Counting Cycles to Efficiently Certify Sparse Hypergraph Quasirandomness

John Lenz *
University of Illinois at Chicago
lenz@math.uic.edu

Dhruv Mubayi †
University of Illinois at Chicago
mubayi@math.uic.edu

September 5, 2013

Abstract

One surprising property of Chung, Graham, and Wilson's characterization of dense quasirandom graphs is a polynomial-time verifiable property $\texttt{Cycle}_4$, which states that the number of copies of the cycle of length four is what one would expect in a random graph of the same density. Targeting problems like random k-SAT, this algorithm has been extended in several ways to sparse quasirandomness by several researchers. In this note, we show how the spectral hypergraph quasirandom property defined by the authors can be used to create an efficient algorithm that certifies if a sparse hypergraph is quasirandom. Compared to the existing algorithms, our algorithm certifies a different version of quasirandomness and has a faster running time, but does not improve upon the current best bounds for applications like random k-SAT.

1 Introduction

The study of quasirandom or pseudorandom graphs was initiated by Thomason [22, 23] and then refined by Chung, Graham, and Wilson [8], resulting in a list of equivalent (deterministic) properties of dense graph sequences which are inspired by $G(n, p)$ for p a fixed constant. One of the more interesting properties is $\texttt{Cycle}_4$, the property that the number of copies of the cycle of length four is what one would expect in a random graph with the same density. This property can be tested in an n-vertex graph in polynomial time, leading to an efficient algorithm to detect if a dense graph is quasirandom. Extending these algorithms to certify sparse quasirandom graphs in polynomial time has several applications; perhaps the most striking is to random k-SAT. Friedman, Goerdt, and Krivelevich [13] used sparse quasirandom graphs and techniques of Chung, Graham, and Wilson [8] to develop a refutation

*Research partly supported by NSA Grant H98230-13-1-0224.
†Research supported in part by NSF Grants 0969092 and 1300138.

algorithm for random k-SAT. Since then, several researchers [9, 10, 12, 15] have investigated algorithms that work with progressively smaller and smaller densities. All of these algorithms have a quasirandom flavor and some certify a version of hypergraph quasirandomness.

Hypergraph quasirandomness has been studied by many researchers [2, 3, 4, 5, 6, 7, 11, 14, 16, 17, 18, 19, 21] and there are many distinct hypergraph quasirandom properties (see [21]). As a consequence of the spectral hypergraph quasirandomness properties developed by the authors [19, 20, 21], we obtain an algorithm which certifies hypergraph quasirandomness in polynomial time. Compared to the existing algorithms [9, 10, 15], our algorithm certifies a different notion of quasirandomness and is more efficient. Applying our algorithm to the random k-SAT problem essentially matches but does not improve the best known results. To state our algorithm, we require a few definitions.

Definition. Let k be an integer. A *proper partition* π of k is an unordered list of at least two positive integers whose sum is k. For $\pi = k_1 + k_2$, a partition of k into two parts, define the cycle of type π and length 2ℓ, denoted $C_{\pi,2\ell}$, as follows. Let $A_1, \dots, A_\ell$ be sets of size k_1 and let $B_1, \dots, B_\ell$ be sets of size k_2. Let $V(C_{\pi,2\ell}) = A_1 \dot\cup \cdots \dot\cup A_\ell \dot\cup B_1 \dot\cup \cdots \dot\cup B_\ell$. For every $1 \le i \le \ell$, let $A_i \cup B_i$ be a hyperedge of $C_{\pi,2\ell}$. In addition, for every $1 \le i \le \ell - 1$, let $B_i \cup A_{i+1}$ be a hyperedge. Finally, let $B_\ell \cup A_1$ be a hyperedge. [19] gave a general definition of $C_{\pi,2\ell}$ for all proper partitions π; the definition given here is a specialization to partitions into two parts and is all that is required in this paper.

Our algorithm works by counting $C_{\lceil k/2 \rceil + \lfloor k/2 \rfloor, 2\ell}$ and is similar to the algorithm of Hán, Person, and Schacht [15] which counted $C_{1+\dots+1,4}$. The cycle $C_{\lceil k/2 \rceil + \lfloor k/2 \rfloor, 2\ell}$ has significantly fewer vertices that $C_{1+\dots+1,4}$ even for large ℓ. Additionally, the spectral techniques developed in [19, 20] lead to a fast method of counting cycles by computing the trace of a matrix.

