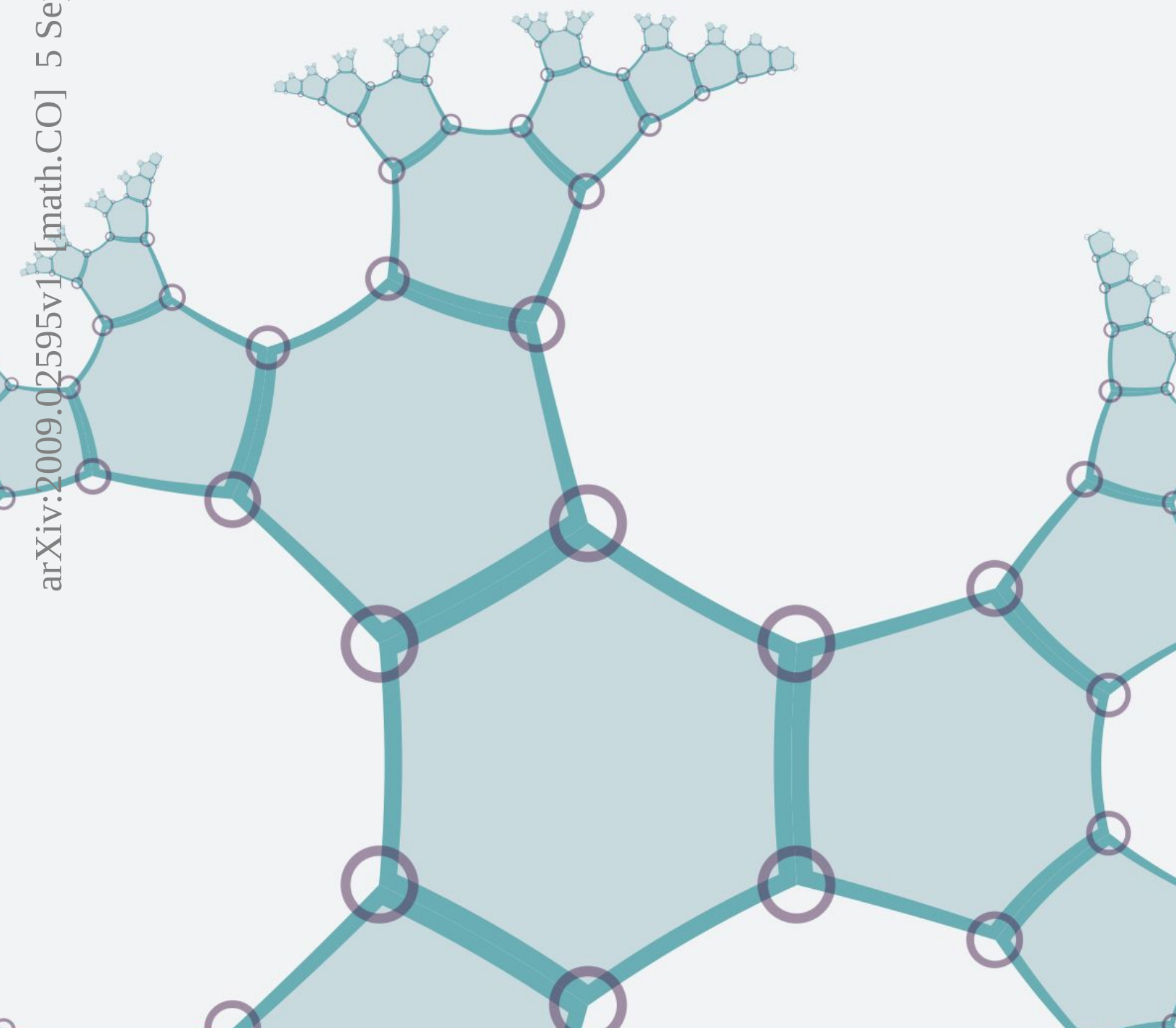
Explicit Near-Fully X-Ramanujan Graphs

Ryan O'Donnell Xinyu Wu

arXiv:2009.02595v1 [math.CO] 5 Sep 2020



Explicit near-fully X-Ramanujan graphs*

Ryan O'Donnell† Xinyu Wu†

September 5, 2020

Abstract

Let $p(Y_1, \dots, Y_d, Z_1, \dots, Z_e)$ be a self-adjoint noncommutative polynomial, with coefficients from $\mathbb{C}^{r\times r}$, in the indeterminates $Y_1, \dots, Y_d$ (considered to be self-adjoint), the indeterminates $Z_1, \dots, Z_e$, and their adjoints $Z_1^*, \dots, Z_e^*$. Suppose $Y_1, \dots, Y_d$ are replaced by independent random $n \times n$ matching matrices, and $Z_1, \dots, Z_e$ are replaced by independent random $n \times n$ permutation matrices. Assuming for simplicity that p's coefficients are 0-1 matrices, the result can be thought of as a kind of random rn-vertex graph $\boldsymbol{G}$. As $n \to \infty$, there will be a natural limiting infinite graph $\mathfrak{X}$ that covers any finite outcome for $\boldsymbol{G}$. A recent landmark result of Bordenave and Collins shows that for any $\varepsilon > 0$, with high probability the spectrum of a random $\boldsymbol{G}$ will be ε-close in Hausdorff distance to the spectrum of $\mathfrak{X}$ (once the suitably defined "trivial" eigenvalues are excluded). We say that $\boldsymbol{G}$ is "ε-near fully $\mathfrak{X}$-Ramanujan".

Our work has two contributions: First we study and clarify the class of infinite graphs $\mathfrak{X}$ that can arise in this way. Second, we derandomize the Bordenave–Collins result: for any $\mathfrak{X}$, we provide explicit, arbitrarily large graphs G that are covered by $\mathfrak{X}$ and that have (nontrivial) spectrum at Hausdorff distance at most ε from that of $\mathfrak{X}$. This significantly generalizes the recent work of Mohanty et al., which provided explicit near-Ramanujan graphs for every degree d (meaning d-regular graphs with all nontrivial eigenvalues bounded in magnitude by $2\sqrt{d-1} + \varepsilon$). To give two simple examples:

- For any $d \geqslant c \geqslant 2$ we obtain explicit arbitrarily large (c, d)-biregular graphs whose spectrum (excluding $0, \pm\sqrt{cd}$) is ε-close in Hausdorff distance to $[-(\sqrt{d-1} + \sqrt{c-1}), -(\sqrt{d-1} - \sqrt{c-1})] \cup [(\sqrt{d-1} - \sqrt{c-1}), (\sqrt{d-1} + \sqrt{c-1})]$.
- We obtain explicit arbitrarily large graphs covered by the modular group — i.e., 3-regular graphs in which every vertex participates in a triangle — whose spectrum (excluding 3) is ε-close in Hausdorff distance to $\{-2\} \cup [\frac{1-\sqrt{13+8\sqrt{2}}}{2}, \frac{1-\sqrt{13-8\sqrt{2}}}{2}] \cup \{0\} \cup [\frac{1+\sqrt{13-8\sqrt{2}}}{2}, \frac{1+\sqrt{13+8\sqrt{2}}}{2}]$.

As an application of our main technical theorem, we are also able to determine the "eigenvalue relaxation value" for a wide class of average-case degree-2 constraint satisfaction problems.

*Or, **ABCDEFG — A***dventures with* **B***ordenave–***C***ollins:* **D***erandomization and* **E***xamples of* **F***un* **G***raphs*

†odonnell@cs.cmu.edu, xinyuwu@cmu.edu. Computer Science Department, Carnegie Mellon University. Author order randomized. Supported by NSF grant CCF-1717606. This material is based upon work supported by the National Science Foundation under grant numbers listed above. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation (NSF).

Contents

1 Introduction **4**
- 1.1 Bordenave and Collins's work . . . 5
- 1.2 X-Ramanujan graphs . . . 8
- 1.3 Our results, and comparison with prior work . . . 10
- 1.4 Implications for degree-2 constraint satisfaction problems . . . 12

2 Setup and definitions **16**
- 2.1 Matrix-weighted graphs . . . 16
- 2.2 Matrix polynomials . . . 18
- 2.3 Lifts of matrix polynomials . . . 19
- 2.4 Projections . . . 21

3 On MPL graphs **22**
- 3.1 Examples of MPL graphs . . . 22
- 3.2 Structure, connectivity and geometry of polynomial lifts . . . 28
- 3.3 Benjamini–Schramm convergence and unimodularity . . . 32
- 3.4 Connectedness of finite lifts produced by our algorithm . . . 34

4 Additional analytic setup and definitions **36**
- 4.1 Nonbacktracking operators . . . 36
- 4.2 Lifting permutations and product lifts . . . 37
- 4.3 Signed permutations . . . 38

5 Derandomization tools **40**

6 Random 2-lifts **42**
- 6.1 A net of bouquets . . . 42
- 6.2 Our main technical theorem . . . 43
- 6.3 Elementary graph theory . . . 45
- 6.4 Encoding and decoding . . . 46
  - 6.4.1 Encoding . . . 47
  - 6.4.2 Decoding . . . 48
- 6.5 Counting . . . 49

7 Weakly derandomizing n-lifts **51**
- 7.1 Derandomizing simple arguments about cycles . . . 51
- 7.2 Weakly derandomizing the main argument . . . 52

8 Constructing explicit good lifts **54**

9 Relating the spectra of the adjacency operator and nonbacktracking operator **56**

10 Explicit lifts of matrix polynomials **64**
- 10.1 Constructing lifts for linear polynomials . . . 64
- 10.2 A linearization trick, and lifts of matrix polynomials . . . 66

A Gallery **75**

1 Introduction

Let G be an n-vertex, d-regular graph. Its adjacency matrix A will always have a "trivial" eigenvalue of d corresponding to the eigenvector $\frac{1}{n}(1,1,\ldots,1)$, the stationary probability distribution for the standard random walk on G. Excluding this eigenvalue, a bound on the magnitude λ of the remaining nontrivial eigenvalues can be very useful; for example, λ can be used to control the mixing time of the random walk on G [Mar08], the maximum cut in G [Fie73], and the error in the Expander Mixing Lemma for G [AC88].

The Alon–Boppana theorem [Alo86] gives a lower bound on how small λ can be, namely $2\sqrt{d-1} - o_{n\to\infty}(1)$. This number $2\sqrt{d-1}$ arises from the spectral radius $\rho(\mathbb{T}_d)$ of the infinite d-regular tree $\mathbb{T}_d$, which is the universal cover for all d-regular graphs ($d \geqslant 3$). Celebrated work of Lubotzky–Phillips–Sarnak [LPS88] and Margulis [Mar88] (see also [Iha66]) shows that for infinitely many d, there exists an explicit infinite family of d-regular graphs satisfying $\lambda \leqslant 2\sqrt{d-1}$. Graphs meeting this bound were dubbed *d-regular Ramanujan graphs*, and subsequent constructions [Chi92, Mor94] gave explicit families of d-regular Ramanujan graphs whenever $d-1$ is a prime power. The fact that these graphs are optimal (spectral) expanders, together with the fact that they are *explicit* (constructible deterministically and efficiently), has made them useful in a variety of application areas in computer science, including coding theory [SS96], cryptography [CFL+18], and derandomization [NN93].

The analysis of LPS/Margulis Ramanujan graphs famously relies on deep results in number theory, and it is still unknown whether infinitely many d-regular Ramanujan exist when $d-1$ is not a prime power. On the other hand, if one is willing to settle for *nearly*-Ramanujan graphs, there is a simple though inexplicit way to construct them for any d and n: Friedman's landmark resolution [Fri08] of Alon's conjecture shows that for any $\varepsilon > 0$, a *random* n-vertex d-regular graph has $\lambda \leqslant 2\sqrt{d-1} + \varepsilon$ with high probability (meaning probability $1 - o_{n\to\infty}(1)$). The proof of Friedman's theorem is also very difficult, although it was notably simplified by Bordenave [Bor19]. The distinction between Ramanujan and nearly-Ramanujan does not seem to pose any problem for applications, but the lack of explicitness does, particularly (of course) for applications to derandomization.

There are several directions in which Friedman's theorem could conjecturally be generalized. One major such direction was conjectured by Friedman himself [Fri03]: that for any fixed base graph K with universal cover tree $\mathfrak{X}$, a random n-lift $\boldsymbol{G}$ of K is nearly "$\mathfrak{X}$-Ramanujan" with high probability. Here the term "$\mathfrak{X}$-Ramanujan" refers to two properties: first, $\mathfrak{X}$ *covers* $\boldsymbol{G}$ in the graph theory sense; second, the "nontrivial" eigenvalues of $\boldsymbol{G}$, namely those not in $\operatorname{spec}(K)$, are bounded in magnitude by the spectral radius $\rho(\mathfrak{X})$ of $\mathfrak{X}$. The modifier "nearly" again refers to relaxing $\rho(\mathfrak{X})$ to $\rho(\mathfrak{X}) + \varepsilon$, here. (We remark that for bipartite K, Marcus, Spielman, and Srivastava [MSS15] showed the existence of an exactly $\mathfrak{X}$-Ramanujan n-lift for every n.) An even stronger version of this conjecture would hold that $\boldsymbol{G}$ is near-*fully* $\mathfrak{X}$-Ramanujan with high probability; by this we mean that for every $\varepsilon > 0$, the nontrivial spectrum of $\boldsymbol{G}$ is ε-close in Hausdorff distance to the spectrum of $\mathfrak{X}$ (i.e., every nontrivial eigenvalue of $\boldsymbol{G}$ is within ε of a point in $\mathfrak{X}$'s spectrum, and vice versa).

This stronger conjecture — and in fact much more — was recently proven by Bordenave and Collins [BC19]. Indeed their work implies that for a wide variety of *non-tree* infinite graphs $\mathfrak{X}$, there is a random-lift method for generating arbitrarily large finite graphs, covered by $\mathfrak{X}$, whose non-trivial spectrum is near-fully $\mathfrak{X}$-Ramanujan. However besides universal cover trees, it is not made clear in [BC19] precisely to which $\mathfrak{X}$'s their results apply.

Our work has two contributions. First, we significantly clarify and partially characterize the class of infinite graphs $\mathfrak{X}$ for which the Bordenave–Collins result can be used; we term these *MPL*

graphs. We establish that all free products of finite vertex-transitive graphs [Zno75] (including Cayley graphs of free products of finite groups), free products of finite rooted graphs [Que94], additive products [MO20], and amalgamated free products [VK19], inter alia, are MPL graphs — but also, that MPL graphs must be unimodular, hyperbolic, and of finite treewidth. The second contribution of our work is to derandomize the Bordenave–Collins result: for every MPL graph $\mathfrak{X}$ and every $\varepsilon > 0$, we give a poly(n)-time deterministic algorithm that outputs a graph on $n' \sim n$ vertices that is covered by $\mathfrak{X}$ and whose nontrivial spectrum is ε-close in Hausdorff distance to that of $\mathfrak{X}$.

1.1 Bordenave and Collins's work

Rather than diving straight into the statement of Bordenave and Collins's main theorem, we will find it helpful to build up to it in stages.

d-regular graphs. Let us return to the most basic case of random n-vertex, d-regular graphs. A natural way to obtain such a graph $\boldsymbol{\mathcal{G}}_n$ (provided n is even) is to independently choose d uniformly random matchings $\boldsymbol{\mathcal{M}}_1, \dots \boldsymbol{\mathcal{M}}_d$ on the same vertex set $V_n = [n] = \{1, 2, \dots, n\}$ and to superimpose them. It will be important for us to remember which edge in $\boldsymbol{\mathcal{G}}_n$ came from which matching, so let us think $\boldsymbol{\mathcal{M}}_1, \dots, \boldsymbol{\mathcal{M}}_d$ as being *colored* with colors $1, \dots, d$. Then $\boldsymbol{\mathcal{G}}_n$ may be thought of as a "color-regular graph"; each vertex is adjacent to a single edge of each color.

Moving to linear algebra, the adjacency matrix $\boldsymbol{A}_n$ for $\boldsymbol{\mathcal{G}}_n$ may be thought of as follows: First, we take the formal polynomial $p(Y_1, \dots, Y_d) = Y_1 + \cdots + Y_d$. Next, we obtain $\boldsymbol{A}_n$ by substituting $Y_j = P_{\sigma_j}$ for each $j \in [d]$, where the σ_j's are independent uniformly random matchings on $[n]$ (i.e., permutations in $\mathfrak{S}(n)$ with all cycles of length 2) and where P_σ denotes the permutation matrix associated to σ.

If we fix a vertex $o \in V_n$ and a number $\ell \in \mathbb{N}$, with high probability the radius-ℓ neighborhood of o in $\boldsymbol{\mathcal{G}}_n$ will look like the radius-ℓ neighborhood of the root of an infinite d-color-regular tree (i.e., the infinite d-regular tree in which each vertex is adjacent to one edge of each color). This tree may be identified with the Cayley graph of the free group $V_\infty = \mathbb{Z}_2 \star Z_2 \star \cdots \mathbb{Z}_2$ with generators $g_1, \dots, g_d$. These generators act as permutations on V_∞ by left-multiplication. Indeed, if one writes P_{g_j} for the associated permutation operator on $\ell_2(V_\infty)$, then the adjacency operator for the Cayley graph is $A_\infty = p(P_{g_1}, \dots, P_{g_d})$.

Bordenave and Collins's generalization of Friedman's theorem may thus be viewed as follows: for $p(Y_1, \dots, Y_d) = Y_1 + \cdots + Y_d$ we have that for any $\varepsilon > 0$, if $\boldsymbol{A}_n = p(P_{\sigma_1}, \dots, P_{\sigma_d})$ and $A_\infty = p(P_{g_1}, \dots, P_{g_d})$, then with high probability the "nontrivial" spectrum of $\boldsymbol{A}_n$ is ε-close in Hausdorff distance to the spectrum of A_∞. Here "nontrivial" refers to excluding $p(1, \dots, 1) = d$.

Weighted color-regular graphs. The Bordenave–Collins theorem is more general than this, however. It also applies to *(edge-)weighted* color-regular graphs. Let $a_1, \dots, a_d \in \mathbb{R}$ be real weights associated with the d colors, and consider the more general linear polynomial $p(Y_1, \dots, Y_d) = a_1 Y_1 + \cdots + a_d Y_d$. Then $\boldsymbol{A}_n = p(P_{\sigma_1}, \dots, P_{\sigma_d})$ is the (weighted) adjacency matrix of a random "color-regular" graph in which each vertex is adjacent to one edge each of colors $1, \dots, d$, with edge-weights $a_1, \dots, a_d$ respectively. Similarly, $A_\infty = p(P_{g_1}, \dots, P_{g_d})$ is the adjacency operator on $\ell_2(V_\infty)$ for the version of the d-color-regular infinite tree in which the edges of color j are weighted by a_j. Again, the Bordenave–Collins result implies that for all $\varepsilon > 0$, with high probability the nontrivial spectrum of $\boldsymbol{A}_n$ (meaning, when $p(1, \dots, 1) = \sum_j a_j$ is excluded) is ε-close in Hausdorff distance to the spectrum of A_∞.

There are several examples where this may be of interest. The first is non-standard random walks on color-regular graphs; for example, taking $a_1 = 1/2$, $a_2 = 1/3$, $a_3 = 1/6$ models random walks where one always "takes the red edge with probability $1/2$, the blue edge with probability $1/3$, and the green edges with probability $1/6$". Another example is the case of $a_1 = \cdots = a_{d/2} = +1$, $a_{d/2+1} = \cdots = a_d = -1$. Here $\boldsymbol{A}_n$ is a d-regular random graphs in which each vertex is adjacent to $d/2$ edges of weight $+1$ and $d/2$ edges of weight -1. This is a natural model for random d-regular instances of the *2XOR* constraint satisfaction problem. Studying the maximum-magnitude eigenvalue of $\boldsymbol{A}_n$ is interesting because it commonly used to efficiently compute an upper bound on the optimal CSP solution (which is NP-hard to find in the worst case); see Section 1.4 for further discussion. Conveniently, the "trivial eigenvalue" of $\boldsymbol{A}_n$ is 0, and the spectrum of A_∞ is easily seen to be identical to that of the d-regular infinite tree, $[-2\sqrt{d-1}, 2\sqrt{d-1}]$. Thus this setting is very similar to that of unweighted random d-regular graphs, but without the annoyance of the eigenvalue of d.

Self-loops and general permutations. The Bordenave–Collins theorem is more general than this, however. Here are two more modest generalizations it allows for. First, one can allow "self-loops" in our template polynomials. In other words, one can generalize to polynomials $p(Y_1, \ldots, Y_d) = a_0\mathbb{1} + a_1 Y_1 + \cdots + a_d Y_d$, where $a_0 \in \mathbb{R}$ and $\mathbb{1}$ can be thought of as a new "indeterminate" which is always substituted with the identity operator (both in the finite case of producing $\boldsymbol{A}_n$ and in the infinite case of A_∞). Second, in addition to having Y_i indeterminates that are substituted with random $n \times n$ matching matrices, one may also allow new indeterminates that are substituted with uniformly random $n \times n$ general permutation matrices. One should be careful to create self-adjoint matrices, i.e. undirected (weighted) graphs, though. To this end, Bordenave and Collins consider polynomials of the form

$$p(Y_1, \ldots, Y_d, Z_1, \ldots, Z_e) = a_0\mathbb{1} + a_1 Y_1 + \cdots + a_d Y_d + a_{d+1} Z_1 + \cdots + a_{d+e} Z_e + a_{d+1}^* Z_1^* + \cdots + a_{d+e}^* Z_e^*. \tag{1}$$

Here $a_0, \ldots, a_d \in \mathbb{R}$, $a_{d+1}, \ldots, a_{d+e} \in \mathbb{C}$, and $Z_1, \ldots, Z_e$ are new indeterminates that in the finite case are always substituted with random $n \times n$ general permutation matrices. We say that the above *polynomial* is "self-adjoint", with the indeterminates $\mathbb{1}, Y_1, \ldots, Y_d$ being treated as self-adjoint. Note that the finite adjacency matrix $\boldsymbol{A}_n = p(P_{\sigma_1}, \ldots, P_{\sigma_d}, P_{\sigma_{d+1}}, \ldots, P_{\sigma_{d+e}})$ that is self-adjoint and hence that represents a (weighted) undirected n-vertex graph. (As a reminder, here $\sigma_1, \ldots, \sigma_d$ are random matching permutations and $\sigma_{d+1}, \ldots, \sigma_{d+e}$ are random general permutations.) As for the infinite case, we extend the notation V_∞ to denote $\mathbb{Z}_2^{\star d} \star \mathbb{Z}^{\star e}$, the free product of d copies of $\mathbb{Z}_2$ and e copies of $\mathbb{Z}$. Then $A_\infty = p(P_{g_1}, \ldots, P_{g_d}, P_{g_{d+1}}, \ldots, P_{g_{d+e}})$, where $g_{d+1}, \ldots, g_{d+e}$ denote the generators of the $\mathbb{Z}$ factors (and note that $P_{g_{d+i}}^* = P_{g_{d+i}}^{-1} = P_{g_{d+i}^{-1}}$).

Matrix coefficients. Now comes one of the more dramatic generalizations: the Bordenave–Collins result also allows for *matrix edge-weights/coefficients*. One motivation for this generalization is that it is needed for the "linearization" trick, discussed below. But another motivation is that it allows the theory to apply to non-regular graphs. The setup now is that for a fixed dimension $r \in \mathbb{N}^+$, we will consider color-regular graphs where each color j is now associated with an edge-weight that may be a *matrix* $a_j \in \mathbb{C}^{r \times r}$. The adjacency matrix of an n-vertex graph with $r \times r$ matrix edge-weights is, naturally, the $n \times n$ block matrix whose (u, v) block is the $r \times r$ weight matrix for edge (u, v). (To be careful here, an undirected edge should be thought of as two opposing directed edges; we insist these directed edges get matrix weights that are adjoints of one another, so as to overall preserve self-adjointness.) In case all edges have the same matrix weight, the

resulting adjacency matrix is just the Kronecker product of the original adjacency matrix and the weight. For example, if P_σ is the adjacency matrix of a matching on $[n]$, and each edge in the matching is assigned the weight

$$a = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \qquad \text{(the adjacency matrix of } \rightarrowtail\text{)},$$

then the resulting matrix-weighted graph has adjacency matrix $P_\sigma \otimes a$, an operator on $\mathbb{C}^n \otimes \mathbb{C}^r$.

Note that a matrix weighted graph's adjacency operator on $\mathbb{C}^n \otimes \mathbb{C}^r$ can simultaneously be viewed as an operator on $\mathbb{C}^{nr}$. In this viewpoint, it is the adjacency matrix of an (uncolored) scalar-weighted nr-vertex graph, which we call the *extension* of the underlying matrix-weighted graph. The situation is particularly simple when the matrix edge-weights are 0-1 matrices; in this case, the extension is an ordinary unweighted graph. In our above example, $P_\sigma \otimes a$ is the adjacency matrix of $n/2$ disjoint copies of the graph formed from C_6 by taking two opposing vertices and hanging a pendant edge on each. Notice that this is a non-regular graph, even though the original matching is regular.

The Bordenave–Collins theorem shows that for any self-adjoint polynomial as in Equation (1), where the coefficients a_j are from $\mathbb{C}^{r\times r}$, we again have that for all $\varepsilon > 0$, the resulting random adjacency operator $\boldsymbol{A}_n$ (on $\mathbb{C}^{nr}$) has its nontrivial spectrum ε-close in Hausdorff distance to that of the operator A_∞ (on $\ell_2(V_\infty \times [r])$). Here the "nontrivial spectrum" refers to the $nr - r$ eigenvalues obtained by removing the eigenvalues of $p(1,\dots,1) = \sum_{j=0}^{d} a_j + \sum_{j=1}^{e}(a_j + a_j^*)$ from the spectrum of $\boldsymbol{A}_n$.

The most notable application of this result is the generalized Friedman conjecture about the spectrum of random lifts of a base graph $K = (R, E)$. That result is obtained by taking $r = |R|$, $e = |E|$, $d = 0$, $a_0 = 0$, and $p = \sum_{(u,v)\in E} a_{uv} Z_{uv}$, where (u, v) denotes a directed edge, a_{uv} is the $r \times r$ matrix that has a single 1 in the (u, v) entry (0's elsewhere), and where Z_{vu} denotes Z_{uv}^* when $v > u$. In this case, the random matrices $\boldsymbol{A}_n$ are adjacency matrices of (extension) graphs that are random n-lifts of K, and the operator $\boldsymbol{A}_\infty$ is the adjacency operator for the universal cover tree $\mathfrak{X}$ of K.

Nonlinear polynomials. We now come to Bordenave and Collins's other dramatic generalization: the polynomials p that serve as "recipes" for producing random finite graphs and their infinite covers need not be linear. Restricting to 0-1 matrix weights but allowing for nonlinear polynomials leads to a wealth of possible infinite graphs $\mathfrak{X}$ (not necessarily trees), which we term *MPL (matrix polynomial lift) graphs*. (Note that even if one ultimately only cares about matrix weights in $\{0,1\}^{r\times r}$, the "linearization" reduction produces linear polynomials with general matrix weights.) An example MPL graph is depicted on our title page; it arises from the polynomial

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \cdot \mathbb{1} + a_{14}Z_1 + a_{41}Z_1^* + a_{36}Z_2 + a_{63}Z_2^* + a_{25}Z_2Z_1 + a_{52}Z_1^*Z_2^*,$$

where again a_{uv} denotes the 6×6 matrix with a 1 in the (u, v) entry.

We may now finally state Bordenave and Collins's main theorem:

Theorem 1.1. *Let p be a self-adjoint noncommutative polynomial with coefficients from $\mathbb{C}^{r\times r}$ in the self-adjoint indeterminates $\mathbb{1}, Y_1, \dots, Y_d$ and the indeterminates $Z_1, \dots, Z_e, Z_1^*, \dots, Z_e^*$. Then for all $\varepsilon, \beta > 0$ and sufficiently large n, the following holds:*

Let $\boldsymbol{A}_n$ be the operator on $\mathbb{C}^n \otimes \mathbb{C}^r$ obtained by substituting the $n \times n$ identity matrix for $\mathbb{1}$, independent random $n \times n$ matching matrices for $Y_1, \dots, Y_d$, and independent random $n \times n$ permutation matrices for $Z_1, \dots, Z_e$. Write $\boldsymbol{A}_{n,\perp}$ for the restriction of $\boldsymbol{A}_n$ to the codimension-r subspace orthogonal to $\mathrm{span}\{(1, \dots, 1)\} \otimes \mathbb{C}^r$. Then except with probability at most β, the spectra $\sigma(\boldsymbol{A}_{n,\perp})$ and $\sigma(A_\infty)$ are at Hausdorff distance at most ε.

Here A_∞ is the operator acting $\ell_2(V_\infty) \otimes \mathbb{C}^r$, where $V_\infty = \mathbb{Z}_2^{\star d} \star \mathbb{Z}^{\star e}$ is the free product of d copies of the group $\mathbb{Z}_2$ and e copies of the group $\mathbb{Z}$, obtained by substituting for $Y_1, \dots, Y_d$ and $Z_1, \dots, Z_e$ the left-regular representations of the generators of V_∞.

As discussed further in Section 1.3, our work derandomizes this theorem by providing explicit (deterministically poly(n)-time computable) $n \times n$ permutation matrices $P_{\sigma_1}, \dots, P_{\sigma_d}$ (matchings), $P_{\sigma_{d+1}}, \dots, P_{\sigma_{d+e}}$ (general), for which the conclusion holds. In fact, our result has the stronger property that for fixed constants d, e, r, k, R, and ε, we construct in deterministic poly(n) time $P_{\sigma_1}, \dots, P_{\sigma_{d+e}}$ that have the desired ε-Hausdorff closeness *simultaneously for all polynomials p* (with degree bounded by k and coefficient matrices bounded in norm by R). A very simple but amusing consequence of this is that for every constant $D \in \mathbb{N}^+$ and $\varepsilon > 0$ we get explicit n-vertex matchings $\mathcal{M}_1, \dots, \mathcal{M}_D$ such that $\mathcal{M}_1 + \mathcal{M}_2 + \cdots + \mathcal{M}_d$ is ε-nearly d-regular Ramanujan for each $d \leqslant D$.

1.2 X-Ramanujan graphs

We would like to now rephrase the Bordenave–Collins result in terms of a new definition of "X-Ramanujan" graphs. Over the years, a number of works have raised the question of how to generalize the classic notion of a d-regular Ramanujan graph to the case of non-regular graphs G; see, e.g., [MO20, Sec. 2.2] for an extended discussion. A natural first possibility is simply to compare the "nontrivial" spectral of G to that of its universal cover tree $\mathfrak{X}$. However this idea is rather limited in scope, particularly because it only pertains to locally tree-like graphs. To illustrate the deficiency, consider the sorts of graph G arising from average-case analysis of constraint satisfaction problems. For example, random regular instances of the 3Sat or NAE-3Sat problem lead one to study graphs G composed of triangles, arranged such that every vertex participates in a fixed number of triangles — say, 4, for a concrete example. Such graphs are 8-regular, so in analyzing G's second-largest eigenvalue one might be tempted to compare it to the Alon–Boppana bound $2\sqrt{7}$, inherited from the 8-regular infinite tree. However as the below theorem of Grigorchuk and Żuk shows, the graph G will in fact always have second-largest eigenvalue at least $1 + 2\sqrt{6} - o(1) > 2\sqrt{7}$. The reason is that the finite triangle graph G is *covered* (in the graph-theoretic sense[1]) by the (non-tree) infinite free product graph $\mathfrak{X} = C_3 \star C_3 \star C_3 \star C_3$, which is known to have spectral radius $\rho(\mathfrak{X}) = 1 + 2\sqrt{6}$.

Theorem 1.2. *([GZ99]'s generalization of the Alon–Boppana bound.) Let $\mathfrak{X}$ be an infinite graph and let $\varepsilon > 0$. Then there exists $c > 0$ such that any n-vertex graph G covered by $\mathfrak{X}$ has at least cn eigenvalues at least $\rho(\mathfrak{X}) - \varepsilon$. (In particular, for large enough n the second-largest eigenvalue of G is at least $\rho(\mathfrak{X}) - \varepsilon$.)*

In light of this, and following [GZ99, Cla07, MO20], we instead take the perspective that the property of "Ramanujan-ness" should derive from the nature of the *infinite graph* $\mathfrak{X}$, rather than that of the finite graph G:

[1] For the complete definition of graph covering, see Definition 2.7.

Definition 1.3 (X-Ramanujan, slightly informal[2])**.** Given an infinite graph $\mathfrak{X}$, we say that finite graph G is $\mathfrak{X}$-*Ramanujan* if:

- $\mathfrak{X}$ covers G;
- the "nontrivial eigenvalues" of G are bounded in magnitude by $\rho(\mathfrak{X})$.

If the bound is relaxed to $\rho(\mathfrak{X}) + \varepsilon$, we say that G is ϵ-*nearly* $\mathfrak{X}$-Ramanujan.

Thus the classic definition of G being a "d-regular Ramanujan graph" is equivalent to being $\mathbb{T}_d$-Ramanujan for $\mathbb{T}_d$ the infinite d-regular tree. It was shown in [MO20] (via non-explicit methods) that for a fairly wide variety of $\mathfrak{X}$, infinitely many $\mathfrak{X}$-Ramanujan graphs exist. This wide variety includes all free products of Cayley graphs, and all "additive products" (see Definition 3.7). Friedman's generalized conjecture (proven by Bordenave–Collins) holds that whenever $\mathfrak{X}$ is the universal cover tree of a base graph K, random lifts of K are ε-nearly $\mathfrak{X}$ with high probability (for any fixed $\varepsilon > 0$).

With this perspective in hand, one can be much more ambitious. Take the earlier example of graphs G where every vertex participates in 4 triangles; i.e., graphs covered by $\mathfrak{X} = C_3 \star C_3 \star C_3 \star C_3$, which is known to have spectrum $\sigma(\mathfrak{X}) = [1 - 2\sqrt{6}, 1 + 2\sqrt{6}]$. The above definition of $\mathfrak{X}$-Ramanujan asks for G's nontrivial eigenvalues to be upper-bounded in magnitude by $1 + 2\sqrt{6}$. But it seems natural to ask if G can also have these eigenvalues bounded *below* by $1 - 2\sqrt{6}$. As another example, it well known that the spectrum of the (c, d)-biregular infinite tree $\mathbb{T}_{c,d}$ with $d > c$ is $[-(\sqrt{d-1} + \sqrt{c-1}), -(\sqrt{d-1} - \sqrt{c-1})] \cup \{0\} \cup [(\sqrt{d-1} - \sqrt{c-1}), (\sqrt{d-1} + \sqrt{c-1})]$, which (when 0 is excluded) contains a notable *gap* between $\pm(\sqrt{d-1} - \sqrt{c-1})$. Are there infinitely many (c, d)-biregular finite graphs G with nontrivial spectrum inside these two intervals? (As far as we are aware, the answer to this question is unknown, but if an ε-tolerance is allowed, the Bordenave–Collins theorem gives a positive answer.) Even further we might ask for (c, d)-biregular G's whose nontrivial spectrum does not have any *other* gaps besides the one between $\pm(\sqrt{d-1} - \sqrt{c-1})$. Of course, since G's spectrum is a finite set this is not strictly possible, but we might ask for it to hold up to an ϵ. Taking these questions to their limit leads to following definition:

Definition 1.4 (Near-fully X-Ramanujan, slightly informal)**.** Given an infinite graph $\mathfrak{X}$ and $\varepsilon > 0$, we say that finite graph G is ε-*near fully* $\mathfrak{X}$-*Ramanujan* if:

- $\mathfrak{X}$ covers G;
- every nontrivial eigenvalue of G is within ϵ of a point $\mathfrak{X}$'s spectrum and vice versa — i.e., the Hausdorff distance of G's nontrivial spectrum to that of $\mathfrak{X}$ is at most ε.

Example 1.5. It is known that for every $\varepsilon > 0$, a sufficiently large d-regular LPS/Ramanujan graph G is ε-near fully $\mathbb{T}_d$-Ramanujan; i.e., every point in $[-2\sqrt{d-1}, 2\sqrt{d-1}]$ is within distance ε of $\mathrm{spec}(G)$.[3] Rather remarkably, is has also been shown this is true of *every* sufficiently large d-regular Ramanujan graph [AGV16].

Bordenave and Collins's Theorem 1.1 precisely implies that for any "MPL graph" — i.e., any infinite graph $\mathfrak{X}$ arising from the "infinite lift" of a $\{0,1\}^r$-edge-weighted polynomial — we can use finite random lifts to (inexplicitly) produce arbitrarily large ε-near fully $\mathfrak{X}$-Ramanujan graphs. Our work, describes in the next section, makes this construction explicit, and significantly characterizes *which* graphs are MPL.

[2]Both of the two bullet points in this definition require a caveat: (i) does "covering" allow for disconnected G? (ii) what exactly counts as a "nontrivial eigenvalue"? These points are addressed at the end of this section.

[3]This fact has been attributed to Serre. [Sar20]

Technical definitional matters: "nontrivial" spectrum and connectedness. As soon as we move away from d-regular graphs, it's no longer particularly clear what the correct definition of "nontrivial spectrum" should be. For example, in Friedman's generalized conjecture about random lifts G of a base graph K with universal covering tree $\mathfrak{X}$, then "nontrivial spectrum" of G is taken to be $\mathrm{spec}(G) \setminus \mathrm{spec}(K)$. But note that this definition is not a function just of $\mathfrak{X}$, since many different base graphs K can have $\mathfrak{X}$ as their universal cover tree. Rather, it depends on the "recipe" by which $\mathfrak{X}$ is realized, namely as the "infinite lift" of K. Taking this as our guide, we will pragmatically define "nontrivial spectrum" only in the context of a specific matrix polynomial p whose infinite lift generates $\mathfrak{X}$; as in Bordenave–Collins's Theorem 1.1, the trivial spectrum is precisely $\mathrm{spec}(p(1,\dots,1))$, the spectrum of the "1-lift of p".

We also need to add a word about connectedness in the context of graph covering. Traditionally, to say that "$\mathfrak{X}$ covers G" one requires that both $\mathfrak{X}$ and G be connected. In the context of the classic "d-regular Ramanujan graph" definition, there are no difficulties because $d \leqslant 2$ is typically excluded; note that for $d = 1$ or 2 we have the random d-regular graphs are surely or almost surely disconnected. However when we move away from trees it does not seem to be a good idea to insist on connectedness. For one, there are many MPL graphs consist of multiple disjoint copies of some infinite graph $\mathfrak{X}$; it seems best to admit $\mathfrak{X}$ as an MPL graph in this case. For two, it's a remarkably delicate question as to when (the extension of) a random n-lift of a matrix polynomial is connected. Fortunately, in most cases $\mathfrak{X}$ is "non-amenable" and this implies that the infinite explicit families of near fully $\mathfrak{X}$-Ramanujan graphs we produce are connected; see Section 3.4 for discussions. Nevertheless, for convenience in this work we will make say (see Definition 2.7) that "$\mathfrak{X}$ covers G" provided each connected component of G is covered by some connected component of $\mathfrak{X}$.

1.3 Our results, and comparison with prior work

The first part of our paper is devoted understanding the class of "MPL graphs". Recall these are defined as follows (cf. Definition 2.21): Suppose the Bordenave–Collins Theorem 1.1 is applied with matrix coefficients $a_j \in \{0,1\}^{r\times r}$. Then the resulting operator A_∞ on $\ell_2(V_\infty) \otimes \mathbb{C}^r$ can be viewed as the adjacency operator of an infinite graph on vertex set $V_\infty \times [r]$. We say that $\mathfrak{X}$ is an MPL graph if (one or more disjoint isomorphic copies of) $\mathfrak{X}$ can be realized in this way.

Our main results concerning MPL graphs are as follows:

- (Section 3.1.) All free products of Cayley graphs of finite groups are MPL graphs. More generally, all "additive products" (as defined in [MO20]) are MPL graphs. Additionally, all "amalgamated free products" (as defined in [VK19]) are MPL graphs. (Furthermore, there are still more MPL graphs that do not appear to fit either category; e.g., the graph depicted on the title page.)

- (Example 3.12) Some zig-zag products and replacement products (as defined in [RVW02]) of finite graphs may be viewed as lifts of matrix-coefficient noncommutative polynomials.

- (Proposition 3.17.) All MPL graphs have finite treewidth. (So, e.g., an infinite grid is not an MPL graph.)

- (Proposition 3.26.) Each connected component of an MPL graph is hyperbolic. (So, e.g., an MPL graph's simple cycles are of bounded length.)

- (Proposition 3.31.) All MPL graphs are unimodular. (So, e.g., Trofimov [Tro85]'s "grandparent graph" is not an MPL graph.)

- (Proposition 3.20.) Given an MPL graph (by its generating polynomial), as well as two vertices, it is efficiently decidable whether or not these vertices are connected.

The remainder of our paper is devoted to derandomizing the Bordenave–Collins Theorem 1.1; i.e., obtaining *explicit* (deterministically polynomial-time computable) arbitrarily large ε-near fully $\mathfrak{X}$-Ramanujan graphs. Thanks to the linearization trick utilized in [BC19], it eventually suffice to derandomize Theorem 1.1 in the case of *linear* polynomials with matrix coefficients. This means that one is effectively seeking ε-near fully $\mathfrak{X}$-Ramanujan graphs for $\mathfrak{X}$ being a (matrix-weighted) color-regular infinite tree.

