

HIGH-GIRTH STEINER TRIPLE SYSTEMS

MATTHEW KWAN, ASHWIN SAH, MEHTAAB SAWHNEY, AND MICHAEL SIMKIN

ABSTRACT. We prove a 1973 conjecture due to Erdős on the existence of Steiner triple systems with arbitrarily high girth.

1. INTRODUCTION

Extremal combinatorics is largely concerned with the extent to which global statistics of an object (such as its *volume* or *density*) control the existence of local substructures. Classical examples include Szemerédi’s theorem on arithmetic progressions in dense subsets of the integers, and the Erdős–Stone–Simonovits theorem on the density threshold which guarantees the existence of a given subgraph.

In the extremal theory of *hypergraphs*, one of the most fundamental research directions, initiated by Brown, Erdős, and Sós [11], concerns the density conditions which guarantee the existence of local configurations with specific statistics. To be precise (and specializing to 3-uniform hypergraphs, also known as *triple systems*), a (j, ℓ) -*configuration* is a set of ℓ triples which span at most j vertices. Perhaps the most well-known theorem in this area concerns the so-called $(6, 3)$ -*problem*: it was proved by Rusza and Szemerédi [55] that every N -vertex triple system with no $(6, 3)$ -configuration has only $o(N^2)$ triples. This innocent-sounding fact has a number of surprising consequences, including Roth’s theorem on 3-term arithmetic progressions in dense sets of integers.

In general, problems of this type are formidably difficult (for example, it is a famous open problem to prove more generally that for any fixed $j \in \mathbb{N}$, N -vertex triple systems with no $(j, j - 3)$ -configuration have $o(N^2)$ edges), and it seems far out of reach to precisely characterize the number of triples which forces the existence of a given type of configuration (or even the order of magnitude of this number), in general. However, certain special cases are tractable. In particular, Brown, Erdős, and Sós found the order of magnitude of the solution to the $(j, j - 2)$ -problem, for any $j \geq 4$: there are constants $C_j > c_j > 0$ such that any N -vertex triple system with at least $C_j N^2$ triples contains a $(j, j - 2)$ -configuration, while on the other hand, for any N there is an N -vertex triple system with $c_j N^2$ triples and no $(j, j - 2)$ -configuration¹. They also observed that it is possible to give a more-or-less exact solution to the $(4, 2)$ -problem. Indeed, the property of having no $(4, 2)$ -configuration is equivalent to the property that every pair of vertices is included in at most one triple (which implies that every vertex is included in at most $(N - 1)/2$ triples, and the total number of triples is at most $\binom{N}{2}/3$). As long as $(N - 1)/2$ and $\binom{N}{2}/3$ are integers², it is a classical fact³ that there exists an N -vertex *Steiner triple system*, in which every pair of vertices is included in *exactly* one triple (such a triple system has exactly $\binom{N}{2}/3$ triples).

As a far-reaching generalization of the above two facts, Erdős conjectured [19] that for any $g \in \mathbb{N}$, if N is admissible and sufficiently large in terms of g , then there is a Steiner triple system which has no $(j, j - 2)$ -configuration for any $4 \leq j \leq g$. That is to say, the $(4, 2)$ -problem is the most restrictive of the nontrivial $(j, j - 2)$ -problems, in the very strong sense that given the constraint of having no $(4, 2)$ -configurations, it makes basically no difference to further impose that there is no $(j, j - 2)$ -configuration for any other j . Erdős’ conjecture has been reiterated by many authors in different contexts; see for example [4–6, 13, 14, 17, 24, 25, 29–31, 38, 42, 45, 48, 57, 58]. In this paper we prove Erdős’ conjecture.

Theorem 1.1. *Given $g \in \mathbb{N}$, there is $N_{1.1}(g) \in \mathbb{N}$ such that if $N \geq N_{1.1}(g)$ and N is congruent to 1 or 3 (mod 6), then there exists a Steiner triple system of order N which contains no $(j, j - 2)$ -configuration for any $4 \leq j \leq g$.*

It is natural to view [Theorem 1.1](#) as being about the existence of Steiner triple systems which are “locally sparse” or have “high girth” (recall that the girth of a graph is the length of its shortest cycle, or equivalently the smallest integer g for which there is a set of g vertices containing at least g edges).

Sah and Sawhney were supported by NSF Graduate Research Fellowship Program DGE-1745302. Sah was supported by the PD Soros Fellowship. Simkin was supported by the Center of Mathematical Sciences and Applications at Harvard University.

¹The only other nontrivial class of (j, ℓ) -problems for which the order of magnitude is known are $(2\ell + 1, \ell)$ -problems, which are somewhat degenerate (in a triple system with no $(2\ell + 1, \ell)$ -configuration, every connected component has at most ℓ edges).

²This happens if and only if N is congruent to 1 or 3 (mod 6); such N are said to be *admissible*.

³This is actually one of the oldest theorems in combinatorics, having been proved by Kirkman [40] in 1847.

Definition 1.2. The *girth* of a triple system is the smallest integer $g \geq 4$ for which it has a $(g, g - 2)$ -configuration. If no such configuration exists, the girth is defined to be infinite. If a triple system has girth strictly greater than $r + 2$, it is said to be *r-sparse*.

In the language of Definition 1.2, Theorem 1.1 says that for any r and sufficiently large admissible N , there exists an r -sparse Steiner triple system of order N . Interpreted in this way, Theorem 1.1 is of interest beyond its relevance to extremal hypergraph theory. Indeed, in the intervening years since Erdős made his conjecture, it has become notorious in design theory (roughly speaking, combinatorial designs are set systems with very strong regularity properties; Steiner triple systems are fundamental examples). Also, Erdős’ conjecture has recently been reiterated in the context of the *high-dimensional combinatorics* program spearheaded by Linial and collaborators (see for example [48]), in which a Steiner triple system can be viewed as a 2-dimensional generalization of a graph perfect matching.

1.1. Previous work. In design theory, there has been quite intensive effort towards studying the existence of r -sparse Steiner triple systems for small r . Every Steiner triple system is 2-sparse, and in fact also 3-sparse (one can show that every $(5, 3)$ -configuration contains a $(4, 2)$ -configuration). The 4-sparse case of Erdős’ conjecture was resolved by Grannell, Griggs, and Whitehead [32] following previous partial results by various authors [10, 16, 46] (4-sparse Steiner triple systems are also said to be *anti-Pasch* or *quadrilateral-free*, and exist for all admissible orders except 7 and 13). No further cases of Erdős’ conjecture were known until now, but there had been significant progress in the cases $r = 5$ and $r = 6$: Wolfe [60] proved that 5-sparse Steiner triple systems exist for “almost all” admissible orders, in a certain asymptotic sense (see [15, 22, 47] for previous results), and Forbes, Grannell and Griggs [20] found a recursive construction that produces infinitely many 6-sparse Steiner triple systems. Until now no 7-sparse Steiner triple systems were known to exist.

In a different direction, Glock, Kühn, Lo, and Osthus [27] and Bohman and Warnke [8] recently proved an *approximate* version of Theorem 1.1: namely, for any $r \in \mathbb{N}$ there is an r -sparse triple system with $(1 - o(1))N^2/6$ triples (recall that a Steiner triple system has exactly $N(N - 1)/6$ triples). The ideas in these proofs play a crucial role in our proof of Theorem 1.1, as we will next discuss.

1.2. Overview of key ideas. In this subsection we discuss some of the key ideas in the proof of Theorem 1.1, at a very high level. We provide a more detailed proof outline in Section 2. For the benefit of non-expert readers, we start by briefly examining some of the beautiful and very powerful ideas underlying the two independent proofs of the so-called *existence of designs* conjecture, due to Keevash [37] and Glock, Kühn, Lo, and Osthus [28], which revolutionized design theory and will play an essential role in our proof of Theorem 1.1.

1.2.1. Approximate decompositions and random greedy algorithms. An order- N Steiner triple system is equivalent to a decomposition of the complete graph K_N into edge-disjoint triangles. The starting point for both the existence of designs proofs in [28, 37] (specialized to Steiner triple systems) is the fact that it is possible to find an *approximate* triangle-decomposition of K_N : a collection of edge-disjoint triangles that cover a $(1 - o(1))$ -fraction of the edges of K_N .

Famously, this fact may be proved via the *probabilistic method*: for example, we can consider the *random greedy* algorithm which builds a collection of triangles one-by-one, at each step choosing a uniformly random triangle among those which are edge-disjoint from previously chosen triangles. Spencer [59] and Rödl and Thoma [54] proved that this process is likely to produce an approximate triangle-decomposition, following a foundational paper of Rödl [51] that considered a slightly different random process (the so-called *Rödl nibble*). These ideas have had an enormous impact throughout combinatorics in the last 30 years.

The aforementioned works of Glock, Kühn, Lo, and Osthus [27] and Bohman and Warnke [8], proving an approximate version of Theorem 1.1, proceeds along similar lines, studying a “high-girth triple process” that iteratively builds an r -sparse Steiner triple system (first suggested by Krivelevich, Kwan, Loh, and Sudakov [42]). It is quite surprising that this is possible, as the set of $(j, j - 2)$ -configurations is diverse and difficult to characterize.

1.2.2. The absorbing method. If one’s goal is to find a Steiner triple system, it is not obvious that an approximate decomposition is actually helpful. Indeed, if we consider the random greedy algorithms described above, the number of choices for the next triangle decreases at each step, and as we are nearing completion the situation becomes very restricted. One can show that these processes are very unlikely to produce a complete Steiner triple system [7].

Surprisingly, it turns out that there is a general scheme that allows one to convert approximate packing/matching/decomposition results into exact ones, called the *absorbing method* (pioneered by Erdős, Gyárfás, and Pyber [18], and later systematized by Rödl, Ruciński, and Szemerédi [53]). Roughly speaking, the idea is to prepare an “absorbing structure” at the start of the proof which is very “flexible” and

can contribute to our desired packing in many different ways. Then, one finds an approximate packing of the complement of this absorbing structure, and uses the special properties of the absorbing structure to transform this into an exact packing.

Although the absorbing method has found success in a range of problems over the years, it is difficult to use it in the setting of designs. A key obstacle is that one must create an absorbing structure which itself can be embedded in a design. Keevash’s breakthrough (which he calls the method of *randomized algebraic construction*) is that it is possible to construct a suitable absorbing structure by randomly sampling from algebraic objects (for Steiner triple systems, the crucial object is the triple system \mathcal{T} consisting of solutions to $x + y + z = 0$ among nonzero elements of an abelian 2-group \mathbb{Z}_2^k ; see the exposition in [35] for more details).

Keevash’s method of randomized algebraic construction is enormously powerful and has since been used to resolve many other problems in design theory [9, 35, 36]. However, it seems to be uniquely unsuitable for Erdős’ conjecture on high-girth Steiner triple systems. Actually, the triple system \mathcal{T} defined above is a Steiner triple system of order $2^k - 1$, and it has the maximum possible number of $(6, 4)$ -configurations among any Steiner triple systems of this order! These $(6, 4)$ -configurations are in fact crucial to the proof; they allow one to make local changes to \mathcal{T} , and this flexibility is essential to the utility of the absorbing structure⁴.

1.2.3. Iterative absorption. A few years after Keevash proved the existence of designs conjecture, Glock, Kühn, Lo, and Osthus [28] discovered a new and completely non-algebraic proof using the method of *iterative absorption* (first introduced by Kühn and Osthus [43] and Knox, Kühn, and Osthus [41], and since refined in various applications to a wide range of different problems; see for example [1–3, 9, 26, 34, 50]). Roughly speaking, the idea is to use bare-hands combinatorics to construct a very limited absorbing structure that can only “fix” very restricted types of defects, and then to iteratively transform our decomposition problem into smaller and smaller decomposition problems until it is small enough to be solved using the absorber.

Specifically, in the setting of Steiner triple systems, it is possible to use a combination of random greedy processes (extending the ideas in Section 1.2.1) to construct an approximate triangle-decomposition in K_N whose leftover edges are “localized”. Namely, all edges uncovered by triangles are contained inside a particular small set of vertices (say, 1% of the vertices of K_N), and moreover almost none of the edges inside this small set are covered by triangles. This works as long as N is sufficiently large (say $N \geq N_0$). Then, one can “zoom in” to our small set of vertices, reducing the situation to a smaller decomposition problem. This can be iterated (about $\log_{0.99}(N_0/N)$ times) until the number of vertices remaining is less than N_0 . That is to say, in a large complete graph K_N we can find a set of edge-disjoint triangles which covers all edges outside of a specific constant-size set of vertices. Since there are only a constant number of possibilities for the set of leftover edges, it suffices to construct “exclusive” absorbers for each: namely, for each possible graph G of uncovered edges, one designs an “absorbing” graph H with the property that both H and $H \cup G$ have a triangle-decomposition.

The method of iterative absorption has certain disadvantages compared to Keevash’s method of randomized algebraic construction (for example, it seems to have much weaker quantitative aspects), but it has the distinct advantage that the requirements for the absorbing structure are much weaker (we only need to be able to handle a specific, tiny, set of vertices). So, it seems plausible that one might be able to design an absorbing structure which is itself high-girth, suitable for proving Erdős’ conjecture on high-girth Steiner triple systems.

1.2.4. Constraint focusing. Iterative absorption is indeed the approach we take in our proof of Theorem 1.1, but there are serious issues that need to be overcome. Most obviously, the property of being r -sparse is not closed under taking unions: if we consider two edge-disjoint sets of triangles, which each correspond to an r -sparse triple system, we cannot guarantee any local sparsity properties of their union. This was highlighted as a key obstacle towards proving Erdős’ conjecture by Glock, Kühn, Lo, and Osthus [27]: indeed, iterative absorption fundamentally depends on being able to augment a partial decomposition with another one on a smaller number of vertices.

Superficially, it may seem like this problem can be overcome with “mere bookkeeping”. Indeed, after the first step of the iteration, there are specific sets of triangles that cannot be used together (as they would create a $(j, j - 2)$ -configuration for some $j \leq r + 2$, together with previously chosen triangles). Such sets can now be viewed as additional “forbidden configurations”, which must be avoided in future steps of the iteration. So, at each step, instead of being presented with the task of finding an r -sparse approximate

⁴It is possible to consider different algebraic structures to avoid $(6, 4)$ -configurations in particular, but the linear-algebraic nature of Keevash’s approach seems to make it fundamentally unsuitable for constructing r -sparse Steiner triple systems for $r \geq 7$. Relatedly, we remark that Fujiwara [23] proved some nonexistence theorems for high-girth Steiner triple systems with certain algebraic properties.

Steiner triple system, we now have the more general task of finding an approximate Steiner triple system avoiding a given collection of forbidden configurations (which one would need to track throughout the process of iterative absorption).

However, there is a fundamental issue with this plan. As we have just explained, at each step of the iteration, we do not only face an r -sparseness constraint, but we are also faced with other constraints inherited from previous steps of the iteration. These constraints build up over the iterations: the basic strategy of iterative absorption has the unfortunate side effect of “focusing” constraints on a particular small set of vertices. Recall that we need about $\log N$ steps of the iteration, but well before this point we will simply have run out of available triangles which we can use without violating the r -sparseness constraint.

1.2.5. Sparsification and efficient absorption. We overcome the constraint focusing issue with a combination of two ideas. First, we observe that constraint focusing is at its most severe when performing iterative absorption on the complete graph K_N , simply because in every step we must take a very large number of triangles (each of which contributes constraints for future steps). If we were instead interested in finding a triangle-decomposition of (say) a sparse random graph, then the constraints arising from each step of the iteration would be negligible⁵.

So, before we begin our iterative absorption procedure, we run the high-girth triple process to find an approximate triangle-decomposition of K_N , in such a way that the uncovered edges resemble a sparse random graph. We then use iterative absorption to find a triangle-decomposition of this graph.

Of course, sparsification comes at a cost: namely, there are fewer edges and triangles to work with (and therefore less freedom to make choices) throughout the iterative absorption process. This issue is amplified by the nature of iterative absorption: in a sparse random graph with density $p = o(1)$, a typical set of $\sqrt{1/p}$ vertices is likely to have almost no edges at all, to say nothing about sets of constant size at the end of the iterative absorption process. The methods in [28] can be quantified, and it is not strictly necessary for the final absorption step to take place on a constant-size set of vertices, but the quantitative dependence is quite poor: the final set is forced to have size at most about $\sqrt{\log N}$ (and therefore the density p is forced to be at least about $1/\sqrt{\log N}$). Actually, this quantitative limitation is one of the key disadvantages of iterative absorption, compared to Keevash’s randomized algebraic construction (which is effective even in quite sparse settings).

In order to overcome the constraint focusing issue, we wish to sparsify quite harshly, and thus we need a more efficient absorbing structure. We accomplish this by taking advantage of short-cycle decompositions obtainable via Euler tours; we believe this type of construction is of independent interest, and we foresee applications to other problems⁶.

1.2.6. Retrospective analysis. There are a significant number of technical challenges one must overcome to implement the strategy described above. For example, a glaring issue is the difficulty of bookkeeping. Generally, in packing arguments involving multiple stages of random choices (including, as far as we know, all applications of iterative absorption), one records “pseudorandomness” or “typicality” properties at each stage, which may then be used in future stages. In our setting, where we need to record the forbidden configurations inherited from previous steps of the iteration, there is a staggering amount of bookkeeping that would be necessary (see Section 2.2.6 for a brief discussion of the relevant subtleties). While we do use notions of typicality in various parts of our proof, the forbidden configurations are largely studied via “retrospective analysis”: instead of attempting to track the complex combinatorics of evolving forbidden configurations, we “remember the randomness” that was used in previous stages of the iterative absorption process, and use this to deduce information about forbidden configurations “on demand”. This analysis is facilitated by our new formalism of *weight systems*, and some general moment estimation lemmas (see Section 3.1) related to *Kim–Vu polynomial concentration* [39] and Rödl and Ruciński’s *deletion method* [52]. We believe these ideas have great potential for further applications.

1.3. Further directions. Strengthening Theorem 1.1, we can also prove a lower bound on the number of r -sparse Steiner triple systems on a given set of vertices. Let erd_j be the number of labeled $(j, j - 2)$ -configurations with girth j on the vertex set $\{1, \dots, j\}$, which contain $\{1, 2, 3\}$ as a triple.

Theorem 1.3. *Given $g \in \mathbb{N}$, there is $c_{1.3}(g) > 0$ such that the following holds. If N is congruent to 1 or 3 (mod 6), then the number of (labeled) Steiner triple systems with girth greater than g , on a specific set of*

⁵This is analogous to the fact that for fixed $j \in \mathbb{N}$, in a random triple system with $m = N^2$ triples, the number of $(j, j - 2)$ -configurations is of the same order m , but in a random triple system with $m = o(N^2)$ triples, the number of $(j, j - 2)$ -configurations is negligible compared to m . Actually, this is the key fact underlying Brown, Erdős, and Sós’ lower bound for the $(j, j - 2)$ -problem.

⁶Related ideas using Euler tours also appear in papers of Keevash [35] and Piga and Sanhueza-Matamala [50].

N vertices, is at least

$$\left(\left(1 - N^{-c_{1.3}(g)} \right) N \exp \left(-2 - \sum_{j=6}^g \frac{\text{erd}_j}{(j-2)!} \right) \right)^{\frac{N^2}{6}}.$$

We remark that it makes no difference whether we consider labeled or unlabeled Steiner triple systems in [Theorem 1.3](#): a factor of $N!$ would be absorbed into the error term. We also remark that the case $g < 6$ provides an estimate for the total number of order- n Steiner triple systems, giving an independent proof of a theorem of Keevash [\[35\]](#) (this was previously not possible with iterative absorption, due to its quantitative limitations).

We believe the lower bound in [Theorem 1.3](#) is best-possible, as conjectured by Glock, Kühn, Lo, and Osthus [\[27\]](#) (see also [\[8\]](#) for the same conjecture in the case $g = 6$). It is not obvious how to prove a matching upper bound, though we remark that the ideas of [\[44\]](#) may be used to prove that for $r \geq 4$, the number of r -sparse Steiner triple systems is at most $(cN)^{N^2/6}$ for some constant $c < e^{-2}$ (that is to say, the probability that a random Steiner triple system is r -sparse is of the form $e^{-\Omega(N^2)}$).

As a different strengthening of [Theorem 1.1](#), we believe our methods are also suitable for showing that a sufficiently dense graph G satisfying certain pseudorandomness and divisibility properties admits an r -sparse triangle-decomposition (for example, in terms of pseudorandomness it should suffice for $G \subseteq K_N$ to have density $p \geq N^{-c_r}$, and for every set A of at most 4 vertices to have $(1 \pm \xi_r)p^{|A|}N$ neighbors, for appropriate constants $\xi_r, c_r > 0$. In terms of divisibility, we need the number of edges in G to be divisible by 3, and for each degree to be even). This would imply a weak version of a conjecture of Glock, Kühn, and Osthus [\[29, Conjecture 7.7\]](#).

Also, it would be of interest to study the optimal dependence between g and $N_{1.1}(g)$ in [Theorem 1.1](#): what is the maximum possible girth of an order- N Steiner triple system? Inspecting our proof of [Theorem 1.1](#), it seems that one obtains a lower bound of $(\log \log N)^c$ for some constant $c > 0$; in any case, $\log \log N$ seems to be a hard barrier for our approach. On the other hand, Lefmann, Phelps, and Rödl [\[45\]](#) proved an upper bound of order $\log N / \log \log N$.

Finally, we remark that it may be of interest to generalize [Theorem 1.1](#) to higher uniformities (e.g. to study Steiner quadruple systems instead of Steiner triple systems), or to consider triple systems other than Steiner triple systems (for example, *Latin squares* can be interpreted as triple systems). Erdős made his conjecture only for Steiner triple systems, but generalizations to higher uniformities have been proposed by Füredi and Ruszinkó [\[24, Conjecture 1.4\]](#), Glock, Kühn, Lo, and Osthus [\[27, Conjecture 7.2\]](#), and Keevash and Long [\[38, Section 3\]](#). Generalizations to Latin squares and 1-factorizations have been proposed by Linial [\[48\]](#). We are not aware of any fundamental obstructions that would prevent one from generalizing the methods in this paper to any of these different settings (in particular, iterative absorption can be applied in the “multipartite” settings of Latin squares and 1-factorizations, see [\[3, 49\]](#)). However, given the level of technicality of this paper, and especially given the considerable additional technicalities that higher-uniformity generalizations would introduce, we hesitate to make any concrete claims.

1.4. Notation. We use standard asymptotic notation throughout, as follows. For functions $f = f(n)$ and $g = g(n)$, we write $f = O(g)$ to mean that there is a constant C such that $|f| \leq C|g|$, $f = \Omega(g)$ to mean that there is a constant $c > 0$ such that $f(n) \geq c|g(n)|$ for sufficiently large n , and $f = o(g)$ to mean that $f/g \rightarrow 0$ as $n \rightarrow \infty$. Subscripts on asymptotic notation indicate quantities that should be treated as constants. Also, following [\[37\]](#), the notation $f = 1 \pm \varepsilon$ means $1 - \varepsilon \leq f \leq 1 + \varepsilon$.

For a real number x , the floor and ceiling functions are denoted $\lfloor x \rfloor = \max\{i \in \mathbb{Z} : i \leq x\}$ and $\lceil x \rceil = \min\{i \in \mathbb{Z} : i \geq x\}$. We will however mostly omit floor and ceiling symbols and assume large numbers are integers, wherever divisibility considerations are not important. All logarithms in this paper are in base e , unless specified otherwise.

Also, where it is not too inconvenient we make an effort to be consistent with fonts: as in [\[27\]](#), calligraphic letters $(\mathcal{A}, \mathcal{B}, \mathcal{C})$ will generally be used to denote sets of triangles (or sets of triples), and Fraktur letters $(\mathfrak{A}, \mathfrak{B}, \mathfrak{C})$ will be used to denote collections of sets of triangles.

1.5. Acknowledgements. The authors thank Yufei Zhao for suggesting a simplification of the proof of [Lemma 6.2](#), Stefan Glock for some helpful discussions and clarifications on the method of iterative absorption, and Zach Hunter for pointing out a number of typographical errors. Michael Simkin thanks Nati Linial for introducing him to the problem and for many discussions on the topic of girth. Matthew Kwan thanks Benny Sudakov for introducing him to the problem and for many discussions on random processes.

2. PROOF OVERVIEW

We prove [Theorem 1.1](#) by describing a probabilistic algorithm that (for fixed $g \in \mathbb{N}$) w.h.p.⁷ produces a Steiner triple system of order N with girth greater than g . Following [\[35\]](#), [Theorem 1.3](#) is proved by counting the number of possible paths the algorithm may take.

2.1. A bird’s-eye view of iterative absorption for Steiner triple systems. The high-level structure of our proof (and our choice of notation, where possible) is modeled after a paper of Barber, Glock, Kühn, Lo, Montgomery, and Osthus [\[1\]](#), in which the machinery of iterative absorption is applied in the special case of Steiner triple systems. For the convenience of the reader, we start by outlining the strategy used to prove [\[1, Theorem 1.3\]](#).

Let $G \subseteq K_N$ be a graph that we wish to decompose into triangles. We assume that G is *triangle-divisible* (the number of edges is divisible by 3, and every degree is even) and *pseudorandom* (say, every set A of at most 100 vertices has about $p^{|A|}N$ common neighbors, for some “not-too-small” p).

We begin by fixing a *vortex* $V(K_N) = U_0 \supseteq U_1 \supseteq \dots \supseteq U_\ell =: X$, in such a way that each $G[U_k]$ is itself pseudorandom (for example, the vortex can be chosen randomly). Then, an absorber graph $H \subseteq G$ is set aside, leaving the graph $G_0 = G \setminus H$. The crucial property of H is that for every triangle-divisible graph L on the vertex set X , the union $L \cup H$ admits a triangle-decomposition. In [\[1\]](#), this graph H is obtained by taking a disjoint union of about $2^{|X|^2}$ “exclusive” absorbers, one for every possible triangle-divisible graph on the vertex set X . (In our proof of [Theorem 1.1](#), we will need a much more efficient construction).

Next is the iterative part: for $0 \leq k < \ell$, we find a set of edge-disjoint triangles \mathcal{M}_k in G_k , covering all edges of $G_k \setminus G_k[U_{k+1}]$. Then, we let G_{k+1} be the subgraph of $G_k[U_{k+1}]$ obtained by removing all edges covered by \mathcal{M}_k . After ℓ iterations, the graph of uncovered edges $G_\ell \subseteq G_0$ lies fully within the smallest vortex set U_ℓ . By the defining property of the absorber there exists a triangle-decomposition \mathcal{T} of $G_\ell \cup H$, so $\mathcal{T} \cup \bigcup_{0 \leq k < \ell} \mathcal{M}_k$ is a triangle-decomposition of G .

The construction of the “exclusive absorbers” comprising H is quite ingenious, but the details are not too relevant to this outline of our proof of [Theorem 1.1](#). The iteration step is really the heart of the proof and consists of two components: “cover down” and “regularity boosting”.

2.1.1. Cover down. Roughly speaking, the “cover down” lemma (cf. [\[1, Lemma 3.8\]](#)) states that if G is a reasonably dense pseudorandom graph on n vertices and $U \subseteq V(G)$ is a not-too-small subset of vertices, then there exists a set of edge-disjoint triangles \mathcal{M} in G , such that all edges not covered by \mathcal{M} lie inside U . Furthermore, \mathcal{M} covers very few of the edges inside U , meaning that the graph of uncovered edges $L := G \setminus E(\mathcal{M})$ (on the vertex set U) remains pseudorandom. This lemma is the key ingredient used to find the sets of triangles \mathcal{M}_k described above.

The cover down lemma is itself proved using a three-stage randomized algorithm. First, we use a random greedy algorithm to find an approximate triangle-decomposition of $G \setminus G[U]$. At this point, apart from the edges in $G[U]$ (which we do not need to cover) there are two types of uncovered edges: those between $V(G) \setminus U$ and U (“crossing edges”) and those inside $V(G) \setminus U$ (“internal edges”).

Second, we consider another random greedy algorithm, whose purpose is to select triangles to cover the leftover internal edges. Indeed, for each uncovered internal edge e (in some arbitrary order), we choose a random triangle consisting of e and two uncovered crossing edges, uniformly at random.

Third, we need to cover the remaining crossing edges. Crucially, this can be viewed as a sequence of *perfect matching* problems: for each vertex $v \in V(G) \setminus U$, we consider the set of vertices $w \in U$ for which vw has not yet been covered; call this set W_v and consider the induced graph $G[W_v]$. An edge in $G[W_v]$ corresponds to a triangle containing v and two uncovered crossing edges, so we can find edge-disjoint triangles to cover the remaining crossing edges by running through each $v \in V(G) \setminus U$ (in some arbitrary order), and finding a perfect matching in $G[W_v]$ (avoiding edges covered in previous stages of this process). For this third step we can take advantage of the well-known fact that pseudorandom graphs contain perfect matchings (as may be proved using Hall’s theorem).

As a technical note, we remark that it is necessary to set aside a small random subset R of the edges in G between $V(G) \setminus U$ and U (this graph R is called the *reserve graph*), before the above three-stage process. If we choose the sampling probability for R appropriately, this ensures that the uncovered edges are (w.h.p.) predominantly comprised of those in R . This means that we will have “plenty of room” for the second stage, and it means that in the third stage it will be very easy to show the pseudorandomness properties that imply existence of perfect matchings.

⁷We say an event holds *with high probability*, or *w.h.p.* for short, if it holds with probability $1 - o(1)$ (asymptotics will be as $N \rightarrow \infty$ or as $n \rightarrow \infty$; this should be clear from context).

2.1.2. *Regularity boosting.* There is a subtle issue with the above “cover down” procedure. When using a random greedy algorithm to find an approximate triangle-decomposition of a graph G , we need to assume that G is pseudorandom (actually, “triangle-regularity” suffices: we need every edge to be contained in roughly the same number of triangles). How close we are able to get to a genuine triangle-decomposition depends on how triangle-regular G is. A simple calculation shows that the dependence is not in our favor: if we were to iterate the above argument naïvely, our regularity would deteriorate very rapidly. In hindsight, this is perhaps unsurprising: it would be quite remarkable if one could do better than a single random greedy algorithm just by chaining together multiple different random greedy algorithms.

In order for the iterative absorption strategy to provide any meaningful gain, it is therefore necessary to “boost” the regularity before each “cover down” step. Specifically, we apply our random greedy algorithm restricted to a certain highly regular subset of triangles. This subset is obtained by randomly sampling from a *fractional* triangle-decomposition. Such a fractional decomposition is obtained in [1] by an ingenious “weight-shifting” argument.

2.2. High-girth iterative absorption. We define an *Erdős j -configuration* in a triple system (or, equivalently, in a set of triangles) to be a subset of triples which span exactly j vertices and have girth exactly j (that is to say, an Erdős configuration is a $(j, j - 2)$ -configuration which contains no proper subconfiguration that is itself an Erdős configuration). Informally, we just use the term “Erdős configuration” to describe Erdős j -configurations for $5 \leq j \leq g$, so our task is to find a triangle-decomposition of K_N containing no Erdős configurations.

In order to prove [Theorem 1.1](#) using the above scheme, we need each \mathcal{M}_k to contain no Erdős configuration, and moreover we need to choose each \mathcal{M}_k in such a way that its union with previous \mathcal{M}_i also has no Erdős configuration. Beyond that, we need to choose all the \mathcal{M}_k in such a way that the final “absorber” triangles in \mathcal{T} (which themselves must contain no Erdős configuration) do not introduce an Erdős configuration at the end.

2.2.1. *Absorbers.* First, we need to choose absorbers in such a way that their corresponding triangle-decompositions never include Erdős configurations. This can actually be done by taking the absorbers in [28] and modifying them in a black-box manner (extending the construction by copies of a gadget we call a “ g -sphere”; see [Definition 4.6](#)).

Recalling [Section 1.2.2](#), the main innovation of our absorber construction is that we are able to choose H in such a way that its number of vertices is only a polynomial function of the size of X (unlike the construction in [1], which has an exponential dependence). Our starting point for this is that in any triangle-divisible graph L (more generally, in any graph with even degrees) we can find an *Euler tour* passing through all the edges, which gives us a decomposition of G into cycles. By appending short paths between every pair of vertices, we can convert this initial cycle-decomposition into a cycle decomposition in which each cycle has bounded size, and if we strategically arrange “exclusive absorbers” of bounded size (as constructed in [1]) on top of this construction, we can convert this cycle-decomposition into a triangle-decomposition.

That is to say, our final absorber graph H is obtained by first carefully arranging a large number of short paths and small exclusive absorbers, and then extending the construction with g -spheres.

2.2.2. *A generalized high-girth process.* The majority of the triangles chosen in iterative absorption arise from instances of the first stage of “cover down”, where we find an approximate triangle-decomposition of an n -vertex graph $G \setminus G[U]$. As discussed in the introduction, Glock, Kühn, Lo, and Osthus [27] and Bohman and Warnke [8] studied⁸ a high-girth (i.e., Erdős configuration-free) random greedy process in the complete graph K_n ; we need to adapt their analysis to our setting, in which we are given a general graph to decompose and a general set of forbidden configurations to avoid. Namely, our forbidden configurations include Erdős configurations, any sets of triangles that would create an Erdős configuration together with previous stages of the iteration, any sets of triangles that would create an Erdős configuration together with a potential outcome of the “absorbing” triangles \mathcal{T} , and also some additional “regularizing” sets of triangles (which we will discuss later).

The definition of our random greedy process is straightforward, and the same as in [8, 27]: we iteratively build a set of edge-disjoint triangles by considering at each step all those triangles consisting of yet-uncovered edges whose addition would not complete a forbidden configuration, and choosing one of these triangles uniformly at random. Though it requires some quite violent changes to the proofs in [8, 27]⁹, it is possible to permit general collections of forbidden configurations, given some statistical assumptions (which we call

⁸Despite the fact that [8, 27] both study the same process with roughly the same methods, ideas exclusive to both of these papers will be useful to us, for different reasons.

⁹For the reader familiar with the proofs in these papers, both papers depend on “extension estimates”, taking advantage of the fact that the forbidden configurations can be interpreted as all the isomorphic copies of a particular family of triple systems.

“well-spreadness”). For example, just as there are at most n^{j-3} Erdős j -configurations (each of which have $j-2$ triples) containing a given triangle T , we need to assume that there are at most $O(n^{j-3})$ forbidden configurations which have $j-2$ triples and contain T . More generally, we need to assume an upper bound on the number of forbidden configurations with a given number of triples which include a given set of triangles in our graph, and we also need some information about *pairs* of forbidden configurations.

Of course, we also need to make some regularity assumptions: the number of forbidden configurations with a particular number of triples which include a triangle T must be about the same for all T . Perhaps surprisingly, it is *not* necessary to assume that this is true for forbidden configurations of a particular isomorphism class; we can treat all forbidden configurations with a given number of triples together, even if they have quite different structure. This is implicit in the analysis by Glock, Kühn, Lo, and Osthus [27] (but not the analysis by Bohman and Warnke [8]), and is crucial for our regularization argument, which we will discuss shortly. Under appropriate assumptions, we can find an approximate triangle-decomposition containing no forbidden configuration in which only a $n^{-\beta_g}$ -fraction of edges remain uncovered, for some constant $\beta_g > 0$ (this polynomial dependence on n is actually rather important, and uses ideas exclusive to the analysis in [8]).

2.2.3. Regularizing the forbidden configurations. As discussed in Section 2.1.2, it is crucial to include a regularity-boosting step between each cover down step. The regularity-boosting argument described in Section 2.1.2 is still relevant, but it is only suitable for triangle-regularity (ensuring that each edge is contained in the same number of triangles). As described above, we also need our forbidden configuration data to be regular. This is basically unavoidable, because the set of “available” triangles that may be selected at each step evolves according to the distribution of forbidden configurations (if a triangle appears in many forbidden configurations, it is likely to quickly become unavailable).

Our regularization task can be viewed as a hypergraph regularization problem: to regularize the collection of dangerous configurations with $j-2$ triples, we consider a certain $(j-2)$ -uniform hypergraph and add hyperedges to it to regularize its degrees. This can be accomplished by an appropriate random hypergraph, where the probability of selecting a hyperedge depends on the degrees of its vertices. Actually, by iterating this idea we are able to obtain a sharp general hypergraph regularization lemma that may be of independent interest.

2.2.4. The remainder of “cover down”. Apart from the approximate packing stage, there are two other stages in the multi-stage cover down process described in Section 2.1.1: we need to cover leftover internal edges with a simple random greedy process, and we need to cover leftover crossing edges with a perfect matching argument. The internal edges can be handled in basically the same way as in [1] (being careful to make choices that avoid forbidden configurations), but it is not as simple to handle the crossing edges.

Indeed, it no longer suffices to find arbitrary perfect matchings in our quasirandom graphs $G[W_v]$; we need to ensure that the union of our perfect matchings is “well-distributed” in that it does not interfere with the well-spreadness properties of the forbidden configurations for the next step of the iteration, and we need to avoid certain configurations of edges that would induce a forbidden configuration of triangles. Hall’s theorem is a very powerful tool to show the existence of perfect matchings, but it gives almost no control over the properties of these matchings.

The solution to this problem is to use an extra round of random sparsification: before applying Hall’s theorem, we randomly sparsify our graphs $G[W_v]$, and delete any edges that could possibly contribute to forbidden configurations. Hall’s theorem is sufficiently powerful that we can still find perfect matchings even after very harsh sparsification, and if we sparsify sufficiently harshly, it turns out that very few deletions will be required (this is closely related to the sparsification idea described in Section 1.2.5, which we will next discuss, and to the Brown–Erdős–Sós lower bound for the $(j, j-2)$ -problem).

2.2.5. Constraint focusing and sparsification. Recall from Section 1.2.4 that there is an issue of “constraint focusing”. Specifically, there is a constant $q_g > 0$ such that if we were to execute the strategy described so far on the complete graph K_N , at each step k w.h.p. about a q_g^k -fraction of triangles in $G[U_k]$ would be excluded from consideration due to choices made at previous steps. The performance of the generalized high-girth triple process has a severe dependence on the density of available triangles, which would restrict how rapidly the sizes of our vortex sets $U_0 \supseteq U_1 \supseteq \dots \supseteq U_\ell$ can decrease. This would cause a “parameter race”: we need our vortex sets to decrease as rapidly as possible, so that we can reach a final set U_ℓ that is small enough for efficient absorption before the situation gets so sparse that there is simply no room to make choices. As far as we can tell this is a losing race, no matter how carefully one studies the high-girth triple process.

As discussed in Section 1.2.5, the solution is to randomly sparsify K_N before we begin the iterative absorption process. Specifically, we start by finding a random set of triangles \mathcal{I} with the high-girth triple process (such that the graph of uncovered edges has density $p = |\mathcal{I}|^{-\nu} = N^{-\Omega(1)}$, for a very small constant

$\nu > 0$). We do this *after* fixing the vortex sets U_k and absorbing graph H ; this order of operations will be important for the “retrospective analysis” described in the next section. Doing things in this order does however introduce some subtleties to the analysis of the high-girth triple process (in particular, we need to analyze the process on “many different scales” to ensure that the graph of uncovered edges in each U_k is pseudorandom).

2.2.6. Weight systems and retrospective analysis. In [Section 2.2.2](#), we mentioned that for the generalized high-girth process to succeed, it suffices for our collection of forbidden configurations to satisfy certain “well-spreadness” properties. In particular, this involves upper bounds (in terms of j and $|\mathcal{R}|$) on the number of forbidden configurations with $j - 2$ triples that contain a particular set of triangles \mathcal{R} . However, statistics of this type do *not* fully capture the evolution of forbidden configurations during the high-girth triple process (in fact, it is not even true that typically all such statistics concentrate around their means).

We do believe that there exists some ensemble of statistics that does concentrate and does fully capture the behavior of the process, but it would be a formidable technical endeavor to characterize these statistics (especially since we need to work with general forbidden configurations, not just Erdős configurations). Of course, this raises the question of how we were able to actually analyze the generalized high-girth triple process. The answer is that it is actually not necessary to actively keep track of all relevant information. Instead, we can “remember the randomness” of the process: at each step, any given triangle was chosen with probability at most about $1/n^3$, and the number of steps so far was at most n^2 . So, for any constant $h \in \mathbb{N}$, the probability that a particular set of h triangles was chosen by the process is at most about $(1/n)^h$. We can use this to control high moments of various statistics, which allows one to prove probabilistic inequalities about forbidden configurations. This type of argument featured prominently in the analysis of Bohman and Warnke [\[8\]](#) (but not the analysis of Glock, Kühn, Lo, and Osthus [\[27\]](#)); actually this is the primary reason why [\[8\]](#) is a shorter paper than [\[27\]](#), despite their main result having stronger quantitative aspects.

In our proof of [Theorem 1.1](#), we take this kind of retrospective analysis much further. The entire iterative absorption proof (consisting of an initial sparsification process, ℓ randomized regularity-boosting steps and ℓ “cover down” steps, each of which consists of multiple random processes) is viewed as one random super-process, and we keep track of the approximate probabilities of the random choices made throughout. For example, we define the “level” $\text{lev}(T)$ of a triangle T to be the largest k for which T is contained within U_k ; it turns out that T is selected by our super-process with probability at most about $p/|U_{\text{lev}(T)}|$.

In order to prove probabilistic inequalities using this distributional information, we introduce the formalism of *weight systems*, which describe the quantities that one needs to estimate in order to control a desired statistic via moment estimation. Our main lemma very closely resembles the Kim–Vu polynomial concentration inequality [\[39\]](#), though our bounds are weaker and we do not require independence (one needs to control “planted” expected values, which can be interpreted as expected partial derivatives of certain polynomials). This lemma is proved using techniques related to Rödl and Ruciński’s deletion method [\[52\]](#).

2.3. Organization of the paper. We start with some preliminaries in [Section 3](#); these include well-known facts in addition to our main *weight systems* lemma, which will be used throughout the proof to keep track of forbidden configurations. In [Section 4](#) we present our high-girth absorber construction; this section includes our new *efficient absorption* ideas. In [Section 5](#) we record a slight generalization of the “regularity boosting” lemma from [\[1\]](#). We also prove a new general-purpose hypergraph regularization lemma in the same section. In [Section 6](#) we record some general lemmas about perfect matchings in random subgraphs of typical graphs.

[Sections 3](#) to [6](#) described above are short and self-contained. From [Section 7](#) onwards, we get into the meat of the proof, introducing special-purpose concepts and notation related to the algorithm described in [Section 2](#), and computing quantities related to this algorithm. In [Section 7](#) we introduce the notion of “well-spreadness” of a collection of forbidden configurations with respect to a vortex $U_0 \supseteq \dots \supseteq U_\ell$ and prove that this well-spreadness property holds for a certain collection of forbidden configurations induced by our absorber construction. Next, in [Section 8](#) we compute “weights” associated with various arrangements of forbidden configurations. These will be inputs to our weight system lemma, and will be used to control various quantities throughout the proof. The reader may wish to skip this section on first reading, as it consists mostly of rather technical computations, the details of which are not important to understand the rest of the proof. In [Section 9](#) we study a generalization of the process studied in [\[8, 27\]](#), which builds an approximate triangle-decomposition avoiding a given collection of forbidden configurations. Finally, in [Sections 10](#) and [11](#) we put everything together according to the strategy described in [Section 2](#), and prove [Theorems 1.1](#) and [1.3](#).

3. PRELIMINARIES

Recall that a Steiner triple system of order N is equivalent to a triangle-decomposition of K_N ; for the rest of the paper, we will only use the language of triangles and triangle-decompositions. In this language, we recall the set of configurations we will need to avoid.

Definition 3.1. For $g \geq 5$, we define an *Erdős g -configuration* to be a configuration of triangles which has g vertices and girth exactly g . That is to say, an Erdős configuration is a “minimal” configuration with two more vertices than triangles.

Note that by this definition, a *diamond* (pair of triangles sharing an edge) is not an Erdős g -configuration, and in fact the triangles in Erdős configurations are edge-disjoint. This is non-standard, but will be convenient for us, as diamonds play a rather special role in the analysis. We see that a Steiner triple system has girth greater than g if and only if it has no Erdős j -configuration for any $5 \leq j \leq g$.

The main fact that we need about Erdős configurations is that they are “minimal”.

Lemma 3.2. *Let \mathcal{E} be an Erdős j -configuration.*

- (1) *Every set of $2 \leq w \leq j - 3$ triangles in \mathcal{E} spans at least $w + 3$ vertices. Equivalently, every set of $1 \leq v \leq j - 4$ vertices of \mathcal{S} touches at least $v + 1$ of the triangles in \mathcal{S} .*
- (2) *Every set of $1 \leq w \leq j - 2$ triangles in \mathcal{E} spans at least $w + 2$ vertices. Equivalently, every set of $0 \leq v \leq j - 2$ vertices of \mathcal{S} touches at least v of the triangles in \mathcal{E} .*

Proof. Property (2) is a direct consequence of property (1). For property (1), observe that an Erdős j -configuration can be equivalently defined as a set of $j - 2$ triangles on j vertices, such that no subset of $3 < v < j$ vertices spans at least $v - 2$ triangles. \square

For the convenience of the reader, we also recall some standard concentration inequalities. The first is Freedman’s inequality. We write $\Delta X(i) = X(i + 1) - X(i)$ to denote one-step differences.

