Revisiting Jerrum's Metropolis Process for the Planted Clique Problem

Elchanan Mossel (MIT), joint work with **Zongchen Chen** (MIT) and **Ilias Zadik** (MIT)

Jun 6, 2023

Oxford Probability Seminar

Supported by Grants:

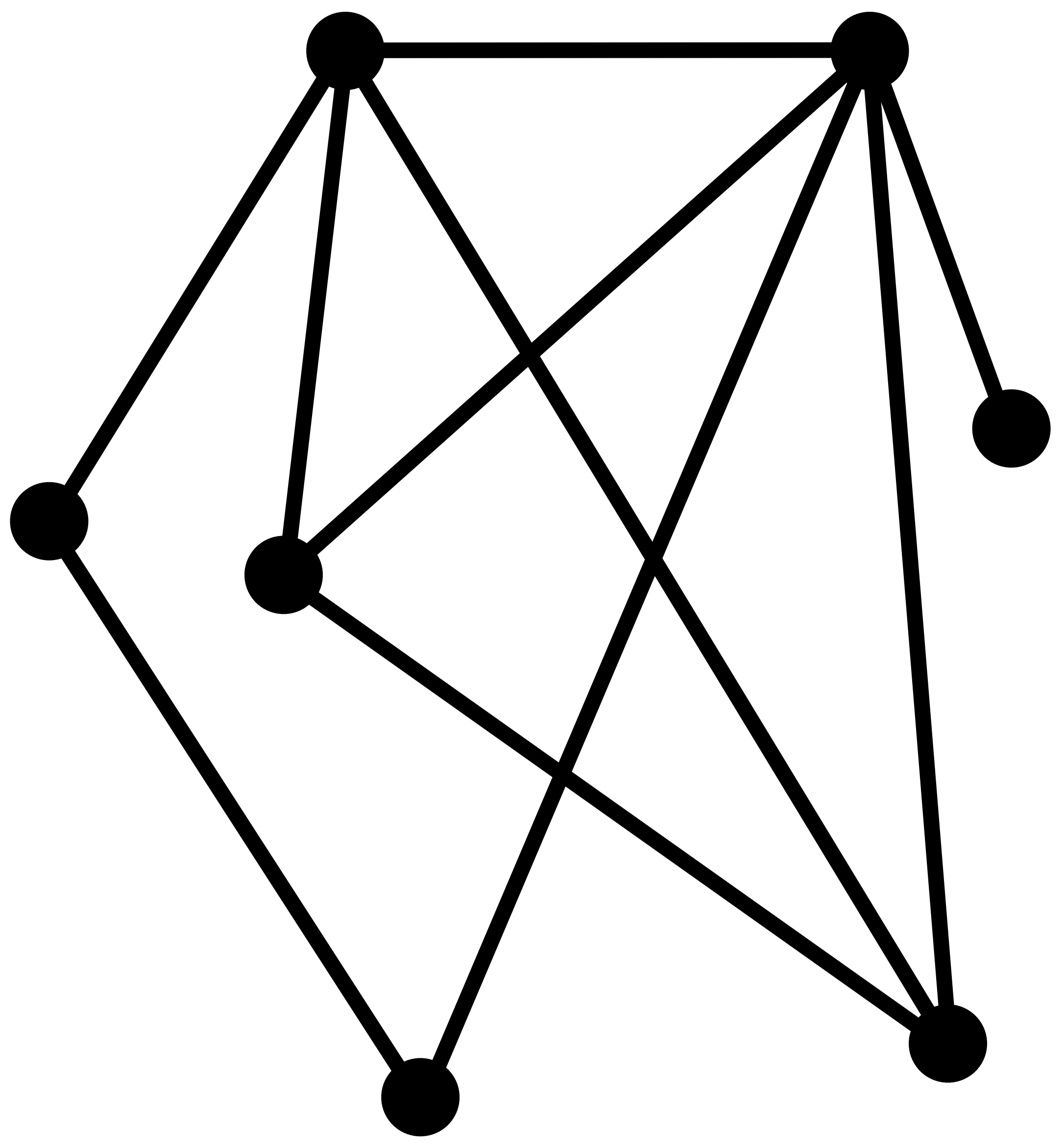
Simons-NSF grant DMS-2031883

Vannevar Bush Faculty Fellowship ONR-N00014-20-1-2826

Simons Investigator award (622132)

