Optimal spread for spanning subgraphs of Dirac hypergraphs

Tom Kelly [a,1], Alp Müyesser [b], Alexey Pokrovskiy [b]

[a] *School of Mathematics, Georgia Institute of Technology, USA*
[b] *Department of Mathematics, University College London, UK*

A R T I C L E   I N F O

A B S T R A C T

Let G and H be hypergraphs on n vertices, and suppose H has large enough minimum degree to necessarily contain a copy of G as a subgraph. We give a general method to randomly embed G into H with good "spread". More precisely, for a wide class of G, we find a randomised embedding $f : G \hookrightarrow H$ with the following property: for every s, for any partial embedding f' of s vertices of G into H, the probability that f extends f' is at most $O(1/n)^s$. This is a common generalisation of several streams of research surrounding the classical Dirac-type problem.

For example, setting $s = n$, we obtain an asymptotically tight lower bound on the number of embeddings of G into H. This recovers and extends recent results of Glock, Gould, Joos, Kühn, and Osthus and of Montgomery and Pavez-Signé regarding enumerating Hamilton cycles in Dirac hypergraphs. Moreover, using the recent developments surrounding the Kahn–Kalai conjecture, this result implies that many Dirac-type results hold robustly, meaning G still embeds into H after a random sparsification of its edge set. This allows us to recover a recent result of Kang, Kelly, Kühn, Osthus, and Pfenninger and of Pham, Sah, Sawhney, and Simkin for perfect matchings, and obtain novel results for Hamilton cycles and factors in Dirac hypergraphs.

Notably, our randomised embedding algorithm is self-contained and does not require Szemerédi's regularity lemma or iterative absorption.

1. Introduction

A central theme in extremal graph theory is computing the minimum-degree thresholds for various properties. A prototypical result is Dirac's theorem [5], which states that n-vertex graphs with minimum degree at least $n/2$ are Hamiltonian. This is of course tight, as graphs with minimum degree less than $n/2$ need not be connected. Results of this flavour, concerned with minimum-degree thresholds for containment of spanning structures, are often referred to as *Dirac-type* theorems.

Although the assumption that $\delta(G) \geq n/2$ cannot be weakened in Dirac's theorem, we can say a lot more about n-vertex graphs with $\delta(G) \geq n/2$ (such graphs are henceforth referred to as Dirac graphs). Sárközy, Selkow, and Szemerédi [33] showed that Dirac graphs in fact contain $\Omega(n)^n$ many distinct Hamilton cycles, which is clearly optimal up to the value of the implied constant factor (the constant terms were later sharpened by Cuckler and Kahn [4]). We refer to results of this type as *enumeration* results.

We also know that the Hamilton cycles in Dirac graphs can survive *random edge deletions* (this phenomenon is often called *robustness*, see the survey of Sudakov [35]). Indeed, an influential result of Krivelevich, Lee, and Sudakov [22] states that for any Dirac graph G, the random subgraph obtained by keeping each edge of G independently with probability p where $p \geq C \log n / n$ (C is an absolute constant) is *asymptotically almost surely* (abbreviated a.a.s.) Hamiltonian. We remark that this result is a common generalisation of Dirac's theorem, as well as the classical result of Pósa [29] that the Erdős-Rényi random graph $\mathcal{G}(n, C \log n / n)$ is a.a.s. Hamiltonian.

Recently, random graph theory was revolutionised by Frankston, Kahn, Narayanan, and Park's [10] proof of the *fractional expectation threshold vs. threshold* conjecture of Talagrand [36] and Park and Pham's [27] proof of the even stronger Kahn–Kalai conjecture [17]. A crucial corollary of Talagrand's conjecture relates the so-called *spread measures* with thresholds. Roughly speaking, for the study of random graphs, this connection implies that if one can find a probability measure on copies of $G \subseteq K_n$ with good *spread* (see Definition 1.16), then G a.a.s. appears in $\mathcal{G}(n, p)$. Simply considering the uniform distribution on copies of $G \subseteq K_n$ already gives startlingly powerful results. An example that barely scratches the surface of this phenomenon is the following: If G is a cycle on n vertices, then the uniform distribution on copies of G in K_n (that is, on Hamilton cycles of K_n) turns out to be well-behaved enough that through Talagrand's conjecture, we can swiftly recover Pósa's [29] result about Hamiltonicity of $\mathcal{G}(n, C \log n / n)$.

Two recent independent papers by Kang, Kelly, Kühn, Osthus, and Pfenninger (KKKOP) [19] and Pham, Sah, Sawhney, and Simkin (PSSS) [28] demonstrated that *spreadness results* can yield common generalisations of enumeration and robustness results. If there is a probability measure on copies of G in H with good spread, where H is for example a hypergraph with minimum degree large enough to contain a copy of G, then Talagrand's conjecture gives for free that H contains G *robustly*, in a similar sense to the aforementioned result of Krivelevich, Lee, and Sudakov (see the next subsection for precise formulations). Moreover, in order for such a probability measure to have good spread, it must have large support; that is, there must be many copies of G in H.

Calculating the spread of copies of $G \subseteq K_n$ sampled uniformly at random is in many cases quite approachable. When H is not a complete graph, copies of $G \subseteq H$ sampled uniformly at random are hard to reason about, and it is unclear how one can compute the spread of such a random subgraph. Both KKKOP [19] and PSSS [28] circumvent this difficulty by employing powerful methods from extremal/probabilistic combinatorics such as Szemerédi's regularity lemma and iterative absorption to construct a probability measure on the desired subgraphs with good spread. The contribution of the current paper is a simpler method that can achieve the same task, in many cases in significantly more generality.

The notion of *vertex spreadness*, introduced by PSSS [28], is crucial. Instead of considering a random copy of G in H and analyzing the probability that a subset of edges of H appears in G, we consider a random *embedding* of $G \hookrightarrow H$ and analyze the probability that a subset of vertices of G are mapped to a subset of vertices of H. Our main results provide a randomized algorithm for embedding $G \hookrightarrow H$ so that the vertex spreadness matches up to a constant factor the vertex spreadness of embedding $G \hookrightarrow K_n$ uniformly at random, when G is a (hypergraph) cycle or the disjoint union of copies of some fixed hypergraph. This allows us to recover and extend many results from [19,28], as well as obtain novel enumeration results which extend recent results of Montgomery and Pavez-Signé [26]. In the next subsections, we provide precise formulations of our results.

1.1. Thresholds

Before stating our results, we make precise some concepts already alluded to in the introduction. For a k-uniform hypergraph H and $d \in [k-1]$, we let $\delta_d(H)$ be the minimum number of edges of H that any set of d vertices of $V(H)$ is contained in.

Definition 1.1 *(Minimum-degree/Dirac threshold).* Let $\mathcal{F}$ be an infinite family of k-uniform hypergraphs. By $\delta_{\mathcal{F},d}$ we denote the smallest real number δ such that for all $\alpha > 0$ and for all but finitely many n the following holds. Let n be such that there exists some $F \in \mathcal{F}$ with $|V(F)| = n$. Let H be an n-vertex k-uniform hypergraph with $\delta_d(H) \geq (\delta + \alpha)\binom{n-d}{k-d}$. Then, there exists some $F \in \mathcal{F}$ with $|V(F)| = n$ such that H contains a copy of F.

Note that for all $\mathcal{F}$ and d, the statement in Definition 1.1 holds for $\delta = 1$ and for the infimum over all such $\delta \in [0, 1]$, so $\delta_{\mathcal{F},d}$ is well-defined.

The notion of thresholds is central in the study of random graphs. Here we present a general definition for the threshold for a binomial random k-uniform hypergraph to contain some spanning subgraph.

Definition 1.2. Let $\mathcal{F}$ be an infinite family of k-uniform hypergraphs. Denote by $p = p_{\mathcal{F}}^*(n)$ the function so that for all $n \in \mathbb{N}$ (such that there exists some $F \in \mathcal{F}$ with $|V(F)| = n$), the binomial random hypergraph $\mathcal{G}^{(k)}(n, p)$ contains a copy of some $F \in \mathcal{F}$ with $|V(F)| = n$ with probability precisely $1/2$.

For example, if $\mathcal{F}$ contains one length-n cycle for every $n \in \mathbb{N}$, then $p_{\mathcal{F}}^*(n)$ is the threshold for $\mathcal{G}(n, p)$ to have a Hamilton cycle.

Definition 1.3 (*p-random sparsification*). Given $p \in [0, 1]$ and some hypergraph H, by H_p, we denote the random graph obtained by keeping every edge of H with probability p, making the decisions independently for each distinct edge of H.

Definition 1.4 (*RD-threshold/Robust Dirac threshold*). Let $\mathcal{F}$ be an infinite family of k-uniform hypergraphs, and let $p = p(n) \in [0, 1]$ be a function. By $\delta_{\mathcal{F},d,p}^{\mathrm{RD}}$, we denote the smallest real number δ such that for all $\alpha, \varepsilon > 0$, there exists $C > 0$ such that for all sufficiently large $n \in \mathbb{N}$ (such that there exists some $F \in \mathcal{F}$ with $|V(F)| = n$), for any n-vertex hypergraph H with $\delta_d(H) \geq (\delta + \alpha)\binom{n-d}{k-d}$, there exists some $F \in \mathcal{F}$ with $|V(F)| = n$ such that the random sparsification $H_{\hat{p}}$ contains a copy of F with probability at least $1 - \varepsilon$ as long as $\hat{p} \geq Cp(n)$. By $\delta_{\mathcal{F},d}^{\mathrm{RD}}$, we denote $\delta_{\mathcal{F},d,p_{\mathcal{F}}^*}^{\mathrm{RD}}$.

It is clear that $\delta_{\mathcal{F},d}^{\mathrm{RD}} \geq \delta_{\mathcal{F},d}$. The other inequality is a lot more interesting, and in the following cases, true.

Theorem 1.5. *For the following values of $\mathcal{F}$ and d, we have that $\delta_{\mathcal{F},d}^{\mathrm{RD}} = \delta_{\mathcal{F},d}$:*

(1) when $\mathcal{F} := \{C_n \colon n \in \mathbb{N}\}$ where C_n is a cycle on n vertices (Krivelevich, Lee, Sudakov [22]);
(2) for each $r \geq 2$, when $\mathcal{F} := \{n \times K_r \colon n \in \mathbb{N}\}$ (*r-clique-factors*) (Pham, Sah, Sawhney, Simkin [28], the $r = 3$ case was obtained by Allen, Böttcher, Corsten, Davies, Jenssen, Morris, Roberts, Skokan [1]);
(3) for each C, when $\mathcal{F}$ is made up of a single tree T_n on n vertices for each $n \in \mathbb{N}$, and $\Delta(T_n) \leq C$ (Pham, Sah, Sawhney, Simkin [28]);
(4) when $\mathcal{F} := \{n \times K_k^{(k)} \colon n \in \mathbb{N}\}$ (*k-uniform matchings*), for any $d \in [k-1]$ (independently Pham, Sah, Sawhney, Simkin [28] and Kang, Kelly, Kühn, Osthus, Pfenninger [19]).

We remark that the first three results concern graphs, so $d = 1$ in each case. Our main contribution in this paper is a unified strategy giving a far-reaching generalisation of the first, second, and fourth result stated above. In particular, this allows us to add hypergraph Hamilton cycles and F-factors for a much broader choice of F (the second result concerns clique factors and the fourth result concerns perfect matchings, which are $K_k^{(k)}$-factors, where $K_n^{(k)}$ denotes the complete k-uniform n-vertex hypergraph) to the list above. It is also conceivable that our methods work in even more generality than we explore in the current paper.

We also remark that the first and the second result above hold in a stronger sense, with *exact* minimum degree conditions as opposed to asymptotic ones, and Kang, Kelly, Kühn, Osthus, Pfenninger [19] obtained an exact result for the fourth result above in the case $d = k - 1$. We refer the reader to the corresponding papers for more details.

1.2. Hypergraph Hamilton cycles and Hamilton connectivity

We recall that for $0 \leq \ell < k$, a k-uniform hypergraph is called an ℓ-*cycle* if its vertices can be ordered cyclically such that each of its edges consists of k consecutive vertices and every two consecutive edges (in the natural order of the edges) share exactly ℓ vertices. In particular, $(k-1)$-cycles and 1-cycles are known as *tight cycles* and *loose cycles* respectively, and 0-cycles are matchings. For $n \in (k - \ell)\mathbb{N}$, we let $C_{n,k,\ell}$ denote the n-vertex ℓ-cycle with vertex set $[n]$ and edge set

$$\{[k + (k - \ell)(i - 1)] \setminus [(k - \ell)(i - 1)] : i \in [n/(k - \ell)]\},$$

where addition above is modulo n. We define k-uniform ℓ-paths $P_{n,k,\ell}$ analogously.

Let $\mathcal{C}_{\ell,k} \coloneqq \{C_{n,k,\ell} : n \in (k - \ell)\mathbb{N}\}$, and let $\mathcal{P}_{\ell,k} \coloneqq \{P_{n,k,\ell} : n \in (k - \ell)\mathbb{N}\}$.

Unlike in the setting of Theorem 1.5(2) and Theorem 1.5(4), we would like to find connected structures in our host graphs. We work with hypergraphs as opposed to graphs, so Pósa-rotation-extension techniques as in Theorem 1.5(1) are unavailable to us. For the same reason, relying on the regularity lemma (and related tools, such as the blow-up lemma) as in Theorem 1.5(3) would be far from straight-forward. Instead, we use as a black box a connection between $\delta_{\mathcal{F},d}$ and the following related parameter.

Definition 1.6 (*Hamilton connectivity threshold*)**.** By $\delta_{k,\ell,d}^{\mathrm{CON}}$ we denote the smallest real number δ such that for all $\alpha > 0$ and for all sufficiently large n the following holds. For any $F \in \mathcal{P}_{\ell,k}$ with $|V(F)| = n$, for any n-vertex k-uniform hypergraph H with $\delta_d(H) \geq (\delta + \alpha)\binom{n-d}{k-d}$, and any disjoint $S, T \subseteq V(H)$ with $|S| = |T| = \ell$, H contains a copy of F where the first and last ℓ vertices of F (in the natural ordering of the vertices induced by the edges) are embedded to S and T, respectively.

It is not hard to see that $\delta_{k,\ell,d}^{\mathrm{CON}} \geq \delta_{\mathcal{C}_{\ell,k},d}$, and it is also known that $\delta_{2,1,1}^{\mathrm{CON}} = \delta_{\mathcal{C}_{1,2},1}$ (see for example [11]), but as far as we are aware, the parameter $\delta_{k,\ell,d}^{\mathrm{CON}}$ has not been studied for other ranges of k, ℓ, d.

Proposition 1.7. *For each of the following range of parameters, we have that $\delta_{k,\ell,d}^{\mathrm{CON}} = \delta_{\mathcal{C}_{\ell,k},d}$.*

1. *$d = k - 1$;*
2. *$1 \leq \ell < k/2$ and $d = k - 2$;*
3. *$\ell = k/2$ and $k/2 < d \leq k - 1$ with k even.*

The first part of the above proposition is proved as Lemma 2.6 in [26]. See Section 7 for the short proof of the second and third part of the proposition. These assertions are a consequence of the more general phenomenon that whenever there is a "standard absorption proof" (as codified in [13]) that $\delta_{\mathcal{C}_{\ell,k},d} \leq \delta$, then $\delta_{k,\ell,d}^{\mathrm{CON}} \leq \delta$ as well (see Proposition 7.1). Therefore, it is quite feasible that Proposition 1.7 remains true for a wider range of the parameters. But we remark that known results imply, for example, that $\delta_{k,\ell,d}^{\mathrm{CON}} > \delta_{\mathcal{C}_{\ell,k},d}$ when $k = 3$, $\ell = 2$ and $d = 1$ (see [30]).

