

SHARP THRESHOLDS FOR NONLINEAR HAMILTONIAN CYCLES IN HYPERGRAPHS

BHARGAV NARAYANAN AND MATHIAS SCHACHT

ABSTRACT. For positive integers $r > \ell$, an r -uniform hypergraph is called an ℓ -cycle if there exists a cyclic ordering of its vertices such that each of its edges consists of r consecutive vertices, and such that every pair of consecutive edges (in the natural ordering of the edges) intersect in precisely ℓ vertices. Such cycles are said to be *linear* when $\ell = 1$, and *nonlinear* when $\ell > 1$. We determine the sharp threshold for nonlinear Hamiltonian cycles and show that for all $r > \ell > 1$, the threshold $p_{r,\ell}^*(n)$ for the appearance of a Hamiltonian ℓ -cycle in the random r -uniform hypergraph on n vertices is sharp and is $p_{r,\ell}^*(n) = \lambda(r, \ell) \left(\frac{e}{n}\right)^{r-\ell}$ for an explicitly specified function λ . This resolves several questions raised by Dudek and Frieze in 2011.

§1. INTRODUCTION

A basic problem in probabilistic combinatorics concerns locating the critical density at which a substructure of interest appears inside a random structure (with high probability). In the context of random graph theory, the question of when a random graph contains a *Hamiltonian cycle* has received considerable attention. Indeed, from the foundational works of Pósa [13], Komlós and Szemerédi [11], Bollobás [2], and Ajtai, Komlós and Szemerédi [1], we have a very complete picture, understanding not only the *sharp threshold* for this problem but the *hitting time* as well. Since these early breakthroughs, there have been a number of papers locating thresholds, both asymptotic and sharp, for various spanning subgraphs of interest (see, e.g., [12, 14] and the references therein for various related results).

In contrast, threshold results for spanning structures in the context of random hypergraph theory have been somewhat harder to come by. Indeed, even the basic question of locating the asymptotic threshold at which a random r -uniform hypergraph (or r -graph, for short) contains a matching, i.e., a spanning collection of disjoint edges, proved to be a major challenge, resisting the efforts of a number of researchers up until the breakthrough work of Johansson, Kahn and Vu [9]; more recently, both the sharp threshold as well as the hitting time for this problem have been obtained by Kahn [10]. In the light of these developments for matchings, we study what is perhaps the next most natural question in this setting,

2010 *Mathematics Subject Classification.* Primary 05C80; Secondary 05C65, 05C45.



FIGURE 1.1. A 7-uniform 4-cycle.

namely that of when a random r -graph contains a Hamiltonian cycle; our main contribution is to resolve the sharp threshold problem for nonlinear Hamiltonian cycles.

There are multiple notions of cycles in hypergraphs, so let us recall the relevant definitions: given positive integers $r > \ell \geq 1$, an r -graph is called an ℓ -cycle if there exists a cyclic ordering of its vertices such that each of its edges consists of r consecutive vertices in the ordering, and such that every pair of consecutive edges (in the natural ordering of the edges) intersect in precisely ℓ vertices (see Figure 1.1 for an example). A *Hamiltonian* ℓ -cycle is then an ℓ -cycle spanning the entire vertex set; of course, an r -graph on n vertices may only contain a Hamiltonian ℓ -cycle when $(r - \ell) \mid n$, and such a cycle then has precisely $n/(r - \ell)$ edges. Finally, by convention, an ℓ -cycle is called *linear* (or *loose*) when $\ell = 1$, *nonlinear* when $\ell > 1$, and *tight* when $\ell = r - 1$.

Given $r > \ell \geq 1$, we set

$$\lambda(r, \ell) = t! \cdot (s - t)!,$$

where $s = r - \ell$ and $1 \leq t \leq s$ is the unique integer satisfying $t \equiv r \pmod{s}$, and define

$$p_{r,\ell}^*(n) = \frac{\lambda(r, \ell)e^s}{n^s}.$$

Writing $G^{(r)}(n, p)$ for the binomial random r -graph on n vertices, where each possible r -set of vertices appears as an edge independently with probability p , our main result is as follows.

Theorem 1.1. *For all integers $r > \ell > 1$ and all $\varepsilon > 0$, as $n \rightarrow \infty$ with $(r - \ell) \mid n$, we have*

$$\mathbb{P}(G^{(r)}(n, p) \text{ contains a Hamiltonian } \ell\text{-cycle}) \rightarrow \begin{cases} 1 & \text{if } p > (1 + \varepsilon)p^*, \text{ and} \\ 0 & \text{if } p < (1 - \varepsilon)p^*, \end{cases}$$

where we abbreviate $p^* = p_{r,\ell}^*(n)$.

