Oriented discrepancy of Hamilton cycles

Lior Gishboliner[*]   Michael Krivelevich[†]   Peleg Michaeli[‡]

November 17, 2022

Abstract

We propose the following extension of Dirac's theorem: if G is a graph with $n \geq 3$ vertices and minimum degree $\delta(G) \geq n/2$, then in every orientation of G there is a Hamilton cycle with at least $\delta(G)$ edges oriented in the same direction. We prove an approximate version of this conjecture, showing that minimum degree $\frac{n+8k}{2}$ guarantees a Hamilton cycle with at least $(n+k)/2$ edges oriented in the same direction. We also study the analogous problem for random graphs, showing that if the edge probability $p = p(n)$ is above the Hamiltonicity threshold, then, with high probability, in every orientation of $G \sim G(n,p)$ there is a Hamilton cycle with $(1 - o(1))n$ edges oriented in the same direction.

1 Introduction

Recently there has been much interest in discrepancy-type problems on graphs. In the general setting of combinatorial discrepancy theory, one is given a hypergraph $\mathcal{H}$ and should colour the vertices of $\mathcal{H}$ with two colours so that each edge is as balanced as possible. More precisely, given a colouring $f : V(\mathcal{H}) \to \{-1, 1\}$, define the imbalance of an edge $e \in \mathcal{H}$ to be $D(e) = \left| \sum_{x \in e} f(x) \right|$. The discrepancy of $\mathcal{H}$ is then defined as $\min_{f:V(\mathcal{H}) \to \{-1,1\}} \max_{e \in E(\mathcal{H})} D(e)$. In other words, the discrepancy is the maximum *guaranteed imbalance*. We refer the reader to Chapter 4 in the book of Matoušek [19] for an overview of combinatorial discrepancy.

Recent works considered instances of the above general setting, in which the hypergraph $\mathcal{H}$ originates from a graph, that is, $V(\mathcal{H})$ is the edge-set of some graph G, and $E(\mathcal{H})$ is a family of certain subgraphs of G. The goal is to identify conditions on G which guarantee that there exists a high-discrepancy subgraph of the specific type. For example, Balogh, Csaba, Jing and Pluhár [1] proved that if G has n vertices and minimum degree $(3/4 + \varepsilon)n$, then in every 2-colouring of $E(G)$, there is a Hamilton cycle with at least $(1/2 + \varepsilon/64)n$ edges of the same colour (and hence discrepancy at least $\varepsilon n/32$). This was generalised to an arbitrary number of colours (with a suitable definition of multicolour discrepancy) in [11] and independently [8]. This result (for two colours) was extended to K_r-factors and r-powers of Hamilton cycles in [2] and [4], respectively. In [10], the authors determined the multicolour discrepancy of Hamilton cycles in the complete graph, and showed that if p is above the Hamiltonicity threshold, then the discrepancy of Hamilton cycles in $G(n,p)$ is essentially as large as in the complete graph.

In the present work, we consider an *oriented* version of graph discrepancy. Here, instead of colouring the edges of the graph with two colours, one considers orientations of the edges. Throughout this paper, we use $\{u, v\}$ to denote an unordered edge (or pair), and (u, v) to denote the directed edge from u to v. Hamilton cycles are naturally suitable for the oriented setting since they have a notion of direction. The goal is thus to find a Hamilton cycle with as many edges as possible which are oriented in the direction of the cycle. More precisely, for an n-vertex graph G, an orientation D of G and a Hamilton cycle $C = v_1, \ldots, v_n, v_1$ in G, we say that the edge $\{v_i, v_{i+1}\}$ (with indices taken modulo n) is a *forward edge of C* if $(v_i, v_{i+1}) \in E(D)$. Note that we always fix one of the two directions of the Hamilton cycle C. Let $\vec{e}_D(C)$ be the number of forward edges. The *oriented discrepancy of Hamilton cycles* in G is the maximum k such that for every orientation D of G, there is a Hamilton cycle C in G such that $\vec{e}_D(C) \geq \frac{n+k}{2}$. Generalising Dirac's theorem [5], we propose the conjecture that if $\delta(G) \geq \frac{n+k}{2}$ then the oriented discrepancy of Hamilton cycles in G is at least k. It is convenient to state this conjecture in the following form:

Conjecture 1.1. *Let $n \geq 3$ and let G be an n-vertex graph with $\delta(G) \geq n/2$. Then, in any orientation of the edges of G, there exists a Hamilton cycle in which at least $\delta(G)$ of the edges are pointing forward.*

If true, Conjecture 1.1 would be best possible. To see this, fix $n \geq 3$ and $\delta \geq n/2$, and consider the graph G with vertex-partition $V(G) = A \cup B$, where $|A| = n - \delta$ and $|B| = \delta$. The edge set is the set of all pairs which touch B. It is easy to see that $\delta(G) = \delta$. Now, orient all edges from A to B (and orient the edges inside B arbitrarily). Consider any Hamilton cycle in G, and fix a cyclic direction. Any edge pointing forward (with respect to that direction) ends at a distinct vertex of B; hence there are at most $|B| = \delta$ such edges. We note that Conjecture 1.1 holds when G is a complete graph due to the fact that every tournament has a directed Hamilton path.

Our first result is an approximate version of Conjecture 1.1.

Theorem 1.2. *Let $k \geq 0$ and let G be an n-vertex graph with $n \geq 30 + 4(k-1)$ and $\delta(G) \geq \frac{n+8k}{2}$. Then in every orientation of G, there is a Hamilton cycle with at least $\frac{n+k}{2}$ edges in the same direction.*

As mentioned above, Conjecture 1.1 can be viewed as an extension of Dirac's theorem, stating that Dirac's theorem holds in a "robust way". Namely, not only do graphs with large minimum degree (above $n/2$) have Hamilton cycles, but the collection of Hamilton cycles is robust enough to guarantee the existence of an unbalanced Hamilton cycle in every edge-orientation. It is worth mentioning that there is by now an extremely rich literature (certainly too rich to be comprehensively surveyed here) on extensions and generalisations of Dirac's theorem in various directions. We refer the reader to the surveys of Gould [12–14] and of Kühn and Osthus [17,18], as well as the survey of Sudakov [23] focusing on robustness-type results.

