ with $d(v_0, \gamma(0)) \le \varepsilon/2$, and for a random walk in Λ_n started from v_0, let $\mathcal{E} = \mathcal{E}_{n,v_0,\varepsilon,\gamma}$ be the event that the walk reaches a point within $\frac{\varepsilon}{2}$ of $\gamma(1)$ before ever reaching a vertex at distance $> \varepsilon$ from the curve γ. For all sufficiently large n,*

$$\Pr[\mathcal{E}_{n,v_0,\varepsilon,\gamma}] \ge \rho^{20T/\varepsilon}$$

The proof is reminiscent of the proof of Lemma 17. The differences are that now we consider a sequence of circles that the random walk must escape from rather than squares, and the direction to which we hope it escapes will be determined by the path γ, not just always going upward (see Figure 6).

Proof. Choose n large enough so that neighboring vertices in Λ_n are at distance at most $\frac{\varepsilon}{20}$. For any $v \in \Lambda_n$, we may partition $C_{\varepsilon/2,v}$ (the circle of radius $\varepsilon/2$ around v) into arcs $A_1, A_2, \ldots, A_s$ as in Definition 18 (c) applied to $\frac{\varepsilon}{2}$. For any $t \in [0,1]$ such that v is within Euclidean distance $\varepsilon/2$ of $\gamma(t)$ but not within Euclidean distance $\varepsilon/2$ of $\gamma(1)$, that is,

$$d(v, \gamma(t)) \le \frac{\varepsilon}{2} \le d(v, \gamma(1)), \tag{14}$$

we define $A^*(v,t)$ to be the union of all arcs A_i such that the next time γ reaches $C_{\varepsilon/2,v}$ it is in the arc A_i, that is, all arcs A_i such that

$$\gamma(\inf\{t' \in [t,1] \mid d(v, \gamma(t')) = \varepsilon/2\}) \in A_i.$$

Note the infimum is well-defined by (14), recalling that $d(\cdot)$ is Euclidean distance, not lattice distance. We then define $\mathcal{L}_{v,t}$ to be the event that a random walk in Λ_n from v first crosses $C_{\varepsilon/2,v}$ through $A^*(v,t)$.

Assume that v_0 is not already within $\frac{\varepsilon}{2}$ of $\gamma(1)$, otherwise there is nothing to show. Then let t_1 be the first time such that $d(v_0, \gamma(t_1)) = \frac{\varepsilon}{2}$. By property (c), we know that $\mathcal{L}_{v_0,0}$ will occur with probability at least ρ. Assuming this happens, suppose the random walk exits to a point $v_1 \in \Lambda_n \setminus B_{\frac{\varepsilon}{2}}(v_0)$ from a previous point $u_1 \in \Lambda_n \cap B_{\frac{\varepsilon}{2}}(v_0)$, where the line segment between u_1 and v_1 passes through a point $a_1 \in A^*(v_0, 0)$ as in Figure 6. We know that the arc length along this circle from a_1 to $\gamma(t_1)$ is at most $\frac{1}{8} 2\pi \frac{\varepsilon}{2} = \frac{\pi}{8}\varepsilon$, so

$$d(v_1, \gamma(t_1)) \le d(v_1, a_1) + d(a_1, \gamma(t_1)) \le d(v_1, u_1) + \frac{\pi}{8}\varepsilon \le \frac{\varepsilon}{20} + \frac{\pi}{8}\varepsilon \le \frac{9}{20}\varepsilon. \tag{15}$$

If $d(v_1, \gamma(1)) \le \frac{\varepsilon}{2}$, we are done and have shown $\Pr(\mathcal{E}) \ge \rho$. If, instead, $d(v_1, \gamma(1)) > \frac{\varepsilon}{2}$, then both inequalities in (14) hold. We then let $t_2 > t_1$ be the first time (later than t_1) such that $d(v_1, \gamma(t_2)) = \frac{\varepsilon}{2}$ and observe that, with probability at least ρ, the event $\mathcal{L}_{v_1,t_1}$ will occur, at which point the random walk will exit to some vertex $v_2 \in B_{\frac{\varepsilon}{2}}(\gamma(t_2))$. Continuing inductively until $d(v_L, \gamma(1)) \le \frac{\varepsilon}{2}$, we see that we will continue to follow the curve γ for a sequence of L steps,

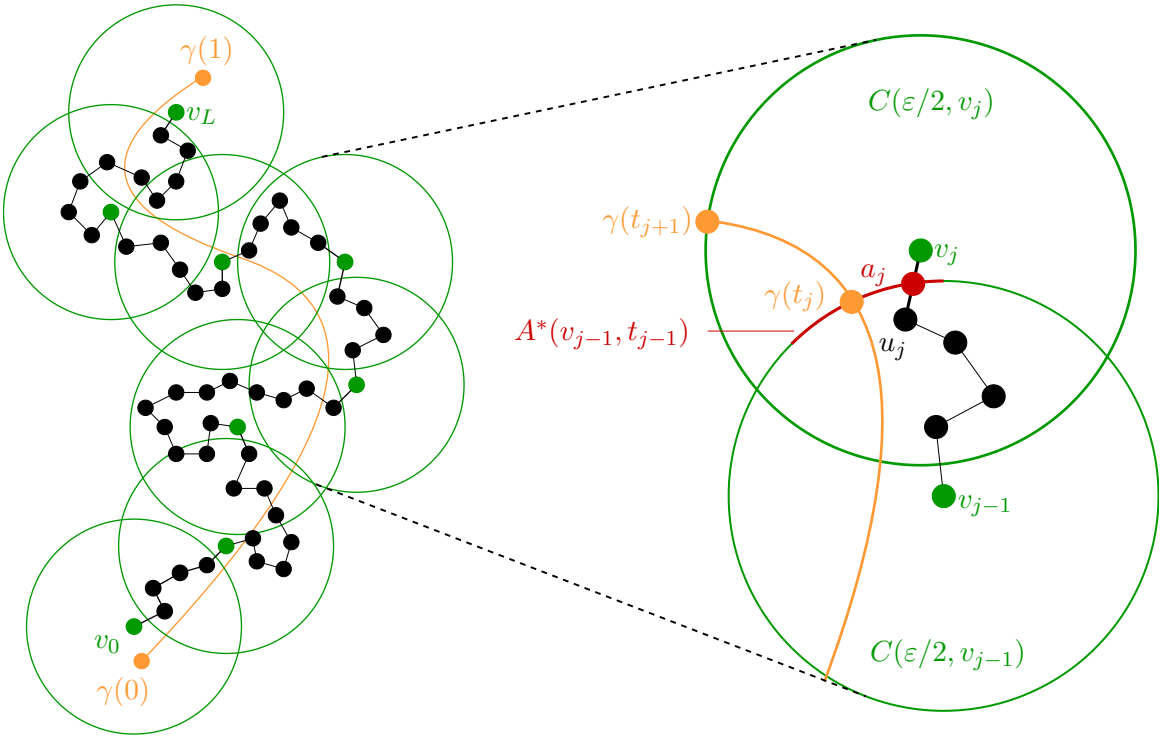


Figure 6: Illustration accompanying the proof of Lemma 23. A random walk from v_0 in the dual lattice eventually reaches a vertex v_L within $\frac{\varepsilon}{2}$ of $\gamma(1)$ while staying within ε of the curve γ because each of the events $\mathcal{L}_{v_0,0}, \mathcal{L}_{v_1,t_1}, \mathcal{L}_{v_2,t_2}, \ldots, \mathcal{L}_{v_{L-1},t_{L-1}}$ occur. A key step of the proof is lower-bounding the distance between $\gamma(t_j + 1)$ and $\gamma(t_j)$, which we accomplish by observing that v_j is closer to $\gamma(t_j)$ than $\gamma(t_{j+1})$, since it is close to a_j and the arc $A^*(v_{j-1}, t_{j-1})$ is small.

through events $\mathcal{L}_{v_0,0}, \mathcal{L}_{v_1,t_1}, \mathcal{L}_{v_2,t_2}, \ldots, \mathcal{L}_{v_{L-1},t_{L-1}}$, with probability at least ρ^L. Observe that the bound in (15) applies to each round, choosing a_j to be the point on the circle between v_j and the previous step u_j. Thus, we may bound the distance between consecutive points on the path as

$$d(\gamma(t_{j+1}), \gamma(t_j)) \geq d(\gamma(t_{j+1}), v_j) - d(v_j, \gamma(t_j)) \geq \frac{\varepsilon}{2} - \frac{9\varepsilon}{20} = \frac{\varepsilon}{20}.$$

Since T is total length of the curve γ, it follows that we only need at most $L = \frac{20T}{\varepsilon}$ steps of the correct events $\mathcal{L}_{v_j,t_j}$ occurring until we reach a vertex within $\frac{\varepsilon}{2}$ of $\gamma(1)$. All the while, we know that the path is always within a ball of radius $\frac{\varepsilon}{2}$ that contains parts of γ, so it never is more than ε away from γ. Thus, we have

$$\Pr[\mathcal{E}] \geq \rho^{20T/\varepsilon}. \qquad \square$$

We will also require the following slightly stronger version of this lemma with some extra conditions about how the path ends:

Lemma 24. *Let $(\{\Lambda_n\}, \rho)$ be a lattice sequence, let γ_1, γ_2 be curves in the plane of positive length, where γ_1 has length T and $\gamma_2(0) = \gamma_1(1)$. For all sufficiently small $\varepsilon > 0$, for any $v_0 \in \Lambda_n$ with $d(v_0, \gamma(0)) \leq \varepsilon/2$, and for a random walk in Λ_n started from v_0, let $\mathcal{E}' = \mathcal{E}'_{n,v_0,\gamma_1,\gamma_2}$ be the event that the walk traverses an edge which intersects the curve γ_2 at a point within $\varepsilon/2$ of $\gamma_1(1)$ before ever reaching a vertex at distance $> \varepsilon$ from the curve γ_1. For all sufficiently large n,*

$$\Pr[\mathcal{E}'_{n,v_0,\gamma_1,\gamma_2}] \geq \rho^{100T/\varepsilon+200}.$$

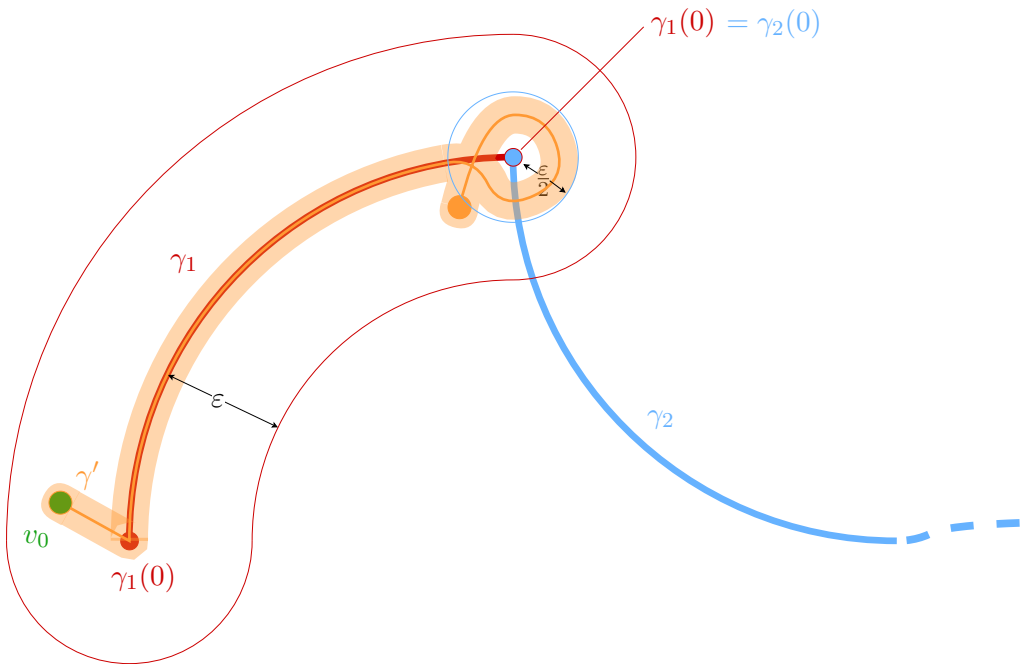


Figure 7: Illustration accompanying the proof of Lemma 24. We apply Lemma 23 to the orange path γ' and $\frac{\varepsilon}{5}$, which is the radius of the orange tube around γ'. The quantities $\frac{\varepsilon}{5}$ and $\frac{3\varepsilon}{2}$ from the proof are crude bounds on the width and length of the last part of the tube that must encircle $\gamma_2(0)$ while staying within the blue circle of radius $\frac{\varepsilon}{2}$.