1.3. New robustness results for Hamilton cycles and factors

An important consequence of our main theorem (Theorem 1.14, to be stated in the next subsection) is that $\delta_{\mathcal{C}_{\ell,k},d}^{\mathrm{RD}} \leq \delta_{k,\ell,d}^{\mathrm{CON}}$ for *all* ranges of the parameters. This is implied by the combination of the following two theorems.

Theorem 1.8. *For every $\alpha > 0$ and $k \in \mathbb{N}$, there exists $C = C_{1.8}(\alpha, k)$ such that the following holds for every $d \in [k - 1]$. Let H be an n-vertex k-uniform hypergraph such that $\delta_d(H) \geq (\delta_{k,1,d}^{\mathrm{CON}} + \alpha)\binom{n-d}{k-d}$. If $(k-1) \mid n$ and $p \geq C \log n / n^{k-1}$, then a.a.s. a random subhypergraph H_p contains a loose Hamilton cycle.*

Theorem 1.9. *For every $\alpha > 0$ and $k \in \mathbb{N}$, the following holds for every $d \in [k - 1]$ and $\ell \in \{2, \ldots, k - 1\}$. Let H be an n-vertex k-uniform hypergraph such that $\delta_d(H) \geq (\delta_{k,\ell,d}^{\mathrm{CON}} + \alpha)\binom{n-d}{k-d}$. If $(k-\ell) \mid n$ and $p = \omega(1/n^{k-\ell})$, then a.a.s. a random subhypergraph H_p contains a Hamilton ℓ-cycle.*

We chose to state these two theorems separately, to highlight the difference in the corresponding values of p we need to work with for the cases of $\ell \leq 1$ and $\ell > 1$ cases for the assertion that $\delta_{\mathcal{C}_{\ell,k},d}^{\mathrm{RD}} \leq \delta_{k,\ell,d}^{\mathrm{CON}}$. Indeed, the function $p_{\mathcal{C}_{\ell,k}}^*$ is $\Theta(\log n / n^{k-1})$ when $\ell \leq 1$ whereas it is $\Theta(1/n^{k-\ell})$ when $\ell > 1$ (see [6,7]). Roughly speaking, this difference is caused by isolated vertices being the bottleneck for random hypergraphs to contain perfect matchings or loose Hamilton cycles ($\ell \leq 1$), whereas for $\ell > 1$, the bottleneck comes from a first-moment argument. Due to this subtlety, deriving Theorem 1.9 from Theorem 1.14 is more challenging. To circumvent this difficulty, we benefit from a recent generalisation of Spiro [34] of Talagrand's conjecture.

Combining this with the connection between the parameters $\delta_{k,\ell,d}^{\mathrm{CON}}$ and $\delta_{\mathcal{C}_{\ell,k},d}$, we derive the following, which considerably extends the families of hypergraphs listed in Theorem 1.5.

Corollary 1.10. *For each k, ℓ and d such that $\delta_{k,\ell,d}^{\mathrm{CON}} = \delta_{\mathcal{C}_{\ell,k},d}$, we have that $\delta_{\mathcal{C}_{\ell,k},d}^{\mathrm{RD}} = \delta_{\mathcal{C}_{\ell,k},d}$. In particular, for the ranges of k, ℓ and d described in Proposition 1.7, we have that $\delta_{\mathcal{C}_{\ell,k},d}^{\mathrm{RD}} = \delta_{\mathcal{C}_{\ell,k},d}$.*

Independently of our work, Anastos, Chakraborti, Kang, Methuku, and Pfenninger [2] also obtained that $\delta_{\mathcal{C}_{\ell,k},d}^{\mathrm{RD}} = \delta_{\mathcal{C}_{\ell,k},d}$ in the regime where $d = k - 1$ and $\ell < k/2$. In fact, in this regime they obtained a stronger result working with *exact* minimum degree thresholds. We conjecture that the hypothesis $\delta_{k,\ell,d}^{\mathrm{CON}} = \delta_{\mathcal{C}_{\ell,k},d}$ is not required in the above statement. See Section 8 for more details.

For a hypergraph F, let $\mathcal{F}_F$ be a family of disjoint unions of copies of F containing one element with $i|V(F)|$ vertices for every $i \in \mathbb{N}$, and let $\delta_{F,d} := \delta_{\mathcal{F}_F,d}$. Hence, if an n-vertex hypergraph H satisfies $\delta_d(H) \geq (\delta_{F,d} + \alpha)\binom{n-d}{k-d}$ and n is sufficiently large and divisible by $|V(F)|$, then H contains an F-*factor*; that is, H contains a spanning subgraph where every component is a copy of F. Our next theorem implies $\delta_{\mathcal{F}_F,d}^{\mathrm{RD}} = \delta_{F,d}$ when F is a *strictly* 1-*balanced* uniform hypergraph. For a hypergraph H with at least two vertices, the 1-*density* of H is $d_1(H) := |E(H)|/(|V(H)| - 1)$. A hypergraph F is 1-*balanced* if $d_1(F') \leq d_1(F)$ for every $F' \subseteq F$ and *strictly* 1-*balanced* if $d_1(F') < d_1(F)$ for every $F' \subsetneq F$. For strictly 1-balanced k-uniform F, Johansson, Kahn, and Vu [15] proved that the threshold for a binomial random k-uniform hypergraph to contain an F-factor is $\Theta(n^{-1/d_1(F)} \log^{1/|E(F)|} n)$.

Theorem 1.11. *For every $\alpha > 0$ and $k, r \in \mathbb{N}$, there exists $C = C_{1.11}(\alpha, k, r) > 0$ such that the following holds for every $d \in [k-1]$. Let F be a k-uniform r-vertex strictly 1-balanced hypergraph, and let H be an n-vertex k-uniform hypergraph such that $\delta_d(H) \geq (\delta_{F,d} + \alpha)\binom{n-d}{k-d}$. If $r \mid n$ and $p \geq C n^{-1/d_1(F)} \log^{1/|E(F)|} n$, then a.a.s. a random subhypergraph H_p contains an F-factor.*

Complete graphs are strictly 1-balanced, so Theorem 1.11 implies Theorem 1.5(2). A k-uniform k-vertex hypergraph consisting of a single edge is also strictly 1-balanced, so Theorem 1.11 implies Theorem 1.5(4).

1.4. Vertex-spread thresholds

The following definition, originally introduced in [28], is at the heart of the paper.

Definition 1.12. Let X and Y be finite sets, and let μ be a probability distribution over injections $\varphi : X \to Y$. For $q \in [0,1]$, we say that μ is q-*vertex-spread* if for every two sequences of distinct vertices $x_1, \ldots, x_s \in X$ and $y_1, \ldots, y_s \in Y$,

$$\mu\left(\{\varphi : \varphi(x_i) = y_i \text{ for all } i \in [s]\}\right) \leq q^s.$$

A *hypergraph embedding* $\varphi : G \hookrightarrow H$ of a hypergraph G into a hypergraph H is an injective map $\varphi : V(G) \to V(H)$ that maps edges of G to edges of H, so there is an embedding of G into H if and only if H contains a subgraph isomorphic to G. Note that the uniformly random embedding $\varphi : G \hookrightarrow H$ when G and H have the same vertex set and H is complete is just a uniformly random permutation of $V(H)$, which is $(e/|V(H)|)$-vertex-spread (using Stirling's approximation).

Definition 1.13 *(VSD-threshold/Vertex-spread Dirac threshold).* Let $\mathcal{F}$ be an infinite family of k-uniform hypergraphs. By $\delta_{\mathcal{F},d}^{\mathrm{VSD}}$ we denote the smallest real number δ such that for all $\alpha > 0$, there exists $C > 0$ such that the following holds for all but finitely many $F \in \mathcal{F}$. If $n = |V(F)|$ and H is any n-vertex k-uniform hypergraph with $\delta_d(H) \geq (\delta + \alpha)\binom{n-d}{k-d}$, then there is a (C/n)-vertex-spread distribution on embeddings of $F \hookrightarrow H$.

Note that for every family $\mathcal{F}$, we have that $\delta_{\mathcal{F},d} \leq \delta_{\mathcal{F},d}^{\mathrm{RD}}, \delta_{\mathcal{F},d}^{\mathrm{VSD}} \leq 1$.

The main results of this paper are that for every $k \in \mathbb{N}$,

1. $\delta_{\mathcal{C}_{k,\ell},d}^{\mathrm{RD}} \leq \delta_{\mathcal{C}_{k,\ell},d}^{\mathrm{VSD}} \leq \delta_{k,\ell,d}^{\mathrm{CON}}$ for every $\ell, d \in [k-1]$ and
2. $\delta_{\mathcal{F}_F,d}^{\mathrm{VSD}} = \delta_{F,d}$ for every k-uniform F and $d \in [k-1]$ and moreover $\delta_{\mathcal{F}_F,d}^{\mathrm{RD}} = \delta_{F,d}$ if F is strictly 1-balanced.

The following theorem implies $\delta_{\mathcal{C}_{k,\ell},d}^{\mathrm{VSD}} \leq \delta_{k,\ell,d}^{\mathrm{CON}}$, for all ranges of the parameters.

Theorem 1.14. *For every $\alpha > 0$ and $k \in \mathbb{N}$, there exists $C = C_{1.14}(\alpha, k)$ such that the following holds for every $\ell, d \in [k-1]$, and every sufficiently large n for which $k - \ell$ divides n. If H is an n-vertex k-uniform hypergraph such that $\delta_d(H) \geq (\delta_{k,\ell,d}^{\mathrm{CON}} + \alpha)\binom{n-d}{k-d}$, then there is a (C/n)-vertex-spread distribution on embeddings $C_{n,k,\ell} \hookrightarrow H$.*

Note a consequence of the above theorem is that whenever, $\delta_{k,\ell,d}^{\mathrm{CON}} = \delta_{\mathcal{C}_{k,\ell},d}$, we have that $\delta_{k,\ell,d}^{\mathrm{CON}} = \delta_{\mathcal{C}_{k,\ell},d} = \delta_{\mathcal{C}_{k,\ell},d}^{\mathrm{VSD}}$.

The following theorem implies $\delta_{\mathcal{F}_F,d}^{\mathrm{VSD}} = \delta_{F,d}$ for every k-uniform F and $d \in [k-1]$.

Theorem 1.15. *For every $\alpha > 0$ and $k, r \in \mathbb{N}$ there exists $C = C_{1.15}(\alpha, k, r) > 0$ such that the following holds for all $d \in [k-1]$ and all sufficiently large n for which r divides n. Let F be a k-uniform r-vertex hypergraph, and let G be the union of n/r disjoint copies of F. If H is an n-vertex k-uniform hypergraph such that $\delta_d(H) \geq (\delta_{F,d} + \alpha)\binom{n-d}{k-d}$, then there is a (C/n)-vertex-spread distribution on embeddings $G \hookrightarrow H$.*

1.4.1. The enumeration aspect

Note that if there is a (C/n)-vertex-spread distribution on embeddings $G \hookrightarrow H$, then

$$1 = \sum_{\varphi:G\hookrightarrow H} \mu(\{\varphi\}) \leq |\{\varphi : G \hookrightarrow H\}| \left(\frac{C}{n}\right)^n = |\{\varphi : G \hookrightarrow H\}| \exp\left(n\log C - n\log n\right),$$

so there are at least $\exp(n\log n - O(n))$ embeddings and at least $\exp(n\log n - O(n))/|\mathrm{Aut}(G)|$ copies of G in H. In particular, Theorems 1.14 and 1.15 imply that under their respective assumptions, the hypergraph H contains at least

- $\exp\left(n\log n - O(n)\right)$ Hamilton ℓ-cycles (for $\ell \neq 0$) and
- $\exp\left((1 - 1/|V(F)|)n\log n - O(n)\right)$ F-factors,

respectively, which strengthen recent results of Montgomery and Pavez-Signé [26] and Glock, Gould, Kühn, and Osthus [12]. This statement can be interpreted as a far-reaching generalisation of the result of Sárközy, Selkow, and Szemerédi [33] stated in the introduction.

We also remark that Ferber, Hardiman, and Mond [9] sharpened the asymptotics for ℓ-cycles for $\ell < k - 1$ and $d = k - 1$ by determining the leading constant in the $O(n)$ term. We refer the reader to [26] for more in depth discussion of the enumeration aspect of Dirac-type results.

1.5. Spreadness

As previously mentioned, deriving Theorem 1.8 and Theorem 1.9 from Theorem 1.14 requires us to pass through the fractional expectation thresholds breakthrough of Frankston, Kahn, Narayanan, and Park [10], and in fact, a generalisation of this breakthrough by Spiro [34]. A hypergraph is called *r-bounded* if each edge has size at most r.

Definition 1.16. Let $q \in [0,1]$, and let $r \in \mathbb{N}$. Let $(V, \mathcal{H})$ be an r-bounded hypergraph, and let μ be a probability distribution on $\mathcal{H}$. We say μ is *q-spread* if the following holds:

$$\mu\left(\{A \in \mathcal{H} : A \supseteq S\}\right) \leq q^{|S|} \text{ for all } S \subseteq V.$$

Frankston, Kahn, Narayanan, and Park (FKNP) [10] proved that if $(V, \mathcal{H})$ is an r-bounded hypergraph and $\mathcal{H}$ supports a q-spread distribution, then a p-random subset of V contains an edge in $\mathcal{H}$ a.a.s. if $p \geq Kq\log r$ as $r \to \infty$ (here, K is some absolute constant). We will consider hypergraphs of the form $(V, \mathcal{H})$ where $V \coloneqq E(H)$ and $\mathcal{H}$ is the set of (edge sets of) Hamilton ℓ-cycles in some hypergraph H. In this case, we sometimes abuse notation and write $\mu(A)$ instead of $\mu(\{A\})$ where $A \in E(H)$ or $\mu(F)$ instead of $\mu(\{E(F)\})$ when $F \subseteq H$.

Let $m_1(H) = \max_{H' \subseteq H : |V(H')| > 1} d_1(H)$. The following proposition allows us to connect spread distributions and vertex-spread distributions.

Proposition 1.17. *For every $C, k, \Delta > 0$, there exists $C' = C_{1.17}(C, k, \Delta) > 0$ such that the following holds for all sufficiently large n. Let H and G be n-vertex k-uniform hypergraphs. If there is a (C/n)-vertex-spread distribution on embeddings $G \hookrightarrow H$ and $\Delta(G) \leq \Delta$, then there is a $\left(C'/n^{1/m_1(G)}\right)$-spread distribution on subgraphs of H which are isomorphic to G.*

In the case when H is complete, since the uniform distribution on embeddings of G into H is $(e/|V(H)|)$-vertex-spread, Proposition 1.17 provides an upper bound on the spreadness for copies of G in H. For many natural choices for G (such as in the case of F-factors or Hamilton ℓ-cycles), this bound is in fact best possible due to the duality between spreadness and fractional expectation thresholds.

