Combinatorics, Probability and Computing (2021), **30**, pp. 631–653
doi:10.1017/S0963548320000619

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Counting Hamilton cycles in Dirac hypergraphs

Stefan Glock[1], Stephen Gould[2], Felix Joos[3]*, Daniela Kühn[2] and Deryk Osthus[2]

[1]ETH Institute for Theoretical Studies, Clausiusstrasse 47, 8092 Zürich, Switzerland; [2]School of Mathematics, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK; [3]Institut für Informatik, Heidelberg University, Germany
*Corresponding author. Email: joos@informatik.uni-heidelberg.de

(Received 20 November 2019; revised 30 October 2020; accepted 24 October 2020; first published online 17 December 2020)

Abstract

A tight Hamilton cycle in a k-uniform hypergraph (k-graph) G is a cyclic ordering of the vertices of G such that every set of k consecutive vertices in the ordering forms an edge. Rödl, Ruciński and Szemerédi proved that for $k \geqslant 3$, every k-graph on n vertices with minimum codegree at least $n/2 + o(n)$ contains a tight Hamilton cycle. We show that the number of tight Hamilton cycles in such k-graphs is $\exp(n \ln n - \Theta(n))$. As a corollary, we obtain a similar estimate on the number of Hamilton ℓ-cycles in such k-graphs for all $\ell \in \{0, \dots, k-1\}$, which makes progress on a question of Ferber, Krivelevich and Sudakov.

2020 MSC Codes: 05C65, 05C81, 05C45

1. Introduction

1.1 Counting Hamilton cycles in graphs

The problem of determining sufficient conditions for the existence of Hamilton cycles in graphs is one of the central topics in graph theory, and has given rise to extensive research. A classical result of Dirac [10] states that graphs on $n \geqslant 3$ vertices with minimum degree at least $n/2$ (*Dirac graphs*) contain a Hamilton cycle, and there are natural families of graphs which show that $n/2$ is best possible.

Bollobás [3] and Bondy [4] asked for an asymptotic estimate for the number of distinct Hamilton cycles in Dirac graphs. In 2003, Sárközy, Selkow and Szemerédi [29] made substantial progress on this question by showing that n-vertex Dirac graphs contain $\exp(n \ln n - \Theta(n))$ Hamilton cycles. They also posed the question of whether this is the right order of magnitude for graphs satisfying other conditions known to ensure Hamiltonicity, like those of Ore, Pósa and Chvátal (see [4]). Further, they conjectured that the minimum number of Hamilton cycles in n-vertex Dirac graphs is $\exp(n \ln n - n(1 + \ln 2) - o(n))$.

Cuckler and Kahn [8] analysed a self-avoiding random walk on the vertices of Dirac graphs to verify this conjecture as a consequence of a more precise result. Moreover, in a separate paper [7], they used entropy considerations to provide an upper bound for the number of Hamilton cycles

†This project has received partial funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 786198, D. Kühn and D. Osthus). The research leading to these results was also partially supported by the EPSRC, grants EP/N019504/1 (S. Glock and D. Kühn) and EP/S00100X/1 (D. Osthus), and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), grants 339933727 and 428212407 (F. Joos), as well as the Royal Society and the Wolfson Foundation (D. Kühn).

© The Author(s), 2020. Published by Cambridge University Press.

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

in Dirac graphs. More precisely, writing $\Psi(G)$ to denote the number of distinct Hamilton cycles of a graph G, the main results of [7] and [8] together state that for any n-vertex Dirac graph G, we have $\log_2 \Psi(G) = 2H(G) - n\log_2 e - o(n)$, where $H(G)$ is the *entropy* of G. Combined with the result [8, Theorem 1.3] that Dirac graphs G on n vertices satisfy $H(G) \geqslant (n/2)\log_2 \delta(G)$, this confirms the conjecture of [29]. Moreover, the parameter $H(G)$ is the maximum of a concave function subject to linear constraints, and can thus be efficiently estimated. This yields an efficient algorithm for estimating $\Psi(G)$ for Dirac graphs G, to within subexponential factors.

1.2 Hamilton cycles in hypergraphs

The study of Hamilton cycles in hypergraphs was initiated in a 1976 paper of Bermond, Germa, Heydemann and Sotteau [2]. For k-uniform hypergraphs (k-*graphs*), we may sensibly define a cycle in a number of ways (see *e.g.* [21], [24] and [30]). Let $k \geqslant 2$ be an integer and let $\ell \in \{0, \dots, k-1\}$. We say that a k-uniform hypergraph C is an ℓ-cycle if there exists a cyclic ordering $x_1, \dots, x_t$ of the vertices of C such that $(k-\ell)$ divides t and $\{x_{(i-1)(k-\ell)+1}, \dots, x_{(i-1)(k-\ell)+k}\}_{i\in[t/(k-\ell)]}$ is the edge set of C (with indices taken modulo t). A *Hamilton ℓ-cycle* of a k-graph G is a subgraph $C \subseteq G$, where C is a k-uniform ℓ-cycle with $V(C) = V(G)$. Thus, if G contains a Hamilton ℓ-cycle, then $k-\ell$ divides $|V(G)|$. Moreover, if $\ell = 0$ then a Hamilton ℓ-cycle is just a perfect matching of G. We usually call a $(k-1)$-cycle a *tight cycle*, and we say that a Hamilton $(k-1)$-cycle of a k-graph G is a *tight Hamilton cycle* of G.

We wish to generalize the study of Hamilton cycles in Dirac graphs to the setting of hypergraphs, and so we now need a natural hypergraph generalization of the notion of degree. Given a k-graph G and a set $S \subseteq V(G)$ of $k-1$ vertices, we say that the *codegree* of S in G, denoted $d_G(S)$ (or simply $d(S)$ when G is clear from the context), is the number of edges of G containing S. For a k-graph G, we write $\delta(G)$ for the minimum codegree over all $(k-1)$-sets $S \subseteq V(G)$, and refer to this quantity as the *minimum codegree* of G.

Katona and Kierstead [16] gave a sufficient condition on the minimum codegree for k-graphs to have a tight Hamilton cycle. Further, they conjectured that for all integers $k \geqslant 2$, a minimum codegree of at least $n/2$ suffices for n-vertex k-graphs. Rödl, Ruciński and Szemerédi proved an asymptotic version [25] of the $k = 3$ case of this conjecture, and then an exact version for large n [28]. The work of [25] was shortly afterwards generalized to all integers $k \geqslant 3$ by the same authors [26]. Further results on tight Hamilton cycles can be found in [1] and [23], for example. For $(k-\ell) \nmid k$, Kühn, Mycroft and Osthus [18] asymptotically determined the threshold for the existence of a Hamilton ℓ-cycle (this generalized previous results in [13], [17] and [19]). Subsequently several exact results were proved in [9] and [14]. It turns out that the threshold is significantly below $n/2$ if $(k-\ell) \nmid k$. For all other cases it follows from the result of [26] that the threshold is asymptotically $n/2$.

1.3 Our main result

Ferber, Krivelevich and Sudakov [12] were the first to generalize the study of counting Hamilton cycles to the hypergraph setting (and also considered perfect matchings). They proved for $1 \leqslant \ell < k/2$ that if a k-graph G on n vertices with $(k-\ell) \mid n$ satisfies $\delta(G) \geqslant \alpha n$ for some $\alpha > 1/2$, then G contains

$$(1 - o(1))^n \cdot n! \cdot \left(\frac{\alpha}{\ell!(k-2\ell)!}\right)^{n/(k-\ell)}$$

Hamilton ℓ-cycles. As a natural question, they asked whether this can be generalized to all ℓ.

We adapt some ideas from the random walk analysis of [8] to show that any large k-graph whose minimum codegree is slightly above $n/2$ contains a large number of tight Hamilton cycles.

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

Theorem 1.1. *For a fixed integer $k \geqslant 2$ and a fixed constant $\gamma > 0$, the number of tight Hamilton cycles of a k-graph G on n vertices with $\delta(G) \geqslant (1/2 + \gamma)n$ is* $\exp(n \ln n - \Theta(n))$.

Notice that we claim this number of tight Hamilton cycles holds with equality, up to the exponential error bound $\exp(-\Theta(n))$. We discuss this error bound further in the concluding remarks. It will suffice to show that the lower bound holds, since any k-graph on n vertices trivially has at most $(n-1)!/2$ distinct tight Hamilton cycles. Theorem 1.1 easily yields the following corollary about the number of Hamilton ℓ-cycles in a k-graph whose codegrees are slightly above $n/2$, for each $\ell \in \{0, \dots, k-1\}$.

Corollary 1.1. *For a fixed integer $k \geqslant 2$ and a fixed constant $\gamma > 0$, the number of Hamilton ℓ-cycles of a k-graph G on n vertices with $(k-\ell) \mid n$ and $\delta(G) \geqslant (1/2 + \gamma)n$ is*

(i) $\exp((1 - 1/k)n \ln n - \Theta(n))$ *if* $\ell = 0$;
(ii) $\exp(n \ln n - \Theta(n))$, *if* $\ell \in [k-1]$.

This addresses the above-mentioned question of Ferber, Krivelevich and Sudakov (though our result is less precise than theirs for $\ell < k/2$). We remark that the $-\frac{1}{k}$-term for the case of perfect matchings is missing in [12, Theorem 1.1], but follows from their proof. Finally, recall that the minimum codegree threshold for the existence of Hamilton ℓ-cycles can be below $n/2$ when $\ell < k-1$. It would thus be a natural question to extend the counting results to this larger range. For the rest of the paper, we focus on counting tight Hamilton cycles.

2. Sketch of the proof of Theorem 1.1

In this section we provide a rough sketch of the proof of our main result.

2.1 Basic notation

We first need to introduce some notation that we use throughout the paper. For a set V and a natural number ℓ, we write $\binom{V}{\ell}$ to denote the set of all unordered ℓ-subsets of distinct elements of V. We write $(V)_\ell$ to denote the set of all ordered ℓ-subsets of distinct elements of V, so that $|(V)_\ell| = \ell! \left|\binom{V}{\ell}\right|$. We generally use boldface capital letters to denote unordered subsets $\mathbf{S} \in \binom{V}{\ell}$ of the fixed size ℓ, and we exclusively use boldface capital letters with arrows above to denote ordered subsets $\overrightarrow{\mathbf{S}} \in (V)_\ell$. When an ordered tuple $\overrightarrow{\mathbf{S}} \in (V)_\ell$ is first given, the arrow will exclusively point to the right. We may subsequently drop the arrow to denote the unordered version of this ℓ-set, so that if $\overrightarrow{\mathbf{S}}$ is the ordered sequence of ℓ distinct elements $(x_1, \dots, x_\ell)$, then $\mathbf{S}$ subsequently used without the arrow denotes the unordered set $\{x_1, \dots, x_\ell\}$. Moreover, we write $\overleftarrow{\mathbf{S}}$ to denote the ordered ℓ-tuple obtained by reversing the ordering of $\overrightarrow{\mathbf{S}}$, so that $\overleftarrow{\mathbf{S}} = (x_\ell, x_{\ell-1}, \dots, x_1)$. Let $G = (V, E)$ be a hypergraph and let $U \subseteq V(G)$. Then the sub(hyper)graph of G *induced* by U, denoted $G[U]$, is the hypergraph $H = (V(H), E(H))$, where $V(H) = U$, and $E(H)$ is precisely the set of all edges of G containing only vertices in U. We write $G - U$ to denote the hypergraph $G' \subseteq G$ obtained from G by deleting the vertices in U and all edges of G containing any vertex in U. We say that a k-graph P is a *k-uniform tight path* (or simply *tight path* if k is clear from the context) if P admits an ordering of its vertices $V(P) = \{v_1, \dots, v_m\}$ such that

$$E(P) = \{\{v_i, \dots, v_{i+k-1}\} \colon 1 \leqslant i \leqslant m - (k-1)\}.$$

The *ends* of P are the ordered $(k-1)$-tuples $(v_1, \dots, v_{k-1})$ and $(v_m, \dots, v_{m-k+2})$. We also say that P *connects* the ends of P. We say that a tight path P with m edges (and thus with

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

$m + (k-1)$ vertices) is an *m-path*, and has *length* m. For a k-graph G and an integer $t \geqslant k$ we say that a sequence $(v_1, v_2, \ldots, v_t)$ of (not necessarily distinct) vertices is a *walk* in G if every set of k consecutive vertices in the sequence forms an edge. Let $\gamma > 0$ be a constant. A k-graph G on n vertices is called γ*-Dirac* if $\delta(G) \geqslant (1/2 + \gamma)n$. Finally, given a hypergraph G, we say a weighting of the edges $\mathbf{x}\colon E(G) \to \mathbb{R}^+$ is a *fractional matching* if we have $\sum_{e \ni v} \mathbf{x}(e) \leqslant 1$ for every $v \in V(G)$, and we say that $\mathbf{x}$ is *perfect* if $\sum_{e \ni v} \mathbf{x}(e) = 1$ for every $v \in V(G)$.

2.2 Outline of the argument

Let $\gamma > 0$, and let G be an n-vertex k-graph satisfying $\delta(G) \geqslant (1/2 + \gamma)n$, where $k \geqslant 2$ and n is sufficiently large. The main step of our proof is to count tight paths of length $n - o(n)$ in G. Using the framework of Rödl, Ruciński and Szemerédi [26], which is based on the absorption technique, we can complete each such long path into a tight Hamilton cycle of G. The key lemma (Lemma 5.1) in the proof of Theorem 1.1 states that we can find many paths of length $\sqrt{n}$ in G, all starting at the same ordered $(k-1)$-tuple $\overrightarrow{\mathbf{S}} \in (V(G))_{k-1}$, such that for each such path the remainder of G still has minimum codegree at least $(1/2 + \gamma - n^{-2/3})(n - \sqrt{n})$. The proof of this 'iteration lemma' is the sole focus of Section 5, and involves the analysis of a self-avoiding random walk $\mathcal{X}$ on the vertices of G. In order to prove the iteration lemma, we first need to show that G admits a perfect fractional matching which is 'normal', which means that each edge of G has weight $\Theta(n^{-k+1})$. We construct such a normal perfect fractional matching $\mathbf{x}$ in Section 4 via a probabilistic argument based on switchings (it is not clear how to generalize the entropy-based approach of [8] to the hypergraph setting).

