

Discrete metric spaces: structure, enumeration, and 0-1 laws

Dhruv Mubayi *
University of Illinois at Chicago
mubayi@uic.edu

Caroline Terry
University of Illinois at Chicago
cterry3@uic.edu

February 15, 2016

Abstract

Fix an integer $r \geq 3$. We consider metric spaces on n points such that the distance between any two points lies in $\{1, \dots, r\}$. Our main result describes their approximate structure for large n . As a consequence, we show that the number of these metric spaces is

$$\left\lceil \frac{r+1}{2} \right\rceil^{\binom{n}{2} + o(n^2)}.$$

Related results in the continuous setting have recently been proved by Kozma, Meyerovitch, Peled, and Samotij [34]. When r is even, our structural characterization is more precise, and implies that almost all such metric spaces have all distances at least $r/2$. As an easy consequence, when r is even we improve the error term above from $o(n^2)$ to $o(1)$, and also show a labeled first-order 0-1 law in the language \mathcal{L}_r , consisting of r binary relations, one for each element of $[r]$. In particular, we show the almost sure theory T is the theory of the Fraïssé limit of the class of all finite simple complete edge-colored graphs with edge colors in $\{r/2, \dots, r\}$.

Our work can be viewed as an extension of a long line of research in extremal combinatorics to the colored setting, as well as an addition to the collection of known structures that admit logical 0-1 laws.

1 Introduction

Given integers $n, r \geq 3$, define $M_r(n)$ to be the set of all metric spaces with underlying set $[n] := \{1, \dots, n\}$ and distances in $\{1, \dots, r\}$. The goal of this paper is to investigate the approximate structure of most elements of $M_r(n)$ for fixed r and large n , and in the case when r is even, to prove that $M_r(n)$ has a labeled first-order 0-1 law.

1.1 Background

A graph is a set equipped with a symmetric irreflexive binary relation. Given $n \in \mathbb{N}$ and a collection \mathcal{H} of graphs, let $\text{Forb}_n(\mathcal{H})$ denote the set of graphs with vertex set $[n]$ which do not contain any element of \mathcal{H} as a subgraph. There is a long line of research in extremal combinatorics which investigates the structure of graphs in $\text{Forb}_n(\mathcal{H})$ for various \mathcal{H} . One of the first such results is due to Erdős, Kleitman, and Rothschild [20], which states that if $\mathcal{H} = \{K_3\}$, then almost all graphs in $\text{Forb}_n(\mathcal{H})$ are bipartite. More precisely, if $B(n)$ is the set of bipartite graphs on $[n]$, then

$$\lim_{n \rightarrow \infty} \frac{|\text{Forb}_n(\{K_3\})|}{|B(n)|} = 1.$$

*Research supported in part by NSF Grant DMS 1300138

In [30], Kolaitis, Promel, and Rothschild extend this result to the case when $\mathcal{H} = \{K_l\}$ for integers $l \geq 3$, showing that almost all K_l -free graphs are $(l-1)$ -partite. These fundamental combinatorial results have been extended and generalized in numerous ways. For instance, in the graph setting, [5, 8, 35, 37] contain similar results about $\text{Forb}_n(\mathcal{H})$ for specific collections \mathcal{H} , and [3, 4, 19, 28, 44] contain results which apply to $\text{Forb}_n(\mathcal{H})$ for \mathcal{H} satisfying general properties. Results of this spirit for other types of structures include, for example, [11, 12, 29] for partial orders, [42, 45, 47] for directed graphs, and [6, 7, 43] for hypergraphs.

In some cases, the structural information obtained about $\text{Forb}_n(\mathcal{H})$ from such investigation is enough to prove a labeled first-order 0-1 law, which we now define. Suppose \mathcal{L} is a finite first-order language and $F = \bigcup_{n \in \mathbb{N}} F_n$, where F_n is a set of \mathcal{L} -structures with underlying set $[n]$. For each \mathcal{L} -sentence ψ , set $\mu_n(\psi)$ to be the proportion of elements in F_n which satisfy ψ . Then the *asymptotic probability* of ψ is $\mu(\psi) = \lim_{n \rightarrow \infty} \mu_n(\psi)$ (if it exists). We say F has a *labeled first-order limit law* if for each \mathcal{L} -sentence ψ , $\mu(\psi)$ exists, and we say F has a *labeled first-order 0-1 law* if moreover, for each \mathcal{L} -sentence ψ , we have $\mu(\psi) \in \{0, 1\}$. The *almost sure theory* of F is the set of \mathcal{L} -sentences ψ such that $\mu(\psi) = 1$. In [30], Kolaitis, Prömel, and Rothschild use the structural information they obtain about $\text{Forb}(\{K_l\}) = \bigcup_{n \in \mathbb{N}} \text{Forb}_n(\{K_l\})$ for $l \geq 3$ to show that each such family has a labeled first-order 0-1 law in the language of graphs and to give an axiomatization of its almost sure theory.

Given a set X , let $\binom{X}{2} = \{Y \subset X : |Y| = 2\}$ and $2^X = \{Y : Y \subset X\}$. An r -graph G is a pair (V, c) , where V is a (vertex) set, and $c : \binom{V}{2} \rightarrow 2^{[r]}$; we call G a *simple complete r -graph* if $|c(xy)| = 1$ for all $xy \in \binom{V}{2}$. Elements of $M_r(n)$ are naturally viewed as simple complete r -graphs by assigning edge colors corresponding to distances. Given a set \mathcal{H} of r -graphs, let $\text{Forb}_n^r(\mathcal{H})$ be the set of simple complete r -graphs with vertex set $[n]$ which contain no element of \mathcal{H} as a substructure. By taking \mathcal{H} to be the set of simple complete r -graphs on three vertices which contain violations of the triangle inequality, we see that $M_r(n) = \text{Forb}_n^r(\mathcal{H})$. In this way, we can view $M_r(n)$ as an edge-colored analogue of $\text{Forb}_n(\mathcal{H})$. This analogy suggests that one could prove similar results as in [20] and [30] about $M_r(n)$. In this paper we show that this is indeed the case, utilizing techniques from graph theory to describe the approximate structure of most elements of $M_r(n)$ for large n .

We may view elements of $M_r(n)$ as first-order structures in the language \mathcal{L}_r consisting of r binary predicates, one for each edge color. In this setting, as a corollary of our structural results, we are able to prove in the case when r is even, that there is a labeled first-order 0-1 law for $M_r = \bigcup_{n \in \mathbb{N}} M_r(n)$ and to give an axiomatization of its almost sure theory. In this paper, we consider only $r \geq 3$ for the following reason. There is no way to violate the triangle inequality using distances in $\{1, 2\}$, so $M_2(n)$ consists of the set of *all* simple complete 2-graphs. This means that given a pair x, y of distinct elements of $[n]$, the distance between x and y is equal to 1 in exactly half of the elements of $M_2(n)$. For each $G \in M_2(n)$, associate a graph \mathcal{G} with vertex set $[n]$ such that for each $x, y \in [n]$, there is an edge between x and y in \mathcal{G} if and only if the distance between x and y is equal to 1 in G . Under this association, we see that $M_2(n)$ behaves exactly like the random graph $G(n, 1/2)$, the structural properties of which have been studied extensively (see [10]), and which is known to have a labeled first-order 0-1 law [21, 24].

The results of this paper may be of interest to both combinatorialists and model theorists. From the combinatorial perspective, our work appears to be the first extension of the classical enumeration results in extremal graph theory to the edge-colored setting. The proofs of our main results will rely on a stability theorem which is proved using a multi-color version of the Szemerédi regularity lemma [2]. While our proof techniques bear some resemblance to the classical results in [19, 20, 29], we need several new ideas that are motivated by work on weighted Turán-type problems [23]. Our contributions also add to existing results that study metric spaces as combinatorial objects

[14, 34, 39, 40]. In particular, [40] and [34] address questions similar to ours in the continuous setting. In [40], Mascioni investigates the following problem. Given an integer n and a fixed set X of n points, if we assign i.i.d. uniform real numbers in $[0, 1]$ to the elements of $\binom{X}{2}$, what is the probability we get a metric space? It is shown in [40] that this probability p satisfies

$$\left(\frac{1}{2}\right)^{\binom{n}{2}} \leq p \leq \left(\frac{1}{2}\right)^{\lfloor n/2 \rfloor} \left(\frac{2}{3}\right)^{\lfloor n/2 \rfloor (\lceil n/2 \rceil - 2)}, \quad (1)$$

where the lower bound is obtained by noting that any assignment of distances from $[\frac{1}{2}, 1]$ yields a metric space. In more recent work, Kozma, Meyerovitch, Peled, and Samotij [34] identify the set of metric spaces on $[n]$ having all distances in $[0, 1]$ with elements in the cube $[0, 1]^{\binom{n}{2}}$. Let M_n be the subset of $[0, 1]^{\binom{n}{2}}$ which corresponds to the set of metric spaces on $[n]$. Then [34] shows that there are constants c , c' , and C such that $c > 0$ and

$$\frac{1}{2} + \frac{c'}{\sqrt{n}} \leq (\text{vol}M_n)^{1/\binom{n}{2}} \leq \frac{1}{2} + \frac{C}{n^c}. \quad (2)$$

They also prove that with high probability, all distances are between $1/2 - n^{-c}$ and 1. The upper bound in (2) implies that $p^{1/\binom{n}{2}}$ approaches $1/2$ as $n \rightarrow \infty$, where $p = \text{vol}M_n$ is the probability from (1). Given a fixed even $r \geq 4$, our results about $M_r(n)$ can be translated into results about metric spaces on $[n]$ with all distances in $\{\frac{1}{r}, \dots, \frac{r-1}{r}, 1\}$. In this setting, our Theorem 1.2 says that almost all such metric spaces (as $n \rightarrow \infty$) have all of their distances in $[\frac{1}{2}, 1]$, therefore capturing a similar phenomenon as the results of [34] (for odd r the situation is slightly more complicated). If it were possible to generalize our results to the setting where $r \rightarrow \infty$ and n is fixed, then they could apply to the continuous setting.

From the model theory perspective, we provide a new example which may aid in understanding further why some classes of structures have labeled first-order limit laws and others do not. There has been much investigation into finding sufficient conditions for when a class of finite structures has various types of logical limit laws. One type of sufficient conditions, first introduced by Compton in [15, 16], requires that the number of structures of size n does not grow too quickly as $n \rightarrow \infty$. The theorems in [15, 16] and various extensions of them (for instance [9, 13]) provide a large number of examples of logical limit laws. However, these theorems only apply to families for which the number of structures of size n grows subexponentially in n . Yet there are many examples of families with logical limit laws for which the number of structures of size n does grow exponentially. For instance $\text{Forb}(\{K_l\})$ for $l \geq 3$ grows exponentially, but has a labeled first-order 0-1 law [30]. For all $r \geq 3$, we will show the number of structures in M_r of size n also grows exponentially, and therefore fails the sufficient conditions in [15, 16, 9, 13]. In [33] Koponen presents conditions which cover more known examples. In particular, it is shown in [33] that the family of almost l -partite graphs for $l \geq 2$ has a logical limit law. Koponen combines this with the main result of [28] to prove the existence of logical limit laws for $\text{Forb}(\{\mathcal{H}\})$ when \mathcal{H} is a complete $(l+1)$ -partite graph with parts of sizes $1, s_1, \dots, s_l$, for some $1 \leq s_1 \leq \dots \leq s_l$. When $s_1 = \dots = s_l$, $\mathcal{H} = K_{l+1}$, so this generalizes the 0-1 law proved in [30] for $\text{Forb}(\{K_l\})$, $l \geq 3$. More results on logical limit laws for various families of graphs appear in [25, 26, 32, 36]. However, these results do not apply to M_r , as elements of M_r are not graphs.

In [31], Koponen studies the asymptotic probability of extension axioms in families of structures in finite relational languages satisfying certain general requirements. This generality allows the results to be applied to structures other than graphs. For example, Koponen combines results of [31] with the main results of [6] and [43] to show certain families of hypergraphs with forbidden

configurations have labeled first-order 0-1 laws (see Example 10.7 of [31]). Another paper which studies logical limit laws for more general languages is [1] by Ahlman and Koponen, which focuses on families of structures in finite relational languages satisfying certain colorability requirements and with an underlying pregeometry. While none of these results apply directly to M_r , a result of [31] does imply that a subfamily C_r of M_r , (which will be defined later) has a labeled first-order 0-1 law. Our results will show that when r is even, almost all elements of M_r are in C_r , which will yield that M_r has a labeled first-order 0-1 law. Therefore, this paper provides the combinatorial argument required to reduce the existence of a labeled first-order 0-1 law for M_r to the existence of one for C_r , while the fact that C_r has a labeled first-order 0-1 law follows from known results, and is in fact very easy to prove directly. Part of the motivation for this work is the idea that having more examples of logical limit laws in languages other than that of graphs, and seeing the techniques used to prove them, will improve our general understanding of when a family of finite structures has a logical limit law.

1.2 New Results

In this section we state the results in this paper. First we give some necessary definitions and notation. Given positive integers r, s and a set X , set $[r] = \{1, \dots, r\}$, $\binom{X}{s} = \{Y \subset X : |Y| = s\}$, and $2^X = \{Y : Y \subset X\}$. Recall that an r -graph G is a pair (V, c) , where V is a set, and $c : \binom{V}{2} \rightarrow 2^{[r]}$. We call V the *vertex set* of G and c the *coloring* of G . In the case when $|c(e)| \leq 1$ for every $e \in \binom{V}{2}$, we say that (V, c) is *simple*, and when $c(e) \neq \emptyset$ for each $e \in \binom{V}{2}$, we say G is *complete*. Given integers $r, n \geq 3$, we consider $M_r(n)$ as the set of simple complete r -graphs $([n], c)$ satisfying the triangle inequality, i.e, for every three pairwise distinct elements x, y, z of $[n]$,

$$\text{if } \{i\} = c(xy), \{j\} = c(yz), \text{ and } \{k\} = c(xz), \text{ then we have } i \leq j + k.$$

Given a set X and $\{x, y\} \in \binom{X}{2}$, we will write xy to mean $\{x, y\}$. Given integers $i < j$, set $[i, j] = \{i, i + 1, \dots, j\}$.

Definition 1.1. *For an even integer $r \geq 4$ and any integer n , let $C_r(n)$ be the set of all simple complete r -graphs $G = ([n], c)$ such that $c(e) \subset [\frac{r}{2}, r]$ for all $e \in \binom{[n]}{2}$.*

When r is even, there is no way to violate the triangle inequality using distances in $[\frac{r}{2}, r]$, so $C_r(n) \subset M_r(n)$. The strongest structural result we will prove (Theorem 1.2 below) says that when $r \geq 4$ is even, almost all elements in $M_r(n)$ are in $C_r(n)$.

Theorem 1.2. *Let $r \geq 4$ be an even integer. Then there is $\beta > 0$ and $M > 0$ such that for all $n \geq M$,*

$$|C_r(n)| \geq |M_r(n)|(1 - 2^{-\beta n}).$$

When r is even, $|C_r(n)| = (\frac{r}{2} + 1)^{\binom{n}{2}}$. Therefore Theorem 1.2 yields that when r is even,

$$\left(\frac{r}{2} + 1\right)^{\binom{n}{2}} \leq |M_r(n)| \leq \left(\frac{1}{1 - 2^{-\beta n}}\right) \left(\frac{r}{2} + 1\right)^{\binom{n}{2}}$$

for some positive β and sufficiently large n . We obtain the following Corollary.

Corollary 1.3. *Let $r \geq 4$ be an even integer. Then $|M_r(n)| = (\frac{r}{2} + 1)^{\binom{n}{2} + o(1)}$.*

When r is odd, we still obtain a result on the approximate structure of most elements of $M_r(n)$ (Theorem 1.5 below), however the situation in this case is more complicated.

Definition 1.4. Let $r \geq 3$ be an odd integer. Define $C_r(n)$ to be the set of simple complete r -graphs $G = ([n], c)$ such that there is a partition $V_1 \cup \dots \cup V_t$ of $[n]$ and for every $xy \in \binom{[n]}{2}$,

$$c(xy) \subset \begin{cases} \left[\frac{r-1}{2}, r-1 \right] & \text{if } xy \in \binom{V_i}{2} \text{ for some } i \in [t] \\ \left[\frac{r+1}{2}, r \right] & \text{if } x \in V_i, y \in V_j \text{ for some } i \neq j \in [t]. \end{cases}$$

It is easy to see that for r odd, $C_r(n) \subset M_r(n)$. Given $\delta > 0$, two r -graphs $G = (V, c)$ and $G' = (V, c')$ with the same vertex set V are δ -close if $|\{e \in \binom{V}{2} : c(e) \neq c'(e)\}| \leq \delta|V|^2$. Set

$$C_r^\delta(n) = \{G \in M_r(n) : \text{there is } G' \in C_r(n) \text{ such that } G \text{ and } G' \text{ are } \delta\text{-close}\}.$$

We now state our structure theorem which holds for all $r \geq 3$. Informally, it states that most members of $M_r(n)$ are in $C_r^\delta(n)$ for small δ and n large enough depending on δ .

Theorem 1.5. Let $r \geq 3$ be an integer. Then for all $\delta > 0$, there exists an M and $\beta > 0$ such that $n > M$ implies

$$\frac{|M_r(n) \setminus C_r^\delta(n)|}{|M_r(n)|} \leq \frac{|M_r(n) \setminus C_r^\delta(n)|}{\lceil \frac{r+1}{2} \rceil \binom{n}{2}} \leq 2^{-\beta n^2}.$$

We will prove the following enumeration theorem as a corollary of Theorem 1.5 in Section 3.

Corollary 1.6. Let $r \geq 3$ be an integer. Then $|M_r(n)| = \lceil \frac{r+1}{2} \rceil \binom{n}{2}^{+o(n^2)}$.

