Received: 19 August 2021 | Accepted: 14 October 2021

DOI: 10.1002/jgt.22794

ARTICLE

WILEY

Counting hypergraphs with large girth

Sam Spiro | **Jacques Verstraëte**

Department of Mathematics, University of California, San Diego, La Jolla, California, USA

Correspondence
Sam Spiro, Department of Mathematics, University of California, San Diego, 9500 Gilman Dr, La Jolla, CA 92093-0112, USA.
Email: sspiro@ucsd.edu

Funding information
ETH Zürich Foundation; National Science Foundation

Abstract

Morris and Saxton used the method of containers to bound the number of n-vertex graphs with m edges containing no ℓ-cycles, and hence graphs of girth more than ℓ. We consider a generalization to r-uniform hypergraphs. The *girth* of a hypergraph H is the minimum $\ell \geq 2$ such that there exist distinct vertices $v_1, \ldots, v_\ell$ and hyperedges $e_1, \ldots, e_\ell$ with $v_i, v_{i+1} \in e_i$ for all $1 \leq i \leq \ell$. Letting $N_m^r(n, \ell)$ denote the number of n-vertex r-uniform hypergraphs with m edges and girth larger than ℓ and defining $\lambda = \lceil (r-2)/(\ell-2) \rceil$, we show

$$N_m^r(n, \ell) \leq N_m^2(n, \ell)^{r-1+\lambda},$$

which is tight when $\ell - 2$ divides $r - 2$ up to a $1 + o(1)$ term in the exponent. This result is used to address the extremal problem for subgraphs of girth more than ℓ in random r-uniform hypergraphs.

KEYWORDS
Berge, cycle, hypergraph

1 | INTRODUCTION

Let $\mathcal{F}$ be a family of r-uniform hypergraphs, or r-graphs for short. Define $N^r(n, \mathcal{F})$ to be the number of $\mathcal{F}$-free r-graphs on $[n] := \{1, \ldots, n\}$, and define $N_m^r(n, \mathcal{F})$ to be the number of $\mathcal{F}$-free r-graphs on $[n]$ with exactly m hyperedges. If $\mathrm{ex}(n, \mathcal{F})$ denotes the maximum number of hyperedges in an $\mathcal{F}$-free r-graph on $[n]$, then it is not difficult to see that for $1 \leq m \leq \mathrm{ex}(n, \mathcal{F})$,

$$\left(\frac{\mathrm{ex}(n,\mathcal{F})}{m}\right)^m \le \binom{\mathrm{ex}(n,\mathcal{F})}{m} \le N_m^r(n,\mathcal{F}) \le \binom{\binom{n}{r}}{m} \le \left(\frac{en^r}{m}\right)^m,$$

and summing over m one obtains $2^{\Omega(\mathrm{ex}(n,\mathcal{F}))} = N^r(n,\mathcal{F}) = 2^{O(\mathrm{ex}(n,\mathcal{F})\log n)}$. The state-of-the-art for bounding $N^r(n,\mathcal{F})$ is the work of Ferber, McKinley, and Samotij [9] which shows that if F is an r-uniform hypergraph with $\mathrm{ex}(n,F) = O(n^\alpha)$ and α not too small, then

$$N^r(n,F) = 2^{O(n^\alpha)},$$

and this result encompasses many of the earlier results in the area [3,4,6,17].

There are relatively few families for which effective bounds for $N_m^r(n,\mathcal{F})$ are known. One family where results are known is $\mathcal{C}_{[\ell]} = \{C_3, C_4, \ldots, C_\ell\}$, the family of all graph cycles of length at most ℓ. Morris and Saxton implicitly proved the following in this setting:

Theorem 1.1 (Morris and Saxton [17]). *For $\ell \ge 3$ and $k = \lfloor \ell/2 \rfloor$, there exists a constant $c = c(\ell) > 0$ such that if n is sufficiently large and $m \ge n^{1+1/(2k-1)}(\log n)^2$, then*

$$N_m^2(n,\mathcal{C}_{[\ell]}) \le e^{cm}(\log n)^{(k-1)m}\left(\frac{n^{1+1/k}}{m}\right)^{km}.$$

In the appendix we give a formal proof of this result. Theorem 1.1 generalizes earlier results of Füredi [11] when $\ell = 4$ and of Kohayakawa, Kreuter, and Steger [15]. Erdős and Simonovits [8] conjectured for $\ell \ge 3$ and $k = \lfloor \ell/2 \rfloor$,

$$\mathrm{ex}(n,\mathcal{C}_{[\ell]}) = \Omega(n^{1+1/k}) \tag{1}$$

which is only known to hold for $\ell \in \{3,4,5,6,7,10,11\}$—see Füredi and Simonovits [12] and also [24] for details. The truth of this conjecture would imply that the upper bound in Theorem 1.1 is tight up to the exponent of $(\log n)^m$.

In this paper we extend Theorem 1.1 to r-graphs. For $\ell \ge 2$, an r-graph F is a *Berge ℓ-cycle* if there exist distinct vertices $v_1, \ldots, v_\ell$ and distinct hyperedges $e_1, \ldots, e_\ell$ with $v_i, v_{i+1} \in e_i$ for all $1 \le i \le \ell$. In particular, a hypergraph H is said to be *linear* if it contains no Berge 2-cycle. We denote by $\mathcal{C}_\ell^r$ the family of all r-uniform Berge ℓ-cycles. If H is an r-graph containing a Berge cycle, then the *girth* of H is the smallest $\ell \ge 2$ such that H contains a Berge ℓ-cycle. Let $\mathcal{C}_{[\ell]}^r = \mathcal{C}_2^r \cup \mathcal{C}_3^r \cup \cdots \cup \mathcal{C}_\ell^r$ denote the family of all r-uniform Berge cycles of length at most ℓ. With this $\mathcal{C}_{[\ell]}^2 = \mathcal{C}_{[\ell]}$, and an r-graph has girth larger than ℓ if and only if it is $\mathcal{C}_{[\ell]}^r$-free. We again emphasize that hypergraphs with girth $\ell \ge 2$ are all linear. We write $N_m^r(n,\ell) := N_m^r(n,\mathcal{C}_{[\ell]}^r)$ for the number of n-vertex r-graphs with m edges and girth larger than ℓ and $N^r(n,\ell) := N^r(n,\mathcal{C}_{[\ell]}^r)$ for the number of n-vertex r-graphs with girth larger than ℓ.

Balogh and Li [2] proved for all $\ell, r \ge 3$ and $k = \lfloor \ell/2 \rfloor$,

$$N^r(n,\ell) = 2^{O(n^{1+1/k})}.$$

This upper bound would be tight up to an $n^{o(1)}$ term in the exponent if the following is true:

Conjecture 1. *For all $\ell \geq 3$ and $r \geq 2$ and $k = \lfloor \ell/2 \rfloor$,*

$$\mathrm{ex}\left(n, \mathcal{C}^r_{[\ell]}\right) = n^{1+1/k-o(1)}.$$

Conjecture 1 holds for $\ell = 3, 4$ and $r \geq 3$—see [7,16,22,23]—but is open and evidently difficult for $\ell \geq 5$ and $r \geq 3$. Györi and Lemons [13] proved $\mathrm{ex}(n, \mathcal{C}^r_\ell) = O(n^{1+1/k})$ with $k = \lfloor \ell/2 \rfloor$, so the conjecture concerns constructions of dense r-graphs of girth more than ℓ. The conjecture for $r = 2$ without the $o(1)$ is (1), and for each $r \geq 3$ is stronger than (1), as can be seen by forming a graph from an extremal n-vertex r-graph of girth more than ℓ whose edge set consists of an arbitrary pair of vertices from each hyperedge. We emphasize that the $o(1)$ term in Conjecture 1 is necessary for $\ell = 3$, due to the Ruzsa–Szemerédi theorem [7,22], and for $\ell = 5$, due to the work of Conlon, Fox, Sudakov, and Zhao [5].

