Hamiltonicity in randomly perturbed graphs and hypergraphs

Michael Krivelevich, Matthew Kwan, Benny Sudakov

December 7, 2014

Abstract

We show that adding linearly many random hyperedges to a dense hypergraph ensures the (aymptotically almost sure) existence of a loose Hamilton cycle. This extends a corresponding theorem for graphs and digraphs due to Bohman, Frieze and Martin; we also give an alternative simple proof of their theorem. We then prove a similar theorem about Hamiltonicity of randomly perturbed tournaments.

Informally speaking, we say that a graph on n vertices is *dense* if its number of edges is comparable to n^2, or more strongly if the degree of each vertex is comparable to n. We say that a graph is *Hamiltonian* if it has a *Hamilton cycle*: a simple cycle containing every vertex in the graph. It is an early result (following from [KS75], say) that "almost all" dense graphs are Hamiltonian. This is usually expressed probabilistically: if we fix $\alpha > 0$ and choose a uniformly random (lablled) graph on n vertices with αn^2 edges, then the probability that our graph is Hamiltonian is $1 - o(1)$. (Here and from now on, asymptotics are as $n \to \infty$, and we implicitly round large quantities to integers).

It is not true that every sufficiently large dense graph is Hamiltonian, even using the strong notion of density where we require each vertex to have degree at least αn. For example, the complete bipartite graph with partition sizes $n/3$ and $2n/3$ has minimum degree $n/3$ but requires the addition of at least $n/3$ edges to become Hamiltonian. However, it was proved in [BFM03] that dense non-Hamiltonian graphs are relatively "fragile" in that adding linearly many edges will almost always make them Hamiltonian. To be precise, [BFM03, Theorem 1] says that for every $\alpha > 0$ there is $c(\alpha)$ such that if we add $c(\alpha)n$ random edges to any graph with minimum degree at least αn, then we a.a.s. end up with a Hamiltonian graph. By contrast, if we start from the empty graph then by [Pós76, Theorem 3] we need $\Omega(n \log n)$ random edges to ensure the existence a.a.s. of a Hamilton cycle.

The random graph model where we start with a fixed graph and add random edges is interesting in its own right, and has been further studied in a number of different situations (for example, [BFKM04] and [KST06]). We can also interpret the result of [BFM03, Theorem 1] in the framework of *smoothed analysis* of algorithms, introduced in [ST04], where one studies the performance of an algorithm under the assumption that the input is noisy. The result of [BFM03, Theorem 1] still holds when we allow the deletion as well as addition of random edges, and tells us that a "sufficiently noisy" dense graph is likely to be Hamiltonian.

More abstractly and non-probabilistically, the result of [BFM03, Theorem 1] gives a certain stronger notion of what it means for "almost all" dense graphs to be Hamiltonian. Not only are the non-Hamiltonian dense graphs few in number, they are topologically "sparse", in that every small (Hamming) neighbourhood in the space of dense graphs contains few non-Hamiltonian graphs. An analogy can be drawn to the measure-theoretic and Baire-category-theoretic notions of "almost all" in analysis.

We have a few different contributions to make. First, in Section 1 we give an alternative simpler proof of [BFM03, Theorem 1] and the corresponding theorem for digraphs ([BFM03, Theorem 3]). The proof is based on a key lemma which may be of independent interest, showing that Hamiltonicity follows from a certain pseudorandom property. We also give a slight extension to *pancyclicity*: that is, not only does the perturbed graph have a Hamilton cycle, it has cycles of all possible lengths.

In Section 2 we generalize to "loose" Hamilton cycles in randomly perturbed hypergraphs. The techniques used in [BFM03, Theorem 3] and Section 1 for graphs seem to be ineffective for higher-order hypergraphs, so we give a different proof involving a reduction to a lemma about perfect matchings in bipartite graphs. This lemma can also be used to prove a similar theorem about perfect matchings in randomly peturbed hypergraphs.

Finally, in Section 3 we study Hamiltonicity in randomly perturbed tournaments (a *tournament* is a complete graph with a single orientation assigned to each edge). A modest random perturbation is in fact able to ensure the existence of multiple edge-disjoint Hamilton cycles.

0.1 Different models of random perturbation

Most results concerning randomly perturbed graphs are stated for random graphs of the form $G \cup R$, where G is a fixed graph and R is a random graph. Usually R is distributed as $\mathbb{G}(n,m)$ (the uniform distribution on n-vertex, m-edge graphs), or $\mathbb{G}(n,p)$ (the Erdős- Rényi distribution where each edge is present with independent probability p). In practice, if $p \approx m/\binom{n}{2}$ then the choice between these two random graph models usually makes little difference. In our case Hamiltonicity is a so-called "monotone" property, so the equivalence between the two models is given by [JŁR00, Corollary 1.16].

As suggested in [ST03, Definition 1], we could also consider random graphs of the form $G \triangle R$, where we can add as well as delete random edges. For our purposes, this is mostly equivalent to the model where we just add edges. For example, if $R \in \mathbb{G}(n,p)$ then $G \triangle R$ has the same distribution as $(G \backslash R') \cup R$, where $R' \in \mathbb{G}(n,p')$ with p' satisfying $1-p'(1-p) = 1-p$. If $p = o(1)$ then $p' \sim p$ and the deletion of the edges in R' a.a.s. does not damage the denseness of G (this is because R' a.a.s. has sublinear maximum degree, as can be shown with a Chernoff bound and a union bound).

Yet another potential model is to start with a fixed graph G and "make it more random" by interpolating slightly towards $\mathbb{G}(n, 1/2)$. For example, we could randomly designate a linear number of pairs of vertices for "resampling" and then decide whether the corresponding edges should be present uniformly and independently at random. This is mostly equivalent to the symmetric difference model.

All these models can be equally well defined for digraphs and hypergraphs, and the considerations mentioned here hold equally true. Let $\mathbb{D}(n,m)$ be the uniform distribution on (not necessarily oriented) n-vertex, m-arc digraphs, and let $\mathbb{D}(n,p)$ be the distribution where each arc is present with independent probability p. Similarly, let $\mathbb{H}_k(n,m)$ be uniform on n-vertex, m-hyperedge k-uniform hypergraphs, and let $\mathbb{H}_k(n,p)$ have each 3-hyperedge present with probability p.

It does not make sense to add arcs to a tournament (we want the result of the random perturbation to remain a tournament), but the other two types of models mentioned here are suitable and more or less equivalent: we can flip the direction a linear number of random arcs, or we can designate a linear number of arcs to be resampled then choose their direction randomly.

