Efficiently Constructible Huge Graphs that Preserve First Order Properties of Random Graphs

Moni Naor [⋆], Asaf Nussboim [⋆⋆], and Eran Tromer

Department of Computer Science and Applied Mathematics,
Weizmann Institute of Science, Rehovot 76100, Israel
{moni.naor,asaf.nussbaum,eran.tromer}@weizmann.ac.il

Abstract. We construct efficiently computable sequences of random-looking graphs that preserve properties of the canonical random graphs $G(2^n, p(n))$. We focus on first-order graph properties, namely properties that can be expressed by a formula ϕ in the language where variables stand for vertices and the only relations are equality and adjacency (e.g. having an isolated vertex is a first-order property $\exists x \forall y (\neg \text{EDGE}(x, y))$). Random graphs are known to have remarkable structure w.r.t. first order properties, as indicated by the following 0/1 law: for a variety of choices of $p(n)$, any *fixed* first-order property ϕ holds for $G(2^n, p(n))$ with probability tending either to 0 or to 1 as n grows to infinity.

We first observe that similar 0/1 laws are satisfied by $G(2^n, p(n))$ even w.r.t. sequences of formulas $\{\phi_n\}_{n\in\mathbb{N}}$ with bounded quantifier depth, $depth(\phi_n) \leq \frac{n}{\lg(1/p(n))}$. We also demonstrate that 0/1 laws do not hold for random graphs w.r.t. properties of significantly larger quantifier depth. For most choices of $p(n)$, we present efficient constructions of huge graphs with edge density nearly $p(n)$ that emulate $G(2^n, p(n))$ by satisfying $\Theta(\frac{n}{\lg(1/p(n))})$-0/1 laws. We show both probabilistic constructions (which also have other properties such as K-wise independence and being computationally indistinguishable from $G(N, p(n))$), and deterministic constructions where for each graph size we provide a specific graph that captures the properties of $G(2^n, p(n))$ for slightly smaller quantifier depths.

1 Introduction

We deal with small families of graphs that resemble large ones. In general we think of our graphs as being huge so they are not represented explicitly, but rather by a procedure that evaluates edge-queries using a succinct representation (a seed) of the graph. Such small families are sampled by randomly picking the succinct representation.

We attempt to capture a large class of properties of truly random graphs $G(N, p)$ where $N = 2^n$ vertices are fixed and the edges are independently picked

⋆ Partly supported by a grant from the Israel Science Foundation.
⋆⋆ Partly supported by the Minerva Foundation 2-8495.

each with probability $p = p(n)$. A prominent class of properties is that of first order properties, namely those that can be expressed by a formula ϕ in the language where variables stand for vertices and the only relations are equality and adjacency (e.g containing a triangle is a first order property of quantifier depth 3 written as $\exists x \exists y \exists z\ (\text{EDGE}(x,y)) \bigwedge (\text{EDGE}(x,z)) \bigwedge (\text{EDGE}(y,z))$. Random graphs are known to exhibit remarkable structure w.r.t. first order properties, namely the famed 0/1 law: any fixed first-order property ϕ holds for $G(N,p)$ with probability tending either to 0 or to 1 as N grows to infinity[1]. Thus one can view this work as dealing with graphs that look random to distinguishers that are expressible as first order properties.

We show that for sufficiently large k, any *exact* k-wise independent graphs (defined below) preserve the 0/1 law of random graphs (this is not true for *almost* k-wise independent graphs). We also show a construction of computationally pseudo-random graphs that satisfy the 0/1 law of random graphs (note that in general, computational pseudo-randomness does not imply such combinatorial properties). Finally, we provide for each graph size a single graph that captures the first order properties of $G(N,p)$, and is efficiently computable. Those results can be extended to first-order properties of quantifier depth up to $\frac{n}{\log(1/p)}$.

On the other hand we show that no efficiently constructed family of graphs can achieve $D(n)$-equivalence to random graphs w.r.t. to an arbitrarily large polynomial $D(n)$. Before elaborating on our main results, we review other notions that capture aspects of the structure of random graphs.

1.1 Random-Looking Graphs

Several characterizations for the concept of a "random-looking" graph have been extensively studied and are known to have a wealth of applications in combinatorics and computer science:

$K(n)$-wise independent graphs. These are a relaxation of $G(N,p(n))$ in the sense that each edge appears w.p. $p(n)$, and the distribution of any fixed $K(n)$ potential edges is mutually independent. Efficient constructions of n^c-wise independent graphs are known for all fixed c and a wide variety of densities $p(n)$ (e.g., [2]).

Combinatorial pseudo-random graphs. This term refers to a collection of definitions that consider a single graph g_n for each size n and intend to capture the edge distribution of $G(N,p)$ by requiring that any induced subgraph of g_n has density $\approx p$. Two of the variants are Thomason's jumbled graphs where for each vertex set U, $\left||E(U)| - p\binom{|U|}{2}\right| \leq \alpha|U|$, where $\alpha = \sqrt{pN}$ is the desired accuracy achieved by $G(N,p)$ and $E(U)$ is the set of vertices in the subgraph induced by U (see [23]). A weaker (yet very useful) definition is *quasi-random graphs*, which requires only that $\forall U \left||E(U)| - p\binom{|U|}{2}\right| \leq o(N^2)$. Quasi-random graphs were shown by Chung, Graham and Wilson ([6]) to be equivalent to the

[1] Note that despite the term "law", the 0/1 law is actually a characteristic that may or may not hold for specific families of graphs.

surprisingly innocent condition that the number of labeled cycles of length 4 is $(pN)^4(1 \pm o(1))$ when $E(g_n) = (p \pm o(1))\binom{N}{2}$. Several deterministic constructions are known for such quasi-random and jumbled graphs (see a recent survey by Krivelevich and Sudakov [16]).

Computationally pseudo-random graphs. These are defined as graphs which are computationally indistinguishable from random graphs [13], in the sense of [12]. Namely, no polynomial-time distinguishing algorithm that performs edge-queries of its choice can tell apart a pseudo-random graph from a random graph $G(N, p(n))$. Explicit constructions of computationally pseudo-random graphs are easily derived from pseudo-random functions. The latter are known to exist iff one-way functions exist [12, 15].

Graphs that preserve specific combinatorial properties of random graphs. Random graphs are known to exhibit a remarkable combinatorial structure (see Bollobás's survey [4]). For instance, consider $G(2^n, 1/2)$ which is the same as the uniform distribution on all 2^n-vertices graphs, and let $N = 2^n$. Then for some value $s(N) \approx 2 \lg N$, it holds that with overwhelming probability $G(2^n, 1/2)$ is:

1. Connected, Hamiltonian, and has a perfect matching.
2. Has clique number and independence number precisely $s(N) \pm 1$.
3. Has chromatic number $\frac{N}{s(N)}(1 \pm \frac{1}{\sqrt{\lg(N)}})$.
4. Has maximal and minimal degree $\frac{1}{2}N(1 \pm 2\sqrt{\frac{\lg(N)}{N}})$.
5. Has connectivity number $\frac{1}{2}N(1 \pm 2\sqrt{\frac{\lg(N)}{N}})$.

Some, but not all, of these properties are met by $poly(n)$-wise independent graphs, and by combinatorial pseudo-random graphs. It was shown in [13][17] that there are efficient constructions of graphs which are simultaneously: computationally pseudo-random (w.r.t. $G(2^n, 1/2)$), almost n^c-wise independent, preserve properties 1–3 above, and approximate properties 4 and 5.

Our work. While the constructions of [13][17] are tailor-made to preserve a *fixed* number of prescribed properties (some of which are provably more complex then first-order properties), the current work constructs small families of graphs that preserve *arbitrary* first-order properties of random graphs and in addition may be computationally pseudo-random (w.r.t. $G(N, p)$) and n^c-wise independent. Alternatively we construct a single graph that satisfies *arbitrary* first-order properties of random graphs.

1.2 Preserving First-Order Properties of Random Graphs

First-order properties are graph properties that can be expressed in first order language, where the variables stand for vertices and the only relations are equality and adjacency. For instance, having an isolated vertex can be written as $\exists x \forall y \neg \textsc{edge}(x, y)$ (see section 2 for definitions).

From the first-order lens, random graphs exhibit a remarkable structure (see Spencer's [19] for an excellent survey). The following 0/1 law is known to hold for $G(N, p)$: every first order property ψ holds with probability tending either to 0 or to 1 as the size of the graph grows to infinity. The case where p is constant is due to Fagin [9] and independently Glebskii et al [11]. The other known case where $p(n) = 2^{-\alpha n}$ for an irrational α is due to Shelah and Spenser [22].

Can one efficiently construct random-looking graphs that resemble $G(N, p(n))$ and satisfy this 0/1 law? The answer is positive, but we shall actually consider graphs that meet a much stronger requirement.

Generalized 0/1 Laws. Rather than fixing a single first-order formula, we shall consider sequences of formulas $\Phi = \{\phi_n\}_{n\in\mathbb{N}}$. Such a sequence can express much richer properties than a single formula. For instance, containing a clique of size $\lg n$ can be expressed by the sequence where $\phi_n = \exists x_1 ... \exists x_{\lg n} \bigwedge_{i \neq j}((x_i \neq x_j) \bigwedge \text{EDGE}(x_i, x_j))$, and the quantifier depth is $depth(\phi_n) = \lg n$ (quantifier depths are formally defined in section 2).

