Spectrum Preserving Short Cycle Removal on Regular Graphs

Pedro Paredes
Computer Science Department, Carnegie Mellon University, Pittsburgh, PA, USA

Abstract

We describe a new method to remove short cycles on regular graphs while maintaining spectral bounds (the nontrivial eigenvalues of the adjacency matrix), as long as the graphs have certain combinatorial properties. These combinatorial properties are related to the number and distance between short cycles and are known to happen with high probability in uniformly random regular graphs.

Using this method we can show two results involving high girth spectral expander graphs. First, we show that given $d \geqslant 3$ and n, there exists an explicit distribution of d-regular $\Theta(n)$-vertex graphs where with high probability its samples have girth $\Omega(\log_{d-1} n)$ and are ϵ-near-Ramanujan; i.e., its eigenvalues are bounded in magnitude by $2\sqrt{d-1}+\epsilon$ (excluding the single trivial eigenvalue of d). Then, for every constant $d \geqslant 3$ and $\epsilon > 0$, we give a deterministic poly(n)-time algorithm that outputs a d-regular graph on $\Theta(n)$-vertices that is ϵ-near-Ramanujan and has girth $\Omega(\sqrt{\log n})$, based on the work of [26].

2012 ACM Subject Classification Mathematics of computing → Spectra of graphs; Theory of computation → Expander graphs and randomness extractors

Keywords and phrases Ramanujan Graphs, High Girth Regular Graphs

Digital Object Identifier 10.4230/LIPIcs.STACS.2021.55

Related Version *Full Version*: `https://arxiv.org/abs/2002.07211`

Funding *Pedro Paredes*: 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).

Acknowledgements I am very grateful to Ryan O'Donnell for numerous comments and suggestions, as well as very thorough feedback on an earlier draft of this paper.

1 Introduction

Let's consider d-regular graphs of n vertices. The study of short cycles and *girth* (defined as the length of the shortest cycle of a graph) in such graphs dates back to at least the 1963 paper of Erdős and Sachs [10], who showed that there exists an infinite family with girth at least $(1-o_n(1))\log_{d-1} n$. On the converse side, a simple path counting argument known as the "Moore bound" shows that this girth is upper bounded by $(1+o_n(1))2\log_{d-1} n$. Though simple, this is the best known upper bound. Given these bounds, it is common to call an infinite family of d-regular n-vertex graphs *high girth* if their girth is $\Omega(\log_{d-1} n)$.

The first explicit construction of high girth regular graphs is attributed to Margulis [23], who gave a construction of graphs that achieve girth $(1-o_n(1))\frac{4}{9}\log_{d-1} n$. A series of works initiated by Lubotzky-Phillips-Sarnak [21] and then improved by several other people [24, 27, 19] culminated in the work of Dahan [9], who proves that for all large enough d there are explicit d-regular n-vertex graphs of girth $(1-o_n(1))\frac{4}{3}\log_{d-1} n$.

Another relevant problem consists of generating random distributions that produce regular graphs with high girth. Results regarding the probabilistic aspects of certain structures (like cycles) in graphs often give us tools to count the number of graphs that satisfy certain conditions, like how many regular graphs have girth at least some value. The distribution of short cycles in uniformly random regular graphs was first studied by Bollobás [7], who proved, that for a fixed k the random variables representing the number of cycles of length exactly k in a uniformly random d-regular graph are asymptotically independent Poisson with mean $(d-1)^k/2k$. Subsequently, McKay-Wormald-Wysocka [25] gave a more precise description of this by finding the asymptotic probability of a random d-regular graph having a certain number of cycles of any length up to $c \log_{d-1} n$, for $c < 1/2$. More recently, Linial and Simkin [20] showed that a random greedy algorithm that is given $d \geqslant 3$, $c \in (0,1)$ and an even n, produces a d-regular n-vertex graph with girth at least $c \log_{d-1} n$ with high probability.

The literature of regular graphs with high girth is closely connected to the literature of *spectral expanders*. Before defining this, let's consider some notation.

▶ **Definition 1.** *Let G be an n-vertex d-regular multigraph. We write $\lambda_i = \lambda_i(G)$ for the eigenvalues of its adjacency matrix A_G, and we always assume they are ordered with $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$. A basic fact is that $\lambda_1 = d$ always; this is called the trivial eigenvalue and corresponds to the all ones vector. We also write $\lambda(G) = \max\{\lambda_2, |\lambda_n|\}$.*

Roughly, a graph with good spectral expansion properties is a graph that has small λ. More formally, an infinite sequence (G_n) of d-regular graphs is called a family of expanders if there is a constant $\delta > 0$ such that $\lambda(G) \leqslant (1-\delta)d$ for all n, or in other words, all eigenvalues are strictly separated from the trivial eigenvalue. This terminology was first introduced by [29] and later it was shown [1] that uniformly random d-regular graphs are spectral expanders with high probability.

The celebrated Alon-Boppana bound shows that λ cannot be arbitrarily small:

▶ **Theorem 2** ([1, 28, 11])**.** *For any d-regular n-vertex graph G we have that $\lambda_2(G) \geqslant 2\sqrt{d-1} - O(1/\log^2 n)$.*

Using some number-theoretic ideas, Lubotzky-Phillips-Sarnak [21], and independently Margulis [24], proved this bound is essentially tight by showing the existence of infinite families of d-regular graphs that meet the bound $\lambda(G) \leqslant 2\sqrt{d-1}$, if $d-1$ is an odd prime. In light of this, Lubotzky-Phillips-Sarnak introduced the following definition:

▶ **Definition 3** (Ramanujan graphs)**.** *A d-regular graph G is called Ramanujan whenever $\lambda(G) \leqslant 2\sqrt{d-1}$.*

These results were improved by Morgenstern [27], who showed the same for all d where $d-1$ is a prime power.

It is still open whether there exist infinite families of Ramanujan graphs for all d. However, if one relaxes this to only seek ϵ-near-Ramanujan graphs (graphs that satisfy $\lambda \leqslant 2\sqrt{d-1}+\epsilon$), then the answer is positive. Friedman [12] proved that uniformly random d-regular n-vertex graphs satisfy $\lambda \leqslant 2\sqrt{d-1} + o_n(1)$ with high probability. This proof was recently simplified by Bordenave [8].

▶ **Theorem 4** ([12, 8])**.** *Fix any $d > 3$ and $\epsilon > 0$ and let G be a uniformly random d-regular n-vertex graph. Then*

$$\boldsymbol{Pr}\left[\lambda(G) \leqslant 2\sqrt{d-1} + \epsilon\right] \geqslant 1 - o_n(1).$$

In fact [8], G achieves the subconstant $\epsilon = \widetilde{O}(1/\log^2 n)$ with probability at least $1 - 1/n^{.99}$.

Recently, it was shown how to achieve a result like the above but deterministically [26]. We write a more precise statement of this below.

▶ **Theorem 5** ([26]). *Given any n, $d \geqslant 3$ and $\epsilon > 0$, there is a deterministic polynomial-time algorithm that constructs a d-regular N-vertex graph with the following properties:*
- $N = n(1 + o_n(1))$;
- $\lambda(G) \leqslant 2\sqrt{d-1} + \epsilon$;

We refer the reader interested in a more thorough history of the literature of Ramanujan graphs to the introduction of [26]. Also, for a comprehensive list of applications and connections of Ramanujan graphs and expanders to computer science and mathematics, see [14].

In this work we concern ourselves with bridging these two worlds, looking for families of regular graphs that are both good spectral expanders and also have high girth. This bridge can be seen in several of the aforementioned works. The explicit construction of high girth regular graphs by Margulis [23] was a motivator to his work on Ramanujan graphs [24]. Additionally, the constructions of [21] and [27] produce graphs that are both Ramanujan and have girth $(1 - o_n(1))\frac{4}{3}\log_{d-1} n$, according to the previously stated restrictions on d.

More recently, Alon-Ganguly-Srivastava [3] showed that for a given d such that $d-1$ is prime and $\alpha \in (0, 1/6)$, there is a construction of infinite families of graphs with girth at least $(1 - o_n(1))(2/3)\alpha \log_{d-1} n$ and λ at most $(3/\sqrt{2})\sqrt{d-1}$ with many eigenvalues localized on small sets of size $O(n^\alpha)$. Their motivation comes from the theory of quantum ergodicity in graphs, which relates high-girth expanding graphs to delocalized eigenvectors. See [3] for more on this. Our main result is based on some of the techniques of this work.

One other motivation to search for graphs with simultaneous good spectral expansion and high girth is its application to the theory of error-correcting codes, particularly for *Low Density Parity Check* or *LDPC* codes, originally introduced by Gallager [13]. The connection with high girth regular graphs was first pointed out by Margulis in [23]. The property of high-girth is desirable since the decoding of such codes relies on an iterative algorithm whose performance is worse in the presence of short cycles. Additionally, using graphs with good spectral properties to generate these codes heuristically seems to lead to good performance, as pointed out by several works [30, 18, 22].