As our second result, we study oriented discrepancy of Hamilton cycles in random graphs. We show that if the edge probability p is above the Hamiltonicity threshold, then with high probability[1] (**whp** for short), in every orientation of $G(n, p)$ there is a Hamilton cycle with *almost all* edges oriented in the same direction.

Theorem 1.3. *Let n be an integer and let $p \geq (\log n + \log \log n + \omega(1))/n$. Then $G \sim G(n, p)$ is **whp** such that in any orientation of its edges there exists a Hamilton cycle with at least $(1 - o(1))n$ edges in the same direction.*

---

[1]i.e., with probability tending to 1 is n tends to infinity

The proof of Theorem 1.3 combines machinery from our earlier paper [10] with a result on the oriented discrepancy of almost-spanning paths in sparse expanders (see Theorem 3.1). In the special case of random graphs, this result is as follows.

Theorem 1.4. *For every $\delta > 0$ there is $C > 0$ such that $G \sim G(n, C/n)$ **whp** satisfies the following: in any orientation of the edges of G there exists a path of length at least $(1 - \delta)n$ with at most δn edges going backwards.*

Theorem 1.4 has the following interesting implication to unbalanced Hamilton cycles in tournaments. It is a well-known fact [22] that every tournament has a directed Hamilton path. Hence, in any tournament, there is a Hamilton cycle (of the underlying complete graph) in which all edges but at most one go in the same direction. It may be, however, that a tournament has only one directed Hamilton path (this is the case in the transitive tournament). It is then natural to relax the notion of "imbalance" by considering Hamilton cycles which have up to δn edges in the opposite direction, and ask how many such unbalanced Hamilton cycles are guaranteed to exist. The following corollary addresses this question.

Corollary 1.5. *For every $\delta > 0$ there is $C > 0$ such that every tournament T on n vertices contains at least $C^{-n}n^n$ Hamilton cycles (of the underlying complete graph), each of which has at least $(1 - \delta)n$ edges in the same direction.*

Corollary 1.5 is optimal up to the value of the constant C. Indeed, let T be the random tournament where each edge is oriented randomly (with probability $1/2$ for each direction). For every Hamilton cycle C in the underlying complete graph of T, the probability that C has at least $(1/2 + \varepsilon)n$ forward edges is at most $e^{-\Omega(\varepsilon^2 n)}$, by the Chernoff bound. Hence, the expected number of Hamilton cycles having at least $0.51n$ (say) forward edges, is at most $e^{-\Omega(n)}n!$. By a similar consideration, the expected number of Hamilton cycles having at least $(1 - \delta)n$ edges in the same direction is at most $c^{-n}n!$ where $c = 2^{1-O(\delta \log(1/\delta))}$. Determining the correct dependence of C on δ in Corollary 1.5 would be interesting.

Theorem 1.2 is proven in the next section, Theorems 1.3 and 1.4 are proven in Section 3 and Corollary 1.5 is proven in Section 4.

2   Unbalanced Hamilton cycles in Dirac graphs

A **diamond** is the graph with vertices a, b, c, d and edges $\{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}$. We call c, d the **ends** of the diamond, and a, b the **centre** of the diamond. Given an orientation of its edges, a diamond is called **good** (with respect to that orientation) if c and d relate in the same way to a, b. (The direction of the edge $\{a, b\}$ will not be important, as the roles of a, b are the same. For convenience, we assume that $a \to b$.) There are four good diamonds, depicted in Fig. 1.

We say that a path is **positive** (with respect to an orientation) if it consists of more forward edges than backward edges when traversing the path from its first vertex to its last. Similarly, it is **negative** if it consists of more backward edges than forward edges. The following is the key property of good diamonds that we will use. It can easily be verified by hand (see Fig. 1).

Observation 2.1. *Let D be a good diamond with ends c, d. Then D contains a positive path of length 3 from c to d, as well as a positive path of length 3 from d to c.*

Lemma 2.2. *Let G be a graph with $|V(G)| = n \geq 30$ vertices $e(G) \geq \lfloor \frac{n^2}{4} \rfloor + 1$. Then every orientation of G contains a good diamond.*
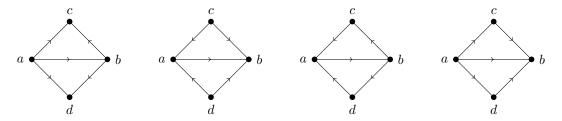
Figure 1: The four good diamonds.

Proof. Edwards [6] and independently Khadžiivanov and Nikiforov [15] proved that a graph with n vertices and $\lfloor \frac{n^2}{4} \rfloor + 1$ edges has an edge which is contained in at least $n/6$ triangles (earlier, Erdős [7] proved a bound of $n/6 - O(1)$). So let $\{a, b\} \in E(G)$ be an edge contained in at least $n/6$ triangles, and let $x_1, \ldots, x_m$, $m \geq n/6 \geq 5$ be such that a, b, x_i is a triangle in G. Fix any orientation of G. By the pigeonhole principle, there are $1 \leq i < j \leq m$ such that x_i, x_j relate in the same way (in the orientation) to a, b. Then a, b, $c := x_i$ and $d := x_j$ form a good diamond in G, as required. $\qquad\square$

Lemma 2.3. *Let $k \geq 1$, and let G be a graph with $|V(G)| = n \geq 30 + 4(k-1)$ and $e(G) \geq \frac{n^2}{4} + 2(k-1)n - 4k^2 + 6k - 1$. Then every orientation of G contains k vertex-disjoint good diamonds.*

