Random Directed Graphs Are Robustly Hamiltonian*

Dan Hefetz,[1] Angelika Steger,[2] Benny Sudakov[3]
[1] *School of Mathematics, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK; e-mail: d.hefetz@bham.ac.uk*
[2] *Institute of Theoretical Computer Science, ETH, Zürich CH-8092, Switzerland; e-mail: steger@inf.ethz.ch*
[3] *Department of Mathematics, ETH, 8092 Zurich, Switzerland; e-mail: benjamin.sudakov@math.ethz.ch*

Received 18 April 2014; revised 21 April 2015; accepted 10 August 2015
Published online 24 December 2015 in Wiley Online Library (wileyonlinelibrary.com).
DOI 10.1002/rsa.20631

ABSTRACT: A classical theorem of Ghouila-Houri from 1960 asserts that every directed graph on n vertices with minimum out-degree and in-degree at least $n/2$ contains a directed Hamilton cycle. In this paper we extend this theorem to a random directed graph $\mathcal{D}(n,p)$, that is, a directed graph in which every ordered pair (u,v) becomes an arc with probability p independently of all other pairs. Motivated by the study of resilience of properties of random graphs, we prove that if $p \gg \log n/\sqrt{n}$, then a.a.s. every subdigraph of $\mathcal{D}(n,p)$ with minimum out-degree and in-degree at least $(1/2+o(1))np$ contains a directed Hamilton cycle. The constant $1/2$ is asymptotically best possible. Our result also strengthens classical results about the existence of directed Hamilton cycles in random directed graphs. © 2015 Wiley Periodicals, Inc. Random Struct. Alg., 49, 345–362, 2016

Keywords: random digraph; Hamilton cycle; resilience

1. INTRODUCTION

A Hamilton cycle of a graph G is a cycle which passes through every vertex of G exactly once. A graph is said to be *Hamiltonian* if it admits a Hamilton cycle. Hamiltonicity is one of the most central notions in graph theory, and has been intensively studied by numerous researchers for many years. The problem of deciding whether a graph is Hamiltonian or not is one of the NP-complete problems that Karp listed in his seminal paper [24], and, accordingly, unless $P = NP$, one cannot hope for a simple classification of such graphs. It is

Correspondence to: Dan Hefetz
*Supported by EPSRC (to D.H.) (EP/K033379/1); SNSF (to B.S.) (200021–149111).

thus important to find general, applicable sufficient conditions for graphs to be Hamiltonian and in the last 60 years many interesting results were obtained in this direction. One of the first results of this type is the celebrated theorem of Dirac [8], asserting that every graph on $n \geq 3$ vertices with minimum degree at least $n/2$ (such graphs are called Dirac graphs) is Hamiltonian.

Dirac's Theorem provides a natural and easy to check sufficient condition for the Hamiltonicity of graphs with very high minimum degree. On the other hand, there are of course Hamiltonian graphs with minimum degree 2. Therefore, while Dirac's Theorem is sharp in general, one would like to have sufficient conditions for the Hamiltonicity of sparser graphs. When looking for such sufficient conditions, it is natural to consider random graphs with an appropriate edge probability. Erdős and Rényi [9] raised the question of what the threshold probability of Hamiltonicity in random graphs is. After a series of efforts by various researchers, including Korshunov [28] and Pósa [34], the problem was finally solved by Komlós and Szemerédi [27] and independently by Bollobás [5], who proved that if $p \geq (\log n + \log\log n + \omega(1))/n$, where $\omega(1)$ tends to infinity with n arbitrarily slowly, then $G(n,p)$ is asymptotically almost surely (or a.a.s. for brevity) Hamiltonian. Note that this is best possible since for $p \leq (\log n + \log\log n - \omega(1))/n$ a.a.s. there are vertices of degree at most one in $G(n,p)$ (see, e.g. [4]). An even stronger result was obtained by Bollobás [5]. He proved that for the random graph process, the hitting time for Hamiltonicity is exactly the same as the hitting time for having minimum degree 2, that is, a.a.s. the very edge which increases the minimum degree to 2 also makes the graph Hamiltonian.

In recent years there has been a lot of interest in proving that certain families of graphs, like Dirac graphs or random graphs, are Hamiltonian in some robust sense. Results in this direction include showing that such graphs admit not only one Hamilton cycle but many (see, e.g. [6, 7, 18, 22]), that they admit many pairwise edge-disjoint Hamilton cycles (see, e.g. [11, 14, 25, 30, 31]), that a player can claim all edges of a Hamilton cycle of these graphs, even when facing an optimal adversary (see, e.g. [1, 10, 20, 29, 35]), and many more.

The measure of how robust a graph is with respect to Hamiltonicity that we use in this paper is via the notion of *local resilience*, which was introduced by Vu and the third author in [36]. Let G be a simple graph and let $\mathcal{P}$ be a monotone increasing graph property. The *local resilience* of G with respect to $\mathcal{P}$, denoted by $r_\ell(G,\mathcal{P})$, is the smallest non-negative integer r such that one can obtain a graph which does not satisfy $\mathcal{P}$ by deleting at most r edges at every vertex of G. Namely,

$$r_\ell(G,\mathcal{P}) = \min\{r : \exists H \subseteq G \text{ such that } \Delta(H) = r \text{ and } G \setminus H \notin \mathcal{P}\}.$$

Let $\mathcal{H}$ denote the graph property of being Hamiltonian. Using the notion of local resilience, one can restate the aforementioned result of Dirac [8] as $r_\ell(K_n,\mathcal{H}) = \lfloor n/2 \rfloor = (1/2 + o(1))n$. Following a series of results (see [2, 3, 15, 36]), it was proved by Lee and Sudakov [32] that a.a.s. $r_\ell(G(n,p),\mathcal{H}) = (1/2 + o(1))np$ for every $p \gg \log n/n$. This is a far reaching generalization of Dirac's Theorem, since the complete graph on n vertices is also a random graph $G(n,p)$ with $p = 1$.

In this paper we aim to prove analogous results for directed graphs (or digraphs for brevity). Similarly to the case of undirected graphs, we define the local resilience of a digraph D with respect to a monotone increasing digraph property $\mathcal{P}$ to be the smallest non-negative integer r such that one can obtain a digraph which does not satisfy $\mathcal{P}$ by deleting at most r out-going arcs and at most r in-going arcs at every vertex of D. Namely,

$$r_\ell(D,\mathcal{P}) = \min\{r : \exists H \subseteq D \text{ such that } \Delta^+(H) \leq r, \Delta^-(H) \leq r \text{ and } D \setminus H \notin \mathcal{P}\}.$$

For a positive integer n and $0 \leq p = p(n) \leq 1$, let $\mathcal{D}(n,p)$ denote the probability space of random labeled *directed* graphs with vertex set $[n] := \{1, 2, \ldots, n\}$. That is, for every ordered pair (u, v) with $1 \leq u \neq v \leq n$ we flip a coin, all coin flips being mutually independent. With probability p we include the arc (u, v) in our digraph and with probability $1 - p$ we do not. By abuse of notation we sometimes use $\mathcal{D}(n,p)$ to denote a single element of the space $\mathcal{D}(n,p)$. We also use $\mathcal{H}$ to denote the *digraph* property of being Hamiltonian, that is, a digraph D satisfies $D \in \mathcal{H}$ if and only if D admits a directed Hamilton cycle.

Similarly to the aforementioned results of Komlós and Szemerédi and of Bollobás regarding the Hamiltonicity of random undirected graphs, results of McDiarmid [33] and Frieze [13] imply that a.a.s. $\mathcal{D}(n,p)$ is Hamiltonian for every $p \geq (1 + o(1)) \log n/n$ but admits no Hamilton cycles when $p \leq (1 - o(1)) \log n/n$. Moreover, a classical analog of Dirac's Theorem for directed graphs was proved in 1960 by Ghouila-Houri [19]. It asserts that every directed graph on n vertices with minimum out-degree and in-degree at least $n/2$ contains a directed Hamilton cycle. Stating this result in terms of local resilience we have that $r_\ell(\mathcal{D}(n, 1), \mathcal{H}) = \lfloor n/2 \rfloor$. Hence it is natural to ask whether one can generalize the theorem of Ghouila-Houri to sparse random directed graphs, similarly to the generalization of Dirac's Theorem proved by Lee and Sudakov in [32]. In this paper we obtain such a result for every p which is not too small.

Theorem 1.1. *For every fixed $\alpha > 0$, if the arc probability of the random digraph $\mathcal{D}(n,p)$ satisfies $\log n/\sqrt{n} \ll p = p(n) \leq 1$, then a.a.s.*

$$(1/2 - \alpha)np \leq r_\ell(\mathcal{D}(n,p), \mathcal{H}) \leq (1/2 + \alpha)np.$$

Our proof of the upper bound in Theorem 1.1 is very simple. On the other hand, our proof of the lower bound in Theorem 1.1 is quite involved and very technical. Therefore, following the suggestion of a referee, for the sake of clarity we will only prove it for constant p. The interested reader can find the complete proof in [21]. Moreover, in Section 5 we will briefly indicate the main changes which are needed in order to adapt our proof to the case $p = o(1)$.