Our technique is directly inspired by the recent work of Mohanty et al. [MOP20a], which obtained an analogous derandomization of Friedman's theorem, based on Bordenave's proof [Bor19]. Although the underlying idea (dating further back to [BL06]) is the same, the technical details are significantly more complex, in the same way that [BC19] is significantly more complex than [Bor19]. (See the discussion toward the end of [BC19, Sec. 4.1] for more on this comparison.) Some distinctions include the fact that the edge-weights no longer commute, one needs the spectral radius of the nonbacktracking operator to directly arise in the trace method calculations (as opposed to its square-root arising as a proxy for the graph growth rate), and one needs to simultaneously handle a net of all possible matrix edge-weights.

Similar to [MOP20a], our key technical theorem Theorem 6.5 concerns random edge-signings (essentially equivalent to random 2-lifts) of sufficiently "bicycle-free" color-regular graphs. Here bicycle-freeness (also referred to as "tangle-freeness") refers to the following:

Definition 1.6 (Bicycle-free)**.** An undirected multigraph is said to be *λ-bicycle free* provided that the distance-λ neighborhood of every vertex has at most one cycle.

We also use this terminology for an "n-lift' — i.e., a sequence of permutations $\sigma_1, \dots, \sigma_d$ (matchings), $\sigma_{d+1}, \dots, \sigma_{d+e}$ (general permutations) on $V_n = [n]$ — when the multigraph with adjacency matrix $P_{\sigma_1} + \cdots + P_{\sigma_d} + P_{\sigma_{d+1}} + P^*_{\sigma_{d+1}} + \cdots + P_{\sigma_{d+e}} + P^*_{\sigma_{d+e}}$ is λ-bicycle free.

Let us state our key technical theorem in an informal way (for the full statement, see Theorem 6.5):

Theorem 1.7. *(Informal statement.) Let G be an n-vertex color-regular graph with matrix weights $a_1, \dots, a_d, a_{d+1}, \dots, a_{d+e}, a^*_{d+1}, \dots, a^*_{d+e} \in \mathbb{C}^{r \times r}$, and assume G is λ-bicycle free for $\lambda \gg (\log\log n)^2$. Consider a uniformly random edge-signing of G, and let $\boldsymbol{B}_n$ denote the nonbacktracking operator of the result. Then for any $\varepsilon > 0$, with high probability we have $\rho(\boldsymbol{B}_n) \leqslant \rho(B_\infty) + \varepsilon$, where B_∞ denotes the nonbacktracking operator of the color-regular infinite tree with matrix weights $a_1, \dots, a^*_{d+e}$.*

As in [MOP20a], although our proof of Theorem 1.7 is similar to, and inspired by, the proof of the key technical theorem of Bordenave–Collins [BC19, Thm. 17], it does not follow from it in a black-box way; we needed to fashion our own variant of it. Incidentally, this (non-derandomized) theorem on random edge-signings is also needed for our applications to random CSPs; see Section 1.4.

After proving Theorem 1.7, and carefully upgrading it so that the conclusion holds simultaneously for *all* weight sets (a_j) (of bounded norm), the overall derandomization task *is* a straightforward recapitulation of the method from [BL06, MOP20a]. Namely, we first run through the proof of the Bordenave–Collins theorem to establish that $O(\log n)$-wise uniform random permutations are sufficient to derandomize it. Constructing these requires $n^{O(\log n)}$ deterministic time, but we apply them with $n = N_0 = 2^{\Theta(\sqrt{\log N})}$, where N is (roughly) the size of the final graph we wish to construct. Thus in $N_0^{O(\log N_0)} = \mathrm{poly}(N)$ deterministic time we obtain a "good" N_0-lift. We also show that this derandomized N_0-lift will preserve the property

of a truly random lift, that its associated graph is (with high probability) λ-bicycle free for $\lambda = \Theta(\log N_0) = \Theta(\sqrt{\log N}) \gg (\log\log N)^2$. Next we show that $1/\operatorname{poly}(N)$-almost $O(\log N)$-wise uniform random *bit-strings* are sufficient to derandomize Theorem 1.7, and recall that these can be constructed in $\operatorname{poly}(N)$ deterministic time. It then remains to repeatedly apply this 2-lifts arising from the derandomized Theorem 1.7 to obtain an explicit "good" N'-lift, for $N' \sim N_0$. We remark that, as in [BL06, MOP20a], this final N'-lift is not "strongly explicit', although it does have the intermediate property of being "probabilistically strongly explicit" (see Section 8 for details). In the end we obtain the following theorem (informally stated; see Theorem 8.1 for the full statement):

Theorem 1.8. *(Informal statement.) For fixed constants d, e, r, R, and $\varepsilon > 0$, there is a deterministic algorithm that, on input n, runs in $\operatorname{poly}(n)$ time and outputs an n'-vertex unweighted color-regular graph ($n' \sim n$) such that the following holds: For all ways of choosing edge-weights $a_1, \ldots, a_{d+e}^* \in \mathbb{C}^{r\times r}$ for the colors with Frobenius-norm bounds $\|a_j\|_{\mathrm{F}}, \|a_j^{-1}\|_{\mathrm{F}} \leqslant R$, the resulting color-regular graph's nonbacktracking operator $B_{N'}$ has its nontrivial eigenvalues bounded in magnitude by $\rho(B_\infty) + \varepsilon$, where B_∞ denotes the nonbacktracking operator for the analogously weighted color-regular infinite tree.*

At this point, it would seem that we are essentially done, and we need only apply the (non-random) results from [BC19] that let them go from nonbacktracking operator spectral radius bounds for linear polynomials (as in Theorem 1.8) to adjacency operator Hausdorff-closeness for general polynomials (as in Theorem 1.1), taking a little care to make sure the parameters in these reductions only depend on d, e, r, R, ε, and k (the degree of the polynomial), and not on the polynomial coefficients a_j themselves. The tools needed for these reductions include: (i) a version of the Ihara–Bass formula for matrix-weighted (possibly infinite) color-regular graphs, to pass from nonbacktracking operators to adjacency operators ([BC19, Prop. 9, Prop. 10]); (ii) a reduction from bounding the spectral radius of nonbacktracking operators to obtaining Hausdorff closeness for linear polynomials ([BC19, Thm. 12, relying on Prop. 10]); (iii) a way to ensure that this reduction does not blow up the norm of the coefficients a_j involved; (iv) the linearization trick to reduce Hausdorff closeness for general polynomials to that for linear polynomials. Unfortunately, and not to put too fine a point on it, there are bugs in the proof of each of (i), (ii), (iii) in [BC19]. Correcting these is why the remainder of our paper (Sections 9 and 10) still requires new material.[4] In brief, the bug in (i) involves a missing case for the spectrum of non-self-adjoint operators B; the bug in (ii) arises because their reduction converts self-adjoint linear polynomials to non-self-adjoint ones, to which their Theorem 17 does not apply. We fill in the former gap, and derive an alternative reduction for (ii) preserving self-adjointness. Finally, the bug in (iii) seems to require more serious changes. We patched it by first establishing only norm bounds for linear polynomials, and then appealing to an alternative version of Anderson's linearization [And13], namely Pisier's linearization [Pis18], the quantitative ineffectiveness of which required some additional work on our part.

1.4 Implications for degree-2 constraint satisfaction problems

In this section we discuss applications of our main technical theorem on random edge-signings, Theorem 1.7, to the study of constraint satisfaction problems (CSPs). In fact, putting this theorem

[4] After consultation with the authors, we are hopeful they will soon be able to published amended proofs. The bug in (i) is not too serious with multiple ways to fix it. The bug in (ii) is fixed satisfactorily in our work, and the authors of [BC19] outlined to alternate fix involving generalizing [BC19, Thm. 17] to non-self-adjoint polynomials. The bug in (iii) is perhaps the most serious, but the authors may have an alternative patch in mind.

together with the finished Bordenave–Collins theorem implies the following variant (cf. [MOP20a, Thm. 1.15]), whose full statement appears as Theorem 10.10:

Theorem 1.9. *(Informal statement.) In the setting of Theorem 1.1, if $\boldsymbol{A}_n$ is the operator produced by substituting random ± 1-signed permutations into the matrix polynomial p, then the Hausdorff distance conclusion holds for the full spectrum $\sigma(\boldsymbol{A}_n)$ vis-a-vis $\sigma(A_\infty)$; i.e., we do not have to remove any "trivial eigenvalues" from $\boldsymbol{A}_n$.*

This theorem will allow us to determine the "eigenvalue relaxation value" for a wide class of average-case CSPs. Roughly speaking, we consider random regular instances of Boolean valued CSPs where the constraints are expressible as degree-2 polynomials (with no linear term). Our work determines the typical eigenvalue relaxation bound for these CSPs; recall that this is a natural, efficiently-computable upper bound on the optimum value of Boolean quadratic programs (and on the SDP/quantum relaxation value). This generalizes previous work [MS16, DMO$^+$19, MOP20b] on random Max-Cut, NAE3-Sat, and "2-eigenvalue 2XOR-like CSPs", respectively. We remark again that our results here do not require the derandomization aspect of our work, but they *do* rely on Theorem 1.9 concerning random *signed* lifts, which is not derivable in a black-box fashion from the work of Bordenave–Collins.

We will be concerned throughout this section with Boolean CSPs: optimization problems over a Boolean domain, which we take to be $\{\pm 1\}$ (equivalent to $\{0,1\}$ or $\{\mathsf{True}, \mathsf{False}\}$). The hallmark of a CSP is that it is defined by a collection of *local* constraints of similar type. Our work is also general enough to handle certain *valued CSPs*, meaning ones where the constraints are not simply predicates (which are satisfied/unsatisfied) but are real-valued functions. These may be thought of giving as "score" for each assignment to the variables in the constraint's scope.

Definition 1.10 (Degree-2 valued CSPs)**.** A *Boolean valued CSP* is defined by a set Ψ of *constraint types* ψ. Each ψ is a function $\psi : \{\pm 1\}^r \to \mathbb{R}$, where r is the *arity*. We will say such a CSP is *degree*-2 if each ψ can be represented as a degree-2 polynomial, with no linear terms, in its inputs.

An *instance* $\mathcal{I}$ of such a CSP is defined by a set of n variables V and a list of m constraints $C = (\psi, S)$, where each $\psi \in \Psi$ and each S is an r-tuple of distinct variables from V. More generally, in an instance with *literals allowed*, each constraint is of the form $C = (\psi, S, \ell)$, where $\ell \in \{\pm 1\}^r$.

The computational task associated with $\mathcal{I}$ is to determine the value of the optimal *assignment* $x : V \to \{\pm 1\}$; i.e.,

$$\mathrm{Opt}(\mathcal{I}) = \max_{x:V\to\{\pm 1\}} \mathrm{obj}(x), \qquad \text{where } \mathrm{obj}(x) = \sum_{C=(\psi,S,\ell)\in\mathcal{I}} \psi(\ell_1 x(S_1), \dots, \ell_r x(S_r)). \tag{2}$$

Note that for a degree-2 CSP, the objective $\mathrm{obj}(x)$ may be considered as a degree-2 homogeneous polynomial (plus a constant term) over the Boolean cube $\{\pm 1\}^n$.

Example 1.11. Let us give several examples where Ψ contains just a single constraint ψ.

When $r = 2$ and $\psi(x_1, x_2) = \frac{1}{2} - \frac{1}{2}x_1x_2$ we obtain the Max-Cut CSP. If we furthermore allow literals here, we obtain the 2XOR CSP. If literals are disallowed, and ψ is changed to $\psi(x_1, x_2, x_3, x_4) = \frac{1}{2} - \frac{1}{2}x_1x_2 + \frac{1}{2} - \frac{1}{2}x_3x_4$, we get a version of the 2XOR CSP in which instances must have an equal number of "equality" and "unequality" constraints.

When $r = 3$ and $\psi(x_1, x_2, x_3) = \frac{3}{4} - \frac{1}{4}(x_1x_2 + x_1x_3 + x_2x_3)$, the CSP becomes 2-coloring a 3-uniform hypergraph; if literals are allowed here, the CSP is known as NAE-3Sat.

The case of the predicate $\psi(x_1, x_2, x_3, x_4) = \frac{1}{2} + \frac{1}{4}(x_1x_2 + x_2x_3 + x_3x_4 - x_1x_4)$ with literals allowed yields the Sort_4 CSP; the predicate here is equivalent to "CHSH game" from quantum mechanics [CHSH69].

Remark 1.12. As additive constants are irrelevant for the task of optimization, we will henceforth assume without loss of generality that degree-2 CSPs involve *homogeneous* degree-2 polynomial constraints.

Definition 1.13 (Instance graph)**.** Given an instance $\mathcal{I}$ of a degree-2 CSP as above, we may associate an *instance graph* G, with adjacency matrix A. This G is the undirected, edge-weighted graph with vertex set $V = [n]$ where, for each nonzero monomial $w_{ij}x_ix_j$ appearing in $\mathrm{obj}(x)$ from Equation (2), G contains edge $\{i, j\}$ with weight $\frac{1}{2}w_{ij}$. The factor of $\frac{1}{2}$ is included so that given an assignment $x \in \{\pm 1\}^n$, we have $\mathrm{obj}(x) = x^\top Ax = \langle A, xx^\top \rangle$, where $\langle A, B \rangle$ denotes the matrix inner product.

For almost all Boolean CSPs, exactly solving the quadratic program to determine Opt is NP-hard; hence computationally tractable relaxations of the problem are interesting. Perhaps the simplest and most natural such relaxation is the *eigenvalue bound* — the generalization of the Fiedler bound [Fie73] for Max-Cut:

Definition 1.14 (Eigenvalue bound)**.** Let $\mathcal{I}$ be a degree-2 CSP instance and let A be the adjacency matrix of its instance graph. The *eigenvalue bound* is

$$\mathrm{Eig}(\mathcal{I}) = \max_{\substack{x \in \ell_2(V) \\ \|x\|_2^2 = n}} \langle A, xx^\top \rangle = \max_{\substack{X \in \mathbb{R}^{n \times n} \text{ PSD} \\ \mathrm{tr}(X) = n}} \langle A, X \rangle = n \cdot \lambda_{\max}(A).$$

It is clear that

$$\mathrm{Opt}(\mathcal{I}) \leqslant \mathrm{Eig}(\mathcal{I})$$

always, and $\mathrm{Eig}(\mathcal{I})$ can be computed (to arbitrary precision) in polynomial time. For many average-case optimization instances, this kind of spectral certificate provides the best known efficiently-computable bound on the instance's optimal value; it is therefore of great interest to characterize $\mathrm{Eig}(\mathcal{I})$ for random instances. See, e.g., [MOP20b] for further discussion. As we will describe below, for a natural model of random instances $\boldsymbol{\mathcal{I}}$ of a degree-2 CSP, our Theorem 1.9 allows us to determine the typical value of $\mathrm{Eig}(\boldsymbol{\mathcal{I}})$ for large instances. Often one can show this exceeds the typical value of $\mathrm{Opt}(\boldsymbol{\mathcal{I}})$ for these random instances, thus leading to a potential *information-computation gap* for the certification task.

We should also mention another efficiently-computable upper bound on $\mathrm{Opt}(\mathcal{I})$:

Definition 1.15 (SDP bound)**.** Let $\mathcal{I}$ be a degree-2 CSP instance and let A be the adjacency matrix of its instance graph. The *SDP bound* is

$$\mathrm{SDP}(\mathcal{I}) = \max_{\substack{X \in \mathbb{R}^{n \times n} \text{ PSD} \\ X_{ii} = 1 \ \forall i}} \langle A, X \rangle.$$

This quantity can also be computed (to arbitrary precision) in polynomial time, and it is easy to see that $\mathrm{Opt}(\mathcal{I}) \leqslant \mathrm{SDP}(\mathcal{I}) \leqslant \mathrm{Eig}(\mathcal{I})$ always. Thus the SDP value can only be a better efficiently-computable upper bound on $\mathrm{Opt}(\mathcal{I})$, and it would be of interest also to characterize its typical value for random degree-2 CSPs, as was done in [MS16, DMO+19, MOP20b]. Proving that $\mathrm{SDP}(\boldsymbol{\mathcal{I}}) \sim \mathrm{Opt}(\boldsymbol{\mathcal{I}})$ with high probability (as happened in those previous works) seems to require that the CSP has certain symmetry properties that do not hold in our present very general setting. We leave investigation of this to future work.

We now describe a model of random degree-2 CSPs for which Bordenave–Collins's Theorem 1.1 and our Theorem 1.9 lets one determine the (high probability) value of the eigenvalue relaxation bound. It is based on the "additive lifts" construction from [MOP20b].

Suppose we have a degree-2 Boolean CSP with constraints $\Psi = \{\psi_1, \ldots, \psi_t\}$ of arity r. We wish to create random "constraint-regular" instances defined by numbers c_{jk} ($1 \leqslant j \leqslant r$, $1 \leqslant k \leqslant t$) in which each variable appears in the jth position of c_{jk} constraints of type ψ_k. Let $c = \sum_{jk} c_{jk}$, and for notational simplicity let $\mathcal{I}_0 = (\varphi_1, \ldots, \varphi_c)$ denote a minimal such CSP instance on a set $[r]$ of variables, where each φ_i stands for some ψ_k with permuted variables, and every scope is considered to be $(1, \ldots, r)$. For each constraint φ_i, we associate an *atom* graph A_i, which is the instance graph on vertex set $[r]$ defined by the single constraint φ_i applied to the r variables.

Now given $n \in \mathbb{N}^+$, we will construct a random CSP on nr variables with nc constraints as a "random lift". We begin with a *base graph*: the complete bipartite graph $K_{r,c}$, with the r vertices in one part representing the variables, and the c vertices in the other part representing the constraints. We call any n-lift of this base graph a *constraint graph*; we can view it as an instance of the CSP where the edges encode which variables participate in which constraints for the random CSP. When we do a random n-lift of this $K_{r,c}$, we create r groups of n variables each, and c groups of n constraints, with a random matching being placed between every group of variables and every group of constraints. In this random bipartite graph, each variable participates in exactly c constraints, one in each group, while each constraint gets exactly one variable from each group.

To obtain the instance graph from the constraint graph, we start with an empty graph with nr vertices corresponding to the variables, and then, for each constraint vertex in the constraint graph from some group j, we place a copy of the atom A_j on the r vertices to which the constraint vertex is adjacent.

With this lift-based model for generating random constraint-regular CSP instances, it is important to allow for *random literals* in the lifted instance. (Otherwise one may easily generate trivially satisfiable CSPs. Consider, for example, random NAE-3Sat instances., as in [DMO+19]. Without literals, all lifted instances will be trivially satisfiable due to the partitioned structure of the $3n$ variables: one can just assigning two of the three n-variable groups the label $+1$ and the other n-variable group the label -1. Hence the need for random literals.) Thus instead of doing a random n-lift of the base $K_{r,c}$ graph, we do a random *signed* n-lift, placing a random ± 1 sign on each edge in the lifted random graph. Then, if u and v are variables and they are connected to a constraint vertex y in the constraint graph, then, in the instance graph the weight of the edge $\{u, v\}$ (if it exists in the atom) will pick up an additional sign of $\text{sign}(\{u, y\}) \cdot \text{sign}(\{v, y\})$. This has the effect of uniformly randomly negating a variable when a predicate is applied to it.

Next, we will see how this model of constructing a random regular CSP is equivalent to a random lift of a particular matrix polynomial. The coefficients of the polynomial will be in $\mathbb{C}^{r \times r}$, indexed by the r variables which the atoms act on. The rc indeterminates in the polynomial are indexed by the edges of the base constraint graph; i.e., we have one (non-self-adjoint) indeterminate $X_{u,j}$ for each variable u and each constraint j. We construct the polynomial p iteratively: for each constraint j, and every pair of variables u and v, if $\{u, v\}$ is an edge in in A_j then we add the terms

$$\frac{1}{2} w_{vu} \, |v\rangle\langle u| \, X_{v,j} X_{u,j}^* + \frac{1}{2} w_{uv} \, |u\rangle\langle v| \, X_{u,j} X_{v,j}^*$$

to p[5]. Then, a random signed lift fits precisely into the model of our Theorem 1.9, and we conclude that for this model of random regular general degree-2 CSPs, the eigenvalue relaxation

[5]Note that this is the same construction as with the additive lifts defined in Example 3.6

bound is, with high probability, $n \cdot (\lambda_{\max(A_\infty)} \pm \varepsilon)$ for any $\varepsilon > 0$.

2 Setup and definitions

The general framework of color-regular matrix-weighted graphs and polynomials lifts was introduced in [BC19]; however we will add some additional terminology. We will also make use of Dirac bra-ket notation.

2.1 Matrix-weighted graphs

Definition 2.1 (Matrix-weighted graph)**.** Let $G = (V, E)$ be a directed multigraph, with V countable and E locally finite. We say that G is *matrix-weighted* if, for some $r \in \mathbb{N}^+$, each directed edge $e \in E$ is given an associated nonzero "weight" $a_e \in \mathbb{C}^{r\times r}$. We say that G is an *undirected* matrix-weighted graph if (with a minor exception) its directed edges are partitioned into pairs e and e^*, where e^* is the reverse of e and where $a_{e^*} = a_e^*$. We call each such pair an undirected edge. The minor exception is that we allow any subset of the self-loops in E to be unpaired, provided each unpaired self-loop e has a self-adjoint weight, $a_e^* = a_e$. (In the terminology of Friedman [Fri93], such self-loops are called "half-loops", in contrast with paired self loops which are called "whole-loops".) The adjacency operator A for G, acting on $\ell_2(V) \otimes \mathbb{C}^r$, is given by

$$\sum_{e=(u,v)\in E} |v\rangle\langle u| \otimes a_e.$$

It can be helpful to think to think of A in matrix form, as an $|V| \times |V|$ matrix whose entries are themselves $r \times r$ edge-weight matrices. Note that if G is undirected then A will be self-adjoint, $A = A^*$.

Definition 2.2 (Extension of a matrix-weighted graph)**.** Let $G = (V, E)$ be a matrix-weighted graph. We will write $\widetilde{G}$ for the *extension* of G, the scalar-weighted multigraph on vertex set $V \times [r]$ formed as follows: for each $e = (u, v) \in E$ with weight $a_e = \sum_{i,j=1}^r c_{ij} |i\rangle\langle j|$, we include into $\widetilde{G}$ the edge $((u, i), (v, j))$ with scalar weight c_{ij} for all i, j such that $c_{ij} \neq 0$. The scalar-weighted graph $\widetilde{G}$ will be undirected when G is, and the two graphs have the same adjacency operator when $\ell_2(V) \otimes \mathbb{C}^r$ is identified with $\ell_2(V \times [r])$. We will be most interested in the case when G is undirected with weight matrices $a_e \in \{0, 1\}^{r\times r}$; in this case, the extension $\widetilde{G}$ will be an ordinary unweighted, undirected (multi)graph.

Definition 2.3 (Index/color set)**.** Given parameters $d, e \in \mathbb{N}$, the associated *index set*, or *color set*, is defined to be $\mathscr{I} = \{0, 1, 2, \ldots, d, d + 1, \ldots, d + 2e\}$, together with the involution $* : \mathscr{I} \to \mathscr{I}$ that has $j^* = j$ for $j \leqslant d$ and $j^* = j + e$ for $d < j \leqslant d + e$. Index 0 is called the *identity-index*, indices $1, \ldots, d$ are called the *matching-indices*, and $d + 1, \ldots, d + 2e$ are called the *permutation-indices*. We also refer to indices as "colors". Finally, we sometimes allow an index set to not include an identity-index.

Definition 2.4 (Color-regular graph)**.** Fix an index set $\mathscr{I}$ and a sequence $(a_j)_{j\in\mathscr{I}}$ of nonzero matrices in $\mathbb{C}^{r\times r}$ with $a_{j^*} = a_j^*$. We say that an unweighted matrix-weighted graph $\mathcal{G}$ is *color-regular with weights* $(a_j)_{j\in\mathscr{I}}$ if each directed edge has an associated color from the set $\mathscr{I}$, and the following properties hold:

- Each vertex v has exactly one outgoing edge of each color.

- If $0 \in \mathscr{I}$, then each edge colored 0 is a half-loop.
- If (e, e^*) is an undirected edge, and e is colored j, then e^* is colored j^*.
- The weight of every j-colored edge is a_j.

We will find it convenient to make the following definition.

Definition 2.5 (Matrix bouquet)**.** A *(matrix) bouquet* $\mathcal{K}$ on index set $\mathscr{I}$ is a color-regular graph on one vertex. Note that $\mathcal{K}$ is completely specified by the matrices $(a_j)_{j \in \mathscr{I}}$.

Figure 1 gives some examples of matrix bouquets.

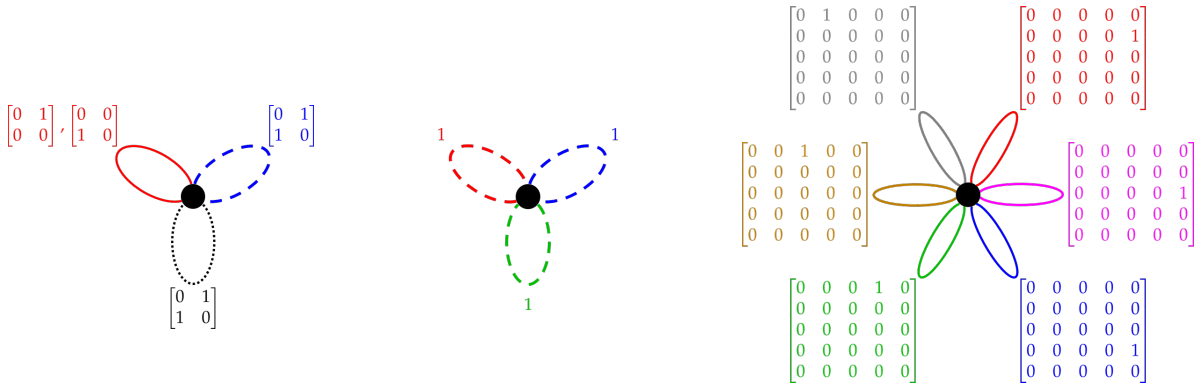


Figure 1: Example matrix bouquets. In each, identity-indices are pictured dotted, matching-indices are depicted dashed, and permutation-indices are depicted solid. All matrix edge-weights a_i happen to have 0-1 entries, though this is not in general necessary. In the left bouquet, $d = 1$, $e = 1$, $r = 2$, and a_0, a_1, a_2, a_3 are $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. In the middle bouquet, $d = 3$, $e = 0$, $r = 1$, and a_1, a_2, a_3 are $1, 1, 1$ (there is no identity-index in this example). In the right bouquet, $d = 0$, $e = 6$, $r = 5$, there is no identity-index, and we have only depicted one of the a, a^* pairs on each whole-loop.

Definition 2.6 (The bouquet of an ordinary graph)**.** Let $G = (V, E)$ be an ordinary unweighted, undirected multigraph (for simplicity, without half-loops), having $|V| = r$ vertices and $|E| = e$ edges. We define the associated *bouquet of* G to be the matrix bouquet $\mathcal{K}_G$ with no identity-index, $d = 0$, and $a_1, \ldots, a_{2e}$ being "0-1 indicator matrices" for the directed edges of G. In other words, if the ith edge in E is $\{u, v\}$, then $a_i = |v\rangle\langle u|$ and $a_{i+e} = |u\rangle\langle v|$. Another description is that $\mathcal{K}_G$ is the one-vertex color-regular graph with 0-1 matrix-weights whose extension, $\widetilde{\mathcal{K}}_G$, is G. We remark that $\sum_{i \in \mathscr{I}} a_i$ is the adjacency matrix of G.

The third bouquet in Figure 1 is an example of this construction; it is the bouquet $\mathcal{K}_G$ associated to the graph G in Figure 2 (whose vertices have been numbered and whose edges have been colored, for illustrative purposes). This example also illustrates the general point that if $\widetilde{\mathcal{G}}$ is the extension of a color-regular graph $\mathcal{G}$ with 0-1 matrix-weights, the ordinary graph $\widetilde{\mathcal{G}}$ need not be regular.

We now recall the standard definitions of multigraph coverings (see, e.g., [Mak15, Sec. 2.3]). We include certain extensions to allow for weighted graphs and possibly-disconnected graphs.

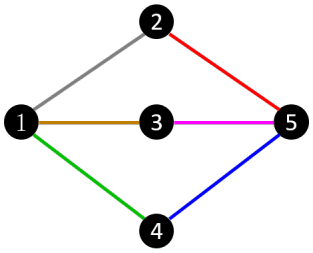


Figure 2: An undirected graph H with 5 vertices and 6 edges; its associated matrix bouquet is the third one depicted in Figure 1.

Definition 2.7 (Matrix-weighted graph covering)**.** Let X and G be connected undirected matrix-weighted graphs. A *homomorphism* from X to G is a pair of maps $f_V : V(X) \to V(G)$ and $f_E : E(X) \to E(G)$ such that if $e = (u, v)$ is an edge in X with weight a, then $f_E(e)$ is $(f_V(u), f_V(v))$ and has weight a in G. We say such a homomorphism is a *covering* (and we say that X *covers* G) if:

- the preimage under f_E of every undirected edge (i.e., pair of opposing edges e, e^*) in $E(G)$ is a collection of undirected edges in X;
- for every vertex $v \in V(X)$, out-edges are mapped bijectively by f_E to the out-edges of $f_V(v)$, and similarly for v's in-edges.

We extend this definition to allow X and/or G to be disconnected; in this case we stipulate that X covers G provided each connected component of G is covered by some connected component of X and, vice versa, each connected component of X covers some connected component of G.

We make the following observations:

Fact 2.8. *Let $\mathcal{K} = (a_j)_{j \in \mathscr{J}}$ be a matrix bouquet, and let $\mathcal{G}$ be a color-regular graph with weights $(a_j)_{j \in \mathscr{J}}$. Then $\mathcal{G}$ covers $\mathcal{K}$.*

Fact 2.9. *Suppose X and G are undirected matrix-weighted graphs with X covering G. Then $\widetilde{X}$ covers $\widetilde{G}$.*

2.2 Matrix polynomials

We now define *matrix polynomials*, which may be thought of as recipes for producing for color-regular graphs.

Definition 2.10 (Unreduced matrix polynomial)**.** Given an index set $\mathscr{J}$ (necessarily including the identity-index 0) and a dimension $r \in \mathbb{N}^+$, an *(unreduced) matrix polynomial* is a noncommutative polynomial over indeterminates $(X_j)_{j \in \mathscr{J}}$ with coefficients from $\mathbb{C}^{r \times r}$; we explicitly disallow the empty monomial. More formally, the matrix polynomials are the free left $\mathbb{C}^{r \times r}$-module with basis given by all words W of positive length over the alphabet $(X_j)_{j \in \mathscr{J}}$. We write X^W for the monomial associated to word W.

We make the matrix polynomials into a noncommutative ring by specifying that $(a_{W_1} X^{W_1}) \times (a_{W_2} X^{W_2}) = a_{W_1} a_{W_2} X^{W_1 W_2}$, where $W_1 W_2$ denotes concatenation. This ring does not have a multiplicative identity, but see Definition 2.12 below.

We further make the matrix polynomials into a $*$-ring by specifying that $(a X_{j_1} \cdots X_{j_k})^* = a^* X_{j_k}^* \cdots X_{j_1}^*$, where X_j^* is a synonym for X_{j^*} (recalling the involution associated to $\mathscr{J}$).

We will be mainly interested in polynomials p that are *self-adjoint*, meaning $p^* = p$. We will also be particularly interested in *linear* matrix polynomials p, meaning $p = \sum_{j \in \mathscr{J}} a_j X_j$.

Definition 2.11 (Evaluating a matrix polynomial)**.** We will be "evaluating" matrix polynomials at bounded operators on a complex Hilbert space $\mathcal{H}$. Given bounded operators $(U_j)_{j \in \mathscr{J}}$, the evaluation of monomial $aX_{j_1} \cdots X_{j_k}$ is defined to be $U_{j_1} \cdots U_{j_k} \otimes a$, an operator $\mathcal{H} \otimes \mathbb{C}^r$.[6] Furthermore, in this paper we will *only* ever evaluate matrix polynomials at bounded operators satisfying the following conditions:

- $U_0 = \mathbb{1}$, the identity operator on $\mathcal{H}$.
- $U_{j*} = U_j^*$.
- All U_j's are unitary, $U_j^* = U_j^{-1}$.

In light of the first two conditions above, we need not explicitly specify $U_0, U_{d+e+1}, \ldots, U_{d+2e}$, and hence may just write $p(U_1, \ldots, U_d, U_{d+1}, \ldots, U_{d+e})$ for a polynomial evaluation. Also note that $p^*(U_1, \ldots, U_{d+e}) = p(U_1, \ldots, U_{d+e})^*$, and hence the evaluation of any self-adjoint polynomial will be a self-adjoint operator.

Definition 2.12 (Reduced matrix polynomial)**.** Given the restrictions on evaluations we imposed in Definition 2.11, we may somewhat simplify the ring of matrix polynomials with which we work. First, we will let $\mathbb{1}$ be a synonym for the indeterminate X_0. Second, we may take $X_j X_j^* = X_j^* X_j = \mathbb{1}$ and $\mathbb{1} X_j = X_j \mathbb{1} = X_j$ as "relations", resulting in a quotient ring which we term the *(reduced) matrix polynomials*. Here monomials correspond to "reduced words" in $\mathbb{Z}_2^{\star d} \star \mathbb{Z}^{\star e}$, the free product of d copies of the group $\mathbb{Z}_2$ and e copies of the group $\mathbb{Z}$. The empty reduced word corresponds to the monomial $\mathbb{1}$, and we will sometimes abbreviate the monomial $a_0 \mathbb{1}$ just as a_0. Notice that if this a_0 is itself the identity operator $\mathbb{1}_{r \times r}$, then the monomial becomes a multiplicative identity for the quotient ring.

Remark 2.13. In the remainder of this work, a "matrix polynomial" will mean a reduced matrix polynomial unless otherwise specified. We will write $p = \sum_w a_w X^w$ for a generic such polynomial, where a_w is nonzero for only finitely many $w \in \mathbb{Z}_2^{\star d} \star \mathbb{Z}^{\star e}$.

2.3 Lifts of matrix polynomials

Definition 2.14 (n-lift)**.** Fix an index set $\mathscr{J} = \{0, 1, \ldots, d + 2e\}$. Let $n \in \mathbb{N}^+$ and write $V_n = \{1, 2, \ldots, n\}$. We define an *$n$-lift* to be a sequence $\mathfrak{L}_n$ of permutations $\sigma_0, \sigma_1, \ldots, \sigma_{d+2e} \in \mathfrak{S}(V_n) \cong \mathfrak{S}(n)$ satisfying

$$\sigma_{j*} = \sigma_j^{-1} \text{ for } 0 \leqslant j \leqslant d + 2e, \quad \text{with } \sigma_0 = \mathbb{1} \text{ and } \sigma_j \text{ being a matching for } 1 \leqslant j \leqslant d.$$

Here a permutation is said to be a *matching* when all its cycles have length 2. (We tacitly disallow simultaneously having $n > 1$ odd and $d > 0$.) Given a permutation $\sigma \in \mathfrak{S}(V_n)$, we will write P_σ for the associated permutation operator acting on the Hilbert space $\ell_2(V_n) \cong \mathbb{C}^n$, namely

$$P_\sigma = \sum_{u \in V_n} |\sigma(u)\rangle\langle u| \,. \tag{3}$$

[6] The reason for this strange-looking convention is that: (i) we wish to have coefficients to the left of monomials in our polynomials, as is standard; (ii) we wish to think of the extension graph $\widetilde{G}$ as being formed from G by replacing each vertex in V with a small "cloud" of r vertices, and connecting edges between clouds — but this forces us to tensor/Kronecker-product the $\mathbb{C}^{r \times r}$ coefficients on the right when forming adjacency matrices.

A special case occurs when $n = 1$; the unique 1-lift $\mathfrak{L}_1$ is has $\sigma_0, \sigma_1, \ldots, \sigma_{d+2e}$ all equal to the identity permutation $\mathbb{1}$.

Definition 2.15 (∞-lift)**.** We extend the definition of an n-lift to the case of $n = \infty$, as follows. Let V_∞ denote the group $\mathbb{Z}_2^{\star d} \star \mathbb{Z}^{\star e}$, with its components generated by $g_1, \ldots, g_{d+e}$. Each of these generators acts as a permutation on V_∞ by left-multiplication; we write $\sigma_1, \ldots, \sigma_{d+e}$ for these permutations. Writing also σ_0 for the identity permutation on V_∞, and $\sigma_{j*} = \sigma_j^{-1}$ for $d < j \leqslant d + e$, we define $\mathfrak{L}_\infty = (\sigma_0, \ldots, \sigma_{d+2e})$ to be "the" ∞-lift associated to index set $\mathscr{I}$. We continue to use the P_σ notation from Equation (3) for the permutation operator acting on $\ell_2(V_\infty)$ associated to σ.

Definition 2.16 (Polynomial lift)**.** Let p be a self-adjoint matrix polynomial over index set $\mathscr{I} = \{0, 1, \ldots, d + 2e\}$ with coefficients in $\mathbb{C}^{r\times r}$. Write $p = \sum_{w\in\mathscr{T}} a_w X^w$, where $\mathscr{T}$ is the finite set of reduced words on which p is supported; for $w \in \mathscr{T}$ we call $a_w X^w$ the associated *term*. The $*$-operation is an involution on these terms, since p is self-adjoint. Thus we may consider $\mathscr{T}$ to be a color set, with each $w \in \mathscr{T}$ being designated a matching-index or a permutation-index depending on whether $w = w^*$ or not. (If $\mathscr{T}$ contains the empty word, we treat that as the identity-index.) Now given an n-lift $\mathfrak{L}_n = (\mathbb{1}, \sigma_1, \ldots, \sigma_{d+2e})$, with $n \in \mathbb{N}^+ \cup \{\infty\}$, we define the associated *polynomial lift* to be the $\mathscr{T}$-color-regular graph $\mathscr{G}_n(\mathfrak{L}_n, p)$ on vertex set V_n defined as follows: for each vertex $u \in V_n$ and each term $a_w X^w$, we include a directed edge from u to $\sigma^w(u)$, with matrix-weight a_w. Here σ^w denotes the permutation formed from the monomial X^w by substituting X_j with σ_j for each $j \in \mathscr{I}$ (and it denotes the identity permutation if w is the empty word).

Notation 2.17. Given a polynomial lift $\mathcal{G}_n = \mathscr{G}_n(\mathfrak{L}_n, p)$ as in the preceding definition, we write

$$A_n(\mathfrak{L}_n, p) = \sum_{w\in\mathscr{T}} P_{\sigma^w} \otimes a_w$$

for its adjacency operator on $\ell_2(V_n) \otimes \mathbb{C}^r$. As noted earlier, this is also the adjacency operator of its extension $\widetilde{\mathcal{G}}_n$.

We will be specifically interested in two kinds of polynomial lifts. The first is the case when the polynomial p is *linear*. As Bordenave and Collins [BC19] show, thanks to the "linearization trick", in order to understand the spectrum of general polynomial lifts, it suffices to understand the spectrum of linear matrix polynomial lifts (and indeed linear lifts with no "constant term" $a_0\mathbb{1}$). Because of the importance of this case, we extend the "bouquet" terminology:

Definition 2.18 (Lifts of linear polynomials/bouquets)**.** Given a linear matrix polynomial $\sum_{j=0}^{d+2e} a_j X_j$, we may associate it with a matrix bouquet in the natural way, deleting from the index set any j with $a_j = 0$. Conversely, given any matrix bouquet $\mathcal{K} = (a_j)_{j\in\mathscr{I}}$, we will identify it with the linear polynomial $\sum_{j\in\mathscr{I}} a_j X_j$ (extending the index set $\mathscr{I}$ to include 0 if necessary, and putting $a_0 = 0$ in this case).[7] Given this identification, we may write $\mathscr{G}_n(\mathfrak{L}_n, \mathcal{K})$ for the n-lift of this polynomial. When $n = 1$, we will simply denote the 1-lift $\mathscr{G}_1(\mathfrak{L}_1, \mathcal{K})$ as $\mathscr{G}_1(\mathcal{K})$. In particular, the ∞-lift $\mathscr{G}_\infty(\mathfrak{L}_\infty, \mathcal{K})$ is a color-regular infinite tree of degree $|\mathscr{I}\backslash\{0\}|$ (with self-loops, if $0 \in \mathscr{I}$). This graph is the *universal cover* of the bouquet $\mathcal{K}$. For notational simplicity, we will henceforth denote it simply by $\mathscr{G}_\infty(\mathcal{K})$.

[7] It is *almost* the case that the one-vertex graph bouquet is the 1-lift of this linear polynomial; the only catch is that, following [BC19], we have insisted that a "matching" permutation has no self-loops. An alternative inelegancy would be to allow matchings to have self-loops, as in the 1-regular configuration model.