Proof. The proof is by induction on k. Fix any orientation of G. It is easy to check that $e(G) \geq n^2/4 + 1$ by the assumptions of the lemma. By Lemma 2.2, G contains a good diamond D (this also establishes the base case $k = 1$). Let G' be the subgraph of G induced by $V(G) \setminus V(D)$. We have $|V(G')| = n - 4 \geq 30 + 4(k-2)$ and

$$e(G') \geq e(G) - 6 - 4(n-4) = e(G) - (4n - 10) \geq \frac{n^2}{4} + 2(k-1)n - 4k^2 + 6k - 1 - (4n - 10)$$

$$= \frac{(n-4)^2}{4} + 2(k-2)(n-4) - 4(k-1)^2 + 6(k-1) - 1.$$

By the induction hypothesis, G' contains $k-1$ vertex-disjoint good diamonds, finishing the proof. $\qquad\square$

We remark that Lemma 2.3 is optimal in the following sense. Consider the complete n-vertex tripartite graph $G = (A \cup B \cup C, E)$ where $|A| = k - 1$ and $|B|, |C|$ are nearly equal. It is not hard to check that G has $n^2/4 + \Theta(kn)$ edges but at most $k - 1$ vertex-disjoint diamonds.

The last ingredient in the proof of Theorem 1.2 is the following lemma due to Pósa.

Lemma 2.4 ([20]). *Let $t \geq 0$ and let G be a graph with n vertices and minimum degree at least $\frac{n+t}{2}$. Let $E \subseteq E(G)$ be the edge-set of a path forest[2] of size at most t. Then there exists a Hamilton cycle in G which uses all edges in E.*

Proof of Theorem 1.2. If $k = 0$ then the assertion follows immediately from Dirac's theorem. So suppose from now on that $k \geq 1$. Fix any orientation of G. We have $e(G) \geq \frac{n}{2} \cdot \frac{n+8k}{2} = \frac{n^2}{4} + 2kn \geq \frac{n^2}{4} + 2(k-1)n - 4k^2 + 6k - 1$, hence we may apply Lemma 2.3.

Let $D_1, \ldots, D_k$ be vertex-disjoint good diamonds in G. Let us denote by c_i, d_i the ends of D_i. For each $1 \leq i \leq k$, let $P_i \subseteq D_i$ be a positive path of length 3 from c_i to d_i, and let $Q_i \subseteq D_i$ be a positive path of length 3 from d_i to c_i; such paths exists by Observation 2.1. Let

---

[2]A path forest is a forest in which every connected component is a path.

4

$E = \bigcup_{i=1}^{k} E(P_i)$. So E is the edge-set of a path forest, and $|E| = 3k$. By Lemma 2.4, there is a Hamilton cycle H of G which contains all edges in E; in other words, H contains $P_1, \ldots, P_k$ as subpaths. Fix a direction of H which matches the orientation of the majority of the edges in $E(H) \smallsetminus E$. We now define a Hamilton cycle H' as follows.

For $i = 1, \ldots, k$, we say that D_i *matches the direction of* H if, when traversing $P_i \subseteq H$ along the fixed direction of H, we go from c_i to d_i (and not from d_i to c_i). Recall that P_i is positive when traversing it from c_i to d_i (and hence negative when traversing it from d_i to c_i). Let I be the set of $i = 1, \ldots, k$ such that P_i does *not* match the direction of H. Now let H' be the Hamilton cycle with edge-set $E(H') = \left(E(H) \smallsetminus \bigcup_{i \in I} E(P_i) \right) \cup \bigcup_{i \in I} E(Q_i)$. In other words, for each $i \in I$, we replace P_i with Q_i. It is easy to see that H' is indeed a Hamilton cycle, and that it has at least $\frac{n+k}{2}$ edges in the same direction. $\qquad\square$

3 Unbalanced Hamilton cycles in sparse expanders

To prove Theorem 1.3 we first prove Theorem 3.1 below, which extends Theorem 1.4 to the more general setting of graphs with certain expansion properties. More precisely, for $\beta > 0$, a graph G is called a β-*graph* if there is an edge in G between every pair of disjoint $U, W \subseteq V(G)$ with $|U|, |W| \geq \beta |V(G)|$. This notion was introduced in [9].

Theorem 3.1. *For every $\delta > 0$ there exists $\beta > 0$ such that the following holds. Let G be a β-graph on n vertices. Then, in any orientation of the edges of G there exists a path of length at least $(1 - \delta)n$ with at most δn edges going backwards.*

Theorem 1.4 follows from Theorem 3.1 by the well-known and simple fact that for every $\beta > 0$ there exists $C > 0$ such that $G(n, C/n)$ is **whp** a β-graph. For a proof (of a stronger statement), see e.g. [10, Lemma 3.4].

To prove Theorem 3.1, we will need the following result from [3].

Lemma 3.2 (Lemma 4.4 in [3]). *Let $m, k \geq 1$, let F be a directed graph on m vertices, and suppose that for every pair of disjoint $A, B \subseteq V(F)$ with $|A|, |B| \geq k$, there is an edge of F from A to B. Then F has a directed path of length at least $m - 2k + 1$.*

Proof of Theorem 3.1. Fix any orientation of G. Let $\beta = \beta(\delta)$ be small enough (to be chosen later). Set $\ell := \lfloor \frac{1}{\sqrt{2\beta}} \rfloor - 1$ and $m := \lceil \frac{n}{\ell+2} \rceil$, noting that $m \geq n/(\frac{1}{\sqrt{2\beta}} + 1) = \sqrt{2\beta}n/(1 + \sqrt{2\beta})$ and $m \leq \frac{2n}{\ell+2} \leq 3\sqrt{\beta}n$. We claim that G contains a collection of m vertex-disjoint directed paths of length ℓ each. We find these paths one by one. Suppose we already found paths $P_1, \ldots, P_j$, $j < m$. Let G' be the graph obtained from G by deleting $V(P_1) \cup \cdots \cup V(P_j)$. Then $|V(G')| = n - j \cdot (\ell + 1) \geq n - (m - 1)(\ell + 1) \geq \frac{n}{\ell+2}$. Observe that G' has no independent set of size $2\beta n$ because G is a β-graph. It follows that $\chi(G') \geq \frac{|V(G')|}{2\beta n} \geq \frac{1}{2\beta(\ell+2)} > \ell$, where the last inequality uses that $\ell(\ell + 2) < (\ell + 1)^2 \leq \frac{1}{2\beta}$. So $\chi(G') \geq \ell + 1$. By the Gallai–Hasse–Roy–Vitaver theorem (see e.g., [24, Theorem 5.1.21]), every orientation of a graph with chromatic number k contains a directed path of length $k - 1$. Hence, G' contains a directed path of length ℓ. This proves our claim about the existence of $P_1, \ldots, P_m$.