Moreover, Proposition 1.17 implies $\delta_{\mathcal{C}_{k,\ell},d}^{\mathrm{RD}} \leq \delta_{\mathcal{C}_{k,\ell},d}^{\mathrm{VSD}}$ for every $\ell \in \{0, 1\}$ and $d \in [k-1]$. Combined with Theorem 1.15, it yields a $(O(1)/n^{k-1})$-spread distribution on perfect matchings in sufficiently dense n-vertex k-uniform hypergraphs, which was proved independently by KKKOP [19, Theorem 1.5] and PSSS [28, Theorem 1.5] (in the $k \mid n$ case). Proposition 1.17 combined with Theorem 1.14 also yields a $(O(1)/n^{k-1})$-spread distribution on loose Hamilton cycles (see Lemma 6.5 for $m_1(C_{n,k,1})$) in sufficiently dense n-vertex k-uniform hypergraphs. Combined with the FKNP theorem, this result implies Theorem 1.8.

To prove Theorem 1.9, we need Spiro's [34] strengthening of the FKNP theorem – see Section 6.2. To prove Theorem 1.11, we combine Proposition 1.17 with a coupling argument of Riordan [31] – see Section 6.3.

Remark. In independent work, Joos, Lang, and Sanhueza-Matamala [16] also proved Theorems 1.8, 1.9, and 1.11. They also obtained these results from stronger results concerning spreadness, but they did not consider vertex spreadness. To prove Theorem 1.9, they used a result of Espuny Díaz and Person [8], which is generalized by the result of Spiro [34].

2. Preliminaries

We use standard notation for "hierarchies" of constants, writing $x \ll y$ to mean that there is a non-decreasing function $f : (0,1] \to (0,1]$ such that the subsequent statements hold for $x \leq f(y)$. Hierarchies with multiple constants are defined similarly. We omit rounding signs where they are not crucial.

We need the following well-known version of the Chernoff bounds.

Lemma 2.1 (Chernoff bound). *Let $X := \sum_{i=1}^{m} X_i$ where $(X_i)_{i \in [m]}$ is a sequence of independent indicator random variables, and let $\mathbb{E}[X] = \mu$. For every $\gamma \in (0,1)$, we have $\mathbb{P}[|X - \mu| \geq \gamma \mu] \leq 2e^{-\mu \gamma^2/3}$.*

We will also use a result of Gupta, Hamann, Müyesser, Parczyk, and Sgueglia [13], which is proved with a standard concentration inequality.

Lemma 2.2 *([13], Lemma 3.5). Let $k, \ell, d \in \mathbb{N}$, $0 < \delta' < \delta < 1$ and $1/n, 1/\ell \ll 1/k, \delta - \delta'$. Let H be a k-uniform n-vertex hypergraph with vertex set V and suppose that $\deg(D, V) \geq \delta n^{k-d}$ for each $D \in \binom{V}{d}$. If $A \subseteq V$ is a vertex set of size ℓ chosen uniformly at random, then for every $D \in \binom{V}{d}$ we have*

$$\mathbb{P}\left[\deg(D, A) < \delta' \ell^{k-d}\right] \leq 2\exp(-\ell(\delta - \delta')^2/2).$$

Finally, we need the following result due to McDiarmid, which appears in the textbook of Molloy and Reed [25, Chapter 16.2]. Here, a *choice* is the position that a particular element gets mapped to in a permutation.

Lemma 2.3 *(McDiarmid's inequality for random permutations). Let X be a non-negative random variable determined by a random permutation π of $[n]$ such that the following holds for some $c, r > 0$:*

1. *Interchanging two elements of π can affect the value of X by at most c*
2. *For any s, if $X \geq s$ then there is a set of at most rs choices whose outcomes certify that $X \geq s$.*

Then, for any $0 \leq t \leq \mathbb{E}[X]$,

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq t + 60c\sqrt{r\mathbb{E}[X]}) \leq 4\exp(-t^2/(8c^2 r\mathbb{E}[X])).$$

3. Proof overview

In this section we overview the proofs of Theorems 1.14 and 1.15. The proof of Theorem 1.15 is simpler, so we begin with this one. In fact, for the sake of simplicity, suppose $k = 2$, $d = 1$ and F is just a triangle, and the host graph H is a graph on n vertices, n is divisible by 3, and $\delta(H) \geq (2/3 + \alpha)n$. Allen, Böttcher, Corsten, Davies, Jenssen, Morris, Roberts, Skokan [1] recently proved Theorem 1.11 for this case (in fact, they proved the stronger result for $\alpha = 0$), but our result (in the $\alpha > 0$ case) is stronger in that it concerns vertex spreadness and also has a shorter proof. Thus, this special case is still of independent interest, and our proof generalises easily to other graph factors (additional ideas are required to handle connected spanning subgraphs, as we will explain later). For now, let us also assume that n is divisible by some integer C that is in turn divisible by 3.

Our goal is to embed a triangle factor in H in a $O(1/n)$-vertex-spread manner. The main idea is quite simple: we wish to first partition $V(H)$ into random "clusters" $U_1, \ldots, U_m$ each inducing a graph with good minimum degree, and then embed a

triangle factor in each cluster U_i in a deterministic fashion. For the latter task, we hope to rely only on $\delta(U_i) \geq (2/3 + \alpha/2)|U_i|$ and $|U_i|$ being divisible by 3 as a black box to argue that an embedding exists. This step is not randomised, and thus gains us no spreadness whatsoever, so we have to rely exclusively on the randomness of the partition to obtain spreadness on this simple algorithm. Let us first suppose that the partition is chosen uniformly at random given that each U_i has size C. A simple calculation shows that for every set of distinct vertices $x_1, \ldots, x_s \in V(H)$ and every function $f \colon [s] \to [m]$,

$$\mathbb{P}\left[x_i \in U_{f(i)} \text{ for each } i \in [s]\right] \leq \left(\frac{C}{n}\right)^s. \tag{1}$$

Suppose that with probability at least $1/2$, each $H[U_i]$ contains a triangle factor. Then, (1) is actually all we need, and by conditioning on each $H[U_i]$ containing a triangle factor, we obtain a $(2C/n)$-vertex-spread distribution on embeddings of triangle factors in H.

If $C \gg \log n$, then it is straightforward to show that indeed $\delta(H[U_i]) \geq (2/3+\alpha/2)|U_i|$ for each i with probability at least $1/2$ by an application of Chernoff's bound and a union bound over the clusters. By the Corrádi–Hajnal theorem [3], each $H[U_i]$ contains a triangle factor, as desired. This argument already implies the desired result up to a logarithmic factor, but to obtain the desired result, we need $C = O(1)$. In this regime, it may be the case that with high probability, a small percentage of the $H[U_i]$ contain isolated vertices, and these clusters cannot possibly contain triangle factors.

A central idea in the proof, replacing the more intricate iterative absorption strategies employed in [19,28], is a *random redistribution* argument. If C is a large constant, it will still be the case that most of the U_i have good minimum degree ($\delta(H[U_i]) \geq (2/3 + \alpha/2)|U_i|$); in this case, we call the cluster *good* and otherwise call it *bad*. The key idea is that we can randomly redistribute vertices of the bad clusters to the good ones while preserving (1) and the minimum degree property.

With high probability, it will be the case that each vertex v can be added to many other good clusters U_i while ensuring that $H[U_i \cup \{v\}]$ still has good minimum degree. One way to randomly redistribute would be the following. For each vertex $v \in V(H)$ living in a bad cluster, among the good clusters U_i so that $U_i + v$ is also good, choose some U_i uniformly at random, and re-define $U_i := U_i + v$. This actually would not break (1), but it would cost us the property that each $|U_i|$ is divisible by three, which is necessary for finding a triangle factor. More importantly, it would also cost us the property that each random set has the same size, which is surprisingly critical while showing that the corresponding random embedding is vertex-spread (see the proof of Lemma 5.3, in particular, Observation 5.4).

Hence, we would like to do the redistribution step while maintaining some control over the sizes of the clusters. We achieve this by ensuring each good cluster receives exactly one new element by adding the following step to the algorithm. (Thus, in the triangle factor case, we initially choose each U_i to satisfy $|U_i| \equiv 2 \mod 3$). We first define an auxiliary bipartite graph B between vertices v which need to be redistributed

and good clusters U_i, having $v \sim_B U_i$ only when $U_i + v$ also has good minimum degree. By potentially redistributing more vertices than necessary, we can ensure that this is a balanced bipartite graph, and it is not too difficult to check that B will have high enough minimum degree to contain a perfect matching with high probability. We want to redistribute randomly, in particular we want a redistribution that preserves (1). So in B, we are looking for a random perfect matching M so that the probability that M extends a given matching of size s is $O(1/n)^s$. In other words, we want to find a $O(1/n)$-spread (not vertex-spread) probability distribution on perfect matchings in B. Such a distribution on perfect matchings does exist, and in fact this assertion is just a weaker version of results already present in [28]. However the statement we need here has such a short proof that we include the simple outline below (a similar proof due to Pham is recorded in an earlier version of the paper [28] available on arXiv).

Lemma 3.1 *(Pham, Sah, Sawhney, Simkin [28]). There exists an absolute constant $C_{3.1}$ with the following property. If G is a balanced bipartite graph on $2n$ vertices with $\delta(G) \geq 3n/4$, then there exists a $(C_{3.1}/n)$-spread distribution on perfect matchings of G.*

Proof sketch. Consider a random subgraph of G' obtained by sampling C edges incident to each vertex of G uniformly at random (we allow edges to be selected twice). We claim that with probability at least $99/100$, G' contains a perfect matching. One can easily verify this by checking Hall's condition for G'. Now, condition G' on the event that it satisfies Hall's condition, and consider an arbitrary perfect matching of G'. This defines M, which is a random perfect matching of G. We claim the distribution of M is $O(1/n)$-spread. Indeed, consider a matching of M' of G of size s. Each edge e of M' can be included in M only if $e = \{x, y\}$ is sampled from the side of x or y, this event has probability at most $2C/(3n/4) = O(1/n)$. Furthermore, for disjoint edges e and e', these events are independent. This implies that M extends M' with probability at most $O(1/n)^s$, implying the desired spreadness. $\square$

In Section 4, the random partitioning aspect of the proof is treated rigorously, culminating in Lemma 4.1, the "random clustering lemma". This lemma might be of independent interest, and can potentially be used as a black box to solve adjacent problems. There are two aspects of Lemma 4.1 that we have not discussed. Firstly, we cannot assume that n is divisible by C for some constant $C = O(1)$, and therefore, we cannot guarantee that *all* clusters in our partition have the same size. We handle this difficulty by having a single exceptional cluster which has a different size, and luckily, it turns out that this is good enough for the applications. For the triangle-factor case, we can use that, if all of the non-exceptional clusters U_i have size divisible by three, then so does the exceptional one. This argument generalises and is sufficient for the proof of Theorem 1.15.

The more important aspect of Lemma 4.1 that we have not discussed is that in Theorem 1.14 (and possibly in future applications), we want to find structures which are

connected (such as Hamilton cycles), so we need to think about edges that go across the clusters. To reason about this, consider an auxiliary graph Q where vertices represent clusters U_i, and we have $U_i \sim_Q U_j$ whenever the bipartite graph between U_i and U_j has good minimum degree. Before the redistribution, Q itself may have many isolated vertices. Thus, in the redistribution step, this will be another aspect we have to consider, so that we can ensure that Q itself is well-connected. If Q itself has a Hamilton cycle, then we can use this cycle to reorder the clusters, and use this new ordering to embed a Hamilton ℓ-cycle in H path-by-path. That is, we can first find a Hamilton path in each cluster, and then connect them together along a Hamilton cycle of Q to find the desired Hamilton cycle. See Fig. 1.

In fact, we need access to *many* Hamilton cycles of Q. Indeed, if we were only working with a fixed Hamilton cycle of Q, we could guarantee at most $C!^{n/C}$ Hamilton cycles for any choice of $U_1, \ldots, U_m$, since each U_i of size C has at most $C!$ Hamilton paths, and we have at most n/C random sets. Also, it is easy to see there are at most $n!/(n/C)!$ choices for $U_1, \ldots, U_m$ (here we consider two choices the same if they are identical up to relabelling the U_i). Thus, in total, this approach could yield at most $C!^{n/C}n!/(n/C)!$ distinct Hamilton cycles. On the other hand, the definition of $O(1/n)$-vertex-spread requires us to find at least $(n/C)^n$ distinct Hamilton cycles, which is much larger than the previous quantity, as a simple application of Stirling's approximation shows.

More thought reveals that to maintain vertex spreadness in the algorithm, we need a vertex-spread distribution on Hamilton cycles of Q. This may seem a bit circular, but it will be the case that Q is almost complete, so the Dirac condition is satisfied with lots of room to spare. Furthermore, the problem we reduce to is a problem about graphs, whereas our general theorem concerns hypergraphs. The toolkit we have to find Hamilton cycles on graphs is significantly more extensive compared to hypergraphs. To conclude, to find a good ordering on Q, we rely on the following result.

Lemma 3.2. *There exists an absolute constant $C_{3.2}$ with the following property. If G is a graph with $\delta(G) \geq 3n/4$ and $e(G) \geq (1 - 1/C_{3.2})\binom{n}{2}$, then there exists a $(C_{3.2}/n)$-vertex-spread distribution on Hamilton cycles of G (that is, on embeddings $\varphi : C_{n,2,1} \hookrightarrow G$).*

There are elementary approaches to this problem, but here we will just give a short derivation of the lemma from [28, Lemma 7.3], which guarantees $O(1/n)$-vertex-spread distributions on Hamilton paths under the same hypotheses. The task is to just modify the distribution so that it gives Hamilton cycles instead.

Proof of Lemma 3.2. Set $C' = C(2, 1/4)$ from [28, Lemma 7.3]. We claim that $C_{3.2} = 10C'$ has the property in the lemma. First note that by [28, Lemma 7.3], there exists a (C'/n)-vertex-spread distribution μ on Hamilton paths of G (that is, on embeddings $\varphi : P_n \hookrightarrow G$, where P_n is $C_{n,2,1}$ with the edge $\{1, n\}$ removed). Let ϕ be a random embedding sampled according to μ, and let μ' be the resulting distribution when μ is conditioned on the embedding ϕ mapping the end points of P_n to adjacent vertices

of G, i.e. $\{\phi(1), \phi(n)\} \in E(G)$. In the distribution given by μ, for every fixed $u, v \in V(G)$, the probability that $\phi(1) = u$ and $\phi(n) = v$ is at most $(C'/n)^2$ by the vertex-spreadness of μ. Furthermore, by a union bound over all non-adjacent $u, v \in V(G)$, we have that $\mu(\{\phi(1), \phi(n)\} \notin E(G)) \le (C'/n)^2 (1/C_{3.2})n^2 \le 1/10$. So μ' is obtained from μ by conditioning on an event with probability at least $9/10$. Therefore, μ' is $(2C'/n)$-vertex-spread, so also $(C_{3.2}/n)$-vertex-spread. □

4. Random clustering lemma

This section is devoted to proving the following "random clustering lemma".