A natural generalization of the basic 0/1 law is the $D(n)$-0/1 law which is satisfied by huge graphs $\mathcal{G}$ if for any sequence Φ having quantifier depth $depth(\phi_n) \leq D(n)$ it holds that

$$\lim_{n\to\infty} \Pr[\mathcal{G}_n \models \phi_n] \in \{0, 1\}. \tag{1}$$

Choosing the quantifier depth as the complexity measure for Φ, rather than the entire length of the formulas, will be well-motivated by the discussed results. Some relaxation of this definition is required, however, since for any sequence Φ satisfying the limit condition in (1), if we negate all formulas for odd n then the limit no longer exists. This shows that with the above definition can never be satisfied, even when $D(n) = 1$. This is overcome by requiring (instead of condition (1)) that for each sequence Φ satisfying $depth(\phi_n) \leq D(n)$ there exists a similar sequence Φ' s.t. $\phi'_n \in \{\phi_n, \neg\phi_n\}$, and $\Pr[\mathcal{G}_n \models \phi'_n] \stackrel{n\to\infty}{\longrightarrow} 1$.

Alas, it can be easily seen that with the above relaxation, the $D(n)$-0/1 laws no longer imply the basic 0/1 law. Thus, to reinstate this implication we explicitly also require that for any fixed formula $\phi \in \Phi$ the limit $\lim_{n\to\infty} \Pr[\mathcal{G}_n \models \phi]$ should exist. Note that with this final definition, satisfying the basic 0/1 law is identical to satisfying the $D(n)$-0/1 law for all $D(n) = \Theta(1)$.

Next, recall that we wish to formalize the notion of some huge graphs $\mathcal{G}^1$ preserving the first-order properties of $\mathcal{G}^2 = G(N, p)$. Having a 0/1 law hold for both $\mathcal{G}^1$ and $\mathcal{G}^2$ may not suffice as it might be the case that $\Pr[\mathcal{G}^1_n \models \phi_n] \stackrel{n\to\infty}{\longrightarrow} 1$, whereas $\Pr[\mathcal{G}^2_n \models \phi_n] \stackrel{n\to\infty}{\longrightarrow} 0$. Therefore the following definition is introduced: $\mathcal{G}^1$ and $\mathcal{G}^2$ are said to be $D(n)$-equivalent, if for any sequence Φ having quantifier depth $depth(\phi_n) \leq D(n)$, it holds that $\lim(\Pr[\mathcal{G}^1_n \models \phi_n] - \Pr[\mathcal{G}^2_n \models \phi_n]) \stackrel{n\to\infty}{\longrightarrow} 0$.

1.3 Our Results

Maximal 0/1 laws for random graphs. We start by establishing the maximal depth, $D(n)$, for which $G(N, p(n))$ satisfies $D(n)$-0/1 laws. For any choice of $p(n)$,[2] we set $D^* = D^*(n, p(n)) = \frac{n(1-o(1))}{\lg(1/p(n))}$ and show that $G(N, p(n))$ satisfies the D^*-0/1 law. On the other hand, we show that for any $p(n)$ there exists $p'(n) = p(n)(1 - o(1))$ s.t. $G(N, p'(n))$ defies the $2D^*$-0/1 law as long as $p(n) \geq 2^{o(\sqrt{n})}$.

A probabilistic construction. For D^* as above, we show that arbitrary n^3-wise independent graphs satisfy the D^*-0/1 law and are D^* equivalent to $G(N, p(n))$. Since for any non-trivial[3] density $p(n)$ there are explicit efficient constructions of n^3-wise independent graphs $\mathcal{G}$ with density $p'(n) = p(n)(1 - o(1))$, our goal is accomplished. A modification of the construction for $\mathcal{G}$ can guarantee (in addition to the above), the computational indistinguishability of $\mathcal{G}$ from $G(N, p(n))$, if one-way functions exist.

Deterministic construction using Paley graphs. We show that for every n and p there exists a specific efficiently computable graph of size $N' = 2^{\Theta(n)}$ and edge density $p' = p \pm \epsilon$, which is $D(n)$-equivalent to $G(N', p')$. Here $D(n)$ depends on ϵ; for example, for $\epsilon(n) > \Theta(1/n)$ we get $D(n) > \frac{n}{2\lg(1/\epsilon)}(1-o(1))$. For the special case $p = 1/2$ we obtain edge density exactly p and $D(n) = \Theta(n)$ which is optimal up to a factor of $4 + o(1)$.

Negative results. While the above positive results are close to optimal, one may still consider the case where $D(n)$ equivalence to random graphs is desired for $D(n)$ so large that $D(n)$-0/1 laws no longer hold for $G(N, p)$. We obtain the following negative result: efficiently constructed graphs $\mathcal{G}$ with seed length $m(n)$ are never $D(n)$-equivalent to $G(2^n, \frac{1}{2})$, for $D(n) = \omega(n + \sqrt{m(n)})$. If one wishes to separate $\mathcal{G}$ from $G(2^n, \frac{1}{2})$ by sequences that have $poly(n)$ total length, then a similar negative result holds for $D(n) = 2m + \omega(n + \sqrt{m(n)})n$. Similar results can be obtained for various choices of p.

1.4 Relationships Among Concepts of Random-Looking Graphs.

Figure 1 summarizes the relationships between the main notions of random-looking graphs for a given density $p(n)$. A black arrow stands for implication, while a dotted one implies that implication fails to hold (the bottom left square refers to the conjunction of the properties). Interestingly, while no notion implies all the others, a single construction achieves all four requirements simultaneously (assuming that one-way functions exist).

We sketch the references to the information given in the table. The two following facts are well known. Any computationally pseudo-random graphs with seed length n^c are statistically far from any n^{c+1}-wise independent graphs. On the other hand, n^c-wise independent graphs generated via polynomials of degree n^c, are easily distinguished from random graphs using only $n^c + 1$ edge queries.

Next, quasi-randomness, $D(n)$-0/1 laws and $D(n)$-equivalence to random graphs may hold even for a single graph per size, and consequently, these con-

[2] Throughout this subsection we assume that $p(n) \leq \frac{1}{2}$. Otherwise each term $p(n)$ concerning quantifier depths should be replaced by $\min\{p(n), 1 - p(n)\}$).

[3] A trivial density is one for which the graph is empty w.p. $1 - o(1)$.

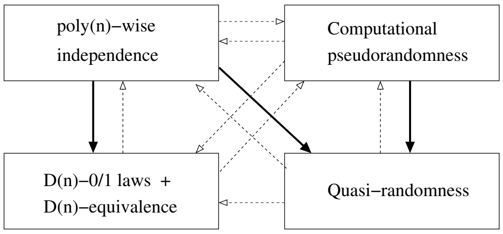

Fig. 1. Relation between notions of random-looking graphs

ditions do not imply neither K-wise independence nor computational pseudo-randomness.

Using the equivalent condition for quasi-randomness concerning the number of 4-cycles, it is easy to show that quasi-randomness is guaranteed by either computational pseudo-randomness or by $\Theta(1)$-wise independence.

Next, it can be seen that computational pseudo-randomness, and (consequently by the above) also quasi-randomness, both fail to imply even depth-2 0/1 laws and depth-2 equivalence to random graphs. Indeed, assuming the existence of one-way functions, by [13] there exist pseudorandom graphs that have an isolated vertex for odd n but are connected for even n.

Although we can provide graphs satisfying $D(n)$-0/1 laws without achieving quasi-randomness, it is not clear whether $D(n)$-0/1 laws combined with $D(n)$-equivalence to random graphs implies quasi-randomness or not.

Finally, our probabilistic construction shows that n^3-wise independence ensures optimal 0/1 laws and optimal equivalence to random graphs. When this construction is strengthen to maintain computational pseudo-randomness (assuming that one-way functions exist), we achieve a single construction which simultaneously meets all 4 criteria for a being random-looking graph.

2 Preliminaries

2.1 First Order Logic on Graphs

Formally, the alphabet of first order logic on graphs is made of:

1. Infinitely many variable symbols such as 'x','y','z' .
2. The binary relation symbols '=' and 'EDGE'.
3. The quantifier symbols '$\forall$' and '$\exists$', the connective symbols '$\neg$', '$\bigvee$', '$\bigwedge$', and the signs '(' and ')'.

A first order formula is a formula written in graphs' first order logic. A first order property is a graph property that can be expressed by a first order formula where the variables x, y, z stand for vertices, '=' stands for equality and 'EDGE' stands for adjacency.

The quantifier depth $depth(\phi)$ of a formula ϕ is defined inductively:

1. For atomic expressions, $depth(x = y) = depth(\text{EDGE}(x, y)) = 0$.
2. $depth(\neg\phi) = depth((\phi)) = depth(\phi)$.
3. $depth(\phi \bigvee \psi) = depth(\phi \bigwedge \psi) = max\{depth(\phi), depth(\psi)\}$
4. $depth(\exists x\phi) = depth(\forall x\phi) = depth(\phi) + 1$.