The terminology "universal cover" stems from the following observation (cf. Fact 2.8):

Fact 2.19. *Let $\mathcal{G}$ be a color-regular graph with weights $(a_j)_{j\in\mathscr{J}}$ and let $\mathcal{K}$ be the associated bouquet. Then $\mathscr{G}_\infty(\mathcal{K})$ covers $\mathcal{G}$.*

More generally, we have the following key observation:

Fact 2.20. *Let p be a self-adjoint matrix polynomial over index set $\mathscr{J} = \{0, 1, \dots, d+2e\}$ with coefficients in $\mathbb{C}^{r\times r}$. Let $\mathfrak{L}_n = (\mathbb{1}, \sigma_1, \dots, \sigma_{d+2e})$ be an n-lift, $n \in \mathbb{N}^+$. Then $\mathscr{G}_\infty(p)$ covers $\mathscr{G}_n(\mathfrak{L}_n, p)$ and hence (Fact 2.9) also $\mathfrak{X} = \widetilde{\mathscr{G}}_\infty(p)$ covers $G = \widetilde{\mathscr{G}}_n(\mathfrak{L}_n, p)$.*

The second kind of polynomial lift that will concern us is the case when the polynomial p's coefficient matrices have 0-1 entries. In this case, the extended ∞-lift $\mathfrak{X}$ described in Fact 2.20 will be an ordinary unweighted infinite graph, and the extended n-lift G will be an ordinary unweighted finite graph that is covered by $\mathfrak{X}$. The main theorem of Bordenave and Collins [BC19] implies that when the n-lift $\mathfrak{L}_n$ is chosen uniformly at random, the resulting G will be $\mathfrak{X}$-Ramanujan (cf. Definition 1.3) with high probability. Our work has two aspects. First, we derandomize the Bordenave–Collins result, provided deterministic poly(n)-time algorithms for producing n-lifts $\mathfrak{L}_n$ such that $G = \widetilde{\mathscr{G}}_n(\mathfrak{L}_n, p)$ is $\mathfrak{X}$-Ramanujan (starting in Section 4). Second, we explore and partly characterize the kinds of infinite graphs $\mathfrak{X}$ that may arise as $\mathfrak{X} = \widetilde{\mathscr{G}}_\infty(p)$ (in Section 3). As we will be significantly investigating these graphs, we will give them a name:

Definition 2.21 (MPL graph)**.** We say an (undirected, unweighted, multi-)graph $\mathfrak{X}$ is *an MPL graph* if there is a matrix polynomial p with coefficient matrices in $\{0,1\}^{r\times r}$ such that $\widetilde{\mathscr{G}}_\infty(p)$ consists of disjoint copies of $\mathfrak{X}$.

We allow disjoint copies for two reasons: (i) in some cases, we only know how to generate multiple copies of $\mathfrak{X}$ via polynomial lifts; (ii) if $\mathfrak{X}'$ consists of disjoint copies of $\mathfrak{X}$, then the notions of "$\mathfrak{X}$-Ramanujan" and "$\mathfrak{X}'$-Ramanujan" coincide (since $\mathfrak{X}'$ has the same spectrum — indeed, spectral measure — as $\mathfrak{X}$, and since $\mathfrak{X}'$ covers G if and only if $\mathfrak{X}$ covers G).

2.4 Projections

Notation 2.22 (Projection to the nontrivial subspace)**.** For $n < \infty$, we define the following unit vector:

$$|+\rangle_n = \tfrac{1}{\sqrt{n}} \sum_{u\in V_n} |u\rangle \in \ell_2(V_n).$$

We sometimes identify this vector with its 1-dimensional span, and we write $|+\rangle_n^\perp \leqslant \ell_2(V_n)$ for its $(n-1)$-dimensional orthogonal complement. Every permutation matrix P_σ for $\sigma \in \mathfrak{S}(V_n)$ preserves both $|+\rangle_n$ and $|+\rangle_n^\perp$; thus we may write

$$P_\sigma = |+\rangle_n\langle+|_n + (0 \oplus P_{\sigma,\perp}), \tag{4}$$

where $P_{\sigma,\perp}$ denotes the action of P_σ on $|+\rangle_n^\perp$ (i.e., the standard group representation of σ) and the 0 is operating on $|+\rangle_n$. We may analogously define $A_{n,\perp}(\mathfrak{L}_n, p)$, and for linear polynomials, $B_{n,\perp}(\mathfrak{L}_n, p)$ for the action of adjacency/nonbacktracking operators on $|+\rangle_n^\perp$.

We refer to the eigenvalues in $\operatorname{spec}(A_n)\setminus\operatorname{spec}(A_{n,\perp})$ as the "trivial" eigenvalues. These trivial eigenvalues are precisely the r eigenvalues of the 1-lift $A_1(p)$.

Proposition 2.23. *The following multiset identity holds:*

$$\operatorname{spec}(A_n(\mathfrak{L}_n, p)) = \operatorname{spec}(A_1(p)) \cup \operatorname{spec}(A_{n,\perp}(\mathfrak{L}_n, p)).$$

Proof. From Equation (4),

$$A_n(\mathfrak{L}_n, p) = \sum_{w \in \mathscr{T}} |+\rangle\langle+| \otimes a^w + \sum_{w \in \mathscr{T}} (0 \oplus P_{\sigma^w,\perp}) \otimes a^w. \qquad \square$$

3 On MPL graphs

The goal of this section is to illustrate a wide variety of infinite graphs that can be realized as MPL graphs, and to prove some partial characterizations of MPL graphs. In this section we will freely switch between writing X_j^* and X_j^{-1} for the adjoint of X_j. We may also sometimes return to the convention (from the introduction) of writing Y_j for self-adjoint indeterminates and Z_j, Z_j^* for the remaining adjoint pairs. We remind the reader of the convention (arising because we multiply matrices on the left) that a term like $X_i X_j$ means "first do X_j, then do X_i".

3.1 Examples of MPL graphs

In this section we will give several examples of MPL graphs, and demonstrate that generalize a number of graph products found in the literature, including free products of finite vertex transitive graphs [Zno75], free products of finite rooted graphs [Que94], additive products [MO20], and amalgamated free products [VK19].

For finite lifts, we additionally show how some replacement products and zig-zag products [RVW02] may be expressed as matrix polynomial lifts.

Example 3.1 ($r = 1$, linear polynomials)**.** The simplest example of MPL graphs occurs when $r = 1$ and the polynomial is linear. In this case, $p = Y_1 + \cdots + Y_d + Z_1 + Z_1^{-1} + \cdots + Z_e^{-1}$ for some d, e, and the resulting MPL graph, $\mathfrak{X} = \widetilde{\mathscr{G}}_\infty(p)$, is the $(d + 2e)$-regular infinite tree.

Example 3.2 ($r = 1$, a general polynomial)**.** In fact, more interesting MPL graphs can already be created with $r = 1$. For instance, with $p = Y + Z + Z^{-1} + Z^{-1}YZ$ we obtain that $\widetilde{\mathscr{G}}_\infty(p)$ is $C_4 \star C_4$, the free product of two 4-cycles. See Figure 7 for an illustration, and Example 3.5 for a generalization to arbitrary free products of Cayley graphs.

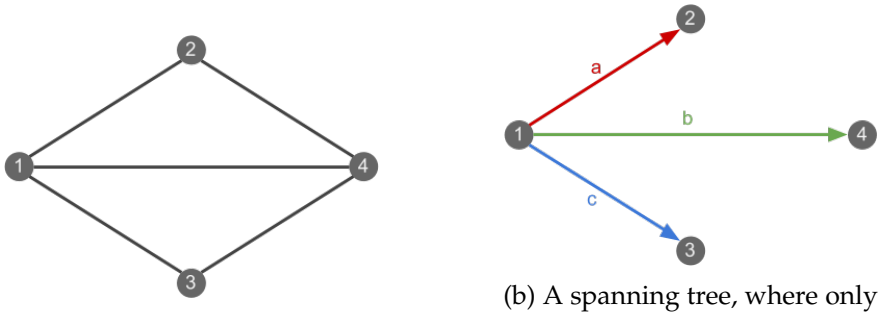
Example 3.3 (Lifts and universal covers)**.** As we have already seen in Section 2.3, when we take $\mathcal{K}$ to be a matrix bouquet of a finite graph $G = (V, E)$ (a linear polynomial where each coefficient has only 1 nonzero entry), the extended n-lift $\widetilde{\mathscr{G}}_n(\mathfrak{L}_n, \mathcal{K})$ is a lift of G in the sense of Amit and Linial [AL02]. Moreover, the infinity-lift $\widetilde{\mathscr{G}}_\infty(\mathcal{K})$ contains $|V|$ copies of the universal covering tree of G.

Using matrix bouquets we can easily obtain non-regular graphs. For example, if we let G be a $(3, 4)$-bipartite complete graph, the infinity-lift of the matrix bouquet of G contains 7 copies of the infinite $(3, 4)$-biregular tree.

Example 3.4 (Adding cycles by polynomial terms)**.** Next, we give a somewhat esoteric example, where the extension of the infinite lift $\widetilde{\mathscr{G}}_\infty$ contains infinitely many copies of a finite graph. This construction will not be as useful for applying Theorem 1.1, but will be illustrative for further examples of various graph products.

When the infinite lift does have infinite copies of a finite graph, the spectrum is equal to the finite graph's spectrum (but with infinite multiplicity).

To create infinitely many copies of a finite undirected graph $G = (V, E)$, we construct the polynomial iteratively. We first start off with a spanning tree of G, which we call $H = (V, E')$, and let p be the linear polynomial corresponding to the matrix bouquet of H. When we create the matrix bouquet, p has $|E|$ pairs of adjoint indeterminates Z_i and Z_i^*, one associated with each undirected edge in H. We further have an involution on the indices such that $Z_i^* = Z_{i*}$. Since the universal cover of a tree is itself, it's clear that $\widetilde{\mathscr{G}}_\infty(p)$ now contains countably infinite copies of H. Next, we "add" p the terms which will generate the edges of G which are not in H. If $\{u, v\}$ is an edge in G but not H, there is a sequence of directed edges bringing u to v which corresponds to a monomial $Z_{i_1} Z_{i_2} \dots Z_{i_k}$. We then add the terms $|v\rangle\langle u|\, Z_{i_k} \dots Z_{i_1} + |u\rangle\langle v|\, Z_{i_1^*}^* \dots Z_{i_k^*}^*$ to the polynomial. An example is in Figure 3.



(a) The graph we want to create

(b) A spanning tree, where only one of the directed edges is shown

Figure 3: An example of creating arbitrary graphs with cycles in the infinite lift. First we find a spanning tree (shown in Figure 3b. The edges have been colored and oriented for illustration purposes; the direction of the red edge indicates the directed edge corresponding to Z_a, the opposite edge corresponds to Z_a^*). The spanning tree corresponds to the linear polynomial $|2\rangle\langle 1|\, Z_a + |4\rangle\langle 1|\, Z_b + |3\rangle\langle 1|\, Z_c + |1\rangle\langle 2|\, Z_{a*} + |1\rangle\langle 4|\, Z_{b*} + |1\rangle\langle 3|\, Z_{c*}$. The two non-tree edges are given by $|4\rangle\langle 2|\, Z_b Z_{a*} + |2\rangle\langle 4|\, Z_a Z_{b*}$ and $|3\rangle\langle 4|\, Z_c Z_{b*} + |4\rangle\langle 3|\, Z_b Z_{c*}$. Adding all these together gives the nonlinear polynomial that creates copies of the graph in Figure 3a in the infinite polynomial lift.

Example 3.5 (Free products of finite vertex transitive graphs)**.** The construction in Example 3.4 is a helpful building block in creating graph products, for instance the free product of vertex transitive graphs (as defined by Znoĭko [Zno75]). For instance, this construction includes Cayley graphs of free products finite groups.

Let G and H be finite vertex transitive graphs (e.g. those of Cayley graphs of finite groups; the particular generating set used does not matter here). We will construct the free product $G \star H$ as an MPL graph. Let $\overline{G}$ and $\overline{H}$ be spanning trees of G and H respectively. Then, let p be the linear polynomial corresponding to the matrix bouquet of $\overline{G} \times \overline{H}$ (the Cartesian product), and let q be the sum of the polynomial terms corresponding to edges in $G \times H$ but not in $\overline{G} \times \overline{H}$ (constructed in the same way as in Example 3.4). Then, $p + q$ is a polynomial whose infinite lift contains (multiple copies of) the free product of G and H. An example for creating $C_3 \star C_4$ is shown in Figure 4.

Example 3.6 (Additive lifts and products)**.** Additive lifts and products are a graph product for non-vertex-transitive finite graphs defined by Mohanty and O'Donnell in in [MO20].

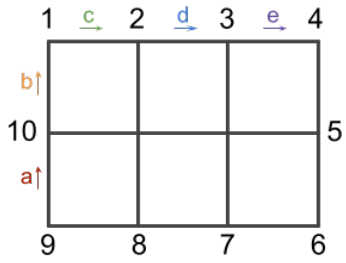


Figure 4: Here is an example of how to create $C_3 \star C_4$ as an MPL graph. We first get the matrix bouquet of the Cartesian product of a spanning tree of C_3 and C_4 (which are simple paths of length 2 and 3 respectively), creating the grid shown above. Then, when creating the matrix bouquet, we assign generators to each edge (some omitted in the drawing), which orients the edges as shown. Finally, we add terms corresponding to the edges in $C_3 \times C_4$ but not in the grid, for example, $|1\rangle\langle 9| Z_b Z_a + |9\rangle\langle 1| Z_{a^*} Z_{b^*}$ for the edge $\{1, 9\}$ and $|1\rangle\langle 4| Z_e Z_d Z_c + |4\rangle\langle 1| Z_{c^*} Z_{d^*} Z_{e^*}$ for the edge $\{1, 4\}$. We similarly add terms corresponding to $\{10, 5\}$, $\{2, 8\}$, $\{3, 7\}$ and $\{4, 6\}$.

The components of the additive product are finite, unweighted, undirected graphs called *atoms*, which are defined on a common vertex set.

Definition 3.7 (Additive products [MO19, Definition 3.4])**.** Let $A_1, \ldots, A_c$ be atoms on a common vertex set $[r]$. Assume that the sum graph $G = A_1 + \cdots + A_c$ is connected; letting $\underline{A}_j$ denote A_j with isolated vertices removed, we also assume that each $\underline{A}_j$ is nonempty and connected. We now define the (typically infinite) *additive product* graph $A_1 \dotplus \cdots \dotplus A_c := (V, E)$ where V and E are constructed as follows.

Let v_1 be a fixed vertex in $[r]$; let V be the set of strings of the form $v_{k+1} C_k v_k C_{k-1} \ldots v_2 C_1 v_1$ for $k \geqslant 0$ such that:

i) each v_i is in $[r]$ and each C_i is in $[c]$,

ii) $C_i \neq C_{i+1}$ for all $i < k$,

iii) v_i and v_{i+1} are both in $\underline{A}_{C_i}$ for all $i \leqslant k$;

and, let E be the set of edges on vertex set V such that for each string $s \in V$,

i) we let $\{uCs, vCs\}$ be in E if $\{u, v\}$ is an edge in $\underline{A}_C$,

ii) we let $\{uCs, vC'uCs\}$ be in E if $\{u, v\}$ is an edge in $\underline{A}_{C'}$, and

iii) we let $\{v_1, vCv_1\}$ be in E if $\{v_1, v\}$ is an edge in $\underline{A}_C$.

We show two ways to construct additive products using polynomial lifts, which illustrates that constructions using polynomial lifts are in general not unique. Let the atoms be $A_1, \ldots, A_m$ on r vertices each.

For the first construction, for each A_i, let $\overline{A}_i$ be a spanning tree. Let H be the sum of $\overline{A}_1, \ldots, \overline{A}_m$, including parallel edges (in other words, sum the adjacency matrices of $\overline{A}_1, \ldots, \overline{A}_m$). Then, we start with the matrix bouquet of H. Then, for each A_i, for each edge not in the spanning tree $\overline{A}_i$, we add the term corresponding to that edge to the polynomial similar to Example 3.4.

The second construction has a pair of adjoint indeterminates $Z_{A,v}$ and $Z^*_{A,v}$ for each atom A and each vertex v which is not an isolated vertex in A. We construct the polynomial p iteratively.

For each edge $\{u, v\}$ in the atom A, we add the term $|v\rangle\langle u|\, Z_{A,v} Z_{A,u}^* + |u\rangle\langle v|\, Z_{A,u} Z_{A,v}^*$ to the polynomial p. We repeat this for for every edge in each every atom A_i, $1 \leqslant i \leqslant m$. Then $\widetilde{\mathscr{G}}_\infty(p)$ consists of (r copies of) the additive lift $A_1 \dotplus \cdots \dotplus A_m$.

The finite graphs arising from the n-lifts of the second construction also has a nice interpretation. Mohanty and O'Donnell make the following definition of *additive lifts*,

Definition 3.8 (additive lifts)**.** Let $A_1, \ldots, A_c$ be atoms on a common vertex set $[r]$. The *additive n-lift* of $A_1, \ldots, A_c$ is the following:

i) It has vertex set $[n] \times [r]$.

ii) For each vertex v in each atom A_i, let $\sigma_{A_i,v}$ be a permutation on $[n]$.

iii) For each atom A_i, for each original edge (u, v), add the matching $\sigma_{A_i,v}^{-1} \sigma_{A_i,u}$ to the vertices $[n] \times \{u\}$ and $[n] \times \{v\}$.

One can check that an additive n-lift is the same as an n-lift of the polynomial described in the second construction above. Applying Theorem 1.1, we can moreover deduce that as $n \to \infty$ the spectrum of a uniformly random additive n-lift (with the trivial eigenvalues of $A_1 + \cdots + A_m$ removed) is close in Hausdorff distance to the corresponding additive product with high probability.

Example 3.9 (Amalgamated free products)**.** Vargas and Kulkarni [VK19] define another graph product based on amalgamated free products from free probability. These generalize the notion of free products of graphs defined in [Que94], and can be used to express Cayley graphs of amalgamated free group products. These graph products are defined on *rooted* graphs, i.e. graphs (V, E) with a distinguished vertex r. We denote these rooted graphs as a triple (V, E, r).

Definition 3.10 ([VK19])**.** Let $G_1 = (V_1, E_1, o_1), \ldots, G_n = (V_n, E_n, o_n)$ be finite rooted undirected graphs. Assume that each G_i comes equipped with an edge coloring $c_i : E_i \to C_i$ such that $C_i \cap C_j = \varnothing$ for every $i \neq j$. Let $C := \bigcup_{i=1}^n C_i$ and let $G = (V, E, r)$ be a rooted graph $G = (V, E, r)$ together with an edge coloring $c : E \to C$, and where $V = [k]$. We call (G, c) the *relator graph*.

We construct the *free product of the (G_i, c_i) with amalgamation* over (G, c), which we denote by $*_{G,c}\{(G_i, c_i)\}_{i=1}^n$. Let V_0 be the set of strings of the form $v_k v_{k-1} \ldots v_2 v_1 e$ for $k > 0$ where

i) o is one of $o_1, \ldots, o_n$,

ii) $v_1, \ldots, v_k$ are elements of $V_1 \backslash o_1, \ldots, V_n \backslash o_n$,

iii) v_i, v_{i+1} do not belong to the same $V_j \backslash o_j$ for $1 \leqslant i \leqslant k-1$.

The vertex set is $\{(r, o)\} \cup [k] \times V_0$. The edges are defined such that $\{(i, vu), (i', v'u)\}$ is an edge if

i) $\{i, i'\}$ is in E,

ii) u is the empty string or some element of V_0,

iii) $\{v, v'\}$ is an edge in some V_j for $1 \leqslant j \leqslant n$, and we treat any of $o_1, \ldots, o_n$ appearing as one of v or v' as the empty string, except when both u and v or v' are the empty string, in which case we let uv or uv' be o,

iv) finally, $c(\{i, i'\}) = c_j(\{v, v'\})$.

This can be constructed as an MPL graph in the following way: Let r be the number of colors k. For each graph G_i in the product and for each non-root vertex v in G_i, let $Y_{G_i,v}$ be a self-adjoint indeterminate. Let $\{u, v\}$ be an edge in G_i where neither u nor v is o_i. We then add the term $(|k_1\rangle\langle k_0| + |k_0\rangle\langle k_1|)Y_{G_i,u}Y_{G_i,v}$ to the polynomial for every edge $\{k_0, k_1\}$ in the relator graph G such that $c(\{k_0, k_1\}) = c_i(\{u, v\})$. When $\{o_i, v\}$ is an edge in G_i, we add the term $(|k_1\rangle\langle k_0| + |k_0\rangle\langle k_1|\, Y_{G_i,v}$ to the polynomial for every edge $\{k_0, k_1\}$ in the relator graph G such that $c(\{k_0, k_1\}) = c_i(\{o_i, v\})$.

As an example (derived from Example 6.1 of [VK19]), let

$$p = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} Y_{1,1} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} Y_{2,1} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} Y_{2,2} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} Y_{2,1}Y_{2,2}.$$

Then, $\widetilde{\mathscr{G}}_\infty(p)$ is the Cayley graph of $SL(2, \mathbb{Z})$ with the group presentation $\langle a, b \,|\, a^4 = b^6 = 1\rangle$.

Example 3.11 (Constant terms; beyond additive products, and amalgamated free products)**.** The class of MPL graphs also include graphs beyond additive products and amalgamated free products. To construct such graphs, one key observation is that we have yet to use the constant term a_0 in the polynomials. One use case of adding the constant term is that we can create MPL graphs where the extension of the ∞-lift contains only one component. For instance, one can add edges which connect the different copies of the universal covering tree created by lifting a matrix bouquet. For example, the following polynomial

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} Z + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} Z^*$$

gives rise to a graph which looks like a infinite ladder (Figure 5a). In general, there may be multiple, possibly very different, ways to construct the same graph. We can also create ladders without using the constant term a_0. For example, we can construct (6 copies of) the graph in Figure 5b by starting off with the matrix bouquet of the graph in Figure 6 and adding further edges to create cycles. The vertices are labeled $1, \ldots, 6$ and the edges are labeled $a, \ldots, f$. Let

$$\begin{aligned} p = {} & |1\rangle\langle 5|\, Z_a + |2\rangle\langle 1|\, Z_b + |3\rangle\langle 2|\, Z_c + |4\rangle\langle 3|\, Z_d + |5\rangle\langle 4|\, Z_e + |2\rangle\langle 6|\, Z_f \\ & + |3\rangle\langle 5|\, Z_c Z_b Z_a + |6\rangle\langle 1|\, Z_a Z_e Z_d Z_c Z_f. \end{aligned}$$

Then $\widetilde{\mathscr{G}}_\infty(p + p^*)$ is the graph depicted in Figure 5b.

(a) The infinite ladder graph

(b) A ladder with alternating hexagons and squares

Figure 5: Examples of MPL graphs

Another example of a graph which, as far as we know, requires a constant term to express as an MPL graph is depicted in Figure 8.

Example 3.12 (Replacement products and zig-zag products)**.** It is also possible to view some graphs arising from polynomial lifts as the result of replacement products and zig-zag products [RVW02].

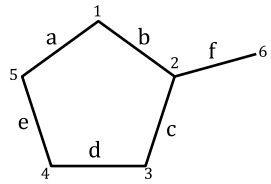


Figure 6: An illustration of the graph whose matrix bouquet serves as a base for Figure 5b

Definition 3.13 (Replacement product)**.** Let G be a D-regular undirected, unweighted graph on n vertices, and let H be a d-regular undirected, unweighted graph on D vertices. Further, G is equipped with a *rotation map* $\mathrm{Rot}_G : [n] \times [D] \to [n] \times [D]$ is a permutation where $\mathrm{Rot}_G(u, i) = (v, j)$ if the i^{th} outgoing edge from u is the j^{th} outgoing edge from v. This can be thought of as a coloring on the edges.

The *replacement product* of G and H, denoted $G \textcircled{r} H$, is a graph on nD vertices which we identify with $[n] \times [D]$. The graph is constructed by first making n copies of H, labeling the vertices with $[n] \times [D]$, and then joining every $(u, i) \in [n] \times [D]$ to $\mathrm{Rot}_G(u, i)$. This results in a $(d + 1)$-regular graph.

With some further restrictions on G, we can view the replacement product as a matrix polynomial lift. Namely, we require that G is the sum of e permutations and d' matchings on $[n]$. In particular, we fix an involution $*$ on $\{0, 1, \ldots, d + 2e\}$ as in a matrix bouquet, and require that G is the graph of an n-lift $\mathfrak{L}_n = (\sigma_0, \sigma_1, \ldots, \sigma_{d+2e'})$ (satisfying $\sigma_{i*} = \sigma_i^{-1}$). In terms of the rotation map, this results in $\mathrm{Rot}_G(u, i) = (\sigma_i(u), i^*)$. Now let A_H be the adjacency matrix of H, and define

$$p := A_H + \sum_{i=1}^{d+2e'} |i\rangle\langle i^*| \, X_i.$$

The X_i terms are self adjoint whenever the corresponding permutation σ_i is a matching, and are part of an adjoint pair X_i, X_{i*} otherwise. Then, $A_n(\mathfrak{L}_n, p)$ is the adjacency matrix of the replacement product $G \textcircled{r} H$.

Definition 3.14. Similar to replacement products, the *zig-zag product* of a D-regular graph G on n vertices, and a d-regular graph H on D vertices is a graph on $[n] \times [D]$, denoted $G \textcircled{z} H$. First, we create n *clouds* of D vertices, each corresponding to a copy of H. An edge exists in $G \textcircled{z} H$ from (u, i) to (v, j) if there exist i' and $j' \in [D]$ such that (i, i') is an edge in H, $\mathrm{Rot}_G(u, i') = (v, j')$, and (j', j) is an edge in H. This can be thought of as taking a step within a cloud (with edges defined by H), then a step between clouds (with edges defined by G), and then a step within a cloud. This construction results in a d^2-regular graph.

Again, we can express zig-zag products as the lift of a matrix polynomial when G is the graph of an n-lift $\mathfrak{L}_n$. Let A_H be the adjacency matrix of H, and define

$$p := \sum_{i=1}^{d+2e'} A_H^2 X_i.$$

Then, $A_n(\mathfrak{L}_n, p)$ is the adjacency matrix of the zig-zag product $G \textcircled{z} H$.

3.2 Structure, connectivity and geometry of polynomial lifts

In this section, we will examine some properties that infinite polynomial lifts must satisfy. We will take p to be a self-adjoint polynomial with $\{0,1\}^{r\times r}$ coefficients, in d self-adjoint indeterminates and e indeterminates and their adjoints. In this section, whenever we refer to a matrix polynomial p we mean with these constraints unless otherwise specified.

Though our main interest is with MPL graphs, our results in this sections apply more generally to the extensions of infinite lifts $\widetilde{\mathscr{G}}_\infty(p)$ of such polynomials (which, in general, consist of the union of possibly non-isomorphic finite or infinite graphs). We will refer to any connected component of an extension of a infinite lift as a "graph arising from a polynomial lift".

Recall that the vertex set of $\widetilde{\mathscr{G}}_\infty(p)$ is $V_\infty \times [r]$, where V_∞ is the free product of d copies of $\mathbb{Z}_2$ and e copies of $\mathbb{Z}$. We think of the vertices of V_∞ as corresponding to the *reduced words* formed by d self-inverse generators and e pairs of generators and their adjoints, named $g_1, \ldots, g_{d+2e}$. Following our convention of left-multiplying permutations, e.g. we think of the word g_2g_1 as g_1 followed by g_2. For a word w, we write $\overline{w}$ to denote its reduced word. In this section we use the notation $\mathfrak{L}_\infty = (o, g_1, \ldots, g_{d+2e})$ where o is the identity element of V_∞, $g_1 \ldots g_{d+2e}$ are generators with $g_1 \ldots g_d$ being self-inverse. Given some term $a_w X^w$ in p, we write g^w to denote substituting the generators into X^w. We label the vertices of $\widetilde{\mathscr{G}}_\infty(p)$ by $(w,i) \in V_\infty \times [r]$, where w is a word of generators, and i indexes into the cloud of r vertices corresponding to w.

The graphs of infinite polynomial lifts are clearly locally finite, and they additionally are constrained to look "tree-like", i.e. their structure and geometry are similar to those of trees. We can formalize this in terms of the *treewidth* and *hyperbolicity*. The treewidth measures how close the graph is to a tree structurally, while the hyperbolicity of a graph measures how close the graph distance metric is to that of a tree metric.

In particular, we will find that the graphs that arise as the infinite lifts of polynomials all have finite treewidth. Let us recall some definitions:

Definition 3.15. Let $G = (V, E)$ be a graph (possibly with infinite vertices). A *tree decomposition* of G is a tree T whose vertices are sets $W = \bigcup_{i\in\Sigma} W_i$ indexed by some set Σ. Each W_i is a subset of V. (T, W) satisfies the following properties:

1. Each $v \in V$ is in at least one W_i.
2. If $(u, v) \in E$, then there exists some $i \in \Sigma$ such that both u and v are in some W_i.
3. If $u \in V$ is in W_i and also W_j, then it is in every W_k for W_k in the unique path in T from W_i to W_j.

Definition 3.16. The *treewidth* of G is the minimum of $\max_{i\in\Sigma} |W_i| - 1$ over all tree decompositions (T, W).

Proposition 3.17. *Let p be a matrix polynomial, and let m be the sum of the degrees of the terms of p. Then, the treewidth of the extension of the ∞-lift $\widetilde{\mathscr{G}}_\infty(p)$ is bounded by $(m+1)r$.*

Proof. We can construct a tree decomposition. Each $v \in V_\infty$ is associated to a cloud of r vertices in $\widetilde{\mathscr{G}}_\infty(p)$. Our vertex sets $\{W_v\}_{v\in V_\infty}$ in the tree decomposition will be indexed by V_∞, and the tree structure on $\{W_v\}$ is also inherited from V_∞. Each W_v contains a copy of the r vertices in the cloud of $v \in V_\infty$, and for each polynomial term $a_w X^w$ in p, W_v also contains the r vertices in the cloud of $u \in V_\infty$ for every u along the path from v to $g^w v$. □

Corollary 3.18. *In particular, graphs which arise from infinite lifts of matrix polynomials must have finite treewidth. An example of a graph that does not have finite treewidth is an infinite grid. Therefore, we cannot derive grids from noncommutative polynomials.*

We have seen from the previous section that, in general, the infinite lifts of a polynomial can contain many connected components. It is also easy to see that these need not be isomorphic. Nevertheless, given the labels of two vertices in $\widetilde{\mathscr{G}}_\infty(p)$, based on p it is easy to decide if they belong in the same connected component.

In the scalar case where $r = 1$, we can decide connectivity with a deterministic finite automaton by utilizing Stallings foldings [Sta83]. The following is from Section 2 of Kapovich and Myasnikov [KM02], which we refer the reader to for full details. We reproduce a sketch of the proof here.

Proposition 3.19. *Let T be a set of words from a free group which is closed under inverse. We say a word w is* reachable *by T if the reduced word $\overline{w}$ can be formed by concatenating an arbitrary combination of words from T, possibly with repeats, and then reducing. Then, the language L of such words is regular.*

Proof sketch. We construct a deterministic finite automaton which, given an input reduced word w, accepts if w is reachable by T. Since the reduced words are also a regular language, and regular languages are closed under intersection, this shows that L is regular.

We construct the automaton iteratively as a directed graph labeled with the generators (the direction of the edge indicates to use the generator or its inverse). We start off with a single node z which will serve as both the initial and final state. For each word $t \in T$, we add a directed loop of length $|t|$ starting and ending at z, such that traversing the loop recovers t. Then, we apply a "folding" process: whenever a node has two outgoing edges to nodes x and y with the same direction and the same label, we replace the nodes x and y with a single new node with incident edges equal to the union of incident edges of x and y. The process terminates when every node has at most one incident node of every label. □

Given a scalar-coefficient polynomial p, the vertices of $\widetilde{\mathscr{G}}_\infty(p)$ correspond to words of the free group. For the connectivity question, the coefficients of p are irrelevant, so we can assume they are all 1. We can take T to be the set of words of generators corresponding to the terms of p. Then, u and v in V_∞ are connected if and only if $\overline{vu^{-1}}$ is in the language defined in Proposition 3.19. It is easy to see that the construction can be modified to allow for self-inverse generators by making the edges labeled by those generators undirected.

With a reduction to the scalar case, we can also understand connectivity for matrix-coefficient polynomials.

Proposition 3.20. *Let p be a matrix polynomial. Let (u, i) and $(v, j) \in V_\infty \times [r]$ be vertices in $\widetilde{\mathscr{G}}_\infty(p)$. It is efficiently decidable whether (u, i) and (v, j) belong to the same connected component.*

Proof. We first create r additional generators, each corresponding to one of the vertices in a cloud. Call these generators $h_1, \ldots, h_r$ and $h_1^{-1} \ldots h_r^{-1}$.

p can be written as a sum of terms of the form $|i\rangle\langle j| \, X^w$ (i.e. terms whose coefficients have exactly one nonzero entry). For each such term, add the word $h_j g^w h_i^{-1}$ to T. Since p is self-adjoint, T is closed under inverse. Then, (u, i) and (v, j) are connected if and only if $h_j v u^{-1} h_i^{-1}$ is reachable by T, which we can check using Proposition 3.19. □

We now move on to describe the *hyperbolicity* of graphs arising from infinite lifts of polynomials. Hyperbolicity is another measure of how tree-like a graph is. Gromov [Gro87] defined

this notion of hyperbolicity for groups; it generalizes a notion of how much a space is like a Riemannian manifold with negative curvature [Gro83]. Hyperbolicity is also interesting for finite graphs [BRS11, BRSV13], including random graphs [CFHM12]. It is related to other combinatorial properties of the graph such as chordality [BKM01, WZ11], independence number and max degree [RS12], and the circumference and girth [HPR19]. It is also useful describing real-world graphs, and has applications in networks and routing (e.g. [MSV11, Kle07]).

In the context of geometric group theory, hyperbolicity has also been studied for infinite graphs, for instance those arising from the free products of groups [Hor16], and tessellations of the Euclidean plane [Car17]. More generally, simplicial complexes arising from free groups such as the *complex of free factors* [BF14] and the *free splitting complex* [HM13] have been proven to be hyperbolic.

We use the following definition for hyperbolicity of (possibly infinite) graphs.

Definition 3.21. Let G be an unweighted, undirected graph. We say that G is δ-hyperbolic if for every 3 vertices u, v, w, the shortest paths (u, v), (u, w) and (v, w) satisfy that for every node $x \in (u, v)$ there exists a node y in (u, w) or (v, w) such that the graph distance (length of the shortest path) $\text{dist}(x, y) \leqslant \delta$.

As an example, a tree is 0-hyperbolic. We are interested in graphs where δ is finite and constant.

In contrast to treewidth, which captures information about the local structure of the graph, hyperbolicity is a global property which captures global information about distances between vertices. Hyperbolicity and treewidth are in general not comparable; for instance, a n-cycle has treewidth 2 but hyperbolicity $\lceil \frac{n}{4} \rceil$. Meanwhile, the Cayley graph of the fundamental group of the torus $G = \langle a, b, c, d \mid aba^{-1}b^{-1}cdc^{-1}d^{-1} = 0 \rangle$ has arbitrarily large grid minors, and hence infinite treewidth, but it has finite hyperbolicity.

For graphs with multiple connected components, δ is typically taken to be infinity. Hence, we are concerned only with the hyperbolicity of connected components within infinite polynomial lifts. In our application of spectral approximations of polynomial lifts, we typically only consider cases where the infinite polynomial lifts is a finite number of isomorphic copies of some infinite graph.

We next show that all connected components of a infinite polynomial lift are δ-hyperbolic for δ a constant depending on the polynomial p. Similar to the case of connectivity, we will eventually make a reduction to the scalar case of $r = 1$. We first prove the statement in the case of $r = 1$, hence we take p to have all coefficients 1 in the following. Given the extension of the infinity lift $\widetilde{\mathscr{G}}_\infty(p)$, we will want to consider the graph given by the polynomials and the Cayley graph of the subgroup of the free group with vertex set V_∞ at the same time. When possible we will use $\widetilde{\mathscr{G}}_\infty(p)$ to refer to the polynomial graph, but V_∞ to refer to the Cayley graph (switching between the two views when necessary). To proceed, we start with some definitions:

Definition 3.22. Let $u, v \in V_\infty$. We define the *tree path* between u and v to be the sequence of generators in $\overline{vu^{-1}}$ (i.e. the edges along the (unique) path in the Cayley graph of the free group V_∞). We denote the *tree distance* $\text{treedist}(u, v)$ as the length of the tree path.

Definition 3.23. We define the *d-neighborhood* of a vertex v to be the set of vertices $w \in V_\infty$ such that $\text{treedist}(v, w) \leqslant d$. We also define the d-neighborhood of the tree path from u to v to be the set of vertices $w \in V_\infty$ such that there exists x on the tree path from u to v with $\text{treedist}(x, w) \leqslant d$.

Definition 3.24. Let p be a polynomial. Let $u, v \in \widetilde{\mathscr{G}}_\infty(p)$ be in the same connected component. We define a *monomial path* between u and v to be a shortest path between u and v in $\widetilde{\mathscr{G}}_\infty(p)$.

Further let the *monomial distance* $\mathrm{dist}(u, v)$ be the length of the monomial path. Note that there may not be a unique shortest path.

The following observations are key for the hyperbolicity of the scalar polynomial lifts.

Lemma 3.25. *Let p be a polynomial with $\{0, 1\}$ coefficients and the degree of any term is at most m. Let $u, v \in \widetilde{\mathscr{G}}_\infty(p)$ be in the same connected component. Now fix a monomial path from u to v and consider traversing it. The following hold:*

i) Every time the monomial path leaves the m-neighborhood of the tree path between u and v, it returns to the m-neighborhood at a point at most tree distance $2m$ away.

ii) For every vertex w on the tree path from u to v, there is a vertex x in the m-neighborhood of w such that x is on the monomial path from u to v.

Proof. We first prove (i). Suppose that the monomial path leaves the m-neighborhood of the tree path at a point x. Note that $\mathrm{treedist}(x, v) > m$ since x has to be outside a m-neighborhood of v. Let y be the vertex such that $\mathrm{treedist}(x, y) = m$ and y is on the tree path between x and v, and let y' be one step further than y on the tree path between x and v. Also note that y' must be on the tree path between u and v.

When removed, the edge (y, y') disconnects V_∞. In particular, after removing the edge we can think of y and y' as the roots of two subtrees, the union of which includes all the vertices of V_∞. u is in the subtree of y and v is in the subtree of y'. Since the maximum degree of p is m, any monomial-step can only get to a vertex at most m in tree distance away from x. Therefore, any monomial step from x remains in the subtree of y.

In order to reach v, the monomial path has to enter the subtree of y'. Let z be the last vertex visited by the monomial path in the subtree of y before entering the subtree of y'. Since each monomial-step can traverse at most m tree-steps, z has to be tree-distance at most $m - 1$ from y, and is hence in the m-neighborhood of y'. Finally, we note that $\mathrm{treedist}(x, z) \leqslant \mathrm{treedist}(x, y) + \mathrm{treedist}(y, z) \leqslant 2m$.

We observe (ii) as a consequence of (i). At any vertex on the monomial path from u to v, the next step either goes to a vertex at most m away in the m-neighborhood of the tree path from u to v, or it leaves the m neighborhood and rejoins at a point at most $2m$ away. □

We can now prove the hyperbolicity of the connected components, where connectivity is defined based on the monomial paths.

Proposition 3.26. *A graph arising from a polynomial lift has finite (depending on the polynomial p) hyperbolicity in every connected component.*

Proof. We first prove the statement for scalar-coefficient p, then make a reduction for general p.

Fix a scalar-coefficient p with maximum degree m, and let C be the maximum monomial distance between any two vertices which are connected and tree distance at most $2m$ apart.

Let u, v, w be vertices in V_∞ from the same connected component. Let x be on a monomial path from u to v. By part (i) of Lemma 3.25, x is either in the m-neighborhood of the tree path from u to v, or x is on a path which leaves the m-neighborhood of the tree path but will rejoin at a point at most $2m$ away from where it left. In either case, x is at most monomial-distance C away from a vertex x' which is in the m-neighborhood of the tree path from u to v and is on the monomial path from u to w. Let y be the point which is on the tree path from u to v and is closest in tree distance to x'.

Since the union of the tree paths between u and v, v and w and u and w is a tree, y lies on either the tree path from u to w or from v to w. Suppose without loss of generality y is on the tree path from u to w. Now fix a monomial path between u to w. From part (ii) of Lemma 3.25 there is a vertex y' such that y' is on the monomial path between u to w and $\mathrm{treedist}(y', y) \leqslant m$. x' and y' are in the same connected component since they are on the monomial path between u and v and the monomial path between u and w respectively. We also know $\mathrm{treedist}(x', y') \leqslant \mathrm{treedist}(x', y) + \mathrm{treedist}(y, y') = 2m$, therefore $\mathrm{dist}(x', y') \leqslant C$, so $\mathrm{dist}(x, y') \leqslant 2C$. Therefore, the connected component that u, v, and w belong to is $2C$-hyperbolic.

Finally, for matrix-coefficient polynomials, we make a similar reduction as in Proposition 3.20, by adding a generator and its inverse for each $i \in [r]$, which we call $h_1, \ldots h_r$ and $h_1^{-1}, \ldots, h_r^{-1}$. We use V'_∞ to denote the free product of V_∞ with another r copies of $\mathbb{Z}$. Next, we make a transformation of p to a scalar-coefficient polynomial p'. We write p as a sum of terms of the form $|i\rangle\langle j|\, X^w$, and for each such term, we add $X_j X^w X_i^{-1}$ to p'. Given (u, i), (v, j) and $(w, k) \in V_\infty \times [r]$, we transform these vertices to o, $\overline{h_j u^{-1} h_i^{-1}}$ and $\overline{h_k w u^{-1} h_i^{-1}}$ in V'_∞. This increases all tree distances by 2, and leaves monomial distances unchanged. Therefore, the connected components of $\widetilde{\mathscr{G}}_\infty(p)$ are also hyperbolic. □

3.3 Benjamini–Schramm convergence and unimodularity

Next, we study some properties of the automorphism group of the graph $\widetilde{\mathscr{G}}_\infty(p)$.