Lemma 3.3 (Freedman’s inequality [21]). *Let $(X(0), X(1), \dots)$ be a supermartingale with respect to a filtration $\mathbf{F}(0) \subseteq \mathbf{F}(1) \subseteq \mathbf{F}(2) \subseteq \dots$. Suppose that $|\Delta X(i)| \leq K$ for all i and let $V(i) = \sum_{j=0}^{i-1} \mathbb{E}[(\Delta X(j))^2 | \mathbf{F}(j)]$. Then for any $t, v > 0$ we have*

$$\Pr[X(i) \geq X(0) + t \text{ and } V(i) \leq v \text{ for some } i] \leq \exp\left(-\frac{t^2}{2v + 2Kt}\right).$$

We will really only use the following immediate corollary of Freedman’s inequality (to deduce this, note that for any random variable Y , if we have $|Y| \leq K$ then $\mathbb{E}Y^2 \leq K\mathbb{E}|Y|$).

Corollary 3.4. *Let $(X(0), \dots, X(m))$ be a supermartingale with respect to a filtration $\mathbf{F}(0) \subseteq \mathbf{F}(1) \subseteq \mathbf{F}(2) \subseteq \dots$. Suppose that $|\Delta X(i)| \leq K$ and $\mathbb{E}[|\Delta X(i)| | \mathbf{F}(i)] \leq D$ for all i . Then for any $t > 0$ we have*

$$\Pr[X(m) \geq X(0) + t] \leq \exp\left(-\frac{t^2}{2mKD + 2Kt}\right).$$

We also state a Chernoff bound for binomial and hypergeometric distributions (see for example [33, Theorems 2.1 and 2.10]). This will be used very frequently throughout the paper.

Lemma 3.5 (Chernoff bound). *Let X be either:*

- *a sum of independent random variables, each of which take values in $\{0, 1\}$, or*
- *hypergeometrically distributed (with any parameters).*

Then for any $\delta > 0$ we have

$$\Pr[X \leq (1 - \delta)\mathbb{E}X] \leq \exp(-\delta^2\mathbb{E}X/2), \quad \Pr[X \geq (1 + \delta)\mathbb{E}X] \leq \exp(-\delta^2\mathbb{E}X/(2 + \delta)).$$

3.1. Weight systems and moments. We will need some non-standard probabilistic machinery. Specifically, at many points in this paper, it will be necessary to prove crude high-probability upper bounds on certain random variables, using estimates on their moments. In this subsection we introduce the formalism of “weight systems” and use it to state a general lemma which we will use for all of these moment arguments. This general lemma is closely related to the Kim–Vu polynomial concentration inequality [39] (the quantities we need to control are almost the same), but the estimates in the proof more closely resemble the *deletion method* [52] of Rödl and Ruciński.

Definition 3.6. Fix a finite set \mathcal{W} . We say that a subset of \mathcal{W} is a *configuration*. A *weight system* $\vec{\pi}$ with ground set \mathcal{W} is an assignment of a nonnegative real number π_T to each $T \in \mathcal{W}$. Let \mathfrak{X} be a multiset of configurations. We define the *weight* of a subset $\mathcal{H} \subseteq \mathcal{W}$ with respect to \mathfrak{X} :

$$\psi^{(\vec{\pi})}(\mathfrak{X}, \mathcal{H}) = \sum_{\mathcal{H} \subseteq \mathcal{S} \in \mathfrak{X}} \prod_{T \in \mathcal{S} \setminus \mathcal{H}} \pi_T,$$

where the sum over \mathcal{S} is with multiplicity. We also define the *maximum weight* of \mathfrak{X} as

$$\kappa^{(\vec{\pi})}(\mathfrak{X}) = \max_{\mathcal{H} \subseteq \mathcal{W}} \psi^{(\vec{\pi})}(\mathfrak{X}, \mathcal{H}).$$

We suppress dependence on $\vec{\pi}$ where there is no ambiguity. For notational convenience notation we write $\psi(\mathfrak{X}, \emptyset) = \psi(\mathfrak{X})$.

Note that if \mathcal{H} is larger than all the sets in \mathfrak{X} , then $\psi(\mathfrak{X}, \mathcal{H}) = 0$. So, in the definition of $\kappa(\mathfrak{X})$ it suffices to take a maximum over all sets \mathcal{H} whose size is at most $\max_{\mathcal{S} \in \mathfrak{X}} |\mathcal{S}|$.

Remark. If each $\pi_T \leq 1$, we can consider a random subset of \mathcal{W} where each element $T \in \mathcal{W}$ is included with probability π_T independently. Then, $\psi(\mathfrak{X})$ is the expected number of elements of \mathfrak{X} (with multiplicity) which are fully included in this random subset. More generally $\psi(\mathfrak{X}, \mathcal{H})$ can be interpreted as a “planted” expectation, where we condition on the elements of \mathcal{H} lying in our random subset. Or, alternatively, we can interpret these weights in the language of polynomials (as in Kim–Vu polynomial concentration). Indeed, we can encode \mathfrak{X} as a multilinear polynomial $f_{\mathfrak{X}}$ in the variables $(x_T)_{T \in \mathcal{W}}$, with nonnegative integer coefficients. In this language, weights of the form $\psi(\mathfrak{X}, \mathcal{H})$ can be interpreted as expected values arising from partial derivatives of $f_{\mathfrak{X}}$ (where we differentiate with respect to the variables $(x_T)_{T \in \mathcal{H}}$).

For now, the reader may think of \mathcal{W} as being a set of triangles, so \mathfrak{X} is a set of configurations of triangles¹⁰. Then, $\psi(\mathfrak{X})$ is the expected number of configurations we expect to see in an appropriate random set of triangles. However, since we are interested in proving upper bounds that hold with very high probability, we need to account for the fact that “planting” a small number of triangles could dramatically increase this expected value, which leads us to the definition of $\kappa(\mathfrak{X})$.

In particular, the following lemma says that if a random subset $\mathcal{R} \subseteq \mathcal{W}$ is in some sense “bounded” by a weight system $\vec{\pi}$ for \mathcal{W} , then the number of $\mathcal{S} \in \mathfrak{X}$ which are fully included in \mathcal{R} is unlikely to be much greater than $\kappa^{(\vec{\pi})}(\mathfrak{X})$.

Lemma 3.7. *Fix integers $d, s \geq 1$ and a real number $C > 0$. Consider a finite set \mathcal{W} , a multiset \mathfrak{X} of subsets of \mathcal{W} that have size at most d , and a weight system $\vec{\pi}$ with ground set \mathcal{W} . Let \mathcal{R} be a random subset of \mathcal{W} such that*

$$\Pr[\mathcal{T} \subseteq \mathcal{R}] \leq C \prod_{T \in \mathcal{T}} \pi_T$$

for all $\mathcal{T} \subseteq \mathcal{W}$ with size at most ds . Let $X = X(\mathfrak{X})$ be the number of configurations $\mathcal{S} \in \mathfrak{X}$ such that $\mathcal{S} \subseteq \mathcal{R}$. Then, for any $\gamma > 0$ we have

$$\Pr[X \geq \gamma \kappa(\mathfrak{X})] \leq \frac{C(ds)^{ds}}{\gamma^s}.$$

Proof. It suffices to prove that

$$\mathbb{E}X^s \leq C(ds)^{ds} \kappa(\mathfrak{X})^s; \quad (3.1)$$

the desired result then follows from an application of Markov’s inequality to the random variable X^s . To prove (3.1), let $x_{\mathcal{S}} \in \mathbb{Z}_{\geq 0}$ be the multiplicity of a set \mathcal{S} in \mathfrak{X} , write

$$M_t = \sum_{\substack{\mathcal{S}_1, \dots, \mathcal{S}_t \subseteq \mathcal{W}: \\ \text{each } |\mathcal{S}_i| \leq d}} \prod_{i=1}^t x_{\mathcal{S}_i} \prod_{T \in \mathcal{S}_1 \cup \dots \cup \mathcal{S}_t} \pi_T,$$

and note that

$$\mathbb{E}X^s = \sum_{\substack{\mathcal{S}_1, \dots, \mathcal{S}_s \subseteq \mathcal{W}: \\ \text{each } |\mathcal{S}_i| \leq d}} \mathbb{E} \left[\mathbb{1}_{\mathcal{S}_1 \cup \dots \cup \mathcal{S}_s \subseteq \mathcal{R}} \prod_{i=1}^s x_{\mathcal{S}_i} \right] \leq C M_s.$$

Now, we recursively bound M_t . Note that $M_1 = \psi(\mathfrak{X}) \leq \kappa(\mathfrak{X})$, and for $t \geq 1$ we can sum over all the ways \mathcal{S}_t can intersect $\mathcal{S}_1 \cup \dots \cup \mathcal{S}_{t-1}$ to obtain

$$\begin{aligned} M_t &= \sum_{\substack{\mathcal{S}_1, \dots, \mathcal{S}_{t-1} \subseteq \mathcal{W}: \\ \text{each } |\mathcal{S}_i| \leq d}} \left(\prod_{i=1}^{t-1} x_{\mathcal{S}_i} \prod_{T \in \mathcal{S}_1 \cup \dots \cup \mathcal{S}_{t-1}} \pi_T \right) \left(\sum_{\mathcal{L} \subseteq \mathcal{S}_1 \cup \dots \cup \mathcal{S}_{t-1}} \sum_{\substack{\mathcal{S}_t \subseteq \mathcal{W}: |\mathcal{S}_t| \leq d, \\ \mathcal{L} = \mathcal{S}_t \cap (\mathcal{S}_1 \cup \dots \cup \mathcal{S}_{t-1})}} x_{\mathcal{S}_t} \prod_{T \in \mathcal{S}_t \setminus \mathcal{L}} \pi_T \right) \\ &\leq \sum_{\substack{\mathcal{S}_1, \dots, \mathcal{S}_{t-1} \subseteq \mathcal{W}: \\ \text{each } |\mathcal{S}_i| \leq d}} \left(\prod_{i=1}^{t-1} x_{\mathcal{S}_i} \prod_{T \in \mathcal{S}_1 \cup \dots \cup \mathcal{S}_{t-1}} \pi_T \right) \left(\sum_{\mathcal{L} \subseteq \mathcal{S}_1 \cup \dots \cup \mathcal{S}_{t-1}} \psi(\mathfrak{X}, \mathcal{L}) \right). \end{aligned} \quad (3.2)$$

¹⁰This mostly describes the settings in which we will apply the machinery of weight systems, though we will sometimes consider more general situations (where \mathcal{W} is a mixture of edges and triangles).

Note that $|\mathcal{S}_1 \cup \dots \cup \mathcal{S}_{t-1}| \leq d(t-1)$, so there are at most $(d(t-1) + 1)^d \leq (dt)^d$ subsets $\mathcal{L} \subseteq \mathcal{S}_1 \cup \dots \cup \mathcal{S}_{t-1}$ which have size at most d (if $|\mathcal{L}| > d$ then $\psi(\mathfrak{X}, \mathcal{L}) = 0$). Also, recall that $\psi(\mathfrak{X}, \mathcal{L}) \leq \kappa(\mathfrak{X})$, so $M_t \leq (dt)^d \kappa(\mathfrak{X}) M_{t-1}$, from which we deduce that $M_s \leq (ds)^{ds} \kappa(\mathfrak{X})^s$. Then, (3.1) follows. \square

We also state a convenient asymptotic corollary of Lemma 3.7.

Corollary 3.8. *Consider a finite set \mathcal{W} of size at most $n^{O(1)}$ and a multiset \mathfrak{X} of subsets of \mathcal{W} each of which has size $O(1)$. Consider a weight system $\vec{\pi}$ with ground set \mathcal{W} and let \mathcal{R} be a random subset of \mathcal{W} such that*

$$\Pr[\mathcal{T} \subseteq \mathcal{R}] = O\left(\prod_{T \in \mathcal{T}} \pi_T + n^{-\omega(1)}\right)$$

for all $\mathcal{T} \subseteq \mathcal{W}$. Let $X = X(\mathfrak{X})$ be the number of configurations $\mathcal{S} \in \mathfrak{X}$ such that $\mathcal{S} \subseteq \mathcal{R}$. Then, with probability $1 - n^{-\omega(1)}$ we have $X \leq n^{o(1)} \kappa(\mathfrak{X})$.

We conclude the section with a simple application of Lemma 3.7, corresponding to the special case where \mathfrak{X} is a collection of singleton sets and the weights are all equal (our weight systems machinery is not really necessary for such a simple case).

Corollary 3.9. *Consider a finite set \mathcal{W} of size at most $n^{O(1)}$, and let \mathcal{R} be a random subset of \mathcal{W} , such that for every choice of distinct $w_1, \dots, w_t \in \mathcal{W}$ we have $\Pr[w_1, \dots, w_t \in \mathcal{R}] \leq p^t + n^{-\omega(1)}$. Then, with probability $1 - n^{-\omega(1)}$ we have $|\mathcal{R}| \leq n^{o(1)} \max(p|\mathcal{W}|, 1)$.*

4. EFFICIENT HIGH-GIRTH ABSORBERS

Recall that a graph is *triangle-divisible* if all its degrees are even and its number of edges is divisible by 3. This is a necessary but insufficient condition for triangle-decomposability. In this section we explicitly define a high-girth “absorbing structure” that will allow us to find a triangle-decomposition extending any triangle-divisible graph on a specific vertex set. Importantly, the size of this structure is only polynomial in the size of the distinguished vertex set, which is needed for our proof strategy (recall the discussion in Section 1.2.2). For a set of triangles \mathcal{R} , let $V(\mathcal{R})$ be the set of all vertices in these triangles.

The following theorem encapsulates the properties of our absorbing structure. Apart from the fact that we can find a high-girth triangle-decomposition extending any triangle-divisible graph on a specific vertex set, we also need a rather technical property (Ab2) stating that the triangle-decompositions can be chosen in a way such that there are only a small number of ways for Erdős configurations to intersect the triangles in these decompositions. This is needed so we can ensure that at the very end, when the absorber is used to complete the final high-girth system, no Erdős configurations are introduced.

Theorem 4.1. *There is $C_{4.1} \in \mathbb{N}$ so that for $g \in \mathbb{N}$ there exists $M_{4.1}(g) \in \mathbb{N}$ such that the following holds. For any $m \geq 1$, there is a graph H with at most $M_{4.1}(g)m^{C_{4.1}}$ vertices containing a distinguished independent set X of size m and satisfying the following properties.*

Ab1 *For any triangle-divisible graph L on the vertex set X there exists a triangle-decomposition \mathcal{S}_L of $L \cup H$ which has girth greater than g .*

Ab2 *Let $\mathcal{B} = \bigcup_L \mathcal{S}_L$ (where the union is over all triangle-divisible graphs L on the vertex set X) and consider any graph K containing H as a subgraph. Say that a triangle in K is nontrivially H -intersecting if it is not one of the triangles in \mathcal{B} , but contains a vertex in $V(H) \setminus X$.*

Then, for every set of at most g triangles \mathcal{R} in K , there is a subset $\mathcal{L}_{\mathcal{R}} \subseteq \mathcal{B}$ of at most $M_{4.1}(g)$ triangles such that every Erdős configuration \mathcal{E} on at most g vertices which includes \mathcal{R} must either satisfy $\mathcal{E} \cap \mathcal{B} \subseteq \mathcal{L}_{\mathcal{R}}$ or must contain a nontrivially H -intersecting triangle $T \notin \mathcal{R}$.

Remark. Note that Ab1, with L as the empty graph on the vertex set X , implies that H itself is triangle-decomposable (hence triangle-divisible).

We will prove Theorem 4.1 by chaining together some special-purpose graph operations.

Definition 4.2 (Path-cover). Let the *path-cover* $\wedge X$ of a vertex set X be the graph obtained as follows. Start with the empty graph on X . Then, for every unordered pair of u, v of distinct vertices in X , add $6|X|^2$ new paths of length 2 between u and v , introducing a new vertex for each (so in total, we introduce $6|X|^2 \binom{|X|}{2}$ new vertices). We call these new length-2 paths *augmenting paths*.

The key point is that for any graph L on the vertex set X with even degrees, the edges of $L \cup \wedge X$ can be decomposed into short cycles.

Lemma 4.3. *If a graph L on X has even degrees, then $L \cup \wedge X$ can be decomposed into cycles of length at most 5, such that there are more cycles of length 4 than of length 5. Additionally, if L is triangle-divisible, then so is $L \cup \wedge X$.*

Proof. By finding an Euler tour in each connected component, we decompose L into edge-disjoint cycles. There are fewer than $|X|^2$ cycles in this decomposition. Now, consider one of these cycles C , and write v_1, \dots, v_ℓ for its vertices. For each $2 \leq i \leq \ell - 1$, add two of the augmenting paths between v_1 and v_i , to obtain an augmented graph $C' \subseteq L \cup \wedge X$ containing C . By design, C' has an edge-decomposition into cycles of lengths $3, 5, 5, \dots, 5, 3$. We can do this edge-disjointly for each of the cycles in the cycle-decomposition of L , since each cycle only uses at most 2 augmenting paths between any pair. The remaining unused edges of $L \cup \wedge X$ consist of augmenting paths between vertices. There are an even number of such augmenting paths between each pair of vertices since the above edge-decompositions involve adding pairs of augmenting paths at a time. Therefore, we can decompose the remaining edges into cycles of length 4. Since we included sufficiently many augmenting paths in $\wedge X$, it follows that a majority of cycles in our final decomposition have length 4.

Finally, the last statement follows from the fact that $\wedge X$ is triangle-divisible. \square

It is proved in [1, Lemma 3.2] that for any triangle-divisible graph H , there is a graph $A(H)$ containing the vertex set of H as an independent set such that $A(H)$ and $A(H) \cup H$ are both triangle-decomposable.

Definition 4.4 (Cycle-cover). Let \mathcal{H} be the (finite) set of all graphs which are a triangle, or whose edges decompose into one cycle of length 4 and one cycle of length 5, or into three cycles of length 4. All such graphs are triangle-divisible (but might not be triangle-decomposable).

Let the *cycle-cover* of a vertex set Y be the graph ΔY obtained as follows. Beginning with Y , for every $H \in \mathcal{H}$ and every injection $f : V(H) \rightarrow Y$ we add a copy of $A(H)$, such that each $v \in V(H) \subseteq V(A(H))$ in the copy of $A(H)$ coincides with $f(v)$ in Y . We do this in a vertex-disjoint way, introducing $|V(A(H)) \setminus V(H)|$ new vertices each time. (Think of the copy of $A(H)$ as being “rooted” on a specific set of vertices in Y , and otherwise being disjoint from everything else.)

Lemma 4.5. *Let $Y = V(\wedge X)$ be the vertex set of the graph $\wedge X$. If a graph L on X is triangle-divisible, then $L \cup \wedge X \cup \Delta Y$ admits a triangle-decomposition.*

Proof. By Lemma 4.3, $L \cup \wedge X$ is triangle-divisible and can be decomposed into cycles of length 3, 4, and 5, with more cycles of length 4 than 5. We group the cycles into individual triangles, pairs of a 4-cycle and 5-cycle, and triples of 4-cycles, all of which are contained within Y . (Triangle-divisibility of $L \cup \wedge X$ implies that there will be no leftover 4-cycles.) To obtain a triangle-decomposition we use the defining property of the graphs $A(H)$ and the definition of ΔY : for every $H \in \mathcal{H}$ and injection $f : V(H) \rightarrow Y$, decompose the corresponding image of $A(H) \cup H$ into triangles if $f(H)$ comprises one of our groups, and otherwise decompose the corresponding image of $A(H)$ into triangles. \square

If $|X| = m$, then Lemma 4.5 implies that $\wedge X \cup \Delta(V(\wedge X))$ can “absorb” any triangle-divisible graph on X , though not necessarily in a high-girth manner. The next transformation will provide us with a girth guarantee.

Definition 4.6 (Sphere-cover). Let the *g-sphere-cover* of a vertex set Z be the graph $\circ_g Z$ obtained by the following procedure. For every triple T of distinct vertices of Z , arbitrarily label these vertices as a, b_1, b_2 . Then, append a “g-sphere” to the triple. Namely, first add $2g - 1$ new vertices b_3, \dots, b_{2g}, c . Then add the edges ab_j for $3 \leq j \leq 2g$, the edges cb_j for $1 \leq j \leq 2g$, the edges $b_j b_{j+1}$ for $2 \leq j \leq 2g - 1$, and the edge $b_{2g} b_1$.

Note that every such g -sphere Q itself has a triangle-decomposition: specifically, we define the *out-decomposition* to consist of the triangles

$$cb_2 b_3, ab_3 b_4, cb_4 b_5, ab_5 b_6, \dots, cb_{2g} b_1.$$

We also identify a particular triangle-decomposition of the edges $Q \cup T$: the *in-decomposition* consists of the triangles

$$cb_1 b_2, ab_2 b_3, cb_3 b_4, ab_4 b_5, \dots, ab_{2g} b_1.$$

For a triple T in Z , let $\mathcal{B}^\circ(T)$ be the set of all triangles in the in- and out-decompositions of the g -sphere associated to T . We emphasize that $T \notin \mathcal{B}^\circ(T)$.

Remark. The g -sphere-cover depends on an ordering of every triple of vertices in Z , so is technically not unique, but this is of little consequence to us. Also, we remark that the object we call a g -sphere is usually called a *cycle of length $2g$* in design theory; we prefer to avoid this terminology as our construction also involves graph cycles.

The key properties of the g -sphere-cover are encapsulated in the following lemma.

Lemma 4.7. *Fix an integer $g > 2$. Consider a vertex set Z and its g -sphere-cover $\circ_g Z$, and a triple of vertices T in Z . Let $\mathcal{E} \subseteq \mathcal{B}^\circ(T)$ be a non-empty set of at most g edge-disjoint triangles.*

- (1) There is $v \in V(\mathcal{B}^\circ(T)) \setminus V(T)$ which appears in exactly one of the triangles in \mathcal{E} .
(2) $|V(\mathcal{E}) \setminus V(T)| \geq |\mathcal{E}|$.

Proof. Write a, b_1, b_2 for the three vertices of T , and let b_3, \dots, b_{2g}, c be as in [Definition 4.6](#).

Consider the cycle graph C_{2g} on the vertices b_1, \dots, b_{2g} (running through these vertices in order). By inspecting [Definition 4.6](#), we see that every triangle in $\mathcal{E} \subseteq \mathcal{B}^\circ(T)$ contains an edge of C_{2g} , and no edge of C_{2g} appears in more than one triangle of \mathcal{E} (since the triangles in \mathcal{E} are edge-disjoint). So, we may identify each triangle of \mathcal{E} with a unique edge of C_{2g} , to give a nonempty subgraph $S \subseteq C_{2g}$.

For (1), we observe that S has at least two degree-1 vertices (since $g < 2g$). If the only degree-1 vertices of S are b_1 and b_2 , then $\mathcal{E} = \{cb_1b_2\}$ (since $g < 2g - 1$), in which case c satisfies the desired conclusion. Otherwise, there is some degree-1 vertex different from b_1, b_2 in S , and this vertex satisfies the desired conclusion.

For (2), note that S has more vertices than edges, so the desired result follows immediately unless S contains both b_1 and b_2 . But if S contains both these vertices then either it is disconnected (in which case it contains at least two more vertices than edges), or S contains the edge b_1b_2 (in which case S additionally contains c). In all cases, we are done. \square

[Lemma 4.7\(1\)](#) implies that $\circlearrowright_g Z$ can “transform” arbitrary triangle-decompositions on Z into high-girth triangle-decompositions, as follows.

Lemma 4.8. *Let $\mathcal{C} = \bigcup_{T \in \binom{Z}{3}} \mathcal{B}^\circ(T)$. If a graph L on the vertex set Z is triangle-decomposable, then $L \cup \circlearrowright_g Z$ has a triangle-decomposition using only triangles of \mathcal{C} , with girth greater than g .*

Proof. We have a triangle-decomposition of L . Classify every triple of vertices in L as “in” or “out” depending on whether it is in the triangle-decomposition or not. Let $Q(T)$ be the g -sphere associated with a triple T .

For every “in” triple T , remove it from our decomposition and instead cover the edges of $T \cup Q(T)$ using the in-decomposition of $Q(T)$. For every “out” triple T , cover the edges of $Q(T)$ using the out-decomposition of $Q(T)$. Either way, we have only used triangles in $\mathcal{B}^\circ(T)$.

We claim that \mathcal{C} contains no Erdős j -configuration \mathcal{E} for $5 \leq j \leq g$. Suppose for the sake of contradiction that there were such a configuration $\mathcal{E} \subseteq \mathcal{C}$. Observe that \mathcal{E} cannot have a vertex which appears in exactly one triangle (recall [Lemma 3.2\(1\)](#) and that $j \geq 5$). Consider some triple T in Z such that $\mathcal{E} \cap \mathcal{B}^\circ(T)$ is nonempty. Erdős configurations contain edge-disjoint triangles, so [Lemma 4.7\(1\)](#) implies there is a vertex $v \in V(\mathcal{B}^\circ(T)) \setminus V(T) = V(\mathcal{B}^\circ(T)) \setminus Z$ in exactly one triangle of $\mathcal{E} \cap \mathcal{B}^\circ(T)$.

However, $\mathcal{E} \subseteq \mathcal{C}$ and thus the triangles in $\mathcal{E} \setminus \mathcal{B}^\circ(T)$ are vertex-disjoint from $V(\mathcal{B}^\circ(T)) \setminus Z$ by construction. We conclude that \mathcal{E} has exactly one triangle containing v . This contradicts our earlier observation! \square

We are ready to prove [Theorem 4.1](#).

Proof of [Theorem 4.1](#). We start with a vertex set X of size m . Let $Y = V(\wedge X)$ and $Z = V(\triangle Y)$, and take $H = \wedge X \cup \triangle Y \cup \circlearrowright_g Z$. [Ab1](#) is a direct consequence of [Lemmas 4.5](#) and [4.8](#). It is also easy to see that H has the claimed size, and that $\mathcal{B} \subseteq \bigcup_{T \in \binom{Z}{3}} \mathcal{B}^\circ(T)$.

Next, we consider [Ab2](#). For a triple of vertices T in Z , recall that $\mathcal{B}^\circ(T)$ is the collection of all triangles in the in-decomposition and the out-decomposition associated with the g -sphere on T , and let $W(T) = V(\mathcal{B}^\circ(T)) \setminus V(T) \subseteq V(H) \setminus Z$. Note that the $\mathcal{B}^\circ(T)$ are disjoint from each other (having no triangles in common), and the $W(T) = V(\mathcal{B}^\circ(T)) \setminus V(T)$ are vertex-disjoint from each other.

Now, given a collection of at most g triangles \mathcal{R} in K , let $\mathcal{T}_{\mathcal{R}}$ be the collection of triples T within Z such that $W(T)$ shares a vertex with \mathcal{R} . Let $\mathcal{L}_{\mathcal{R}} = \bigcup_{T \in \mathcal{T}_{\mathcal{R}}} \mathcal{B}^\circ(T)$. We claim that $\mathcal{L}_{\mathcal{R}}$ satisfies the desired property, noting that its size is bounded by a constant depending only on g .

Consider an Erdős j -configuration \mathcal{E} , for $5 \leq j \leq g$ (consisting of triangles in K), which includes \mathcal{R} and contains some $T^* \in \mathcal{B} \setminus \mathcal{L}_{\mathcal{R}}$. Suppose for the sake of contradiction that every $T' \in \mathcal{E} \setminus (\mathcal{R} \cup \mathcal{B})$ is disjoint from $V(H) \setminus X$. There is a unique triple T_0 within Z with $T^* \in \mathcal{B}^\circ(T_0)$. Let $\mathcal{K} = \mathcal{E} \cap \mathcal{B}^\circ(T_0)$. Note that \mathcal{K} is nonempty (as $T^* \in \mathcal{K}$), and it contains at most g triangles, all edge-disjoint, so by [Lemma 4.7\(1\)](#) there is $v \in W(T_0)$ which appears in exactly one of the triangles in \mathcal{K} . Let $\mathcal{W} \subseteq \mathcal{E}$ be the subset of triangles of \mathcal{E} which contain v .

By assumption, $v \in V(H) \setminus X$ cannot be a vertex in any of the triangles in $\mathcal{E} \setminus (\mathcal{R} \cup \mathcal{B})$, so $\mathcal{W} \subseteq \mathcal{E} \cap (\mathcal{R} \cup \mathcal{B})$. If v was in one of the triangles in \mathcal{R} , then by definition we would have $T_0 \in \mathcal{T}_{\mathcal{R}}$, which would imply $T^* \in \mathcal{L}_{\mathcal{R}}$, a contradiction. Hence we deduce $\mathcal{W} \subseteq \mathcal{E} \cap \mathcal{B}$. But now note that all of the triangles in $\bigcup_{T \neq T_0} \mathcal{B}^\circ(T)$ have all their vertices in $Z \cup \bigcup_{T \neq T_0} W(T)$, which is disjoint from $W(T_0)$ by construction. So $\mathcal{W} \subseteq \mathcal{E} \cap \mathcal{B}^\circ(T_0) = \mathcal{K}$. Finally, v is in exactly one triangle of \mathcal{K} , so we deduce that $|\mathcal{W}| = 1$. This contradicts [Lemma 3.2\(1\)](#) (recalling that $j \geq 5$): every vertex of an Erdős configuration must be in at least 2 triangles. \square

5. REGULARIZATION

5.1. Triangle regularization. Given a set of triangles with suitable regularity and “extendability” properties, the following lemma finds a subset which is substantially more regular. This lemma is closely related to [1, Lemma 4.2] and the proof is almost the same (the idea is to sample from a fractional triangle-decomposition), however we need to make certain adjustments in our setting.

Lemma 5.1. *There is $n_{5.1} \in \mathbb{N}$ such that the following holds for all $n \geq n_{5.1}$. Suppose $C \geq 2$ and $p \in (n^{-1/6}, 1)$, let G be a graph on n vertices and let \mathcal{T} be a collection of triangles of G , satisfying the following properties.*

- (1) *Every edge $e \in E(G)$ is in $(1 \pm 1/(12C^5))p^2n$ triangles of \mathcal{T} .*
- (2) *For every set $S \subseteq V(G)$ with $2 \leq |S| \leq 4$ forming a clique in G , there are between $C^{-1}p^{|S|}n$ and $Cp^{|S|}n$ vertices $u \in V(G) \setminus S$ which form a triangle in \mathcal{T} with every distinct pair $v, w \in S$.*

Then, there is a subcollection $\mathcal{T}' \subseteq \mathcal{T}$ such that every edge $e \in E(G)$ is in $(1 \pm n^{-1/4})p^2n/4$ triangles of \mathcal{T}' .

Proof. Let \mathcal{T}_5 be the collection of all copies of K_5 in G , such that all triangles in this K_5 are present in \mathcal{T} . For every edge e in $E(G)$ let $\mathcal{T}(e) \subseteq \mathcal{T}$ be the subcollection of triangles in \mathcal{T} which contain e , and let $\mathcal{T}_5(e) \subseteq \mathcal{T}_5$ be the subcollection of 5-cliques in \mathcal{T}_5 which contain e .

Also, let

$$c_e = \frac{p^2n - |\mathcal{T}(e)|}{|\mathcal{T}_5(e)|}.$$

Iteratively applying assumption 2 in the lemma statement (and then compensating for double-counting), we see that $|\mathcal{T}_5(e)| \geq (C^{-1}p^2n)(C^{-1}p^3n)(C^{-1}p^4n)/6$ for all $e \in E(G)$, implying that

$$|c_e| \leq \frac{p^2n/(12C^5)}{C^{-3}p^9n^3/6} \leq \frac{1}{2C^2p^7n^2}.$$

For $e \in E(G)$ and $J \in \mathcal{T}_5(e)$, define $\psi_{e,J}: \mathcal{T} \rightarrow \mathbb{R}$ by

$$\psi_{e,J}(T) = \begin{cases} -1/6 & \text{if } |V(T) \cap e| = 1 \text{ and } T \subseteq J, \\ 1/3 & \text{if } |V(T) \cap e| \neq 1 \text{ and } T \subseteq J, \\ 0 & \text{if } T \not\subseteq J. \end{cases}$$

The key point of this construction is that it provides a simple encoding of certain delta functions: for any $e, e' \in E(G)$ and $J \in \mathcal{T}_5(e)$, one can check (via case analysis) that

$$\sum_{T \in \mathcal{T}(e')} \psi_{e,J}(T) = \mathbb{1}_{e'=e}.$$

Now, define $\psi: \mathcal{T} \rightarrow \mathbb{R}$ by

$$\psi(T) = \frac{1}{4} + \frac{1}{4} \sum_{e \in E(G)} c_e \sum_{J \in \mathcal{T}_5(e)} \psi_{e,J}(T).$$

For a given $T \in \mathcal{T}$, there are at most $(Cp^3n)(Cp^4n)/2$ many $J \in \mathcal{T}_5$ such that $T \subseteq J$ (to see this, iteratively apply assumption 2 in the lemma statement). We deduce that $\psi(T) \in [0, 1]$, since

$$\begin{aligned} |\psi(T) - 1/4| &\leq \frac{1}{4} \sum_{e \in E(G), J \in \mathcal{T}_5(e)} |c_e| |\psi_{e,J}(T)| \leq \frac{1}{4} \sum_{e \in E(G), J \in \mathcal{T}_5(e): T \subseteq J} \frac{1}{2C^2p^7n^2} \cdot \frac{1}{3} \\ &\leq \left(\frac{1}{2}C^2p^7n^2\right) \cdot \binom{5}{2} \cdot \frac{1}{24C^2p^7n^2} \leq \frac{1}{4}. \end{aligned}$$

Also, for any $e \in E(G)$ we have

$$\begin{aligned} \sum_{T \in \mathcal{T}(e)} \psi(T) &= \frac{1}{4} |\mathcal{T}(e)| + \frac{1}{4} \sum_{e' \in E(G)} c_{e'} \sum_{J \in \mathcal{T}_5(e')} \sum_{T \in \mathcal{T}(e)} \psi_{e',J}(T) \\ &= \frac{1}{4} |\mathcal{T}(e)| + \frac{1}{4} \sum_{e' \in E(G)} c_{e'} \sum_{J \in \mathcal{T}_5(e')} \mathbb{1}_{e=e'} = \frac{1}{4} |\mathcal{T}(e)| + \frac{1}{4} c_e |\mathcal{T}_5(e)|. \end{aligned}$$

By the definition of c_e , it follows that

$$\sum_{T \in \mathcal{T}(e)} \psi(T) = \frac{1}{4} p^2n.$$

Now, we define a random subset $\mathcal{T}' \subseteq \mathcal{T}$ by including each $T \in \mathcal{T}$ with probability $\psi(T)$ independently. For every $e \in E(G)$, the expected number of $T \in \mathcal{T}'$ containing e is exactly $p^2n/4$. By a Chernoff bound,

e is in $(1 \pm n^{-1/4})p^2n/4$ of the triangles in \mathcal{T}' , with probability at least $1 - \exp(-\Omega(n^{-1/2} \cdot p^2n))$. Since $p \geq n^{-1/6}$, a union bound over at most $\binom{n}{2}$ edges e concludes the proof. \square

5.2. General hypergraph regularization. In this section we prove a general lemma stating that for any not-too-dense hypergraph \mathfrak{G} , we can add a small number of random hyperedges to \mathfrak{G} to obtain a nearly-regular hypergraph. Moreover, we can specify a sparse hypergraph \mathfrak{H} of hyperedges we would like to avoid adding. For the proof of [Theorem 1.1](#), we will apply this lemma to an auxiliary hypergraph \mathfrak{G} whose vertices correspond to triangles in some graph G , and whose hyperedges correspond to forbidden configurations of triangles.

In the following lemma statement, the term “degree” always refers to the degree of a vertex, i.e., the number of hyperedges containing that vertex (in hypergraphs there are “higher-order” notions of degree, which will not concern us here).

It strikes us that this lemma may be of independent interest, so we prove a stronger lemma statement than is strictly necessary for the proof of [Theorem 1.1](#) (in particular, the proof is iterative, but only one step of this iteration is actually necessary for us).

Lemma 5.2. *Fix a constant $k \in \mathbb{N}$. Consider n -vertex k -uniform hypergraphs $\mathfrak{G}, \mathfrak{H}$, such that $\mathfrak{G} \subseteq \mathfrak{H}$ and such that the maximum degree of \mathfrak{H} is at most $\binom{n-1}{k-1}/(36 \cdot 2^k)$. Let d_{\max} be the maximum degree of \mathfrak{G} . Let $\mathbb{G}^{(k)}(n, p)$ be a random k -uniform hypergraph on the same vertex set as \mathfrak{G} , where each of the $\binom{n}{k}$ possible edges are present with probability p independently. There is a random simple k -uniform hypergraph \mathfrak{G}' , containing no edge of \mathfrak{H} , satisfying the following properties.*

- (1) *The maximum degree of $\mathfrak{G} \cup \mathfrak{G}'$ is at most $9d_{\max}$ with probability $1 - n^{-\omega(1)}$;*
- (2) *The vertex degrees of $\mathfrak{G} \cup \mathfrak{G}'$ all differ by at most $(\log n)^2$ with probability $1 - n^{-\omega(1)}$;*
- (3) *\mathfrak{G}' can be coupled as a subgraph of $\mathbb{G}^{(k)}(n, p)$, where $p = 2 \cdot 2^k d_{\max} / \binom{n-1}{k-1}$.*

Proof. We will iteratively construct a sequence of random hypergraphs $\mathfrak{G}(1), \dots, \mathfrak{G}(\tau)$, where the number of hypergraphs $\tau \geq 0$ is itself random (we can think of τ as a stopping time). Specifically, given outcomes of $\mathfrak{G}(1), \dots, \mathfrak{G}(t)$, for some $t \geq 0$, either we will take $\tau = t$ and stop, or we will define a random hypergraph $\mathfrak{G}(t+1)$ whose distribution depends on the outcomes of $\mathfrak{G}(1), \dots, \mathfrak{G}(t)$. This distribution will be chosen in such a way as to typically reduce the difference between the maximum and minimum degree, without introducing too many new edges. At the end, we will take $\mathfrak{G}' = (\mathfrak{G}(1) \cup \dots \cup \mathfrak{G}(\tau)) \setminus \mathfrak{H}$.

So, consider outcomes of $\mathfrak{G}(1), \dots, \mathfrak{G}(t)$. Write $F(t)$ for the difference between the maximum and the minimum degrees of $\mathfrak{G} \cup ((\mathfrak{G}(1) \cup \dots \cup \mathfrak{G}(t)) \setminus \mathfrak{H})$ (so $F(0)$ is the difference between the maximum and minimum degrees of \mathfrak{G}). If one of the following three criteria is met, we take $\tau = t$ and stop.

- (A) $F(t) \leq (\log n)^2$;
- (B) $F(t) \geq F(t-1)/2$;
- (C) $\mathfrak{G}(t)$ has maximum degree greater than $4F(t-1)$.

The idea is that if (A) occurs then we have successfully regularized our hypergraph, while if (B) or (C) occurs then we have failed to sufficiently reduce $F(t)$ or have introduced too many new hyperedges (in which case we give up; we will show that this is unlikely). Note that the process cannot continue for more than $\lceil \log_2(d_{\max}/(\log n)^2) \rceil$ steps (before this time, either (A) or (B) must occur).

If none of (A–C) hold at step t , then we define the next hypergraph $\mathfrak{G}(t+1)$ as follows. For each vertex v let $d_v(t)$ be the degree of v in $\mathfrak{G} \cup ((\mathfrak{G}(1) \cup \dots \cup \mathfrak{G}(t)) \setminus \mathfrak{H})$ and let $d_{\max}(t) = \max_{v \in V(\mathfrak{G})} d_v(t)$. For each vertex v , define its “weight” $w_v = d_{\max}(t) + F(t) - d_v(t)$. Hypothetically speaking, if we were able to choose $\mathfrak{G}(t+1)$ in such a way that the degree of each vertex v in $\mathfrak{G}(t+1) \setminus (\mathfrak{G}(1) \cup \dots \cup \mathfrak{G}(t) \cup \mathfrak{H})$ were precisely equal to w_v , then $\mathfrak{G} \cup ((\mathfrak{G}(1) \cup \dots \cup \mathfrak{G}(t+1)) \setminus \mathfrak{H})$ would be regular, i.e., $F(t+1) = 0$. We will not quite be able to manage this, but we will be able to get close by randomly sampling according to the weights w_v .

Note that $F(t) \leq w_v \leq 2F(t)$, so the weights of different vertices differ by a multiplicative factor of at most 2. Define $W = \sum_{S \in \binom{V(\mathfrak{G})}{k-1}} \prod_{v \in S} w_v$ (where $\binom{V(\mathfrak{G})}{k-1}$ denotes the collection of all $(k-1)$ -sets of vertices in \mathfrak{G}), and define the next random hypergraph $\mathfrak{G}(t+1)$ by including each k -set of vertices S with probability $(\prod_{v \in S} w_v)/W \leq 2^k F(t) / \binom{n-1}{k-1}$ independently.

We now prove (3). Note that, conditional on outcomes of $\mathfrak{G}(1), \dots, \mathfrak{G}(t)$ such that (A–C) do not hold, the random hypergraph $\mathfrak{G}(t+1)$ can be coupled as a subgraph of $\mathbb{G}^{(k)}(n, p_t)$, where $p_t = 2^k(2^{-t}F(0)) / \binom{n-1}{k-1}$ (here we are using that $F(t) \leq 2^{-t}F(0)$, since $\tau > t$). It follows that $\bigcup_{t=1}^{\tau} \mathbb{G}^{(k)}(n, p_t)$ can be coupled as a subgraph of $\mathbb{G}^{(k)}(n, p)$, where

$$p = \sum_{t=1}^{\infty} p_t \leq 2p_1 \leq 2 \cdot \frac{2^k d_{\max}}{\binom{n-1}{k-1}} \leq 1.$$

So, (3) holds.

Now, we claim that (conditioning on any outcomes of $\mathfrak{G}(1), \dots, \mathfrak{G}(t)$, but subject to the randomness of $\mathfrak{G}(t+1)$), with probability $1 - n^{-\omega(1)}$ we have that $F(t+1) \leq F(t)/2$, and that $\mathfrak{G}(t+1)$ has maximum degree at most $4F(t)$. Taking a union bound over at most $t \leq \lceil \log_2(d_{\max}/(\log n)^2) \rceil$ steps, it will follow that the probability (B) or (C) ever occur is at most $n^{-\omega(1)}$. Note that if (B) and (C) never occur then we have $F(0) + \dots + F(\tau) \leq F(0)(1 + 1/2 + 1/4 + \dots + 1/2^\tau) \leq 2d_{\max}$ and it is easy to check that (1–2) hold. Thus these claims will finish the proof.

So, condition on any outcomes of $\mathfrak{G}(1), \dots, \mathfrak{G}(t)$, and consider $\mathfrak{G}(t+1)$ as defined above. Let E_v be the set of edges containing v in $\mathfrak{G}(t+1)$. First, we have

$$\mathbb{E}|E_v| = \frac{w_v \left(\sum_{S \in \binom{V(\mathfrak{G}) \setminus \{v\}}{k-1}} \prod_{u \in S} w_u \right)}{W} \leq w_v \leq 2F(t).$$

Noting that $F(t) > (\log n)^2$ (since (A) did not occur), it follows from a Chernoff bound (Lemma 3.5) that with probability $1 - n^{-\omega(1)}$, the maximum degree of $\mathfrak{G}(t+1)$ is at most $4F(t)$.

Note that only a $\binom{n-1}{k-2}/\binom{n-1}{k-1} \leq k/n$ fraction of size- $(k-1)$ subsets of $V(\mathfrak{G})$ contain v , so the above expression for $\mathbb{E}|E_v|$ also implies that $\mathbb{E}|E_v| \geq w_v(1 - 2^k(k/n))$. Since $\mathfrak{G}(1) \cup \dots \cup \mathfrak{G}(t) \cup \mathfrak{H}$ has maximum degree at most $8d_{\max} + \binom{n-1}{k-1}/(36 \cdot 2^k) \leq \binom{n-1}{k-1}/(4 \cdot 2^k)$, we additionally compute

$$\mathbb{E}|E_v \cap (\mathfrak{G}(1) \cup \dots \cup \mathfrak{G}(t) \cup \mathfrak{H})| \leq \frac{\binom{n-1}{k-1}}{4 \cdot 2^k} \cdot \frac{w_v 2^{k-1}}{\binom{n-1}{k-1}} \leq w_v/8.$$

If n is sufficiently large with respect to k , it follows that $\mathbb{E}|E_v \setminus (\mathfrak{G}(1) \cup \dots \cup \mathfrak{G}(t) \cup \mathfrak{H})| \in [13w_v/16, w_v]$. By a Chernoff bound (Lemma 3.5) and the fact that $w_v \geq F(t) > (\log n)^2$, with probability $1 - n^{-\omega(1)}$ every vertex v has

$$|E_v \setminus (\mathfrak{G}(1) \cup \dots \cup \mathfrak{G}(t) \cup \mathfrak{H})| - w_v \leq w_v/4 \leq F(t)/2,$$

in which case $F(t+1) \leq F(t)/2$, as desired. \square

6. SUBSAMPLED TYPICAL GRAPHS AND PERFECT MATCHINGS

In this section we prove that graphs which are “random-like” have perfect matchings, in a certain robust sense. Our notion of being random-like is as follows.