1.1. Notation and Preliminaries

For the sake of simplicity and clarity of presentation, we do not make a particular effort to optimize some of the constants obtained in our proofs. We also omit floor and ceiling signs whenever these are not crucial. Most of our results are asymptotic in nature and whenever necessary we assume that n is sufficiently large. Throughout the paper, log stands for the natural logarithm, unless explicitly stated otherwise. We say that a graph property $\mathcal{P}$ holds *asymptotically almost surely*, or a.a.s. for brevity, if the probability of satisfying $\mathcal{P}$ tends to 1 as the number of vertices n tends to infinity. Our graph-theoretic notation is standard and follows that of [38]. In particular, we use the following.

For a directed graph (or *digraph* for brevity) D, let $V(D)$ and $E(D)$ denote its sets of vertices and arcs respectively, and let $v(D) = |V(D)|$ and $e(D) = |E(D)|$. For disjoint sets $A, B \subseteq V(D)$, let $E_D(A, B)$ denote the set of arcs of D which are oriented from some vertex of A towards some vertex of B, and let $e_D(A, B) = |E_D(A, B)|$. Let $d_D(A, B) = \frac{e_D(A,B)}{|A||B|}$ denote the *directed density* of the ordered pair (A, B) in D. For a vertex $u \in V(D)$ and a set $Y \subseteq V(D)$ let $N_D^+(u, Y) = \{y \in Y : (u, y) \in E(D)\}$ denote the set of *out-neighbors* of u in Y, and let $\deg_D^+(u, Y) = |N_D^+(u, Y)|$. Similarly, let $N_D^-(u, Y) = \{y \in Y : (y, u) \in E(D)\}$

denote the set of *in-neighbors* of u in Y, and let $\deg_D^-(u,Y) = |N_D^-(u,Y)|$. We abbreviate $\deg_D^+(u,V(D))$ to $\deg_D^+(u)$ and $\deg_D^-(u,V(D))$ to $\deg_D^-(u)$. Moreover, $\deg_D^+(u)$ is referred to as the out-degree of u and $\deg_D^-(u)$ is referred to as the in-degree of u. Let $\delta^+(D) = \min_{u\in V(D)} \deg_D^+(u)$ and $\delta^-(D) = \min_{u\in V(D)} \deg_D^-(u)$ denote the minimum out-degree of D and the minimum in-degree of D, respectively. Similarly, let $\Delta^+(D) = \max_{u\in V(D)} \deg_D^+(u)$ and $\Delta^-(D) = \max_{u\in V(D)} \deg_D^-(u)$ denote the maximum out-degree of D and the maximum in-degree of D, respectively. Sometimes, if there is no risk of confusion, we discard the subscript D in the above notation.

Throughout the paper we will make use of the following well-known bounds on the lower and upper tails of the binomial distribution due to Chernoff (see e.g. [23]).

Theorem 1.2 (Chernoff bounds). *Let $X \sim Bin(n,p)$ and let $0 \le \varepsilon \le 1$. Then*

i. $Pr(X \le (1-\varepsilon)np) \le \exp\left\{-\frac{\varepsilon^2 np}{2}\right\}$.

ii. $Pr(X \ge (1+\varepsilon)np) \le \exp\left\{-\frac{\varepsilon^2 np}{3}\right\}$.

iii. $Pr(|X-np| \ge \varepsilon np) \le 2\exp\left\{-\frac{\varepsilon^2 np}{3}\right\}$.

The rest of this paper is organized as follows. In Section 2 we briefly outline some of the main ideas of our proof of Theorem 1.1. In Section 3 we discuss some tools that will be used in the proof of Theorem 1.1, most notably, the diregularity Lemma. In Section 4 we prove Theorem 1.1 for constant p. In Section 5 we discuss some of the changes we need to make in the proof of Theorem 1.1 in case $p = o(1)$. Finally, in Section 6 we present some open problems.

2. A SHORT OUTLINE OF THE PROOF OF OUR MAIN RESULT FOR CONSTANT *p*

Since our proof of the lower bound in Theorem 1.1 is quite involved (even in the special case of constant p), we first sketch very briefly some of its main ideas. Some of the concepts and tools we use, will only be stated precisely and proved in the following sections.

Let $D \in \mathcal{D}(n,p)$ and let $D' = (V,E)$ be a digraph obtained from D by deleting at each vertex $u \in V(D)$ at most $(1/2-\alpha)\deg_D^+(u)$ out-going arcs and at most $(1/2-\alpha)\deg_D^-(u)$ in-going arcs. Note that both the out-degree and the in-degree in D of every vertex $u \in V$ is a.a.s. roughly np and therefore both the out-degree and the in-degree in D' of every vertex $u \in V$ is a.a.s. at least $(1/2+\alpha-o(1))np$.

Let $\varepsilon > 0$ and β be real numbers satisfying $\varepsilon \ll \beta \ll \alpha$. Apply the Directed Regularity Lemma (see Section 3 for more details) to $D' = (V,E)$ with appropriate parameters. Let $\{V_0, V_1, \ldots, V_k\}$ be the corresponding ε-regular partition. For some appropriately chosen $d > 0$ let $R = R(D',d)$ be the corresponding regularity digraph; that is, the directed graph with vertex set $\{v_1, v_2, \ldots, v_k\}$ such that for every $1 \le i \ne j \le k$, (v_i, v_j) is an arc of R if and only if (V_i, V_j) is ε-regular with directed density at least d.

We claim that R contains a subgraph R' on at least $(1-\beta)k$ vertices and with minimum out-degree and in-degree at least $|R'|/2$. Indeed, suppose not. Then by recursively deleting vertices whose out-degree or in-degree is strictly smaller than half the number of vertices, we would delete at least βk vertices. By symmetry we may assume that half of them have too small an in-degree. Since $\beta \gg \varepsilon$ by assumption, it follows that the majority of these missing arcs correspond to ε-regular pairs and thus have to have density less than d. Since $\beta \ll \alpha$ by assumption, one can check that it follows that in order to obtain D' from D we have deleted strictly more than $(1/2 - \alpha)\deg_D^-(u)$ in-going arcs at some vertex $u \in V(D)$, contrary to our assumption. By Ghouila-Houri's Theorem [19] (see Theorem 4.1) the subgraph R' is Hamiltonian. Equivalently, there exists an almost spanning cycle $C_R : v_1, v_2, \ldots, v_r, v_1$ of R. This corresponds to a directed "cycle" $C : V_1, V_2, \ldots, V_r, V_1$ of D'. Note that, by the definition of R, the pair $(V_i, V_{(i \mod r)+1})$ is ε-regular with positive directed density for every $1 \leq i \leq r$.

In order to build a directed Hamilton cycle of D', we will first build an almost spanning cycle C_1 and then use it to absorb all the remaining vertices. In order to add some vertex u to C_1 we will find an arc $(x, y) \in E(C_1)$ such that $(x, u) \in E(D')$ and $(u, y) \in E(D')$ and will then remove (x, y) from C_1 and add to it (x, u) and (u, y). In order for this to work, when building C_1 we will have to ensure that there exists a pairing (in the sense suggested above) of all vertices of $V \setminus V(C_1)$ with certain arcs of C_1; we refer to this as our *main task*. From now on we focus on building C_1.

We build C_1 such that it includes almost all vertices of V_i for every $1 \leq i \leq r$, as well as a few other problematic vertices. We do so by building a long path and then closing it to a cycle. We add vertices to the path one by one. Except when dealing with problematic vertices, this is done in the order indicated by C, that is, if the last vertex added to the path is $x \in V_s$, then the next vertex added to the path will be an out-neighbor of x in $V_{(s \mod r)+1}$. Except when dealing with problematic vertices, we choose only *nice* vertices, that is, vertices that have roughly the right number of out-neighbors in the next set along C. Thinking ahead to the moment at which we will want to close the directed path we are building into a directed cycle, we start building C_1 at a vertex $v_0 \in V_1$ which is nice and has many in-neighbors in V_r. We refrain from touching some predetermined subset of those in-neighbors until we attempt to close the directed path we built into a directed cycle.

In the digraph D a *typical* vertex u has roughly $|V_i|p$ in-neighbors and roughly $|V_i|p$ out-neighbors in V_i for every $1 \leq i \leq r$. A simple calculation shows that a.a.s. there is only a very small number of atypical vertices. Our first task is to build a directed path which includes all of these atypical vertices. Let u be an arbitrary atypical vertex which we wish to add to the path we have built thus far. It follows by the aforementioned lower bounds on the minimum in-degree and the minimum out-degree of D' that there must exist indices $1 \leq j_1, j_2 \leq r$ such that u has many in-neighbors in V_{j_1} and many out-neighbors in V_{j_2}. We walk along C (always choosing nice vertices as described above) until we reach V_{j_1}. Using regularity we can ensure that we enter V_{j_1} in an in-neighbor of u. We can thus add u to the path and proceed to a nice vertex of V_{j_2}. Once all atypical vertices are included in the path we focus on our main task.

We continue moving along C as before except that, at every step, the new vertex we add to the path is chosen uniformly at random from all nice vertices. Using this randomness we wish to show that for every vertex u which will not be included in C_1, there will be many times in which we claim an in-neighbor of u followed by an out-neighbor of u; each such time is referred to as a *successful trial*. In our analysis we use known bounds on the tail of the binomial distribution. Hence, in order to ensure the independence of trials which is needed for the binomial distribution and in order to bound from below the probability that a single trial is successful, we will only consider arcs of C_1 which are far from each other, as trials for each specific vertex $u \in V \setminus V(C_1)$.

Once the path we built is sufficiently long, we close it into a cycle. In order to do so we simply walk along the cycle C until we reach V_{r-2}. Using regularity we can now extend the path by two more arcs such that the second arc touches some $x \in V_r$ which is an in-neighbor of v_0. Claiming (x, v_0) completes the cycle C_1.

