Estimating peer effects in noisy, low-rank networks via network smoothing

Alex Hayes[1, 2] and Keith Levin[2]

[1]Department of Economics, Stanford University, USA
[2]Department of Statistics, University of Wisconsin-Madison, USA

April 10, 2026

Peer effect estimation requires precise network measurement, yet most empirical networks are noisy, rendering standard estimators inconsistent. To address measurement error in networks, we propose a method to estimate peer effects in networks whose expected adjacency matrix is low-rank. Our key result shows that peer effects over a true unobserved network are asymptotically equivalent to peer effects over the expected adjacency matrix. This result reduces peer effect estimation in noisy networks to low-rank matrix estimation targeting the expected adjacency matrix. We develop our theory for weighted networks observed with additive noise, but the approach can be applied whenever there is a low-rank estimation method suited to the noise structure. We demonstrate via simulations that our approach applies to egocentric samples, aggregated relational data, and networks with missing edges, each requiring a different low-rank estimation method.

1 Introduction

One of the fundamental challenges in estimating social effects, such as contagion, is measuring and quantifying friendships in a reliable manner. These measurement issues pose a fundamental problem, as most current techniques to estimate peer effects in social networks rely on data that precisely encodes potential social dependencies. For instance, in the social sciences, a popular approach is to use the "name generator" method, where each study participant is asked to name their friends or other social ties. This method is replete with complications: participants report different friends depending on the precise language of the name generator (Shakya, Christakis, and Fowler, 2017; Bidart and Charbonneau, 2011), participants often forget day-to-day social contacts (Smieszek et al., 2012), and reports from different participants can disagree (De Bacco et al., 2021).

When friendship measurements are noisy, traditional peer effect estimation runs into a dilemma. Standard network autoregressive models assume that we observe a true, noiseless network, which precisely encodes pathways for peer influence. However, if we only observe a network via noisy measurement, these estimators may produce inconsistent or misleading results. We propose a solution to this problem: replace noisy measurements of adjacency matrices with smoothed, de-noised estimates of network structure. Our approach is motivated by a key theoretical insight: under low-rank network models, contagion over a true, noiseless network is asymptotically equivalent to contagion over a smoothed, latent adjacency matrix.

To formalize matters, let us consider two network autoregressive models. The first is a standard *peer contagion model*, where influence operates over the degree-normalized adjacency matrix:

$$Y_i = \beta_0 + \mathbf{W}_i \boldsymbol{\beta}_\text{w} + \mathbf{X}_i \boldsymbol{\beta}_\text{x} + \beta_\text{y} \sum_{j \neq i} \frac{\mathbf{A}_{ij}}{d_i} Y_j + \varepsilon_i, \tag{1}$$

where Y_i is an outcome for node i, $\mathbf{W}_i$ are observed covariates, $\mathbf{X}_i$ is a latent vector encoding the behavior of node i, $\mathbf{A}_{ij} \in \mathbb{R}$ is the strength of the edge between nodes i and j, and $d_i = \sum_j \mathbf{A}_{ij}$ is the degree of node i. The second model is a *latent contagion model*, where influence operates over a smoothed, expected adjacency matrix:

$$Y_i = \theta_0 + \mathbf{W}_i \boldsymbol{\theta}_\text{w} + \mathbf{X}_i \boldsymbol{\theta}_\text{x} + \theta_\text{y} \sum_{i \neq j} \frac{\mathbb{E}\left[\mathbf{A}_{ij} \,\middle|\, \mathbf{X}\right]}{\mathbb{E}[d_i \mid \mathbf{X}_i]} Y_j + \varepsilon_i. \tag{2}$$

Under a random dot product graph (RDPG; Athreya et al., 2018) model, where $\mathbb{E}\left[\mathbf{A}_{ij} \,\middle|\, \mathbf{X}\right] = \mathbf{X}_i^\top \mathbf{X}_j$, influence is inversely proportional to the cosine distance between nodes in latent space. We develop two-stage least squares estimators for both of these models, and prove that they are consistent under misspecification. That is, estimators derived under the latent contagion model consistently recover peer contagion parameters when the true data generating process follows the peer contagion model, and vice versa. In a formal and quantifiable sense, contagion in peer and latent spaces are (asymptotically) equivalent under a random dot product graph model.

This equivalence has important practical implications. The latent contagion estimators we propose are functions of the nodal covariates and the latent positions $\mathbf{X}$, but not the adjacency matrix $\mathbf{A}$ itself. This means that even when $\mathbf{A}$ is observed with noise, our estimators remain consistent provided we can obtain a sufficiently good estimate $\widehat{\mathbf{X}}$ of $\mathbf{X}$. Crucially, in random dot product graphs, it is often possible to obtain high-quality estimates of $\mathbf{X}$ even when the network is observed with measurement error or when data is missing. We leverage standard tools from spectral network analysis—specifically, the adjacency spectral embedding—to estimate latent positions. Under a sub-gamma network model, which includes binary, weighted, and count-valued networks as special cases, we show that our estimators achieve $\sqrt{n}$ convergence rates and are asymptotically normal. Since these results hold when the observed network contains sub-gamma noise, this enable inference when network measurements are corrupted by random errors. The flexibility of our approach extends beyond the specific case of sub-gamma noise. While we focus on this setting in our theoretical results, the latent contagion framework opens the door to a wide range of robust estimation strategies. Any method that can reliably estimate the low-rank structure of $\mathbb{E}[\mathbf{A}]$—whether through matrix completion for missing data, debiasing for measurement error, or other spectral methods—can be combined with our latent contagion model to estimate peer effects in challenging data settings. We demonstrate this flexibility through simulations and an empirical application studying the contagiousness of smoking in an adolescent social network.

Notation

For a positive integer n, we let $[n] = \{1, 2, \ldots, n\}$. We denote the identity matrix, zero matrix, and the square matrix of all ones by $\mathbf{I}$, $\mathbf{0}$, and $\mathbf{J}$, respectively. For an $n_1 \times n_2$ matrix $\mathbf{H}$, we denote by $\mathbf{H}_{\cdot j}$ the column vector formed by the j-th column of $\mathbf{H}$; and we denote by $\mathbf{H}_{i\cdot}$ the row vector formed by the i-th row of $\mathbf{H}$. For a slight abuse of notation, we also let $\mathbf{H}_i \in \mathbb{R}^{n_2}$ denote the column vector formed by transposing the i-th row of $\mathbf{H}$, that is, $\mathbf{H}_i = (\mathbf{H}_{i\cdot})^\top$. Given any suitably specified ordering on eigenvalues of a square matrix $\mathbf{H}$, we let $\lambda_i(\mathbf{H})$ denote the i-th eigenvalue (under such an ordering) of $\mathbf{H}$ and $\sigma_i(\mathbf{H})$ the i-th singular value of $\mathbf{H}$. We let $\|\mathbf{H}\|$ denote the spectral norm of H and $\|\mathbf{H}\|_F$ denote the Frobenius norm. We let $\|\mathbf{H}\|_{2\to\infty}$ denote the maximum of the Euclidean norms of the rows of H, so that $\|\mathbf{H}\|_{2\to\infty} = \max_i \|\mathbf{H}_i\|$. We denote matrix sequences by $\mathbf{H}_n$, and row vectors by $\mathbf{H}_i$. When referring to the i-th row of a matrix $\mathbf{H}_n$, we write $(\mathbf{H}_n)_i$. We use standard Landau notation for asymptotic rates. Throughout, $C > 0$ will denote a constant independent of n that may change from line to line.

2 Related work

Our work builds on several strands of literature. Most directly related are previous works that model contagion in latent spaces. Sweet and Adhikari (2020) suggested a Hoff model for social influence in a latent space, while Chen, Fan, and Zhu (2023) model correlations between stock returns via the latent space of a stochastic blockmodel in a high-dimensional time series setting. A large literature studies network autoregressions with endogeneity, where the network depends on nodal features. Hayes and Levin (2025) considers identification and asymptotic signal-to-noise ratios in network autoregression models. McFowland and Shalizi (2021) establishes consistency of ordinary least squares in network autoregression models unrolled in time, which Chang and Paul (2024) extends to show asymptotic normality in the longitudinal setting. Paul, Nath, and Warren (2022) presents a network autoregressive model that accounts for estimation error in latent positions, but not in the network itself, considering a quasi-maximum likelihood estimator very similar to ours. Johnsson and Moon (2019) proposes a non-parametric approach to account for endogeneity in cross-sectional networks using sieve estimators, and Egami and Tchetgen Tchetgen (2021) proposes a generalized method of moments estimator using double negative controls to account for contextual effects and homophily.

Most relevant to our focus on measurement error is a growing literature on linear-in-means models with network mismeasurement. Griffith (2022) and Griffith and J. Kim (2024) consider bias when individuals are capped at reporting a maximum number of friends. Boucheron, Lugosi, and Massart (2013) considers linear-in-means models when only partial network data is available but the network distribution is known. Lewbel, Qu, and Tang (2024b) shows that two-stage least squares estimators remain consistent under small amounts of network mismeasurement, while Lewbel, Qu, and Tang (2024a) proposes an adjustment for larger amounts of measurement error. Wenrui Li, Sussman, and Kolaczyk (2022) shows how multiple measurements of a network can be used to estimate and adjust for measurement error rates. Zhang (2024) presents an estimator for mismeasured networks with bounded degrees, and Wei Li, Chakraborty, and Lunde (2025) considers ordinary least squares with peer effects under misspecification.

Related work on causal inference with network misspecification includes Sävje (2024), which considers causal estimation when exposure maps are misspecified, and Hardy et al. (2024), who propose a mixture model over potentially misspecified treatments in linear-in-means models. These works highlight the close relationship between misspecified treatments and mismeasured networks, which can be thought of as conceptual duals. Lewbel, Qu, and Tang (2023) and C. L. Yu et al. (2022) consider estimating spillover effects when the network is entirely unobserved. Spohn, Henckel, and Maathuis (2023), Chin (2019), and Leung (2022) similarly consider linear models for causal interference.

Finally, our work connects to the broader literature on network autoregressive models. LeSage and Pace (2009, Chapter 2) describe how the typical marginal effect interpretation does not apply to linear-in-means models due to non-linearity, and present methods to compute impact scores that retain this interpretation. Vazquez-Bare (2023) discusses when coefficients have a causal interpretation (see also Leung 2022; McFowland and Shalizi 2021). Estimation approaches are given by Ord (1975), Harry H Kelejian and Ingmar R Prucha (1998), Lee (2002), Lee (2003), Lee (2004), Harry H. Kelejian and Ingmar R. Prucha (2007), Lee, Liu, and Lin (2010), Su (2012), Drukker, Egger, and Ingmar R. Prucha (2013), and Lin and Lee (2010) and surveyed in Bivand, Millo, and Piras (2021). Key identification results were given by Bramoullé, Djebbari, and Fortin (2009), and Bramoullé, Djebbari, and Fortin (2020) surveys identification in network autoregressive and linear-in-means models.

3 Models and Estimators

In this section, we formalize two network autoregressive models for peer influence: a standard peer contagion model and our proposed latent contagion model. We then develop two-stage least squares estimators for both models, and establish our main theoretical results showing that these estimators remain consistent under model misspecification.

3.1 Setup

Consider a network with n nodes encoded by a symmetric adjacency matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$. For each node $i \in [n]$, we observe an outcome $Y_i \in \mathbb{R}$ and covariates $\mathbf{W}_i \in \mathbb{R}^p$. We assume each node has an unobserved latent position $\mathbf{X}_i \in \mathbb{R}^d$ that governs how node i forms connections to other nodes. Let $d_i = \sum_{j \neq i} \left|\mathbf{A}_{ij}\right|$ denote the degree of node i. While our development focuses on undirected networks with symmetric adjacency matrices, the framework extends naturally to directed networks. For notational convenience, we define $\mathbf{D} = \text{diag}(d_1, d_2, \ldots, d_n)$ and the degree-normalized adjacency matrix $\mathbf{G} = \mathbf{D}^{-1}\mathbf{A} \in \mathbb{R}^{n\times n}$. If node i is isolated with $d_i = 0$, we set $\mathbf{G}_{i\cdot} = \mathbf{0}$. Similarly, in the latent space, we define $\mathbf{P} = \mathbb{E}[\mathbf{A} \mid \mathbf{X}]$, $\tilde{d}_i = \sum_j \mathbf{P}_{ij} = \mathbb{E}[d_i \mid \mathbf{X}_i]$, $\widetilde{\mathbf{D}} = \text{diag}(\tilde{d}_1, \ldots, \tilde{d}_n)$, and $\widetilde{\mathbf{G}} = \widetilde{\mathbf{D}}^{-1}\mathbf{P} \in \mathbb{R}^{n\times n}$.

3.2 Two Models of Contagion

We now present two complementary specifications for how peer influence propagates through social networks.

Peer contagion model. The first specification is a standard spatial autoregressive model adapted to networks via latent positions:

$$Y_i = \beta_0 + \mathbf{W}_i\boldsymbol{\beta}_{\text{w}} + \mathbf{X}_i\boldsymbol{\beta}_{\text{x}} + \beta_{\text{y}} \sum_{j\neq i} \frac{\mathbf{A}_{ij}}{d_i} Y_j + \varepsilon_i \tag{3}$$

$$\mathbf{Y} = \left(\mathbf{I} - \beta_{\text{y}}\mathbf{G}\right)^{-1} (\mathbf{1}_n\beta_0 + \mathbf{W}\boldsymbol{\beta}_{\text{w}} + \mathbf{X}\boldsymbol{\beta}_{\text{x}} + \boldsymbol{\varepsilon}) \tag{4}$$

Here, $\beta_{\text{y}} \in \mathbb{R}$ measures how peer outcomes influence focal outcomes through the observed network structure. The latent positions $\mathbf{X}_i$ model homophily, which is crucial for both statistical identification (Hayes and Levin, 2025) and causal inference (Shalizi and Thomas, 2011). We call this the *peer contagion* model because influence operates over the degree-normalized adjacency matrix $\mathbf{G}$.

Latent contagion model. Our proposed alternative replaces the observed adjacency matrix with its conditional expectation:

$$Y_i = \theta_0 + \mathbf{W}_i\boldsymbol{\theta}_{\text{w}} + \mathbf{X}_i\boldsymbol{\theta}_{\text{x}} + \theta_{\text{y}} \sum_{j\neq i} \frac{\mathbb{E}\left[\mathbf{A}_{ij} \,\middle|\, \mathbf{X}\right]}{\mathbb{E}[d_i \mid \mathbf{X}_i]} Y_j + \varepsilon_i \tag{5}$$

$$\mathbf{Y} = \left(\mathbf{I} - \theta_{\text{y}}\widetilde{\mathbf{G}}\right)^{-1} (\mathbf{1}_n\theta_0 + \mathbf{W}\boldsymbol{\theta}_{\text{w}} + \mathbf{X}\boldsymbol{\theta}_{\text{x}} + \boldsymbol{\varepsilon}) \tag{6}$$

Under a random dot product graph model, where $\mathbb{E}\left[\mathbf{A}_{ij} \,\middle|\, \mathbf{X}_i, \mathbf{X}_j\right] = \mathbf{X}_i^\top \mathbf{X}_j$, the latent contagion model takes a particularly interpretable form:

$$Y_i = \theta_0 + \mathbf{W}_i\boldsymbol{\theta}_{\text{w}} + \mathbf{X}_i\boldsymbol{\theta}_{\text{x}} + \theta_{\text{y}} \sum_{j\neq i} \frac{\mathbf{X}_i^\top \mathbf{X}_j}{\sum_k \mathbf{X}_i^\top \mathbf{X}_k} Y_j + \varepsilon_i. \tag{7}$$

In this specification, peer influence is inversely proportional to the cosine distance between nodes in latent space. All pairs of nodes exert influence on one another, with the strength of influence determined by latent proximity rather than the realization of individual edges. The latent contagion model reflects a fundamentally different view of social influence: rather than contagion traveling along realized edges, it diffuses based on the propensity for connection. Nodes close in latent space exert strong influence regardless of whether a specific edge forms, while distant nodes exert negligible influence.

Both the peer and latent contagion models require a stochastic model for how the network $\mathbf{A}$ is generated. We adopt a flexible sub-gamma framework that encompasses binary, weighted, and count-valued networks.

Definition 1 (Sub-gamma network model)**.** Let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be a random symmetric matrix. Let $\mathbf{P} = \mathbb{E}[\mathbf{A} \,|\, \mathbf{X}]$ be the expectation of $\mathbf{A}$ conditional on $\mathbf{X} \in \mathbb{R}^{n \times d}$, which has independent and identically distributed rows $\mathbf{X}_1, \dots, \mathbf{X}_n$. Assume $\mathbf{P}$ has rank d and is positive semi-definite with eigenvalues $\mathbf{s}_1 \geq \mathbf{s}_2 \geq \cdots \geq \mathbf{s}_d > 0 = \mathbf{s}_{d+1} = \cdots = \mathbf{s}_n$. Conditional on $\mathbf{X}$, the upper-triangular elements of $\mathbf{A} - \mathbf{P}$ are independent (ν_n, b_n)-sub-gamma random variables.

