Robust Estimation for the Erdös-Rényi Model

Heon Lee, George Chemmala, Arjan Chakravarthy
Brown University
heon_lee@brown.edu george_chemmala@brown.edu arjan_chakravarthy@brown.edu

Abstract

We study the robust estimation of the edge probability p in the Erdös-Rényi random graph model under adversarial perturbation of vertices. We define a new class of adversarial models, the (q, ε)-adversarial model, which naturally generalizes to corresponding definitions for (q, ε)-oblivious and (q, ε)-omniscient adversaries, analogous to the ε-adversarial framework. This new model allows us to explore robustness across varying levels of adversarial strengths while maintaining a unified framework for analysis.

In the presence of a (q, ε)-oblivious adversary, we propose a novel algorithm that runs in $O(n)$ time and guarantees $|p - \hat{p}| \leq O(1/\sqrt[4]{n})$ with probability at least $1 - \frac{1}{\sqrt{n}}$, representing a significant improvement in both efficiency and accuracy compared to existing methods designed for ε-omniscient adversaries. Existing methods either have a slower runtime with a comparable error guarantee or a similar runtime with a worse error guarantee. Furthermore, we introduce an iterative algorithm that uses variance-based filtering to identify and remove corrupted vertices. This approach empirically demonstrates strong performance and achieves a runtime of $O(\varepsilon n^3)$. Work is ongoing to establish theoretical guarantees for this method.

1. Introduction

The Erdös-Rényi graph model is a fundamental framework in network theory and probability, commonly used for studying random graphs [2]. It is constructed by independently connecting each pair of n vertices with probability p [3]. Despite its simplicity and inability to fully represent the complexity of real-world networks, the Erdös-Rényi model serves as a foundational tool for analyzing network properties. Applications span various domains, including social, biological, and communication networks [2]. This paper investigates a scenario where an adversary corrupts an ε-fraction of the vertices in an Erdös-Rényi graph. The objective is to estimate the original edge formation probability p given the corrupted graph.

1.1. Problem Setup

In this paper, we are primarily concerned with the Erdös-Rényi random graph model which was first formally defined by Edgar Gilbert in 1959 and independently rediscovered by its namesakes Erdös and Rényi later that year [3] [4]. Notice that the p value of the Erdös-Rényi graph has a direct impact on the degree distribution of the graph which follows a Binomial distribution $B(n-1, p)$ [5]. The expected degree of any given vertex in the graph is $(n-1)p$, since each vertex is connected to $n-1$ other vertices with probability p.

Refer to Figure 1 for an example of an Erdös-Rényi graph with $n = 10$ vertices and $p = 1/2$.

Definition 1.1. *The Erdös-Rényi random graph model $G(n, p)$ is a probability distribution over graphs with n vertices, where each edge is included independently with probability p.*

The central challenge of this paper is to estimate the true edge probability p of a graph $G \sim G(n, p)$ in the presence of adversarial corruption. Specifically, we consider an adversary $\mathcal{A}$ capable of corrupting up to an ε-fraction of the vertices and producing a corrupted graph, $\mathcal{A}(G)$. The adversary can alter the structure of the graph by modifying any edge that is incident to a corrupted vertex, including both adding and removing an edge.

With this setup, we build on two ε-adversaries proposed by Acharya, Jain, Kamath, Suresh, and Zhang [1] with our own (q, ε)-adversary.

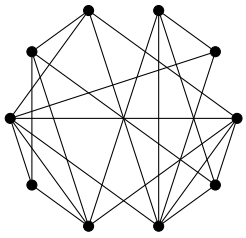

Figure 1. Example of an Erdös-Rényi graph $G(10, 1/2)$

ε-*omniscient adversary*: The adversary knows the true value of the edge probability p and observes the realization of the graph $G \sim G(n, p)$. They then choose the set of corrupted vertices and rewire the edges of these corrupted vertices.

ε-*oblivious adversary*: The adversary knows the true value of the edge probability p. They must choose the set of corrupted vertices and the distribution of edges of the corrupted vertices without the realization of the graph G.

(q, ε)-*omniscient adversary*: The adversary knows the true value of the edge probability p and observes the realization of the graph $G \sim G(n, p)$. They then choose the set of corrupted vertices, and the edges of the corrupted vertices are rewired with a new edge probability q.

(q, ε)-*oblivious adversary*: The adversary knows the true value of the edge probability p. They must choose the set of corrupted vertices and the distribution of edges of the corrupted vertices without the realization G. However, the adversary can choose to rewire the edges adjacent to the corrupted vertices with a new edge probability q.

Despite this adversarial perturbation, we aim to estimate the true edge probability p. While this might initially seem straightforward, given that the adversary is limited to corrupting an ε-fraction of corrupted vertices, the impact is more complex. Specifically, corrupting a single vertex alters the degree distribution of the uncorrupted vertices. In other words, when the adversary perturbs the edges connected to corrupted vertices, it indirectly creates or removes edges incident to uncorrupted vertices as well. Consequently, the observed p value in the corrupted graph can deviate significantly from the true p value of the original graph prior to corruption.

The primary metric we seek to optimize is

$$|p - \hat{p}|,$$

the difference between the theoretical p value of the graph and the robust estimation on the corrupted graph, $\hat{p}$.

1.2. Related Work

On our particular problem, we examined a paper by Acharya, Jain, Kamath, Suresh, and Zhang [1] that introduced the ε-omniscient adversary and proposed the naive mean, naive median, the Prune algorithms, for both the mean and the median, and the Spectral algorithm.

The naive mean and median algorithms estimate the edge probability p using the degree distribution of the corrupted graph. In an uncorrupted graph, the mean and median degrees of the vertices are expected to be $(n-1)p$. The naive algorithm calculates the mean or median degree of the vertices in the corrupted graph and normalizes this value by dividing it by $(n-1)$, yielding an estimate for p.

The Prune algorithm removes an ε-fraction of the vertices from the top and bottom of the degree distribution—the extremal vertex degrees—and then estimates the edge probability p based on the remaining vertices with the mean and median of the pruned degree distribution. By removing the extremal vertex degrees, Acharya et.al proves that the pruned mean would have an error proportional to $O(\varepsilon^2)$ and the pruned median would have an error proportional to $O(\varepsilon)$. The algorithm has a runtime of $O(n \log(n\varepsilon))$ to find the ε-fraction of vertices with the highest and lowest degrees [1].

