Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

Constructing Hamilton cycles and perfect matchings efficiently

(Extended abstract)

Michael Anastos *

Abstract

Starting with the empty graph on $[n]$, at each round, a set of $K = K(n)$ edges is presented chosen uniformly at random from the ones that have not been presented yet. We are then asked to choose at most one of the presented edges and add it to the current graph. Our goal is to construct a Hamiltonian graph with $(1 + o(1))n$ edges within as few rounds as possible.

We show that in this process, one can build a Hamiltonian graph of size $(1+o(1))n$ in $(1+o(1))(1+(\log n)/2K)n$ rounds w.h.p. The case $K = 1$ implies that w.h.p. one can build a Hamiltonian graph by choosing $(1 + o(1))n$ edges in an online fashion as they appear along the first $(0.5 + o(1))n \log n$ rounds of the random graph process. This answers a question of Frieze, Krivelevich and Michaeli. Observe that the number of rounds is asymptotically optimal as the first $0.5n \log n$ edges do not span a Hamilton cycle w.h.p. The case $K = \Theta(\log n)$ implies that the Hamiltonicity threshold of the corresponding Achlioptas process is at most $(1+o(1))(1+(\log n)/2K)n$. This matches the $(1-o(1))(1+(\log n)/2K)n$ lower bound due to Krivelevich, Lubetzky and Sudakov and resolves the problem of determining the Hamiltonicity threshold of the Achlioptas process with $K = \Theta(\log n)$.

We also show that in the above process one can construct a graph G that spans a matching of size $\lfloor V(G)/2 \rfloor$ and $(0.5+o(1))n$ edges within $(1+o(1))(0.5+(\log n)/2K)n$ rounds w.h.p.

Our proof relies on a robust Hamiltonicity property of the strong 4-core of the binomial random graph which we use as a black-box. This property allows it to absorb

*Institute of Science and Technology Austria, Klosterneurburg 3400, Austria. E-mail: michael.anastos@ist.ac.at. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101034413  .

paths covering vertices outside the strong 4-core into a cycle.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-005`

1 Introduction

Let $G_0, G_1, ..., G_N$, $N = \binom{n}{2}$ be the random graph process. That is, G_0 is the empty graph on $[n]$ and G_{i+1} is formed by adding to G_i an edge chosen uniformly at random from the non-present ones, for $0 \leq i < N$. Equivalently let $e_1, e_2, ..., e_N$ be a permutation of the edges of the complete graph K_n chosen uniformly at random and set $G_i = ([n], \{e_1, ..., e_i\})$, $0 \leq i < N$. Let τ_2 be the minimum i such that G_i has minimum degree 2 and τ_H be the minimum i such that G_i is Hamiltonian. Building upon work of Pósa [13] and Korshunov [11], Bollobás [6] and independently Ajtai, Komlós and Szemerédi [1] proved that $\tau_2 = \tau_H = 0.5n(\log n + (1 + o(1)) \log \log n)$ w.h.p.[1] Thus, to achieve Hamiltonicy, one has to wait until the minimum degree becomes 2. Unfortunately, this necessary condition is satisfied w.h.p. only by graphs of the random graphs process that have at least $0.5n \log n$ edges, while a Hamilton cycle uses only n of them. This raises the following question. Can one built a Hamiltonian subgraph of G_t that spans $(1 + o(1))n$ edges in an online fashion for some $t = (1 + o(1))\tau_2$?

