Private Edge Density Estimation for Random Graphs: Optimal, Efficient and Robust

Hongjie Chen
ETH Zürich

Jingqiu Ding
ETH Zürich

Yiding Hua
ETH Zürich

David Steurer
ETH Zürich

Abstract

We give the first polynomial-time, differentially node-private, and robust algorithm for estimating the edge density of Erdős-Rényi random graphs and their generalization, inhomogeneous random graphs. We further prove information-theoretical lower bounds, showing that the error rate of our algorithm is optimal up to logarithmic factors. Previous algorithms incur either exponential running time or suboptimal error rates.

Two key ingredients of our algorithm are (1) a new sum-of-squares algorithm for robust edge density estimation, and (2) the reduction from privacy to robustness based on sum-of-squares exponential mechanisms due to Hopkins et al. (STOC 2023).

1 Introduction

Privacy has nowadays become a major concern in large-scale data processing. Releasing seemingly harmless statistics of a dataset could unexpectedly leak sensitive information of individuals (see e.g. [NS09, DSSU17] for privacy attacks). Differential privacy (DP) [DMNS06] has emerged as a by-now standard technique for protecting the privacy of individuals with rigorous guarantees. An algorithm is said to be differentially private if the distribution of its output remains largely unchanged under the change of a single data point in the dataset.

For datasets represented by graphs (e.g. social networks), two notions of differential privacy have been investigated in the literature: edge differential privacy [NRS07, KRSY11], where each edge is regarded as a data point; and node differential privacy [BBDS13, KNRS13], where each node along with its incident edges is regarded as a data point. Node differential privacy is an arguably more desirable notion than edge differential privacy. On the other hand, node differential privacy is also in general more difficult to achieve without compromising on utility, as many graph statistics usually have high sensitivity in the worst case. It turns out that many graph statistics can have significantly smaller sensitivity on typical graphs under natural distributional assumptions. Several recent works could thus manage to achieve optimal or nearly-optimal utility guarantees in a number of random graph parameter estimation problems [BCS15, BCSZ18, SU19, CDd+24].

In this paper, we continue this line of work and study perhaps the most elementary statistical task in graph data analysis: Given an n-node Erdős-Rényi random graph of which each edge is present with probability p° independently, output an estimate $\hat{p}$ of the edge density parameter p°, subject to node differential privacy. We consider the error metric $|\hat{p}/p^\circ - 1|$ which can reflect the fact that, the task is more difficult for smaller p°.

38th Conference on Neural Information Processing Systems (NeurIPS 2024).

Without privacy requirement, the empirical edge density[1] $\hat{p}$ achieves the information theoretically optimal error rate $|\hat{p}/p^\circ - 1| \leqslant \tilde{O}(1/(n\sqrt{p^\circ}))$. The standard way to achieve ε-differential node privacy is to add Laplace noise with standard deviation $\Theta(1/(\varepsilon n))$ to the empirical edge density $\hat{p}$. This will incur an additional privacy cost of $\Theta(1/(\varepsilon n p^\circ))$ which dominates the non-private error $\tilde{O}(1/(n\sqrt{p^\circ}))$. Surprisingly, Borgs et al. [BCSZ18] gave an algorithm with privacy cost only $\tilde{O}(1/(\varepsilon n\sqrt{np^\circ}))$ which is negligible to the non-private error for any $\varepsilon \gg 1/\sqrt{n}$. However, their algorithm is based on a general Lipschitz extension technique that has exponential running time. Later, Sealfon and Ullman [SU19] provided a polynomial-time algorithm based on smooth sensitivity with privacy cost $\tilde{O}(1/(\varepsilon n\sqrt{np^\circ}) + 1/(\varepsilon^2 n^2 p^\circ))$, which is much greater than that of [BCSZ18] for $\varepsilon \ll 1/(\sqrt{np^\circ})$. Moreover, [SU19] gives evidence that their approach is inherently prohibited from achieving better privacy cost. On the other hand, known lower bounds in [BCSZ18, SU19] are not for Erdős-Rényi random graphs. This leads us to the following question:

Is there a polynomial-time, differentially node-private, and rate-optimal edge density estimation algorithm for Erdős-Rényi random graphs?

We essentially settled this question in this paper. Specifically, we give a polynomial-time and differentially node-private algorithm with privacy cost $\tilde{O}(1/(\varepsilon n\sqrt{np^\circ}))$. Moreover, we show this error rate is optimal up to a logarithmic factor by proving an information-theoretical lower bound of $\Omega(1/(\varepsilon n\sqrt{np^\circ}))$. Our algorithm actually works for the more general inhomogeneous random graphs [BJR07]. The inhomogeneous random graph model encompasses any random graph model where edges appear independently (after conditioning on node labels). Notable examples include the stochastic block model [HLL83], the latent space model [HRH02], and graphon [BC17].

Our algorithm largely exploits the close connection between differential privacy and adversarial robustness in statistics. This connection dates back to [DL09] and has witnessed significant progress in the past few years [LKKO21, LKO22, KMV22, GH22, AUZ23, HKM22, HKMN23, AKT+23, CCAd+23, CDd+24]. In particular, a very recent line of works [HKM22, HKMN23, CDd+24] could efficiently achieve optimal or nearly-optimal accuracy guarantees in a number of high-dimensional statistical tasks, by integrating two powerful tools — sum-of-squares method [RSS18] and exponential mechanisms [MT07]— in robustness and privacy respectively. Our algorithm extends this line of work. The key technical ingredients of our algorithm are (1) a new sum-of-squares algorithm for robust edge density estimation and (2) an exponential mechanism whose score function is based on the sum-of-squares program. As a consequence, our private algorithm is also robust to adversarial corruptions.

1.1 Results

To state our results formally, we need the following definitions.

Definition 1.1 (Node distance, neighboring graphs)**.** Let $n \in \mathbb{N}$. The node distance between two n-node graphs G and G', denoted by $\mathrm{dist}(G, G')$, is the minimum number of nodes in G that need to be rewired to obtain G'. Moreover, we say G and G' are neighboring graphs if $\mathrm{dist}(G, G') \leqslant 1$.

Definition 1.2 (Node differential privacy)**.** Let $\mathcal{G}$ be the set of graphs. A randomized algorithm $\mathcal{A} : \mathcal{G} \to \mathbb{R}$ is ε-differentially (node-)private if for every neighboring graphs G, G' and every $S \subseteq \mathbb{R}$, we have

$$\mathbb{P}[\mathcal{A}(G) \in S] \leqslant e^{\varepsilon} \cdot \mathbb{P}[\mathcal{A}(G') \in S]\,.$$

Definition 1.3 (Node corruption model)**.** Let $n \in \mathbb{N}$ and $\eta \in [0,1]$. For an n-node graph G, we say an n-node graph G' is an η-corrupted version of G if $\mathrm{dist}(G, G') \leqslant \eta n$.

Erdős-Rényi random graphs. We provide a polynomial-time, differentially node-private and robust edge density estimation algorithm for Erdős-Rényi random graphs.

[1]The (empirical) edge density of an n-node graph equals the number of edges divided by $n(n-1)/2$.

Theorem 1.4 (Erdős-Rényi random graphs, combination of Theorem D.1 and Theorem F.1)**.** *There are constants C_1, C_2, C_3 such that the following holds. For any $\eta \leqslant C_1$, $\varepsilon \geqslant C_2 \log(n)/n$, and $p^\circ \geqslant C_3/n$, there exists a polynomial-time ε-differentially node-private algorithm which, given an η-corrupted Erdős-Rényi random graph $\mathbb{G}(n, p^\circ)$, outputs an estimate $\tilde{p}$ satisfying*

$$\left|\frac{\tilde{p}}{p^\circ} - 1\right| \leqslant O\left(\frac{\sqrt{\log n}}{n\sqrt{p^\circ}} + \frac{\log^2 n}{\varepsilon n\sqrt{np^\circ}} + \frac{\eta \log n}{\sqrt{np^\circ}}\right),$$

with probability $1 - n^{-\Omega(1)}$.

The first term $O(\sqrt{\log n}/(n\sqrt{p^\circ}))$ is the sampling error that is necessary even without privacy or robustness. The second term $O(\log^2(n)/(\varepsilon n\sqrt{np^\circ}))$ is the privacy cost of our algorithm, which matches the exponential-time algorithm in [BCSZ18]. The third term $O(\eta \log n/\sqrt{np^\circ})$ is the robustness cost of our algorithm, which matches the information-theoretical lower bound $\Omega(\eta/\sqrt{np^\circ})$ in [AJK$^+$22, Theorem 1.5] up to a $\log n$ factor.

Moreover, we provide the following lower bound which shows that the privacy cost of our algorithm is optimal up to a $\log n$ factor.[2]

Theorem 1.5 (Privacy lower bound for Erdős-Rényi random graphs)**.** *Suppose there is an ε-differentially node-private algorithm that, given an Erdős-Rényi random graph $\mathbb{G}(n, p^\circ)$, outputs an estimate $\tilde{p}$ satisfying $|\tilde{p}/p^\circ - 1| \leqslant \alpha$ with probability $1 - \beta$. Then we must have*

$$\alpha \geqslant \Omega\left(\frac{\log(1/\beta)}{\varepsilon n\sqrt{np^\circ}}\right).$$

Inhomogeneous random graphs. Given an n-by-n edge connection probability matrix Q°, the inhomogeneous random graph model $\mathbb{G}(n, Q^\circ)$ defines a distribution over n-node graphs where each edge $\{i, j\}$ is present with probability $(Q^\circ)_{ij}$ independently.

We provide a polynomial-time, differentially node-private and robust edge density estimation algorithm for inhomogeneous random graphs.

Theorem 1.6 (Inhomogeneous random graphs, combination of Theorem D.1 and Theorem E.1)**.** *Let Q° be an n-by-n edge connection probability matrix and let $p^\circ := \sum_{i,j} Q^\circ_{ij}/(n^2 - n)$. Suppose $\|Q^\circ\|_\infty \leqslant Rp^\circ$ for some R. There is a sufficiently small constant c such that the following holds. For any η such that $\eta \log(1/\eta)R \leqslant c$, there exists a polynomial-time ε-differentially node-private algorithm which, given an η-corrupted inhomogeneous random graph $\mathbb{G}(n, Q^\circ)$, outputs an estimate $\tilde{p}$ satisfying*

$$\left|\frac{\tilde{p}}{p^\circ} - 1\right| \leqslant O\left(\frac{\sqrt{\log n}}{n\sqrt{p^\circ}} + \frac{R\log^2 n}{\varepsilon n} + R\eta\log(1/\eta)\right),$$

with probability $1 - n^{-\Omega(1)}$.

We improve on the previous private edge density estimation algorithm for inhomogeneous random graphs by Chen et al. [CDd$^+$24, Lemma 4.10]. Their algorithm is based on [SU19] and has privacy cost $\tilde{O}(R/(\varepsilon n) + 1/(\varepsilon^2 nd^\circ))$, while our algorithm only has privacy cost $\tilde{O}(R/(\varepsilon n))$. To the best of our knowledge, even without privacy requirement and in the special case of Erdős-Rényi random graphs, no previous algorithm can match our guarantees in the sparse regime. Specifically, when $d^\circ \ll \log n$ and $\eta \geqslant \Omega(1)$, our algorithm can provide a constant-factor approximation of d°, while the best previous robust algorithm [AJK$^+$22] can not.

We also provide matching lower bounds, showing that the guarantee of our algorithm in Theorem 1.6 is optimal up to logarithmic factors.

[2] Borgs et al. [BCSZ18] proved a lower bound for a variant of Erdős-Rényi random graphs. However, it is not clear whether their proof technique can be easily extended to Erdős-Rényi random graphs.

Theorem 1.7 (Robustness lower bound for inhomogeneous random graphs)**.** *Suppose there is an algorithm satisfies the following guarantee for any symmetric matrix $Q^\circ \in [0,1]^{n\times n}$. Given an η-corrupted inhomogeneous random graph $\mathbb{G}(n, Q^\circ)$, the algorithm outputs an estimate $\hat{p}$ satisfying $|\hat{p}/p^\circ - 1| \leqslant \alpha$ with probability at least 0.99, where $p^\circ = \sum_{i,j} Q^\circ_{ij}/(n^2-n)$. Then we must have $\alpha \geqslant \Omega(R\eta)$, where $R = \max_{i,j} Q^\circ_{ij}/p^\circ$.*

Theorem 1.8 (Privacy lower bound for inhomogeneous random graphs)**.** *Suppose there is an ε-differentially node-private algorithm satisfies the following guarantee for any symmetric matrix $Q^\circ \in [0,1]^{n\times n}$. Given an inhomogeneous random graph $\mathbb{G}(n, Q^\circ)$, the algorithm outputs an estimate $\hat{p}$ satisfying $|\hat{p}/p^\circ - 1| \leqslant \alpha$ with probability $1-\beta$, where $p^\circ = \sum_{i,j} Q^\circ_{ij}/(n^2-n)$. Then we must have*

$$\alpha \geqslant \Omega\left(\frac{R\log(1/\beta)}{n\varepsilon}\right),$$

where $R = \max_{i,j} Q^\circ_{ij}/p^\circ$.

1.2 Techniques

We give an overview of the key techniques used to obtain our algorithm. As our techniques for Erdős-Rényi random graphs can be easily extended to the more general inhomogeneous random graph model, we will focus on Erdős-Rényi random graphs to avoid a proliferation of notation. Specifically, given an η-corrupted Erdős-Rényi random graph $\mathbb{G}(n, d^\circ/n)$, our goal is to output a private estimate of d°.

Reduction from privacy to robustness. Hopkins et al. [HKMN23] and Asi et al. [AUZ23] independently discovered the following black-box reduction from privacy to robustness. Given a robust algorithm $\mathcal{A}_{\text{robust}}$, one can directly obtain a private algorithm via applying the exponential mechanism [MT07] with the following score function,

$$\operatorname{score}(d; A) := \min_{A'}\left\{\operatorname{dist}(A', A) \,:\, |\mathcal{A}_{\text{robust}}(A') - d| \leqslant 1/\operatorname{poly}(n)\right\}, \tag{1.1}$$

where A is the adjacency matrix of input graph and d is a candidate estimate. For privacy analysis, note that the sensitivity of the above score function is bounded by 1, as the node distance between neighboring graphs is at most 1. For utility analysis, when the input graph is a typical Erdős-Rényi random graph, the exponential mechanism will with high probability output a $\hat{d}$ of score $O(\log(n)/\varepsilon)$. Then we can argue that such a $\hat{d}$ is close to d° using the robustness of $\mathcal{A}_{\text{robust}}$. For example, if we plug in the robust algorithm in [AJK+22, Theorem 1.3], then the corresponding exponential mechanism will only incur a privacy cost of $\tilde{O}(1/(\varepsilon n\sqrt{d^\circ}))$.

However, directly plugging in the robust algorithm in [AJK+22] will lead to an exponential-time algorithm, as a single evaluation of the score function requires enumerating all n-node graphs. To obtain a polynomial-time algorithm, we develop a new robust algorithm via the *sum-of-squares* method.[3]

Robust algorithm via sum-of-squares. The sum-of-squares method uses convex programming (in particular, semidefinite programming) to solve polynomial programming. It is a very powerful tool for designing polynomial-time robust estimators (see [RSS18]). To obtain a robust algorithm via sum-of-squares, we first identify a set of polynomial constraints that a typical (uncorrupted) Erdős-Rényi random graph would satisfy. Specifically, these polynomial constraints encode the following regularity conditions: (1) the degrees of the nodes are highly concentrated, and (2) the centered adjacency matrix is spectrally bounded.

[3] In general, the black-box reduction by [HKMN23, AUZ23] does not provide guarantees in terms of computational complexity. For the problem of robust edge density estimation under node corruption, there is no known sum-of-squares algorithm before our work, and we are only aware of the iterative algorithm [AJK+22]. For such algorithms not based on convex relaxation, it is completely unclear how to use the aforementioned connection between private and robust estimation towards an efficient private algorithm.

We also include the constraint that at most η fraction of the nodes in the graph are corrupted. Then we give a proof that if a graph satisfies the above constraints, then its average degree will be close to d°, even when η fraction of nodes in the input graph are arbitrarily corrupted. Importantly, the proof is simple enough that it is captured by the sum-of-squares proof system (see [FKP+19]). This allows us to extend the utility guarantee of the polynomial program to its semidefinite programming relaxation, which results in a polynomial-time robust algorithm.

Sum-of-squares exponential mechanism. Given the above robust algorithm, we then use the sum-of-squares exponential mechanism developed in [HKM22, HKMN23] to obtain a private algorithm. More specifically, we apply the exponential mechanism with the sum-of-squares relaxation of the score function in Eq. (1.1). In this way, we obtain a private algorithm that is also robust to adversarial corruptions.

1.3 Notation

We introduce some notation used throughout this paper. We write $f \lesssim g$ to denote the inequality $f \leqslant C \cdot g$ for some absolute constant $C > 0$. We write $O(f)$ and $\Omega(f)$ to denote quantities f_- and f_+ satisfying $f_- \lesssim f$ and $f \lesssim f_+$ respectively. We use boldface to denote random variables, e.g., $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}$. For a matrix M, we use $\|M\|_{\mathrm{op}}$ for the spectral norm of M. Let $\mathbb{1}$ and $\mathbb{0}$ denote the all-one and all-zero vector respectively, of which the size will be clear from the context. We use a graph G and its adjacency matrix $A = A(G)$ interchangeably when there is no ambiguity. For an n-by-n matrix M, we use $d(M)$ to denote its average row/column sum, i.e., $d(M) = \sum_{i,j} M_{ij}/n$. For any matrices (or vectors) M, N of the same shape, we use $M \odot N$ to denote the element-wise product (aka Hadamard product) of M and N.

1.4 Organization

The rest of the paper is organized as follows. In Section 2, we give a proof overview of our results and defer full proofs to the appendices. The appendices are organized as follows. We provide some sum-of-squares background in Appendix A and some concentration inequalities for random graphs in Appendix B. In Appendix C, we present a general sum-of-squares exponential mechanism that all of our private algorithms in this paper are based on. In Appendix D, we present our coarse estimation algorithm and give a full proof of its guarantees (Theorem D.1). In Appendix E, we present our fine estimation algorithm for inhomogeneous random graphs and give a full proof of its guarantees (Theorem E.1). In Appendix F, we present our fine estimation algorithm for Erdős-Rényi random graphs and give a full proof of its guarantees (Theorem F.1). All lower bounds are proved in Appendix G.

2 Private and robust algorithm for Erdős-Rényi random graphs

In this section, we describe our private and robust algorithm for Erdős-Rényi random graphs. We also give an overview of the analysis of our algorithm and sketch the proof of our lower bounds.

Our overall algorithm consists of two stages. In the first stage, we compute a coarse estimate that approximates the edge density parameter within constant factors. In the second stage, we improve the accuracy of this coarse estimate to the optimum. Since our algorithm is private in both stages, it is also private overall by the composition theorem of differential privacy (see [DR14, Section 3.5]).

We remark that for the Erdős-Rényi random graph model $\mathbb{G}(n, p^\circ)$, estimating its edge density parameter p° is equivalent to estimating its expected average degree $d^\circ := np^\circ$.[4] For the convenience of notation, we set our goal as estimating the expected average degree d° throughout this section.

[4] Strictly speaking, the expected average degree of $\mathbb{G}(n, p^\circ)$ should be $(n-1)p^\circ$. Here we call np° the expected average degree just for notational convenience. In the end, $(n-1)p^\circ = (1 - 1/n) \cdot np^\circ$.

2.1 General algorithm framework

Given an n-by-n symmetric matrix A and a scalar $\gamma \in [0,1]$, let $\mathcal{T}(Y,z;A,\gamma)$ be a polynomial system with indeterminates $Y = (Y_{ij})_{i,j\in[n]}$ and $z = (z_i)_{i\in[n]}$ that encodes the node distance between Y and A:

$$\mathcal{T}(Y,z;A,\gamma) := \begin{cases} z \odot z = z,\ \langle \mathbb{1}, z\rangle \geqslant (1-\gamma)n \\ 0 \leqslant Y \leqslant \mathbb{1}\mathbb{1}^\top,\ Y = Y^\top \\ Y \odot zz^\top = A \odot zz^\top \end{cases} \Bigg\} . \tag{2.1}$$

Let $\mathcal{R}(Y)$ be A polynomial system that encodes regularity conditions of Erdős-Rényi random graphs. The key observation here is that, for any $Y \in \{0,1\}^{n\times n}$ and $z \in \{0,1\}^n$ that satisfy constraints in $\mathcal{T}(Y,z;A,\gamma) \cup \mathcal{R}(Y)$, Y is a graph that behaves like Erdős-Rényi random graphs (in the sense of the regularity conditions) and is within node distance γn to A where they agree on $\{i \in [n] : z_i = 1\}$.

The key ingredient of our result is that, given proper regularity conditions $\mathcal{R}(Y)$, we can give degree-8 sum-of-squares proofs: for any Y that satisfies constraints in $\mathcal{T}(Y,z;A,\gamma) \cup \mathcal{R}(Y)$, the average degree of Y is close to the expected average degree d°, even when the input graph A is a γ-corrupted Erdős-Rényi random graph $\mathbb{G}(n, d^\circ/n)$. As a result of the sum-of-squares proofs-to-algorithms framework (see Theorem A.6), we can get an efficient and robust estimator $\tilde{\mathbb{E}}[d(Y)]$, where $\tilde{\mathbb{E}}$ is a pseudo-expectation obtained by solving level-8 sum-of-squares relaxation of $\mathcal{T}(Y,z;A,\gamma) \cup \mathcal{R}(Y)$.

Based on the above identifiability proof for robust estimation, we design a private and robust algorithm by applying the exponential mechanism[5] with the following score function:

$$\begin{aligned} \text{sos-score}(d;A) := \min_{0\leqslant\gamma\leqslant 1} \gamma n \text{ s.t. } \exists \text{ level-8 pseudo-expectation } \tilde{\mathbb{E}} \text{ satisfying} \\ \mathcal{T}(Y,z;A,\gamma) \cup \mathcal{R}(Y) \cup \left\{|d(Y)-d| \leqslant 1/\mathrm{poly}(n)\right\}. \end{aligned} \tag{2.2}$$

Similar to Eq. (1.1), it is easy to observe this exponential mechanism is private.

Lemma 2.1 (Privacy)**.** *Consider the distribution $\mu_{A,\varepsilon}$ with support $[0,n]$ and density*

$$\mathrm{d}\mu_{A,\varepsilon}(d) \propto \exp(-\varepsilon \cdot \textit{sos-score}(d;A))\,, \tag{2.3}$$

where sos-score$(d;A)$ is defined in Eq. (2.2). A sample from $\mu_{A,\varepsilon}$ is 2ε-differentially private.