The Spectral algorithm is based on the fact that the metric $\|(A - p_S)_{S \times S}\|_2$ is small for when S is a set of uncorrupted vertices. Here $p_S := (\sum_{i,j \in S} A_{i,j})/|S|^2$; this is the measured p value of the sub-graph G_S—the graph that consists of vertices in S. Intuitively, this metric should be small for uncorrupted subsets because the expected value of the entries of the adjacency matrix is p for uncorrupted vertices. The Spectral algorithm estimates the edge probability p by minimizing this norm by removing vertices in S that contribute to increasing this norm. We can find these vertices by computing the top eigenvector of $(A - p_S)$. Eventually, we will get a set of vertices that appear to be uncorrupted and estimate the edge probability p based on the mean of the degree distribution of these vertices.

1.3. Contributions

This paper introduces the (q, ε)-adversarial model, extending the ε-adversarial framework to encompass varying levels of adversarial strength under a unified analysis. We propose three robust algorithms for estimating the edge probability p: the Mean-Adjusted Median, the Bias-Corrected Mean-Adjusted Median, and the Variance-Based Filtering method. Our contributions include:

1. Developing a novel Mean-Adjusted Median algorithm with a runtime of $O(n)$ and theoretical guarantees on its error bound.
2. Enhancing the Mean-Adjusted Median with a Bias-Correction mechanism, leading to improved accuracy.
3. Designing a Variance-Based Filtering algorithm for identifying corrupted vertices, supported by empirical evidence.

2. Notation

$G(n, p)$: Erdös-Rényi random model with n vertices and edge probability p.

$\mathcal{A}(G)$: Perturbed graph with a fraction ε of corrupted vertices.

$\bar{p}$ and $\tilde{p}$: Normalized mean and median of the degree distribution.

$\bar{d}$ and $\tilde{d}$: Mean and median of the degree distribution.

s^2: Variance of the degrees of the graph.

$\hat{\sigma}^2$: Variance of the degrees of the perturbed graph.

$\deg(v)$: Degree of vertex v

3. Mean-Adjusted Median

In this section, we propose an algorithm based on the median and mean of the degree distribution for robustly estimating the edge probability p. Since the median is more robust to outliers than the mean for (q, ε)-oblivious adversary, the perturbed median of the degree distribution consistently tends closer to the median of the distribution prior to adversarial corruption than the corresponding mean values. Thus, we derive our Mean-Adjusted Median algorithm by analyzing this difference between the median and the mean.

Plotting out a histogram (Figure 2) of the degree distribution of the graph altered by the (q, ε)-adversary, we can see that the median of the degree distribution is roughly two times closer to the mean of the original degree distribution than the mean. Since the mean of the original degree distribution is $(n-1)p$, in order to derive $\hat{p}$, we divide an estimator of the original degree distribution by a factor of $n-1$. This leads us to consider using the difference between the mean and the median in order to find p.

The resulting algorithm for the Mean-Adjusted median is described in Algorithm 1.

Algorithm 1 Mean-Adjusted Median

Require: Laplacian matrix L, epsilon ε
$D \leftarrow$Degrees of vertices using L
$\bar{p} \leftarrow$ normalized mean
$\tilde{p} \leftarrow$ normalized median
return $\frac{(2-\varepsilon)\tilde{p}-\bar{p}}{1-\varepsilon}$

For the empirical and theoretical results presented in this paper, we adopt the (q, ε)-oblivious adversary.

Definition 3.1. *Let $G \sim G(n, p)$. The (q, ε)-**oblivious adversary** randomly selects a subset $B \subset V$ of corrupted vertices with $|B| = \varepsilon n$ and rewires edges adjacent to B according to a new edge probability q, such that:*

$$\mathbb{P}((u,v) \in E') = \begin{cases} q, & \text{if } u \in B \text{ or } v \in B, \\ 1, & \text{if } u, v \notin B \text{ and } (u,v) \in E \\ 0, & \text{if } u, v \notin B \text{ and } (u,v) \notin E \end{cases}$$

where E' is the edge set of $\mathcal{A}(G)$.

While the (q, ε)-adversary is not as powerful as the ε-omniscient adversary, it can still significantly alter the graph's structure and simulate major shifts in connectivity similar to those caused by the ε-omniscient adversary. Therefore, evaluating whether the Mean-Adjusted Median algorithm can accurately estimate the edge probability p under the influence of a (q, ε)-adversary serves as a robust test of the algorithm's resilience and offers a valuable direction for further study.

Theorem 3.2. *Given a graph perturbed by a (q, ε)-adversary with $\varepsilon < 0.5$, the Mean-Adjusted Median algorithm has a runtime of $O(n)$ and a guaranteed error of*

$$|p - \hat{p}| \leq O\left(\frac{1}{\sqrt[4]{n}}\right)$$

with probability at least $1 - \frac{1}{\sqrt{n}}$.

Proof. Let D' be the degree distribution of the perturbed graph. We then calculate the normalized mean $\bar{p}'$ and the normalized median $\tilde{p}'$ of the D' by taking the mean and median of D' and dividing by $n-1$. Both the mean and median can be computed in $O(n)$ time. Finally, we run an $O(1)$ operation to return the mean-adjusted median. The algorithm has a runtime of $O(n)$.

The error of the algorithm is given by the absolute difference between the mean-adjusted median and the true edge probability p. We start by examining the degrees of the vertices in the graph, which are directly proportional to the edge probability p of the graph.

Let $\tilde{d}'$ and $\bar{d}'$ be the median and mean of D'. Let $\eta = p - q$, $\hat{d} := \frac{(2-\varepsilon)\tilde{d}' - \bar{d}'}{1-\varepsilon}$, and $\hat{p} := \frac{\hat{d}}{n-1}$.