The critical density $p_{r,\ell}^*$ appearing in our result corresponds to the so-called ‘expectation threshold’, namely the density above which the expected number of Hamiltonian ℓ -cycles in $G^{(r)}(n, p)$ begins to diverge. A moment’s thought should convince the reader that, unlike in the case of linear Hamiltonian cycles where one has to worry about isolated vertices, there are no ‘coupon collector type’ obstacles to the presence of nonlinear Hamiltonian cycles; therefore, the conclusion of Theorem 1.1 should not come as a surprise. Indeed, the

problem of whether something like Theorem 1.1 ought to hold was raised by Dudek and Frieze [4, 5]. Towards such a result, they showed that $p_{r,\ell}^*$ is an asymptotic threshold for all nonlinear Hamiltonian cycles, that $p_{r,\ell}^*$ is a sharp threshold for tight Hamiltonian cycles when $r \geq 4$, and that $p_{r,\ell}^*$ is a semi-sharp threshold for all $r > \ell \geq 3$.

The main difficulty in proving Theorem 1.1 is that, with the exception of the case of tight Hamiltonian cycles with $r \geq 4$ mentioned earlier, the second moment method is in itself not sufficient to prove the result; for instance, it is easy to verify, even in the simple case of $r = 3$ and $\ell = 2$ (i.e., tight Hamiltonian cycles in 3-graphs) that the requisite second moment is too large to yield our result. To prove Theorem 1.1, we shall combine a careful second moment estimate, which necessitates working modulo various symmetries, with a powerful theorem of Friedgut [6] characterising coarse thresholds.

This paper is organised as follows. We gather the tools we require in Section 2. The proof of Theorem 1.1 follows in Section 3. We conclude in Section 4 with a discussion of some open problems.

§2. PRELIMINARIES

We begin with some background on thresholds. Recall that a *monotone r -graph property* W is a sequence $(W_n)_{n \geq 0}$ of families of r -graphs, where W_n is a family of r -graphs on n vertices closed under the addition of edges and invariant under r -graph isomorphism.

Given a monotone r -graph property $W = (W_n)_{n \geq 0}$, a function $p^*(n)$ is said to be a *threshold* or *asymptotic threshold* for W if $\mathbb{P}(G^{(r)}(n, p) \in W_n)$ tends, as $n \rightarrow \infty$, either to 1 or 0 as p/p^* tends either to ∞ or 0 respectively, and a function $p^*(n)$ is said to be a *sharp threshold* for W if $\mathbb{P}(G^{(r)}(n, p) \in W_n)$ tends, as $n \rightarrow \infty$, either to 1 or 0 as p/p^* remains bounded away from 1 either from above or below respectively. Of course, thresholds and sharp thresholds are not unique, but following common practice, we will often say ‘the’ threshold or sharp threshold when referring to the appropriate equivalence class of functions. Finally, a function $p^*(n)$ is said to be a *semi-sharp threshold* for W if there exist constants $C_0 \leq 1 \leq C_1$ such that $\mathbb{P}(G^{(r)}(n, p) \in W_n)$ tends, as $n \rightarrow \infty$, either to 1 or 0 as p/p^* remains bounded below by C_1 or above by C_0 respectively; while we do not need this notion ourselves, we give this definition to place existing results around our main result in the appropriate context.

That every monotone property has an asymptotic threshold follows from a (much more general) result of Bollobás and Thomason [3]. Unlike with asymptotic thresholds, a monotone property need not necessarily have a sharp threshold; such properties are said to have *coarse thresholds*. We shall make use of a powerful characterisation of monotone properties that have coarse thresholds due to Friedgut [6] which says, roughly, that such

properties are ‘approximable by a local property’; a concrete formulation at a level of generality sufficient for our purposes, see [7], is as follows.

Proposition 2.1. *Fix $r \in \mathbb{N}$ and let $W = (W_n)_{n \geq 0}$ be a monotone r -graph property that has a coarse threshold. Then there exists a constant $\alpha > 0$, a threshold function $\hat{p} = \hat{p}(n)$ with*

$$\alpha < \mathbb{P}(G^{(r)}(n, \hat{p}) \in W_n) < 1 - 3\alpha$$

for all $n \in \mathbb{N}$, a constant $\beta > 0$ and a fixed r -graph F such that the following holds: for infinitely many $n \in \mathbb{N}$, there exists an r -graph on n vertices $H_n \notin W_n$ such that

$$\mathbb{P}(H_n \cup G^{(r)}(n, \beta \hat{p}) \in W_n) < 1 - 2\alpha,$$

where the random r -graph $G^{(r)}(n, \beta \hat{p})$ is taken to be on the same vertex set as H_n , and

$$\mathbb{P}(H_n \cup \tilde{F} \in W_n) > 1 - \alpha,$$

where \tilde{F} denotes a random copy of F on the same vertex set as H_n . □

We shall also require the Paley–Zygmund inequality.

Proposition 2.2. *If X is a non-negative random variable, then*

$$\mathbb{P}(X > 0) \geq \frac{\mathbb{E}[X]^2}{\mathbb{E}[X^2]}. \quad \square$$

Finally, we collect together some standard estimates for factorials and binomial coefficients.

Proposition 2.3. *For all $n \in \mathbb{N}$, we have*

$$\sqrt{2\pi n} \left(\frac{n}{e}\right)^n \leq n! \leq e\sqrt{n} \left(\frac{n}{e}\right)^n,$$

and for all positive integers $1 \leq k \leq n$, we have

$$\binom{n}{k} \leq \left(\frac{en}{k}\right)^k. \quad \square$$

§3. PROOF OF THE MAIN RESULT

In this section, we shall prove Theorem 1.1. We begin by setting up some notational conventions that we shall adhere to in the sequel.