For each $1 \leq i \leq m$, let x_i (resp. y_i) be the first (resp. last) vertex of P_i. Define an auxiliary directed graph F on $[m]$ in which $i \rightarrow j$ if and only if $\{y_i, x_j\} \in E(G)$. We claim that for every pair of disjoint $A, B \subseteq [m]$ with $|A|, |B| \geq \beta n$, there is an edge of F from A to B. Indeed, set $X := \{x_i : i \in A\}$ and $Y := \{y_i : i \in B\}$. Since G is a β-graph, there is an edge in G between X and Y; namely there are $j \in A, i \in B$ with $\{y_i, x_j\} \in E(G)$. This gives that $i \rightarrow j$ in F.

By Lemma 3.2 with $k = \lceil \beta n \rceil$, F has a directed path of length $t := m - 2 \cdot \lceil \beta n \rceil + 1 \geq m - 2\beta n - 1$. Without loss of generality, let us assume that this path is $1, \ldots, t+1$. Then $\{y_i, x_{i+1}\} \in E(G)$ for every $1 \leq i \leq t$. Now observe that

$$x_1 \xrightarrow{P_1} y_1 \to x_2 \to \cdots \to y_t \to x_{t+1} \xrightarrow{P_{t+1}} y_{t+1}$$

is a path of length $(t+1) \cdot \ell + t \geq (t+1)\ell$ with all but at most t edges in the same direction. We have

$$(t+1)\ell \geq (m - 2\beta n) \cdot \left(\frac{1}{\sqrt{2\beta}} - 2 \right) \geq \left(\frac{\sqrt{2\beta}n}{1 + \sqrt{2\beta}} - 2\beta n \right) \cdot \left(\frac{1}{\sqrt{2\beta}} - 2 \right) \geq (1 - \delta)n,$$

provided that β is small enough in terms of δ. Also, $t \leq m \leq 3\sqrt{\beta}n \leq \delta n$. This completes the proof. $\qquad \square$

The second main tool used in the proof of Theorem 1.3 is the following lemma, which slightly generalises [10, Lemma 4.1]. This lemma allows us to absorb a long path in $G \sim G(n, p)$ into a Hamilton cycle (for p above the Hamiltonicity threshold). The precise statement is as follows:

Lemma 3.3. *For every $\delta > 0$ there exists $0 < \varepsilon \leq \delta$ for which the following holds. Let $p = (\log n + \log \log n + \omega(1))/n$ and let $G \sim G(n, p)$. Then, **whp**, there exists a partition $V(G) = V^\star \cup V'$ with $|V^\star| \leq \varepsilon n$ such that every path $P \subseteq V'$ with $|V(P)| \leq (1 - \delta)n$ can be extended to a Hamilton cycle in G.*

The proof of Lemma 3.3 is essentially the same as that of [10, Lemma 4.1]. For completeness, we give the proof in the appendix. We can now derive Theorem 1.3 from Theorem 3.1 and Lemma 3.3.

Proof of Theorem 1.3. Let $\delta > 0$ and $p = (\log n + \log \log n + \omega(1))/n$, and let $G \sim G(n, p)$. (We assume where needed that n is large enough.) By Lemma 3.3 there exist $0 < \varepsilon \leq \delta$ and, **whp**, a partition $V(G) = V^\star \cup V'$ such that $n' = |V'| \geq (1 - \varepsilon)n$, and any path $P \subseteq V'$ with $|V(P')| \leq (1 - \delta)n$ can be extended to a Hamilton cycle of G. Let β be the constant obtained from Theorem 3.1 by plugging in δ. As mentioned above, G is **whp** a β'-graph for $\beta' = \beta(1-\varepsilon)$. If this happens then $G' = G[V']$ is a β-graph. Now fix any orientation of G. By Theorem 3.1 we know that there exists a path P in G' of length at least $(1-\delta)n'$ (but also at most $(1-\delta)n$) with at most $\delta n'$ edges going backwards. By the guarantees of Lemma 3.3 we can, **whp**, extend P to a Hamilton cycle of G, having at most $2\delta n' + \varepsilon n \leq 3\delta n$ edges going backwards. The result follows by taking $\delta \to 0$. $\qquad \square$

4 Proof of Corollary 1.5

Proof of Corollary 1.5. Set $\delta' = \delta/2$, and assume for simplicity that $\delta'n$ is an integer. Let $\mathcal{P}$ be the set of all paths of length $(1 - \delta')n$ in T in which all but at most $\delta'n$ edges are in the same direction. Let C' be the constant given by applying Theorem 1.4 with δ'. Let $G \sim G(n, C'/n)$. The expected number of paths from $\mathcal{P}$ which are present in G is $(C'/n)^{(1-\delta')n} \cdot |\mathcal{P}|$. On the other hand, at least one such path is present with probability $1 - o(1)$ by Theorem 1.4. Hence, $(C'/n)^{(1-\delta')n} \cdot |\mathcal{P}| \geq 1/2$, and so $|\mathcal{P}| \geq 1/2 \cdot (n/C')^{(1-\delta')n}$. Each path in $\mathcal{P}$ can be extended in $(\delta'n - 1)!$ ways into a Hamilton cycle in which at least $(1 - \delta)n$ edges are in the same direction,

and each Hamilton cycle is counted at most n times this way. Hence, the number of such Hamilton cycles is at least

$$(\delta'n - 1)! \cdot |\mathcal{P}|/n \geq \frac{1 - o(1)}{2n} \cdot \left(\frac{\delta'n}{3}\right)^{\delta'n} \cdot \left(\frac{n}{C'}\right)^{(1-\delta')n} \geq C^{-n}n^n,$$

for a large enough constant C depending on C' and δ. $\qquad \square$

Acknowledgements: We thank the anonymous referees for a careful reading of the paper and useful comments.

