Lectures 3 and 4
Hamiltonicity threshold in random graphs

The main aim of these two lectures is to establish the Hamiltonicity threshold in the probability space $G(n,p)$, this is the minimum value of the edge probability $p(n)$, for which a random graph G drawn from $G(n,p)$ is **whp** Hamiltonian. By doing so we will prove a classical result of Komlós and Szemerédi [3] and independently of Bollobás [2].

Let us start by providing some intuition on where this threshold is expected to be located. It is well known that the threshold probability for connectivity in $G(n,p)$ is $p = \frac{\ln n}{n}$. More explicitly, one can prove (and we leave this as an exercise) that for any function $\omega(n)$ tending to infinity arbitrarily slowly with n, if $p = \frac{\ln n - \omega(n)}{n}$, then **whp** $G \sim G(n,p)$ is not connected, whereas for $p = \frac{\ln n + \omega(n)}{n}$ **whp** $G \sim G(n,p)$ is connected. Perhaps more importantly, the main reason for the threshold for connectivity to be around $\ln n/n$ is that precisely at this value of probability the last isolated vertex in $G(n,p)$ typically ceases to exist. Of course, the graph cannot be connected while having isolated vertices, and this is the easy part of the connectivity threshold statement; the hard(er) part is to prove that if $p(n)$ is such that $\delta(G) \geq 1$ **whp**, then G is **whp** connected.

If so, we can suspect that the threshold for Hamiltonicity of $G(n,p)$ coincides with that of non-existence of vertices of degree at most one, the latter being an obvious necessary condition for Hamiltonicity. This is exactly what was proven in [3, 2]. Let us therefore set our goal by stating first a fairly accessible result about the threshold for $\delta(G) \geq 2$, both in $G(n,p)$ and in $G(n,m)$.

Proposition 1 *Let $\omega(n)$ be any function tending to infinity arbitrarily slowly with n. Then:*

- *in the probability space $G(n,p)$,*
  1. *if $p(n) = \frac{\ln n + \ln\ln n - \omega(n)}{n}$, then $G \sim G(n,p)$ **whp** satisfies $\delta(G) \leq 1$;*
  2. *if $p(n) = \frac{\ln n + \ln\ln n + \omega(n)}{n}$, then $G \sim G(n,p)$ **whp** satisfies $\delta(G) \geq 2$;*
- *in the probability space $G(n,m)$,*
  1. *if $m(n) = \frac{(\ln n + \ln\ln n - \omega(n))n}{2}$, then $G \sim G(n,m)$ **whp** satisfies $\delta(G) \leq 1$;*
  2. *if $m(n) = \frac{(\ln n + \ln\ln n + \omega(n))n}{2}$, then $G \sim G(n,m)$ **whp** satisfies $\delta(G) \geq 2$.*

Proof Straightforward application of the first (for proving $\delta(G) \geq 2$) and the second (for proving $\delta(G) \leq 1$) methods in both probability spaces; left as an exercise. □

Hence our goal will be to prove that for $p(n) = \frac{\ln n + \ln\ln n + \omega(n)}{n}$ and for $m(n) = \frac{(\ln n + \ln\ln n + \omega(n))n}{2}$ the random graphs $G(n,p)$ and $G(n,m)$ respectively are **whp** Hamiltonian.

We will actually prove a stronger, and a much more delicate, result about the hitting time for Hamiltonicity in random graph processes. Let us first define this notion formally. Let $\sigma : E(K_n) \to [N]$ be a permutation of the edges of the complete graph K_n on n vertices, we can write $\sigma = (e_1, \ldots, e_N)$, where $N = \binom{n}{2}$. A *graph process* $\tilde{G} = \tilde{G}(\sigma)$ is a nested sequence $\tilde{G} = (G_i)_{i=0}^N$, where the graph G_i has $[n]$ as its vertex set and $\{e_1, \ldots, e_i\}$, the prefix of σ of length i, as its edge set. The sequence (G_i) thus starts with the empty graph on n vertices, finishes with the complete graph on n vertices, and its i-th element G_i has exactly i edges; moreover, it is nested, as for $i \geq 1$ the graph G_i is obtained from its predecessor G_{i-1} by adding the i-th edge e_i of σ. We can view $\tilde{G}(\sigma)$ as a graph process (as the name indicates suggestively) or as an evolutionary process, unraveling from the empty graph to the complete graph, guided by σ.

