UNIVERSITY OF CALIFORNIA,
IRVINE

Flow-Based Decomposition for Geometric and Combinatorial Markov Chain Mixing

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Computer Science

by

Daniel Frishberg

Dissertation Committee:
Distinguished Professor David Eppstein, Chair
Distinguished Professor Michael Goodrich
Associate Professor Milena Mihail

2023

DEDICATION

I dedicate this dissertation to the world's cats, including Marvin and Tammi, a continuing source of companionship and joy, and Marmalade, Hopi, Wave, and Orion, whom I remember often.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

Much of this dissertation, namely the results in Chapter 2 and Appendix A, is reproduced, with modifications, from a work that will appear in the Proceedings of the 50th International Colloquium on Automata, Languages and Programming (ICALP), 2023, under the title "Improved mixing for the convex polygon triangulation flip walk." These results are also on arXiv [26]. The material in Chapter 3 and Appendix B is on arXiv [25]. All of this material is licensed under Creative Commons CC BY 4.0. David Eppstein, the co-author listed in this publication, directed and supervised research which forms the basis for the dissertation.

I acknowledge helpful conversations on this topic with Hadi Khodabandeh, Milena Mihail, Ioannis Panageas, Eric Vigoda, Charlie Carlson, Prasad Tetali, Vedat Alev, Michail Sarantis, Zongchen Chen, Alexandre Stauffer, Lionel Pournin, Jean Cardinal, Karthik Gajulapalli, and Pedro Matias. I thank Milena Mihail, Ioannis Panageas, Eric Vigoda, and Prasad Tetali for further nuturing my interest in the topic of mixing times. I thank the colleagues, labmates, and students who formed the intellectual community that made the journey of graduate school both stimulating and fun. Special thanks are due to Nil Mamano, an early guide who helped me build confidence as a researcher during my first year of graduate school.

David Eppstein pointed me to geometric graph connectivity problems, taught me some of the machinery used in this setting, and helped me develop my interest in rapid mixing, providing me freedom to explore my interests, crucial encouragement, and permission to ruminate. My mentor at the University of Maryland, College Park, Bill Gasarch, helped me build an appreciation of the research landscape. My teaching faculty mentor, Michael Shindler, has helped me find my way as a theory educator, sharing pedagogical strategies and materials, giving me feedback on my teaching, and encouraging me to develop my passion for teaching. Danny Mann has also been a crucial teaching mentor, helping me develop my teaching style and introducing me to an amazing pedagogical community. Michael Goodrich gave me an early opportunity to spread my wings as a teaching assistant for a thoroughly enjoyable theory of computation class. Erik Sudderth taught me the joy of graphical models.

I thank the many teachers and mentors who have kindled my love of math, algorithms, research, pedagogy, and other interests over the years: Dalia Anik, Dave Weaver, Ann Maffrey, Danielle Bostick, Ms. Ashbury, Anthony Muehlberger, Philip Bagley, Ellis Wisner, Richard Salter, Bob Geitz, Sarah Manchester, Roseanne Jones, Larry March, DeSales Harrison, Karen Rosenbaum, and many others. I thank the many friends who have offered encouragement along the way. I thank my supportive colleagues at my workplace prior to graduate school—especially Mark, Eddie, Aaron, and Amit—whose encouragement helped me navigate a big career change. My roommate, Will Gallagher, cooked many meals for me as I was studying for the GRE. I thank my parents, Dina and David, who taught me curiosity and scientific skepticism, recognized and nurtured my interests in math and computation at an early age—and who continue to be a source of wisdom and support. My brother, Nathan, and my sister-in-law, Madison, have been incredible in their encouragement. I thank my granny Sylvia for nurturing my out-of-bounds Sagittarian moon, my Bubbe Catherine for her example of moral courage, and my Zayde Emanuel for modeling kindness and generosity.

VITA

Daniel Frishberg

EDUCATION

Doctor of Philosophy in Computer Science **2023**
University of California, Irvine *Irvine, California*

Master of Science in Computer Science **2020**
University of California, Irvine *Irvine, California*

Bachelor of Arts in Computer Science **2007**
Oberlin College *Oberlin, Ohio*

RESEARCH EXPERIENCE

Graduate Student Researcher **2019**
University of California, Irvine *Irvine, California*

TEACHING EXPERIENCE

Teaching Assistant **2018-2023**
University of California, Irvine *Irvine, California*

Instructor of Record **2022**
University of California, Irvine *Irvine, California*

REFEREED JOURNAL PUBLICATIONS

David Eppstein, Daniel Frishberg, and Martha C. Osegueda. Angles of arc-polygons and Lombardi drawings of cacti. **2023**
Computational Geometry

David Eppstein, Daniel Frishberg, and William Maxwell. On the treewidth of Hanoi graphs. **2022**
Theoretical Computer Science

REFEREED CONFERENCE PUBLICATIONS

David Eppstein and Daniel Frishberg. Improved mixing for the convex polygon triangulation flip walk. **2023**
Will appear, Proc. 50th International Colloquium on Automata, Languages and Programming (ICALP).

David Eppstein, Daniel Frishberg, and Martha C. Osegueda. Angles of arc-polygons and Lombardi drawings of cacti. **2021**
Proc. 33rd Canadian Conference on Computational Geometry (CCCG).

David Eppstein, Daniel Frishberg, and William Maxwell. On the treewidth of Hanoi graphs. **2020**
Proc. 10th International Conference on Fun with Algorithms (FUN).

Simplifying activity-on-edge graphs. **2020**
Proc. 17th Scandinavian Symposium and Workshops on Algorithm Theory (SWAT).

New applications of nearest-neighbor chains: Euclidean TSP and Motorcycle Graphs. N. Mamano, A. Efrat, D. Eppstein, D. Frishberg, M. T. Goodrich, S.G. Kobourov, P. Matias, V. Polishchuk. **2019**
Proc. 30th International Symposium on Algorithms and Computation (ISAAC).

PREPRINTS

Charlie Carlson, Daniel Frishberg, and Eric Vigoda. Improved distributed algorithms for random colorings. **2023**
Submitted for publication.

David Eppstein and Daniel Frishberg. Rapid mixing of the hardcore Glauber dynamics and other Markov chains in bounded-treewidth graphs. **2021**
arXiv preprint.

ABSTRACT OF THE DISSERTATION

Flow-Based Decomposition for Geometric and Combinatorial Markov Chain Mixing

By

Daniel Frishberg

Candidate for Doctor of Philosophy in Computer Science

University of California, Irvine, 2023

David Eppstein, Chair

We prove that the well-studied *triangulation flip walk* on a convex point set mixes in time $O(n^3 \log^3 n)$, the first progress since McShine and Tetali's $O(n^5 \log n)$ bound in 1997. In the process we give lower and upper bounds of respectively $\Omega(1/(\sqrt{n} \log n))$ and $O(1/\sqrt{n})$ —asymptotically tight up to an $O(\log n)$ factor—for the *expansion* of the *associahedron* graph K_n—the first $o(1)$ expansion result for this graph. We show quasipolynomial mixing for the k-*angulation* flip walk on a convex point set, for fixed $k \geq 4$, and a *treewidth* result for the flip graph on $n \times n$ lattice triangulations.

We show that the *hardcore Glauber dynamics*—a random walk on the independent sets of an input graph—mixes rapidly in graphs of bounded treewdith for all fixed values of the standard parameter $\lambda > 0$, giving a simple alternative to existing sampling algorithms for these structures. We also show rapid mixing for analogous Markov chains on dominating sets and b-edge covers (for fixed $b \geq 1$ and $\lambda > 0$) in bounded-treewidth graphs, and for Markov chains on the b-matchings (for fixed $b \geq 1$ and $\lambda > 0$), the maximal independent sets, and the maximal b-matchings of a graph (for fixed $b \geq 1$), in graphs of bounded carving width.

To obtain these results, we introduce a decomposition framework for showing rapid Markov chain mixing. This framework is a purely combinatorial analogue that in some settings gives better results than the *projection-restriction* technique of Jerrum, Son, Tetali, and Vigoda.

Chapter 1

Introduction

The study of *mixing times*—the art and science of proving upper and lower bounds on the efficiency of Markov chain Monte Carlo sampling methods—is a well-established area of research, of interest for graph-theoretic sampling problems, spin systems in statistical physics, probability, and the study of subset systems. Work in this area brings together techniques from spectral graph theory, combinatorics, and probability, and dates back decades; for a comprehensive survey of classic methods, results, and open questions see the canonical text by Levin, Wilmer, and Peres [48]. Recent breakthroughs [2, 3, 4, 16, 17, 18, 43, 47]—incorporating techniques from the theory of abstract simplicial complexes—have led to a recent slew of results for the mixing times of graph-theoretic chains for sampling independent sets, matchings, Ising model configurations, and a number of other structures in graphs, injecting renewed energy into an already active area.

1.1 k-angulations

In Chapter 2, we focus on a class of *geometric* sampling problems that has received considerable attention from the counting and sampling [1, 41] and mixing time [54, 56, 71, 13] research communities over the last few decades, but for which tight bounds have been elusive: sampling *triangulations.* A triangulation is a maximal set of non-crossing edges connecting pairs of points (see Figure 2.1) in a given n-point set. Every pair of triangles sharing an edge forms a quadrilateral. A triangulation *flip* consists of removing such an edge, and replacing it with the only other possible diagonal within the same quadrilateral. Flips give a natural Markov chain (the *flip walk*): one selects a uniformly random diagonal from a given triangulation and (if possible) flips the diagonal.

McShine and Tetali gave a classic result in a 1997 paper [54], showing that in the special case of a convex two-dimensional point set (a convex n-gon), the flip walk *mixes* (converges to approximately uniform) in time $O(n^5 \log n)$, improving on the best-known prior (and first polynomial) upper bound, $O(n^{25})$, by Molloy, Reed, and Steiger [56]. McShine and Tetali applied a Markov chain comparison technique due to Diaconis and Saloff-Coste [21] and to Randall and Tetali [63] to obtain their bound, using a bijection between triangulations and a structure known as *Dyck paths.* They noted that they could not improve on this bound using this bijection. Furthermore, they believed that an earlier *lower* bound of $\Omega(n^{3/2})$, also by Molloy, Reed, and Steiger [56], should be tight. We show the following result (see Section 2.3 for the precise definition of mixing time):

Theorem 1.1. *The triangulation flip walk on the convex $n+2$-point set mixes in time $O(n^3 \log^3 n)$.*

Prior to the present work, no progress had been made either on upper or lower bounds for this chain in 25 years—even as new polynomial upper bounds and exponential lower bounds were given for other geometric chains, from lattice point set triangulations [71, 13] to quadrangulations of planar maps [14], and despite many breakthroughs using the newer techniques for other problems.

1.1.1 Decomposition framework

To prove our result, we develop a general decomposition framework that applies to a broad class of Markov chains, as an alternative to prior work by Jerrum, Son, Tetali, and Vigoda [40] that used spectral methods. We obtain our new mixing result for triangulations, then generalize our technique to obtain the first nontrivial mixing result for k-angulations. In Chapter 3 we further generalize this work to obtain the first rapid mixing bounds for Markov chains for sampling independent sets, dominating sets, and *b-edge covers* (generalizing edge covers) in graphs of bounded treewidth, and for maximal independent sets, b-matchings, and *maximal b-matchings* in graphs of bounded treewidth and degree. In that work we also strengthen existing results [35, 23] for proper q-colorings in graphs of bounded treewidth and degree.

The key observation that unifies these chains is that, when viewing their state spaces as graphs (exponentially large graphs relative to the input), they all admit a recursive decomposition satisfying key properties. First, each such graph, called a "flip graph," can be partitioned into a small number of induced subgraphs, where each subgraph is a *Cartesian product* of smaller graphs that are structurally similar to the original graph—and thus can be partitioned again into even smaller product graphs. Second, at each level of recursion, pairs of subgraphs are connected by large matchings. Intuitively, we can "slice" a flip graph into subgraphs that are well connected to each other, then "peel" apart the subgraphs using their Cartesian product structure, and repeat the process recursively. Each recursive level of slicing cuts through many edges (the large matchings), and indeed the peeling also disconnects many mutually well-connected subgraphs from one another.

Prior work has applied this "slicing" and "peeling" paradigm, also known as *projection-*

restriction: Jerrum, Son, Tetali, and Vigoda [40] gave a decomposition theorem (Theorem 2.2) for obtaining bounds on the *spectral gap* of a chain. The spectral gap of a chain is the difference between the two largest eigenvalues of the transition matrix of the chain (equivalently, the adjacency matrix of the flip graph, up to normalization factors). They defined, with respect to any decomposition of the state space of a chain into subgraphs ("slicing"), a *projection* chain—whose states are identified with the subgraphs and whose transitions correspond to the edges between pairs of the subgraphs—and a *restriction* chain within each of the subgraphs. They showed that the gap of the overall chain can be bounded from below by the product of the gaps of the projection chain and the restriction chain, up to some loss factors. Using standard inequalities that relate the spectral gap to the mixing time of a chain, their technique then gives bounds on the mixing time.

One of our contributions is to unify prior applications [40, 35, 23] of projection-restriction into a sufficient set of conditions—given in Lemma 2.9—under which one can apply the spectral decomposition theorem. A more substantial technical contribution is our Theorem 2.1, an analogue to Jerrum, Son, Tetali, and Vigoda's Theorem 2.2 that uses *multicommodity flows* instead of the spectral gap to bound congestion. One can bound the mixing time of a chain by constructing a multicommodity flow in the corresponding flip graph, and bounding the *congestion* of the flow.

One can use our flow-based theorem in place of the spectral theorem and, in some cases, obtain better mixing bounds. In particular, in the case of triangulations, we obtain polynomial mixing via an adaptation of our (combinatorial) technique (Lemma 2.12)—and it is not clear how to adapt the existing spectral theorem to get even a polynomial bound. In the case of k-angulations, our theorem gives a bound that has better dependence on the parameter k.

1.2 Glauber dynamics on graph-theoretic structures

The *Glauber dynamics* on independent sets in a graph—motivated in part by modeling systems in statistical physics—is a Markov chain in which one starts at an arbitrary independent set, then repeatedly chooses a vertex at random and, with probability that depends on a fixed parameter $\lambda > 0$, either removes the vertex from the set (if it is in the set), or adds it to the set (if it is not in the set and has no neighbor in the set). This chain, also known as the *hardcore model*, has seen recent rapid mixing results under various conditions. In addition to independent sets, similar dynamics have been studied for a number of other structures—including, for example, q-colorings, matchings, and edge covers (more generally, b-matchings and b-edge covers).

1.2.1 Our contribution

In Chapter 3, we prove that the hardcore Glauber dynamics mixes rapidly on graphs of bounded treewidth for all fixed $\lambda > 0$, and that the Glauber dynamics on partial q-colorings (for all $\lambda > 0$) of a graph of bounded treewidth, and on q-colorings of a graph of bounded carving width, mix rapidly. Marc Heinrich proved the latter result, namely for q-colorings, in a 2020 preprint [35]. Heinrich's result applies to all graphs of bounded treewidth; however, for graphs of bounded carving width whose degree is less than quadratic in their treewidth, we improve on Heinrich's upper bound—provided that q is fixed. We also prove that the analogous dynamics on the b-edge covers (when b is bounded) and the dominating sets of a graph of bounded treewidth mix rapidly for all $\lambda > 0$. In a similar vein, we prove that three additional chains—on b-matchings (when $\lambda > 0$), on maximal independent sets, and on maximal b-matchings—mix rapidly in graphs where carving width is bounded. (For the latter two chains we consider only the unbiased version.)

To prove our results, we apply the framework we introduce in Chapter 2. As we discuss in Chapter 2, the framework makes progress towards unifying prior work on similar Glauber dynamics with prior work on probabilistic graphical models. The application to *k-angulations* (proving quasipolynomial mixing) illustrates the applicability of the framework beyond graphical models and sampling problems in graphs.

1.3 Main Results

1.3.1 Geometric Results

To obtain Theorem 1.1, we show the following result for the expansion of the associahedron:

Theorem 1.2. *The expansion of the associahedron $K_{3,n+2}$ is $\Omega(1/(\sqrt{n}\log n))$ and $O(1/\sqrt{n})$.*

We will prove the lower bound in Section 2.6 and Section A.1 using the *multicommodity flow*-based machinery we introduce in Section 2.4, after giving intuition in Section 2.2. Combining this result with the connection between flows and mixing [67]—with some additional effort in Section A.1—gives our new $O(n^3 \log^3 n)$ bound (Theorem 1.1) for triangulation mixing.

Although the expansion lower bound is more interesting for the sake of rapid mixing, the upper bound in Theorem 1.2—which we prove in Section A.3—recovers Molloy, Reed, and Steiger's $\Omega(n^{3/2})$ mixing lower bound [56]. It is also the first result showing that the associahedron has combinatorial expansion $o(1)$. By contrast, Anari, Liu, Oveis Gharan, and Vinzant recently proved [4, 3], settling a conjecture of Mihail and Vazirani [55], that matroids have expansion one. (Mihail and Vazirani in fact conjectured that all graphs realizable as the 1-skeleton of a *0-1 polytope* have expansion one.) Although the set of convex n-gon triangulations is not a matroid, it is an important subset system—and this work shows that it does not have expansion one. More generally, we give the following quasipolynomial bound for k-angulations:

Theorem 1.3. *For every fixed $k \geq 3$, the k-angulation flip walk on the convex $(k-2)n+2$-point set mixes in time $n^{O(k \log n)}$.*

In Section 2.8, we give a lower bound on the *treewidth* of the $n \times n$ integer lattice point set triangulation flip graph:

Theorem 1.4. *The treewidth of the triangulation flip graph F_n on the $n \times n$ integer lattice point set is $\Omega(N^{1-o(1)})$, where $N = |V(F_n)|$.*

1.3.2 Graph-Theoretic Chain Results

Our main results are the following (see Section 3.1 for relevant definitions).

Theorem 1.5. *The hardcore Glauber dynamics mixes in time $n^{O(t)}$ on graphs of treewidth t for all fixed $\lambda > 0$.*

Theorem 1.6. *The (unbiased) Glauber dynamics on q-colorings (when $q \geq \Delta + 2$ is fixed) mixes in time $n^{O(t)}$ on graphs of treewidth t and degree Δ. The Glauber dynamics on partial q-colorings (when $q \geq \Delta + 2$ is fixed) mixes in time $n^{O(t)}$ on graphs of treewidth t for all fixed $\lambda > 0$.*

Theorem 1.7. *The Glauber dynamics on b-edge covers mixes in time $n^{O(t^2)}$ on graphs of treewidth t, for all fixed b and fixed $\lambda > 0$. The Glauber dynamics on dominating sets mixes in time $n^{O(t)}$ on graphs of treewidth t for all fixed $\lambda > 0$. The Glauber dynamics on b-matchings mixes in time $n^{O(t)}$ on graphs of treewidth t, fixed b, and fixed degree Δ for all fixed $\lambda > 0$.*

Theorem 1.8. *There exist flip chains on maximal independent sets and maximal b-matchings, whose stationary distributions are uniform, that mix in time $n^{O(t)}$ on graphs of treewidth t and fixed degree Δ.*

1.4 Organization

1.4.1 Chapter 2 Organization: Geometric Chains

In Chapter 2 we define the Markov chains we are analyzing. In Section 2.2, we give intuition for the decomposition by describing its application to triangulations. In Section 2.4 we present our general decomposition meta-theorems, and compare our contribution to prior work by Jerrum, Son, Tetali, and Vigoda [40]. In particular, we discuss why our purely combinatorial machinery is needed for obtaining new bounds in the case of triangulations. In Section 2.6 we prove a general result that gives a coarse bound on triangulation mixing. Improving this bound to near tightness requires some technical optimizations, which we defer to Appendix A.1; we give a matching upper bound (up to logarithmic factors) in Appendix A.3. In Section 2.7, we show that general k-angulations admit a decomposition satisfying a relaxation (Lemma 2.11) of our general theorem that implies quasipolynomial-time mixing. We analyze the particular quasipolynomial bound we obtain, and show that our combinatorial technique (Theorem 2.1) gives a better dependence on k than one would obtain with the prior decomposition theorem. In Section 2.5 we prove our general combinatorial decomposition theorem, Theorem 2.1. In Section 2.8 we prove a theorem about lattice triangulations; in Appendix A.4 we fill in a few remaining proof details.

1.4.2 Chapter 3 Organization: Graph-Theoretic Chains

In Section 3.1, we give some additional relevant definitions and background for our graph-theoretic chains, including defining the Glauber dynamics and the hardcore model (the chain on independent sets). In Section 3.2.3, we use the chain on independent sets to illustrate what we call a "non-hierarchical" version of the framework (actually the version we give in Chapter 2). This non-hierarchical version works on this chain when carving width is bounded, and in Sections 3.4.1 and Section 3.4.2 we describe how to apply it respectively to q-colorings, and to b-edge covers and b-matchings.

To fully prove Theorem 1.5 and Theorem 1.7, we need to deal with unbounded-degree graphs. In Section 3.3, we discuss how to modify the framework to accomplish this, proving Theorem 1.5 for $\lambda = 1$. We defer some of the details of this proof to Section 3.5, in which we also finish the proof of Theorem 1.6 for $\lambda = 1$.

We prove the general case $\lambda > 0$ of Theorems 1.5 and 1.6 in Appendix 3.6. We finish the proofs of Theorems 1.7 and 1.8 in Section 3.7: applying the framework to the relevant chains requires a further refinement of the framework.

In all of the above, we prove rapid mixing but defer derivation of specific upper bounds to Appendix B.1.

Chapter 2

Improved mixing for the convex polygon triangulation flip walk

2.1 Background

2.1.1 Triangulations of convex point sets and lattice point sets

Let P_n be the regular polygon with n vertices. Every triangulation t of P_{n+2} has $n-1$ diagonals, and every diagonal can be *flipped*: every diagonal D belongs to two triangles forming a convex quadrilateral, so D can be removed and replaced with the diagonal D' lying in the same quadrilateral and crossing D. The set of all triangulations of P_{n+2}, for $n \geq 1$, is the vertex set of a graph that we denote K_n. (This notation is standard, but unfortunate, as it coincides with the notation for a complete graph.) The edges of this graph are the flips between adjacent triangulations. The graph K_n is realizable as the 1-skeleton of an $n-1$-dimensional polytope [49] called the *associahedron* (we also use this name for the graph itself). It is also isomorphic to the rotation graph on the set of all binary plane trees with $n+1$ leaves [69], and equivalently the set of all parenthesizations of an algebraic expression with $n+1$ terms, with "flips" defined as applications of the associative property of multiplication. union The structure of this graph depends only on the convexity and the number of vertices of the polygon, and not on its precise geometry. That is, P_{n+2} need not be regular for K_n to be well defined.

McShine and Tetali [54] showed that the *mixing time* (see Section 2.3) of the uniform random walk on K_n is $O(n^5 \log n)$, following Molloy, Reed, and Steiger's [56] lower bound of $\Omega(n^{3/2})$. Standard inequalities [67] then imply that the *expansion* of K_n is $\Omega(1/(n^4 \log n))$ and $O(n^{1/4})$. It is easy to generalize triangulations to *k-angulations* of a convex polygon $P_{(k-2)n+2}$, and to generalize the definition of a flip between triangulations to a flip between k-angulations: a k-angulation is a maximal division of the polygon into k-gons, and a flip consists of taking a pair of k-gons that share a diagonal, removing that diagonal, and replacing it with one of the other diagonals in the resulting $2k-2$-gon. One can then define the *k-angulation flip walk* on the k-angulations of $P_{(k-2)n+2}$. An analogous graph to the associahedron is defined over the triangulations of the integer lattice (grid) point set with n rows of points and n columns. Substantial prior work has been done on bounds for the number of triangulations in this graph ([1], [41]), as well as characterizing the mixing time of random walks on the graph, when the walks are weighted by a function of the lengths of the edges in a triangulation ([13] [12]).

2.1.2 Convex triangulation flip walk and mixing time

Consider the following random walk on the triangulations of the convex $n+2$-gon:

for $t = 1, 2, \ldots$ **do**

Begin with an arbitrary triangulation t.

Flip a fair coin.

If the result is tails, do nothing.

Else, select a diagonal in t uniformly at random, and flip the diagonal.

end for

(The "do nothing" step is a standard MCMC step that enforces a technical condition known as *laziness*, required for the arguments that bound mixing time.)

At any given time step, this walk induces a probability distribution π over the triangulations of the $n+2$-gon. Standard spectral graph theory shows that π converges to the uniform distribution in the limit. Formally, what McShine and Tetali showed [54] is that the number of steps before π is within *total variation distance* $1/4$ of the uniform distribution is bounded by $O(n^5 \log n)$—in other words, that the *mixing time* is $O(n^5 \log n)$. Any polynomial bound means the walk *mixes rapidly*. We formally define total variation distance:

The *total variation distance* between two probability distributions μ and ν over the same set Ω is defined as

$$d(\mu, \nu) = \frac{1}{2} \sum_{S \in \Omega} |\pi(S) - \pi^*(S)|.$$

Consider a Markov chain with state space Ω with transition matrix P. Given a starting state $S \in \Omega$, the chain induces a probability distribution π_t at each time step t. Suppose the chain is *irreducible*: it connects every pair of states. Suppose further that the chain is *lazy*: it has constant probability of remaining at any given state. Then the distribution converges in the limit to a *stationary* distribution π^*. Furthermore, if the transition probabilities are symmetric (as is the case for the k-angulation flip walk), then the stationary distribution is uniform. The *mixing time* is defined as follows: Given an arbitrary $\varepsilon > 0$, the *mixing time*, $\tau(\varepsilon)$, of a Markov chain with state space Ω and stationary distribution π^* is the minimum time t such that, regardless of starting state, we always have

$$d(\pi_t, \pi^*) < \varepsilon.$$

Suppose that the chain belongs to a family of chains, whose size is parameterized by a value n. (It may be that Ω is exponential in n.) If $\tau(\varepsilon)$ is upper bounded by a function that is polynomial in $\log(1/\varepsilon)$ and in n, say that the chain is *rapidly mixing*. It is common to omit the parameter ε, assuming its value to be the arbitrary constant 1/4.

2.2 Decomposing the convex point set triangulation flip graph

2.2.1 Bounding mixing via expansion

We have a Markov chain that is in fact a random walk on the associahedron K_n. We wish to bound the mixing time of this walk. It turns out that one way to do this is by lower-bounding the *expansion* of the same graph K_n. Intuitively, expansion concerns the extent to which "bottlenecks" exist in a graph. More precisely, it measures the "sparsest" cut—the minimum ratio of the number of edges in a cut divided by the number of vertices on the smaller side of the cut:

The *edge expansion* (or simply *expansion*), $h(G)$, of a graph $G = (V, E)$ is the quantity

$$\min_{S \subseteq V : |S| \leq |V|/2} |\partial S| / |S|,$$

where $\partial S = \{(s,t) | s \in S, t \notin S\}$ is the set of edges across the $(S, V \setminus S)$ cut. A lower bound on edge expansion leads to an upper bound on mixing [39, 67]:

Lemma 2.1. *The mixing time of the Markov chain whose transition matrix is the normalized* *adjacency matrix of a* Δ*-regular graph* G *is*

$$O\left(\frac{\Delta^2 \log(|V(G)|)}{(h(G))^2}\right).$$

One can do better [22, 67] if the paths in a *multicommodity flow* are not too long (Section 2.3).

2.2.2 "Slicing and peeling"

We would like to show that there are many edges in every cut, relative to the number of vertices on one side of the cut. We partition the triangulations $V(K_n)$ into n equivalence classes, each inducing a subgraph of K_n. We show that many edges exist between each pair of the subgraphs. Thus the partitioning "slices" through many edges. After the partitioning, we show that each of the induced subgraphs has large expansion. To do so, we show that each such subgraph decomposes into many copies of a smaller flip graph K_i, $i < n$. This inductive structure lets us assume that K_i has large expansion—then show that the copies of the smaller flip graph are all well connected to one another. We call this "peeling," because one must peel the many K_i copies from one another—removing many edges—to isolate each copy. Molloy, Reed, and Steiger [56] obtained their $O(n^{25})$ mixing *upper* bound via a different decomposition, namely using the *central* triangle, via a non-flow-based method. That decomposition is the one we use for our quasipolynomial bound for general k-angulations in Appendix 2.7. However, we use a different decomposition here, one with a structure that lets us obtain a nearly tight bound, via a multicommodity flow construction. We formalize the slicing step now:

Fix a "special" edge e^* of the convex $n+2$-gon P_{n+2}. For each triangle T having e^* as one of its edges, define the *oriented class* $\mathcal{C}^*(T)$ to be the set of triangulations of P_{n+2} that include T as one of their triangles. Let $\mathcal{T}_n$ be the set of all such triangles; let $\mathcal{S}_n$ be the set of all classes $\{\mathcal{C}^*(T) | T \in \mathcal{T}_n\}$.

Orient P_{n+2} so that e^* is on the bottom. Then say that T (respectively $\mathcal{C}^*(T)$) is to the *left* of T' (respectively $\mathcal{C}^*(T')$) if the topmost vertex of T lies counterclockwise around P_{n+2} from the topmost vertex of T'. Say that T' lies to the *right* of T. Write $T < T'$ and $T' > T$.

See Figure 2.1.

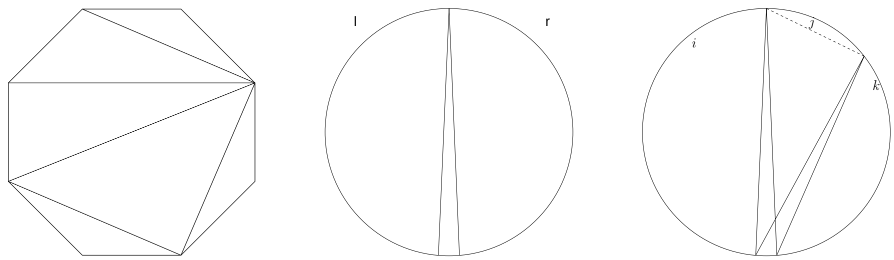

Figure 2.1: Left: A triangulation of the regular octagon. Center: a class $\mathcal{C}^*(T) \in \mathcal{S}_n$, represented schematically by the triangle T that induces it. We depict the regular $n+2$-gon as a circle (which it approximates as $n \to \infty$), for ease of illustration. Each triangulation $t \in \mathcal{C}^*(T)$ consists of T (the triangle shown), and an arbitrary triangulation of the two polygons on either side of T. Notice that $\mathcal{C}^*(T) \cong K_l \square K_r$, where T partitions the $n+2$-gon into an l-gon and an r-gon. Right: the matching $\mathcal{E}^*(T, T')$ between classes $\mathcal{C}^*(T) \cong K_i \square K_{j+k}$ and $\mathcal{C}^*(T') \cong K_{i+j} \square K_k$, is in bijection with the triangulations in $K_i \square K_j \square K_k$ (induced by the quadrilateral containing T and T'). Therefore, $|\mathcal{E}^*(T, T')| = C_i C_j C_k$.

We make observations about the structure of each class as an induced subgraph of K_n

The *Cartesian product* graph $G \square H$ of graphs G and H has vertices $V(G) \times V(H)$ and edges

$$\{((u,v),(u',v)) | (u,u') \in E(G), v \in V(H)\} \cup \{((u,v),(u,v')) | (v,v') \in E(H), u \in V(G)\}.$$

Given a vertex $w = (u,v) \in V(G) \times V(H)$, call u the *projection* of w onto G, and similarly call v the projection of w onto H. (Applying the obvious associativity of the Cartesian product operator, one can naturally define the product $G_1 \square G_2 \square \cdots \square G_k = \square_{i=1}^k G_i$.)

We can now characterize the structure of each class as an induced subgraph of K_n:

Lemma 2.2. *Each class $\mathcal{C}^*(T)$ is isomorphic to a Cartesian product of two associahedron* *graphs K_l and K_r, with $l + r = n - 1$.*

Proof. Each triangle T partitions the $n+2$-gon into two smaller convex polygons with side lengths $l+1$ and $r+1$, such that $l + r = n - 1$. Thus each triangulation in $\mathcal{C}^*(T)$ can be identified with a tuple of triangulations of these smaller polygons. The Cartesian product structure then follows from the fact that every flip between two triangulations in $\mathcal{C}^*(T)$ can be identified with a flip in one of the smaller polygons. □

Lemma 2.2 will be central to the peeling step. For the slicing step, building on the idea in Lemma 2.2 will help us characterize the edge sets between classes:

Given classes $\mathcal{C}^*(T), \mathcal{C}^*(T') \in \mathcal{S}_n$, denote by $\mathcal{E}^*(T, T')$ the set of edges (flips) between $\mathcal{C}^*(T)$ and $\mathcal{C}^*(T')$. Let $\mathcal{B}^*_{n,T'}(T)$ and $\mathcal{B}^*_{n,T}(T')$ be the *boundary sets*—the sets of endpoints of edges in $\mathcal{E}^*(T, T')$—that lie respectively in $\mathcal{C}^*(T)$ and $\mathcal{C}^*(T')$.

Lemma 2.3. *For each pair of classes $\mathcal{C}^*(T)$ and $\mathcal{C}^*(T')$, the boundary set $\mathcal{B}^*_{n,T'}(T)$ induces a subgraph of $\mathcal{C}^*(T)$ isomorphic to a Cartesian product of the form $K_i \square K_j \square K_k$, for some $i + j + k = n - 2$.*

Proof. Each flip between triangulations in adjacent classes $\mathcal{C}^*(T)$ involves flipping a diagonal of T to transform the triangulation $t \in \mathcal{C}^*(T)$ into triangulation $t' \in \mathcal{C}^*(T')$. Whenever this is possible, there must exist a quadrilateral Q, sharing two sides with T (the sides that are not flipped), such that both t and t' contain Q. Furthermore, every $t \in \mathcal{C}^*(T)$ containing Q has a flip to a distinct $t' \in \mathcal{C}^*(T')$. The set of all such boundary vertices $t \in \mathcal{C}^*(T)$ can be identified with the Cartesian product described because Q partitions P_{n+2} into three smaller polygons, so that each triangulation in $\mathcal{B}^*_{n,T'}(T)$ consists of a tuple of triangulations in each of these smaller polygons, and such that every flip between triangulations in $\mathcal{B}^*_{n,T'}(T)$ consists of a flip in one of these smaller polygons. □

Lemma 2.4. *The set $\mathcal{E}^*(T, T')$ of edges between each pair of classes $\mathcal{C}^*(T)$ and $\mathcal{C}^*(T')$ is a nonempty matching. Furthermore, this edge set is in bijection with the vertices of a Cartesian product $K_i \square K_j \square K_k, i + j + k = n - 2$.*

Proof. The claim follows from the reasoning in Lemma 2.3 and from the observation that each triangulation in $\mathcal{B}^*_{n,T'}(T)$ has exactly one flip (namely, flipping a side of the triangle T) to a neighbor in $\mathcal{B}^*_{n,T}(T')$. □

Lemma 2.4 characterizes the structure of the edge sets (namely matchings) between classes; we would also like to know the sizes of the matchings. We will use the following formula:

Let C_n be the nth *Catalan number*, defined as $C_n = \frac{1}{n+1}\binom{2n}{n}$.

Lemma 2.5. *[46, 36]: The number of vertices in the associahedron K_n is C_n, and this number grows as $\frac{1}{\sqrt{\pi} \cdot n^{3/2}} \cdot 2^{2n}$.*

We will prove the following in Section A.4:

Lemma 2.6. *For every $T, T' \in \mathcal{T}_n$,*

$$|\mathcal{E}^*(T, T')| \geq \frac{|\mathcal{C}^*(T)||\mathcal{C}^*(T')|}{C_n}.$$

Lemma 2.6—which states that the number of edges between a pair of classes is at least equal to the product of the cardinalities of the classes, divided by the total number of vertices in the graph $|V(K_n)| = C_n$—is crucial to our results. To explain why this is, we will need to present our multicommodity flow construction (Appendix 2.6). We will give intuition in Section 2.4. For now, it suffices to say that Lemma 2.6 implies that there are many edges between a given pair of classes, justifying (intuitively) the slicing step. For the peeling step, we need the fact that Cartesian graph products preserve the well-connectedness of the graphs in the product [31]:

Lemma 2.7. *Given graphs $G_1, G_2, \ldots, G_k$, Cartesian product $G_1 \square G_2 \square \cdots \square G_k$ satisfies*

$$h(G_1 \square G_2 \square \cdots \square G_k) \geq \frac{1}{2} \min_i h(G_i).$$

Lemma 2.2 says that each of the classes $\mathcal{C}^*(T) \in \mathcal{S}_n$ is a Cartesian graph product of associahedron graphs K_l, K_r, $l < n, r < n$, allowing us to "peel" (decompose) $\mathcal{C}^*(T)$ into graphs that can then be recursively sliced into classes and peeled. Lemma 2.7 implies that the peeling must disconnect many edges, as it involves splitting a Cartesian product graph into many subgraphs (copies of K_l).

We will make all of this intuition rigorous in Section 2.6 by constructing our flow. The choice of paths through which to route flow will closely trace the edges in this recursive "slicing and peeling" decomposition. We will then show that, with this choice of paths, the resulting *congestion*—the maximum amount of flow carried along an edge—is bounded by a suitable polynomial factor. This will provide a lower bound on the expansion.

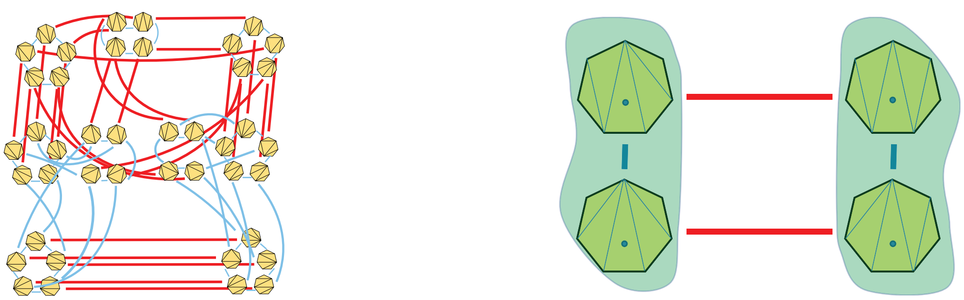

Figure 2.2: **Left**: The associahedron graph K_5, with each vertex representing a triangulation of the regular heptagon. Flips are shown with edges (in blue and red). The vertex set $V(K_n)$ is partitioned into a set $\mathcal{S}_n$ of five equivalence classes (of varying sizes). Within each class, all triangulations share the same triangle containing the bottom edge e^*. Flips (edges) between triangulations in the same class are shown in blue. Flips between triangulations in different classes are shown in red. To "slice" K_5 into its subgraphs, one must cut through these red matchings. **Right**: A class $\mathcal{C}^*(T)$ from the graph K_5 on the left-hand side, viewed as an induced subgraph of K_5. The identifying triangle T is marked with a blue dot. This subgraph is isomorphic to a Cartesian product of two K_2 graphs; each copy of K_2 induced by fixing the rightmost diagonal is outlined in green. "Peeling" apart this product requires disconnecting the two red edges connecting the K_2 copies.

2.3 Bounding expansion via multicommodity flows

The way we will lower-bound expansion is by using *multicommodity flows* [67, 42]. A *multicommodity flow* ϕ in a graph $G = (V, E)$ is a collection of functions $\{f_{st} : A \to \mathbb{R} \mid s, t \in V\}$, where $A = \bigcup_{\{u,v\} \in E} \{(u, v), (v, u)\}$, combined with a *demand function* $D : V \times V \to \mathbb{R}$.

Each f_{st} is a flow sending $D(s, t)$ units of a commodity from vertex s to vertex t through the edges of G. We consider the capacities of all edges to be infinite. Let $f_{st}(u, v)$ be the amount of flow sent by f_{st} across the arc (u, v). (It may be that $f_{st}(u, v) \neq f_{st}(v, u)$.) Let

$$f(u, v) = \frac{1}{|V|} \sum_{s,t \in V \times V} f_{st}(u, v),$$

and let $\rho = \max_{(u,v) \in A} f(u, v)$. Call ρ the *congestion*. Unless we specify otherwise, we will mean by "multicommodity flow" a *uniform multicommodity flow*, i.e. one in which $D(s, t) = 1$ for all s, t. The following is well established and enables the use of multicommodity flows as a powerful lower-bounding technique for expansion:

Lemma 2.8. *Given a uniform multicommodity flow f in a graph $G = (V, E)$ with congestion ρ, the expansion $h(G)$ is at least $1/(2\rho)$.*

Lemma 2.8, combined with Lemma 2.1, gives an automatic upper bound on mixing time given a multicommodity flow with an upper bound on congestion—but with a quadratic loss. As we will discuss in Appendix 2.5, one can do better if the paths used in the flow are short [22, 67].

2.4 Our framework

In addition to the new mixing bounds for triangulations and for general k-angulations, we make general technical contributions, in the form of three meta-theorems, which we present in this section. First, Theorem 2.1 provides a general recursive mechanism for analyzing the expansion of a flip graph in terms of the expansion of its subgraphs. Equivalently, viewing the random walk on such a flip graph as a Markov chain, this theorem provides a mechanism for analyzing the mixing time of a chain, in terms of the mixing times of smaller *restriction* chains into which one decomposes the original chain—and analyzing a *projection* chain over these smaller chains. We obtain, in certain circumstances such as the k-angulation walk, better mixing time bounds than one obtains applying similar prior decomposition theorems—which used a different underlying machinery.

The second theorem, Lemma 2.9, observes and formalizes a set of conditions satisfied by a number of chains (equivalently, flip graphs) under which one can apply either our Theorem 2.1, or prior decomposition techniques, to obtain rapid mixing reuslts. Depending on the chain, one may then obtain better results either by applying Theorem 2.1, or by applying the prior techniques. It is Lemma 2.9 that we apply to a number of other chains in Chapter 3 The case of general k-angulations satisfies a relaxation of this theorem (Lemma 2.11), giving a quasipolynomial bound. This bound will come from incurring a polynomial loss over logarithmic recursion depth.

The third theorem, Lemma 2.12, adapts the machinery in Theorem 2.1 to eliminate this multiplicative loss altogether, assuming that a chain satisfies certain properties. One such key property is the existence of large matchings in Lemma 2.6 in Section 2.2. Another property, which we will discuss further after presenting Lemma 2.12, is that the *boundary sets*—the vertices in one class (equivalently, states in a restriction chain) having neighbors in another class—are well connected to the rest of the first class. When these properties are satisfied, one can apply our flow machinery to overcome the multiplicative loss and obtain a polynomial bound. However, the improvement relies on observations about congestion that do not obviously translate to the spectral setting.

2.4.1 Markov chain decomposition via multicommodity flow

In this section we state our first general theorem. To place our contribution in context with prior work, we cast our flip graphs in the language of Markov chains. As we discussed in Section 2.1.2, any Markov chain satisfying certain mild conditions has a *stationary* distribution π^* (which in the case of our triangulation walks is uniform). We can view such a chain as a random walk on a graph $\mathcal{M}$ (an unweighted graph in the case of the chains we consider, which have uniform distributions and regular transition probabilities). In the case of convex polygon triangulations, we have $\mathcal{M} = K_n$.

The flip graph $\mathcal{M}$ has vertex set Ω and (up to normalization by degree) adjacency matrix P—and we abuse notation, identifying the Markov chain $\mathcal{M}$ with this graph. When π^* is not uniform, it is easy to generalize the flip graph to a *weighted* graph, with each vertex (state) t assigned weight $\pi(t)$, and each transition (edge) (t, t') assigned weight $\pi(t)P(t, t') = \pi(t')P(t', t)$. We assume here that this latter equality holds, a condition on the chain $\mathcal{M}$ known as *reversibility*. We then replace a uniform multicommodity flow with one where $D(t, t') = \pi(t)\pi(t')$ (up to normalization factors).

