

---

# Signal Recovery from Random Dot-Product Graphs Under Local Differential Privacy

---

Siddharth Vishwanath  
University of California San Diego

Jonathan Hehir  
Penn State University

## Abstract

We consider the problem of recovering latent information from graphs under  $\varepsilon$ -edge local differential privacy where the presence of relationships/edges between two users/vertices remains confidential, even from the data curator. For the class of generalized random dot-product graphs, we show that a standard local differential privacy mechanism induces a specific geometric distortion in the latent positions. Leveraging this insight, we show that consistent recovery of the latent positions is achievable by appropriately adjusting the statistical inference procedure for the privatized graph. Furthermore, we prove that our procedure is nearly minimax-optimal under local edge differential privacy constraints. Lastly, we show that this framework allows for consistent recovery of geometric and topological information underlying the latent positions, as encoded in their persistence diagrams. Our results extend previous work from the private community detection literature to a substantially richer class of models and inferential tasks.

## 1 Introduction

Graphs are essential for modeling many systems and are used in a variety of applications, e.g., relationships between people in a social network, associations between individuals based on email records, financial transactions, or website browsing activity. In many cases, the existence, lack thereof, or the nature of such relationships may be sensitive. Without a formal privacy framework, malicious actors can exploit this sensitive information [7, 50, 51].

Differential privacy (DP) [19] has emerged as the standard for ensuring formal privacy, allowing for population-level inference, while limiting the risk of exposing the contribution of any single individual in a database. This disclosure risk is controlled by the *privacy budget*,  $\varepsilon > 0$ , where smaller values of  $\varepsilon$  provide stronger privacy guarantees. For graphs, DP comes in several forms, each with subtle differences between them [44]. These differences are, perhaps, best illuminated by answers to the following questions: (i) *What sensitive information needs protection?*, (ii) *How stringent should the privacy guarantees be?*, and (iii) *Who, if anyone, can be trusted with the data?*

The first question leads to two categories: *node DP* and *edge DP*. In node DP, the identities of the vertices and all their incident relationships are considered sensitive; whereas, in edge DP, the identities of the vertices are publicly available, and the sensitive information is in the presence/absence of relationships between vertices. The second question distinguishes *pure DP* from *approximate DP*. Pure DP, or  $\varepsilon$ -DP, strictly controls disclosure risk using  $\varepsilon$ ; whereas, approximate DP, or  $(\varepsilon, \delta)$ -DP, allows for a small failure probability  $\delta$ . The third question leads to *central DP* vs. *local DP*. In central DP, a *trusted curator* holds the full database and ensures DP in the data release. In local DP [16, 39, 61], each individual obfuscates their own data before sharing, removing the need for a trusted curator.

Deciding between node vs. edge DP is constrained by the nature of what is publicly available and what constitutes sensitive information, e.g., in social networks where individuals' identities are public. The choice between pure vs. approximate DP and central vs. local DP depends on considerations of privacy, utility, and practicality. We focus on  $\varepsilon$ -edge local differential privacy ( $\varepsilon$ -edge LDP), where the vertex identities are public, but the strongest form of privacy for the relationships is required.

Commonly referred to as the *privacy-utility tradeoff*, if  $\varepsilon$  is too large there is risk of leaking sensitive information, while if  $\varepsilon$  is too small the private information precludes meaningful inference. Addressing this requires

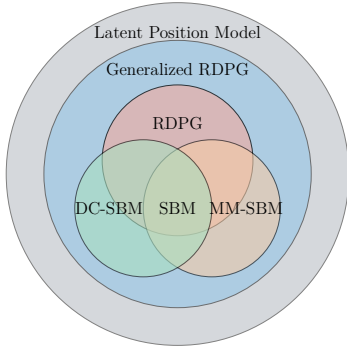


Figure 1: Hierarchy of network models.

(i) a judicious choice of  $\varepsilon$ , (ii) optimal algorithms for sanitizing sensitive information, and (iii) appropriate adjustments in the resulting statistical inference to account for privacy. It is well known that there is no free lunch in DP, and that “*any meaningful privacy-utility guarantees must come with reasonable assumptions on the data generating mechanism*” [40, Section 3]. In this work, we consider graphs generated from a flexible, nonparametric framework.

**Generalized random dot-product graphs.** Random dot-product graphs (RDPGs; [6]) and generalized random dot-product graphs (GRDPGs; [57]) provide a flexible framework for analyzing graphs. They cover a wide range of models, including the stochastic block model (SBM; [31]) and its extensions, such as the degree-corrected SBM (DC-SBM; [35]) and mixed-membership SBM (MM-SBM; [4]). GRDPGs belong to the family of *latent position models* [30], as illustrated in Figure 1. For example, the GRDPG representation of a SBM consists of a mixture of Dirac masses associated with the block memberships. Similarly, MM-SBMs and DC-SBMs admit GRDPGs which lie, respectively, on a simplex and on projective space. Recent work has shown that the adjacency embeddings of GRDPGs recovers meaningful information from its latent positions [46, 56–58].

**Contributions.** Given a graph  $\mathbf{A}$ , we consider a simple  $\varepsilon$ -edge DP mechanism called `edgeFlip` that outputs a differentially private synthetic graph,  $\mathcal{M}_\varepsilon(\mathbf{A})$ . `edgeFlip` has independently appeared in prior work under various aliases [e.g., 29, 33, 36, 49, 54]. We study the impact of `edgeFlip` on the quality of inference from GRDPGs. Our main contributions are as follows:

- We show that the class of GRDPGs are closed under `edgeFlip`, i.e., if  $\mathbf{A}$  is a GRDPG then  $\mathcal{M}_\varepsilon(\mathbf{A})$  is also a GRDPG (Theorem 3.1).
- Using this insight, if  $\mathbf{A}$  is a GRDPG with latent positions  $\mathbf{X} \in \mathbb{R}^{n \times d}$ , we derive minimax lower bounds for uniformly estimating the latent positions under  $\varepsilon$ -edge LDP in the  $\ell_{2,\infty}$ -norm. Our result highlights the interplay between the privacy budget  $\varepsilon$ , the sample size  $n$ , and the graph sparsity  $\rho_n$  (Theorem 3.2).

- For sufficiently dense graphs, by adjusting the estimation procedure to account for `edgeFlip` we show that a privacy-adjusted spectral embedding (Algorithm 1) provides consistent estimates of the latent positions, with near minimax-optimal accuracy up to an  $O(\sqrt{\log n})$  factor (Theorem 3.3). This same gap exists in the best-known bounds for the non-private setting [57, 63].

- Finally, we show how existing results in community detection under  $\varepsilon$ -edge DP can be extended to the broader problem of recovering topological information underlying the latent positions. To this end, we use tools from topological data analysis (TDA) and establish convergence rates for persistence diagram estimation under  $\varepsilon$ -edge LDP.

**Related Work.** There is an extensive body of work on differential privacy (DP) in graphs, but theoretical work relating to the privacy-utility tradeoff tends to focus on releasing graph statistics, such as degree sequences, triangle counts, and subgraph counts under edge DP [e.g., 27, 33, 34, 37, 38].

The closest related work to our setting comes from the emerging literature on private community detection [15, 28, 29, 48]. In [29], community detection for `edgeFlip` under  $\varepsilon$ -edge LDP is studied for SBMs and DC-SBMs without sparsity patterns, with utility measured via misclassification rates in approximate clustering. In a similar setting, [11] consider estimating the membership probabilities for DC-MM-SBMs under  $\varepsilon$ -edge LDP. In the central  $\varepsilon$ -edge DP setting, [48] consider community detection for SBMs (where the average degree grows as  $\log n$ ), focusing on *exact recovery* using semi-definite programming. In [15], the analysis is extended to SBMs whose expected degree is flexible, providing bounds for, both, *exact* and *weak recovery*. This line of analysis has also been extended to more general models which relax the SBM assumptions [28, 52].

Although these works apply similar spectral methods and privacy mechanisms, there are several key distinctions. First, their settings are more focused and restrictive. SBMs and DC-SBMs are specific cases of the more general GRDPGs that we study. Second, they assess utility based on cluster recovery accuracy, whereas we evaluate utility by the  $\ell_{2,\infty}$  error in recovering latent positions. The broader framework allows for more complex structures beyond simple assortative clusters, such as hierarchical or topological patterns. Finally, all of our results are in the local DP setting, where there is no centralized trusted curator releasing queries. The price to be paid for this generality is that our upper bounds apply only to a denser regime where the average degree grows at rate  $\omega(\log^4 n)$ , as opposed

to  $\omega(\log n)$  which corresponds to the information-theoretic threshold. The former is the regime in which the best known recovery bounds for GRDPGs operate.

## 2 Background

**Notation.** For  $\mathbf{u} \in \mathbb{R}^d$ ,  $\|\mathbf{u}\|$  is the Euclidean norm, and for  $\mathbf{B} \in \mathbb{R}^{n \times m}$ ,  $\|\mathbf{B}\|_{\text{op}}$  and  $\|\mathbf{B}\|_{2,\infty}$  denote the  $\ell_2$ - and  $\ell_{2,\infty}$ -operator norm.  $\mathbf{X} = [X_1 | \dots | X_n]^\top \in \mathbb{R}^{n \times d}$  denotes the  $n \times d$  matrix whose row vectors are observations  $\mathbf{X}_n = \{X_1, \dots, X_n\} \subset \mathbb{R}^d$ . For fixed known integers  $d, p, q$  with  $p + q = d$ ,  $\mathbb{I}_{p,q} = \text{diag}(\mathbf{1}_p, -\mathbf{1}_q)$  is the indefinite identity matrix, and

$$\mathcal{O}(p, q) = \{\mathbf{Q} \in \mathbb{R}^{d \times d} : \mathbf{Q}^\top \mathbb{I}_{p,q} \mathbf{Q} = \mathbf{Q} \mathbb{I}_{p,q} \mathbf{Q}^\top = \mathbb{I}_{p,q}\}$$

is the indefinite orthogonal group with signature  $(p, q)$ . When  $q = 0$ ,  $\mathcal{O}(d)$  is the usual orthogonal group.  $\mathbf{U}(n, d) = \{\mathbf{U} \in \mathbb{R}^{n \times d} : \mathbf{U}^\top \mathbf{U} = \mathbb{I}_d\}$  are  $n \times d$  matrices with orthonormal columns, and  $\mathbb{B}(n) \subset \{0, 1\}^{n \times n}$  is the set of binary symmetric  $n \times n$  matrices.

For a probability distribution  $\mathbb{P}$  on  $\mathbb{R}^d$  and measurable  $f_\# \mathbb{P}$  is the pushforward measure satisfying  $f_\# \mathbb{P}(A) = \mathbb{P}(f^{-1}(A))$  for  $A \subseteq \mathbb{R}^k$ . We use standard asymptotic notation;  $a_n = O(b_n)$  or  $b_n = \Omega(a_n)$  if  $\limsup_n |a_n/b_n| \leq C$ ;  $a_n = o(b_n)$  or  $b_n = \omega(a_n)$  if  $\limsup_n |a_n/b_n| = 0$ .

### 2.1 Differential Privacy for Graphs

In the  $\varepsilon$ -edge DP framework, given a graph  $\mathbf{G} = (V, E)$  with adjacency matrix  $\mathbf{A} \in \mathbb{B}(n)$ , the sensitive records correspond the edges of the graph,  $E$ . The privacy mechanism,  $\mathcal{A}_\varepsilon$ , takes  $\mathbf{A}$  as input and produces a random output  $\mathcal{A}_\varepsilon(\mathbf{A})$  taking values in  $\mathcal{Z}$  such that the presence or absence of any single edge is not revealed.

**Definition 2.1 ( $\varepsilon$ -edge DP).** For  $\varepsilon > 0$ ,  $\mathcal{A}_\varepsilon : \mathbb{B}(n) \rightarrow \mathcal{Z}$  satisfies  $\varepsilon$ -edge differential privacy if, for all  $\mathbf{A}', \mathbf{A}'' \in \mathbb{B}(n)$  differing on a single edge, i.e.,  $\text{Ham}(\mathbf{A}', \mathbf{A}'') = 2$ , and for all  $S \subseteq \mathcal{Z}$ ,

$$\mathbb{P}(\mathcal{A}_\varepsilon(\mathbf{A}') \in S) \leq e^\varepsilon \cdot \mathbb{P}(\mathcal{A}_\varepsilon(\mathbf{A}'') \in S). \quad (1)$$

Equivalently, from [17],  $\mathcal{A}_\varepsilon$  defines a conditional probability distribution  $\mathcal{Q}_\varepsilon$  such that for  $\mathbf{Z} = \mathcal{A}_\varepsilon(\mathbf{A})$ ,

$$\mathcal{Q}_\varepsilon(\mathbf{Z} \in S | \mathbf{A} = \mathbf{A}') \leq e^\varepsilon \cdot \mathcal{Q}_\varepsilon(\mathbf{Z} \in S | \mathbf{A} = \mathbf{A}'').$$

In the local DP setting, the privacy mechanism is applied to each edge independently; see, e.g., [29, 43]

**Definition 2.2 ( $\varepsilon$ -edge LDP).** The privacy mechanism  $\mathcal{A}_\varepsilon = \{\mathcal{A}_{ij}^\varepsilon : 1 \leq i < j \leq n\}$  is  $\varepsilon$ -edge locally DP if, for each  $i < j \in [n]$  and  $\mathbf{Z}_{ij} = \mathcal{A}_\varepsilon(\mathbf{A}_{ij}) \in \mathcal{Z}$ , the resulting conditional distributions  $\{\mathcal{Q}_{ij}^\varepsilon : 1 \leq i < j \leq n\}$  are such that: for all  $S \subset \mathcal{Z}$  and  $x, x' \in \{0, 1\}$ ,

$$\mathcal{Q}_{ij}(\mathbf{Z}_{ij} \in S | \mathbf{A}_{ij} = x) \leq e^\varepsilon \cdot \mathcal{Q}_{ij}(\mathbf{Z}_{ij} \in S | \mathbf{A}_{ij} = x').$$

The privacy mechanism we focus on in this work is the symmetric edgeFlip [see, e.g., 22, 29, 33, 36, 48, 49,

54] produces an  $\varepsilon$ -edge LDP synthetic graph  $\mathcal{M}_\varepsilon(\mathbf{A})$  for each  $\mathbf{A} \in \mathbb{B}(n)$ , i.e., for edgeFlip,  $\mathcal{Z} = \mathbb{B}(n)$ . This is closely related to Warner's randomized response [61]. The following definition of edgeFlip is due to [33] and adapted from [29, Definition 3.5].

**Definition 2.3 (edgeFlip).** Given  $\varepsilon > 0$ , let  $\pi(\varepsilon) \doteq 1/(e^\varepsilon + 1)$  and  $\mathcal{F} : \{0, 1\} \rightarrow \{0, 1\}$  be given by

$$\mathcal{F}(x) = \begin{cases} x & \text{w.p. } 1 - \pi(\varepsilon) \\ 1 - x & \text{w.p. } \pi(\varepsilon) \end{cases}$$

For each  $i \in [n]$ , let  $\mathbf{A}_{i,*}$  be the edges connected to vertex  $v_i$ , and define the randomized response mechanism  $\mathcal{M}_{i,\varepsilon}(\mathbf{A}_{i,*}) = [\mathcal{F}(\mathbf{A}_{i,i+1}) | \mathcal{F}(\mathbf{A}_{i,i+2}) | \dots | \mathcal{F}(\mathbf{A}_{i,n})]$ , and the full randomized response for all vertices,

$$T_\varepsilon(\mathbf{A}) = [\mathcal{M}_{1,\varepsilon}(\mathbf{A}_{1,*}) | \dots | \mathcal{M}_{n,\varepsilon}(\mathbf{A}_{n,*})]^\top.$$

Then edgeFlip,  $\mathcal{M}_\varepsilon : \mathbb{B}(n) \rightarrow \mathbb{B}(n)$ , is given by

$$\mathcal{M}_\varepsilon(\mathbf{A}) = T_\varepsilon(\mathbf{A}) + T_\varepsilon(\mathbf{A})^\top.$$

By the closure property of DP mechanisms under post-processing, since edgeFlip satisfies  $\varepsilon$ -edge DP any analysis performed on  $\mathcal{M}_\varepsilon(\mathbf{A})$  after edgeFlip will also satisfy the same privacy guarantee [19]. Given a privacy budget  $\varepsilon > 0$ , and the edge flip probability in Definition 2.3, we define

$$\begin{aligned} \sigma(\varepsilon) &\doteq \sqrt{1 - 2\pi(\varepsilon)} = \sqrt{e^\varepsilon - 1/e^\varepsilon + 1}, \\ \tau(\varepsilon) &\doteq \sqrt{\pi(\varepsilon)} = \sqrt{1/e^\varepsilon + 1}. \end{aligned} \quad (2)$$

### 2.2 Spectral Embedding and GRDPGs

Statistical inference on graphs often begins by embedding the vertices in Euclidean space. Of numerous ways to achieve this [e.g., 9, 24, 53], we consider the adjacency spectral embedding of  $\mathbf{A}$  [6, 57].

**Definition 2.4.** For  $\mathbf{A} \in \mathbb{B}(n)$  and fixed  $d < n$ , let

$$\mathbf{A} = \mathbf{U} \mathbf{\Lambda}(\mathbf{A}) \mathbf{U}^\top + \mathbf{V} \mathbf{\Omega}(\mathbf{A}) \mathbf{V}^\top, \quad (3)$$

be the spectral decomposition of  $\mathbf{A}$ , where  $|\mathbf{\Lambda}| \equiv |\mathbf{\Lambda}(\mathbf{A})|$  are the top- $d$  eigenvalues by magnitude and  $\mathbf{U} \in \mathbf{U}(n, d)$  are the corresponding eigenvectors. Then the adjacency spectral embedding of  $\mathbf{A}$  is given by

$$\hat{\mathbf{X}} = \text{sp}(\mathbf{A}; d) = \mathbf{U} |\mathbf{\Lambda}(\mathbf{A})|^{1/2} \in \mathbb{R}^{n \times d}. \quad (4)$$

The spectral embedding is unique only up to orthogonal transformations; indeed, for any  $\mathbf{Q} \in \mathcal{O}(d)$ ,  $\mathbf{U}' = \mathbf{U} \mathbf{Q} \in \mathbf{U}(n, d)$  and leaves (3) invariant. Moreover, if  $\mathbf{\Lambda}$  contains  $p$  positive and  $q$  negative eigenvalues, then  $\mathbf{\Lambda} = |\mathbf{\Lambda}|^{1/2} \mathbb{I}_{p,q} |\mathbf{\Lambda}|^{1/2}$ . GRDPGs can be seen as probabilistic graphical models which explicitly account for indefinite  $\mathbf{\Lambda}$  in the spectral decomposition.

To this end, a probability distribution  $\mathbb{P}$  on  $\mathbb{R}^d$  is said to be  $(p, q)$ -admissible if  $\text{supp}(\mathbb{P})$  is compact

and  $0 \leq \mathbf{x}^\top \mathbb{I}_{p,q} \mathbf{y} \leq 1$ , for all  $\mathbf{x}, \mathbf{y} \in \text{supp}(\mathbb{P})$ . In other words, samples from  $\mathbb{P}$  produce valid probabilities w.r.t. the indefinite inner product.

**Definition 2.5 (GRDPG).** Let  $\mathbb{P}$  be  $(p, q)$ -admissible and let  $\rho_n \leq 1$ . The random graph  $\mathbf{A} \in \mathcal{B}(n)$  is a generalized random dot-product graph with signature  $(p, q)$ , sparsity  $\rho_n$ , and latent positions  $\mathbf{X} \in \mathbb{R}^{n \times d}$  if

- (i) The rows of  $\mathbf{X}$  are observed i.i.d. from  $\mathbb{P}$ , and
- (ii)  $\mathbf{A}_{ij} | \mathbf{X} \sim_{\text{i.i.d.}} \text{Ber}(\rho_n \cdot \mathbf{X}_i^\top \mathbb{I}_{p,q} \mathbf{X}_j) \quad \forall 1 \leq i < j \leq n$ .

More succinctly,  $(\mathbf{A}, \mathbf{X}) \sim \mathcal{G}(\mathbb{P}, \rho_n; p, q)$ . Conditional on  $\mathbf{X} \in \mathbb{R}^{n \times d}$ , we write  $\mathbf{A} \sim \mathcal{G}(\mathbf{X}, \rho_n; p, q)$ .

Some examples of GRDPGs include the following.

(Ex<sub>1</sub>) When  $\mathbb{P} = \delta_{\mathbf{x}}$  is a Dirac mass at point  $\mathbf{x} \in \mathbb{R}^d$  with  $\pi \doteq \|\mathbf{x}\| \in (0, 1)$ , then  $\mathcal{G}(\mathbb{P}, \rho_n; d, 0)$  is an Erdős-Rényi graph with probability  $\pi \rho_n$ . When  $\|\mathbf{x}\| = 1$  and  $\rho_n = \log n/n$ , the resulting graph corresponds to the sharp connectivity threshold for Erdős-Rényi graphs.

(Ex<sub>2</sub>) Let  $\mathbb{P} = \frac{1}{2}(\delta_{\mathbf{x}_1} + \delta_{\mathbf{x}_2})$  where, for  $\gamma > 0$ ,  $\|\mathbf{x}_1\| = \|\mathbf{x}_2\| = \sqrt{1 + \gamma}$ ,  $\theta = \angle(\mathbf{x}_1, \mathbf{x}_2) = \arccos\left(\frac{1-\gamma}{1+\gamma}\right)$ .

Then  $(\mathbf{A}, \mathbf{X}) \sim \mathcal{G}(\mathbb{P}, \rho_n; 2, 0)$  is a stochastic block model  $\text{SBM}(n; \gamma, \rho_n)$ . The intra- and inter-community edge probabilities are  $\rho_n(1 + \gamma)$  and  $\rho_n(1 - \gamma)$ . On the other hand, when  $(p, q) = (1, 1)$ , the SBM is disassortative.

(Ex<sub>3</sub>) If, instead,  $\mathbf{x}_1 = \beta_1 \mathbf{x}$  and  $\mathbf{x}_2 = \beta_2 \mathbf{x}$  for a fixed  $\mathbf{x} \in \mathbb{R}^d$  and  $\beta_1, \beta_2 > 0$ , then  $\mathcal{G}(\mathbb{P}, \mathbf{X}, \rho_n; 2, 0)$  is a DC-SBM. On the other hand, if  $\mathbb{P} = \alpha \delta_{\mathbf{x}_1} + (1 - \alpha) \delta_{\mathbf{x}_2}$  for  $\alpha \in (0, 1)$ , then the SBM has unbalanced communities.

**Remark 2.1.** GRDPGs are invariant to  $\mathcal{O}(p, q)$  transformations and equivariant w.r.t. scaling transformations. To see this, let  $\mathbf{Q} \in \mathcal{O}(p, q)$  and  $s > 0$  be a scale parameter. For  $f(\mathbf{x}) = \mathbf{Q}(s \cdot \mathbf{x})$ , it follows that

$$\rho_n f(\mathbf{X}) \mathbb{I}_{p,q} f(\mathbf{X})^\top = s^2 \rho_n \mathbf{X} \mathbb{I}_{p,q} \mathbf{X}^\top, \quad (5)$$

and,  $\mathcal{G}(\mathbb{P}, \rho_n) \stackrel{d}{=} \mathcal{G}(f_\# \mathbb{P}, \rho_n/s^2)$ .

In other words, the latent positions of  $\mathcal{G}(\mathbb{P}, \rho_n; p, q)$  are identifiable only up to an  $\mathcal{O}(p, q)$  transformation, and  $\rho_n$  is identifiable only up to scaling by a constant factor. In order to address the first invariance, we consider metrics on  $\mathbb{R}^{n \times d}$  which are  $\mathcal{O}(p, q)$  invariant (see Definition 3.1). For the second source of non-identifiability, we fix the scale of  $\mathbb{P}$  to satisfy

$$\mathbb{E}(\xi_1^\top \mathbb{I}_{p,q} \xi_2) = 1 \quad \text{for } \xi_1, \xi_2 \sim \mathbb{P}. \quad (6)$$

See, e.g., [2, Corollary 2] and [45, line after Eq. (2)], where the same condition is used to fix the scale of  $\mathbb{P}$ .

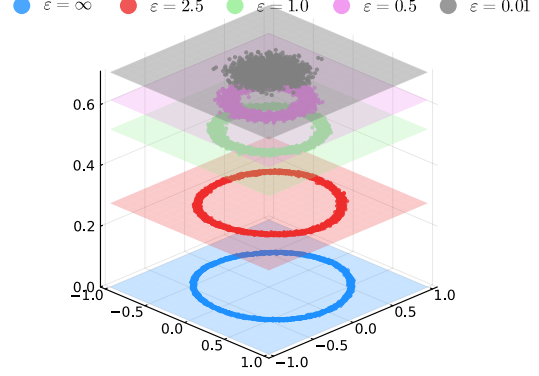


Figure 2: Illustration of Theorem 3.1. Spectral embedding of  $\mathcal{M}_\varepsilon(\mathbf{A})$  after edgeFlip when  $\mathbf{X} \sim \mathbb{P} = \text{Unif}(\mathbb{S}^1)$  and  $\rho_n \equiv 1$ .

### 3 Main Results

This section presents the main results, and the proofs for the main results are collected in Appendix C.