References

[1] József Balogh, Béla Csaba, Yifan Jing and András Pluhár, *On the discrepancies of graphs*, Electronic Journal of Combinatorics **27** (2020), no. 2, Paper No. 2.12, 14. MR4245067 ↑1

[2] József Balogh, Béla Csaba, András Pluhár and Andrew Treglown, *A discrepancy version of the Hajnal-Szemerédi theorem*, Combinatorics, Probability and Computing **30** (2021), no. 3, 444–459. MR4247634 ↑1

[3] Ido Ben-Eliezer, Michael Krivelevich and Benny Sudakov, *The size Ramsey number of a directed path*, Journal of Combinatorial Theory. Series B **102** (2012), no. 3, 743–755. MR2900815 ↑5

[4] Domagoj Bradač, *Powers of Hamilton cycles of high discrepancy are unavoidable*, Electronic journal of combinatorics **29** (2022). ↑1

[5] Gabriel A. Dirac, *Some theorems on abstract graphs*, Proceedings of the London Mathematical Society. Third Series **2** (1952), 69–81. MR47308 ↑2

[6] Christopher S. Edwards, *A lower bound for the largest number of triangles with a common edge*, Unpublished manuscript (1977). ↑4

[7] Paul Erdős, *On a theorem of Rademacher-Turán*, Illinois Journal of Mathematics **6** (1962), 122–127. MR137661 ↑4

[8] Andrea Freschi, Joseph Hyde, Joanna Lada and Andrew Treglown, *A note on color-bias hamilton cycles in dense graphs*, SIAM Journal on Discrete Mathematics **35** (2021), no. 2, 970–975. ↑1

[9] Limor Friedman and Michael Krivelevich, *Cycle lengths in expanding graphs*, Combinatorica **41** (2021), no. 1, 53–74. MR4235314 ↑5

[10] Lior Gishboliner, Michael Krivelevich and Peleg Michaeli, *Colour-biased Hamilton cycles in random graphs*, Random Structures and Algorithms **60** (2022), 289–307. ↑1, 3, 5, 6, 8, 9, 10

[11] Lior Gishboliner, Michael Krivelevich and Peleg Michaeli, *Discrepancies of spanning trees and Hamilton cycles*, Journal of Combinatorial Theory. Series B **154** (2022), 262–291. ↑1

[12] Ronald J. Gould, *Updating the Hamiltonian problem—a survey*, Journal of Graph Theory **15** (1991), no. 2, 121–157. MR1106528 ↑2

[13] Ronald J. Gould, *Advances on the Hamiltonian problem—a survey*, Graphs and Combinatorics **19** (2003), no. 1, 7–52. MR1974368 ↑2

[14] Ronald J. Gould, *Recent advances on the Hamiltonian problem: Survey III*, Graphs and Combinatorics **30** (2014), no. 1, 1–46. MR3143857 ↑2

[15] Nikolaĭ G. Khadžiivanov and Vladimir S. Nikiforov, *Solution of a problem of P. Erdős about the maximum number of triangles with a common edge in a graph*, Doklady Bolgarskoĭ Akademii Nauk. Comptes Rendus de l'Académie Bulgare des Sciences **32** (1979), no. 10, 1315–1318. (In Russian). MR562936 ↑4

[16] Michael Krivelevich, *Long paths and hamiltonicity in random graphs*, Random graphs, geometry and asymptotic structure (Nikolaos Fountoulakis and Dan Hefetz, eds.), Cambridge University Press, 2016, pp. 4–27. ↑8

[17] Daniela Kühn and Deryk Osthus, *A survey on Hamilton cycles in directed graphs*, European Journal of Combinatorics **33** (2012), no. 5, 750–766. MR2889513 ↑2

[18] Daniela Kühn and Deryk Osthus, *Hamilton cycles in graphs and hypergraphs: an extremal perspective*, Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. IV, 2014, pp. 381–406. MR3727617 ↑2

[19] Jiří Matoušek, **Geometric discrepancy**, Algorithms and Combinatorics, vol. 18, Springer-Verlag, Berlin, 1999. MR1697825 ↑1

[20] Lajos Pósa, *On the circuits of finite graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **8** (1963), 355–361. MR197355 ↑4

[21] Lajos Pósa, *Hamiltonian circuits in random graphs*, Discrete Mathematics **14** (1976), no. 4, 359–364. MR389666 ↑8

[22] László Rédei, *Ein kombinatorischer satz*, Acta Litt. Szeged **7** (1934), no. 39-43, 97. ↑3

[23] Benny Sudakov, *Robustness of graph properties*, Surveys in combinatorics 2017, Cambridge Univ. Press, Cambridge, 2017, pp. 372–408. MR3728112 ↑2

[24] Douglas B. West, **Introduction to graph theory**, Vol. 2, Prentice Hall Upper Saddle River, 2001. ↑5

Appendix A   Proof of Lemma 3.3

As mentioned earlier, the proof is almost identical to the proof of [10, Lemma 4.1]. For the proof, we would need the following lemmas and definitions. At its core, the proof uses Pósa's rotation–extension technique (for an overview, see [16]). The following is an immediate consequence of Pósa's lemma.