Proposition 3.27. *Let p be a matrix polynomial with coefficients in $\{0,1\}^{r\times r}$. $\widetilde{\mathscr{G}}_\infty(p)$ has at most r vertex orbits.*

Proof. Consider the graph of $\widetilde{\mathscr{G}}_\infty(p)$, and recall that each vertex can be labeled (v, c) for $v \in V_\infty$ and $c \in [r]$. The natural covering map from $\widetilde{\mathscr{G}}_\infty(p)$ to $\widetilde{\mathscr{G}}_1(p)$ maps any two vertices labeled with the same c in $\widetilde{\mathscr{G}}_\infty(p)$ onto the same vertex in $\widetilde{\mathscr{G}}_1(p)$. For any two vertices (u, c) and (v, c), there must be some sequence g a product of generators corresponding to the polynomial terms of p that such that $gu = v$. Then, the map $P_g \otimes I_{r\times r}$ applied to $A_\infty(p)$ brings (u, c) to (v, c), so these two vertices are in the same orbit. Therefore, there can be at most r orbits. □

We move on to another property that the automorphism groups of these infinite graphs have. We say that a graph is *unimodular* the Haar measure on the automorphism group is both left invariant and right invariant[8]. Intuitively, this condition means that the infinite graphs "looks the same" in any direction. Unimodularity can be a useful property in characterizing graphs, for example, all unimodular trees are the universal covering tree of some finite graph [BK90].

Definition 3.28. Let $G = (V, E)$ be a graph. We say that $f : V \times V \to [0, \infty]$ is *diagonally invariant* if for all automorphisms $g \in \mathrm{Aut}(G)$, we have $f(gu, gv) = f(u, v)$. In the case where we let the graph be a parameter of f, we also say f is diagonally invariant if $f(G; u, v) = f(G; gu, gv)$ for all automorphisms $g \in \mathrm{Aut}(G)$.

The following is the Mass-Transport Principle of Benjamini, Lyons, Peres and Schramm [BLPS99], and in a vertex transitive graph, satisfying this condition is equivalent to unimodularity.

Definition 3.29. Let $G = (V, E)$ be a vertex-transitive graph. Fix some vertex $o \in V$. We say that G satisfies the *Mass-Transport Principle* if for all diagonally invariant $f : V \times V \to [0, \infty]$,

$$\sum_{x\in V} f(o, x) = \sum_{x\in V} f(x, o).$$

[8] unrelated to totally unimodular matrices

The Mass-Transport Principle can be extended to non-vertex-transitive graphs. For finite graphs, we can choose $o \in V$ uniformly at random and require that the equation hold in expectation. We can extend this to *quasi-transitive* graphs (graphs which have a finite number of vertex orbits) as well. To do so, we have to introduce some machinery. For a graph G, we call the pair (G, o) with $o \in G$ a *rooted* graph. (G, o) is isomorphic to (G', o') if there is an isomorphism from G to G' bringing o to o'. We write $[G, o]$ to denote the isomorphism class of (G, o). Let $\mathcal{G}$ be the space of isomorphism classes of locally finite graphs. We sometimes also need to consider isomorphism classes of graphs with two roots, which we will denote $[G; a, b]$. We call $\mathcal{G}_\star$ the space of these isomorphism classes.

We can define a metric on $\mathcal{G}$,

$$d([G, o], [G', o']) = \sup_{r>0} \{2^{-r} : B([G, o], r) \text{ is isomorphic to } B([G', o'], r)\}.$$

We write $\mathcal{G}_D$ for the subspace of $\mathcal{G}$ where the degrees of the vertices are uniformly bounded by D. Under the above metric, $\mathcal{G}_D$ is compact metric space. In the following, we consider probability measures on the Borel σ-algebra generated by the metric.

To formulate the Mass-Transport Principle, one can try to extend the notion of picking a vertex uniformly at random to infinite graphs. The natural way to do this for a quasi-transitive graph is to pick a isomorphism class with probability proportional to the density of the type of the root. This is a probability measure μ on $\mathcal{G}$. This form of the Mass-Transport Principle is known as the Intrinsic Mass-Transport Principle due to Benjamini and Schramm [BS01].

Definition 3.30. We say that a probability measure μ on $\mathcal{G}$ satisfies the Mass-Transport Principle or is *unimodular* if for all Borel $f : \mathcal{G}_\star \to [0, \infty]$,

$$\mathop{\mathbf{E}}_{[G,o]\sim\mu} \left[\sum_{x \in V(G)} f(G; o, x) \right] = \mathop{\mathbf{E}}_{[G,o]\sim\mu} \left[\sum_{x \in V(G)} f(G; x, o) \right].$$

For quasi-transitive graphs, the Intrinsic Mass-Transport Principle is equivalent to unimodularity [AL07, Theorem 3.1], which explains the terminology. For a finite graph G, the measure supported on $[G, o]$ for $o \in G$ defined by

$$\mu_n[G, o] = \frac{|\operatorname{Aut}(G) \cdot o|}{|V(G)|}$$

is unimodular.

One example of a non-unimodular graph is Trofimov's grandparent graph [Tro85]. The construction is as follows: Start with a 3-regular tree, and pick an end ξ. For each vertex, call the vertex one step in the direction of ξ its parent, and the vertex two steps in the direction of ξ its grandparent. Join each vertex to its grandparent with an undirected edge. This produces a 6-regular graph. Another class of examples are the Diestel–Leader graphs [DL01].

Our next goal is to show that the graphs arising from matrix polynomials. This gives us some slightly surprising examples of graphs that cannot be expressed by non-commutative polynomials. For instance, the grandparent graph appears to have a repeating structure, and is vertex transitive, but nevertheless non-commutative matrix polynomials cannot express it.

A nice consequence of defining unimodularity via the Mass-Transport principle is that we can now consider limits of graphs as limits of probability measures on $\mathcal{G}$. For a single finite graph G, we can take the probability measure on equivalence classes of $[G, o]$ for $o \in G$ chosen uniformly

at random, or in other words, $[G, o]$ is weighted according to the size of the vertex isomorphism class of o. The notion of convergence we will use is that of weak convergence of measures. Recall that we say a sequence of probability measures $\{\mu_n\}_{n=1}^{\infty}$ *converges weakly* or *converges in distribution* to μ if for every continuous function f,

$$\lim_{n\to\infty} \mathop{\mathbf{E}}_{\mu_n}[f] = \mathop{\mathbf{E}}_{\mu}[f].$$

In the literature, μ (or G on which μ is supported) is sometimes known as the "local limit" or the convergence is known as "convergence in the sense of Benjamini–Schramm". By having the support of μ on a random collection of G, we can also consider the limit of random graphs.

If each μ_n is supported on some finite graph G_n, this form of convergence intuitively says that with high probability, each radius-r-neighborhood of G_n looks like a radius-r neighborhood of G appearing with probability according to μ, where $r \to \infty$ as $n \to \infty$.

An important open question is whether every unimodular measure is such a limit of finite graphs [AL07, Question 10.1].

Rooted graphs of depth at most k (for $k > 0$) which we call $\mathcal{G}^{(k)}$ forms a base of $\mathcal{G}$ (as a metric space). The finite graphs are all isolated points in this metric, thus these base elements are discrete. To show weak convergence in $\mathcal{G}_D$ (i.e. when the degree is uniformly bounded), it suffices to show convergence in measure on every set in the base.

To show convergence on the base sets, it suffices to show that for every finite rooted graph Γ and $k > 0$,

$$\lim_{n\to\infty} \mathop{\mathbf{Pr}}_{[G,o]\sim\mu_n}[B(o,k) \cong \Gamma] = \mathop{\mathbf{Pr}}_{[G,o]\sim\mu}[B(o,k) \cong \Gamma],$$

where $B(o,k)$ denotes a k-neighborhood around the root of G.

One useful property is that unimodularity is preserved under weak convergence [AL07], and we will use this to show the following:

Proposition 3.31. *Graphs arising from infinite lifts of matrix polynomials are unimodular.*

Proof. We will do so by showing that the random finite lifts weakly converge to the measure on the infinite graph. Let μ_n be the uniform distribution over randomly-rooted n-lifts. Since this is a uniform distribution over unimodular measures, it's unimodular as well.

From Proposition 3.27, there are at most r vertex orbits in the infinite graph. Thus, picking a random root can be done by picking a representative corresponding to one of the r base vertices uniformly. We call this probability measure μ.

For every $k > 0$, with high probability as $n \to \infty$ none of the permutations have cycles smaller than k. In this case, for all finite rooted graphs Γ, $\mathbf{Pr}_{[G,o]\sim\mu_n}[B(o,k) \cong \Gamma] = \mathbf{Pr}_{[G,o]\sim\mu}[B(o,k) \cong \Gamma]$. Therefore, $\mu_n \to \mu$ in measure on the base sets. □

3.4 Connectedness of finite lifts produced by our algorithm

We now turn to the question of whether the finite lifts produced from our derandomization of Theorem 1.1 (with the full theorem statement in Theorem 10.13) result in connected (extension) graphs when applied to a polynomial. To do so, we consider random walks on the graphs.

Definition 3.32. Let A be the adjacency matrix of a locally finite multigraph G (indexed by pairs of vertices). We define the *random walk matrix* of G to be P_G such that

$$(P_G)_{uv} := A_{uv}/\deg(u),$$

where $\deg(u)$ is the degree of the vertex u.

We need the following fact about P_G.

Fact 3.33. *For a finite multigraph G, the largest eigenvalue of P_G is 1 and it has multiplicity 1 if and only if G is connected.*

For spectral computations, it is convenient to analyze a symmetrized version of P_G.

Fact 3.34. *Let A be the adjacency matrix of a locally finite multigraph G. Let D be the* degree matrix, *i.e. the diagonal matrix where D_{uu} is the degree of vertex u in A. Then, $P_G = D^{-1}A$, and we define the symmetric matrix $S_G := D^{-1/2}AD^{-1/2}$.*

One can check that if λ is an eigenvalue of S_G with eigenvector v, then λ is an eigenvalue of P_G with eigenvector $D^{1/2}v$, so in particular their spectra are equal.

For an infinite graph G, it turns out that whether $\rho(P_G) = 1$ is characterized by an isoperimetric property: whether the graph is *amenable*. Amenability was first defined as a property of groups, see, e.g. [Woe00, Ch. 12]. The definitions for amenability of Cayley graphs can also be extended to general graphs. There are many equivalent definitions for amenability in graphs; the following is a statement about the *isoperimetric constant* of the graph.

Definition 3.35. Let $G = (V, E)$ be a locally finite undirected graph. Let $S \subset V$ be a set of vertices, and we define the *boundary* $\partial S \subset E$ to be the set of edges $\{u, v\}$ with $u \in S$ and $v \notin S$. Let $m(S)$ be the sum of the vertex degrees in S. We say that G is *amenable* if

$$\inf_{S \subset V, |S| < \infty} \frac{|\partial S|}{m(S)} = 0.$$

The quantity $\inf_{S \subset V, |S| < \infty} |\partial S|/|S|$ is known as the *isoperimetric constant* or *isoperimetric number*.

As some examples, every finite graph is amenable, an infinite path is amenable, while a d-regular infinite tree for $d \geqslant 3$ is not amenable.

The following was shown by Gerl [Ger88] (which also contains other equivalent characterizations of amenability), and is an extension the result on Cayley graphs by Kesten [Kes59]. The result for d-regular graphs was also shown in [BMST88].

Proposition 3.36. *Let G be a locally finite undirected graph. Then, G is amenable if and only if $\rho(P_G) = 1$.*

We can therefore conclude:

Proposition 3.37. *Let p be a matrix polynomial with r-dimensional coefficients. Suppose that the extension of the ∞-lift $\widetilde{\mathscr{G}}_\infty(p)$ is non-amenable, and that $\widetilde{\mathscr{G}}_1(p)$ is connected. Then, there exists $\varepsilon > 0$ such that applying Theorem 10.13 with k the total degree of p and R the maximum Frobenius norm of the coefficients of p gives a lift $\mathfrak{L}_n$ such that $A_n(\mathfrak{L}_n, p)$ satisfies the conclusion of Theorem 10.13 and $\widetilde{\mathscr{G}}_n(\mathfrak{L}_n, p)$ is connected.*

Proof. Write $p = \sum_{w \in \mathscr{T}} a_w X_w$, $A_n = A_n(\mathfrak{L}_n, p)$ and let D_n be the degree matrix of A_n. We will show that $S_{\widetilde{\mathscr{G}}_n} = D_n^{-1/2} A_n D_n^{-1/2}$ projected onto $|+\rangle_n$ has the eigenvalue 1 with multiplicity 1, and $\rho(S_{\widetilde{\mathscr{G}}_n, \perp}) < 1$ (where $S_{\widetilde{\mathscr{G}}, \perp}$ denotes the projection onto the orthogonal complement of $|+\rangle_n$).

Recall that $\mathscr{G}_n(\mathfrak{L}_n, p)$ is a $\mathscr{T}$-regular graph. Therefore, each group of vertices in $\widetilde{\mathscr{G}}_n$ corresponding to the same vertex in $\mathscr{G}_n$ has the same set of degrees. Hence, D_n is a block-diagonal matrix with identical blocks, in fact $D_n = I_n \otimes D_1$, where D_1 is the degree matrix of $\widetilde{\mathscr{G}}_1(p)$. Therefore, we can express $S_{\widetilde{\mathscr{G}}_n}$ as the lift of another polynomial. If we define $p' = \sum_{w \in \mathscr{T}} D_1^{-1/2} a_w D_1^{-1/2} X^w$, then

$S_{\widetilde{\mathscr{G}}_n} = A_n(\mathfrak{L}_n, p')$. Note that p' has the same degree as p, each of the coefficients has Frobenius norm at most that of the corresponding coefficient in p, and that $\widetilde{\mathscr{G}}_n(\mathfrak{L}_n, p')$ is connected if and only if $\widetilde{\mathscr{G}}_n(\mathfrak{L}_n, p)$ is connected.

By Facts 3.33 and 3.34, $\max \sigma(S_{\widetilde{\mathscr{G}}_n}) = 1$. Since we assume $\widetilde{\mathscr{G}}_1$ is connected, $S_{\widetilde{\mathscr{G}}_1}$ has maximum eigenvalue 1 with multiplicity 1. Using Proposition 2.23, the eigenvalue 1 is in $\operatorname{spec}(S_{\widetilde{\mathscr{G}}_n}) \setminus \operatorname{spec}(S_{\widetilde{\mathscr{G}}_n,\perp})$. Hence, showing that $\left\|S_{\widetilde{\mathscr{G}}_n,\perp}\right\| < 1$ ensures that $S_{\widetilde{\mathscr{G}}_n}$ has eigenvalue 1 with multiplicity 1.

Let $\varepsilon > 0$ be a parameter to be chosen later. We apply Theorem 10.13 with r being the dimension of the coefficients of p, k the total degree and R the maximum norm of the coefficients of p, and N of our choice, obtaining an n-lift $\mathfrak{L}_n$. From Proposition 3.36, the assumption that $\widetilde{\mathscr{G}}_\infty$ is non-amenable implies $\rho(A_\infty(\mathfrak{L}_\infty, p')) < 1$. If we choose ε small enough that $\rho(A_\infty(\mathcal{L}_\infty, p')) + \varepsilon < 1$, then $\left\|S_{\widetilde{\mathscr{G}}_n,\perp}\right\| < 1$, and we conclude that $\widetilde{\mathscr{G}}_n(\mathfrak{L}_n, p')$ is connected, hence $\widetilde{\mathscr{G}}_n(\mathfrak{L}_n, p)$ is connected. □

We will not attempt to characterize which polynomial lifts are non-amenable here, but refer the reader to some previous work done on amenable graphs. The isoperimetric constant, and hence determining if a graph is amenable, is related to the *growth rate* [Moh88] (in particular, non-amenable graphs have exponential growth). Amenability also has connections to percolation processes [BLPS99], and other spectral properties and properties of the random walk [Ger88, Kai92, Dod84, Sal92]. There are also some combinatorial characterizations e.g. on trees [Woe00, Thm. 10.9], [FMK17], and on vertex-transitive graphs [SW90].

4 Additional analytic setup and definitions

4.1 Nonbacktracking operators

Throughout this section, let $\mathscr{I}_0 = \{0, 1, \ldots, d + 2e\}$ be an index set, and write $\mathscr{I} = \mathscr{I}_0 \setminus \{0\}$ for the version with no identity-index. Let $p = \sum_{w \in \mathscr{T}} a_w X^w$ be a self-adjoint matrix polynomial over $\mathscr{I}_0$, with coefficients in $\mathbb{C}^{r \times r}$, let $\mathcal{K} = a_0 + a_1 X_1 + \cdots + a_{d+2e} X_{d+2e}$ be a self-adjoint linear matrix polynomial over $\mathscr{I}_0$, with coefficients in $\mathbb{C}^{r \times r}$, and let $\mathfrak{L}_n = (\mathbb{1}, \sigma_1, \ldots, \sigma_{d+2e})$ be an n-lift, $n \in \mathbb{N}^+ \cup \{\infty\}$. Let $\mathcal{G}_n = \mathcal{G}_n(\mathfrak{L}_n, \mathcal{K})$ be the resulting n-lifted color-regular graph on vertex set V_n. Recall that the adjacency operator of $\mathcal{G}_n$ is

$$A_n = A_n(\mathfrak{L}_n, \mathcal{K}) = \sum_{i \in \mathscr{I}_0} P_{\sigma_i} \otimes a_i.$$

Definition 4.1 (Nonbacktracking operator)**.** The *nonbacktracking operator* of $\mathcal{G}_n$ is the following (in general non-Hermitian) bounded operator on $\ell_2(V_n) \otimes \ell_2(\mathscr{I}) \otimes \mathbb{C}^r$:

$$B_n = B_n(\mathfrak{L}_n, \mathcal{K}) = \sum_{i,j=1}^{d+2e} \mathbf{1}[j \neq i^*] \cdot P_{\sigma_i} \otimes |j\rangle\langle i| \otimes a_j.$$

As $\ell_2(V_n) \otimes \ell_2(\mathscr{I}) \cong \ell_2(V_n \times \mathscr{I})$, we may think of B_n as arising from nonbacktracking steps along the colored arcs $V_n \times \mathscr{I}$ of $\mathcal{G}_n$.

Remark 4.2. We only define the nonbacktracking operators of $\mathcal{G}_n$ arising as lifts of linear polynomials, and not general polynomials.

Remark 4.3. A warning: the summation defining B_n above does not include $i, j = 0$; i.e., it excludes the identity-loop color in $\mathcal{G}_n$. Hence B_n does not depend on a_0. The reason for this convention (chosen by Bordenave–Collins) is that in the needed technical theorems involving B_n, we will be in a setup where $a_0 = 0$ anyway. In case $a_0 = 0$, the operator B_n corresponds to the "usual" nonbacktracking operator of the extended graph $\widetilde{\mathcal{G}}_n$.

Notation 4.4. We use $B_\infty(\mathcal{K}) = B_\infty(\mathfrak{L}_\infty, \mathcal{K})$ to denote the nonbacktracking operator of the ∞-lift of $\mathcal{K}$ $A_\infty(\mathcal{K})$.

Notation 4.5 (Projection to the nontrivial subspace)**.** As in Section 2.4, for a linear polynomial p we write $B_{n,\perp}(\mathfrak{L}_n, p)$ for the action of adjacency/nonbacktracking operators on $|+\rangle_n^\perp$.

Finally, we record a kind of formula for the spectral radius of B_∞. The following fact is proven within "Proof of Lemma 14" in Bordenave and Collins's work [BC19]:[9]

Proposition 4.6. *For $i \in [d+2e]$, define the $r^2 \times r^2$ matrix $A_i = a_i \otimes \overline{a}_i$, where $\overline{a}_i$ denotes the matrix formed by taking the complex conjugate of each entry in a_i. Also, define L to be the $(d+2e) \times (d+2e)$ block matrix, with blocks of size $r^2 \times r^2$, whose (i,j)th block is $\mathbb{1}[j \neq i^*] \cdot A_i$. Then $\rho(B_\infty) = \sqrt{\rho(L)}$.*

Corollary 4.7. *If $\|a_i\|_{\mathrm{F}} \leqslant R$ for all $i \in [d+2e]$, then $\rho(B_\infty) \leqslant \sqrt{d+2e}R$.*

Proof. The claim follows from $\rho(B_\infty)^4 = \rho(L)^2 \leqslant \mathrm{tr}(LL^*)$ and

$$\mathrm{tr}(LL^*) = \sum_{i=1}^{d+2e} \sum_{j \neq i^*} \mathrm{tr}(A_i A_i^*) = \sum_{i=1}^{d+2e} \sum_{j \neq i^*} \|a_i\|_{\mathrm{F}}^2 \|\overline{a}_i\|_{\mathrm{F}}^2 \leqslant (d+2e)^2 R^4. \qquad \square$$

4.2 Lifting permutations and product lifts

We continue with the notation from the previous section, assuming henceforth that $n < \infty$.

Definition 4.8 (The graph $G_{\mathfrak{L}_n}$ associated to a lift)**.** Let $\sigma \in \mathfrak{S}(n)$ be a general permutation. Then $P_\sigma + P_{\sigma^{-1}}$ is the adjacency matrix of an n-edge undirected graph on vertex set $[n]$, which we will denote by G_σ. Similarly, suppose $\sigma \in \mathfrak{S}(n)$ is a matching permutation. Then P_σ alone is the adjacency matrix of an $n/2$-edge undirected graph on vertex set $[n]$; abusing notation slightly, we will also denote it by G_σ. Now supposing $\mathfrak{L}_n = (\mathbb{1}, \sigma_1, \ldots, \sigma_{d+2e})$ is an n-lift, we define an associated undirected (multi)graph $G_{\mathfrak{L}_n} = G_{\sigma_1} + \cdots + G_{\sigma_d} + G_{\sigma_{d+1}} + \cdots + G_{\sigma_{d+2e}}$. More precisely, $G_{\mathfrak{L}_n}$ is the graph on vertex set V_n whose adjacency matrix is $\sum_{i=1}^{d+2e} P_{\sigma_i}$. Note that we have omitted the identity-index $i = 0$ here, so $G_{\mathfrak{L}_n}$ has no self-loops. Equivalent definitions are that $G_{\mathfrak{L}_n} = \widetilde{\mathscr{G}}_n(\mathfrak{L}_n, \mathcal{K})$ where $\mathcal{K}$ is the linear matrix polynomial $\sum_{j \in \mathscr{J}} X_i$ (with $r = 1$).

Definition 4.9 (The lift of a permutation/matching)**.** Suppose again that $\sigma \in \mathfrak{S}(n)$ is a general permutation. Given a sequence $\mathfrak{T}_m$ of permutations $\tau_0 = \mathbb{1}, \tau_1, \ldots, \tau_{2n} \in \mathfrak{S}(m)$ with $\tau_{j+n} = \tau_j^{-1}$ (for $j \in [n]$), we can form the lifted extension graph $\widetilde{\mathscr{G}}_m(\mathfrak{T}_m, \mathcal{K}_{G_\sigma})$. This graph is of the form G_ρ for a permutation $\rho \in \mathfrak{S}(mn)$, and we refer to (ρ, ρ^{-1}) as the *permutation pair* (σ, σ^{-1}) *lifted by* $\mathfrak{T}_m$. Identifying $\mathbb{C}^{mn} = \mathbb{C}^m \otimes \mathbb{C}^n$ we have $P_\rho = \sum_{j=1}^n P_{\tau_j} \otimes |\sigma(j)\rangle\langle j|$.

If $\sigma \in \mathfrak{S}(n)$ is a matching permutation, we may do something similar. In this case, P_σ alone is the adjacency matrix of an $n/2$-edge undirected graph on vertex set $[n]$, which we also denote

[9] Therein they write "We deduce from Theorem 16(ii) that $\rho(B_\star)$ is equal to the spectral radius of L", but this is clearly a typo for "the *square-root of* the spectral radius of L"; personal communication with the authors confirms this.

by G_σ. Given a sequence $\mathfrak{T}_m$ of permutations $\tau_0 = \mathbb{1}, \tau_1, \ldots, \tau_n \in \mathfrak{S}(m)$ with $\tau_{j+n/2} = \tau_i^{-1}$ (for $j \in [n/2]$), we can again form the lifted extension graph $\widetilde{\mathscr{G}}_m(\mathfrak{T}_m, \mathcal{K}_{G_\sigma})$. This graph is a perfect matching, and hence may be viewed as G_ρ for $\rho \in \mathfrak{S}(kn)$ a matching permutation. We again refer to ρ as the *matching permutation σ lifted by $\mathfrak{T}_m$*.

Definition 4.10 (Product lift)**.** Let $\mathfrak{L}_n = (\mathbb{1}, \sigma_1, \ldots, \sigma_{d+2e})$ denote an n-lift. For $m \in \mathbb{N}^+$, we can produce an mn-lift $(\rho_0 = \mathbb{1}, \rho_1, \ldots, \rho_{d+2e})$ as follows. For for $1 \leqslant i \leqslant d$ assume that $\mathfrak{T}_m^{(i)} = (\tau_1^{(i)}, \ldots, \tau_n^{(i)})$ is a sequence of permutations in $\mathfrak{S}(m)$ with $\tau_{j+n/2}^{(i)} = (\tau_j^{(i)})^{-1}$; and, for $d+1 \leqslant i \leqslant d+e$ assume that $\mathfrak{T}_m^{(i)} = (\tau_1^{(i)}, \ldots, \tau_{2n}^{(i)})$ is a sequence of permutations in $\mathfrak{S}(m)$ with $\tau_{j+n}^{(i)} = (\tau_j^{(i)})^{-1}$. Now for $1 \leqslant i \leqslant d$, let ρ_i be the matching permutation σ_i lifted by $\mathfrak{T}_m^{(i)}$; and, for $d+1 \leqslant i \leqslant d+e$, let (ρ_i, ρ_{i+e}) be the permutation pair (σ_i, σ_{i+e}) lifted by $\mathfrak{T}_m^{(i)}$. We write $\mathfrak{T}_m \otimes \mathfrak{L}_n$ for the mn-lift $(\rho_i)_i$, calling it the *product* of the m-lifts $\mathfrak{T}_m^{(i)}$ with the n-lift $\mathfrak{L}_n$. Note that $G_{\mathfrak{T}_m \otimes \mathfrak{L}_n}$ is an m-fold graph lift of $G_{\mathfrak{L}_n}$ (of the usual sort).

4.3 Signed permutations

Notation 4.11. We continue in the setup of the previous section, but restrict attention to the case of $m = 2$. In this case, given the 2-lifts $\mathfrak{T}_2^{(i)} = (\tau_j^{(i)})_j$ (for $1 \leqslant i \leqslant d$, $1 \leqslant j \leqslant n$ and $d+1 \leqslant i \leqslant d+e$, $1 \leqslant j \leqslant 2n$) we may write

$$P_{\tau_j^{(i)}} = \frac{1}{2}\begin{bmatrix} +1 & +1 \\ +1 & +1 \end{bmatrix} + \chi_i(j) \cdot \frac{1}{2}\begin{bmatrix} +1 & -1 \\ -1 & +1 \end{bmatrix} = |+\rangle\langle+| + \chi_i(j) \cdot |-\rangle\langle-|\,,$$

where $\chi_1, \ldots, \chi_d : [n] \to \{\pm 1\}$ and $\chi_{d+1}, \ldots, \chi_{d+e} : [2n] \to \{\pm 1\}$ and where we use the notation

$$|+\rangle = |+\rangle_2 = \tfrac{1}{\sqrt{2}}(|1\rangle + |2\rangle), \qquad |-\rangle = \tfrac{1}{\sqrt{2}}(|1\rangle - |2\rangle).$$

Definition 4.12 (edge-signing/signed permutation)**.** Recall that we have $\chi_i(j+n) = \chi_i(j)$ for $d+1 \leqslant i \leqslant d+e$ and thus it is natural to treat these χ_i's as functions $[n] \to \{\pm 1\}$. Similarly for $1 \leqslant i \leqslant d$ we have $\chi_i(j+n/2) = \chi_i(j)$ and thus it is natural to treat these χ_i's as functions $[n/2] \to \{\pm 1\}$. In this form, we will identify $\chi_1, \ldots, \chi_d$ as *edge-signings* of the 1-regular matching graphs $G_{\sigma_1}, \ldots, G_{\sigma_d}$, and identify $\chi_{d+1}, \ldots, \chi_{d+e}$ as edge-signings of the 2-regular graphs $G_{\sigma_{d+1}}, \ldots, G_{\sigma_{d+e}}$. Relatedly, we may think of each $\chi_i\sigma_i$ as a *signed permutation* (i.e., a member of the hyperoctahedral group of order n). Collectively, $\chi = (\chi_1, \ldots, \chi_d, \chi_{d+1}, \ldots, \chi_{d+e})$ is an edge-signing of the graph $G_{\mathfrak{L}_n}$.

Notation 4.13. We will use $P_{\chi_i\sigma_i}$ to denote the signed permutation matrix naturally associated to the signed permutation $\chi_i\sigma_i$. In the context of a matrix polynomial p, given a monomial X^w, we use $P_{\chi\sigma^w}$ to denote the signed permutation matrix associated to the signed permutation $(\chi\sigma)^w$ given by substituting X_j with $\chi_j\sigma_j$ for each $j \in \mathscr{I}$. We also write χ^w to denote the signing obtained by substituting X_j with χ_j for each $j \in \mathscr{I}$.

Notation 4.14. Let $p = \sum_{w \in \mathscr{T}} a_w X^w$ be a matrix polynomial. We write $A_n(\chi\mathfrak{L}_n, p)$ for the *signed adjacency operator* of the *sign-lifted color-regular graph* $\mathscr{G}_n = \mathscr{G}_n(\chi\mathfrak{L}_n, p)$, the variant of $\mathscr{G}_n(\mathfrak{L}_n, p)$ in which the w-colored edge emanating from $u \in V_n$ is also signed by $\chi^w(u)$.

Remark 4.15. One might also interpret this as changing the matrix-weight on the edge from a^w to $\chi^w(u)a^w$, but one must be careful *not* to use this interpretation for the signed nonbacktracking operator, as we will see.

Notice that for $1 \leqslant i \leqslant d+2e$, the ith permutation ρ_i in the composition lift $\mathfrak{T}_2 \otimes \mathfrak{L}_n$ satisfies

$$\begin{aligned} P_{\rho_i} = \sum_{j=1}^{n} P_{\tau_j^{(i)}} \otimes |\sigma_i(j)\rangle\langle j| &= \sum_{j=1}^{n} \Big(|+\rangle\langle +| + \chi_i(j) \cdot |-\rangle\langle -| \Big) \otimes |\sigma_i(j)\rangle\langle j| \\ &= \sum_{j=1}^{n} |+\rangle\langle +| \otimes |\sigma_i(j)\rangle\langle j| + \sum_{j=1}^{n} \chi_i(j) \cdot |-\rangle\langle -| \otimes |\sigma_i(j)\rangle\langle j| \\ &= |+\rangle\langle +| \otimes P_{\sigma_i} + |-\rangle\langle -| \otimes P_{\chi_i \sigma_i}. \end{aligned}$$

Thus

$$\begin{aligned} A_{2n}(\mathfrak{T}_2 \otimes \mathfrak{L}_n, p) &= |+\rangle\langle +| \otimes \left(\sum_{w \in \mathscr{T}} P_{\sigma^w} \otimes a_w \right) + |-\rangle\langle -| \otimes \left(\sum_{w \in \mathscr{T}} P_{\chi_i \sigma^w} \otimes a_i \right) \\ &= |+\rangle\langle +| \otimes A_n(\mathfrak{L}_n, \mathcal{K}) + |-\rangle\langle -| \otimes A_n(\chi \mathfrak{L}_n, \mathcal{K}), \end{aligned} \tag{5}$$

From these, we deduce

Proposition 4.16. *The following multiset identities hold:*

1. $\operatorname{spec}(A_{2n}(\mathfrak{T}_2 \otimes \mathfrak{L}_n, p)) = \operatorname{spec}(A_n(\mathfrak{L}_n, p)) \cup \operatorname{spec}(A_{n,\chi}(\mathfrak{L}_n, p))$,
2. $\operatorname{spec}(A_{2n,\perp}(\mathfrak{T}_2 \otimes \mathfrak{L}_n, p)) = \operatorname{spec}(A_{n,\perp}(\mathfrak{L}_n, p)) \cup \operatorname{spec}(A_{n,\chi}(\mathfrak{L}_n, p))$.

Proof. Item 1 follows immediately from Equation (5). Using $|+\rangle \otimes |+\rangle_n = |+\rangle_{2n}$, we have

$$|+\rangle\langle +| \otimes A_n(\mathfrak{L}_n, p) = \sum_{w \in \mathscr{T}} (|+\rangle\langle +| \otimes P_{\sigma^w}) \otimes a_i = \sum_{w \in \mathscr{T}} |+\rangle_{2n} \langle +|_{2n} \otimes a^w + \sum_{w \in \mathscr{T}} (0 \oplus P_{\sigma^w, \perp}) \otimes a^w, \tag{6}$$

and combining this with Item 1 gives us Item 2 □

Similarly, in the case of a linear polynomial $\mathcal{K}$, we have

Proposition 4.17. *The following multiset identities hold:*

1. $\operatorname{spec}(B_{2n}(\mathfrak{T}_2 \otimes \mathfrak{L}_n, \mathcal{K})) = \operatorname{spec}(B_n(\mathfrak{L}_n, \mathcal{K})) \cup \operatorname{spec}(B_{n,\chi}(\mathfrak{L}_n, \mathcal{K}))$,
2. $\operatorname{spec}(B_{2n,\perp}(\mathfrak{T}_2 \otimes \mathfrak{L}_n, \mathcal{K})) = \operatorname{spec}(B_{n,\perp}(\mathfrak{L}_n, \mathcal{K})) \cup \operatorname{spec}(B_{n,\chi}(\mathfrak{L}_n, \mathcal{K}))$.

Proof. Similarly to the computations for A_n, we have

$$\begin{aligned} B_{2n} = B_{2n}(\mathfrak{T}_2 \otimes \mathfrak{L}_n, \mathcal{K}) &= |+\rangle\langle +| \otimes \left(\sum_{i,j=1}^{d+2e} \mathbf{1}[j \neq i^*] \cdot P_{\sigma_i} \otimes |j\rangle\langle i| \otimes a_j \right) \\ &\quad + |-\rangle\langle -| \otimes \left(\sum_{i,j=1}^{d+2e} \mathbf{1}[j \neq i^*] \cdot P_{\chi_i \sigma_i} \otimes |j\rangle\langle i| \otimes a_j \right) \\ &= |+\rangle\langle +| \otimes B_n(\mathfrak{L}_n, \mathcal{K}) + |-\rangle\langle -| \otimes B_n(\chi \mathfrak{L}_n, \mathcal{K}), \end{aligned} \tag{7}$$

where here we've introduced the notation

$$B_{n,\chi} = B_n(\chi \mathfrak{L}_n, \mathcal{K}) = \sum_{i,j=1}^{d+2e} \mathbf{1}[j \neq i^*] \cdot P_{\chi_i \sigma_i} \otimes |j\rangle\langle i| \otimes a_j \tag{8}$$

for the *signed nonbacktracking operator* of the sign-lifted color-regular graph $\mathscr{G}_n(\chi\mathfrak{L}_n, \mathcal{K})$. Notice here that the action of $B_n(\chi\mathfrak{L}_n, \mathcal{K})$ on a colored arc (u,i) picks up the *sign* $\chi_i(u)$ on the "stepped-from" arc (u,i), but picks up the *matrix-weights* a_j on the "stepped-to" arcs (v,j) (for $v = \sigma_i(u)$ and $j \neq i^*$). (This is why one shouldn't automatically think of an edge-signing simply as changing the matrix-weight on colored edge (u,i) from a_i to $\chi_i(u)a_i$.) Equation (7) allows us to conclude Item 1. We can further compute,

$$|+\rangle\langle+| \otimes B_n(\mathfrak{L}_n, \mathcal{K}) = \sum_{i,j=1}^{d+2e} \mathbf{1}[j \neq i^*] \cdot \left(|+\rangle\langle+| \otimes P_{\sigma_i}\right) \otimes |j\rangle\langle i| \otimes a_j$$
$$= \sum_{i,j=1}^{d+2e} \mathbf{1}[j \neq i^*] \cdot |+\rangle_{2n}\langle+|_{2n} \otimes |j\rangle\langle i| \otimes a_j \quad + \quad |+\rangle\langle+| \otimes \sum_{i,j=1}^{d+2e} \mathbf{1}[j \neq i^*] \cdot (0 \oplus P_{\sigma_i,\perp}) \otimes |j\rangle\langle i| \otimes a_j,$$

and hence from Item 1 we may further conclude Item 2. □

5 Derandomization tools

We require the following standard derandomization tools:

Definition 5.1. Let $\delta \in [0,1]$ and $t \in \mathbb{N}^+$. A sequence of random bits $\boldsymbol{x} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \in \{\pm 1\}^n$ is said to be (δ, t)-*wise uniform* if, for every $S \subseteq [n]$ with $0 < |S| \leqslant t$, it holds that $|\mathbf{E}[\prod_{i \in S} \boldsymbol{x}_i]| \leqslant \delta$.

The key property of such sequences we use is the following:

Fact 5.2. *Let $q \in \mathbb{C}[X_1, \ldots, X_n]$ be a (usual) polynomial of degree at most t such that the sum of the magnitudes of the coefficients of q is M. Then if $\boldsymbol{x} \in \{\pm 1\}^n$ is (δ, t)-wise uniform and $\boldsymbol{u} \in \{\pm 1\}^n$ is truly uniformly random, then $|\mathbf{E}[q(\boldsymbol{x})] - \mathbf{E}[q(\boldsymbol{u})]| \leqslant \delta M$.*

A classic result is that (δ, t)-wise uniform bits can be strongly explicitly constructed from a truly random seed of length $O(\log t + \log\log n + \log(1/\delta))$:

Theorem 5.3. *([NN93, AGHP92, Sho90].) There is a deterministic algorithm that, given δ, t, and n, runs in time $\mathrm{poly}(n/\delta)$ and outputs a multiset $X \subseteq \{\pm 1\}^n$ of cardinality $S = \mathrm{poly}(t \log(n)/\delta)$ (a power of 2) such that, for $\boldsymbol{x} \sim X$ chosen uniformly at random, the sequence $\boldsymbol{x}$ is (δ, t)-wise uniform. Indeed, if the algorithm is additionally given $1 \leqslant s \leqslant S$ and $1 \leqslant i \leqslant n$ (written in binary), it can output the ith bit of the sth string in X in deterministic time $\mathrm{polylog}(n/\delta)$.*

Definition 5.4. Let $t \in \mathbb{N}^+$ and let $[n]_t$ denote the set of all sequences of t distinct indices from $[n]$. A random permutation $\boldsymbol{\pi} \in \mathfrak{S}(n)$ is said to be *t-wise uniform* if, for every sequence $(i_1, \ldots, i_t) \in [n]_t$, the distribution of $(\boldsymbol{\pi}(i_1), \ldots, \boldsymbol{\pi}(i_t))$ is uniform on $[n]_t$. For $\delta \in [0,1]$, we say that $\boldsymbol{\pi}$ is *δ-almost t-wise uniform* if its distribution is δ-close in total variation distance to that of a truly t-wise uniform distribution $\boldsymbol{\pi}'$.

Again, the key property of such permutations we use is the following:

Fact 5.5. *Let q be a (usual) polynomial of degree at most t in indicator random variables $(1[\boldsymbol{\pi}(i) = j])_{i,j=1}^n$. Then q has the same expectation under a t-wise uniform $\boldsymbol{\pi} \in \mathfrak{S}(n)$ as it has under a truly uniform $\boldsymbol{\pi} \in \mathfrak{S}(n)$.*

Combining the strongly explicit (δ, t)-wise uniform permutations of [Kas07, KNR09] with the main theorem in [AL13], we obtain the following theorem (called "Corollary 2.6" in [MOP20a]):

Theorem 5.6. *([KNR09, Kas07, AL13]) There is a deterministic algorithm that, given t and n, runs in time $\mathrm{poly}(n^t)$ and outputs a multiset $\Pi \subseteq \mathfrak{S}(n)$ (closed under inverses) of cardinality $S = \mathrm{poly}(n^t)$ (a power of 2) such that, when $\boldsymbol{\pi} \sim \Pi$ is chosen uniformly at random, $\boldsymbol{\pi}$ is n^{-100t}-almost t-wise uniform. Indeed, if the algorithm is additionally given $1 \leqslant s \leqslant S$ and $1 \leqslant i \leqslant n$ (written in binary), it can output $\pi_s(i)$ and $\pi_s^{-1}(i)$ (where π_s is the sth permutation in Π) in deterministic time $\mathrm{poly}(t \log(n/\delta))$.*

Definition 5.7. For n even, we say a general permutation $\pi \in \mathfrak{S}(n)$ has the *associated matching* $\sigma \in \mathfrak{S}(n)$, where σ matches (i.e., has as a 2-cycle) the pairs $(\pi(1), \pi(2)), \ldots, (\pi(n-1), \pi(n))$. We note that when $\boldsymbol{\pi} \sim \mathfrak{S}(n)$ is uniformly random, its associated matching $\boldsymbol{\sigma}$ is uniformly distributed among matching permutations.

Remark 5.8. If $\pi \in \mathfrak{S}(n)$ has associated matching $\sigma \in \mathfrak{S}(n)$, each indicator $1[\sigma(k) = \ell]$ is a degree-2 polynomial in the indicators $1[\pi(i) = j]$.

Definition 5.9. Given an index set $\mathscr{I} = \{0, 1, \ldots, d+2e\}$, we say a random n-lift $\boldsymbol{\mathfrak{L}}_n = (\mathbb{1}, \boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_{d+2e})$ is *δ-almost t-wise uniform* if:

- the permutations $\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_d, \boldsymbol{\sigma}_{d+1}, \ldots, \boldsymbol{\sigma}_{d+e}$ are independent;
- each of $\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_d$ is the associated matching of a $\frac{\delta}{d+e}$-almost $2t$-wise uniform permutation;
- each of $\boldsymbol{\sigma}_{d+1}, \ldots, \boldsymbol{\sigma}_{d+e}$ is a $\frac{\delta}{d+e}$-almost t-wise uniform permutation.