Cliques in Random Graphs

- Erdős–Rényi random graph $\mathcal{G}(n,1/2)$
  - n vertices, every pair connected with prob $1/2$ independently
- Max clique of $\mathcal{G}(n,1/2)$ has size $\approx 2\log n$ w.h.p.
- Best known algorithm finds a clique of size $\approx \log n$ w.h.p.



Q: Can we find a $(1+\varepsilon)\log n$ clique in $\mathcal{G}(n,1/2)$ efficiently?

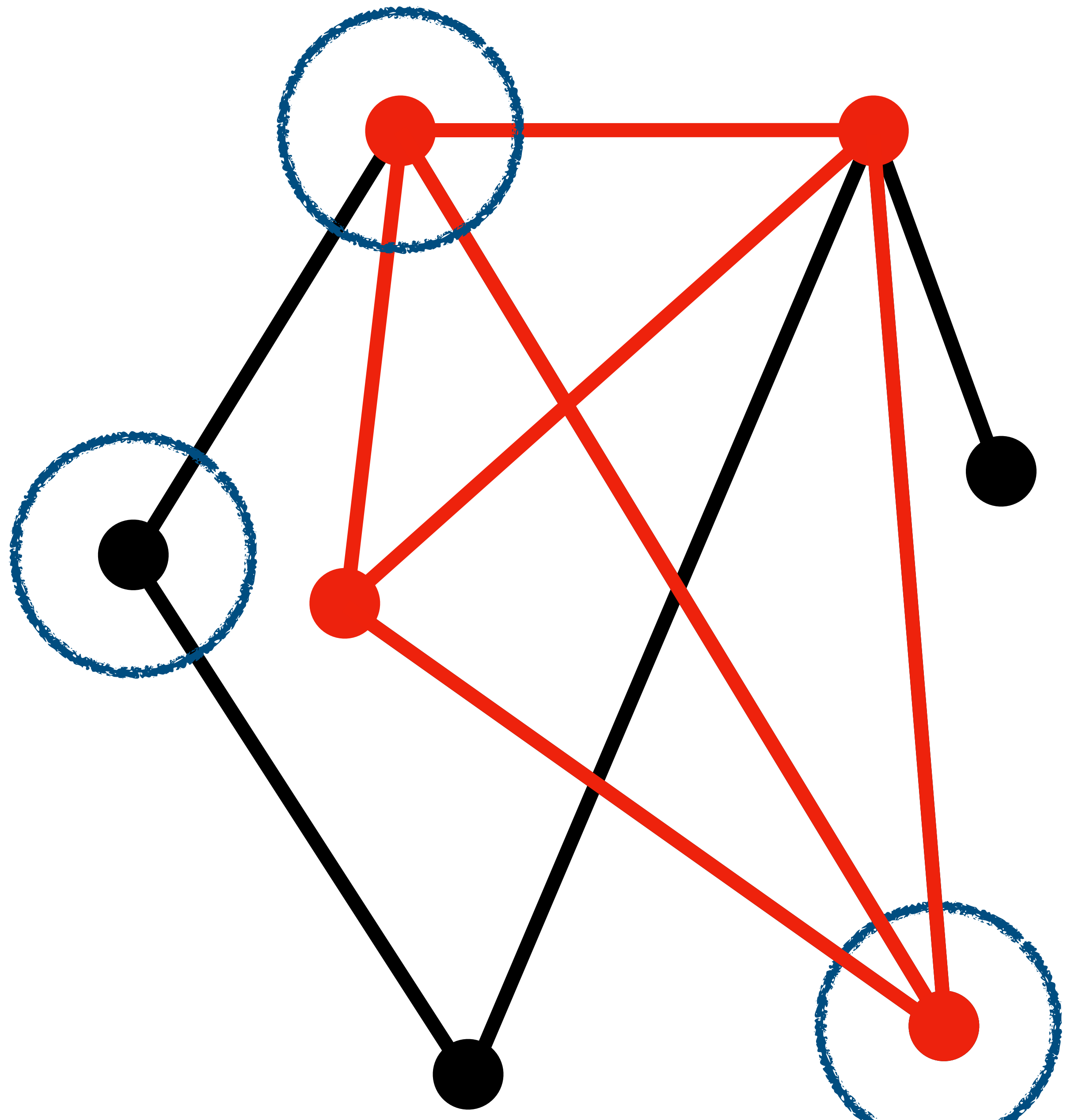
- Can do this in $n^{\Theta(\log n)}$ time by exhaustive search