1 Hamiltonicity and pancyclicity in dense graphs and digraphs

A graph or digraph on n vertices is *pancyclic* if it has a cycle of each length ranging from 3 to n. It is *vertex pancyclic* if every vertex is contained in a cycle of every such length. It was observed by Bondy [Bon75] that almost any non-trivial condition on a graph G that ensures Hamiltonicity also ensures pancyclicity. In accordance with Bondy's observation, we prove the following theorem. (Let $\delta(D)$ be the minimum of the in- and out- degrees among the vertices in D).

Theorem 1.1. *For each $\alpha > 0$, there is $c(\alpha)$ such that if D is a digraph with $\delta(D) \geq \alpha n$, and $R \in \mathbb{D}(n, c(\alpha)n)$, then $D \cup R$ is a.a.s. vertex pancyclic.*

This theorem immediately implies the corresponding theorems for undirected graphs, pancyclicity and Hamiltonicity. If we are only interested in one of these weaker cases, our proofs can be modified to give better values for $c(\alpha)$. We will mention several such possible improvements, but we make no attempt to fight for the sharpest possible quantitative bounds.

1.1 Proof of Theorem 1.1

The proofs in [BFM03] can give an exponentially small bound on the probability of failure to be Hamiltonian, so it is possible to use a union bound and some small adjustments to prove Theorem 1.1 as a corollary. However, we prefer to give an alternative proof, which we feel is more direct, and provides better quantitative bounds (for most values of α). The key to our proof is the following lemma, which gives a general pseudorandomness condition for pancyclicity.

Lemma 1.2. *Let D be a directed graph on n vertices with $\delta(D) \geq 8k$, and suppose for all disjoint $A, B \subseteq V(D)$ with $|A| = |B| \geq k$, there is an arc from A to B. Then D is vertex pancyclic.*

Lemma 1.2 will be a corollary of the following lemma.

Lemma 1.3. *Let D be a directed graph with $\delta(D) \geq 4k$, and suppose for all disjoint $A, B \subseteq V(D)$ with $|A| = |B| \geq k$, there is an arc from A to B. Then D is Hamiltonian.*

Remark 1.4. With some effort, the ideas in the proof of Lemma 1.3 can be used directly to prove Lemma 1.2 with a weaker condition on $\delta(D)$. We do not know whether the condition can be weakened all the way to $\delta(D) \geq 4k$, as Bondy's observation would suggest. The constants in Lemmas 1.2 and 1.3 can both be halved for the undirected case, just by simplifying the main argument in the proof of Lemma 1.3.

Proof of Lemma 1.3. The idea of the proof is to start with a longest path P and manipulate it into a cycle C on the same vertex set. We will show that D is strongly connected, so if C were not Hamiltonian, there would be an arc from $V(C)$ to its complement, which could be combined with C to give a longer path than P, contradicting maximality.

This type of argument goes back to a classical theorem by Dirac [Dir52, Theorem 3]. It also bears some resemblance to the "rotation-extension" idea introduced in [Pós76], and a variation for directed graphs in [FK05, Section 4.3].

First we acknowledge some immediate consequences of the condition on D. Note that if A and B are disjoint sets with size at least k, then in fact there are $|A| - k$ vertices of A with an arc into B. To see this, note that for any fewer number of such vertices in A, we can delete those vertices and at least k will remain, one of which has an arc to B. Also, D is strongly connected. To see this, note that for any v, w, both of $N^+(v)$ and $N^-(w)$ have size at least $4k > k$. If they intersect then there is a length-2 path from v to w; otherwise there must be an arc from $N^+(v)$ to $N^-(w)$ giving a length-3 path.

Let $P = u \dots w$ be a maximal-length directed path in D. We will use the notation v^+ (respectively v^-) for the successor (respectively predecessor) of a vertex v on P, and also write V^+, V^- for the set of successors or predecessors of a set of vertices V.

By maximality, $N^+(w) \subset P$ and $N^-(u) \subset P$. Let U_1 be the first $3k$ elements of $N^-(u)$ on P, and let U_2 be the last k (note $U_1 \cap U_2 = \varnothing$). Similarly let W_1 be the first k and W_2 the last $3k$ elements of $N^+(w)$. We will now show that there is a cycle on the vertex set $V(P)$.

First, consider the case where each vertex of W_1 precedes each vertex of U_2. If wu is in D then we can immediately close P into a cycle. Otherwise, $\left|W_1^-\right| = |W_1| = \left|U_2^+\right| = |U_2| = k$, so there is an arc w_1u_2 from W_1^- to U_2^+. This is enough to piece together a cycle on $V(P)$: start at u_2 and move along P to w, from where there is a shortcut back to w_1^+. Now move along P from w_1^+ to u_2^-, from where we can jump back to u, then move along P to w_1, then jump to u_2. See Figure Figure 1 for an illustration.

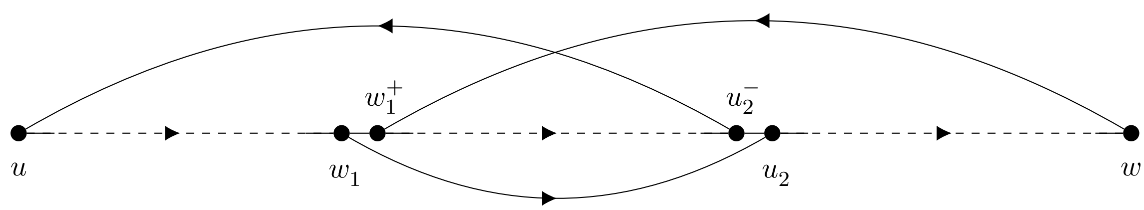

Figure 1. The case where the vertices of W_1 precede the vertices of U_2. The horizontal line through the center is P; the broken lines indicate subpaths.

Otherwise, each vertex of U_1 precedes each vertex of W_2. Let U_{12} contain the k elements of U_1 furthest down the path. Let U_{11} be the set of vertices among the first $2k$ vertices of P which have an arc to U_{12}^+. By the discussion at the beginning of the proof, $|U_{11}| \geq k$. Similarly, let W_{21} contain the k elements of W_2 first appearing on the path, and let W_{22} be the set of at least k vertices among the last $2k$ on P which have an arc from W_{21}^-. By the condition on D, there is an arc $w_{22}u_{11}$ from W_{22}^- to U_{11}^+. By definition, there is $u_{12} \in U_{12}^+$ and $w_{21} \in W_{21}^-$ such that the arcs $u_{11}^-u_{12}$ and $w_{21}w_{22}^+$ are in D. We can piece everything together to get a cycle on the vertices of P: start at u_{11}, move along P until u_{12}^-, then jump back to u. Move along P until u_{11}^-, then take the shortcut to u_{12}. Continue along P to w_{21}, jump to w_{22}^+, continue to w, jump back to w_{21}^+, and continue to w_{22}. From here there is a shortcut back to u_{11}. See Figure Figure 2.