Proof. Since the node distance between neighboring graphs is at most 1, the sensitivity of the following score function is bounded by 1:

$$\text{score}(d;A) := \min_{0\leqslant\gamma\leqslant 1} \gamma n \text{ s.t. } \mathcal{T}(Y,z;A,\gamma) \cup \mathcal{R}(Y) \cup \left\{|d(Y)-d| \leqslant 1/\mathrm{poly}(n)\right\} \text{ is feasible.}$$

One can show that such sensitivity bound is inherited by its sum-of-squares relaxation sos-score as defined in Eq. (2.2). By a standard sensitivity-to-privacy argument (see e.g. [DR14, Theorem 3.10]), the exponential mechanism is 2ε-differentially private. □

To analyze the utility of the private algorithm, we use the robustness of the score function. Assume the input graph is uncorrupted for simplicity. For a typical Erdős-Rényi random graph $A^\circ \sim \mathbb{G}(n, d^\circ/n)$, we have sos-score$(d^\circ, A^\circ) = 0$. By a standard volume argument (see e.g. [DR14, Theorem 3.11]), the exponential mechanism with high probability outputs a scalar d satisfying sos-score$(d;A^\circ) \lesssim \log(n)/\varepsilon$. By the definition of our score function in Eq. (2.2), this implies that there exists a level-8 pseudo-distribution satisfying $\mathcal{T}(Y,z;A^\circ,\gamma) \cup \mathcal{R}(Y)$ with $\gamma \lesssim \log(n)/(\varepsilon n)$. The utility then follows from the above identifiability proof for robust estimation.

[5]To efficiently implement this exponential mechanism, we note that the score function Eq. (2.2) can be evaluated in polynomial time by combining binary search and semidefinite programming. By discretizing $[0,n]$ with step size $1/\mathrm{poly}(n)$, one can sample from the distribution Eq. (2.3) with a polynomial number of queries to the score function. For more detailed discussions, see Remark C.1 and Remark C.2.

2.2 Coarse estimation

In this part, we describe a private and robust algorithm that can estimate the expected average degree d° within a constant approximation ratio.

Theorem 2.2 (Coarse estimation algorithm, informal restatement of Theorem D.1)**.** *For η smaller than some constant, there is a polynomial-time ε-differentially node-private algorithm which, given an η-corrupted Erdős-Rényi random graph $\mathbb{G}(n, d^\circ/n)$, outputs an estimate $\hat{d}$ such that $|\hat{d} - d^\circ| \leqslant 0.5d^\circ$.*

We give a proof sketch of Theorem 2.2 at the end of this subsection. The formal theorem and proofs are deferred to Appendix D.

Identifiability proof for robust estimation. We first give a polynomial system that can identify the expected average degree d° up to constant factors, even when η-fraction of nodes are corrupted. Consider the following regularity condition on degrees:

$$\mathcal{R}(Y) := \left\{(Y\mathbb{1})_i \leqslant 2\log(1/\eta) \cdot d(Y)\,, \quad \forall i \in [n]\right\}. \tag{2.4}$$

The following lemma shows that Erdős-Rényi random graphs satisfy $\mathcal{T}(Y, z; A, 2\eta) \cup \mathcal{R}(Y)$ with high probability.

Lemma 2.3 (Feasibility)**.** *Let $A^\circ \sim \mathbb{G}(n, d^\circ/n)$ and let A be an η-corrupted version of A°. With high probability, there exists a graph Y that satisfies the constraints in $\mathcal{T}(Y, z; A, 2\eta) \cup \mathcal{R}(Y)$.*

Proof sketch. For $d^\circ \gg \log(n)$, the maximum degree of A° is of order $O(d^\circ)$. Therefore, the uncorrupted graph A° satisfies the constraints. For $d^\circ \ll \log n$, using concentration properties of random graphs, we can show that the number of high degree nodes is bounded by ηn. A feasible graph can then be obtained from the uncorrupted graph A° by trimming these highest degree nodes. □

Next, we show that these polynomial constraints give an identifiability proof for the expected average degree d°.

Lemma 2.4 (Identifiability)**.** *Let $A^\circ \sim \mathbb{G}(n, d^\circ/n)$ and let A be an η-corrupted version of A°. For η smaller than some constant and $\gamma \leqslant O(\eta)$, with high probability there is a degree-8 sum-of-squares proof that, if Y satisfies $\mathcal{T}(Y, z; A, \gamma) \cup \mathcal{R}(Y)$, then $|d(Y) - d^\circ| \leqslant 0.001d^\circ$.*

Proof sketch. We first assume that $d^\circ \gg \log(n)$, for which the proof is simpler. By the degree-bound constraint $\mathcal{R}(Y)$, we have $n|d(Y) - d(A^\circ)| \leqslant 2\log(1/\eta) \cdot (d(Y) + d^\circ) \cdot \mathrm{dist}(Y, A^\circ)$. Using the constraints $Y \odot zz^\top = A \odot zz^\top$ and $\langle 1, z\rangle \geqslant (1-\gamma)n$, we have $\mathrm{dist}(Y, A) \leqslant \gamma n$. Since $\mathrm{dist}(A, A^\circ) \leqslant \eta n$, by triangle inequality, we have $\mathrm{dist}(Y, A^\circ) \leqslant (\gamma + \eta)n$. Therefore, we have $|d(Y) - d(A^\circ)| \leqslant 0.0001d^\circ$ when γ, η are at most some small constants. Finally, by random graph concentration, we have $|d^\circ - d(A^\circ)| \leqslant o(d^\circ)$ with high probability. Therefore, we have $|d(Y) - d(A^\circ)| \leqslant 0.001d^\circ$.

To deal with the sparse regime where $d^\circ \ll \log n$, we need to truncate the nodes of A° with degree $\Omega(\log(1/\eta)d^\circ)$. Our key observation is that, the average degree of the graph before and after truncation only differ by a constant factor. Therefore, we can still get $|d(Y) - d(A^\circ)| \leqslant 0.001d^\circ$.

Furthermore, it can be shown that this proof is a degree-8 sum-of-squares proof. □

Robust algorithm via sum-of-squares. Consider the algorithm that finds a level-8 pseudo-expectation satisfying $\mathcal{T}(Y, z; A, 2\eta) \cup \mathcal{R}(Y)$ —with $\mathcal{R}(Y)$ given in Eq. (2.4)— and outputs $\tilde{\mathbb{E}}[d(Y)]$. By Lemma 2.3, such a pseudo-expectation $\tilde{\mathbb{E}}$ exists with high probability. It follows from the sum-of-squares identifiability proof in Lemma 2.4 that $|\tilde{\mathbb{E}}[d(Y)] - d^\circ| \leqslant 0.001d^\circ$. Moreover, the algorithm can be implemented by semidefinite programming and run in polynomial time.

Private and robust algorithm via sum-of-squares exponential mechanism. We present our private and robust algorithm in Algorithm 2.5 and give a proof sketch of Theorem 2.2.

Algorithm 2.5 (Private coarse estimation for Erdős-Rényi random graphs)**.**
Input: η-corrupted Erdős-Rényi random graph A.

Privacy parameter: ε.

Output: A sample from the distribution $\mu_{A,\varepsilon}$ with support $[0, n]$ and density

$$\mathrm{d}\mu_{A,\varepsilon}(d) \propto \exp(-\varepsilon \cdot \text{sos-score}(d; A)) , \tag{2.5}$$

where

$$\begin{aligned} \text{sos-score}(d; A) := \min_{0 \leqslant \gamma \leqslant 1} \gamma n \text{ s.t. } \exists \text{ level-8 pseudo-expectation } \tilde{\mathbb{E}} \text{ satisfying} \\ \mathcal{T}(Y, z; A, \gamma) \cup \mathcal{R}(Y) \cup \left\{ |d(Y) - d| \leqslant 1/\mathrm{poly}(n) \right\} , \end{aligned} \tag{2.6}$$

with $\mathcal{R}(Y)$ given in Eq. (2.4).

Proof sketch of Theorem 2.2. Privacy. By Lemma 2.1, Algorithm 2.5 is 2ε-differentially private.

Utility. For simplicity, we consider the case when there is no corruption (i.e. $\eta = 0$). The analysis for the case when $\eta > 0$ is similar. Let $A^\circ \sim \mathbb{G}(n, d^\circ/n)$. Then with high probability $\text{sos-score}(d^\circ; A^\circ) = 0$. By a standard volume argument, Algorithm 2.5 outputs a scalar d that satisfies $\text{sos-score}(d; A^\circ) \lesssim \log(n)/\varepsilon$ with high probability. By the definition of sos-score in Eq. (2.6), this implies that there exists a level-8 pseudo-distribution satisfying $\mathcal{T}(Y, z; A, \gamma) \cup \mathcal{R}(Y) \cup \{|d(Y) - d| \leqslant 1/\mathrm{poly}(n)\}$ with $\gamma \lesssim \log(n)/(\varepsilon n)$. When $\log(n)/(\varepsilon n)$ is at most a small constant, it follows from our sum-of-squares identifiability proof in Lemma 2.4 that, Algorithm 2.5 outputs a constant-factor approximation of d° with high probability. □

2.3 Fine estimation

From Section 2.2, we know how to obtain a constant-factor approximation of d° privately and robustly. In this section, we show how to improve the accuracy to the optimum.

Theorem 2.6 (Fine estimation algorithm, informal restatement of Theorem F.1)**.** *Let $0.5d^\circ \leqslant \hat{d} \leqslant 2d^\circ$. For η smaller than some constant, there is a polynomial-time ε-differentially node-private algorithm which, given an η-corrupted Erdős-Rényi random graph $\mathbb{G}(n, d^\circ/n)$ and $\hat{d}$, outputs an estimate $\tilde{d}$ such that*

$$\left| \frac{\tilde{d}}{d^\circ} - 1 \right| \leqslant \tilde{O}\left(\frac{1}{\sqrt{nd^\circ}} + \frac{1}{\varepsilon n \sqrt{d^\circ}} + \frac{\eta}{\sqrt{d^\circ}} \right) .$$

We give a proof sketch of Theorem 2.6 at the end of this section. The formal theorem and proofs are deferred to Appendix F.

Identifiability proof for robust estimation. We first give a polynomial system which can identify the expected average degree d° with optimal error rate, when provided with a coarse estimate $\hat{d}$. Consider the following regularity conditions on degrees and eigenvalues:

$$\mathcal{R}(Y) := \left\{ \begin{array}{ll} |(Y\mathbb{1})_i - d(Y)| \leqslant \sqrt{\hat{d}} \log n , & \forall i \in [n] \\ \left\| Y - \frac{d(Y)}{n} \mathbb{1}\mathbb{1}^\top \right\|_{\mathrm{op}} \leqslant \sqrt{\hat{d} \log n} & \end{array} \right\} . \tag{2.7}$$

Lemma 2.7 (Feasibility)**.** *Let $A^\circ \sim \mathbb{G}(n, d^\circ/n)$ and let A be an η-corrupted version of A°. Suppose $d^\circ/2 \leqslant \hat{d} \leqslant 2d^\circ$. Then with high probability, there exists a graph Y that satisfies the constraints in $\mathcal{T}(Y, z; A, \eta) \cup \mathcal{R}(Y)$.*

Proof. By Chernoff bound, with high probability, the degree of each node in A° deviates from d° by at most $O(\sqrt{d^\circ} \log n)$. By the concentration of the spectral norm of random matrices [BvH16], with high probability, we have $\|A^\circ - \frac{d(A^\circ)}{n}\mathbb{1}\mathbb{1}^\top\|_{\text{op}} \lesssim \sqrt{d^\circ \log n}$. Hence, $\mathcal{T}(Y, z; A, \eta) \cup \mathcal{R}(Y)$ is satisfied by $Y = A^\circ$ and $z = z^\circ$ where z° is the indicator vector for uncorrupted nodes. □

Next we give a sum-of-squares identifiability proof for expected average degree estimation with optimal accuracy.

Lemma 2.8 (Identifiability). *Let $A^\circ \sim \mathbb{G}(n, d^\circ/n)$ and let A be an η-corrupted version of A°. Suppose $d^\circ/2 \leqslant \hat{d} \leqslant 2d^\circ$. For η smaller than some constant and $\gamma \leqslant O(\eta)$, with high probability there is a degree-8 sum-of-squares proof that, if Y satisfies $\mathcal{T}(Y, z; A, \eta) \cup \mathcal{R}(Y)$, then*

$$\left|\frac{d(Y)}{d^\circ} - 1\right| \leqslant \tilde{O}\left(\frac{1}{\sqrt{nd^\circ}} + \frac{\eta}{\sqrt{d^\circ}}\right).$$

Proof sketch. Let Y_1, Y_2 be two graphs satisfying the regularity condition $\mathcal{R}(Y_1)$ and $\mathcal{R}(Y_2)$ as described in Eq. (2.7), respectively. We give sum-of-squares proof that, if $\text{dist}(Y_1, Y_2) \leqslant \zeta n$ and ζ is at most some small constant, then $|d(Y_1) - d(Y_2)| \leqslant \zeta\sqrt{\hat{d}}\log n$.

Let $w \in \{0,1\}^n$ be the indicator vector for the shared induced subgraph between Y_1 and Y_2, i.e $Y_1 \odot ww^\top = Y_2 \odot ww^\top$. When $\text{dist}(Y_1, Y_2) \leqslant \zeta n$, we have $\langle w, \mathbb{1}\rangle \geqslant (1-\zeta)n$. We have

$$\begin{aligned}
n\Big(d(Y_1) - d(Y_2)\Big) =& \langle Y_1 - Y_2, \mathbb{1}\mathbb{1}^\top\rangle \\
=& \langle Y_1 - Y_2, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle \\
=& \langle Y_1 - \frac{d(Y_1)}{n}\mathbb{1}\mathbb{1}^\top + \frac{d(Y_1)}{n}\mathbb{1}\mathbb{1}^\top - \frac{d(Y_2)}{n}\mathbb{1}\mathbb{1}^\top + \frac{d(Y_2)}{n}\mathbb{1}\mathbb{1}^\top - Y_2, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle \\
=& \langle Y_1 - \frac{d(Y_1)}{n}\mathbb{1}\mathbb{1}^\top, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle + \langle \frac{d(Y_2)}{n}\mathbb{1}\mathbb{1}^\top - Y_2, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle \\
&+ \langle \frac{d(Y_1)}{n}\mathbb{1}\mathbb{1}^\top - \frac{d(Y_2)}{n}\mathbb{1}\mathbb{1}^\top, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle\,.
\end{aligned}$$

By rearranging terms, we can get

$$\frac{\langle \mathbb{1}, w\rangle^2}{n}\Big(d(Y_1) - d(Y_2)\Big) = \langle Y_1 - \frac{d(Y_1)}{n}\mathbb{1}\mathbb{1}^\top, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle + \langle \frac{d(Y_2)}{n}\mathbb{1}\mathbb{1}^\top - Y_2, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle\,.$$

For the first term $\langle Y_1 - \frac{d(Y_1)}{n}\mathbb{1}\mathbb{1}^\top, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle$, we have

$$\langle Y_1 - \frac{d(Y_1)}{n}\mathbb{1}\mathbb{1}^\top, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle = 2\langle Y_1 - \frac{d(Y_1)}{n}\mathbb{1}\mathbb{1}^\top, \mathbb{1}(\mathbb{1}-w)^\top\rangle + \langle \frac{d(Y_1)}{n}\mathbb{1}\mathbb{1}^\top - Y_1, (\mathbb{1}-w)(\mathbb{1}-w)^\top\rangle\,.$$

From constraints $|(Y_1\mathbb{1})_i - d(Y_1)| \leqslant \sqrt{\hat{d}}\log(n)$ for all $i \in [n]$, we have

$$\langle Y_1 - \frac{d(Y_1)}{n}\mathbb{1}\mathbb{1}^\top, \mathbb{1}(\mathbb{1}-w)^\top\rangle = \langle Y_1\mathbb{1} - d(Y_1)\mathbb{1}, \mathbb{1} - w\rangle \leqslant \zeta n \log(n)\sqrt{\hat{d}}\,.$$

From constraints $\left\|Y_1 - \frac{d(Y_1)}{n}\mathbb{1}\mathbb{1}^\top\right\|_{\text{op}} \leqslant \delta\sqrt{\hat{d}}$, we have

$$\langle \frac{d(Y_1)}{n}\mathbb{1}\mathbb{1}^\top - Y_1, (\mathbb{1}-w)(\mathbb{1}-w)^\top\rangle \leqslant \left\|Y_1 - \frac{d(Y_1)}{n}\mathbb{1}\mathbb{1}^\top\right\|_{\text{op}} \|\mathbb{1} - w\|_2^2 \leqslant \zeta n\sqrt{\hat{d}}\log(n)\,,$$

The same bounds also apply for the second term $\langle Y_2 - \frac{d(Y_2)}{n}\mathbb{1}\mathbb{1}^\top, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle$. Since $\langle \mathbb{1}, w\rangle \geqslant \Omega(n)$, it follows that $|d(Y_1) - d(Y_2)| \leqslant \tilde{O}(\zeta\sqrt{\hat{d}}) \leqslant \tilde{O}(\zeta\sqrt{d^\circ})$.

Since the original uncorrupted graph satisfies the regularity conditions, this gives the identifiability proof that $|d(Y) - d(A^\circ)| \leqslant \tilde{O}(\zeta\sqrt{d^\circ})$. By random graph concentration, with high probability, we have $|d^\circ - d(A^\circ)| \leqslant \tilde{O}(\sqrt{d^\circ/n})$. The claim thus follows. □

Robust algorithm via sum-of-squares. Consider the algorithm that finds a level-8 pseudo-expectation satisfying $\mathcal{T}(Y, z; A, \eta) \cup \mathcal{R}(Y)$ —with $\mathcal{R}(Y)$ given in Eq. (2.7)— and outputs $\tilde{\mathbb{E}}[d(Y)]$. By Lemma 2.7, such a pseudo-expectation $\tilde{\mathbb{E}}$ exists with high probability. It follows from the sum-of-squares identifiability proof in Lemma 2.8 that $|\tilde{\mathbb{E}}[d(Y)]/d^\circ - 1| \leqslant \tilde{O}(1/\sqrt{nd^\circ} + \eta/\sqrt{d^\circ})$. Moreover, the algorithm can be implemented by semidefinite programming and run in polynomial time.

Private and robust algorithm via sum-of-squares exponential mechanism. We present our private and robust algorithm in Algorithm 2.9 and give a proof sketch of Theorem 2.6.

Algorithm 2.9 (Private fine estimation for Erdős-Rényi random graphs)**.**

Input: η corrupted random graph A, ε-differentially private coarse estimate $\hat{d}$.

Privacy parameter: ε.

Output: A sample from the distribution $\mu_{A,\varepsilon}$ with support $[0, n]$ and density

$$\mathrm{d}\mu_{A,\varepsilon}(d) \propto \exp(-\varepsilon \cdot \text{sos-score}(d; A))\,, \tag{2.8}$$

where sos-score$(d; A)$ is defined as

$$\begin{aligned}\text{sos-score}(d; A) := \min_{0 \leqslant \gamma \leqslant 1} \gamma n \text{ s.t. } &\exists \text{ level-8 pseudo-expectation } \tilde{\mathbb{E}} \text{ satisfying} \\ &\mathcal{T}(Y, z; A, \gamma) \cup \mathcal{R}(Y) \cup \left\{|d(Y) - d| \leqslant 1/\mathrm{poly}(n)\right\},\end{aligned} \tag{2.9}$$

with $\mathcal{R}(Y)$ given in Eq. (2.7).

Proof sketch of Theorem 2.6. Privacy. By Lemma 2.1, Algorithm 2.9 is 2ε-differentially private. *Utility.* For simplicity, we consider the case when there is no corruption (i.e. $\eta = 0$). The analysis for the case when $\eta > 0$ is similar. Let $A^\circ \sim \mathbb{G}(n, d^\circ/n)$. Then with high probability sos-score$(d^\circ; A^\circ) = 0$. By a standard volume argument, Algorithm 2.9 outputs a scalar d that satisfies sos-score$(d; A^\circ) \lesssim \log(n)/\varepsilon$ with high probability. By the definition of sos-score in Eq. (2.9), this implies that with high probability there exists a level-8 pseudo-distribution satisfying $\mathcal{T}(Y, z; A, \gamma) \cup \mathcal{R}(Y)$ with $\gamma \lesssim \log(n)/(\varepsilon n)$. Taking $\eta = \log(n)/(\varepsilon n)$ in Lemma 2.8, it follows that Algorithm 2.9 outputs an estimate $\tilde{d}$ such that $|\tilde{d}/d^\circ - 1| \leqslant \tilde{O}(1/\sqrt{nd^\circ} + 1/(\varepsilon n\sqrt{d^\circ}))$ with high probability. □

2.4 Lower bound

We sketch the proof of Theorem 1.5. Let $\alpha \in [0, 1]$ and $d = (1 - \alpha)d^\circ$. We can construct a coupling ω between the distributions $\mathbb{G}(n, d/n)$ and $\mathbb{G}(n, d^\circ/n)$ with the following property. For $(\boldsymbol{G}, \boldsymbol{G}') \sim \omega$, we have dist$(\boldsymbol{G}, \boldsymbol{G}')$ bounded by $\tilde{O}(\alpha n\sqrt{d^\circ})$ with overwhelmingly high probability. By the definition of differential privacy, when $\varepsilon\alpha n\sqrt{d^\circ} \leqslant 1/\mathrm{polylog}(n)$, the output of an ε-differentially private algorithm are indistinguishable under $\mathbb{G}(n, d/n)$ and $\mathbb{G}(n, d^\circ/n)$. Therefore, by setting $\alpha = \tilde{O}(1/\varepsilon n\sqrt{d^\circ})$, we conclude that no ε-differentially private algorithm can achieve error rate better than $\tilde{O}(1/\varepsilon n\sqrt{d^\circ})$. This provides a matching lower bound for our private edge density estimation algorithm.

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 815464). We thank the anonymous reviewers for constructive feedback.