Consider a Markov chain $\mathcal{M}$ with finite state space Ω and probability transition matrix P, and stationary distribution π. Consider a partition of the states of Ω into classes $\Omega_1, \Omega_2, \ldots, \Omega_k$. Let the *restriction* chain, for $i = 1, \ldots, k$, be the chain with state space Ω_i, probability distribution π_i, with $\pi_i(x) = \pi(x)/(\sum_{y\in\Omega_i} \pi(y))$, for $x \in \Omega_i$, and transition probabilities $P_i(x, y) = P(x, y)/(\sum_{z\in\Omega_i} P(x, z))$. Let the *projection* chain be the chain with state space $\bar{\Omega} = \{1, 2, \ldots, k\}$, stationary distribution $\bar{\pi}$, with $\bar{\pi}(i) = \sum_{x\in\Omega_i} \pi(i)$, and transition probabilities $\bar{P}(i,j) = \sum_{x\in\Omega_i, y\in\Omega_j} P(x,y)$.

Theorem 2.1. *Let $\mathcal{M}$ be a reversible Markov chain with finite state space Ω probability transition matrix P, and stationary distribution π^*. Suppose $\mathcal{M}$ is connected (irreducible). Suppose $\mathcal{M}$ can be decomposed into a collection of restriction chains $(\Omega_1, P_1), (\Omega_2, P_2), \ldots, (\Omega_k, P_k)$, and a projection chain $(\bar{\Omega}, \bar{P})$. Suppose each restriction chain admits a multicommodity flow (or canonical paths) construction with congestion at most $\rho_{\max}$. Suppose also that there exists a multicommodity flow construction in the projection chain with congestion at most $\bar{\rho}$. Then there exists a multicommodity flow construction in $\mathcal{M}$ (viewed as a weighted graph in the natural way) with congestion*

$$(1 + 2\bar{\rho}\gamma\Delta)\rho_{\max},$$

where $\gamma = \max_{i\in[k]} \max_{x\in\Omega_i} \sum_{y\notin\Omega_i} P(x,y)$, and Δ is the degree of $\mathcal{M}$.

The proof of Theorem 2.1 is in Section 2.5. Jerrum, Son, Tetali, and Vigoda [40] presented an analogous decomposition theorem, which we restate below as Theorem 2.2, and which has become a standard tool in mixing time analysis. The key difference between our theorem and theirs is that our theorem uses multicommodity flows, while their theorem uses the so-called *spectral gap*—another parameter that can use to bound the mixing time of a chain. Often, the spectral gap gives tighter mixing bounds than combinatorial methods. Their Theorem 2.2 gave bounds analogous to our Theorem 2.1, but with the multicommodity flow congestion replaced with the *spectral gap* of a chain—and with a 3γ term in place of our 2γ. (They also gave an analogous version for the *log-Sobolev* constant—yet another parameter for bounding mixing times.) The spectral gap of a chain $\mathcal{M} = (\Omega, P)$, which we denote λ, is the difference between the two largest eigenvalues of the transition matrix P (which we can view as the normalized adjacency matrix of the corresponding weighted graph). The key point is that while on the one hand the mixing time τ satisfies $\tau \le \lambda^{-1}\log|\Omega|$, the bound on mixing using expansion in Lemma 2.1 comes from passing through the spectral gap: $\lambda \ge \frac{(h(\mathcal{M}))^2}{2\Delta^2}$, where Δ is the degree of the flip graph and $h(\mathcal{M})$ is the expansion of $\mathcal{M}$. The quadratic loss in passing from expansion to mixing is not incurred when bounding the spectral gap directly, so one can obtain better bounds via the spectral gap. Jerrum, Son, Tetali, and Vigoda gave a mechanism for doing precisely this:

Theorem 2.2. *[40] Let $\mathcal{M}$ be a reversible Markov chain with finite state space Ω probability transition matrix P, and stationary distribution π^*. Suppose $\mathcal{M}$ is connected (irreducible). Suppose $\mathcal{M}$ can be decomposed into a collection of restriction chains (Ω_1, P_1), (Ω_2, P_2), $\ldots, (\Omega_k, P_k)$, and a projection chain $(\bar{\Omega}, \bar{P})$. Suppose each restriction chain has spectral gap at least $\lambda_{\min}$. Suppose also that the projection chain has spectral gap at least $\bar{\lambda}$. Then $\mathcal{M}$ has gap at least*

$$\min\left\{\frac{\lambda_{\min}}{3}, \frac{\bar{\lambda}\lambda_{\min}}{3\gamma + \bar{\lambda}}\right\},$$

where γ is as in Theorem 2.1.

Our Theorem 2.1 has a relatively simple proof (Section 2.5) and shows that the earlier spectral machinery can be replaced with a purely combinatorial technique. We also obtain a tighter bound on expansion than would result from a black-box application of Theorem 2.2. The cost to our improvement is in passing from expansion to mixing via the spectral gap. Nonetheless, we will show that in the case of triangulations, our Theorem 2.1 can be adapted to give a new mixing bound whereas, by contrast, it is not clear how to obtain even a polynomial bound adapting Jerrum, Son, Tetali, and Vigoda's spectral machinery. We will also show that for general k-angulations, one can, with our technique, use a combinatorial insight to eliminate the γ factor in our decomposition in favor of a Δ^{-1} factor (for k-angulations we have $\gamma = k/\Delta$)—whereas it is not clear how to do so with the spectral decomposition.

2.4.2 General pattern for bounding projection chain congestion

Our second decomposition theorem, which we will apply to general k-angulations, states that if one can recursively decompose a chain into restriction chains in a particular fashion, and if the projection chain is well connected, then Theorem 2.1 gives an expansion bound:

Lemma 2.9. *Let $\mathcal{F} = \{\mathcal{M}_1, \mathcal{M}_2, \dots\}$ be a family of connected graphs, parameterized by a value n. Suppose that every graph $\mathcal{M}_n = (\mathcal{V}_n, \mathcal{E}_n) \in \mathcal{F}$, for $n \geq 2$, can be partitioned into a set $\mathcal{S}_n$ of classes satisfying the following conditions:*

1. *Each class in $\mathcal{S}_n$ is isomorphic to a Cartesian product of one or more graphs $\mathcal{C}(T) \cong \mathcal{M}_{i_1} \square \cdots \mathcal{M}_{i_k}$, where for each such graph $\mathcal{M}_{i_j} \in \mathcal{F}$, $i_j \leq n/2$.*
2. *The number of classes is $O(1)$.*
3. *For every pair of classes $\mathcal{C}(T), \mathcal{C}(T') \in \mathcal{S}_n$ that share an edge, the number of edges between the two classes is $\Omega(1)$ times the size of each of the two classes.*
4. *The ratio of the sizes of any two classes is $\Theta(1)$.*

Suppose further that $|\mathcal{V}_1| = 1$. Then the expansion of $\mathcal{M}_n$ is $\Omega(n^{-O(1)})$.

Lemma 2.9 follows from applying induction to Theorem 2.1. An analogue in terms of spectral gap follows from applying induction to Theorem 2.2. Furthermore, as we will prove in Appendix 2.5, a precise statement of the bounds given by Lemma 2.9 is as follows:

Lemma 2.10. *Suppose a flip graph $\mathcal{M}_n = (\mathcal{V}_n, \mathcal{E}_n)$ belongs to a family $\mathcal{F}$ of graphs satisfying the conditions of Lemma 2.9. Suppose further that every graph $\mathcal{M}_k = (\mathcal{V}_k, \mathcal{E}_k) \in \mathcal{F}$, $k < n$, satisfies*

$$|\mathcal{V}_k|/|\mathcal{E}_{k,\min}| \leq f(k),$$

for some function $f(k)$, where $\mathcal{E}_{k,\min}$ is the smallest edge set between adjacent classes $\mathcal{C}(T), \mathcal{C}(T') \in \mathcal{S}_k$, where $\mathcal{S}_k$ is as in Lemma 2.9. Then the expansion of $\mathcal{M}_n$ is

$$\Omega(1/(2f(n))^{\log n})).$$

Proof. Constructing an arbitrary multicommodity flow (or set of canonical paths) in the projection graph at each inductive step gives the result claimed. The term $|\mathcal{V}_k|/|\mathcal{E}_{k,\min}|$ bounds the (normalized) congestion in any such flow because the total amount of flow exchanged by all pairs of vertices (states) combined is $|\mathcal{V}_k|^2$, and the minimum weight of an edge in the projection graph is $|\mathcal{E}_{k,\min}|$.

Notice that we do not incur a $\gamma\Delta$ term here, because even if a state (vertex) in $\Omega_i \subseteq \mathcal{V}_k$ has neighbors $x \in \Omega_j, y \in \Omega_l$, z still only receives no more than $|\mathcal{V}_k|^2/\mathcal{E}_{k,\min}\}$ flow across the edges (z, x) and (z, y) combined. □

Remark 2.1. *The $\gamma\Delta$ factor in Theorem 2.1, which does not appear in Lemma 2.10, does appear in a straightforward appliation of Jerrum, Son, Tetali, and Vigoda's Theorem 2.2.*

We will show that k-angulations (with fixed $k \geq 4$) satisfy a relaxation of Lemma 2.9:

Lemma 2.11. *Suppose a family $\mathcal{F}$ of graphs satisfies the conditions of Lemma 2.9, with the $\Omega(1)$, $O(1)$, and $\Theta(1)$ factors in Conditions 3, 2, and 4 respectively replaced by $\Omega(n^{-O(1)})$, $O(n^{O(1)})$, and $\Theta(n^{O(1)})$. Then for every $\mathcal{M}_n \in \mathcal{F}$, the expansion of $\mathcal{M}_n$ is $\Omega(n^{-O(\log n)})$.*

Lemma 2.9 enables us to relate a number of chains admitting a certain decomposition process in a black-box fashion, unifying prior work applying Theorem 2.2 separately to individual chains. Marc Heinrich [35] presented a similar but less general construction for the Glauber dynamics on q-colorings in bounded-treewidth graphs; other precursors exist, including for the hardcore model on certain trees [40] and a general argument for a class of graphical models [20]. In Chapter 3, we apply Lemma 2.9 to chains for sampling independent sets and dominating sets in bounded-treewidth graphs, as well as chains on q-colorings, maximal independent sets, and several other structures, in graphs whose treewidth and degree are bounded.

2.4.3 Eliminating inductive loss: nearly tight conductance for triangulations

We now give the meta-theorem that we will apply to triangulations. Lemma 2.9—using either Theorem 2.1 or Theorem 2.2—gives a merely quasipolynomial bound when applied straightforwardly to k-angulations, including the case of triangulations—simply because the $f(n)$ term in Lemma 2.10 is $\omega(1)$ and thus the overall congestion is $\omega(1)^{\log n}$ (not polynomial). However, it turns out that the large matchings given by Lemma 2.6 between pairs of classes in the case of triangulations (but not general k-angulations), combined with some additional structure in the triangulation flip walk, satisfy an alternative set of conditions that suffice for rapid mixing. The conditions are:

Lemma 2.12. *Let $\mathcal{F} = \{\mathcal{M}_1, \mathcal{M}_2, \dots\}$ be an infinite family of connected graphs, parameterized by a value n. Suppose that for every graph $\mathcal{M}_n = (\mathcal{V}_n, \mathcal{E}_n) \in \mathcal{F}$, for $n \geq 2$, the vertex set $\mathcal{V}_n$ can be partitioned into a set $\mathcal{S}_n$ of classes inducing subgraphs of $\mathcal{M}_n$ that satisfy the following conditions:*

1. *Each subgraph is isomorphic to a Cartesian product of one or more graphs $\mathcal{C}(T) \cong \mathcal{M}_{i_1} \Box \cdots \mathcal{M}_{i_k}$, where for each such graph $\mathcal{M}_{i_j} \in \mathcal{F}$, $i_j < n$.*
2. *The number of classes is $n^{O(1)}$.*
3. *For every pair of classes $\mathcal{C}(T), \mathcal{C}(T') \in \mathcal{S}_n$, the set of edges between the subgraphs induced by the two classes is a matching of size at least $\frac{|\mathcal{C}(T)||\mathcal{C}(T')|}{|\mathcal{V}_n|}$.*
4. *Given a pair of classes $\mathcal{C}(T), \mathcal{C}(T') \in \mathcal{S}_n$, there exists a graph $\mathcal{M}_i$ in the Cartesian product $\mathcal{C}(T)$, and a class $\mathcal{C}(U) \in \mathcal{S}_i$ within the graph $\mathcal{M}_i$, such that the set of vertices in $\mathcal{C}(T)$ having a neighbor in $\mathcal{C}(T')$ is precisely the set of vertices in $\mathcal{C}(T)$ whose projection onto $\mathcal{M}_i$ lies in $\mathcal{C}(U)$. Furthermore, no class $\mathcal{C}(U)$ within $\mathcal{M}_i$ is the projection of more than one such boundary.*

Suppose further that $|\mathcal{V}_1| = 1$. Then the expansion of $\mathcal{M}_n$ is $\Omega(1/(\kappa(n)n))$, where $\kappa(n) = \max_{1 \leq i \leq n} |\mathcal{C}(S_i)|$ is the maximum number of classes in any $\mathcal{M}_i, i \leq n$.

Unlike Lemma 2.9, this lemma requires a purely combinatorial construction; it is not clear how to apply spectral methods to obtain even a polynomial bound. Condition 4 is crucial. To give more intuition for this condition, we state and prove the following fact about the triangulation flip graph (visualized in Figure 2.3):

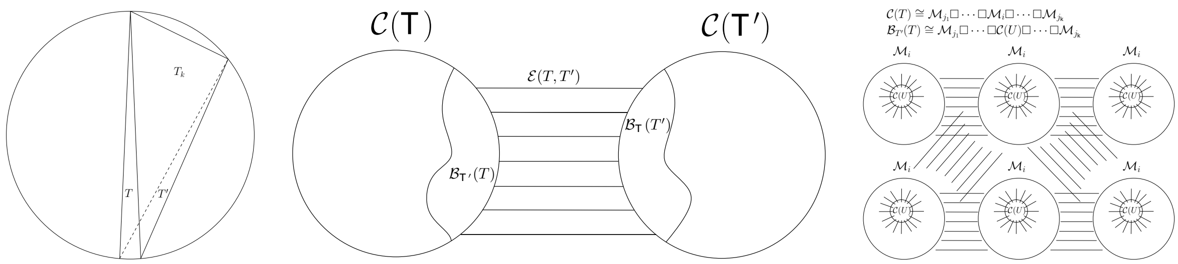

Figure 2.3: **Left**: (Lemma 2.13) The set of edges $\mathcal{E}^*(T, T')$ has $K_i \square \mathcal{C}^*(T_k)$ as its set of boundary vertices in $\mathcal{C}^*(T)$. **Center**: An illustration of Condition 3 in Lemma 2.12, showing a large matching $\mathcal{E}(T, T')$ between two classes (subgraphs) $\mathcal{C}(T)$ and $\mathcal{C}(T')$. **Right**: An illustration of Conditions 1 and 4 in Lemma 2.12: $\mathcal{C}(T)$ as a Cartesian product of smaller graphs $\mathcal{M}_{j_1}, \ldots, \mathcal{M}_i, \ldots, \mathcal{M}_{j_k}$ in the family $\mathcal{F}$. The schematic view shows this Cartesian product as a collection of copies of $\mathcal{M}_i$, connected via perfect matchings between pairs of the copies—with the pairs to connect determined by the structure of the Cartesian product. The boundary $\mathcal{B}_{T'}(T)$ (center) is isomorphic to a class $\mathcal{C}(U)$ (right) within $\mathcal{M}_i$, a graph in the product. Within each copy of $\mathcal{M}_i$, many edges connect $\mathcal{C}(U)$ to the rest of $\mathcal{M}_i$.

Lemma 2.13. *Given $T, T' \in \mathcal{T}_n$, suppose T' lies to the right of T. Then the subgraph of $\mathcal{C}^*(T)$ induced by $\mathcal{B}^*_{n,T'}(T)$ is isomorphic to a Cartesian product $K_l \square \mathcal{C}^*(T_k)$, where $l + r = n - 1$, and where T_k has as an edge the right diagonal of T, and as the vertex opposite this edge the topmost vertex of T'. A symmetric fact holds for $\mathcal{B}^*_{n,T}(T')$.*

Proof. Every triangulation in $\mathcal{B}^*_{n,T'}(T)$ (i) includes the triangle T and (ii) is a single flip away from including the triangle T'. As we observed in the proof of Lemma 2.3, this implies that $\mathcal{B}^*_{n,T'}(T)$ consists of the set of triangulations in $\mathcal{C}^*(T)$ containing a quadrilateral Q. Specifically, Q shares two sides with T: one of these is e^*, and the other is the left side of T.

One of the other two sides of Q is the right side of $\mathcal{C}^*(T')$. Combining this side with the "top" side of Q and with the right side of T, one obtains the triangle T_k, proving the claim. □

Lemma 2.13 implies that there are many edges between the boundary set $\mathcal{B}^*_{n,T'}(T)$ and the rest of $\mathcal{C}^*(T)$: $\mathcal{C}^*(T) \cong K_l \square K_r$, where K_l and K_r are smaller associahedron graphs, so $\mathcal{C}^*(T)$ is a collection of copies of K_r, with pairs of copies connected by perfect matchings. Each K_r copy can itself be decomposed into a set $\mathcal{S}_r$ of classes, one of which, namely $\mathcal{C}^*(T_k)$, is the intersection of $\mathcal{B}^*_{n,T'}(T)$ with the K_r copy. Applying Condition 3 to the K_r copy implies that there are many edges between boundary vertices in $\mathcal{C}^*(T_k)$ to other subgraphs (classes) in the K_r copy. That is, the boundary set $\mathcal{B}^*_{n,T'}(T)$ is well connected to the rest of $\mathcal{C}^*(T)$.

Figure 2.3 visualizes this situation in general terms for the framework. We have now proven:

Lemma 2.14. *The associahedron graph K_n, along with the oriented partition, satisfies the conditions of Lemma 2.12.*

Proof. The graph K_n is connected [54]. Conditions 1 and 3 follow from Lemma 2.2, Lemma 2.4, and Lemma 2.6. Concerning the boundary sets, Condition 4 follows from Lemma 2.13 and from the discussion leading to this lemma. □

Together with Lemma 2.1 and the fact that K_n is a $\Theta(n)$-regular graph, Lemma 2.14 implies rapid mixing, pending the proof of Lemma 2.12—which we prove in Appendix 2.6.

2.4.4 Intuition for the flow construction for triangulations

We will prove Lemma 2.12 in Appendix 2.6, from which a coarse expansion lower bound for triangulations—and a corresponding coarse (but polynomial) upper bound for mixing—will be immediate by Lemma 2.14. We give some intuition now for the flow construction we will give in the proof of Lemma 2.12, and in particular for the centrality of Condition 3 and

Condition 4 (corresponding respectively to Lemma 2.6 and Lemma 2.13 for triangulations). Consider the case of triangulations, for concreteness. Every $t \in \mathcal{C}^*(T), t' \in \mathcal{C}^*(T')$ must exchange a unit of flow. This means that a total of $|\mathcal{C}^*(T)||\mathcal{C}^*(T')|$ flow must be sent across the matching $\mathcal{E}^*(T, T')$. To minimize congestion, it will be optimal to equally distribute this flow across all of the boundary matching edges. We can decompose the overall problem of routing flow from each $t \in \mathcal{C}^*(T)$ to each $t' \in \mathcal{C}^*(T')$ into three subproblems: (i) *concentrating* flow from every triangulation in $\mathcal{C}^*(T)$ within the boundary set $\mathcal{B}^*_{n,T'}(T)$, (ii) routing flow across the matching edges $\mathcal{E}^*(T, T')$, i.e. from $\mathcal{B}^*_{n,T'}(T) \subseteq \mathcal{C}^*(T)$ to $\mathcal{B}^*_{n,T}(T') \subseteq \mathcal{C}^*(T')$, and (iii) *distributing* flow from the boundary $\mathcal{B}^*_{n,T}(T')$ to each $t' \in \mathcal{C}^*(T')$. Now, the amount of

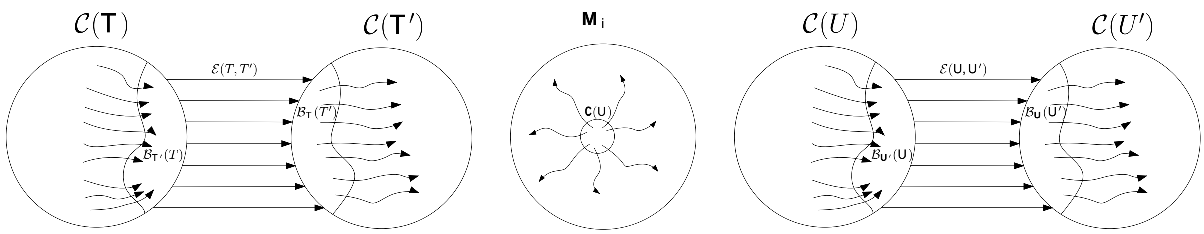

Figure 2.4: **Left**: The problem of sending flow from each $t \in \mathcal{C}^*(T)$ to each $t' \in \mathcal{C}^*(T')$, decomposed into subproblems: (i) *concentrating* flow within $\mathcal{B}^*_{n,T'}(T)$, (ii) *transmitting* the flow across the boundary matching $\mathcal{E}^*(T, T')$, and (iii) *distributing* the flow from $\mathcal{B}^*_{n,T}(T')$ throughout $\mathcal{C}^*(T')$. **Center**: Within each copy of $\mathcal{M}_i$ in the product $\mathcal{C}^*(T') \cong \mathcal{M}_{j_1} \square \cdots \square \mathcal{M}_i \square \cdots \square \mathcal{M}_{j_k}$, the distribution problem in Figure 2.4 induces the problem of distributing flow from a class $\mathcal{C}^*(U)$—namely the projection of $\mathcal{B}^*_{n,T}(T')$ onto $\mathcal{M}_i$—throughout the rest of $\mathcal{M}_i$. **Right**: The problem in the center figure induces subproblems in which $\mathcal{C}^*(U) \subseteq \mathcal{M}_i$ must send flow to each $\mathcal{C}^*(U') \subseteq \mathcal{M}_i$. These subproblems are of the same form as the original $\mathcal{C}^*(T), \mathcal{C}^*(T')$ problem (left), and can be solved recursively. The large matchings $\mathcal{E}^*(T, T'), \mathcal{E}^*(U, U')$ guaranteed by Condition 3 prevent any recursive congestion increase.

flow that must be concentrated from $\mathcal{C}^*(T)$ at *each* boundary triangulation $u \in \mathcal{B}^*_{n,T'}(T)$ (and symmetrically distributed from *each* $v \in \mathcal{B}^*_{n,T}(T')$ throughout $\mathcal{C}^*(T')$) is equal to

$$\frac{|\mathcal{C}^*(T)||\mathcal{C}^*(T')|}{|\mathcal{B}^*_{n,T'}(T)|} = \frac{|\mathcal{C}^*(T)||\mathcal{C}^*(T')|}{|\mathcal{B}^*_{n,T}(T')|} = \frac{|\mathcal{C}^*(T)||\mathcal{C}^*(T')|}{|\mathcal{E}^*(T, T')|} \leq C_n,$$

where we have used the equality $|\mathcal{B}^*_{n,T'}(T)| = |\mathcal{B}^*_{n,T}(T')| = |\mathcal{E}^*(T, T')|$ by Lemma 2.3 and Lemma 2.4, and where the inequality follows from Lemma 2.6. As a result, in the "concentration" and "distribution" subproblems (i) and (iii), at most C_n flow is concentrated at or distributed from any given triangulation (Figure 2.4). This bound yields a recursive structure: the concentration (respectively distribution) subproblem decomposes into a flow problem within $\mathcal{C}^*(T)$ (respectively $\mathcal{C}^*(T')$), in which, by the inequality, each triangulation has C_n total units of flow it must receive (or send). We will then apply Condition 4, observing (see Figure 2.4) that the concentration (symmetrically) distribution of this flow can be done entirely between pairs of classes $\mathcal{C}^*(U), \mathcal{C}^*(U')$ within copies of a smaller flip graph $\mathcal{M}_i$ in the Cartesian product $\mathcal{C}^*(T') \cong \mathcal{M}_{j_1} \square \cdots \square \mathcal{M}_i \square \cdots \square \mathcal{M}_{j_k}$.

The $\mathcal{C}^*(U), \mathcal{C}^*(U')$ subproblem is of the same form as the original $\mathcal{C}^*(T), \mathcal{C}^*(T')$ problem (Figure 2.4), and we will show that the C_n bound on the flow (normalizing to congestion one) across the $\mathcal{E}^*(T, T')$ edges will induce the same C_n bound across the $\mathcal{E}^*(U, U')$ edges in the induced subproblem. We further decompose the $\mathcal{C}^*(U), \mathcal{C}^*(U')$ problem into concentration, transmission, and distribution subproblems without any gain in overall congestion. To see this, view the initial flow problem in K_n as though every triangulation $t \in V(K_n)$ is initially "charged" with $|V(K_n)| = C_n$ total units of flow to distribute throughout K_n. Similarly, in the induced distribution subproblem within each copy of $\mathcal{M}_i = K_i$ in the product $\mathcal{C}^*(T')$, each vertex on the boundary $\mathcal{B}^*_{n,T'}(T)$ is initially "charged" with C_n total units to distribute throughout K_i. Just as the original problem in K_n results in each $\mathcal{E}^*(T, T')$ carrying at most C_n flow across each edge, similarly (we will show in Section 2.6) the induced problem in K_i results in each $\mathcal{E}^*(U, U')$ carrying at most C_n flow across each edge. This preservation of the bound C_n under the recursion avoids any congestion increase.

One must be cautious, due to the linear recursion depth, not to accrue even a constant-factor loss in the recursive step (the coefficient 2 in Theorem 2.1). In Theorem 2.1, it turns out that this loss comes from routing *outbound* flow within a class $\mathcal{C}^*(T)$—flow that must be sent to other classes—and then also routing *inbound* flow. The combination of these steps involves two "recursive invocations" of a uniform multicommodity flow that is inductively assumed to exist within $\mathcal{C}^*(T)$. We will show in Section 2.6 that one can avoid the second "invocation" with an initial "shuffling" step: a uniform flow within $\mathcal{C}^*(T)$ in which each triangulation $t \in \mathcal{C}^*(T)$ distributes all of its outbound flow evenly throughout $\mathcal{C}^*(T)$.

It is here that Jerrum, Son, Tetali, and Vigoda's spectral Theorem 2.2 breaks down, giving a 3-factor loss at each recursion level, due to applying the Cauchy-Schwarz inequality to a *Dirichlet form* that is decomposed into expressions over the restriction chains. Although Jerrum, Son, Tetali, and Vigoda gave circumstances for mitigating or eliminating their multiplicative loss, this chain does not satisfy those conditions in an obvious way.

2.5 Proof that the conditions of Lemma 2.9 imply rapid mixing

In this section we prove Theorem 2.1. Lemma 2.9 will then follow by way of Lemma 2.10.

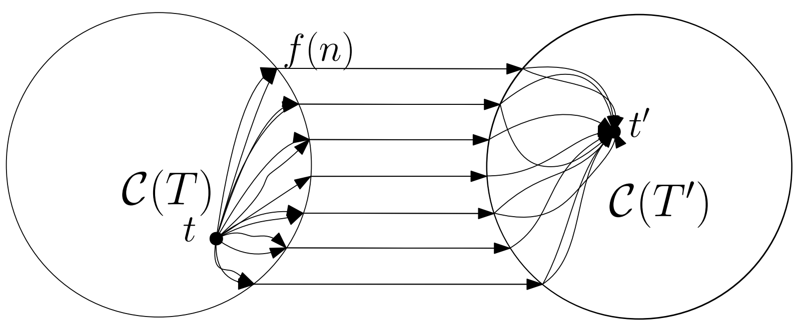

Figure 2.5: In the flow construction we use for quasipolynomial mixing (Theorem 1.3), we first find a flow in $\mathcal{C}(T)$ (similarly $\mathcal{C}(T')$) and bound its congestion. (Actually, we assume such a flow exists, for the inductive hypothesis.) We then reuse the paths from this flow in routing the flow between t and t'. Reusing these paths results in compounding the amount of flow across each path by $f(n)$, where $f(n)$ is the amount of flow across an edge between the two classes.

Theorem 2.1. *Let $\mathcal{M}$ be a reversible Markov chain with finite state space Ω probability transition matrix P, and stationary distribution π^*. Suppose $\mathcal{M}$ is connected (irreducible). Sup-*

pose $\mathcal{M}$ can be decomposed into a collection of restriction chains $(\Omega_1, P_1), (\Omega_2, P_2), \ldots, (\Omega_k, P_k)$, and a projection chain $(\bar{\Omega}, \bar{P})$. Suppose each restriction chain admits a multicommodity flow (or canonical paths) construction with congestion at most $\rho_{\max}$. Suppose also that there exists a multicommodity flow construction in the projection chain with congestion at most $\bar{\rho}$. Then there exists a multicommodity flow construction in $\mathcal{M}$ (viewed as a weighted graph in the natural way) with congestion

$$(1 + 2\bar{\rho}\gamma\Delta)\rho_{\max},$$

where $\gamma = \max_{i \in [k]} \max_{x \in \Omega_i} \sum_{y \notin \Omega_i} P(x, y)$, and Δ is the degree of $\mathcal{M}$.

Here, instead of a uniform flow in which each pair of states exchanges a single unit, it will be convenient to use a specification of demands and definition of congestion that are closer to standard in the analysis of Markov chains: the demands are $D(z, w) = \pi(z)\pi(w)$, and the congestion across an edge (x, y) produced by a multicommodity flow f that satisfies the demands $\{D(z, w) | z, w \in \Omega\}$ is $\rho(x, y) = f(x, y)/(\Delta \cdot Q(x, y))$, where $Q(x, y) = \pi(x)P(x, y) = \pi(y)P(y, x)$ (by reversibility). One can check that in the uniform case, this definition is equivalent to our definition in Section 2.3. Furthermore, Lemma 2.8 and Lemma 2.1 work for the weighted case [67] with this adjusted definition of congestion.

The proof of Theorem 2.1 is in fact not difficult to describe intuitively: if one finds a flow (collection of fractional paths) through the projection graph between every pair of classes (restriction chains), this flow induces a subproblem in each class Ω_i, in which each "boundary vertex"—each vertex (state) $z \in \Omega_i$ that brings in flow from a neighbor $w \in \Omega_j$—must route the flow it receives throughout Ω_i. The state z may bring in an amount of flow up to $\bar{\rho}\gamma\Delta$ from such neighbors, and z must route this flow (which we will show in the proof is at most $\bar{\rho}\gamma\Delta\pi(z)$) throughout Ω_i. By assumption, it is possible for z to route $\pi(z)$ flow throughout Ω_i with congestion at most $\rho_{\max}$, and therefore z can route the $\bar{\rho}\gamma\Delta\pi(z)$ flow throughout Ω_i with congestion at most $\bar{\rho}\gamma\Delta\rho_{\max}$. The factor of 2 in the term $1 + 2\bar{\rho}\gamma\Delta$ comes from applying the above reasoning twice: once for "inbound flow" that z brings into Ω_i, and once for "outbound flow" that z must route from Ω_i to other classes. Finally, the factor of 1 comes from routing flow between pairs of states within Ω_i.

We now make this reasoning precise:

Proof. (Proof of Theorem 2.1) Let $\{f_i, i = 1, \ldots, k\}$ be a collection of flow functions over the restriction chains with congestion $\rho_i \leq \rho_{max}$, as supposed in the theorem statement. Suppose we have a flow $\bar{f}$ with congestion $\bar{\rho}$ in the projection chain.

We construct a multicommodity flow f in the overall chain $\mathcal{M}$ as follows: for every edge $e = (x, y), x \in \Omega_i, y \in \Omega_j, i \neq j$ between restriction state spaces, let $f_{xy} = \bar{f}(i, j)Q(x, y)/\bar{Q}(i, j)$. For pairs of states $x, y \in \Omega_i$, simply use the same (fractional) paths to send flow as in f_i. Now, for *non-adjacent* pairs of states $x \in \Omega_i, y \in \Omega_j$, $i \neq j$, we will use the flow $\bar{f}$ to route the $x - y$ flow, perhaps through one or more intermediate restriction spaces. We need to consider how to route the flow *through* each intermediate restriction space. This induces a collection of subproblems over each restriction space Ω_i in which each state $z \in \Omega_i$ "brings in" and similarly "sends out" at most $\sum_{j \neq i} \sum_{w \in \Omega_j} \bar{\rho}\bar{Q}(i, j)Q(z, w)/\bar{Q}(i, j) \leq \bar{\rho}\pi(z)\gamma\Delta$ units of flow. We reuse the (fractional) paths that produce the flow with congestion ρ_{max}, scaling the resulting congestion by $\bar{\rho}\gamma\Delta$. More precisely, if $e = (x, y)$ is an edge internal to a restriction space Ω_i, let

$$\hat{f}_{zu}(x, y) = (D_{in}(z, u) + D_{out}(u, z))(f_{i,zu}(x, y) \cdot \frac{\bar{\pi}(i)}{\pi(z)\pi(u)})$$

and denote

$$\hat{f}(x, y) = \sum_{zu} \hat{f}_{zu}(x, y) = \sum_{z,u \in \Omega_i} (D_{in}(z, u) + D_{out}(u, z))(f_{i,zu}(x, y) \cdot \frac{\bar{\pi}(i)}{\pi(z)\pi(u)})$$

and $\hat{\rho}(x,y) = \frac{\hat{f}(x,y)}{Q(x,y)}$, where

$$D_{in}(z,u) = (\sum_{j \neq i} \sum_{w \in \Omega_j} \bar{f}(i,j) Q(z,w) / \bar{Q}(i,j)) \cdot \frac{\pi(u)}{\bar{\pi}(i)} \leq \bar{\rho} \frac{\pi(z)\pi(u)}{\bar{\pi}(i)} \gamma \Delta$$

is the share of the demand brought in by z to Ω_i that must be sent to u, and $D_{out}(u,z)$ is similar. This definition $\hat{f}_{zu}(x,y)$ indeed satisfies the demands $D_{in}(z,u)$ and $D_{out}(u,z)$: $f_{i,zu}$ is defined as sending $\frac{\pi(z)\pi(u)}{\bar{\pi}(i)}$ units of flow along a set of fractional paths from z to u, so the function

$$\hat{f}_{zu}(x,y) = (D_{in}(z,u) + D_{out}(u,z))(f_{i,zu}(x,y) \cdot \frac{\bar{\pi}(i)}{\pi(z)\pi(u)})$$

sends $D_{in}(z,u) + D_{out}(u,z)$ units of flow along the same set of fractional paths.

Now, we know by the definition of the congestion ρ_i produced by f_i that

$$\sum_{z,u \in \Omega_i} f_{i,zu}(x,y) \leq \rho_i Q_i(x,y),$$

where $Q_i(x,y) = \frac{\pi(x)P(x,y)}{\bar{\pi}(i)}$. Therefore

$$\hat{\rho}(x,y) = \frac{\hat{f}(x,y)}{Q(x,y)} = \frac{1}{Q(x,y)} \sum_{z,u \in \Omega_i} (D_{in}(z,u) + D_{out})(f_{i,zu}(x,y) \cdot \frac{\bar{\pi}(i)}{\pi(z)\pi(u)})$$

$$\leq \frac{1}{Q(x,y)} \sum_{z,u \in \Omega_i} 2\bar{\rho}\gamma\Delta \frac{\pi(z)\pi(u)}{\bar{\pi}(i)} (f_{i,zu}(x,y) \cdot \frac{\bar{\pi}(i)}{\pi(z)\pi(u)}) = 2\bar{\rho}\gamma\Delta \sum_{z \in \Omega_i} \frac{f_{i,zu}(x,y)}{Q(x,y)} \leq 2\bar{\rho}\gamma\Delta\rho_{max}.$$

Now, for $x, y \in \Omega_i$ and for $u \in \Omega_i, v \in \Omega_j \neq \Omega_i$, we let

$$f_{vu}(x,y) = \sum_{z \in \Omega_i} \hat{f}_{v,zu}(x,y),$$

where

$$\hat{f}_{v,zu}(x,y) = (D_{v,in}(z,u) + D_{v,out}(u,z))(f_{i,zu}(x,y) \cdot \frac{\bar{\pi}(i)}{\pi(z)\pi(u)},$$

where

$$D_{v,in}(z,u) = \frac{\pi(u)}{\bar{\pi}(i)} \sum_{k:\exists w \in \Omega_k, w \sim z} \bar{f}_{j,i}(k,i) \cdot \frac{Q(\Omega_k, z)}{\bar{Q}(k,i)} \cdot \frac{\pi(v)}{\bar{\pi}(j)}$$

and $D_{v,out}(u,z)$ is symmetric. The function f_{vu} indeed is a valid flow sending $\pi(u)\pi(v)$ units from v to u, and also that

$$\sum_{j \neq i} \sum_{v \in \Omega_j} \hat{f}_{v,zu} = \hat{f}_{zu}.$$

Therefore

$$\rho(x,y) = \sum_{u \in \Omega_i} \sum_{v \notin \Omega_i} \frac{f_{vu}(x,y)}{Q(x,y)} = \sum_{u,z \in \Omega_i} \sum_{v} \frac{\hat{f}_{v,zu}(x,y)}{Q(x,y)} = \sum_{u,z \in \Omega_i} \frac{\hat{f}_{zu}(x,y)}{Q(x,y)} = \hat{\rho}(x,y) \leq 2\bar{\rho}\gamma\Delta\rho_{max}.$$

Finally, in the term $\rho(x,y)$ we have only considered u, v flow where u, v lie in different classes. Adding the congestion $\rho_i \leq \rho_{max}$ produced by reusing the flow f_i for pairs $u, v \in \Omega_i$ justifies the expression

$$(1 + 2\bar{\rho}\gamma\Delta)\rho_{max}.$$

□

Lemma 2.9 and Lemma 2.10 are now immediate, as is Lemma 2.11.

2.6 Proof that the conditions of Lemma 2.12 imply rapid mixing

In this section we prove Lemma 2.12:

Lemma 2.12. *Let $\mathcal{F} = \{\mathcal{M}_1, \mathcal{M}_2, \ldots\}$ be an infinite family of connected graphs, parameterized by a value n. Suppose that for every graph $\mathcal{M}_n = (\mathcal{V}_n, \mathcal{E}_n) \in \mathcal{F}$, for $n \geq 2$, the vertex set $\mathcal{V}_n$ can be partitioned into a set $\mathcal{S}_n$ of classes inducing subgraphs of $\mathcal{M}_n$ that satisfy the* *following conditions:*

1. *Each subgraph is isomorphic to a Cartesian product of one or more graphs* $\mathcal{C}(T) \cong \mathcal{M}_{i_1} \square \cdots \mathcal{M}_{i_k}$*, where for each such graph* $\mathcal{M}_{i_j} \in \mathcal{F}$*,* $i_j < n$*.*

2. *The number of classes is* $n^{O(1)}$*.*

3. *For every pair of classes* $\mathcal{C}(T), \mathcal{C}(T') \in \mathcal{S}_n$*, the set of edges between the subgraphs induced by the two classes is a matching of size at least* $\frac{|\mathcal{C}(T)||\mathcal{C}(T')|}{|\mathcal{V}_n|}$*.*

4. *Given a pair of classes* $\mathcal{C}(T), \mathcal{C}(T') \in \mathcal{S}_n$*, there exists a graph* $\mathcal{M}_i$ *in the Cartesian product* $\mathcal{C}(T)$*, and a class* $\mathcal{C}(U) \in \mathcal{S}_i$ *within the graph* $\mathcal{M}_i$*, such that the set of vertices in* $\mathcal{C}(T)$ *having a neighbor in* $\mathcal{C}(T')$ *is precisely the set of vertices in* $\mathcal{C}(T)$ *whose projection onto* $\mathcal{M}_i$ *lies in* $\mathcal{C}(U)$*. Furthermore, no class* $\mathcal{C}(U)$ *within* $\mathcal{M}_i$ *is the projection of more than one such boundary.*

Suppose further that $|\mathcal{V}_1| = 1$*. Then the expansion of* $\mathcal{M}_n$ *is* $\Omega(1/(\kappa(n)n))$*, where* $\kappa(n) = \max_{1 \le i \le n} |\mathcal{C}(S_i)|$ *is the maximum number of classes in any* $\mathcal{M}_i, i \le n$*.*

We will use the fact that one can prove an analogue of Lemma 2.7 for multicommodity flows—namely one that does not lose a factor of two. We prove this in Appendix A.4:

Lemma 2.15. *Let* $J = G \square H$*. Given multicommodity flows* g *and* h *in* G *and* H *respectively with congestion at most* ρ*, there exists a multicommodity flow* f *for* J *with congestion at most* ρ*.*

We will construct a "good flow"—that is, a uniform multicommodity flow with polynomially bounded congestion—in any $\mathcal{M}_n \in \mathcal{F}$ satisfying the conditions of Lemma 2.12, via an inductive process. The base case, $|\mathcal{V}| = 1$, is trivial. For the inductive hypothesis, we assume that for all $i < n$, there exists a good flow in $\mathcal{M}_i$. For the inductive step, we begin by combining Lemma 2.15 with Condition 1 to obtain a good flow in each $\mathcal{C}(T)$: since each class is a product of smaller graphs $\{\mathcal{M}_i\}$ in the same family, the inductive assumption that those smaller graphs have good flows carries through to $\mathcal{C}(T)$ by Lemma 2.15.

The more difficult part of the inductive step is then to route flow between pairs of vertices that lie in different classes. We now introduce machinery, in the form of *multi-way single-commodity flows*, that we will apply to the boundary set structure in Condition 4 to find the right paths for these pairs.

Define a *multi-way single-commodity flow* (MSF), given a graph $G = (V, E)$, with *source* set $S \subseteq V$ and *sink* set $T \subseteq V$, and a set of "surplus" and "deficit" amounts $\sigma : S \to \mathbb{R}$ and $\delta : T \to \mathbb{R}$, as a flow $f : A(E) \to \mathbb{R}$ in G, such that:

1. the net flow out of each vertex $s \in S \setminus T$ is $\sigma(s)$,
2. the net flow into each vertex $t \in T \setminus S$ is $\delta(t)$,
3. the net flow out of each vertex $u \in S \cap T$ is $\sigma(u) - \delta(u)$, and
4. the net flow into (out of) each vertex $u \in V \setminus (S \cup T)$ is zero.

Denote the MSF as the tuple $\rho = (f, S, T, \sigma, \delta)$. (Here $A(E)$ is the directed arc set obtained by creating directed arcs (u, v) and (v, u) for each edge $\{u, v\} \in E$.) When σ and δ are constant functions, abuse notation and denote by σ and δ their values. Intuitively, an MSF describes sending flow from some set of vertices (the source set) in a graph to another set (the sink set). It differs from a multicommodity flow in that it is not important that every vertex in S send flow to every vertex in T. For instance, in a bipartite graph, if the source set and sink set are the two sides of the bipartition, and all surpluses and demands are one, it suffices to direct the flow across a matching.

It will also be useful to talk about an *MSF problem*, in which we are given surpluses and demands but need to find the actual flow function. Define a *multi-way single-commodity flow problem* (MSF problem) as a tuple $\pi = (S, T, \sigma, \delta)$, where S, T, σ, δ are as in the definition of an MSF, but no flow function f is specified.

(One could alternatively formulate an MSF problem as a more familiar $s - t$ flow problem by adding extra vertices and edges. However, the definition of an MSF will make our flow construction more convenient.)

The main lemma of this section is as follows:

Lemma 2.16. *Let a graph $\mathcal{M}_n \in \mathcal{F}$ be given, with $n > 1$ and $\mathcal{F}$ satisfying the conditions of Lemma 2.12. Suppose that for all $1 \leq i < n$, the graph $\mathcal{M}_i$ has a uniform multicommodity flow with congestion at most ρ, for some $\rho > 0$. Then there exists a uniform multicommodity flow in $\mathcal{M}_n$ with congestion at most $\rho + \kappa$, where $\kappa = |\mathcal{S}_n|$ is the number of classes in the partition described in Lemma 2.12.*

Lemma 2.16 forms the inductive step of an argument that proves Lemma 2.12.