Proof. Assume that ε is small enough so that γ_2 eventually leaves $B_{\frac{\varepsilon}{2}}(\gamma_2(0))$. We apply Lemma 23 with $\frac{\varepsilon}{5}$ to the path γ' drawn in Figure 7. The path consists of 3 parts: a straight line from v_0 $\gamma_1(0)$, then γ_1 until it enters $B_\varepsilon(\gamma_2(0))$, then a loop that crosses itself, encircling $\gamma_2(0)$ contained within $B_{\frac{\varepsilon}{2}}(\gamma_2(0))$. Then observe that $\mathcal{E} \subseteq \mathcal{E}'$, where $\mathcal{E}$ is the event from Lemma 23 applied to this new path using $\varepsilon/5$. This is because any random walk in Λ_n satisfying $\mathcal{E}$ stays within ε of γ_1 and then traverses a cycle of edges encircling $\gamma_2(0)$ while staying within ε of $\gamma_2(0)$. Since we chose ε such that γ_2 eventually leaves $B_{\frac{\varepsilon}{2}}(\gamma_2(0))$, one of these edges must intersect γ_2. We may upper bound the length of γ' as

$$T' \leq \frac{\varepsilon}{2} + T + \frac{3\varepsilon}{2} = T + 2\varepsilon,$$

so the bound follows. □

Note that when a disconnected plane graph D has bounded faces $\phi_1, \ldots, \phi_k$ and satisfies the hypotheses of Theorem 20, we can add curves to D to create a connected plane graph D' whose bounded faces satisfy $\phi_i' \subseteq \phi_i$ for all i. We have then that $\mathcal{E}_{D',\Lambda_n,\varepsilon} \subseteq \mathcal{E}_{D,\Lambda_n,\varepsilon}$, and thus it suffices to prove Theorem 20 in the case where D is connected. Therefore, we assume that D is connected for the rest of the proof.

By compactness of the curves $\gamma \in \Gamma$ for the plane graph $D = (V, \Gamma)$, we have the following:

Lemma 25. *For $0 < \delta < \varepsilon$ sufficiently small, we have that:*

(A1) The distance between any two vertices $u, v \in V$ is at least 3ε.

(A2) *If a_γ denotes the last time t at which $\gamma(t)$ is in the closed ball of radius ε about $\gamma(0)$, and b_γ denotes the first time t at which $\gamma(t)$ is in the closed ball of radius ε about $\gamma(1)$, the restricted curves $\bar{\gamma} = \gamma_{|[a_\gamma, b_\gamma]}$ are all at distance at least 3ε each other.*

(A3) *For any two points p, q in a common face of D and both at distance at least ε from any curve of D, there is a curve γ (not a curve of D) joining p to q whose distance to every curve of D is at least δ.*

This is a straightforward consequence of compactness of the curves; we include a proof in the appendix for completeness.

4.4 Multiplicative Approximate Balance

From here on, we let ε, δ be as promised by Lemma 25. We call a curve of D an *outer curve* if every point of the curve lies on the boundary of the outer face, and an *inner curve* if no point does, other than possibly its endpoints. Note that every curve must be one of these two types. We let Γ^I and Γ^O denote set of inner curves and outer curves, respectively. Since D is connected, we can order the inner curves of D as $\gamma_1^I, \ldots, \gamma_{m_i}^I$ such that for all ℓ, the plane graph D_ℓ of D with edges

$$\Gamma_\ell = \Gamma^O \cup \{\gamma_1^I, \ldots, \gamma_\ell^I\}$$

and vertex set $V_\ell = \{\{\gamma(x) \mid \gamma \in \Gamma_\ell, x \in \{0,1\}\}$ is connected (see Figure 8). Moreover, without loss of generality we assume that the orientation of each curve is such that $\gamma_\ell^I(1)$ is a vertex of $\Gamma_{\ell-1}$ for all $\ell = 1, \ldots, m_i$ (the curves are oriented "towards the outer face").

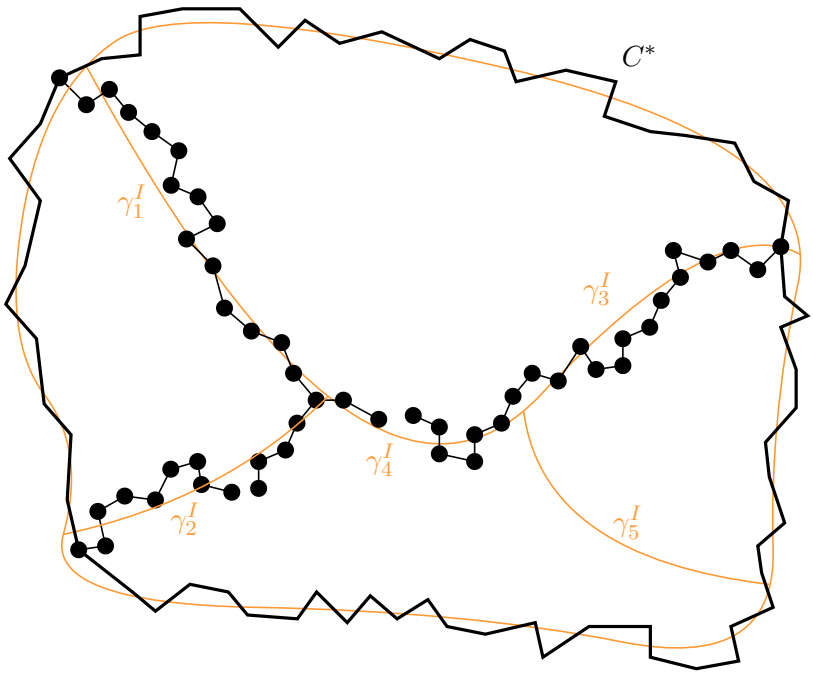


Figure 8: A run of the first four phases of Wilson's algorithm approximating the first four inner curves $\gamma_1^I, \gamma_2^I, \gamma_3^I, \gamma_4^I$. Note that paths corresponding to γ_2^I and γ_4^I each are missing an edge, and in particular, the edges present do not disconnect the interior of C^*. These missing edges are dual to the edges in the corresponding primal spanning tree whose removal would disconnect the tree into components approximating the faces of this drawing.

We prove the theorem by analyzing how Wilson's algorithm constructs spanning trees of Ω^*_{D,Λ_n} with wired boundary conditions, where the whole boundary cycle C^* is used as the root of Wilson's algorithm. In particular, it is equivalent to view Wilson's algorithm as building a unicylic graph, initialized with the boundary cycle of Ω^*_{D,Λ_n}. We take advantage of the freedom to choose the starts of loop-erased random walks afforded by Wilson's algorithm by using the following implementation, which determines starts by using additional random walks. We construct trees in rounds, where for tree T_i^j the subscript i denotes the phase and the superscript j denotes the step within phase i. Within phase i, we alternate a loop-erased random walk from a vertex outside T_i^j to a vertex in T_i^j that gets added to the tree to obtain T_i^{j+1} (one step of Wilson's algorithm) with a random walk among the vertices in T_i^{j+1} until a vertex outside T_i^{j+1} is reached (choosing the starting point for the next step of Wilson's algorithm). Here we describe the procedure for one phase of this process; the particular choices of source vertex and target subsets that will be useful for our purposes will be specified below.

1. At the beginning of each phase i, we have an existing tree T_{i-1} that has already been built. (For $i = 1$, T_0 consists of just the root vertex). We choose a source vertex v and target subset $U_i \subseteq T_{i-1}$ for this phase, and initialize $T_{i-1}^0 = T_{i-1}$.

2. We begin each step of this phase with T_{i-1}^{j-1} (at the beginning of the phase, for $j = 1$) and a source vertex. We do one of two things according to whether the source belongs to the tree T_{i-1}^{j-1}:

   (a) If the source vertex is not in T_{i-1}^{j-1}, we conduct a loop-erased random walk from the source until it hits T_{i-1}^{j-1} at a vertex u. This loop-erased random walk is added to T_{i-1}^{j-1} to create T_{i-1}^j. If $u \in U_i$ this phase ends, and we set $T_i = T_{i-1}^j$. Otherwise, we increment j, and continue this phase with u as the new source vertex (we will be in case (b) next).

   (b) If the source vertex is in T_{i-1}^{j-1}, we take a random walk from the source until we reach a vertex u outside of T_{i-1}, and then increment j and restart this step from the vertex u as the new source vertex, and $T_{i-1}^j = T_{i-1}^{j-1}$ (we will be in case (a) next).

3. The previous loop continues until either the target is eventually hit by an instance of the loop-erased random walk, or the entire spanning tree is completed.

Note that we can use a single random walk W_i from $v \notin T_{i-1}$ to implement each phase of the algorithm (with loop erasure while in case (2a), and without loop erasure while in case (2b)). In particular, with this implementation, we have the following observation for general graphs:

Observation 26. *Suppose that we run the implementation of Wilson's algorithm above on a graph G, and have built the tree T_ℓ after the first ℓ phases.*

(a) For a connected set of vertices S, if the walk $W_{\ell+1}$ begins from a vertex $v \in S$ and ends phase $\ell + 1$ by hitting T_ℓ for the first time at the target $U_{\ell+1} \subseteq S$ and without leaving S, then after this phase, there is a path P in $T_{\ell+1}$ joining v to $U_{\ell+1}$.

(b) For a connected set of vertices S, suppose the walk $W_{\ell+1}$ begins from a vertex in S that is adjacent to $v \in T_\ell \cap S$ and ends phase $\ell + 1$ by hitting the target $U_{\ell+1} \subseteq S$ without leaving S. Then after this phase, there is a path P from a vertex $v' \in (S \cap T_\ell) \setminus U_{\ell+1}$ to $U_{\ell+1}$, all of whose vertices belong to S, and all but one of whose edges belong to $T_{\ell+1}$.

Proof. In (a), the path P consists of loop-erased W_{l+1} from v until the first time it reaches T_ℓ, and then any path within T_ℓ to $U_{\ell+1}$. Both segments of this path are in $T_{\ell+1}$.