For instance, the property of being either an empty graph or containing a triangle is a first order property that can be expressed by the following formula of quantifier depth 3: $(\forall u \forall v \neg \text{EDGE}(u, v)) \bigvee (\exists x \exists y \exists z\ (\text{EDGE}(x, y)) \bigwedge (\text{EDGE}(x, z)) \bigwedge (\text{EDGE}(y, z))$.

2.2 Distributions on Huge Graphs

Definition 1 (Distributions on Huge Graphs). *Let $\ell : \mathbb{N} \longrightarrow \mathbb{N}$ be a poly(n)-bounded length function. Distributions on huge graphs with vertex sets $\{V_n\}_{n\in\mathbb{N}}$, $V_n \subseteq \{0,1\}^{\ell(n)}$ are a sequence of distributions $\mathcal{G} = \{\mathcal{G}_n\}_{n\in\mathbb{N}}$, where each $\mathcal{G}_n$ is taken over the set of simple, labeled undirected graphs over V_n.*

For our probabilistic constructions the vertex sets are simply $V_n = \{0,1\}^n$. For our deterministic constructions the distributions $\mathcal{G}_n$ are degenerate (i.e., have support of size 1). We often abbreviate the term "distributions over huge graphs" and refer to "huge graphs" instead.

Definition 2 (Efficiently constructible huge graphs). *Let $\ell_1, \ell_2 : \mathbb{N} \longrightarrow \mathbb{N}$ be poly(n)-bounded length functions. Distributions on huge graphs $\mathcal{G} = \{\mathcal{G}_n\}_{n\in\mathbb{N}}$ with vertex sets $\{V_n\}_{n\in\mathbb{N}}$, $V_n \subseteq \{0,1\}^{\ell_1(n)}$ are efficiently constructible if there exists a deterministic polynomial-time evaluation algorithm E such that for all $n \in N$: for uniformly drawn $s \in \{0,1\}^{\ell_2(n)}$, the distribution of graph*

$$(V_n,\ \{(u, v) : E(s, u, v) = 1\})$$

is identical to $\mathcal{G}_n$.

Note that for our deterministic constructions, $\mathcal{G}_n$ is degenerate and $\ell_2(n) = 0$.

Definition 3 ($(p(n), K(n))$-wise independent graphs). *Let $p : \mathbb{N} \to [0,1]$, and $K : \mathbb{N} \to \mathbb{R}^+$. Huge graphs $\mathcal{G} = \{\mathcal{G}_n\}_{n\in\mathbb{N}}$ are $(p(n), K(n))$-wise independent if in $\mathcal{G}_n$ every potential edge appears w.p. $p(n)$, and the distribution of any $K(n)$ potential edges is mutually independent.*

Computational indistinguishability between distributions over huge graphs is defined exactly like (standard) computational indistinguishability between distributions over functions, with function evaluation replaced by graph edge queries. (For more details the reader may consult [13] [17].)

2.3 New Definitions: Generalized 0/1 Laws

Definition 4 ($D(n)$-0/1 law). *Let $\mathcal{G}$ be huge graphs, and let $D : \mathbb{N} \to \mathbb{N}$. The $D(n)$-0/1 law holds for $\mathcal{G}$ if for any sequence of formulas Φ having quantifier depth $depth(\phi_n) \leq D(n)$ the following conditions are satisfied:*

- *There exist a sequence $\Phi' = \{\phi'_n\}_{n \in \mathbb{N}}$, such that $\phi'_n \in \{\phi_n, \neg\phi_n\}$, and $\Pr[\mathcal{G}_n \models \phi'_n] \stackrel{n\to\infty}{\longrightarrow} 1$.*
- *For any single formula $\phi \in \Phi$ the limit $\lim_{n\to\infty} \Pr[\mathcal{G}_n \models \phi]$ exists.*

For the motivation of this definition, see Section 1.2. Note that meeting the basic 0/1 law is precisely the same as satisfying the $D(n)$-0/1 law for all $D(n) = \Theta(1)$.

Definition 5 ($D(n)$-equivalence of huge graphs). *Let $D : \mathbb{N} \to \mathbb{N}$. Two huge graphs $\mathcal{G}^1, \mathcal{G}^2$ are $D(n)$-equivalent if for any sequence of formulas Φ having quantifier depth $depth(\phi_n) \leq D(n)$ it holds that $\lim(\Pr[\mathcal{G}^1_n \models \phi_n] - \Pr[\mathcal{G}^2_n \models \phi_n]) \stackrel{n\to\infty}{\longrightarrow} 0$.*

3 Extension properties and 0/1 laws

We now describe extension properties, which were used by Fagin as a sufficient condition for his basic 0/1 law [9]. These extension properties will be used for two purposes: first, to establish the maximal depth, $D(n)$, for which $D(n)$-0/1 laws are satisfied by $G(N, p(n))$, and later, for proving $D(n)$-0/1 laws for efficiently constructed graphs.

Definition 6 (Extension Properties).

- *A single graph g maintains the t-extension property P^{EXT}_t if for all distinct vertices $v_1, ..., v_t$ and any bits $b_1, ..., b_t$ there exists an extending vertex $u \notin \{v_1, ..., v_t\}$ s.t. the edge $\{u, v_i\}$ appears in g iff $b_i = 1$.*
- *Let $T : \mathbb{N} \to \mathbb{N}$. A sequence of huge graphs $\mathcal{G} = \{\mathcal{G}_n\}_{n \in \mathbb{N}}$ achieves the $T(n)$-extension property if $\Pr[\mathcal{G}_n \models P^{EXT}_{T(n)}] \stackrel{n\to\infty}{\longrightarrow} 1$.*

We first state the sufficiency of $D(n)$-extension to $D(n)$-0/1 laws. We remark that although Spencer considers only the case of a single formula (rather then a sequence of formulas), the following Theorem is actually proved in Spencer's [19, Section 2.5]:

Theorem 1. *Let $\mathcal{G}$ be huge graphs, and let $D : \mathbb{N} \to \mathbb{N}$ be an arbitrary increasing function. If $\mathcal{G}$ achieves $D(n)$-extension, then $\mathcal{G}$ satisfies the $D(n)$-0/1 law.*

We next prove that any pair of huge graphs that achieve $D(n)$-extension are $D(n)$-equivalent.

Theorem 2. *Let $\mathcal{G}^1$ and $\mathcal{G}^2$ be huge graphs, and let $D : \mathbb{N} \to \mathbb{N}$ be an arbitrary increasing function. If both $\mathcal{G}^1$ and $\mathcal{G}^2$ achieve $D(n)$-extension, then $\mathcal{G}^1$ and $\mathcal{G}^2$ are $D(n)$-equivalent.*

Proof. Assume towards contradiction that $\mathcal{G}^1$ and $\mathcal{G}^2$ (as above) are not $D(n)$-equivalent. By Theorem 1, both $\mathcal{G}^1$ and $\mathcal{G}^2$ satisfy the $D(n)$-0/1 law. Therefore our negation assumption implies that there exist an infinite subset $N \subset \mathbb{N}$ and a sequence $\Phi = \{\phi_n\}_{n \in N}$ having quantifier depth $d_\Phi(n) \leq D(n)$, s.t. $\Pr[\mathcal{G}^1_n \models \phi_n] \overset{n \in N, n \to \infty}{\longrightarrow} 1$, whereas $\Pr[\mathcal{G}^2_n \models \phi_n] \overset{n \in N, n \to \infty}{\longrightarrow} 0$.

Consider a third distribution $\mathcal{G}^3 = \frac{1}{2}\mathcal{G}^1 + \frac{1}{2}\mathcal{G}^2$. Namely, we construct two graphs g_1, g_2 according to $\mathcal{G}^1, \mathcal{G}^2$ resp. and then toss a fair coin to choose the final graph $g_3 \in \{g_1, g_2\}$. We get $\Pr[\mathcal{G}^3_n \models \phi_n] \overset{n \in N, n \to \infty}{\longrightarrow} 1/2$. On the other hand, $\mathcal{G}^3$ clearly achieves the $T(n)$-extension property, so Theorem 1 implies that $\mathcal{G}^3$ satisfies the $D(n)$-0/1 law. This contradiction completes the proof. ∎

We next claim that the maximal extension achieved by $G(N, p(n))$ is approximately $\frac{n}{\lg(1/p(n))}$ (the proof is omitted in this preliminary version):

Theorem 3. *For arbitrary $p : \mathbb{N} \to (0,1)$, set $p'(n) = min\{p(n), 1-p(n)\}$, and let $T(n) = \frac{n - 2\lg n}{\lg(1/p'(n))}$. Then $G(2^n, p(n))$ achieves the $T(n)$-extension property, and does not achieve the $(1+\Delta)T(n)$-extension property for any constant $\Delta > 0$.*

An interesting consequence of Theorem 3 is that from the lens of first order logic, very sparse graphs and very dense graphs look the same. Formally, this is expressed by the fact that by Theorem 3 $G(N,p)$ and $G(N, 1-p)$ have the same extension. This coincides with the intuition that for $p < 1/2$, finding an extending vertex for the hardest requirement that "all edges must appear" is just as hard as finding an extending vertex for the requirement that "all edges must not appear" when $p' = 1 - p$. For instance, we get that depth-$\frac{n}{10}$ properties can not distinguish between $G(2^n, 0.001)$ and $G(2^n, 0.999)$.