1.1 Our results

We can now state our results and put them in perspective. Let's first introduce some useful definitions and notation.

▶ **Definition 6** (Bicycle-free at radius r). *A multigraph is said to be bicycle-free at radius r if the distance-r neighborhood of every vertex has at most one cycle.*

▶ **Definition 7** ((r, τ)-graph). *Let r and τ be a positive integers. Then, we call a graph G a (r, τ)-graph if it satisfies the following conditions:*

- *G is bicycle-free at radius at least r;*
- *The number of cycles of length at most r is at most τ.*

Our main result is the following short cycle removal theorem:

▶ **Theorem 8.** *There exists a deterministic polynomial-time algorithm* fix *that, given as input a d-regular n-vertex (r, τ)-graph G satisfying $r \leqslant (2/3)\log_{d-1}(n/\tau) - 5$ outputs a graph* $\mathrm{fix}(G)$ *satisfying*

- $\mathrm{fix}(G)$ *is a* d*-regular graph with* $n + O(\tau \cdot (d-1)^{r/2+1})$ *vertices;*
- $\lambda(\mathrm{fix}(G)) \leqslant \max\{\lambda(G), 2\sqrt{d-1}\} + O_d(1/r)$*;*
- $\mathrm{fix}(G)$ *has girth at least* r*.*

The key fact in our proof of this statement is a theorem proved by Kahale [15], originally used to construct Ramanujan graphs with better expansion of sublinear sized subsets. See also [3] and [2] for other applications of this technique. We will prove this theorem in Section 2.

The preconditions of this theorem are not arbitrary. Even though random uniformly n-vertex d-regular graphs have constant girth with high probability, they are bicycle-free at radius $\Omega(\log_{d-1} n)$ and the number of cycles of length at most $c \log_{d-1} n$ (for small enough c) is $o(n)$ with high probability. Recall that from Theorem 4 we also know that being near-Ramanujan is also a property that occurs with high probability in random regular graphs. So a statement like the above can be used to produce distributions over regular graphs that have high girth and are near-Ramanujan with high probability. With this in mind, we introduce the following definition:

▶ **Definition 9** ((Λ, g)-good graphs)**.** *We call a graph* G *a* (Λ, g)*-good graph if* $\lambda(G) \leqslant \Lambda$ *and girth*$(G) \geqslant g$*.*

Let $\mu_d(n)$ *be a distribution over* d*-regular graphs with* $\sim n$ *vertices. We say* $\mu_d(n)$ *is* (Λ, g)*-good if* $G \sim \mu_d(n)$ *is* (Λ, g)*-good with probability at least* $1 - o_n(1)$*.*

Additionally, we call the distribution explicit if sampling an element is doable in polynomial time.

We shall prove the following using Theorem 8 in Appendix A:

▶ **Theorem 10.** *Given* $d \geqslant 3$ *and* n*, let* G *be a uniformly random* d*-regular* n*-vertex graph. For any* $c < 1/4$ *and* $\epsilon > 0$*,* $\mathrm{fix}(G)$ *is a* $(2\sqrt{d-1} + \epsilon, c \log_{d-1} n)$*-good explicit distribution.*

Recall that the upper bound on the girth of a regular graph is $(1 + o_n(1))2 \log_{d-1} n$, so this distribution has optimal girth up to a constant. Based on our proof of the above and using some classic results about the number of d-regular n-vertex graphs, we can show a lower bound on the number of $(2\sqrt{d-1} + \epsilon,\ c \log_{d-1} n)$-good graphs in some range.

▶ **Corollary 11.** *Let* $d \geqslant 3, n$ *be integers and* $\epsilon > 0, c > 1/4$ *reals. The number of* d*-regular graphs with number of vertices in* $[n, n + O(n^{3/8})]$*, which are* $(2\sqrt{d-1} + \epsilon,\ c \log_{d-1} n)$*-good, is at least*

$$\Omega\left(\left(\frac{d^d n^d}{e^d (d!)^2}\right)^{n/2}\right).$$

We prove both of these results in Appendix A.

Finally, we show a slightly stronger version of result of [26] by plugging our short cycle removal theorem into their construction.

▶ **Theorem 12.** *Given any integer* n *and constants* $d \geqslant 3$*,* $\epsilon > 0$ *and* c*, there is a deterministic polynomial-time (in* n*) algorithm that constructs a* d*-regular* N*-vertex graph with the following properties:*

- $N = n(1 + o_n(1))$*;*
- $\lambda(G) \leqslant 2\sqrt{d-1} + \epsilon$*;*
- G *has girth at least* $c\sqrt{\log n}$*.*

Note that this only works for large enough n. Also, the running time from the theorem above has an exponential dependency on d, ϵ and c. The proof of this statement as well as the precise dependencies on these constants will be worked out in Appendix B.

1.2 Models of random regular graphs

We will introduce some classic models of random regular graphs, which we will use throughout the paper.

▶ **Definition 13** ($\mathcal{G}_d(n)$)**.** *Let $\mathcal{G}_d(n)$ denote the set of d-regular n-vertex graphs. We write $G \sim \mathcal{G}_d(n)$ to denote that G is sampled uniformly at random from $\mathcal{G}_d(n)$.*

Sampling from $\mathcal{G}_d(n)$ is not easy a priori; the standard way to do so is using the *configuration model*, which was originally defined by Bollobás [7].

▶ **Definition 14** (Configuration model)**.** *Given integers $n > d > 0$ with nd even, the configuration model produces a random n-vertex, d-regular undirected multigraph (with loops) G. This multigraph is induced by a uniformly random matching on the set of "half-edges", $[n] \times [d] \cong [nd]$ (where $(v, i) \in [n] \times [d]$ is thought of as half of the ith edge emanating from vertex v). Given a matching, the multigraph G is formed by "attaching" the matched half-edges.*

This model corresponds exactly to the uniform distribution on not necessarily simple d-regular n-vertex graphs. It also not hard to see that the conditional distribution of the d-regular n-vertex configuration model when conditioned on it being a simple graph is exactly the uniform distribution on $\mathcal{G}_d(n)$. The probability that the sampled graph is simple is $\Omega_d(1)$.

The configuration model has the advantage that is easy to sample and to analyze. For reference, the proof of Theorem 4 was done in terms of the configuration model and so the theorem also applies to it.

2 Short cycles removal

In this section we prove Theorem 8. Recall that we are given a d-regular n-vertex (r, τ)-graph G with the constraint specified in Theorem 8 and we wish to find some d-regular graph $\mathrm{fix}(G)$ on $\sim n$ vertices such that $\lambda(\mathrm{fix}(G)) \leqslant \lambda(G) + o_r(1)$ and its girth is at least r.

Briefly, the algorithm that achieves this works by removing one edge per small cycle from G, effectively breaking apart all such cycles, and then fixing the resulting off degree vertices by adding d-ary trees in a certain way. We will now more carefully outline this method and then proceed to fill in some details as well as show it works as desired.

Before starting, we introduce some notation which will be helpful.

▶ **Definition 15** ($\mathrm{Cyc}_g(G)$)**.** *Given a graph G, let $\mathrm{Cyc}_g(G)$ denote the collection of all cycles in G of length at most g. Recall that if $\mathrm{Cyc}_g(G)$ is empty then G is said to have girth exceeding g.*

▶ **Definition 16** ($B_\delta(S)$)**.** *Given a set of vertices S in a graph G, let $B_\delta(S)$ denote the collection of vertices in G within distance δ of S. We will occasionally abuse this notation and write $B_\delta(v)$ instead of $B_\delta(\{v\})$ for a vertex v.*

Let E_c be a set containing exactly one arbitrary edge per cycle in $\mathrm{Cyc}_r(G)$. Note that the bicycle-freeness property implies E_c is a matching. Let H_t be a graph with the same vertex set as G obtained by removing all edges in E_c from G. To prevent ambiguity, whenever we pick something arbitrarily let's suppose the algorithm fix uses the lexicographical order of node labels as a tiebreaker. We also partition the endpoints of each edge as described in the following definition:

▶ **Definition 17** ($V_i(E)$). *Given a matching E, we let $V_1(E)$ and $V_2(E)$ be two disjoint sets of vertices constructed as follows: for all $e = (u, v) \in E$ place u in $V_1(E)$ and v in $V_2(E)$ (so each endpoint is in exactly one of the two sets).*

Note that according to the above definition we have $|V_1(E_c)| = |V_2(E_c)| = |E_c| \leqslant \tau$. For ease of notation we also define:

▶ **Definition 18** ($\phi_E(v)$). *Given a matching E and $(u, v) \in E$ such that $u \in V_1(E)$ and $v \in V_2(E)$, we denote by ϕ_E the function that maps endpoints to endpoints, so we have $\phi_E(u) = v$ and $\phi_E(v) = u$.*

We will often abuse notation and drop the E from ϕ_E when it is clear from context.