When $\delta = 0$, we simply say that $\boldsymbol{\mathfrak{L}}_n$ is *t-wise uniform*. Note that a δ-almost t-wise uniform n-lift is δ-close, in total variation distance, to a t-wise uniform n-lift.

The $2t$-wise requirement in the second bullet of this definition was chosen so that, in light of Remark 5.8, we could make the following observation:

Fact 5.10. *In the notation of Definition 5.9, suppose q is a polynomial of degree at most t in the indicator random variables $(1[\boldsymbol{\sigma}_i(j) = k])_{i \in \mathscr{I}_0}^{j,k=1\ldots n}$. Then q has the same expectation under a t-wise uniform n-lift $\boldsymbol{\mathfrak{L}}_n$ as it has under a truly uniform n-lift $\boldsymbol{\mathfrak{L}}_n$.*

Finally, applying Theorem 5.6 and adjusting constants, we conclude the following:

Theorem 5.11. *Fix an index set $\mathscr{I}$. There is a deterministic algorithm that, given t and n, runs in time $\mathrm{poly}(n^t)$ and outputs a multiset Λ of n-lifts (indexed by $\mathscr{I}$) such that, when $\boldsymbol{\mathfrak{L}}_n \sim \Lambda$ is chosen uniformly at random, $\boldsymbol{\mathfrak{L}}_n$ is n^{-100t}-close in total variation distance to being a t-wise uniform random n-lift. Indeed, if the algorithm is additionally given $i \in \mathscr{I}_0$, $1 \leqslant s \leqslant S$, and $1 \leqslant j \leqslant n$ (written in binary), it can output $\sigma_i^{(s)}(j)$ in deterministic time $\mathrm{poly}(t \log(n/\delta))$ (where $\sigma_i^{(s)}$ is the ith permutation of the sth lift in Λ).*

(Note that we rely on Theorem 5.6's ability to compute permutation-inverses strongly explicitly, not just for computing $\sigma_i^{(s)}$ with $d+e+1 \leqslant i \leqslant d+2e$, but also for computing the matching permutations $\sigma_i^{(s)}$ with $1 \leqslant i \leqslant d$ strongly explicitly. E.g., for such a matching index, we may efficiently compute $\sigma_i(j)$ from its underlying associated general permutation π_i by first letting $k = \pi_i^{-1}(j)$, then setting k' to be $k+1$ if k is odd or $k-1$ if k is even, and finally computing $\sigma_i(j) = \pi_i(k')$.)

6 Random 2-lifts

Throughout this section we fix an index set $\mathscr{I} = \{1, \dots, d+2e\}$ with no identity-index (as we will be working exclusively with nonbacktracking operators). We will be considering linear matrix polynomials over this index set, $a_1 X_1 + \cdots + a_{d+2e} X_{d+2e}$, where $a_j \in \mathbb{C}^{r\times r}$. We will abuse our own terminology very slightly by referring to these as "bouquets" $\mathcal{K} = (a_1, \dots, a_{d+2e})$, the abuse being the possibility of $a_i = 0$. In fact, it will not even be an abuse, because we will focus on bouquets whose matrix-weights are bounded and have bounded inverses (and therefore are nonzero).

6.1 A net of bouquets

Definition 6.1. Given $R > 0$ (and implicitly $\mathscr{I}$ and r), we define $\mathfrak{K}_R$ to be the collection of all matrix bouquets $(a_1, \dots, a_{d+2e}) \in (\mathbb{C}^{r\times r})^{d+2e}$ with the property that each a_i is *R-bounded*, meaning that both $\|a_i\|_{\mathrm{F}} \leqslant R$ and $\|a_i^{-1}\|_{\mathrm{F}} \leqslant R$.

The below ε-net result is mostly proven in [BC19, Sec. 4.5]. We remark that for our main Theorem 6.5 we will only need the $\lambda = 1$ case, but for the weak derandomization of [BC19] in Section 7 we need the $\lambda = \Theta(\log n)$ case.

Proposition 6.2. *Given constants $d, e, r, \varepsilon > 0$ and $R > 1$, there is a large constant $\kappa = \kappa(d, e, r, R, \varepsilon)$ such that the following holds: Given $\lambda \in \mathbb{N}^+$, there is an (efficiently computable) net $\Xi_\lambda \subset \mathfrak{K}_{2R}$ with the following properties:*

- *$|\Xi_\lambda| \leqslant \kappa^\lambda$, and each element of Ξ_λ has "encoding length"[10] at most $\kappa\lambda$.*
- *For every R-bounded bouquet $\mathcal{K} \in \mathfrak{K}_R$, there exists $\underline{\mathcal{K}} \in \Xi_\lambda$ such that both:*
  - *$\left| \left\| B_n(\chi\mathfrak{L}_n, \mathcal{K})^\lambda \right\|_{\mathrm{op}} - \left\| B_n(\chi\mathfrak{L}_n, \underline{\mathcal{K}})^\lambda \right\|_{\mathrm{op}} \right| \leqslant \varepsilon^\lambda$ holds for every finite signed lift $\chi\mathfrak{L}_n$;*
  - *$\left| \rho(B_\infty(\mathcal{K})) - \rho(B_\infty(\underline{\mathcal{K}})) \right| \leqslant \varepsilon$.*

As this exact statement does not precisely appear in [BC19], we provide a proof (which closely follows that in [BC19, Sec. 4.5]).

Proof. Let $\mathfrak{F}_R$ denote the set of all matrices in $\mathbb{C}^{r\times r}$ with Frobenius norm at most R, and let $\mathfrak{F}_R^H$ denote the subset of Hermitian such matrices. For any fixed rational $\delta > 0$ we can create a "grid" $\mathfrak{G}_\delta$ of all matrices in $\mathfrak{F}_R$ with entries being complex integer multiples of δ, such that: (i) $|\mathfrak{G}_\delta| \leqslant O(rR/\delta)^{r^2}$; (ii) for every $b \in \mathfrak{F}_R$ there is $\underline{b} \in \mathfrak{G}_\delta$ with $\|b - \underline{b}\|_{\mathrm{F}} \leqslant \delta$. (Also, (ii) holds for $\mathfrak{F}_R^H$ when restricting to $\mathfrak{G}_\delta^H = \mathfrak{G}_\delta \cap \mathfrak{F}_R^H$.) Suppose now that $b^{-1} \in \mathfrak{F}_R$ as well. Then provided $\delta < \frac{1}{2R}$ (which we will certainly ensure) we have, writing $\Delta = \underline{b} - b$,

$$\begin{aligned}\left\| b^{-1} - \underline{b}^{-1} \right\|_{\mathrm{F}} &= \left\| b^{-1}(\Delta b^{-1} + (\Delta b^{-1})^2 + (\Delta b^{-1})^3 + \cdots \right\|_{\mathrm{F}} \\ &\leqslant \|b^{-1}\|_{\mathrm{F}}(\|\Delta\|_{\mathrm{F}}\|b^{-1}\|_{\mathrm{F}})/(1 - \|\Delta\|_{\mathrm{F}}\|b^{-1}\|_{\mathrm{F}}) \leqslant 2\delta R^2 \leqslant R;\end{aligned}$$

hence $\|\underline{b}^{-1}\|_{\mathrm{F}} \leqslant 2R$, and therefore $\underline{b}^{-1} \in \mathfrak{F}_{2R}$. For a suitable $\delta > 0$ to be chosen later we will define

$$\Xi_\lambda = \{(\underline{b}_1, \dots, \underline{b}_d, \underline{b}_{d+1}, \dots, \underline{b}_{d+e}, \underline{b}_{d+1}^*, \dots, \underline{b}_{d+e}^*) : \underline{b}_1, \dots, \underline{b}_d \in \mathfrak{G}_\delta^H \text{ and } \underline{b}_{d+1}, \dots, \underline{b}_{d+e} \in \mathfrak{G}_\delta\}. \quad (9)$$

[10] Meaning, number of bits needed to represent the matrix, with the real and complex part of each entry being stored as the ratio of two integers.

A key observation is now that for any bouquet $\mathcal{K} = (a_1, \dots, a_{d+2e}) \in \mathfrak{K}_R$ there is a bouquet $\underline{\mathcal{K}} = (\underline{a}_1, \dots, \underline{a}_{d+2e}) \in \Xi$ with

$$\|a_i - \underline{a}_i\|_{\mathrm{op}} \leqslant \|a_i - \underline{a}_i\|_{\mathrm{F}} \leqslant \delta \quad \forall i \in [d+2e];$$

hence from the definition Equation (8) and the bounds $\|P_{\chi_i \sigma_i}\|_{\mathrm{op}} \leqslant 1$ and $\| \, |j\rangle\langle i| \, \|_{\mathrm{op}} \leqslant 1$, we conclude

$$\left| \|B_n(\chi\mathfrak{L}_n, \mathcal{K})\|_{\mathrm{op}} - \|B_n(\chi\mathfrak{L}_n, \underline{\mathcal{K}})\|_{\mathrm{op}} \right| \leqslant (d+2e)^2 \cdot \delta$$

for every finite signed lift $\chi\mathfrak{L}_n$. This bound already suffices for the $\lambda = 1$ case; more generally, using $\|B^\lambda - \underline{B}^\lambda\| \leqslant \lambda \cdot \max\{\|B\|, \|\underline{B}\|\}^{\lambda-1} \cdot \|B - \underline{B}\|$, and also the bounds

$$\|B_n(\chi\mathfrak{L}_n, \mathcal{K})\|_{\mathrm{op}}, \ \|B_n(\chi\mathfrak{L}_n, \underline{\mathcal{K}})\|_{\mathrm{op}} \leqslant (d+2e)^2 \cdot R,$$

we conclude

$$\left| \left\| B_n(\chi\mathfrak{L}_n, \mathcal{K})^\lambda \right\|_{\mathrm{op}} - \left\| B_n(\chi\mathfrak{L}_n, \underline{\mathcal{K}})^\lambda \right\|_{\mathrm{op}} \right| \leqslant \lambda((d+2e)^2 R)^\lambda \cdot \delta. \tag{10}$$

Finally, it is not hard to verify that for the matrix "L" defined in Proposition 4.6, and the analogously defined "$\underline{L}$", we have

$$\|L - \underline{L}\|_{\mathrm{op}} \leqslant \|L - \underline{L}\|_{\mathrm{F}} \leqslant (d+2e) \cdot 2\delta R,$$

and also that $\|L\|_{\mathrm{op}}, \|\underline{L}\|_{\mathrm{op}} \leqslant (d+2e)R^2$. Hence by [Bha97, Thm. VIII.1.1] we get

$$|\rho(L) - \rho(\underline{L})| \leqslant (2(d+2e)R^2)^{1-1/r}((d+2e) \cdot 2\delta R)^{1/r} \leqslant 2(d+2e)R \cdot \delta^{1/r}$$
$$\implies \quad \left| \rho(B_\infty(\mathcal{K})) - \rho(B_\infty(\underline{\mathcal{K}})) \right| \leqslant \sqrt{2(d+2e)R} \cdot \delta^{1/(2r)}. \tag{11}$$

In light of Equation (9) and inequalities (10) and (11), we may be complete the proof by taking $\delta = c(d, e, r, R, \varepsilon)^\lambda$ for a suitably small rational constant $c(d, e, r, R, \varepsilon) > 0$, and then taking $\kappa = \kappa(c, d, e, r, R, \varepsilon)$ a large enough constant. □

6.2 Our main technical theorem

Our goal for the remainder of the section is to prove Theorem 6.5 below, which we state after making a few preliminary definitions.

Definition 6.3. A *color sequence* γ is a sequence $i_1, \dots, i_t \in \mathscr{I}$. Its *reverse*, denoted γ^*, is the color sequence $i_t^*, \dots, i_1^*$. We call γ a *nonbacktracking (n.b.) color sequence* if $i_{s+1} \neq i_s^*$ for all $s \in [t-1]$.

Definition 6.4. Given a color sequence $\gamma = (i_1, \dots, i_t)$ and matrix-weights $a_1, \dots, a_{d+2e}$, we introduce the notation $a(\gamma) = a_{i_t} a_{i_{t-1}} \cdots a_{i_1}$. Note that $a(\gamma^*) = a(\gamma)^*$ and hence $\|a(\gamma^*)\|_{\mathrm{F}} = \|a(\gamma)\|_{\mathrm{F}}$.

Theorem 6.5. *Let $d, e, r, R, \bar{\varepsilon}, p > 0$ be constants, and let $\Xi_1 \subset \mathfrak{K}_{2R}$ and κ be as in Proposition 6.2, with $\lambda = 1$ and $\varepsilon = \bar{\varepsilon}/3$. Then there are large enough constants C_1, C_2 such that the following holds:*

Let $\mathfrak{L}_n$ be any n-lift ($n \geqslant C_1$) that is λ-bicycle-free, where $\lambda = C_2(\log\log n)^2$. Then for a uniformly random edge-signing $\boldsymbol{\chi}$ of $G_{\mathfrak{L}_n}$, except with probability at most n^{-p} the following holds simultaneously for all matrix bouquets $\mathcal{K} \in \Xi_1$:

$$\|B_n(\boldsymbol{\chi}\mathfrak{L}_n, \mathcal{K})\|_{\mathrm{op}} \leqslant \rho(B_\infty(\mathcal{K})) + \bar{\varepsilon}/3. \tag{12}$$

When this happens we may easily deduce from Proposition 6.2 that

$$\rho(B_n(\boldsymbol{\chi}\mathfrak{L}_n, \mathcal{K})) \leqslant \|B_n(\boldsymbol{\chi}\mathfrak{L}_n, \mathcal{K})\|_{\mathrm{op}} \leqslant \rho(B_\infty(\mathcal{K})) + \bar{\varepsilon} \tag{13}$$

holds simultaneously for all R-bounded bouquets $\mathcal{K} \in \mathfrak{K}_R$.

Fix any $\mathfrak{L}_n$ and any $\mathcal{K} = (a_1, \dots, a_{d+2e}) \in \mathcal{K}_{2R}$. It suffices to show that Inequality (12) holds for this fixed $\mathcal{K}$ except with probability at most n^{-2p} over the choice of $\boldsymbol{\chi}$; we may then take a union bound over the at most κ bouquets $\mathcal{K} \in \Xi_1$, and enlarge C_1 if necessary so that $\kappa n^{-2p} \leqslant n^{-p}$.

Let us write $\boldsymbol{B}_n = B_n(\boldsymbol{\chi}\mathfrak{L}_n, \mathcal{K})$ for brevity and write $\overline{\rho} = \rho(B_\infty(\mathcal{K})) + \bar{\varepsilon}/6$. The main characterization of $\overline{\rho}$ that we will need is the below proposition of Bordenave–Collins:[11]

Proposition 6.6. *([BC19, eqn. (50)].) There is a constant $c \geqslant 1$ depending only on d, e, r, R, $\bar{\varepsilon}$ (and not on the specific a_i's) such that for all $t \in \mathbb{N}$,*

$$\sum_{\substack{\text{n.b. color sequences} \\ \gamma = (i_1, \dots, i_t)}} \|a(\gamma)\|_{\mathrm{F}}^2 \leqslant c\overline{\rho}^{2t}.$$

In particular, $\|a(\gamma)\|_{\mathrm{F}} \leqslant c\overline{\rho}^t$ for all n.b. color sequences of length t.

It now remains to prove that except with probability at most n^{-2p}:

$$\|\boldsymbol{B}_n\|_{\mathrm{op}} \leqslant \overline{\rho} + \bar{\varepsilon}/6. \tag{14}$$

The key to proving this is to use the Trace Method in expectation. In preparation for this, we make some definitions (borrowing some terminology from [MOP20a]):

Definition 6.7. In the color-regular graph $\mathcal{G}_n = \mathscr{G}_n(\mathfrak{L}_n, \mathcal{K})$, a length-$t$ walk may be naturally specified by giving a starting vertex $u \in V_n$ and a sequence $i_1, \dots, i_t \in \mathscr{I}$ of colors. We call the walk *closed* if the final vertex is equal to u. We define an *ℓ-hike* to be a closed, length-2ℓ walk composed of two consecutive n.b. color sequences $(i_1, \dots, i_\ell)$ and $(j_{\ell+1}, \dots, j_{2\ell})$; equivalently, it is defined by a length-2ℓ color sequence that is nonbacktracking except possibly at the midpoint of the walk. Finally, we call an ℓ-hike *even* if each (undirected) edge is traversed an even number of times (in either direction); more generally, we call it *singleton-free* if no edge is traversed exactly once.

Fix

$$\ell = \lceil C_3 \log n \rceil, \tag{15}$$

where C_3 is a large constant depending only on d, e, r, R, $\bar{\varepsilon}$, p. (The constant C_2 will also depend on C_3, and the constant C_1 will depend on C_2 and C_3.) Write $\boldsymbol{B}_n = B_n(\boldsymbol{\chi}\mathfrak{L}_n, \mathcal{K})$, and define the random variable

$$\mathbf{Trace} = \mathrm{tr}\Big(\boldsymbol{B}_n^\ell (\boldsymbol{B}_n^\ell)^*\Big),$$

which is nonnegative since $\boldsymbol{B}_n^\ell (\boldsymbol{B}_n^\ell)^*$ is positive semidefinite. Temporarily abbreviating $\widetilde{\sigma}_i = P_{\chi_i \sigma_i}$, one can check that

$$\boldsymbol{B}_n^\ell (\boldsymbol{B}_n^\ell)^* = \sum \widetilde{\sigma}_{j_{2\ell-1}} \widetilde{\sigma}_{j_{2\ell-2}} \dots \widetilde{\sigma}_{j_\ell} \widetilde{\sigma}_{i_\ell} \widetilde{\sigma}_{i_{\ell-1}} \dots \widetilde{\sigma}_{i_1} \otimes |j_{2\ell}\rangle\langle i_0^*| \otimes a_{j_{2\ell}} a_{j_{2\ell-1}} \dots a_{j_{\ell+1}} a_{i_{\ell-1}} a_{i_{\ell-2}} \dots a_{i_0},$$

where the sum is over n.b. color sequences $(i_0, i_1, \dots, i_\ell)$ and $(j_\ell, \dots, j_{2\ell})$ such that $i_\ell^* = j_\ell$. Taking the trace, we obtain

$$\mathbf{Trace} = \mathrm{tr}\Big(\boldsymbol{B}_n^\ell (\boldsymbol{B}_n^\ell)^*\Big) = \sum_{\substack{\text{special } (\ell+1)\text{-hikes } (u,\gamma) \\ \gamma = (i_0, \dots, i_\ell, j_\ell, \dots, j_{2\ell})}} \boldsymbol{\chi}[(u,\gamma)] \, \mathrm{tr}(a(\widehat{\gamma})),$$

where:

[11] We remark that this is the one place in this section where the inverse norm-bound $\|a_i^{-1}\|_{\mathrm{op}} \leqslant 2R$ is used; it takes care of a "fencepost error" in the natural proof of this proposition. In fact, a personal communication from the authors of [BC19] suggests that a variation of their argument can eliminate the need to bound the values $\|a_i^{-1}\|_{\mathrm{op}}$.

- the adjective "special" denotes that $i_\ell^* = j_\ell$ and $i_0^* = j_{2\ell}$;
- the notation $\chi[(u,\gamma)]$ denotes the product of the edge-signs from χ for each edge traversed by the hike (u,γ);[12]
- $\widehat{\gamma}$ denotes that the middle two elements i_ℓ and j_ℓ are omitted from γ.

Taking the expectation over χ kills all non-even hikes, and we conclude

$$\mathbf{E}[\mathbf{Trace}] = \sum_{\substack{\text{even, special} \\ (\ell+1)\text{-hikes } (u,\gamma)}} \operatorname{tr}(a(\widehat{\gamma})) \leqslant \sqrt{r} \sum_{\substack{\text{even, special} \\ (\ell+1)\text{-hikes } (u,\gamma)}} \|a(\widehat{\gamma})\|_{\mathrm{F}}. \tag{16}$$

Note that if we delete the first/last steps, and also the middle two steps, of a special $(\ell+1)$-hike (u,γ) (these two step-pairs both being a step and its reverse), we obtain an $(\ell-1)$-hike (v,γ'), where $v = \sigma_{i_0}(u)$. We have $a(\widehat{\gamma}) = a_{j_{2\ell}} a(\gamma') a_{i_0}$, and hence $\|a(\widehat{\gamma})\|_{\mathrm{F}} \leqslant (2R)^2 \|a(\gamma')\|_{\mathrm{F}}$. For each $(\ell-1)$-hike (v,γ'), there are at most $(d+2e)^2$ ways to add a step and its reverse to the beginning and the middle of the hike to get a special $(\ell+1)$-hike. Replacing also the condition "even" by the broader condition "singleton-free", we finally conclude:

$$\mathbf{E}[\mathbf{Trace}] \leqslant 4\sqrt{r} R^2 (d+2e)^2 \sum_{\substack{\text{singleton-free} \\ (\ell-1)\text{-hikes } (v,\gamma)}} \|a(\gamma)\|_{\mathrm{F}}. \tag{17}$$

Our goal is now to bound the right-hand side of Inequality (17) using only the λ-bicycle-free property of $\mathcal{G}_n$.

6.3 Elementary graph theory

We begin with some elementary graph theory. In the below definitions, by "graph" we mean an undirected multigraph, possibly with self-loops.

Definition 6.8. The *excess* of a graph H is $\operatorname{exc}(H) = |E(H)| - |V(H)|$.

Definition 6.9. Given a degree-2 vertex v in a graph H, *smoothing* v refers to the operation of deleting v from H and adding an edge between its two (former) neighbors. The *smoothing* of a graph H, which we will denote $\underline{H}$, is the graph obtained by iteratively smoothing degree-2 vertices (it is easy to see that the order does not matter). Each vertex in $\underline{H}$ has the same degree in $\underline{H}$ as it had in H; hence the vertex set of $\underline{H}$ consists of all vertices that originally had degree other than 2 in H. The edges in $\underline{H}$ correspond to paths in H whose internal nodes all had degree 2. We call these paths *stretches*. Finally, it is easy to see that $\operatorname{exc}(\underline{H}) = \operatorname{exc}(H)$.

We will also need the following lemma of elementary graph theorem from [MOP20a]:

Lemma 6.10. *([MOP20a, Thm. 2.13].) Let H be a k-vertex graph that is λ-bicycle free, where $\lambda \geqslant 10 \ln k$. Then $\operatorname{exc}(H) \leqslant \frac{\ln(ek)}{\lambda} \cdot k$.*

[12] Note that the first and last edge-sign do not actually arise in **Trace**, but it is okay that we have artificially inserted them: in any special hike, the first and last step are always along the same edge and hence contribute equal signs; thus their product is always 1.

Given an $(\ell-1)$-hike (v,γ), let $\mathcal{H} = \mathcal{H}_{(v,\gamma)}$ be the edge-colored subgraph of $\mathcal{G}_n$ induced by the edges on which the hike walks. This $\mathcal{H}$ is λ-bicycle free (since $\mathcal{G}_n$ is), and it has $k \leqslant 2\ell-2$ vertices. Provided C_2 and then C_1 are taken large enough given C_3, we will have

$$10\ln(2\ell-2) \leqslant \lambda.$$

(The left-hand side is $\Theta(\log\log n)$, the right-hand side is $\Theta((\log\log n)^2)$.) Thus Lemma 6.10 implies $\mathrm{exc}(\mathcal{H}) \leqslant \frac{\ln(ek)}{\lambda}\cdot k$. Due to the nonbacktracking nature of the hike, every vertex in $\mathcal{H}$ has degree at least 2, except possibly its initial vertex v and its midpoint vertex, which we call w. Abusing notation slightly, let us write $\underline{\mathcal{H}} = (\underline{\mathcal{V}}, \underline{\mathcal{E}})$ for the smoothing of $\mathcal{H}$ where we do *not* smooth v or w if they have degree 2. We still have $\mathrm{exc}(\underline{\mathcal{H}}) = \mathrm{exc}(\mathcal{H})$. Writing c_i for the number of vertices of degree i in $\underline{\mathcal{H}}$, we have

$$\begin{aligned}\frac{\ln(ek)}{\lambda}\cdot k \geqslant \mathrm{exc}(\mathcal{H}) = \mathrm{exc}(\underline{\mathcal{H}}) = |\underline{\mathcal{E}}| - |\underline{\mathcal{V}}| &= \tfrac{1}{2}(c_1 + 2c_2 + 3c_3 + \cdots) - (c_1 + c_2 + c_3 + \cdots)\\ &\geqslant -\tfrac{2}{3}c_1 - \tfrac{1}{3}c_2 + \tfrac{1}{3}\cdot\tfrac{1}{2}(c_1 + 2c_2 + 3c_3 + 4c_4 + \cdots)\\ &= -\tfrac{2}{3}c_1 - \tfrac{1}{3}c_2 + \tfrac{1}{3}|\underline{\mathcal{E}}|.\end{aligned}$$

We have $c_1 + c_2 \leqslant 2$ (since v, w are the only vertices in $\underline{\mathcal{H}}$ that may have degree smaller than 3). Hence:

Corollary 6.11. *For any $(\ell-1)$-hike (v,γ), the smoothing $\underline{\mathcal{H}}$ of $\mathcal{H}_{(v,\gamma)}$ has*

$$|\underline{\mathcal{E}}| \leqslant \tfrac{3\ln(ek)}{\lambda}\cdot k + 4 \leqslant 10\tfrac{\ln\ell}{\lambda}\cdot\ell.$$

(The latter inequality uses $\lambda \leqslant \frac{1}{4}\ell$, which holds provided C_2 and then C_1 are taken large enough given C_3; recall $\lambda = \Theta((\log\log n)^2)$ and $\ell = \Theta(\log n)$.) That is, $\mathcal{H}_{(v,\gamma)}$ consists of the vertices $\underline{\mathcal{V}}$, together with at most $10\frac{\ln\ell}{\lambda}\cdot\ell$ stretches (disjoint paths) linking these vertices.

6.4 Encoding and decoding

For the purpose of analyzing the right-hand side of Inequality (17), we will now introduce two algorithms called Encode and Decode. The Encode algorithm is inspired by the analysis in [BC19, Lem. 26], and is a significant elaboration of a similar kind of algorithm in [MOP20a]. It will take as input a singleton-free $(\ell-1)$-hike (v,γ), and will output v together with a few "data structures", which we now define.

Definition 6.12. An *outline data structure* Outline consists of:

- The *trek count* T: an integer between 1 and $50\frac{\ln\ell}{\lambda}\cdot\ell$.
- The *length data* $\ell_1,\ldots,\ell_T$: a sequence of positive numbers adding to $2(\ell-1)$.
- The *type data*: A subset of $[T]$ called $\mathcal{F}_1$ (the letter $\mathcal{F}$ stands for "fresh"), a disjoint subset of $[T]$ called $\mathcal{F}_2$ with $|\mathcal{F}_2| = |\mathcal{F}_1|$, and the remaining subset of $[T]$ called $\mathcal{S}$ ("stale"); thus $[T] = \mathcal{F}_1 \sqcup \mathcal{F}_2 \sqcup \mathcal{S}$.
- The *matching data*: this consists of a bijection $\tau : \mathcal{F}_2 \to \mathcal{F}_1$ such that $\tau(t) < t$ for each $t \in \mathcal{F}_2$ (that is, τ matches each $t \in \mathcal{F}_2$ with an "earlier" $t' \in \mathcal{F}_1$); and also, a mapping $\rho : \mathcal{F}_2 \to \{\text{forward}, \text{backward}\}$.

Definition 6.13. A *fresh data structure* FRESH consists of a list $C_1, \ldots, C_f$, where each C_i is an n.b. color sequence. We say that FRESH is *compatible* with an outline data structure OUTLINE (as above) if $f = |\mathcal{F}_1|$ and $|C_j| = \ell_{i_j}$ when the elements of $\mathcal{F}_1$ in increasing order are denoted $i_1, \ldots, i_f$.

Definition 6.14. A *stale data structure* STALE consists of a list $(r_1, h_1), \ldots, (r_s, h_s)$, where each $0 \leqslant r_i < 2(\ell - 1)$ is an integer and each h_i is in $\mathscr{I} \cup \{0\}$. We say that STALE is *compatible* with an outline data structure OUTLINE (as above) if $s = |\mathcal{S}|$.

6.4.1 Encoding

As mentioned, the ENCODE algorithm takes as input a singleton-free $(\ell - 1)$-hike (v, γ), and it outputs $(v, \text{OUTLINE}, \text{FRESH}, \text{STALE})$, where OUTLINE is an outline data structure and FRESH/STALE are compatible fresh/stale data structures (respectively). Recall the notation $\underline{\mathcal{H}} = (\underline{\mathcal{V}}, \underline{\mathcal{E}})$ for the smoothing of $\mathcal{H} = \mathcal{H}_{(v,\gamma)}$. The algorithm ENCODE analyzes how the hike traverses stretches from $\mathcal{H}$. Crucially, by the singleton-free property, every stretch is traversed (in either direction) *at least twice*.

To explain in detail how the ENCODE algorithm works, imagine a hiker walking along the edges of the hike (v, γ) from beginning to end. At various points the hiker will pause at some vertex in $\underline{\mathcal{V}}$, look at the upcoming edges of the hike, designate some initial portion of them one or more *treks*, record information about the trek(s) into the data structures (in particular, increasing the trek count T), then walk the trek(s) and pause again. The first pause occurs at the beginning of the hike before any edges have been traversed. In general, when the hiker has paused after completing $t - 1$ treks, the hiker looks ahead to the next stretch $\underline{e} \in \underline{\mathcal{E}}$ that will be traversed and performs the following three rules:

1. First traversals. If this will be the *first* time the hiker traverses $\underline{e}$ (in either direction), then it will be designated as the single next trek. Its length (number of edges) ℓ_t is appended to the length data of OUTLINE, and the trek index t is included into $\mathcal{F}_1$ in the type data. Finally, the n.b. color sequence C is appended to the fresh data structure FRESH.

2. Second traversals. If this will be the *second* time the hiker traverses $\underline{e}$, then again it will be designated the single next trek. Its length ℓ_t is appended to the length data of OUTLINE, and the trek index t is included into $\mathcal{F}_2$ in the type data. Furthermore, matching data is recorded: $\tau(t)$ is set to the earlier trek index on which $\underline{e}$ was traversed; and, $\rho(t)$ is set to "forward" if $\underline{e}$ will be traversed in the same direction as last time, and "backward" otherwise.

3. Third or higher traversals. If this will be the *third or higher* time the hiker traverses $\underline{e}$, the hiker begins preparing a "stale supertrek". The hiker looks further ahead at upcoming stretches to be traversed. All upcoming "third-time-or-higher" stretches $\underline{e}', \underline{e}'', \ldots$ are added to the stale supertrek, up to (but not including) the next stretch that will be a first-time or second-time traversal.

Next, the stale supertrek is broken down into multiple "stale subtreks" of length equal to the bicycle-free radius λ, with the necessary exception that the last of these subtreks allowed to have length less than λ. Note that while the first vertex of the first subtrek and the last vertex of the last subtrek will be in $\underline{\mathcal{V}}$, the intermediate starting/ending vertices may be any vertices in $\mathcal{H}$ (i.e., they may be "inside" a stretch from $\underline{\mathcal{E}}$). There is one additional proviso: if some stale subtrek encompasses the exact midpoint of the hike (i.e., both the $(\ell - 1)$th and the ℓth steps) then this

stale subtrek is further broken into two subtreks at this the midpoint. This proviso is included to ensure that the hiker traverses each stale subtrek in a nonbacktracking fashion.

Now the hiker designates these stale subtreks as the next few treks, and data about them is encoded into the data structures. First, their lengths $\ell_t, \ell_{t+1}, \ldots$ are added to Outline's length data. Next, their trek indices are included into the type data's $\mathcal{S}$. Finally, for each stale subtrek Q, an (r, h) pair is appended into the stale data structure Stale. The role of the (r, h) pair is to allow the Decode algorithm to reconstruct the stale subtrek Q. Let y be the vertex of $\mathcal{H}$ at which Q begins and let z be the vertex of $\mathcal{H}$ at which Q ends. We know the hike must have reached z at an earlier time (in fact, at least two earlier times). The parameter r is set to the first "time" (between 0 and $2(\ell-1)$) at which the hike reached z. As for defining h, let B be the subgraph of $\mathcal{H}$ induced by the radius-λ neighborhood of x. By the bicycle-free property, B contains at most one cycle. If there is such a cycle, and the trek walked along it for at least one step, then h is set to the color from $\mathscr{I}$ of the first such cycle-step. Otherwise, if the trek does not walk on B's cycle, or if B does not even have a cycle, then h is conveniently set to 0.

This completes the description of the Encode algorithm. Finally, we need to briefly verify two things about the Outline data structure output by Encode. First, as we noted, the singleton-free property implies that each stretch that is traversed is traversed at least twice; this means that the type data will indeed have $|\mathcal{F}_2| = |\mathcal{F}_1|$ and a well-defined bijection τ. Second, and more importantly, we need to verify that the trek count T will indeed be at most $50\frac{\ln \ell}{\lambda} \cdot \ell$ as promised. From Corollary 6.11 we know that there are at most $10\frac{\ln \ell}{\lambda} \cdot \ell$ stretches. These can contribute at most $20\frac{\ln \ell}{\lambda} \cdot \ell$ "fresh treks" in total. Let us now upper-bound the possible number of "stale treks" (ignoring the $+1$ we may get due to potentially traversing the hike's midpoint). Note that when the hiker produces a stale supertrek of length k, it gets decomposed into $\lceil k/\lambda \rceil \leqslant k/\lambda + 1$ stale subtreks. If we sum this quantity over all stale supertreks, we get

$$(\#\text{ of steps involved in stale supertreks})/\lambda + (\#\text{ of stale supertreks}) \leqslant 2(\ell-1)/\lambda + 20\tfrac{\ln \ell}{\lambda} \cdot \ell,$$

the last figure because each stale supertrek is preceded by a fresh trek. Now including the at most $20\frac{\ln \ell}{\lambda} \cdot \ell$ fresh treks, as well as the potential $+1$ for the midpoint, we get a total trek count of at most

$$20\tfrac{\ln \ell}{\lambda} \cdot \ell + 2(\ell-1)/\lambda + 20\tfrac{\ln \ell}{\lambda} \cdot \ell + 1 \leqslant 50\tfrac{\ln \ell}{\lambda} \cdot \ell,$$

as claimed. (Here we again used $\lambda \leqslant \frac{1}{4}\ell$.)

6.4.2 Decoding

We now describe the Decode algorithm, which has the following properties:

1. It takes as input a vertex v, an outline data structure Outline, and compatible fresh and stale data structures Fresh and Stale.
2. It either outputs a singleton set $\{(v, \gamma)\}$, where (v, γ) is an $(\ell-1)$-hike in $\mathcal{G}_n$; or, it outputs the empty set $\varnothing$ ("invalid").
3. If $(v, \text{Outline}, \text{Fresh}, \text{Stale})$ is the output of the Encode algorithm applied to some singleton-free $(\ell-1)$-hike (v, γ) in $\mathcal{G}_n$, then Decode will output $\{(v, \gamma)\}$.

In other words, a singleton-free hike may be recovered from the data output by Encode. We will not explicitly prove Item 3 above; rather, it will become clear as we describe the Decode algorithm. Also, implicit in our description is that whenever the Decode algorithm cannot continue,

or is required to do something nonsensical, or produces an invalid singleton-free $(\ell-1)$-hike, it simply outputs $\varnothing$.

The Decode algorithm will (attempt to) recover a singleton-free hike $\{(v,\gamma)\}$ in a trek-by-trek fashion from its input data. The algorithm will always maintain a "current vertex" from $\underline{\mathcal{V}}$. The initial value of the "current vertex" is of course the vertex v which is part of Decode's input. For $t = 1 \dots T$, the Decode algorithm recovers the tth trek by first considering whether t is in $\mathcal{F}_1$, $\mathcal{F}_2$, or $\mathcal{S}$. If $t \in \mathcal{F}_1$, then Decode will take the first unused n.b. color sequence C from Fresh (which by compatibility will have length ℓ_t) and will follow C from the current vertex. The final vertex reached becomes the new "current vertex". At this point, the Decode algorithm will have "learned" a new stretch $\underline{e} \in \underline{\mathcal{E}}$, which it may associate to this tth trek.

If $t \in \mathcal{F}_2$, then Decode will first let $t' = \tau(t) \in \mathcal{F}_1$. Since $t' < t$, the Decode algorithm will have already learned the stretch $\underline{e} \in \underline{\mathcal{E}}$ associated to t'. The Decode algorithm can therefore typically recover the tth trek: it just follows $\underline{e}$ again. If $\underline{e}$ is a self-loop in $\underline{\mathcal{H}}$, there is an ambiguity about the direction in which the trek traverses $\underline{e}$, but this is resolved using $\rho(t)$ from the matching data.

Finally, if $t \in \mathcal{S}$, then Decode will recover the stale trek with the aid of the associated stale data (r,h). Let y denote the current vertex (i.e., the initial vertex for the stale trek) and ℓ_t the trek length. As this trek is stale, it will take place entirely within a subgraph of $\mathcal{H}$ that the Decode algorithm has already "learned". In fact, it will take place within a further subgraph B, the one of radius ℓ_t centered at y. The Decode algorithm can infer the ending vertex $z \in B$ of the trek using the stale datum r; the rth vertex visited so far on the hike is z. Next, since $\ell_t \leqslant \lambda$ (in the properly encoded case), B will have at most one cycle. It is not hard to see that in this unicyclic B, a nonbacktracking walk from y to z of length ℓ_t is *almost* uniquely defined. The only possible ambiguity that may arise comes from the direction in which B's cycle (should it exist) is traversed. This ambiguity is resolved using the stale datum h; it supplies the first color used by the trek when the cycle is entered.

This completes the description of the Decode algorithm.

6.5 Counting

We now bound the right-hand side of Inequality (17). By virtue of the Item 3 property of the Decode algorithm, we have

$$\sum_{\substack{\text{singleton-free} \\ (\ell-1)\text{-hikes } (v,\gamma)}} \|a(\gamma)\|_{\mathrm{F}} \leqslant \sum_{v,\ \textsc{Outline}} \sum_{\text{compatible } \textsc{Fresh},\ \textsc{Stale}} \sum_{(v,\gamma)\in \textsc{Decode}(v,\textsc{Outline},\textsc{Fresh},\textsc{Stale})} \|a(\gamma)\|_{\mathrm{F}} \tag{18}$$

where on the right-hand side we sum over all possible v, Outline, then all possible Fresh, Stale compatible with Outline, then all hikes in the set Decode$(v,$ Outline, Fresh, Stale$)$ (which set, recall, either has cardinality 0 or 1). A key property of this new summation is that there is a natural way to decompose and bound $\|a(\gamma)\|_{\mathrm{F}}$ as a function of Outline. Given the outline data structure Outline with trek count T and length data $\ell_1, \dots, \ell_T$, we can naturally break up any length-$2(\ell-2)$ color sequence γ into parts $\gamma_1, \dots, \gamma_T$ of length $\ell_1, \dots, \ell_T$. Then by submultiplicativity of norms we have

$$\|a(\gamma)\|_{\mathrm{F}} \leqslant \|a(\gamma_1)\|_{\mathrm{F}} \cdots \|a(\gamma_T)\|_{\mathrm{F}}.$$

We then bound each of these factors depending on the type data in Outline. For a factor $\|a(\gamma_i)\|_{\mathrm{F}}$ with i in the stale subset $\mathcal{S}$, we use Proposition 6.6 to bound it by $c\overline{\rho}^{\ell_i}$.