1.1 | Counting r-graphs of large girth

In this study we simplify and refine the arguments of Balogh and Li [2] to prove effective and almost tight bounds on $N^r_m(n, \ell)$ relative to $N^2_m(n, \ell)$.

Theorem 1.2. *Let $\ell, r \geq 3$ and $\lambda = \lceil (r-2)/(\ell-2) \rceil$. Then for all $m, n \geq 1$,*

$$N^r_m(n, \ell) \leq N^2_m(n, \ell)^{r-1+\lambda}. \tag{2}$$

We note that (2) corrects a bound[1] which appears in [20]. The inequality (2) is essentially tight when $\ell - 2$ divides $r - 2$, due to standard probabilistic arguments (see, e.g., Janson, Łuczak, and Rucinski [14]): it is possible to show that when $m \leq n^{1+1/(\ell-1)}$, the uniform model of random n-vertex r-graphs with m edges has girth larger than ℓ with probability at least a^{-m} for some constant $a > 1$ depending only on ℓ and r. In particular, there exist some constants $b, c > 1$ such that for $m \leq n^{1+1/(\ell-1)}$ we have

$$N^r_m(n, \ell) \geq a^{-m}\binom{\binom{n}{r}}{m} \geq b^{-m}(n^r/m)^m \geq b^{-m}(n^2/m)^{\left(r-1+\frac{r-2}{\ell-2}\right)m} \geq c^{-m} \cdot N^2_m(n, \ell)^{r-1+\frac{r-2}{\ell-2}}, \tag{3}$$

where the third inequality used $m \leq n^{1+1/(\ell-1)}$ and the last inequality used the trivial bound $N^2_m(n, \ell) \leq (en^2/m)^m$. This shows that the bound of Theorem 1.2 is best possible when $\ell - 2$ divides $r - 2$ up to a multiplicative error of c^{-m} for some constant $c > 1$. We believe that (3) should define the optimal exponent, and propose the following conjecture:

Conjecture 2. *For all $r \geq 2$, $\ell \geq 3$ and $m, n \geq 1$,*

[1]Theorem 20 of [20] claims a stronger upper bound for $N^r_m(n, 4)$ than what we prove in Theorem 1.2, but we have confirmed with the authors that there was a subtle error in their proof.

$$N_m^r(n,\ell) \le N_m^2(n,\ell)^{r-1+\frac{r-2}{\ell-2}}.$$

Theorem 1.2 shows that this conjecture is true when $\ell - 2$ divides $r - 2$, so the first open case of Conjecture 2 is when $\ell = 4$ and $r = 3$.

In the case that Berge ℓ-cycles are forbidden instead of all Berge cycles of length at most ℓ, we can prove an analog of Theorem 1.2 with weaker quantitative bounds. To this end, let $N_{[m]}^r(n,\mathcal{F})$ denote the number of n-vertex $\mathcal{F}$-free r-graphs on at most m hyperedges.

Theorem 1.3. *For each $\ell, r \ge 3$, there exists $c = c(\ell, r)$ such that*

$$N_m^r\left(n, C_\ell^r\right) \le 2^{cm} \cdot N_{[m]}^2(n, C_\ell)^{r!/2}.$$

We suspect that this result continues to hold with $N_{[m]}^2(n, C_\ell)$ replaced by $N_m^2(n, C_\ell)$.

1.2 | Subgraphs of random r-graphs of large girth

Denote by $H_{n,p}^r$ the r-graph obtained by including each hyperedge of K_n^r independently and with probability p. Given a family of r-graphs $\mathcal{F}$, let $\mathrm{ex}(H_{n,p}^r, \mathcal{F})$ denote the size of a largest $\mathcal{F}$-free subgraph of $H_{n,p}^r$. Recall that a statement depending on n holds *asymptotically almost surely* or a.a.s. if it holds with probability tending to 1 as $n \to \infty$. A hypergraph of girth at least three is a linear hypergraph, and it is not hard to show by a simple first moment calculation that if $p \ge n^{-r}\log n$, then a.a.s.

$$\mathrm{ex}\left(H_{n,p}^r, C_{[2]}^r\right) = \Theta(\min\{pn^r, n^2\}).$$

Our first result essentially determines the a.a.s. behavior of the number of edges in an extremal subgraph of $H_{n,p}^r$ of girth four. In this theorem we omit the case $p < n^{-r+\frac{3}{2}}$, as it is straightforward to show that a.a.s. $\mathrm{ex}(H_{n,p}^r, C_{[3]}^r) = \Theta(pn^r)$ when $p \ge n^{-r}\log n$ in this range.

Theorem 1.4. *Let $r \ge 3$. If $p \ge n^{-r+\frac{3}{2}}(\log n)^{2r-3}$, then a.a.s.*

$$p^{\frac{1}{2r-3}}n^{2-o(1)} \le \mathrm{ex}\left(H_{n,p}^r, C_{[3]}^r\right) \le p^{\frac{1}{2r-3}}n^{2+o(1)}.$$

Due to Theorems 1.2 and 1.4, the number of linear triangle-free r-graphs with n vertices and m edges where $n^{3/2+o(1)} \le m \le \mathrm{ex}(n, C_{[3]}^r) = o(n^2)$ and $r \ge 3$ is

$$N_m^r(n,3) = N_m^2(n,3)^{2r-3+o(1)} = \left(\frac{n^2}{m}\right)^{(2r-3)m+o(m)}.$$

The authors and Nie et al. [19] obtained bounds for r-uniform loose triangles,[2] where for $r = 3$ the same essentially tight bounds as in Theorem 1.4 were obtained, but for $r > 3$

[2]The loose triangle is the Berge triangle whose edges pairwise intersect in exactly one vertex.

there remains a significant gap. In the case of subgraphs of girth larger than four, Theorem 1.2 allows us to generalize results of Morris and Saxton [17] and earlier results of Kohayakawa, Kreuter, and Steger [15] giving subgraphs of large girth in random graphs in the following way:

Theorem 1.5. *Let $\ell \geq 4$ and $r \geq 2$, and let $k = \lfloor \ell/2 \rfloor$ and $\lambda = \lceil (r-2)/(\ell-2) \rceil$. Then a.a.s.*

$$\mathrm{ex}\left(H^r_{n,p}, C^r_{[\ell]}\right) \leq \begin{cases} n^{1+\frac{1}{\ell-1}+o(1)} & n^{-r+1+\frac{1}{\ell-1}} \leq p < n^{\frac{-(r-1+\lambda)(k-1)}{2k-1}} (\log n)^{(r-1+\lambda)k}, \\ p^{\frac{1}{(r-1+\lambda)k}} n^{1+\frac{1}{k}+o(1)} & n^{\frac{-(r-1+\lambda)(\ell-1-k)}{\ell-1}} (\log n)^{(r-1+\lambda)k} \leq p \leq 1. \end{cases}$$

If Conjecture 1 is true, then

$$\mathrm{ex}\left(H^r_{n,p}, C^r_{[\ell]}\right) \geq \begin{cases} n^{1+\frac{1}{\ell-1}+o(1)} & n^{-r+1+\frac{1}{\ell-1}} \leq p < n^{\frac{-(r-1)(\ell-1-k)}{\ell-1}}, \\ p^{\frac{1}{(r-1)k}} n^{1+\frac{1}{k}-o(1)} & n^{\frac{-(r-1)(\ell-1-k)}{\ell-1}} \leq p \leq 1. \end{cases}$$

We emphasize that there is a significant gap in the bounds of Theorem 1.5 due to the presence of λ in the exponent of p in the upper bound and its absence in the lower bound, and this gap is closed by Theorem 1.4 when $\ell = 3$ by an improvement to the lower bound. A similar phenomenon appears in the recent work of Mubayi and Yepremyan [18], who determined the a.a.s. value of the extremal function for loose even cycles in $H^r_{n,p}$ for all but a small range of p. It seems likely that the following conjecture is true:

Conjecture 3. *Let $\ell, r \geq 3$ and $k = \lfloor \ell/2 \rfloor$. Then there exists $\gamma = \gamma(\ell, r)$ such that a.a.s.*

$$\mathrm{ex}\left(H^r_{n,p}, C^r_{[\ell]}\right) = \begin{cases} n^{1+\frac{1}{\ell-1}+o(1)} & n^{-r+1+\frac{1}{\ell-1}} \leq p < n^{-\frac{\gamma(\ell-1-k)}{\ell-1}}, \\ p^{\frac{1}{\gamma k}} n^{1+\frac{1}{k}+o(1)} & n^{-\frac{\gamma(\ell-1-k)}{\ell-1}} \leq p \leq 1. \end{cases}$$

Conjecture 2 suggests the possible value $\gamma(\ell, r) = r - 1 + (r-2)/(\ell-2)$, which is the correct value for $\ell = 3$ by Theorem 1.4. We are not certain that this is the right value of γ in general, even when $r = 3$ and $\ell = 4$, and more generally, Conjecture 1 is an obstacle for $r \geq 3$ and $\ell \geq 5$. Theorem 1.5 shows that if γ exists, then $(r-1)k \leq \gamma \leq (r-1+\lambda)k$ provided Conjecture 1 holds.

Letting $f(n,p) = \mathrm{ex}(H^3_{n,p}, C^3_{[4]})$, we plot the bounds of Theorem 1.5 in Figure 1, where the upper bound is in blue and the lower bound is in green. The truth of Conjecture 2 for $\ell = 4$ would imply the slightly better upper bound $f(n,p) \leq p^{1/5} n^{3/2+o(1)}$.

Notation: A set of size k will be called a *k-set*. As much as possible, when working with a k-graph G and an r-graph H with $k < r$, we will refer to elements of $E(G)$ as edges and elements of $E(H)$ as hyperedges. Given a hypergraph H on $[n]$, we define the *k-shadow* $\partial^k H$ to be the k-graph on $[n]$ consisting of all k-sets e which lie in a hyperedge of $E(H)$. If $G_1, \ldots, G_q$ are k-graphs on $[n]$, then $\bigcup G_i$ denotes the k-graph G on $[n]$ which has edge set $\bigcup E(G_i)$.

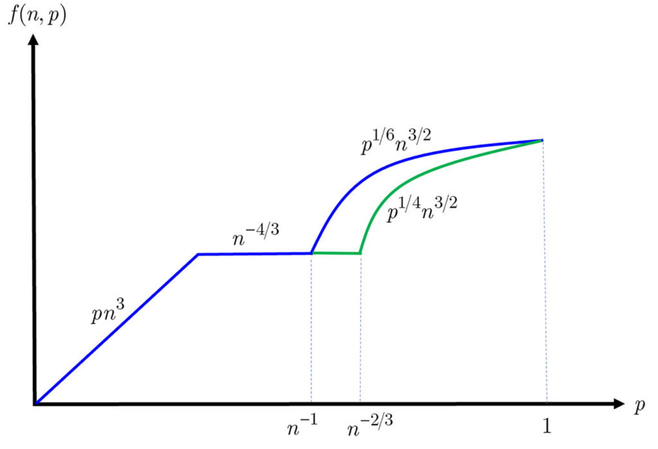

FIGURE 1 Subgraphs of $H^3_{n,p}$ of girth five

2 | PROOF OF THEOREM 1.2

As Balogh and Li [2] observed, if $\ell \geq 3$ and H has girth larger than ℓ, then H is uniquely determined by $\partial^2 H$, which we can view as the graph obtained by replacing each hyperedge of H by a clique. A key insight in proving Theorem 1.2 is that we can replace each hyperedge of H with a sparser graph B and still uniquely recover H from this graph. To this end, we say that a graph B is a *book* if there exist cycles $F_1, \ldots, F_k$ and an edge xy such that $B = \bigcup F_i$ and $E(F_i) \cap E(F_j) = \{xy\}$ for all $i \neq j$. In this case we call the cycles F_i the *pages* of B and we call the common edge xy the *spine* of B. The following lemma shows that if we replace each hyperedge in H by a book on r vertices which has small pages, then the vertex sets of books in the resulting graph are exactly the hyperedges of H.

Lemma 2.1. *Let H be an r-graph of girth larger than ℓ. If $\partial^2 H$ contains a book B on r vertices such that every page has length at most ℓ, then there exists a hyperedge $e \in E(H)$ such that $V(B) = e$.*

Proof. Let F be a cycle in $\partial^2 H$ with $V(F) = \{v_1, \ldots, v_p\}$ such that $v_i v_{i+1} \in E(\partial^2 H)$ for $i < p$ and $v_1 v_p \in E(\partial^2 H)$. If $p \leq \ell$ we claim that there exists an $e \in E(H)$ such that $V(F) \subseteq e$. Indeed, by definition of $\partial^2 H$ there exists some hyperedge $e_i \in E(H)$ with $v_i, v_{i+1} \in e_i$ for all $i < p$ and some hyperedge e_p with $v_1, v_p \in e_p$. If all of these e_i hyperedges are equal then we are done, so we may assume $e_1 \neq e_p$. Define i_1 to be the largest index such that $e_i = e_1$ for all $i \leq i_1$, define i_2 to be the largest index so that $e_i = e_{i_1+1}$ for all $i_1 < i \leq i_2$, and so on up to $i_q = p$, and note that $2 \leq q \leq p$ since $e_1 \neq e_p$. If all the e_{i_j} hyperedges are distinct, then they form a Berge q-cycle in H since $v_{1+i_j} \in e_{i_j} \cap e_{1+i_j} = e_{i_j} \cap e_{i_{j+1}}$ for all j, a contradiction. Thus we can assume $e_{i_j} = e_{i_{j'}}$ for some $j < j'$. We can further assume that $e_{i_s} \neq e_{i_{s'}}$ for any $j \leq s < s' < j'$, as otherwise we could replace j, j' with s, s'. Finally note that $j < j' - 1$, as otherwise we would have $e_{i_j} = e_{i_{j'}} = e_{i_j+1}$, contradicting the maximality of i_j. We conclude that the distinct hyperedges $e_{i_j}, e_{i_{j+1}}, \ldots, e_{i_{j'-1}}$ form a Berge $(j' - j)$-cycle with $2 \leq j' - j \leq \ell$ in H, a contradiction. This proves the claim.

Now let B be a book with spine xy and pages $F_1, \ldots, F_k$ of length at most ℓ. By the claim there exist hyperedges $e_1, \ldots, e_k \in E(H)$ such that $V(F_i) \subseteq e_i$ for all i, and in particular

$x, y \in e_i$ for all i. Because H is linear, this implies that all of these hyperedges are equal and we have $V(B) \subseteq e_1$. If B has r vertices, then we further have $V(B) = e_1$. □

We now complete the proof of Theorem 1.2. With $\lambda := \lceil (r-2)/(\ell-2) \rceil$ we observe for all $\ell, r \geq 3$ that there exists a book graph B on r vertices $\{x_1, ..., x_r\}$ with $r-1+\lambda$ edges $f_1, ..., f_{r-1+\lambda}$. Indeed if $\ell - 2$ divides $r-2$ one can take λ copies of C_ℓ which share a common edge, and otherwise one can take $\lambda - 1$ copies of C_ℓ and a copy of C_p with $p = r - (\lambda-1)(\ell-2) \geq 3$. From now on we let B denote this book graph. If $f_i = \{x_j, x_{j'}\} \in E(B)$ and $e = \{v_1, ..., v_r\} \subseteq [n]$ is any r-set with $v_1 < \cdots < v_r$, define $\phi_i(e) = \{v_j, v_{j'}\}$. If H is an r-graph on $[n]$, define $\phi_i(H)$ to be the graph on $[n]$ which has all edges of the form $\phi_i(e)$ for $e \in E(H)$; so in particular $\bigcup \phi_i(H)$ is the graph obtained by replacing each hyperedge of H with a copy of B.