Now, we introduce the element of randomness in the above definition. Suppose the permutation σ is drawn uniformly at random from the set of all $N!$ permutations of the edges of K_n. Then the corresponding process $\tilde{G}(\sigma)$ is called a *random graph process*. We can describe it in the following equivalent way: start by setting G_0 to be the empty graph on n vertices, and for each $1 \leq i \leq N$, obtain G_i by choosing an edge e_i of K_n missing in G_{i-1} uniformly at random and adding it to G_{i-1}. This very nice and natural probability space models a random evolutionary process in graphs; here too we proceed from the empty graph to the complete graph, but in a random fashion.

Random graph processes are so important not just because they model evolution very nicely; in fact, they embed the probability spaces $G(n,m)$ for various m; due to standard connections between $G(n,m)$ and $G(n,p)$ one can also claim they "contain" $G(n,p)$ as well. Observe that running a random process $\tilde{G}$ and stopping it (or taking a *snapshot*) at time m produces the probability distribution $G(n,m)$. Indeed, every graph G with vertex set $[n]$ and exactly m edges is the m-th element of the same number of graph processes, namely, of $m!(N-m)!$ of them. Thus, understanding random graph processes usually leads to immediate consequences for $G(n,m)$, and then for $G(n,p)$, and Hamiltonicity is not exceptional in this sense.

Let $\mathcal{P}$ be a property of graphs on n vertices; assume that P is monotone increasing (i.e., adding edges preserves it), and that the complete graph K_n possesses $\mathcal{P}$ (you can think of $\mathcal{P}$ as being the property of Hamiltonicity). Then, given permutation $\sigma : E(K_n) \to [N]$ and the corresponding graph process $\tilde{G}(\sigma)$, we can define the first moment i when the i-th element G_i of $\tilde{G}$ has $\mathcal{P}$. This is the so called *hitting time* of $\mathcal{P}$, denoted by $\tau_{\mathcal{P}}(\tilde{G}(\sigma))$:

$$\tau_{\mathcal{P}}(\tilde{G}(\sigma)) = \min\{i \geq 0 : G_i \text{ has } \mathcal{P}\}\,.$$

Of course, due to the monotonicity of $\mathcal{P}$ from this point till the end of the process the graphs G_i all have $\mathcal{P}$. When $\tilde{G}$ is a random graph process, the hitting time $\tau_{\mathcal{P}}(\tilde{G})$ becomes a random variable,

and one can study its typical behavior. A related task is compare two hitting times, and to try to bundle them, deterministically or probabilistically.

We now state the main result of this part of the course, due to Ajtai, Komlós and Szemeredi [1], and to Bollobás [2].

Theorem 2 *Let $\tilde{G}$ be a random graph process on n vertices. Denote by $\tau_2(\tilde{G})$ and $\tau_{\mathcal{H}}(\tilde{G})$ the hitting times of the properties of having minimum degree at least 2, and of Hamiltonicity, respectively. Then* **whp***:*

$$\tau_2(\tilde{G}) = \tau_{\mathcal{H}}(\tilde{G})\,.$$

In words, for a typical graph process, Hamiltonicity arrives *exactly* at the very moment the last vertex of degree less than two disappears. Of course, it cannot arrive earlier deterministically, so the main point of the above theorem is to prove that typically is does not arrive later either.

As we indicated above, random graph process results are usually more powerful than those for concrete random graph models. Here too we are able to derive the results for $G(n,m)$ and $G(n,p)$ easily from the above theorem.