In Section 5 we use $\mathbf{x}$ to define the transition probabilities of the random walk $\mathcal{X}$. We construct $\mathcal{X}$ such that an outcome of $\mathcal{X}$ corresponds to a tight path in G of length $\sqrt{n}$ which starts at some given $\overrightarrow{\mathbf{S}} \in (V(G))_{k-1}$. We wish to count the number of outcomes of $\mathcal{X}$ which essentially leave the γ-Dirac property of the remaining graph intact. Such outcomes of $\mathcal{X}$ are called *good* walks. It will suffice to show that $\mathcal{X}$ is good with probability at least $1/2$. To do this, we will show that it is likely that the vertices that $\mathcal{X}$ visits look roughly like a uniformly random subset of the vertices of G, of appropriate size. We will show that the behaviour of $\mathcal{X}$ over a small number of steps can be assumed to be very close to the behaviour of a modified version of $\mathcal{X}$, in which the walk is allowed to revisit vertices. We use the normality property of $\mathbf{x}$ to show that the modified walk mixes rapidly, and we use the fact that $\mathbf{x}$ is a perfect fractional matching to show that, under the stationary distribution, each vertex is essentially visited with the same probability. Thus, roughly speaking, the distribution of the vertices for $\mathcal{X}$ to visit at any step is close to uniform on $V(G)$. We give a more thorough sketch of the proof of the iteration lemma in Section 5.2.

In Section 6 we focus on repeatedly applying the iteration lemma to obtain many long paths in G. Let P be a $\sqrt{n}$-path in G obtained from the first iteration of the iteration lemma, let $\mathbf{T}$ be the unordered set consisting of the final $k-1$ vertices of P, and let $G_P \coloneqq G - (V(P) \setminus \mathbf{T})$. The idea is that since the γ-Dirac property is essentially intact in G_P, we may find a new normal perfect fractional matching $\mathbf{x}_P\colon E(G_P) \to \mathbb{R}^+$ and can thus apply the iteration lemma to G_P. In this second iteration we insist that all the walks $\mathcal{X}$ start at the final ordered $(k-1)$-tuple of P. Then we may attach any of the paths P' from the second iteration onto P to obtain a longer tight path in G which still leaves the γ-Dirac property of the remaining graph essentially intact. We show that we may iterate this process until fewer than $n^{7/8}$ vertices of G remain, and we multiplicatively use the count of paths given by the iteration lemma to deduce that the number of resulting long paths of G is essentially the number given in the statement of Theorem 1.1. (Observe that each combination of paths yields a different concatenated path.) Finally, we complete the proof of Theorem 1.1 by absorbing the vertices left over by each such long path into a tight Hamilton cycle of G.

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

3. Preliminaries

In the following section we collect further notation, as well as some results that we will use throughout the paper.

3.1 Notation

Let $\gamma > 0$ be a constant. We say that a k-graph G is an (n, k, γ)-*graph* if G has n vertices and G is γ-Dirac. When G is clear from the context, we often write V instead of $V(G)$. For a k-graph G and $\mathbf{S} \in \binom{V}{k-1}$, we write

$$N_G(\mathbf{S}) \coloneqq \{v \in V \colon \mathbf{S} \cup \{v\} \in E(G)\}.$$

We say that $\mathbf{S}$ is *isolated* if $N_G(\mathbf{S}) = \emptyset$, and that $\mathbf{S}$ is *non-isolated* if $\mathbf{S}$ is not isolated. For a positive integer ℓ, we say that a walk $(v_1, \dots, v_{\ell+k-1})$ on the vertices of G is an ℓ-*walk*. Let $\overrightarrow{\mathbf{S}}, \overrightarrow{\mathbf{T}} \in (V)_{k-1}$. We say an ℓ-walk $(v_1, \dots, v_{\ell+k-1})$ in G is an ℓ-*walk from* $\overrightarrow{\mathbf{S}}$ *to* $\overrightarrow{\mathbf{T}}$ if

$$\overrightarrow{\mathbf{S}} = (v_1, \dots, v_{k-1}) \quad \text{and} \quad \overrightarrow{\mathbf{T}} = (v_{\ell+1}, \dots, v_{\ell+k-1}).$$

A *matching* M of a k-graph G is a collection of vertex-disjoint edges of G, and we say that M is *perfect* if every vertex $v \in V(G)$ is included in some edge of M. Let M be a matching in a k-graph G. Where it has no effect on the argument, we sometimes abuse notation and identify M with the subgraph $M' \subseteq G$ satisfying $E(M') = M$ and $V(M') = \bigcup_{e \in M} e$. For finite sets $U \subseteq V$ and a function $f \colon V \to \mathbb{R}$, we define $f(U) \coloneqq \sum_{u \in U} f(u)$, and $\|f\|_\infty \coloneqq \max_{v \in V} f(v)$. We write $\mathbb{1}_U \colon V \to \{0, 1\}$ to be the *indicator function* for U, defined by $\mathbb{1}_U(x) = 1$ if $x \in U$, and $\mathbb{1}_U(x) = 0$ otherwise. For an event $\mathcal{E}$ in a probability space, we write $\mathcal{E}^c$ to denote the complement of $\mathcal{E}$. We write $\log x$ to mean $\log_2 x$, and we write $\ln x$ to mean $\log_e x$. We also write $a = (1 \pm b)c$ to mean $(1 - b)c < a < (1 + b)c$. For a natural number n we write $[n] \coloneqq \{1, \dots, n\}$. We write $x \ll y$ to mean that for any $y \in (0, 1]$ there exists an $x_0 \in (0, 1)$ such that for all $0 < x \leqslant x_0$ the subsequent statement holds. Hierarchies with more constants are defined similarly and should be read from right to left. Constants in hierarchies will always be real numbers in $(0, 1]$. Moreover, if $1/x$ appears in a hierarchy, this implicitly means that x is a natural number. More precisely, $1/x \ll y$ means that for any $y \in (0, 1]$, there exists an $x_0 \in \mathbb{N}$ such that for all $x \in \mathbb{N}$ with $x \geqslant x_0$ the subsequent statement holds. We assume large numbers to be integers if this does not affect the argument.

3.2 Probabilistic tools

In this subsection we collect some probabilistic definitions and results that we will need throughout the paper.

The *total variation distance* between two probability measures μ and ν on a finite set S is

$$d_{TV}(\mu, \nu) \coloneqq \sup\{|\mu(T) - \nu(T)| \colon T \subseteq S\}.$$

It is well known that the total variation distance satisfies

$$d_{TV}(\mu, \nu) = \frac{1}{2} \sum_{s \in S} |\mu(s) - \nu(s)| = \inf\{\mathbb{P}[X \neq Y]\}, \tag{3.1}$$

where the infimum is taken over coupled random variables X and Y having laws μ and ν respectively (see [11, page 119] for more details). We write $d_{TV}(X, Y)$ for the total variation distance between the laws of the random variables X and Y.

Next, we need an inequality of [8], which follows easily from Azuma's inequality.

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

Lemma 3.1 ([8, Lemma 5.3]). *Let $X_0, X_1, \dots$ be random variables taking values in a set V, and let $g\colon V \to \mathbb{R}$. Then, for any $t > 0$ and any $p, q \in \mathbb{N}$, we have*

$$\mathbb{P}\left[\left|\sum_{k=0}^{p}\Big(g(X_{k+q}) - \mathbb{E}[g(X_{k+q})|X_0,\dots,X_k]\Big)\right| > t\|g\|_\infty\sqrt{pq}\right] < 2qe^{-t^2/2}.$$

We will need the following Chernoff-type bound (see *e.g.* [5] and [15]).

Lemma 3.2. *Let X be a random variable with a binomial or hypergeometric distribution. Suppose $\mathbb{E}[X] > 0$ and let $t > 0$. Then $\mathbb{P}[X \leqslant \mathbb{E}[X] - t] \leqslant e^{-t^2/(2\mathbb{E}[X])}$.*

We conclude this section with a result which shows that most small sets in an (n, k, γ)-graph inherit the Dirac condition. We omit the proof; for details, see the appendix of the arXiv version of the paper.

Proposition 3.3. *Let $1/n \ll 1/m \ll \gamma, 1/k, 1/t, 1/\ell$, where $\ell \mid n$, and let G be an (n, k, γ)-graph. Let $\mathcal{P}$ be a partition of V into ℓ-sets and let $\mathcal{P}_0 \subseteq \mathcal{P}$ be of size $|\mathcal{P}_0| = t$. Pick $\mathcal{P}' \subseteq \mathcal{P} \setminus \mathcal{P}_0$ of size m uniformly at random. Then*

$$\mathbb{P}\big[G\big[\bigcup(\mathcal{P}_0 \cup \mathcal{P}')\big] \text{ is } \gamma/2\text{-Dirac}\big] \geqslant 1 - e^{-\sqrt{m}}.$$

3.3 Tight Hamilton-connectedness

Let G be an (n, k, γ)-graph and let P be a tight path in G. We say that P is a *tight Hamilton path* of G if $V(P) = V$. We say that G is *tight Hamilton-connected* if, for any disjoint $\overrightarrow{\mathbf{S}}, \overrightarrow{\mathbf{T}} \in (V)_{k-1}$, there is a tight Hamilton path of G which connects $\overrightarrow{\mathbf{S}}$ and $\overrightarrow{\mathbf{T}}$. We will deduce from the results in [26] that large (n, k, γ)-graphs are tight Hamilton-connected for $k \geqslant 3$. This will be important in the absorption step of our main argument, and also in the mixing part of our random walk analysis. We begin by stating the main theorem of [26].

Theorem 3.1 ([26, Theorem 1.1]). *Let $1/n \ll \gamma, 1/k$, where $k \geqslant 3$, and let G be an (n, k, γ)-graph. Then G contains a tight Hamilton cycle.*

The next lemma ensures the existence of an 'absorbing path' A, which can absorb small sets of vertices into its interior.

Lemma 3.4 ([26, Lemma 2.1]). *Let $1/n \ll \gamma, 1/k$, where $k \geqslant 3$, suppose that $\gamma \leqslant 1/(32k)$, set $\beta \coloneqq 2^{k-4}\gamma^{2k}n$, and let G be an (n, k, γ)-graph. Then there exists a tight path A in G with $|V(A)| \leqslant 16k\gamma^{k-1}n$ such that for every subset $U \subseteq V \setminus V(A)$ of size $|U| \leqslant \beta$, there is a tight path A_U in G with $V(A_U) = V(A) \cup U$ and such that A_U has the same ends as A.*

The next lemma will enable us to find constant-length tight paths between any disjoint pair of ordered $(k-1)$-sets of vertices.

Lemma 3.5 ([26, Lemma 2.4]). *Let $1/n \ll \gamma, 1/k$, where $k \geqslant 3$, and let G be an (n, k, γ)-graph. Then, for every $\overrightarrow{\mathbf{S}}, \overrightarrow{\mathbf{T}} \in (V)_{k-1}$ with $\mathbf{S} \cap \mathbf{T} = \emptyset$, there is an ℓ-path P in G with $\ell \leqslant 2k/\gamma^2$ that connects $\overrightarrow{\mathbf{S}}$ and $\overrightarrow{\mathbf{T}}$.*

We are now ready to prove that large (n, k, γ)-graphs are tight Hamilton-connected.

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

Lemma 3.6. *Let $1/n \ll \gamma, 1/k$, where $k \geqslant 2$, and let G be an (n,k,γ)-graph. Then G is tight Hamilton-connected.*

Proof. First note that this result follows easily from Dirac's theorem for the case $k=2$. Now suppose $k \geqslant 3$ and suppose without loss of generality that $\gamma > 0$ is sufficiently small in comparison to k. Let $\overrightarrow{\mathbf{S}}, \overrightarrow{\mathbf{T}} \in (V)_{k-1}$ be disjoint, and write $\overrightarrow{\mathbf{S}} = (s_1, \dots, s_{k-1})$. Set $\gamma' := 3\gamma/4$, so that $G' := G - (\mathbf{S} \cup \mathbf{T})$ is γ'-Dirac, and set $n' := n - 2(k-1)$. We apply Lemma 3.4 to G' to obtain a tight path A in G' with $|V(A)| \leqslant 16k(\gamma')^{k-1}n'$, with the properties as stated in Lemma 3.4. Choose a set $W \subseteq V \setminus (\mathbf{S} \cup \mathbf{T} \cup V(A))$ of size $(\gamma')^{3k}n'$ uniformly at random among all sets of that size. Then a simple application of Lemma 3.2 shows that with high probability, for all $\mathbf{M} \in \binom{V}{k-1}$ we have

$$|N_G(\mathbf{M}) \cap W| \geqslant (1/2 + \gamma/2)|W|,$$

and thus in particular, $G[W]$ is $\gamma/2$-Dirac. We fix such a choice of W. Set

$$G'' := G - (\mathbf{S} \cup \mathbf{T} \cup V(A) \cup W),$$

and notice that G'' is $\gamma/2$-Dirac.