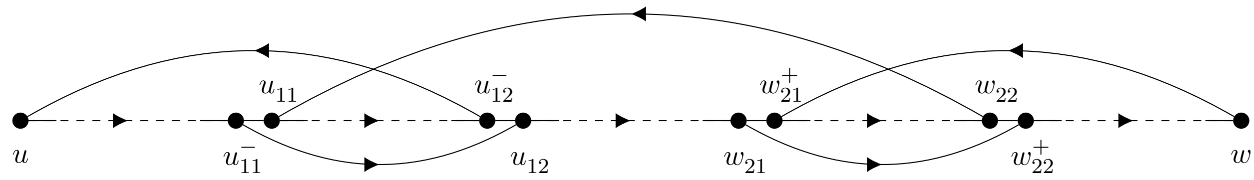

Figure 2. The case where the vertices of U_1 precede the vertices of W_2.

□

Proof of Lemma 1.3. Fix a vertex v. Let U^+ and U^- be arbitrary k-subsets of $N^+(v)$ and $N^-(v)$ respectively. There is an arc from U^+ to U^- which immediately gives a 3-cycle.

Next, let W^+ be the set of (fewer than k) vertices with no arc into U^+, and similarly let W^- be the set of vertices with no arc from U^-. Now consider the induced digraph D' obtained from D by deleting v and the vertices in U^+, U^-, W^+, W^-. Since we have removed fewer than $4k$ vertices, D' satisfies the conditions of Lemma 1.3 so has a Hamilton cycle. In particular, for every ℓ satisfying $4 \le \ell \le n - 4k$, there is a path $P_\ell = u_\ell \dots w_\ell$ in D' of length $\ell - 4$. By construction, there is an arc from U^+ to u_ℓ and from w_ℓ to U^-, which we can combine with arcs to and from v to get a cycle of length ℓ.

Finally, for every $\ell > n - 4k$, arbitrarily delete vertices from D to obtain an induced digraph D'' with ℓ vertices which satisfies the conditions of Lemma 1.3. Since D'' has a Hamilton cycle, D has a cycle of length ℓ. □

Proof of Theorem 1.1. In view of the discussion in Section 0.1, we assume $R \in \mathbb{D}(n,p)$ with $p = 2c(\alpha)/n$.

If $A, B \subseteq V(D)$ are disjoint sets with $|A| = |B| = \alpha n/8$, the probability that there are no arcs from A to B in $D \cup R$ is at most

$$(1-p)^{(\alpha n/8)^2} \le e^{-p\alpha^2 n^2/64} = e^{-c(\alpha)\alpha^2 n/32}.$$

The number of choices of such pairs of disjoint sets A, B is

$$\binom{n}{\alpha n/4}\binom{\alpha n/4}{\alpha n/8} = O\left(\frac{1}{\sqrt{n}} e^{n(H(\alpha/4)+\alpha H(1/2)/4)}\right),$$

where $H(\alpha) = -\alpha \log \alpha - (1-\alpha)\log(1-\alpha)$ (this can be proved with Stirling's approximation). By a union bound, the probability that $D \cup R$ does not satisfy the condition of Lemma 1.3 is at most

$$O\left(\frac{1}{\sqrt{n}} e^{n\left(H(\alpha/4)+\alpha H(1/2)/4 - c(\alpha)\alpha^2 n/32 + o(1)\right)}\right).$$

This converges to zero if $c(\alpha)$ is chosen to be sufficiently large. □

Remark 1.5. If we are only interested in Hamiltonicity, we can use Lemma 1.3 directly in the proof of Theorem 1.1, which gives a better constant $c(\alpha)$ than in [BFM03, Theorem 3], for all α. If we make the adjustment for undirected graphs mentioned in Remark 1.4, then we also beat the corresponding constant in [BFM03, Theorem 1] for large values of α.

2 Perfect Matchings and Hamilton Cycles in Hypergraphs

For hypergraphs, there is no single most natural notion of a cycle or of minimum degree.

Definition 2.1. We say that a k-uniform hypergraph is a *cycle of type g* if there exists a cyclic ordering of the vertices of C such that every hyperedge consists of k consecutive vertices and every pair of consecutive hyperedges has $k - g$ vertices in common. In particular, a *loose cycle* is a cycle of type $k - 1$, and a matching is a cycle of type k.

Definition 2.2. Let H be a k-uniform hypergraph. The *degree* of a set of vertices $s \subseteq V(H)$ is the number of hyperedges that include s. The *minimum q-degree* $\delta_q(H)$ of H is the minimum degree among all $s \subseteq V(H)$ with $|s| = q$.

We say graphs are dense if they have many edges (have large δ_0), or more strongly if they have large minimum degree (have large δ_1). The idea of minimum q-degree generalizes this for hypergraphs. Indeed, for a k-regular hypergraph H, a double-counting argument shows that if $q \leq p$ then

$$\delta_q(H) \geq \delta_p(H) \binom{n-q}{p-q} \bigg/ \binom{k-q}{p-q}.$$

So, imposing that a k-uniform hypergraph has large $(k-1)$-degree ensures that it has large q-degrees for all q.

Definition 2.3. Let $\mathbb{H}_k(n, m)$ be the random k-uniform hypergraph on the vertex set $[n]$ chosen uniformly from all 3-uniform hypergraphs with m hyperedges.

In this section we will prove the following two theorems.

Theorem 2.4. *There is $c_k^{\mathrm{M}}(\alpha)$ such that if H is a k-uniform hypergraph on $[kn]$ with $\delta_{k-1}(H) \geq \alpha n$, and $R \in \mathbb{H}\big(kn, c_k^{\mathrm{M}}(\alpha)n\big)$, then $H \cup R$ a.a.s. has a perfect matching.*

Theorem 2.5. *There is $c_k^{\mathrm{H}}(\alpha)$ such that if H is a k-uniform hypergraph on $[(k-1)n]$with $\delta_{k-1}(H) \geq \alpha n$, and $R \in \mathbb{H}\big((k-1)n, c_k^{\mathrm{H}}(\alpha)n\big)$, then $H \cup R$ a.a.s. has a loose Hamilton cycle.*

Our requirement $\delta_{k-1}(H) = \Omega(n)$ actually implies $\delta_q(H) = \Omega\big(n^{k-q}\big)$ for all q. We do not know whether such requirements on the q-degrees with $q < k-1$ suffice for the conclusions of our theorems. We also do not know how many random edges are necessary for other (non-loose) types of Hamilton cycles.