Is the $D(n)$-extension property not only a sufficient but also a *necessary* condition for $D(n)$-0/1 laws? While for general graphs the answer is no (we can show examples where $2^{\omega(n)}$-0/1 laws are satisfied without achieving even 2-extension), we now show that for $G(N, p(n))$ the maximal extension and the maximal depth of 0/1 laws are roughly the same in the following sense: for any choice of $p(n)$ there exists $p'(n) \approx p(n)$ s.t. $G(N, p'(n))$ cannot achieve $D(n)$-0/1 laws for $D(n)$ larger than twice its maximal extension:

Theorem 4. *Let $p : \mathbb{N} \to (0,1)$, s.t. $\frac{1}{p(n)}$ and $\frac{1}{1-p(n)} = 2^{o(\sqrt{n})}$. Then there exists $p' : \mathbb{N} \to (0,1)$ where $p'(n) = p(n)(1 \pm o(1))$, s.t. $G(2^n, p'(n))$ defies the $D(n)$-0/1 law for $D(n) = (2 \pm o(1))\frac{n}{\lg(\frac{1}{p'(n)})}$.*

Proof. The claim will follow by presenting $p'(n)$ as above and a sequence of first-order formulas $\Phi = \{\phi_n\}_{n \in \mathbb{N}}$ having $depth(\phi_n) = (2 \pm o(1))\frac{n}{\lg(1/p'(n))}$ s.t.

1. For sufficiently large n, $1/4 \leq \Pr[G(2^n, p'(n)) \models \phi_n] \leq 3/4$.
2. The limit $\lim_{n \to \infty} \Pr[G(2^n, p'(n)) \models \phi_n]$ does *not* exist.

We use formulas ϕ_n that state the existence of a clique of size $\approx 2\frac{n}{\lg(1/p(n))}$ in the graph. We assume w.l.o.g. that $p(n) \leq 1/2$ (otherwise, let $p(n) > 1/2$, ϕ_n states the existence of independent sets that size).

By the classical analysis of Bollobás and Erdös concerning cliques in random graphs [5], there exists an integer $S^* = S^*(n, p(n)) = (2 - o(1))\frac{n}{\lg(1/p(n))}$ s.t. S^*-cliques appear in $G(2^n, p(n))$ almost surely. Namely, for $\phi_n = \exists v_1 ... v_{S^*} \bigwedge_{i \neq j}((v_i \neq v_j) \bigwedge \text{EDGE}(v_i, v_j))$, we have $\Pr[G(2^n, p(n)) \models \phi_n] = 1 - o(1)$.

Fix a sufficiently large n s.t. $\Pr[G(2^n, p(n)) = \phi_n] \geq 3/4$, and define H as follows:

$$H(q) = \Pr[G(2^n, q) \models \phi] =$$

$$\Sigma_{g \models \phi} \Pr[G(2^n, q) = g] = \Sigma_{g \models \phi} q^{E(g)}(1 - q)^{\binom{2^n}{2} - E(g)},$$

where $E(g)$ denotes the number of edges in g. Clearly, $H(\cdot)$ is continuous in q, and ϕ_n is a monotone property[4]. Thus, for any choice of $1/4 \leq \mu(n) \leq 3/4$ there exists (a unique) $p'(n) \leq p(n)$ s.t. $\Pr[G(2^n, p'(n)) = \phi_n] = \mu(n)$. In particular, we can take $\{\mu(n)\}_{n \in \mathbb{N}}$ s.t. the sequence has no limit. We thus get that $G(2^n, p'(n))$ defies the $2\frac{n}{\lg(1/p(n))}$-0/1 law.

We need to prove that the $\approx 2\frac{n}{\lg(1/p'(n))}$ -0/1 law is also defied by $G(2^n, p'(n))$, so to complete the entire proof we will show that:

1. $\lg(1/p(n)) = \lg(1/p'(n))(1 \pm o(1))$, and
2. $p'(n) = p(n)(1 - o(1))$.

Indeed, fix n so $p = p(n), p' = p'(n), \mu = \mu(n)$ and $\phi = \phi_n$, and let $\delta = \delta(n)$ be defined by $p' = p(1 - \delta)$. Let $\mathbb{E}_{S,n,q}$ denote the expected number of S-cliques in $G(2^n, q)$. Again, by [5] $\mathbb{E}_{S^*,n,p} \leq 2^{(2+o(1))n}$. Next, Markov's inequality gives:

$$1/4 \leq \mu \stackrel{\text{def}}{=} \Pr[G(2^n, p') = \phi] \leq \mathbb{E}_{S^*,n,p'} = \binom{2^n}{S^*} \times p'^{\binom{S^*}{2}} = \binom{2^n}{S^*} p^{\binom{S^*}{2}} (1 - \delta)^{\binom{S^*}{2}}$$

$$\leq 2^{(2+o(1))n}(1 - \delta)^{\binom{S^*}{2}} = 2^{(2+o(1))n} e^{-\Theta(\delta(S^*)^2)} = 2^{(2+o(1))n} e^{-\Theta(\delta(\frac{n}{\lg(\frac{1}{p})})^2)}$$

Thus $\delta(n) = o(1)$ *iff* $(\lg \frac{1}{p})^2 = o(n)$ but the latter condition is met since the conditions of the theorem include $\frac{1}{p} = 2^{o(\sqrt{n})}$. This proves that $p'(n) \stackrel{\text{def}}{=} p(n)(1 - \delta(n)) = p(n)(1 - o(1))$.

Finally, as $0 < \delta \leq 1/2$ we get $\frac{1}{1-\delta} = 1 + \frac{\delta}{1-\delta} \leq 1 + 2\delta \leq e^{2\delta}$. Consequently,

$$\frac{\lg \frac{1}{p'}}{\lg(1/p)} = \frac{\lg \frac{1}{p(1-\delta)}}{\lg(1/p)} = \frac{\lg \frac{1}{p} + \lg \frac{1}{1-\delta}}{\lg(1/p)} \leq 1 + \frac{\lg e^{2\delta}}{\lg(1/p)} = 1 + \Theta(\frac{\delta}{\lg(1/p)}) = 1 + o(1),$$

since $\delta(n) = o(1)$. The claim follows. ■

An immediate corollary of Theorems 1, 3, and 4 is that the maximal depth $D^*(n)$ for which $G(N, p)$ satisfies $D^*(n)$-0/1 laws is $\Theta(\frac{n}{\lg(\frac{1}{p})})$:

Theorem 5. *Let $p : \mathbb{N} \to (0, 1)$. Then*

1. *$G(2^n, p(n))$ satisfies the $[\frac{n - 2\lg n}{\lg(1/p(n))}]$-0/1 law.*

[4] Namely, if $g \models \phi$ and g' is obtained by adding edges to g, then $g' \models \phi$ as well.

2. *If* $\frac{1}{p(n)}, \frac{1}{1-p(n)} \leq 2^{o(\sqrt{n})}$*, then there exists* $p' : \mathbb{N} \to (0,1)$ *s.t. ,* $p'(n) = p(n)(1 \pm o(1))$*, and* $G(2^n, p'(n))$ *defies the* $[\frac{2n}{\lg(1/p'(n))}]$*-0/1 law.*

In light of Theorem 5, our aim becomes to efficiently construct graphs that satisfy $\Theta(\frac{n}{\lg(\frac{1}{p})})$-0/1 laws and are $\Theta(\frac{n}{\lg(\frac{1}{p})})$-equivalent to $G(N,p)$.

4 Computational and k-wise Independent Graphs and Equivalence

Given the tight relationship between extensions and first-order graph properties, constructing computational and k-wise independent graphs satisfying the 0/1-laws is simple. The next theorem shows that n^3-wise independence in graphs guarantees the optimal $\frac{n}{\lg(\frac{1}{p})}$-0/1 laws and thus $\frac{n}{\lg(\frac{1}{p})}$-equivalence to random graphs.