In what follows, we fix $r, \ell \in \mathbb{N}$ with $r > \ell > 1$, set $s = r - \ell$, take t to be the unique integer satisfying $t \equiv r \pmod{s}$ with $1 \leq t \leq s$, and set $\lambda = t!(s-t)!$. We shall henceforth assume that n is a large integer divisible by s , and we set $m = n/s$ so that m is the number of edges in an ℓ -cycle on n vertices. Finally, all r -graphs on n vertices in the sequel will implicitly be assumed to be on the vertex set $[n] = \{1, 2, \dots, n\}$.

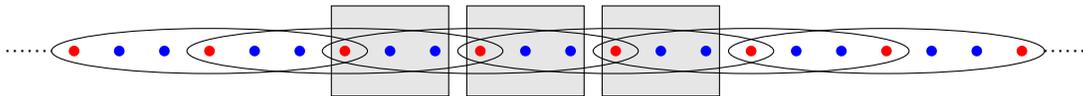


FIGURE 3.1. Blocks and subblocks of a 7-uniform 4-cycle.

To deal with r -graph cycles on the vertex set $[n]$, we shall define an equivalence relation on S_n , the symmetric group of permutations of $[n] = \{1, 2, \dots, n\}$; we shall ignore the group structure of S_n for the most part, so for us, a permutation $\sigma \in S_n$ is just an arrangement $\sigma(1), \sigma(2), \dots, \sigma(n)$ of the elements of $[n]$ (namely, vertices), at locations indexed by $[n]$.

We divide $[n]$ into m *blocks* of size s , where for $0 \leq i < m$, the i -th such block is comprised of the interval $\{is + 1, is + 2, \dots, is + s\}$ of vertices, and we further divide each such block into two *subblocks*, where the t -subblock of a block consists of the first t vertices in the block, and the $(s - t)$ -subblock of a block consists of the last $s - t$ vertices in the block. Now, define an equivalence relation on S_n by saying that two permutations σ and τ are *subblock equivalent* if τ may be obtained from σ by only rearranging vertices within subblocks; in other words, an equivalence class of this equivalence relation may be viewed as an element of the quotient $Q_n = S_n / (S_t \times S_{s-t})^m$.

The definition of the above equivalence relation is motivated by the natural ℓ -cycle associated with a permutation: given $\sigma \in S_n$, consider the r -graph H_σ on $[n]$ with m edges, where for $0 \leq i < m$, the i -th edge of H_σ is the r -set $\{\sigma(is + 1), \sigma(is + 2), \dots, \sigma(is + r)\}$, of course with indices being considered cyclically modulo n . It is easy to verify both that H_σ is an ℓ -cycle for each $\sigma \in S_n$, and that if σ and τ are subblock equivalent, then $H_\sigma = H_\tau$; see Figure 3.1 for an illustration. Hence, in what follows, we shall abuse notation and call the elements of Q_n permutations (when strictly speaking, they are equivalence classes of permutations), and for $\sigma \in Q_n$, we write H_σ for the natural ℓ -cycle associated with σ .

We parameterise $p = Cp_{r,\ell}^*(n) = C\lambda e^s/n^s$ for some constant $C > 0$, and work with $G = G^{(r)}(n, p)$, where we take the vertex set of G to be $[n]$. Therefore, our goal is to show that G contains a Hamiltonian ℓ -cycle with high probability when $C > 1$ (namely, the 1-statement), and that G does not contain a Hamiltonian ℓ -cycle with high probability when $C < 1$ (namely, the 0-statement).

In what follows, constants suppressed by asymptotic notation are allowed to depend on fixed parameters (quantities depending only on r, ℓ, C , etc.) but not on variables that depend on n , which we send to infinity along multiples of s . We also adopt the standard convention that an event holds with high probability if the probability of the event in question is $1 - o(1)$ as $n \rightarrow \infty$.

We shall focus our attention on the random variable X that counts the number of $\sigma \in Q_n$ for which the ℓ -cycle H_σ is contained in G , noting that G contains a Hamiltonian ℓ -cycle if and only if $X > 0$.

We start by computing the first moment of X .

Lemma 3.1. *We have $\mathbb{E}[X] = |Q_n|p^m = n!(p/\lambda)^m$, so that*

$$\mathbb{E}[X] \longrightarrow \begin{cases} \infty & \text{if } C > 1, \text{ and} \\ 0 & \text{if } C < 1. \end{cases}$$

Proof. This follows from noting that $|Q_n| = n!/\lambda^m$, estimating $n!$ using Proposition 2.3, and using the fact that $n = ms$. \square

In particular, the above first moment estimate, combined with Markov's inequality, establishes the 0-statement. To establish the 1-statement, the following second moment estimate will be crucial.