Metropolis Process

- [Jerrum'92] considered the Metropolis Process (MP) for finding a $(1+\varepsilon)\log n$ clique in $\mathcal{G}(n,1/2)$

> ▸ Initialization: a clique X_0
>
> ▸ At time t, generate X_t from X_{t-1} as follows:
>
> - Pick a vertex v uniformly at random:
> - If $v \notin X_{t-1}$, let $X_t = X_{t-1} \cup \{v\}$ if it is a clique, and $X_t = X_{t-1}$ otherwise
> - If $v \in X_{t-1}$, let $X_t = \begin{cases} X_{t-1} \backslash \{v\}, & \text{w.p. } e^{-\beta} \\ X_{t-1}, & \text{w.p. } 1 - e^{-\beta} \end{cases}$ $\quad \beta \geq 0$

Metropolis Process: Example



$X_{t-1} \cup \{v\}$ is a clique, so $X_t = X_{t-1} \cup \{v\}$

$X_{t-1} \cup \{v\}$ is not a clique, so $X_t = X_{t-1}$

$$X_t = \begin{cases} X_{t-1} \setminus \{v\}, & \text{w.p. } e^{-\beta} \\ X_{t-1}, & \text{w.p. } 1 - e^{-\beta} \end{cases}$$

Metropolis Process for $\mathcal{G}(n,1/2)$

- As t grows, the distribution of X_t converges to stationary distribution π

$$\pi(C) \propto e^{\beta|C|}, \forall \text{ clique } C$$

  - $\beta = 0$: π is uniform distribution over all cliques
  - $\beta = \Theta(1)$: $C \sim \pi$ has size $\approx \log n$ w.h.p.
  - $\beta = \Theta(\log n)$: $C \sim \pi$ has size $\approx (1+\varepsilon)\log n$ w.h.p.

$C \sim \pi$: a random clique drawn from π

Hope: X_t converges to π quickly (poly-time), and we get a $(1+\varepsilon)\log n$ clique!

[Jerrum'92]: For any $\beta \geq 0$, MP fails to find a $(1+\varepsilon)\log n$ clique in $\mathcal{G}(n,1/2)$, *even if we "plant" a large clique of size* $k = n^{\alpha}$, $\alpha < 1/2$ 🙁

Planted Clique Model

- Planted clique model $\mathscr{G}(n, 1/2, k)$ [Jerrum'92, Kučera'95]

Step 1: G_0 is an Erdős–Rényi $\mathscr{G}(n, 1/2)$

Step 2: Pick a subset of k vertices u.a.r. and form a planted k-clique $\mathscr{PC}$