Lemma 4.1. *Let $1/n \ll 1/C' \ll 1/C \ll \alpha, \varepsilon, 1/k, 1/d, 1/t \le 1$ where $n, k, d, t, C \in \mathbb{N}$, and let $\delta \in [0, 1]$. If H is a k-uniform hypergraph on n vertices with $\delta_d(H) \ge (\delta + \alpha)\binom{n-d}{k-d}$, then there exists a random partition $\mathcal{U} := \{U_1, U_2, \ldots, U_m\}$ of $V(H)$ with the following properties:*

(1) $|U_i| = C$ for each $2 \le i \le m$ and $|U_1|$ equals $(C-1)C$ plus the remainder when n is divided by $(C-1)C$;
(2) $\delta_d(H[U_i]) \ge (\delta + \alpha/2)\binom{|U_i|}{k-d}$ for each $i \in [m]$;
(3) for every $i \in [m]$, there exist $u_i \in U_i$, $T_i \subseteq U_i \setminus \{u_i\}$ of size t, and $N_i \subseteq \mathcal{U} \setminus \{U_i\}$ of size at least $(1 - \varepsilon)m$ such that for every $U_j \in N_i$,

$$\delta_d(H[U_i \cup T_j \setminus \{u_i\}]) \ge (\delta + \alpha/2)\binom{|U_i \cup T_j|}{k - d} \quad and$$

$$\delta_d(H[U_j \cup T_i \setminus \{u_j\}]) \ge (\delta + \alpha/2)\binom{|U_j \cup T_i|}{k - d};$$

(4) for every set of distinct vertices $y_1, \ldots, y_s \in V(H)$ and every function $f \colon [s] \to [m]$,

$$\mathbb{P}\left[y_i \in U_{f(i)} \text{ for each } i \in [s]\right] \le \left(\frac{C'}{n}\right)^s.$$

For the remainder of this section, fix $1/n \ll 1/C' \ll 1/C \ll \alpha, 1/k, 1/d, 1/t \le 1$ where $n, k, d, t, C \in \mathbb{N}$ as in Lemma 4.1. Let r be $(C-1)C$ plus the remainder when n is divided by $(C-1)C$. Let $W_1 := [r]$, and for $i \ge 2$, let $W_i := [r + (C-1)(i-1)] \setminus \bigcup_{j=1}^{i-1} W_j = [r + (C-1)(i-2) + 1, r + (C-1)(i-1)]$. Note that $\{W_1, \ldots, W_m\}$ is a partition of $[n]$ for $m := (n - r)/(C-1) + 1$. Let H be a k-uniform hypergraph with vertex set $V := \{v_1, \ldots, v_n\}$ satisfying $\delta_d(H) \ge (\delta + \alpha)\binom{n-d}{k-d}$. Let $\pi : [n] \to [n]$ be a uniformly random permutation of $[n]$, and let $V_i := \{v_j : \pi(j) \in W_i\}$.

As described in Section 3, the partition of $V(H)$ into $V_1, \ldots, V_m$ is in a sense "close" to the one we want in Lemma 4.1 because the fourth condition holds, and the second and third condition hold apart from a few exceptions. Lemmas 4.3, 4.4, and 4.5 set up

precisely the conditions we need for our *random redistribution* argument to succeed. Each of these lemmas rely on the following lemma.

Lemma 4.2. *For every $T \subseteq V$ of size at most t and every $i \in [m]$,*

$$\mathbb{P}\left[\delta_d(H[V_i \cup T]) \geq (\delta + 2\alpha/3)\binom{|V_i \cup T|}{k-d}\right] \geq 1 - \exp(-C\alpha^2/200).$$

Moreover, the same inequality holds for random sets $T := \{v_j : \pi(j) \in W\}$ for every $W \subseteq [n]$ of size at most t.

Proof. Let $T \subseteq V$ of size at most t. Note that each V_i has the same distribution as a uniformly random subset of V of size $C - 1$ for every $i \in \{2, \ldots, m\}$, and similarly, V_1 has the same distribution as a uniformly random subset of V of size r. Moreover, for every $W \subseteq [n]$ and $D \subseteq V$ with $|W| = |D|$, in the distribution conditional on the event that $D = \{v_j : \pi(j) \in W\}$, each set $V_i \setminus D$ has the same distribution as a uniformly random subset of V of size $|W_i \setminus W|$.

For every $D' \subseteq T$ of size at most d and $W' \subseteq W_i$ of size $d - |D'|$, let $E_{D',W'}$ be the event that $|D| = d$ and $\deg(D, V_i) < (\delta + 3\alpha/4)\binom{|V_i|}{k-d}$, where $D := \{v_j : \pi(j) \in W'\} \cup D'$. Note that if $\overline{E_{D',W'}}$ holds for every such D' and W', then $\delta_d(H[V_i \cup T]) \geq (\delta + 3\alpha/4)\binom{|V_i|}{k-d} \geq (\delta + 2\alpha/3)\binom{|V_i \cup T|}{k-d}$. We claim that for every such D' and W', we have

$$\mathbb{P}[E_{D',W'}] \leq \exp(-C\alpha^2/100).$$

Indeed, by the law of total probability, it suffices to show

$$\mathbb{P}[E_{D',W'} \mid \{v_j : \pi(j) \in W'\} = D''] \leq \exp(-C\alpha^2/100)$$

for every $D'' \subseteq V \setminus D'$ of size $d - |D'|$, which follows from Lemma 2.2 with $D' \cup D''$ playing the role of D and $V_i \setminus D''$ playing the role of A. Since there are $\binom{|T| + |W_i|}{d} \leq (3C^2)^d$ (using $1/C \ll 1/t$ and $|W_i| \leq 2C^2$) choices for D' and W', the result follows by the union bound, since $1/C \ll \alpha, 1/d$.

If instead $T := \{v_j : \pi(j) \in W\}$ for some $W \subseteq [n]$ of size at most t, then the proof is essentially the same, so we omit it. $\square$

The next lemma will be used to ensure (3) holds. When applied with $T = \{v\}$ for a vertex v in a "bad" cluster, it also ensures that there are many good clusters to which v can be added in the random redistribution step.

Lemma 4.3. *With probability at least $99/100$, for every $T \subseteq V$ of size at most t, there are at least $(1 - 1/C^2)m$ sets V_i for which $\delta_d(H[V_i \cup T]) \geq (\delta + 2\alpha/3)\binom{|V_i \cup T|}{k-d}$ and for every $W \subseteq [n]$ of size at most t, the same property holds for the random set $T := \{v_j : \pi(j) \in W\}$.*

Proof. For every $T \subseteq V$ of size at most t, let X_T denote the number of random sets V_i such that the minimum degree condition above holds. By linearity of expectation and Lemma 4.2, we have $\mathbb{E}[X_T] \geq (1 - \exp(-C\alpha^2/200))m$. Note that interchanging two elements of the random permutation π either has no effect on $V_1, \ldots, V_m$ or changes the values of two random sets V_i and V_j and therefore can affect the value of X_T by at most 2. Also, if $X_T \geq s$, this can be certified by at most $2C^2s$ choices of the random permutation. Therefore, by Lemma 2.3 applied with $c = 2$, $r = 2C^2$, and $t = m/C^3$, we have

$$\mathbb{P}\left[X_T < (1 - 1/C^2)m\right] \leq 4\exp\left(-\frac{(m/C^3)^2}{64C^2m}\right) \leq \exp(-n/C^{10}).$$

By a union bound, we have that $X_T \geq (1 - 1/C^2)m$ for every $T \subseteq V$ of size at most t with probability at least $99/100$, as desired.

If instead $T := \{v_j : \pi(j) \in W\}$ for some $W \subseteq [n]$ of size at most t, then the proof is essentially the same, so we omit it. $\quad\square$

The next lemma ensures that most clusters are good. Moreover, in our random redistribution argument, we will assign every good cluster a vertex from a bad cluster. The lemma also ensures that for every good cluster (except V_1), there are many options for this vertex.

Lemma 4.4. *With probability at least* $99/100$, *for all but at most* $\exp\left(-\alpha^2 C/1000\right)m$ *many* $i \in \{2, 3, \ldots, m\}$, *there are at least* $(1 - \exp\left(-\alpha^2 C/500\right))n$ *vertices* $v \in V$ *such that* $\delta_d(H[V_i \cup \{v\}]) \geq (\delta + 2\alpha/3)\binom{|V_i \cup \{v\}|}{k-d}$.

Proof. By Lemma 4.2, we have that $\mathbb{P}\left[\delta_d(H[V_i \cup \{v\}]) < (\delta + 2\alpha/3)\binom{|V_i \cup \{v\}|}{k-d}\right] \leq \exp\left(-\alpha^2 C/200\right)$ for any $v \in V$ and any random set V_i. Call v *bad* for V_i if this low probability event holds. So for any random set V_i, we have $\mathbb{E}\left[|\{v \in V : v \text{ is bad for } V_i\}|\right] \leq \exp\left(-\alpha^2 C/200\right)n$ by linearity of expectation. By Markov's inequality, for any random set V_i, we have $\mathbb{P}\left[|\{v \in V : v \text{ is bad for } V_i\}| \geq \exp\left(-\alpha^2 C/500\right)n\right] \leq \exp\left(-\alpha^2 C/500\right)$. Let X denote the number of V_i such that $|\{v \in V : v \text{ is bad for } V_i\}| \geq \exp\left(-\alpha^2 C/500\right)n$. By linearity of expectation, $\mathbb{E}[X] \leq \exp\left(-\alpha^2 C/500\right)m$. By Markov's inequality, $\mathbb{P}\left[X \geq \exp\left(-C\alpha^2/1000\right)m\right] \leq \exp\left(-C\alpha^2/1000\right) \leq 1/100$, implying the desired statement. $\quad\square$

For every $i \in [m]$, choose some $W_i' \subseteq W_i$ of size t arbitrarily (irrespective of π), let $T_i := \{v_j : \pi(j) \in W_i'\}$, let N_i^+ be the set of $V_j \in \{V_1, \ldots, V_m\}$ for $j \neq i$ such that $\delta_d(H[V_j \cup T_i]) \geq (\delta + 2\alpha/3)\binom{|V_j \cup T_i|}{k-d}$, and let N_i^- be the set of V_j for $j \neq i$ such that $V_i \in N_j^+$. Note that $T_i \subseteq V_i$ for all $i \in [m]$. In order to ensure (1) holds, we will need to have $U_1 = V_1$. The next lemma ensures that U_1 (which will be V_1) satisfies (2) and (3).

Lemma 4.5. *With probability at least* $99/100$, $|N_1^-| \geq (1 - \exp(-\alpha^2 C/500))(m - 1)$.

Proof. By the "moreover" part of Lemma 4.2, we have that

$$\mathbb{P}\left[\delta_d(H[V_1 \cup T_i]) < (\delta + 2\alpha/3)\binom{|V_1 \cup T_i|}{k-d}\right] \leq \exp\left(-\alpha^2 C/200\right)$$

for every $i \geq 2$, so

$$\mathbb{P}\left[V_i \in N_1^-\right] \geq 1 - \exp\left(-\alpha^2 C/200\right).$$

By linearity of expectation, $\mathbb{E}\left[|N_1^-|\right] \geq (1 - \exp\left(-\alpha^2 C/200\right))(m-1)$. By Markov's inequality,

$$\mathbb{P}\left[m - 1 - |N_1^-| \geq \exp\left(-C\alpha^2/500\right)(m-1)\right] \leq \exp\left(-C\alpha^2/500\right) \leq 1/100,$$

implying the desired statement. □

Now we prove Lemma 4.1, by considering the distribution on $V_1, \ldots, V_m$ conditional on the events in the previous three lemmas and applying the random redistribution argument.

Proof of Lemma 4.1. Let E_1 be the event that the property in Lemma 4.3 holds, let E_2 be the event that the property in Lemma 4.4 holds, and let E_3 be the event that the property in Lemma 4.5 holds. We will condition on $E_1 \cap E_2 \cap E_3$, which holds with probability at least $97/100$.

For each permutation $\pi : [n] \rightarrow [n]$, define a set of "bad" random sets $\mathcal{F}_\pi \subseteq \{V_2, \ldots, V_m\}$ to include the V_i for which we have any of

(A1) $\delta_d(H[V_i]) < (\delta + 2\alpha/3)\binom{|V_i|}{k-d}$,
(A2) $\delta_d(H[V_i \cup \{v\}]) < (\delta + 2\alpha/3)\binom{|V_i \cup \{v\}|}{k-d}$ for at least $\exp(-\alpha^2 C/500)n$ vertices $v \in V$,
      and
(A3) $|N_i^-| < (1 - 1/\sqrt{C})m$.

Let $m' := |V_2 \cup \cdots \cup V_m|/C$, and note that m' is a positive integer by the choice of r. If necessary, add extra elements of $\{V_2, \ldots, V_m\}$ arbitrarily to $\mathcal{F}_\pi$ to ensure $\mathcal{F}_\pi$ has size at least $m - 1 - m' = m - 1 - m(C-1)/C = m/C - 1$. We claim that if $\pi \in E_1 \cap E_2 \cap E_3$, then $|\mathcal{F}_\pi| = m - 1 - m'$. Indeed, there are at most m/C^2 of type (A1) by Lemma 4.3 applied with $T = \emptyset$ assuming $\pi \in E_1$, there are at most $\exp\left(-\alpha^2 C/1000\right)m$ of type (A2) by Lemma 4.4 assuming $\pi \in E_2$, and there are at most $m/C^{4/3}$ of type (A3) since $\sum_{i=1}^m |N_i^-| = \sum_{i=1}^m |N_i^+| \geq (1 - 1/C^2)m^2$ by Lemma 4.3 assuming $\pi \in E_1$.

By possibly relabelling the sets $V_1, \ldots, V_m$, we assume without loss of generality that $\mathcal{F}_\pi = \{V_i : i \in [m] \setminus [m+1-|\mathcal{F}_\pi|]\}$. Now consider random sets V_i given by a random permutation π conditioned on $E_1 \cap E_2 \cap E_3$.

We define a bipartite graph G_π between $\bigcup_{V_i \in \mathcal{F}_\pi} V_i$ and $\{V_2, \ldots, V_m\} \setminus \mathcal{F}_\pi$. Note that the first part is a set of vertices and the second part is a set of subsets of vertices, and by the choice of m', the first and the second part both have size m', as $(m-1-m')(C-1) = m'$. We put an edge between v and V_i whenever $\delta_d(H[V_i \cup \{v\}]) \geq (\delta + 2\alpha/3)\binom{|V_i \cup \{v\}|}{k-d}$. Conditioning on E_1 and considering $T = \{v\}$, we have $d_G(v) \geq (1 - 1/C^2)m - |\mathcal{F}_\pi| \geq 3m'/4$ for every $v \in \bigcup_{V_i \in \mathcal{F}_\pi} V_i$. By the choice of $\mathcal{F}_\pi$ to contain all sets satisfying (A2), for every $V_i \in \{V_2, \ldots, V_m\} \setminus \mathcal{F}_\pi$, we have $d_G(V_i) \geq m' - \exp(-\alpha^2 C/500)n \geq 3m'/4$. Thus, $\delta(G) \geq 3m'/4$, and by Lemma 3.1, there is a $(C_{3.1}/m)$-spread distribution on perfect matchings of G_π.