The sub-gamma family is broad, including Bernoulli, Poisson, Exponential, Gamma, and Gaussian distributions, as well as all bounded distributions (Boucheron, Lugosi, and Massart, 2013; Vershynin, 2020). We provide a formal definition of sub-gamma random variables along with a handful of related technical results in Appendix A. The generality of this framework allows us to handle diverse edge types: binary friendships, weighted interaction frequencies, or count-valued communication volumes. As a consequence of this generality, our results require a comparatively high network density; in parallel work we show density requirements are less stringent in binary networks.

3.3 Identification

Identification in spatial autoregressive models is subtle, due to their conditional specification via $\mathbf{Y}_i \mid \mathbf{Y}_{-i}, \mathbf{W}, \mathbf{X}, \mathbf{A}$. The model can equivalently be written in its total law form (Equations 4 and 6), as is standard in the Markov random field literature (Besag, 1974; Rue, 2005).

Lemma 1. *If* $|\beta| < 1$*, then* $\mathbf{I} - \beta\mathbf{G}$ *is invertible with all eigenvalues in the interval* $(1 - \beta, 1 + \beta)$*.*

Proof. Since $\mathbf{G} = \mathbf{D}^{-1}\mathbf{A}$ is row stochastic, all its eigenvalues have absolute value at most 1 by the Gershgorin circle theorem. Therefore, all eigenvalues of $\beta\mathbf{G}$ have absolute value at most $|\beta|$, implying all eigenvalues of $\mathbf{I} - \beta\mathbf{G}$ lie in $1 \pm \beta$ and are bounded away from zero. □

The following identification result is crucial for both models.

Lemma 2 (Martellosio 2022)**.** *Consider a network with fixed n nodes. For the peer contagion model* (4)*, suppose* $\mathbb{E}[\boldsymbol{\varepsilon} \mid \mathbf{W}, \mathbf{X}, \mathbf{G}] = \mathbf{0}$ *and* $\left|\beta_{\text{y}}\right| < 1$*. Then* $\beta_0, \boldsymbol{\beta}_{\text{w}}, \boldsymbol{\beta}_{\text{x}}, \beta_{\text{y}}$ *are identified if and only if*

$$\operatorname{rank}\left[\mathbf{1}_n \ \mathbf{W} \ \mathbf{X} \ \mathbf{GW} \ \mathbf{GX}\right] > \operatorname{rank}\left[\mathbf{1}_n \ \mathbf{W} \ \mathbf{X}\right]$$

An analogous result holds for the latent contagion model (6) *with* $\widetilde{\mathbf{G}}$ *replacing* $\mathbf{G}$*, under the additional assumption that that* $d \geq 2$*, which rules out collinearity between* $\mathbf{X}$ *and* $\widetilde{\mathbf{G}}\mathbf{X}$*.*

Remark 1. The latent positions $\mathbf{X}$ are only identified up to an orthogonal transformation $\mathbf{Q}$, which implies $\boldsymbol{\beta}_{\text{x}}$ is also only identified up to $\mathbf{Q}$ (Hayes and Levin, 2025). In some models, $\mathbf{1}_n$ may lie in the column space of $\mathbf{X}$, in which case the intercept should be dropped.

Remark 2. Linear-in-means models can suffer from asymptotic degeneracy when $\mathbf{GY}$ (or $\widetilde{\mathbf{G}}\mathbf{Y}$) becomes collinear with $\left[\mathbf{1}_n \ \mathbf{W} \ \mathbf{X}\right]$, leading to parameters that are identified but inestimable (Hayes and Levin, 2025). We assume throughout that the network structure prevents such degeneracy. For random dot product graphs, this requires either sufficient sparsity to prevent concentration of $\mathbf{GY}$ around its expectation, or sufficient degree heterogeneity to ensure linear independence. See Example 4 of Hayes and Levin (2025) for details.

3.4 Estimators

To construct feasible estimators, we require an estimate $\widehat{\mathbf{X}}$ of the latent positions $\mathbf{X}$. We use the adjacency spectral embedding (Sussman, Minh Tang, and Priebe, 2014), a spectral estimate appropriate both for precisely observed networks and networks observed with additive noise. Other types of measurement error will require distinct estimators of $\mathbf{X}$; we leave detailed investigation of these other settings to future work.

Definition 2 (Adjacency spectral embedding)**.** Given a network $\mathbf{A}$, the d-dimensional *adjacency spectral embedding* is $\widehat{\mathbf{X}} = \widehat{\mathbf{U}}\hat{\mathbf{S}}^{1/2} \in \mathbb{R}^{n\times d}$, where $\widehat{\mathbf{U}}\hat{\mathbf{S}}\widehat{\mathbf{V}}^\top$ is the rank-d truncated singular value decomposition of $\mathbf{A}$.

The adjacency spectral embedding provides a consistent estimate of $\mathbf{X}$ under mild sparsity conditions (Athreya et al., 2018).

Lemma 3 (Levin, Athreya, et al. 2019)**.** *Under suitable regularity conditions, there exists a $d \times d$ orthogonal matrix $\mathbf{Q}$ such that*

$$\max_{i\in[n]} \left\|\mathbf{Q}\widehat{\mathbf{X}}_i - \mathbf{X}_i\right\| = o_p(1).$$

We now define plug-in estimators that replace $\mathbf{X}$ with $\widehat{\mathbf{X}}$ wherever it appears.

Definition 3 (Peer contagion estimator)**.** Let $\breve{\mathbf{Z}} = \left[\mathbf{1}_n \ \mathbf{W} \ \widehat{\mathbf{X}} \ \mathbf{Gy}\right] \in \mathbb{R}^{n\times(p+d+2)}$ and $\breve{\mathbf{H}} = \left[\mathbf{W} \ \widehat{\mathbf{X}} \ \mathbf{GW} \ \mathbf{G}\widehat{\mathbf{X}} \ \mathbf{G}^2\mathbf{W} \ \mathbf{G}^2\widehat{\mathbf{X}}\right] \in \mathbb{R}^{n\times(3p+3d)}$. Let $\breve{\mathbf{M}} = \breve{\mathbf{H}}(\breve{\mathbf{H}}^\top\breve{\mathbf{H}})^{-1}\breve{\mathbf{H}}^\top$ denote the projection matrix onto the column space of $\breve{\mathbf{H}}$. Then we define our peer contagion estimator according to

$$\hat{\boldsymbol{\beta}} = (\breve{\mathbf{Z}}^\top\breve{\mathbf{M}}\breve{\mathbf{Z}})^{-1}\breve{\mathbf{Z}}^\top\breve{\mathbf{M}}\mathbf{y}. \tag{8}$$

For the latent contagion estimator, we construct $\hat{\mathbf{P}} = \widehat{\mathbf{X}}\widehat{\mathbf{X}}^\top$, $\hat{d}_i = \sum_j \hat{\mathbf{P}}_{ij}$, $\widehat{\mathbf{D}} = \operatorname{diag}(\hat{d}_1, \dots, \hat{d}_n)$, and

$$\widehat{\mathbf{G}} = \widehat{\mathbf{D}}^{-1}\hat{\mathbf{P}} \in \mathbb{R}^{n\times n}. \tag{9}$$

Definition 4 (Latent contagion estimator)**.** Let $\widehat{\mathbf{Z}} = \left[\mathbf{1}_n \ \mathbf{W} \ \widehat{\mathbf{X}} \ \widehat{\mathbf{G}}\mathbf{y}\right] \in \mathbb{R}^{n\times(p+d+2)}$ and $\widehat{\mathbf{H}} = \left[\mathbf{W} \ \widehat{\mathbf{X}} \ \widehat{\mathbf{G}}\mathbf{W} \ \widehat{\mathbf{G}}\widehat{\mathbf{X}} \ \widehat{\mathbf{G}}^2\mathbf{W} \ \widehat{\mathbf{G}}^2\widehat{\mathbf{X}}\right] \in \mathbb{R}^{n\times(3p+3d)}$. Let $\widehat{\mathbf{M}} = \widehat{\mathbf{H}}(\widehat{\mathbf{H}}^\top\widehat{\mathbf{H}})^{-1}\widehat{\mathbf{H}}^\top$ denote the projection matrix onto the column space of $\widehat{\mathbf{H}}$. Then our latent contagion estimator is given by

$$\hat{\boldsymbol{\theta}} = (\widehat{\mathbf{Z}}^\top\widehat{\mathbf{M}}\widehat{\mathbf{Z}})^{-1}\widehat{\mathbf{Z}}^\top\widehat{\mathbf{M}}\mathbf{y}. \tag{10}$$

In some cases, $\breve{\mathbf{H}}$ and $\widehat{\mathbf{H}}$ will have collinear columns; any subset of columns with rank $p + d + 1$ or greater is sufficient for identification (see Bramoullé, Djebbari, and Fortin, 2009, for details on the construction of the instruments matrix).

3.5 Main Results

We begin with some preliminary theoretical results, which show that under correctly-specified models, peer and latent contagion models that adjust for latent positions can be estimated by plugging in $\widehat{\mathbf{X}}$ as an estimate for $\mathbf{X}$. These results shows that it is possible to distinguish between localized effects on a network, as parameterized by $\boldsymbol{\beta}_{\mathrm{x}}$ and $\boldsymbol{\theta}_{\mathrm{x}}$, and diffusions, as parameterized by β_{y} and θ_{y}.

Theorem 1 (Peer contagion estimators under peer contagion). *Suppose the data are generated according to peer contagion as in Equation (4) and* $\mathbf{A}$ *follows a sub-gamma model as in Definition 1. Under Assumptions 1, 2, 3, and 5, detailed in the Appendix, there exists a sequence of orthogonal matrices* $\mathbf{Q}_{\mathbf{Z}} \in \mathbb{R}^{(p+d+2)\times(p+d+2)}$ *such that*

$$\sqrt{n}\left(\mathbf{Q}_{\mathbf{Z}}\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \to \mathcal{N}\left(\mathbf{0}, \Sigma_{\beta}\right)$$

where $\boldsymbol{\beta} = (\beta_0, \boldsymbol{\beta}_{\mathrm{w}}^\top, \boldsymbol{\beta}_{\mathrm{x}}^\top, \beta_{\mathrm{y}})^\top$.

Theorem 2 (Latent contagion estimators under latent contagion). *Suppose the data are generated according to latent contagion as in Equation (6) and* $\mathbf{A}$ *follows a sub-gamma model as in Definition 1. Under Assumptions 1, 2, 3, 4, and 6, detailed in the Appendix, there exists a sequence of orthogonal matrices* $\mathbf{Q}_{\mathbf{Z}} \in \mathbb{R}^{(p+d+2)\times(p+d+2)}$ *such that*

$$\sqrt{n}\left(\mathbf{Q}_{\mathbf{Z}}\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) \to \mathcal{N}(\mathbf{0}, \Sigma_{\theta}),$$

where $\boldsymbol{\theta} = (\theta_0, \boldsymbol{\theta}_{\mathrm{w}}^\top, \boldsymbol{\theta}_{\mathrm{x}}^\top, \theta_{\mathrm{y}})^\top$.

Our central theoretical contribution establishes that estimators derived under one model remain consistent when the true data generating process follows the other model. This equivalence holds under random dot product graphs.

Theorem 3 (Latent contagion estimators under peer contagion). *Suppose the data are generated according to peer contagion as in Equation (4) but we use latent contagion estimators from Definition 4. Suppose* $\mathbf{A}$ *follows a sub-gamma model as in Definition 1. Under Assumptions 1, 2, 3, 4 and 5, detailed in the Appendix, there exists a sequence of orthogonal matrices* $\mathbf{Q}_{\mathbf{Z}} \in \mathbb{R}^{(p+d+2)\times(p+d+2)}$ *such that*

$$\sqrt{n}\left(\mathbf{Q}_{\mathbf{Z}}\hat{\boldsymbol{\theta}} - \boldsymbol{\beta}\right) \to \mathcal{N}(\mathbf{0}, \Sigma_{\theta}).$$

Theorem 4 (Peer contagion estimators under latent contagion). *Suppose the data are generated according to latent contagion as in Equation (6) but we use peer contagion estimators from Definition 3. Suppose* $\mathbf{A}$ *follows a sub-gamma model as in Definition 1. Under Assumptions 1, 2, 3 and 6, detailed in the Appendix, there exists a sequence of orthogonal matrices* $\mathbf{Q}_{\mathbf{Z}} \in \mathbb{R}^{(p+d+2)\times(p+d+2)}$ *such that*

$$\sqrt{n}\left(\mathbf{Q}_{\mathbf{Z}}\hat{\boldsymbol{\beta}} - \boldsymbol{\theta}\right) \to \mathcal{N}\left(\mathbf{0}, \Sigma_{\beta}\right).$$

The proceeding two theorems establish that ordinary least squares and two-stage least squares estimators for peer effects in the peer and latent contagion models are both consistent and asymptotically normal estimators.

Proofs of Theorems 1, 2, 3 and 4 can be found in the Appendix. The general proof strategy is to show that $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$ are close to "oracle" estimates based on using the true latent positions. Appendix H discusses convergence of these oracle estimators to the true parameters. Appendices I and J show convergence of our estimators defined above to these oracle estimators.

We are, in fact, able to generalize these results to a statement about asymptotic equivalence of the corresponding functionals. Recall that the popular design matrices of the peer and latent contagion models are, respectively,

$$\mathbf{Z} = \left[\mathbf{1}_n \ \mathbf{W} \ \mathbf{X} \ \mathbf{G}\mathbf{y}\right] \ \text{ and } \ \widetilde{\mathbf{Z}} = \left[\mathbf{1}_n \ \mathbf{W} \ \mathbf{X} \ \widetilde{\mathbf{G}}\mathbf{y}.\right]. \tag{11}$$

Consider the corresponding "projection parameters", or the population coefficients corresponding to projection of outcomes onto these design matrices, and letting F be an appropriately supported cumulative distribution function (see Wei Li, Chakraborty, and Lunde 2025 for a discussion of these parameters in the network regression context),

$$\tau(F) = \operatorname*{argmin}_b \mathbb{E}_F\left[(\mathbf{Y} - \mathbf{Z}\, b)^2\right] = \mathbb{E}_F\left[\mathbf{Z}^\top\mathbf{Z}\right]^{-1}\mathbb{E}_F\left[\mathbf{Z}^\top\mathbf{Y}\right], \text{ and} \tag{12}$$

$$\tilde{\tau}(F) = \operatorname*{argmin}_t \mathbb{E}_F\left[\left(\mathbf{Y} - \widetilde{\mathbf{Z}}\, t\right)^2\right] = \mathbb{E}_F\left[\widetilde{\mathbf{Z}}^\top\widetilde{\mathbf{Z}}\right]^{-1}\mathbb{E}_F\left[\widetilde{\mathbf{Z}}^\top\mathbf{Y}\right]. \tag{13}$$

Let F_β denote the law of the peer contagion process and F_θ the law of the latent contagion process. When the expectation is taken with respect to the true generating process, the projection parameters simplify to the corresponding regression coefficients. Letting $\mathbb{E}_\beta$ denote expectations under F_β and $\mathbb{E}_\theta$ denote expectations under F_θ,

$$\begin{aligned} \tau(F_\beta) &= \mathbb{E}_\beta\left[\mathbf{Z}^\top\mathbf{Z}\right]^{-1}\mathbb{E}_\beta\left[\mathbf{Z}^\top\mathbb{E}_\beta\left[\mathbf{Y}\,|\,\mathbf{Z}\right]\right] = \mathbb{E}_\beta\left[\mathbf{Z}^\top\mathbf{Z}\right]^{-1}\mathbb{E}_\beta\left[\mathbf{Z}^\top\mathbf{Z}\boldsymbol{\beta}\right] = \boldsymbol{\beta}, \quad \text{and} \\ \tilde{\tau}(F_\theta) &= \mathbb{E}_\theta\left[\widetilde{\mathbf{Z}}^\top\widetilde{\mathbf{Z}}\right]^{-1}\mathbb{E}_\theta\left[\widetilde{\mathbf{Z}}^\top\mathbb{E}_\theta\left[\mathbf{Y}\,|\,\widetilde{\mathbf{Z}}\right]\right] = \mathbb{E}_\theta\left[\widetilde{\mathbf{Z}}^\top\widetilde{\mathbf{Z}}\right]^{-1}\mathbb{E}_\theta\left[\widetilde{\mathbf{Z}}^\top\widetilde{\mathbf{Z}}\boldsymbol{\theta}\right] = \boldsymbol{\theta}. \end{aligned} \tag{14}$$

These projection parameters are not, in general, equal under the latent and peer contagion models. However, fixing either the peer or latent contagion model, the following theorem shows that they are asymptotically equivalent, in the sense that the difference between the two parameters is negligible relative to the typical estimation error.