To prove Lemma 2.16, we will start by partitioning $\mathcal{M}_n$ into the classes $\mathcal{S}_n$ as described in Lemma 2.12. Now consider any vertex $s \in \mathcal{C}(T)$, for a given class $\mathcal{C}(T) \in \mathcal{S}_n$, and consider any other class $\mathcal{C}(T') \neq \mathcal{C}(T)$. Consider a multi-way single-commodity flow problem

$$\pi_s = (\{s\}, \mathcal{C}(T'), \sigma_s = |\mathcal{C}(T')|, \delta_s = 1).$$

We will "solve" this problem—construct a flow function f_s that satisfies the surpluses and demands of the problem. Notice that to solve π_s is to send a unit of flow from s to every $t \in \mathcal{C}(T')$. Thus if we construct such a function f_s for every $s \in \mathcal{C}(T)$, and construct similar flows for every pair of classes $\mathcal{C}(T), \mathcal{C}(T')$, we will have constructed a uniform multicommodity flow in $\mathcal{M}_n$. We will do precisely this, then analyze the congestion of the sum of these flow functions.

To construct f_s, we will express the problem π_s as the composition of four MSF problems

$$\begin{aligned}
\pi_{shuf} &= (\{s\}, \mathcal{C}(T), \sigma_{shuf} = \sigma_s = |\mathcal{C}(T')|, \delta_{shuf} = \frac{|\mathcal{C}(T')|}{|\mathcal{C}(T)|}),\\
\pi_{conc} &= (\mathcal{C}(T), \mathcal{B}_{T'}(T), \sigma_{conc} = \delta_{shuf}, \delta_{conc} = \frac{|\mathcal{C}(T')|}{|\mathcal{B}_{T'}(T)|}),\\
\pi_{tran} &= (\mathcal{B}_{T'}(T), \mathcal{B}_{T}(T'), \sigma_{tran} = \delta_{tran} = \delta_{conc} = \frac{|\mathcal{C}(T')|}{|\mathcal{B}_{T'}(T)|} = \frac{|\mathcal{C}(T')|}{|\mathcal{B}_{T}(T')|},\\
\pi_{dist} &= (\mathcal{B}_{T}(T'), \mathcal{C}(T'), \sigma_{dist} = \delta_{tran}, \delta_{dist} = \delta_s = 1).
\end{aligned}$$

(Here we have defined the matching $\mathcal{E}(T, T')$ and the boundary set $\mathcal{B}_{T'}(T)$ for the general family $\mathcal{F}$ in the same way we defined $\mathcal{E}^*(T, T')$ and $\mathcal{B}^*_{n,T'}(T)$ for the associahedron. We have implicitly used the equality $|\mathcal{B}_{T'}(T)| = |\mathcal{E}(T, T')| = |\mathcal{B}_{T}(T')|$, which follows from the assumption in Condition 3 that these boundary edges form a matching.)

Remark 2.2. *Comparing σ and δ values and comparing source and sink sets shows that if one specifies flow functions solving the four subproblems $\pi_{shuf}, \pi_{conc}, \pi_{tran}, \pi_{dist}$, one can take the arc-wise sum of these functions as a solution to the original MSF problem π_s.*

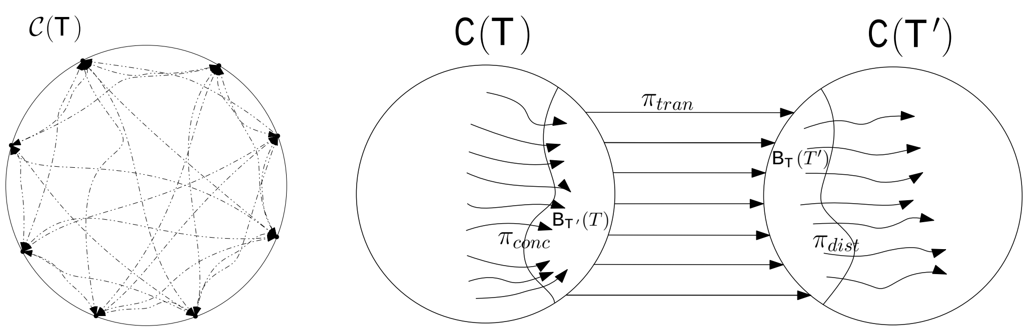

Figure 2.6: The decomposition of the MSF problem π_s. Left: π_{shuf}, solved in aggregate for all $s \in \mathcal{C}(T)$ by a uniform multicommodity flow in $\mathcal{C}(T)$. Right: the problems π_{conc}, π_{tran}, and π_{dist}, in which the (single) commodity from $s \in \mathcal{C}(T)$ begins uniformly spread throughout $\mathcal{C}(T)$. The flow must then be concentrated on the boundary $\mathcal{B}_{T'}(T)$ (for π_{conc}), sent to $\mathcal{C}(T')$ (for π_{tran}), and distributed uniformly throughout $\mathcal{C}(T')$ (for π_{dist}).

Intuitively, π_{shuf} describes the problem of "shuffling," or distributing evenly throughout $\mathcal{C}(T)$, the flow that s must send to vertices in $\mathcal{C}(T')$. We solve this subproblem in aggregate

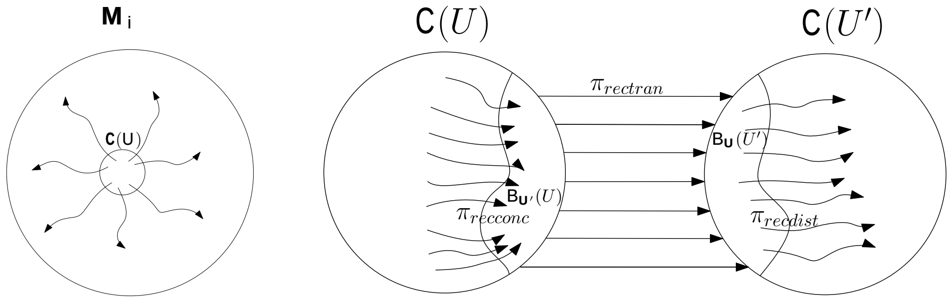

Figure 2.7: Left: An illustration of π_{rec}, to which we reduce π_{dist} in Lemma 2.19, in which $\mathcal{C}(U)$ must distribute its flow throughout $\mathcal{M}_i$, inducing a corresponding distribution of flow from $\mathcal{B}_T(T')$ throughout $\mathcal{C}(T)$, by the isomorphism in Condition 4.

Right: a decomposition of the flow $\pi_{rec,U,U'}$ from Lemma 2.20, which decomposes into $\pi_{conc}, \pi_{tran}, \pi_{dist}$, which are similar to $\pi_{conc}, \pi_{tran}, \pi_{dist}$ and thus admit a recursive decomposition (Lemma 2.21).

for every $s \in \mathcal{C}(T)$ by applying the inductive hypothesis and Lemma 2.15, obtaining a uniform multicommodity flow f_T in $\mathcal{C}(T)$ with combined congestion at most ρ. We then let $f_{shuf} = f_{s,shuf}$ be the part of f_T that sends flow just for s—since f_T can be written as a sum $\sum_{s \in \mathcal{C}(T)} f_{s,shuf}$, where $f_s = \sum_{s' \in \mathcal{C}(T)} f_{s,s'}$, where $f_{s,s'}$ is the single-commodity flow function already defined for the s, s' pair.

Thus we prove the following:

Lemma 2.17. *The MSF subproblem π_{shuf} as defined in this section for any two classes $\mathcal{C}(T), \mathcal{C}(T') \in \mathcal{S}_n$, with $\mathcal{S}_n$ partitioning $\mathcal{M}_n \in \mathcal{F}$, $n > 1$, with $\mathcal{F}$ satisfying the conditions of Lemma 2.12, can be solved in aggregate for all $s \in \mathcal{C}(T)$ and for all $\mathcal{C}(T') \neq \mathcal{C}(T)$, while generating at most congestion ρ—where ρ is as in the statement of Lemma 2.16.*

Proof. As in the discussion leading to this lemma, the uniform multicommodity flow f_T in $\mathcal{C}(T)$ given by the application of the inductive hypothesis and Lemma 2.15 has congestion at most ρ. More precisely, in this uniform multicommodity flow, the *un-normalized* congestion (as we have previously defined), is at most $\rho|\mathcal{C}(T)|$. Under the definition of $\sigma_{shuf} = |\mathcal{C}(T')|$, and summing over all s and over all $\mathcal{C}(T')$, what we in fact need is a *scaled* version of f_T—in which the amount of flow sent between each pair of vertices $s, s' \in \mathcal{C}(T)$, and therefore the overall congestion across each edge within $\mathcal{C}(T)$, is scaled so that each s sends to each s'

$$\frac{|\mathcal{V}_n|}{|\mathcal{C}(T)|}$$

units of flow, instead of just one unit.

Thus we increase the un-normalized congestion from $\rho|\mathcal{C}(T)|$ to $\rho|\mathcal{V}_n|$. However, since we are now considering congestion within the graph $\mathcal{M}_n$ instead of the induced subgraph $\mathcal{C}(T)$, the *normalized* congestion ρ does not change. □

We define f_{tran}—solving the problem π_{tran} of transmitting the flow from the boundary edges $\mathcal{B}_{T'}(T) \subseteq \mathcal{C}(T)$ to $\mathcal{B}_T(T') \subseteq \mathcal{C}(T')$ in the natural way: for each directed arc $(u, v) \in \mathcal{E}(T, T')$, let $f(u, v) = \sigma_{tran} = \delta_{tran}$. Summing the resulting flow over every $s \in \mathcal{C}(T)$ gives (normalized) congestion

$$\frac{1}{|\mathcal{V}_n|}|\mathcal{C}(T)|\sigma_{tran} = \frac{|\mathcal{C}(T)||\mathcal{C}(T')|}{|\mathcal{E}(T,T')||\mathcal{V}_n|} \leq 1,$$

where the inequality follows from Condition 3 of Lemma 2.12.

Thus we have proven:

Lemma 2.18. *The MSF subproblem π_{tran} as defined in this section for a given pair of classes $\mathcal{C}(T), \mathcal{C}(T')$ can be solved by a function f_{tran} while generating at most congestion one—when summing over all $s \in \mathcal{C}(T)$.*

It remains to solve π_{conc} and π_{dist}. We observe that these two problems are of the same form up to reversal of flows: π_{conc} describes beginning with flow from a single commodity distributed equally throughout $\mathcal{C}(T)$, and ending with that flow concentrated (uniformly) within the boundary $\mathcal{B}_{T'}(T)$. On the other hand, π_{dist} describes just the reverse process within $\mathcal{C}(T')$. We will construct π_{dist} within $\mathcal{C}(T')$, in aggregate, for all $s \in \mathcal{C}(T)$; the form of this construction will give a symmetric construction for π_{conc} within $\mathcal{C}(T)$.

Our construction is recursive, and it is here that we use the boundary set structure in Condition 4: we use this condition to reduce the problem π_{dist} to a problem

$$\pi_{rec} = (\mathcal{C}(U) \in \mathcal{S}_i, \mathcal{M}_i, \sigma_{rec} = \sigma_{dist} = \frac{|\mathcal{C}(T')|}{|\mathcal{B}_T(T')|}, \delta_{rec} = \delta_s = 1).$$

We obtain a reduction that allows us to pass from the problem π_{dist} to the problem π_{rec}: by Condition 4, we have that the projection of $\mathcal{B}_T(T')$ onto some $\mathcal{M}_i$ in the Cartesian product $\mathcal{C}(T) \cong \Box_i \mathcal{M}_i$ is precisely $\mathcal{C}(U)$, for some $\mathcal{C}(U) \in \mathcal{S}_i$. Therefore, if one views π_{dist} as a process of distributing flow throughout $\mathcal{C}(T')$, the flow is initially uniform within every copy of $\mathcal{M}_j$, for all graphs $\mathcal{M}_j$ in the product other than $\mathcal{M}_i$. It therefore suffices to distribute the flow within each copy of $\mathcal{M}_i$, in which it is initially concentrated uniformly within $\mathcal{C}(U)$.

Thus we prove:

Lemma 2.19. *The problem π_{dist} described in this section can be solved by any flow function f_{rec} that solves the MSF problem π_{rec} as described in this section. Furthermore, if f_{rec} generates congestion at most ρ, then f_{dist} also generates congestion at most ρ. The problem π_{conc} is of the same form as the reversal of π_{dist} and therefore is solved by a flow function similar to f_{rec}, also with congestion at most ρ.*

Proof. The first part of the lemma statement—the reduction—is justified by the discussion leading to this lemma. That is, we can construct a flow function f_{dist} that solves π_{dist} as the arc-wise sum of many separate (but identical) functions f_{rec}—one such function within each copy of $\mathcal{M}_i$ in the Cartesian product $\mathcal{C}(T')$.

The preservation of the congestion bound ρ follows from the fact that these functions are defined over disjoint sets of arcs, since the copies of $\mathcal{M}_i$ are all mutually disjoint.

Finally, the symmetry of π_{conc} and π_{dist} follows from the discussion leading to this lemma. □

Furthermore, notice that in π_{rec}, we have the problem of flow that is initially concentrated uniformly within a class $\mathcal{C}(U) \in \mathcal{S}_i$, such that an equal amount must be distributed to each vertex $t \in \mathcal{C}(U')$, for every class $\mathcal{C}(U') \in \mathcal{S}_i$. Let $\pi_{rec,U,U'}$ be this problem of sending the flow that is bound for vertices in $\mathcal{C}(U')$. We now have:

Lemma 2.20. *The problem π_{rec}, defined with respect to $s \in \mathcal{C}(T)$ and $\mathcal{C}(U) \in \mathcal{M}_i$, can be decomposed into a collection of problems $\pi_{rec,U,U'}$, one for each $\mathcal{C}(U') \in \mathcal{S}_i$.*

Proof. Following the discussion leading to this lemma, it suffices to define

$$\pi_{rec,U,U'} = (\mathcal{C}(U), \mathcal{C}(U'), \sigma_{rec,U,U'} = \sigma_{rec}\frac{|\mathcal{C}(U')|}{|\mathcal{V}_i|}, \delta_{rec,U,U'} = \delta_{rec}).$$

The definitions of $\sigma_{rec,U,U'}$ and $\delta_{rec,U,U'}$ are indeed correct (achieve the decomposition of π_{rec} stated in the lemma): $\delta_{rec,U,U'} = \delta_{rec}$ obviously agrees with π_{rec}, and one can check that

$$\sum_{U'} \sigma_{rec,U,U'} = \sigma_{rec},$$

as needed. □

Furthermore, since $\mathcal{M}_i$ is in the family $\mathcal{F}$ and thus satisfies the conditions of Lemma 2.12, the problem $\pi_{rec,U,U'}$ is of the same form as our original problem ($\pi_{conc}, \pi_{tran}, \pi_{dist}$) of sending flow that was uniformly concentrated within $\mathcal{C}(T)$ to vertices in $\mathcal{C}(T')$, where $\mathcal{C}(T), \mathcal{C}(T') \in \mathcal{S}_n$ were classes in the original graph $\mathcal{M}_n$.

That is, just as we decomposed the original problem π_s into the "concentration" problem π_{conc}, the "transmission" problem π_{tran}, and the "distribution" problem π_{dist}, we can recursively decompose $\pi_{rec,U,U'}$ in the same fashion. In particular, we can solve the resulting transmission problem, in the same fashion as before. Furthermore, recall that the original problem π_s was defined with respect to a single $s \in \mathcal{C}(T)$. We claim that even after solving the transmission problem for *all* $s \in \mathcal{C}(T)$, we obtain congestion at most one.

Furthermore:

Remark 2.3. *Summing* σ_{dist} *over all* $s \in \mathcal{C}(T)$ *produces*

$$|\mathcal{C}(T)|\sigma_{dist} = \frac{|\mathcal{C}(T)||\mathcal{C}(T')|}{|\mathcal{B}_{T'}(T)}| \leq |\mathcal{V}_n|$$

flow "concentrated" within each boundary vertex.

These facts, we claim, indicate that the congestion does not increase as we pass from one level of recursion to the next. Remark 2.3 implies that in this reduction, we have within $\mathcal{M}_i$ a problem similar to the original problem in $\mathcal{M}_n$: that is, in the original problem, the overall flow construction, we have a collection of MSF problems $\{\{\pi_s\}|s \in \mathcal{V}_n\}$, in which each $s \in \mathcal{V}_n$ is "charged" with initial surplus values $\sum_{T'} |\mathcal{C}(T')| = |\mathcal{V}_n|$. What we have now is a single MSF problem, in $\mathcal{M}_i$, in which each $u \in \mathcal{B}_T(T') \cap \mathcal{M}_i = \mathcal{C}(U)$ has a surplus (summing over all $s \in \mathcal{C}(T)$) of $|\mathcal{C}(T)|\sigma_{dist} \leq |\mathcal{V}_n|$, by Remark 2.3. Furthermore, just as the original problem of distributing $|\mathcal{C}(T)||\mathcal{V}_n|$ outbound flow from vertices $s \in \mathcal{C}(T)$ throughout $\mathcal{M}_n$ induces the subproblem of sending $|\mathcal{C}(T)||\mathcal{C}(T')|$ flow from $\mathcal{C}(T)$ to $\mathcal{C}(T')$ (and thus by Condition 3 producing $\leq |\mathcal{V}_n|$ flow across each $\mathcal{E}(T, T')$ edge), similarly the subproblem of distributing $|\mathcal{V}_n|$ flow from each $u \in \mathcal{C}(T)$ throughout $\mathcal{M}_i$ induces the subproblem of sending

$$|\mathcal{C}(U)||\mathcal{V}_n|\frac{|\mathcal{C}(U')|}{|V(\mathcal{M}_i)|}$$

flow from $\mathcal{C}(U)$ to each $\mathcal{C}(U')$ in $\mathcal{M}_i$, since each $\mathcal{C}(U')$ receives a portion of the $|\mathcal{C}(U)||\mathcal{V}_n|$ outbound flow from $\mathcal{C}(U)$ that is proportional to the cardinality of $\mathcal{C}(U')$ within $V(\mathcal{M}_i)$. This generates at most

$$\frac{|\mathcal{C}(U)||\mathcal{V}_n|\frac{|\mathcal{C}(U')|}{|V(\mathcal{M}_i)|}}{|\mathcal{E}(U, U')} \leq \frac{|V(\mathcal{M}_i)||\mathcal{V}_n|}{|V(\mathcal{M}_i)|} = |\mathcal{V}_n|$$

flow across the matching edges $\mathcal{E}(U, U')$, producing (normalized) congestion one, and matching the flow across $\mathcal{E}(T, T')$. (Here, in the first inequality, we have applied Condition 3 to the matching $\mathcal{E}(U, U')$.) Thus we have a recursive decomposition in which the congestion does not increase in the recursion.

Lemma 2.21. *Let the problem $\pi_{rec,U,U'}$ be defined as in this section, with respect to $s \in \mathcal{C}(T)$, class $\mathcal{C}(U), \mathcal{C}(U')$ being classes in $\mathcal{M}_i$, with $\mathcal{M}_i$ a graph in the Cartesian product $\mathcal{C}(T)$.*

Then $\pi_{rec,U,U'}$ can be recursively decomposed into $\pi_{recconc}, \pi_{rectran}$, and $\pi_{recdist}$, with each problem solved by a respective flow $f_{recconc}, f_{rectran}, f_{recdist}$, such that:

(i) The sum total congestion incurred by all of the $f_{rectran}$ subproblems induced by all $s \in \mathcal{C}(T)$, is at most one, and

(ii) $\pi_{recccone}$ and $\pi_{recdist}$ are similar to the problems π_{dist} and π_{conc} described in this section and thus admit a recursive decomposition as in Lemma 2.19, and

(iii) the demand $\delta_{recconc}$ is upper-bounded by δ_{conc}, the surplus value in the original concentration problem π_{conc}; similarly, $\sigma_{recdist} \leq \sigma_{dist}$.

Proof. We prove (ii) first: define

$$\pi_{recconc} = \left(\mathcal{C}(U), \mathcal{B}_{U'}(U), \sigma_{recconc} = \sigma_{rec}, \delta_{recconc} = \sigma_{recconc}\frac{|\mathcal{C}(U)|}{|\mathcal{B}_{U'}(U)|}\right),$$
$$\pi_{rectran} = (\mathcal{B}_{U'}(U), \mathcal{B}_U(U'), \sigma_{rectran} = \delta_{rectran} = \delta_{recconc}),$$
$$\pi_{recdist} = \left(\mathcal{B}_U(U'), \mathcal{C}(U'), \sigma_{recdist} = \delta_{rectran}, \delta_{recdist} = \sigma_{recdist}\frac{|\mathcal{B}_U(U')|}{|\mathcal{C}(U')|}\right).$$

Comparing source and sink sets, and comparing σ and δ functions shows that π_{rec} decomposes into $\pi_{recconc}, \pi_{rectran}$, and $\pi_{recdist}$. Each class $\mathcal{C}(U)$ and $\mathcal{C}(U')$ decomposes as a Cartesian product satisfying Condition 1 in Lemma 2.12, and similarly the boundary sets $\mathcal{B}_{U'}(U), \mathcal{B}_U(U')$ satisfy Condition 4. Thus exactly the same form of decomposition used to reduce the original π_{dist} and π_{conc} to π_{rec} also works for $\pi_{recconc}$ and $\pi_{recdist}$. We can thus recursively construct $f_{recconc}$ and $f_{recdist}$, proving (ii).

For (i), we need to define $f_{rectran}$ and to bound the resulting congestion.

Define $f_{rectran}$ in the same natural way we defined f_{tran}: simply assign $\sigma_{rectran} = \delta_{rectran}$ to each arc.

We observe that

$$\sigma_{recconc} = \sigma_{rec,U,U'} = \sigma_{rec}\frac{|\mathcal{C}(U')|}{|\mathcal{V}_i|} = \sigma_{dist}\frac{|\mathcal{C}(U')|}{|\mathcal{V}_i|} = \frac{|\mathcal{C}(T')|}{|\mathcal{B}_T(T')|}\cdot\frac{|\mathcal{C}(U')|}{|\mathcal{V}_i|},$$

by the definitions of the MSF problems we have given in this section. Also,

$$\sigma_{rectran} = \delta_{rectran} = \delta_{recconc} = \sigma_{recconc}\frac{|\mathcal{C}(U)|}{|\mathcal{B}_{U'}(U)|}.$$

Combining these facts gives

$$\sigma_{rectran} = \frac{|\mathcal{C}(T')|}{|\mathcal{B}_T(T')|}\cdot\frac{|\mathcal{C}(U')|}{|\mathcal{V}_i|}\cdot\frac{|\mathcal{C}(U)|}{|\mathcal{B}_{U'}(U)|} \leq \frac{|\mathcal{C}(T')|}{|\mathcal{E}(T',T)|} = \sigma_{tran},$$

where the inequality follows from the fact that the matching $\mathcal{E}(U,U')$ satisfies Condition 3 of Lemma 2.12.

Now, to obtain the un-normalized congestion $\bar{\rho}_{rectran}$ that results from $f_{rectran}$, we sum over all $s \in \mathcal{C}(T)$, scaling the above quantity by a factor of $|\mathcal{C}(T)|$, giving

$$\bar{\rho}_{rectran} = |\mathcal{C}(T)|\sigma_{rectran} = |\mathcal{C}(T)|\sigma_{tran} = \frac{|\mathcal{C}(T)||\mathcal{C}(T')|}{|\mathcal{E}(T,T')|} \leq |\mathcal{V}_n|,$$

where we have again applied Condition 3 of Lemma 2.12.

Thus we obtain *normalized* congestion at most $\frac{|\mathcal{V}_n|}{|\mathcal{V}_n|} \leq 1$, proving (i).

For (iii), claim (i) also implies that the congestion does not increase in the recursive decomposition given by (ii)—that is, passing from π_{dist}, to π_{rec}, to $\pi_{rec,U,U'}$, to $\pi_{recdist}$, preserves the bound

$$\sigma_{recdist} \leq \sigma_{dist}.$$

The analogous fact for $\sigma_{recconc}$ is symmetric. □

We now have all the pieces we need to prove Lemma 2.16:

Lemma 2.16. *Let a graph $\mathcal{M}_n \in \mathcal{F}$ be given, with $n > 1$ and $\mathcal{F}$ satisfying the conditions of Lemma 2.12. Suppose that for all $1 \leq i < n$, the graph $\mathcal{M}_i$ has a uniform multicommodity flow with congestion at most ρ, for some $\rho > 0$. Then there exists a uniform multicommodity flow in $\mathcal{M}_n$ with congestion at most $\rho + \kappa$, where $\kappa = |\mathcal{S}_n|$ is the number of classes in the partition described in Lemma 2.12.*

Proof. To construct the desired uniform multicommodity flow, it suffices to construct, for every $\mathcal{C}(T), \mathcal{C}(T') \in \mathcal{S}_n$ and for every $s \in \mathcal{C}(T)$, the flow f_s solving the MSF problem π_s. As shown in this section, π_s decomposes (Remark 2.2) as the subproblems $\pi_{shuf}, \pi_{conc}, \pi_{tran}$, and π_{dist}.

For π_{shuf}, summing over all $s \in \mathcal{C}(T)$ and over all $s \in \mathcal{C}(T')$, the sum of the f_{shuf} flows given by the inductive hypothesis and the Cartesian flow structure (Lemma 2.15) of $\mathcal{C}(T)$ gives congestion at most ρ, by Lemma 2.17.

For a given $\mathcal{C}(T), \mathcal{C}(T')$ pair, again summing over all $s \in \mathcal{C}(T')$, we obtain flows f_{tran} for π_{tran} whose sum is congestion one, by Lemma 2.18.

Dividing π_{dist} (and symmetrically π_{conc}) into copies of the π_{rec} problem as in Lemma 2.19, and further dividing each π_{rec} into problems $\pi_{rec,U,U'}$ (by Lemma 2.20), each of which we further divide into $\pi_{recconc}, \pi_{rectran}$, and $\pi_{recdist}$. Furthermore, by Lemma 2.21, these subproblems are of the same form as π_{conc}, π_{tran}, and π_{dist}, with the natural solution $f_{rectran}$ to the "transmission" problem π_{tran} being of the same form as f_{tran} and producing, like f_{tran}, overall congestion one after summing over all $s \in \mathcal{C}(T)$.

We then recursively decompose $\pi_{recconc}$ and $\pi_{recdist}$ in the same fashion as we did π_{conc} and π_{dist}, with, by Lemma 2.21, congestion one in the transmission problems at each level of recursion. Since all flow produced by solving the subproblems in this decomposition is counted by the transmission flows, and since (it is easy to see) each arc occurs in only one such transmission flow, we obtain overall congestion one for $\pi_{rec,U,U'}$.

Recall that $\pi_{rec,U,U'}$ is defined with respect to a given $\mathcal{C}(T), \mathcal{C}(T')$ pair, where $\mathcal{C}(U)$ is determined by $\mathcal{C}(T)$, as a class within the graph $\mathcal{M}_i$, within the Cartesian product $\mathcal{C}(T) \cong \Box_j \mathcal{M}_j$. Thus we must sum this bound of congestion one for $f_{rec,U,U'}$ over all $\mathcal{C}(U') \in \mathcal{S}_i$. By assumption $|\mathcal{S}_i| \leq \kappa$, so we obtain κ flows each with congestion one, giving overall congestion at most κ.

One may worry that the κ^2 pairs of classes exchanging flow may produce κ^2 congestion, since we do obtain κ^2 subproblems. Fortunately, we can justify the κ bound as follows: consider κ MSF problems instead of κ^2 problems. In each of the κ MSF problems, a given class $\mathcal{C}(T)$ must send flow to *all* other classes. This introduces some asymmetry, as the concentration flow within $\mathcal{C}(T)$ involves only a single commodity, while the distribution flow within $\mathcal{C}(T)$ involves $\kappa - 1$ commodities. Thus we can break this distribution flow into $\kappa - 1$ recursive distribution flows that each involve a single commodity distributed throughout $\mathcal{C}(T)$ from $\mathcal{B}_{T'}(T)$ for some $\mathcal{C}(T')$.

The concentration flow takes slightly more work: it involves a single commodity but induces a subproblem in which every pair of subclasses within $\mathcal{C}(T)$ must exchange a unit of flow. Consider the boundary sets $\mathcal{B}_{T'}(T)$ and $\mathcal{B}_{T''}(T)$ along which $\mathcal{C}(T)$ must send flow to any two of the other classes $\mathcal{C}(T')$ and $\mathcal{C}(T'')$. By Condition 4, we know that all of this flow occurs between subclasses within copies of smaller flip graphs. Say these subclasses are $\mathcal{C}(U')$

and $\mathcal{C}(U'')$. Notice that we do not need to send flow in both directions, *because we have only a single commodity*. Only the *amount* of flow sent matters. This observation gives us a convenient subproblem in which for each pair of subclasses $\mathcal{C}(U'), \mathcal{C}(U'')$, one class sends to the other an amount of flow that, by Condition 3, generates congestion at most one, producing appropriate recursive subproblems without an increase in congestion. □

Lemma 2.16 forms the inductive step of Lemma 2.12 (with a trivial base case), and thus we have proven Lemma 2.12.

2.7 k-angulations of convex point sets: quasipolynomial mixing

2.7.1 Generalizing triangulations

As we stated in the introduction, one can generalize triangulations to k-angulations. We do so in more detail here. A *quadrangulation* of a point set is a maximal subdivision of the point set into quadrilaterals, where each quadrilateral has all of its vertices in the point set. Consider P_{2n+2}, the regular polygon with $2n+2$ vertices. We denote by $K_{4,2n+2}$ the graph whose vertex set is the set of all quadrangulations of P_{2n+2}, and whose edges are the flips between quadrilaterals. Here, a flip is defined as follows: each diagonal belongs to two quadrilaterals, which together form a hexagon. Replace the diagonal with one of the other two diagonals in the hexagon. (Thus each diagonal in a quadrangulation can be flipped in two possible ways [14].)

There is a polytope, analogous to the associahedron, known as the *accordiohedron* [38, 5], whose vertices and edges are those of a *subgraph* of $K_{4,2n+2}$. However, we ignore this polytope and just consider the graph $K_{4,2n+2}$.

We refer to a *k-angulation* of a point set as a maximal subdivision of the point set into k-gons, each of whose vertices all belong to the point set. A bijection exists [36] between the k-angulations of $P_{(k-2)n+2}$ and the set of all $k-1$-ary plane trees with n internal nodes.

It is easy to generalize the definition of a flip between triangulations or quadrangulations to a flip between k-angulations: each diagonal in a k-angulation belongs to two k-gons, which together form a $2k-2$-gon. A flip then consists of replacing this diagonal—which connects two opposite vertices in the $2k-2$-gon—with one of the $k-2$ other such diagonals.

We generalize the associahedron graph K_n as follows: Define the *k-angulation flip graph* $K_{k,(k-2)n+2}$ as the graph whose vertices represent the k-angulations of $P_{(k-2)n+2}$, and whose edges represent the flips between k-angulations.

Define the *k-angulation flip walk* as the natural Markov chain whose state space is $K_{k,(k-2)n+2}$.

2.7.2 (Generalized) Catalan numbers

The usual notation for Catalan numbers is simply C_n; we will now consider a generalization: [62, 46, 36] Let $C_{k,n} = \frac{1}{(k-2)n+1}\binom{(k-1)n}{n}$. These numbers, which generalize Catalan numbers, are similar but not identical to the *Fuss-Catalan* numbers.

We will use the following fact in proving that the random walk on k-angulations mixes in quasipolynomial time:

Lemma 2.22. *[46, 36] The number of k-angulations of the convex $(k-2)n+2$-gon is counted by $C_{k,n}$.*

One can show using Stirling's formula, and in particular a result by Robbins [64], that:

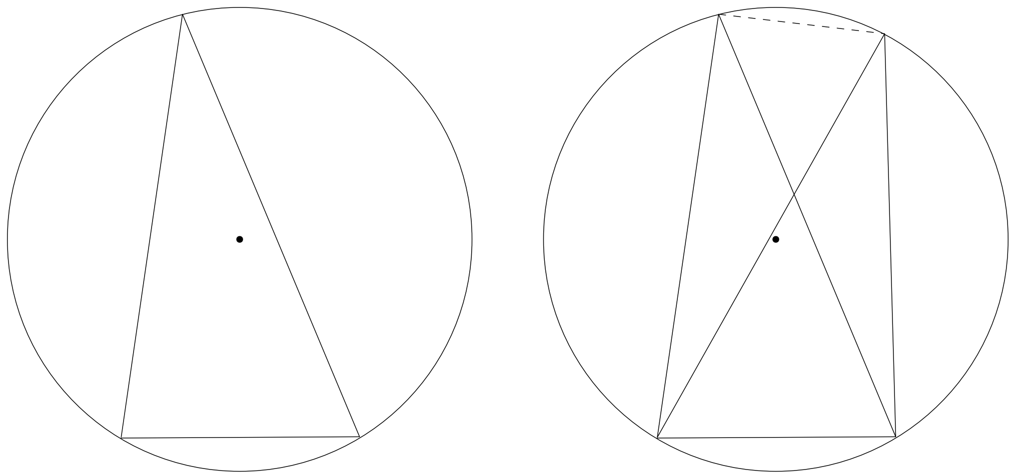

Figure 2.8: Left: a class $\mathcal{C}(T)$ in $K_{3,n+2}$. Each triangulation in $\mathcal{C}(T)$ contains the central triangle depicted. We depict the polygon P_{n+2} as a circle for simplicity. Right: the set of edges $\mathcal{E}(T,T')$ (which form a matching) between two classes.

Lemma 2.23. *For all $k \geq 3$ and $n \geq 1$, $e^{-1/6\frac{k-2}{k-1}} f(k,n) \leq C_{k,n} \leq e^{1/12} \cdot f(k,n)$, where*

$$f(k,n) = \frac{\sqrt{k-1}}{\sqrt{2\pi}((k-2)n)^{3/2}} \cdot \frac{(k-1)^{(k-1)n}}{(k-2)^{(k-2)n}}.$$

We will prove the following:

Lemma 2.24. *The flip graph $K_{k,(k-2)n+2}$, along with the partition $\mathcal{S}_{k,(k-2)n+2}$, satisfies Lemma 2.11.*

Theorem 1.3, as we will show in Appendix 2.7, will follow from tracing the particular quasipolynomial factors in the proof of Lemma 2.24.

To prove Lemma 2.24, we will partition $K_{k,(k-2)n+2}$ into a set of classes $\mathcal{S}_k$ in a suitable fashion. We will define a partition that generalizes one by Molloy, Reed, and Steiger [56]. In order to define $\mathcal{S}_k$, we need some observations about the structure of the graph $K_{k,(k-2)n+2}$.

2.7.3 Partition into classes

Given a k-gon T containing the center of the regular $(k-2)n+2$-gon $P_{(k-2)n+2}$ and sharing all of its vertices with $P_{(k-2)n+2}$, identify T with the class $\mathcal{C}(T)$ of k-angulations $v \in V(K_{k,(k-2)n+2})$ such that T forms one of the k-gons in the k-angulation v. Let $\mathcal{S}_{k,(k-2)n+2}$ be the set of all such $\mathcal{C}(T)$ classes. (If P_{n+2} has an even number of edges, we perturb the center slightly so that every triangulation lies in some class.)

Remark 2.4. *The set $\mathcal{S}_{k,(k-2)n+2}$ is a partition of $V(K_{k,(k-2)n+2})$, because no pair of k-gons whose endpoints are polygon vertices can contain the origin without crossing.*

(This generalizes the partition of Molloy, Reed, and Steiger [56].)

Given classes $\mathcal{C}(T), \mathcal{C}(T') \in \mathcal{S}_{k,(k-2)n+2}$, let $\mathcal{E}(T,T')$ be the set of edges between with one endpoint in $\mathcal{C}(T)$ and one endpoint in $\mathcal{C}(T')$. Let $\mathcal{B}_{T'}(T)$ denote the set of vertices in $\mathcal{C}(T)$ that have at least one neighbor in $\mathcal{C}(T')$. See Figure 2.8.

Remark 2.5. *The set of edge sets of the form $\mathcal{E}(T,T')$ is a partition of all edges between pairs of vertices in different classes.*

Cardinalities of classes and of edge sets

We make some observations about the nature and cardinalities of the classes in $\mathcal{S}_{k,(k-2)n+2}$, and of the sets and numbers of edges between the classes.

Lemma 2.25. *Each k-gonal class in $\mathcal{S}_{k,(k-2)n+2}$ induces a subgraph of $K_{k,(k-2)n+2}$ that is isomorphic to the Cartesian product $K_{k,(k-2)i_1+2} \square K_{k,(k-2)i_2+2} \square \cdots K_{k,(k-2)i_k+2}$, for some $1 \leq i_1 \leq \cdots \leq i_k \leq n/2$, $i_1 + \cdots + i_k = n-1$.*

Proof. Each k-gon T partitions the regular $(k-2)n+2$-gon into smaller convex polygons with side lengths $(k-2)i_1+2, (k-2)i_2+2, \ldots, (k-2)i_k+2$. Thus each k-angulation in

$\mathcal{C}(T)$ can be identified with a tuple of k-angulations of these smaller polygons. The Cartesian product structure then follows from the fact that every flip between two k-angulations in $\mathcal{C}(T)$ can be identified with a flip in one of the smaller polygons. $\square$

Lemma 2.26. *For each pair of classes $\mathcal{C}(T)$ and $\mathcal{C}(T')$, the boundary set $\mathcal{B}_{T'}(T)$ induces a subgraph of $\mathcal{C}(T)$ isomoprhic to a union of Cartesian products of the form $K_{(k-2)i_1+2}\Box K_{(k-2)i_2+2}\Box\cdots\Box K_{(k-2)i_{2k-2}+2}$, for some $i_1 \leq \cdots \leq i_{2k-2} \leq n/2$, $i_1 + \cdots + i_{2k-2} = n-2$.*

Proof. Each flip between k-angulations in adjacent classes $\mathcal{C}(T)$ involves flipping a diagonal of the k-gon T to transform k-angulation $t \in \mathcal{C}(T)$ into k-angulation $t' \in \mathcal{C}(T')$. Whenever this is possible, there must exist a $2k-2$-gon Q, sharing $k-1$ sides with T (the $k-1$ sides that are not flipped), such that both t and t' contain Q. Furthermore, every $t \in \mathcal{C}(T)$ containing Q has a flip to a distinct $t' \in \mathcal{C}(T')$. The set of all such boundary vertices $t \in \mathcal{C}(T)$ can be identified with the Cartesian product described because Q partitions $P_{(k-2)n+2}$ into a collection of smaller polygons, so that each k-angulation in $\mathcal{B}_{T'}(T)$ consists of a tuple of k-angulations in each of these smaller polygons, and such that every flip between k-angulations in $\mathcal{B}_{T'}(T)$ consists of a flip in one of these smaller polygons. (There may be many such $2k-2$-gons for a given pair of classes, but the claim holds as a lower bound.) $\square$

Lemma 2.27. *Each set of edges between classes in $\mathcal{S}_{k,(k-2)n+2}$ is in bijection with the vertices of a union of Cartesian products of the form $K_{(k-2)i_1+2}\Box K_{(k-2)i_2+2}\Box\cdots\Box K_{(k-2)i_{2k-2}+2}$, for $i_1 \leq \cdots \leq i_{2k-2} \leq n/2$, $i_1 + \cdots + i_{2k-2} = n-2$. Furthermore, no two edges in any such edge set share a vertex, i.e. the edge set is a matching.*

Proof. The claim follows from the reasoning in Lemma 2.26. $\square$

Corollary 2.1. *Each k-gonal class in $\mathcal{S}_{k,(k-2)n+2}$ has cardinality $C_{k,i_1}C_{k,i_2}\cdots C_{k,i_k}$, and each edge set between classes has cardinality at least $C_{k,i_1}C_{k,i_2}\cdots C_{k,i_{2k-2}}$. Here, $i_1,\ldots,i_{2k-2}$ are as in Lemmas 2.25 and 2.27.*

2.7.4 Applying the framework

We are almost ready to prove that $K_{k,(k-2)n+2}$ satisfies the conditions of Lemma 2.11, but first we need the following known fact:

Lemma 2.28. *$K_{k,(k-2)n+2}$ is connected.*

One way to prove Lemma 2.28 is via the isomorphism [36] between flips on k-angulations and rotations on $k-1$-ary plane trees. One can prove that the rotation graph on $k-1$-ary plane trees is connected as follows: find a path from any given tree to a "spine," where all internal nodes belong to a simple path via left children from the root to the leftmost leaf [19]. (This path consists of repeated left rotations.) Every non-spine tree has some internal node at which a left rotation can be performed. Furthermore, when no such operation is still possible, one has a spine.

Nakamoto, Kawatani, Matsumoto, and Urrutia [58] also gave a proof of connectedness for the special case $k = 4$. Sleator, Tarjan, and Thurston proved [69] that the diameter of $K_{3,n+2}$ is at most $2n - 6$ for $n \geq 11$.

We now prove Lemma 2.24:

Lemma 2.24. *The flip graph $K_{k,(k-2)n+2}$, along with the partition $\mathcal{S}_{k,(k-2)n+2}$, satisfies Lemma 2.11.*

Proof. By Lemma 2.25 and the observation that there are at most $\binom{(k-2)n+2}{k}$ classes, the partition $\mathcal{S}_{k,(k-2)n+2}$ meets Conditions 1 and 5 of the framework, with the modification to Condition 1 that the $O(1)$ term is replaced with $O(n^{O(1)})$, and Condition 6 follows from the identification of each class with a k-gon containing the center of the $(k-2)n+2$-gon.

Corollary 2.1 gives a formula for the size of each class and each edge set between classes. Lemma 2.23 then gives a polynomial bound on the ratio of $N = |V(K_{k,(k-2)n+2})|$ to the size of the smallest class (similarly the smallest edge set). Conditions 2, 3, and 4 follow, with the modification that the $O(1)$ terms are replaced with $O(n^{O(1)})$ terms. $\square$

To derive the specific quasipolynomial bound in Theorem 1.3, we first observe the following:

Remark 2.6. *The smallest edge set between classes in $\mathcal{S}_{k,(k-2)n+2}$ has size at least*

$$\begin{aligned}
C_{k,i_1} \cdots C_{k,i_{2k-2}} &\geq N \cdot \frac{f(k,i_1)\cdots f(k,i_{2k-2})}{e^{(2k-2)/6+1/12}((k-1)/(k-2))^{2k-2} f(k,n)} \\
&\geq N e^{(3-4k)/12} \cdot \frac{(k-2)^{k-3/2}}{(k-1)^{3k-5/2}} \cdot \frac{1}{(2\pi)^{k-3/2}} \cdot \frac{1}{n^{3k}}.
\end{aligned}$$

The next fact we need comes from Lemma 2.10:

Lemma 2.10. *Suppose a flip graph $\mathcal{M}_n = (\mathcal{V}_n, \mathcal{E}_n)$ belongs to a family $\mathcal{F}$ of graphs satisfying the conditions of Lemma 2.9. Suppose further that every graph $\mathcal{M}_k = (\mathcal{V}_k, \mathcal{E}_k) \in \mathcal{F}$, $k < n$, satisfies*

$$|\mathcal{V}_k|/|\mathcal{E}_{k,\min}| \leq f(k),$$

for some function $f(k)$, where $\mathcal{E}_{k,\min}$ is the smallest edge set between adjacent classes $\mathcal{C}(T), \mathcal{C}(T') \in \mathcal{S}_k$, where $\mathcal{S}_k$ is as in Lemma 2.9. Then the expansion of $\mathcal{M}_n$ is

$$\Omega(1/(2f(n))^{\log n})).$$

Applying Lemma 2.10, and using the fact that $K_{k,(k-2)n+2}$ is a $\leq (k-2)n$-regular graph with

$\log N \le (k-1)n\log(k-1)$, gives mixing time

$$\begin{aligned}
&O((2N/\mathcal{E}_{\min})^{2\log n}(k-1)^3(\log(k-1))n^3) \\
&= O((k-1)^3(\log(k-1))n^3(2e^{(4k-3)/12}\cdot\frac{(k-1)^{3k-5/2}}{(k-2)^{k-3/2}}\cdot(2\pi)^{k-3/2}\cdot n^{3k})^{2\log n}) \\
&= O((k-1)^3(\log(k-1))n^3(2e^{(4k)/12}\cdot(k-1)^{3k}\cdot(2\pi)^{k}\cdot n^{3k})^{2\log n}) \\
&= O((k-1)^3(\log(k-1))\cdot n^{2(3k\log(k-1)+k(1+\log\pi)+3k\log n+k)+5}).
\end{aligned}$$

Here we have implicitly used Lemma 2.1 to pass from the expansion bound given by Lemma 2.10 to a mixing bound. Actually, we can do better using the following standard lemma, which allows for passing from congestion to mixing without a quadratic loss:

Lemma 2.29. *[22, 67] Suppose a uniform multicommodity flow f exists in a graph $G=(V,E)$ with congestion ρ, in which for every $s,t\in V$,*

$$\max_{P\in\Gamma_{st}}|P|\le l,$$

for some $l>0$, where Γ_{st} s the set of (simple) paths in G from s to t, and where we use the shorthand $f_{st}(P)$ to denote the fraction of the s,t commodity that f sends along the path P. Then the mixing time of the uniform random walk on G is

$$O\left(\frac{\rho l}{d}\log(|V(G)|)\right),$$

where d is the maximum degree of G.