Lemma A.1 (Pósa's lemma [21])**.** *Let G be a graph and let $P = v_0, \ldots, v_t$ be a longest path in G. Then there exists a set $R \subseteq V(P) \setminus \{v_0\}$ with the following properties:*

1. For every $v \in R$ there is a path P' with $V(P') = V(P)$ and with endpoints v_0 and v.

2. $|N(R)| \leq 2|R| - 1$.

Proof. Take R to be the set of endpoints of paths obtained from P via a sequence of rotations fixing v_0. Pósa's lemma [21] (see also [16]) states that $|N(R)| \leq 2|R| - 1$. ☐

Recall that a non-edge $\{x, y\}$ of G is called a booster if adding $\{x, y\}$ to G creates a graph which is either Hamiltonian or whose longest path is longer than that of G. For a positive integer k and a positive real α we say that a graph $G = (V, E)$ is a (k, α)**-expander** if $|N(U)| \geq \alpha|U|$ for every set $U \subseteq V$ of at most k vertices. The following is a widely-used fact stating that connected $(k, 2)$-expanders have many boosters. For a proof, see e.g. [16].

Lemma A.2. *Let G be a connected $(k, 2)$-expander which contains no Hamilton cycle. Then G has at least $(k + 1)^2/2$ boosters.*

Let us now introduce some additional definitions. For a pair of disjoint vertex-sets U, W in a graph, the **density** of (U, W) is defined as $d(U, W) := |E(U, W)|/(|U||W|)$. For $\beta, p \in (0, 1]$, we say that $G = (V, E)$ is (β, p)**-pseudorandom** if for any two disjoint $U, W \subseteq V$ with $|U|, |W| \geq \beta|V|$ we have $|d(U, W) - p| < \beta p$. Note that a (β, p)-pseudorandom graph is, in particular, a β-graph.

We now move on to establish some useful properties satisfied **whp** by $G(n, p)$, where p is above the Hamiltonicity threshold.

Lemma A.3 ([10, Lemmas 4.4 and 3.4])**.** *Let $\varepsilon > 0$ be sufficiently small, let $p = (\log n + \log \log n + \omega(1))/n$, let $G \sim G(n, p)$ and let $\beta > 0$. Then, **whp**,*

(P1) $\delta(G) \geq 2$ *and* $\Delta(G) \leq 10 \log n$.

(P2) *No vertex $v \in V(G)$ with $d(v) < \log n/10$ is contained in a 3- or a 4-cycle, and every two distinct vertices $u, v \in V(G)$ with $d(u), d(v) < \log n/10$ are at distance at least 5 apart.*

(P3) *Every set $U \subseteq V(G)$ of size at most $\varepsilon n/100$ spans at most $\varepsilon |U| \log n/10$ edges.*

(P4) *There exist disjoint sets $U_1, U_2 \subseteq V(G)$ with $|U_1|, |U_2| \leq \varepsilon n$ for which the following hold for every $v \in V(G)$:*

    (a) *If $d(v) \geq \log n/10$ then $d(v, U_1), d(v, U_2) \geq \varepsilon \log n/100$;*

    (b) *If $d(v) \leq \log n/10$ then v and all of its neighbours are in U_1.*

(P5) *G is (β, p)-pseudorandom.*

Lemma A.4. *For every $\zeta > 0$ there exists $c > 0$ such that the following holds. Let $p = (\log n + \log \log n + \omega(1))/n$ and let $G \sim G(n, p)$. Then, **whp**, G satisfies the following: for every $W \subseteq V(G)$ of size $|W| \geq \zeta n$ and for every $(|W|/4, 2)$-expander H on W which is a subgraph of G and has at most $cn \log n$ edges, if H is not Hamiltonian then G contains a booster with respect to H.*

Proof. The proof follows that of [10, Lemma 4.5]. We use a first moment argument. Evidently, the number of choices for the set W is at most 2^n. Let us fix a choice of W. For each t, the number of choices of H for which $|E(H)| = t$ is at most

$$\binom{\binom{|W|}{2}}{t} \leq \binom{n^2}{t} \leq \left(\frac{en^2}{t}\right)^t.$$

Now let H be a non-Hamiltonian $(|W|/4, 2)$-expander on W, and set $t := |E(H)|$. It is known and easy to show that if a graph H is a $(|V(H)|/4, 2)$-expander then H is connected. By Lemma A.2, H has at least $(|W|/4)^2/2 = |W|^2/32 \geq \zeta^2 n^2/32$ boosters. Now, the probability that G contains H, but no booster thereof is at most

$$p^t \cdot (1-p)^{\zeta^2 n^2/32} \leq p^t \cdot \left(1 - \frac{\log n}{n}\right)^{\zeta^2 n^2/32} \leq \left(\frac{2\log n}{n}\right)^t \cdot \exp\left(-\zeta^2 n \log n/32\right).$$

Summing over all choices of W and H, we see that the probability that the assertion of the lemma does not hold is at most

$$2^n \cdot \exp\left(-\zeta^2 n \log n/32\right) \cdot \sum_{t=1}^{cn \log n} \left(\frac{2en \log n}{t}\right)^t. \tag{1}$$

Setting $g(t) := (2en \log n/t)^t$, we note that $g'(t) = g(t) \cdot (\log(2en \log n/t) - 1) > 0$ for every t in the range of the sum in (1), assuming $c < 1$, say. Thus, this sum is not larger than

$$cn \log n \cdot (2e/c)^{cn \log n} = \exp\left(\left(\log(2e/c) \cdot c + o(1)\right)n \log n\right).$$

Now, if c is small enough so that $\log(2e/c) \cdot c < \zeta^2/32$, we get that (1) tends to 0 as n tends to infinity. This completes the proof. $\square$

    The following lemma ([10, Lemma 4.6]) states that a graph possessing certain simple properties is necessarily an expander. Such statements are common in the study of Hamiltonicity of random graphs.