The remaining factors may be naturally grouped into pairs according to Outline's matching data. Suppose that $t \in \mathcal{F}_1$, $t' \in \mathcal{F}_2$, and $\tau(t') = t$. It is a property of the decoding algorithm that whenever $(v, \gamma) \in \textsc{Decode}(v, \textsc{Outline}, \textsc{Fresh}, \textsc{Stale})$, we have either $\gamma_t = \gamma_{t'}$ or $\gamma_t = \gamma_{t'}^*$ (depending on $\rho(t')$). Recalling Definition 6.4 we see that either way, $\|a(\gamma_t)\|_{\mathrm{F}} = \|a(\gamma_{t'})\|_{\mathrm{F}}$. Furthermore, this norm *only* depends on the fresh data structure Fresh; specifically, if t is the jth element of $\mathcal{F}_1$ (in increasing order), then $\|a(\gamma_t)\|_{\mathrm{F}} = \|a(\gamma_{t'})\|_{\mathrm{F}} = \|a(C_j)\|_{\mathrm{F}}$. As a consequence, for a given $(v, \textsc{Outline}, \textsc{Fresh}, \textsc{Stale})$ the following bound holds:

$$\sum_{(v,\gamma)\in\textsc{Decode}(v,\textsc{Outline},\textsc{Fresh},\textsc{Stale})} \|a(\gamma)\|_{\mathrm{F}} \leqslant \left(\prod_{i\in\mathcal{S}} c\overline{\rho}^{\ell_i}\right)\left(\prod_{j=1}^{f}\|a(C_j)\|_{\mathrm{F}}^2\right) = c^{|\mathcal{S}|}\overline{\rho}^{\ell_{\mathcal{S}}}\cdot\prod_{j=1}^{f}\|a(C_j)\|_{\mathrm{F}}^2,$$

where we have written $\ell_{\mathcal{S}} = \sum_{i\in\mathcal{S}} \ell_i$ for the number of stale steps. Notice that the above bound does not depend on v, Decode or, on Stale, just on Outline and Fresh. Substituting it into Inequality (18) yields

$$\sum_{\substack{\text{singleton-free}\\ (\ell-1)\text{-hikes } (v,\gamma)}} \|a(\gamma)\|_{\mathrm{F}} \leqslant n\cdot\sum_{\textsc{Outline}} c^{|\mathcal{S}|}\overline{\rho}^{\ell_{\mathcal{S}}}\cdot\#\{\text{compatible }\textsc{Stale}\}\cdot\sum_{\text{compatible }\textsc{Fresh}}\prod_{j=1}^{f}\|a(C_j)\|_{\mathrm{F}}^2. \quad (19)$$

Now for a fixed Outline we can interchange the final sum and product in the above:

$$\begin{aligned}\sum_{\text{compatible }\textsc{Fresh}}\prod_{j=1}^{f}\|a(C_j)\|_{\mathrm{F}}^2 &= \prod_{t\in\mathcal{F}_1}\sum_{\text{n.b. color sequences } C \text{ of length } \ell_t}\|a(C)\|_{\mathrm{F}}^2\\ &\leqslant \prod_{t\in\mathcal{F}_1}(c\overline{\rho}^{2\ell_t}) = c^{|\mathcal{F}_1|}\overline{\rho}^{\ell_{\mathcal{F}}},\end{aligned}$$

where the inequality is Proposition 6.6 and we have written $\ell_{\mathcal{F}} = 2\sum_{i\in\mathcal{F}_1}\ell_i = \sum_{i\in\mathcal{F}_1\sqcup\mathcal{F}_2}\ell_i$. Putting this into Inequality (19) yields

$$\begin{aligned}\sum_{\substack{\text{singleton-free}\\ (\ell-1)\text{-hikes } (v,\gamma)}}\|a(\gamma)\|_{\mathrm{F}} &\leqslant n\cdot\sum_{\textsc{Outline}} c^{|\mathcal{F}_1|+|\mathcal{S}|}\overline{\rho}^{\ell_{\mathcal{F}}+\ell_{\mathcal{S}}}\cdot\#\{\text{compatible }\textsc{Stale}\}\\ &\leqslant n\cdot c^T\cdot\overline{\rho}^{2(\ell-1)}\cdot\#\{\text{compatible }(\textsc{Outline},\textsc{Stale})\}.\end{aligned}$$

It remains to bound the number of compatible (Outline, Stale) pairs, and it suffices to do this very crudely. For a given choice of T, the number of possibilities for the length data is at most $(2\ell)^T$. The number of possibilities for the type data is at most 3^T. The number of choices for the matching data is at most $(2T)^T$. And the number of choices for compatible stale data is at most $(2D\ell)^T$ for $D = d + 2e$. Thus we have at most $(24D\ell T)^T$ possibilities, given T. Recalling $T \leqslant 50\frac{\ln\ell}{\lambda}\ell$ and also Inequality (17), we finally conclude

$$\mathbf{E}[\mathbf{Trace}] \leqslant n\cdot\exp\left(O\left(\tfrac{\log^2\ell}{\lambda}\ell\right)\right)\cdot\overline{\rho}^{2\ell}, \quad (20)$$

where we have now begun using $O(\cdot)$ notation to hide constants depending on d, e, r, R, $\bar{\varepsilon}$, p. Markov's inequality now implies that except with probability at most n^{-2p},

$$\|\boldsymbol{B}_n\|_{\mathrm{op}}^{2\ell} \leqslant \mathrm{tr}\left(\boldsymbol{B}_n^{\ell}(\boldsymbol{B}_n^{\ell})^*\right) = \mathbf{Trace} \leqslant n^{2p+1}\cdot\exp\left(O\left(\tfrac{\log^2\ell}{\lambda}\ell\right)\right)\cdot\overline{\rho}^{2\ell}, \quad (21)$$

and in this case

$$\|\boldsymbol{B}_n\|_{\mathrm{op}} \leqslant \exp\Big(O\Big(\tfrac{\log n}{\ell} + \tfrac{\log^2 \ell}{\lambda}\Big)\Big) \cdot \overline{\rho} \leqslant \overline{\rho} + O\Big(\tfrac{\log n}{\ell}\Big) + O\Big(\tfrac{\log^2 \ell}{\lambda}\Big) \tag{22}$$

(where the second inequality used that $\overline{\rho} \leqslant O(1)$ thanks to Corollary 4.7). By taking C_3, then C_2, then C_1 large enough, this shows Inequality (14) holds for our fixed choice of $\mathcal{K}$, except with probability at most n^{-2p}. This completes the proof of Theorem 6.5.

Corollary 6.15. *In the setting of Theorem 6.5, suppose the edge signs $\boldsymbol{\chi}$ are not independent and uniformly random, but are merely drawn from a $(\delta, 2\ell)$-wise uniform distribution, where $\delta = 1/n^{C_0}$ for a sufficiently large constant $C_0 = C_0(d, e, r, R, \bar{\varepsilon}, p)$, and where $\ell = \lceil C_3 \log n \rceil$ as in Equation (15). Then the conclusion of the theorem continues to hold.*

Proof. The only place where the independence of the edge signs was used was in Equation (16), to say that non-even hikes dropped out of the expectation. When $\boldsymbol{\chi}$ is merely $(\delta, 2\ell)$-wise uniform, the inequality therein is not exact; however, since **Trace** is a degree-2ℓ polynomial in $\boldsymbol{\chi}$, the error is (Fact 5.2) at most an additive

$$\delta \cdot \sum_{\text{special } (\ell+1)\text{-hikes } (u,\gamma)} |\mathrm{tr}(a(\widehat{\gamma}))| \leqslant \delta \cdot n(d+2e)^{2\ell} \cdot r(2R)^{2\ell},$$

where we used naive upper bounds on the number of special $(\ell+1)$-hikes and on the absolute trace of the hike weight (the latter using $\mathcal{K} \in \mathcal{A}_R$). In turn, by choosing the constant C_0 large enough (in particular, large enough vis-a-vis C_3), we can ensure this error is much smaller than the right-hand side of Inequality (20) — for concreteness, say at most half of it. (We use here the naive lower bound $\overline{\rho} \geqslant \bar{\varepsilon}/6$.) Thus the subsequent bound Inequality (21) is affected only up to a constant factor, and in the final bound Inequality (22) only the hidden constant in the $O(\frac{\log n}{\ell})$ is affected. □

7 Weakly derandomizing *n*-lifts

In this section we fix the same notation and terminology as in the beginning of Section 6. We remark that except for the analysis of the net, this section is not substantially different from the analogous [MOP20a, Sec. 4].

7.1 Derandomizing simple arguments about cycles

In the course of proving their main theorem on random n-lifts, Bordenave and Collins need to show that the graph underlying a random n-lift is bicycle-free for radius $\lambda = \Omega(\log n)$ and also that it has at least one vertex with an acyclic λ-neighborhood. These are straightforward proofs, and it is also straightforward to see that they are weakly derandomizable using $\Theta(\log n)$-wise uniform permutations; indeed, this was essentially already done in [MOP20a, Prop. 4.3]. As the exact statement needed is not literally in [MOP20a, Prop. 4.3], we provide a statement and briefly sketch a proof here.

Proposition 7.1. *Given constants d, e, there are constants C_0, C_1 such that the following holds: Assume $n, \lambda \in \mathbb{N}^+$ satisfy $n \geqslant C_0$ and $\lambda \leqslant (\log n)/C_1$. Then for a 4λ-wise uniform n-lift $\mathfrak{L}_n$, except with probability at most $n^{-.999}$ we have both of the following:*

1. *$G_{\mathfrak{L}_n}$ has at least one vertex whose radius-λ neighborhood is acyclic;*

2. *$G_{\mathfrak{L}_n}$ is λ-bicycle free.*

Proof. We will mainly rely on [BC19, Lem. 23]. For a fixed graph H on m edges, let $\boldsymbol{X}_H$ denote the number of copies of H in $G_{\mathfrak{L}_n}$. In [BC19, Lem. 23], the quantity $\mathbf{E}[\boldsymbol{X}_H]$ is computed (and then bounded) assuming the lift $\mathfrak{L}_n$ is truly uniformly random. Now it is easy to see that $\boldsymbol{X}_H$ is a polynomial of degree m in the indicator random variables $(1[\sigma_i(j) = k])_{i\in\mathscr{G}}^{j,k=1\dots n}$ for $\mathfrak{L}_n$. Thus Fact 5.10 tells us, whenever $m \leqslant 4\lambda$, that $\mathbf{E}[\boldsymbol{X}_H]$ has the same value under our 4λ-wise uniform $\mathfrak{L}_n$ as it has in [BC19, Lem. 23].

In [BC19, Lem. 23], the following is first shown: provided $\ell \leqslant \sqrt{n}$, the expected number of cycles of length ℓ in $G_{\mathfrak{L}_n}$ is at most $O((d+2e)^\ell)$. Using this fact for all $\ell \leqslant 2\lambda$, it is directly derived in [BC19, Cor. 8] that the expected number of vertices $v \in G_{\mathfrak{L}_n}$ that have a cycle in their λ-neighborhood is at most $O((d+2e)^{3\lambda})$. Now for sufficiently large constants C_0, C_1 we will have $2\lambda \leqslant \sqrt{n}$ and $O((d+2e)^{3\lambda}) \leqslant \frac{1}{2}n^{.001}$. Thus we conclude this expectation is at most $\frac{1}{2}n^{.001}$ under our 4λ-wise uniform $\mathfrak{L}_n$, and hence by Markov, Item 1 holds except with probability at most $\frac{1}{2}n^{-.999}$.

Regarding Item 2, it is noted in [BC19, Lem. 23] that if $G_{\mathfrak{L}_n}$ is *not* λ-bicycle free then it must contain a "witness subgraph" to this fact (either a "handcuffs graph" or a "theta graph") of at most 4λ edges. Provided $4\lambda \leqslant \sqrt{n}$, they bound the expected number of such witness graphs (for truly uniform $\mathfrak{L}_n$) by $O(\lambda^3(d+2e)^{4\lambda}/n)$. Similar to before, for sufficiently large constants C_0, C_1 we will have $4\lambda \leqslant \sqrt{n}$ and $O(\lambda^3(d+2e)^{4\lambda}/n) \leqslant \frac{1}{2}n^{-.999}$. Thus we conclude the expected number of witness graphs is at most $\frac{1}{2}n^{-.999}$ even under our 4λ-wise uniform $\mathfrak{L}_n$, and thus the probability of $G_{\mathfrak{L}_n}$ not being λ-bicycle free is at most $\frac{1}{2}n^{-.999}$. □

7.2 Weakly derandomizing the main argument

In this setting, let us restate (a slight variant of) Bordenave and Collins's main theorem on the nonbacktracking operator of random n-lifts [BC19, Thm. 17]:

Theorem 7.2. *Let d, e, r, R, $\bar{\varepsilon} > 0$ be constants. Then there are constants C_0, $C_1 > 0$ such that the following holds:*

Let $n \geqslant C_0$, let $\lambda = \lfloor(\log n)/C_1\rfloor$, and let $\Xi_\lambda \subset \mathfrak{K}_{2R}$ be as in Proposition 6.2, with $\varepsilon = \bar{\varepsilon}/3$. Then except with probability at most $n^{-.99}$ over the choice of a uniformly random n-lift $\mathfrak{L}_n$, the following all hold: Items 1 and 2 of Proposition 7.1, and also for all matrix bouquets $\mathcal{K} \in \Xi_\lambda$,

$$\left\| B_{n,\perp}(\mathfrak{L}_n, \mathcal{K})^\lambda \right\|_{\mathrm{op}} \leqslant (\rho(B_\infty(\mathcal{K})) + \bar{\varepsilon}/3)^\lambda. \tag{23}$$

When this last event happens we may easily deduce from Proposition 6.2 that

$$\rho(B_{n,\perp}(\mathfrak{L}_n, \mathcal{K}))^\lambda \leqslant \left\| B_{n,\perp}(\mathfrak{L}_n, \mathcal{K})^\lambda \right\|_{\mathrm{op}} \leqslant (\rho(B_\infty(\mathcal{K})) + \bar{\varepsilon})^\lambda \tag{24}$$

holds simultaneously for all R-bounded bouquets $\mathcal{K} \in \mathfrak{K}_R$.

Our goal in this section is to show the following:

Theorem 7.3. *There is a constant $C = C(d, e, r, R, \bar{\varepsilon})$ such that Theorem 7.2 continues to hold only assuming that $\mathfrak{L}_n$ is $(C \log n)$-wise uniform.*

The key insight is that the only probabilistic component of Theorem 7.2 (besides Proposition 7.1) is a bound ([BC19, Prop. 22]) on the expectation of certain polynomials of degree $O(\log n)$

in the lift's indicator random variables $1[\sigma_i(j) = k]$. Unfortunately, to carefully verify this requires recapping details from [BC19] in a somewhat black-box fashion.

We start by remarking that the way [BC19, Thm. 17] is originally stated only involves the (λth root of) Inequality (24); i.e., it only involves bounding $\rho(B_{n,\perp}(\mathfrak{L}_n, \mathcal{K}))$. However, their proof immediately begins (see [BC19, (27)]) by passing to bounding $\left\|B_{n,\perp}(\mathfrak{L}_n, \mathcal{K})^\lambda\right\|_{\mathrm{op}}$ for the given choice of λ (which they call "ℓ"). The bulk of the work is of course to bound $\left\|B_{n,\perp}(\mathfrak{L}_n, \mathcal{K})^\lambda\right\|_{\mathrm{op}}$ for a *fixed* $\mathcal{K} = (a_1, \dots, a_{d+2e}) \in \mathfrak{K}_{2R}$ (where our $2R$ essentially corresponds to their "ε^{-1}"). After this, the claim for all $\mathcal{K} \in \Xi_\lambda$ follows from a union bound. Indeed, for fixed $\mathcal{K} \in \mathfrak{K}_{2R}$ we will be able to show that Inequality (23) holds except with probability $O(n^{-.999})$, provided C_0, C_1 are large enough — even for $(C \log n)$-wise uniform $\mathfrak{L}_n$. Since $|\Xi_\lambda| \leqslant \kappa(d, e, r, R, \bar{\varepsilon})^\lambda$, by increasing C_1 (and C_0) if necessary, we can ensure that $|\Xi_\lambda| \cdot O(n^{-.999}) \leqslant n^{-.99}$, completing the proof.

Summarizing, our goal is to show:

$$\begin{aligned}&\text{For fixed } \mathcal{K} = (a_1, \dots, a_{d+2e}) \in \mathfrak{K}_{2R},\\ &\text{Inequality (23) holds except with probability } O(n^{-.999}) \text{ when } \mathfrak{L}_n \text{ is } (C \log n)\text{-wise uniform.}\end{aligned} \tag{25}$$

The operator $B_{n,\perp}(\mathfrak{L}_n, \mathcal{K})^\lambda$ occurring in Inequality (23) is hard to control directly. But using a key idea from [Bor19], it is shown in [BC19, Sec. 4.2] that there are certain related operators $\underline{\boldsymbol{B}}^{(\lambda)}$ and $\boldsymbol{R}_k^{(\lambda)}$ (defined in [BC19, (32), (34)]) such that the following holds: *if* $G_{\mathfrak{L}_n}$ is λ-bicycle free, then

$$\left\|B_{n,\perp}(\mathfrak{L}_n, \mathcal{K})^\lambda\right\|_{\mathrm{op}} \leqslant \|\underline{\boldsymbol{B}}^{(\lambda)}\|_{\mathrm{op}} + \frac{1}{n}\sum_{k=1}^{\lambda} \|\boldsymbol{R}_k^{(\lambda)}\|_{\mathrm{op}}.$$

By Proposition 7.1 (and using $C \log n \gg 4\lambda$), we may indeed assume $G_{\mathfrak{L}_n}$ is λ-bicycle free for the purposes of proving Equation (25) (and Proposition 7.1 also establishes that Items 1 and 2 also hold with the required probability, as needed for Theorem 7.3). Thus, writing $\rho^* = \rho(B_\infty(\mathcal{K}))$, it remains to show that

$$\|\underline{\boldsymbol{B}}^{(\lambda)}\|_{\mathrm{op}} + \frac{1}{n}\sum_{k=1}^{\lambda} \|\boldsymbol{R}_k^{(\lambda)}\|_{\mathrm{op}} \leqslant (\rho^* + \bar{\varepsilon}/3)^\lambda \tag{26}$$

except with probability $O(n^{-.999})$ for $(C \log n)$-wise uniform $\mathfrak{L}_n$.[13]

As mentioned earlier, the only probabilistic component of Bordenave–Collins's proof of this (for truly uniform $\mathfrak{L}_n$) is the below proposition (a slightly restated version of [BC19, Prop. 22]); we leave some of the terminology in it undefined.

Proposition 7.4. *([BC19, Prop. 22], following directly from [Bor19, Props. 11, 28].) For a fixed sequence $f = (f_1, \dots, f_\tau)$ of potential "colored edges", $f_t = (x_t, y_t) \in [n] \times [n]$ of color $i_t \in \mathscr{I}$, and assuming $\tau_0 \leqslant \tau \leqslant \sqrt{n}$, if $\mathfrak{L}_n = (\mathbb{1}, \boldsymbol{\sigma}_1, \dots, \boldsymbol{\sigma}_{d+2e})$ is a random lift, then*

$$\left|\mathbf{E}\left[\prod_{t=1}^{\tau_0}(1[\sigma_{i_t}(x_t) = y_t] - 1/n) \prod_{t=\tau_0+1}^{\tau} 1[\sigma_{i_t}(x_t) = y_t]\right]\right| \leqslant c2^b(1/n)^m(3\tau/\sqrt{n})^{m_1},$$

where c is a universal constant, b is the number of "inconsistent" edges in f, m is the number of distinct edges in f, and m_1 is the number of "consistent edges" of multiplicity 1 in f.

The only thing we need to take away from this is the following:

[13] In fact, we can show it holds except with probability n^{-100}. The bottleneck for showing low failure probability is the probability of $G_{\mathfrak{L}_n}$ failing to be λ-bicycle free.

Fact 7.5. *Proposition 7.4 holds equally well if $\mathfrak{L}_n$ is merely τ-wise uniform, by Fact 5.10.*

Bordenave and Collins only ever apply Proposition 7.4 (twice) with $\tau = 2\lambda m$, where "m" is an integer parameter they eventually set to $\Theta((\log n)/(\log\log n))$. To save on the amount of t-wise uniformity needed, though, we will eventually set m to be a large constant depending on $d, e, r, R, \bar{\varepsilon}$, and also C_1. With this setting, we will indeed have $\tau = 2\lambda m \leqslant \sqrt{n}$ (provided C_0 is large enough) and also that $(C\log n)$-wise uniformity of $\mathfrak{L}_n$ suffices.

Inspecting the proof of [BC19, Props. 24, 28], particularly [BC19, (52), (56)][14], we see that it can use Proposition 7.4 to deduce (for a constant $c = c(d, e, r, R, \bar{\varepsilon})$) that

$$\mathbf{E}[\|\underline{\boldsymbol{B}}^{(\lambda)}\|_{\mathrm{op}}^{2m}] \leqslant n(c\lambda m)^{10m} \cdot (\rho^* + \bar{\varepsilon}/4)^{2\lambda m}, \qquad \mathbf{E}[\|\underline{\boldsymbol{R}}_k^{(\lambda)}\|_{\mathrm{op}}^{2m}] \leqslant (c\lambda m)^{32m} \cdot c^{2\lambda m},$$

provided $n \geqslant n_0(d, e, r, R, \bar{\varepsilon})$ and $\lambda \leqslant \log n$ (both of which we are assuming) and also provided $(c\lambda m)^{12m}/n \leqslant 1/2$ (which also holds given our choice of m, assuming C_0 is large enough). Thus by Markov's inequality, except with probability at most $(\lambda+1)n^{-1} = O(n^{-.999})$ over the choice of $(C\log n)$-wise uniform $\mathfrak{L}_n$, we have all of

$$\|\underline{\boldsymbol{B}}^{(\lambda)}\|_{\mathrm{op}}^{2m} \leqslant n^2(c\lambda m)^{10m} \cdot (\rho^* + \bar{\varepsilon}/4)^{2\lambda m}, \qquad \|\underline{\boldsymbol{R}}_k^{(\lambda)}\|_{\mathrm{op}}^{2m} \leqslant n(c\lambda m)^{32m} \cdot c^{2\lambda m} \quad \forall k \leqslant \lambda$$

$$\implies \|\underline{\boldsymbol{B}}^{(\lambda)}\|_{\mathrm{op}} \leqslant n^{1/m}(c\lambda m)^5 \cdot (\rho^* + \bar{\varepsilon}/4)^{\lambda} \qquad \implies \frac{1}{n}\sum_{k=1}^{\lambda} \|\underline{\boldsymbol{R}}_k^{(\lambda)}\|_{\mathrm{op}} \leqslant n^{1/(2m)}(c\lambda m)^{16} \cdot \frac{\lambda c^{\lambda}}{n}$$

$$= \left(n^{1/(\lambda m)}(c\lambda m)^{5/\lambda}\right)^{\lambda} \cdot (\rho^* + \bar{\varepsilon}/4)^{\lambda}.$$

Now for any constant $\delta = \delta(d, e, r, R, \bar{\varepsilon}) > 0$, and given the constant C_1, we can ensure $\|\underline{\boldsymbol{B}}^{(\lambda)}\|_{\mathrm{op}} \leqslant (1+\delta)^{\lambda} \cdot (\rho^* + \bar{\varepsilon}/4)^{\lambda}$ by taking m a sufficiently large constant (and C_0 large enough). Further, we can ensure $\frac{1}{n}\sum_{k=1}^{\lambda} \|\underline{\boldsymbol{R}}_k^{(\lambda)}\|_{\mathrm{op}} \leqslant \delta^{\lambda}$ by taking C_1 large enough; then irrespective of m's constant value, it suffices to take C_0 large enough. We conclude that by setting constants appropriately, we can ensure

$$\|\underline{\boldsymbol{B}}^{(\lambda)}\|_{\mathrm{op}} + \frac{1}{n}\sum_{k=1}^{\lambda} \|\underline{\boldsymbol{R}}_k^{(\lambda)}\|_{\mathrm{op}} \leqslant (1+\delta)^{\lambda}(\rho^* + \bar{\varepsilon}/4)^{\lambda} + \delta^{\lambda},$$

which is enough for Inequality (26) if δ is taken small enough (recalling that Corollary 4.7 upper-bounds ρ^* by a constant).

8 Constructing explicit good lifts

We remark that except for the analysis of the net, the below theorem is not substantially different from its analogue in [MOP20a, Sec. 5].

Theorem 8.1. *Fix an index set $\mathscr{I} = \{1, \ldots, d+2e\}$. Fix also constants $r \in \mathbb{N}^+$ and $R, \varepsilon > 0$. Then for any constant $C' > 0$, there is a $\mathrm{poly}(N)$-time deterministic algorithm that, on input N, outputs an N'-lift $\mathfrak{L}_{N'}$ (with $N \leqslant N' \leqslant N + o(N)$) such that:*

- *$G_{\mathfrak{L}_{N'}}$ has at least one vertex whose radius-$C'\sqrt{\log N}$ neighborhood is acyclic;*
- *for every R-bounded bouquet $\mathcal{K} \in \mathfrak{K}_R$ with r'-dimensional matrix-weights ($r' \leqslant r$), we have*

$$\rho(B_{N',\perp}(\mathfrak{L}_{N'}, \mathcal{K})) \leqslant \rho(B_{\infty}(\mathcal{K})) + \varepsilon.$$

[14] Though (52) has a typo; its exponent on "ρ" should be "$2\ell m$" not "ℓm".

Proof. We may assume R and ε are rational, and it suffices to assume (as we will tacitly do a couple of times) that N is at least a sufficiently large constant. We will prove the result just for r-dimensional bouquets; the fact that it works simultaneously for all $r' \leqslant r$ is a small twist we will comment on at the end of the proof.

Given N, the algorithm will choose an even integer n_0 on the order of $2^{\Theta(\sqrt{\log N})}$ and then define $n_i = 2^i n_0$ for each $1 \leqslant i \leqslant t = \lceil \log_2(N/n_0) \rceil$. The value N' will be equal to n_t; it is clear that $N' \geqslant N$, and it is easy for the algorithm to arrange that $N' \leqslant N(1 + 2^{-\Theta(\sqrt{\log N})})$ by adjusting n_0's value by at most a factor of 2. The idea is that the algorithm will sequentially produce n_i-lifts $\mathfrak{L}_{n_i}$ for $i = 0 \dots t$. The 0th one $\mathfrak{L}_{n_0}$ will be produced using Theorem 7.3 and each subsequent $\mathfrak{L}_{n_i} = \mathfrak{L}_{2n_{i-1}}$ will be produced via Corollary 6.15 as an edge-signing of $\mathfrak{L}_{n_{i-1}}$ (recalling from Section 4.3 that edge-signing a lift corresponds to forming a product with 2-lifts). As these lifts are formed, the algorithm will maintain the following invariants:

1. $G_{\mathfrak{L}_{n_i}}$ has at least one vertex whose radius-$C'\sqrt{\log N}$ neighborhood is acyclic;
2. $G_{\mathfrak{L}_{n_i}}$ is $C'\sqrt{\log N}$-bicycle free;
3. for every R-bounded bouquet $\mathcal{K} \in \mathfrak{K}_R$ with r-dimensional matrix-weights we have

$$\rho(B_{n_i,\perp}(\mathfrak{L}_{n_i}, \mathcal{K})) \leqslant \rho(B_\infty(\mathcal{K})) + \varepsilon.$$

Note that as soon as invariant Items 1 and 2 hold for n_0, they automatically hold for all subsequent n_i. The reasons are that the 2-lift of any an acyclic radius-λ neighborhood in $G_{\mathfrak{L}_{n_{i-1}}}$ becomes two acyclic radius-λ neighborhoods in $G_{\mathfrak{L}_{n_i}}$; and, every 2-lift of a λ-bicycle free graph is λ-bicycle free ([MOP20a, Prop. 2.12]). Thus it suffices for the algorithm to ensure that all three invariants hold for the 0th lift $\mathfrak{L}_{n_0}$, and then that the 3rd invariant holds for all subsequent lifts.

The 0th lift. As mentioned, the 0th lift $\mathfrak{L}_{n_0}$ is constructed via Theorem 7.3, with its "$\bar{\varepsilon}$" parameter set to $\varepsilon/2$ and its "n" parameter set to $n_0 = 2^{\Theta(\sqrt{\log N})}$ large enough so that the resulting value of λ is at least $C'\sqrt{\log N}$ (and so that $n_0 \geqslant C_0$). The theorem tells us that if $\mathfrak{L}_{n_0}$ is a t-wise uniform n_0-lift, $t = C \log n_0 = O(\sqrt{\log N})$, then except with probability at most $n_0^{-.99} = 2^{-\Theta(\sqrt{\log N})}$, invariant Items 1 and 2 hold, as does the following:

$$\forall \mathcal{K} \in \Xi_\lambda, \qquad \left\| B_{n_0,\perp}(\mathfrak{L}_{n_0}, \mathcal{K})^\lambda \right\|_{\mathrm{op}} \leqslant (\rho(B_\infty(\mathcal{K})) + \varepsilon/6)^\lambda. \tag{27}$$

As noted in Theorem 7.2, Inequality (27) has invariant Item 3 as a consequence (even if it only has $\varepsilon/3$ in place of $\varepsilon/6$). A crucial point is that a deterministic poly(N)-time algorithm can *check*, given a realization of $\mathfrak{L}_{n_0}$, that invariant Items 1 and 2 hold (just use breadth-first search), and it can also *check* that Inequality (27) holds (with $\varepsilon/3$ in place of $\varepsilon/6$). To see this last claim, note first from Proposition 6.2 that the bouquets $\mathcal{K}$ in the net Ξ_λ can be enumerated in poly(N) time (in fact, in subpolynomial time, $O(1)^\lambda = 2^{O(\sqrt{\log N})}$). Next note that each matrix $B = B_{n_0,\perp}(\mathfrak{L}_{n_0}, \mathcal{K})^\lambda$ can straightforwardly be computed exactly, as can be $S = B^*B$. Observe that $\|B\|_{\mathrm{op}} \leqslant \beta$ iff $\|S\|_{\mathrm{op}} \leqslant \beta^2$ iff $\beta^2 \mathbb{1} - S \succeq 0$, and the last condition (testing if a Hermitian matrix is positive semidefinite) can be checked efficiently [GLS93, pg. 295] for any given β. Finally, note that each $\rho(B_\infty(\mathcal{K}))$ can be efficiently computed to additive accuracy $\varepsilon/6$, by Proposition 4.6 (in constant time, in fact). Having established that Inequality (27) can be efficiently and deterministically checked (up to replacing $\varepsilon/6$ with $\varepsilon/3$, which is sufficient for invariant Item 3), it remains to note that using

Theorem 5.11 (with its "t" parameter set to $C \log n_0$), the algorithm can enumerate all lifts $\mathfrak{L}_{n_0} \in \Lambda$ in $\mathrm{poly}(n_0^t) = \mathrm{poly}(N)$ time, check all invariants for each, and we are assured that at least one (in fact, all but an $n_0^{-.99} + n_0^{-100t}$ fraction) will satisfy the invariants.

The subsequent lifts. As noted, we only need to show that for each $1 \leqslant i \leqslant t$, the algorithm can deterministically and efficiently find a 2-lift $\mathfrak{L}_{n_i}$ of $\mathfrak{L}_{n_{i-1}}$ such that invariant Item 3 holds, presuming that it holds for all smaller i. By Proposition 4.17, Item 2, this is equivalent to finding an edge-signing χ of $\mathfrak{L}_{n_{i-i}}$ such that

$$\rho(B_{n_{i-1}}(\chi\mathfrak{L}_{n_{i-1}}, \mathcal{K})) \leqslant \rho(B_\infty(\mathcal{K})) + \varepsilon \qquad \forall \mathcal{K} \in \mathfrak{K}_R. \tag{28}$$

To do this, the algorithm employs Corollary 6.15 with, say, $\bar{\varepsilon} = \varepsilon/6$ and $p = 1$. A key point is that $G_{\mathfrak{L}_{n_{i-1}}}$ is $C'\sqrt{\log N}$-bicycle free by invariant Item 3, and this is at least $C_2(\log\log n_i)^2$ even for the largest $n_t = N'$, as needed for Corollary 6.15 (presuming C' is taken large enough). Similarly, even for the largest n_t, Corollary 6.15 only requires $(1/\mathrm{poly}(N), O(\log N))$-wise uniform signings of $O(N)$ bits, and by Theorem 5.3 the algorithm can enumerate over a $\mathrm{poly}(N)$-size set of such signings. Corollary 6.15 tells us that for all but an $n_i^{-p} \leqslant n_0^{-1} = 2^{-\Theta(\sqrt{\log N})}$ fraction of such signings, the resulting $\chi\mathfrak{L}_{n_{i-1}}$ has

$$\|B_{n_{i-1}}(\chi\mathfrak{L}_{n_{i-1}}, \mathcal{K}))\|_{\mathrm{op}} \leqslant \rho(B_\infty(\mathcal{K})) + \varepsilon/6$$

for all $\mathcal{K} \in \Xi_1$. As with the 0th signing, the algorithm can deterministically and efficiently verify this (with $\varepsilon/3$ in place of $\varepsilon/6$) by enumerating over all $\mathcal{K} \in \Xi_1$ and calculating; and again, this in turn implies Inequality (28).

This concludes the proof, except for verifying the algorithm can find a lift that works simultaneously for all $r' \leqslant r$. But this only requires the algorithm to check all the nets in dimensions $1 \leqslant r' \leqslant r$ (only a constant-factor running time increase) and to note that in the failure probability analysis, union-bounding over all r' has a negligible effect. □

In the above proof, note that if we use one set of (δ, t)-wise uniform bits for all the edge-signings ($1 \leqslant i \leqslant t$), then by a union-bound the probability of any failure is $2^{-\Theta(\sqrt{\log N})}$. Furthermore, the total "seed-length" used is $O(\log N)$. Thus using the "strongly explicit" aspect of Theorems 5.6 and 5.11, and omitting all the deterministic "checks", we obtain the following "probabilistically strongly explicit" construction, just as in [MOP20a, App. B]:

Theorem 8.2. *In the setting of Theorem 8.1, there is also an algorithm that takes as input a number N and a seed $s \in \{0,1\}^{O(\log N)}$, and in deterministic $\mathrm{polylog}(N)$ time outputs a binary circuit $\mathfrak{C}$ that implements the adjacency list of a "color-regular lift graph" $G_{\mathfrak{L}_{N'}}$ (with $N \leqslant N' \leqslant N + o(N)$). Furthermore, with high probability over the choice of a uniformly random s (namely, except with probability $2^{-\Theta(\sqrt{\log N})}$), the resulting $G_{\mathfrak{L}_{N'}}$ satisfies the conclusions of Theorem 8.1.*

9 Relating the spectra of the adjacency operator and nonbacktracking operator

In this section we will relate the spectra of A_n and B_n. We require the following fact about spectra of possibly infinite-dimensional operators on Hilbert spaces (see e.g. [DS88a, VII.5]).

Fact 9.1. *Let $T : \mathscr{H} \to \mathscr{H}$ be a bounded linear operator on a Hilbert space $\mathscr{H}$. The spectrum of T can be decomposed into three disjoint sets. Suppose $\lambda \in \sigma(T)$. Then one of the following hold:*

1. *λ is in the* point spectrum *where $\lambda I - T$ is not injective. In this case, λ is an eigenvalue and there exists an eigenvector $v \in \mathscr{H}$ such that $\lambda v = Tv$.*
2. *λ is in the* continuous spectrum *where $\lambda I - T$ is not surjective, but the range of $\lambda I - T$ is dense in $\mathscr{H}$. In this case, for every $\varepsilon > 0$ there exists an approximate eigenvector $v_\varepsilon \in \mathscr{H}$ with $\|v_\varepsilon\| = 1$ such that $\|\lambda v_\varepsilon - Tv_\varepsilon\| \leqslant \varepsilon$, where $\|\cdot\|$ denotes the Hilbert space norm.*
3. *λ is in the* residual spectrum *where $\lambda I - T$ is not surjective and its range is not dense in $\mathscr{H}$. In this case, $\overline{\lambda}$ is an eigenvalue of the adjoint T^*, i.e. there exists $v \in \mathscr{H}$ such that $\overline{\lambda} v = T^* v$.*

The following is essentially [BC19, Prop. 9].

Proposition 9.2. *Let $\mathcal{K} = (a_0, \dots, a_{d+2e})$ be a matrix bouquet, and let $\mathfrak{L}_n$ be an n-lift (possibly with $n = \infty$). Let $\lambda \in \mathbb{C}$ such that $\lambda^2 \notin \sigma(a_{i*} a_i)$ for all $1 \leqslant i \leqslant d + 2e$. We define $\mathcal{K}_\lambda = (a_0(\lambda), \dots, a_{d+2e}(\lambda))$ where the coefficients are defined*

$$a_0(\lambda) = -1 - \sum_{i=1}^{d+2e} a_i(\lambda^2 - a_{i*} a_i)^{-1} a_{i*} \quad \text{and} \quad a_i(\lambda) = \lambda a_i (\lambda^2 - a_{i*} a_i)^{-1}.$$

Note that $\mathcal{K}_\lambda$ does not satisfy the symmetry conditions of matrix bouquets. We can nevertheless define the n-lift $A_n(\mathfrak{L}_n, \mathcal{K}_\lambda)$. Then, $\lambda \in \sigma(B_n(\mathfrak{L}_n, \mathcal{K}))$ if and only if $0 \in \sigma(A_n(\mathfrak{L}_n, \mathcal{K}_\lambda))$. Moreover, for $n < \infty$, $\lambda \in \sigma(B_{n,\perp}(\mathfrak{L}_n, \mathcal{K}))$ if and only if $0 \in \sigma(A_{n,\perp}(\mathfrak{L}_n, \mathcal{K}_\lambda))$.

Proof. In this proof, when A_n (an operator on $\ell_2(V_n) \otimes \mathbb{C}^r$) acts on a vector v, we think of v as being indexed by V_n, and each entry is an element in $\mathbb{C}^r$. We write v_x for $x \in V_n$ to indicate the block indexed by x. Similarly, when B_n acts on a vector v, we write $v_{x,i} \in \mathbb{C}^r$ for $x \in V_n$ and $i \in \mathscr{I}$ to index into v.

Let $\lambda \in \sigma(B_n(\mathfrak{L}_n, \mathcal{K}))$. We prove the statement assuming there exists some $i \in \mathscr{I}$ such that a_i is invertible, and use a density argument. Suppose no a_i is invertible. Then, let $\varepsilon > 0$ and let $i \in \mathscr{I}$ be an index where a_i is nonzero. Suppose $\lambda \in \sigma(B_n(\mathfrak{L}_n, \mathcal{K}_\lambda))$, and let $A_n = A_n(\mathfrak{L}_n, \mathcal{K}_\lambda)$ be the n-lift of $\mathcal{K}_\lambda$. Let $\mathcal{K}_\varepsilon$ be defined the same as $\mathcal{K}_\lambda$ but with $a_i + \varepsilon I_r$ in place of a_i. Let $A_\varepsilon = A_n(\mathfrak{L}_n, \mathcal{K}_\varepsilon) - A_n(\mathfrak{L}_n, \mathcal{K}_\lambda)$. Let $\delta > 0$. In our proof below, in all cases we will construct a vector $u_\delta \in V_n \otimes \mathbb{C}^r$ such that $\|u_\delta\| \leqslant C$ for a universal constant C and $\|A_\varepsilon u_\delta\| \leqslant \delta$. Since $\|A_\varepsilon u_\delta\| \leqslant \|A_n u_\delta\| + O(\varepsilon)$, taking $\varepsilon \to 0$ shows 0 is in the spectrum of A_n.

We split the proof into cases based on the classification in Fact 9.1.

Case 1: First suppose λ is in the point spectrum of B_n. Then there exists v such that $\lambda v = B_n v$. Our strategy is to construct a $u \in V_n \otimes \mathbb{C}^r$ such that u is in the kernel of $A_n(\mathfrak{L}_n, \mathcal{K}_\lambda)$, which shows 0 is an eigenvalue of $A_n(\mathfrak{L}_n, \mathcal{K}_\lambda)$.

Fix $x \in V_n$ and $i \in \mathscr{I}$ such that a_i is invertible. By definition of B_n,

$$\lambda v_{x,i} = \langle i | \langle x | B_n v = \langle i | \langle x | \sum_{y \in V_n} \sum_{j,k=1}^{d+2e} \mathbf{1}[k \neq j^*] \, |\sigma_j(y)\rangle \, |k\rangle \otimes a_k v_{y,j} = \sum_{j \neq i^*} a_i v_{\sigma_{i*}(x),j}. \tag{29}$$

Our candidate for u, defined coordinate-wise, is

$$u_x = \sum_{j=1}^{d+2e} v_{x,j}. \tag{30}$$

We can then write

$$\lambda v_{x,i} = a_i u_{\sigma_{i*}(x)} - a_i v_{\sigma_{i*}(x),i*}.$$

Plugging in $(\sigma_{i*}(x), i^*)$ in place of (x, i) in the above, we get

$$\lambda v_{\sigma_{i*}(x),i*} = a_{i*} u_x - a_{i*} v_{x,i*}.$$

Multiplying Equation (29) by λ, we get

$$\lambda^2 v_{x,i} = \lambda a_i u_{\sigma_{i*}(x)} - \lambda a_i v_{\sigma_{i*}(x),i*} = \lambda a_i u_{\sigma_{i*}(x)} - a_i a_{i*} u_x + a_i a_{i*} v_{x,i}.$$

By assumption, $(\lambda^2 - a_i a_{i*})$ is invertible. Therefore we can rearrange the above to

$$v_{x,i} = \lambda(\lambda^2 - a_i a_{i*})^{-1} a_i u_{\sigma_{i*}(x)} - (\lambda^2 - a_i a_{i*})^{-1} a_i a_{i*} u_x. \tag{31}$$

Now let $y = \sigma_{i*}(x)$. Substituting Equation (31) into Equation (29), we get

$$\begin{aligned}
\lambda v_{x,i} &= \lambda^2(\lambda^2 - a_i a_{i*})^{-1} a_i u_{\sigma_{i*}(x)} - \lambda(\lambda^2 - a_i a_{i*})^{-1} a_i a_{i*} u_x \\
&= \sum_{j \neq i^*} a_i v_{y,j} \\
&= \sum_{j \neq i^*} a_i \lambda(\lambda^2 - a_j a_{j*})^{-1} a_j u_{\sigma_{j*}(y)} - \sum_{j \neq i^*} a_i(\lambda^2 - a_j a_{j*})^{-1} a_j a_{j*} u_y.
\end{aligned} \tag{32}$$

We now note the following identities:

$$\begin{aligned}
(\lambda^2 - a_i a_{i*})^{-1} a_i &= (\lambda^2 - a_i a_{i*})^{-1} a_i (\lambda^2 - a_{i*} a_i)(\lambda^2 - a_{i*} a_i)^{-1} \\
&= (\lambda^2 - a_i a_{i*})^{-1} (\lambda^2 - a_i a_{i*}) a_i (\lambda^2 - a_{i*} a_i)^{-1} \\
&= a_i (\lambda^2 - a_{i*} a_i)^{-1},
\end{aligned} \tag{33}$$

and using the above,

$$\begin{aligned}
\lambda^2(\lambda^2 - a_i a_{i*})^{-1} - a_i(\lambda^2 - a_{i*} a_i)^{-1} a_{i*} &= \lambda^2(\lambda^2 - a_i a_{i*})^{-1} - a_i a_{i*}(\lambda^2 - a_i a_{i*})^{-1} \\
&= (\lambda^2 - a_i a_{i*})(\lambda^2 - a_i a_{i*})^{-1} = 1.
\end{aligned} \tag{34}$$

In the two terms on the right hand side of Equation (32), if we set $j = i^*$ and use Equation (33) we get

$$a_i \lambda(\lambda^2 a_{i*} a_i)^{-1} a_i^* u_{\sigma_i(y)} = \lambda(\lambda^2 - a_i a_{i*})^{-1} a_i a_{i*} u_x$$

and by using Equation (34) we get

$$a_i(\lambda^2 - a_{i*} a_i)^{-1} a_{i*} a_i u_y = a_i a_{i*}(\lambda^2 - a_i a_{i*})^{-1} a_i u_y = \lambda^2(\lambda^2 - a_i a_{i*})^{-1} a_i u_y - a_i u_y.$$

Therefore, using Equations (33) and (34), we can rearrange Equation (32) to get

$$\lambda^2(\lambda^2 - a_i a_{i*})^{-1} a_i u_{\sigma_{i*}(x)} = \sum_j a_i \lambda(\lambda^2 - a_j a_{j*})^{-1} a_j u_{\sigma_{j*}(y)} - \sum_{j \neq i^*} a_i(\lambda^2 - a_j a_{j*})^{-1} a_j a_{j*} u_y,$$

and then

$$a_i \left(1 + \sum_j a_j(\lambda^2 - a_{j*} a_j)^{-1} a_{j*}\right) u_y = a_i \sum_j \lambda a_j(\lambda^2 - a_{j*} a_j)^{-1} u_{\sigma_{j*}(y)}.$$

Since we assumed a_i is invertible, this implies

$$\left(1 + \sum_j a_j(\lambda^2 - a_{j*}a_j)^{-1}a_{j*}\right)u_y = \sum_j \lambda a_j(\lambda^2 - a_{j*}a_j)^{-1}u_{\sigma_{j*}(y)}.$$

This holds for all $x \in V_n$. Observing that

$$\langle x|\, A_n(\mathfrak{L}_n, \mathcal{K}_\lambda)u = \sum_{i=0}^{d+2e} a_i(\lambda)u_{\sigma_{i*}(x)},$$

we conclude that 0 is in the kernel of $A_n(\mathfrak{L}_n, \mathcal{K}_\lambda)$.