Lemma 3.2. *For $C > 1$, we have $\mathbb{E}[X^2] = O(\mathbb{E}[X]^2)$.*

Let us point out that Lemma 3.2 does not make the stronger promise that

$$\mathbb{E}[X^2] = (1 + o(1))\mathbb{E}[X]^2,$$

and indeed, such an estimate does not hold generally for an arbitrary pair of integers $r > \ell > 1$.

Proof of Lemma 3.2. To estimate the second moment of X , it will be convenient to make the following definition: for $0 \leq b \leq m$, let $N(b)$ denote, for any fixed permutation $\sigma \in Q_n$, the number of permutations $\tau \in Q_n$ meeting σ in b edges, by which we mean that H_τ intersects H_σ in exactly b edges. With this definition in place, using the trivial fact that $N(0) \leq |Q_n|$, we have

$$\begin{aligned} \mathbb{E}[X^2] &= \sum_{\sigma, \tau \in Q_n} \mathbb{P}(H_\sigma \cup H_\tau \subset G) \\ &= |Q_n|p^m \sum_{b=0}^m N(b)p^{m-b} \\ &\leq |Q_n|^2 p^{2m} + |Q_n|p^m \sum_{b=1}^m N(b)p^{m-b} \\ &= \mathbb{E}[X]^2 + \mathbb{E}[X] \sum_{b=1}^m N(b)p^{m-b}, \end{aligned}$$

whence it follows that

$$\frac{\mathbb{E}[X^2]}{\mathbb{E}[X]^2} \leq 1 + \sum_{b=1}^m \frac{N(b)p^{-b}}{|Q_n|},$$

so to prove Lemma 3.2, it suffices to show that the sum

$$\Gamma = \sum_{b=1}^m \frac{N(b)p^{-b}}{|Q_n|}$$

satisfies the estimate $\Gamma = O(1)$ when $C > 1$.

The rough plan of attack now is similar to that adopted by Dudek and Frieze [4], but we shall require a more careful two-stage analysis since we require stronger estimates: first, we shall control the ‘canonical’ contributions to Γ , and subsequently bound the ‘non-canonical’ contributions in terms of the aforementioned ‘canonical’ ones; we make precise these notions below.

An r -graph is called an ℓ -*path* if there exists a linear ordering of its vertices such that each of its edges consists of r consecutive vertices, and such that every pair of consecutive edges (in the natural ordering of the edges) intersect in precisely ℓ vertices. Given a permutation $\sigma \in Q_n$, we say that $\tau \in Q_n$ meets σ *canonically* if H_τ meets H_σ in a family of vertex-disjoint ℓ -paths, and we otherwise say that τ meets σ *non-canonically*.

For $1 \leq b \leq m$, let $N_c(b)$ to be the number of permutations $\tau \in Q_n$ which canonically meet a fixed permutation $\sigma \in Q_n$ in b edges, set $N'(b) = N(b) - N_c(b)$, and decompose $\Gamma = \Gamma_c + \Gamma'$, where naturally

$$\Gamma_c = \sum_{b=1}^m \frac{N_c(b)p^{-b}}{|Q_n|}$$

and

$$\Gamma' = \Gamma - \Gamma_c = \sum_{b=1}^m \frac{N'(b)p^{-b}}{|Q_n|}.$$

First, we bound the canonical contributions to Γ .

Claim 3.3. *For $C > 1$, we have $\Gamma_c = O(1)$.*

Proof. Fix a permutation $\sigma \in Q_n$ and for $1 \leq a \leq b$, write $N_c(b, a)$ for the number of permutations $\tau \in Q_n$ which meet σ canonically in b edges which together form a vertex-disjoint ℓ -paths in H_σ . We now proceed to estimate $N_c(b, a)$.

Given σ , a (b, a) -configuration in σ is a collection of b edges in H_σ which together form a vertex-disjoint ℓ -paths; clearly, a (b, a) -configuration covers $sb + la$ vertices. The number of ways to choose a (b, a) -configuration in σ is clearly at most

$$\binom{m}{a} \binom{b}{a},$$

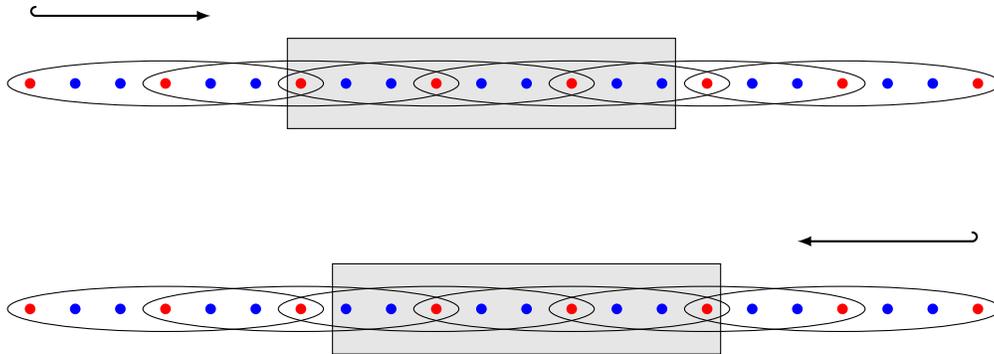


FIGURE 3.2. The rigid interior of a 7-uniform 4-path in the two possible directions of embedding.

since there are at most $\binom{m}{a}$ ways of locating the leftmost edge in each of the a ℓ -paths in H_σ , and the number of ways to subsequently choose the number of edges in each of these a paths so that there are b edges in total is clearly at most the number of solutions to the equation $x_1 + x_2 + \cdots + x_a = b$ over the positive integers, which is $\binom{b-1}{a-1} \leq \binom{b}{a}$.