2.1 Proofs of Theorems 2.4 and 2.5

Our theorems will follow from a sequence of lemmas. The first step is to show that R almost gives the structure of interest on its own. Let a *sub-cycle* be a hypergraph which can be extended to a loose Hamilton cycle by adding edges.

Lemma 2.6. *For any $\varepsilon > 0$, there is $\beta_k'(\varepsilon)$ such that $R \in \mathbb{H}_k(kn, \beta_k'(\varepsilon))$ a.a.s. has a matching of size $(1 - \varepsilon)n$.*

Lemma 2.7. *For any $\varepsilon > 0$, there is $\gamma_k'(\varepsilon)$ such that $R \in \mathbb{H}_k((k-1)n, \gamma_k'(\varepsilon))$ a.a.s. has a sub-cycle of size $(1-\varepsilon)n$.*

Proof of Lemma 2.6. To realize the distribution $\mathbb{H}_k(kn, \beta_k'(\varepsilon)n)$, imagine that independently uniformly random hyperedges are repeatedly dropping into R, until we have seen $\beta_k'(\varepsilon)$ distinct hyperedges. We iteratively build a matching M greedily. If a new hyperedge e drops into R which does not share any vertices with the hyperedges in M so far, then add e to M. If we have already added i hyperedges to M, then the probability we will add the next random hyperedge is

$$p_i = \binom{kn-ki}{k} \Big/ \binom{kn}{k} = \Omega\left(\left(\frac{n-i}{n}\right)^k\right).$$

The number of random edges required to build a matching of size $(1-\varepsilon)n$ therefore has the distribution of

$$X = \sum_{i=1}^{(1-\varepsilon)n} X_i,$$

where each X_i is independently geometrically distributed with success probability p_i. By Chebyshev's inequality or a Chernoff bound, X is a.a.s. close to its expectation, which is at most

$$O\left(\sum_{i=1}^{(1-\varepsilon)n} n^k/(n-i)^k\right) = O\left(n^k \int_{\varepsilon n}^{\infty} \frac{1}{x^k} dx\right) = O(n).$$

So, if $\beta_k'(\varepsilon)$ is large enough then a.a.s. $|M| \geq (1-\varepsilon)n$. □

Proof of Lemma 2.7. A sub-cycle is a loose Hamilton cycle or a disjoint union of loose paths, but a union of disjoint paths may not be a sub-cycle if it has so many paths that there are not enough vertices left to link them together. A union of j paths (which have length greater than zero) with a total of i hyperedges is a sub-cycle precisely when $i + j \leq n$. Call such sub-cycles *proper* sub-cycles.

As in the proof of Lemma 2.6, imagine that independently uniformly random hyperedges are repeatedly dropping into R. We build a proper sub-cycle S greedily. Suppose we are at a point where S has i hyperedges and j paths, with $i \leq (1-\varepsilon)n$. There are $(k-1)i + j \leq (k-2)i + n$ vertices in the paths in S.

There are 3 kinds of hyperedges we can add to S: a hyperedge that starts a new loose path ("type 1"), a hyperedge that extends an existing path ("type 2"), or a hyperedge that joins two paths together ("type 3"). We can only add a type 1 hyperedge when $i + j \leq n - 2$, and can only add a type 2 hyperedge when $i + j \leq n - 1$.

If $i + j \leq n - 2$ then there are at least

$$\binom{(k-1)n - ((k-2)i + n)}{k} = \binom{(k-2)(n-i)}{k} = \Omega\Big((n-i)^k\Big)$$

type-1 hyperedges that could possibly be added. Otherwise, we must have $j > n - 2 - i \geq \varepsilon n - 2$, so (identifying one vertex in an extremal hyperedge of each path as "eligible to be joined"), there are at least

$$\binom{j}{2}\binom{(k-2)(n-i)}{k-2} = \Omega\Big(n^2(n-i)^{k-2}\Big) = \Omega\Big((n-i)^k\Big)$$

type-3 hyperedges that could be added. Combining both cases, we can see that if S has $i \le (1-\varepsilon)n$ hyperedges, then the next edge will be added with probability at least

$$\Omega\Big((n-i)^k\Big) \Big/ \binom{(k-1)n}{k} = \Omega\left(\left(\frac{n-i}{n}\right)^k\right).$$

We conclude with the same reasoning as the proof of Lemma 2.6. □

The second step to prove Theorems 2.4 and 2.5 is to show that a dense hypergraph plus a large partial structure a.a.s. gives the structure we are looking for. For both theorems, we will be able to reduce this step to the following lemma.

Lemma 2.8. *There is $\xi(\alpha) > 0$ such that the following holds. Let G be a bipartite graph with partitions A, B of equal size n, and suppose $\delta(G) \ge \alpha n$. Let M be a uniformly random matching with $(1-\xi(\alpha))n$ edges between A and B. Then $G \cup M$ has a perfect matching.*

Proof. We use Hall's marriage theorem: we need to show that a.a.s. $|N_{G\cup M}(W)| \ge |W|$ for all $W \subseteq A$. If $|W| \le \alpha n$, then $|N_{G\cup M}(W)| \ge |N_G(W)| \ge \alpha n \ge |W|$ by the degree condition on G. Similarly, if $|W| \ge (1-\alpha)n$ then every $b \in B$ has an edge to W in G, so $|N_{G\cup M}(W)| = |B| \ge |W|$. The difficult case is where $\alpha n \le |W| \le (1-\alpha)n$. It is relatively straightfoward to prove that for any individual such W, a.a.s. $|N_{G\cup M}(W)| \ge |W|$. However the probability of failure is not small enough for a union bound over all choices of W. Our approach is to use Szemerédi's regularity lemma on G and to show that the random matching a.a.s. distributes very evenly over the resulting clusters. Then, ε-regularity and the even distribution of the added edges are enough to ensure that $|N_{G\cup M}(W)| \ge |W|$ for each W.