Then we obtain mixing time

$$O((2N/\mathcal{E}_{\min})^{\log n}(k-1)^2(\log(k-1))n^2\cdot l),$$

where l is the maximum length of a path in the flow construction. It is not difficult to see that since the diameter of the projection graph is at most k, we obtain a recurrence

$$l = T(n) = k + 2kT(n/2) = O(n^{\log_2 k+1}),$$

giving total mixing

$$O((2N/\mathcal{E}_{\min})^{\log n}(k-1)^2(\log(k-1))n^2 \cdot n^{\log_2 k+1}),$$

As noted in Remark 2.1, we did not incur a term $\gamma\Delta$ in this calculation. Furthermore, it is not clear how one could avoid this $\gamma\Delta = k$ factor using the spectral decomposition technique (Theorem 2.2). That technique would give a mixing bound of

$$O(((3N/\mathcal{E}_{\min}) \cdot k \cdot k)^{\log n} \cdot (k-2)n).$$

(Here we have ignored an additional $\log|\Omega|$ term, as one might be able to reduce this term via, for instance, the log-Sobolev version of the Jerrum/Son/Tetali/Vigoda decomposition.) Comparing the two expressions above shows that our technique gives an improvement of

$$\Omega\left(\frac{n^{\log_2(k^2)}}{n^{\log_2 k+1} \cdot kn\log k}\right) = \Omega(n^{\log_2 k-2}/(k\log k)).$$

2.8 Integer lattice triangulation flip graphs

2.8.1 Definition

The integer lattice triangulation flip graph, studied extensively in prior work ([1, 13, 12, 41, 71]), is analogous to the associahedron and is defined as follows:

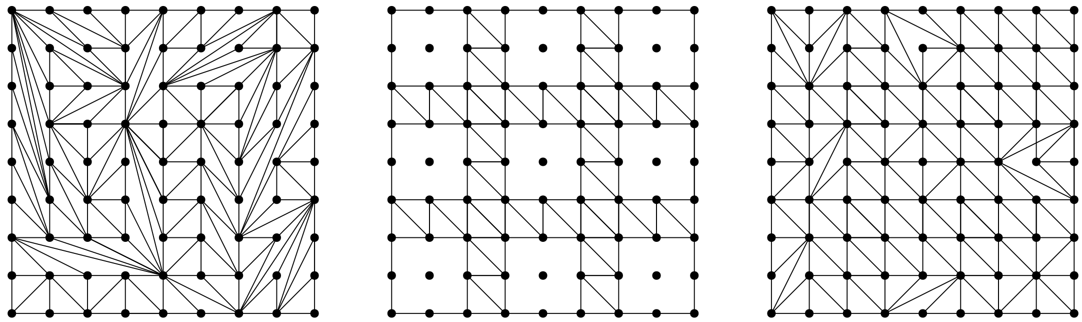

Figure 2.9: Left: a triangulation of the 9x9 integer lattice. Center: a division of the lattice into 9 3x3 sections, as described in the proof of Theorem 1.4. Right: a full triangulation compatible with the division of the lattice.

Let the *integer lattice triangulation flip graph* be the graph F_n whose vertices are the triangulations of the $n \times n$ integer lattice point set (integer grid), and whose edges are the pairs of triangulations that differ by exactly one diagonal.

It will be useful to define notation for the number of triangulations in this graph: Let $g(n)$ be the number of triangulations of the $n \times n$ integer lattice point set.

In fact, $g(n)$ is unknown in general, though much progress has been made on upper and lower bounds, including the following result of Kaibel and Ziegler [41]:

Lemma 2.30. *For $n \geq 1$,*

$$h(n) = \Theta(2^{cn^2}),$$

for some constant c.

As discussed in the introduction, in recent years progress has been made ([13, 12, 71]) in studying the mixing properties of the natural flip walk on the integer lattice triangulation flip graph. However, this work has focused on biased versions of the flip walk, in which a real parameter $\lambda > 0$ induces a weight function on the triangulations of the lattice, and in which the random walk is biased in favor of triangulations with larger weights. The case of $\lambda = 1$ is the unbiased version of the walk. It is now known [12] that when $\lambda > 1$, the walk does not mix rapidly, but that rapid mixing does occur for certain values of λ smaller than one.

However, the question is open for the biased version.

We do not settle the question—which would equate to showing that the integer lattice triangulation flip graph has expansion at least $1/p(n)$ for some polynomial function p, but we do show a weaker result in a similar spirit: that the flip graph has large subgraphs with large expansion. Expansion measures the extent to which bottlenecks exist in a graph: large expansion corresponds to a graph that does not have bottlenecks, roughly speaking. Thus, even if bottlenecks exist in F_n—that is, if rapid mixing does *not* occur, i.e. if the expansion is too small—then there still exist regions of the graph that are not prone to bottlenecks, and thus internally induce rapidly mixing walks. Although far from clear evidence of large expansion in F_n itself, one might hope that if bottlenecks exist, this result may suggest places to look for them.

2.8.2 Additional preliminaries: treewidth, separators, and vertex expansion

The *treewidth* of a graph G is a different density parameter from expansion. There are many equivalent definitions of treewidth; one of the standard definitions is in terms of a so-called *tree decomposition*.

Closely related to treewidth are *vertex separators*: A *vertex separator* for a graph G is a subset $X \subseteq V(G)$ of the vertices of G such that $G \setminus X$ is disconnected. X is a *balanced separator* if $G \setminus X$ consists of two subgraphs, A and B, such that no edge exists between A and B, and such that $|V(G)|/3 \leq |V(A)| \leq |V(B)| \leq 2|V(G)|/3$. We also say, if $|X| \leq s$ for a given $s \geq 1$, that X is an *s-separator.*

With respect to an integer $s \geq 1$, a graph G is *recursively s-separable* if either $|V(G)| \leq 1$, or G has a balanced s-separator X such that the two mutually disconnected subgraphs induced by removing X from G are both recursively s-separable. The following relates treewidth and recursive separability [27]:

Lemma 2.31. *For every $t \geq 1$, every graph with treewidth at most t is recursively $t+1$-separable.*

Treewidth in general is of interest in large part because many NP-hard problems become tractable on graphs of bounded treewidth. For a survey of this phenomenon, known as *fixed-parameter tractability*, see [9]. Our interest in treewidth, however, is mainly in its role as a density parameter, in particular for Theorem 1.4.

Treewidth, as a density parameter, is weaker than vertex expansion, in the sense that a high vertex expansion implies a high treewidth, but not vice versa. The following following corollary to Lemma 2.31 makes this precise:

Corollary 2.2. *If the vertex expansion of a family of graphs $G(N)$ on N vertices is at least $h_v(N)$, then the treewidth $t(N)$ of the family is $\Omega(N \cdot h_v(N))$.*

Proof. Suppose $G(N)$ has vertex expansion at least $h_v(N)$. Then every balanced separator X is of size at least

$$|X| \geq h_v(N) \cdot N/3,$$

by the definition of a balanced separator and the definition of vertex expansion. □

In this section we prove Theorem 1.4.

Theorem 1.4. *The treewidth of the triangulation flip graph F_n on the $n \times n$ integer lattice point set is $\Omega(N^{1-o(1)})$, where $N = |V(F_n)|$.*

Proof. We will show that F_n has a large induced subgraph with large expansion, which will imply large treewidth. Partition the points of the $n \times n$ grid into n grids of size $\sqrt{n} \times \sqrt{n}$. (If

n is not a perfect square, we can take $\sqrt{\lfloor n \rfloor}$.) That is, fill in a partial triangulation as follows: let each point in the grid have coordinates (i, j), where $1 \le i \le n$ and $1 \le j \le n$. Fill in all vertical edges connecting two consecutive points with the same j coordinate whenever $j \equiv 0 \pmod{\sqrt{n}}$ or $j \equiv 1 \pmod{\sqrt{n}}$, and fill in all horizontal edges connecting two consecutive points with the same i coordinate whenever $i \equiv 0 \pmod{\sqrt{n}}$ or $i \equiv 1 \pmod{\sqrt{n}}$.

Fill in also all unit horizontal and unit vertical edges connecting vertices in adjacent $\sqrt{n} \times \sqrt{n}$ sub-grids, and fill in all unit diagonals with negative slope inside the resulting squares. See Figure 2.9, center. (The choice of these edges and diagonals to fill in between subgrids is arbitrary, but must be consistent.)

Now consider the subgraph H_n of F_n induced by restricting $V(F_n)$ to the triangulations that extend this partial triangulation. That is, the vertices of H_n are the triangulations that consist of separately triangulating each of the $\sqrt{n} \times \sqrt{n}$ grids. H_n is the Cartesian product of n graphs that are each isomorphic to $F_{\sqrt{n}}$. See Figure 2.9, right, for an example of such a triangulation.

Assuming the smallest possible expansion for a graph on $g(\sqrt{n})$ vertices, $F_{\sqrt{n}}$ graph has expansion $\Omega(1/g(\sqrt{n}))$. The degree of F_n is $O(n^2)$. Now, by Lemma 2.7, H_n has expansion $\Omega(1/g(\sqrt{n}))$.

Therefore, by Corollary 2.2, H_n has treewidth $\Omega(g(n)/(n^2 g(\sqrt{n})))$.

Now, by Lemma 2.30, H_n has treewidth

$$\Omega(g(n)/(n^2 g(\sqrt{n}))) = \Omega(2^{c \cdot n^2 - cn - 2\log n}) = \Omega(2^{cn^2(1-o(1))}) = \Omega(N^{1-o(1)}),$$

proving the theorem. □

Chapter 3

Rapid mixing for the hardcore Glauber dynamics and other Markov chains in bounded-treewidth graphs

In this chapter we apply the machinery we developed in Chapter 2 to a number of graph-theoretic chains. Recall from Chapter 2 that a *multicommodity flow* in a graph $\mathcal{M}$ is a set of $|V(\mathcal{M})|^2$ flows, one flow for each ordered pair of vertices (u, v), where each flow sends one unit of a commodity from u to v. If a flow exists in $\mathcal{M}$ with small *congestion*—i.e. one in which no edge carries too much flow—then the natural Markov chain whose states are the vertices of $\mathcal{M}$ mixes rapidly. (We make this precise in Section 3.1.4.)

All of the chains we analyze can be seen as natural random walks on a "Glauber graph" $\mathcal{M}(G)$ whose vertices are subsets of an underlying set. (For our purposes, this underlying set is either the vertex set or the edge set of the input graph G.) Thus each of these random walks is performed on a graph that may be exponentially large with respect to the size of the input graph. In Chapter 2, we showed that when all of a certain set of conditions hold, we can construct a multicommodity flow in $\mathcal{M}(G)$ with congestion polynomial in $n = |V(G)|$, implying that the unbiased random walk on $\mathcal{M}(G)$ mixes rapidly.

Recall that the conditions specify that $\mathcal{M}(G)$ can be partitioned into a small number of induced subgraphs, all of which are approximately the same size, with large numbers of edges between pairs of the subgraphs. The conditions also require that each of these induced subgraphs have a special structure, allowing for the decomposition of each induced subgraph into smaller Glauber graphs that are similar in structure to $\mathcal{M}(G)$. This self similarity allows for the inductive construction of a multicommodity flow, by assembling flows on smaller Glauber graphs together into a flow in $\mathcal{M}(G)$ with small congestion.

3.0.1 Prior work and our contribution

Prior work on rapid mixing of Markov chains on subset systems includes the special case of matroid polytopes. For this case, recent results [4, 3] have partly solved a 30-year-old conjecture of Mihail and Vazirani [55]. Other prior work uses multicommodity flows (and the essentially equivalent *canonical paths* technique) to obtain polynomial mixing upper bounds on structures of exponential size, including matchings and 0/1 knapsack solutions [57, 32]. Madras and Randall [52] used a decomposition of the hardcore model state space to prove rapid mixing under different conditions. We also decompose the state space, but our approach is different from that of Madras and Randall and is more similar to Heinrich's [35] application of the projection-restriction technique (see Chapter 2) pioneered by Jerrum, Son, Tetali, and Vigoda [40]. Essentially, the projection-restriction technique involves partitioning the state space of a chain into a collection of sub-state spaces, each of which internally mixes rapidly, and all of which are well connected to one another. Heinrich used the vertex separation properties of bounded-treewidth graphs to obtain an inductive argument: the resulting sub-spaces are themselves Cartesian products of chains on smaller graphs, and thus mix rapidly. (See Lemma 2.15.)

We partition the state space recursively using the same vertex separation properties, and indeed the framework conditions from Chapter 2, which we restate, in Section 3.2.3, in terms more conducive to our graph-theoretic chains, suffice for rapid mixing using the projection-restriction technique. Thus part of our contribution is simply to *observe* that these conditions suffice. The main contribution in this chapter is to extend the framework to the chains we analyze—for most of which this is not trivial. Thus our main contribution in this chapter is to give conditions under which rapid mixing occurs in the *projection* chains. In particular, in the case of independent sets, Jerrum, Son, Tetali, and Vigoda [40] applied their technique to a special case of the hardcore model, namely regular trees. However, it was not clear how to generalize this application to non-regular trees or to bounded-treewidth graphs. We resolve this with the *hierarchical* version of our framework, and show that this version also gives an analogous result for dominating sets.

Our framework solves another key problem that arises in applying the projection-restriction technique: to apply that technique in a straightforward fashion, one needs each of the state spaces in the partition to be a Cartesian product of chains on smaller spaces. For four of our eight chains, the sub-spaces obtained in the decomposition are not Cartesian products but may be non-disjoint unions of Cartesian products. In some cases, the sub-spaces may not even be mutually disjoint and may induce non-ergodic restriction chains. We solve this problem by using the structure of the state spaces of Glauber dynamics as graphs to obtain multicommodity flows with bounded congestion. We will discuss this further in Section 3.8. Our approach is also inspired by Kaibel's [42] construction of a flow with bounded congestion in any graph whose vertices are hypercube vertices and whose edges can be partitioned into bipartite graphs in a hierarchical fashion.

3.0.2 Application to graphical models

Prior work [11, 30] has shown that related chains, including *softcore models*—in which the sampled sets need not be independent—mix rapidly on graphs of bounded treewidth. However, all of the Glauber dynamics we consider pertain to graph-theoretic sampling problems, in which one is sampling a subset of either the vertices or the edges of a graph, where the subsets must obey certain constraints, e.g. independence. As a result, and as Bordewich and Kang [11] note, their technique does not extend to these models.

Similarly, in the setting of probabilistic graphical models, De Sa, Zhang, Olukotun, and Ré [20] considered graphs with bounded *hierarchy width*. They showed—via arguments similar to the projection-restriction technique [40]—that graphs with logarithmically bounded hierarchy width admit rapid mixing for the Glauber dynamics on models with bounded *maximum factor weight*. It is straightforward to apply their argument to the *Ising* and *Potts* models with fixed parameters, on graphs of bounded carving width. This case of these models also admits application of projection-restriction (and in the special case of the path graph Jerrum, Sinclair, Tetali, and Vigoda observed this for the Ising model [40]), and it fits our framework. Since our framework does not give a substantial improvement on existing results for these models, we do not address them in detail in this paper; we simply note that the framework we developed in our prior work [26, 25] applies to these cases and to every undirected graphical model having only pairwise and unary factors, bounded maximum factor weights, constantly many values for each random variable, and bounded carving width. This shows that the framework unifies these models—in which all states have positive probability and which prior work has addressed in these graphs—with graph-theoretic chains where some states have zero probability—for which our results are new. We give a brief sketch of how to apply our framework in Section 3.6.2. See Bordewich, Greenhill, and Patel [10] and Chen, Liu, and Vigoda [17] for definitions of and results for these models.

3.0.3 Further discussion of prior work

Sly [70] showed that, except for restricted values of $\lambda < 1$, the hardcore Glauber dynamics does not mix rapidly on general graphs unless RP = NP, and in fact showed that approximately sampling from the corresponding distribution is hard unless RP = NP. However, Anari, Liu, and Gharan [2] used a technique known as *spectral independence* to obtain rapid mixing for the hardcore Glauber dynamics when λ is below the so-called *uniqueness threshold* that depends on the maximum degree of the input graph. They showed, by exhibiting an infinite family of examples, that the technique they used could not be further improved (namely beyond the uniqueness threshold) even for trees. By contrast, we show that rapid mixing, for all fixed values of λ, indeed holds not only for trees but for all graphs of bounded treewidth. Chen, Galanis, Štefankovič, and Vigoda [16] and Feng, Guo, Yin, and Zhang [28] generalized this work and applied it to graph colorings.

Other results exist for trees beyond the uniqueness threshold, however: Martinelli, Sinclair, and Weitz [53] showed that the Glauber dynamics on the hardcore model mixes in $O(n \log n)$ time on the complete $\Delta - 1$-ary tree with n nodes. They also showed that the dynamics on q-colorings ($q \geq \Delta + 2$) mixes in $O(n \log n)$ time on the same trees. Lucier, Molloy, and Peres [51] showed that the dynamics mixes rapidly on general trees of bounded degree, namely in time $O(n^{O(1+\Delta/(q \log \Delta))})$.

Prior work also exists for q-colorings of bounded-treewidth graphs: Berger, Kenyon, Mossel, and Peres [6] showed rapid mixing for q-colorings of trees. Tetali, Vera, Vigoda, and Yang [72] gave upper and lower bounds for complete trees. Vardi [73] showed that the so-called *single-flaw* dynamics—a variaton on the Glauber dynamics in which at most one monochromatic edge is permitted in a valid state—mixes rapidly on bounded-treewidth graphs when $q \geq (1 + \varepsilon)\Delta$, for any fixed parameter $\varepsilon > 0$. The proof used the vertex separaton properties of bounded-treewidth graphs to construct a multicommodity flow with bounded congestion, although the construction was substantally different from our divide-and-conquer approach. Dyer, Goldberg, and Jerrum [23] showed rapid mixing when the degree of the graph is at least $2t$ and $q \geq 4t$, where t is the treewidth. On the other hand, Heinrich [35] showed that the Glauber dynamics on q-colorings of a bounded-treewidth graph mixes rapidly when $q \geq \Delta + 2$. Our construction, as we will discuss in more detail in Section 3.0.1, bears some similarity to Heinrich's. We also require that q (and therefore Δ) be bounded. However, due to a more general analysis of the state spaces of Glauber dynamics as graphs, we obtain a more general framework that holds for a greater variety of chains.

Planar graphs have unbounded but sublinear treewidth. For planar graphs, Hayes [33] showed that the Glauber dynamics on q-colorings of a planar graph of maximum degree Δ mixes rapidly when $q \geq \Delta + O(\sqrt{\Delta})$. Later, Hayes, Vera, and Vigoda [34] proved rapid mixing for q-colorings of planar graphs when $q = \Omega(\Delta / \log \Delta)$, generalizing further to a spectral condition on the adjacency matrix of the graph.

Bezáková and Sun showed [8] that the hardcore model mixes rapidly in chordal graphs with bounded-size separators. Lastly, Chen, Galanis, Štefankovič, and Vigoda applied the spectral independence technique to prove that the Glauber dynamics on the q-colorings of a triangle-free graph with dgree Δ mixes rapidly provided that $q \geq \alpha\Delta + 1$, where α is greater than a threshold approximately equal to 1.763. We show that when the carving width of G is bounded, G need not be triangle free, and it suffices that $q \geq \Delta + 2$ be bounded. We prove a similar result for the natural Glauber dynamics on *partial* q-colorings.

Although our mixing results are new, Wan, Tu, Zhang, and Li showed [74] that exact counting of independent sets is fixed-parameter tractable in treewidth. Furthermore, our result does not technically constitute a proof of fixed-parameter tractability, as the treewidth appears in the exponent of the polynomial we obtain. For this problem and all the other problems we consider, the problem of exact counting—and therefore also uniform sampling—has already been solved on the graphs we consider by an extension of Courcelle's theorem [61]. In fact, the standard reduction from approximate sampling to approximate counting [68] gives a somewhat different rapidly mixing Markov chain on a larger state space. Nonetheless, our result does settle the question of rapid mixing for a natural chain, and it implies a simpler scheme for approximately sampling independent sets than one would obtain via this reduction.

Such a scheme is known as a *fully polynomial randomized approximation scheme (FPRAS)*. Huang, Lu, and Zhang provided an FPRAS for sampling b-edge covers in general graphs when $b \leq 2$, and for sampling b-matchings when $b \leq 7$ [37]. This FPRAS relied on a rapid mixing argument for a somewhat different Markov chain than ours. Existing dominating set results for certain regular graphs are also known [7].

Exact counting of maximal independent sets—which would give an FPRAS by the equivalence of counting and sampling—was shown in [60] to be hard for chordal graphs but is known [15] to be tractable in graphs of bounded treewidth. However, again our result improves on the simplicity of existing algorithms.

3.1 Preliminaries

In this section we define the hardcore Glauber dynamics. We also define in this section the standard notion of carving width and the additional graph-theoretic chains we explore. See Chapter 2 for definitions of the standard notions of expansion, multicommodity flows, and treewidth.

3.1.1 Rapid mixing and Glauber dynamics

As we discussed in Chapter 2, rapid mixing is of interest in the random generation of certain graph-theoretic objects, including such subset systems as the set of matchings in a graph. To generate, approximately uniformly at random, an object of a given class—say, an independent set in a given graph—it suffices to conduct a random walk on a graph whose vertices are the objects of interest, and whose edges are *flips* between the objects, under some suitable definition of a flip. (For technical reasons, self loops need to be added to the graph in a standard fashion.) Basic spectral graph theory shows that, under mild conditions, the walk converges to the uniform distribution in the limit. It is of interest for efficient sampling algorithms to determine how rapid the convergence is. In the case of subset systems such as those we consider, the walk takes place over an exponentially large number of subsets defined over an underlying set of size n. If the convergence, or *mixing time*, of the walk is polynomial in n, then the random walk is said to be *rapidly mixing*.

Recall that the mixing time, denoted τ, is the minimum number of steps in the random walk before convergence is guaranteed, regardless of the starting point of the walk. Convergence is measured via the *total variation distance* [67] between the distribution over states induced by the walk at a given time step, and the uniform distribution. One can obtain convergence to other distributions by adding weights to the vertices and edges of the graph—see Section 3.6.1.

Of interest for our mixing results is the hardcore Glauber dynamics, defined as follows: The *hardcore Glauber dynamics* on the independent sets of a graph G is the following chain, defined with respect to a fixed real parameter $\lambda > 0$:

1. Let X_0 be an arbitrary independent set in G.

2. For $t \geq 0$, select a vertex $v \in V(G)$ uniformly at random.

3. If $v \notin X_t$ and $X_t \cup \{v\}$ is not a valid independent set, do nothing.

4. Otherwise:

    Let $X_{t+1} = X_t \cup \{v\}$ with probability $\lambda/(\lambda + 1)$.

    Let $X_{t+1} = X_t \setminus \{v\}$ with probability $1/(\lambda + 1)$.

3.1.2 Carving width

The *carving width* of a graph is a density parameter that is weaker than treewidth, in the sense that high treewidth implies high carving width, but the converse is not true. Carving width is defined with respect to a so-called *carving decomposition* [24] of a given graph G—in short, a binary tree T whose leaves are identified with the vertices of G. Each node $X \in T$ is identified with the subgraph of G induced by the vertices of G (leaves of T) having X as an ancestor in T. Each edge of T induces a cut in T; this cut induces a partition of the leaves of T (vertices of G) into two sets. This partition is naturally identified with a cut in G.

The *width* of a carving decomposition is the maximum number of edges of G across any such cut, where the maximum is taken over all edges in T. The *carving width* of G is the minimum width of a carving decomposition of G. See Seymour and Thomas [66] for a detailed treatment. For our purposes, carving width is of interest due to its relationship to the treewidth and degree of a graph. Specifically, Eppstein [24] observed the following fact that follows from results of Nestoridis and Thilikos [59] and of Robertson and Seymour [65]:

Lemma 3.1. *Given a graph G with maximum degree Δ, let $\mathrm{tw}(G)$ denote the treewidth of G, and let $\mathrm{cw}(G)$ denote the carving width of G. For every graph G, $(2/3)(\mathrm{tw}(G)+1) \leq \mathrm{cw}(G) \leq \Delta(\mathrm{tw}(G)-1)$.*

It follows from the definition of carving width that every graph with bounded carving width also has bounded degree. Combining this fact with Lemma 3.1 implies the following:

Corollary 3.1. *A graph has bounded degree and treewidth if and only if it has bounded carving width.*

3.1.3 Dominating sets, b-matchings, and b-edge covers

A *dominating set* in a graph $G = (V, E)$ is a set $S \subseteq V$ of vertices such that for every vertex

$v \in V$, either $v \in S$ or there exists some vertex $u \in S$ such that $(u, v) \in E$.

b-matchings [45] and b-edge covers [29, 44] generalize the definitions of matchings and edge covers respectively: Let $G = (V, E)$ be a graph. Let $b : V \to \mathbb{Z}^{\geq 0}$ be any function assigning a nonnegative integer to each vertex. A *b-matching* in a graph $G = (V, E)$ is a set $S \subseteq E$ of edges such that every $v \in V$ has at most $b(v)$ incident edges in S.

Let $G = (V, E)$ be a graph. Let $b : V \to \mathbb{Z}^{\geq 0}$ be any function assigning a nonnegative integer to each vertex. A *b-edge cover* in a graph $G = (V, E)$ is a set $S \subseteq E$ of edges such that every $v \in V$ has at least $b(v)$ incident edges in S.

Sometimes, as in the result by Huang, Lu, and Zhang [37], b-edge covers and b-matchings are defined so that b is a constant, i.e. $b(u) = b(v)$ for all $u, v \in V$.

For dominating sets, b-edge covers, and b-matchings, we consider a chain similar to the hardcore dynamics, except that in the case of b-edge covers and b-matchings, we are of course selecting edges instead of vertices. Also, in the case of dominating sets and b-edge covers, instead of verifying independence before *adding* a vertex (or edge), we verify validity of a set (e.g. domination) before *dropping* a vertex (or edge).

We also consider the Glauber dynamics on *q-colorings*: A *q-coloring* of a graph G is an assignment of a color from the list $[q] = \{1, 2, \ldots, q\}$ to each vertex of G, such that no two adjacent vertices have the same color. A *partial q-coloring* of a graph G is an assignment of a color from $[q]$ to each of a subset of the vertices of G, such that no two adjacent vertices have the same color.

The Glauber dynamics on the partial q-colorings of G is as follows: Let the Glauber dynamics on the partial q-colorings of a graph G be the following chain defined with respect to $\lambda > 0$:

1. Let X_0 be an arbitrary partial q-coloring of G.

2. For $t \geq 0$, select a vertex $v \in V(G)$ uniformly at random, and select a color $c \in [q+1]$ uniformly at random.

3. If $c = q+1$, then:

   If v is already colored in X_t, remove the coloring of v with probability $1/(\lambda+1)$.

   Otherwise, let $X_{t+1} = X_t$.

4. If $c \leq q$, then:

   If v is *not* already colored with c in X_t, set the color of v to c with probability $\lambda/(\lambda+1)$.

   Otherwise, let $X_{t+1} = X_t$.

Finally, the Glauber dynamics on the (complete) q-colorings of G is as follows (for this chain we do not define a biased version): Let the Glauber dynamics on the q-colorings of a graph G be the following chain:

1. Let X_0 be an arbitrary q-coloring of G.

2. For $t \geq 0$, select a vertex $v \in V(G)$ uniformly at random, and select a color $c \in [q]$—other than the color of v—uniformly at random.

3. If v has no neighbor with color c, then change the color of v to c with probability $1/2$ to obtain X_{t+1}.

4. Otherwise, do nothing, i.e. let $X_{t+1} = X_t$.

We define a graph whose vertices are the maximal independent sets of an underlying graph G, and then define the flip chain as a random walk on this graph: Given a graph $G = (V, E)$, let the *maximal independent set Glauber graph* be the graph $\mathcal{M}_{\text{MIS}}(G)$ whose vertices are the maximal independent sets of G, and whose edges are the pairs of maximal independent sets that differ by one *flip*, where a flip is defined as:

1. adding one vertex v to a given independent set $S \subseteq V$,
2. removing every $u \in S$ such that $(u, v) \in E$, and
3. adding a subset of the vertices at distance two in G from v.

Since $\mathcal{M}_{\mathrm{MIS}}(G)$ is undirected, we also define the reversal of a flip as a flip. See Figure 3.5 for an example of a flip.

Lemma 3.2. *The graph $\mathcal{M}_{\mathrm{MIS}}(G)$ is connected.*

Proof. The proof relies on a greedy transformation argument and is in Section B.2. □

For maximal b-matchings, we define a Glauber graph similar to the maximal independent set Glauber graph, except that we are of course selecting edges instead of vertices in our sets. A flip consists of adding some edge $e = (u, v)$ to the b-matching, then removing edges incident to u and v as needed until a valid b-matching is obtained, then adding edges incident to neighbors of u and v as needed to obtain maximality.

3.1.4 Glauber dynamics with parameter $\lambda > 0$

Formally, the Glauber dynamics is defined as follows:

The *Glauber dynamics* is a Markov chain, parameterized by $\lambda > 0$, with state space $\Omega = V(\mathcal{M}(G))$ and probability matrix P, where for $S, S' \in V(\mathcal{M}(G))$ with $S \neq S'$,

$$P(S, S') = \lambda/(\Delta_{\mathcal{M}}(\lambda + 1))$$

when $|S' \setminus S| = 1$, and

$$P(S, S') = 1/(\Delta_{\mathcal{M}}(\lambda + 1))$$

when $|S \setminus S'| = 1$. If $S = S'$, then $P(S, S') = 1 - \sum_{S'' \neq S} P(S, S'')$.

Here $\Delta_{\mathcal{M}}$ is the maximum degree of the Glauber graph—i.e. the maximum number of neighboring states that a state S can have.

The two cases described for $S \neq S'$ are exhaustive for all of the chains that we have parameterized by λ. For all of the chains for which λ is not defined, the transition probability is in every case $1/(2\Delta_{\mathcal{M}})$.

3.2 $\lambda = 1$: Bounded carving width

To build up to the proof of Theorem 1.5, we first show a weaker result: that the unbiased Glauber dynamics on independent sets mixes rapidly in graphs of bounded carving width. The full proof of Theorem 1.5, even in the unbiased case, requires the non-hierarchical framework.

The main technical lemma in this section, Lemma 3.10, comes from Chapter 2. Our contribution in this chapter is the application to independent sets in graphs of bounded carving width—which we strengthen to graphs of bounded treewidth in Section 3.3.

The independent set flip chain (the hardcore model) is the natural random walk on what we will call the *independent set Glauber graph*: Given a graph G, let the *independent set Glauber graph* $\mathcal{M}_{\text{IS}}(G)$ be the graph whose vertices are the independent sets of G, and whose edges are the pairs of independent sets S, S' such that $|S \oplus S'| = 1$. The following is known, but we give a proof:

Lemma 3.3. *The independent set Glauber graph is connected.*

Proof. Consider the empty independent set $\emptyset$. Every independent set $S \in V(\mathcal{M}_{\text{IS}}(G))$ has a path of length $|S|$ to $\emptyset$, formed by removing each vertex in S in arbitrary order. □

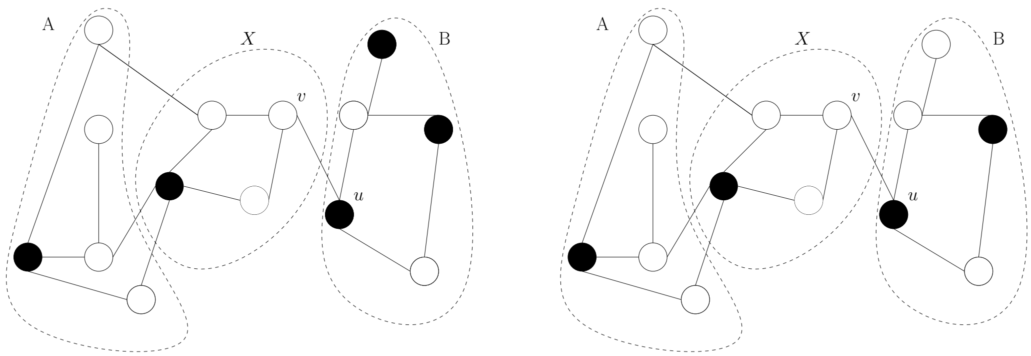

Figure 3.1: Two independent sets in a graph G, belonging to the same class, induced by the restriction of the sets to X.

3.2.1 Partitioning the vertices of $\mathcal{M}_{\mathrm{IS}}(G)$ into classes

The vertices of the Glauber graph $\mathcal{M}_{\mathrm{IS}}(G)$ are subsets of the vertices of an underlying graph G. When G has bounded treewidth, we can choose a small separator X that partitions $V(G) \setminus X$ into two mutually disconnected vertex subsets, A and B, neither of which is too large. Consider the problem of sampling an independent set S from G. Given a separator X for G, partition the independent sets in G into equivalence classes as follows:

Let $G = (V, E)$ be a graph. Let $\mathcal{M}_{\mathrm{IS}}(G)$ be as we have defined. Let $X \subseteq V$ be a vertex separator for G. Let $\mathcal{S}_{\mathrm{IS}}(G)$ be the set of equivalence classes of $V(\mathcal{M}_{\mathrm{IS}}(G))$ in which two independent sets S and S' are in the same class if $S \cap X = S' \cap X$. Let $T = S \cap X$, and call the corresponding class $\mathcal{C}_{\mathrm{IS}}(T)$.

See Figure 3.1 for an example of a partitioning and a class.

Let A and B be the mutually disconnected vertex subsets into which the removal of X partitions $G[V \setminus X]$. Given a fixed independent subset $T \subseteq X$, identify the independent sets in $\mathcal{C}_{\mathrm{IS}}(T)$ with the pairs of the form (S_A, S_B), where S_A is an independent set in $A \setminus N_A(T)$, and S_B is an independent set in $B \setminus N_B(T)$, where $N_A(T)$ and $N_B(T)$ denote the union of the neighborhoods of vertices in T, in A and B respectively. That is, identify each independent set in $\mathcal{C}_{\mathrm{IS}}(T)$ with a pair of an independent set in A that avoids neighbors of vertices in T,

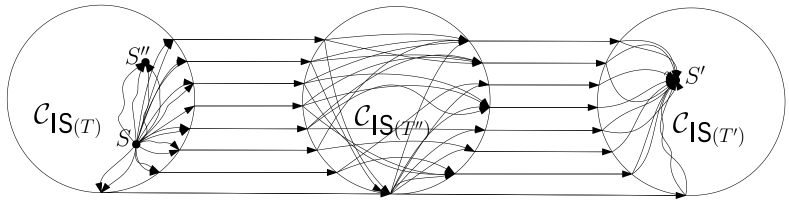

Figure 3.2: A schematic view of three classes in the independent set Glauber graph $\mathcal{M}_{\text{IS}}(G)$. The large circles denote classes under the partition described in Section 3.2.1. The curved arrows illustrate the construction of a flow in $\mathcal{M}_{\text{IS}}(G)$ from an independent set $S \in \mathcal{C}_{\text{IS}}(T)$ to another independent set $S'' \in \mathcal{C}_{\text{IS}}(T)$—and also to an independent set $S' \in \mathcal{C}_{\text{IS}}(T')$. Here, $\mathcal{C}_{\text{IS}}(T)$ and $\mathcal{C}_{\text{IS}}(T'')$ are adjacent classes in $\mathcal{M}_{\text{IS}}(G)$, connected by a large number of edges, and similarly $\mathcal{C}_{\text{IS}}(T')$ and $\mathcal{C}_{\text{IS}}(T'')$ are adjacent. In Section 3.2.2 we formalize this flow.

and a similar independent set in B. Consider the two Glauber graphs $\mathcal{M}_{\text{IS}}(A \setminus N_A(T))$ and $\mathcal{M}_{\text{IS}}(B \setminus N_B(T))$, whose vertices are respectively the independent sets in $G[A \setminus N_A(T)]$, and those in $G[B \setminus N_B(T)]$. If two independent sets $S = (S_A, S_B)$ and $S' = (S'_A, S'_B)$ belong to the same class, then a flip exists between S and S' in $\mathcal{M}_{\text{IS}}(G)$ precisely when a flip exists between the restrictions of S and S' to either $\mathcal{M}_{\text{IS}}(A \setminus N_A(T))$ or $\mathcal{M}_{\text{IS}}(B \setminus N_B(T))$ (but not both). See Figure 3.1. Therefore, each class induces, in $\mathcal{M}_{\text{IS}}(G)$, a subgraph that is isomorphic to a Cartesian product of two smaller Glauber graphs:

Lemma 3.4. *Given a graph G and a vertex separator X that partitions $V(G)$ into subgraphs A and B, for every class $T \in \mathcal{S}_{\text{IS}}(G)$,*

$$\mathcal{C}_{\text{IS}}(T) \cong \mathcal{M}_{\text{IS}}(A \setminus N_A(T)) \square \mathcal{M}_{\text{IS}}(B \setminus N_B(T)).$$

(Here we identify the class $\mathcal{C}_{\text{IS}}(T)$ with the subgraph it induces in $\mathcal{M}_{\text{IS}}(G)$.)

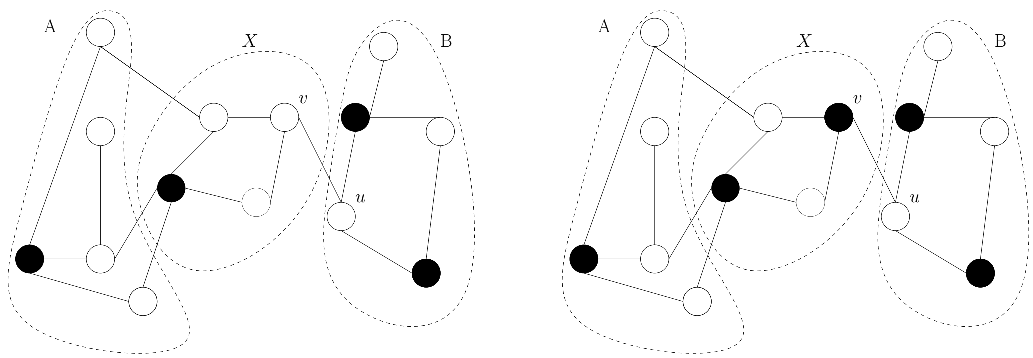

Figure 3.3: Two independent sets in a graph G: S (left) and S' (right), belonging to distinct classes. S and S' differ by a flip, with the separator X inducing the classes to which the sets belong. S' results from adding v to S. $|S'| < |S|$, since S' excludes those independent sets that contain the vertex u.

3.2.2 Rapid mixing of the independent set flip chain when G has bounded carving width

As described in Section 3.2.1, we use a small vertex separator X in G to give a decomposition of $\mathcal{M}_{\text{IS}}(G)$ into subgraphs, each of which has a Cartesian product structure—in which both factor graphs in the product are themselves Glauber graphs. Since Cartesian products preserve flow congestion upper bounds (see Lemma 2.15), this decomposition provides a crucial inductive structure. We analyze this structure in this section.

Lemma 3.5. *Let G be a graph with bounded treewidth t, let $\mathcal{M}_{\text{IS}}(G)$ be as we have defined, and let $\mathcal{S}_{\text{IS}}(G)$ be as we have defined with respect to a small balanced separator X with $|X| \leq t+1$. The number of classes in $\mathcal{S}_{\text{IS}}(G)$ is $O(1)$.*

Proof. The lemma follows from the fact that $|\mathcal{S}_{\text{IS}}(G)| \leq 2^{|X|} \leq 2^{t+1} = O(1)$, where the first inequality is true because each class is identified with a subset of the vertices in X. □

Lemma 3.6. *Let G be a graph with bounded treewidth t and bounded degree Δ, let $\mathcal{M}_{\text{IS}}(G)$ be as we have defined, and let $\mathcal{S}_{\text{IS}}(G)$ be as we have defined with respect to a small balanced separator X with $|X| \leq t+1$. For every pair of classes $\mathcal{C}_{\text{IS}}(T), \mathcal{C}_{\text{IS}}(T') \in \mathcal{S}_{\text{IS}}(G)$, $|\mathcal{C}_{\text{IS}}(T)| = \Theta(1)|\mathcal{C}_{\text{IS}}(T')|$.*

Lemma 3.7. *Let G be a graph, let $\mathcal{M}_{\text{IS}}(G)$ be as we have defined, and let $\mathcal{S}_{\text{IS}}(G)$ be as we have defined with respect to a separator X. Let $\mathcal{C}_{\text{IS}}(T), \mathcal{C}_{\text{IS}}(T') \in \mathcal{S}_{\text{IS}}(G)$ be two classes. No independent set in $\mathcal{C}_{\text{IS}}(T)$ has more than $O(1)$ flips to independent sets in $\mathcal{C}_{\text{IS}}(T')$.*

Lemma 3.8. *Let G be a graph with bounded treewidth t and bounded degree Δ, let $\mathcal{M}_{\text{IS}}(G)$ be as we have defined, and let $\mathcal{S}_{\text{IS}}(G)$ be as we have defined with respect to a small balanced separator X with $|X| \leq t+1$. Let $\mathcal{C}_{\text{IS}}(T), \mathcal{C}_{\text{IS}}(T') \in \mathcal{S}_{\text{IS}}(G)$ be two classes. Suppose there exists at least one flip between an independent set in $\mathcal{C}_{\text{IS}}(T)$ and an independent set in $\mathcal{C}_{\text{IS}}(T')$. Then there exist at least $\Omega(1)|\mathcal{C}_{\text{IS}}(T)|$ flips between independent sets in $\mathcal{C}_{\text{IS}}(T)$ and independent sets in $\mathcal{C}_{\text{IS}}(T')$.*

The proofs of Lemma 3.6, Lemma 3.7, and Lemma 3.8 are in Section B.2. We will use these facts to prove the following, applying the framework from Chapter 2, in Section 3.2:

Lemma 3.9. *Given a graph G with bounded carving width, the natural random walk on the independent set Glauber graph $\mathcal{M}_{\text{IS}}(G)$ has mixing time $\tau(n) = O(n^c)$, where $c = O(1)$.*

To prove Theorem 1.5, however, we need to get rid of the assumption that degree is bounded. We address this issue in Section 3.3.

3.2.3 Abstraction into framework conditions

The observations in Lemmas 3.5 through 3.8 correspond to the set of conditions we gave in Lemma 2.10 (Chapter 2).

The conditions are, given a connected Glauber graph $\mathcal{M}(G)$, on some set of structures over an underlying graph G with n vertices:

1. The vertices of $\mathcal{M}(G)$ can be partitioned into a set $\mathcal{S}$ of classes, where $|\mathcal{S}| = O(1)$.

2. The ratio of the sizes of any two classes in $\mathcal{S}$ is $\Theta(1)$.

3. Given two classes $\mathcal{C}(T), \mathcal{C}(T') \in \mathcal{S}$, no vertex in $\mathcal{C}(T)$ has more than $O(1)$ edges to vertices in $\mathcal{C}(T')$.

4. For every pair of classes that share at least one edge, the number of edges between the two classes is $\Theta(1)$ times the size of each of the two classes.

5. Each class in $\mathcal{S}$ is the Cartesian product of two Glauber graphs $\mathcal{M}(G_1)$ and $\mathcal{M}(G_2)$, each of which can be recursively partitioned in the same way as $\mathcal{M}(G)$.

6. The recursive partitioning mentioned in Condition 5 reaches the base case (graphs with one or zero vertices) in $O(\log n)$ steps.