Our random procedure for building (the main part of) the directed path (see above) ensures that a.a.s. there will be strictly more than $|V \setminus V(C_1)|$ successful trials for every vertex $u \in V \setminus V(C_1)$. We can therefore greedily add all the vertices of $V \setminus V(C_1)$ to our directed cycle.

3. THE DIREGULARITY LEMMA AND PROPERTIES OF RANDOM DIRECTED GRAPHS

The Diregularity Lemma is a version of Szemerédi's Regularity Lemma (see [37]) for directed graphs. Before stating the lemma, we introduce the relevant terminology. Let H be a directed bipartite graph with bipartition $V(H) = A \cup B$ and let $\varepsilon > 0$. We say that the *ordered* pair (A, B) is *ε-regular* if $|d_H(X, Y) - d_H(A, B)| \leq \varepsilon$ holds for every $X \subseteq A$ and every $Y \subseteq B$ such that $|X| \geq \varepsilon|A|$ and $|Y| \geq \varepsilon|B|$.

Let $D = (V, E)$ be a digraph. A partition $\{V_0, V_1, \ldots, V_k\}$ of V in which the, possibly empty, set V_0 has been singled out as an exceptional set, is called an *ε-regular partition* if it satisfies the following conditions:

i. $|V_0| \leq \varepsilon|V|$;
ii. $|V_1| = \ldots = |V_k|$;
iii. all but at most εk^2 of the pairs (V_i, V_j), where $1 \leq i \neq j \leq k$, are ε-regular.

Remark 3.1. *It follows from Property* (*iii*) *above that, if* $\{V_0, V_1, \ldots, V_k\}$ *is an* ε-regular partition, *then there are at most* $\sqrt{\varepsilon}k$ *indices* $1 \leq i \leq k$ *for which there are at least* $\sqrt{\varepsilon}k$ *indices* $1 \leq j \neq i \leq k$ *such that* (V_i, V_j) *is not* ε*-regular. Similarly, there are at most* $\sqrt{\varepsilon}k$ *indices* $1 \leq i \leq k$ *for which there are at least* $\sqrt{\varepsilon}k$ *indices* $1 \leq j \neq i \leq k$ *such that* (V_j, V_i) *is not* ε*-regular.*

Lemma 3.2 (Diregularity Lemma). *For every* $\varepsilon > 0$ *and every* $m \geq 1$ *there exists an integer* $M = M(m, \varepsilon)$*, such that every digraph of order at least* m *admits an* ε*-regular partition* $\{V_0, V_1, \ldots, V_k\}$*, where* $m \leq k \leq M$.

Let $D = (V, E)$ be a directed graph and let $d > 0$ be a parameter. Given an ε-regular partition $\{V_0, V_1, \ldots, V_k\}$ of V, we define the *regularity digraph* $R = R(D, d)$ to be the directed graph with vertex set $\{v_1, v_2, \ldots, v_k\}$ such that for every $1 \leq i \neq j \leq k$, (v_i, v_j) is an arc of R if and only if (V_i, V_j) is ε-regular with directed density at least d.

The following simple lemma is an immediate corollary of the definition of ε-regularity.

Lemma 3.3. *Let* (A, B) *be an* ε*-regular pair with directed density* d*. Let* $X \subseteq A$ *and* $Y \subseteq B$ *be sets of size* $|X| \geq \varepsilon|A|$ *and* $|Y| \geq \varepsilon|B|$*. Then the following properties hold*

1. $deg^+(x, Y) \leq (1 + \varepsilon)d|Y|$ *for all but at most* $\varepsilon|A|$ *vertices* $x \in A$*;*
2. $deg^+(x, Y) \geq (1 - \varepsilon)d|Y|$ *for all but at most* $\varepsilon|A|$ *vertices* $x \in A$*;*
3. $deg^-(y, X) \leq (1 + \varepsilon)d|X|$ *for all but at most* $\varepsilon|B|$ *vertices* $y \in B$*;*

4. *$\deg^-(y, X) \geq (1-\varepsilon)d|X|$ for all but at most $\varepsilon|B|$ vertices $y \in B$.*

We omit the straightforward proof.

Definition 3.4. *Let $D = (V, E)$ be a digraph and let A, B and X be subsets of V such that $A \cap B = \emptyset$. Assume that the pair (A, B) is ε-regular with directed density d. A vertex $u \in A$ is said to be* nice with respect to X and B *if $\deg_D^+(u, B \setminus X) \geq (1-\varepsilon)d|B \setminus X|$.*

The purpose of the set X in Definition 3.4 and in the next two lemmas is to make them more flexible. This will be useful in the proof of our main result where we will consider nice vertices with respect to a set X which will constantly change.

Lemma 3.5. *Let $D = (V, E)$ be a digraph and let A and B be disjoint subsets of V of size ℓ such that the pair (A, B) is ε-regular with directed density d. Let X be a subset of V such that $|B \setminus X| \geq \varepsilon\ell$. Let $w \in V \setminus (A \cup B)$ be a vertex such that $\deg_D^+(w, A \setminus X) > \varepsilon\ell$. Then there exists a vertex $y \in N_D^+(w, A \setminus X)$ which is nice with respect to X and B.*

Proof. Let $S = \{y \in A : \deg_D^+(y, B \setminus X) < (1-\varepsilon)d|B \setminus X|\}$. Since the pair (A, B) is ε-regular with directed density d and since $|B \setminus X| \geq \varepsilon\ell$, it follows by Lemma 3.3 that $|S| \leq \varepsilon|A|$. Since, moreover, $\deg_D^+(w, A \setminus X) > \varepsilon\ell$, it follows that $N_D^+(w, A \setminus X) \setminus S \neq \emptyset$. This completes the proof of the lemma as, by the definition of S, any vertex of $N_D^+(w, A \setminus X) \setminus S$ is nice with respect to X and B. ∎

Lemma 3.6. *Let ε and d be positive real numbers such that $d \geq 100\varepsilon$. Let ℓ be a positive integer and let $D = (V, E)$ be a digraph. Let V_1, V_2, V_3, V_4 and V_5 be pairwise disjoint subsets of V of size ℓ such that, for every $1 \leq i \leq 4$, the pair (V_i, V_{i+1}) is ε-regular with directed density at least d. Let X be a subset of V such that $|V_i \setminus X| \geq \ell/2$ holds for every $2 \leq i \leq 5$. Let $Y \subseteq V_3 \setminus X$ and $Z \subseteq V_4 \setminus X$ be sets of size at least $d\ell$. Let $v_1 \in V_1$ be a nice vertex with respect to X and V_2. Assume we choose uniformly and independently at random vertices v_2, v_3 and v_4 such that $v_2 \in N_D^+(v_1, V_2 \setminus X)$ is nice with respect to X and V_3, $v_3 \in N_D^+(v_2, V_3 \setminus X)$ is nice with respect to X and V_4, and $v_4 \in N_D^+(v_3, V_4 \setminus X)$ is nice with respect to X and V_5. Then $Pr(v_3 \in Y$ and $v_4 \in Z) \geq d^5/20$.*

Proof. Let $N_4 = \{z \in Z : z$ is nice with respect to X and $V_5\}$. Since the pair (V_4, V_5) is ε-regular with directed density at least d and $|V_5 \setminus X| \geq \ell/2 \geq \varepsilon\ell$, it follows by Lemma 3.3 that $|N_4| \geq |Z| - \varepsilon\ell \geq (d-\varepsilon)\ell$.

Let $N_3 = \{y \in Y : y$ is nice with respect to X and $V_4\}$ and let $N_3' = \{x \in N_3 : \deg_D^+(x, N_4) \geq (1-\varepsilon)d|N_4|\}$. Since the pair (V_3, V_4) is ε-regular with directed density at least d and $|V_4 \setminus X| \geq \ell/2 \geq \varepsilon\ell$, it follows by Lemma 3.3 that $|N_3| \geq |Y| - \varepsilon\ell \geq (d-\varepsilon)\ell$. Since, moreover, the pair (V_3, V_4) is ε-regular with directed density at least d and $|N_4| \geq (d-\varepsilon)\ell \geq \varepsilon\ell$, it follows by Lemma 3.3 that $|N_3'| \geq |N_3| - \varepsilon\ell \geq (d-2\varepsilon)\ell$.

Let $N_2 = \{y \in N_D^+(v_1, V_2 \setminus X) : y$ is nice with respect to X and $V_3\}$ and let $N_2' = \{x \in N_2 : \deg_D^+(x, N_3') \geq (1-\varepsilon)d|N_3'|\}$. Since the pair (V_1, V_2) is ε-regular with directed density at least d, $|V_2 \setminus X| \geq \ell/2$ and v_1 is nice with respect to X and V_2, it follows that $|N_D^+(v_1, V_2 \setminus X)| \geq (1-\varepsilon)d|V_2 \setminus X| \geq d\ell/3$. Since, moreover, the pair (V_2, V_3) is ε-regular with directed density at least d and $|V_3 \setminus X| \geq \ell/2 \geq \varepsilon\ell$, it follows by Lemma 3.3 that $|N_2| \geq |N_D^+(v_1, V_2 \setminus X)| - \varepsilon\ell \geq d\ell/4$. Finally, since the pair (V_2, V_3) is ε-regular with directed density at least d and $|N_3'| \geq (d-2\varepsilon)\ell \geq \varepsilon\ell$, it follows by Lemma 3.3 that $|N_2'| \geq |N_2| - \varepsilon\ell \geq d\ell/5$.