Let $\mathcal{H}_{m,n}$ denote the set of r-graphs on $[n]$ with m hyperedges and girth more than ℓ, and let $\mathcal{G}_{m,n}$ be the set of graphs on $[n]$ with m edges and girth more than ℓ. We claim that ϕ_i maps $\mathcal{H}_{m,n}$ to $\mathcal{G}_{m,n}$. Indeed, if $H \in \mathcal{H}_{m,n}$ then each hyperedge of H contributes a distinct edge to $\phi_i(H)$ since H is linear, so $e(\phi_i(H)) = e(H) = m$. One can show that if $\phi_i(e_1), ..., \phi_i(e_p)$ form a p-cycle in $\phi_i(H)$, then $e_1, ..., e_p$ form a Berge p-cycle in H; so $H \in \mathcal{H}_{m,n}$ implies $\phi_i(H)$ does not contain a cycle of length at most ℓ.

Let $\mathcal{G}_{m,n}^t = \{(G_1, G_2, ..., G_t) : G_i \in \mathcal{G}_{m,n}\}$. Then we define a map $\phi : \mathcal{H}_{m,n} \to \mathcal{G}_{m,n}^{r-1+\lambda}$ by

$$\phi(H) = (\phi_1(H), ..., \phi_{r-1+\lambda}(H)).$$

We claim that this map is injective. Indeed, fix some $H \in \mathcal{H}_{m,n}$ and let $\mathcal{B}(G)$ denote the set of books B in the graph $G := \bigcup \phi_i(H) \subseteq \partial^2 H$. By definition of ϕ we have $E(H) \subseteq \mathcal{B}(G)$ for all H. Moreover, if $H \in \mathcal{H}_{m,n}$ then Lemma 2.1 implies $\mathcal{B}(G) \subseteq E(H)$. Thus $E(H)$ (and hence H) is uniquely determined by G, which is itself determined by $\phi(H)$, so the map is injective. In total we conclude

$$N_m^r(n, \ell) = |\mathcal{H}_{m,n}| \leq |\mathcal{G}_{m,n}^{r-1+\lambda}| = N_m^2(n, \ell)^{r-1+\lambda},$$

proving Theorem 1.2. □

3 | PROOF OF THEOREM 1.3

For arbitrary hypergraphs H, the map $\phi(H) = \partial^{r-1}H$ (let alone the map to $\partial^2 H$) is not injective. However, we will show that this map is "almost" injective when considering H which are $\mathcal{C}_\ell^r$-free. To this end, we say that a set of vertices $\{v_1, ..., v_r\}$ is a *core set* of an r-graph H if there exist distinct hyperedges $e_1, ..., e_r$ with $\{v_1, ..., v_r\} \backslash \{v_i\} \subseteq e_i$ for all i. The following observation shows that core sets are the only obstruction to $\phi(H) = \partial^{r-1}H$ being injective.

Lemma 3.1. *Let H be an r-graph. If $\{v_1, ..., v_r\}$ induces a K_r^{r-1} in $\partial^{r-1}H$, then either $\{v_1, ..., v_r\} \in E(H)$ or $\{v_1, ..., v_r\}$ is a core set of H.*

Proof. By assumption of $\{v_1, ..., v_r\}$ inducing a K_r^{r-1} in $\partial^{r-1}H$, for all i there exist $e_i' \in E(\partial^{r-1}H)$ with $e_i' = \{v_1, ..., v_r\} \backslash \{v_i\}$. By definition of $\partial^{r-1}H$, this means there exist (not necessarily distinct) $e_i \in E(H)$ with $e_i \supseteq e_i' = \{v_1, ..., v_r\} \backslash \{v_i\}$. Given this, either

$e_i = \{v_1, ..., v_r\}$ for some i, or all of the e_i distinct, in which case $\{v_1, ..., v_r\}$ is a core set of H. In either case we conclude the result. □

We next show that C_ℓ^r-free r-graphs have few core sets.

Lemma 3.2. *Let $\ell, r \geq 3$ and let H be a C_ℓ^r-free r-graph with m hyperedges. The number of core sets in H is at most $\ell^2 r^2 m$.*

Proof. We claim that H contains no core sets if $\ell \leq r$. Indeed, assume for contradiction that H contained a core set $\{v_1, ..., v_r\}$ with distinct hyperedges $e_i \supseteq \{v_1, ..., v_r\} \backslash \{v_i\}$. It is not difficult to see that the hyperedges $e_1, ..., e_\ell$ form a Berge ℓ-cycle, a contradiction to H being C_ℓ^r-free. Thus from now on we may assume $\ell > r$.

Let $\mathcal{A}_1$ denote the set of core sets in H, and for any $\mathcal{A}' \subseteq \mathcal{A}_1$ and $(r-1)$-set S, define $d_{\mathcal{A}'}(S)$ to be the number of core sets $A \in \mathcal{A}'$ with $S \subseteq A$. Observe that $d_{\mathcal{A}_1}(S) > 0$ for at most $\binom{r}{r-1}m = rm(r-1)$-sets S, since in particular S must be contained in a hyperedge of H.

Given $\mathcal{A}_i$, define $\mathcal{A}_i' \subseteq \mathcal{A}_i$ to be the core sets $A \in \mathcal{A}_i$ which contain an $(r-1)$-set S with $d_{\mathcal{A}_i}(S) \leq \ell r$, and let $\mathcal{A}_{i+1} = \mathcal{A}_i \backslash \mathcal{A}_i'$. Observe that $|\mathcal{A}_i'| \leq \ell r \cdot rm$ since each $(r-1)$-set S with $d_{\mathcal{A}_i}(S) > 0$ is contained in at most ℓr elements of $\mathcal{A}_i'$. In particular,

$$|\mathcal{A}_1| \leq (\ell - r) \cdot \ell r^2 m + |\mathcal{A}_{\ell-r+1}| \leq \ell^2 r^2 m + |\mathcal{A}_{\ell-r+1}|. \tag{4}$$

Assume for the sake of contradiction that $\mathcal{A}_{\ell-r+1} \neq \emptyset$. We prove by induction on $r \leq i \leq \ell$ that one can find distinct vertices $v_1, ..., v_i$ and distinct hyperedges $e_1, ..., e_{i-1}, \tilde{e}_i$ such that $v_j, v_{j+1} \in e_j$ for $1 \leq j < i$ and $v_1, v_i \in \tilde{e}_i$, and such that $\{v_i, v_{i-1}, ..., v_{i-r+2}, v_1\} \in \mathcal{A}_{\ell-i+1}$. For the base case, consider any $\{v_r, v_{r-1}, ..., v_1\} \in \mathcal{A}_{\ell-r+1}$. As this is a core set, there exist distinct hyperedges $e_j \supseteq \{v_1, ..., v_r\} \backslash \{v_{j+2}\}$ and $\tilde{e}_r \supseteq \{v_1, ..., v_r\} \backslash \{v_2\}$, proving the base case of the induction.