Frieze, Krivelevich and Michaeli studied a generalization of this question in the following setting [9]. Once again let $e_1, e_2, ..., e_N$ be a permutation of $E(K_n)$ chosen uniformly at random. The sequence $e_1, e_2, ..., e_N$ is revealed, one edge at a time. Starting with the empty graph on $[n]$, as soon as an edge is revealed we must decide, immediately and irrevocably, whether to choose and add it to our graph. Let B_i be the graph constructed after the ith edge has been revealed. Let $\mathcal{B}'_{HAM}$ be the set of pairs (t, b) for which there exists an algorithm that builds a Hamiltonian graph of size at most b within the first t rounds of the above process w.h.p. Clearly, as $B_i \subseteq G_i$ for all i and $\tau_2 > 0.5n \log n$ w.h.p., a necessary condition for $(t, b) \in \mathcal{B}'_{HAM}$ is that $t \geq 0.5n \log n$ and $b \geq n$. Frieze, Krivelevich and Michaeli proved that for every $\epsilon > 0$ there exists $C > 0$ such that if $t \geq (0.5 + \epsilon)n \log n$ and $b \geq 9n$ or $t \geq Cn \log n$ and $b \geq (1 + \epsilon)n$ then $(t, b) \in \mathcal{B}'_{HAM}$. They also asked whether there exist $\epsilon > 0$ and a pair t, b such that $t \leq (0.5 + \epsilon)n \log n$, $b \leq (1 + \epsilon)n$ and $(t, b) \notin \mathcal{B}'_{HAM}$. Theorem 1.1 answers this question.

A second way to generalize our question is within the framework of the Achlioptas processes. Inspired by the "power of two choices" paradigm Achiloptas proposed the following process. Starting with the empty graph on $[n]$, at each round, a set of $K = K(n)$ edges is presented chosen uniformly at random from the ones that have not been presented yet (or from all $\binom{n}{2}$ possible ones). We are then asked to choose one of them to add to the current graph, immediately and irrevocably. The aim of the Achlioptas process is to accelerate or delay a given graph property. For example, Bohman and Frieze proved that there exist $\epsilon > 0$ and a strategy that w.h.p. ensure that one can construct a graph with no component

[1] We say that a sequence of events $\{\mathcal{E}_n\}_{n \geq 1}$ holds *with high probability* if $\lim_{n \to \infty} \mathbf{Pr}(\mathcal{E}_n) = 1 - o(1)$.

of size $\Omega(n)$ after $(1+\epsilon)n/2$ rounds, thus delaying the appearance of the giant [4]. Krivelevich, Lubetzky and Sudakov studied $\tau_H(K)'$, the minimum number of rounds needed to construct a Hamiltonian graph in the above process [12]. They proved that w.h.p.

$$(1+o(1))\left(1+\frac{\log n}{2K}\right)n \leq \tau_H(K) \leq (1+o(1))\left(3+\frac{\log n}{K}\right)n. \tag{1}$$

To obtain the upper bound, they constructed a random 3-out graph which is known to be Hamiltonian [5]. For the lower bound they proved that for any algorithm $\mathcal{A}$ and any $\epsilon > 0$, after $(1-\epsilon)(1+0.5\log n/K)n$ rounds there exist $n^{\epsilon/2}$ vertices of degree smaller than 2 w.h.p. Their argument goes as follows. After $0.5(1-\epsilon)n$ rounds, the graph constructed so far by $\mathcal{A}$ contains at least ϵn vertices of degree smaller than 2, deterministically. From those vertices, at least $n^{\epsilon/2}$ will not be incident to any edge that will be presented in the next $0.5(1-\epsilon)n(\log n)/K$ rounds w.h.p. Any such vertices have degree at most 1 in the graph constructed so far.

Krivelevich, Lubetzky and Sudakov also proved that the lower bound in (1) is the correct one, in the sense that it is equal to $(1+o(1))\tau_H(K)$ w.h.p., in the regimes $K = o(\log n)$ and $K = \omega(\log n)$. In these regimes the lower bound reduces to $(1+o(1))(n\log n)/2K$ and $(1+o(1))n$ respectively. Theorem 1.1 implies that the lower bound in (1) is always the correct one. The problem of improving the bounds in (1) is also stated as Problem 43 in Frieze's bibliography on Hamilton cycles in random graphs [7].