Note that

$$Pr(v_3 \in N_3') \geq Pr(v_3 \in N_3' \text{ and } v_2 \in N_2') = Pr(v_2 \in N_2') \cdot Pr(v_3 \in N_3' \mid v_2 \in N_2')$$
$$\geq \frac{|N_2'|}{|N_2|} \cdot \frac{\deg_D^+(v_2, N_3')}{\deg_D^+(v_2, V_3 \setminus X)} \geq \frac{|N_2'|}{|V_2|} \cdot \frac{(1-\varepsilon)d|N_3'|}{|V_3|} \geq \frac{d\ell/5}{\ell} \cdot \frac{d^2\ell/2}{\ell} = d^3/10\,.$$

Therefore

$$Pr(v_3 \in Y \text{ and } v_4 \in Z) \geq Pr(v_3 \in N_3' \text{ and } v_4 \in N_4) = Pr(v_3 \in N_3') \cdot Pr(v_4 \in N_4 \mid v_3 \in N_3')$$
$$\geq \frac{d^3}{10} \cdot \frac{\deg_D^+(v_3, N_4)}{\deg_D^+(v_3, V_4 \setminus X)} \geq \frac{d^3}{10} \cdot \frac{(1-\varepsilon)d|N_4|}{|V_4|}$$
$$\geq \frac{d^3}{10} \cdot \frac{d^2\ell/2}{\ell} = d^5/20$$

as claimed. ∎

Lemma 3.7. *Let n be a positive integer and let $\log n/n \ll p = p(n) \leq 1$. Let $\varepsilon > 0$ be arbitrarily small, let $c > 0$ be a constant, and let $D = (V, E) \in \mathcal{D}(n,p)$. For a set $Y \subseteq V$, let B_Y denote the set of all vertices $u \in V \setminus Y$ for which $|deg_D^+(u, Y) - |Y|p| \geq \varepsilon|Y|p$ or $|deg_D^-(u, Y) - |Y|p| \geq \varepsilon|Y|p$. Let $b = \max\{|B_Y| : Y \subseteq V, \ |Y| \geq cn\}$, then a.a.s. $b \leq p^{-1}\log n$.*

Proof. Let $u \in V$ be any vertex and let $Y \subseteq V \setminus \{u\}$ be any set of size at least cn. Note that $\deg_D^+(u, Y) \sim Bin(|Y|, p)$ and similarly $\deg_D^-(u, Y) \sim Bin(|Y|, p)$. Thus $\mathbb{E}(\deg_D^+(u, Y)) = \mathbb{E}(\deg_D^-(u, Y)) = |Y|p$. It follows by Theorem 1.2 (iii) that $Pr\left(|\deg_D^+(u, Y) - |Y|p| \geq \varepsilon|Y|p\right)$ is bounded from above by $2\exp\left\{-\frac{\varepsilon^2}{3} \cdot |Y|p\right\} \leq e^{-\varepsilon^2 c' np}$, where $c' > 0$ is an appropriate constant, and the same holds for $Pr\left(|\deg_D^-(u, Y) - |Y|p| \geq \varepsilon|Y|p\right)$ as well. Hence,

$$Pr(b \geq p^{-1}\log n) = Pr(\exists Y \subseteq V \text{ of size at least } cn \text{ such that } |B_Y| \geq p^{-1}\log n)$$
$$\leq 2^n \binom{n}{p^{-1}\log n} \left(2e^{-\varepsilon^2 c' np}\right)^{p^{-1}\log n} = o(1)\,.$$ ∎

4. PROOF OF THE MAIN RESULT

We start with the upper bound in Theorem 1.1; we will in fact prove the following stronger result. Let $D \in \mathcal{D}(n,p)$, where $p \gg \log n/n$. Then a.a.s. one can delete at most $\left(1/2 + 10\sqrt{\frac{\log n}{np}}\right)\deg_D^+(u)$ of the out-going arcs and at most $\left(1/2 + 10\sqrt{\frac{\log n}{np}}\right)\deg_D^-(u)$ of the in-going arcs at every vertex $u \in V(D)$ so that the resulting digraph is non-Hamiltonian; note that $10\sqrt{\frac{\log n}{np}} = o(1)$ for $p \gg \log n/n$.

Proof of Theorem 1.1 (upper bound). Let $V = V_1 \cup V_2$ be an arbitrary partition of $[n]$ into two parts of equal size (that is, $||V_1| - |V_2|| \leq 1$). Let $D = ([n], E) \in \mathcal{D}(n,p)$. It follows by Theorem 1.2 (iii) and union bound, that a.a.s. every $v \in [n]$ satisfies $|\deg_D^+(v) - np| \leq 4\sqrt{np\log n}$ and $|\deg_D^-(v) - np| \leq 4\sqrt{np\log n}$. Let $u \in V_1$ be an arbitrary vertex, then clearly

$\deg_D^+(u, V_2) \sim Bin(|V_2|, p)$. In particular, $(n-1)p/2 \leq \mathbb{E}(\deg_D^+(u, V_2)) \leq (n+1)p/2$. Since a.a.s. $\deg_D^+(u) \geq np - 4\sqrt{np\log n}$, it follows by Theorem 1.2 (ii) that

$$\begin{aligned}
&Pr\left(\deg_D^+(u, V_2) \geq \left(1/2 + 10\sqrt{\frac{\log n}{np}}\right)\deg_D^+(u)\right)\\
&\leq Pr\left(\deg_D^+(u, V_2) \geq \left(1 + 20\sqrt{\frac{\log n}{np}}\right)(np/2 - 2\sqrt{np\log n})\right)\\
&\leq Pr\left(\deg_D^+(u, V_2) \geq \left(1 + 15\sqrt{\frac{\log n}{np}}\right)\mathbb{E}(\deg^+(u, V_2))\right)\\
&\leq e^{-\frac{225\log n}{3np}\cdot\frac{(n-1)p}{2}} = o(1/n)\,.
\end{aligned}$$

Taking the union bound over all vertices of V_1, we conclude that a.a.s. for every $u \in V_1$ we have $\deg_D^+(u, V_2) \leq \left(1/2 + 10\sqrt{\frac{\log n}{np}}\right)\deg_D^+(u)$. An analogous argument shows that a.a.s. $\deg_D^-(w, V_1) \leq \left(1/2 + 10\sqrt{\frac{\log n}{np}}\right)\deg_D^-(w)$ for every $w \in V_2$. Our claim now follows since one can obtain a non-Hamiltonian digraph by deleting all arcs of D that are oriented from V_1 to V_2. In particular, $r_\ell(\mathcal{D}(n,p), \mathcal{H}) \leq \left(1/2 + 10\sqrt{\frac{\log n}{np}}\right)np$ a.a.s. ∎

The remainder of this section is devoted to the proof of the lower bound. Namely, we will prove that a.a.s. even if an adversary deletes at most $(1/2 - \alpha)\deg_D^+(u)$ of the out-going arcs and at most $(1/2 - \alpha)\deg_D^-(u)$ of the in-going arcs at every vertex $u \in V(D)$, where $\alpha > 0$ is an arbitrarily small constant, there is still a Hamilton cycle in the resulting digraph.

Let $D \in \mathcal{D}(n,p)$. Note that a.a.s. $|\deg_D^+(u) - np| \leq 4\sqrt{np\log n}$ and $|\deg_D^-(u) - np| \leq 4\sqrt{np\log n}$ hold for every vertex $u \in V(D)$. Hence, we will assume throughout the proof that D satisfies these properties. Let $D' = (V, E)$ be a digraph obtained from D by deleting at most $(1/2 - \alpha)\deg_D^+(u)$ of the out-going arcs and at most $(1/2 - \alpha)\deg_D^-(u)$ of the in-going arcs at every vertex $u \in V(D)$.

Let ε and ξ be positive real numbers and let m be a positive integer such that $m^{-1} \ll \varepsilon \ll \xi \ll \alpha$ and $\varepsilon \ll \alpha^4\xi^{12}p^{14}$, where for positive real numbers a and b the notation $a \ll b$ means that a/b is a sufficiently small real number.

Apply Lemma 3.2 to $D' = (V, E)$ with parameters ε and m. Let $\{V_0, V_1, \ldots, V_k\}$ be the corresponding ε-regular partition, and let $R = R(D', d)$ be the corresponding regularity digraph, where $d = \xi p$. It follows by the definition of R that the ordered pair (V_i, V_j) is ε-regular with directed density at least d whenever $(v_i, v_j) \in E(R)$. Let ℓ denote the common size of $V_1, \ldots, V_k$; note that $(1 - \varepsilon)n/k \leq \ell \leq n/k$.

We first show that R contains an almost spanning cycle; our proof will use the following immediate corollary of a classical theorem of Ghouila-Houri [19].

