Tight Inapproximability of Max Independent Set in Triangle-Free Graphs

Édouard Bonnet
Univ Lyon, CNRS, ENS de Lyon, Université Claude Bernard Lyon 1, LIP UMR5668, France

Abstract

For every $\varepsilon > 0$, it is NP-hard to $n^{1-\varepsilon}$-approximate Max Independent Set in n-vertex graphs [Håstad '96, Zuckerman '07]. In triangle-free graphs, a simple argument gives a polynomial-time $n^{\frac{1}{2}}$-approximation algorithm, whereas, for every $\varepsilon > 0$, an $n^{\frac{1}{4}-\varepsilon}$-approximation algorithm would imply that NP $\subseteq$ BPP [Bonnet, Thomassé, Tran, Watrigant; ESA '20].

In this note, we close this gap by proving the corresponding hardness against $n^{\frac{1}{2}-\varepsilon}$-approximation algorithms. The reduction is very simple and uses the Moser–Tardos resampling algorithm to make the constructed graphs triangle-free. The soundness uses a result of Haeupler, Saha, and Srinivasan building on the proof of Moser and Tardos, to upper-bound the probability that a fixed relatively large subset is an independent set after the Moser–Tardos algorithm terminates.

We generalize this scheme and show that, for any nonempty finite family $\mathcal{F}$ of graphs, each containing at least one cycle, for any $\varepsilon > 0$, an $n^{\mu(\mathcal{F})-\varepsilon}$-approximation algorithm for Max Independent Set in graphs excluding every member of $\mathcal{F}$ as a subgraph implies that NP $\subseteq$ BPP, where

$$\mu(\mathcal{F}) := 1 - \max_{H \in \mathcal{F}} \min_{\substack{U \subseteq V(H) \\ H[U] \text{ contains a cycle}}} \frac{|U|-2}{|E(H[U])|-1}.$$

1 Introduction

Given a simple graph G, the Max Independent Set problem (or MIS for short) asks for a largest *independent set* in G, that is, a subset of vertices that are pairwise nonadjacent. This problem is notoriously inapproximable: For every $\varepsilon > 0$, it is NP-hard to $n^{1-\varepsilon}$-approximate in n-vertex graphs [4, 9].

It is conjectured in [1] that the class of all graphs is the only *hereditary* (i.e., closed under taking induced subgraphs) graph class on which Max Independent Set is that inapproximable. More formally, the conjecture is that for every graph H, there exists a constant $\delta_H \in [0, 1)$ such that MIS admits a (randomized) polynomial-time $\mathcal{O}(n^{\delta_H})$-approximation algorithm in H-free graphs (i.e., the graphs excluding H as an induced subgraph). This is shown to be a weakening of an algorithmic version of the celebrated Erdős–Hajnal conjecture.

We are here interested in tight approximability of MIS in hereditary classes, starting with H-free classes and $\mathcal{F}$-free classes (i.e., excluding every graph in $\mathcal{F}$ as an induced subgraph) for a finite family $\mathcal{F}$. More specifically, we wish to match an $\mathcal{O}(n^{\delta_H})$-approximation algorithm with the conditional lower bound that, for any constant $\varepsilon > 0$, an $\mathcal{O}(n^{\delta_H-\varepsilon})$-approximation algorithm contradicts some standard complexity-theoretic assumption. Prior to our work, no such result was known for any graph H.

For triangle-free graphs (that is, when H is a triangle), there is a straightforward $\mathcal{O}(n^{\frac{1}{2}})$-approximation algorithm. If the input graph has a vertex of degree at least $n^{\frac{1}{2}}$, its neighborhood, which by triangle-freeness is an independent set, yields the desired approximate solution. Otherwise the graph has maximum degree at most $n^{\frac{1}{2}}$, and every maximal independent set is large enough. On the complexity side, it was shown that, for every $\varepsilon > 0$, a polynomial-time $n^{\frac{1}{4}-\varepsilon}$-approximation algorithm would imply that NP $\subseteq$ BPP [1]. We close this gap and provide the first example of polynomial-factor tightness for MIS in H-free graphs.

▶ **Theorem 1.** *For every constant $\varepsilon > 0$, there is no polynomial-time $n^{\frac{1}{2}-\varepsilon}$-approximation algorithm for* Max Independent Set *on n-vertex triangle-free graphs unless* NP $\subseteq$ BPP.

The reduction behind Theorem 1 starts with a hard general MIS-instance H, blows every vertex into an independent set of size n, and independently samples (i.e., includes) every edge of the intermediate blow-up F with probability $\Theta(1/n)$. While the current graph contains a triangle, the algorithm resamples the three variables corresponding to its edges, thereby following the Moser–Tardos algorithm. In expectation, this process terminates after $n^{\mathcal{O}(1)}$ steps, by the work of Moser and Tardos [6, Theorem 1.2] (see Theorem 3). The obtained graph G has n^2 vertices and is, by design, triangle-free.