Algorithm 1. *On input $k \ge 4$, $\epsilon, \eta > 0$, and a k-uniform hypergraph H with m edges and n vertices,*

- *Let $\pi = \lfloor k/2 \rfloor + \lceil k/2 \rceil$.*
- *Define ℓ as follows:*

$$\ell = \ell_\epsilon = \begin{cases} \left\lceil \frac{k}{4\epsilon} \right\rceil & \text{if } k \text{ is even,} \\ \left\lceil \frac{k+1}{4\sqrt{k}-2} \right\rceil & \text{otherwise.} \end{cases} \tag{1}$$

- *Let N be the number of labeled circuits of type π and length 4ℓ in H.*
- *If $N < (1+\eta)(k!)^{4\ell} m^{4\ell} n^{-2\ell k}$ then output **`Quasirandom`**, otherwise output **`Unknown`**.*

The main result of this paper is the following.

Theorem 1. *Let $k \ge 4$ and $\epsilon, \delta > 0$ be given. There exists a constant $\eta > 0$ depending only on δ and k such that the following holds. If on input k, ϵ, η, and a k-uniform, n-vertex, m-edge hypergraph H Algorithm 1 outputs **`Quasirandom`**, then for all $S_1 \subseteq \binom{V(H)}{\lceil k/2 \rceil}$ and all*

$S_2 \subseteq \binom{V(H)}{\lfloor k/2 \rfloor}$,

$$\left| e(S_1, S_2) - \frac{k!m}{n^k}|S_1||S_2| \right| \leq \frac{\delta m}{n^{k/2}}\sqrt{|S_1||S_2|}. \tag{2}$$

In addition, if

$$p \geq \begin{cases} n^{-k/2+\epsilon} & \text{if } k \text{ is even,} \\ n^{-k/2+\sqrt{k}} & \text{otherwise,} \end{cases} \tag{3}$$

then with probability going to one as n goes to infinity, for a hypergraph H drawn from the distribution $G^{(k)}(n,p)$, Algorithm 1 on input k, ϵ, η, and H outputs **`Quasirandom`**. *Finally, on input of a n-vertex hypergraph Algorithm 1 runs in time $O(n^{k\omega} \operatorname{polylog} n)$ where ω is the exponent for matrix multiplication.*

2 Algorithm Correctness

In this section we prove that Algorithm 1 runs in time $O(n^{k\omega} \operatorname{polylog} n)$ and additionally prove that if Algorithm 1 outputs `Quasirandom` then (2) holds. The proof is very similar to [20, Section 4] and rather than repeat the (extensive) definitions of [19, 20], we will just use identical notation and definitions as [19, 20]. The definitions we require are the definitions of the adjacency map and the first and second largest eigenvalues [19, Section 3] and the definitions of the powers of the adjacency map [19, Section 6.1]. We also require a proposition on counting circuits [19, Proposition 7], two algebraic properties of multilinear maps ([20, Lemma 9] and [20, Proposition 2]), and the hypergraph expander mixing lemma [19, Theorem 4].

Proof of the running time of Algorithm 1. If we count labeled circuits directly, then Algorithm 1 runs in time $O(n^{2k\ell})$ since $C_{\pi,4\ell}$ has $2k\ell$ vertices. For k even this is $n^{O(k^2)}$ and for k odd is $n^{O(k^{3/2})}$. Using [19, Proposition 7], an alternate way of counting labeled circuits is to compute the trace of $A[\tau^2_{\vec{\pi}}]^{2\ell}$. The matrix $A = A[\tau^2_{\vec{\pi}}]$ can be computed in time $O(n^{2k} \operatorname{polylog} n)$ since there are $n^{2\lfloor k/2 \rfloor}$ basis vectors on which to evaluate $\tau_{\vec{\pi}} * \tau_{\vec{\pi}}$ and each evaluation takes $O(n^{\lceil k/2 \rceil})$ multiplications. Note that the maximum size of entries of A is $n^{2k\ell}$. Since A is an $n^{2\lfloor k/2 \rfloor} \times n^{2\lfloor k/2 \rfloor}$-matrix, using repeated squaring we can compute $A^{2\ell}$ in time $O(n^{k\omega} \log \ell \operatorname{polylog} n)$, where ω is the exponent for matrix multiplication. Thus Algorithm 1 runs in time $O(n^{k\omega} \operatorname{polylog} n)$. □

Proof of Algorithm 1 correctness. We prove that if Algorithm 1 outputs `Quasirandom` then equation (2) holds. Let k, ϵ, δ, π, and ℓ be defined as in the theorem and let H be any k-uniform hypergraph with m edges and n vertices. We need to prove that there exists an $\eta > 0$ so that if the number of labeled circuits of type π and length 4ℓ is at most $(1+\eta)(k!)^{4\ell} m^{4\ell} n^{-2\ell k}$, then (2) holds. This proof is very similar to the proof that $\texttt{Cycle}_{4\ell}[\pi] \Rightarrow \texttt{Eig}[\pi]$ and we use the same notation. Let $\hat{1}$ denote the all-ones vector scaled to unit length, let $A = A[\tau_{\vec{\pi}}^2]$, and let $\mu_1, \ldots, \mu_d$ be the eigenvalues of A arranged so that $|\mu_1| \geq \cdots \geq |\mu_d|$. Since $\mu_i^{2\ell} \geq 0$, [19, Proposition 7] implies that

$$\mu_1^{2\ell} \leq \mu_1^{2\ell} + \mu_2^{2\ell} \leq \operatorname{Tr}\left[A^{2\ell}\right] = \#\{\text{possibly degenerate } C_{\pi,4\ell} \text{ in } H\} \leq (1+\eta)\frac{(k!)^{4\ell} m^{4\ell}}{n^{2\ell k}}. \quad (4)$$