We apply Theorem 3.1 to G'' to obtain a Hamilton cycle C of G''. Delete $k-1$ consecutive edges of C to obtain a Hamilton path P of G'' with ends $\overrightarrow{\mathbf{X}}$ and $\overrightarrow{\mathbf{Y}}$. We use the property that all $(k-1)$-tuples in G have high codegree in W to find, for each $i \in [k-1]$ in turn, a vertex $v_i \in W \setminus \{v_1, \dots, v_{i-1}\}$ such that $\{s_i, \dots, s_{k-1}, v_1, \dots, v_i\}$ is an edge. Let $\overrightarrow{\mathbf{S}'} := (v_{k-1}, \dots, v_1)$, so that we have found a $(k-1)$-path P_S with ends $\overrightarrow{\mathbf{S}}$ and $\overrightarrow{\mathbf{S}'}$. Let $\overrightarrow{\mathbf{A}}_1$ and $\overrightarrow{\mathbf{A}}_2$ be the ends of A. We similarly find mutually disjoint $\overrightarrow{\mathbf{X}'}, \overrightarrow{\mathbf{Y}'}, \overrightarrow{\mathbf{A}'_1}, \overrightarrow{\mathbf{A}'_2}, \overrightarrow{\mathbf{T}'} \in (W)_{k-1}$ and $(k-1)$-paths $P_X, P_Y, P_{A_1}, P_{A_2}, P_T$ with the corresponding pairs of ends. Since $G[W]$ is $\gamma/2$-Dirac, we can apply Lemma 3.5 to obtain a path P_{SX} of length at most $8k/\gamma^2$ in $G[W]$ which connects $\overleftarrow{\mathbf{S}'}$ and $\overleftarrow{\mathbf{X}'}$. Since P_{SX} contains so few vertices, we can repeat the process to find disjoint paths P_{YA_1} and P_{A_2T} in $G[W]$ with ends $\overleftarrow{\mathbf{Y}'}$ and $\overleftarrow{\mathbf{A}'_1}$, and $\overleftarrow{\mathbf{A}'_2}$ and $\overleftarrow{\mathbf{T}'}$, respectively. Let

$$W' := W \setminus (V(P_{SX}) \cup V(P_{YA_1}) \cup V(P_{A_2T})).$$

We apply the absorbing property of A to obtain a tight path $A_{W'}$ in G with $V(A_{W'}) = V(A) \cup W'$, such that $A_{W'}$ has ends $\overrightarrow{\mathbf{A}_1}$ and $\overrightarrow{\mathbf{A}_2}$.

Then $P_S \cup P_{SX} \cup P_X \cup P \cup P_Y \cup P_{YA_1} \cup P_{A_1} \cup A_{W'} \cup P_{A_2} \cup P_{A_2T} \cup P_T$ is a tight Hamilton path in G which connects $\overrightarrow{\mathbf{S}}$ and $\overrightarrow{\mathbf{T}}$. □

4. Normal perfect fractional matchings

Let $k \geqslant 2$ and let G be a k-graph on n vertices. We say that an edge weighting $\mathbf{x}\colon E(G) \to \mathbb{R}^+$ is *C-normal* if

$$\frac{1}{Cn^{k-1}} \leqslant \mathbf{x}(e) \leqslant \frac{C}{n^{k-1}} \quad \text{for each } e \in E(G). \tag{4.1}$$

In this section we adapt some ideas of [6] to show that an (n,k,γ)-graph G admits a normal perfect fractional matching (see Lemma 4.2). This will be an essential tool in our random walk analysis for showing that the random walk is roughly equally likely to visit any vertex. The idea is to construct a perfect fractional matching of G in which the weight of any edge e is set to be the probability that e is included in a uniformly random perfect matching of G. (A sufficiently large (n,k,γ)-graph with $k \mid n$ has at least one perfect matching [20, 27].) A crucial feature of this approach is that any edge e is roughly equally likely to be included in a uniformly random perfect

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

matching of G. We show this using the so-called 'switching method' in a similar way as in [6]. Let $k \geqslant 2$, let G be a k-graph, let $e \in E(G)$, and let M_ℓ be a perfect matching of G containing precisely ℓ edges intersecting e. Supposing $0 \leqslant \ell \leqslant k-1$, we define an (e, M_ℓ)-*upswitching* to be a matching Y of G satisfying

(i) $e \subseteq V(Y)$,
(ii) Y contains precisely $\ell+1$ edges intersecting e,
(iii) for all $e' \in M_\ell$, we have either $e' \subseteq V(Y)$ or $e' \cap V(Y) = \emptyset$.

Supposing instead that $\ell \in [k]$, we define an (e, M_ℓ)-*downswitching* to be a matching Y of G satisfying

(i) $e \subseteq V(Y)$,
(ii) Y contains precisely $\ell-1$ edges intersecting e,
(iii) for all $e' \in M_\ell$, we have either $e' \subseteq V(Y)$ or $e' \cap V(Y) = \emptyset$.

Note that if Y is an (e, M_ℓ)-upswitching, then we can obtain a new perfect matching M' from M_ℓ by replacing $M_\ell[V(Y)]$ with Y. Then M' contains exactly $\ell+1$ edges intersecting e. Similarly, if Y is an (e, M_ℓ)-downswitching, then M' has exactly $\ell-1$ edges intersecting e.

Lemma 4.1. *Let $1/n \ll 1/C \ll \gamma, 1/k$, where $k \geqslant 2$ and $k \mid n$. Let G be an (n, k, γ)-graph, and let M be a uniformly random perfect matching of G. Then, for each $e \in E(G)$, we have*

$$\frac{1}{Cn^{k-1}} \leqslant \mathbb{P}[e \in M] \leqslant \frac{C}{n^{k-1}}. \tag{4.2}$$

Proof. Choose new integers m and B satisfying $1/n \ll 1/C \ll 1/B \ll 1/m \ll \gamma, 1/k$, and fix $e \in E(G)$. For each integer $\ell \in [k]$, let $\mathcal{M}_\ell$ be the set of perfect matchings of G containing precisely ℓ edges intersecting e. Note that

$$\mathbb{P}[e \in M] = |\mathcal{M}_1|/(|\mathcal{M}_1| + \cdots + |\mathcal{M}_k|)$$

and recall that there is at least one perfect matching of G since G is γ-Dirac (so the denominator here is non-zero). We first bound $|\mathcal{M}_\ell|/|\mathcal{M}_{\ell+1}|$ from above and below for each $\ell \in [k-1]$, and (4.2) will follow quickly. Let $\ell \in [k-1]$. We define an auxiliary bipartite multigraph $G^{\uparrow}_{e,\ell}$ with vertex bipartition $(\mathcal{M}_\ell, \mathcal{M}_{\ell+1})$. For each $M_\ell \in \mathcal{M}_\ell$ and each (e, M_ℓ)-upswitching Y of *size* m (containing precisely m edges), we add an edge in $G^{\uparrow}_{e,\ell}$ from M_ℓ to the matching $M_{\ell+1} \in \mathcal{M}_{\ell+1}$ obtained by replacing $M_\ell[V(Y)]$ with Y. Write $\delta^{e,\uparrow}_{\mathcal{M}_\ell}$ to denote the minimum degree in $G^{\uparrow}_{e,\ell}$ over all $M_\ell \in \mathcal{M}_\ell$, and write $\Delta^{e,\uparrow}_{\mathcal{M}_{\ell+1}}$ to denote the maximum degree in $G^{\uparrow}_{e,\ell}$ over all $M_{\ell+1} \in \mathcal{M}_{\ell+1}$. By double-counting $|E(G^{\uparrow}_{e,\ell})|$, we obtain

$$|\mathcal{M}_\ell|/|\mathcal{M}_{\ell+1}| \leqslant \Delta^{e,\uparrow}_{\mathcal{M}_{\ell+1}}/\delta^{e,\uparrow}_{\mathcal{M}_\ell}.$$

To bound $\Delta^{e,\uparrow}_{\mathcal{M}_{\ell+1}}$, we fix $M_{\ell+1} \in \mathcal{M}_{\ell+1}$ and bound the number of pairs (M_ℓ, Y), where $M_\ell \in \mathcal{M}_\ell$ and Y is an (e, M_ℓ)-upswitching of size m that produces $M_{\ell+1}$. Note that any such Y must contain all vertices in the $\ell+1$ edges of $M_{\ell+1}$ intersecting e, and there are at most $n^{m-\ell-1}$ choices for the other $m-\ell-1$ edges of $M_{\ell+1}$ whose vertices are to be included in $V(Y)$. Once $V(Y)$ is fixed, there are at most $(mk)!$ choices for $M_\ell[V(Y)]$ (and hence for M_ℓ). Thus we have

$$\Delta^{e,\uparrow}_{\mathcal{M}_{\ell+1}} \leqslant (mk)!n^{m-\ell-1}.$$

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

To bound $\delta^{e,\uparrow}_{\mathcal{M}_\ell}$, we fix $M_\ell \in \mathcal{M}_\ell$ and bound the number of (e, M_ℓ)-upswitchings of size m from below. Let

$$U(M_\ell) \coloneqq \{e' \in M_\ell \colon e \cap e' \neq \emptyset\}.$$

Note that any (e, M_ℓ)-upswitching Y of size m must include all the vertices in $U(M_\ell)$, and there are $\binom{n/k-\ell}{m-\ell}$ choices for the remaining $m-\ell$ edges of M_ℓ whose vertices are to be included in $V(Y)$. We apply Proposition 3.3 (with $\mathcal{P} = M_\ell$, $\mathcal{P}_0 = U(M_\ell)$, and with $m-\ell, k, \ell$ playing the roles of m, ℓ, t, respectively) to deduce that there are at least

$$(1 - e^{-\sqrt{m-\ell}})\binom{n/k-\ell}{m-\ell} \geqslant (mk)^{-m} n^{m-\ell}$$

choices of $X \subseteq M_\ell \setminus U(M_\ell)$ of size $m-\ell$ such that $G[V(X \cup U(M_\ell))]$ is $\gamma/2$-Dirac. Note that for each such X, we may first choose a matching U' of size $\ell+1$ in $G[V(X \cup U(M_\ell))]$ such that $e \subseteq V(U')$ and e intersects every edge in U', and then choose a perfect matching Y' of $G[V(X \cup U(M_\ell)) \setminus V(U')]$. Then $Y \coloneqq Y' \cup U'$ is an (e, M_ℓ)-upswitching of size m, unique to this choice of X. We deduce that $\delta^{e,\uparrow}_{\mathcal{M}_\ell} \geqslant (mk)^{-m} n^{m-\ell}$, and conclude that $|\mathcal{M}_\ell|/|\mathcal{M}_{\ell+1}| \leqslant (mk)!(mk)^m/n \leqslant B/n$.

We now bound the terms $|\mathcal{M}_\ell|/|\mathcal{M}_{\ell+1}|$ from below analogously. Let $\ell \in [k-1]$. We define an auxiliary bipartite multigraph $G^{\downarrow}_{e,\ell+1}$ with vertex bipartition $(\mathcal{M}_\ell, \mathcal{M}_{\ell+1})$. For each $M_{\ell+1} \in \mathcal{M}_{\ell+1}$ and each $(e, M_{\ell+1})$-downswitching Y of size m, we add an edge in $G^{\downarrow}_{e,\ell+1}$ from $M_{\ell+1}$ to the matching $M_\ell \in \mathcal{M}_\ell$ obtained by replacing $M_{\ell+1}[V(Y)]$ with Y. Let $\delta^{e,\downarrow}_{\mathcal{M}_{\ell+1}}$ denote the minimum degree in $G^{\downarrow}_{e,\ell+1}$ among all $M_{\ell+1} \in \mathcal{M}_{\ell+1}$, and let $\Delta^{e,\downarrow}_{\mathcal{M}_\ell}$ denote the maximum degree in $G^{\downarrow}_{e,\ell+1}$ among all $M_\ell \in \mathcal{M}_\ell$. It is easy to see that $\Delta^{e,\downarrow}_{\mathcal{M}_\ell} \leqslant (mk)! n^{m-\ell}$. Now fix $M_{\ell+1} \in \mathcal{M}_{\ell+1}$ and let

$$U(M_{\ell+1}) \coloneqq \{e' \in M_{\ell+1} \colon e \cap e' \neq \emptyset\}.$$

We apply Lemma 3.3 again (with $\mathcal{P} = M_{\ell+1}$, $\mathcal{P}_0 = U(M_{\ell+1})$, and with $m-\ell-1, k, \ell+1$ playing the roles of m, ℓ, t, respectively) to deduce that

$$\delta^{e,\downarrow}_{\mathcal{M}_{\ell+1}} \geqslant (mk)^{-m} n^{m-\ell-1},$$

and thus

$$|\mathcal{M}_\ell|/|\mathcal{M}_{\ell+1}| \geqslant 1/((mk)!(mk)^m n) \geqslant 1/(Bn).$$

Finally, note that

$$\mathbb{P}[e \in M] = \frac{|\mathcal{M}_1|}{|\mathcal{M}_1| + \cdots + |\mathcal{M}_k|} \leqslant \frac{|\mathcal{M}_1|}{|\mathcal{M}_k|} = \frac{|\mathcal{M}_1|}{|\mathcal{M}_2|} \cdot \frac{|\mathcal{M}_2|}{|\mathcal{M}_3|} \cdots \frac{|\mathcal{M}_{k-1}|}{|\mathcal{M}_k|} \leqslant \frac{B^{k-1}}{n^{k-1}} \leqslant \frac{C}{n^{k-1}},$$

and similarly

$$\mathbb{P}[e \in M] \geqslant |\mathcal{M}_1|/(k|\mathcal{M}_k|) \geqslant 1/(kB^{k-1}n^{k-1}) \geqslant 1/(Cn^{k-1}).$$ □

Finally, we use Lemma 4.1 to show that an (n, k, γ)-graph admits a normal perfect fractional matching.