Theorem 4.1. *Let D be a digraph on n vertices. If $\delta^+(D) \geq n/2$ and $\delta^-(D) \geq n/2$, then D admits a directed Hamilton cycle.*

Lemma 4.2. *The regularity digraph R contains a directed cycle of length $r \geq (1-2\sqrt{\varepsilon})k$.*

Proof. Let $1 \leq i \leq k$ be an index for which there are at most $\sqrt{\varepsilon}k$ indices $1 \leq j \neq i \leq k$ such that (V_i, V_j) is not ε-regular (recall that by Remark 3.1, at least $(1 - \sqrt{\varepsilon})k$ indices

$1 \le i \le k$ have this property). Let $1 \le j \ne i \le k$ be such that (V_i, V_j) is an ε-regular pair but $(v_i, v_j) \notin E(R)$. Since $n \ll \ell^2 p$, we can use Theorem 1.2 (i) and union bound to show that a.a.s. $e_D(V_i, V_j) \ge (1-\alpha/5)\ell^2 p$. Since $(v_i, v_j) \notin E(R)$ even though (V_i, V_j) is ε-regular, it must hold that $d_{D'}(V_i, V_j) < d$. Hence, recalling that $d = \xi p$, we conclude that at least $(1 - \xi - \alpha/5)\ell^2 p$ arcs of $E_D(V_i, V_j)$ were deleted from D in order to obtain D'. Recall that $\xi, \sqrt{\varepsilon} \ll \alpha$. If $\deg_R^+(v_i) < k/2 + 2\sqrt{\varepsilon}k$, then at least

$$\begin{aligned}\left(k - 1 - \sqrt{\varepsilon}k - (k/2 + 2\sqrt{\varepsilon}k)\right)(1 - \xi - \alpha/5)\,\ell^2 p &\ge \left(1/2 - 4\sqrt{\varepsilon} - \xi - \alpha/5\right) k\ell^2 p \\ &> (1/2 - \alpha/3)\,(1 - \varepsilon) n\ell p \\ &> (1/2 - \alpha/2)\, n\ell p\end{aligned}$$

arcs of $E_D(V_i, [n] \setminus V_i)$ were deleted from D to obtain D'. Since a.a.s. the maximum out-degree of D is at most $np + 4\sqrt{np \log n}$, it follows that there exists some vertex $u \in V_i$ such that strictly more than $(1/2 - \alpha)\deg_D^+(u)$ out-going arcs which are incident with u in D were deleted to obtain D', contrary to our assumption. Therefore $\deg_R^+(v_i) \ge k/2 + 2\sqrt{\varepsilon}k$. Since the same argument applies to every $1 \le i \le k$ for which there are at most $\sqrt{\varepsilon}k$ indices $1 \le j \ne i \le k$ such that (V_i, V_j) is not ε-regular, it follows by Remark 3.1 that at least $(1 - \sqrt{\varepsilon})k$ vertices of R have out-degree at least $k/2 + 2\sqrt{\varepsilon}k$ each. An analogous argument shows that at least $(1 - \sqrt{\varepsilon})k$ vertices of R have in-degree at least $k/2 + 2\sqrt{\varepsilon}k$ each.

Let R' be the graph obtained from R by successively deleting vertices whose out-degree or in-degree is strictly smaller than $k/2$. It follows by the previous paragraph that there are at least $(1 - 2\sqrt{\varepsilon})k$ vertices, each with in-degree and out-degree at least $k/2 + 2\sqrt{\varepsilon}k$. It is evident that none of these vertices will be deleted and thus $(1 - 2\sqrt{\varepsilon})k \le |V(R')| \le k$. Moreover, $\min\{\delta^+(R'), \delta^-(R')\} \ge k/2 \ge |V(R')|/2$ holds by the definition of R'. Applying Theorem 4.1 to R' completes the proof of the lemma. ∎

Assume without loss of generality that $C_R : v_1, v_2, \dots, v_r, v_1$ is a cycle of R of length $r \ge (1 - 2\sqrt{\varepsilon})k$. Note that it follows from the definition of R that the pair $(V_i, V_{(i \mod r)+1})$ is ε-regular with directed density at least d, for every $1 \le i \le r$. For the sake of simplicity of presentation we will discard the "mod r" in the rest of the proof. Hence V_{i+1} will mean V_1 in case $i = r$ and V_{i-1} will mean V_r in case $i = 1$.

We now show how one can find a Hamilton cycle of D'. This is done in four stages. In the first stage we build a path P_1 of D' that includes certain "problematic" vertices. In the second stage we extend the path that was built in the first stage to an "almost spanning" path P_2 such that, for every $u \in V \setminus P_2$, there are "many" arcs $(x, y) \in E(P_2)$ for which $(x, u) \in E(D')$ and $(u, y) \in E(D')$. In the third stage we close the path that was built in the second stage into a cycle. Finally, in the fourth stage we extend this cycle to a Hamilton cycle by adding all remaining vertices.

4.1. Stage 1: Absorbing Problematic Vertices into a Short Path

Let B denote the set of vertices $u \in V$ for which there exists an index $1 \le i \le r$ such that $|\deg_D^+(u, V_i) - \ell p| \ge \varepsilon \ell p$ or $|\deg_D^-(u, V_i) - \ell p| \ge \varepsilon \ell p$. It follows by Lemma 3.7 that a.a.s.

$$|B| \le rp^{-1} \log n\,. \tag{1}$$

Since the vertices of B have atypical degrees into some sets V_i, it might be hard to absorb them in the fourth stage into the cycle we will build in the first three stages. Therefore, in Stage 1 we will build a directed path P_1 of $D' = (V, E)$ that includes all the vertices of B (by abuse of notation, P_1 will denote the path we build at any point during Stage 1; moreover, we will use P_1 to denote the path as well as its vertex set).

Let $v_0 \in V_1$ be an arbitrary vertex such that $\deg^+_{D'}(v_0, V_2) \geq (1-\varepsilon)d\ell$ and $\deg^-_{D'}(v_0, V_r) \geq 2\varepsilon\ell$. Let $A_0 \subseteq N^-_{D'}(v_0, V_r)$ be an arbitrary set of size $2\varepsilon\ell$; in Stage 3 we will use a subset of A_0 to close the path we are building into a cycle. Since the pairs (V_r, V_1) and (V_1, V_2) are ε-regular with directed density at least $d \gg \varepsilon$, such a vertex v_0 exists by Lemma 3.3. Initially, $P_1 = \{v_0\}$. As soon as $B \subseteq P_1$, Stage 1 is over and we proceed to Stage 2. For as long as $B \setminus P_1 \neq \emptyset$, we repeat the following procedure. Let $v \in B \setminus P_1$ be an arbitrary vertex. Let $1 \leq j_1, j_2 \leq r$ be indices such that $\deg^-_{D'}(v, V_{j_1}) \geq \ell p/3$ and $\deg^+_{D'}(v, V_{j_2}) \geq \ell p/3$. Let x be the last vertex added to P_1 and let $1 \leq s \leq r$ be such that $x \in V_s$. If $s \neq j_1 - 2$, then we extend P_1 by an arc (x, y), where $y \in N^+_{D'}(V_{s+1} \setminus P_1)$ is a nice vertex with respect to P_1 and V_{s+2}, and repeat this process. Otherwise, we extend P_1 by 4 arcs (x, y_1), (y_1, y_2), (y_2, v) and (v, z), where $y_1 \in V_{j_1-1}, y_2 \in V_{j_1}$ and $z \in V_{j_2}$ is a nice vertex with respect to P_1 and V_{j_2+1}.

It remains to prove that the above procedure is indeed feasible. We begin by justifying the existence of the indices j_1 and j_2.

Claim 4.3. *For every $v \in V$ there are indices $1 \leq j_1, j_2 \leq r$ such that $deg^-_{D'}(v, V_{j_1}) \geq \ell p/3$ and $deg^+_{D'}(v, V_{j_2}) \geq \ell p/3$.*

Proof. We will prove the existence of j_1; the existence of j_2 can be proved by an analogous argument. Since $(1 - o(1))np \leq \deg^-_D(v) \leq (1 + o(1))np$, it follows by the definition of D' that

$$\deg^-_{D'}(v) \geq (1/2 - o(1))np \,. \tag{2}$$

Suppose for a contradiction that $\deg^-_{D'}(v, V_i) < \ell p/3$ for every $1 \leq i \leq r$. Since, moreover, $r \geq (1 - 2\sqrt{\varepsilon})k$, it follows that

$$\deg^-_{D'}(v) < |V_0| + |V \setminus (V_0 \cup V_1 \cup \ldots \cup V_r)| + \sum_{i=1}^{r} \deg^-_{D'}(v, V_i) \leq \varepsilon n + 2\sqrt{\varepsilon} k\ell + r\ell p/3$$
$$\leq \varepsilon n + 2\sqrt{\varepsilon} n + np/3 < 2np/5 \,,$$

which contradicts (2). ∎

Our next goal is to prove that the arcs we wish to add to P_1 at each step exist. In order to do so, we will want $|V_i \cap P_1|$ to be small for every $1 \leq i \leq r$. This is an immediate corollary of Lemma 3.7 and the following claim.