Using [20, Lemma 9] and (4), we have that

$$\frac{k!m}{n^{k/2}} = \tau_{\vec{\pi}}(\hat{1}, \ldots, \hat{1}) \leq \|\tau_{\vec{\pi}}\| \leq \sqrt{\mu_1} \leq (1+\eta)^{1/4\ell} \frac{k!m}{n^{k/2}}.$$

Combining this with (4), we obtain

$$\frac{k!^2 m^2}{n^k} \leq \mu_1 \leq (1+\eta)^{1/2\ell} \frac{k!^2 m^2}{n^k} \qquad \text{and} \qquad \mu_2 \leq \eta^{1/2\ell} \frac{k!^2 m^2}{n^k}.$$

Since $\tau_{\vec{\pi}}(\hat{1}, \ldots, \hat{1})^2 = \frac{k!^2 m^2}{n^k}$, [20, Proposition 2] implies that given $\delta > 0$, it is possible to choose $\eta > 0$ such that $\lambda_{2,\pi}(H) = \|\tau_{\vec{\pi}} - qJ_{\vec{\pi}}\| \leq \delta \frac{m}{n^{k/2}}$. Finally, the Hypergraph Expander Mixing Lemma [19, Theorem 4] shows that for all $S_1 \subseteq \binom{V(H)}{\lceil k/2 \rceil}$ and all $S_2 \subseteq \binom{V(H)}{\lfloor k/2 \rfloor}$

$$\left| e(S_1, S_2) - \frac{k!m}{n^k}|S_1||S_2| \right| \leq \lambda_{2,\pi}(H)\sqrt{|S_1||S_2|} \leq \frac{\delta m}{n^{k/2}}\sqrt{|S_1||S_2|}.$$

□

3 Counting Circuits in $G^{(k)}(n,p)$

All that remains to prove Theorem 1 is to prove that for almost all k-uniform hypergraphs with probability satisfying (3), Algorithm 1 outputs `Quasirandom`. This is proved in two steps; let G be a hypergraph drawn from the distribution $G^{(k)}(n,p)$. We first prove that the expected number of labeled circuits in G is $(1+o(1))p^{4\ell}n^{2k\ell}$ (note that $C_{\pi,4\ell}$ has 4ℓ edges and $2k\ell$ vertices, but the proof is complicated by the presence of degenerate cycles). Second, we use the second moment method to prove that with high probability the number of labeled circuits in G is close to the expectation. This in turn will imply that w.h.p. Algorithm 1 outputs `Quasirandom`, since w.h.p. $|E(G)| = (1+o(1))p\binom{n}{k}$ so that w.h.p. $(1+o(1))p^{4\ell}n^{2k\ell} = (1+o(1))(k!)^{4\ell}|E(G)|^{4\ell}n^{-2k\ell}$, which is the bound checked by Algorithm 1. A *degenerate cycle of type* π *and length* 4ℓ is a k-uniform hypergraph H for which there exists an edge-preserving surjection $\phi : V(C_{\pi,4\ell}) \to V(H)$ where ϕ is not an injection.

Lemma 2. *Let $k \geq 2$, $\ell \geq 1$, and $\pi = \lfloor k/2 \rfloor + \lceil k/2 \rceil$. Let H be a degenerate cycle of type π and length 4ℓ with v vertices and e edges. Then*

$$e \geq \begin{cases} \frac{v}{\lceil k/2 \rceil} & \text{if } v > 2\ell\left\lceil \frac{k}{2} \right\rceil + \left\lfloor \frac{k}{2} \right\rfloor, \\ \frac{v}{\lceil k/2 \rceil} - 1 & \text{otherwise.} \end{cases}$$