Lemma 4.2. *Let $1/n \ll 1/C \ll \gamma, 1/k$, where $k \geqslant 2$, and let G be an (n, k, γ)-graph. Then there exists a C-normal perfect fractional matching of G.*

Proof. Let i be the unique integer in $\{0, 1, \ldots, k-1\}$ satisfying $n \equiv i \bmod k$. For each $\mathbf{S} \in \binom{V}{i}$, let $G_{\mathbf{S}} \coloneqq G - \mathbf{S}$. We define an edge weighting $\mathbf{x}_{\mathbf{S}} \colon E(G_{\mathbf{S}}) \to \mathbb{R}^+$ by setting $\mathbf{x}_{\mathbf{S}}(e) \coloneqq \mathbb{P}[e \in M_{\mathbf{S}}]$ for each

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

$e \in E(G_{\mathbf{S}})$, where $M_{\mathbf{S}}$ is a uniformly random perfect matching in $G_{\mathbf{S}}$. We define an edge weighting $\mathbf{x}\colon E(G) \to \mathbb{R}^+$ by setting

$$\mathbf{x}(e) \coloneqq \binom{n-1}{i}^{-1} \sum_{\mathbf{S} \in \binom{V}{i}} \mathbf{x}_{\mathbf{S}}(e)$$

for each $e \in E(G)$, where we set $\mathbf{x}_{\mathbf{S}}(e)$ to be 0 for each $\mathbf{S}$ such that $e \notin E(G_{\mathbf{S}})$. Then, by Lemma 4.1, $\mathbf{x}$ is the desired C-normal perfect fractional matching of G. □

5. Counting short paths

The aim of this section is to prove the following lemma, which guarantees many short tight paths in a γ-Dirac k-graph G, such that the γ-Dirac property of the graph G' obtained from deleting any such path is still essentially intact.

Lemma 5.1 (Iteration Lemma). *Let $1/n \ll c \ll \gamma, 1/k$ where $k \geqslant 2$, let G be an (n,k,γ)-graph, and let $\overrightarrow{\mathbf{S}} \in (V)_{k-1}$. There exists a set $\mathcal{P}$ of $\sqrt{n}$-paths in G such that:*

(i) $|\mathcal{P}| \geqslant (cn)^{\sqrt{n}}$,
(ii) $\overrightarrow{\mathbf{S}}$ *is an end of each* $P \in \mathcal{P}$,
(iii) *if* $P \in \mathcal{P}$ *and* $\overrightarrow{\mathbf{T}}$ *is the non-*$\overrightarrow{\mathbf{S}}$ *end of* P, *then* $G' \coloneqq G - (V(P) \setminus \mathbf{T})$ *satisfies* $\delta(G') \geqslant (1/2 + \gamma - n^{-2/3})(n - \sqrt{n})$.

We now provide some important definitions and sketch the proof of Lemma 5.1. We then collect together a number of technical lemmas, and finally use these results to prove Lemma 5.1.

5.1 Random walk notation

We first define some random walks which will be of central importance to the proof of Lemma 5.1. Let $k \geqslant 2$ be an integer, let G be a k-graph, and let $\mathbf{x}\colon E(G) \to \mathbb{R}^+$ be a positive edge weighting function. Each of our random walks $\mathcal{Z} = (Z_{-(k-2)}, Z_{-(k-3)}, \dots)$ on V will begin with an ordered $(k-1)$-tuple $(Z_{-(k-2)}, \dots, Z_0)$ chosen according to some probability distribution $\mu\colon (V)_{k-1} \to \mathbb{R}^+$. We say that μ is the *initial distribution* of $\mathcal{Z}$. Random vertices will then be added one by one to each $\mathcal{Z}$ according to the transition probabilities of $\mathcal{Z}$. Suppose we are given a random walk $\mathcal{Z} = (Z_{-(k-2)}, \dots, Z_{j-1})$ on V, up to time $j-1$. Then we say that *semiviable* vertices for step j are those vertices $v \in V$ satisfying $\{Z_{j-(k-1)}, \dots, Z_{j-1}\} \cup \{v\} \in E(G)$. We say that *viable* vertices for step j are those vertices $v \in V$ which are semiviable and satisfy $v \notin \{Z_{-(k-2)}, \dots, Z_{j-1}\}$. Let Q_j and R_j denote the sets of semiviable and viable vertices for step j, given the random walk up to time $j-1$, respectively.

We say that a random walk $\mathcal{X} = (X_{-(k-2)}, X_{-(k-3)}, \dots)$ on the vertices of G, with any initial distribution μ, is a *self-avoiding* $\mathbf{x}$-*walk* to mean that the transition probabilities of $\mathcal{X}$ for $j \geqslant 1$ are defined for all $v \in V$ by

$$\mathbb{P}[X_j = v \mid X_{-(k-2)}, \dots, X_{j-1}] \coloneqq \frac{\mathbf{x}(\{X_{j-(k-1)}, \dots, X_{j-1}\} \cup \{v\}) \mathbb{1}_{R_j}(v)}{\sum_{w \in R_j} \mathbf{x}(\{X_{j-(k-1)}, \dots, X_{j-1}\} \cup \{w\})},$$

whenever R_j is non-empty, otherwise we terminate the walk. Here and throughout, we define $\mathbf{x}(S) \coloneqq 0$ for any $S \notin E(G)$. Note that $\mathcal{X} = (X_{-(k-2)}, X_{-(k-3)}, \dots)$ is equivalent to the random walk $\mathcal{X}^{(k-1)} = (\overrightarrow{\mathbf{X}}_0, \overrightarrow{\mathbf{X}}_1, \dots)$, where each $\overrightarrow{\mathbf{X}}_i$ is the ordered $(k-1)$-tuple $\overrightarrow{\mathbf{X}}_i = (X_{i-(k-2)}, \dots, X_i)$.

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

We thus refer to both $\mathcal{X}$ and $\mathcal{X}^{(k-1)}$ as the self-avoiding $\mathbf{x}$-walk on G with initial distribution μ, since they are reformulations of each other.

Suppose now that G has no isolated $(k-1)$-tuples (this will always be true for us). We say that a random walk $\mathcal{Y} = (Y_{-(k-2)}, Y_{-(k-3)}, \dots)$ on the vertices of G (or $\mathcal{Y}^{(k-1)} = (\overrightarrow{\mathbf{Y}}_0, \overrightarrow{\mathbf{Y}}_1, \dots)$ on $(V)_{k-1}$, where $\overrightarrow{\mathbf{Y}}_i \coloneqq (Y_{i-(k-2)}, \dots, Y_i)$), with any initial distribution μ, is a *simple* $\mathbf{x}$*-walk* to mean that the transition probabilities of $\mathcal{Y}$ for $j \geqslant 1$ are defined by

$$\mathbb{P}[Y_j = v \mid Y_{-(k-2)}, \dots, Y_{j-1}] \coloneqq \frac{\mathbf{x}(\{Y_{j-(k-1)}, \dots, Y_{j-1}\} \cup \{v\})\mathbb{1}_{Q_j}(v)}{\sum_{w \in Q_j} \mathbf{x}(\{Y_{j-(k-1)}, \dots, Y_{j-1}\} \cup \{w\})} \tag{5.1}$$

for all $v \in V$. Note that $\mathcal{Y}^{(k-1)}$ is a Markov chain on $(V)_{k-1}$ because the transition probabilities at any time depend only on the current state. When the stationary distribution of $\mathcal{Y}^{(k-1)}$ exists and is unique, we denote it by π, and we say that a simple $\mathbf{x}$-walk $\mathcal{W}^{(k-1)} = (\overrightarrow{\mathbf{W}}_0, \overrightarrow{\mathbf{W}}_1, \dots)$ on the ordered $(k-1)$-tuples of V is the *stationary* $\mathbf{x}$*-walk* on G if the initial distribution is π. Again we have that $\mathcal{W}^{(k-1)} = (\overrightarrow{\mathbf{W}}_0, \overrightarrow{\mathbf{W}}_1, \dots)$ is equivalent to the walk $\mathcal{W} = (W_{-(k-2)}, W_{-(k-3)}, \dots)$ on the vertices of G, where each $\overrightarrow{\mathbf{W}}_i$ is the ordered $(k-1)$-tuple $\overrightarrow{\mathbf{W}}_i = (W_{i-(k-2)}, \dots, W_i)$. We also call $\mathcal{W}$ the stationary $\mathbf{x}$-walk, and use the vertex (or tuple) version whenever it is more convenient. Finally, whenever the initial distribution μ of $\mathcal{X}$ (or $\mathcal{Y}$) satisfies $\mu(\overrightarrow{\mathbf{S}}) = 1$ for some $\overrightarrow{\mathbf{S}} \in (V)_{k-1}$, we say that $\mathcal{X}$ (or $\mathcal{Y}$) has *starting tuple* $\overrightarrow{\mathbf{S}}$.

5.2 Further notation and sketch of the proof of Lemma 5.1

We now describe our approach to proving Lemma 5.1. Introduce a new constant C satisfying $1/n \ll c \ll 1/C \ll \gamma, 1/k$, and let G be an (n, k, γ)-graph. By Lemma 4.2, there exists a C-normal perfect fractional matching $\mathbf{x}$ of G. We fix such a C-normal $\mathbf{x}$ throughout this proof sketch. We will analyse a self-avoiding $\mathbf{x}$-walk $\mathcal{X}$ on G with starting tuple $\overrightarrow{\mathbf{S}} \in (V)_{k-1}$. We stop the walk after time $\kappa \coloneqq \sqrt{n}$, so that we may write $\mathcal{X} = (X_{-(k-2)}, \dots, X_\kappa)$. Note that each outcome of $\mathcal{X}$ will correspond to a tight κ-path in G, with $\overrightarrow{\mathbf{S}}$ as one end.

We define

$$V_j \coloneqq V \setminus \{X_{-(k-2)}, \dots, X_{j-(k-1)}\}$$

to be the set of all vertices of G except for all vertices $\mathcal{X}^{(k-1)}$ has visited strictly before $\overrightarrow{\mathbf{X}}_j$. We say that V_j is the *residual vertex set* of G at time j. We also define $G_j \coloneqq G[V_j]$ and say that G_j is the *residual graph* at time j. We also write $\mathcal{X}(j)$ to denote the walk $\mathcal{X}$ up to time j, specifically $\mathcal{X}(j) \coloneqq (X_{-(k-2)}, \dots, X_j)$.

We will show that it is likely that the γ-Dirac property of the residual graph G_κ is still essentially intact, by showing that it is likely that the vertices that $\mathcal{X}$ visits look roughly like a uniformly random subset of V (see Lemma 5.7). For this, we will use the following 'tracking functions' to monitor the progress of $\mathcal{X}$, with respect to how the codegree of each $(k-1)$-tuple in the residual graph deteriorates over time.

For each unordered $(k-1)$-tuple $\mathbf{S} \in \binom{V}{k-1}$, we define a function $g_{\mathbf{S}} \colon V \to \mathbb{R}^+$ by setting $g_{\mathbf{S}}(v) \coloneqq \mathbb{1}_{N_G(\mathbf{S})}(v)$ for each $v \in V$, so that, in particular, if $\mathbf{S} \subseteq V_j$ then $g_{\mathbf{S}}(V_j) = d_{G_j}(\mathbf{S})$. We call the set $\mathcal{F} \coloneqq \{g_{\mathbf{S}} \colon \mathbf{S} \in \binom{V}{k-1}\}$ the set of *tracking functions* of G. We say that $\mathcal{X} = (X_{-(k-2)}, \dots, X_\kappa)$ is *good* if

$$\sum_{i=-(k-2)}^{\kappa} g_{\mathbf{S}}(X_i) = \frac{\kappa}{n} g_{\mathbf{S}}(V) \pm n^{3/10} \quad \text{for all } g_{\mathbf{S}} \in \mathcal{F}. \tag{5.2}$$

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

Thus, to say that $\mathcal{X}$ is good is to say that the set of $\kappa + k - 1$ vertices that $\mathcal{X}$ visits look roughly like a uniformly random subset of V, with respect to the codegrees of all $(k-1)$-tuples. In particular, for all $(k-1)$-tuples $\mathbf{S} \in \binom{V}{k-1}$, the proportion of vertices of $N_G(\mathbf{S})$ visited by $\mathcal{X}$ is approximately κ/n, and this is the property that will allow us to deduce condition (iii) of Lemma 5.1.

Let $(\overrightarrow{\mathbf{X}}_a, \dots, \overrightarrow{\mathbf{X}}_b)$ be a suitable interval of $\mathcal{X}^{(k-1)}$, and let $\mathcal{Y}^{(k-1)} = (\overrightarrow{\mathbf{Y}}_0 = \overrightarrow{\mathbf{X}}_a, \overrightarrow{\mathbf{Y}}_1, \dots)$ be the simple $\mathbf{x}$-walk on G_a with starting tuple $\overrightarrow{\mathbf{X}}_a$. To show that the walk $\mathcal{X}$ is likely to be good, the following will be the main steps.

(i) First we show that the behaviour of $(\overrightarrow{\mathbf{X}}_a, \dots)$ follows closely that of $\mathcal{Y}^{(k-1)}$ by exhibiting a coupling of the two walks such that the probability of $\mathcal{X}^{(k-1)}$ and $\mathcal{Y}^{(k-1)}$ being different is acceptably small, provided $b - a$ is small.
(ii) Next we see that $\mathcal{Y}^{(k-1)}$ *mixes* (converges to its stationary distribution π) rapidly.
(iii) We also show that the stationary $\mathbf{x}$-walk $\mathcal{W} = (W_0, W_1, \dots)$ satisfies $\mathbb{P}[W_i = v] \approx 1/|V_a|$ for each $v \in V_a$.