#### 3.1 Closure of GRDPGs under edgeFlip

The key insight in the remaining sections comes from the following result, which states that the class of GRDPGs is closed under edgeFlip.

**Theorem 3.1.** Suppose  $(\mathbf{A}, \mathbf{X}) \sim \mathcal{G}(\mathbb{P}, \rho_n; p, q)$ , and for  $\varepsilon \geq 0$ , let  $\mathcal{M}_\varepsilon(\mathbf{A})$  denote the  $\varepsilon$ -edge LDP graph under edgeFlip. Let  $\varphi_\varepsilon : \mathbb{R}^d \rightarrow \mathbb{R}^{d+1}$  be the map given by  $\varphi_\varepsilon(\mathbf{x}) = \tau(\varepsilon) \oplus \sigma(\varepsilon) \rho_n^{1/2} \mathbf{x}$ , i.e.,

$$\varphi_\varepsilon(\mathbf{x}) = \begin{pmatrix} \tau(\varepsilon) \\ \sigma(\varepsilon) \rho_n^{1/2} \mathbf{x} \end{pmatrix} \in \mathbb{R}^{d+1}.$$

Then,  $(\mathcal{M}_\varepsilon(\mathbf{A}), \varphi_\varepsilon(\mathbf{X})) \sim \mathcal{G}(\varphi_{\varepsilon\#} \mathbb{P}, 1; p+1, q)$ .

Put simply, when edgeFlip is applied to a GRDPG with signature  $(p, q)$ , the resulting random graph is also a GRDPG with signature  $(p+1, q)$ , and its latent positions are embedded into  $\mathbb{R}^{d+1}$  via  $\varphi_\varepsilon$ . See Figure 2 for an illustration.  $\varphi_\varepsilon$  scales the latent positions by a factor  $\sigma(\varepsilon)$ ; as  $\varepsilon$  decreases,  $\sigma(\varepsilon)$  decreases. This results in the squeezing effect in Figure 2. Notably,  $\mathcal{M}_\varepsilon(\mathbf{A})$  loses the sparsity pattern of  $\mathbf{A}$  whenever  $\tau(\varepsilon) = \omega(\rho_n^{1/2})$ , i.e.,  $\varepsilon = o(\log 1/\rho_n)$ . Nevertheless,  $\mathcal{M}_\varepsilon(\mathbf{A})$  still lends itself to important inference tasks, as we will see in the coming sections.

The addition of the extra dimension in Theorem 3.1 be viewed as the extra source of randomness introduced via edgeFlip. To see this, let  $p_{ij} = \mathbf{X}_i^\top \mathbb{I}_{p,q} \mathbf{X}_j$ , and recall from (1) that  $\mathbf{A}' = \mathcal{M}_\varepsilon(\mathbf{A})$  is such that

$$\mathbb{P}(\mathbf{A}'_{ij} = 1) = \pi(\varepsilon)(1 - p_{ij}) + (1 - \pi(\varepsilon))p_{ij},$$

i.e., the probability of an edge  $\mathbf{A}'_{ij}$  is a convex mixture of the original connection probabilities of  $\mathbf{A}$ . [57, Appendix B] establishes that GRDPGs are, essentially,

the only random graph model which reproduce mixtures of connection probabilities as convex combinations in the latent space. Under `edgeFlip`, this manifests as a convex combination of the original latent positions  $\mathbf{X}$  (when  $\varepsilon = \infty$ ) and a single point in the extra  $(d+1)$ th dimension (when  $\varepsilon = 0$ ), which corresponds to the latent position of an Erdős-Rényi graph.

**Corollary 3.1.** *When  $\varepsilon = 0$  in Theorem 3.1,  $\varphi_\varepsilon(\mathbf{x}) \equiv (\sqrt{1/2}, \mathbf{0}_d^\top)^\top$  and  $\mathcal{M}_\varepsilon(\mathbf{A}) \sim \text{Erdős-Rényi}(1/2)$ .*

By the preceding discussion, it would seem as though applying  $\mathcal{M}_\varepsilon$  successively with  $\varepsilon_1, \varepsilon_2$  would result in a GRDPG with signature  $(p+2, q)$ —one extra dimension for each application of  $\mathcal{M}_\varepsilon$ . However, probabilistically speaking, applying `edgeFlip` with probabilities  $\pi(\varepsilon_1), \pi(\varepsilon_2)$  is equivalent to performing a single `edgeFlip` with adjusted probability

$$\pi = \pi(\varepsilon_1)(1 - \pi(\varepsilon_2)) + (1 - \pi(\varepsilon_1))\pi(\varepsilon_2).$$

The next result addresses this by showing that the geometric interpretation is consistent with the probabilistic intuition.

**Proposition 3.1** (paraphrased; see Appendix C.2). *Let  $\mathbf{A}' = \mathcal{M}_{\varepsilon_2} \circ \mathcal{M}_{\varepsilon_1}(\mathbf{A})$  and  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}^{d+2}$  where*

$$\varphi(\mathbf{x}) = (\tau_2, \sigma_2\tau_1, \sigma_2\sigma_1\mathbf{x}^\top)^\top \in \mathbb{R}^{d+2},$$

*where  $\tau_i = \tau(\varepsilon_i)$  and  $\sigma_i = \sigma(\varepsilon_i)$  for  $i = 1, 2$ . Then, there exists  $\mathbf{Q} \in \mathcal{O}(p+2, q)$  and  $a, b > 0$  such that*

$$\mathbf{Q}\varphi(\mathbf{x}) = (0, a, b\mathbf{x}^\top)^\top \quad \forall \mathbf{x} \in \mathbb{R}^d.$$

By invariance under  $\mathcal{O}(p+2, q)$  transformations from (5), Proposition 3.1 shows there is a minimal  $\mathbb{R}^{d+1}$  representation of the latent positions that encodes the GRDPG structure after multiple applications of `edgeFlip` which faithfully only has one extra dimension.

### 3.2 Minimax rates for GRDPGs

Before considering the general problem of estimation from GRDPGs, it is instructive to first consider how Theorem 3.1 simplifies analysis by considering the case when  $\mathbf{A} \sim \text{SBM}(n; \gamma, \rho_n)$ . Recall from (E $\mathbf{x}_2$ ) that the latent positions are  $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^d$  with  $\|\mathbf{x}_1\| = \|\mathbf{x}_2\| = \sqrt{1+\gamma}$  and  $\theta = \arccos(1 - \gamma/(1+\gamma))$ . From Theorem 3.1,  $\mathcal{M}_\varepsilon(\mathbf{A})$  is also an SBM with latent positions  $\mathbf{y}_i = \varphi_\varepsilon(\mathbf{x}_i)$ ; see Figure 3. The intra- and inter-community edge probabilities become

$$\begin{aligned} a_\varepsilon &= \mathbf{y}_i^\top \mathbb{I}_{d+1} \mathbf{y}_i = \tau(\varepsilon)^2 + \sigma(\varepsilon)^2 \rho_n(1 + \gamma), \\ b_\varepsilon &= \mathbf{y}_1^\top \mathbb{I}_{d+1} \mathbf{y}_2 = \tau(\varepsilon)^2 + \sigma(\varepsilon)^2 \rho_n(1 - \gamma). \end{aligned}$$

For small  $\varepsilon \leq 1$ , an application of [65, Theorem 2.1] shows that exact recovery is possible only when

$$\frac{(a_\varepsilon - b_\varepsilon)^2}{a_\varepsilon} > \frac{2 \log n}{n} \implies \gamma \rho_n^2 \sigma(\varepsilon)^4 = \Omega\left(\frac{\log n}{n}\right). \quad (7)$$

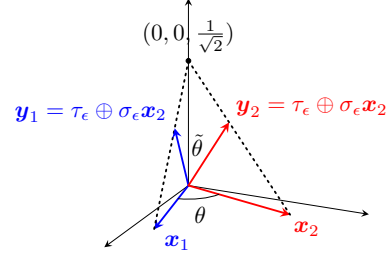


Figure 3: Illustration of `edgeFlip` for  $\text{SBM}(n; \gamma, \rho_n)$

In the absence of privacy, however, the detection threshold for exact recovery requires

$$\gamma \rho_n = \Omega\left(\frac{\log n}{n}\right).$$

Compared to (7), the nature of `edgeFlip` leads to an amplification of the effect due to sparsity  $\rho_n$ , and coincides with the observations in [29, p. 13] for *weak recovery*. A similar analysis can be carried out for  $\varepsilon$ -edge LDP exact recovery from SBMs with  $k$ -communities and for DC-SBMs using the results from [23].

We now turn our attention to the problem of estimating the latent positions of GRDPGs under  $\varepsilon$ -edge LDP. Since the latent positions are identifiable only up to  $\mathcal{O}(p, q)$  transformations by (5); any metric for assessing the uniform error in the latent positions of GRDPGs must take into account this non-identifiability. To this end, [63] define the metric  $d_{2,\infty}$  as follows.

**Definition 3.1** ( $d_{2,\infty}$  metric). *For  $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{n \times d}$ ,*

$$d_{2,\infty}(\mathbf{X}, \mathbf{Y}) = \min_{\mathbf{O} \in \mathcal{O}(d) \cap \mathcal{O}(p,q)} \|\mathbf{Y}\mathbf{Q}_\mathbf{Y}^{-1} - \mathbf{X}\mathbf{Q}_\mathbf{X}^{-1}\mathbf{O}\|_{2,\infty},$$

*where  $\mathbf{Q}_\mathbf{X}, \mathbf{Q}_\mathbf{Y} \in \mathcal{O}(p, q)$  are the spectral alignments such that  $\mathbf{X}\mathbf{Q}_\mathbf{X}^{-1} = \mathbf{U}_\mathbf{X}|\Lambda(\mathbf{X})|^{1/2}$ , and similarly for  $\mathbf{Y}$ .*

For fixed  $p, q, d$ , let  $\mathcal{X} \subset \mathbb{R}^d$  be the set of all regular  $(p, q)$ -admissible latent positions:

$$\mathcal{X} \doteq \left\{ \mathbf{X} \in \mathbb{R}^{n \times d} : \mathbf{X}\mathbb{I}_{p,q}\mathbf{X}^\top \in [0, 1]^{n \times n} \right\}.$$

For  $\mathbf{A} \sim \mathcal{G}(\mathbf{X}, \rho_n; p, q)$ , let  $\hat{\mathbf{X}} : \mathbb{B}(n) \rightarrow \mathbb{R}^{n \times d}$  be an estimator of the latent positions  $\mathbf{X} \in \mathcal{X}$ . The minimax risk for estimating the latent positions  $\mathbf{X}$  is given by

$$\mathfrak{R}_n(\mathcal{X}) = \inf_{\hat{\mathbf{X}}} \sup_{\mathbf{X} \in \mathcal{X}} \mathbb{E}_{\mathbf{A} \sim \mathcal{G}(\mathbf{X}, \rho_n; p, q)} \left[ d_{2,\infty}(\hat{\mathbf{X}}(\mathbf{A}), \mathbf{X}) \right],$$

where the infimum is taken over all estimators  $\hat{\mathbf{X}}$ . In the absence of privacy and when  $n\rho_n = \omega(\log n)$ , the minimax rate from [63, Theorem 1 & Corollary 1] is

$$\mathfrak{R}_n(\mathcal{X}) = \Omega\left(\sqrt{\frac{\log n}{n\rho_n}}\right). \quad (8)$$

We point out that the rate in (8) implicitly accounts for the  $\rho_n^{-1/2}$  rescaling needed to recover the original latent positions  $\mathbf{X}$  from  $\mathbf{A} \sim \mathcal{G}(\mathbf{X}, \rho_n; p, q)$ .

In the  $\varepsilon$ -edge LDP setting, let  $\mathbb{A}_\varepsilon$  denote the set of all  $\varepsilon$ -edge LDP mechanisms satisfying Definition 2.2. Under the additional constraint of privacy, we restrict our attention to estimators  $\hat{X} : \mathcal{Z} \rightarrow \mathbb{R}^{n \times d}$  which use the private output  $\mathbf{Z} = \mathcal{A}_\varepsilon(\mathbf{A}) \in \mathcal{Z}$ ; in this setting the minimax risk is given by

$$\mathfrak{R}_n(\mathcal{X}, \varepsilon) = \inf_{\mathcal{A}_\varepsilon \in \mathbb{A}_\varepsilon} \inf_{\hat{X}} \sup_{\mathbf{X} \in \mathcal{X}} \mathbb{E} \left[ d_{2,\infty}(\hat{X}(\mathcal{A}_\varepsilon(\mathbf{A})), \mathbf{X}) \right], \quad (9)$$

where the expectation is taken over the randomness in  $\mathbf{A}$  and in  $\mathcal{A}_\varepsilon \in \mathbb{A}_\varepsilon$ . The following result provides a lower bound on the minimax risk for estimating the latent positions of GRDPGs under  $\varepsilon$ -edge LDP.

**Theorem 3.2.** *Let  $\mathcal{X}$  be as defined above, and let  $n\rho_n = \omega(\log n)$ . Then, for  $0 < \varepsilon < 3/8\rho_n$  and for sufficiently large  $n$ , the minimax risk satisfies*

$$\mathfrak{R}_n(\mathcal{X}, \varepsilon) = \Omega \left( \sqrt{\frac{\log n}{n\sigma(\varepsilon)^4 \rho_n^2}} \right). \quad (10)$$

In other words, Theorem 3.2 characterizes the worst-case risk any estimator of the latent positions of a GRDPG can achieve under *all possible*  $\varepsilon$ -edge LDP mechanisms satisfying Definition 2.2. In the *high-privacy regime* where  $\varepsilon \rightarrow 0$  as  $n \rightarrow \infty$ , the privacy-utility tradeoff under  $\varepsilon$ -edge LDP differs from the non-private setting in (8) in two key ways:

(i) The effective sample size reduces from  $n$  to  $\sigma^4(\varepsilon)n \ll n$ . This phenomenon is well-known in the literature on  $\varepsilon$ -LDP estimation [see, e.g., 5, 8, 17, 55].

(ii) Moreover, in contrast to (8), there is an amplification of the effect due to sparsity in (10), i.e.,  $\rho \mapsto \rho_n^2$ . In [43], a similar phenomenon was observed for change-point localization in dynamic networks.

Note that in order for the lower bound in (10) to not be vacuous, we require that  $\sigma(\varepsilon)^4 \rho_n^2 = \Omega(\log n/n)$ . By further noting that  $\sigma(\varepsilon)^2 = \tanh(\varepsilon/2)$ , this is equivalent to

$$\varepsilon = \Omega \left( \tanh^{-1} \left( \sqrt{\frac{\log n}{n\rho_n^2}} \right) \right), \quad (11)$$

In other words, if  $\varepsilon \rightarrow 0$ , consistent recovery of the latent positions is only possible for sufficiently dense graphs. For instance, while graphs with sparsity  $\rho_n \asymp (\log n/n)^{2/3}$  admit consistent estimation in the non-private setting ((8)), in the minimax sense no consistent estimators exist at the same sparsity under  $\varepsilon$ -edge LDP ((9)). We also note that (11) determines the detection threshold for exact recovery when  $\mathcal{G}(\mathbb{P}, \rho_n; 2, 0) \stackrel{d}{=} \text{SBM}(\gamma, \rho_n)$  as noted in (7).

### 3.3 Privacy-Adjusted Spectral Embedding

Having established the thresholds for estimating latent positions of GRDPGs under  $\varepsilon$ -edge LDP, we now focus on constructing an optimal estimator. Without privacy constraints, for  $(\mathbf{A}, \mathbf{X}) \sim \mathcal{G}(\mathbb{P}, \rho_n; p, q)$  the spectral embedding  $\hat{\mathbf{X}} = \text{sp}(\mathbf{A}; d)$  serves as an estimator for the latent positions after global scaling by the sparsity factor, i.e.,  $\rho_n^{1/2} \hat{\mathbf{X}}$  [57, Theorem 1]. In the regime where  $\rho_n = o(1)$ ,  $\hat{\mathbf{X}}$  estimates a quantity that shrinks to zero as  $n$  grows. Rescaling by the sparsity factor—specifically, using  $\rho_n^{-1/2} \hat{\mathbf{X}}$ —yields a consistent estimator for  $\mathbf{X}$  up to  $\mathcal{O}(p, q)$  transformations [57, Theorem 1] and [2, Theorem 4]. Under privacy constraints, however, a privacy mechanism  $\mathcal{A} \in \mathbb{A}_\varepsilon$  may alter  $\mathbf{A}$  in ways that the spectral embedding does not explicitly account for. Consequently, accurate statistical inference from the  $\varepsilon$ -edge LDP graph requires choosing an appropriate privacy mechanism and compensating for its geometric distortion in the estimation procedure.

To this end, we select  $\mathcal{A} = \mathcal{M}_\varepsilon$  as the `edgeFlip` mechanism and introduce a *privacy-adjusted spectral embedding* procedure in Algorithm 1. The intuition behind Algorithm 1 is as follows. By Theorem 3.1, the latent positions of  $\mathcal{M}_\varepsilon(\mathbf{A})$  and  $\mathbf{A}$  are connected via the mapping  $\varphi_\varepsilon$ . Since the privacy parameter  $\varepsilon$  can be disclosed publicly without disclosure risk [18, 20], we can reverse the effect of  $\varphi_\varepsilon$  on the spectral estimate by applying the inverse transformation to  $\mathcal{M}_\varepsilon(\mathbf{A})$ . The remaining challenge is that the sparsity parameter  $\rho_n$  is not private. In principle, one might aim to release a private estimate of  $\rho_n$  by optimally allocating the privacy budget  $\varepsilon$ , for example using the MWEM procedure of [26] or one of its variants; however, this approach is incompatible with local-DP constraints. In lieu of a private estimate, we mitigate the influence of  $\rho_n$  by computing an estimate  $\check{\rho}_n$ , as described in Step 2 of Algorithm 1. The following result establishes that this procedure yields a consistent estimator for the latent positions  $\mathbf{X}$ , up to  $\mathcal{O}(p, q)$  transformations.

**Theorem 3.3.** *Suppose  $(\mathbf{X}, \mathbf{A}) \sim \mathcal{G}(\mathbb{P}, \rho_n; p, q)$  where  $\mathbb{E}(X_1^\top \mathbb{I}_{p,q} X_2) = 1$ . For  $\varepsilon > 0$  let  $\mathcal{M}_\varepsilon(\mathbf{A})$  be the  $\varepsilon$ -edge LDP graph under `edgeFlip`, and let  $\check{\mathbf{X}}$  be the privacy-adjusted spectral embedding using Algorithm 1. If  $n\rho_n = \omega(\log n)$ , then with probability  $1 - O(n^{-1})$ ,*

$$d_{2,\infty} \left( \frac{\check{\mathbf{X}}}{\sqrt{\check{\rho}_n}}, \mathbf{X} \right) = O \left( \frac{\log n}{\sqrt{n\sigma(\varepsilon)^4 \rho_n^2}} \right). \quad (12)$$

In light of Theorem 3.2, we see that the convergence rate of the privacy-adjusted spectral embedding  $\check{\mathbf{X}}$  matches the lower bound in (9) up to an  $O(\sqrt{\log n})$  gap. A similar discrepancy exists in the best-known bounds in the absence of privacy constraints. In [57, Theorem 1] and [2, Theorem 4], the convergence rate



**Algorithm 1:** Privacy-Adjusted Spectral Embedding**Input:**  $\mathcal{M}_\varepsilon(\mathbf{A})$  and privacy budget  $\varepsilon$ 

- 1: Compute the privacy-adjusted adjacency matrix

$$\check{\mathbf{A}} = \frac{1}{\sigma^2(\varepsilon)} \left( \mathcal{M}_\varepsilon(\mathbf{A}) - \tau^2(\varepsilon) \mathbf{1}_n \mathbf{1}_n^\top \right).$$

- 2: Estimate the sparsity
- $\check{\rho}_n$

$$\check{\rho}_n = \binom{n}{2}^{-1} \sum_{i < j} \check{\mathbf{A}}_{ij}.$$

- 3: Spectral embedding

$$\check{\mathbf{X}} = \text{sp}(\check{\mathbf{A}}; d).$$

**return**  $\check{\mathbf{X}} \in \mathbb{R}^{n \times d}$  and  $\check{\rho}_n \in [0, 1]$ for estimating  $\mathbf{X}$  using  $\hat{\mathbf{X}} = \text{sp}(\mathbf{A}; d)$  is

$$d_{2,\infty} \left( \frac{\hat{\mathbf{X}}}{\sqrt{\rho_n}}, \mathbf{X} \right) = O_p \left( \frac{\log n}{\sqrt{n \rho_n}} \right). \quad (13)$$

In contrast to (8), the rate in (13) has precisely the same  $O(\sqrt{\log n})$  gap. Therefore, modulo  $\log n$  factors, for the problem of estimating the latent positions of GRDPGs under  $\varepsilon$ -edge LDP, we see that: (i) **edge-Flip** is the optimal privacy mechanism, and (ii) the privacy-adjusted spectral embedding  $\check{\mathbf{X}}$  is the optimal estimator. Moreover, the range of parameters satisfying  $\rho_n^2 \sigma(\varepsilon)^4 = \tilde{\Theta}(1/n)$  determines a phase-transition boundary for consistent estimation.

We also point out that, as noted in the discussion preceding (6),  $\mathbb{E}(X_1^\top \mathbb{I}_{p,q} X_2) = 1$  is fixed in Theorem 3.3 in order to ensure  $\rho_n$  is identifiable and that Step 2 of Algorithm 1 produces a consistent estimate of  $\rho_n$ . If, instead,  $\mathbb{E}(X_1^\top \mathbb{I}_{p,q} X_2) = \mu$ , then the results in both (12) and (13) will hold for  $\mathbf{X}/\sqrt{\mu}$  in place of  $\mathbf{X}$ .

**Experiment 3.1.** *In order to examine the bound in Theorem 3.3, we generate  $(\mathbf{A}, \mathbf{X}) \sim \mathcal{G}(\mathbb{P}, \rho_n; p, q)$  where  $\mathbb{P} = \text{Unif}(S^1)$  is the uniform distribution on the circle,  $\rho_n = \log^4 n / \sqrt{n}$ , and  $p, q = 1$ . For each  $\varepsilon, n$ , the  $\varepsilon$ -edge LDP graph  $\mathcal{M}_\varepsilon(\mathbf{A})$  is generated via **edge-Flip** and  $\check{\mathbf{X}}$  is computed using Algorithm 1. The error  $d_{2,\infty}(\check{\mathbf{X}}, \mathbf{X})$  is averaged across 10 iterations. The results are shown in Figure 4. The shaded curves, given by  $\sigma(\alpha\varepsilon)^2 = \log n / \sqrt{n \rho_n^2}$  for  $\alpha \in \{25, 35, 55\}$ , are contours where (12) is expected to be constant. As Figure 4 shows, the bound is fairly tight.*

**3.4 Topological inference under edgeFlip**

While the previous sections have focused on estimating the latent positions of a GRDPG under  $\varepsilon$ -edge LDP in the  $d_{2,\infty}$  metric, in this section we examine how  $\mathcal{M}_\varepsilon(\mathbf{A})$  can be used to extract meaningful topological and geometric information underlying  $\mathbf{X}$ . To this

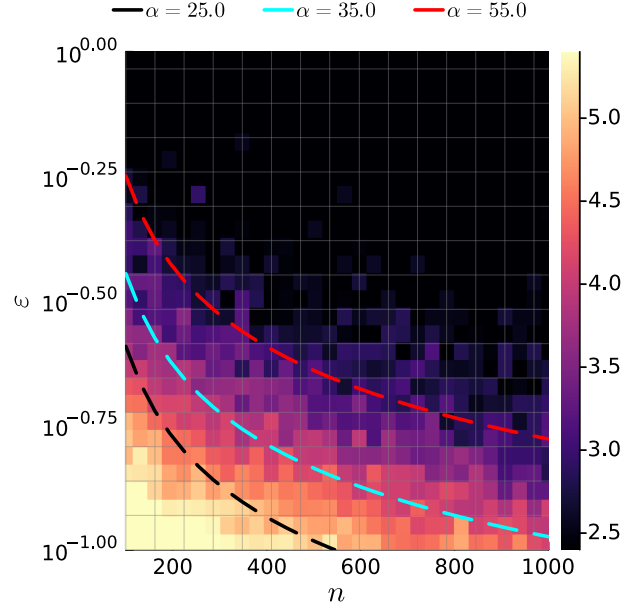


Figure 4: Illustration of the bounds in Theorem 3.3. Heatmap of  $d_{2,\infty}(\check{\mathbf{X}}, \mathbf{X})$  error for  $\varepsilon$  vs.  $n$  in Experiment 3.1.

end, we use tools from topological data analysis (TDA). Owing to length constraints, we provide an accessible overview of TDA in Appendix D. The main insight for this section comes from the following result, which is a simple consequence of Theorem 3.1.

**Corollary 3.2.** *Under the conditions of Theorem 3.1, the map  $\varphi_\varepsilon$  is a homeomorphism, and  $\varphi_\varepsilon(\mathbf{X}) \simeq \mathbf{X}$ .*

In simpler terms, Corollary 3.2 states that, in view of Theorem 3.1, the latent positions of  $\mathcal{M}_\varepsilon(\mathbf{A})$  are related to those of  $\mathbf{A}$  by deformations which only allow stretching/squeezing/bending. The result is immediately obvious since  $\varphi_\varepsilon : \mathbb{R}^d \rightarrow \mathbb{R}^{d+1}$  is a linear map. Notably, homeomorphisms preserve topological structures, as shown for spectral embeddings in Figure 2.