Proof. Let $\phi : V(C_{\pi,4\ell}) \to V(H)$ be an edge-preserving surjection. Let $E_1, \dots, E_{4\ell}$ be the edges of $C_{\pi,4\ell}$, so that $\phi(E_1), \dots, \phi(E_{4\ell})$ are the edges of H (possibly with repetitions). Since ϕ is edge-preserving, $|\phi(E_1) \cap \phi(E_2)| \geq \lfloor k/2 \rfloor$ so if $\phi(E_2)$ is a distinct edge from $\phi(E_1)$ then $\phi(E_2)$ has at most $\lceil k/2 \rceil$ vertices outside $\phi(E_1)$. Similarly, if $\phi(E_3)$ is an edge of H distinct from $\phi(E_1)$ and $\phi(E_2)$ then it can have at most $\lceil k/2 \rceil$ vertices of H outside $\phi(E_1) \cup \phi(E_2)$ since $\phi(E_3)$ must intersect $\phi(E_2)$ in at least $\lfloor k/2 \rfloor$ vertices. In general, if $\phi(E_i)$ is an edge distinct from $\phi(E_1), \dots, \phi(E_{i-1})$ then it can have at most $\lceil k/2 \rceil$ vertices of H outside $\cup_{j<i} \phi(E_j)$. Since H has no isolated vertices (ϕ is an edge-preserving surjection),

$$v = |V(H)| = \left| \cup \phi(E_i) \right| \leq |\phi(E_1)| + (e-1) \left\lceil \frac{k}{2} \right\rceil = k + (e-1) \left\lceil \frac{k}{2} \right\rceil \tag{5}$$

$$\Leftrightarrow \quad v - \lfloor k/2 \rfloor \leq e \lceil k/2 \rceil$$

$$\Leftrightarrow \quad \frac{v}{\lceil k/2 \rceil} - 1 \leq \frac{v - \lfloor k/2 \rfloor}{\lceil k/2 \rceil} \leq e.$$

Now assume that $v > 2\ell \left\lceil \frac{k}{2} \right\rceil + \left\lfloor \frac{k}{2} \right\rfloor$. We show that the lower bound on e from the previous paragraph can be improved by one. First, assume that for all i, there exists a $j \neq i$ such that $\phi(E_i) = \phi(E_j)$. Since each edge of H has at least two edges of $C_{\pi,4\ell}$ mapped to it, $e = |E(H)| \leq \frac{1}{2} |E(C_{\pi,4\ell})| = 2\ell$. Define disjoint vertex sets $A_1, \dots, A_{4\ell} \subseteq V(C_{\pi,4\ell})$ such that $E_i = A_i \cup A_{i+1}$ for $i < 4\ell$ and $E_{4\ell} = A_{4\ell} \cup A_1$. Now define an auxiliary graph G as follows. First, set $V(G) = \{\phi(A_i) : 1 \leq i \leq 4\ell\}$. (Note this is somewhat subtle: for each set $\phi(A_i)$ we create a vertex of G, except if for some $i \neq j$ we have $\phi(A_i) = \phi(A_j)$, in which case the same vertex of G is used. On the other hand, if $\phi(A_i) \cap \phi(A_j) \neq \emptyset$ but $\phi(A_i) \neq \phi(A_j)$, then separate vertices of G are created.) For each hyperedge $\phi(E_i)$ of H with $i < 4\ell$, add a graph edge between $\phi(A_i)$ and $\phi(A_{i+1})$; in addition, add an edge between $\phi(A_{4\ell})$ and $\phi(A_1)$. The number of edges of G is the same as the number of hyperedges of H so $|E(G)| = e \leq 2\ell$. Also, G is connected since H is a degenerate cycle. Since G is connected with at most 2ℓ edges, G has at most $2\ell + 1$ vertices. Each vertex of G translates into at most $\lceil k/2 \rceil$ vertices of H, but if k is odd there must be at least one vertex of G representing a set of size $\lfloor k/2 \rfloor$. Thus $|V(H)| \leq 2\ell \lceil k/2 \rceil + \lfloor k/2 \rfloor$, which contradicts our assumption that $v > 2\ell \lceil k/2 \rceil + \lfloor k/2 \rfloor$.

Thus we can assume that there exists some E_i such that for all $j \neq i$, $\phi(E_i) \neq \phi(E_j)$. By symmetry, relabel the edges of the cycle so that $i = 1$, i.e. for all $j \neq 1$, $\phi(E_1) \neq \phi(E_j)$. One of the two (or both) of the following occur:

- $\phi(A_1) \subseteq \bigcup_{i>1} \phi(A_i)$,
- $\phi(E_{4\ell})$ is distinct as an edge of H from $\phi(E_1), \dots, \phi(E_{4\ell-1})$.

Indeed, assume that $\phi(A_1)$ is not contained in $\cup_{i>1} \phi(A_i)$. Since $\phi(E_{4\ell})$ includes $\phi(A_1)$ and $\phi(E_2) \cup \cdots \cup \phi(E_{4\ell-1}) = \cup_{i>1} \phi(A_i)$, $\phi(E_{4\ell})$ is distinct from $\phi(E_2), \dots, \phi(E_{4\ell-1})$. But by assumption, $\phi(E_1) \neq \phi(E_{4\ell})$, so that $\phi(E_{4\ell})$ is distinct as a hyperedge of H.