Theorem 5. *Under Assumptions 6, 7, 8 and 9, and supposing the latent contagion model in Equation (6) holds, then*

$$\|\tilde{\tau}(F_\theta) - \tau(F_\theta)\| = \|\boldsymbol{\theta} - \tau(F_\theta)\| = \left\|\mathbb{E}_\theta\left[\widetilde{\mathbf{Z}}^\top\widetilde{\mathbf{Z}}\right]^{-1}\mathbb{E}_\theta\left[\widetilde{\mathbf{Z}}^\top\mathbf{Y}\right] - \mathbb{E}_\theta\left[\mathbf{Z}^\top\mathbf{Z}\right]^{-1}\mathbb{E}_\theta\left[\mathbf{Z}^\top\mathbf{Y}\right]\right\| = o(n^{-1/2}).$$

If the peer contagion model in Equation (4) holds instead of the model in Equation (6), then, with Assumption 5 in place of Assumption 6,

$$\left\|\tau(F_\beta) - \tilde{\tau}(F_\beta)\right\| = \left\|\boldsymbol{\beta} - \tilde{\tau}(F_\beta)\right\| = \left\|\mathbb{E}_\beta\left[\mathbf{Z}^\top\mathbf{Z}\right]^{-1}\mathbb{E}_\beta\left[\mathbf{Z}^\top\mathbf{Y}\right] - \mathbb{E}_\beta\left[\widetilde{\mathbf{Z}}^\top\widetilde{\mathbf{Z}}\right]^{-1}\mathbb{E}_\beta\left[\widetilde{\mathbf{Z}}^\top\mathbf{Y}\right]\right\| = o(n^{-1/2}).$$

A proof can be found in Appendix F. In the above theorem, we use the "assumption lean" representation of regression coefficients, or the "projection estimands", which are the definitions of τ and $\tilde{\tau}$ given in (12) and (13), respectively. These parameters are non-parametric. That is, $\tilde{\tau}$ is a projection parameter that one might want to estimate even if the true data generating model is the peer contagion model and $\boldsymbol{\theta}$ does not index the generating model, such that $\boldsymbol{\theta}$ is undefined. The key idea of Theorem 5 is that the true parameter and the projection parameter are within parametric estimation error of one another, so they are functionally indistinguishable in the asymptotic limit.

This result has substantial implications for estimating contagion effects in noisy networks. Suppose that outcomes are generated according to the peer contagion model in Equation (3), but the network $\mathbf{A}$ is unobserved and one only has access to a noisy variant $\mathbf{N}$. In this case, treating $\mathbf{N}$ as the true network and plugging it into typical estimators such as $\hat{\boldsymbol{\beta}}$ will ignore measurement error in $\mathbf{N}$ and typically lead to inconsistent estimation, amongst other issues. In particular, it is challenging to target the parameter $\boldsymbol{\beta}$, because this requires projecting onto $\mathbf{Z}$, which is a function of an unknown network $\mathbf{A}$. Theorem 5, however, suggests a path forward: we can instead target the parameter $\tilde{\tau}$, which does not equal $\boldsymbol{\beta}$, but is functionally indistinguishable. Crucially, to estimate $\tilde{\tau}$ we do not need to project outcomes onto $\mathbf{Z}$, but rather onto $\widetilde{\mathbf{Z}}$, depends on the low-rank expectation $\mathbf{P}$ rather than the precise network $\mathbf{A}$. This is a substantially easier task: we do not need $\mathbf{A}$ proper, only its principal subspace. And, conveniently, given a noisy observation of a network $\mathbf{N}$, there are many off-the-shelf methods to estimate the principal subspace of $\mathbf{A}$.

This leads to a potentially generic recipe for estimating contagion effects in noisy networks: find a subspace estimator tailored to the noise process, and target $\tilde{\tau}$ using an estimator like $\hat{\boldsymbol{\theta}}$. The variance of the corresponding estimator will depend on how fast the subspace estimator converges. For sufficiently fast estimators, such as the adjacency spectral embedding, there is no asymptotic variance penalty due to estimating the principal subspace. In fact, this approach and our results thus far are sufficient to develop an estimator for contagion effects in noisy weighted networks. Suppose that $\mathbf{N} = \mathbf{A} + \mathbf{E}$, where $\mathbf{E}$ is mean-zero sub-gamma noise. In this case, $\mathbf{N}$ still satisfies the assumptions of the sub-gamma network model (Definition 1), and thus the adjacency spectral embedding is still a consistent estimator of the principal subspace of $\mathbf{A}$, with only a slight adjustment to the sub-gamma parameters in the convergence rate. Thus, $\hat{\boldsymbol{\theta}}$ immediately accommodates sub-gamma noise in the adjacency matrices.

Figure 1 offers some visual intuition for the generic nature of the network smoothing strategy, showing estimates of $\mathbf{P}$ obtained from networks with various types of corruption. Even when 10% of edges are missing, matrix completion methods produce estimates $\hat{\mathbf{P}}$ that closely approximate the true latent structure. In the subsequent section, we provide evidence via simulations that network smoothing is a viable strategy to estimate contagion in networks with missing edges, in ego-centrically sampled networks, and in networks where only aggregated relational data is available.

4 Simulation study

We now verify via simulation that $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\beta}}$ are consistent and asymptotically normal estimates of contagion effects (Fig. 2), and that they obtain the expected $n^{-1/2}$ convergence rate predicted by our theoretical results. All networks in our simulations below are gener-

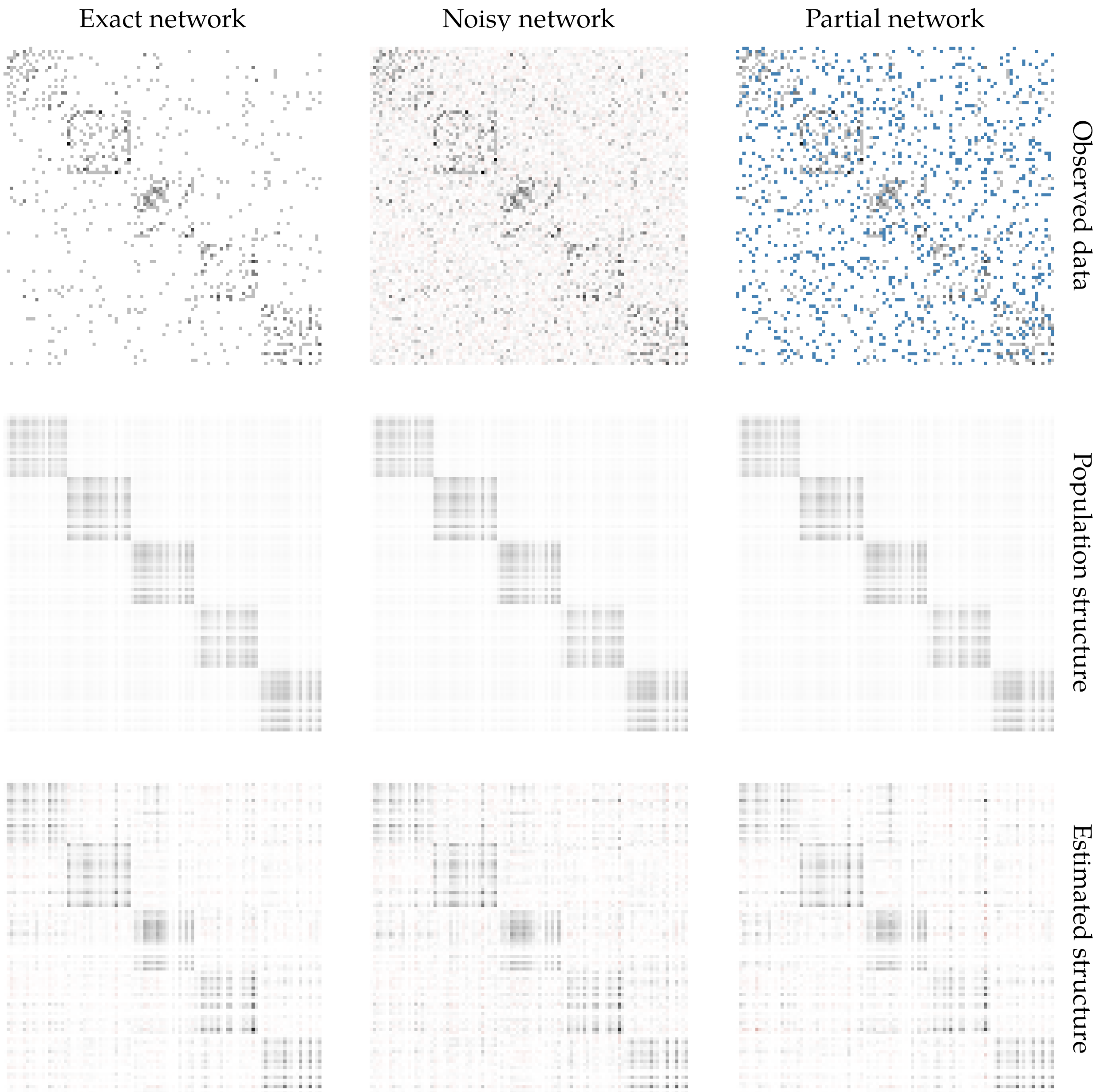

Figure 1: Adjacency matrices, corresponding expectations, and estimated expectations for observed network data corresponding to the same underlying population structure in a random dot product graph, as well as various estimates of the latent structure, based on the adjacency spectral embedding in the left and middle columns, and nuclear-norm based matrix completion in the right column. Blue entries in the right column indicate unobserved entries of **A**.

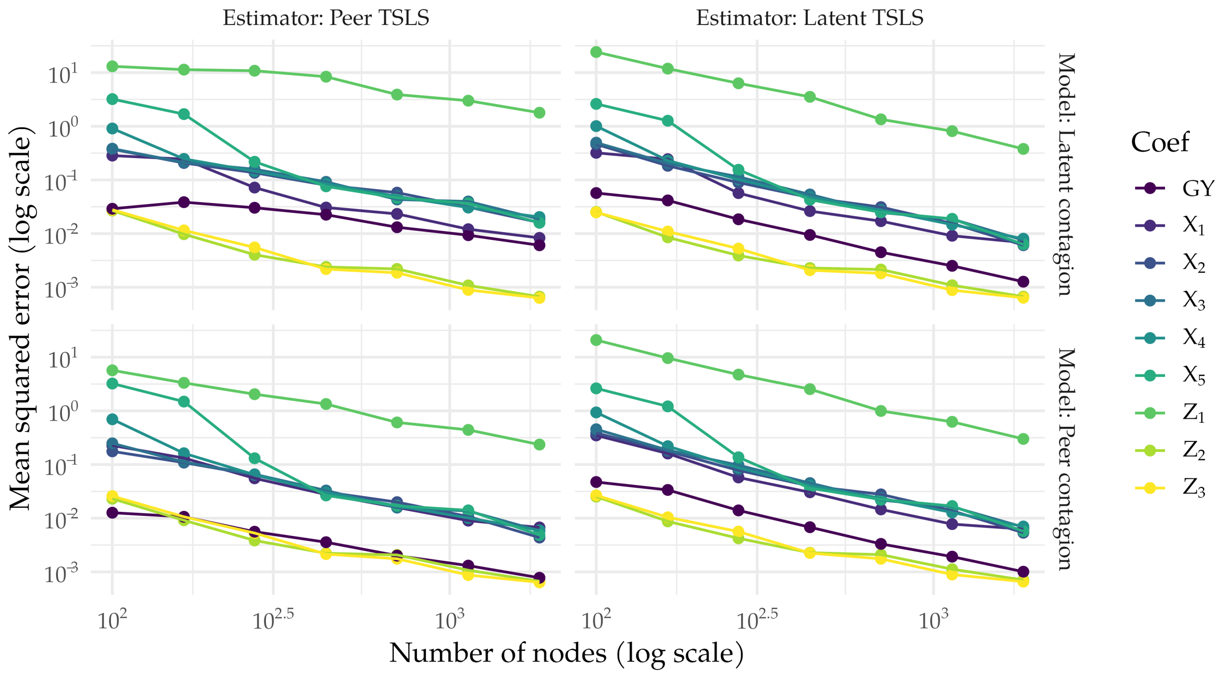

Figure 2: Monte Carlo estimates of mean squared error of $\hat{\boldsymbol{\theta}}$, and $\hat{\boldsymbol{\beta}}$. The average degree in these simulations is $n^{3/4}$.

ated from a Poisson degree-corrected stochastic blockmodel (Definition 5) with n nodes and five equally probable blocks.

Definition 5 (Poisson Degree-Corrected Stochastic Blockmodel)**.** The Poisson degree-corrected stochastic blockmodel (Rohe et al., 2018; Karrer and Newman, 2011) is an undirected model of community membership, with d communities. Each node, indexed by $i = 1, 2, \dots, n$, is assigned a block $z_i \in \{1, 2, \dots, d\}$ with probability $\Pr(z_i = k) = \pi_k$, and a degree-correction parameter ξ_i, which describes the propensity of vertex i to connect with other nodes. Conditional on block memberships and degree-correction parameters, edges are generated independently between every pair of vertices in the network according to a Poisson distribution. The expected number of edges between two vertices depends on their community memberships, their degree correction parameters, a positive semi-definite matrix $\mathbf{B} \in [0, 1]^{d \times d}$ of inter-block edge formation probabilities, and a scaling factor $\rho_n \in [0, 1]$, which may vary with n. That is,

$$\mathbb{E}\Big[\mathbf{A}_{ij} = 1 \,\Big|\, z_i, z_j, \xi_i, \xi_j\Big] = \rho_n \, \xi_i \mathbf{B}_{z_i, z_j} \xi_j.$$

For our simulation study, we take there to be $d = 5$ blocks, and sample diagonal elements of $\mathbf{B}$ from a $\text{Uniform}(0.75, 0.85)$ distribution and off-diagonal elements of $\mathbf{B}$ from a $\text{Uniform}(0.01, 0.05)$ distribution, such that networks are strongly assortatively, mostly forming edges within blocks. We sample degree correction parameters according to $\xi_i \sim \text{Exponential}(1/3) + 1$. Shifting the distribution of ξ_i away from zero limits the number of isolated nodes. The sparsity parameter ρ_n is set so that the expected mean degree of the network is $n^{3/4}, n^{1/2}$ or $n^{1/4}$. Once we have generated these parameters, we compute $\mathbf{X} = \mathbf{U}\mathbf{S}^{1/2}$ where $\mathbf{U}\mathbf{S}\mathbf{U}^\top$ is the eigendecomposition of the low-rank expectation $\mathbb{E}[A \mid z_1, z_2, \ldots, z_n, \theta]$. Errors $\boldsymbol{\varepsilon}$ are sampled from a standard normal distribution, and we include three covariates $\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3$, also sampled from a standard normal. We set $\beta_0 = \theta_0 = 0$, $\beta_\mathrm{y} = \theta_\mathrm{y} = 0.2$, $\boldsymbol{\beta}_\mathrm{w} = \boldsymbol{\theta}_\mathrm{w} = (5, 5, 5)$ and $\boldsymbol{\beta}_\mathrm{x} = \boldsymbol{\theta}_\mathrm{x} = (2, 2, 2, 2, 2)$. Then $\mathbf{Y}$ is generated according to eq. (4) and eq. (6), respectively.

For both models, we compute $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$, and measure the estimation error to the corresponding model coefficients. Since $\boldsymbol{\beta}_\mathrm{x}$ is only identified up to orthogonal rotation, we Procrustes align $\widehat{\mathbf{X}}$ with $\mathbf{X}$ to investigate recover of $\boldsymbol{\beta}_\mathrm{x}$ (we set $\beta_0 = \theta_0 = 0$ to simplify this Procrustes alignment). In our experiments, we vary the sample size n (i.e., the number of vertices) on a logarithmic scale, considering $n \in \{100, 163, 264, 430, 698, 1135, 1845, 3000\}$, and replicate our experiments 100 times for each simulation setting. Figure 2 shows the mean squared error of the estimated coefficients as a function of the number of nodes. Mean squared error for $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\beta}}$ decreases at $n^{-1/2}$ rates in both the latent and peer contagion models, exactly as dictated by our theory.

4.1 Noisy networks

We additionally investigate how network smoothing performs when the network is mismeasured. We consider the exact same simulation setting as before, but now observe a noisy adjacency matrix $\mathbf{N}$ rather than $\mathbf{A}$. We consider estimators $\hat{\boldsymbol{\theta}}$. In some cases, we use an alternative estimator of the singular vectors $\mathbf{U}$ and singular values $\mathbf{S}$ that is more appropriate for the particular noise process. We consider the following noise processes and corresponding estimators. In the first four cases, existing estimators are capable of recovering the singular values and singular vectors of $\mathbf{P}$. In the last two settings, we are unaware of principal subspace estimators, and expect difficulties with estimation.