It is immediate that $\alpha(G) \geqslant \alpha(H) \cdot n$. On the other hand, we show that, with high probability, $\alpha(G) = \mathcal{O}(\alpha(H) \cdot n \log n)$. This uses a follow-up work by Haeupler, Saha, and Srinivasan [3, Theorem 2.2] (see Theorem 4), building on [6], in order to upper-bound the probability, for any fixed set S of $\Omega(\alpha(H) \cdot n \log n)$ vertices of G, that S is an independent set in G. We then conclude with a union bound over all sets of $\Theta(\alpha(H) \cdot n \log n)$ vertices of G. Finally, the inapproximability of MIS by Håstad [4] and Zuckerman [9] (see Theorem 5) yields Theorem 1.

We believe that the elementary proof of Theorem 1 is an accessible application of the Moser–Tardos algorithm/theorem, which can be taught within various related courses (approximation algorithms, randomness in computer science, graph independent sets, etc.). We decided that the current paper would primarily present this particular and simpler case, but the proof scheme naturally generalizes to $\mathcal{F}$-free graphs. Besides, Theorem 1 remains the only example of polynomial-factor tight approximability for MIS in graphs excluding finitely many induced subgraphs.

For every nonempty finite family $\mathcal{F}$ of graphs, each with a cycle, we set

$$\mu(\mathcal{F}) := 1 - \max_{H \in \mathcal{F}} \min_{\substack{U \subseteq V(H) \\ H[U] \text{ contains a cycle}}} \frac{|U|-2}{|E(H[U])|-1}.$$

The proof technique sketched after Theorem 1 works for $\mathcal{F}$-free graphs, and even for $\mathcal{F}$-subgraph-free graphs, i.e., those excluding every graph of $\mathcal{F}$ as a mere subgraph. We obtain the following theorem.

▶ **Theorem 2.** *Let $\mathcal{F}$ be a nonempty finite family of graphs, each containing a cycle. For every constant $\varepsilon > 0$, there is no polynomial-time $n^{\mu(\mathcal{F})-\varepsilon}$-approximation algorithm for* Max Independent Set *on n-vertex $\mathcal{F}$-subgraph-free graphs unless* NP $\subseteq$ BPP.

This theorem has some interesting consequences. For $\mathcal{F} = \{K_t\}$, where K_t is the t-vertex clique and $t \geqslant 3$,

$$\mu(\mathcal{F}) = 1 - \frac{t-2}{\binom{t}{2}-1} = 1 - \frac{t-2}{\frac{t(t-1)}{2}-1} = 1 - \frac{t-2}{\frac{(t-2)(t+1)}{2}} = 1 - \frac{2}{t+1} = \frac{t-1}{t+1}.$$

On the positive side, MIS admits an $n^{\frac{t-2}{t-1}}$-approximation algorithm in K_t-free graphs [1]. For every integer $t \geqslant 4$, there is still a gap between the hardness exponent $\frac{t-1}{t+1} - \varepsilon$ and the achieved exponent $\frac{t-2}{t-1}$; note that for $t = 3$, the triangle-free case, it happens that $\frac{t-1}{t+1} = \frac{t-2}{t-1}$.

For $\mathcal{F} = \{K_{s,t}\}$, where $K_{s,t}$ is the biclique with sides of size s and t, with $2 \leqslant s \leqslant t$,

$$\mu(\mathcal{F}) = 1 - \frac{s+t-2}{st-1} = \frac{st-s-t+1}{st-1} = \frac{(s-1)(t-1)}{st-1} = \frac{s-1}{s} \cdot \frac{t-1}{t-\frac{1}{s}}.$$

This nearly matches (in fact, asymptotically matches when t goes to infinity) the existing $\mathcal{O}_{s,t}(n^{\frac{s-1}{s}})$-approximation algorithms [2, 1].

For $\mathcal{F} = \{C_3, C_4, \dots, C_{\gamma-1}\}$ where $\gamma \geqslant 4$ and C_ℓ is the ℓ-vertex cycle, the $\mathcal{F}$-free graphs are precisely the graphs of girth at least γ, and

$$\mu(\mathcal{F}) = 1 - \frac{\gamma-3}{\gamma-2} = \frac{1}{\gamma-2}.$$

Thus, Theorem 2 improves the hardness exponent $\frac{1}{2\gamma-2} - \varepsilon$ shown in [1] for graphs of girth at least γ, but does not settle the approximability of MIS on these classes. Indeed, the exponent of the best approximation factor is $\frac{2}{\gamma-1}$ when γ is odd, and $\frac{2}{\gamma}$ when γ is even [5, 8]. We leave the closing of these gaps as open problems.

2 Preliminaries

We denote by $V(G)$ and $E(G)$ the vertex and edge sets, respectively, of a graph G. For every $X \subseteq V(G)$, $G[X]$ is the subgraph of G induced by X, i.e., the graph obtained from G after removing all the vertices not in X. We denote by $\alpha(G)$ the *independence number* of G, that is, the largest cardinality of an *independent set* of G.

Following the notation in [6, 3], we let $\mathcal{P}$ be a finite family of mutually independent random variables, and $\mathcal{A}$ be a finite family of events, informally referred to as "bad" events, determined by these variables. For any event B (not necessarily in $\mathcal{A}$) determined by $\mathcal{P}$, let $\mathrm{vbl}(B)$ be the smallest subset of $\mathcal{P}$ determining B, and define

$$\mathit{\Gamma}_{\mathcal{A}}(B) := \big\{A \in \mathcal{A} \setminus \{B\} \mid \mathrm{vbl}(A) \cap \mathrm{vbl}(B) \neq \varnothing\big\}.$$

When $\mathcal{A}$ is clear from the context, we simply write $\mathit{\Gamma}(B)$.