Putting the above together, we see that even for small q, the distribution of each X_i, given the walk to time $i - q$, is typically close to the uniform distribution on V_{i-q}, and thus, for each tracking function $g \in \mathcal{F}$, we have

$$\mathbb{E}[g(X_i)] \approx \frac{1}{n-(i-q)} g(V_{i-q}) \approx \frac{1}{n} g(V).$$

Lastly, then:

(iv) We show that the actual values of the quantities $\sum_{a \leqslant i \leqslant b} g(X_i)$ are very likely to be close to their expectations, by using Lemma 3.1.

This completes the sketch of the proof that $\mathcal{X}$ is likely to be good. It only remains to count the number of good walks (outcomes of) $\mathcal{X}$. The count will be obtained by simply dividing a lower bound for the probability that $\mathcal{X}$ is good, by an upper bound for the probability of obtaining any specific outcome of $\mathcal{X}$. This completes the sketch of the proof of Lemma 5.1.

5.3 Random walk analysis

In this subsection we collect some of the tools that we will use to prove Lemma 5.1. We first define some convenient terminology for edge weightings. Let $\mathbf{x}\colon E(G) \to \mathbb{R}^+$ be a positive edge weighting of a k-graph G. We say that $\mathbf{x}$ is a_1-*lower-balanced* if, for all non-isolated $\mathbf{S} \in \binom{V}{k-1}$ and all $v \in N_G(\mathbf{S})$, we have that

$$\mathbf{x}(\mathbf{S} \cup \{v\}) / \left(\sum_{v' \in V} \mathbf{x}(\mathbf{S} \cup \{v'\}) \right) \geqslant a_1.$$

That is, all possible $\mathbf{x}$-walk transition probabilities are bounded below by a_1. Similarly, we say that $\mathbf{x}$ is a_2-*upper-balanced* if, for all non-isolated $\mathbf{S} \in \binom{V}{k-1}$ and all $v \in N_G(\mathbf{S})$, we have

$$\mathbf{x}(\mathbf{S} \cup \{v\}) / \left(\sum_{v' \in V} \mathbf{x}(\mathbf{S} \cup \{v'\}) \right) \leqslant a_2.$$

We say that $\mathbf{x}$ is (a_1, a_2)-*balanced* if $\mathbf{x}$ is a_1-lower-balanced and a_2-upper-balanced. We now give a simple result which shows that we may couple the self-avoiding $\mathbf{x}$-walk and the simple $\mathbf{x}$-walk to behave very similarly over small distances.

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

Lemma 5.2. *Let $k \geqslant 2$, let G be a k-graph, let $\overrightarrow{\mathbf{M}} \in (V)_{k-1}$, and let $\mathbf{x}\colon E(G) \to \mathbb{R}^+$ be a positive r-upper-balanced edge weighting. Let $\mathcal{X} = (X_{-(k-2)}, \dots)$ and $\mathcal{Y} = (Y_{-(k-2)}, \dots)$ be, respectively, the self-avoiding $\mathbf{x}$-walk and the simple $\mathbf{x}$-walk on G, each with starting tuple $\overrightarrow{\mathbf{M}}$. Then, for any positive integer $q \leqslant \delta(G)$, we have*

$$d_{TV}(X_q, Y_q) \leqslant q^2 r.$$

Proof. If $1 \leqslant i \leqslant \delta(G)$, and $\mathcal{X}$ and $\mathcal{Y}$ agree up to time $i-1$, say $X_j = Y_j = v_j \in V$ for each $j \in \{-k+2, \dots, i-1\}$, then we couple at the next step so that X_i coincides with Y_i whenever the choice of Y_i is a viable choice for X_i, which is to say that Y_i is not a vertex already seen. So with this coupling, for any positive integer $i \leqslant \delta(G)$ we have

$$\mathbb{P}[X_i \neq Y_i \mid X_j = Y_j \text{ for all } j \in \{-k+2, \dots, i-1\}] = \mathbb{P}[Y_i \in \{v_{-(k-2)}, \dots, v_{i-k}\}] \leqslant ir.$$

Thus, for any positive integer $q \leqslant \delta(G)$, we obtain

$$\mathbb{P}[X_q \neq Y_q] \leqslant \sum_{i=1}^{q} \mathbb{P}[X_i \neq Y_i \mid X_j = Y_j \text{ for all } j \in \{-k+2, \dots, i-1\}] \leqslant q^2 r.$$

The desired result now follows from (3.1). □

We now aim to show that $\mathcal{Y}^{(k-1)}$ mixes rapidly. The key part of the proof will be the following argument that $\mathcal{Y}^{(k-1)}$ has many different choices for how to arrive at a specified target ordered tuple $\overrightarrow{\mathbf{T}} \in (V)_{k-1}$ in a fixed number of steps.

Lemma 5.3. *Let $1/n \ll \zeta \ll 1/\ell \ll \gamma, 1/k$, where $k \geqslant 2$, and let G be an (n, k, γ)-graph. For any $\overrightarrow{\mathbf{S}}, \overrightarrow{\mathbf{T}} \in (V)_{k-1}$, there are at least $\zeta n^{\ell-(k-1)}$ ℓ-walks from $\overrightarrow{\mathbf{S}}$ to $\overrightarrow{\mathbf{T}}$ in G. Further, if $\mathbf{S} \cap \mathbf{T} = \emptyset$, then there are at least $\zeta n^{\ell-(k-1)}$ ℓ-paths in G which connect $\overrightarrow{\mathbf{S}}$ and $\overleftarrow{\mathbf{T}}$.*

Proof. Let $\overrightarrow{\mathbf{S}}, \overrightarrow{\mathbf{T}} \in (V)_{k-1}$, set $a \coloneqq \ell - (k-1)$, let

$$\mathcal{Q} \coloneqq \binom{V \setminus (\mathbf{S} \cup \mathbf{T})}{a},$$

let $G_Q \coloneqq G[\mathbf{S} \cup Q \cup \mathbf{T}]$ for each $Q \in \mathcal{Q}$, and define

$$\mathcal{Q}' \coloneqq \{Q \in \mathcal{Q}\colon G_Q \text{ is } (\gamma/2)\text{-Dirac}\}.$$

We apply Proposition 3.3 (with $\mathcal{P} = \{\{v\}\colon v \in V\}$, $\mathcal{P}_0 = \{\{v\}\colon v \in \mathbf{S} \cup \mathbf{T}\}$, and with a, 1, $|\mathbf{S} \cup \mathbf{T}|$ playing the roles of m, ℓ, t, respectively) to deduce that $|\mathcal{Q}'| \geqslant |\mathcal{Q}|/2$. Fix $Q \in \mathcal{Q}'$ and write $\overrightarrow{\mathbf{T}} = (t_1, \dots, t_{k-1})$. Since $G[Q \cup \mathbf{T}]$ is $\gamma/4$-Dirac, for each $i \in [k-1]$ in turn we may find a vertex $v_i \in Q \setminus \{v_1, \dots, v_{i-1}\}$ such that $\{v_1, \dots, v_i\} \cup \{t_1, \dots, t_{k-i}\}$ is an edge. Write $\overrightarrow{\mathbf{T}'} \coloneqq (v_{k-1}, \dots, v_1)$. Thus we obtain a $(k-1)$-path P_Q^2 in G_Q which connects $\overrightarrow{\mathbf{T}'}$ and $\overleftarrow{\mathbf{T}}$. Since $G[\mathbf{S} \cup Q]$ is $\gamma/4$-Dirac, we may apply Lemma 3.6 (with $\gamma/4$ playing the role of γ) to find an a-path P_Q^1 in $G[\mathbf{S} \cup Q]$ which connects $\overrightarrow{\mathbf{S}}$ and $\overleftarrow{\mathbf{T}'}$, and the obvious concatenation of P_Q^1 and P_Q^2 is an ℓ-walk W_Q from $\overrightarrow{\mathbf{S}}$ to $\overrightarrow{\mathbf{T}}$ in G_Q. It is clear that these ℓ-walks W_Q are distinct for different choices of $Q \in \mathcal{Q}'$. It thus suffices to observe that

$$|\mathcal{Q}'| \geqslant \frac{1}{2}|\mathcal{Q}| \geqslant \zeta n^{\ell-(k-1)}.$$

If $\mathbf{S} \cap \mathbf{T} = \emptyset$, then the walks W_Q do not revisit vertices and thus correspond to ℓ-paths in G which connect $\overrightarrow{\mathbf{S}}$ and $\overleftarrow{\mathbf{T}}$. □

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

Lemma 5.3 shows that in an (n,k,γ)-graph G, the Markov chain $\mathcal{Y}^{(k-1)}$ is irreducible, and thus there is a unique stationary distribution π of $\mathcal{Y}^{(k-1)}$. We will use this fact without stating it from now on. We note that it also follows from Lemma 5.3 that $\mathcal{Y}^{(k-1)}$ is aperiodic, which implies that the distribution of $\overrightarrow{\mathbf{Y}}_t$ converges to π as $t\to\infty$. However, we need something stronger, namely that this convergence occurs quickly. This is achieved by the following lemma, which shows that $\mathcal{Y}^{(k-1)}$ mixes rapidly.

Lemma 5.4 (Mixing Lemma). *Let $1/n \ll 1/\lambda \ll \gamma,\tau,1/k$, where $k\geqslant 2$, let G be an (n,k,γ)-graph, let $\mu\colon (V)_{k-1}\to\mathbb{R}$ be a probability distribution, and let $\mathbf{x}\colon E(G)\to\mathbb{R}^+$ be a positive (τ/n)-lower-balanced edge weighting. Let $\mathcal{Y}=(Y_{-(k-2)},\dots)$ be the simple $\mathbf{x}$-walk on G with initial distribution μ, and let $\mathcal{W}=(W_{-(k-2)},\dots)$ be the stationary $\mathbf{x}$-walk on G. For any $\eta>0$, if $q\geqslant \lambda\ln(1/\eta)$, then $d_{TV}(Y_q,W_q)<\eta$.*

Proof. Choose new constants ζ and ℓ satisfying $1/n\ll 1/\lambda\ll\zeta\ll 1/\ell\ll\gamma,\tau,1/k$. We proceed by using Lemma 5.3 to show that for two simple $\mathbf{x}$-walks $\mathcal{Z}^{(k-1)}$ and $\mathcal{Z}'^{(k-1)}$ on G given any initial distributions, we can find a coupling such that $\mathcal{Z}^{(k-1)}$ and $\mathcal{Z}'^{(k-1)}$ are relatively likely to meet after ℓ steps. Using this, we then exhibit a coupling of $\mathcal{Y}^{(k-1)}$ and $\mathcal{W}^{(k-1)}$ that will allow us to use (3.1) to upper-bound $d_{TV}(Y_q,W_q)$.

Let $\mathcal{Z}^{(k-1)}$ be a simple $\mathbf{x}$-walk on G with any initial distribution ϕ, and fix any $\overrightarrow{\mathbf{T}}\in(V)_{k-1}$. By Lemma 5.3, G has at least $\zeta n^{\ell-(k-1)}$ ℓ-walks from $\overrightarrow{\mathbf{S}}$ to $\overrightarrow{\mathbf{T}}$, for each $\overrightarrow{\mathbf{S}}\in(V)_{k-1}$, and thus

$$\begin{aligned}\mathbb{P}\big[\overrightarrow{\mathbf{Z}}_\ell=\overrightarrow{\mathbf{T}}\big] &= \sum_{\overrightarrow{\mathbf{S}}\in(V)_{k-1}}\phi(\overrightarrow{\mathbf{S}})\mathbb{P}\big[\overrightarrow{\mathbf{Z}}_\ell=\overrightarrow{\mathbf{T}}\mid\overrightarrow{\mathbf{Z}}_0=\overrightarrow{\mathbf{S}}\big]\\ &\geqslant \sum_{\overrightarrow{\mathbf{S}}\in(V)_{k-1}}\phi(\overrightarrow{\mathbf{S}})\zeta n^{\ell-(k-1)}(\tau/n)^\ell\\ &= \zeta\tau^\ell/n^{k-1}.\end{aligned}$$

Thus we may construct a coupling of any pair of simple $\mathbf{x}$-walks $\mathcal{Z}^{(k-1)}$ and $\mathcal{Z}'^{(k-1)}$ such that

$$\mathbb{P}\big[\overrightarrow{\mathbf{Z}}_\ell=\overrightarrow{\mathbf{Z}}'_\ell=\overrightarrow{\mathbf{T}}\big]\geqslant\zeta\tau^\ell/n^{k-1}\quad\text{for all }\overrightarrow{\mathbf{T}}\in(V)_{k-1}.$$

Under this coupling, we have

$$\mathbb{P}\big[\overrightarrow{\mathbf{Z}}_\ell=\overrightarrow{\mathbf{Z}'}_\ell\big]\geqslant\frac{\zeta\tau^\ell}{n^{k-1}}|(V)_{k-1}|\geqslant\zeta\tau^\ell\left(\frac{n-(k-2)}{n}\right)^{k-1}\geqslant\zeta\tau^\ell/2. \tag{5.3}$$

We now construct a coupling of $\mathcal{Y}^{(k-1)}$ and $\mathcal{W}^{(k-1)}$ as follows. We partition the time steps into consecutive intervals of length ℓ. In the first interval, we couple $\mathcal{Y}^{(k-1)}$ and $\mathcal{W}^{(k-1)}$ as in (5.3), so that

$$\mathbb{P}\big[\overrightarrow{\mathbf{Y}}_\ell=\overrightarrow{\mathbf{W}}_\ell\big]\geqslant\zeta\tau^\ell/2.$$