We use a bipartite version of Szemerédi's regularity lemma (which can be deduced from say [Tao05, Theorem 2.3] in a similar way to [KS96, Theorem 1.10]). Let $\alpha' = \alpha/2$ and let $\varepsilon > 0$ be a small constant depending on α that will be determined later (assume for now that $\varepsilon < \alpha/8$). There is a large constant K depending only on α such that there exist partitions $A = V_0^1 \cup \cdots \cup V_k^1$ and $B = V_0^2 \cup \cdots \cup V_k^2$ with $k \le K$, in such a way that the following conditions are satisfied. The "exceptional" clusters V_0^1 and V_0^2 both have fewer than εn vertices, and the non-exceptional clusters in A and B have equal size: $\left|V_1^1\right| = \cdots = \left|V_k^\ell\right| = an$ and $\left|V_1^1\right| = \cdots = \left|V_k^\ell\right| = bn$. There is a subgraph $G' \subseteq G$ with minimum degree at least $(\alpha' + \varepsilon)n$ such that each pair of distinct clusters V_i^1, V_j^2 $(i, j \ge 1)$ is ε-regular in G' with density zero or at least 2ε. Define the cluster graph $\mathfrak{G}$ as the bipartite graph whose vertices are the non-exceptional clusters V_i^ℓ, and whose edges are the pairs of clusters between which there is nonzero density in G'.

Fix some V_i^1 and V_j^2, with $i, j \ge 1$. The vertices of each V_i^1 are matched by M to at most an vertices; conditioning on their number, they are uniformly randomly chosen from B. Let X_b be the indicator random variable for the event that the vertex $b \in B$ is matched with a vertex in V_i^1. The random variables X_b are said to be *negatively associated* (see [JDP83, Application 3.1(c)]). This means in particular that we can apply Chernoff bounds to $\sum_{b\in V_j^2} X_b$ even though the X_b are not independent. By a union bound, a.a.s. there are at most $(ab + \xi(\alpha))n$ edges of M between every pair V_i^1, V_j^2. We assume this holds for the remainder of the proof.

Now, in G' each V_i^ℓ has at least $(\alpha' + \varepsilon)n\left|V_i^\ell\right|$ edges to other clusters. There are at most $\varepsilon n\left|V_i^\ell\right|$ edges to the exceptional cluster V_0 and at most $\left|V_i^\ell\right|^2$ edges to each other cluster. So, $d_{\mathfrak{G}}(V_i^\ell) \ge \left((\alpha'+\varepsilon)n\left|V_i^\ell\right| - \varepsilon n\left|V_i^\ell\right|\right)/\left|V_i^\ell\right|^2 \ge \alpha' k$ and $\mathfrak{G}$ has minimum degree at least $\alpha' k$.

Consider any $W \subset A$ with $\alpha n \le |W| \le (1-\alpha)n$. For each i let $\pi_i = \left|V_i^1 \cap W\right|/(an)$, and let $\mathfrak{H}$ be the set of clusters V_i^1 with $\pi_i \ge \varepsilon$. If ε is small, $\mathfrak{H}$ must be nonempty. Now, if $V_j^2 \in N_{\mathfrak{G}}(\mathfrak{H})$ then by ε-regularity there are edges in G' from W to at least $(1-\varepsilon)bn$ vertices of V_j^2. It follows that if $|N_{\mathfrak{G}}(\mathfrak{H})| = k$ (that is, every V_i^ℓ is a neighbour of $\mathfrak{H}$) then $|N_{G\cup M}(W)| \ge |N_{G'}(W)| \ge (1-\varepsilon)\big(n - \left|V_0^2\right|\big) \ge |W|$ for small ε, and we are done.

Otherwise, choose $\mathfrak{S} \subseteq N(\mathfrak{H})$ with $|\mathfrak{S}| = \alpha' k$ and let S be the vertices in the clusters of $\mathfrak{S}$. By the same ε-regularity argument, $|N_{G'}(W) \cap S| \ge \alpha' k(1-\varepsilon)bn \ge (1-\varepsilon)^2\alpha' n$. We are assuming there is V_j^2 outside $N_{\mathfrak{G}}(\mathfrak{H})$, and this V_j^2 must have $\alpha' k$ neighbours outside $\mathfrak{H}$ in $\mathfrak{G}$, so $|\mathfrak{H}| \le (1-\alpha')k$. Since $|M| = (1-\xi(\alpha))n$, each $V_i^1 \cap W$ has at least $(\pi_i a - \xi(\alpha))n$ neighbours in M, and at most $\alpha' k(ab + \xi(\alpha))n$ of these are in S. So,

$$\begin{aligned} |N_M(W)\backslash S| &\ge \sum_{V_i^1 \in \mathfrak{H}} (\pi_i a - \xi(\alpha))n - |\mathfrak{H}|\big(\alpha' k(ab + \xi(\alpha))\big)n \\ &\ge |W| - \varepsilon n - \big(1-\alpha'\big)\alpha'(1-\varepsilon)^2 - k^2\xi(\alpha)n \end{aligned}$$

and

$$|N_{G\cup M}(W)| \ge |W| + \Big(\alpha'(1-\varepsilon)^2 - \varepsilon - \big(1-\alpha'\big)\alpha'(1-\varepsilon)^2 - k^2\xi(\alpha)\Big)n.$$

If ε is chosen to be small enough relative to α and $\xi(\alpha)$ chosen to be small enough relative to K, then this gives $|N_{G\cup M}(W)| \ge |W|$. $\square$

Now we describe the reduction of Theorems 2.4 and 2.5 to Lemma 2.8. Consider a k-uniform hypergraph H. Suppose A is a set of n vertices and B is a $(k-1)$-uniform hypergraph on the remaining vertices. Then we define a bipartite graph $G_{A,B}(H)$ as follows. The vertices of $G_{A,B}(H)$ are the vertices in A, as well as the edges in B (we abuse notation and identify the hypergraph B with its edge set). We put an edge between $a \in A$ and $\{b_1, \dots, b_{k-1}\} \in B$ if $\{a, b_1, \dots, b_{k-1}\}$ is a hyperedge in H.

The significance of $G_{A,B}(H)$ is that if B is a perfect matching or loose Hamilton cycle on $V(H)\backslash A$, and if $G_{A,B}(H)$ has a large matching M, then the edges of M correspond to a large matching (respectively, a large sub-cycle) in H. Conversely, if H has a large matching or Hamilton cycle (as provided by Lemmas 2.6 and 2.7), then there is a set A and a perfect matching (respectively, loose Hamilton cycle) B such that $G_{A,B}(H)$ has a large matching.

By the symmetry of the distribution of R, we can assume A and B are uniformly random. We give one final pair of lemmas, ensuring that the linear minimum $(k-1)$-degree of H corresponds a.a.s. to linear minimum degree in $G_{A,B}(H)$.

