

ARTICLE

On directed analogues of expander and hyperfinite graph sequences

Endre Csóka^{1,†} and Łukasz Grabowski^{2,‡,*}

¹Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Budapest, Hungary and ²Department of Mathematics and Statistics, Lancaster University, Lancaster, UK

*Corresponding author. Email: lukasz.grabowski@lancaster.ac.uk

(Received 20 August 2020; revised 6 April 2021; accepted 26 April 2021; first published online 6 July 2021)

Abstract

We introduce and study analogues of expander and hyperfinite graph sequences in the context of directed acyclic graphs, which we call ‘extender’ and ‘hypershallow’ graph sequences, respectively. Our main result is a probabilistic construction of non-hypershallow graph sequences.

Keywords: graph limits, expander, extender, hyperfinite

2020 MSC Codes: Primary 05C20; Secondary 94C15

1. Introduction

Hyperfinite and expander graph sequences are perhaps the two most fundamental concepts studied in the theory of sparse graph limits. Hyperfinite graph sequences were explicitly introduced in [2] (and implicitly they are present in earlier works, e.g. in [5]). Expander graph sequences (frequently informally referred to as ‘expander graphs’) have been studied since at least the 70s in many different branches of mathematics and computer science (see [4] for a survey with some historical information). Both notions (or their close relatives) are broadly used in combinatorics, group theory, ergodic theory, and operator algebras.

In this article, we study the analogues of hyperfinite and expander graph sequences in the context of oriented graphs, particularly directed acyclic graphs. We call these analogues ‘hypershallow’ and ‘extender’ graph sequences, respectively. Our main result (see Theorem 5 below) is a stochastic construction of graph sequences which are not hypershallow (we do not know any deterministic construction of such graph sequences). The question whether non-hypershallow graph sequences exist was partially motivated by the techniques presented in [1] and in [5] for obtaining conditional lower bounds in circuit complexity. We will discuss this in Section 4.

Let us now precisely define hypershallow graph sequences and state our main result.

Basic conventions. The set of natural numbers is $\mathbb{N} := \{0, 1, 2, \dots\}$. We use the shorthand (X_n) for denoting a sequence $(X_n)_{n=0}^\infty$.

A graph is a pair $G = (V, E)$ where V is a non-empty finite set, and $E \subset V \times V$ is a subset which is disjoint from the diagonal. We say that G is *undirected* if E is a symmetric subset of $V \times V$. A

[†]E.Cs. was supported by the ERC Synergy grant No. 810115.

[‡]L.G. was supported by the ERC under the European Union’s Horizon 2020 research and innovation programme (Grant agreement No. 805495).

path of length D in a graph $G = (V, E)$ is a tuple $(x_0, \dots, x_D) \in V^{D+1}$, such that for $i < D$ we have either $(x_i, x_{i+1}) \in E$ or $(x_{i+1}, x_i) \in E$.

A path (x_0, \dots, x_D) is *simple* if $x_i \neq x_j$ for $i \neq j$. It is a *directed* path if for all $i < D$ we have $(x_i, x_{i+1}) \in E(G)$. A *cycle* is a path (x_0, \dots, x_D) such that $x_0 = x_D$. We say that G is a *dag* (which stands for *directed acyclic graph*) if it does not have directed cycles.

We are now ready to define *hypershallow* graph sequences.

- Definition 1.**
1. Let G be a graph and let $S \subsetneq V(G)$ be a proper subset. We define $(S;G) \in \mathbb{N}$ as the maximal D such that there exists a directed simple path (x_0, \dots, x_D) in G disjoint from S .
 2. Let (G_n) be a sequence of dags with uniformly bounded in-degrees. We say that (G_n) is *hypershallow* if $\forall \varepsilon > 0, \exists D \in \mathbb{N}, \exists (S_n)$ with $S_n \subsetneq V(G_n)$ and $|S_n| < \varepsilon |V(G_n)|$, such that $\forall n \in \mathbb{N}$ we have $(S_n;G_n) \leq D$.

Remark 2. Let us take a moment to explicitly state the analogy between the definitions of hypershallow and hyperfinite graph sequences.

1. We first recall the definition of hyperfinite graph sequences. If G is an undirected graph and $S \subsetneq V(G)$, then we note that $\text{codepth}(S;G)$ is the maximum of lengths of simple paths disjoint from S .

We define a sequence (G_n) of undirected graphs with uniformly bounded degrees to be *hyperfinite* if $\forall \varepsilon > 0, \exists D \in \mathbb{N}, \exists (S_n)$ with $S_n \subsetneq V(G_n)$ and $|S_n| < \varepsilon |V(G_n)|$, such that $\forall n \in \mathbb{N}$ we have $(S_n;G_n) \leq D$.

This is easily seen to be equivalent to the definition of hyperfiniteness in [2].

From this point of view, and with our convention that undirected graphs form a subclass of all graphs, within the class of bounded degree undirected graphs the hypershallow sequences are exactly the same as hyperfinite sequences.

2. Let us explain the choice of the word ‘hypershallow’, again by analogy with the word ‘hyperfinite’. One of the simplest classes of undirected graph sequences consists of those sequences (G_n) which have uniformly finite connected components, that is, $\exists D$ such that $\forall n$ we have that the connected components of G_n are of size at most D . We recall that the expression ‘hyperfinite graph sequence’ is meant to suggest that we are dealing with ‘the next simplest thing’: informally, a sequence (G_n) is hyperfinite if it is possible to obtain from (G_n) a sequence with uniformly finite connected components by removing an arbitrarily small proportion of vertices from (G_n) .

The motivation to use the word ‘hypershallow’ is similar. For a dag G , let $\text{depth}(G)$ denote the maximum of lengths of directed paths in G . One of the simplest classes of dag sequences with uniformly bounded in-degrees consists of the ‘uniformly shallow’ sequences, that is, $\exists D$ such that $\forall n$ we have $(G_n) \leq D$. The name ‘hypershallow graph sequence’ is meant to suggest that we are dealing with ‘the next simplest thing’: after removing a small proportion of vertices, we get a sequence which is uniformly shallow.^a

The following definition allows us, informally speaking, to capture ‘how badly’ a sequence of graphs fails at being hypershallow.

- Definition 3.**
1. Let G be a dag, let $\varepsilon, \rho > 0$. We say that G is an (ε, ρ) -extender if for every $S \subsetneq V(G)$ with $|S| \leq \varepsilon |V(G)|$ we have $(S;G) \geq \rho$.

^aUniformly shallow sequences are much more frequently called ‘bounded depth sequences’. However, the authors think that ‘hypershallow’ sounds much better than ‘hyper-bounded-depth’.

- Let (G_n) be a sequence of dags with uniformly bounded in-degrees, and let (ρ_n) be a sequence of positive real numbers with $\lim_{n \rightarrow \infty} \rho_n = \infty$. We say that (G_n) is a (ρ_n) -extender sequence if $\lim_{n \rightarrow \infty} |V(G_n)| = \infty$ and $\exists \varepsilon > 0, \exists C > 0, \forall n \in \mathbb{N}: G_n$ is an $(\varepsilon, C\rho_{|V(G_n)|})$ -extender.