Given $r \geq 3$, define $\mathcal{L}_r = \{R_1, \dots, R_r\}$ where each R_i is a binary relation symbol. Given $n \in \mathbb{N}$, we can naturally consider elements $G = ([n], c) \in M_r(n)$ as \mathcal{L}_r -structures by interpreting for each $(x, y) \in [n]^2$, $R_i^G(x, y) \Leftrightarrow xy \in \binom{[n]}{2}$ and $c(xy) = \{i\}$. We will prove as a consequence of Theorem 1.2 that, when r is even, $M_r = \bigcup_{n \in \mathbb{N}} M_r(n)$ has a labeled first-order 0-1 law. In fact, we will define a set T_r of \mathcal{L}_r -sentences such that the following holds (see Section 2 for the definition of T_r).

Theorem 1.7. Let $r \geq 4$ be an even integer and consider the elements of $M_r(n)$ and $C_r(n)$ as \mathcal{L}_r -structures. Then $M_r = \bigcup_{n \in \mathbb{N}} M_r(n)$ and $C_r = \bigcup_{n \in \mathbb{N}} C_r(n)$ have labeled first-order 0-1 laws. Moreover they have the same almost sure theory, and this theory is axiomatized by T_r .

When r is odd, the error term in Corollary 1.6 cannot be improved from $o(n^2)$ to $O(n)$, and moreover, Theorem 1.2 does not hold (see Section 7 for a detailed discussion). This leads us to make the following conjecture.

Conjecture 1.8. Let $r \geq 3$ be an odd integer and consider elements of $M_r(n)$ as \mathcal{L}_r -structures as in Theorem 1.7. Then $M_r = \bigcup_{n \in \mathbb{N}} M_r(n)$ has a labeled first-order limit law, but does not have a labeled first-order 0-1 law.

1.3 Notation and outline

Throughout the paper, we will omit floors and ceilings where they are unimportant to the argument. Let $r \geq 3$ be an integer and let G be an r -graph. We will write $V(G)$ to denote the vertex set of G and c^G to denote its coloring. For simplicity of notation we set $E(G) = \binom{V(G)}{2}$, and for subsets $X, Y \subset V(G)$, set $E(X, Y) = \{xy \in E(G) : x \in X, y \in Y\}$, and $E(X) = E(X, X)$. Given a simple complete r -graph G , we define $d^G : E(G) \rightarrow [r]$ to be the function sending $xy \in E(G)$ to the unique $i \in [r]$ such that $c^G(xy) = \{i\}$. We will sometimes also wish to discuss graphs, meaning a set equipped with a single binary, symmetric, irreflexive relation. In order to avoid confusion, these

graphs will be denoted by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where \mathcal{V} is the vertex set of \mathcal{G} and $\mathcal{E} \subset \binom{\mathcal{V}}{2}$ is the edge set of \mathcal{G} . Given a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and $v \in \mathcal{V}$, we will write $\mathcal{DEG}(v) = |\{u : uv \in \mathcal{E}\}|$.

By a *violating triple* we will mean a tuple $(i, j, k) \in \mathbb{N}^3$ such that $|i - j| \leq k \leq i + j$ is false. By a *violating triangle*, we will mean an r -graph H such that $V(H) = \{x, y, z\}$, and for some violating triple (i, j, k) , $i \in c^H(xy)$, $j \in c^H(yz)$, and $k \in c^H(xz)$. Given two r -graphs H and G , we say H is a *sub- r -graph* of G if $V(H) \subset V(G)$ and for all $xy \in \binom{V(H)}{2}$, $c^H(xy) \subset c^G(xy)$. We say G *omits* H if H is not a sub- r -graph of G . Define a *metric r -graph* to be an r -graph $G = (V, c)$ which omits all violating triangles. We will say an r -graph *contains a violating triangle* if it contains a violating triangle as a sub- r -graph. Given two finite r -graphs G and G' with $V(G) = V(G')$, set

$$\Delta(G, G') = \{xy \in E(G) : c^G(xy) \neq c^{G'}(xy)\}.$$

In this notation, given $\delta > 0$, G and G' are δ -close if $|\Delta(G, G')| \leq \delta|V(G)|^2$. Given a set of finite r -graphs S and a finite r -graph G , say that G is δ -close to S if G is δ -close to some element of S . Given $r \geq 3$, set

$$m(r) = \left\lceil \frac{r+1}{2} \right\rceil.$$

A subset $A \subset [r]$ is called a *metric set* if A^3 contains no violating triples. Note that when r is even, $[\frac{r}{2}, r]$ is a metric subset of $[r]$ of size $m(r)$. When r is odd, both $[\frac{r-1}{2}, r-1]$ and $[\frac{r+1}{2}, r]$ are metric subsets of $[r]$ of size $m(r)$. As remarked earlier, any r -graph meeting the requirements in the definition of $C_r(n)$ is already in $M_r(n)$. In particular, $C_r(n)$ contains all simple complete metric r -graphs with distances in $[m(r), r]$, therefore $|C_r(n)| \geq m(r) \binom{n}{2}$. These observations yield the following fact we will use throughout the paper.

Remark 1.9. *Let $n, r \geq 3$ be integers. Then*

$$|M_r(n)| \geq |C_r(n)| \geq m(r) \binom{n}{2},$$

and if r is even, then $|C_r(n)| = m(r) \binom{n}{2}$.

We now give an outline of the paper. In Section 2 we introduce the notion of a labeled first-order 0-1 law, and prove as a consequence of Theorem 1.2 that Theorem 1.7 is true, i.e. when $r \geq 4$ is an even integer, M_r has a labeled first-order 0-1 law in the language consisting of r binary predicates. In Section 3 we prove Corollary 1.6, which provides an asymptotic enumeration of $M_r(n)$ as a consequence of Theorem 1.5. In Section 4 we provide preliminaries and notation regarding a multi-color version of Szemerédi's regularity lemma, then we prove Theorem 4.13, which is a stability result needed to prove Theorem 1.5. In Section 5 we prove Theorem 1.5, and in Section 6 we prove Theorem 1.2. Finally, in Section 7, we explain why Corollary 1.3 and Theorem 1.2 do not hold when r is odd, then discuss open questions concerning $M_r(n)$ when r is odd.

2 Proof of logical 0-1 law

In this section we assume Theorem 1.2 and prove Theorem 1.7, which says that for even integers $r \geq 4$, the family $M_r = \bigcup_{n \in \mathbb{N}} M_r(n)$ has a labeled first-order 0-1 law in the language \mathcal{L}_r consisting of r binary relation symbols. The outline of the argument is as follows. Theorem 1.2 allows us to reduce Theorem 1.7 to showing the existence of a labeled first-order 0-1 law for the subfamily $C_r = \bigcup_{n \in \mathbb{N}} C_r(n)$. The existence of a labeled first-order 0-1 law for C_r follows from a standard argument. In particular, it follows from a theorem in [31] which generalizes the method in [21].

We assume familiarity with the theory of Fraïssé limits. For an introduction to this subject, see chapter 7 of [27]. For a survey on logical 0-1 laws see [49]. We begin with the required terminology concerning 0-1 laws.

Definition 2.1. Let \mathcal{L} be a finite first-order language. For each n , suppose V_n is a set of \mathcal{L} -structures on $[n]$, and $V = \bigcup_{i \in \mathbb{N}} V_i$.

1. $\mu_n^V : V_n \rightarrow [0, 1]$ is the probability measure defined by setting $\mu_n^V(G) = \frac{1}{|V_n|}$ for each $G \in V_n$.
2. Given a first-order \mathcal{L} -sentence ψ , set $\mu_n^V(\psi) = \mu_n^V(\{G \in V_n : G \models \psi\})$ and $\mu^V(\psi) = \lim_{n \rightarrow \infty} \mu_n^V(\psi)$. When $\mu^V(\psi)$ exists, it is called the labeled asymptotic probability of ψ .
3. The almost sure theory of V is $T_{as}^V = \{\psi : \psi \text{ is an } \mathcal{L}\text{-sentence and } \lim_{n \rightarrow \infty} \mu_n^V(\psi) = 1\}$.
4. V has a labeled first-order 0-1 law if for each first-order \mathcal{L} -sentence ψ , $\mu^V(\psi)$ exists and is 0 or 1.

It is straightforward to show that V has a labeled first-order 0-1 law if and only if T_{as}^V is a complete, consistent theory with infinite models.

Fix an even integer $r \geq 4$ for the rest of the section. Define $\mathcal{L}_r = \{R_1(x, y), \dots, R_r(x, y)\}$, where each $R_i(x, y)$ is a binary relation symbol. Given an r -graph G , make G into an \mathcal{L}_r -structure by interpreting for all $(x, y) \in V(G)^2$,

$$R_i^G(x, y) \Leftrightarrow xy \in E(G) \text{ and } i \in c^G(xy).$$

From here on, all r -graphs will be considered as \mathcal{L}_r -structures in this way. We now prove that as a consequence of Theorem 1.2, M_r has a labeled first-order 0-1 law if and only if C_r does.

Lemma 2.2. For all \mathcal{L}_r -sentences ψ , if $\mu^{C_r}(\psi)$ exists, then $\mu^{M_r}(\psi)$ exists, and moreover, $\mu^{C_r}(\psi) = \mu^{M_r}(\psi)$.

Proof. Assume $\mu^{C_r}(\psi)$ exists. For all n ,

$$\mu_n^{M_r}(\psi) = \frac{|\{G \in M_r(n) \setminus C_r(n) : G \models \psi\}|}{|M_r(n)|} + \frac{|\{G \in C_r(n) : G \models \psi\}|}{|M_r(n)|}. \quad (3)$$

By Theorem 1.2, there is $\beta > 0$ such that for sufficiently large n ,

$$|M_r(n) \setminus C_r(n)| \leq 2^{-\beta n} |M_r(n)| \text{ and } |C_r(n)| \leq |M_r(n)| \leq (1 + 2^{-\beta n}) |C_r(n)|,$$

where the second inequality is because for all n , $C_r(n) \subset M_r(n)$. Thus for sufficiently large n ,

$$\frac{|\{G \in C_r(n) : G \models \psi\}|}{|C_r(n)|(1 + 2^{-\beta n})} \leq \frac{|\{G \in C_r(n) : G \models \psi\}|}{|M_r(n)|} \leq \frac{|\{G \in C_r(n) : G \models \psi\}|}{|C_r(n)|}.$$

and

$$\frac{|\{G \in M_r(n) \setminus C_r(n) : G \models \psi\}|}{|M_r(n)|} \leq 2^{-\beta n}.$$

Therefore

$$\lim_{n \rightarrow \infty} \frac{|\{G \in M_r(n) \setminus C_r(n) : G \models \psi\}|}{|M_r(n)|} = 0$$

and

$$\lim_{n \rightarrow \infty} \frac{|\{G \in C_r(n) : G \models \psi\}|}{|M_r(n)|} = \lim_{n \rightarrow \infty} \frac{|\{G \in C_r(n) : G \models \psi\}|}{|C_r(n)|} = \mu^{C_r}(\psi).$$

Combining these with (3) yields that $\mu^{M_r}(\psi) = \mu^{C_r}(\psi)$. \square

Towards stating the definition of T_r from Theorem 1.7, we now fix some notation. Fix an integer $k \geq 2$. Given $A \in M_r(k)$, write $x_1 \dots x_k \equiv A$ as short hand for the \mathcal{L}_r -formula which says that sending $x_i \mapsto i$ makes $x_1 \dots x_k$ isomorphic to A . Explicitly we mean the formula $\psi(x_1, \dots, x_k)$ given by

$$\bigwedge_{1 \leq i < j \leq k} \left(R_{d^A(i,j)}(x_i, x_j) \wedge \bigwedge_{s \neq d^A(i,j)} \neg R_s(x_i, x_j) \right).$$

Given $A \in M_r(k)$ and $A' \in M_r(k+1)$, write $A \prec A'$ to denote that for all $ij \in \binom{[k]}{2}$, $d^A(ij) = d^{A'}(ij)$. Given such a pair $A \prec A'$, let $\sigma_{A'/A}$ be the following sentence:

$$\forall x_1 \dots \forall x_k ((x_1 \dots x_k \equiv A) \rightarrow \exists y (x_1 \dots x_k y \equiv A')).$$

Sentences of this form are called *extension axioms*. Let T be a set of \mathcal{L}_r -sentences axiomatizing an infinite metric space with distances all in $[\frac{r}{2}, r]$,

$$T' = \bigcup_{k \in \mathbb{N}} \{ \sigma_{A'/A} : A \in C_r(k), A' \in C_r(k+1), A \prec A' \}, \text{ and}$$

$$T_r = T \cup T'.$$

Given an \mathcal{L}_r -structure M , let $Th(M)$ denote the set of \mathcal{L}_r -sentences true in M . Given a set of \mathcal{L}_r -sentences Γ , we will write $M \models \Gamma$ to denote that $M \models \phi$ for all $\phi \in \Gamma$.

Proof of Theorem 1.7. Lemma 2.2 implies that if C_r has a labeled first-order 0-1 law then so does M_r , and moreover $T_{as}^{C_r} = T_{as}^{M_r}$. Therefore, it suffices to show C_r has a labeled first-order 0-1 law and to show T_r axiomatizes $T_{as}^{C_r}$. Let \mathcal{C}_r be the class of \mathcal{L}_r -structures obtained by closing C_r under isomorphism. Given $M \in \mathcal{C}_r$, we will write $V(M)$ to denote the underlying set of M . That \mathcal{C}_r is a Fraïssé class is straightforward to see. For the sake of completeness we verify that \mathcal{C}_r has the amalgamation property. Given $X, Y \in \mathcal{C}_r$, an isometry $f : X \rightarrow Y$ is an injective map from $V(X)$ into $V(Y)$ such that for all $xy \in E(X)$, $d^X(x, y) = d^Y(f(x), f(y))$. Suppose $A, B, C \in \mathcal{C}_r$ and $f : C \rightarrow A$, $g : C \rightarrow B$ are isometries. Without loss of generality, assume that f and g are inclusion maps and $V(A) \cap V(B) = V(C)$. To verify the amalgamation property, we want to find $D \in \mathcal{C}_r$ and isometries $h : A \rightarrow D$ and $s : B \rightarrow D$ such that for all $c \in V(C)$, $s(c) = h(c)$. We do this by setting $V(D) = V(A) \cup V(B)$ and for $xy \in \binom{V(D)}{2}$, setting

$$d^D(x, y) = \begin{cases} d^A(x, y) & \text{if } xy \in E(A), \\ d^B(x, y) & \text{if } xy \in E(B) \setminus E(A), \\ r & \text{if } x \in (V(A) \setminus V(C)), y \in (V(B) \setminus V(C)). \end{cases} \quad (4)$$

D is a simple complete r -graph with $d^D(x, y) \in [\frac{r}{2}, r]$ for all $xy \in E(D)$, so $D \in \mathcal{C}_r$. Define $h : A \rightarrow D$ and $s : B \rightarrow D$ to be the inclusion maps. Then for all $c \in V(C)$, $h(c) = s(c) = c$, as desired, and \mathcal{C}_r has the amalgamation property. Note that we could have chosen any color in $[\frac{r}{2}, r]$ to assign the edges in the third case of (4), as there are no forbidden configurations in \mathcal{C}_r . We leave the rest of the verification that \mathcal{C}_r is a Fraïssé class to the reader.

Let $FL(\mathcal{C}_r)$ be the Fraïssé limit of \mathcal{C}_r and make $FL(\mathcal{C}_r)$ into an \mathcal{L}_r -structure by interpreting, for each $(x, y) \in FL(\mathcal{C}_r)^2$, $R_i(x, y)$ if and only if $d^{C_r}(x, y) = i$. It is a standard exercise to see that $FL(\mathcal{C}_r) \models T_r$ and further that T_r axiomatizes $Th(FL(\mathcal{C}_r))$. Therefore T_r is a complete, consistent \mathcal{L}_r -theory, so to show C_r has a labeled first-order 0-1 law, it suffices to show that for each $\psi \in T_r$, $\mu^{C_r}(\psi) = 1$. For $\psi \in T$, this is obvious. Because there are no forbidden configurations in \mathcal{C}_r , a

straightforward counting argument shows that for $\psi \in T'$, $\mu^{C_r}(\neg\psi) = 0$, and therefore $\mu^{C_r}(\psi) = 1$. An example of such an argument applied to graphs is the proof of Lemma 2.4.3 of [38]. The proof in our case is only slightly more complicated, so we omit it. We also point out that this fact (that for all $\psi \in T'$, $\mu^{C_r}(\psi) = 1$) follows directly from a much more general result, Theorem 3.15 of [31]. Because this theorem is much more powerful than what our example requires, we leave it to the interested reader to verify it applies to C_r and $\psi \in T'$. \square

We end this section by showing that while there is a Fraïssé limit naturally associated to M_r , its theory is very different from the almost sure theory we obtain from M_r . Let \mathcal{M}_r be the class of finite metric spaces obtained by closing M_r under isomorphism, that is, \mathcal{M}_r is the class of all finite metric spaces with distances all in $[r]$. It is well known that \mathcal{M}_r is a Fraïssé class. For instance, this is a simple case of general results contained in [18], which tell us when, given $S \subset \mathbb{R}$, the class of finite metric spaces with distances all in S forms a Fraïssé class. For completeness we verify the amalgamation property for our case, that is, when $S = [r]$.