In (b), we construct the desired path backwards from $U_{\ell+1}$. We first add to P the last loop-erased part of $W_{\ell+1}$ before it reaches $U_{\ell+1}$, corresponding to an entire step as described in Case (a) of our implementation of Wilson's algorithm. Note this entire path is in $T_{\ell+1}$, and it is entirely contained in S because it is a subset of $W_{\ell+1}$ and all vertices of $W_{\ell+1}$ are in S. Let y be the first vertex of this path. If y is the first vertex of $W_{\ell+1}$, it must be adjacent to v and adding v to the start of this path produces a path with the desired properties with $v' = v$. Otherwise, let x be the vertex preceding y in $W_{\ell+1}$. It must be that $x \in T_{\ell+1}$ and there is no edge in $T_{\ell+1}$ between x and y, as x is the penultimate vertex in a step (as described in case (b) of our implementation of Wilson's algorithm) of $W_{\ell+1}$ that is entirely contained in $T_{\ell+1}$ until its last vertex, y, is (at that point) outside $T_{\ell+1}$. Because $W_{\ell+1}$ doesn't reach $U_{\ell+1}$ until after it visits v', and there must be some component of $S \cap T_\ell$ in the same component of $T_{\ell+1}$ as x because otherwise x wouldn't have been added to $T_{\ell+1}$, it follows that x is in the same component of $T_{\ell+1} \cap S$ as some other vertex $v' \in (S \cap T_\ell) \setminus U_{\ell+1}$. We then let path P consists of the path from v' to x in $T_{\ell+1} \cap S$, the edge $\{x, y\}$, and the final loop-erased part of $W_{\ell+1}$ from y to $U_{\ell+1}$. All vertices of this path are in S, and all edges except $\{x, y\}$ are in $T_{\ell+1}$. □

In particular, applied to our situation, using as S the set of vertices close to a given curve, we obtain the following:

Observation 27. *Suppose that we run the implementation of Wilson's algorithm above on the dual graph Ω^*_{D,Λ_n}, and have built the tree T_ℓ after the first ℓ phases. Then:*

(a) If the walk $W_{\ell+1}$ begins from a vertex v and ends phase $\ell+1$ by hitting T_ℓ for the first time at the target $U_{\ell+1} = T_\ell \cap B(\delta, \gamma_{\ell+1}(1))$ while also staying within distance δ of the curve $\gamma_{\ell+1}$, then after this phase, there is a path P in $T_{\ell+1}$ joining v to $U_{\ell+1}$ whose vertices are all within distance δ from the curve $\gamma_{\ell+1}$.

(b) Suppose the walk $W_{\ell+1}$ begins from a vertex adjacent to a vertex $v \in T_\ell \cap B(\delta, \gamma_{\ell+1}(0))$ and ends phase $\ell+1$ by hitting the target $U_{\ell+1} = T_\ell \cap B(\delta, \gamma_{\ell+1}(1))$ while also staying within distance δ of the curve $\gamma_{\ell+1}$ throughout the phase. Suppose all vertices of T_ℓ that are within δ of the curve $\gamma_{\ell+1}$ are in either $B(\delta, \gamma_{\ell+1}(0))$ or $B(\delta, \gamma_{\ell+1}(1))$. Then after this phase, there is a path P in the dual graph joining some vertex in $T_\ell \cap B(\delta, \gamma_{\ell+1}(0))$ to $U_{\ell+1}$ whose vertices are all within distance δ from the curve $\gamma_{\ell+1}$, and such that all but at most one edge of P belongs to the tree $T_{\ell+1}$.

Proof. This follows from Observation 26, with $G = \Omega^*_{D,\Lambda_n}$ and S to be all the vertices of Ω^*_{D,Λ_n} within distance δ of the curve $\gamma_{\ell+1}$. □

Using Observation 27, we complete the proof as follows. For a plane graph D, let $f(D)$ be the number of interior faces of D. We say a (not-necessarily-spanning) tree T in the dual graph Ω^*_{D,Λ_n} *δ-corresponds* to a plane graph D with inner curves $\gamma^I_1, \ldots, \gamma^I_\ell$ if there are paths $P_1, \ldots, P_\ell$, such that for each i:

(a) P_i and P_j intersect if and only if γ^I_i and γ^I_j share a common endpoint.

(b) For any point p, if $\gamma^I_{i_1}, \ldots, \gamma^I_{i_s}$ are all the curves of D that have p as an endpoint, then the union of the paths $P_{i_1}, \ldots, P_{i_s}$ is a tree.

(c) Every vertex of P_i is within distance δ of the curve γ^I_i.

(d) For $f(D)-1$ of the paths, all but one edge of P_i belongs to the tree T, while for the rest of the paths, the whole path belongs to T. Here $f(D)$ is the number of faces of D.

We then have the following lemma, which applies to each plane graph D_ℓ in the sequence constructed above.

Lemma 28. *Let T be a tree in the dual graph Ω^*_{D,Λ_n}, and let H_T be the spanning subgraph of the primal graph $\Omega_{D,h}$ obtained by removing from Ω_{D,Λ_n} all the edges e for which $e^* \in T$. If T δ-corresponds to D_ℓ, then there are $f(D_\ell)-1$ edges of H whose removal results in connected components $C_1,\dots,C_k$ whose induced partition of Ω_{D,Λ_n} is ε-compatible with D_ℓ.*

Let us first use Lemma 28 and induction on the sequence of plane graphs D_ℓ to prove the Theorem. Initially, tree T_0 consisting of just the root vertex of Ω^*_{D,Λ_n} (equivalently, of just the wired boundary cycle of Ω^*_{D,Λ_n}) trivially δ-corresponds to the plane graph D_0 with no interior curves. Having already constructed a tree $T_{\ell-1}$ that δ-corresponds to $D_{\ell-1}$, we begin another phase of Wilson's algorithm from a vertex $v \in \Omega_{D,\Lambda_n}$. If $T_{\ell-1} \cap B(\delta, \gamma^I_\ell(0)) = \emptyset$, we begin from a vertex v at minimum distance from $\gamma^I_\ell(0)$ (call this Case A), and note v will not be in $T_{\ell-1}$. Otherwise, if $T_{\ell-1} \cap B(\delta, \gamma^I_\ell(0)) = \emptyset$, we begin at a vertex v adjacent to any vertex in this set (Case B). In both cases, we use the target $U_\ell = T_{\ell-1} \cap B(\delta, \gamma^I_\ell(1))$. Note that when $\gamma^I_\ell(1)$ is incident on the outer face of D, the target U_ℓ contains a portion of the boundary cycle of Ω^*_{D,Λ_n} within distance δ of $\gamma^I_\ell(1)$. By Lemma 24, with constant probability, the walk W_ℓ for this phase will hit the target U_ℓ while staying within distance δ of γ^I_ℓ. Thus, by Observation 27, and the fact that we are in Case A instead of Case B if and only if $f(D_\ell) = f(D_{\ell-1})$ (as in Euler's formula), at the end of the phase, the tree T_ℓ δ-corresponds to D_ℓ.

In particular, after phase $\ell = m_i$ (recall m_i is the total number of interior curves in D) we have that with constant probability, our tree T_{m_i} consists only of the root and vertices within distance δ of the internal curves of D, and that there is a collection of paths $P_1,\dots,P_{m_i}$ satisfying properties (a),(b),(c),(d) above, and additionally that all but precisely $f(D)-1 = k-2$ of the paths belong entirely to three T_{m_i}.

From here, we complete Wilson's algorithm with arbitrary choices for starting vertices to produce a final tree T, which still contains all but at most one edge of each path P_i, and the whole path in all but precisely $k-1$ cases. In particular, in the primal graph, the tree T corresponds to a tree from which $k-1$ edges can be deleted, to produce a partition that is ε-compatible with the drawing D. □

It remains to prove Lemma 28.

Proof of Lemma 28. Let $k = f(D_\ell)$ and let the paths $P_1,\dots,P_\ell$ be as in the definition of δ-correspondence. Note the lemma is trivial if $k = 1$, so we assume $k > 1$. If we remove all the edges e from Ω_{D,Λ_n} for which $e^* \in T$, the result is a connected spanning subgraph of of Ω_{D,Λ_n} and thus removing any additional $k-1$ edges results in graph with at most k components. Thus, to prove the Lemma, it suffices to show that if we remove all edges e from Ω_{D,Λ_n} for which the corresponding e^* belong to any of the paths P_i, the resulting graph contains k components $C_1,\dots,C_k$ that induce a partition that is ε-compatible with D_ℓ. Indeed, if after removing the edges of these paths, we then remove the whole tree, we are still left with the same connected components $C_1,\dots,C_k$, by (d).

This is easy to show if we know that as a plane graph, the union of $P_1,\dots,P_\ell$ has the same number of faces as D. Indeed, any two points that belong to a common face of D and lie at distance greater than ε from every other face must be joined by a curve which is at distance greater than δ

from every curve in D by condition (A3) as ensured by our choice of ε, δ via Lemma 25, and thus disjoint the union of the paths $P_1, \ldots, P_\ell$, by (c).

To check the number of faces of the union of $P_1, \ldots, P_\ell$, note that for each vertex of p of D, we have from (a) and (b) that the union of the paths $\{P_i\}$ corresponding to curves $\gamma_i \in \Gamma(D)$ incident to p form a tree. By our choice of δ, ε in (A1), (A2), (A3) and (c), we have that all vertices which belong to more than one of the P_i of this tree lie within distance δ of p. Thus we can define a tree τ_p which is a union of one segment from each P_i, in each case containing an endpoint of P_i, such that the whole tree τ_p lies within distance δ of the point p. Considering the union of all of these τ_p as a plane graph, it has $|V(D)|$ connected components and a single face.

Each path $P_1, \ldots, P_\ell$ consists of two end segments which belong to trees τ_p, τ_q, and a middle segment which joins two such trees. Note that no two middle segments intersect, by (c) and the choice of δ. In particular, when we add these middle segments to our drawing one-by-one, at each step we decrease the number of components or increase the number of faces, as in Euler's formula. As the same is true when we add curves γ_i to build D one curve at a time, the two drawings have precisely the same number of faces. □

Corollary 21 now follows with some additional observations about how the parameters of Theorem 20 applies to grids.

Proof of Corollary 21. Let D be the plane graph consisting of a unit rectangle divided vertically into k even pieces. Consider the lattice sequence $(\{\Lambda_i\}, 1/8)$ that has $\Lambda_i = (1/i)(n\mathbb{Z} \times m\mathbb{Z})$. This is a valid lattice sequence by Definition 18.

Note Λ_{mn} is the graph $(1/m)\mathbb{Z} \times (1/n)\mathbb{Z})$, and it is on this graph we focus. In this case, $\Omega_{D,\Lambda_{mn}}$ can be chosen to be exactly the $m \times n$ grid graph. Note that, as $\varepsilon > 1/m > 1/n$, conditions (a), (b), and (c) of Definition 18 hold for Λ_{mn}, indicating that mn is sufficiently large for our purposes.

The proof of Lemma 23 assumes the stronger condition that neighboring vertices are at distance at most $\varepsilon/20$; this holds for Λ_{mn} by the hypothesis $\varepsilon \geq 20/n$. Provided $\varepsilon \leq 1/(3k)$, Lemma 25 holds for ε and any $\delta < \varepsilon$, for example for $\delta = \varepsilon/2$.