Lemma 2.9. *There is $\beta^{\mathrm{M}}(\alpha) > 0$ such that the following holds. Let H satisfy the conditions of Theorem 2.4, let A be a uniformly random set of n vertices, and let B be a uniformly random perfect matching on $V(H)\backslash A$. Then a.a.s. $G_{A,B}(H)$ has minimum degree $\beta^{\mathrm{M}}(\alpha)n$.*

Lemma 2.10. *There is $\beta^{\mathrm{H}}(\alpha) > 0$ such that the following holds. Let H satisfy the conditions of Theorem 2.5, let A be a uniformly random set of n vertices, and let B be a uniformly random loose Hamilton cycle on $V(H)\backslash A$. Then a.a.s. $G_{A,B}(H)$ has minimum degree $\beta^{\mathrm{H}}(\alpha)n$.*

Proof of Lemma 2.9. It is helpful to conceptualize the randomness in a certain way. Let

$$a_1, \dots, a_n,\ b_1^1, \dots, b_k^1,\ b_1^2, \dots, b_k^n$$

be a uniformly random ordering of $V(H)$. Let $A = \{a_1, \dots, a_n\}$, let $b^i = \{b_1^i, \dots, b_k^i\}$ and let $B = \{b^1, \dots, b^n\}$.

First, condition on some $b = b^i$. Note that $\Pr(b \cup \{a_1\} \in E(H)) \geq \delta_{k-1}(H)/n$ and more generally

$$\Pr(b \cup \{a_j\} \in E(H) \mid a_1, \dots, a_j) \geq (\delta_{k-1}(H) - j)/n \geq \alpha/2$$

for $j \leq \alpha n/2$. By a Chernoff bound, $d_{G_H}(b) \geq \alpha^2 n/8$ with probability $1 - e^{-\Omega(n)}$. With a union bound, this a.a.s. holds for all b.

Now, $\delta_1(H) \geq \alpha' n^{k-1}$ for some α' depending on α. Condition on some $a = a_i$. Each b_j^i shares at most $\binom{n}{k-2}$ hyperedges with a_i, so if $j \leq 2\sqrt{\beta^{\mathrm{M}}(\alpha)}n$ for sufficently small $\beta^{\mathrm{M}}(\alpha)$ and large n, then

$$\Pr\big(\{a\} \cup b^{j+1} \in E(H) \mid b^1, \dots, b^j\big) \geq \left(\delta_1(H) - kj\binom{n}{k-2}\right) \Big/ \binom{n}{k-1} \geq \sqrt{\beta^{\mathrm{M}}(\alpha)}.$$

By a Chernoff bound and a union bound, a.a.s. each $d_{G_H}(a) \geq \beta^{\mathrm{M}}(\alpha)$. $\square$

Proof of Lemma 2.10. We give essentially the same proof as for Lemma 2.9. Let

$$a_1, \dots, a_n, b_0, \dots, b_{(k-2)n-1}$$

be a uniformly random ordering of $V(H)$. Let $A = \{a_1, \dots, a_n\}$, let $b^i = \{b_{(k-2)i}, b_{(k-2)i+1}, \dots, b_{(k-2)(i+1)}\}$ (where the subscripts are interpreted modulo $(k-2)n$), and let $B = \{b^1, \dots, b^n\}$.

With exactly the same proof as for Lemma 2.9, there is small $\beta^{\mathrm{H}}(\alpha)$ such that a.a.s. $d_{G_H}(b) \geq \beta^{\mathrm{H}}(\alpha)n$ for all b. Next, condition on some $a = a_i$. For small $\zeta(\alpha)$ and $j \leq 2\sqrt{\beta^{\mathrm{H}}(\alpha)}n$,

$$\Pr\big(\{a\} \cup b^j \in E(H) \mid b^0, \dots, b^{j-1}\big) \geq \left(\delta_2(H) - (k-2)j\binom{n}{k-3}\right) \Big/ \binom{n}{k-2} \geq \sqrt{\beta^{\mathrm{H}}(\alpha)},$$

so a.a.s. each $d_{G_H}(a) \geq \beta^{\mathrm{H}}(\alpha)n$. $\square$

We have established that $G_{A,B}(H)$ is a bipartite graph with linear minimum degree. If we condition on A and B, then $G_{A,B}(R)$ a.a.s. provides a uniformly random large matching between A and B, so Lemma 2.8 ensures the existence of a perfect matching in $G_{A,B}(H \cup R)$, corresponding to a perfect matching or loose Hamilton cycle in $H \cup R$.

3 Multiple Hamilton cycles in tournaments

There are several seemingly different conditions that are equivalent to Hamiltonicity for tournaments (see [Moo68, Chapters 2-3]). A tournament is Hamiltonian if and only if it is irreducible (cannot be divided into two partitions with all arcs between the two partitions in the same direction), if and only if it is strongly connected (has a directed path from every vertex to every other), if and only if it is pancyclic (contains cycles of all lengths). It was first proved in [MM62] that a uniformly random tournament is a.a.s. irreducible, hence Hamiltonian.

It was more recently proved in [KLOP14] that if a tournament is t-strongly connected (it remains strongly connected after the deletion of $t-1$ vertices), then it has $\Omega\big(\sqrt{t}/\log t\big)$ arc-disjoint Hamilton cycles (this was improved to $\Omega\big(\sqrt{t}\big)$ in [Pok14]). It is not difficult to show that a random tournament is a.a.s. t-connected for fixed t, which implies that a random tournament contains q arc-disjoint Hamilton cycles for $q = O(1)$.

Just as for graphs, it is not true that every sufficiently large tournament is Hamiltonian. For example, a *transitive* tournament (corresponding to a linear order on the vertices) is never strongly connected, and to create a path against the direction of the linear order, we must change a linear number of edges. However, flipping linearly many random edges cannot possibly suffice to make a transitive tournament Hamiltonian, because this is not enough to a.a.s. flip one of the arcs pointing away from the least element of the linear order. However, in [ŁRG96, Section 3] it was proved that any superlinear number of flips suffices: flipping each arc with probability $\omega(1/n)$ will a.a.s. result in a Hamiltonian tournament.

Unsurprisingly, we find that transitive tournaments are a "worst case", in that flipping a superlinear number of arcs suffices for all T. We also find that the vertices in a transitive tournament with only outgoing or only incoming arcs are a "bottleneck" for strong connectivity, and a gentler random perturbation suffices for tournaments which do not have such "imbalanced" vertices. Recall that $\delta(T)$ is the minimum of the in- and out- degrees among the vertices in T.