We define the random sets U_i and vertices u_i as follows. First, sample π from the uniform distribution on permutations of $[n]$ conditional on $E_1 \cap E_2 \cap E_3$. Then, sample M_π from the $(C_{3.1}/m)$-spread distribution on perfect matchings of G_π. For each $V_i \notin \mathcal{F}_\pi$, let u_i be the vertex that it is assigned to in M_π, and let $U_i = V_i \cup \{u_i\}$. Letting $U_1 = V_1$, and picking u_1 to be an arbitrary vertex in $V_1 \setminus T_1$ concludes the algorithm that defines $\mathcal{U} = \{U_1, \ldots, U_{m'+1}\}$ with m in the lemma statement being $m' + 1$. We now check the required conditions.

(1) The condition on $|U_1|$ holds by construction, and for $i \geq 2$, we have $|U_i| = |V_i| + 1 = C$, as required.

(2) For $i = 1$, this condition holds because we condition on E_3, and for $i \geq 2$, the condition holds by the definition of the edges in the bipartite graph G_π.

(3) For every $i \in [m]$, let $N_i = \{U_j : V_j \in N_i^+ \cap N_i^-\}$. Note that $|N_i^+| \geq (1 - 1/C^2)m$ since we are conditioning on E_1. Note that $|N_1^-| \geq (1 - \exp(-\alpha^2 C/500))(m-1)$ since we are conditioning on E_3. Note that for $i \in [2, m'+1]$, we have $|N_i^-| \geq (1 - 1/\sqrt{C})m$ since these V_i are good sets and so satisfy the reverse of (3). Combining these gives $|N_i| \geq m' - m/C^2 - m/\sqrt{C} \geq (1 - \varepsilon)(m' + 1)$. For $i \geq 2$, note that $\delta_d(H[U_i \cup T_j \setminus \{u_i\}]) = \delta_d(H[V_i \cup T_j])$. Also note $\delta_d(H[U_1 \cup T_j \setminus \{u_1\}]) = \delta_d(H[V_1 \cup T_j] \setminus \{u_1\}) \geq \delta_d(H[V_1 \cup T_j]) - \binom{|V_1 \cup T_j| - 1}{k-d-1} \geq \delta_d(H[V_1 \cup T_j]) - \alpha \binom{|V_1 \cup T_j|}{k-d}/10$. Set $s_{i,j} := \alpha\binom{|V_1 \cup T_j|}{k-d}/10$. We thus have

$$\delta_d(H[U_i \cup T_j \setminus \{u_i\}]) \geq \delta_d(H[V_i \cup T_j]) - s_{i,j} \geq (\delta + 2\alpha/3)\binom{|V_i \cup T_j|}{k-d} - s_{i,j}$$

$$\geq (\delta + \alpha/2)\binom{|U_i \cup T_j|}{k-d}$$

for $i \geq 1$ and $U_j \in N_i$ because $V_j \in N_i^+$, and similarly

$$\delta_d(H[U_j \cup T_i \setminus \{u_j\}]) \geq \delta_d(H[V_j \cup T_i]) - s_{i,j} \geq (\delta + 2\alpha/3)\binom{|V_j \cup T_i|}{k-d} - s_{i,j}$$

$$\geq (\delta + \alpha/2)\binom{|U_j \cup T_i|}{k-d}$$

for $i \geq 1$ and $U_j \in N_i$ because $V_j \in N_i^-$.

(4) Let $y_1, \ldots, y_s \in V(H)$, and let $f : [s] \to [m'+1]$ as in the statement. For each $i \in [s]$, let D_i be the event that $y_i \in U_{f(i)}$, let D_i^1 be the event $y_i \in V_{f(i)}$, and let D_i^2 be the event $y_i = u_i$. Note that D_i^1 is determined only by the permutation π, and the event D_i^2 only holds if y_i is matched to V_i in G_π by M_π. Thus, for every $S \subseteq [s]$, we have

$$\mathbb{P}\left[\bigcap_{i \in S} D_i^1\right] \leq \left(\frac{C^{2|S|}(n-|S|)!}{n!}\right) \Big/ \mathbb{P}\left[E_1 \cap E_2 \cap E_3\right] \leq \frac{100}{97}\left(\frac{eC^2}{n}\right)^{|S|},$$

and for every $\pi' \in E_1 \cap E_2 \cap E_3$,

$$\mathbb{P}\left[\bigcap_{i \in [s] \setminus S} D_i^2 \;\middle|\; \pi = \pi'\right] \leq \left(\frac{C_{3.1}}{m'}\right)^{s-|S|}.$$

Therefore, for every $S \subseteq [s]$, we have

$$\mathbb{P}\left[\bigcap_{i \in S} D_i^1 \cap \bigcap_{i \in [s] \setminus S} D_i^2\right] = \mathbb{P}\left[\bigcap_{i \in S} D_i^1\right] \mathbb{P}\left[\bigcap_{i \in [s] \setminus [S]} D_i^2 \;\middle|\; \bigcap_{i \in S} D_i^1\right]$$

$$\leq \frac{100}{97}\left(\frac{eC^2}{n}\right)^{|S|}\left(\frac{C_{3.1}}{m'}\right)^{s-|S|} \leq \left(\frac{C'}{2n}\right)^s.$$

Finally, the result follows by a union bound over the 2^s choices of $S \subseteq [s]$, since

$$\bigcap_{i \in [s]} D_i = \bigcap_{i \in [s]} (D_i^1 \cup D_i^2) = \bigcup_{S \subseteq [s]} \left(\bigcap_{i \in S} D_i^1 \cap \bigcap_{i \in [s] \setminus S} D_i^2\right). \qquad \square$$

5. Proof of Theorems 1.14 and 1.15

In this section, we prove Theorems 1.14 and 1.15. Since the proof of Theorem 1.14 is the most challenging, we do not include all of the details for the proof of Theorem 1.15 and instead describe at the end of this section how to modify the proof of Theorem 1.14 to prove Theorem 1.15.

Recall that a k-uniform Hamilton ℓ-path with s edges occupies $(s-1)(k-\ell) + k = s(k-\ell) + \ell$ vertices. We call a set S (k,ℓ)-*path-divisible* if $|S|$ has the right divisibility condition to contain a k-uniform Hamilton ℓ-path, that is, $k-\ell$ divides $|S|-\ell$. We call a set S (k,ℓ)-*internal-path-divisible* if $|S|+2\ell$ has the right divisibility condition to contain a k-uniform Hamilton ℓ-path, so $k-\ell$ divides $|S|+\ell$. We call an integer q path-divisible (or internal-path-divisible) if a set of size q is path-divisible (or internal-path-divisible). Let $f(k,\ell) := \lceil (2k-\ell)/(k-\ell) \rceil$, and note that the final ℓ vertices of an ℓ path with $f(k,\ell)$ edges are disjoint from the first ℓ.

We now describe the random Hamilton ℓ-cycle embedding algorithm. In this algorithm, we first find a Hamilton cycle in H by gluing together many ℓ-paths. We assume each of these paths comes with a natural ordering of its vertices. For technical reasons, it is more convenient to find an embedding of $C'_{n,k,\ell}$ into H, where $C'_{n,k,\ell}$ is the hypergraph isomorphic to $C_{n,k,\ell}$ where edges are shifted to the left by ℓ; that is, $C'_{n,k,\ell}$ has vertex set $[n]$, $\{n-\ell+1,\ldots,n,1,\ldots,k-\ell\} \in E(C'_{n,k,\ell})$, and all other edges are "$(k-\ell)$-shifts" of this edge. See Fig. 1 for an illustration of the embedding algorithm.

Definition 5.1 *(Random Hamilton cycle embedding algorithm).* Let $\alpha > 0$, let $k \in \mathbb{N}$, and let $d, \ell \in [k-1]$. Let $1/n \ll 1/C' \ll 1/C \ll \alpha, 1/k$ as in Lemma 4.1, where $k-\ell$ divides n, and $C-(k-\ell)f(k,\ell)-\ell$ is (k,ℓ)-internal-path divisible, meaning that $k-\ell$ divides C. Given a hypergraph H on n vertices with $\delta_d(H) \geq (\delta^{\mathrm{CON}}_{k,\ell,d} + \alpha)\binom{n-d}{k-d}$, we define $\psi : C'_{n,k,\ell} \hookrightarrow H$, a random embedding of a Hamilton ℓ-cycle in H, as follows.

Step 1: Sample random clusters. Let $\mathcal{U} = \{U_1,\ldots,U_m\}$ be a random partition of $V(H)$ obtained by applying Lemma 4.1 with $10\ell k$ playing the role of t, $1/5$ playing the role of ε, and $\delta^{\mathrm{CON}}_{k,\ell,d}$ playing the role of δ (the choice of the other variables is as above). Define an auxiliary graph G with vertex set $[m]$ where i is adjacent to j if $U_j \in N_i$ or $U_i \in N_j$. Note that by the choice of ε, we have $\delta(G) \geq 4m/5$.

Let $\phi : C_{m,2,1} \hookrightarrow G$ be a random embedding of a Hamilton cycle on G coming from Lemma 3.2, and note that ϕ is a permutation of $[m]$. Throughout, we view $[m]$ cyclically, meaning occurrences of the indices $m+1$ and 0 are to be read as 1 and m, respectively. Let $z = \phi^{-1}(1)$ so that $U_1 = U_{\phi(z)}$, the largest random set.

Step 2: Find connecting paths between clusters. For each $i \in [m]$, find an ℓ-path $P_{\phi(i),\phi(i+1)}$ of length $f(k,\ell)$ contained in $U_{\phi(i)} \cup U_{\phi(i+1)}$ with the following properties.

1. The paths $P_{\phi(i),\phi(i+1)}$ are pairwise vertex-disjoint.
2. $P_{\phi(i),\phi(i+1)}$ has its first ℓ vertices in $U_{\phi(i)}$, and the last $|V(P_{\phi(i),\phi(i+1)})| - \ell \geq \ell$ vertices in $U_{\phi(i+1)}$ (that is, $P_{\phi(i),\phi(i+1)}$ does not alternate between clusters).

Step 3: Find Hamilton paths in the leftover part of each cluster. For each $i \in [m]$, call the ℓ-set that is an endpoint of $P_{\phi(i-1),\phi(i)}$ in $U_{\phi(i)}$ as S and call the ℓ-set that is an endpoint of $P_{\phi(i),\phi(i+1)}$ in $U_{\phi(i)}$ as T. Find a Hamilton ℓ-path in $H[U_{\phi(i)} \setminus (V(P_{\phi(i),\phi(i+1)}) \cup V(P_{\phi(i-1),\phi(i)})) \cup S \cup T]$ with endpoints S and T, and denote this Hamilton path as $P_{\phi(i)}$.

Defining ψ. Recall that $z = \phi^{-1}(1)$. To define $\psi : [n] \to V(H)$, we define $\psi(1)$ to be the first vertex of $P_{\phi(z-1),\phi(z)}$ that is contained in $U_{\phi(z)}$ and define the remaining $\psi(i)$ for $i > 1$ to be consistent with the ordering given by $P_{\phi(1)} \cup P_{\phi(1),\phi(2)} \cup P_{\phi(2)} \cup \cdots \cup P_{\phi(m)} \cup P_{\phi(m),\phi(1)}$.
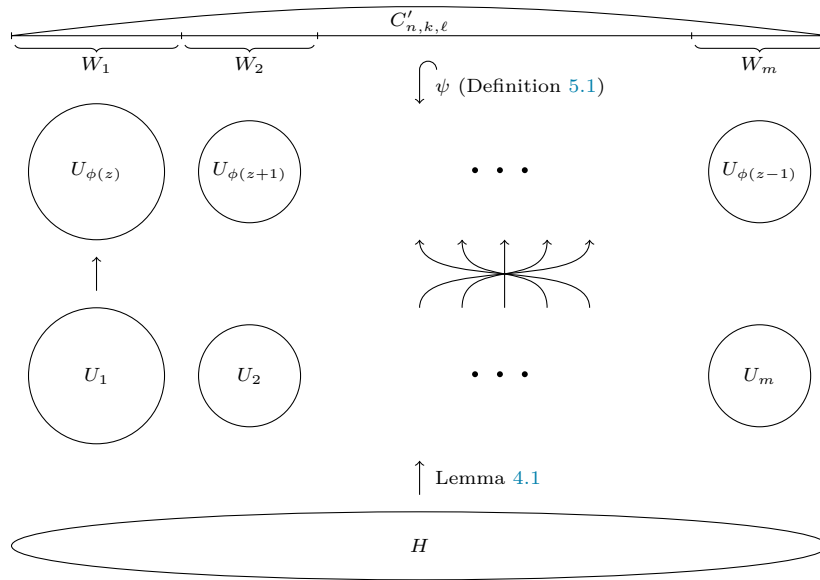
Fig. 1. An illustration of the random Hamilton cycle embedding algorithm.

For the next two lemmas, fix $\alpha > 0$, $k \in \mathbb{N}$, $d, \ell \in [k-1]$, and $1/n \ll 1/C' \ll 1/C \ll \alpha, 1$ as in Definition 5.1, and let H be an n-vertex hypergraph with $\delta_d(H) \geq (\delta_{k,\ell,d}^{\mathrm{CON}} + \alpha)\binom{n-d}{k-d}$. Theorem 1.14 follows immediately from the following two lemmas (note that producing embeddings of $C'_{n,k,\ell}$ is equivalent to producing those of $C_{n,k,\ell}$).

Lemma 5.2. *The random embedding $\psi : C'_{n,k,\ell} \hookrightarrow H$ in Definition 5.1 is well-defined.*

Lemma 5.3. *The distribution produced by the algorithm in Definition 5.1 is $(C'C_{3.2}/n)$-vertex-spread.*

Proof of Lemma 5.2. By construction, $P_{\phi(1)} \cup P_{\phi(1),\phi(2)} \cup P_{\phi(2)} \cup \cdots \cup P_{\phi(m)} \cup P_{\phi(m),\phi(1)}$ is a Hamilton ℓ-cycle of H. Furthermore, by Step 2, the first vertex of $P_{\phi(z-1),\phi(z)}$ that is contained in $U_{\phi(z)}$ is in fact the $(\ell+1)$th vertex of $P_{\phi(z-1),\phi(z)}$, so ψ is an embedding of $C'_{n,k,\ell}$ in H, as required. Therefore it suffices to show that each step in Definition 5.1 can be performed.

Step 1. This step can be performed because of Lemmas 4.1 and 3.2.

Step 2. Let $i \in [m]$ and suppose paths satisfying the two properties have been found for each $i' < i$, and let J denote the set of vertices used by these previously found paths. Consider $U_{\phi(i)}$ and $U_{\phi(i+1)}$. Note that from each set, J uses at most $\ell + (k-\ell)f(k,\ell) \leq 3k$ vertices. Consider the $T_{\phi(i)} \subseteq U_{\phi(i)}$ guaranteed by Lemma 4.1(3). Note that $t - 3k \geq \ell$, so we may fix an ℓ-subset $T' \subseteq T_{\phi(i)}$ disjoint from J. Let $U \subseteq U_{\phi(i+1)} \setminus (J \cup \{u_{\phi(i+1)}\})$ be obtained by deleting at most k elements from the latter superset so that $U \cup T'$ is (k, ℓ)-path divisible. Observe that $\delta_d(H[U \cup T']) \geq (\delta_{k,\ell,d}^{\mathrm{CON}} + \alpha/10)\binom{|U \cup T'|}{k-d}$ (by Lemma 4.1(3)); hence, by definition of $\delta_{k,\ell,d}^{\mathrm{CON}}$, $H[U \cup T']$ contains a Hamilton ℓ-path with one endpoint

being the ℓ-set T', and the other endpoint being some ℓ-subset of U that can be chosen arbitrarily. We truncate this path (keeping the side closer to T' intact) to keep only its first $f(k,\ell)$ edges. The resulting path satisfies the desired properties, showing that Step 2 can be executed for each $i \in [m]$.