Remark 4. It is easy to check that a sequence (G_n) of dags with uniformly bounded in-degrees is not hypershallow if and only if it contains a subsequence which is a (ρ_n) -extender for some (ρ_n) with $\lim_{i \rightarrow \infty} \rho_n = \infty$.

We are now ready to state our main theorem.

Theorem 5. *There exists a sequence of directed acyclic graphs with uniformly bounded degrees which is an (n^δ) -extender sequence, with $\delta \approx 0.019$.*

Our proof of this theorem is probabilistic. The most important part of the proof consists of studying the random graphs G_n^d which will be introduced in Section 3. We do not know of a non-probabilistic way of constructing a non-hypershallow sequence of dags with uniformly bounded degrees.

On the other hand, we can ask how fast the sequence (ρ_n) can grow, provided that there exists a (ρ_n) -extender sequence. In this direction, we have the following result.

Theorem 6. *Let (δ_n) be a sequence of numbers in $[0, 1]$ such that $\lim_{n \rightarrow \infty} \delta_n = 1$. If (G_n) is a sequence of directed acyclic graphs with uniformly bounded in-degrees, then (G_n) is not an (n^{δ_n}) -extender sequence.*

Remark 7. Theorem 6 implies, for example, that there are no $(\frac{n}{\log(n)})$ -extender sequences. However, we do not know whether there exists an (n^δ) -extender sequence for every $\delta < 1$.

In Section 2, we list some standard definitions and conventions, and we discuss a variant of Pinsker’s inequality which involves the Shannon entropy (Proposition 12). Pinsker’s inequality is the most important external result in our analysis of the random graphs G_n^d .

The main part of this article is Section 3 where we introduce the random graphs G_n^d and use them to prove Theorem 5.

We conclude this article with Section 4, where we present the proof of Theorem 6, and we discuss our initial motivations for studying hypershallow and extender graph sequences which are related to the theory of boolean circuits.

2. Preliminaries

We use the following conventions. If $n \in \mathbb{N}$, then $n = \{0, 1, \dots, n - 1\}$. If X is a set, then $\text{Pow}(X)$ denotes the power set of X , that is, the set of all subsets of X .

2.1 Graphs

Definition 8. Let $G = (V, E)$ be a graph, and let $v, w \in V$.

- $\text{In}(v;G) := \{x \in V : (x, v) \in G\}$, $\text{indeg}(v;G) := |\text{In}(v;G)|$,
- $\text{Out}(v;G) := \{x \in V : (v, x) \in G\}$, $\text{outdeg}(v;G) := |\text{Out}(v;G)|$,
- $\text{deg}(v;G) := \text{indeg}(v;G) + \text{outdeg}(v;G)$,
- $\text{In}(G) := \{v \in V : \text{indeg}(v;G) = 0\}$, $\text{Out}(G) := \{v \in V : \text{outdeg}(v;G) = 0\}$,

- 5. $\text{maxindeg}(G) := \max_{v \in V} \text{indeg}(v; G)$, $\text{maxoutdeg}(G) := \max_{v \in V} \text{outdeg}(v; G)$,
 $\text{maxdeg}(G) := \max_{v \in V} \text{deg}(v; G)$,

Definition 9. Let (G_n) be a sequence of graphs. We say that (G_n) has, respectively, *bounded degree*, *bounded in-degree*, or *bounded out-degree*, if, respectively, $\max_{n \in \mathbb{N}} \text{maxdeg}(G_n) < \infty$, $\max_{n \in \mathbb{N}} \text{maxindeg}(G_n) < \infty$, or $\max_{n \in \mathbb{N}} \text{maxoutdeg}(G_n) < \infty$.

2.2 Probability

Definition 10. 1. If μ is a probability measure on \mathbb{N} , then we also use the symbol μ for the function $\mathbb{N} \rightarrow \mathbb{R}$ which sends $k \in \mathbb{N}$ to $\mu(\{k\})$ (so in particular we can write $\mu(k)$ instead of $\mu(\{k\})$), and we let

$$H(\mu) := - \sum_{i \in \mathbb{N}} \mu(i) \log(\mu(i)),$$

where by convention $0 \log(0) = 0$.

- 2. A *random variable* on a standard probability space (X, μ) with values in a standard Borel space Y is a Borel function $f: X \rightarrow Y$. The *law* of f is the push-forward measure $f^*(\mu)$ on Y , that is, for $U \subset Y$ we let $f^*(\mu)(U) := \mu(f^{-1}(U))$.
- 3. If f is an \mathbb{N} -valued random variable and α is its law, then we define $H(f) := H(\alpha)$.
- 4. If f and g are random variables with values in a standard Borel space Z , then we define a new random variable $f \sqcup g$ with values in Z by, informally speaking, choosing between f and g with probability $\frac{1}{2}$.
 Formally, suppose that f and g are defined on (X, μ) and (Y, ν) , respectively. The probability space on which $f \sqcup g$ is defined is $(X \sqcup Y, \omega)$, where ω is the unique measure on $X \sqcup Y$ such that $\omega(U) = \frac{\mu(U)}{2}$ when $U \subset X$ and $\omega(U) = \frac{\nu(U)}{2}$ when $U \subset Y$. We let $f \sqcup g(x) := f(x)$ for $x \in X \subset X \sqcup Y$ and $f \sqcup g(y) := g(y)$ for $y \in Y \subset X \sqcup Y$.
- 5. For $\alpha: \mathbb{N} \rightarrow \mathbb{R}$, we let $\|\alpha\|_1 := \sum_{i \in \mathbb{N}} |\alpha(i)|$.

Lemma 11. a. If f and g are random variables with values in the same space Y , with laws α and β , respectively, then the law of $f \sqcup g$ is $\frac{\alpha + \beta}{2}$.
 b. If f is a random variable with values in $\{0, 1, \dots, k - 1\}$ then $H(f) \leq \log(k)$.

Proof. a. Follows directly from the definitions.
 b. See e.g. [3, Section 2.3] □

The main point of the following proposition is contained in its second item. Informally, it allows us to say the following: if f and g are \mathbb{N} -valued random variables with laws α and β , respectively, and $(x, y) \in \mathbb{N}^2$ is chosen according to the law $\alpha \times \beta$, then either it is roughly as probable that $x > y$ as it is that $y > x$, or the entropy of $f \sqcup g$ is substantially larger than the average of the entropies of f and g .

Proposition 12. Let f and g be \mathbb{N} -valued random variables with laws α and β , respectively.