Theorem 6. *Let* $p : \mathbb{N} \to (0,1)$*, and set* $p'(n) = min\{p(n), 1-p(n)\}$*. Let* $D(n) = \frac{\frac{n}{2} - 2\lg n}{\lg(1/p'(n))}$ *and let* $K(n) = 2nD^2(n)$*. Let* $\mathcal{G}$ *be* $(p(n), K(n))$*-wise independent huge graphs (see definition 3). Then* $\mathcal{G}$ *satisfies the* $D(n)$*-0/1 law and is* $D(n)$*-equivalent to* $G(2^n, p(n))$*.*

The proof is via the extension property and we omit it in this preliminary version. Recall that for arbitrary $p(n)$, one can construct (based on [13, 17]), $poly(n)$-wise independent graphs that are also computationally pseudo-random w.r.t. $G(2^n, p(n))$. Combining this with Theorem 6 one can show the following:

Theorem 7. *Let* $c > 0, p : \mathbb{N} \to [0,1]$*. Then there exist an explicit efficient construction of huge graphs* $\mathcal{G}$ *that for some* $D(n) = \frac{n}{\lg(1/p(n))}(1-o(1))$ *are:*

1. $(p'(n), n^c)$*-wise independent for some* $p'(n)$ *s.t.* $|p'(n) - p(n)| \leq 2^{-3n}$*.*
2. *Satisfy the* $D(n)$*-0/1 law and are* $D(n)$ *equivalent to* $G(2^n, p(n))$*.*
3. *Computationally indistinguishable from* $G(2^n, p(n))$ *if one-way functions exist.*

5 A Single Graph Equivalent to Random Graphs

In this section we demonstrate a single huge graph (for each size) that is deterministically constructible and "behaves like $G(N, p(n))$": the sequence is $D(n)$-equivalent to $G(N, p(n))$ and have edge density $p(n) \pm \epsilon$. The construction is based on Paley graphs, which are known to preserve a variety of properties of random graphs [2]. We employ the following generalized definition:

Definition 7 (Paley graph). *Let* $\mathcal{F}$ *be a finite field of size* N*, let* $M \in \mathbb{N}$ *such that* $2M \mid (N-1)$*, and let* $p \in \left\{\frac{1}{M}, \frac{2}{M}, \ldots, \frac{M-1}{M}\right\}$*. Let* $Z \subset \left\{a \in \mathcal{F} : a^M = 1\right\}$ *with* $|Z| = pM$*. Then the Paley graph* $G_{\mathcal{F},M,p,Z} = (\mathcal{F}, E_{\mathcal{F},M,p,Z})$ *is given by*

$$E_{\mathcal{F},M,p,Z} = \left\{\{u,v\} :\ u,v \in \mathcal{F},\ (u-v)^{(N-1)/M} \in Z\right\} \tag{2}$$

It is readily verified that every node has exactly $p(N-1)$ neighbors, and that the graph is undirected since the exponent in (2) is even.

The rest of this section is structured as follows. First, as a technical aid we define sets of linear equalities that contain certificates to "$x \not\equiv 0 \pmod{M}$", and observe that for certain M these sets can be small. Then, we show that the $D(n)$-0/1 properties of a Paley graph $G_{\mathcal{F},M,p,Z}$ are related to the size of the smallest such certifying set for M. Next, we show that for appropriate parameters we can efficiently compute edge queries in $G_{\mathcal{F},M,p,Z}$. Finally, we describe two concrete sequences of Paley graphs, and invoke the aforementioned lemmas to derive their efficient computability and $D(n)$-0/1 properties.

Definition 8 (nonzero-certifying set). *A set $C \subset \mathbb{N}\times\mathbb{Z}$ is* nonzero-certifying modulo M *if $\sum_{(y,z)\in C} y < M$ and for all $x \in \mathbb{Z}$:*

$$x \not\equiv 0 \pmod{M} \qquad \textit{iff} \qquad \exists (y,z) \in C:\ y_j x \equiv z_j \pmod{M} \tag{3}$$

For example, for any $M \in \mathbb{N}$ the set $\{(1,r)\}_{r\in\{1,\ldots,M-1\}}$ is nonzero-certifying modulo M. Smaller sets can be obtained by the following:

Lemma 1. *Let $M = q_1^{e_1} q_2^{e_2} \cdots q_s^{e_s}$ for distinct primes q_i and $e_i \in \mathbb{N}$. Then there exists a set C which is nonzero-certifying modulo M and $|C| = \sum_{t=1}^{s} e_t(q_t-1)$.*

Proof (sketch). Denote $\pi_t = \prod_{t'=t+1}^{s} q_{t'}^{e_{t'}}$, and set $C = \left\{ \left(\pi_t q_t^i\ ,\ \pi_t (M/q_t) r \right) \right\}_{t,i,r}$ where $t \in \{1,\ldots,s\}$, $i \in \{0,\ldots,e_t-1\}$, $r \in \{1,\ldots,q_t-1\}$. Then $|C| = \sum_{t=1}^{s} e_t(q_t-1) < \sum_{t=1}^{s} (\lg q_t^{e_t})(B-1) = (B-1)\lg M$, and it is readily verified that $\sum_{(y,z)\in C} y = M-1$. To show that (3) indeed holds, show that it holds modulo each $q_t^{e_t}$ by considering the q_t-ary representation of $z \bmod q_t^{e_t}$; then apply the Chinese Remainder Theorem.[5] ■

The next lemma shows that Paley graphs satisfy $D(n)$-0/1 laws with $D(n)$ that is related to the size of nonzero-certifying sets. The analysis follows Graham and Spencer's proof of the connection between similar Paley graphs (restricted to $M = 2$) and tournament problems [14][3]. Recall that for a finite field $\mathcal{F}$, a character $\chi : \mathcal{F} \to \mathbb{C}$ of order M is a multiplicative homomorphism from $\mathcal{F}^*$ onto the M-th roots of unity, extended with $\chi(0) = 0$; such χ exist whenever $M \mid (N-1)$. We will invoke Weil's theorem:

Theorem 8 (Weil). *Let $\mathcal{F}$ be a finite field, let $N = |\mathcal{F}|$, and let χ be a character of order M. Let $f(x) \in \mathcal{F}[x]$ be a monic polynomial that is not an M-th root of any polynomial in $\mathcal{F}[x]$. Then:*

$$\left| \sum_{u\in\mathcal{F}} \chi(f(u)) \right| < (\deg F - 1)\sqrt{N}$$

[5] Essentially, we are forming a system of linear equations which expresses a special case of the additive analogue of the Pohlig-Hellman-Silver algorithm [18].

Lemma 2. *Let $\mathcal{G} = \{G_{\mathcal{F},M,p,Z}\}_n$ be a sequence of Paley graphs with $\mathcal{F} = \mathcal{F}(n)$, $M = M(n)$, $p = p(n)$, $Z = Z(n)$, $N = |\mathcal{F}(n)|$ such that $N > M^{\omega(1)}$. Let $\ell = \ell(n)$, and suppose that for every n there exist a set of size ℓ which is nonzero-certifying modulo M. Then $\mathcal{G}$ satisfies the $D(n)$-0/1 law for $D(n) = \frac{\lg N}{2\ell}(1-o(1))$.*

Proof. By Theorem 11, it suffices to show that $G_{\mathcal{F},M,p,Z}$ satisfies the $D(n)$-extension law. Denote $d = D(n)$, $\ell = \ell(n)$. Let $C = \{(y_j, z_j)\}_{j=1}^{\ell}$ be nonzero-certifying modulo M , and let $\chi : \mathcal{F} \to \mathbb{C}$ be a character of order M.

Let $v_1, \ldots, v_d \in \mathcal{F}$ be arbitrary vertices, and let $b_1, \ldots, b_d \in \{0, 1\}$. We wish to show that there exists an extending vertex $u \in \mathcal{F} \setminus \{v_1, \ldots, v_d\}$ such that $\{u, v_i\} \in E_{\mathcal{F},M,p,Z}$ iff $b_i = 1$ for all $i = 1, \ldots, d$. Let $w_1, \ldots, w_d \in \mathcal{F}$ be chosen arbitrarily subject to $w_i^{(N-1)/M} \in Z$ iff $b_i = 1$, for $i = 1, \ldots d$. Then by definition of $E_{\mathcal{F},M,p,Z}$, it suffices to show that there exists a vertex $u \notin \{v_1, \ldots, v_d\}$ such that $(u - v_i)^{(N-1)/M} = w_i^{(N-1)/M}$ for all i. This further reduces to $\chi(u - v_i) = \chi(w_i)$, since in this case $\mu_i = (u - v_i)/w_i$ is in $\mathrm{Ker}_\chi = \chi^{-1}(1)$ so the order of μ_i divides $|\mathrm{Ker}_\chi| = (N - 1)/M$, whence $(u - v_i)^{(N-1)/M}/w_i^{(N-1)/M} = \mu_i^{(N-1)/M} = 1$.

It thus suffices to show that there exists $u \in \mathcal{F}$ such that $\chi(u - v_i) = \chi(w_i)$ for all i. Let α be a generator of $\mathcal{F}^*$, and denote:

$$h(u) = \prod_{i=1}^{d} h_i(u) \qquad \text{where} \qquad h_i(u) = \prod_{j=1}^{\ell} \left(1 - \frac{\chi(u - v_i)^{y_j}}{\chi(w_i^{y_j}\alpha^{z_j})}\right) \quad (i = 1, \ldots, d)$$

Note that $h_i(u) = 0$ iff there exists $j \in \{1, \ldots, \ell\}$ such that $\chi(u-v_i)^{y_j}/\chi(w_i^{y_j}\alpha^{z_j}) = 1$. Since $\chi(\alpha)$ is a generator of the multiplicative group of M-th roots of unity in $\mathbb{C}$, which has order M, for $u \neq v_i$ we can take discrete logs to base $\chi(\alpha)$. Then:

$$h_i(u) = 0 \qquad \text{iff} \qquad \exists j \in \{1, \ldots, \ell\} : y_j \log_{\chi(\alpha)}((u - v_i)/w_i) \equiv z_j \pmod{M}$$

Since C_n is nonzero-certifying modulo M, by considering $x = \log_{\chi(a)}((u - v_i)/w_i)$ we get that $h_i(u) = 0$ iff $x \equiv 0 \pmod{M}$, i.e., iff $\chi(u - v_i) \neq \chi(w_i)$. Our task is thus reduced to showing the existence of an "extending vertex" $u \in \mathcal{F} \setminus \{v_1, \ldots, v_d\}$ such that $h(u) \neq 0$.