Corollary 3 *Let $m(n) = \frac{(\ln n + \ln\ln n + \omega(n))n}{2}$. Then a random graph $G \sim G(n,m)$ is* **whp** *Hamiltonian.*

Proof Generate a random graph G distributed according to $G(n,m)$ by running a random graph process $\tilde{G}$ and stopping it at time m. By Proposition 1 we know that $\tau_2(\tilde{G}) \leq m$. Theorem 2 implies that typically $\tau_2(\tilde{G}) = \tau_{\mathcal{H}}(\tilde{G})$, and thus the graph of the process has become Hamiltonian not later than m. Hence G is **whp** Hamiltonian as well. □

Corollary 4 *Let $p(n) = \frac{\ln n + \ln\ln n + \omega(n)}{n}$. Then a random graph $G \sim G(n,m)$ is* **whp** *Hamiltonian.*

Proof Observe that generating a random graph $G \sim G(n,p)$ and conditioning on its number of edges being exactly equal to m produces the distribution $G(n,m)$. Let $\omega_1(n) = \omega(n)/3$. Denote $I = [Np - n\omega_1(n), Np + n\omega_1(n)]$. Observe that for every $m \in I$, the random graph $G \sim G(n,m)$ is **whp** Hamiltonian by Corollary 3. Also, the number of edges in $G(n,p)$ is distributed binomially with parameters N and p and has thus standard deviation less than $\sqrt{Np} \ll n\omega_1(n)$. Applying Chebyshev we derive that **whp** $|E(G)| \in I$. Hence

$$\begin{aligned} Pr[G \sim G(n,p) \text{ is not Hamiltonian}] &= \sum_{m=0}^{N} Pr[|E(G)| = m] \cdot Pr[G \text{ is not Hamiltonian}\,|\,|E(G) = m] \\ &\leq Pr[|E(G)| \notin I] + \sum_{m \in I} Pr[|E(G)| = m]\, Pr[G \text{ is not Hamiltonian}\,|\,|E(G) = m] \\ &= o(1) + \sum_{m \in I} Pr[|E(G)| = m]\, Pr[G \sim G(n,m) \text{ is not Hamiltonian}| = o(1) \cdot Pr[Bin(N,p) \in I] \\ &= o(1)\,. \end{aligned}$$

□

Now we start proving Theorem 2. The proof is somewhat technical, so before diving into its details, we outline its main idea briefly. Recall that our goal is to prove that for a typical random process $\tilde{G}$, we have $\tau_2(\tilde{G}) = \tau_{\mathcal{H}}(\tilde{G})$. In order to prove this, we will take a very close look at the snapshot G_{τ_2} of $\tilde{G}$, aiming to prove that this graph is **whp** Hamiltonian. By definition, the minimum degree of G_{τ_2} is exactly two, and it is thus quite reasonable to expect that this graph is typically a $(k, 2)$-expander for $k = \Theta(n)$. This is true indeed, however such expansion by itself does not quite guarantee Hamiltonicity. As indicated in Lecture 1, expanders form a very convenient backbone for augmenting a graph to a Hamiltonian one – according to Corollary 3.5 of Lecture 1 every connected non-Hamiltonian $(k, 2)$-expander has $\Omega(k^2)$ boosters. Observe though that since we aim to prove a hitting time result, we cannot allow ourselves to sprinkle few random edges on top on our expander - a Hamilton cycle should appear at the very moment the minimum degree in the random graph process becomes two. We will circumvent this difficulty in the following way: we will argue that the snapshop G_{τ_2} typically is not only a good expander by itself, but also contains a subgraph Γ_0 who is about as good an expander as G_{τ_2} is, but contains only a small positive proportion on its edges. Having obtained such Γ_0 we will start looking for boosters relative to Γ_0, but already contained in our graph G_{τ_2} – thus avoiding the need for sprinkling. We will argue that G_{τ_2} is typically such that it is contains a booster with respect to every sparse expander in it. If this is the case, then we will be able to start with Γ_0 and to update it sequentially by adding a booster after a booster (at most n boosters will need to be added by definition), until we will finally reach Hamiltonocity – all within G_{τ_2}; observe crucially that at each step of this augmentation procedure the updated backbone Γ_i, obtained by adjoining to Γ_0 the previously added boosters, has at most n more edges than Γ_0 and is thus still a sparse subgraph of G_{τ_2}; of course the required expansion is inherited from an iteration to iteration. Then our claim about G_{τ_2} typically containing a booster with respect to every sparse expander within is applicable, and we can push the process through. This is quite a peculiar proof idea – it appears that the random graph is helping itself to become Hamiltonian!