- 1. We have $2H(f \sqcup g) - H(f) - H(g) \geq 0$.
- 2. We have

$$\alpha \times \beta (\{(x, y) \in \mathbb{N}^2 : x > y\}) \leq \frac{1}{2} + 2\sqrt{2H(f \sqcup g) - H(f) - H(g)}. \tag{1}$$

Proof. By the previous lemma, we have that the law of $f \sqcup g$ is $\gamma := \frac{\alpha + \beta}{2}$. As such the first item follows from Jensen’s inequality.

The second item is a simple corollary of Pinsker’s inequality (see e.g. [6, Theorem 2.16] for the statement and proof of Pinsker’s inequality). To derive it, we start by stating the following two special cases of Pinsker’s inequality:

$$\|\alpha - \gamma\|_1^2 \leq 2D(\alpha \parallel \gamma)$$

and

$$\|\beta - \gamma\|_1^2 \leq 2D(\beta \parallel \gamma),$$

where

$$D(\alpha \parallel \gamma) := \sum_{i \in \mathbb{N}} \alpha(i) \log \left(\frac{\alpha(i)}{\gamma(i)} \right),$$

and similarly for $D(\beta \parallel \gamma)$. By convention, we set $0 \log(0) = 0 \log\left(\frac{0}{0}\right) = 0$ in the definitions of $D(\alpha \parallel \gamma)$ and $D(\beta \parallel \gamma)$.

Noting that $\|\alpha - \gamma\|_1 = \|\beta - \gamma\|_1$, summing the two inequalities above gives

$$2\|\beta - \gamma\|_1^2 = 2\|\alpha - \gamma\|_1^2 \leq 2D(\alpha \parallel \gamma) + 2D(\beta \parallel \gamma).$$

A direct computation shows that $D(\alpha \parallel \gamma) + D(\beta \parallel \gamma) = 2H(\gamma) - H(\alpha) - H(\beta)$, so together with the triangle inequality we deduce that

$$\|\alpha - \beta\|_1 \leq 2\sqrt{2H(\gamma) - H(\alpha) - H(\beta)}. \tag{2}$$

On the other hand, the left-hand side of (1) is equal to

$$\begin{aligned} \sum_{i,j \in \mathbb{N}: i < j} \alpha(i)\beta(j) &\leq \sum_{i,j \in \mathbb{N}: i < j} \alpha(i)(\alpha(j) + \|\alpha - \beta\|_1) \\ &\leq \alpha \times \alpha (\{(x, y) \in \mathbb{N}^2 : x > y\}) + \|\alpha - \beta\|_1 \\ &\leq \frac{1}{2} + \|\alpha - \beta\|_1 \end{aligned}$$

which together with (2) finishes the proof. □

3. Existence of non-hypershallow sequences

In this section, we will describe a probabilistic construction of non-hypershallow sequences of dags. They will be in fact n^δ -expander sequences for $\delta \approx 0.019$.

We will construct a sequence of random graphs \mathbf{G}_n^d which asymptotically almost surely forms, after small modifications, an n^δ -extender sequence. The graphs \mathbf{G}_n^d will be essentially defined as follows. The vertices are $\{1, 2, \dots, n\}$ and for every $i < j$, we add an edge (i, j) independently with probability proportional to $\frac{1}{j-i}$. In order to simplify the proof, we will slightly change the probabilities when we define \mathbf{G}_n^d in Subsection 3.2.

We start with the definition and discussion of *depth functions* in Subsection 3.1, as they provide a convenient way of characterising the property of being an (ϵ, ρ) -extender, which will be crucial in the analysis of the random graphs \mathbf{G}_n^d in Subsection 3.3.

3.1 Depth functions

Given a graph G and $S \subset V(G)$, we can associate to it a function which ‘measures the maximal distance to $S \cup \text{Out}(G)$ ’. More precisely, we define $\delta_S : V(G) \rightarrow \mathbb{N}$ by setting $\delta_S(x) = 0$ when $x \in$

$S \cup \text{Out}(G)$, and for $x \notin S \cup \text{Out}(G)$ we let $\delta_S(x)$ to be the maximal $l \in \mathbb{N}$ for which there exists a directed simple path x_0, \dots, x_l with $x_0 = x, x_l \in S \cup \text{Out}(G)$, and $x_i \notin S$ when $0 \leq i < l$. Let us start by abstracting some properties of δ_S into the notion of a *depth function* as follows.

Definition 13. Let G be a graph.

1. A *depth function* for G is a function $f: V(G) \rightarrow \mathbb{N}$ such that the following conditions hold
 - a. For every $(a, b) \in E(G)$, we have either $f(a) > f(b)$ or $f(a) = 0$
 - b. For every $a \in V(G)$ such that $f(a) \neq 0$, there exists $b \in V(G)$ such that $(a, b) \in E(G)$ and $f(b) = f(a) - 1$.
2. Let $\varepsilon > 0$ and let $\rho \in \mathbb{N}$. An (ε, ρ) -*depth function* for G is a depth function f for G such that for all $v \in V(G)$ we have $f(v) \leq \rho$ and $|f^{-1}(0) \setminus \text{Out}(G)| \leq \varepsilon|V(G)|$.

Example 14. It is straightforward to verify that if $S \subset V(G)$ then δ_S is a

$$\left(\frac{|S \setminus \text{Out}(G)|}{|V(G)|}, \text{codepth}(S;G) + 1 \right)$$

-depth function.

Lemma 15. *If f is a (ε, ρ) -depth function, then $\text{codepth}(f^{-1}(0) \setminus \text{Out}(G);G) \leq \rho$.*

Proof. Let (x_0, x_1, \dots, x_k) be a simple path disjoint from $f^{-1}(0) \setminus \text{Out}(G)$. We have $f(x_0) \leq \rho$ and $f(x_{i+1}) < f(x_i)$ for all $i < k$, by 1(a) in Definition 13. The only vertex, in any simple path, which can be in $\text{Out}(G)$ is the last vertex, so we deduce that $x_{k-1} \notin f^{-1}(0)$, i.e. $f(x_{k-1}) > 0$. This shows that $k - 1 < \rho$, and thus $k \leq \rho$. This shows that any simple path disjoint from $f^{-1}(0) \setminus \text{Out}(G)$ has length at most ρ , which proves the lemma. □

This lemma allows us to characterise extender graphs as follows.

Corollary 16. *Let $\varepsilon, \rho > 0$ and let G be a directed acyclic graph.*

- a. *If G is an (ε, ρ) -extender, then there are no (ε, ρ) -depth functions for G .*
- b. *If there are no $(\varepsilon, \rho + 1)$ -depth functions for G , then G is an (ε, ρ) -extender.*

Proof. a. Let $\varepsilon' > 0$ and suppose that f is a (ε', ρ) -depth function for G .

By the previous lemma, we have $\text{codepth}(f^{-1}(0) \setminus \text{Out}(G);G) \leq \rho$. Therefore, since we assume that G is an (ε, ρ) -extender, we have $|f^{-1}(0) \setminus \text{Out}(G)| > \varepsilon|V(G)|$. By the definition of being a (ε', ρ) -depth function, we have $|f^{-1}(0) \setminus \text{Out}(G)| \leq \varepsilon'|V(G)|$, which shows $\varepsilon' > \varepsilon$.