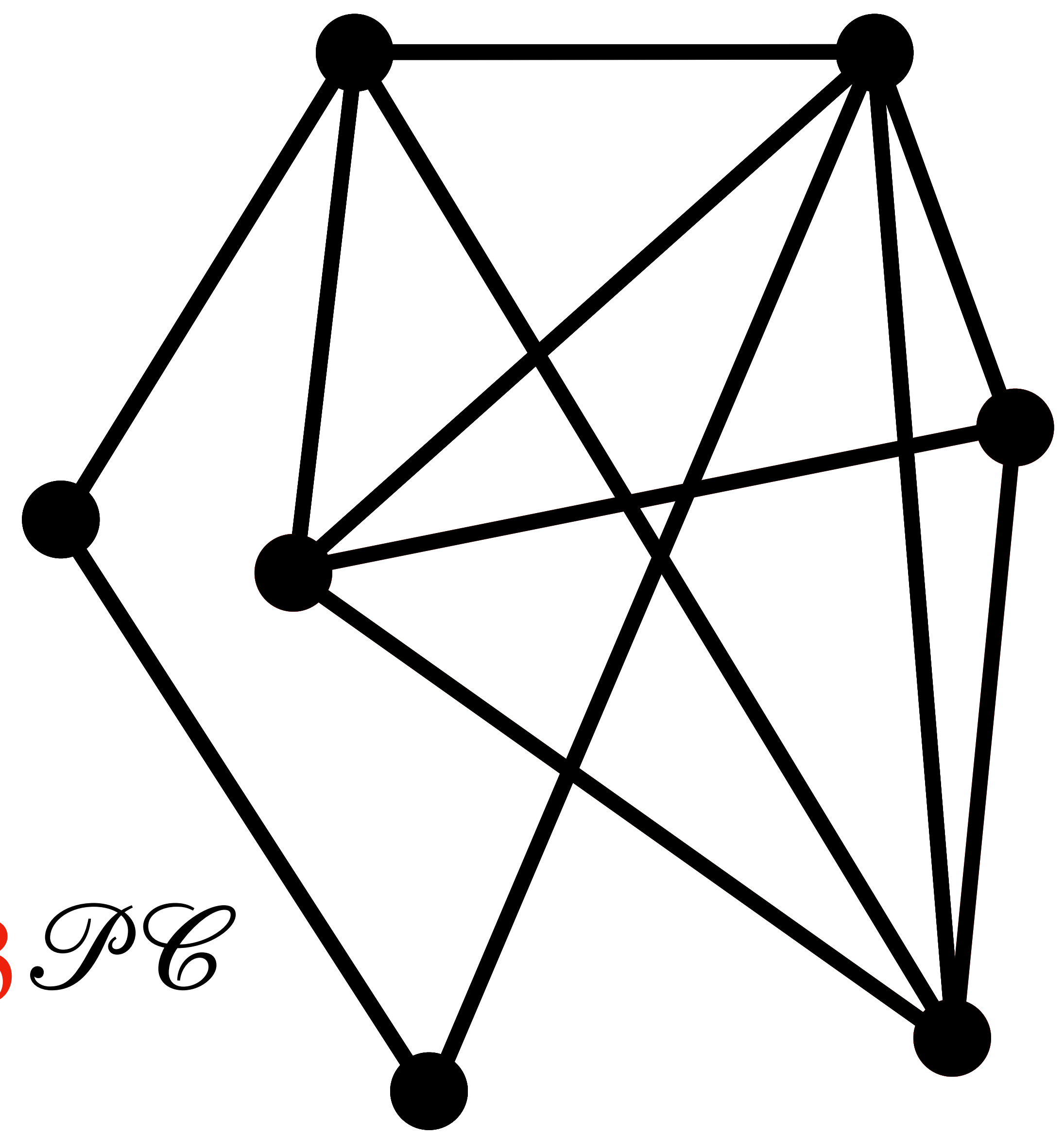
➡ $G = G_0 \cup \mathscr{PC}$

Goal: Recover $\mathscr{PC}$ from observing $G \sim \mathscr{G}(n, 1/2, k)$

Q: How large does k need to be?

to (efficiently) find the clique?