Since we break apart each cycle in $\mathrm{Cyc}_r(G)$, we can conclude that H_t has girth greater than r. However, note that in removing edges, H_t is no longer d-regular.

To fix this, consider the following object which we refer to as a d-regular tree of height h: a finite rooted tree of height h where the root has d children but all other non-leaf vertices have $d-1$ children. This definition implies that every non-leaf vertex in a d-regular tree has degree d.

We shall add two d-regular trees to H_t in order to fix the off degrees, while maintaining the desired girth and bound on λ. The idea of using d-regular trees is based on the degree-correction gadget used in [3] for their construction of high-girth near-Ramanujan graphs with localized eigenvectors. As such, we will use some of the tools used in their proofs.

Let h be an integer parameter we shall fix later. Let T_1 and T_2 be two d-regular trees of height h and let L_1 and L_2 be the sets of leaves of each one. Note that $|L_1| = |L_2| = d(d-1)^{h-1} \approx (d-1)^h$. We shall add the two trees to H_t and then pair up elements of $V_1(E_c)$ with elements of L_1 (and analogously for $V_2(E_c)$ and L_2) and merge the paired up vertices. However, we have to deal with two potential issues:

- $|L_i| \neq |V_i(E_c)|$, in which case we cannot get an exact pairing between these sets;
- This procedure might result in the creation of small cycles (potentially even cycles of length $O(1)$).

To expand on the latter point, we describe a potential problematic instance. Suppose we can somehow pick h such that $|L_i| = |V_i(E_c)|$ and then arbitrarily pair up their elements. Suppose there are two edges in E_C corresponding to two cycles of constant length and denote their endpoints by $v_1 \in V_1(E_C), v_2 \in V_2(E_C)$ and $u_1 \in V_1(E_C), u_2 \in V_2(E_C)$. If the distance in T_1 of v_1 and u_1 given by the pairing of $V_1(E_c)$ and L_1 is small (constant, for example) and the same applies to the distance in T_2 of v_2 and u_2, then there is a cycle of small length (constant, for example) in the graph resulting from adding the two trees to H_t.

To address this issue we remove some extra edges from G that are somehow "isolated" and group them with edges from E_C. The goal is to have the endpoints of any two edges in E_C be far apart in T_1 and T_2 distance, but close to some of the endpoints of the extra edges. With this in mind, we set $h = \lceil \log_{d-1} \tau \rceil + \lceil r/2 \rceil + 1$ so that $|L_i| \approx \tau \cdot (d-1)^{r/2+1}$, which is close to the number of extra edges we want to remove. This choice will also be helpful later when we analyze the spectral properties of the construction.

Formally, this leads us to the following proposition:

▶ **Proposition 19.** *There is a set of edges E_t of G such that the following is true for $i \in \{1, 2\}$:*

- $|V_i(E_t) \cup V_i(E_c)| = d(d-1)^{h-1}$;
- *for all distinct $u, v \in V_i(E_t) \cup V_i(E_c)$, we have $v \notin B_r(u)$ and $u \notin B_r(v)$.*

Additionally, we can find such a set in polynomial time.

Proof. We will describe the efficient algorithm that does this.

We are going to incrementally grow our set E_t, one edge at the time, until $|V_i(E_t) \cup V_i(E_c)| = d(d-1)^{h-1}$, so suppose E_t is initially an empty set. We start by, for all $e = (u, v) \in E_c$, marking all vertices in $B_{1+r}(\{v, u\})$. Note that we marked at most $\tau \cdot (d(d-1)^r) \leqslant 2\tau(d-1)^{r+1}$ vertices.

Notice that, since we marked all vertices at distance $1 + r$ from any vertex in $V_i(E_c)$, we can safely pick any unmarked vertex and an arbitrary neighbor and add that edge to E_t.

We can now describe a procedure to add a single edge to E_t:

- Pick an unmarked vertex u and an arbitrary neighbor v of u;
- Add (u, v) to E_t;
- Mark all vertices in $B_{1+r}(\{u, v\})$.

By the same reasoning as before, as long as we have an unmarked vertex, this procedure works. If we repeat the above t times, we are left with at least $n - 2\tau(d-1)^{r+1} - 2t(d-1)^{r+1}$ unmarked vertices. We claim the procedure can be successfully repeated at least $2\tau(d-1)^{r/2+2}$ times. In such a case, the number of unmarked vertices left is at least:

$$n - 2\tau(d-1)^{r+1} - 4\tau(d-1)^{r/2+2}(d-1)^{r+1} \geqslant n - 6\tau(d-1)^{3r/2+3},$$

which is always greater than 0 when $r \leqslant \frac{2}{3}\log_{d-1}(n/\tau) - 5$. Hence, we always have at least one unmarked vertex to pick throughout the procedure.

Note that the number of repetitions we require exactly matches the size of $|E_t|$ so we need this to be exactly $d(d-1)^{h-1} - \tau \leqslant 2\tau(d-1)^{r/2+2}$, which means our algorithm always succeeds. ◀

We will state some simple properties of this construction that will be relevant later on.

▶ **Fact 20.** $|V_i(E_t)| \geqslant \tau \cdot (d-1)^{\lceil r/2 \rceil}$

Proof. We simply have: $|V_i(E_t)| = |E_t| = d(d-1)^{h-1} - \tau \geqslant \tau \cdot (d-1)^{\lceil r/2 \rceil}$. ◀

▶ **Fact 21.** *For all $e \in E_t$, there is at most one cycle in $B_r(e)$ in G and if there is a cycle it has length greater than r.*

Proof. That there is at most one cycle in $B_r(e)$ is obvious since G is bicycle-free at radius r. So, let's suppose there is a cycle C in $B_r(e)$ with length less than or equal to r. Then, there is at least one edge $e' \in C$ that is also in E_c, but in that case $e' \in B_r(e)$, which contradicts the definition of E_t. ◀

We can now extend our definition of H_t. Let H be the graph obtained from G by removing all edges in E_c and in E_t.

Recall our plan to add T_1 and T_2, two d-regular trees of height h (recall $h = \lceil \log_{d-1} \tau \rceil + \lceil r/2 \rceil + 1$), to H while pairing up elements of L_i with endpoints of removed edges. We will now describe a pairing process that achieves high girth (and later we will see how it also achieves low λ).

First, consider a canonical ordering of L_1 and L_2 based on visit times from a breath-first search, as illustrated in Figure 1 for $d = 3$. Given this ordering, the following is easy to see:

▶ **Fact 22.** *The tree distance between two leaves with indices i and j is at least $2(1 + \log_{d-1}((|i-j|+1)/d))$.*

Proof. Let's show that the lowest common ancestor of the two leaves is at least $1+\log_{d-1}((|i-j|+1)/d)$, this proves the claim since we need to travel this distance twice, from the ith indexed leaf to the ancestor and then back to the jth indexed leaf. Let V_0 be the set of $|i-j|+1$ leaves with indices between i and j. Let's construct the smallest subtree that includes V_0 from bottom up and compute its height, which is an upper bound to the desired lowest common ancestor. First, group elements of V_0 in groups of at most $d-1$ consecutive indices and add one representative of each group to a set V_1. Each group corresponds to a node that parents all of its elements. There are at most $|V_0|/(d-1)$ such groups, so $|V_1| \leqslant |V_0|/(d-1)$. Repeat the same procedure until $|V_a| \leqslant 1$, in which case a is an upper bound to the height of the goal subtree, and by induction we have that $|V_{i+1}| \leqslant |V_i|/(d-1)$, so $a \geqslant \log_{d-1}|V_0|$.

This is not quite right because if the last grouping corresponds to the root of the tree, we need to group elements in d groups, because this is the degree of the root, so by accounting for this we have $a \geqslant 1+\log_{d-1}(|V_0|/d)$. ◀

Now, consider the following pairing of elements in L_1 and $V_1(E_t) \cup V_1(E_c)$: pick an arbitrary element of $V_1(E_c)$ and pair it up with the first leaf of L_1. Now pick $(d-1)^{\lceil r/2 \rceil}$ distinct elements of $V_1(E_t)$ and pair them up with the next leaves of L_1. Repeat this procedure, of pairing one element of $V_1(E_c)$ with $(d-1)^{\lceil r/2 \rceil}$ elements of $V_1(E_t)$ with a contiguous block of leaves until we exhaust all elements of $V_1(E_c)$. Note that by Fact 20, there always are enough elements in E_t to perform this pairing. Pair up any remaining leaves with the remaining elements of $V_1(E_t)$ arbitrarily. Now repeat the same procedure but for L_2 and $V_2(E_t) \cup V_2(E_c)$ with the same groupings (so the endpoints of an edge in either E_t or E_c are mapped to the same leaves of L_1 and L_2). This pairing procedure is pictured in Figure 2 below.