Claim 4.4. *For every $1 \leq i \leq r$ and every $v \in B$, absorbing v in P_1 via the aforementioned procedure enlarges $|V_i \cap P_1|$ by at most 5.*

Proof. Fix some $1 \leq i \leq r$ and assume we aim to absorb $v \in B$ in P_1. Starting with $x \in V_s$ we walk along the cycle of clusters until we reach V_{j_1-2}, thus adding at most one vertex of V_i to P_1. The result now follows since clearly $|V_i \cap \{y_1, y_2, v, z\}| \leq 4$. ∎

We can now prove that the required arcs exist. Let v, j_1, j_2, x and s be as above. If $s \neq j_1 - 2$, then, since x is nice with respect to P_1 and V_{s+1} and since $|V_{s+1} \cap P_1|, |V_{s+2} \cap P_1| \leq 5rp^{-1} \log n$ by (1) and by Claim 4.4, the arc (x, y) exists by Lemma 3.5. Assume now that $s = j_1 - 2$. Since x is nice with respect to P_1 and V_{j_1-1} and since $\deg^-_{D'}(v, V_{j_1}) \geq \ell p/3 \geq \varepsilon\ell$ holds by the definition of j_1, there are vertices $y_1 \in N^+_{D'}(x, V_{j_1-1} \setminus P_1)$ and $y_2 \in N^-_{D'}(v, V_{j_1} \setminus P_1)$ such that $(y_1, y_2) \in E(D')$. Moreover, since $\deg^+_{D'}(v, V_{j_2}) \geq \ell p/3 \geq \varepsilon\ell$ holds by the definition of j_2 and since $|V_{j_2+1} \cap P_1| \leq 5rp^{-1}\log n$, it follows by Lemma 3.5 that we can find a vertex $z \in N^+_{D'}(v, V_{j_2} \setminus P_1)$ which is nice with respect to P_1 and V_{j_2+1}. Finally, since z is nice with respect to P_1 and V_{j_2+1}, we can repeat this procedure.

4.2. Stage 2: Extending the Path to an Almost Spanning One

In this subsection we extend P_1 to an almost spanning path of D' which satisfies certain desirable properties. Let $A'_0 \subseteq A_0 \setminus P_1$ be an arbitrary set of size $\varepsilon\ell$; such a set exists since $|A_0| = 2\varepsilon\ell$ by definition and $|P_1| \leq 5rp^{-1}\log n \leq \varepsilon\ell$ by (1) and by Claim 4.4. In Stage 3 we will use A'_0 to close the path we are building into a cycle; we thus refrain from adding any of its vertices to the path during Stage 2. Throughout this stage we denote the current path by P_2 and let $L_2 = A'_0 \cup P_2$. Initially $P_2 = P_1$.

For as long as $|V_i \setminus L_2| > 3(1-\varepsilon)^{-1}d^{-1}\varepsilon\ell$ holds for every $1 \leq i \leq r$, we repeat the following process. Let x be the last vertex added to P_2 and let $1 \leq s \leq r$ be such that $x \in V_s$. Let

$$N = \{y \in N^+_{D'}(x, V_{s+1} \setminus L_2) : y \text{ is nice with respect to } L_2 \text{ and } V_{s+2}\}.$$

We extend P_2 by an arc (x, y), where $y \in N$ is chosen uniformly at random, independently from all previous choices. In order to prove that this process is feasible, it suffices to show that $N \neq \emptyset$ throughout Stage 2. This holds by Lemma 3.5 since x is nice with respect to L_2 and V_{s+1} and since $|V_{s+1} \setminus L_2|, |V_{s+2} \setminus L_2| > 3(1-\varepsilon)^{-1}d^{-1}\varepsilon\ell$ by assumption.

The remainder of this subsection is dedicated to the proof of the following lemma which will play a crucial role in Stage 4.

Lemma 4.5. *Asymptotically almost surely, at the end of Stage 2,* $|\{(x,y) \in E(P_2) : (x,u) \in E(D') \text{ and } (u,y) \in E(D')\}| \geq 10\sqrt{\varepsilon}d^{-1}n$ *holds for every* $u \in V \setminus P_2$.

In order to prove Lemma 4.5 we will need the following claim.

Claim 4.6. *For every* $u \in V \setminus P_2$ *there exists a set* $I_u \subseteq [r]$ *of size* $|I_u| \geq \alpha r/20$ *which satisfies the following two properties:*

1. $\deg^-_{D'}(u, V_{i+2}) \geq \alpha\ell p/2$ *and* $\deg^+_{D'}(u, V_{i+3}) \geq \alpha\ell p/2$ *for every* $i \in I_u$.
2. $(i-j) \mod r \geq 4$ *and* $(j-i) \mod r \geq 4$ *for every* $i \neq j \in I_u$.

Proof. Fix some $u \in V \setminus P_2$; note that $u \notin B$ holds by the description of Stage 1. Let I'_u denote the set of all indices $1 \leq i \leq r$ for which $\deg^-_{D'}(u, V_{i+2}) \geq \alpha\ell p/2$ and $\deg^+_{D'}(u, V_{i+3}) \geq \alpha\ell p/2$. We claim that $|I'_u| \geq \alpha r/4$. Indeed, assume for the sake of contradiction that $|I'_u| < \alpha r/4$. Fix some $i \in [r] \setminus I'_u$. Since $u \notin B$, it follows that $\deg^-_D(u, V_{i+2}) \geq (1-\varepsilon)\ell p$ and $\deg^+_D(u, V_{i+3}) \geq (1-\varepsilon)\ell p$. Therefore, by the definition of I'_u we must have $|(v,u) \in E(D) \setminus E(D') : v \in V_{i+2}| \geq (1-\varepsilon-\alpha/2)\ell p$ or $|(u,v) \in E(D) \setminus E(D') : v \in V_{i+3}| \geq (1-\varepsilon-\alpha/2)\ell p$. Hence, in order to obtain D' from D, one has to delete at least

$$(r - \alpha r/4)(1-\varepsilon-\alpha/2)\ell p \geq (1-\alpha)np > (1/2-\alpha)\deg^+_D(u) + (1/2-\alpha)\deg^-_D(u)\,.$$

arcs which are incident with u, contrary to our assumption.

Now, let I_u be obtained from I'_u by keeping every fourth element, that is, if $I'_u = \{i_1, i_2, \ldots, i_{|I'_u|}\}$, where $i_1 < i_2 < \ldots < i_{|I'_u|}$, then $I_u = \{i_j : j \equiv 0 \mod 4\}$. It is evident that $|I_u| \geq \alpha r/20$ and that I_u satisfies Properties (1) and (2) of the claim. ∎

Proof of Lemma 4.5. Fix some $u \in V \setminus P_2$, let $I_u \subseteq [r]$ be as in Claim 4.6 and let $i \in I_u$ be an arbitrary index. Consider the path P_2 at the moment $|(P_2 \setminus P_1) \cap V_s| \geq \alpha \ell p/4$ first occurs for some $1 \leq s \leq r$. Recall that $|A'_0| = \varepsilon \ell$ and note that $|V_{i+2} \cap P_1| \leq 5rp^{-1} \log n$ holds by (1) and Claim 4.4. It thus follows by Claim 4.6 that

$$\begin{aligned} \deg^-_{D'}(u, V_{i+2} \setminus L_2) &\geq \deg^-_{D'}(u, V_{i+2}) - |V_{i+2} \cap L_2| \\ &\geq \alpha \ell p/2 - |V_{i+2} \cap P_1| - |V_{i+2} \cap (P_2 \setminus P_1)| - |A'_0| \\ &\geq \alpha \ell p/2 - \alpha \ell p/4 - 5rp^{-1} \log n - \varepsilon \ell \geq \alpha \ell p/5 \end{aligned} \tag{3}$$

holds until this moment.

An analogous argument shows that

$$\deg^+_{D'}(u, V_{i+3} \setminus L_2) \geq \alpha \ell p/5 \tag{4}$$

holds as well.

At the moment $|(P_2 \setminus P_1) \cap V_s| \geq \alpha \ell p/4$ first occurs for some $1 \leq s \leq r$, let $E_u = |\{(x,y) \in E(P_2 \setminus P_1) : \exists i \in I_u \text{ such that } x \in N^-_{D'}(u, V_{i+2}) \text{ and } y \in N^+_{D'}(u, V_{i+3})\}|$, that is, E_u is a random variable which counts some of the arcs of P_2 that can absorb u. Using a union bound argument, in order to complete the proof of the lemma, it suffices to prove that $Pr(E_u < 10\sqrt{\varepsilon} d^{-1} n) = o(1/n)$.

Let (v_1, v_2), (v_2, v_3) and (v_3, v_4) be three consecutive arcs of $P_2 \setminus P_1$, where $v_m \in V_{i+m-1}$ for every $1 \leq m \leq 4$. Assume that $|(P_2 \setminus P_1) \cap V_j| \leq \alpha \ell p/4$ was still true for every $1 \leq j \leq r$ immediately after the arc (v_3, v_4) was added to P_2. Since $|A'_0| = \varepsilon \ell$ and $|V_j \cap P_1| \leq 5rp^{-1} \log n$ holds for every $1 \leq j \leq r$ by (1) and Claim 4.4, it follows that $|V_{i+m} \setminus L_2| \geq \ell/2$ for every $1 \leq m \leq 4$. Since, moreover, $\deg^-_{D'}(u, V_{i+2} \setminus L_2), \deg^+_{D'}(u, V_{i+3} \setminus L_2) \geq \alpha \ell p/5 \gg d\ell$ hold by (3) and (4), it follows by Lemma 3.6 (with $X = L_2$, $Y = N^-_{D'}(u, V_{i+2} \setminus L_2)$ and $Z = N^+_{D'}(u, V_{i+3} \setminus L_2)$) that

$$Pr(v_3 \in N^-_{D'}(u, V_{i+2}) \text{ and } v_4 \in N^+_{D'}(u, V_{i+3})) \geq d^5/20\,. \tag{5}$$