$G = G_0 \cup \mathscr{PC}$

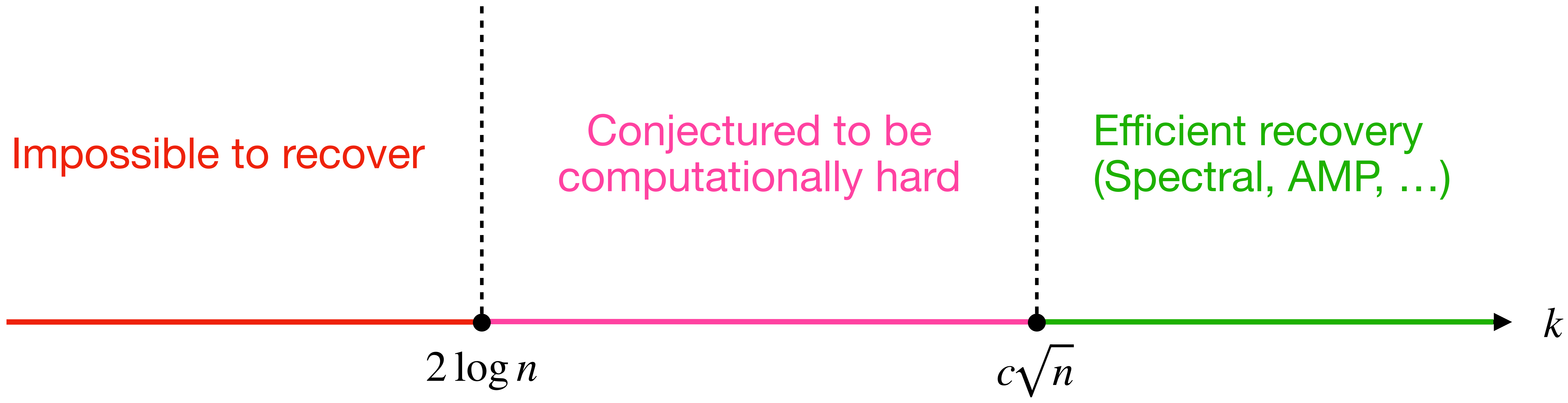


Recovering Planted Clique

Goal: Recover $\mathscr{PC}$ from observing $G \sim \mathscr{G}(n,1/2,k)$

- $k \geq (2+\varepsilon)\log n$: $n^{\Theta(\log n)}$ time by exhaustive search
- $k = \Omega(\sqrt{n \log n})$: $\mathrm{poly}(n)$ time by degree counting
- $k = \Omega(\sqrt{n})$: $\mathrm{poly}(n)$ time
  - Spectral method, approximate message passing, and more… [AKS'98, FR'10, DM'13, DGGP'14]
  - If $k = o(\sqrt{n})$, many algorithms fail: MP [Jer'92], Sum-of-Squares hierarchy [BHK+'16], statistical-query algorithms [FGR+'17], …

The Planted Clique Conjecture



- Computational hardness implies same for other important problems: compressed sensing, sparse PCA, property testing, cryptography…

MP for Planted Clique Model

- Suppose $k = |\mathscr{PC}| = n^{\alpha}$ where $0 < \alpha < 1$
- X_t converges to π, where $\pi(C) \propto e^{\beta|C|}$, $\forall$ clique C
- $C \sim \pi$ is contained in $\mathscr{PC}$ w.h.p.
  - # Cliques inside $\mathscr{PC}$ $= 2^k = 2^{n^{\alpha}} \gg n^{\Theta(\log n)} =$ # Cliques outside $\mathscr{PC}$

Hope: X_t converges to π in poly time, and we see a significant portion of $\mathscr{PC}$!

[Jerrum'92]: For any $\alpha < 1/2$ and $\beta \geq 0$, MP requires $n^{\Theta(\log n)}$ time to find a $(1+\varepsilon)\log n$ clique under worst-case initialization X_0 🙁

Revisiting Jerrum's Result

[Jerrum'92]: For any $\alpha < 1/2$ and $\beta \geq 0$, MP requires $n^{\Theta(\log n)}$ time to find a $(1+\varepsilon)\log n$ clique under worst-case initialization X_0 🙁

(a) Why $\alpha < 1/2$? Does MP work when $1/2 \leq \alpha < 1$?