Let us get to work. As outlined before, we run a random graph process $\tilde{G}$ and take a snapshot at the hitting time $\tau_2 = \tau_2(\tilde{G})$. Denote

$$\begin{aligned} m_1 &= \frac{n \ln}{2} \\ m_2 &= n \ln n \,. \end{aligned}$$

Observe that by Proposition 1 we have that **whp** $m_1 \leq \tau_2 \leq m_2$. Let

$$d_0 = \lfloor \delta_0 \ln n \rfloor$$

where $\delta_0 > 0$ is a sufficiently small constant to be chosen later, and denote, for a graph G on n vertices,

$$SMALL(G) = \{v \in V(G) : d(v) < d_0\} \,.$$

Observe that for $G \sim G(n,m)$ with $m \geq m_1$, the expected vertex degree is asymptotically equal to $\ln n$. Thus falling into $SMALL(G)$ is a rather rare event, and we can expect the vertices of $SMALL(G)$ to be few and far apart in the graph. In addition, such G should have typically very nice edge distribution, with no small and dense vertex subsets, and many edges crossing between any tow large disjoint subsets. This is formalized in the following lemma.

Lemma 5 *Let $\tilde{G} = (G_i)_{i=0}^N$ be a random graph process on n vertices. Denote $G = G_{\tau_2}$, where $\tau_2 = \tau_2(\tilde{G})$ is the hitting time for having minimum degree two in $\tilde{G}$. Then* **whp** *G has the following properties:*

(P1) $\Delta(G) \leq 10 \ln n$; $\delta(G) \geq 2$;

(P2) $|SMALL(G)| \leq n^{0.3}$;

(P3) *G does not contain a non-empty path of length at most 4 such that both of its (possibly identical) endpoints lie in $SMALL(G)$;*

(P4) *every vertex subset $U \subset [n]$ of size $|U| \leq \frac{n}{\ln^{1/2} n}$ spans at most $|U| \cdot \ln^{3/4} n$ edges in G;*

(P5) *for every pair of disjoint vertex subsets U, W of sizes $|U| \leq \frac{n}{\ln^{1/2} n}$, $|W| \leq |U| \cdot \ln^{1/4} n$, the number of edges of G crossing between U and W is at most $\frac{d_0|U|}{2}$;*

(P6) *for every pair of disjoint vertex subsets U, W of size $|U| = |W| = \left\lceil \frac{n}{\ln^{1/2} n} \right\rceil$, G has at least $0.5n$ edges between U and W.*

Proof The proof is a fairly standard (though tedious) manipulation with binomial coefficients. We will thus prove several of the above items, leaving the proof of remaining ones to the reader.

(P1): Observe that since **whp** $\tau_2 \leq m_2$, it is enough to prove that in $G \sim G(n, m_2)$ there are **whp** no vertices of degree at least $10 \ln n$. For a given vertex $v \in [n]$, the probability that v has degree at least $10 \ln n$ in $G(n, m_2)$ is at most

$$\binom{n-1}{10\ln n} \frac{\binom{N-10\ln n}{m_2-10\ln n}}{\binom{N}{m_2}} \leq \left(\frac{en}{10\ln n}\right)^{10\ln n} \left(\frac{m_2}{N}\right)^{10\ln n},$$

by the standard estimates on binomial coefficients stated in Lecture 1. After cancellations we see that the above estimate is at most $(2en/10(n-1))^{10\ln n} = o(1/n)$. Applying the union bound we obtain that typically at time m_2 and thus at $\tau_2 \leq m_2$ as well there are no vertices of degree at least $10 \ln n$. The bound on $\delta(G)$ is immediate from the definition of τ_2.