Next, given a (b, a) -configuration \mathcal{P} in σ , let us count the number of choices for $\tau \in Q_n$ for which H_τ contains \mathcal{P} . We do this in two steps. First, we count the number of ways in which the vertices covered by \mathcal{P} can be embedded into τ , and then estimate the number of ways in which the vertices not covered by \mathcal{P} can be ordered in τ , ensuring at all times that we only count up to subblock equivalence.

Now, there are at most $a! \binom{m}{a}$ ways to choose the starting blocks of the leftmost edges of the a distinct ℓ -paths of \mathcal{P} in τ . Once the left endpoint of one of these ℓ -paths has been fixed in τ , we observe that there are only $O(1)$ ways, up to subblock equivalence, to embed the remaining vertices of this ℓ -path into τ ; indeed, the relative ordering of all the vertices in an ℓ -path, with $O(1)$ exceptions at the left and right extremes, is rigid up to subblock equivalence, up to a reversal of the direction of embedding (left-to-right or right-to-left), as shown in Figure 3.2. Consequently, once the location of the a leftmost edges have been determined in τ , the number of ways of embedding the rest of \mathcal{P} into τ is at most L^a for some $L = L(r, \ell)$. We conclude that the number of ways to embed \mathcal{P} is at most

$$\binom{m}{a} a! L^a.$$

Once we have embedded \mathcal{P} into τ , there are $(n - sb - \ell a)!$ ways to arrange the remaining vertices uncovered by \mathcal{P} , without accounting for subblock equivalence. It is easy to check that any embedding of \mathcal{P} covers at most $b + \ell a$ blocks in τ , so the number of choices

for $\tau \in Q_n$ with a given embedding of \mathcal{P} is at most

$$\frac{(n - sb - \ell a)!}{\lambda^{m-b-\ell a}}.$$

From the above estimates and using Proposition 2.3 to bound binomial coefficients, we conclude that

$$\begin{aligned} N_c(b, a) &\leq \binom{m}{a} \binom{b}{a} \binom{m}{a} a! L^a \frac{(n - sb - \ell a)!}{\lambda^{m-b-\ell a}} \\ &\leq \exp(O(a)) \frac{n^{2a} b^a (n - sb - \ell a)!}{a^{2a} \lambda^{m-b}}, \end{aligned} \quad (3.1)$$

where, as remarked upon before, constants suppressed by the asymptotic notation depend only on r and ℓ .

To finish the proof of the lemma, we now use, in order, the above bound (3.1), the fact that $\ell \geq 2$, Proposition 2.3, and the fact that $1 + x \leq e^x$ for all $x \in \mathbb{R}$ to show that

$$\begin{aligned} \Gamma_c &= \sum_{b=1}^m \sum_{a=1}^b \frac{N_c(b, a) p^{-b}}{|Q_n|} \\ &\leq \sum_{b=1}^m \sum_{a=1}^b \exp(O(a)) \frac{n^{2a} b^a (n - sb - \ell a)! \lambda^m}{a^{2a} \lambda^{m-b} n!} \frac{n^{sb}}{C^b \lambda^b e^{sb}} \\ &\leq \sum_{b=1}^m \sum_{a=1}^b C^{-b} \exp(O(a)) \frac{n^{2a} b^a (n - sb - 2a)! n^{sb}}{a^{2a} n! e^{sb}} \\ &\leq \sum_{b=1}^m \sum_{a=1}^b C^{-b} \exp(O(a)) \frac{n^{sb+2a} b^a (n - sb - 2a)^{n-sb-2a}}{a^{2a} n^n} \\ &\leq \sum_{b=1}^m \sum_{a=1}^b C^{-b} \exp(O(a)) \frac{b^a}{a^{2a}} \left(1 - \frac{sb + 2a}{n}\right)^{n-sb-2a} \\ &\leq \sum_{b=1}^m \sum_{a=1}^b C^{-b} \exp(a \log b - 2a \log a - sb - 2a + (sb + 2a)^2/n + O(a)). \end{aligned}$$

We uniformly have $(4a^2 + 4sab)/n = O(n)$, since $a \leq b \leq m \leq n$, so the above estimate reduces to

$$\Gamma_c \leq \sum_{b=1}^m \left(C^{-b} \exp(-sb + (sb)^2/n) \sum_{a=1}^b \exp(a \log b - 2a \log a + O(a)) \right). \quad (3.2)$$

Finally, since $sb \leq sm = n$, we uniformly have

$$\exp(-sb + (sb)^2/n) \leq 1$$

for all $1 \leq b \leq m$, and it is straightforward to verify that we uniformly have

$$\exp(a \log b - 2a \log a + O(a)) = \exp(o(b))$$

for all $1 \leq a \leq b$. Using these two bounds in (3.2), it follows that for $C > 1$, we have

$$\Gamma_c \leq \sum_{b=1}^m C^{-b} b \exp(o(b)) = \sum_{b=1}^m C^{-b+o(b)} = O(1),$$

proving the claim. \square

The second and final step in the proof of Lemma 3.2 is to estimate the non-canonical contributions to Γ .