Case 2: If λ is in the continuous spectrum of B_n, then for every $\varepsilon > 0$ there exists a $v_\varepsilon \in V_n \otimes \mathscr{I} \otimes \mathbb{C}^r$ such that $\|v_\varepsilon\| = 1$ and $\|\lambda v_\varepsilon - B_n v_\varepsilon\| \leqslant \varepsilon$. The above construction applied to v_ε gives us a vector u_ε such that $\|A_n u_\varepsilon\| \leqslant \varepsilon$, and by scaling u appropriately we can conclude that 0 is in the continuous spectrum of $A_n(\mathfrak{L}_n, \mathcal{K}_\lambda)$.

Case 3: Suppose λ is in the residual spectrum of B_n. Then there exists v such that $\overline{\lambda}v = B_n^* v$. A similar proof as Case 1 holds here. Fix $x \in V_n$ and $i \in \mathscr{I}$. We find that

$$\overline{\lambda}v_{x,i} = \langle i|\langle x|\, B_n v = \langle i|\langle x| \sum_{y\in V_n}\sum_{j,k=1}^{d+2e} \mathbf{1}[k \neq j^*]\, |\sigma_{j*}(y)\rangle\, |j\rangle \otimes a_{k*}v_{y,k} = \sum_{j\neq i^*} a_{j*}v_{\sigma_i(x),j}. \tag{35}$$

Our candidate for u here is

$$u_x = \sum_{j=1}^{d+2e} a_{j*}v_{x,j}. \tag{36}$$

Similar to Case 1, we get

$$v_{x,i} = \overline{\lambda}(\lambda^2 - a_i a_{i*})^{-1}u_{\sigma_i(x)} - (\lambda^2 - a_i a_{i*})^{-1}a_i u_x. \tag{37}$$

Now let $y = \sigma_i(x)$. Substituting Equation (37) into Equation (35), we get

$$\begin{aligned}\overline{\lambda}v_{x,i} &= \lambda^2(\lambda^2 - a_i a_{i*})^{-1}u_{\sigma_i(x)} - \overline{\lambda}(\lambda^2 - a_i a_{i*})^{-1}a_i u_x = \sum_{j\neq i^*} a_{j*}v_{y,j} \\ &= \sum_{j\neq i^*} a_{j*}\overline{\lambda}(\lambda^2 - a_j a_{j*})^{-1}u_{\sigma_j(y)} - \sum_{j\neq i^*} a_{j*}(\lambda^2 - a_j a_{j*})^{-1}a_j u_y. \end{aligned}\tag{38}$$

Therefore, using Equations (33) and (34), we can rearrange Equation (38) to get

$$1 + \sum_j a_{j*}(\lambda^2 - a_j a_{j*})^{-1}a_j u_y = \sum_j \overline{\lambda}a_{j*}(\lambda^2 - a_j a_{j*})^{-1}u_{\sigma_j(y)}.$$

This holds for all $x \in V_n$. Observing that

$$\langle x|\, (A_n(\mathfrak{L}_n, \mathcal{K}_\lambda))^* u = \sum_{i=0}^{d+2e} a_i(\lambda)^* u_{\sigma_i(x)},$$

and using that $a_i^* = a_{i*}$ we conclude that 0 is in the kernel of $(A_n(\mathfrak{L}_n, \mathcal{K}_\lambda))^*$ (and hence in the kernel of $A_n(\mathfrak{L}_n, \mathcal{K}_\lambda)$ since $A_n(\mathfrak{L}_n, \mathcal{K}_\lambda)$ is a normal operator).

Projection onto $|+\rangle_n^\perp$ Finally, we can check that the computations in Case 1 still hold with v and u projected onto $|+\rangle_n^\perp$ (this is the only case to check for finite dimensional operators). Moreover, Equation (31) requires that $|+\rangle\langle+|\,u$ is nonzero when $|+\rangle\langle+|\,v$ is nonzero, hence $|+\rangle\langle+|\,u$ is in the kernel of $A_{n,\perp}(\mathfrak{L}_n, \mathcal{K}_\lambda)$. □

Next, we relate the spectral radius of the nonbacktracking operator to that of the adjacency operator. To begin, we recall some standard definitions.

Let $\mathcal{K}$ be a matrix bouquet and let $A_\infty = A_\infty(\mathfrak{L}_\infty, \mathcal{K})$ be the ∞-lift of $\mathcal{K}$. The *resolvent* of A_∞ is defined as

$$G(\mu) := (\mu I - A_\infty)^{-1}.$$

$G(\mu)$ is an operator on $\ell_2(V_\infty) \otimes \mathbb{C}^r$ defined for all $\mu \in \mathbb{C} \backslash \sigma(A_\infty)$. Since A_∞ is self-adjoint, we will henceforth restrict ourselves to μ being real. G is self-adjoint since A is self-adjoint. We can index into $G(\mu)$ in the following sense: Given $x, y \in V_\infty$, define

$$G_{xy}(\mu) := \langle x|\, G(\mu)\, |y\rangle \in \mathbb{C}^{r\times r}.$$

For $\mu \notin [-\rho(A_\infty), \rho(A_\infty)]$, we define an auxiliary matrix bouquet $\widehat{\mathcal{K}}(\mu)$ based on the resolvent. Then, studying the nonbacktracking operator of $\widehat{\mathcal{K}}(\mu)$ will tell us whether μ is in $\sigma(A_\infty)$. Let

$$\hat{a}_i(\mu) = G_{oo}(\mu)^{-1/2} G_{og_i}(\mu) G_{oo}(\mu)^{-1/2} \qquad \text{for} \quad 1 \leqslant i \leqslant d + 2e, \tag{39}$$

where $o \in V_\infty$ represents the origin, and g_i is a generator of $\mathbb{Z}$ or $\mathbb{Z}_2$. $\widehat{\mathcal{K}}(\mu)$ is the matrix bouquet with the same index set and involution as $\mathcal{K}$ but with $(\hat{a}_1(\mu), \hat{a}_2(\mu), \ldots, \hat{a}_{d+2e}(\mu))$ as the coefficients. Note that if $\mathcal{K}$ is self-adjoint, then so is $\widehat{\mathcal{K}}(\mu)$.

Before proving that $\hat{a}_i$'s are well-defined, we need a recursive characterization of G_{oo}. We define the operator $A_\infty^{(o)}$ to be A_∞ but with the edges around the root removed, i.e.

$$A_\infty^{(o)} = A_\infty - \sum_{i \in \mathscr{I}_0} a_i\, |g_i\rangle\langle o| + a_{i*}\, |o\rangle\langle g_i|\,.$$

Let $V_\infty^{(i)}$ be the subset of V_∞ where the first word in the free product of $\mathbb{Z}$'s and $\mathbb{Z}_2$'s is g_i, the i'th generator. $\mathbb{C}^r \otimes \ell^2(V_\infty)$ then decomposes into the direct sum $\mathbb{C}^r \otimes |o\rangle \oplus_i \mathbb{C}^r \otimes \ell^2(V_\infty^{(i)})$, and $A_\infty^{(o)}$ acts on each part of the direct sum independently. Let $A_\infty^{(i)}$ be $A_\infty^{(o)}$ restricted to $\mathbb{C}^r \otimes \ell^2(V_\infty^{(i)})$. We define $G^{(o)}$ to be the resolvent of $A_\infty^{(o)}$ and $G^{(i)}$ to be the resolvent of $A_\infty^{(i)}$. Note that $A_\infty^{(i)}$ is self-adjoint, and further $\rho(A_\infty^{(i)}) \leqslant \rho(A_\infty)$. This is because, for every $x \in \mathbb{C}^r \otimes \ell^2(V_\infty^{(i)})$ with $\|x\| = 1$, we can extend it to an x' in $\mathbb{C}^r \otimes \ell^2(V_\infty)$ by filling the other coordinates with 0, and $\|x'\| = 1$, so

$$x^T A_\infty^{(i)} x = x'^T A_\infty x' \leqslant \|A_\infty\|_{\mathrm{op}} = \rho(A_\infty).$$

Note that $G^{(o)}_{g_i g_i} = G^{(i)}_{oo}$. We will write $\gamma_i(\mu)$ to refer to either of these.

Where the expressions are well-defined, the resolvents satisfy the following recursion relations. This is almost identical to Lemma 11 from [BC19], but since our notation is slightly different, we reproduce the proof below.

Lemma 9.3. *Let $\mathcal{K} = (a_0, \ldots, a_{d+2e})$ be a matrix bouquet. Let $\mu \notin \sigma(A_\infty) \cup \sigma(A_\infty^{(o)})$, and let $G_{oo}(\mu)$ and $\gamma_i(\mu)$ be invertible for all $i \in \mathscr{I}_0$. Then, the following hold (with the dependence of on μ suppressed in the notation):*

1. $G_{og_i} = G_{oo}a_{i*}\gamma_i = \gamma_{i*}a_{i*}G_{oo}$, *hence* $\hat{a}_i = G_{oo}^{1/2}a_{i*}\gamma_i G_{oo}^{-1/2}$,
2. $G_{oo}^{-1} = \mu I_r - a_0 - \sum_{i\in\mathscr{I}} G_{oo}^{-1}G_{og_i}a_i$,
3. $G_{oo}^{1/2}a_{i*}G_{oo}^{1/2} = G_{oo}^{-1/2}G_{og_i}\left(I_r - G_{oo}^{-1}G_{og_{i*}}G_{oo}^{-1}G_{og_i}\right)^{-1}G_{oo}^{-1/2}$.

Proof. In this proof we also suppress the dependence on μ for the resolvents. Applying the resolvent identity on G and $G^{(o)}$, we have

$$G = G^{(o)} + G(A_\infty - A_\infty^{(o)})G^{(o)} = G^{(o)} + G^{(o)}(A_\infty - A_\infty^{(o)})G. \tag{40}$$

We note that for $x, y \in V_\infty$ in different connected components of $A_\infty^{(o)}$, we have $G_{xy}^{(o)} = 0$. Then, composing the first equality in Equation (40) with $\langle o|\cdot|g_i\rangle$ gives us $G_{og_i} = G_{oo}a_{i*}\gamma_i$, and composing the second expression with $\langle g_i|\cdot|o\rangle$ gives $G_{g_io} = \gamma_i a_i G_{oo}$ which implies Item 1 by symmetry.

Next, composing Equation (40) with $\langle o|\cdot|o\rangle$ gives

$$G_{oo} = G_{oo}^{(o)} + \sum_{j\in\mathscr{I}_0} G_{og_i}a_j G_{oo}^{(o)}.$$

We then use that $G_{oo}^{(}o) = (\mu I_r - a_0)^{-1}$ (since o is an isolated vertex in $A^{(o)}$), which is well defined since $\|a_0\| \leqslant \|A_\infty\|$. Rearranging the previous equation, we get Item 2

$$G_{oo}\left(\mu I_r - a_0 - \sum_{j\in\mathscr{I}_0} a_{j*}\gamma_j a_j\right) = I_r.$$

Finally, we fix $i \in \mathscr{I}_0$ and remove the edges around both o and $g_i \in V_\infty$ by defining

$$A_\infty^{(og_i)} = A_\infty^{(o)} - \sum_{j\neq i*} a_j \left|g_i g_j\right\rangle\!\left\langle g_i\right| - a_{j*}\left|g_i\right\rangle\!\left\langle g_i g_j\right|.$$

Using the resolvent identity similarly as before, and using that $G^{(og_i)}_{g_ig_j,g_ig_j} = \gamma_j$ by translation invariance, we obtain

$$\gamma_i^{-1} = \mu I_r - a_o - \sum_{j\neq i*} a_{j*}\gamma_j a_{j*}.$$

Using this, we find that

$$\gamma_i(I_r - a_i\gamma_{i*}a_{i*}\gamma_i)^{-1} = (\gamma_i^{-1} - a_i\gamma_{i*}a_{i*})^{-1} = \left(\mu I_r - a_0 - \sum_j a_{j*}\gamma_j a_j\right)^{-1} = G_{oo}.$$

Multiplying on the left by $G^{1/2}a_{i*}$, substituting in $a_{i*}\gamma_i = G_{oo}^{-1}G_{og_i}$ and then multiplying on the right by $G^{-1/2}$ gives Item 3. □

Next, we show that outside $[-\rho(A_\infty), \rho(A_\infty)]$, $G_{oo}(\mu)$ and $G_{oo}^{(i)}(\mu)$ are well-defined and bounded, hence $\hat{a}_i$ is well-defined.

Lemma 9.4. *Let $\varepsilon > 0$, and let $\mathcal{K} = (a_0, \ldots, a_{d+2e})$ be a matrix bouquet with coefficients satisfying $\|a_i\|_{\mathrm{op}} \leqslant 1$ for $i \in \mathscr{I}_0$, and let $\mu \in \mathbb{R}$ with $|\mu| > \rho(A_\infty) + \varepsilon$. Then, $\hat{a}_i(\mu)$ is well-defined and satisfies $\|\hat{a}_i(\mu)\|_{\mathrm{op}} \leqslant \varepsilon^{-3/2}(|\mu| + (d+2e+1))$. Moreover, if there is $\delta > 0$ such that $\|a_i^{-1}\| \leqslant 1/\delta$ for $i \in \mathscr{I}$, then we can also conclude that $\hat{a}_i(\mu)$ is non-singular and satisfies $\|\hat{a}_i(\mu)^{-1}\|_{\mathrm{op}} \leqslant \varepsilon\delta(|\mu| + (d+2e+1))^{-3/2}$.*

Proof. In this proof we will need several facts about spectral measures, which can be found for instance in [DS88b, Ch. X].

Given A_∞, there exists a unique positive operator-valued measure E (the *resolution to the identity*) supported on $\sigma(A_\infty)$ such that

$$\int_{\sigma(A_\infty)} dE = I, \quad \int_{\sigma(A_\infty)} \lambda \, dE(\lambda) = A_\infty \quad \text{and} \quad \int_{\sigma(A_\infty)} \frac{1}{\mu - \lambda} \, dE(\lambda) = G(\mu).$$

We define ν_o to be the positive operator-valued measure $\langle \delta_o, E(\cdot)\delta_o \rangle$. In particular,

$$\int_{\sigma(A_\infty)} \frac{1}{\mu - \lambda} \, d\nu_o(\lambda) = G_{oo}(\mu).$$

We will use the fact that since ν_o is positive-valued, for any function $f \geqslant 0$ pointwise,

$$\int_{\sigma(A_\infty)} f \, d\nu_o \geqslant 0 \quad \text{in Loewner order.}$$

One way to see this is, for any vector $x \in \mathbb{C}^r$, $\langle x, \nu_o(\cdot)x \rangle$ defines a nonnegative real-valued measure which is equal to the usual spectral measure with respect to the vector $x \otimes |o\rangle \in \mathbb{C}^r \otimes \ell^2(V_\infty)$, therefore

$$\left\langle x, \left(\int_{\sigma(A_\infty)} f \, d\nu_o \right) x \right\rangle = \int_{\sigma(A_\infty)} f \, d\langle x, \nu_o(\cdot)x \rangle \geqslant 0.$$

We can bound $\rho(A_\infty) = \|A_\infty\|_{\mathrm{op}} \leqslant \sum_{i \in \mathscr{I}_0} \|a_i\|_{\mathrm{op}} \leqslant d + 2e + 1$. For a real $\mu > \rho(A_\infty)$, for $\lambda \in \sigma(A_\infty)$ we can lower bound

$$\frac{1}{\mu + (d+2e+1)} \leqslant \frac{1}{\mu - \lambda}$$

and therefore

$$G_{oo}(\mu) = \int_{\sigma(A_\infty)} \frac{1}{\mu - \lambda} \, d\nu_o \geqslant \int_{\sigma(A_\infty)} \frac{1}{\mu + (d+2e+1)} \, d\nu_o = (\mu + (d+2e+1))^{-1} I.$$

Similarly, for $\mu < -\rho(A_\infty)$,

$$|G_{oo}(\mu)| = \int_{\sigma(A_\infty)} \frac{1}{\lambda - \mu} \, d\nu_o \geqslant \int_{\sigma(A_\infty)} \frac{1}{|\mu| + (d+2e+1)} \, d\nu_o = (|\mu| + (d+2e+1))^{-1} I.$$

We can upper bound

$$\|G_{oo}(\mu)\|_{\mathrm{op}} \leqslant \|G(\mu)\|_{\mathrm{op}} \leqslant \frac{1}{\operatorname{dist}(\mu, \sigma(A_\infty))} \leqslant \varepsilon^{-1}.$$

Therefore, we have shown that

$$(|\mu| + (d+2e+1))^{-1} I_r \leqslant |G_{oo}(\mu)| \leqslant \varepsilon^{-1} I_r$$

which in particular shows that $G_{oo}(\mu)^{-1/2}$ is well-defined. Since $\rho(A_\infty^{(i)}) \leqslant \rho(A_\infty)$, we get an identical bound for each $\gamma_i(\mu)$. Now, combining with $\|a_i\|_{\mathrm{op}} \leqslant 1$ and $\|a_i^{-1}\| \leqslant 1/\delta$ and using Lemma 9.3, Item 1 we get

$$\|\hat{a}_i(\mu)\|_{\mathrm{op}} \leqslant \varepsilon^{-3/2}(|\mu| + (d+2e+1)) \quad \text{and} \quad \|\hat{a}_i(\mu)^{-1}\|_{\mathrm{op}} \leqslant \varepsilon\delta(|\mu| + (d+2e+1))^{-3/2}. \qquad \square$$

Proposition 9.5. *Let $\mathcal{K}$ be a matrix bouquet with coefficients satisfying $\|a_i\|_{\mathrm{op}} < 1$ for all $i \in \mathscr{I}_0$, and let $A = A_\infty(\mathfrak{L}_\infty, \mathcal{K})$. Let $\mu \in \mathbb{R}$ with $|\mu| > \rho(A)$. Let $\widehat{\mathcal{K}}(\mu)$ be defined as in Equation (39). Let $\widehat{B}_\mu = B_\infty(\mathfrak{L}_\infty, \widehat{\mathcal{K}}(\mu))$ be the corresponding nonbacktracking operator of the ∞-lift of $\widehat{\mathcal{K}}(\mu)$. Then, $1 \notin \sigma(\widehat{B}_\mu)$.*

Moreover, the above also holds for any n-lift $\mathfrak{L}_n$ with $A = A_{n,\perp}(\mathfrak{L}_n, \mathcal{K})$ and $\widehat{B}_\mu = B_{n,\perp}(\mathfrak{L}_n, \widehat{\mathcal{K}}(\mu))$.

Proof. We will apply Proposition 9.2 with $\lambda = 1$. Let $\mathcal{K}' = (a'_0, \dots, a'_{d+2e})$ be the matrix bouquet obtained from applying Proposition 9.2 to $\widehat{\mathcal{K}}(\mu)$, and let $A' = A_\infty(\mathfrak{L}_\infty, \mathcal{K}')$ be the corresponding lift. We know that $0 \in A'$ if and only if $1 \in \sigma(\widehat{B}_\mu)$, so showing that $A' = (G_{oo}(\mu)^{1/2} \otimes I)(A - \mu I)(G_{oo}(\mu)^{1/2})$ shows that if $|\mu| > \rho(\widehat{A}_\mu)$ then $1 \notin \sigma(\widehat{B}_\mu)$. Note that $G_{oo}(\mu)^{1/2}$ is well-defined by Lemma 9.4.

Let $i \in \mathscr{I}$. We check that (omitting the dependence on μ),

$$\begin{aligned} a'_i &= \hat{a}_i(1 - \hat{a}_{i*}\hat{a}_i)^{-1} = G_{oo}^{-1/2} G_{og_i} G_{oo}^{-1/2} \left(1 - G_{oo}^{-1/2} G_{og_{i*}} G_{oo}^{-1} G_{og_i} G_{oo}^{-1/2}\right)^{-1} \\ &= G_{oo}^{-1/2} G_{og_i} \left(G_{oo}^{1/2} - G_{oo}^{-1/2} G_{og_{i*}} G_{oo}^{-1} G_{og_i}\right)^{-1} = G_{oo}^{-1/2} G_{og_i} \left(1 - G_{oo}^{-1} G_{og_{i*}} G_{oo}^{-1} G_{og_i}\right)^{-1} G_{oo}^{-1/2}, \end{aligned}$$

which is equal to $G_{oo}(\mu)^{1/2} a_{i*} G_{oo}(\mu)^{1/2}$ by Lemma 9.3, Item 3.

For a'_0, we check that

$$\begin{aligned} a'_0 &= -1 - \sum_{i \in \mathscr{I}} \hat{a}_i (1 - \hat{a}_{i*}\hat{a}_i)^{-1} \hat{a}_{i*} \\ &= -1 - \sum_{i \in \mathscr{I}} G_{oo}^{-1/2} G_{og_i} G_{oo}^{-1/2} \left(1 - G_{oo}^{-1/2} G_{og_{i*}} G_{oo}^{-1} G_{og_i} G_{oo}^{-1/2}\right)^{-1} G_{oo}^{-1/2} G_{og_{i*}} G_{oo}^{-1/2} \\ &= -1 - \sum_{i \in \mathscr{I}} G_{oo}^{-1/2} G_{og_i} \left(G_{oo}^{1/2} - G_{oo}^{-1/2} G_{og_{i*}} G_{oo}^{-1} G_{og_i}\right)^{-1} G_{oo}^{-1/2} G_{og_{i*}} G_{oo}^{-1/2} \\ &= -1 - \sum_{i \in \mathscr{I}} G_{oo}^{-1/2} G_{og_i} \left(1 - G_{oo}^{-1} G_{og_{i*}} G_{oo}^{-1} G_{og_i}\right)^{-1} G_{oo}^{-1} G_{og_{i*}} G_{oo}^{-1/2}. \end{aligned}$$

Now applying Lemma 9.3, Item 3, we get

$$a'_0 = -1 - \sum_{i \in \mathscr{I}} G_{oo}^{1/2} a_{i*} G_{og_{i*}} G_{oo}^{-1/2}.$$

Now using $G_{og_{i*}} = \gamma_i a_i G_{oo}$ from Lemma 9.3, Item 1, we have

$$\begin{aligned} a'_0 &= -1 - \sum_{i \in \mathscr{I}} G_{oo}^{1/2} a_{i*} \gamma_i a_i G_{oo}^{1/2} = -1 - \sum_{i \in \mathscr{I}} G_{oo}^{-1/2} G_{og_i} a_i G_{oo}^{1/2} \\ &= G_{oo}^{1/2} \left(-G_{oo}^{-1} - \sum_{i \in \mathscr{I}} G_{oo}^{-1} G_{og_i} a_i\right) G_{oo}^{1/2}, \end{aligned}$$

and by applying Lemma 9.3, Item 2, we get

$$a'_0 = G_{oo}^{1/2} (a_0 - \mu I_r) G_{oo}^{1/2}.$$

Now, notice that if we swap each a_i with a_{i*}, all lifts of the resulting matrix bouquet remain the same. Therefore we conclude that $A' = G_{oo}^{1/2}(A - \mu I) G_{oo}^{1/2}$.

Repeating the argument with $A = A_{n,\perp}(\mathfrak{L}_n, \mathcal{K})$, $\widehat{A}_\mu = A_{n,\perp}(\mathfrak{L}_n, \widehat{\mathcal{K}}(\mu))$ and $\widehat{B}_\mu = B_{n,\perp}(\mathfrak{L}_n, \widehat{\mathcal{K}}(\mu))$ gives the last part of the proposition. □

Finally, we can prove the following, which is essentially Theorem 12 of [BC19].

Theorem 9.6. *Let $\mathcal{K} = (a_0, \dots, a_{d+2e})$ be a matrix bouquet with $\max_{i \in \mathscr{I}_0} \|a_i\|_{\mathrm{op}} < 1$ and let $\mathfrak{L}_n$ be an n-lift. For any $\varepsilon > 0$, there exists $\delta > 0$ depending only on e, d, ε and not on the particular a_i's such that if for all $\mu \in \mathbb{R}$ with $|\mu| > \|A_\infty(\mathfrak{L}_\infty, \mathcal{K})\|_{\mathrm{op}} + \varepsilon$ it holds that*

$$\rho(B_{n,\perp}(\mathfrak{L}_n, \widehat{\mathcal{K}}(\mu))) \leqslant \rho(B_\infty(\mathfrak{L}_\infty, \widehat{\mathcal{K}}(\mu))) + \delta, \tag{41}$$

where $\widehat{\mathcal{K}}(\mu)$ is defined via Equation (39), then $\|A_{n,\perp}(\mathfrak{L}_n, \mathcal{K})\|_{\mathrm{op}} \leqslant \|A_\infty(\mathfrak{L}_\infty, \mathcal{K})\|_{\mathrm{op}} + \varepsilon$.

Proof. We write $A_\infty = A_\infty(\mathfrak{L}_\infty, \mathcal{K})$, $A_{n,\perp} = A_{n,\perp}(\mathfrak{L}_n, \mathcal{K})$, $\widehat{B}_\mu = B_\infty(\mathfrak{L}_\infty, \widehat{\mathcal{K}}(\mu))$, and $\widehat{B}_{n,\mu} = B_{n,\perp}(\mathfrak{L}_n, \widehat{\mathcal{K}}(\mu))$.

Fix $\varepsilon > 0$. First we show that we have $\rho(\widehat{B}_\mu) < 1$ for all $\mu \in \mathbb{R}$ with $|\mu| > \|A_\infty\|_{\mathrm{op}} + \varepsilon$. To do so, suppose for the sake of contradiction that $\rho(\widehat{B}_\mu) \geqslant 1$ for some $|\mu| > \|A_\infty\|_{\mathrm{op}} + \varepsilon$. As we take $\mu \to \infty$ if $\mu > 0$ or $\mu \to -\infty$ if $\mu < 0$, Lemma 9.4 implies that $\|\hat{a}_i(\mu)\|_{\mathrm{op}} \lesssim_{e,d,\varepsilon} |\mu|^{-1/2}$, therefore $\rho(\widehat{B}_\mu) \leqslant \|\widehat{B}_\mu\|_{\mathrm{op}} \lesssim_{e,d,\varepsilon} |\mu|^{-1/2}$ as well, in particular $\rho(\widehat{B}_\mu) \to 0$ as $|\mu| \to \infty$. Using the continuity of $\rho(\widehat{B}_\mu)$ as a function of μ (due to the $\hat{a}_i(\mu)$'s being continuous as a function of μ and Proposition 4.6), there must exist some μ' such that $\rho(\widehat{B}_{\mu'})_{\mathrm{op}} = 1$. Lemma 13 in [BC19] says that in this case, $1 \in \sigma(\widehat{B}_{\mu'})$. Since we assumed $|\mu| > \|A_\infty\|_{\mathrm{op}} + \varepsilon$, we know $|\mu'| > \|A_\infty\|_{\mathrm{op}} + \varepsilon$ as well, but this contradicts Proposition 9.5, so we conclude that $\rho(\widehat{B}_\mu) < 1$.

Again using that $\rho(\widehat{B}_\mu) \lesssim |\mu|^{-1/2}$ as $|\mu| \to \infty$, there exists some $D > \|A_\infty\|_{\mathrm{op}} + \varepsilon$ depending on e, d, ε such that $\rho(\widehat{B}_\mu) < 1/2$ for $|\mu| > D$. On $[-D, -\|A_\infty\|_{\mathrm{op}} - \varepsilon] \cup [\|A_\infty\|_{\mathrm{op}} + \varepsilon, D]$, $\rho(\widehat{B}_\mu)$ is uniformly continuous as a function of μ, hence there exists $\delta > 0$ such that $\rho(\widehat{B}_\mu) < 1 - \delta$ for all $|\mu| > \|A_\infty\|_{\mathrm{op}} + \varepsilon$. Hence, if we have Inequality (41) for all $|\mu| > \|A_\infty\|_{\mathrm{op}} + \varepsilon$ then $\rho(\widehat{B}_{n,\mu}) < 1$ for those μ, which by Proposition 9.5 implies $|\mu| > \rho(A_{n,\perp})$. □

10 Explicit lifts of matrix polynomials

In this section we combine Section 9 with Theorem 8.2 to get closeness of the spectra of $A_n(\mathfrak{L}_n, p)$ and $A_\infty(p)$ for a given p and $\mathfrak{L}_n$ from Theorem 8.2. The proof follows two main steps: first we will use Section 9 to bound the norm of $A_n(\mathfrak{L}_n, \mathcal{K})$ for all linear polynomials $\mathcal{K}$, and then using some operator theoretic tools in the literature, we can upgrade this bound for linear polynomials into Hausdorff-closeness in spectra for all polynomials.

10.1 Constructing lifts for linear polynomials

In this section our goal is to prove, roughly, that given $\varepsilon > 0$, for a linear polynomial $\mathcal{K}$ with bounded coefficients we can construct an $N' \sim N$ lift in $\mathrm{poly}(N)$ time such that $\|A_\infty(\mathcal{K})\|_{\mathrm{op}} - \varepsilon \leqslant \|A_{N',\perp}(\mathfrak{L}_{N'}, \mathcal{K})\|_{\mathrm{op}} \leqslant \|A_\infty(\mathcal{K})\|_{\mathrm{op}} + \varepsilon$.

To get the lower bound, we will require the following deterministic condition by Bordenave–Collins that ensures $\sigma(A_\infty)$ is in an ε-neighborhood of $\sigma(A_{n,\perp})$.

Proposition 10.1 ([BC19, Prop. 7]). *Let $\varepsilon, R > 0$. There exists $h \in \mathbb{N}^+$ such that for every $n \in \mathbb{N}^+$, if an n-lift $\mathfrak{L}_n$ has at least one vertex whose h-neighborhood is acyclic, then every matrix bouquet $\mathcal{K}$ with $\max_i \|a_i\|_{\mathrm{F}} \leqslant R$ satisfies*

$$\sigma(A_\infty(\mathcal{K})) \subseteq \sigma(A_{n,\perp}(\mathfrak{L}_n, \mathcal{K})) + [-\varepsilon, \varepsilon].$$

Remark 10.2. The proposition within Bordenave–Collins is stated for a particular matrix bouquet $\mathcal{K}$, but it suffices to prove the proposition for an ε-net of matrix bouquets, and use the stability of the spectra (Theorem 10.3).

The following theorem about the stability of the spectra of self-adjoint bounded linear operators under symmetric perturbations can be found in Kato, Ch. 5, Thm. 4.10 [Kat95]:

Theorem 10.3. *Let T be a self-adjoint bounded linear operator on a Hilbert space $\mathscr{H}$ and let $A : \mathscr{H} \to \mathscr{H}$ be a symmetric bounded operator. Let $S = T + A$. Then S is self-adjoint and* $\operatorname{dist}_H(\sigma(S), \sigma(T)) \leqslant \|A\|_{\mathrm{op}}$.

We also need a version of Proposition 10.1 for random signed lifts.

Proposition 10.4. *Let $\varepsilon, R > 0$. There exists $h \in \mathbb{N}^+$ such that for every $n \in \mathbb{N}^+$, if an n-lift $\mathfrak{L}_n$ has at least one vertex whose h-neighborhood is acyclic, and $\mathcal{K}$ is a matrix bouquet with $\max_i \|a_i\|_{\mathrm{F}} \leqslant R$, then any signing χ satisfies*

$$\sigma(A_\infty(\mathcal{K})) \subseteq \sigma(A_n(\boldsymbol{\chi}\mathfrak{L}_n, \mathcal{K})) + [-\varepsilon, \varepsilon].$$

Remark 10.5. The proof of this is essentially the same as Proposition 10.1, by noting that for $x \in V_n$ containing no cycle in its h-neighborhood, for $0 \leqslant k \leqslant 2h$,

$$\int_{\mathbb{R}} \lambda^k \, d\mu^x_{A_n} = \frac{1}{r} \sum_{\gamma} \operatorname{tr} \prod_{t=1}^{k} \chi_{i_t} a_{i_t} = \frac{1}{r} \sum_{\gamma} \operatorname{tr} \prod_{t=1}^{k} a_{i_t} = \int_{\mathbb{R}} \lambda^k \, d\mu^o_{A_\infty},$$

where γ is a closed length-k walk starting and ending at x. Or in words, the sum of the product of the weights on closed length-k walks at x is equal to that of the infinite graph A_∞.

Theorem 10.6. *Fix an index set $\mathscr{I}_0 = \{0, 1, \ldots, d + 2e\}$ and involution $*$ for a matrix bouquet. Fix also constants $r \in \mathbb{N}^+$ and $R, \varepsilon > 0$. Then, there is a* $\operatorname{poly}(N)$*-time deterministic algorithm that, on input N, outputs an N'-lift $\mathfrak{L}_{N'}$ (with $N \leqslant N' \leqslant N + o(N)$) such that for every matrix bouquet $\mathcal{K} = (a_0, \ldots, a_{d+2e})$ satisfying $\max_{i \in \mathscr{I}_0} \|a_i\|_{\mathrm{F}} \leqslant R$ with r'-dimensional matrix weights ($r' \leqslant r$), we have that $\|A_\infty(\mathcal{K})\|_{\mathrm{op}} - \varepsilon \leqslant \|A_{N',\perp}(\mathfrak{L}_{N'}, \mathcal{K})\|_{\mathrm{op}} \leqslant \|A_\infty(\mathcal{K})\|_{\mathrm{op}} + \varepsilon$.*

Moreover, for $h \in \mathbb{N}^+$ given by applying Proposition 10.1, $\mathcal{L}_{N'}$ also has at least one vertex whose h-neighborhood is acyclic.

Proof. By homogeneity, we first reduce to the case where $\max_{i \in \mathscr{I}_0} \|a_i\|_{\mathrm{F}} \leqslant \max_i \|a_i\|_{\mathrm{op}} < 1$ by scaling and replacing ε with $\varepsilon/2R$.

The lift $\mathfrak{L}_{N'}$ we construct will satisfy the hypothesis of Proposition 10.1, hence guaranteeing that $\|A_\infty(\mathcal{K})\|_{\mathrm{op}} \leqslant \|A_{N',\perp}(\mathfrak{L}_{N'}, \mathcal{K})\|_{\mathrm{op}} + \varepsilon$.

Next, we show that

$$\|A_{N',\perp}(\mathfrak{L}_{N'}, \mathcal{K})\|_{\mathrm{op}} \leqslant \|A_\infty(\mathcal{K})\|_{\mathrm{op}} + \varepsilon.$$

Given a matrix bouquet $\mathcal{K} = (a_0, \ldots, a_{d+2e})$ with $\max_i \|a_i\|_{\mathrm{op}} < 1$, we construct another matrix bouquet $\mathcal{K}' = (a'_0, \ldots, a'_{d+2e})$ where the coefficients have bounded inverse. For $n \in \mathbb{N}_+$, denote $A_{n,\perp} := A_{n,\perp}(\mathfrak{L}_n, \mathcal{K})$ and $A'_{n,\perp} := A'_{n,\perp}(\mathfrak{L}_n, \mathcal{K}')$. We will construct $\mathcal{K}'$ such that for any n-lift $\mathfrak{L}_n$, $\|A_n - A'_n\|_{\mathrm{op}} \leqslant \varepsilon/3$. Then, if we show that $\|A'_{n,\perp}\|_{\mathrm{op}} \leqslant \|A'_\infty\|_{\mathrm{op}} + \varepsilon/3$, it implies

$$\|A_{n,\perp}\|_{\mathrm{op}} \leqslant \|A'_{n,\perp}\|_{\mathrm{op}} + \|A_n - A'_n\|_{\mathrm{op}} \leqslant \|A'_\infty\|_{\mathrm{op}} + 2\varepsilon/3 \leqslant \|A_\infty\|_{\mathrm{op}} + \varepsilon.$$

Note it suffices to consider $a_0, \ldots, a_{d+2e}$ such that each a_i is self-adjoint (by replacing a_i with $(a_i + a_{i*})/2$, this results in the same lift $A_n(\mathfrak{L}_n, \mathcal{K})$). For each a_i, $i \in \mathscr{I}$, we can diagonalize

$a_i = U_i D_i U_i^*$. Let $R' = 3\varepsilon^{-1}(d+2e)$. Let D' be D but replace each diagonal entry in $[0, 1/R')$ with $1/R'$ and replace each diagonal entry in $(-1/R', 0)$ with $-1/R'$, and define $a_i' := U_i D_i' U_i^*$. Then, a_i' satisfies that $\|a_i'\|_{\text{op}} < 1$, $\|a_i'^{-1}\|_{\text{op}} \leqslant R'$ and $\|a_i' - a_i\|_{\text{op}} \leqslant 1/R'$. Furthermore, $\|A_n - A_n'\|_{\text{op}} \leqslant \sum_{i \in \mathscr{I}_0} \|a_i - a_i'\| \leqslant \varepsilon/3$.

Let $\mu \in \mathbb{R}$ with $|\mu| > \|A_\infty'\|_{\text{op}} + \varepsilon/3$. If $|\mu| > (d + 2e + 1)$, then certainly $|\mu| > \|A_n'\|_{\text{op}}$. Otherwise, we have $|\mu| \leqslant (d + 2e + 1)$ and we construct $\widehat{\mathcal{K}}(\mu)$ from $\mathcal{K}'$ as in Equation (39). By Lemma 9.4, we know for each $\hat{a}_i \in \mathscr{I}$,

$$\|\hat{a}_i\|_{\text{op}} \leqslant 2(\varepsilon/3)^{-3/2}(d + 2e + 1),$$

and

$$\|\hat{a}_i^{-1}\|_{\text{op}} \leqslant 2(\varepsilon/3)R'^{-1}(d + 2e + 1)^{-3/2}.$$

Let δ be from Theorem 9.6 with $\varepsilon/3$ in place of ε. Let $h \in \mathbb{N}^+$ be from applying Proposition 10.1, with $R = \max\{2(\varepsilon/3)^{-3/2}(d + 2e + 1), 2(\varepsilon/3)R'^{-1}(d + 2e + 1)^{-3/2}\}$, and ε. We then apply Theorem 8.1 with the same R and $\varepsilon = \delta$ and $C' = 1$, taking N' to be large enough that $\sqrt{\log N'} \geqslant h$, to obtain an N'-lift $\mathfrak{L}_{N'}$ that simultaneously for all $\mu \in \mathbb{R}$ with $|\mu| > \|A_\infty'\|_{\text{op}} + \varepsilon/3$, $\rho(B_{N',\perp}(\mathfrak{L}_{N'}, \widehat{\mathcal{K}}(\mu)) \leqslant \rho(B_\infty(\widehat{\mathcal{K}}(\mu)) + \delta$. By applying Theorem 9.6 we conclude that $\|A_{N',\perp}'\|_{\text{op}} \leqslant \|A_\infty'\|_{\text{op}} + \varepsilon/3$.

