Almost all Steiner Triple Systems have perfect matchings

Matthew Kwan *

Abstract

We show that for any n divisible by 3, almost all order-n Steiner triple systems have a perfect matching (also known as a parallel class). In fact, we prove a general upper bound on the number of perfect matchings in a Steiner triple system and show that almost all Steiner triple systems essentially attain this maximum. We accomplish this via a general theorem comparing a uniformly random Steiner system to the outcome of a random clique removal process, which we hope will be useful for other problems. We believe our methods can be easily adapted to other types of designs, for example to show that almost all Latin squares have transversals.

1 Introduction

A *Steiner system* of order n with parameters (q, r), $q > r$, is a collection S of size-q subsets of $[n] = \{1, \dots, n\}$ (that is, a q-uniform hypergraph on the vertex set $[n]$), such that every size-r subset of $[n]$ is included in exactly one element (hyperedge) of S. Steiner systems are among the most fundamental types of combinatorial designs, and have strong connections to a wide range of different subjects, ranging from group theory, to finite geometry, to experimental design, to the theory of error-correcting codes. See [26, Chapter 19] for an introduction to the subject.

A $(q, 1)$-Steiner system is a partition of $[n]$ into subsets of size q. In the language of hypergraphs, it is precisely a q-uniform perfect matching. Steiner systems with all other combinations of parameters are notoriously difficult to study. The simplest case is that of $(3, 2)$-Steiner systems, known as *Steiner triple systems*. A Steiner triple system is actually nothing more than a triangle-decomposition of the edges of the complete graph K_n, so Steiner triple systems are natural "symmetric" counterparts to *Latin squares*, which can be defined as triangle-decompositions of the complete tripartite graph $K_{n,n,n}$.

In 1974 Wilson [28] used estimates for the number of Latin squares to prove a coarse estimate for the number of Steiner triple systems. Babai [3] used this estimate to prove that a $(1 - o(1))$-proportion of Steiner triple systems have trivial automorphism group (equivalently, a uniformly random order-n Steiner triple system a.a.s.[1] has trivial automorphism group). We believe this is the only nontrivial property known to hold a.a.s. for random Steiner systems with any combination of parameters. Following Erdős and Rényi's seminal paper [6] on random graphs and Erdős' popularization of the probabilistic method, there have been great developments in the theory of random combinatorial structures of all kinds, but essentially none of the tools developed seem to be applicable to Steiner systems. Steiner systems lack independence or any kind of recursive structure, which rules out many of the techniques used to study Erdős-Rényi random graphs and random permutations, and there is basically no freedom to make local changes, which precludes the use of "switching" techniques often used in the study of random regular graphs (see for example [15]). It is

*Department of Mathematics, ETH, 8092 Zürich. Email: matthew.kwan@math.ethz.ch.

[1] By "asymptotically almost surely", or "a.a.s.", we mean that the probability of an event is $1 - o(1)$. Here and for the rest of the paper, asymptotics are as $n \to \infty$.

not even clear how to study random Steiner systems empirically; in an attempt to find an efficient algorithm to generate a random Steiner triple system, Cameron [5] designed a Markov chain on Steiner triple systems, but he was not able to determine whether this chain was connected.

In a huge breakthrough that will surely revolutionize design theory, Keevash [12] very recently proved that "partial" Steiner systems satisfying a certain "quasirandomness" condition can be completed into Steiner systems. He used this to settle a longstanding existence question for Steiner systems, and shortly afterwards [13] he proved new estimates for the number of Steiner systems with each combination of parameters. In particular he proved that there are

$$\big(n/e^2 + o(n)\big)^{n^2/6} \tag{1}$$

Steiner triple systems of order n, as long as n satisfies a necessary divisibility condition (Steiner triple systems can only exist if n is 1 or 3 mod 6).

Of course, this new estimate makes it possible, in theory, to prove new a.a.s. properties of random Steiner triple systems just by giving an estimate asymptotically smaller than (1) for the number of Steiner triple systems not satisfying a certain property. However, for most properties it is not at all clear how to prove such estimates. Instead, we introduce a way to use Keevash's methods to show that a uniformly random Steiner triple system can in some sense be approximated by the outcome of a random process called the *triangle removal process*. We remark that actually Keevash's lower bound is proved via a randomized construction that involves the triangle removal process, so many properties that hold a.a.s. in the triangle removal process trivially hold a.a.s. in this random construction. Such results have been proved in [17, Proposition 3.1] and [18]. However, the Steiner triple systems obtainable by Keevash's construction comprise a negligible proportion of the set of Steiner triple systems, so a more delicate approach is required to study a uniformly random Steiner triple system. We give the details in Section 2.

As an application of our new method, we prove that if $3 \mid n$ (that is, if $n \equiv 3 \bmod 6$) then almost all order-n Steiner triple systems have many perfect matchings. The existence of perfect matchings is one of the most central questions in the theory of graphs and hypergraphs; in particular, one of the most celebrated recent developments in the field is the Fulkerson-prize-winning work of Johansson, Kahn and Vu [11] on perfect matchings in random hypergraphs. A perfect matching in a Steiner triple system is also called a *parallel class*, and has particular significance. One of the oldest problems in combinatorics, famously solved by Ray-Chaudhuri and Wilson [23], asks whether for all $n \equiv 3 \bmod 6$ there exists an order-n Steiner triple system which can be partitioned into hyperedge-disjoint perfect matchings (a *Kirkman triple system*). Alon, Kim and Spencer [1] proved that every Steiner triple system has an almost-perfect matching covering all but $o\Big(\sqrt{n}\log^{3/2} n\Big)$ vertices, and Bryant and Horsley [4] proved that for infinitely many $n \equiv 3 \bmod 6$ there exist Steiner triple systems with no perfect matching.

Theorem 1.1. *Let $n \equiv 3 \bmod 6$ and let $\boldsymbol{S}$ be a uniformly random order-n Steiner triple system. Then a.a.s. $\boldsymbol{S}$ contains*

$$\Big((1-o(1))\frac{n}{2e^2}\Big)^{n/3}$$

perfect matchings.

We remark that if $3 \nmid n$ (that is, if $n \equiv 1 \bmod 6$) then obviously no order-n Steiner triple system can have a perfect matching, but exactly the same proof can be used to show that there is a matching covering all but one vertex.

We prove Theorem 1.1 in Section 3 using our new method combined with the so-called *absorbing method*, which was introduced as a general method by Rödl, Ruciński and Szemerédi [24] (though

the basic idea had been used earlier, for example by Krivelevich [14]). Basically, we prove the a.a.s. existence of certain small substructures that allow us to complete an almost-perfect matching into a perfect one.

Up to the error term, a random Steiner triple system actually has the maximum possible number of perfect matchings: we also prove the following upper bound.

Theorem 1.2. *Any Steiner triple system has at most*

$$\Big((1+o(1))\frac{n}{2e^2}\Big)^{n/3}$$

perfect matchings.

We give a short proof of Theorem 1.2 in Section 4, using the so-called *entropy method* due to Radhakrishnan [22].

1.1 Latin squares

An order-n Latin square is usually defined as an $n\times n$ array of the numbers between 1 and n (we call these *symbols*), such that each row and column contains each symbol exactly once. As mentioned earlier, this is equivalent to a 3-uniform hypergraph whose triples comprise a triangle-decomposition of the edges of the complete tripartite graph $K_{n,n,n}$ (the three parts correspond to the rows, columns and symbols, so a triangle (i,j,k) corresponds to putting the symbol k in the cell (i,j)). A perfect matching in a Latin square is called a *transversal* and the property of containing a transversal is of great interest. In particular, the famous Ryser-Brualdi-Stein conjecture speculates that every odd-order Latin square has a transversal, and every even-order Latin square has a partial transversal (matching) of size $n-1$. See the survey of Wanless [27] for more information.

We are quite confident that our methods can be adapted to the setting of Latin squares. The main obstacle is that Keevash's completion results have not yet been adapted to Latin squares. We believe such an adaptation should be quite straightforward, but it would be well outside the scope of this paper to include the details here. Modulo a Keevash-type completion result we can prove the following theorem, answering a question of Glebov and Luria [8] and proving that the Ryser-Brualdi-Stein conjecture holds for almost all Latin squares.

"Theorem" 1.3. *Let $\boldsymbol{L}$ be a uniformly random order-n Latin square. Then a.a.s. $\boldsymbol{L}$ contains*

$$\Big((1-o(1))\frac{n}{e^2}\Big)^{n}$$

transversals.

We note that the counterpart of Theorem 1.2 for Latin squares (that a Latin square can have at most $\big((1+o(1))n/e^2\big)^n$ transversals) was first proved by Taranenko [25]. In Section 5 we outline how the proof of Theorem 1.1 can be adapted to prove "Theorem" 1.3, conditional on a Keevash-type completion result.

1.2 Structure of the paper

The structure of this paper is as follows. In Section 2 we explain our method for studying random Steiner triple systems (in fact, we do everything in the general setting of (q,r)-Steiner systems, because this does not introduce any difficulties). In Section 3 we define absorbers and prove Theorem 1.1, in Section 4 we prove Theorem 1.2, and in Section 5 we sketch how "Theorem" 1.3 may

be proved. In Section 6 we have some concluding remarks, including a long list of open problems. Finally, we have two appendices; in Appendix A we provide a straightforward but necessary adaptation of Keevash's proof of (1), to estimate the number of completions of a partial Steiner system, and in Appendix B we have a very simple analysis of a clique removal process generalizing the triangle removal process.

1.3 Acknowledgements

The author would like to thank Asaf Ferber and Benny Sudakov for very helpful discussions about random designs and the intricacies of the absorbing method. Asaf and Benny could both have deservedly been coauthors on this paper, but graciously declined the opportunity.

2 Random Steiner systems via clique removal processes

In this section we describe our method for comparing random Steiner triple systems with the outcome of the triangle removal process. Actually, the situation for general (q,r)-Steiner systems is essentially the same, so we state everything in general terms in case this is useful in the future. For this entire section, q and r should be considered fixed, and whenever we say "Steiner system" we mean "(q,r)-Steiner system". Let $Q = \binom{q}{r}$ and let $N = \binom{n}{r}/Q = (1+o(1))((q-r)!/q!)n^r$ be the number of hyperedges in a Steiner system. We also assume n satisfies the condition that $\binom{q-i}{r-i}$ divides $\binom{n-i}{r-i}$ for every $0 \le i \le r-1$, which is necessary for the existence of an order-n Steiner system.

We will shortly state a general theorem for comparing random Steiner systems with clique removal processes, but first we need several definitions.

Definition 2.1 (partial Steiner systems)**.** A *partial Steiner system* (or *partial system* for short) is a q-uniform hypergraph on $[n]$ in which every r-set of vertices is included in no more than one hyperedge. Let $\mathcal{S}_m$ be the set of partial systems with m hyperedges. We will also want to consider partial systems equipped with an ordering on their hyperedges. Let $\mathcal{O}$ be the set of ordered Steiner systems, and let $\mathcal{O}_m$ be the set of ordered partial systems with m hyperedges. For $S \in \mathcal{O}_m$ and $i \le m$, let S_i be the partial system consisting of the first i hyperedges of S. For a (possibly ordered) partial system S, let $G(S)$ be the r-uniform hypergraph with an edge for every pair of vertices which does not appear in any hyperedge of S. So, if S has m hyperedges, then $G(S)$ has $\binom{n}{r} - Qm$ edges.

Definition 2.2 (quasirandomness)**.** For an r-graph G with m edges, let $d(G) = m/\binom{n}{r}$ denote its density. For an $(r-1)$-set of vertices W in an r-graph G, the *neighbourhood* $N_G(W)$ is the set of vertices v such that $\{v\} \cup W$ is an edge of G. We say G is (ε, h)-*quasirandom* if for every set A of $(r-1)$-sets of vertices with $|A| \le h$, we have $\left|\bigcap_{W\in A} N_G(W)\right| = (1\pm\varepsilon)d(G)^{|A|}n$. (The notation $f = 1\pm\varepsilon$ means $1-\varepsilon \le f \le 1+\varepsilon$). Let $\mathcal{S}_m^{\varepsilon,h} \subseteq \mathcal{S}_m$ be the set of partial systems $S \in \mathcal{S}_m$ such that $G(S)$ is (ε,h)-quasirandom, and let $\mathcal{O}_m^{\varepsilon,h} \subseteq \mathcal{O}_m$ be the set of ordered partial systems $S \in \mathcal{O}_m$ such that $S_i \in \mathcal{S}_i^{\varepsilon,h}$ for each $i \le m$.