By Lemma A.5 in the Appendix A, $\hat{p}$ is a random variable such that $\mathbb{E}[\hat{p}] \approx p - \frac{\varepsilon\eta + (2\varepsilon - \varepsilon^2)\frac{\text{Sign}(\eta)}{2a}}{(1-\varepsilon)(n-1)}$ where

$$a = \frac{1-\varepsilon}{\sqrt{2\pi}\sigma_X} + \frac{\varepsilon}{\sqrt{2\pi}\sigma_Y} e^{-\frac{(\mu_X - \mu_Y)^2}{2\sigma_Y^2}}.$$

Here,

$$\begin{aligned} \mu_X &= ((1-\varepsilon)n - 1)p + \varepsilon n q; \\ \sigma_X^2 &= ((1-\varepsilon)n - 1)p(1-p) + \varepsilon n q(1-q); \\ \mu_Y &= (n-1)q; \\ \sigma_Y^2 &= (n-1)q(1-q). \end{aligned}$$

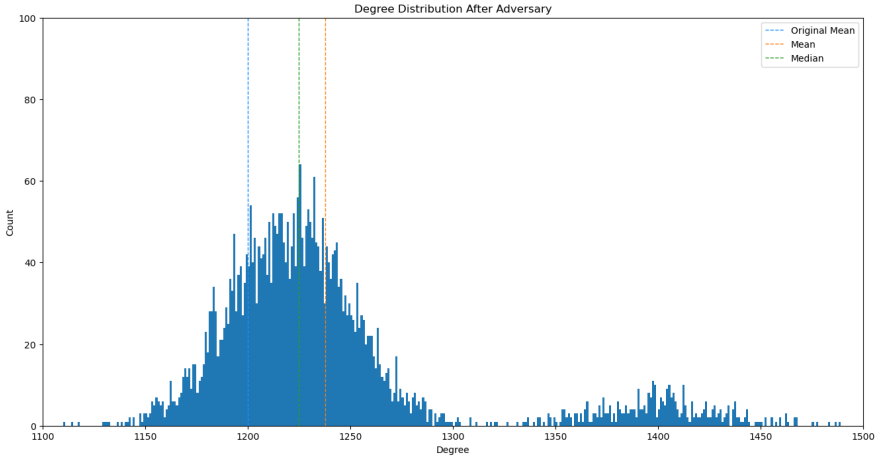

Figure 2. $(n-1)p$ plotted along with the mean and median of the altered distribution

Then we observe that

$$p - \hat{p} = (p - \mathbb{E}[\hat{p}]) + (\mathbb{E}[\hat{p}] - \hat{p}) \approx (\mathbb{E}[\hat{p}] - \hat{p}) + \frac{\varepsilon\eta + (2\varepsilon - \varepsilon^2)\frac{\mathrm{Sign}(\eta)}{2a}}{(1-\varepsilon)(n-1)}.$$

By the triangle inequality,

$$|p - \hat{p}| \leq |\mathbb{E}[\hat{p}] - \hat{p}| + \left| \frac{\varepsilon\eta + (2\varepsilon - \varepsilon^2)\frac{\mathrm{Sign}(\eta)}{2a}}{(1-\varepsilon)(n-1)} \right| \leq |\mathbb{E}[\hat{p}] - \hat{p}| + \frac{\varepsilon + \frac{2\varepsilon - \varepsilon^2}{2a}}{(1-\varepsilon)(n-1)}.$$

In the proof of Lemma A.6, we showed that $a \geq \frac{1}{2\sqrt{\pi n}}$, which implies $\frac{1}{a} \in [0, 2\sqrt{\pi n}]$. As a result,

$$|p - \hat{p}| \leq |\mathbb{E}[\hat{p}] - \hat{p}| + \frac{\varepsilon + (2\varepsilon - \varepsilon^2)\sqrt{\pi n}}{(1-\varepsilon)(n-1)} \leq |\mathbb{E}[\hat{p}] - \hat{p}| + \frac{8\varepsilon}{\sqrt{n}}$$

Moreover, as $n \to \infty$, the bias $\frac{\varepsilon\eta + (2\varepsilon - \varepsilon^2)\frac{\mathrm{Sign}(\eta)}{2a}}{(1-\varepsilon)(n-1)} \in \left[\frac{\varepsilon\eta}{(1-\varepsilon)(n-1)}, \frac{\varepsilon\eta + (2\varepsilon - \varepsilon^2)\sqrt{\pi n}}{(1-\varepsilon)(n-1)} \right]$ goes to zero. Thus, the mean of $\hat{p}$ approaches p.

By Lemma A.6, we have $\mathrm{Var}(\hat{p}) \leq 64\frac{1+\sqrt{n}+n}{n^2} \leq \frac{256}{n}$ since $n \geq 1$. Using Chebyshev's inequality, for any $\delta > 0$,

$$\Pr(|\hat{p} - \mathbb{E}[\hat{p}] \geq \delta) \leq \frac{\mathrm{Var}(\hat{p})}{\delta^2} \leq \frac{256}{\delta^2 n}.$$

Hence, with probability at least $1 - \frac{256}{\delta^2 n}$,

$$|p - \hat{p}| \leq |\mathbb{E}[\hat{p}] - \hat{p}| + \frac{8\varepsilon}{\sqrt{n}} \leq \delta + \frac{8\varepsilon}{\sqrt{n}}.$$

Letting $\delta = \frac{16}{\sqrt[4]{n}}$, we conclude that with probability at least $1 - \frac{1}{\sqrt{n}}$,

$$|p - \hat{p}| \leq \frac{16}{\sqrt[4]{n}} + \frac{8\varepsilon}{\sqrt{n}} \leq O\left(\frac{1}{\sqrt[4]{n}}\right)$$

as desired.

□

4. Bias-Corrected Mean-Adjusted Median

In the previous section, letting $\hat{p}$ denote the output of the Mean-Adjusted Median algorithm, we showed that the expected value of $\hat{p}$ is approximately

$$\mathbb{E}[\hat{p}] \approx p - \frac{\varepsilon\eta + (2\eta - \eta^2)\frac{\mathrm{Sign}(\eta)}{2a}}{(1-\varepsilon)(n-1)}.$$

Although the bias diminishes as n increases, we aim to remove this bias entirely. However, the bias term depends on ε, p, q, and n. While ε and n are known, p and q are unknown quantities.

Due to the relative weakness of the (q, ε)-oblivious adversary, we can compute $\mathbb{E}[\bar{p'}]$ as a function of p, q, ε, and n. By Lemma A.1, we have

$$\mathbb{E}[\bar{p'}] = \frac{\mathbb{E}[\bar{d'}]}{n-1} = p - 2\varepsilon\eta + \varepsilon^2\eta - \frac{\eta(\varepsilon - \varepsilon^2)}{n-1}.$$

Rewriting this,

$$\mathbb{E}[\bar{p'}] = p\left(1 - 2\varepsilon + \varepsilon^2 - \frac{\varepsilon - \varepsilon^2}{n-1}\right) + q\left(2\varepsilon - \varepsilon^2 + \frac{\varepsilon - \varepsilon^2}{n-1}\right).$$