Assume that we have proven the result for $i < \ell$. By assumption of $\{v_i, v_{i-1}, ..., v_{i-r+2}, v_1\} \in \mathcal{A}_{\ell-i+1}$, we have $\{v_i, v_{i-1}, ..., v_{i-r+2}, v_1\} \notin \mathcal{A}_{\ell-i}'$, so there exists a set of vertices $\{u_1, ..., u_{\ell r+1}\}$ such that $\{v_i, v_{i-1}, ..., v_{i-r+3}, v_1, u_j\} \in \mathcal{A}_{\ell-i}$ for all j. Because $|\bigcup_{k=1}^{i-1} e_k| \leq \ell r$, there exists some j such that $u_j \notin \bigcup_{k=1}^{i-1} e_k$. For this j, let $v_{i+1} := u_j$ and let $e_i, \tilde{e}_{i+1}$ be distinct hyperedges containing v_i, v_{i+1} and v_1, v_{i+1}, respectively, which exist by the assumption of this being a core set. Note that v_{i+1} is distinct from every other $v_{i'}$ since $v_{i'} \in \bigcup_{k=1}^{i-1} e_k$ for $i' \leq i$, and similarly the hyperedges $e_i, \tilde{e}_{i+1}$ are distinct from every hyperedge $e_{i'}$ with $i' < i$ since these new hyperedges contain $v_{i+1} \notin \bigcup_{k=1}^{i-1} e_k$. This proves the inductive step and hence the claim. The $i = \ell$ case of this claim implies that H contains a Berge ℓ-cycle, a contradiction. Thus $\mathcal{A}_{\ell-r+1} = \emptyset$, and the result follows by (4). □

Combining these two lemmas gives the following result, which allows us to reduce from r-graphs to $(r-1)$-graphs. We recall that $N_{[m]}^r(n, \mathcal{F})$ denotes the number of n-vertex $\mathcal{F}$-free r-graphs on at most m hyperedges.

Proposition 3.3. *For each $\ell, r \geq 3$, there exists $c = c(\ell, r)$ such that*

$$N^r_{[m]}\left(n, \mathcal{C}^r_\ell\right) \le 2^{cm} \cdot N^r_{[m]}\left(n, \mathcal{C}^{r-1}_\ell\right)^r.$$

Proof. If $e = \{v_1, v_2, \ldots, v_r\} \subseteq [n]$ is any r-set with $v_1 < v_2 < \cdots < v_r$, let $\phi_i(e) = \{v_1, \ldots, v_r\} \setminus \{v_i\}$. Given an r-graph H on $[n]$, let $\phi_i(H)$ be the $(r-1)$-graph on $[n]$ with edge set $\{\phi_i(e) : e \in E(H)\}$, and define $\phi(H) = (\phi_1(H), \phi_2(H), \ldots, \phi_r(H))$ and $\psi(H) = (\phi(H), E(H))$. Observe that $\bigcup \phi_i(H) = \partial^{r-1} H$. Let $\mathcal{H}_{[m],n}$ denote the set of all r-graphs on $[n]$ with at most m hyperedges which are $\mathcal{C}^r_\ell$-free, and let $\phi(\mathcal{H}_{[m],n})$, $\psi(\mathcal{H}_{[m],n})$ denote the image sets of $\mathcal{H}_{[m],n}$ under these respective maps. Observe that ψ is injective since it records $E(H)$, so it suffices to bound how large $\psi(\mathcal{H}_{[m],n})$ can be.

Let $\mathcal{G}_{[m],n}$ denote the set of $(r-1)$-graphs on $[n]$ which have at most m edges and which are $\mathcal{C}^{r-1}_\ell$-free. It is not difficult to see that $\phi(\mathcal{H}_{[m],n}) \subseteq \mathcal{G}^r_{[m],n}$. We observe by Lemmas 3.1 and 3.2 that for any $(G_1, G_2, \ldots, G_r) \in \phi(\mathcal{H}_{[m],n})$, say with $\phi(H) = (G_1, \ldots, G_r)$, there are at most $(1 + \ell^2 r^2)m$ copies of K^{r-1}_r in $\bigcup G_i = \partial^{r-1} H$. We also observe that if $((G_1, G_2, \ldots, G_r), E) \in \psi(\mathcal{H}_{[m],n})$, then E is a set of at most m copies of K^{r-1}_r in $\bigcup G_i$. Thus given any $(G_1, \ldots, G_r) \in \phi(\mathcal{H}_{[m],n}) \subseteq \mathcal{G}^r_{[m],n}$, there are at most $2^{(1+\ell^2 r^2)m}$ choices of E such that $((G_1, \ldots, G_r), E) \in \psi(\mathcal{H}_{[m],n})$. We conclude that

$$N_{[m]}\left(n, \mathcal{C}^r_\ell\right) = |\mathcal{H}_{[m],n}| \le |\mathcal{G}_{[m],n}|^r \cdot 2^{(1+\ell^2 r^2)m} = N^r_{[m]}\left(n, \mathcal{C}^{r-1}_\ell\right)^r \cdot 2^{(1+\ell^2 r^2)m},$$

proving the result. □

Applying this proposition repeatedly gives $N^r_{[m]}(n, \mathcal{C}^r_\ell) \le 2^{cm} N^2_{[m]}(n, \mathcal{C}_\ell)^{r!/2}$. Combining this with the trivial inequality $N^r_m(n, \mathcal{C}^r_\ell) \le N^r_{[m]}(n, \mathcal{C}^r_\ell)$ gives Theorem 1.3.

4 | PROOF OF THEOREMS 1.4 AND 1.5

To prove that our bounds hold a.a.s., we use the Chernoff bound [1].

Proposition 4.1 (Alon and Spencer [1]). *Let X denote a binomial random variable with N trials and probability p of success. For any $\epsilon > 0$ we have* $\Pr[|X - pN| > \epsilon pN] \le 2\exp(-\epsilon^2 pN/2)$.

Proof of the upper bounds in Theorem 1.5. Let

$$p_0 = n^{-\frac{(r-1+\lambda)(k-1)}{2k-1}} (\log n)^{(r-1+\lambda)k}.$$

For $p \ge p_0$, define

$$m = p^{\frac{1}{(r-1+\lambda)k}} n^{1+\frac{1}{k}} \log n,$$

and note that this is large enough to apply Theorem 1.1 for $p \ge p_0$. Let Y_m denote the number of subgraphs of $H^r_{n,p}$ which are $\mathcal{C}^r_{[\ell]}$-free and have exactly m edges, and note that

$\mathrm{ex}(H^r_{n,p}, \mathcal{C}^r_{[\ell]}) \geq m$ if and only if $Y_m \geq 1$. By Markov's inequality, Theorem 1.2, and Theorem 1.1:

$$\begin{aligned}\Pr[Y_m \geq 1] &\leq \mathbb{E}[Y_m] = p^m \cdot N^r_m(n, \ell) \\ &\leq p^m \cdot N^2_m(n, \ell)^{r-1+\lambda} \\ &\leq \left(p^{\frac{1}{r-1+\lambda}} e^c (\log n)^{k-1} \left(\frac{n^{1+\frac{1}{k}}}{m} \right)^k \right)^{m(r-1+\lambda)} \\ &= \left(\frac{e^c}{\log n} \right)^{m(r-1+\lambda)}.\end{aligned}$$

The right-hand side converges to zero, so for $p \geq p_0$, a.a.s.

$$\mathrm{ex}\left(H^r_{n,p}, \mathcal{C}^r_{[\ell]}\right) < m.$$

As $\mathbb{E}[\mathrm{ex}(H^r_{n,p}, \mathcal{C}^r_{[\ell]})]$ is nondecreasing in p, the bound

$$\mathrm{ex}\left(H^r_{n,p}, \mathcal{C}^r_{[\ell]}\right) < n^{1+\frac{1}{\ell-1}} (\log n)^2$$

continues to hold a.a.s. for all $p < p_0$. □

Proof of the upper bound in Theorem 1.4. This proof is almost identical to the previous, so we omit some of the redundant details. Let $m = p^{\frac{1}{2r-3}} n^2 \log n$ and let Y_m denote the number of subgraphs of $H^r_{n,p}$ which are $\mathcal{C}^r_{[\ell]}$-free and have exactly m edges. By Markov's inequality, Theorem 1.2, and the trivial bound $N^2_m(n, 3) \leq \binom{n^2}{m}$ which is valid for all m, we find for all p

$$\Pr[Y_m \geq 1] \leq p^m (en^2/m)^{(2r-3)m} = (e/\log n)^m.$$

This quantity converges to zero, so we conclude the result by the same reasoning as in the previous proof. □

This proof shows that for all p we have $\mathbb{E}[\mathrm{ex}(H^r_{n,p}, \mathcal{C}^r_{[\ell]})] < p^{\frac{1}{2r-3}} n^2 \log n$. However, for $p \leq n^{-r+3/2}$ this is weaker than the trivial upper bound $\mathbb{E}[\mathrm{ex}(H^r_{n,p}, \mathcal{C}^r_{[\ell]})] \leq p\binom{n}{r}$.