Persistent homology forms the backbone of most TDA routines, where multiscale geometric and topological features underlying a collection of points  $\mathbf{X} \in \mathbb{R}^{n \times d}$  are encoded in an object called a *persistence diagram*  $\mathfrak{Dgm}(\mathbf{X})$ . In a nutshell,  $\mathfrak{Dgm}(\mathbf{X})$  is a collection of points in the upper-half plane  $\Omega = \{(b, d) \in \mathbb{R}^2 : b \leq d\}$ , where each point  $(b, d) \in \mathfrak{Dgm}(\mathbf{X})$  with  $b < d$  represents the birth time,  $b$ , and death time,  $d$ , of a topological feature in  $\mathbf{X}$ . The topological features are indexed by the dimension of the feature, e.g.,  $0d$  features are connected components or ‘clusters’,  $1d$  features are loops,  $2d$  features are holes, and so on. Given two persistence diagrams  $D_1, D_2 \subset \Omega$ , the *bottleneck distance*  $W_\infty(D_1, D_2)$  is quantifies the similarity between the topological features in the two diagrams.

The following result shows that the persistence diagram  $\mathfrak{Dgm}(\check{\mathbf{X}}/\check{\rho}_n)$  is a consistent estimator of  $\mathfrak{Dgm}(\mathbf{X})$ .

**Theorem 3.4.** *Under the conditions of Theorem 3.3, let  $D_n = \mathfrak{Dgm}(\mathbf{X})$  and  $\check{D}_n = \mathfrak{Dgm}(\check{\mathbf{X}}/\sqrt{\rho_n})$ . Then, with probability at least  $1 - O(n^{-1})$ ,*

$$W_\infty(\check{D}_n, D_n) = O\left(\frac{\log n}{\sqrt{n\sigma(\varepsilon)^4\rho_n^2}}\right). \quad (14)$$

Theorem 3.4 guarantees that  $\check{\mathbf{X}}$  consistently recovers the essential topological features of  $\mathbf{X}$ ; this includes connected components, loops, and other salient geometric and topological features. Since the 0th order persistence diagram precisely captures the clusters in the underlying data and by noting that the rate in (14) is the same as in (12), the privacy-adjusted spectral embedding  $\check{\mathbf{X}}$  provides the same privacy-utility trade-off guarantees as in Section 3.3 and can recover more general topological features of  $\mathbf{X}$ .

Although we do not pursue it in this work, the convergence in  $d_{2,\infty}$  metric in Theorem 3.3 can be used to show that “flat clustering” methods like  $k$ -means are consistent under  $\varepsilon$ -edge LDP when the latent positions form well-separated clusters (see, e.g., [46, Theorem 6]). Similarly, we believe that it should be possible to establish the consistency of DBSCAN when the distribution generating the latent positions,  $\mathbb{P}$ , is sufficiently regular, e.g., admits a Hölder density. In other words, Theorem 3.3 could be used to show the consistency of other downstream inferential procedures under  $\varepsilon$ -edge LDP when necessary conditions are met.

## 4 Experiments

**Experiment 4.1.** *To examine recovery of latent positional information in real-world data, we use the OpenFlights dataset where  $\mathbf{A}$  represents direct flight connections between airports. Figure 5(a) displays the scatterplot of  $\mathfrak{P}(\mathbf{X})$ , obtained using spectral embedding of  $\mathbf{A}$  followed by UMAP—which uses topological information to perform unsupervised dimension reduction [47, Section 2]. Each airport  $\{i\}$  is colored based on its timezone (see Appendix D for details). Although the unweighted nature of the graph implies that all connections are treated equally, and disregards any additional information such as distance, flight duration, or frequency— $\mathfrak{P}(\mathbf{X})$  recovers the subtle geographic information associated with the airports, e.g., Japan is close to China and Italy is close to the United Kingdom. Since  $\sigma(\varepsilon) \approx \varepsilon$  in the admissible regime of (14), taking  $\rho_n$  to be the sparsity of  $\mathbf{A}$ ,  $\beta \in \{\frac{1}{3}, \frac{2}{3}, \frac{11}{10}\}$ , and*

$$\varepsilon_n = \log^2 n / \rho_n^2 n^\beta,$$

*Figures 5(b,c,d) show the plots of  $\mathfrak{P}(\check{\mathbf{X}})$  obtained using  $\mathcal{M}_\varepsilon(\mathbf{A})$ . For moderately small  $\varepsilon$ , Figures 5(b,c) show that  $\mathfrak{P}(\check{\mathbf{X}})$  preserves (to a large extent) the salient relationships in Figure 5(a). However, for very small  $\varepsilon$ , as shown in Figure 5(d),  $\mathfrak{P}(\check{\mathbf{X}})$  no longer captures this information.*

**Experiment 4.2.** *In this experiment, we illustrate the benefit of the topological perspective by employing persistence diagrams to perform  $\varepsilon$ -edge LDP community detection. We generate  $N = 4n$  points for  $\mathbf{X}$  from the following shape: (i)  $2n$  are sampled uniformly from a circle, (ii)  $n$  points sampled from a Lemniscate (which looks like “ $\infty$ ”), and (iii)  $n$  points sampled from a cluster inside the circle but disjoint from the rest. The shape contains 3 non-trivial connected components, and has three loops (order-1 homological features): one from the circle, and two from the Lemniscate. See Figure 6(a). We then generate  $\mathbf{A} \sim \mathcal{G}(\mathbf{X}, 1; 2, 0)$ .*

*For a range of  $\varepsilon$ , we obtain the persistence diagram  $\check{D}^\varepsilon$  from  $\mathcal{M}(\mathbf{A})$  and compute the bottleneck distance  $W_\infty(D, \check{D}^\varepsilon)$ . The results are averaged across 10 iterations. Figure 6(b) shows the convergence in the bottleneck distance for  $n \in \{500, \dots, 1500\}$ , and the black dashed line plots the r.h.s. of (14) when  $n = 1500$ .*

*Clustering for  $\mathbf{X}$  by itself is particularly challenging for algorithms like  $k$ -means since  $\mathbf{X}$  doesn’t admit “flat” clusters. To address this limitation, we use a simplified variant of the algorithm in [41] as outlined in Algorithm 2 in Appendix D. Figure 6(c) plots the Adjusted RAND Index between the true labels and the predicted labels obtained using: (i) the topological approach, (ii)  $k$ -means when  $k = 3$  is provided as input, and (iii) using DBSCAN with  $\text{minPts}$  to be the size of the smallest cluster. The results illustrate how Theorem 3.4 enables access to using more tailored topological approaches to recover the underlying clusters.*

**Experiment 4.3.** *As noted in Section 3.3, the spectral approach is suboptimal at the information-theoretic thresholds for sparse graphs, and several SDP-based methods have been proposed to address this limitation for specific models [1]. Nevertheless, this approach comes at the expense of computational time and a careful selection of hyperparameters. To benchmark the performance, we generate  $\mathbf{A}$  from an imbalanced 2-class SBM with imbalance ratio 1/3. See Figure 7(a).*

*SDP-solvers and “flat clustering” methods, e.g.,  $k$ -means, are particularly well suited for SBMs. We use two SDP-approaches: (i)  $\text{SDP}_1$  from [25] which works for balanced SBMs, and (ii)  $\text{SDP}_2$  from [52, Eq. (2)] for imbalanced SBMs with the imbalance ratio input to 1/3. Figure 7(b) plots the Adjusted RAND Index between the true labels and the predicted labels. Here, the spectral approaches do about as well as  $\text{SDP}_2$  when the imbalance ratio is known. On the other hand  $\text{SDP}_1$  is suboptimal compared to the other methods. Figure 7(c), shows the runtime of the methods. Notably,  $k$ -means and Algorithm 2 use sparse eigensolvers and scale well, whereas SDP methods can become prohibitively expensive for large graphs.*



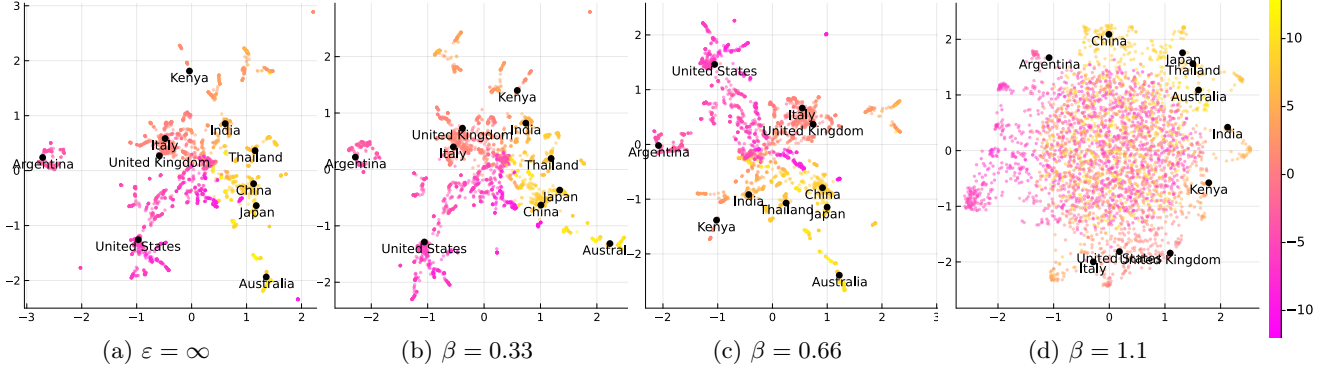


Figure 5: Spectral embedding followed by UMAP for the OpenFlights network in Experiment 4.1.

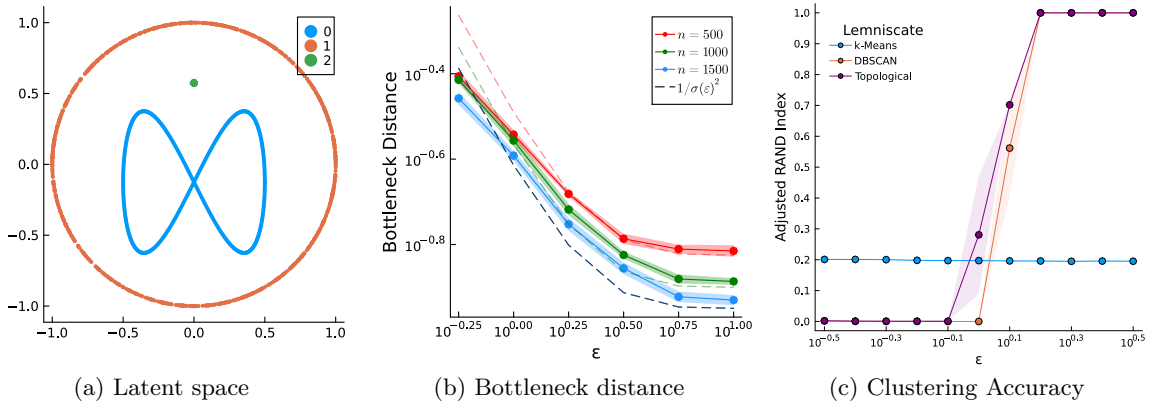


Figure 6: Recovery of topological information for the setup in Experiment 4.2.

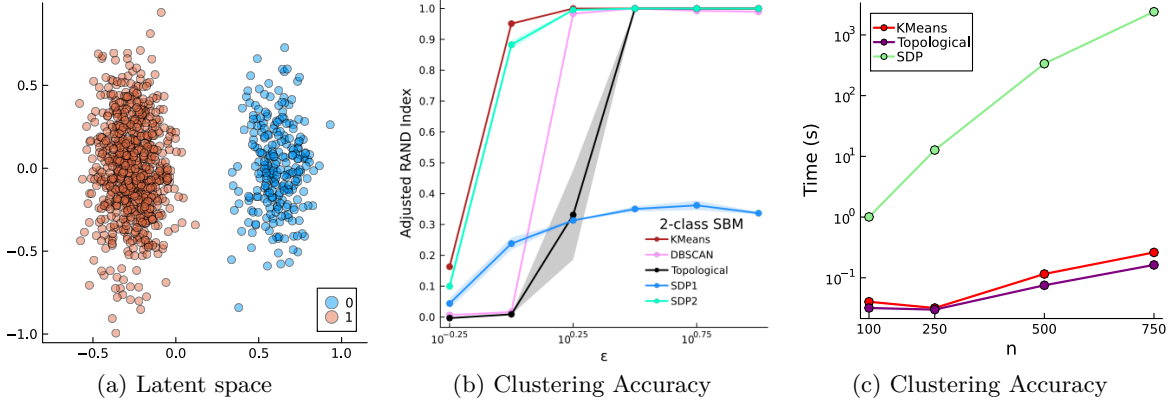


Figure 7: Performance of spectral methods alongside SDP-based methods in Experiment 4.3.

## 5 Conclusion

In this work, we address the problem of estimating the latent positions of a GRDPG under  $\varepsilon$ -edgeLDP constraints. If  $(\mathbf{A}, \mathbf{X}) \sim \mathcal{G}(\mathbb{P}, \rho_n; p, q)$  is a GRDPG with latent positions  $\mathbf{X}$ , we showed that the  $\varepsilon$ -edge LDP synthetic graph  $\mathcal{M}_\varepsilon(\mathbf{A})$  obtained via `edgeFlip` can be used to construct a privacy-adjusted spectral embedding  $\tilde{\mathbf{X}}$  that achieves near-minimax optimal rates for sufficiently dense graphs. Additionally, we showed that

the class of GRDPGs is closed under `edgeFlip`, with the latent positions of  $\mathcal{M}_\varepsilon(\mathbf{A})$  related to those of  $\mathbf{A}$  by a simple geometric transformation. This perspective enabled us to tackle more nuanced inferential tasks, such as recovering topological information through persistence diagrams. For future work, it would be interesting to investigate the optimality of recovering the latent positions  $\mathbf{X}$  under central  $\varepsilon$ -DP constraints, and to extend the analysis to more general graphical models such as graphons.

## References

- [1] Abbe, E. (2017). Community detection and stochastic block models: recent developments. *Journal of Machine Learning Research*, 18(1):6446–6531.
- [2] Agterberg, J., Tang, M., and Priebe, C. (2020a). Nonparametric two-sample hypothesis testing for random graphs with negative and repeated eigenvalues. *arXiv preprint arXiv:2012.09828*.
- [3] Agterberg, J., Tang, M., and Priebe, C. E. (2020b). On two distinct sources of nonidentifiability in latent position random graph models. *arXiv preprint arXiv:2003.14250*.
- [4] Airoldi, E. M., Blei, D. M., Fienberg, S. E., and Xing, E. P. (2008). Mixed membership stochastic blockmodels. *Journal of Machine Learning Research*.
- [5] Asoodeh, S., Aliakbarpour, M., and Calmon, F. P. (2021). Local differential privacy is equivalent to contraction of an  $f$ -divergence. In *2021 IEEE International Symposium on Information Theory (ISIT)*, pages 545–550. IEEE.
- [6] Athreya, A., Fishkind, D. E., Tang, M., Priebe, C. E., Park, Y., Vogelstein, J. T., Levin, K., Lyzinski, V., and Qin, Y. (2017). Statistical inference on random dot product graphs: A survey. *The Journal of Machine Learning Research*, 18(1):8393–8484.
- [7] Backstrom, L., Dwork, C., and Kleinberg, J. (2007). Wherefore art thou r3579x? anonymized social networks, hidden patterns, and structural steganography. In *Proceedings of the 16th international conference on World Wide Web*, pages 181–190.
- [8] Barber, R. F. and Duchi, J. C. (2014). Privacy and statistical risk: Formalisms and minimax bounds. *arXiv preprint arXiv:1412.4451*.
- [9] Belkin, M. and Niyogi, P. (2003). Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, 15(6):1373–1396.
- [10] Cape, J., Tang, M., and Priebe, C. E. (2019). The two-to-infinity norm and singular subspace geometry with applications to high-dimensional statistics. *The Annals of Statistics*, 47(5):2405–2439.
- [11] Chakraborty, A., Chatterjee, S., and Nandy, S. (2024). PriME: Privacy-aware membership profile estimation in networks. *arXiv preprint arXiv:2406.02794*.
- [12] Chazal, F., De Silva, V., Glisse, M., and Oudot, S. (2016). *The Structure and Stability of Persistence Modules*. Springer.
- [13] Chazal, F., Guibas, L. J., Oudot, S. Y., and Skraba, P. (2013). Persistence-based clustering in Riemannian manifolds. *Journal of the ACM (JACM)*, 60(6):1–38.
- [14] Chazal, F. and Michel, B. (2017). An introduction to topological data analysis: Fundamental and practical aspects for data scientists. *arXiv preprint arXiv:1710.04019*.
- [15] Chen, H., Cohen-Addad, V., dOrsi, T., Epasto, A., Imola, J., Steurer, D., and Tiegel, S. (2023). Private estimation algorithms for stochastic block models and mixture models. *Advances in Neural Information Processing Systems*, 36:68134–68183.
- [16] Duchi, J. C., Jordan, M. I., and Wainwright, M. J. (2013). Local privacy and statistical minimax rates. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 429–438. IEEE.
- [17] Duchi, J. C., Jordan, M. I., and Wainwright, M. J. (2018). Minimax optimal procedures for locally private estimation. *Journal of the American Statistical Association*, 113(521):182–201.
- [18] Dwork, C., Kohli, N., and Mulligan, D. (2019). Differential privacy in practice: Expose your ep-silons! *Journal of Privacy and Confidentiality*, 9(2).
- [19] Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography Conference*, pages 265–284. Springer.
- [20] Dwork, C., Roth, A., et al. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4):211–407.
- [21] Edelsbrunner, H. and Harer, J. (2010). *Computational Topology: An Introduction*. American Mathematical Society.
- [22] Eden, T., Liu, Q. C., Raskhodnikova, S., and Smith, A. D. (2023). Triangle counting with local edge differential privacy. In *International Colloquium on Automata, Languages, and Programming*.
- [23] Gao, C., Ma, Z., Zhang, A. Y., and Zhou, H. H. (2018). Community detection in degree-corrected block models. *The Annals of Statistics*, 46(5):2153–2185.
- [24] Grover, A. and Leskovec, J. (2016). node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 855–864.
- [25] Hajek, B., Wu, Y., and Xu, J. (2016). Achieving exact cluster recovery threshold via semidefinite programming. *IEEE Transactions on Information Theory*, 62(5):2788–2797.
- [26] Hardt, M., Ligett, K., and McSherry, F. (2012). A simple and practical algorithm for differentially

- private data release. *Advances in neural information processing systems*, 25.
- [27] Hay, M., Li, C., Miklau, G., and Jensen, D. (2009). Accurate estimation of the degree distribution of private networks. In *2009 Ninth IEEE International Conference on Data Mining*, pages 169–178. IEEE.
- [28] He, W., Fichtenberger, H., and Peng, P. (2024). A differentially private clustering algorithm for well-clustered graphs. In *The Twelfth International Conference on Learning Representations*.
- [29] Hehir, J., Slavković, A., and Niu, X. (2022). Consistent spectral clustering of network block models under local differential privacy. *Journal of Privacy and Confidentiality*, 12(2).
- [30] Hoff, P. D., Raftery, A. E., and Handcock, M. S. (2002). Latent space approaches to social network analysis. *Journal of the American Statistical Association*, 97(460):1090–1098.
- [31] Holland, P. W., Laskey, K. B., and Leinhardt, S. (1983). Stochastic blockmodels: First steps. *Social Networks*, 5(2):109–137.
- [32] Horn, R. A. and Johnson, C. R. (2012). *Matrix Analysis*. Cambridge University Press.
- [33] Imola, J., Murakami, T., and Chaudhuri, K. (2021). Locally differentially private analysis of graph statistics. In *30th USENIX Security Symposium (USENIX Security ’21)*.
- [34] Jiang, H., Pei, J., Yu, D., Yu, J., Gong, B., and Cheng, X. (2021). Applications of differential privacy in social network analysis: A survey. *IEEE Transactions on Knowledge and Data Engineering*.
- [35] Karrer, B. and Newman, M. E. (2011). Stochastic blockmodels and community structure in networks. *Physical Review E*, 83(1):016107.
- [36] Karwa, V., Krivitsky, P. N., and Slavković, A. B. (2017). Sharing social network data: Differentially private estimation of exponential family random-graph models. *Journal of the Royal Statistical Society. Series C: Applied Statistics*, 66(3):481–500.
- [37] Karwa, V., Raskhodnikova, S., Smith, A., and Yaroslavtsev, G. (2011). Private analysis of graph structure. *Proceedings of the VLDB Endowment*, 4(11):1146–1157.
- [38] Karwa, V. and Slavković, A. (2016). Inference using noisy degrees: Differentially private  $\beta$ -model and synthetic graphs. *The Annals of Statistics*, 44(1):87–112.
- [39] Kasiviswanathan, S. P., Lee, H. K., Nissim, K., Raskhodnikova, S., and Smith, A. (2011). What can we learn privately? *SIAM Journal on Computing*, 40(3):793–826.
- [40] Kifer, D. and Machanavajjhala, A. (2011). No free lunch in data privacy. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of Data*, pages 193–204.
- [41] Kurlin, V. (2016). A fast persistence-based segmentation of noisy 2d clouds with provable guarantees. *Pattern Recognition Letters*, 83:3–12.
- [42] Lei, J. and Rinaldo, A. (2015). Consistency of spectral clustering in stochastic block models. *The Annals of Statistics*, 43(1):215–237.
- [43] Li, M., Berrett, T., and Yu, Y. (2022). Network change point localisation under local differential privacy. *Advances in Neural Information Processing Systems*, 35:15013–15026.
- [44] Li, Y., Purcell, M., Rakotoarivelo, T., Smith, D., Ranbaduge, T., and Ng, K. S. (2023). Private graph data release: A survey. *ACM Computing Surveys*, 55(11):1–39.
- [45] Lunde, R. and Sarkar, P. (2023). Subsampling sparse graphons under minimal assumptions. *Biometrika*, 110(1):15–32.
- [46] Lyzinski, V., Sussman, D. L., Tang, M., Athreya, A., and Priebe, C. E. (2014). Perfect clustering for stochastic blockmodel graphs via adjacency spectral embedding. *Electronic Journal of Statistics*, 8:2905–2922.
- [47] McInnes, L., Healy, J., and Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*.
- [48] Mohamed, M. S., Nguyen, D., Vullikanti, A., and Tandon, R. (2022). Differentially private community detection for stochastic block models. In *International Conference on Machine Learning*, pages 15858–15894. PMLR.
- [49] Mülle, Y., Clifton, C., and Böhm, K. (2015). Privacy-integrated graph clustering through differential privacy. In *EDBT/ICDT Workshops*, pages 247–254.
- [50] Narayanan, A., Shi, E., and Rubinstein, B. I. (2011). Link prediction by de-anonymization: How we won the Kaggle social network challenge. In *The 2011 International Joint Conference on Neural Networks*, pages 1825–1834. IEEE.
- [51] Narayanan, A. and Shmatikov, V. (2008). Robust de-anonymization of large sparse datasets. In *2008 IEEE Symposium on Security and Privacy (sp 2008)*, pages 111–125.
- [52] Nguyen, D. and Vullikanti, A. K. (2024). Differentially private exact recovery for stochastic block models. In *Forty-first International Conference on Machine Learning*.

- [53] Perozzi, B., Al-Rfou, R., and Skiena, S. (2014). Deepwalk: Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 701–710.
- [54] Qin, Z., Yu, T., Yang, Y., Khalil, I., Xiao, X., and Ren, K. (2017). Generating synthetic decentralized social graphs with local differential privacy. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, pages 425–438.
- [55] Rohde, A. and Steinberger, L. (2020). Geometrizing rates of convergence under local differential privacy constraints. *The Annals of Statistics*, 48(5):2646–2670.
- [56] Rubin-Delanchy, P. (2020). Manifold structure in graph embeddings. *Advances in Neural Information Processing Systems*, 33.
- [57] Rubin-Delanchy, P., Cape, J., Tang, M., and Priebe, C. E. (2022). A statistical interpretation of spectral embedding: The generalised random dot product graph. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 84(4):1446–1473.
- [58] Solanki, V., Rubin-Delanchy, P., and Gallagher, I. (2019). Persistent homology of graph embeddings. *arXiv preprint arXiv:1912.10238v1*.
- [59] Tropp, J. A. (2015). An introduction to matrix concentration inequalities. *Foundations and Trends® in Machine Learning*, 8(1-2):1–230.
- [60] Tsybakov, A. B. (2008). *Introduction to Nonparametric Estimation*. Springer Publishing Company, Incorporated, 1st edition.
- [61] Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American statistical association*, 60(309):63–69.
- [62] Wasserman, L. (2018). Topological data analysis. *Annual Review of Statistics and Its Application*, 5:501–532.
- [63] Yan, H. and Levin, K. (2023). Minimax rates for latent position estimation in the generalized random dot product graph. *arXiv preprint arXiv:2307.01942*.
- [64] Yu, B. (1997). Assouad, Fano, and Le Cam. In *Festschrift for Lucien Le Cam: Research papers in probability and statistics*, pages 423–435. Springer.
- [65] Zhang, A. Y. and Zhou, H. H. (2016). Minimax rates of community detection in stochastic block models. *The Annals of Statistics*, 44(5):2252–2280.