Step 3. Similarly to the previous step, we need to check the relevant divisibility and minimum degree conditions to show that a path of the desired form exists for each $i \in [m]$.

We first show that $U_{\phi(i)} \setminus (V(P_{\phi(i),\phi(i+1)}) \cup V(P_{\phi(i-1),\phi(i)}))$ is (k,ℓ)-internal-path-divisible for every $i \in [n]$. Indeed, for $i \neq \phi^{-1}(1)$, we have $|U_{\phi(i)} \setminus (V(P_{\phi(i),\phi(i+1)}) \cup V(P_{\phi(i-1),\phi(i)}))| = C - \ell - f(k,\ell)(k-\ell)$, so the internal path divisibility follows from our choice of C. For $i = \phi^{-1}(1)$, we have $|U_{\phi(i)} \setminus (V(P_{\phi(i),\phi(i+1)}) \cup V(P_{\phi(i-1),\phi(i)}))| = n - \ell - f(k,\ell)(k-\ell) - (m-1)C$, so the internal path divisibility follows since $k - \ell$ divides both n and C.

Now we claim that $\delta_d(H[U_{\phi(i)} \setminus (V(P_{\phi(i),\phi(i+1)}) \cup V(P_{\phi(i-1),\phi(i)}))]) \geq (\delta_{k,\ell,d} + \alpha/10)\binom{|U_{\phi(i)}|}{k-d}$. This is a consequence of Lemma 4.1(2) since $|V(P_{\phi(i),\phi(i+1)}) \cup V(P_{\phi(i-1),\phi(i)})| \leq 10k$, and $1/C \ll 1/k$. Therefore, the desired Hamilton path with the desired endpoints exists by the definition of $\delta_{k,\ell,d}^{\mathrm{CON}}$, meaning that Step 3 can be executed. $\square$

Proof of Lemma 5.3. Let $\psi : C'_{n,k,\ell} \hookrightarrow H$ be the random embedding from Definition 5.1. To show the distribution is $(C'C_{3.2}/n)$-vertex-spread, we show that for every $s \in [n]$ and every two sequences of distinct $x_1, \ldots, x_s \subseteq [n]$ and $y_1, \ldots, y_s \in V(H)$, we have $\mathbb{P}[\psi(x_i) = y_i \text{ for all } i \in [s]] \leq (C'C_{3.2}/n)^s$.

To that end, let r be $(C-1)C$ plus the remainder when n is divided by $(C-1)C$. Let $W_1 := [r]$, and for $i \geq 2$, let $W_i := [r+C(i-2)+1, r+C(i-1)] = [r+C(i-1)] \setminus \bigcup_{j=1}^{i-1} W_j$. (Note that these are slightly different from those defined in the previous section.) We refer to each W_i as a "window". Let $w : [s] \to [m]$ where $w(i)$ is the unique index such that $x_i \in W_{w(i)}$. We refer to $W_{w(i)}$ as "x_i's window". Recall $\mathcal{U} = \{U_1, \ldots, U_m\}$ is the collection of random sets and ϕ is the random embedding of $C_{m,2,1}$ in the auxiliary graph G as defined in Step 1, and $z = \phi^{-1}(1)$. See Fig. 1. We will frequently use the following.

Observation 5.4. The embedding ψ induces bijections between and W_{i+1} and $U_{\phi(z+i)}$ for each $i \geq 0$ (where the subscripts of W and the inputs to ϕ are to be interpreted cyclically). In particular, for every $i \in [s]$, we have $\psi(x_i) = y_i$ only if $y_i \in U_{\phi(z+w(i)-1)}$.

For each function $f : [s] \to [m]$, let E_f be the event that $y_i \in U_{f(i)}$ for each $i \in [s]$. Let $\mathcal{F}$ be the family of functions $f : [s] \to [m]$ satisfying the following:

- If $w(i) = 1$ (that is, x_i is in the biggest window), then $f(i) = 1$.
- If $w(i) = w(j)$ (that is, x_i and x_j are in the same window), then $f(i) = f(j)$.
- If $w(i) \neq w(j)$ (that is, x_i and x_j are in different windows), then $f(i) \neq f(j)$.

Let b denote the number of $i \in [m]$ for which $\{x_1, \ldots, x_s\} \cap W_i \neq \emptyset$; that is, b is the number of windows that are the window of some x_i for $i \in [s]$. Define $\hat{b}$ as b if $W_1 \cap \{x_1, \ldots, x_s\} = \emptyset$ and define $\hat{b}$ as $b - 1$ otherwise.

We claim the following:

(a) $\mathbb{P}[E_f] \leq (C'/n)^s$ for every $f : [s] \to [m]$;

(b) if $f : [s] \to [m]$ is not in $\mathcal{F}$, then $\mathbb{P}[\psi(x_i) = y_i \text{ for all } i \in [s] \mid E_f] = 0$;

(c) $|\mathcal{F}| = \prod_{i=0}^{\hat{b}-1}(m - i) \leq m^{\hat{b}}$;

(d) $\hat{b} \leq b \leq s$;

(e) $\mathbb{P}[\psi(x_i) = y_i \text{ for all } i \in [s] \mid E_f] \leq (C_{3.2}/m)^{\hat{b}}$ for every $f \in \mathcal{F}$.

Indeed, (a) follows immediately from Lemma 4.1(4), (b) follows from Observation 5.4, and (c) and (d) are straightforward from the definitions. To prove (e), first note that by Observation 5.4,

$$\mathbb{P}[\psi(x_i) = y_i \text{ for all } i \in [s] \mid E_f] \leq \mathbb{P}[\phi(z + w(i) - 1) = f(i) \text{ for all } i \in [s]].$$

Let $x_1', \ldots, x_{\hat{b}}'$ and $y_1', \ldots, y_{\hat{b}}'$ be the elements of $\{z + w(i) - 1 : i \in [s]\} \setminus \{z\}$ and $\{f(i) : i \in [s]\} \setminus \{1\}$, respectively. By the vertex spreadness of ϕ with $x_1', \ldots, x_{\hat{b}}'$ and $y_1', \ldots, y_{\hat{b}}'$ playing the roles of $\{x_1, \ldots, x_s\}$ and $\{y_1, \ldots, y_s\}$, respectively, we have

$$\mathbb{P}[\phi(z + w(i) - 1) = f(i) \text{ for all } i \in [s]] \leq (C_{3.2}/m)^{\hat{b}}.$$

The two inequalities above together imply (e).

Combining (a)–(e), we have

$$
\begin{aligned}
\mathbb{P}[\psi(x_i) = y_i \ \forall i \in [s]] &\overset{(b)}{=} \sum_{f \in \mathcal{F}} \mathbb{P}[\psi(x_i) = y_i \ \forall i \in [s] \mid E_f]\, \mathbb{P}[E_f] \\
&\overset{(a)}{\leq} \sum_{f \in \mathcal{F}} \mathbb{P}[\psi(x_i) = y_i \ \forall i \in [s] \mid E_f]\left(\frac{C'}{n}\right)^s \\
&\overset{(e)}{\leq} \left(\frac{C'}{n}\right)^s \sum_{f \in \mathcal{F}} \left(\frac{C_{3.2}}{m}\right)^{\hat{b}} = \left(\frac{C'}{n}\right)^s |\mathcal{F}|\left(\frac{C_{3.2}}{m}\right)^{\hat{b}} \\
&\overset{(c)}{\leq} \left(\frac{C'}{n}\right)^s m^{\hat{b}}\left(\frac{C_{3.2}}{m}\right)^{\hat{b}} = \left(\frac{C'}{n}\right)^s (C_{3.2})^{\hat{b}} \\
&\overset{(d)}{\leq} \left(\frac{C'}{n}\right)^s (C_{3.2})^s = \left(\frac{C'C_{3.2}}{n}\right)^s,
\end{aligned}
$$

as desired. $\quad\square$

We conclude this section by explaining how to modify the proof of Theorem 1.14 to prove Theorem 1.15. First we describe the random embedding algorithm.

Definition 5.5 *(Random F-factor embedding algorithm).* Let $\alpha > 0$, let $k, r \in \mathbb{N}$, and let $d \in [k-1]$. Let $1/n \ll 1/C' \ll 1/C \ll \alpha, 1$ as in Lemma 4.1, where r divides n and C. Given a hypergraph H on n vertices with $\delta_d(H) \geq (\delta_{F,d} + \alpha)\binom{n-d}{k-d}$, we define $\psi : G \hookrightarrow H$, a random embedding of an F-factor in H, as follows.

    Step 1: Sample random clusters. Let $\mathcal{U} = \{U_1, \ldots, U_m\}$ be a random partition of $V(H)$ obtained by applying Lemma 4.1 with $\delta_{F,d}$ playing the role of δ (the choice of the other variables is as above – ε and t are unimportant and can be set to 1).

    Step 2: Find an F-factor in each cluster. For each $i \in [m]$, find an F-factor in $H[U_i]$. Label the vertices of U_i as $u_{i,1}, \ldots, u_{i,|U_i|}$ such that $H[\{u_{i,(j-1)r+1}, \ldots, u_{i,jr}\}] \cong F$ for every $j \in [|U_i|/r]$.

    Defining ψ. Assume without loss of generality that $V(G) = [n]$ and $G[\{(j-1)r + 1, \ldots, jr\}] \cong F$ for every $j \in [n/r]$. For every $i \in [n]$, let $\psi(i) := u_{x,y}$, where x is the largest integer such that $i > \sum_{j=1}^{x-1} |U_i|$ and $y = i - \sum_{j=1}^{x-1} |U_i|$.

To prove Theorem 1.15, we need analogues of Lemmas 5.2 and 5.3. The analogue of Lemma 5.2 is straightforward, so we omit it. For the analogue of Lemma 5.3, we note the following changes:

- since there is no ϕ in Definition 5.5, it can be interpreted as the identity, so $z = 1$;
- the family of functions $\mathcal{F}$ consists of the single function w;
- we do not need to define $\hat{b}$ or b, and we do not need (d);
- (a) and (b) are unchanged, but for (c), we have $|\mathcal{F}| = 1$, and for (e), we have simply that this probability is at most 1.

Altogether, we obtain a (C'/n)-vertex-spread distribution, as desired.

6. Spreadness from vertex spreadness

In this section we prove Proposition 1.17, which, combined with Theorem 1.14 and the FKNP theorem, implies Theorem 1.8. We also prove Theorem 1.9 using Theorem 1.14 and Spiro's strengthening of the FKNP theorem, and we prove Theorem 1.11 by combining Theorem 1.15 with the FKNP theorem and a coupling argument of Riordan [31].

6.1. Proof of Proposition 1.17

We begin with the proof of Proposition 1.17. First we need a lemma, upper bounding the number of "partial embeddings" of G into H.

Lemma 6.1. *Let H and G be n-vertex k-uniform hypergraphs. If $F \subseteq H$ has v vertices and c components, then there are at most*

$$n^c(k\,\Delta(G))^{v-c}$$

hypergraph embeddings $\varphi : G[X] \hookrightarrow H[V(F)]$ where $X \in \binom{V(G)}{v}$ and $F \subseteq H[\varphi(X)]$.

Proof. Let $F_1, \ldots, F_c$ be the components of F, and for each $i \in [c]$, let $V(F_i) = \{v_{i,1}, \ldots, v_{i,|V(F_i)|}\}$, where for every $j \in [|V(F_i)|] \setminus \{1\}$, there exists $j' < j$ such that $v_{i,j}$ and $v_{i,j'}$ are contained in a common edge of F_i. Each embedding $\varphi : G[X] \hookrightarrow H[V(F)]$ with $F \subseteq H[\varphi(X)]$ is then determined by

1. the preimages $\varphi^{-1}(v_{1,1}), \ldots, \varphi^{-1}(v_{c,1})$ of the "roots" and
2. for each $i \in [c]$, a sequence in $[k\,\Delta(G)]^{|V(F_i)|-1}$, where the jth term in the sequence determines $\varphi^{-1}(v_{i,j+1})$ based on $\varphi^{-1}(v_{i,j'})$, where $j' \leq j$ and $v_{i,j'}$ and $v_{i,j+1}$ are contained in a common edge of F_i.

Note that there are at most n^c choices for the preimages of the roots and at most $(k\,\Delta(G))^{v-c}$ choices for the sequences. Combining these choices yields the desired bound. □

Proof of Proposition 1.17. Let $1/n \ll 1/C' \ll 1/C, 1/k, 1/\Delta \leq 1$. Let H and G be n-vertex k-uniform hypergraphs, where $\Delta(G) \leq \Delta$, and suppose there is a (C/n)-vertex-spread distribution μ on embeddings $G \hookrightarrow H$. For every $F \subseteq H$ isomorphic to G, let

$$\mu'(F) := \mu\left(\{\varphi : \{\varphi(e) : e \in E(G)\} = E(F)\}\right),$$

and note that μ' is a probability distribution on subgraphs of H which are isomorphic to G.

We prove that μ' is $\left(C'/n^{1/m_1(G)}\right)$-spread. To that end, let $S \subseteq E(H)$, and let $T \subseteq H$ have edge set S and subject to that, the fewest number of vertices. We may assume T is isomorphic to a subgraph of G, or else $\mu'(\{F \subseteq H : E(F) \supseteq S\}) = 0$. We may also assume $S \neq \emptyset$. Let v and c be the number of vertices and components of T, respectively. By Lemma 6.1, the number of embeddings $\varphi : G[X] \hookrightarrow H[V(T)]$ where $X \in \binom{V(G)}{v}$ and $T \subseteq H[\varphi(X)]$ is at most $n^c(k\Delta^{v-c})$, so since μ is (C/n)-vertex-spread,

$$\mu'(\{F \subseteq H : E(F) \supseteq S\}) \leq n^c(k\Delta)^{v-c}\left(\frac{C}{n}\right)^v \leq \left(\frac{(C')^{m_1(G)}}{n}\right)^{v-c}$$

$$= \left(\frac{(C')^{m_1(G)}}{n}\right)^{|S|(v-c)/|S|}.$$

Let $T_1, \ldots, T_c$ be the components of T. Since each T_i is isomorphic to a subgraph of G, for every $i \in [c]$ we have $|E(T_i)|/(|V(T_i)| - 1) \leq m_1(G)$. In particular,

$$|S| = \sum_{i=1}^{c} |E(T_i)| \le m_1(G) \sum_{i=1}^{c} (|V(T_i)| - 1) = m_1(G)(v - c),$$

so $(v - c)/|S| \ge 1/m_1(G)$. Therefore

$$\mu'(\{F \subseteq H : E(F) \supseteq S\}) \le \left(\frac{(C')^{m_1(G)}}{n} \right)^{|S|(v-c)/|S|} \le \left(\frac{C'}{n^{1/m_1(G)}} \right)^{|S|},$$

as desired. $\square$

6.2. Stronger spreadness: proof of Theorem 1.9

Next we prove Theorem 1.9. As mentioned, we need a result of Spiro [34], which requires the following stronger notion of spreadness.