(P2): Notice that since adding edges can only decrease the size of $SMALL(G)$, it is enough to prove that typically already at time m_1 $|SMALL(G_{m_1})| \leq n^{0.3}$. Let $G \sim G(n, m_1)$. If $|SMALL(G)| \geq n^{0.3}$,

then G contains a subset $V_0 \subset V$, $|V_0| = k = \lceil n^{0.3} \rceil$ such that $e_G(V_0, V - V_0) \leq d_0 k$. The probability of this to happen in $G(n, m_1)$ is at most:

$$\binom{n}{k} \sum_{i \leq d_0 k} \binom{k(n-k)}{i} \cdot \frac{\binom{N-k(n-k)}{m_1-i}}{\binom{N}{m_1}} \leq \binom{n}{k} \sum_{i \leq d_0 k} \binom{kn}{i} \cdot \frac{\binom{N-k(n-k)}{m_1-i}}{\binom{N-i}{m_1-i}} \cdot \frac{\binom{N-i}{m_1-i}}{\binom{N}{m_1}}$$

$$\leq \left(\frac{en}{k}\right)^k \sum_{i \leq d_0 k} \left(\frac{ekn}{i}\right)^i \cdot e^{-\frac{(m_1-i)(k(n-k)-i)}{N-i}} \cdot \left(\frac{m_1}{N}\right)^i \leq \left(\frac{en}{k}\right)^k \sum_{i \leq d_0 k} \left(\frac{ekm_1 n}{iN}\right)^i \cdot e^{-\frac{0.9km_1 n}{N}}$$

$$\leq \left(\frac{en}{k}\right)^k \left[(d_0 k + 1) \cdot \left(\frac{ekm_1 n}{d_0 kN}\right)^{d_0 k} \cdot e^{-\frac{0.9km_1 n}{N}}\right]$$

$$\leq (d_0 k + 1) \left[3n^{0.7} \left(\frac{3 \ln n}{d_0}\right)^{d_0} \cdot e^{-0.8 \ln n}\right]^k = o(1)\,,$$

for δ_0 small enough.

(P3): Since **whp** $m_1 \leq \tau_2 \leq m_2$, it is enough to prove the following statement:**whp** every two (possibly identical) vertices of $SMALL(G_{m_1})$ are not connected by a path of length at most 4 in G_{m_2}.

Let us prove first that **whp** there is no such path in $G_{m_1} \sim G(n, m_1)$. First we treat the case where the endpoints of the path are distinct. Fix $1 \leq r \leq 4$, a sequence P of distinct vertices $v_0, \ldots, v_r$ in $[n]$ and denote by $\mathcal{A}_P$ the event $(v_i, v_{i+1}) \in E(G_{m_1}$ for every $0 \leq i \leq r-1$. Then

$$Pr[\mathcal{A}_P] = \frac{\binom{N-r}{m_1-r}}{\binom{N}{m_1}} \leq \left(\frac{m_1}{N}\right)^r = \left(\frac{\ln n}{n-1}\right)^r .$$

If we now condition on $\mathcal{A}_P$, then the two edges (v_0, v_1) and (v_{r-1}, v_r) are present in G_{m_1}. Thus in order for both v_0, v_r to fall into $SMALL(G_{m_1})$, out of $2n-4$ potential edges between $\{v_0, v_r\}$ and the rest of the graph (the edges $(v_0, v_1), (v_{r-1}, v_r)$ are excluded from the count), only at most $2d_0 - 2$ are present in G_{m_1}. Hence:

$$\begin{aligned}
Pr[v_0, v_r \in SMALL(G_{m_1}) | \mathcal{A}_P] &\leq \sum_{i=0}^{2d_0-2} \binom{2n-4}{i} \cdot \frac{\binom{N-r-2n+4}{m_1-r-i}}{\binom{N-r}{m_1-r}} \\
&\leq (2d_0 - 1)\binom{2n-4}{2d_0-2} \cdot \frac{\binom{N-r-2n+4}{m_1-r-2d_0+2}}{\binom{N-r}{m_1-r}} \\
&\leq 2d_0 \binom{2n-4}{2d_0-2} \cdot \frac{\binom{N-r-2n+4}{m_1-r-2d_0+2}}{\binom{N-r-2d_0+2}{m_1-r-2d_0+2}} \cdot \frac{\binom{N-r-2d_0+2}{m_1-r-2d_0+2}}{\binom{N-r}{m_1-r}} \\
&\leq 2d_0 \cdot \left(\frac{en}{d_0-1}\right)^{2d_0-2} \cdot e^{-\frac{(m_1-r-2d_0+2)(2n-2d_0-2)}{N-r-2d_0+2}} \cdot \left(\frac{m_1-r}{N-r}\right)^{2d_0-2} \\
&\leq 2d_0 \cdot \left(\frac{em_1 n}{(d_0-1)N}\right)^{2d_0-2} \cdot e^{-\frac{1.9m_1 n}{N}} \leq n^{-1.8}\,,
\end{aligned}$$