- b. Let $S \subset V(G)$ be a set with $|S| \leq \varepsilon|V(G)|$. Then by Example 14, we have that δ_S is a $(\varepsilon, \text{codepth}(S;G) + 1)$ -depth function. Since we assume that there are no $(\varepsilon, \rho + 1)$ -depth functions, we deduce $\text{codepth}(S;G) + 1 > \rho + 1$, and hence $\text{codepth}(S;G) > \rho$. This shows that G is an (ε, ρ) -extender and finishes the proof. □

It will be useful to restate the above corollary for the case of graph sequences.

Corollary 17. *Let (G_n) be a bounded in-degree sequence of directed acyclic graphs and let (ρ_n) be a sequence of positive real numbers with $\lim_{n \rightarrow \infty} \rho_n = \infty$. The following conditions are equivalent.*

- a. *The sequence (G_n) is a (ρ_n) -extender sequence*
- b. *There exists $C > 0, \varepsilon > 0$ such that for all $n \in \mathbb{N}$ we have that G_n does not admit a $(\varepsilon, C \cdot \rho_{|V(G_n)|})$ -depth function.*

One of the steps in the proof of Theorem 5, Proposition 22, requires bounding the number of depth functions on a graph. We finish this subsection with a lemma which is used to count depth functions in the proof of Proposition 22. First, we need the following definition.

Definition 18. For $D \subset E(G)$, we define $\delta'_D: V(G) \rightarrow \mathbb{N} \cup \{\infty\}$ by setting $\delta'_D(v)$ to be equal to the maximal length of a directed simple path in the graph $(V(G), D)$ which connects v to a vertex in $\text{Out}((V(G), D))$.

In other words, δ'_D is the ‘standard’ depth function for the graph $(V(G), D)$. While it is not true that for every $D \subset E(G)$, we have that δ'_D is a depth function for the graph G , the following lemma shows in particular that for every depth function f we can find D such that $f = \delta'_D$.

Lemma 19. *Let G be a graph and f be a depth function for G . Then there exists $D \subset E(G)$ such that $|D| \leq |V(G)|$ and $\delta'_D = f$.*

Proof. By Condition 1b) of Definition 13, there exists a function $n: V(G) \setminus f^{-1}(0) \rightarrow V(G)$ such that for every $v \in V(G) \setminus f^{-1}(0)$ we have $f(n(v)) = f(v) - 1$. We let $D := \{(v, n(v)) : v \in V(G) \setminus f^{-1}(0)\}$. It is straightforward to check that D has the desired properties. □

3.2 Definition and basic properties of the random graphs G_n^d

In this article, a *random graph* is a pair (V, \mathbf{E}) where V is a non-empty finite set and \mathbf{E} is a random variable with values in $\text{Pow}(V \times V)$ such that \mathbf{E} is disjoint from the diagonal in $V \times V$ almost surely.

For $n \in \mathbb{N}$ let $\mathbb{Z}_n := \mathbb{Z}/n\mathbb{Z}$. For $a, b \in \mathbb{Z}_n$, we write $a > b$ if $a = \bar{a} + n\mathbb{Z}, b = \bar{b} + n\mathbb{Z}$, where $\bar{a}, \bar{b} \in \{0, 1, \dots, n - 1\}$ and $\bar{a} > \bar{b}$. We also let $R(n) := \lfloor \log(n) \rfloor$.

We start by defining a random variable J_n with values in $\mathbb{Z}_n \times \mathbb{Z}_n$, as follows. We first choose $v \in \mathbb{Z}_n$ uniformly at random, then we choose $r \in R(n)$ uniformly at random, and we choose (x, y) uniformly at random in

$$\{v, v + 1, \dots, v + 2^r - 1\} \times \{v + 2^r, v + 2^r + 1, \dots, v + 2^{r+1} - 1\} \subset \mathbb{Z}_n \times \mathbb{Z}_n.$$

The law of J_n will be denoted with ι_n .

Now for $d, n \in \mathbb{N}$, we define a random graph G_n^d as follows: we let $V(G_n^d) = \mathbb{Z}_n$, and the random variable \mathbf{E}_n^d with values in $\text{Pow}(\mathbb{Z}_n \times \mathbb{Z}_n)$ is defined by choosing dn elements of $\mathbb{Z}_n \times \mathbb{Z}_n$ independently at random according to the law ι_n . This finishes the definition of G_n^d .

Let us note that G_n^d is typically neither a dag nor of bounded degree, but the following lemma implies that with high probability G_n^d becomes a bounded degree dag after removing a small amount of vertices.

Lemma 20. *Let $\varepsilon > 0, d, n, \Delta \in \mathbb{N}$, and let $\mathbf{E} := \mathbf{E}_n^d$. We have*

$$\mathbb{P}_{\mathbf{E}}(|\{v \in \mathbb{Z}_n : \text{deg}(v; (\mathbb{Z}_n, \mathbf{E})) \geq 2\Delta\}| \geq \varepsilon \cdot n) \leq \frac{2d^\Delta}{\Delta! \cdot \varepsilon} \tag{3}$$

and

$$\mathbb{P}_{\mathbf{E}}(|\{(x, y) \in \mathbf{E} : x > y\}| \geq \varepsilon \cdot dn) \leq \frac{2d}{\varepsilon \cdot R(n)} \tag{4}$$

Proof. Note that we have the following $(\mathbb{Z}_n \times \mathbb{Z}_n)$ -valued random variable J'_n whose law is the same as the law of J_n , that is, it is equal to ι_n . We choose $(v, r) \in \mathbb{Z}_n \times R(n)$ uniformly at random, then we choose $(a, b) \in 2^r \times 2^r$ uniformly at random and we choose the edge $(v, v + 2^r + b - a)$.

Therefore, if we fix $v \in \mathbb{Z}_n$, then

$$\mathbb{P}_{\mathbf{E}}(\text{outdeg}(v; (\mathbb{Z}_n, \mathbf{E})) \geq \Delta) \leq \binom{nd}{\Delta} \frac{1}{n^\Delta} \leq \frac{d^\Delta}{\Delta!},$$

where the last inequality is obtained by writing

$$\binom{nd}{\Delta} \frac{1}{n^\Delta} = \frac{nd \cdot \dots \cdot (nd - \Delta + 1)}{\Delta!} \frac{1}{n^\Delta} \leq \frac{(nd)^\Delta}{\Delta!} \frac{1}{n^\Delta} = \frac{d^\Delta}{\Delta!}.$$

Similarly, for a fixed $v \in \mathbb{Z}_n$, we have

$$\mathbb{P}_{\mathbf{E}}(\text{indeg}(v; (\mathbb{Z}_n, \mathbf{E})) \geq \Delta) \leq \frac{d^\Delta}{\Delta!},$$

and hence

$$\mathbb{P}_{\mathbf{E}}(\text{deg}(v; (\mathbb{Z}_n, \mathbf{E})) \geq 2\Delta) \leq \frac{2d^\Delta}{\Delta!}.$$