References

[AJ06] José A Adell and Pedro Jodrá, *Exact kolmogorov and total variation distances between some familiar discrete distributions*, Journal of Inequalities and Applications **2006** (2006), 1–8. 38

[AJK+22] Jayadev Acharya, Ayush Jain, Gautam Kamath, Ananda Theertha Suresh, and Huanyu Zhang, *Robust estimation for random graphs*, Conference on Learning Theory, PMLR, 2022, pp. 130–166. 3, 4, 20, 37

[AKT+23] Daniel Alabi, Pravesh K Kothari, Pranay Tankala, Prayaag Venkat, and Fred Zhang, *Privately estimating a gaussian: Efficient, robust, and optimal*, Proceedings of the 55th Annual ACM Symposium on Theory of Computing, 2023, pp. 483–496. 2

[AUZ23] Hilal Asi, Jonathan Ullman, and Lydia Zakynthinou, *From robustness to privacy and back*, International Conference on Machine Learning, PMLR, 2023, pp. 1121–1146. 2, 4

[BBDS13] Jeremiah Blocki, Avrim Blum, Anupam Datta, and Or Sheffet, *Differentially private data analysis of social networks via restricted sensitivity*, Proceedings of the 4th conference on Innovations in Theoretical Computer Science, 2013, pp. 87–96. 1

[BC17] Christian Borgs and Jennifer Chayes, *Graphons: A nonparametric method to model, estimate, and design algorithms for massive networks*, Proceedings of the 2017 ACM Conference on Economics and Computation, 2017, pp. 665–672. 2

[BCS15] Christian Borgs, Jennifer T. Chayes, and Adam Smith, *Private graphon estimation for sparse graphs*, 2015. 1

[BCSZ18] Christian Borgs, Jennifer Chayes, Adam Smith, and Ilias Zadik, *Revealing network structure, confidentially: Improved rates for node-private graphon estimation*, 2018. 1, 2, 3

[BJR07] Béla Bollobás, Svante Janson, and Oliver Riordan, *The phase transition in inhomogeneous random graphs*, Random Structures & Algorithms **31** (2007), no. 1, 3–122. 2

[BS14] Boaz Barak and David Steurer, *Sum-of-squares proofs and the quest toward optimal algorithms*, Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. IV, Kyung Moon Sa, Seoul, 2014, pp. 509–533. MR 3727623 14

[BS16] ______, *Proofs, beliefs, and algorithms through the lens of sum-of-squares*, lecture notes, 2016. 14

[BvH16] Afonso S. Bandeira and Ramon van Handel, *Sharp nonasymptotic bounds on the norm of random matrices with independent entries*, Ann. Probab. **44** (2016), no. 4, 2479–2506. MR 3531673 9, 18

[CCAd+23] Hongjie Chen, Vincent Cohen-Addad, Tommaso d'Orsi, Alessandro Epasto, Jacob Imola, David Steurer, and Stefan Tiegel, *Private estimation algorithms for stochastic block models and mixture models*, Advances in Neural Information Processing Systems **36** (2023), 68134–68183. 2

[CDd+24] Hongjie Chen, Jingqiu Ding, Tommaso d'Orsi, Yiding Hua, Chih-Hung Liu, and David Steurer, *Private graphon estimation via sum-of-squares*, arXiv preprint arXiv:2403.12213 (2024). 1, 2, 3, 19

[DL09] Cynthia Dwork and Jing Lei, *Differential privacy and robust statistics*, Proceedings of the forty-first annual ACM symposium on Theory of computing, 2009, pp. 371–380. 2

[DMNS06] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith, *Calibrating noise to sensitivity in private data analysis*, Theory of Cryptography: Third Theory of Cryptography Conference, TCC 2006, New York, NY, USA, March 4-7, 2006. Proceedings 3, Springer, 2006, pp. 265–284. 1

[DR14] Cynthia Dwork and Aaron Roth, *The algorithmic foundations of differential privacy*, Foundations and Trends® in Theoretical Computer Science **9** (2014), no. 3–4, 211–407. 5, 6

[DSSU17] Cynthia Dwork, Adam Smith, Thomas Steinke, and Jonathan Ullman, *Exposed! a survey of attacks on private data*, Annual Review of Statistics and Its Application **4** (2017), 61–84. 1

[FKP+19] Noah Fleming, Pravesh Kothari, Toniann Pitassi, et al., *Semialgebraic proofs and efficient algorithm design*, Foundations and Trends® in Theoretical Computer Science **14** (2019), no. 1-2, 1–221. 5

[GH22] Kristian Georgiev and Samuel Hopkins, *Privacy induces robustness: Information-computation gaps and sparse mean estimation*, Advances in Neural Information Processing Systems (S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, eds.), vol. 35, Curran Associates, Inc., 2022, pp. 6829–6842. 2

[HKM22] Samuel B Hopkins, Gautam Kamath, and Mahbod Majid, *Efficient mean estimation with pure differential privacy via a sum-of-squares exponential mechanism*, Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing, 2022, pp. 1406–1417. 2, 5

[HKMN23] Samuel B. Hopkins, Gautam Kamath, Mahbod Majid, and Shyam Narayanan, *Robustness implies privacy in statistical estimation*, Proceedings of the 55th Annual ACM Symposium on Theory of Computing (New York, NY, USA), STOC 2023, Association for Computing Machinery, 2023, p. 497–506. 2, 4, 5, 18

[HL18] Samuel B. Hopkins and Jerry Li, *Mixture models, robustness, and sum of squares proofs*, Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing (New York, NY, USA), STOC 2018, Association for Computing Machinery, 2018, p. 1021–1034. 14

[HLL83] Paul W Holland, Kathryn Blackmond Laskey, and Samuel Leinhardt, *Stochastic blockmodels: First steps*, Social networks **5** (1983), no. 2, 109–137. 2

[Hop20] Samuel B Hopkins, *Mean estimation with sub-gaussian rates in polynomial time*, The Annals of Statistics **48** (2020), no. 2, 1193–1213. 14

[HRH02] Peter D Hoff, Adrian E Raftery, and Mark S Handcock, *Latent space approaches to social network analysis*, Journal of the american Statistical association **97** (2002), no. 460, 1090–1098. 2

[KMV22] Pravesh Kothari, Pasin Manurangsi, and Ameya Velingker, *Private robust estimation by stabilizing convex relaxations*, Proceedings of Thirty Fifth Conference on Learning Theory (Po-Ling Loh and Maxim Raginsky, eds.), Proceedings of Machine Learning Research, vol. 178, PMLR, 02–05 Jul 2022, pp. 723–777. 2

[KNRS13] Shiva Prasad Kasiviswanathan, Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith, *Analyzing graphs with node differential privacy*, Theory of Cryptography: 10th Theory of Cryptography Conference, TCC 2013, Tokyo, Japan, March 3-6, 2013. Proceedings, Springer, 2013, pp. 457–476. 1

[KRSY11] Vishesh Karwa, Sofya Raskhodnikova, Adam Smith, and Grigory Yaroslavtsev, *Private analysis of graph structure*, Proceedings of the VLDB Endowment **4** (2011), no. 11, 1146–1157. 1

[KSS18] Pravesh K Kothari, Jacob Steinhardt, and David Steurer, *Robust moment estimation and improved clustering via sum of squares*, Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, 2018, pp. 1035–1046. 14

[LKKO21] Xiyang Liu, Weihao Kong, Sham Kakade, and Sewoong Oh, *Robust and differentially private mean estimation*, Advances in neural information processing systems **34** (2021), 3887–3901. 2

[LKO22] Xiyang Liu, Weihao Kong, and Sewoong Oh, *Differential privacy and robust statistics in high dimensions*, Conference on Learning Theory, PMLR, 2022, pp. 1167–1246. 2

[MT07] Frank McSherry and Kunal Talwar, *Mechanism design via differential privacy*, 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07), IEEE, 2007, pp. 94–103. 2, 4

[NRS07] Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith, *Smooth sensitivity and sampling in private data analysis*, Proceedings of the thirty-ninth annual ACM symposium on Theory of computing, 2007, pp. 75–84. 1

[NS09] Arvind Narayanan and Vitaly Shmatikov, *De-anonymizing social networks*, 2009 30th IEEE symposium on security and privacy, IEEE, 2009, pp. 173–187. 1

[PS17] Aaron Potechin and David Steurer, *Exact tensor completion with sum-of-squares*, Proceedings of the 2017 Conference on Learning Theory, 2017. 14

[RSS18] Prasad Raghavendra, Tselil Schramm, and David Steurer, *High dimensional estimation via sum-of-squares proofs*, Proceedings of the International Congress of Mathematicians: Rio de Janeiro 2018, World Scientific, 2018, pp. 3389–3423. 2, 4, 14

[SU19] Adam Sealfon and Jonathan Ullman, *Efficiently estimating erdos-renyi graphs with node differential privacy*, Advances in Neural Information Processing Systems **32** (2019). 1, 2, 3

A Sum-of-squares background

A.1 Sum-of-squares hierarchy

In this paper, we use the sum-of-squares semidefinite programming hierarchy [BS14, BS16, RSS18] for both algorithm design and analysis. The sum-of-squares proof-to-algorithm framework has been proven useful in many optimal or state-of-the-art results in algorithmic statistics [HL18, KSS18, PS17, Hop20]. We provide here a brief introduction to pseudo-distributions, sum-of-squares proofs, and sum-of-squares algorithms.

Pseudo-distribution. We can represent a finitely supported probability distribution over $\mathbb{R}^n$ by its probability mass function $\mu\colon \mathbb{R}^n \to \mathbb{R}$ such that $\mu \geqslant 0$ and $\sum_{x\in\mathrm{supp}(\mu)} \mu(x) = 1$. We define pseudo-distributions as generalizations of such probability mass distributions by relaxing the constraint $\mu \geqslant 0$ to only require that μ passes certain low-degree non-negativity tests.

Definition A.1 (Pseudo-distribution)**.** A *level-ℓ pseudo-distribution* μ over $\mathbb{R}^n$ is a finitely supported function $\mu : \mathbb{R}^n \to \mathbb{R}$ such that $\sum_{x\in\mathrm{supp}(\mu)} \mu(x) = 1$ and $\sum_{x\in\mathrm{supp}(\mu)} \mu(x) f(x)^2 \geqslant 0$ for every polynomial f of degree at most $\ell/2$.

We can define the expectation of a pseudo-distribution in the same way as the expectation of a finitely supported probability distribution.

Definition A.2 (Pseudo-expectation)**.** Given a pseudo-distribution μ over $\mathbb{R}^n$, we define the *pseudo-expectation* of a function $f : \mathbb{R}^n \to \mathbb{R}$ by

$$\tilde{\mathbb{E}}_{\mu} f := \sum_{x\in\mathrm{supp}(\mu)} \mu(x) f(x)\,. \tag{A.1}$$

The following definition formalizes what it means for a pseudo-distribution to satisfy a system of polynomial constraints.

Definition A.3 (Constrained pseudo-distributions)**.** Let $\mu : \mathbb{R}^n \to \mathbb{R}$ be a level-ℓ pseudo-distribution over $\mathbb{R}^n$. Let $\mathcal{A} = \{f_1 \geqslant 0, \ldots, f_m \geqslant 0\}$ be a system of polynomial constraints. We say that μ *satisfies* $\mathcal{A}$ at level r, denoted by $\mu \models_{r} \mathcal{A}$, if for every multiset $S \subseteq [m]$ and every sum-of-squares polynomial h such that $\deg(h) + \sum_{i\in S} \max\{\deg(f_i), r\} \leqslant \ell$,

$$\tilde{\mathbb{E}}_{\mu} h \cdot \prod_{i\in S} f_i \geqslant 0\,. \tag{A.2}$$

We say μ satisfies $\mathcal{A}$ and write $\mu \models \mathcal{A}$ (without further specifying the degree) if $\mu \models_{0} \mathcal{A}$.

We remark that if μ is an actual finitely supported probability distribution, then we have $\mu \models \mathcal{A}$ if and only if μ is supported on solutions to $\mathcal{A}$.

Sum-of-squares proof. We introduce sum-of-squares proofs as the dual objects of pseudo-distributions, which can be used to reason about properties of pseudo-distributions. We say a polynomial p is a sum-of-squares polynomial if there exist polynomials (q_i) such that $p = \sum_i q_i^2$.

Definition A.4 (Sum-of-squares proof)**.** A *sum-of-squares proof* that a system of polynomial constraints $\mathcal{A} = \{f_1 \geqslant 0, \ldots, f_m \geqslant 0\}$ implies $q \geqslant 0$ consists of sum-of-squares polynomials $(p_S)_{S\subseteq[m]}$ such that[6]

$$q = \sum_{\text{multiset } S\subseteq[m]} p_S \cdot \prod_{i\in S} f_i\,.$$

If such a proof exists, we say that $\mathcal{A}$ *(sos-)proves* $q \geqslant 0$ within degree ℓ, denoted by $\mathcal{A} \vdash_{\ell} q \geqslant 0$. In order to clarify the variables quantified by the proof, we often write $\mathcal{A}(x) \vdash^{x}_{\ell} q(x) \geqslant 0$.

[6]Here we follow the convention that $\prod_{i\in S} f_i = 1$ for $S = \emptyset$.

We say that the system $\mathcal{A}$ *sos-refuted* within degree ℓ if $\mathcal{A} \mathrel{|\!\!\frac{\;\;}{\ell}} -1 \geqslant 0$. Otherwise, we say that the system is *sos-consistent* up to degree ℓ, which also means that there exists a level-ℓ pseudo-distribution satisfying the system.

The following lemma shows that sum-of-squares proofs allow us to deduce properties of pseudo-distributions that satisfy some constraints.

Lemma A.5. *Let μ be a pseudo-distribution, and let $\mathcal{A}, \mathcal{B}$ be systems of polynomial constraints. Suppose there exists a sum-of-squares proof $\mathcal{A} \mathrel{|\!\!\frac{\;\;}{r'}} \mathcal{B}$. If $\mu \mathrel{|\!\!\frac{\;\;}{r}}\!\!= \mathcal{A}$, then $\mu \mathrel{|\!\!\frac{\;\;}{r \cdot r' + r'}}\!\!= \mathcal{B}$.*

Sum-of-squares algorithm. Given a system of polynomial constraints, the *sum-of-squares algorithm* searches through the space of pseudo-distributions that satisfy this polynomial system by semidefinite programming.

Since semidefinite programing can only be solved approximately, we can only find pseudo-distributions that approximately satisfy a given polynomial system. We say that a level-ℓ pseudo-distribution *approximately satisfies* a polynomial system, if the inequalities in Eq. (A.2) are satisfied up to an additive error of $2^{-n^\ell} \cdot \|h\| \cdot \prod_{i \in S} \|f_i\|$, where $\|\cdot\|$ denotes the Euclidean norm[7] of the coefficients of a polynomial in the monomial basis.

Theorem A.6 (Sum-of-squares algorithm)**.** *There exists an $(n+m)^{O(\ell)}$-time algorithm that, given any explicitly bounded[8] and satisfiable system[9] $\mathcal{A}$ of m polynomial constraints in n variables, outputs a level-ℓ pseudo-distribution that satisfies $\mathcal{A}$ approximately.*

Remark A.7 (Approximation error and bit complexity)*.* For a pseudo-distribution that only approximately satisfies a polynomial system, we can still use sum-of-squares proofs to reason about it in the same way as Lemma A.5. In order for approximation errors not to amplify throughout reasoning, we need to ensure that the bit complexity of the coefficients in the sum-of-squares proof are polynomially bounded.

A.2 Useful sum-of-squares lemmas

Lemma A.8.

$$\{x^2 = x\} \mathrel{|\!\!\frac{x}{2}} 0 \leqslant x \leqslant 1\,.$$

Proof. The first inequality is trivial due to $\{x^2 = x\} \mathrel{|\!\!\frac{x}{2}} x = x^2 \geqslant 0$. For the second inequality, it follows that

$$\{x^2 = x\} \mathrel{|\!\!\frac{x}{2}} x \leqslant \frac{x^2}{2} + \frac{1}{2} = \frac{x}{2} + \frac{1}{2}\,.$$

Rearranging the terms, we get

$$\{x^2 = x\} \mathrel{|\!\!\frac{x}{2}} x \leqslant 1\,.$$

□

Lemma A.9.

$$\{x^2 = x, y^2 = y\} \mathrel{|\!\!\frac{x,y}{4}} 1 - xy \leqslant (1-x) + (1-y)\,.$$

Proof. By Lemma A.8, it follows that

$$\{x^2 = x, y^2 = y\} \mathrel{|\!\!\frac{x,y}{2}} 0 \leqslant x, y \leqslant 1\,.$$

Therefore, we have

$$\{x^2 = x, y^2 = y\} \mathrel{|\!\!\frac{x,y}{4}} (1-y)(1-x) \geqslant 0$$

[7] The choice of norm is not important here because the factor 2^{-n^ℓ} swamps the effects of choosing another norm.

[8] A system of polynomial constraints is *explicitly bounded* if it contains a constraint of the form $\|x\|^2 \leqslant M$.

[9] Here we assume that the bit complexity of the constraints in $\mathcal{A}$ is $(n+m)^{O(1)}$.

$$\overset{x,y}{\underset{4}{\vdash}} 1 - x - y \geqslant -xy$$
$$\overset{x,y}{\underset{4}{\vdash}} 2 - x - y \geqslant 1 - xy\,.$$

□

Lemma A.10. *Given constant C, we have*

$$\{-C \leqslant x \leqslant C\} \overset{x}{\underset{2}{\vdash}} x^2 \leqslant C^2\,.$$

Proof.

$$\begin{aligned}\{-C \leqslant x \leqslant C\} &\overset{x}{\underset{2}{\vdash}} (C-x)(C+x) \geqslant 0\\ &\overset{x}{\underset{2}{\vdash}} C^2 - x^2 \geqslant 0\\ &\overset{x}{\underset{2}{\vdash}} C^2 \geqslant x^2\,.\end{aligned}$$

□

Lemma A.11. *Given constant C, we have*

$$\{x^2 \leqslant C^2\} \overset{x}{\underset{2}{\vdash}} -C \leqslant x \leqslant C\,.$$

Proof. For the first inequality, we have

$$\{x^2 \leqslant C^2\} \overset{x}{\underset{2}{\vdash}} x \geqslant -\frac{x^2}{2C} - \frac{C}{2} \geqslant -\frac{C^2}{2C} - \frac{C}{2} = -C\,.$$

For the second inequality, we have

$$\{x^2 \leqslant C^2\} \overset{x}{\underset{2}{\vdash}} x \leqslant \frac{x^2}{2C} + \frac{C}{2} \leqslant \frac{C^2}{2C} + \frac{C}{2} = C\,.$$

□

B Concentration inequalities

Lemma B.1 (Average degree concentration)**.** *Let Q° be an n-by-n edge connection probability matrix and let $d^\circ := d(Q^\circ)$. Let $\boldsymbol{A} \sim \mathbb{G}(n, Q^\circ)$. Then for any $\delta \in (0,1)$,*

$$\mathbb{P}(|d(\boldsymbol{A}) - d^\circ| \geqslant \delta d^\circ) \leqslant 2\exp\left(-\frac{\delta^2 n d^\circ}{6}\right),$$

Proof. Let $\mu := \mathbb{E}\sum_{i<j} \boldsymbol{A}_{ij} = \sum_{i<j} p_{ij}$. Using Chernoff bound, for $\delta \in (0,1)$,

$$\mathbb{P}\left(\left|\sum_{i<j} \boldsymbol{A}_{ij} - \mu\right| \geqslant \delta\mu\right) \leqslant 2\exp\left(-\frac{\delta^2\mu}{3}\right),$$
$$\mathbb{P}(|d(\boldsymbol{A}) - d^\circ| \geqslant \delta d^\circ) \leqslant 2\exp\left(-\frac{\delta^2 n d^\circ}{6}\right).$$

□

Lemma B.2 (Degree distribution)**.** *Let Q° be an n-by-n edge connection probability matrix. Let d be a parameter such that $d \geqslant 5$ and $\|Q^\circ\|_\infty \leqslant d/n$. Then for every $t \in [2e^2, \log n]$, an inhomogeneous random graph $\mathbb{G}(n, Q^\circ)$ has at least $e^{-t}n$ nodes with degree at least td with probability at most* $\exp(-te^{-t}nd/4)$.

Proof. Let $\boldsymbol{m}_k$ denote the number of nodes with degree at least k in $\mathbb{G}(n, Q^\circ)$. Then for every $\gamma \in [0,1]$,

$$\begin{aligned}\mathbb{P}(\boldsymbol{m}_{td} \geqslant \gamma n) &\leqslant \binom{n}{\gamma n}\binom{\gamma n^2}{\gamma ntd/2}\left(\frac{d}{n}\right)^{\gamma ntd/2} \\ &\leqslant \left(\frac{e}{\gamma}\right)^{\gamma n}\left(\frac{2e}{t}\right)^{\gamma ntd/2} \\ &= \exp\left(-\gamma n\left(\frac{td}{2}\log\frac{t}{2e} - \log\frac{e}{\gamma}\right)\right).\end{aligned}$$

Plugging in $\gamma = e^{-t}$ gives

$$\mathbb{P}\big(\boldsymbol{m}_{td} \geqslant e^{-t}n\big) \leqslant \exp\left(-te^{-t}n\left(\frac{d}{2}\log\frac{t}{2e} - 1 - 1/t\right)\right).$$

For $t \in [2e^2, \log n]$ and $d \geqslant 5$,

$$\mathbb{P}\big(\boldsymbol{m}_{td} \geqslant e^{-t}n\big) \leqslant \exp\left(-te^{-t}n\left(\frac{d}{2} - \frac{5}{4}\right)\right) \leqslant \exp\big(-te^{-t}nd/4\big)\,.$$

□

Lemma B.3 (Degree pruning)**.** *Let Q° be an n-by-n edge connection probability matrix. Let d be a parameter such that $d \geqslant 5$ and $\|Q^\circ\|_\infty \leqslant d/n$. Then with probability at least $1 - n^{1-d/4}$, an inhomogeneous graph $\mathbb{G}(n, Q^\circ)$ has the following property. For all $t \in [2e^2, \log n]$, the number of edges incident to nodes with degree at least td is at most $2te^{-t}nd$;*

Proof. Let $\boldsymbol{m}_k$ denote the number of nodes with degree at least k in $\mathbb{G}(n, Q^\circ)$. By Lemma B.2, for any $t \in [2e^2, \log n]$ and $d \geqslant 5$,

$$\mathbb{P}\big(\boldsymbol{m}_{td} \geqslant e^{-t}n\big) \leqslant \exp\big(-te^{-t}nd/4\big) \leqslant n^{-d/4}\,.$$

Applying union bound, the event that $m_k \leqslant e^{-k/d}n$ for any integer $k \in [2e^2 d, (\log n)d]$ happens with probability at least $1 - n^{1-d/4}$. We condition our following analysis on this event.

Fix a $t \in [2e^2, \log n]$. The number of edges incident to nodes with degree at least td is at most

$$\sum_{i=0}(t+i+1)d \cdot e^{-(t+i)}n = nd\sum_{i=t}(i+1)e^{-i} = nde^{-t}\left(\frac{e}{e-1}t + \frac{e^2}{(e-1)^2}\right) \leqslant 2te^{-t}nd\,.$$

□

Lemma B.4 (Degree-truncated subgraph)**.** *Let Q° be an n-by-n edge connection probability matrix and let $d^\circ := d(Q^\circ)$. Let d be a parameter such that $d \geqslant 5$ and $\|Q^\circ\|_\infty \leqslant d/n$. For $\delta \in (0,1)$, an inhomogeneous graph $\boldsymbol{A} \sim \mathbb{G}(n, Q^\circ)$ has the following property with probability at least $1 - n^{1-d/4} - \exp(-\delta^2 nd^\circ/6)$. For every $t \in [2e^2, \log n]$, A contains an n-node subgraph $\tilde{A}$ of such that*

- *$(\tilde{A}\mathbb{1})_i \leqslant td$ for any $i \in [n]$;*
- *$(1-\delta)d^\circ - 4te^{-t}d \leqslant d(\tilde{A}) \leqslant (1+\delta)d^\circ$.*

Proof. By Lemma B.1 and Lemma B.3, $\boldsymbol{A} \sim \mathbb{G}(n, Q^\circ)$ has the following two properties with probability at least $1 - n^{1-d/4} - \exp(-\delta^2 nd^\circ/6)$.