Suppose $A, B, C \in \mathcal{M}_r$ and $f : C \rightarrow A$, $g : C \rightarrow B$ are isometries. Without loss of generality, assume that f and g are inclusion maps and $V(A) \cap V(B) = V(C)$. To verify the amalgamation property, we want to find $D \in \mathcal{M}_r$ and isometries $h : A \rightarrow D$ and $s : B \rightarrow D$ such that for all $c \in V(C)$, $s(c) = h(c)$. Given $s, t \in [r]$, let $t \dot{+} s = \min\{r, t + s\}$. Set $V(D) = V(A) \cup V(B)$ and for $xy \in \binom{V(D)}{2}$, set

$$d^D(x, y) = \begin{cases} d^A(x, y) & \text{if } xy \in E(A), \\ d^B(x, y) & \text{if } xy \in E(B) \setminus E(A), \\ \max\{d^A(x, c) \dot{+} d^B(c, y) : c \in V(C)\} & \text{if } x \in (V(A) \setminus V(C)), y \in (V(B) \setminus V(C)). \end{cases} \quad (5)$$

We leave it to the reader to verify that the assigned distances do not violate the triangle inequality, and therefore, that D is in \mathcal{M}_r . Define $h : A \rightarrow D$ and $s : B \rightarrow D$ to be the inclusion maps. Then for all $c \in V(C)$, $h(c) = s(c) = c$, as desired, and \mathcal{M}_r has the amalgamation property. Note that unlike in the proof of the amalgamation property for \mathcal{C}_r , the distance in the third line of (5) must be chosen carefully, as there are many forbidden configurations in \mathcal{M}_r .

Let $FL(\mathcal{M}_r)$ be the Fraïssé limit of \mathcal{M}_r . It is a standard exercise that the theory of $FL(\mathcal{M}_r)$ is axiomatized by the axioms for an infinite metric space with distances all in $[r]$ and the collection of all extension axioms of the form $\sigma_{A'/A}$ for some $A \in M_r(k)$, $A' \in M_r(k+1)$ with $A \prec A'$, and $k \geq 0$. We can see now that $\text{Th}(FL(\mathcal{M}_r))$ and $\text{Th}(FL(\mathcal{C}_r))$ are different. For instance, let ψ be the sentence

$$\exists x \exists y R_1(x, y).$$

Then $\psi \in \text{Th}(FL(\mathcal{M}_r))$, while clearly $\text{Th}(FL(\mathcal{C}_r)) \models \neg\psi$. Model theoretically, $\text{Th}(FL(\mathcal{C}_r))$ is simple (in the sense of Definition 7.2.1 in [48]). This can be seen by adapting the argument used to prove the theory of the random graph is simple, as \mathcal{C}_r is just an edge-colored version of the random graph (see Corollary 7.3.14 in [48] for a proof that the theory of the random graph is simple). On the other hand, a straightforward adjustment of the construction in Theorem 5.5(b) of [17] shows that $\text{Th}(FL(\mathcal{M}_r))$ has the *r-strong order property* (SOP_r), a measure of the complexity of a first-order theory defined in [46]. It is shown in [46] that for all $n \geq 3$, a theory with SOP_n is not simple. In sum, when $r \geq 4$ is even, we have a family of labeled finite structures, M_r , associated to two theories which differ in model theoretic complexity:

- $\text{Th}(FL(\mathcal{M}_r))$ where \mathcal{M}_r is obtained by closing M_r under isomorphism. This theory has SOP_r (and therefore is not simple).
- $T_{as}^{M_r} = T_{as}^{C_r} = \text{Th}(FL(\mathcal{C}_r))$, where $C_r \subset M_r$ is a special subfamily, and \mathcal{C}_r is obtained by closing C_r under isomorphism. This theory is simple.

3 Asymptotic Enumeration

In this section we assume Theorem 1.5 and prove Corollary 1.6, which asymptotically enumerates $M_r(n)$ for all $r \geq 3$. Recall that for all integers $r \geq 3$, $m(r) = \lceil \frac{r+1}{2} \rceil$.

Proof of Corollary 1.6. Fix an integer $r \geq 3$. All logs will be base $m(r)$ unless otherwise stated. Remark 1.9 implies that $|M_r(n)| \geq m(r) \binom{n}{2}$, so it suffices to show that for all $0 < \gamma < 1$, there is M such that $n > M$ implies $|M_r(n)| < m(r) \binom{n}{2} + \gamma n^2$.

Fix $0 < \gamma < 1$. Let $H(x) = -x \log_2 x - (1-x) \log_2(1-x)$. Recall that $\binom{n}{xn} \leq 2^{H(x)n}$ for all $n \in \mathbb{N}$ and $0 < x \leq \frac{1}{2}$ (see for instance page 427 of [22]). Basic calculus shows $H(x) \rightarrow 0$ as $x \rightarrow 0$. Choose $\delta > 0$ small enough so that

$$(H(\delta) + \delta) \log 2 + \delta \log r < \frac{\gamma}{4}.$$

Theorem 1.5 implies there exists a $\beta = \beta(\delta) > 0$ and $M_1 = M_1(\delta)$ such that $n > M_1$ implies

$$|M_r(n) \setminus C_r^\delta(n)| \leq 2^{-\beta n^2} m(r) \binom{n}{2}.$$

Choose $M > M_1$ large enough so that $n > M$ implies $\frac{\gamma}{4} n^2 + n \log n < \frac{\gamma}{2} n^2$ and $\frac{\gamma}{2} n^2 + \log 2 \leq \gamma n^2$. We now assume $n > M$ and bound the size of $C_r^\delta(n)$. All elements $G \in C_r^\delta(n)$ can be constructed as follows:

- Choose an element of $G' \in C_r(n)$. There are $|C_r(n)|$ ways to do this. If r is even, then $|C_r(n)| = m(r) \binom{n}{2}$. If r is odd, we must find an upper bound for $|C_r(n)|$. When r is odd, we can construct any element of $C_r(n)$ by first choosing a partition of $[n]$, then assigning a color to each edge in a way compatible with the partition. There are at most $n^n m(r) \binom{n}{2}$ ways to do this.
- Choose at most δn^2 edges to be in $\Delta(G, G')$. There are at most $\binom{n^2}{\delta n^2} 2^{\delta n^2} \leq 2^{(H(\delta) + \delta)n^2}$ ways to do this.
- Assign a color to each edge in $\Delta(G, G')$. There are at most $r^{\delta n^2}$ ways to do this.

Thus

$$|C_r^\delta(n)| \leq n^n m(r) \binom{n}{2} 2^{(H(\delta) + \delta)n^2} r^{\delta n^2} = m(r) \binom{n}{2} + n^2((H(\delta) + \delta) \log 2 + \delta \log r) + n \log n.$$

By our assumptions on δ and M , this is at most $m(r) \binom{n}{2} + \frac{\gamma}{4} n^2 + n \log n < m(r) \binom{n}{2} + \frac{\gamma}{2} n^2$. Therefore, since $M_r(n) = (M_r(n) \setminus C_r^\delta(n)) \cup C_r^\delta(n)$ we have

$$|M_r(n)| \leq m(r) \binom{n}{2} - n^2 \beta \log 2 + m(r) \binom{n}{2} + \frac{\gamma}{2} n^2 \leq 2m(r) \binom{n}{2} + \frac{\gamma}{2} n^2 = m(r) \binom{n}{2} + \frac{\gamma}{2} n^2 + \log 2 \leq m(r) \binom{n}{2} + \gamma n^2,$$

where the last inequality is by the choice of M . □

4 Stability Theorem

In this section we prove a stability theorem which implies that for all integers $r \geq 3$, for large enough n , if $G \in M_r(n)$ has close to the maximal number of different distances occurring between its vertices, then it is structurally close to an element of $C_r(n)$. This is a crucial step in the proofs of Theorems 1.2 and 1.5. Before proceeding further, we require some definitions and notation.

4.1 Regularity Lemmas and Preliminaries

In this section we state a version of Szemerédi's Regularity Lemma which applies to r -graphs. We will also prove easy consequences of this for our situation.

Definition 4.1. *Let $r \geq 3$ be an integer. Fix a finite r -graph G and disjoint subsets $X, Y \subset V(G)$.*

1. *We call a partition $\mathcal{A} = \{A_1, \dots, A_m\}$ of $V(G)$ an equipartition of order m if $||A_i| - |A_j|| \leq 1$ for all $i \neq j$. A refinement of \mathcal{A} is a partition $\mathcal{B} = \{B_1, \dots, B_k\}$ such that for each $i \in [k]$, there is $j \in [m]$ such that $B_i \subset A_j$.*
2. *For $l \in [r]$, set*

$$e_l^G(X, Y) := |\{xy \in E(X, Y) : l \in c^G(xy)\}| \text{ and}$$

$$\rho_l^G(X, Y) := \frac{e_l(X, Y)}{|X||Y|}.$$

3. *The density vector of (X, Y) in G is $(\rho_1^G, \dots, \rho_r^G)$ where $\rho_i^G = \rho_i^G(X, Y)$.*
4. *(X, Y) is ϵ -regular for G if for all $X' \subset X$ and $Y' \subset Y$ with $|X'| \geq \epsilon|X|$ and $|Y'| \geq \epsilon|Y|$, for all $l \in [r]$,*

$$|\rho_l^G(X, Y) - \rho_l^G(X', Y')| \leq \epsilon.$$

5. *A partition $\mathcal{B} = \{B_1, \dots, B_k\}$ of $V(G)$ is called ϵ -regular for G if it is an equipartition of $V(G)$, and for all but at most ϵk^2 of the pairs $ij \in \binom{[k]}{2}$, (B_i, B_j) is ϵ -regular for G .*

We now state the multi-color version of the Szemerédi Regularity Lemma and one of its corollaries we will use in this paper. Both results appear in [2].

Theorem 4.2. *(Regularity Lemma) Fix an integer $r \geq 2$. For every $\epsilon > 0$ and positive integer m , there is an integer $CM = CM(m, \epsilon)$ such that if G is a finite r -graph with at least CM vertices, and \mathcal{A} is an equipartition of G of order m , then there k such that $m \leq k \leq CM$ and a refinement \mathcal{B} of \mathcal{A} of order k which is ϵ -regular for G .*

Theorem 4.3. *(Embedding Lemma) Fix an integer $r \geq 2$. For every $0 < d < 1$ and $k \in \mathbb{N} \setminus \{0\}$, there is a positive $\gamma = \gamma_{el}(d, k) \leq d$ and $\delta = \delta_{el}(d, k)$ such that the following holds. Suppose that H and G are r -graphs and $V(H) = \{v_1, \dots, v_k\}$. Suppose V_1, \dots, V_k are pairwise disjoint subsets of $V(G)$ such that for every $ij \in \binom{[k]}{2}$, (V_i, V_j) is γ -regular for G , and for each $l \in [r]$, $l \in c^H(v_i v_j)$ implies $\rho_l^G(V_i, V_j) \geq d$. Then there are at least $\delta \prod_{i=1}^k |V_i|$ k -tuples $(w_1, \dots, w_k) \in V_1 \times \dots \times V_k$ such that for each $ij \in \binom{[k]}{2}$, $c^H(v_i v_j) \subset c^G(w_i w_j)$.*

We will apply these theorems to what are called *reduced r -graphs*, which we define below. Recall that a metric r -graph is an r -graph with no violating triangles.

Definition 4.4. *Let $r \geq 2$ be an integer, G a finite r -graph, and $0 < \eta \leq d \leq 1$.*

1. *Suppose $\mathcal{P} = \{V_1, \dots, V_t\}$ is an η -regular partition for G . Let $R(G, \mathcal{P}, d)$ be the r -graph R with vertex set $[t]$ such that $s \in c^R(ij)$ if and only if (V_i, V_j) is η -regular for G and $\rho_s(V_i, V_j) \geq d$. We say R is a reduced r -graph obtained from G with parameters η and d .*

2. Let $\tilde{M}_r(t)$ be the set of metric r -graphs on $[t]$ and set

$$Q_{\eta,d,t}(G) = \{R(G, \mathcal{P}, d) : \mathcal{P} \text{ is an } \eta\text{-regular equipartition for } G \text{ and } \mathcal{P} \text{ has order } t\}, \text{ and}$$

$$Q_{\eta,d}(G) = \bigcup_{t=\frac{1}{\eta}}^{CM(\frac{1}{\eta}, \eta)} Q_{\eta,d,t}(G).$$

We emphasize that the difference between $\tilde{M}_r(t)$ and $M_r(t)$ is that r -graphs in $\tilde{M}_r(t)$ need not be simple and need not be complete. The following two lemmas will be needed.

Lemma 4.5. *Let $r \geq 2$ be an integer, $0 < d < 1$, $0 < \gamma \leq \gamma_{el}(d, 3)$, and $\delta \leq \delta_{el}(d, 3)$. Let $(i, j, k) \in [r]^3$ be a violating triple. Suppose $G \in M_r(n)$ and $V_1, V_2, V_3 \subset V(G)$ are pairwise disjoint and pairwise γ -regular for G with $\delta|V_1||V_2||V_3| \geq 1$. If $\{X, Y, Z\} = \{V_1, V_2, V_3\}$, then*

$$\min\{\rho_i^G(X, Y), \rho_j^G(Y, Z), \rho_k^G(X, Z)\} < d. \quad (6)$$

Proof. Suppose for contradiction that $\{X, Y, Z\} = \{V_1, V_2, V_3\}$ and (6) fails. By Theorem 4.3 there exists at least $\delta|V_1||V_2||V_3| \geq 1$ tuples $(x, y, z) \in X \times Y \times Z$ such that $i \in c^G(xy)$, $j \in c^G(yz)$ and $k \in c^G(xz)$. But now $\{x, y, z\}$ is a violating triangle in G , a contradiction. \square

Lemma 4.6. *Let $0 < d < 1$ and $0 < \eta \leq \gamma_{el}(d, 3)$. There is an M such that $n > M$ implies that for all $G \in M_r(n)$, $\emptyset \neq Q_{\eta,d}(G) \subset \bigcup_{t=\frac{1}{\eta}}^{CM(\frac{1}{\eta}, \eta)} \tilde{M}_r(t)$. In other words, any reduced r -graph obtained from G with parameters d and η omits all violating triangles.*

Proof. Let $M = \frac{2CM(\frac{1}{\eta}, \eta)}{\delta_{el}(d, 3)^{\frac{1}{3}}}$. Suppose $n > M$ and $G \in M_r(n)$. As $n > CM(\frac{1}{\eta}, \eta)$, there is t with $\frac{1}{\eta} \leq t \leq CM(\frac{1}{\eta}, \eta)$ and $\mathcal{P} = \{V_1, \dots, V_t\}$ an η -regular partition for G . Therefore $Q_{\eta,d,t}(G) \neq \emptyset$, so $Q_{\eta,d}(G) \neq \emptyset$. Let $R = R(G, \mathcal{P}, d) \in Q_{\eta,d,t}(G)$. We will show that $R \in \tilde{M}_r(t)$. Note that for all $V_i, V_j, V_k \in \mathcal{P}$,

$$\delta_{el}(d, 3)|V_i||V_j||V_k| \geq \delta_{el}(d, 3) \left(\frac{n}{t} - 1\right)^3 > \delta_{el}(d, 3) \left(\frac{n}{2t}\right)^3 \geq \delta_{el}(d, 3) \frac{n^3}{8CM(\frac{1}{\eta}, \eta)^3} > 1,$$

by assumption on M . Thus by Lemma 4.5, R contains no violating triangle, so $R \in \tilde{M}_r(t)$. \square

We spend the rest of this section stating various definitions and facts we will need for our proofs. We will work with the following subset $\tilde{C}_r(n) \subset \tilde{M}_r(n)$ which is an analogue of $C_r(n) \subset M_r(n)$.

Definition 4.7. *Let $r \geq 3$ be an integer. Set $\tilde{C}_r(t)$ to be the the set of complete r -graphs R with $V(R) = [t]$ such that*

(i) *if r is even, then for all $xy \in E(R)$, $c^R(xy) = [\frac{r}{2}, r]$.*

(ii) *if r is odd, then there is a partition $[t] = V_1 \cup \dots \cup V_s$ such that for all $xy \in \binom{[t]}{2}$,*

$$c^R(xy) = \begin{cases} [\frac{r-1}{2}, r-1] & \text{if } xy \in \binom{V_i}{2} \text{ for some } i \in [s] \\ [\frac{r+1}{2}, r] & \text{if } xy \in E(V_i, V_j) \text{ for some } i \neq j \in [s]. \end{cases}$$

Note that elements of $\tilde{C}_r(t)$ contain no violating triangles, so $\tilde{C}_r(t) \subset \tilde{M}_r(t)$. Further, note that when r is even, $\tilde{C}_r(t)$ consists of a single r -graph. The following weight function defined on metric r -graphs is crucial to our proof.

Definition 4.8. Let $t \geq 2$ and $r \geq 3$ be integers and let $R \in \tilde{M}_r(t)$. For $ij \in \binom{[t]}{2}$, set

$$f^R(i, j) = \max\{|c^R(ij)|, 1\} \quad \text{and} \quad W(R) = \prod_{ij \in \binom{[t]}{2}} f^R(i, j).$$

Note that for integers $r \geq 3$ and $t \geq 2$, any r -graph R with t vertices has $W(R) \leq r^{\binom{t}{2}}$. Recall that when r is even $m(r) = \lfloor \frac{r}{2}, r \rfloor$ and when r is odd, $m(r) = \lfloor \frac{r-1}{2}, r-1 \rfloor = \lfloor \frac{r+1}{2}, r \rfloor$, so for any integers $r \geq 3$ and $t \geq 2$, for all $R \in \tilde{C}_r(t)$ and $ij \in \binom{[t]}{2}$, $f^R(i, j) = m(r)$, and thus $W(R) = m(r)^{\binom{t}{2}}$.

We now state a lemma which restricts how many colors we can assign to the edges of a triangle $\{i, j, k\}$ in an r -graph without creating a violating triangle. The proof of this lemma is elementary but somewhat tedious, and for this reason is relegated to the Appendix.

Lemma 4.9. Fix an integer $r \geq 3$. Let A, B , and C be nonempty subsets of $[r]$ such that $|A| \geq |B| \geq |C|$, $|A| > m(r)$, and $|B| \geq m(r)$. Set $x = |A| - m(r)$ and $y = |B| - m(r)$, and suppose

$$|C| \geq \begin{cases} \max\{m(r) - x - y, 1\} & \text{if } r \text{ is even} \\ \max\{m(r) - x - y + 2, 1\} & \text{if } r \text{ is odd.} \end{cases}$$

Then there is a violating triple $(a, b, c) \in A \times B \times C$.