Lemma A.5 ([10, Lemma 4.6]). *Let $m, d \geq 1$ be integers and let H be a graph on $h \geq 4m$ vertices satisfying the following properties:*

1. $\delta(H) \geq 2$;

2. *No vertex $v \in V(H)$ with $d(v) < d$ is contained in a 3- or a 4-cycle, and every two distinct vertices $u, v \in V(H)$ with $d(u), d(v) < d$ are at distance at least 5 apart;*

3. *Every set $U \subseteq V(H)$ of size at most $5m$ contains at most $d|U|/10$ edges;*

4. *There is an edge between every pair of disjoint sets $U_1, U_2 \subseteq V(H)$ of size m each.*

Then H is an $(h/4, 2)$-expander.

Lemma A.6. *Let $\varepsilon, \zeta > 0$. For $p = (\log n + \log \log n + \omega(1))/n$, the random graph $G \sim G(n, p)$ satisfies the following **whp**. Let $W \subseteq V(G)$ be such that $|W| \geq \zeta n$, and for every $v \in W$ it holds that $d(v, W) \geq \min\{d(v), \varepsilon \log n\}$. Then for every $w \in W$ there exists $Y \subseteq W$ with $|Y| \geq \zeta n/4$ such that for each $y \in Y$, there is a Hamilton path in $G[W]$ whose endpoints are w and y.*

Proof. The proof follows that of [10, Lemma 4.7]. We may and will assume ε is sufficiently small (it is enough to have $\varepsilon \leq \min\{1/10, 125\zeta, c/2\}$, where $c = c(\zeta)$ is the constant from Lemma A.4). We will assume that the events defined in Lemmas A.3 and A.4 hold, where we use (P5) in Lemma A.3 with $\beta = \varepsilon/500$. These events hold **whp**. We will show that then, deterministically, the assertion of Lemma A.6 holds.

It will be convenient to set $d_0 := \varepsilon \log n$. Let $W \subseteq V(G)$ be as in the statement of Lemma A.6. We select a random spanning subgraph H of $G[W]$ as follows. For each $v \in W$, if $d(v, W) < d_0$ then add to H all edges of $G[W]$ incident to v. Otherwise, namely if $d(v, W) \geq d_0$, then randomly select a set of d_0 edges of $G[W]$ incident to v and add these to H. Note that $|E(H)| \leq |W| \cdot d_0 \leq \varepsilon n \log n$. On the other hand, our assumption that $d(v, W) \geq \min\{d(v), \varepsilon \log n\}$ for every $v \in W$ implies that $\delta(H) \geq \min\{\delta(G), d_0\}$. Hence, as $d_0 \geq 2$ (for large enough n), we have $\delta(H) \geq 2$ by Property (P1) of Lemma A.3.

We claim that with positive probability (in fact, **whp**), H is a $(|W|/4, 2)$-expander. In light of Lemma A.5, it is sufficient to show that with positive probability, H satisfies Conditions 1–4 in that lemma. Here, we will choose the parameters of Lemma A.5 as $d := d_0$ and $m := \varepsilon n/500$. Note that $|V(H)| = |W| = \zeta n \geq \varepsilon n/125 = 4m$, as required by Lemma A.5. We already showed that $\delta(H) \geq 2$ (which is Condition 1 in Lemma A.5). Let us prove Condition 2. Observe that if $v \in V(H) = W$ has degree less than d_0 in H, then also $d(v, W) < d_0$ in G (by the definition of H). But then, since $d(v, W) \geq \min\{d(v), d_0\}$ by the assumption of Lemma A.6, we have $d(v) < d_0 = \varepsilon \log n \leq \log n/10$. Now, by (P2) in Lemma A.3, there is no 3- or 4-cycle in G containing a vertex v whose degree in H is less than d_0, and every two such vertices are at distance at least 5 apart.

Condition 3 in Lemma A.5 holds because H is a subgraph of G and due to Property (P3) in Lemma A.3 (note that $5m = \varepsilon n/100$). Let us now prove that Condition 4 holds. Let $U_1, U_2 \subseteq V(H) = W$ be disjoint sets satisfying $|U_1|, |U_2| = m = \varepsilon n/500$. Since G is (β, p)-pseudorandom with $\beta = \varepsilon/500$, we have

$$|E_G(U_1, U_2)| \geq (1 - \beta)p \cdot |U_1||U_2| \geq \frac{|U_1||U_2| \log n}{2n} \geq \frac{\varepsilon^2 n \log n}{500000} = \Omega(n \log n) . \tag{2}$$

Now, let us bound (from above) the probability that $|E_H(U_1, U_2)| = 0$ (where the randomness is with respect to the choice of H). Recall that H is defined by choosing, for each $v \in W$, a random

10

set $E(v)$ of $\min\{d(v,W), d_0\}$ edges of $G[W]$ incident to v, with all choices made uniformly and independently, and letting $E(H) = \bigcup_{v \in W} E(v)$. Fix any $u_1 \in U_1$ with $d(u_1, U_2) \geq 1$, and let $\mathcal{A}_{u_1}$ be the event that there is no edge in $E(u_1)$ with an endpoint in U_2. Observe that if $d(u_1, W) < d_0$ then $\mathbb{P}(\mathcal{A}_{u_1}) = 0$, and otherwise

$$\mathbb{P}(\mathcal{A}_{u_1}) = \binom{d(u_1,W) - d(u_1,U_2)}{d_0} \Big/ \binom{d(u_1,W)}{d_0} = \prod_{i=0}^{d_0-1} \frac{d(u_1,W) - d(u_1,U_2) - i}{d(u_1,W) - i}$$

$$\leq \left(1 - \frac{d(u_1,U_2)}{d(u_1,W)}\right)^{d_0} \leq \left(1 - \frac{d(u_1,U_2)}{\Delta(G)}\right)^{d_0} \leq e^{-d(u_1,U_2)\cdot\frac{d_0}{\Delta(G)}} \leq e^{-\varepsilon d(u_1,U_2)/10}.$$