Definition 2.3 (the clique removal processs)**.** The clique removal process is defined as follows. Start with the complete r-graph $K_n^{(r)}$ and iteratively delete a copy of $K_q^{(r)}$ chosen uniformly at random from all copies of $K_q^{(r)}$ in the remaining r-graph. If we continue this process for m steps, the deleted copies of $K_q^{(r)}$ (in order) can be interpreted as an ordered partial system in $\mathcal{O}_m$. It is also possible that the process aborts (because there are no copies of $K_q^{(r)}$ left) before m steps, in which case we say it returns the value "$*$". We denote by $\mathbb{R}(n,m)$ the resulting distribution on $\mathcal{O}_m \cup \{*\}$.

Now, our general theorem is as follows.

Theorem 2.4. *There is $h_0 \in \mathbb{N}$ such that for fixed $h \geq h_0$ and sufficiently small a there is $b(a) > 0$ such that the following holds. Fix $\alpha \in (0,1)$, let $\mathcal{P} \subseteq \mathcal{O}_{\alpha N}$ be a property of ordered partial systems, let $\mathcal{Q} \supseteq \mathcal{O}_{\alpha N}^{n^{-a},h}$, let $\boldsymbol{S} \in \mathcal{O}$ be uniformly random and let $\boldsymbol{S}' \in \mathbb{R}(n, \alpha N)$. If*

$$\Pr\big(\boldsymbol{S}' \notin \mathcal{P} \,\big|\, \boldsymbol{S}' \in \mathcal{Q}\big) \leq \exp\Big(-n^{r-b}\Big)$$

then

$$\Pr(\boldsymbol{S}_{\alpha N} \notin \mathcal{P}) \leq \exp\big(-\Omega\big(n^{1-2a}\big)\big).$$

Note that (as we prove in Appendix B), the clique removal process is likely to produce quasirandom graphs; that is, $\Pr(\boldsymbol{S}' \in \mathcal{Q}) = 1 - o(1)$. However, as we will see in Section 3.1.1, the conditioning in Theorem 2.4 can still be useful because the probabilities under consideration are so small (it is certainly not true that $\Pr(\boldsymbol{S}' \notin \mathcal{Q})$ is anywhere near as small as $\exp\big(-\Omega\big(n^2\big)\big)$).

The proof of Theorem 2.4 follows from a sequence of several lemmas. The most important is the following: we can use Keevash's methods to estimate the number of ways to complete any $S \in \mathcal{O}_{\alpha N}^{\varepsilon,h}$, and show that it does not vary too much between choices of S.

Lemma 2.5. *There is $h \in \mathbb{N}$ such that for any fixed $a > 0$, there is $b = b(a) > 0$ such that the following holds. For any fixed $\alpha \in (0,1)$, any $\varepsilon \leq n^{-a}$, and any $S, S' \in \mathcal{O}_{\alpha N}^{\varepsilon,h}$,*

$$\frac{\big|\mathcal{O}^{\mathrm{ext}}(S)\big|}{\big|\mathcal{O}^{\mathrm{ext}}(S')\big|} \leq \exp\Big(n^{r-b}\Big).$$

Lemma 2.5 can be proved with only very slight adaptations to the proof of (1) in [13]. The details are in Appendix A. The point of Lemma 2.5 is that if we can prove some property holds with extremely high probability (say $1 - \exp(-\Omega(n^r))$) in a uniformly random $\boldsymbol{S} \in \mathcal{O}_{\alpha N}^{\varepsilon,h}$, then it also holds with essentially the same probability in $\boldsymbol{S}_{\alpha N}$, for a uniformly random $\boldsymbol{S} \in \mathcal{O}$ conditioned on the event $\boldsymbol{S}_{\alpha N} \in \mathcal{O}_{\alpha N}^{\varepsilon,h}$.

The next step is to show that the event $\boldsymbol{S}_{\alpha N} \in \mathcal{O}_{\alpha N}^{\varepsilon,h}$ is very likely. In fact, this event occurs a.a.s. for a random ordering of any given Steiner system. We prove the following lemma in Section 2.1.

Lemma 2.6. *The following holds for any fixed $h \in \mathbb{N}$, $\alpha \in (0,1)$ and $a \in (0,1/2)$. Let $\varepsilon = n^{-a}$, consider any Steiner system S, and uniformly at random order its hyperedges to obtain an ordered Steiner system $\boldsymbol{S} \in \mathcal{O}$. Then $\Pr\Big(\boldsymbol{S}_{\alpha N} \notin \mathcal{O}_{\alpha N}^{\varepsilon,h}\Big) = \exp\big(-\Omega\big(n^{1-2a}\big)\big)$.*

Therefore, if we can prove a property holds with extremely high probability in a uniformly random $\boldsymbol{S} \in \mathcal{O}_{\alpha N}^{\varepsilon,h}$ for sufficiently small ε and sufficiently large h, then that property also holds a.a.s. in the first αN edges of a uniformly random $\boldsymbol{S} \in \mathcal{O}$.

Next, the following lemma says that each $S \in \mathcal{O}_{\alpha N}^{\varepsilon,h}$ is roughly equally likely to be produced by the clique removal process, so that $\mathbb{R}(n, \alpha N)$ approximates the uniform distribution on $S \in \mathcal{O}_{\alpha N}^{\varepsilon,h}$. It is proved in Section 2.2.

Lemma 2.7. *For any $a > 0$ and $\alpha \in [0,1]$, and for $S, S' \in \mathcal{O}_{\alpha N}^{n^{-a},q-1}$ and $\boldsymbol{S} \in \mathbb{R}(n, \alpha N)$,*

$$\frac{\Pr(\boldsymbol{S} = S)}{\Pr(\boldsymbol{S} = S')} \leq \exp\big(O\big(n^{r-a}\big)\big).$$

We can finally combine everything to prove Theorem 2.4.

Proof of Theorem 2.4. We have

$$\begin{aligned}\Pr(\boldsymbol{S}_{\alpha N} \notin \mathcal{P}) &\le \Pr\Big(\boldsymbol{S}_{\alpha N} \notin \mathcal{P} \,\Big|\, \boldsymbol{S}_{\alpha N} \in \mathcal{O}_{\alpha N}^{n^{-a},h}\Big) + \Pr\Big(\boldsymbol{S}_{\alpha N} \notin \mathcal{O}_{\alpha N}^{n^{-a},h}\Big)\\ &\le \exp\big(O\big(n^{r-c}\big)\big)\Pr\Big(\boldsymbol{S}' \notin \mathcal{P} \,\Big|\, \boldsymbol{S}' \in \mathcal{O}_{\alpha N}^{n^{-a},h}\Big) + \exp\big(-\Omega\big(n^{1-2a}\big)\big),\end{aligned}$$

where $c = b(a)$ in the notation of Lemma 2.5. But, if a is small enough then Theorem B.1 guarantees that $\Pr\Big(\boldsymbol{S}' \in \mathcal{O}_{\alpha N}^{n^{-a},h}\Big) = 1 - o(1)$ so

$$\Pr\Big(\boldsymbol{S}' \notin \mathcal{P} \,\Big|\, \boldsymbol{S}' \in \mathcal{O}_{\alpha N}^{n^{-a},h}\Big) \le \Pr\big(\boldsymbol{S}' \notin \mathcal{P} \,\big|\, \boldsymbol{S}' \in \mathcal{Q}\big)\frac{\Pr(\boldsymbol{S}' \in \mathcal{Q})}{\Pr\Big(\boldsymbol{S}' \in \mathcal{O}_{\alpha N}^{n^{-a},h}\Big)} = (1+o(1))\Pr\big(\boldsymbol{S}' \notin \mathcal{P} \,\big|\, \boldsymbol{S}' \in \mathcal{Q}\big).$$

If $b < c$ this completes the proof. □

In Section 2.1 we prove Lemma 2.6 and in Section 2.2 we prove Lemma 2.7. Also, in Lemma 2.10 we prove some lemmas which are useful tools for applying Theorem 2.4 in practice.

2.1 Randomly ordered Steiner systems

In this subsection we prove Lemma 2.6.

Proof. Fix $m \le \alpha N$. Note that $\boldsymbol{S}_m$ (as an unordered partial system) is a uniformly random subset of m hyperedges of S. We can obtain an almost equivalent random partial system by including each hyperedge of S with independent probability m/N. Let $\boldsymbol{S}'$ denote the partial system so obtained, and let $\boldsymbol{G}' = G(\boldsymbol{S}')$. Now, fix a set A of $(r-1)$-sets of vertices with $|A| \le h$. It suffices to prove

$$\left|\bigcap_{W\in A} N_{\boldsymbol{G}'}(W)\right| = \big(1+O\big(n^{-a}\big)\big)\Big(1-\frac{m}{N}\Big)^{|A|}n, \tag{2}$$

with probability $1-\exp\big(-\Omega\big(n^{1-2a}\big)\big)$. Indeed, the so-called Pittel inequality (see [9, p. 17]) would imply that the same estimate holds with essentially the same probability if we replace $\boldsymbol{S}'$ with $\boldsymbol{S}_m$, and then we can apply the union bound over all $m \le \alpha N$ and all choices of A.

Note that there are at most $\binom{|A|}{2} = O(1)$ hyperedges of S that include more than one set in A (this follows from the Steiner property, because the intersection of two sets in A has size at least r). Let U be the set of vertices involved in these atypical hyperedges, plus the vertices that appear in sets in A, so that $|U| = O(1)$. Let $\boldsymbol{N} = \big|\big(\bigcap_{W\in A} N_{\boldsymbol{G}'}(W)\big)\backslash U\big|$. For every $v \notin U$ and $W \in A$ there is exactly one hyperedge e_v^W in S containing v and W, whose presence in $\boldsymbol{S}'$ would prevent v from contributing to $\boldsymbol{N}$. For each $v \notin U$ the hyperedges e_v^W are distinct, so

$$\Pr\left(v \in \bigcap_{W\in A} N_{\boldsymbol{G}'}(W)\right) = \Big(1-\frac{m}{N}\Big)^{|A|},$$

and by linearity of expectation $\mathbb{E}\boldsymbol{N} = (1-m/N)^{|A|}(n-O(1))$. Now, $\boldsymbol{N}$ is determined by the presence of at most $(n-|U|)|A| = O(n)$ hyperedges in $\boldsymbol{S}'$, and changing the presence of each affects $\boldsymbol{N}$ by at most $q-1 = O(1)$. So, by the Azuma-Hoeffding inequality (see [9, Section 2.4]),

$$\begin{aligned}\Pr\left(\left|\boldsymbol{N} - \Big(1-\frac{m}{N}\Big)^{|A|}n\right| > n^{-a}\Big(1-\frac{m}{N}\Big)^{|A|}n\right) &\le \exp\left(-\Omega\left(\frac{\Big(n^{-a}(1-\alpha)^h n\Big)^2}{n}\right)\right)\\ &= \exp\big(-\Omega\big(n^{1-2a}\big)\big).\end{aligned}$$

Finally, we recall that $\big|\big(\bigcap_{x\in X} N_{\boldsymbol{G}'}(x)\big)\big| = \boldsymbol{N} + O(1)$, which completes the proof of (2). □

2.2 Approximate uniformity of the clique removal process

In this subsection we prove Lemma 2.7. We first make a simple observation about small subgraph statistics in quasirandom hypergraphs, which we will use at several points in the paper.