The Moser–Tardos resampling algorithm proceeds as follows. Initially sample all variables in $\mathcal{P}$ independently according to their respective distributions. While some event $A \in \mathcal{A}$ holds, resample all variables in $\mathrm{vbl}(A)$.

▶ **Theorem 3** ([6, Theorem 1.2]). *Let $\lambda \colon \mathcal{A} \to (0,1)$ be a map such that for every $A \in \mathcal{A}$,*

$$\Pr[A] \leqslant \lambda(A) \prod_{B \in \mathit{\Gamma}(A)} \big(1 - \lambda(B)\big). \tag{1}$$

Then the expected number of resampling steps of the Moser–Tardos algorithm is at most

$$\sum_{A \in \mathcal{A}} \frac{\lambda(A)}{1-\lambda(A)}.$$

We will also need to bound the probability that an event outside $\mathcal{A}$ holds at some point during the execution of the Moser–Tardos algorithm. Haeupler, Saha, and Srinivasan observe that the following can be derived from the proof of Moser and Tardos.

▶ **Theorem 4** ([3, Theorem 2.2]). *Assume $\mathcal{A}, \lambda$ satisfy the hypotheses of Theorem 3. Let B be any event determined by the variables in $\mathcal{P}$. The probability that B is true at least once during the execution of the Moser–Tardos algorithm is at most*

$$\Pr[B] \prod_{A \in \mathit{\Gamma}(B)} \big(1 - \lambda(A)\big)^{-1}. \tag{2}$$

A fortiori, the probability that B holds when the algorithm terminates is at most the quantity in (2). Finally, we will rely on the inapproximability of Max Independent Set in general graphs.

▶ **Theorem 5** (promise version of [9, Theorem 1.1] on complements)**.** *For every constant $\delta \in (0, 1/2)$, it is NP-hard to distinguish, given an n-vertex graph H, between the following two cases: $\alpha(H) \leqslant n^\delta$ and $\alpha(H) \geqslant n^{1-\delta}$.*

3 Construction of the triangle-free graph

The reduction is very simple. Given an n-vertex graph H, we build an n^2-vertex triangle-free graph G as follows. First replace every vertex $v \in V(H)$ by an independent set I_v of size $n \geqslant 1$. This yields the intermediate graph F where $xy \in E(F)$ if and only if $x \in I_u$ and $y \in I_v$ for some $uv \in E(H)$.

For every $e \in E(F)$, we introduce an independent Bernoulli variable

$$P_e \sim \text{Bernoulli}(p) \quad \text{with} \quad p := \frac{1}{10n},$$

i.e., $P_e = 1$ with probability p, and $P_e = 0$ otherwise. The family

$$\mathcal{P} := \{P_e \mid e \in E(F)\}$$

is our set of independent random variables (as used in Section 2).

For every triangle xyz in F, let A_{xyz} be the "bad" event that $P_{xy} = P_{yz} = P_{xz} = 1$. Set

$$\mathcal{A} := \{A_{xyz} \mid xy, yz, xz \in E(F)\}.$$

Let $F\langle\mathcal{P}\rangle$ be the graph with vertex set $V(F)$ and edge set $\{e \in E(F) \mid P_e = 1\}$.

We run the Moser–Tardos algorithm on $\mathcal{P}, \mathcal{A}$. We set $G := F\langle\mathcal{P}_{\text{out}}\rangle$, where $\mathcal{P}_{\text{out}}$ is the assignment of the variables in $\mathcal{P}$ when the algorithm terminates; see Figure 1. By design, we have:

▶ **Observation 6.** *G is triangle-free.*

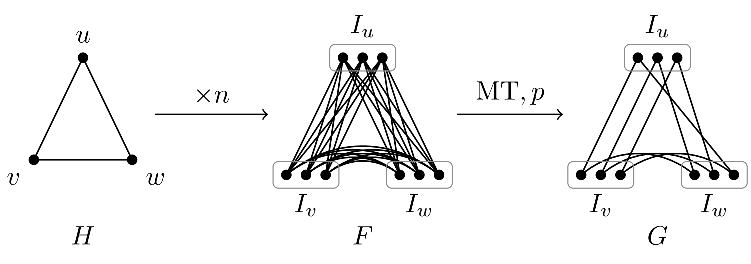

Figure 1 Construction of the triangle-free G from a graph H; here H is a triangle.

This randomized reduction does indeed terminate and take expected polynomial time.

▶ **Lemma 7.** *In expectation, it takes $\mathcal{O}(n^9)$ time to build G.*

Proof. One can build F from H in $\mathcal{O}(n^4)$ time. Set $\lambda := 2p^3 \in (0, 1)$ and, by a slight abuse of notation, also use λ for the constant map $A \mapsto 2p^3$ on $\mathcal{A}$. Let us check that the pair $\mathcal{A}, \lambda$ satisfies Equation (1), so that we can apply Theorem 3.