1. **Baseline**: $\mathbf{N} = \mathbf{A}$ and $\mathbf{U}$ and $\mathbf{S}$ are estimated via singular value decomposition of $\mathbf{N}$, as in $\hat{\boldsymbol{\theta}}$. The estimated singular vectors and values are simply substituted for the relevant population quantities during estimation: $\widehat{\mathbf{X}} = \widehat{\mathbf{U}}\hat{\mathbf{S}}$, $\hat{\mathbf{P}} = \widehat{\mathbf{U}}\hat{\mathbf{S}}\widehat{\mathbf{U}}^\top$, $\hat{d}_i = \sum_{j \neq i} \hat{\mathbf{P}}$ and $\widehat{\mathbf{G}} = \widehat{\mathbf{D}}^{-1}\hat{\mathbf{P}}$.

2. **Gaussian noise**: $\mathbf{N}_{ij} = \mathbf{A}_{ij} + \boldsymbol{\varepsilon}_{ij}$ where $\boldsymbol{\varepsilon}_{ij}$ is i.i.d $\mathcal{N}(0, \sigma^2)$ noise and $\sigma^2 = 0.1$. $\mathbf{U}$ and $\mathbf{S}$ are again estimated via singular value decomposition of $\mathbf{N}$.

3. **Missing edges**: $\mathbf{N} = \mathbf{A}$, except 30% of entries of $\mathbf{N}$ are missing at random, with missingness independent of the network. $\mathbf{U}$ and $\mathbf{S}$ are estimated via matrix completion, in particular the `AdaptiveImpute` algorithm of Cho, D. Kim, and Rohe (2019). Estimated singular values and vectors are plugged in as the baseline case to construct estimates of $\mathbf{P}$ and $\widetilde{\mathbf{G}}$.

4. **Aggregated relational data**: $\mathbf{N} = \mathbf{A}\mathbf{W}$ where $\mathbf{W}$ are traits (also observed) that are correlated with latent $\mathbf{X}$. $\mathbf{W}$ has the same dimensions as $\mathbf{X}$, and each column of $\mathbf{W}$ is sampled from a multivariate normal distribution with correlation 0.8 to the corresponding column of $\mathbf{X}$. In the aggregated relational data setting, $\mathbf{U}$ is estimated by the left singular vectors of $\mathbf{N}$, denoted $\bar{\mathbf{U}}$. $\mathbf{S}$ is then estimated via $\bar{\mathbf{U}}^\top \mathbf{Y}\mathbf{W}^\top \bar{\mathbf{U}} (\bar{\mathbf{U}}^\top \mathbf{W}\mathbf{W}^\top \bar{\mathbf{U}})^{-1}$,

and these estimates are used as plug-in estimates in the usual way. Theory and motivation for these estimators is under development in forthcoming work by Hayes, Chandrasekhar, McCormick and Breza.

5. **Ego-centric data**: In ego-centric sampling, half of the nodes in the network are selected at random, and only edges incident to those nodes are observed. Unobserved edges are imputed using Algorithm 1 of Chan and T. Li (2023), and then $\mathbf{P}$ is estimated via the full matrix recovery technique presented in the same manuscript.

6. **Degree capped**: $\mathbf{N}$ is a censored version of $\mathbf{A}$, where, for each node i, at most twenty edges incident to i are available. When $d_i > 20$, edges indicent to i are removed uniformly at randomly until i only has twenty incident edges. $\mathbf{U}$ and $\mathbf{S}$ are estimated via singular value decomposition of $\mathbf{N}$.

7. **Edges flipped**: $\mathbf{N}$ is a version of $\mathbf{A}$ when 15% of edges in the network have been flipped. To preserve the total number of edges in the network, this is implemented via random edge swapping, where 15% of edges in the network are swapped with a random edge with different edge value. $\mathbf{U}$ and $\mathbf{S}$ are estimated via singular value decomposition of $\mathbf{N}$.

The results of these simulations are in Figure 3, which compares mean squared error for β_{y} and θ_{y} when using the estimates of $\mathbf{X}$ and $\widetilde{\mathbf{G}}$ defined above. We see that the network smoothing approach is able to recover both β_{y} and θ_{y} in where $\mathbf{A}$ is contaminated with Gaussian noise, when $\mathbf{A}$ has missing edges, and when only an aggregated relational data variant of $\mathbf{A}$ is observed. This matches our expectations that consistent estimation of peer effects is possible whenever principal subspace estimation is possible. However, under degree censoring and edge flip noise mechanisms, principal subspace estimation is challenging and the singular value decomposition is a poor estimator, such that β_{y} and θ_{y} cannot be recovered.

4.2 Sparsity

We additionally investigate how $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\beta}}$ perform in sparser, correctly observed networks. In Figure 4 we report mean squared error when the average degree is $n^{1/2}$, and in Figure 5 we report mean squared error when the average degree is $n^{1/4}$. In both cases, we observed worse performance relative to the denser $n^{3/4}$ average degree case. When the average degree is $n^{1/2}$, mean squared error is more volatile, but Figure 4 does still suggest that $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\beta}}$ are consistent under peer contagion. Under latent contagion, $\hat{\boldsymbol{\theta}}$ appears consistent, but $\hat{\boldsymbol{\beta}}$ struggle to recover the intercept (coefficient Z_1) and the peer effect. We suspect that this is primarily due to finite sample collinearity issues, as degree heterogeneity is necessary to differentiate the intercept from the contagion term, and degree heterogeneity is less pronounced in sparser graphs. In the sparsest setting, with average degree $n^{1/4}$, the impact of sparsity is more severe. Figure 5 suggests that all estimators experience slower rates of convergence and may not even be consistent, under both models, although $\hat{\boldsymbol{\beta}}$ is potentially consistent under the peer contagion model. We believe that this degradation in performance is primarily attributable to noise in the estimates $\widehat{\mathbf{X}}$ around $\mathbf{X}$.

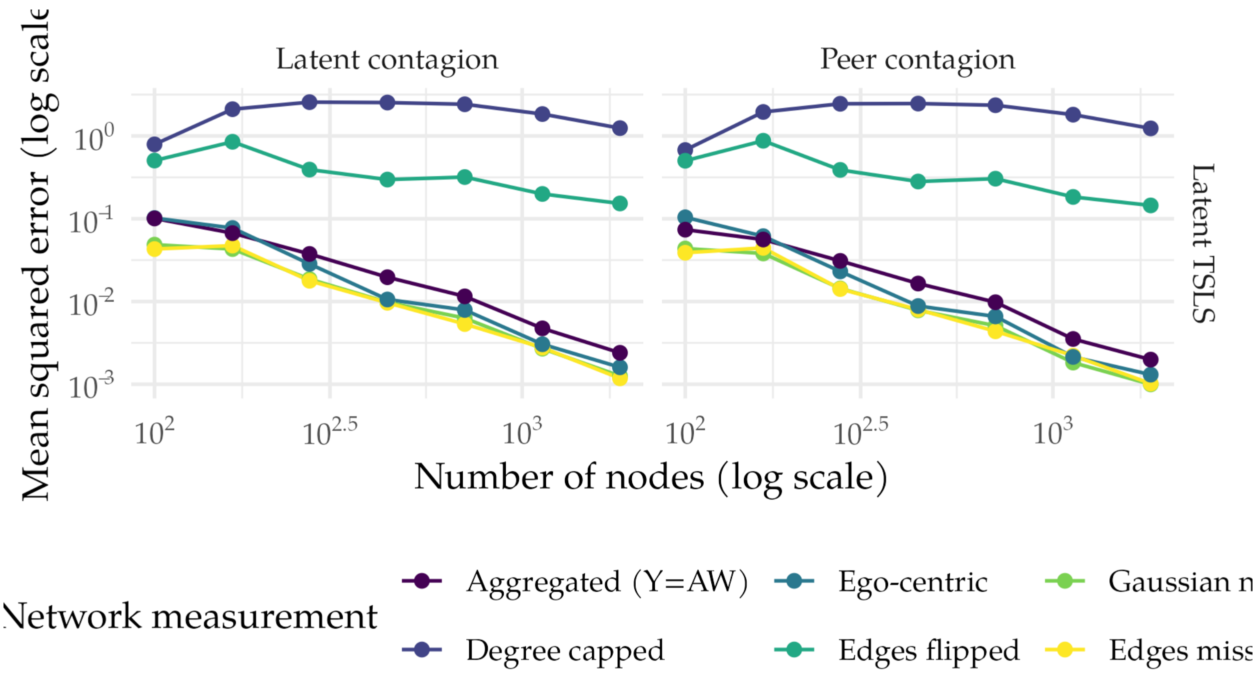

Figure 3: Monte Carlo estimates of mean squared error of $\hat{\boldsymbol{\theta}}$ under various noise models for the network. The average degree in these simulations is $n^{3/4}$.

These simulations show that the sparsity can degrade the performance of the estimators that we have proposed, via two mechanisms: sparsity can reduce degree heterogeneity, leading to collinearity issues, and it may degrade the accuracy of the plug-in estimate $\widehat{\mathbf{X}}$ of $\mathbf{X}$. These simulations suggest that applied analyses using our estimators should assess the quality of estimates $\widehat{\mathbf{X}}$ (namely, assess whether the estimates $\widehat{\mathbf{X}}$ correspond to meaningful structure in the network) and should consider the possibility of variance inflation in regression estimates due to collinearity. In sparse networks with strong block diagonal structure and limited degree heterogenity, the intercept and $\mathbf{X}$ may be collinear, and it may be reasonable to drop a column for collinearity reasons.

5 Data Application

To demonstrate our methods, we applied our estimators to network data collected during the *Teenage Friends and Lifestyle Study*, reported in L. Michell and West (1996), Lynn Michell and Amos (1997), Lynn Michell (1997), and M. P. Michell L. (2000). Recently, Hayes, Fredrickson, and Levin (2025) and Di Maria, Abbruzzo, and Lovison (2022) investigated network-mediation using this data, studying how sex influenced network position, and how network position subsequently influenced smoking behaviors in adolescents. These analyses assumed that there were no peer effects on smoking after accounting for latent network position. We re-analyzed the same data to investigate if smoking exhibits spillovers in addition to being localized within the adolescent social network.

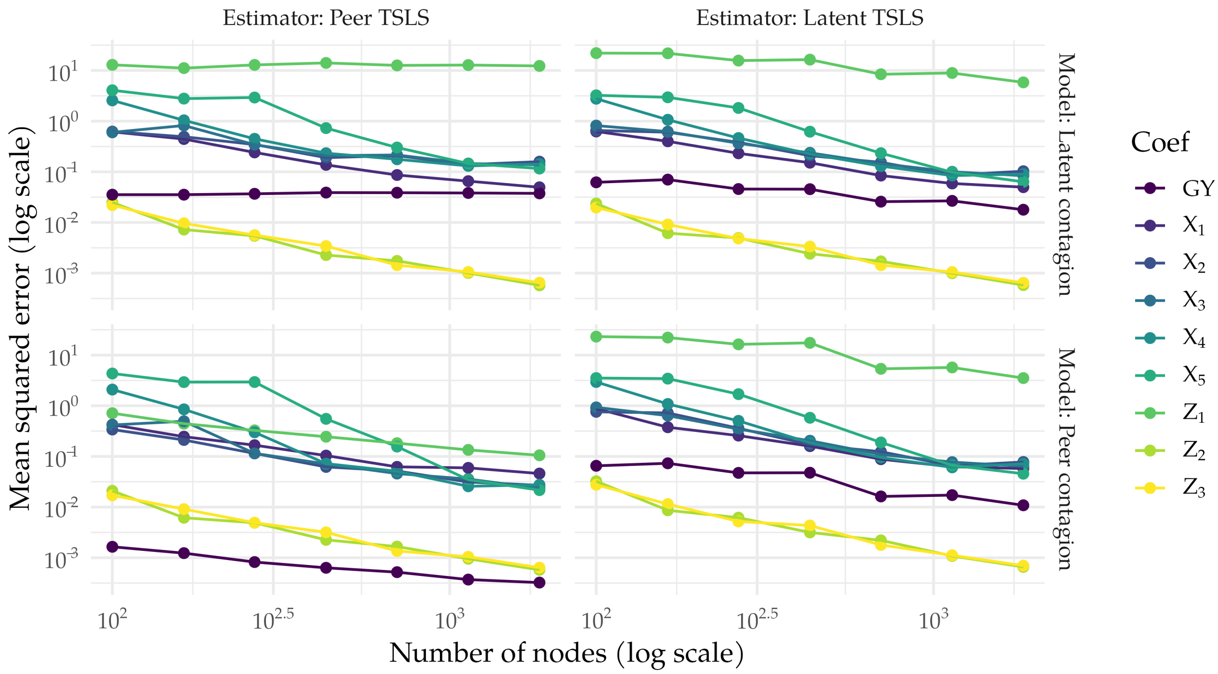

Figure 4: Monte Carlo estimates of mean squared error of $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\beta}}$. The average degree in these simulations is $n^{1/2}$.

5.1 Data

The *Teenage Friends and Lifestyle Study* collected three waves of survey data in a secondary school in Glasgow, beginning in January 1995. Students in the study filled out a questionnaire about their lifestyle and risk-taking behaviors, including alcohol, tobacco and drug use, and additionally were asked to list six of their friends. M. P. Michell L. (2000) found that smoking was mostly concentrated in friend groups composed of popular girls, unpopular students, and trouble-makers: "risk taking behaviour was heavily polarized within social categories so that, for instance, groups of individuals (and their peripherals) were in general either risk-taking or non-risk-taking".

The social network was collected by asking students "who are your best friends", and allowing adolescents to list up to six responses. We considered only data from the first wave of the survey, which included 153 adolescents. Sex and tobacco use were self-reported as nominal features with levels "Male" and "Female"; and "Never", "Occasional," and "Regular," respectively. To match the analyses of Hayes, Fredrickson, and Levin (2025) and Di Maria, Abbruzzo, and Lovison (2022), for the tobacco use measure we combined "Occasional" and "Regular" into a single level, and compared smokers with non-smokers. Also to match the analysis of Hayes, Fredrickson, and Levin (2025), we treated age (continuous) and church attendance (nominal) as possible confounders and thus included these variables as controls.

We computed the adjacency spectral embedding of the social network A. In the Glasgow data, the social network is directed: an edge $i \to j$ indicates that student i listed student j as friend. This directedness means that students have two distinct co-embeddings

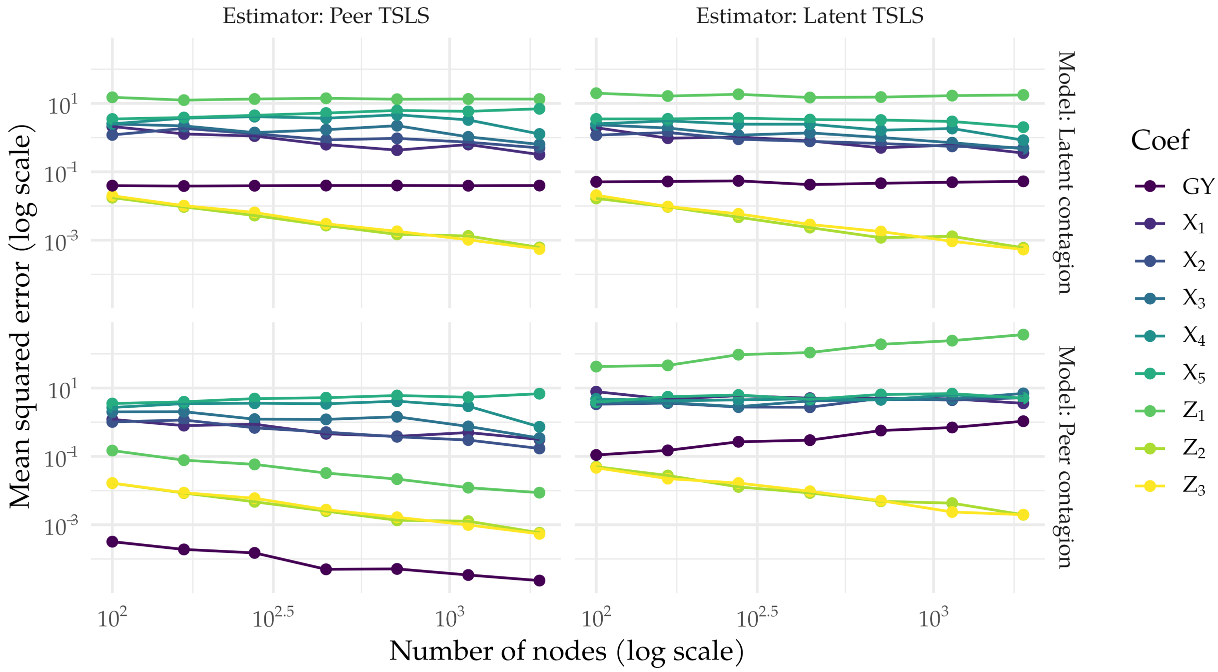

Figure 5: Monte Carlo estimates of mean squared error of $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\beta}}$. The average degree in these simulations is $n^{1/4}$.