Proposition 2.8. *Let H be a fixed r-graph with an identified vertex subset U, and let $|\mathrm{Aut}(H,U)|$ be the number of automorphisms of H fixing U. Let G be an $(\varepsilon, v(H)-1)$-quasirandom r-graph on n vertices, and let F be a copy of the induced r-graph $H[U]$ in G. Then, the number of completions of F to a copy of H is*

$$(1\pm O(\varepsilon))\left(1-\frac{i}{N}\right)^{e(H)-e(F)}\frac{n^{q-|U|}}{|\mathrm{Aut}(H,U)|}.$$

In particular, the number of copies of H in G is

$$(1\pm O(\varepsilon))\left(1-\frac{i}{N}\right)^{e(H)}\frac{n^{q}}{|\mathrm{Aut}(H)|}.$$

Proof. It suffices to show that the number of embeddings of H (as a labelled object) into G extending F is

$$(1\pm O(\varepsilon))\left(1-\frac{i}{N}\right)^{e(H)-e(F)}n^{q-|V|}.$$

We prove this by induction on the number of vertices in $V(H)\backslash U$. The base case is where $U=V(H)$, which is trivial. Fix a vertex $v\in V(H)\backslash U$; by induction there are

$$(1\pm O(\varepsilon))\left(1-\frac{i}{N}\right)^{e(H)-e(F)-d_H(v)}n^{q-|V|-1}$$

embeddings of $H-v$ extending F, where $d_H(v)$ is the number of edges of H containing v. For each such embedding, by $(\varepsilon, v(H)-1)$-quasirandomness, there are $(1\pm\varepsilon)\big(1-\frac{i}{N}\big)^{d_H(v)}n$ ways to choose a vertex of G with the right adjacencies to complete the embedding of H. □

Now we are ready to prove Lemma 2.7.

Proof. Each $G(S_i)$ has

$$\big(1\pm O\big(n^{-a}\big)\big)\left(1-\frac{i}{N}\right)^{Q}\frac{n^q}{q!}$$

copies of $K_q^{(r)}$, by $(n^{-a}, q-1)$-quasirandomness and Proposition 2.8. We therefore have

$$\Pr(\boldsymbol{S}=S)=\prod_{i=0}^{\alpha N-1}\frac{1}{(1\pm O(n^{-a}))(1-i/N)^Q n^q/q!},$$

and a similar expression holds for $\Pr(\boldsymbol{S}=S')$. Taking quotients term-by-term gives

$$\begin{aligned}\frac{\Pr(\boldsymbol{S}=S)}{\Pr(\boldsymbol{S}=S')}&\le\big(1+O\big(n^{-a}\big)\big)^{\alpha N}\\&\le\exp\big(O\big(n^{r-a}\big)\big)\end{aligned}$$

as desired. □

2.3 A coupling lemma and a concentration inequality

In this subsection we prove two lemmas that will be useful in combination with Theorem 2.4. First, after some necessary definitions we will show how to couple the clique removal process with a simpler random hypergraph distribution.

Definition 2.9. For a partial system S, let $\mathbb{G}(S,p)$ be the random distribution on q-uniform hypergraphs where each hyperedge not conflicting with S (that is, not intersecting a hyperedge of S in more than $r-1$ vertices) is included with probability p. So, if $\varnothing$ is the empty order-n partial system, then $\mathbb{G}(\varnothing,p) := \mathbb{G}(n,p)$ is the standard binomial random r-graph. Let $\mathbb{G}^*(S,p)$ be the distribution on partial systems obtained from $\mathbb{G}(S,p)$ by considering all hyperedges which intersect another hyperedge in more than $r-1$ vertices, and deleting all these hyperedges. Let $\mathbb{R}(S,m)$ be the partial system distribution obtained with m steps of the clique removal process starting from $G(S)$.

Let $P = n^{r-q}(q-r)!$ so that $\binom{n}{q}P = (1+o(1))N$. For small $\alpha > 0$, we can view $\mathbb{G}^*(S,\alpha P)$ as a "bite" of a "nibbling" process (see for example [2, Section 4.7]), that should be similar to $\mathbb{R}(S,\alpha N)$.

Lemma 2.10. *let $\mathcal{P}$ be a property of unordered partial systems that is monotone increasing in the sense that $S \in \mathcal{P}$ and $S' \supseteq S$ implies $S' \in \mathcal{P}$. For fixed $h \geq q-1$ and $a \in (0,1)$ there is $b(a,h) > 0$ such that the following holds. Fix $\alpha \in (0,1)$ and $S \in \mathcal{O}_m^{n^{-a},h}$ for some $m \leq N - \alpha N$ and let $\mathcal{Q} = \left\{S \in \mathcal{O}_{\alpha N+m}^{n^{-b},h} : S_m \in \mathcal{O}_m^{n^{-a},h}\right\} \supseteq \mathcal{O}_{\alpha N+m}^{n^{-a},h}$. Let $\boldsymbol{S} \in \mathbb{R}(S,\alpha N)$ and $\boldsymbol{S}^* \in \mathbb{G}^*(S,\alpha P)$. Then*

$$\Pr(S \cup \boldsymbol{S} \notin \mathcal{P} \,|\, S \cup \boldsymbol{S} \in \mathcal{Q}) = O(1)\Pr(S \cup \boldsymbol{S}^* \notin \mathcal{P}).$$

Proof. Let $\boldsymbol{S}^* \in \mathbb{G}^*(S,\alpha P)$ be obtained from $\boldsymbol{G} \in \mathbb{G}(S,\alpha P)$. Note that conditioning on the number of edges in $\boldsymbol{G}$, its edges comprise a uniformly random subset of its size, of the set of all possible edges. With probability $\Omega(1)$ the number of edges in $\boldsymbol{G}$ is at most $\binom{n}{q}\alpha P \leq \alpha N$, in which case $\boldsymbol{S}^*$ can be coupled as a subset of $\boldsymbol{S}$. Indeed, a random ordering of $\boldsymbol{G}$ can be viewed as the first few elements of a random ordering of the set of all possible edges, and the clique removal process with this ordering produces a superset of $\boldsymbol{S}^*$.

Now, let b be $a\,b(a,h)$ in the notation of Theorem B.1, so that $\Pr(S \cup \boldsymbol{S} \in \mathcal{Q}) = 1 - o(1)$. We have

$$\begin{aligned}\Pr(S \cup \boldsymbol{S} \notin \mathcal{P} \,|\, S \cup \boldsymbol{S} \in \mathcal{Q}) &\leq \Pr(S \cup \boldsymbol{S}^* \notin \mathcal{P} \,|\, S \cup \boldsymbol{S} \in \mathcal{Q} \text{ and } e(\boldsymbol{G}) \leq \alpha N) \\ &\leq \Pr(S \cup \boldsymbol{S}^* \notin \mathcal{P}) / \Pr(S \cup \boldsymbol{S} \in \mathcal{Q} \text{ and } e(\boldsymbol{G}) \leq \alpha N) \\ &= \Pr(S \cup \boldsymbol{S}^* \notin \mathcal{P}) / \Omega(1). \qquad \square\end{aligned}$$

In this subsection we also state and prove a bounded-differences inequality with Bernstein-type tails which can be used to analyse $\mathbb{G}^*(S,\alpha P)$. Standard bounded-difference inequalities such as the Azuma-Hoeffding inequality do not provide strong enough tail bounds to apply Theorem 2.4.

Theorem 2.11. *Let $\boldsymbol{\omega} = (\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_n)$ be a sequence of independent, identically distributed random variables with $\Pr(\boldsymbol{\omega}_i = 1) = p$ and $\Pr(\boldsymbol{\omega}_i = 0) = 1-p$. Let $f(\boldsymbol{\omega})$ satisfy the Lipschitz condition $|f(\boldsymbol{\omega}) - f(\boldsymbol{\omega}')| \leq K$ for all pairs $\boldsymbol{\omega}, \boldsymbol{\omega}' \in \{0,1\}^n$ differing in exactly one coordinate. Then*

$$\Pr(|f(\boldsymbol{\omega}) - \mathbb{E}f(\boldsymbol{\omega})| > t) \leq \exp\left(-\frac{t^2}{4K^2np + 2Kt}\right).$$

Proof. We use Freedman's inequality (Lemma B.2), with the Doob martingale $\boldsymbol{X}(0), \dots, \boldsymbol{X}(n)$ defined by $\boldsymbol{X}(i) = \mathbb{E}[f(\boldsymbol{\omega}) \,|\, \boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_i]$. Note that $\boldsymbol{X}(0) = \mathbb{E}f(\boldsymbol{\omega})$ and $\boldsymbol{X}(n) = f(\boldsymbol{\omega})$. It suffices to show that $V(n) = \sum_{i=0}^{n-1} \mathbb{E}\Big[(\Delta \boldsymbol{X}(i))^2 \,\Big|\, \boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_i\Big] \le 2K^2 np$ with probability 1.

Condition on $\boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_i$ (thereby conditioning on $\boldsymbol{X}(i)$). Let X^0 and X^1 be the values of $\boldsymbol{X}(i+1)$ in the cases $\boldsymbol{\omega}_{i+1} = 0$ and $\boldsymbol{\omega}_{i+1} = 1$, respectively. We have

$$\begin{aligned} \boldsymbol{X}(i) &= pX^1 + (1-p)X^0, \\ \big|\boldsymbol{X}(i) - X^0\big| &= p\big|X^1 - X^0\big| \le Kp. \end{aligned}$$

So,

$$\begin{aligned} \mathbb{E}\Big[(\Delta \boldsymbol{X}(i))^2 \,\Big|\, \boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_i\Big] &= p\big(\boldsymbol{X}(i) - X^1\big)^2 + (1-p)\big(\boldsymbol{X}(i) - X^0\big)^2 \\ &\le K^2 p + (1-p)K^2 p^2 \\ &\le 2K^2 p. \end{aligned}$$

The desired bound on $V(n)$ follows. □

3 Perfect matchings via absorbers

In this section we introduce our absorbers, discuss how they can be used to find perfect matchings, and prove that absorbers can be found in the partial Steiner triple system produced by the triangle removal process. This will culminate in the proof of Theorem 1.1.

Definition 3.1. An *absorber* for an ordered triple (x, y, z) is a set of edges of the form

$$\{\{x, x_1, x_2\}, \{y, y_1, y_2\}, \{z, z_1, z_2\}, \{w_x, w_y, w_z\}\} \cup \{\{x_1, y_2, w_z\}, \{y_1, z_2, w_x\}, \{z_1, x_2, w_y\}\}.$$

We call x, y, z the *rooted* vertices and we call the other 9 vertices the *external* vertices. Note that an absorber already has a perfect matching on its 12 vertices (the first 4 edges; we call this the *covering* matching), but it also has a perfect matching on its 9 vertices not including x, y, z (the last 3 edges; we call this the *non-covering* matching). See Figure 1.

Now we explain how absorbers are put together. The relative positions of the absorbers will be determined by a "template" structure, as follows.

Lemma 3.2. *For any sufficiently large n, there exists a 3-uniform hypergraph T with $10n$ vertices, at most $160n$ hyperedges and a "flexible set" Z of $2n$ vertices, such that if we remove any n vertices from Z, the resulting hypergraph has a perfect matching. We call this H a* resilient template.

To prove Lemma 3.2 we use the following lemma of Montgomery [20, Lemma 2.8]

Lemma 3.3. *For any sufficiently large n, there exists a bipartite graph H on vertex classes X and $Y \sqcup Z$ with $|X| = 3n$, $|Y| = |Z| = 2n$, and maximum degree 40, such that if we remove any n vertices from Z, the resulting bipartite graph has a perfect matching.*

Proof of Lemma 3.2. Consider the bipartite graph H from Lemma 3.3 on the vertex set $X \sqcup Y \sqcup Z$. Add a set W of $|X|$ new vertices and put a perfect matching between W and X, to obtain a $10n$-vertex tripartite graph H'. Now, define our resilient template with flexible set Z by putting a hyperedge with the vertices of each of the (at most $160n$) paths of length 2 in H'. □

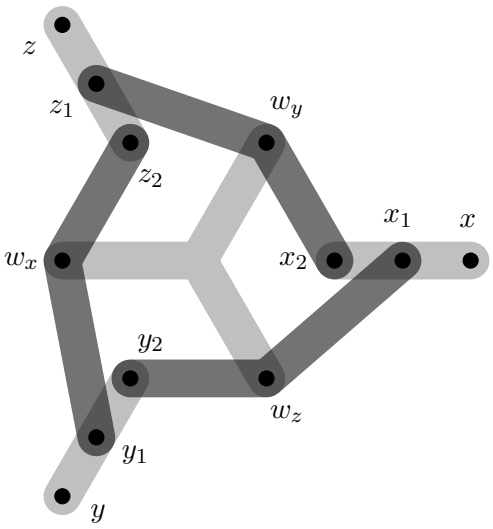

Figure 1. An illustration of an absorber for (x, y, z). The light edges are the covering matching and the dark edges are the non-covering matching.