Conditions 1 through 4 correspond respectively to Lemmas 3.5 through 3.8; Condition 5 corresponds to Lemma 3.4. Condition 6 is satisfied by our chains due to the assumption of bounded treewidth, which ensures a recursive decomposition in which the classes are isomorphic to Cartesian products of Glauber graphs over an underlying graph whose number of vertices has been reduced by a constant factor. Thus all of these conditions are satisfied by the chain on independent sets in graphs of bounded carving width, and Lemma 3.9 will follow from the claim that the above conditions imply rapid mixing. Indeed, we proved the following in Chapter 2 (Lemma 2.9):

Lemma 3.10. *Given a graph $\mathcal{M}(G)$ satisfying the conditions in this section, the expansion of $\mathcal{M}(G)$ is $\Omega(1/n^c)$, where $c = O(1)$.*

We revisit the proof sketch in Section 2.5, restating it in terms more conducive to the Glauber dynamics in this chapter:

Proof Sketch. The idea of the proof is first to partition $\mathcal{M}(G)$ into classes as we have described. By Lemma 3.4, each class $\mathcal{C}(T) \in \mathcal{S}(G)$ is isomorphic to the Cartesian product

$\mathcal{M}(A \setminus N_A(T)) \Box \mathcal{M}(B \setminus N_B(T))$. We make an inductive argument, in which the inductive hypothesis assumes that for each such Cartesian product, the graphs $\mathcal{M}(A \setminus N_A(T))$ and $\mathcal{M}(B \setminus N_B(T))$ have multicommodity flows with congestion $\rho_A \leq c^{\log|V(G)|-1}$, $\rho_B \leq c^{\log|V(G)|-1}$ respectively, for an appropriate constant c. Lemma 2.15 then implies that $\mathcal{C}(T)$ has a flow f_T with congestion $\rho_T \leq c^{\log|V(G)|-1}$.

The inductive step is then to combine the f_T flows for all of the classes, constructing a multicommodity flow f in $\mathcal{M}(G)$ with small congestion. We need to route flow between every pair of Glauber graph vertices S and S' in $\mathcal{M}(G)$. If S and S' belong to the same class $\mathcal{C}(T)$, this is easy: use the same flow that S uses to send its unit to S' in f_T. If $S \in \mathcal{C}(T)$ and $S' \in \mathcal{C}(T') \neq \mathcal{C}(T)$ belong to different classes, we do the following (see Figure 3.2):

1. Find a path from $\mathcal{C}(T)$ to $\mathcal{C}(T')$, where each pair of consecutive classes on the path share at least one edge. Let this path be

$$\mathcal{C}(T) = \mathcal{C}(T_1), \mathcal{C}(T_2), \ldots, \mathcal{C}(T_k) = \mathcal{C}(T').$$

2. Let S send an equal fraction of the S-S' unit (through paths in $\mathcal{C}(T)$) to each $Z \in \mathcal{C}(T)$ that has a neighbor $Y \in \mathcal{C}(T_2)$.

3. For $i = 2, \ldots, k-1$, within $\mathcal{C}(T_i)$, for every Y that receives flow from a neighbor in $\mathcal{C}(T_{i-1})$, let Y send an equal fraction of its S-S' unit (using paths in $\mathcal{C}(T_i)$) to every $Z \in \mathcal{C}(T_i)$ with a neighbor in $\mathcal{C}(T_{i+1})$.

4. For $i = 1, \ldots, k-1$, let each $Z \in \mathcal{C}(T_i)$ having a neighbor $Y \in \mathcal{C}(T_{i+1})$ send its fraction of the $S - S'$ unit to Y across the edge (Z, Y).

5. Let S' receive an equal fraction of the S-S' unit (through paths in $\mathcal{C}(T')$) from each $Y \in \mathcal{C}(T')$ that has a neighbor in $\mathcal{C}(T_{k-1})$.

We specified in Chapter 2 how to route the flow at each intermediate step, making use of the existing flows within each class guaranteed by the inductive hypothesis. We then derived upper bounds on the amount of flow resulting from this routing across any given edge within a class, as well as on the amount of flow across each boundary edge e between classes. We showed that the latter is $O(N)$, where $N = |V(\mathcal{M}(G))|$, i.e. $\rho(e) = \frac{1}{N}O(N) = O(1)$; we showed that the former is at most an $O(1)$ factor times the existing congestion ρ_T.

This leads to a total congestion of $O(1)^l$, where l is the number of levels of induction. The fact that X is a balanced separator implies that $l = O(\log n)$; the lemma now follows from Lemma 2.1. □

Lemma 3.9 now follows.

We will use the phrase "non-hierarchical framework" to describe this set of conditions—which apply to the chains we study when the underlying graph G has bounded carving width.

Although Jerrum, Son, Tetali, and Vigoda [40] did not consider bounded-treewidth graphs generally, these conditions do allow their projection-restriction technique to be applied. In effect, Lemma 3.10 characterizes a sufficient set of conditions for applying Jerrum, Son, Tetali, and Vigoda's technique.

The first main technical contribution of this chapter is in Section 3.3, in which we give an alternative set of conditions—which we will call our "hierarchical framework"—that allows us to handle underlying graphs of unbounded degree (though treewidth still must be bounded), and to handle chains other than the hardcore model. This will allow us to complete the proofs of Theorems 1.5, 1.6, and 1.7.

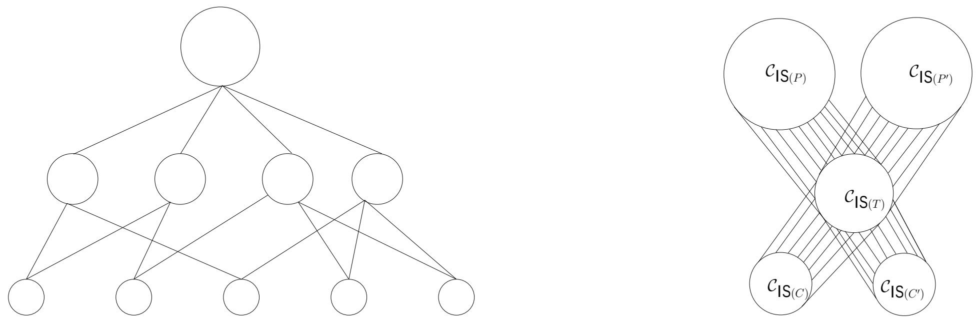

Figure 3.4: Left: a schematic representation of the classes in the independent set Glauber graph and edges between them when degree is unbounded. Right: a class $\mathcal{C}_{\mathrm{IS}}(T)$, with two parents, $\mathcal{C}_{\mathrm{IS}}(P)$ and $\mathcal{C}_{\mathrm{IS}}(P')$, and two children, $\mathcal{C}_{\mathrm{IS}}(C)$ and $\mathcal{C}_{\mathrm{IS}}(C')$. (Classes with larger cardinality are drawn larger.) The parallel edges depict the fact that a child class always has every one of its vertices adjacent to a vertex in a given parent class, and that the edges between any given pair of classes are vertex-disjoint.

3.3 $\lambda = 1$: Unbounded degree

3.3.1 Hierarchical framework

We now sketch a set of "hierarchical" framework conditions that guarantee rapid mixing in the case of unbounded degree (when treewidth is bounded). Several of the chains we consider satisfy these conditions so long as the treewidth of the underlying graph is bounded.

In the original framework, we assumed that the classes were approximately the same size. Although all of the Glauber graphs to which we apply this hierarchical framework satisfy this condition in graphs with bounded carving width, this is not the case when the degree is unbounded. Fortunately, in the case of independent sets, partial q-colorings, dominating sets, and b-edge covers, we can solve this problem with some modifications to the framework.

3.3.2 Independent sets

In the proof of Lemma 3.10, the assumption that the classes were approximately the same size allowed us to argue that even in the worst case, any given class $\mathcal{C}_{\mathrm{IS}}(T)$ can route flow for all pairs of vertices without being too congested, because $\mathcal{C}_{\text{IS}}(T)$ is sufficiently large. Once we discard this assumption, we need to be more explicit in specifying the path through which any given $\mathcal{C}_{\text{IS}}(T)$ routes flow to any given $\mathcal{C}_{\text{IS}}(T')$. Namely, we show that one can engineer the flow so that for any such $\mathcal{C}_{\text{IS}}(T), \mathcal{C}_{\text{IS}}(T')$ pair, every intermediate class $\mathcal{C}_{\text{IS}}(T'')$ that handles flow between sets $S \in \mathcal{C}_{\text{IS}}(T)$ and $S' \in \mathcal{C}_{\text{IS}}(T')$ has a larger cardinality than one of $\mathcal{C}_{\text{IS}}(T)$ or $\mathcal{C}_{\text{IS}}(T')$. This allows us to bound the number of pairs of sets, relative to $|\mathcal{C}_{\text{IS}}(T'')|$, for which $\mathcal{C}_{\text{IS}}(T'')$ carries flow.

To accomplish this, we observe that for any pair of classes $\mathcal{C}_{\text{IS}}(T)$ and $\mathcal{C}_{\text{IS}}(T')$, if there exists *one* flip between an independent set in $\mathcal{C}_{\text{IS}}(T')$ and an independent set in $\mathcal{C}_{\text{IS}}(T)$, then without loss of generality *every* independent set in $\mathcal{C}_{\text{IS}}(T')$ has a flip to some independent set in $\mathcal{C}_{\text{IS}}(T)$. Namely, this flip consists of dropping some vertex v from $T' \subseteq X$ to obtain T. In this case we call $\mathcal{C}_{\text{IS}}(T)$ a *parent* of $\mathcal{C}_{\text{IS}}(T')$, and $\mathcal{C}_{\text{IS}}(T')$ a *child* of $\mathcal{C}_{\text{IS}}(T)$. See Figure 3.4. Since the set of these edges is vertex disjoint, this implies that $|\mathcal{C}_{\text{IS}}(T)| \geq |\mathcal{C}_{\text{IS}}(T')|$. In fact, whenever $T \subseteq T'$, we have $|\mathcal{C}_{\text{IS}}(T)| \geq |\mathcal{C}_{\text{IS}}(T')|$.

Thus for any pair of classes $\mathcal{C}_{\text{IS}}(T)$ and $\mathcal{C}_{\text{IS}}(T')$, one can find paths from the two classes to a "common ancestor", and route flow along these paths, through the common ancestor. Since for every class $\mathcal{C}_{\text{IS}}(T'')$ on this path, either $|\mathcal{C}_{\text{IS}}(T'')| \geq |\mathcal{C}_{\text{IS}}(T)|$ or $|\mathcal{C}_{\text{IS}}(T'')| \geq |\mathcal{C}_{\text{IS}}(T')|$, we are still able to bound the congestion in a fashion similar to the non-hierarchical framework. We make this precise and derive the resulting congestion bounds in Section 3.5.

Recall that in the proof sketch of Lemma 3.10 (Section 3.2.2), for every pair of Glauber graph vertices $S \in \mathcal{C}(T), S' \in \mathcal{C}(T') \neq \mathcal{C}(T)$, we found a sequence of classes $\mathcal{C}(T) = \mathcal{C}(T_1), \mathcal{C}(T_2), \ldots, \mathcal{C}(T_{k-1}), \mathcal{C}(T_k) = \mathcal{C}(T')$, through which to route the $S - S'$ flow. As discussed in Section 3.3, when degree is unbounded, the classes are no longer nearly the same size, and thus if this sequence is chosen carelessly, some $\mathcal{C}(T_i)$ may carry flow for too many $S - S'$ pairs.

The solution is to choose the sequences carefully. This is possible provided that there exists a partial order $<$ on the classes with a unique maximal element, where $\mathcal{C}(T) > \mathcal{C}(T')$ implies $|\mathcal{C}(T)| \geq |\mathcal{C}(T')|$. Under this condition, we can simply choose our sequence of classes so that for some i with $1 \leq i \leq k$, $|\mathcal{C}(T_1)| \leq |\mathcal{C}(T_2)| \leq \cdots \leq |\mathcal{C}(T_i)| \geq |\mathcal{C}(T_{i+1})| \geq \cdots |\mathcal{C}(T_{k-1})| \geq |\mathcal{C}(T_k)|$.

3.3.3 Hierarchical Framework Conditions

The conditions are as follows. Conditions 2 through 4 are new and concern the partial order described above; Condition 1 and Conditions 5 through 7 are as in the non-hierarchical framework.

1. The vertices of $\mathcal{M}(G)$ can be partitioned into a set $\mathcal{S}$ of classes, where $|\mathcal{S}| = O(1)$.

2. There exists a partial order $<$ on the classes in $\mathcal{S}$, such that whenever $\mathcal{C}(T), \mathcal{C}(T') \in \mathcal{S}$ and $\mathcal{C}(T) > \mathcal{C}(T')$, we have $|\mathcal{C}(T)| \geq |\mathcal{C}(T')|$.

3. The partial order $<$ has a unique maximal element.

4. Whenever an edge exists between vertices in $\mathcal{C}(T)$ and $\mathcal{C}(T')$ with $\mathcal{C}(T) > \mathcal{C}(T')$, the number of such edges is $|\mathcal{C}(T')|$.

5. For every pair of classes $\mathcal{C}(T)$ and $\mathcal{C}(T')$ that share an edge, the maximum degree, in $\mathcal{C}(T)$, of a vertex in $\mathcal{C}(T')$, is $O(1)$, and the maximum degree, in $\mathcal{C}(T')$, of a vertex in $\mathcal{C}(T)$, is $O(1)$.

6. Each class in $\mathcal{S}$ is the Cartesian product of two Glauber graphs $\mathcal{M}(G_1)$ and $\mathcal{M}(G_2)$, each of which can be recursively partitioned in the same way as $\mathcal{M}(G)$.

7. The recursive partitioning mentioned in Condition 6 reaches the base case (graphs with one or zero vertices) in $O(\log n)$ levels of recursion.

Lemma 3.11. *Given a graph $\mathcal{M}(G)$ satisfying the conditions in Sectionx 3.3.3, the expansion of $\mathcal{M}(G)$ is $\Omega(1/n^c)$, where $c = O(1)$.*

We defer the proof of Lemma 3.11 to Section 3.5.

3.4 Bounded carving width: application of framework beyond independent sets

3.4.1 q-colorings

We now apply the non-hierarchical framework to q-colorings in graphs of bounded carving width. For reasons that will soon become apparent, we need to generalize to *list colorings*: A *list coloring* of a graph $G = (V, E)$, given a function $L : V \to 2^{[q]}$ assigning a list of colors to each vertex in V, is a coloring of G consistent with L. A *partial list coloring* is a coloring of some of the vertices of G consistent with L.

We consider the Glauber graph $\mathcal{M}_{\text{COL}}(G, L)$, defined as follows: Let the Glauber graph $\mathcal{M}_{\text{COL}}(G, L)$, given an input graph G and a set of colors $[q]$ and a function L as in the definition of list colorings, be the graph whose vertices are the list colorings of G consistent with L, and whose edges are the pairs of list colorings C, C' that differ by a color assignment to exactly one vertex $v \in V(G)$.

The Glauber dynamics is clearly the natural random walk on $\mathcal{M}_{\text{COL}}(G, L)$, with self-loops added in the standard fashion. The following lemma therefore suffices to prove the first claim in Theorem 1.6:

Lemma 3.12. *$\mathcal{M}_{\text{COL}}(G, L)$, defined over a graph G and a list $L : V(G) \to 2^{[q]}$, with $L(v) \geq \delta(v) + 2$ for every $v \in V(G)$, satisfies the conditions of the non-hierarchical framework* *whenever G has bounded carving width and q is fixed.*

Proof. We partition $V(\mathcal{M}_{\text{COL}}(G, L))$ into classes induced by a small balanced separator X, where each class is identified with a list coloring T of $G[X]$. This partitioning satisfies Condition 5 since each class $\mathcal{C}_{\text{COL}}(T)$ consists of the tuples of the form (C_A, C_B), where C_A is a valid list coloring of $G[A]$, and C_B is a valid list coloring of $G[B]$—with A and B being the mutually disconnected subsets of $V(G)$ resulting from the removal of X. Here, we adjust the list $L(u)$ for each $u \in N_A(X) \cup N_B(X)$, removing from $L(u)$ every color that is assigned to a neighbor of u in X under the coloring C.

The subproblems on A and B are independent, and that a flip within $\mathcal{C}_{\text{COL}}(T)$ corresponds to a flip within either A or B but not both. Furthermore, the condition that $L(u) \geq \delta(u) + 2$ is preserved even after L is modified, since every color removed from $L(u)$ corresponds to a neighbor of u in X—i.e. a neighbor that is not part of the subproblem on A or B. Condition 5 follows.

Condition 1 follows from the fact that $|X|$ and q are bounded. Condition 2 can be seen from the bounded carving width of G by considering the following mapping $f : V(\mathcal{M}_{\text{COL}}(G, L)) \rightarrow \mathcal{C}_{\text{COL}}(T)$ for any T: given a list coloring $C \in V(\mathcal{M}_{\text{COL}}(G, L))$, let $C' = f(C)$ be the list coloring that (i) agrees with T on its restriction to X, (ii) agrees with C on its assignment of colors to all vertices having no neighbor in X, and (iii) is consistent with both (i) and (ii) on its assignment of colors to neighbors of vertices in X.

We can always satisfy (iii) because for each $u \in N_A(X) \cup N_B(X)$, we have $|L(u)| \geq \delta(u) + 2$. (There may be multiple list colorings satisfying (iii); resolve ambiguity in defining $f(C)$ via an arbitrary ordering on the list colorings of G.) Condition 4 follows from a similar mapping.

Condition 3 is follows from the definition of a flip; Condition 6 follows from the bounded carving width of G. □

3.4.2 b-edge covers and b-matchings

For b-edge covers and b-matchings, we now apply the non-hierarchical framework in graphs of bounded carving width. As with independent sets, dealing with unbounded degree in b-edge covers requires the hierarchical framework.

Lemma 3.13. *Given an input graph G of bounded carving width, the Glauber dynamics on b-matchings and on b-edge covers satisfy the conditions of the non-hierarchical framework, when the maximum value of the function b is bounded.*

Proof. The proof involves verifying that the chain on b-edge covers satisfies the non-hierarchical framework conditions, as we have done for independent sets. There are a few additional details, however.

In defining a b-edge cover, we are selecting subsets of edges instead of vertices. Thus, to define our flip chain on b-edge covers, we modify the flip chain on independent sets in the natural way: dropping or adding edges instead of vertices. The corresponding Glauber graph $\mathcal{M}_{\text{BEC}}(G)$ is connected, since every b-edge cover has a path in $\mathcal{M}_{\text{BEC}}(G)$ to the trivial b-edge cover (where every edge is selected). We identify each class $\mathcal{C}_{\text{BEC}}(T)$ with the set T of edges chosen incident to vertices in X. Since degree is bounded and $|X| \leq t$, there are $O(1)$ classes, satisfying Condition 1.

Given a class $\mathcal{C}_{\text{BEC}}(T)$, we pass recursively to subproblems on A and B, where we update $b(v)$ for each $v \in A \cup B$ according to the number of edges in T incident to v. That is, for each vertex u selected in T, and for each edge (u, v) with $v \in A$ (similarly $v \in B$), decrement $b(v)$ when passing to the subproblem on A (similarly B). The choices made in the A subproblem and the B subproblem are independent, giving the required Cartesian product structure for Condition 5, and there are still $O(\log n)$ levels of recursion, satisfying Condition 6. For Condition 3, the proof is the same as for independent sets. Conditions 2 and 4 follow from a similar mapping argument to that in the proof of Lemma 3.6.

The proof for b-matchings is similar to that for b-edge covers. □

3.4.3 Maximal independent sets and maximal b-matchings

The main idea of applying the framework to maximal independent sets and maximal b-matchings is similar to that for independent sets, b-matchings, and b-edge covers, but some adaptation is required: the definition of a flip is somewhat different, and the proof that classes have the required Cartesian product structure has a few more details. We thus defer dealing with these chains to Section 3.7.6.

3.5 Hierarchical framework

In this section we complete the proof of the unbiased case of Theorem 1.5 and Theorem 1.6, by fully specifying the hierarchical framework, and showing that the chain on independent sets satisfies the conditions. Fully proving Theorem 1.7 and Theorem 1.8 requires some adaptation of the framework, which we defer to Section 3.7.

3.5.1 Proof that conditions of the hierarchical framework imply rapid mixing

We are ready to prove the counterpart of Lemma 3.10 for the hierarchical framework, from which the unbiased case of Theorem 1.5 will follow.

Lemma 3.11. *Given a graph $\mathcal{M}(G)$ satisfying the conditions in Sectionx 3.3.3, the expansion of $\mathcal{M}(G)$ is $\Omega(1/n^c)$, where $c = O(1)$.*

Proof. We use the scheme in the proof of Lemma 3.10, with the following specification: when routing flow from $S \in \mathcal{C}(T)$ to $S' \in \mathcal{C}(T') \neq \mathcal{C}(T)$, we find a sequence of classes $\mathcal{C}(T) = \mathcal{C}(T_1), \mathcal{C}(T_2), \ldots, \mathcal{C}(T_{k-1}), \mathcal{C}(T_k) = \mathcal{C}(T')$ as before, where each consecutive pair of classes in the sequence shares an edge in $\mathcal{M}(G)$. In the proof of Lemma 3.10, this sequence was arbitrary; we now require that, under the partial order $<$ in Condition 2, for some $1 \leq i \leq k$, $\mathcal{C}(T_1) < \cdots < \mathcal{C}(T_i) > \mathcal{C}(T_{i+1}) > \cdots > \mathcal{C}(T_k)$; Condition 3 guarantees that this requirement can be satisfied.

We now bound the resulting congestion. As in the proof of Lemma 3.10, for $i = 2, \ldots, k-1$, the congestion added to edges in $\mathcal{C}(T_i)$ in the inductive step is at most $N^2/(|Y_i||Z_i|) \cdot c^{\log n - 1}$. Unfortunately, without assuming that the classes are approximately the same size, we can no longer say that $|Y_i| = \Omega(N)$ or $|Z_i| = \Omega(N)$. Instead, we argue as follows: thanks to the choice of our sequence, for every pair of classes $\mathcal{C}(T)$ and $\mathcal{C}(T')$ that use a given class T_i to route flow, either $|\mathcal{C}(T_i)| \geq |\mathcal{C}(T)|$ (and $|\mathcal{C}(T_i)| \geq |\mathcal{C}(T_{i-1})|$) or $|\mathcal{C}(T_i)| \geq |\mathcal{C}(T')|$ (and $|\mathcal{C}(T_i)| \geq |\mathcal{C}(T_{i+1})|$). Assume the former case without loss of generality. For every pair of classes $\mathcal{C}(T)$ and $\mathcal{C}(T')$ that use the edges between $\mathcal{C}(T_{i-1})$ and $\mathcal{C}(T_i)$, $|\mathcal{C}(T)| \leq |\mathcal{C}(T_{i-1})|$, and therefore the number of pairs S, S' of Glauber graph vertices that use these edges is at most

$$\sum_{T,T':|\mathcal{C}(T)| \leq |\mathcal{C}(T_{i-1})|} |\mathcal{C}(T)||\mathcal{C}(T')| \leq N|\mathcal{S}||\mathcal{C}(T_{i-1})| = O(1)N|\mathcal{C}(T_{i-1})|.$$

Therefore, since there are $|\mathcal{C}(T_{i-1})|$ edges between $\mathcal{C}(T_{i-1})$ and $\mathcal{C}(T_i)$ (by Condition 4), each such edge carries at most $N|\mathcal{S}||\mathcal{C}(T_{i-1})|/|\mathcal{C}(T_{i-1})| = O(N)$ units of flow, giving $O(1)$ congestion.

To bound congestion within $\mathcal{C}(T_i)$, we specify the routing of flow from $\mathcal{Y}_i$ (the set of vertices on the $\mathcal{C}(T_{i-1}), \mathcal{C}(T_i)$ boundary) to $\mathcal{Z}_i$ (the set of vertices on the $\mathcal{C}(T_i), \mathcal{C}(T_{i+1})$ boundary) as follows: first let each $Y \in \mathcal{Y}_i$ send an equal fraction of its flow—of which it receives $O(N)$ units from each of $O(1)$ edges—to *every vertex in* $\mathcal{C}(T_i)$, using the flow that is assumed to exist within $\mathcal{C}(T_i)$ to route the flow. Then let each $Z \in \mathcal{Z}_i$ receive its flow similarly from all vertices in $\mathcal{C}(T_i)$. The resulting congestion across each edge is at most

$$(2O(N)/|\mathcal{C}(T_i)|) \cdot c^{\log n - 1}|\mathcal{C}(T_i)|/N \leq c^{\log n},$$

for a constant c. This gives the desired congestion bound, proving the lemma. □

We now prove Lemma 3.14 by tracing the polynomial factors in the proof of Lemma 3.11:

Lemma 3.14. *Suppose a Glauber graph $\mathcal{M}(G)$ satisfies the conditions of the hierarchical framework. Then the mixing time of the corresponding Glauber dynamics is*

$$O(((2(K+1))^{2\log n}) \cdot \Delta_{\mathcal{M}}^2 \log N),$$

where $\Delta_{\mathcal{M}}$ is the maximum degree of the Glauber graph $\mathcal{M}(G)$, $n = |V(G)|$, K is the number of classes in the partition, and $N = |V(\mathcal{M}(G))|$.

Proof. The analysis is similar to the proof of Lemma 2.10, with the following modifications: each edge set $\mathcal{E}(T, T')$ from $\mathcal{C}(T)$ to a parent $\mathcal{C}(T')$ has $|\mathcal{E}(T, T')| = |\mathcal{C}(T)|$. Therefore, outbound flow along each edge in such an edge set is at most $N|\mathcal{C}(T)|/|\mathcal{E}(T, T')| = N$: each vertex (all vertices in $\mathcal{C}(T)$ are boundary vertices) then receives from each other vertex at most $N/|\mathcal{C}(T)|$ units. As we will show shortly (see analysis of through flow below), edges to children each carry at most $K|\mathcal{C}(T)|$. Thus we will count the flow resulting from edges to children with through flow.

Inbound flow is symmetric. The result is to scale the amount of flow across each edge internal to $\mathcal{C}(T)$ by a factor of $2N/|\mathcal{C}(T)|$.

For through flow (including the outbound flow to children as described above), each boundary vertex in $\mathcal{C}(T)$ carrying flow from (or to) a set of child classes $\{\mathcal{C}(T_1'), \ldots, \mathcal{C}(T_k')\}$ carries at most $\sum_{i=1}^{k} N|\mathcal{C}(T_i')|K_i/(|\mathcal{C}(T_i')|)$ units, where K_i is the number of classes descendent from $\mathcal{C}(T_i')$, including $\mathcal{C}(T_i')$ itself. This sum is at most NK. Each boundary vertex carrying flow from (or to) an ancestor similarly carries at most NK units. Thus through flow contributes a factor of $2NK/|\mathcal{C}(T)|$.

The resulting overall congestion is therefore at most

$$(2(K+1))^{\log n},$$

and applying Lemma 2.8 and Lemma 2.1 gives the resulting mixing bound. □

3.5.2 Independent sets

We now finish the proof of the unbiased case of Theorem 1.5.

Verification of conditions

To show that the chain on independent sets satisfies the conditions of the hierarchical framework when treewidth is bounded (but degree is unbounded), we first define a partial order $<$ on the classes in $\mathcal{S}_{\text{IS}}(G)$. Recall that these are the classes induced by the separator X in the underlying graph G. For $\mathcal{C}_{\text{IS}}(T), \mathcal{C}_{\text{IS}}(T') \in \mathcal{S}_{\text{IS}}(G)$, let $\mathcal{C}_{\text{IS}}(T) < \mathcal{C}_{\text{IS}}(T')$ if $T \subseteq T'$ and $T \neq T'$. Call $\mathcal{C}_{\text{IS}}(T)$ an *ancestor* of $\mathcal{C}_{\text{IS}}(T')$, and $\mathcal{C}_{\text{IS}}(T')$ a *descendant* of $\mathcal{C}_{\text{IS}}(T)$. If $\mathcal{C}_{\text{IS}}(T)$ covers $\mathcal{C}_{\text{IS}}(T')$ in this relation, call $\mathcal{C}_{\text{IS}}(T)$ a *parent* of $\mathcal{C}_{\text{IS}}(T')$, and $\mathcal{C}_{\text{IS}}(T')$ a *child* of $\mathcal{C}_{\text{IS}}(T)$.

We now prove that the chain on independent sets satisfies the conditions of the hierarchical framework on graphs of bounded treewidth.

Lemma 3.15. *Given a graph G with fixed treewidth $t-1$, the hardcore Glauber dynamics on the independent sets of G satisfies the conditions of the hierarchical framework.*

Proof. Let $\mathcal{M}_{\text{IS}}(G)$, X, and $\mathcal{S}_{\text{IS}}(G)$ be as previously defined. We have already proven Condition 1 and Conditions 4 through 7 in Lemmas 3.5 through 3.8.

The partial order we have defined satisfies Condition 2 because for every parent class $\mathcal{C}_{\text{IS}}(T)$ and child class $\mathcal{C}_{\text{IS}}(T')$, the recursive subproblems in the Cartesian product comprising $\mathcal{C}_{\text{IS}}(T')$ are at least as constrained as the subproblems in the product comprising $\mathcal{C}_{\text{IS}}(T)$. That is, $\mathcal{C}_{\text{IS}}(T)$ and $\mathcal{C}_{\text{IS}}(T')$ are each a Cartesian product of two smaller Glauber graphs on the independent sets in subgraphs A_T and B_T of G, and subgraphs $A_{T'}$ and $B_{T'}$ of G respectively. We have $V(A_{T'}) \subseteq V(A_T)$ and $V(B_{T'}) \subseteq V(B_T)$, where the set $V(A_T) \setminus V(A_{T'})$ consists of the vertices in A that have a neighbor in T' but not in T.

Condition 3 follows from the fact that the empty independent set is the unique set that is an ancestor of all other independent sets.

□

It now follows by Lemma 2.8 that $\mathcal{M}(G)$ has expansion $\Omega(1/n^{O(1)})$, and Theorem 1.5 follows from this fact and from Lemma 2.1. More precisely, observing that the number of classes in the partition is at most 2^{t+1} and applying Lemma 3.14 gives the bound claimed in Theorem 1.5, namely

$$O(((1+\hat{\lambda})\hat{\lambda})^2(1+\log\hat{\lambda})n^{2(t+2)(1+\log\hat{\lambda})+5}),$$

where $\hat{\lambda} = \max\{\lambda, 1/\lambda\}$. (We will give the machinery that justifies the terms involving $\lambda, \hat{\lambda}$ in Section 3.6.)

3.5.3 Partial q-colorings

We now prove the unbiased case of the claim about partial colorings in Theorem 1.6:

Let $\mathcal{M}_{\text{PCOL}}(G, L)$, given an input graph G and function $L : V(G) \to 2^{[q]}$, be the graph whose vertices are the partial list colorings of G, and whose edges are the pairs of partial list colorings that differ by the removal or addition of a color assignment to a single vertex.

We show that this Glauber graph satisfies the conditions of the hierarchical framework:

Lemma 3.16. *Given a graph G with bounded carving width and list function $L : V(G) \to 2^{[q]}$, where $q \geq \Delta+2$ is fixed and $L(v) \geq \delta(v)+2$ for all $v \in V(G)$, the Glauber graph $\mathcal{M}_{\text{PCOL}}(G, L)$ has expansion $\Omega(1/n^c)$, where $c = O(1)$.*

Proof. The partitioning is the same as in the proof of Lemma 3.12, except that we allow each class to be identified with a partial list coloring of X. Condition 1, Condition 5, Condition 6, and Condition 7 can be seen as before. For Conditions 2 and 3, the partial order is analogous to the partial order for independent sets: given partial list colorings C and C' of X, let C be a parent of C' if C and C' agree except for a single vertex to which C' assigns a coloring and C does not. Condition 4 follows from this definition. The lemma follows. □

We obtain the bound in Theorem 1.6 via the observations in Section B.1.

3.6 All $\lambda > 0$

Until now, we have only considered the unbiased versions of our chains. In this section we complete the proof of Theorem 1.5, for arbitrary fixed $\lambda > 0$. To do so, we need to introduce the standard notion of *conductance* [67], which extends the definition of expansion in the natural way to the setting of a weighted graph.

3.6.1 (Weighted) Conductance

The conductance is defined with respect to a *stationary* distribution π induced by a random walk. The stationary distribution is the distribution to which the random walk converges in the limit. The convergence requires mild conditions: (i) that walk be *ergodic*, meaning that the Glauber graph is connected; (ii) that the walk be *reversible*; and (iii) that the walk be *lazy*.

Laziness means that with constant probability the walk stays at the current vertex at any step; reversibility means that for every pair of sets $S, S' \in \mathcal{M}_{\text{IS}}(G)$, we have

$$\pi(S)P(S, S') = \pi(S')P(S', S),$$

where $P(S, S')$ denotes the probability that $X_{t+1} = S'$, given that $X_t = S$.

The Glauber dynamics on independent sets satisfies these conditions, and our other Glauber dynamics satisfy them as well.

In the case of the Glauber dynamics on independent sets, the stationary distribution π evaluates to

$$\pi(S) = \lambda^{|S|}/Z(\mathcal{M}_{\text{IS}}(G)),$$

where for each of our Glauber graphs $\mathcal{M}(G)$,

$$Z(\mathcal{M}(G)) = \sum_{S \subseteq V(\mathcal{M}(G))} \lambda^{|S|}$$

is the normalizing constant. For all independent sets S in G, and for all S' such that $|S \setminus S'| = 1$,

$$\pi(S)P(S, S') = \pi(S')P(S', S) = \left(\frac{1}{n(1+\lambda)}\right)\left(\frac{\lambda^{|S|}}{Z(\mathcal{M}_{\text{IS}}(G))}\right),$$

where $n = |V(G)|$.

For dominating sets and partial q-colorings, we define the same distribution; for b-edge covers we define the analogous distribution over edges.

Remark 3.1. *For each of our Glauber graphs $\mathcal{M}(G)$, the probability transition function $P(S, S')$, viewed as a matrix, is in fact the adjacency matrix of an edge-weighted version of $\mathcal{M}(G)$, ignoring self loops.*

That is: Given a Glauber graph $\mathcal{M}(G)$ and a Markov chain on $\mathcal{M}(G)$ with stationary distribution π and probability transition function P, assign the weight $\pi(S)$ to each vertex S of $\mathcal{M}(G)$, and assign the weight $Q(S, S') = \pi(S)P(S, S')$ to each edge (S, S').

Extend the definition of a Cartesian graph product to the weighted graphs described in this section, so that for vertices $g \in V(G), h \in V(H)$, the weight of the tuple $(g, h) \in V(G \Box H)$ is $\pi(g, h) = \pi_G(g)\pi_H(h)$, where π_G and π_H are the vertex weight functions for G and H respectively. Let the weight of each edge e between (g, h) and (g', h') be

$$Q(e) = \pi_H(h)(\Delta_G Q_G(g, g'))/(\Delta_G + \Delta_H),$$

if $g \neq g'$ and $h = h'$, and

$$Q(e) = \pi_G(g)(\Delta_H Q_H(h, h'))/(\Delta_G + \Delta_H),$$

if $g = g'$ and $h \neq h'$, where Q_G and Q_H are the edge weight functions for G and H, and Δ_G and Δ_H are the maximum degrees of G and H.

For the self loop $e = ((g, h), (g, h))$, let

$$Q(e) = \pi(g,h) - \sum_{(g'',h):g''\neq g} Q((g,h),(g'',h)) - \sum_{(g,h''):h''\neq h} Q((g,h),(g,h'')),$$

Lemma 3.17. *Given the extended definition of Cartesian products, the stationary distribution π in the discussion leading to Remark 3.1, and the resulting vertex and edge weights as in the weighted definition of a Glauber graph, for each of our Glauber graphs $\mathcal{M}(G)$ and for each class $\mathcal{C}(T) \cong \mathcal{M}(A)\Box\mathcal{M}(B)$, and for each $S \in V(\mathcal{M}(A)), S' \in V(\mathcal{M}(B))$, the following facts hold:*

1. *The vertex weight of $S \cup S' \cup T$ in $\mathcal{M}(G)$ is equal to*

$$\pi_{\mathcal{M}(G)}(S \cup S' \cup T) = \pi_{\mathcal{C}(T)}(S, S')\pi_{\mathcal{M}(G)}(\mathcal{C}(T)),$$

where $\pi_{\mathcal{M}(G)}(\mathcal{C}(T))$ is defined as $\sum_{U\in\mathcal{C}(T)} \pi_{\mathcal{M}(G)}(U)$, and

2. *For all S'' with $|S \setminus S''| = 1$, the weight in $\mathcal{M}(G)$ of the edge e between $S \cup S' \cup T$ and $S'' \cup S' \cup T$ is*

$$Q_{\mathcal{M}(G)}(e) = Q_{\mathcal{C}(T)}(e)\pi_{\mathcal{M}(G)}(\mathcal{C}(T))\frac{|V(A)| + |V(B)|}{|V(G)|}.$$

Proof. We have

$$\frac{\lambda^{|S|+|S'|+|T|}}{Z(\mathcal{M}(G))} = \pi_{\mathcal{M}(A)}(S)\pi_{\mathcal{M}(B)}(S')\lambda^{|T|} \cdot \frac{Z(\mathcal{M}(A))Z(\mathcal{M}(B))}{Z(\mathcal{M}(G))} = \pi_{\mathcal{C}(T)}(S, S')\pi_{\mathcal{M}(G)}(\mathcal{C}(T)),$$

and

$$\begin{aligned}
Q(e) &= \frac{1}{|V(G)|(\lambda+1)} \cdot \frac{\lambda^{|S|+|S'|+|T|}}{Z(\mathcal{M}(G))} \\
&= Q_{\mathcal{M}(A)}(S, S'')\pi_{\mathcal{M}(B)}(S')\lambda^{|T|} \cdot \frac{Z(\mathcal{M}(A))Z(\mathcal{M}(B))}{Z(\mathcal{M}(G))} \cdot \frac{|V(A)|}{|V(G)|} \\
&= Q_{\mathcal{C}(T)}(e)\lambda^{|T|} \cdot \frac{Z(\mathcal{M}(A))Z(\mathcal{M}(B))}{Z(\mathcal{M}(G))} \cdot \frac{|V(A)|+|V(B)|}{|V(G)|} \\
&= Q_{\mathcal{C}(T)}(e)\pi_{\mathcal{M}(G)}(\mathcal{C}(T))\frac{|V(A)|+|V(B)|}{|V(G)|}.
\end{aligned}$$

□

Given a lazy, reversible, ergodic random walk on a weighted graph $\mathcal{M} = (\mathcal{V}, \mathcal{E})$ with stationary distribution π and probability matrix $P : \mathcal{V} \times \mathcal{V} \to [0, 1]$, the *conductance* is the quantity

$$\phi(\mathcal{M}) = \min_{\mathcal{S} \subseteq \mathcal{V} : 0 < \pi(\mathcal{S}) \leq 1/2} \frac{Q(\mathcal{S}, \mathcal{V} \setminus \mathcal{S})}{\pi(\mathcal{S})},$$

where for sets $\mathcal{S}, \mathcal{S}' \subseteq \mathcal{V}$,

$$\pi(\mathcal{S}) = \sum_{S \in \mathcal{S}} \pi(S),$$

and

$$Q(\mathcal{S}, \mathcal{S}') = \sum_{S \in \mathcal{S}, S' \in \mathcal{S}'} Q(S, S'),$$

and where $Q(S, S') = \pi(S)P(S, S')$ given $S, S' \in \mathcal{V}$.

We now extend the definitions of multicommodity flows and congestion: Let a multicommodity flow f in a graph $\mathcal{M} = (\mathcal{V}, \mathcal{E})$ be defined as before, except that each pair of vertices $S, S' \in \mathcal{V}$ exchanges $\pi(S)\pi(S')$ units of flow in each direction. Let

$$\rho = \sum_{e \in \mathcal{E}} \frac{f(e)}{Q(e)}.$$

The following generalizations of Lemma 2.8 and Lemma 2.1 relate the conductance, congestion, and mixing time [67]:

Theorem 3.1. *Given a multicommodity flow with congestion ρ in a graph $\mathcal{M}$, the conductance ϕ satisfies $\phi \geq 1/(2\rho)$.*

Theorem 3.2. *The mixing time of a Markov chain with state space Ω, stationary distribution π^*, and conductance at least ϕ is at most*

$$\tau = O(\phi^{-2} \log(1/(\pi^*_{\min}))),$$

where

$$\pi^*_{\min} = \min_{u \in \Omega} \pi^*(u).$$

3.6.2 Analysis of flow construction

We now complete the proof of Theorem 1.5. It suffices to show the following lemma:

Lemma 3.18. *The flow f constructed in $\mathcal{M}_{\text{IS}}(G)$ in the proof of Lemma 3.11, adjusted so that $\mathcal{M}_{\text{IS}}(G)$ is weighted according to the parameter $\lambda > 0$, and so that each pair of sets S, S' exchanges $\pi(S)\pi(S')$ units of flow, results in a congestion factor gain of at most $\rho = O(1)$ at each of the $O(\log n)$ levels of induction, resulting in at most polynomial overall congestion. The same holds for the flip chain on partial q-colorings.*

Proof. We use the same inductive argument, with the following adjustment: if $\mathcal{C}(T)$ is a descendant of $\mathcal{C}(T_{i-1})$, and $\mathcal{C}(T_{i-1})$ is a child of $\mathcal{C}(T_i)$, where $\mathcal{C}(T)$ uses the edges between $\mathcal{C}(T_{i-1})$ and $\mathcal{C}(T_i)$ to send flow to $\mathcal{C}(T')$, then distribute this flow as before across these edges, but now let each edge carry flow in proportion to its weight. We have $\pi(\mathcal{C}(T)) = O(1)\pi(\mathcal{C}(T_{i-1}))$—because for every independent set $S \in \mathcal{C}(T)$ there exists a distinct independent set $S' \in \mathcal{C}(T_{i-1})$ with $\pi(S') = (1/\lambda^{|T \setminus T_{i-1}|})\pi(S)$, namely $S' = S \setminus (T \setminus T_{i-1})$.

Each edge (S, S') with $S \in \mathcal{C}(T_i), S' \in \mathcal{C}(T_{i-1})$, satisfies $Q(S, S') = \Omega(\pi(S)/n)$ (where the constant-factor difference depends on λ). Thus the congestion along these edges is still $O(n)$. We then allow each vertex $S \in \mathcal{C}(T_i)$, having received at most $O(\pi(S))$ units along each of $O(1)$ incoming edges from child classes, to distribute these units to all other vertices in $\mathcal{C}(T_i)$ according to their weight. That is, let S send $O(\pi(S)\pi(S'')/\pi(\mathcal{C}(T_i)))$ units to each $S'' \in \mathcal{C}(T_i)$. By the inductive hypothesis and Lemma 2.15, a flow f_{T_i} already exists under which S sends $\pi(S)\pi(S'')/(\pi(\mathcal{C}(T_i)))^2$ units to S'' at a congestion cost of $O(nc^{\log n - 1})$, for appropriate constant c. Thus letting S send $O(\pi(S)\pi(S'')/\pi(\mathcal{C}(T_i)))$ units to S'' reduces the amount sent across each edge by at least a factor of $1/\pi(\mathcal{C}(T_i))$, while the weight of each edge increases when passing from the factor graphs of $\mathcal{C}(T_i)$ to $\mathcal{M}(G)$ by at most the same factor—up to the change in degree of the flip graph—by Lemma 3.17. This gives congestion cost at most $O(nc^{\log n})$. □

Specific polynomial bounds

We now revisit the discussion in Section B.1. In Theorem 1.5, the $1 + \log \hat{\lambda}$ term in the exponent comes from observing that, in Lemma 3.14, we can replace the $K + 1$ term with $K\lambda^{t+1} + 1$ —since in the proof of Lemma 3.14, this is the factor by which the flow carried into a class from a child class increases when adjusting for the weights induced by the parameter λ. A similar analysis gives the result for partial q-colorings in Theorem 1.6.

The Ising and Potts models

As we discussed in the introduction, one can apply our framework to the Ising and Potts models when the parameters of these models are fixed. We do not give the standard definitions of these models or a detailed proof. Instead, we simply observe that in a graph of bounded carving width, the same decomposition technique given in the non-hierarchical framework applies, with the following modification. Since all assignments of spins are possible, so instead of considering the cardinalities of the sets we consider weights of configurations under the standard (exponentiated) energy functions. One can verify the conditions under this modification, using the insight that the weights of the classes differ by a constant factor from one another, since this factor is determined only by evaluating the energy function at a constant number of edges and vertices.