Definition 6.2. Let $q \in [0, 1]$, and let $r_0, \ldots, r_\ell \in \mathbb{N}$ be a decreasing sequence of positive integers. Let $(V, \mathcal{H})$ be an r_0-bounded hypergraph, and let μ be a probability distribution on $\mathcal{H}$. We say μ is $(q; r_0, \ldots, r_\ell)$-*spread* if the following holds for all $i \in [\ell]$:

$$\mu\left(\{A \in \mathcal{H} : |A \cap S| = t\}\right) \le q^t \text{ for all } t \in [r_i, r_{i-1}] \text{ and } S \in \bigcup_{i=r_i}^{r_{i-1}} \left\{ \binom{E}{i} : E \in \mathcal{H} \right\}.$$

As noted by Spiro, if μ is a q-spread measure on an r_0-bounded hypergraph $(V, \mathcal{H})$, then it is also $(4q; r_0, \ldots, r_\ell)$ spread where $r_i := \lceil r_{i-1}/2 \rceil$ for $i \in [\ell]$, so the following result with $\ell = \Theta(\log r)$ implies the FKNP theorem for uniform hypergraphs. (Spiro [34, Theorem 3.1] also proved a slightly stronger result in the non-uniform setting which implies the FKNP theorem, but we do not need this result here).

Theorem 6.3 *(Spiro [34]). There exists a constant $K_{6.3} > 0$ such that the following holds for all $K > K_{6.3}$. Let $q \in [0, 1]$, let $r_0, \ldots, r_\ell \in \mathbb{N}$ be a decreasing sequence of positive integers, where $r_\ell = 1$, and let $(V, \mathcal{H})$ be an r_0-uniform hypergraph. If there exists a $(q; r_0, \ldots, r_\ell)$-spread probability distribution on $\mathcal{H}$ and $p \ge K\ell q$, then a p-random subset of V contains an edge of $\mathcal{H}$ with probability at least $1 - K_{6.3}/(K\ell)$.*

We apply Theorem 6.3 with 1 playing the role of ℓ in Definition 6.2. This special case of Theorem 6.3 is in some sense the "base case" of Spiro's proof [34, Lemma 2.5] and can be proved directly with the second moment method. To apply Theorem 6.3, we also need the following analogue of Proposition 1.17 with this notion of spreadness.

Proposition 6.4. *For every $C > 0$ and $k \in \mathbb{N}$, there exists $C' = C_{6.4}(k, C) > 0$ such that the following holds for all $\ell \in \{2, \ldots, k - 1\}$ and all sufficiently large n. Let H be an n-vertex k-uniform hypergraph. If there is a (C/n)-vertex-spread distribution on embeddings $C_{n,k,\ell} \hookrightarrow H$, then there is a $(C'/n^{k-\ell}; n/(k - \ell), 1)$-spread distribution on Hamilton ℓ-cycles of H.*

Theorem 6.3 and Proposition 6.4 together imply that $\delta^{\mathrm{RD}}_{\mathcal{C}_{k,\ell},d} \leq \delta^{\mathrm{VSD}}_{\mathcal{C}_{k,\ell},d}$ for every $\ell \in \{2, \ldots, k-1\}$ and $d \in [k-1]$. Theorem 1.9 follows immediately from Theorems 6.3 and 1.14 and Proposition 6.4, so the remainder of this subsection is devoted to the proof of Proposition 6.4. This proof is inspired by the proof of Kahn, Narayanan, and Park [18] determining the threshold for $\mathcal{G}(n,p)$ to contain the square of a Hamilton cycle.

Lemma 6.5. *Every connected subgraph of $C_{n,k,\ell}$ with v vertices and t edges satisfies*

$$v \geq \min\{(k-\ell)t + \ell, n\}.$$

In particular, if $\ell > 0$, then

$$m_1(C_{n,k,\ell}) = \frac{1}{k-\ell} \cdot \frac{n}{n-1}.$$

Proof. Let $F \subseteq C_{n,k,\ell}$ be a connected subgraph with v vertices and t edges. We may assume without loss of generality that F has vertex set $[v]$ where $v < n$. For some $I \subseteq [n/(k-\ell)]$ of size t, the edge set of F is

$$\{f_i := [k + (k-\ell)(i-1)] \setminus [(k-\ell)(i-1)] : i \in I\}.$$

Therefore,

$$v \geq \sum_{i=1}^{n/(k-\ell)} \mathbb{1}_{i \in I} \left| f_i \setminus \bigcup_{j=1}^{i-1} f_j \right| \geq k + (k-\ell)(t-1) = (k-\ell)t + \ell,$$

as desired. To see that $m_1(C_{n,k,\ell}) = n/((k-\ell)(n-1))$ for $\ell > 0$, first note $d_1(C_{n,k,\ell}) = n/((k-\ell)(n-1))$. Moreover, every spanning subgraph $F \subseteq H$ satisfies $d_1(F) \leq d_1(H)$. If $F \subseteq H$ has fewer than n vertices, then $|V(F)| \geq (k-\ell)|E(F)| + \ell$, so

$$d_1(F) = \frac{|E(F)|}{|V(F)|-1} \leq \frac{|V(F)|-\ell}{(k-\ell)(|V(F)|-1)} \leq \frac{1}{k-\ell},$$

as desired. $\quad\square$

Lemma 6.6. *If S is a subset of edges of $C_{n,k,\ell}$, then the number of subgraphs of $C_{n,k,\ell}$ with t edges, all in S, no isolated vertices, and c components, is at most*

$$\binom{k|S|}{c} \left(2 \cdot 16^k\right)^t.$$

Proof. Let $f_i := [k + (k-\ell)(i-1)] \setminus [(k-\ell)(i-1)]$ for each $i \in [n/(k-\ell)]$ as in Lemma 6.5. Each subgraph $F \subseteq H$ with t edges, all in S, no isolated vertices, and c components $F_1, \ldots, F_c$, is determined by

1. c distinct vertices $\{v_1, \ldots, v_c\} \in \binom{\bigcup_{e \in S} e}{c}$ to serve as the "roots" of $F_1, \ldots, F_c$,
2. a c-composition $t_1, \ldots, t_c$ of t (that is, $t_1 + \cdots + t_c = t$), where $|E(F_i)| = t_i$ for all $i \in [c]$, and
3. for each $i \in [c]$, the edges $E(F_i) \subseteq S \cap \{f_j : |k + (k - \ell)(j - 1) - v_i| \leq 2kt_i\}$.

Note that there are at most $\binom{k|S|}{c}$ choices for the roots, at most $\binom{t-1}{c-1} \leq 2^t$ choices for the c-composition, and at most $\prod_{i=1}^{c} 2^{4kt_i} = \left(16^k\right)^t$ choices for the edges of each F_i. Combining these choices yields the desired bound. $\quad\square$

Proof of Proposition 6.4. Let $1/n \ll 1/C' \ll 1/C, 1/k \leq 1$. Let H be an n-vertex k-uniform hypergraph, and suppose there is a (C/n)-vertex-spread distribution μ on embeddings $C_{n,k,\ell} \hookrightarrow H$. Let $\mathcal{H}$ be the set of edge sets of Hamilton ℓ-cycles of H. Define a probability distribution μ' on $\mathcal{H}$ as in the proof of Proposition 1.17, with $C_{n,k,\ell}$ playing the role of G. By Proposition 1.17 and Lemma 6.5, μ' is $\left(2^{-k}C'/n^{k-\ell}\right)$-spread.

We prove that μ' is $(C'/n^{k-\ell}; n/(k - \ell), 1)$-spread. To that end, let $t \in [n/(k - \ell)]$, and let $S \subseteq E(H)$ be a subset of a Hamilton ℓ-cycle of H. Let $F \subseteq H$ have edge set S and subject to that, the fewest number of vertices.

If $t \geq n/k$, then since μ' is $(2^{-k}C'/n^{k-\ell})$-spread, we have

$$\mu'\left(\{A \in \mathcal{H} : |A \cap S| = t\}\right) \leq 2^{|S|}\left(\frac{2^{-k}C'}{n^{k-\ell}}\right)^t \leq \left(\frac{C'}{n^{k-\ell}}\right)^t,$$

as desired. Therefore, we may assume $t < n/k$, and in particular, every subgraph of F has fewer than n vertices. By Lemmas 6.1, 6.5, and 6.6, since μ is (C/n)-vertex-spread,

$$\mu'\left(\{A \in \mathcal{H} : |A \cap S| = t\}\right) \leq \sum_{c=1}^{t} \binom{k|S|}{c}(2 \cdot 16^k)^t n^c \left(\frac{2k^2 C}{n}\right)^{(k-\ell)t+\ell c}$$

$$\leq (2 \cdot 16^k)^t \sum_{c=1}^{t} \left(\frac{ekn \cdot 2k^2 C}{c}\right)^c \left(\frac{2k^2 C}{n}\right)^{(k-\ell)t+(\ell-1)c}$$

$$\leq \left(\frac{2 \cdot 16^k (2k^2 C)^{k-\ell}}{n^{k-\ell}}\right)^t \sum_{c=1}^{t} \left(\frac{4ek^4 C^2}{c}\right)^c \left(\frac{2k^2 C}{n}\right)^{(\ell-2)c}$$

$$\leq \left(\frac{C'/2}{n^{k-\ell}}\right)^t \sum_{c=1}^{\infty} \left(\frac{1}{c}\right)^c \leq \left(\frac{C'}{n^{k-\ell}}\right)^t,$$

as desired. $\quad\square$

6.3. Proof of Theorem 1.11

We conclude this section with the proof of Theorem 1.11. With Theorem 1.15 in hand, the proof of Theorem 1.11 is similar to the PSSS proof of Theorem 1.5(2) assuming a

similar result for K_r-factors [28, Theorem 1.8]. For hypergraphs F and H, the F-*complex* of H, denoted H_F, is the $|V(F)|$-uniform multi-hypergraph with vertex set $V(H)$ in which every copy of F in H corresponds to a distinct hyperedge of H on the same set of vertices. Note that H has an F-factor if and only if H_F contains a perfect matching. We let $\mathcal{G}_F(n, p)$ be the binomial random multi-hypergraph on n vertices where every edge of the F-complex of the complete hypergraph is included independently with probability p.

To prove Theorem 1.11, we need following result of Riordan [31, Theorem 18].

Theorem 6.7. *For every $k, r \in \mathbb{N}$, there exists $a = a_{6.7}(k, r) \in (0, 1]$ such that the following holds. If F is a strictly 1-balanced k-uniform r-vertex hypergraph and $p = p(n) \leq \log^2(n)/n^{1/d_1(F)}$, then for some $\pi = \pi(n) \sim ap^{|E(F)|}$, we may couple $G \sim \mathcal{G}^{(k)}(n, p)$ with $G_F \sim \mathcal{G}_F(n, \pi)$ such that, a.a.s. for every F-edge present in H_F, the corresponding copy of F is present in G.*

Proof of Theorem 1.11. Let $1/n \ll 1/C \ll 1/C' \ll 1/C'' \ll 1/k, 1/r, a, \alpha < 1$, where $a = a_{6.7}(k, r)$. It suffices to prove the result for $p = Cn^{-1/d_1(F)} \log^{1/|E(F)|} n$. By Theorem 6.7, H_p contains an F-factor if $(H_F)_\pi$ contains a perfect matching where $\pi = ap^{|E(F)|} = aC^{|E(F)|}n^{-(|V(F)|-1)} \log n$.

By Theorem 1.15, there exists a (C''/n)-vertex-spread distribution on embeddings of $G \hookrightarrow H$ where G is n/r disjoint copies of F. Since an F-factor in H corresponds to a perfect matching in H_F, an embedding of G into H is also an embedding of G' into H' where G' and H' are the "simplifications" of G_F and H_F, respectively (G' is just an r-uniform perfect matching because a strictly 1-balanced hypergraph is connected). Hence, Proposition 1.17 implies that H' supports a (C'/n^{r-1})-spread distribution on its perfect matchings. Therefore, the Frankston–Kahn–Narayanan–Park theorem implies that H'_π contains a perfect matching a.a.s. if $\pi > KC' \log(n/r)/n^{r-1}$. Indeed, this inequality holds by the choice of C. Since H'_π a.a.s. contains a perfect matching, so does $(H_F)_\pi$. Thus, as mentioned, by Theorem 6.7, H_p a.a.s. contains an F-factor, as desired. □

7. Proof of Proposition 1.7

The following two properties were introduced in [13] to codify in abstract what it means for there to be a standard absorption proof for hypergraphs which are obtained from so-called $\mathcal{A}$-chains. In this paper, we are concerned only with hypergraph Hamilton cycles, so the $\mathcal{A}$-chains are obtained by setting $\mathcal{A}$ to be a single edge. Therefore, we state the following properties in the language of hypergraph Hamilton cycles. Note also that in contrast to [13], we use $\binom{n-d}{k-d}$ terms instead of n^{k-d} terms, as this was more convenient in the current paper.

Ab For any $\alpha > 0$, there exist $0 < \tau, \eta \leq \alpha$ and $n_0 \in \mathbb{N}$ so that if H is a k-uniform hypergraph on $n \geq n_0$ vertices with $\delta_d(H) \geq (\delta + \alpha)\binom{n-d}{k-d}$, then there exists $A \subseteq V(H)$ of size at most τn with the following property.

For any $L \subseteq V(H) \setminus A$ of size at most ηn such that $|L| + |A|$ is (k, ℓ)-path-divisible, there exists an embedding of a k-uniform ℓ-path to H with vertex set $A \cup L$. Furthermore, the embedding of the sets of ℓ first and ℓ last endpoints of the path does not depend on the subset L, meaning that there exist disjoint ℓ-sets $A_1, A_2 \subseteq A$ such that for all L, the ℓ-path covering $A \cup L$ has endpoints A_1 and A_2.

Con For any $\alpha > 0$, there exist a positive integer c and $n_0 \in \mathbb{N}$ so that if H is a k-uniform hypergraph H on $n \geq n_0$ vertices with $\delta_d(H) \geq (\delta + \alpha)\binom{n-d}{k-d}$, then the following holds.

For every $S, T \subseteq V(H)$ of vertex-disjoint ℓ-sets, H contains an embedding of a k-uniform ℓ-path of length at most c with start on S and end on T.

As noted in [13] (see Section 6, Table 1), both of these properties are known to hold with $\delta_{\mathcal{C}_{\ell,k,d}}$ playing the role of δ for the parameters ℓ, k, d as listed in Proposition 1.7(2-3). The following therefore immediately implies Proposition 1.7(2-3).

Proposition 7.1. *Let $\delta \in [0, 1]$, let $k \in \mathbb{N}$, and let $\ell, d \in [k - 1]$. If $\delta \geq \delta_{\mathcal{C}_{\ell,k,d}}$ and the properties **Ab** and **Con** hold for δ, k, ℓ, and d, then $\delta_{k,\ell,d}^{\mathrm{CON}} \leq \delta$.*

Proof. Let $\alpha > 0$, let n be sufficiently large and (k, ℓ)-path-divisible, and consider a k-uniform hypergraph H on n vertices with $\delta_d(H) \geq (\delta + \alpha)\binom{n-d}{k-d}$. Let S and T be disjoint vertex subsets of size ℓ each. We wish to find a Hamilton ℓ-path of H with end sets S and T.