for δ_0 small enough. Hence, applying the union bound over all such sequences of $r+1$ vertices, we conclude that the probability that there exists a path in G_{m_1} of length at most 4, connecting two

distinct vertices from $SMALL(G_{m_1})$ is at most $\sum_{r\leq 4} n^{r+1}\cdot\left(\frac{\ln n}{n-1}\right)^r\cdot n^{-1.8}=o(1)$. The case where the endpoitns of the path are identical is treated similarly.

In light of the above, we can assume that after m_1 steps of the random graph process the current graph does not have a forbidden short path between the vertices of $SMALL$. Moreover, by **(P2)** we can assume that $|SMALL(G_{m_1})|\leq n^{0.3}$. Now, let us run the process between m_1 and m_2. In order for the i-th edge of the process, $m_1<i\leq m_2$, to close a short path between the vertices of $SMALL(G_{m_1})$, it should fall inside a current set U of vertices at distance at most 3 from $SMALL(G_{m_1})$. We have proven (property **(P1)**) that in fact **whp** the maximum degree of G_{m_2} as well is at most $10\ln n$. Hence, **whp** in this time interval, the set U has size at most $|SMALL(G_{m_1}|\cdot(10\ln n)^3$, and thus the probability of the i-th edge of the process to fall inside U is at most $\frac{\binom{|U|}{2}}{N-m_2}=o(n^{-1.3})$. Taking the union bound over all such i in the interval $(m_1,m_2]$, we establish the desired property.

Properties **(P4)**–**(P6)** can be proven quite similarly (and are in fact simpler to prove), and we spare the reader from the (perhaps somewhat boring...) task of reading their proofs.

□

The above stated properties **(P1)**–**(P6)** are sufficient to prove that G_{τ_2} is a very good expander by itself. Our goal is somewhat different though – we aim to prove that G_{τ_2} contains a much sparser, but still fairly good expander. For this purpose, assume that a graph $G=(V,E)$ has properties **(P1)**–**(P6)**. Form a random subgraph Γ_0 of G as follows. For every $v\in V-SMALL(G)$, choose a set $E(v)$ of d_0 edges of G incident to v uniformly at random; for every $v\in SMALL(G)$, define $E(v)$ to be the set of *all* edges of G touching v. Finally, define Γ_0 to be the spanning subgraph of G, whose edge set is:

$$E(\Gamma_0)=\bigcup_v E(v)\,.$$

In words, in order to form Γ_0 we retain all edges touching the vertices of $SMALL(G)$, and sparsify randomly other edges.

Lemma 6 *With high probability (over the choices of $E(v)$) the subgraph Γ_0 is a $(k,2)$-expander with at most $d_0 n$ edges, where $k=\frac{n}{4}$.*

Proof Since by definition $|E(v)|\leq d_0$ for every $v\in V$, it follows immediately that $|E(\Gamma_0|\leq d_0 n$. We now prove that typically Γ_0 has the following property:

(P7) For every pair of disjoint sets U,W of size $|U|=|W|=\left\lceil\frac{n}{\ln^{1/2}n}\right\rceil$, Γ_0 has at least one edge between U and W.

Fix sets U,W as above. We know by **(P6)** that G has at least $0.5n$ edges between U and W. For a vertex $u\in U$, the probability that none of the edges between u and W falls into $E(u)$ is at most

$$\frac{\binom{d_G(u)-d_G(u,W)}{d_0}}{\binom{d_G(u)}{d_0}}\leq e^{-\frac{d_0\cdot d_G(u,W)}{d_G(u)}}\leq e^{-\frac{d_0}{10\ln n}\cdot d_G(u,W)}$$

by **(P1)**. Hence the probability that none of the vertices u from U chooses an edge between u and W to be put into its set $E(u)$ is at most:

$$\prod_{u\in U} e^{-\frac{d_0}{10\ln n}\cdot d_G(u,W)} = e^{-\frac{d_0}{10\ln n}\cdot e_G(U,W)} = e^{-\Theta(n)}\,.$$

Applying the union bound over all choices of U, W gives the desired claim.