Let $F^j_u : j \in I_u$ be independent random variables, where $F^j_u \sim Bin(\alpha \ell p/5, d^5/20)$ for every $j \in I_u$. Let $F_u = \sum_{j \in I_u} F^j_u$, then $F_u \sim Bin(|I_u| \cdot \alpha \ell p/5, d^5/20)$. We claim that E_u dominates F_u, that is, that $Pr(E_u < K) \leq Pr(F_u < K)$ for every K. Indeed, note that the inequality (5) holds regardless of the choice of v_1 (as long as it is nice with respect to L_2 and V_{i+1}). Therefore, whenever we add to P_2 an arc (x,y), where $x \in V_{j+2}$ and $y \in V_{j+3}$ for some $j \in I_u$, we can imagine that a coin is tossed with the probability of *success*, that is, the probability that $x \in N^-_{D'}(u, V_{j+2})$ and $y \in N^+_{D'}(u, V_{j+3})$, being at least $d^5/20$. Moreover, for every $j \in I_u$, we consider all arcs $(x,y) \in E_{D'}(V_{j+2}, V_{j+3})$, added to P_2 during Stage 2 until $|(P_2 \setminus P_1) \cap V_s| \geq \alpha \ell p/4$ first occurred for some $1 \leq s \leq r$. Since $||(P_2 \setminus P_1) \cap V_a| - |(P_2 \setminus P_1) \cap V_b|| \leq 1$ holds for every $1 \leq a, b \leq r$ by the description of Stage 2, for every $j \in I_u$, we thus consider at least $\alpha \ell p/4 - 1 \geq \alpha \ell p/5$ arcs (x,y) such that $x \in V_{j+2}$ and $y \in V_{j+3}$, that is, there are at least $\alpha \ell p/5 \cdot |I_u| \geq \alpha^2 r \ell p/100 \geq \alpha^2 np/150$ trials. It thus follows by Theorem 1.2 (i) that

$$Pr(E_u < 10\sqrt{\varepsilon}d^{-1}n) \le Pr(F_u < 10\sqrt{\varepsilon}d^{-1}n) \le Pr(F_u < \alpha^2 npd^5/6000)$$
$$\le Pr(F_u \le \mathbb{E}(F_u)/2) \le \exp\left\{-\frac{1}{8}\cdot\frac{\alpha^2 np}{150}\cdot\frac{d^5}{20}\right\} = o(1/n)$$

as claimed. ∎

4.3. Stages 3 and 4: Closing the Path into a Cycle and Absorbing All Remaining Vertices

This subsection consists of two simple parts, namely Stage 3 and Stage 4. In Stage 3 we will close the path P_2 which was built in Stage 2 into a cycle C. In Stage 4 we will use Lemma 4.5 to absorb all of the remaining vertices into C, thus creating a Hamilton cycle.

Stage 3: In this stage we close P_2 into a directed cycle C, by adding a few more arcs to P_2 as follows. Throughout this stage, we denote the current path by P_3; initially $P_3 = P_2$. Moreover, we set $L_3 = P_3 \cup A'_0$. Let x denote the last vertex added to P_3 and let $1 \le s \le r$ be such that $x \in V_s$. If $s \ne r-2$, then we extend P_3 by an arc (x,y), where $y \in N^+_{D'}(x, V_{s+1} \setminus L_3)$ is a nice vertex with respect to L_3 and V_{s+2}, and repeat this process. Otherwise, we extend P_3 by 3 arcs (x,y), (y,z) and (z,v_0), where $y \in V_{r-1}$ and $z \in A'_0$, thus producing the required cycle C.

It is evident that, if it works, this process produces a cycle C of D'. It thus remains to prove that, at every step of the process, the arcs we wish to add to the current path exist. If $s \ne r-2$, then the arc (x,y) exists by Lemma 3.5 since x is nice with respect to L_3 and V_{s+1} and since $|V_{s+1} \setminus L_3|, |V_{s+2} \setminus L_3| \ge 3(1-\varepsilon)^{-1}d^{-1}\varepsilon\ell - 1$ by the description of Stages 2 and 3. If $s = r-2$, then $\deg^+_{D'}(x, V_{r-1} \setminus L_3) \ge (1-\varepsilon)d|V_{r-1} \setminus L_3| \ge \varepsilon\ell$ since x is nice with respect to L_3 and V_{r-1} and since $|V_{r-1} \setminus L_3| \ge 3(1-\varepsilon)^{-1}d^{-1}\varepsilon\ell - 1$ by the description of Stages 2 and 3. Since, moreover, $|A'_0| \ge \varepsilon\ell$ by definition and the pair (V_{r-1}, V_r) is ε-regular, we can find vertices $y \in N^+_{D'}(x, V_{r-1} \setminus L_3)$ and $z \in A'_0$ such that $(y,z) \in E(D')$. Finally, note that $(z, v_0) \in E(D')$ by the definition of A'_0.

Stage 4: In this final stage, we extend the directed cycle C we built in Stage 3 to a Hamilton cycle of D' by absorbing all remaining vertices. Let $v_1, \dots, v_t$ be an arbitrary ordering of the vertices of $V \setminus V(C)$. Let $\{(x_i, y_i) : 1 \le i \le t\}$ be a set of pairwise distinct arcs of C such that $(x_i, v_i) \in E(D')$ and $(v_i, y_i) \in E(D')$ hold for every $1 \le i \le t$. By replacing (x_i, y_i) with (x_i, v_i) and (v_i, y_i) for every $1 \le i \le t$ we obtain a directed Hamilton cycle of D'.

In order to complete the proof, it remains to show that the set $\{(x_i, y_i) : 1 \le i \le t\}$ exists. By Lemma 4.5, it suffices to show that $t < 10\sqrt{\varepsilon}d^{-1}n$. We will in fact prove the slightly stronger inequality $|V \setminus P_2| < 10\sqrt{\varepsilon}d^{-1}n$. As noted above, it follows by the description of Stage 2 that $||V_i \cap (P_2 \setminus P_1)| - |V_j \cap (P_2 \setminus P_1)|| \le 1$ holds for every $1 \le i,j \le r$. Hence, at the end of Stage 2, we have $3(1-\varepsilon)^{-1}d^{-1}\varepsilon\ell \le |V_i \setminus L_2| \le 3(1-\varepsilon)^{-1}d^{-1}\varepsilon\ell + 1 + \max\{|V_j \cap P_1| : 1 \le j \le r\} \le 4(1-\varepsilon)^{-1}d^{-1}\varepsilon\ell$ for every $1 \le i \le r$, where the last inequality holds by Claim 4.4 and by (1). Therefore

$$\begin{aligned}|V \setminus P_2| &\le |V_0| + |V \setminus (V_0 \cup V_1 \cup \ldots \cup V_r)| + \sum_{i=1}^{r} |V_i \setminus P_2| \\ &\le \varepsilon n + 2\sqrt{\varepsilon}k\ell + \sum_{i=1}^{r} |V_i \setminus L_2| + |A'_0| \\ &\le \varepsilon n + 2\sqrt{\varepsilon}k\ell + 4(1-\varepsilon)^{-1}d^{-1}\varepsilon r\ell + \varepsilon\ell < 10\sqrt{\varepsilon}d^{-1}n\end{aligned}$$

as claimed.

5. ROBUST HAMILTONICITY OF SPARSE RANDOM DIGRAPHS

Our proof of the lower bound in Theorem 1.1 for $p = o(1)$ follows the same general lines described in Section 4, but is much more involved. Below is a brief account of some of the new difficulties which arise in this case and of our way to overcome them. Since many of the changes are highly technical, for the sake of simplicity our account will be rather informal and not the most accurate; the interested reader can find all the exact details in [21].

The first major change we face is the need for a version of the Directed Regularity Lemma which applies to sparse digraphs. Such a lemma exists (see [26]) but, since $p \ll \varepsilon$, using it is less straightforward. Hence, in order to be able to move along the cycle of clusters in Stages 1,2 and 3, we have to ensure regularity of very small sets. Indeed, a typical vertex $x \in V_s$ has at most $2\ell p \ll \varepsilon\ell$ out-neighbors in V_{s+1} and so it is possible that none of them are nice with respect to V_{s+2}. There are technical tools to overcome this problem (see, e.g. [16, 17]); unfortunately these tools lead to the introduction of more "problematic" vertices (in addition to those denoted by B).

Informally speaking, $x \in V_s$ is said to be *a bad vertex of type I* if the pair $(N^+_{D'}(x, V_{s+1}), V_{s+2})$ is not regular. Similarly, u is said to be *a bad vertex of type II* if there are many indices $1 \le i \le r$ for which the pair $(N^-_{D'}(u, V_i), N^+_{D'}(u, V_{i+1}))$ is not regular. Let B_1 denote the set of bad vertices of type I and let B_2 denote the set of bad vertices of type II. Using the aforementioned tools from [16], the fact that D' is a subdigraph of a random directed graph and the first moment method, one can prove that B_2 is small and that $B_1 \cap V_i$ is small for every $1 \le i \le r$.

When moving along the cycle of clusters, we do not want to ever reach a bad vertex of type I as we might not be able to continue. Therefore, we intentionally avoid these vertices throughout Stages 1,2 and 3. On the other hand, bad vertices of type II might be hard to absorb in Stage 4 (since if $(N^-_{D'}(u, V_i), N^+_{D'}(u, V_{i+1}))$ is not regular, we cannot find a good lower bound on the probability that an arc $(x, y) \in E_{D'}(V_i, V_{i+1})$ we add to P_2 is actually in $E_{D'}(N^-_{D'}(u, V_i), N^+_{D'}(u, V_{i+1})))$ and so we include them already in the path we build in Stage 1.