A straightforward consequence of this is that $m(r)$ is the largest size of a metric subset of $[r]$. Another important consequence is the following.

Corollary 4.10. Let $r, t \geq 3$ be integers and let $R \in \tilde{M}_r(t)$. Suppose $uv, vw, uw \in E(R)$, and $f^R(u, v) \geq f^R(v, w) > m(r)$. Then $f^R(u, w) < m(r)$ and $\max\{f^R(u, v)f^R(u, w), f^R(v, w)f^R(u, w)\} \leq m(r)^2 - 1$.

Proof. For $xy \in \binom{[t]}{2}$, set $f(x, y) = f^R(x, y)$. Given $A, B, C \subset [r]$ and $x, y \in [r]$, write $P(A, B, C, x, y)$ if A, B, C, x, y satisfy the hypotheses of Lemma 4.9, that is, if the following hold.

- $|A| \geq |B| \geq |C|$,
- $x = |A| - m(r)$ and $y = |B| - m(r)$,
- $|A| > m(r)$ and $|B| \geq m(r)$.

Set $A = c^R(uv)$, $B = c^R(vw)$, $C = c^R(uw)$, $x = |A| - m(r)$, and $y = |B| - m(r)$. We show $P(A, B, C, x, y)$ holds. By definition of x and y , the second bullet holds. Since $|A| = f(u, v)$, $|B| = f(v, w)$, and $|C| = f(u, w)$, our assumptions imply $|A| \geq |B| > m(r)$, so the last bullet holds. We now show the first bullet holds, that is, that $|A| \geq |B| \geq |C|$. Suppose for a contradiction that $|C| > |B|$. Let $z = |C| - m(r)$ and note our assumptions imply that either $P(A, C, B, x, z)$ or $P(C, A, B, z, x)$ holds. In either case, $|B| > m(r) \geq m(r) - x - z + 2$ implies by Lemma 4.9 that there is a violating triple $(a, b, c) \in A \times B \times C$. Now $\{u, v, w\}$ is a violating triangle in R , a contradiction. Thus $|A| \geq |B| \geq |C|$.

Consequently, $P(A, B, C, x, y)$ holds, so if $|C| \geq m(r) - x - y + 2$ were true, Lemma 4.9 would imply that there is a violating triple $(a, b, c) \in A \times B \times C$, making $\{u, v, w\}$ a violating triangle in

R , a contradiction. Therefore, we must have $|C| < m(r) - x - y + 2$. Our assumptions imply that $x, y \geq 1$, so in fact, $|C| < m(r)$. Further, we have shown that

$$|B||C| = f(v, w)f(u, w) \leq (m(r)+y)(m(r)-x-y+1) \leq (m(r)+y)(m(r)-y) = m(r)^2 - y^2 \leq m(r)^2 - 1,$$

and

$$|A||C| = f(u, v)f(u, w) \leq (m(r)+x)(m(r)-x-y+1) \leq (m(r)+x)(m(r)-x) = m(r)^2 - x^2 \leq m(r)^2 - 1,$$

as desired. \square

4.2 Two Lemmas

In this section, we prove two lemmas toward our stability result. The first lemma bounds the size of $W(R)$ for $R \in \tilde{M}_r(t)$. We will frequently use the following inequality which holds for all integers $r \geq 3$:

$$m(r)^2 - 1 \geq r. \tag{7}$$

Lemma 4.11. *Let $t, r \geq 3$ be integers and $R \in \tilde{M}_r(t)$. Let $a_R = |\{ij \in E(R) : f^R(i, j) > m(r)\}|$. Then*

$$W(R) \leq m(r)^{\binom{t}{2} + t + 5} \left(\frac{m(r)^2 - 1}{m(r)^2} \right)^{a_R}.$$

Proof. Fix an integer $r \geq 3$. Given an integer t and $R \in \tilde{M}_r(t)$, set $g(R) = m(r)^{\binom{t}{2} + t + 5} \left(\frac{m(r)^2 - 1}{m(r)^2} \right)^{a_R}$.

We proceed by induction on t . Assume $t = 3$ and fix $R \in \tilde{M}_r(t)$. In this case $a_R \leq 3$, so $g(R) \geq m(r)^5(m(r)^2 - 1)^3$. It is straightforward to verify that $r^3 \leq m(r)^5$, as $r \geq 3$. Therefore,

$$W(R) \leq r^3 \leq m(r)^5(m(r)^2 - 1)^3 \leq g(R).$$

Assume now that $t > 3$ and the claim holds for all t' with $3 \leq t' < t$. Fix $R \in \tilde{M}_r(t)$, set $a = a_R$, and for $xy \in \binom{[t]}{2}$, set $f(x, y) = f^R(x, y)$. If $a = 0$ then $W(R) \leq m(r)^{\binom{t}{2}} \leq g(R)$ trivially. So assume $a > 0$.

Choose $uv \in E(R)$ such that $f(u, v)$ is maximum, and note that $a > 0$ implies $f(u, v) > m(r)$. Define R' to be the r -graph with $V(R') = [t] \setminus \{u, v\}$ and for each $xy \in E(R')$, $c^{R'} = c^R|_{V(R')}$. Let $a' = a_{R'}$,

$$Y = \{z \in V(R') : \max\{f(u, z), f(v, z)\} > m(r)\},$$

and set $s = |Y|$. For all $z \in Y$, because $\max\{f(u, z), f(v, z)\} > m(r)$ and $f(u, v) > m(r)$, Corollary 4.10 implies $\min\{f(u, z), f(v, z)\} < m(r)$ and $f(u, z)f(v, z) \leq m(r)^2 - 1$. By the definition of Y , for all $z \in V(R') \setminus Y$, $\max\{f(u, z), f(v, z)\} \leq m(r)$, so $f(u, z)f(v, z) \leq m(r)^2$. Combining these facts we have

$$\begin{aligned} W(R) &= W(R')f(u, v) \left(\prod_{z \in Y} f(u, z)f(v, z) \right) \left(\prod_{z \notin Y} f(u, z)f(v, z) \right) \\ &\leq W(R')f(u, v)(m(r)^2 - 1)^s m(r)^{2(t-2-s)} \leq W(R')r(m(r)^2 - 1)^s m(r)^{2(t-2-s)}. \end{aligned}$$

Using (7), we can upper bound this by

$$W(R')(m(r)^2 - 1)^{s+1} m(r)^{2(t-2-s)} = W(R') \left(\frac{m(r)^2 - 1}{m(r)^2} \right)^{s+1} m(r)^{2t-2}.$$

By the induction hypothesis, this is at most

$$m(r)^{\binom{t-2}{2}+t-2+5} \left(\frac{m(r)^2-1}{m(r)^2} \right)^{a'} \left(\frac{m(r)^2-1}{m(r)^2} \right)^{s+1} m(r)^{2t-2} = m(r)^{\binom{t}{2}+t+4} \left(\frac{m(r)^2-1}{m(r)^2} \right)^{a'+s+1}.$$

Note that $a = a' + |\{zu : z \in Y \text{ and } f(u, z) > m(r)\} \cup \{vz : z \in Y \text{ and } f(v, z) > m(r)\} \cup \{uv\}|$. Because for each $z \in Y$ exactly one of $f(u, z)$ or $f(v, z)$ is strictly greater than $m(r)$, this shows $a = a' + s + 1$. Therefore,

$$W(R) \leq m(r)^{\binom{t}{2}+t+4} \left(\frac{m(r)^2-1}{m(r)^2} \right)^a < g(R).$$

This completes the proof. \square

We now fix some notation. Suppose $r \geq 3$ is an integer, $0 < \epsilon < 1$, R is an r -graph, $u \in V(R)$, and $t = |V(R)|$. For $i \in [r]$, set

$$\begin{aligned} \Gamma_i^R(u) &= \{v \in V(R) : f^R(u, v) = i\}, \\ \mu_i^R(u) &= |\Gamma_i^R(u)|, \text{ and} \\ V_0^R(\epsilon) &= \{v \in V(R) : \mu_{m(r)}^R(v) < (1 - \sqrt{\epsilon})(t - 1)\}. \end{aligned}$$

We now prove the second lemma.

Lemma 4.12. *For every integer $r \geq 3$ there are C_1, C_2, C_3 , depending only on r such that for every $0 < \epsilon < 1$, there is M such that if $t > M$ the following holds. Suppose $R \in \tilde{M}_r(t)$ with $W(R) > m(r)^{(1-\epsilon)\binom{t}{2}}$. Let $a_R = |\{ij \in E(R) : f^R(i, j) > m(r)\}|$ and $b_R = |\{ij \in E(R) : f^R(i, j) < m(r)\}|$. Then*

1. $a_R \leq C_1 \epsilon t^2$,
2. $b_R \leq C_2 \epsilon t^2$, and
3. $|V_0^R(\epsilon)| \leq \sqrt{\epsilon} C_3 t$.

Proof. Let $r, t \geq 3$ be integers. Fix $\epsilon > 0$ and suppose $R \in \tilde{M}_r(t)$ is such that $W(R) > m(r)^{(1-\epsilon)\binom{t}{2}}$. Set $a = a_R$ and $b = b_R$. All logs in this proof are base $m(r)$. Our assumptions and Lemma 4.11 imply $m(r)^{(1-\epsilon)\binom{t}{2}} < W(R) \leq m(r)^{\binom{t}{2}+t+5} \left(\frac{m(r)^2-1}{m(r)^2} \right)^a$. Consequently,

$$\left(\frac{m(r)^2}{m(r)^2-1} \right)^a < m(r)^{\epsilon \binom{t}{2} + t + 5}. \quad (8)$$

Suppose M_1 is large enough so that $t > M_1$ implies $t(1 - \frac{\epsilon}{2}) + 5 < \frac{\epsilon t^2}{4}$, and assume $t > M_1$. Taking log of both sides of (8) we obtain

$$a \log \left(\frac{m(r)^2}{m(r)^2-1} \right) \leq \epsilon \binom{t}{2} + t + 5 < \frac{\epsilon}{2} t^2 + \frac{\epsilon}{4} t^2 = \frac{3\epsilon t^2}{4},$$

where the last inequality is by assumption on M_1 . Therefore $a \leq C_1 \epsilon t^2$, for appropriate choice of $C_1 = C_1(r)$. This proves (1). For (2), note that by the definitions of $W(R)$, a , and b we have

$$W(R) \leq (m(r) - 1)^b r^a m(r)^{\binom{t}{2} - a - b}.$$

Thus our assumptions and part (1) imply that,

$$m(r)^{(1-\epsilon)\binom{t}{2}} < (m(r)-1)^b r^{C_1 \epsilon t^2} m(r)^{\binom{t}{2}-a-b} \leq (m(r)-1)^b r^{C_1 \epsilon t^2} m(r)^{\binom{t}{2}-b} = \left(\frac{m(r)-1}{m(r)} \right)^b r^{C_1 \epsilon t^2} m(r)^{\binom{t}{2}}.$$

Consequently,

$$\left(\frac{m(r)}{m(r)-1} \right)^b < m(r)^{\epsilon \binom{t}{2}} r^{C_1 \epsilon t^2}.$$

Taking log of both sides, we obtain

$$b \log \left(\frac{m(r)}{m(r)-1} \right) < \epsilon \binom{t}{2} + C_1 \epsilon t^2 \log r < \left(\frac{1}{2} + C_1 \log r \right) \epsilon t^2,$$

from which (2) follows directly for an appropriate choice of $C_2 = C_2(r)$. For (3), parts (1) and (2) yield

$$|\{ij \in E(R) : f(i, j) = m(r)\}| = \binom{t}{2} - a - b \geq \binom{t}{2} - (C_1 + C_2) \epsilon t^2 = \left(\frac{1}{2} - (C_1 + C_2) \epsilon \right) t^2 - \frac{t}{2}.$$

Setting $m = |V_0^R(\epsilon)| = |\{u \in V(R) : \mu_{m(r)}^R(u) < (1 - \sqrt{\epsilon})(t - 1)\}|$, it is clear that

$$\sum_{v \in V(R)} \mu_{m(r)}^R(v) \leq m(1 - \sqrt{\epsilon})(t - 1) + (t - m)(t - 1) = t^2 - t - \sqrt{\epsilon} m t + \sqrt{\epsilon} m.$$

On the other hand, let \mathcal{G} be the graph with vertex set $\mathcal{V} = [t]$ and edge set $\mathcal{E} = \{ij \in \binom{\mathcal{V}}{2} : f^R(ij) = m(r)\}$. Then

$$\sum_{v \in V(R)} \mu_{m(r)}^R(v) = \sum_{v \in \mathcal{V}} \mathcal{DEG}(v) = 2|\mathcal{E}| \geq 2 \left(\left(\frac{1}{2} - \epsilon(C_1 + C_2) \right) t^2 - \frac{t}{2} \right) = (1 - 2\epsilon(C_1 + C_2))t^2 - t.$$

Consequently $(1 - 2\epsilon(C_1 + C_2))t^2 - t \leq t^2 - t - \sqrt{\epsilon} m t + \sqrt{\epsilon} m$. Simplifying this we obtain

$$m \leq \frac{2\epsilon(C_1 + C_2)t^2}{\sqrt{\epsilon}(t - 1)} = 2\sqrt{\epsilon}(C_1 + C_2) \frac{t^2}{t - 1}.$$

Set $C_3 = 3(C_1 + C_2)$. It is now clear that there is M_2 such that if $t > M_2$, then $m \leq \sqrt{\epsilon} C_3 t$, so (3) holds. Therefore if $t > M = \max\{M_1, M_2\}$, (1), (2), and (3) hold. \square

4.3 Proof of the Stability Result

In this section we will prove our stability result below.

Theorem 4.13. *Fix an integer $r \geq 3$. For all $\delta > 0$ there is $0 < \epsilon < 1$ and M such that for all $t > M$ the following holds. If $R \in \tilde{M}_r(t)$ and $W(R) > m(r)^{(1-\epsilon)\binom{t}{2}}$, then R is δ -close to $\tilde{C}_r(t)$.*

The following is a consequence of Lemma 4.9, so its proof appears in the appendix along with the proof of Lemma 4.9.

Lemma 4.14. *Suppose $r \geq 3$ is an integer and $A, B, C \subset [r]$ are such that $|A| = |B| = |C| = m(r)$ and there is no violating triple $(a, b, c) \in A \times B \times C$. Then one of the following holds:*

1. r is even and $A = B = C = [m(r) - 1, r]$.
2. r is odd and for some relabeling $\{A, B, C\} = \{D, E, F\}$ one of the following holds:
 - (a) $D = F = E = [m(r) - 1, r - 1]$.
 - (b) $D = F = [m(r), r]$, $E \subset [m(r) - 1, \dots, r]$.

An immediate corollary of this is the following.

Corollary 4.15. *Suppose $r, t \geq 3$ are integers, $R \in \tilde{M}_r(t)$, and $xy, yz, xz \in \binom{[t]}{2}$ are such that $f^R(x, y) = f^R(y, z) = f^R(x, z) = m(r)$. Then one of the following holds:*

1. r is even and $c^R(xy) = c^R(yz) = c^R(xz) = [m(r) - 1, r]$.
2. r is odd and for some relabeling $\{x, y, z\} = \{u, v, z\}$ one of the following holds:
 - (a) $c^R(uv) = c^R(uw) = c^R(vw) = [m(r) - 1, r - 1]$.
 - (b) $c^R(uv) = c^R(uw) = [m(r), r]$, $c^R(vw) \subset [m(r) - 1, r]$.

Proof. $R \in \tilde{M}_r(t)$ implies there is no violating triple $(a, b, c) \in c^R(uv) \times c^R(uw) \times c^R(vw)$. Thus the corollary follows immediately by applying Lemma 4.14 to $A = c^R(uv)$, $B = c^R(uw)$ and $C = c^R(vw)$. \square

We will use the following consequence of Corollary 4.15.

Lemma 4.16. *For all integers $r \geq 3$ and $0 < \epsilon < 1$, there is M such that $t > M$ and $R \in \tilde{M}_r(t)$ implies the following. Let $V = [t]$. Then one of the following holds.*

- (i) r is even and for all $xy \in \binom{V}{2} \setminus E(V, V_0^R(\epsilon))$, $f^R(x, y) = m(r)$ implies $c^R(xy) = [m(r) - 1, r]$.
- (ii) r is odd and for all $xy \in \binom{V}{2} \setminus E(V, V_0^R(\epsilon))$, $f^R(x, y) = m(r)$ implies both of the following:
 - (a) either $r \in c^R(xy)$ or $c^R(xy) = [m(r) - 1, r - 1]$.
 - (b) either $m(r) - 1 \in c^R(xy)$ or $c^R(xy) = [m(r), r]$.

Proof. Fix an integer $r \geq 3$ and $0 < \epsilon < 1$. Choose M large enough so that $t > M$ implies $t - 2 - 2\sqrt{\epsilon}(t - 1) \geq 1$ and fix $R \in \tilde{M}_r(t)$. Suppose $xy \in \binom{V}{2} \setminus E(V_0^R(\epsilon), V)$ and $f^R(x, y) = m(r)$. Since $x, y \notin V_0^R(\epsilon)$, $\min\{\mu_{m(r)}^R(x), \mu_{m(r)}^R(y)\} \geq (1 - \sqrt{\epsilon})(t - 1)$. Therefore

$$|V \cap \Gamma_{m(r)}^R(x) \cap \Gamma_{m(r)}^R(y)| \geq t - 2 - 2\sqrt{\epsilon}(t - 1) \geq 1,$$

where the last inequality holds by our assumption on M . Thus there is $z \in V \setminus \{x, y\}$ such that $f^R(x, y) = f^R(y, z) = f^R(x, z) = m(r)$. If r is even, part (1) of Corollary 4.15 implies $c^R(xy) = [m(r) - 1, r]$, so (i) holds. If r is odd, part (2) of Corollary 4.15 implies $c^R(xy) \subset [m(r) - 1, r]$. Recall that since r is odd, $|[m(r) - 1, r]| = m(r) + 1$. Therefore, since $|c^R(xy)| = f^R(x, y) = m(r)$ and $c^R(xy) \subset [m(r) - 1, r]$, $m(r) - 1 \notin c^R(xy)$ implies $c^R(xy) = [m(r), r]$ (so (i) holds), and $r \notin c^R(xy)$ implies $c^R(xy) = [m(r) - 1, r - 1]$ (so (ii) holds). \square

We now fix some notation. Suppose $r \geq 3$ is an integer, R is an r -graph, and $u \in V(R)$. For $i \in [r]$, set

$$N_i^R(u) = \{v \in V(R) : i \in c^R(uv)\} \text{ and} \\ \deg_i^R(u) = |N_i^R(u)|.$$

Proof of Theorem 4.13. Let $r \geq 3$ be an integer, and fix $\delta > 0$. Let C_1, C_2, C_3 be as in Lemma 4.12. We will consider the cases when r is even and odd separately.