Formally the process that we consider is the following one. Starting with the empty graph on $[n]$, at each round, a set of $K = K(n)$ edges is presented chosen uniformly at random from the ones that have not been presented yet. We are then asked to choose *at most* one of them to add to the current graph immediately and irrevocably. We let B_i be the graph constructed after i rounds. We let $\mathcal{B}_{HAM} = \mathcal{B}_{HAM}(K)$ be the set of pairs $(t,b) = (t(K), b(K))$ for which there exists an algorithm that builds a Hamiltonian graph of size at most b within the first t rounds of the above process w.h.p. Similarly, we let $\mathcal{B}_{PM} = \mathcal{B}_{PM}(K)$ be the set of pairs (t,b) for which there exists an algorithm that builds a graph of size at most b that spans a matching of size $\lfloor n/2 \rfloor$ within the first t rounds of the above process w.h.p.

Theorem 1.1. *Let $K = K(n) = O(\log n)$. Then,*

$$\left(\left(1+\frac{250}{\log\log n}\right)\left(1+\frac{\log n}{2K}\right)n, \left(1+\frac{11}{\log\log n}\right)n\right) \in \mathcal{B}_{HAM}.$$

The case $K = \omega(\log n)$ of the above theorem follows from Theorem 1.2 of [12]. Once again, as G_t has minimum degree 0 for $t \leq 0.5n\log n$ w.h.p., one has that $(t,b) \in \mathcal{B}_{PM}$ only if $t \geq 0.5n\log n$ and $b \geq n/2$.

Theorem 1.2.

$$\left(\left(1+\frac{250}{\log\log n}\right)\left(0.5+\frac{\log n}{2K}\right)n, \left(0.5+\frac{11}{\log\log n}\right)n\right) \in \mathcal{B}_{PM}.$$

Ramark 1.3. *Frieze, Krivelevich and Michaeli gave an alternative proof to Theorem 1.2 for the case $K = 1$ (See Theorem 4 of [9]).*

Ramark 1.4. *One may consider the variations of the process where at every round, the K edges that are presented are chosen uniformly at random from all $\binom{n}{2}$ possible edges or from the ones that are missing from the graph that is constructed so far. Theorems 1.1 and 1.2 as stated also hold for these variations.*

In this note we sketch the proof of Theorem 1.1. Theorem 1.2 can be proven in a similar manner. Both proofs are based on structural properties of the strong 4-core of a random graph which we describe in the next section.

2 The strong k-core

For a graph G we define the *strong k-core* of G to be the maximal subset S of $V(G)$ with the property that every vertex in $S \cup N(S)$ has at least k neighbors in S. By $N(S)$ we denote the set of vertices in $V(G) \setminus S$ that are adjacent to S. Observe that if the sets $S_1, S_2 \subset V(G)$ have this property, then so does the set $S_1 \cup S_2$. Thus the strong k-core of a graph is well-defined. It also naturally partitions the vertex set of a graph G into 3 sets which we denote by $V_{k,black}(G), V_{k,blue}(G)$ and $V_{k,red}(G)$ where $V_{k,black}(G)$ is the strong k-core of G, $V_{k,blue}(G)$ is its neighborhood and $V_{k,red}(G)$ is the rest i.e. $V_{k,red}(G) = V(G) \setminus (V_{k,black}(G) \cup N(V_{k,black}(G))$. In our knowledge, the strong 3-core was first used in [3] for finding the longest cycle in sparse random graphs while the concept of the strong k-core was first formalized in [2]. There it was observed that the strong 4-core of $G(n, c/n)$ is robustly Hamiltonian for $c \geq 20$ as described below. For a graph G and $U \subseteq V(G)$ denote by $G[U]$ the subgraph of G induced by U. By $G(n, p)$ we denote the binomial random graph i.e., the random graph on $[n]$ where every edge appears independently with probability p.