Definition 6.1. An n -vertex graph G is (p, ξ) -typical if every vertex has $(1 \pm \xi)pn$ neighbors, and every pair of distinct vertices have $(1 \pm \xi)p^2n$ common neighbors.

Unless p decays very rapidly with n , it is straightforward to prove that a random graph $\mathbb{G}(n, p)$ is likely to be $(p, o(1))$ -typical. Thus, typical graphs resemble random graphs in a certain statistical sense.

In the case where $p \in (0, 1)$ is a constant (the “dense” case), typicality implies a slightly weaker property called *quasirandomness*, which famously has several equivalent definitions. One of these is that between any two disjoint vertex sets S, T there are about $p|S||T|$ edges (see for example [12]). We will need a quantitative version of this implication, where p is allowed to decay with n ; the proof is virtually the same as in the dense case, but we include it for completeness as we could not find this specific statement in the literature.

Lemma 6.2. Let $G = (V, E)$ be an n -vertex (p, ξ) -typical graph, for $p, \xi \in (0, 1)$. For every pair of disjoint vertex sets $S, T \subseteq V$, the number of edges $e(S, T)$ between S and T satisfies

$$|e(S, T) - p|S||T|| \leq 2(\xi^{1/2}pn + \sqrt{pn})\sqrt{|S||T|}.$$

Proof. We write $\text{codeg}(v, w)$ for the number of common neighbors of v and w , and we write $\text{deg}_T(s)$ for the number of neighbors of a vertex s in a set T . Note that

$$\begin{aligned} |e(S, T) - p|S||T||^2 &= \left| \sum_{s \in S} (\text{deg}_T(s) - p|T|) \right|^2 \leq |S| \sum_{s \in S} (\text{deg}_T(s) - p|T|)^2 \leq |S| \sum_{s \in V} (\text{deg}_T(s) - p|T|)^2 \\ &= |S| \sum_{s \in V} \text{deg}_T(s)^2 - 2p|T||S| \sum_{s \in V} \text{deg}_T(s) + p^2n|T|^2|S| \\ &= |S| \sum_{t_1, t_2 \in T} \text{codeg}(t_1, t_2) - 2p|T||S| \sum_{t \in T} \text{deg}(t) + p^2n|T|^2|S| \\ &\leq |S|(|T|(1 + \xi)pn + |T|(|T| - 1)(1 + \xi)p^2n) - 2p|T|^2|S|(1 - \xi)pn + p^2n|T|^2|S| \\ &\leq 2|S||T|pn + 3\xi|T|^2|S|p^2n \leq 2|S||T|pn + 3\xi|T||S|p^2n^2, \end{aligned}$$

and thus

$$|e(S, T) - p|S||T|| \leq 2(\xi^{1/2}pn + \sqrt{pn})\sqrt{|S||T|}. \quad \square$$

We next give a convenient Hall-type criterion for a bipartite graph to have a perfect matching. It is an immediate consequence of the main theorem in [56].

Lemma 6.3. *Let $G = (X \cup Y, E)$ be a bipartite graph with $|X| = |Y| = n$. Suppose that for every $S \subseteq X$, $S' \subseteq Y$ with $|S'| < |S| \leq \lceil n/2 \rceil$ we have $e(S, Y \setminus S') \neq 0$, and that for every $T' \subseteq X$, $T \subseteq Y$ with $|T'| < |T| \leq \lceil n/2 \rceil$ we have $e(T, X \setminus T') \neq 0$. Then G has a perfect matching.*

We now prove that with high probability, a random subgraph of a typical graph is “robustly matchable”, in the sense that it is not possible to destroy all perfect matchings by deleting a subgraph with small maximum degree. We accomplish this by applying Lemma 6.3 with a random bipartition of our random subgraph.

Lemma 6.4. *There is an absolute constant $\xi = \xi_{6.4} > 0$ such that the following holds. Fix $\gamma \in (0, 1/2)$, and let n be an even number which is sufficiently large in terms of γ . Let G be an n -vertex graph which is (p, ξ) -typical with $p \in [n^{-1/3}, 1]$.*

Now, let R be a random subgraph of G obtained by keeping each edge with probability $n^\gamma/(pn)$ independently. With probability at least $1 - \exp(-\Omega(n^\gamma))$, this random subgraph R has the property that for any subgraph $F \subseteq R$ with maximum degree at most ξn^γ , there is a perfect matching in $R \setminus F$.

Proof. Independently from the randomness of R , consider a uniformly random bipartition $X \cup Y$ of the vertex set of G , into two parts of size exactly $n/2$. Let G' be the bipartite subgraph of G consisting of edges between X and Y . Every vertex v has degree $(1 \pm \xi)pn$ in G , so by the Chernoff bound for hypergeometric random variables (Lemma 3.5) and a union bound, with probability $1 - n \exp(-\Omega(\xi pn))$ we have $\deg_{G'}(v) = (1 \pm 2\xi)p(n/2)$ for each vertex v . Now consider some $S \subseteq X, S' \subseteq Y$ satisfying $\lceil n/4 \rceil \geq |S| > |S'|$. By Lemma 6.2, we have

$$e_{G'}(S, Y \setminus S') = \sum_{v \in S} \deg_{G'}(v) - e_G(S, S') \geq (1 - 2\xi)p(n/2)|S| - p|S||S'| - 4\xi^{1/2}pn\sqrt{|S||S'|} \geq pn|S|/10$$

for large n and small ξ . Similarly, if $T' \subseteq X, T \subseteq Y$ with $|T'| < |T| \leq \lceil n/4 \rceil$ then $e_{G'}(T, X \setminus T') \geq pn|T|/10$.

Now, fix any outcome of G' satisfying the above properties, and let $R' = G' \cap R$. We observe that with probability at least $1 - \exp(-\Omega(n^\gamma))$, we have $e_{R'}(S, Y \setminus S') \geq n^\gamma|S|/20$ for every pair of sets $S \subseteq X, S' \subseteq Y$ with $\lceil n/4 \rceil \geq |S| > |S'|$. Indeed, by a union bound over S, S' and the Chernoff bound, the probability that this fails to hold is at most

$$\sum_{k=1}^{\lceil n/4 \rceil} \binom{n/2}{k} \sum_{\ell=0}^{k-1} \binom{n/2}{\ell} \exp(-\Omega(kn^\gamma)) \leq \sum_{k=1}^{\lceil n/4 \rceil} n^{2k} \exp(-\Omega(kn^\gamma)) \leq \exp(-\Omega(n^\gamma)).$$

Assuming the above event holds, to make $e_{R'}(S, Y \setminus S') = 0$ we need to delete at least $n^\gamma/20$ edges incident to some vertex of S . By symmetry, the same is true (with probability at least $1 - \exp(-\Omega(n^\gamma))$) when switching the roles of X and Y . The desired result then follows from Lemma 6.3, as long as $\xi < 1/20$. \square

7. WELL-SPREAD FORBIDDEN CONFIGURATIONS

For our generalized version of the high-girth triple process, we will need to make certain statistical “well-spreadness” assumptions about our family of forbidden configurations, which mimic some basic statistical properties of the family of Erdős configurations. We will have to define these assumptions in a sufficiently loose way that “sparse enough” random sets of triangles are likely to satisfy them, since such sets will arise during the regularization step of our proof (recall the outline in Section 2). In particular, we cannot make assumptions that rely too closely on the specific way that triangles are arranged with respect to each other in every forbidden configuration (whereas such information is used rather crucially in [8, 27]).

For example, note that in the complete graph K_n , the number of Erdős j -configurations containing a given set \mathcal{R} of triangles is $O(n^{j-v(\mathcal{R})})$, where $v(\mathcal{R})$ is the number of vertices spanned by the triangles in \mathcal{R} . We will make an assumption of this type, for general forbidden configurations with $j - 2$ triangles, without demanding that each of the configurations have exactly j vertices. This type of assumption is probably the most important for us, but we will also need to make further assumptions about *pairs* of forbidden configurations. A subtle property of Erdős j -configurations is that their vertex set is determined by any set of $j - 3$ triangles (equivalently, deleting a triangle does not delete any vertices), by Lemma 3.2(1). This has implications for the way that certain pairs of Erdős configurations are distributed. For example, if we fix a pair of distinct triangles T, T' , then the number of Erdős j -configurations containing T is $\Theta(n^{j-3})$, but the number of pairs of Erdős j -configurations $\mathcal{E}, \mathcal{E}'$ with $T \in \mathcal{E}, T' \in \mathcal{E}'$ and $\mathcal{E} \setminus \{T\} = \mathcal{E}' \setminus \{T'\}$ is only $O(n^{j-4})$ (because both \mathcal{E} and \mathcal{E}' must have the same vertex set containing all $v(T \cup T') \geq 4$ vertices of T and T').

It turns out that assumptions of the above type (concerning the numbers of Erdős configurations and pairs of Erdős configurations containing given sets of triangles) are sufficient for a basic generalization of

the theorems in [8, 27]. However, we will actually need rather more complicated assumptions: we need a general theorem which we can apply iteratively, “moving down” a vortex $U_0 \supseteq U_1 \supseteq \dots \supseteq U_\ell$ (recall the discussion in Section 2.1). We therefore need our family of forbidden configurations to be well-spread, in a somewhat technical way, with respect to our vortex.

Definition 7.1 (Well-spread). First, for any nonempty set of at most $j - 2$ triangles \mathcal{R} , define

$$v^j(\mathcal{R}) = \begin{cases} |\mathcal{R}| + 2 & \text{if } |\mathcal{R}| \in \{1, j - 2\} \\ |\mathcal{R}| + 3 & \text{if } 1 < |\mathcal{R}| < j - 2. \end{cases}$$

The idea is that if \mathcal{R} is a nonempty subset of an Erdős j -configuration, then it has at least $v^j(\mathcal{R})$ vertices (this follows from Lemma 3.2 applied to \mathcal{R}).

Now, fix a descending sequence of subsets $V(K_N) = U_0 \supseteq \dots \supseteq U_k$, and let $n = |U_k|$. Fix some $y, z \in \mathbb{R}$ (which should be thought of as “error parameters”). For a triple T in K_N , let $\text{lev}(T)$ be the maximum i such that T has all its vertices in U_i . For $4 \leq j \leq g$, say that a collection of sets of $j - 2$ triangles \mathfrak{F}_j is (y, z) -well-spread (with respect to our descending sequence of subsets) if the following conditions hold.

WS0 Every $\mathcal{E} \in \mathfrak{F}_j$ is an edge-disjoint collection of triangles.

WS1 For every nonempty set of triangles \mathcal{R} and any sequence t_0, \dots, t_{k-1} , the number of $\mathcal{E} \in \mathfrak{F}_j$ which include \mathcal{R} and, for each $0 \leq i < k$, contain t_i triangles $T \notin \mathcal{R}$ with $\text{lev}(T) = i$ is at most

$$zn^{j-(t_0+\dots+t_{k-1})-v^j(\mathcal{R})} \prod_{0 \leq i < k} |U_i|^{t_i}.$$

WS2 Fix a pair of triangles T, T' and any sequence t_0, \dots, t_{k-1} . Consider all the pairs of distinct $\mathcal{E}, \mathcal{E}' \in \mathfrak{F}_j$ for which $T \in \mathcal{E}$, $T' \in \mathcal{E}'$ and $\mathcal{E} \setminus \{T\} = \mathcal{E}' \setminus \{T'\}$. The number of such pairs $\mathcal{E}, \mathcal{E}'$ which contain t_i triangles $T'' \in \mathcal{E} \setminus \{T\}$ with $\text{lev}(T'') = i$, for $0 \leq i < k$, is at most

$$zn^{j-(t_0+\dots+t_{k-1})-4} \prod_{0 \leq i < k} |U_i|^{t_i}.$$

WS3 If $j = 4$, then for every triangle T and edge $e \notin E(T)$, there are at most z configurations $\mathcal{E} \in \mathfrak{F}_j$ which contain T and a second triangle T' with $\text{lev}(T') = k$ and $e \in E(T')$.

WS4 For every triangle T and any sequence t_0, \dots, t_{k-1} , the number of $\mathcal{E} \in \mathfrak{F}_j$ which contain T and for each $0 \leq i < k$ contain t_i triangles $T' \neq T$ with $\text{lev}(T') = i$ is at most

$$yn^{j-(t_0+\dots+t_{k-1})-3} \prod_{0 \leq i < k} |U_i|^{t_i}.$$

Finally, we say that a collection of sets of $j - 2$ triangles \mathfrak{F}_j is (z, z) -well-spread if it is (z, z) -well-spread.

Remark. We make a few important comments regarding the above definition. First, note that we allow $j \in \{4, 5\}$, despite the fact that Erdős 4-configurations and Erdős 5-configurations do not exist. Second, note that **WS4** is simply **WS1** for $|\mathcal{R}| = 1$, but with an adjusted error parameter. In certain applications, it will be crucial that $y = o(z)$. Third, note that being well-spread for $U_0 \supseteq \dots \supseteq U_{k+1}$ is not strictly stronger than being well-spread for $U_0 \supseteq \dots \supseteq U_k$ (since the definition of n changes); similarly, the definition of $\text{lev}(T)$ depends on our sequence of sets. These are not serious issues, though we must use care when applying the definitions and results in this section. Finally, we note that the definition of being well-spread still makes sense when $k = 0$ (in which case the sequences t_0, \dots, t_{k-1} are empty).

It is not too difficult to prove that for any descending sequence of subsets $V(K_N) = U_0 \supseteq \dots \supseteq U_k$, the collection of Erdős j -configurations in K_N is $O_{j,k}(1)$ -well-spread. However, we will need a version of this fact that incorporates the influence of the absorbing structure defined in Section 4; we are not only interested in Erdős configurations, but more generally in sets of triangles which would complete an Erdős configuration when combined with some absorber triangles. Furthermore, we will need to account for the additional (random) forbidden configurations which may be added for purposes of regularization. We first tackle the impact of the absorbing structure, which is the more difficult of these considerations.

Lemma 7.2. *Let $k \geq 0$, consider a descending sequence of subsets $V(K_N) = U_0 \supseteq \dots \supseteq U_k \supseteq X$, and let $n = |U_k|$. Consider an absorbing structure H (with distinguished set X) satisfying **Ab1** and **Ab2** in Theorem 4.1 for $g \in \mathbb{N}$. Suppose H is embedded in a complete graph K_N , and assume that (if $k > 0$) we have $V(H) \setminus X \subseteq U_0 \setminus U_1$.*

*For any triangle-divisible graph L on the vertex set X let \mathcal{S}_L be the associated triangle-decomposition of $L \cup H$ (as in **Ab1**), and let $\mathcal{B} = \bigcup_L \mathcal{S}_L$. Suppose that $|\mathcal{B}|^{2g} \leq N^\beta$, for some $\beta \leq 1$. For $j \geq 4$, let $\mathfrak{F}_j^\mathcal{B}$ be the collection of all $(j - 2)$ -sets of triangles such that one can obtain an Erdős j' -configuration for some $5 \leq j' \leq g$ by adding $j' - j$ triangles from \mathcal{B} . Then $\mathfrak{F}_j^\mathcal{B}$ is $(C_{7.2}(g, k), C_{7.2}(g, k)n^\beta)$ -well-spread with respect to $U_0 \supseteq \dots \supseteq U_k$.*

Proof. The configurations in $\mathfrak{F}_j^{\mathcal{B}}$ are edge-disjoint by definition. Consider $j \geq 4$ and sets of at most $j - 2$ triangles \mathcal{R}, \mathcal{K} in K_N with $\mathcal{R} \neq \emptyset$, and consider a sequence of nonnegative integers t_0, \dots, t_{k-1} . Let $\#(\mathcal{R}, \mathcal{K}, j, t_0, \dots, t_{k-1})$ be the number of Erdős configurations \mathcal{E} with $\mathcal{E} \cap \mathcal{B} = \mathcal{K}$, with $\mathcal{R} \subseteq \mathcal{E} \setminus \mathcal{K}$, which have $j + |\mathcal{K}| - 2$ triangles (so $j + |\mathcal{K}|$ vertices) in total, with $|\mathcal{K}| \leq g - j$, and which contain t_i triangles $T \notin \mathcal{K} \cup \mathcal{R}$ with $\text{lev}(T) = i$. Let

$$\#(\mathcal{R}, j, t_0, \dots, t_{k-1}) = \sum_{\substack{\mathcal{K} \subseteq \mathcal{B} \\ |\mathcal{K}| \leq g-j}} \#(\mathcal{R}, \mathcal{K}, j, t_0, \dots, t_{k-1}),$$

which is an upper bound for the quantity of interest in **WS1** for $\mathfrak{F}_j^{\mathcal{B}}$, and for **WS4** in the cases with $|\mathcal{R}| = 1$.

We start by proving an upper bound on $\#(\mathcal{R}, \mathcal{K}, j, t_0, \dots, t_{k-1})$. Suppose \mathcal{E} contributes to this quantity. For $0 \leq i \leq k$, let $v_i = |(V(\mathcal{E}) \setminus V(\mathcal{K} \cup \mathcal{R})) \cap (U_i \setminus U_{i+1})|$ be the number of vertices in \mathcal{E} but not in $\mathcal{K} \cup \mathcal{R}$ which lie in $U_i \setminus U_{i+1}$ (here we take U_{k+1} to be the empty set). Recalling the definition of v^j from **Definition 7.1**, note that

$$v_0 + \dots + v_k = j + |\mathcal{K}| - v(\mathcal{K} \cup \mathcal{R}) \leq j - v^j(\mathcal{R}). \quad (7.1)$$

To see why the inequality holds, note that if $|\mathcal{R}| \in \{1, j - 2\}$, the inequality is equivalent to $v(\mathcal{K} \cup \mathcal{R}) \geq |\mathcal{K} \cup \mathcal{R}| + 2$, which holds by **Lemma 3.2(2)**. Otherwise, if $1 < |\mathcal{R}| < j - 2$, the inequality is equivalent to $v(\mathcal{K} \cup \mathcal{R}) \geq |\mathcal{K} \cup \mathcal{R}| + 3$. This holds by **Lemma 3.2(1)**, since $1 < |\mathcal{K} \cup \mathcal{R}| < j + |\mathcal{K}| - 2$.

Now, since $|V(\mathcal{R})| \geq 3$, we may apply **Lemma 3.2(2)** to the vertex set $(V(\mathcal{E}) \setminus V(\mathcal{K} \cup \mathcal{R})) \cap (U_0 \setminus U_{i+1})$ to obtain $v_0 + \dots + v_i \leq t_0 + \dots + t_i$ for $0 \leq i < k$. Therefore, the contribution to $\#(\mathcal{R}, \mathcal{K}, j, t_0, \dots, t_{k-1})$ from a particular choice of $(v_i)_{0 \leq i < k}$ is at most

$$\begin{aligned} \prod_{0 \leq i \leq k} |U_i|^{v_i} &= |U_k|^{v_0 + \dots + v_k} \prod_{0 \leq i < k} \left(\frac{|U_i|}{|U_{i+1}|} \right)^{v_0 + \dots + v_i} \leq n^{j - v^j(\mathcal{R})} \prod_{0 \leq i < k} \left(\frac{|U_i|}{|U_{i+1}|} \right)^{t_0 + \dots + t_i} \\ &= n^{j - v^j(\mathcal{R}) - (t_0 + \dots + t_{k-1})} \prod_{0 \leq i < k} |U_i|^{t_i}. \end{aligned} \quad (7.2)$$

Note that there are only $O_{g,k}(1)$ possibilities for $(v_i)_{0 \leq i < k}$, so

$$\#(\mathcal{R}, \mathcal{K}, j, t_0, \dots, t_{k-1}) \leq O_{g,k}(1) \cdot n^{j - v^j(\mathcal{R}) - (t_0 + \dots + t_{k-1})} \prod_{0 \leq i < k} |U_i|^{t_i}. \quad (7.3)$$

We will now distinguish some cases to prove that

$$\#(\mathcal{R}, j, t_0, \dots, t_{k-1}) \leq O_{g,k}(1) \cdot n^{\beta + j - v^j(\mathcal{R}) - (t_0 + \dots + t_{k-1})} \prod_{0 \leq i < k} |U_i|^{t_i}, \quad (7.4)$$

which will show that $\mathfrak{F}_j^{\mathcal{B}}$ satisfies **WS1** with parameter $z = O_{g,k}(n^\beta)$.

- *Case 1:* $n \geq N^{1/2}$. There are at most $|\mathcal{B}|^g \leq N^{\beta/2} \leq n^\beta$ ways to choose \mathcal{K} , so we can simply sum (7.3) over all such choices to obtain (7.4). Note this covers all cases when $k = 0$.
- *Case 2:* $n < N^{1/2}$ and $t_0 = 0$. We claim there are actually only $O_g(1)$ choices for \mathcal{K} in this case, so we can again sum over all choices. Recall **Ab2**. Either $\mathcal{K} \subseteq \mathcal{L}_{\mathcal{R}}$, which provides $O_g(1)$ choices for \mathcal{K} , or there must be $T \in \mathcal{E} \setminus (\mathcal{R} \cup \mathcal{B})$ such that T contains a vertex of $V(H) \setminus X$. In the latter case, note that $\mathcal{K} \subseteq \mathcal{B}$ and $V(H) \setminus X \subseteq U_0 \setminus U_1$. Then $t_0 > 0$ follows, which is a contradiction.
- *Case 3:* $n < N^{1/2}$ and $t_0 > 0$. Configurations with $\mathcal{K} = \emptyset$ are bounded by (7.3). Otherwise, we obtain a strict inequality $v_0 + \dots + v_i < t_0 + \dots + t_i$ for all $0 \leq i < k$. Indeed, if $|(V(\mathcal{E}) \setminus V(\mathcal{K} \cup \mathcal{R})) \cap (U_0 \setminus U_{i+1})| = v_0 + \dots + v_i = 0$ then this is trivial since $t_0 > 0$. Otherwise apply **Lemma 3.2(1)** to this nonempty vertex set, noting that $\mathcal{K} \cup \mathcal{R}$ contains at least 4 vertices total. There are at most $|\mathcal{B}|^g$ ways to choose \mathcal{K} , and for each fixed \mathcal{K} and $(v_i)_{0 \leq i < k}$, the number of choices for \mathcal{E} is at most

$$\prod_{0 \leq i \leq k} |U_i|^{v_i} \leq (n/N) n^{j - v^j(\mathcal{R}) - (t_0 + \dots + t_{k-1})} \prod_{0 \leq i < k} |U_i|^{t_i},$$

mimicking the proof of (7.2). Since $|\mathcal{B}|^g \leq N^{1/2} \leq N/n$, the desired bound follows.

We have now established (7.4) in all cases, proving that $\mathfrak{F}_j^{\mathcal{B}}$ satisfies **WS1**. Next, we observe that in the case $|\mathcal{R}| = 1$ and $\mathcal{K} \neq \emptyset$, the inequality in (7.1) is strict (due to **Lemma 3.2(1)** and $1 < |\mathcal{K} \cup \mathcal{R}| < j + |\mathcal{K}| - 2$). So, in this case, the proof of (7.4) can be seen to save a factor of n , and actually shows

$$\#(\mathcal{R}, \mathcal{K}, j, t_0, \dots, t_{k-1}) \leq O_{g,k}(1) \cdot n^{\beta + j - 4 - (t_0 + \dots + t_{k-1})} \prod_{0 \leq i < k} |U_i|^{t_i}. \quad (7.5)$$

Then, (7.5) (for the case $\mathcal{K} \neq \emptyset$) and (7.3) (for the case $\mathcal{K} = \emptyset$) together imply

$$\#(\mathcal{R}, j, t_0, \dots, t_{k-1}) \leq O_{g,k}(1) \cdot n^{j-3 - (t_0 + \dots + t_{k-1})} \prod_{0 \leq i < k} |U_i|^{t_i},$$

for $|\mathcal{R}| = 1$, which yields **WS4** of $(C_{7.2}(g, k), C_{7.2}(g, k)n^\beta)$ -well-spreadness. Also, (7.5) proves **WS3** (with room to spare; we do not need to consider the edge e at all).

It remains to prove property **WS2** with the given parameters. Fix T, T' . We are counting the number of pairs $\mathcal{E}, \mathcal{E}' \in \mathfrak{F}_j^\mathcal{B}$ satisfying a certain property, and such that there are t_i triangles $T'' \in \mathcal{E} \setminus \{T\}$ with $\text{lev}(T'') = i$ for $0 \leq i < k$. First, the contribution coming from pairs in which \mathcal{E} is a “derived” forbidden configurations (i.e., not a complete Erdős configuration, induced by some nonempty subset $\mathcal{K} \subseteq \mathcal{B}$) is bounded as desired by (7.5), taking $\mathcal{R} = \{T\}$. A symmetric argument for \mathcal{E}' shows we only need to consider the contribution coming from the case where $\mathcal{E}, \mathcal{E}'$ are both genuine Erdős configurations. In this case $\mathcal{E}' \setminus \{T'\}$ has the same vertex set as \mathcal{E}' , so \mathcal{E}' has to contain all $v(T \cup T') \geq 4$ vertices of T and T' . To count the number of such configurations let $v_i = |(V(\mathcal{E}) \setminus V(T \cup T')) \cap (U_i \setminus U_{i+1})|$ for $0 \leq i \leq k$ (once again taking $U_{k+1} = \emptyset$), and note that

$$v_0 + \cdots + v_k \leq j - 4.$$

With this modified definition of v_i , we still have $v_0 + \cdots + v_i \leq t_0 + \cdots + t_i$ for $0 \leq i < k$ by Lemma 3.2(2) and the number of choices for \mathcal{E} given a particular choice of $(v_i)_{0 \leq i < k}$ is at most

$$\prod_{i \leq k} |U_i|^{v_i} = |U_k|^{v_0 + \cdots + v_i} \prod_{i < k} \left(\frac{|U_i|}{|U_{i+1}|} \right)^{v_0 + \cdots + v_i} \leq n^{j-4} \prod_{i < k} \left(\frac{|U_i|}{|U_{i+1}|} \right)^{t_0 + \cdots + t_i} = n^{j-4 - (t_0 + \cdots + t_{k-1})} \prod_{i < k} |U_i|^{t_i}.$$

Summing over $O_{g,k}(1)$ sequences $(v_i)_{0 \leq i < k}$ and $O_g(1)$ ways to choose \mathcal{E}' from triangles on $V(\mathcal{E})$ completes the proof. \square

Finally, the following lemma says that if one adds appropriate *random* configurations to a well-spread collection of configurations, the result is still well-spread.

Lemma 7.3. *Fix $\beta \in (0, 1/3)$. Let \mathfrak{F}_j be an n^β -well-spread collection of sets of $j - 2$ triangles with respect to $V(K_N) = U_0 \supseteq \cdots \supseteq U_k$. Let $\mathfrak{F}_j^{\text{rand}}$ be a random collection of sets of vertex-disjoint triangles in $K_N[U_k] \cong K_n$, where every set of $j - 2$ triangles is included in $\mathfrak{F}_j^{\text{rand}}$ with probability $p_j = n^\beta n^{-2j+6}$. With probability $1 - n^{-\omega(1)}$, the collection of sets of triangles $\mathfrak{F}_j \cup \mathfrak{F}_j^{\text{rand}}$ is $C_{7.3}(g, k)n^\beta$ -well-spread.*

Proof. First, we claim that with probability $1 - n^{-\omega(1)}$, $\mathfrak{F}_j^{\text{rand}}$ satisfies property **WS1** of $C_{7.3}(g, k)n^\beta$ -well-spreadness. Indeed, note that no configuration in $\mathfrak{F}_j^{\text{rand}}$ ever contains a triangle T with $\text{lev}(T) < k$. Fixing nonempty \mathcal{R} , there are at most $(n^3)^{j-2-|\mathcal{R}|}$ possible configurations including \mathcal{R} , and each is present independently with probability p_j . The expected number of such configurations in $\mathfrak{F}_j^{\text{rand}}$ is at most $n^{\beta+j-3|\mathcal{R}|} \leq n^{\beta+j-v^j(\mathcal{R})}$, so a Chernoff bound proves the claim. Note in this situation, **WS4** is a special case of **WS1**.

Next, we claim that with probability $1 - n^{-\omega(1)}$, if $j = 4$ then $\mathfrak{F}_j^{\text{rand}}$ satisfies property **WS3**. Indeed, for a triangle T and an edge $e \notin E(T)$, there are at most n triangles T' with $\text{lev}(T') = k$ containing e . So, the expected number of configurations in $\mathfrak{F}_4^{\text{rand}}$ containing T and such a triangle T' is at most $np_4 \leq n^{\beta-1}$, implying (via a Chernoff bound) that property **WS3** holds for $\mathfrak{F}_4^{\text{rand}}$ with probability $1 - n^{-\omega(1)}$.

For property **WS2**, we need to consider separately the case where $\mathcal{E}, \mathcal{E}' \in \mathfrak{F}_j^{\text{rand}}$ and the case where $\mathcal{E} \in \mathfrak{F}_j^{\text{rand}}, \mathcal{E}' \in \mathfrak{F}_j$. In both cases we can again use a simple Chernoff bound. The expected number of suitable pairs in the former case is at most $(n^3)^{j-3} p_j^2 \leq n^{2\beta-j+3} = o(1)$, choosing $j - 3$ disjoint triangles for $\mathcal{E} \setminus \{T\}$. Then, the expected number of suitable pairs in the latter case is at most $O_g(n^{\beta+j-3} p_j) \leq n^{2\beta-j+3} = o(1)$. Here we are using that $\mathcal{E}' \setminus \{T'\} = \mathcal{E} \setminus \{T\}$ must have all vertices in U_k , so **WS4** applied to \mathfrak{F}_j implies there are at most $n^{\beta+j-3}$ choices for \mathcal{E}' . The result follows. \square

8. BOUNDS FOR WEIGHT SYSTEMS

At many points in the proof of Theorem 1.1, we will prove upper bounds on random variables by defining a weight system $\vec{\pi}$ and a set of configurations \mathfrak{X} , and applying Lemma 3.7. In order to apply Lemma 3.7 it is necessary to estimate the maximum weight $\kappa(\vec{\pi})(\mathfrak{X})$, and in many of these applications the case analysis involved in this estimation is quite involved (though mostly quite mechanical). We collect all such computations in this section: in particular, we prove a number of abstract lemmas providing upper bounds on maximum weights in various situations, tailored for various applications in the proof of Theorem 1.1. The lemmas in this section are likely to seem unmotivated on first reading; the reader is encouraged to skip over this section and refer to it only as various estimates are needed in later sections.

First, for our most straightforward applications of Lemma 3.7, our ground set \mathcal{W} is a set of triangles in a graph, and our weight system $\vec{\pi}$ encodes the approximate probability that a given triangle T ends up being chosen in one of the (random) stages of the proof of Theorem 1.1. This probability depends on the position of T with respect to the vortex $U_0 \supseteq \cdots \supseteq U_\ell$ (recall the outline in Section 2). The multiset of configurations with which we will apply Lemma 3.7 will be defined in terms of a well-spread collection of

forbidden configurations (for example, we might be interested in a high-probability upper bound on the number of forbidden configurations containing a given triangle T , of which a certain number of triangles have so far been chosen, in which case we would apply [Lemma 3.7](#) to a multiset of configurations obtained by taking appropriate-size subsets of forbidden configurations containing T). In the following lemma we collect some general weight estimates which are suitable for studying collections of configurations of this type.

Recall that for a triangle T in K_N and a vortex $V(K_N) = U_0 \supseteq \dots \supseteq U_k$ we write $\text{lev}(T)$ for the maximum i such that T has all its vertices in U_i . The functions $v^j(\cdot)$ used throughout the section were defined in [Definition 7.1](#).

Lemma 8.1. *Fix positive real numbers w, y, z , fix a descending sequence of subsets $V(K_N) = U_0 \supseteq \dots \supseteq U_k$, and let $n = |U_k|$. Let $\mathfrak{F}_4, \dots, \mathfrak{F}_g$ be collections of sets of triangles of K_N , where each configuration in \mathfrak{F}_j contains $j - 2$ triangles and each \mathfrak{F}_j is (y, z) -well-spread with respect to our sequence of sets. Let \mathcal{W} be the set of triangles in K_N , and let $\bar{\pi}$ be the weight system defined by $\pi_T = w/|U_{\text{lev}(T)}|$ for all $T \in \mathcal{W}$.*

- (1) *Fix a nonempty set of triangles \mathcal{Q} of K_N , and integers j and f such that $4 \leq j \leq g$ and $f \geq 0$. Let $\mathfrak{A}_{\mathcal{Q},j,f}^{(1)}$ be the multiset of sets of triangles constructed as follows. Consider every $\mathcal{E} \in \mathfrak{F}_j$ and every partition $\mathcal{E} = \mathcal{Q} \cup \mathcal{Z} \cup \mathcal{F}$ such that $|\mathcal{F}| = f$ and all triangles in \mathcal{Z} are contained in U_k . For each such choice of $(\mathcal{E}, \mathcal{F}, \mathcal{Z})$, we add a copy of \mathcal{F} to $\mathfrak{A}_{\mathcal{Q},j,f}^{(1)}$. Then*

$$\psi(\mathfrak{A}_{\mathcal{Q},j,f}^{(1)}) = \begin{cases} O_{g,k}(zw^f n^{j-v^j(\mathcal{Q})-f}) & \text{for all } \mathcal{Q}, \\ O_{g,k}(yw^f n^{j-v^j(\mathcal{Q})-f}) & \text{if } |\mathcal{Q}| = 1. \end{cases}$$

- (2) *Fix sets of triangles $\mathcal{Q}, \mathcal{Q}'$ of K_N with $\mathcal{Q} \neq \emptyset$, fix j, j', v', f such that $4 \leq j, j' \leq g$ and $v', f \geq 0$, and let $v = v^j(\mathcal{Q})$. Let $\mathfrak{A}_{\mathcal{Q},\mathcal{Q}',j,j',v',f}^{(2)}$ be the multiset of sets of triangles constructed as follows. Consider each $\mathcal{E} \in \mathfrak{F}_j, \mathcal{E}' \in \mathfrak{F}_{j'}$ and every choice of partitions $\mathcal{E} = \mathcal{Q} \cup \mathcal{Z} \cup \mathcal{F}$ and $\mathcal{E}' = \mathcal{Q}' \cup \mathcal{Z}' \cup \mathcal{F}'$ such that $\mathcal{F} \cup \mathcal{F}'$ is disjoint from $\mathcal{Q} \cup \mathcal{Z} \cup \mathcal{Q}' \cup \mathcal{Z}'$, $\mathcal{Q}' \cup (\mathcal{E} \cap \mathcal{E}') \neq \emptyset$, $v' = v^{j'}(\mathcal{E}' \cap (\mathcal{E} \cup \mathcal{Q}'))$, $|\mathcal{F} \cup \mathcal{F}'| = f$, and all the triangles in $\mathcal{Z} \cup \mathcal{Z}'$ are contained in U_k . (Note v' is well-defined as $\mathcal{E}' \cap (\mathcal{E} \cup \mathcal{Q}') = \mathcal{Q}' \cup (\mathcal{E} \cap \mathcal{E}')$.) For each such choice of $(\mathcal{E}, \mathcal{E}', \mathcal{F}, \mathcal{F}', \mathcal{Z}, \mathcal{Z}')$, we add a copy of $\mathcal{F} \cup \mathcal{F}'$ to $\mathfrak{A}_{\mathcal{Q},\mathcal{Q}',j,j',v',f}^{(2)}$. Then*

$$\psi(\mathfrak{A}_{\mathcal{Q},\mathcal{Q}',j,j',v',f}^{(2)}) = O_{g,k}(z^2 w^f n^{j+j'-v-v'-f}).$$

- (3) *Fix a pair of (not necessarily distinct) triangles T, T' of K_N , and fix some $4 \leq j \leq g$ and $f \geq 0$. Let $\mathfrak{A}_{T,T',j,f}^{(3)}$ be the multiset of sets of triangles constructed as follows. Consider each pair of distinct $\mathcal{E}, \mathcal{E}' \in \mathfrak{F}_j$ with $\mathcal{E} \setminus \{T\} = \mathcal{E}' \setminus \{T'\}$, and every choice of partition $\mathcal{E} \setminus \{T\} = \mathcal{Z} \cup \mathcal{F}$ such that $|\mathcal{F}| = f$ and all the triangles in \mathcal{Z} are contained in U_k . For each such choice of $(\mathcal{E}, \mathcal{E}', \mathcal{F}, \mathcal{Z})$, we add a copy of \mathcal{F} to $\mathfrak{A}_{T,T',j,f}^{(3)}$. Then*

$$\psi(\mathfrak{A}_{T,T',j,f}^{(3)}) = O_{g,k}(zw^f n^{j-4-f}).$$

Proof. First we consider (1). For some choice of $\mathcal{E}, \mathcal{F}, \mathcal{Z}$ and some $0 \leq i \leq k-1$, let t_i be the number of triangles $T \in \mathcal{F}$ with $\text{lev}(T) = i$. For any particular profile $(t_i)_{0 \leq i < k}$, the contribution to $\psi(\mathfrak{A}_{\mathcal{Q},j,f}^{(1)})$ from choices of $\mathcal{E}, \mathcal{F}, \mathcal{Z}$ with that profile is

$$O_g(1) \cdot \left(zn^{j-(t_0+\dots+t_{k-1})-v^j(\mathcal{Q})} \prod_{i < k} |U_i|^{t_i} \right) \left(w^f |U_k|^{-(f-t_0-\dots-t_{k-1})} \prod_{i < k} |U_i|^{-t_i} \right) = O_g(zw^f n^{j-v^j(\mathcal{Q})-f}).$$

To obtain this estimate, we have applied [WS1](#) (in the definition of well-spreadness, in [Definition 7.1](#)) to $\mathcal{R} = \mathcal{Q} \neq \emptyset$ (since every $T \in \mathcal{Z}$ satisfies $\text{lev}(T) = k$) in order to count choices of \mathcal{E} , with an extra factor depending only on g to account for the number of choices of \mathcal{F} . The first bound in (1) follows by summing over $O_{g,k}(1)$ possibilities for the profile $(t_i)_{0 \leq i < k}$. For the alternate bound when $|\mathcal{Q}| = 1$, we simply use [WS4](#) instead of [WS1](#).

Next, (2) is very similar, but with slightly more complicated notation. For $0 \leq i < k$, let t_i be the number of triangles $T \in \mathcal{F}$ with $\text{lev}(T) = i$, and let t'_i be the number of triangles $T \in \mathcal{F}' \setminus \mathcal{E} = \mathcal{F}' \setminus \mathcal{F}$ with $\text{lev}(T) = i$. The contribution to $\psi(\mathfrak{A}_{\mathcal{Q},\mathcal{Q}',j,j',v',f}^{(2)})$ coming from a particular choice of the profiles $(t_i)_{0 \leq i < k}, (t'_i)_{0 \leq i < k}$ is at most

$$\begin{aligned} & O_g(1) \cdot \left(zn^{j-(t_0+\dots+t_{k-1})-v} \prod_{i < k} |U_i|^{t_i} \right) \left(zn^{j'-(t'_0+\dots+t'_{k-1})-v'} \prod_{i < k} |U_i|^{t'_i} \right) \\ & \cdot \left(w^f |U_k|^{-(f-t_0-\dots-t_{k-1}-t'_0-\dots-t'_{k-1})} \prod_{i < k} |U_i|^{-(t_i+t'_i)} \right) = O_g(z^2 w^f n^{j+j'-v-v'-f}). \end{aligned}$$

Here we have used **WS1** twice to count choices for $\mathcal{E}, \mathcal{E}'$, first applied to $\mathcal{R} = \mathcal{Q} \neq \emptyset$ and then to $\mathcal{R} = \mathcal{E}' \cap (\mathcal{E} \cup \mathcal{Q}') = \mathcal{Q}' \cup (\mathcal{E} \cap \mathcal{E}') \neq \emptyset$ (there are $O_g(1)$ choices for such an \mathcal{R}). The desired result follows upon summing over possibilities for the profiles $(t_i)_{0 \leq i < k}$ and $(t'_i)_{0 \leq i < k}$.

Finally, (3) is essentially identical to (1). For $0 \leq i < k$, let t_i be the number of triangles $T'' \in \mathcal{F}$ with $\text{lev}(T'') = i$. The contribution to $\psi(\mathfrak{A}_{T, T', j, f}^{(3)})$ coming from a particular choice of the profile $(t_i)_{0 \leq i < k}$ is at most

$$O_g(1) \cdot \left(zn^{j-(t_0+\dots+t_{k-1})-4} \prod_{i < k} |U_i|^{t_i} \right) \left(w^f n^{-(f-t_0-\dots-t_{k-1})} \prod_{i < k} |U_i|^{-t_i} \right) = O_g(zw^f n^{j-4-f}),$$

where we have used **WS2**. As before the desired bound is obtained by summing over choices of $(t_i)_{0 \leq i < k}$. \square

Using the general estimates in [Lemma 8.1](#), we now estimate some maximum weights $\kappa(\mathfrak{R}) = \max_{\mathcal{H}} \psi(\mathfrak{R}, \mathcal{H})$ (recall the notation introduced [Definition 3.6](#)), in certain rather specific settings.

Each of the estimates in the following lemma corresponds to a specific quantity that we will need to control when we analyze a generalized high-girth triple process in [Section 9](#).

Lemma 8.2. *Fix positive real numbers $w, z \geq 1$, fix a descending sequence of subsets $V(K_N) = U_0 \supseteq \dots \supseteq U_k$, and set $n = |U_k|$. Let $\mathfrak{F}_4, \dots, \mathfrak{F}_g$ be collections of sets of triangles of K_N , where each set in \mathfrak{F}_j contains $j-2$ triangles and each \mathfrak{F}_j is z -well-spread with respect to our sequence of sets. Let \mathcal{W} be the set of triangles in K_N , and let $\vec{\pi}$ be the weight system defined by $\pi_T = w/|U_{\text{lev}(T)}|$ for all $T \in \mathcal{W}$.*

- (1) *For $4 \leq j \leq g$, $0 \leq c \leq j-5$, and distinct triangles T, T' in K_N , let $\mathfrak{R}_{T, T', j, c}^{(1)}$ be the multiset of sets of triangles constructed as follows. Consider each $\mathcal{S} \in \mathfrak{F}_j$ such that $T, T' \in \mathcal{S}$, and consider each subset $\mathcal{O} \subseteq \mathcal{S} \setminus \{T, T'\}$ containing exactly $j-c-4$ triangles, all of which are within U_k . For each such $(\mathcal{S}, \mathcal{O})$, add a copy of $\mathcal{S} \setminus (\mathcal{O} \cup \{T, T'\})$ to $\mathfrak{R}_{T, T', j, c}^{(1)}$. Then $\kappa(\mathfrak{R}_{T, T', j, c}^{(1)}) = O_{g,k}(zw^g n^{j-c-5})$.*
- (2) *For an edge e and triangle T in K_N with $e \not\subseteq T$, let $\mathfrak{R}_{e, T}^{(2)}$ be the multiset of sets of triangles constructed as follows. Consider each triangle $T' \neq T$, fully within U_k and containing the edge e , and consider each $\mathcal{S} \in \bigcup_{j=4}^g \mathfrak{F}_j$ such that $T, T' \in \mathcal{S}$. For each such (T', \mathcal{S}) , add a copy of $\mathcal{S} \setminus \{T, T'\}$ to $\mathfrak{R}_{e, T}^{(2)}$. Then $\kappa(\mathfrak{R}_{e, T}^{(2)}) = O_{g,k}(zw^g)$.*
- (3) *For (not necessarily distinct) triangles T, T' in K_N , let $\mathfrak{R}_{T, T'}^{(3)}$ be the multiset of sets of triangles constructed as follows. Consider each triangle $T^* \notin \{T, T'\}$ fully within U_k , and consider each pair of distinct configurations $\mathcal{S}, \mathcal{S}' \in \bigcup_{j=4}^g \mathfrak{F}_j$ such that $\{T, T', T^*\} \cap \mathcal{S} = \{T, T^*\}$ and $\{T, T', T^*\} \cap \mathcal{S}' = \{T', T^*\}$. For each such $(T^*, \mathcal{S}, \mathcal{S}')$, add a copy of $(\mathcal{S} \cup \mathcal{S}') \setminus \{T, T', T^*\}$ to $\mathfrak{R}_{T, T'}^{(3)}$. Then $\kappa(\mathfrak{R}_{T, T'}^{(3)}) = O_{g,k}(z^2 w^{2g})$.*
- (4) *For a triangle T in $K_N[U_k]$, $4 \leq j \leq g$, and $1 \leq c \leq j-4$, let $\mathfrak{R}_{T, j, c-1}^{(4)}$ be the multiset of sets of triangles constructed as follows. Consider each $\mathcal{S} \in \mathfrak{F}_j$ containing T , and consider each subset $\mathcal{O} \subseteq \mathcal{S} \setminus \{T\}$ containing exactly $j-c-2$ triangles, all of which are within U_k . Then, consider each $\mathcal{S}' \in \bigcup_{j'=5}^g \mathfrak{F}_{j'}$ such that $|\mathcal{S}' \cap (\mathcal{O} \cup \{T\})| = 2$ and $\mathcal{S}' \not\subseteq \mathcal{S}$. For each such $(\mathcal{S}, \mathcal{S}', \mathcal{O})$, add a copy of $(\mathcal{S} \cup \mathcal{S}') \setminus (\mathcal{O} \cup \{T\})$ to $\mathfrak{R}_{T, j, c-1}^{(4)}$. Then $\kappa(\mathfrak{R}_{T, j, c-1}^{(4)}) = O_{g,k}(z^2 w^{2g} n^{j-c-3})$.*

Proof of [Lemma 8.2\(1\)](#). Fix any set of triangles \mathcal{H} . We wish to prove an upper bound on the weight $\psi(\mathfrak{R}_{T, T', j, c}^{(1)}, \mathcal{H})$. Recall the representation of each $\mathcal{K} \in \mathfrak{R}_{T, T', j, c}^{(1)}$ in the form $\mathcal{K} = \mathcal{S} \setminus (\mathcal{O} \cup \{T, T'\})$, and note that we can write $\psi(\mathfrak{R}_{T, T', j, c}^{(1)}, \mathcal{H}) = \psi(\mathfrak{R}')$, where $\mathfrak{R}' = \{\mathcal{K} \setminus \mathcal{H} : \mathcal{K} \in \mathfrak{R}_{T, T', j, c}^{(1)}, \mathcal{K} \supseteq \mathcal{H}\}$.