If $\overrightarrow{\mathbf{Y}}_\ell=\overrightarrow{\mathbf{W}}_\ell$, then we couple $\mathcal{Y}^{(k-1)}$ and $\mathcal{W}^{(k-1)}$ such that $\overrightarrow{\mathbf{Y}}_t=\overrightarrow{\mathbf{W}}_t$ for all $t\geqslant\ell$. Otherwise we again couple $\mathcal{Y}^{(k-1)}$ and $\mathcal{W}^{(k-1)}$ in the second time interval, as in (5.3), so that

$$\mathbb{P}\big[\overrightarrow{\mathbf{Y}}_{2\ell}=\overrightarrow{\mathbf{W}}_{2\ell}\mid\overrightarrow{\mathbf{Y}}_\ell\neq\overrightarrow{\mathbf{W}}_\ell\big]\geqslant\zeta\tau^\ell/2.$$

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

One can easily check that repeating this process yields a valid coupling of $\mathcal{Y}^{(k-1)}$ and $\mathcal{W}^{(k-1)}$. Note that with this coupling and for any q which is sufficiently large compared to ℓ, we have

$$\mathbb{P}\big[\overrightarrow{\mathbf{Y}}_q \neq \overrightarrow{\mathbf{W}}_q\big] \leqslant \prod_{k=1}^{q/\ell} \mathbb{P}\Bigg[\overrightarrow{\mathbf{Y}}_{k\ell} \neq \overrightarrow{\mathbf{W}}_{k\ell} \,\Bigg|\, \bigcap_{j\leqslant k-1} \{\overrightarrow{\mathbf{Y}}_{j\ell} \neq \overrightarrow{\mathbf{W}}_{j\ell}\}\Bigg] \leqslant (1-\zeta\tau^{\ell}/2)^{q/\ell}.$$

We use this coupling and apply (3.1) to obtain

$$d_{TV}(Y_q, W_q) \leqslant \mathbb{P}\big[Y_q \neq W_q\big] \leqslant \mathbb{P}\big[\overrightarrow{\mathbf{Y}}_q \neq \overrightarrow{\mathbf{W}}_q\big] \leqslant (1-\zeta\tau^{\ell}/2)^{q/\ell} \leqslant \exp(-\zeta\tau^{\ell}q/2\ell) < \eta,$$

provided $q \geqslant \lambda \ln(1/\eta)$. □

It will be useful to have an explicit formula for the stationary distribution π of $\mathcal{Y}^{(k-1)}$, and so we obtain this now. (Observe that the simple $\mathbf{x}$-walk $\mathcal{Y}^{(k-1)}$ on $(V)_{k-1}$ is in general not a symmetric Markov chain.)

Proposition 5.5. *Let $1/n \ll \gamma \ll 1/k$ where $k \geqslant 2$, and let G be an (n,k,γ)-graph. Let $\mathbf{x}\colon E(G) \to \mathbb{R}^+$ be a positive edge weighting of G, and for each $\overrightarrow{\mathbf{M}} \in (V)_{k-1}$ define*

$$\pi(\overrightarrow{\mathbf{M}}) \coloneqq \frac{\sum_{v\in V} \mathbf{x}(\mathbf{M}\cup\{v\})}{\sum_{\overrightarrow{\mathbf{B}}\in(V)_{k-1}} \sum_{v\in V} \mathbf{x}(\mathbf{B}\cup\{v\})}. \tag{5.4}$$

Then π is the unique stationary distribution of the simple $\mathbf{x}$-walk $\mathcal{Y}^{(k-1)}$ on $(V)_{k-1}$.

Proof. By standard results on the stationary distribution of a Markov chain (see *e.g.* [22, Proposition 1.20]), it suffices to prove that $\pi\colon (V)_{k-1} \to \mathbb{R}^+$ as defined in (5.4) is a probability distribution on $(V)_{k-1}$, and that

$$\sum_{\overrightarrow{\mathbf{S}}\in(V)_{k-1}} \pi(\overrightarrow{\mathbf{S}})P(\overrightarrow{\mathbf{S}},\overrightarrow{\mathbf{T}}) = \pi(\overrightarrow{\mathbf{T}}) \quad \text{for all } \overrightarrow{\mathbf{T}} \in (V)_{k-1}, \tag{5.5}$$

where $P(\overrightarrow{\mathbf{S}},\overrightarrow{\mathbf{T}})$ denotes the (one-step) transition probability of $\mathcal{Y}^{(k-1)}$ from $\overrightarrow{\mathbf{S}}$ to $\overrightarrow{\mathbf{T}}$. It follows quickly from (5.4) that π is a probability distribution on $(V)_{k-1}$, and (5.5) follows from applying (5.1) and (5.4). □

Next we show that, provided the edge weighting $\mathbf{x}\colon E(G) \to \mathbb{R}^+$ is an 'almost-perfect' fractional matching, the stationary $\mathbf{x}$-walk on G is such that each vertex of G is roughly equally likely to be the current vertex at any time. Let $\varepsilon \in (0,1)$ and $k \geqslant 2$. For a k-graph G, we say that a fractional matching $\mathbf{x}\colon E(G) \to \mathbb{R}^+$ is *ε-almost-perfect* if $\sum_{e\ni v} \mathbf{x}(e) \geqslant 1-\varepsilon$ for all $v \in V$.

Lemma 5.6. *Let $1/n \ll \gamma, 1/k$, where $k \geqslant 2$, and let G be an (n,k,γ)-graph. Suppose that $\mathbf{x}\colon E(G) \to \mathbb{R}^+$ is an $n^{-2/5}$-almost-perfect fractional matching of G and let*

$$\mathcal{W} = (W_{-(k-2)}, W_{-(k-3)}, \dots)$$

be the stationary $\mathbf{x}$-walk on G. Then, for each $i \in \mathbb{N}$ and each $v \in V$, we have

$$\mathbb{P}[W_i = v] = (1 \pm n^{-1/3}) \cdot \frac{1}{n}. \tag{5.6}$$

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

Proof. We reformulate $\mathcal{W}$ as $\mathcal{W} = (\overrightarrow{\mathbf{W}}_0, \overrightarrow{\mathbf{W}}_1, \dots)$, where $\overrightarrow{\mathbf{W}}_i \coloneqq (W_{i-(k-2)}, \dots, W_i)$. Let $v \in V$ and $i \in \mathbb{N}$. By the law of total probability, (5.1) and Proposition 5.5, we obtain

$$\begin{aligned}
\mathbb{P}[W_i = v] &= \sum_{\overrightarrow{\mathbf{S}} \in (V)_{k-1}} \mathbb{P}\big[\overrightarrow{\mathbf{W}}_{i-1} = \overrightarrow{\mathbf{S}}\big]\mathbb{P}\big[W_i = v \mid \overrightarrow{\mathbf{W}}_{i-1} = \overrightarrow{\mathbf{S}}\big] \\
&= \sum_{\overrightarrow{\mathbf{S}} \in (V)_{k-1}} \frac{\sum_{v' \in V} \mathbf{x}(\mathbf{S} \cup \{v'\})}{\sum_{\overrightarrow{\mathbf{B}} \in (V)_{k-1}} \sum_{v' \in V} \mathbf{x}(\mathbf{B} \cup \{v'\})} \cdot \frac{\mathbf{x}(\mathbf{S} \cup \{v\})}{\sum_{v' \in V} \mathbf{x}(\mathbf{S} \cup \{v'\})} \\
&= \frac{\sum_{\mathbf{S} \in \binom{V}{k-1}} \mathbf{x}(\mathbf{S} \cup \{v\})}{\sum_{\mathbf{B} \in \binom{V}{k-1}} \sum_{v' \in V} \mathbf{x}(\mathbf{B} \cup \{v'\})} \\
&= \frac{\sum_{\mathbf{S} \in \binom{V}{k-1}} \mathbf{x}(\mathbf{S} \cup \{v\})}{\sum_{v' \in V} \sum_{\mathbf{B} \in \binom{V}{k-1}} \mathbf{x}(\mathbf{B} \cup \{v'\})} \\
&= \frac{\sum_{e \ni v} \mathbf{x}(e)}{\sum_{v' \in V} \sum_{e \ni v'} \mathbf{x}(e)}.
\end{aligned}$$

Applying

$$1 - n^{-2/5} \leqslant \sum_{e \ni v} \mathbf{x}(e) \leqslant 1 \quad \text{for all } v \in V,$$

we obtain (5.6). □

We now show the crucial fact that, for any large set $U \subseteq V$, the probability that the self-avoiding $\mathbf{x}$-walk $\mathcal{X}$ is in U after a small number of steps is roughly $|U|/n$. We will apply this fact to neighbourhoods of $(k-1)$-tuples in the proof that $\mathcal{X}$ is likely to be good. This in turn will be used to show that $\mathcal{X}$ will, in expectation, behave roughly uniformly with respect to the codegrees of all $(k-1)$-tuples.

Lemma 5.7 (Uniformity Lemma). *Let $1/n \ll 1/C \ll \gamma, 1/k$, where $k \geqslant 2$, and let G be an (n, k, γ)-graph. Let $\mathbf{x}\colon E(G) \to \mathbb{R}^+$ be a C-normal, $n^{-2/5}$-almost-perfect fractional matching, let $\overrightarrow{\mathbf{S}} \in (V)_{k-1}$, and let*

$$\mathcal{X} = (X_{-(k-2)}, X_{-(k-3)}, \dots)$$

be the self-avoiding $\mathbf{x}$-walk on G with starting tuple $\overrightarrow{\mathbf{S}}$. Then, for any $q \in [(\ln n)^2, n^{1/5}]$ and any $U \subseteq V$ of size $|U| \geqslant n^{3/4}$, we have

$$\mathbb{P}[X_q \in U] = (1 \pm n^{-3/10})\frac{|U|}{n}. \tag{5.7}$$

Proof. We first argue that $\mathbf{x}$ is $(1/C^2 n, 2C^2/n)$-balanced. Indeed, since $\mathbf{x}$ is C-normal and G is γ-Dirac, we have

$$\frac{\max_{e \in E(G)} \mathbf{x}(e)}{\min_{\mathbf{S} \in \binom{V}{k-1}} \sum_{v \in V} \mathbf{x}(\mathbf{S} \cup \{v\})} \leqslant \frac{C}{n^{k-1}} \cdot \frac{Cn^{k-1}}{(1/2 + \gamma)n} \leqslant \frac{2C^2}{n}.$$

The lower bound follows similarly. Now let $U \subseteq V$ be of size $|U| \geqslant n^{3/4}$, and fix $q \in [(\ln n)^2, n^{1/5}]$. Let

$$\mathcal{Y} = (Y_{-(k-2)}, Y_{-(k-3)}, \dots)$$

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

be the simple $\mathbf{x}$-walk on G with starting tuple $\overrightarrow{\mathbf{S}}$ and let

$$\mathcal{W} = (W_{-(k-2)}, W_{-(k-3)}, \dots)$$

be the stationary $\mathbf{x}$-walk on G. We will show that X_q is distributed similarly to Y_q since q is not too large, and that Y_q is distributed similarly to W_q since q is large enough, and finally we will use Lemma 5.6 to show that $\mathbb{P}[W_q \in U]$ is roughly $|U|/n$.

Setting

$$c_1 \coloneqq |\mathbb{P}[X_q \in U] - \mathbb{P}[Y_q \in U]|$$

and applying Lemma 5.2 (with $2C^2/n$ playing the role of r), we obtain

$$\begin{aligned} c_1 &= \left|\sum_{v\in U} (\mathbb{P}[X_q = v] - \mathbb{P}[Y_q = v])\right| \\ &\leqslant \sum_{v\in V} |\mathbb{P}[X_q = v] - \mathbb{P}[Y_q = v]| \\ &\overset{(3.1)}{=} 2d_{TV}(X_q, Y_q) \\ &\leqslant 4C^2q^2/n \\ &< \frac{1}{3}n^{-13/10}|U|. \end{aligned}$$

Consider the term

$$c_2 \coloneqq |\mathbb{P}[Y_q \in U] - \mathbb{P}[W_q \in U]|.$$

We apply Lemma 5.4 (with $1/C^2$, $1/n^2$ playing the roles of τ, η respectively, and using $q \geqslant (\ln n)^2 = ((\ln n)/2)\cdot \ln(1/\eta))$ to obtain

$$c_2 \leqslant \sum_{v\in V} |\mathbb{P}[Y_q = v] - \mathbb{P}[W_q = v]| \overset{(3.1)}{=} 2d_{TV}(Y_q, W_q) < 2n^{-2} < \frac{1}{3}n^{-13/10}|U|.$$

We now apply Lemma 5.6 to the term

$$c_3 \coloneqq |\mathbb{P}[W_q \in U] - |U|/n|$$

to obtain

$$c_3 = \left|\sum_{v\in U}\left(\mathbb{P}[W_q = v] - \frac{1}{n}\right)\right| \leqslant \sum_{v\in U}\left|\mathbb{P}[W_q = v] - \frac{1}{n}\right| \leqslant n^{-4/3}|U|.$$

Finally, by the triangle inequality we have

$$|\mathbb{P}[X_q \in U] - |U|/n| \leqslant c_1 + c_2 + c_3 < n^{-13/10}|U|,$$

which is equivalent to (5.7). □

5.4 The walk is likely to be good

The aim of this section is to prove the following lemma, which states that under our assumptions, a self-avoiding $\mathbf{x}$-walk of length $\sqrt{n}$ is good with high probability.

Lemma 5.8. *Let $1/n \ll 1/C \ll \gamma, 1/k$, where $k \geqslant 2$, and let G be an (n,k,γ)-graph. Let $\mathbf{x}\colon E(G) \to \mathbb{R}^+$ be a C-normal perfect fractional matching, let $\overrightarrow{\mathbf{S}} \in (V)_{k-1}$, let $\kappa \coloneqq \sqrt{n}$,*

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

and let $\mathcal{X} = (X_{-(k-2)}, \dots, X_\kappa)$ *be a self-avoiding* $\mathbf{x}$*-walk on* G *with starting tuple* $\overrightarrow{\mathbf{S}}$*. Then* $\mathbb{P}[\mathcal{X} \text{ is good}] \geqslant 1 - 1/n$.