- $|d(A) - d^\circ| \leqslant \delta d^\circ$.
- For all $t \in [2e^2, \log n]$, the number of edges incident to nodes with degree at least td is at most $2te^{-t}nd$.

Consider a graph A with the above two properties. Fix a $t \in [2e^2, \log n]$. By removing at most $2te^{-t}nd$ edges from A, we can obtain a graph $\tilde{A}$ such that the maximum degree of $\tilde{A}$ is at most td. Moreover,

$$\begin{aligned} d(\tilde{A}) &\geqslant d(A) - 4te^{-t}d \geqslant (1-\delta)d^\circ - 4te^{-t}d\,, \\ d(\tilde{A}) &\leqslant d(A) \leqslant (1+\delta)d^\circ\,. \end{aligned}$$

□

Lemma B.5 (Spectral bound [BvH16]). *Let $\boldsymbol{A} \sim \mathbb{G}(n, p_0)$ and suppose $np_0 \geqslant 5$. Then with probability at least $1 - n^{-\Omega(1)}$,*

$$\left\| \boldsymbol{A} - p_0(\mathbb{1}\mathbb{1}^\top - \mathrm{Id}) \right\|_{\mathrm{op}} \leqslant O\left(\sqrt{np_0 \log n}\right).$$

C Sum-of-squares exponential mechanism

In this section, we present our sum-of-squares exponential mechanism and prove its properties in a general setting that incorporates all special cases in Appendix D, Appendix E and Appendix F.

Setup. Let $\mathcal{D} \subset \mathbb{R}^N$. Given an n-by-n symmetric matrix A, our goal is to output an element d from $\mathcal{D}$ privately. We say two symmetric matrices are neighboring if they differ in at most one row and one column. The utility of an element $d \in \mathcal{D}$ is quantified by a score function defined as follows.

Score function. For an n-by-n symmetric matrix A and a scalar γ, consider the following polynomial system with indeterminates $(Y_{ij})_{i,j\in[n]}$, $(z_i)_{i\in[n]}$ and coefficients that depend on A, γ:

$$\mathcal{Q}_1(Y, z; A, \gamma) := \begin{cases} z \odot z = z,\ \langle \mathbb{1}, z\rangle \geqslant (1-\gamma)n \\ 0 \leqslant Y \leqslant \mathbb{1}\mathbb{1}^\top,\ Y = Y^\top \\ Y \odot zz^\top = A \odot zz^\top \end{cases}. \tag{C.1}$$

For an element $d \in \mathcal{D}$, let $\mathcal{Q}_2(Y; d)$ be a polynomial system with coefficients depending on d (and independent of A, γ). Then for a matrix A and an element $d \in \mathcal{D}$, we define the score of d with regard to A to be

$$s(d; A) := \min_{0\leqslant\gamma\leqslant 1} \gamma n \text{ s.t. } \begin{cases} \exists \text{ level-}\ell \text{ pseudo-expectation } \tilde{\mathbb{E}} \text{ satisfying} \\ \mathcal{Q}_1(Y, z; A, \gamma) \cup \mathcal{Q}_2(Y; d)\,. \end{cases} \tag{C.2}$$

For $s(d; A)$ to be well-defined, we assume that for every $d \in \mathcal{D}$ there exists a symmetric matrix $A^* \in [0,1]^{n\times n}$ such that $\mathcal{Q}_2(A^*; d)$ is true.

Remark C.1 (Score function computation). Observe that a level-ℓ pseudo-expectation satisfying $\mathcal{Q}_1(Y, z; A, \gamma) \cup \mathcal{Q}_2(Y; d)$ is also a level-ℓ pseudo-expectation satisfying $\mathcal{Q}_1(Y, z; A, \gamma') \cup \mathcal{Q}_2(Y; d)$ for any $\gamma' \geqslant \gamma$. Thus we can compute $s(d; A)$ using binary search. Given a scalar γ, checking if there exists a level-ℓ pseudo-expectation satisfying $\mathcal{Q}_1(Y, z; A, \gamma) \cup \mathcal{Q}_2(Y; d)$ is equivalent to checking if a semidefinite program of size $n^{O(\ell)}$ is feasible. Since we only have efficient algorithms for semidefinite programming up to a given precision, we can only efficiently search for pseudo-distributions that *approximately* satisfy a given polynomial system. In spite of this, as long as the bit complexity of the coefficients in our sum-of-squares proof are polynomially bounded, the analysis of our algorithm based on sum-of-squares proofs will still work due to our discussion in Remark A.7. We refer interested readers to [HKMN23] for a formal (and quite technical) treatment of approximate pseudo-expectations.

Exponential mechanism. Given an n-by-n symmetric matrix A, our sos exponential mechanism with privacy parameter ε outputs a sample from the distribution $\mu_{A,\varepsilon}$ that is supported on $\mathcal{D}$ and has density

$$\mathrm{d}\mu_{A,\varepsilon}(d) \propto \exp(-\varepsilon \cdot s(d; A))\,. \tag{C.3}$$

Remark C.2 (Sampling). To efficiently sample from $\mu_{A,\varepsilon}$, we can use the following straightforward discretization scheme. More specifically, given a discretization parameter δ, we output an element $d \in \{0, \delta, 2\delta, \ldots, \lfloor n/\delta \rfloor \delta\}$ with probability proportional to $\exp(-\varepsilon \cdot \text{sos-score}(d; A))$. As the error introduced by discretization is at most δ and our target estimation error is $\omega(1/n)$, we can choose $\delta = 1/n$ and the discretization error is then negligible. Moreover, our algorithm requires at most n^2 evaluations of score functions.

Properties. The following lemma shows that the sensitivity of score function $s(d; A)$ is at most 1.

Lemma C.3 (Sensitivity bound)**.** *For any $d \in \mathcal{D}$ and any two n-by-n symmetric matrices A, A' that differ in at most one row and one column, the score function defined in Eq. (C.2) satisfies*

$$|s(d; A) - s(d; A')| \leqslant 1\,.$$

Proof. Without loss of generality, we assume that A and A' differ in the first row and column. Consider the linear functions (ℓ_i) where $\ell_1(z) = 0$ and $\ell_i(z) = z_i$ for $i \geqslant 2$. Then for every polynomial inequality $q(Y, z) \geqslant 0$ in $\mathcal{Q}_1(Y, z; A', \gamma + 1/n) \cup \mathcal{Q}_2(Y; d)$,

$$\mathcal{Q}_1(Y, z; A, \gamma) \cup \mathcal{Q}_2(Y; d) \,\Big|\frac{Y,z}{\deg(q)}\; q(Y, \ell(z)) \geqslant 0\,.$$

The same argument also holds for polynomial equalities. Then by [CDd+24, Lemma 8.1], $s(d; A') \leqslant s(d; A) + 1$. Due to symmetry of A and A', we also have $s(d; A) \leqslant s(d; A') + 1$. Therefore, $|s(d; A) - s(d; A')| \leqslant 1$. □

The following privacy guarantee of our sos exponential mechanism is a direct corollary of Lemma C.3.

Lemma C.4 (Privacy)**.** *The exponential mechanism defined in Eq. (C.3) is 2ε-differentially node private.*

Lemma C.5 (Volume of low-score points)**.** *Let $A \in \mathbb{R}^{n\times n}$ and $\varepsilon > 0$. Consider the distribution $\mu_{A,\varepsilon}$ defined by Eq. (C.3). Suppose $(Y = A^*, z = z^*)$ is a solution to $\mathcal{Q}_1(Y, z; A, \gamma^*)$. Then for any $t \geqslant 0$,*

$$\mathop{\mathbb{P}}_{\boldsymbol{d}\sim\mu_{A,\varepsilon}}\left(s(\boldsymbol{d}; A) \geqslant \gamma^* n + \frac{t \log n}{\varepsilon}\right) \leqslant \frac{\mathrm{vol}(\mathcal{D})}{\mathrm{vol}(\mathcal{G}(A^*))} \cdot n^{-t}\,,$$

where $\mathcal{G}(A^*) := \{d \in \mathcal{D} : \mathcal{Q}_2(A^*; d) \text{ is true}\}$.

Proof. Note $(Y = A^*, z = z^*)$ is also a solution to $\mathcal{Q}_1(Y, z; A, \gamma^*) \cup \mathcal{Q}_2(Y; d)$ for any d such that $\mathcal{Q}_2(A^*; d)$ is true. Let $\mathcal{G}(A^*) := \{d \in \mathcal{D} : \mathcal{Q}_2(A^*; d) \text{ is true}\}$. Thus $s(d; A) \leqslant \gamma^* n$ for any $d \in \mathcal{G}(A^*)$. For $t \geqslant 0$,

$$\mathop{\mathbb{P}}_{\boldsymbol{d}\sim\mu_{A,\varepsilon}}\left(s(\boldsymbol{d}; A) \geqslant \gamma^* n + \frac{t \log n}{\varepsilon}\right) \leqslant \frac{\mathrm{vol}(\mathcal{D}) \cdot \exp\big(-\varepsilon\gamma^* n - t \log n\big)}{\mathrm{vol}(\mathcal{G}(A^*)) \cdot \exp(-\varepsilon\gamma^* n)} = \frac{\mathrm{vol}(\mathcal{D})}{\mathrm{vol}(\mathcal{G}(A^*))} \cdot n^{-t}\,.$$

□

D Coarse estimation

In this section, we describe our coarse estimation algorithm that achieves constant multiplicative approximation of the expected average degree d°.

Theorem D.1 (Coarse estimation for inhomogeneous random graphs)**.** *Let Q° be an n-by-n edge connection probability matrix and let $d^\circ := d(Q^\circ)$. Suppose $\|Q^\circ\|_\infty \leqslant Rd^\circ/n$ for some R. There are constants C_1, C_2, C_3 such that the following holds. For any $\eta, \varepsilon, d^\circ$ such that $\eta \log(1/\eta) R \leqslant C_1$, $\varepsilon \geqslant C_2 \log^2(n) R/n$, and $d^\circ \geqslant C_3$, there exists a polynomial-time ε-differentially node private algorithm which, given an η-corrupted inhomogeneous random graph $\mathbb{G}(n, Q^\circ)$, outputs an estimate $\hat{d}$ satisfying $|\hat{d}/d^\circ - 1| \leqslant 0.5$ with probability $1 - n^{-\Omega(1)}$.*

We make a few remarks on Theorem D.1.

- Our algorithm in Theorem D.1 is a sum-of-squares exponential mechanism. R, η, ε are parameters given as inputs to our algorithm.
- We can get a constant estimate of p° by taking $\hat{p} = \frac{\hat{d}}{n-1}$. Since $\frac{\hat{p}}{p^\circ} = \frac{\hat{d}}{d^\circ}$, it follows that $|\frac{\hat{p}}{p^\circ} - 1| \leqslant 0.5$.
- When $Q^\circ = p^\circ(\mathbb{1}\mathbb{1}^\top - \mathrm{Id})$, the inhomogeneous random graph $\mathbb{G}(n, Q^\circ)$ is just the Erdős-Rényi random graph $\mathbb{G}(n, p^\circ)$. Thus, by setting $R = \frac{n}{n-1}$ in Theorem D.1, we directly obtain a coarse estimation result for Erdős-Rényi random graphs.
- The utility guarantee of our algorithm holds in the constant-degree regime (i.e. $d^\circ \geqslant \Omega(1)$). To the best of our knowledge, even without privacy requirement and in the special case of Erdős-Rényi random graphs, no previous algorithm can match our guarantees in the constant-degree regime. Specifically, when $d^\circ \ll \log n$ and $\eta \geqslant \Omega(1)$, the robust algorithm in [AJK+22] can not provide a constant-factor approximation of d°.

In Appendix D.1, we set up polynomial systems that our algorithm uses and prove useful sos inequalities. In Appendix D.2, we show that we can easily obtain a robust algorithm via sos proofs in Appendix D.1. Then in Appendix D.3, we describe our algorithm and prove Theorem D.1.

D.1 Sum-of-squares

For an adjacency matrix A and two nonnegative scalars γ and σ, consider the following polynomial systems with indeterminates $Y = (Y_{ij})_{i,j\in[n]}$, $z = (z_i)_{i\in[n]}$ and coefficients that depend on A, γ, σ:

$$\mathcal{P}_1(Y, z; A, \gamma) := \begin{cases} z \odot z = z,\ \langle \mathbb{1}, z\rangle \geqslant (1-\gamma)n \\ 0 \leqslant Y \leqslant \mathbb{1}\mathbb{1}^\top,\ Y = Y^\top \\ Y \odot zz^\top = A \odot zz^\top \end{cases}\Bigg\}, \tag{D.1}$$

$$\mathcal{P}_2(Y; \sigma) := \begin{cases} d(Y) = \langle Y, \mathbb{1}\mathbb{1}^\top\rangle / n \\ (Y\mathbb{1})_i \leqslant \sigma d(Y) \qquad \forall i \in [n] \end{cases}\Bigg\}. \tag{D.2}$$

For convenience of notation, we will consider the following combined polynomial system in remaining of the section

$$\mathcal{C}(Y, z; A, \gamma, \sigma) := \mathcal{P}_1(Y, z; A, \gamma) \cup \mathcal{P}_2(Y; \sigma). \tag{D.3}$$

Lemma D.2. *If (A^*, z^*) is a feasible solution to $\mathcal{C}(Y, z; A, \gamma^*, \sigma)$ and $1 - 2\gamma\sigma - 2\gamma^*\sigma > 0$, then it follows that*

$$\mathcal{C}(Y, z; A, \gamma, \sigma) \Big|\frac{Y,z}{8}\ (1 - 2\gamma\sigma - 2\gamma^*\sigma) d(A^*) \leqslant d(Y) \leqslant \frac{1}{1 - 2\gamma\sigma - 2\gamma^*\sigma} d(A^*).$$

Proof. Let $w = z \odot z^*$, by constraint $Y \odot zz^\top = A \odot zz^\top$ and $A^* \odot z^*(z^*)^\top = A \odot z^*(z^*)^\top$, we have

$$\begin{aligned} \mathcal{C} \Big|\frac{Y,z}{4}\ Y \odot ww^\top &= Y \odot zz^\top \odot z^*(z^*)^\top \\ &= A \odot zz^\top \odot z^*(z^*)^\top \\ &= A \odot z^*(z^*)^\top \odot zz^\top \\ &= A^* \odot z^*(z^*)^\top \odot zz^\top \\ &= A^* \odot ww^\top. \end{aligned}$$

Applying this equality, it follows that

$$\begin{aligned} \mathcal{C} \Big|\frac{Y,z}{4}\ n \cdot d(Y) &= \left\langle Y, \mathbb{1}\mathbb{1}^\top \right\rangle \\ &= \left\langle Y, ww^\top \right\rangle + \left\langle Y, \mathbb{1}\mathbb{1}^\top - ww^\top \right\rangle \end{aligned}$$

$$= \left\langle A^*, ww^\top \right\rangle + \left\langle Y, 2(\mathbb{1} - w)\mathbb{1}^\top \right\rangle - \left\langle Y, (\mathbb{1} - w)(\mathbb{1} - w)^\top \right\rangle .$$

For the first term, since $A^*_{i,j} \in [0,1]$, $z^*_i \in \{0,1\}$ and $\mathcal{C} \,\Big|\frac{Y,z}{2}\, 0 \leqslant z_i \leqslant 1$ for all $i, j \in [n]$, we have

$$\mathcal{C} \,\Big|\frac{Y,z}{4}\, A^*_{i,j} w_i w_j = A^*_{i,j} z^*_i z^*_j z_i z_j \leqslant A^*_{i,j} .$$

Therefore, it follows that

$$\mathcal{C} \,\Big|\frac{Y,z}{4}\, \left\langle A^*, ww^\top \right\rangle \leqslant \left\langle A^*, \mathbb{1}\mathbb{1}^\top \right\rangle \leqslant n \cdot d(A^*) . \tag{D.4}$$

For the second term, we have

$$\begin{aligned}
\mathcal{C} \,\Big|\frac{Y,z}{4}\, \left\langle Y, 2(\mathbb{1} - w)\mathbb{1}^\top \right\rangle &= \langle Y\mathbb{1}, 2(\mathbb{1} - w) \rangle \\
&= \sum_{i\in[n]} 2(1 - w_i) \cdot (Y\mathbb{1})_i \\
&= \sum_{i\in[n]} 2(1 - z_i z^*_i) \cdot (Y\mathbb{1})_i \\
&\leqslant \sum_{i\in[n]} 2(1 - z_i) \cdot (Y\mathbb{1})_i + \sum_{i\in[n]} 2(1 - z^*_i) \cdot (Y\mathbb{1})_i ,
\end{aligned}$$

where the last inequality is due to Lemma A.9. From constraints $\sum_{i\in[n]} 1 - z^*_i \leqslant \gamma^* n$, $\sum_{i\in[n]} 1 - z_i \leqslant \gamma n$ and $(Y\mathbb{1})_i \leqslant \sigma d(Y)$ for all $i \in [n]$, it follows that

$$\begin{aligned}
\mathcal{C} \,\Big|\frac{Y,z}{4}\, \left\langle Y, 2(\mathbb{1} - w)\mathbb{1}^\top \right\rangle &\leqslant \sum_{i\in[n]} 2(1 - z_i) \cdot \sigma d(Y) + \sum_{i\in[n]} 2(1 - z^*_i) \cdot \sigma d(Y) \\
&= 2\sigma d(Y) \cdot \left(\sum_{i\in[n]} 1 - z_i \right) + 2\sigma d(Y) \cdot \left(\sum_{i\in[n]} 1 - z^*_i \right) \\
&\leqslant 2\gamma n \sigma d(Y) + 2\gamma^* n \sigma d(Y) .
\end{aligned} \tag{D.5}$$

For the third term, since $\mathcal{C} \,\Big|\frac{Y,z}{2}\, Y_{i,j} \geqslant 0$ and $\mathcal{C} \,\Big|\frac{Y,z}{2}\, 1 - w_i \geqslant 0$ for all $i, j \in [n]$, it follows that

$$\mathcal{C} \,\Big|\frac{Y,z}{8}\, \left\langle Y, (\mathbb{1} - w)(\mathbb{1} - w)^\top \right\rangle \geqslant 0 . \tag{D.6}$$

Combining Eq. (D.4), Eq. (D.5) and Eq. (D.6), we can get

$$\begin{aligned}
\mathcal{C} \,\Big|\frac{Y,z}{8}\, & n \cdot d(Y) \leqslant n \cdot d(A^*) + 2\gamma n \sigma d(Y) + 2\gamma^* n \sigma d(Y) \\
\Big|\frac{Y,z}{8}\, & d(Y) \leqslant \frac{d(A^*)}{1 - 2\gamma\sigma - 2\gamma^*\sigma} .
\end{aligned}$$

Swapping the roll of A^* and Y, we can use the same proof to get

$$\begin{aligned}
\mathcal{C} \,\Big|\frac{Y,z}{8}\, & n \cdot d(A^*) \leqslant n \cdot d(Y) + 2\gamma n \sigma d(A^*) + 2\gamma^* n \sigma d(A^*) \\
\Big|\frac{Y,z}{8}\, & (1 - 2\gamma\sigma - 2\gamma^*\sigma) d(A^*) \leqslant d(Y) .
\end{aligned}$$

This completes the proof. □

Lemma D.3. *Let Q° be an n-by-n edge connection probability matrix and $d^\circ := d(Q^\circ)$. Suppose $\|Q^\circ\|_\infty \leqslant Rd^\circ/n$ for $R \in \mathbb{R}$. Let A be an η-corrupted adjacency matrix of a random graph $\boldsymbol{G}^\circ \sim \mathbb{G}(n, Q^\circ)$. Suppose $\eta \log(1/\eta) R \leqslant C_1$ for some constant C_1 that is small enough. With probability $1 - n^{-\Omega(1)}$, there exists A^* and z^* such that*

1. *$|d(A^*) - d^\circ| \leqslant 0.1 d^\circ$.*
2. *(A^*, z^*) is a feasible solution to $\mathcal{C}(Y, z; A, \gamma, \sigma)$ with $\gamma = 2\eta$ and $\sigma = 2\log(1/\eta)R$.*

Proof. Let $\boldsymbol{A}^\circ$ be the adjacency matrix of $\boldsymbol{G}^\circ$ and $z^\circ \in \{0,1\}^n$ denote the set of uncorrupted nodes ($z_i^\circ = 1$ if and only if node i is uncorrupted).

By Lemma B.2 and Lemma B.3, we know that, with probability $1 - n^{-\Omega(1)}$, there exists a degree-pruned adjacency matrix $\tilde{A}$ such that

1. $\|\tilde{A}\mathbb{1}\|_\infty \leqslant \log(1/\eta)Rd^\circ$.
2. At most ηn nodes are pruned.
3. At most $2\eta\log(1/\eta)nRd^\circ$ edges are pruned.

Let $\tilde{z} \in \{0,1\}^n$ denote the set of unpruned nodes ($z_i^\circ = 1$ if and only if node i is not pruned). We will show that $A^* = \tilde{A}$ and $z^* = z^\circ \odot \tilde{z}$ satisfies the lemma.

Guarantee 1. By Lemma B.1, we know that, with probability $1 - n^{-\Omega(1)}$,

$$|d(\boldsymbol{A}^\circ) - d^\circ| \leqslant 10\sqrt{\frac{d^\circ \log n}{n}}\,. \tag{D.7}$$

From degree pruning guarantee (3), we have that

$$|d(\tilde{A}) - d(\boldsymbol{A}^\circ)| \leqslant 2\eta\log(1/\eta)nRd^\circ\,. \tag{D.8}$$

Combining Eq. (D.7) and Eq. (D.8), for some constant C_1 that is small enough, we have

$$\begin{aligned}|d(\tilde{A}) - d^\circ| &\leqslant |d(\tilde{A}) - d(\boldsymbol{A}^\circ)| + |d(\boldsymbol{A}^\circ) - d^\circ| \\ &\leqslant 10\sqrt{\frac{d^\circ \log n}{n}} + 2\eta\log\frac{1}{\eta}Rd^\circ \\ &\leqslant 10\sqrt{\frac{\log n}{n}}d^\circ + 2C_1 d^\circ \\ &\leqslant 0.1 d^\circ\,.\end{aligned} \tag{D.9}$$

Guarantee 2. It is easy to check that $z^* \odot z^* = z^*$, $0 \leqslant A^* \leqslant \mathbb{1}\mathbb{1}^\top$ and $A^* = (A^*)^\top$. Since $\langle \mathbb{1}, \tilde{z}\rangle \geqslant 1 - \eta n$ by degree pruning condition (2) and $\langle \mathbb{1}, z^\circ\rangle \geqslant 1 - \eta n$ by corruption rate, it is easy to verify that

$$\langle \mathbb{1}, z^*\rangle \geqslant 1 - 2\eta n\,.$$

Moreover, we have $A^* \odot z^*(z^*)^\top = A \odot z^*(z^*)^\top$ due to

$$\tilde{A} \odot \tilde{z}\tilde{z}^\top \odot z^\circ(z^\circ)^\top = \boldsymbol{A}^\circ \odot \tilde{z}\tilde{z}^\top \odot z^\circ(z^\circ)^\top = \boldsymbol{A}^\circ \odot z^\circ(z^\circ)^\top \odot \tilde{z}\tilde{z}^\top = A \odot z^\circ(z^\circ)^\top \odot \tilde{z}\tilde{z}^\top\,.$$

From Eq. (D.9), we can get that $d^\circ \leqslant 2d(\tilde{A})$. Plugging this into degree pruning condition (1), we get

$$\|\tilde{A}\mathbb{1}\|_\infty \leqslant \log(1/\eta)Rd^\circ \leqslant 2\log(1/\eta)Rd(\tilde{A})\,.$$

Therefore, we have

$$(A^*\mathbb{1})_i \leqslant 2\log(1/\eta)Rd(A^*)\,.$$

for all $i \in [n]$.