Therefore the proof of Theorem 20 applies, not just in the limit of the lattice sequence $(\{\Lambda_i\}, 1/8)$, but already to the lattice Λ_{mn}. The length of all interior curves in D is $k-1$, a constant, implying that for a uniformly random spanning tree of Λ_{mn}, the probability there are $k-1$ edges whose removal results in a forest that is ε-compatible with D only depends on k and ε. □

4.5 Additive Approximate Balance

We now turn to the proof of Theorem 22. We begin by observing several useful facts implied by the definition of a uniform lattice sequence.

Observation 29. *Let $(\{\Lambda_n\}, \rho, R, C_1, C_2)$ be a uniform lattice sequence. There exist positive constants $c_{\min}, c_{\max}$ such that, for all sufficiently large w, h, n, any $\frac{w}{n} \times \frac{h}{n}$ rectangle contains between $c_{\min} \cdot w \cdot h$ and $c_{\max} \cdot w \cdot h$ vertices in Λ_n.*

Proof. Let n be large enough so that the conditions of Definition 19 apply, and let $h, w \geq 4R$.

This rectangle contains at least $\lfloor w/(2R) \rfloor \times \lfloor h/(2R) \rfloor$ disjoint circles of radius R/n placed in an evenly-spaced grid, and each must contain at least C_1 points. Note when $w \geq 4R$, then $\lfloor w/(2R) \rfloor \geq w/(2R) - 1 \geq w/(2R) - w/(4R) = w/(4R)$, and the same is true for h. Therefore the number of points in this rectangle is least $C_1 \lfloor w/(2R) \rfloor \lfloor h/(2R) \rfloor \geq C_1 \left(\frac{w}{4R}\right)\left(\frac{h}{4R}\right)$. Picking $c_{\min} = C_1/16R^2$ ensures the desired lower bound is true.

This rectangle can be completely covered by the union of $\lceil w/R \rceil \times \lceil h/R \rceil$ balls of radius of R/n placed in an evenly-spaced grid. As each ball contains at most C_2 vertices, the number of vertices

in this box is at most $\lceil w/R \rceil \cdot \lceil h/R \rceil \cdot C_2 \leq \left(\frac{2w}{R}\right)\left(\frac{2h}{R}\right) C_2$. Picking $c_{\max} = 4C_2/R^2$ ensures the desired upper bound is true. □

Given $c_{\min}$ and $c_{\max}$, which depend only on the uniform lattice sequence, we define constants

$$c_1 := 1 - \left(\frac{1}{50} \cdot \frac{c_{\min}}{c_{\max}}\right),$$
$$c_2 := \frac{1}{50} \cdot \frac{c_{\min}}{c_{\max}}.$$

Term c_1 is meant to describe the rate at which a sequence of rectangles constructed in the proof are shrinking. Term c_2 is meant to describe the width of the interval through which we must leave each rectangle we consider. While $c_1 = 1 - c_2$, this is simply a consequence of trying to choose constants that are convenient to work with; we will continue to use c_1 and c_2 in accordance with their conceptual meanings, as just described, rather than substituting one for another.

Observation 30. *Let $(\{\Lambda_n\}, \rho, R, C_1, C_2)$ be a uniform lattice sequence, and let $c_{\min}, c_{\max}, c_1, c_2$ be as above. There exist positive constants $R', R''\rho', \rho''$ such that the following hold for all sufficiently large n.*

(B1) Observation 29 holds for any $w, h \geq \frac{(1-c_1)R'}{2}$.

(B2) $R' \geq 50R$.

(B3) $R' \geq \frac{c_1}{1-c_1} R$.

(B4) The distance between any two adjacent vertices in Λ_n is at most $\frac{R'}{2n}(1 - c_1)$.

(B5) A random walk in Λ_n begun from the point in a $\frac{2(1-c_1)R'}{n} \times \frac{(5-c_1)R'}{2n}$ rectangle that is horizontally-centered and at distance $\frac{2R'}{n}$ from the top will first leave the rectangle along the top with probability at least ρ'.

(B6) For any $w \geq 2(1-c_1)R'$ and $h \in [w - 2R, w]$, a random walk in Λ_n begun from the center of a $\frac{w}{n} \times \frac{h}{n}$ rectangle will first leave to a vertex above any given segment of the top of the rectangle of width $\frac{2c_2 w}{n}$ with probability at least ρ'.

(B7) With probability at least ρ'', a random walk in Λ_n^* begun from the center v of a circle of radius at least $\frac{R''}{n}$ first exits the circle to a point v' such that the ray from v to v' passes through the arc of angle $\frac{\pi}{6}$ that is centered at the top of the circle.

Proof. Because Observation 29 holds whenever $w, h \geq 4R$, (B1) holds whenever R' is chosen large enough so that $R' \geq 400Rc_{\max}/c_{\min}$. Choosing R' this large also suffices to imply (B2) and (B3). And (B4) follows immediately from axiom (b) in the definition of a uniform lattice sequence; we only bother to state this weaker bound here because we will use it later.

Note that (B5) is just a claim about random walk in a finite subgraph induced by the vertices in the given rectangle, and so would hold if we knew this finite graph was connected. Although connectivity of this subgraph does not hold in general, we can deduce (B5) by repeated applications of part (c) of the definition of uniform lattice sequence, with $r = R/n$. Property (c) implies that, with positive probability, we exit such such a circle for the first time closer to the top and further from the left side, and also with positive probability, closer to the top and further from the right side. Repeated applications thus give that we exit the given rectangle at the top. (B6) follows from a similar argument, choosing a constant number of circles of radius $\frac{c_2 w}{2n}$.

Likewise, (B7) is another strengthening of property (c) from the definition of a uniform lattice sequence, with a specific region of the circle identified. It is not a direct consequence of property (c) for the same radius $\frac{R}{n}$ and probability ρ, but it is straightforward to get such a guarantee for larger $R'' > R$ by using a small constant number of circles of radius $\frac{R}{n}$. □

Proof of Theorem 22. We show that there is a $1/\text{poly}(n)$ probability that a random spanning tree of $\Omega^S_{\Lambda_n}$ can be cut into pieces differing in size by an additive constant by examining the probability the tree resembles a particular plane partition consisting of the unit square divided into k vertical pieces. For each n, we may use a slightly different vertical partition, chosen to ensure equal numbers of vertices in each region.

Consider $\Omega^S_{\Lambda_n}$. We draw $k-1$ vertical lines $L_1, L_2, \ldots, L_{k-1}$ to partition the unit square into k vertical strips each containing exactly the same number of vertices of $\Omega^S_{\Lambda_n}$, up to an error of ± 1 (which is only necessary because the number of points in $\Omega^S_{\Lambda_n}$ may not be divisible by k). We can find the location of each L_i by sweeping from left to right until we have the correct number of points. (If necessary, we can first perturb the embedding slightly so that no two vertices are vertically aligned.) Observe that there is an absolute bound on how close these lines can be to each other and to the left and right sides of the unit square for all sufficiently large n. This is because Observation 29 implies that if one strip has area $\frac{c_{\max}}{c_{\min}}$ times the area of another one, it definitely contains more vertices. So let $d_0 > 0$ be such that the horizontal distance between all vertical lines is at least $2d_0$. Let $\varepsilon := 4(1-c_1)^2 d_0^2 < d_0$.

We run the ordinary version of Wilson's algorithm (not the modified version described in Section 4.4), initiating the first $k-1$ random walks $W_1, W_2, \ldots, W_{k-1}$ from the bottom of the unit square, lower-bounding the probability each walk W_i makes it to the top while roughly following L_i. Figure 9 depicts a typical walk satisfying these properties, as well as several of the geometric objects described shortly in the proof to aid in the probability analysis.

Each walk W_i starts with a short segment β_i from a dual vertex adjacent to the boundary cycle C^*_n and proceeds inward, reaching a distance of at least $\frac{R}{n} + \frac{R'}{n}$ above the bottom of the unit square while staying within ε of L_i, as depicted in the bottom-left of Figure 9. From the definition of a uniform lattice sequence, it is straightforward to see that such a path must exist, and all such paths must have constant length. Since uniform lattice sequences have at most a constant degree at every vertex, there is a constant probability that W_i takes path β_i.

From there, we next claim that there is a $\frac{1}{\text{poly}(n)}$ probability that W_i reaches a point within $\frac{\varepsilon}{2}$ of the point labeled $\gamma(0)$ in Figure 9, which is the point on L_i of distance $\varepsilon + \frac{R}{n}$ from the bottom of the unit square. We repeatedly apply (B7) to a sequence of circles of exponentially increasing radius. Each circle is centered at the current position of the walk (the bottom three green points in Figure 9) and extends downward to be tangent to the horizontal line which is a distance of $\frac{R}{n}$ to the bottom of the unit square, ensuring that the walk does not hit the boundary cycle C^*_n. At each step, the walk leaves the circle within the top arc of angle $\frac{\pi}{6}$ with probability at least ρ'', thus increasing its vertical distance from the $\frac{R}{n}$ line by a multiplicative factor of $1 + \cos(\pi/12)$ while remaining within the shaded triangular region, which is defined to have bottom angle $\pi/6$ and contain the initial vertex. After

$$s(n) := \frac{\log(n\varepsilon/R'')}{\log(1+\cos(\pi/12))}$$

circles, the elevation will have increased from its initial value of $\frac{R''}{n}$ to ε, at which point it is clear from Figure 9 that the walk must have visited some point within the red circle around $\gamma(0)$ of