Let $\mathcal{Q} = \mathcal{H} \cup \{T, T'\}$ and recall the notation in [Lemma 8.1\(1\)](#). Observe that we have the multiset inclusion $\mathfrak{R}' \subseteq \mathfrak{A}_{\mathcal{Q}, j, c-|\mathcal{H}|}^{(1)}$: we can witness this via the mapping

$$(\mathcal{S}, \mathcal{O}) \mapsto (\mathcal{E}, \mathcal{F}, \mathcal{Z}) = (\mathcal{S}, \mathcal{S} \setminus (\mathcal{Q} \cup \mathcal{O}), \mathcal{O}).$$

To be explicit, this is an injective mapping from the data defining \mathfrak{R}' to the data defining $\mathfrak{A}_{\mathcal{Q}, j, c-|\mathcal{H}|}^{(1)}$, and $(\mathcal{S} \setminus (\mathcal{O} \cup \{T, T'\})) \setminus \mathcal{H} = \mathcal{F}$ means that the set corresponding to $(\mathcal{S}, \mathcal{O})$ in \mathfrak{R}' and the set corresponding to $(\mathcal{E}, \mathcal{F}, \mathcal{Z})$ in $\mathfrak{A}_{\mathcal{Q}, j, c-|\mathcal{H}|}^{(1)}$ are the same. Furthermore, we easily check that the conditions for $(\mathcal{E}, \mathcal{F}, \mathcal{Z})$ in the definition of $\mathfrak{A}_{\mathcal{Q}, j, c-|\mathcal{H}|}^{(1)}$ are satisfied.

Let $v = v^j(\mathcal{H} \cup \{T, T'\}) = (|\mathcal{H}| + 2) + 3$ (here we are using that $\mathcal{H} \cup \{T, T'\} \subsetneq \mathcal{S}$, since $|\mathcal{O}| = j - c - 4 \geq 1$ and \mathcal{O} is disjoint from $\mathcal{H} \cup \{T, T'\}$). By our multiset inclusion and [Lemma 8.1\(1\)](#) we have

$$\psi(\mathfrak{R}_{T, T', j, c}^{(1)}, \mathcal{H}) \leq \psi(\mathfrak{A}_{\mathcal{Q}, j, c-|\mathcal{H}|}^{(1)}) = O_{g,k}(zw^g n^{j-v-(c-|\mathcal{H}|)}) = O_{g,k}(zw^g n^{j-c-5}).$$

The desired result follows. \square

Proof of Lemma 8.2(2). Fix any \mathcal{H} , and recall the representation of each $\mathcal{K} \in \mathfrak{R}_{e,T}^{(2)}$ in the form $\mathcal{S} \setminus \{T, T'\}$. We write

$$\psi(\mathfrak{R}_{e,T}^{(2)}, \mathcal{H}) = \sum_{j=4}^g \psi(\mathfrak{R}'_j), \text{ where } \mathfrak{R}'_j = \{\mathcal{K} \setminus \mathcal{H} : \mathcal{K} \in \mathfrak{R}_{e,T}^{(2)}, |\mathcal{K}| = j - 4, \mathcal{K} \supseteq \mathcal{H}\}$$

(that is to say, $\psi(\mathfrak{R}'_j)$ is the contribution to $\psi(\mathfrak{R}_{e,T}^{(2)})$ arising from data (T', \mathcal{S}) with $\mathcal{S} \in \mathfrak{F}_j$). We will prove that $\psi(\mathfrak{R}'_j) = O_{g,k}(zw^g)$ for each j ; the desired result will follow.

Case 1: $|\mathcal{H}| = j - 4$. In this case \mathfrak{R}'_j consists of copies of the empty set. If $j = 4$, the number of such copies is $\psi(\mathfrak{R}'_j) \leq z = O_{g,k}(zw^g)$, by **WS3** applied to T and e . Otherwise, if $j > 4$ then the number of copies is at most z , by **WS1** applied with $\mathcal{R} = \mathcal{H} \cup \{T\}$.

Case 2: $|\mathcal{H}| < j - 4$. In this case we will further sum over possibilities for T' : let $\mathfrak{R}'_{j,T'}$ be the submultiset of \mathfrak{R}'_j arising from a particular choice of T' . Let $\mathcal{Q} = \mathcal{H} \cup \{T, T'\}$, let $v = v^j(\mathcal{H} \cup \{T, T'\}) \geq (|\mathcal{H}| + 2) + 3$, and note that $\mathfrak{R}'_{j,T'} \subseteq \mathfrak{A}_{\mathcal{Q},j,j-4-|\mathcal{H}|}^{(1)}$: the inclusion is witnessed by the injective mapping $\mathcal{S} \mapsto (\mathcal{E}, \mathcal{F}, \mathcal{Z}) = (\mathcal{S}, \mathcal{S} \setminus \mathcal{Q}, \emptyset)$. So, by **Lemma 8.1(1)** we have $\psi(\mathfrak{R}'_{j,T'}) = O_{g,k}(zw^g n^{j-v-(j-4-|\mathcal{H}|)}) = O_{g,k}(zw^g/n)$. The desired bound follows by summing over at most n choices of $T' \supseteq e$. \square

Proof of Lemma 8.2(3). Fix any \mathcal{H} , and recall the representation of each $\mathcal{K} \in \mathfrak{R}_{T,T'}^{(3)}$ in the form $\mathcal{S} \cup \mathcal{S}' \setminus \{T, T', T^*\}$. Write $\mathcal{S} \preceq_{\mathcal{H}} \mathcal{S}'$ when $|\mathcal{H} \cap \mathcal{S}| > |\mathcal{H} \cap \mathcal{S}'|$ or when $|\mathcal{H} \cap \mathcal{S}| = |\mathcal{H} \cap \mathcal{S}'|$ and $|\mathcal{S}| \leq |\mathcal{S}'|$. We slightly modify the definition of $\mathfrak{R}_{T,T'}^{(3)}$ to only consider choices of $(T^*, \mathcal{S}, \mathcal{S}')$ such that $\mathcal{S} \preceq_{\mathcal{H}} \mathcal{S}'$ (due to the symmetry between \mathcal{S} and \mathcal{S}' , this affects $\psi(\mathfrak{R}_{T,T'}^{(3)}, \mathcal{H})$ by a factor of at most two). Let $\mathfrak{R}' = \{\mathcal{K} \setminus \mathcal{H} : \mathcal{K} \in \mathfrak{R}_{T,T'}^{(3)}, \mathcal{K} \supseteq \mathcal{H}\}$, so $\psi(\mathfrak{R}_{T,T'}^{(3)}, \mathcal{H}) = \psi(\mathfrak{R}')$. Keep in mind that for a triple $(T^*, \mathcal{S}, \mathcal{S}')$ contributing to \mathfrak{R}' , we always have $\mathcal{H} \subseteq \mathcal{S} \cup \mathcal{S}' \setminus \{T, T', T^*\}$.

Unlike the previous parts of **Lemma 8.2**, it is not quite as convenient to divide the proof into logically independent cases. It is easier to prove two different bounds by applying different parts of **Lemma 8.1**, and to patch these together to deduce the desired bound $\psi(\mathfrak{R}') = O_{g,k}(z^2 w^{2g})$.

Claim 1. Let $\mathfrak{R}'_{\mathcal{Q}, \mathcal{Q}', j, j', v', q}$ be the submultiset of \mathfrak{R}' arising from data $(T^*, \mathcal{S}, \mathcal{S}')$ such that $\mathcal{S} \in \mathfrak{F}_j$ and $\mathcal{S}' \in \mathfrak{F}_{j'}$, and such that

$$\mathcal{Q} = (\mathcal{S} \cap \mathcal{H}) \cup \{T\}, \quad \mathcal{Q}' = (\mathcal{S}' \cap \mathcal{H}) \cup \{T'\}, \quad v' = v^{j'}(\mathcal{S}' \cap (\mathcal{S} \cup \mathcal{Q}')), \quad q = |\mathcal{S} \cap \mathcal{S}' \setminus \{T, T', T^*\}|.$$

Let $v = v^j(\mathcal{Q})$. Then $\psi(\mathfrak{R}'_{\mathcal{Q}, \mathcal{Q}', j, j', v', q}) = O_{g,k}(z^2 w^{2g} n^{8+q+|\mathcal{H}|-v-v'})$.

Claim 2. Suppose that $\mathcal{H} = \emptyset$. Let \mathfrak{R}''_j be the submultiset of \mathfrak{R}' arising from data $(T^*, \mathcal{S}, \mathcal{S}')$ such that $\mathcal{S}, \mathcal{S}' \in \mathfrak{F}_j$ and $\mathcal{S} \setminus \{T\} = \mathcal{S}' \setminus \{T'\}$. Then $\psi(\mathfrak{R}''_j) = O_{g,k}(zw^g)$.

Before proving these claims, we show how they can be combined to prove that $\psi(\mathfrak{R}') = O_{g,k}(z^2 w^{2g})$. Consider any $\mathcal{K} \in \mathfrak{R}'$ arising from data $(T^*, \mathcal{S}, \mathcal{S}')$, where $\mathcal{S} \in \mathfrak{F}_j$ and $\mathcal{S}' \in \mathfrak{F}_{j'}$. We will show that either

- (A) $v^j((\mathcal{S} \cap \mathcal{H}) \cup \{T\}) + v^{j'}(\mathcal{S}' \cap (\mathcal{H} \cup \{T'\} \cup \mathcal{S})) \geq 8 + |\mathcal{S} \cap \mathcal{S}' \setminus \{T, T', T^*\}| + |\mathcal{H}|$, or
- (B) $j = j'$ and $\mathcal{S} \setminus \{T\} = \mathcal{S}' \setminus \{T'\}$.

It will follow that $\psi(\mathfrak{R}')$ is upper-bounded by a sum of $O_g(1)$ terms of the form $\psi(\mathfrak{R}'_{\mathcal{Q}, \mathcal{Q}', j, j', v', q})$ with $v + v' \geq q + |\mathcal{H}| + 8$, plus a sum of $O_g(1)$ terms of the form $\psi(\mathfrak{R}''_j)$. The conclusion of the lemma will then follow from Claims 1 and 2.

So, we show that either (A) or (B) must hold. Let $\mathcal{S} \in \mathfrak{F}_j$, $\mathcal{S}' \in \mathfrak{F}_{j'}$. We can write $\mathcal{S}' \cap (\mathcal{H} \cup \{T'\} \cup \mathcal{S})$ as the disjoint union $\{T', T^*\} \cup (\mathcal{S}' \cap \mathcal{H}) \cup ((\mathcal{S}' \cap \mathcal{S}) \setminus (\mathcal{H} \cup \{T', T^*\}))$, so

$$\begin{aligned} |(\mathcal{S} \cap \mathcal{H}) \cup \{T\}| + |\mathcal{S}' \cap (\mathcal{H} \cup \{T'\} \cup \mathcal{S})| &= |\mathcal{S} \cap \mathcal{H}| + |\mathcal{S}' \cap \mathcal{H}| + |(\mathcal{S}' \cap \mathcal{S}) \setminus (\mathcal{H} \cup \{T', T^*\})| + 3 \\ &= |\mathcal{H}| + |(\mathcal{S} \cap \mathcal{S}') \setminus \{T, T', T^*\}| + 3. \end{aligned}$$

Recalling the definitions of $v^j(\cdot)$ and $v^{j'}(\cdot)$, and noting that $T^* \notin (\mathcal{S} \cap \mathcal{H}) \cup \{T\}$ and $T', T^* \in \mathcal{S}' \cap (\mathcal{H} \cup \{T'\} \cup \mathcal{S})$, we see that (A) holds unless $|(\mathcal{S} \cap \mathcal{H}) \cup \{T\}| = 1$ and $|\mathcal{S}' \cap (\mathcal{H} \cup \{T'\} \cup \mathcal{S})| = j' - 2$. So, assume that this is the case. It follows from $|(\mathcal{S} \cap \mathcal{H}) \cup \{T\}| = 1$ that $\mathcal{S} \cap \mathcal{H} = \emptyset$, and since we are assuming that $\mathcal{S} \preceq_{\mathcal{H}} \mathcal{S}'$, we further deduce that $\mathcal{H} = \emptyset$ and $j \leq j'$. But $|\mathcal{S}' \cap (\mathcal{H} \cup \{T'\} \cup \mathcal{S})| = j' - 2$ means that $\mathcal{S}' \setminus \{T'\} \subseteq \mathcal{S} \setminus \{T\}$, meaning that $j' \leq j$. So, we have $j = j'$, and our assumption $|\mathcal{S}' \cap (\mathcal{H} \cup \{T'\} \cup \mathcal{S})| = j' - 2$ implies that (B) holds.

Now, to complete the proof of the lemma it suffices to prove Claims 1 and 2. For Claim 2, we consider the injective mapping $(T^*, \mathcal{S}, \mathcal{S}') \mapsto (\mathcal{E}, \mathcal{E}', \mathcal{F}, \mathcal{Z}) = (\mathcal{S}, \mathcal{S}', \mathcal{S} \setminus \{T, T^*\}, \{T^*\})$ to see that $\mathfrak{R}'_{j,j'} \subseteq \mathfrak{A}_{T,T',j,j-4}^{(3)}$. So, the desired result follows from **Lemma 8.1(3)**.

For Claim 1, we have $(\mathcal{S} \setminus (\mathcal{Q} \cup \{T^*\})) \cup (\mathcal{S}' \setminus (\mathcal{Q}' \cup \{T^*\})) = (\mathcal{S} \cup \mathcal{S}') \setminus (\mathcal{H} \cup \{T, T', T^*\})$ (recall that if $T \neq T'$ then $T \notin \mathcal{S}'$ and $T' \notin \mathcal{S}$, by the definition of $\mathfrak{K}_{T, T'}^{(3)}$) and

$$\begin{aligned} |(\mathcal{S} \cup \mathcal{S}') \setminus (\mathcal{H} \cup \{T, T', T^*\})| &= |\mathcal{S} \setminus \{T, T^*\}| + |\mathcal{S}' \setminus \{T', T^*\}| - |(\mathcal{S} \setminus \{T, T^*\}) \cap (\mathcal{S}' \setminus \{T', T^*\})| - |\mathcal{H}| \\ &= (j-4) + (j'-4) - q - |\mathcal{H}|. \end{aligned}$$

Let $f = j + j' - 8 - q - |\mathcal{H}|$; we claim that $\mathfrak{K}_{\mathcal{Q}, \mathcal{Q}', j, j', v', q} \subseteq \mathfrak{A}_{\mathcal{Q}, \mathcal{Q}', j, j', v, v', f}^{(2)}$. The inclusion is given by the injective mapping

$$(T^*, \mathcal{S}, \mathcal{S}') \mapsto (\mathcal{E}, \mathcal{E}', \mathcal{F}, \mathcal{F}', \mathcal{Z}, \mathcal{Z}') = (\mathcal{S}, \mathcal{S}', \mathcal{S} \setminus (\mathcal{Q} \cup \{T^*\}), \mathcal{S}' \setminus (\mathcal{Q}' \cup \{T^*\}), \{T^*\}, \{T^*\}).$$

The conditions for $\mathfrak{A}_{\mathcal{Q}, \mathcal{Q}', j, j', v, v', f}^{(2)}$ are easily checked; note in particular that under this mapping $\mathcal{F} \cup \mathcal{F}' = (\mathcal{S} \cup \mathcal{S}') \setminus (\mathcal{H} \cup \{T, T', T^*\})$ and $\mathcal{Q} \cup \mathcal{Z} \cup \mathcal{Q}' \cup \mathcal{Z}' = \mathcal{H} \cup \{T, T', T^*\}$ are disjoint. The desired result then follows from Lemma 8.1(2). \square

Proof of Lemma 8.2(4). Fix any \mathcal{H} and recall the representation of each $\mathcal{K} \in \mathfrak{K}_{T, j, c-1}^{(4)}$ in the form $(\mathcal{S} \cup \mathcal{S}') \setminus (\mathcal{O} \cup \{T\})$. Let $\mathfrak{K}' = \{\mathcal{K} \setminus \mathcal{H} : \mathcal{K} \in \mathfrak{K}_{T, j, c-1}^{(4)}, \mathcal{K} \supseteq \mathcal{H}\}$, so $\psi(\mathfrak{K}_{T, j, c-1}^{(4)}, \mathcal{H}) = \psi(\mathfrak{K}')$. We may assume $T \notin \mathcal{H}$, for otherwise $\psi(\mathfrak{K}') = 0$. Similarly to the proof of Lemma 8.2(3), we prove three different bounds by applying different parts of Lemma 8.1, and patch these together to deduce the desired bound $\psi(\mathfrak{K}') = O_{g,k}(z^2 w^{2g} n^{j-c-3})$.

Claim 1. Let $\mathfrak{K}'_{j', q, v, v'}$ be the submultiset of \mathfrak{K}' arising from data $(\mathcal{S}, \mathcal{S}', \mathcal{O})$ such that $\mathcal{S}' \in \mathfrak{F}_{j'}$ and such that

$$v = v^j((\mathcal{S} \cap \mathcal{H}) \cup \{T\}), \quad q = |(\mathcal{S} \cap \mathcal{S}') \setminus (\mathcal{O} \cup \{T\})|, \quad v' = v^{j'}(\mathcal{S}' \cap (\mathcal{S} \cup \mathcal{H})).$$

Then $\psi(\mathfrak{K}'_{j', q, v, v'}) = O_{g,k}(z^2 w^{2g} n^{j-v-v'-((c-1)-4-q-|\mathcal{H}|)})$.

Claim 2. Suppose that $\mathcal{H} \neq \emptyset$. Let $\mathfrak{K}''_{j', v, v'}$ be the submultiset of \mathfrak{K}' arising from data $(\mathcal{S}, \mathcal{S}', \mathcal{O})$ such that $\mathcal{S}' \in \mathfrak{F}_{j'}$, $\mathcal{S}' \subseteq \mathcal{S} \cup \mathcal{H}$, $\mathcal{H} \cap \mathcal{S} = \emptyset$, and

$$v = v^j(\{T\} \cup (\mathcal{S}' \cap \mathcal{S})), \quad v' = v^{j'}(\mathcal{H} \cup (\mathcal{S}' \cap \{T\})).$$

Then $\psi(\mathfrak{K}''_{j', v, v'}) = O_{g,k}(z^2 w^g n^{j'+j-v'-v-(c-1)})$.

Claim 3. Suppose that $|\mathcal{H}| = 1$. Let \mathfrak{K}''' be the submultiset of \mathfrak{K}' arising from data $(\mathcal{S}, \mathcal{S}', \mathcal{O})$ such that $\mathcal{S}' \setminus \mathcal{H} = \mathcal{S} \setminus \{T\}$. Then $\psi(\mathfrak{K}''') = O_{g,k}(z^2 w^{2g} n^{j-c-3})$.

Before proving the above three claims, we show how they can be used together to prove that $\psi(\mathfrak{K}') = O_{g,k}(z^2 w^{2g} n^{j-c-3})$. Consider any $\mathcal{K} \in \mathfrak{K}'$ arising from data $(\mathcal{S}, \mathcal{S}', \mathcal{O})$, where $\mathcal{S}' \in \mathfrak{F}_{j'}$. We will show that either

- (A) $v^j((\mathcal{S} \cap \mathcal{H}) \cup \{T\}) + v^{j'}(\mathcal{S}' \cap (\mathcal{S} \cup \mathcal{H})) \geq |(\mathcal{S} \cap \mathcal{S}') \setminus (\mathcal{O} \cup \{T\})| + |\mathcal{H}| + 8$, or
- (B) $\mathcal{H} \neq \emptyset$, $\mathcal{H} \cap \mathcal{S} = \emptyset$, $\mathcal{S}' \subseteq \mathcal{S} \cup \mathcal{H}$, and $v^j(\{T\} \cup (\mathcal{S}' \cap \mathcal{S})) + v^{j'}(\mathcal{H} \cup (\mathcal{S}' \cap \{T\})) \geq j' + 4$, or
- (C) $|\mathcal{H}| = 1$ and $\mathcal{S}' \setminus \mathcal{H} = \mathcal{S} \setminus \{T\}$.

It will follow that $\psi(\mathfrak{K}')$ is upper-bounded by a sum of $O_g(1)$ terms of the form $\psi(\mathfrak{K}'_{j', q, v, v'})$ with $v + v' \geq q + |\mathcal{H}| + 8$, plus a sum of $O_g(1)$ terms of the form $\psi(\mathfrak{K}''_{j', v, v'})$ with $v + v' \geq j' + 4$, plus $\psi(\mathfrak{K}''')$. The conclusion of the lemma will then follow from Claims 1–3.

So, we show that one of (A–C) must hold. Noting that $T \notin \mathcal{H}$, we have

$$|(\mathcal{S} \cap \mathcal{H}) \cup \{T\}| + |\mathcal{S}' \cap (\mathcal{S} \cup \mathcal{H})| = 1 + |\mathcal{H}| + |\mathcal{S}' \cap \mathcal{S}| = 1 + |\mathcal{H}| + |(\mathcal{S} \cap \mathcal{S}') \setminus (\mathcal{O} \cup \{T\})| + |\mathcal{S}' \cap (\mathcal{O} \cup \{T\})|.$$

Recalling the definitions of $v^j(\cdot)$ and $v^{j'}(\cdot)$, and recalling that $|\mathcal{S}' \cap (\mathcal{O} \cup \{T\})| = 2$, we see that (A) holds unless $(\mathcal{S} \cap \mathcal{H}) \cup \{T\}$ has size in $\{1, j-2\}$ and $\mathcal{S}' \cap (\mathcal{S} \cup \mathcal{H})$ has size in $\{1, |\mathcal{S}'|\}$. Assume this is the case. Note that $\mathcal{S}' \cap (\mathcal{S} \cup \mathcal{H})$ contains at least the two triangles in $\mathcal{S}' \cap (\mathcal{O} \cup \{T\})$, so by the second of the two assumptions we have just made, $\mathcal{S}' \subseteq \mathcal{S} \cup \mathcal{H}$. But by the definition of $\mathfrak{K}_{T, j, c-1}^{(4)}$ we have $\mathcal{S}' \not\subseteq \mathcal{S}$, so $\mathcal{H} \neq \emptyset$. Also, \mathcal{O} is disjoint from $\mathcal{H} \cup \{T\}$, so $\mathcal{S}' \not\subseteq \mathcal{H} \cup \{T\}$. So, by the first of our two assumptions, $|(\mathcal{S} \cap \mathcal{H}) \cup \{T\}| = 1$, meaning that $\mathcal{S} \cap \mathcal{H} = \emptyset$.

Now, using the facts that $\mathcal{S}' \subseteq \mathcal{S} \cup \mathcal{H}$ and $\mathcal{S} \cap \mathcal{H} = \emptyset$, we have

$$|\{T\} \cup (\mathcal{S}' \cap \mathcal{S})| + |\mathcal{H} \cup (\mathcal{S}' \cap \{T\})| = |\mathcal{S}' \cap \mathcal{S}| + |\mathcal{H}| + 1 = |\mathcal{S}' \setminus \mathcal{H}| + |\mathcal{H}| + 1 = (|\mathcal{S}'| + 2) - 1.$$

It follows that (B) holds unless $|\mathcal{H} \cup (\mathcal{S}' \cap \{T\})| \in \{1, |\mathcal{S}'|\}$ and $|\{T\} \cup (\mathcal{S}' \cap \mathcal{S})| \in \{1, j-2\}$. Assume this is the case. We have $\mathcal{H} \subseteq \mathcal{S}'$, and since $|\mathcal{S}' \cap (\mathcal{O} \cup \{T\})| = 2$ it is not possible to have $\mathcal{H} \cup (\mathcal{S}' \cap \{T\}) = \mathcal{S}'$, so $|\mathcal{H} \cup (\mathcal{S}' \cap \{T\})| = 1$. Since \mathcal{H} is nonempty it follows that $|\mathcal{H}| = 1$ and $T \notin \mathcal{S}'$. Also, $\{T\} \cup (\mathcal{S}' \cap \mathcal{S})$ contains two triangles in $\mathcal{O} \cup \{T\}$, so $\{T\} \cup (\mathcal{S}' \cap \mathcal{S}) = \mathcal{S}$. Combining all the observations and assumptions we have made so far, we see that (C) must hold.

To complete the proof of the lemma it suffices to prove Claims 1–3. For Claim 1, we consider the injective mapping

$$(\mathcal{S}, \mathcal{S}', \mathcal{O}) \mapsto (\mathcal{E}, \mathcal{E}', \mathcal{F}, \mathcal{F}', \mathcal{Z}, \mathcal{Z}') = (\mathcal{S}, \mathcal{S}', \mathcal{S} \setminus (\mathcal{O} \cup \mathcal{H} \cup \{T\}), \mathcal{S}' \setminus (\mathcal{O} \cup \mathcal{H} \cup \{T\}), \mathcal{O}, \mathcal{S}' \cap (\mathcal{O} \cup \{T\})).$$

With $j' = |\mathcal{S}'|$ and $q = |(\mathcal{S} \cap \mathcal{S}') \setminus (\mathcal{O} \cup \{T\})|$ we have

$$\begin{aligned} |\mathcal{F} \cup \mathcal{F}'| &= |\mathcal{S} \cup \mathcal{S}'| - |\mathcal{O} \cup \mathcal{H} \cup \{T\}| = (j - 2 + j' - 2 - |\mathcal{S} \cap \mathcal{S}'|) - (j - c - 2 + |\mathcal{H}| + 1) \\ &= j' + c - 3 - |\mathcal{S} \cap \mathcal{S}' \cap (\mathcal{O} \cup \{T\})| - q - |\mathcal{H}| \\ &= (c - 1) + (j' - 4) - q - |\mathcal{H}|, \end{aligned}$$

and we can check that $\mathfrak{K}'_{j',q,v,v'} \subseteq \mathfrak{A}_{\mathcal{Q},\mathcal{Q}',j,j',v,v',f}^{(2)}$, where $\mathcal{Q} = (\mathcal{S} \cap \mathcal{H}) \cup \{T\}$, $\mathcal{Q}' = \mathcal{S}' \cap \mathcal{H}$ and $f = (c - 1) + (j' - 4) - q - |\mathcal{H}|$. (Here we use the fact that $T \in K_N[U_k]$ to verify that the triangles of \mathcal{Z}' are all within U_k .) The statement of Claim 1 then follows from [Lemma 8.1\(2\)](#).

For Claim 2, we consider the “reverse order” mapping

$$(\mathcal{S}, \mathcal{S}', \mathcal{O}) \mapsto (\mathcal{E}, \mathcal{E}', \mathcal{F}, \mathcal{F}', \mathcal{Z}, \mathcal{Z}') = (\mathcal{S}', \mathcal{S}, \mathcal{S}' \setminus (\mathcal{O} \cup \mathcal{H} \cup \{T\}), \mathcal{S} \setminus (\mathcal{O} \cup \mathcal{H} \cup \{T\}), \mathcal{S}' \cap \mathcal{O}, \mathcal{O}).$$

In the setting of Claim 2, recalling that $\mathcal{S}' \subseteq \mathcal{S} \cup \mathcal{H}$ (implying $\mathcal{F} \subseteq \mathcal{F}'$), $\mathcal{H} \neq \emptyset$, and $\mathcal{H} \cap \mathcal{S} = \emptyset$, we have

$$|\mathcal{F} \cup \mathcal{F}'| = |\mathcal{F}'| = |\mathcal{S}' \setminus (\mathcal{O} \cup \mathcal{H} \cup \{T\})| = |\mathcal{S}' \setminus (\mathcal{O} \cup \{T\})| = c - 1,$$

and we can check that $\mathfrak{K}''_{j',v,v'} \subseteq \mathfrak{A}_{\mathcal{Q},\{T\},j',j,v',v,c-1}^{(2)}$, where $\mathcal{Q} = \mathcal{H} \cup (\mathcal{S}' \cap \{T\})$. The statement of Claim 2 then follows from [Lemma 8.1\(2\)](#).

Finally, for Claim 3 we let T' be the single triangle in \mathcal{H} , consider the injective mapping

$$(\mathcal{S}, \mathcal{S}', \mathcal{O}) \mapsto (\mathcal{E}, \mathcal{E}', \mathcal{F}, \mathcal{Z}) = (\mathcal{S}, \mathcal{S}', \mathcal{S} \setminus (\{T\} \cup \mathcal{O}), \mathcal{O}),$$

check that $\mathfrak{K}''' \subseteq \mathfrak{A}_{T,T',j,c-1}^{(3)}$, and apply [Lemma 8.1\(3\)](#). \square

Recall from the outline in [Section 2.2.5](#) that we need an initial “sparsification” step which involves a special instance of our high-girth triple process. In order to analyze this process we need a variation on [Lemma 8.2\(2\)](#), in which we prove a sharper weight bound but in a simpler setting without a descending sequence of sets $U_0 \supseteq \dots \supseteq U_k$ (therefore we do not need [Lemma 8.1](#) or the notion of well-spreadness).

Lemma 8.3. *Consider an absorbing structure H (with distinguished set X) satisfying [Ab1](#) and [Ab2](#) in [Theorem 4.1](#) for $g \in \mathbb{N}$, embedded in K_N . Let \mathcal{B} and $\mathfrak{F}_j^{\mathcal{B}}$ be as in [Lemma 7.2](#), and suppose $|\mathcal{B}|^{2g} \leq N^\beta$ for some $\beta \leq 1$. Let \mathcal{W} be the set of triangles in K_N and let $\vec{\pi}$ be the weight system defined by $\pi_T = 1/N$ for all $T \in \mathcal{W}$.*

For a set of vertices U and an edge e , both disjoint from $V(H) \setminus X$, and a triangle T , let $\tilde{\mathfrak{K}}_{e,T,U}$ be the multiset of sets of triangles constructed as follows. Consider each triangle $T' \neq T$ containing the edge e with its third vertex in U , and consider each $\mathcal{S} \in \bigcup_{j=4}^g \mathfrak{F}_j^{\mathcal{B}}$ such that $T, T' \in \mathcal{S}$. For each such (T', \mathcal{S}) , add a copy of $\mathcal{S} \setminus \{T, T'\}$ to $\tilde{\mathfrak{K}}_{e,T,U}$. Then $\kappa(\tilde{\mathfrak{K}}_{e,T,U}) = O_g(1 + |U|N^{\beta-1})$.

Proof. The proof proceeds in basically the same way as [Lemma 8.2\(2\)](#). Fix \mathcal{H} and define $\tilde{\mathfrak{K}}'_j$ as in the proof of [Lemma 8.2\(2\)](#), using $\tilde{\mathfrak{K}}_{e,T,U}$ in place of $\mathfrak{K}_{e,T}^{(2)}$. Explicitly, write

$$\psi(\tilde{\mathfrak{K}}_{e,T,U}, \mathcal{H}) = \sum_{j=4}^g \psi(\tilde{\mathfrak{K}}'_j), \text{ where } \tilde{\mathfrak{K}}'_j = \{\mathcal{K} \setminus \mathcal{H} : \mathcal{K} \in \tilde{\mathfrak{K}}_{e,T,U}, |\mathcal{K}| = j - 4, \mathcal{K} \supseteq \mathcal{H}\}.$$

Case 1: $|\mathcal{H}| = j - 4$. In this case $\tilde{\mathfrak{K}}'_j$ consists of copies of the empty set. The number of such copies is at most the number of Erdős configurations of the form $\mathcal{E} = \mathcal{M} \cup \mathcal{H} \cup \{T, T'\}$, where T' contains no vertices of $V(H) \setminus X$ and $\mathcal{M} \subseteq \mathcal{B}$. By [Ab2](#) with $\mathcal{R} = \mathcal{H} \cup \{T\}$, we must have $\mathcal{M} \subseteq \mathcal{L}_{\mathcal{R}}$, so there are only $O_g(1)$ choices for \mathcal{M} . But $\mathcal{M} \cup \mathcal{R}$ determines the vertex set of \mathcal{E} , by [Lemma 3.2\(1\)](#). That is to say, $\psi(\tilde{\mathfrak{K}}'_j) = O_g(1)$.

Case 2: $|\mathcal{H}| < j - 4$. In this case we further sum over possibilities for T' : let $\tilde{\mathfrak{K}}'_{j,T'}$ be the submultiset of $\tilde{\mathfrak{K}}'_j$ arising from a particular choice of T' . For a particular choice of T' , let $v = v^j(\mathcal{H} \cup \{T, T'\}) \geq (|\mathcal{H}| + 2) + 3$. Using [Lemma 7.2](#) (with $k = 0$) and [WS1](#) arising from it applied to $\mathcal{H} \cup \{T, T'\}$, we compute $\psi(\tilde{\mathfrak{K}}'_{j,T'}) = O_{g,k}(N^{\beta+j-v-(j-4-|\mathcal{H}|)}) = O_{g,k}(N^{\beta-1})$. Summing over at most $|U|$ choices of $T' \supseteq e$ yields the desired bound $\psi(\tilde{\mathfrak{K}}'_j) = O_{g,k}(|U|N^{\beta-1})$. \square

The estimate in the following lemma will be used to prove a high-probability upper bound on the number of forbidden configurations containing a given triangle at a given stage of the proof of [Theorem 1.1](#). Although the computations will be less complicated than [Lemma 8.2](#), here we need a more precise estimate, taking into account both the triangles that have been selected at previous stages, and the edges that were not

covered by the initial sparsification step. To encode this information, the ground set of our weight system will contain a mixture of triangles and edges.

Henceforth, we will be less detailed in our applications of [Lemma 8.1](#); as in the proofs of the different parts of [Lemma 8.2](#), whenever applying [Lemma 8.1](#) we should formally define an injective map which witnesses a multiset inclusion, and deduce an inequality of weights. From now on all applications of [Lemma 8.1](#) will be comparatively simple. In particular we will only need [Lemma 8.1\(1\)](#).

The following lemma will be used in [Section 10.1](#).

Lemma 8.4. *Fix positive real numbers $p \leq 1$ and y, z , a pair of integers $4 \leq j \leq j' \leq g$, a descending sequence of subsets $V(K_N) = U_0 \supseteq \dots \supseteq U_k$, and let $n = |U_k|$.*

- *Let $\mathfrak{F}_{j'}$ be a collection of sets of $j' - 2$ triangles in K_N , which is (y, z) -well-spread with respect to our sequence of sets.*
- *Let \mathcal{W}_1 be the set of triangles in K_N , and let \mathcal{W}_2 be the set of edges of $K_N[U_k]$.*
- *Let $\bar{\pi}$ be the weight system for $\mathcal{W} = \mathcal{W}_1 \cup \mathcal{W}_2$ defined by $\pi_T = 1/|U_{\text{lev}(T)}|$ for each $T \in \mathcal{W}_1$ and $\pi_e = p$ for each $e \in \mathcal{W}_2$.*

Let T be a triangle in $K_N[U_k]$, and let $\mathfrak{B}_{T,j,j'}$ be the multiset (of sets of edges and triangles) constructed as follows. Consider every $\mathcal{S} \in \mathfrak{F}_{j'}$ and every partition $\mathcal{S} = \{T\} \cup \mathcal{Z} \cup \mathcal{F}$, such that $|\mathcal{Z}| = j - 3$ and every triangle in \mathcal{Z} is within U_k . For each such $(\mathcal{S}, \mathcal{F}, \mathcal{Z})$, add a copy of $E(\mathcal{Z}) \cup \mathcal{F}$ (i.e., $3(j - 3)$ edges and $j' - j$ triangles) to $\mathfrak{B}_{T,j,j'}$.

If $zn^{-1} \leq yp^{3(j-3)}$ then $\kappa(\mathfrak{B}_{T,j,j'}) = O_{g,k}(yp^{3(j-3)}n^{j-3})$.

Proof. First note that $\psi(\mathfrak{B}_{T,j,j'}, \emptyset) = O_{g,k}(yp^{3(j-3)}n^{j-3})$. Indeed, if we were to ignore the edges in $E(\mathcal{Z})$ in the definition of $\mathfrak{B}_{T,j,j'}$, we would have a bound of the form $\psi(\mathfrak{A}_{\{T\},j',j'-j}^{(1)}) = O_{g,k}(yn^{j-3})$ by [Lemma 8.1\(1\)](#). Then, the $3(j - 3)$ edges in the triangles in \mathcal{Z} contribute a factor of $p^{3(j-3)}$.

So it suffices to study weights of the form $\psi(\mathfrak{B}_{T,j,j'}, \mathcal{H}^2 \cup \mathcal{H}^3)$, where $\mathcal{H}^2 \subseteq \mathcal{W}_2$ is a set of edges and $\mathcal{H}^3 \subseteq \mathcal{W}_1$ is a set of triangles, and $\mathcal{H}^2 \cup \mathcal{H}^3 \neq \emptyset$. Fix such $\mathcal{H}^2, \mathcal{H}^3$. Actually, in this case where $\mathcal{H}^2 \cup \mathcal{H}^3 \neq \emptyset$, we will be able to get away with quite a crude bound that completely ignores the edge-weights π_e (only taking the triangle-weights π_T into account). Let $\bar{\pi}'$ be the modified weight system where we set $\pi'_e = 1$ for all $e \in \mathcal{W}_2$ and $\pi'_S = \pi_S$ for all $S \in \mathcal{W}_1$ (i.e., we ignore the influence of the edges in \mathcal{W}_2), and note that

$$\psi^{(\bar{\pi})}(\mathfrak{B}_{T,j,j'}, \mathcal{H}^2 \cup \mathcal{H}^3) \leq \psi^{(\bar{\pi}')}(\mathfrak{B}_{T,j,j'}, \mathcal{H}^2 \cup \mathcal{H}^3).$$

For convenience, from now on we write ψ' instead of $\psi^{(\bar{\pi}')}$, and seek to bound $\psi'(\mathfrak{B}_{T,j,j'}, \mathcal{H}^2 \cup \mathcal{H}^3)$.

Case 1: $\mathcal{H}^3 \neq \emptyset$. In this case, we apply [Lemma 8.1\(1\)](#) to obtain

$$\psi'(\mathfrak{B}_{T,j,j'}, \mathcal{H}^2 \cup \mathcal{H}^3) \leq \psi'(\mathfrak{A}_{\{T\},j',j'-j-|\mathcal{H}^3|}^{(1)}) = O_{g,k}(zn^{j-3-|\mathcal{H}^3|}) = O_{g,k}(zn^{j-4}).$$

The desired bound then follows from the assumption $z \leq yp^{3(j-3)}n$.

Case 2: $\mathcal{H}^3 = \emptyset$ and $j' > 4$. In this case we fix an edge $e \in \mathcal{H}^2$ (recall that $\mathcal{H}^2 \cup \mathcal{H}^3 \neq \emptyset$) and again apply [Lemma 8.1\(1\)](#) to obtain

$$\psi'(\mathfrak{B}_{T,j,j'}, \mathcal{H}^2 \cup \mathcal{H}^3) \leq \psi'(\mathfrak{B}_{T,j,j'}, \{e\}) \leq \sum_{\substack{T' \supseteq e \\ T' \subseteq U_k}} \psi'(\mathfrak{A}_{\{T,T'\},j',j'-j}^{(1)}) = n \cdot O_{g,k}(zn^{j-v^{j'}(\{T,T'\})}).$$

In the second inequality, we may impose $T' \subseteq U_k$ since each term in $\psi'(\mathfrak{B}_{T,j,j'}, \{e\})$ comes from data $(\mathcal{S}, \mathcal{F}, \mathcal{Z})$ for which $e \in \mathcal{H}^2 \subseteq E(\mathcal{Z})$ (meaning that $T' \in \mathcal{Z}$ is within U_k). Now, since we are assuming $j' > 4$, we have $v^{j'}(\{T, T'\}) = 5$, and the desired bound follows (again using that $z \leq yp^{3(j-3)}n$).

Case 3: $\mathcal{H}^3 = \emptyset$ and $j' = 4$. In this case we must have $j = j' = 4$, and $\psi'(\mathfrak{B}_{T,j,j'}, \mathcal{H}^2 \cup \mathcal{H}^3)$ is simply the number of $\mathcal{S} \in \mathfrak{F}_j$ which consist of T and a second triangle $T' \supseteq e$ within U_k . This number is at most $z \leq yp^{3(j-3)}n$ by [WS3](#) applied to T and e . \square

For the remaining lemmas in this section it will be convenient to have some general weight estimates concerning the set of triangles containing a given edge.

Definition 8.5. For an edge e and a set of vertices S in a graph K_N , let $\mathcal{N}_e(S)$, be the set of triangles in K_N which contain e and a third vertex $v \in S$. We say $\mathcal{N}_e(S)$ is the *fan* of e with respect to S .

Lemma 8.6. *Fix positive real numbers $z, w \geq 1$, a positive integer $4 \leq j \leq g$, a descending sequence of subsets $V(K_N) = U_0 \supseteq \dots \supseteq U_k$, and let $n = |U_k|$. Let \mathfrak{F}_j be a z -well-spread collection of sets of $j - 2$ triangles in K_N , with respect to our sequence of sets. Let \mathcal{W} be the set of triangles in K_N , and let $\bar{\pi}$ be the weight system for \mathcal{W} defined by $\pi_T = w/|U_{\text{lev}(T)}|$ for each $T \in \mathcal{W}$.*

- (1) Fix an edge $e \subseteq U_k$ and a nonempty set \mathcal{Q} of triangles in K_N . Let $\mathfrak{C}_{\mathcal{Q},e,j}^{(1)}$ be the multiset of sets of triangles constructed as follows. Consider every $\mathcal{S} \in \mathfrak{F}_j$ and every partition $\mathcal{S} = \{T\} \cup \mathcal{Q} \cup \mathcal{F}$ such that $|\mathcal{F}| = j - 3 - |\mathcal{Q}|$ and $T \in \mathcal{N}_e(U_k)$. For each such $(\mathcal{S}, \mathcal{F}, T)$, add a copy of \mathcal{F} to $\mathfrak{C}_{\mathcal{Q},e,j}^{(1)}$. Then $\psi(\mathfrak{C}_{\mathcal{Q},e,j}^{(1)}) = O_{g,k}(zw^{g-1})$.
- (2) Fix an edge $e \subseteq U_k$ and a nonempty set \mathcal{Q} of triangles in K_N . Let $\mathfrak{C}_{\mathcal{Q},e,j}^{(2)}$ be the multiset of sets of triangles constructed as follows. As in (1), consider every $\mathcal{S} \in \mathfrak{F}_j$ and every partition $\mathcal{S} = \{T\} \cup \mathcal{Q} \cup \mathcal{F}$ such that $|\mathcal{F}| = j - 3 - |\mathcal{Q}|$ and $T \in \mathcal{N}_e(U_k)$. For each such $(\mathcal{S}, \mathcal{F}, T)$, add a copy of $\mathcal{F} \cup \{T\}$ (instead of \mathcal{F} , as in (1)) to $\mathfrak{C}_{\mathcal{Q},e,j}^{(2)}$. Then $\psi(\mathfrak{C}_{\mathcal{Q},e,j}^{(2)}) = O_{g,k}(zw^g/n)$.
- (3) Fix distinct edges $e, e' \subseteq U_k$. Let $\mathfrak{C}_{e,e',j}^{(3)}$ be the multiset of sets of triangles constructed as follows. Consider every $\mathcal{S} \in \mathfrak{F}_j$ such that e' appears in one of the triangles in \mathcal{S} , and consider every partition $\mathcal{S} = \{T\} \cup \mathcal{F}$ such that $|\mathcal{F}| = j - 3$, $T \in \mathcal{N}_e(U_k)$. For each such $(\mathcal{S}, \mathcal{F}, T)$, add a copy of \mathcal{F} to $\mathfrak{C}_{e,e',j}^{(3)}$. Then $\psi(\mathfrak{C}_{e,e',j}^{(3)}) = O_{g,k}(zw^{g-1})$.
- (4) Fix distinct edges $e, e' \subseteq U_k$. Let $\mathfrak{C}_{e,e',j}^{(4)}$ be defined in the same way as $\mathfrak{C}_{e,e',j}^{(3)}$, except that for each $(\mathcal{S}, \mathcal{F}, T)$ as in (3), we add a copy of \mathcal{S} to $\mathfrak{C}_{e,e',j}^{(4)}$ (instead of \mathcal{F} , as in (3)). Then $\psi(\mathfrak{C}_{e,e',j}^{(4)}) = O_{g,k}(zw^g/n)$.