When building P_1 in Stage 1, we make sure it includes all vertices of $B \cap B_2$. This is a little more technical but similar to Stage 1 in the dense case. However, unlike the dense case, $|P_1 \cap V_i| \gg \ell p$ might hold at the end of Stage 1 for some $1 \le i \le r$. Therefore, at the end of Stage 1, there might be some vertex $z \in V \setminus P_1$ such that $N^-_{D'}(z, V) \subseteq P_1$ or $N^+_{D'}(z, V) \subseteq P_1$. Clearly, it will now be impossible to absorb z into the cycle we aim to build. In order to avert this problem, as soon as there exists a vertex $z \in V \setminus P_1$ for which $\deg^-_{D'}(z, P_1)$ or $\deg^+_{D'}(z, P_1)$ is too large, we declare z to be *dangerous* and make it our top priority to include z in P_1.

Since we use this more careful version of Stage 1, we do not need to make any major changes to the description of Stage 2. However, the assertion of Lemma 4.5 becomes weaker and its proof more complicated. In particular, when proving that the random variable E_u behaves similarly to the binomial random variable F_u, we will use the fact that D' is a subdigraph of a random directed graph. Moreover, when proving the existence of the sets I_u, we will also need to use the fact that $B_2 \subseteq P_1$.

Stage 3 is essentially the same. Since, as noted above, the assertion of Lemma 4.5 is weaker, we will no longer be able to greedily absorb all the remaining vertices in Stage 4. Instead, in order to prove the existence of the required matching between $V \setminus V(C)$ and $E(C)$, we will use Lemma 4.5, Hall's Theorem and the fact that D' is a subdigraph of a random directed graph.

6. CONCLUDING REMARKS AND OPEN PROBLEMS

We have proved that a.a.s. $(1/2-\alpha)np \le r_\ell(\mathcal{D}(n,p),\mathcal{H}) \le (1/2+\alpha)np$, where $\alpha > 0$ is an arbitrarily small constant, provided that $p \gg \log n/\sqrt{n}$ (the complete proof can be found in [21]). For *undirected* random graphs it was proved in [32] that $r_\ell(G(n,p),\mathcal{H}) = (1/2 + o(1))np$ holds a.a.s. for every $p \gg \log n/n$. This is essentially tight since for $p < \log n/n$ a.a.s. $G(n,p)$ contains no Hamilton cycle. Since it is also known (see [33] and [13]) that for $p = \Omega(\log n/n)$ directed random graphs are a.a.s. Hamiltonian, it is natural to ask the following question.

Question 6.1. *Is it true that for $p \gg \log n/n$, a.a.s. every subdigraph of $\mathcal{D}(n,p)$ with minimum out-degree and in-degree at least $(1/2 + o(1))np$ contains a directed Hamilton cycle?*

Recall that our proof of the upper bound in Theorem 1.1 does hold for every $p \gg \log n/n$. On the other hand, since our proof method for the lower bound relies on the existence of linearly many pairwise arc disjoint triangles in $\mathcal{D}(n,p)$, each sharing one arc with a given cycle (recall Stage 4), it cannot be used when $p = o(n^{-1/2})$, and hence some new ideas are required.

Note added in proof: Very recently Ferber et al. [12] came close to answering Question 6.1 in the affirmative. Using a different approach than the one appearing in this paper, they proved that every subdigraph of $\mathcal{D}(n,p)$ with minimum out-degree and in-degree at least $(1/2 + o(1))np$ contains a directed Hamilton cycle, provided that $p \gg \log^8 n/n$.

ACKNOWLEDGMENT

We would like to thank Asaf Ferber for many stimulating discussions.

REFERENCES

[1] S. Ben-Shimon, A. Ferber, D. Hefetz, and M. Krivelevich, Hitting time results for Maker-Breaker games, Random Struct Algorithms 41 (2012), 23–46.

[2] S. Ben-Shimon, M. Krivelevich, and B. Sudakov, Local resilience and Hamiltonicity Maker-Breaker games in random-regular graphs, Combin Probab Comput 20 (2011), 173–211.

[3] S. Ben-Shimon, M. Krivelevich, and B. Sudakov, On the resilience of Hamiltonicity and optimal packing of Hamilton cycles in random graphs, SIAM J Discrete Math 25 (2011), 1176–1193.

[4] B. Bollobás, Random graphs, Academic Press, London, 1985.

[5] B. Bollobás, The evolution of sparse graphs, Graph theory and combinatorics (Cambridge, 1983), Academic Press, London, 1984, pp. 35–57.

[6] C. Cooper and A. Frieze, On the number of Hamilton cycles in a random graph, J Graph Theory 13 (1989), 719–735.

[7] B. Cuckler and J. Kahn, Hamiltonian cycles in Dirac graphs, Combinatorica 29 (2009), 299–326.

[8] G. A. Dirac, Some theorems on abstract graphs, Proc London Math Soc 2 (1952), 69–81.

[9] P. Erdős and A. Rényi, On the evolution of random graphs, Bull Inst Stat Tokyo 38 (1961), 343–347.

[10] A. Ferber, R. Glebov, M. Krivelevich, and A. Naor, Biased games on random boards, 46 (2015), 651–676.

[11] A. Ferber, M. Krivelevich, and B. Sudakov, Counting and packing Hamilton cycles in dense graphs and oriented graphs, (in press).

[12] A. Ferber, R. Nenadov, A. Noever, U. Peter, and N. Škorić, Robust hamiltonicity of random directed graphs, Proceedings of the 26th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 15) (2015), 1752–1758.

[13] A. M. Frieze, An algorithm for finding Hamilton cycles in random directed graphs, J Algorithms 9 (1988), 181–204.

[14] A. Frieze and M. Krivelevich, On packing Hamilton cycles in ε-regular graphs, J Combin Theory Ser B 94 (2005), 159–172.

[15] A. Frieze and M. Krivelevich, On two Hamiltonian cycle problems in random graphs, Israel J Math 166 (2008), 221–234.

[16] S. Gerke, Y. Kohayakawa, V. Rödl, and A. Steger, Small subsets inherit sparse ε-regularity, J. Combin Theory Ser B 97 (2007), 34–56.

[17] S. Gerke and A. Steger, The sparse regularity lemma and its applications, In B. Webb, editor, Surveys in combinatorics, Cambridge University Press, Cambridge, 2005, pp. 227–258.

[18] R. Glebov and M. Krivelevich, On the number of Hamilton cycles in sparse random graphs, SIAM J Discrete Math 27 (2013), 27–42.

[19] A. Ghouila-Houri, Une condition suffisante d'existence d'un circuit hamiltonien, C R Acad Sci Paris 25 (1960), 495–497.

[20] D. Hefetz, M. Krivelevich, M. Stojaković, and T. Szabó, A sharp threshold for the Hamilton cycle Maker-Breaker game, Random Struct Algorithms 34 (2009), 112–122.

[21] D. Hefetz, A. Steger, and B. Sudakov, Random directed graphs are robustly Hamiltonian, arXiv:1404.4734v1.

[22] S. Janson, The numbers of spanning trees, Hamilton cycles and perfect matchings in a random graph, Combin Probab Comput 3 (1994), 97–126.

[23] S. Janson, T. Łuczak, and A. Ruciński, Random graphs, Wiley, New York, 2000.

[24] R. Karp, Reducibility among combinatorial problems, Complexity of computer computations, Plenum, New York, 1972, pp. 85–103.

[25] F. Knox, D. Kühn, and D. Osthus, Edge-disjoint Hamilton cycles in random graphs, 46 (2015), 397–445.

[26] Y. Kohayakawa, Szemerédi's regularity lemma for sparse graphs, Foundations of computational mathematics (Rio de Janeiro, 1997), Springer, Berlin, 1997, pp. 216–230.

[27] J. Komlós and E. Szemerédi, Limit distributions for the existence of Hamilton circuits in a random graph, Discrete Math 43 (1983), 55–63.

[28] A. D. Korshunov, Solution of a problem of Erdős and Rényi on Hamilton cycles in non-oriented graphs, Soviet Math Dokl 17 (1976), 760–764.

[29] M. Krivelevich, C. Lee, and B. Sudakov, Robust Hamiltonicity of Dirac graphs, Trans Am Math Soc 366 (2014), 3095–3130.

[30] M. Krivelevich and W. Samotij, Optimal packings of Hamilton cycles in sparse random graphs, SIAM J Discrete Math 26 (2012), 964–982.

[31] D. Kühn, J. Lapinskas, and D. Osthus, Optimal packings of Hamilton cycles in graphs of high minimum degree, Combin Probab Comput 22 (2013), 394–416.

[32] C. Lee and B. Sudakov, Dirac's theorem for random graphs, Random Struct Algorithms 41 (2012), 293–305.

[33] C. J. H. McDiarmid, Clutter percolation and random graphs, Math Program Stud 13 (1980), 17–25.

[34] L. Pósa, Hamiltonian circuits in random graphs, Discrete Math 14 (1976), 359–364.

[35] M. Stojaković and T. Szabó, Positional games on random graphs, Random Struct Algorithms 26 (2005), 204–223.

[36] B. Sudakov and V. Vu, Local resilience of graphs, Random Struct Algorithms 33 (2008), 409–433.

[37] E. Szemerédi, Regular partitions of graphs, In J. C. Bermond, J. C. Fournier, M. Las Vergnas, and D. Sotteau, editors, Proceedings of Colloques International CNRS, 1978, pp. 399–401.

[38] D. B. West, Introduction to graph theory, 2nd edition, Upper Saddle river, NJ 07458, USA, 2001.