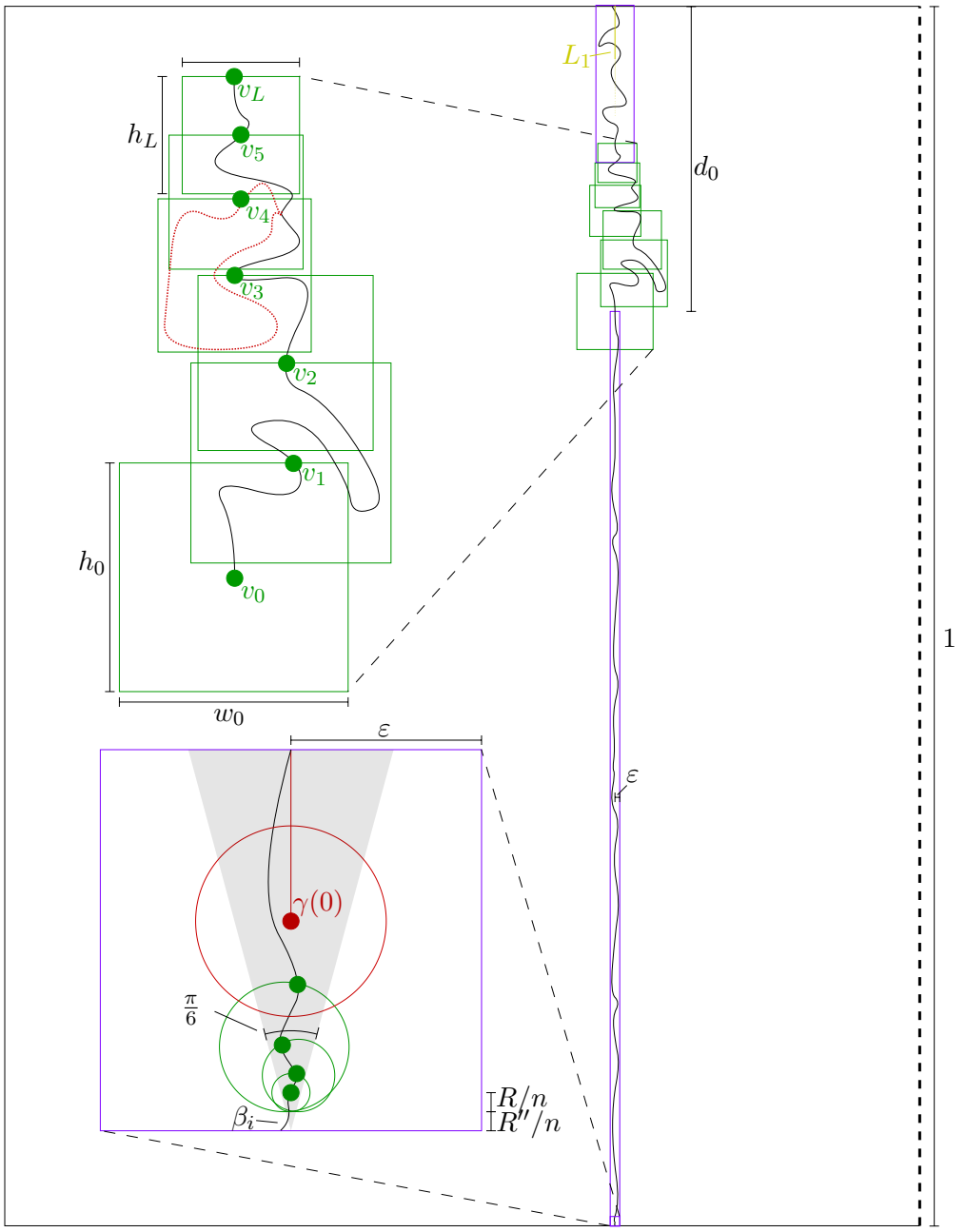


Figure 9: Illustration accompanying the proof of Theorem 22, which gives an inverse polynomial lower bound on the probability of the first $k-1$ walks of Wilson's algorithm follow paths like the windy thin black line on the right. In this diagram, we suppose $k = 2$ and there is roughly uniform density of dual vertices in the graph. Thus, there is only one such path, and it is supposed to split the unit square (which is the boundary of the figure, but also extending beyond the dashed line to the right) in two equal-area pieces. Parts of the path are magnified on the left. The dashed red curve is an erased loop, which we assume was traversed clockwise before continuing on the black path. The figure is drawn to-scale with the following parameters from the proof: $c_1 = \frac{7}{8}$, $c_2 = \frac{1}{8}$, $c_3 = 1.8$, $d_0 = \frac{1}{4}$, $L = 6$, $\varepsilon = \frac{1}{256}$, $\frac{R}{n} = \frac{R'}{n} = \frac{1}{2048}$, $\frac{R'}{n} = \frac{1}{16}$. Aside from c_1 being too small and c_2 being too large, these are all valid values for the parameters.

radius $\frac{\varepsilon}{2}$. This happens with probability at least

$$(\rho'')^{s(n)} = \Omega\left(\frac{1}{\text{poly}(n)}\right).$$

As soon as this happens, we apply Lemma 23 (which is valid as a uniform lattice sequence is indeed a lattice sequence) to the curve beginning at $\gamma(0)$ and proceeding upward to conclude that W_i stays within a radius of ε of L_i until it is at a distance of d_0 from the top of the unit square with constant probability. Let us call the total probability of all of these events occurring $p_0(n)$, which we have shown is $\Omega\left(\frac{1}{\text{poly}(n)}\right)$.

At this point, W_i has reached the vertex v_0 in Figure 9; this is the first vertex encountered in W_i that is of distance at most d_0 from the top. More generally, letting $d_j := (c_1)^j d_0$, for each $j = 1, 2, \ldots$ let v_j be the first vertex encountered in W_i that is of distance at most d_j from the top, assuming W_i makes it that far. At each step, there is a non-self-intersecting path $P_{i,j}$ from the start of W_i to v_j that will define the ultimate partition of the dual tree assuming parts of it are not erased. Let $\text{imb}(j)$ denote the imbalance of the partition of the vertices of $\Omega^S_{\Lambda_n}$ obtained by extending $P_{i,j}$ straight upward until it his the top of the unit square. Here imbalance is defined as the difference between the number of vertices on the left side and the target size of $\frac{i}{k}\left|\Omega^S_{\Lambda_n}\right|$.

Assuming W_i makes it to each vertex v_j, we let S_j be the rectangle centered at v_j with width $w_j := 2(1-c_1)d_j$ and height $h_j := 2(1-c_1)d_j - 2e_j$, where e_j is the difference between d_j and the vertical distance between v_j and the top of the unit square, which is at most d_j. For convenience, we define v_j, w_j, h_j, and S_j for $j = -1, -2$ as well, in the obvious way. Note that v_j and S_j will still be contained within the unit square since we know $d_0 < \frac{1}{2}$. We prove the following claim by induction on j:

Claim 4. *Let $0 \le j \le L$, where L is such that $d_L = (c_1)^L d_0 \ge \frac{R'}{n}$. With probability at least $p_0(n)(\rho')^j$, all of the following events jointly occur:*

(C1) The walk W_i eventually reaches a first vertex v_j that is distance at most d_j from the top.

(C2) Each of $v_0, v_1, \ldots, v_j$ are within

$$\varepsilon'_j := \varepsilon + \left(\frac{1}{4} + c_2\right)(w_0 + w_1 + w_2 + \cdots + w_{j-1})$$

of line L_i.

(C3) Each vertex encountered so far in W_i has been within $\varepsilon'_L + w_0$ of L_i.

(C4)

$$|\text{imb}(j)| \le c_{\max} n^2 \left(c_2 w_{j-1} d_j + w_{j-2} h_{j-2} + w_{j-1}\left(h_{j-1} + \frac{R}{n}\right)\right).$$

(C5) For any vertex v previously visited by W_i within rectangle S_j, the sub-path of $P_{i,j}$ from v to v_j stays within S_{j-1} until reaching v_j.

For the base case ($j = 0$), we have already noted that the walk makes it to v_0 while staying within the more stringent bound of ε with probability at least $p_0(n)$. This satisfies (C1), (C2), and (C3). When this happens, the maximum possible magnitude of imbalance is bounded by the initial imbalance of one plus the number of vertices in the ε-neighborhood of L_i, which is a rectangle of

dimensions $2\varepsilon \times (1 - d_0 + e_0)$. Assuming n is sufficiently large so that Observation 29 applies, we may thus bound the absolute value of the imbalance by

$$\begin{aligned} 1 + c_{\max}(2\varepsilon n)((1 - d_0 + e_0)n) &= 1 + 2c_{\max}n^2 \cdot 4(1 - c_1)^2 d_0^2 \cdot (1 - d_0 + e_0) \\ &\leq 2c_{\max}n^2 w_0 h_0 \leq c_{\max}n^2(w_{-2}h_{-2} + w_{-1}h_{-1}), \end{aligned}$$

which establishes (C4). Finally, (C5) holds by the proof of Lemma 23, that is, the way in which we've assured that our walk is staying within ε of line L_i. This proof uses balls of small radius at most $\varepsilon/2 = 2(1-c_1)^2 d_0^2$ and bounds the probability the walk always makes progress toward a point further along L_i, that is, a point higher up along L_i. If such a path went below the bottom of S_{-1}, it would have had to backtrack through several circles of radius at most $\varepsilon/2$ (which is much smaller than $h_{-1} = c_1^{-1} d_0$) in the wrong direction, and the fact that it does not do so is already captured by our probability $p_0(n)$. This concludes the base case $j = 0$.

For the inductive case, suppose the statement holds for $j - 1$. Suppose the random walk W_i has just reached vertex v_{j-1} in a way that satisfies all five properties (C1-5), which happens with probability at least $p_0(n)(\rho')^{j-1}$. We have from (B6) that, with probability at least ρ', W_i will next leave S_{j-1} to a vertex within horizontal distance $c_2 \frac{w_j}{n}$ of any given target t within the top of S_{j-1}. Since $p_0(n)(\rho')^{j-1} \cdot \rho' = p_0(n)(\rho')^j$, it suffices to show that there is some target t (that is, an x-coordinate) for which this event implies (C1-5) hold after step j.

Consider how $\mathrm{imb}(j)$ changes as we vary t. Recall the imbalance is calculated by extending $P_{i,j}$ straight upward until it hits the top of the unit square. We first focus on the contribution to the change in the imbalance from this extension of $P_{i,j}$ from v_j to the top of the unit rectangle. Specifically, if we change t to be $\frac{1}{4}w_{j-1}$ to the left or right, applying Observation 29 to the rectangle above v_j to the top of the unit square, we know the imbalance will change by at least

$$c_{\min}n^2 \cdot \frac{1}{4} w_{j-1} \left(d_j - 2\frac{R}{n} \right).$$

Here the $-2\frac{R}{n}$ comes from the fact that the height of the precise rectangle we choose is shorter by up to $\frac{R}{n}$ on the bottom because that is how far v_j might be above the top of S_{j-1}, and shorter by $\frac{R}{n}$ on the top because the vertices in Λ_n within the top band of height $\frac{R}{n}$ might contain vertices that are outside of the boundary cycle C^*. Using (B2), we know that

$$d_j \geq d_L \geq \frac{R'}{n} \geq 50\frac{R}{n},$$

so it follows that a lower bound on the imbalance change is given by

$$c_{\min}n^2 \cdot \frac{1}{4} w_{j-1} \left(d_j - 2\frac{R}{n} \right) \geq \frac{1}{4} \cdot \frac{48}{50} \cdot c_{\min}n^2 w_{j-1} d_j = \frac{6}{25} c_{\min}n^2 w_{j-1} d_j.$$

On the other hand we are assuming the previous absolute imbalance is

$$\begin{aligned} |\mathrm{imb}(j-1)| &\leq c_{\max}n^2 \left(c_2 w_{j-2} d_{j-1} + w_{j-3} h_{j-3} + w_{j-2} \left(h_{j-2} + \frac{R}{n} \right) \right) \\ &\leq c_{\max}n^2 \left(c_2 w_{j-2} d_{j-1} + w_{j-3} \cdot 2(1 - c_1) d_{j-3} + w_{j-2} \left(2(1 - c_1) d_{j-2} + \frac{R}{n} \right) \right) \\ &= c_{\max}n^2 \left(\frac{1}{c_1^2} c_2 w_{j-1} d_j + \frac{2(1 - c_1)}{c_1^5} w_{j-1} d_j + \frac{1}{c_1^2} w_{j-1} \left(2(1 - c_1) d_{j-1} + c_1 \frac{R}{n} \right) \right) \end{aligned}$$

From how c_1 was chosen, we know, for example, that $1/c_1^2 \le 1/c_1^5 \le 2$:

$$\le c_{\max} n^2 \left(2c_2 w_{j-1} d_j + 4(1-c_1) w_{j-1} d_j + 2w_{j-1} \left(2(1-c_1) d_j + c_1 \frac{R}{n} \right) \right)$$