For every $A_{xyz} \in \mathcal{A}$, we have $|\Gamma(A_{xyz})| < d := 3n^2$. Indeed, there are fewer than n^2 triangles of F containing the edge xy (and likewise for yz and xz). By Bernoulli's inequality,

$$(1-\lambda)^d \geqslant 1 - d\lambda = 1 - 3n^2 \cdot 2p^3 = 1 - \frac{3}{500n} > \frac{1}{2}.$$

Therefore, for every $A \in \mathcal{A}$,

$$\lambda(1-\lambda)^d > \frac{\lambda}{2} = p^3 = \Pr[A],$$

and Equation (1) holds. By Theorem 3, if R is the number of resampling steps, then

$$\mathbb{E}[R] \leqslant s := \sum_{A \in \mathcal{A}} \frac{\lambda(A)}{1-\lambda(A)} = |\mathcal{A}| \cdot \frac{\lambda}{1-\lambda}.$$

We have $|\mathcal{A}| \leqslant n^6$ and $(1-\lambda)^{-1} < 2$, so

$$s < 2|\mathcal{A}|\lambda = 4|\mathcal{A}|p^3 \leqslant \frac{n^3}{250},$$

and hence the expected number of resamplings is $\mathcal{O}(n^3)$; since F has fewer than n^6 triangles and each search for a violated event can test them all in $\mathcal{O}(n^6)$ time, the total expected running time is $\mathcal{O}(n^9)$. ◀

By Markov's inequality, we get the following corollary.

▶ **Corollary 8.** *G is built in time $\mathcal{O}(n^{10})$ with high probability.*

4 Completeness and soundness

If $S \subseteq V(H)$ is an independent set of H, then $\bigcup_{v \in S} I_v$ is an independent set of F, hence of G. In particular, since $|\bigcup_{v \in S} I_v| = |S|n$, we have

▶ **Observation 9.** *$\alpha(G) \geqslant \alpha(H) \cdot n$.*

We now prove the following lemma.

▶ **Lemma 10.** *$\alpha(G) \leqslant \lceil 320\,\alpha(H) \cdot n \ln n \rceil$ with high probability.*

We fix a nonempty set $S \subseteq V(F)$. Our goal is to upper-bound the probability that S is an independent set in the output graph G produced when the Moser–Tardos algorithm terminates. We will then use a union bound to prove Lemma 10.

We set $s_v := |S \cap I_v|$ for every $v \in V(H)$. Therefore,

$$s := |S| = \sum_{v \in V(H)} s_v.$$

Let $m(S)$ denote the number of edges in $F[S]$. Thus,

$$m(S) = \sum_{uv \in E(H)} s_u s_v.$$

We finally set $z_v := s_v/s$.

We have $\sum_{v \in V(H)} z_v = \frac{1}{s} \sum_{v \in V(H)} s_v = 1$ and $z_v \geqslant 0$ for every $v \in V(H)$. Therefore, the Motzkin–Straus theorem applied to $\overline{H}, (z_v)_{v \in V(\overline{H})}$ yields the following inequality.

▶ **Lemma 11** ([7, Theorem 1] applied to the complement $\overline{H}$). $\sum_{uv \in E(\overline{H})} z_u z_v \leqslant \frac{1}{2}\left(1 - \frac{1}{\alpha(H)}\right)$.

We use this lemma to lower-bound the number of edges of $F[S]$ when S is relatively large.

▶ **Lemma 12.** *If* $s \geqslant 2n \cdot \alpha(H)$*, then* $m(S) \geqslant \frac{s^2}{4\alpha(H)}$.

Proof. Since $\sum_{v \in V(H)} z_v = 1$, we have

$$1 = \left(\sum_{v \in V(H)} z_v\right)^2 = \left(\sum_{v \in V(H)} z_v^2\right) + 2 \sum_{\{u,v\} \in \binom{V(H)}{2}} z_u z_v.$$

Together with Lemma 11, this identity yields

$$\sum_{uv \in E(H)} z_u z_v = \sum_{\{u,v\} \in \binom{V(H)}{2}} z_u z_v - \sum_{uv \in E(\overline{H})} z_u z_v \geqslant \frac{1}{2}\left(1 - \sum_{v \in V(H)} z_v^2\right) - \frac{1}{2}\left(1 - \frac{1}{\alpha(H)}\right).$$

Multiplying by s^2, we get

$$m(S) = \sum_{uv \in E(H)} s_u s_v \geqslant \frac{s^2}{2\alpha(H)} - \frac{1}{2} \sum_{v \in V(H)} s_v^2.$$

Since $s_v \leqslant n$ for every $v \in V(H)$,

$$m(S) \geqslant \frac{s^2}{2\alpha(H)} - \frac{n}{2} \sum_{v \in V(H)} s_v = \frac{s^2}{2\alpha(H)} - \frac{ns}{2}.$$

Thus, if $s \geqslant 2n \cdot \alpha(H)$, we have $\frac{ns}{2} \leqslant \frac{s^2}{4\alpha(H)}$, and we conclude that

$$m(S) \geqslant \frac{s^2}{2\alpha(H)} - \frac{s^2}{4\alpha(H)} = \frac{s^2}{4\alpha(H)}.$$ ◀

Let $\mathcal{I}_S$ be the event that $P_e = 0$ for every $e \in E(F[S])$. We have

$$\Pr[\mathcal{I}_S] = (1-p)^{m(S)}.$$

Let $\mathcal{J}_S$ be the event that $\mathcal{I}_S$ holds for the output assignment, equivalently, that S is an independent set in G.