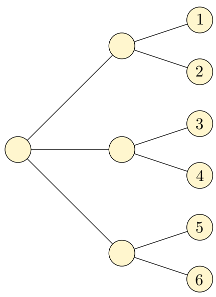

Figure 1 Leaf ordering for $d = 3$.

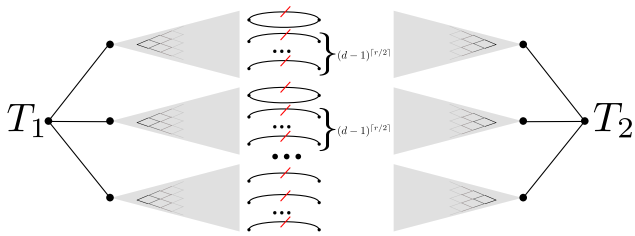

Figure 2 Example pairing.

Let $\mathrm{fix}(G)$ be defined as the graph resulting from applying the method described in the previous paragraph to fix the degrees of H. It is now obvious that $\mathrm{fix}(G)$ is a d-regular graph and we only add $|T_1|+|T_2| = O(\tau \cdot (d-1)^{r/2+1})$ new vertices, so it has $n+O(\tau \cdot (d-1)^{r/2+1})$ total vertices. We will now analyze the resulting girth and λ value and prove Theorem 8 in the process.

2.1 Analyzing the girth of $\mathrm{fix}(G)$

Here we prove that the girth of $\mathrm{fix}(G)$ is at least r. Let's start by supposing, for the sake of contradiction, that there is a cycle C of length less than r. We know that the girth of H is more than r by definition, so C has to use an edge from T_1 or T_2. Without loss of generality, let's assume that C contains at least one edge from T_1. Since T_1 is a tree, C has to eventually exit T_1 and use some edges from H, so in particular it uses some vertex $v \in L_1$. We will show that in this case C has length at least r, which is a contradiction. Thus, we have to handle two cases: $v \in V_1(E_c)$ and $v \in V_1(E_t)$.

Let us start with the $v \in V_1(E_c)$ case. Let's follow C starting in v and show that to loop back to v, C would require to traverse at least r edges. So, we start in v and go into T_1 by following the only edge in T_1 that connects to v. Then, the cycle C has to use some edges from T_1 and finally exit through some other vertex in L_1 before eventually looping back to v. Suppose that $u \in L_1$ is such a vertex. Due to our grouping of elements in E_t with $(d-1)^{\lceil r/2 \rceil}$ elements in E_c, if u is in $V_1(E_c)$, we know that the tree indices of v and u differ by at least $(d-1)^{\lceil r/2 \rceil}$. Hence, plugging this into the bound from Fact 22, the tree distance between v and u is at least $r-1$, which would imply C has length at least r. So u has to be in $V_1(E_t)$.

Continuing our traversal of C, we now exit T_1 through u and need to loop back to v. From our construction in Proposition 19 we know that the distance in H between v and u is at least r, so any short path in $\mathrm{fix}(G)$ between these vertices has to go through T_1 or T_2. Again, our Proposition 19 construction gives that the distance in H between v and any other vertex in L_1 is at least r, so such a short path will have to use some edges in T_2.

Finally, we claim that the distance from u to any vertex w in L_2 is at least r. If $w \neq \phi(u)$, we know from our Proposition 19 construction that the distance between u and w is at least r. Otherwise, if there is a path P of length less than r from u to w, then the cycle $P + uw$ has length at most r and is in $B_r(\{u, w\})$, which contradicts Fact 21. In conclusion, it is not possible to loop back to v using less than r steps, which concludes the proof of the $v \in V_1(E_c)$ case.

The proof for the $v \in V_1(E_t)$ case is already embedded in the previous proof, so we will just sketch it. Using the same argument we start by following C into T_1 and eventually exiting through some vertex $u \in V_1(E_t)$. As we saw before, the H distance between u and v is at least r and the H distance between u and any other vertex in L_1 or any vertex in L_2 is at least r, so we cannot loop back to v from u, which concludes the proof of this case.

2.2 Bounding $\lambda(\mathrm{fix}(G))$

We finally analyze the spectrum of $\mathrm{fix}(G)$ by proving that $\lambda(\mathrm{fix}(G)) \leqslant \lambda(G) + O_d(1/r)$. This argument is similar to the proof in Section 4 of [3], but adapted to our construction.

First, observe that the adjacency matrix of $\mathrm{fix}(G)$, which we will denote by simply A, can be written in the following way: $A = A_G - A_{E_c} - A_{E_t} + A_{T_1} + A_{T_2}$, where A_G is the adjacency matrix of G defined on the vertex set of $\mathrm{fix}(G)$ (which is to say G with a few isolated vertices from the added trees), A_{E_c} is the adjacency matrix of the cycle edges removed, and so on. Also, let V_G be the set of vertices from G, V_1 the set of vertices from T_1 and V_2 the set of vertices from T_2, so $V = V_G \cup V_1 \cup V_2$. In this section we will prove $\lambda(A) \leqslant \lambda(G) + O_d(1/r)$.

Let g be any unit eigenvector of A orthogonal to the all ones vector, so $\sum_{v \in V} g_v^2 = 1$ and $\sum_{v \in V} g_v = 0$. We have that $|\sum_{v \in V_1 \cup V_2} g_v| \leqslant \sqrt{2|T_i|}$ by Cauchy-Schwarz (since this vector is supported on only $2|T_i|$ entries), which in turn implies that $|\sum_{v \in V_G} g_v| \leqslant \sqrt{2|T_i|}$.

It suffices to show that $|g^T A g| \leqslant \lambda(G) + O_d(1/r)$. To do so, we shall analyze the contributions of A_G, A_{E_c}, A_{E_t}, A_{T_1} and A_{T_2} to $|g^T A g|$.

To bound the contribution of A_{T_1} and A_{T_2}, we use a lemma proved by Alon-Ganguly-Srivastava:

▶ **Lemma 23** ([3, Lemma. 4.1]). *Let W_i be the set of non-leaf vertices of T_i. Then for any vector f we have:*

$$|f^T A_{T_i} f| \leqslant 2\sqrt{d-1} \sum_{w \in W_i} f_w^2 + \sqrt{d-1} \sum_{v \in L_i} f_v^2.$$

Recall that the edges in $E_t \cup E_c$ define a perfect matching between L_1 and L_2, so we have the following:

$$|g^T(A_{E_c} + A_{E_t})g| = \left|\sum_{uv \in E_t \cup E_c} 2g_u g_v\right| \leqslant \sum_{v \in L_1 \cup L_2} g_v^2.$$

Finally, let g_G be the projection of g to the subspace spanned by V_G. Observe that $|g^T A_G g| = |g_G^T A_G g_G|$. Now, let $\mathbf{1}_G$ be the all ones vector supported on the set V_G and $g_\perp$ be a vector orthogonal to $\mathbf{1}_G$ such that $g_G = a\mathbf{1}_G + g_\perp$, for some constant a. We have that $\mathbf{1}_G^T g_G = a\mathbf{1}_G^T \mathbf{1}_G$, which implies

$$|a| = \left|\frac{\sum_{v \in V_G}(g_G)_v}{n}\right| \leqslant \frac{\sqrt{2|T_i|}}{n}.$$

Now observe:

$$|g_G^T A_G g_G| \leqslant |g_\perp^T A_G g_\perp| + |(a\mathbf{1}_G)^T A_G(a\mathbf{1}_G)| \leqslant \lambda(G)\sum_{v \in V_G} g_v^2 + \frac{2|T_i|d}{n}.$$

Note that $\sum_{v \in V_G} g_v^2 \leqslant 1$. We claim that the term $\frac{2|T_i|d}{n}$ is $O_d(1/r)$. We have $|T_i| = O(\tau \cdot (d-1)^{r/2+1})$ and we know from the problem constraints that $r \leqslant (2/3)\log_{d-1}(n/\tau) - 5$ which implies $\tau \cdot (d-1)^{r/2+1}/n \leqslant O((d-1)^{-r}) = O_d(1/r)$.

We can now plug everything together and apply Lemma 23 to obtain:

$$|g^T Ag| \leqslant \lambda(G) + (\sqrt{d-1}+1)\sum_{v \in L_1 \cup L_2} g_v^2 + O_d(1/r).$$

We will conclude our proof by showing that $\sum_{v \in L_1 \cup L_2} g_v^2$ is $O(1/r)$. It should be clear from the symmetry of our construction that we only need to prove $\sum_{v \in L_1} g_v^2 = O(1/r)$, since the same is analogous for L_2.