Now we can describe our absorbing structure in its entirety.

Definition 3.4. An *absorbing structure* is a 3-uniform hypergraph of the following form. Consider a resilient template and put externally vertex-disjoint absorbers on each edge of the template, introducing 9 new vertices for each. (Note that the template just describes the relative positions of the absorbers, its edges are not actually in the absorbing structure).

Note that an absorbing structure with a flexible set of size n has $10n + 9 \times 160n = O(n)$ vertices and at most $7 \times 160n = O(n)$ edges. Now, we show how absorbers can be used to find perfect matchings.

Lemma 3.5. *Consider a 3-uniform hypergraph S (with vertex set V, $|V| = n$) satisfying the following properties for some $\delta = \delta(n) = o(1)$ and fixed $\beta > 0$.*

1. *There is an absorbing structure H in S with at most δn vertices and a flexible set Z of size $\delta^2 n$.*
2. *For at most δn of the vertices $v \in V$, we have $|\{\{x, y\} \subseteq Z : \{v, x, y\} \in S\}| < 2\delta^3 n$. That is to say, few vertices have low degree into the flexible set Z, in S.*
3. *Every vertex subset $W \subseteq V$ with $|W| \geq \delta^3 n$ induces at least $(1 - \beta)|W|^3/(6n)$ hyperedges*

Then S has

$$\left(\frac{n}{2e^2}(1 - \beta - O(\delta \log n))\right)^{n/3}$$

perfect matchings.

Proof. We can assume $\delta < 1/4$, so that we can greedily choose a δn-vertex matching in $V \backslash V'$ containing each of the vertices violating the second condition. There are $n' = n - v(H) - \delta n \geq (1 - 2\delta)n$ vertices in $V \backslash V(H)$ remaining unmatched.

Next, repeatedly choose an edge in the remaining unmatched vertices in $V \backslash V(H)$ until there are only $\delta^3 n$ such vertices remaining unmatched. (This means we take a matching with $m =$

$\big(n' - \delta^2 n\big)/3 = n/3 - O(\delta n)$ edges). The number of ways to make this ordered sequence of choices is at least

$$\prod_{i=1}^{m}(1-\beta)\frac{(n'-3i)^3}{6n} = \left(\frac{1-\beta}{6n}\right)^m \exp\left(\sum_{i=1}^{m}\log\big(n'-3i\big)^3\right) \tag{3}$$
$$= \left(\frac{(1-\beta)(n')^3}{6n}\right)^m \exp\left(n'\sum_{i=1}^{m}\frac{3}{n'}\log\left(1-3\frac{i}{n'}\right)\right).$$

Since $\log(1-3t)$ is a decreasing function of t, we have

$$\sum_{i=1}^{m}\frac{1}{n'}\log\left(1-3\frac{i}{n'}\right) \le \int_0^{m/n'} 3\log(1-3t)\,\mathrm{d}\,t \le \sum_{i=1}^{n'/3}\frac{1}{n'}\log\left(1-3\frac{i-1}{n'}\right).$$

Noting that

$$\sum_{i=1}^{m}\left(\log\left(1-3\frac{i-1}{n'}\right)-\log\left(1-3\frac{i}{n'}\right)\right) = \sum_{i=1}^{m}\left(\log\left(1+\frac{3}{n'-3i}\right)\right)$$
$$\le 3\sum_{i=1}^{m}\frac{1}{n'-3i} = O(\log n)$$

and that $\int \log s\,\mathrm{d}\,s = s(\log s - 1)$, we have

$$\sum_{i=1}^{n'/3}\frac{3}{n'}\log\left(1-3\frac{i}{n'}\right) = \int_0^{1/3-O(\delta)} 3\log(1-3t)\,\mathrm{d}\,t + O\left(\frac{\log n}{n}\right)$$
$$= \int_{O(\delta)}^{1}\log s\,\mathrm{d}\,s + O\left(\frac{\log n}{n}\right)$$
$$= -1 - O(\delta\log(n)).$$

So, the expression in (3) is equal to

$$\left((1-\beta)\frac{n^2}{6}\right)^{n'/3} e^{-n} n^{O(\delta n)} = \left(\frac{n^2}{6e^3}\right)^{n/3} n^{O(\delta n)}.$$

Dividing by $m! = (n/(3e))^{n/3} n^{O(\delta n)}$ (because we were counting ordered matchings), there are at least $\big(n/\big(2e^2\big)\big)^{n/3} n^{O(\delta n)}$ ways to choose a matching in $V\backslash V(H)$ covering all but $\delta^3 n$ vertices, none of which violate the second condition. We can greedily extend such a matching to cover half of Z using the second and third conditions. By the defining property of a resilient template, there is a perfect matching in the subgraph of the template induced by the vertices uncovered so far. For each hyperedge of this perfect matching, use the covering matching of the corresponding absorber; for each other hyperedge in the template use the non-covering matching. □

3.1 Absorbing properties in the triangle removal process

In this section we prove that the properties in Lemma 3.5 (for say $\delta = 1/\log^2 n$ and arbitrarily small β) hold in the triangle removal process, conditioned on a certain quasirandomess event, with probability $1 - \exp\Big(-\tilde{\Omega}\big(n^2\big)\Big)$. (A tilde above asymptotic nation indicates that polylogarithmic

factors are being ignored). This will allow us to use Theorem 2.4 to deduce that the same is true for a uniformly random Steiner triple system, proving Theorem 1.1. Recall that Steiner triple systems have $q = 3$ and $r = 2$, so in the notation of Section 2 (and Section 2.3) we have $Q = 3$, $N = \binom{n}{2}/3 = (1 + o(1))n^2/6$ and $P = 1/n$.

Fix a large constant $h \in \mathbb{N}$ (we will see later exactly how large it should be), and consider arbitrarily small $\alpha > 0$. We will treat α as a constant, but asymptotics in this section will be mostly uniform over α.

3.1.1 Absorbers

First we find absorbers. Note that K_n is $\big(n^{1/2}, h\big)$-quasirandom, and let $a = a\,b(1/2, h)$ in the notation of Theorem B.1, so that $\mathcal{O}_{\alpha N}^{n^{-a},h}$ is nonempty. Let $b = b(a)$ in the notation of Lemma 2.10. As in Lemma 2.10, let $\mathcal{Q} = \Big\{S \in \mathcal{O}_{2\alpha N}^{n^{-b},h} : S_{\alpha N} \in \mathcal{O}_{\alpha N}^{n^{-a},h}\Big\} \supseteq \mathcal{O}_{2\alpha N}^{n^{-a},h}$. Let $\boldsymbol{S} \in \mathbb{R}(n, 2\alpha n)$, conditioned on the event $\boldsymbol{S} \in \mathcal{Q}$, and condition on any $\boldsymbol{S}_{\alpha N} = S \in \mathcal{O}_{\alpha N}^{n^{-a},h}$. We will use Lemma 2.10 to analyse $\boldsymbol{S} \backslash \boldsymbol{S}_{\alpha N} \in \mathbb{R}(S, \alpha N)$ via $\mathbb{G}^*(S, \alpha/N)$. So, let $\boldsymbol{S}^* \in \mathbb{G}^*(S, \alpha/N)$ be obtained from $\boldsymbol{G} \in \mathbb{G}(S, \alpha/N)$.

By quasirandomness, every vertex has degree $(1 \pm n^{-a})(1 - \alpha)n$ $(1 \pm n^{-a})\alpha n/2 = \Omega(\alpha n)$ in S. Consider vertices x, y, z. Say an *absorber-extension* is a collection of 4 hyperedges which forms an absorber on x, y, z when combined with three edges of S incident to x, y, z. Let $\boldsymbol{Y}$ be the maximum size of an edge-disjoint collection of absorber-extensions in $\boldsymbol{S}^*$. That is, it is the maximal size of a collection of hyperedge-disjoint isolated absorber-extensions in $\boldsymbol{G}$, where we say a subgraph of $\boldsymbol{G}$ is *isolated* if none of its hyperedges intersect another hyperedge of $\boldsymbol{G}$ in more than one vertex. Now, adding a hyperedge to a hypergraph can invalidate at most three absorber-extensions in a maximal collection, and removing a hyperedge can remove at most one absorber-extension in a maximal collection. So, changing the presence of one edge of $\boldsymbol{G}$ can change $\boldsymbol{Y}$ by at most 3. By Theorem 2.11,

$$\Pr(\boldsymbol{Y} \le \mathbb{E}\boldsymbol{Y} - t) \le \exp\left(-\Omega\left(\frac{t^2}{3^2\binom{n}{3}\alpha/n + 3t}\right)\right) = \exp\left(-\Omega\left(\frac{t^2}{\alpha n^2 + t}\right)\right). \tag{4}$$

Let $\boldsymbol{X}$ be the number of isolated absorber-extensions in $\boldsymbol{G}$ and let $\boldsymbol{Z}$ be the number of pairs of hyperedge-intersecting absorber-extensions in $\boldsymbol{G}$. We can obtain a collection of disjoint isolated absorber-extensions by considering the collection of all isolated absorber-extensions and deleting one from each intersecting pair, so $\boldsymbol{Y} \ge \boldsymbol{X} - \boldsymbol{Z}$ and $\mathbb{E}\boldsymbol{Y} \ge \mathbb{E}\boldsymbol{X} - \mathbb{E}\boldsymbol{Z}$.

If h is large enough and α is small enough, there are $\Theta\big(\alpha^3 n^6\big)$ possible absorber-extensions not conflicting with S. Indeed, to choose such a candidate absorber-extension, first choose three hyperedges e_x, e_y, e_z incident to x, y, z (there are $(\alpha n)^3$ ways to do this). Then, consider the graph H obtained by removing the three rooted hyperedges of an absorber and replacing the remaining hyperedges with triangles. The number of ways to complete our candidate absorber-extension is precisely the number of copies of H rooted on the vertices of e_x, e_y, e_z in the obvious way, in $G(S)$. By Proposition 2.8 this number is $(1 - \alpha)^{O(1)} n^3 = \Theta\big(n^3\big)$. The probability each possible absorber-extension occurs and is isolated in $\boldsymbol{G}$ is $\Theta\Big((\alpha/n)^4(1 - \alpha/n)^{O(n)}\Big) = \Theta\big(\alpha^4 n^{-4}\big)$. So, $\mathbb{E}\boldsymbol{X} = \Theta\big(\alpha^7 n^2\big)$.

Now, there are several possibilities for a hyperedge-intersecting pair of distinct absorber-extensions.

- they could intersect in the unrooted hyperedge of the covering matching of the absorber. There are $\Theta\Big((\alpha n)^6 n^3\Big)$ possibilities for such a pair of absorber-extensions, and each occurs with probability $\Theta\Big((\alpha/n)^7\Big)$, so the expected number of such is $\Theta\big(\alpha^{13} n^2\big)$.

- They could intersect in the unrooted hyperedge of the covering matching as above, and also in one hyperedge of the non-covering matching". There are $\Theta\Big((\alpha n)^4 n^3\Big)$ possible such pairs, and each occurs with probability $\Theta\Big((\alpha/n)^6\Big)$, so the expected number of such is $\Theta\big(\alpha^{10} n\big)$.
- They could intersect in one edge of the non-covering matching. There are $\Theta\Big((\alpha n)^4 n^5\Big)$ possible such pairs, and each occurs with probability $\Theta\Big((\alpha/n)^7\Big)$, so the expected number of such is $\Theta\big(\alpha^{11} n^2\big)$.
- They could intersect in two edges of the non-covering matching. There are $\Theta\Big((\alpha n)^3 n^4\Big)$ possible such pairs, and each occurs with probability $\Theta\Big((\alpha/n)^6\Big)$, so the expected number of such is $\Theta\big(\alpha^9 n\big)$.

In summary (for small α), we have $\mathbb{E}\boldsymbol{Z} = O\big(\alpha^{11} n^2\big)$, so $\mathbb{E}\boldsymbol{Y} = \Theta\big(\alpha^7 n^2\big)$. Considering α as fixed, (4) gives $\boldsymbol{Y} = \Omega\big(n^2\big)$ with probability $1 - \exp\big(-\Omega\big(n^2\big)\big)$.