Assume that $\phi(A_1) \subseteq \cup_{i>1} \phi(A_i)$, and consider the argument from the first paragraph: consider edges of the cycle one by one starting from E_1. Each $\phi(E_j)$ which is distinct can add at most $\lceil k/2 \rceil$ new vertices to the union $\cup_{1<i\leq j}\phi(A_i)$, since $\phi(E_j)$ must share at least $\lfloor k/2 \rfloor$ with $\phi(E_{j-1})$. Since $\phi(A_1)$ is a subset of $\cup_{i>1}\phi(A_i)$, the inequality in (5) can be improved to

$$v = |\cup_{i>1}\phi(A_i)| \leq e\left\lceil \frac{k}{2} \right\rceil.$$

Now assume that $\phi(E_{4\ell})$ is distinct as an edge of H from $\phi(E_1), \ldots, \phi(E_{4\ell-1})$. By definition, $\phi(E_{4\ell})$ shares $\phi(A_{4\ell})$ with $\phi(E_{4\ell-1})$ and shares $\phi(A_1)$ with $\phi(E_1)$. Thus $\phi(E_{4\ell})$ is a hyperedge of H distinct from $\phi(E_1), \ldots, \phi(E_{4\ell-1})$ which does not use any new vertices. We can therefore improve the bound on the number of edges by one. □

Lemma 3. *Let k be an even integer at least 4, $\epsilon > 0$, $\pi = k/2 + k/2$, and ℓ as defined in (1). For $p \geq n^{-k/2+\epsilon}$, the expected number of labeled, degenerate circuits of type π and length 4ℓ in $G^{(k)}(n,p)$ is $o(p^{4\ell}n^{2k\ell})$.*

Proof. We need to prove that for every degenerate cycle H of type π and length 4ℓ we have $p^{|E(H)|}n^{|V(H)|} = o(p^{4\ell}n^{2k\ell})$. Indeed, the expected number of labeled copies of H in $G^{(k)}(n,p)$ is $p^{|E(H)|}n^{|V(H)|}$ and there are constantly many degenerate cycles since ℓ and k are constants independent of n and p. Thus if $p^{|E(H)|}n^{|V(H)|} = o(p^{4\ell}n^{2k\ell})$ for all H, the expected number of degenerate cycles is $o(p^{4\ell}n^{2k\ell})$.

Let H be a degenerate cycle of type π and length 4ℓ with v vertices and e edges. We need to prove that $p^e n^v = o(p^{4\ell}n^{2k\ell})$. Substituting in p, we need to prove that

$$n^{v-\frac{ek}{2}+e\epsilon} = o(n^{2k\ell-\frac{4\ell k}{2}+4\ell\epsilon}),$$

i.e. we need to prove that

$$v - e\left(\frac{k}{2} - \epsilon\right) < 4\ell\epsilon. \tag{6}$$

Case 1: $v > \ell k + \frac{k}{2}$. By Lemma 2, $e \geq \frac{2v}{k}$. If $e = 4\ell$, then H has the same number of edges as $C_{\pi,4\ell}$ but fewer vertices so trivially $p^e n^v = o(p^{4\ell}n^{2k\ell})$. Thus $e < 4\ell$ which implies that $v \leq 2k\ell - \frac{k}{2}$, since the map $\phi : V(C_{\pi,4\ell}) \to V(H)$ must map at least two sets of size $k/2$ in $V(C_{\pi,4\ell})$ to the same vertex set in H. Inserting these bounds into (6) and simplifying, we obtain

$$v - e\left(\frac{k}{2} - \epsilon\right) \leq v - \frac{2v}{k}\left(\frac{k}{2} - \epsilon\right) = \frac{2v\epsilon}{k} \leq \frac{2\epsilon}{k}\left(2k\ell - \frac{k}{2}\right) = 4\ell\epsilon - \epsilon < 4\ell\epsilon.$$

Case 2: $v \leq \ell k + \frac{k}{2}$. By Lemma 2, we have that $e \geq \frac{2v}{k} - 1$. Inserting these bounds into (6), we obtain

$$\begin{aligned} v - e\left(\frac{k}{2} - \epsilon\right) \leq v - \left(\frac{2v}{k} - 1\right)\left(\frac{k}{2} - \epsilon\right) = \frac{2v\epsilon}{k} + \frac{k}{2} - \epsilon \leq \frac{2\epsilon}{k}\left(\ell k + \frac{k}{2}\right) + \frac{k}{2} - \epsilon \\ = 2\ell\epsilon + \frac{k}{2}. \end{aligned}$$