The following lemma can be proved using a known method by Kahale [15, Lemma 5.1]. This statement is similar to one found in [2, Lemma 3.2] and its proof is also very similar. For completeness, we present a self-contained proof of that based on the one from [2].

▶ **Lemma 24.** *Let v be some vertex of V. Let l be a positive integer such that $B_l(v)$ forms a tree. Let X_i be the set of all vertices at distance exactly i from v in $\mathrm{fix}(G)$, so $X_0 = \{v\}$. Let f be any non zero eigenvector with eigenvalue $|\mu| \geqslant 2\sqrt{d-1}$. Then, for $1 \leqslant i \leqslant l$:*

$$\sum_{u \in X_i} f^2(u) \geqslant \sum_{u \in X_{i-1}} f^2(u)$$

Proof. We will proceed by induction on i. First of all, let's establish the $i = 1$ case. Note that we have $\sum_{u \in X_1} f(u) = \mu f(v)$.

By Cauchy-Schwarz we get $d \cdot \sum_{u \in X_1} f^2(u) \geqslant \mu^2 f^2(v)$, and using the fact that $|\mu| \geqslant 2\sqrt{d-1}$ we obtain the desired:

$$\sum_{u \in X_1} f^2(u) \geqslant \frac{\mu^2}{d} f^2(v) \geqslant f^2(v).$$

Let's now assume that the statement is true for $i-1$ and prove that this implies it is true for i. Let u be some vertex in X_{i-1}. Recall that $B_l(v)$ is a tree and let u' be its parent in X_{i-2} and $w_1, \ldots w_{d-1}$ be its children in X_i. We have $f(u') + \sum_{i=1}^{d-1} f(w_i) = \mu f(u)$. Note that $f(u') = \sqrt{d-1}f(u')/\sqrt{d-1}$ and apply Cauchy-Schwarz to obtain:

$$\left(\frac{f^2(u')}{d-1} + \sum_{i=1}^{d-1} f^2(w_i)\right)(2d-2) \geqslant \mu^2 f^2(u),$$

which implies

$$\frac{f^2(u')}{d-1} + \sum_{i=1}^{d-1} f^2(w_i) \geqslant \frac{\mu^2}{2d-2} f^2(u) \geqslant 2f^2(u),$$

where the last inequality follows from the fact that $|\mu| \geqslant 2\sqrt{d-1}$.

We can finally sum the above for all $u \in X_{i-1}$, noting that from the fact that $B_l(v)$ is a tree we know that each element in X_{i-2} appears $d-1$ times (as the parent of $d-1$ vertices) and each element in X_i appears once:

$$\sum_{u \in X_{i-2}} f^2(u) + \sum_{u \in X_i} f^2(u) \geqslant 2 \sum_{u \in X_{i-1}} f^2(u).$$

We now apply the induction hypothesis and obtain the result:

$$\sum_{u \in X_i} f^2(u) \geqslant 2 \sum_{u \in X_{i-1}} f^2(u) - \sum_{u \in X_{i-2}} f^2(u) \geqslant \sum_{u \in X_{i-1}} f^2(u).$$ ◀

Our plan is to pick the parameters l and v from Lemma 24 and use it to show that $\sum_{v \in L_1} g_v^2 = O(1/r)$. Let μ be the eigenvalue associated with g and suppose that $|\mu| > 2\sqrt{d-1}$, otherwise $|\mu| \leqslant \lambda(G)$, which would imply the result. Set v to be the root of T_1. We will show that if we pick $l = h + \lfloor r/2 \rfloor$, where $h = \lceil \log_{d-1} \tau \rceil + \lceil r/2 \rceil + 1$ is the height of T_1 and T_2, then $B_l(v)$ forms a tree.

Note that $B_h(v)$ is exactly T_1, so it obviously forms a tree. To observe what happens in $B_l(v) \setminus B_h(v)$, we first prove the following proposition, whose proof uses some of the ideas of Section 2.1:

▶ **Proposition 25.** *Let u be a vertex in L_1. Let $\mathcal{P}(u)$ be the set of non-empty paths that start in u and whose first step does not go into T_1. Then, the shortest path in $\mathcal{P}(u)$ that ends in any vertex in L_1 has length at least r.*

Proof. As in the previous girth proof, we have two cases, $u \in V_1(E_c)$ and $u \in V_1(E_t)$. The latter case is obvious from the proof in Section 2.1, since if $u \in V_1(E_t)$ then the H distance to any node in L_1 is at least r (from Proposition 19) and the H distance to any node in L_2 is also at least r (from Fact 21). So, suppose $u \in V_1(E_c)$.

Let's follow the same proof strategy as before, so let $P \in \mathcal{P}(u)$ be the shortest path and let's follow P starting in u. Again, from Proposition 19 the H distance of u to any node in L_1 is at least r. However, u might reach $\phi(u)$ in a short number of steps (namely, if the cycle corresponding to $(u, \phi(u))$ is short). So, let's follow P to $\phi(u)$ and into T_2. We are now in the exact same situation as in the setup of the proof in Section 2.1 (but starting in T_2), so the result follows. ◀

Let u be some vertex in L_1. Let's say a vertex w is at $\mathcal{P}$-distance δ from u if the shortest path $P \in \mathcal{P}(u)$ that ends in w has length δ. Additionally, let $S_\delta(u)$ be the set of vertices that are at a $\mathcal{P}$-distance of at most δ from u. From Proposition 25, we know that for all distinct $u, w \in L_1$, the sets $S_{\lfloor r/2 \rfloor}(u)$ and $S_{\lfloor r/2 \rfloor}(w)$ are disjoint. Thus, we have that for $u \in L_1$ the vertices in $S_{\lfloor r/2 \rfloor}(u)$ form disjoint trees rooted at u, which shows that $B_l(v)$ forms a tree.

We can now apply Lemma 24 and conclude that for all $1 \leqslant i \leqslant l$, we have $\sum_{u \in X_i} g_u^2 \geqslant \sum_{u \in X_{i-1}} g_u^2$. So the sequence $(\sum_{u \in X_i} g_u^2)_i$ is an increasing sequence. Note that $X_h = L_1$, so $\sum_{u \in X_h} g_u^2 = \sum_{u \in L_1} g_u^2$. Additionally, we know that the total sum of $(\sum_{u \in X_i} g_u^2)_i$ is at most one (since g is a unit vector and the X_i are disjoint), so we have that $\lfloor r/2 \rfloor \cdot \sum_{u \in X_h} g_u^2 \leqslant \sum_{i=h}^{l} \sum_{u \in X_i} g_u^2 \leqslant 1$ and finally $\sum_{u \in L_1} g_u^2 = \sum_{u \in X_h} g_u^2 \leqslant 1/\lfloor r/2 \rfloor = O(1/r)$.

This concludes the proof of Theorem 8.

3 Open problems

- Can we improve Theorem 12 to obtain high girth?
  Something like this could be proved by showing that when 2-lifting a graph with large enough girth, with sufficiently high probability the girth of the resulting graph increases. This would boost the girth of the graph generated by the first step of the construction of [26] during the repeated 2-lift step. However, it is unclear if this can be done. Alternatively, one could show that bicycle-freeness increases with good probability as we 2-lift, but this is also unclear.
  A different strategy would be to find a different way to derandomize Theorem 4 such that the we can generate a starter graph of larger size. However, it is unclear if this strategy could work since the tool used to derandomize this, namely (δ, k)-wise uniform permutations (defined in Appendix C), cannot be improved to derandomize this to the required extent.
- Can we obtain Theorem 10 for higher values of c; for example, can we build a distribution that is $(2\sqrt{d-1}+\epsilon,\ .99\log_{d-1} n)$-good?
  One promising strategy would be to show that the graphs produced by the distribution described in [20], which were shown to have girth at least $.99\log_{d-1} n$ with high probability, are also near-Ramanujan with high probability. Numerical calculations seem to indicate that the answer is positive, as pointed out in one of the open problems given in [20].