## Checklist

1. For all models and algorithms presented, check if you include:
  - (a) A clear description of the mathematical setting, assumptions, algorithm, and/or model. [**Yes**]
  - (b) An analysis of the properties and complexity (time, space, sample size) of any algorithm. [**No**] *The time/space complexity of the algorithms are standard.*
  - (c) (Optional) Anonymized source code, with specification of all dependencies, including external libraries. [**Yes**]
2. For any theoretical claim, check if you include:
  - (a) Statements of the full set of assumptions of all theoretical results. [**Yes**]
  - (b) Complete proofs of all theoretical results. [**Yes**]
  - (c) Clear explanations of any assumptions. [**Yes**]
3. For all figures and tables that present empirical results, check if you include:
  - (a) The code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL). [**Yes**]
 

<https://github.com/sidv23/grdpg-ldp.git>
  - (b) All the training details (e.g., data splits, hyperparameters, how they were chosen). [**Yes**]
  - (c) A clear definition of the specific measure or statistics and error bars (e.g., with respect to the random seed after running experiments multiple times). [**Yes**]
  - (d) A description of the computing infrastructure used. (e.g., type of GPUs, internal cluster, or cloud provider). [**No**]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets, check if you include:
  - (a) Citations of the creator If your work uses existing assets. [**Not Applicable**]
  - (b) The license information of the assets, if applicable. [**Not Applicable**]
  - (c) New assets either in the supplemental material or as a URL, if applicable. [**Not Applicable**]
  - (d) Information about consent from data providers/curators. [**Not Applicable**]

---

# Supplementary Materials

## Signal Recovery from Random Dot-Product Graphs Under Differential Privacy

1	Introduction	1
2	Background	3
3	Main Results	4
4	Experiments	8
5	Conclusion	9
A	Spectral alignment matrices	14
B	Auxiliary Results	15
C	Proofs	20
D	Topological Data Analysis	30

A summary of notations is provided in Table 1. Throughout, we use  $a_n \lesssim b_n$  and  $a_n = O(b_n)$  to denote  $a_n \leq Cb_n$  for some constant  $C > 0$  which may change from line to line but *does not* depend on  $\varepsilon$  or  $n$ . We use  $O_p(b_n)$  and  $o_p(b_n)$  to denote the usual Mann-Wald asymptotic order for random variables, i.e.,  $X_n = O_p(a_n)$  if  $\mathbb{P}(|X_n/a_n| > C_r) \leq n^{-r}$ , and  $X_n = o_p(a_n)$  if  $\limsup_n \mathbb{P}(|X_n/a_n| > C) = 0$  for all  $C > 0$ .

Table 1: Table of notations

$\mathcal{O}(d), \mathbb{I}_d$	The group of $d \times d$ orthogonal matrices, and the identity matrix
$\mathcal{O}(p, q), \mathbb{I}_{p, q}$	The group of $d \times d$ indefinite orthogonal matrices, and the indefinite identity matrix
$\mathbb{B}(n)$	The set of binary, symmetric $n \times n$ matrices $\subset \{0, 1\}^{n \times n}$
$\mathbb{U}(n, d)$	The set (Stiefel manifold) of $n \times d$ matrices with orthonormal columns satisfying $\mathbf{U}^\top \mathbf{U} = \mathbb{I}_d$
$\Lambda(\mathbf{A})$	The diagonal matrix of eigenvalues of a square matrix $\mathbf{A}$
$\ \mathbf{X}\ _{\text{op}}$	The operator/spectral norm of matrix $\mathbf{X}$ given by $\sqrt{\lambda_{\max}(\mathbf{X}^\top \mathbf{X})}$
$\ \mathbf{X}\ _{2, \infty}$	The two-to-infinity norm of matrix $\mathbf{X} \in \mathbb{R}^{n \times d}$ given by $\max_{1 \leq i \leq n} \ \mathbf{X}_{i, *}\ _2$
$\ \mathbf{u}\ $	The $\ell_2$ -norm of a vector $\mathbf{u} \in \mathbb{R}^d$
$(\mathbf{A}, \mathbf{X}) \sim \mathcal{G}(\mathbb{P}, \rho_n; p, q)$	GRDPG with $(p, q)$ -admissible measure $\mathbb{P}$ , sparsity $\rho_n \leq 1$ and latent positions $\mathbf{X} \in \mathbb{R}^{n \times d}$
$\mathbf{P} \equiv \mathbf{P}_{\mathbf{X}}$	The expected adjacency matrix of $\mathbf{A}$ , i.e., $\mathbf{P} \equiv \mathbf{P}_{\mathbf{X}} = \mathbf{X} \mathbb{I}_{p, q} \mathbf{X}^\top$
$\Delta_{\mathbf{X}}$	The <i>empirical</i> second moment matrix of $\mathbf{X}$ , i.e., $\Delta_{\mathbf{X}} =$ for $\xi \sim \mathbb{P}$
$\Delta_{\xi}, \Delta_{\mathbb{P}}$	The <i>population</i> second moment matrix of $\mathbb{P}$ , i.e., $\Delta_{\xi} \equiv \Delta_{\mathbb{P}} = \mathbb{E}(\xi \xi^\top)$ for $\xi \sim \mathbb{P}$
$\hat{\mathbf{X}} = \text{sp}(\mathbf{A}; d)$	The adjacency spectral embedding of $\mathbf{A}$ in $\mathbb{R}^d$ ; see (4)
$\mathbf{O}_{\mathbf{X}}$	The $\mathcal{O}(d)$ matrix aligning $\hat{\mathbf{X}}$ to $\mathbf{U}_{\mathbf{P}}  \Lambda_{\mathbf{P}} ^{1/2}$ ; see Fact A.1
$\mathbf{Q}_{\mathbf{X}}$	The $\mathcal{O}(p, q)$ matrix aligning $\mathbf{X}$ to $\mathbf{U}_{\mathbf{P}}  \Lambda_{\mathbf{P}} ^{1/2}$ ; see Fact A.2
$\mathbf{Q}_{\xi}$	The population analogue of $\mathbf{Q}_{\mathbf{X}}$ ; see Fact A.4
$\mathbf{O}_{\xi}$	The matrix aligning $\mathbf{Q}_{\mathbf{X}}$ to $\mathbf{Q}_{\xi}$ ; see Fact A.5
$\pi(\varepsilon), \sigma(\varepsilon), \tau(\varepsilon)$	Parameters of <code>edgeFlip</code> ; see (2)
$\mathcal{M}_{\varepsilon}(\mathbf{A})$	The $\varepsilon$ -edge LDP synthetic copy of $\mathbf{A}$ under <code>edgeFlip</code>
$\check{\mathbf{A}}, \check{\mathbf{X}}, \check{\rho}_n$	The privacy-adjusted adjacency matrix, spectral embedding and estimated sparsity
$\mathfrak{D}_{\text{gm}}(\mathbf{X})$	Persistence diagram associated with the rows of $\mathbf{X}$
$W_{\infty}(\mathbf{D}_1, \mathbf{D}_2)$	Bottleneck distance between two persistence diagrams $\mathbf{D}_1, \mathbf{D}_2$
$d_H(\mathbb{X}, \mathbb{Y})$	Hausdorff distance between two compact sets $\mathbb{X}$ and $\mathbb{Y}$

## A Spectral alignment matrices

There are several spectral alignment matrices which appear in the proofs of the main results. This section provides a brief overview of these matrices and their properties. See, also, [2, Tables 1 & 2] and [3, Section 3] for a comprehensive overview.

Let  $\xi \sim \mathbb{P}$  be a random vector in  $\mathbb{R}^d$  with distribution  $\mathbb{P}$  and let  $\mathbf{X} \in \mathbb{R}^{n \times d}$  whose rows  $X_1, X_2, \dots, X_n$  are i.i.d. copies of  $\xi$ . Following the setup in Section 3:

- (i) Let  $\Delta_\xi := \mathbb{E}(\xi\xi^\top) \in \mathbb{R}^{d \times d}$  be the second-moment matrix associated with  $\xi$  and let

$$\Delta_{\mathbf{X}} \doteq \frac{1}{n} \mathbf{X}^\top \mathbf{X} \in \mathbb{R}^{d \times d}$$

be the *empirical* second-moment matrix associated with  $\mathbf{X}$ .

- (ii) For the edge-probability matrix  $\mathbf{P} = \rho_n \mathbf{X} \mathbb{I}_{p,q} \mathbf{X}^\top \in [0, 1]^{n \times n}$  under sparsity  $\rho_n$ , let  $\mathbf{P} \doteq \mathbf{U}_{\mathbf{P}} \Lambda_{\mathbf{P}} \mathbf{U}_{\mathbf{P}}^\top$  be its spectral decomposition where  $\Lambda_{\mathbf{P}}$  contains the  $d$  eigenvalues of  $\mathbf{P}$  such that  $\text{sgn}(\Lambda_{\mathbf{P}}) = \mathbb{I}_{p,q}$  and  $\mathbf{U}_{\mathbf{P}} \in \mathbb{U}(n, d)$ . Define

$$\tilde{\mathbf{X}} \doteq \mathbf{U}_{\mathbf{P}} |\Lambda_{\mathbf{P}}|^{1/2} \in \mathbb{R}^{n \times d} \quad (15)$$

as a surrogate for  $\rho_n^{1/2} \mathbf{X}$ .

- (iii) For the random graph  $\mathbf{A} \in \mathbb{B}(n)$  such that  $\mathbb{P}(\mathbf{A}_{ij} = 1) = \mathbf{P}_{ij}$ , let

$$\mathbf{A} = \mathbf{U} \Lambda \mathbf{U}^\top + \mathbf{U}_\perp \Lambda_\perp \mathbf{U}_\perp^\top$$

be its spectral decomposition of  $\mathbf{A}$  where  $\Lambda \equiv \Lambda_{\mathbf{A}}$  contains the top- $d$  eigenvalues of  $\mathbf{A}$  *by magnitude* and  $\mathbf{U} \in \mathbb{U}(n, d)$  contains the corresponding eigenvectors. From Definition 2.4,

$$\hat{\mathbf{X}} \doteq \mathbf{U} |\Lambda|^{1/2} \in \mathbb{R}^{n \times d} \quad (16)$$

is the *adjacency spectral embedding* of  $\mathbf{A}$ .

**Remark A.1.** Note that the matrices  $\hat{\mathbf{X}}$  and  $\tilde{\mathbf{X}}$  depend on the sparsity via rescaling by a factor of  $\rho_n^{1/2}$  whereas the matrix  $\mathbf{X}$  does not.

Comparing the expressions in (15) and (16) and by noting that  $\mathbb{E}(\mathbf{A}) = \mathbf{P}$ , one would expect that  $\hat{\mathbf{X}}$  is close to  $\tilde{\mathbf{X}}$ . However, since  $\mathbf{U}_{\mathbf{X}}, \mathbf{U} \in \mathbb{U}(n, d)$  arise from their respective spectral decompositions, they are unique only up to orthogonal transformations. The matrix  $\mathbf{O}_{\mathbf{X}}$  is the orthogonal matrix that aligns  $\hat{\mathbf{X}}$  to  $\tilde{\mathbf{X}}$ .

**Fact A.1.**  $\mathbf{O}_{\mathbf{X}} \in \mathbb{O}(d)$  is the matrix which solves the orthogonal Procrustes problem:

$$\mathbf{O}_{\mathbf{X}} \doteq \arg \min_{\mathbf{O} \in \mathbb{O}(d)} \|\hat{\mathbf{X}} - \tilde{\mathbf{X}} \mathbf{O}\|_F^2.$$

The positions  $\tilde{\mathbf{X}} \in \mathbb{R}^{n \times d}$  act as a surrogate for the rescaled latent positions  $\rho_n^{1/2} \mathbf{X}$ . Since  $\tilde{\mathbf{X}}$  arises from  $\mathbf{P}$ —which admits an indefinite spectral decomposition,  $\tilde{\mathbf{X}}$  and  $\rho_n^{1/2} \mathbf{X}$  are related by an  $\mathbb{O}(p, q)$  transformation.

**Fact A.2** ([57, p. 1457], [3, Section 4] and [2, Section 3.1]). The matrix  $\mathbf{Q}_{\mathbf{X}} \in \mathbb{O}(p, q)$  aligns  $\tilde{\mathbf{X}}$  to  $\rho_n^{1/2} \mathbf{X}$ , i.e.,

$$\rho_n^{1/2} \mathbf{X} \mathbf{Q}_{\mathbf{X}}^{-1} = \tilde{\mathbf{X}} = \mathbf{U}_{\mathbf{P}} |\Lambda_{\mathbf{P}}|^{1/2}.$$

From [3, Eq. (1)], the expression for  $\mathbf{Q}_{\mathbf{X}}$  is given by

$$\mathbf{Q}_{\mathbf{X}} \doteq \left( \frac{1}{n \rho_n} |\Lambda_{\mathbf{P}}| \right)^{-1/2} \mathbf{V}^\top \Delta_{\mathbf{X}}^{1/2}, \quad (17)$$

where  $\mathbf{V} \in \mathbb{O}(d)$  are the eigenvectors in the spectral decomposition  $\Delta_{\mathbf{X}}^{1/2} \mathbb{I}_{p,q} \Delta_{\mathbf{X}}^{1/2} = \mathbf{V} \Lambda_2 \mathbf{V}^\top$ .



The expression in (17) is slightly different from [3, Eq. (1)] since the authors don't consider the rescaling by the sparsity factor  $\rho_n^{1/2}$ . The stated expression holds by noting the following relationship between the eigenvalues of  $\mathbf{P}$  and  $\Delta_{\mathbf{X}}$ .

**Fact A.3.** *The matrices  $\frac{1}{n\rho_n}\mathbf{P}$ ,  $\Delta_{\mathbf{X}}\mathbb{I}_{p,q}$  and  $\Delta_{\mathbf{X}}^{1/2}\mathbb{I}_{p,q}\Delta_{\mathbf{X}}^{1/2}$  have the same eigenvalues, i.e.,*

$$\frac{1}{n\rho_n}\Lambda_{\mathbf{P}} = \Lambda(\rho_n\Delta_{\mathbf{X}}\mathbb{I}_{p,q}) = \Lambda\left((\rho_n\Delta_{\mathbf{X}})^{1/2}\mathbb{I}_{p,q}(\rho_n\Delta_{\mathbf{X}})^{1/2}\right).$$

*Proof.* The proof is a simple consequence of the property that the eigenvalues associated with the product of compatible matrices is invariant to cyclic permutations (up to a collection of repeated zero eigenvalues). Specifically, if  $\mu$  is a non-zero eigenvalue of the matrix product  $\mathbf{BCD}$  with eigenvector  $\mathbf{u}$ , then  $\mathbf{D}\mathbf{u}$  is an eigenvector of  $\mathbf{DBC}$  with eigenvalue  $\mu$ , i.e.,  $\mathbf{DBC}(\mathbf{D}\mathbf{u}) = \mathbf{D}(\mathbf{BCD}\mathbf{u}) = \mathbf{D}\mu\mathbf{u} = \mu(\mathbf{D}\mathbf{u})$ . Therefore, it follows that  $\Lambda(\frac{1}{n\rho_n}\mathbf{P}) = \Lambda(\frac{1}{n}\mathbf{X}\mathbb{I}_{p,q}\mathbf{X}^T) = \Lambda(\mathbb{I}_{p,q}\frac{1}{n}\mathbf{X}^T\mathbf{X}) = \Lambda(\mathbb{I}_{p,q}\Delta_{\mathbf{X}}) = \Lambda(\Delta_{\mathbf{X}}^{1/2}\mathbb{I}_{p,q}\Delta_{\mathbf{X}}^{1/2})$ . ■

The matrix  $\mathbf{Q}_{\mathbf{X}} \in \mathcal{O}(p, q)$  from Fact A.2 has a population analogue,  $\mathbf{Q}_{\xi}$ , which is characterized as follows.

**Fact A.4** ([2, Lemma 2]). *For  $\Delta_{\xi} = \mathbb{E}(\xi\xi^T)$ , consider the spectral decomposition of  $\Delta_{\xi}^{1/2}\mathbb{I}_{p,q}\Delta_{\xi}^{1/2}$ ,*

$$\Delta_{\xi}^{1/2}\mathbb{I}_{p,q}\Delta_{\xi}^{1/2} = \mathbf{W}\Lambda_{\xi}\mathbf{W}^T \quad \text{where } \mathbf{W} \in \mathcal{O}(d).$$

*Then, the population analogue of  $\mathbf{Q}_{\mathbf{X}}$  is the matrix  $\mathbf{Q}_{\xi} \in \mathcal{O}(p, q) \cap \mathcal{O}(d)$  given by*

$$\mathbf{Q}_{\xi} \doteq |\Lambda_{\xi}|^{-1/2}\mathbf{W}^T\Delta_{\xi}^{1/2}. \quad (18)$$

Since  $\mathbf{Q}_{\mathbf{X}}$  and  $\mathbf{Q}_{\xi}$  are determined by the matrices  $\mathbf{V}, \mathbf{W} \in \mathcal{O}(d)$  arising from spectral decompositions, similar to Fact A.1 it follows that  $\mathbf{Q}_{\mathbf{X}}$  and  $\mathbf{Q}_{\xi}$  are also unique only up to orthogonal transformations. The matrix  $\mathbf{O}_{\xi}$  aligns  $\mathbf{Q}_{\mathbf{X}}$  to  $\mathbf{Q}_{\xi}$ .

**Fact A.5.** *The matrix  $\mathbf{O}_{\xi} \in \mathcal{O}(d)$  is the matrix which aligns  $\mathbf{Q}_{\mathbf{X}}$  to  $\mathbf{Q}_{\xi}$  and is given by*

$$\mathbf{O}_{\xi} \doteq \arg \min_{\mathbf{O} \in \mathcal{O}(d)} \left\| \mathbf{Q}_{\mathbf{X}}^{-1}\mathbf{O} - \mathbf{Q}_{\xi}^{-1} \right\|_F^2$$

Finally, the following lemma shows that the matrices  $\mathbf{Q}_{\mathbf{X}}$  and  $\mathbf{Q}_{\xi}$  are invariant to scale transformations.

**Lemma A.1.** *For  $\mathbf{Q}_{\mathbf{X}}, \mathbf{Q}_{\xi}$  as defined in (17) and (18), respectively,  $\mathbf{Q}_{t\mathbf{X}} = \mathbf{Q}_{\mathbf{X}}$  and  $\mathbf{Q}_{t\xi} = \mathbf{Q}_{\xi}$  for all  $t > 0$ .*

*Proof.* For  $t > 0$  and  $\mathbf{Y} := t\mathbf{X}$ ,

$$\Delta_{\mathbf{Y}} = \frac{1}{n}\mathbf{Y}^T\mathbf{Y} = \frac{1}{n}t^2\mathbf{X}^T\mathbf{X} = t^2\Delta_{\mathbf{X}}.$$

Similarly, we also have  $\mathbf{P}_{\mathbf{Y}} = t^2\mathbf{P}_{\mathbf{X}}$ . Therefore,  $\Lambda_{\mathbf{P}_{\mathbf{Y}}} = t^2\Lambda_{\mathbf{P}_{\mathbf{X}}}$ , and, for the spectral decomposition in Fact A.2,  $\mathbf{V}_{\Delta_{\mathbf{Y}}} = \mathbf{V}_{\Delta_{\mathbf{X}}}$ . Plugging these into the expression for  $\mathbf{Q}_{\mathbf{Y}}$ , we get

$$\mathbf{Q}_{\mathbf{Y}} = \left( \frac{1}{n\rho_n} |\Lambda_{\mathbf{P}_{\mathbf{Y}}}| \right)^{-1/2} \mathbf{V}_{\Delta_{\mathbf{Y}}}^T \Delta_{\mathbf{Y}}^{1/2} = \left( \frac{t^2}{n\rho_n} |\Lambda_{\mathbf{P}_{\mathbf{X}}}| \right)^{-1/2} \mathbf{V}_{\Delta_{\mathbf{X}}}^T (t^2\Delta_{\mathbf{X}})^{1/2} = \mathbf{Q}_{\mathbf{X}}.$$

A similar argument also shows that  $\mathbf{Q}_{t\xi} = \mathbf{Q}_{\xi}$ . ■

## B Auxiliary Results

In this section, we collect some results which are used in the proofs presented in Appendix C.

### B.1 Properties of $\mathbf{P}$ and $\Delta$

For  $\varepsilon > 0$  and  $\mathbf{X} \sim_{\text{i.i.d.}} \mathbb{P}$ , let the map  $\varphi_\varepsilon : \mathbb{R}^d \rightarrow \mathbb{R}^{d+1}$  be the map in Theorem 3.1 given by  $\varphi_\varepsilon(\mathbf{x}) = \tau(\varepsilon) \oplus \sigma(\varepsilon)\mathbf{x}$ . Define

$$\mathbf{Y} \doteq \varphi_\varepsilon(\mathbf{X}) \in \mathbb{R}^{n \times (d+1)}, \quad \mathbf{Q} \doteq (\varphi_\varepsilon)_\# \mathbb{P} \quad \text{and} \quad \eta \sim \mathbf{Q}.$$

We establish some properties of  $\mathbf{P}_\mathbf{Y}$  and  $\Delta_\mathbf{Q}$  in relation to  $\mathbf{P}_\mathbf{X}$  and  $\Delta_\mathbf{P}$ , which are used extensively in Appendix C. The following result characterizes the spectral properties of  $\Delta_\mathbf{Q}$ .

**Lemma B.1.** *For  $\varepsilon > 0$ , the following properties hold for  $\Delta_\mathbf{P}$  and  $\Delta_\mathbf{Q}$ .*

1.  $\Delta_\mathbf{P}$  and  $\Delta_\mathbf{Q}$  are both positive definite.
2. If  $\lambda_1(\Delta_\mathbf{P}) \geq \lambda_2(\Delta_\mathbf{P}) \geq \dots \geq \lambda_d(\Delta_\mathbf{P})$  are the eigenvalues of  $\Delta_\mathbf{P}$ , and  $\lambda_1(\Delta_\mathbf{Q}) \geq \lambda_2(\Delta_\mathbf{Q}) \geq \dots \geq \lambda_{d+1}(\Delta_\mathbf{Q})$  are the eigenvalues of  $\Delta_\mathbf{Q}$ , then

$$\lambda_1(\Delta_\mathbf{Q}) \geq \sigma^2(\varepsilon)\lambda_1(\Delta_\mathbf{P}) \geq \dots \geq \sigma(\varepsilon)^2\lambda_d(\Delta_\mathbf{P}) \geq \lambda_{d+1}(\Delta_\mathbf{Q}).$$

3. For sufficiently small  $\varepsilon > 0$ , there exists  $C_1 > 0$  such that

$$\lambda_1(\Delta_\mathbf{Q}) \leq \tau(\varepsilon)^2 + C_1\sigma(\varepsilon)\tau(\varepsilon).$$

4. For sufficiently small  $\varepsilon > 0$ , there exists  $C_2 > 0$  such that

$$\lambda_{d+1}(\Delta_\mathbf{Q}) \geq C_2\sigma^2(\varepsilon).$$

*Proof.* For notational simplicity, throughout the proof we take  $\sigma = \sigma(\varepsilon)$  and  $\tau = \tau(\varepsilon)$ .

*Part 1.* For  $\xi \sim \mathbb{P}$ ,  $\xi\xi^\top$  is positive definite a.e.- $\mathbb{P}$ . To see this, note that for any  $\mathbf{x} \in \mathbb{R}^d$ ,

$$\mathbf{x}^\top (\xi\xi^\top) \mathbf{x} = \|\mathbf{x}^\top \xi\| \geq 0 \quad \text{a.e.} - \mathbb{P}.$$

It follows that  $\mathbb{E}(\mathbf{x}^\top (\xi\xi^\top) \mathbf{x}) = \mathbf{x}^\top \mathbb{E}(\xi\xi^\top) \mathbf{x} = \mathbf{x}^\top \Delta_\mathbf{P} \mathbf{x} \geq 0$ . Since  $\Sigma = \text{Cov}(\xi)$  is assumed to be full-rank, it follows that  $\mathbf{x}^\top \Delta_\mathbf{P} \mathbf{x} > 0$  for all  $\mathbf{x} \in \mathbb{R}^d$ . Therefore,  $\Delta_\mathbf{P}$  is positive definite. Let  $\eta = \phi(\xi) = \tau(\varepsilon) \oplus \sigma(\varepsilon)\xi$ . Because  $\Delta_\mathbf{Q} = \mathbb{E}(\eta\eta^\top)$ , from a similar argument it follows that  $\mathbf{x}^\top \Delta_\mathbf{Q} \mathbf{x} \geq 0$  for all  $\mathbf{x} \in \mathbb{R}^{d+1}$ . It remains to show that  $\mathbf{x}^\top \Delta_\mathbf{Q} \mathbf{x} > 0$ . To this end, note that  $\Delta_\mathbf{Q}$  is the block matrix given by

$$\Delta_\mathbf{Q} = \begin{bmatrix} \tau^2 & \sigma\tau\mathbb{E}(\xi)^\top \\ \sigma\tau\mathbb{E}(\xi) & \sigma^2\Delta_\mathbf{P} \end{bmatrix}. \quad (19)$$

It is easy to verify that the determinant of  $\Delta_\mathbf{Q}$  is given by

$$\det(\Delta_\mathbf{Q}) = \sigma^{2d}\tau^2 \det(\Delta_\mathbf{P}) \left(1 - \mathbb{E}(\xi)^\top \Delta_\mathbf{P}^{-1} \mathbb{E}(\xi)\right).$$

If we can show that  $\det(\Delta_\mathbf{Q}) > 0$ , or, equivalently that  $\mathbb{E}(\xi)^\top \Delta_\mathbf{P}^{-1} \mathbb{E}(\xi) < 1$ , then by Sylvester's criterion [32, Theorem 7.2.5], it will follow that  $\Delta_\mathbf{Q}$  is positive definite. With this in mind, let  $\mathbf{z} = \mathbb{E}(\xi)$ . Note that

$$\Delta_\mathbf{P} = \mathbb{E}(\xi\xi^\top) = \text{Cov}(\xi) + \mathbb{E}(\xi)\mathbb{E}(\xi)^\top = \Sigma + \mathbf{z}\mathbf{z}^\top,$$

where  $\Sigma$  is positive definite. Then, using the Sherman-Morrison-Woodbury formula [32, Section 0.7.2],

$$\begin{aligned} \mathbf{z}^\top \Delta_\mathbf{P}^{-1} \mathbf{z} &= \mathbf{z}^\top \left( \Sigma^{-1} - \frac{\Sigma^{-1} \mathbf{z} \mathbf{z}^\top \Sigma^{-1}}{1 + \mathbf{z}^\top \Sigma^{-1} \mathbf{z}} \right) \mathbf{z} \\ &= \mathbf{z}^\top \Sigma^{-1} \mathbf{z} - \frac{1 + \mathbf{z}^\top \Sigma^{-1} \mathbf{z}}{(\mathbf{z}^\top \Sigma^{-1} \mathbf{z})^2} \end{aligned}$$

$$= \frac{\mathbf{z}^\top \Sigma^{-1} \mathbf{z}}{1 + \mathbf{z}^\top \Sigma^{-1} \mathbf{z}} < 1.$$

This implies that  $1 - \mathbf{z}^\top \Delta_{\mathbf{P}}^{-1} \mathbf{z} > 0$ , and, therefore,  $\Delta_{\mathbf{Q}}$  is full-rank and positive definite.