We now claim that for every graph G, and every its subgraph Γ_0 of minimum degree 2 satisfying properties **(P2)**, **(P3)**, **(P4)**, **(P5)**, **(P7)**, the subgraph Γ_0 is an $(n/4, 2)$-expander. In order to verify this claim, let $S \subset [n]$ be a subset of size $|S| \leq n/4$. Denote $S_1 = S \cap SMALL(G)$, $S_2 = S - SMALL(G)$. Consider first the case where $|S_2| \leq \frac{n}{\ln^{1/2} n}$. Since $\delta(\Gamma_0) \geq 2$, and all vertices from $SMALL$ are at distance more than 4 from each other by **(P3)**, we obtain: $|N_{\Gamma_0}(S_1)| \geq 2|S_1|$. As for vertices from S_2, they are all of degree at least d_0 in Γ_0. The set S_2 spans at most $|S_2| \cdot \ln^{3/4} n$ edges in G, and thus in Γ_0, according to **(P4)**. It thus follows that $e_{\Gamma_0}(S_2, V - S_2) \geq d_0|S_2| - 2e_{\Gamma_0}(S_2) > \frac{d_0|S_2|}{2}$. Hence $|N_{\Gamma_0}(S_2)| \geq |S_2| \cdot \ln^{1/4} n$, by **(P5)**. Finally, notice that, due to the non-existence of short paths connecting $SMALL(G)$ again, the set $N_{\Gamma_0}(S_1)$ contains only one vertex from $u \cup N_{\Gamma_0}$ for every $u \in S_2$ (here we use the fact the forbidden paths have length at most 4). Therefore, $|S_1 \cap (S_2 \cup N_{\Gamma_0}(S_2))| \leq |S_2|$. Altogether,

$$\begin{aligned} |N_{\Gamma_0}(S)| &= |N_{\Gamma_0}(S_2)| + |N_{\Gamma_0}(S_1) - (S_2 \cup N_{\Gamma_0}(S_2))| \\ &\leq |S_2| \cdot \ln^{1/4} n + 2|S_1| - |S_2| \leq 2(|S_1| + |S_2|) = 2|S|\,, \end{aligned}$$

as required. The complementary case $\frac{n}{\ln^{1/2} n} \leq |S_2| \leq \frac{n}{4}$ is very simple: by property **(P7)**, such S_2 misses at most $n/\ln^{1/2} n$ vertices in its neighborhood in Γ_0, also $|S_1| \leq |SMALL(G)| \leq n^{0.3}$ by **(P2)**. In follows that $|N_{\Gamma_0}(S)| \geq n - \frac{n}{\ln^{1/2} n} - |S_2| - |SMALL(G)| \geq \frac{n}{2}$. □

Notice that every $(\frac{n}{4}, 2)$-expander Γ on n vertices is necessarily connected. Indeed, if such Γ is not connected, then consider its component C of size $|C| \leq \frac{n}{2}$, and take U to be an arbitrary subset of C of size $|U| = \min\{\lfloor \frac{n}{4} \rfloor, |C|\}$. Then the external neighborhood of U in Γ has size at least $2|U| > |C - U|$ by our expansion assumption, and falls entirely within C – a contradiction.

As we stated already in this lecture, expanders are not necessarily Hamiltonian themselves, but they are amenable to reaching Hamiltonicity by adding extra (random) edges, as they contain many boosters. However, in our circumstances we do not have extra time for sprinkling, and the required boosters should come from within the already existing edges of the random graph. Fortunately, a random graph $G(n, m)$ with $m = m(n)$ in the relevant range has **whp** a booster with respect to any sparse expander it contains, as given by the following lemma.