Case 1: r is even. Fix $0 < \epsilon < 1$ small enough so that $\max\{\sqrt{\epsilon}C_3, (C_1 + C_2)\epsilon\} < \frac{\delta}{2}$. Apply Lemma 4.12 to ϵ to obtain M_1 , and apply Lemma 4.16 to ϵ to obtain M_2 . Set $M = \max\{M_1, M_2\}$. Fix $t > M$ and $R \in \tilde{M}_r(t)$ such that $W(R) \geq m(r)^{(1-\epsilon)\binom{t}{2}}$. Set $V = [t]$. Let R' be the unique element of $\tilde{C}_r(t)$, that is, R' is the complete r -graph with vertex set V such that for all $xy \in \binom{V}{2}$, $c^{R'}(xy) = [m(r) - 1, r]$. We show $|\Delta(R, R')| \leq \delta t^2$.

Let $V_0 = V_0^R(\epsilon)$ and $V_1 = V \setminus V_0$. Define $A = E(V_0, V) \cup \{xy \in \binom{V}{2} : f^R(x, y) \neq m(r)\}$. Suppose $xy \in \binom{V}{2} \setminus A$. Then $xy \in \binom{V}{2} \setminus E(V, V_0)$ and $f^R(x, y) = m(r)$, so Lemma 4.16 (i) implies $c^R(xy) = [m(r) - 1, r]$. Thus $c^R(xy) = c^{R'}(xy)$ and $xy \notin \Delta(R, R')$. We have shown $\Delta(R, R') \subset A$, and consequently $|\Delta(R', R)| \leq |A|$.

We now bound $|A|$. The definition of A and parts (1), (2), and (3) of Lemma 4.12 imply

$$|A| \leq |V||V_0| + a_R + b_R \leq (\sqrt{\epsilon}C_3 + (C_1 + C_2)\epsilon)t^2.$$

By assumption on ϵ , $(\sqrt{\epsilon}C_3 + (C_1 + C_2)\epsilon)t^2 < (\frac{\delta}{2} + \frac{\delta}{2})t^2 = \delta t^2$, and consequently, $|\Delta(R, R')| \leq \delta t^2$ as desired.

Case 2: r is odd. Fix $0 < \epsilon < 1$ small enough so that $\max\{\sqrt{\epsilon}C_3, (C_1 + C_2)\epsilon, 2\sqrt{\epsilon}\} < \frac{\delta}{5}$. Apply Lemma 4.12 to ϵ to obtain M_1 and apply Lemma 4.16 to ϵ to obtain M_2 . Choose $M > \max\{M_1, M_2\}$ large enough so that $t > M$ implies $\frac{2}{\sqrt{\epsilon}t} < \frac{\delta}{5}$, $\sqrt{\epsilon}t^2 + t \leq 2\sqrt{\epsilon}t^2$, and $\frac{t^2}{t-1} < 2t$. Fix $t > M$ and $R \in \tilde{M}_r(t)$ such that $W(R) \geq m(r)^{(1-\epsilon)\binom{t}{2}}$ and set $V = [t]$. We construct an element $R' \in \tilde{C}_r(t)$, then show $|\Delta(R, R')| \leq \delta t^2$. First we choose integers $k \geq 1$, $\ell \in \{0, \dots, k\}$, and a partition $V_0, V_1, \dots, V_\ell, \dots, V_k$ of V with the following properties:

- $|V_0| \leq \sqrt{\epsilon}C_3t$,
- If $0 < l$, then for each $1 \leq i \leq l$, there is $u_i \in V$ and $B_i \subset V$ such that $V_i = (N_{m(r)-1}^R(u_i) \cap B_i) \cup \{u_i\}$,
- If $l < k$, then V_{l+1}, \dots, V_k are singletons.

Step 1: Let $V_0 = V_0^R(\epsilon)$. Note that part (3) of Lemma 4.12 implies $|V_0| \leq \sqrt{\epsilon}C_3t$. Define $B_1 = V \setminus V_0$. If there exists $u \in B_1$ such that $|N_{m(r)-1}^R(u) \cap B_1| \geq \sqrt{\epsilon}(t-1)$, then choose u_1 to be any $u \in B_1$ with $|N_{m(r)-1}^R(u) \cap B_1|$ maximal, and set $V_1 = (N_{m(r)-1}^R(u_1) \cap B_1) \cup \{u_1\}$. If $V \setminus (V_0 \cup V_1) = \emptyset$, set $k = l = 1$ and end the construction. If not, go to step 2. If no u exists in B_1 such that $|N_{m(r)-1}^R(u) \cap B_1| \geq \sqrt{\epsilon}(t-1)$, then put each element of B_1 into its own part and end the construction. This means we set $l = 0$, $k = t - |V_0|$, and let V_1, \dots, V_k partition B_1 into singletons.

Step $i+1$: Suppose $i \geq 1$ and we have chosen V_i, B_i , and u_i such that $V_i = (N_{m(r)-1}^R(u_i) \cap B_i) \cup \{u_i\}$ and $V \setminus \bigcup_{j=0}^i V_j \neq \emptyset$. Set $B_{i+1} = V \setminus \bigcup_{j=0}^i V_j$. If there exists $u \in B_{i+1}$ such that $|N_{m(r)-1}^R(u) \cap B_{i+1}| \geq \sqrt{\epsilon}(t-1)$, choose u_{i+1} to be any $u \in B_{i+1}$ with $|N_{m(r)-1}^R(u) \cap B_{i+1}|$ maximal, and set $V_{i+1} = (N_{m(r)-1}^R(u_{i+1}) \cap B_{i+1}) \cup \{u_{i+1}\}$. If $V \setminus \bigcup_{j=0}^{i+1} V_j = \emptyset$, set $k = l = i+1$ and end the construction. Otherwise go to step $i+2$. If no u exists in B_{i+1} such that $|N_{m(r)-1}^R(u) \cap B_{i+1}| \geq$

$\sqrt{\epsilon}(t-1)$, then put each element of B_{i+1} into its own part and end the construction. This means we set $l = i$, $k = t - |\bigcup_{j=0}^i V_j|$, and let V_{i+1}, \dots, V_k partition B_{i+1} into singletons. This completes the construction of the partition $V_0, V_1, \dots, V_l, \dots, V_k$. Given $xy \in \binom{V}{2}$, define

$$c^{R'}(xy) = \begin{cases} [m(r) - 1, r - 1] & \text{if } xy \in \binom{V_i}{2} \text{ some } 0 \leq i \leq l \\ [m(r), r] & \text{otherwise.} \end{cases}$$

This completes our construction of R' . We now bound $|\Delta(R, R')|$. Set

$$A = E(V_0, V) \cup \left\{ xy \in \binom{V}{2} : f^R(x, y) \neq m(r) \right\} \cup \bigcup_{i=1}^l E(\{u_i\}, V) \cup E(V_i, V \setminus \Gamma_{m(r)}^R(u_i)).$$

We first bound $|A|$, then $|\Delta(R, R') \setminus A|$. By parts (1), (2), and (3) of Lemma 4.12,

$$\left| E(V_0, V) \cup \left\{ xy \in \binom{V}{2} : f^R(x, y) \neq m(r) \right\} \right| \leq |V||V_0| + a_R + b_R \leq C_3\sqrt{\epsilon}t^2 + C_1\epsilon t^2 + C_2\epsilon t^2.$$

By construction, for each $1 \leq i \leq l$, $|V_i| \geq \sqrt{\epsilon}(t-1)$, therefore $l \leq \frac{t}{\sqrt{\epsilon}(t-1)}$. Thus

$$\left| \bigcup_{i=1}^l E(\{u_i\}, V) \right| \leq lt \leq \frac{t^2}{\sqrt{\epsilon}(t-1)} \leq \frac{2t}{\sqrt{\epsilon}},$$

where the last inequality is by assumption on M . By construction, for each $1 \leq i \leq l$, $u_i \notin V_0$ implies $|V \setminus \Gamma_{m(r)}^R(u_i)| \leq \sqrt{\epsilon}(t-1) + 1$. Therefore

$$\left| \bigcup_{i=1}^l E(V_i, V \setminus \Gamma_{m(r)}^R(u_i)) \right| \leq \sum_{i=1}^l |V_i||V \setminus \Gamma_{m(r)}^R(u_i)| \leq (\sqrt{\epsilon}(t-1)+1) \sum_{i=1}^l |V_i| \leq (\sqrt{\epsilon}(t-1)+1)t \leq 2\sqrt{\epsilon}t^2,$$

where the last inequality is by assumption on M . Combining all of this yields that

$$|A| \leq \left(\sqrt{\epsilon}C_3 + (C_1 + C_2)\epsilon + \frac{2}{\sqrt{\epsilon}t} + 2\sqrt{\epsilon} \right) t^2.$$

We now bound $|\Delta(R, R') \setminus A|$. An edge $xy \in \Delta(R, R') \setminus A$ is contained in one of the following:

- $X = \bigcup_{l+1 \leq i < j \leq k} \{xy \in E(V_i, V_j) \setminus A : c^R(xy) \neq [m(r), r]\}$.
- For some $1 \leq i \leq l$, $Y_i = \{xy \in E(V_i) \setminus A : c^R(xy) \neq [m(r) - 1, r - 1]\}$.
- For some $1 \leq i \leq \ell$ and $i < j \leq k$, $Z_{ij} = \{xy \in E(V_i, V_j) \setminus A : c^R(xy) \neq [m(r), r]\}$.

We now bound $|X|$. Define \mathcal{G} to be the graph with vertex set $\mathcal{V} = \bigcup_{j=l+1}^k V_j$ and edge set

$$\mathcal{E} = \left\{ xy \in \binom{\mathcal{V}}{2} : m(r) - 1 \in c^R(xy) \right\}.$$

By definition of X , for all $xy \in X$ we have $xy \in \binom{V}{2} \setminus E(V_0, V)$, $f^R(x, y) = m(r)$, and $c^R(xy) \neq [m(r), r]$, so Lemma 4.16 (ii)(b) implies $m(r) - 1 \in c^R(xy)$, and therefore $X \subset \mathcal{E}$. By construction, for all $u \in \mathcal{V}$, $\mathcal{DEG}(v) = |N_{m(r)-1}^R(u) \cap \mathcal{V}| < \sqrt{\epsilon}(t-1)$, thus

$$|X| \leq |\mathcal{E}| = \frac{1}{2} \sum_{v \in \mathcal{V}} \mathcal{DEG}(v) < \frac{\sqrt{\epsilon}t^2}{2}.$$

We now show each Y_i is empty. If $l = 0$ this is vacuous, so assume $l \geq 1$. Suppose for a contradiction that for some $1 \leq i \leq l$, $Y_i \neq \emptyset$. Then there is $xy \in E(V_i)$ such that $f^R(x, y) = m(r)$ and $c^R(xy) \neq [m(r) - 1, r - 1]$. By Lemma 4.16 (ii)(a), $r \in c^R(xy)$. But by construction, $m(r) - 1 \in c^R(u_i x) \cap c^R(u_i y)$. Now $(r, m(r) - 1, m(r) - 1) \in c^R(xy) \times c^R(u_i x) \times c^R(u_i y)$ is a violating triple, making $\{x, y, u_i\}$ a violating triangle, a contradiction.

We now show each Z_{ij} is empty. Suppose for a contradiction that for some $1 \leq i \leq \ell$ and some $i < j \leq k$, there is $xy \in Z_{ij}$, say with $x \in V_i$, $y \in V_j$. Then $f^R(x, y) = f^R(u_i, y) = m(r)$ and $c^R(xy) \neq [m(r), r]$. By Lemma 4.16 (ii)(b), $m(r) - 1 \in c^R(xy)$, and by construction $m(r) - 1 \in c^R(xu_i)$. Also by construction, $m(r) - 1 \notin c^R(u_i y)$, so Lemma 4.16 (ii)(b) implies $c^R(u_i y) = [m(r), r]$. But now $(r, m(r) - 1, m(r) - 1) \in c^R(u_i y) \times c^R(u_i x) \times c^R(xy)$ is a violating triple, making $\{u_i, x, y\}$ a violating triangle, a contradiction.

Combining all of this yields that $|\Delta(R, R') \setminus A| < \frac{\sqrt{\epsilon} t^2}{2}$, so

$$|\Delta(R, R')| \leq (\sqrt{\epsilon} C_3 + (C_1 + C_2)\epsilon + \frac{2}{\sqrt{\epsilon} t} + 2\sqrt{\epsilon} + \frac{\sqrt{\epsilon}}{2})t^2.$$

By our assumptions on ϵ and because $\frac{2}{\sqrt{\epsilon} t} < \frac{\delta}{5}$, $(\sqrt{\epsilon} C_3 + (C_1 + C_2)\epsilon + \frac{2}{\sqrt{\epsilon} t} + 2\sqrt{\epsilon} + \frac{\sqrt{\epsilon}}{2})t^2 < 5\frac{\delta}{5}t^2 = \delta t^2$, and $|\Delta(R, R')| < \delta t^2$ as desired. \square

5 Proof of Theorem 1.5

In this section we prove Theorem 1.5, which says that for all integers $r \geq 3$ and all $\delta > 0$, almost all elements of $M_r(n)$ are δ -close to $C_r(n)$. We begin with some key definitions. For $n, r, s \geq 3$ integers, and $\delta, \eta, d, \epsilon \geq 0$, set

$$\tilde{C}_r^\delta(s) = \{R \in \tilde{M}_r(s) : R \text{ is } \delta\text{-close to } \tilde{C}_r(s)\},$$

$$D_r(n, \delta, \eta, d) = \{G \in M_r(n) : Q_{\eta, d}(G) \neq \emptyset \text{ and for all } R \in Q_{\eta, d}(G), R \in \tilde{C}_r^\delta(t) \text{ where } t = |V(R)|\},$$

$$\tilde{E}_r(s, \epsilon) = \{R \in \tilde{M}_r(s) : W(R) \geq m(r)^{(1-\epsilon)\binom{s}{2}}\}, \text{ and}$$

$$E_r(n, \epsilon, \eta, d) = \{G \in M_r(n) : \text{for all } R \in Q_{\eta, d}(G), R \in \tilde{E}_r(t, \epsilon) \text{ where } t = |V(R)|\},$$

and recall that $C_r^\delta(n) = \{G \in M_r(n) : G \text{ is } \delta\text{-close to } C_r(n)\}$. Theorem 1.5 follows from two lemmas that we now prove. The first lemma below informally states that r -graphs in $M_r(n)$ with reduced r -graphs close to $\tilde{C}_r(t)$ are themselves close to $C_r(n)$. We will use throughout the following observations.

Fact 5.1. *Suppose $3 \leq t \leq n$. Then $\binom{\lceil n/t \rceil}{2} \leq \binom{n/t+1}{2} \leq \frac{n^2}{t^2}$ and $\lceil n/t \rceil^2 \leq (n/t + 1)^2 \leq 4n^2/t^2$.*

Lemma 5.2. *Let $r, n \geq 3$ be integers. For all $\delta > 0$, there is d_0 such that for all positive $d \leq d_0$ and $\eta \leq \gamma_{el}(d, 3)$,*

$$D_r(n, \delta/8, \eta, d) \subset C_r^\delta(n).$$

Proof. Fix $\delta > 0$ and set $d_0 = \frac{\delta}{2(2r+5)}$. Fix $d \leq d_0$ and $\eta \leq \gamma_{el}(d, 3)$, and suppose $G \in D_r(n, \delta/8, \eta, d)$. Then by definition of $D_r(n, \delta/8, \eta, d)$, $G \in M_r(n)$ and there is $R \in Q_{\eta, d}(G)$ which is $\frac{\delta}{8}$ -close to $\tilde{C}_r(t)$ where $t = |V(R)|$. Let $R' \in \tilde{C}_r(t)$ be such that R is $\frac{\delta}{8}$ -close to R' . We will build an element $G' \in C_r(n)$ such that G is δ -close to G' .