Since (B3) implies $c_1 \frac{R}{n} \le (1-c_1)\frac{R'}{1} \le (1-c_1) d_L \le (1-c_1) d_j$, we see this is

$$\begin{aligned}
&\le c_{\max} n^2 \left(2c_2 w_{j-1} d_j + 4(1-c_1) w_{j-1} d_j + 2w_{j-1} \left(3(1-c_1) d_j \right) \right) \\
&= c_{\max} n^2 w_{j-1} d_j \left(2c_2 + 10(1-c_1) \right) \\
&= c_{\max} n^2 w_{j-1} d_j \left(\frac{c_{\min}}{25 c_{\max}} + \frac{c_{\min}}{5 c_{\max}} \right) \qquad \text{(from how } c_1 \text{ and } c_2 \text{ were chosen)} \\
&= \frac{6}{25} c_{\min} n^2 w_{j-1} d_j.
\end{aligned}$$

Thus, there must be some target t in the middle half of the top boundary of S_{j-1} for which the imbalance contribution of vertices within $d_j - e_j$ of the top of the unit square exactly cancels out $\text{imb}(j-1)$ (up to a single vertex, perhaps). In other words, if W_i were to proceed from v_{j-1} to a vertex v_j with x-coordinate exactly at the target t, taking a path that went straight upward to the top of S_{j-1}, then horizontally along the top of the square, then straight up to v_j, the partition would be balanced up to one vertex. However, the path it takes through S_{j-1} can introduce further imbalance of the following magnitudes:

- An imbalance of at most $c_{\max} n^2 c_2 w_{j-1} d_j$ coming from the fact that W_i might exit S_{j-1} a distance of up to $\pm c_2$ slightly to either side of the target t. (If there is a one-vertex imbalance from before, it can be accounted for in this term as well.)
- An imbalance of $c_{\max} n^2 w_{j-1}(h_{j-1} + e_j) \le c_{\max} w_{j-1} \left(h_{j-1} + \frac{R}{n} \right)$ coming from the fact that we can make an arbitrary partition of the vertices in S_{j-1}, where the e_j term accounts for including the vertices above S_{j-1} and below v_j.
- An imbalance of $c_{\max} n^2 w_{j-2} h_{j-2}$ coming from loop erasures that affect the previous rectangle S_{j-2} as well. For example, this happens in Figure 9 when the red dashed loop is erased shortly after the walk visits v_4. Though we do not need to consider any rectangles further back in the sequence: By the inductive hypothesis of (C5), such a loop cannot extend below the bottom of S_{j-2} (which, in Figure 9, is the green rectangle centered at v_3).

Summing these up, we obtain the upper bound on $|\text{imb}(j)|$ from (C4).

The other four properties are easy to check. (C1) is immediate because points above the top of S_{j-1} are within d_j of the top of the unit square. (C2) holds because we exit within a c_2 fraction of w_{j-1} from the target, which needs to be at most a $\frac{1}{4}$ fraction away from the current position, so the total additional drift left or right is at most $(\frac{1}{4} + c_2) w_{j-1}$. (C3) is simply a crude upper-bound on how far to the side any walk could have traveled while staying within the left and right boundaries of the rectangles S_j, the biggest rectangle having width w_0. For (C5), note that S_j does not overlap with S_{j-2}, so any such vertex v must be in $S_{j-1} \setminus S_{j-2}$. Hence, v must have been visited by W_i after v_{j-1}, so the path forward from v must first exit S_{j-1} to v_j. Thus, by induction, Claim 4 holds for all $0 \le j \le L$.

Returning to the proof of the theorem, we pick L so that $\frac{R'}{n} \le d_L \le \frac{2R'}{n}$; let $c_3 \in [1,2]$ be such that $d_L = (c_1)^L d_0 = \frac{c_3 R'}{n}$. Then with probability $p_0(n)(\rho')^L$, W_i will satisfy (C1-5) with $j = L$. From vertex v_L, we know from (B5) that, with probability at least ρ', W_i will continue all the

way to the top of the unit square, hitting C^* before exiting the top of a final $\frac{2(1-c_1)R'}{n} \times \frac{(5-c_1)R'}{2n}$ rectangle, which is shown in purple at the very top of Figure 9.[2] Observe that the imbalance before the final part of the walk starts is at most

$$\begin{aligned}
|\text{imb}(L)| &\leq c_{\max} n^2 \left(c_2 w_{L-1} d_L + w_{L-2} h_{L-2} + w_{L-1} \left(h_{L-1} + \frac{R}{n} \right) \right) \\
&\leq \frac{1}{c_1^4} c_{\max} n^2 \left(c_2 2(1-c_1) d_L^2 + (2(1-c_1))^2 d_L^2 + 2(1-c_1) d_L^2 + 2(1-c_1) d_L \frac{R}{n} \right) \\
&\leq \frac{1}{c_1^4} c_{\max} n^2 \left(c_2 2(1-c_1) \frac{4R'^2}{n^2} + (2(1-c_1))^2 \frac{4R'^2}{n^2} + 2(1-c_1) \frac{4R'^2}{n^2} + 2(1-c_1) \frac{2R'R}{n^2} \right).
\end{aligned}$$

Since the n^2 terms cancel and c_1, c_2, R', and R are all constant, this is a constant. To bound the additional imbalance generated from this final part of the walk, observe that the final walk stays within a rectangle that does not overlap with S_{L-2}. This is because the final walk starts at least a distance of $\frac{h_{L-1}}{2}$ above the top of S_{L-2} and the distance from the start of the final walk to the bottom of the final rectangle is

$$\frac{(5-c_1)R'}{2n} - \frac{2R'}{n} = \frac{R'}{n}(1-c_1) - \frac{R'}{2n}(1-c_1) \leq \frac{R'}{n}(1-c_1) - e_{L-1} \leq d_{L-1}(1-c_1) - e_{L-1} = \frac{h_{L-1}}{2},$$

where the first inequality follows from (B4). Thus, we may bound the imbalance by applying (B1) to just the final two rectangles. The imbalance is at most

$$c_{\max} n^2 w_{L-1} h_{L-1} + c_{\max} \cdot (5-c_1)(1-c_1) R'^2 \leq c_{\max} n^2 \left(\frac{2(1-c_1)}{c_1} \cdot \frac{2R'}{n} \right)^2 + c_{\max} \cdot (5-c_1)(1-c_1) R'^2,$$

which is also a constant. Summing these imbalances, we know that the total imbalance is at most a constant, not depending on n. Also, applying (C2) and (C3), the maximum horizontal deviation of W_i from L_i is at most

$$\begin{aligned}
\varepsilon'_L + w_0 &\leq \varepsilon + \left(\frac{1}{4} + c_2 \right) \left(\sum_{j=0}^{\infty} (c_1)^j w_0 \right) + w_0 \\
&= 4(1-c_1)^2 d_0^2 + \left(\frac{1}{4} + c_2 \right) \left(\frac{w_0}{1-c_1} \right) + w_0 \\
&= 4(1-c_1)^2 d_0^2 + \left(\frac{1}{4} + c_2 \right) (2d_0) + 2(1-c_1) d_0 \\
&\leq d_0 \left(\frac{4}{50^2} + \frac{1}{2} + \frac{2}{50} + \frac{2}{50} \right) \\
&\leq d_0,
\end{aligned}$$

so by the way we chose d_0, no walk will hit the left or right sides of the unit square, and no two neighboring walks will collide with each other.

[2]Note that this part of the figure is a bit misleading because it appears rather large in the figure, wider than the longer purple rectangle at the bottom. This is just due to the specific parameters we chose to make the figure readable. As n goes to infinity, the top rectangle will actually shrink down to a point (which is necessary to get an additive constant imbalance!) while the bottom rectangle will remain relatively the same size and shape.

We have shown that all $k-1$ walks will make it to the top of the unit square and have constant imbalance with probability at least

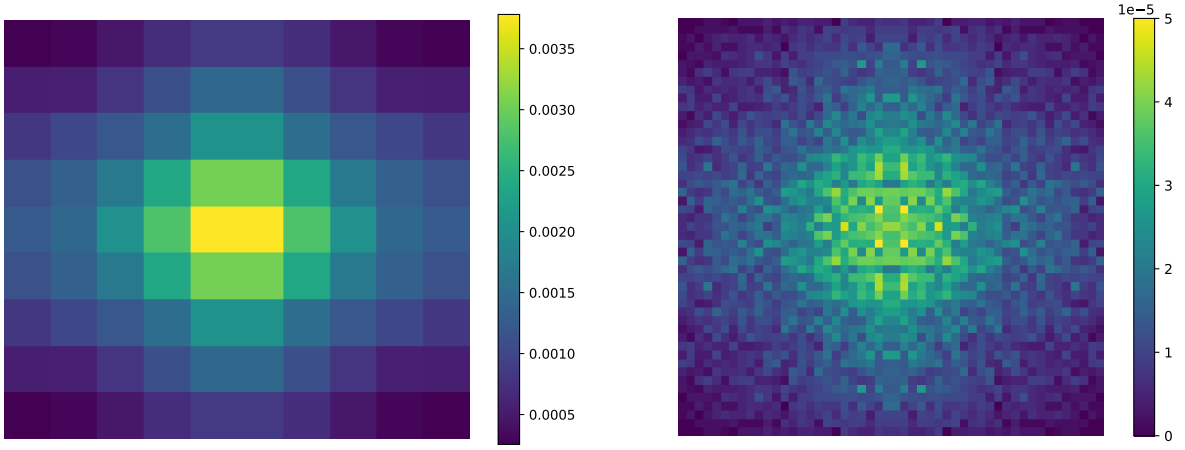
$$\left(p_0(n)(\rho')^{L+1}\right)^{k-1} = \left(p_0(n)\rho'\left(\frac{c_3 R'}{d_0 n}\right)^{(\log(\rho')/\log(c_1))}\right)^{k-1} = \Omega\left(\frac{1}{p_0(n)^{k-1}n^q}\right),$$

where

$$q = \frac{(k-1)\log(\rho')}{\log(c_1)}.$$

Note that this is $\frac{1}{\text{poly}(n)}$ since $p_0(n)$ is a polynomial function of n and q is a constant that does not depend on n. When this happens, the primal tree must be splittable into k pieces that are balanced up to an additive constant. □

5 Experiments



(a) 10×10 grid.

(b) 50×50 grid.

Figure 10: Heatmap of likelihoods that each vertical edge of the 10×10 and 50×50 grid graph is an exactly balanced split edge of a uniformly random spanning tree. Each pixel represents a vertical edge. Symmetry is enforced manually; for instance, the center two vertical edges of the 10×10 grid are exactly the same color not by coincidence, but because the experiment was only run for one of them, with both edges receiving the same color. 1,000,000 trials were run for each symmetry class of edges in each graph. In the 10×10 grid, the number of those trials yielding exact balance ranged from 255 (in the corners) to 3,781 (in the center). In the 50×50 grid, they ranged from 0 to 50. In particular, this plot is still noisy at this number of trials.

We visualize some of the results in this paper by running computational experiments with $k = 2$ on 10×10, 50×50, and 100×100 grid graphs. By running the first random walks of Wilson's algorithm on the dual grid, we empirically estimate the probability of various edges being contained in a random spanning tree and splitting it into two balanced pieces. The only concrete lower bounds we give on these probabilities is for central edges of the grid (in Lemma 12). However, Figure 10 suggests that, even though central edges have the highest probability, similar bounds, possibly worse by another factor of n, should apply for other edges as well.