▶ **Lemma 13.** *If* $s \geqslant 2n \cdot \alpha(H)$*, then* $\Pr[\mathcal{J}_S] \leqslant \exp\left(-\frac{s^2}{80n\,\alpha(H)}\right)$.

Proof. Again, set $\lambda := 2p^3$ and, by a slight abuse of notation, also use λ for the constant map $A \mapsto 2p^3$ on $\mathcal{A}$. We already checked in the proof of Lemma 7 that $\mathcal{A}$ and λ satisfy the hypotheses of Theorem 3. Thus, applying Theorem 4 to $\mathcal{I}_S$ and evaluating this event on the output assignment gives

$$\Pr[\mathcal{J}_S] \leqslant \Pr[\mathcal{I}_S] \cdot \prod_{A \in \Gamma(\mathcal{I}_S)} (1 - \lambda(A))^{-1} = \Pr[\mathcal{I}_S] \cdot (1-\lambda)^{-|\Gamma(\mathcal{I}_S)|}.$$

There are fewer than $m(S)n^2$ triangles of F with at least one edge in $F[S]$. Therefore, $|\Gamma(\mathcal{I}_S)| \leqslant m(S)n^2$. We thus obtain that

$$\Pr[\mathcal{J}_S] \leqslant (1-p)^{m(S)} \cdot (1-\lambda)^{-m(S)n^2}.$$

If $\Pr[\mathcal{J}_S] = 0$, the claimed bound is immediate. Otherwise, taking natural logarithms gives

$$\ln \Pr[\mathcal{J}_S] \leqslant m(S)\ln(1-p) - m(S)n^2 \ln(1-\lambda).$$

Since $p, \lambda \in (0, 1/2)$, we have $\ln(1-p) \leqslant -p$ and $-\ln(1-\lambda) \leqslant 2\lambda$, hence

$$\ln \Pr[\mathcal{J}_S] \leqslant -m(S)p + 2\lambda m(S)n^2 = -\frac{m(S)}{10n} + \frac{m(S)}{250n} \leqslant -\frac{m(S)}{20n}.$$

If $s \geqslant 2n \cdot \alpha(H)$, Lemma 12 implies that $m(S) \geqslant \frac{s^2}{4\alpha(H)}$, and thus

$$\Pr[\mathcal{J}_S] \leqslant \exp\left(-\frac{s^2}{80n\,\alpha(H)}\right). \quad ◀$$

We can now prove the main result of this section.

Proof of Lemma 10. We set the threshold $s_0 := \lceil 320\,\alpha(H)\,n \ln n \rceil$. Our goal is to show that

$$\Pr[\alpha(G) \leqslant s_0] = 1 - o(1).$$

This holds trivially if $s_0 > n^2$; we therefore assume that $s_0 \leqslant n^2$.

Note that $\Pr[\alpha(G) \geqslant s_0]$ is equal to the probability that there is at least one independent set of size s_0 in G. For all sufficiently large n, we have $s_0 \geqslant 2n\,\alpha(H)$, so a union bound and Lemma 13 give

$$\Pr[\alpha(G) \geqslant s_0] \leqslant \binom{n^2}{s_0} \exp\left(-\frac{{s_0}^2}{80n\,\alpha(H)}\right) \leqslant \exp\left(2s_0 \ln n - \frac{s_0}{80n\,\alpha(H)} \cdot s_0\right)$$

$$\leqslant \exp\left(2s_0 \ln n - (4\ln n)\cdot s_0\right) = \exp\left(-2s_0 \ln n\right) = o(1),$$

and we conclude. ◀

5 Wrapping up

We can now complete the proof of Theorem 1.

▶ **Theorem 1.** *For every constant $\varepsilon > 0$, there is no polynomial-time $N^{\frac{1}{2}-\varepsilon}$-approximation algorithm for* Max Independent Set *on N-vertex triangle-free graphs unless* NP $\subseteq$ BPP.

Proof. The claim is trivial for $\varepsilon > \frac{1}{2}$, so we may assume that $\varepsilon \in (0, \frac{1}{2}]$. Set $\delta := \frac{\varepsilon}{2} \in (0, \frac{1}{2})$.

By Theorem 5, it is NP-hard to distinguish n-vertex graphs H with $\alpha(H) \geqslant n^{1-\delta}$ from those with $\alpha(H) \leqslant n^{\delta}$. Given H, run the construction of Section 3 for at most n^{10} computation steps, where $N := n^2$. If the cutoff is reached, let G be the edgeless N-vertex graph; otherwise, let G be the output of the construction. By Lemma 7 and Markov's inequality, the cutoff is reached with probability $\mathcal{O}(n^{-1}) = o(1)$.