We can then express q as a function of $p, \varepsilon, \mathbb{E}[\bar{p'}]$, and n:

$$q = \frac{\mathbb{E}[\bar{p'}] - p\left(1 - 2\varepsilon + \varepsilon^2 - \frac{\varepsilon - \varepsilon^2}{n-1}\right)}{2\varepsilon - \varepsilon^2 + \frac{\varepsilon - \varepsilon^2}{n-1}} = p + \frac{\mathbb{E}[\bar{p'}] - p}{2\varepsilon - \varepsilon^2 + \frac{\varepsilon - \varepsilon^2}{n-1}}.$$

For large n, the sample mean, $\bar{p'}$, which we can compute given the perturbed graph, closely approximates $\mathbb{E}[\bar{p'}]$, and $\hat{p}$ is a good estimation of p with high probability as established in Theorem 3.2, $\hat{p}$. Thus, we define an estimator for q:

$$\hat{q} = \hat{p} + \frac{\bar{p'} - \hat{p}}{2\varepsilon - \varepsilon^2 + \frac{\varepsilon - \varepsilon^2}{n-1}}.$$

Using $\hat{p}, \hat{q}, \varepsilon$, and n, we estimate η as $\hat{\eta} := \hat{p} - \hat{q}$ and a by $\hat{a}$. Using $\hat{\eta}$ and $\hat{a}$, we are then able to estimate the bias term $\frac{\varepsilon\eta + (2\eta - \eta^2)\frac{\text{Sign}(\eta)}{2a}}{(1-\varepsilon)(n-1)}$. Finally, instead of returning $\hat{p}$, we return the adjusted estimator

$$\hat{p}^* := \hat{p} + \frac{\varepsilon\hat{\eta} + (2\hat{\eta} - \hat{\eta}^2)\frac{\text{Sign}(\hat{\eta})}{2\hat{a}}}{(1-\varepsilon)(n-1)}.$$

The resulting algorithm for the Bias-Corrected Mean-Adjusted Median is described in Algorithm 2.

Algorithm 2 Bias-Corrected Mean-Adjusted Median

Require: Laplacian matrix L, epsilon ε
$D \leftarrow$Degrees of vertices in L
$\bar{p'} \leftarrow$ normalized mean
$\hat{p} \leftarrow$ MEAN-ADJUSTED-MEDIAN(L, ε)
$\hat{q} \leftarrow \hat{p} + \frac{\bar{p'} - \hat{p}}{2\varepsilon - \varepsilon^2 + \frac{\varepsilon - \varepsilon^2}{n-1}}$
$\hat{\eta}, \hat{a} \leftarrow$ estimation of η, a using $\hat{p}, \hat{q}, \varepsilon, n$
return $\hat{p} + \frac{\varepsilon\hat{\eta} + (2\hat{\eta} - \hat{\eta}^2)\frac{\text{Sign}(\hat{\eta})}{2\hat{a}}}{(1-\varepsilon)(n-1)}$

While $\hat{p}^*$ is not strictly unbiased due to the approximations, it is effective in reducing the bias.

We empirically test the Bias-Corrected Mean-Adjusted Median algorithm and find that it consistently outperforms the original Mean-Adjusted Median algorithm, achieving improved accuracy in estimating p. The analysis can be found in Section 6.

5. Variance-Based Filtering

In this section, we propose a variance-based algorithm for robustly estimating the parameter p on an adversarially corrupted graph $\mathcal{A}(G)$. The key idea is that the variance of the degree distribution of the vertices in $\mathcal{A}(G)$ can provide insight into which vertices have been corrupted. Thus, we can study the difference in variance between the perturbed sample graph and the expected variance, assuming a binomial distribution. Even in an adversarially perturbed graph, the degree distribution of the subgraph consisting of the uncorrupted vertices follows a binomial distribution. Therefore, the observed variance of such a subgraph should equal the theoretical variance. However, the degree distribution of the vertices in $\mathcal{A}(G)$ may not follow a binomial distribution. The variance-based filtering algorithm seeks to uncover this subgraph of uncorrupted vertices by minimizing the difference in the theoretical variance and the observed variance. Formally, we define $s^2 = \frac{1}{n-1}\sum_{v \in V}(\deg(v) - \bar{d})^2$ where V denotes the set of vertices. Similarly, we define $\hat{\sigma}^2 = n\hat{p}(1-\hat{p})$ where $\hat{p}$ denotes the mean of the degree distribution assuming a binomial distribution. On an unperturbed graph, since the degrees of the vertices roughly follow a binomial distribution, $\left|s^2 - \hat{\sigma}^2\right|$ will be small, with high probability. We exploit this fact in Algorithm 3.

This iterative algorithm seeks to remove vertices that cause the greatest difference to the variance. Assuming we start with a graph $G \sim G(n,p)$, we remove the vertex that maximizes:

$$\max_{v \in V} \left| s^2_{G_{(V \setminus \{v\})}} - \hat{\sigma}^2_{G_{(V \setminus \{v\})}} \right|.$$

Since we know that in an unperturbed graph, $|s^2 - \hat{\sigma}^2|$ is small with high probability, removing the vertex that causes the greatest deviation with this term would enable us to reduce the difference in the observed and expected variance, ultimately converging on a set of vertices where this term is minimized. It is important to note, though, that the vertices that minimize this difference may not always converge to the true set of uncorrupted vertices. We implement this removal process εn times in order to remove an ε-fraction of the vertices, simulating removing the corrupted vertices. However, since we minimize the difference between the variances, the subgraph that remains has a degree distribution that is the best candidate among all

Algorithm 3 Variance Algorithm

Require: Laplacian matrix L, epsilon ε
  $n \leftarrow$number of vertices
  $V \leftarrow$set of all vertices
  while $t = 0, t < \varepsilon n$ **do**
    for vertex v in V **do**
      Compute subgraph after removing v using L
      Compute and store $\left| s^2_{G_{(V\setminus\{v\})}} - \hat{\sigma}^2_{G_{(V\setminus\{v\})}} \right|$ for the given vertex v
    end for
    $x \leftarrow \max_{v\in V} \left| s^2_{G_{(V\setminus\{v\})}} - \hat{\sigma}^2_{G_{(V\setminus\{v\})}} \right|$
    $V \leftarrow V \setminus x$
  end while

subgraphs to follow a binomial distribution after εn vertex removals. The time complexity of this variance-based filtering algorithm is $O(\varepsilon n^3)$ since we iterate εn times and in each iteration, we simulate removing every vertex, which takes $O(n)$ time, and calculating the variance on this new subgraph, also $O(n)$ time.

6. Results

To evaluate these different models on estimating the parameter p, we compare the mean-squared error across the methods on the (q, ε)-adversary. Since the spectral method is bounded by $\varepsilon < 1/60$, we evaluate the experiments in Figure 5 with an $\varepsilon = 0.01$ and the experiments in Figures 3, 4, 6 with an $\varepsilon = 0.1$ [1]. For each data point, we run each method on 20 trials and take the average of the estimates for a holistic measure of the mean-squared error.