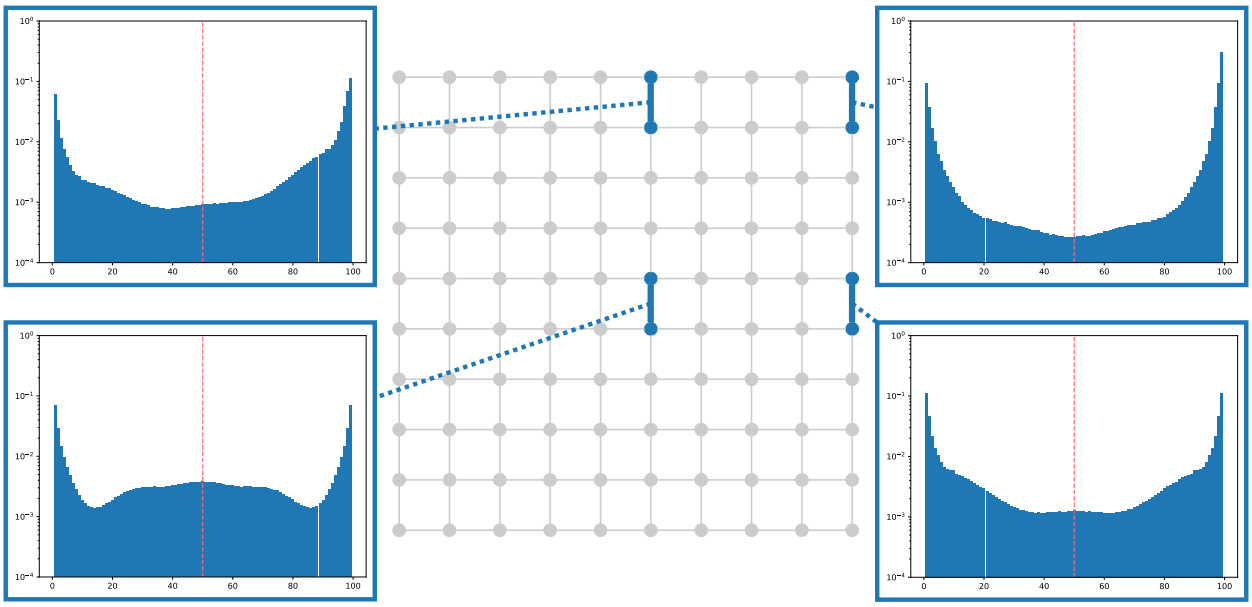


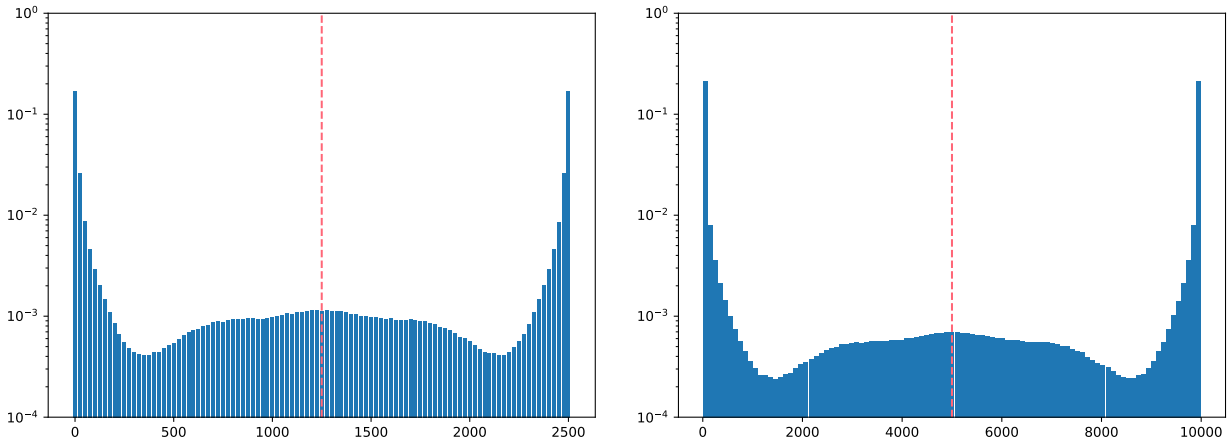
Figure 11: Distributions over component sizes after removing an edge from a random spanning tree for four specific edges in the 10×10 grid. The x-axis represents the size of the component below the edge, which can be any number between 1 and 99. Each bar represents a single value, with the one under the dashed red line representing exact balance. All plots are on the same logarithmic scale, with probability on the y-axis. Note that the probabilities sum to strictly less than one because sometimes the edge is not even included in the tree. The distribution is asymmetric for the two edges in the top row and symmetric for the edges in the middle row.

When a given edge is in the tree but does not give a balanced split, how imbalanced is it? Figure 11 shows the empirical distribution over component sizes for various edges in the 10×10 grid. As one can see, the most common component sizes are heavily imbalanced. Curiously though, at least for central edges on the 10×10 grid, being exactly balanced is more likely than being moderately balanced. Figure 12 shows that this trend holds for 50×50 and 100×100 grids as well.

Comparing these figures also illustrates the main point of Corollary 21. In each histogram, summing the area of a constant number of bars near the dashed line yields the probability that the parts are approximately balanced. As one can see, this probability does not appear to be vanishing as we increase the grid resolution; it should be lower-bounded by a constant.

Acknowledgements

We thank Jacob Calvert for pointing out Lemma 14 could be proved more simply using the Cauchy-Schwarz inequality. This material is based upon work supported by the National Science Foundation under Grant No. DMS-1928930 and by the Alfred P. Sloan Foundation under grant G-2021-16778, while the authors were in residence at the Simons Laufer Mathematical Sciences Institute (formerly MSRI) in Berkeley, California, during the Fall 2023 semester. S. Cannon is supported in part by NSF CCF-2104795. W. Pegden is supported in part by NSF DMS-2054503. J. Tucker-Foltz is supported in part by a Google PhD Fellowship.



(a) 50×50 grid, bin size 25.

(b) 100×100 grid, bin size 101.

Figure 12: Distribution over component sizes for the central edges in the 50×50 and 100×100 grids, similar to the bottom-left histogram in Figure 11. In the 50×50 grid, each bar (except the ones on the very end) represents 25 possible values, and in the 100×100 grid they represent 101 possible values. These plots are on the same logarithmic scales and have the same numbers of bars as in Figure 11.

Appendix

This section includes proofs omitted from the main body of the paper.

To prove Proposition 7, we will use the following propositions:

Proposition 31. *For vertices u, v in a graph G with M edges, the commute time $\tau_u(v) + \tau_v(u)$ between u and v is equal to $2M\mathcal{R}_{u,v}$, where $\mathcal{R}_{u,v}$ denotes the effective resistance between u and v.* □

The above is proved, e.g., in [9]. We will use the following bound on the effective resistance of the grid, which can be found, e.g., in Proposition 9.16 in [24]:

Proposition 32. *The effective resistance between opposite corners of a $k \times k$ grid graph is at most $2 \log k$.* □

Finally we recall the following monotonicity principles for effective resistance:

Proposition 33. *If G is a graph and G' is obtained from G by adding edges and/or gluing vertices, the effective resistance between any pair of vertices in G' is at most the effective resistance between the corresponding pair of vertices in G.* □

Proof of Proposition 7. Suppose the G is an $m \times n$ grid graph, with $n \leq m$. By Propositions 6 and 31, it suffices that the effective resistance between any vertex v in the dual G^* and the root r corresponding to the outer face of G satisfies $\mathcal{R}_{v,r} \leq 2 \log n$. We show this in the following way. Let $\tilde{G}$ be an $(m+1) \times (n+1)$ grid graph. Observe that G^* can be obtained from $\tilde{G}$ by gluing all its boundary vertices (i.e., those incident with its outer face in the canonical drawing) to form a single vertex, which is then the root r of G^*. By Proposition 33, for any vertex $v \in G^*$, an upper bound on the effective resistance between v and r in G^* can be obtained by considering the

effective resistance between v and any boundary vertex in $\tilde{G}$. To this end, simply consider a largest square subgrid S of $\tilde{G}$ which has v as its corner vertex. The opposite corner u will necessarily be a boundary vertex of $\tilde{G}$. Even just in S, the effective resistance between v and u is at most $2\log n$, and thus by Proposition 33, it is at most $2\log n$ in $\tilde{G}$ as well, proving the proposition. □

Proof of Lemma 8. Note every k-forest of G can be obtained by removing $k-1$ edges from some spanning tree of G. For each spanning tree, the number of ways in which to do this is exactly $\binom{N-1}{k-1}$. This means there are at most $\binom{N-1}{k-1} \cdot \mathrm{sp}(G)$ k-forests of G.

If the probability a uniformly random spanning tree of G is k-splittable is at least α, this means there are at least $\alpha \cdot \mathrm{sp}(G)$ splittable spanning trees of G. For a splittable spanning tree, exactly one of the ways of removing $k-1$ edge produces a balanced k-forest. On the other hand, each forest can be obtained from a spanning tree in at most

$$\binom{M-(N-1)}{k-1} \leq (M-N+1)^{k-1}$$

ways, since this is a bound for the number of choices for the additional $k-1$ edges which belong to a spanning tree which contains the forest. This means there are at least $\alpha \cdot \mathrm{sp}(G)/(M-N+1)^{k-1}$ balanced k-forests of G.

The probability a uniformly random k-forest is balanced is the number of balanced k-forests divided by the total number of k-forests. This is at least

$$\frac{\alpha \cdot \mathrm{sp}(G)/(M-N+1)^{k-1}}{\binom{N-1}{k-1}\mathrm{sp}(G)} \geq \frac{\alpha}{N^{k-1}(M-N+1)^{k-1}}.$$ □

Proof of Theorem 9. We will refer to each time we begin again at Step 1 as a *round* of this algorithm. Note in Step 2, if there exist $k-1$ edges whose removal disconnects T into k components of equal size, this collection of edges is unique; there are never two distinct ways to divide a tree into k components of equal size.

Correctness: Let P be a particular balanced k-partition of G. Let $P_1, \ldots, P_k$ be the districts of P. In any given round of this algorithm, the probability we sample a spanning tree that can be divided to produce exactly P is

$$\frac{\mathrm{sp}(G/P)\prod_{i=1}^{k}\mathrm{sp}(P_i)}{\mathrm{sp}(G)}.$$

We then accept this sample with probability $1/s = 1/\mathrm{sp}(G/P)$. In all, the probability we sample partition P is proportional to $\prod_{i=1}^{k}\mathrm{sp}(P_i)$, meaning we are sampling exactly from the spanning tree distribution.

Runtime: In the interest of runtime, we can implement the first step by running Wilson's algorithm on the dual graph, with the vertex corresponding to the outer face of the grid as the root. In this way, Wilson's algorithm takes $O(N\log N)$ steps to produce a uniformly random spanning tree of G. The resulting uniformly random spanning tree of the dual graph can be converted to a uniformly random spanning tree of the grid graph in linear time, via the bijection in Lemma 5.

Step 2 can be easily implemented using a greedy approach and depth-first search, which runs in time $O(N)$. Creating G/P takes at most $O(N)$ time, and computing its spanning tree count is constant time whenever k is a constant, as G/P has only k vertices. This means the runtime of each round of this algorithm is $O(N\log N)$ in expectation.