Proof of Lemma 8.6(1–2). First note that (1) implies (2). Indeed, the terms in $\psi(\mathfrak{C}_{\mathcal{Q},e,j}^{(2)})$ exactly correspond to those in $\psi(\mathfrak{C}_{\mathcal{Q},e,j}^{(1)})$, except that each has an extra factor of $\pi_T = w/|U_{\text{lev}(T)}| = w/n$ for some $T \in \mathcal{N}_e(U_k)$. So, it suffices to prove (1). *Case 1:* $1 \leq |\mathcal{Q}| \leq j - 4$. Let \mathfrak{C}'_T be the submultiset of $\mathfrak{C}_{\mathcal{Q},e,j}^{(1)}$ corresponding to a particular choice of $T \in \mathcal{N}_e(U_k)$, so that

$$\begin{aligned} \psi(\mathfrak{C}_{\mathcal{Q},e,j}^{(1)}) &= \sum_{T \in \mathcal{N}_e(U_k)} \psi(\mathfrak{C}'_T) \leq \sum_{T \in \mathcal{N}_e(U_k)} \psi(\mathfrak{A}_{\mathcal{Q} \cup \{T\}, j, j-3-|\mathcal{Q}|}^{(1)}) \\ &= n \cdot O_{g,k}(zw^{g-1}n^{j-v^j(\mathcal{Q} \cup \{T\})-(j-3-|\mathcal{Q}|)}) = n \cdot O_{g,k}(zw^{g-1}n^{-1}) = O_{g,k}(zw^{g-1}), \end{aligned}$$

by Lemma 8.1(1).

Case 2: $|\mathcal{Q}| = j - 3 \geq 2$. In this case, $\psi(\mathfrak{C}_{\mathcal{Q},e,j}^{(1)})$ is simply the number of choices of $\mathcal{S} \in \mathfrak{F}_j$ and $T \in \mathcal{N}_e(U_k)$ such that $\mathcal{S} = \mathcal{Q} \cup \{T\}$. This number is at most $zn^{j-v^j(\mathcal{Q})} \leq z$ by **WS1** (here we are using that $e \subseteq U_k$).

Case 3: $j = 4$ and $|\mathcal{Q}| = 1$. In this case, let T^* be the single triangle in \mathcal{Q} and note that $\psi(\mathfrak{C}_{\mathcal{Q},e,j}^{(1)})$ is simply the number of $\mathcal{S} \in \mathfrak{F}_j$ which consist of T^* and a second triangle $T \supseteq e$ which is within U_k . This number is at most z by **WS3** applied to T^* and e . \square

Proof of Lemma 8.6(3–4). Similarly to the last proof, we see that (3) implies (4), so it suffices to prove (3). First, let \mathfrak{C}' be the submultiset of $\mathfrak{C}_{e,e',j}^{(3)}$ arising from data $(\mathcal{S}, \mathcal{F}, T)$ for which e, e' are both contained in the same triangle of \mathcal{S} (then e and e' uniquely determine this triangle, which must be T). We have

$$\psi(\mathfrak{C}') \leq \psi(\mathfrak{A}_{\{T\}, j, j-3}^{(1)}) = O_{g,k}(zw^{g-1}n^{j-v^j(\{T\})-(j-3)}) = O_{g,k}(zw^{g-1})$$

by Lemma 8.1(1).

Let $\mathfrak{C}'' = \mathfrak{C}_{e,e',j}^{(3)} \setminus \mathfrak{C}'$ be the complementary submultiset of $\mathfrak{C}_{e,e',j}^{(3)}$ arising from data $(\mathcal{S}, \mathcal{F}, T)$ for which e, e' belong to different triangles of \mathcal{S} . It now suffices to prove that $\psi(\mathfrak{C}'') = O_{g,k}(zw^{g-1})$.

Case 1: $j \geq 5$. Let $\mathfrak{C}''_{T,T'}$ be the submultiset of \mathfrak{C}'' corresponding to a particular choice of $T \in \mathcal{N}_e(U_k)$, and for which e' is contained in a particular triangle $T' \in \mathcal{F}$. Note that $\psi(\mathfrak{C}''_{T,T'}) \leq \pi_{T'}\psi(\mathfrak{A}_{\{T,T'\}, j, j-4}^{(1)})$, and recall that $\pi_{T'} = w/|U_{\text{lev}(T')}|$, so using Lemma 8.1(1), we have

$$\begin{aligned} \psi(\mathfrak{C}'') &= \sum_{T \in \mathcal{N}_e(U_k)} \sum_{T' \supseteq e'} \psi(\mathfrak{C}''_{T,T'}) \leq \sum_{T \in \mathcal{N}_e(U_k)} \sum_{i=0}^k \sum_{\substack{T' \supseteq e' \\ \text{lev}(T')=i}} \frac{w}{|U_i|} \psi(\mathfrak{A}_{\{T,T'\}, j, j-4}^{(1)}) \\ &= n \cdot \sum_{i=0}^k |U_i| \cdot \frac{w}{|U_i|} \cdot O_{g,k}(zn^{j-v^j(\{T,T'\})-(j-4)}) = O_{g,k}(zw). \end{aligned}$$

Case 2: $j = 4$. Let $\mathfrak{C}''_{T'}$ be the submultiset of \mathfrak{C}'' for which e' is contained in a particular triangle $T' \in \mathcal{F}$. Then $\psi(\mathfrak{C}''_{T'})$ is $\pi_{T'}$ times the number of choices of $\mathcal{S} \in \mathfrak{F}_j$ which consist of T' and a second triangle $T \supseteq e$

which is within U_k . This number is at most z by **WS3** applied to T' and e , so

$$\psi(\mathfrak{C}'') \leq \sum_{T' \supseteq e'} \psi(\mathfrak{C}'_{T'}) \leq \sum_{i=0}^k \sum_{\substack{T' \supseteq e' \\ \text{lev}(T')=i}} \frac{w}{|U_i|} \cdot z \leq \sum_{i=0}^k |U_i| \cdot \frac{w}{|U_i|} \cdot z = O_k(zw),$$

as desired. \square

We now use **Lemma 8.6** to prove some technical maximum weight estimates. As in **Lemma 8.4**, we will need to consider weight systems whose ground sets consist of mixtures of triangles and edges.

In the following lemma \mathcal{W}_1 and \mathcal{W}_2 will be two copies of the same set of triangles but their weights will be different. The reason for this is that in the proof of **Theorem 1.1** we will need to distinguish between triangles chosen during the initial sparsification step, and triangles chosen during the iterative ‘‘vortex’’ procedure. This lemma will be used in **Section 10.2**.

Lemma 8.7. *Fix positive real numbers $y, z \geq 1$ and $r, p \leq 1$. Also fix a positive integer $4 \leq j \leq g$, a descending sequence of subsets $V(K_N) = U_0 \supseteq \dots \supseteq U_k \supseteq U_{k+1}$, and let $n = |U_k|$.*

- Let $U_k^* = U_k \setminus U_{k+1}$.
- Let \mathfrak{F}_j be a collection of sets of $j-2$ triangles in K_N which is (y, z) -well-spread with respect to the sequence $U_0 \supseteq \dots \supseteq U_k$ (not including U_{k+1}).
- Let \mathcal{W}' be the set of triangles in K_N , and let $\mathcal{W}_1, \mathcal{W}_2$ be two disjoint ‘‘marked’’ copies of \mathcal{W}' . Specifically, for $T \in \mathcal{W}'$, we write $(T, 1)$ for its copy in \mathcal{W}_1 , and $(T, 2)$ for its copy in \mathcal{W}_2 .
- Let \mathcal{W}_3 be the set of edges in K_N between U_k^* and U_{k+1} .
- Let $\bar{\pi}$ be the weight system for $\mathcal{W} = \mathcal{W}_1 \cup \mathcal{W}_2 \cup \mathcal{W}_3$ defined by $\pi_{T,1} = 1/N$ and $\pi_{T,2} = p/|U_{\text{lev}(T)}|$ for each triangle T , and $\pi_e = rp$ for each $e \in \mathcal{W}_3$.

(In the last bullet point, and for the rest of the lemma statement and proof, we define $\text{lev}(\cdot)$ with respect to $U_0 \supseteq \dots \supseteq U_k$, without U_{k+1} .)

Let e be an edge in $K_N[U_k^*]$, and let $\mathfrak{L}_{e,j}$ be the multiset (of sets of edges and ‘‘marked’’ triangles) constructed as follows. Consider each $\mathcal{S} \in \mathfrak{F}_j$ and $T \in \mathcal{S} \cap \mathcal{N}_e(U_{k+1})$, and consider each ‘‘marking’’ function $\Phi : \mathcal{S} \setminus \{T\} \rightarrow \{1, 2\}$ that gives at least one triangle the mark ‘‘2’’. For each such (\mathcal{S}, T, Φ) we add to $\mathfrak{L}_{e,j}$ a copy of

$$(E(T) \setminus \{e\}) \cup \{(T', \Phi(T')) : T' \in \mathcal{S} \setminus \{T\}\}.$$

If $z/|U_{k+1}| \leq yr^2p^3$ then $\kappa(\mathfrak{L}_{e,j}) = O_{g,k}(yr^2p^3|U_{k+1}|)$.

Proof. The proof will follow a similar strategy to **Lemma 8.4**. Let $\bar{\pi}'$ be the weight system on the ground set $\mathcal{W}' \cup \mathcal{W}_3$ where we set $\pi'_e = 1$ for all $e \in \mathcal{W}_3$ (i.e., we ignore the influence of the edges in \mathcal{W}_3), and set $\pi'_{T'} = 1/|U_{\text{lev}(T')}|$. So, for each triangle $T \in \mathcal{W}'$ we have $\pi_{T,1} \leq \pi'_{T'}$ and $\pi_{T,2} = p \cdot \pi'_{T'}$. For convenience we write ψ' instead of $\psi(\bar{\pi}')$ (so ψ with no superscripts will be reserved for weights with respect to $\bar{\pi}$).

First we study $\psi(\mathfrak{L}_{e,j}, \emptyset)$. Let \mathfrak{L}'_T be the submultiset of $\mathfrak{L}_{e,j}$ corresponding to a particular choice of T , and note that $\psi(\mathfrak{L}'_T) \leq 2^{j-3}p(rp)^2\psi'(\mathfrak{A}_{\{T\},j,j-3}^{(1)})$. Indeed, first note that for any \mathcal{S}, T there are $2^{j-3} - 1 \leq 2^{j-3}$ choices of Φ . For each such (\mathcal{S}, T, Φ) contributing to $\mathfrak{L}_{e,j}$ we have $\pi_{T',\Phi(T')} \leq \pi'_{T'}$ for all $T' \in \mathcal{S} \setminus \{T\}$, and $\pi_{T',\Phi(T')} = \pi_{T',2} = p \cdot \pi'_{T'}$ for at least one $T' \in \mathcal{S} \setminus \{T\}$. The factor of $(rp)^2$ is due to the weights $\pi_{e'} = rp$ for the two edges $e' \in E(T) \setminus \{e\}$. We then have

$$\psi(\mathfrak{L}_{e,j}, \emptyset) \leq \sum_{T \in \mathcal{N}_e(U_{k+1})} \psi(\mathfrak{L}'_T) \leq |U_{k+1}| 2^{j-3} p (rp)^2 \cdot O_{g,k}(yn^{j-3-(j-3)}) = O_{g,k}(|U_{k+1}| r^2 p^3 y),$$

by **Lemma 8.1**(1).

So, it now suffices to study weights of the form $\psi(\mathfrak{L}_{e,j}, \mathcal{H}^2 \cup \mathcal{H}^3)$, where $\mathcal{H}^2 \subseteq \mathcal{W}_3 \setminus \{e\}$ is a set of edges and $\mathcal{H}^3 \subseteq \mathcal{W}_1 \cup \mathcal{W}_2$ is a set of ‘‘marked’’ triangles, and $\mathcal{H}^2 \cup \mathcal{H}^3 \neq \emptyset$. Actually, in this case where $\mathcal{H}^2 \cup \mathcal{H}^3 \neq \emptyset$, we will be able to get away with a crude bound that completely ignores the edge-weights π_e and ignores the distinction between \mathcal{W}_1 and \mathcal{W}_2 . To be precise, let $\mathcal{H}^{3'} \subseteq \mathcal{W}'$ be the set of triangles obtained by removing the marks from the triangles in \mathcal{H}^3 , and let $\mathfrak{L}''_{\mathcal{H}^2}$ be the multiset of sets of triangles obtained by including a copy of $\mathcal{S} \setminus \{T\}$ for each $\mathcal{S} \in \mathfrak{F}_j$ and $T \in \mathcal{S} \cap \mathcal{N}_e(U_{k+1})$ such that $\mathcal{H}^2 \subseteq E(\mathcal{S})$. Note that $\psi(\mathfrak{L}_{e,j}, \mathcal{H}^2 \cup \mathcal{H}^3) \leq 2^{j-3}\psi'(\mathfrak{L}''_{\mathcal{H}^2}, \mathcal{H}^{3'})$. So, recalling our assumption $z/|U_{k+1}| \leq yr^2p^3$, it suffices to prove that $\psi'(\mathfrak{L}''_{\mathcal{H}^2}, \mathcal{H}^{3'}) = O_{g,k}(z)$.

Case 1: $\mathcal{H}^3 \neq \emptyset$. In this case, we observe that $\mathcal{N}_e(U_{k+1}) \subseteq \mathcal{N}_e(U_k)$ and apply **Lemma 8.6**(1) to see that

$$\psi'(\mathfrak{L}''_{\mathcal{H}^2}, \mathcal{H}^{3'}) \leq \psi'(\mathfrak{C}_{\mathcal{H}^{3'}, e, j}^{(1)}) = O_{g,k}(z).$$

Case 2: $\mathcal{H}^2 \neq \emptyset$ and $\mathcal{H}^3 = \emptyset$. In this case we fix $e' \in \mathcal{H}^2$ and apply **Lemma 8.6**(3) to see that

$$\psi'(\mathfrak{L}''_{\mathcal{H}^2}, \mathcal{H}^{3'}) \leq \psi'(\mathfrak{C}_{e', e, j}^{(3)}) = O_{g,k}(z). \quad \square$$

In the following lemma we have *three* distinguished sets of triangles: one corresponds to the initial sparsification step, one corresponds to the triangles chosen in previous steps of the “vortex” procedure, and the other corresponds to triangles that survive an additional sparsification in preparation for the “covering crossing edges” stage of the proof (recall the outline in [Section 2](#)). This lemma will be used in [Section 10.3](#).

Lemma 8.8. *Fix positive real numbers y, z, r, p, γ with $y, z \geq 1$ and $r, p \leq 1$. Also fix a positive integer $4 \leq j \leq g$, a descending sequence of subsets $V(K_N) = U_0 \supseteq \dots \supseteq U_k \supseteq U_{k+1}$, and let $n = |U_k|$.*

- Let $U_k^* = U_k \setminus U_{k+1}$.
- Let \mathfrak{F}_j be a collection of sets of $j-2$ triangles in K_N which is (y, z) -well-spread with respect to the sequence $U_0 \supseteq \dots \supseteq U_k$ (not including U_{k+1}).
- Let \mathcal{W}' be the set of triangles in K_N . For $i \in \{1, 2\}$ let $\mathcal{W}_i = \{(T, i) : T \in \mathcal{W}'\}$, and let

$$\mathcal{W}_3 = \{(T, 3) : T \in \mathcal{W}', T \text{ has one vertex in } U_k^* \text{ and two vertices in } U_{k+1}\}.$$

- Let \mathcal{W}_4 be the set of edges in $K_N[U_k]$ which are not between two vertices in U_k^* .
- Let $\bar{\pi}$ be the weight system for $\mathcal{W} = \mathcal{W}_1 \cup \mathcal{W}_2 \cup \mathcal{W}_3 \cup \mathcal{W}_4$ defined as follows. Let $\pi_{T,1} = 1/N$ and $\pi_{T,2} = p/|U_{\text{lev}(T)}|$ for each $T \in \mathcal{W}'$, and $\pi_{T,3} = n^\gamma/(rp^2|U_{k+1}|)$ for each $(T, 3) \in \mathcal{W}_3$. Then, let $\pi_{e'} = rp$ for each edge e' between U_k^* and U_{k+1} , and let $\pi_{e'} = p$ for each edge e' between two vertices of U_{k+1} .

(In the last bullet point, and for the rest of the lemma statement and proof, we define $\text{lev}(\cdot)$ with respect to $U_0 \supseteq \dots \supseteq U_k$, without U_{k+1} .)

Let e be an edge of K_N between U_k^* and U_{k+1} , and let $\mathfrak{M}_{e,j}$ be the multiset (of sets of edges and “marked” triangles) constructed as follows. Consider each $\mathcal{S} \in \mathfrak{F}_j$ and $T \in \mathcal{S} \cap \mathcal{N}_e(U_{k+1})$, such that at least one of the triangles in $\mathcal{S} \setminus \{T\}$ lies within U_k , and consider each marking function $\Phi : \mathcal{S} \setminus \{T\} \rightarrow \{1, 2, 3\}$ that gives at least one triangle a mark different from “1”, and for which $\Phi(T') = 3$ implies $(T', 3) \in \mathcal{W}_3$. For each such (\mathcal{S}, T, Φ) we add to $\mathfrak{M}_{e,j}$ a copy of

$$(E(T) \setminus \{e\}) \cup \bigcup_{\substack{T' \in \mathcal{S} \setminus T \\ \Phi(T')=3}} E(T') \cup \{(T', \Phi(T')) : T' \in \mathcal{S} \setminus \{T\}\} \cup \{(T, 3)\}.$$

If $z(n^\gamma|U_k|/(rp^2|U_{k+1}|))^{g+1}/n \leq y$ and $rn^\gamma \leq p|U_{k+1}|/|U_k|$ and $pn^\gamma \leq 1$, then $\kappa(\mathfrak{M}_{e,j}) = O_{g,k}(y)$.

Proof. We proceed in a similar way to the proofs of [Lemmas 8.4](#) and [8.7](#), but here there are more cases to consider. Let $\bar{\pi}'$ be the weight system on the ground set \mathcal{W}' where we let $\pi_{T'}' = 1/|U_{\text{lev}(T')}|$ for each $T' \in \mathcal{W}'$. We write ψ' instead of $\psi(\bar{\pi}')$. Note that for each $T' \in \mathcal{W}'$ we have $\pi_{T',1} \leq \pi_{T'}'$ and $\pi_{T',2} = p \cdot \pi_{T'}'$, and for each $(T', 3) \in \mathcal{W}_3$ we have

$$\pi_{T',3} \prod_{e' \in E(T')} \pi_{e'} = \frac{n^\gamma}{rp^2|U_{k+1}|} \cdot p(rp)^2 = \frac{rpn^\gamma}{|U_{k+1}|} \leq \frac{p^2}{|U_k|} \leq \frac{p}{|U_{\text{lev}(T')}|} = p \cdot \pi_{T'}'.$$

Let \mathfrak{M}'_T be the subcollection of \mathfrak{M} corresponding to a particular choice of T , and note that

$$\psi(\mathfrak{M}'_T) \leq 3^{j-3} \cdot \pi_{T,3} \cdot p(rp) \cdot p \cdot \psi'(\mathfrak{A}_{\{T\},j,j-3}^{(1)}) = \frac{3^{j-3}pn^\gamma}{|U_{k+1}|} \psi'(\mathfrak{A}_{\{T\},j,j-3}^{(1)}).$$

Indeed, here we are using that for each \mathcal{S}, T there are at most 3^{j-3} choices for Φ , we are using that $\prod_{e' \in E(T) \setminus \{e\}} \pi_{e'} = p(rp)$, and we are using that at least one of the triangles in $\mathcal{S} \setminus \{T\}$ has mark different from “1” (from which we gain a factor of p). We deduce using [Lemma 8.1\(1\)](#) that

$$\begin{aligned} \psi(\mathfrak{M}_{e,j}, \emptyset) &= \sum_{T \in \mathcal{N}_e(U_{k+1})} \psi(\mathfrak{M}'_T) \leq \frac{3^{j-3}pn^\gamma}{|U_{k+1}|} \sum_{T \in \mathcal{N}_e(U_{k+1})} \psi'(\mathfrak{A}_{\{T\},j,j-3}^{(1)}) \\ &= \frac{pn^\gamma}{|U_{k+1}|} \cdot |U_{k+1}| \cdot O_{g,k}(yn^{j-(1+2)-(j-3)}) = O_{g,k}(ypn^\gamma) = O_{g,k}(y). \end{aligned}$$

At this point, similarly to the proofs of [Lemmas 8.4](#) and [8.7](#), we will break into cases and use much cruder estimates to handle $\psi(\mathfrak{M}_{e,j}, \mathcal{H})$ for $\mathcal{H} \neq \emptyset$. However, in the upcoming *Case 2*, we will need to assume a certain technical condition, so first we explicitly consider a special situation which we have to handle quite precisely.

In particular, suppose that $\mathcal{H}^3 \subseteq \mathcal{W}_1 \cup \mathcal{W}_2 \cup \mathcal{W}_3$ consists of a single (marked) triangle and that $\{e\} \cup \mathcal{H}^2 \subseteq E(\mathcal{H}^3)$. We call this the “exceptional case”. In this case, we need only consider data (\mathcal{S}, T, Φ) for which $\mathcal{H}^3 = \{T\}$ and we see

$$\psi(\mathfrak{M}_{e,j}, \mathcal{H}^2 \cup \mathcal{H}^3) \leq 3^{j-2} \psi'(\mathfrak{A}_{\{T\},j,j-3}^{(1)}) = O_{g,k}(y),$$

again using [Lemma 8.1\(1\)](#).

Now, fix $\mathcal{H}^2 \subseteq \mathcal{W}_4 \setminus \{e\}$ and $\mathcal{H}^3 \subseteq \mathcal{W}_1 \cup \mathcal{W}_2 \cup \mathcal{W}_3$ such that $\mathcal{H}^2 \cup \mathcal{H}^3 \neq \emptyset$ and such that the exceptional case does not hold; it suffices to bound $\psi(\mathfrak{M}_{e,j}, \mathcal{H}^2 \cup \mathcal{H}^3)$. We can afford to be rather crude with our estimates. Let $w = n^\gamma |U_k| / (rp^2 |U_{k+1}|) \geq 1$, let $\mathcal{H}^{3'} \subseteq \mathcal{W}$ be the set of unmarked triangles underlying \mathcal{H}^3 , and let $\mathfrak{M}_{\mathcal{H}^2}''$ be the collection of all $\mathcal{S} \in \mathfrak{F}_j$ so that there is some $T \in \mathcal{S} \cap \mathcal{N}_e(U_{k+1})$ (if such a T exists, it is unique) for which at least one of the triangles in $\mathcal{S} \setminus \{T\}$ lies within U_k and for which $\mathcal{H}^2 \subseteq E(\mathcal{S})$. Note that for $(T', 3) \in \mathcal{W}_3$ we have $\pi_{T',3} = w \cdot \pi_{T'}$, which implies the crude bound

$$\psi(\mathfrak{M}_{e,j}, \mathcal{H}^2 \cup \mathcal{H}^3) \leq 3^{j-3} \cdot w^{j-2-|\mathcal{H}^3|} \cdot \psi'(\mathfrak{M}_{\mathcal{H}^2}'', \mathcal{H}^{3'}). \quad (8.1)$$

We now consider five different cases for \mathcal{H} and j . In most cases we will prove $\psi'(\mathfrak{M}_{\mathcal{H}^2}'', \mathcal{H}^{3'}) = O_{g,k}(z/n)$, from which the desired bound $\psi(\mathfrak{M}_{e,j}, \mathcal{H}^2 \cup \mathcal{H}^3) = O_{g,k}(y)$ follows, using (8.1) and our assumption $zw^{g+1}/n \leq y$. (Case 5 is slightly different.)

Case 1: $\mathcal{H}^3 = \emptyset$ and $\mathcal{H}^2 \neq \emptyset$. Fix any $e' \in \mathcal{H}^2$. By Lemma 8.6(4),

$$\psi'(\mathfrak{M}_{\mathcal{H}^2}'', \mathcal{H}^{3'}) \leq \psi'(\mathfrak{C}_{e,e',j}^{(4)}) = O_{g,k}(z/n).$$

Case 2: $|\mathcal{H}^3| = 1$ and $\mathcal{H}^{3'} \cap \mathcal{N}_e(U_k) \neq \emptyset$. In this case, we need only consider data (\mathcal{S}, T) for which T is the single triangle in $\mathcal{H}^{3'} \cap \mathcal{N}_e(U_k)$. We are assuming that the exceptional case does not hold, meaning $\{e\} \cup \mathcal{H}^2 \not\subseteq E(\mathcal{H}^3)$, so $\mathcal{H}^2 \not\subseteq E(T)$, i.e., \mathcal{H}^2 contains an edge e' not in T . Now, we proceed similarly to the proof of Lemma 8.6(1). We consider two subcases. First, if $j = 4$ then each $\mathcal{S} \in \mathfrak{M}_{\mathcal{H}^2}''$ consists of T and a second triangle $T' \supseteq e'$ satisfying $\text{lev}(T') = k$ (recall that at least one of the triangles in $\mathcal{S} \setminus \{T\}$ must lie within U_k). So, by WS3 applied to T and e' we have

$$\psi'(\mathfrak{M}_{\mathcal{H}^2}'', \mathcal{H}^{3'}) \leq \frac{1}{|U_k|} \cdot z = O(z/n).$$

On the other hand, if $j > 4$ then by Lemma 8.1(1) we have

$$\psi'(\mathfrak{M}_{\mathcal{H}^2}'', \mathcal{H}^{3'}) \leq \sum_{i=0}^k \sum_{\substack{T' \supseteq e': \\ \text{lev}(T')=i}} \frac{1}{|U_i|} \cdot \psi'(\mathfrak{A}_{\{T,T'\},j,j-4}^{(1)}) = O_{g,k}(zn^{j-v^j(\{T,T'\})-(j-4)}) = O_{g,k}(z/n).$$

Case 3: $\mathcal{H}^3 \neq \emptyset$ but $\mathcal{H}^{3'} \cap \mathcal{N}_e(U_k) = \emptyset$. By Lemma 8.6(2) we have

$$\psi'(\mathfrak{M}_{\mathcal{H}^2}'', \mathcal{H}^{3'}) \leq \psi'(\mathfrak{C}_{\mathcal{H}^{3'},e,j}^{(2)}) = O_{g,k}(z/n).$$

Case 4: $\mathcal{H}^{3'} \cap \mathcal{N}_e(U_k) \neq \emptyset$ and $2 \leq |\mathcal{H}^3| \leq j-3$. Any $\mathcal{S} \in \mathfrak{F}_j$ contains at most one triangle in $\mathcal{N}_e(U_k)$, so we may assume $\mathcal{H}^{3'} \cap \mathcal{N}_e(U_k)$ contains a single triangle (and we need only consider data (\mathcal{S}, T) for which that single triangle is T). Then by Lemma 8.1(1) we have

$$\psi'(\mathfrak{M}_{\mathcal{H}^2}'', \mathcal{H}^{3'}) \leq \psi'(\mathfrak{A}_{\mathcal{H}^{3'},j,j-2-|\mathcal{H}^3|}^{(1)}) = O_{g,k}(zn^{j-(|\mathcal{H}^3|+3)-(j-2-|\mathcal{H}^3|)}) = O_{g,k}(z/n).$$

Case 5: $|\mathcal{H}^3| = j-2$. Finally, this case is trivial: we have $\psi'(\mathfrak{M}_{\mathcal{H}^2}'', \mathcal{H}^{3'}) \in \{0, 1\}$ since the only possible element is $\mathcal{S} = \mathcal{H}^3$. Then (8.1) and $|\mathcal{H}^3| = j-2$ demonstrate

$$\psi(\mathfrak{M}_{e,j}, \mathcal{H}^2 \cup \mathcal{H}^3) \leq 3^{j-3} = O_{g,k}(y). \quad \square$$

Finally, the following lemma will be used in Section 10.4, applied at the end of stage k to all $i \geq k+1$.

Lemma 8.9. Fix positive real numbers $z, w \geq 1$ and $p \leq 1$, integers $h \geq 1$ and $4 \leq j \leq g$, and a descending sequence of subsets $V(K_N) = U_0 \supseteq \dots \supseteq U_i \supseteq U_{i+1}$.

- Let \mathfrak{F}_j be a collection of sets of $j-2$ triangles in K_N which is (y, z) -well-spread with respect to the sequence $U_0 \supseteq \dots \supseteq U_i$ (not including U_{i+1}).
- Let \mathcal{W}' be the set of triangles in K_N , and for $a \in \{1, 2\}$ let $\mathcal{W}_a = \{(T, a) : T \in \mathcal{W}'\}$. Let \mathcal{W}_3 be the set of edges in $K_N[U_i]$.
- Let $\bar{\pi}$ be the weight system for $\mathcal{W} = \mathcal{W}_1 \cup \mathcal{W}_2 \cup \mathcal{W}_3$ defined by $\pi_{T,1} = 1/N$ and $\pi_{T,2} = p/|U_{\text{lev}(T)}|$ for each triangle T , and $\pi_e = p$ for each edge e .

(In the last bullet point, and for the rest of the lemma statement and proof, we define $\text{lev}(\cdot)$ with respect to $U_0 \supseteq \dots \supseteq U_i$, without U_{i+1} .)

Let Q be a set of edges in $K_N[U_i]$ spanning $|V(Q)| \leq h$ vertices, and fix $i^* \in \{i, i+1\}$. Let $\mathfrak{R}_{Q,i^*,j}$ be the multiset (of sets of edges and marked triangles) constructed as follows. Consider each $e \in Q$, each $\mathcal{S} \in \mathfrak{F}_j$, and $T \in \mathcal{S} \cap \mathcal{N}_e(U_{i^*})$. Also consider each marking function $\Phi : \mathcal{S} \setminus \{T\} \rightarrow \{1, 2\}$ that gives at least one triangle the mark “2”. For each such $(e, \mathcal{S}, T, \Phi)$, let u be the single vertex in $T \setminus e$ and add to $\mathfrak{R}_{Q,i^*,j}$ a copy of

$$\{uv \in \mathcal{W}_3 : v \in V(Q)\} \cup \{(T', \Phi(T')) : T' \in \mathcal{S} \setminus \{T\}\}.$$

If $zp^{-h-1}/|U_{i^*}| \leq y$ then $\kappa(\mathfrak{R}_{Q,i^*,j}) = O_{g,i,h}(yp^{V(Q)+1}|U_{i^*}|)$.

Proof. Let \mathfrak{R}'_e be the submultiset of $\mathfrak{R}_{Q,i^*,j}$ corresponding to a particular choice of e . There are fewer than $h^2 = O_h(1)$ choices of e , so it suffices to show that $\kappa(\mathfrak{R}'_e) = O_{g,i}(yp^{V(Q)+1}|U_{i^*}|)$. We proceed in a similar way to the proofs of [Lemmas 8.4, 8.7, and 8.8](#). Let $\bar{\pi}'$ be the weight system on the ground set \mathcal{W}' where we let $\pi'_{T'} = 1/|U_{\text{lev}(T')}|$ for each $T' \in \mathcal{W}'$, and write ψ' instead of $\psi^{(\bar{\pi}'')}$.

First, by [Lemma 8.1\(1\)](#), we have

$$\psi(\mathfrak{R}'_e, \emptyset) \leq 2^{j-3} \cdot p \cdot p^{|V(Q)|} \cdot \sum_{T \in \mathcal{N}_e(U_{i^*})} \psi'(\mathfrak{A}_{\{T\},j,j-3}) = O_{g,i}(|U_{i^*}|p^{V(Q)+1}y).$$

(The factor of p comes from the fact that we are only considering (\mathcal{S}, T, Φ) for which at least one triangle in $\mathcal{S} \setminus \{T\}$ has the mark “2”.)

Now, fix $\mathcal{H}^2 \subseteq \mathcal{W}_3$ and $\mathcal{H}^3 \subseteq \mathcal{W}_1 \cup \mathcal{W}_2$ such that $\mathcal{H}^2 \cup \mathcal{H}^3 \neq \emptyset$. Let $\mathcal{H}^{3'}$ be the set of unmarked triangles underlying \mathcal{H}^3 , and let $\mathfrak{R}''_{e,\mathcal{H}^2}$ be obtained by including a copy of $\mathcal{S} \setminus \{T\}$ for every \mathcal{S} and $T \in \mathcal{S} \cap \mathcal{N}_e(U_{i^*})$ such that all edges in \mathcal{H}^2 contain the single vertex in $T \setminus \{e\}$. We have $\psi(\mathfrak{R}'_e, \mathcal{H}^2 \cup \mathcal{H}^3) \leq 2^{j-3} \psi'(\mathfrak{R}''_{e,\mathcal{H}^2}, \mathcal{H}^{3'})$. Recalling our assumption $zp^{-h-1}/|U_{i^*}| \leq y$, it suffices to prove that $\psi'(\mathfrak{R}''_{e,\mathcal{H}^2}, \mathcal{H}^{3'}) = O_{g,i}(z)$.

Case 1: $\mathcal{H}^3 \neq \emptyset$. In this case, [Lemma 8.6\(1\)](#) implies

$$\psi'(\mathfrak{R}''_{e,\mathcal{H}^2}, \mathcal{H}^{3'}) \leq \psi'(\mathfrak{C}_{\mathcal{H}^{3'},e,j}^{(1)}) = O_{g,i}(z).$$

Case 2: $\mathcal{H}^3 = \emptyset$ and $\mathcal{H}^2 \neq \emptyset$. Fix an edge $e'' \in \mathcal{H}^2$. Each set in $\mathfrak{R}''_{e,\mathcal{H}^2}$ is specified by a pair (\mathcal{S}, T) , where T is a triangle containing e and another vertex $u \in e''$. That is to say, there are at most two possibilities for T . For each such T , let \mathfrak{R}'''_T be the submultiset of $\mathfrak{R}''_{e,\mathcal{H}^2}$ corresponding to that choice of T , let e' be one of the two edges of $E(T) \setminus e$, and observe that

$$\psi'(\mathfrak{R}'''_T, \mathcal{H}^{3'}) \leq \psi'(\mathfrak{C}_{e',e'',j}^{(3)}) = O_{g,i}(z)$$

by [Lemma 8.6\(3\)](#). □

9. A GENERALIZED HIGH-GIRTH TRIPLE PROCESS

In this section we study a generalization of the high-girth triple process studied in [\[8, 27\]](#). We show that in quite a general setting it is possible to find a set of edge-disjoint triangles avoiding a set of “forbidden configurations”. Where possible, we make some effort to use similar notation as in [\[27\]](#).

As sketched in [Section 2](#), it is important that at every stage of the main iteration, we “remember the randomness” of prior stages. So, for several of the lemmas in this paper, one of the inputs will be a random set of triangles coming from the results of the iteration so far. The following definition captures the properties we need from this random set of triangles: they are consistent with having arisen from the iteration so far, and in particular the probability that any set of triangles has been chosen can be approximated by a product of weights. This will be crucial in order to apply the estimates in [Section 8](#). Later, in the proof of [Theorem 1.1](#) we will need a more refined version of this definition ([Definition 10.2](#)).

Definition 9.1 (Well-distributedness). Fix a descending sequence of subsets $V(K_N) = U_0 \supseteq \dots \supseteq U_k$, and an “error factor” $w \geq 1$. Say a random set of triangles \mathcal{D} in K_N is (w, b) -well-distributed (with respect to our descending sequence of subsets) if for any $s \in \mathbb{N}$ and distinct triangles D_1, \dots, D_s we have

$$\Pr[D_1, \dots, D_s \in \mathcal{D}] \leq w^s \left(\prod_{i=1}^s \frac{1}{|U_{\text{lev}(D_i)}|} + b \right).$$

We will always take $b = n^{-\omega(1)}$ to be super-polynomially small in $n = |U_k|$, so the condition in [Definition 9.1](#) will be meaningful for all bounded $s = O(1)$.

We are nearly ready to state the main theorem in this section. It is convenient to separately define the assumptions of this theorem, as we will want to refer to them multiple times.

Definition 9.2 (Goodness). Consider a descending sequence of subsets $V(K_N) = U_0 \supseteq \dots \supseteq U_k$, and consider collections $\mathfrak{F}_4^{\text{sup}}, \dots, \mathfrak{F}_g^{\text{sup}}$ of triangles in K_N , where each $\mathcal{S} \in \mathfrak{F}_j^{\text{sup}}$ consists of $j - 2$ edge-disjoint triangles. We also consider some initial data (which are themselves allowed to be random):

- let \mathcal{D} be a random “already chosen” set of edge-disjoint triangles in K_N ;
- let $G \subseteq K_N[U_k] \cong K_n$ be a random “remainder” graph (the graph which we wish to decompose into triangles), with edge set E disjoint from $E(\mathcal{D})$;
- let \mathcal{A} be a (random) subset of the triangles in G (the “available” triangles which we may use in our triangle-decomposition);

- let $\mathfrak{J}_4, \dots, \mathfrak{J}_g$ be random collections of sets of triangles in \mathcal{A} (the “forbidden configurations” we would like to avoid), such that for each $\mathcal{S} \in \bigcup_{j=4}^g \mathfrak{J}_j$ there is $\mathcal{S}' \in \bigcup_{j=4}^g \mathfrak{F}_j^{\text{sup}}$ with $\mathcal{S} \subseteq \mathcal{S}'$ and $\mathcal{S}' \setminus \mathcal{S} \subseteq \mathcal{D}$. (That is to say, each of our forbidden configurations is a set of triangles which together with some triangles in \mathcal{D} forms one of the configurations in $\bigcup_{j=4}^g \mathfrak{F}_j^{\text{sup}}$.)

We say that all the above data (namely, the sets U_i , the collections of triangles $\mathfrak{F}_j^{\text{sup}}$, and the random objects $\mathcal{D}, G, \mathcal{A}, \mathfrak{J}_4, \dots, \mathfrak{J}_g$) are (C, β, b) -good if:

- each $\mathfrak{F}_j^{\text{sup}}$ is n^β -well-spread ([Definition 7.1](#));
- \mathcal{D} is (n^β, b) -well-distributed ([Definition 9.1](#));
- with probability 1:
 - every $e \in E$ is in $(1 \pm n^{-1/C})3|\mathcal{A}|/|E|$ triangles of \mathcal{A} ;
 - every $T \in \mathcal{A}$ is contained in $(j-2)|\mathfrak{J}_j|/|\mathcal{A}| \pm n^{-1/C}|\mathcal{A}|^{j-3}/|E|^{j-3}$ of the configurations in \mathfrak{J}_j ;
 - $|\mathcal{A}| \geq n^{1-\beta}|E|$;
 - $|E| \geq n^{2-\beta}$;
 - $|\mathfrak{J}_j| \leq C|\mathcal{A}|^{j-2}/|E|^{j-3}$;
 - There is no $\mathcal{S} \in \bigcup_{j=4}^g \mathfrak{J}_j$ which is a subset of another $\mathcal{S}' \in \bigcup_{j=4}^g \mathfrak{J}_j$.

We remark that the inequality $|E| \geq n^{2-\beta}$ can actually be derived from the inequality $|\mathcal{A}| \geq n^{1-\beta}|E|$, so the former inequality is not strictly speaking necessary (though we will not actually need this deduction anywhere in the paper).

The role of the collections $\mathfrak{F}_j^{\text{sup}}$ is that, while the collections of forbidden configurations \mathfrak{J}_j are allowed to depend on the random data, they are always “induced” from the collections $\mathfrak{F}_j^{\text{sup}}$, which are well-spread and do not depend on the random data. In our proof of [Theorem 1.1](#) we will take $\mathfrak{F}_j^{\text{sup}}$ to contain all Erdős configurations, all forbidden configurations induced by absorber triangles (see [Lemma 7.2](#)), and in addition some random sets of triangles that will be used for “regularization”.

Now, we present the main theorem in this section. In addition to guaranteeing the existence of an almost-triangle-decomposition, we require the existence of a probability distribution over such triangle-decompositions, satisfying certain properties; we will “remember the randomness” of this distribution for future stages of the main iterations.

Theorem 9.3. *Fix a constant $g \in \mathbb{N}$. For $C > 0$ there are $\beta = \beta_{9.3}(g, C) > 0$ and $\alpha = \alpha_{9.3}(g, C) > 0$ such that the following holds. Consider (C, β, b) -good data, as defined in [Definition 9.2](#). Define $h(b)$ by setting $h(b) = \sqrt{b^\alpha + \exp(-n^\alpha)}$ if $b \leq N^{-1/\beta}$, and $h(b) = 1$ otherwise. There is a random set of edge-disjoint triangles $\mathcal{M} \subseteq \mathcal{A}$ (depending on our data) such that:*

- no $\mathcal{S} \in \bigcup_j \mathfrak{J}_j$ is fully included in \mathcal{M} , and
- with probability at least $1 - h(b)$ over the randomness of the initial data $\mathcal{D}, G, \mathcal{A}, \mathfrak{J}_4, \dots, \mathfrak{J}_g$: for any $s_1, s_2 \geq 0$ and any triangles $D_1, \dots, D_{s_1} \in \mathcal{A}$ and edges $e_1, \dots, e_{s_2} \in E$ we have

$$\Pr[D_1, \dots, D_{s_1} \in \mathcal{M} \text{ and } e_1, \dots, e_{s_2} \notin E(\mathcal{M}) \mid \mathcal{D}, G, \mathcal{A}, \mathfrak{J}_4, \dots, \mathfrak{J}_g] \leq (O_{g,C}(|E|/|\mathcal{A}|))^{s_1} (O(n^{-\beta}))^{s_2} + 2h(b).$$

The particular form of the error term $h(b)$ is not very important; all that is important for the proof of [Theorem 1.1](#) is that if $b = n^{-\omega_{g,C}(1)}$ then $h(b) = n^{-\omega_{g,C}(1)}$ as well. We prove [Theorem 9.3](#) via analysis of a random process, as follows.

Definition 9.4 (Generalized high-girth triple process). Let G be an n -vertex graph with edge set $E = E(0)$, let $\mathcal{A} = \mathcal{A}(0)$ be a set of triangles in G , and for $4 \leq j \leq g$ let \mathfrak{J}_j be a collection of sets of $j-2$ edge-disjoint triangles which are forbidden from being used together¹¹. All these objects (“initial data”) may be random.

Now, consider the following random process (using fresh independent randomness). Let $\mathcal{C}(0) = \emptyset$ and let $t = 0$. While $\mathcal{A}(t) \neq \emptyset$:

- (1) choose a uniformly random triangle $T^*(t) \in \mathcal{A}(t)$;
- (2) let $\mathcal{C}(t+1) = \mathcal{C}(t) \cup \{T^*(t)\}$ (add $T^*(t)$ to the set of “chosen triangles”);
- (3) let $\mathcal{A}(t+1)$ be the set of triangles in $\mathcal{A}(t)$ which do not share an edge with $T^*(t)$ and would not complete any $\mathcal{E} \in \bigcup_{j=4}^g \mathfrak{J}_j$ when combined with the triangles in $\mathcal{C}(t+1)$;
- (4) increment t to $t+1$.

We also define some variables that track the evolution of the process:

¹¹The reason for our perhaps unnatural-seeming indexing is that, in practice, the configurations in \mathfrak{J}_j will “behave like” Erdős j -configurations, in the sense of [Definition 7.1](#). This indexing convention also makes our analysis more consistent with the analysis in [\[27\]](#). However, one important difference between our analysis and the analysis in [\[27\]](#) is that in [\[27\]](#), the set \mathfrak{J}_4 has a very different role: it is the set of *diamonds* (pairs of triangles that share an edge). Recall that our forbidden configurations always consist of edge-disjoint triangles, and we separately enforce an edge-disjointness condition in [Definition 9.4](#).