Proof of the lower bounds in Theorem 1.5. We use homomorphisms similar to Foucaud, Krivelevich, and Perarnau [10] and Perarnau and Reed [21]. If F and F' are hypergraphs and $\chi : V(F) \to V(F')$ is any map, we let $\chi(e) = \{\chi(u) : u \in e\}$ for any $e \in E(F)$. For two r-graphs F and F', a map $\chi : V(F) \to V(F')$ is a *homomorphism* if $\chi(e) \in E(F')$ for all $e \in E(F)$, and χ is a *local isomorphism* if χ is a homomorphism and $\chi(e) \neq \chi(f)$ whenever $e, f \in E(F)$ with $e \cap f \neq \emptyset$. A key lemma is the following: □

Lemma 4.2. *If $F \in \mathcal{C}^r_{[\ell]}$ and $\chi : F \to F'$ is a local isomorphism, then F' has girth at most ℓ.*

Proof. Let F be a Berge p-cycle with $p \leq \ell$ and $E(F) = \{e_1, e_2, ..., e_p\}$. Then there exist distinct vertices $v_1, v_2, ..., v_p$ such that $v_i \in e_i \cap e_{i+1}$ for $i < p$ and $v_p \in e_p \cap e_1$. First assume there exists $i \neq j$ such that $\chi(e_i) = \chi(e_j)$. By reindexing, we can assume $\chi(e_1) = \chi(e_k)$ for some $k > 1$, and further that $\chi(e_i) \neq \chi(e_j)$ for any $1 \leq i < j < k$. Note that $k \geq 3$ since $e_1 \cap e_2 \neq \emptyset$ and χ is a local isomorphism. If we also have $\chi(v_i) \neq \chi(v_j)$ for all $1 \leq i < j < k$, then $\chi(v_i) \in \chi(e_i) \cap \chi(e_{i+1})$ for $i < k - 1$ and $\chi(v_{k-1}) \in \chi(e_{k-1}) \cap \chi(e_1)$, so $\chi(e_1), \chi(e_2), ..., \chi(e_{k-1})$ is the edge set of a Berge $(k-1)$-cycle in F' as required.

Suppose $\chi(v_i) = \chi(v_j)$ for some $1 \leq i < j < k$, and as before we can assume there exists no $i \leq i' < j' < j$ with $\chi(v_{i'}) = \chi(v_{j'})$. Then $\chi(v_i), \chi(v_{i+1}), ..., \chi(v_{j-1})$ are distinct vertices with $\chi(v_h) \in \chi(e_h) \cap \chi(e_{h+1})$ for $i \leq h < j-1$ and $\chi(v_{j-1}) \in \chi(e_{j-1}) \cap \chi(e_1)$. Note that $\chi(v_i) \neq \chi(v_{i+1})$ since this would imply $|\chi(e_i)| < r$, contradicting that χ is a homomorphism, so $j > i+1$. Thus the hyperedges $\chi(e_i), \chi(e_{i+1}), ..., \chi(e_{j-1})$ form a Berge $(j-i)$-cycle in F' with $j - i \geq 2$ as desired.

This proves the result if $\chi(e_i) = \chi(e_j)$ for some $i \neq j$. If this does not happen and the $\chi(v_i)$ are all distinct, then F' is a Berge p-cycle, and if $\chi(v_i) = \chi(v_j)$ then the same proof as above gives a Berge $(j-i)$-cycle in F'. □

The following lemma allows us to find a relatively dense subgraph of large girth in any r-graph whose maximum i-degree is not too large, where the *i-degree* of an i-set S is the number of hyperedges containing S.

Lemma 4.3. *Let $\ell, r \geq 3$ and let H be an r-graph with maximum i-degree Δ_i for each $i \geq 1$. If $t \geq r^2 4^r \Delta_i^{1/(r-i)}$ for all $i \geq 1$, then H has a subgraph H' of girth larger than ℓ with*

$$e(H') \geq \mathrm{ex}\left(t, \mathcal{C}^r_{[\ell]}\right) t^{-r} \cdot e(H).$$

Proof. Let J be an extremal $\mathcal{C}^r_{[\ell]}$-free r-graph on t vertices and $\chi : V(H) \to V(J)$ chosen uniformly at random. Let $H' \subseteq H$ be the random subgraph which keeps the hyperedge $e \in E(H)$ if

(1) $\chi(e) \in E(J)$, and
(2) $\chi(e) \neq \chi(f)$ for any other $f \in E(H)$ with $|e \cap f| \neq 0$.

We claim that H' is $\mathcal{C}^r_{[\ell]}$-free. Indeed, assume H' contained a subgraph F isomorphic to some element of $\mathcal{C}^r_{[\ell]}$. Let F' be the subgraph of J with $V(F') = \{\chi(u) : u \in V(F)\}$ and $E(F') = \{\chi(e) : e \in E(F)\}$, and note that $F \subseteq H'$ implies that each hyperedge of F satisfies (1), so every element of $E(F')$ is a hyperedge in J. By conditions (1) and (2), χ is a local isomorphism from F to F'. By Lemma 4.2, $F' \subseteq J$ contains a Berge cycle of length at most ℓ, a contradiction to J being $\mathcal{C}^r_{[\ell]}$-free.

It remains to compute $\mathbb{E}[e(H')]$. Given $e \in E(H)$, let A_1 denote the event that (1) is satisfied, let $E_i = \{f \in E(H) : |e \cap f| = i\}$, and let A_2 denote the event that $\chi(f) \not\subseteq \chi(e)$ for any $f \in \bigcup_i E_i$, which in particular implies (2) for the hyperedge e. It is not too difficult to see that $\Pr[A_1] = r! e(J) t^{-r}$, and that for any $f \in E_i$ we have $\Pr[\chi(f) \subseteq \chi(e) | A_1] = (r/t)^{r-i}$. Note for each $i \geq 1$ that $|E_i| \leq 2^r \Delta_i$, as e has at most 2^r subsets of size i each of i-degree at most Δ_i. Taking a union bound we find

$$\Pr[A_2|A_1] \geq 1 - \sum_{i=1}^{r}|E_i|(r/t)^{r-i} \geq 1 - \sum_{i=1}^{r} 2^r\Delta_i(r/t)^{r-i} \geq 1 - \sum_{i=1}^{r} r^{-1}2^{-r} \geq \frac{1}{2},$$

where the second to last inequality used $(r4^r)^{i-r} \geq r^{-1}4^{-r}$ for $i \leq r$. Consequently

$$\Pr[e \in E(H')] = \Pr[A_1] \cdot \Pr[A_2|A_1] \geq r!e(J)t^{-r} \cdot \frac{1}{2} \geq e(J)t^{-r},$$

and linearity of expectation gives $\mathbb{E}[e(H')] \geq e(J)t^{-r} \cdot e(H) = \text{ex}(t, \mathcal{C}^r_{[\ell]})t^{-r} \cdot e(H)$. Thus there exists some $\mathcal{C}^r_{[\ell]}$-free subgraph $H' \subseteq H$ with at least $\text{ex}(t, \mathcal{C}^r_{[\ell]})t^{-r} \cdot e(H)$ hyperedges. □

By the Chernoff bound one can show for

$$p \geq p_1 := n^{\frac{-(r-1)(\ell-1-k)}{\ell-1}}$$

that a.a.s. $H^r_{n,p}$ has maximum i-degree at most $\Theta(pn^{r-i})$ for all i. If Conjecture 1 is true, then a.a.s. for $p \geq p_1$ Lemma 4.3 with $t = \Theta(p^{1/(r-1)}n)$ gives

$$\text{ex}\left(H^r_{n,p}, \mathcal{C}^r_{[\ell]}\right) = \Omega\left(t^{-r}\text{ex}\left(t, \mathcal{C}^r_{[\ell]}\right)pn^r\right) = p^{\frac{1}{(r-1)k}}n^{1+\frac{1}{k}-o(1)}.$$

This proves the lower bound in Theorem 1.5. □

Proof of the lower bound in Theorem 1.4. We use the following variant of Lemma 4.3:

Lemma 4.4. *Let H be an r-graph and let $R_{\ell,v}(H)$ be the number of Berge ℓ-cycles in H on v vertices. For all $t \geq 1$, H has a subgraph H' of girth larger than 3 with*

$$e(H') \geq \left(e(H)t^{2-r} - \sum_{\ell=2}^{3}\sum_{v} t^{2-v}R_{\ell,v}(H)\right)e^{-c\sqrt{\log t}},$$

where $c > 0$ is an absolute constant.