Thus, (A^*, z^*) is a feasible solution to $\mathcal{C}(Y, z; A, \gamma, \sigma)$ with $\gamma = 2\eta$ and $\sigma = 2\log(1/\eta)R$. □

D.2 Robust algorithm

In this section, we show that the following algorithm based on sum-of-squares proofs in Appendix D.1 obtains a robust constant multiplicative approximation of d°.

Algorithm D.4 (Robust coarse estimation algorithm).
Input: η-corrupted adjacency matrix A, corruption fraction η and parameter R.

Algorithm: Obtain level-8 pseudo-expectation $\tilde{\mathbb{E}}$ by solving sum-of-squares relaxation of program $\mathcal{C}(Y,z;A,\gamma,\sigma)$ (defined in Eq. (D.3)) with A, $\gamma = 2\eta$ and $\sigma = 2\log(1/\eta)R$.

Output: $\tilde{\mathbb{E}}[d(Y)]$

Theorem D.5 (Robust coarse estimation). *Let Q° be an n-by-n edge connection probability matrix and let $d^\circ := d(Q^\circ)$. Suppose $\|Q^\circ\|_\infty \leqslant Rd^\circ/n$ for some R. Let A be an η-corrupted adjacency matrix of a random graph $\boldsymbol{G}^\circ \sim \mathbb{G}(n, Q^\circ)$. Suppose $\eta\log(1/\eta)R \leqslant c$ for some constant c that is small enough. With probability $1 - n^{-\Omega(1)}$, Algorithm D.4 outputs an estimate $\hat{d}$ satisfying $|\frac{\hat{d}}{d^\circ} - 1| \leqslant 0.5$.*

Proof. By Lemma D.2 and Lemma D.3, we know that

$$\mathcal{C}(Y,z;A,\gamma,\sigma) \,\Big|\frac{Y,z}{8}\, (1-4\gamma\sigma)d(A^*) \leqslant d(Y) \leqslant \frac{1}{1-4\gamma\sigma} d(A^*)\,,$$

and,

$$|d(A^*) - d^\circ| \leqslant 0.1 d^\circ\,.$$

Therefore, we have

$$\mathcal{C}(Y,z;A,\gamma,\sigma) \,\Big|\frac{Y,z}{8}\, 0.9(1-4\gamma\sigma)d^\circ \leqslant d(Y) \leqslant \frac{1.1}{1-4\gamma\sigma} d^\circ\,.$$

Consider $4\gamma\sigma$, for constant c that is small enough, we have

$$4\gamma\sigma = 8\eta\log(1/\eta)R \leqslant 8c \leqslant 0.1\,.$$

This implies that $0.9(1-4\gamma\sigma) \geqslant \frac{1}{2}$ and $\frac{1.1}{1-4\gamma\sigma} \leqslant \frac{11}{9} \leqslant \frac{3}{2}$. Therefore, we have

$$\mathcal{C}(Y,z;A,\gamma,\sigma) \,\Big|\frac{Y,z}{8}\, \frac{1}{2}d^\circ \leqslant d(Y) \leqslant \frac{3}{2}d^\circ\,.$$

Thus, the level-8 pseudo-expectation $\tilde{\mathbb{E}}$ satisfies

$$\frac{1}{2}d^\circ \leqslant \tilde{\mathbb{E}}[d(Y)] \leqslant \frac{3}{2}d^\circ\,,$$

which implies that

$$\left|\frac{\tilde{\mathbb{E}}[d(Y)]}{d^\circ} - 1\right| \leqslant \frac{1}{2}\,.$$

□

D.3 Private algorithm

In this section, we present our algorithm and prove Theorem D.1. Our algorithm instantiates the sum-of-squares exponential mechanism in Appendix C.

Score function. For an n-by-n symmetric matrix A and a scalar d, we define the score of d with regard to A to be

$$s(d;A) := \min_{0\leqslant\gamma\leqslant 1} \gamma n \text{ s.t.} \begin{cases} \exists \text{ level-8 pseudo-expectation } \tilde{\mathbb{E}} \text{ satisfying} \\ \mathcal{C}(Y,z;A,\gamma,\sigma) \cup \{|d(Y) - d| \leqslant \alpha d\}\,, \end{cases} \tag{D.10}$$

where $\mathcal{C}(Y,z;A,\gamma,\sigma)$ is the polynomial system defined in Eq. (D.3), and σ, α are fixed parameters whose values will be decided later. Note that $(Y = \frac{d}{n}\mathbb{1}\mathbb{1}^\top, z = \mathbb{0})$ is a solution to the polynomial system $\mathcal{C}(Y,z;A,1,\sigma) \cup \{|d(Y) - d| \leqslant \alpha d\}$ for any $A \in \mathbb{R}^{n\times n}$, $d \in [0,n]$, and $\sigma \geqslant 1$.

To efficiently compute $s(d;A)$, we can use the scheme as described in Remark C.1.

Exponential mechanism. Given a privacy parameter ε and an n-by-n symmetric matrix A, our algorithm is the exponential mechanism with score function Eq. (D.10) and range $[0, n]$.

Algorithm D.6 (Coarse estimation)**.**
Input: Graph A.

Parameters: $\varepsilon, \sigma, \alpha$.

Output: A sample from the distribution $\mu_{A,\varepsilon}$ with support $[0, n]$ and density

$$\mathrm{d}\mu_{A,\varepsilon}(d) \propto \exp(-\varepsilon \cdot s(d; A)), \tag{D.11}$$

where $s(d; A)$ is defined in Eq. (D.10).

To efficiently sample from $\mu_{A,\varepsilon}$, we can use the scheme as described in Remark C.2.

Privacy. The following privacy guarantee of our algorithm is a direct corollary of Lemma C.4.

Lemma D.7 (Privacy)**.** *Algorithm D.6 is 2ε-differentially node private.*

Utility. The utility guarantee of our algorithm is stated in the following lemma.

Lemma D.8 (Utility)**.** *Let Q° be an n-by-n edge connection probability matrix and let $d^\circ := d(Q^\circ)$. Suppose $\|Q^\circ\|_\infty \leqslant Rd^\circ/n$ for some R. There are constants C_1, C_2, C_3 such that the following holds. For any $\eta, \varepsilon, d^\circ$ such that $\eta \log(1/\eta)R \leqslant C_1$, $\varepsilon \geqslant C_2 \log^2(n)R/n$, and $d^\circ \geqslant C_3$, given an η-corrupted inhomogeneous random graph $\mathbb{G}(n, Q^\circ)$, Algorithm D.6 outputs an estimate $\hat{d}$ satisfying $|\hat{d} - d^\circ| \leqslant 0.5d^\circ$ with probability $1 - n^{-\Omega(1)}$.*

Before proving Lemma D.8, we need the following two lemmas.

Lemma D.9 (Volume of low-score points)**.** *Let $A \in \mathbb{R}^{n\times n}$ and $\varepsilon > 0$. Consider the distribution $\mu_{A,\varepsilon}$ defined by Eq. (D.11). Suppose $(Y = A^*, z = z^*)$ is a solution to $\mathcal{C}(Y, z; A, \gamma^*, \sigma)$ and $d(A^*) \geqslant 2$. Then for any $t \geqslant 0$,*

$$\mathop{\mathbb{P}}_{\boldsymbol{d}\sim\mu_{A,\varepsilon}}\left(s(\boldsymbol{d}; A) \geqslant \gamma^* n + \frac{t \log n}{\varepsilon}\right) \leqslant \frac{n^{-t+1}}{\alpha}.$$

Proof. Apply Lemma C.5 with $\mathcal{D} = [0, n]$ and

$$\mathcal{G}(A^*) = \left\{d \in \mathcal{D} \,:\, \frac{d(A^*)}{1+\alpha} \leqslant d \leqslant \frac{d(A^*)}{1-\alpha}\right\}.$$

As $[d(A^*)/(1+\alpha), d(A^*)] \subseteq \mathcal{G}(A^*)$ and $d(A^*) \geqslant 2 \geqslant 1+\alpha$, we have $\operatorname{vol}(\mathcal{G}(A^*)) \geqslant \alpha$. □

Lemma D.10 (Low score implies utility)**.** *Let $A \in \mathbb{R}^{n\times n}$ and consider the score function $s(\cdot\,; A)$ defined in Eq. (D.10). Suppose $(Y = A^*, z = z^*)$ is a solution to $\mathcal{C}(Y, z; A, \gamma^*, \sigma)$. For a scalar d such that $s(d; A) \leqslant \tau n$ and $(\gamma^* + \tau)\sigma \leqslant 0.1$,*

$$\frac{0.8}{1+\alpha} d(A^*) \leqslant d \leqslant \frac{1.25}{1-\alpha} d(A^*).$$

Proof. Applying Lemma D.2 with $(\gamma^* + \tau)\sigma \leqslant 0.1$, we have

$$\mathcal{C}(Y, z; A, \tau, \sigma) \left|\frac{Y,z}{8}\right. 0.8d(A^*) \leqslant d(Y) \leqslant 1.25d(A^*)$$

Thus,

$$\mathcal{C}(Y, z; A, \tau, \sigma) \cup \{|d(Y) - d| \leqslant \alpha d\} \left|\frac{Y,z}{8}\right. \frac{0.8}{1+\alpha} d(A^*) \leqslant d \leqslant \frac{1.25}{1-\alpha} d(A^*).$$

□

Now we are ready to prove Lemma D.8.

Proof of Lemma D.8. Let A be a realization of η-corrupted $\mathbb{G}(n, Q^\circ)$. By Lemma D.3, the following event happens with probability $1 - n^{-\Omega(1)}$. There exists a solution $(Y = A^*, z = z^*)$ to $\mathcal{C}(Y, z; A, \gamma^*, \sigma)$ with $\gamma^* = 2\eta$, $\sigma = 2\log(1/\eta)R$, and $0.9d^\circ \leqslant d(A^*) \leqslant 1.1d^\circ$.

As $d(A^*) \geqslant 0.9d^\circ \geqslant 2$, then it follows by setting $t = 10$ and $\alpha = 0.01$ in Lemma D.9 that,

$$\mathop{\mathbb{P}}_{\boldsymbol{d} \sim \mu_{A,\varepsilon}} (s(\boldsymbol{d}; A) \leqslant \tau n) \geqslant 1 - n^{-9} \text{ where } \tau := 2\eta + 10\log(n)/(\varepsilon n)\,.$$

As an η-corrupted graph is actually uncorrupted when $\eta < 1/n$, we can assume $\eta \geqslant 1/(2n)$ without loss of generality. Thus,

$$(2\eta + \tau)\sigma \leqslant 8\eta\log(1/\eta)R + \frac{20\log^2(n)R}{n\varepsilon}\,.$$

For $\eta\log(1/\eta)R$ and $\log^2(n)R/(\varepsilon n)$ smaller than some constant, we have $(2\eta+\tau)\sigma \leqslant 0.1$. Let $\hat{d}$ be a scalar such that $s(\hat{d}; A) \leqslant \tau n$. Then by Lemma D.10,

$$\frac{0.8}{1+\alpha} d(A^*) \leqslant \hat{d} \leqslant \frac{1.25}{1-\alpha} d(A^*)\,.$$

Plugging in $\alpha \leqslant 0.01$ and $0.9d^\circ \leqslant d(A^*) \leqslant 1.1d^\circ$, we have

$$0.5d^\circ \leqslant \hat{d} \leqslant 1.5d^\circ\,.$$

□

Proof of Theorem D.1. By Lemma D.7 and Lemma D.8.

E Fine estimation for inhomogeneous random graphs

From Appendix D, we have a constant multiplicative approximation of the expected average degree d°. In this section, we show how to use this coarse estimate to obtain our fine estimator for inhomogeneous random graphs.

Theorem E.1 (Fine estimation for inhomogeneous random graphs)**.** *Let Q° be an n-by-n edge connection probability matrix and let $d^\circ := d(Q^\circ)$. Suppose $\|Q^\circ\|_\infty \leqslant Rd^\circ/n$ for some R. There is a sufficiently small constant c such that the following holds. For any η such that $\eta\log(1/\eta)R \leqslant c$, there exists a polynomial-time ε-differentially node private algorithm which, given an η-corrupted inhomogeneous random graph $\mathbb{G}(n, Q^\circ)$ and a constant-factor approximation of d°, outputs an estimate $\tilde{d}$ satisfying*

$$\left|\frac{\tilde{d}}{d^\circ} - 1\right| \leqslant O\left(\sqrt{\frac{\log n}{d^\circ n}} + \frac{R\log^2 n}{\varepsilon n} + \eta\log(1/\eta)R\right),$$

with probability $1 - n^{-\Omega(1)}$.

We make a few remarks on Theorem E.1.

- Our algorithm in Theorem E.1 is a sum-of-squares exponential mechanism. R, η, ε are parameters given as input to our algorithm.
- We can get an estimate of p° by taking $\tilde{p} = \frac{\tilde{d}}{n-1}$. Since $\frac{\tilde{p}}{p^\circ} = \frac{\tilde{d}}{d^\circ}$, it follows that
$$\left|\frac{\tilde{p}}{p^\circ} - 1\right| \leqslant O\left(\sqrt{\frac{\log n}{d^\circ n}} + \frac{R\log^2 n}{\varepsilon n} + \eta\log(1/\eta)R\right).$$
- Combining Theorem D.1 and Theorem E.1 gives us an efficient, private, and robust edge density estimation algorithm for inhomogeneous random graphs whose utility guarantee is information-theoretically optimal up to a factor of $\log n$ and $\log(1/\eta)$.

In Appendix E.1, we set up polynomial systems that our algorithm uses and prove useful sos inequalities. In Appendix E.2, we show that we can easily obtain a robust algorithm via sos proofs in Appendix E.1. Then in Appendix E.3, we describe our algorithm and prove Theorem E.1.

E.1 Sum-of-squares

For an adjacency matrix A and nonnegative scalars γ, σ and $\hat{d}$, consider the following polynomial systems with indeterminates $Y = (Y_{ij})_{i,j\in[n]}$, $z = (z_i)_{i\in[n]}$ and coefficients that depend on $A, \gamma, \sigma, \hat{d}$:

$$\mathcal{P}_1(Y, z; A, \gamma) := \begin{Bmatrix} z \odot z = z,\ \langle \mathbb{1}, z\rangle \geqslant (1-\gamma)n \\ 0 \leqslant Y \leqslant \mathbb{1}\mathbb{1}^\top,\ Y = Y^\top \\ Y \odot zz^\top = A \odot zz^\top \end{Bmatrix}, \tag{E.1}$$

$$\mathcal{P}_3(Y; \sigma, \hat{d}) := \left\{ (Y\mathbb{1})_i \leqslant \sigma\hat{d} \quad \forall i \in [n] \right\}. \tag{E.2}$$

For convenience of notation, we will consider the following combined polynomial system in the remaining of this section

$$\mathcal{D}(Y, z; A, \gamma, \sigma, \hat{d}) := \mathcal{P}_1(Y, z; A, \gamma) \cup \mathcal{P}_3(Y; \sigma, \hat{d}). \tag{E.3}$$

Lemma E.2. *If (A^*, z^*) is a feasible solution to $\mathcal{D}(Y, z; A, \gamma^*, \sigma, \hat{d})$, then it follows that*

$$\mathcal{D}(Y, z; A, \gamma, \sigma, \hat{d}) \left|\frac{Y,z}{8}\right. |d(Y) - d(A^*)| \leqslant 2(\gamma + \gamma^*)\sigma\hat{d}.$$

Proof. Let $w = z \odot z^*$. Using similar analysis as in the proof of Lemma D.2, it follows that

$$\mathcal{D} \left|\frac{Y,z}{4}\right. Y \odot ww^\top = A^* \odot ww^\top,$$

and,

$$\begin{aligned} \mathcal{D} \left|\frac{Y,z}{8}\right. n \cdot d(Y) &= \left\langle Y, \mathbb{1}\mathbb{1}^\top \right\rangle \\ &= \left\langle A^*, ww^\top \right\rangle + \left\langle Y, 2(\mathbb{1} - w)\mathbb{1}^\top \right\rangle - \left\langle Y, (\mathbb{1} - w)(\mathbb{1} - w)^\top \right\rangle \\ &\leqslant \left\langle A^*, \mathbb{1}\mathbb{1}^\top \right\rangle + \langle Y\mathbb{1}, 2(\mathbb{1} - w)\rangle \\ &\leqslant n \cdot d(A^*) + 2(\gamma + \gamma^*)\sigma n\hat{d}. \end{aligned}$$

By rearranging the terms, we have

$$\mathcal{D} \left|\frac{Y,z}{8}\right. d(Y) - d(A^*) \leqslant 2(\gamma + \gamma^*)\sigma\hat{d}.$$

Swapping the roll of Y and A^*, we can also get

$$\mathcal{D} \left|\frac{Y,z}{8}\right. d(A^*) - d(Y) \leqslant 2(\gamma + \gamma^*)\sigma\hat{d}.$$

This completes the proof. □

Lemma E.3. *Let Q° be an n-by-n edge connection probability matrix and $d^\circ := d(Q^\circ)$. Suppose $\|Q^\circ\|_\infty \leqslant Rd^\circ/n$ for $R \in \mathbb{R}$. Let A be an η-corrupted adjacency matrix of a random graph $\boldsymbol{G}^\circ \sim \mathbb{G}(n, Q^\circ)$. Suppose $\eta \log(1/\eta)R \leqslant C_1$ for some constant C_1 that is small enough. With probability $1 - n^{-\Omega(1)}$, there exists A^* and z^* such that*

- $|d(A^*) - d^\circ| \leqslant 10\sqrt{\frac{d^\circ \log n}{n}} + 2\eta\log(1/\eta)Rd^\circ$.
- *(A^*, z^*) is a feasible solution to $\mathcal{D}(Y, z; A, \gamma, \sigma, \hat{d})$ with η-corrupted A, $\gamma = 2\eta$, $\sigma = 10\log(1/\eta)R$ and $\hat{d} \geqslant \frac{1}{2}d^\circ$.*

Proof. Let $\boldsymbol{A}^\circ$ be the adjacency matrix of $\boldsymbol{G}^\circ$ and $z^\circ \in \{0,1\}^n$ denote the set of uncorrupted nodes ($z_i^\circ = 1$ if and only if node i is uncorrpted).

By Lemma B.2 and Lemma B.3, we know that, with probability $1 - n^{-\Omega(1)}$, there exists a degree-pruned adjacency matrix $\tilde{A}$ such that

1. $\left\|\tilde{A}\mathbb{1}\right\|_\infty \leqslant \log(1/\eta)Rd^\circ$.

2. At most ηn nodes are pruned.

3. At most $2\eta \log(1/\eta) n R d^\circ$ edges are pruned.

Let $\tilde{z} \in \{0,1\}^n$ denote the set of unpruned nodes ($\tilde{z}_i = 1$ if and only if node i is not pruned). We will show that $A^* = \tilde{A}$ and $z^* = z^\circ \odot \tilde{z}$ satisfies the lemma.

Guarantee 1. By Lemma B.1, we know that, with probability $1 - n^{-\Omega(1)}$,

$$|d(\boldsymbol{A}^\circ) - d^\circ| \leqslant 10\sqrt{\frac{d^\circ \log n}{n}} \,. \tag{E.4}$$

From degree pruning guarantee (3), we have that

$$|d(\tilde{A}) - d(\boldsymbol{A}^\circ)| \leqslant 2\eta \log(1/\eta) R d^\circ \,. \tag{E.5}$$

Combining Eq. (E.4) and Eq. (E.5), we have

$$\begin{aligned} |d(\tilde{A}) - d^\circ| &\leqslant |d(\tilde{A}) - d(A^\circ)| + |d(A^\circ) - d^\circ| \\ &\leqslant 10\sqrt{\frac{d^\circ \log n}{n}} + 2\eta \log(1/\eta) R d^\circ \,. \end{aligned}$$

Guarantee 2. It is easy to check that $z^* \odot z^* = z^*$, $0 \leqslant A^* \leqslant \mathbb{1}\mathbb{1}^\top$ and $A^* = (A^*)^\top$. Since $\langle \mathbb{1}, \tilde{z}\rangle \geqslant 1 - \eta n$ by degree pruning condition (2) and $\langle \mathbb{1}, z^\circ \rangle \geqslant 1 - \eta n$ by corruption rate, it is easy to verify that

$$\langle \mathbb{1}, z^* \rangle \geqslant 1 - 2\eta n \,.$$

Moreover, we have $A^* \odot z^*(z^*)^\top = A \odot z^*(z^*)^\top$ due to

$$\tilde{A} \odot \tilde{z}\tilde{z}^\top \odot z^\circ(z^\circ)^\top = \boldsymbol{A}^\circ \odot \tilde{z}\tilde{z}^\top \odot z^\circ(z^\circ)^\top = \boldsymbol{A}^\circ \odot z^\circ(z^\circ)^\top \odot \tilde{z}\tilde{z}^\top = A \odot z^\circ(z^\circ)^\top \odot \tilde{z}\tilde{z}^\top \,.$$

By degree pruning condition (1), we have

$$(A^* \mathbb{1})_i \leqslant \log(1/\eta) R d^\circ \leqslant \sigma \hat{d} \,.$$

for all $i \in [n]$.

Thus, (A^*, z^*) is a feasible solution to $\mathcal{D}(Y, z; A, \gamma, \sigma, \hat{d})$ with $\gamma = 2\eta$, $\sigma = 10 \log(1/\eta) R$ and $\hat{d} \geqslant \frac{1}{2} d^\circ$. □

E.2 Robust algorithm

In this section, we show that the following algorithm based on sum-of-squares proofs in Appendix E.1 obtains a robust approximation of d° that is optimal up to logarithmic factors.

Algorithm E.4 (Robust fine estimation algorithm for inhomogeneous random graphs)**.**
Input: η-corrupted adjacency matrix A, corruption fraction η and parameter R.

Algorithm:

1. Obtain coarse estimator $\hat{d}$ by applying Algorithm D.4 with A, η, R as input.

2. Obtain level-8 pseudo-expectation $\tilde{\mathbb{E}}$ by solving sum-of-squares relaxation of program $\mathcal{D}(Y, z; A, \gamma, \sigma, \hat{d})$ (defined in Eq. (E.3)) with A, $\gamma = 2\eta$, $\sigma = 10 \log(1/\eta) R$ and $\hat{d}$.