Lemma 7 *Let $\tilde{G} = (G_i)_{i=0}^N$ be a random graph process on n vertices. Denote $G = G_{\tau_2}$, where $\tau_2 = \tau_2(\tilde{G})$ is the hitting time for having minimum degree two in $\tilde{G}$. Assume the constant δ_0 is small enough. Then* **whp** *for every $(n/4, 2)$-expander $\Gamma \subset G$ with $V(\Gamma) = V(G)$ and $|E(\Gamma)| \leq d_0 n + n$, Γ is Hamiltonian , or G contains at least one booster with respect to Γ.*

Proof Recall that every connected $(k,2)$-expander Γ has at least $k^2/2$ boosters, by Corollary 3.5 of Lecture 1. In order for a random graph G to violate the assertion of the lemma, G should contain some $(n/4,2)$-expander Γ with few edges, but none of at least as many as $n^2/32$ boosters relative to Γ (note that the required connectivity of G is delivered by the expansion of Γ, as explained above). Since we cannot pinpoint the exact location of τ_2, we instead take the union bound over all $m_1 \le m \le m_2$, as **whp** τ_2 is located in this interval. So the estimate is:

$$\sum_{m=m_1}^{m_2} \sum_{i \le d_0 n + n} \frac{\binom{N}{i} \cdot \binom{N-i-\frac{n^2}{32}}{m-i}}{\binom{N}{m}} + o(1) \tag{1}$$

(we sum over all relevant values of m, adding in the end $o(1)$ to account for the probability that τ_2 falls outside the interval $[m_1, m_2]$; then we sum over all possible values i of $|E(\Gamma)|$; then we bound from above by $\binom{N}{i}$ the number of $(n/4,2)$-expanders with i edges in the complete graph on n vertices, and finally we require the edges of Γ to be present in $G(n,m)$, but all at least $n^2/32$ boosters relative to Γ to be omitted). The ratio of the binomial coefficients above can be estimated as follows:

$$\frac{\binom{N-i-\frac{n^2}{32}}{m-i}}{\binom{N}{m}} \le \frac{e^{-\frac{\frac{n^2}{32}(m-i)}{N-i}} \binom{N-i}{m-i}}{\binom{N}{m}} \le e^{-\frac{m}{17}} \left(\frac{m}{N}\right)^i ,$$

assuming that δ_0 in the definition of d_0 is small enough. We can thus estimate the i-th summand in (1) as follows:

$$\binom{N}{i} e^{-\frac{m}{17}} \left(\frac{m}{N}\right)^i \le \left(\frac{eN}{i} \cdot \frac{m}{N}\right)^i \cdot e^{-\frac{m}{17}} = \left(\frac{em}{i}\right)^i \cdot e^{-\frac{m}{17}} \le \left(\frac{em}{d_0 n + n}\right)^{d_0 n + n} \cdot e^{-\frac{m}{17}} = o(n^{-3}) ,$$

again for δ_0 small enough (it is even exponentially, and not just polynomially, small in n). Summing over all $i \le d_0 n + n$ and then over all $m_1 \le m \le m_2$ establishes the required claim. □

The stage is now set to deliver the final (and crushing!) blow in the proof of Theorem 2. Recall that our goal is to prove that for a random graph process $\tilde{G}$, **whp** the graph $G = G_{\tau_2}$ at the very moment τ_2 when the minimum degree becomes 2 is already Hamiltonian. First, observe that by Lemma 6 **whp** G contains an $(n/4,2)$-expander Γ_0 with at most $d_0 n$ edges. We start with this sparse expander Γ_0 and keep adding boosters to it until the current graph Γ_i becomes Hamiltonian; obviously at most n steps (edge additions) will be needed to reach Hamiltonicity. If we ever get stuck before reaching Hamiltonicity, say at step $i \ge 0$, then the current graph Γ_i is still an $(n/4,2)$-expander, is connected and non-Hamiltonian, has at most $d_0 n + n$ edges, but the graph G has no boosters with respect to Γ_i. This however does not happen typically due to Lemma 7. If so, the process of edge addition eventually completes with a subgraph $\Gamma_i \subset G$, which is Hamiltonian. The proof is complete!

References

[1] M. Ajtai, J. Komlós, E. Szemerédi, First occurrence of Hamilton cycles in random graphs, Cycles in graphs (Burnaby, B.C., 1982), North-Holland Mathematical Studies **115**, North-Holland, Amsterdam (1985), 173–178.

[2] B. Bollobás, *The evolution of sparse graphs*, Graph Theory and Combinatorics, Academic Press, London (1984), 35–57.

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