Here, in the last inequality we used Property (P1) in Lemma A.3. Note that the events $(\mathcal{A}_{u_1} : u_1 \in U_1)$ are independent, and that if $E_H(U_1, U_2) = \varnothing$ then $\mathcal{A}_{u_1}$ occurred for every $u_1 \in U_1$ with $d(u_1, U_2) \geq 1$. It now follows that

$$\mathbb{P}(E_H(U_1,U_2) = \varnothing) \leq \exp\left(-\frac{\varepsilon}{10} \cdot \sum_{u_1 \in U_1} d(u_1,U_2)\right) = \exp\left(-\frac{\varepsilon}{10} \cdot |E_G(U_1,U_2)|\right) \leq e^{-\Omega(n \log n)},$$

where in the last inequality we used (2). By taking the union bound over all at most 2^{2n} choices of U_1, U_2, we see that with high probability, $E_H(U_1, U_2) \neq \varnothing$ for every pair of disjoint sets $U_1, U_2 \subseteq W$ of size m each.

Finally, we apply Lemma A.5 to conclude that **whp** H is a $(|W|/4, 2)$-expander. From now on, we fix such a choice of H. Before establishing the assertion of the lemma, we first show that $G[W]$ is Hamiltonian. To find a Hamilton cycle in $G[W]$, we define a sequence of graphs H_i, $i \geq 0$, as follows. To begin, set $H_0 = H$. For each $i \geq 0$, if H_i is Hamiltonian then stop, and otherwise take a booster of H_i contained in $G[W]$ and add it to H_i to obtain H_{i+1}. That such a booster exists is guaranteed by Lemma A.4, as we will always have $|E(H_i)| \leq |E(H)| + |W| \leq |E(H)| + n \leq \varepsilon n \log n + n \leq c/2 \cdot n \log n + n \leq cn \log n$, provided that ε is smaller than $c/2$, where c is the constant appearing in Lemma A.4. Note also that H_i is a subgraph of $G[W]$ for each $i \geq 0$. Evidently, this process has to stop (because as long as H_i is not Hamiltonian, the maximum length of a path in H_i is longer than in H_{i-1}), thus showing that $G[W]$ must contain a Hamilton cycle, as claimed.

Now let $w \in W$. As $G[W]$ is Hamiltonian, there exists a Hamilton path P of $G[W]$ such that w is one of the endpoints of P. Evidently, P is a longest path in $G[W]$. Furthermore, note that $G[W]$ is a $(|W|/4, 2)$-expander because H, a subgraph of $G[W]$, is such an expander. By Lemma A.1 (applied to $G[W]$), there exists a set $R \subseteq V(P) = W$ such that for every $y \in R$ there is a Hamilton path in $G[W]$ with endpoints w and y, and such that $|N_{G[W]}(R)| \leq 2|R|-1$. Now, since $G[W]$ is a $(|W|/4, 2)$-expander, it must be the case that $|R| > |W|/4 \geq \zeta n/4$. So we see that the assertion of the lemma holds with $Y = R$. This completes the proof. $\square$

Proof of Lemma 3.3. Set $\varepsilon = \delta/3$. For convenience, we show the existence of a partition $V(G) = V^\star \cup V'$ with $|V^\star| \leq 2\varepsilon n$ instead of $|V^\star| \leq \varepsilon n$ (this clearly does not matter). We assume that G satisfies the properties detailed in Lemma A.3 for $\beta < \delta/40$ and the assertion of Lemma A.6. These events happen **whp**. Let U_1, U_2 be disjoint subsets of $V = V(G)$ satisfying (P4). Set $V^\star = U_1 \cup U_2$ and $V' = V \smallsetminus V^\star$, and let $P \subseteq V'$ be a path with $|V(P)| \leq (1-\delta)n$ and endpoints a_1, a_2. In particular, $|V^\star| \leq 2\varepsilon n$. Our goal is to extend P to a Hamilton cycle of G. Write $V'' = V' \smallsetminus V(P)$, partition $V'' = V_1'' \cup V_2''$ as equally as possible, so $|V_i''| \geq \lfloor (\delta-2\varepsilon)n/2 \rfloor \geq \delta n/10$. For $i = 1, 2$, let $W_i = V_i'' \cup U_i$ and choose a neighbour w_i of a_i in W_i; this is possible since
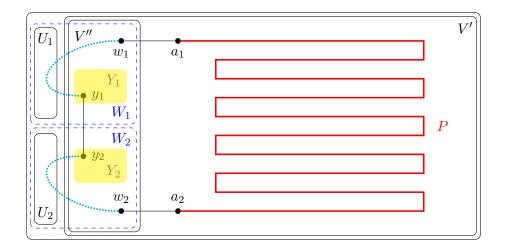
11

Figure 2: Outline of the proof of Lemma 3.3.

$d(a_i, U_i) \geq \varepsilon \log n / 100$ by (P4). Note that $|W_i| \geq \delta n / 10$ and for every $v \in W_i$ it holds that $d(v, W_i) \geq d(v, U_i) \geq \min\{d(v), \varepsilon \log n / 100\}$ by (P4). Hence, by Lemma A.6 (with parameters $\zeta = \delta/10$ and $\varepsilon/100$), there exists a set $Y_i \subseteq W_i$ with $|Y_i| \geq \delta n / 40$ such that for every $y \in Y_i$ there is a Hamilton path spanning W_i from w_i to y. Since G is (β, p)-pseudorandom for $\beta < \delta/40$, there is an edge e in G between Y_1 and Y_2 with endpoints $y_i \in Y_i$, say. For $i = 1, 2$, denote by Q_{y_i} the Hamilton path between w_i and y_i. We now construct a Hamilton cycle of G containing P as follows (as depicted in Fig. 2):

$$a_1 \to w_1 \xrightarrow{Q_{y_1}} y_1 \xrightarrow{e} y_2 \xrightarrow{Q_{y_2}} w_2 \to a_2 \xrightarrow{P} a_1. \qquad \square$$