Output: $\tilde{\mathbb{E}}[d(Y)]$

Theorem E.5 (Robust fine estimation for inhomogeneous random graphs)**.** *Let Q° be an n-by-n edge connection probability matrix and let $d^\circ := d(Q^\circ)$. Suppose $\|Q^\circ\|_\infty \leqslant R d^\circ / n$ for some R. Let A be an η-corrupted adjacency matrix of a random graph $\boldsymbol{G}^\circ \sim \mathbb{G}(n, Q^\circ)$. Suppose*

$\eta\log(1/\eta)R \leqslant c$ for some constant c that is small enough. With probability $1-n^{-\Omega(1)}$, Algorithm E.4 outputs an estimate $\tilde{d}$ satisfying

$$\left|\frac{\tilde{d}}{d^\circ}-1\right| \leqslant O\left(\sqrt{\frac{\log n}{d^\circ n}}+\eta\log(1/\eta)R\right).$$

Proof. By Theorem D.5, we have $\frac{1}{2}d^\circ \leqslant \hat{d} \leqslant \frac{3}{2}d^\circ$. Let $\gamma^* = 2\eta$, by Lemma E.2 and Lemma E.3, it follows that

$$\begin{aligned}\mathcal{D}(Y,z;A,\gamma,\sigma,\hat{d}) \left|\frac{Y,z}{8}\right. |d(Y)-d(A^*)| &\leqslant 2(\gamma+\gamma^*)\sigma\hat{d} \\ &\leqslant 2\cdot 4\eta\cdot 10\log(1/\eta)R\cdot\frac{3}{2}d^\circ \\ &= 120\eta\log(1/\eta)Rd^\circ\,.\end{aligned}$$

and,

$$|d(A^*)-d^\circ| \leqslant 10\sqrt{\frac{d^\circ\log n}{n}}+2\eta\log(1/\eta)Rd^\circ\,.$$

Therefore, we have

$$\mathcal{D}(Y,z;A,\gamma,\sigma,\hat{d}) \left|\frac{Y,z}{O(1)}\right. |d(Y)-d^\circ| \leqslant 200\eta\log(1/\eta)Rd^\circ+10\sqrt{\frac{d^\circ\log n}{n}}\,.$$

Thus, the level-8 pseudo-expectation $\tilde{\mathbb{E}}$ satisfies

$$\left|\tilde{\mathbb{E}}[d(Y)]-d^\circ\right| \leqslant 200\eta\log(1/\eta)Rd^\circ+10\sqrt{\frac{d^\circ\log n}{n}}\,,$$

which implies that

$$\left|\frac{\tilde{\mathbb{E}}[d(Y)]}{d^\circ}-1\right| \leqslant O\left(\sqrt{\frac{\log n}{d^\circ n}}+\eta\log(1/\eta)R\right).$$

□

E.3 Private algorithm

In this section, we present our algorithm and prove Theorem E.1. Our algorithm instantiates the sum-of-squares exponential mechanism in Appendix C.

Score function. For an n-by-n symmetric matrix A and a scalar d, we define the score of d with regard to A to be

$$s(d;A) := \min_{0\leqslant\gamma\leqslant 1}\gamma n \text{ s.t. } \begin{cases}\exists \text{ level-8 pseudo-expectation } \tilde{\mathbb{E}} \text{ satisfying} \\ \mathcal{D}(Y,z;A,\gamma,\sigma,\hat{d}) \cup \{|d(Y)-d| \leqslant \alpha d\}\,,\end{cases} \tag{E.6}$$

where $\mathcal{D}(Y,z;A,\gamma,\sigma,\hat{d})$ is the polynomial system defined in Eq. (E.3), $\hat{d}$ is a coarse estimate, and σ,α are fixed parameters whose values will be decided later. Note that $(Y=\frac{d}{n}\mathbb{1}\mathbb{1}^\top, z=\mathbb{0})$ is a solution to the polynomial system $\mathcal{D}(Y,z;A,1,\sigma,\hat{d})\cup\{|d(Y)-d|\leqslant\alpha d\}$ for any $A\in\mathbb{R}^{n\times n}$ and any d such that $0\leqslant d\leqslant\min\{\sigma\hat{d},n\}$.

To efficiently compute $s(d;A)$, we can use the scheme as described in Remark C.1.

Exponential mechanism. Given a privacy parameter ε and an n-by-n symmetric matrix A, our private algorithm in Theorem E.1 is the exponential mechanism with score function Eq. (E.6) and range $[0,\min\{\sigma\hat{d},n\}]$.

Algorithm E.6 (Fine estimation for inhomogeneous random graphs)**.**
Input: Graph A, coarse estimate $\hat{d}$.

Parameters: $\varepsilon, \sigma, \alpha$.

Output: A sample from the distribution $\mu_{A,\varepsilon}$ with support $[0, \min\{\sigma\hat{d}, n\}]$ and density

$$\mathrm{d}\mu_{A,\varepsilon}(d) \propto \exp(-\varepsilon \cdot s(d; A))\,, \tag{E.7}$$

where $s(d; A)$ is defined in Eq. (E.6).

To efficiently sample from $\mu_{A,\varepsilon}$, we can use the scheme as described in Remark C.2.

Privacy. The following privacy guarantee of our algorithm is a direct corollary of Lemma C.4.

Lemma E.7 (Privacy)**.** *Algorithm E.6 is 2ε-differentially node private.*

Utility. The utility guarantee of our algorithm is stated in the following lemma.

Lemma E.8 (Utility)**.** *Let Q° be an n-by-n edge connection probability matrix and let $d^\circ := d(Q^\circ)$. Suppose $\|Q^\circ\|_\infty \leqslant Rd^\circ/n$ for some R. There is a sufficiently small constant c such that the following holds. For any η such that $\eta\log(1/\eta)R \leqslant c$, given an η-corrupted inhomogeneous random graph $\mathbb{G}(n, Q^\circ)$ and a coarse estimate $\hat{d}$ such that $0.5d^\circ \leqslant \hat{d} \leqslant 2d^\circ$, Algorithm E.6 outputs an estimate $\tilde{d}$ satisfying*

$$\left|\frac{\tilde{d}}{d^\circ} - 1\right| \leqslant O\left(\sqrt{\frac{\log n}{d^\circ n}} + \frac{R\log^2 n}{\varepsilon n} + \eta\log(1/\eta)R\right),$$

with probability $1 - n^{-\Omega(1)}$.

Before proving Lemma E.8, we need the following two lemmas.

Lemma E.9 (Volume of low-score points)**.** *Let $A \in \mathbb{R}^{n\times n}$ and $\varepsilon > 0$. Consider the distribution $\mu_{A,\varepsilon}$ defined by Eq. (E.7). Suppose $(Y = A^*, z = z^*)$ is a solution to $\mathcal{D}(Y, z; A, \gamma^*, \sigma, \hat{d})$ and $2 \leqslant d(A^*) \leqslant \sigma\hat{d}$. Then for any $t \geqslant 0$,*

$$\mathop{\mathbb{P}}_{\boldsymbol{d}\sim\mu_{A,\varepsilon}}\left(s(\boldsymbol{d}; A) \geqslant \gamma^* n + \frac{t\log n}{\varepsilon}\right) \leqslant \frac{n^{-t+1}}{\alpha}\,.$$

Proof. Apply Lemma C.5 with $\mathcal{D} = [0, \min\{\sigma\hat{d}, n\}]$ and

$$\mathcal{G}(A^*) = \left\{d \in \mathcal{D} \,:\, \frac{d(A^*)}{1+\alpha} \leqslant d \leqslant \frac{d(A^*)}{1-\alpha}\right\}.$$

As $[d(A^*)/(1+\alpha), d(A^*)] \subseteq \mathcal{G}(A^*)$ and $d(A^*) \geqslant 2 \geqslant 1+\alpha$, we have $\mathrm{vol}(\mathcal{G}(A^*)) \geqslant \alpha$. □

Lemma E.10 (Low score implies utility)**.** *Let $A \in \mathbb{R}^{n\times n}$ and consider the score function $s(\cdot\,; A)$ defined in Eq. (E.6). Suppose $(Y = A^*, z = z^*)$ is a solution to $\mathcal{D}(Y, z; A, \gamma^*, \sigma, \hat{d})$. For a scalar d such that $s(d; A) \leqslant \tau n$,*

$$\frac{d(A^*) - 2(\gamma^* + \tau)\sigma\hat{d}}{1+\alpha} \leqslant d \leqslant \frac{d(A^*) + 2(\gamma^* + \tau)\sigma\hat{d}}{1-\alpha}\,.$$

Proof. By Lemma E.2,

$$\mathcal{D}(Y, z; A, \tau, \sigma, \hat{d}) \;\Big|\frac{Y,z}{8}\; |d(Y) - d(A^*)| \leqslant 2(\gamma^* + \tau)\sigma\hat{d}\,.$$

Thus,

$$\begin{aligned} &\mathcal{D}(Y, z; A, \tau, \sigma, \hat{d}) \cup \{|d(Y) - d| \leqslant \alpha d\} \\ &\Big|\frac{Y,z}{8}\; \frac{d(A^*) - 2(\gamma^* + \tau)\sigma\hat{d}}{1+\alpha} \leqslant d \leqslant \frac{d(A^*) + 2(\gamma^* + \tau)\sigma\hat{d}}{1-\alpha}\,. \end{aligned}$$

□

Now we are ready to prove Lemma E.8.

Proof of Lemma E.8. Let A be a realization of η-corrupted $\mathbb{G}(n, Q^\circ)$. By Lemma E.3, the following event happens with probability at least $1 - n^{-\Omega(1)}$. There exists a solution $(Y = A^*, z = z^*)$ to $\mathcal{D}(Y, z; A, \gamma^*, \sigma, \hat{d})$ with $\gamma^* = 2\eta$, $\sigma = 10\log(1/\eta)R$, and

$$|d(A^*) - d^\circ| \leqslant 10\sqrt{d^\circ \log(n)/n} + 2\eta\log(1/\eta)Rd^\circ .$$

For $\eta\log(1/\eta)R$ smaller than some constant, we have $0.9d^\circ \leqslant d(A^*) \leqslant 1.1d^\circ$. Note that $d(A^*) \geqslant 0.9d^\circ \geqslant 2$ and $d(A^*) \leqslant 1.1d^\circ \leqslant \sigma\hat{d}$. Then it follows by setting $t = 10$ and $\alpha = n^{-2}$ in Lemma E.9 that,

$$\mathop{\mathbb{P}}_{\boldsymbol{d}\sim\mu_{A,\varepsilon}}(s(\boldsymbol{d}; A) \leqslant \tau n) \geqslant 1 - n^{-7} \text{ where } \tau := 2\eta + 10\log(n)/(\varepsilon n) .$$

Let $\tilde{d}$ be a scalar such that $s(\tilde{d}; A) \leqslant \tau n$. Then by Lemma E.10,

$$\frac{d(A^*) - 2(2\eta+\tau)\sigma\hat{d}}{1+\alpha} \leqslant \tilde{d} \leqslant \frac{d(A^*) + 2(2\eta+\tau)\sigma\hat{d}}{1-\alpha} .$$

Plugging in everything, we have

$$\left|\frac{\tilde{d}}{d^\circ} - 1\right| \leqslant O\left(\sqrt{\frac{\log n}{d^\circ n}} + \frac{R\log(1/\eta)\log n}{\varepsilon n} + R\eta\log(1/\eta)\right) .$$

As an η-corrupted graph is actually uncorrupted when $\eta < 1/n$, we can assume $\eta \geqslant 1/(2n)$ without loss of generality. Therefore,

$$\left|\frac{\tilde{d}}{d^\circ} - 1\right| \leqslant O\left(\sqrt{\frac{\log n}{d^\circ n}} + \frac{R\log^2 n}{\varepsilon n} + R\eta\log(1/\eta)\right) .$$

□

Proof of Theorem E.1. By Lemma E.7 and Lemma E.8.

F Fine estimation for Erdős-Rényi random graphs

From Appendix D, we have a a constant multiplicative approximation of the expected average degree d°. In this section, we show how to use this coarse estimate to obtain our fine estimate for Erdős-Rényi random graphs.

Theorem F.1 (Fine estimation for Erdős-Rényi random graphs)**.** *There are constants C_1, C_2, C_3 such that the following holds. For any $\eta \leqslant C_1$, $\varepsilon \geqslant C_2\log(n)/n$, and $d^\circ \geqslant C_3$, there exists a polynomial-time ε-differentially node private algorithm which, given an η-corrupted Erdős-Rényi random graph $\mathbb{G}(n, d^\circ/n)$ and a constant-factor approximation of d°, outputs an estimate $\tilde{d}$ satisfying*

$$\left|\frac{\tilde{d}}{d^\circ} - 1\right| \leqslant O\left(\sqrt{\frac{\log n}{d^\circ n}} + \frac{\log^2 n}{\sqrt{d^\circ}\varepsilon n} + \frac{\eta\log n}{\sqrt{d^\circ}}\right),$$

with probability $1 - n^{-\Omega(1)}$.

We make a few remarks on Theorem F.1.

- Our algorithm in Theorem F.1 is an sum-of-squares exponential mechanism. R, η, ε are parameters given as input to our algorithm.
- We can get an estimate of p° by taking $\hat{p} = \frac{\hat{d}}{n-1}$. Since $\frac{\hat{p}}{p^\circ} = \frac{\hat{d}}{d^\circ}$, it follows that

$$\left|\frac{\hat{p}}{p^\circ} - 1\right| \leqslant O\left(\sqrt{\frac{\log n}{d^\circ n}} + \frac{\log^2 n}{\sqrt{d^\circ}\varepsilon n} + \frac{\eta\log n}{\sqrt{d^\circ}}\right).$$

- Combining Theorem D.1 and Theorem F.1 gives us an efficient, private, and robust edge density estimation algorithm for Erdős-Rényi random graphs whose utility guarantee is information-theoretically optimal up to a factor of $\log n$.

In Appendix F.1, we set up polynomial systems that our algorithm uses and prove useful sos inequalities. In Appendix F.2, we show that we can easily obtain a robust algorithm via sos proofs in Appendix F.1. Then in Appendix F.3, we describe our algorithm and prove Theorem F.1.

F.1 Sum-of-squares

For an adjacency matrix A and nonnegative scalars γ, σ and $\hat{d}$, consider the following polynomial systems with indeterminates $Y = (Y_{ij})_{i,j\in[n]}$, $z = (z_i)_{i\in[n]}$ and coefficients that depend on $A, \gamma, \sigma, \delta, \hat{d}$:

$$\mathcal{P}_1(Y, z; A, \gamma) := \begin{Bmatrix} z \odot z = z,\ \langle \mathbb{1}, z\rangle \geqslant (1-\gamma)n \\ 0 \leqslant Y \leqslant \mathbb{1}\mathbb{1}^\top,\ Y = Y^\top \\ Y \odot zz^\top = A \odot zz^\top \end{Bmatrix}, \tag{F.1}$$

$$\mathcal{P}_4(Y; \sigma, \delta, \hat{d}) := \begin{Bmatrix} |(Y\mathbb{1})_i - d(Y)| \leqslant \sigma\sqrt{\hat{d}} & \forall i \in [n] \\ \left\| Y - \frac{d(Y)}{n}\mathbb{1}\mathbb{1}^\top \right\|_{\mathrm{op}} \leqslant \delta\sqrt{\hat{d}} & \end{Bmatrix}. \tag{F.2}$$

For convenience of notation, we will consider the following combined polynomial system in remaining of the section

$$\mathcal{E}(Y, z; A, \gamma, \sigma, \delta, \hat{d}) := \mathcal{P}_1(Y, z; A, \gamma) \cup \mathcal{P}_4(Y; \sigma, \delta, \hat{d}). \tag{F.3}$$

Lemma F.2. *If (A^*, z^*) is a feasible solution to $\mathcal{E}(Y, z; A, \gamma^*, \sigma, \delta, \hat{d})$ and $\gamma + \gamma^* < 1$, then it follows that*

$$\mathcal{E}(Y, z; A, \gamma, \sigma, \delta, \hat{d}) \left|\frac{Y,z}{8}\right. |d(Y) - d(A^*)| \leqslant \frac{4(\gamma+\gamma^*)\sigma\sqrt{\hat{d}} + 2(\gamma+\gamma^*)\delta\sqrt{\hat{d}}}{(1-\gamma-\gamma^*)^2}.$$

Proof. Let $w = z \odot z^*$. Notice that, by Lemma A.9, we have $\mathcal{E} \left|\frac{z}{4}\right. 1 - w_i \leqslant 2 - z_i - z_i^*$ for all $i \in [n]$. Moreover, using similar analysis as in the proof of Lemma D.2, it follows that

$$\mathcal{E} \left|\frac{Y,z}{4}\right. Y \odot ww^\top = A^* \odot ww^\top.$$

Therefore, we can get

$$\begin{aligned} \mathcal{E} \left|\frac{Y,z}{4}\right. n\Big(d(Y) - d(A^*)\Big) &= \langle Y - A^*, \mathbb{1}\mathbb{1}^\top\rangle \\ &= \langle Y - A^*, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle \\ &= \langle Y - \frac{d(Y)}{n}\mathbb{1}\mathbb{1}^\top + \frac{d(Y)}{n}\mathbb{1}\mathbb{1}^\top - \frac{d(A^*)}{n}\mathbb{1}\mathbb{1}^\top + \frac{d(A^*)}{n}\mathbb{1}\mathbb{1}^\top - A^*, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle \\ &= \langle Y - \frac{d(Y)}{n}\mathbb{1}\mathbb{1}^\top, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle + \langle \frac{d(A^*)}{n}\mathbb{1}\mathbb{1}^\top - A^*, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle \\ &\quad + \langle \frac{d(Y)}{n}\mathbb{1}\mathbb{1}^\top - \frac{d(A^*)}{n}\mathbb{1}\mathbb{1}^\top, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle \\ &= \langle Y - \frac{d(Y)}{n}\mathbb{1}\mathbb{1}^\top, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle + \langle \frac{d(A^*)}{n}\mathbb{1}\mathbb{1}^\top - A^*, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle \\ &\quad + \Big(d(Y) - d(A^*)\Big)\Big(n - \frac{1}{n}\langle \mathbb{1}, w\rangle^2\Big). \end{aligned} \tag{F.4}$$

By rearranging terms, we can get

$$\mathcal{E} \left|\frac{Y,z}{8}\right. \frac{\langle \mathbb{1}, w\rangle^2}{n}\Big(d(Y) - d(A^*)\Big) = \langle Y - \frac{d(Y)}{n}\mathbb{1}\mathbb{1}^\top, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle + \langle \frac{d(A^*)}{n}\mathbb{1}\mathbb{1}^\top - A^*, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle. \tag{F.5}$$

We bound the two terms on the right-hand side separately. For the first term $\langle Y - \frac{d(Y)}{n}\mathbb{1}\mathbb{1}^\top, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle$, we have

$$\mathcal{E}\Big|\frac{Y,z}{8}\ \langle Y-\frac{d(Y)}{n}\mathbb{1}\mathbb{1}^\top, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle = 2\langle Y-\frac{d(Y)}{n}\mathbb{1}\mathbb{1}^\top, \mathbb{1}(\mathbb{1}-w)^\top\rangle + \langle \frac{d(Y)}{n}\mathbb{1}\mathbb{1}^\top - Y, (\mathbb{1}-w)(\mathbb{1}-w)^\top\rangle\,. \tag{F.6}$$

From constraints $|(Y\mathbb{1})_i - d(Y)| \leqslant \sigma\sqrt{\hat{d}}$ for all $i \in [n]$, $\langle \mathbb{1}, z\rangle \geqslant (1-\gamma)n$ and $\langle \mathbb{1}, z^*\rangle \geqslant (1-\gamma^*)n$, we have

$$\begin{aligned}\mathcal{E}\Big|\frac{Y,z}{8}\ \langle Y - \frac{d(Y)}{n}\mathbb{1}\mathbb{1}^\top, \mathbb{1}(\mathbb{1}-w)^\top\rangle &= \langle Y\mathbb{1} - d(Y)\mathbb{1}, \mathbb{1} - w\rangle \\ &\leqslant \sum_{i\in[n]} (1-w_i)\sigma\sqrt{\hat{d}} \\ &\leqslant \sigma\sqrt{\hat{d}}\cdot\left(\sum_{i\in[n]} 2 - z_i - z_i^*\right) \\ &\leqslant (\gamma+\gamma^*)n\sigma\sqrt{\hat{d}}\,.\end{aligned} \tag{F.7}$$

From constraints $\left\|Y - \frac{d(Y)}{n}\mathbb{1}\mathbb{1}^\top\right\|_{\mathrm{op}} \leqslant \delta\sqrt{\hat{d}}$, $\langle \mathbb{1}, z\rangle \geqslant (1-\gamma)n$ and $\langle \mathbb{1}, z^*\rangle \geqslant (1-\gamma^*)n$, we have

$$\begin{aligned}\mathcal{E}\Big|\frac{Y,z}{8}\ \langle \frac{d(Y)}{n}\mathbb{1}\mathbb{1}^\top - Y, (\mathbb{1}-w)(\mathbb{1}-w)^\top\rangle &\leqslant \left\|Y - \frac{d(Y)}{n}\mathbb{1}\mathbb{1}^\top\right\|_{\mathrm{op}} \|\mathbb{1}-w\|_2^2 \\ &\leqslant \delta\sqrt{\hat{d}}\cdot\left(\sum_{i\in[n]}(1-w_i)^2\right) \\ &= \delta\sqrt{\hat{d}}\cdot\left(\sum_{i\in[n]} 1 - w_i\right) \\ &\leqslant \delta\sqrt{\hat{d}}\cdot\left(\sum_{i\in[n]} 2 - z_i - z_i^*\right) \\ &\leqslant (\gamma+\gamma^*)n\delta\sqrt{\hat{d}}\,,\end{aligned} \tag{F.8}$$

where the equality is because $\mathcal{E}\Big|\frac{z}{2}\ (1-w_i)^2 = (1-z_iz_i^*)^2 = 1 - z_iz_i^* = 1 - w_i$.