Proof. By the work of Ruzsa and Szemeredi [22] and Erdős, Frankl, and Rödl [7], it is known for all t that there exists a $\mathcal{C}^r_{[3]}$-free r-graph J on t vertices with $t^2e^{-c\sqrt{\log t}}$ hyperedges. Choose a map $\chi : V(H) \to V(J)$ uniformly at random and define $H'' \subseteq H$ to be the subgraph which keeps a hyperedge $e = \{v_1, ..., v_r\} \in E(H)$ if and only if $\chi(e) \in E(J)$.

We claim that if e_1, e_2, e_3 form a Berge triangle in H'', then $\chi(e_1) = \chi(e_2) = \chi(e_3)$. Observe that if v_1, v_2, v_3 are vertices with $v_i \in e_i \cap e_{i+1}$, then we must have, for example, as otherwise $|\chi(e_2)| < r$. Because J is linear we must have $|\chi(e_i) \cap \chi(e_j)| \in \{1, r\}$. These hyperedges cannot all intersect in 1 vertex since this together with the distinct vertices $\chi(v_1), \chi(v_2), \chi(v_3)$ defines a Berge triangle in H'', so we must have to say $\chi(e_1) = \chi(e_2)$. But this means $\chi(v_3), \chi(v_2)$ are distinct vertices in $\chi(e_1) = \chi(e_2)$ and $\chi(e_3)$, so $|\chi(e_1) \cap \chi(e_3)| > 1$ and we must have $\chi(e_1) = \chi(e_3)$ as desired.

The probability that a given Berge triangle C on v vertices in H maps to a given hyperedge in J is at most $(r/t)^v$ (since this is the probability that every vertex of C maps into the edge of J). By linearity of expectation, H'' contains at most $\sum_v R_{3,v}(H)e(J)(r/t)^v$ Berge triangles in expectation. An identical proof shows that H'' contains at most $\sum_v R_{2,v}(H)e(J)(r/t)^v$ Berge 2-cycles in expectation. We can then delete a hyperedge from each of these Berge cycles in H'' to find a subgraph H' with

$$\mathbb{E}[e(H')] \geq e(J)t^{-r} \cdot e(H) - \sum_{\ell=2}^{3} \sum_{v} R_{\ell,v}(H)e(J)(r/t)^v.$$

The result follows since $e(J) = t^2 e^{-c\sqrt{\log t}}$. □

We now prove the lower bound in Theorem 1.4. By Markov's inequality one can show that a.a.s. $R_{3,3r-3}(H_{n,p}^r) = O(p^3 n^{3r-3})$. By the Chernoff bound we have a.a.s. that $e(H_{n,p}^r) = \Omega(pn^r)$, so if we take $t = p^{2/(2r-3)} n(\log n)^{-1}$, then a.a.s. $t^{5-3r}R_{3,3r-3}(H_{n,p}^r)$ is significantly smaller than $t^{2-r}e(H_{n,p}^r)$. A similar result holds for each term $t^{2-v}R_{\ell,v}(H_{n,p}^r)$ with $\ell = 2, 3$ and $v \leq \ell(r-1)$, so by Lemma 4.4 we conclude $\mathrm{ex}(H_{n,p}^r, \mathcal{C}_{[3]}^r)\ ] \geq p^{1/(2r-3)} n^{2-o(1)}$ a.a.s., proving the lower bound in Theorem 1.4.

We note that the proof of Lemma 4.4 fails for larger ℓ. In particular, a Berge 4-cycle can appear in H'' by mapping onto two edges in J intersecting at a single vertex, and with this the bound becomes ineffective.

5 | CONCLUDING REMARKS

- In this paper, we extended ideas of Balogh and Li to bound the number of n-vertex r-graphs with m edges and girth more than ℓ in terms of the number of n-vertex graphs with m edges and girth more than ℓ. The reduction is best possible when $m = \Theta(n^{\ell/(\ell-1)})$ and $\ell - 2$ divides $r - 2$. Theorem 1.3 shows that similar reductions can be made when forbidding a single family of Berge cycles.

By using variations of our method, we can prove the following generalization. For a graph F, a hypergraph H is a *Berge-F* if there exists a bijection $\phi : E(F) \to E(H)$ such that $e \subseteq \phi(e)$ for all $e \in E(F)$. Let $\mathcal{B}^r(F)$ denote the family of r-uniform Berge-F. We can prove the following extension of Theorem 1.3: if there exists a vertex $v \in V(F)$ such that $F - v$ is a forest, then there exists $c = c(F, r)$ such that

$$N_m^r(n, \mathcal{B}^r(F)) \leq 2^{cm} \cdot N_{[m]}^2(n, F)^{r!/2}.$$

For example, this result applies when F is a theta graph. We do not believe that the exponent $r!/2$ is optimal in general, and we propose the following problem.

Problem 1. Let $\ell, r \geq 3$. Determine the smallest value $\beta = \beta(\ell, r) > 0$ such that there exists a constant $c = c(\ell, r)$ so that, for all $m, n \geq 1$,

$$N_m^r\left(n, \mathcal{C}_\ell^r\right) \leq 2^{cm} \cdot N_{[m]}^2(n, C_\ell)^\beta.$$

Theorem 1.3 shows that $\beta \le r!/2$ for all ℓ, r, but in principle we could have $\beta = O_\ell(r)$. We claim without proof that it is possible to use variants of our methods to show $\beta(3,r), \beta(4,r) \le \binom{r}{2}$, but beyond this we do not know any nontrivial upper bounds on β.

• We proposed Conjecture 3 on the extremal function for subgraphs of large girth in random hypergraphs: for some constant $\gamma = \gamma(\ell, r)$, a.a.s.

$$\mathrm{ex}\left(H^r_{n,p}, C^r_{[\ell]}\right) = \begin{cases} n^{1+\frac{1}{\ell-1}+o(1)} & n^{-r+1+\frac{1}{\ell-1}} \le p < n^{-\frac{\gamma(\ell-1-k)}{\ell-1}}, \\ p^{\frac{1}{\gamma k}} n^{1+\frac{1}{k}+o(1)} & n^{-\frac{\gamma(\ell-1-k)}{\ell-1}} \le p \le 1. \end{cases}$$

For $\ell = 3$, this conjecture is true with $\gamma = 2r - 3$, and Conjecture 2 suggests perhaps $\gamma = r - 1 + (r-2)/(\ell-2)$, although we do not have enough evidence to support this (see also the work of Mubayi and Yepremyan [18] on loose even cycles). It would be interesting as a test case to know if $\gamma(3,4) = 5/2$:

Problem 2. Prove or disprove that Conjecture 3 holds with $\gamma(3,4) = 5/2$.