- Let $E(t) \subseteq E(0)$ be the set of edges not appearing in any triangle in $\mathcal{C}(t)$ (i.e., the set of “uncovered” edges). Note that $|E(t)| = |E(0)| - 3t$.
- For some step t and an edge $e \in E(t)$, let $\mathcal{X}_e(t)$ be the set of triangles in $\mathcal{A}(t)$ which contain the edge e .
- For a triangle $T \in \mathcal{A}(t)$ let $\mathcal{T}_T(t) \subseteq \mathcal{A}(t)$ be the set of triangles $T^* \neq T$ that “threaten” T , in the sense that choosing T^* as the next triangle in the process would cause $T \notin \mathcal{A}(t+1)$ (i.e., $E(T) \cap E(T^*) \neq \emptyset$ or $\{T, T^*\} \cup \mathcal{C}(t)$ contains some element of $\bigcup_{j=4}^g \mathfrak{J}_j$).

We also define some functions that describe the typical behavior of the process. (They are defined in terms of the initial data, so if those data are random, then the following functions are random as well.)

Definition 9.5 (Trajectories). Given $|E(0)|, |\mathcal{A}(0)|, |\mathfrak{J}_4|, \dots, |\mathfrak{J}_g|$, we define

$$p(t) = \frac{|E(0)| - 3t}{|E(0)|}, \quad \rho(t) = \sum_{j=4}^g \frac{(j-2)t^{j-3}|\mathfrak{J}_j|}{|\mathcal{A}(0)|^{j-2}},$$

$$f_{j,c}(t) = \binom{j-3}{c} \left(\frac{t}{|\mathcal{A}(0)|} \right)^c (p(t)^3 e^{-\rho(t)})^{j-3-c} \left(\frac{(j-2)|\mathfrak{J}_j|}{|\mathcal{A}(0)|} \right) \text{ for } 4 \leq j \leq g, 0 \leq c \leq j-4,$$

$$f_{\text{edge}}(t) = p(t)^2 e^{-\rho(t)} \left(\frac{3|\mathcal{A}(0)|}{|E(0)|} \right), \quad f_{\text{threat}}(t) = 3f_{\text{edge}}(t) + \sum_{j=4}^g f_{j,j-4}(t).$$

Remark. The reader can interpret $p(t)$ as the fraction of the graph G which has not yet been covered at time t . As we will see in the next result, $f_{\text{edge}}(t)$ and $f_{\text{threat}}(t)$ describe the trajectories of the random processes $|\mathcal{X}_e(t)|$ and $|\mathcal{T}_T(t)|$, respectively. We can also interpret $f_{j,c}(t)$ as describing the trajectory of a quantity that evolves with the process: namely, it describes the approximate number of forbidden configurations $\mathcal{S} \in \mathfrak{J}_j$, containing a fixed triangle T , such that c of the other $j-3$ triangles in \mathcal{S} have already been chosen, and the other $j-3-c$ triangles are still available. Finally, $\rho(t)$ is a slightly less intuitive parameter which captures the Poissonian rate at which triangles become unavailable due to forbidden configurations.

Roughly speaking, for a triangle $T \in \mathcal{A}(0)$, the probability that its three edges are in $E(t)$ is about $p(t)^3$, and given this, the probability T is in $\mathcal{A}(t)$ is about $e^{-\rho(t)}$.

We will prove the following theorem.

Theorem 9.6. Fix a constant g . For $C > 0$ there are $\beta = \beta_{9.6}(g, C) > 0$ and $\alpha = \alpha_{9.6}(g, C) > 0$ such that the following holds. Fix (C, β, b) -good data as in [Definition 9.2](#) with $b \leq N^{-1/\beta}$. Then with probability at least $1 - b^\alpha - \exp(-n^\alpha)$ (over all randomness, including that of the initial data): the process in [Definition 9.4](#) runs for $\tau_{\text{cut}} := \lceil (1 - n^{-\beta})|E|/3 \rceil$ steps without terminating, and moreover for each $0 \leq t \leq \tau_{\text{cut}}$, in the notation of [Definition 9.5](#),

- $|\mathcal{X}_e(t) - f_{\text{edge}}(t)| \leq n^{-\beta} f_{\text{edge}}(t)$ for each $e \in E(t)$;
- $|\mathcal{T}_T(t) - f_{\text{threat}}(t)| \leq n^{-\beta} f_{\text{threat}}(t)$ for each $T \in \mathcal{A}(t)$;
- $|\mathcal{T}_T(t) \cap \mathcal{T}_{T'}(t)| \leq n^{1/2}$ for each pair of edge-disjoint $T, T' \in \mathcal{A}(t)$;
- $|\mathcal{T}_T(t) \cap \mathcal{X}_e(t)| \leq n^{1/2}$ for each $T \in \mathcal{A}(t)$ and $e \in E(t)$ with $e \not\subseteq T$.

Before proving [Theorem 9.6](#), we deduce [Theorem 9.3](#) from it.

Proof of [Theorem 9.3](#). We use the same α, β as in [Theorem 9.6](#), except that we additionally enforce $\alpha < \beta/3$. That is, we take $\beta_{9.3}(g, C) = \beta_{9.6}(g, C)$ and $\alpha_{9.3}(g, C) = \min\{\alpha_{9.6}(g, C), \beta_{9.6}(g, C)/4\}$.

Let \mathbf{E} be the event that the conclusion of [Theorem 9.6](#) holds, so $\Pr[\mathbf{E}] \geq 1 - h(b)^2$. Fix an outcome of the initial data $\mathcal{D}, G, \mathcal{A}, \mathfrak{J}_4, \dots, \mathfrak{J}_g$ such that $\Pr[\mathbf{E} \mid \mathcal{D}, G, \mathcal{A}, \mathfrak{J}_4, \dots, \mathfrak{J}_g] \geq 1 - h(b)$ (the initial data satisfies this property with probability $1 - h(b)$ by Markov’s inequality). For all probabilities in the rest of this proof, we implicitly condition on our fixed outcome of $\mathcal{D}, G, \mathcal{A}, \mathfrak{J}_4, \dots, \mathfrak{J}_g$.

For this proof, we need to consider the random set of triangles obtained by running the process in [Definition 9.4](#) for τ_{cut} steps. However, due to an extension to [Theorem 9.3](#) we will need to prove later, it will be helpful to consider the more general case where we fix $\tau \leq \tau_{\text{cut}}$ and consider the result of running the process for τ steps. Let \mathcal{M} be the corresponding random set of triangles. Fix s_1, s_2 , distinct triangles D_1, \dots, D_{s_1} and distinct edges e_1, \dots, e_{s_2} . Define

$$P(D_1, \dots, D_{s_1}; e_1, \dots, e_{s_2}) = \Pr[D_1, \dots, D_{s_1} \in \mathcal{M} \text{ and } e_1, \dots, e_{s_2} \notin E(\mathcal{M}) \mid \mathcal{D}, G, \mathcal{A}, \mathfrak{J}_4, \dots, \mathfrak{J}_g].$$

We wish to prove that $P(D_1, \dots, D_{s_1}; e_1, \dots, e_{s_2}) \leq (O_{g,C}(|E|/|\mathcal{A}|))^{s_1} (O(p(\tau)))^{s_2} + 2h(b)$ (note that $p(\tau_{\text{cut}}) \approx n^{-\beta}$).

Before proceeding, we observe that it essentially suffices to consider the case where s_1, s_2 are quite small. Indeed, let $r_1 = \min\{s_1, \lfloor n^{\beta/3} \rfloor\}$ and $r_2 = \min\{s_2, \lfloor n^{\beta/3} \rfloor\}$. If we can prove that

$$P(D_1, \dots, D_{r_1}; e_1, \dots, e_{r_2}) \leq (O_{g,C}(|E|/|\mathcal{A}|))^{r_1} (O(p(\tau)))^{r_2} + h(b) \quad (9.1)$$

it would follow that

$$\begin{aligned} P(D_1, \dots, D_{s_1}; e_1, \dots, e_{s_2}) &\leq P(D_1, \dots, D_{r_1}; e_1, \dots, e_{r_2}) \leq (O_{g,C}(|E|/|\mathcal{A}|))^{r_1} (O(p(\tau)))^{r_2} + h(b) \\ &\leq 2^{s_2} (O_{g,C}(|E|/|\mathcal{A}|))^{s_1} (O(p(\tau)))^{s_2} + 2h(b), \end{aligned}$$

as desired (since $\alpha < \beta/3$).

So, it suffices to prove (9.1). Fix steps $t_1, \dots, t_{r_1} \leq \tau_{\text{cut}}$. We will estimate the probability that $T^*(t_1) = D_1, \dots, T^*(t_{r_1}) = D_{r_1}$ and $e_1, \dots, e_{r_2} \notin E(\mathcal{M})$; we will then sum this expression over choices of t_1, \dots, t_{r_1} . Note that we may assume that no e_i appears in $D_1 \cup \dots \cup D_{r_1}$, and in fact the D_j are edge-disjoint (as otherwise the relevant probability is zero).

First, note that (as long as the process does not terminate before time t) we always have $|E(t)| = p(t)|E(0)|$, so if the conclusion of Theorem 9.6 holds then for each $t \leq \tau$ we have $|\mathcal{A}(t)| = (1 \pm n^{-\beta})A(t)$, where $A(t) = p(t)|E(0)|f_{\text{edge}}(t)/3 = p(t)^3 e^{-\rho(t)}|\mathcal{A}(0)|$.

Next, we observe that for any $e \neq e'$ we always have $|\mathcal{X}_e \cap \mathcal{X}_{e'}| \leq 1 \leq \sqrt{n}$, so if the conclusion of Theorem 9.6 holds, then for each $t \leq \tau$ we have

$$\begin{aligned} \left| \left(\bigcup_{i:t < t_i} \mathcal{T}_{D_i}(t) \right) \cup \left(\bigcup_{i=1}^{r_2} \mathcal{X}_{e_i}(t) \right) \right| &\geq \sum_{i:t < t_i} |\mathcal{T}_{D_i}(t)| + \sum_{i=1}^{r_2} |\mathcal{X}_{e_i}| - \binom{r_1 + r_2}{2} \sqrt{n} \\ &\geq |\{i : t < t_i\}| (1 \pm n^{-\beta}) f_{\text{threat}}(t) + r_2 (1 \pm n^{-\beta}) f_{\text{edge}}(t) - n^{1-5\beta} \\ &\geq (1 - 2n^{-\beta})(r_2 + 3|\{i : t < t_i\}|) f_{\text{edge}}(t), \end{aligned}$$

provided β is sufficiently small (here we write $\{i : t < t_i\}$ for the set of all $i \in \{1, \dots, r_1\}$ such that $t < t_i$). In the last inequality, we implicitly used $f_{\text{threat}}(t) \geq 3f_{\text{edge}}(t) = \Omega_{g,C}(n^{1-3\beta})$ which follows from the given bounds on the initial data.

Now, to have $T^*(t_i) = D_i$, at each step $t < t_i$ we must choose a triangle not threatening D_i (i.e., not in $\mathcal{T}_{D_i}(t)$), out of a total of $|\mathcal{A}(t)|$ possibilities. To have $e_i \notin E(\mathcal{M})$, at each step t we must choose a triangle not in $\mathcal{X}_{e_i}(t)$. It follows that

$$\begin{aligned} &\Pr[T^*(t_1) = D_1, \dots, T^*(t_{r_1}) = D_{r_1} \text{ and } e_1, \dots, e_{r_2} \notin E(\mathcal{M}) \text{ and } \mathbf{E} \text{ holds}] \\ &\leq \prod_{t \in \{0, \dots, \tau\} \setminus \{t_1, \dots, t_{r_1}\}} \left(1 - \frac{(1 - 2n^{-\beta})(r_2 + 3|\{i : t < t_i\}|) f_{\text{edge}}(t)}{(1 + n^{-\beta})A(t)} \right) \prod_{i=1}^{r_1} \frac{1}{(1 - n^{-\beta})A(t_i)}. \end{aligned} \quad (9.2)$$

We then compute

$$\sum_{t=0}^{\tau} \frac{f_{\text{edge}}(t)}{A(t)} = \sum_{t=0}^{\tau} \frac{1}{|E(0)|/3 - t} = \log \left(\frac{|E(0)|/3}{|E(0)|/3 - \tau} \right) + O(1) = \log(1/p(\tau)) + O(1)$$

and similarly

$$\sum_{t=0}^{\tau} \frac{|\{i : t < t_i\}| f_{\text{edge}}(t)}{A(t)} = \sum_{i=1}^{r_1} (\log(1/p(t_i)) + O(1)).$$

Also, we have $f_{\text{edge}}(t_1)/A(t_1) + \dots + f_{\text{edge}}(t_{r_1})/A(t_{r_1}) \leq r_1 n^\beta / |E(0)| \leq 1/\sqrt{n}$ for small β . Using the inequality $1 - x \leq e^{-x}$, we deduce that

$$\Pr[T^*(t_1) = D_1, \dots, T^*(t_{r_1}) = D_{r_1} \text{ and } e_1, \dots, e_{r_2} \notin E(\mathcal{M})] \leq (O(p(\tau)))^{r_2} \prod_{i=1}^{r_1} O\left(\frac{e^{\rho(t_i)}}{|\mathcal{A}|}\right) + h(b).$$

Noting that $\rho(t) = O_{g,C}(1)$ for all t , and summing over all $O(|E|^{r_1})$ choices of t_1, \dots, t_{r_1} , we obtain the desired inequality (9.1). \square

Now we prove Theorem 9.6. An important part of this is to carefully track the evolution of the sets $\mathcal{X}_e(t)$. In order to study how these random sets evolve, there is one other class of random variables we will need to track.

Definition 9.7. Recall the process and notation defined in Definition 9.4. For some step t and a triangle $T \in \mathcal{A}(t)$, and some $4 \leq j \leq g$ and $0 \leq c \leq j - 4$, let $\mathfrak{X}_{T,j,c}(t)$ be the set of $\mathcal{E} \in \mathfrak{J}_j$ such that $T \in \mathcal{E}$, such that c triangles of \mathcal{E} are already in $\mathcal{C}(t)$, and such that the remaining $j - 3 - c$ triangles are in $\mathcal{A}(t)$.

The configurations in $\mathfrak{X}_{T,j,j-4}(t)$ are especially important because they specify a pair of triangles $\{T, T^*\}$ such that choosing T^* next would make T unavailable (T^* threatens T). The idea is that each $|\mathcal{X}_e(t)|$ tends to be close to $f_{\text{edge}}(t)$, and each $|\mathfrak{X}_{T,j,c}(t)|$ tends to be close to $f_{j,c}(t)$. We define some hitting times that measure the first time we leave these trajectories.

Definition 9.8. In the setting of [Theorem 9.6](#), let B be a large constant (depending on g and C ; large enough to satisfy a certain inequality that will arise in the proof of [Theorem 9.6](#)). Recalling the notation $p(t) = (|E(0)| - 3t)/|E(0)|$ in [Definition 9.5](#), we define “error thresholds”¹²

- $e_{\text{edge}}(t) = p(t)^{-B} n^{1-1/(2C)}$;
- $e_{j,c}(t) = e_{\text{edge}}(t)(p(t)^2 |\mathcal{A}(0)| / |E(0)|)^{j-4-c}$.

Now, recall the functions $f_{\text{edge}}, f_{j,c}$ defined in [Definition 9.5](#):

$$p(t) = \frac{|E(0)| - 3t}{|E(0)|}, \quad \rho(t) = \sum_{j=4}^g \frac{(j-2)t^{j-3} |\mathfrak{J}_j|}{|\mathcal{A}(0)|^{j-2}},$$

$$f_{j,c}(t) = \binom{j-3}{c} \left(\frac{t}{|\mathcal{A}(0)|} \right)^c (p(t)^3 e^{-\rho(t)})^{j-3-c} \left(\frac{(j-2) |\mathfrak{J}_j|}{|\mathcal{A}(0)|} \right) \text{ for } 4 \leq j \leq g, \ 0 \leq c \leq j-4,$$

$$f_{\text{edge}}(t) = p(t)^2 e^{-\rho(t)} \left(\frac{3|\mathcal{A}(0)|}{|E(0)|} \right), \quad f_{\text{threat}}(t) = 3f_{\text{edge}}(t) + \sum_{j=4}^g f_{j,j-4}(t).$$

and also recall the random sets $\mathcal{X}_e(t), \mathfrak{X}_{T,j,c}(t)$ defined in [Definitions 9.4](#) and [9.7](#). We then define τ_{traj} to be the first t for which

$$|\mathcal{X}_e(t) - f_{\text{edge}}(t)| > e_{\text{edge}}(t) \quad \text{or} \quad |\mathfrak{X}_{T,j,c}(t) - f_{j,c}(t)| > e_{j,c}(t)$$

for some $e \in E(t)$ or some j, c and some $T \in \mathcal{A}(t)$ (i.e., the point when one of our statistics leaves our predicted trajectory)¹³. If there is no t for which this happens, we write $\tau_{\text{traj}} = \infty$.

Our main goal is to prove that $\tau_{\text{traj}} \geq \tau_{\text{cut}} = \lceil (1-n^{-\beta})|E|/3 \rceil$ (i.e., that our statistics stay on our predicted trajectory until our desired cutoff point). We will prove this with a martingale concentration inequality and the fact that $|\mathcal{X}_e| - f_{\text{edge}}$ and $|\mathfrak{X}_{T,j,c}| - f_{j,c}$ are each approximately martingales, with bounded differences. To make this rigorous, we will need crude bounds on certain auxiliary statistics.

Definition 9.9. Using the process and notation defined in [Definition 9.4](#), we make the following additional definitions.

- For distinct $T, T' \in \mathcal{A}(t)$, and any $4 \leq j \leq g$ and $0 \leq c \leq j-5$, let $Z_{T,T',j,c}(t)$ be the number of $\mathcal{E} \in \mathfrak{X}_{T,j,c}(t)$ with $T' \in \mathcal{E}$.
- For $e \in E(t)$ and $T \in \mathcal{A}(t)$ with $e \not\subseteq T$, let $Z_{e,T}(t)$ be the number of $\mathcal{E} \in \bigcup_{j=4}^g \mathfrak{J}_j$ such that $\mathcal{E} \cap \mathcal{A}(t) = \{T, T'\}$ for some triangle T' containing e , and $\mathcal{E} \setminus \mathcal{A}(t) \subseteq \mathcal{C}(t)$.
- For (not necessarily distinct) $T, T' \in \mathcal{A}(t)$, let $Z_{T,T'}(t)$ be the number of distinct $\mathcal{E}, \mathcal{E}' \in \bigcup_{j=4}^g \mathfrak{J}_j$ such that $\mathcal{E} \cap \mathcal{A}(t) = \{T, T^*\}$ and $\mathcal{E}' \cap \mathcal{A}(t) = \{T', T^*\}$ for some $T^* \in \mathcal{A}(t) \setminus \{T, T'\}$, and $(\mathcal{E} \cup \mathcal{E}') \setminus \mathcal{A}(t) \subseteq \mathcal{C}(t)$ (that is, all other triangles in $\mathcal{E}, \mathcal{E}'$ have been chosen during the process). This implies that T^* threatens both T and T' (or if $T = T'$, that it threatens in more than one way).
- For $T \in \mathcal{A}(t)$, some $4 \leq j \leq g$ and $1 \leq c \leq j-4$, let $Z_{T,j,c-1}(t)$ be the number of distinct pairs $\mathcal{E}, \mathcal{E}'$ with $\mathcal{E} \in \mathfrak{X}_{T,j,c-1}(t)$ and $\mathcal{E}' \in \bigcup_{j'=5}^g \mathfrak{J}_{j'}$, such that $|\mathcal{E}' \cap \mathcal{A}(t)| = 2$, $\mathcal{E}' \cap \mathcal{A}(t) \subseteq \mathcal{E} \cap \mathcal{A}(t)$, and $(\mathcal{E} \cup \mathcal{E}') \setminus \mathcal{A}(t) \subseteq \mathcal{C}(t)$ (in such a case \mathcal{E} would actually be redundant; avoiding \mathcal{E}' is technically a stronger constraint than avoiding \mathcal{E} now).

Lemma 9.10. *In the setting of [Theorem 9.6](#), if β, α are sufficiently small then with probability $1 - b^\alpha - \exp(-n^\alpha)$, for each $t < \tau_{\text{traj}} \wedge \tau_{\text{cut}}$, with the definitions in [Definition 9.9](#) we have*

- (1) For $4 \leq j \leq g$ and $0 \leq c \leq j-5$, $Z_{T,T',j,c}(t) \leq n^{j-c-5+100g\beta}$.
- (2) $Z_{e,T}(t) \leq n^{100g\beta}$ for every $e \in E(t)$ and $T \in \mathcal{A}(t)$.
- (3) $Z_{T,T'}(t) \leq n^{100g\beta}$ for (not necessarily distinct) $T, T' \in \mathcal{A}(t)$.
- (4) $Z_{T,j,c-1}(t) \leq n^{j-c-3+100g\beta}$ for each $4 \leq j \leq g$, $1 \leq c \leq j-4$, and each $T \in \mathcal{A}(t)$.

Proof. If β is sufficiently small in terms of B, C then it follows $e_{\text{edge}}(t)/f_{\text{edge}}(t) \leq 1/2$ for all $t < \tau_{\text{traj}} \wedge \tau_{\text{cut}}$. Therefore for such t ,

$$|\mathcal{A}(t)| = \frac{1}{3} \sum_{e \in E(t)} |\mathcal{X}_e(t)| = \Omega((n^{-\beta}|E|) \cdot f_{\text{edge}}(\tau_{\text{cut}})) = \Omega(n^{-4\beta} |\mathcal{A}|).$$

¹²These error thresholds are completely different from those in [\[27\]](#), and similar to those in [\[8\]](#).

¹³This deviates slightly from the notation in [\[27\]](#) (the authors of that paper write τ_{violated} for something similar but not exactly the same).

We immediately deduce that $\mathcal{C}(\tau_{\text{traj}} \wedge \tau_{\text{cut}})$ contains s specific triangles with probability $(O(n^{4\beta}|E|/|\mathcal{A}|))^s$. Using $|\mathcal{A}| \geq n^{1-\beta}|E|$, we see this set of triangles is $(O(n^{5\beta}), 0)$ -well-distributed.

Recall that the randomness of our generalized high-girth triple process is conditional on the randomness of the initial data (including \mathcal{D}), and recall that we are assuming \mathcal{D} is (n^β, b) -well-distributed. It follows that $\mathcal{D} \cup \mathcal{C}(\tau_{\text{traj}} \wedge \tau_{\text{cut}})$ is $(n^{6\beta}, b)$ -well-distributed (for any particular set of s triangles, we sum over all 2^s ways to designate those triangles as appearing in \mathcal{D} or in $\mathcal{C}(\tau_{\text{traj}})$).

Given a multiset \mathfrak{K} of sets of triangles, write $X(\mathfrak{K})$ for the random number of sets in \mathfrak{K} that are fully included in $\mathcal{D} \cup \mathcal{C}(\tau_{\text{traj}} \wedge \tau_{\text{cut}})$. Let $\mathfrak{K}_{T,T',j+r,c+r}^{(1)}$, etc., be defined as in Lemma 8.2 with subset sequence $U_0 \supseteq \dots \supseteq U_k$ and with $\mathfrak{J}_j = \mathfrak{J}_j^{\text{sup}}$ for $4 \leq j \leq g$. The desired bounds now follow from Lemmas 3.7 and 8.2 (with say $s = \min(\frac{1}{10} \log_n(1/b), n^\beta)$ in Lemma 3.7), provided α is sufficiently small relative to β :

- (1) Note that $Z_{T,T',j,c}(t) \leq \sum_{r=0}^{g-j} X(\mathfrak{K}_{T,T',j+r,c+r}^{(1)})$ for $T, T' \in \mathcal{A}(t)$ (since \mathfrak{J}_j is “induced” from $\bigcup_{j'=4}^g \mathfrak{J}_{j'}^{\text{sup}}$). (1) then follows from Lemma 3.7 and Lemma 8.2(1), and a union bound over $O_g(n^6)$ choices of T, T', j, c .
- (2) Note that $Z_{e,T}(t) \leq X(\mathfrak{K}_{e,T}^{(2)})$ for $e \in E(t)$ and $T \in \mathcal{A}(t)$. (2) then follows from Lemma 3.7 and Lemma 8.2(2), and a union bound over $O(n^5)$ choices of e, T .
- (3) Note that $Z_{T,T'}(t) \leq X(\mathfrak{K}_{T,T'}^{(3)})$ for $T, T' \in \mathcal{A}(t)$, due to the relation between \mathfrak{J}_j and $\mathfrak{J}_j^{\text{sup}}$. (3) then follows from Lemma 3.7 and Lemma 8.2(3), and a union bound over $O(n^6)$ choices of T, T' .
- (4) Note that $Z_{T,j,c-1}(t) \leq \sum_{r=0}^{g-j} X(\mathfrak{K}_{T,j+r,c+r-1}^{(4)})$ for $T \in \mathcal{A}(t)$ (recall that we assume no configuration in $\bigcup_{j=4}^g \mathfrak{J}_j$ is contained in another). (4) then follows from Lemma 3.7 and Lemma 8.2(4), and a union bound over $O_g(n^3)$ choices of T, j, c . \square

Proof of Theorem 9.6. In this proof we write “w.s.h.p.” (short for “with sufficiently high probability”) to mean that an event holds with probability at least $1 - b^\alpha - \exp(-n^\alpha)$, for α that will be chosen sufficiently small to satisfy certain inequalities. Let τ_{crude} be the first time t for which any of the events in Lemma 9.10 fail, and $\tau_{\text{crude}} = \infty$ if there is no such t . Note that for $t < \tau_{\text{crude}}$, $T, T' \in \mathcal{A}(t)$ edge-disjoint, and $e \in E(t)$ with $e \not\subseteq T$ we have

$$X_{T,T'}(t) := |\mathcal{T}_T(t) \cap \mathcal{T}_{T'}(t)| \leq Z_{T,T'}(t) + \sum_{e' \subseteq T} Z_{e',T'}(t) + \sum_{e'' \subseteq T'} Z_{e'',T}(t) \leq n^{200g\beta}, \quad (9.3)$$

$$X_{e,T}(t) := |\mathcal{X}_e(t) \cap \mathcal{T}_T(t)| \leq Z_{e,T}(t) \leq n^{100g\beta},$$

for sufficiently small β . This verifies the last two of the claims in Theorem 9.6 for such t . Let $\tau_{\text{stop}} := \tau_{\text{traj}} \wedge \tau_{\text{crude}} \wedge \tau_{\text{cut}}$. Our main goal is to show that w.s.h.p. we have $\tau_{\text{stop}} = \tau_{\text{cut}}$, meaning that the process does not terminate and stays on-trajectory until the cutoff time. This also implies the claims about the sets $\mathcal{X}_e(t)$ and \mathcal{T}_T . Indeed, for $t < \tau_{\text{stop}}$ we have

$$|\mathcal{X}_e(t)| = f_{\text{edge}}(t) \pm e_{\text{edge}}(t) \quad (9.4)$$

and

$$|\mathcal{T}_T(t)| = \sum_{e \in T} |\mathcal{X}_e(t)| - 3 + \sum_{j=4}^g |\mathfrak{X}_{T,j,j-4}(t)| \pm O_g(Z_{T,T}(t)) = f_{\text{threat}}(t) \pm O_g(e_{\text{edge}}(t)). \quad (9.5)$$

Similarly:

$$|\mathcal{A}(t)| = \frac{1}{3} \sum_{e \in E(t)} |\mathcal{X}_e(t)| \stackrel{(9.4)}{=} \frac{1}{3} |E(t)| (f_{\text{edge}}(t) \pm e_{\text{edge}}(t)) = \frac{1}{3} |E(0)| p(t) (f_{\text{edge}}(t) \pm e_{\text{edge}}(t)). \quad (9.6)$$

For β sufficiently small in terms of C it holds that $e_{\text{edge}}(t) = o(n^{-\beta} f_{\text{edge}}(t))$ and $e_{\text{edge}}(t) = o(n^{-\beta} f_{\text{threat}}(t))$ with $t \leq \tau_{\text{cut}}$. Thus it remains to show that w.s.h.p. $\tau_{\text{stop}} = \tau_{\text{cut}}$.

For an edge $e \in E(0)$, let τ_e be the first t for which $e \notin E(t)$ (i.e., the “point when it gets covered”), and let $\tau_e^{\text{freeze}} = \tau_{\text{stop}} \wedge (\tau_e - 1)$. For a triangle $T \in \mathcal{A}(0)$, let τ_T be the first t for which $T \notin \mathcal{A}(t)$ (i.e., the “point when it becomes unavailable”) and let $\tau_T^{\text{freeze}} = \tau_{\text{stop}} \wedge (\tau_T - 1)$. Define

$$X_e^\pm(t) = \pm |\mathcal{X}_e(t \wedge \tau_e^{\text{freeze}})| \mp f_{\text{edge}}(t \wedge \tau_e^{\text{freeze}}) - e_{\text{edge}}(t \wedge \tau_e^{\text{freeze}}),$$

$$X_{T,j,c}^\pm(t) = \pm |\mathfrak{X}_{T,j,c}(t \wedge \tau_T^{\text{freeze}})| \mp f_{j,c}(t \wedge \tau_T^{\text{freeze}}) - e_{j,c}(t \wedge \tau_T^{\text{freeze}}).$$

Here, each of these definitions is really shorthand for two separate definitions, one with a superscript “+” and one with a superscript “−”. By Lemma 9.10 (the conclusion of which occurs with sufficiently high probability) and the definition of τ_{traj} , in order to show that $\tau_{\text{stop}} = \tau_{\text{cut}}$ w.s.h.p., it suffices to show that $X_e^s(t) \leq 0$ and $X_{T,j,c}^s(t) \leq 0$ for all e, T, j, c, t and all $s \in \{-, +\}$ with probability at least $1 - \exp(-n^\alpha)$.

For β sufficiently small in terms of C , we have $X_e^s(0) \leq -n^{1-2/C}$ and $X_{T,j,c}^s(0) \leq -n^{j-3-c-2/C}$ for all e, T, j, c, s by the initial regularity conditions in the definition of (C, β, b) -goodness. For $0 \leq t \leq \tau_{\text{stop}}$ we

define $\mathcal{C}^+(t) = (\mathcal{C}(t), \mathcal{D}, G, \mathcal{A}, \mathcal{J}_4, \dots, \mathcal{J}_g)$. That is, $\mathcal{C}^+(t)$ consists of all data contributing to the process up to step t . Recalling the notation $\Delta X(t) = X(t+1) - X(t)$, we will show for all such t, e, T, j, c, s that:

- (A) $\mathbb{E}[\Delta X_e^s(t) | \mathcal{C}^+(t)], \mathbb{E}[\Delta X_{T,j,c}^s(t) | \mathcal{C}^+(t)] \leq 0$. That is to say, X_e^s and $X_{T,j,c}^s$ are supermartingales.
- (B) $\mathbb{E}[|\Delta X_e^s(t)| | \mathcal{C}^+(t)] \leq n^{-1+1/3}$ and $\mathbb{E}[|\Delta X_{T,j,c}^s(t)| | \mathcal{C}^+(t)] \leq n^{j-5-c+1/3}$.
- (C) $|\Delta X_e^s(t)| \leq n^{1/3}$ and $|\Delta X_{T,j,c}^s(t)| \leq n^{j-c-4+1/3}$.

The desired result will then follow from Freedman's martingale concentration inequality ([Corollary 3.4](#)) as long as C is larger than some absolute constant (which we may assume without loss of generality).

We note that $\Delta X_e^s(t) = \Delta X_{T,j,c}^s(t) = 0$ trivially for $t \geq \tau_{\text{stop}}$. Similarly $\Delta X_e^s(t) = 0$ if $t \geq \tau_e - 1$ and $\Delta X_{T,j,c}^s(t) = 0$ if $t \geq \tau_T - 1$. Thus, we may fix an outcome $\mathcal{C}^+(t)$ of the process for which $t < \tau_{\text{stop}}$ and when $\mathcal{X}_e(t)$ is involved we condition such that $e \in E(t)$ and $t+1 < \tau_e$, while when $\mathfrak{X}_{T,j,c}(t)$ is involved we condition such that $T \in \mathcal{A}(t)$ and $t+1 < \tau_T$.

Step 1: Expected changes. Let $e \in E(t)$. We note that $\tau_e > t+1$ if and only if $T^*(t) \notin \mathcal{X}_e(t)$. Conditioning on this event, and in light of the fact that $|\mathcal{X}_e(t)| \leq 4|\mathcal{A}(t)|/|E(t)|$, we see that $T^*(t)$ is chosen uniformly at random from $|\mathcal{A}(t)| - |\mathcal{X}_e(t)| = (1 \pm 4/|E(t)|)|\mathcal{A}(t)|$ triangles. Additionally, each $T \in \mathcal{X}_e(t)$ will not be in $\mathcal{A}(t+1)$ if and only if $T^*(t+1) \in \mathcal{T}_T(t) \setminus \mathcal{X}_e(t)$. Therefore:

$$\begin{aligned} \mathbb{E}[\Delta |\mathcal{X}_e(t)| | \mathcal{C}^+(t), t < \tau_e^{\text{freeze}}] &= -\frac{1}{|\mathcal{A}(t)| - |\mathcal{X}_e(t)|} \sum_{T \in \mathcal{X}_e(t)} (|\mathcal{T}_T(t)| - |\mathcal{X}_e(t)|) \\ &\stackrel{(9.4), (9.5)}{=} -\frac{1}{(1 \pm 4/|E(t)|)|\mathcal{A}(t)|} \sum_{T \in \mathcal{X}_e(t)} (f_{\text{threat}}(t) - f_{\text{edge}}(t) \pm O_g(e_{\text{edge}}(t))) \\ &= -\frac{|\mathcal{X}_e(t)|(f_{\text{threat}}(t) - f_{\text{edge}}(t) \pm O_g(e_{\text{edge}}(t)))}{(1 \pm 4/|E(t)|)|\mathcal{A}(t)|} \\ &\stackrel{(9.4)}{=} -\frac{(f_{\text{edge}}(t) \pm e_{\text{edge}}(t))(f_{\text{threat}}(t) - f_{\text{edge}}(t) \pm O_g(e_{\text{edge}}(t)))}{(1 \pm 4/|E(t)|)|\mathcal{A}(t)|} \\ &= -\frac{f_{\text{edge}}(t)(f_{\text{threat}}(t) - f_{\text{edge}}(t))}{|\mathcal{A}(t)|} \pm O_g\left(\frac{e_{\text{edge}}(t)(f_{\text{edge}}(t) + f_{\text{threat}}(t))}{|\mathcal{A}(t)|}\right) \\ &\stackrel{(9.6)}{=} -\frac{f_{\text{edge}}(t)(f_{\text{threat}}(t) - f_{\text{edge}}(t))}{|E(0)|p(t)(f_{\text{edge}}(t) \pm e_{\text{edge}}(t))/3} \pm O_g\left(\frac{e_{\text{edge}}(t)(f_{\text{edge}}(t) + f_{\text{threat}}(t))}{|\mathcal{A}(t)|}\right) \\ &= -\frac{3(f_{\text{threat}}(t) - f_{\text{edge}}(t))}{|E(0)|p(t)} \pm O_g\left(\frac{e_{\text{edge}}(t)(f_{\text{edge}}(t) + f_{\text{threat}}(t))}{|\mathcal{A}(t)|}\right). \end{aligned}$$

We observe that for $t < \tau_{\text{cut}}$ we have $f_{\text{threat}}(t) = O_{g,C}(f_{\text{edge}}(t)) = O_{g,C}(|\mathcal{A}(t)|/|E(t)|)$. Thus:

$$\mathbb{E}[\Delta |\mathcal{X}_e(t)| | \mathcal{C}^+(t), t < \tau_e^{\text{freeze}}] = -\frac{3(f_{\text{threat}}(t) - f_{\text{edge}}(t))}{|E(0)|p(t)} \pm O_{g,C}\left(\frac{e_{\text{edge}}(t)}{|E(t)|}\right). \quad (9.7)$$

Next, we calculate $\mathbb{E}[\Delta |\mathfrak{X}_{T,j,c}(t)|]$ for $T \in \mathcal{A}(t)$, $4 \leq j \leq g$ and $0 \leq c \leq j-4$. In contrast with $\mathcal{X}_e(t)$, which always decreases with t , sets of triangles can be added as well as removed from $\mathfrak{X}_{T,j,c}$ in any given time step. Specifically, a configuration $\mathcal{S} \in \mathfrak{X}_{T,j,c-1}(t)$ becomes a member of $\mathfrak{X}_{T,j,c}(t+1)$ if $T^*(t+1) \in \mathcal{S} \setminus \{T\}$ — but only if $T^*(t+1)$ does not threaten an additional triangle in \mathcal{S} . On the other hand, a configuration $\mathcal{S} \in \mathfrak{X}_{T,j,c}(t)$ is not in $\mathfrak{X}_{T,j,c}(t+1)$ if $T^*(t+1) \in \mathcal{S}$ or if any of the triangles in \mathcal{S} becomes unavailable (i.e., for some $T' \in \mathcal{S}$ we have $T^*(t+1) \in \mathcal{T}_{T'}(t) \cup \mathcal{S}$). With this in mind (and accounting for double counting) we write

$$\mathbb{E}[\Delta |\mathfrak{X}_{T,j,c}(t)| | \mathcal{C}^+(t), t < \tau_T^{\text{freeze}}] = \frac{N^{\text{loss}} + N^{\text{gain}}}{|\mathcal{A}(t)| - |\mathcal{T}_T(t)| - 1},$$

where

$$N^{\text{loss}} = -\sum_{\mathcal{S} \in \mathfrak{X}_{T,j,c}(t)} \left(\sum_{T' \in \mathcal{S} \cap \mathcal{A}(t)-T} |\mathcal{T}_{T'}(t)| \pm O\left(1 + \sum_{T', T'' \in \mathcal{S} \cap \mathcal{A}(t)} X_{T', T''}\right) \right),$$

and

$$N^{\text{gain}} = (|\mathfrak{X}_{T,j,c-1}(t)| - O_g(Z_{T,j,c-1}(t)))(j-2-c).$$

(Here, and in the following calculations, some of the random variables, such as N^{gain} and $f_{j,c-1}(t)$, are undefined when $c=0$. In these cases we define them as 0, without further comment).

We observe that since $t < \tau_{\text{traj}}$ we have $|\mathcal{T}_T(t)| = O_{g,C}(f_{\text{edge}}(t))$ and so

$$|\mathcal{A}(t)| - |\mathcal{T}_T(t)| = \left(1 \pm O_{g,C}\left(\frac{1}{|E(t)|}\right)\right)|\mathcal{A}(t)| = \left(1 \pm O\left(\frac{e_{\text{edge}}(t)}{f_{\text{edge}}(t)}\right)\right)|E(0)|p(t)f_{\text{edge}}(t)/3.$$

Similarly, applying (9.3),

$$\begin{aligned} N^{\text{loss}} &= -(j-3-c)f_{j,c}(t)f_{\text{threat}}(t) \pm O_g(e_{j,c}(t)f_{\text{threat}}(t) + n^{200g\beta}f_{j,c}(t)) \\ &= -(j-3-c)f_{j,c}(t)f_{\text{threat}}(t) \pm O_{g,C}\left(\frac{e_{j,c}(t)|\mathcal{A}(t)|}{p(t)|E(0)|}\right) \end{aligned}$$

and

$$N^{\text{gain}} = (j-2-c)f_{j,c-1}(t) \pm O_g(e_{j,c-1}(t) + Z_{T,j,c-1}(t)) = (j-2-c)f_{j,c-1}(t) \pm O_{g,C}\left(\frac{e_{j,c}(t)|\mathcal{A}(t)|}{p(t)|E(0)|}\right).$$

The last inequality comes from $Z_{T,j,c-1}(t) \leq n^{j-c-3+100g\beta}$ by Lemma 9.10.

Therefore:

$$\mathbb{E}[\Delta|\mathfrak{X}_{T,j,c}(t)| \mid \mathcal{C}^+(t), t < \tau_T^{\text{freeze}}] = \frac{(j-2-c)f_{j,c-1}(t) - (j-3-c)f_{j,c}(t)f_{\text{threat}}(t)}{|E(0)|p(t)f_{\text{edge}}(t)/3} \pm O_{g,C}\left(\frac{e_{j,c}(t)}{p(t)|E(0)|}\right).$$

Since $\Delta|\mathcal{X}_e(t)|$ is always negative, (9.7) implies that

$$\mathbb{E}\left[\left|\Delta|\mathcal{X}_e(t)|\right| \mid \mathcal{C}^+(t), t < \tau_e^{\text{freeze}}\right] \leq n^{-1+1/6}. \quad (9.8)$$

Considering separately the contributions from N^{loss} and N^{gain} , we see that

$$\mathbb{E}\left[\left|\Delta|\mathfrak{X}_{T,j,c}(t)|\right| \mid \mathcal{C}^+(t), t < \tau_T^{\text{freeze}}\right] \leq n^{j-5-c+1/6}. \quad (9.9)$$

Step 2: Functional differences. Direct computation reveals that

$$f'_{\text{edge}}(t) = e^{-\rho(t)}\left(\frac{3|A(0)|}{|E(0)|}\right)(2p(t)p'(t) - \rho'(t)p(t)^2) = -\frac{3}{|E(0)|p(t)}(f_{\text{threat}}(t) - f_{\text{edge}}(t)).$$

Furthermore

$$f''_{\text{edge}}(t) = e^{-\rho(t)}\left(\frac{3|A(0)|}{|E(0)|}\right)(2(p'(t))^2 - 4p(t)p'(t)\rho'(t) + (-\rho''(t) + (\rho'(t))^2)p(t)^2).$$

We note that $p'(t), \rho'(t) = O_{g,C}(1/|E(0)|)$ and $\rho''(t) = O_{g,C}(1/|E(0)|^2)$. Therefore, for every $0 \leq t \leq \tau_{\text{cut}}$ we have $|f''_{\text{edge}}(t)| = O_{g,C}(|A(0)|/|E(0)|^3) = O_{g,C}(n^3/n^{3(2-\beta)}) = O(e_{\text{edge}}(t)/(|E(0)|p(t)))$. Thus, by applying Taylor's theorem:

$$\Delta f_{\text{edge}}(t) = f'_{\text{edge}}(t) \pm O\left(\frac{e_{\text{edge}}(t)}{|E(0)|p(t)}\right).$$

Similarly

$$f'_{j,c}(t) = \frac{(j-2-c)f_{j,c-1}(t) - (j-3-c)f_{j,c}(t)f_{\text{threat}}(t)}{|E(0)|p(t)f_{\text{edge}}(t)/3}$$

and

$$|f''_{j,c}(t)| = O_{g,C}\left(f_{j,c}(t)\left(\frac{\mathbb{1}_{c \geq 2}}{t^2} + \frac{1}{(|E(0)|p(t))^2} + \rho''(t) + (\rho'(t))^2 + \frac{\mathbb{1}_{c \geq 1}}{t|E(0)|p(t)} + \frac{\rho'(t)\mathbb{1}_{c \geq 1}}{t} + \frac{\rho'(t)}{|E(0)|p(t)}\right)\right).$$

Therefore $|f''_{j,c}(t)| = O_{g,C}(|A(0)|/|E(0)|^{j-3-c}/(|E(0)|p(t)^2)) = O(e_{j,c}(t)/(|E(0)|p(t)))$. Applying Taylor's theorem:

$$\Delta f_{j,c}(t) = f'_{j,c}(t) \pm O\left(\frac{e_{j,c}(t)}{|E(0)|p(t)}\right).$$

We also compute that if $B \geq 2g$ then

$$\Delta e_{\text{edge}}(t) \geq \Omega\left(\frac{Be_{\text{edge}}(t)}{p(t)|E(0)|}\right), \quad \Delta e_{j,c}(t) \geq \Omega\left(\frac{Be_{j,c}(t)}{p(t)|E(0)|}\right).$$

Hence, if B is sufficiently large in terms of g, C then we have $\mathbb{E}[\Delta X_e^\pm(t) \mid \mathcal{C}^+(t), t+1 < \tau_e^{\text{freeze}}] \leq 0$ and $\mathbb{E}[\Delta X_{T,j,c}^\pm(t) \mid \mathcal{C}^+(t), t+1 < \tau_T^{\text{freeze}}] \leq 0$. Next, recall as shown earlier that if $t \geq \tau_e^{\text{freeze}}$ or $t \geq \tau_T^{\text{freeze}}$, respectively, then $\Delta X_e^s(t) = 0$ or $\Delta X_{T,j,c}^s(t) = 0$. This justifies (A).