To prove Lemma 5.8, we will need some results on the behaviour of the residual graphs G_j as the walk $\mathcal{X}$ progresses, so that we can apply Lemma 5.7 to each G_j. We define $\mathbf{x}|_{G_j}$ to be the restriction of $\mathbf{x}$ to $E(G_j)$, so that $\mathbf{x}|_{G_j} \colon E(G_j) \to \mathbb{R}^+$ is a (not necessarily perfect) fractional matching of G_j.

Proposition 5.9. *Suppose the assumptions of Lemma 5.8 hold. Let* $\mathcal{F}$ *be the set of tracking functions of* G*. Then, for any* $g \in \mathcal{F}$ *and any* $j \in \{0, \dots, \kappa\}$*, the following conditions hold deterministically:*

(i) $g(V_j) = (1 \pm n^{-1/4})((n-j)/n)g(V)$,
(ii) G_j *is* $\gamma/2$*-Dirac,*
(iii) $\mathbf{x}|_{G_j}$ *is* $2C$*-normal,*
(iv) $\mathbf{x}|_{G_j}$ *is* $n^{-2/5}$*-almost-perfect.*

Proof. It suffices to prove that conditions (i)–(iv) hold for any $j \in \{0, \dots, \kappa\}$ and any outcome $x(j) = (x_{-(k-2)}, \dots, x_j)$ of $\mathcal{X}(j)$. Throughout the proof, we let $j \in \{0, \dots, \kappa\}$ be fixed, and we let $x(j)$ be a fixed outcome of $\mathcal{X}(j)$, thus determining V_j and G_j.

(i) Fix $g \in \mathcal{F}$. It is clear that $g(V) - \kappa \leqslant g(V_j) \leqslant g(V)$. Relaxing the upper bound and recalling that $g(V) \geqslant n/2$, we obtain $g(V_j) = (1 \pm 2\kappa/n)g(V)$. Note that

$$2\kappa/n = 2n^{-1/2} < n^{-1/4}(1 - \kappa/n) - \kappa/n \leqslant n^{-1/4}(1 - j/n) - j/n,$$

which implies (i).

(ii) Let $\mathbf{M} \in \binom{V_j}{k-1}$. By (i), we have

$$\begin{aligned} d_{G_j}(\mathbf{M}) &= g_{\mathbf{M}}(V_j) \\ &\geqslant (1 - n^{-1/4})\frac{n-j}{n} g_{\mathbf{M}}(V) \\ &\geqslant (1 - n^{-1/4})\left(\frac{1}{2} + \gamma\right)(n-j) \\ &\geqslant \left(\frac{1}{2} + \frac{\gamma}{2}\right)|V_j|. \end{aligned} \tag{5.8}$$

Since (5.8) holds for all $\mathbf{M} \in \binom{V_j}{k-1}$, we conclude that G_j is $\gamma/2$-Dirac.

The calculations for (iii) and (iv) are straightforward. □

For any $j \in \{0, \dots, \kappa\}$ and any fixed outcome $x(j) = (x_{-(k-2)}, \dots, x_j)$ of $\mathcal{X}(j)$, we write $\mathbb{P}_{x(j)}$ for the probability measure in the conditional probability space where we have fixed $\mathcal{X}(j) = x(j)$, so that $\mathbb{P}_{x(j)}[\,\cdot\,] = \mathbb{P}[\,\cdot \mid \mathcal{X}(j) = x(j)]$. We are now ready to prove Lemma 5.8.

Proof of Lemma 5.8. We need to prove that with probability at least $1 - 1/n$,

$$\mathrm{Error}_g(\mathcal{X}) \coloneqq \left| \left(\sum_{j=-(k-2)}^{\kappa} g(X_j) \right) - \frac{\kappa}{n} g(V) \right| < n^{3/10} \tag{5.9}$$

holds simultaneously for every $g \in \mathcal{F}$. It will suffice to prove that (5.9) holds for any fixed $g \in \mathcal{F}$ with probability at least $1 - 1/n^k$, say, since $|\mathcal{F}| \leqslant n^{k-1}$. Fix $g \in \mathcal{F}$ and set $q \coloneqq (\ln n)^2$. By

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

breaking up $\mathrm{Error}_g(\mathcal{X})$ and repeatedly applying the triangle inequality, we obtain

$$\begin{aligned}\mathrm{Error}_g(\mathcal{X}) &= \left| \sum_{j=-(k-2)}^{0} g(X_j) + \sum_{j=1}^{q-1}\left(g(X_j) - \frac{g(V)}{n}\right) + \sum_{j=q}^{\kappa}\left(g(X_j) - \frac{g(V)}{n}\right)\right| \\ &\leqslant 2q + \left|\sum_{j=0}^{\kappa-q}\left(g(X_{j+q}) - \frac{g(V)}{n}\right)\right| \\ &\leqslant 2q + \left|\sum_{j=0}^{\kappa-q}\left(g(X_{j+q}) - \mathbb{E}[g(X_{j+q}) \mid \mathcal{X}(j)]\right)\right| \\ &\quad + \sum_{j=0}^{\kappa-q}\left|\mathbb{E}[g(X_{j+q}) \mid \mathcal{X}(j)] - \frac{g(V_j)}{n-j}\right| + \sum_{j=0}^{\kappa-q}\left|\frac{g(V_j)}{n-j} - \frac{g(V)}{n}\right|.\end{aligned}$$

We now prove an upper bound for each of the three sums in the final expression above. To this end, fix $j \in \{0, \dots, \kappa - q\}$, fix an outcome $x(j) = (x_{-(k-2)}, \dots, x_j)$ of $\mathcal{X}(j)$, and let $\mathbf{T} \in \binom{V}{k-1}$ be such that $g = g_{\mathbf{T}}$. We apply Proposition 5.9 to deduce that G_j is $\gamma/2$-Dirac, and that $\mathbf{x}|_{G_j}$ is $2C$-normal and $n^{-2/5}$-almost-perfect. We can now apply Lemma 5.7 to G_j to deduce that

$$\begin{aligned}\left|\mathbb{E}[g(X_{j+q}) \mid \mathcal{X}(j) = x(j)] - \frac{g(V_j)}{n-j}\right| &= \left|\mathbb{P}_{x(j)}[X_{j+q} \in N_{G_j}(\mathbf{T})] - \frac{|N_{G_j}(\mathbf{T})|}{n-j}\right| \\ &< (n-j)^{-13/10}|N_{G_j}(\mathbf{T})| \\ &< n^{-5/4}g(V).\end{aligned}$$

We deduce that

$$|\mathbb{E}[g(X_{j+q}) \mid \mathcal{X}(j)] - g(V_j)/(n-j)| < n^{-5/4}g(V) \quad \text{for each } j \in \{0, \dots, \kappa - q\}.$$

Next, we apply Proposition 5.9(i) to obtain that, for each $j \in \{0, \dots, \kappa - q\}$, we have

$$|g(V_j)/(n-j) - g(V)/n| \leqslant n^{-5/4}g(V).$$

Finally, applying Lemma 3.1 with $\log n$ playing the role of t, and using $\|g\|_\infty = 1$ (there are no isolated $(k-1)$-tuples), we see that with probability at least $1 - 2q\exp(-(\log n)^2/2)$, we have

$$\left|\sum_{j=0}^{\kappa-q}\left(g(X_{j+q}) - \mathbb{E}[g(X_{j+q}) \mid \mathcal{X}(j)]\right)\right| \leqslant \log n\sqrt{q(\kappa - q)},$$

so that altogether, with probability at least $1 - 1/n^k$, we have

$$\mathrm{Error}_g(\mathcal{X}) \leqslant 2q + \log n\sqrt{q(\kappa-q)} + 2(\kappa - q + 1)n^{-5/4}g(V) < n^{3/10},$$

completing the proof of the lemma. □

We now have all the tools we need to prove Lemma 5.1.

Proof of Lemma 5.1. Choose a new constant C satisfying $1/n \ll c \ll 1/C \ll \gamma, 1/k$, and let $\mathbf{x}\colon E(G) \to \mathbb{R}^+$ be a C-normal perfect fractional matching (such an $\mathbf{x}$ exists by Lemma 4.2). Write $\kappa \coloneqq \sqrt{n}$, and let $\mathcal{X} = (X_{-(k-2)}, \dots, X_\kappa)$ be the self-avoiding $\mathbf{x}$-walk on G with starting tuple $\overrightarrow{\mathbf{S}}$. It is clear from the definition of a self-avoiding $\mathbf{x}$-walk that any outcome of $\mathcal{X}$ corresponds to a κ-path in G with $\overrightarrow{\mathbf{S}}$ as one end (note that the walk does not stop before time κ, since all codegrees are large enough). We argue now that good outcomes of $\mathcal{X}$ also satisfy condition (iii), where we say

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

an outcome $X = (X_{-(k-2)}, \dots, X_\kappa)$ of $\mathcal{X}$ is a *good outcome* if X satisfies (5.2). Let P be a tight path in G corresponding to a good outcome X of $\mathcal{X}$, let $\overrightarrow{\mathbf{T}}$ be the non-$\overrightarrow{\mathbf{S}}$ end of P, and let G_κ denote the residual graph of G at time κ of X. Thus $G_\kappa = G - (V(P) \setminus \mathbf{T})$ and $|V_\kappa| = n - \kappa$. Let $\mathbf{M} \in \binom{V_\kappa}{k-1}$ and let $g_{\mathbf{M}} \in \mathcal{F}$ be the tracking function of G corresponding to $\mathbf{M}$. Since X is good, we obtain

$$\begin{aligned} d_{G_\kappa}(\mathbf{M}) &= g_{\mathbf{M}}(V_\kappa) \\ &= g_{\mathbf{M}}(V) - \sum_{j=-(k-2)}^{\kappa} g_{\mathbf{M}}(X_j) + \sum_{j=\kappa-(k-2)}^{\kappa} g_{\mathbf{M}}(X_j) \\ &\overset{(5.2)}{\geqslant} \frac{n-\kappa}{n} g_{\mathbf{M}}(V) - n^{3/10} \\ &\geqslant \left(\frac{1}{2} + \gamma - n^{-2/3}\right)|V_\kappa|. \end{aligned} \tag{5.10}$$

Since (5.10) holds for all $\mathbf{M} \in \binom{V_\kappa}{k-1}$, we conclude that

$$\delta(G_\kappa) \geqslant (1/2 + \gamma - n^{-2/3})|V_\kappa|.$$

Lastly, it suffices to count the number of good outcomes of $\mathcal{X}$. We begin by finding an upper bound for the probability that $\mathcal{X}$ yields any particular fixed tight path. For any $j \in \{0, \dots, \kappa\}$, we have by Proposition 5.9(ii)–(iii) that G_j is $\gamma/2$-Dirac and $\mathbf{x}|_{G_j}$ is $2C$-normal. It follows that $\mathbf{x}|_{G_j}$ is $8C^2/(n-j)$-upper-balanced. In particular, setting $p := 16C^2/n$, we have that all transition probabilities of $\mathcal{X}$ are bounded from above by p. Let $Q = (q_{-(k-2)}, \dots, q_\kappa)$ be a fixed κ-path in G with $\overrightarrow{\mathbf{S}} = (q_{-(k-2)}, \dots, q_0)$ as one end. Then

$$\mathbb{P}[\mathcal{X} = Q] = \prod_{j=1}^{\kappa} \mathbb{P}\left[X_j = q_j \,\middle|\, \bigcap_{i=-(k-2)}^{j-1} \{X_i = q_i\}\right] \leqslant p^\kappa.$$

By Lemma 5.8, we have that $\mathbb{P}[\mathcal{X} \text{ is good}] \geqslant 1/2$, so we conclude that the number of good outcomes of $\mathcal{X}$ (and thus the number of tight κ-paths in G satisfying (ii) and (iii)) is at least $(1/2)/p^\kappa \geqslant (cn)^\kappa$, which completes the proof of the lemma. □

6. Counting and absorbing long paths

In this section we show how to iterate Lemma 5.1 to construct tight paths in G which use almost all of the vertices of G, and we count the number of choices that can be made in this process to obtain a lower bound for the number of these long paths. Finally, we prove Theorem 1.1 by showing how these paths can be completed into tight Hamilton cycles of G.