Thus to prove (6), we require

$$2\ell\epsilon + \frac{k}{2} < 4\ell\epsilon \quad \Leftrightarrow \quad \frac{k}{4\epsilon} < \ell,$$

which is true by the definition of ℓ. □

Lemma 4. *Let k be an odd integer at least 5, $\pi = \lfloor k/2 \rfloor + \lceil k/2 \rceil$, and ℓ as defined in (1). For $p \geq n^{-k/2+\sqrt{k}}$, the expected number of labeled, degenerate cycles of type π and length 4ℓ in $G^{(k)}(n,p)$ is $o(p^{4\ell} n^{2k\ell})$.*

Proof. Similar to the previous proof, we need to prove that $p^{|E(H)|} n^{|V(H)|} = o(p^{4\ell} n^{2k\ell})$ for all degenerate cycles H. Let H be a degenerate cycle of type π and length 4ℓ with v vertices and e edges. Substituting in the value of p, we need to prove that

$$v - e\left(\frac{k}{2} - \sqrt{k}\right) < 4\ell\sqrt{k}. \tag{7}$$

Case 1: $v > 2\ell \lceil k/2 \rceil + \lfloor k/2 \rfloor$. In this case, Lemma 2 implies that $e \geq v / \lceil k/2 \rceil$. If $e = 4\ell$, then H has the same number of edges as $C_{\pi,4\ell}$ but fewer vertices so trivially $p^e n^v = o(p^{4\ell} n^{2k\ell})$. Thus $e < 4\ell$ which implies that $v \leq 2k\ell - \lfloor k/2 \rfloor$, since the map $\phi : V(C_{\pi,4\ell}) \to V(H)$ must map at least two sets of size $\lfloor k/2 \rfloor$ in $V(C_{\pi,4\ell})$ to the same vertex set in H. Starting from (7), plugging in these two bounds on v and e, and using that $k \geq 5$ we obtain

$$\begin{aligned} v - e\left(\frac{k}{2} - \sqrt{k}\right) &\leq v - \frac{v}{\lceil k/2 \rceil}\left(\frac{k}{2} - \sqrt{k}\right) = v\left(1 - \frac{2}{k+1}\left(\frac{k}{2} - \sqrt{k}\right)\right) \\ &\leq \left(2\ell k - \frac{k-1}{2}\right)\left(1 - \frac{k}{k+1} + \frac{2\sqrt{k}}{k+1}\right). \end{aligned} \tag{8}$$

Thus to prove (7), we require that (8) is less than $4\ell\sqrt{k}$. Solving for ℓ, we obtain

$$\left(2\ell k - \frac{k-1}{2}\right)\left(1 - \frac{k}{k+1} + \frac{2\sqrt{k}}{k+1}\right) < 4\ell\sqrt{k} \quad \Leftrightarrow \quad \ell < \frac{2k^{3/2} + k - 2\sqrt{k} - 1}{4k - 8\sqrt{k}}. \tag{9}$$

Therefore, to complete the proof we need to check that the definition of ℓ from from (1) makes this inequality true for all $k \geq 5$.

Since $8k\sqrt{k} < 4k^2$ and $8\sqrt{k} < 26k$ for $k \geq 5$,

$$8k\sqrt{k} + 8\sqrt{k} < 4k^2 + 26k + 2.$$

This implies that

$$4k^2 + 8k\sqrt{k} - 36k + 8\sqrt{k} < 8k^2 - 10k + 2$$

which implies that

$$(k+4\sqrt{k}-1)(4k-8\sqrt{k}) < (2k^{3/2}+k-2\sqrt{k}-1)(4\sqrt{k}-2)$$
$$\frac{k+1}{4\sqrt{k}-2}+1 < \frac{2k^{3/2}+k-2\sqrt{k}-1}{4k-8\sqrt{k}}.$$

Since $\ell = \left\lceil \frac{k+1}{4\sqrt{k}-2} \right\rceil$, (9) holds.

Case 2: $v \leq 2\ell \lceil k/2 \rceil + \lfloor k/2 \rfloor$. In this case, Lemma 2 implies that $e \geq v/\lceil k/2 \rceil - 1$. Plugging these two bounds into (7), we obtain

$$\begin{aligned} v - e\left(\frac{k}{2}-\sqrt{k}\right) &\leq v - \left(\frac{v}{\lceil k/2 \rceil}-1\right)\left(\frac{k}{2}-\sqrt{k}\right) \\ &\leq \left(2\ell\left\lceil\frac{k}{2}\right\rceil + \left\lfloor\frac{k}{2}\right\rfloor\right)\left(1-\frac{1}{\lceil k/2\rceil}\left(\frac{k}{2}-\sqrt{k}\right)\right)+\frac{k}{2}-\sqrt{k}. \end{aligned} \tag{10}$$