Figure 3 displays a comparison of the mean-squared error across six different methods: the naive mean, naive median, prune-then-mean, prune-then-median, mean-adjusted median, and bias-corrected mean-adjusted median algorithms. From Figure 3, we observe that as the number of vertices in the Erdös-Rényi graph increases, the empirical error of the naive mean and median algorithms remains fairly stable. This is because the error of these algorithms scales with regards to ε and is not correlated with n [1].

Figure 4 shows a more focused depiction of the more robust mean and median-based algorithms from 3. Empirically, all of these algorithms scale inversely with respect to n since the mean-squared error decreases as the number of vertices n increases. We notice the Bias-Corrected Mean-Adjusted Median algorithm performs significantly better than the plain Mean-Adjusted Median algorithm, empirically demonstrating that we can successfully remove the bias term from Mean-Adjusted Median algorithm.

Figure 5 compares the Mean-Adjusted Median algorithms with the Spectral Method and the Variance-Based filtering method. Due to computational reasons, we limit running the Spectral method on graphs with less than 500 nodes. We notice that the Spectral Method and the Variance-based Filtering method perform better than the Mean-Adjusted Median method.

Finally, figure 6 displays all of the datapoints for the variance-based filtering algorithm with the 20 trials shown for a given n. The line plot displayed takes the mean error of these 20 trials to illustrate a general trend in the error as the number of vertices increases. It is important to note that the error in this figure is the absolute error, rather than the mean-squared error.

Methods	**Runtime**	**Authors**
Mean/Median	O(n)	Acharya et al.
Prune then Mean/Median	$O(n\log(\varepsilon n))$	Acharya et al.
Spectral Method	$\tilde{O}\left(\varepsilon n^3\right)$	Acharya et al.
Mean-Adjusted Median	$O(n)$	Lee et al.
Bias-Corrected Mean-Adjusted Median	$O(n)$	Lee et al.
Variance Method	$O(\varepsilon n^3)$	Lee et al.

7. Conclusion

In this work, we introduced a novel adversarial framework, the (q, ε)-adversary model, and proposed three algorithms for robust estimation of the edge probability p in adversarially perturbed Erdős-Rényi random graphs: the Mean-Adjusted

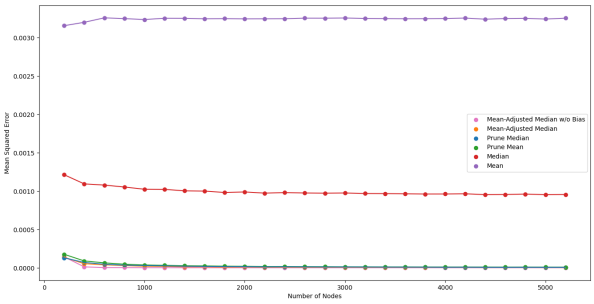

Figure 3. $\varepsilon = 0.1$, $p = 0.3$, $q = 0.6$

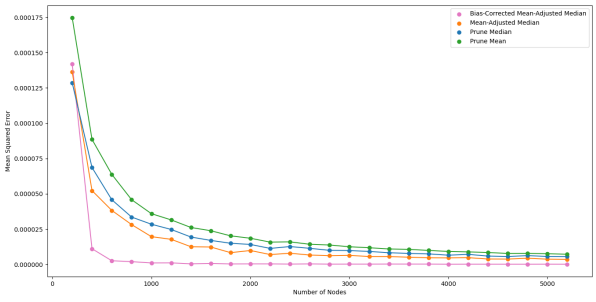

Figure 4. $\varepsilon = 0.1$, $p = 0.3$, $q = 0.6$

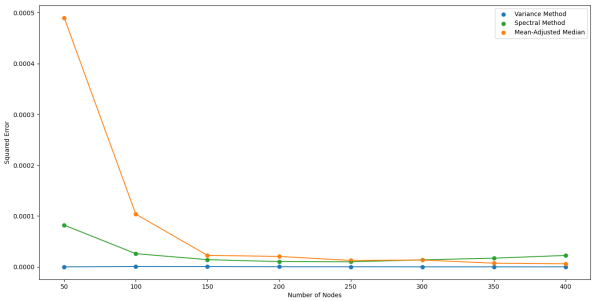

Figure 5. $\varepsilon = 0.01$, $p = 0.3$, $q = 0.6$

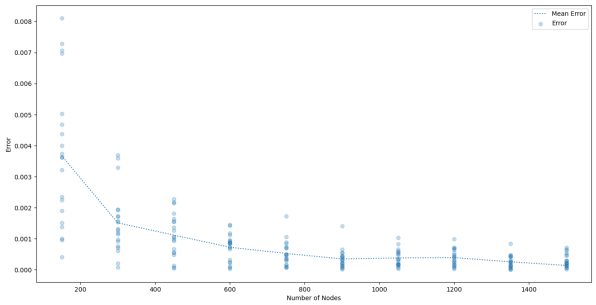

Figure 6. Each dot represents a trial; $\varepsilon = 0.1$, $p = 0.3$, $q = 0.6$

Median, Bias-Corrected Mean-Adjusted Median, and Variance-Based Filtering algorithms. By extending the adversarial models proposed in prior research, our models tackle adversarially corrupted graphs using the mean, median, and the variance.

The theoretical analysis in Theorem 3.2 establishes that the Mean-Adjusted Median algorithm guarantees a bounded error of $O\left(\frac{1}{\sqrt[4]{n}}\right)$ with high probability, while maintaining an efficient runtime of $O(n)$. This gurantee highlights that on the (q, ε)-oblivious adversary, the Mean-Adjusted Median algorithm outperforms the naive mean and median algorithms. The Bias-Corrected Mean-Adjusted Median builds on this foundation by addressing bias, leading to improved accuracy in empirical evaluations. Additionally, the Variance-Based Filtering algorithm effectively identifies corrupted vertices, albeit with a higher computational cost of $O(\varepsilon n^3)$.

We present empirical evaluations of the mean-squared errors and runtime comparisons of these algorithms, focusing on their performance against the (q, ε)-adversary. Our results confirm the theoretical upper bound for the Mean-Adjusted Median, showing $|p - \hat{p}| \leq O(1/\sqrt[4]{n})$, while highlighting the practical advantages of our proposed methods over baseline techniques. Furthermore, the empirical results demonstrate that these methods outperform existing techniques in practice, such as prune-then-mean/median and the spectral method by Acharya et al., in both error and runtime efficiency when applied to (q, ε)-oblivious adversaries.