*Part 2.* Using Cauchy's interlacing theorem Horn and Johnson [32, Theorem 4.3.17] for the block-matrix representation of  $\Delta_{\mathbf{Q}}$  in Eq. (19), we obtain

$$\lambda_1(\Delta_{\mathbf{Q}}) \geq \lambda_1(\sigma^2 \Delta_{\mathbf{P}}) \geq \lambda_2(\Delta_{\mathbf{Q}}) \geq \lambda_2(\sigma^2 \Delta_{\mathbf{P}}) \geq \dots \geq \lambda_d(\sigma^2 \Delta_{\mathbf{P}}) \geq \lambda_{d+1}(\Delta_{\mathbf{Q}}).$$

By noting that  $\lambda_i(\sigma^2 \Delta_{\mathbf{P}}) = \sigma^2 \lambda_i(\Delta_{\mathbf{P}})$ , the result for (2) follows.

*Part 3.* The Geršgorin disk theorem [32, Theorem 6.1.1] for  $\Delta_{\mathbf{Q}}$  asserts that for all  $1 \leq k \leq d+1$ , the collection of eigenvalues of  $\Delta_{\mathbf{Q}}$  satisfy

$$\lambda_k(\Delta_{\mathbf{Q}}) \in \bigcup_{1 \leq i \leq d+1} \left[ (\Delta_{\mathbf{Q}})_{ii} - \sum_{j \neq i} (\Delta_{\mathbf{Q}})_{ij}, (\Delta_{\mathbf{Q}})_{ii} + \sum_{j \neq i} (\Delta_{\mathbf{Q}})_{ij} \right].$$

This implies that  $\lambda_{\max}(\Delta_{\mathbf{Q}}) = \lambda_1(\Delta_{\mathbf{Q}}) \leq \tau^2 + \sigma\tau \sum_{1 \leq i \leq d} \mathbb{E}(\xi)$ . Taking  $C_1 = \mathbf{1}_d^\top \mathbb{E}(\xi)$ , the result follows.

*Part 4.* Taking  $\omega = 1 - \mathbb{E}(\xi)^\top \Delta_{\mathbf{P}}^{-1} \mathbb{E}(\xi) > 0$  and using the fact that  $\det(\Delta_{\mathbf{Q}}) = \prod_{1 \leq i \leq d+1} \lambda_i(\Delta_{\mathbf{Q}})$ , we obtain

$$\begin{aligned} \lambda_{d+1}(\Delta_{\mathbf{Q}}) &= \frac{\det(\Delta_{\mathbf{Q}})}{\lambda_1(\Delta_{\mathbf{Q}}) \times \prod_{j=2}^d \lambda_j(\Delta_{\mathbf{Q}})} \\ &\stackrel{(i)}{=} \frac{\sigma^{2d} \tau^2 \det(\Delta_{\mathbf{P}}) (1 - \mathbb{E}(\xi)^\top \Delta_{\mathbf{P}}^{-1} \mathbb{E}(\xi))}{\lambda_1(\Delta_{\mathbf{Q}}) \times \prod_{j=2}^d \lambda_j(\Delta_{\mathbf{Q}})} \\ &= \sigma^{2d} \tau^2 \omega \frac{\prod_{j=1}^d \lambda_j(\Delta_{\mathbf{P}})}{\lambda_1(\Delta_{\mathbf{Q}}) \times \prod_{j=2}^d \lambda_j(\Delta_{\mathbf{Q}})} \\ &= \frac{\lambda_d(\Delta_{\mathbf{P}}) \tau^2 \omega}{\lambda_1(\Delta_{\mathbf{Q}})} \times \sigma^2 \times \left( \prod_{j=1}^{d-1} \sigma^2 \lambda_j(\Delta_{\mathbf{P}}) \middle/ \prod_{j=2}^d \lambda_j(\Delta_{\mathbf{Q}}) \right) \\ &\stackrel{(ii)}{\geq} \left( \frac{\lambda_d(\Delta_{\mathbf{P}}) \tau^2 \omega}{\tau^2 + \sigma\tau C_1} \right) \sigma^2 \\ &\stackrel{(iii)}{\geq} \sigma^2 C_2, \end{aligned}$$

where (i) follows from the definition of  $\det(\Delta_{\mathbf{Q}})$  from Eq. (19), (ii) uses the interlacing property of the eigenvalues in part 2, i.e.  $\lambda_j(\Delta_{\mathbf{Q}}) \leq \sigma^2 \lambda_{j-1}(\Delta_{\mathbf{P}})$  for  $j \in \{2 \dots d\}$ , and (iii) follows from taking  $C_2 = \lambda_d(\Delta_{\mathbf{P}}) \tau^2 \omega / (\tau^2 + \sigma\tau C_1) > 0$ . ■

The next lemma characterizes the spectral properties of  $\mathbf{P}_{\mathbf{Y}}$ .

**Lemma B.2.** *For  $\varepsilon > 0$ , the following properties hold for  $\mathbf{P}_{\mathbf{X}}$  and  $\mathbf{P}_{\mathbf{Y}}$ .*

1. *Up to a collection of repeated zero eigenvalues,*

$$\Lambda(\mathbf{P}_{\mathbf{Y}}) = \Lambda(\mathbf{Y}^\top \mathbf{Y} \mathbb{I}_{p+1,q}) = \Lambda \left( \left( \mathbf{Y}^\top \mathbf{Y} \right)^{-1/2} \mathbb{I}_{p+1,q} \left( \mathbf{Y}^\top \mathbf{Y} \right)^{-1/2} \right).$$

2. *With probability greater than  $1 - 2(d+1)/n$ ,*

$$\left\| \mathbf{Y}^\top \mathbf{Y} \mathbb{I}_{p+1,q} - n \Delta_{\mathbf{Q}} \mathbb{I}_{p+1,q} \right\| \leq C \sqrt{n \log n}.$$

3.  $|\lambda_{d+1}(\mathbf{P}_{\mathbf{Y}})| = \Omega_{\mathbf{P}} \left( n \sigma^2(\varepsilon) \right).$

*Proof.* As before, for notational simplicity throughout the proof we take  $\sigma = \sigma(\varepsilon)$  and  $\tau = \tau(\varepsilon)$ .

*Part 1.* The first claim follows from the same argument as in Fact A.3.

*Part 2.* First, we note that  $\mathbf{Y}^\top \mathbf{Y}$  can be written as the sum of iid random matrices

$$\mathbf{Y}^\top \mathbf{Y} = \sum_{i=1}^n \mathbf{Y}_i^\top \mathbf{Y}_i \in \mathbb{R}^{d+1 \times d+1},$$

with  $\mathbb{E}(\mathbf{Y}_i^\top \mathbf{Y}_i) = \Delta_Q$ . Since  $\mathbb{P}$  has compact (and, therefore, bounded) support, each  $\mathbf{X}_i \sim \mathbb{P}$  has  $\|\mathbf{X}_i\| \leq \text{diam}(\mathbb{X}) = L < \infty$  a.e.- $\mathbb{P}$ . Consequently,  $\|\mathbf{Y}_i\| = \sqrt{\mathbf{Y}_i^\top \mathbf{Y}_i} = \|\tau \oplus \sigma \mathbf{X}_i\| \leq \sqrt{\tau^2 + L^2} = M < \infty$  a.e. For  $t > 0$ , we can use the matrix Bernstein inequality [59, Theorem 6.1.1] to obtain the tail bound

$$\mathbb{P}\left(\left\|\mathbf{Y}^\top \mathbf{Y} - n\Delta_Q\right\| > t\right) \leq 2d \exp\left\{\frac{-t^2}{2v(\mathbf{Y}^\top \mathbf{Y}) + \frac{2}{3}Mt}\right\}, \quad (20)$$

where  $v(\mathbf{Y}^\top \mathbf{Y}) = n\|\mathbb{E}(\mathbf{Y}_i^\top \mathbf{Y}_i \mathbf{Y}_i^\top \mathbf{Y}_i)\| \leq n \max\{\tau^2 + \sigma\tau C_1, M^2\} = nk$  for  $k = \max\{\tau^2 + \sigma\tau C_1, M^2\}$ . This follows from the fact that  $E(\mathbf{Y}_i^\top \mathbf{Y}_i) = \Delta_Q$  is bounded in spectral norm from Lemma B.1(3), and  $E(\mathbf{Y}_i \mathbf{Y}_i^\top) \leq \|\mathbf{Y}_i\|^2 \leq M^2$ . For  $\delta > 0$ , Eq. (20) is equivalent to

$$\mathbb{P}\left(\left\|\mathbf{Y}^\top \mathbf{Y} - n\Delta_Q\right\| \leq \sqrt{2k\delta n} + \frac{2M}{3}\delta\right) \geq 1 - 2d e^{-\delta}.$$

Taking  $\delta = \log n$  and  $C$  sufficiently large, it follows that

$$\mathbb{P}\left(\left\|\mathbf{Y}^\top \mathbf{Y} - n\Delta_Q\right\| \leq C\sqrt{n \log n}\right) \geq 1 - \frac{2(d+1)}{n}.$$

By noting that

$$\left\|\mathbf{Y}^\top \mathbf{Y} \mathbb{I}_{p+1,q} - n\Delta_Q \mathbb{I}_{p,q}\right\| \leq \left\|\mathbf{Y}^\top \mathbf{Y} - n\Delta_Q\right\| \cdot \|\mathbb{I}_{p+1,q}\| = \left\|\mathbf{Y}^\top \mathbf{Y} - n\Delta_Q\right\|,$$

we obtain the desired result, i.e. with probability greater than  $1 - 2(d+1)n^{-1}$

$$\left\|\mathbf{Y}^\top \mathbf{Y} \mathbb{I}_{p+1,q} - n\Delta_Q \mathbb{I}_{p,q}\right\| \leq C\sqrt{n \log n}. \quad (21)$$

*Part 3.* We begin by noting that

$$\begin{aligned} \left| \lambda_i(\mathbf{Y}^\top \mathbf{Y} \mathbb{I}_{p+1,q}) - \lambda_i(n\Delta_Q \mathbb{I}_{p+1,q}) \right| &\stackrel{(i)}{\leq} \left| \lambda_i(\mathbf{Y}^\top \mathbf{Y} \mathbb{I}_{p+1,q}) - \lambda_i(n\Delta_Q \mathbb{I}_{p+1,q}) \right| \\ &\stackrel{(ii)}{\leq} \left\| \mathbf{Y}^\top \mathbf{Y} \mathbb{I}_{p+1,q} - n\Delta_Q \mathbb{I}_{p+1,q} \right\|, \end{aligned} \quad (22)$$

where (i) follows from the reverse triangle inequality, and (ii) is a consequence of Weyl's perturbation theorem [32, Theorem 4.3.1]. Therefore, the eigenvalues of  $\mathbf{Y}^\top \mathbf{Y} \mathbb{I}_{p+1,q}$  can be controlled by the eigenvalues of  $n\Delta_Q \mathbb{I}_{p+1,q}$ . We now obtain a lower bound on  $\lambda_i(n\Delta_Q \mathbb{I}_{p+1,q}) = n\lambda_i(\Delta_Q \mathbb{I}_{p+1,q})$ .

From *Part (1)* we have that  $\lambda_i(\Delta_Q \mathbb{I}_{p+1,q}) = \lambda_i(\mathbb{I}_{p+1,q} \Delta_Q)$ . Furthermore, the singular-values of  $\mathbb{I}_{p+1,q} \Delta_Q$  are given by

$$\varsigma_i(\mathbb{I}_{p+1,q} \Delta_Q) = \sqrt{\lambda_i\left((\mathbb{I}_{p+1,q} \Delta_Q)^\top \mathbb{I}_{p+1,q} \Delta_Q\right)} = \sqrt{\lambda_i(\Delta_Q^2)} = |\lambda_i(\Delta_Q)|.$$

From Horn and Johnson [32, Theorem 5.6.9], it follows that

$$|\lambda_{\min}(\Delta_Q \mathbb{I}_{p+1,q})| = |\lambda_{\min}(\mathbb{I}_{p+1,q} \Delta_Q)| \geq \varsigma_{\min}(\mathbb{I}_{p+1,q} \Delta_Q) = |\lambda_{\min}(\Delta_Q)| \geq \sigma^2 C_2,$$

where the last inequality follows from Lemma B.1. Therefore, we obtain the lower bound

$$|\lambda_{\min}(n\Delta_Q \mathbb{I}_{p+1,q})| \geq n\sigma^2 C_2. \quad (23)$$

Similarly, we can also find an upper bound for  $|\lambda_{\max}(\Delta_Q \mathbb{I}_{p+1,q})|$  using Horn and Johnson [32, Theorem 5.6.9] by observing that

$$|\lambda_{\max}(\Delta_Q \mathbb{I}_{p+1,q})| \leq \varsigma_{\max}(\Delta_Q \mathbb{I}_{p+1,q}) \leq |\lambda_{\max}(\Delta_Q)| = \tau^2 + \sigma\tau C_1.$$

This implies that  $|\lambda_{\max}(n\Delta_Q \mathbb{I}_{p+1,q})| = \Theta(n)$ . Moreover, combining Eq. (22) with Eq. (21) we obtain

$$\mathbb{P}\left(\left|\lambda_{d+1}(\mathbf{Y}^\top \mathbf{Y} \mathbb{I}_{p+1,q})\right| > |\lambda_{d+1}(n\Delta_Q \mathbb{I}_{p+1,q})| - C\sqrt{n \log n}\right) \geq 1 - \frac{2(d+1)}{n},$$

which, when combined with Eq. (23) yields

$$\mathbb{P}\left(\left|\lambda_{d+1}(\mathbf{Y}^\top \mathbf{Y} \mathbb{I}_{p+1,q})\right| > n\sigma^2 C_2 - C\sqrt{n \log n}\right) \geq 1 - \frac{2(d+1)}{n}.$$

When  $\sigma > \frac{2C}{C_2} \sqrt{\log n/n}$ , or, alternatively, when  $\sigma = \omega\left(\sqrt{\log n/n}\right)$ , the lower bound is non-trivial. Therefore, when  $\sigma = \omega\left(\sqrt{\log n/n}\right)$  it follows that

$$\mathbb{P}\left(\left|\lambda_{d+1}(\mathbf{Y}^\top \mathbf{Y} \mathbb{I}_{p+1,q})\right| > \frac{C_2}{2} n\sigma^2\right) \geq 1 - \frac{2(d+1)}{n}.$$

The result follows by noting that this is equivalent to the statement that  $|\lambda_{d+1}(\mathbf{P}_\mathbf{Y})| = \Omega_P(n\sigma^2)$ .  $\blacksquare$

We collect the main findings from Lemma B.1 and Lemma B.2 in the following corollary.

**Corollary B.1.** *For  $\mathbf{P}_\mathbf{Y}$ ,  $\Delta_Q$  and their associated eigenvalues  $\lambda_i(\mathbf{P}_\mathbf{Y})$  and  $\lambda_i(\Delta_Q)$ :*

1.  $\lambda_{\min}(\Delta_Q) = \Omega(\sigma(\varepsilon)^2)$  and  $\lambda_{\max}(\Delta_Q) = O(1)$ .
2.  $\lambda_{\min}(\mathbf{P}_\mathbf{Y}) = \Omega(n\sigma(\varepsilon)^2)$  and  $\lambda_{\max}(\mathbf{P}_\mathbf{Y}) = \Theta(\lambda_{\max}(\mathbf{P}_\mathbf{X}))$ .

For  $\mathcal{M}_\varepsilon(\mathbf{A}) =: \mathbf{A}_\mathbf{Y} \sim \mathcal{G}(\mathbf{Y}; p+1, q)$ , next lemma establishes a tail bound for  $\|\mathbf{A}_\mathbf{Y} - \mathbf{P}_\mathbf{Y}\|$  using a straightforward application of Lei and Rinaldo [42, Theorem 5.2].

**Lemma B.3.** *Under the conditions of Theorem B.1, there exists a constant  $C > 0$  such that with probability greater than  $1 - 1/n$ ,*

$$\|\mathbf{A}_\mathbf{Y} - \mathbf{P}_\mathbf{Y}\| \leq C\sqrt{n\gamma},$$

where  $\gamma = \sigma^2(\varepsilon) \max_{ij}(\mathbf{P}_\mathbf{X})_{ij} + \tau^2(\varepsilon) \leq 1$ .

*Proof.* We may directly apply Lei and Rinaldo [42, Theorem 5.2] to obtain the claim provided we verify the following two conditions: (i)  $\mathbb{E}(\mathcal{M}_\varepsilon(\mathbf{A})) = \mathbf{P}_\mathbf{Y}$ , and (ii)  $n \max_{i,j}(\mathbf{P}_\mathbf{Y})_{ij} < n\gamma$ . Note that the additional requirement that  $n\gamma > n\tau(\varepsilon)^2 > c_0 \log n$  for a fixed  $c_0 > 0$  is trivially satisfied. To verify the first condition, for  $1 \leq i < j \leq n$  we have that

$$\begin{aligned} \mathbb{E}\left(\mathcal{M}_\varepsilon(\mathbf{A}_{ij})\right) &\stackrel{(i)}{=} \mathbb{P}\left(\mathcal{M}_\varepsilon(\mathbf{A}_{ij}) = 1\right) \\ &\stackrel{(ii)}{=} \mathbb{P}\left(\mathcal{M}_\varepsilon(\mathbf{A}_{ij}) = 1 \mid \mathbf{A}_{ij=1}\right) \mathbb{P}(\mathbf{A}_{ij} = 1) + \mathbb{P}\left(\mathcal{M}_\varepsilon(\mathbf{A}_{ij}) = 1 \mid \mathbf{A}_{ij=0}\right) \mathbb{P}(\mathbf{A}_{ij} = 0) \\ &= (1 - \pi(\varepsilon)) \mathbf{X}_i^\top \mathbb{I}_{p,q} \mathbf{X}_j + \pi(\varepsilon)(1 - \mathbf{X}_i^\top \mathbb{I}_{p,q} \mathbf{X}_j) \\ &= \sigma^2(\varepsilon) \mathbf{X}_i^\top \mathbb{I}_{p,q} \mathbf{X}_j + \tau^2(\varepsilon) = (\mathbf{P}_\mathbf{Y})_{ij}, \end{aligned}$$

where (i) uses the fact that  $\mathcal{M}_\varepsilon(\mathbf{A}_{ij})$  is from a Bernoulli distribution, and (ii) uses the definition of `edgeFlip`. For the second condition, note that

$$\max_{i,j \in [n]} (\mathbf{P}_\mathbf{Y})_{ij} = \max_{i,j \in [n]} \sigma^2(\varepsilon) \mathbf{X}_i^\top \mathbb{I}_{p,q} \mathbf{X}_j + \tau^2(\varepsilon) = \sigma^2(\varepsilon) \left( \max_{i,j \in [n]} \mathbf{X}_i^\top \mathbb{I}_{p,q} \mathbf{X}_j \right) + \tau^2(\varepsilon) = \gamma \leq 1.$$

Therefore, the claim that  $\|\mathbf{A}_\mathbf{Y} - \mathbf{P}_\mathbf{Y}\| = O_p(\sqrt{n})$  follows from Lei and Rinaldo [42, Theorem 5.2].  $\blacksquare$

### B.1.1 Upper bounds for GRDPG Estimation

In this section, recall some key results from [2, 57, 63] which are used in the proofs of our main results. The alignment matrices appearing here are described in Appendix A.

Recall that for  $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{n \times d}$ , the  $\ell_{2 \rightarrow \infty}$  metric between  $\mathbf{X}$  and  $\mathbf{Y}$  which reproduces the non-identifiability of the GRDPG is given by

$$d_{2,\infty}(\mathbf{X}, \mathbf{Y}) = \inf_{\mathbf{O} \in \mathcal{O}(p,q) \cap \mathcal{O}(d)} \left\| \mathbf{Y} \mathbf{Q}_{\mathbf{Y}}^{-1} - \mathbf{X} \mathbf{Q}_{\mathbf{X}}^{-1} \mathbf{O} \right\|_{2,\infty}. \quad (24)$$

where the matrices  $\mathbf{Q}_{\mathbf{X}}, \mathbf{Q}_{\mathbf{Y}}$  are the alignment matrices for  $\mathbf{X}$  and  $\mathbf{Y}$ , respectively, as defined in Fact A.2. Equivalently,

$$d_{2,\infty}(\mathbf{X}, \mathbf{Y}) = \inf_{\mathbf{O} \in \mathcal{O}(p,q) \cap \mathcal{O}(d)} \left\| \mathbf{U}_{\mathbf{Y}} |\mathbf{\Lambda}_{\mathbf{Y}}|^{1/2} - \mathbf{U}_{\mathbf{X}} |\mathbf{\Lambda}_{\mathbf{X}}|^{1/2} \mathbf{O} \right\|_{2,\infty}.$$

The following result from [2] is a restatement of the main result from [57] for estimating the latent positions of GRDPGs under the  $d_{2,\infty}$  metric.

**Proposition B.1** (Theorem 4 of [2]). *Given  $(p, q)$ -admissible  $\mathbb{P}$ , let  $(\mathbf{A}, \mathbf{X}) \sim \mathcal{G}(\mathbb{P}; p, q, \rho_n)$ , and let  $\widehat{\mathbf{X}} = \text{sp}(\mathbf{A}; d) \in \mathbb{R}^{n \times d}$  be the adjacency spectral embedding of  $\mathbf{A}$ . If  $n\rho_n = \omega(\log^4 n)$ , then with probability  $1 - O(n^{-2})$ ,*

$$d_{2,\infty}(\widehat{\mathbf{X}}(\mathbf{A}), \rho_n^{1/2} \mathbf{X}) = O\left(\frac{\log n}{\sqrt{n}}\right).$$

Equivalently, for estimating  $\mathbf{X}$ , it follows that with probability  $1 - O(n^{-2})$ ,

$$d_{2,\infty}(\rho_n^{-1/2} \widehat{\mathbf{X}}(\mathbf{A}), \mathbf{X}) = O\left(\frac{\log n}{\sqrt{n\rho_n}}\right).$$

In addition to the above result, we also need the following key property of the matrices  $\mathbf{Q}_{\mathbf{X}}$  and its population analogue  $\mathbf{Q}_{\xi}$ . The expressions for  $\mathbf{Q}_{\xi}, \mathbf{Q}_{\xi}$  are given in Fact A.2 and Fact A.5, respectively. The following result is an improvement of [58, Lemma 16] using the matrix concentration result from [2, Lemma 2].

**Lemma B.4.** *For  $\Delta \equiv \Delta_{\mathbb{P}}$  and  $\mathbf{X} \sim_{\text{i.i.d.}} \mathbb{P}$ , let  $\mathbf{Q}_{\mathbf{X}} \in \mathcal{O}(p, q)$  and  $\mathbf{Q}_{\xi} \in \mathcal{O}(p, q) \cap \mathcal{O}(d)$  be the alignment matrices described in Fact A.2 and Fact A.4, respectively. Then, there exists a block orthogonal matrix  $\mathbf{O} \in \mathcal{O}(d) \cap \mathcal{O}(p, q)$  such that*

$$\left\| \mathbf{Q}_{\mathbf{X}}^{-1} - \mathbf{Q}_{\xi}^{-1} \mathbf{O} \right\| = O_p\left(\sqrt{\frac{\log n}{n}}\right).$$

Moreover, by Lemma A.1, the result holds true for arbitrary scale transformations  $t\mathbf{X}$  and  $t\xi$  for  $t > 0$ .