Note that if there are $\Omega\big(n^2\big)$ edge-disjoint absorber-extensions then there must in fact be $\Omega(n)$ vertex-disjoint absorber-extensions. We can find these greedily; the degree in $\boldsymbol{S}$ of each vertex is $O(n)$, so deleting $O(1)$ vertices can delete at most $O(n)$ edges. By the union bound and Lemma 2.10, it follows that with probability $1-\exp\big(-\Omega\big(n^2\big)\big)$, every triple of vertices has $\Omega(n)$ externally vertex-disjoint absorbers in $\boldsymbol{S}$.

If $\boldsymbol{S}$ has this property, then we can greedily build our absorbing structure, as follows. Recalling that $\delta = o(1)$, choose a resilient template H with $O\big(\delta^2 n\big)$ hyperedges, on $O\big(\delta^2 n\big)$ vertices of $\boldsymbol{S}$, such that the flexible set is $Z = \big[\delta^2 n\big]$ for some fixed $c > 0$. Noting that an absorber has $O(1)$ non-rooted vertices, we can greedily find an absorber in $\boldsymbol{S}$ on each edge of H, in an externally vertex-disjoint fashion. The entire absorbing structure then has $O\big(\delta^2 n\big) \le \delta n$ vertices. This proves that the first property of Lemma 3.5 holds with probability $1 - \exp\big(-\Omega\big(n^2\big)\big)$ in $\boldsymbol{S}$ (conditioned on $\mathcal{Q}$), and by Theorem 2.4 it therefore holds a.a.s. in a uniformly random Steiner triple system. (Actually, what we have proved is slightly stronger, because we can specify the location of the flexible set of the absorbing structure).

3.1.2 High degree into the flexible set

Recall that $\delta = 1/\log^2 n = \tilde{\Theta}(1)$. With $\boldsymbol{S}^* \in \mathbb{G}^*(n, \alpha/N)$ obtained from $\boldsymbol{G} \in \mathbb{G}(n, \alpha/N)$, fix a set W with δn vertices and let $\boldsymbol{Y}$ be the number of edges of $\boldsymbol{S}^*$ with one vertex in W and two vertices in the set $Z = \big[\delta^2 n\big]$. That is, $\boldsymbol{Y}$ the number of such edges in $\boldsymbol{G}$ that are "isolated". There are $\Theta\Big((\delta n)^2 n\Big) = \Theta\Big(n^3 \log^{-O(1)} n\Big)$ possible edges and each is present and isolated with probability $(\alpha/n)(1-\alpha/n)^{O(n)} = \Theta\big(n^{-1}\big)$, so $\mathbb{E}\boldsymbol{Y} = \tilde{\Theta}\big(n^2\big)$ and just as in the preceding argument, Theorem 2.11 shows that for some δ, with probability $1-\exp\Big(-\tilde{\Omega}\big(n^2\big)\Big)$ we have $\boldsymbol{Y} \ge 2\delta^3 n^2/4$. This is not possible unless there is a vertex v in W with degree $2\delta^3 n$ in Z, in $\boldsymbol{S}^*$. Since there are fewer than $2^n = e^{o\left(n^2\right)}$ choices for W, the union bound and Lemma 2.10 prove that with probability $1-\exp\Big(-\tilde{\Omega}\big(n^2\big)\Big)$ every set W has such a vertex (and therefore the second condition of Lemma 3.5 holds), in an instance of the triangle removal process $\mathbb{R}(n, \alpha N)$. Applying Theorem 2.4 with $S = \varnothing$ proves that the same holds a.a.s. in a uniformly random Steiner triple system.

3.1.3 Density in subsets

Fix a set $W \subseteq V$ with $|W| \geq \delta^3 n$. With $\boldsymbol{S}^*$ and $\boldsymbol{G}$ as in the previous subsection, redefine $\boldsymbol{Y}$ to be the number of edges of $\boldsymbol{S}^*$ included in W. There are $(1+o(1))|W|^3/6$ possible edges, and each is present and isolated in $\boldsymbol{G}$ with probability $(\alpha/n)(1-\alpha/n)^{O(n)} = (\alpha/n)(1-O(\alpha))$, so with the same reasoning as in the previous two sections, Theorem 2.11 shows that with probability $1-\exp\bigl(-\tilde{\Omega}\bigl(n^2\bigr)\bigr)$ we have $\boldsymbol{Y} \geq \alpha(1-O(\alpha))|W|^3/(6n)$. The union bound and Lemma 2.10 proves that this holds for any W with probability $1-\exp\bigl(-\tilde{\Omega}\bigl(n^2\bigr)\bigr)$, in an instance of the triangle removal process $\mathbb{R}(n,\alpha N)$. Therefore by Theorem 2.4 it holds a.a.s. in $\boldsymbol{S}_{\alpha N}$ for a uniformly random $\boldsymbol{S} \in \mathcal{S}$. By symmetry, this property also holds a.a.s. in $\boldsymbol{S}_{k\alpha N} \backslash \boldsymbol{S}_{(k-1)\alpha N}$ for any $k \leq 1/\alpha$ (we impose that $1/\alpha$ is an integer). So the total number of edges in $\boldsymbol{S}$ induced by any W is $(1-O(\alpha))|W|^3/(6n)$. For β a large multiple of α, the third condition in Lemma 3.5 is then satisfied.

4 An upper bound for the number of perfect matchings

In this section we prove Theorem 1.2, with the entropy method. See [16] for a brief introduction to the notion of entropy as necessary for proving upper bounds of this type.

Proof. Let $\mathcal{M}$ be the set of perfect matchings in S. Consider a uniformly random $\boldsymbol{M} \in \mathcal{M}$ and let $H(\boldsymbol{M}) = \log|\mathcal{M}|$ be the entropy of $\boldsymbol{M}$. Let $\boldsymbol{M}_v$ be the hyperedge of $\boldsymbol{M}$ containing the vertex v, so that the sequence $(\boldsymbol{M}_v)_{v\in[n]}$ determines $\boldsymbol{M}$. For any ordering on the vertices of S,

$$H(M) = \sum_{v\in V} H\bigl(\boldsymbol{M}_v \,\big|\, \boldsymbol{M}_{v'} : v' < v\bigr).$$

Now, a sequence $\lambda \in [0,1]^n$ with all λ_v distinct induces an ordering on $[n]$, with $v' < v$ when $\lambda_{v'} > \lambda_v$. Let $\boldsymbol{R}_v(\lambda)$ be 1 plus the number of 3-edges $e \neq \boldsymbol{M}_v$ such that $\lambda_{v'} < \lambda_v$ for all $v' \in \bigcup_{x\in X} \boldsymbol{M}_x$. (So, $\boldsymbol{R}_v(\lambda) = 1$ if $\lambda_{v'} > \lambda_v$ for some $v' \in \boldsymbol{M}_v \backslash \{v\}$). This is an upper bound on the number of possible values for $\boldsymbol{M}_v$ given the information $(\boldsymbol{M}_{v'} : \lambda_{v'} > \lambda_v)$. Let

$$R_v^{M,\lambda} = \mathbb{E}\bigl[\boldsymbol{R}_v(\boldsymbol{\lambda}) \,\big|\, \boldsymbol{M} = M,\, \boldsymbol{\lambda}_v = \lambda,\, \lambda_{v'} < \lambda_v \text{ for all } v' \in \boldsymbol{M}_v \backslash \{v\}\bigr].$$

(Note that $\boldsymbol{\lambda}_v = \lambda$ occurs with probability zero, so formally we should condition on $\boldsymbol{\lambda}_e = \lambda \pm \mathrm{d}\,\lambda$ and take limits in what follows, but there are no continuity issues so we will ignore this detail). By linearity of expectation, we have

$$R_v^{M,\lambda} = ((n-1)/2 - 1)\lambda^6,$$

because there are $(n-1)/2$ edges in S containing v, and for each such edge e other than $\boldsymbol{M}_v$ the two non-v vertices must be matched to different 3-edges in $\boldsymbol{M}$ (if they were matched to the same edge that edge would intersect e in two vertices, violating the defining property of a Steiner triple system). By Jensen's inequality,

$$\mathbb{E}\bigl[\log \boldsymbol{R}_v(\boldsymbol{\lambda}) \,\big|\, \boldsymbol{M} = M,\, \boldsymbol{\lambda}_v = \lambda,\, \lambda_{v'} < \lambda_v \text{ for all } v' \in \boldsymbol{M}_v \backslash \{v\}\bigr] \leq \log R_v^{M,\lambda},$$

and

$$\Pr\bigl(\lambda_{v'} < \lambda_v \text{ for all } v' \in \boldsymbol{M}_v \backslash \{v\}\bigr) = \lambda^2$$

so

$$\mathbb{E}[\log \boldsymbol{R}_v(\boldsymbol{\lambda}) \,|\, \boldsymbol{M} = M,\, \boldsymbol{\lambda}_v = \lambda] \leq \lambda^2 \log R_v^{M,\lambda}.$$

For any $M \in \mathcal{M}$ we then have

$$\begin{aligned}\mathbb{E}[\log \boldsymbol{R}_v(\boldsymbol{\lambda}) \,|\, \boldsymbol{M} = M] &\le \mathbb{E}\Big[\boldsymbol{\lambda}_v^2 \log R_v^{M,\boldsymbol{\lambda}_i}\Big] \\ &= \int_0^1 \lambda^2 \log\big(((n-1)/2-1)\lambda^6\big)\,\mathrm{d}\,\lambda \\ &= \frac{1}{3}(\log((1+o(1))n/2) - 2),\end{aligned}$$

using the fact that $\int_0^1 \lambda^{A-1}\log\big(C\lambda^B\big)\,\mathrm{d}\,t = A^{-1}\log C - A^{-2}B$ for any $A, B, C > 0$. We conclude that

$$\begin{aligned}\log|\mathcal{M}| &\le H(\boldsymbol{M}) \\ &\le \sum_{i\in[n]} \mathbb{E}[\log \boldsymbol{R}_i(\boldsymbol{\lambda})] \\ &\le \frac{n}{3}(\log((1+o(1))n/2) - 2)\end{aligned}$$

which is equivalent to the theorem statement. □

5 Latin squares

In this section we briefly sketch how one should adapt the methods in this paper to prove "Theorem" 1.3.

The main difference from the case of Steiner triple systems is that instead of considering triangles in graphs (subgraphs of K_n), we consider triangles in subgraphs of the complete tripartite graph $K_{n,n,n}$. In this section we use the notation $V = V_1 \sqcup V_2 \sqcup V_3$ for the vertex partition of $K_{n,n,n}$. We say a subgraph $G \subseteq K_{n,n,n}$ with m edges between each pair of parts is (ε, h)-quasirandom if for each $i \in \{1,2,3\}$, every set $A \subseteq V\backslash V_i$ with $|A| \le h$ has $(1 \pm \varepsilon)\big(m/n^2\big)^{|A|}n$ common neighbours in V_i. We believe that the following result should follow from a slight adaptation of the proof in [13].

Conjecture 1. *There are $h \in \mathbb{N}$, $\varepsilon_0, a \in (0,1)$ and $n_0, \ell \in \mathbb{N}$ such that if $n \ge n_0$, $m/n^2 \ge n^{-a}$ and $\varepsilon \le \varepsilon_0\big(m/n^2\big)^\ell$, and $G \subseteq K_{n,n,n}$ is (ε, h)-quasirandom with m edges between each pair of parts, then the edges of G can be decomposed into triangles.*

Redefining $\mathcal{O}$, $\mathcal{O}_m$ and $\mathcal{O}_m^{\varepsilon,h}$ in the obvious way for ordered (partial) Latin squares and redefining $\mathbb{R}(n,m)$ to be the distribution on $\mathcal{O}_m$ obtained with m steps of the triangle removal process starting from $K_{n,n,n}$, we can then prove a counterpart to Theorem 2.4 for random Latin squares, with virtually the same proof.