Our findings provide valuable insights into designing robust algorithms for graph parameter estimation under adversarial perturbations. Future work could focus on extending the theoretical guarantees of the proposed algorithms to ε-omniscient adversaries. Appendix B includes proofs of lemmas that could serve as a foundation for establishing these guarantees. Another promising direction involves investigating applications of these methods in real-world networks.

References

[1] J. Acharya, A. Jain, G. Kamath, A. T. Suresh, and H. Zhang. Robust estimation for random graphs, 2022.

[2] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47–97, Jan. 2002.

[3] P. Erdős and A. Rényi. On random graphs. i. *Publicationes Mathematicae Debrecen*, 6(3–4):290–297, July 2022.

[4] E. N. Gilbert. Random graphs. *The Annals of Mathematical Statistics*, 30(4):1141–1144, Dec. 1959.

[5] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. Random graphs with arbitrary degree distributions and their applications. *Physical Review E*, 64(2), July 2001.

A. Proofs for Mean-Adjusted Median Algorithm

In this section, let D' be the degree distribution of the graph perturbed by the (q, ε)-oblivious adversary. Let $\bar{d}'$ and $\tilde{d}'$ be the mean and median of D', respectively.

Lemma A.1. $\mathbb{E}[\bar{d}'] = (n-1)p + \varepsilon(1-2n)\eta + \varepsilon^2 n\eta$ *where* $\eta := p - q$.

Proof. By definition, for $v \in R'$,

$$\mathbb{E}[\deg(v)] = \mathbb{E}\left[\sum_{u \in R', u \neq v} X_{uv} + \sum_{u \in S'} X_{uv}\right] = ((1-\varepsilon)n - 1)p + \varepsilon nq$$

and for $v \in S'$,

$$\mathbb{E}[\deg(v)] = (n-1)q.$$

Therefore,

$$\begin{aligned}
\mathbb{E}[\bar{d}'] &= \frac{1}{n}\left(\sum_{v \in R'} \mathbb{E}[\deg(v)] + \sum_{v \in S'} \mathbb{E}[\deg(v)]\right) \\
&= \frac{1}{n}\left(\sum_{v \in R'} (((1-\varepsilon)n - 1)p + \varepsilon nq) + \sum_{v \in S'} (n-1)q\right) \\
&= (n-1)p + \varepsilon(1-2n)\eta + \varepsilon^2 n\eta.
\end{aligned}$$

□

Letting $D' = \{D'_1, \ldots, D'_n\}$, let $D'_1, \ldots, D'_{(1-\varepsilon)n}$ be the degrees of the uncorrupted vertices and $D'_{(1-\varepsilon)n}, \ldots, D'_n$ be the degrees of the corrupted vertices.

Let $X = U + W$ where $U \sim \text{Binomial}(\varepsilon n, q)$ and $W \sim \text{Binomial}((1-\varepsilon)n - 1, p)$, and $Y \sim \text{Binomial}(n-1, q)$. Then the uncorrupted sample is drawn from the distribution of X and the corrupted sample is drawn from the distribution of Y. As n grows large, the dependence of pairwise sample asymptotically approaches 0, so we may assume that the samples are drawn iid from X and Y, respectively.

We now observe that for large n, X and Y can be approximated using the normal distribution due to the Central Limit Theorem. Hence,

$$X \sim N(\mu_X, \sigma_X^2), Y \sim N(\mu_Y, \sigma_Y^2)$$

where

$$\begin{aligned}
\mu_X &= ((1-\varepsilon)n - 1)p + \varepsilon nq; \\
\sigma_X^2 &= ((1-\varepsilon)n - 1)p(1-p) + \varepsilon nq(1-q); \\
\mu_Y &= (n-1)q; \\
\sigma_Y^2 &= (n-1)q(1-q).
\end{aligned}$$

For the remainder of our analysis, we proceed under this simplifying assumption.

Thus, we have a sample of size n drawn from the mixture distribution

$$Z := \begin{cases} X & \text{w.p. } (1-\varepsilon) \\ Y & \text{w.p. } \varepsilon. \end{cases}$$

Lemma A.2. $\text{Var}(\bar{d}') = \frac{(1-\varepsilon)\sigma_X^2 + \varepsilon\sigma_Y^2 + \varepsilon(1-\varepsilon)(\mu_X - \mu_Y)^2}{n}$

Proof. We have

$$\text{Var}(\bar{d}') = \frac{\text{Var}(Z)}{n} = \frac{(1-\varepsilon)\text{Var}(X) + \varepsilon\text{Var}(Y) + \varepsilon(1-\varepsilon)(\mu_X - \mu_Y)^2}{n}$$

as desired. □

Lemma A.3.

$$\mathbb{E}[\tilde{d'}] \approx \mu_X - \varepsilon \frac{\text{Sign}(\eta)}{2a}$$

where $a := (1-\varepsilon)/\sqrt{2\pi\sigma_X^2} + \varepsilon \frac{\varepsilon}{\sqrt{2\pi}\sigma_Y} e^{-\frac{(\mu_X-\mu_Y)^2}{2\sigma_Y^2}}$.

Proof. We know that the expected median $\mathbb{E}[\tilde{d'}]$ of D' satisfies:

$$(1-\varepsilon)F_X(\mathbb{E}[\tilde{d'}]) + \varepsilon F_Y(\mathbb{E}[\tilde{d'}]) = 0.5 \tag{1}$$

where F_X and F_Y are the CDFs of X and Y, respectively.