## C Proofs

In this section, we present the proofs for the main results in Section 3. We refer the reader to Table 1 for a summary of notations used in this section. The proofs also make use of the supporting results in Appendix B.

Throughout the proofs, for  $\varepsilon > 0$  we use  $\sigma \doteq \sigma(\varepsilon)$  and  $\tau \doteq \tau(\varepsilon)$  for brevity. For the map  $\varphi \doteq \varphi_{\varepsilon} : \mathbb{R}^d \rightarrow \mathbb{R}^{d+1}$  given by  $\varphi_{\varepsilon}(\mathbf{x}) = \tau \oplus \sigma \rho_n^{1/2} \mathbf{x}$ , we define  $\mathbf{Y} = \varphi(\mathbf{X})$  and  $\mathbb{Q} = (\varphi)_{\#} \mathbb{P}$  for the pushforward measure, and  $\eta \sim \mathbb{Q}$  to denote an arbitrary random variable distributed according to  $\mathbb{Q}$ .

### C.1 Proof of Theorem 3.1

For  $(\mathbf{A}, \mathbf{X}) \sim \mathcal{G}(\mathbb{P}, \rho_n; p, q)$ , let  $\mathbf{A}^{\circ} = \mathcal{M}_{\varepsilon}(\mathbf{A})$  and  $\mathbf{A}^{\bullet} \sim \mathcal{G}(\varphi_{\#} \mathbb{P}, 1; p+1, q)$ . Since each entry  $\mathbf{A}_{ij}^{\circ}, \mathbf{A}_{ij}^{\bullet} \in \{0, 1\}$  are independent, it suffices to show that

$$\mathbf{Y} \sim_{\text{i.i.d.}} \mathbb{Q} \quad \text{and} \quad \mathbf{A}_{ij}^{\circ} \stackrel{d}{=} \mathbf{A}_{ij}^{\bullet} \quad \text{for all } 1 \leq i < j \leq n.$$



For the first claim, note that for measurable sets  $A_1 \cdots A_n \subset \mathbb{R}^{d+1}$ ,

$$\mathbb{Q}(Y_1 \in A_1, \dots, Y_n \in A_n) = \mathbb{P}(X_1 \in \varphi^{-1}(A_1), \dots, X_n \in \varphi^{-1}(A_n)) = \prod_{i=1}^n \mathbb{P}(\varphi^{-1}(A_i)) = \prod_{i=1}^n \mathbb{Q}(A_i),$$

where the third equality follows from the fact that  $X_i$  are iid and the final equality follows from the definition of the pushforward  $\varphi_{\#}\mathbb{P}$ .

For the second claim, from Definition 2.3, observe that

$$\begin{aligned} \mathbb{P}(\mathbf{A}_{ij}^{\circ} = 1 \mid \mathbf{X}) &= \mathbb{P}\left(\{\mathcal{F}(\mathbf{A}_{ij}) = 1 - \mathbf{A}_{ij}\} \cap \{\mathbf{A}_{ij} = 0\} \mid \mathbf{X}\right) + \mathbb{P}\left(\{\mathcal{F}(\mathbf{A}_{ij}) = \mathbf{A}_{ij}\} \cap \{\mathbf{A}_{ij} = 1\} \mid \mathbf{X}\right) \\ &= \pi(\varepsilon)\mathbb{P}(\mathbf{A}_{ij} = 0 \mid \mathbf{X}) + (1 - \pi(\varepsilon))\mathbb{P}(\mathbf{A}_{ij} = 1 \mid \mathbf{X}) \\ &\stackrel{(i)}{=} \pi(\varepsilon)\left(1 - \rho_n X_i^{\top} \mathbb{I}_{p,q} X_j\right) + (1 - \pi(\varepsilon))\rho_n X_i^{\top} \mathbb{I}_{p,q} X_j \\ &= \pi(\varepsilon) + (1 - 2\pi(\varepsilon))\rho_n X_i^{\top} \mathbb{I}_{p,q} X_j \\ &= \tau^2(\varepsilon) + \sigma^2(\varepsilon)\rho_n X_i^{\top} \mathbb{I}_{p,q} X_j \\ &\stackrel{(ii)}{=} \begin{pmatrix} \tau(\varepsilon) & \sigma(\varepsilon)\rho_n^{1/2} X_i^{\top} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \mathbb{I}_{p,q} \end{pmatrix} \begin{pmatrix} \tau(\varepsilon) \\ \sigma(\varepsilon)\rho_n^{1/2} X_j \end{pmatrix} \\ &= Y_i \mathbb{I}_{p+1,q} Y_j \\ &= \mathbb{P}(\mathbf{A}_{ij}^{\bullet} = 1 \mid \mathbf{Y}), \end{aligned}$$

where (i) follows from the fact that  $\mathbf{A} \sim \mathcal{G}(\mathbb{P}, \rho_n; p, q)$ , and (ii) follows from the inner-product in  $\mathbb{R}^{d+1}$  w.r.t. the indefinite identity matrix  $\mathbb{I}_{p+1,q}$  with signature  $(p+1, q)$ . ■

## C.2 Composition of edgeFlips and Proposition 3.1

Given  $(\mathbf{A}, \mathbf{X}) \sim \mathcal{G}(\mathbb{P}, \rho_n; p, q)$ , let  $\mathcal{M}_{\varepsilon_2}(\mathcal{M}_{\varepsilon_1}(\mathbf{A}))$  be edge-DP graph obtained after applying `edgeFlip` successively with privacy budgets  $\varepsilon_1, \varepsilon_2 > 0$ . Let  $\tau_i = \tau(\varepsilon_i)$ ,  $\sigma_i = \sigma(\varepsilon_i)$  and  $\pi_i = \pi(\varepsilon_i)$  for  $i = 1, 2$ .

On the one hand, taking  $\mathbf{A}^{(1)} = \mathcal{M}_{\varepsilon_1}(\mathbf{A})$  and  $\mathbf{A}^{(2)} = \mathcal{M}_{\varepsilon_2}(\mathbf{A}^{(1)})$ , from Theorem 3.1 it follows that

$$(\mathbf{A}^{(1)}, \varphi_1(\mathbf{X})) \sim \mathcal{G}(\varphi_{1\#}\mathbb{P}, 1; p+1, q) \quad \text{and} \quad (\mathbf{A}^{(2)}, \varphi_2 \circ \varphi_1(\mathbf{X})) \sim \mathcal{G}((\varphi_2 \circ \varphi_1)_{\#}\mathbb{P}, 1; p+2, q),$$

where  $\varphi_1 : \mathbb{R}^d \rightarrow \mathbb{R}^{d+1}$  and  $\varphi_2 : \mathbb{R}^{d+1} \rightarrow \mathbb{R}^{d+2}$  are given by  $\varphi_1(\mathbf{x}) = \tau_1 \oplus \sigma_1 \mathbf{x}$  and  $\varphi_2(\mathbf{y}) = \tau_2 \oplus \sigma_2 \mathbf{y}$ . Specifically, for  $\varphi = \varphi_2 \circ \varphi_1$  given by

$$\varphi(\mathbf{x}) = \left( \tau_2, \sigma_2 \tau_1, \sigma_2 \sigma_1 \mathbf{x}^{\top} \right)^{\top} \in \mathbb{R}^{d+2},$$

the latent positions of  $\mathbf{A}^{(2)}$  are  $\varphi(\mathbf{X}) \in \mathbb{R}^{n \times (d+2)}$ . On the other hand, we have

$$\begin{aligned} \mathbb{P}(\mathbf{A}_{ij}^{(2)} = 1) &= (1 - \pi_2) \cdot \mathbb{P}(\mathbf{A}_{ij}^{(1)} = 1) \\ &\quad + \pi_2 \cdot \mathbb{P}(\mathbf{A}_{ij}^{(1)} = 0) \\ &= (1 - \pi_2) \cdot \left\{ (1 - \pi_1) \cdot \mathbb{P}(\mathbf{A}_{ij} = 1) + \pi_1 \cdot \mathbb{P}(\mathbf{A}_{ij} = 0) \right\} \\ &\quad + \pi_2 \cdot \left\{ \pi_1 \cdot \mathbb{P}(\mathbf{A}_{ij} = 1) + (1 - \pi_1) \cdot \mathbb{P}(\mathbf{A}_{ij} = 0) \right\} \\ &= \left( 1 - \pi_1 - \pi_2 + 2\pi_1\pi_2 \right) \cdot \mathbb{P}(\mathbf{A}_{ij} = 1) \end{aligned}$$

$$\begin{aligned}
 & + \left( \pi_1 + \pi_2 - 2\pi_1\pi_2 \right) \cdot \mathbb{P}(\mathbf{A}_{ij} = 0) \\
 & = (1 - \pi') \cdot \mathbb{P}(\mathbf{A}_{ij} = 1) + \pi' \cdot \mathbb{P}(\mathbf{A}_{ij} = 0),
 \end{aligned}$$

where  $\pi' \doteq (\pi_1 + \pi_2 - 2\pi_1\pi_2)$ . In other words, using Eq. (2), for  $\varepsilon'$  given by

$$\varepsilon' = \log \left( \frac{1}{\pi_1 + \pi_2 - 2\pi_1\pi_2} - 1 \right),$$

it follows that  $\mathcal{M}_{\varepsilon'}(\mathbf{A}) \stackrel{d}{=} \mathcal{M}_{\varepsilon_2} \circ \mathcal{M}_{\varepsilon_1}(\mathbf{A})$ , i.e.,  $\mathbf{A}^{(2)}$  is equivalently the graph obtained by applying `edgeFlip` to  $\mathbf{A}$  with privacy budget  $\varepsilon'$ . Using Theorem 3.1 to this characterization would imply that the latent positions of  $\mathbf{A}$  are given by  $\varphi'(\mathbf{X}) \in \mathbb{R}^{d+1}$  where

$$\varphi'(\mathbf{x}) = \tau(\varepsilon') \oplus \sigma(\varepsilon')\mathbf{x}.$$

The following result is a more general version of Proposition 3.1, and shows that the two characterizations above are equivalent. In particular, only one extra dimension is needed to faithfully encode the latent positions of the graph after several applications of `edgeFlip`.

**Proposition C.1.** *Under the same setting as Theorem 3.1, for  $\varepsilon_1 \cdots \varepsilon_m > 0$ , let  $\mathbf{A}^{(m)} = \mathcal{M}_{\varepsilon_1} \circ \cdots \circ \mathcal{M}_{\varepsilon_m}(\mathbf{A})$  be the graph obtained after applying `edgeFlip`  $m$  times. Let  $\varphi_{\varepsilon_i} : \mathbb{R}^{m+i-1} \rightarrow \mathbb{R}^{m+i}$  be given by  $\varphi_{\varepsilon_i}(\mathbf{y}) = \tau(\varepsilon_i) \oplus \sigma(\varepsilon_i)\mathbf{y}$  and let*

$$\varphi^{(m)} = \varphi_{\varepsilon_1} \circ \varphi_{\varepsilon_2} \circ \cdots \circ \varphi_{\varepsilon_m} : \mathbb{R}^d \rightarrow \mathbb{R}^{d+m}.$$

Then, there exists  $\mathbf{Q} \in \mathbb{O}(p+m, q)$  and  $a, b > 0$  such that

$$\mathbf{Q}\varphi^{(m)}(\mathbf{x}) = (\mathbf{0}_{m-1}, a, b\mathbf{x}^\top)^\top \quad \forall \mathbf{x} \in \mathbb{R}^d,$$

and for  $\psi : \mathbb{R}^d \rightarrow \mathbb{R}^{d+1}$  given by  $\psi(\mathbf{x}) = (a, b\mathbf{x}^\top)^\top$ , it follows that

$$(\mathbf{A}^{(m)}, \psi(\mathbf{X})) \sim \mathcal{G}(\psi_\# \mathbb{P}, 1; p+1, q).$$

*Proof of Proposition C.1.* Let  $\varphi_i = \varphi_{\varepsilon_i}$ ,  $\tau_i = \tau(\varepsilon_i)$  and  $\sigma_i = \sigma(\varepsilon_i)$  for all  $i \in [m]$ . The proof follows by induction on  $m$ . For  $m = 1$ , the result holds trivially from Theorem 3.1. By induction hypothesis, suppose the claim holds for  $\mathbf{A}^{(m-1)}$ ,  $m \geq 2$ , i.e., there exists  $\mathbf{W} \in \mathbb{O}(p+m-1, q)$  and  $\alpha, \beta > 0$  such that

$$\mathbf{W}\varphi^{(m-1)}(\mathbf{x}) = (\mathbf{0}_{m-2}, \alpha, \beta\mathbf{x}^\top)^\top \in \mathbb{R}^{d+m-1} \quad \text{for all } \mathbf{x} \in \mathbb{R}^d.$$

For  $\mathbf{x} \in \mathbb{R}^d$ , let  $\mathbf{y} = \varphi^{(m-1)}(\mathbf{x}) = (y_1, y_2, \dots, y_{m-1}, \gamma\mathbf{x}^\top)^\top \in \mathbb{R}^{d+m-1}$  where<sup>1</sup>  $\gamma > 0$ . By definition of  $\varphi_m$ , we have

$$\varphi_m(\mathbf{y}) = \left( \tau_m, \sigma_m \mathbf{y}^\top \right)^\top = \left( \tau_m, \sigma_m y_1, \sigma_m \mathbf{y}_{2:m-1}, \sigma_m \gamma \mathbf{x}^\top \right)^\top.$$

Let  $R \in \mathbb{O}(2)$  be the rotation matrix such that  $R(\tau_m, \sigma_m y_1)^\top = (0, \delta y_1)^\top$ , i.e.,

$$R = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad \text{for } \theta = \arctan \left( \frac{\tau_m}{\sigma_m y_1} \right),$$

and let  $\mathbf{Q}_1$  be the block-orthogonal matrix given by

$$\mathbf{Q}_1 = \begin{pmatrix} R & \mathbf{0}_{2 \times (d+m-2)} \\ \mathbf{0}_{(d+m-2) \times 2} & \mathbb{I}_{d+m-2} \end{pmatrix} \in \mathbb{R}^{(d+m) \times (d+m)}.$$

By construction,  $\mathbf{Q}_1 \varphi_m(\mathbf{y}) = (0, \delta y_1, \sigma_m \mathbf{y}_{2:m-1}, \sigma_m \gamma \mathbf{x}^\top)^\top$ . Similarly, let  $\mathbf{Q}_2$  be the matrix given by

$$\mathbf{Q}_2 = \begin{pmatrix} 1 & \mathbf{0}_{(p+m-1)}^\top \\ \mathbf{0}_{(p+m-1)} & \mathbf{W} \end{pmatrix} \in \mathbb{R}^{(p+m) \times (p+m)}.$$

<sup>1</sup>A simple calculation shows that  $y_1 = \tau_{m-1}$  and  $\gamma = \prod_{i=1}^{m-1} \sigma_i$ .

From the induction hypothesis, it follows that  $\mathbf{Q}_2 \mathbf{Q}_1 \varphi_m(\mathbf{y}) = (0, \mathbf{0}_{m-2}, \sigma_m \alpha, \sigma_m \beta \mathbf{x}^\top)^\top =: (\mathbf{0}_{m-1}, a, b \mathbf{x}^\top)^\top$ . Therefore, for  $\mathbf{Q} \doteq \mathbf{Q}_2 \mathbf{Q}_1$  and  $\varphi^{(m)} = \varphi_m \circ \varphi^{(m-1)}$ ,

$$\mathbf{Q} \varphi^{(m)}(\mathbf{x}) = (\mathbf{0}_{m-1}, a, b \mathbf{x}^\top)^\top \in \mathbb{R}^{d+m}.$$

It remains to show that  $\mathbf{Q} = \mathbf{Q}_2 \mathbf{Q}_1 \in \mathcal{O}(p+m, q)$ . This follows by noting that  $\mathbf{Q}_2 \in \mathcal{O}(p+m, q)$  since

$$\mathbf{Q}_2^\top \mathbb{I}_{p+m, q} \mathbf{Q}_2 = \begin{pmatrix} 1 & \mathbf{0}_{(p+m-1)}^\top \\ \mathbf{0}_{(p+m-1)} & \mathbf{W}^\top \mathbb{I}_{p+m-1, q} \mathbf{W} \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0}_{(p+m-1)}^\top \\ \mathbf{0}_{(p+m-1)} & \mathbb{I}_{p+m-1, q} \end{pmatrix} = \mathbb{I}_{p+m, q},$$

and  $\mathbf{Q}_1 \in \mathcal{O}(p+m, q)$  since

$$\mathbf{Q}_1^\top \mathbb{I}_{p+m, q} \mathbf{Q}_1 = \begin{pmatrix} R^\top \mathbb{I}_{2,0} R & \mathbf{O}_{2 \times (d+m-2)} \\ \mathbf{O}_{(d+m-2) \times 2} & \mathbb{I}_{p+m-2, q} \end{pmatrix} = \mathbb{I}_{p+m, q}$$

Therefore, it follows that

$$(\mathbf{Q}_2 \mathbf{Q}_1)^\top \mathbb{I}_{p+m, q} (\mathbf{Q}_2 \mathbf{Q}_1) = \mathbf{Q}_1^\top (\mathbf{Q}_2^\top \mathbb{I}_{p+m, q} \mathbf{Q}_2) \mathbf{Q}_1 = \mathbf{Q}_1^\top \mathbb{I}_{p+m, q} \mathbf{Q}_1 = \mathbb{I}_{p+m, q}.$$

■

### C.3 Proof of Theorem 3.2

The outline of the proof is as follows:

- ① **Finite Reduction:** We consider a finite set of latent positions  $\mathcal{X}_0, \mathcal{X}_1 \subset \mathcal{X}$  which are  $\eta$ -separated in the  $d_{2, \infty}$  metric to reduce the minimax risk over  $\mathcal{X}$  to the minimax risk over a finite subset.
- ② **Le Cam's Lemma:** Next, we apply Le Cam's lemma to lower bound the minimax risk over  $\mathcal{X}_0 \cup \mathcal{X}_1$  in terms of the total variation metric between mixture distributions of the private outputs under any  $\varepsilon$ -edgeLDP mechanism satisfying Definition 2.2.
- ③ **Bounding via  $\chi^2$ -divergence:** We bound the total variation metric from above using the  $\chi^2$ -divergence and simplify the expression in terms of the induced marginal distributions of the private outputs.
- ④ **Simplifying  $\chi^2$ -divergence:** We further simplify the expression for the  $\chi^2$ -divergence using the structure of the probability matrices generated by the latent positions in  $\mathcal{X}_0 \cup \mathcal{X}_1$  and the constraints imposed by  $\varepsilon$ -edgeLDP. The proof in this step is fairly technical but uses standard techniques.
- ⑤ **Setting the optimal separation  $\eta$ :** We set the optimal  $\eta$  based on the above steps to obtain the result.

**Step ①.** Let  $\mathcal{X}$  be the set of all  $(p, q)$ -admissible latent positions, and fix  $\mathbf{X}_0 \in \mathcal{X}$  as follows. Let  $\mathbf{U}_0 \in \mathbb{U}(n, d)$  be the matrix guaranteed by [63, Lemma 4] satisfying:

$$\mathbf{U}_0 \mathbf{1}_d = \frac{1}{\sqrt{n}} \mathbf{1}_n, \quad \sqrt{\frac{d}{n}} \leq \|\mathbf{U}_0\|_{2, \infty} \leq \sqrt{\frac{d}{n-2d}}, \quad \text{and} \quad \max_{i \in [n], j \in [d]} |\mathbf{U}_{0, ij}| \leq \frac{1}{\sqrt{n-2d}}.$$

i.e., the first column of  $\mathbf{U}_0$  is proportional to the vector of 1s and has incoherence parameter  $\mu(\mathbf{U}_0) = nd/(n-2d)$ . For  $\lambda_1 = n/2$  and  $\lambda_2 = \dots = \lambda_d = n/12d$ , let  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_d)$ , and set  $\mathbf{X}_0 \doteq \mathbf{U}_0 \mathbf{\Lambda}^{1/2}$ . By construction, when  $n \geq 4d$  it is easy to see that  $\mathbf{X}_0 \in \mathcal{X}$  since the rows  $\mathbf{x}_{0,1}, \dots, \mathbf{x}_{0,n} \in \mathbb{R}^d$  of  $\mathbf{X}_0$  are such that

$$\begin{aligned} \mathbf{x}_{0,i}^\top \mathbb{I}_{p,q} \mathbf{x}_{0,j} &= \sum_{\ell} \omega_{\ell} \lambda_{\ell} u_{i\ell} u_{j\ell} \leq \lambda_1 \frac{1}{n} + \sum_{\ell=2}^d \lambda_{\ell} |u_{i\ell}| |u_{j\ell}| \leq \frac{1}{2} + \frac{n}{12(n-2d)} \leq \frac{2}{3} \\ \mathbf{x}_{0,i}^\top \mathbb{I}_{p,q} \mathbf{x}_{0,j} &= \sum_{\ell} \omega_{\ell} \lambda_{\ell} u_{i\ell} u_{j\ell} \geq \lambda_1 \frac{1}{n} - \sum_{\ell=2}^d \lambda_{\ell} |u_{i\ell}| |u_{j\ell}| \geq \frac{1}{2} - \frac{n}{12(n-2d)} \geq \frac{1}{3}, \end{aligned} \tag{25}$$

where  $\omega_{\ell} \in \{-1, +1\}$  are the diagonal entries of  $\mathbb{I}_{p,q}$ . Let  $\mathbf{P}_0 = \mathbf{X}_0 \mathbb{I}_{p,q} \mathbf{X}_0^\top$  be the probability matrix with sparsity  $\rho_n \leq 1$  and take  $\mathcal{X}_0 = \{\mathbf{X}_0\}$ . Consider the following set of  $n$  latent positions:

$$\mathcal{X}_1 = \left\{ \mathbf{X}_k \in \mathbb{R}^{n \times d} : \mathbf{X}_k = \mathbf{X}_0 + \eta e_k \mathbf{v}(k)^\top, \quad \mathbf{v}(k) = \mathbf{x}_{0,k} / \|\mathbf{x}_{0,k}\| \in \mathbb{S}^{d-1} \right\}.$$

In other words,  $\mathbf{X}_k$  is obtained from  $\mathbf{X}_0$  by scaling the  $k$ -th row of  $\mathbf{X}_0$  by  $\eta$ . It is easy to see that  $\mathbf{X}_k \in \mathcal{X}$  for all  $\eta < \sqrt{3/2} - 1$  sufficiently small. By construction, we have  $d_{2,\infty}(\mathbf{X}_k, \mathbf{X}_0) \geq \eta/2$ . Let  $\mathbf{P}_k = \mathbf{X}_k \mathbb{I}_{p,q} \mathbf{X}_k^\top$  be the probability matrix generated by each  $\mathbf{X}_k$  with sparsity  $\rho_n \leq 1$ , and define  $\mathbf{\Gamma}_k$  to be the difference between the probability matrices  $\mathbf{P}_k$  and  $\mathbf{P}_0$  given by

$$\mathbf{\Gamma}^k = \mathbf{P}_k - \mathbf{P}_0 = \rho_n \left( \eta \mathbf{u}(k) e_k^\top + \eta e_k \mathbf{u}(k)^\top + \eta^2 e_k e_k^\top \right), \quad \text{for } \mathbf{u}(k) = \mathbf{X}_0 \mathbb{I}_{p,q} \mathbf{v}(k). \quad (26)$$

**Step ②.** Consider the GRDPGs  $\mathbf{A}_0 \sim \mathcal{G}(\mathbf{X}_0, \rho_n; p, q)$  and  $\mathbf{A}_k \sim \mathcal{G}(\mathbf{X}_k, \rho_n; p, q)$ , i.e.,  $\mathbf{A}_0 \sim \text{Ber}(\mathbf{P}_0)$  and  $\mathbf{A}_k \sim \text{Ber}(\mathbf{P}_k)$ , respectively, and let  $F_0$  and  $F_k$  denote their joint distributions on  $\mathbb{B}(n)$ . For any  $\varepsilon$ -edgeLDP mechanism  $\mathcal{A} \in \mathbb{A}_\varepsilon$  satisfying Definition 2.2, let  $\mathbf{Z}_0 = \mathcal{A}(\mathbf{A}_0)$  and  $\mathbf{Z}_k = \mathcal{A}(\mathbf{A}_k)$  be the private outputs with distributions  $\mathcal{A}_\# F_0$  and  $\mathcal{A}_\# F_k$  on  $\mathcal{Z}$ , respectively, i.e., for all  $\ell \in \{0\} \cup [n]$  and  $A = (a_{ij}) \in \mathbb{B}(n)$ ,

$$F_\ell(A) = \prod_{i < j} F_{\ell,ij}(a_{ij}) = \prod_{i < j} \mathbf{P}_{\ell,ij}^{a_{ij}} (1 - \mathbf{P}_{\ell,ij})^{1-a_{ij}}$$

and for  $Z = (z_{ij})$ ,

$$\mathcal{A}_\# F_\ell(Z) = \int_{A \in \mathbb{B}(n)} \mathcal{Q}(Z \mid \mathbf{A} = A) dF_k(A) = \prod_{i < j} \int_{a_{ij} \in \{0,1\}} \mathcal{Q}_{ij}(z_{ij} \mid \mathbf{A}_{ij} = a_{ij}) dF_{\ell,ij}(a_{ij}) =: \psi_{\ell,ij}(z_{ij}). \quad (27)$$

Let  $\overline{\mathcal{A}_\# F} = \frac{1}{n} \sum_k \mathcal{A}_\# F_k \in \text{co}(\{F_1, \dots, F_n\})$  be the mixture distribution in the convex hull of  $\{\mathcal{A}_\# F_1, \dots, \mathcal{A}_\# F_n\}$ . From Le Cam's lemma [64, Lemma 1], it follows that

$$\mathfrak{R}_n(\mathcal{X}, \varepsilon) \doteq \inf_{\mathcal{A}_\varepsilon \in \mathbb{A}_\varepsilon} \inf_{\hat{\mathbf{X}}} \sup_{\mathbf{X} \in \mathcal{X}} \mathbb{E} \left[ d_{2,\infty}(\hat{\mathbf{X}}(\mathcal{A}_\varepsilon(\mathbf{A})), \mathbf{X}) \right] \geq \inf_{\mathcal{A}_\varepsilon \in \mathbb{A}_\varepsilon} \frac{\eta}{4} \left( 1 - \text{D}_{\text{TV}}(\overline{\mathcal{A}_\# F}, \mathcal{A}_\# F_0) \right), \quad (28)$$

where  $\text{D}_{\text{TV}}(\cdot, \cdot)$  is the total variation metric.