Denote $S = \sum_{u\in\mathcal{F}} h(u)$. By the triangle inequality:

$$|S| \le \sum_{\substack{u\in\mathcal{F}\\ h(u)\neq 0}} \prod_{i=1}^{d}\prod_{j=1}^{\ell} \left(1 + \left|\frac{\chi(u - v_i)^{y_j}}{\chi(w_i^{y_j}\alpha^{z_j})}\right|\right) \le \sum_{\substack{u\in\mathcal{F}\\ h(u)\neq 0}} 2^{d\ell} = d2^{d\ell} + \sum_{\substack{u\in\mathcal{F}\setminus\{v_1,\ldots,v_d\}\\ h(u)\neq 0}} 2^{d\ell} \tag{4}$$

Thus, if $|S| > d2^{d\ell}$ then there exists an extending vertex. To lower bound $|S|$, we first expand the product over i and j. Denote $\mathcal{I} = \{1, \ldots, d\} \times \{1, \ldots \ell\}$. Then:

$$S = \sum_{u\in\mathcal{F}} \prod_{i=1}^{d}\prod_{j=1}^{\ell} \left(1 + \frac{\chi(u - v_i)^{y_j}}{-\chi(w_i^{y_j}\alpha^{z_j})}\right) = \sum_{u\in\mathcal{F}} \sum_{I\subseteq\mathcal{I}} \prod_{(i,j)\in I} \frac{\chi(u - v_i)^{y_j}}{-\chi(w_i^{y_j}\alpha^{z_j})}$$

$$= \sum_{u\in\mathcal{F}} \sum_{I\subseteq\mathcal{I}} P_I \left(\prod_{(i,j)\in I} \chi(u - v_i)^{y_j}\right) \qquad \text{where} \quad P_I = \prod_{(i,j)\in I} \frac{1}{-\chi(w_i^{y_j}\alpha^{z_j})}$$

By separating the case $I = \emptyset$ and, changing order of summation and using the multiplicativity of χ, we then obtain:

$$S = N + \sum_{\substack{I \subseteq \mathcal{I} \\ I \neq \emptyset}} P_I \sum_{u \in \mathcal{F}} \chi(f_I(u)) \qquad \text{where} \quad f_I(u) = \prod_{(i,j) \in I} (u - v_i)^{y_j}$$

For all $I \subseteq \mathcal{I}$ with $I \neq \emptyset$, $f_I(u)$ has at least one root v_i and the multiplicity of any root v_i is at most $\sum_{j=1}^{\ell} y_i < M$ by Definition 8, so $f_I(u)$ is not an M-th power of any polynomial in $\mathcal{F}[u]$. Also, $\deg f_I \leq d(M-1)$. Invoking Weil's theorem, we obtain for all such I:

$$\left| \sum_{u \in \mathcal{F}} \chi(f_I(u)) \right| \leq (d(M-1)-1)\sqrt{N}$$

Then by the triangle inequality,

$$|S| \geq N - \sum_{\substack{I \subseteq \mathcal{I} \\ I \neq \emptyset}} P_I \left| \sum_{u \in \mathcal{F}} \chi(f_I(u)) \right| > N - 2^{d\ell} d(M-1)\sqrt{N}$$

By (4), there remains to show that $2^{d\ell} d \geq N - 2^{d\ell} d(M-1)\sqrt{N}$. Indeed:

$$\left(N - 2^{d\ell} d(M-1)\sqrt{N}\right) - 2^{d\ell} d \geq \sqrt{N}\left(\sqrt{N} - 2^{d\ell} dM\right)$$

and the latter is greater than 0 when $\lg N > 2\,(d\ell + \lg d + \lg M)$, i.e., when $d > \frac{\lg N - 2\lg M}{(2+o(1))\ell} > \frac{\lg N - 2\lg N/\omega(1)}{(2+o(1))\ell} = \frac{\lg N}{2\ell}(1-o(1))$. ∎

Remark 1. Since the choice $w_1, \ldots, w_d \in \mathcal{F}$ in the above proof was arbitrary, we have actually shown a stronger result: for the same parameters as in Lemma 2, there exists an edge labeling $L : \mathcal{F} \times \mathcal{F} \to \{1, \ldots, M\}$ of the full graph of size N, such that for any d vertices $v_1, \ldots, v_d$ and labels $a_1, \ldots a_d$ there exists a vertex $u \in \mathcal{F} \setminus \{v_1, \ldots, v_d\}$ such that $L(u, v_i) = a_i$ for all $i = 1, \ldots, d$. ∎

Recall that $M \in \mathbb{N}$ is called B-smooth if no prime divisor of M is larger than B.

Corollary 1 *Let $\mathcal{G} = \{G_{\mathcal{F},M,p,Z}\}_n$ be a sequence of Paley graphs with $\mathcal{F} = \mathcal{F}(n)$, $M = M(n)$, $p = p(n)$, $Z = Z(n)$, $N = |\mathcal{F}(n)|$ such that $N > M^{\omega(1)}$ and M is B-smooth for $B = B(n)$. Then $\mathcal{G}$ satisfies the $D(n)$-0/1 law for*

$$D(n) = \frac{\lg N}{2(B-1)\lg M}(1-o(1))$$

Proof. Let $M = q_1^{e_1} q_2^{e_2} \cdots q_s^{e_s}$ for distinct primes $q_i \leq B$ and $e_i \in \mathbb{N}$. Then by Lemma 1, there exists a set C which is nonzero-certifying modulo M and $\ell(n) = |C| = \sum_{t=1}^{s} e_t(q_t - 1) < \sum_{t=1}^{s} (\lg q_t^{e_t})(B-1) = (B-1)\lg M$. The claim follows by Lemma 2. ∎

We now address the issue of efficient computability. The following lemma shows that there are sequences of Paley graphs in which edge queries can be computed efficiently, under constraints which will be addressed by the concrete sequences described later.

Lemma 3. *There exists a deterministic algorithm A which, for any $\mathcal{F}$, N, M and p as in Definition 7, evaluates edge queries in a Paley graph $G_{\mathcal{F},M,p,Z}$ in the following sense: given an oracle $\mathcal{O}_{\mathcal{F}}$ which computes the basic operations in $\mathcal{F}$, and given an element $g \in \mathcal{F}$ of order M in $\mathcal{F}^*$, there exists Z as in Definition 7 such that $A^{\mathcal{O}_{\mathcal{F}}}(N, M, p, g, u, v) = 1$ iff $(u, v) \in E_{\mathcal{F},M,p,Z}$. Moreover, if M is B-smooth then A runs in time poly$(\log N, B)$.*

Proof. Note that $\langle g \rangle = \{a \in \mathcal{F} : a^M = 1\}$, and set $Z = \{a \in \langle g \rangle : \log_g a < pM\}$. For $u \neq v$, to test whether $a = (u-v)^{(N-1)/M}$ fulfills $a \in Z$, it suffices to compute discrete logarithms in the group $\langle g \rangle$, whose order is B-smooth. This can be done deterministically in time *poly*$(\log N, B, |C_{\mathcal{F}}|)$ using the Pohlig-Hellman-Silver algorithm [18]. ■

We can now proceed to describe two specific efficiently computable huge graphs based on sequences of Paley graphs. As we have seen, it suffices to find a deterministically computable sequence of pairs (N, M) such that N is a prime power, $2M|(N-1)$, M is highly smooth, and we can deterministically find an efficient representation of the finite field $\mathcal{F} = \mathrm{GF}(N)$ and an element $g \in \mathcal{F}^*$ of order M. Moreover, we wish the sequence to be dense: for every $n \in \mathbb{N}$ there should be (N, M) fulfilling $M = 2^{\Theta(n)}$.

Recall the following results about finite fields, from [20] and [21].

Theorem 9 (Shoup). (a) *Let q be prime and $m \in \mathbb{N}$. Then there exists a deterministic algorithm that computes an irreducible polynomial $I(X)$ of degree m in $\mathrm{GF}(q)[X]$ in time poly(q, m).* **(b)** *Let $I(X)$ be any an irreducible polynomial of degree m in $\mathrm{GF}(q)[X]$, and let $\mathcal{F} = \mathrm{GF}(q)[X]/(I(X))$. There exists a deterministic algorithm which, given $I(X)$, runs in time poly(q, m) and outputs a set of elements in $\mathcal{F}$ which contains at least one generator of $\mathcal{F}^*$.*

The following is an explicit construction which approximates any desired edge density $p(n)$ up to an additive term of $\epsilon(n) < \Theta(1/n)$, and achieves $D(n)$ which is optimal up to a constant. Here, we choose N and M using Euler's theorem.