3.7 Dealing with non-independence

The Glauber dynamics on independent sets induces a Glauber graph, $\mathcal{M}_{\mathrm{IS}}(G)$, that behaves well when partitioned into classes. That is, each class $\mathcal{C}_{\mathrm{IS}}(T)$ is isomorphic to the Cartesian product of two Glauber graphs on subgraphs of G. As we will see in Section 3.5, the Glauber dynamics on partial q-colorings is similarly well-behaved. Unfortunately, as we will discuss in Section 3.7, this does not hold for dominating sets or, in the unbounded-degree case, for b-edge covers. In these problems, the selection $T \subseteq X$ of vertices (or edges) in the separator X with which the class $\mathcal{C}(T)$ is identified imposes constraints on what vertices (or edges) can be chosen in the two subgraphs $A \cup B = G \setminus X$—and choices in A may invalidate those in B.

In the case of maximal independent sets and maximal b-matchings, the situation is worse: the classes induced by selection of $T \subseteq X$ may not even be internally connected.

We address both of these problems by relaxing the framework condition that each class $\mathcal{C}(T)$ be a Cartesian product of Glauber graphs, and instead require that each class $\mathcal{C}(T)$ be the (not necessarily disjoint) union of Cartesian products of Glauber graphs, satisfying certain conditions. We fully specify this condition, and show that the remaining chains satisfy it, this section. That is, we complete the proofs of Theorem 1.7 and Theorem 1.8. The principal problem is that when attempting to partition the Glauber graph into classes as we did for independent sets, the resulting classes are not isomorphic to Cartesian products of Glauber graphs. For instance, in the case of b-edge covers, we wish to identify a class of b-edge covers with the set T of edges selected within the subgraph of G induced by the separator X. Unfortunately, the resulting subproblems on A and B (where, as before, $A \cup B = V(G) \setminus X$) are not independent. This is because for each vertex $x \in X$, the sum of the number of incident edges selected in A and those in B must be at least $b(x)$, so the choices made in A depend on those made in B, and vice versa.

The solution is as follows: we divide each class into smaller (not necessarily disjoint) "subclasses", each of which is a Cartesian product of smaller Glauber graphs on b-edge covers. We detail this in Section 3.7.2.

We encounter a similar problem in the case of dominating sets, with an additional challenge that will require us to generalize the definition of a dominating set into what we call the "constrained Steiner dominating set" problem. We give the full details in Section 3.7.3.

For maximal independent sets and maximal b-matchings (Section 3.7.6), the non-hierarchical framework is more natural, as we require bounded degree. The challenge is twofold: first, we need to define the Glauber graphs and show that they are connected. Secondly, we need to deal with non-independence as with b-edge covers and dominating sets—with the additional challenge, as we will see, that the classes are not necessarily internally connected.

3.7.1 Framework relaxation to allow non-independence

Lemma 3.19. *Suppose a Glauber graph $\mathcal{M}(G)$ satisfies the conditions of the hierarchical framework in Section 3.3.3, except for Condition 6. Suppose further that each class $\mathcal{C}(T) \in \mathcal{S}$ is the union of at most $O(1)$ subclasses $\mathcal{C}(T) = \mathcal{C}(T_1) \cup \mathcal{C}(T_2) \cup \cdots \mathcal{C}(T_k)$, where for $i = 1, \ldots, k-1$:*

1. $|\mathcal{C}(T_i)| = \Theta(1)|\mathcal{C}(T_{i+1})|$, *and*

2. $\mathcal{C}(T_i)$ and $\mathcal{C}(T_{i+1})$ share at least $\Omega(1)|\mathcal{C}(T_i)| = \Omega(1)|\mathcal{C}(T_{i+1})|$ vertices.

Suppose further that for $i = 1, \ldots, k$, $\mathcal{C}(T_i)$ is isomorphic to the Cartesian product of two Glauber graphs $\mathcal{M}(G_1)$ and $\mathcal{M}(G_2)$, each of which can be recursively partitioned in the same way as $\mathcal{M}(G)$.

Then the expansion of $\mathcal{M}(G)$ is $\Omega(1/n^c)$, where $c = O(1)$.

Proof. It suffices to construct a multicommodity flow among the subclasses in $\mathcal{S}(T)$ and bound its congestion. By the inductive hypothesis and Lemma 2.15, each subclass $\mathcal{C}(T_i)$ has an internal flow f_{T_i} with congestion $\rho_{T_i} \leq c^{\log n - 1}$. We would like to derive a flow f_T with congestion $\rho_T \leq c^{\log n - 1}$; this will allow the rest of the proof of Lemma 3.11 to be applied.

The solution is to follow the proof sketch of Lemma 3.9: for Glauber graph vertices $S \in \mathcal{C}(T_i), S' \in \mathcal{C}(T_j) \neq \mathcal{C}(T_i)$, send the $S - S'$ flow through the classes $\mathcal{C}(T_{i+1}), \mathcal{C}(T_{i+1}), \ldots, \mathcal{C}(T_{j-1})$. For $l = i, \ldots, j$, let $\mathcal{Y}_l$ and $\mathcal{Z}_l$ be as in the proof of Lemma 3.9, except that $Z_l = Y_{l+1}$. That is, the boundary vertices in consecutive pairs of classes on the path are shared between the two classes. The routing of flow within each class on the path is the same as in Lemma 3.9.

The resulting congestion bound is the same as in Lemma 3.9. The only concern is that since the subclasses may not be disjoint, each edge within a subclass may incur congestion from multiple steps on the path. However, because the number of classes is $O(1)$, there are $O(1)$ such steps, and thus the factor by which this increases congestion is $O(1)$. □

3.7.2 b-edge covers in the relaxed hierarchical framework

To finish the proof of Theorem 1.7, it now suffices to show that the chains on dominating sets and b-edge covers satisfy the conditions of the hierarchical framework when treewidth is bounded. We begin with b-edge covers.

Let $\mathcal{C}_{\text{BEC}}(T)$ be defined as in Section 3.4.2, with the following modification: define each class $\mathcal{C}_{\text{BEC}}(T)$ so that T is identified with the set of selected edges *both of whose endpoints* are in X, instead of including all edges incident to vertices in X.

We now divide each class $\mathcal{C}_{\text{BEC}}(T)$ into subclasses. For each $x \in X$, let $\delta_T(x)$ be the number of edges incident to x (from neighbors in X) that are selected in T. Let $b'(x) = b(x) - \delta_T(x)$, i.e. decrease $b(x)$ by the number of edges incident to x selected in X. Clearly for each $x \in X$, every valid b-edge cover in T includes numbers of edges from neighbors in A and B that sum to at least $b'(x)$.

We will define a subclass of $\mathcal{C}_{\text{BEC}}(T)$ for each possible assignment of b-values to the vertices in X in the subproblems on A and B. (The number of these subclasses, since $|X| = O(1)$ and b is bounded, is still $O(1)$.)

Formally: Define functions β and $\overline{\beta}$ as any assignments of b-values, in the subproblems on A and B respectively, to all vertices $x \in X$, such that the β and $\overline{\beta}$ values sum to $b'(x)$ for each x.

There are many degrees of freedom in defining β. Consider each possible choice of β and $\overline{\beta}$.

Define the subclass $\mathcal{C}_{\text{BEC}}(T_\beta)$ as the set of all b-edge covers that consist of a β-edge cover in A and a $\overline{\beta}$-edge cover in B.

That is, in class $\mathcal{C}_{\text{BEC}}(T_\beta)$, for each x, the number of incident edges selected in A is at least $\beta(x)$, and the number of incident edges in B is at least $\overline{\beta}(x)$.

Each of these subclasses $\mathcal{C}_{\text{BEC}}(T_\beta)$ is a Cartesian product of b-edge cover Glauber graphs, over subgraphs A and B of G, and thus internally has a good flow f_β; thus it suffices to combine flows within these subclasses together to design a flow f_T in $\mathcal{C}_{\text{BEC}}(T)$. We can then apply the hierarchical framework to obtain the desired flow in $\mathcal{M}_{\text{BEC}}(G)$.

Lemma 3.20. *Given a graph G and corresponding Glauber graph $\mathcal{M}_{\text{BEC}}(G)$, each class $\mathcal{C}_{\text{BEC}}(T)$ of b-edge covers in $\mathcal{M}_{\text{BEC}}(G)$ satisfies the conditions of Lemma 3.19.*

Proof. The number of subclasses is clearly $O(1)$. The subclasses are also all within an $O(1)$ size factor of one another. To see this, compare $|\mathcal{C}_{\text{BEC}}(T)|$ and $|\mathcal{C}_{\text{BEC}}(T_\beta)|$, for any β. Fix some lexicographic ordering of the edges of G. For every b-edge cover $S \in \mathcal{C}_{\text{BEC}}(T)$, there exists a b-edge cover $S' \in \mathcal{C}_{\text{BEC}}(T_\beta)$ that includes the lexicographically first $\beta(x)$ edges in A incident to x, for each $x \in X$, and also includes the first $\overline{\beta}(x)$ edges in B incident to x. (Let S' agree with S on all other edges.) Clearly this is a 2^b-to-1 mapping, i.e. an $O(1)$-to-1 mapping.

Finally, every pair of subclasses overlaps in at least $\Omega(1)|\mathcal{C}_{\text{BEC}}(T)|$ vertices: consider the set of all b-edge covers in $\mathcal{C}_{\text{BEC}}(T)$ in which for each $x \in X$, x has $\min\{b'(x), \delta_A(x)\}$ incident edges selected in A, and $\min\{b'(x), \delta_B(x)\}$ incident edges selected in B. The number of such b-edge covers is $\Omega(1)|\mathcal{C}_{\text{BEC}}(T)|$, by similar reasoning to the above; furthermore, every pair of subclasses of $\mathcal{C}_{\text{BEC}}(T)$ clearly contains this set of b-edge covers. The lemma follows. $\square$

The rest of the hierarchical framework conditions are easy to verify, and thus the result in Theorem 1.7 for the unbiased case of b-edge covers follows. The specific bound follows from the following observations: first, constructing the flow within a class $\mathcal{C}_{\text{BEC}}(T)$ incurs a factor of

$$2L|\mathcal{C}_{\text{BEC}}(T)|/(\nu(\min_{\beta} |\mathcal{C}_{\text{BEC}}(T_\beta)|)),$$

where ν is the minimum fraction of vertices shared by a pair of adjacent classes whose intersection is used in the flow, L is the maximum number of subclasses that a vertex can belong to, and $\min_\beta |\mathcal{C}_{\text{BEC}}(T_\beta)|$ is the smallest subclass of $\mathcal{C}_{\text{BEC}}(T)$. L is at most the number of subclasses, which is upper-bounded by $(b+1)^{t+1}$; the smallest subclass has size at least $|\mathcal{C}_{\text{BEC}}(T)|/2^{b(t+1)}$; and $\nu \geq 1/2$. (The latter two facts follow readily from observations made in the proof of Lemma 3.20.)

Thus the construction of the flow within $\mathcal{C}_{\mathrm{BEC}}(T)$ incurs a factor of at most $4(b+1)^{t+1} \cdot 2^{b(t+1)}$. In the biased chain the overlap between adjacent subclasses is at least $\lambda/(1+\lambda)$ instead of $1/2$, and we need to adjust the ratio $|\mathcal{C}_{\mathrm{BEC}}(T)|/|\mathcal{C}_{\mathrm{BEC}}(T_\beta)|$ by a factor of $\hat{\lambda}^{b(t+1)}$. Therefore this expression becomes

$$2\frac{1+\lambda}{\lambda}(b+1)^{t+1} \cdot (2\hat{\lambda})^{b(t+1)}.$$

The rest of the inductive step is as in the case of independent sets, i.e. we apply Lemma 3.14, using $K \leq 2^{t(t+1)/2}$, $\Delta_{\mathcal{M}} \leq m$, and $N \leq 2^m$. (In the weighted case, $\pi^*_{\min} \leq (2\hat{\lambda})^m$. When considering K in the application of Lemma 3.14, we need to weight K by a factor of $\hat{\lambda}^{t(t+1)/2}$.

Thus we obtain an additional factor of

$$2(K\hat{\lambda}^{t(t+1)/2} + 1) \leq 2((2\hat{\lambda})^{t(t+1)/2+1}).$$

Altogether, these two flow constructions combined, in each iteration, result in a factor of at most

$$4(\frac{1+\lambda}{\lambda})(b+1)^{t+1} \cdot (2\hat{\lambda})^{(t+1)(b+t/2)+1}.$$

The resulting mixing time is therefore at most

$$O(((1+\hat{\lambda})\hat{\lambda})^2 m^3 (\log\hat{\lambda}+1) n^{2(2+\log(1+\lambda)-\log(\lambda)+(t+1)\log(b+1)+((t+1)(b+t/2)+1)(1+\log\hat{\lambda}))}).$$

We have ignored one detail: technically the number of levels of induction is greater than $\log n$, because the $t+1$ vertices in the separator are included in the independent subproblems within each subclass. Furthermore, we cannot assume that we have two connected components of size at most $|V(G)|/2$ at each level of decomposition, so the base of the log is $3/2$ and not 2.

However, for every $\varepsilon < 1/2$, we have for all $n \geq (t+1)/\varepsilon$ that $2n/3 + t + 1 \leq n(2/3 + \varepsilon)$.

Thus at the cost of a base case for the induction of $(2\hat{\lambda})^{((t+1)/\varepsilon)^2}$, we adjust the $\log n$ exponent in the congestion term to $\log_{1/(2/3+\varepsilon)} n$. Letting $\varepsilon = 1/6$, we obtain the mixing bound claimed in Theorem 1.7, namely

$$O\left(((2\hat{\lambda})^{36(t+1)^2})^2 \left(\frac{\hat{\lambda}}{(1+\hat{\lambda})}\right)^2 m^3(\log\hat{\lambda}+1)n^{\frac{2(3+\log(1+\lambda)-\log(\lambda)+(t+1)\log(b+1)+((t+1)(b+t/2)+1)(1+\log\hat{\lambda}))}{\log(6/5)}}\right).$$

3.7.3 Dominating sets in the relaxed hierarchical framework

To finish the proof of Theorem 1.7 in the unbiased case, we now deal with dominating sets.

As with b-edge covers, defining classes in the same way as in the case of independent sets does not result in Cartesian products of dominating set Glauber graphs, because it may be that some vertices in X are not dominated by vertices in $T \cap X$; these vertices must then be dominated by vertices in A or in B.

Therefore, to preserve the recursive structure of the problem and thus complete the proof of Theorem 1.7, we define the *constrained Steiner dominating set problem* as a generalization of the dominating set problem, in which there are three types of vertices:

1. "normal" vertices, which must be dominated and may be selected in a dominating set,
2. "Steiner" vertices, which need not be dominated and may be selected, and
3. "forbidden" vertices, which must be dominated and must not be selected.

Now, we let X be a balanced vertex separator in G as before. We would like to define each class $\mathcal{C}_{\text{DOM}}(T)$ by a choice of vertices in X. In the resulting subproblem in A (similarly B), we then designate each vertex $v \in A \cup B$ having a neighbor selected in $T \subseteq X$ as

Steiner. However, there may be vertices in X that are not selected or dominated by any vertex in T. To obtain a valid dominating set, some neighbor of each such vertex w must be chosen in either A or B. Thus we have non-independent subproblems, which ruins the Cartesian product structure needed for the divide-and-conquer argument. To resolve this non-independence, we divide $\mathcal{C}_{\text{DOM}}(T)$ into subclasses as follows:

Given a graph G, separator X, and CSDS Glauber graph $\mathcal{M}_{\text{DOM}}(G)$, and class $\mathcal{C}_{\text{DOM}}(T)$ of CSDS's in $\mathcal{M}_{\text{DOM}}(G)$, let U be the set of undominated vertices in X in class $\mathcal{C}_{\text{DOM}}(T)$. For each subset $W \subseteq U$, let the subclass $\mathcal{C}_{\text{DOM}}(T_W)$ be the set of all CSDS's that consist of a union of a CSDS on $A \cup W$, and a CSDS on $B \cup U \setminus W$—in which each $w \in W$ is a forbidden vertex in the A subproblem, and each $\overline{w} \in U \setminus W$ is a forbidden vertex in the B subproblem.

There are at most $2^t = O(1)$ such subclasses. Technically, as with b-edge covers, these are not equivalence classes, as some CSDS solutions may belong to multiple classes. We will address this shortly, but first we specify how to route flow among the subclasses within $\mathcal{C}_{\text{DOM}}(T)$. Once we have specified this flow, we can simply apply the flow described in the proof of Lemma 3.11 to route flow among the "main" classes.

Lemma 3.21. *Given a graph G, corresponding CSDS Glauber graph $\mathcal{M}_{\text{DOM}}(G)$, and a class $\mathcal{C}_{\text{DOM}}(T)$ of CSDS's in $\mathcal{M}_{\text{DOM}}(G)$, the partition into subclasses given for dominating sets satisfies the conditions of Lemma 3.19.*

Proof. Clearly there are $O(1)$ classes. We observe that the subclasses are all within an $O(1)$ size factor of one another. To see this, compare the sizes of $\mathcal{C}_{\text{DOM}}(T)$ and $\mathcal{C}_{\text{DOM}}(T_W)$. Since $\mathcal{C}_{\text{DOM}}(T_W) \subseteq \mathcal{C}_{\text{DOM}}(T)$, $|\mathcal{C}_{\text{DOM}}(T_W)| \leq \mathcal{C}_{\text{DOM}}(T)$. On the other hand, consider the mapping that sends every CSDS $s \in \mathcal{C}_{\text{DOM}}(T)$ to a CSDS $s' \in \mathcal{C}_{\text{DOM}}(T_W)$ in which at least one neighbor (say, the first lexicographically) of each $w \in W$ is selected in A, and in which the first neighbor of each $\overline{w} \in U \setminus W$ is selected in B. This mapping is clearly $2^{|U|}$-to-1 $= O(1)$-to-1, and thus the size factor difference is $O(1)$.

Now, since the subclasses are not equivalence classes, many pairs of subclasses overlap. In particular, let T and U be as before, and suppose for some $u \in U$, $W' = W \cup \{u\}$. Then $\mathcal{C}_{\mathrm{DOM}}(T_W)$ and $\mathcal{C}_{\mathrm{DOM}}(T_{W'})$ overlap at those CSDS's in which some neighbor of u in A is selected, and some neighbor of u in B is selected. For every such pair of subclasses $\mathcal{C}_{\mathrm{DOM}}(T_W)$ and $\mathcal{C}_{\mathrm{DOM}}(T_{W'})$, at least half of the CSDS's in $\mathcal{C}_{\mathrm{DOM}}(T_W)$ and at least half of those in $\mathcal{C}_{\mathrm{DOM}}(T_{W'})$ are part of the overlap.

It is clear that $\mathcal{C}_{\mathrm{DOM}}(T)$ is internally connected via these overlaps: every CSDS in $\mathcal{C}_{\mathrm{DOM}}(T)$ has a path to the trivial CSDS in which every non-forbidden vertex of $A \cup B$ is selected. Thus the conditions of Lemma 3.19 are satisfied. □

As in the discussion following Lemma 3.20, we derive the bound in Theorem 1.7 as follows: the flow within a class incurs a congestion factor of

$$2\frac{1+\lambda}{\lambda} \cdot (2\hat{\lambda})^{t+1} |\mathcal{C}_{\mathrm{DOM}}(T)| / (\min_W |\mathcal{C}_{\mathrm{DOM}}(T_W)|) \leq 2\frac{1+\lambda}{\lambda}(4\hat{\lambda})^{t+1}.$$

The application of Lemma 3.14 contributes a $2(K\hat{\lambda}^{t+1} + 1)$ factor to the inductive step, with $K = 2^{t+1}$.

Thus the factor for one iteration is at most

$$4\frac{1+\lambda}{\lambda}(8\hat{\lambda}^2)^{t+2}.$$

The inclusion of the $t+1$ separator vertices in the subproblems, as with b-edge covers, increases the induction depth, and an analogous analysis gives a base case of $(2\hat{\lambda})^{6(t+1)}$ using $\varepsilon = 1/6$. Putting all this together with the fact that $\Delta_{\mathcal{M}} \leq n$ and $1/\pi^*_{\min} \leq (2\hat{\lambda})^n$ gives a mixing bound of

$$
\begin{aligned}
&O(((2\hat{\lambda})^{6(t+1)})^2(\frac{\hat{\lambda}}{1+\hat{\lambda}})^2(1+\log\hat{\lambda})\cdot n^{2\log(4\frac{1+\lambda}{\lambda}(8\hat{\lambda}^2)^{t+2})/\log(6/5)+3})\\
&=O(((2\hat{\lambda})^{12(t+1)})(\frac{\hat{\lambda}}{1+\hat{\lambda}})^2(1+\log\hat{\lambda})\cdot n^{2(2+\log(\frac{1+\lambda}{\lambda})+(t+2)(3+\log\hat{\lambda}))/\log(6/5)+3}).
\end{aligned}
$$

3.7.4 Rapid mixing in the relaxed hierarchical framework for all $\lambda > 0$

We now generalize Lemma 3.19 to all $\lambda > 0$, finishing the proof of Theorems 1.6 and 1.7.

Lemma 3.22. *For the Glauber graphs $\mathcal{M}_{\text{BEC}}(G)$ and $\mathcal{M}_{\text{DOM}}(G)$, with classes defined as in Lemma 3.19, and with stationary distribution π induced by parameter λ as in the discussion in Section 3.6.1, the flow construction in Lemma 3.19 results in a congestion factor gain of at most $\rho = O(1)$ at each of the $O(\log n)$ levels of induction, resulting in at most polynomial overall congestion.*

Proof. We need to show that the flow construction within a class $\mathcal{C}(T)$ in Lemma 3.19 produces at most an $O(1)$-factor increase in congestion; the rest of the argument is similar to the proof of Lemma 3.18. For the case of dominating sets, consider a pair of $\mathcal{C}_{\text{DOM}}(T_W) \subseteq \mathcal{C}_{\text{DOM}}(T)$ and $\mathcal{C}_{\text{DOM}}(T_{W'}) \subseteq \mathcal{C}_{\text{DOM}}(T)$. For every such pair, consider the intersection $\mathcal{I}$ of the two subclasses, namely the set of dominating sets in which for every input graph vertex $v \in W \cup W'$, some neighbor of v is selected in A, and for every vertex $w \in U \setminus (W \cap W')$, some neighbor of w is selected in B. There exists an $O(1)$-to-1 mapping from $\mathcal{C}_{\text{DOM}}(T_W)$ to $\mathcal{I}$—found by adding $O(1)$ neighbors of vertices in U as described above to each dominating set $S \in \mathcal{C}_{\text{DOM}}(T_W)$—under which the image S' of S has $|S' \setminus S| = O(1)$, and therefore $\pi(S') = \lambda^{|S' \setminus S|}\pi(S) = \Theta(1)\pi(S)$. This shows that $\pi(\mathcal{I}) = \Theta(\pi(\mathcal{C}_{\text{DOM}}(T_W))) = \Theta(\pi(\mathcal{C}_{\text{DOM}}(T_{W'})))$.

Thus we use the overlaps between classes to route flow along a path of classes as in the proof of Lemma 3.19. As before, at each class in the path, the internal routing produces an $O(1)$ factor increase in the congestion within the class. The concern, again, is that due to overlap, there may be edges belonging to multiple classes that thus incur congestion multiple times in the routing of the flow; as before, this is not a problem as there are $O(1)$ pairs of classes for which this occurs.

The argument for $\mathcal{M}_{\text{BEC}}(G)$ is similar, with the intersection $\mathcal{I}$ found by selecting sufficiently many edges incident to each vertex $x \in X$ to satisfy membership in both subclasses $\mathcal{C}_{\text{BEC}}(T_\beta)$ and $\mathcal{C}_{\text{BEC}}(T_{\beta'})$. $\square$

3.7.5 Rapid mixing of the Glauber dynamics on b-matchings for all $\lambda > 0$

For the claim about b-matchings in Theorem 1.7, we do not need the relaxed framework; in fact it suffices to combine Lemma 3.13 with the following lemma:

Lemma 3.23. *One can adapt the proof of Lemma 3.18 to the hierarchical framework, proving the claim about the Glauber dynamics on b-matchings in Theorem 1.7 for all* $\lambda > 0$.

Proof. The proof of Lemma 3.18 uses a simple mapping argument to show that for every ancestor $\mathcal{C}(T_i)$ of a class $\mathcal{C}(T)$, $\pi(\mathcal{C}(T_i)) = \Theta(1)\pi(\mathcal{C}(T))$, then allows each boundary edge e between classes to carry $O(1)Q(e)$ units of flow across the boundary, by ensuring that each boundary vertex S carries flow in proportion to its weight $\pi(S)$. Since all pairs of classes $\mathcal{C}(T), \mathcal{C}(T')$ have a common ancestor in the case of b-matchings, we in fact have $\pi(\mathcal{C}(T)) = \Theta(1)\pi(\mathcal{C}(T'))$ for every pair of classes $\mathcal{C}(T), \mathcal{C}(T')$. The bound on flow across the boundary therefore still holds; the argument for bounding congeston factor increase wthin a class is the same as in the proof of Lemma 3.18. $\square$

For the specific mixing upper bound for b-matchings, we use Lemma 3.14, and observe that $K \leq 2^{\Delta(t+1)}$; the parameter λ contributes at most a $\hat{\lambda}^{\Delta(t+1)}$ factor, $1/\pi^*_{\min} \leq (2\hat{\lambda})^m$, and $\Delta_{\mathcal{M}} \leq m$, so we have a mixing upper bound of

$$\begin{aligned} &O(((1+\hat{\lambda})\hat{\lambda})^2(2((2\hat{\lambda})^{\Delta(t+1)}+1))^{2\log n}m^3(1+\log\hat{\lambda})) \\ &\quad = O(((1+\hat{\lambda})\hat{\lambda})^2(1+\log\hat{\lambda})m^3 n^{2\Delta(t+2)(1+\log\hat{\lambda})+2}). \end{aligned}$$

3.7.6 Maximal independent sets and maximal b-matchings in the non-hierarchical framework

Dealing with internally disconnected classes

As noted in Section 3.7, we use the non-hierarchical framework and assume bounded carving width for the chains on maximal independent sets and maximal b-matchings. Once we have defined these chains, we will see that partitioning the Glauber graph for each chain in the natural way will result in classes that are not necessarily internally connected. The solution will be to relax the framework conditions so that the classes need not be disjoint—but then require that every pair of overlapping classes must overlap in a large number of vertices. More precisely:

Lemma 3.24. *Suppose a Glauber graph $\mathcal{M}(G)$ satisfies the conditions of the non-hierarchical framework in Section 3.2.3, except that:*

1. *The $O(1)$ classes are not necessarily disjoint.*
2. *Each pair of classes $\mathcal{C}(T)$ and $\mathcal{C}(T')$ sharing at least one vertex shares $\Theta(1)|\mathcal{C}(T)| = \Theta(1)|\mathcal{C}(T')|$ vertices.*

Then the expansion of $\mathcal{M}(G)$ is $\Omega(1/n^c)$, where $c = O(1)$.

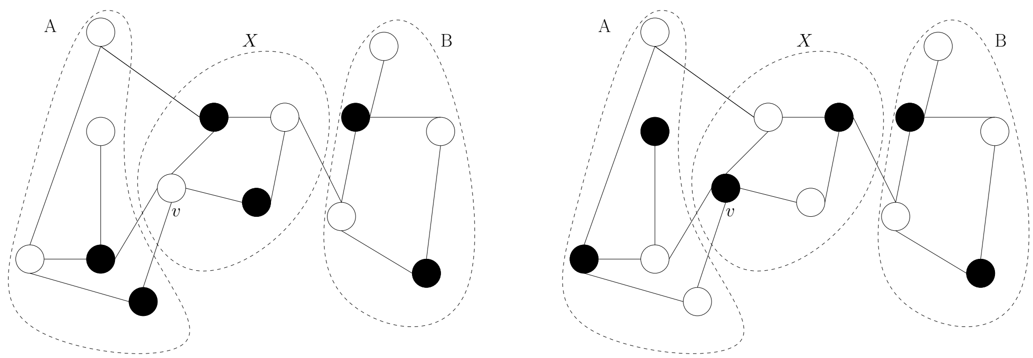

Figure 3.5: Two maximal independent sets in a graph G: S (left) and S' (right). S and S' differ by a flip, with the separator X inducing the classes to which the sets belong. S' results from adding v to S, removing the neighbors of v, and adding some of the neighbors of the removed vertices.

Proof. The multicommodity flow construction is as in the proof of Lemma 3.9, except that when sending flow from $S \in \mathcal{C}(T)$ to $S' \in \mathcal{C}(T') \neq \mathcal{C}(T)$ via a path through intermediate classes, we now have some pairs of intermediate classes that share boundary vertices, instead of sharing boundary edges. The flow is the same as before, except that there is no need to send flow across a boundary in these cases.

The congestion analysis is the same as in the proof of Lemma 3.19. $\square$

Maximal independent sets

We now apply the non-hierarchical framework to the flip chain on maximal independent sets. We first define the maximal independent set mixing problem, then show that it meets the criteria of the framework, up to the conditions in Lemma 3.24.

We return to the Glauber graph defined for maximal independent sets.

Lemma 3.25. *The maximum degree of $\mathcal{M}_{\text{MIS}}(G)$ is at most $n \cdot 2^{\Delta^2+\Delta}$, where $n = |V(G)|$ and Δ is the maximum degree of any vertex in G.*

Proof. The number of neighbors of a given maximal independent set S is the same as the number of ways to choose a vertex v to add to or drop from S, along with a subset of the vertices at distance at most two from v to add or drop. □

Define the *maximal independent set flip chain* on a graph G with Glauber graph $\mathcal{M}_{\text{MIS}}(G)$ as the following Markov chain (let Δ_M be the maximum degree of $\mathcal{M}_{\text{MIS}}(G)$):

1. Let X_0 be an arbitrary maximal independent set in $V(\mathcal{M}_{\text{MIS}}(G))$.

2. For $t \geq 0$, define X_{t+1} as follows:

   With probability $(1/2)(\delta(X_t)/\Delta_M)$, let X_{t+1} be a neighbor in $\mathcal{M}_{\text{MIS}}(G)$ of X_t, selected uniformly at random from the neighbors of X_t.

   With probability $1 - (1/2)(\delta(X_t)/\Delta_M)$ let $X_{t+1} = X_t$.

For technical reasons, the following observation is necessary for obtaining a rapid mixing bound from an expansion bound on $\mathcal{M}_{\text{MIS}}(G)$.

Remark 3.2. *This is the standard Markov chain on $\mathcal{M}_{\text{MIS}}(G)$, with appropriate self loops added in the standard way. Furthermore, by Lemma 3.25, if G has bounded degree, then the degree-based weighting does not cause the spectral expansion of the chain to differ by more than a polynomial factor from the edge (or vertex) expansion of $\mathcal{M}_{\text{MIS}}(G)$.*

Verification of conditions for maximal independent sets

We show how to apply the non-hierarchical version of the framework when the carving width of G is bounded. First, to satisfy Condition 1, we would *like* to use a partition analogous to that defined for independent sets: each class $\mathcal{C}_{\text{MIS}}(T)$ is the set of maximal independent sets that agree on their restriction to the vertex separator X for G. However, a subtlety arises when considering the Cartesian product structure of the Glauber graphs on A and B within a class T: in the independent set Glauber graph, $\mathcal{C}_{\text{IS}}(T)$ was a Cartesian product of two independent set Glauber graphs $\mathcal{M}_{\text{MIS}}(G_A)$ and $\mathcal{M}_{\text{MIS}}(G_B)$, respectively defined on the independent sets in $A \setminus N_A(T)$ and on those in $B \setminus N_B(T)$. Here, however, the independent sets chosen in A and in B need to give a *maximal* independent set when their union is taken with the set chosen in X.

It may be that the independent set in X identified with $\mathcal{C}_{\text{MIS}}(T)$ is not maximal. For a simple example, suppose X is a path of length three, consisting of vertices u, v, and w and edges (u, v) and (v, w), with u having neighbors only in A, w having neighbors only in B, and v having only u and w as neighbors. Suppose $\mathcal{C}_{\text{MIS}}(T)$ is identified with the independent set $\{u\} \subseteq X$. Then every maximal independent set $S \in T$ has some neighbor of w in B chosen, or else S would not be maximal. Furthermore, one can show that defining classes in this way would result in internally disconnected classes. Thus we cannot simply eliminate vertices in $N_A(T)$ from A and $N_B(T)$ from B and define smaller maximal independent set Glauber graphs. Instead, we define the classes—which, per the conditions of Lemma 3.24, need not be disjoint—as follows:

Given a graph G with maximal independent set Glauber graph $\mathcal{M}_{\text{MIS}}(G)$ and a class $\mathcal{C}_{\text{MIS}}(T)$, let $U \subseteq X$ be the set of all unselected vertices in T that have no neighbor selected in T. For each independent subset of the vertices in $N_A(U) \cup N_B(U)$ that *covers* all of U—that is, for each independent subset $C \subseteq N_A(U) \cup N_B(U)$ such that every $x \in U$ has some neighbor $z \in C$, let $\mathcal{C}_{\text{MIS}}(T_C)$ be the class of all independent sets in G that agree with T on X, and that include all of the vertices in C.

The smaller Glauber graphs on A and B are now independent for a given class $\mathcal{C}_{\text{MIS}}(T_C)$: for each $z \in C$, remove z and all neighbors in $N_A(z) \cup N_B(z)$ from the graph, and consider the resulting maximal independent set Glauber graphs on $A \setminus (C \cup N_A(C))$ and on $B \setminus (C \cup N_B(C))$. Each class $\mathcal{C}_{\text{MIS}}(T_C)$ is a Cartesian product of two such graphs. It suffices to show that this definition obeys the conditions of Lemma 3.24:

Lemma 3.26. *Given a graph G with bounded carving width and corresponding maximal independent set Glauber graph $\mathcal{M}_{\text{MIS}}(G)$, the definition of classes given satisfies the conditions of Lemma 3.24.*

Proof. The Cartesian product structure of $\mathcal{C}_{\text{MIS}}(T_C)$ and the fact that X is a balanced separator satisfy Conditions 5 and 6 of the non-hierarchical framework.

The classes do not partition $\mathcal{C}_{\text{MIS}}(T)$. However, clearly there are $O(1)$ classes. The classes are also within an $O(1)$ size factor of one another. To see this, define the following mapping f from the set of all maximal independent sets in $\mathcal{M}_{\text{MIS}}(G)$ to the set of maximal independent sets in a class $\mathcal{C}_{\text{MIS}}(T_C)$. For each maximal independent set $S \in V(\mathcal{M}_{\text{MIS}}(G))$, let $S' = f(S) \in \mathcal{C}_{\text{MIS}}(T_C)$ be the following maximal independent set: (i) let S' agree with T on all vertices in X; (ii) let S' agree with T_C on all vertices in $N_A(T) \cup N_B(T)$; (iii) let S' agree with S on all vertices not in $X \cup N_A(T) \cup N_B(T)$ and having no neighbor in $X \cup N_A(T) \cup N_B(T)$; (iv) add vertices to S, if needed, to obtain maximality.

The symmetric difference $f(S) \oplus S$ is of size at most $\Delta^2|X| \leq \Delta^2(t+1) = O(1)$, where t is the (bounded) treewidth of G and Δ is the (bounded) degree; thus f is an $O(1)$-to-1 mapping. By similar reasoning, the number of shared maximal independent sets between any two overlapping classes $\mathcal{C}_{\text{MIS}}(T_C)$ and $\mathcal{C}_{\text{MIS}}(T_{C'})$ is at least $\Omega(1)|\mathcal{C}_{\text{MIS}}(T_C)|$, and the number of edges between any two adjacent classes $\mathcal{C}_{\text{MIS}}(T_C)$ and $\mathcal{C}_{\text{MIS}}(T'_{C'})$ is at least $\Omega(1)|\mathcal{C}_{\text{MIS}}(T_C)|$. The lemma follows.

□

Maximal b-matchings in the non-hierarchical framework

We return to the maximal b-matching Glauber graph defined in Section 3.1.

The argument that the graph is connected is similar to the proof of Lemma 3.2.

It suffices to define a partition and verify the conditions. We would *like* to identify each class $\mathcal{C}_{\text{MBM}}(T)$ with the chosen subset of the edges that have at least one endpoint in the small balanced separator X. However, as with maximal independent sets, the maximality requirement introduces non-independent subproblems. To modify the definition of the classes, we first need to introduce the notion of a *saturated* vertex:

Given a b-matching in a graph G, consider a vertex v *saturated* if $b(v)$ edges incident to v are selected in the matching.

It may be that a vertex $v \in X$ is not saturated in a maximal b-matching, and thus the choice of edges inducing $\mathcal{C}_{\text{MBM}}(T)$ does not saturate v. In this case, we have a constraint on the subproblems in A and B. Namely, it must be that some neighbor of v, $u \in N(X)$, is saturated, or else the edge (u, v) could be added to the matching. We use this fact to define the subclasses of a class $\mathcal{C}_{\text{MBM}}(T)$:

Given a graph G with separator X, maximal b-matching Glauber graph $\mathcal{M}_{\text{MBM}}(G)$, and a set T of edges selected whose endpoints all lie in X, let $U \subseteq X$ be the set of unsaturated vertices in X induced by T; let C be a minimal set of edges such that, after adding C to T, some neighbor of v is saturated for every $v \in U$. Define the class $\mathcal{C}_{\text{MBM}}(T_C)$ as the set of all maximal b-matchings in T that contain all edges in C.

As in Section 3.7.6, these classes are not equivalence classes, because they overlap. Again, however, each b-matching belongs to at most $O(1)$ subclasses, and thus this overlap does not interfere with the proof.

We now verify that this definition of classes satisfies Lemma 3.24:

Lemma 3.27. *Given a graph G with bounded carving width, corresponding maximal b-matching Glauber graph $\mathcal{M}_{\text{MBM}}(G)$, the division into classes satisfies the conditions of*

Lemma 3.24.

Proof. The argument is similar to the proof of Lemma 3.26: again we have a Cartesian product structure in each class $\mathcal{C}_{\mathrm{MBM}}(T_C)$—where the resulting maximal b-matching subproblems on A and B result from (i) removing each edge $(u, w) \in C$ from $G[A]$ and $G[B]$, and (ii) decreasing $b(u)$ and $b(w)$ accordingly.

Clearly the number of classes is $O(1)$, due to the bounded carving width of G. The classes differ by an $O(1)$ size factor, and the overlaps are large; the argument, along with the resulting flow, is similar to that in the proof of Lemma 3.26. □

Specific mixing upper bounds for maximal independent sets and maximal b-matchings

For the derivation of the specific bounds stated in Theorem 1.8, we apply Lemma 2.10, with the modification that the $\mathcal{E}_{\min}$ term must be replaced by the term $\min\{\mathcal{E}_{\min}, \mathcal{O}_{\min}\}$, where $\mathcal{O}_{\min}$ is the size of the smallest overlap between a pair of classes that share at least one vertex. For maximal independent sets, $\mathcal{E}_{\min} \geq N/2^{7\Delta^6(t+1)}$ and $\mathcal{O}_{\min} \geq N/2^{3\Delta^2(t+1)}$, so $\min\{\mathcal{E}_{\min}, \mathcal{O}_{\min}\} \geq N/2^{7\Delta^6(t+1)}$. We also gain at each level of induction an additional factor of $K \leq 2^{(\Delta+1)(t+1)}$ due to overlaps. Combining this with the fact that $\Delta_{\mathcal{M}} \leq 2^{3\Delta^2}n$ and $N \leq 2^n$ gives a total mixing bound of

$$
\begin{aligned}
&O((2 \cdot 2^{7\Delta^6(t+1)} \cdot 2^{(\Delta+1)(t+1)})^{2\log n} \cdot 2^{6\Delta^2} n^3) \\
&\qquad = O(2^{6\Delta^2} n^{2(t+1)(7\Delta^6+\Delta+1)+5})
\end{aligned}
$$

as claimed. (The $\log(3/2)$ term in the theorem statement comes from the fact that the base of the log in the induction is $3/2$.)

A similar argument for maximal b-matchings gives the result claimed in Theorem 1.8, with

$\Delta_{\mathcal{M}} \leq 2^{6\Delta^2} m$, $N \leq 2^m$, $K \leq 2^{3\Delta^2(t+1)}$, $\mathcal{E}_{\min} \geq N/2^{8\Delta^7(t+1)}$, and $\mathcal{O}_{\min} \geq N/2^{4\Delta^3(t+1)}$.

I.e., we have

$$\begin{aligned} &O((2 \cdot 2^{8\Delta^7(t+1)} \cdot 2^{3\Delta^2(t+1)})^{2\log n} \cdot 2^{12\Delta^2} m^3) \\ &\qquad = O(2^{12\Delta^2} m^3 n^{2(t+1)(8\Delta^7+3\Delta^2)+2}). \end{aligned}$$

3.8 Open Questions

We have developed a framework whose application shows rapid mixing for several natural flip chains on graph-theoretic structures in graphs of bounded treewidth. However, some work is required in showing that each of the structures satisfies the conditions of the framework. We hope that a more robust version of the framework can be developed that further unifies these techniques.

In particular, all of the structures we have analyzed satisfy the conditions of Courcelle's theorem, as noted previously. It would be interesting to determine whether the framework can be extended to all structures satisfying these conditions.

The fact that our results hold for all values of $\lambda > 0$ is not especially surprising, as Ioannis Panageas has observed, since the limiting case $\lambda = \infty$ corresponds to the optimization version of each problem, and the case $\lambda = 1$ corresponds to uniform sampling; as stated in the introduction, both of these problems are already known to be fixed-parameter tractable in treewidth. (In fact, as we noted in the introduction, the extension of Courcelle's theorem, combined with the reduction from sampling to counting, applies to all values of $\lambda > 0$.) Nonetheless, our result does settle a missing case of the mixing question in some generality, through purely combinatorial methods.

In all of our mixing bounds, the dependency on the parameters—treewidth and degree—is bad. It would be interesting to see whether some refinement of our methods could give a truly fixed-parameter tractable result, in which the treewidth and degree do not appear in the exponent of n.

Bibliography

[1] An upper bound for the number of planar lattice triangulations. *Journal of Combinatorial Theory, Series A*, 103(2):383 – 386, 2003. doi:10.1016/S0097-3165(03)00097-9.

[2] Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1319–1330, 2020. doi:10.1109/FOCS46700.2020.00125.

[3] Nima Anari, Kuikui Liu, Shayan Oveis Gharan, and Cynthia Vinzant. Log-concave polynomials II: High-dimensional walks and an FPRAS for counting bases of a matroid. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing (STOC 2019)*, New York, NY, USA, 2019. Association for Computing Machinery. doi:10.1145/3313276.3316385.

[4] Nima Anari, Shayan Oveis Gharan, and Cynthia Vinzant. Log-concave polynomials, entropy, and a deterministic approximation algorithm for counting bases of matroids. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 35–46, 2018. doi:10.1109/FOCS.2018.00013.

[5] Pinaki Banerjee. Scattering amplitudes from positive geometries. In Prafulla Kumar Behera, Vipin Bhatnagar, Prashant Shukla, and Rahul Sinha, editors, *XXIII DAE High Energy Physics Symposium*, pages 421–427, Singapore, 2021. Springer Singapore.

[6] Noam Berger, Claire Kenyon, Elchanan Mossel, and Yuval Peres. Glauber dynamics on trees and hyperbolic graphs. *Probability Theory and Related Fields*, 131:311 – 340, 2005. doi:s00440-004-0369-4.

[7] Ivona Bezáková, Andreas Galanis, Leslie Ann Goldberg, Heng Guo, and Daniel Štefankovič. Approximation via correlation decay when strong spatial mixing fails. *SIAM Journal on Computing*, 48(2):279–349, 2019.

[8] Ivona Bezáková and Wenbo Sun. Mixing of Markov chains for independent sets on chordal graphs with bounded separators. In *Computing and Combinatorics*, pages 664–676. Springer International Publishing, 2020.