Theorem 3.1. *Consider a tournament T with n vertices and $\delta(T) \geq d$. Independently designate each arc as "random" with probability p satisfying $np(d+1) \to \infty$, then choose the orientations of the chosen edges independently and uniformly at random. The resulting perturbed tournament P a.a.s. has q arc-disjoint Hamilton cycles, for $q = O(1)$.*

This theorem is sharp, in the sense that if we do not have $np(d+1) \to \infty$ then P might not be a.a.s. Hamiltonian. To see this, consider a "transitive cluster-tournament" T on $n = rk$ vertices defined as follows. Let R be a regular tournament on $2d+1$ vertices (this means the indegree and outdegree are equal for each vertex). To construct T, start with r disjoint copies $R_1, \ldots, R_r$ of R, then put an arc from v to w for every $v \in R_i$, $w \in R_j$ with $i < j$. In order for the perturbed tournament P to be Hamiltonian, there must be an arc entering R_1, so one of the $O(n(d+1))$ arcs exiting R_1 must be resampled. This will not happen a.a.s. unless $pn(d+1) \to \infty$.

3.1 Proof of Theorem 3.1

Fix t; we prove that P is t-strongly connected. The idea of the proof is to choose a set S of t vertices with a large indegree and outdegree. Then we show that with high probability almost every vertex has many paths to and from each vertex in S.

Lemma 3.2. *In any tournament, there are less than k vertices with indegree (respectively outdegree) less than $(k-1)/2$.*

Proof. The sum of indegrees (respectively outdegrees) of a tournament on k vertices is $\binom{k}{2}$, because each arc contributes 1 to this sum. Therefore in every set of k vertices of a tournament, there is a vertex of outdegree (indegree) at least $(k-1)/2$ in the induced tournament. $\square$

It follows from Lemma 3.2 that the set of all vertices with indegree (respectively outdegree) less than $n/6$ has size smaller than $n/3$. So, there are at least $n/3$ vertices with indegree and outdegree at least $n/6$. Since $t = O(1) = o(n/6)$, there is a set S of t such vertices.

Lemma 3.3. *If a vertex v has outdegree (respectively indegree) k in T, then it has outdegree (respectively indegree) at least $k/3$ in P, with probability $e^{-\Omega(\sqrt{n})}$ uniformly over k.*

Proof. We only prove the statement where v has outdegree k; the indegree case is identical. If $k = 0$ the lemma is trivial, so assume $k \geq 1$. First consider the cases where $k > \sqrt{n}$ or $p < 1/\sqrt{n}$. There are k arcs pointing away from v in T; let X be the number of those arcs that are changed by the perturbation. Since $X \in \mathrm{Bin}(k, p/2)$, a Chernoff bound gives

$$\Pr(X > 2k/3) \leq \Pr(X - \mathbb{E}X > k/6) \leq e^{-\Omega\left(k^2/\mathbb{E}X\right)} = e^{-\Omega\left(\sqrt{n}\right)}.$$

That is, with the required probability, $k/3$ of the original out-neighbours survive the perturbation.

Otherwise (for large n), there is a set of $n/2$ arcs pointing towards v in T. Let Y be the number of these arcs that are changed by the perturbation, so $Y \in \mathrm{Bin}(n/2, p/2)$ with $\mathbb{E}Y \geq \sqrt{n}/2$. Then,

$$\Pr(Y < k/3) \leq \Pr(X - \mathbb{E}Y < -\mathbb{E}Y/3) \leq e^{-\Omega\left(\sqrt{n}\right)},$$

so with the required probability, $k/3$ new out-neighbours are added by the perturbation. $\square$

Lemma 3.4. *Suppose w has outdegree (respectively indegree) at least $n/6$, and $v \neq w$ is a vertex with indegree (respectively outdegree) at least k. Then with probability $1 - e^{-\Omega\left(\sqrt{n}\right)} - e^{-\Omega(np(k+1))}$ (uniformly over k), there are $t' = 3t$ internally vertex-disjoint paths of length at most 3 from w to v (respectively from v to w).*

Proof. We will only prove the statement where v has indegree at least k; the other case is identical. By independence, we can condition on the outcome of the perturbation on individual arcs. Condition on the outcome for all arcs adjacent to w, and let $N_P^+(w)$ be the set of vertices to which there is an arc from w in P. By Lemma 3.3, we can assume $\left|N_P^+(w)\right| \geq n/18$.

We first prove the lemma for the case where $k \leq 12t'$. There are at least $n' = n/18 - 1$ arcs between $N_P^+(w)$ and v, each of which will be pointing towards v in P with independent probability at least $p/2$. Let $Z \in \mathrm{Bin}(n', p/2)$. The probability that less than t' arcs will point from $N_P^+(w)$ to v in P is at most

$$\Pr\big(Z < t'\big) \leq e^{-\Omega(\mathbb{E}Z - t')} = e^{-\Omega(np)} = e^{-\Omega(np(k+1))}.$$

So, with the required probability there are t' suitable length-2 paths from w to v.

We can now assume $k > 12t'$. Condition on the result of the perturbation for the arcs adjacent to v. Let $N_P^-(v)$ be the set of vertices from which there is an edge into v in P; by Lemma 3.3, we can assume $\left|N_P^-(v)\right| \geq k/3$.

Now, if $\left|N_P^+(w) \cap N_P^-(v)\right| \geq t'$ then there are t' disjoint length-2 paths from w to v and we are done. So we can assume $U^+ = N_P^+(w) \backslash \left(N_P^-(v) \cup \{v\}\right)$ has at least $n' = n/18 - t' - 1$ vertices, and $U^- = N_P^-(v) \backslash \left(N_P^+(w) \cup \{w\}\right)$ has at least $k' = k/6 - t' - 1$ vertices (note $k' \geq t'$ by assumption).

Now, we would like to show that with the required probability there is a set of t' independent arcs from U^+ into U^- in P, which will give t' suitable length-3 paths. Partition U^+(respectively U^-) into subsets $U_1^+, \dots, U_{t'}^+$ (respectively $U_1^-, \dots, U_{t'}^-$), such that each $\left|U_1^+\right| > n'/(2t')$ and each $\left|U_1^-\right| > k'/(2t')$. For each i, the probability that there is no arc from U_i^+ into U_i^- after the perturbation is at most

$$(1 - p/2)^{n'k'/(4t')} \leq e^{-\Theta(np(k+1))},$$

so with the required probability there a set of t' suitable independent arcs, each between a pair U_i^+, U_i^-. □

Lemma 3.5. *Fix some $w \in S$. In P, there are a.a.s. t disjoint paths from w to each other vertex (respectively, from each other vertex to w).*

Proof. We only prove there are paths from w to each other vertex; the reverse case is identical. If there are $3t$ disjoint paths of length at most 3 from w to v then we say v is *safe*. It follows from Lemma 3.4 that a vertex with indegree k is safe with probability $1 - e^{-\Omega\left(\sqrt{n}\right)} - e^{-\Omega(np(k+1))}$.