Theorem 10. *Let $p = p(n) \in (0, 1)$ and let $\epsilon = \epsilon(n) \geq c_0/n$ for a certain constant $c_0 > 0$. Then there exists a deterministically efficiently computable huge graph $\mathcal{G} = \{g_n\}_n$ which satisfies the $D(n)$-0/1 law for $D(n) = \frac{n}{2\log(1/\epsilon)}(1-o(1))$, and g_n has size $2^{\theta(n)}$ and edge density $p'(n)$ such that $|p'(n) - p(n)| < \epsilon(n)$.*

Proof. Set $c_0 = 2\lg 3$. Let $N = 3^{n'}$ where $n' = 2^k$ and $k = \lceil \lg(n/\lg 3) \rceil$. Let $M = 2^{\lceil \lg(1/\epsilon) \rceil}$. Note that $2^n < N \leq 2^{2n}$, and that $M < 2^{\lg(1/\epsilon)+1} < 2^{\lg(n/2\lg 3)+1} = 2^{\lg(n/\lg 3)} \leq n'$, so $M \mid n'$. Since 3 is relatively prime to $2n'$, Euler's theorem yields $3^{\varphi(2n')} \equiv 1 \pmod{2n'}$, where $\varphi(2n') = n'$. Hence $2M \mid (N-1)$. We have $\epsilon/2 < \frac{1}{M} \leq \epsilon$, and can choose $p'(n) \in \{\frac{1}{M}, \frac{2}{M}, \ldots, \frac{M-1}{M}\}$ such that $|p'(n) - p(n)| \leq \frac{1}{M} \leq \epsilon$.

By Theorem 9(a), we can deterministically compute an irreducible polynomial of degree n' in $\mathrm{GF}(3)[X]$ in time *poly*$(n') =$ *poly*(n), and can thus efficiently calculate in the field $\mathcal{F} = \mathrm{GF}(3^{n'})$.[6] To deterministically find an element of order M in time *poly*(n), run the algorithm of Theorem 9(b) and, for each output element β, directly test whether $\gamma = \beta^{(N-1)/S}$ has order M by computing the first M powers of γ. Note that when β generates $\mathcal{F}^*$, γ indeed has order M.

By the above and Lemma 3 there exists a set Z such that $G_{\mathcal{F},M,p,Z}$ is a Paley graph whose edge queries can be computed deterministically in time *poly*$(\log N)$ = *poly*(n). Then $\mathcal{G} = \left\{G_{\mathcal{F}(n),M(n),p'(n),Z(n)}\right\}_n$ is a deterministically efficiently computable huge graph with density $p' = p \pm \epsilon$. Since M is 2-smooth, by Corollary 1 $\mathcal{G}$ satisfies the $D(n)$-0/1 law for $D(n) = \frac{\lg N}{2 \lg M}(1-o(1)) \geq \frac{n}{2\lg(1/\epsilon)}(1-o(1))$. ∎

The above allows only $\epsilon(n) > \Theta(1/n)$, which means we cannot meaningfully approximate graphs with density $p \ll 1/n$. To enable better approximation ϵ, and also to obtain N closer to 2^n (albeit at some cost in the extension $D(n)$), we will replace Euler's totient function $\varphi(\cdot)$ with Carmichael's function $\lambda(\cdot)$, which likewise satisfies that $b^{\lambda(a)} \equiv 1 \pmod{a}$ for any relatively prime $a, b \in \mathbb{N}$. The benefit is that $\lambda(a)$ occasionally assumes much smaller values than $\varphi(a)$ (cf. [8]). For square-free $a \in \mathbb{N}$, $\lambda(a) = \mathrm{lcm}\{q-1 : q \text{ prime}, q \mid a\}$. For $b \in \mathbb{N}$, let $\eta(b) = \prod_{q \text{ prime},\ q-1|b} q$. Note that $\lambda(\eta(b)) = b$. Then by [1]:

Theorem 11 (Pomerance, Odlyzko). *There exists a constant $c_1 > 0$ such that for all sufficiently large A, there exists $b < (\ln A)^{c_1 \ln\ln\ln A}$ s.t. $\eta(b) > A$.*

Theorem 12. *Let $p = p(n) \in (0,1)$ and let $\epsilon > 2^{-n^{1/c_2 \ln\ln n}}$ for a constant $c_2 > 0$. Then there exists a deterministically efficiently computable huge graph $\mathcal{G} = \{g_n\}_n$ which satisfies the $D(n)$-0/1 law for $D(n) = n/\log(1/\epsilon)^{\Theta(\log\log\log(1/\epsilon))}$, and g_n has size $2^{n(1+o(1))}$ and edge density $p'(n)$ such that $|p'(n) - p(n)| < \epsilon(n)$.*

Proof. We first find appropriate N, M. Let $B = (\ln(6/\epsilon))^{c_1 \ln\ln\ln(6/\epsilon)}$. Then by Theorem 11, for sufficiently large n there exists $b < B$ such that $\eta(b) > 6/\epsilon$. We can deterministically find such b by exhaustive search in time *poly*$(B) <$ *poly*(n). Fix any c_2 larger than c_1. It is readily verified that $\log(6/\epsilon)^{c_1 \ln\ln n} < n/\sqrt{\ln n}$ for sufficiently large n, and since $n > \ln(6/\epsilon)$ we get $B < n/\sqrt{\lg n}$ and thus $b < n/\sqrt{\ln n} = o(n)$. Let n' be the smallest multiple of b that is larger than n, and let $N = 3^{n'}$. Then $2^n \leq N \leq 2^{n(1+o(1))}$.

Let $M = \prod_{\text{prime } q|\eta(b),\ q>\kappa} q$ where κ is the largest such that $M \geq 1/\epsilon$. Note that $M \mid \eta(b)$ and $2 \mid \eta(b)$ but $2 \nmid M$, so $2M \mid \eta(b)$, and and from the definition of λ we get $\lambda(2M) \mid \lambda(\eta(b)) = b$. Thus $\lambda(2M) \mid n'$, and since $3 \nmid M$ we get $3^{2^r b} \equiv 1 \pmod{2M}$, i.e., $2M | (N-1)$. Also note that all prime factors of M are at most $b+1$, so M is $(B+1)$-smooth and $M < (B+1)/\epsilon = (1/\epsilon)^{1+o(1)}$. Since $\frac{1}{M} \leq \epsilon$, we can choose $p'(n) \in \left\{\frac{1}{M}, \frac{2}{M}, \ldots, \frac{M-1}{M}\right\}$ such that $|p'(n) - p(n)| \leq \frac{1}{M} \leq \epsilon$.

Conclude as in Theorem 10, with two differences. First, to test whether $\gamma = \beta^{(N-1)/M}$ is of order M, use the fact that M is $(B+1)$-smooth and square-free: by the Chinese Remainder, γ has order M iff $\gamma^{S/q} \neq 1$ (and thus $\gamma^{M/q}$ has order q) for every prime $q \mid M$, and this can be checked in time *poly*$(B \lg M)$ = *poly*(n). Second, M is $(B+1)$-smooth so we get $D(n) = \frac{n}{2B \lg M}(1-o(1)) = n/B^{1+o(1)}$. ∎

---

[6] Alternatively replace 5 with 3, and by [10], $X^{2^k} - 2$ is irreducible in $\mathrm{GF}(3)[X]$.

6 The Limits of Small Families

We now argue that no small and efficient family can be $D(n)$-equivalent to $G(2^n, 1/2)$ once $D(n)$ is an arbitrary polynomial in n. We can generalize the theorem to hold for various choices of p.

Theorem 13. *Let $\mathcal{G}$ be an efficiently constructed distribution on huge graphs with seed length $m(n)$, and let $D : \mathbb{N} \to \mathbb{N}$, s.t. $D(n) = 2m(n) + \omega(\sqrt{m(n)} + n)n$. Then $\mathcal{G}$ is* not *$D(n)$-equivalent to $G(2^n, 1/2)$.*

Proof. Intuitively, the theorem stems from the fact that any efficiently constructed graph has a low Kolmogorov complexity (KC), whereas random graphs exhibit a high KC. The claim will follow once we provide a sequence of separating formulas $\Phi = \{\phi_n\}_{n\in\mathbb{N}}$ which have depth $depth(\phi_n) = D(n)$ and length $|\phi_n| = n^{\Theta(1)}$ s.t. $\Pr[\mathcal{G}_n \models \phi_n] = 1$, but $\Pr[G(2^n, \frac{1}{2}) \models \phi_n] \stackrel{n\to\infty}{\longrightarrow} 0$.

Fix n and let $m = m(n), d = D(n)$. Let E be the evaluating algorithm of $\mathcal{G}$. Namely, to each graph g in the support of $\mathcal{G}$ there corresponds a seed $s = s(g) \in \{0,1\}^m$ s.t. for any vertex pair $u, v \in \{0,1\}^n$, it holds that $E(s, u, v) = 1$ when the edge $\{u, v\}$ appears in g, and $E(s, u, v) = 0$ otherwise. The standard reduction from Turing machines to Boolean circuits implies the existence of a *poly*(n)-size Boolean formula $\psi_{E,n}$ s.t. $\psi_{E,n}(s, u, v) = E(s, u, v)$ for all inputs s, u, v of appropriate length.