"Theorem" 5.1. *There is $h_0 \in \mathbb{N}$ such that for fixed $h \ge h_0$ and sufficiently small a there is $b(a) > 0$ such that the following holds. Fix $\alpha \in (0,1)$, let $\mathcal{P} \subseteq \mathcal{O}_{\alpha N}$ be a property of ordered partial Latin squares, let $\mathcal{Q} \supseteq \mathcal{O}_{\alpha N}^{n^{-a},h}$, let $\boldsymbol{S} \in \mathcal{O}$ be uniformly random and let $\boldsymbol{S}' \in \mathbb{R}(n, \alpha N)$. If*

$$\Pr\big(\boldsymbol{S}' \notin \mathcal{P} \,\big|\, \boldsymbol{S}' \in \mathcal{Q}\big) < \exp\Big(-n^{2-b}\Big)$$

then

$$\Pr(\boldsymbol{S}_{\alpha N} \notin \mathcal{P}) = \exp\big(-\Omega\big(n^{1-2a}\big)\big).$$

Now we outline what should be adapted from the arguments in Section 3 for the Latin squares case. The definition of an absorber can remain the same, noting that the hypergraph in Definition 3.1 is tripartite (if x, y, z are in different parts, then an absorber on (x, y, z) can be chosen with each 3-edge having a vertex in each part). The definition of a resilient template should be adapted slightly: a resilient template is a tripartite hypergraph H with a flexible set Z satisfying $|Z \cap V_1| = |Z \cap V_2| = |Z \cap V_3|$, such that if half the vertices are removed from each $Z \cap V_i$ the remaining hypergraph has a perfect matching. To prove a counterpart of Lemma 3.2 we can just use three copies (one for each V_i) of the construction in the proof of Lemma 3.2. The counterpart of Lemma 3.5 is as follows (with virtually the same proof).

Lemma 5.2. *Consider a 3-uniform tripartite hypergraph $L \subseteq K^{(3)}_{n,n,n}$ satisfying the following properties for some $\delta = \delta(n) = o(1)$ and fixed $\beta > 0$.*

1. *There is an absorbing structure H in L with at most δn vertices and a flexible set Z intersecting each V_i in $\delta^2 n$ vertices.*
2. *For at most δn of the vertices $v \in V_1$, we have $|\{(x, y) \in V_2 \times V_3 : (v, x, y) \in L\}| < 3\delta^3 n$, and similarly most $v \in V_2$ and $v \in V_3$ have high degree into $V_1 \times V_3$ and $V_1 \times V_2$ respectively.*
3. *For any choice of $W_i \subseteq V_i$ such that each $|W_i| \geq \delta^3 n$, there are at least $(1-\beta)|W_1||W_2||W_3|/n$ hyperedges in $W_1 \times W_2 \times W_3$.*

Then L has

$$\left(\frac{n}{e^2}(1 - \beta - O(\delta \log n))\right)^n$$

transversals.

One can then use Lemma 5.2 and "Theorem" 5.1 to prove "Theorem" 1.3 in basically the same way as the proof of Theorem 1.1 in Section 3.1.

6 Concluding remarks

In this paper we introduced a new method for analysing random Steiner systems, and we used it to prove that almost all Steiner triple systems have many perfect matchings. There are many interesting open questions about random Steiner (triple) systems and perfect matchings that remain.

- We believe the most interesting problem that seems approachable by our methods is to prove that almost all Steiner triple systems (and Latin squares) can be decomposed, or at least approximately decomposed, into disjoint perfect matchings (transversals). The proof of Theorem 1.1 can be easily modified to prove that almost all Steiner triple systems have $\Omega(n)$ disjoint perfect matchings, but to find $(1 - o(1))n/2$ disjoint perfect matchings would require a new idea. For Latin squares, the property of being decomposable into transversals is equivalent to the important property of having an *orthogonal mate*, which has a long history dating back to Euler. More details can be found in [27].
- There is the obvious question of proving that almost all (q, r)-Steiner systems have perfect matchings, for all choices of (q, r). We believe this should be possible using the methods in this paper, and the main difficulty probably lies in coming up with a suitable absorber. If $r > 2$ then there is the more general (and much more difficult) question of whether (q, r)-Steiner systems typically contain (q, r')-Steiner systems, for all $r' < r$. It might even be true that almost all (q, r)-Steiner systems can be "completely decomposed" in the sense that

they can be partitioned into disjoint $(q, r-1)$-Steiner systems, which can be partitioned into $(q, r-2)$-Steiner systems, and so on. A similar phenomenon occurs for regular subgraphs of random regular graphs (see [9, Section 9.5]).

- Another interesting question about random Steiner triple systems is whether they contain Steiner triple subsystems on fewer vertices. McKay and Wanless [19] proved that almost all Latin squares have many small Latin subsquares, but it was conjectured by Quackenbush [21] that most Steiner triple systems do not have proper subsystems. It seems unlikely that the methods in this paper will be able to prove or disprove this conjecture without substantial new ideas; actually by consideration of the random 3-graph $\mathbb{G}(n, 1/n)$ we suspect the expected number of 7-vertex Steiner triple systems (Fano planes) in a random Steiner triple system is $\Theta(1)$, and that the distribution of this number is asymptotically Poisson.

- We could ask more generally about containment and enumeration of subgraphs. Is it true that every fixed hypergraph H whose every subgraph has more vertices than edges, appears a.a.s. in a random Steiner triple system? Can we show that moreover the number of copies of H is concentrated? The methods in this paper can probably be used to prove a lower bound for the number of copies of H when every subgraph of H has at least 2 more vertices than edges, but due to the "infamous upper tail" issue (see [10]), an upper bound for the number of copies of H is likely to be more difficult.

- One of the most fundamental properties of random graphs and hypergraphs is that they have low *discrepancy*, meaning that every sufficiently large subset of vertices has about the expected number of edges. In Section 3.1.3 we effectively proved a very weak discrepancy bound, but it is not clear how to use our methods to reach anywhere near optimal discrepancy. See [17] for some conjectures about discrepancy of Latin squares and their high-dimensional relatives.

References

[1] N. Alon, J.-H. Kim, and J. Spencer, *Nearly perfect matchings in regular simple hypergraphs*, Israel Journal of Mathematics **100** (1997), no. 1, 171–187.

[2] N. Alon and J. H. Spencer, *The probabilistic method*, John Wiley & Sons, 2004.

[3] L. Babai, *Almost all Steiner triple systems are asymmetric*, Annals of Discrete Mathematics **7** (1980), 37–39.

[4] D. Bryant and D. Horsley, *Steiner triple systems without parallel classes*, SIAM Journal on Discrete Mathematics **29** (2015), no. 1, 693–696.

[5] P. J. Cameron, *A Markov chain for Steiner triple systems*, (2002), `http://www.maths.qmul.ac.uk/~pjc/csgnotes/random.pdf`.

[6] P. Erdős and A. Rényi, *On random graphs I*, Publ. Math. Debrecen **6** (1959), 290–297.

[7] D. A. Freedman, *On tail probabilities for martingales*, the Annals of Probability (1975), 100–118.

[8] R. Glebov and Z. Luria, *On the maximum number of Latin transversals*, Journal of Combinatorial Theory, Series A **141** (2016), 136–146.

[9] S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, Cambridge University Press, 2000.

[10] S. Janson and A. Rucinski, *The infamous upper tail*, Random Structures & Algorithms **20** (2002), no. 3, 317–342.

[11] A. Johansson, J. Kahn, and V. Vu, *Factors in random graphs*, Random Structures & Algorithms **33** (2008), no. 1, 1–28.

[12] P. Keevash, *The existence of designs*, arXiv preprint arXiv:1401.3665 (2014).

[13] ———, *Counting designs*, arXiv preprint arXiv:1504.02909 (2015).

[14] M. Krivelevich, *Triangle factors in random graphs*, Combinatorics, Probability and Computing **6** (1997), no. 03, 337–347.

[15] M. Krivelevich, B. Sudakov, V. H. Vu, and N. C. Wormald, *Random regular graphs of high degree*, Random Structures & Algorithms **18** (2001), no. 4, 346–363.

[16] N. Linial and Z. Luria, *An upper bound on the number of Steiner triple systems*, Random Structures & Algorithms **43** (2013), no. 4, 399–406.

[17] ———, *Discrepancy of high-dimensional permutations*, arXiv preprint arXiv:1512.04123 (2015).

[18] A. Lubotzky, Z. Luria, and R. Rosenthal, *Random Steiner systems and bounded degree coboundary expanders of every dimension*, arXiv preprint arXiv:1512.08331 (2015).

[19] B. D. McKay and I. M. Wanless, *Most Latin squares have many subsquares*, Journal of Combinatorial Theory, Series A **86** (1999), no. 2, 323–347.

[20] R. Montgomery, *Embedding bounded degree spanning trees in random graphs*, arXiv preprint arXiv:1405.6559 (2014).

[21] R. W. Quackenbush, *Algebraic speculations about Steiner systems*, Annals of Discrete Mathematics **7** (1980), 25–35.

[22] J. Radhakrishnan, *An entropy proof of Brégman's theorem*, Journal of Combinatorial Theory, Series A **77** (1997), no. 1, 161–164.

[23] D. K. Ray-Chaudhuri and R. M. Wilson, *Solution of Kirkman's schoolgirl problem*, Proc. symp. pure Math, vol. 19, 1971, pp. 187–203.

[24] V. Rödl, A. Ruciński, and E. Szemerédi, *Perfect matchings in large uniform hypergraphs with large minimum collective degree*, Journal of Combinatorial Theory, Series A **116** (2009), no. 3, 613–636.

[25] A. Taranenko, *Multidimensional permanents and an upper bound on the number of transversals in Latin squares*, Journal of Combinatorial Designs **23** (2015), no. 7, 305–320.

[26] J. H. van Lint and R. M. Wilson, *A course in combinatorics*, Cambridge university press, 2001.

[27] I. M. Wanless, *Transversals in Latin squares: a survey*, Surveys in Combinatorics 2011 (R. Chapman, ed.), London Mathematical Society Lecture Note Series, vol. 392, Cambridge University Press, 2011, pp. 403–437.

[28] R. M. Wilson, *Nonisomorphic Steiner triple systems*, Mathematische Zeitschrift **135** (1974), no. 4, 303–313.

A Counting completions of Steiner systems

In this section we prove Lemma 2.5. This is accomplished with minor adaptations of a proof by Keevash [13]. As in Section 2, we consider q and r to be fixed, let $Q = \binom{q}{r}$, let $N = \binom{n}{r}/Q$ and impose that $\binom{q-i}{r-i}$ divides $\binom{n-i}{r-i}$ for every $0 \le i \le r-1$.

For a partial system $S \in \mathcal{S}_{\alpha N}^{n^{-a}}$, let $\mathcal{S}^{\mathrm{ext}}(S)$ be the number of Steiner systems that contain S. We want to determine $\big|\mathcal{O}^{\mathrm{ext}}(S)\big| = (N-\alpha N)!\big|\mathcal{S}^{\mathrm{ext}}(S)\big|$ up to a factor of $e^{n^{2-b}}$ (for some b).

First, we can get an upper bound via the entropy method. This proof is almost exactly the same to the proof of [13, Theorem 6.1], and is very similar to the proof that appears in a paper by Linial and Luria [16] about Steiner triple systems. The reader may refer to that paper for more detailed exposition and a brief introduction to the notion of entropy.

Theorem A.1. *For any $a > 0$, any $\alpha \in [0,1]$, and any $S^* \in \mathcal{S}_{\alpha N}^{n^{-a},q-1}$,*

$$\big|\mathcal{S}^{\mathrm{ext}}(S^*)\big| \le \left(\big(1+O\big(n^{-a}\big)\big)\left(\frac{1-\alpha}{e}\right)^{Q-1}\binom{n-r}{q-r}\right)^{N(1-\alpha)}.$$

Proof. Let $\boldsymbol{S} \in \mathcal{S}^{\mathrm{ext}}(S^*)$ be a uniformly random completion of S. Let $H(\boldsymbol{S}) = \log\big|\mathcal{S}^{\mathrm{ext}}(S^*)\big|$ be the entropy of $\boldsymbol{S}$.