Theorem 2.1 (Theorem 3.3 of [2])**.** *Let $G \sim G(n, c/n)$, $c \geq 20$. Let G' be the subgraph of G induced by $V_{4,black}(G) \cup V_{4,blue}(G)$. Then for every $U \subseteq V_{4,blue}(G)$ and matching M on $V_{4,blue} \setminus U$ w.h.p. we have that $G'[V(G') \setminus U] \cup M$ has a Hamilton cycle that spans M.*

Theorem 2.1 enable us to use the strong 4-core of $G(n, 20/n)$ as an absorber for finding large cycles. Indeed, assume that a graph G contains $G' \sim G(n, 20/n)$ as a subgraph. In addition assume that there exists a set of vertex disjoint paths $\mathcal{P}$ that do not intersect $V_{4,black}(G') \cup V_{4,blue}(G')$ internally and whose endpoints lie in $V_{4,blue}(G')$. Then, given G' and $\mathcal{P}$, one can contract each path of $\mathcal{P}$ into an edge. This results to a matching M on $V_{4,blue}(G')$. Theorem 2.1 then gives that $G' \cup M$ spans a Hamilton cycle which spans all the edges in M. Replacing the edges in M with the corresponding paths in $\mathcal{P}$ gives a cycle of G whose vertex set consists of $V_{4,black}(G')$, $V_{4,black}(G')$ and the set of vertices spanned by the paths in $\mathcal{P}$. This will be our main strategy in proving Theorem 2.1.

The next lemma will also be used in the proof of Theorem 1.1. For its proof see Lemma 3.3 of [2].

Lemma 2.2. *Let $G \sim G(n, c/n)$, $c \geq 20$. Then $|V_{4,blue}(G)| \geq 0.1 \cdot (2c)^3 e^{-2c} n$ w.h.p.*

3 Constructing a Hamilton cycle online, efficiently

We now sketch the proof of Theorem 1.1. To simplify its description we only consider the case $K = 1$. Thus at round i we are presented with an edge e_i chosen uniformly at random from the ones that have not been presented yet, for $i \in [N]$. For its proof we describe an algorithm $\mathcal{A}$ that chooses $(1 + 11/\log\log n)n$ edges within the first $(1 + 250/\log\log n)(1 + \log n/2)n$ rounds and constructs a Hamiltonian graph w.h.p. Let

$$n' = \frac{n}{\log\log n}, \quad t_\epsilon = \left(\frac{50}{\log\log n}\right)\left(1 + \frac{\log n}{2}\right)n,$$

$t_0 = 0$, $t_1 = t_\epsilon$, $t_2 = t_1 + t_\epsilon + n$, $t_3 = t_2 + t_\epsilon$, $t_4 = t_3 + t_\epsilon + n(\log n/2)$ and $t_5 = t_4 + t_\epsilon$. $\mathcal{A}$ consists of 5 phases. Its ith phase starts when $e_{t_{i-1}+1}$ is presented and ends once $\mathcal{A}$ decides whether to keep the edge e_{t_i}.

During its first phase, $\mathcal{A}$ picks the first $10n'$ edges that are spanned by $[n']$. Let G_1 be the graph $\mathcal{A}$ constructed during Phase 1 of $\mathcal{A}$, $U = V_{4,black}(G) \cup V_{4,blue}(G')$, $W = V_{4,blue}(G')$ and $Z = [n] \setminus U$. Lemma 2.2 implies that $|U| = \Omega(n')$ w.h.p. The rest of the phases of $\mathcal{A}$ aim to cover the vertices in Z by a set $\mathcal{P}'$ of vertex disjoint paths with endpoints in W that do not internally intersect U. To do so, during its second phase, $\mathcal{A}$ greedily covers Z with at most $n/(\log\log n)^2$ vertex disjoint paths, each of length at most $\log n$. Here we allow paths of length 0 which correspond to single vertices. Let $\mathcal{P}$ be the set of these paths. Then, during Phase 3, $\mathcal{A}$ greedily matches the endpoints of the paths in $\mathcal{P}$ to W, each path $P \in \mathcal{P}$ is therefore potentially extended to a path with a pair of unique endpoints in W. Let $End(\mathcal{P})$ be the set of endpoints of paths in $\mathcal{P}$ that lie in Z (are left unmatched). During Phase 4, $\mathcal{A}$ attempts to match the vertices in $End(\mathcal{P})$ to $\log^{0.8} n$ many vertices in the interior of distinct paths in $\mathcal{P}$. This is possible as $t_4 - t_3 = t_\epsilon + 0.5n\log n$, which implies that each vertex in $End(\mathcal{P})$ is incident to $\omega(\log^{0.8} n)$ edges in $\{e_{t_3+1}, ..., e_{t_4}\}$ whose other endpoint lies in $[n] \setminus U$. Finally, during Phase 5, using the edges selected during Phase 4, $\mathcal{A}$ reroutes the paths in $\mathcal{P}$ with an endpoint in $End(\mathcal{P})$ through the rest of the paths. Such a rerouting may look as follows. Let $Q = v_1, v_2, ..., v_k$ and $P = u_1, u_2, ..., u_r$ be vertex disjoint paths with $v_1, v_k, u_r \in W$ and $u_1 \in Z$. In such a case, adding the edges u_1v_i and $v_{i+1}v$ with $v \in W$, $1 \le i \le k-1$ (selected during phases 4 and 5 respectively) and removing the edge v_iv_{i+1} from $E(P) \cup E(Q)$ results to 2 vertex disjoint paths that cover $V(P) \cup V(Q)$ and have their endpoints in W.