Now by linearity of expectation, we have

$$\mathbb{E}_{\mathbf{E}}(|\{v \in \mathbb{Z}_n : \text{deg}(v; (\mathbb{Z}_n, \mathbf{E})) \geq 2\Delta\}|) = \sum_{v \in \mathbb{Z}_n} \mathbb{P}_{\mathbf{E}}(\text{deg}(v; (\mathbb{Z}_n, \mathbf{E})) \geq 2\Delta),$$

and the right-hand side is bounded from above by $\frac{2nd^\Delta}{\Delta!}$. Thus, by Markov’s inequality we have

$$\mathbb{P}_{\mathbf{E}}(|\{v \in \mathbb{Z}_n : \text{deg}(v; (\mathbb{Z}_n, \mathbf{E})) \geq 2\Delta\}| \geq \varepsilon \cdot n) < \frac{2nd^\Delta}{\Delta! \varepsilon n} = \frac{2d^\Delta}{\Delta! \varepsilon},$$

which finishes the proof of (3).

In order to prove (4), we start by bounding $\iota_n(\{(x, y) \in \mathbb{Z}_n \times \mathbb{Z}_n : x > y\})$ from above. By the definition of J'_n , the only way in which J'_n might take a value (x, y) with $x > y$ is when we start by choosing $(v, r) \in \mathbb{Z}_n \times R(n)$ such that $v > n - 2^{r+1}$. As such we have

$$\iota_n(\{(x, y) \in \mathbb{Z}_n \times \mathbb{Z}_n : x > y\}) \leq \frac{1}{R(n)} \sum_{r < R(n)} \frac{1}{n} |\{a \in \mathbb{Z}_n : a > n - 2^{r+1}\}|,$$

which is bounded from above by

$$\frac{1}{nR(n)} \sum_{r < R(n)} 2^{r+1} < \frac{1}{nR(n)} 2^{R(n)+1} \leq \frac{2n}{nR(n)} = \frac{2}{R(n)}$$

Therefore, we have

$$\mathbb{E}_{\mathbf{E}}(|\{(a, b) \in \mathbf{E} : a > b\}|) < \frac{2nd}{R(n)}$$

and Markov’s inequality again gives us the desired bound. □

3.3 Construction of an n^δ -extender sequence from \mathbf{G}_n^d

The key lemma which we need is the following.

Lemma 21. *Let $d, n \in \mathbb{N}$, let $\varepsilon > 0$, let $k := \lfloor n^{\varepsilon^3} \rfloor$, and let $l: \mathbb{Z}_n \rightarrow k$. We have*

$$\iota_n(\{(a, b) \in \mathbb{Z}_n \times \mathbb{Z}_n : l(a) > l(b)\}) < \frac{1}{2} + 4\varepsilon \tag{5}$$

Let us discuss the intuition behind the proof of Lemma 21. First, let us discuss the meaning of the left-hand side of (5). We first choose $(v, r) \in \mathbb{Z}_n \times R(n)$ uniformly at random, then we look at

the distribution of $l(v), l(v + 1), \dots, l(v + 2^r - 1)$ on the one side and the distribution of $l(v + 2^r), l(v + 2^r + 1), \dots, l(v + 2^{r+1} - 1)$ on the other side. We sample an element of $\{0, \dots, k - 1\}$ from the first distribution and an element of $\{0, \dots, k - 1\}$ from the second distribution. Then the left-hand side of (5) is the probability that the first element is larger than the second element.

If the distribution of $l(v), l(v + 1), \dots, l(v + 2^r - 1)$ is very close to the distribution of $l(v + 2^r), l(v + 2^r + 1), \dots, l(v + 2^{r+1} - 1)$, then for a random edge between the two vertex sets, l increases or decreases with approximately the same probability. But if the two distributions are not very close, then the entropy of the distribution of the union $l(v), l(v + 1), \dots, l(v + 2^{r+1} - 1)$ is larger than the average of the two entropies (this statement is formalised by Proposition 12).

As the entropy of the distribution of $l(v), \dots, l(v + 2^{R(n)})$ is bounded from above by $\log(k)$ (by Lemma 11(b)), it should be clear that when we choose k sufficiently small, then for a fixed $v \in \mathbb{Z}_n$ there will be only a small amount of r 's for which the distribution of $l(v), l(v + 1), \dots, l(v + 2^r - 1)$ and $l(v + 2^r), l(v + 2^r + 1), \dots, l(2^{r+1} - 1)$ is very different.

Proof of Lemma 21. For $v \in \mathbb{Z}_n, r < R(n)$, let $X_{v,r}$ denote the restriction of l to $[v, v + 2^r - 1] \subset \mathbb{Z}_n$, and let $Y_{v,r}$ denote the restriction of l to $[v + 2^r, v + 2^{r+1} - 1]$. We consider $X_{v,r}$ and $Y_{v,r}$ as k -valued random variables.

Note that $X_{v,r+1} = X_{v,r} \sqcup Y_{v,r}$. As such, by the first item of Proposition 12, for all v, r we have

$$2 \cdot H(X_{v,r+1}) - H(X_{v,r}) - H(Y_{v,r}) \geq 0.$$

On the other hand, we have $\mathbb{E}_{v,r}(H(X_{v,r})) = \mathbb{E}_{v,r}(H(Y_{v,r}))$, where (v, r) is chosen uniformly at random from $\mathbb{Z}_n \times R(n)$. Hence,

$$\mathbb{E}_{v,r}(2 \cdot H(X_{v,r+1}) - H(X_{v,r}) - H(Y_{v,r})) = 2\mathbb{E}_{v,r}(H(X_{v,r+1})) - 2\mathbb{E}_{v,r}(H(X_{v,r})),$$

and so

$$\begin{aligned} \mathbb{E}_{v,r}(2 \cdot H(X_{v,r+1}) - H(X_{v,r}) - H(Y_{v,r})) &= \frac{2}{R(n)} \mathbb{E}_v \left(\sum_{r < R(n)} H(X_{v,r+1}) - H(X_{v,r}) \right) \\ &= \frac{2}{R(n)} \mathbb{E}_v(H(X_{v,R(n)}) - H(X_{v,0})) \\ &\leq \frac{2 \log(k)}{R(n)}. \end{aligned}$$

Now Markov's inequality shows that

$$\mathbb{P}_{v,r}(2 \cdot H(X_{v,r+1}) - H(X_{v,r}) - H(Y_{v,r}) \geq \varepsilon^2) \leq \frac{2 \log(k)}{\varepsilon^2 \cdot R(n)} \leq 2\varepsilon.$$