Let $\mathcal{P} = \{V_1, \dots, V_t\}$ be an η -regular partition for G such that $R = R(G, \mathcal{P}, d)$. Define

$$A = \Delta(R, R') \cup \{ij \in \binom{[t]}{2} : (V_i, V_j) \text{ is not } \eta\text{-regular for } G\}.$$

Note that $|A| \leq \frac{\delta}{8}t^2 + \eta t^2$. Define G' by $V(G') = V(G) = [n]$ and for $xy \in \binom{[n]}{2}$,

$$d^{G'}(x, y) = \begin{cases} r-1 & \text{if } xy \in E(V_i) \text{ for some } i \in [t] \\ r-1 & \text{if } xy \in E(V_i, V_j) \text{ for some } ij \in \binom{[t]}{2} \text{ such that either } ij \in A \text{ or } d^G(x, y) \notin c^{R'}(ij) \\ d^G(x, y) & \text{if } xy \in E(V_i, V_j) \text{ for some } ij \in \binom{[t]}{2} \setminus A \text{ and } d^G(x, y) \in c^{R'}(ij). \end{cases}$$

Set

$$U_r = \begin{cases} [m(r), r] & \text{if } r \text{ is odd} \\ [m(r) - 1, r] & \text{if } r \text{ is even} \end{cases} \quad \text{and} \quad L_r = \begin{cases} [m(r) - 1, r - 1] & \text{if } r \text{ is odd} \\ [m(r) - 1, r] & \text{if } r \text{ is even.} \end{cases}$$

Note that $r-1 \in U_r \cap L_r$. By the definition of $\tilde{C}_r(t)$, there is a partition $\tilde{W}_1, \dots, \tilde{W}_s$ of $[t]$ such that for all $ij \in \binom{[t]}{2}$,

$$c^{R'}(ij) = \begin{cases} L_r & \text{if } ij \in E(\tilde{W}_u) \text{ some } u \in [s] \\ U_r & \text{if } ij \in E(\tilde{W}_u, \tilde{W}_v) \text{ some } uv \in \binom{[s]}{2}. \end{cases}$$

Define a new partition W_1, \dots, W_s of $[n]$ by setting $W_u = \bigcup_{i \in \tilde{W}_u} V_i$ for each $u \in [s]$. Then by construction, for all $xy \in \binom{[n]}{2}$,

$$d^{G'}(x, y) \in \begin{cases} L_r & \text{if } xy \in E(W_u) \text{ some } u \in [s] \\ U_r & \text{if } xy \in E(W_u, W_v) \text{ some } uv \in \binom{[s]}{2}. \end{cases}$$

Therefore, $G' \in C_r(n)$ by definition. We now show $|\Delta(G, G')| \leq \delta n^2$. Recall that by definition of $Q_{\eta, d}(G)$ and $\eta, 3 \leq \frac{1}{\eta} \leq t \leq n$. Edges $xy \in \Delta(G, G')$ fall into the following categories:

- $xy \in E(V_i)$ for some $i \in [t]$. There are at most $t \binom{\lceil n/t \rceil}{2} \leq t \frac{n^2}{t^2} = \frac{n^2}{t} \leq \eta n^2$ such edges.
- $xy \in E(V_i, V_j)$ for some $ij \in A$. The number of such edges is at most

$$|A| \binom{\lceil n/t \rceil}{2} \leq |A| \frac{4n^2}{t^2} \leq \left(\frac{\delta}{8}t^2 + \eta t^2 \right) \frac{4n^2}{t^2} = \left(\frac{\delta}{2} + 4\eta \right) n^2.$$

- $xy \in E(V_i, V_j)$ for some $ij \in \binom{[t]}{2} \setminus A$ such that $d^G(x, y) \notin c^{R'}(ij)$. This means (V_i, V_j) is η -regular for G and $c^{R'}(ij) = c^R(ij)$. Because $R = R(G, \mathcal{P}, d)$, for each $l \in [r] \setminus c^R(ij)$ we have that $e_l^G(V_i, V_j) \leq d|V_i||V_j|$. Therefore there are at most $dr \binom{t}{2} \left\lceil \frac{n}{t} \right\rceil^2 \leq dr \frac{t^2}{2} \frac{4n^2}{t^2} = 2dr n^2$ such edges.

Combining these bounds with the fact that $\eta \leq d \leq d_0 = \frac{\delta}{2(2r+5)}$ yields

$$|\Delta(G, G')| \leq n^2 \left(5\eta + \frac{\delta}{2} + 2dr \right) \leq n^2 \left(5d_0 + \frac{\delta}{2} + 2d_0r \right) = n^2 \left(\frac{\delta}{2} + d_0(2r+5) \right) = \delta n^2. \quad \square$$

We now prove the second lemma. Informally, it says that most graphs in $M_r(n)$ have all their reduced graphs R with $W(R)$ quite large.

Lemma 5.3. *For all $\epsilon > 0$, there is $\beta = \beta(\epsilon)$ and $d_0 = d_0(\epsilon) > 0$, such that for all positive $d \leq d_0$ and $\eta \leq \gamma_{el}(d, 3)$, there is M such that $n \geq M$ implies*

$$\frac{|M_r(n) \setminus E_r(n, \epsilon, \eta, d)|}{m(r) \binom{n}{2}} \leq 2^{-\beta n^2}. \quad (9)$$

Proof. All logs in this proof are base 2. Fix $\epsilon > 0$ and set $\beta = \frac{\epsilon \log m(r)}{8}$. Define

$$F(x) = 5x \log r + 4r(H(x) + x) - 2\beta,$$

and choose $d_0 < \frac{1}{3}$ small enough so that $F(d_0) < -\beta$. Recall that for $0 \leq y \leq x \leq \frac{1}{2}$, $H(y) \leq H(x)$, so for any $0 \leq y \leq x \leq d_0$, $F(y) \leq F(x)$. Fix $d \leq d_0$ and $\eta \leq \gamma_{el}(d, 3) \leq d$. Set $N = CM(\frac{1}{\eta}, \eta)$ and define

$$C = \log(N - \frac{1}{\eta} + 1) + \log(2^r) \binom{N}{2} + (H(\eta) + \eta)N^2,$$

$$C' = \log N + \frac{\log m(r)}{2} + N^2 \log r, \text{ and}$$

$$C'' = 5\eta \log r + 4r(H(d) + d) - 4\beta.$$

For any integer n , define

$$S(n) = n \log(N) + (H(\eta) + \eta)N^2 + (5\eta \log r)n^2 + 4r(H(d) + d)n^2 + nN^2 \log r,$$

$$S'(n) = S(n) + \log(N - \frac{1}{\eta} + 1) + \binom{N}{2} \log(2^r), \text{ and}$$

$$S''(n) = S'(n) - 4\beta n^2 + \frac{\log m(r)}{2} n.$$

Notice that $S''(n) = C + C'n + C''n^2$ and $C'' \leq F(d) - 2\beta$. Choose $M \geq N$ large enough so that $n > M$ implies $S''(n) < (C'' + 2\beta)n^2 \leq F(d)n^2$. We show $n > M$ implies (9) holds. Fix $n > M$. Our assumptions on $d \leq d_0$ and M imply $S''(n) < F(d)n^2 \leq F(d_0)n^2 < -\beta n^2$, so it suffices to show

$$\frac{|M_r(n) \setminus E_r(n, \epsilon, \eta, d)|}{m(r) \binom{n}{2}} \leq 2^{S''(n)}. \quad (10)$$

By definition of $E(n, \epsilon, \eta, d)$, we have $G \in M_r(n) \setminus E_r(n, \epsilon, \eta, d)$ if and only if there is $\frac{1}{\eta} \leq t \leq N$ and $R \in \tilde{M}_r(t)$ such that $R \in Q_{\eta, d}(G)$ and $W(R) < m(r)^{(1-\epsilon)\binom{t}{2}}$. We give an upper bound for the number of such G .

Fix some $\frac{1}{\eta} \leq t \leq N$ and $R \in \tilde{M}_r(t)$ such that $W(R) < m(r)^{(1-\epsilon)\binom{t}{2}}$. All $G \in M_r(n)$ such that $R \in Q_{\eta, d}(G)$ can be constructed as follows:

- Choose an equipartition of $[n]$ into t pieces V_1, \dots, V_t . There are at most $t^n \leq N^n$ such partitions. Note that for each $i \in [t]$, $|V_i| \leq \lceil n/t \rceil$.
- Choose $J \subset \binom{[t]}{2}$ to be the set of ij such that (V_i, V_j) is not η -regular for G . There are at most $\binom{\binom{t}{2}}{\eta \binom{t}{2}} 2^{\eta \binom{t}{2}} \leq 2^{H(\eta)t^2 + \eta t^2} \leq 2^{(H(\eta) + \eta)N^2}$ ways to do this.
- Choose $d^G(x, y)$ for each $xy \in E(V_i)$ and $i \in [t]$. There are at most $r^{t \binom{\lceil n/t \rceil}{2}} \leq r^{t \frac{n^2}{t^2}} = r^{\frac{n^2}{t}} \leq r^{\eta n^2}$ ways to do this.
- Choose $d^G(x, y)$ for each $xy \in E(V_i, V_j)$ where $ij \in J$. The number of ways to do this is at most $(r^{\lceil n/t \rceil^2})^{\eta t^2} \leq r^{\frac{4n^2}{t^2} \eta t^2} = r^{4\eta n^2}$.
- Choose $d^G(x, y)$ for each $xy \in E(V_i, V_j)$ where $ij \in I = \binom{[t]}{2} \setminus J$. For each $ij \in I$, (V_i, V_j) is η -regular, so the colors for edges in $E(V_i, V_j)$ can be chosen as follows:

- (a) For each $s \notin c^R(ij)$, choose a subset of $E(V_i, V_j)$ of size at most $d|V_i||V_j|$ to have color s . The number of ways to do this is at most

$$\left(\binom{\lceil n/t \rceil^2}{d \lceil n/t \rceil^2} 2^{d \lceil n/t \rceil^2} \right)^r \leq 2^{r(H(d)\lceil n/t \rceil^2 + d \lceil n/t \rceil^2)} \leq 2^{\frac{4rn^2}{t^2}(H(d)+d)}.$$

- (b) Assign colors from $c^R(ij)$ to the rest of the edges in $E(V_i, V_j)$. The number of ways to do this is at most $f^R(i, j)^{\lceil n/t \rceil^2} \leq f^R(i, j)^{\frac{n^2}{t^2} + \frac{2n}{t} + 1} \leq f^R(i, j)^{\frac{n^2}{t^2}} r^{\frac{2n}{t} + 1} \leq f^R(i, j)^{\frac{n^2}{t^2}} r^n$ ways to do this.

Therefore, the total number of ways to choose $d^G(x, y)$ for $xy \in E(V_i, V_j)$ where $ij \in I$ is at most

$$\prod_{ij \in I} 2^{(H(d)+d)\frac{4rn^2}{t^2}} f^R(i, j)^{\frac{n^2}{t^2}} r^n \leq 2^{4r(H(d)+d)n^2} r^{nN^2} \left(\prod_{ij \in I} f^R(i, j)^{\frac{n^2}{t^2}} \right) \leq 2^{4r(H(d)+d)n^2} r^{nN^2} W(R)^{\frac{n^2}{t^2}}.$$

By the assumption, $W(R) < m(r)^{(1-\epsilon)\binom{t}{2}}$. Therefore

$$W(R)^{\frac{n^2}{t^2}} < m(r)^{(1-\epsilon)\binom{t}{2}\frac{n^2}{t^2}} < m(r)^{(1-\epsilon)\left(\binom{n}{2} + \frac{n}{2}\right)}.$$

Combining the above yields that the number of $G \in M_r(n)$ with $R \in Q_{\eta, d}(G)$ is at most

$$N^n 2^{(H(\eta)+\eta)N^2} r^{5\eta n^2} 2^{4r(H(d)+d)n^2} r^{nN^2} m(r)^{(1-\epsilon)\left(\binom{n}{2} + \frac{n}{2}\right)} = 2^{S(n)} m(r)^{(1-\epsilon)\left(\binom{n}{2} + \frac{n}{2}\right)}.$$

The number of $R \in \tilde{M}_r(t)$ with $\frac{1}{\eta} \leq t \leq N$ is at most $(N - \frac{1}{\eta} + 1)|\tilde{M}_r(N)|$, so

$$\begin{aligned} |M_r(n) \setminus E_r(n, \epsilon, \eta, d)| &< (N - \frac{1}{\eta} + 1)|\tilde{M}_r(N)| 2^{S(n)} m(r)^{(1-\epsilon)\left(\binom{n}{2} + \frac{n}{2}\right)} \\ &< (N - \frac{1}{\eta} + 1)(2^r)^{\binom{N}{2}} 2^{S(n)} m(r)^{(1-\epsilon)\left(\binom{n}{2} + \frac{n}{2}\right)} = 2^{S'(n)} m(r)^{(1-\epsilon)\left(\binom{n}{2} + \frac{n}{2}\right)}. \end{aligned}$$

Thus

$$\frac{|M_r(n) \setminus E_r(n, \epsilon, \eta, d)|}{m(r)^{\binom{n}{2}}} < \frac{2^{S'(n)} m(r)^{(1-\epsilon)\left(\binom{n}{2} + \frac{n}{2}\right)}}{m(r)^{\binom{n}{2}}} = 2^{S''(n)}.$$

We have shown that $n > M$ implies (10) holds, so we are done. \square

Proof of Theorem 1.5. Fix $\delta > 0$. Apply Theorem 4.13 to $\frac{\delta}{8}$ to obtain ϵ and $M_{4.13}$. Apply Lemma 5.2 to δ to obtain $(d_0)_{5.2}$. Apply Lemma 5.3 to ϵ to obtain β and $(d_0)_{5.3}$. Let $d_0 = \min\{(d_0)_{5.2}, (d_0)_{5.3}\}$. Apply Lemma 5.3 to $d = d_0 \leq (d_0)_{5.3}$ and $\eta = \min\{\gamma_{el}(d, 3), \frac{1}{M_{4.13}}\}$ to obtain $M_{5.3}$. Set $M = \max\{CM(\eta, \frac{1}{\eta}), M_{5.3}\}$ and fix $n > M$. Lemma 5.3 implies

$$\frac{|M_r(n) \setminus E_r(n, \epsilon, \eta, d)|}{m(r)^{\binom{n}{2}}} \leq 2^{-\beta n^2}. \quad (11)$$

We now show $E_r(n, \epsilon, \eta, d) \subset D_r(n, \delta/8, \eta, d)$. Suppose $G \in E_r(n, \epsilon, \eta, d)$. We need to show that $Q_{\eta, d}(G) \neq \emptyset$ and for all $R \in Q_{\eta, d}(G)$, $R \in \tilde{C}_r^{\delta/8}(t)$ where $t = |V(R)|$. As $n > CM(\eta, \frac{1}{\eta})$, by Theorem 4.2, we have $Q_{\eta, d}(G) \neq \emptyset$. Suppose $R \in Q_{\eta, d}(G)$ and set $t = |V(R)|$. By definition of $E_r(n, \epsilon, \eta, d)$, $R \in \tilde{E}_r(t, \epsilon)$. Theorem 4.13 and our assumptions on η imply that $R \in \tilde{C}_r^{\delta/8}(t)$,

so $E_r(n, \epsilon, \eta, d) \subset D_r(n, \delta/8, \eta, d)$. Lemma 5.2 implies $D_r(n, \delta/8, \eta, d) \subset C_r^\delta(n)$. Combining these inclusions with (11) we have that

$$\frac{|M_r(n) \setminus C_r^\delta(n)|}{m(r) \binom{n}{2}} \leq 2^{-\beta n^2}.$$

By Remark 1.9, $|M_r(n)| \geq m(r) \binom{n}{2}$, so

$$\frac{|M_r(n) \setminus C_r^\delta(n)|}{|M_r(n)|} \leq \frac{|M_r(n) \setminus C_r^\delta(n)|}{m(r) \binom{n}{2}} \leq 2^{-\beta n^2},$$

which completes our proof of Theorem 1.5. \square

6 Proof of Theorem 1.2

In this section we prove Theorem 1.2, which says that for all even integers $r \geq 4$, almost all G in $M_r(n)$ are in $C_r(n)$. The outline of the proof is as follows. Given $\epsilon > 0$ and integers $r, n \geq 3$, define

$$\begin{aligned} A_r(n, \epsilon) &= \{G \in M_r(n) : \exists x \in [n] \text{ such that for some } l \in [m(r) - 2], |N_l^G(x)| \geq \epsilon n\}, \\ A'_r(n, \epsilon) &= \{G \in M_r(n) \setminus A_r(n, \epsilon) : \exists xy \in \binom{[n]}{2} \text{ with } d^G(x, y) \in [m(r) - 2]\}. \end{aligned}$$

For all $\epsilon > 0$, $n \in \mathbb{N}$, and even integers $r \geq 4$, we have that $M_r(n) = C_r(n) \cup A_r(n, \epsilon) \cup A'_r(n, \epsilon)$, and thus $M_r(n) \setminus C_r(n) = A_r(n, \epsilon) \cup A'_r(n, \epsilon)$. We will show that when r is even, there are $\epsilon > 0$ and $\beta > 0$ such that for large n , $|A_r(n, \epsilon) \cup A'_r(n, \epsilon)| \leq 2^{-\beta n} |M_r(n)|$, from which Theorem 1.2 will follow. We do this in two lemmas, one for each of the sets A_r and A'_r defined above. The first lemma will apply to all $r \geq 3$, while the second will apply only to even $r \geq 4$.

Lemma 6.1. *For all integers $r \geq 3$ and all $\epsilon > 0$ there is $\beta > 0$ and M such that $n > M$ implies*

$$|A_r(n, \epsilon)| \leq 2^{-\beta n^2} |C_r(n)|. \quad (12)$$

Proof. Let $r \geq 3$ be an integer and fix $\epsilon > 0$. By Remark 1.9, it suffices to find $\beta > 0$ and M such that $n > M$ implies

$$|A_r(n, \epsilon)| \leq 2^{-\beta n^2} m(r) \binom{n}{2}.$$

Choose $T > 0$ large enough so that $\frac{\epsilon^2 T^2}{64} - \frac{\epsilon T}{8} \geq 1$, then choose $0 < \delta < \min\{\frac{1}{T}, \frac{\epsilon^2}{128}\}$. Apply Theorem 4.13 to δ to obtain $\epsilon_{4.13}$ and $M_{4.13}$. Apply Lemma 5.3 to $\epsilon_{4.13}$ to obtain d_0 and $\beta > 0$. Choose $d \leq d_0$ and $\eta < \min\{\delta, \gamma_{el}(d, 3), \frac{\epsilon}{2}, d, \frac{1}{M_{4.13}}\}$. Apply Lemma 5.3 to this d and η to obtain $M_{5.3}$. Choose $M \geq \max\{M_{5.3}, CM(\frac{1}{\eta}, \eta)\}$. Lemma 5.3 implies that for all $n > M$,

$$\frac{|M_r(n) \setminus E_r(n, \epsilon_{4.13}, \eta, d)|}{m(r) \binom{n}{2}} \leq 2^{-\beta n^2}.$$

Therefore, it suffices to prove that $n > M$ implies that $A_r(n, \epsilon) \subset M_r(n) \setminus E_r(n, \epsilon_{4.13}, \eta, d)$. Fix $n > M$ and suppose for a contradiction that there is some $G \in A_r(n, \epsilon) \cap E_r(n, \epsilon_{4.13}, \eta, d)$. Since $G \in A_r(n, \epsilon)$, there is $x \in [n]$ and $l \in [m(r) - 2]$ such that $|N_l^G(x)| \geq \epsilon n$. Because $n > CM(\frac{1}{\eta}, \eta)$, by Theorem 4.2, there is $R \in Q_{\eta, d}(G)$. Also, $G \in E_r(n, \epsilon_{4.13}, \eta, d)$ implies that $W(R) \geq m(r)^{(1-\epsilon_{4.13})\binom{t}{2}}$ where $t = |V(R)|$. Then $t \geq \frac{1}{\eta} > M_{4.13}$ implies that there is $R' \in \tilde{C}_r(t)$ such that $|\Delta(R, R')| \leq \delta t^2$.