Since $1 - 1/\lceil k/2 \rceil (k/2 - \sqrt{k}) = \frac{2\sqrt{k}+1}{k+1}$, (10) continues as

$$\begin{aligned} \left(\ell(k+1)+\frac{k-1}{2}\right)\left(\frac{2\sqrt{k}+1}{k+1}\right)+\frac{k}{2}-\sqrt{k} &< 2\ell\sqrt{k}+\ell+\sqrt{k}+\frac{1}{2}+\frac{k}{2}-\sqrt{k} \\ &= 2\ell\sqrt{k}+\ell+\frac{k}{2}+\frac{1}{2}. \end{aligned}$$

Therefore, to prove (7) we must prove that

$$2\ell\sqrt{k}+\ell+\frac{k}{2}+\frac{1}{2} \leq 4\ell\sqrt{k} \quad \Leftrightarrow \quad \ell \geq \frac{k+1}{4\sqrt{k}-2},$$

which is true by the definition of ℓ. □

Proof of Theorem 1. We need to prove that for all $\eta > 0$ and for p defined as in (3), with probability going to one as n goes to infinity, a graph G drawn from the distribution $G^{(k)}(n,p)$ has the following property: the number of labeled circuits of type π and length 4ℓ in G is at most $(1+\eta)(k!)^{4\ell}|E(G)|^{4\ell}n^{-2k\ell}$. With high probability, $|E(G)| = (1 \pm \frac{\eta}{16\ell})\frac{pn^k}{k!}$. Therefore, we need to prove that with high probability the number of labeled circuits in G is at most $(1+\frac{\eta}{2})p^{4\ell}n^{2k\ell}$.

We prove this in two stages. First, we prove that with high probability the number of labeled, degenerate cycles is upper bounded by $\frac{\eta}{4}p^{4\ell}n^{2k\ell}$. This is just the first moment method as follows. Let Y be the number of labeled, degenerate cycles in G. By Markov's Inequality,

$$\mathbb{P}\left[Y > \frac{\eta p^{4\ell}n^{2k\ell}}{4}\right] < \frac{4\,\mathbb{E}[Y]}{\eta p^{4\ell}n^{2k\ell}}.$$

Lemmas 3 and 4 prove that the right hand side goes to zero as n goes to infinity, so with high probability the number of labeled, degenerate cycles in G is at most $\frac{\eta}{4}p^{4\ell}n^{2k\ell}$.

We now use the second moment method to prove that with high probability, the number of labeled (non-degenerate) cycles is concentrated around its expectation. First by the definition of p, the expected number of cycles goes to infinity. For each list W of $2k\ell$ distinct vertices of G, define an event A_W as "W forms a labeled $C_{\pi,4\ell}$ in G". Using the techniques in [1, Section 4.3], the second moment method then comes down to showing that for a list W,

$$\Delta^* = \sum_{W' \sim W} \mathbb{P}[A_{W'}|A_W] = o(p^{4\ell}n^{2k\ell}),$$

where $W' \sim W$ means that the events A_W and $A_{W'}$ are dependent.

Assume that $\hat{W}$ and $\hat{W}'$ are cycles in $K_n^{(k)}$ which share at least one edge and let W and W' be the lists of vertices of $\hat{W}$ and $\hat{W}'$ respectively. Then we have that A_W and $A_{W'}$ are dependent. Let v be the number of vertices of $\hat{W}'$ which do not appear in $V(\hat{W})$. Since the cycle is two regular, there must be at least $\frac{2v}{k}$ edges of $\hat{W}'$ which are not in $E(\hat{W})$. Thus

$$\mathbb{P}[A_{W'}|A_W] \leq p^{2v/k}.$$

There are at most n^v choices for v vertices outside $V(\hat{W})$, so

$$\Delta^* \leq \sum_{v=1}^{2k\ell-k} p^{2v/k}n^v. \tag{11}$$

The algebra in Case 1 of Lemmas 3 and 4 computed that if e and v are integers such that $e \geq v/\lceil k/2 \rceil$ and $v \leq 2k\ell - \lfloor k/2 \rfloor$, then $p^e n^v = o(p^{4\ell}n^{2k\ell})$. Since W and W' must share at least k vertices, $v \leq 2k\ell - k \leq 2k\ell - \lfloor k/2 \rfloor$, so if we let $e = 2v/k \geq v/\lceil k/2 \rceil$, then the computations in Case 1 of Lemmas 3 and 4 imply that $p^{2v/k}n^v = o(p^{4\ell}n^{2k\ell})$ for all $1 \leq v \leq 2k\ell - k$. Since there are constantly many terms in the sum (11), $\Delta^* = o(p^{4\ell}n^{2k\ell})$ which implies that with probability going to one as n goes to infinity, the number of cycles is at most $(1 + \frac{\eta}{4})p^{4\ell}n^{2k\ell}$. Combined with the fact that with high probability the number of degenerate cycles is at most $\frac{\eta}{4}p^{4\ell}n^{2k\ell}$, the number of labeled circuits in G is at most $(1 + \frac{\eta}{2})p^{4\ell}n^{2k\ell}$ with high probability, completing the proof. □

References

[1] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Inc., Hoboken, NJ, third edition, 2008. With an appendix on the life and work of Paul Erdős.