By Lemma 3.2, there are at most $(2d + 1)$ vertices with outdegree d, and the vertex with the $2k$th smallest outdegree has outdegree at least $k - 1$. Let Q be the set of non-safe vertices, and note

$$\begin{aligned}
\mathbb{E}|Q| &= (2d+1)e^{-\Omega(np(d+1))} + \sum_{k=d+1}^{n} 2e^{-\Omega(np(k+1))} + ne^{-\Omega\left(\sqrt{n}\right)} \\
&= e^{-\Omega(np(d+1))}\left(O(d+1) + \frac{1}{1 - e^{-\Omega(np)}}\right) + o(1) \\
&= e^{-\Omega(np(d+1))}O(d+1).
\end{aligned}$$

where we have used the geometric series formula, the inequality $1 - e^{-x} \geq (x \wedge 1)/2$ (for positive x) and the fact that $d + 1 = \Omega(1)$.

By Markov's inequality, a.a.s. $|Q| \leq \sqrt{e^{-\Omega(pn(d+1))}}O(d+1) = o(d+1)$. If $d \leq 4t$ then $|Q| = 0$ for large n. Otherwise, for large n, $|Q| \leq 3d$, so every vertex $v \in Q$ has $3t$ safe neighbours $v_1, \dots, v_{3t}$. Now, fix a maximal set M of disjoint paths from w to v. If $|M| < t$ then the paths in M collectively use less than $3t$ vertices other than t, so there is some v_i not in any path of M, and there is some path from w to v_i that does not contain any vertex of M. But then $w \dots v_i v$ is a path from w to v disjoint with every path in M, contradicting maximality. □

It follows from Lemma 3.5 that a.a.s. every vertex outside S has t disjoint paths to and from every vertex in S. If we delete $t - 1$ vertices, then there is at least one vertex w of S remaining, and w has least one path to and from every other vertex. That is, P is strongly connected. (In fact, we have also proved that P has diameter at most 8).

References

[BFKM04] Tom Bohman, Alan Frieze, Michael Krivelevich, and Ryan Martin, *Adding random edges to dense graphs*, Random Structures & Algorithms **24** (2004), no. 2, 105–117.

[BFM03] Tom Bohman, Alan Frieze, and Ryan Martin, *How many random edges make a dense graph Hamiltonian?*, Random Structures & Algorithms **22** (2003), no. 1, 33–42.

[Bon75] JA Bondy, *Pancyclic graphs: Recent results*, Infinite and finite sets: to Paul Erdős on his 60th birthday, Volume 1 (A Hajnal, R Rado, and Vera T Sos, eds.), Colloquia mathematica Societatis János Bolyai, vol. 10, North-Holland Publishing Company, Amsterdam-London, 1975, pp. 181–187.

[Dir52] Gabriel Andrew Dirac, *Some theorems on abstract graphs*, Proceedings of the London Mathematical Society **3** (1952), no. 1, 69–81.

[FK05] Alan Frieze and Michael Krivelevich, *On packing Hamilton cycles in ε-regular graphs*, Journal of Combinatorial Theory, Series B **94** (2005), no. 1, 159–172.

[JDP83] Kumar Joag-Dev and Frank Proschan, *Negative association of random variables with applications*, The Annals of Statistics (1983), 286–295.

[JŁR00] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński, *Random graphs*, Cambridge University Press, 2000.

[KLOP14] Daniela Kühn, John Lapinskas, Deryk Osthus, and Viresh Patel, *Proof of a conjecture of thomassen on Hamilton cycles in highly connected tournaments*, Proceedings of the London Mathematical Society (2014), pdu019.

[KS75] J Komlós and E Szemerédi, *Hamilton cycles in random graphs*, Infinite and finite sets: to Paul Erdős on his 60th birthday, Volume 2 (A Hajnal, R Rado, and Vera T Sos, eds.), Colloquia mathematica Societatis János Bolyai, vol. 10, North-Holland Publishing Company, Amsterdam-London, 1975, pp. 1003–1010.

[KS96] J Komlós and M Simonovits, *Szemerédi's regularity lemma and its applications in graph theory*, Combinatorics, Paul Erdős is eighty (D Miklós, VT Sos, and T Szonyi, eds.), Bolyai society mathematical studies, vol. 2, János Bolyai Mathematical Society, Budapest, 1996, pp. 295–352.

[KST06] Michael Krivelevich, Benny Sudakov, and Prasad Tetali, *On smoothed analysis in dense graphs and formulas*, Random Structures & Algorithms **29** (2006), no. 2, 180–193.

[ŁRG96] Tomasz Łuczak, Andrzej Ruciński, and Jacek Gruszka, *On the evolution of a random tournament*, Discrete Mathematics **148** (1996), no. 1, 311–316.

[MM62] JW Moon and L Moser, *Almost all tournaments are irreducible*, Canadian Mathematical Bulletin **5** (1962), 61–65.

[Moo68] JW Moon, *Topics on tournaments*, Holt, Rinehart and Winston, 1968.

[Pok14] Alexey Pokrovskiy, *Edge disjoint Hamiltonian cycles in highly connected tournaments*, arXiv preprint arXiv:1406.7556 (2014).

[Pós76] L Pósa, *Hamiltonian circuits in random graphs*, Discrete Mathematics **14** (1976), no. 4, 359–364.

[ST03] Daniel A Spielman and Shang-Hua Teng, *Smoothed analysis: Motivation and discrete models*, Algorithms and data structures: 8th International Workshop, WADS 2003, Ottawa, Ontario, Canada, July 30 - August 1, 2003, Proceedings (Frank Dehne, Jörg Rüdiger Sack, and Michie Smid, eds.), Lecture Notes in Computer Science, vol. 2748, Springer, 2003, pp. 256–270.

[ST04] ———, *Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time*, Journal of the ACM **51** (2004), no. 3, 385–463.

[Tao05] Terence Tao, *Szemerédi's regularity lemma revisited*, arXiv preprint math/0504472 (2005).