[9] Hans L. Bodlaender. Fixed-parameter tractability of treewidth and pathwidth. volume 7370 of *Lecture Notes in Computer Science*, pages 196–227. Springer, 2012. doi:10.1007/978-3-642-30891-8_12.

[10] Magnus Bordewich, Catherine Greenhill, and Viresh Patel. Mixing of the glauber dynamics for the ferromagnetic potts model. *Random Structures & Algorithms*, 48(1):21–52, 2016. doi:10.1002/rsa.20569.

[11] Magnus Bordewich and Ross J. Kang. Subset Glauber dynamics on graphs, hypergraphs and matroids of bounded tree-width. *Electronic Journal of Combinatorics*, 21(4), 2014. doi:10.37236/4195.

[12] Pietro Caputo, Fabio Martinelli, Alistair Sinclair, and Alexandre Stauffer. Dynamics of lattice triangulations on thin rectangles. *Electronic Journal of Probability*, 21, 05 2015. doi:10.1214/16-EJP4321.

[13] Pietro Caputo, Fabio Martinelli, Alistair Sinclair, and Alexandre Stauffer. Random lattice triangulations: structure and algorithms. *Annals of Applied Probability*, 25:1650–1685, 2015.

[14] Alessandra Caraceni and Alexandre Stauffer. Polynomial mixing time of edge flips on quadrangulations. *Probability Theory and Related Fields*, 176(1):35–76, Feb 2020. doi:10.1007/s00440-019-00913-5.

[15] Yijia Chen and Jörg Flum. The parameterized complexity of maximality and minimality problems. *Annals of Pure and Applied Logic*, 151(1):22 – 61, 2008. doi:10.1016/j.apal.2007.09.003.

[16] Zongchen Chen, Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. Rapid mixing for colorings via spectral independence. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1548–1557. doi:10.1137/1.9781611976465.94.

[17] Zongchen Chen, Kuikui Liu, and Eric Vigoda. Rapid mixing of Glauber dynamics up to uniqueness via contraction, 2020. arXiv:2004.09083.

[18] Zongchen Chen, Kuikui Liu, and Eric Vigoda. Optimal mixing of glauber dynamics: Entropy factorization via high-dimensional expansion. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1537–1550, 2021.

[19] Karel Culik II and Derick Wood. A note on some tree similarity measures. *Information Processing Letters*, 15(1):39–42, 1982.

[20] Christopher M De Sa, Ce Zhang, Kunle Olukotun, and Christopher Ré. Rapidly Mixing Gibbs Sampling for a Class of Factor Graphs Using Hierarchy Width. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc., 2015.

[21] Persi Diaconis and Laurent Saloff-Coste. Comparison theorems for reversible markov chains. *The Annals of Applied Probability*, 3(3):696–730, 1993.

[22] Persi Diaconis and Daniel Stroock. Geometric Bounds for Eigenvalues of Markov Chains. *The Annals of Applied Probability*, 1(1):36 – 61, 1991. doi:10.1214/aoap/1177005980.

[23] Martin Dyer, Leslie Ann Goldberg, and Mark Jerrum. Matrix norms and rapid mixing for spin systems. *The Annals of Applied Probability*, 19(1):71–107, 2021/10/28/ 2009. Full publication date: Feb., 2009. URL: http://www.jstor.org/stable/30243572.

[24] David Eppstein. The effect of planarization on width. In *Graph Drawing and Network Visualization*, pages 560–572. Springer International Publishing, 2018.

[25] David Eppstein and Daniel Frishberg. Rapid mixing of the hardcore Glauber dynamics and other Markov chains in bounded-treewidth graphs. 2021. arXiv:2111.03898, doi:10.48550/arXiv.2111.03898.

[26] David Eppstein and Daniel Frishberg. Improved mixing for the convex polygon triangulation flip walk. 2022. URL: https://arxiv.org/abs/2207.09972, doi:10.48550/ARXIV.2207.09972.

[27] Jeff Erickson. Computational topology: Treewidth. Lecture Notes, 2009. URL: http://jeffe.cs.illinois.edu/teaching/comptop/2009/notes/treewidth.pdf.

[28] Weiming Feng, Heng Guo, Yitong Yin, and Chihao Zhang. Rapid mixing from spectral independence beyond the boolean domain. In *Proceedings of the Thirty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA 2021, page 1558–1577, USA, 2021. Society for Industrial and Applied Mathematics.

[29] S. M. Ferdous, A. Khan, and A. Pothen. Parallel algorithms through approximation: B-edge cover. In *2018 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, pages 22–33. IEEE Computer Society, May 2018. doi:10.1109/IPDPS.2018.00013.

[30] Q Ge and D Štefankovič. A graph polynomial for independent sets of bipartite graphs. *Combinatorics, Probability and Computing*, 21(5):695–714, 2012.

[31] F. Graham and P. Tetali. Isoperimetric inequalities for cartesian products of graphs. *Comb. Probab. Comput.*, 7:141–148, 1998.

[32] Venkatesan Guruswami. Rapidly mixing markov chains: A comparison of techniques (a survey). *ArXiv*, 2016. URL: https://arxiv.org/abs/1603.01512.

[33] Thomas P Hayes. A simple condition implying rapid mixing of single-site dynamics on spin systems. In *2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)*, pages 39–46. IEEE, 2006.

[34] Thomas P. Hayes, Juan C. Vera, and Eric Vigoda. Randomly coloring planar graphs with fewer colors than the maximum degree. *Random Structures & Algorithms*, 47(4):731–759, 2015. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/rsa.20560, doi:10.1002/rsa.20560.

[35] Marc Heinrich. Glauber dynamics for colourings of chordal graphs and graphs of bounded treewidth, 2020. URL: https://arxiv.org/abs/2010.16158, arXiv:2010.16158.

[36] Peter J. Hilton and Jean J. Pedersen. Catalan numbers, their generalization, and their uses. *The Mathematical Intelligencer*, 13:64–75, 1991.

[37] Lingxiao Huang, Pinyan Lu, and Chihao Zhang. Canonical paths for MCMC: from art to science. In *Proceedings of the twenty-seventh annual ACM-SIAM symposium on Discrete algorithms*, pages 514–527. SIAM, 2016.

[38] Mrunmay Jagadale and Alok Laddha. Towards positive geometry of multi scalar field amplitudes : Accordiohedron and effective field theory. 2021. URL: `https://arxiv.org/abs/2104.04915`.

[39] Mark Jerrum and Alistair Sinclair. Conductance and the rapid mixing property for Markov chains: The approximation of permanent resolved. In *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing*, STOC '88, 1988. `doi:10.1145/62212.62234`.

[40] Mark Jerrum, Jung-Bae Son, Prasad Tetali, and Eric Vigoda. Elementary bounds on Poincaré and log-Sobolev constants for decomposable Markov chains. *The Annals of Applied Probability*, 14(4):1741–1765, 2004. URL: `http://www.jstor.org/stable/4140446`.

[41] V. Kaibel and G. Ziegler. Counting lattice triangulations. *arXiv: Combinatorics*, 2002.

[42] Volker Kaibel. On the expansion of graphs of 0/1-polytopes. In *The Sharpest Cut: The Impact of Manfred Padberg and His Work*, pages 199–216. SIAM, 2004.

[43] Tali Kaufman and Izhar Oppenheim. High order random walks: Beyond spectral gap. *Combinatorica*, 40(2):245–281, 2020.

[44] Arif Khan and Alex Pothen. A new 3/2-approximation algorithm for the b-edge cover problem. In *2016 Proceedings of the Seventh SIAM Workshop on Combinatorial Scientific Computing*, pages 52–61. SIAM, 2016.

[45] Arif Khan, Alex Pothen, Md. Mostofa Ali Patwary, Nadathur Rajagopalan Satish, Narayanan Sundaram, Fredrik Manne, Mahantesh Halappanavar, and Pradeep Dubey. Efficient approximation algorithms for weighted b-matching. *SIAM Journal on Scientific Computing*, 38(5):S593–S619, 2016. `doi:10.1137/15M1026304`.

[46] David A. Klarner. Correspondences between plane trees and binary sequences. *Journal of Combinatorial Theory*, 9(4):401–411, 1970. URL: `https://www.sciencedirect.com/science/article/pii/S002198007080093X`, `doi:10.1016/S0021-9800(70)80093-X`.

[47] Vedat Levi Alev and Lap Chi Lau. Improved analysis of higher order random walks and applications. *arXiv e-prints*, pages arXiv–2001, 2020.

[48] David A Levin, Yuval Peres, and Elizabeth Wilmer. *Markov chains and mixing times*, volume 107. American Mathematical Soc., 2017.

[49] J. Loday. The multiple facets of the associahedron. In *Proc. 2005 Academy Coll. Series*, 2005.

[50] László Lovász and Ravi Kannan. Faster mixing via average conductance. In *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing*, STOC '99, page 282–287, New York, NY, USA, 1999. Association for Computing Machinery. doi:10.1145/301250.301317.

[51] Brendan Lucier, Michael Molloy, and Yuval Peres. The Glauber dynamics for colourings of bounded degree trees. In Irit Dinur, Klaus Jansen, Joseph Naor, and José Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 631–645, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

[52] Neal Madras and Dana Randall. Markov chain decomposition for convergence rate analysis. *The Annals of Applied Probability*, 12(2):581 – 606, 2002. doi:10.1214/aoap/1026915617.

[53] Fabio Martinelli, Alistair Sinclair, and Dror Weitz. Fast mixing for independent sets, colorings, and other models on trees. *Random Structures & Algorithms*, 31(2):134–172, 2007. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/rsa.20132, doi:10.1002/rsa.20132.

[54] Lisa McShine and P. Tetali. On the mixing time of the triangulation walk and other catalan structures. In *Randomization Methods in Algorithm Design*, 1997.

[55] Milena Mihail and Umesh Vazirani. On the expansion of 0-1 polytopes. *Journal of Combinatorial Theory, Series B*, 1989.

[56] Michael Molloy, Bruce Reed, and William Steiger. On the mixing rate of the triangulation walk. *Randomization Methods in Algorithm Design*, 1997.

[57] Ben Morris and Alistair Sinclair. Random walks on truncated cubes and sampling 0-1 knapsack solutions. *SIAM journal on computing*, 34(1):195–226, 2004.

[58] Atsuhiro Nakamoto, Gen Kawatani, Naoki Matsumoto, and Jorge Urrutia. Geometric quadrangulations of a polygon. *Electronic Notes in Discrete Mathematics*, 68:59–64, 2018. Discrete Mathematics Days 2018. doi:10.1016/j.endm.2018.06.011.

[59] Nestor V. Nestoridis and Dimitrios M. Thilikos. Square roots of minor closed graph classes. *Discrete Applied Mathematics*, 168:34–39, 2014. doi:10.1016/j.dam.2013.05.026.

[60] Yoshio Okamoto, Takeaki Uno, and Ryuhei Uehara. Linear-time counting algorithms for independent sets in chordal graphs. In *Graph-Theoretic Concepts in Computer Science*, pages 433–444. Springer Berlin Heidelberg, 2005.

[61] Reinhard Pichler, Stefan Rümmele, and Stefan Woltran. Counting and enumeration problems with bounded treewidth. In *Logic for Programming, Artificial Intelligence, and Reasoning*, pages 387–404. Springer Berlin Heidelberg, 2010.

[62] Prashanth Raman. The positive geometry for ϕp interactions. *Journal of High Energy Physics*, 2019(10):271, Oct 2019. doi:10.1007/JHEP10(2019)271.

[63] Dana Randall and Prasad Tetali. Analyzing glauber dynamics by comparison of markov chains. In *Latin American Symposium on Theoretical Informatics*, pages 292–304. Springer, 1998.

[64] Herbert E. Robbins. A remark on stirling's formula. *American Mathematical Monthly*, 62:402–405, 1955.

[65] Neil Robertson and P.D Seymour. Graph minors. X. obstructions to tree-decomposition. *Journal of Combinatorial Theory, Series B*, 52(2):153–190, 1991. doi:10.1016/0095-8956(91)90061-N.

[66] P. D. Seymour and R. Thomas. Call routing and the ratcatcher. *Combinatorica*, 14(2):217–241, Jun 1994. doi:10.1007/BF01215352.

[67] Alistair Sinclair. Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combinatorics, Probability and Computing*, 1(4):351–370, 1992. doi:10.1017/S0963548300000390.

[68] Alistair Sinclair and Mark Jerrum. Approximate counting, uniform generation and rapidly mixing Markov chains. *Information and Computation*, 82(1):93–133, 1989. doi:10.1016/0890-5401(89)90067-9.

[69] Daniel D Sleator, Robert E Tarjan, and William P Thurston. Rotation distance, triangulations, and hyperbolic geometry. *Journal of the American Mathematical Society*, 1(3):647–681, 1988.

[70] Allan Sly. Computational transition at the uniqueness threshold. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pages 287–296, 2010. doi:10.1109/FOCS.2010.34.

[71] Alexandre Stauffer. A Lyapunov function for Glauber dynamics on lattice triangulations. *Probability Theory and Related Fields*, 169:469–521, 2015.

[72] Prasad Tetali, Juan C. Vera, Eric Vigoda, and Linji Yang. Phase transition for the mixing time of the glauber dynamics for coloring regular trees. In *Proceedings of the 2010 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2010. doi:10.1137/1.9781611973075.134.

[73] Shai Vardi. Randomly coloring graphs of logarithmically bounded pathwidth. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018.

[74] Pengfei Wan, Jianhua Tu, Shenggui Zhang, and Binlong Li. Computing the numbers of independent sets and matchings of all sizes for graphs with bounded treewidth. *Applied Mathematics and Computation*, 332:42 – 47, 2018. doi:10.1016/j.amc.2018.03.017.

Appendix A

Missing details for triangulations

A.1 Nearly tight conductance for triangulations: lower bound

Lemma 2.14 and Lemma 2.12, as we showed in Appendix 2.6, imply the known result that the flip walk on triangulations of the convex polygon mixes rapidly. However, the bound given by Lemma 2.14 is $O(n^2)$ congestion, giving $O(n^7)$ mixing time by Lemma 2.1. Through a more careful flow construction, one can further improve this bound to $O(n^3 \log^3 n)$. For the more careful construction, we will define a different decomposition, via the *central triangle*: Given a triangle T containing the center of the regular $n+2$-gon P_{n+2} and sharing all of its vertices with P_{n+2}, identify T with the class $\mathcal{C}(T)$ of triangulations $t \in V(K_n)$ such that T forms one of the triangles in t. Let $\mathcal{S}_n$ be the set of all such $\mathcal{C}(T)$ classes. (If P_{n+2} has an even number of edges, we perturb the center slightly so that every triangulation lies in some class.)

Remark A.1. *The set $\mathcal{S}_n$ is a partition of $V(K_n)$, because no pair of triangles whose endpoints are polygon vertices can contain the origin without crossing.*

Molloy, Reed, and Steiger [56] defined this same partition in their work.

See Figure A.1.

We will combine this central-triangle decomposition with the oriented decomposition we defined earlier. What we gain from using the central-triangle decomposition is that the number of levels of induction will now be $O(\log n)$, by the fact that using the central triangle to partition the classes divides the $n+2$-gon into smaller polygons of size $\leq n/2$. What we *lose*, however, is that we no longer have matchings between every pair of classes, nor are all of the matchings between adjacent pairs sufficiently large to obtain a polynomial bound. Thus if we were to use just this decomposition on its own, we would be stuck with the quasipolynomial bound (which in fact is what we obtain for general k-angulations in Appendix 2.7).

Fortunately, we will show how to combine the two decompositions. With suitable care, this will allow us to eliminate one of the factors of n—which we incurred in the n levels of induction via Lemma 2.12. Some further optimizations will give us the claimed congestion bound $O(\sqrt{n})$:

Lemma A.1. *Suppose that for all* $1 \leq i \leq n/2$, *a uniform multicommodity flow exists with congestion* $O(\sqrt{i}\log i)$ *in* K_i. *Then a uniform multicommodity flow exists in* K_n *with congestion* $O(\sqrt{n}\log n)$.

(Here of course the constant hidden in the O notation is independent of the number of induction levels.) Once we prove this lemma, then clearly Theorem 1.1 follows via simple induction and an application of Lemma 2.1. (The base case in the induction is trivial.)

What we will do is, before routing the flow between triangulations in two different classes, to do the same shuffling step as before—this time within each class in the *central-triangle* partition, instead of within each class in the *oriented* partition. This is simply a "scaled-up" uniform multicommodity flow in each class and produces no increase in congestion, using the same analysis as before. We then have an MSF problem for each pair of classes, in which the flow is routed through a set of intermediate classes, and:

Remark A.2. *The boundary set $\mathcal{B}_{T'}(T)$ between every pair of central-triangle-induced classes is isomorphic to a Cartesian product $\mathcal{C}^*(U)\Box K_j\Box K_l$, where $\mathcal{C}^*(U)$ is an* oriented *class in K_i, and where $\mathcal{C}(T) \cong K_i\Box K_j\Box K_l$, $i+j+l=n-2$, $i,j,l \leq n/2$.*

In other words, even though we are now using the *central-triangle* decomposition, our boundary classes are, as before, Cartesian products of *oriented* classes with associahedron graphs. Therefore:

Remark A.2, combined with our earlier congestion analysis for concentration and distribution flows, implies the following:

Lemma A.2. *Suppose it is possible to construct a multicommodity flow f in K_n in which the total congestion across edges between a pair of classes is at most $\bar{\rho}$. Then the total congestion produced by f is at most $2\bar{\rho}n$.*

Proof. Routing flow through an intermediate class $\mathcal{C}(T'')$, say, that originates at class $\mathcal{C}(T)$ and is bound for $\mathcal{C}(T')$, can be accomplished with the combination of a uniform multicommodity flow in $\mathcal{C}(T)$ (a "shuffling flow"), scaled as in the construction in Appendix 2.6, and an MSF with source set $\mathcal{C}(T)$ and sink set $\mathcal{C}(T'')$. This MSF then induces in $\mathcal{C}(T'')$ both a concentration flow and a distribution flow. Notice that $\mathcal{C}(T'')$ has $O(n)$ distinct neighboring classes. Therefore there are $O(n)$ such concentration flows and $O(n)$ such distribution flows. Since each concentration flow and each distribution flow produces no increase in congestion relative to the amount across each edge between classes, the claim follows. □

We will use this idea of combining decompositions to obtain our $O(\sqrt{n}\log n)$ congestion bound. We will exhibit a flow with $\bar{\rho} = O(\sqrt{n})\log n$, and will show how to avoid the $O(n)$

gain in Lemma A.2.

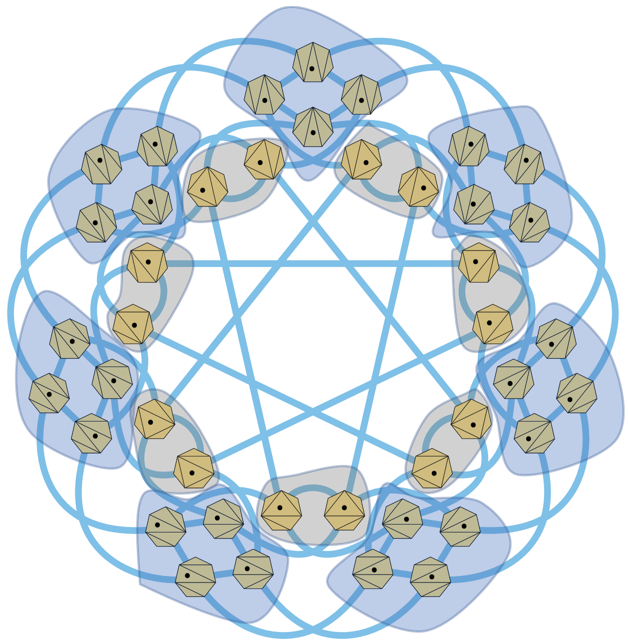

Figure A.1: An alternative partitioning of the associahedron graph K_5, with each vertex representing a triangulation of the regular heptagon. Flips are shown with edges (in blue). The vertex set $V(K_n)$ is partitioned into a set $\mathcal{S}_n$ of equivalence classes. Within each class, all triangulations share the same *central triangle*—contrast with the oriented partition depicted in Figure 2.2.

We do so by choosing carefully a good set of "paths" between each pair of classes, where each path consists of a sequence of intermediate classes through which to route flow. One first attempt might be, given classes $\mathcal{C}(S)$ and $\mathcal{C}(U)$, to consider the P_{n+2} vertices by which S and U differ. Route flow from $\mathcal{C}(S)$ to $\mathcal{C}(S')$, where S' is a triangle formed by replacing some vertex of $S \setminus U$ with a vertex of $U \setminus S$. This results in routing flow through at most two intermediate classes. Unfortunately, if we do this for all S, U pairs, then some of these intermediate classes will end up routing flow for too many S, U pairs, and the congestion improvement will be insufficient for our purposes.

Roughly speaking, and perhaps counterintuitively, it turns out that the large congestion under the scheme described above results from using paths that are too short: there exist many pairs of large classes $\mathcal{C}(S), \mathcal{C}(U)$ such that the intermediate classes found under this scheme are much smaller than $\mathcal{C}(S)$ and $\mathcal{C}(U)$, and thus cannot effectively "spread out" the congestion that results from sending the $\mathcal{C}(S), \mathcal{C}(U)$ flow. Instead we will find slightly longer paths.

To define these longer paths, we first need to organize the classes into a (non-disjoint) union of larger classes, which we call *regions*: Let the *length* of a diagonal connecting two n-gon vertices a and b be the number of edges on the n-gon lying between a and b, when traversing the shorter of the two distances around the n-gon from a to b. Mark 24 equally spaced points on the convex polygon P_{n+2}, in counterclockwise order. Define the following 24 *regions* as collections of the classes: let $\mathcal{U}_i$, $0 \leq i \leq 23$, be the set of all classes $\mathcal{C}(T)$ such that the vertex *opposite the shortest edge* of T lies in the (inclusive) interval $[i \cdot n/24, (i+2) \cdot n/24)$.

This is not a partition of the classes $\{\mathcal{C}(T)\}$, since it is possible for a triangle T to have a vertex in two of the intervals described. However:

Remark A.3. *All of the regions are of equal size, and the regions each have cardinality* $C_n/12$. *Also, the regions form a cycle, in which each consecutive pair of regions shares an overlap of size* $C_n/24$.

The idea now is that we will establish the existence of a flow, within each of the 24 regions, that has congestion $O(\sqrt{n/2} + \log(n/2))$. Once this is accomplished, we will use the constant-factor intersections of the classes to route flow between classes with additional (additive) congestion $O(\sqrt{n} + \log n)$.

We will further partition the central-triangle-induced classes into two additional levels of classes.

Given a central triangle T, let the *apex* of T be the vertex of T opposite the shortest side. (If T has no unique shortest side, break ties in some arbitrary fashion; this will not change the asymptotics.) Let the *second* vertex of T be the first of the two non-apex vertices that succeeds the apex in counterclockwise order; let the remaining vertex be the *third* vertex. Given a vertex $p \in [0, n+1]$ of the $n+2$-gon, with the vertices labeled in counterclockwise order, let $\mathcal{A}_p$ be the set of classes $\mathcal{C}(T)$ with p as the apex of T. Let $\mathcal{L}_{pq}$ be the set of classes $\mathcal{C}(T)$ in $\mathcal{A}_p$ with q as the second vertex of T.

Remark A.4. *Every $\mathcal{L}_{pq}$ is a collection of central-triangle classes sharing an edge, namely the diagonal pq. Every class $\mathcal{C}(T) \in \mathcal{L}_{pq}$ is, therefore, a Cartesian product $K_{q-p-1} \Box \mathcal{C}^*(T)$ of a flip graph K_{q-p-1} over the $q-p+1$-gon on one side of the diagonal pq, and a class $\mathcal{C}^*(T) \subseteq V(K_{n-q+p-1})$ in the* oriented *partition induced by the edge pq in the $n-q+p+1$-gon on the other side of pq.*

Lemma A.3. *Within $\mathcal{L}_{pq}$, for all $p \in [0, n+1]$, it is possible to route a unit of flow between every ordered pair of triangulations $t, t' \in \mathcal{L}_{pq}$ while producing congestion one across the edges between any pair of central-triangle-induced classes.*

Proof. By Remark A.4, we can apply Lemma 2.6 to conclude that every pair of classes $\mathcal{C}(T), \mathcal{C}(T') \in \mathcal{L}_{pq}$ can exchange one unit of a commodity with congestion at most

$$\frac{|\mathcal{C}(T)||\mathcal{C}(T')|}{|\mathcal{E}(T,T')||V(K_n)|} \leq \frac{|V(K_{q-p-1})\Box\mathcal{C}^*(T)||V(K_{q-p-1})\Box\mathcal{C}^*(T')|}{|V(K_{q-p-1})\Box\mathcal{E}^*(T,T')||V(K_n)|} \leq \frac{|\mathcal{C}^*(T)||\mathcal{C}^*(T')|}{|\mathcal{E}^*(T,T')|} \cdot \frac{C_{q-p-1}}{C_n},$$

with the first expression describing the product of the cardinalities of the classes divided by the number of edges between them, and a normalization factor $|V(K_n)| = C_n$ according to the definition of congestion. The first inequality comes from Remark A.4, and the second from rearranging terms. Applying Lemma 2.6 gives an upper bound of

$$\frac{C_{n-q+p-q}C_{q-p-1}}{C_n} \leq 1.$$

(Actually, this quantity is not only at most one but at most $\frac{C_{q-p-1}C_{n-q+p}}{C_n} = O(1/(q-p)^{3/2})$.) $\square$

Lemma A.3 describes only the flow across edges *between pairs of central-triangle classes.*

The construction in Appendix 2.6 shows how to obtain polynomial congestion *within* classes from this bound. However, as we observed in the proof of Lemma 2.16, one suffers a loss accounting for flow from $\kappa = O(n)$ classes. We now show how to improve this $O(n)$ loss to $O(\sqrt{n})$.

For the purpose of analyzing congestion, a uniform multicommodity flow in K_n is equivalent to the sum of $|V(K_n)| = C_n$ *single*-commodity flows, one "originating" at (having sink set as) a single triangulation $t \in V(K_n)$. Furthermore, given the *oriented* partition $\mathcal{S}_n$, and considering the classes $\mathcal{C}^*(T_1), \mathcal{C}^*(T_2), \ldots, \mathcal{C}^*(T_n) \in \mathcal{S}_n$, a uniform multicommodity flow in K_n is equivalent to the sum of n uniform multicommodity flows, one *within* each class $\mathcal{C}^*(T_i)$, *added* to n *multi-way single-commodity flows* (MSFs), each of which distributes flow from one class $\mathcal{C}^*(T_i)$ to the rest of the graph K_n.

The congestion bound one obtains for the MSFs (ignoring the flows within the classes) from the analysis in the proof of Lemma 2.16 is then $\kappa = O(n)$. The following lemma states that we can do better: we can solve these n MSF problems with congestion $O(\sqrt{n})$ by improving the flow construction. Intuitively, given two triangles T_i, T_j with third vertices i and j on the $n+2$-gon, the size of the matching $\mathcal{E}^*(T_i, T_j)$ between $\mathcal{C}^*(T_i)$ and $\mathcal{C}^*(T_j)$ is large when $|j-i|$ is small, and small when $|j-i|$ is large. When T_i and T_j are far apart ($|j-i|$ is large), we will route some of the $\mathcal{C}^*(T_i) \to \mathcal{C}^*(T_j)$ flow through a sequence of intermediate classes $\{\mathcal{C}^*(T_k)\}$, $i < k < j$, taking advantage of the larger matchings between $\mathcal{C}^*(T_i)$ and $\mathcal{C}^*(T_k)$, and between $\mathcal{C}^*(T_k)$ and $\mathcal{C}^*(T_j)$.

In particular, we will first group the classes into pairs of consecutive classes $\mathcal{C}^*(T_i), \mathcal{C}^*(T_{i+1})$ (with, say, $i = 0 \pmod 2$), and let the two classes within a given pair exchange flow, so that all of the flow originating at either class in the pair is uniformly distributed throughout the pair $\mathcal{C}^*(T_i) \cup \mathcal{C}^*(T_{i+1})$. That way, subsequent flow sent by the two classes can now be considered as a single commodity. We will then group these pairs of classes into sets of four classes, then sets of eight, and so on—reaching $O(\log n)$ hierarchical levels of sets, until all $n(n-1)$ ordered pairs of classes have exchanged flow.

Lemma A.4. *Given the flip graph K_n over the $n+2$-gon and the special edge e^*, consider the triangles $T_1, T_2, \ldots, T_n$ that include e^* as an edge, such that $T_1, T_2, \ldots, T_n$ occur in consecutive order according to their third vertex. Consider the n MSF problems $\pi_1, \pi_2, \ldots, \pi_n$, one for each* oriented *class $\mathcal{C}^*(T_i)$, $i = 1, 2, \ldots, n$. Suppose each π_i has source set $\mathcal{C}^*(T_i)$ and sink set $V(K_n)$, with uniform surplus and demand functions $\sigma_i = |V(K_n)| = C_n, \delta_i = |\mathcal{C}^*(T_i)|$. Then $\pi_1, \pi_2, \ldots, \pi_n$ can be reduced to an alternative collection of MSF problems that can be solved with congestion $O(\sqrt{n})$.*

Proof. Assume for simplicity that n is a power of two; it is easy to modify the solution if not. Group the classes hierarchically as described in the disussion preceding this lemma. Let $\pi_{[i,j]}$ be the problem, defined over the subgraph of K_n induced by the classes $\mathcal{C}^*(T_i) \cup \cdots \cup \mathcal{C}^*(T_j)$, of distributing flow from the "left half" of the classes $\mathcal{C}^*(T_i) \cup \cdots \cup \mathcal{C}^*(T_{i+\frac{j-i+1}{2}-1})$ to the "right half" $\mathcal{C}^*(T_{i+\frac{j-i+1}{2}}) \cup \cdots \cup \mathcal{C}^*(T_j)$. Define $\bar{\pi}_{[i,j]}$ symmetrically.

As discussed, the original collection of MSF problems $\{\pi_i\}$ reduces to a collection $\{\pi_{[i,j]}, \bar{\pi}_{[i,j]}\}$, where the pairs $[i,j]$ are those induced by hierarchically partitioning the classes—first into problems $\pi_{[1,n]}$ and $\bar{\pi}_{[1,n]}$, then into (on the "left-hand side") $\pi_{[1,n/2]}, \bar{\pi}_{[1,n/2]}$ and (on the "right-hand side") $\pi_{[n/2+1,n]}, \bar{\pi}_{[n/2+1,n]}$, then into four pairs of problems, and so on.

Now, for a given pair of problems $\pi_{[i,j]}, \bar{\pi}[i,j]$, each class $\mathcal{C}^*(T_l)$ on the "left-hand side" $\mathcal{C}^*(T_i) \cup \cdots \cup \mathcal{C}^*(T_{i+\frac{j-i+1}{2}-1})$, $i \leq l \leq i + \frac{j-i+1}{2} - 1$, must distribute $C_n|\mathcal{C}^*(T_l)|$ units of flow—that is, the demand $\sigma_{[i,j]}$ times the size of the source set $\mathcal{C}^*(T_l)$ of $\pi_{[i,j]}$—to the right-hand side, and vice versa. Each class $\mathcal{C}^*(T_r)$ on the right-hand side receives a $\frac{|\mathcal{C}^*(T_r)|}{|\sum_{r' \in [i,j]} |\mathcal{C}^*(T_{r'})|}$ factor of this flow, and the flow must be distributed across the matching $|\mathcal{E}^*(T_l, T_r)|$.

This produces congestion at most

$$O\left(\frac{C_n|\mathcal{C}^*(T_l)||\mathcal{C}^*(T_r)|}{|\mathcal{E}^*(T_l,T_r)|\sum_{r'\in[i,j]}|\mathcal{C}^*(T_{r'})|C_n}\right)=O\left(\frac{|\mathcal{C}^*(T_l)||\mathcal{C}^*(T_r)|}{|\mathcal{E}^*(T_l,T_r)|\sum_{r'\in[i,j]}|\mathcal{C}^*(T_{r'})|}\right).$$

We will bound this quantity as $O(\sqrt{j-i})$, by showing that $\frac{|\mathcal{C}^*(T_l)|}{|\mathcal{E}^*(T_l,T_r)|}=O((j-i)^{3/2})$ and that $\frac{|\mathcal{C}^*(T_r)|}{|\sum_{r'\in[i,j]}|\mathcal{C}^*(T_{r'})|}=O(1/(j-i))$.

The first inequality is true because, by Lemma 2.4, $\mathcal{E}^*(T_l,T_r)$ is in bijection with the vertex set of a Cartesian product $K_{l-1}\square K_{r-l-1}\square K_{n-r}$ graph, whereas $\mathcal{C}^*(T_l)\cong K_{l-1}\square K_{n-l}$. Thus $\frac{|\mathcal{C}^*(T_l)|}{|\mathcal{E}^*(T_l,T_r)|}\leq\frac{C_{n-l}}{C_{r-l-1}n-r}$. We can assume without loss of generality that $1\leq l\leq r\leq n/2$ (since T_l and T_r send one another the same amount of flow), and therefore this quantity is at most $O((r-l)^{3/2})=O((j-i)^{3/2})$.

The second inequality can be seen by noticing that for all $r'\in[i,j]$, $\mathcal{C}^*(T_{r'})\cong K_{r'-1}\square K_{n-r'}$, so $|\mathcal{C}^*(T_{r'})|=C_{r'-1}C_n-r'=\frac{C_n}{\Theta(r'^{3/2})}$. Since this is a decreasing function of r', we have

$$\sum_{i\leq r'\leq j}|\mathcal{C}^*(T_{r'})|\geq\sum_{i\leq r'\leq r}|\mathcal{C}^*(T_{r'})|\geq(r-i+1)|\mathcal{C}^*(T_r)|\geq\frac{(j-i)}{2}|\mathcal{C}^*(T_r)|,$$

which implies certainly that $|\mathcal{C}^*(T_r)||\sum_{r'\in[i,j]}|=O(j-i)$.

Since every MSF pair $\pi_{[i,j]},\hat{\pi}_{[i,j]}$ is solvable with congestion $O\left(\sqrt{j-i}\right)$, where $j-i=1,2,4,8,\ldots,n$, the overall congestion is $\sum_{k=0}^{\log n}\sqrt{2^k}=O(\sqrt{n})$, as claimed.

Finally, one may worry that there may be a factor $(j-i+1)/2$ gain in congestion for each $\pi_{[i,j]},\bar{\pi}_{[i,j]}$ pair, since $\mathcal{C}^*(T_r)$ must receive flow from $(j-i+1)/2$ classes—just as we had a κ-factor gain in the proof of Lemma 2.16. However, that gain occurred because we had κ separate MSF problems. Here, however, we only have two MSF problems, inducing two flows. The same construction we used in that proof then gives $O(1)$ congestion per $\pi_{[i,j]},\bar{\pi}_{[i,j]}$ pair.

□

Corollary A.1. *Within $\mathcal{L}_{pq}$, for a given $p \in [0, n+1]$, consider a collection of $q-p-1$ MSF problems, each of which corresponds to one class $\mathcal{C}(T) \in \mathcal{L}_{pq}$ and describes distributing a single commodity with surplus value C_n, initially concentrated in $\mathcal{C}(T)$, throughout the rest of $\mathcal{L}_{pq}$. All of these problems can be solved while producing total congestion $O(\sqrt{q-p})$.*

Proof. It suffices to combine the constructions in Lemma A.3 and Lemma A.4. The exchange in Lemma A.3, that is, induces MSF subproblems that can be viewed, by Remark A.4, as an exchange between pairs of *oriented* subclasses of K_{n-q+p} (in copies of K_{q-p-1})—in which the surplus values are all $C_{q-p-1}C_{n-q+p}$. We can then apply Lemma A.4 to obtain congestion $O(\sqrt{n})\frac{C_{q-p-1}C_{n-q+p}}{C_n}$.

Here we need to be careful. First, it may be that this bound exceeds $O(\sqrt{n})$, in particular if $\frac{C_{q-p-1}C_{n-q+p}}{C_n} \geq \omega(1/\sqrt{n})$. Fortunately, by the proof of Lemma A.4, the $O(\sqrt{n})$ bound in that lemma can be sharpened to $O(\sqrt{q-p})$ (by noticing that at no level of the hierarchical partitioning do we have $j-i > q-p$). Thus we have the congestion bound $O(\sqrt{q-p})\frac{C_{q-p-1}C_{n-q+p}}{C_n}$.

Finally, we have assumed surplus values of $C_{q-p-1}C_{n-q+p}$. Actually, however, the present claim concerns surplus values C_n. Scaling by $\frac{C_n}{C_{q-p-1}C_{n-q+p}}$ gives congestion $O(\sqrt{q-p})$. □

Lemma A.5. *Within a given $\mathcal{A}_p$, $p \in [0, n+1]$, consider a collection of MSF problems, one for each $\mathcal{L}_{pq} \in \mathcal{A}_p$ (with source set $\mathcal{L}_{pq}$, with surplus values C_n), with flow that must be distributed uniformly throughout $\mathcal{A}_p$. It is possible to solve these problems while producing total congestion $O(\sqrt{n}\log n)$.*

Proof. To route flow between pairs of central-triangle classes lying in distinct second-vertex classes, i.e. between $\mathcal{C}(T_{pqr}) \in \mathcal{L}_{pq}, \mathcal{C}(T_{pq'r'}) \in \mathcal{L}_{pq'}$, $q \neq q'$, we will use the same trick as in the hierarchical grouping in Lemma A.4. Assume that $p = 0$ without loss of generality and for simplicity. For all $\mathcal{L}_{pq}, \in \mathcal{A}_p$, it holds that $n/4 \leq q < n/2$: $q \geq n/4$ since the triangle edge pq must be at least as long as the edge qr in any $T_{pqr}, \mathcal{C}(T_{pqr}) \in \mathcal{L}_{pq}$ by definition of $\mathcal{L}_{pq}$ and of $\mathcal{A}_p$, and $q \leq n/2$ since for every $\mathcal{C}(T_{pqr})$, T_{pqr} is a central triangle. It will turn out to be convenient to include in the grouping only the classes $\{\mathcal{L}_{pq} | 7n/24 \leq q < n/2\}$. Order the second-vertex-induced classes $\{\mathcal{L}_{pq}\}$ with $q \in [7n/24, n/2-1]$ in increasing order. Group pairs of adjacent second-vertex classes, then group these pairs into adjacent pairs, and so on.

Suppose $q < q'$ without loss of generality. At the $j-i$ level of the grouping, i.e. the level at which the number of second-vertex classes on the left- and right-hand sides combined is $j-i$, the amount of flow to be exchanged between $\mathcal{L}_{pq}$ and $\mathcal{L}_{pq'}$ lying on respectively the left and right-hand sides of the group, in each direction, is

$$\frac{C_n |\mathcal{L}_{pq}||\mathcal{L}_{pq'}|}{\sum_{i \leq q'' \leq j} |\mathcal{L}_{pq''}|},$$

where $[i,j]$ is the interval of classes $\mathcal{L}_{pi}, \ldots, \mathcal{L}_{pj}$ defining the group.

Let $\mathcal{E}(L_{pq}, L_{pq'}) = \bigcup_{\mathcal{C}(T_{pqr}) \in \mathcal{L}_{pq}} \mathcal{E}(T_{pqr}, T_{pq'r})$ denote the matching connecting $\mathcal{L}_{pq}$ and $\mathcal{L}_{pq'}$. The resulting congestion is at most

$$\frac{|\mathcal{L}_{pq}||\mathcal{L}_{pq'}|}{\sum_{i \leq q'' \leq j} |\mathcal{L}_{pq''}||\mathcal{E}(L_{pq}, L_{pq'})|} \leq \frac{O((q'-q)^{3/2})}{(j-i)},$$

where the inequality holds because $|\mathcal{L}_{pq''}| \geq |\mathcal{L}_{pq}|$ for $i+(j-i)/2 \leq q'' \leq j$, so that $\frac{|\mathcal{L}_{pq}|}{\sum_{i \leq q'' \leq j} |\mathcal{L}_{pq''}|} \leq \frac{1}{(j-i)/2}$, and because $\frac{|\mathcal{L}_{pq'}|}{|\mathcal{E}(L_{pq}, L_{pq'})|} = O((q'-q)^{3/2})$. The latter fact can be seen as follows: first,

$$\mathcal{E}(L_{pq}, L_{pq'}) = \bigcup_{T_{pqr} \in \mathcal{L}_{pq}} \mathcal{E}(T_{pqr}, T_{pq'r}).$$

Every $\mathcal{C}(T_{pqr})$ has a nonempty matching to its neighboring class $\mathcal{C}(T_{pq'r})$, and indeed $\mathcal{C}(T_{pq'r})$ lies in $\mathcal{A}_p$ and in $\mathcal{L}_{pq'}$. On the other hand, due to the constraint for membership in $\mathcal{A}_p$ that pq and pq' be the shortest edges of their respective central triangles, there may exist some values of r for which $\mathcal{T}_{pq'r} \in \mathcal{L}_{pq'}$ but for which there is no neighbor of $\mathcal{T}_{pq'r} \in \mathcal{L}_{pq}$. Fortunately:

(i) Since by assumption $q \geq 7n/24$, we get that $\mathcal{C}(T_{pqr}) \in \mathcal{L}_{pq}$ (i.e. the edge qr is indeed shorter than the edges pq and rp) for $r = n/2+1, \ldots, 2 \cdot 7n/24 = 7n/12$, and thus there are at least $n/12$ central-triangle classes in $\mathcal{L}_{pq}$ (and thus at least as many in $\mathcal{L}_{pq'}$.

(ii) In $\mathcal{L}_{pq}$ (and similarly $\mathcal{L}_{pq'}$), the central-triangle classes occur in decreasing order of size (up to asymptotic order) as r increases.

Facts (i) and (ii) imply that an $\Omega(1)$ factor of the triangulations in $\mathcal{L}_{pq'}$ lie in central-triangle classes having a neighboring class in $\mathcal{L}_{pq}$, and thus

$$|\mathcal{E}(L_{pq}, L_{pq'})| = \Omega(1)\Omega\left(\frac{1}{(q'-q)^{3/2}}\right)|\mathcal{C}(L_{pq'})|.$$

Now, the $\frac{O((q'-q)^{3/2})}{(j-i)} \leq \sqrt{j-i} \leq \sqrt{n}$ congestion that occurs across the boundary matching $\mathcal{E}(L_{pq}, L_{pq'})$ for a given q, q' pair occurs for a *single commodity*, at a single level in the hierarchical grouping. We need to distribute this flow evenly first throughout each class $\mathcal{C}(T_{pqr})$ that receives it, and then throughout $\mathcal{L}_{pq}$. By the same reasoning as in the proof of Lemma A.4, this flow can be distributed throughout a given class $\mathcal{C}(T_{pqr}) \in \mathcal{L}_{pq}$ with no asymptotic congestion gain. Summing over all levels of the grouping produces

$$\sum_{s=0}^{\log(n/2-7n/24)} O(\sqrt{2^s}) = O(\sqrt{n})$$

congestion within each $\mathcal{C}(T_{pqr})$.

To distribute the flow received by $\mathcal{C}(T_{pqr})$ throughout $\mathcal{L}_{pq}$, first notice that the total amount of (normalized by a factor of C_n) flow received by $\mathcal{C}(T_{pqr})$ is at most $O(\log(r-q))|\mathcal{C}(T_{pqr})|$ from classes $\mathcal{C}(T_{pq'r}), q < q' < (r-q)/2$, because each vertex in each boundary set $\mathcal{B}_{T_{pq'r}}(T_{pqr})$

receives $O(\sqrt{q'-q})$ flow and $|\mathcal{B}_|(T_{pqr}) = \Theta(\frac{1}{(q'-q)^{3/2}})|\mathcal{C}(T_{pqr})|$, so the total is

$$|\mathcal{C}(T_{pqr})| \sum_{k=1}^{(r-q)/2} \frac{\sqrt{k}}{k^{3/2}} = O(\log(r-q))|\mathcal{C}(T_{pqr})|.$$

The analysis is similar for classes with $q' < q$.