Claim 3.4. *For $C > 1$, we have $\Gamma' = O(1)$.*

Proof. We shall prove the claim by means of a comparison argument: we shall demonstrate how we may group summands in Γ' so as to get estimates analogous to those that we obtained for Γ in the proof of Claim 3.3.

For any $\sigma, \tau \in Q_n$, we may decompose the intersection of H_σ and H_τ into a collection of vertex-disjoint *weak paths*, where a *weak path* is just a sequence of edges in which every consecutive pair of edges intersect.

We fix a permutation $\sigma \in Q_n$ for the rest of the argument. Given a weak path P' in H_σ , notice that there is a unique minimal ℓ -path P in H_σ covering precisely the same set of vertices as P' ; we call P the *minimal cover* of P' . Now, given any $\tau \in Q_n$, there is a unique minimal *covering configuration in σ* associated with τ obtained by taking the minimal covers of each of the weak paths in which H_τ meets H_σ . To prove the claim, we shall show that the contributions to Γ' from all those $\tau \in Q_n$ whose covering configuration is \mathcal{P} is comparable to the contributions to Γ_c from all those $\tau \in Q_n$ meeting σ canonically in \mathcal{P} .

We fix a (b, a) configuration \mathcal{P} in σ consisting of b edges in total distributed across a ℓ -paths, and we consider the set of permutations $\tau \in Q_n$ with minimal cover \mathcal{P} that meet σ non-canonically; we additionally parametrise this set by $1 \leq k \leq b$, writing $Q(\mathcal{P}, k)$ for the set of such permutations τ for which there are k edges of \mathcal{P} missing from the intersection \mathcal{P}' of H_τ and H_σ .

We claim that the number of ways to select a configuration \mathcal{P}' as above, and then embed the vertices covered by \mathcal{P}' into a permutation $\tau \in Q(\mathcal{P}, k)$ in such a way that \mathcal{P}' is contained in H_τ is, up to subblock equivalence, at most

$$\binom{b}{k} \binom{m}{a} a! R^{a+k}$$

for some $R = R(r, \ell)$.

We may verify the estimate above as follows. The number of possible choices for \mathcal{P}' , namely the number of ways to choose k edges from \mathcal{P} such that each of the ℓ -paths of \mathcal{P} remains a weak path after these k edges are removed, may be crudely bounded above

by $\binom{b}{k}$. As in the proof of Claim 3.3, there are $\binom{m}{a}a!$ ways to choose the starting blocks of the leftmost edges of the a distinct weak paths of \mathcal{P}' in τ . Assume now that we have fixed \mathcal{P}' , the starting blocks in τ of the leftmost edges of the a weak paths of \mathcal{P}' , and the directions of embedding of these weak paths into τ (for which there are 2^a choices). Now, it is easy to see from the linear structure of a weak path that the relative order of vertices in disjoint edges of a weak path must be preserved in any embedding of that weak path into τ , so in particular, there are only $O(1)$ choices for the location in τ of any particular vertex covered by \mathcal{P}' (once endpoints and directions of embedding have been fixed, as we have assumed). Furthermore, it follows from the rigidity of an ℓ -path (as in the proof of Claim 3.3) that any vertex covered by \mathcal{P}' , with $O(1)$ exceptions at the left and right extremities of each of the a weak paths, which possesses potential embedding locations in more than one subblock must necessarily be within $O(1)$ distance (in σ) of some edge present in \mathcal{P} but not in \mathcal{P}' ; clearly there $O(k)$ such vertices in total. These facts taken together demonstrate the validity of the bound claimed above.

Now, noting that the contribution of any $\tau \in Q(\mathcal{P}, k)$ to Γ' is a factor of p^k times the contribution to Γ_c from any $\tau \in Q_n$ meeting σ canonically in \mathcal{P} , we may mimic the proof of Claim 3.3 to show that

$$\Gamma' \leq \sum_{b=1}^m \sum_{a=1}^b \left(\exp(O(a)) \frac{n^{2a} b^a}{a^{2a}} \frac{(n - sb - \ell a)! \lambda^m}{\lambda^{m-b}} \frac{n^{sb}}{n!} \frac{1}{C^b \lambda^b e^{sb}} \sum_{k=1}^b \binom{b}{k} R^k p^k \right).$$

Observing that

$$\sum_{k=1}^b \binom{b}{k} R^k p^k = \sum_{k=1}^b \exp(O(k)) \frac{b^k}{k^k n^{sk}} = O(1),$$

we are left with an estimate for Γ' of the same form as the one for Γ which we showed to be $O(1)$ when $C > 1$ in the proof of Claim 3.3; the claim follows. \square

The two claims above together imply that $\Gamma = O(1)$ when $C > 1$, from which it follows that $\mathbb{E}[X^2] = O(\mathbb{E}[X]^2)$ when $C > 1$; the result follows. \square

With our moment estimates in hand, we are now ready to prove our main result.