[2] T. Austin and T. Tao. Testability and repair of hereditary hypergraph properties. *Random Structures Algorithms*, 36(4):373–463, 2010.

[3] F. Chung. Quasi-random hypergraphs revisited. *Random Structures Algorithms*, 40(1):39–48, 2012.

[4] F. R. K. Chung. Quasi-random classes of hypergraphs. *Random Structures Algorithms*, 1(4):363–382, 1990.

[5] F. R. K. Chung and R. L. Graham. Quasi-random hypergraphs. *Random Structures Algorithms*, 1(1):105–124, 1990.

[6] F. R. K. Chung and R. L. Graham. Quasi-random set systems. *J. Amer. Math. Soc.*, 4(1):151–196, 1991.

[7] F. R. K. Chung and R. L. Graham. Cohomological aspects of hypergraphs. *Trans. Amer. Math. Soc.*, 334(1):365–388, 1992.

[8] F. R. K. Chung, R. L. Graham, and R. M. Wilson. Quasi-random graphs. *Combinatorica*, 9(4):345–362, 1989.

[9] A. Coja-Oghlan, C. Cooper, and A. Frieze. An efficient sparse regularity concept. *SIAM J. Discrete Math.*, 23(4):2000–2034, 2009/10.

[10] A. Coja-Oghlan, A. Goerdt, A. Lanka, and F. Schädlich. Techniques from combinatorial approximation algorithms yield efficient algorithms for random $2k$-SAT. *Theoret. Comput. Sci.*, 329(1-3):1–45, 2004.

[11] D. Conlon, H. Hàn, Y. Person, and M. Schacht. Weak quasi-randomness for uniform hypergraphs. *Random Structures Algorithms*, 40(1):1–38, 2012.

[12] U. Feige and E. Ofek. Easily refutable subformulas of large random 3CNF formulas. In *Automata, languages and programming*, volume 3142 of *Lecture Notes in Comput. Sci.*, pages 519–530. Springer, Berlin, 2004.

[13] J. Friedman, A. Goerdt, and M. Krivelevich. Recognizing more unsatisfiable random k-SAT instances efficiently. *SIAM J. Comput.*, 35(2):408–430, 2005.

[14] W. T. Gowers. Quasirandomness, counting and regularity for 3-uniform hypergraphs. *Combin. Probab. Comput.*, 15(1-2):143–184, 2006.

[15] H. Hàn, Y. Person, and M. Schacht. Note on strong refutation algorithms for random k-SAT formulas. In *LAGOS'09—V Latin-American Algorithms, Graphs and Optimization Symposium*, volume 35 of *Electron. Notes Discrete Math.*, pages 157–162. Elsevier Sci. B. V., Amsterdam, 2009.

[16] P. Keevash. A hypergraph regularity method for generalized Turán problems. *Random Structures Algorithms*, 34(1):123–164, 2009.

[17] Y. Kohayakawa, B. Nagle, V. Rödl, and M. Schacht. Weak hypergraph regularity and linear hypergraphs. *J. Combin. Theory Ser. B*, 100(2):151–160, 2010.

[18] Y. Kohayakawa, V. Rödl, and J. Skokan. Hypergraphs, quasi-randomness, and conditions for regularity. *J. Combin. Theory Ser. A*, 97(2):307–352, 2002.

[19] J. Lenz and D. Mubayi. Eigenvalues and linear quasirandom hypergraphs. submitted. http://arxiv.org/abs/1208.4863.

[20] J. Lenz and D. Mubayi. Eigenvalues of non-regular linear quasirandom hypergraphs. online at http://www.math.uic.edu/ lenz/nonregular-art.pdf.

[21] J. Lenz and D. Mubayi. The poset of hypergraph quasirandomness. accepted in Random Structures and Algorithms. http://arxiv.org/abs/1208.5978.

[22] A. Thomason. Pseudorandom graphs. In *Random graphs '85 (Poznań, 1985)*, volume 144 of *North-Holland Math. Stud.*, pages 307–331. North-Holland, Amsterdam, 1987.

[23] A. Thomason. Random graphs, strongly regular graphs and pseudorandom graphs. In *Surveys in combinatorics 1987 (New Cross, 1987)*, volume 123 of *London Math. Soc. Lecture Note Ser.*, pages 173–195. Cambridge Univ. Press, Cambridge, 1987.