Thus, G is always triangle-free, is built in time $\mathcal{O}(n^{10}) = \mathcal{O}(N^5)$, and satisfies Observation 9. Suppose, for contradiction, that there is a polynomial-time $N^{\frac{1}{2}-\varepsilon}$-approximation algorithm for Max Independent Set on N-vertex triangle-free graphs. Let S be the independent set returned by this algorithm on G. A union bound on the cutoff event and the failure event in Lemma 10 gives, except with probability $o(1)$,

$$|S| \leqslant \alpha(G) \leqslant \lceil 320\,\alpha(H) n \ln n \rceil.$$

The approximation guarantee and Observation 9 give

$$|S| \geqslant \frac{\alpha(G)}{N^{\frac{1}{2}-\varepsilon}} \geqslant \alpha(H)n \cdot N^{\varepsilon-\frac{1}{2}} = \alpha(H)n^{2\varepsilon}.$$

For all sufficiently large n, if $\alpha(H) \leqslant n^{\delta}$, then, except with probability $o(1)$,

$$|S| \leqslant \lceil 320 n^{1+\delta} \ln n \rceil < n^{1+3\delta}.$$

If $\alpha(H) \geqslant n^{1-\delta}$, then

$$|S| \geqslant n^{1-\delta+2\varepsilon} = n^{1+3\delta}.$$

Hence, accepting exactly when $|S| \geqslant n^{1+3\delta}$ distinguishes the two cases with probability $1-o(1)$. The finitely many smaller inputs can be handled by brute force. Since the distinction is NP-hard, this yields NP $\subseteq$ BPP. ◀

The same result holds for randomized polynomial-time $N^{\frac{1}{2}-\varepsilon}$-approximation algorithms with success probability at least $\frac{2}{3}$, since $\mathcal{O}(\log n)$ independent repetitions, retaining the largest valid independent set returned, amplify the success probability to $1 - o(1)$.

6 Generalization to $\mathcal{F}$-free graphs

We recall that

$$\mu(\mathcal{F}) := 1 - \max_{K\in\mathcal{F}} \min_{\substack{U\subseteq V(K)\\ K[U] \text{ contains a cycle}}} \frac{|U|-2}{|E(K[U])|-1}.$$

Note that if every graph in $\mathcal{F}$ contains a cycle, then $\mu(\mathcal{F}) \in (0,1)$. The proof of the following theorem closely follows the triangle-free case.

▶ **Theorem 2.** *Let $\mathcal{F}$ be a nonempty finite family of graphs, each containing a cycle. For every constant $\varepsilon > 0$, there is no polynomial-time $N^{\mu(\mathcal{F})-\varepsilon}$-approximation algorithm for Max Independent Set on N-vertex $\mathcal{F}$-subgraph-free graphs unless* NP $\subseteq$ BPP*.*

Proof. Set $\mu := \mu(\mathcal{F})$. For every $K \in \mathcal{F}$, choose $U_K \subseteq V(K)$ attaining the inner minimum in the definition of μ, and set

$$J_K := K[U_K], \qquad v_K := |V(J_K)|, \qquad e_K := |E(J_K)|.$$

Thus,

$$\frac{v_K - 2}{e_K - 1} \leqslant 1 - \mu. \tag{3}$$

Let H be an n-vertex graph. Put

$$\ell := \left\lceil n^{\frac{1-\mu}{\mu}} \right\rceil \qquad \text{and} \qquad N := n\ell.$$

As in Section 3, replace every vertex $v \in V(H)$ by an independent set I_v of size ℓ, and denote the resulting complete blow-up by F. For every edge $e \in E(F)$, introduce an independent variable $P_e \sim \text{Bernoulli}(p)$, where

$$p := cN^{-(1-\mu)} \qquad \text{and} \qquad c := \frac{1}{\sqrt{16 e_*^2 |\mathcal{F}|}},$$

with $e_* := \max_{K \in \mathcal{F}} e_K$. Note that $p, c \in (0,1)$, and c depends only on $\mathcal{F}$.

For every $K \in \mathcal{F}$ and every copy Q of J_K in F, not necessarily induced, let A_Q be the event that $P_e = 1$ for every $e \in E(Q)$. Let $\mathcal{A}$ be the family of all these bad events, and define

$$\lambda(A_Q) := 2p^{e_K} \in (0,1)$$

whenever Q is a copy of J_K. A fixed edge of F belongs to at most $2e_K N^{v_K-2}$ copies of J_K. Hence, for every bad event A_Q,

$$\sum_{A \in \Gamma(A_Q)} \lambda(A) \leqslant e_* \sum_{K \in \mathcal{F}} 2e_K N^{v_K-2} \cdot 2p^{e_K} \leqslant 4e_*^2 \sum_{K \in \mathcal{F}} N^{v_K-2} p^{e_K} \leqslant 4e_*^2 \sum_{K \in \mathcal{F}} N^{v_K-2} p^{e_K-1}.$$

Moreover,

$$N^{v_K-2} p^{e_K-1} = c^{e_K-1} N^{v_K-2-(1-\mu)(e_K-1)} \leqslant c^{e_K-1} \leqslant c^2, \tag{4}$$

where the first inequality follows from (3), and the second follows from $c \leqslant 1$ and $e_K \geqslant 3$. Therefore,

$$\sum_{A \in \Gamma(A_Q)} \lambda(A) \leqslant 4e_*^2 |\mathcal{F}| c^2 \leqslant \frac{1}{2}.$$