corresponding to their propensity to send out-edges and receive in-edges. Letting $\widehat{\mathbf{A}} \approx \widehat{\mathbf{U}}\hat{\mathbf{S}}\widehat{\mathbf{V}}^T$ be the truncated singular value decomposition of $\mathbf{A}$, the left co-embedding $\hat{\mathbf{L}} = \widehat{\mathbf{U}}\hat{\mathbf{S}}^{1/2}$ described how students in the network send edges (i.e., claim friends), and the right co-embedding $\widehat{\mathbf{X}} \equiv \widehat{\mathbf{V}}\hat{\mathbf{S}}^{1/2}$ described how students receive edges (i.e., are claimed as friends). Our results used the right co-embeddings $\widehat{\mathbf{X}}$. We did not select any particular dimension d for the latent space. Instead, we repeated our analysis for many values of d, to investigate the sensitivity of our results to the dimension of the latent space. Once we obtained embeddings $\widehat{\mathbf{X}}$ via the singular value decomposition, we performed a multiverse analysis, estimating regression coefficients using $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\beta}}$. For each estimator, we considered two variants, one including $\widehat{\mathbf{X}}$ as covariates, to adjust estimates for latent positions in the network, and one without including $\widehat{\mathbf{X}}$ as covariates. That is, we used two-stage least squares estimator derived under all four of the following generative models:

$$\mathbf{Y} = \mathbf{1}_n\beta_0 + \mathbf{W}\boldsymbol{\beta}_{\mathrm{w}} + \mathbf{G}\mathbf{Y}\beta_{\mathrm{y}} + \boldsymbol{\varepsilon} \tag{15}$$

$$\mathbf{Y} = \mathbf{1}_n\theta_0 + \mathbf{W}\boldsymbol{\theta}_{\mathrm{w}} + \widetilde{\mathbf{G}}\mathbf{Y}\theta_{\mathrm{y}} + \boldsymbol{\varepsilon} \tag{16}$$

$$\mathbf{Y} = \mathbf{1}_n\beta_0 + \mathbf{W}\boldsymbol{\beta}_{\mathrm{w}} + \mathbf{X}\boldsymbol{\beta}_{\mathrm{x}} + \mathbf{G}\mathbf{Y}\beta_{\mathrm{y}} + \boldsymbol{\varepsilon} \tag{17}$$

$$\mathbf{Y} = \mathbf{1}_n\theta_0 + \mathbf{W}\boldsymbol{\theta}_{\mathrm{w}} + \mathbf{X}\boldsymbol{\theta}_{\mathrm{x}} + \widetilde{\mathbf{G}}\mathbf{Y}\theta_{\mathrm{y}} + \boldsymbol{\varepsilon} \tag{18}$$

Recall that $\mathbf{W}$ consists of age and church attendance. The estimators $\hat{\boldsymbol{\theta}}$ required an estimate of the dimension d of the latent positions $\mathbf{X}$, as does $\hat{\boldsymbol{\beta}}$ when $\mathbf{X}$ are included as covariates in the model. We repeated our analysis for $2 \leq d \leq 25$ in order to understand how the dimension of the embedding affected estimates. For each estimate, we reported a 95%

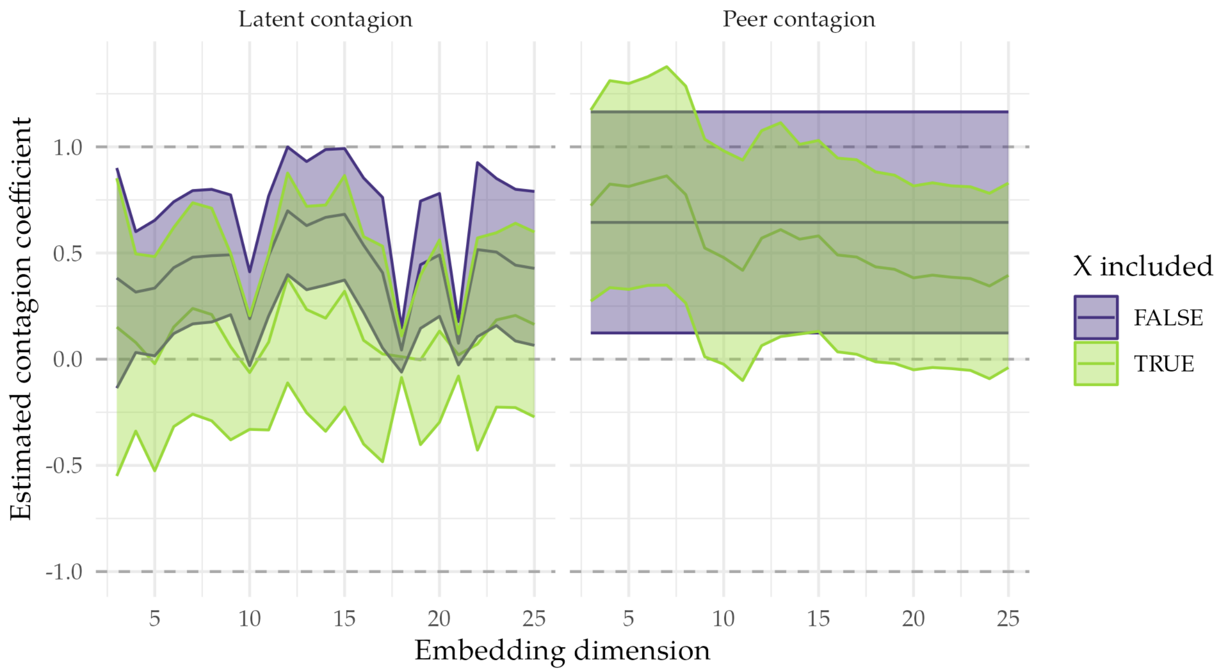

Figure 6: Point estimates and asymptotic 95% confidence intervals for β_{y} and θ_{y} in the Glasgow adolescent social network. β_{y} and θ_{y} measure the contagiousness of smoking.

asymptotic confidence interval in Fig. 6.

5.2 Results

Estimates of the smoking spillovers varied moderately across the multiverse analysis. The simplest and most consistent story emerged amongst estimates that do not adjust for latent homophily by including $\mathbf{X}$ in the regression specification. These estimates, visualized in blue, were consistently large and statistically differentiated from zero, suggesting that smoking does exhibit spillover effects. However, estimates that adjusted for latent homophily by including $\mathbf{X}$ told a different story. In most cases, including $\mathbf{X}$ in the regression reduced the point estimate of the contagion coefficient, and the associated confidence interval often contained zero. However, these confidence intervals remained wide, indicating substantially uncertainty about the presence or absence of a spillover effect after accounting for localized smoking behavior in the network via the $\widehat{\mathbf{X}}$ terms.

Another set of comparisons, between the latent contagion estimators and the peer contagion estimators, was also informative. In particular, Theorems 1, 2, 3, and 4 prove that latent and peer contagion estimators are asymptotically equivalent under the random dot product graph, provided that the network is precisely observed. However, we observed some deviation between the latent contagion and peer contagion estimators, with latent contagion estimates indicating lower levels outcome spillover across most embedding dimensions. There numerous possible explanations for the difference between the latent and peer contagion estimates: (1) the peer network might have been observed with noise, (2) a random dot product model may not haven been appropriate for network (in which case we would prefer the estimates based on the peer contagion model), or (3) the network may have been too small (recall $n = 153$ nodes) for asymptotics results to applicable.

Lastly, we observed that estimates under peer contagion model were fairly stable as a function of the embedding dimension d, but estimates under the latent contagion model were somewhat more volatile in the embedding dimension, with estimates of the contagion coefficient collapsing towards for the specific values $d = 10, 18, 21$. We suspect this volatility as a function of embedding dimension was related to the small sample size the limited number of smokers in the network.

Altogether, our analysis indicated that there was substantial uncertainty about the presence and scale of the contagiousness of smoking in the adolescent social network, after accounting for the localized natural of smoking in the network by control from $\widehat{\mathbf{X}}$. This uncertainty mirrored previous results suggesting that peer effects and homophily are challenging to distinguish in social networks (Hayes and Levin, 2025; Shalizi and Thomas, 2011). A possibly next step to differentiate between these effects in the adolescent social network would be to consider longitudinal models of smoking behavior, such as those proposed in (Chang and Paul, 2025).

6 Discussion

We have shown that under low-rank network models, contagion over a true network is asymptotically equivalent to contagion over a smoothed, latent adjacency matrix. This equivalence enables consistent peer effects estimation even when the observed network contains measurement error, provided we can reliably estimate the network's eigenspace. Any method that reliably estimates the eigenspace of $\mathbb{E}[\mathbf{A}]$—whether through spectral embedding for sub-gamma noise, matrix completion for missing edges, or debiasing techniques for systematic measurement error—can be combined with our latent contagion framework to estimate peer effects in challenging data settings.

While this paper focuses on parameter estimation rather than causal identification, the equivalence we establish has implications for causal inference under interference. Under appropriate conditional ignorability assumptions, network autoregressive parameters can have causal interpretations (Vazquez-Bare, 2023; Leung, 2022; McFowland and Shalizi, 2021). Our results suggest that when peer influence operates through latent structure, causal effects may be more accurately estimated by defining exposures based on latent proximity rather than observed edges.

Acknowledgements

We thank Hyunseung Kang, Ralph Trane, Karl Rohe, Edward McFowland, Arun Chandrasekhar, Michael Newton, Tyler McCormick and the attendees of the University of Wisconsin-Madison IFDS Ideas Seminar for their helpful comments and suggestions. Support for this research was provided by the University of Wisconsin-Madison, Office of the Vice Chancellor for Research and Graduate Education with funding from the Wisconsin Alumni Research Foundation, as well as NSF grants DMS 2052918 and DMS 2023239. Support for this research was also provided by American Family Insurance through a research partnership with the University of Wisconsin-Madison's American Family Insurance Data Science Institute.

Here we collect proofs of our main theoretical results, stated in Theorems 1, 2, 3 and 4, as well as our projection equivalence result in Theorem 5. In the main text, these are listed as holding under "suitable regularity conditions". We begin by listing the assumptions on the covariates and model parameters that constitute these conditions. The first of our assumptions ensures that the spectrum of $\mathbf{P} = \mathbf{X}\mathbf{X}^\top$ is suitably well-behaved.

Assumption 1. The expected adjacency matrix $\mathbf{P}$ satisfies $d = \operatorname{rank} \mathbf{P} = O_p(1)$ and

$$\mathbf{s}_d = \Omega_p(1). \tag{19}$$

Further, the edge-level variance does not grow too quickly compared to the signal in $\mathbf{s}_d$:

$$\mathbf{s}_d = \omega_p\Big(\sqrt{\nu + b^2}\sqrt{n}\log n\Big). \tag{20}$$

Letting $\kappa(\mathbf{P}) = \mathbf{s}_1/\mathbf{s}_d$ be the condition number of $\mathbf{P}$, we require

$$\frac{\kappa(\mathbf{P})(\nu + b^2) n \log^2 n}{\mathbf{s}_d^{3/2}} = o_p(1). \tag{21}$$

We also require assumptions on the behavior of the minimum degree of the weighted adjacency matrix $\mathbf{P}$. Denoting the degree of vertex i in $\mathbf{P}$ by

$$\tilde{d}_i = \mathbb{E}\left[d_i \mid \mathbf{X}\right] = [\mathbf{P}\mathbf{1}]_i = \sum_{j=1}^{n} \mathbf{X}_i^\top \mathbf{X}_j, \tag{22}$$

we require the following assumption.

Assumption 2. The degrees of the expected adjacency matrix $\mathbf{P}$ are such that

$$\min_{i\in[n]} \tilde{d}_i^2 = \omega\left(\kappa(\mathbf{P})(\nu + b^2) n^{3/2} \log^2 n\right), \tag{23}$$

$$\mathbf{s}_d \min_{i\in[n]} \tilde{d}_i = \omega\left(\kappa^{3/2}(\mathbf{P})(\nu + b^2) n^{3/2} \log^2 n\right) \quad \text{and} \tag{24}$$

$$\frac{\mathbf{s}_1 \sqrt{\nu + b^2}\sqrt{n}\log n}{\min_{i\in[n]} \tilde{d}_i^2} = o_p\left(\frac{1}{n^{1/4}}\right). \tag{25}$$

We also require regularity assumptions on the latent positions and other covariates.

Assumption 3. The latent positions $\mathbf{X} \in \mathbb{R}^{n\times d}$ are such that

$$\|\mathbf{X}\|_{2,\infty} = O_p\Big(\kappa(\mathbf{P})\sqrt{\nu + b^2}\log n\Big). \tag{26}$$

Our results for the latent contagion estimator $\hat{\boldsymbol{\theta}}$ require a slightly stronger assumption. We note that these bounds are focused on the setting where the variance of the edges (as encoded by the subgamma parameters ν, b) are decoupled from the growth rate of $\mathbf{s}_d$. As such, we expect that these assumptions can be relaxed by extending the bounds in Rubin-Delanchy et al., 2022 to the case of subgamma edge distributions, but we leave this to future work.

Assumption 4. The tail behavior of the edge-level noise is such that

$$\frac{\kappa(\mathbf{P})\sqrt{\nu + b^2 n}\log n}{\mathbf{s}_d} = O_p(1). \tag{27}$$

Assumption 5 (Peer oracle assumptions)**.** Under the peer contagion model of Equation (3), the following conditions hold.

1. All the diagonal elements of $\mathbf{G}$ are zero and $\mathbf{G}$ has uniformly bounded row and column sums.
2. $|\beta_{\mathrm{y}}| < 1$, so that $\left(\mathbf{I} - \beta_{\mathrm{y}}\mathbf{G}\right)^{-1}$ is non-singular.
3. The regressor matrices $\mathbf{W}$ and $\mathbf{X}$ have full column rank (for n large enough), and the elements of $\mathbf{W}$ and $\mathbf{X}$ are uniformly bounded in absolute value almost surely.
4. $\boldsymbol{\varepsilon}$ are i.i.d. with zero mean, variance $\mathbb{E}\left[\varepsilon_i^2\right] = \sigma_\epsilon^2 < \infty$ and all entries of $\boldsymbol{\varepsilon}$ have finite fourth moments.
5. The instrument matrices $\mathbf{H}$ have full column rank, are composed of a subset of linearly independent columns of

   $$\left[\mathbf{W}\;\mathbf{X}\;\mathbf{GW}\;\mathbf{GX}\;\mathbf{G}^2\mathbf{W}\;\mathbf{G}^2\mathbf{X}\right],$$

   where the subset contains $\mathbf{W}$ and $\mathbf{X}$.
6. The limit

   $$\lim_{n\to\infty}\frac{1}{n}\mathbf{H}^\top\mathbf{H}$$

   is finite and non-singular. Further,

   $$\frac{1}{n}\mathbf{H}^\top\mathbf{Z}$$

   converges in probability to a matrix that is finite with full column rank.

Our results for the latent contagion model require similar assumptions.

Assumption 6 (Latent oracle assumptions)**.** Under the model in Equation (5), the conditions of Assumption 5 hold, with $\widetilde{\mathbf{G}}$ in place of $\mathbf{G}$ and $\widetilde{\mathbf{H}}$ in place of $\mathbf{H}$.

Under either Assumption 5 or 6, the matrix of node-level covariates, the design matrix and the instrument matrix are all well-conditioned. Further, it is straightforward under either of these assumptions, the following growth rates hold. These will prove convenient to have for reference in our proofs to follow.

$$\max\{\|\mathbf{X}\|, \|\mathbf{W}\|\} = O_p\left(\sqrt{n}\right), \tag{28}$$

$$\sigma_{\min}(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}) = \Omega(\sqrt{n}) \;\text{ and }\; \|\widetilde{\mathbf{Z}}\| = O_p\left(\sqrt{n}\right), \;\text{ and} \tag{29}$$

$$\sigma_{\min}(\widetilde{\mathbf{H}}) = \Omega(\sqrt{n}) \text{ and } \|\widetilde{\mathbf{H}}\| = O_p(\sqrt{n}). \tag{30}$$

Our projection equivalence result, Theorem 5, does not require the subgamma edge behavior of Definition 1. Instead, we need only that the edge noise has bounded second moments, along with degree growth conditions that hold in expectation, rather than in probability.

Assumption 7. The matrix $\lim_{n\to\infty} \mathbb{E}_\theta \widetilde{\mathbf{Z}}^\top \widetilde{\mathbf{Z}}/n$ exists and is invertible.