The expected number of rounds until a random spanning tree of an N-vertex grid graph G is found to be k-splittable is $O(N^{2k-2})$ by our Theorem 15 when k is constant. Similarly as in the proof of Lemma 8, a trivial upper bound on s for any G/P is $\binom{2N}{k-1} = O(N^{k-1})$, as there are at most $2N$ edges in G and $k-1$ are in any spanning tree of G/P. This means the expected number of rounds we must do until a k-splittable partition is returned is $O(N^{3k-3})$ in expectation. Altogether, this gives an expected running time of $O(N^{3k-2} \log N)$. □

Proof of Theorem 10. We will refer to each time we begin again at Step 1 as a *round* of this algorithm.

Correctness: The up-down walk is known to sample approximately uniformly from the k-forests of G, and walking for longer than its mixing time ensures this is the case. The probability a given partition P is returned is proportional to the number of k-forests whose connected components are the districts of P, which is exactly the spanning tree distribution.

Runtime: The mixing time of the up-down walk on a graph with N vertices is $O(N \log N)$, and each step can be implemented $O(\log N)$ amortized time [10, 1]. It then takes $O(N)$ steps to check if the sampled k-forest is balanced. Thus each round takes total time $O(N \log^2 N)$. By Theorem 2, it takes $O(N^{4k-4})$ rounds in expectation to see a balanced k-forest. Thus the total expected running time is $O(N^{4k-3} \log^2 N)$. □

Proof of Lemma 25. Let ε_1 be the minimum distance between any pair of points in the finite set V. For a curve γ, define a_γ, b_γ and then $\bar{\gamma}$ as in the statement of (A2), but with ε_1 in place of ε. The images of all the curves $\bar{\gamma}$ are compact sets, they are disjoint, and there are finitely many of them, so there is some minimum positive distance ε_2 between them.

We let $\varepsilon = \min(\varepsilon_1, \varepsilon_2)/3$, and choose δ as follows. By uniform continuity of the curves, for any $\gamma \in \Gamma(D)$ there is a δ_γ such that for any s, t, we have that $|s - t| < \delta_\gamma \implies d(\gamma(s) - \gamma(t)) < \varepsilon$. Choose δ_1 to be the minimum of the δ_γ, over all $\gamma \in \Gamma$.

Now, given a compact set $K \subseteq \mathbb{R}^2$, consider the function $f_K : [0,1]^2 \times [0,1]^2 \to \mathbb{R}$ where $f_K(x, y)$ is defined to be the supremum, over all curves from x to y, of the distance between the curve and K:

$$f_K(x, y) = \sup_{\substack{\gamma:\gamma(0)=x,\gamma(1)=y \\ \mathrm{im}(\gamma)\cap K=\varnothing}} d(\gamma, K) \quad \text{for} \quad x, y \notin \mathrm{im}(\gamma). \tag{16}$$

Note that the surpremum is over the infinite family of all curves from x to y, not over the finite family of curves that are the subject of the Lemma. (Note that if x and y lie in different connected components of $\mathbb{R}^2 \setminus K$, we have that $f_K(x, y) = 0$.) f is easily seen to be a continuous function of x and y; for example, a curve from x to y can be linearly extended to nearby alternative endpoints x' and y' when they are closer to x, y than K is.

Observe that if $x, y \notin K$, and x, y belong to the same connected component of the open set $\mathbb{R}^2 \setminus K$, there is a curve γ from x to y that is disjoint from K, since open connected subsets of Euclidean space are path connected. Its image is compact and thus has positive distance from K. This shows that when x, y belong to a common connected component of $\mathbb{R}^2 \setminus K$, $f_K(x, y)$ is positive.

Now fix any inner face ϕ of D and consider the case where $K = K_{\phi,\varepsilon}$ is the compact set consisting of all points in ϕ which are at distance $\geq \varepsilon$ from the boundary of the face. By uniform continuity, there is a minimum value $\delta_{\phi,\varepsilon}$ of f_K over all points $x, y \in K$. We can then define δ to be, say, half the minimum of $\delta_{\phi,\varepsilon}$ over the finitely many choices for the inner face ϕ. □

References

[1] Nima Anari, Kuikui Liu, Shayan Oveis Gharan, Cynthia Vinzant, and Thuy-Duong Vuong. Log-concave polynomials IV: Approximate exchange, tight mixing times, and near-optimal sampling of forests. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2021, page 408–420, New York, NY, USA, 2021. Association for Computing Machinery.

[2] Eric Autry, Daniel Carter, Gregory Herschlag, Zach Hunter, and Jonathan C. Mattingly. Metropolized forest recombination for Monte Carlo sampling of graph partitions. *SIAM Journal on Applied Mathematics*, 83(4):1366–1391, 2023.

[3] Eric A. Autry, Daniel Carter, Gregory Herschlag, Zach Hunter, and Jonathan C. Mattingly. Metropolized multiscale forest recombination for redistricting. *Multiscale Modeling & Simulation*, 19(4):1885–1914, 2021.

[4] Amariah Becker, Daryl R. DeFord, Dara Gold, Sam Hirsch, Mary E. Marshall, and Jessica Ring Amunson. Brief of Computational Redistricting Experts as *amici curiae* in support of Appellees and Respondents. Merrill v. Milligan; Merrill v. Caster; United States Supreme Court, 2022. Nos. 21-1086, 21-1087, 2022.

[5] Amariah Becker, Moon Duchin, Dara Gold, and Sam Hirsch. Computational redistricting and the Voting Rights Act. *Election Law Journal: Rules, Politics, and Policy*, 20(4):407–441, 2021.

[6] Gerdus Benadè, Ruth Buck, Moon Duchin, Dara Gold, and Thomas Weighill. Ranked choice voting and proportional representation. Working Paper. Available at `https://ssrn.com/abstract=3778021`, 2021.

[7] Sophia Caldera, Daryl DeFord, Moon Duchin, Samuel C. Gutekunst, and Cara Nix. Mathematics of nested districts: The case of Alaska. *Statistics and Public Policy*, pages 1–22, 2020.

[8] Sarah Cannon, Moon Duchin, Dana Randall, and Parker Rule. Spanning tree methods for sampling graph partitions. Preprint. Available at `https://arxiv.org/abs/2210.01401`, 2022.

[9] Ashok K Chandra, Prabhakar Raghavan, Walter L Ruzzo, and Roman Smolensky. The electrical resistance of a graph captures its commute and cover times. In *Proceedings of the twenty-first annual ACM symposium on Theory of computing*, pages 574–586, 1989.

[10] Moses Charikar, Paul Liu, Tianyu Liu, and Thuy-Duong Vuong. On the complexity of sampling redistricting plans. Preprint. Available at `https://arxiv.org/abs/2206.04883`.

[11] Jowei Chen, Christopher S. Elmendorf, Ruth Greenwood, Theresa J. Lee, Nicholas O. Stephanolpoulos, and Christopher S. Warshaw. Brief of *amici curiae* Professors Jowei Chen, Christopher S. Elmendorf, Nicholas O. Stephanolpoulos, and Christopher S. Warshaw in support of Appellees/Respondents. Merrill v. Milligan; Merrill v. Caster; United States Supreme Court, 2022. Nos. 21-1086, 21-1087, 2022.

[12] Jowei Chen and Nicholas O. Stephanopoulos. The race-blind future of voting rights. *The Yale Law Journal*, 130(4), 2021.

[13] Jeanne N. Clelland, Nicholas Bossenbroek, Thomas Heckmaster, Adam Nelson, Peter Rock, and Jade VanAusdall. Compactness statistics for spanning tree recombination. Preprint. Available at `https://arxiv.org/abs/2103.02699`, 2021.

[14] Aloni Cohen, Moon Duchin, JN Matthews, and Bhushan Suwal. Census TopDown: The Impacts of Differential Privacy on Redistricting. In Katrina Ligett and Swati Gupta, editors, *2nd Symposium on Foundations of Responsible Computing (FORC 2021)*, volume 192 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 5:1–5:22, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[15] Mary Cryan, Heng Guo, and Giorgos Mousa. Modified log-Sobolev inequalities for strongly log-concave distributions. *The Annals of Probability*, 49(1):506 – 525, 2021.

[16] Daryl DeFord, Moon Duchin, and Justin Solomon. A computational approach to measuring vote elasticity and competitiveness. *Statistics and Public Policy*, 7(1):69–86, 2020.

[17] Daryl DeFord, Moon Duchin, and Justin Solomon. Recombination: A Family of Markov Chains for Redistricting. *Harvard Data Science Review*, 3(1), 2021.

[18] Moon Duchin, Jeanne Clelland, Daryl DeFord, Jordan Ellenberg, Tyler Jarvis, Nestor Guillen, Dmitry Morozov, Elchanan Mossel, Dana Randall, Justin Solomon, Ari Stern, Guy-Uriel Charles, Luis Fuentes-Rohwer, Anna Dorman, Dana Paikowsky, and Robin Tholin. Amicus brief of mathematicians, law professors, and students in support of appellees and affirmance. Rucho v. Common Cause, United States Supreme Court, 2019. Available at `https://mggg.org/SCOTUS-MathBrief.pdf`, 2019.

[19] Moon Duchin and Douglas M. Spencer. Models, Race, and the Law. *The Yale Law Journal*, 130:744–797, 2021.

[20] 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.

[21] Alan Frieze and Wesley Pegden. Subexponential mixing for partition chains on grid-like graphs. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 3317–3329, 2023.

[22] 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):452–479, 2020.

[23] Christopher T. Kenny, Shiro Kuriwaki, Cory McCartan, Evan T. R. Rosenman, Tyler Simko, and Kosuke Imai. The use of differential privacy for census data and its impact on redistricting: The case of the 2020 U.S. Census. *Science Advances*, 7(41):eabk3283, 2021.

[24] David A Levin and Yuval Peres. *Markov chains and mixing times*. American Mathematical Society, second edition edition, 2017.

[25] Jonathan C. Mattingly. Expert report on the North Carolina State Legislature and Congressional redistricting (corrected version). Harper v. Hall, North Carolina Superior Court, Wake County, No. 21-cvs-500085, 2021.

[26] Cory McCartan and Kosuke Imai. Sequential Monte Carlo for sampling balanced and compact redistricting plans. Submitted. Available at `https://arxiv.org/abs/2008.06131`.

[27] Ariel D. Procaccia and Jamie Tucker-Foltz. Compact redistricting plans have many spanning trees. *ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2022.

[28] Arnaud Rousselle. Quenched invariance principle for random walks on Delaunay triangulations. *Electron. J. Probab*, 20(33):1–32, 2015.

[29] Kristopher Tapp. Spanning tree bounds for grid graphs. *arXiv preprint arXiv:2109.05987*, 2021.

[30] Jamie Tucker-Foltz. Locked polyomino tilings. Preprint. Available at `https://arxiv.org/abs/2307.15996`, 2023.

[31] David Bruce Wilson. Generating random spanning trees more quickly than the cover time. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, STOC, page 296–303, 1996.

[32] Zhanzhan Zhao, Cyrus Hettle, Swati Gupta, Jonathan Christopher Mattingly, Dana Randall, and Gregory Joseph Herschlag. Mathematically quantifying non-responsiveness of the 2021 Georgia Congressional districting plan. In *Equity and Access in Algorithms, Mechanisms, and Optimization*, EAAMO '22. Association for Computing Machinery, 2022.