Now notice that the total amount of normalized flow received by $\mathcal{C}(T_{pqr})$ from classes $\mathcal{C}(T_{pq'r})$ with $q' - q \geq (r-q)/2$ is at most

$$\sum_{k=r-n/2}^{(r-q)/2} O\left(\frac{1}{k^{3/2}}\right) \cdot O(\sqrt{(r-q)})|\mathcal{C}(T_{pqr})| = O(\frac{\sqrt{r-q}}{\sqrt{r-n/2}})|\mathcal{C}(T_{pqr})|.$$

Recall that we are dealing with a single commodity. Thus we do not have multiple MSFs to be concerned about. Unfortunately, however, the bound given by the construction in the proof of Lemma 2.16 gives a bound of $O(\frac{\sqrt{r-q}}{\sqrt{r-n/2}} \cdot \sqrt{n})$, insufficient for our purposes.

Fortunately, we can apply the hierarchical grouping trick again within $\mathcal{L}_{pq}$, but we need to take care: first, it is insufficient merely to apply Corollary A.1, as we simply recover the $O(\sqrt{n})$ factor gain mentioned above. Second, unlike in Corollary A.1, we are dealing here with only a single commodity (this will help us). What we do is observe that since the average flow received by a class $\mathcal{C}(T_{pqr})$ is, as stated, $O(\frac{\sqrt{r-q}}{\sqrt{r-n/2}}) = O(\frac{\sqrt{n/2-q}}{\sqrt{r-n/2}})$, and the average over all classes $\mathcal{C}(T_{pqr})$ within the range $r \in [i, j]$ (assuming for the worst case that $i = n/2 + (n/2 - 1) = 1$, since we are only considering flow from classes q' where this holds in the worst case) is at most

$$\sum_{r=n/2+1}^{n/2+j} \frac{(j+n/2-q)^{3/2}\sqrt{n/2-q}}{j(r-q)^{3/2}\sqrt{r-n/2}},$$

where we have used the fact that

$$\sum_{s\in[n/2+1,n/2+j]} |\mathcal{C}(T_{pqs})| \geq j \cdot \frac{\sqrt{n/2-q}}{(j+n/2-q)^{3/2}}|\mathcal{L}_{pq}|,$$

since the left-hand side is a sum of j terms each of which is at least

$$|\mathcal{C}(T_{pq(n/2+j)})| = \frac{\sqrt{n/2-q}}{(j+n/2-q)^{3/2}}|\mathcal{L}_{pq}|,$$

and also the fact that

$$|\mathcal{C}(T_{pqr})| = \frac{O(\sqrt{n/2-q})}{(r-q)^{3/2}}|\mathcal{L}_{pq}|.$$

When $j \leq n/2-q$, we can bound the term

$$\begin{aligned}\sum_{r=n/2+1}^{n/2+j} \frac{(j+n/2-q)^{3/2}\sqrt{n/2-q}}{j(r-q)^{3/2}\sqrt{r-n/2}} &\leq \frac{(j+n/2-q)^{3/2}}{n/2-q}\frac{1}{j}\sum_{r=n/2+1}^{n/2+j}\frac{1}{\sqrt{r-n/2}} \\ &\leq \sqrt{n/2-q}\frac{\sqrt{j}}{j} = \frac{\sqrt{n/2-q}}{\sqrt{j}}\end{aligned}$$

since $r-q \geq n/2-q$ always, and since we are assuming $j \leq n/2-q$.

When $j > n/2-q$, remember that we are considering only the flow to each $\mathcal{C}(T_{pqr})$ from classes $\mathcal{C}(T_{pq'r})$ with $q'-q \geq (r-q)/2$, and therefore with $r \leq n/2-q$. Thus the average never exceeds $\frac{\sqrt{n/2-q}}{\sqrt{j}}$.

Let μ_{j-i} denote this average. Now we can bound the congestion across a given matching $\mathcal{E}(T_{pqr}, T_{pqr'})$, for $\mathcal{C}(T_{pqr}), \mathcal{C}(T_{pqr'}) \in \mathcal{L}_{pq}$ as

$$\begin{aligned}\frac{C_n\mu_{j-i}|\mathcal{C}(T_{pqr})||\mathcal{C}(T_{pqr'})|}{C_n\sum_{s=i}^{j}|\mathcal{C}(T_{pqs})||\mathcal{E}(T_{pqr},T_{pqr'})|} &\leq \frac{(r'-r)^{3/2}\mu_{j-i}}{j-i} \leq \frac{(r'-r)^{3/2}\sqrt{n/2-q}}{(j-i)^{3/2}} \\ &\leq \sqrt{n/2-q} \leq \sqrt{n}\end{aligned}$$

for all $j-i$. Since we are dealing with a single commodity and every class $\mathcal{C}(T_{pqr})$ has at most $O(\sqrt{n})$ surplus or demand in each of its boundary vertices, we can use the same construction as in the proof of Lemma 2.16 to conclude that the overall resulting congestion in $\mathcal{L}_{pq}$ is at most $O(\sqrt{n})$.

It remains to consider the flow received by $\mathcal{C}(T_{pqr})$ from classes $\mathcal{C}(T_{pq'r})$ with $q'-q \leq (r-q)/2$. As we have already observed, the average for each $\mathcal{C}(T_{pqr})$ is at most $O(\log(r-q)) = O(\log n)$, and thus we can simply apply Corollary A.1 to obtain $O(\sqrt{n}\log n)$ congestion.

Lastly, we have only distributed flow so far among the classes $\mathcal{L}_{pq}$ with $q \geq 7n/24$. We need to send flow from classes with $q \geq 7n/24$ to those with $n/4 \leq q < 7n/24$ and vice versa. We will first use the same construction as above to concentrate all of the flow from the $[7n/24, n/2]$ classes within the $[7n/24, n/3]$ classes. Because (as one can show) the $[7n/24, n/3]$ classes constitute a $\Theta(1)$ factor of the triangulations in $\mathcal{A}_p$, this concentration causes at most an $O(1)$ increase in congestion.

Now let the $[7n/24, n/3]$ and the $[n/4, 7n/24]$ classes exchange flow. Once more we apply the hierarchical grouping trick. The challenge is now that the number of central-triangle classes in $\mathcal{L}_{pq}$ is small. Let $\chi(\mathcal{L}_{pq}) = |\{\mathcal{C}(T)_{pqr}|\mathcal{C}(T_{pqr}) \in \mathcal{L}_{pq}\}|$ denote the number of central-triangle classes in $\mathcal{L}_{pq}$. We have $\chi(\mathcal{L}_{pq}) = 2(q-n/4)$ whenever $n/4 \leq q \leq n/3$.

Thus for $n/4 \leq q < q' \leq n/3$ we can bound

$$\frac{C_n|\mathcal{L}_{pq}||\mathcal{L}_{pq'}|}{C_n\sum_{q''=i}^{j}|\mathcal{L}_{pq''}||\mathcal{E}(L_{pq}, L_{pq'})|} \leq \frac{(q'-q)^{3/2}\chi(\mathcal{L}_{pq'})(1/n^3)}{(j-i)/2\chi(\mathcal{L}_{p(q'-(j-i)/2)})(1/n^3)} = O(\sqrt{j-i})$$

for each group, and we are done. □

The following two lemmas will be useful shortly:

Lemma A.6. *Within a given $\mathcal{L}_{pq}$, consider a* single *MSF problem in which the source set* *is a union of central-triangle classes, and each $\mathcal{C}(T) \in \mathcal{L}_{pq}$ has surplus value $O(1)C_n$ (where the values may differ among the classes). It is possible to solve this problem while producing congestion $O(\sqrt{n/2-q})$.*

Proof. Let every $\mathcal{C}(T_{pqr})$ send a $|\mathcal{C}(T_{pqr'})|/|\mathcal{L}_{pq}|$ factor of its surplus to each $\mathcal{C}(T_{pqr'})$. By Lemma 2.6 and by the fact that $|\mathcal{L}_{pq}| = \Theta(1)\frac{C_{q-p}C_{n-(q-p)}}{\sqrt{n/2-q}}$, we have

$$\frac{C_n|\mathcal{C}(T_{pqr})||\mathcal{C}(T_{pqr'})|\sqrt{n/2-q}}{C_n|\mathcal{E}(T,T')|C_{q-p}C_{n-(q-p)}} \leq \sqrt{n/2-q}$$

congestion across each edge in $\mathcal{E}(T,T')$, proving the lemma. □

Lemma A.7. *Within a given $\mathcal{A}_p$, consider a* single *MSF problem in which the source set is a union of central-triangle classes, and each $\mathcal{C}(T) \in \mathcal{A}_p$ has surplus value $O(1)C_n$ (where the values may differ among the classes). It is possible to solve this problem while producing total congestion $O(\sqrt{n})$.*

Proof. Distributing incoming flow throughout each $\mathcal{L}_{pqr}$ causes $O(\sqrt{n})$ congestion within $\mathcal{L}_{pqr}$. Now modify the flow construction in the proof of Lemma A.5 as follows: omit the hierarchical grouping trick when distributing flow among the second-vertex classes, and instead simply let each $\mathcal{L}_{pq}$ send an $|\mathcal{L}_{pq'}|/|\mathcal{A}_p|$ factor its surplus to each $\mathcal{L}_{pq'}$ once. Then the congestion across $\mathcal{E}(\mathcal{L}_{pq}, \mathcal{L}_{pq'})$ when $\mathcal{L}_{pqr}$ sends flow to $\mathcal{L}_{pq'}$ (assuming $q < q'$) is

$$\frac{|\mathcal{L}_{pq}||\mathcal{L}_{pq'}|}{|\mathcal{E}(\mathcal{L}_{pq}, \mathcal{L}_{pq'})|(n/2-q)|\mathcal{L}_{pq}|} \leq \frac{(q'-q)^{3/2}}{n/2-q}.$$

Since (as in the proof of Lemma A.5) a $1/(q'-q)^{3/2}$ factor of the vertices in a given $\mathcal{C}(T_{pq'r})$ receive $\frac{O(q'-q)^{3/2})}{n/2-q}$ flow from a given neighboring class $\mathcal{L}_{pq}$, the resulting congestion within $\mathcal{L}_{pq'}$ —accounting for the $\sqrt{n/2-q'}$ gain incurred in distributing the flow throughout $\mathcal{L}_{pq'}$—is at most

$$\begin{aligned}
&\sum_{q'-q=1}^{n} \frac{(q'-q)^{3/2}}{(q'-q)^{3/2}(n/2-q)} \cdot \sqrt{n/2-q'} \\
\leq &\sum_{q'-q=1}^{n} \frac{1}{\sqrt{n/2-q}} \\
= &O(\sqrt{n})
\end{aligned}$$

(by Lemma A.6 and the fact that $n/2 - q > n/2 - q'$) and the congestion within $\mathcal{L}_{pq}$ is at most

$$\sum_{q'-q=1}^{n/2-q} \frac{(q'-q)^{3/2}}{(q'-q)^{3/2}(n/2-q)} \cdot \sqrt{n/2-q} = O(\sqrt{n}).$$

□

Lemma A.8. *Within every region $\mathcal{U}_i$, it is possible to route a unit of flow between every ordered pair of triangulations $t, t' \in \mathcal{U}_i$ while producing total congestion $O(\sqrt{n}\log n)$.*

Proof. We need a shuffling step first: let each central-triangle class shuffle via a uniform multicommodity flow, scaled so that each triangulation $t \in \mathcal{C}(T)$ in $\mathcal{U}_i$ sends $\frac{|\mathcal{U}_i|}{|\mathcal{C}(T)|}$ units to each $t' \in \mathcal{C}(T)$. By the natural induction we have been using, this can be done with $O(\sqrt{n/2}\log(n/2))$ congestion. We then have a collection of MSFs, each with source set $\mathcal{C}(T_{pqr})$, for each $\mathcal{C}(T_{pqr})$. Apply Lemma A.5 to solve these MSFs with $O(\sqrt{n}\log n)$ additional congestion.

Finally, we need to solve $n/12$ MSFs, one for each apex class in the region $\mathcal{U}_i$. Each MSF has as its source set an apex class. All apex classes are isomorphic to one another and have cardinality C_n/n; the surplus values are all $|\mathcal{U}_i| = \Theta(1)C_n$, and the sink set for each MSF is all of $\mathcal{U}_i$.

We will use the hierarchical grouping trick once more: just as we grouped together central-triangle classes within a second-vertex class in Lemma A.4, and just as we grouped together second-vertex classes in the proof of Lemma A.5, here we group apex classes first into pairs, then into contiguous sequences (in, say, counterclockwise order according to the apex p) of four, then eight, and so on up to $n/3$.

Crucially, whenever $\mathcal{A}_p, \mathcal{A}_{p'}$ lie in a given $\mathcal{U}_i$ (i.e. $|p' - p| \leq n/12$), a $\Theta(1)$ factor of the triangulations in $\mathcal{A}_p$ lie in classes $\mathcal{C}(T_{pqr})$ having a neighboring class $\mathcal{C}(T_{p'qr})$ in $\mathcal{A}_{p'}$ such that $|\mathcal{E}(T_{pqr}, T_{p'qr})| \geq (p' - p)^{3/2}|\mathcal{C}(T_{pqr})|$.

Thus the hierarchical grouping produces

$$\frac{\Theta(1)C_n|\mathcal{A}_p||\mathcal{A}_{p'}|}{\Theta(1)C_n(j-i)|\mathcal{A}_p||\mathcal{E}(A_p, A_{p'})|} = O\left(\frac{(p'-p)^{3/2}}{j-i}\right) = O(\sqrt{j-i})$$

congestion at the $j - i$ level, and $O(\sqrt{n})$ congestion overall.

The MSF subproblems induced in each $\mathcal{A}_p$ each involve distributing a single commodity with surplus $O(\sqrt{j-i})$ throughout a given class $\mathcal{C}(T_{pqr})$, such that the resulting average flow concentrated in $\mathcal{C}(T_{pqr})$ is $O(1)C_n$, then distributing this flow throughout $\mathcal{A}_p$. There are $O(\log n)$ such subproblems. This produces $O(\sqrt{n}\log^2 n)$ congestion within each $\mathcal{A}_p$. However, we can eliminate a log factor and obtain $O(\sqrt{n}\log n)$ total congestion by applying Lemma A.7 to each of the $O(\log n)$ single-commodity flows within each $\mathcal{A}_p$. □

We have now demonstrated that a flow exists, in the subgraph of K_n induced by the region $\mathcal{U}_i$—in which the additional congestion added in the inductive step is $O(\sqrt{n}\log n)$. We are now ready to prove Theorem 1.1 by way of Lemma A.1, by routing flow among the 24 regions:

Lemma A.1. *Suppose that for all $1 \leq i \leq n/2$, a uniform multicommodity flow exists with congestion $O(\sqrt{i}\log i)$ in K_i. Then a uniform multicommodity flow exists in K_n with congestion $O(\sqrt{n}\log n)$.*

Proof. The first step is to apply Lemma A.8, obtaining a flow f_i within each $\mathcal{U}_i$ in which each pair of triangulations exchanges a unit of flow, and in which each edge carries at most $O(\sqrt{n}\log n)$ congestion.

We do the same for all regions. There is a wrinkle: since some edge classes (and pairs thereof) belong to more than one region, these 24 scaled-up flows result in multiple units of flow being sent between some pairs, as well as a constant-factor increase in congestion. For the former, we simply let pairs in the same class abstain from exchanging flow after the (lexicographically, say) first of the six flows. Clearly, the flows between pairs can never increase the congestion in the network.

For the latter, one may worry that we have lost our "additive advantage" and will now incur a multiplicative penalty in the induction. Fortunately, however, the multiplicative factor is only applied *after* we have applied the inductive hypothesis within each triangular class.

Next, we need to route the $\mathcal{U}_i \to \mathcal{U}_{i+1}$ flow through the triangular classes in the intersection $\mathcal{U}_i \cap \mathcal{U}_{i+1}$. This we accomplish by noting that, by Remark A.3, we can simply concentrate the flow within the intersection between the regions, then send it with $O(1)$ congestion gain. To get from $\mathcal{U}_{i+1}$ to the other 22 classes, we send flow in turn within $\mathcal{U}_{i+1}$, concentrating it on the boundary wtih $\mathcal{U}_{i+2}$, and so on. Upon reaching the destination region, we then distribute the flow in a fashion symmetric to the concentration.

The increases in congestion in this process are all by a constant factor, and crucially, again, these increases are not applied more than once in the induction: our application of the inductive hypothesis occurs only within each *central-triangle* class, and all subsequent routing and redistribution of flow through and within these classes avoids multiplying these factors by the congestion assumed in the inductive hypothesis.

Finally, the overall $O(\sqrt{n}\log n)$ congestion bound claimed now follows from combining the $\log n$ levels of induction with the master theorem. □

Theorem 1.2 is now immediate. A mixing upper bound of $O(n^4 \log^2 n)$ follows from Lemma 2.1; in Section A.2 we will improve this to the $O(n^3 \log^3 n)$ bound claimed in Theorem 1.1.

A.2 Eliminating $\log |V(K_n)|$: mixing time $O(n^3 \log^3 n)$ for triangulations

We have obtained our $O(n^4 \log^2 n)$ bound by showing that the expansion of K_n is $\Omega(1/(\sqrt{n} \log n))$, then applying Lemma 2.1. The loss comes from: (i) normalizing by the degree $\Theta(n)$ of K_n, (ii) squaring the resulting bound per Lemma 2.1, and (iii) multiplying by an additional factor of $\log |V(K_n)| = \Theta(n)$. We show in this section that we can eliminate the $\Theta(n)$ factor in step (iii), obtaining an overall bound of $O(n^3 \log^3 n)$ via a result of Lovász and Kannan:

Lemma A.9. *[50] Given a family of finite, reversible, connected Markov chains $\{\mathcal{M}_n = (\Omega_n, P_n)\}$ parameterized by n, with stationary distribution π, let $\pi_{\min} = \min_{\{t \in \Omega_n\}} \pi(t)$. For all $x \in [1/\pi_{\min}, 1/2]$, define the quantity*

$$\phi(x) = \min_{S:\pi(S) \leq x} \frac{|\partial S|}{\mathrm{vol}(S)},$$

where $\mathrm{vol}(S) = \sum_{t \in S} \frac{\pi(t)}{\Delta}$ is the probability mass of S normalized by the maximum degree Δ of the chain $\mathcal{M}_n$ (viewed as a graph). Then the mixing time of $\mathcal{M}_n$ is at most

$$\tau(n) \leq O(1) \int_{\pi_{\min}}^{1/2} \frac{dx}{(\phi(x))^2 x}.$$

Lemma A.9 implies that in a given flip graph, if small sets have sufficiently larger expansion than large sets, then one can eliminate the $\log |\Omega|$ factor incurred in passing to mixing from squared expansion. This in fact is true for K_n: suppose a set $S \subseteq V(K_n)$ is at most $(C_{n/k})^k/2$, for a given integer $k \in [1, \ldots, n+1]$. S can be partitioned into a collection of subsets of disjoint Cartesian products of the form $K_{i_1} \Box K_{i_2} \Box \cdots \Box K_{i_k}$, where each K_{i_j} is a smaller flip graph with all $i_j \leq \frac{n}{k}$, because of the following fact:

Lemma A.10. *For every integer $1 \leq k \leq n$, every triangulation $t \in V(K_n)$ lies in some Cartesian product of flip graphs $K_{i_1} \Box K_{i_2} \Box \cdots \Box K_{i_k}$, with $i_j \leq \frac{n}{2^{\lfloor \log_3 k \rfloor}}$ for all j.*

Proof. To identify the Cartesian product to which t belongs, partition K_n using the central-triangle partitioning. Each class is a Cartesian product of three smaller flip graphs induced by three smaller polygons; partition each of these classes according to the three central triangles in the three smaller polygons. Repeat this process recursively, in a "breadth-first" fashion, with the triangles placed at a given level in some consistent lexicographic order. . Stop the partitioning after k polygons have been obtained. Now the original n-gon has been partitioned into a collection of smaller polygons, the size of each of which is at most $\max\{1, n/2^{\lfloor \log_3 k \rfloor}\}$. This is because, first, if the recursion depth is d, then the number of "leaf nodes"—polygons at the bottom level of partitioning—is at most 3^d. Second, the breadth-first nature of the partitioning guarantees that each level of partitioning decreases the maximum size of a polygon by at least half, so the largest polygon has size at most $n/2^{\lfloor \log_3 k \rfloor}$.

Now, once the partitioning has stopped, the number of triangulations lying in the resulting partition is at least $(C_{n/k})^k$, because the partition consists of k polygons whose sizes add up to at least n, and because the size of the resulting Cartesian product $C_{l_1} C_{l_2} \cdots C_{l_k}$, $\sum_i l_i \geq n$, is minimized when $l_i = n/k$ for all i. □

The following is now immediate:

Corollary A.2. *For every $S \subseteq V(K_n)$, if $|S| \leq (C_{n/k})^k/2$, for integer $k \in [1, n]$, then $|\partial S|/|S| \geq \Omega(1/((n/2^{\lfloor \log_3 k \rfloor})^{3/2} \log(n/2^{\lfloor \log_3 k \rfloor})))$.*

Proof. The claim follows from noticing that any such set can be partitioned into its intersections with Cartesian products (sets of triangulations) of the form described in Lemma A.10, each of which is at most half full, then noticing that in each such Cartesian product, by Lemma A.10 each graph K_{i_j} in the product has $i_j \leq \frac{n}{2^{\lfloor \log_3 k \rfloor}}$. Appyling Theorem 1.2 then proves the claim. □

We now combine Lemma A.9 with Lemma A.10, then combine Lemma 2.1 with Theorem 1.2 to obtain mixing time $O(n^3 \log^3 n)$ for triangulations, proving Theorem 1.1:

Proof. Proof of Theorem 1.1 We can write, by Lemma A.9,

$$\begin{aligned}
\tau(n) &\leq O(1) \int_{\pi_{\min}}^{1/2} \frac{dx}{(\phi(x))^2 x} \\
&= O(1) \sum_{k=1}^{n} \int_{(C_{n/(k+1)})^{k+1}/C_n}^{(C_{n/k})^k/C_n} O((n/2^{\lfloor \log_3 k \rfloor})^3 \log^2(n/2^{\lfloor \log_3 k \rfloor})) \frac{dx}{x} \\
&\leq O(n^3 \log^2 n) \sum_{k=1}^{n} O((1/2^{\lfloor \log_3 k \rfloor})^3 \int_{(C_{n/(k+1)})^{k+1}/C_n}^{(C_{n/k})^k/C_n} \frac{dx}{x} \\
&= O(n^3 \log^2 n) \sum_{k=1}^{n} O((1/2^{\lfloor \log_3 k \rfloor})^3 \ln\left(\frac{(C_{n/k})^k}{(C_{n/(k+1)})^{k+1}}\right) \\
&= O(n^3 \log^2 n) \sum_{k=1}^{n} O((1/2^{\lfloor \log_3 k \rfloor})^3 \ln(O(n^{3/2})) \\
&= O(n^3 \log^3 n) \sum_{k=1}^{n} O(1/k^{\log_3 8}) \\
&= O(n^3 \log^3 n) \cdot O(1) = O(n^3 \log^3 n).
\end{aligned}$$

□

A.3 Associahedron expansion upper bound

To prove Theorem 1.2, we simply find a sparse cut and apply the definition of expansion. We use the central-triangle partition we used in Appendix 2.7 and Appendix A.1. These are the same classes used to show the $\Omega(n^{3/2})$ mixing lower bound by Molloy, Reed, and Steiger [56].

As we discussed in the introduction, their mixing lower bound does *not* imply the expansion upper bound we give here, but our expansion upper bound does imply their mixing lower bound.

A.3.1 Finding a sparse cut

We will find a cut $(S, \bar{S})$ with $|\partial S|/|S| = O(n^{-1/2})$. We start by partitioning the vertices of the associahedron into central-triangle classes as in Appendix A.1; within any given class, all vertices will be on the same side of the cut. Consider the associahedron K_{n-2} over the n-gon. Draw the regular n-gon in the plane, and label the vertices of the regular n-gon $[0, n-1]$, with 0 as the topmost vertex.

Define the *length* of a diagonal be as in Section A.1, as the (least) number of n-gon edges lying between the two endpoints of the diagonal. Let $\mathcal{C}_l = \{t \in \mathcal{C}(T) | T \text{ has shortest side length} l\}$ be the set of all triangulations whose central triangle's shortest side has length l, $\ l \in [1, n/3]$.

Let $\mathcal{S} = \bigcup_{\mathcal{C}_l | l \in [1,n/6]} \mathcal{C}_l$. Let $\bar{\mathcal{S}} = V(K_{n-2}) \setminus \mathcal{S} = \bigcup_{\mathcal{C}_l | l \in (n/6, n/3]} \mathcal{C}_l$.

Lemma A.11. *The cut $\mathcal{S}$ is indeed a partition of $V(K_{n-2})$, has $|\mathcal{S}| = \Theta(1)|V(K_{n-2})|$ and $|\bar{\mathcal{S}}| = \Theta(1)|V(K_{n-2})|$.*

Proof. Every triangulation lies in exactly one $\mathcal{C}_l$, and that $\mathcal{S}$ and $\bar{\mathcal{S}}$ together partition all of the triangulations. To see that $|\mathcal{S}| = \Theta(1)|V(K_{n-2})| = \Theta(1)C_{n-2}$, we first count the cardinality of each $\mathcal{C}_l, l \in [1, n/6]$. Consider the number of ways to choose a central triangle T so that $\mathcal{C}(T) \subseteq \mathcal{C}_l$, i.e. so that T has shortest side length l. There are n ways to choose the apex of a central triangle (the vertex opposite the shortest side). Conditioned on this choice, and conditioned on a choice of l for the side length opposite the apex, there are l ways to choose the second vertex (the first vertex after the apex in counterclockwise order) so that the center of the n-gon still lies inside the triangle. For all of these choices, the side opposite the apex is indeed shortest. The number of triangulations lying in a class with shortest side length l is $\Theta\left(\frac{1}{n^{3/2}l^{3/2}}\right)C_{n-2}$, and thus the number of triangulations is

$$\sum_{l=1}^{n/6} nl\frac{1}{\Theta(n^{3/2}l^{3/2})} = \Theta\left(\frac{1}{\sqrt{n}}\right)\sum_{l=1}^{n/6}\Theta\left(\frac{1}{\sqrt{l}}\right) = \Theta\left(\frac{\sqrt{n}}{\sqrt{n}}\right) = \Theta(1).$$

Thus $|\mathcal{S}| = \Theta(1)C_{n-2}|$. For $|\bar{\mathcal{S}}|$, notice that for every T with shortest side length $l \in (n/6, n/3]$, there are $\Theta(n^2)$ ways to choose T, by the same argument we used for $l \in [1, n/6]$, and each central-triangle-induced class $\mathcal{C}(T)$ with shortest side $l \in (n/6, n/3]$ has $|\mathcal{C}(T)| = \Theta\left(\frac{1}{n^3}\right)C_{n-2}$. Thus we have the sum

$$\frac{|\bar{\mathcal{S}}|}{C_{n-2}} = \sum_{l=n/6+1}^{n/3}\frac{\Theta(n^2)}{\Theta(n^3)} = \Theta(1).$$

□

Lemma A.12. *The cut* $(\mathcal{S}, \bar{\mathcal{S}})$ *has* $|\partial S|/|S| = O(1/\sqrt{n})$

Proof. In order for a triangulation $t \in \mathcal{C}(T)$, given a central-triangle class T in $\mathcal{C}_l$, $l \in [1, n/6]$, to have a neighbor in $\bar{\mathcal{S}}$, i.e. for t to have a neighboring triangulation $t' \in \mathcal{C}(T')$ with T' having shortest side length $k \geq n/6+1$, the central triangles T and T' must form a quadrilateral in t and in t'. This quadrilateral, in t, consists of T along with a triangle U, where U has shortest side length k. The fraction of triangulations lying in the boundary set $\mathcal{B}_{T'}(T)$ is therefore at most $O\left(\frac{1}{(k-l)^{3/2}}\right)$.

For $l \in [1, n/6]$, let

$$\partial_l S = \{(t,t') \in E(K_{n-2}) | t \in \mathcal{C}_l, t' \in \bar{\mathcal{S}}\} = \bigcup_{\mathcal{C}(T)\subseteq\mathcal{C}_l, \mathcal{C}(T')\subseteq\bar{\mathcal{S}}} \mathcal{E}(T,T')$$

be the set of all cut edges incident to triangulations in $\mathcal{C}_l$.

We will split the sets $\{\mathcal{C}_l\}$ in $\mathcal{S}$ into the cases $l \in [1, n/8]$ and $l \in (n/8, n/6]$. First, for all $l \in [1, n/8]$, by the above reasoning, we have

$$
\begin{aligned}
|\partial_l S|/|\mathcal{C}_l| &= \sum_{k=n/6+1}^{n/3} O\left(\frac{1}{(k-l)^{3/2}}\right) \leq (n/3 - n/6) O\left(\frac{1}{(n/6+1-l)^{3/2}}\right) \\
&\leq (n/3 - n/6) O\left(\frac{1}{(n/6+1-n/8)^{3/2}}\right) = O\left(\frac{n}{n^{3/2}}\right) \\
&= O(1/\sqrt{n}).
\end{aligned}
$$

On the other hand, for $l \in (n/8, n/6]$, we compute the sum

$$
\begin{aligned}
\frac{\sum_{l=n/8+1}^{n/6} |\partial_l S|}{|\mathcal{S}|} &= \sum_{l=n/8+1}^{n/6} \frac{|\partial_l S|}{\Theta(1) C_{n-2}} \leq \sum_{l=n/8+1}^{n/6} |\mathcal{C}_l| \sum_{k=n/6+1}^{n/3} O\left(\frac{1}{(k-l)^{3/2}}\right) \\
&= \int_{l=n/8+1}^{n/6} O\left(\frac{1}{\sqrt{n}\sqrt{l}}\right) \int_{k=n/6+1}^{n/3} O\left(\frac{1}{(k-l)^{3/2}}\right) dk dl,
\end{aligned}
$$

where for the last inequality we have applied the observation from the proof of Lemma A.11 that $|\mathcal{C}_l| = O\left(\frac{1}{\sqrt{n}\sqrt{l}}\right) C_{n-2}$, and have used the fact that asymptotically the summation is equal to a double integral. Evaluating the inner integral we obtain

$$
O\left(\frac{1}{\sqrt{n}}\right) \int_{l=n/8+1}^{n/6} O\left(\frac{1}{\sqrt{l}\sqrt{n/6+1-l}}\right) dl = O\left(\frac{1}{n}\right) \int_{u=1}^{n/6-n/8} O\left(\frac{1}{\sqrt{u}}\right) du,
$$

with the substitution $u = n/6 + 1 - l$ and the observation that when $l \geq n/8$ we have $1/\sqrt{l} = O(1/\sqrt{n})$. Finally, evaluating the integral gives

$$
O\left(\frac{1}{n} \cdot \sqrt{n/6 - n/8}\right) = O(1/\sqrt{n}).
$$

We now have

$$
\begin{aligned}
|\partial S|/|S| &= \sum_{l\in[1,n/6]} |\partial_l S|/|S| \\
&= \sum_{l\in[1,n/8]} |\partial_l S|/|S| + \sum_{l\in[n/8+1,n/6]} |\partial_l S|/|S| = O(1/\sqrt{n}) + O(1/\sqrt{n}) \\
&= O(1/\sqrt{n}),
\end{aligned}
$$

as claimed. □

A.4 Missing details from k-angulation walk proofs

Lemma 2.6. *For every $T, T' \in \mathcal{T}_n$,*

$$|\mathcal{E}^*(T,T')| \geq \frac{|\mathcal{C}^*(T)||\mathcal{C}^*(T')|}{C_n}.$$

Proof. $\mathcal{C}^*(T)$ and $\mathcal{C}^*(T')$ are Cartesian products of the form $\mathcal{C}^*(T) = K_i \square K_{j+k}$ and $T' = K_{i+j}\square K_k$, where $|\mathcal{E}^*(T,T')| = K_i\square K_j \square K_k$. Therefore, $|\mathcal{C}^*(T)| = C_{i-1}C_{j+k-1}$, $|\mathcal{C}^*(T')| = C_{i+j-1}C_{k-1}$, and $|\mathcal{E}^*(T,T')| = C_{i-1}C_{j-1}C_{k-1}$. Thus we have

$$\frac{|\mathcal{C}^*(T)||\mathcal{C}^*(T')|}{|\mathcal{E}^*(T,T')|C_{n-1}} \leq \frac{C_{j+k-1}C_{i+j-1}}{C_{j-1}C_{n-1}}.$$

This ratio increases as j increases, for any fixed i (similarly, for any fixed k). This is because, if i is fixed, maximizing the ratio is equivalent to maximizing

$$\frac{C_{i+j-1}}{C_{j-1}}.$$

It suffices to show that C_{i+j-1}/C_{j-1} increases whenever j increases by one, i.e.

$$\frac{C_{i+j}/C_j}{C_{i+j-1}/C_{j-1}} > 1.$$

I.e., it suffices to show that

$$\frac{C_{i+j}}{C_{i+j-1}} > \frac{C_j}{C_{j-1}},$$

i.e.

$$\frac{i+j}{i+j+1}\frac{(2(i+j))!(i+j-1)!^2}{(2(i+j-1))!(i+j)!^2} > \frac{j}{j+1}\frac{(2j)!(j-1)!^2}{(2(j-1))!j!^2},$$

i.e.

$$\frac{2(i+j)-1}{i+j+1} > \frac{2j-1}{j+1}.$$

The latter inequality clearly holds for all $i \geq 1$.

Therefore, the ratio in the lemma statement is maximized when j is maximized, i.e. $j = n-2$ and $i = k = 1$. Thus we have

$$\frac{|\mathcal{C}^*(T)||\mathcal{C}^*(T')|}{|\mathcal{E}^*(T,T')|C_{n-1}} \leq \frac{C_{n-2}C_{n-2}}{C_{n-3}C_{n-1}}.$$

It is immediate from the definition of Catalan numbers that $C_{n-1}/C_{n-2} \geq C_{n-2}/C_{n-3}$, so this ratio is at most one, and the claim follows. □

Lemma 2.15. *Let $J = G \square H$. Given multicommodity flows g and h in G and H respectively with congestion at most ρ, there exists a multicommodity flow f for J with congestion at most ρ.*

Proof. Let g and h be as stated; we construct f as follows:

1. Within each copy of H in J, construct the flow internally according to h. Similarly, use g internal to each G copy for each pair of vertices within the G copy.

2. Order the copies of H arbitrarily $H_1, \ldots, H_{|V(G)|}$. For each pair of H copies H_r and H_s, $s < r$, and for each vertex $h_r \in H_r, h_s \in H_s$, let the flow from h_r to h_s go through (i) the h flow in H_r from h_r to the counterpart vertex $u \in H_r$ of h_s, then through (ii) the g flow that goes from u to h_s (in the G copy that h_s and u both belong to).

Part 1 generates no additional flow. Part 2 generates at most $|V(H)|$ extra flow through each existing g flow, and at most $|V(G)|$ extra flow through each existing h flow. This results in scaling the amount of g flow using any given edge in a G copy by a factor of $|V(H)|$—while replacing the $\frac{1}{|V(G)|}$ term in the congestion definition by $\frac{1}{|V(J)|} = \frac{1}{|V(G)||V(H)|}$—and similarly scaling the amount of h flow using an edge in an H copy by $|V(G)|$. The result follows. □

Appendix B

Missing details for Glauber dynamics

B.1 Derivation of upper bounds in main theorems

We now analyze the specific polynomial upper bounds that we obtain from each version of the framework.

In the following bounds, we consider all logarithms to be base two, unless otherwise stated. The $\log n$ terms in the exponents of these bounds come from the balanced separators guaranteed by bounded treewidth. Technically, as we have defined balanced separators, one of the two mutually disconnected subgraphs obtained by removing a balanced separator may have size greater than $n/2$. However, one can show [27] that no *connected component* of the resulting disconnected graph has size greater than $n/2$. It is straightforward to modify many of our proofs to account for Cartesian products with multiple factor graphs, by iterating Lemma 2.15. When this is not possible, we will explicitly state the base we use.

We proved the following in Chapter 2:

Lemma 2.10. *Suppose a flip graph $\mathcal{M}_n = (\mathcal{V}_n, \mathcal{E}_n)$ belongs to a family $\mathcal{F}$ of graphs satisfying* *the conditions of Lemma 2.9. Suppose further that every graph $\mathcal{M}_k = (\mathcal{V}_k, \mathcal{E}_k) \in \mathcal{F}$, $k < n$, satisfies*

$$|\mathcal{V}_k|/|\mathcal{E}_{k,\min}| \leq f(k),$$

for some function $f(k)$, where $\mathcal{E}_{k,\min}$ is the smallest edge set between adjacent classes $\mathcal{C}(T), \mathcal{C}(T') \in \mathcal{S}_k$, where $\mathcal{S}_k$ is as in Lemma 2.9. Then the expansion of $\mathcal{M}_n$ is

$$\Omega(1/(2f(n))^{\log n})).$$

In the case of q-colorings, tracing the constant factors in the proof of Lemma 3.12, we see that $N/\mathcal{E}_{\min} \leq q^{2\Delta(t+1)}$, that $\Delta_{\mathcal{M}} \leq (q-1)n$, and that $N \leq q^n$. Combining this with Lemma 2.1 and Lemma 2.8 gives the bound claimed in Theorem 1.6. More precisely, the bound is

$$O((q-1)^2 \log q \cdot n^{4(t+1)\Delta \log q + 5}).$$

Recall Lemma 3.14, which we proved in Section 3.5:

Lemma 3.14. *Suppose a Glauber graph $\mathcal{M}(G)$ satisfies the conditions of the hierarchical framework. Then the mixing time of the corresponding Glauber dynamics is*

$$O(((2(K+1))^{2\log n}) \cdot \Delta_{\mathcal{M}}^2 \log N),$$

where $\Delta_{\mathcal{M}}$ is the maximum degree of the Glauber graph $\mathcal{M}(G)$, $n = |V(G)|$, K is the number of classes in the partition, and $N = |V(\mathcal{M}(G))|$.

The unbiased case of the bound in Theorem 1.5 now follows from combining Lemma 3.14 with the observation that for this chain, $\log N \leq n$, $\Delta_{\mathcal{M}} = n$, and $K \leq 2^{t+1}$. Similarly, as we will see, the bound for the unbiased case of partial q-colorings in Theorem 1.6 will follow from the fact that $K \leq (q+1)^{t+1}$, $\Delta_{\mathcal{M}} \leq qn$, and $N \leq (q+1)^n$, so that

$$(2(K+1))^{2\log n}\Delta_{\mathcal{M}}^2 \log N = n^{2\log 2(K+1)}\Delta_{\mathcal{M}}^2 \log N \leq q^2 \log(q+1) \cdot n^{2(t+2)\log(q+1)+5}.$$

B.2 Deferred Proof Details

Lemma 3.6. *Let G be a graph with bounded treewidth t and bounded degree Δ, let $\mathcal{M}_{\text{IS}}(G)$ be as we have defined, and let $\mathcal{S}_{\text{IS}}(G)$ be as we have defined with respect to a small balanced separator X with $|X| \leq t+1$. For every pair of classes $\mathcal{C}_{\text{IS}}(T), \mathcal{C}_{\text{IS}}(T') \in \mathcal{S}_{\text{IS}}(G)$, $|\mathcal{C}_{\text{IS}}(T)| = \Theta(1)|\mathcal{C}_{\text{IS}}(T')|$.*

Proof. Consider the class $\mathcal{C}_{\text{IS}}(T_r)$ whose vertex set in X is the empty set, and consider any class $\mathcal{C}_{\text{IS}}(T) \neq \mathcal{C}_{\text{IS}}(T_r)$. $\mathcal{C}_{\text{IS}}(T_r)$ consists of the set of all pairs (S_A, S_B), where S_A is an independent set in A, and S_B is an independent set in S_B. $\mathcal{C}_{\text{IS}}(T)$ consists of the set of all pairs (S'_A, S'_B), where S'_A is an independent set in $A \setminus N_A(T)$, and S'_B is an independent set in $B \setminus N_B(T)$.

Clearly every independent set S'_A in $A \setminus N_A(T)$ is also an independent set in A (and the situation is the same for S'_B), so a trivial injective mapping exists from the sets in $\mathcal{C}_{\text{IS}}(T)$ to the sets in $\mathcal{C}_{\text{IS}}(T_r)$. For the other direction, consider the mapping $f : \mathcal{P}(A) \to \mathcal{P}(A \setminus N_A(T))$ that sends every independent set $S_A \subseteq A$ to its restriction $S'_A = S_A \setminus N_A(T)$. Because the degree Δ of G is bounded, $|N_A(T)| \leq t\Delta = O(1)$, and thus f is at worst a $2^{t\Delta} = O(1)$-to-one mapping. This shows that the classes differ in size by a factor of $O(1)$, proving the lemma. □

Lemma 3.7. *Let G be a graph, let $\mathcal{M}_{\text{IS}}(G)$ be as we have defined, and let $\mathcal{S}_{\text{IS}}(G)$ be as we have defined with respect to a separator X. Let $\mathcal{C}_{\text{IS}}(T), \mathcal{C}_{\text{IS}}(T') \in \mathcal{S}_{\text{IS}}(G)$ be two classes. No independent set in $\mathcal{C}_{\text{IS}}(T)$ has more than $O(1)$ flips to independent sets in $\mathcal{C}_{\text{IS}}(T')$.*

Proof. Each edge (S, S') between independent sets $S \in \mathcal{C}_{\text{IS}}(T)$ and $S' \in \mathcal{C}_{\text{IS}}(T') \neq \mathcal{C}_{\text{IS}}(T)$ consists of flipping a single vertex $v \in X$. It is clear that S has no other flips to independent sets in $\mathcal{C}_{\text{IS}}(T')$. □

Lemma 3.8. *Let G be a graph with bounded treewidth t and bounded degree Δ, let $\mathcal{M}_{\text{IS}}(G)$ be as we have defined, and let $\mathcal{S}_{\text{IS}}(G)$ be as we have defined with respect to a small balanced separator X with $|X| \leq t+1$. Let $\mathcal{C}_{\text{IS}}(T), \mathcal{C}_{\text{IS}}(T') \in \mathcal{S}_{\text{IS}}(G)$ be two classes. Suppose there exists at least one flip between an independent set in $\mathcal{C}_{\text{IS}}(T)$ and an independent set in $\mathcal{C}_{\text{IS}}(T')$. Then there exist at least $\Omega(1)|\mathcal{C}_{\text{IS}}(T)|$ flips between independent sets in $\mathcal{C}_{\text{IS}}(T)$ and independent sets in $\mathcal{C}_{\text{IS}}(T')$.*

Proof. T and T' differ by exactly one vertex; call it v. (Or else no flip could exist between $\mathcal{C}_{\text{IS}}(T)$ and $\mathcal{C}_{\text{IS}}(T')$.) Suppose $v \in T$ and $v \notin T'$; then every independent set in $\mathcal{C}_{\text{IS}}(T)$ has a flip to some independent set in $\mathcal{C}_{\text{IS}}(T')$. (See Figure 3.3.) Thus the number of edges from $\mathcal{C}_{\text{IS}}(T)$ to $\mathcal{C}_{\text{IS}}(T')$ is $|\mathcal{C}_{\text{IS}}(T)|$; the lemma now follows from Lemma 3.6. □

Lemma 3.2. *The graph $\mathcal{M}_{\text{MIS}}(G)$ is connected.*

Proof. Let $S \neq S'$ be maximal independent sets, and consider the symmetric difference $S \oplus S'$: if $|S \oplus S'| > 0$, choose some $v \in S' \setminus S$. Obtain a new set S'' by adding v to S and removing all neighbors of v from S, then greedily adding neighbors of neighbors of v until an maximal independent set is obtained. Repeat this process with a new vertex $v' \in S' \setminus S''$, and so on, for every vertex in $S' \setminus S$, obtaining a sequence of sets $S_1 = S, S_2 = S'', S_3, \ldots, S_k$. Crucially, once a vertex v is selected from S' in this process, giving set S_i, we have $v \in S_j$ for all $i \leq j \leq k$. This is because the only way for a vertex to be removed in the process is for one of its neighbors to be selected from S'. However, since S' is an independent set, no neighbor of v belongs to S'.

Thus we have $S_k = S'$, proving that there is a path in $\mathcal{M}_{\text{MIS}}(G)$ between every pair of maximal independent sets. □