- First evidence of "hardness" for planted clique problem when $k = o(\sqrt{n})$ is commonly attributed to the failure of MP in [Jerrum'92]

(b) Why $(1+\varepsilon)\log n$ clique?

- Can we first find $\gamma \log n$ vertices from $\mathscr{PC}$, and then recover $\mathscr{PC}$ easily?

(c) Why worst-case initialization? Same is true for many lower bounds of MCMC

- Can we use simple and nature "empty clique" initialization $X_0 = \emptyset$?

Our Results

$k = |\mathscr{PC}| = n^{\alpha}$

[Chen-Mossel-Zadik'23]: For any $\alpha < 1$, MP requires $n^{\omega(1)}$ time to reach:

- Either a clique of size $(1 + \varepsilon)\log n$
- Or a clique of intersection $\gamma \log n$ with $\mathscr{PC}$

When (i) under worst-case initialization and $\beta \geq 0$

(ii) under empty clique initialization and $\beta = o(\log n)$ or $\omega(\log n)$

- Big failure of MP for the planted clique problem
- Contrary to common sense predictions: no strong evidence of hardness

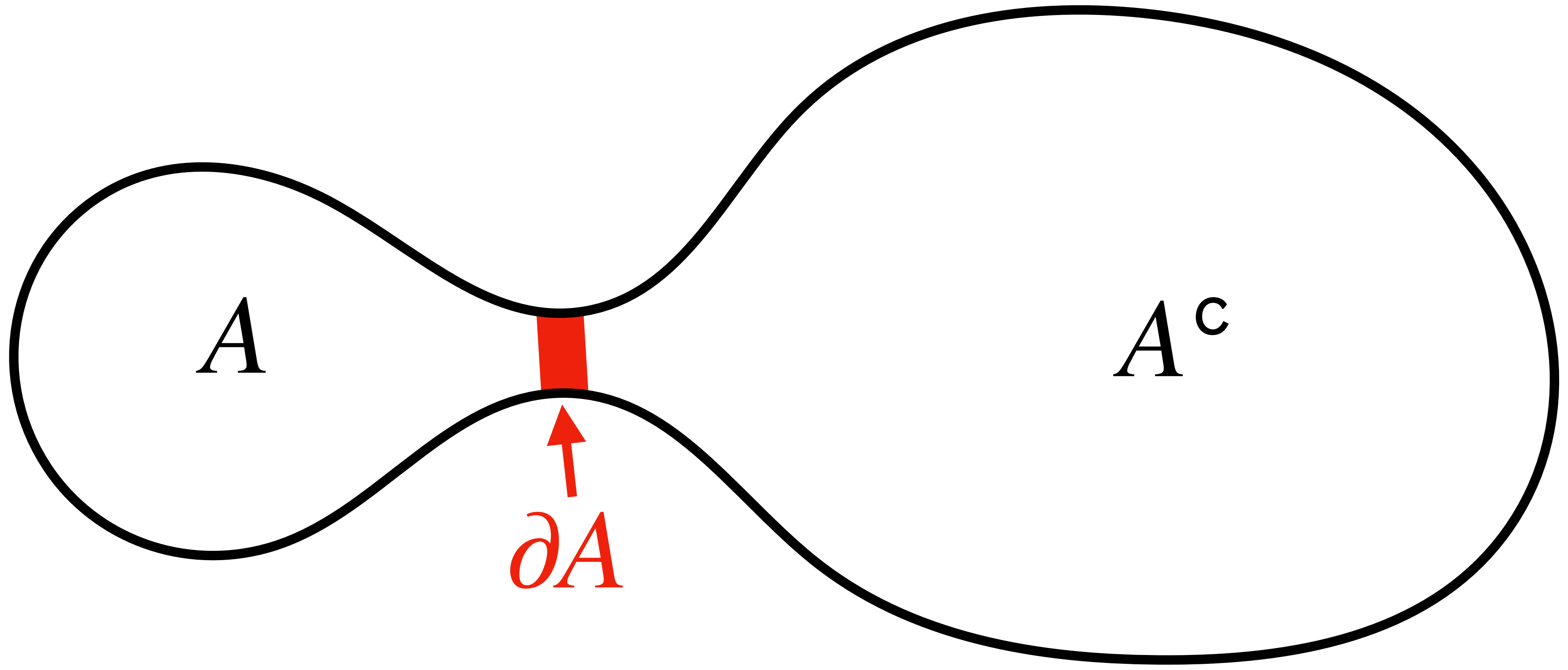
Proof Approach: Worst-case Initialization