We wish the separating formulas to hold for a graph g *iff* g is in the support of $\mathcal{G}$, namely, when there exists a seed $s = s(g)$ s.t. all the edges of g are correctly evaluated by $\psi_{E,n}(s, \cdot, \cdot)$. However, to reduce the quantifier depth of ϕ_n, we only attempt that ϕ_n expresses the following condition where $r = r(n)$ is specified later:

Condition 1 *Every subgraph on r vertices $v_1, ..., v_r$ is isomorphic to some subgraph correctly evaluated by $\psi_{E,n}$ using some seed $s \in \{0,1\}^m$.*

Condition 1 can be expressed as follows (here $u_{i_1}...u_{i_n}$ denote the bits of a vertex $u_i \in \{0,1\}^n$):

$$\forall v_1, v_2 \ldots v_r \exists u_{1_1} \ldots u_{1_n}, \ldots, u_{r_1} ... u_{r_n} \exists s = s_1 ... s_m$$
$$\bigwedge_{i \neq j} \text{EDGE}(v_i, v_j) \Leftrightarrow \psi_{E,n}(s_1...s_m, u_{1_1}...u_{1_n}, u_{r_1}...u_{r_n}).$$

This expression is, however, not a first-order sentence on graphs. In first-order language the variables stand for vertices, whereas in the above expression $\psi_{E,n}(s_1...s_m, u_{1_1}...u_{1_n}, u_{r_1}...u_{r_n})$ actually refers to the bits s_i and u_{i_j}. This is resolved as each bit can be encoded using a single edge (or non-edge). Indeed, a string $x = x_1...x_\ell \in \{0,1\}^\ell$ is encoded using 2ℓ (not necessarily distinct) vertices $\bar{x}_1, ..., \bar{x}_\ell, \bar{x}'_1, ..., \bar{x}'_\ell$ s.t. $Enc(x) = \text{EDGE}(\bar{x}_1, \bar{x}'_1)...\text{EDGE}(\bar{x}_\ell, \bar{x}'_\ell)$. Note that for any string x, a valid encoding exists as long as the graph contains both edges and non-edges.[7]

[7] Note that one cannot write a first order expression that states the validity of the encoding, namely, the requirement that indeed $\text{EDGE}(\bar{x}_i, \bar{x}'_i) = x_i$. Yet, this encoding will suffice for our needs.

We recall that all the encodings in ψ_n are valid as long as the graph is neither the complete nor the empty graph. Thus we define the separating formula ϕ_n by $\phi_n = \psi_n \bigvee \gamma \bigvee \gamma'$, where γ, γ' are two fixed formulas which state that the graph is either complete or empty ($\gamma = \forall u, v(u \neq v \Rightarrow \text{EDGE}(u,v))$, and $\gamma' = \forall u, v(\neg\text{EDGE}(u,v)))$.

We finally prove that ϕ_n indeed separates $\mathcal{G}_n$ from $G(2^n, \frac{1}{2})$. We first note that $\Pr[\mathcal{G}_n \models \phi_n] = 1$. Indeed for any single graph g in the support of $\mathcal{G}_n$, if the graph is either complete or empty we are done. Otherwise, each vertex in g has a valid encoding. Since all the encodings in ψ_n are valid, clearly $g \models \psi_n$.

On the other hand $G(2^n, \frac{1}{2})$ is complete or empty with only vanishing probability. Hence it suffices to show that w.h.p. $G(2^n, \frac{1}{2}) \not\models \psi_n$. Indeed assume for a fixed graph g, that $g \models \psi_n$. This implies that for any subgraph on r vertices g_r of g the following holds: there exist strings $\bar{s} \in \{0,1\}^m$, and $\bar{v}_i \in \{0,1\}^n, i = 1, ..., r$ s.t. when the evaluator $E = E_{\mathcal{G}}$ is given all $\binom{r}{2}$ inputs in lexicographic order, then $E(\bar{s}, \bar{v}_i, \bar{v}_j)$ is exactly the adjacency string of g_r. In particular this implies that g_r has Kolmogorov complexity $\text{KC}(g_r) \leq m + rn + \Theta(1)$. Since with overwhelming probability a r-subgraph of a random graph has $\text{KC}(g_r) \geq \Omega(r^2)$, we get that ψ_n rarely holds for random graphs when $m + rn \leq o(r^2)$, namely when we set $r = \omega(\sqrt{m} + n)$. As the depth of ϕ_n is clearly $r + 2nr + 2m = 2m + \omega(\sqrt{m} + n)n$ the claim follows and this concludes the proof. ■

Remark 2. The above can be strengthened to show that $\mathcal{G}$ is not $D(n)$-equivalent to $G(2^n, 1/2)$ even for $D(n) = \omega(\sqrt{m(n)} + n)$, at the expense of using seperating formulas of size exponential in n.

Acknowledgments. The second author wishes to thank Nati Linial and Avi Wigderson for helpful discussions and Daniel Reichman for referring him to [16]. We thank Ronen Gradwohl, Eran Ofek, Guy Rothblum, Tal Sagiv and Udi Wieder for helpful comments on an earlier draft.

References

1. L. M. Adleman, C. Pomerance and R. S. Rumely, *On Distinguishing Prime Numbers from Composite Numbers*, Annals of Mathematics, vol. 117, no. 1, 173–206, 1983.
2. N. Alon and J. H. Spencer, *The Probabilistic Method*, John Wiley and Sons, 1992.
3. L. Babai, *Character Sums, Weil's Estimates, and Paradoxical Tournaments*, lecture notes, http://people.cs.uchicago.edu/∼laci/reu02.dir/paley.pdf
4. B. Bollobás. *Random Graphs*, Academic Press, 1985.
5. B. Bollobás and P. Erdös, *Cliques in Random Graphs*, Cambridge Philosophical Society Mathematical Proc., vol. 80, 419–427, 1976.
6. F. R. K. Chung, R. L. Graham and R. M. Wilson, *Quasi-random graphs*, Combinatorica, vol. 9, 345–362, 1989.
7. A. Ehrenfeucht, *An Application of Games to the Completeness Problem for Formalized Theories*, Fundamenta Mathematicae, vol. 49, 129–141, 1961.

8. P. Erdös, C. Pomerance and E. Schmutz, *Carmichael's Lambda Function*, Acta Arithmetica, vol. 58, 363-385, 1991.
9. R. Fagin, *Probabilities in Finite Models*, Journal of Symbolic Logic, vol. 41, 50–58, 1969.
10. S. Gao and D. Panario, *Tests and Constructions of Irreducible Polynomials Over Finite Fields*, Foundations of Computational Mathematics (F. Cucker, M. Shub, Eds.), 346–361, Springer, 1997.
11. Y. V. Glebskii, D. I. Kogan, M. I. Liagonkii, V. A. Talanov, *Range and Degree of Realizability of Formulas in the Restricted Predicate Calculus*, Cybernetics, vol. 5, 142–154, 1976.
12. O. Goldreich, S. Goldwasser, S. Micali, *How to Construct Random Functions*, Journal of the ACM, vol. 33, no. 4, 276–288, 1985.
13. O. Goldreich, S. Goldwasser, A. Nussboim, *On the Implementation of Huge Random Objects*, proc. 44th IEEE Symposium on Foundations of Computer Science, 68–79, 2003.
14. R. L. Graham and J. H. Spencer, *A Constructive Solution to a Tournament Problem*, Canadian Math Bulletin, vol. 14, 45–48, 1971.
15. J. Håstad, R. Impagliazzo, L.A. Levin and M. Luby, *A Pseudo-Random Generator from any One-Way Function*, SIAM Journal on Computing, vol. 28, num. 4, 1364–1396, 1999.
16. M. Krivelevich and B. Sudakov, *Pseudo-random Graphs*, preprint, http://www.math.princeton.edu/∼bsudakov/papers.html
17. A. Nussboim. *Huge Pseudo-Random Graphs that Preserve Global Properties of Random Graphs*, M.Sc. Thesis, Advisor: S. Goldwasser, Weizmann Institute of Science, 2003. http://www.wisdom.weizmann.ac.il/∼asafn/psdgraphs.ps
18. S. C. Pohlig and M. E. Hellman, *An Improved Algorithm for Computing Logarithms Over* GF(p) *and Its Cryptographic Significance*, IEEE Transactions on Information Theory, Vol. IT-24, 106–110, 1978.
19. J. H. Spencer. *The Strange Logic of Random Graphs.* Springer Verlag, 2001.
20. V. Shoup, *New Algorithms for Finding Irreducible Polynomials over Finite Fields*, Mathematics of Computation, vol. 54, 435–447, 1990.
21. V. Shoup, *Searching for primitive roots in finite fields*, Mathematics of Computation, vol. 58, 369–380, 1992.
22. J. H. Spencer and S. Shelah, *Zero-One Laws for Sparse Random Graphs*, Journal of the American Mathematical Society, vol. 1, 97–115, 1988.
23. A. Thomason, *Pseudo-random graphs*, Proceedings of Random Graphs, Annals of Discrete Mathematics 33, 307–331, 1987.