By the second item of Proposition 12, if for some r, v, ε we have $H(X_{v,r+1}) - H(X_{v,r}) - H(Y_{v,r}) < \varepsilon^2$, then $\mathbb{P}_{x,y}(X_{v,r}(x) > Y_{v,r}(y)) \leq \frac{1}{2} + 2\varepsilon$. Thus by the definition of ι_n , we have

$$\begin{aligned} \iota_n(\{(a, b) \in \mathbb{Z}_n \times \mathbb{Z}_n : l(a) > l(b)\}) &< \\ &< \frac{1}{2} + 2\varepsilon + \mathbb{P}_{v,r}(2 \cdot H(X_{v,r+1}) - H(X_{v,r}) - H(Y_{v,r}) \geq \varepsilon^2) \\ &\leq \frac{1}{2} + 4\varepsilon, \end{aligned}$$

which finishes the proof. □

Proposition 22. Let $d, n \in \mathbb{N}, d \geq 3$, let $\mathbf{E} := \mathbf{E}_n^d$, let $\varepsilon \in (0, 1)$, and let $k := \lfloor n^{\varepsilon^3} \rfloor$,

$$\mathbb{P}_{\mathbf{E}}((\mathbb{Z}_n, \mathbf{E}) \text{ admits a } (\varepsilon, k)\text{-depth function}) < 2^{H(\frac{1}{d})dn} \left(\frac{1}{2} + 4\varepsilon\right)^{(d-1)n}, \tag{6}$$

where for $x \in (0, 1)$ we set $H(x) = -x \log(x) - (1 - x) \log(1 - x)$.

Proof. Clearly it is enough to show that

$$\mathbb{E}_{\mathbf{E}}(\text{number of } (\varepsilon, k)\text{-depth functions for } (\mathbb{Z}_n, \mathbf{E})) < 2^{H(\frac{1}{d})dn} \left(\frac{1}{2} + 4\varepsilon\right)^{(d-1)n}. \tag{7}$$

By Lemma 19, the left-hand side of (7) is bounded above by

$$\mathbb{E}_{\mathbf{E}}\left(\left|\left\{D \subset \mathbf{E}: |D| \leq n \text{ and } \delta'_D \text{ is an } (\varepsilon, k)\text{-depth function for } (\mathbb{Z}_n, \mathbf{E})\right\}\right|\right). \tag{8}$$

Given $I \subset dn$, let $\text{Set}_I: (\mathbb{Z}_n \times \mathbb{Z}_n)^{dn} \rightarrow \text{Pow}(\mathbb{Z}_n \times \mathbb{Z}_n)$ be defined by

$$\text{Set}_I(x_0, \dots, x_{dn-1}) := \{x_i: i \in I\},$$

and let $\text{Set} := \text{Set}_{dn}$. Furthermore, if G is a graph and $D \subset E(G)$ is such that δ'_D is a (ε, k) -depth function, then let us say that D is an (ε, k) -depth set for G .

Recall that the law of \mathbf{E} is the push-forward of ι_n^{dn} through the map Set . As such, we deduce that (8) is bounded above by:

$$\sum_{I \subset dn: |I| \leq n} \iota_n^{dn}\left(\left\{(e_i)_{i \in dn} \in (\mathbb{Z}_n \times \mathbb{Z}_n)^{dn}: \text{Set}_I((e_i)_{i \in dn}) \text{ is a } (\varepsilon, k)\text{-depth set for } (\mathbb{Z}_n, \text{Set}((e_i)_{i \in dn}))\right\}\right) \tag{9}$$

Let us first estimate the number of summands in (9). Recall that for $0 < \alpha \leq \frac{1}{2}$ and $m \in \mathbb{N}$ we have $\sum_{i \leq \alpha m} \binom{m}{i} \leq 2^{H(\alpha)m}$ (see e.g. [3, Theorem 3.1]). Since $\varepsilon < 1$ and $d \geq 3$, we see that the number of summands in (9) is therefore at most $2^{H(\frac{1}{d})dn}$.

To estimate each summand, let us fix $I \subset dn$, and let us fix $(e_i)_{i \in I} \in (\mathbb{Z}_n \times \mathbb{Z}_n)^I$. Let $D := \{e_i: i \in I\}$ and let l be the depth function δ'_D on the graph (\mathbb{Z}_n, D) . The probability that l will still be a depth function after we add $nd - |I|$ remaining edges is, by Lemma 21, at most

$$\left(\frac{1}{2} + 4\varepsilon\right)^{dn-|I|} \leq \left(\frac{1}{2} + 4\varepsilon\right)^{(d-1)n}.$$

As such, we have that (9) is bounded above by

$$2^{H(\frac{1}{d})dn} \left(\frac{1}{2} + 4\varepsilon\right)^{(d-1)n},$$

and hence (7) and (6) hold true. This finishes the proof. □

We are now ready to prove Theorem 5. Clearly, it follows from the following theorem.

Theorem 23. *Let $\varepsilon \in (0, \frac{1}{8})$. Then there exists a bounded degree sequence (H_n) of directed acyclic graphs which is an (n^{ε^3}) -extender sequence.*

Proof. Let $\delta > 0$ be such that $p := \frac{1}{2} + 4(\varepsilon + 2\delta) < 1$. Let d be such that

$$2^{H(\frac{1}{d})} \cdot \left(\frac{1}{2} + 4(\varepsilon + 2\delta)\right)^{\frac{d-1}{d}} < 1. \tag{10}$$

It is possible to choose such d since $H(x) \rightarrow 0$ as $x \rightarrow 0$. Let $\Delta \in \mathbb{N}$ be such that $\frac{2d^\Delta}{\Delta! \cdot \delta} < \delta$, and let n_0 be such that for $n > n_0$ we have $\frac{2d^2}{\delta R(n)} < \delta$ and

$$2^{H\left(\frac{1}{d}\right)dn} \left(\frac{1}{2} + 4(\varepsilon + 2\delta)\right)^{(d-1)n} < 1 - 2\delta,$$

which is possible by (10).

Therefore, by Proposition 22, we have for $n > n_0$ that

$$\mathbb{P}_{\mathbf{E}}(\{v \in \mathbb{Z}_n : \deg(v; \mathbb{Z}_n, \mathbf{E}) \geq 2\Delta\} \geq \delta \cdot n) < 1 - 2\delta.$$

Furthermore, by Lemma 20, we have

$$\mathbb{P}_{\mathbf{E}}(|\{v \in \mathbb{Z}_n : \deg(v; \mathbb{Z}_n, \mathbf{E}) \geq 2\Delta\}| \geq \delta \cdot n) \leq \frac{2d^\Delta}{\Delta! \cdot \delta} < \delta$$

and

$$\mathbb{P}_{\mathbf{E}}(|\{(a, b) \in \mathbf{E} : a > b\}| \geq \frac{\delta}{d} \cdot dn = \delta n) \leq \frac{2d}{\delta R(n)} < \delta.$$