Assumption 8. The entries of $\mathbf{A} - \mathbf{P}$ are, conditionally on $\mathbf{X}$, mean zero and independent (up to symmetry), and obey

$$\max_{i,j} \mathbb{E}_\theta \left[(\mathbf{A} - \mathbf{P})_{ij}^2 \mid \mathbf{X} \right] \leq \nu_n. \tag{31}$$

Assumption 9. The expected degrees $\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_n$ are such that

$$\nu_n \sum_{i=1}^n \mathbb{E}_\theta \frac{1}{\tilde{d}_i^2} = o(n^{-1/2}), \tag{32}$$

where ν_n is the variance parameter is Equation (31) above.

A Technical Results

Here we collect basic results, largely related to concentration inequalities, which we will use to establish our technical results in the sequel.

Definition 6. Let Z be a mean-zero random variable with cumulant generating function $\psi_Z(t) = \log \mathbb{E}\left[e^{tZ}\right]$.

1. Z is sub-Gaussian(ν) for $\nu > 0$ if $\psi_Z(t) \leq t^2\nu/2$ for all $t \in \mathbb{R}$.
2. Z is sub-gamma(ν, b) for $\nu, b \geq 0$ if $\psi_Z(t) \leq \frac{t^2\nu}{2(1-bt)}$ and $\psi_{-Z}(t) \leq \frac{t^2\nu}{2(1-bt)}$ for all $t < 1/b$.

Lemma 4 (Boucheron, Lugosi, and Massart, 2013 Chapter 2)**.** *Suppose that Z is a (ν, b)-subgamma random variable. Then for all $t > 0$,*

$$\Pr\left[|X| > \sqrt{2\nu t} + bt \right] \leq \exp\{-t\}$$

The following is a basic result concerning subgamma random variables, which we prove for the sake of completeness.

Lemma 5. *Let $Z_1, Z_2, \dots, Z_n$ be a collection of independent (ν, b)-subgamma random variables and let $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R}$ be nonnegative. Then, defining $S_n = \sum_i \alpha_i Z_i$, for any $t > 0$, for any constant $c > 0$, for suitably-chosen constant C_0, it holds with probability at least $1 - C_0 n^{-c}$ that*

$$|S_n| \leq C\sqrt{\nu + b^2} \left(\sum_{i=1}^n \alpha_i^2 \right)^{1/2} \log n \tag{33}$$

and

$$|S_n| = O_p\left(\sqrt{\nu + b^2}\sqrt{\sum_{i=1}^{n} \alpha_i^2}\right). \tag{34}$$

Proof. By a basic property of subgamma random variables (see Boucheron, Lugosi, and Massart, 2013, Chapter 2), $\alpha_i Z_i$ is (ν_i, b_i)-subgamma, where $\nu_i = \alpha_i^2 \nu$ and $b_i = \alpha_i b_i$, and the random sum $S_n = \sum_i \alpha_i Z_i$ is a subgamma random variable with parameters

$$\bar{\nu} = \sum_{i=1}^{n} \nu_i = \nu \sum_{i=1}^{n} \alpha_i^2$$

$$\bar{b} = \max_{i \in [n]} b_i = b \max_{i \in [n]} \alpha_i \le b\sqrt{\sum_{i=1}^{n} \alpha_i^2}.$$

Thus, applying Lemma 4, for any $t > 0$,

$$\Pr\left[|S_n| > \sqrt{2\bar{\nu}t} + \bar{b}t\right] \le \exp\{-t\}.$$

Setting t to be any constant yields Equation (34). Setting $t = C \log n$ for suitably large $C > 0$ and noting that $\bar{b} \le C\bar{\nu}^{1/2}$ for suitably-chosen constant $C > 0$, it follows that

$$\Pr\left[|S_n| > C(\bar{\nu}^{1/2} \log^{1/2} n + \bar{b} \log n\right] \le 2n^{-c}.$$

Observing that

$$\bar{\nu}^{1/2} \log^{1/2} n + \bar{b} \log n \le \sqrt{\nu + b^2}\left(\sum_{i=1}^{n} \alpha_i^2\right)^{1/2} \log n$$

establishes Equation (33). □

Lemma 6. *Let $Z_1, Z_2, \dots, Z_n$ be a collection of independent (ν, b)-subgamma random variables and let $c > 0$ be a constant. Then it holds with probability at least $1 - Cn^{-c}$ that*

$$\max_{i \in [n]} |Z_i| \le C\sqrt{\nu + b^2} \log n,$$

Proof. For $t \ge 0$, applying a union bound followed by Lemma 4,

$$\Pr[\max_i |Z_i| > \sqrt{2\nu t} + bt] \le \sum_{i=1}^{n} \Pr[|Z_i| > \sqrt{2\nu t} + bt] \le n \exp\{-t\}.$$

Taking $t = C \log n$ for $C > 0$ chosen suitably large, it holds that with probability at least $1 - n^{-c}$,

$$\max_i |Z_i| \le \sqrt{2C\nu \log n} + Cb \log n \le C\sqrt{\nu + b^2} \log n,$$

as we set out to show. □

A similar result to the one below appeared in Hayes, Fredrickson, and Levin, 2025. We restate it here with slightly adapted notation for the sake of completeness.

Lemma 7. *Suppose that $\boldsymbol{\varepsilon} \in \mathbb{R}^n$ is a vector of independent mean-zero random variables with bounded second moments,*

$$\max_{i\in[n]} \mathbb{E}\left[\boldsymbol{\varepsilon}_i^2\right] \le B$$

for some $B > 0$ not depending on n. Let $H \in \mathbb{R}^{n\times n}$ be a (possibly random) matrix with $\boldsymbol{\varepsilon}$ independent of H. Then

$$\|H\boldsymbol{\varepsilon}\| = O_p\left(\sqrt{B \operatorname{tr} H^T H}\right).$$

In particular, taking $H = I$, $\|\boldsymbol{\varepsilon}\| = O_p\left(\sqrt{Bn}\right)$.

Proof. We observe that

$$\mathbb{E}\left[\|H\boldsymbol{\varepsilon}\|^2\right] = \mathbb{E}\left[\boldsymbol{\varepsilon}^T H^T H \boldsymbol{\varepsilon}\right] \le B \operatorname{tr} H^T H.$$

Let $\delta > 0$ be a constant. Applying Markov's inequality, for any $t > 0$,

$$\Pr\left[\frac{\|H\boldsymbol{\varepsilon}\|^2}{t} > \delta\right] \le \frac{\mathbb{E}\left[\|H\boldsymbol{\varepsilon}\|^2\right]}{t\delta} \le \frac{B \operatorname{tr} H^T H}{t\delta}.$$

Let r_n be any function of n growing such that $r_n = \omega(B \operatorname{tr} H^T H)$. Then taking $t = r_n$,

$$\lim_{n\to\infty} \Pr\left[\frac{\|H\boldsymbol{\varepsilon}\|^2}{r_n} > \delta\right] = 0.$$

Thus, $\|H\boldsymbol{\varepsilon}\|^2 = o_p(r_n)$ for any $r_n = \omega(B \operatorname{tr} H^T H)$, and it follows that

$$\|H\boldsymbol{\varepsilon}\|^2 = O_p(B \operatorname{tr} H^T H).$$

Taking square roots completes the proof. □

Lemma 8. *With notation as above, for $\beta \in (-1, 1)$ fixed and $\boldsymbol{\varepsilon}$ independent of $\widetilde{\mathbf{G}}$, define*

$$\tilde{\epsilon} = \left(I - \beta\widetilde{\mathbf{G}}\right)^{-1} \boldsymbol{\varepsilon} \in \mathbb{R}^n.$$

Then, if the entries of $\boldsymbol{\varepsilon}$ have bounded second moments as in Lemma 6,

$$\|\tilde{\epsilon}\| = O_p\left(\sqrt{\frac{Bn}{(1-\beta)^2}}\right).$$

Proof. Observe that by Lemma 1,

$$\operatorname{tr}\left(\left(\left(I - \beta\widetilde{\mathbf{G}}\right)^{-1}\right)^{\top} \left(I - \beta\widetilde{\mathbf{G}}\right)^{-1}\right) \le \frac{n}{(1-\beta)^2}.$$

Applying Lemma 7 with $H = \left(I - \beta\widetilde{\mathbf{G}}\right)^{-1}$, it follows that

$$\|\tilde{\epsilon}\| = O_p\left(\sqrt{\frac{Bn}{(1-\beta)^2}}\right),$$

as we set out to show. □

Our final result in this section concerns concentration of the degrees $d_1, d_2, \ldots, d_n$ of the observed network $\mathbf{A}$ about their conditional expectations, defined in Equation (22).

Lemma 9. *Suppose that* $\mathbf{A}$ *follows a sub-gamma model as in Definition 1. Then*

$$\max_{i\in[n]} \left| d_i - \tilde{d}_i \right| = O_p\left(\sqrt{\nu + b^2}\sqrt{n}\log n \right).$$

Proof. Fix $i \in [n]$. We observe that

$$d_i - \tilde{d}_i = \sum_{j\in[n]\setminus\{i\}} (A_{ij} - \rho_n X_i^T X_j)$$

is, conditional on $\mathbf{X}$, a sum of (ν, b)-subgamma random variables. Applying Lemma 5 with suitably chosen constants, it holds with probability at least $1 - 2n^{-3}$ that

$$\left| d_i - \tilde{d}_i \right| \le C\sqrt{\nu + b^2}\sqrt{n}\log n. \tag{35}$$

A union bound over $i \in [n]$ completes the proof. □

Lemma 10. *Suppose that* $\mathbf{A}$ *follows a sub-gamma model as in Definition 1 and that the model parameters grow in such a way that Equation (23) holds. Then with probability at least* $1-O(n^{-2})$, *it holds uniformly over all* $i \in [n]$ *that*

$$\left| \frac{1}{d_i} - \frac{1}{\tilde{d}_i} \right| \le \frac{C\sqrt{\nu + b^2}\sqrt{n}\log n}{\tilde{d}_i^2}.$$

Proof. Defining $\gamma_n = C\sqrt{\nu + b^2}\sqrt{n}\log n$, using the fact that $a^{-1} - b^{-1} = b^{-1}(a-b)a^{-1}$ and applying Lemma 9 twice, it holds with high probability that

$$\left| \frac{1}{d_i} - \frac{1}{\tilde{d}_i} \right| \le \frac{C\gamma_n}{\tilde{d}_i(\tilde{d}_i - \gamma_n)} \le \frac{C\sqrt{\nu + b^2}\sqrt{n}\log n}{\tilde{d}_i^2},$$

where the last inequality follows from our growth assumption in Equation (23). □

B Estimating the Latent Positions

Here we collect results relating the latent position estimates $\widehat{\mathbf{X}}$ to the true latent positions $\mathbf{X}$. Many of the results in this section can be found elsewhere in the spectral methods literature (see, for example, Lyzinski et al., 2017; Levin, Athreya, et al., 2019; Levin, Lodhia, and Levina, 2022; Hayes, Fredrickson, and Levin, 2025). We include these results, with notation adjusted to the current setting, for the sake of convenience.

Lemma 11. *Suppose that* $\mathbf{A}$ *follows a sub-gamma model as in Definition 1. Then with high probability,*

$$\|\mathbf{A} - \mathbf{P}\| \le C\sqrt{\nu + b^2}\sqrt{n}\log n.$$

Proof. This result appears as Lemma 5 in Levin, Lodhia, and Levina, 2022, setting $N = 1$ in the notation of that work. □

Lemma 12. *Suppose that* $\mathbf{A}$ *follows a sub-gamma model as in Definition 1 and that Assumption 1 holds. There exists a constant* $C > 0$ *such that with high probability,*

$$\left\|\hat{\mathbf{S}}^{-1/2}\right\| \leq C\mathbf{s}_d^{-1/2} \quad \text{and} \quad \left\|\hat{\mathbf{S}}^{1/2}\right\| \leq C\mathbf{s}_1^{1/2},$$

where $\hat{\mathbf{S}}$ *is as in Definition 2.*

Proof. Both of these facts are shown in the course of proving Lemma 4 of Levin, Lodhia, and Levina, 2022. In particular, see Equations (28) and (32) in that work. □

The following two results are standard in the spectral embeddings literature (see, e.g., Lyzinski et al., 2017) once we include our assumption that d is order a constant. For the first, see Lemma 27 in Hayes, Fredrickson, and Levin, 2025 or Proposisiton 19 in Levin, Lodhia, and Levina, 2022. For the second, see Lemma 38 in Hayes, Fredrickson, and Levin, 2025 or Proposition 20 in Levin, Lodhia, and Levina, 2022.

Lemma 13. *Suppose that* $\mathbf{A}$ *follows a sub-gamma model as in Definition 1 and that Assumption 1 holds. Then there exists a sequence of orthogonal matrices* $\mathbf{Q} \in \mathbb{R}^{d \times d}$ *such that*

$$\left\|\mathbf{U}^\top \widehat{\mathbf{U}} - \mathbf{Q}\right\|_F \leq \frac{C(\nu + b^2) n \log^2 n}{\mathbf{s}_d^2}.$$

Lemma 14. *Under the same assumptions as Lemma 13,*

$$\left\|\widehat{\mathbf{U}} - \mathbf{U}\mathbf{U}^\top \widehat{\mathbf{U}}\right\|_F \leq \frac{C\sqrt{\nu + b^2}\sqrt{n}\log n}{\mathbf{s}_d}.$$

Furthermore, with $\mathbf{Q}$ *the matrix guaranteed by Lemma 13,*

$$\left\|\mathbf{Q}\hat{\mathbf{S}} - \mathbf{S}\mathbf{Q}\right\|_F \leq \frac{C\mathbf{s}_1 \left(\nu + b^2\right) n \log^2 n}{\mathbf{s}_d^2} + C\sqrt{\nu + b^2}\log n,$$

$$\left\|\mathbf{Q}\hat{\mathbf{S}}^{1/2} - \mathbf{S}^{1/2}\mathbf{Q}\right\|_F \leq \frac{C\mathbf{s}_1 \left(\nu + b^2\right) n \log^2 n}{\mathbf{s}_d^{5/2}} + \frac{C\sqrt{\nu + b^2}\log n}{\mathbf{s}_d^{1/2}}$$

and

$$\left\|\mathbf{Q}\hat{\mathbf{S}}^{-1/2} - \mathbf{S}^{-1/2}\mathbf{Q}\right\|_F \leq \frac{C\mathbf{s}_1 \left(\nu + b^2\right) n \log^2 n}{\mathbf{s}_d^{7/2}} + \frac{C\sqrt{\nu + b^2}\log n}{\mathbf{s}_d^{3/2}}.$$

The following result, which generalizes Lemma 40 in Hayes, Fredrickson, and Levin, 2025, is central to proving our main results.

Lemma 15. *Suppose that* $\mathbf{A}$ *follows a sub-gamma model as in Definition 1 and that Assumption 1 holds. Let* $\mathbf{B} \in \mathbb{R}^{n \times r}$ *be a matrix with* $\mathbf{A} - \mathbf{P}$ *independent of* $\mathbf{B}$ *conditional on* $\mathbf{X}$*. Then there exists* $\mathbf{Q} \in \mathbb{R}^{d \times d}$ *such that*

$$\left\| \left(\widehat{\mathbf{X}} \mathbf{Q}^\top - \mathbf{X} \right)^\top \mathbf{B} \right\| \le \frac{C \sqrt{r} \sqrt{\nu + b^2} \|\mathbf{B}\| \log n}{\sqrt{\mathbf{s}_d}} + \frac{C \kappa(\mathbf{P}) (\nu + b^2) \|\mathbf{B}\| n \log^2 n}{\mathbf{s}_d^{3/2}}.$$

Proof. Take $\mathbf{Q} \in \mathbb{R}^{d \times d}$ to be the orthogonal matrix guaranteed by Lemma 13. Applying a standard decomposition for the adjacency spectral embedding (see, for example Lyzinski et al., 2017; Levin, Athreya, et al., 2019; Hayes, Fredrickson, and Levin, 2025), writing $\mathbf{E} = \mathbf{A} - \mathbf{P}$ for ease of notation,

$$\begin{aligned} \left(\widehat{\mathbf{X}} \mathbf{Q}^\top - \mathbf{X} \right)^\top \mathbf{B} &= \mathbf{Q} \left(\widehat{\mathbf{U}} \hat{\mathbf{S}}^{1/2} - \mathbf{U} \mathbf{S}^{1/2} \mathbf{Q} \right)^\top \mathbf{B} \\ &= \mathbf{Q} \mathbf{S}^{-1/2} \mathbf{U}^\top \mathbf{E} \mathbf{B} + \mathbf{Q} \left(\mathbf{Q} \hat{\mathbf{S}}^{-1/2} - \mathbf{S}^{-1/2} \mathbf{Q} \right)^\top \mathbf{U}^\top \mathbf{E} \mathbf{B} \\ &\quad + \mathbf{Q} \hat{\mathbf{S}}^{-1/2} \mathbf{Q}^\top \mathbf{U}^\top \mathbf{E} \mathbf{U} \mathbf{U}^\top \mathbf{B} + \mathbf{Q} \hat{\mathbf{S}}^{1/2} \left(\mathbf{U} \mathbf{U}^\top \widehat{\mathbf{U}} - \mathbf{U} \mathbf{Q} \right)^\top \mathbf{B} \\ &\quad + \mathbf{Q} \left(\mathbf{Q} \hat{\mathbf{S}}^{1/2} - \mathbf{S}^{1/2} \mathbf{Q} \right)^\top \mathbf{U}^\top \mathbf{B} + \mathbf{Q} \hat{\mathbf{S}}^{-1/2} (\widehat{\mathbf{U}} - \mathbf{U} \mathbf{Q})^\top \mathbf{E} \left(\mathbf{I} - \mathbf{U} \mathbf{U}^\top \right) \mathbf{B}. \end{aligned} \tag{36}$$

We will bound each of the six right-hand terms, after which the triangle inequality will yield our result.