Plugging Eq. (F.7) and Eq. (F.8) into Eq. (F.6), it follows that

$$\mathcal{E}\Big|\frac{Y,z}{8}\ \langle Y - \frac{d(Y)}{n}\mathbb{1}\mathbb{1}^\top, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle \leqslant 2(\gamma+\gamma^*)n\sigma\sqrt{\hat{d}} + (\gamma+\gamma^*)n\delta\sqrt{\hat{d}}\,. \tag{F.9}$$

For the second term $\langle \frac{d(A^*)}{n}\mathbb{1}\mathbb{1}^\top - A^*, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle$, we can apply the same proof as above to get

$$\mathcal{E}\Big|\frac{Y,z}{8}\ \langle \frac{d(A^*)}{n}\mathbb{1}\mathbb{1}^\top - A^*, \mathbb{1}\mathbb{1}^\top - ww^\top\rangle \leqslant 2(\gamma+\gamma^*)n\sigma\sqrt{\hat{d}} + (\gamma+\gamma^*)n\delta\sqrt{\hat{d}}\,. \tag{F.10}$$

Plugging Eq. (F.9) and Eq. (F.10) into Eq. (F.5), it follows that

$$\mathcal{E}\Big|\frac{Y,z}{8}\ \frac{\langle \mathbb{1}, w\rangle^2}{n}\Big(d(Y) - d(A^*)\Big) \leqslant 4(\gamma+\gamma^*)n\sigma\sqrt{\hat{d}} + 2(\gamma+\gamma^*)n\delta\sqrt{\hat{d}}\,.$$

Using the same proof strategy, we can also get

$$\mathcal{E}\Big|\frac{Y,z}{8}\ \frac{\langle \mathbb{1}, w\rangle^2}{n}\Big(d(A^*) - d(Y)\Big) \leqslant 4(\gamma+\gamma^*)n\sigma\sqrt{\hat{d}} + 2(\gamma+\gamma^*)n\delta\sqrt{\hat{d}}\,.$$

Applying Lemma A.10, it follows that

$$\mathcal{E}\left|\frac{Y,z}{8}\ \frac{\langle \mathbb{1}, w\rangle^4}{n^2}\Big(d(Y)-d(A^*)\Big)^2 \leqslant \Big(4(\gamma+\gamma^*)n\sigma\sqrt{\hat{d}}+2(\gamma+\gamma^*)n\delta\sqrt{\hat{d}}\Big)^2 . \tag{F.11}$$

Now, we would like to lower bound $\langle \mathbb{1}, w\rangle^4$. By Lemma A.9, we have $\mathcal{E}\left|\frac{z}{4}\ w_i \geqslant z_i + z_i^* - 1\right.$ for all $i \in [n]$. Therefore,

$$\mathcal{E}\left|\frac{Y,z}{8}\ \langle \mathbb{1}, w\rangle = \sum_{i\in[n]} w_i \geqslant \sum_{i\in[n]} (z_i + z_i^* - 1) \geqslant (1-\gamma-\gamma^*)n .\right.$$

Since $\gamma+\gamma^* < 1$, we have $1-\gamma-\gamma^* > 0$, and, therefore,

$$\mathcal{E}\left|\frac{Y,z}{8}\ \langle \mathbb{1}, w\rangle^4 \geqslant (1-\gamma-\gamma^*)^4 n^4 .\right.$$

Plugging this into Eq. (F.11), we have

$$\mathcal{E}\left|\frac{Y,z}{8}\ (1-\gamma-\gamma^*)^4 n^2\Big(d(Y)-d(A^*)\Big)^2 \leqslant \Big(4(\gamma+\gamma^*)n\sigma\sqrt{\hat{d}}+2(\gamma+\gamma^*)n\delta\sqrt{\hat{d}}\Big)^2\right.$$

$$\left|\frac{Y,z}{8}\ \Big(d(Y)-d(A^*)\Big)^2 \leqslant \frac{\Big(4(\gamma+\gamma^*)\sigma\sqrt{\hat{d}}+2(\gamma+\gamma^*)\delta\sqrt{\hat{d}}\Big)^2}{(1-\gamma-\gamma^*)^4} .\right.$$

Applying Lemma A.11, it follows that

$$\mathcal{E}\left|\frac{Y,z}{8}\ |d(Y)-d(A^*)| \leqslant \frac{4(\gamma+\gamma^*)\sigma\sqrt{\hat{d}}+2(\gamma+\gamma^*)\delta\sqrt{\hat{d}}}{(1-\gamma-\gamma^*)^2} .\right.$$

□

Lemma F.3. *Let A be an η-corrupted adjacency matrix of a random graph $\boldsymbol{G}^\circ \sim \mathbb{G}(n, \frac{d^\circ}{n})$. With probability $1-n^{-\Omega(1)}$, there exists A^* and z^* such that*

- $|d(A^*)-d^\circ| \leqslant 10\sqrt{\frac{d^\circ \log n}{n}}$.
- *(A^*, z^*) is a feasible solution to $\mathcal{E}(Y, z; A, \gamma, \sigma, \delta, \hat{d})$ with $\gamma = \eta$, $\sigma = 4\log n$, $\delta = 4C\sqrt{\log n}$ for some constant C and $\hat{d} \geqslant \frac{1}{2}d^\circ$.*

Proof. Let $\boldsymbol{A}^\circ$ be the adjacency matrix of $\boldsymbol{G}^\circ$ and $z^\circ \in \{0,1\}^n$ denote the set of uncorrupted nodes ($z_i^\circ = 1$ if and only if node i is uncorrupted).We will show that $A^* = \boldsymbol{A}^\circ$ and $z^* = z^\circ$ satisfies the lemma.

Guarantee 1. By Lemma B.1, we know that, with probability $1-n^{-\Omega(1)}$,

$$|d(\boldsymbol{A}^\circ)-d^\circ| \leqslant 10\sqrt{\frac{d^\circ \log n}{n}} . \tag{F.12}$$

Guarantee 2. It is easy to check that $z^* \odot z^* = z^*$, $0 \leqslant A^* \leqslant \mathbb{1}\mathbb{1}^\top$, $A^* = (A^*)^\top$ and $\langle \mathbb{1}, z^*\rangle \geqslant 1-\eta n$. By Lemma B.2, we know that, with probability $1-n^{-\Omega(1)}$,

$$\|\boldsymbol{A}^\circ \mathbb{1} - d^\circ \mathbb{1}\|_\infty \leqslant \sqrt{d^\circ}\log n . \tag{F.13}$$

Combining Eq. (F.12) and Eq. (F.13), we have

$$\begin{aligned}\|\boldsymbol{A}^\circ \mathbb{1} - d(\boldsymbol{A}^\circ)\mathbb{1}\|_\infty &\leqslant \|\boldsymbol{A}^\circ \mathbb{1} - d^\circ \mathbb{1}\|_\infty + \|d^\circ \mathbb{1} - d(\boldsymbol{A}^\circ)\mathbb{1}\|_\infty \\ &\leqslant \sqrt{d^\circ}\log n + 10\sqrt{\frac{d^\circ \log n}{n}} \\ &\leqslant 2\log n\sqrt{d^\circ} .\end{aligned} \tag{F.14}$$

Therefore, for $\sigma = 4\log n$ and $\hat{d} \geqslant \frac{1}{2}d^\circ$, it follows that

$$|(A^*\mathbb{1})_i - d(A^*)| \leqslant 2\log n\sqrt{d^\circ} \leqslant \sigma\sqrt{\hat{d}}\,,$$

for all $i \in [n]$.

By Lemma B.5, we know that, with probability $1 - n^{-\Omega(1)}$, for some universal constant C,

$$\left\|A^\circ - \frac{d^\circ}{n}\mathbb{1}\mathbb{1}^\top\right\|_{\mathrm{op}} \leqslant C\sqrt{d^\circ \log n}\,. \tag{F.15}$$

Combining Eq. (F.12) and Eq. (F.15), we have

$$\begin{aligned}\left\|A^\circ - \frac{d(A^\circ)}{n}\mathbb{1}\mathbb{1}^\top\right\|_{\mathrm{op}} &\leqslant \left\|A^\circ - \frac{d(A^\circ)}{n}\mathbb{1}\mathbb{1}^\top\right\|_{\mathrm{op}} + \left\|\frac{d(A^\circ)}{n}\mathbb{1}\mathbb{1}^\top - \frac{d^\circ}{n}\mathbb{1}\mathbb{1}^\top\right\|_{\mathrm{op}} \\ &\leqslant C\sqrt{d^\circ \log n} + 10\sqrt{\frac{d^\circ \log n}{n}} \\ &\leqslant 2C\sqrt{d^\circ \log n}\,.\end{aligned} \tag{F.16}$$

Therefore, for $\delta = 4C\sqrt{\log n}$ and $\hat{d} \geqslant \frac{1}{2}d^\circ$, it follows that

$$\left\|A^* - \frac{d(A^*)}{n}\mathbb{1}\mathbb{1}^\top\right\|_{\mathrm{op}} \leqslant 2C\sqrt{d^\circ \log n} \leqslant \delta\sqrt{\hat{d}}\,.$$

Thus, (A^*, z^*) is a feasible solution to $\mathcal{E}(Y, z; A, \gamma, \sigma, \delta, \hat{d})$ with $\gamma = \eta$, $\sigma = 4\log n$, $\delta = 4C\sqrt{\log n}$ and $\hat{d} \geqslant \frac{1}{2}d^\circ$. □

F.2 Robust algorithm

In this section, we show that the following algorithm based on sum-of-squares proofs in Appendix F.1 obtains a robust approximation of d° that is optimal up to logarithmic factors.

Algorithm F.4 (Robust fine estimation algorithm for Erdős-Rényi random graphs)**.**
Input: η-corrupted adjacency matrix A and corruption fraction η.

Algorithm:

1. Obtain coarse estimator $\hat{d}$ by applying Algorithm D.4 with $A, \eta, R = 1$ as input.
2. Obtain level-8 pseudo-expectation $\tilde{\mathbb{E}}$ by solving sum-of-squares relaxation of program $\mathcal{E}(Y, z; A, \gamma, \sigma, \delta, \hat{d})$ (defined in Eq. (F.3)) with A, $\gamma = \eta$, $\sigma = 4\log n$, $\delta = 4C\sqrt{\log n}$ and $\hat{d}$.

Output: $\tilde{\mathbb{E}}[d(Y)]$

Theorem F.5 (Robust fine estimation for Erdős-Rényi random graphs)**.** *Let A be an η-corrupted adjacency matrix of a random graph $\boldsymbol{G}^\circ \sim \mathbb{G}(n, \frac{d^\circ}{n})$. With probability $1 - n^{-\Omega(1)}$, Algorithm F.4 outputs an estimate $\tilde{d}$ satisfying*

$$\left|\frac{\tilde{d}}{d^\circ} - 1\right| \leqslant O\left(\sqrt{\frac{\log n}{d^\circ n}} + \frac{\eta \log n}{\sqrt{d^\circ}}\right).$$

Proof. By Theorem D.5, we have $\frac{1}{2}d^\circ \leqslant \hat{d} \leqslant \frac{3}{2}d^\circ$. Let $\gamma^* = \eta$, by Lemma F.2 and Lemma F.3, it follows that

$$\mathcal{E}(Y, z; A, \gamma, \sigma, \delta, \hat{d}) \,\Big|\frac{Y, z}{O(1)}\; |d(Y) - d(A^*)| \leqslant \frac{4(\gamma + \gamma^*)\sigma\sqrt{\hat{d}} + 2(\gamma + \gamma^*)\delta\sqrt{\hat{d}}}{(1 - \gamma - \gamma^*)^2}$$

$$\leqslant \frac{40\eta \log n \sqrt{d^\circ} + 40C\eta\sqrt{d^\circ \log n}}{(1-2\eta)^2}$$
$$\leqslant C'\eta \log n \sqrt{d^\circ}\,,$$

for some constant C', and,

$$|d(A^*) - d^\circ| \leqslant 10\sqrt{\frac{d^\circ \log n}{n}}\,.$$

Therefore, we have

$$\mathcal{E}(Y, z; A, \gamma, \sigma, \delta, \hat{d}) \vdash^{Y,z}_{O(1)} |d(Y) - d^\circ| \leqslant C'\eta \log n \sqrt{d^\circ} + 10\sqrt{\frac{d^\circ \log n}{n}}\,.$$

Thus, the level-8 pseudo-expectation $\tilde{\mathbb{E}}$ satisfies

$$\left|\tilde{\mathbb{E}}[d(Y)] - d^\circ\right| \leqslant C'\eta \log n \sqrt{d^\circ} + 10\sqrt{\frac{d^\circ \log n}{n}}\,,$$

which implies that

$$\left|\frac{\tilde{\mathbb{E}}[d(Y)]}{d^\circ} - 1\right| O\left(\sqrt{\frac{\log n}{d^\circ n}} + \frac{\eta \log n}{\sqrt{d^\circ}}\right).$$

□

F.3 Private algorithm

In this section, we present our algorithm and prove Theorem F.1. Our algorithm instantiates the sum-of-squares exponential mechanism in Appendix C.

Score function. For an n-by-n symmetric matrix A and a scalar d, we define the score of d with regard to A to be

$$s(d; A) := \min_{0 \leqslant \gamma \leqslant 1} \gamma n \text{ s.t. } \begin{cases} \exists \text{ level-8 pseudo-expectation } \tilde{\mathbb{E}} \text{ satisfying} \\ \mathcal{E}(Y, z; A, \gamma, \sigma, \delta, \hat{d}) \cup \{|d(Y) - d| \leqslant \alpha d\}\,, \end{cases} \tag{F.17}$$

where $\mathcal{E}(Y, z; A, \gamma, \sigma, \delta, \hat{d})$ is the polynomial system defined in Eq. (F.3), $\hat{d}$ is a coarse estimate, and σ, δ, α are fixed parameters whose values will be decided later. Note that $(Y = \frac{d}{n}\mathbb{1}\mathbb{1}^\top, z = \mathbb{0})$ is a solution to the polynomial system $\mathcal{E}(Y, z; A, 1, \sigma, \delta, \hat{d}) \cup \{|d(Y)/d - 1| \leqslant \alpha\}$ for any $A \in \mathbb{R}^{n\times n}$ and any $d \in [0, n]$.

To efficiently compute $s(d; A)$, we can use the scheme as described in Remark C.1.

Exponential mechanism. Given a privacy parameter ε and an n-by-n symmetric matrix A, our private algorithm in Theorem F.1 is the exponential mechanism with score function Eq. (F.17) and range $[0, n]$.

Algorithm F.6 (Fine estimation for Erdős-Rényi random graphs)**.**
Input: Graph A, coarse estimate $\hat{d}$.

Parameters: $\varepsilon, \sigma, \delta, \alpha$.

Output: A sample from the distribution $\mu_{A,\varepsilon}$ with support $[0, n]$ and density

$$\mathrm{d}\mu_{A,\varepsilon}(d) \propto \exp(-\varepsilon \cdot s(d; A))\,, \tag{F.18}$$

where $s(d; A)$ is defined in Eq. (F.17).

To efficiently sample from $\mu_{A,\varepsilon}$, we can use the scheme as described in Remark C.2.

Privacy. The following privacy guarantee of our algorithm is a direct corollary of Lemma C.4.

Lemma F.7 (Privacy)**.** *Algorithm F.6 is 2ε-differentially node private.*

Utility. The utility guarantee of our algorithm is stated in the following lemma.

Lemma F.8 (Utility). *There are constants C_1, C_2, C_3 such that the following holds. For any $\eta \leqslant C_1$, $\varepsilon \geqslant C_2 \log(n)/n$, and $d^\circ \geqslant C_3$, given an η-corrupted Erdős-Rényi random graph $\mathbb{G}(n, d^\circ/n)$ and a coarse estimate $\hat{d}$ such that $0.5d^\circ \leqslant \hat{d} \leqslant 2d^\circ$, Algorithm F.6 outputs an estimate $\tilde{d}$ satisfying*

$$\left|\frac{\tilde{d}}{d^\circ} - 1\right| \leqslant O\left(\sqrt{\frac{\log n}{d^\circ n}} + \frac{\log^2 n}{\sqrt{d^\circ}\varepsilon n} + \frac{\eta \log n}{\sqrt{d^\circ}}\right),$$

with probability $1 - n^{-\Omega(1)}$.

Before proving Lemma F.8, we need the following two lemmas.

Lemma F.9 (Volume of low-score points). *Let $A \in \mathbb{R}^{n\times n}$ and $\varepsilon > 0$. Consider the distribution $\mu_{A,\varepsilon}$ defined by Eq. (F.18). Suppose $(Y = A^*, z = z^*)$ is a solution to $\mathcal{E}(Y, z; A, \gamma^*)$ and $d(A^*) \geqslant 2$. Then for any $t \geqslant 0$,*

$$\mathop{\mathbb{P}}_{\boldsymbol{d}\sim\mu_{A,\varepsilon}}\left(s(\boldsymbol{d}; A) \geqslant \gamma^* n + \frac{t \log n}{\varepsilon}\right) \leqslant \frac{n^{-t+1}}{\alpha}.$$

Proof. Apply Lemma C.5 with $\mathcal{D} = [0, n]$ and

$$\mathcal{G}(A^*) = \left\{d \in \mathcal{D} : \frac{d(A^*)}{1+\alpha} \leqslant d \leqslant \frac{d(A^*)}{1-\alpha}\right\}.$$

As $[d(A^*)/(1+\alpha), d(A^*)] \subseteq \mathcal{G}(A^*)$ and $d(A^*) \geqslant 2 \geqslant 1+\alpha$, we have $\mathrm{vol}(\mathcal{G}(A^*)) \geqslant \alpha$. □

Lemma F.10 (Low score implies utility). *Let $A \in \mathbb{R}^{n\times n}$ and consider the score function $s(\cdot\,; A)$ defined in Eq. (F.17). Suppose $(Y = A^*, z = z^*)$ is a solution to $\mathcal{E}(Y, z; A, \gamma^*)$. For a scalar d such that $s(d; A) \leqslant \tau n$ and $\gamma^* + \tau \leqslant 0.1$,*

$$\frac{1}{1+\alpha}\left(d(A^*) - 5(\gamma^* + \tau)(\sigma + \delta)\sqrt{\hat{d}}\right) \leqslant d \leqslant \frac{1}{1-\alpha}\left(d(A^*) + 5(\gamma^* + \tau)(\sigma + \delta)\sqrt{\hat{d}}\right).$$

Proof. Applying Lemma F.2 with $\gamma^* + \tau \leqslant 0.1$, we have

$$\mathcal{E}(Y, z; A, \tau) \;\Big|\!\frac{Y,z}{8}\; |d(Y) - d(A^*)| \leqslant 5(\gamma^* + \tau)(\sigma + \delta)\sqrt{\hat{d}}.$$

Thus,

$$\begin{aligned}&\mathcal{E}(Y, z; A, \tau) \cup \{|d(Y) - d| \leqslant \alpha d\}\\ \Big|\!\frac{Y,z}{8}\;& \frac{1}{1+\alpha}\left(d(A^*) - 5(\gamma^* + \tau)(\sigma + \delta)\sqrt{\hat{d}}\right) \leqslant d \leqslant \frac{1}{1-\alpha}\left(d(A^*) + 5(\gamma^* + \tau)(\sigma + \delta)\sqrt{\hat{d}}\right).\end{aligned}$$

□

Now we are ready to prove Lemma F.8.

Proof of Lemma F.8. Let A be a realization of η-corrupted $\mathbb{G}(n, d^\circ/n)$. By Lemma F.3, the following event happens with probability at least $1 - n^{-\Omega(1)}$. There exists a solution $(Y = A^*, z = z^*)$ to $\mathcal{E}(Y, z; A, \gamma^*, \sigma, \delta, \hat{d})$ where $\gamma^* = \eta$, $\sigma \leqslant O(\log n)$, $\delta \leqslant O(\sqrt{\log n})$, and $|d(A^*) - d^\circ| \leqslant O\left(\sqrt{d^\circ \log(n)/n}\right)$.

As $d(A^*) \geqslant 0.9d^\circ \geqslant 2$, then it follows by setting $t = 10$ and $\alpha = n^{-2}$ in Lemma F.9 that,

$$\mathop{\mathbb{P}}_{\boldsymbol{d}\sim\mu_{A,\varepsilon}}(s(\boldsymbol{d}; A) \leqslant \tau n) \geqslant 1 - n^{-7} \text{ where } \tau := 2\eta + 10\log(n)/(\varepsilon n).$$

Let $\tilde{d}$ be a scalar such that $s(\tilde{d}; A) \leqslant \tau n$. For η and $\log(n)/(\varepsilon n)$ smaller than some constant, we have $2\eta + \tau \leqslant 0.1$. Then by Lemma F.10,

$$\frac{1}{1+\alpha}\left(d(A^*) - 5(\eta + \tau)(\sigma + \delta)\sqrt{\hat{d}}\right) \leqslant \tilde{d} \leqslant \frac{1}{1-\alpha}\left(d(A^*) + 5(\eta + \tau)(\sigma + \delta)\sqrt{\hat{d}}\right).$$

Plugging in everything, we have

$$\left|\frac{\tilde{d}}{d^\circ}-1\right| \leqslant O\left(\sqrt{\frac{\log n}{d^\circ n}}+\frac{\log^2 n}{\sqrt{d^\circ}\varepsilon n}+\frac{\eta \log n}{\sqrt{d^\circ}}\right).$$

□

Proof of Theorem F.1. By Lemma F.7 and Lemma F.8.

G Lower bounds

In this section, we prove Theorem 1.5, Theorem 1.7, and Theorem 1.8.

G.1 Lower bound for Erdős-Rényi random graphs

In this section, we prove Theorem 1.5.

Theorem (Restatement of Theorem 1.5)**.** *Suppose there is an ε-differentially node-private algorithm that, given an Erdős-Rényi random graph $\mathbb{G}(n,p^\circ)$, outputs an estimate $\tilde{p}$ satisfying $|\tilde{p}/p^\circ - 1| \leqslant \alpha$ with probability $1-\beta$. Then we must have*

$$\alpha \geqslant \Omega\left(\frac{\log(1/\beta)}{\varepsilon n\sqrt{np^\circ}}\right).$$

We leave the formal proof of Theorem 1.5 to the end of this section. Now we sketch the proof idea. One natural idea to prove this theorem is to construct a coupling ω of $\mathbb{G}(n,p^\circ)$ and $\mathbb{G}(n,(1-2\alpha)p^\circ)$ such that for $(\boldsymbol{G},\boldsymbol{G}')\sim\omega$, the typical distance between $\boldsymbol{G}$ and $\boldsymbol{G}'$ can be well controlled. However, such a coupling is tricky to construct directly, as the node degrees in an Erdős-Rényi random graph are not independent. To avoid dealing with such dependence, we instead consider the directed Erdős-Rényi random graphs, which is inspired by the proof of [AJK+22, Theorem 1.5]. The directed Erdős-Rényi random graph model, denoted by $\tilde{\mathbb{G}}(n,p^\circ)$, is a distribution over n-node directed graphs where each edge (i,j) is present with probability p° independently. Since the outdegrees in a directed Erdős-Rényi random graph are i.i.d. Binomial random variables, it is not so difficult to construct a coupling of $\tilde{\mathbb{G}}(n,p^\circ)$ and $\tilde{\mathbb{G}}(n,(1-2\alpha)p^\circ)$. Then we can convert such a coupling into a coupling of $\mathbb{G}(n,p^\circ)$ and $\mathbb{G}(n,(1-2\alpha)p^\circ)$.