• It seems likely that $N^r_m(n, \mathcal{F})$ controls the a.a.s. behavior of $\mathrm{ex}(H^r_{n,p}, \mathcal{F})$ as $n \to \infty$. Specifically, it is clear that if $\mathcal{F}$ is a family of finitely many r-graphs and $p = p(n)$ and $m = m(n)$ are defined so that $p^m N^r_m(n, \mathcal{F}) \to 0$ as $n \to \infty$, then a.a.s. as $n \to \infty$, $H^r_{n,p}$ contains no $\mathcal{F}$-free subgraph with at least m edges. It would be interesting to determine when $H^r_{n,p}$ a.a.s. contains an $\mathcal{F}$-free subgraph with at least m edges. In particular, we leave the following problem:

Problem 3. Let $m = m(n)$ and $p = p(n)$ so that $p^m N^r_m(n, \ell) \to \infty$ as $n \to \infty$. Then $H^r_{n,p}$ a.a.s. contains a subgraph of girth more than ℓ with at least m edges.

In particular, perhaps one can obtain good bounds on the variance of $N^r_m(n, \ell)$ in $H^r_{n,p}$.

ACKNOWLEDGMENTS

We thank the referees for their helpful comments.

REFERENCES

1. N. Alon and J. H. Spencer, *The probabilistic method*. John Wiley & Sons, Hoboken, New Jersey, 2004.
2. J. Balogh and L. Li, *On the number of linear hypergraphs of large girth*, J. Graph Theory. **93** (2020), no. 1, 113–141.
3. J. Balogh, B. Narayanan, and J. Skokan, *The number of hypergraphs without linear cycles*, J. Combin. Theory Ser. B. **134** (2019), 309–321.
4. J. Balogh and W. Samotij, *The number of $K_{s,t}$-free graphs*, J. Lond. Math. Soc. **83** (2011), no. 2, 368–388.
5. D. Conlon, J. Fox, B. Sudakov, and Y. Zhao, *The regularity method for graphs with few 4-cycles*, J. London Math. Soc. **104** (2021), no. 5, 2376-2401.
6. J. Corsten and T. Tran, *Balanced supersaturation for degenerate hypergraphs*, J. Graph Theory. arXiv preprint arXiv:1707.03788, 2021.
7. P. Erdős, P. Frankl, and V. Rödl, *The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent*, Graphs Combin. **2** (1986), no. 1, 113–121.
8. P. Erdős and M. Simonovits, *Supersaturated graphs and hypergraphs*, Combinatorica. **3** (1983), no. 2, 181–192.

9. A. Ferber, G. McKinley, and W. Samotij, *Supersaturated sparse graphs and hypergraphs*, Int. Math. Res. Not. **2020** (2020), no. 2, 378–402.
10. F. Foucaud, M. Krivelevich, and G. Perarnau, *Large subgraphs without short cycles*, SIAM J. Discrete Math. **29** (2015), no. 1, 65–78.
11. Z. Füredi, *Random Ramsey graphs for the four-cycle*, Discrete Math. **126** (1994), no. 1–3, 407–410.
12. Z. Füredi and M. Simonovits, *The history of degenerate (bipartite) extremal graph problems*, (L. Lovasz, I. Z. Ruzsa, and V. T. Sos eds.), Erdős Centennial, Springer, Berlin, 2013, pp. 169–264.
13. E. Győri and N. Lemons, *Hypergraphs with no cycle of a given length*, Combin. Probab. Comput. **21** (2012), no. 1–2, 193.
14. S. Janson, T. Łuczak, and A. Rucinski, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, (R. L. Graham, and J. K. Lenstra eds.), Wiley-Interscience, New York, 2000.
15. Y. Kohayakawa, B. Kreuter, and A. Steger, *An extremal problem for random graphs and the number of graphs with large even-girth*, Combinatorica. **18** (1998), no. 1, 101–120.
16. F. Lazebnik and J. Verstraëte, *On hypergraphs of girth five*, Electron. J. Combin. **10** (2003), no. 1, R25.
17. R. Morris and D. Saxton, *The number of $C_{2\ell}$-free graphs*, Adv. Math. **298** (2016), 534–580.
18. D. Mubayi and L. Yepremyan, *Random Turán theorem for hypergraph cycles*, arXiv preprint arXiv:2007.10320, 2020.
19. J. Nie, S. Spiro, and J. Verstraëte, *Triangle-free subgraphs of hypergraphs*, Graphs Combin. **37** (2021), no. 6, 2555–2570.
20. C. Palmer, M. Tait, C. Timmons, and A. Z. Wagner, *Turán numbers for Berge-hypergraphs and related extremal problems*, Discrete Math. **342** (2019), no. 6, 1553–1563.
21. G. Perarnau and B. Reed, *Existence of spanning $\mathcal{F}$-free subgraphs with large minimum degree*, Combin. Probab. Comput. **26** (2017), no. 3, 448–467.
22. I. Z. Ruzsa and E. Szemerédi, *Triple systems with no six points carrying three triangles*, Combinatorics (Proceedings of the Fifth Hungarian Colloquium Keszthely, 1976), vol. II, vol. **18**, Colloquium Mathematical Society, János Bolyai, 1978, pp. 939–945.
23. C. Timmons and J. Verstraëte, *A counterexample to sparse removal*, European J. Combin. **44** (2015), 77–86.
24. J. Verstraëte, *Extremal problems for cycles in graphs*, Recent Trends in Combinatorics, IMA Vol. Math. Appl., vol. **159**, Springer, Cham, 2016, pp. 83–116.

How to cite this article: S. Spiro and J. Verstraëte, *Counting hypergraphs with large girth*, J. Graph Theory. 2022;100:543–558. https://doi.org/10.1002/jgt.22794

APPENDIX A: PROOF OF THEOREM 1.1

Here we give a formal proof of Theorem 1.1. The key tool will be the following theorem of Morris and Saxton.

Theorem 1 (Morris and Saxton [17, Theorem 5.1]). *For each $k \geq 2$, there exists a constant $C = C(k)$ such that the following holds for sufficiently large $t, n \in \mathbb{N}$ with $t \leq n^{(k-1)^2/k(2k-1)}/(\log n)^{k-1}$. There exists a collection $\mathcal{G}_k(n, t)$ of at most*

$$\exp(Ct^{-1/(k-1)}n^{1+1/k}\log t)$$

graphs on $[n]$ such that $e(G) \leq tn^{1+1/k}$ for all $G \in \mathcal{G}_k(n, t)$ and such that every C_{2k}-free graph is a subgraph of some $G \in \mathcal{G}_k(n, t)$.

Recall that we wish to prove that for $\ell \geq 3$ and $k = \lfloor \ell/2 \rfloor$, there exists a constant $c > 0$ such that if n is sufficiently large and $m \geq n^{1+1/(2k-1)}(\log n)^2$, then

$$N_m^2(n, \mathcal{C}_{[\ell]}) \leq e^{cm}(\log n)^{(k-1)m}\left(\frac{n^{1+1/k}}{m}\right)^{km}.$$

The bound is trivial if $\ell = 3$ since $N_m^2(n, C_3) \leq \binom{n^2}{m}$, so we may assume $\ell \geq 4$ from now on. Because $N_m^2(n, \mathcal{C}_{[\ell]}) \leq N_m^2(n, C_{2k})$ for all $\ell \geq 4$, it suffices to prove this bound for $N_m^2(n, C_{2k})$. For any integer $t \leq n^{(k-1)^2/k(2k-1)}/(\log n)^{k-1}$ and n sufficiently large, Theorem 1 implies

$$N_m^2(n, C_{2k}) \leq |\mathcal{G}_k(n, t)| \cdot \binom{tn^{1+1/k}}{m} \leq \exp(Ct^{-1/(k-1)}n^{1+1/k}\log t) \cdot (etn^{1+1/k}/m)^m, \tag{A1}$$

with the first inequality using that every C_{2k}-free graph on m edges is an m-edged subgraph of some $G \in \mathcal{G}_k(n, t)$. By taking $t = (n^{1+1/k}\log n/m)^{k-1}$, which is sufficiently small to apply (A1) provided $m \geq n^{1+1/(2k-1)}(\log n)^2$, we see that $N_m^2(n, C_{2k})$ satisfies the desired inequality.