Proof of Theorem 1.1. As mentioned earlier, the 0-statement, namely that $G^{(r)}(n, p)$ does not contain a Hamiltonian ℓ -cycle with high probability if $p < (1 + \varepsilon)p_{r,\ell}^*(n)$ follows immediately from Lemma 3.1 and Markov's inequality.

We prove the 1-statement, namely that $G^{(r)}(n, p)$ contains a Hamiltonian ℓ -cycle with high probability if $p > (1 + \varepsilon)p_{r,\ell}^*(n)$, by showing that the property of containing a Hamiltonian ℓ -cycle has a sharp threshold, and that this sharp threshold must (asymptotically) necessarily be $p_{r,\ell}^*(n)$.

If $p > (1 + \varepsilon)p_{r,\ell}^*(n)$, then it follows from Lemma 3.2 and the Paley–Zygmund inequality, i.e., Proposition 2.2, that $G^{(r)}(n, p)$ contains a Hamiltonian ℓ -cycle with probability at least $\delta > 0$ for some $\delta = \delta(\varepsilon, r, \ell)$; consequently, if the property of containing a Hamiltonian ℓ -cycle has a sharp threshold, this sharp threshold is necessarily asymptotic to $p_{r,\ell}^*(n)$.

It remains to prove that the monotone r -graph property $W = (W_n)_{n \geq 0}$ of containing a Hamiltonian ℓ -cycle has a sharp threshold, so suppose for the sake of a contradiction that W has a coarse threshold.

It follows from Proposition 2.1 that there is a fixed r -graph F and a threshold function $\hat{p} = \hat{p}(n)$ with the property that for infinitely many $n \in \mathbb{N}$, there is an r -graph $H_n \notin W_n$ on n vertices such that adding a random copy of F to H_n is significantly more likely to make the resulting graph contain a Hamiltonian ℓ -cycle than adding a random collection of edges of density about \hat{p} ; concretely, for some universal constants $\alpha, \beta > 0$, we have

$$\mathbb{P}(H_n \cup G^{(r)}(n, \beta\hat{p}) \in W_n) < 1 - 2\alpha, \quad (3.3)$$

where the random r -graph $G^{(r)}(n, \beta\hat{p})$ is taken to be on the same vertex set as H_n , and

$$\mathbb{P}(H_n \cup \tilde{F} \in W_n) > 1 - \alpha, \quad (3.4)$$

where \tilde{F} denotes a random copy of F on the same vertex set as H_n .

Now, the only way F can help induce a Hamiltonian ℓ -cycle in H_n is through some sub-hypergraph of itself that appears in all large enough ℓ -cycles, so by pigeonholing (and adding extra edges if necessary), we conclude from (3.4) that there exists a fixed ℓ -path P , say with k edges on $\ell + sk$ vertices, with the property that, for some universal constant $\gamma > 0$, we have

$$\mathbb{P}(H_n \cup \tilde{P} \in W_n) > \gamma,$$

where \tilde{P} again denotes a random copy of P on the same vertex set as H_n . In other words, a positive fraction of all the possible ways to embed P into the vertex set of H_n are *useful* and end up completing Hamiltonian ℓ -cycles.

Since \hat{p} is an asymptotic threshold for W , clearly $\hat{p}(n) = \Theta(p_{r,\ell}^*(n)) = \Theta(n^{-s})$, since $p_{r,\ell}^*$ is also an asymptotic threshold for W , as can be read off from the proof of Lemma 3.2. On the other hand, the expected number of useful copies of P created by the addition of a $\beta\hat{p} = \Theta(n^{-s})$ density of random edges to H_n is

$$\Omega\left(\binom{n}{\ell + sk} (n^{-s})^k\right) = \Omega(n^\ell),$$

and a routine application of the second moment method (indeed, ℓ -paths are suitably ‘balanced’) shows that adding a $\beta\hat{p}$ density of random edges to H_n must, with high

probability, create at least one useful copy of P in H_n and complete a Hamiltonian ℓ -cycle, contradicting (3.3).

We have now shown that W has a sharp threshold, and that this threshold must be asymptotic to $p_{r,\ell}^*(n)$; the 1-statement follows, completing the proof. \square

§4. CONCLUSION

There are two basic questions that our work raises; we conclude this paper by discussing these problems.

First, now that we have identified the sharp threshold for the appearance of nonlinear Hamiltonian cycles, one can and should ask about the ‘width’ of the critical window. Since the sharp threshold corresponds to the expectation threshold, we do not expect the hitting time to be of interest. Nonetheless, it is plausible that the expectation threshold is much sharper than what we have shown, and we conjecture the following.