For the first right-hand term in Equation (36), using submultiplicativity and the fact that $\mathbf{Q}$ is orthogonal,

$$\|\mathbf{Q} \mathbf{S}^{-1/2} \mathbf{U}^\top \mathbf{E} \mathbf{B}\| \le \frac{\|\mathbf{U}^\top \mathbf{E} \mathbf{B}\|}{\sqrt{\mathbf{s}_d}}.$$

Taking the singular value decomposition of $\mathbf{B}$, then using submultiplicativity and applying Bernstein's inequality (Boucheron, Lugosi, and Massart, 2013; Vershynin, 2020), recalling that $\mathbf{E}$ is independent of $\mathbf{B}$ conditional on $\mathbf{X}$,

$$\|\mathbf{Q} \mathbf{S}^{-1/2} \mathbf{U}^\top \mathbf{E} \mathbf{B}\| \le \frac{C \sqrt{\nu + b^2} \sqrt{r} \|\mathbf{B}\| \log n}{\sqrt{\mathbf{s}_d}}. \tag{37}$$

For the second term in Equation (36), submultiplicativity followed by Bernstein's inequality and Lemma 14 yield

$$\begin{aligned} \left\| \mathbf{Q} \left(\mathbf{Q} \hat{\mathbf{S}}^{-1/2} - \mathbf{S}^{-1/2} \mathbf{Q} \right)^\top \mathbf{U}^\top \mathbf{E} \mathbf{B} \right\| &\le \left\| \mathbf{Q} \hat{\mathbf{S}}^{-1/2} - \mathbf{S}^{-1/2} \mathbf{Q} \right\| \|\mathbf{U}^\top \mathbf{E} \mathbf{B}\| \\ &\le C \left(\frac{\mathbf{s}_1 \sqrt{\nu + b^2}\, n \log n}{\mathbf{s}_d^2} + d \right) \frac{(\nu + b^2) \sqrt{r} \|\mathbf{B}\| \log^2 n}{\mathbf{s}_d^{3/2}}. \end{aligned}$$

Using the growth assumptions in Equations (28), (20), (21) and (19),

$$\left\| \mathbf{Q} \left(\mathbf{Q} \hat{\mathbf{S}}^{-1/2} - \mathbf{S}^{-1/2} \mathbf{Q} \right)^\top \mathbf{U}^\top \mathbf{E} \mathbf{B} \right\| \le \frac{C \sqrt{\nu + b^2} \sqrt{r} \|\mathbf{B}\| \log n}{\sqrt{\mathbf{s}_d}}. \tag{38}$$

Considering the third right-hand term in Equation (36), Bernstein's inequality and Lemma 12 yield

$$\left\| \mathbf{Q}\hat{\mathbf{S}}^{-1/2}\mathbf{Q}^\top \mathbf{U}^\top \mathbf{E}\mathbf{U}\mathbf{U}^\top \mathbf{B} \right\| \leq \frac{C\sqrt{\nu + b^2}\|\mathbf{B}\| \log n}{\sqrt{\mathbf{s}_d}}. \tag{39}$$

For the fourth right-hand term in Equation (36), Lemmas 12 and 13 yield

$$\left\| \mathbf{Q}\hat{\mathbf{S}}^{1/2} \left(\mathbf{U}\mathbf{U}^\top \widehat{\mathbf{U}} - \mathbf{U}\mathbf{Q}\right)^\top \mathbf{B} \right\| \leq \left\| \hat{\mathbf{S}}^{1/2} \right\| \left\| \mathbf{U}^\top \widehat{\mathbf{U}} - \mathbf{Q} \right\| \|\mathbf{B}\| \leq \frac{C(\nu + b^2)\|\mathbf{B}\| n \log^2 n}{\mathbf{s}_d^{3/2}}. \tag{40}$$

Similarly, for the fifth term in Equation (36), submultiplicativity and Lemma 14 yield

$$\begin{aligned} \left\| \mathbf{Q}\left(\mathbf{Q}\hat{\mathbf{S}}^{1/2} - \mathbf{S}^{1/2}\mathbf{Q}\right)^\top \mathbf{U}^\top \mathbf{B} \right\| &\leq \left\| \mathbf{Q}\hat{\mathbf{S}}^{1/2} - \mathbf{S}^{1/2}\mathbf{Q} \right\| \|\mathbf{B}\| \\ &\leq \frac{C\kappa(\mathbf{P})(\nu + b^2)\|\mathbf{B}\| n \log^2 n}{\mathbf{s}_d^{3/2}} + \frac{C\sqrt{\nu + b^2}\|\mathbf{B}\| \log n}{\sqrt{\mathbf{s}_d}}. \end{aligned} \tag{41}$$

Finally, to control the sixth right-hand term in Equation (36), adding and subtracting appropriate quantities yields

$$\begin{aligned} \mathbf{Q}\hat{\mathbf{S}}^{-1/2}(\widehat{\mathbf{U}} - \mathbf{U}\mathbf{Q})^\top \mathbf{E}\left(\mathbf{I} - \mathbf{U}\mathbf{U}^\top\right)\mathbf{B} = \mathbf{Q}\hat{\mathbf{S}}^{-1/2}&\left(\widehat{\mathbf{U}} - \mathbf{U}\mathbf{U}^\top \widehat{\mathbf{U}}\right)^\top \mathbf{E}\left(\mathbf{I} - \mathbf{U}\mathbf{U}^\top\right)\mathbf{B} \\ &+ \mathbf{Q}\hat{\mathbf{S}}^{-1/2}\left(\mathbf{U}\mathbf{U}^\top \widehat{\mathbf{U}} - \mathbf{U}\mathbf{Q}\right)^\top \mathbf{E}\left(\mathbf{I} - \mathbf{U}\mathbf{U}^\top\right)\mathbf{B}. \end{aligned} \tag{42}$$

By submultiplicativity followed by Lemmas 11 and 14,

$$\begin{aligned} \left\| \mathbf{Q}\hat{\mathbf{S}}^{-1/2}\left(\widehat{\mathbf{U}} - \mathbf{U}\mathbf{U}^\top \widehat{\mathbf{U}}\right)^\top \mathbf{E}\left(\mathbf{I} - \mathbf{U}\mathbf{U}^\top\right)\mathbf{B} \right\| &\leq \|\hat{\mathbf{S}}^{-1/2}\| \left\| \widehat{\mathbf{U}} - \mathbf{U}\mathbf{U}^\top \widehat{\mathbf{U}} \right\| \|\mathbf{E}\| \|\mathbf{B}\| \\ &\leq \frac{C\|\mathbf{B}\|(\nu + b^2) n \log^2 n}{\mathbf{s}_d^{3/2}}. \end{aligned} \tag{43}$$

Similarly, submultiplicativity followed by Lemmas 11, 12 and 13,

$$\begin{aligned} \left\| \mathbf{Q}\hat{\mathbf{S}}^{-1/2}\left(\mathbf{U}\mathbf{U}^\top \widehat{\mathbf{U}} - \mathbf{U}\mathbf{Q}\right)^\top \mathbf{E}\left(\mathbf{I} - \mathbf{U}\mathbf{U}^\top\right)\mathbf{B} \right\| &\leq \frac{C\|\mathbf{B}\|(\nu + b^2)^{3/2} n^{3/2} \log^3 n}{\mathbf{s}_d^{5/2}} \\ &\leq \frac{C(\nu + b^2)\|\mathbf{B}\| n \log^2 n}{\mathbf{s}_d^{3/2}}. \end{aligned} \tag{44}$$

where the second inequality follows from our growth assumption in Equation (20). Applying the triangle inequality to Equation (42) and using Equations (43) and (44),

$$\left\| \mathbf{Q}\hat{\mathbf{S}}^{-1/2}(\widehat{\mathbf{U}} - \mathbf{U}\mathbf{Q})^\top \mathbf{E}\left(\mathbf{I} - \mathbf{U}\mathbf{U}^\top\right)\mathbf{B} \right\| \leq \frac{C(\nu + b^2)\|\mathbf{B}\| n \log^2 n}{\mathbf{s}_d^{3/2}}. \tag{45}$$

Applying the triangle inequality to Equation (36), followed by Equations (37), (38), (39), (40), (41), and (45), completes the proof. ☐

Our final result in this section concerns the behavior of

$$\hat{d}_i = \sum_{j=1}^{n} \widehat{\mathbf{X}}_i^\top \widehat{\mathbf{X}}_j, \tag{46}$$

which estimates the "latent" degrees $\tilde{d}_i$ given in Equation (22).

Lemma 16. *Suppose that* $\mathbf{A}$ *follows a sub-gamma model as in Definition 1 and that Assumptions 1 through 4 hold. Then, with high probability, it holds uniformly over all* $i \in [n]$ *that*

$$\left|\hat{d}_i - \tilde{d}_i\right| \le C\sqrt{\kappa(\mathbf{P})}\sqrt{\nu + b^2}\sqrt{n}\log n. \tag{47}$$

Further,

$$\max_{i\in[n]}\left|\frac{1}{\hat{d}_i} - \frac{1}{\tilde{d}_i}\right| \le \frac{C\sqrt{\kappa(\mathbf{P})}\sqrt{\nu + b^2}\sqrt{n}\log n}{\min_i \delta_i^2}. \tag{48}$$

Proof. Fix $i \in [n]$. Recalling the definitions of $\hat{d}_i$ and $\tilde{d}_i$ from Equations (46) and (22), respectively, writing $\mathbf{e}_i \in \mathbb{R}^n$ for the i-th standard basis vector,

$$\begin{aligned}\left|\hat{d}_i - \tilde{d}_i\right| &= \left|\mathbf{e}_i^\top \left(\widehat{\mathbf{P}} - \mathbf{P}\right) 1\right| \\ &\le \left|\mathbf{e}_i^\top \left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)\mathbf{X}^\top 1\right| + \left|\mathbf{e}_i^\top \mathbf{X}\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)^\top 1\right| + \left|\mathbf{e}_i^\top \left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)^\top 1\right|.\end{aligned} \tag{49}$$

By Cauchy-Schwarz and Lemma 15,

$$\begin{aligned}\left|\mathbf{e}_i^\top \left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)\mathbf{X}^\top 1\right| &\le C\sqrt{n}\|\mathbf{X}\|\left(\frac{\sqrt{\nu + b^2}\log n}{\sqrt{\mathbf{s}_d}} + \frac{\kappa(\mathbf{P})(\nu + b^2)n\log^2 n}{\mathbf{s}_d^{3/2}}\right) \\ &\le C\sqrt{n}\left(\sqrt{\kappa(\mathbf{P})}\sqrt{\nu + b^2}\log n + \frac{\kappa^{3/2}(\mathbf{P})(\nu + b^2)n\log^2 n}{\mathbf{s}_d}\right).\end{aligned}$$

Applying our growth bound in Equation (27),

$$\left|\mathbf{e}_i^\top \left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)\mathbf{X}^\top 1\right| \le C\sqrt{\kappa(\mathbf{P})}\sqrt{\nu + b^2}\sqrt{n}\log n.$$

Similarly, trivially upper bounding $\|\mathbf{e}_i^\top \mathbf{X}\| \le \|\mathbf{X}\|$,

$$\left|\mathbf{e}_i^\top \mathbf{X}\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)^\top 1\right| \le C\sqrt{\kappa(\mathbf{P})}\sqrt{\nu + b^2}\sqrt{n}\log n,$$

Applying Lemma 15 twice more, once with $\mathbf{B} = \mathbf{e}_i$ and once with $\mathbf{B} = 1$,

$$\begin{aligned}\left|\mathbf{e}_i^\top \left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)^\top 1\right| &\le C\sqrt{n}\left(\frac{(\nu + b^2)\log^2 n}{\mathbf{s}_d} + \frac{\kappa^2(\mathbf{P})(\nu + b^2)^2 n^2\log^4 n}{\mathbf{s}_d^3}\right) \\ &\le C\sqrt{\kappa(\mathbf{P})}\sqrt{\nu + b^2}\sqrt{n}\log n,\end{aligned}$$

where we have used our growth assumptions in Equations (19), (20), (28), (21) and (27). Applying the above three bounds to Equation (49) yields Equation (47) after noting that the right-hand side does not depend on our choice of $i \in [n]$.

To see Equation (48), note that for any $i \in [n]$,

$$\left|\frac{1}{\hat{d}_i} - \frac{1}{\tilde{d}_i}\right| \le \frac{\left|\hat{d}_i - \tilde{d}_i\right|}{\hat{d}_i \tilde{d}_i}.$$

Using Equation (49) in combination with our assumption in Equation (23), it holds uniformly over all $i \in [n]$ that

$$\left|\frac{1}{\hat{d}_i} - \frac{1}{\tilde{d}_i}\right| \le \frac{C\left|\hat{d}_i - \tilde{d}_i\right|}{\tilde{d}_i^2}.$$

A second application of Equation (49) yields Equation (48) and completes the proof. □

C Controlling the Averaging Operators

Lemma 17. *Suppose that* $\mathbf{A}$ *follows a sub-gamma model as in Definition 1 and suppose that Assumptions 1 through 3 hold and either of Assumptions 5 or 6 hold. Then with high probability,*

$$\left\|\mathbf{G} - \widetilde{\mathbf{G}}\right\| \le C\left(1 + \frac{\mathbf{s}_1}{\min_{i\in[n]} \tilde{d}_i}\right) \frac{\sqrt{\nu + b^2}\sqrt{n}\log n}{\min_{i\in[n]} \tilde{d}_i}.$$

Proof. Recalling the definitions of $\mathbf{G}$ and $\widetilde{\mathbf{G}}$ from Section 3.1, the triangle inequality yields

$$\left\|\mathbf{G} - \widetilde{\mathbf{G}}\right\| \le \left\|\left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1}\right)\mathbf{P}\right\| + \left\|\mathbf{D}^{-1}\left(\mathbf{A} - \mathbf{P}\right)\right\|. \tag{50}$$

By submultiplicativity of the norm and Lemma 9,

$$\left\|\left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1}\right)\mathbf{P}\right\| \le \mathbf{s}_1 \max_{i\in[n]} \frac{\left|d_i - \tilde{d}_i\right|}{d_i \tilde{d}_i} \le \frac{C\mathbf{s}_1\sqrt{\nu + b^2}\sqrt{n}\log n}{\min_{i\in[n]} \tilde{d}_i^2}, \tag{51}$$

where we have used our growth assumption in Equation (23).