References

1. Noga Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986.
2. Noga Alon. Explicit expanders of every degree and size. *arXiv preprint*, 2020. `arXiv:2003.11673`.
3. Noga Alon, Shirshendu Ganguly, and Nikhil Srivastava. High-girth near-ramanujan graphs with localized eigenvectors. *arXiv preprint*, 2019. `arXiv:1908.03694`.
4. 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.
5. Edward A. Bender and E. Rodney Canfield. The asymptotic number of labeled graphs with given degree sequences. *J. Combinatorial Theory Ser. A*, 24(3):296–307, 1978. `doi:10.1016/0097-3165(78)90059-6`.
6. Yonatan Bilu and Nathan Linial. Lifts, discrepancy and nearly optimal spectral gap. *Combinatorica*, 26(5):495–519, 2006. `doi:10.1007/s00493-006-0029-7`.
7. Béla Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.*, 1(4):311–316, 1980. `doi:10.1016/S0195-6698(80)80030-8`.
8. Charles Bordenave. A new proof of Friedman's second eigenvalue theorem and its extension to random lifts. Technical Report 1502.04482v4, arXiv, 2019. To appear in Annales scientifiques de l'École normale supérieure. `arXiv:1502.04482v4`.
9. Xavier Dahan. Regular graphs of large girth and arbitrary degree. *Combinatorica*, 34(4):407–426, 2014. `doi:10.1007/s00493-014-2897-6`.
10. Paul Erdős and Horst Sachs. Reguläre graphen gegebener tailenweite mit minimaler knollenzahl. *Wiss. Z. Univ. Halle-Willenberg Math. Nat.*, 12:251–258, 1963.
11. Joel Friedman. Some geometric aspects of graphs and their eigenfunctions. *Duke Mathematical Journal*, 69(3):487–525, 1993.
12. Joel Friedman. A proof of Alon's second eigenvalue conjecture and related problems. *Memoirs of the American Mathematical Society*, 195(910):viii+100, 2008.
13. R. G. Gallager. Low-density parity-check codes. *IRE Trans.*, IT-8:21–28, 1962. `doi:10.1109/tit.1962.1057683`.

14 Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *American Mathematical Society Bulletin*, 43(4):439–561, 2006.

15 Nabil Kahale. Eigenvalues and expansion of regular graphs. *Journal of the ACM (JACM)*, 42(5):1091–1106, 1995.

16 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.

17 Martin Kassabov. Symmetric groups and expander graphs. *Inventiones Mathematicae*, 170(2):327–354, 2007.

18 John Lafferty and Dan Rockmore. Codes and iterative decoding on algebraic expander graphs. In *the Proceedings of ISITA*. Citeseer, 2000.

19 Felix Lazebnik and Vasiliy A. Ustimenko. Explicit construction of graphs with an arbitrary large girth and of large size. *Discrete Appl. Math.*, 60(1-3):275–284, 1995. ARIDAM VI and VII (New Brunswick, NJ, 1991/1992). `doi:10.1016/0166-218X(94)00058-L`.

20 Nati Linial and Michael Simkin. A randomized construction of high girth regular graphs. *arXiv preprint*, 2019. `arXiv:1911.09640`.

21 A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988. `doi:10.1007/BF02126799`.

22 Mohammad M Mansour and Naresh R Shanbhag. Construction of ldpc codes from ramanujan graphs. In *36th Annu. Conf. on Information Sciences and Systems*, 2002.

23 G. A. Margulis. Explicit constructions of graphs without short cycles and low density codes. *Combinatorica*, 2(1):71–78, 1982. `doi:10.1007/BF02579283`.

24 G. A. 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.

25 Brendan D. McKay, Nicholas C. Wormald, and Beata Wysocka. Short cycles in random regular graphs. *Electron. J. Combin.*, 11(1):Research Paper 66, 12, 2004. URL: `http://www.combinatorics.org/Volume_11/Abstracts/v11i1r66.html`.

26 Sidhanth Mohanty, Ryan O'Donnell, and Pedro Paredes. Explicit near-ramanujan graphs of every degree. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 510–523, 2020.

27 Moshe Morgenstern. Existence and explicit constructions of $q+1$ regular Ramanujan graphs for every prime power q. *J. Combin. Theory Ser. B*, 62(1):44–62, 1994. `doi:10.1006/jctb.1994.1054`.

28 A. Nilli. On the second eigenvalue of a graph. *Discrete Mathematics*, 91(2):207–210, 1991.

29 Mark S Pinsker. On the complexity of a concentrator. In *7th International Telegraffic Conference*, volume 4, pages 1–318. Citeseer, 1973.

30 Joachim Rosenthal and Pascal O Vontobel. Constructions of ldpc codes using ramanujan graphs and ideas from margulis. In *in Proc. of the 38-th Allerton Conference on Communication, Control, and Computing*. Citeseer, 2000.

A A near-Ramanujan graph distribution of girth $\Omega(\log_{d-1} N)$

Recall Theorem 4, which says that uniformly random d-regular graphs are near-Ramanujan. We will combine this result with our machinery of Section 2 to show Theorem 10, namely that there exists a distribution over graphs that is $(2\sqrt{d-1}+\epsilon, c\log_{d-1} n)$-good for any $\epsilon > 0$ and $c < 1/4$, which we will show is the distribution resulting from applying algorithm fix to a sample of $\mathcal{G}_d(n)$.

First, we note that $\mathcal{G}_d$ has nice bicycle-freeness. We quote the relevant result from [8], which we restate below:

▶ **Lemma 26** ([8, Lemma 9]). *Let $d \geqslant 3$ and r be positive integers. Then $G \sim \mathcal{G}_d(n)$ is bicycle-free at radius r with probability $1 - O((d-1)^{4r}/n)$.*

An obvious corollary of this is that for any constant $c < 1/4$, we have that $G \sim \mathcal{G}_d(n)$ is bicycle free at radius $c \log_{d-1} n$ with high probability.

To bound the number of short cycles in $\mathcal{G}_d(n)$ we use a classic result that very accurately estimates the number of short cycles in random regular graphs.

▶ **Lemma 27** ([25, Section 2]). *Let $G \sim \mathcal{G}_d(n)$ and X_i be the random variable that denotes the number of cycles of length i in G. Let $R_i = \max\{(d-1)^i/i, \log n\}$. Then*

$$\boldsymbol{Pr}\left[X_i \leqslant R_i,\ \textit{for all}\ 3 \leqslant i \leqslant 1/4 \log_{d-1} n\right] = 1 - o_n(1).$$

Given the above, we obtain the following bound, for all $c < 1/4$:

$$\sum_{i=1}^{c \log_{d-1} n} \max\{(d-1)^i/i, \log n\} = O(n^c).$$

So we obtain the following proposition:

▶ **Proposition 28.** *For any $c < 1/4$ and any $\epsilon > 0$, $G \sim \mathcal{G}_d(n)$ is a $(c \log_{d-1} n, O(n^c))$-graph and satisfies $\lambda(G) \leqslant 2\sqrt{d-1} + \epsilon$ with probability $1 - o_n(1)$.*

Finally, we want to apply Theorem 8, so first we need to verify its preconditions. For all $c < 1/4$ we have that $(2/3)\log_{d-1}(n/n^c) = (2/3)(1-c)\log_{d-1} n \geqslant c \log_{d-1} n$. Also note that $n^c(d-1)^{c/2 \log_{d-1} n + 1} = n^{3c/2} = O(n^{3/8})$, so when applying Theorem 8 the resulting graph has $n + O(n^{3/8}) = n(1 + o_n(1))$ vertices. Thus, we obtain Theorem 10.

▶ Remark 29. Recall that $\mathcal{G}_d(n)$ is the same as the conditional distribution of the d-regular n-vertex configuration model when conditioned on it being a simple graph. Indeed, a graph drawn from the d-regular n-vertex configuration model is simple with probability $\Omega_d(1)$. A result very similar to Lemma 27 also holds for the configuration model and thus the results of this section also hold for the configuration model.

A.1 Counting near-Ramanujan graphs with high girth

We will briefly prove Corollary 11 using the result we just proved. For simplicity, we are going to work with the configuration model, using the observation of Remark 29.

Our proof will use a classic result on the number of not necessarily simple d-regular n-vertex graphs, which is the same as the number of graphs in the n-vertex d-regular configuration model. It is easy to show [5] that for nd even, the number of such graphs is

$$\sim \left(\left(\frac{d^d n^d}{e^d (d!)^2}\right)^{n/2}\right).$$

Hence, the core claim we need to prove, is the following:

▶ **Proposition 30.** *Let G_1 and G_2 be distinct graphs that follow the preconditions of Theorem 8. Then $\operatorname{fix}(G_1)$ and $\operatorname{fix}(G_2)$ are also distinct.*

This proposition implies that given any two good d-regular n-vertex graphs, applying fix produces two distinct graphs. From our proof of Theorem 10 we also know that the result of applying fix adds at most $O(n^{3/8})$ vertices. Finally, since a $(1 - o_n(1))$ fraction of the graphs are good an thus when we apply fix they result in $(2\sqrt{d-1} + \epsilon,\ c \log_{d-1} n)$-good graphs, the result follows. For briefness, we will not give a detailed proof but only a sketch of the proof.

Proof sketch of Proposition 30. Recall the H graph from the description of fix and let H_1 be such graph corresponding to G_1 and define H_2 analogously. If H_1 and H_2 are distinct, then $\text{fix}(G_1)$ and $\text{fix}(G_2)$ are also distinct. This follows from the fact that the vertices of the two added trees will have to be matched up in an isomorphism between $\text{fix}(G_1)$ and $\text{fix}(G_2)$.