Let $G = G(S^*)$. For each $e \in G$, let $\boldsymbol{S}_e$ be the q-edge that includes e in $\boldsymbol{S}$. So, the sequence $(\boldsymbol{S}_e)_{e\in G}$ determines $\boldsymbol{S}$. For any ordering on the r-edges of G, we have

$$H(\boldsymbol{S}) = \sum_{e\in G} H\big(\boldsymbol{S}_e \,\big|\, (\boldsymbol{S}_{e'} : e' < e)\big).$$

Now, a sequence $\lambda \in [0,1]^{E(G)}$ with all λ_e distinct induces an ordering on the r-edges of G, with $e' < e$ when $\lambda_{e'} > \lambda_e$. Let $\boldsymbol{R}_e(\lambda)$ be an upper bound on the number of possible values for $\boldsymbol{S}_e$ given the information $(\boldsymbol{S}_{e'} : \lambda_{e'} > \lambda_e)$, defined as follows. $\boldsymbol{R}_e = 1$ if $\lambda_{e'} > \lambda_e$ for any of the r-subsets $e' \subseteq \boldsymbol{S}_e$ (because in this case $\boldsymbol{S}_e$ is determined), and otherwise $\boldsymbol{R}_e$ is 1 plus the number of q-sets of vertices $X \ne \boldsymbol{S}_e$ containing e such that

- $f \in G$ for each r-subset $f \subseteq X$, and
- $\lambda_{e'} < \lambda_e$ for each of the r-edges $e' \in G$ included in any of the q-edges of $\boldsymbol{S}$ that include the r-subsets in $\{f : f \subseteq X, |f| = r\} \backslash \boldsymbol{S}_e$.

Note that the number of edges e' considered by the second bullet point is $Q(Q-1)$ unless $|X \cap B| > r$ for some q-edge B in $\boldsymbol{S}$ (in this case we say X is "bad"). By Proposition 2.8, the number of choices for X is

$$\big(1 \pm O\big(n^{-a}\big)\big)(1-\alpha)^{Q-1}\frac{n^{q-r}}{(q-r)!} = \big(1 \pm O\big(n^{-a}\big)\big)(1-\alpha)^{Q-1}\binom{n-r}{q-r}$$

and the number of those that are bad is at most $\sum_{i=r+1}^{q} N\binom{n}{q-r-i} = O\big(n^{q-r-1}\big)$.

Now, note that for any λ we have

$$H(\boldsymbol{S}_e \,|\, (\boldsymbol{S}_{e'} : \lambda_{e'} > \lambda_{e'})) \le \mathbb{E}[\log \boldsymbol{R}_e(\lambda)].$$

Let $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_e)_{e\in G}$ be a sequence of independent random variables, where each $\boldsymbol{\lambda}_e$ has the uniform distribution in $[0,1]$. (Note that almost surely each $\boldsymbol{\lambda}_e$ is distinct). By linearity of expectation and the tower law,

$$H(\boldsymbol{S}) = \sum_{e\in G} \mathbb{E}[\log \boldsymbol{R}_e(\boldsymbol{\lambda})].$$

Next, for any $S \in \mathcal{S}^{\mathrm{ext}}(S^*)$ and $\lambda \in [0,1]$, let

$$R_e^{S,\lambda} = \mathbb{E}\big[\boldsymbol{R}_e(\boldsymbol{\lambda}) \,\big|\, \boldsymbol{S} = S,\, \boldsymbol{\lambda}_e = \lambda,\, \boldsymbol{\lambda}_{e'} < \boldsymbol{\lambda}_e \text{ for all } e' \subseteq \boldsymbol{S}_e \backslash \{e\}\big].$$

By the discussion above, and linearity of expectation, we have

$$\begin{aligned} R_e^{S,\lambda} &= \left(\big(1 \pm O\big(n^{-a}\big)\big)(1-\alpha)^{Q-1}\binom{n-r}{q-r} - O\big(n^{q-r-1}\big)\right)\lambda^{Q(Q-1)} \\ &= \big(1 \pm O\big(n^{-a}\big)\big)(1-\alpha)^{Q-1}\binom{n-r}{q-r}\lambda^{Q(Q-1)}. \end{aligned}$$

Now, by Jensen's inequality,

$$\mathbb{E}\big[\log \boldsymbol{R}_e(\boldsymbol{\lambda}) \,\big|\, \boldsymbol{S} = S,\, \boldsymbol{\lambda}_e = \lambda,\, \boldsymbol{\lambda}_e = \lambda,\, \boldsymbol{\lambda}_{e'} < \boldsymbol{\lambda}_e \text{ for all } e' \subseteq \boldsymbol{S}_e \backslash \{e\}\big] \le \log R_e^{S,\lambda},$$

and

$$\Pr\big(\boldsymbol{\lambda}_e = \lambda,\, \boldsymbol{\lambda}_{e'} < \boldsymbol{\lambda}_e \text{ for all } e' \subseteq \boldsymbol{S}_e \backslash \{e\} \,\big|\, \boldsymbol{\lambda}_e = \lambda\big) = \lambda^{Q-1},$$

so

$$\mathbb{E}[\log \boldsymbol{R}_e(\boldsymbol{\lambda}) \,|\, \boldsymbol{S} = S,\, \boldsymbol{\lambda}_e = \lambda] \le \lambda^{Q-1} \log R_e^{S,\lambda} + \big(1 - \lambda^{Q-1}\big)\log 1 = \lambda^2 \log R_e^{S,\lambda}.$$

For any $S \in \mathcal{S}^{\mathrm{ext}}(S^*)$, we then have

$$\begin{aligned} \mathbb{E}[\log \boldsymbol{R}_e(\boldsymbol{\lambda}) \,|\, \boldsymbol{S} = S] &\le \mathbb{E}\Big[\lambda^{Q-1} \log R_e^{S,\boldsymbol{\lambda}_e}\Big] \\ &= \int_0^1 \lambda^{Q-1} \log\left(\big(1 \pm O\big(n^{-a}\big)\big)(1-\alpha)^{Q-1}\binom{n-r}{q-r}\lambda^{Q(Q-1)}\right) \mathrm{d}\,\lambda \\ &= \frac{1}{Q}\left(\log\left(\big(1 \pm O\big(n^{-a}\big)\big)(1-\alpha)^{Q-1}\binom{n-r}{q-r}\right) - (Q-1)\right), \end{aligned}$$

using the fact that $\int_0^1 \lambda^{A-1} \log\big(C\lambda^B\big)\,\mathrm{d}\,t = A^{-1}\log C - A^{-2}B$ for any $A, B, C > 0$. We conclude

$$\begin{aligned} \log\big|\mathcal{S}^{\mathrm{ext}}(S^*)\big| &\le H(\boldsymbol{S}) \\ &\le \sum_{e \in G} \mathbb{E}[\log \boldsymbol{R}_e(\boldsymbol{\lambda})] \\ &\le (N - \alpha N)\frac{1}{Q}\left(\log\left(\big(1 \pm O\big(n^{-a}\big)\big)(1-\alpha)^{Q-1}\binom{n-r}{q-r}\right) - (Q-1)\right), \end{aligned}$$

which is equivalent to the theorem statement. □

For the lower bound, we will count ordered Steiner systems.

Theorem A.2. *There is $h \in \mathbb{N}$ such that for any $a > 0$, there is $b = b(a)$ such that the following holds. For any $\alpha \in (0,1)$ and any $S^* \in \mathcal{O}_{\alpha N}^{n^{-a},h}$,*

$$\big|\mathcal{O}^{\mathrm{ext}}(S^*)\big| \le \big|\mathcal{S}^{\mathrm{ext}}(S^*)\big| \le \left(\big(1 + O\big(n^{-a}\big)\big)\left(\frac{1-\alpha}{e}\right)^{Q-1}\binom{n-r}{q-r}\right)^{N(1-\alpha)}(N - \alpha N)!.$$

To prove Theorem A.2 we will need an analysis of the clique removal process (see Appendix B) and the following immediate consequence of [13, Theorem 2.1].

Theorem A.3. *There are $h \in \mathbb{N}$, $\varepsilon_0, a \in (0,1)$ and $n_0, \ell \in \mathbb{N}$ such that if $S \in \mathcal{S}_m^{\varepsilon,h}$ is a partial system with $n \geq n_0$, with $d(G(S)) = 1 - m/N \geq n^{-a}$ and $\varepsilon \leq \varepsilon_0 d(G)^\ell$, then S can be completed to an Steiner system.*

Sketch of proof of Theorem A.2. Let h be as in Theorem A.3. Fix small c (smaller than $a\,b(a,h)$ in the notation of Theorem B.1), let $\varepsilon = n^{-c}$ and let $M = (1-\varepsilon)N$. Let $\boldsymbol{S} \in \mathbb{R}(S, M-m)$ be the result of running the clique removal process on $G(S)$ to build a partial system extending S, until there are M hyperedges. Let $\mathcal{O}^*$ be the set of M-hyperedge (ε,h)-quasirandom ordered partial systems $S \in \mathcal{O}_M^{\varepsilon,h}$ extending S. By Proposition 2.8, for each $S \in \mathcal{O}^*$, the number of copies of $K_q^{(r)}$ in each $G(S_i)$ is $(1 \pm O(n^{-c}))(1-i/N)^Q n^q/q!$ by quasirandomness. So,

$$\Pr(\boldsymbol{S} = S) \leq \prod_{i=\alpha N}^{M} \frac{1}{(1-O(n^{-c}))(1-i/N)^Q n^q/q!}.$$

By Theorem B.1 we have

$$\sum_{S \in \mathcal{O}^*} \Pr(\boldsymbol{S} = S) = 1 - o(1),$$

so

$$\begin{aligned}|\mathcal{O}^*| &\geq (1-o(1)) \prod_{i=\alpha N}^{M} \big(1-O\big(n^{-c}\big)\big)\left(1-\frac{i}{N}\right)^Q n^q/q!. \\ &= \left(\big(1-O\big(n^{-c}\big)\big)\frac{n^q}{q!}\right)^{(1-\alpha)N} \exp\left(Q\sum_{i=\alpha N}^{M} \log\left(1-\frac{i}{N}\right)\right).\end{aligned}$$

Now, note that

$$\sum_{i=\alpha N}^{M} \frac{1}{N}\log\left(1-\frac{i+1}{N}\right) \leq \int_{\alpha}^{(1-\varepsilon)} \log(1-t)\,\mathrm{d}\,t \leq \sum_{i=\alpha N}^{M} \frac{1}{N}\log\left(1-\frac{i}{N}\right).$$

We compute

$$\begin{aligned}\sum_{i=\alpha N}^{M}\left(\log\left(1-\frac{i}{N}\right)-\log\left(1-\frac{i+1}{N}\right)\right) &= \sum_{i=\alpha N}^{M}\log\left(1+\frac{1}{N-(i+1)}\right) \\ &\leq \sum_{i=\alpha N}^{M}\frac{1}{N-(i+1)} \\ &= O(\log n).\end{aligned}$$

so, noting that $\int \log p\,\mathrm{d}\,p = p(\log p - 1)$,

$$\begin{aligned}Q\sum_{i=\alpha N}^{M}\log\left(1-\frac{i}{N}\right) &= QN\int_{\alpha}^{(1-\varepsilon)}\log(1-t)\,\mathrm{d}\,t + O(\log n) \\ &= QN\int_{\varepsilon}^{(1-\alpha)}\log p\,\mathrm{d}\,p + O(\log n) \\ &= QN((1-\alpha)(\log(1-\alpha)-1) - \varepsilon(\log\varepsilon - 1)) + O(\log n),\end{aligned}$$

$$\exp\left(Q\sum_{i=\alpha N}^{M}\log\left(1-\frac{i}{N}\right)\right) = \left(\big(1+O\big(n^{-c}\log n\big)\big)\frac{1-\alpha}{e}\right)^{QN(1-\alpha)}.$$

For $b < c$, it follows that

$$|\mathcal{O}^*| \geq \left(\left(1 - O\left(n^{-b}\right)\right)\frac{n^q(1-\alpha)^Q}{q!e^Q}\right)^{(1-\alpha)N}$$
$$= \left(\left(1 - O\left(n^{-b}\right)\right)\left(\frac{1-\alpha}{e}\right)^{Q-1}\binom{n-r}{q-r}\right)^{(1-\alpha)N}(N-\alpha N)!$$

By Theorem A.3 (assuming b is small enough) it follows that

$$\left|\mathcal{O}^{\mathrm{ext}}(S^*)\right| \geq \left(\frac{n(1-\alpha)^2}{e^2}\left(1 - O\left(n^{-b}\right)\right)\right)^{(1-\alpha)N}(N-\alpha N)!. \qquad \square$$

Lemma 2.5 then immediately follows from Theorem A.1 and Theorem A.2.