[Chen-Mossel-Zadik'23]: For any $\alpha < 1$ and $\beta \geq 0$, MP requires $n^{\Omega(\log n)}$ time to reach:

- Either a clique of size $(1+\varepsilon)\log n$
- Or a clique of intersection $\gamma \log n$ with $\mathscr{PC}$

under worst-case initialization

"Bottleneck argument": If $\pi(\partial A)/\pi(A) = n^{-\Omega(\log n)}$, then MP requires $n^{\Omega(\log n)}$ time to escape A (reach A^{c}) when started from $X_0 \sim \pi(\,\cdot \mid A)$

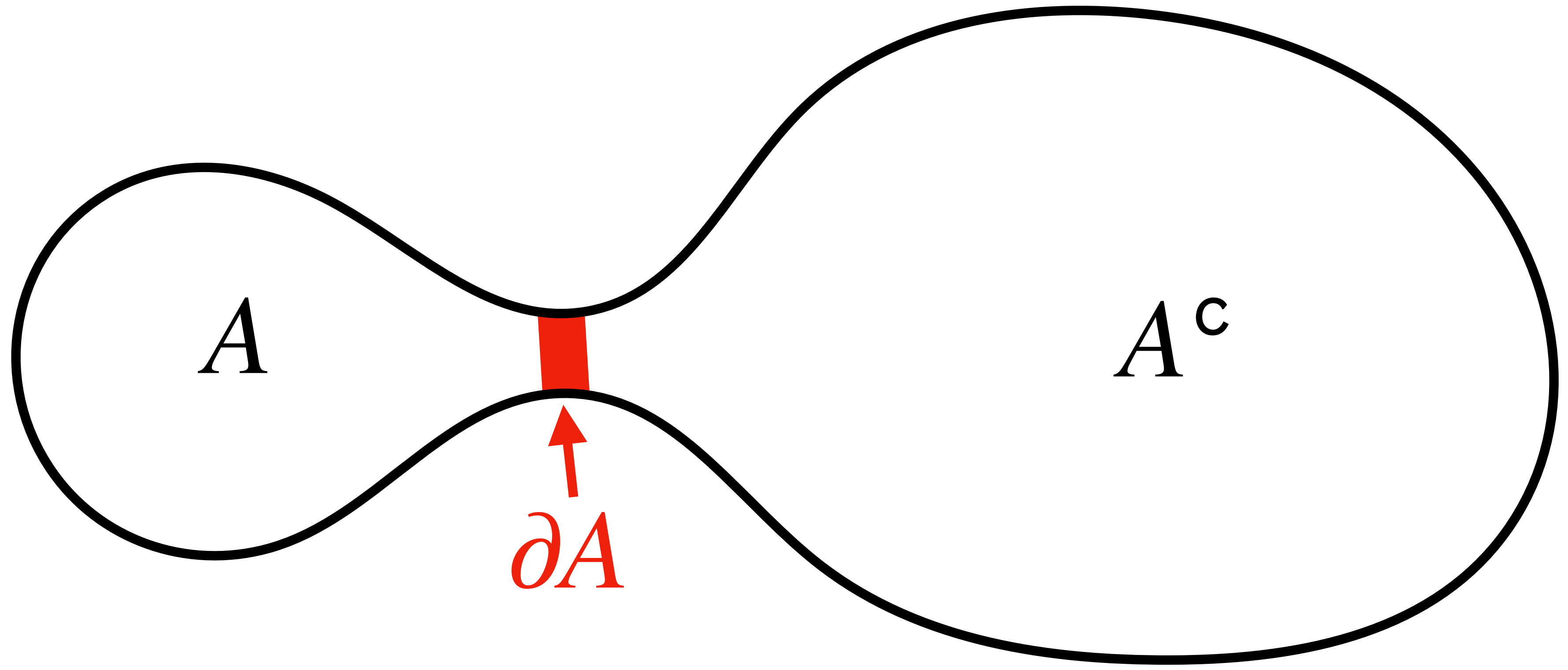


A: a subset of cliques
∂A: boundary cliques of A

Bottleneck for Large Intersection

- $A = \{C : |C \cap \mathcal{PC}| \leq \gamma \log n\}$
- $\partial A = \{C : |C \cap \mathcal{PC}| = \gamma \log n\}$
- Can show $\frac{|\partial A|}{|A|} \approx \frac{\mathbb{E}|\partial A|}{\mathbb{E}|A|} = n^{-\Omega(\log n)}$ w.h.p.
  - $\beta = 0$: $\frac{\pi(\partial A)}{\pi(A)} = \frac{|\partial A|}{|A|}$
  - For general β: $\frac{\pi(\partial A)}{\pi(A)} \approx \frac{\max_q e^{\beta q} |\mathcal{C}_{q,\gamma \log n}|}{\max_q e^{\beta q} |\mathcal{C}_q|} \leq \frac{|\mathcal{C}_{q^*,\gamma \log n}|}{|\mathcal{C}_{q^*}|} \approx n^{-\Omega(\log n)}$

"Bottleneck argument": If $\pi(\partial A)/\pi(A) = n^{-\Omega(\log n)}$, then MP requires $n^{\Omega(\log n)}$ time to escape A (reach A^{c}) when started from $X_0 \sim \pi(\cdot \mid A)$



$\mathcal{C}_q = \{C : |C| = q\}$ $\mathcal{C}_{q,r} = \{C : |C| = q, |C \cap \mathcal{PC}| = r\}$

Bottleneck for Large Size