**Step ③.** Let  $\varphi_0$  and  $\varphi_k$  denote the density functions associated with the measures  $\mathcal{A}_\# F_0$  and  $\mathcal{A}_\# F_k$  on  $\mathcal{Z}$ , i.e.,  $d\mathcal{A}_\# F_k / d\mathcal{A}_\# F_0 = \varphi_k / \varphi_0$ , and let  $\bar{\varphi} = \frac{1}{n} \sum_k \varphi_k$ . From the properties of the Chi-squared divergence  $\text{D}_{\chi^2}(\cdot, \cdot)$  in [60, Lemma 2.27 and Eq. (2.15)], we get

$$\begin{aligned} \text{D}_{\text{TV}}(\overline{\mathcal{A}_\# F}, \mathcal{A}_\# F_0)^2 &\leq \text{D}_{\chi^2}(\overline{\mathcal{A}_\# F}, \mathcal{A}_\# F_0) \\ &= \int \left( \frac{d\mathcal{A}_\# F_k}{d\mathcal{A}_\# F_0} \right)^2 d\mathcal{A}_\# F_0 - 1 \\ &= \int \left( \frac{\bar{\varphi}}{\varphi_0} \right)^2 d\mathcal{A}_\# F_0 - 1 = \mathbb{E} \left[ \left( \frac{\bar{\varphi}(\mathbf{Z}_0)}{\varphi_0(\mathbf{Z}_0)} \right)^2 \right] - 1. \end{aligned}$$

where the expectation is taken over  $\mathbf{Z}_0 \sim \mathcal{A}_\# F_0$ . Moreover, by expanding the square and using (27), we have

$$\left( \frac{\bar{\varphi}(Z)}{\varphi_0(Z)} \right)^2 = \frac{1}{n^2} \sum_{k,\ell} \frac{\varphi_k(Z) \varphi_\ell(Z)}{\varphi_0(Z)^2} = \frac{1}{n^2} \sum_{k,\ell} \prod_{i < j} \frac{\psi_{k,ij}(z_{ij}) \psi_{\ell,ij}(z_{ij})}{\psi_{0,ij}(z_{ij}) \psi_{0,ij}(z_{ij})},$$

and it follows that

$$1 + \text{D}_{\chi^2}(\overline{\mathcal{A}_\# F}, \mathcal{A}_\# F_0) = \frac{1}{n^2} \sum_{k,\ell} \prod_{i < j} \mathbb{E}_{\mathbf{Z}_{0,ij}} \left[ \frac{\psi_{k,ij}(\mathbf{Z}_{0,ij}) \psi_{\ell,ij}(\mathbf{Z}_{0,ij})}{\psi_{0,ij}(\mathbf{Z}_{0,ij}) \psi_{0,ij}(\mathbf{Z}_{0,ij})} \right] = \frac{1}{n^2} \sum_{k,\ell} \prod_{i < j} \int \frac{\psi_{k,ij}(z_{ij}) \psi_{\ell,ij}(z_{ij})}{\psi_{0,ij}(z_{ij}) \psi_{0,ij}(z_{ij})} dz_{ij}. \quad (29)$$

**Step ④.** The simplification in this step is similar to the steps in [43, Appendix B]. Consider the integral in (29). We begin by noting that

$$\begin{aligned} \psi_{k,ij}(z) &= \int_{a_{ij} \in \{0,1\}} \mathcal{Q}_{ij}(z \mid \mathbf{A}_{ij} = a_{ij}) dF_{k,ij}(a_{ij}) \\ &= \mathcal{Q}_{ij}(z \mid \mathbf{A}_{ij} = 1) \mathbf{P}_{k,ij} + \mathcal{Q}_{ij}(z \mid \mathbf{A}_{ij} = 0) (1 - \mathbf{P}_{k,ij}) \\ &= \mathcal{Q}_{ij}(z \mid \mathbf{A}_{ij} = 1) (\mathbf{P}_{0,ij} + \mathbf{\Gamma}_{k,ij}) + \mathcal{Q}_{ij}(z \mid \mathbf{A}_{ij} = 0) (1 - \mathbf{P}_{0,ij} - \mathbf{\Gamma}_{k,ij}) \end{aligned}$$

$$\begin{aligned}
 &= (\mathbf{P}_{0,ij} + \mathbf{\Gamma}_{k,ij}) \left( \mathcal{Q}_{ij}(z \mid \mathbf{A}_{ij} = 1) - \mathcal{Q}_{ij}(z \mid \mathbf{A}_{ij} = 0) \right) + \mathcal{Q}_{ij}(z \mid \mathbf{A}_{ij} = 0) \\
 &= (\mathbf{P}_{0,ij} + \mathbf{\Gamma}_{k,ij}) \left( \xi_1(z) - \xi_0(z) \right) + \xi_0(z),
 \end{aligned}$$

where we used the fact that  $\mathbf{P}_{k,ij} = \mathbf{P}_{0,ij} + \mathbf{\Gamma}_{k,ij}$  and defined  $\xi_1(z) \doteq \mathcal{Q}_{ij}(z \mid \mathbf{A}_{ij} = 1)$  and  $\xi_0(z) \doteq \mathcal{Q}_{ij}(z \mid \mathbf{A}_{ij} = 0)$ . Plugging in the above expression into the integral in (29) and simplifying, we get

$$\begin{aligned}
 \frac{\psi_{k,ij}(z)\psi_{\ell,ij}(z)}{\psi_{0,ij}(z)} &= \frac{\left[ (\mathbf{P}_{0,ij} + \mathbf{\Gamma}_{k,ij}) \left( \xi_1(z) - \xi_0(z) \right) + \xi_0(z) \right] \cdot \left[ (\mathbf{P}_{0,ij} + \mathbf{\Gamma}_{\ell,ij}) \left( \xi_1(z) - \xi_0(z) \right) + \xi_0(z) \right]}{\mathbf{P}_{0,ij} \left( \xi_1(z) - \xi_0(z) \right) + \xi_0(z)} \\
 &= \overbrace{\left[ \mathbf{P}_{0,ij} \left( \xi_1(z) - \xi_0(z) \right) + \xi_0(z) \right]}^{=:f_1(z)} + \overbrace{\left[ (\mathbf{\Gamma}_{k,ij} + \mathbf{\Gamma}_{\ell,ij}) \left( \xi_1(z) - \xi_0(z) \right) \right]}^{=:f_2(z)} \\
 &\quad + \underbrace{\left[ \frac{\mathbf{\Gamma}_{k,ij}\mathbf{\Gamma}_{\ell,ij}(\xi_1(z) - \xi_0(z))^2}{\mathbf{P}_{0,ij} \left( \xi_1(z) - \xi_0(z) \right) + \xi_0(z)} \right]}_{=:f_3(z)}.
 \end{aligned}$$

Note that

$$\begin{aligned}
 \int f_1(z)dz &= \mathbf{P}_{0,ij} \int \xi_1(z)dz + (1 - \mathbf{P}_{0,ij}) \int \xi_0(z)dz = 1 \\
 \int f_2(z)dz &= (\mathbf{\Gamma}_{k,ij} + \mathbf{\Gamma}_{\ell,ij}) \int (\xi_1(z) - \xi_0(z))dz = 0,
 \end{aligned}$$

and,

$$\int f_3(z)dz = \mathbf{\Gamma}_{k,ij}\mathbf{\Gamma}_{\ell,ij} \int \frac{(\xi_1(z) - \xi_0(z))^2}{\mathbf{P}_{0,ij} \left( \xi_1(z) - \xi_0(z) \right) + \xi_0(z)} dz.$$

From (25) we have  $\mathbf{P}_{0,ij} \leq 2\rho_n/3$ , and for  $\varepsilon \leq \min\{1, 3/8\rho_n\} \leq \min\{1, 1/4\mathbf{P}_{0,ij}\}$  using [43, Lemma B.1], we have

$$\mathbf{P}_{0,ij} \left( \xi_1(z) - \xi_0(z) \right) + \xi_0(z) \geq \xi_0(z)/2.$$

Plugging this back into the integral of  $f_3(z)$ ,

$$\begin{aligned}
 \int f_3(z)dz &\leq \mathbf{\Gamma}_{k,ij}\mathbf{\Gamma}_{\ell,ij} \int \frac{\xi_0(z)^2(\xi_1(z)/\xi_0(z) - 1)^2}{\xi_0(z)/2} dz = 2\mathbf{\Gamma}_{k,ij}\mathbf{\Gamma}_{\ell,ij} \int \left( \frac{\xi_1(z)}{\xi_0(z)} - 1 \right)^2 \xi_0(z)dz \\
 &\leq 2\mathbf{\Gamma}_{k,ij}\mathbf{\Gamma}_{\ell,ij}(e^\varepsilon - 1)^2,
 \end{aligned}$$

where the final inequality follows from the definition of  $\varepsilon$ -edgeLDP in Definition 2.2 which requires that

$$\frac{\xi_1(z)}{\xi_0(z)} = \frac{\mathcal{Q}_{ij}(z \mid \mathbf{A}_{ij} = 1)}{\mathcal{Q}_{ij}(z \mid \mathbf{A}_{ij} = 0)} \leq e^\varepsilon.$$

Since  $\varepsilon \leq 1$ , it follows that  $(e^\varepsilon - 1) = \sigma(\varepsilon)^2(e^\varepsilon + 1) \leq 4\sigma(\varepsilon)^2$ , and

$$\int \frac{\psi_{k,ij}(z)\psi_{\ell,ij}(z)}{\psi_{0,ij}(z)} dz \leq 1 + 2(e^\varepsilon - 1)^2 \leq 1 + c_1\mathbf{\Gamma}_{k,ij}\mathbf{\Gamma}_{\ell,ij} \cdot \sigma(\varepsilon)^4 \leq \exp(c_1\mathbf{\Gamma}_{k,ij}\mathbf{\Gamma}_{\ell,ij} \cdot \sigma(\varepsilon)^4),$$

where  $c_1 = 16$ . Plugging this bound back into (29), we get

$$1 + \mathbf{D}_{\chi^2}(\overline{\mathcal{A}_\# F}, \mathcal{A}_\# F_0) = \frac{1}{n^2} \sum_{k,\ell} \prod_{i < j} \int \frac{\psi_{k,ij}(z)\psi_{\ell,ij}(z)}{\psi_{0,ij}(z)} dz$$

$$\begin{aligned}
 &\leq \frac{1}{n^2} \sum_{k,\ell} \prod_{i < j} \exp(c_1 \sigma(\varepsilon)^4 \mathbf{\Gamma}_{k,ij} \mathbf{\Gamma}_{\ell,ij} \cdot \sigma(\varepsilon)^4) \\
 &= \frac{1}{n^2} \sum_{k,\ell} \exp \left( c_1 \sigma(\varepsilon)^4 \sum_{i < j} \mathbf{\Gamma}_{k,ij} \mathbf{\Gamma}_{\ell,ij} \right) \\
 &= \frac{1}{n^2} \sum_{k,\ell} \exp \left( c_1 \sigma(\varepsilon)^4 \langle \mathbf{\Gamma}_k, \mathbf{\Gamma}_\ell \rangle_F \right) \\
 &= \frac{1}{n^2} \sum_k \exp \left( c_1 \sigma(\varepsilon)^4 \|\mathbf{\Gamma}_k\|_F^2 \right) + \frac{1}{n^2} \sum_{k \neq \ell} \exp \left( c_1 \sigma(\varepsilon)^4 \langle \mathbf{\Gamma}_k, \mathbf{\Gamma}_\ell \rangle_F \right) \\
 &\leq \frac{1}{n} \exp \left( c_1 \sigma(\varepsilon)^4 \max_k \|\mathbf{\Gamma}_k\|_F^2 \right) + \frac{n(n-1)}{n^2} \exp \left( c_1 \sigma(\varepsilon)^4 \max_{k \neq \ell} \langle \mathbf{\Gamma}_k, \mathbf{\Gamma}_\ell \rangle_F \right) \\
 &\leq \exp \left( c_1 \sigma(\varepsilon)^4 \max_k \|\mathbf{\Gamma}_k\|_F^2 - \log n \right) + \exp \left( c_1 \sigma(\varepsilon)^4 \max_{k \neq \ell} \langle \mathbf{\Gamma}_k, \mathbf{\Gamma}_\ell \rangle_F \right). \tag{30}
 \end{aligned}$$

**Step ⑤.** For  $\mathbf{\Gamma}_k$  in (26), using the triangle inequality and by noting that  $\mathbf{u}(k) = \mathbf{X}_0 \mathbb{I}_{p,q} v(k)$ , we have

$$\begin{aligned}
 \|\mathbf{\Gamma}_k\|_F &\leq \rho_n \left\| \eta \mathbf{u}(k) e_k^\top + \eta e_k \mathbf{u}(k)^\top + \eta^2 e_k e_k^\top \right\|_F \\
 &\leq \rho_n (\eta^2 + 2\eta \|u(k)\|) \\
 &\leq \rho_n (\eta^2 + 2\eta \|\mathbf{X}_0\|_{\text{op}}) \leq 4\rho_n \eta \sqrt{n},
 \end{aligned}$$

since  $\eta^2 < \eta < 1$ . It follows that

$$\max_k \|\mathbf{\Gamma}_k\|_F^2 \leq 16\rho_n^2 \eta^2 n.$$

Similarly, note that  $\mathbf{\Gamma}_k$  has non-zero entries only in the  $k$ -th row and column, and for  $k \neq \ell$  we have a total of two non-zero entries in  $\mathbf{\Gamma}_k^\top \mathbf{\Gamma}_\ell$ , and, therefore,

$$\langle \mathbf{\Gamma}_k, \mathbf{\Gamma}_\ell \rangle_F = \text{tr}(\mathbf{\Gamma}_k^\top \mathbf{\Gamma}_\ell) = 2\rho_n^2 \eta^2 \cdot u(k)_\ell \cdot u(\ell)_k \leq 2\rho_n^2 \eta^2,$$

where  $u(k)_\ell = \mathbf{x}_{0,\ell}^\top \mathbb{I}_{p,q} v(k) \leq 1$  and similarly for  $u(\ell)_k$ . Plugging these bounds back into  $D_{\chi^2}(\cdot, \cdot)$  in (30), we get

$$1 + D_{\chi^2}(\overline{\mathcal{A}_\# F}, \mathcal{A}_\# F_0) \leq \exp \left( 16c_1 \sigma(\varepsilon)^4 \rho_n \eta^2 n - \log n \right) + \exp \left( 8c_1 \sigma(\varepsilon)^4 \rho_n^2 \eta^2 \right).$$

For a suitably large absolute constant  $C > 16c_1$ , choosing

$$\eta = \frac{1}{C} \sqrt{\frac{\log n}{n\rho_n^2 \sigma(\varepsilon)^4}}$$

results in  $D_{\chi^2}(\overline{\mathcal{A}_\# F}, \mathcal{A}_\# F_0) = o(1)$ , and by Le Cam's lemma in (28), we have

$$\mathfrak{R}_n(\mathcal{X}, \varepsilon) \geq \frac{\eta}{4} (1 - D_{\chi^2}(\overline{\mathcal{A}_\# F}, \mathcal{A}_\# F_0)) = \Omega \left( \sqrt{\frac{\log n}{n\rho_n^2 \sigma(\varepsilon)^4}} \right).$$

This completes the proof of Theorem 3.2. ■

#### C.4 Proof of Theorem 3.3

First, let  $\check{\mathbf{A}}$  be the *privacy-adjusted adjacency matrix* from Algorithm 1 and let  $\overline{\mathbf{A}}$  be the matrix given by

$$\overline{\mathbf{A}} = \mathcal{M}_\varepsilon(\mathbf{A}) - \tau^2 \mathbf{1}\mathbf{1}^\top.$$

with spectral decomposition  $\overline{\mathbf{A}} = \overline{\mathbf{U}} \overline{\mathbf{\Lambda}} \overline{\mathbf{U}}^\top$  corresponding to the leading  $d$  eigenvalues by absolute value. Let  $\overline{\mathbf{X}} = \overline{\mathbf{U}} |\overline{\mathbf{\Lambda}}|^{1/2} \in \mathbb{R}^{n \times d}$  be the spectral embedding of  $\overline{\mathbf{A}}$ . The proof outline is as follows:



- ① We first show that  $\bar{\mathbf{A}}$  is close to the expected adjacency matrix  $\bar{\mathbf{P}}$  of a GRDPG, and establish the convergence rate of  $\bar{\mathbf{X}}$  to  $t_n \mathbf{X}$  in the  $d_{2,\infty}$  metric.
- ② We use the result in ① to quantify the convergence rate of  $\rho_n^{1/2} \tilde{\mathbf{X}}$  to  $\mathbf{X}$  in the  $d_{2,\infty}$  metric.
- ③ Finally, use  $\check{\rho}_n$  from Algorithm 1 as a plug-in estimator to get the final bound.

**Step ①.** In a similar fashion to the sparse GRDPGs considered in Rubin-Delanchy et al. [57, Theorem 5], let  $\bar{\mathbf{P}}$  be the expected adjacency matrix for the random dot product graph  $(\mathbf{B}, \sigma \mathbf{X}) \sim \mathcal{G}(\sigma_{\#} \mathbb{P}, \rho_n; p, q)$  given by

$$\bar{\mathbf{P}} = \rho_n \cdot (\sigma \mathbf{X}) \mathbb{I}_{p,q} (\sigma \mathbf{X})^\top = (t_n^{1/2} \mathbf{X}) \mathbb{I}_{p,q} (t_n^{1/2} \mathbf{X})^\top,$$

where, for notational simplicity,  $t_n^{1/2} \doteq \sigma \rho_n^{1/2}$  denotes the *effective sparsity parameter* under privacy. Although  $\bar{\mathbf{A}}$  is not the adjacency matrix of a GRDPG, it is close to  $\mathbf{B}$ . Indeed, for all  $1 \leq i < j \leq n$ ,

$$\begin{aligned} \mathbb{E}(\bar{\mathbf{A}}_{ij}) &= \mathbb{E}(\mathcal{M}_\varepsilon(\mathbf{A})_{ij} - \tau^2) \stackrel{(i)}{=} Y_i^\top \mathbb{I}_{p,q} Y_j - \tau^2 \\ &\stackrel{(ii)}{=} (\sigma \rho_n^{1/2} X_i)^\top \mathbb{I}_{p,q} (\sigma \rho_n^{1/2} X_j) + \tau^2 - \tau^2 \\ &= \sigma^2 \rho_n X_i^\top \mathbb{I}_{p,q} X_j \\ &= \bar{\mathbf{P}}_{ij}, \end{aligned}$$

where in (i) we used the fact that  $(\mathcal{M}_\varepsilon(\mathbf{A}), \mathbf{Y}) \sim \mathcal{G}(\varphi_{\#} \mathbb{P}, 1; p+1, q)$  from Theorem 3.1 and (ii) follows from the definition of  $\varphi$ . Using [42, Theorem 5.2] (see Lemma B.3), with probability  $1 - O(n^{-1})$  it follows that

$$\|\bar{\mathbf{A}} - \bar{\mathbf{P}}\| = \|\mathbf{A}_{\mathbf{Y}} - \mathbf{P}_{\mathbf{Y}}\| = O(\sqrt{n}).$$

Furthermore, for  $\bar{\mathbf{U}}$  in the spectral decomposition  $\bar{\mathbf{P}} = \bar{\mathbf{U}} \bar{\mathbf{\Lambda}} \bar{\mathbf{U}}^\top$ , from<sup>2</sup> [57, Lemma 12], it follows that

$$\|(\bar{\mathbf{A}} - \bar{\mathbf{P}}) \bar{\mathbf{U}}\|_{2,\infty} = O\left(\frac{\sqrt{n \log n}}{\sqrt{n}}\right),$$

with probability  $1 - O(n^{-1})$ . Next, we prepare to quantify  $d_{2,\infty}(\bar{\mathbf{X}}, t_n \mathbf{X})$ . In order to do so, we need to characterize the matrices which align  $\bar{\mathbf{X}}$  to  $t_n \mathbf{X}$  as per Definition 3.1.

To this end, let  $\mathbf{Q}_{\sqrt{t_n} \mathbf{X}}$  be the alignment matrix satisfying  $\bar{\mathbf{U}} |\bar{\mathbf{\Lambda}}|^{1/2} = t_n^{1/2} \mathbf{X} \mathbf{Q}_{\sqrt{t_n} \mathbf{X}}^{-1}$  (see Fact A.2); by invariance of  $\mathbf{Q}_{\mathbf{X}}$  to scale transformations (Lemma A.1), it follows that  $\mathbf{Q}_{\sqrt{t_n} \mathbf{X}} = \mathbf{Q}_{\mathbf{X}}$ . Let  $\mathbf{Q}_n = \mathbf{O}_{\mathbf{X}}^\top \mathbf{Q}_{\sqrt{t_n} \mathbf{X}} = \mathbf{O}_{\mathbf{X}}^\top \mathbf{Q}_{\mathbf{X}}$  where  $\mathbf{O}_{\mathbf{X}} \in \mathcal{O}(d)$  is described in Fact A.1. From Rubin-Delanchy et al. [57, Eq. 15], it follows that

$$\bar{\mathbf{X}} - t_n^{1/2} \mathbf{X} \mathbf{Q}_n^{-1} = (\bar{\mathbf{A}} - \bar{\mathbf{P}}) \bar{\mathbf{U}} |\bar{\mathbf{\Lambda}}|^{-1/2} \mathbb{I}_{p,q} \mathbf{O}_{\mathbf{X}} + \mathbf{R}.$$

From the definition of  $d_{2,\infty}$  from (24) in Definition 3.1,

$$d_{2,\infty}(\bar{\mathbf{X}}, t_n \mathbf{X}) \leq \left\| \bar{\mathbf{X}} - t_n^{1/2} \mathbf{X} \mathbf{Q}_n^{-1} \right\|_{2,\infty}.$$

From [57, Eq. (14)] and [2, Lemma 3], the matrix  $\mathbf{R}$  is such that with probability  $1 - O(n^{-1})$ ,

$$\|\mathbf{R}\|_{2,\infty} = O_p\left(\sqrt{\frac{\log^4 n}{n^2 t_n}}\right),$$

and, for the first term on the r.h.s., using Cape et al. [10, Propositions 6.3 & 6.5], we obtain

$$\begin{aligned} \left\| (\bar{\mathbf{A}} - \bar{\mathbf{P}}) \bar{\mathbf{U}} |\bar{\mathbf{\Lambda}}(\bar{\mathbf{P}})|^{-1/2} \mathbb{I}_{p,q} \right\|_{2,\infty} &\leq \left\| (\bar{\mathbf{A}} - \bar{\mathbf{P}}) \bar{\mathbf{U}} \right\|_{2,\infty} \left\| |\bar{\mathbf{\Lambda}}(\bar{\mathbf{P}})|^{-1/2} \mathbb{I}_{p,q} \right\| \\ &= \left\| (\bar{\mathbf{A}} - \bar{\mathbf{P}}) \bar{\mathbf{U}} \right\|_{2,\infty} \left\| |\bar{\mathbf{\Lambda}}(\bar{\mathbf{P}})|^{-1/2} \right\| \left\| \mathbb{I}_{p,q} \right\| \end{aligned}$$

<sup>2</sup>Note that the  $\log^c n$  dependence in [57, Lemma 12] is replaced with  $\log n$  using [3, Theorem 4].

$$\begin{aligned} &\leq \|(\bar{\mathbf{A}} - \bar{\mathbf{P}})\bar{\mathbf{U}}\|_{2,\infty} \left\| |\boldsymbol{\Lambda}(\bar{\mathbf{P}})|^{-1/2} \right\| \\ &\stackrel{(i)}{=} O\left(\frac{\log n}{\sqrt{nt_n}}\right), \end{aligned}$$

with probability  $1 - O(n^{-1})$ , where (i) uses the fact that  $\lambda_i(\bar{\mathbf{P}}) = \Theta(nt_n)$ . Plugging in the value for  $t_n$ , we obtain

$$\left\| \bar{\mathbf{X}} - \sigma \rho_n^{1/2} \mathbf{X} \mathbf{Q}_n^{-1} \right\|_{2,\infty} = O\left(\frac{\log n}{\sqrt{n\sigma^2\rho_n}} + \sqrt{\frac{\log^4 n}{n^2\sigma^2\rho_n}}\right) = O\left(\frac{\log n}{\sqrt{n\sigma^2\rho_n}}\right). \quad (31)$$