We now approximate the CDFs using a Taylor expansion. Since ε is small, we linearize around the median of X. Since X is normal, the median of X is μ_X. Letting $\Delta = \mathbb{E}[\tilde{d'}] - \mu_X$, then

$$F_X(\mathbb{E}[\tilde{d'}]) = 0.5 + \sum_{k=0}^{\infty} f_X^{(k)}(\mu_X)\Delta^k$$

and

$$F_Y(\mathbb{E}[\tilde{d'}]) = F_Y(\mu_X) + f_Y(\mu_X)\Delta + O(\Delta^2) \approx F_Y(\mu_X) + f_Y(\mu_X)\Delta$$

where f_X and f_Y are the PDFs of X and Y. Since X is normal, we know that the $f^{(k)}(\mu_X) = 0$ for $k > 0$. Hence,

$$F_X(\mathbb{E}[\tilde{d'}]) = 0.5 + \frac{1}{\sqrt{2\pi\sigma_X^2}}\Delta.$$

Substituting the expansion into Equation 1, we know

$$(1-\varepsilon)\left(0.5 + \frac{\Delta}{\sqrt{2\pi\sigma_X^2}}\right) + \varepsilon\left(F_Y(\mu_X) + f_Y(\mu_X)\Delta\right) = 0.5.$$

Simplifying and solving for Δ, we have

$$\Delta = -\varepsilon \frac{F_Y(\mu_X) - 0.5}{a}$$

where

$$a := (1-\varepsilon)/\sqrt{2\pi\sigma_X^2} + \varepsilon f_Y(\mu_X) = \frac{1-\varepsilon}{\sqrt{2\pi\sigma_X^2}} + \frac{\varepsilon}{\sqrt{2\pi}\sigma_Y} e^{-\frac{(\mu_X-\mu_Y)^2}{2\sigma_Y^2}},$$

implying that

$$\mathbb{E}[\tilde{d'}] \approx \mu_X - \varepsilon \frac{F_Y(\mu_X) - 0.5}{a}. \tag{2}$$

We now express $F_Y(\mu_X) - 0.5$ in terms of standard normal variables. Letting

$$z = \frac{\mu_X - \mu_Y}{\sigma_Y} = \frac{(n(1-\varepsilon)-1)\eta}{\sqrt{nq(1-q)}},$$

then $F_Y(\mu_X) = \Phi(z)$ where Φ is the standard normal CDF. Then

$$F_Y(\mu_X) - 0.5 = \Phi(z) - 0.5.$$

Substituting back into Equation 2, then $\mathbb{E}[\tilde{d'}]$ is

$$\mathbb{E}[\tilde{d'}] \approx \mu_X - \varepsilon \frac{\Phi(z) - 0.5}{a}.$$

We now observe that assuming $q \in (0,1)$, $z = \frac{(n(1-\varepsilon)-1)\eta}{\sqrt{nq(1-q)}} \approx \text{Sign}(\eta)C\sqrt{n}$ for some constant $C > 0$. For $z \geq 3$, we know $\Phi(z) \in [0.99, 1]$, and for $z \leq 3$, we know that $\Phi(z) \in [0, 0.01]$. Consequently, $\Phi(z) - 0.5 \approx \begin{cases} 0.5 & \text{if } p > q \\ -0.5 & \text{if } p < q \\ 0 & \text{if } p = q, \end{cases}$ with the approximation being tighter for increasing n. Then,

$$\Phi(z) - 0.5 \approx \frac{\text{Sign}(\eta)}{2}.$$

Therefore,

$$\mathbb{E}[\tilde{d}'] \approx \mu_X - \varepsilon\frac{\Phi(z) - 0.5}{a} \approx \mu_X - \varepsilon\frac{\text{Sign}(\eta)}{2a}$$

as desired. □

Lemma A.4.

$$\text{Var}(\tilde{d}') \approx \frac{1}{4na^2}$$

where $a = (1-\varepsilon)/\sqrt{2\pi\sigma_X^2} + \varepsilon f_Y(\mu_X)$.

Proof. Again, assuming that our samples are drawn i.i.d., for large n, the sample median of i.i.d. draws from a continuous distribution with PDF f_Z at its median $\mathbb{E}[\tilde{d}']$ (where $F_Z(\mathbb{E}[\tilde{d}']) = 0.5$) has an asymptotic normal distribution. Using the delta method on the central limit theorem, we know that

$$\text{Var}\left(\tilde{d}'\right) \approx \frac{1}{4nf_Z(\mathbb{E}[\tilde{d}'])^2} \approx \frac{1}{4nf_Z(\mu_X)^2} = \frac{1}{4na^2}.$$

□

We claim that $\hat{p} := \frac{\hat{d}}{n-1}$ where $\hat{d} := \frac{(2-\varepsilon)\tilde{d}' - \bar{d}'}{1-\varepsilon}$ is a good estimator of p. We first show that the expected value of $\hat{p}$ is close to p and that the variance is also small.

Lemma A.5. $\mathbb{E}[\hat{p}] \approx p - \frac{\varepsilon\eta + (2\varepsilon - \varepsilon^2)\frac{\text{Sign}(\eta)}{2a}}{(1-\varepsilon)(n-1)}$.

Proof. We first find $\mathbb{E}[\hat{d}]$. By Lemmas A.1 and A.3, we have

$$\begin{aligned}
\mathbb{E}[\hat{d}] &= \mathbb{E}\left[\frac{(2-\varepsilon)\tilde{d}' - \bar{d}'}{1-\varepsilon}\right] \\
&= \frac{(2-\varepsilon)\mathbb{E}[\tilde{d}'] - \mathbb{E}[\bar{d}']}{1-\varepsilon} \\
&\approx \frac{(2-\varepsilon)\left[\mu_X - \varepsilon\frac{\text{Sign}(\eta)}{2a}\right] - [(n-1)p + \varepsilon(1-2n)\eta + \varepsilon^2 n\eta]}{1-\varepsilon} \\
&= (n-1)p - \frac{\varepsilon\eta + (2\varepsilon - \varepsilon^2)\frac{\text{Sign}(\eta)}{2a}}{1-\varepsilon}.
\end{aligned}$$

Because $\hat{p} := \frac{\hat{d}}{n-1}$, then

$$\begin{aligned}
\hat{p} &\approx \frac{(n-1)p - \frac{\varepsilon\eta + (2\varepsilon - \varepsilon^2)\frac{\text{Sign}(\eta)}{2a}}{1-\varepsilon}}{n-1} \\
&= p - \frac{\varepsilon\eta + (2\varepsilon - \varepsilon^2)\frac{\text{Sign}(\eta)}{2a}}{(1-\varepsilon)(n-1)}
\end{aligned}$$

as desired. □

Lemma A.6. $\mathrm{Var}(\hat{p}) \leq 64\frac{1+\sqrt{n}+n}{n^2}$.

Proof. We first find bounds for $\sigma_X^2, \sigma_Y^2, (\mu_X - \mu_Y)^2$, and a. We know that $\varepsilon \in [0, 0.5)$ and $p, q \in [0, 1]$. Then $\varepsilon(1-\varepsilon), p(1-p), q(1-q) \leq 0.25$.