As such, by the union bound, we get for each $n > n_0$ a graph G_n with $V(G_n) = \mathbb{Z}_n$ such that G_n does not admit a $(\varepsilon + 2\delta, n^{(\varepsilon+2\delta)^3})$ -depth function, and furthermore

$$|\{v \in \mathbb{Z}_n : \deg(v; G_n) \geq 2\Delta\}| \leq \delta \cdot n$$

and

$$|\{(a, b) \in E(G_n) : a > b\}| \leq \delta \cdot n.$$

Let $B \subset \mathbb{Z}_n$ be the union of

$$\{v \in \mathbb{Z}_n : \deg(v; G_n) \geq 2\Delta\}$$

and

$$\{a \in \mathbb{Z}_n : \exists b \in \mathbb{Z}_n \text{ such that } (a, b) \in E(G_n) \text{ and } a > b\}.$$

We let H_n be the subgraph of G_n induced by the set of vertices $\mathbb{Z}_n \setminus B$. Clearly, H_n is a sequence of bounded degree dags, and since $|B| \leq 2\delta n$, we see that H_n does not admit a $(\varepsilon, n^{(\varepsilon+2\delta)^3})$ -depth function, and hence it also does not admit a $(\varepsilon, n^{\varepsilon^3})$ -depth function. By Corollary 17, this finishes the proof. □

4. Final remarks

Let us proceed with the proof of Theorem 6. Clearly, Theorem 6 follows from the following proposition.

Proposition 24. *Let (G_n) be a sequence of bounded in-degree directed acyclic graphs and let (δ_n) be a sequence of real numbers in the interval $(0, 1]$ such that $\lim_{n \rightarrow \infty} \delta_n = 1$. For every $\varepsilon > 0$, there exists a sequence (S_n) with $S_n \subsetneq V(G_n)$ such that $|S_n| < \varepsilon |V(G_n)|$ and*

$$\lim_{n \rightarrow \infty} \frac{\text{codepth}(S_n; G_n)}{|V(G_n)|^{\delta_n}} = 0 \tag{11}$$

Proof. Let $m_n := |V(G_n)|$, let $d \in \mathbb{N}$ be such that $\text{maxindeg}(G_n) \leq d$ for all $n \in \mathbb{N}$, and let us fix $\varepsilon > 0$. Since the graphs G_n are dags, we may assume that $V(G_n) = \{0, \dots, m_n - 1\}$ and that $(x, y) \in E(G_n)$ implies $x < y$.

Let us informally describe our strategy for constructing the sets S_n : first, we will include in S_n all vertices adjacent to an edge of length between $m_n^{c_n}$ and $m_n^{c_n+\varepsilon}$ for a suitable c_n . This way any directed path disjoint from S_n will have edges of length either less than $m_n^{c_n}$ ('short edges') or larger than $m_n^{c_n+\varepsilon}$ ('long edges').

The number of the long edges in a directed path is at most $m_n^{1-c_n-\varepsilon}$. To bound the number of short edges, we will also include in S_n all vertices which are congruent to at most $m_n^{c_n}$ modulo $\frac{m_n^{c_n}}{\varepsilon}$. This way any path disjoint from S_n and consisting only of short edges must be contained in an interval of length $\frac{m_n^{c_n}}{\varepsilon}$, and so in particular its length is at most $\frac{m_n^{c_n}}{\varepsilon}$.

These bounds on the total number of long edges, and the maximal number of consecutive short edges allow us to obtain the desired bound on $\text{codepth}(S_n; G_n)$. Let us now make it precise.

Since $|E(G_n)| \leq dm_n$, by the pigeon hole principle we may find $c_n \in [0, 1)$ such that the set:

$$X_n := \{(x, y) \in E(G_n) : y - x \in [m_n^{c_n}, m_n^{c_n+\varepsilon})\}$$

has cardinality at most $\varepsilon \cdot dm_n$. Let $A := \lfloor \frac{m_n^{c_n}}{\varepsilon} \rfloor$, let $B := \lfloor m_n^{c_n} \rfloor$, and let

$$Y_n := \{x \in V(G_n) : x \equiv k \pmod A \text{ for some } k \text{ with } 0 \leq k < B\}.$$

Finally, we let

$$S_n := \{x \in V(G_n) : \exists y \in V(G_n) \text{ such that } (x, y) \in X_n\} \cup Y_n.$$

Clearly, we have $|S_n| \leq |X_n| + |Y_n| \leq \varepsilon dm_n + \varepsilon m_n = \varepsilon(d + 1)m_n$. Thus, since ε was arbitrary, in order to finish the proof it is enough to argue that (11) holds.

In order to estimate $\text{codepth}(S_n; G_n)$ let us fix $n \in \mathbb{N}$ and let

$$(x_0, \dots, x_l) \tag{12}$$

be a directed path in G_n disjoint from S_n . By the definition of X_n , and since S_n contains all starting vertices of edges in X_n , we see that for all $i < l$ we have either $x_i - x_{i+1} < m_n^{c_n}$ or $x_i - x_{i+1} \geq m_n^{c_n+\varepsilon}$.

Let $(x_j, x_{j+1}, \dots, x_{j+M})$ be a maximal subpath of (12) such that for all $k < M$ we have $x_{j+k+1} - x_j < m_n^{c_n}$. Since $Y_n \subset S_n$, we see that $M \leq \frac{m_n^{c_n}}{\varepsilon}$. On the other hand, the maximal number of edges in the path (12) with length at least $m_n^{c_n+\varepsilon}$ is at most $\frac{m_n}{m_n^{c_n+\varepsilon}}$.

In other words, there are at most $n^{c_n+\varepsilon}$ segments in (12), where each segment consists of at most M 'short' edges and a single 'long' edge. It follows that the length of the path (12) is bounded by:

$$(M + 1) \cdot \frac{m_n}{m_n^{c_n+\varepsilon}} \leq \left(\frac{m_n^{c_n}}{\varepsilon} + 1 \right) \frac{m_n}{m_n^{c_n+\varepsilon}} \leq \left(\frac{2m_n^{c_n}}{\varepsilon} \right) \frac{m_n}{m_n^{c_n+\varepsilon}} = \frac{2}{\varepsilon} m_n^{1-\varepsilon},$$

and hence

$$\text{codepth}(S_n, G_n) \leq \frac{2}{\varepsilon} m_n^{1-\varepsilon}.$$

In particular, since $m_n = |V(G_n)|$, this establishes (11) and finishes the proof. □

Circuit complexity. We finish this article by explaining some conjectural applications of hyper-shallow graph sequences to the theory of boolean circuits. As this is not of crucial importance for this article, we allow ourselves to be a little bit less precise for the sake of brevity.

If X is a set, then 2^X is the set of all functions from X to $2 = \{0, 1\}$. This leads to the following notational clash: for $n \in \mathbb{N}$, the symbol 2^n can either denote a number (and hence a set of numbers) or the set of all functions from $\{0, 1, \dots, n - 1\}$ to $\{0, 1\}$. We believe that resolving this ambiguity will not cause any difficulty for the reader.