**Step ②.** Note that  $\check{\mathbf{A}} = \frac{1}{\sigma^2} \bar{\mathbf{A}}$ , and the spectral decomposition of  $\check{\mathbf{A}}$  corresponding to the leading  $d$  eigenvalues by magnitude becomes

$$\check{\mathbf{A}} = \check{\mathbf{U}} \check{\boldsymbol{\Lambda}} \check{\mathbf{U}}^\top = \bar{\mathbf{U}} \left( \frac{1}{\sigma^2} \bar{\boldsymbol{\Lambda}} \right) \bar{\mathbf{U}}^\top.$$

Therefore,  $\check{\mathbf{X}} = \check{\mathbf{U}} |\check{\boldsymbol{\Lambda}}|^{1/2} = \frac{1}{\sigma} \bar{\mathbf{X}}$ , and from (31), it follows that

$$\left\| \check{\mathbf{X}} - \rho_n^{1/2} \mathbf{X} \mathbf{Q}_n^{-1} \right\|_{2,\infty} = O\left(\frac{\log n}{\sqrt{n\sigma^4\rho_n}}\right) \quad \text{and} \quad \left\| \rho_n^{-1/2} \check{\mathbf{X}} - \mathbf{X} \mathbf{Q}_n^{-1} \right\|_{2,\infty} = O\left(\frac{\log n}{\sqrt{n\sigma^4\rho_n^2}}\right), \quad (32)$$

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

**Step ③.** Let  $\mathbf{Z} = \mathcal{M}_\varepsilon(\mathbf{A})$  and consider the estimator  $\check{\rho}_n$ , given by

$$\check{\rho}_n = \binom{n}{2}^{-1} \sum_{i < j} \check{\mathbf{A}}_{ij} = \frac{1}{\sigma^2} \left\{ \binom{n}{2}^{-1} \sum_{i < j} \mathbf{Z}_{ij} - \tau^2 \right\} = \frac{1}{\sigma^2} (\hat{\theta} - \tau^2).$$

Since the entries  $\mathcal{M}_{ij}$  are i.i.d.,  $\hat{\theta} \doteq \binom{n}{2}^{-1} \sum_{i < j} \mathbf{M}_{ij}$  is a second-order  $U$ -statistic. The expected value of  $\hat{\theta}$  is given by

$$\mathbb{E}(\hat{\theta} \mid \mathbf{X}) = \mathbb{E}(\mathbf{Z}_{ij} \mid \mathbf{X}) = Y_i^\top \mathbb{I}_{p+1,q} Y_j = \rho_n \sigma^2 (X_i^\top \mathbb{I}_{p,q} X_j) + \tau^2,$$

and, by assumption that  $\mathbb{E}[X_i^\top \mathbb{I}_{p,q} X_j] = 1$ , it follows that

$$\mathbb{E}(\hat{\theta}) = \mathbb{E}_{\mathbf{X}} [\mathbb{E}(\hat{\theta} \mid \mathbf{X})] = \rho_n \sigma^2 \mathbb{E}[X_i^\top \mathbb{I}_{p,q} X_j] + \tau^2 = \rho_n \sigma^2 + \tau^2.$$

Therefore,

$$\mathbb{E}(\check{\rho}_n \mid \mathbf{X}) = \rho_n \cdot \binom{n}{2}^{-1} \sum_{i < j} \mathbf{P}_{ij} \quad \text{and} \quad \mathbb{E}(\check{\rho}_n) = \rho_n.$$

By the standard protocol for bounding  $U$ -statistics by using Hoeffding's inequality (see, e.g., [2, Lemma 5]), it follows that with probability  $1 - O(n^{-2})$ ,

$$\frac{1}{\sqrt{\check{\rho}_n}} = \frac{1}{\sqrt{\rho_n}} \left( 1 + O\left(\sqrt{\frac{\log n}{n\rho_n}}\right) \right). \quad (33)$$

Combining the results from (32) and (33), we get that when  $n\rho_n = \omega(\log n)$ ,

$$\left\| \check{\rho}_n^{-1/2} \check{\mathbf{X}} - \mathbf{X} \mathbf{Q}_n^{-1} \right\|_{2,\infty} = \left\| \rho_n^{-1/2} \check{\mathbf{X}} - \mathbf{X} \mathbf{Q}_n^{-1} \right\|_{2,\infty} (1 + o(1)) = O\left(\frac{\log n}{\sqrt{n\sigma^4\rho_n^2}}\right). \quad (34)$$

Since  $\mathbf{Q}_n^{-1} = \mathbf{Q}_{\mathbf{X}}^{-1} \mathbf{O}_{\mathbf{X}}$ , it follows that with probability  $1 - O(n^{-1})$ ,

$$d_{2,\infty}(\check{\rho}_n^{-1/2} \check{\mathbf{X}}, \mathbf{X}) \leq \left\| \check{\rho}_n^{-1/2} \check{\mathbf{X}} - \mathbf{X} \mathbf{Q}_n^{-1} \right\|_{2,\infty} = O\left(\frac{\log n}{\sqrt{n\sigma^4\rho_n^2}}\right).$$

■

### C.5 Proof of Theorem 3.4

Let  $\check{\mathbf{X}}, \check{\rho}_n$  be the privacy-adjusted spectral embedding and the estimated sparsity parameter from Algorithm 1, respectively, and let  $\mathbf{Q}_{\mathbf{X}}, \mathbf{Q}_{\xi}, \mathbf{O}_{\mathbf{X}}, \mathbf{O}_{\xi}$  be as defined in Fact A.2, Fact A.4 and Fact A.5, respectively, and let  $\mathbf{Q}_n = \mathbf{O}_{\mathbf{X}}^T \mathbf{Q}_{\mathbf{X}}$  as in the proof of Theorem 3.3 in Appendix C.4.

Given  $D_n = \mathfrak{Dgm}(\mathbf{X})$  and  $\check{D}_n = \mathfrak{Dgm}(\check{\mathbf{X}}/\sqrt{\check{\rho}_n})$ , by the triangle inequality we have

$$\begin{aligned} W_{\infty}(\check{D}_n, D_n) &\leq W_{\infty}(\check{D}_n, \mathfrak{Dgm}(\mathbf{X}\mathbf{Q}_n^{-1})) \\ &\quad + W_{\infty}(\mathfrak{Dgm}(\mathbf{X}\mathbf{Q}_n^{-1}), \mathfrak{Dgm}(\mathbf{X}\mathbf{Q}_{\mathbf{X}}^{-1})) \\ &\quad + W_{\infty}(\mathfrak{Dgm}(\mathbf{X}\mathbf{Q}_{\mathbf{X}}^{-1}), \mathfrak{Dgm}(\mathbf{X}\mathbf{Q}_{\xi}^{-1}\mathbf{O}_{\xi})) \\ &\quad + W_{\infty}(\mathfrak{Dgm}(\mathbf{X}\mathbf{Q}_{\xi}^{-1}\mathbf{O}_{\xi}), D_n). \end{aligned}$$

By the stability of persistence diagrams (Theorem D.1), for  $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{n \times d}$ ,

$$W_{\infty}(\mathfrak{Dgm}(\mathbf{X}), \mathfrak{Dgm}(\mathbf{Y})) \leq d_H(\mathbf{X}, \mathbf{Y}) = \|\mathbf{X} - \mathbf{Y}\|_{2,\infty},$$

where  $d_H(\mathbf{X}, \mathbf{Y})$  denotes the Hausdorff distance between  $\mathbf{X}$  and  $\mathbf{Y}$ . Therefore, we have the following bounds:

$$\begin{aligned} W_{\infty}(\check{D}_n, D_n) &\leq \|\check{\rho}_n \check{\mathbf{X}} - \mathbf{X}\mathbf{Q}_n^{-1}\|_{2,\infty} & (= T_1) \\ &\quad + W_{\infty}(\mathfrak{Dgm}(\mathbf{X}\mathbf{Q}_n^{-1}), \mathfrak{Dgm}(\mathbf{X}\mathbf{Q}_{\mathbf{X}}^{-1})) & (= T_2) \\ &\quad + \|\mathbf{X}\mathbf{Q}_{\mathbf{X}}^{-1} - \mathbf{X}\mathbf{Q}_{\xi}^{-1}\mathbf{O}_{\xi}\|_{2,\infty} & (= T_3) \\ &\quad + W_{\infty}(\mathfrak{Dgm}(\mathbf{X}\mathbf{Q}_{\xi}^{-1}\mathbf{O}_{\xi}), D_n) & (= T_4) \end{aligned}$$

The proof proceeds by bounding each term  $T_1, T_2, T_3, T_4$  in the following steps. From Theorem 3.3, and in particular, from (34), it follows that with probability  $1 - O(n^{-1})$ ,

$$T_1 = \|\check{\rho}_n \check{\mathbf{X}} - \mathbf{X}\mathbf{Q}_n^{-1}\|_{2,\infty} = O\left(\frac{\log n}{\sqrt{n\sigma^4\rho_n^2}}\right).$$

For the second term,  $\mathbf{Q}_n^{-1} = \mathbf{Q}_{\mathbf{X}}^{-1}\mathbf{O}_{\mathbf{X}}$  where  $\mathbf{O}_{\mathbf{X}} \in \mathcal{O}(d)$ . Since persistence diagrams are invariant to orthogonal transformations (35), it follows that

$$T_2 = 0$$

Similarly, for the fourth term, since  $\mathbf{Q}_{\xi} \in \mathcal{O}(p, q) \cap \mathcal{O}(d)$  (see Fact A.4) and  $\mathbf{O}_{\xi} \in \mathcal{O}(d)$ , it follows that  $\mathbf{Q}_{\xi}^{-1}\mathbf{O}_{\xi} \in \mathcal{O}(d)$ . By invariance to  $\mathcal{O}(d)$  transformations, it also follows that

$$T_4 = 0.$$

Finally, for the third term, using [10, Propositions 6.3 & 6.5],

$$T_3 = \|\mathbf{X}\mathbf{Q}_{\mathbf{X}}^{-1} - \mathbf{X}\mathbf{Q}_{\xi}^{-1}\mathbf{O}_{\xi}\|_{2,\infty} = \|\mathbf{X}\|_{2,\infty} \|\mathbf{Q}_{\mathbf{X}}^{-1} - \mathbf{X}\mathbf{Q}_{\xi}^{-1}\mathbf{O}_{\xi}\|_{\text{op}} = O_p\left(\frac{\log n}{\sqrt{n}}\right),$$

where the final bound follows from Lemma B.4 and the fact that  $\|\mathbf{X}\|_{2,\infty} \leq C$  for some constant  $C$  when  $\mathbf{X}$  is  $(p, q)$ -admissible [2, Theorem 3].

Combining the bounds for  $T_1, T_2, T_3, T_4$ ,

$$W_{\infty}(\check{D}_n, D_n) = O\left(\frac{\log n}{\sqrt{\sigma^2 n \rho_n}}\right),$$

and we obtain the desired result. ■

## D Topological Data Analysis

We present the necessary background for Topological Data Analysis (TDA) here. We refer the reader to [14, 62] for an accessible overview and to [21] for a comprehensive treatment.

**Persistence Diagrams.** Given a collection of points  $\mathbf{X} \in \mathbb{R}^{n \times d}$  where  $\{\mathbf{x}_1, \dots, \mathbf{x}_n\} \subset \mathbb{R}^d$ , persistent homology sheds important insights on the geometric and topological features underlying  $\mathbf{X}$ . The underlying philosophy is as follows. The shape of  $\mathbf{X}$  at resolution  $t > 0$  is encoded in the union of balls  $B(\mathbf{X}, t) \doteq \cup_{\mathbf{x} \in \mathbf{X}} B(\mathbf{x}, t)$  centered at the sample points  $\mathbf{X}$ . Typically, this information is extracted using a geometric object  $\mathcal{K}(\mathbf{X}, t)$ , called the *simplicial complex*.

The collection  $\{B(\mathbf{X}, t) : t \geq 0\}$ , called a *filtration*, is a sequence of nested topological spaces, i.e. for  $t_1 < t_2 < \dots < t_m$ ,  $B(\mathbf{X}, t_1) \subseteq B(\mathbf{X}, t_2) \subseteq \dots \subseteq B(\mathbf{X}, t_m)$ . As  $t$  varies, the evolution of topology is encoded in this filtration. Roughly speaking, new cycles (i.e. connected components, loops, holes, etc.) appear, or existing cycles disappear (get filled out). Specifically, each  $k$ -dimensional feature in  $B(\mathbf{X}, t)$  can be represented as an element in a vector space,  $H_k(\mathbf{X}, t)$ , called its *homology*. When a new  $k$ -dimensional feature appears in  $B(\mathbf{X}, b)$  for some  $b > 0$ , then a non-trivial  $k$ -cycle appears in  $H_k(\mathbf{X}, b)$ , and the dimension of this vector space increases by one. In this case, the feature is said to be *born* at  $b$ . The same  $k$ -dimensional feature *dies* at resolution  $d > 0$ , if this  $k$ -cycle is absent in  $H_k(\mathbf{X}, d + \delta)$ , for all  $\delta > 0$ . *Persistent homology*,  $\mathbf{PH}_*(\mathbf{X})$ , is an *algebraic module* which tracks these persistent pairs  $(b, d)$  of births and deaths across the entire filtration for every  $k$ . The persistent pairs in  $\mathbf{PH}_*(\mathbf{X})$  are summarized in a *persistence diagram*

$$\mathfrak{Dgm}(\mathbf{X}) = \{(b, d) \in \mathbb{R}^2 : 0 \leq b < d \leq \infty \text{ and } (b, d) \in \mathbf{PH}_*(\mathbf{X})\}.$$

See Figure 7 for an illustration. The space of persistence diagrams  $\Omega = \{(x, y) \in \mathbb{R}^2 : x < y\}$ , is endowed with a collection of Wasserstein metrics  $\{W_p(\cdot, \cdot)\}_{p \geq 1}$ , and the special case of  $W_\infty(\cdot, \cdot)$  is referred to as the *bottleneck distance*.

**Definition D.1.** For two persistence diagrams  $\mathbf{D}_1$  and  $\mathbf{D}_2$ , the bottleneck distance is given by

$$W_\infty(\mathbf{D}_1, \mathbf{D}_2) \doteq \inf_{\phi \in \Phi} \sup_{\mathbf{x} \in \mathbf{D}_1 \cup \partial\Omega} \|\mathbf{x} - \phi(\mathbf{x})\|_\infty,$$

where  $\Phi$  is the collection of bijections between  $\mathbf{D}_1$  and  $\mathbf{D}_2$  including the boundary  $\partial\Omega = \{(x, y) \in \mathbb{R}^2 : x = y\}$  with infinite multiplicity.

We assume throughout this work that the probability distribution  $\mathbb{P}$  underlying the random dot product graph has compact support. Compactness ensures that the resulting persistence diagrams are *pointwise finite dimensional*, and satisfy a stability property w.r.t. the Hausdorff distance  $d_H(\cdot, \cdot)$ .

**Theorem D.1** (Stability of Persistence Diagrams, 12). For compact sets  $\mathbf{X}, \mathbf{Y} \subset (\mathbb{M}, \rho)$  from a metric space,

$$W_\infty(\mathfrak{Dgm}(\mathbf{X}), \mathfrak{Dgm}(\mathbf{Y})) \leq d_H(\mathbf{X}, \mathbf{Y}) \doteq \inf_{\mathbf{x} \in \mathbf{X}} \sup_{\mathbf{y} \in \mathbf{Y}} \rho(\mathbf{x}, \mathbf{y}).$$

and between two finite point clouds  $\mathbf{X} \in \mathbb{R}^{n \times d}$  and  $\mathbf{Y} \in \mathbb{R}^{m \times d}$ ,

$$W_\infty(\mathfrak{Dgm}(\mathbf{X}), \mathfrak{Dgm}(\mathbf{Y})) \leq d_H(\mathbf{X}, \mathbf{Y}) \doteq \max \left\{ \min_{i \in [n]} \max_{j \in [m]} \|\mathbf{x}_i - \mathbf{y}_j\|, \min_{j \in [m]} \max_{i \in [n]} \|\mathbf{x}_i - \mathbf{y}_j\| \right\}.$$

Although they encode subtle features underlying the data, an important property of persistence diagrams is that they are invariant to  $O(d)$  transformations, i.e. for all  $\mathbf{X} \in \mathbb{R}^{n \times d}$  and  $\mathbf{O} \in O(d)$ ,

$$\mathfrak{Dgm}(\mathbf{X}\mathbf{O}) = \mathfrak{Dgm}(\mathbf{X}). \tag{35}$$

This makes persistence diagrams particularly useful for analyzing the latent structure underlying the spectral embedding of a graph. This line of investigation was first initiated by [58]. In contrast, however, persistence diagrams are not invariant to some other group actions, e.g. scale transformations.

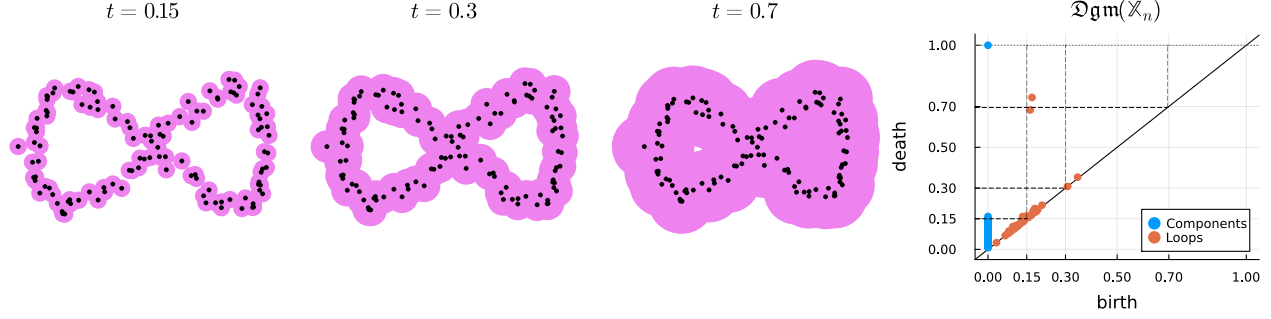


Figure 8: Illustration of a  $\mathcal{Dgm}(\mathbf{X})$  computed on a point-cloud in  $\mathbf{X} \in \mathbb{R}^{n \times d}$  using the Čech filtration.

### D.1 Additional details on Topology-aware clustering

Chazal et al. [13] proposed ToMATo clustering as an algorithm to overcome the drawbacks of  $k$ -means clustering by forming clusters based on topological persistence. However, their methodology [13, Algorithm 1] requires the following tuning parameters: (1) A Kernel  $k$  for density filtering, (2) a bandwidth  $\sigma$ , (2) the cutoff level  $\tau$  of the density filtration, and (3) the resolution  $\delta$  for the Rips complex. While their method is useful for robust clustering in the presence of outliers; we note that under differential privacy, no extraneous outliers are introduced. Therefore, we propose Algorithm 2 as a simple clustering algorithm inspired by ToMATo, which requires just one tuning parameter  $q$ .

Taking  $q = 10$ , we may filter the relevant points in the persistence diagram  $\mathcal{Dgm}(\hat{\mathbf{X}}_n)$ . The filtered points in the  $0^{th}$  order persistence diagram are as shown in Figure 8(a). Similarly, the filtered points in the  $1^{st}$  order persistence diagram are shown in Figure 8(b).

Figure 9(a) shows the representative cycles in the  $1^{st}$  order persistence diagram obtained using these filtered points from Figure 9(b), and corresponds to the two cycles present in the Two Circles latent positions in Figure 7. We can use filtered points in the persistence diagram to cluster the points using Algorithm 2, as shown in Figure 9(b).

---

#### Algorithm 2: Topology-aware Clustering

---

- 1: **Input:** points  $\mathbf{X}_n$  and a cutoff quantile  $q$
  - 2: **Output:** Clusters  $\check{C}$
  - 3: Compute the persistence diagram  $\Delta = \mathcal{Dgm}(\mathbf{X}_n)$  **for**  $i \in |\Delta|$  **do**
  - 4:     **end**  
       Compute the total persistence  $\delta_i = d_i - b_i$  for the birth/death pair  $(b_i, d_i) \in \Delta$
  - 5: Compute the *local outlier factors*  $\mathcal{L} = \{L(\delta_i) : i \in |\Delta|\}$
  - 6: Initialize clusters  $\check{C} \leftarrow \emptyset$  **for**  $i \in |\Delta|$  **do**  
    |      $L(\delta_i) > \text{median}(\mathcal{L}) + q \cdot \text{mad}(\mathcal{L})$   
    **end**
  - 7: Filter the simplex  $\sigma$  with  $(b_i, d_i) = (\text{birth}(\sigma), \text{death}(\sigma))$
  - 8: Extract the vertices  $V(\sigma) = \{\mathbf{x} \in \mathbf{X}_n : \mathbf{x} \subset \sigma\}$
  - 9:  $\check{C} \leftarrow \check{C} \cup V(\sigma)$
  - 10: **return**  $\check{C}$
-

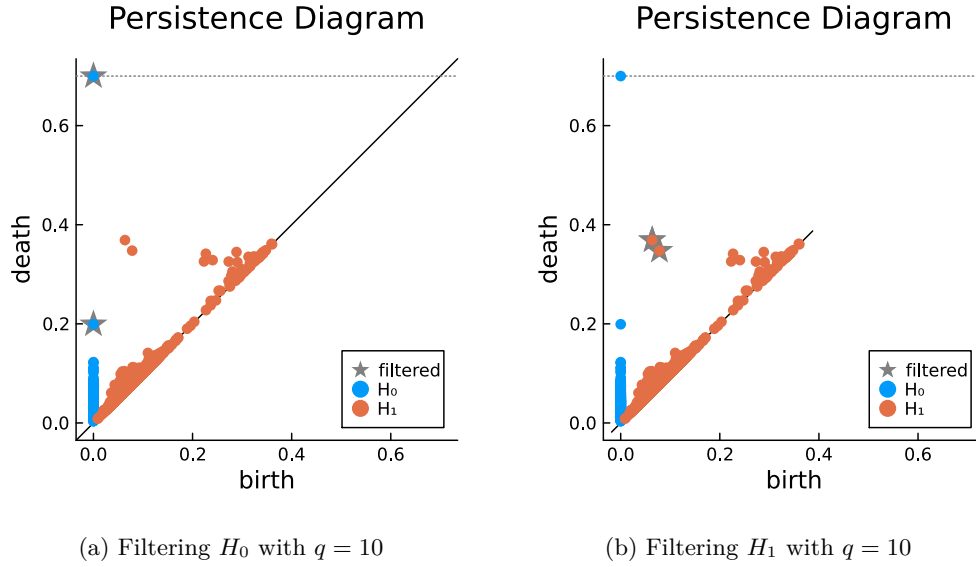


Figure 9: Filtered points in the persistence diagrams for topology-aware clustering for the Two Circles data using Algorithm 2.

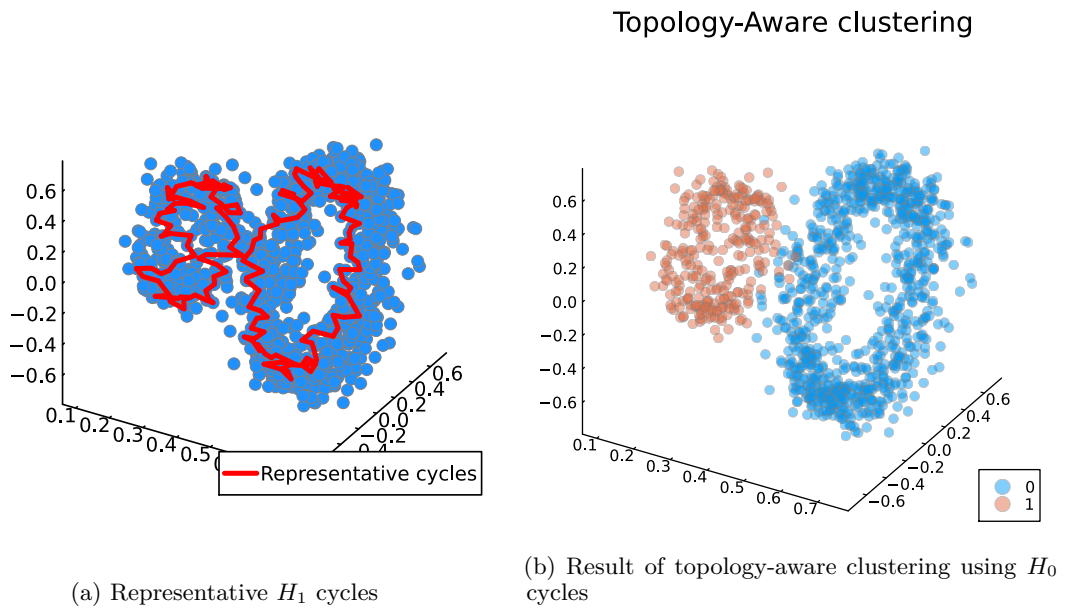


Figure 10: Representative cycles and topology-aware clustering for the Two Circles data.