We claim that if G_1 and G_2 are distinct, then H_1 and H_2 are distinct. Let S_i be the set of vertices that were endpoints of edges removed from cycles in G_1 and G_2, respectively. Note that there are at least two such vertices in S_i and also we cannot remove multiple edges adjacent to one vertex since this would imply the existence of two cycles in a small neighborhood, breaking the bicycle-freeness assumption. We can ignore the other removed edges since the local neighborhoods of edges removed from cycles are necessarily distinct from the local neighborhoods of the other removed edges. Now, the edges removed from G_i form a perfect matching on S_i that adds exactly $|S_i|/2$ cycles to H_i. Also, there is exactly one perfect matching that adds $|S_i|/2$ cycles to H_i to recover G_i. That means that there is only one G_i that could have generated H_i, which implies the claim. ◀

B Explicit near-Ramanujan graphs of girth $\Omega(\sqrt{\log n})$

In this section we prove Theorem 12, building on the construction in the proof of Theorem 5. We note that the original construction has no guarantees on the girth of the constructed graph other than a constant girth. We will briefly recap the main tools and ideas from the paper.

B.1 Review of constructing explicit near-Ramanujan graphs

Given a d-regular n-vertex graph $G = (V, E)$, let $w \in \{\pm 1\}^E$ be an edge-signing of G. The *2-lift of G given w* is defined as the following d-regular $2n$-vertex graph $G_2 = (V_2, E_2)$:

$$V_2 = V \times \{\pm 1\} \qquad E_2 = \{\{(u, \sigma), (v, \sigma \cdot w(u, v))\} : (u, v) \in E, \sigma \in \{\pm 1\}\}.$$

It was observed in [6] that the spectrum of G_2 is given by the union of the spectra of G and $\widetilde{G}_w$, where the latter refers to the eigenvalues of the adjacency matrix of G signed according to w, where each nonzero entry is $w(u, v)$ for $\{u, v\} \in E$.

This connection between the spectrum of an edge-signing of a graph and a 2-lift gave rise to the following theorem, which was proved in [26]. Below we write $\rho(G) = \max\{|\lambda_i| : i \in [n]\}$ for the *spectral radius* of G.

▶ **Theorem 31** ([26, Theorem 3.1]). *Let $G = (V, E)$ be an arbitrary d-regular n-vertex graph ($d \geqslant 3$). Assume G is bicycle-free at radius $r \gg (\log \log n)^2$. Then for a uniformly random edge-signing w, except with probability at most n^{-100} we have:*

$$\rho(\widetilde{G}_w) \leqslant 2\sqrt{d-1} \cdot \left(1 + \frac{(\log \log n)^4}{r^2}\right).$$

Furthermore, this can be derandomized: given a constant C there is a generator $h : \{0,1\}^s \to \{\pm 1\}^E$ computable in time $\text{poly}(N^{C \log d})$, with seed length $s = O(\log(2C) + \log \log n + C \cdot \log(d) \cdot \log(n))$, such that for $u \in \{0,1\}^s$ chosen uniformly at random, with probability at most n^{-100} we have:

$$\rho(\widetilde{G}_{h(u)}) \leqslant 2\sqrt{d-1} \cdot \left(1 + \frac{(\log \log n)^4}{r^2}\right) + \frac{\sqrt{d}}{C^2}.$$

This theorem is a powerful tool that, combined with the above observation, allows one to double the number of vertices in a near-Ramanujan graph while keeping it near-Ramanujan, as long as the bicycle-freeness is good enough. It is easy to show that if G is bicycle-free at radius r, then *any* 2-lift of G is also bicycle-free at radius r. So, the strategy employed by [26] is to start with a graph with a smaller number of vertices that is bicycle-free at a big enough radius and 2-lift it enough times until the graph has the required number of vertices.

To generate this starting graph, the authors first showed how out to weakly derandomize [8]. Formally, the following is proved:

▶ **Theorem 32** ([26, Theorem 4.8]). *For a large enough universal constant α and any integer $n > 0$, given d, ϵ and c such that:*

$$3 \leqslant d \leqslant \alpha^{-1}\sqrt{\log n},\ \alpha^3 \cdot \left(\frac{\log\log n}{\log_{d-1} n}\right)^2 \leqslant \epsilon \leqslant 1,\ c < 1/4.$$

Let G be chosen from the d-regular n-vertex uniform configuration model. Then, except with probability at most $n^{-.99}$, the following hold:

- *G is bicycle-free at radius $c\log_{d-1} n$;*
- *$\lambda(G) \leqslant 2\sqrt{d-1} \cdot (1+\epsilon)$;*

Furthermore, this can be derandomized: there is a generator $h : \{0,1\}^s \to \mathcal{G}_d(n)$, with seed length $s = O(\log^2(n)/\sqrt{\epsilon})$ computable in time $\mathrm{poly}(n^{\log(n)/\sqrt{\epsilon}})$, such that for $u \in \{0,1\}^s$ chosen uniformly at random, with probability at most $n^{-.99}$ we have that the above statements remain true for $G = h(u)$.

Using these two theorems we can setup the construction of [26]. So, first assume we are given $n, d \geqslant 3$ and $\epsilon > 0$ and we wish to construct a d-regular graph G with n vertices with $\lambda(G) \leqslant 2\sqrt{d-1} + \epsilon$. The construction is now the following:

1. Use Theorem 32 to construct a d-regular graph G_0 with a small number of vertices $n_0 = n_0(n)$. If we pick n_0 to be $2^{O(\sqrt{\log n})}$ then the generator seed length is $O(\log(n)/\sqrt{\epsilon})$ and is computable in time $\mathrm{poly}(n^{1/\sqrt{\epsilon}})$, so we can enumerate over all possible seeds and find at least one that produces a graph that is bicycle-free at radius $\Omega(\log(n_0)) = \Omega(\sqrt{\log n}) \gg (\log\log n)^2$ and has $\lambda(G_0) \leqslant 2\sqrt{d-1} \cdot (1+\epsilon)$ in $\mathrm{poly}(n)$ time.
2. Next, we can repeatedly apply Theorem 31 to double the number of vertices of G_0, by choosing C to be $\sim d^{1/4}/\sqrt{\epsilon}$. We then enumerate over all seeds until we find one that produces a good graph, which only requires $\mathrm{poly}(n)$ time. On each application the bicycle-freeness radius is maintained (so we can keep applying Theorem 31) and the number of vertices of doubles. After roughly $\log(n/n_0)$ applications, the resulting graph has $n(1 + o_n(1))$ vertices and $\lambda(G) \leqslant 2\sqrt{d-1} \cdot (1+\epsilon)$.

B.2 Improving the girth of the construction

We are finally ready to prove Theorem 12. We are going to apply a similar strategy as the one from Appendix A. Instead of derandomizing Lemma 27 we are going to obtain a simpler bound, which is good enough to obtain the desired. We note however, that Lemma 27 can be derandomized and for completeness we show how to in Appendix C.

We start by proving the following lemma:

▶ **Lemma 33.** *Let G be a d-regular n-vertex graph with $\lambda(G) \geqslant 2\sqrt{d-1}$ and such that G is bicycle-free at radius $\alpha\log_{d-1} n$, for $\alpha \leqslant 2$. Then we can apply* fix *to G and obtain a graph such that:*

- $\mathrm{fix}(G)$ *is d-regular and has $n(1+o_n(1))$ vertices;*
- $\lambda(\mathrm{fix}(G)) \leqslant \lambda(G) + o_n(1)$*;*
- $\mathrm{fix}(G)$ *has girth $(\alpha/3)\log_{d-1} n$.*

Before proving this lemma, we prove a core proposition in a slightly more generic way.

▶ **Proposition 34.** *Let G be a d-regular graph that is bicycle-free at radius $2r$, then*

$$|\mathrm{Cyc}_r(G)| \leqslant n/(d-1)^r.$$

Proof. Pick one vertex per cycle in $\mathrm{Cyc}_r(G)$ and place it in a set S. We claim that for every distinct $u, v \in S$, $B_r(u) \cap B_r(v) = \emptyset$. Suppose this wasn't the case and suppose there is some w such that $w \in B_r(u) \cap B_r(v)$, for some pair u, v. Note that $B_{2r}(w)$ includes the two length r cycles that correspond to u and v, which contradicts bicycle-freeness in G.

Given the above, we have that the sets $B_r(u)$ for $u \in S$ are pairwise disjoint and also we know that $|B_r(u)| = d(d-1)^{r-1}$. Hence we have:

$$|\mathrm{Cyc}_r(G)| \cdot d(d-1)^{r-1} \leqslant n,$$

which implies the desired result. ◀

And we can prove the above lemma.