Lemma G.1 (Coupling)**.** *Let $p^\circ \in [0,1]$, $\alpha \in [0,1/2]$, and $p' := (1-2\alpha)p^\circ$. There exists a coupling ω of $\mathbb{G}(n,p^\circ)$ and $\mathbb{G}(n,p')$ with the following property. For $(\boldsymbol{G},\boldsymbol{G}')\sim\omega$, the distribution of $\mathrm{dist}(\boldsymbol{G},\boldsymbol{G}')$ is the binomial distribution $\mathrm{Bin}(n,\Delta)$ where $\Delta = \mathrm{TV}(\mathrm{Bin}(n,p^\circ),\mathrm{Bin}(n,p'))$. Moreover, if $p^\circ \leqslant c$ and $\alpha \leqslant c'/\sqrt{np^\circ}$ for some constants c, c', then $\Delta \lesssim \alpha\sqrt{np^\circ}$.*

Proof. We first show that it suffices to construct a coupling of $\tilde{\mathbb{G}}(n,p^\circ)$ and $\tilde{\mathbb{G}}(n,p')$. For a directed graph $\tilde{G}$, it can be converted into an undirected graph $U(\tilde{G})$ by letting $\{i,j\}\in U(\tilde{G})$ iff $i \leqslant j$ and $(i,j)\in\tilde{G}$. It is easy to see that[10] $\mathrm{dist}(\tilde{G},\tilde{G}') = \mathrm{dist}(U(\tilde{G}),U(\tilde{G}'))$. Also observe that if $\tilde{\boldsymbol{G}}\sim\tilde{\mathbb{G}}(n,p^\circ)$ then $U(\tilde{\boldsymbol{G}})\sim\mathbb{G}(n,p^\circ)$. Therefore, a coupling $\tilde{\omega}$ of $\tilde{\mathbb{G}}(n,p^\circ)$ and $\tilde{\mathbb{G}}(n,p')$ can be easily converted in to a coupling ω of $\mathbb{G}(n,p^\circ)$ and $\mathbb{G}(n,p')$ such that, for $(\tilde{\boldsymbol{G}},\tilde{\boldsymbol{G}}')\sim\tilde{\omega}$ and $(\boldsymbol{G},\boldsymbol{G}')\sim\omega$, we have

$$\mathrm{dist}\big(\tilde{\boldsymbol{G}},\tilde{\boldsymbol{G}}'\big) \overset{\mathrm{d}}{=} \mathrm{dist}(\boldsymbol{G},\boldsymbol{G}').$$

Now we construct a coupling of $\tilde{\mathbb{G}}(n,p^\circ)$ and $\tilde{\mathbb{G}}(n,p')$. Instead of sampling each edge independently, an equivalent way to sample from $\tilde{\mathbb{G}}(n,p^\circ)$ is as follows. For each $i\in[n]$:

[10] For a directed graph $\tilde{G}$, we define its adjacency matrix $\tilde{A}$ to be $\tilde{A}(i,j) := \mathbb{1}\{(i,j)\in\tilde{G}\}$. The (node) distance between two n-node directed graphs $\tilde{G},\tilde{G}'$, denoted by $\mathrm{dist}(\tilde{G},\tilde{G}')$, is number of nonzero rows of $\tilde{A}-\tilde{A}'$.

- Sample an outdegree $\boldsymbol{d} \sim \mathrm{Bin}(n, p^\circ)$.
- Sample a uniformly random subset $\boldsymbol{S} \subseteq [n]$ of size $\boldsymbol{d}$. For each $j \in \boldsymbol{S}$, add an edge from i to j.

Then it is easy to see there exists a coupling $\tilde{\omega}$ of $\tilde{\mathbb{G}}(n, p^\circ)$ and $\tilde{\mathbb{G}}(n, p')$ such that if $(\tilde{\boldsymbol{G}}, \tilde{\boldsymbol{G}}') \sim \tilde{\omega}$ then $\mathrm{dist}(\tilde{\boldsymbol{G}}, \tilde{\boldsymbol{G}}') \sim \mathrm{Bin}(n, \Delta)$ where

$$\Delta = \mathrm{TV}(\mathrm{Bin}(n, p^\circ), \mathrm{Bin}(n, p'))\,.$$

We have the following bound on the total variation between binomial distributions (see e.g. [AJ06, Equation (2.15)]). For $0 < p < 1$ and $0 < x < 1 - p$,

$$\mathrm{TV}(\mathrm{Bin}(N, p), \mathrm{Bin}(N, p + x)) \leqslant \frac{\sqrt{e}}{2} \frac{\tau(x)}{(1 - \tau(x))^2}\,,$$

where $\tau(x) := x\sqrt{\frac{N+2}{2p(1-p)}}$, provided $\tau(x) < 1$. Plugging in $N = n$, $p = p^\circ$, and $x = 2\alpha p^\circ$, we have

$$\Delta = \mathrm{TV}(\mathrm{Bin}(n, p^\circ), \mathrm{Bin}(n, p')) \lesssim \alpha\sqrt{np^\circ}\,,$$

provided $p^\circ \leqslant c$ and $\alpha \leqslant c'/\sqrt{np^\circ}$ for sufficiently small absolute constants c, c'. □

Proof of Theorem 1.5. Let $\mathcal{A}$ be an algorithm satisfying the theorem's assumptions. Let $p' := (1 - 2\alpha)p^\circ$. Let ω be a coupling of $\mathbb{G}(n, p^\circ)$ and $\mathbb{G}(n, p')$ as guaranteed by Lemma G.1. Then for $(\boldsymbol{G}, \boldsymbol{G}') \sim \omega$, we have $\mathrm{dist}(\boldsymbol{G}, \boldsymbol{G}') \sim \mathrm{Bin}(n, \Delta)$ where $\Delta = \mathrm{TV}(\mathrm{Bin}(n, p^\circ), \mathrm{Bin}(n, p'))$.

By the utility assumption of algorithm $\mathcal{A}$,

$$\mathop{\mathbb{P}}_{\mathcal{A}, \mathbb{G}(n, p^\circ)}(|\mathcal{A}(\boldsymbol{G}) - p^\circ| < \alpha p^\circ) \geqslant 1 - \beta\,.$$

As algorithm $\mathcal{A}$ is ε-DP, we have for any graphs G, G' that,

$$\mathop{\mathbb{P}}_{\mathcal{A}}(|\mathcal{A}(G') - p'| < \alpha p^\circ) \leqslant e^{\varepsilon \cdot \mathrm{dist}(G, G')} \cdot \mathop{\mathbb{P}}_{\mathcal{A}}(|\mathcal{A}(G) - p'| < \alpha p^\circ)\,.$$

Taking expectation w.r.t. the coupling ω on both sides gives

$$\begin{aligned}
\mathop{\mathbb{E}}_{\omega} \mathop{\mathbb{E}}_{\mathcal{A}} \mathbb{1}\{|\mathcal{A}(\boldsymbol{G}') - p'| < \alpha p^\circ\} &\leqslant \mathop{\mathbb{E}}_{\omega} e^{\varepsilon \cdot \mathrm{dist}(\boldsymbol{G}, \boldsymbol{G}')} \cdot \mathop{\mathbb{E}}_{\mathcal{A}} \mathbb{1}\{|\mathcal{A}(\boldsymbol{G}) - p'| < \alpha p^\circ\}\,, \\
\mathop{\mathbb{P}}_{\mathcal{A}, \tilde{\mathbb{G}}(n, p')}(|\mathcal{A}(\boldsymbol{G}') - p'| < \alpha p^\circ) &\leqslant \mathop{\mathbb{E}}_{\omega, \mathcal{A}} e^{\varepsilon \cdot \mathrm{dist}(\boldsymbol{G}, \boldsymbol{G}')} \cdot \mathbb{1}\{|\mathcal{A}(\boldsymbol{G}) - p'| < \alpha p^\circ\}\,. \qquad \text{(G.1)}
\end{aligned}$$

By the utility assumption of algorithm $\mathcal{A}$ and $p' < p^\circ$, the left-hand side of Eq. (G.1) is at least $1 - \beta$. Using the Cauchy-Schwartz inequality, the right-hand side of Eq. (G.1) can be upper bounded as follows,

$$\begin{aligned}
\mathop{\mathbb{E}}_{\omega, \mathcal{A}} e^{\varepsilon \cdot \mathrm{dist}(\boldsymbol{G}, \boldsymbol{G}')} \cdot \mathbb{1}\{|\mathcal{A}(\boldsymbol{G}) - p'| < \alpha p^\circ\} &\leqslant \sqrt{\mathop{\mathbb{E}}_{\omega, \mathcal{A}} e^{2\varepsilon \cdot \mathrm{dist}(\boldsymbol{G}, \boldsymbol{G}')}} \sqrt{\mathop{\mathbb{E}}_{\omega, \mathcal{A}} \mathbb{1}\{|\mathcal{A}(\boldsymbol{G}) - p'| < \alpha p^\circ\}} \\
&\leqslant \sqrt{\mathop{\mathbb{E}}_{\mathrm{Bin}(n, \Delta)} e^{2\varepsilon \cdot \boldsymbol{X}}} \sqrt{\mathop{\mathbb{P}}_{\mathcal{A}, \tilde{\mathbb{G}}(n, p^\circ)}(|\mathcal{A}(\boldsymbol{G}) - p'| < \alpha p^\circ)}\,.
\end{aligned}$$

By squaring both sides of Eq. (G.1) and plugging in the above two bounds, we have

$$(1 - \beta)^2 \leqslant \mathop{\mathbb{E}}_{\mathrm{Bin}(n, \Delta)} \left[e^{2\varepsilon \cdot \boldsymbol{X}}\right] \cdot \mathop{\mathbb{P}}_{\mathcal{A}, \tilde{\mathbb{G}}(n, p^\circ)}(|\mathcal{A}(\boldsymbol{G}) - p'| < \alpha p^\circ)\,.$$

Using the formula for the moment generating function of binomial distributions, we have

$$\mathop{\mathbb{E}}_{\mathrm{Bin}(n, \Delta)} \left[e^{2\varepsilon \cdot \boldsymbol{X}}\right] = \left(1 + \Delta(e^{2\varepsilon} - 1)\right)^n \leqslant e^{n\Delta(e^{2\varepsilon} - 1)}\,.$$

Then

$$\mathop{\mathbb{P}}_{\mathcal{A}, \mathbb{G}(n, p^\circ)}(|\mathcal{A}(\boldsymbol{G}) - p'| < \alpha p^\circ) \geqslant (1 - \beta)^2 \cdot e^{-n\Delta(e^{2\varepsilon} - 1)}\,.$$

Since $p' - p^\circ = 2\alpha p^\circ$, the two events $\{\hat{p} : |\hat{p} - p^\circ| < \alpha p^\circ\}$ and $\{\hat{p} : |\hat{p} - p'| < \alpha p^\circ\}$ are disjoint. Thus,

$$\mathbb{P}_{\mathcal{A},\mathbb{G}(n,p^\circ)}(|\mathcal{A}(\boldsymbol{G}) - p'| < \alpha p^\circ) \leqslant 1 - \mathbb{P}_{\mathcal{A},\mathbb{G}(n,p^\circ)}(|\mathcal{A}(\boldsymbol{G}) - p^\circ| < \alpha p^\circ) \leqslant \beta\,.$$

Therefore, we have the following lower bound

$$\Delta \geqslant \frac{2\log(1-\beta) + \log(1/\beta)}{n(e^{2\varepsilon} - 1)}\,,$$

which is $\Delta \gtrsim \frac{\log(1/\beta)}{n\varepsilon}$ for samll enough ε and β.

By Lemma G.1, if $p^\circ \leqslant c$ and $\alpha \leqslant c'/\sqrt{np^\circ}$ for some constants c, c', then $\Delta \lesssim \alpha\sqrt{np^\circ}$. Combined with the lower bound $\Delta \gtrsim \frac{\log(1/\beta)}{n\varepsilon}$, we have

$$\alpha \gtrsim \frac{\log(1/\beta)}{n\varepsilon\sqrt{np^\circ}}\,.$$

□

G.2 Lower bound for inhomogeneous random graphs

In this section, we prove Theorem 1.7 and Theorem 1.8.

We first show the lower bound for the error rate of robust estimation.

Theorem (Restatement of Theorem 1.7)**.** *Suppose there is an algorithm satisfies the following guarantee for any symmetric matrix $Q^\circ \in [0,1]^{n\times n}$. Given an η-corrupted inhomogeneous random graph $\mathbb{G}(n, Q^\circ)$, the algorithm outputs an estimate $\hat{p}$ satisfying $|\hat{p}/p^\circ - 1| \leqslant \alpha$ with probability at least* 0.99*, where $p^\circ = \sum_{i,j} Q^\circ_{ij}/(n^2 - n)$. Then we must have $\alpha \geqslant \Omega(R\eta)$, where $R = \max_{i,j} Q^\circ_{ij}/p^\circ$.*

Proof. Let $p^\circ \in [0,1]$, and let $Q^\circ \in [0,1]^{n\times n}$ be the matrix, in which all entries are p°, except for the rows and columns corresponding to a set of ηn nodes setting to be Rp°. Let Q be the matrix, in which all entries are p°, except for the rows and columns corresponding to a set of ηn nodes setting to be 0.

We construct the following pair of distributions $\mathcal{D}_0$ and $\mathcal{D}_1$:

- $\mathcal{D}_0$: The distribution of $\boldsymbol{G} \sim \mathbb{G}(Q^\circ)$.
- $\mathcal{D}_1$: The distribution of $\boldsymbol{G} \sim \mathbb{G}(Q)$.

Then we have $\frac{1}{n^2}|\|Q^\circ\|_1 - \|Q\|_1| \geqslant \Omega(R\eta n^2 p^\circ)$.

On the other hand, there is a coupling between $\tilde{G} \sim \mathbb{G}(Q^\circ)$ and $\tilde{G}' \sim \mathbb{G}(Q)$ such that $\mathrm{dist}(\tilde{G}, \tilde{G}') \leqslant \eta n$. Therefore, the two distributions are indistinguishable under the η-corruption model. Since the edge density of $\mathbb{G}(Q^\circ)$ differs from $\mathbb{G}(Q)$ by $\Omega(R\eta p^\circ)$, no algorithm can achieve error rate $o(R\eta p^\circ)$ with probability $1 - o(1)$ for both distributions under the corruption of η-fraction of the nodes. □

Theorem (Restatement of Theorem 1.8)**.** *Suppose there is an ε-differentially node-private algorithm satisfies the following guarantee for any symmetric matrix $Q^\circ \in [0,1]^{n\times n}$. Given an inhomogeneous random graph $\mathbb{G}(n, Q^\circ)$, the algorithm outputs an estimate $\hat{p}$ satisfying $|\hat{p}/p^\circ - 1| \leqslant \alpha$ with probability $1 - \beta$, where $p^\circ = \sum_{i,j} Q^\circ_{ij}/(n^2 - n)$. Then we must have*

$$\alpha \geqslant \Omega\left(\frac{R\log(1/\beta)}{n\varepsilon}\right),$$

where $R = \max_{i,j} Q^\circ_{ij}/p^\circ$.

Proof of Theorem 1.8. We will prove the lower bound by constructing a pair of distributions $\mathcal{D}_0$ and $\mathcal{D}_1$ such that the total variation distance between them is small, but the difference in edge density is significant. Then since ε-differentially node-private algorithm needs to have similar distributions in the output, it could not succeed in accurately estimating the edge density accurately under both distributions.

Let $\eta \in [0, 0.001)$. Let $p^\circ \in [0,1]$, and let $Q^\circ \in [0,1]^{n\times n}$ be the matrix, in which all entries are p°, except for the rows and columns corresponding to a set of ηn nodes setting to be 0. Let Q be the matrix, in which all entries are p°, except for the rows and columns corresponding to a set of ηn nodes setting to be Rp°.

We construct the following pair of distributions $\mathcal{D}_0$ and $\mathcal{D}_1$:

- $\mathcal{D}_0$: The distribution of $\boldsymbol{G} \sim \mathbb{G}(Q^\circ)$.
- $\mathcal{D}_1$: The distribution of $\boldsymbol{G} \sim \mathbb{G}(Q)$.

Let $p' = \frac{\|Q^\circ\|_1}{n^2}$ and $p = \frac{\|Q\|_1}{n^2}$. We have $|p - p'| \geqslant R\eta p^\circ$.

On the other hand, there is a coupling between $\tilde{G} \sim \mathbb{G}(Q)$ and $\tilde{G}' \sim \mathbb{G}(Q^\circ)$ such that $\operatorname{dist}(\tilde{G}, \tilde{G}') \leqslant \eta n$. Taking expectation w.r.t. the coupling ω on both sides gives

$$\mathop{\mathbb{E}}_{\omega}\mathop{\mathbb{E}}_{\mathcal{A}} \mathbb{1}\left\{\left|\mathcal{A}(\tilde{\boldsymbol{G}}') - p\right| < \frac{R\eta}{2}p^\circ\right\} \leqslant \mathop{\mathbb{E}}_{\omega} e^{\varepsilon\cdot \operatorname{dist}(\tilde{G},\tilde{G}')} \cdot \mathop{\mathbb{E}}_{\mathcal{A}} \mathbb{1}\left\{\left|\mathcal{A}(\tilde{\boldsymbol{G}}) - p\right| < \frac{R\eta}{2}p^\circ\right\},$$
$$\mathop{\mathbb{P}}_{\mathcal{A},\tilde{\mathbb{G}}(Q^\circ)}\left(\left|\mathcal{A}(\tilde{\boldsymbol{G}}') - p\right| < \frac{R\eta}{2}p^\circ\right) \leqslant \mathop{\mathbb{E}}_{\omega,\mathcal{A}} e^{\varepsilon\cdot \operatorname{dist}(\tilde{G},\tilde{G}')} \cdot \mathbb{1}\left\{\left|\mathcal{A}(\tilde{\boldsymbol{G}}) - p\right| < \frac{R\eta}{2}p^\circ\right\}.$$

By the utility assumption of algorithm $\mathcal{A}$ and $p < p^\circ$, the left-hand side is at least $1-\beta$. Using the Cauchy-Schwartz inequality, the right-hand side can be upper bounded as follows,

$$\begin{aligned}\mathop{\mathbb{E}}_{\omega,\mathcal{A}} e^{\varepsilon\cdot \operatorname{dist}(\tilde{G},\tilde{G}')} \cdot \mathbb{1}\left\{\left|\mathcal{A}(\tilde{\boldsymbol{G}}) - p\right| < \frac{R\eta}{2}p^\circ\right\} &\leqslant \sqrt{\mathop{\mathbb{E}}_{\omega,\mathcal{A}} e^{2\varepsilon\cdot \operatorname{dist}(\tilde{G},\tilde{G}')}}\sqrt{\mathop{\mathbb{E}}_{\omega,\mathcal{A}} \mathbb{1}\left\{\left|\mathcal{A}(\tilde{\boldsymbol{G}}) - p\right| < \frac{R\eta}{2}p^\circ\right\}} \\ &\leqslant \exp(\varepsilon\eta n)\sqrt{\mathop{\mathbb{P}}_{\mathcal{A},\tilde{\mathbb{G}}(Q)}\left(\left|\mathcal{A}(\tilde{\boldsymbol{G}}) - p\right| < \frac{R\eta}{2}p^\circ\right)}.\end{aligned}$$

Thus

$$(1-\beta)^2 \leqslant \exp(\varepsilon\eta n) \cdot \mathop{\mathbb{P}}_{\mathcal{A},\tilde{\mathbb{G}}(Q)}\left(\left|\mathcal{A}(\tilde{\boldsymbol{G}}) - p\right| < \frac{R\eta}{2}p^\circ\right).$$

Then

$$\mathop{\mathbb{P}}_{\mathcal{A},\tilde{\mathbb{G}}(Q)}\left(\left|\mathcal{A}(\tilde{\boldsymbol{G}}) - p\right| < \frac{R\eta}{2}p^\circ\right) \geqslant (1-\beta)^2 \cdot \exp(-\varepsilon\eta n).$$

Since $|p - p^\circ| \geqslant R\eta p^\circ$, the two events $\{\hat{p} : |\hat{p} - p^\circ| < \frac{R\eta}{2}p^\circ\}$ and $\{\hat{p} : |\hat{p} - p| < \frac{R\eta}{2}p^\circ\}$ are disjoint, which implies

$$\mathop{\mathbb{P}}_{\mathcal{A},\tilde{\mathbb{G}}(Q)}\left(\left|\mathcal{A}(\tilde{\boldsymbol{G}}) - p\right| < \frac{R\eta}{2}p^\circ\right) \leqslant 1 - \mathop{\mathbb{P}}_{\mathcal{A},\tilde{\mathbb{G}}(Q)}\left(\left|\mathcal{A}(\tilde{\boldsymbol{G}}) - p^\circ\right| < \frac{R\eta}{2}p^\circ\right) \leqslant \beta.$$

Therefore, we have $\beta \geqslant (1-\beta)^2 \exp(-\varepsilon\eta n)$. As result, we need to have $\eta \geqslant \Omega\left(\frac{\log(\beta)}{\varepsilon n}\right)$. Thus we have

$$|p - p'| \geqslant \Omega\left(\frac{R\log(\beta)p^\circ}{\varepsilon n}\right).$$

Since $p^\circ \geqslant p'$, it follows that

$$|p - p'| \geqslant \Omega\left(\frac{R\log(\beta)p'}{\varepsilon n}\right),$$

which finishes the proof. □

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: We have formal proofs for what we claim in the abstract and introduction.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: We prove matching upper and lower bounds of error rate. We show polynomial running time of our algorithm.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory Assumptions and Proofs**

Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: We state everything formally and provide full proofs. When we say something intuitive and informal, we always have formal counterparts in the appendices.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental Result Reproducibility**

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [NA]

Justification: Our paper does not include experiments.

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).

(d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [NA]

   Justification: This paper does not include experiments requiring code.

   Guidelines:

   - The answer NA means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (https://nips.cc/public/guides/CodeSubmissionPolicy) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (https://nips.cc/public/guides/CodeSubmissionPolicy) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental Setting/Details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: [NA]

   Justification: This is a theoretical paper that does not include experiments.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment Statistical Significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [NA]

   Justification: This is a theoretical paper that does not include experiments.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments Compute Resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [NA]

   Justification: This is a theoretical paper that does not include experiments.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code Of Ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics https://neurips.cc/public/EthicsGuidelines?

   Answer: [Yes]

   Justification: The authors conform, in every respect, with the NeurIPS Code of Ethics while writing this paper.

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader Impacts**

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [NA]

Justification: This is a theoretical paper on random graph estimation that does not have societal impact.

Guidelines:

- The answer NA means that there is no societal impact of the work performed.
- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Justification: This is a theoretical paper that does not poses such risks.

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [NA]

Justification: This is a theoretical paper that does not use existing assets.

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, paperswithcode.com/datasets has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New Assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [NA]

    Justification: This is a theoretical paper that does not involve new assets.

    Guidelines:

    - The answer NA means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and Research with Human Subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [NA]

    Justification: This is a theoretical paper that does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional Review Board (IRB) Approvals or Equivalent for Research with Human Subjects**

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [NA]

Justification: This is a theoretical paper that does not involve crowdsourcing nor research with human subjects.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.