Let $\mathcal{P} = \{V_1, \dots, V_t\}$ be an η -regular partition for G such that $R = R(G, \mathcal{P}, d)$, and define $\Sigma = \{i \in [t] : |N_l^G(x) \cap V_i| \geq \frac{\epsilon}{2}|V_i|\}$. We have that

$$\epsilon n \leq |N_l^G(x)| = \sum_{i \in \Sigma} |N_l^G(x) \cap V_i| + \sum_{i \notin \Sigma} |N_l^G(x) \cap V_i| \leq |\Sigma| \left\lceil \frac{n}{t} \right\rceil + (t - |\Sigma|) \frac{\epsilon}{2} \left\lceil \frac{n}{t} \right\rceil. \quad (13)$$

Note that $n \geq 3$ implies $\lceil n/t \rceil \leq \frac{4n}{3t}$, so

$$|\Sigma| \left\lceil \frac{n}{t} \right\rceil + (t - |\Sigma|) \frac{\epsilon}{2} \left\lceil \frac{n}{t} \right\rceil = |\Sigma| \left(1 - \frac{\epsilon}{2}\right) \left\lceil \frac{n}{t} \right\rceil + \frac{\epsilon t}{2} \left\lceil \frac{n}{t} \right\rceil \leq |\Sigma| \left(1 - \frac{\epsilon}{2}\right) \frac{4n}{3t} + \frac{\epsilon t}{2} \frac{4n}{3t} = |\Sigma| \left(1 - \frac{\epsilon}{2}\right) \frac{4n}{3t} + \frac{2\epsilon n}{3}.$$

Combining this with (13), yields that $\epsilon n \leq |\Sigma| \left(1 - \frac{\epsilon}{2}\right) \frac{4n}{3t} + \frac{2\epsilon n}{3}$. Rearranging this, we obtain that $|\Sigma| \geq (\epsilon n/3) / ((4n/3t)(1 - \frac{\epsilon}{2})) = \frac{\epsilon t}{4(1 - \frac{\epsilon}{2})} \geq \frac{\epsilon t}{4}$. Set

$$I = \{ij \in E(\Sigma) : (V_i, V_j) \text{ is } \eta\text{-regular for } G \text{ and } c^R(ij) = c^{R'}(ij)\}.$$

Applying that \mathcal{P} is an η -regular partition for G , that $|\Delta(R, R')| \leq \delta t^2$, and that $\frac{\epsilon t}{4} \leq |\Sigma|$ yields

$$|I| \geq \binom{\frac{\epsilon t}{4}}{2} - \eta t^2 - \delta t^2 = t^2 \left(\frac{\epsilon^2}{32} - \eta - \delta \right) - \frac{\epsilon t}{8} \geq t^2 \left(\frac{\epsilon^2}{32} - 2\delta \right) - \frac{\epsilon t}{8}, \quad (14)$$

where the last inequality is because $\eta < \delta$. By our assumptions, $t \geq \frac{1}{\delta} \geq T$ and $\delta < \frac{\epsilon^2}{128}$. These facts imply the right hand side of (14) is at least $\frac{\epsilon^2 T}{64} - \frac{\epsilon T}{8} \geq 1$. Thus $I \neq \emptyset$.

Take $ij \in I$ and let $W_i = N_l^G(x) \cap V_i$ and $W_j = N_l^G(x) \cap V_j$. Since $\eta < \frac{\epsilon}{2}$ and (V_i, V_j) is η -regular for G , we have $\rho_{r-1}^G(W_i, W_j) \geq \rho_{r-1}^G(V_i, V_j) - \eta$. Because $c^R(ij) = c^{R'}(ij)$, we have that $r - 1 \in c^R(ij)$. Therefore, by definition of R , $\rho_{r-1}^G(V_i, V_j) \geq d$, so $\rho_{r-1}^G(W_i, W_j) \geq d - \eta > 0$, where the last inequality is by the assumption on η . Therefore, there is $(x_i, x_j) \in W_i \times W_j$ such that $d^G(x_i, x_j) = r - 1$. But now $d^G(x, x_i) = l$, $d^G(x, x_j) = l$, and $d^G(x_i, x_j) = r - 1$ implies that $\{x, x_i, x_j\}$ is a violating triangle in G , a contradiction. This finishes the proof that $A_r(n, \epsilon) \subset M_r(n) \setminus E_r(n, \epsilon_{4.13}, \eta, d)$, so we are done. \square

Lemma 6.2. *Let $r \geq 4$ be an even integer. There are $\epsilon, \beta > 0$ and N such that $n > N$ implies*

$$|A'_r(n, \epsilon)| \leq 2^{N^2 - \beta n} |C_r(n)|. \quad (15)$$

Proof. All logs are base 2. Set $\beta = \frac{1}{2}(\log(m(r)^2) - \log(m(r)^2 - 2))$ and choose $\epsilon > 0$ small enough so that

$$2r(H(\epsilon) + \epsilon) - 2\beta < -\frac{3\beta}{2}. \quad (16)$$

Given an integer k , set

$$F(k) = \log \binom{k}{2} + \log(m(r) - 2) - 2 \log(m(r)^2 - 2) + 2rk(H(\epsilon) + \epsilon) \text{ and}$$

$$F'(k) = F(k) + 3 \log m(r).$$

By Corollary 1.6, there is n_0 such that $n > n_0$ implies

$$|M_r(n)| \leq 2^{(n-1)^2 - \beta n} m(r) \binom{n}{2} = 2^{(n-1)^2 - \beta n} |C_r(n)|. \quad (17)$$

By (16) and definition of $F'(n)$, there is n_1 such that $n > n_1$ implies

$$F'(n) - 2\beta n + 5 < -\beta n. \quad (18)$$

Apply Lemma 6.1 to ϵ to obtain $M_{6.1}$ and $\beta_{6.1}$. Choose $N > \max\{M_{6.1}, n_0, n_1\}$ large enough so $\beta_{6.1}(N-2)^2 > 1$. We show by induction that for all $n \geq N$, (15) holds. We begin with the base cases $n = N$ and $n = N+1$. Combining (17) with the fact that for all n , $A'_r(n, \epsilon) \subset M_r(n)$ yields

$$\begin{aligned} |A'_r(N, \epsilon)| &\leq |M_r(N)| \leq 2^{(N-1)^2 - \beta N} |C_r(N)| < 2^{N^2 - \beta N} |C_r(N)| \text{ and} \\ |A'_r(N+1, \epsilon)| &\leq |M_r(N+1)| \leq 2^{N^2 - \beta(N+1)} |C_r(N+1)|. \end{aligned}$$

Therefore (15) holds for $n = N$ and $n = N+1$. Suppose now $n \geq N+2$ and (15) holds for all m such that $N \leq m \leq n-1$. We show it holds for n . We can construct any element G of $A'_r(n, \epsilon)$ as follows.

- Choose a pair of elements $xy \in \binom{[n]}{2}$. There are $\binom{n}{2}$ ways to do this.
- Choose $d^G(x, y) \in [m(r) - 2]$. There are $m(r) - 2$ ways to do this.
- Put a structure on $[n] \setminus \{x, y\}$. There are at most $|M_r(n-2)|$ ways to do this.
- For each $l \in [m(r) - 2]$, choose $N_l(x)$ and $N_l(y)$. Since G is not in $A_r(n, \epsilon)$, for each $l \in [m(r) - 2]$, $\max\{|N_l(x)|, |N_l(y)|\} \leq \epsilon n$. Therefore, there are at most $\binom{n}{\epsilon n} 2^{\epsilon n} 2^{(m(r)-2)} \leq 2^{2rn(H(\epsilon)+\epsilon)}$ ways to do this.
- For each $z \in [n] \setminus (\{x, y\} \cup \bigcup_{l=1}^{m(r)-2} N_l(x) \cup N_l(y))$, choose $d^G(x, z)$ and $d^G(y, z)$. Note that $(d^G(x, z), d^G(y, z))$ must be chosen from $[m(r) - 1, r] \times [m(r) - 1, r] \setminus \{(m(r) - 1, r), (r, m(r) - 1)\}$, so there are at most $m(r)^2 - 2$ choices.

Combining all of this we obtain that $|A'_r(n, \epsilon)|$ is at most

$$\binom{n}{2} (m(r) - 2) 2^{2rn(H(\epsilon)+\epsilon)} (m(r)^2 - 2)^{n-2} |M_r(n-2)| = 2^{F(n)} (m(r)^2 - 2)^n |M_r(n-2)|. \quad (19)$$

Because $M_r(n-2) = C_r(n-2) \cup A_r(n-2, \epsilon) \cup A'_r(n-2, \epsilon)$,

$$|M_r(n-2)| = |C_r(n-2)| + |A_r(n-2, \epsilon)| + |A'_r(n-2, \epsilon)|.$$

Lemma 6.1 implies $|A_r(n-2, \epsilon)| \leq |C_r(n-2)| 2^{-\beta_{6.1}(n-2)^2}$, and our induction hypothesis implies $|A'_r(n-2, \epsilon)| \leq |C_r(n-2)| 2^{N^2 - \beta(n-2)}$. Remark 1.9 implies $|C_r(n)| = m(r)^{2n-3} |C_r(n-2)|$. Combining these facts with (19), we obtain that

$$\begin{aligned} |A'_r(n, \epsilon)| &\leq 2^{F(n)} (m(r)^2 - 2)^n (1 + 2^{-\beta_{6.1}(n-2)^2} + 2^{N^2 - \beta(n-2)}) |C_r(n-2)| \\ &= 2^{F(n)} (m(r)^2 - 2)^n m(r)^{-2n+3} (1 + 2^{-\beta_{6.1}(n-2)^2} + 2^{N^2 - \beta(n-2)}) |C_r(n)| \\ &= 2^{F'(n) - 2\beta n} (1 + 2^{-\beta_{6.1}(n-2)^2} + 2^{N^2 - \beta(n-2)}) |C_r(n)|. \end{aligned} \quad (20)$$

By assumption on N , $-\beta_{6.1}(n-2)^2 < -1$, so we have that

$$1 + 2^{-\beta_{6.1}(n-2)^2} + 2^{N^2 - \beta(n-2)} \leq 2 + 2^{N^2 - \beta(n-2)} \leq \begin{cases} 4 & \text{if } N^2 - \beta(n-2) \leq 1, \\ 2(2^{N^2 - \beta(n-2)}) & \text{if } N^2 - \beta(n-2) > 1. \end{cases}$$

Combining this with (20) yields that

$$|A'_r(n, \epsilon)| \leq \begin{cases} 2^{F'(n)-2\beta n+2}|C_r(n)| & \text{if } N^2 - \beta(n-2) \leq 1 \text{ and} \\ 2^{F'(n)-3\beta n+N^2+5}|C_r(n)| & \text{if } N^2 - \beta(n-2) > 1. \end{cases}$$

In both cases we have $|A'_r(n, \epsilon)| \leq 2^{N^2+F'(n)-2\beta n+5}|C_r(n)|$, so by (18), $|A'_r(n, \epsilon)| \leq 2^{N^2-\beta n}|C_r(n)|$. This completes the induction. \square

Proof of Theorem 1.2. Fix $r \geq 4$ an even integer. Apply Lemma 6.2 to obtain $\epsilon_{6.2}$, $\beta_{6.2}$ and $N_{6.2}$. Apply Lemma 6.1 to $\epsilon_{6.2}$ to obtain $\beta_{6.1}$ and $M_{6.1}$. Set $\epsilon = \epsilon_{6.2}$ and $\beta = \frac{1}{2}\beta_{6.2}$. Let M' be large enough so that $n > M'$ implies $2^{-\beta_{6.1}n^2} + 2^{N_{6.2}^2-\beta_{6.2}n} < 2^{-\beta n}$. Set $M = \max\{M_{6.1}, N_{6.2}, M'\}$. For all n , by definition, $M_r(n) \setminus C_r(n) = A_r(n, \epsilon) \cup A'_r(n, \epsilon)$. Therefore, when $n > M$ our assumptions imply

$$|M_r(n) \setminus C_r(n)| = |A_r(n, \epsilon)| + |A'_r(n, \epsilon)| \leq (2^{-\beta_{6.1}n^2} + 2^{N_{6.2}^2-\beta_{6.2}n})|C_r(n)| < 2^{-\beta n}|C_r(n)|.$$

Rearranging yields that $|C_r(n)| \geq |M_r(n)|(1 - 2^{-\beta n})$, as desired. \square

7 Concluding remarks

- When r is odd, the error term in Corollary 1.6 cannot be strengthened from $o(n^2)$ to $o(1)$ (or even to $O(n)$), as in Corollary 1.3. This can be seen by constructing a large collection of elements of $M_r(n)$, which will show that $|M_r(n)|$ is at least $m(r) \binom{n}{2}^{\Omega(n \log_{m(r)} n)}$. Fix n a sufficiently large integer. Define a *matching* to be a set $S \subset \binom{[n]}{2}$ such that no two elements of S have nonempty intersection. Given a matching S , define $A(S)$ to be the set of simple complete r -graphs G such that for each $xy \in S$, $d^G(x, y) = m(r) - 1$ and for each $xy \in \binom{[n]}{2} \setminus S$, $d^G(x, y) \in [m(r), r]$. One can easily verify that for any matching S , no element of $A(S)$ contains a violating triangle, so $A(S) \subset M_r(n)$, and that given another matching $S' \neq S$, $A(S) \cap A(S') = \emptyset$. Further, it is clear that $|A(S)| = m(r) \binom{n}{2}^{-|S|}$ and $|S| \leq \frac{n}{2}$, so $|A(S)| \geq m(r) \binom{n}{2}^{-\frac{n}{2}}$. Finally, note that there are at least $(\frac{n}{2})!$ distinct matchings on $[n]$. This and Stirling's approximation yields that

$$|M_r(n)| \geq \left(\frac{n}{2}\right)! m(r) \binom{n}{2}^{-\frac{n}{2}} = m(r) \binom{n}{2}^{\Omega(n \log_{m(r)} n)}.$$

Combining this with Theorem 1.6, the best bounds we have obtained for $|M_r(n)|$ are

$$m(r) \binom{n}{2}^{\Omega(n \log n)} \leq |M_r(n)| \leq m(r) \binom{n}{2}^{o(n^2)}.$$

We conjecture that in fact, $|M_r(n)| = m(r) \binom{n}{2}^{\Theta(n \log n)}$.

- It is impossible to extend Theorem 1.2 to the case when r is odd. Indeed, one can show that

$$|C_r(n)| \leq (1 - r^{-66r^2})|M_r(n)|.$$

The proof of this (see the appendix of [41]) in fact shows that there is a \mathcal{L}_r -sentence ψ such that for all n , $C_r(n) \subset \{G \in M_r(n) : G \models \neg\psi\}$, and

$$|C_r(n)| \leq r^{65r^2} |\{G \in M_r(n) : G \models \psi\}|. \quad (21)$$

Suppose we knew that for some $\alpha > 0$, $|C_r(n)| \geq \alpha|M_r(n)|$ for all sufficiently large n . Then since for all $G \in C_r(n)$, $G \models \neg\psi$ we would know that

$$|\{G \in M_r(n) : G \models \neg\psi\}| \geq \alpha|M_r(n)|$$

Dividing both sides of this by $|M_r(n)|$ gives us that $\mu^{M_r}(\neg\psi) \geq \alpha$, and therefore $\mu^{M_r}(\psi) \leq 1 - \alpha$. By dividing the quantities in (21) by $|M_r(n)|$, we obtain that $|C_r(n)|/|M_r(n)| \leq \mu^{M_r}(\psi)r^{65r^2}$, and therefore $\alpha/r^{65r^2} \leq \mu^{M_r}(\psi)$. Combining these inequalities, we would have that

$$0 < \frac{\alpha}{r^{65r^2}} \leq \mu^{M_r}(\psi) \leq 1 - \alpha < 1,$$

that is, $\mu^{M_r}(\psi) \notin \{0, 1\}$. Therefore, if we could show such an α existed, we would know that M_r had no labeled first-order 0-1 law. However, we do not know that such an α exists. In fact it seems likely to the authors that instead, $\lim_{n \rightarrow \infty} |C_r(n)|/|M_r(n)| = 0$.

Acknowledgment. We are very grateful to a referee for carefully reading the paper and providing many detailed comments that helped improve the presentation.

8 Appendix

Proof of Lemma 4.9. Given an integer $r \geq 3$, subsets $A, B, C \subset [r]$, and integers x, y , write $H_r(A, B, C, x, y)$ to mean A, B, C, x, y satisfy the hypotheses of the lemma for r . We show by induction on r that for all $r \geq 3$, $A, B, C \subset [r]$, and $x, y \in \mathbb{N}$, $H_r(A, B, C, x, y)$ implies $A \times B \times C$ contains a violating triple.