- Jerrum's bottleneck ∂B: cliques of size $(1+2\varepsilon/3)\log n$ expandable to size $(1+\varepsilon)\log n$
- Work only when $\alpha < 1/2$
- When $1/2 \leq \alpha < 1$, $\dfrac{\pi(\partial B)}{\pi(B)}$ is large since B and ∂B are mostly cliques contained in $\mathscr{P}\mathscr{C}$
- Need to take "combined bottleneck" $\approx A \cap B$ (A is previous bottleneck for large intersection)

Proof Approach: Empty Clique Initialization

- $\beta = \omega(\log n)$: Probability of removing a vertex $= e^{-\beta} = n^{-\omega(1)}$
  - MP $\approx$ Randomized Greedy algorithm (pick a random vertex and add if possible)
- $\beta = o(\log n)$: Consider the "projected process" over $\mathbb{N}^+$ for sizes of cliques: $0 = |X_0| \to |X_1| \to |X_2| \to \cdots$ (MP: $\emptyset = X_0 \to X_1 \to X_2 \to \cdots$)
- Use an auxiliary birth and death process $\{Y_t\}_t$ to bound $\{|X_t|\}_t$
- Can show large hitting time when $\beta = o(\log n)$

Conclusion and Future Problems

[Chen-Mossel-Zadik'23]: For any $\alpha < 1$, MP requires $n^{\omega(1)}$ time to reach:

- Either a clique of size $(1+\varepsilon)\log n$
- Or a clique of intersection $\gamma \log n$ with $\mathscr{PC}$

When (i) under worst-case initialization and $\beta \geq 0$

(ii) under empty clique initialization and $\beta = o(\log n)$ or $\omega(\log n)$

- Failure of MP under empty clique initialization and for $\beta = \Theta(\log n)$?
- General tools for analyzing MCMC algorithms under natural initialization?
- Efficient MCMC algorithms for recovering the planted clique?

Thank you!