Proof of Lemma 33. By plugging G into Proposition 34 we can conclude that G is a $(\alpha \log_{d-1} n, n^{1-\alpha/2})$-graph. We wish to apply Theorem 8 so first recall its preconditions. By definition $\lambda(G) \geqslant 2\sqrt{d-1}$. However, the precondition on the radius of bicycle-freeness does not hold, since $(2/3)\log_{d-1}(n/n^{1-\alpha/2}) = (\alpha/3)\log_{d-1} n$ which is less than $\alpha \log_{d-1} n$. If we instead use the fact that G is also trivially a $((\alpha/3)\log_{d-1} n, n^{1-\alpha/2})$-graph, then the precondition is satisfied.

Thus, we can apply Theorem 8 and we obtain that $\mathrm{fix}(G)$ satisfies all the required conditions, which concludes the proof. ◀

Given this lemma, we will modify the first step of the construction of [26] to produce a graph G_0 with girth $c\sqrt{\log n}$. Note that, similarly to bicycle-freeness, the girth of a graph can only increase when applying any 2-lift, so this strategy guarantees that after step 2 of the construction, the final graph has the desired girth, which would imply Theorem 12.

First, when enumerating over all seeds to generate G_0 in step 1, we look for one that guarantees that G_0 is bicycle-free at radius $(1/5)\log_{d-1} n_0$ (recall that by Theorem 32 a $1 - o_n(1)$ fraction of the seeds satisfy this). Next, we apply Lemma 33 and obtain $\mathrm{fix}(G_0)$ with girth $(1/15)\log_{d-1} n_0$ and the desired value of $\lambda(G_0)$. Let $\kappa = 15c/\log_{d-1} 2$. We can set n_0 to $2^{\kappa\sqrt{\log n}}$, in which case G_0 has girth $c\sqrt{\log n}$.

Note that the above only works as long as $\kappa \leqslant \sqrt{\log n}$, otherwise $n_0 > n$. Also, from Theorem 31 and Theorem 32, we need $d \leqslant (\log n)^{1/8}/C$ and $\epsilon \gg \sqrt{d}(\log\log n)^4/(\log n)$ (the details on how to obtain these can be found on [26]).

Finally, we can precisely determine the running time of this algorithm. From Theorem 32, constructing G_0 takes time $\mathrm{poly}(n_0^{\log(n_0)/\sqrt{\epsilon}}) = \mathrm{poly}(n^{\log(c/\log_{d-1}(2))/\sqrt{\epsilon}})$ and using Theorem 31 with the appropriate choice of C takes time $\mathrm{poly}(n^{d^{1/4}\log(d)/\sqrt{\epsilon}})$.

C Derandomizing the number of short cycles

To make the statement of this section more precise, we will first define a known derandomization tool.

▶ **Definition 35** ((δ, k)-wise uniform permutations). *Let $\delta \in [0,1]$ and $k \in \mathbb{N}^+$. Let $[n]_k$ denote the set of all sequences of k distinct indices from $[n]$. A random permutation $\pi \in S_n$ is said to be (δ, k)-wise uniform if, for every sequence $(i_1, \ldots, i_k) \in [n]_k$, the distribution of $(\pi(i_1), \ldots, \pi(i_k))$ is δ-close in total variation distance from the uniform distribution on $[n]_k$. When $\delta = 0$, we simply say that the permutation is (truly) k-wise uniform.*

Kassabov [17] and Kaplan–Naor–Reingold [16] independently obtained a deterministic construction of (δ, k)-wise uniform permutations with seed length $O(k \log n + \log(1/\delta))$.

▶ **Theorem 36** ([16, 17]). *There is a deterministic algorithm that, given δ, k, and n, runs in time $\mathrm{poly}(n^k/\delta)$ and outputs a multiset $\Pi \subseteq S_n$ (closed under inverses) of cardinality $S = \mathrm{poly}(n^k/\delta)$ (a power of 2) such that, for $\pi \sim \Pi$ chosen uniformly at random, π is a (δ, k)-wise uniform permutation.*

This theorem is required to obtain the generator mentioned in Theorem 32 and is the reason why (δ, k)-wise uniform permutations are useful tools to apply here. We will also need a convenient theorem of Alon and Lovett [4]:

▶ **Theorem 37** ([4]). *Let $\boldsymbol{\pi} \in S_n$ be a (δ, k)-wise uniform permutation. Then one can define a (truly) k-wise uniform permutation $\boldsymbol{\pi}' \in S_n$ such that the total variation distance between $\boldsymbol{\pi}$ and $\boldsymbol{\pi}'$ is $O(\delta n^{4k})$.*

We can now define a "derandomized" version of the configuration model, using this tool.

▶ **Definition 38.** *Recall how the configuration model is defined by a perfect matching of a set $[nd]$ of "half-edges".*

Let's denote this matching by M and define a way to generate it using random permutations. First a uniformly random permutation $\pi \in S_{nd}$ is chosen; then we set $M_{\pi(j),\pi(j+1)} = M_{\pi(j+1),\pi(j)} = 1$ for each odd $j \in [nd]$.

We can write the adjacency matrix A of G as the sum, over all $i, i' \in [d]$, of $M_{(v,i),(v',i')}$. Hence

$$\boldsymbol{A}_{v,v'} = \sum_{i,i'=1}^{d} \sum_{\substack{odd\\ j \in [nd]}} (1[\pi(j) = (v,i)] \cdot 1[\pi(j+1) = (v',i')] + 1[\pi(j) = (v',i')] \cdot 1[\pi(j+1) = (v,i)]).$$

The d-regular n-vertex (δ, k)-wise uniform configuration model is defined by using (δ, k)-wise uniform permutations instead. Similarly, we define the d-regular n-vertex k-wise uniform configuration model.

We can now describe the proposition we wish to prove.

▶ **Proposition 39.** *Fix $d \geqslant 3$, n and $k \geqslant c \log_{d-1} n$, where $c < 1/4$. Let G be drawn from the d-regular n-vertex $4k$-wise configuration model and X_i be the random variable that denotes the number of cycles of length i in G. Let $R_i = \max\{(d-1)^i/i, \log n\}$. Then*

$$\boldsymbol{Pr}\left[X_i \leqslant R_i, \text{ for all } 1 \leqslant i \leqslant 1/4 \log_{d-1} n\right] = 1 - o_n(1).$$

By Theorem 37, these statements remain true in the $(\delta, 4k)$-wise uniform versions of the model, $\delta \leqslant 1/n^{16k+1}$.

Proof. The proof follows almost directly from the proof of Lemma 27. First, note that X_i can be written as a polynomial of degree at most i in the entries of G's adjacency matrix, by summing over the products of the edge indicators of all possible cycles of length i in G. Thus, from our formula in Definition 38, it can be written as a polynomial of degree at most $2k$ in the permutation indicators $1[\pi(j) = (v,i)]$. So we can compute $\mathbf{E}[X_i]$ assuming that X_i is drawn from the fully uniform configuration model. Similarly, X_i^2 can be written as a polynomial of degree at most $4k$ in the permutation indicators, so we can compute $\mathbf{Var}[X_i]$ assuming that X_i is drawn from the fully uniform configuration model.

From [25] we have the following estimates, that only apply when $(d-1)^{2i-1} = o(n)$:

$$\mathbf{E}[X_i] = \frac{(d-1)^i}{2i}(1 + O(i(i+d)/n)) \qquad \mathbf{Var}[X_i] = \mathbf{E}[X_i] + O(i(i+d)/n)\mathbf{E}[X_i]^2 .$$

By applying Chebyshev's inequality to each X_i, just like in [25], we get the desired result. ◀

We can finally rewrite Theorem 32 in the language of the d-regular n-vertex (δ, k)-wise uniform configuration model and tack on the result we just proved.

▶ **Theorem 40.** *For a large enough universal constant α and any integer $n > 0$, fix $3 \leqslant d \leqslant \alpha^{-1}\sqrt{\log n}$ and $c < 1/4$, and let $\varepsilon \leqslant 1$ and k satisfy*

$$\varepsilon \geqslant \alpha^3 \cdot \left(\frac{\log\log n}{\log_{d-1} n}\right)^2, \qquad k \geqslant \alpha \log(n)/\sqrt{\varepsilon}.$$

Let G be chosen from the d-regular n-vertex k-wise uniform configuration model. Then except with probability at most $1/n^{.99}$, the following hold:

- *G is bicycle-free at radius $c \log_{d-1} n$;*
- *The total number of cycles of length at most $c \log_{d-1} n$ is $O(n^c)$;*
- *$\lambda(G) \leqslant 2\sqrt{d-1} \cdot (1+\varepsilon)$.*

Finally, by Theorem 37, these statements remains true in the (δ, k)-wise uniform configuration model, $\delta \leqslant 1/n^{16k+1}$.