To verify (B) we note that:

$$\begin{aligned} \mathbb{E}[\left|\Delta X_e^s(t)\right| \mid \mathcal{C}^+(t), t < \tau_e^{\text{freeze}}] &\leq \mathbb{E}\left[\left|\Delta|\mathcal{X}_e(t)|\right| \mid \mathcal{C}^+(t), t < \tau_e^{\text{freeze}}\right] + |\Delta f_{\text{edge}}(t)| + |\Delta e_{\text{edge}}(t)| \\ &\stackrel{(9.8)}{=} O_{g,C}\left(\frac{n^{1/6}}{n}\right) = o\left(\frac{n^{1/3}}{n}\right). \end{aligned}$$

Similarly,

$$\begin{aligned} \mathbb{E}[|\Delta X_{T,j,c}^s(t)| \mid \mathcal{C}^+(t), t < \tau_T^{\text{freeze}}] &\leq \mathbb{E}\left[|\Delta \mathfrak{X}_{T,j,c}(t)| \mid \mathcal{C}^+(t), t < \tau_T^{\text{freeze}}\right] + |\Delta f_{j,c}(t)| + |\Delta e_{j,c}(t)| \\ &\stackrel{(9.9)}{=} O_{g,C}\left(n^{j-5-c+1/6}\right) = o\left(n^{j-5-c+1/3}\right). \end{aligned}$$

Finally, again use that if $t \geq \tau_e^{\text{freeze}}$ or $t \geq \tau_T^{\text{freeze}}$, respectively, then $\Delta X_e^s(t) = 0$ or $\Delta X_{T,j,c}^s(t) = 0$. This justifies (B).

Step 3: Boundedness. It remains to verify (C). We have $|\Delta \mathcal{X}_e(t)| \leq 0$ and for $t < \tau_e^{\text{freeze}} \leq \tau_{\text{crude}}$ we have

$$-\Delta |\mathcal{X}_e(t)| \leq 1 + \max_{T^* \in \mathcal{A}(t) \setminus \mathcal{X}_e(t)} Z_{e,T^*}(t) \leq 2n^{100g\beta},$$

using the definition of τ_{crude} as the first time the outcomes of Lemma 9.10 are violated. Additionally, for $t < \tau_T^{\text{freeze}}$:

$$|\Delta \mathfrak{X}_{T,j,c}(t)| \leq \max_{T^* \in \mathcal{A}(t) \setminus \{T\}} Z_{T,T^*,j,c-1}(t) \leq n^{j-c-4+100g\beta}.$$

Finally:

$$\begin{aligned} -\Delta |\mathfrak{X}_{T,j,c}(t)| &\leq \begin{cases} \max_{T^* \in \mathcal{A}(t) \setminus \{T\}} \sum_{T' \in \mathcal{T}_{T^*}(t) \cup \{T^*\} \setminus \{T\}} Z_{T,T',j,c}(t) & \text{if } c < j - 4 \\ \max_{T^* \in \mathcal{A}(t) \setminus \{T\}} X_{T,T^*}(t) & \text{if } c = j - 4. \end{cases} \\ &\leq n^{j-c-4+100g\beta}. \end{aligned}$$

We can then deduce (C) (provided β is sufficiently small in terms of g). \square

9.1. Tracking extension statistics, starting from an almost-complete graph. Recall from the outline in Section 2 that we need our generalized high-girth triple process in two different settings, for the proof of Theorem 1.1. First, right after planting our absorber we run this process on the remaining edges of the complete graph K_N , to “sparsify” to some density p . Second, we run the process as the main part of the iterative step of the proof of Theorem 1.1; in stage k of the iteration we use our process to cover almost all edges of $K_N[U_k]$ which do not lie in U_{k+1} .

While Theorem 9.3 is sufficient in the latter setting, for the initial sparsification step we need to track a bit more information about the outcome of our generalized high-girth triple process; namely, we need to understand certain “extension statistics” between the “vortex” sets U_k . The following lemma tracks such statistics, but only in the relatively simple setting of the initial sparsification step.

Proposition 9.11. *For constant $g \in \mathbb{N}$, there is $\beta = \beta_{9.11}(g) > 0$ such that the following holds. Consider an absorbing structure H (with distinguished set X) satisfying Ab1 and Ab2 in Theorem 4.1, embedded in K_N . Let \mathcal{B} and $\mathfrak{F}_j^\mathcal{B}$ be as in Lemma 7.2, with $|\mathcal{B}|^{2g} \leq N^\beta$. We can then derive data suitable for running the generalized high-girth triple process defined in Definition 9.4:*

- Let $n = N$ and let $G \subseteq K_n$ be the graph of edges which do not appear in H .
- Let \mathcal{A} be the set of triangles in G which do not create an Erdős configuration when added to \mathcal{B} .
- Let $\mathfrak{J}_4, \dots, \mathfrak{J}_g$ be defined as follows. First, let $\mathfrak{J}'_4, \dots, \mathfrak{J}'_g$ be obtained from $\mathfrak{F}_4^\mathcal{B}, \dots, \mathfrak{F}_g^\mathcal{B}$ by removing all configurations which contain a triangle not in \mathcal{A} . Then, obtain $\mathfrak{J}_4, \dots, \mathfrak{J}_g$ by “removing redundancies” from $\mathfrak{J}'_4, \dots, \mathfrak{J}'_g$ (to be precise, we remove from \mathfrak{J}'_j all configurations \mathcal{S} which fully include some configuration $\mathcal{S}' \in \bigcup_{j' < j} \mathfrak{J}'_{j'}$).

Then, for some $\nu = \nu_{9.11}(g) > 0$, with probability $1 - n^{-\omega(1)}$: the process in Definition 9.4 runs for $\tau_{\text{cut}} := \lceil (1 - n^{-\nu})|E|/3 \rceil$ steps without terminating, and for $0 \leq t \leq \tau_{\text{cut}}$,

- $|\mathcal{X}_e(t) - f_{\text{edge}}(t)| \leq n^{-\beta} f_{\text{edge}}(t)$ for each $e \in E(t)$;
- $|\mathcal{T}_T(t) - f_{\text{threat}}(t)| \leq n^{-\beta} f_{\text{threat}}(t)$ for each $T \in \mathcal{A}(t)$.
- $|\mathcal{T}_T(t) \cap \mathcal{T}_{T'}(t)| \leq n^{1/2}$ for each pair of edge-disjoint $T, T' \in \mathcal{A}(t)$.
- $|\mathcal{T}_T(t) \cap \mathcal{X}_e(t)| \leq n^{1/2}$ for each $T \in \mathcal{A}(t)$ and $e \in E(t)$ with $e \not\subseteq T$.

Moreover, there is a constant $\tilde{B} = \tilde{B}_{9.11}(g) > 0$ such that the following hold for any $\zeta \in (0, 1/2)$.

- (1) Fix a vertex v and a vertex subset $U \subseteq V(G)$ such that either

- (i) $|U| \geq n^{1/2}$, or
- (ii) U is disjoint from $V(H) \setminus X$, and $v \notin V(H)$, and $|U| \geq n^\zeta$.

Let $\tilde{\tau}_{\text{cut}} = (1 - |U|^{-1/\tilde{B}})|E|/3$. Then with probability $1 - n^{-\omega_\zeta(1)}$, for each $0 \leq t \leq \tilde{\tau}_{\text{cut}}$: there are $(1 + o(1))p(t)|U|$ edges in $E(t)$ between v and U .

- (2) Fix a set of at most 6 edges $Q \subseteq E(G)$ and a vertex subset $U \subseteq V(G)$ such that either

- (i) $|U| \geq n^{1/2}$, or
- (ii) U is disjoint from $V(H) \setminus X$, no edge in Q contains a vertex of $V(H)$, and $|U| \geq n^\zeta$.

Let $\tilde{\tau}_{\text{cut}} = (1 - |U|^{-1/\tilde{B}})|E|/3$. Then with probability $1 - n^{-\omega_\zeta(1)}$, for each $0 \leq t \leq \tilde{\tau}_{\text{cut}}$: there are $(1 + o(1))p(t)^{|V(Q)|}e^{-|Q|\rho(t)}|U|$ vertices $u \in U$ such that $xyu \in \mathcal{A}(t)$ for each edge $xy \in Q$.

Note that τ_{cut} and $\tilde{\tau}_{\text{cut}}$ are different: if U is very small then we can't run the process for very long.

To prove [Proposition 9.11](#) we will extend the analysis in the proof of [Theorem 9.6](#).

Proof. We will take $\beta < \beta_{9.6}(g, O_g(1))$ (and also assume that β is small enough to satisfy certain inequalities throughout the proof).

Let $\mathcal{D} = \emptyset$, $k = 0$, $U_0 = V(K_n)$, $b = 0$ and $\mathfrak{F}_j^{\text{sup}} = \mathfrak{J}_j$ for each $4 \leq j \leq g$. We first claim that these data, together with $G, \mathcal{A}, \mathfrak{J}_4, \dots, \mathfrak{J}_g$, are $(O_g(1), \beta, n^{-\omega(1)})$ -good, as defined in [Definition 9.2](#). Given this claim, all parts of [Proposition 9.11](#) except (1) and (2) will directly follow from [Theorem 9.6](#).

Intuitively, this goodness claim follows from the fact that the absorber H is so small that G is very nearly just K_N , and \mathcal{A} is very nearly just the set of triangles in K_N , and \mathfrak{J}_j is very nearly just the set of j -vertex Erdős configurations (by symmetry, every triangle is included in the same number of Erdős configurations). To be precise, $(2, \beta, n^{-\omega(1)})$ -goodness follows from the following facts.

- For every edge $e \in G$, the number of triangles in \mathcal{A} including e is at least $n - v(H) = (1 - n^{-1/3})|\mathcal{A}|/|E(G)|$. ([Lemma 3.2\(1\)](#) implies every vertex in an Erdős configuration is contained in at least two triangles of the configuration, so every triangle of G not in \mathcal{A} must have all its vertices in $V(H)$.)
- For every triangle T in G , the number of configurations $\mathcal{S} \in \mathfrak{J}_j$ including T which are *not* Erdős configurations is at most

$$\sum_{i=1}^{g-j} |\mathcal{B}|^i n^{j+i-(i+1+3)} = O_g(|\mathcal{B}|^g n^{j-4}) \leq n^{-2/3} |\mathcal{A}|^{j-3} / |E(G)|^{j-3}$$

(here we are summing over possibilities for a configuration $\mathcal{S} \in \mathfrak{J}_j$ which arises from an Erdős $(i+j)$ -configuration including a set \mathcal{Q} of $i \geq 1$ triangles of \mathcal{B} , and observing that in such a situation we have $v(\mathcal{Q} \cup \{T\}) \geq i + 1 + 3$ by [Lemma 3.2\(1\)](#)).

- For every triangle T in G , the number of Erdős j -configurations containing T which do not appear in \mathfrak{J}_j is at most

$$\sum_{j'=4}^g \sum_{i=1}^{g-j'} |\mathcal{B}|^i n^{j'+i-(i+1+3)} \cdot n^{j-((j'-2)+3)} = O_g(|\mathcal{B}|^g n^{j-5}) \leq n^{-2/3} |\mathcal{A}|^{j-3} / |E(G)|^{j-3}$$

(here we are summing over Erdős j -configurations which fully include some $\mathcal{S} \in \mathfrak{J}_{j'}$ which arises from an Erdős $(i+j')$ -configuration, for $i \geq 1$).

- For every triangle T in G , the number of Erdős j -configurations containing T which also contain a triangle not in \mathcal{A} is at most $(|E(H)|n + |V(H)|^3) \cdot n^{j-5} \leq n^{-2/3} |\mathcal{A}|^{j-3} / |E(G)|^{j-3}$ (since two edge-disjoint triangles span at least 5 vertices, and from earlier a triangle of G not in \mathcal{A} must have all vertices in $V(H)$).

Now, it suffices to consider (1) and (2). The proofs of (1) and (2) are very similar (both involve defining an auxiliary supermartingale and using similar analysis as in [Theorem 9.6](#)), so we handle them together.

First, for (1), we note that v has at least $(1 - n^{-1/3})|U|$ neighbors in U , with respect to G . This is true for different reasons in case (i) and (ii): in case (i), we observe that v is adjacent to all vertices of U , and in case (ii), we use that $|V(H)|$ is tiny compared to $|U|$.

For (2), we similarly note that there are at least $(1 - n^{-1/3})|U|$ vertices $u \in U$ such that $e \cup \{u\} \in \mathcal{A}$ for each $e \in Q$. Again, this is true for different reasons in cases (i) and (ii). In case (ii), we recall from [Lemma 3.2\(1\)](#) that every vertex in an Erdős configuration is contained in at least two triangles of the configuration, so every triangle not in \mathcal{A} must have all its vertices in $V(H)$. That is to say, every vertex in U forms a valid triangle with every edge in Q . For (i), we again use that $|V(H)|$ is tiny compared to $|U|$.

Recall the definitions of τ_{traj} and τ_{crude} from the proof of [Theorem 9.6](#) (taking $C = 2$), and let $\tilde{Z}_{e,T}(t) \leq Z_{e,T}(t)$ be the number of forbidden configurations $\mathcal{S} \in \bigcup_{j=4}^g \mathfrak{J}_j$ such that $\mathcal{S} \cap \mathcal{A}(t) = \{T, T'\}$ for some triangle T' consisting of e and a vertex of U . We claim that $\tilde{Z}_{e,T}(t) \leq |U|^{1/3}$ for all $t \leq \tau_{\text{traj}} \wedge \tilde{\tau}_{\text{cut}} \wedge \tau_{\text{crude}}$, with probability $1 - n^{-\omega_\zeta(1)}$. Indeed, in case (i) this is true with probability 1 (we have $\tilde{Z}_{e,T}(t) \leq Z_{e,T}(t) = n^{100g\beta} \leq |U|^{1/3}$ for all $t \leq \tau_{\text{crude}}$, provided β is sufficiently small). In case (ii), we use basically the same proof as for [Lemma 9.10\(3\)](#), applying [Lemma 8.3](#) in place of [Lemma 8.2\(2\)](#). Specifically, for each $t \leq \tau_{\text{traj}} \wedge \tilde{\tau}_{\text{cut}}$, we have $|\mathcal{A}(t)| = \Omega(|U|^{-4/\tilde{B}}|\mathcal{A}|)$ (where \tilde{B} is sufficiently large), and therefore any s given triangles appear in $\mathcal{C}(\tau_{\text{traj}} \wedge \tilde{\tau}_{\text{cut}})$ with probability $(O(|U|^{4/\tilde{B}}|E|/|\mathcal{A}|))^s$. Noting that $\mathcal{D} = \emptyset$, we can then deduce the desired fact from [Lemma 8.3](#) and [Lemma 3.7](#).

Let $\tilde{\tau}_{\text{crude}}$ be the first time t for which $\tilde{Z}_{e,T}(t) \geq |U|^{1/3}$. Define the error threshold $\tilde{e}(t) = p(t)^{-\tilde{B}/8}|U|n^{-1/6}$.

For (1), let $f(t) = p(t)|U|$ and let $Y(t)$ be the set of vertices $u \in U$ for which $uv \in E(t)$. For (2), let $f(t) = p(t)^{|V(Q)|} e^{-|Q|\rho(t)} |U|$ and let $Y(t)$ be the set of vertices $u \in U$ such that $xyu \in \mathcal{A}(t)$ for each $xy \in Q$. For both (1) and (2), we let $\tilde{\tau}_{\text{traj}}$ be the first t for which $||Y(t)| - f(t)| > \tilde{e}(t)$, and let $\tilde{\tau}_{\text{stop}} = \tilde{\tau}_{\text{cut}} \wedge \tau_{\text{stop}} \wedge \tilde{\tau}_{\text{crude}} \wedge \tilde{\tau}_{\text{traj}}$ (recalling the definition of τ_{stop} from the proof of [Theorem 9.6](#), with $C = 2$). Then, define

$$X^\pm(t) = \pm |Y(t \wedge \tilde{\tau}_{\text{stop}})| \mp f(t \wedge \tilde{\tau}_{\text{stop}}) - \tilde{e}(t \wedge \tilde{\tau}_{\text{stop}}).$$

It now suffices to show that, for $s \in \{+, -\}$,

- (A) $\mathbb{E}[\Delta X^s(t) | \mathcal{C}(t)] \leq 0$,
- (B) $\mathbb{E}[|\Delta X^s(t)| | \mathcal{C}(t)] \leq |U|^{1+1/3}/n^2$,
- (C) $|\Delta X^s(t)| \leq |U|^{1/3}$.

Indeed, the desired result will then follow from [Corollary 3.4](#).

Step 1: expected changes. Fix t and an outcome of $\mathcal{C}^+(t)$ such that $t \leq \tilde{\tau}_{\text{stop}} \leq \tau_{\text{stop}}$. Recall that τ_{stop} is defined in terms of error functions $e_{\text{edge}}, e_{j,c}$, and that $e_{\text{edge}}(t), e_{j,j-4}(t) = p(t)^{-B} n^{1-1/4}$ for some constant B . We may assume that \tilde{B} is large relative to B , so $(e_{\text{edge}}(t) + e_{j,j-4}(t))/f_{\text{edge}}(t)$ is much smaller than $\tilde{e}(t)/f(t)$.

Let $A(t) = f_{\text{edge}}(t)|E(t)|/3$. For (1), we compute

$$\mathbb{E}[\Delta |Y(t)| | \mathcal{C}^+(t)] = -\frac{\sum_{u \in Y(t)} |\mathcal{X}_{vu}|}{|\mathcal{A}(t)|} = \frac{-f(t)f_{\text{edge}}(t)}{A(t)} \pm O\left(\frac{\tilde{e}(t)}{p(t)|E(0)|}\right).$$

For (2), for each $u \in U$, let $\mathcal{Z}_{u,Q}$ be the set of $|Q|$ triangles of the form xyu for $xy \in Q$. We have

$$\begin{aligned} \mathbb{E}[\Delta |Y(t)| | \mathcal{C}^+(t)] &= -\frac{\sum_{u \in Y(t)} \left(\sum_{x \in V(Q)} |\mathcal{X}_{xu}(t)| + \sum_{T \in \mathcal{Z}_{u,Q}} \sum_{j=4}^g |\mathfrak{X}_{T,j,j-4}(t)| - O(\varepsilon_{V(Q), \mathcal{Z}_{u,Q}}^u(t)) \right)}{|\mathcal{A}(t)|} \\ &= \frac{-|Y(t)|(|V(Q)|f_{\text{edge}}(t) + |Q|\sum_{j=4}^g f_{j,j-4}(t))}{A(t)} \pm O_g\left(\frac{\tilde{e}(t)}{p(t)|E(0)|}\right), \end{aligned}$$

where

$$\varepsilon_{V, \mathcal{Z}}^u(t) = \sum_{T, T' \in \mathcal{Z}} X_{T, T'}(t) + \sum_{T \in \mathcal{Z}, x \in V} Z_{ux, T}(t) \leq O(n^{200g\beta}) \leq e_{\text{edge}}(t)$$

(recalling the definition of $X_{T, T'}(t)$ from [Theorem 9.6](#)).

Step 2: computing functional differences. For (1), we compute

$$\Delta f(t) = \frac{-f(t)f_{\text{edge}}(t)}{A(t)}.$$

For (2), we compute

$$\Delta f(t) = \frac{-|Y(t)|(|V(Q)|f_{\text{edge}}(t) + |Q|\sum_{j=4}^g f_{j,j-4}(t))}{A(t)} + O\left(\frac{|U|}{|E(0)|^2}\right)$$

We also compute $\Delta \tilde{e}(t) \geq \Omega(\tilde{B}\tilde{e}(t)/(p(t)|E(0)|))$, so if \tilde{B} is sufficiently large then $\mathbb{E}\Delta X_e^\pm(t), \mathbb{E}\Delta X_{T,j,c}^\pm(t) \leq 0$, verifying (A). We can also verify (B) with the calculations so far.

Step 3: boundedness. Let $t \leq \tilde{\tau}_{\text{stop}}$. Note that $\Delta |Y(t)| \leq 0$. For (1) we have $-\Delta |Y(t)| \leq 2$ and for (2) we have

$$-\Delta |Y(t)| \leq |Q| \max_T \tilde{Z}_{e,T}(t) \leq 6|U|^{1/3},$$

which implies (C). □

10. THE MASTER ITERATION LEMMA

As sketched in [Section 2](#), the key ‘‘cover down’’ lemma driving iterative absorption takes as input an appropriate ‘‘vortex’’ of sets $K_N = U_0 \supseteq U_1 \supseteq \dots \supseteq U_\ell$, and a set of edge-disjoint triangles satisfying various properties (in particular, leaving uncovered a ‘‘typical’’ subgraph of edges in $K_N[U_k]$, for some $k < \ell$). Then, it extends this set of triangles in such a way that all the edges in $K_N[U_k]$ are covered except some of those in $K_N[U_{k+1}]$ (and the remaining uncovered edges in $K_N[U_{k+1}]$ form a typical subgraph). So, if certain assumptions are satisfied at the beginning (e.g., the uncovered edges form a typical subgraph of $K_N = K_N[U_0]$), then it is possible to ‘‘feed this lemma into itself’’ ℓ times, to obtain a set of edge-disjoint triangles covering all edges not in $K_N[U_\ell]$. In this section we state and prove a master iteration lemma in our setting (using the definitions and results in the previous sections).

The statement of our master iteration lemma will be a little technical, as it must take quite a lot of data as input. First, we require that our graph ‘‘looks like a random set of triangles in a random graph’’ in terms

of degrees and rooted subgraph statistics, and with respect to the descending sequence of subsets defining our vortex.

Definition 10.1 (Iteration-typicality). Fix a descending sequence of subsets $V(K_n) = U_k \supseteq \dots \supseteq U_\ell$. Consider a graph $G \subseteq K_n$ and a set of triangles \mathcal{A} in G . We say that (G, \mathcal{A}) is (p, q, ξ, h) -iteration-typical (with respect to our sequence of subsets) if:

- for every $k \leq i < \ell$, every vertex in U_i is adjacent to a $(1 \pm \xi)p$ fraction both of the vertices in U_i , and of the vertices in U_{i+1} (with respect to G), and
- for any i, i^* with $k \leq i < \ell$ and $i^* \in \{i, i+1\}$, and any edge subset $Q \subseteq G[U_i]$ spanning $|V(Q)| \leq h$ vertices, a $(1 \pm \xi)p^{|V(Q)}q^{|Q|}$ -fraction of the vertices $u \in U_{i^*}$ are such that $uvw \in \mathcal{A}$ for all $vw \in Q$.

In our application, $p \geq n^{-\nu}$ will be a sufficiently decaying function of n , and $q = \Omega_g(1)$ will exceed some positive constant (i.e., a constant fraction of the triangles in our graph are available). Also, for our application ξ just needs to be sufficiently small in terms of various constants (but we will take $\xi = o(1)$ for convenience), and h just needs to be at least 4 (this will be necessary to apply Lemma 5.1).

We will also need a stronger notion of well-distributedness than that in Definition 9.1, concerning a random pair of sets of triangles \mathcal{I}, \mathcal{D} (one of which will arise from our initial “sparsification” process, and one of which will arise from the master iteration lemma itself). We consider not only the probability that a given set of triangles are present in \mathcal{I} and \mathcal{D} , but also the probability that a given set of edges is uncovered by the triangles in \mathcal{I} . It is necessary to record all this information so that it can be used as input for the weight estimate lemmas in Section 8.

Definition 10.2 (Strong well-distributedness). Fix a descending sequence of subsets $V(K_N) = U_0 \supseteq \dots \supseteq U_k$, and $p \in \mathbb{R}$. Say that a random pair of sets of triangles $(\mathcal{I}, \mathcal{D})$ is *strongly* (p, C, b) -well-distributed with respect to our sequence of subsets if for any $s, t, r \geq 0$, any distinct triangles $I_1, \dots, I_s, D_1, \dots, D_t$ in K_N , and any distinct edges e_1, \dots, e_r in K_N , we have

$$\Pr[I_1, \dots, I_s \in \mathcal{I} \text{ and } D_1, \dots, D_t \in \mathcal{D} \text{ and } e_1, \dots, e_r \notin E(\mathcal{I})] \leq C^{s+t+r} \left(p^r N^{-s} \prod_{j=1}^t \frac{p}{|U_{\text{lev}(T_j)}|} + b \right),$$

where $\text{lev}(\cdot)$ is defined with respect to $U_0 \supseteq \dots \supseteq U_k$.

The reader should think of the “ p ” in Definition 10.2 as being the same as the “ p ” in Definition 10.1, of C as being a large constant “error factor”, and of b as being a very small additive error term (of the form $n^{-\omega(1)}$). To explain (informally) the meaning of the term “ $p/|U_{\text{lev}(T_j)}|$ ”: at stage k , we select about $p|U_k|^2$ triangles, out of about $p^3|U_k|^3$ available triangles (up to a constant factor depending only on g). So, each available triangle is selected with probability about $1/(p^2|U_k|)$. Each triangle in K_N is available with probability about p^3 , so the overall probability that a triangle is selected at stage k is at most about $p/|U_k|$.

Next, the following definitions describe the initial data (the vortex and the set of forbidden configurations), the data that is recorded at each stage (including the random set of triangles constructed so far), and the assumptions that we need to make about these data.

Definition 10.3 (Initial data). Consider a descending sequence of subsets $V(K_N) = U_0 \supseteq \dots \supseteq U_\ell$, and collections of sets of triangles $\mathfrak{F}_4, \dots, \mathfrak{F}_g$ in K_N . We say that the data $(U_0, \dots, U_\ell, \mathfrak{F}_4, \dots, \mathfrak{F}_g)$ are (ρ, β, ℓ) -structured if the following properties hold.

- $|U_k| = \lfloor |U_{k-1}|^{1-\rho} \rfloor$ for each $1 \leq k \leq \ell$.
- Each \mathfrak{F}_j is a collection of sets of $j-2$ edge-disjoint triangles, which is $(O_{g,\ell}(1), |U_k|^\beta)$ -well-spread with respect to the truncated sequence $U_0 \supseteq \dots \supseteq U_k$ for each $0 \leq k \leq \ell$.

Definition 10.4 (Data for stage k). Fix a descending sequence of subsets $V(K_N) = U_0 \supseteq \dots \supseteq U_\ell$, and fix collections of sets of triangles $\mathfrak{F}_4, \dots, \mathfrak{F}_g$. Now, for some $0 \leq k < \ell$ consider a random graph $G \subseteq K_N[U_k]$, a random set of triangles \mathcal{A} in G , and a random pair of sets of triangles \mathcal{I}, \mathcal{D} . We say that the data $(G, \mathcal{A}, \mathcal{I}, \mathcal{D})$ is (p, q, ξ, C, b) -iteration-good for stage k , with respect to the initial data $(U_0, \dots, U_\ell, \mathfrak{F}_4, \dots, \mathfrak{F}_g)$, if:

IG0 All degrees of G are even,

IG1 $(\mathcal{I}, \mathcal{D})$ is strongly (p, C, b) -well-distributed with respect to the truncated sequence $U_0 \supseteq \dots \supseteq U_k$, and if the following three properties are satisfied with probability $1 - \xi$.

IG2 \mathcal{I} and \mathcal{D} are disjoint, and $\mathcal{I} \cup \mathcal{D}$ is a partial Steiner triple system containing none of the configurations in $\bigcup_{j=4}^g \mathfrak{F}_j$.

IG3 (G, \mathcal{A}) is $(p, q, \xi, 4)$ -iteration-typical with respect to the truncated sequence $U_k \supseteq \dots \supseteq U_\ell$.

IG4 None of the edges of G are covered by the triangles in $\mathcal{I} \cup \mathcal{D}$, and for every $T \in \mathcal{A}$, the set of triangles $\mathcal{I} \cup \mathcal{D} \cup \{T\}$ contains no configuration in $\bigcup_{j=4}^g \mathfrak{F}_j$ (that is to say, the triangles in \mathcal{A} are still “available” for use, in that they would not complete a forbidden configuration on their own).

We also introduce some notation for how our data are updated as we add new triangles to our partial Steiner triple system.

Definition 10.5 (Updating data). Fix collections of sets of triangles $\mathfrak{F}_4, \dots, \mathfrak{F}_g$, let $G \subseteq K_N$ be any graph and \mathcal{D} any set of triangles in K_N , let \mathcal{A} be any set of triangles in G , and let $U \subseteq V(G)$ be any subset of vertices in G . We define $G_U(\mathcal{D}) \subseteq G[U]$ to be the graph of edges of $G[U]$ which do not appear in any of the triangles in \mathcal{D} (i.e., the edges that are uncovered by \mathcal{D}). We also define $\mathcal{A}_U(\mathcal{D}) \subseteq \mathcal{A}$ to be the set of all triangles $T \in \mathcal{A}$ with edges in $G_U(\mathcal{D})$ for which $\mathcal{D} \cup \{T\}$ contains no configuration in $\bigcup_{j=4}^g \mathfrak{F}_j$ that includes T (i.e., the set of triangles that are still available in combination with \mathcal{D}).

Finally, we state our master iteration lemma. Here and in the rest of the paper, g will be treated as a constant for the purpose of all asymptotic notation. We will encounter several other parameters (e.g. β, θ, ρ, ν), depending on g , which will also be treated as constants for the purpose of asymptotic notation.

Proposition 10.6. *There are positive real-valued functions*

$$\beta_{10.6} : \mathbb{N} \rightarrow \mathbb{R}_{>0}, \quad \rho_{10.6} : \mathbb{N} \times \mathbb{R}_{>0} \rightarrow \mathbb{R}_{>0}, \quad \nu_{10.6} : \mathbb{N} \times \mathbb{R}_{>0} \rightarrow \mathbb{R}_{>0}$$

such that the following holds. Fix a constant $g \in \mathbb{N}$, let $0 < \beta \leq \beta_{10.6}(g)$, let $0 < \rho \leq \rho_{10.6}(g, \beta)$, and let $0 < \nu \leq \nu_{10.6}(g, \rho)$. Fix (ρ, β, ℓ) -structured initial data $(U_0, \dots, U_\ell, \mathfrak{F}_4, \dots, \mathfrak{F}_g)$, where $\ell = O(1)$, consider some $k \in \{0, \dots, \ell - 1\}$, and consider data $(G, \mathcal{A}, \mathcal{I}, \mathcal{D})$ that is (p, q, ξ, C, b) -iteration-good for stage k , where

$$|U_k|^{-\nu} \leq p \leq |U_k|^{-\Omega(1)}, \quad q = \Omega(1), \quad \xi = o(1), \quad C = O(1), \quad b = n^{-\omega(1)}.$$

Then, there is a random set of triangles $\mathcal{M} \subseteq \mathcal{A}$, such that with high probability all edges of G are covered by \mathcal{M} except those in $G[U_{k+1}]$, and such that the updated data

$$(G_{U_{k+1}}(\mathcal{M}), \mathcal{A}_{U_{k+1}}(\mathcal{I} \cup \mathcal{D} \cup \mathcal{M}), \mathcal{I}, \mathcal{D} \cup \mathcal{M})$$

is $(p, q, o(1), O(1), n^{-\omega(1)})$ -iteration-good for stage $k + 1$.

We emphasize that the asymptotic notation in the conclusion of [Proposition 10.6](#) depends on the asymptotic notation in the assumptions (i.e., the $o(1)$ term in the conclusion may for example decay like the square root of the $o(1)$ term in the assumptions). We will only apply [Proposition 10.6](#) $\ell = O(1)$ times, so we will not actually have to worry about the dependence.

The assumption that $(G, \mathcal{A}, \mathcal{I}, \mathcal{D})$ are (p, q, ξ, C, b) -iteration-good for stage k tells us that [IG2–IG4](#) hold with probability at least $1 - \xi = 1 - o(1)$. In proving [Proposition 10.6](#), we may (and do) assume that [IG2–IG4](#) actually hold with probability 1. Indeed, we can simply condition on the event that [IG2–IG4](#) hold; this does not have much of an impact on the strong well-distributedness assumption in [IG1](#), because conditioning on an event that holds with probability $1 - o(1)$ increases the probability of any other event by a factor of at most $1 + o(1)$.

The proof of [Proposition 10.6](#) will be rather long, so we split it into subsections. As sketched in [Section 2](#), we find our desired set of triangles \mathcal{M} via a sequence of several steps. First we “regularize” our data using the tools in [Section 5](#), then we find a set of triangles covering almost all the edges not in $G[U_{k+1}]$, using the generalized high-girth process described in [Theorem 9.3](#) (both these steps are in [Section 10.1](#)). Then, we cover the remaining edges not in $G[U_{k+1}]$ with two further steps: a simple greedy algorithm to cover the remaining “internal” edges in $G[U_k \setminus U_{k+1}]$, (in [Section 10.2](#)), and a matching argument to cover the remaining “crossing” edges between U_k and U_{k+1} (in [Section 10.3](#)). Each subsection will begin with a detailed summary of what will be proved in that subsection; the reader may wish to read all these summaries before starting on the details.

10.1. Regularization and approximate covering. We now start the proof of [Proposition 10.6](#). So, consider initial data $(U_0, \dots, U_\ell, \mathfrak{F}_4, \dots, \mathfrak{F}_g)$, and random data $(G, \mathcal{A}, \mathcal{I}, \mathcal{D})$ for stage k , as in the lemma statement. Let $n = |U_k|$. Our goal is to find an appropriate set of triangles \mathcal{M} covering all the edges of G that are not in $G[U_{k+1}]$. In this subsection we use our regularization and high-girth process lemmas to find a random set of triangles $\mathcal{M}^* \subseteq \mathcal{A}$ covering *almost* all the desired edges.

The majority of the uncovered edges outside $G[U_{k+1}]$ will belong to a random set of edges R , which we set aside before applying our regularization and high-girth process lemmas. This random graph R will come in handy in the later steps of the proof of [Proposition 10.6](#), when we need to cover the leftover edges. To be precise, for any outcome of the random data $(G, \mathcal{A}, \mathcal{I}, \mathcal{D})$, let R be a random subgraph of G obtained by including each edge between U_{k+1} and $U_k \setminus U_{k+1}$ with probability $n^{-\theta}$ independently. Here $\theta > 0$ is a constant that will be chosen sufficiently small with respect to g, β (later, we will require ρ to be sufficiently small with respect to θ). Let $G^* = G \setminus (R \cup G[U_{k+1}])$ be the graph consisting of the edges we would like to cover to prove [Proposition 10.6](#), other than the edges in R .

To summarize what we will prove in this subsection, our random set of triangles $\mathcal{M}^* \subseteq \mathcal{A}$ will satisfy the following properties.

- A1** Each triangle in \mathcal{M}^* lies in the graph G^* , and the triangles in \mathcal{M}^* are edge-disjoint.
A2 $\mathcal{I} \cup \mathcal{D} \cup \mathcal{M}^*$ contains no forbidden configuration $\mathcal{S} \in \bigcup_{j=4}^g \mathfrak{F}_j$.
A3 With high probability (over R and the random data $\mathcal{I}, \mathcal{D}, G, \mathcal{A}$): for any $s, t \in \mathbb{N}$, any distinct edges $e_1, \dots, e_s \in G^*$ and triangles F_1^*, \dots, F_t^* , we have

$$\Pr[F_1^*, \dots, F_t^* \in \mathcal{M}^* \text{ and } e_1, \dots, e_s \notin E(\mathcal{M}^*) \mid R, \mathcal{I}, \mathcal{D}, G, \mathcal{A}] \leq (C^* n^{-\beta})^s (C^*/(p^2 n))^t + n^{-\omega(1)},$$

for some $C^* = O(1)$.

- A4** With high probability, for every pair of distinct vertices $u, w \in U_{k+1}$ there are at most $n^{1-2\theta+o(1)}$ vertices $v \in U_k \setminus U_{k+1}$ for which $uv, vw \in G \setminus E(\mathcal{M}^*)$.

10.1.1. *Regularizing the available triangles.* Let $\mathcal{A}_0^* \subseteq \mathcal{A}$ be the subset of available triangles that are in the graph G^* . By **IG3** (iteration-typicality),

- every edge $e \in G$ is in $(1 \pm o(1))p^2 qn$ triangles of \mathcal{A} , and
- for every clique $K \subseteq G$ with $2 \leq s \leq 4$ vertices, there are $(1 \pm o(1))p^s q \binom{s}{2} n$ different vertices u which form a triangle in \mathcal{A} with every edge of K .

Recall that $p \geq n^{-\nu}$, that each edge of R is present with probability $n^{-\theta}$, that $|U_{k+1}| = n^{1-\rho}$ and that $G^* = G \setminus (R \cup G[U_{k+1}])$. If ν is sufficiently small compared to ρ and θ , then using a Chernoff bound over the randomness of R , we can see that w.h.p. G is sufficiently similar to G^* (and \mathcal{A} is sufficiently similar to \mathcal{A}_0^*) that the above two properties also hold with G^* in place of G and \mathcal{A}_0^* in place of \mathcal{A} . So, by [Lemma 5.1](#), we can find $\mathcal{A}^* \subseteq \mathcal{A}_0^*$ for which every edge $e \in G^*$ is in $(1 \pm n^{-1/4})p^2 qn/4$ triangles of \mathcal{A}^* .

10.1.2. *Regularizing the forbidden configurations.* We introduce notation for the (random) configurations which are “locally” dangerous at stage k (i.e., given that we have already chosen the triangles in $\mathcal{I} \cup \mathcal{D}$, what subconfigurations must we now avoid?). Specifically, let \mathfrak{J}_j contain every set $\mathcal{S} \subseteq \mathcal{A}^*$ of $j-2$ triangles which comprises a forbidden configuration (in $\bigcup_{j'=4}^g \mathfrak{F}_{j'}$) together with some (possibly empty) set of triangles in $\mathcal{I} \cup \mathcal{D}$. That is to say, if we can choose \mathcal{M}^* in such a way as to avoid each configuration in $\bigcup_{j=4}^g \mathfrak{J}_j$, then **A2** is satisfied.

Note that each \mathfrak{J}_j can be interpreted as a $(j-2)$ -uniform hypergraph on the vertex set \mathcal{A}^* . We would like to apply [Lemma 5.2](#) to “regularize” these hypergraphs. Unfortunately, we cannot directly apply [Lemma 5.2](#) to each \mathfrak{J}_j . The reason is that these collections of configurations may have some redundancies: a configuration $\mathcal{S} \in \mathfrak{J}_j$ may be a proper subset of a configuration $\mathcal{S}' \in \mathfrak{J}_{j'}$, meaning that avoiding \mathcal{S} is a stronger property than avoiding \mathcal{S}' . Our generalized high-girth process analysis cannot tolerate such redundancies (cf. [Definition 9.2](#)).

Our application of [Lemma 5.2](#) is therefore a little delicate. We introduce some notation to remove redundancies: for collections of sets of triangles $\mathfrak{S}, \mathfrak{S}_1, \dots, \mathfrak{S}_h$, we define $\mathfrak{S}(\mathfrak{S}_1, \dots, \mathfrak{S}_h) \subseteq \mathfrak{S}$ to be the subcollection of \mathfrak{S} obtained by removing all supersets of configurations in $\bigcup_{j=1}^h \mathfrak{S}_j$. Now, we iteratively construct regularized sets of triangles $\mathfrak{J}_4^*, \dots, \mathfrak{J}_g^*$, with no redundancies, such that avoidance of all configurations in $\bigcup_{j=4}^g \mathfrak{J}_j^*$ implies avoidance of all configurations in $\bigcup_{j=4}^g \mathfrak{J}_j$. We will do this in such a way that the following conditions are satisfied (viewing each \mathfrak{J}_j^* as a $(j-2)$ -uniform hypergraph on the vertex set \mathcal{A}^*):

- (1) each \mathfrak{J}_j^* has maximum degree $o(|\mathcal{A}^*|/|E|)^{j-3}$,
- (2) the maximum and minimum degree of each \mathfrak{J}_j^* differ by $n^{o(1)}$,
- (3) There is a $O(n^\beta)$ -well-spread collection of sets of $j-2$ triangles $\mathfrak{F}_j^{\text{sup}}$ which includes $\mathfrak{F}_j \cup (\mathfrak{J}_j^* \setminus \mathfrak{J}_j)$.

So, suppose for some $4 \leq j \leq g$ we have already constructed $\mathfrak{J}_4^*, \dots, \mathfrak{J}_{j-1}^*$ in such a way that (1–3) hold; we describe how to construct \mathfrak{J}_j^* . First, we deduce from [Lemma 8.4](#) and [Corollary 3.8](#) that w.h.p. \mathfrak{J}_j has maximum degree $d_{\max} \leq p^{3(j-3)} n^{j-3+o(1)}$. Here we are applying [Lemma 8.4](#) with $y = O(1)$ and $z = O(n^\beta)$, where the weights of the form $\pi_{T'}$ describe the probability that a triangle T' is in $\mathcal{I} \cup \mathcal{D}$ (which is $O(1/N) + O(p/|U_{\text{lev}(T')}|) = O(1/|U_{\text{lev}(T')}|)$), and the weights of the form π_e describe the probability that an edge e is uncovered by \mathcal{I} . Then, with notation as in [Corollary 3.8](#) and [Lemma 8.4](#), $X(\mathfrak{B}_{T,j,j'})$ is the number of configurations in $\mathfrak{F}_{j'}$ which yield a configuration in \mathfrak{J}_j containing a particular triangle T . The joint probability guarantees that we need to apply [Corollary 3.8](#) are provided by the strong well-distributedness of $(\mathcal{I}, \mathcal{D})$.

Next, by **IG3** (iteration-typicality) and a Chernoff bound over the randomness of R , w.h.p. $|\mathcal{A}^*| = \Theta_g(p^3 n^3)$ and $|E(G^*)| = \Theta(pn^2)$, so in fact $d_{\max} \leq p^{j-3-o(1)} (|\mathcal{A}^*|/|E(G^*)|)^{j-3} = o(|\mathcal{A}^*|/|E(G^*)|)^{j-3}$. Let $\mathfrak{S}_j = \mathfrak{J}_j(\mathfrak{J}_4^*, \dots, \mathfrak{J}_{j-1}^*)$ be the subset of \mathfrak{J}_j obtained by removing redundancies with previously defined \mathfrak{J}_{j-1}^* .

Now, let $\mathfrak{H}^{(1)}$ be the $(j-2)$ -uniform hypergraph with vertex set \mathcal{A}^* , whose hyperedges are those sets of $j-2$ triangles which are not vertex-disjoint. Let \mathfrak{K} be the complete $(j-2)$ -uniform hypergraph with vertex set \mathcal{A}^* , and let $\mathfrak{H}^{(2)}$ be the complement of $\mathfrak{K}(\mathfrak{J}_4^*, \dots, \mathfrak{J}_{j-1}^*)$ (i.e., the hypergraph of supersets of

configurations in $\mathfrak{J}_4^*, \dots, \mathfrak{J}_{j-1}^*$. The maximum degree of $\mathfrak{H}^{(1)}$ is $O(n^{3(j-3)-1})$, and the maximum degree of $\mathfrak{H}^{(2)}$ is at most $\sum_{4 \leq j' < j} n^{3(j-j')} o((|\mathcal{A}^*|/|E|)^{j'-3}) = O(n^{3(j-3)-2})$. So, [Lemma 5.2](#) (applied with $\mathfrak{G} = \mathfrak{G}_j$ and $\mathfrak{H} = \mathfrak{H}^{(1)} \cup \mathfrak{H}^{(2)} \cup \mathfrak{G}_j$, both of which have $|\mathcal{A}|$ vertices) gives us a collection of sets of triangles \mathfrak{G}'_j , such that $\mathfrak{J}_j^* := \mathfrak{G}'_j \cup \mathfrak{G}_j$ w.h.p. satisfies (1–2) above. Also, \mathfrak{G}'_j is stochastically dominated by a binomial random hypergraph $\mathfrak{G}_j^{\text{rand}} \sim \mathbb{G}^{(j-2)}(|\mathcal{A}|, p_j)$, for $p_j = n^\beta n^{-2j+6}$ (assuming ν is sufficiently small compared to β). Let $\mathfrak{F}_j^{\text{sup}} = \mathfrak{G}_j^{\text{rand}} \cup \mathfrak{F}_j$, and note that by [Lemma 7.3](#), w.h.p. this collection of sets of triangles is $O(n^\beta)$ -well-spread. So, there is an outcome of \mathfrak{J}_j^* satisfying (1–3), as desired.

10.1.3. The generalized high-girth triple process. We now apply [Theorem 9.3](#) with the random objects $\mathcal{D} \cup \mathcal{I}, G^*, \mathcal{A}^*, \mathfrak{J}_4^*, \dots, \mathfrak{J}_g^*$. (We assume $\nu < \beta/2$, so $p \geq n^{-\beta/2}$). We thereby obtain a random set of triangles $\mathcal{M}^* \subseteq \mathcal{A}^*$ satisfying [A1](#) to [A3](#).

10.1.4. Counting extensions of edges. Finally, we verify [A4](#). For any $u, w \in U_{k+1}$ and $v \in U_k \setminus U_{k+1}$, the probability that $vu, vw \in G \setminus E(\mathcal{M}^*)$ is at most $n^{-2\theta} + O(n^{-\theta-\beta}) + O(n^{-2\beta}) \leq 2n^{-2\theta}$ by [A3](#) (vu and vw must be in R or uncovered by \mathcal{M}^* , and we are assuming θ is small relative to β). In fact, for any $s \in \mathbb{N}$ and any s choices of v , the probability that $vu, vw \in G \setminus E(\mathcal{M}^*)$ for all these v is at most $(2n^{-2\theta})^s + n^{-\omega(1)}$, so by [Corollary 3.9](#), with probability $1 - n^{-\omega(1)}$, there are at most $n^{1-2\theta+o(1)}$ different v for which this is the case. The desired result then follows from a union bound over choices of u, w .