B Random clique removal

In this section we give a very simple analysis of the clique removal process. The analysis here is rather crude and quite standard, but we could not find an existing source for precisely the result we need. As in Section 2, we consider q and r to be fixed, let $Q = \binom{q}{r}$, let $N = \binom{n}{r}/Q$ and impose that $\binom{q-i}{r-i}$ divides $\binom{n-i}{r-i}$ for every $0 \leq i \leq r-1$.

As introduced in Section 2, the clique removal process is defined as follows. We start with a graph G with say $N - Qm$ edges, then iteratively delete (the edges of) a copy of $K_q^{(r)}$ chosen uniformly at random from all copies of $K_q^{(r)}$ in the remaining graph. Let

$$G = \boldsymbol{G}(m), \boldsymbol{G}(m+1), \dots$$

be the sequence of random graphs generated by this process. This process cannot continue forever, but we "freeze" the process instead of aborting it; if $\boldsymbol{G}(\boldsymbol{M})$ is the first graph in the sequence with no copies of $K_q^{(r)}$, then let $\boldsymbol{G}(i) = \boldsymbol{G}(\boldsymbol{M})$ for $i \geq \boldsymbol{M}$.

Our objective in this section is to show that if G is quasirandom then the clique removal process is likely to maintain quasirandomness and unlikely to freeze until nearly all edges are gone.

Theorem B.1. *For all $h \geq 2$ and $a > 0$ there is $b(a,h) > 0$ such that the following holds. Let $n^{-a} \leq \varepsilon < 1/2$ and suppose G is a (ε, h)-quasirandom graph with $N - 3m = N - 3\alpha N$ edges. Then a.a.s. $\boldsymbol{M} \geq \left(1 - n^{-b}\right)N$ and moreover for each $N - m \leq i \leq \left(1 - \varepsilon^b\right)N$, $\boldsymbol{G}(i)$ is $\left(\varepsilon^b, h\right)$-quasirandom.*

Note that $K_n^{(r)}$ is $(O(1/n), h)$-quasirandom for any h, so in particular when we start the clique removal process from $G = K_n^{(r)}$ it typically runs almost to completion.

To prove Theorem B.1, it will be convenient to use Freedman's inequality [7, Theorem 1.6], as follows. (This was originally stated for martingales, but it also holds for supermartingales with the same proof). Here and in what follows, we write $\Delta X(i)$ for the one-step change $X(i+1) - X(i)$ in a variable X.

Lemma B.2. *Let $\boldsymbol{X}(0), \boldsymbol{X}(1), \dots$ be a supermartingale with respect to a filtration $(\mathcal{F}_i)$. Suppose that $\Delta\boldsymbol{X}(i) \leq K$ for all i, and let $V(i) = \sum_{j=0}^{i-1}\mathbb{E}\left[\left(\Delta\boldsymbol{X}(j)\right)^2 \,\middle|\, \mathcal{F}_j\right]$. Then for any $t, v > 0$,*

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

Proof of Theorem B.1. For a set A of $(r-1)$-sets of vertices with $|A| \le h$, let $\boldsymbol{Y}_A(i) = \big|\bigcap_{W\in A} N_{\boldsymbol{G}(i)}(W)\big|$. Let $p(i) = (1 - i/N)$ (and let $p^k(i) = (1-i/N)^k$), so that $p^{|A|}(i)n$ is the predicted trajectory of each $\boldsymbol{Y}_A(i)$.

Fix some large C and small c to be determined. We will choose $b < c/(C+1)$ so that $e(i) := p(i)^{-C}\varepsilon^c \le \varepsilon^b$ for $i \le N\big(1-\varepsilon^b\big)$. This means that if the conditions

$$\begin{aligned}\boldsymbol{Y}_A(i) &\ge p^{|A|}(i)n(1+e(i)),\\ \boldsymbol{Y}_A(i) &\le p^{|A|}(i)n(1-e(i))\end{aligned}$$

are satisfied for all A, then $\boldsymbol{G}(i)$ is $(e(i),h)$-quasirandom (therefore $\big(\varepsilon^b,h\big)$-quasirandom).

Let $\boldsymbol{T}'$ be the smallest index $i \ge m$ such that for some A, the above equations are violated (let $\boldsymbol{T}' = \infty$ if this never happens). Let $\boldsymbol{T} = \boldsymbol{T}' \wedge N\big(1-\varepsilon^b\big)$. Define the stopped processes

$$\begin{aligned}\boldsymbol{Y}_A^+(i) &= \boldsymbol{Y}_A(i\wedge\boldsymbol{T}) - p^{|A|}(i\wedge\boldsymbol{T})n(1+e(i\wedge\boldsymbol{T})),\\ \boldsymbol{Y}_A^-(i) &= -\boldsymbol{Y}_A(i\wedge\boldsymbol{T}) + p^{|A|}(i\wedge\boldsymbol{T})n(1-e(i\wedge\boldsymbol{T})).\end{aligned}$$

We want to show that for each A and each $s\in\{+,-\}$, the process $\boldsymbol{Y}_A^s = (\boldsymbol{Y}_A^s(i), \boldsymbol{Y}_A^s(i+1),\dots)$ is a supermartingale, and then we want to use Lemma B.2 and the union bound to show that a.a.s. each $\boldsymbol{Y}_A^s$ only takes negative values.

To see that this suffices to prove Theorem B.1, note that if $i < \boldsymbol{T}$ then by Proposition 2.8 the number of copies of $K_q^{(r)}$ is

$$\boldsymbol{R}(i) = (1\pm O(e(i)))p^Q(i)n^q/q! > 0.$$

This means $\boldsymbol{T} \le \boldsymbol{M}$, so the event that each $\boldsymbol{Y}_A^s$ only takes negative values contains the event that each $\boldsymbol{G}(i)$ is non-frozen and sufficiently quasirandom for $i \le N\big(1-\varepsilon^b\big)$.

Let $N_A(i) = \bigcap_{W\in A} N_{\boldsymbol{G}(i)}(W)$, so that $\boldsymbol{Y}_A(i) = |N_A(i)|$. Fix A, and consider $x \in N_A(i)$, for $i < \boldsymbol{T}$. The only way we can have $x \notin N_A(i+1)$ is if we remove a copy of $K_q^{(r)}$ containing an r-edge $\{x\}\cup W$ for some $W\in A$. Now, for each $W\in A$, the number of copies of $K_q^{(r)}$ in $\boldsymbol{G}(i)$ containing the edge $\{x\}\cup W$ is $(1\pm O(e(i)))p^{Q-1}(i)n^{q-r}/(q-r)!$ by Proposition 2.8. The number of copies of $K_q^{(r)}$ containing $\{x\}\cup W$ and $\{x\}\cup W'$ for two different $W, W'\in A$ is at most $O\big(n^{q-r-1}\big)$. So,

$$\begin{aligned}\Pr(x\notin N_A(i+1)) &= \frac{1}{\boldsymbol{R}(i)}\left(\sum_{W\in A}(1\pm O(e(i)))p^{Q-1}(i)n^{q-r}/(q-r)! - O\big(n^{q-r-1}\big)\right)\\ &= |A|(1\pm O(e(i)))p^{-1}(i)/N.\end{aligned}$$

For $i < \boldsymbol{T}$ we have $|N_A(i)| = (1\pm e(i))p^{|A|}(i)n$, so by linearity of expectation

$$\begin{aligned}\mathbb{E}[\Delta\boldsymbol{Y}_A(i)\,|\,\boldsymbol{G}(i)] &= -|A|(1\pm O(e(i)))p^{|A|-1}(i)n/N\\ &= -|A|p^{|A|-1}(i)n/N + O\Big(e(i)p^{|A|-1}(i)/n\Big).\end{aligned}$$

Note also that we have the bound $\Delta\boldsymbol{Y}_A(i) \le q-1 = O(1)$ (with probability 1). Also, for fixed k we

have

$$\begin{aligned}\Delta p^k(i) &= \left(1-\frac{i+1}{N}\right)^k - \left(1-\frac{i}{N}\right)^k \\ &= \left(1-\frac{i}{N}\right)^k\left(\left(\frac{N-i-1}{N-i}\right)^k - 1\right) \\ &= p^k(i)\left(\left(1-\frac{1}{N-i}\right)^k - 1\right) \\ &= p^k(i)\left(-\frac{k}{N-i} + O\left(\frac{1}{(N-i)^2}\right)\right) \\ &= -\frac{kp^{k-1}(i)}{N}\left(1+O\left(\frac{1}{N}p(i)\right)\right) \\ &= -\frac{kp^{k-1}(i)}{N} + o\Big(e(i)p^{k-1}(i)/N\Big),\end{aligned}$$

and with ep^kdenoting the pointwise product $i \mapsto e(i)p^k(i)$, for C much larger than k we have

$$\begin{aligned}\Delta\Big(ep^k\Big)(i) &= \varepsilon^c \Delta p^{k-C}(i) = -\varepsilon^c \Theta\Big(Cp^{k-C-1}(i)/N\Big) \\ &= \Theta\Big(Ce(i)p^{k-1}(i)/N\Big).\end{aligned}$$

For large C it follows that

$$\mathbb{E}\big[\Delta \boldsymbol{Y}_A^+(i) \,\big|\, \boldsymbol{G}(i)\big] = \mathbb{E}[\Delta \boldsymbol{Y}_A(i) \,|\, \boldsymbol{G}(i)] - \Delta p^{|A|}(i)n - \Delta\Big(ep^{|A|}\Big)(i)n \le 0,$$

and similarly

$$\mathbb{E}\big[\Delta \boldsymbol{Y}_A^-(i) \,\big|\, \boldsymbol{G}(i)\big] \le 0$$

for $i < \boldsymbol{T}$. (For $i \ge \boldsymbol{T}$ we trivially have $\Delta \boldsymbol{Y}_A^s(i) = 0$) Since each $\boldsymbol{Y}_A^s$ is a Markov process, it follows that each is a supermartingale. Now, we need to bound $\Delta \boldsymbol{Y}_A^s(i)$ and $\mathbb{E}\Big[(\Delta \boldsymbol{Y}_A^s(i))^2 \,\Big|\, \boldsymbol{G}(i)\Big]$, which is easy given the preceding calculations. First, recalling that $\Delta \boldsymbol{Y}_A(i) = O(1)$ and noting that $\Delta p^k(i), \Delta\big(ep^k\big)(i) = O(1/N)$ we immediately have $|\Delta \boldsymbol{Y}_A^s(i)| = O(1)$. Noting in addition that $\mathbb{E}[\Delta \boldsymbol{Y}_A(i) \,|\, \boldsymbol{G}(i)] = O(1/n)$, we have

$$\mathbb{E}\Big[(\Delta \boldsymbol{Y}_A^s(i))^2 \,\Big|\, \boldsymbol{G}(i)\Big] = O(\mathbb{E}[\Delta \boldsymbol{Y}_A^s(i) \,|\, \boldsymbol{G}(i)]) = O\Big(\frac{n}{N}\Big).$$

Since $\boldsymbol{T} \le N$, it follows that

$$\sum_{i=0}^{\infty} \mathbb{E}\Big[(\Delta \boldsymbol{Y}_A^s(i))^2 \,\Big|\, \boldsymbol{G}(i)\Big] = O\Big(N\frac{n}{N}\Big) = O(n).$$

Provided $c < 1$ (and recalling that $\varepsilon < 1/2$), applying Lemma B.2 with $t = e(m)p^{|A|}(m)n - \varepsilon p^{|A|}(m)n = \Omega(n\varepsilon^c)$ and $v = O(n)$ then gives

$$\Pr(\boldsymbol{Y}_A^s(i) > 0 \text{ for some } i) \le \exp\big(-O\big(n\varepsilon^{2c}\big)\big).$$

So, if $2c < \log_\varepsilon n \le a$, the union bound over all A, s finishes the proof. □