Conjecture 4.1. *For all integers $r > \ell > 1$, if $p = p(n)$ is such that, as $n \rightarrow \infty$, we have $\mathbb{E}[X_\ell] \rightarrow \infty$, then*

$$\mathbb{P}(G^{(r)}(n, p) \text{ contains a Hamiltonian } \ell\text{-cycle}) \rightarrow 1,$$

where X_ℓ is the random variable counting the number of Hamiltonian ℓ -cycles in $G^{(r)}(n, p)$.

Second, it is natural to ask what happens for linear Hamiltonian cycles. The proof of Theorem 1.1 shows that the appearance of a linear Hamiltonian cycle in $G^{(r)}(n, p)$ has a sharp threshold, and we expect this sharp threshold to coincide with the sharp threshold for the disappearance of isolated vertices (i.e., vertices not contained in any edges). For $r \geq 3$, writing

$$p_r^{\text{deg}}(n) = \frac{(r-1)! \log n}{n^{r-1}}$$

to denote the sharp threshold for the disappearance of isolated vertices in $G^{(r)}(n, p)$, we predict the following.

Conjecture 4.2. *For each $r \geq 3$, $p_r^{\text{deg}}(n)$ is the sharp threshold for the appearance of a linear Hamiltonian cycle in $G^{(r)}(n, p)$.*

In the case where $r = 3$, Frieze [8] showed that p_3^{deg} is a semi-sharp threshold for the appearance of a linear Hamiltonian cycle, and Dudek and Frieze [5] showed that p_r^{deg} is an asymptotic threshold for the appearance of a linear Hamiltonian cycle for all $r \geq 3$. Of course, we expect much more than Conjecture 4.2 to be true and naturally expect the hitting time for the appearance of a linear Hamiltonian cycle to coincide with the hitting time for the disappearance of isolated vertices, but even Conjecture 4.2 appears to be out of the reach of existing techniques.

ACKNOWLEDGEMENTS

The first author wishes to acknowledge support from NSF grant DMS-1800521.

REFERENCES

- [1] M. Ajtai, J. Komlós, and E. Szemerédi, *First occurrence of Hamilton cycles in random graphs*, Cycles in graphs (Burnaby, B.C., 1982), North-Holland Math. Stud., vol. 115, North-Holland, Amsterdam, 1985, pp. 173–178. MR821516 ↑1
- [2] B. Bollobás, *The evolution of sparse graphs*, Graph theory and combinatorics (Cambridge, 1983), Academic Press, London, 1984, pp. 35–57. MR777163 ↑1
- [3] B. Bollobás and A. Thomason, *Threshold functions*, Combinatorica **7** (1987), no. 1, 35–38, DOI 10.1007/BF02579198. MR905149 ↑2
- [4] A. Dudek and A. Frieze, *Tight Hamilton cycles in random uniform hypergraphs*, Random Structures Algorithms **42** (2013), no. 3, 374–385, DOI 10.1002/rsa.20404. MR3039684 ↑1, 3
- [5] ———, *Loose Hamilton cycles in random uniform hypergraphs*, Electron. J. Combin. **18** (2011), no. 1, Paper 48, 14. MR2776824 ↑1, 4
- [6] E. Friedgut, *Sharp thresholds of graph properties, and the k -sat problem*, J. Amer. Math. Soc. **12** (1999), no. 4, 1017–1054, DOI 10.1090/S0894-0347-99-00305-7. With an appendix by Jean Bourgain. MR1678031 ↑1, 2
- [7] ———, *Hunting for sharp thresholds*, Random Structures Algorithms **26** (2005), no. 1-2, 37–51, DOI 10.1002/rsa.20042. MR2116574 ↑2
- [8] A. Frieze, *Loose Hamilton cycles in random 3-uniform hypergraphs*, Electron. J. Combin. **17** (2010), no. 1, Note 28, 4. MR2651737 ↑4
- [9] A. Johansson, J. Kahn, and V. Vu, *Factors in random graphs*, Random Structures Algorithms **33** (2008), no. 1, 1–28, DOI 10.1002/rsa.20224. MR2428975 ↑1
- [10] J. Kahn, February 2018. Personal communication. ↑1
- [11] J. Komlós and E. Szemerédi, *Limit distribution for the existence of Hamiltonian cycles in a random graph*, Discrete Math. **43** (1983), no. 1, 55–63, DOI 10.1016/0012-365X(83)90021-3. MR680304 ↑1
- [12] R. Montgomery, *Spanning trees in random graphs*, available at arXiv:1810.03299. Submitted. ↑1
- [13] L. Pósa, *Hamiltonian circuits in random graphs*, Discrete Math. **14** (1976), no. 4, 359–364. MR0389666 ↑1
- [14] O. Riordan, *Spanning subgraphs of random graphs*, Combin. Probab. Comput. **9** (2000), no. 2, 125–148. MR1762785 ↑1

DEPARTMENT OF MATHEMATICS, RUTGERS UNIVERSITY, PISCATAWAY, NJ 08854, USA

Email address: narayanan@math.rutgers.edu

DEPARTMENT OF MATHEMATICS, YALE UNIVERSITY, NEW HAVEN, USA

Email address: mathias.schacht@yale.edu