We then observe that

$$\begin{aligned}\sigma_X^2 &= ((1-\varepsilon)n - 1)p(1-p) + \varepsilon n q(1-q) \\ &\leq np(1-p) + \frac{nq(1-q)}{2} \\ &\leq \frac{n}{4} + \frac{n}{8} \\ &\leq \frac{n}{2}.\end{aligned}$$

We also have that

$$\sigma_Y^2 = (n-1)q(1-q) \leq \frac{n}{4}.$$

We then have that

$$\begin{aligned}(\mu_X - \mu_Y)^2 &= (((1-\varepsilon)n - 1)p + \varepsilon n q - (n-1)q)^2 \\ &\leq (np + nq/2)^2 \\ &\leq 4n^2.\end{aligned}$$

Finally, we observe that

$$\begin{aligned}a &= \frac{1-\varepsilon}{\sqrt{2\pi}\sigma_X} + \frac{\varepsilon}{\sqrt{2\pi}\sigma_Y} e^{-\frac{(\mu_X-\mu_Y)^2}{2\sigma_Y^2}} \\ &\geq \frac{0.5}{\sqrt{2\pi}\sigma_X} \\ [\sigma_X^2 \leq n/2] &\geq \frac{1}{2\sqrt{2\pi}\sqrt{n/2}} \\ &\geq \frac{1}{2\sqrt{\pi n}}.\end{aligned}$$

We now bound $\mathrm{Var}(\tilde{d}')$ and $\mathrm{Var}(\bar{d}')$. Using Lemma A.4 and $a \geq \frac{1}{2\sqrt{\pi n}}$, we observe that

$$\mathrm{Var}(\tilde{d}') = \frac{1}{4na^2} \leq \frac{1}{4n\left(\frac{1}{2\sqrt{\pi n}}\right)^2} = \pi.$$

Using Lemma A.2 and our previously found bounds, we observe that

$$\begin{aligned}\mathrm{Var}(\bar{d}') &= \frac{(1-\varepsilon)\sigma_X^2 + \varepsilon\sigma_Y^2 + \varepsilon(1-\varepsilon)(\mu_X - \mu_Y)^2}{n} \\ &\leq \frac{n/2 + n/8 + n^2}{n} \\ &\leq 1 + n \\ &\leq 2n.\end{aligned}$$

We know that by the Cauchy-Schwarz inequality, $|\mathrm{Cov}(X, Y)| \leq \sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}$ for any random variables X, Y.

Hence,

$$\begin{aligned}
\mathrm{Var}(\hat{d}) &= \mathrm{Var}\left(\frac{(2-\varepsilon)\tilde{d}' - \bar{d}'}{1-\varepsilon}\right) \\
&= \frac{1}{(1-\varepsilon)^2}\mathrm{Var}\left((2-\varepsilon)\tilde{d}' - \bar{d}'\right) \\
&= \frac{\mathrm{Var}((2-\varepsilon)\tilde{d}') + \mathrm{Var}(\bar{d}') + 2\mathrm{Cov}((2-\varepsilon)\tilde{d}', \bar{d}')}{(1-\varepsilon)^2} \\
&\leq \frac{\mathrm{Var}((2-\varepsilon)\tilde{d}') + \mathrm{Var}(\bar{d}') + 2\sqrt{\mathrm{Var}((2-\varepsilon)\tilde{d}')\mathrm{Var}(\bar{d}')}}{(1-\varepsilon)^2} \\
&\leq \frac{(2-\varepsilon)^2\pi + 2n + 2(2-\varepsilon)\sqrt{2\pi n}}{(1-\varepsilon)^2} \\
&\leq 16\pi + 8n + 16\sqrt{2\pi n} \\
&\leq 64(1+\sqrt{n}+n).
\end{aligned}$$

As a result, we know that

$$\mathrm{Var}(\hat{p}) = \mathrm{Var}\left(\frac{\hat{d}}{n-1}\right) = \frac{\mathrm{Var}(\hat{d})}{(n-1)^2} \leq 64\frac{1+\sqrt{n}+n}{n^2}$$

as desired. □

B. Proofs for Variance Algorithm

Lemma B.1. *Let F be the subgraph of uncorrupted vertices. Then $|p_F - p| \leq \frac{1}{\sqrt{n}}$ with probability at least $1 - 2\exp\left(-(1-\varepsilon)^2 n\right)$.*

Proof. Let $N_F := |F| = \binom{(1-\varepsilon)n}{2}$. Since each of the N_F edges of F is included independently with probability p, the number of edges in F, E_F, follows a binomial distribution:

$$E_F \sim \mathrm{Binomial}(N_F, p).$$

Then the empirical edge probability

$$p_F = \frac{E_F}{N_F}.$$

Because $\mu := \mathbb{E}[E_F] = N_F p$ and the binomial distribution is the sum of independent Bernoulli distributions, by the Hoeffding inequality,

$$\Pr(|p_F - p| \geq t) = \Pr(|E_F - \mu|/N_F \geq t) \leq 2\exp(-2N_F t^2).$$

Letting $t := \frac{1}{\sqrt{n}}$, then

$$\begin{aligned}
\Pr\left(|p_F - p| \geq \frac{1}{n}\right) &\leq 2\exp(-2N_F/n) \\
&= 2\exp\left(-\frac{2(1-\varepsilon)n((1-\varepsilon)n-1)}{2n}\right) \\
&= 2\exp(-(1-\varepsilon)((1-\varepsilon)n-1)) \\
&= 2\exp(-(1-\varepsilon)^2 n)\exp(1-\varepsilon) \\
&\leq 2\exp(-(1-\varepsilon)^2 n).
\end{aligned}$$

Therefore,

$$\Pr\left(|p_F - p| \leq \frac{1}{n}\right) \geq 1 - 2\exp(-(1-\varepsilon)^2 n)$$

as desired. □

Lemma B.2. *Let F be the subgraph of uncorrupted vertices. Then $|\hat{\sigma}_F^2 - \sigma_F^2| \leq O(\sqrt{n})$ with probability at least $1 - 2\exp(-(1-\varepsilon)^2 n)$.*

Proof. By Lemma B.1,

$$\begin{aligned}
|\hat{\sigma}_F^2 - \sigma_F^2| &= ((1-\varepsilon)n - 1)|p_F(1-p_F) - p(1-p)| \\
&\leq ((1-\varepsilon)n - 1)\left|\frac{\sqrt{n} - 2p\sqrt{n} - 1}{n}\right| \\
&\leq n\left|\frac{\sqrt{n} - 2p\sqrt{n} - 1}{n}\right| \\
&\leq \left|\sqrt{n} - 2p\sqrt{n} - 1\right| \\
&\leq O(\sqrt{n})
\end{aligned}$$

□