Lemma 6.1. *Let $\gamma > 0$, let $k \geqslant 2$, and let G be an (n, k, γ)-graph. There are at least* $\exp(n \ln n - \Theta(n))$ *tight paths in G of length at least $n - n^{7/8}$.*

Proof. We describe an algorithm on G. Let $\overrightarrow{\mathbf{S}}_0 \in (V)_{k-1}$ be arbitrary, set $G_0 := G$, set $n_0 := n$, and set $\gamma_0 := \gamma$. For each $i \geqslant 0$, set $n_{i+1} := n_i - \sqrt{n_i}$, and set $\gamma_{i+1} := \gamma_i - (n_i)^{-2/3}$. Set L to be the smallest index such that $n_L < n^{7/8}$. Suppose we have already performed i steps of the algorithm, and obtained a k-graph G_i on n_i vertices satisfying $\delta(G_i) \geqslant (1/2 + \gamma_i)n_i$, and we have obtained $\overrightarrow{\mathbf{S}}_i \in (V(G_i))_{k-1}$. If $\gamma_i < \gamma/2$ or $i = L$, then we terminate the algorithm. Otherwise, we apply Lemma 5.1 to G_i to obtain a set $\mathcal{P}_{i+1}$ of $\sqrt{n_i}$-paths, each with chosen starting tuple $\overrightarrow{\mathbf{S}}_i$. Choose

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

$P_{i+1} \in \mathcal{P}_{i+1}$ arbitrarily, let $\overrightarrow{\mathbf{T}}_{i+1}$ be the non-$\overrightarrow{\mathbf{S}}_i$ end of P_{i+1}, and put $\overrightarrow{\mathbf{S}}_{i+1} \coloneqq \overleftarrow{\mathbf{T}}_{i+1}$. Set

$$G_{i+1} \coloneqq G_i - (V(P_{i+1}) \setminus \mathbf{S}_{i+1}).$$

Observe that by Lemma 5.1(iii), we have

$$\delta(G_{i+1}) \geqslant (1/2 + \gamma_{i+1})n_{i+1}.$$

Let

$$r_i \coloneqq \sum_{j=0}^{i-1} (n_j)^{-2/3}.$$

Note that, provided $\gamma_{i-1} \geqslant \gamma/2$, we have $\gamma_i = \gamma - r_i$. We claim that the algorithm does not terminate in the first L steps. To see this, note that it suffices to show that $r_L = o(1)$. Write $\kappa_i \coloneqq \sqrt{n_i}$ and observe that

$$n_{L-1} = n - \sum_{j=0}^{L-2} \kappa_j \leqslant n - (L-1)\kappa_{L-1}.$$

Rearranging, we obtain that $L \leqslant 2n/\kappa_{L-1}$. Using $n_{L-1} \geqslant n^{7/8}$, we obtain that

$$r_L \leqslant \frac{L}{(n_{L-1})^{2/3}} \leqslant \frac{2n}{(n_{L-1})^{7/6}} \leqslant 2n^{-1/48} = o(1),$$

so the algorithm does not terminate in the first L steps, as claimed. When the algorithm terminates, we have obtained tight paths $P_1, \dots, P_L$. By construction, we may concatenate these paths, in order, to obtain a path $Q \coloneqq \bigcup_{i \leqslant L} P_i$ of length $n - n_L \geqslant n - n^{7/8}$. Let N be the number of tight paths of length $n - n_L$ in G. By Lemma 5.1(i), there is a positive constant $c < 1$ such that the number of choices for P_{i+1} is at least $(cn_i)^{\kappa_i}$, for each $i \in \{0, \dots, L-1\}$. Thus we obtain

$$N \geqslant \prod_{i=0}^{L-1} (cn_i)^{\kappa_i} \geqslant c^n \prod_{i=0}^{L-1} \frac{n_i!}{n_{i+1}!} = c^n \frac{n!}{n_L!} \geqslant c^n \frac{n!}{(n^{7/8})!} = \exp(n \ln n - \Theta(n)).$$

□

We are now ready to prove Theorem 1.1.

Proof of Theorem 1.1. The upper bound holds trivially. To prove the lower bound, we choose a set $W \subseteq V$ of size $n^{9/10}$ uniformly at random. A simple application of Lemma 3.2 shows that there is a choice of W such that $|N(\mathbf{S}) \cap W| \geqslant (1/2 + 3\gamma/4)|W|$ for all $\mathbf{S} \in \binom{V}{k-1}$. Fix such a choice of W, set $G' \coloneqq G - W$, and put $n' \coloneqq n - n^{9/10}$. Then G' is $\gamma/2$-Dirac, and we apply Lemma 6.1 to G' (with $\gamma/2$ playing the role of γ) to find a set $\mathcal{P}$ of tight paths of length at least $n' - (n')^{7/8}$ in G', such that

$$|\mathcal{P}| \geqslant \exp(n' \ln n' - \Theta(n')) = \exp(n \ln n - \Theta(n)).$$

Fix $P \in \mathcal{P}$, let $\overrightarrow{\mathbf{S}}_P$ and $\overrightarrow{\mathbf{T}}_P$ be the ends of P, and let $U_P \coloneqq V(G') \setminus V(P)$, so that $|U_P| \leqslant (n')^{7/8} \leqslant n^{7/8}$. Notice that $G[W \cup U_P \cup \mathbf{S}_P \cup \mathbf{T}_P]$ is $\gamma/2$-Dirac. Thus, by Lemma 3.6, there is a tight Hamilton path Q_P of $G[W \cup U_P \cup \mathbf{S}_P \cup \mathbf{T}_P]$ with ends $\overleftarrow{\mathbf{S}}_P$ and $\overleftarrow{\mathbf{T}}_P$. Then $C_P \coloneqq P \cup Q_P$ is a tight Hamilton cycle of G.

Define $\mathcal{C} \coloneqq \{C_P \colon P \in \mathcal{P}\}$ and note that for each $C \in \mathcal{C}$, the number of $P \in \mathcal{P}$ with $C = C_P$ is at most n^2, since P must be a subpath of C. We conclude that $\mathcal{C}$ is a set of at least

$$n^{-2} \exp(n \ln n - \Theta(n)) = \exp(n \ln n - \Theta(n))$$

tight Hamilton cycles in G. □

Finally, we prove Corollary 1.1.

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

Proof of Corollary 1.1. Let $\gamma > 0$ be fixed, let $k \geqslant 2$, $\ell \in \{0, \dots, k-1\}$, $(k-\ell) \mid n$, and let G be a k-graph on n vertices satisfying $\delta(G) \geqslant (1/2+\gamma)n$. First, suppose $\ell = 0$, and recall that the number of Hamilton 0-cycles of G is precisely the number of perfect matchings of G. By considering the number of perfect matchings in the complete k-graph on n vertices, it is easy to see that the upper bound of (i) holds. We now use Theorem 1.1 to show that the lower bound holds. Let $\mathcal{M}$ be the set of perfect matchings of G, and let $\mathcal{C}$ be the set of tight Hamilton cycles of G. Notice that, for any $M \in \mathcal{M}$, there are at most $(n/k)!(k!)^{n/k}$ choices of $C \in \mathcal{C}$ such that $M \subseteq E(C)$, because we may construct all vertex orderings corresponding to possible such C by reordering the edges of M and the vertices within them. By applying Theorem 1.1, we conclude that

$$|\mathcal{M}| \geqslant \frac{|\mathcal{C}|}{(n/k)!(k!)^{n/k}} = \exp\left(\left(1-\frac{1}{k}\right)n\ln n - \Theta(n)\right).$$

For the case $\ell \in [k-1]$, first notice that

$$\frac{k-\ell}{2}(n-1)! = \exp\left(n\ln n - \Theta(n)\right)$$

is a trivial upper bound for the number of Hamilton ℓ-cycles of G. Finally, it suffices to apply Theorem 1.1 to G and observe that every tight Hamilton cycle of G contains $k-\ell$ Hamilton ℓ-cycles (since $(k-\ell) \mid n$), and each Hamilton ℓ-cycle of G is contained in at most $(k!)^{n/(k-\ell)}$ tight Hamilton cycles. □

7. Concluding remarks

Although Theorem 1.1 holds in γ-Dirac k-graphs with equality, we believe that the error bound can be made more precise. More specifically, we believe the following hypergraph version of [8, Theorem 1.1] holds, giving a more accurate lower bound for the number of tight Hamilton cycles in such hypergraphs.

Conjecture 7.1. *For a fixed integer $k \geqslant 2$ and a fixed constant $\gamma > 0$, the number of tight Hamilton cycles of a k-graph G on n vertices with $\delta(G) \geqslant (1/2+\gamma)n$ is at least $(1/2 - o(1))^n n!$.*

It would of course be desirable to obtain a formula for the number of tight Hamilton cycles in γ-Dirac k-graphs G which takes properties of G like the degrees and codegrees into account. We recall that such a formula has already been obtained [8, Theorem 1.3, Theorem 1.5] in terms of the 'entropy of G' in the $k=2$ case, and it would be interesting to see if this (or a similar) notion can be extended to $k \geqslant 3$.

Finally, we note that the results of [8] show that graphs with minimum degree precisely at the threshold for Hamiltonicity in fact have many Hamilton cycles. The exact minimum codegree threshold for existence of a tight Hamilton cycle in k-graphs on n vertices is not yet known for $k \geqslant 4$, but is known to be $\lfloor n/2 \rfloor$ in the case $k=3$ [28, Theorem 1.2], and it is of course a natural question to ask if the conclusions of Theorem 1.1 or indeed Conjecture 7.1 hold in this exact setting.

References

[1] Aigner-Horev, E. and Levy, G. (2020) Tight Hamilton cycles in cherry-quasirandom 3-uniform hypergraphs. *Combin. Probab. Comput.* doi:10.1017/S0963548320000486

[2] Bermond, J.-C., Germa, A., Heydemann, M.-C. and Sotteau, D. (1976) Hypergraphes Hamiltoniens. *Prob. Comb. Theorie Graph Orsay* **260** 39–43.

[3] Bollobás, B. (1995) Extremal graph theory. In *Handbook of Combinatorics*, Vol II (R. L. Graham *et al.*, eds), pp. 1231–1292. Elsevier Science.

[4] Bondy, J. A. (1995) Basic graph theory: paths and circuits. In *Handbook of Combinatorics*, Vol I (R. L. Graham *et al.*, eds), pp. 3–110. Elsevier Science.

[5] Chung, F. and Lu, L. (2006) Concentration inequalities and martingale inequalities: a survey. *Internet Math.* **3** 79–127.

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press

[6] Coulson, M., Keevash, P., Perarnau, G. and Yepremyan, L. (2020) Rainbow factors in hypergraphs. *J. Combin. Theory Ser. A* **172** 105184.
[7] Cuckler, B. and Kahn, J. (2009) Entropy bounds for perfect matchings and Hamiltonian cycles. *Combinatorica* **29** 327–335.
[8] Cuckler, B. and Kahn, J. (2009) Hamiltonian cycles in Dirac graphs. *Combinatorica* **29** 299–326.
[9] Czygrinow, A. and Molla, T. (2014) Tight codegree condition for the existence of loose Hamilton cycles in 3-graphs. *SIAM J. Discrete Math.* **28** 67–76.
[10] Dirac, G. A. (1952) Some theorems on abstract graphs. *Proc. Lond. Math. Soc.* **3** 69–81.
[11] Durrett, R. (2019) *Probability: Theory and Examples*, Vol. 49 of Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.
[12] Ferber, S., Krivelevich, M. and Sudakov, B. (2016) Counting and packing Hamilton ℓ-cycles in dense hypergraphs. *J. Combin.* **7** 135–157.
[13] Hàn, H. and Schacht, M. (2010) Dirac-type results for loose Hamilton cycles in uniform hypergraphs. *J. Combin. Theory Ser. B* **100** 332–346.
[14] Han, J. and Zhao, Y. (2015) Minimum codegree threshold for Hamilton ℓ-cycles in k-uniform hypergraphs. *J. Combin. Theory Ser. A* **132** 194–223.
[15] Janson, S., Łuczak, T. and Ruciński, A. (2000) *Random Graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience.
[16] Katona, G. Y. and Kierstead, H. A. (1999) Hamiltonian chains in hypergraphs. *J. Graph Theory* **30** 205–212.
[17] Keevash, P., Kühn, D., Mycroft, R. and Osthus, D. (2011) Loose Hamilton cycles in hypergraphs. *Discrete Math.* **311** 544–559.
[18] Kühn, D., Mycroft, R. and Osthus, D. (2010) Hamilton ℓ-cycles in uniform hypergraphs. *J. Combin. Theory Ser. A* **117** 910–927.
[19] Kühn, D. and Osthus, D. (2006) Loose Hamilton cycles in 3-uniform hypergraphs of high minimum degree. *J. Combin. Theory Ser. B* **96** 767–821.
[20] Kühn, D. and Osthus, D. (2006) Matchings in hypergraphs of large minimum degree. *J. Graph Theory* **51** 269–280.
[21] Kühn, D. and Osthus, D. (2014) Hamilton cycles in graphs and hypergraphs: an extremal perspective. In *Proceedings of the International Congress of Mathematicians (Seoul 2014)*, Vol. IV, pp. 381–406. Kyung Moon Sa, Seoul.
[22] Levin, D. A. and Peres, Y. (2017) *Markov Chains and Mixing Times*, American Mathematical Society.
[23] Reiher, C., Rödl, V., Ruciński, A., Schacht, M. and Szemerédi, E. (2019) Minimum vertex degree condition for tight Hamiltonian cycles in 3-uniform hypergraphs. *Proc. Lond. Math. Soc.* **119** 409–439.
[24] Rödl, V. and Ruciński, A. (2010) Dirac-type questions for hypergraphs: a survey (or more problems for Endre to solve). In *An Irregular Mind*, pp. 561–590. Springer.
[25] Rödl, V., Ruciński, A. and Szemerédi, E. (2006) A Dirac-type theorem for 3-uniform hypergraphs. *Combin. Probab. Comput.* **15** 229–251.
[26] Rödl, V., Ruciński, A. and Szemerédi, E. (2008) An approximate Dirac-type theorem for k-uniform hypergraphs. *Combinatorica* **28** 229–260.
[27] Rödl, V., Ruciński, A. and Szemerédi, E. (2009) Perfect matchings in large uniform hypergraphs with large minimum collective degree. *J. Combin. Theory Ser. A* **116** 613–636.
[28] Rödl, V., Ruciński, A. and Szemerédi, E. (2011) Dirac-type conditions for Hamiltonian paths and cycles in 3-uniform hypergraphs. *Adv. Math.* **227** 1225–1299.
[29] Sárközy, G. N., Selkow, S. M. and Szemerédi, E. (2003) On the number of Hamiltonian cycles in Dirac graphs. *Discrete Math.* **265** 237–250.
[30] Zhao, Y. (2016) Recent advances on Dirac-type problems for hypergraphs. In *Recent Trends in Combinatorics*, pp. 145–165. Springer.

Cite this article: Glock S, Gould S, Joos F, Kühn D and Osthus D (2021). Counting Hamilton cycles in Dirac hypergraphs. *Combinatorics, Probability and Computing* **30**, 631–653. https://doi.org/10.1017/S0963548320000619

https://doi.org/10.1017/S0963548320000619 Published online by Cambridge University Press