Case $r = 3$: Fix $A, B, C \subset [3]$, and integers x, y such that $H_3(A, B, C, x, y)$. As $m(3) = 2$ and $3 - m(3) = 1$, we have $|A| = 3$, $x = 1$, $|B| \geq 2$, $0 \leq y \leq 1$, and $|C| \geq \max\{2 - 1 - y + 2, 1\} = \max\{3 - y, 1\}$. If $y = 0$, then $|B| = 2$ and $|C| \geq 3 - y = 3$, contradicting that $|B| \geq |C|$. Therefore, $y = 1$, $|B| = 3$, and $|C| \geq 2$. This implies that $A = B = [3]$ and $C \cap \{1, 3\} \neq \emptyset$, so either $(3, 1, 1)$ or $(1, 1, 3)$ is in $A \times B \times C$, and we are done.

Case $r \geq 3$: Let $r > 3$ and suppose by induction that the claim holds for all $3 \leq r' < r$. Fix $A, B, C \subset [r]$ and integers x, y such that $H_r(A, B, C, x, y)$. Notice this implies $x \geq y \geq 0$ and $x \geq 1$. Suppose $A, B, C \subset [r - 1]$. Then

$$|A| = m(r) + x = \begin{cases} m(r-1) + x + 1 & \text{if } r \text{ is even} \\ m(r-1) + x & \text{if } r \text{ is odd,} \end{cases}$$

$$|B| = m(r) + y = \begin{cases} m(r-1) + y + 1 & \text{if } r \text{ is even} \\ m(r-1) + y & \text{if } r \text{ is odd} \end{cases}$$

and

$$|C| \geq \begin{cases} \max\{m(r) - x - y, 1\} = \max\{m(r-1) - (x+1) - (y+1) + 3, 1\} & \text{if } r \text{ is even} \\ \max\{m(r) - x - y + 2, 1\} = \max\{m(r-1) - x - y + 2, 1\} & \text{if } r \text{ is odd.} \end{cases}$$

Thus, $H_{r-1}(A, B, C, x, y)$ holds when r is odd, and $H_{r-1}(A, B, C, x+1, y+1)$ holds when r is even, so we are done by the induction hypothesis. Assume now one of A, B , or C contains r . Let $a = \min A$, $b = \min B$, $c = \min C$, $a' = \max A$, $b' = \max B$, and $c' = \max C$. Our assumptions imply that

$$a \leq r - |A| + 1 = r - (m(r) + x) + 1 = \begin{cases} m(r) - 1 - x & \text{if } r \text{ is even} \\ m(r) - x & \text{if } r \text{ is odd,} \end{cases} \quad (22)$$

and

$$b \leq r - |B| + 1 \leq r - (m(r) + y) + 1 = \begin{cases} m(r) - 1 - y & \text{if } r \text{ is even} \\ m(r) - y & \text{if } r \text{ is odd.} \end{cases}$$

Thus

$$a + b \leq \begin{cases} m(r) - 1 - x + m(r) - 1 - y = r - x - y & \text{if } r \text{ is even} \\ m(r) - x + m(r) - y = r - x - y + 1 & \text{if } r \text{ is odd.} \end{cases}$$

If

$$c' > \begin{cases} r - x - y & \text{if } r \text{ is even} \\ r - x - y + 1 & \text{if } r \text{ is odd,} \end{cases}$$

then (a, b, c') is a violating triple and we are done. So assume

$$c' \leq \begin{cases} r - x - y & \text{if } r \text{ is even} \\ r - x - y + 1 & \text{if } r \text{ is odd.} \end{cases} \quad (23)$$

Note that

$$c \leq \begin{cases} r - x - y - |C| + 1 \leq r - x - y - (m(r) - x - y) + 1 = m(r) - 1 & \text{if } r \text{ is even} \\ r - x - y + 1 - |C| + 1 \leq r - x - y + 1 - (m(r) - x - y + 2) + 1 = m(r) - 1 & \text{if } r \text{ is odd.} \end{cases}$$

Therefore

$$c + a \leq \begin{cases} m(r) - 1 + m(r) - 1 - x = r - x & \text{if } r \text{ is even} \\ m(r) - 1 + m(r) - x = r - x & \text{if } r \text{ is odd.} \end{cases}$$

If $b' > r - x$, then (a, b', c) is a violating triple and we are done. So assume $b' \leq r - x$. Because $x \geq 1$, this implies $r \notin B$. Further,

$$b \leq \begin{cases} r - x - (m(r) + y) + 1 = m(r) - x - y - 1 & \text{if } r \text{ is even} \\ r - x - (m(r) + y) + 1 = m(r) - x - y & \text{if } r \text{ is odd.} \end{cases} \quad (24)$$

Suppose $r \notin C$. Then we must have that $a' = r \in A$. Therefore,

$$a' - b \geq \begin{cases} r - (m(r) - x - y - 1) = m(r) + x + y - 1 & \text{if } r \text{ is even} \\ r - (m(r) - x - y) = m(r) + x + y - 1 & \text{if } r \text{ is odd.} \end{cases}$$

We now have $c \leq m(r) - 1 < m(r) + x + y - 1 \leq a' - b$, so (a', b, c) is a violating triple, and we are done.

Suppose now $c' = r \in C$. By (23), this implies that r is odd, $x = 1$ and $y = 0$. By (24), $b \leq m(r) - 1$. Therefore,

$$c' - b \geq r - (m(r) - 1) = m(r) > m(r) - x,$$

so by (22), (a, b, c') is a violating triple. This completes the induction. \square

Proof of Lemma 4.14. We proceed by induction on $r \geq 3$. The base case $r = 3$ can easily be verified. Suppose now the claim holds for all $3 \leq r' < r$. Set $A' = A \cap [r - 1]$, $B' = B \cap [r - 1]$, and $C' = C \cap [r - 1]$.

Suppose that r is odd. If $A, B, C \subset [r - 1]$, then because $|A| = |B| = |C| = m(r) = m(r - 1)$, the induction hypothesis implies that $A = B = C = [m(r - 1) - 1, r - 1] = [m(r) - 1, r - 1]$, i.e.

case (2)(a) holds. Suppose now at least one of A , B , or C contain r . By relabeling if necessary, we may assume $r \in A$. Let $a' = r \in A$, $b = \min B$ and $c = \min C$. Then $b \leq r - |B| + 1 = m(r)$. Therefore $c \geq a' - b \geq r - m(r) = m(r) - 1$, so $C \subset [m(r) - 1, r]$. Similarly, $c \leq r - |C| + 1 = m(r)$, so $b \geq a' - c \geq r - m(r) = m(r) - 1$ implies $B \subset [m(r) - 1, r]$. If $b = c = m(r) - 1$, then (a', b, c) is a violating triple, a contradiction. Thus at most one of b or c is $m(r) - 1$. Therefore, by relabeling if necessary, we may assume $B \subset [m(r), r]$ and $C \subset [m(r) - 1, r]$. Recall that $|[m(r) - 1, r]| = m(r) + 1 = |B| + 1$, so this implies that $B = [m(r), r]$. Let $a = \min A$. Then $r \in B$ and $c \leq m(r)$ implies $a \geq r - m(r) = m(r) - 1$, so $A \subset [m(r) - 1, r]$. If $C = [m(r), r]$, then we are done. If $C \neq [m(r), r]$, then $c < m(r)$ implies $(m(r) - 1, r, c)$ is a violating triple, so $m(r) - 1 \notin A$. Thus $A \subset [m(r), r]$ and $|A| = |[m(r), r]|$ implies $A = [m(r), r]$ and we are done.

Suppose now that r is even. Note that $\min\{|A'|, |B'|, |C'|\} \geq m(r) - 1$. If two elements of the set $\{|A'|, |B'|, |C'|\}$ are strictly greater than $m(r) - 1 = m(r - 1)$, then Lemma 4.9 implies there is a violating triple in $A' \times B' \times C'$, a contradiction. Therefore by relabeling if necessary, we may assume $|A'| = |B'| = m(r) - 1$, so $r \in A \cap B$. Let $a = \min A$, $b = \min B$, $c = \min C$ and note that $\max\{a, b, c\} \leq r - m(r) + 1 = m(r) - 1$. Now (a, r, c) and (r, b, c) cannot be violating triples, so

$$\begin{aligned} a &\geq r - c \geq r - (m(r) - 1) = m(r) - 1, \\ c &\geq r - b \geq r - (m(r) - 1) = m(r) - 1 \text{ and} \\ b &\geq r - c \geq r - (m(r) - 1) = m(r) - 1. \end{aligned}$$

Thus, $A, B, C \subset [m(r) - 1, r]$. Since $|A| = |B| = |C| = |[m(r) - 1, r]|$, this implies $A = B = C = [m(r) - 1, r]$. \square

References

- [1] Ove Ahlman and Vera Koponen, *Random l -colourable structures with a pregeometry*, arXiv:1207.4936 [math.LO], 2012.
- [2] Maria Axenovich and Ryan Martin, *A version of szemerédi's regularity lemma for multicolored graphs and directed graphs that is suitable for induced graphs*, arXiv: 1106.2871v1 [math.CO], 2011.
- [3] József Balogh, Béla Bollobás, and Miklós Simonovits, *The number of graphs without forbidden subgraphs*, J. Combin. Theory Ser. B **91** (2004), no. 1, 1–24.
- [4] ———, *The typical structure of graphs without given excluded subgraphs*, Random Structures Algorithms **34** (2009), no. 3, 305–318.
- [5] ———, *The fine structure of octahedron-free graphs*, J. Combin. Theory Ser. B **101** (2011), no. 2, 67–84.
- [6] József Balogh and Dhruv Mubayi, *Almost all triple systems with independent neighborhoods are semi-bipartite*, J. Combin. Theory Ser. A **118** (2011), no. 4, 1494–1518.
- [7] ———, *Almost all triangle-free triple systems are tripartite*, Combinatorica **32** (2012), no. 2, 143–169.
- [8] József Balogh and Wojciech Samotij, *The number of $K_{s,t}$ -free graphs*, J. Lond. Math. Soc. (2) **83** (2011), no. 2, 368–388.

- [9] Jason P. Bell, *Sufficient conditions for zero-one laws*, Trans. Amer. Math. Soc. **354** (2002), no. 2, 613–630 (electronic).
- [10] Béla Bollobás, *Random graphs*, second ed., Cambridge Studies in Advanced Mathematics, vol. 73, Cambridge University Press, Cambridge, 2001.
- [11] Graham Brightwell and Sarah Goodall, *The number of partial orders of fixed width*, Order **13** (1996), no. 4, 315–337.
- [12] Graham Brightwell, David A. Grable, and Hans Jürgen Prömel, *Forbidden induced partial orders*, Discrete Math. **201** (1999), no. 1-3, 53–80.
- [13] Stanley Burris and Karen Yeats, *Sufficient conditions for labelled 0-1 laws*, Discrete Math. Theor. Comput. Sci. **10** (2008), no. 1, 147–156.
- [14] Rohan Chabukswar and Aaron Tievsy, *Almost monochromatic triangles and other ramsey problem variants*, <http://users.ece.cmu.edu/~rchabuks/main.pdf>, 2004, preprint.
- [15] Kevin J. Compton, *A logical approach to asymptotic combinatorics. I. First order properties*, Adv. in Math. **65** (1987), no. 1, 65–96.
- [16] ———, *A logical approach to asymptotic combinatorics. II. Monadic second-order properties*, J. Combin. Theory Ser. A **50** (1989), no. 1, 110–131.
- [17] Gabriel Conant and Caroline Terry, *Model theoretic properties of the Urysohn sphere*, Annals of Pure and Applied Logic **167** (2016), no. 1, 49 – 72.
- [18] Christian Delhommé, Claude Laflamme, Maurice Pouzet, and Norbert Sauer, *Divisibility of countable metric spaces*, European J. Combin. **28** (2007), no. 6, 1746–1769.
- [19] Paul Erdős, Péter Frankl, and Vojtěch Rödl, *The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent*, Graphs Combin. **2** (1986), no. 2, 113–121.
- [20] Paul Erdős, Daniel J. Kleitman, and Bruce L. Rothschild, *Asymptotic enumeration of K_n -free graphs*, Colloquio Internazionale sulle Teorie Combinatorie (Rome, 1973), Tomo II, Accad. Naz. Lincei, Rome, 1976, pp. 19–27. Atti dei Convegni Lincei, No. 17.
- [21] Ronald Fagin, *Probabilities on finite models*, J. Symbolic Logic **41** (1976), no. 1, 50–58.
- [22] J. Flum and M. Grohe, *Parameterized complexity theory (texts in theoretical computer science. an eatcs series)*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.
- [23] Zoltán Füredi and André Kündgen, *Turán problems for integer-weighted graphs*, J. Graph Theory **40** (2002), no. 4, 195–225.
- [24] Ju. V. Glebskiĭ, D. I. Kogan, M. I. Liogon’kiĭ, and V. A. Talanov, *Volume and fraction of satisfiability of formulas of the lower predicate calculus*, Kibernetika (Kiev) (1969), no. 2, 17–27.
- [25] Simi Haber and Michael Krivelevich, *The logic of random regular graphs*, J. Comb. **1** (2010), no. 3-4, 389–440.

- [26] Peter Heinig, Tobias Muller, Marc Noy, and Anusch Taraz, *Logical limit laws for minor-closed classes of graphs*, arXiv:1401.7021 [math.CO], 2014.
- [27] Wilfrid Hodges, *Model theory*, Encyclopedia of Mathematics and its Applications, vol. 42, Cambridge University Press, Cambridge, 1993.
- [28] Christoph Hundack, Hans Jürgen Prömel, and Angelika Steger, *Extremal graph problems for graphs with a color-critical vertex*, *Combin. Probab. Comput.* **2** (1993), no. 4, 465–477.
- [29] Daniel J. Kleitman and Bruce L. Rothschild, *Asymptotic enumeration of partial orders on a finite set*, *Trans. Amer. Math. Soc.* **205** (1975), 205–220.
- [30] Phokion G. Kolaitis, Hans J. Prömel, and Bruce L. Rothschild, *K_{l+1} -free graphs: asymptotic structure and a 0-1 law*, *Trans. Amer. Math. Soc.* **303** (1987), no. 2, 637–671.
- [31] Vera Koponen, *Asymptotic probabilities of extension properties and random l -colourable structures*, *Ann. Pure Appl. Logic* **163** (2012), no. 4, 391–438.
- [32] ———, *Random graphs with bounded maximum degree: asymptotic structure and a logical limit law*, *Discrete Math. Theor. Comput. Sci.* **14** (2012), no. 2, 229–254.
- [33] ———, *A limit law of almost l -partite graphs*, *J. Symbolic Logic* **78** (2013), no. 3, 911–936.
- [34] Gady Kozma, Tom Meyerovitch, Ron Peled, and Wojciech Samotij, *Random points in the metric polytope*, 2013, Slides of a talk given at Random Combinatorial Structures and Statistical Mechanics, May 6–10, Venice, Italy.
- [35] Esther Lamken and Bruce L. Rothschild, *The numbers of odd-cycle-free graphs*, *Finite and infinite sets*, Vol. I, II (Eger, 1981), *Colloq. Math. Soc. János Bolyai*, vol. 37, North-Holland, Amsterdam, 1984, pp. 547–553.
- [36] James F. Lynch, *Convergence law for random graphs with specified degree sequence*, *ACM Trans. Comput. Log.* **6** (2005), no. 4, 727–748.
- [37] M. Malliaris and S. Shelah, *Regularity lemmas for stable graphs*, *Trans. Amer. Math. Soc.* **366** (2014), no. 3, 1551–1585.
- [38] David Marker, *Model theory*, Graduate Texts in Mathematics, vol. 217, Springer-Verlag, New York, 2002, An introduction.
- [39] Vania Mascioni, *Equilateral triangles in finite metric spaces*, *Electron. J. Combin.* **11** (2004), no. 1, Research Paper 18, 20 pp. (electronic).
- [40] ———, *On the probability that finite spaces with random distances are metric spaces*, *Discrete Math.* **300** (2005), no. 1-3, 129–138.
- [41] Dhruv Mubayi and Caroline Terry, *Discrete metric spaces: structure, enumeration, and 0-1 laws*, under review, *J. Symbolic Logic*, arXiv:1502.01212 [math.CO], 2015.
- [42] Deryk Osthus, Daniela Kühn, Timothy Townsend, and Yi Zhao, *On the structure of oriented graphs and digraphs with forbidden tournaments or cycles*, arxiv:1404.6178 [math.CO], 2014.
- [43] Yury Person and Mathias Schacht, *Almost all hypergraphs without Fano planes are bipartite*, *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, SIAM, Philadelphia, PA, 2009, pp. 217–226.

- [44] Hans J. Prömel and Angelika Steger, *The asymptotic number of graphs not containing a fixed color-critical subgraph*, *Combinatorica* **12** (1992), no. 4, 463–473.
- [45] Robert W. Robinson, *Counting labeled acyclic digraphs*, *New directions in the theory of graphs* (Proc. Third Ann Arbor Conf., Univ. Michigan, Ann Arbor, Mich., 1971), Academic Press, New York, 1973, pp. 239–273.
- [46] Saharon Shelah, *Toward classifying unstable theories*, *Ann. Pure Appl. Logic* **80** (1996), no. 3, 229–255.
- [47] Richard P. Stanley, *Acyclic orientations of graphs*, *Discrete Math.* **5** (1973), 171–178.
- [48] Katrin Tent and Martin Ziegler, *A course in model theory*, *Lecture Notes in Logic*, vol. 40, Association for Symbolic Logic, La Jolla, CA; Cambridge University Press, Cambridge, 2012.
- [49] Peter Winkler, *Random structures and zero-one laws*, *Finite and infinite combinatorics in sets and logic* (Banff, AB, 1991), *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, vol. 411, Kluwer Acad. Publ., Dordrecht, 1993, pp. 399–420.