10.2. Covering leftover internal edges. So far, we have a random set of triangles \mathcal{M}^* covering almost all of the edges of G^* . We distinguish two types of edges we still need to cover to prove [Proposition 10.6](#): first, we need to handle the remaining uncovered edges in $G[U_k \setminus U_{k+1}]$ (i.e., those fully outside U_{k+1}), and second, we need to handle the remaining uncovered edges between $U_k \setminus U_{k+1}$ and U_{k+1} (some of these are in G^* , but most are in R).

In this subsection, we handle all the edges of the first type. Namely, using a simple random greedy procedure, we augment \mathcal{M}^* with a further random set of triangles $\mathcal{M}^\dagger \subseteq \mathcal{A}$, covering all the remaining edges in $G[U_k \setminus U_{k+1}]$. In the process, we will also cover a few edges of our reserve graph R . To summarize what we will prove in this subsection, our random set of triangles $\mathcal{M}^\dagger \subseteq \mathcal{A}$ will satisfy the following properties.

- B1** Each triangle in \mathcal{M}^\dagger consists of an edge in $G[U_k \setminus U_{k+1}] \setminus E(\mathcal{M}^*)$ and two edges in R , and the triangles in \mathcal{M}^\dagger are edge-disjoint.
- B2** With high probability, no edge in $G[U_k \setminus U_{k+1}]$ is left uncovered by $\mathcal{M}^* \cup \mathcal{M}^\dagger$.
- B3** $\mathcal{I} \cup \mathcal{D} \cup \mathcal{M}^* \cup \mathcal{M}^\dagger$ contains no forbidden configuration $\mathcal{S} \in \bigcup_{j=4}^g \mathfrak{F}_j$.
- B4** Conditioning on any outcome of R, \mathcal{M}^* and the random data $\mathcal{I}, \mathcal{D}, G, \mathcal{A}$: for any $s \in \mathbb{N}$ and any distinct triangles $F_1^\dagger, \dots, F_s^\dagger \in G$ we have

$$\Pr[F_1^\dagger, \dots, F_s^\dagger \in \mathcal{M}^\dagger \mid R, \mathcal{M}^*, \mathcal{I}, \mathcal{D}, G, \mathcal{A}] \leq (C^\dagger / (n^{1-\rho-2\theta} p^2))^s,$$

for some $C^\dagger = O(1)$.

10.2.1. Defining a random greedy procedure. Arbitrarily enumerate the uncovered edges in $G[U_k \setminus U_{k+1}]$ as f_1, \dots, f_m . By iteration-typicality and a Chernoff bound using the randomness of R , for each $i \leq m$ there are $(1 + o(1))n^{-2\theta} p^2 q |U_{k+1}| = (1 + o(1))n^{1-\rho-2\theta} p^2 q$ choices for a triangle $T_i \in \mathcal{A}$ consisting of f_i and two edges of R (call these “candidates for T_i ”). However, we cannot arbitrarily choose such a triangle for each i , as these triangles may edge-intersect with each other, and they may form forbidden configurations (with each other and with \mathcal{M}^* , in combination with the triangles in $\mathcal{I} \cup \mathcal{D}$).

We will choose our triangles T_1, \dots, T_m via the following random greedy procedure. For each $i \leq m$, in order, we will consider all possible triangles $T \in \mathcal{A}$ consisting of f_i and two edges of $R \setminus E(T_1 \cup \dots \cup T_{i-1})$, such that $\mathcal{I} \cup \mathcal{D} \cup \mathcal{M}^* \cup \{T_1, \dots, T_{i-1}, T\}$ contains no forbidden configuration. We will choose such a triangle uniformly at random to take as T_i .

Of course, it is possible that there are no such triangles, in which case we do not define T_i . In fact, we leave T_i undefined (and say that our entire random greedy procedure fails) whenever there are fewer than $n^{1-\rho-2\theta} p^2 q / 2$ choices for T_i at step i (this ensures that at every step there is “a lot of randomness” in the triangles chosen throughout the process, and in particular that [B4](#) holds). Let \mathcal{M}^\dagger be the set of T_i which actually get defined in our process.

10.2.2. A well-distributedness-type property. Recall that we are assuming $(\mathcal{I}, \mathcal{D})$ is strongly $(p, C, n^{-\omega(1)})$ -well-distributed ([IG1](#)), for some $C = O(1)$, and recall the probabilistic guarantee on \mathcal{M}^* from [A3](#). We now claim that $(\mathcal{I}, \mathcal{D} \cup \mathcal{M}^* \cup \mathcal{M}^\dagger)$ is strongly $(p, O(1), n^{-\omega(1)})$ -well-distributed. In fact we need a slightly stronger property incorporating R : for any $s, t, r, u \in \mathbb{N}$, any distinct triangles $I_1, \dots, I_s, D_1, \dots, D_t$ in K_N , and any distinct edges $e_1, \dots, e_r, e'_1, \dots, e'_u \in K_N$, with e'_1, \dots, e'_u being between $U_k \setminus U_{k+1}$ and U_{k+1} , let \mathbf{E}

be the event that $I_1, \dots, I_s \in \mathcal{I}$, and e_1, \dots, e_r are uncovered by \mathcal{I} , and $D_1, \dots, D_t \in \mathcal{D} \cup \mathcal{M}^* \cup \mathcal{M}^\dagger$, and $e'_1, \dots, e'_u \in R$. We claim that there is a constant $C_{10.2.2} = O(1)$ such that

$$\Pr[\mathbf{E}] \leq C_{10.2.2}^{s+t+r+u} \left(p^r (pn^{-\theta})^u N^{-s} \prod_{j=1}^t \frac{p}{|U_{\text{lev}}(D_j)|} + n^{-\omega(1)} \right). \quad (10.1)$$

This will allow us to apply [Lemma 8.7](#) to analyze our random greedy procedure. To verify our claim, it is convenient to partition \mathbf{E} into sub-events, which we will later sum over. Let $\mathcal{T} = \{D_1, \dots, D_t\}$, and consider a partition $\mathcal{T} = \mathcal{T}^0 \cup \mathcal{T}^* \cup \mathcal{T}^\dagger$. We consider the sub-event $\mathbf{E}' \subseteq \mathbf{E}$ that \mathbf{E} occurs, and moreover that $\mathcal{T}^0 \subseteq \mathcal{D}, \mathcal{T}^* \subseteq \mathcal{M}^*, \mathcal{T}^\dagger \subseteq \mathcal{M}^\dagger$. We have

$$\begin{aligned} \Pr[\mathbf{E}'] &\leq C^{s+t+r} \left(p^{r+u+3|\mathcal{T}^*|} N^{-s} \prod_{D \in \mathcal{T}^0} \frac{p}{|U_{\text{lev}}(D)|} + n^{-\omega(1)} \right) \cdot (n^{-\theta})^u \\ &\quad \cdot \left(\left(\frac{C^*}{p^2 n} \right)^{|\mathcal{T}^*|} \left(\frac{C^*}{n^\beta} \right)^{|\mathcal{T}^\dagger|} + n^{-\omega(1)} \right) \cdot \left(\frac{C^\dagger}{n^{-2\theta} p^2 n^{1-\rho}} \right)^{|\mathcal{T}^\dagger|}. \end{aligned}$$

Indeed, for \mathbf{E}' to occur, first we must have $I_1, \dots, I_s \in \mathcal{I}$, then $e_1, \dots, e_r, e'_1, \dots, e'_u$ and the edges in the triangles in \mathcal{T}^* must be uncovered by \mathcal{I} , then we must have $\mathcal{T}^0 \subseteq \mathcal{D}$. The probability of these events (over the randomness of \mathcal{I}, \mathcal{D}) may be bounded using [IG1](#). Then, given an outcome of $(\mathcal{I}, \mathcal{D})$, the edges e'_1, \dots, e'_u must be in R . The probability of this event is $(n^{-\theta})^u$. Then, we must have $\mathcal{T}^* \subseteq \mathcal{M}^*$, and the edges in the triangles in \mathcal{T}^\dagger must be uncovered by \mathcal{M}^* . The probability of these events can be bounded using [A3](#). Finally, we must have $\mathcal{T}^\dagger \subseteq \mathcal{M}^\dagger$, and the probability of this can be bounded by [B4](#) (which we have already proved).

If θ, ν, ρ are sufficiently small with respect to β , then summing this bound for $\Pr[\mathbf{E}']$ over all 3^t choices of the partition $\mathcal{T}^0 \cup \mathcal{T}^* \cup \mathcal{T}^\dagger$ yields the desired bound [\(10.1\)](#), with say $C_{10.2.2} = 3CC^*C^\dagger$.

10.2.3. Bounding the impact of forbidden configurations. We now show that w.h.p. for each $i \leq m$, only a $o(1)$ -fraction of candidates for T_i are forbidden due to the fact that their addition to \mathcal{M}^\dagger would create a forbidden configuration (together with other triangles in $\mathcal{I} \cup \mathcal{D} \cup \mathcal{M}^* \cup \mathcal{M}^\dagger$).

Fix $i \leq m$ and let \mathcal{M}_i^\dagger be the set of triangles we have chosen before the i -th step of our random greedy procedure (so, \mathcal{M}_i^\dagger contains the triangles T_1, \dots, T_{i-1} , unless the procedure has already failed). Since \mathcal{M}_i^\dagger is a subset of \mathcal{M}^\dagger for each i , the well-distributedness-type property described in [\(10.1\)](#) holds with \mathcal{M}_i^\dagger in place of \mathcal{M}^\dagger . We now deduce from [Lemmas 3.7](#) and [8.7](#) that with probability $1 - n^{-\omega(1)}$ at most $n^{-2\theta} p^3 n^{1-\rho+o(1)}$ of the candidates for T_i (that is to say, an $o(1)$ fraction of these candidates) cannot be chosen due to forbidden configurations.

Indeed, here we are applying [Lemma 8.7](#) with $y = O(1)$, $z = O(n^\beta)$, $r = n^{-\theta}$ and $e = f_i$, where the weights of the form $\pi_{T,1}$ and $\pi_{T,2}$ describe the probability that a triangle T is in \mathcal{I} or $\mathcal{D} \cup \mathcal{M}^* \cup \mathcal{M}^\dagger$, respectively, and the weights of the form π_e describe the probability that an edge e is uncovered by \mathcal{I} and appears in R . In the notation of [Corollary 3.8](#) and [Lemma 8.7](#), $X(\mathfrak{F}_{f_i,j})$ is an upper bound on the number of configurations $\mathcal{S} \in \mathfrak{F}_j$ containing a candidate for T_i for which $\mathcal{S} \setminus \{T\} \subseteq \mathcal{I} \cup \mathcal{D} \cup \mathcal{M}^* \cup \mathcal{M}^\dagger$. The joint probability guarantees that we need to apply [Corollary 3.8](#) are provided by [\(10.1\)](#).

10.2.4. Bounding the impact of edge-intersections. We now show that for each $i \leq m$, with probability $1 - n^{-\omega(1)}$, only $n^{o(1)-\beta} p |U_{k+1}|$ of the candidates for T_i (that is, a $o(1)$ -fraction, provided θ, ν are small compared to β) are forbidden due to the fact that they edge-intersect previously chosen triangles in \mathcal{M}^\dagger . Note that a previously chosen triangle $T_j \in \mathcal{M}^\dagger$ can only edge-intersect a candidate for T_i if f_j and f_i share a vertex. So, it suffices to show that with probability $1 - n^{-\omega(1)}$ every vertex $v \in U_k \setminus U_{k+1}$ is incident to at most $n^{o(1)-\beta} p |U_{k+1}|$ of the $(1 + o(1)) p |U_{k+1}|$ edges of G which were uncovered by \mathcal{M}^* . By [A3](#), the probability a given edge $e \in E(G)$ incident to v is uncovered by \mathcal{M}^* is at most $O(n^{-\beta})$, and in fact we can use [A3](#) and [Corollary 3.9](#) to deduce the desired fact.

10.3. Covering leftover crossing edges. At this point, we have found a random set of triangles \mathcal{M}^* covering almost all the edges of G^* , as well as a random set of triangles \mathcal{M}^\dagger covering all the remaining edges in $G[U_k \setminus U_{k+1}]$. In this subsection we will find a final random set of triangles $\mathcal{M}^\ddagger \subseteq \mathcal{A}$ covering all the remaining edges between $U_k \setminus U_{k+1}$ and U_{k+1} (most of which are in the reserve graph R). The problem of covering the remaining edges between U_{k+1} and a given vertex $v \in U_k \setminus U_{k+1}$ can be viewed as a matching problem for a certain auxiliary graph (a ‘‘link graph’’ associated with v), so in this subsection we apply the tools from [Section 6](#).

Let $\gamma > 0$ be a constant which is small with respect to $g, \beta, \theta, \rho, \nu$. To summarize what we will prove in this subsection, our random set of triangles $\mathcal{M}^\ddagger \subseteq \mathcal{A}$ will satisfy the following properties.

- C1** Every triangle in \mathcal{M}^\ddagger consists of a single edge in $G[U_{k+1}]$, and two edges between $U_k \setminus U_{k+1}$ and U_{k+1} . Also, the triangles in $\mathcal{I} \cup \mathcal{D} \cup \mathcal{M}^* \cup \mathcal{M}^\dagger \cup \mathcal{M}^\ddagger$ are edge-disjoint.
- C2** With high probability, no edge in $G \setminus G[U_{k+1}]$ is left uncovered by $\mathcal{M}^* \cup \mathcal{M}^\dagger \cup \mathcal{M}^\ddagger$.
- C3** $\mathcal{I} \cup \mathcal{D} \cup \mathcal{M}^* \cup \mathcal{M}^\dagger \cup \mathcal{M}^\ddagger$ contains no forbidden configuration $\mathcal{S} \in \bigcup_{j=4}^g \mathfrak{F}_j$.
- C4** Conditioning on any outcome of $R, \mathcal{M}^*, \mathcal{M}^\dagger$ and the random data $\mathcal{I}, \mathcal{D}, G, \mathcal{A}$: every set of $s \in \mathbb{N}$ distinct triangles is present in \mathcal{M}^\ddagger with probability at most $(n^\gamma / (n^{-\theta} p^2 n^{1-\rho}))^s$. That is to say, \mathcal{M}^\ddagger is approximately stochastically dominated by a random set of triangles each independently selected with this probability.

After this subsection, it will just remain to verify that the union $\mathcal{M} = \mathcal{M}^* \cup \mathcal{M}^\dagger \cup \mathcal{M}^\ddagger$ satisfies the conclusion of [Proposition 10.6](#).

10.3.1. Defining link graphs. For a vertex $v \in U_k \setminus U_{k+1}$, let $N_{U_{k+1}}(v) \subseteq U_{k+1}$ be the set of vertices $u \in U_{k+1}$ such that $vu \in G$. Let L_v be the graph with vertex set $N_{U_{k+1}}(v)$, and an edge uw whenever $vuw \in \mathcal{A}$. Let W_v be the set of vertices $u \in N_{U_{k+1}}(v)$ such that uv has not yet been covered by $\mathcal{M}^* \cup \mathcal{M}^\dagger$.

Now, our goal is to find a set of triangles \mathcal{M}^\ddagger which covers the remaining edges in $G \setminus G[U_k]$ in pairs (each triangle consists of two edges incident to some vertex $v \in U_k \setminus U_{k+1}$, and an edge in $G[N_{U_{k+1}}(v)]$). The triangles we are allowed to use correspond precisely to the edges in $L_v[W_v]$: in the language of our link graphs, our task is to find a perfect matching in each $L_v[W_v]$. However, not just any perfect matchings will do; we need to avoid forbidden configurations and our perfect matchings need to be edge-disjoint.

Our plan is to randomly sparsify the link graphs, and then delete all edges that could possibly cause problems with forbidden configurations or edge-disjointness. If we sparsify sufficiently harshly, we will be able to show that very few subsequent deletions are actually necessary, which will allow us to apply [Lemma 6.4](#).

10.3.2. Typicality of link graphs. For a vertex $v \in U_k \setminus U_{k+1}$, let $N_R(v)$ be the set of vertices $u \in U_{k+1}$ such that $vu \in R$. By iteration-typicality and a Chernoff bound using the randomness of R , each $|N_R(v)| = (1 \pm o(1))pn^{-\theta}|U_{k+1}| = (1 \pm o(1))pn^{1-\rho-\theta}$ and in each link graph L_v :

- each vertex $u \in N_{U_{k+1}}(v)$ has $(1 \pm o(1))p^2qn^{-\theta}|U_{k+1}| = (1 \pm o(1))pq|N_R(v)|$ neighbors in $N_R(v)$, and
- each pair of vertices $u, w \in N_{U_{k+1}}(v)$ have $(1 \pm o(1))p^3q^2n^{-\theta}|U_{k+1}| = (1 \pm o(1))(pq)^2|N_R(v)|$ common neighbors in $N_R(v)$.

To see this, note that the degree of u in L_v is simply the number of vertices $x \in U_{k+1}$ such that uvx is a triangle, and the number of common neighbors of u and w in L_v is simply the number of vertices $x \in U_{k+1}$ such that uvx and wvx are both triangles. These quantities are both controlled by iteration-typicality. Then, $N_R(v)$ can be interpreted as a binomial random subset of $N_{U_{k+1}}(v)$, where each vertex is present with probability $n^{-\theta}$.

Now, as already observed in [Section 10.2.4](#), it follows from iteration-typicality, [A3](#), and [Corollary 3.9](#) that with probability $1 - n^{-\omega(1)}$, for each vertex $v \in U_k \setminus U_{k+1}$ there are at most $n^{o(1)-\beta}p|U_{k+1}|$ edges in $G \setminus R$ which are incident to v and not covered by \mathcal{M}^* . When this is the case, for each v the symmetric difference $N_R(v) \triangle W_v$ has size at most $n^{o(1)-\beta}p|U_{k+1}|$. So, if θ, ν are small with respect to β , then $L_v[W_v]$ has $(1 \pm o(1))pn^{-\theta}n^{1-\rho}$ vertices and is $(pq, o(1))$ -typical (in the sense of [Definition 6.1](#)).

10.3.3. Sparsification, deletion and matching. Recall that $\gamma > 0$ is a constant which is small with respect to g, β, θ, ρ . For each $v \in U_k$ let S_v be the randomly sparsified subgraph of $L_v[W_v]$ obtained by retaining each edge with probability $n^\gamma / (n^{-\theta} p^2 q n^{1-\rho})$ (independently for each v and each edge).

We delete some edges from the sparsified link graphs S_v . First, let D_1 be the set of edges which appear in multiple S_v . Second, recall that every edge of L_v corresponds to a triangle in \mathcal{A} ; let \mathcal{R} be the collection of sets of edges in $\bigcup_v S_v$ whose corresponding triangles create a forbidden configuration when combined with triangles in $\mathcal{I} \cup \mathcal{D} \cup \mathcal{M}^* \cup \mathcal{M}^\dagger$. Let D_2 be the set of all edges in the configurations in \mathcal{R} . If we can find a perfect matching in each $S_v \setminus (D_1 \cup D_2)$, we will be able to take the corresponding triangles as \mathcal{M}^\ddagger .

In order to apply [Lemma 6.4](#) we need to bound the maximum degrees of each $S_v \cap D_1$ and $S_v \cap D_2$. First, [A4](#) implies that w.h.p. every edge appears in at most $n^{1-2\theta+o(1)}$ of our link graphs $L_v[W_v]$. Condition on an outcome of R, \mathcal{M}^* such that this is the case, and fix a vertex $v \in U_k \setminus U_{k+1}$. Conditioning on the event that an edge e appears in S_v , the probability that e appears in some other $S_{v'}$ is at most $n^{-\theta+2\nu+\gamma+\rho+o(1)} = n^{-\Omega(1)}$, by a union bound over choices of v' (recalling that ν, γ, ρ are small compared to θ). By a Chernoff bound (and the independence of our random sparsifications), it follows that with probability $1 - n^{-\omega(1)}$ the maximum degree of $S_v \cap D_1$ is at most $o(n^\gamma)$.

Similarly, using [Lemma 8.8](#) and [Corollary 3.8](#) we can see that with probability $1 - n^{-\omega(1)}$ the maximum degree of $S_v \cap D_2$ is at most $n^{o(1)} = o(n^\gamma)$. Indeed, for each $w \in N_{U_{k+1}}(v)$ we apply [Lemma 8.8](#) with $y = O(1)$, $z = O(n^\beta)$, $r = n^{-\theta}$ and $e = vw$. The weights of the form $\pi_{T,1}$ and $\pi_{T,2}$ describe the probability

that a triangle T is in \mathcal{I} or $\mathcal{D} \cup \mathcal{M}^* \cup \mathcal{M}^\dagger$, and the weights of the form $\pi_{T,3}$ describe the probability that a triangle $T = \{v', u, u'\}$ is such that $uu' \in S_{v'}$. The weights of the form $\pi_{e'}$, for e' between $U_k \setminus U_{k+1}$ and U_{k+1} describe the probability that e' is uncovered by \mathcal{I} and appears in R , and the weights of the form π_e , for e inside U_{k+1} , simply describe the probability that e' is uncovered by \mathcal{I} . Then, with notation as in [Lemma 8.8](#) and [Corollary 3.8](#), $X(\mathfrak{M}_{e,j})$ is an upper bound on the number of edges in D_2 incident to w . The joint probability guarantees that we need to apply [Corollary 3.8](#) are provided by [\(10.1\)](#), and the independently randomly sparsified link graphs $S_{v'}$. (Here we must also ensure $pn^\gamma \leq 1$ to apply [Lemma 8.8](#), which follows from the assumption $p \leq |U_k|^{-\Omega(1)}$ and choosing γ sufficiently small.)

We conclude that, with probability $1 - n^{-\omega(1)}$, the graphs $S_v \setminus (D_1 \cup D_2)$ are each suitable for application of [Lemma 6.4](#) and we can find the desired perfect matchings. [C4](#) follows from the fact that the desired matchings are stochastically dominated by binomial sampling (conditional on suitability).

10.4. Verifying iteration-goodness. Finally, we complete the proof of [Proposition 10.6](#) by verifying that the conditions of $(p, q, o(1))$ -iteration-goodness are satisfied by including the triangles $\mathcal{M} = \mathcal{M}^* \cup \mathcal{M}^\dagger \cup \mathcal{M}^\ddagger$ (specifically, update the data to $(G_{U_{k+1}}(\mathcal{M}), \mathcal{A}_{U_{k+1}}(\mathcal{I} \cup \mathcal{D} \cup \mathcal{M}), \mathcal{I}, \mathcal{D} \cup \mathcal{M})$). The only properties of iteration-goodness that are not obviously satisfied by definition are [IG1](#) (strong well-distributedness) and [IG3](#) (iteration-typicality).

10.4.1. Strong well-distributedness. To show that $(\mathcal{I}, \mathcal{D} \cup \mathcal{M})$ is strongly $(p, O(1), n^{-\omega(1)})$ -well-distributed, we perform a very similar calculation as in [Section 10.2.2](#). Consider any $s, t, r \in \mathbb{N}$, any distinct triangles $I_1, \dots, I_s, D_1, \dots, D_t$ in K_N , and any distinct edges e_1, \dots, e_r in K_N . Let $\mathcal{T} = \{D_1, \dots, D_t\}$, and consider a partition $\mathcal{T} = \mathcal{T}^0 \cup \mathcal{T}^* \cup \mathcal{T}^\dagger \cup \mathcal{T}^\ddagger$.

We are interested in the probability of the event that I_1, \dots, I_s are contained in \mathcal{I} , and e_1, \dots, e_r are uncovered by \mathcal{I} , and $\mathcal{T}^0 \subseteq \mathcal{D}, \mathcal{T}^* \subseteq \mathcal{M}^*, \mathcal{T}^\dagger \subseteq \mathcal{M}^\dagger, \mathcal{T}^\ddagger \subseteq \mathcal{M}^\ddagger$. This probability is at most

$$C^{s+t+r} \left(p^{r+3|\mathcal{T}^*|} N^{-s} \prod_{T \in \mathcal{T}^0} \frac{p}{|U_{\text{lev}(T)}|} + n^{-\omega(1)} \right) \cdot \left((2n^{-\theta})^{2|\mathcal{T}^\ddagger|} + n^{-\omega(1)} \right) \\ \cdot \left(\left(\frac{C^*}{p^2 n} \right)^{|\mathcal{T}^*|} \left(\frac{C^*}{n^\beta} \right)^{|\mathcal{T}^\dagger|} + n^{-\omega(1)} \right) \cdot \left(\frac{C^\dagger}{n^{-2\theta} p^2 n^{1-\rho}} \right)^{|\mathcal{T}^\dagger|} \cdot \left(\frac{n^\gamma}{n^{-\theta} p^2 n^{1-\rho}} \right)^{|\mathcal{T}^\ddagger|}.$$

(Here we have used that the probability a set of $2|\mathcal{T}^\ddagger|$ edges of G between $U_k \setminus U_{k+1}$ and U_{k+1} are all uncovered by \mathcal{M}^* is at most $(2n^{-\theta})^{2|\mathcal{T}^\ddagger|} + n^{-\omega(1)}$. This incorporates the definition of R , and [A3](#), [B4](#), and [C4](#), and assumes that θ is small compared to β).

Summing over all 4^s choices of the partition $\mathcal{T} = \mathcal{T}^0 \cup \mathcal{T}^* \cup \mathcal{T}^\dagger \cup \mathcal{T}^\ddagger$ verifies the desired well-distributedness claim, provided θ, ν, ρ are small with respect to β , and γ, ν, ρ are small compared to θ .

10.4.2. Iteration-typicality. We are assuming that (G, \mathcal{A}) is $(p, q, o(1), 4)$ -iteration-typical, and we need to prove that the updated data $(G_{U_{k+1}}(\mathcal{M}), \mathcal{A}_{U_{k+1}}(\mathcal{I} \cup \mathcal{D} \cup \mathcal{M}))$ is likely to still be $(p, q, o(1), 4)$ -iteration-typical. We will simply show that the process of updating the data affects iteration-typicality in a negligible fashion.

Specifically, let $G^{\text{diff}} = G[U_{k+1}] \setminus G_{U_{k+1}}(\mathcal{M})$ and let $\mathcal{A}^{\text{diff}}$ be the set of triangles in $\mathcal{A} \setminus \mathcal{A}_{U_{k+1}}(\mathcal{I} \cup \mathcal{D} \cup \mathcal{M})$ which lie in the updated graph $G_{U_{k+1}}(\mathcal{M})$. It suffices to show that with probability $1 - n^{-\omega(1)}$, for each $k+1 \leq i < \ell$ and $i^* \in \{i, i+1\}$:

- (1) every vertex $v \in U_i$ has $o(p|U_i|)$ neighbors in U_i and $o(p|U_{i+1}|)$ neighbors in U_{i+1} with respect to G^{diff} , and
- (2) for any edge subset $Q \subseteq G[U_k]$ spanning $|V(Q)| \leq 4$ vertices:
 - (a) there are $o(p^{|V(Q)}|U_{i^*}|)$ vertices $u \in U_{i^*}$ for which there is a triangle $uvw \in \mathcal{A}^{\text{diff}}$ for some $vw \in Q$, and
 - (b) there are $o(p^{|V(Q)}|U_{i^*}|)$ vertices $u \in U_{i^*}$ for which there is an edge $uv \in G^{\text{diff}}$ for some $v \in V(Q)$.

First, (2a) follows from [Lemma 8.9](#) and [Corollary 3.8](#). Indeed, here we are applying [Lemma 8.9](#) with $y = O(1)$, $z = O(n^\beta)$, where the weights of the form $\pi_{T,1}$ and $\pi_{T,2}$ describe the probability that a triangle T is in \mathcal{I} or $\mathcal{D} \cup \mathcal{M}$, respectively, and the weights of the form π_e describe the probability that an edge e is uncovered by \mathcal{I} . With notation as in [Corollary 3.8](#) and [Lemma 8.9](#), $X(\mathfrak{R}_{Q,i^*,j})$ is an upper bound on the number of configurations $\mathcal{S} \in \mathfrak{F}_j$ which cause a triangle containing an edge of Q to contribute to $\mathcal{A}^{\text{diff}}$. The joint probability guarantees that we need to apply [Corollary 3.8](#) are provided by the strong well-distributedness fact we have just proved in the last subsection.

Then, for (1) and (2b) we simply observe that $G^{\text{diff}} \subseteq G^\ddagger$, where G^\ddagger is the set of edges in $E(\mathcal{M}^\ddagger)$ which are inside U_{k+1} . As in [Section 10.3.3](#), using [A4](#) and [C4](#), we can see that G^\ddagger is stochastically dominated by a random subgraph of $G[U_{k+1}]$ in which every edge is present with probability $n^{-\theta+2\nu+\gamma+\rho+o(1)} = o(p^{|V(Q)|}) =$

$o(p)$ independently (we are assuming that γ, ρ, ν are sufficiently small relative to θ). So, (1) and (2b) both follow from a Chernoff bound.

11. PROOF OF THE MAIN THEOREM

Proof of Theorem 1.1. Fix $\beta < \min\{\beta_{9.3}(g, O_g(1)), \beta_{10.6}(g)\}$ (with the same $O_g(1)$ as in the proof of Proposition 9.11), let $\rho = \rho_{10.6}(g, \beta)$, and choose $\ell = O(1)$ such that $M_{4.1}(g) \cdot (1-\rho)^\ell < \beta/(10g)$. Fix a descending sequence of subsets (a ‘‘vortex’’) $V(K_N) = U_0 \supseteq \cdots \supseteq U_\ell$, with $|U_k| = \lfloor |U_{k-1}|^{1-\rho} \rfloor$ for each $1 \leq k \leq \ell$. Then, let $m = |U_\ell|$, and fix an absorbing structure H (as in Theorem 4.1) whose ‘‘flexible set’’ X has size m . By the choice of ℓ this absorbing structure H has at most $N^{\beta/(10g)}$ vertices; embed it in K_N in such a way that X coincides with U_ℓ and all other vertices of H are in $U_0 \setminus U_1$.

Let $G = K_N \setminus H$, let \mathcal{B} be as in Theorem 4.1, and for each $4 \leq j \leq g$, let $\mathfrak{F}_j^{\mathcal{B}}$ be as defined in Lemma 7.2. Note that $|\mathcal{B}|^{2g} \leq ((N^{\beta/(10g)})^3)^{2g} = N^\beta$, and $\mathfrak{F}_j^{\mathcal{B}}$ is $(O(1), |U_k|^\beta)$ -well-spread with respect to $U_0 \supseteq \cdots \supseteq U_k$ for all $0 \leq k \leq \ell$ by Lemma 7.2. Let \mathcal{A} be the set of triangles in G which do not create an Erdős configuration when added to \mathcal{B} .

Let $\nu = \nu_{10.6}(g, \rho)$ and $\tilde{B} = \tilde{B}_{9.11}(g)$, and let $p = \max(m^{-\nu}, m^{-1/\tilde{B}})$. Now, let \mathcal{I} be the partial Steiner triple system obtained from $(1-p)|E(G)|/3$ steps of the generalized high-girth process in Definition 9.4, and let $n = N$. Applying Proposition 9.11 together with the analysis in the proof of Theorem 9.3 (and taking an appropriate union bound over relevant choices of vertices v , vertex sets U_i and edge sets Q) shows that (\mathcal{I}, \emptyset) is strongly $(p, O(1), n^{-\omega(1)})$ -well-distributed and $(p, q, o(1), 4)$ -iteration-typical for some $q = \Omega_g(1)$. It follows that with probability $1 - n^{-\omega(1)}$, the data $(G_{U_0}(\mathcal{I}), \mathcal{A}_{U_0}(\mathcal{I}), \mathcal{I}, \emptyset)$ are $(p, q, o(1), O(1), n^{-\omega(1)})$ -iteration-good, suitable for applying Proposition 10.6, taking $\mathfrak{F}_j = \mathfrak{F}_j^{\mathcal{B}}$. (The required divisibility condition follows from the fact that $N \equiv 1, 3 \pmod{6}$ and H is triangle-divisible.) We now apply Proposition 10.6 repeatedly (ℓ times), to obtain a set of triangles \mathcal{M} in G that covers all edges except those in $G[X]$ and avoids all configurations in $\bigcup_{j=4}^g \mathfrak{F}_j^{\mathcal{B}}$. Finally, the crucial property Ab1 of our absorbing structure H allows us to transform $\mathcal{I} \cup \mathcal{M}$ into a Steiner triple system with girth greater than g , since the leftover graph at the end is triangle-divisible (it obtained from a triangle-divisible graph by removing a collection of edge-disjoint triangles). By the definitions of \mathcal{A} and $\mathfrak{F}_j^{\mathcal{B}}$, this last step does not introduce any forbidden configurations (all triangles in $\mathcal{I} \cup \mathcal{M}$ come from \mathcal{A} , and the only other triangles in our Steiner triple system come from \mathcal{B}). \square

Proof sketch of Theorem 1.3. We proceed in almost exactly the same way as in [35]. To prove Theorem 1.3, we observe that in the above proof of Theorem 1.1, we are (randomly) constructing an *ordered* Steiner triple system (where the triangles in \mathcal{I} are ordered according to the order they are selected in our generalized high-girth triple process, and we arbitrarily order the other triangles to appear after the triangles in \mathcal{I}). Our construction succeeds with probability $1 - o(1)$, and the probability of any particular successful outcome is at most about

$$\prod_{t=1}^{(1-p)|E(G)|/3} \frac{1}{A(t)},$$

where, as in the proof of Proposition 9.11, we let $A(t) = (1/3)|E(t)|f_{\text{edge}}(t)$ be the approximate number of available triangles at step t of the initial high-girth process. So, the number of possible outcomes of our construction is at least about $\prod_{t=1}^{(1-p)|E(G)|/3} A(t)$. We can approximate the logarithm of this product by an integral, and computations as in [27, Section 6] provide a lower bound on the desired number of ordered Steiner triple systems with girth greater than g . Dividing by the number of possible orderings $((N/2)/3)!$ yields the desired result. \square

REFERENCES

- [1] Ben Barber, Stefan Glock, Daniela Kühn, Allan Lo, Richard Montgomery, and Deryk Osthus, *Minimalist designs*, Random Structures Algorithms **57** (2020), 47–63.
- [2] Ben Barber, Daniela Kühn, Allan Lo, and Deryk Osthus, *Edge-decompositions of graphs with high minimum degree*, Adv. Math. **288** (2016), 337–385.
- [3] Ben Barber, Daniela Kühn, Allan Lo, Deryk Osthus, and Amelia Taylor, *Clique decompositions of multipartite graphs and completion of Latin squares*, J. Combin. Theory Ser. A **151** (2017), 146–201.
- [4] Robert A. Bezer, *Counting configurations in designs*, J. Combin. Theory Ser. A **96** (2001), 341–357.
- [5] Robert A. Bezer, *The girth of a design*, J. Combin. Math. Combin. Comput. **40** (2002), 97–113.
- [6] Zoltán L. Blázsik and Zoltán Lóránt Nagy, *Spreading linear triple systems and expander triple systems*, European J. Combin. **89** (2020), 103155, 16.
- [7] Tom Bohman, Alan Frieze, and Eyal Lubetzky, *Random triangle removal*, Adv. Math. **280** (2015), 379–438.
- [8] Tom Bohman and Lutz Warnke, *Large girth approximate Steiner triple systems*, J. Lond. Math. Soc. (2) **100** (2019), 895–913.
- [9] Candida Bowtell and Peter Keevash, *The n -queens problem*, arXiv preprint arXiv:2109.08083 (2021).

- [10] A. E. Brouwer, *Steiner triple systems without forbidden configurations*, Mathematisch Centrum Amsterdam (1977), Technical Report ZW104/77.
- [11] W. G. Brown, P. Erdős, and V. T. Sós, *Some extremal problems on r -graphs*, New directions in the theory of graphs (Proc. Third Ann Arbor Conf., Univ. Michigan, Ann Arbor, Mich, 1971), 1973, pp. 53–63.
- [12] F. R. K. Chung, R. L. Graham, and R. M. Wilson, *Quasi-random graphs*, *Combinatorica* **9** (1989), 345–362.
- [13] Charles J. Colbourn, *The configuration polytope of l -line configurations in Steiner triple systems*, *Math. Slovaca* **59** (2009), 77–108.
- [14] Charles J. Colbourn and Yuichiro Fujiwara, *Small stopping sets in Steiner triple systems*, *Cryptogr. Commun.* **1** (2009), 31–46.
- [15] Charles J. Colbourn, Eric Mendelsohn, Alexander Rosa, and Jozef Širáň, *Anti-mitre Steiner triple systems*, *Graphs Combin.* **10** (1994), 215–224.
- [16] Jean Doyen, *Linear spaces and Steiner systems*, Geometries and groups (Berlin, 1981), Lecture Notes in Math., vol. 893, Springer, Berlin-New York, 1981, pp. 30–42.
- [17] David Ellis and Nathan Linial, *On regular hypergraphs of high girth*, *Electron. J. Combin.* **21** (2014), Paper 1.54, 17.
- [18] P. Erdős, A. Gyárfás, and L. Pyber, *Vertex coverings by monochromatic cycles and trees*, *J. Combin. Theory Ser. B* **51** (1991), 90–95.
- [19] Paul Erdős, *Problems and results in combinatorial analysis*, Colloquio Internazionale sulle Teorie Combinatorie (Rome, 1973), Tomo II, Accad. Naz. Lincei, 1976, pp. 3–17. Atti dei Convegni Lincei, No. 17.
- [20] A. D. Forbes, M. J. Grannell, and T. S. Griggs, *On 6-sparse Steiner triple systems*, *J. Combin. Theory Ser. A* **114** (2007), 235–252.
- [21] David A. Freedman, *On tail probabilities for martingales*, *Ann. Probability* **3** (1975), 100–118.
- [22] Yuichiro Fujiwara, *Infinite classes of anti-mitre and 5-sparse Steiner triple systems*, *J. Combin. Des.* **14** (2006), 237–250.
- [23] Yuichiro Fujiwara, *Nonexistence of sparse triple systems over abelian groups and involutions*, *J. Algebraic Combin.* **26** (2007), 495–506.
- [24] Zoltán Füredi and Miklós Ruszinkó, *Uniform hypergraphs containing no grids*, *Adv. Math.* **240** (2013), 302–324.
- [25] Stefan Glock, *Triple systems with no three triples spanning at most five points*, *Bull. Lond. Math. Soc.* **51** (2019), 230–236.
- [26] Stefan Glock, Daniela Kühn, Allan Lo, Richard Montgomery, and Deryk Osthus, *On the decomposition threshold of a given graph*, *J. Combin. Theory Ser. B* **139** (2019), 47–127.
- [27] Stefan Glock, Daniela Kühn, Allan Lo, and Deryk Osthus, *On a conjecture of Erdős on locally sparse Steiner triple systems*, *Combinatorica* **40** (2020), 363–403.
- [28] Stefan Glock, Daniela Kühn, Allan Lo, and Deryk Osthus, *The existence of designs via iterative absorption: hypergraph F -designs for arbitrary F* , *Mem. Amer. Math. Soc.* (to appear).
- [29] Stefan Glock, Daniela Kühn, and Deryk Osthus, *Extremal aspects of graph and hypergraph decomposition problems*, *Surveys in combinatorics 2021*, London Math. Soc. Lecture Note Ser., vol. 470, Cambridge Univ. Press, Cambridge, 2021, pp. 235–265.
- [30] David A. Grable and Kevin T. Phelps, *Random methods in design theory: a survey*, *J. Combin. Des.* **4** (1996), 255–273.
- [31] Elliot Granath, András Gyárfás, Jerry Hardee, Trent Watson, and Xiaoze Wu, *Ramsey theory on Steiner triples*, *J. Combin. Des.* **26** (2018), 5–11.
- [32] M. J. Grannell, T. S. Griggs, and C. A. Whitehead, *The resolution of the anti-Pasch conjecture*, *J. Combin. Des.* **8** (2000), 300–309.
- [33] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000.
- [34] Felix Joos, Jaehoon Kim, Daniela Kühn, and Deryk Osthus, *Optimal packings of bounded degree trees*, *J. Eur. Math. Soc. (JEMS)* **21** (2019), 3573–3647.
- [35] Peter Keevash, *Counting designs*, *J. Eur. Math. Soc. (JEMS)* **20** (2018), 903–927.
- [36] Peter Keevash, *The existence of designs II*, arXiv preprint arXiv:1802.05900 (2018).
- [37] Peter Keevash, *The existence of designs*, arXiv preprint arXiv:1401.3665 (2019).
- [38] Peter Keevash and Jason Long, *The Brown–Erdős–Sós conjecture for hypergraphs of large uniformity*, arXiv preprint arXiv:2007.14824 (2020).
- [39] Jeong Han Kim and Van H Vu, *Concentration of multivariate polynomials and its applications*, *Combinatorica* **20** (2000), 417–434.
- [40] Thomas P Kirkman, *On a problem in combinations*, *Cambridge and Dublin Mathematical Journal* **2** (1847), 191–204.
- [41] Fiachra Knox, Daniela Kühn, and Deryk Osthus, *Edge-disjoint Hamilton cycles in random graphs*, *Random Structures Algorithms* **46** (2015), 397–445.
- [42] Michael Krivelevich, Matthew Kwan, Po-Shen Loh, and Benny Sudakov, *The random k -matching-free process*, *Random Structures Algorithms* **53** (2018), 692–716.
- [43] Daniela Kühn and Deryk Osthus, *Hamilton decompositions of regular expanders: a proof of Kelly’s conjecture for large tournaments*, *Adv. Math.* **237** (2013), 62–146.
- [44] Matthew Kwan, Ashwin Sah, and Mehtaab Sawhney, *Large deviations in random Latin squares*, arXiv preprint arXiv:2106.11932 (2021).
- [45] Hanno Lefmann, Kevin T. Phelps, and Vojtěch Rödl, *Extremal problems for triple systems*, *J. Combin. Des.* **1** (1993), 379–394.
- [46] A. C. H. Ling, C. J. Colbourn, M. J. Grannell, and T. S. Griggs, *Construction techniques for anti-Pasch Steiner triple systems*, *J. London Math. Soc.* (2) **61** (2000), 641–657.
- [47] Alan C. H. Ling, *A direct product construction for 5-sparse triple systems*, *J. Combin. Des.* **5** (1997), 443–447.
- [48] Nati Linial, *Challenges of high-dimensional combinatorics*, Laszlo Lovasz 70th birthday conference, Budapest, <https://www.cs.huji.ac.il/~nati/PAPERS/challenges-hdc.pdf>, 2018.
- [49] Richard Montgomery, *Fractional clique decompositions of dense partite graphs*, *Combin. Probab. Comput.* **26** (2017), 911–943.
- [50] Simón Piga and Nicolás Sanhueza-Matamala, *Cycle decompositions in 3-uniform hypergraphs*, arXiv preprint arXiv:2101.12205 (2021).

- [51] Vojtěch Rödl, *On a packing and covering problem*, European J. Combin. **6** (1985), 69–78.
- [52] Vojtěch Rödl and Andrzej Ruciński, *Threshold functions for Ramsey properties*, J. Amer. Math. Soc. **8** (1995), 917–942.
- [53] Vojtěch Rödl, Andrzej Ruciński, and Endre Szemerédi, *A Dirac-type theorem for 3-uniform hypergraphs*, Combin. Probab. Comput. **15** (2006), 229–251.
- [54] Vojtěch Rödl and Luboš Thoma, *Asymptotic packing and the random greedy algorithm*, Random Structures Algorithms **8** (1996), 161–177.
- [55] I. Z. Ruzsa and E. Szemerédi, *Triple systems with no six points carrying three triangles*, Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II, Colloq. Math. Soc. János Bolyai, vol. 18, North-Holland, Amsterdam-New York, 1978, pp. 939–945.
- [56] Eli Shamir and Benny Sudakov, *Two-sided, unbiased version of Hall’s marriage theorem*, Amer. Math. Monthly **124** (2017), 79–80.
- [57] Chong Shangquan and Itzhak Tamo, *Degenerate Turán densities of sparse hypergraphs*, J. Combin. Theory Ser. A **173** (2020), 105228, 25.
- [58] Asaf Shapira and Mykhaylo Tyomkyn, *A Ramsey variant of the Brown-Erdős-Sós conjecture*, Bull. Lond. Math. Soc. **53** (2021), 1453–1469.
- [59] Joel Spencer, *Asymptotic packing via a branching process*, Random Structures Algorithms **7** (1995), 167–172.
- [60] Adam Wolfe, *5-sparse Steiner triple systems of order n exist for almost all admissible n* , Electron. J. Combin. **12** (2005), Research Paper 68, 42.

INSTITUTE OF SCIENCE AND TECHNOLOGY AUSTRIA (IST AUSTRIA), 3400 KLOSTERNEUBURG, AUSTRIA
Email address: `matthew.kwan@ist.ac.at`

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MA 02139, USA
Email address: `{asah,msawhney}@mit.edu`

CENTER OF MATHEMATICAL SCIENCES AND APPLICATIONS, HARVARD UNIVERSITY, CAMBRIDGE, MA 02138, USA
Email address: `msimkin@cmsa.fas.harvard.edu`