One may show that the set of edges selected during the last 4 phases span a set $\mathcal{P}'$ of vertex disjoint paths with endpoints in W that do not internally intersect U w.h.p. Given G_1 and $\mathcal{P}$, one may appeal to Theorem 2.1, as discussed in the previous section, to show the existence of a Hamilton cycle spanned by the constructed graph. Finally note that the edges selected during phases 2 and 3 span a set of paths, thus there are at most n. Therefore, in total, $\mathcal{A}$ selects $10n' + n + |End|(\log^{0.8} n + 2)$ which is equal to $(1 + o(1))n$ in the high probability event that $|End(\mathcal{P})| = o(n/\log n)$.

References

[1] Miklós Ajtai, János Komlós, and Endre Szemerédi. First occurrence of Hamilton cycles in random graphs. *North-Holland Mathematics Studies*, 115(C):173–178, 1985.

[2] Michael Anastos. A note on long cycles in sparse random graphs. *The Electronic Journal of Combinatorics 30.2*, 2023.

[3] Michael Anastos and Alan Frieze. A scaling limit for the length of the longest cycle in a sparse random graph. *Journal of Combinatorial Theory, Series B*, 148:184–208, 2021.

[4] Tom Bohman and Alan Frieze. Avoiding a giant component. *Random Structures & Algorithms*, 19(1):75–85, 2001.

[5] Tom Bohman and Alan Frieze. Hamilton cycles in 3-out. *Random Structures & Algorithms*, 35(4):393–417, 2009.

[6] Béla Bollobás. The evolution of sparse graphs, graph theory and combinatorics, 1984. *MR*, 777163(2):35–57, 1984.

[7] Alan Frieze. Hamilton cycles in random graphs: a bibliography. *arXiv preprint arXiv:1901.07139*, 2019.

[8] Alan Frieze and Michal Karoński. *Introduction to random graphs*. Cambridge University Press, 2016.

[9] Alan Frieze, Michael Krivelevich, and Peleg Michaeli. Fast construction on a restricted budget. *arXiv preprint arXiv:2207.07251*, 2022.

[10] János Komlós and Endre Szemerédi. Limit distribution for the existence of Hamiltonian cycles in a random graph. *Discrete mathematics*, 43(1):55–63, 1983.

[11] Aleksei Dmitrievich Korshunov. Solution of a problem of Erdős and Rényi on Hamiltonian cycles in nonoriented graphs. In *Doklady Akademii Nauk*, volume 228, pages 529–532. Russian Academy of Sciences, 1976.

[12] Michael Krivelevich, Eyal Lubetzky, and Benny Sudakov. Hamiltonicity thresholds in achlioptas processes. *Random Structures & Algorithms*, 37(1):1–24, 2010.

[13] Lajos Pósa. Hamiltonian circuits in random graphs. *Discrete Mathematics*, 14(4):359–364, 1976.