This justifies the second inequality in the following statement, while the first inequality holds by the Weierstrass product inequality:

$$\lambda(A_Q) \prod_{A \in \Gamma(A_Q)} (1 - \lambda(A)) \geqslant 2p^{e_K} \left(1 - \sum_{A \in \Gamma(A_Q)} \lambda(A)\right) \geqslant p^{e_K} = \Pr[A_Q].$$

The hypotheses of Theorem 3 are therefore satisfied. Run the Moser–Tardos algorithm with $(P_e)_{e \in E(F)}, \mathcal{A}$ and let G be the graph defined by its output assignment. (Technically, p should be chosen as a rational number with a polynomial-length representation, sufficiently close to the displayed value that all the inequalities of this proof remain valid.) If R denotes the number of resampling steps, then Theorem 3 and the bound N^{v_K} on the number of copies of J_K give

$$\mathbb{E}[R] \leqslant \sum_{A \in \mathcal{A}} \frac{\lambda(A)}{1 - \lambda(A)} \leqslant 2 \cdot 2p^3 \sum_{K \in \mathcal{F}} N^{v_K} = N^{\mathcal{O}_{\mathcal{F}}(1)},$$

where the second inequality holds since $\lambda(A) \leqslant 2p^3$ and $(1 - \lambda(A))^{-1} \leqslant 2$. All bad events can be enumerated and tested in $N^{\mathcal{O}_{\mathcal{F}}(1)}$ time, so the expected running time is polynomial in N as well. Moreover, G contains no J_K as a subgraph for any $K \in \mathcal{F}$. In particular, G is $\mathcal{F}$-subgraph-free, since every copy of K would contain a copy of $J_K = K[U_K]$.

Completeness is unchanged from the triangle-free case:

$$\alpha(G) \geqslant \ell\alpha(H). \tag{5}$$

We next establish the corresponding soundness bound. Fix a nonempty set $S \subseteq V(F)$, set $s := |S|$, and let $m(S) := |E(F[S])|$. The same calculation as in Lemma 12 (with ℓ now replacing n as the size of each I_v) gives

$$m(S) \geqslant \frac{s^2}{2\alpha(H)} - \frac{\ell s}{2} = \frac{s(s - \alpha(H)\ell)}{2\alpha(H)}.$$

Thus, if $s \geqslant 2\alpha(H)\ell$, then $s - \alpha(H)\ell \geqslant \frac{s}{2}$, so

$$m(S) \geqslant \frac{s^2}{4\alpha(H)}. \tag{6}$$

Let $\mathcal{I}_S$ be the event that $P_e = 0$ for every $e \in E(F[S])$, and let $\mathcal{J}_S$ be the event that S is an independent set in G. For every fixed edge $e \in E(F)$, the copy count above gives

$$\sum_{\substack{A_Q \in \mathcal{A} \\ \text{with } e \in E(Q)}} \lambda(A_Q) \leqslant \sum_{K \in \mathcal{F}} 2e_K N^{v_K - 2} \cdot 2p^{e_K} \leqslant 4e_* p \sum_{K \in \mathcal{F}} N^{v_K-2} p^{e_K - 1} \leqslant p \cdot 4e_* |\mathcal{F}| c^2 \leqslant \frac{p}{4},$$

where the third inequality holds by (4). Every event in $\Gamma(\mathcal{I}_S)$ depends on some variable P_e with $e \in E(F[S])$, and therefore

$$\sum_{A \in \Gamma(\mathcal{I}_S)} \lambda(A) \leqslant \sum_{e \in E(F[S])} \sum_{\substack{A_Q \in \mathcal{A} \\ \text{with } e \in E(Q)}} \lambda(A_Q) \leqslant \frac{p\, m(S)}{4}. \tag{7}$$

Moreover, $\lambda(A) < 1/2$ for every $A \in \mathcal{A}$. We apply Theorem 4 and get

$$\Pr[\mathcal{J}_S] \leqslant \Pr[\mathcal{I}_S] \prod_{A \in \Gamma(\mathcal{I}_S)} (1 - \lambda(A))^{-1} = (1-p)^{m(S)} \prod_{A \in \Gamma(\mathcal{I}_S)} (1 - \lambda(A))^{-1}$$

Since $1 - p \leqslant \exp(-p)$, and $-\ln(1-x) \leqslant 2x$ for every $x \in (0, 1/2)$, we have

$$\Pr[\mathcal{J}_S] \leqslant \exp\left(-pm(S) - \sum_{A \in \Gamma(\mathcal{I}_S)} \ln(1 - \lambda(A)) \right) \leqslant \exp\left(-pm(S) + 2 \sum_{A \in \Gamma(\mathcal{I}_S)} \lambda(A) \right)$$

Therefore, by (7),

$$\Pr[\mathcal{J}_S] \leqslant \exp\left(-\frac{p\, m(S)}{2} \right).$$

Together with (6), this shows that, whenever $s \geqslant 2\alpha(H)\ell$,

$$\Pr[\mathcal{J}_S] \leqslant \exp\left(-\frac{ps^2}{8\alpha(H)} \right). \tag{8}$$

Set $s_0 := \lceil 32\, \alpha(H) p^{-1} \ln N \rceil$. If $s_0 > N$, then $\alpha(G) < s_0$ holds trivially, so assume that $s_0 \leqslant N$.