Let $\eta, \tau, \alpha' \in (0, 1)$ and $c \geq 1$ satisfy $1/n \ll 1/c \ll \eta, \tau \leq \alpha' \ll \alpha$, where η, τ and c satisfy **Ab** and **Con**, respectively, with α' playing the role of α. Let $R \subseteq V(H)$ be an $(\eta n/2)$-subset so that for any 2ℓ-subset $Q \subseteq V(H)$ we have that $\delta_d(H[R \cup Q]) \geq (\delta + \alpha/2)\binom{|R \cup Q|-d}{k-d}$ (such a subset exists by a union bound over applications of Lemma 2.2). Since $\alpha' \ll \alpha$, we have $\delta_d(H \setminus R) \geq (\delta + \alpha/2)\binom{n-d}{k-d}$, so by **Ab** (with α' playing the role of α), there exists an absorbing set $A \subseteq V(H) \setminus R$ of size at most τn, and $A_1, A_2 \subseteq A$ so that for any small L, $A \cup L$ has a Hamilton path with A_1 and A_2 as the first and last ℓ vertices of the path. Similarly, $\delta_d(H \setminus (R \cup A)) \geq (\delta + \alpha/2)\binom{n-d}{k-d}$, so $H \setminus (R \cup A)$ therefore contains two vertex-disjoint ℓ-paths P_0 and P_1 together covering all but at most $3k$ vertices outside $R \cup A$. (To find such paths, first delete a minimal number of vertices from $H \setminus (R \cup A)$ to obtain a hypergraph where the number of vertices is divisible by $k - \ell$. Then, since $\delta \geq \delta_{\mathcal{C}_{\ell,k,d}}$, the resulting hypergraph contains a Hamilton ℓ-cycle, and we again delete a minimum number of vertices from the cycle to obtain the two paths. Note that in the first step, we delete at most k vertices, and in the second step, we delete at most $2k$ vertices.)

Now, using **Con** four times (each time deleting c vertices from the host hypergraph, which does not significantly change the minimum degree condition), we can extend P_0 to P_0' so that it has as endpoints S and A_1, and extend P_1 to P_1' so that it has endpoints A_2 and T. While doing this, we only use vertices coming from the set R (those which have not been previously used). The remaining vertices in R together with the vertices

deleted for divisibility reasons (while finding P_0, P_1) can be absorbed into A, as the number of remaining vertices is at most $\eta n/2 + 3k \leq \eta n$ (since $1/n \ll 1/k$) and have the appropriate divisibility property (otherwise n itself would not be (k, ℓ)-path-divisible). This allows us to find the desired spanning path with endpoints S and T. $\quad \square$

8. Open problems

8.1. Beyond Hamilton connectivity

We conjecture that Theorem 1.14 can be strengthened by replacing $\delta_{k,\ell,d}^{\mathrm{CON}}$ with $\delta_{\mathcal{C}_{k,\ell,d}}$, as follows.

Conjecture 8.1. *For every $\alpha > 0$ and $k \in \mathbb{N}$, there exists $C = C_{8.1}(\alpha, k)$ such that the following holds for every $\ell, d \in [k-1]$ and every sufficiently large n for which $k - \ell$ divides n. If H is an n-vertex k-uniform hypergraph such that $\delta_d(H) \geq (\delta_{\mathcal{C}_{k,\ell,d}} + \alpha)\binom{n-d}{k-d}$, then there is a (C/n)-vertex-spread distribution on embeddings $C_{n,k,\ell} \hookrightarrow H$.*

Together with Theorem 6.3 and Propositions 1.17 and 6.4, if true Conjecture 8.1 would imply $\delta_{\mathcal{C}_{k,\ell,d}} = \delta_{\mathcal{C}_{k,\ell,d}}^{\mathrm{RD}} = \delta_{\mathcal{C}_{k,\ell,d}}^{\mathrm{VSD}}$. For $\ell = 0$, this indeed holds because $\delta_{k,0,d}^{\mathrm{CON}} = \delta_{\mathcal{C}_{k,0,d}}$ for all k and d. However, for $\ell > 0$, this problem seems very difficult. A natural starting point would be the following case of $k = 3$, $\ell = 2$, and $d = 1$: If H is an n-vertex 3-uniform hypergraph such that $\delta_1(H) \geq (5/9 + \alpha)\binom{n}{2}$, then there is a (C/n)-vertex-spread distribution on embeddings $C_{n,3,2} \hookrightarrow H$. Reiher, Rödl, Ruciński, Schacht, and Szemerédi [30] proved that $\delta_{\mathcal{C}_{3,2,1}} = 5/9$ and noted that $\delta_{3,2,1}^{\mathrm{CON}} > 5/9$, so this case does not follow from Theorem 1.14.

8.2. Exact minimum-degree thresholds

It would be interesting to investigate whether Theorem 1.14 holds with the minimum-degree condition replaced by an *exact* one. However, in general, this problem is challenging because exact minimum-degree conditions for the existence of a single Hamilton ℓ-cycle are known only in a few special cases.

For example, Katona and Kierstead [20] conjectured that every n-vertex k-uniform hypergraph H with $\delta_{k-1}(H) \geq \lfloor (n - k + 3)/2 \rfloor$ has a tight Hamilton cycle. This conjecture was confirmed in the case $k = 3$ for sufficiently large n by Rödl, Ruciński, and Szemerédi [32] but remains open for $k \geq 4$. Nevertheless, it seems likely that if the Katona–Kierstead conjecture holds, then *robust* versions of it hold as well.

More generally, we conjecture that if a k-uniform hypergraph has minimum d-degree large enough to guarantee a Hamilton ℓ-cycle, then there is a (C/n)-vertex-spread embedding of $C_{n,k,\ell}$ into it, where C only depends on k. To that end, for every $k \in \mathbb{N}$, $d \in [k-1]$, $\ell \in \{0, \dots, k-1\}$, and $n \in \mathbb{N}$ divisible by $k - \ell$, let $h_{k,\ell,d}(n)$ be the minimum integer $D \geq 0$ such that the following holds: If H is an n-vertex k-uniform hypergraph satisfying $\delta_d(H) \geq D$, then H has a Hamilton ℓ-cycle.

Conjecture 8.2. *For every $k \in \mathbb{N}$, there exists $C = C_{8.2}(k)$ such that the following holds for every $d \in [k-1]$, $\ell \in \{0, \ldots, k-1\}$, and $n \in \mathbb{N}$ for which $k - \ell$ divides n. If H is an n-vertex k-uniform hypergraph such that $\delta_d(H) \geq h_{k,\ell,d}(n)$, then there is a (C/n)-vertex-spread distribution on embeddings of $C_{n,k,\ell} \hookrightarrow H$.*

Note that Conjecture 8.2 strengthens Conjecture 8.1. As mentioned, it seems difficult to prove such a result without knowing the value of $h_{k,\ell,d}(n)$. It would be interesting to confirm Conjecture 8.2 for the case $k = 3$ and $d = \ell = 2$, since the result of Rödl, Ruciński, and Szemerédi [32] implies $h_{3,2,2}(n) = \lfloor n/2 \rfloor$ for all large n. Another interesting case would be $d = k - 1$ and $\ell = 1$, since a result of Han and Zhao [14] implies $h_{k,1,k-1}(n) = \lceil n/(2k-2) \rceil$ for all large n. As mentioned earlier, the robustness result that would be implied by the existence of such a vertex-spread measure has already been confirmed by Anastos, Chakraborti, Kang, Methuku, and Pfenninger [2], independently of our work here.

It would also be interesting to investigate whether Theorem 1.15 holds with an exact minimum-degree condition. Such a result was proved for $d = k - 1$ and $F \cong K_k^{(k)}$ (so an F-factor is a perfect matching) by Kang, Kelly, Kühn, Osthus, and Pfenninger [19] and for $k = 2$ and $F \cong K_r$ by Pham, Sah, Sawhney, and Simkin [28]. For graphs (i.e. $k = 2$), minimum-degree conditions for the existence of F-factors are comparatively well understood (see e.g. [21,24,23]), and we suspect many of these results likely admit robust versions.

8.3. General robustness

Considering all of the research into robustness to date, it seems that *every* Dirac-type result admits a robust version. It is plausible that for every family of k-uniform hypergraphs $\mathcal{F}$ and $d \in [k-1]$, we have $\delta_{\mathcal{F},d} = \delta_{\mathcal{F},d}^{\mathrm{RD}} = \delta_{\mathcal{F},d}^{\mathrm{VSD}}$, which would partially explain this phenomenon and significantly generalise Conjecture 8.1. However, proving such a result seems out of reach at present.

One consequence of such a result (using the $s = n$ case of the definition of vertex-spread) would be that whenever H has minimum degree large enough to necessarily contain a copy of F, H supports at least $(n/C)^n$ many embeddings (as opposed to copies) of F, where $n = |V(H)| = |V(F)|$. Although plausible, it seems quite difficult to approach even this special case without any knowledge about what F looks like.

Data availability

No data was used for the research described in the article.

Acknowledgments

References

[1] P. Allen, J. Böttcher, J. Corsten, E. Davies, M. Jenssen, P. Morris, B. Roberts, J. Skokan, A robust Corrádi-Hajnal theorem, Random Struct. Algorithms 65 (2024) 61–130.
[2] M. Anastos, D. Chakraborti, D.Y. Kang, A. Methuku, V. Pfenninger, Robust hamiltonicity of Dirac hypergraphs, (2023) forthcoming.
[3] K. Corradi, A. Hajnal, On the maximal number of independent circuits in a graph, Acta Math. Acad. Sci. Hung. 14 (1963) 423–439.
[4] B. Cuckler, J. Kahn, Hamiltonian cycles in Dirac graphs, Combinatorica 29 (2009) 299–326.
[5] G.A. Dirac, Some theorems on abstract graphs, Proc. Lond. Math. Soc. (3) 2 (1952) 69–81.
[6] A. Dudek, A. Frieze, Loose Hamilton cycles in random uniform hypergraphs, Electron. J. Comb. 18 (2011), 14pp.
[7] A. Dudek, A. Frieze, Tight Hamilton cycles in random uniform hypergraphs, Random Struct. Algorithms 42 (2013) 374–385.
[8] A. Espuny Díaz, Y. Person, Spanning F-cycles in random graphs, Comb. Probab. Comput. 32 (2023) 833–850.
[9] A. Ferber, L. Hardiman, A. Mond, Counting Hamilton cycles in Dirac hypergraphs, Combinatorica 43 (2023) 665–680.
[10] K. Frankston, J. Kahn, B. Narayanan, J. Park, Thresholds versus fractional expectation-thresholds, Ann. Math. (2) 194 (2021) 475–495.
[11] Daniel G., Minimum degree requirements for a graph to be Hamilton-connected, Mathematics Stack Exchange, https://math.stackexchange.com/q/3950940 (version: 2020-12-16).
[12] S. Glock, S. Gould, F. Joos, D. Kühn, D. Osthus, Counting Hamilton cycles in Dirac hypergraphs, Comb. Probab. Comput. 30 (2021) 631–653.
[13] P. Gupta, F. Hamann, A. Müyesser, O. Parczyk, A. Sgueglia, A general approach to transversal versions of Dirac-type theorems, Bull. Lond. Math. Soc. 55 (2023) 2817–2839.
[14] J. Han, Y. Zhao, Minimum codegree threshold for Hamilton ℓ-cycles in k-uniform hypergraphs, J. Comb. Theory, Ser. A 132 (2015) 194–223.
[15] A. Johansson, J. Kahn, V. Vu, Factors in random graphs, Random Struct. Algorithms 33 (2008) 1–28.
[16] F. Joos, R. Lang, N. Sanhueza-Matamala, Robust hamiltonicity, arXiv:2312.15262, 2023.
[17] J. Kahn, G. Kalai, Thresholds and expectation thresholds, Comb. Probab. Comput. 16 (2007) 495–502.
[18] J. Kahn, B. Narayanan, J. Park, The threshold for the square of a Hamilton cycle, Proc. Am. Math. Soc. 149 (2021) 3201–3208.
[19] D.Y. Kang, T. Kelly, D. Kühn, D. Osthus, V. Pfenninger, Perfect matchings in random sparsifications of Dirac hypergraphs, Combinatorica (2024), https://doi.org/10.48550/arXiv.2211.01325.
[20] G.Y. Katona, H.A. Kierstead, Hamiltonian chains in hypergraphs, J. Graph Theory 30 (1999) 205–212.
[21] J. Komlós, G.N. Sárközy, E. Szemerédi, Proof of the Alon-Yuster conjecture, in: Combinatorics, Prague, 1998, vol. 235, 2001, pp. 255–269.
[22] M. Krivelevich, C. Lee, B. Sudakov, Robust hamiltonicity of Dirac graphs, Trans. Am. Math. Soc. 366 (2014) 3095–3130.
[23] D. Kühn, D. Osthus, Embedding large subgraphs into dense graphs, in: Surveys in Combinatorics 2009, in: London Math. Soc. Lecture Note Ser., vol. 365, Cambridge Univ. Press, Cambridge, 2009, pp. 137–167.
[24] D. Kühn, D. Osthus, The minimum degree threshold for perfect graph packings, Combinatorica 29 (2009) 65–107.
[25] M. Molloy, B. Reed, Graph Colouring and the Probabilistic Method, Algorithms and Combinatorics, vol. 23, Springer-Verlag, Berlin, 2002.
[26] R. Montgomery, M. Pavez-Signé, Counting Spanning Subgraphs in Dense Hypergraphs, 2023.
[27] J. Park, H.T. Pham, A proof of the Kahn-Kalai conjecture, J. Am. Math. Soc. 37 (2024) 235–243.
[28] H.T. Pham, A. Sah, M. Sawhney, M. Simkin, A toolkit for robust thresholds, arXiv:2210.03064, 2022.
[29] L. Pósa, Hamiltonian circuits in random graphs, Discrete Math. 14 (1976) 359–364.
[30] C. Reiher, V. Rödl, A. Ruciński, M. Schacht, E. Szemerédi, Minimum vertex degree condition for tight Hamiltonian cycles in 3-uniform hypergraphs, Proc. Lond. Math. Soc. (3) 119 (2019) 409–439.
[31] O. Riordan, Random cliques in random graphs and sharp thresholds for F-factors, Random Struct. Algorithms 61 (2022) 619–637.

[32] V. Rödl, A. Ruciński, E. Szemerédi, Dirac-type conditions for Hamiltonian paths and cycles in 3-uniform hypergraphs, Adv. Math. 227 (2011) 1225–1299.

[33] G.N. Sárközy, S.M. Selkow, E. Szemerédi, On the number of Hamiltonian cycles in Dirac graphs, Discrete Math. 265 (2003) 237–250.

[34] S. Spiro, A smoother notion of spread hypergraphs, Comb. Probab. Comput. (2023) 1–10.

[35] B. Sudakov, Robustness of graph properties, in: Surveys in Combinatorics 2017, in: London Math. Soc. Lecture Note Ser., vol. 440, Cambridge Univ. Press, Cambridge, 2017, pp. 372–408.

[36] M. Talagrand, Are many small sets explicitly small?, in: Proceedings of the Forty-Second ACM Symposium on Theory of Computing, 2010, pp. 13–36.