Finally, we note that our application of Theorem 8.1 depended only on R, ε, e, d, r and not the coefficients themselves, hence the lift $\mathfrak{L}_{N'}$ works for every matrix bouquet satisfying the given bounds. □

10.2 A linearization trick, and lifts of matrix polynomials

In order to derive norm convergence for polynomials from norm convergence for linear polynomials, we use the following *linearization* result, which appears in [Pis18] as Corollary 11 and the following remarks (and first appeared implicitly in [Pis96]). We will first state these results in terms of their limiting behavior, and derive quantitative bounds later. For complex-valued random variables $(X_n)_{n \in \mathbb{N}}$ and x, we say X_n converges in probability to x if for every $\varepsilon > 0$,

$$\lim_{n \to \infty} \mathbf{Pr}[|X_n - x| \geqslant \varepsilon] = 0.$$

Proposition 10.7. *Let I be a finite index set, and let $(x_j)_{j \in I}$ be free Haar unitaries in a unital C*-algebra A. Let $(X_j^{(N)})_{j \in I}$ be a system of random unitary matrices with a common dimension, for each N. If for every $r \in \mathbb{N}_+$ and any set of matrix coefficients $(a_j)_{j \in I}$ and $a_j \in \mathbb{C}^{r \times r}$ we have*

$$\left\| a_0 \otimes \mathbf{1} + \sum_{j \in I} a_j \otimes X_j^{(N)} + a_j^* \otimes (X_j^{(N)})^* \right\| \xrightarrow{N \to \infty} \left\| a_0 \otimes \mathbf{1} + \sum_{j \in I} a_j \otimes x_j + a_j^* \otimes (x_j)^* \right\| \quad \textit{in probability,}$$

Then, for all matrix-coefficient polynomials P in $|I|$ variables and their adjoints, we have

$$\left\| P((X_j^{(N)})_{j \in I}, (X_j^{(N)*})_{j \in I}) \right\| \xrightarrow{N \to \infty} \left\| P((x_j)_{j \in I}, (x_j^*)_{j \in I}) \right\| \quad \textit{in probability.}$$

Finally, the convergence in norm for all polynomials then implies convergence in Hausdorff distance. The following appears in Proposition 2.1 in [CM14][15].

[15] The proposition there is stated for convergence almost surely, but the same proof works for convergence in probability. Only the first inclusion in Item 2 is shown, but the other direction is also easy to show.

Proposition 10.8. *Let I be a finite index set, and let $(x_j)_{j\in I}$ and $(X_j^{(N)})_{j\in I}$ for $N \in \mathbb{N}_+$ be variables in C^*-probability spaces with faithful states. Then, the following are equivalent:*

1. *For every matrix-coefficient polynomial P in $|I|$ variables and their adjoints,*

$$\left\|P((X_j^{(N)})_{j\in I}, (X_j^{(N)*})_{j\in I})\right\| \xrightarrow{N\to\infty} \left\|P((x_j)_{j\in I}, (x_j^*)_{j\in I})\right\| \quad \text{in probability.}$$

2. *For every matrix-coefficient polynomial P in $|I|$ variables and their adjoints, let Y_N be the random variable $P((X_j^{(N)})_{j\in I}, (X_j^{(N)*})_{j\in I})$, and let Y be $P((x_j)_{j\in I}, (x_j^*)_{j\in I})$. Then, the spectrum of Y_N converges in probability to the spectrum of Y in Hausdorff distance, that is to say, for every $\varepsilon > 0$, with probability going to 1 as N goes to infinity,*

$$\sigma(Y_N) \subseteq \sigma(Y) + [-\varepsilon, \varepsilon]$$

and

$$\sigma(Y) \subseteq \sigma(Y_N) + [-\varepsilon, \varepsilon].$$

Using Theorem 7.2 followed by Theorem 6.5 instead of Theorem 8.1 in the proof of Theorem 10.6, and also using Proposition 10.4 instead of Proposition 10.1, and finally applying Propositions 10.7 and 10.8, we can deduce the following probabilistic statement.

Theorem 10.9. *Let p be a matrix polynomial. Let $\mathfrak{L}_n$ be a uniformly random n-lift, and let $\mathfrak{T}_2^{(1)}, \ldots, \mathfrak{T}_2^{(k)}$ be k independent uniformly random 2-lifts. Then, $\sigma(A_{2^k n,\perp}(\mathfrak{T}_2^{(1)} \otimes \cdots \otimes \mathfrak{T}_2^{(k)} \otimes \mathfrak{L}_n, p))$ converges in probability in Hausdorff distance to $\sigma(A_\infty(p))$ an $n \to \infty$.*

As a consequence, we get the following theorem, which says that the spectrum of a uniformly random signed lift converges to that of the ∞-lift, without the need to project onto the $|+\rangle_n^\perp$ space.

Theorem 10.10. *Let p be a matrix polynomial, and let $\boldsymbol{\chi}\mathfrak{L}_n$ be a uniformly random signed n-lift (i.e. by taking both $\boldsymbol{\chi}$ and $\mathfrak{L}_n$ uniformly random). Then, $\sigma(A_n(\boldsymbol{\chi}\mathfrak{L}_n, p))$ converges in probability in Hausdorff distance to $\sigma(A_\infty(p))$.*

Proof. We take $k = 1$ in Theorem 10.9. Proposition 4.16, Item 2 tells us that

$$\sigma(A_{2n,\perp}(\mathfrak{T}_2 \otimes \mathfrak{L}_n, p)) = \sigma(A_{n,\perp}(\mathfrak{L}_n, p)) \cup \sigma(A_{n,\boldsymbol{\chi}}(\mathfrak{L}_n, p)).$$

We know from Bordenave–Collins' theorem Theorem 1.1 that $\sigma(A_{n,\perp}(\mathfrak{L}_n, p))$ converges in probability in Hausdorff distance to $\sigma(A_\infty(p))$. The above equation tells us that $\sigma(A_{n,\boldsymbol{\chi}}(\mathfrak{L}_n, p)) \subseteq \sigma(A_{2n,\perp}(\mathfrak{T}_2 \otimes \mathfrak{L}_n, p))$, which converges to $\sigma(A_\infty(p))$ by Theorem 10.9. □

Next, we find the derandomized analogue of Theorem 10.9. To do so, we will need to examine the proof of Proposition 10.7, which uses the following factorization. We state this only specialized to our situation of permutations/matchings converging to generators of $\mathbb{Z}$ and $\mathbb{Z}_2$, instead of the more general form appearing in [Pis18, Cor. 4, Cor. 7].

Proposition 10.11. *Let $\mathfrak{L}_n$ be an n-lift (with n possibly being ∞). Let p be a matrix polynomial satisfying $\|A_n(\mathfrak{L}_n, p)\|_{\mathrm{op}} \leqslant 1$. Then, there exists $m \in \mathbb{N}_+$ such that p can be factorized into*

$$p = \alpha_0 D_1 \alpha_1 \ldots D_m \alpha_m,$$

where each α_i is a rectangular complex-valued matrix satisfying $\|\alpha_i\|_{\mathrm{op}} \leqslant 1$, and each D_i is a self-adjoint linear matrix-coefficient polynomial (possibly of different dimensions) satisfying $\|A_n(\mathfrak{L}_n, D_i)\|_{\mathrm{op}} \leqslant 1$.

Using this factorization, we can upgrade the convergence in norm for linear polynomials in Theorem 10.6 to convergence in the entire spectrum in Hausdorff distance, not just for linear polynomials but for all self-adjoint polynomials.

We will need to use the *spectral mapping theorem* (see e.g. [DS88a, Thm VII.3.11]).

Theorem 10.12. *Let $T : \mathscr{H} \to \mathscr{H}$ be a bounded linear operator on a Hilbert space $\mathscr{H}$. Let f be a holomorphic function defined on a neighborhood of $\sigma(T)$. Then, $f(\sigma(T)) = \sigma(f(T))$.*

We will only be using this in the case where f is a polynomial. In this case $f(A_n(\mathfrak{L}_n, p)) = A_n(\mathfrak{L}_n, f(p))$, where we interpret $f(p)$ by way of addition and multiplication in the polynomial ring.

Theorem 10.13. *Fix an index set $\mathscr{I}_0 = \{0, 1, \ldots, d+2e\}$ and involution $*$ for a matrix bouquet. Fix also constants $r, k \in \mathbb{N}^+$ and $R, \varepsilon > 0$. Then, there is a $\mathrm{poly}(N)$-time deterministic algorithm that, on input N, outputs an N'-lift $\mathfrak{L}_{N'}$ (with $N \leqslant N' \leqslant N + o(N)$) such that simultaneously for every self-adjoint matrix polynomial p with total degree at most k, coefficients in $\mathbb{C}^{r' \times r'}$ with $r' \leqslant r$ and the Frobenius norm of the coefficients of p are at most R, we have that $\sigma(A_{n,\perp}(\mathfrak{L}_{N'}, p))$ and $\sigma(A_\infty(p))$ are ε-close in Hausdorff distance.*

Proof. First, we scale p by $1/2R$ so that $\max_i \|a_i\|_F < 1$, and replace ε with $\varepsilon/2R$.

The lift $\mathfrak{L}_{N'}$ we construct will satisfy the hypothesis of Proposition 10.1, hence $\sigma(A_\infty(\mathcal{K})) \subseteq \sigma(A_{n,\perp}(\mathfrak{L}_n, \mathcal{K})) + [-\varepsilon, \varepsilon]$.

To show the other inclusion, we proceed by an ε-net argument over all polynomials with total degree at most k. Let $\delta = \varepsilon/(3(k+1))$. Let $\mathfrak{F}_{r'}$ be the set of matrices in $\mathbb{C}^{r' \times r'}$ with Frobenius norm less than 1, and let $\mathfrak{G}_{r',\delta}$ be a net of matrices in $\mathfrak{F}_{r'}$ (e.g. by taking the entries to be complex integer multiples of δ) with $|\mathfrak{G}_{r',\delta}| \leqslant O(r'/\delta)^{r'^2}$ such that for every $b \in \mathfrak{F}_{r'}$ there is $\underline{b} \in \mathfrak{G}_{r',\delta}$ such that $\|b - \underline{b}\|_F \leqslant \delta$.

Let $\mathfrak{P}_{r'}$ be the set of polynomials with total degree at most k with coefficients from $\mathfrak{G}_{r',\delta}$. Note that $|\mathfrak{P}_{r'}| \leqslant (|\mathfrak{G}_{r',\delta}| + 1)^k (d+2e)^k \leqslant O(r/\delta)^{kr^2} (d+2e)^k$. Let $p = \sum_{w \in \mathscr{T}} a_w X^w$ now be some matrix polynomial with coefficients in $\mathfrak{F}_{r'}$. There exists a polynomial $\underline{p}$ in $\mathfrak{P}_{r'}$ where $\underline{p} = \sum_{w \in \mathscr{T}} \underline{a_w} X^w$ where $\mathscr{T}$ is the same as in p, and for each $w \in \mathscr{T}$, $\|a_w - \underline{a_w}\|_F \leqslant \delta$. Since $|\mathscr{T}| \leqslant k+1$, for any n-lift $\mathfrak{L}_n$, $\|A_{n,\perp}(\mathfrak{L}_n, p) - A_{n,\perp}(\mathfrak{L}_n, \underline{p})\|_{\mathrm{op}} \leqslant \delta(k+1) \leqslant \varepsilon/3$, and also $\|A_\infty(p) - A_\infty(\underline{p})\|_{\mathrm{op}} \leqslant \varepsilon/3$. Using Theorem 10.3, we have $\mathrm{dist}_H(A_n(\mathfrak{L}_n, p), A_n(\mathfrak{L}_n, \underline{p})) \leqslant \varepsilon/3$. Therefore, it suffices to find an N'-lift $\mathfrak{L}_{N'}$ such that $\mathrm{dist}_H(A_{N'}(\mathfrak{L}_{N'}, \underline{p}), A_\infty(\underline{p})) \leqslant \varepsilon/3$ for every $\underline{p} \in \mathfrak{F}_{r'}$ for every $r' \leqslant r$.

Now fix a polynomial $p \in \overline{\mathfrak{P}}_r$. Let $\overline{T}$ be the interval $[-(k+1+\varepsilon/3), (k+1+\varepsilon/3)]$, and let $S \subset T$ be the spectrum of $A_\infty(p)$. Let $f : \mathbb{R} \to \mathbb{R}$ be a smooth function such that $f(x) = -1$ for $x \in S$ and $f(x) = 1$ for $x \in T \backslash (S + [-\varepsilon/3, \varepsilon/3])$. Let q' be a polynomial such that $|f - q'| < 0.1$ on T, and let q be q' scaled so that $\min q(x) = -2$ for $x \in T$.

Using the spectral mapping theorem Theorem 10.12, we constructed q such that $q(S) \subseteq [-2, 0]$ so we know $\|q(A_\infty(p)) - 1\|_{\mathrm{op}} = 1$. We apply Proposition 10.11 on $q(A_\infty(p)) - 1$ to get a factorization $\alpha_0 D_1 \alpha_1 \ldots D_m \alpha_m$ where each α_i is a rectangular scalar matrix, and each D_i is the ∞-lift of some self-adjoint linear polynomial (which we view as a matrix bouquet) β_i. The factorization further satisfies $\|\alpha_i\|_{\mathrm{op}} \leqslant 1$ and $\|D_i\|_{\mathrm{op}} \leqslant 1$.

We repeat this construction and factorization for every $\underline{p} \in \mathfrak{P}_{r'}$ and for every $r' \leqslant r$, and let $\overline{m}$ be the maximum number of factors in the factorization, let $\overline{r}$ be the maximum dimension of the matrix bouquets, and let $\overline{R}$ be the maximum Frobenius norm of the coefficients of the matrix bouquets obtained in any of the factorizations. Let $\overline{\varepsilon} > 0$ be small enough such that $(1+\overline{\varepsilon})^{\overline{m}} \leqslant 1.5$. We then apply Theorem 10.6 with the parameters $\overline{r}, \overline{R}$ and $\overline{\varepsilon}$ to construct an N'-lift $\mathfrak{L}_{N'}$ with

$N \leqslant N' \leqslant N + o(N)$ such that simultaneously for every matrix bouquet with $\max_{i \in \mathscr{I}_0} \leqslant \overline{R}$ and at most $\bar{r}$-dimensional weights, $\|A_{N',\perp}(\mathfrak{L}_{N'}, \mathcal{K})\|_{\mathrm{op}} \leqslant \|A_\infty(\mathcal{K})\|_{\mathrm{op}} + \bar{\varepsilon}$. Using this $\mathfrak{L}_{N'}$ on the factorization of $q(A_\infty(p)) + 1$ (by substituting in $\mathfrak{L}_{N'}$ for $\mathfrak{L}_\infty$, we have $\|q(A_{N'}(\mathfrak{L}_{N'}, p)) - 1\|_{\mathrm{op}} \leqslant 1.5$. Applying the spectral mapping theorem to $q(A_{N'}(\mathfrak{L}_{N'}, p))$, since $q > 1$ on $T \backslash (S + [-\varepsilon/3, \varepsilon/3])$ we conclude that $\sigma(A_{N'}(\mathfrak{L}_{N'}, p)) \subseteq \sigma(A_\infty(p)) + [-\varepsilon/3, \varepsilon/3]$. Furthermore, this holds simultaneously for every $p \in \mathfrak{P}_{r'}$ for any $r \leqslant r'$, hence we conclude as desired. □

Finally, by applying Theorem 8.2 instead of Theorem 8.1 in the proofs of Theorems 10.6 and 10.13, we can obtain

Theorem 10.14. *In the setting of Theorems 10.6 and 10.13, there is also an algorithm that takes as input a number N and a seed $s \in \{0,1\}^{O(\log N)}$, and in deterministic* $\mathrm{polylog}(N)$ *time outputs a binary circuit $\mathfrak{C}$ that implements the adjacency list of a "color-regular lift graph" $G_{\mathfrak{L}_{N'}}$ (with $N \leqslant N' \leqslant N + o(N)$). Furthermore, with high probability over the choice of a uniformly random s (namely, except with probability $2^{-\Theta(\sqrt{\log N})}$), the resulting $G_{\mathfrak{L}_{N'}}$ satisfies the conclusions of Theorems 10.6 and 10.13 respectively.*

Acknowledgments

The authors would like to thank Charles Bordenave and Benoît Collins for several helpful clarifications regarding [BC19]. The authors would also like to thank Sidhanth Mohanty for sharing the second construction in Example 3.6, and for pointing us to [VK19]. Additionally, the authors would like to thank Sidhanth Mohanty and Pedro Paredes for many helpful discussions. Finally, the authors thank Peter Sarnak for telling us of the results mentioned in Example 1.5.

References

[AC88] Noga Alon and Fan Chung. Explicit construction of linear sized tolerant networks. *Discrete Mathematics*, 72(1-3):15–19, 1988. 4

[AGHP92] Noga Alon, Oded Goldreich, Johan Håstad, and René Peralta. Simple constructions of almost k-wise independent random variables. *Random Structures & Algorithms*, 3(3):289–304, 1992. 40

[AGV16] Miklós Abért, Yair Glasner, and Bálint Virág. The measurable Kesten theorem. *The Annals of Probability*, 44(3):1601–1646, 2016. 9

[AL02] Alon Amit and Nathan Linial. Random graph coverings. I. General theory and graph connectivity. *Combinatorica*, 22(1):1–18, 2002. 22

[AL07] David Aldous and Russell Lyons. Processes on unimodular random networks. *Electron. J. Probab.*, 12(54):1454–1508, 2007. 33, 34

[AL13] Noga Alon and Shachar Lovett. Almost k-wise vs. k-wise independent permutations, and uniformity for general group actions. *Theory of Computing*, 9:559–577, 2013. 40, 41

[Alo86] Noga Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986. 4

[And13] Greg W. Anderson. Convergence of the largest singular value of a polynomial in independent Wigner matrices. *Ann. Probab.*, 41(3B):2103–2181, 2013. 12

[BC19] Charles Bordenave and Benoît Collins. Eigenvalues of random lifts and polynomials of random permutation matrices. *Annals of Mathematics*, 190(3):811–875, 2019. 4, 11, 12, 16, 20, 21, 37, 42, 44, 46, 52, 53, 54, 57, 60, 64, 69

[BF14] Mladen Bestvina and Mark Feighn. Hyperbolicity of the complex of free factors. *Adv. Math.*, 256:104–155, 2014. 30

[Bha97] Rajendra Bhatia. *Matrix analysis*. Springer–Verlag, 1997. 43

[BK90] Hyman Bass and Ravi Kulkarni. Uniform tree lattices. *J. Amer. Math. Soc.*, 3(4):843–902, 1990. 32

[BKM01] Gunnar Brinkmann, Jack H. Koolen, and V. Moulton. On the hyperbolicity of chordal graphs. *Ann. Comb.*, 5(1):61–69, 2001. 30

[BL06] Yonatan Bilu and Nathan Linial. Lifts, discrepancy and nearly optimal spectral gap. *Combinatorica*, 26(5):495–519, 2006. 11, 12

[BLPS99] Itai Benjamini, Russell Lyons, Yuval Peres, and Oded Schramm. Group-invariant percolation on graphs. *Geom. Funct. Anal.*, 9(1):29–66, 1999. 32, 36

[BMST88] N. L. Biggs, Bojan Mohar, and John Shawe-Taylor. The spectral radius of infinite graphs. *Bull. London Math. Soc.*, 20(2):116–120, 1988. 35

[Bor19] Charles Bordenave. A new proof of Friedman's second eigenvalue theorem and its extension to random lifts. *Annales scientifiques de l'École normale supérieure*, 2019. arXiv:1502.04482. 4, 11, 53

[BRS11] Sergio Bermudo, José M. Rodríguez, and José M. Sigarreta. Computing the hyperbolicity constant. *Comput. Math. Appl.*, 62(12):4592–4595, 2011. 30

[BRSV13] Sergio Bermudo, José M. Rodríguez, José M. Sigarreta, and Jean-Marie Vilaire. Gromov hyperbolic graphs. *Discrete Math.*, 313(15):1575–1585, 2013. 30

[BS01] Itai Benjamini and Oded Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6:no. 23, 13, 2001. 33

[Car17] Walter Carballosa. Gromov hyperbolicity and convex tessellation graph. *Acta Math. Hungar.*, 151(1):24–34, 2017. 30

[CFHM12] Wei Chen, Wenjie Fang, Guangda Hu, and Michael W. Mahoney. On the hyperbolicity of small-world and tree-like random graphs. In *Algorithms and computation*, volume 7676 of *Lecture Notes in Comput. Sci.*, pages 278–288. Springer, Heidelberg, 2012. 30

[CFL+18] Anamaria Costache, Brooke Feigon, Kristin Lauter, Maike Massierer, and Anna Puskás. Ramanujan graphs in cryptography. *arXiv preprint arXiv:1806.05709*, 2018. 4

[Chi92] Patrick Chiu. Cubic Ramanujan graphs. *Combinatorica*, 12(3):275–285, 1992. 4

[CHSH69] John Clauser, Michael Horne, Abner Shimony, and Richard Holt. Proposed experiment to test local hidden-variable theories. *Physical Review Letters*, 23(15):880–884, 1969. 13

[Cla07] Pete Clark. Ramanujan graphs part II. http://alpha.math.uga.edu/~pete/ramgraphs2.pdf, 2007. Notes from a lecture. 8

[CM14] Benoît Collins and Camille Male. The strong asymptotic freeness of Haar and deterministic matrices. *Ann. Sci. Éc. Norm. Supér. (4)*, 47(1):147–163, 2014. 66

[DL01] Reinhard Diestel and Imre Leader. A conjecture concerning a limit of non-Cayley graphs. *J. Algebraic Combin.*, 14(1):17–25, 2001. 33

[DMO+19] Yash Deshpande, Andrea Montanari, Ryan O'Donnell, Tselil Schramm, and Subhabrata Sen. The threshold for SDP-refutation of random regular NAE-3SAT. In *Proceedings of the 30th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2305–2321. SIAM, Philadelphia, PA, 2019. 13, 14, 15

[Dod84] Jozef Dodziuk. Difference equations, isoperimetric inequality and transience of certain random walks. *Trans. Amer. Math. Soc.*, 284(2):787–794, 1984. 36

[DS88a] Nelson Dunford and Jacob T. Schwartz. *Linear operators. Part I*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1988. General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication. 56, 68

[DS88b] Nelson Dunford and Jacob T. Schwartz. *Linear operators. Part II*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1988. Spectral theory. Selfadjoint operators in Hilbert space, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1963 original, A Wiley-Interscience Publication. 62

[Fie73] Miroslav Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23(98):298–305, 1973. 4, 14

[FMK17] Behrang Forghani and Keivan Mallahi-Karai. Amenability of trees. In *Groups, graphs and random walks*, volume 436 of *London Math. Soc. Lecture Note Ser.*, pages 176–189. Cambridge Univ. Press, Cambridge, 2017. 36

[Fri93] Joel Friedman. Some geometric aspects of graphs and their eigenfunctions. *Duke Math. J.*, 69(3):487–525, 1993. 16

[Fri03] Joel Friedman. Relative expanders or weakly relatively Ramanujan graphs. *Duke Mathematical Journal*, 118(1):19–35, 2003. 4

[Fri08] Joel Friedman. A proof of Alon's second eigenvalue conjecture and related problems. *Memoirs of the American Mathematical Society*, 195(910):viii+100, 2008. 4

[Ger88] Peter Gerl. Random walks on graphs with a strong isoperimetric property. *J. Theoret. Probab.*, 1(2):171–187, 1988. 35, 36

[GLS93] Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*, volume 2 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, second edition, 1993. 55

[Gro83] Mikhael Gromov. Filling Riemannian manifolds. *J. Differential Geom.*, 18(1):1–147, 1983. 30

[Gro87] Mikhael Gromov. Hyperbolic groups. In *Essays in group theory*, volume 8 of *Math. Sci. Res. Inst. Publ.*, pages 75–263. Springer, New York, 1987. 29

[GZ99] Rostislav Grigorchuk and Andrzej Żuk. On the asymptotic spectrum of random walks on infinite families of graphs. In *Random walks and discrete potential theory (Cortona, 1997)*, Sympos. Math., XXXIX, pages 188–204. Cambridge Univ. Press, Cambridge, 1999. 8

[HM13] Michael Handel and Lee Mosher. The free splitting complex of a free group, I: hyperbolicity. *Geom. Topol.*, 17(3):1581–1672, 2013. 30

[Hor16] Camille Horbez. Hyperbolic graphs for free products, and the Gromov boundary of the graph of cyclic splittings. *J. Topol.*, 9(2):401–450, 2016. 30

[HPR19] Verónica Hernández, Domingo Pestana, and José M. Rodríguez. Several extremal problems on graphs involving the circumference, girth, and hyperbolicity constant. *Discrete Appl. Math.*, 263:177–194, 2019. 30

[Iha66] Yasutaka Ihara. On discrete subgroups of the two by two projective linear group over $\mathfrak{p}$-adic fields. *Journal of the Mathematical Society of Japan*, 18:219–235, 1966. 4

[Kai92] Vadim A. Kaimanovich. Dirichlet norms, capacities and generalized isoperimetric inequalities for Markov operators. *Potential Anal.*, 1(1):61–82, 1992. 36

[Kas07] Martin Kassabov. Symmetric groups and expander graphs. *Inventiones Mathematicae*, 170(2):327–354, 2007. 40, 41

[Kat95] Tosio Kato. *Perturbation theory for linear operators*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition. 65

[Kes59] Harry Kesten. Full Banach mean values on countable groups. *Math. Scand.*, 7:146–156, 1959. 35

[Kle07] Robert Kleinberg. Geographic routing using hyperbolic space. In *Proceedings of the IEEE INFOCOM 2007 - 26th IEEE International Conference on Computer Communication*, pages 1902–1909, 2007. 30

[KM02] Ilya Kapovich and Alexei Myasnikov. Stallings foldings and subgroups of free groups. *J. Algebra*, 248(2):608–668, 2002. 29

[KNR09] Eyal Kaplan, Moni Naor, and Omer Reingold. Derandomized constructions of k-wise (almost) independent permutations. *Algorithmica. An International Journal in Computer Science*, 55(1):113–133, 2009. 40, 41

[LPS88] Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988. 4

[Mak15] Aleksandar Makelov. Expansion in lifts of graphs, 2015. Bachelor's thesis, Harvard College. 17

[Mar08] Andrey Markov. Rasprostranenie predel'nyh teorem ischisleniya veroyatnostej na summu velichin svyazannyh v cep'. *Zapiski Akademii nauk po Fiziko-matematicheskomu otdeleniiu, VIII seriya*, 25(3), 1908. 4

[Mar88] Grigory Margulis. Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators. *Problemy Peredachi Informatsii*, 24(1):51–60, 1988. 4

[MO19] Sidhanth Mohanty and Ryan O'Donnell. X-Ramanujan graphs. Technical Report 1904.03500, arXiv, 2019. 24

[MO20] Sidhanth Mohanty and Ryan O'Donnell. X-Ramanujan graphs. In *Proceedings of the 31st Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1226–1243, 2020. 5, 8, 9, 10, 22, 23

[Moh88] Bojan Mohar. Isoperimetric inequalities, growth, and the spectrum of graphs. *Linear Algebra Appl.*, 103:119–131, 1988. 36

[MOP20a] Sidhanth Mohanty, Ryan O'Donnell, and Pedro Paredes. Explicit near-Ramanujan graphs of every degree. In *Proceedings of the 52nd Annual ACM Symposium on Theory of Computing*, 2020. 11, 12, 13, 40, 44, 45, 46, 51, 54, 55, 56

[MOP20b] Sidhanth Mohanty, Ryan O'Donnell, and Pedro Paredes. The SDP value for random two-eigenvalue CSPs. In Christophe Paul and Markus Bläser, editors, *Proceedings of the 37th Annual Symposium on Theoretical Aspects of Computer Science*, volume 154 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 50:1–50:45, Dagstuhl, Germany, 2020. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. 13, 14, 15

[Mor94] Moshe Morgenstern. Existence and explicit constructions of $q+1$ regular Ramanujan graphs for every prime power q. *Journal of Combinatorial Theory. Series B*, 62(1):44–62, 1994. 4

[MS16] Andrea Montanari and Subhabrata Sen. Semidefinite programs on sparse random graphs and their application to community detection. In *Proceedings of the 48th Annual ACM Symposium on Theory of Computing*, pages 814–827. ACM, New York, 2016. 13, 14

[MSS15] Adam Marcus, Daniel Spielman, and Nikhil Srivastava. Interlacing families I: Bipartite Ramanujan graphs of all degrees. *Annals of Mathematics. Second Series*, 182(1):307–325, 2015. 4

[MSV11] Fabien de Montgolfier, Mauricio Soto, and Laurent Viennot. Treewidth and hyperbolicity of the internet. In *2011 IEEE 10th International Symposium on Network Computing and Applications*, pages 25–32, 2011. 30

[NN93] Joseph Naor and Moni Naor. Small-bias probability spaces: efficient constructions and applications. *SIAM Journal on Computing*, 22(4):838–856, 1993. 4, 40

[Pis96] Gilles Pisier. A simple proof of a theorem of Kirchberg and related results on C^*-norms. *J. Operator Theory*, 35(2):317–335, 1996. 66

[Pis18] Gilles Pisier. On a linearization trick. *Enseign. Math.*, 64(3-4):315–326, 2018. 12, 66, 67

[Que94] Gregory Quenell. Combinatorics of free product graphs. In *Geometry of the spectrum (Seattle, WA, 1993)*, volume 173 of *Contemp. Math.*, pages 257–281. Amer. Math. Soc., Providence, RI, 1994. 5, 22, 25

[RS12] José M. Rodríguez and José M. Sigarreta. Bounds on Gromov hyperbolicity constant in graphs. *Proc. Indian Acad. Sci. Math. Sci.*, 122(1):53–65, 2012. 30

[RVW02] Omer Reingold, Salil Vadhan, and Avi Wigderson. Entropy waves, the zig-zag graph product, and new constant-degree expanders. *Ann. of Math. (2)*, 155(1):157–187, 2002. 10, 22, 26

[Sal92] Maura Salvatori. On the norms of group-invariant transition operators on graphs. *J. Theoret. Probab.*, 5(3):563–576, 1992. 36

[Sar20] Peter Sarnak, 2020. Personal communication. 9

[Sho90] Victor Shoup. New algorithms for finding irreducible polynomials over finite fields. *Mathematics of Computation*, 54(189):435–447, 1990. 40

[SS96] Michael Sipser and Daniel Spielman. Expander codes. *Transactions on Information Theory*, 42(6):1710–1722, 1996. 4

[Sta83] John R. Stallings. Topology of finite graphs. *Invent. Math.*, 71(3):551–565, 1983. 29

[SW90] Paolo M. Soardi and Wolfgang Woess. Amenability, unimodularity, and the spectral radius of random walks on infinite graphs. *Math. Z.*, 205(3):471–486, 1990. 36

[Tro85] Vladimir Ivanovich Trofimov. Groups of automorphisms of graphs as topological groups. *Mat. Zametki*, 38(3):378–385, 476, 1985. 10, 33

[VK19] Jorge Garza Vargas and Archit Kulkarni. Spectra of infinite graphs via freeness with amalgamation. Technical Report 1912.10137, arXiv, 2019. 5, 10, 22, 25, 26, 69

[Woe00] Wolfgang Woess. *Random walks on infinite graphs and groups*, volume 138 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2000. 35, 36

[WZ11] Yaokun Wu and Chengpeng Zhang. Hyperbolicity and chordality of a graph. *Electron. J. Combin.*, 18(1):Paper 43, 22, 2011. 30

[Zno75] D. V. Znoĭko. Free products of nets and free symmetrizers of graphs. *Mat. Sb. (N.S.)*, 98(140)(4(12)):518–537, 639, 1975. 5, 22, 23

A Gallery

The appendix contains pictures of a few interesting MPL graphs.

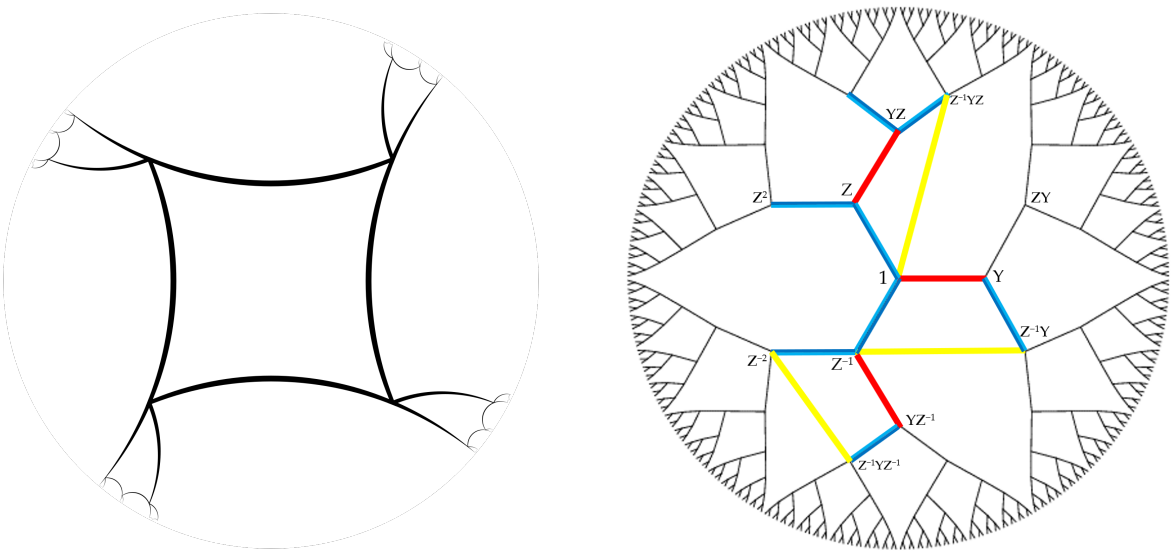


Figure 7: On the left, an illustration of $C_4 \star C_4$, the free product of two 4-cycles. On the right, a hint to how it arises as $\widetilde{\mathscr{G}_\infty}(Y + Z + Z^{-1} + Z^{-1}YZ)$.

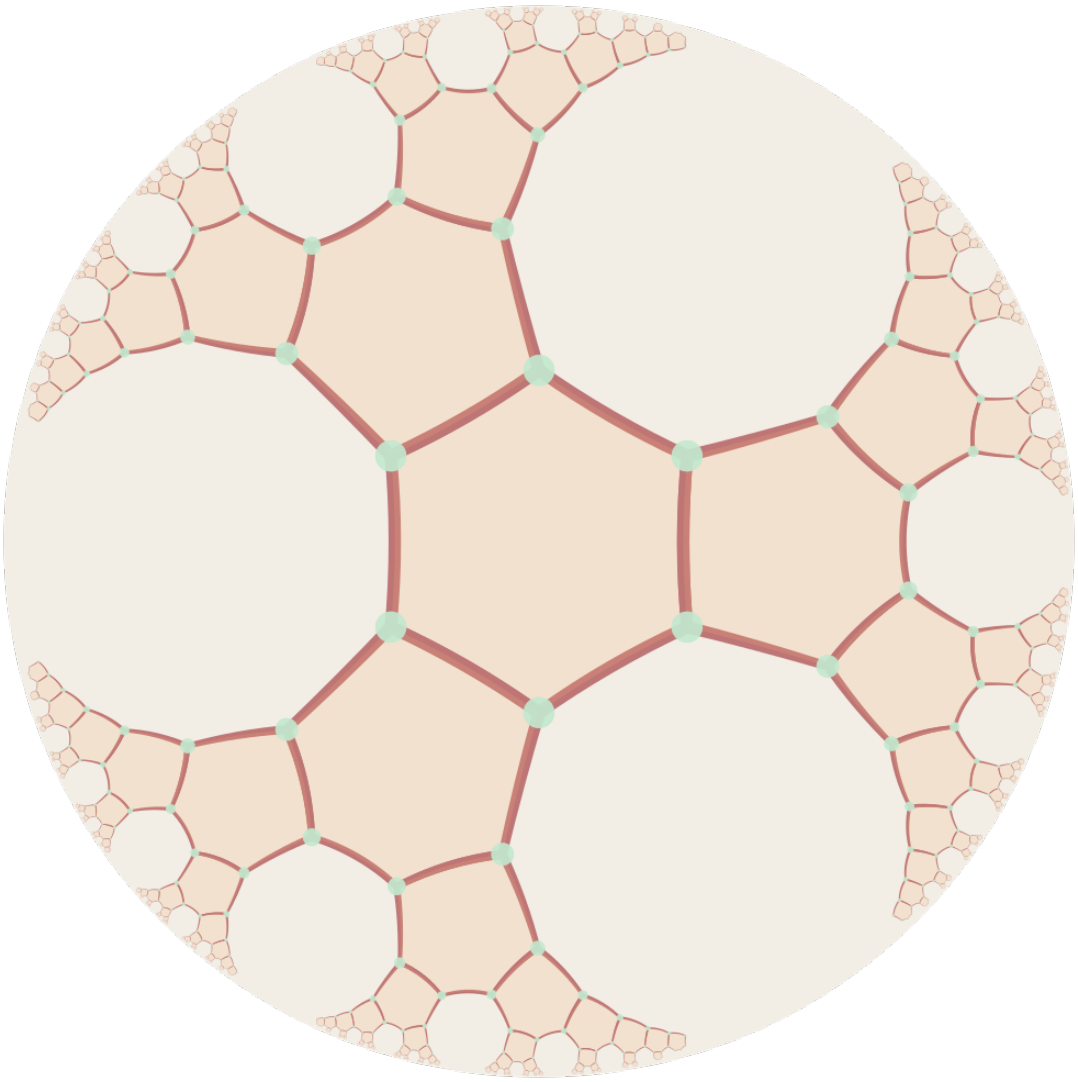


Figure 8: An illustration of $\widetilde{\mathscr{G}}_\infty(p)$ with

$$p = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} + |1\rangle\langle 4|\, Z_1 + |3\rangle\langle 6|\, Z_2 + |2\rangle\langle 5|\, Z_2 Z_1 + |4\rangle\langle 1|\, Z_1^* + |6\rangle\langle 3|\, Z_2^* + |5\rangle\langle 2|\, Z_1^* Z_2^*.$$

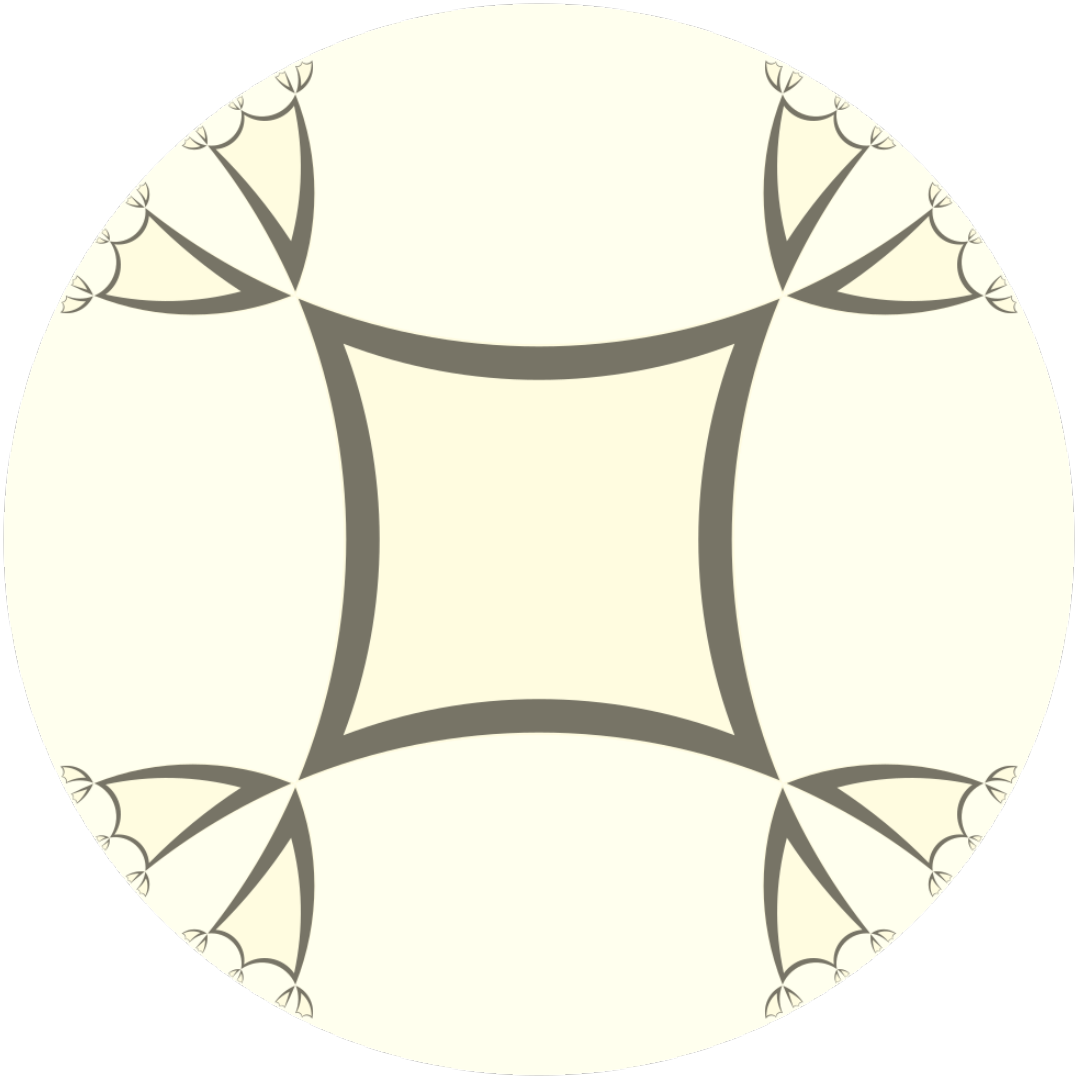


Figure 9: An illustration of $C_4 \star C_4 \star C_4$ as $\widetilde{\mathscr{G}}(p)$ with $r = 4$ and

$$p = \sum_{i=5}^{7}(|i-3\rangle\langle i-4|\, Z_i + |i-4\rangle\langle i-3|\, Z_i^*) + |4\rangle\langle 1|\, Z_7Z_6Z_5 + |1\rangle\langle 4|\, (Z_7Z_6Z_5)^* + \sum_{i=1}^{4}(|i\rangle\langle i|\, Z_i + |i\rangle\langle i|\, Z_i^* + |i\rangle\langle i|\, Y_i + Z_i^*Y_iZ_i).$$