A convenient informal way of thinking about it is that if $k \in 2^n$ then k is both a number smaller than 2^n and a function from $n = \{0, \dots, n - 1\}$ to $2 = \{0, 1\}$, and the translation between the two

interpretations is that the binary expansion of a number k smaller than 2^n can be thought of as a function from $\{0, \dots, n - 1\}$ to $\{0, 1\}$.

A *circuit* is a pair $\mathcal{C} = (G, \text{gate})$, where G is a dag and gate is a function which assigns to each vertex $v \in V(G) \setminus \text{In}(G)$ a function $\text{gate}(v): 2^{\text{In}(v;G)} \rightarrow 2$. We will inherit the notation for \mathcal{C} from the notation for G , thus, for example, we may write $\text{In}(\mathcal{C})$ for $\text{In}(G)$.

For any $f \in 2^{\text{In}(\mathcal{C})}$, there exists exactly one function $F \in 2^{V(\mathcal{C})}$ with the property that for every $v \in V(\mathcal{C}) \setminus \text{In}(\mathcal{C})$ we have $\text{gate}(v)(F|_{\text{In}(v;\mathcal{C})}) = F(v)$. In particular, we think of the restriction of F to $\text{Out}(G)$ as the output of the circuit \mathcal{C} when f is ‘fed’ as the input.

Typically, both $\text{In}(\mathcal{C})$ and $\text{Out}(\mathcal{C})$ have some labels, for example, both $\text{In}(\mathcal{C})$ and $\text{Out}(\mathcal{C})$ are labelled with elements of $\{0, \dots, n - 1\}$, in which case we may consider \mathcal{C} to implement a function $2^n \rightarrow 2^n$.

By a simple counting argument, ‘almost every’ sequence of functions $(f_n: 2^n \rightarrow 2^n)$ cannot be implemented by a sequence of bounded in-degree circuits (\mathcal{C}_n) such that $|V(G_n)| = O(n)$. However, it is notoriously difficult to give ‘explicit’ examples of sequences which cannot be computed by linear-sized circuits.

Following [1], let us state one precise question.

Definition 25. For $i \in \mathbb{N}$ we let $l(i) = \lceil \log(i) \rceil$, and we define $\text{shift}_n: 2^{n \sqcup l(n)} \rightarrow 2^n$ as follows: if $f \in 2^n$ and $k \in 2^{l(n)}$, then for $j < n$ we let

$$\text{shift}_n(f \sqcup k)(j) := f(j - k),$$

where $j - k$ should be understood as an operation modulo n . In other words, $\text{shift}_n(f \sqcup k)$ is equal to ‘ f shifted by k ’.

Question 26. Suppose that (\mathcal{C}_n) is a bounded in-degree sequence of circuits which computes shift_n . Is it the case that $n = o(|V(G_n)|)$?

This innocent-looking question seems difficult to resolve (though there are some conditional results in [1]). The authors of this article came up with the notion of hypershallow graph sequences motivated by the following strategy to attack this question: (1) ‘Clearly’ if (\mathcal{C}_n) is a hypershallow sequence which computes shift_n , then $n = o(|V(\mathcal{C}_n)|)$ and (2) perhaps all graph sequences are hypershallow.

The main result of this paper is that not all graph sequences are hypershallow (Theorem 5). More annoyingly, the authors have not even been able to establish the first point of the above strategy. As such, the following question is also open.

Question 27. Suppose that (\mathcal{C}_n) is a bounded in-degree sequence of circuits which computes shift_n and which is hypershallow. Is it the case that $n = o(|V(\mathcal{C}_n)|)$?

Let us finish this article by stating another question to which positive answer would imply a positive answer to Question 27. We need to start with some definitions.

An *advice circuit* is a circuit \mathcal{C} together with a partition of $\text{In}(\mathcal{C})$ into two disjoint subsets $\text{In}_{\text{std}}(\mathcal{C})$ and $\text{In}_{\text{adv}}(\mathcal{C})$. We think of such a circuit as receiving its input on the vertices in $\text{In}_{\text{std}}(\mathcal{C})$, together with some extra advice tailored specifically for a given input on the vertices in $\text{In}_{\text{adv}}(\mathcal{C})$. This is made precise in the following definition.

Definition 28. Let \mathcal{C} be an advice circuit. We say that \mathcal{C} *computes* $f: 2^{\text{In}_{\text{std}}(\mathcal{C})} \rightarrow 2^{\text{Out}(\mathcal{C})}$ if for every $s \in 2^{\text{In}_{\text{std}}(\mathcal{C})}$ there exists $t \in 2^{\text{In}_{\text{adv}}(\mathcal{C})}$ such that the output of \mathcal{C} on $s \sqcup t$ is equal to $f(s)$.

An ε -advice circuit is an advice circuit C with $|\text{In}_{adv}(C)| \leq \varepsilon |\text{In}(C)|$. With this we are ready to state the following question.

Question 29. Is it true that there exists $\varepsilon > 0$ such that the sequence (shift_n) cannot be computed by a sequence (C_n) of bounded in-degree ε -advice circuits which have depth 1?

It is not difficult to see that the positive answer to this question implies the positive answer to Question 27.

Acknowledgements

We thank one of the anonymous referees for correcting several typos and errors, and for very useful suggestions for improving the readability of this article. We also thank the organisers of the workshop *Measurability, Ergodic Theory and Combinatorics* which took place at Warwick University in July 2019. A very substantial progress on this project happened during that workshop, and we are grateful for inviting both of us and for providing an excellent environment for mathematical collaboration. Finally, we thank the authors of the blog *Gödel's Lost Letter and P=NP* for blogging about the paper [1] and thereby bringing it to our attention (as well as for many other very interesting posts over the years). This was the starting point for this project.

References

- [1] Afshani, P., Freksen, C. B., Kamma, L. and Larsen, K. G. (2019) Lower bounds for multiplication via network coding. In *46th International Colloquium on Automata, Languages, and Programming (ICALP 2019)*, pp. 10:1–10:12.
- [2] Elek, G. (2007) The combinatorial cost. *Enseign. Math. (2)* **53**(3–4) 225–235.
- [3] Galvin, D. (2014) Three tutorial lectures on entropy and counting, available at <https://arxiv.org/abs/1406.7872v1>.
- [4] Hoory, S., Linial, N. and Wigderson, A. (2006) Expander graphs and their applications. *Bull. Am. Math. Soc. (N.S.)* **43**(4), 439–561.
- [5] Lipton, R. J. and Tarjan, R. E. (1980) Applications of a planar separator theorem. *SIAM J. Comput.* **9**(3), 615–627.
- [6] Massart, P. (2007) *Concentration Inequalities and Model Selection*, Vol. 1896 of *Lecture Notes in Mathematics*, Springer, Berlin. Lectures from the 33rd Summer School on Probability Theory held in Saint-Flour, July 6–23, 2003, With a foreword by Jean Picard.