Since $N = \Theta\left(n^{\frac{1}{\mu}}\right)$ and $p^{-1} = \Theta\left(N^{1-\mu}\right) = \Theta(\ell)$, we have $s_0 \geqslant 2\alpha(H)\ell$ for all sufficiently large n. A union bound using (8) gives

$$\Pr[\alpha(G) \geqslant s_0] \leqslant \binom{N}{s_0} \exp\left(-\frac{p{s_0}^2}{8\alpha(H)} \right) \leqslant \exp\left(s_0 \ln N - 4 s_0 \ln N \right) \leqslant \exp(-3 s_0 \ln N) = o(1).$$

Therefore, with high probability,

$$\alpha(G) < s_0 = \mathcal{O}(\alpha(H)\ell \ln N). \tag{9}$$

It remains to turn this construction into a worst-case polynomial-time reduction and compare the gap. Fix a constant q such that the expected running time of the construction is $\mathcal{O}(N^q)$, and truncate it after N^{q+1} computation steps. If the cutoff is reached, output the edgeless N-vertex graph. By Markov's inequality, the cutoff is reached with probability $\mathcal{O}(N^{-1}) = o(1)$.

Since every graph in $\mathcal{F}$ contains a cycle, the resulting graph is $\mathcal{F}$-subgraph-free whether or not the cutoff is reached, and (5) always holds. Set $\delta := \min\{1/4, \varepsilon/(4\mu)\} \in (0, 1/2)$. By Theorem 5, it is NP-hard to distinguish n-vertex graphs H with $\alpha(H) \geqslant n^{1-\delta}$ from those with $\alpha(H) \leqslant n^{\delta}$. Equations (5) and (9) yield the gap

$$\Omega\left(\frac{n^{1-2\delta}}{\ln N}\right) = N^{\mu(1-2\delta)-o(1)}.$$

Since $\mu(1-2\delta) \geqslant \mu - \varepsilon/2 > \mu - \varepsilon$, this is larger than $N^{\mu-\varepsilon}$ for all sufficiently large N. Such an approximation algorithm would therefore distinguish the two cases with probability $1 - o(1)$, and would imply that NP $\subseteq$ BPP. ◀

AI disclosure. Theorem 1 was proven essentially autonomously by GPT-5.6 Sol Pro. The author then suggested the generalization Theorem 2 and the definition of $\mu(\mathcal{F})$. The write-up is due to the author, who tried to make the proof of Theorem 1 a pleasant and effortless read.

References

1 Édouard Bonnet, Stéphan Thomassé, Xuan Thang Tran, and Rémi Watrigant. An algorithmic weakening of the Erdős–Hajnal conjecture. In *28th Annual European Symposium on Algorithms (ESA 2020)*, volume 173 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 23:1–23:18. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2020. `doi:10.4230/LIPIcs.ESA.2020.23`.

2 Pavel Dvořák, Andreas Emil Feldmann, Ashutosh Rai, and Paweł Rzążewski. Parameterized inapproximability of independent set in H-free graphs. *Algorithmica*, 85(4):902–928, 2023. `doi:10.1007/s00453-022-01052-5`.

3 Bernhard Haeupler, Barna Saha, and Aravind Srinivasan. New constructive aspects of the Lovász local lemma. *Journal of the ACM*, 58(6):28:1–28:28, 2011. `doi:10.1145/2049697.2049702`.

4 Johan Håstad. Clique is hard to approximate within $n^{1-\varepsilon}$. In *37th Annual Symposium on Foundations of Computer Science, FOCS 1996, Burlington, Vermont, USA, 14-16 October, 1996*, pages 627–636. IEEE Computer Society, 1996. `doi:10.1109/SFCS.1996.548522`.

5 Burkhard Monien and Ewald Speckenmeyer. Ramsey numbers and an approximation algorithm for the vertex cover problem. *Acta Informatica*, 22(1):115–123, 1985. `doi:10.1007/BF00290149`.

6 Robin A. Moser and Gábor Tardos. A constructive proof of the general Lovász local lemma. *Journal of the ACM*, 57(2):11:1–11:15, 2010. `doi:10.1145/1667053.1667060`.

7 Theodore S. Motzkin and Ernst G. Straus. Maxima for graphs and a new proof of a theorem of Turán. *Canadian Journal of Mathematics*, 17:533–540, 1965. `doi:10.4153/CJM-1965-053-6`.

8 Owen J. Murphy. Computing independent sets in graphs with large girth. *Discrete Applied Mathematics*, 35(2):167–170, 1992. `doi:10.1016/0166-218X(92)90041-8`.

9 David Zuckerman. Linear degree extractors and the inapproximability of Max Clique and Chromatic Number. *Theory of Computing*, 3(6):103–128, 2007. `doi:10.4086/toc.2007.v003a006`.