Similarly, this time using Lemma 11,

$$\left\|\mathbf{D}^{-1}\left(\mathbf{A} - \mathbf{P}\right)\right\| \le \frac{C\sqrt{\nu + b^2}\sqrt{n}\log n}{\min_{i\in[n]} \tilde{d}_i}. \tag{52}$$

Applying the above two displays to Equation (50),

$$\left\|\mathbf{G} - \widetilde{\mathbf{G}}\right\| \le C\left(1 + \frac{\mathbf{s}_1}{\min_{i\in[n]} \tilde{d}_i}\right) \frac{\sqrt{\nu + b^2}\sqrt{n}\log n}{\min_{i\in[n]} \tilde{d}_i},$$

as we set out to show. □

Lemma 18. *Suppose that* $\mathbf{A}$ *follows a sub-gamma model as in Definition 1 and suppose that Assumptions 1 through 4 hold. Then under either Assumption 5 or 6,*

$$\left\|\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right\| \le C\left(1 + \frac{\mathbf{s}_1}{\min_{i\in[n]} \tilde{d}_i}\right) \frac{\sqrt{\kappa(\mathbf{P})}\sqrt{\nu + b^2}\sqrt{n}\log n}{\min_{i\in[n]} \tilde{d}_i}.$$

Proof. Recalling the definitions of $\widehat{\mathbf{G}}$ and $\widetilde{\mathbf{G}}$ from Equation (9) and Section 3.1, respectively, and applying the triangle inequality,

$$\left\|\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right\| \le \left\|\left(\widehat{\mathbf{D}}^{-1} - \widetilde{\mathbf{D}}^{-1}\right)\mathbf{P}\right\| + \left\|\widehat{\mathbf{D}}^{-1}\left(\hat{\mathbf{P}} - \mathbf{P}\right)\right\|. \tag{53}$$

Using submultiplicativity of the norm and Lemma 16,

$$\left\|\left(\widehat{\mathbf{D}}^{-1} - \widetilde{\mathbf{D}}^{-1}\right)\mathbf{P}\right\| \le \frac{C\mathbf{s}_1\sqrt{\kappa(\mathbf{P})}\sqrt{\nu + b^2}\sqrt{n}\log n}{\min_{i\in[n]} \tilde{d}_i^2}. \tag{54}$$

Since $\hat{\mathbf{P}}$ is a truncation of $\mathbf{A}$, we have $\|\hat{\mathbf{P}} - \mathbf{P}\| \le 2\|\mathbf{A} - \mathbf{P}\|$. Submultiplicativity and Lemma 11 thus yield

$$\left\|\widehat{\mathbf{D}}^{-1}\left(\hat{\mathbf{P}} - \mathbf{P}\right)\right\| \le \left\|\widehat{\mathbf{D}}^{-1}\right\| \left\|\hat{\mathbf{P}} - \mathbf{P}\right\| \le \frac{C\sqrt{\nu + b^2}\sqrt{n}\log n}{\min_{i\in[n]} \hat{d}_i}.$$

By Lemma 16 and our assumption in Equation (23), $\min_i \hat{d} \ge C \min_i \tilde{d}_i$, and thus

$$\left\|\widehat{\mathbf{D}}^{-1}\left(\hat{\mathbf{P}} - \mathbf{P}\right)\right\| \le \frac{C\sqrt{\nu + b^2}\sqrt{n}\log n}{\min_{i\in[n]} \tilde{d}_i}. \tag{55}$$

Applying Equations (54) and (55) to Equation (53),

$$\left\|\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right\| \le C\left(1 + \frac{\mathbf{s}_1}{\min_{i\in[n]} \tilde{d}_i}\right) \frac{\sqrt{\nu + b^2}\sqrt{n}\log n}{\min_{i\in[n]} \tilde{d}_i},$$

completing the proof. □

Lemma 19. *Suppose that* $\mathbf{A}$ *follows a sub-gamma model as in Definition 1 and suppose that Assumptions 1 through 4 hold. Then under either Assumption 5 or 6, with high probability,*

$$\left\|\widehat{\mathbf{G}} - \mathbf{G}\right\| \le C\left(1 + \frac{\mathbf{s}_1}{\min_{i\in[n]} \tilde{d}_i}\right) \frac{\sqrt{\nu + b^2}\sqrt{n}\log n}{\min_{i\in[n]} \tilde{d}_i}.$$

Proof. By the triangle inequality,

$$\left\|\widehat{\mathbf{G}} - \mathbf{G}\right\| \le \left\|\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right\| + \left\|\mathbf{G} - \widetilde{\mathbf{G}}\right\|.$$

Applying Lemmas 18 and 17 completes the proof. □

Lemma 20. *Suppose that* $\mathbf{A}$ *follows a sub-gamma model as in Definition 1 and suppose that Assumptions 1 through 3 hold. Let* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ *be such that* $\mathbf{A} - \mathbf{P}$ *is independent of* $\mathbf{v}, \mathbf{u}$ *conditional on* $\mathbf{X}$*. Then, under either Assumption 5 or 6,*

$$\left|\mathbf{u}^\top \left(\mathbf{G} - \widetilde{\mathbf{G}}\right) \mathbf{v}\right| = o_p\Big(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\Big).$$

Proof. Recalling the definitions of $\mathbf{G}$ and $\widetilde{\mathbf{G}}$ from Section 3.1, the triangle inequality yields

$$\left|\mathbf{u}^\top \left(\mathbf{G} - \widetilde{\mathbf{G}}\right) \mathbf{v}\right| \le \left|\mathbf{u}^\top \left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1}\right) \mathbf{A}\mathbf{v}\right| + \left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \left(\mathbf{A} - \mathbf{P}\right) \mathbf{v}\right|. \tag{56}$$

Using the assumption that $\mathbf{A}-\mathbf{P}$ is independent of $\mathbf{u}$ and $\mathbf{v}$ conditional on $\mathbf{X}$, Bernstein's inequality yields

$$\left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \left(\mathbf{A} - \mathbf{P}\right) \mathbf{v}\right| \le C\sqrt{\sum_{i,j} \frac{(\nu + b^2) u_i^2 v_j^2 \log n}{\tilde{d}_i^2}} \le \frac{C\|\mathbf{u}\|\|\mathbf{v}\|\sqrt{\nu + b^2}\log n}{\min_{i\in[n]} \tilde{d}_i}.$$

Using our growth assumption in Equation (23),

$$\left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \left(\mathbf{A} - \mathbf{P}\right) \mathbf{v}\right| = o_p\Big(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\Big). \tag{57}$$

Applying the triangle inequality,

$$\left|\mathbf{u}^\top \left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1}\right) \mathbf{A}\mathbf{v}\right| \le \left|\mathbf{u}^\top \left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1}\right) \mathbf{P}\mathbf{v}\right| + \left|\mathbf{u}^\top \left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1}\right) \left(\mathbf{A} - \mathbf{P}\right) \mathbf{v}\right|. \tag{58}$$

Applying submultiplicativity of the norm followed by Lemmas 11 and 9 and using our growth assumption in Equation (23),

$$\left|\mathbf{u}^\top \left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1}\right) \left(\mathbf{A} - \mathbf{P}\right) \mathbf{v}\right| \le \|\mathbf{u}\|\|\mathbf{v}\| \left\|\mathbf{A} - \mathbf{P}\right\| \max_{i\in[n]} \frac{|d_i - \tilde{d}_i|}{d_i \tilde{d}_i} \le \frac{C\|\mathbf{u}\|\|\mathbf{v}\|(\nu + b^2) n \log^2 n}{\min_{i\in[n]} \tilde{d}_i^2}.$$

Applying our growth assumption in Equation (23),

$$\left|\mathbf{u}^\top \left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1}\right) \left(\mathbf{A} - \mathbf{P}\right) \mathbf{v}\right| = o_p\Big(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\Big). \tag{59}$$

Recalling $\widetilde{\mathbf{G}} = \widetilde{\mathbf{D}}^{-1}\mathbf{P}$, factoring appropriately and applying the triangle inequality,

$$\left|\mathbf{u}^\top \left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1}\right) \mathbf{P}\mathbf{v}\right| \le \left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \left(\mathbf{D} - \widetilde{\mathbf{D}}\right) \widetilde{\mathbf{G}}\mathbf{v}\right| + \left|\mathbf{u}^\top \left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1}\right) \left(\mathbf{D} - \widetilde{\mathbf{D}}\right) \widetilde{\mathbf{G}}\mathbf{v}\right|. \tag{60}$$

By submultiplicativity of the norm and the Cauchy-Schwarz inequality,

$$\left|\mathbf{u}^\top \left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1}\right) \left(\mathbf{D} - \widetilde{\mathbf{D}}\right) \widetilde{\mathbf{G}}\mathbf{v}\right| \le \|\mathbf{u}\|\|\mathbf{v}\| \left\|\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1}\right\| \left\|\mathbf{D} - \widetilde{\mathbf{D}}\right\|,$$

where we have used the fact that $\|\widetilde{\mathbf{G}}\| = 1$. Applying Lemmas 9 and 10 and using our growth assumption in Equation (23),

$$\left|\mathbf{u}^\top \left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1}\right) \left(\mathbf{D} - \widetilde{\mathbf{D}}\right) \widetilde{\mathbf{G}}\mathbf{v}\right| \le \frac{C\|\mathbf{u}\|\|\mathbf{v}\|(\nu + b^2) n \log^2 n}{\min_{i\in[n]} \tilde{d}_i^2} = o_p\Big(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\Big). \tag{61}$$

Expanding the matrix-vector products and rearranging,

$$\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \left(\mathbf{D} - \widetilde{\mathbf{D}}\right) \widetilde{\mathbf{G}}\mathbf{v} = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\left(d_i - \tilde{d}_i\right) u_i}{\tilde{d}_i^2} \mathbf{X}_i^\top \mathbf{X}_j v_j \tag{62}$$

Define the matrix $\mathbf{R} \in \mathbb{R}^{n\times n}$ according to

$$R_{ij} = \frac{u_i \mathbf{X}_i^\top \mathbf{X}_j}{\tilde{d}_i^2}. \tag{63}$$

Recalling that the degrees are the row sums of the adjacency matrices, we can rewrite Equation (62) as

$$\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \left(\mathbf{D} - \widetilde{\mathbf{D}}\right) \widetilde{\mathbf{G}}\mathbf{v} = \left[(\mathbf{A} - \mathbf{P})\, 1\right]^\top \mathbf{R}\mathbf{v} = 1^\top (\mathbf{A} - \mathbf{P})\, \mathbf{R}\mathbf{v}.$$

Thus, expanding out the product,

$$\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \left(\mathbf{D} - \widetilde{\mathbf{D}}\right) \widetilde{\mathbf{G}}\mathbf{v} = \sum_{i=1}^{n} \sum_{j=1}^{n} (\mathbf{A} - \mathbf{P})_{ij} (\mathbf{R}\mathbf{v})_j = \sum_{i<j} (\mathbf{A} - \mathbf{P})_{ij} \left[(\mathbf{R}\mathbf{v})_j + (\mathbf{R}\mathbf{v})_i\right],$$

where we have used the fact that $\mathbf{A} - \mathbf{P}$ is symmetric with zero diagonal by assumption (though note that the case of non-zero diagonal can be handled straightforwardly). Since the entries of $\mathbf{A} - \mathbf{P}$ are (ν, b)-subgamma random variables, conditionally independent (up to symmetry) given $\mathbf{v}$ and $\mathbf{R}$, standard concentration inequalities imply that with high probability,

$$\left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \left(\mathbf{D} - \widetilde{\mathbf{D}}\right) \widetilde{\mathbf{G}}\mathbf{v}\right| \le C\sqrt{(\nu+b^2) \sum_{i<j} \left[(\mathbf{R}\mathbf{v})_j + (\mathbf{R}\mathbf{v})_i\right]^2} \le C\sqrt{(\nu+b^2) \sum_{i=1}^{n} (\mathbf{R}\mathbf{v})_i^2\, n \log n},$$

where we have used the inequality $(x + y)^2 \le 2(x^2 + y^2)$. It follows that, with high probability,

$$\left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1}\left(\mathbf{D}-\widetilde{\mathbf{D}}\right) \widetilde{\mathbf{G}}\mathbf{v}\right| \le C\sqrt{\nu+b^2}\, \|\mathbf{R}\mathbf{v}\| \sqrt{n} \log n \le C\|\mathbf{v}\| \sqrt{\nu+b^2} \sqrt{\sum_{i,j} \frac{u_i^2 (\mathbf{X}_i^\top \mathbf{X}_j)^2}{\tilde{d}_i^4}} \sqrt{n} \log n.$$

Using the Cauchy-Schwarz inequality and the fact that all summands inside the square root are non-negative,

$$\left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \left(\mathbf{D} - \widetilde{\mathbf{D}}\right) \widetilde{\mathbf{D}}^{-1}\mathbf{P}\mathbf{v}\right| \le \frac{C\|\mathbf{u}\|\|\mathbf{v}\| \sqrt{\nu + b^2} \|\mathbf{X}\|_F \|\mathbf{X}\|_{2,\infty} \sqrt{n} \log n}{\min_{i\in[n]} \tilde{d}_i^2} = o_p\left(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\right), \tag{64}$$

where the equality follows from our growth assumptions in Equations (23), (26) and (28).

Applying Equations (61) and (64) to Equation (60),

$$\left|\mathbf{u}^\top \left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1}\right) \mathbf{P}\mathbf{v}\right| = o_p\left(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\right). \tag{65}$$

Applying Equations (59) and (65) to Equation (58),

$$\left|\mathbf{u}^\top\left(\mathbf{D}^{-1}-\widetilde{\mathbf{D}}^{-1}\right)\mathbf{A}\mathbf{v}\right| = o_p\Big(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\Big),$$

and applying this and Equation (57) to Equation (56) completes the proof. □

Lemma 21. *Suppose that* $\mathbf{A}$ *follows a sub-gamma model as in Definition 1 and suppose that Assumptions 1 through 4 hold. Let* $\mathbf{u},\mathbf{v}\in\mathbb{R}^n$ *be such that* $\mathbf{A}-\mathbf{P}$ *is independent of* $\mathbf{v},\mathbf{u}$ *conditional on* $\mathbf{X}$*. Then under either Assumption 5 or 6,*

$$\left|\mathbf{u}^\top\left(\widehat{\mathbf{G}}-\widetilde{\mathbf{G}}\right)\mathbf{v}\right| = o_p\Big(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\Big).$$

Proof. Recalling $\widehat{\mathbf{G}}$ and $\widetilde{\mathbf{G}}$ as given in Equation (9) and Section 3.1, respectively,

$$\left|\mathbf{u}^\top\left(\widehat{\mathbf{G}}-\widetilde{\mathbf{G}}\right)\mathbf{v}\right| \le \left|\mathbf{u}^\top\left(\widehat{\mathbf{D}}^{-1}-\widetilde{\mathbf{D}}^{-1}\right)\hat{\mathbf{P}}\mathbf{v}\right| + \left|\mathbf{u}^\top\widetilde{\mathbf{D}}^{-1}\left(\hat{\mathbf{P}}-\mathbf{P}\right)\mathbf{v}\right|. \tag{66}$$

By definition of $\hat{\mathbf{P}}$ and $\mathbf{P}$, the triangle inequality yields

$$\begin{aligned}\left|\mathbf{u}^\top\widetilde{\mathbf{D}}^{-1}\left(\hat{\mathbf{P}}-\mathbf{P}\right)\mathbf{v}\right| \le{}& \left|\mathbf{u}^\top\widetilde{\mathbf{D}}^{-1}\left(\mathbf{Q}\widehat{\mathbf{X}}-\mathbf{X}\right)\mathbf{X}^\top\mathbf{v}\right| + \left|\mathbf{u}^\top\widetilde{\mathbf{D}}^{-1}\mathbf{X}\left(\mathbf{Q}\widehat{\mathbf{X}}-\mathbf{X}\right)^\top\mathbf{v}\right| \\ &+ \left|\mathbf{u}^\top\widetilde{\mathbf{D}}^{-1}\left(\mathbf{Q}\widehat{\mathbf{X}}-\mathbf{X}\right)\left(\mathbf{Q}\widehat{\mathbf{X}}-\mathbf{X}\right)^\top\mathbf{v}\right|.\end{aligned} \tag{67}$$

By submultiplicativity and Lemma 15,

$$\left|\mathbf{u}^\top\widetilde{\mathbf{D}}^{-1}\left(\mathbf{Q}\widehat{\mathbf{X}}-\mathbf{X}\right)\mathbf{X}^\top\mathbf{v}\right| \le \frac{C\|\mathbf{u}\|\|\mathbf{v}\|\|\mathbf{X}\|}{\min_i \tilde{d}_i}\left(\frac{\sqrt{\nu+b^2}\log n}{\sqrt{\mathbf{s}_d}} + \frac{\kappa(\mathbf{P})(\nu+b^2)n\log^2 n}{\mathbf{s}_d^{3/2}}\right),$$

and our growth assumptions in Equations (23) and (24) yield

$$\left|\mathbf{u}^\top\widetilde{\mathbf{D}}^{-1}\left(\mathbf{Q}\widehat{\mathbf{X}}-\mathbf{X}\right)\mathbf{X}^\top\mathbf{v}\right| = o_p\Big(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\Big).$$

A near-identical argument yields

$$\left|\mathbf{u}^\top\widetilde{\mathbf{D}}^{-1}\mathbf{X}\left(\widehat{\mathbf{X}}\mathbf{Q}-\mathbf{X}\right)^\top\mathbf{v}\right| = o_p\Big(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\Big).$$

Applying the above two bounds to Equation (67),

$$\left|\mathbf{u}^\top\widetilde{\mathbf{D}}^{-1}\left(\hat{\mathbf{P}}-\mathbf{P}\right)\mathbf{v}\right| \le \left|\mathbf{u}^\top\widetilde{\mathbf{D}}^{-1}\left(\mathbf{Q}\widehat{\mathbf{X}}-\mathbf{X}\right)\left(\widehat{\mathbf{X}}\mathbf{Q}-\mathbf{X}\right)^\top\mathbf{v}\right| + o_p\Big(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\Big). \tag{68}$$

By Cauchy-Schwarz and two applications of Lemma 15,

$$\left|\mathbf{u}^\top\widetilde{\mathbf{D}}^{-1}\left(\widehat{\mathbf{X}}\mathbf{Q}-\mathbf{X}\right)\left(\mathbf{Q}\widehat{\mathbf{X}}-\mathbf{X}\right)^\top\mathbf{v}\right| \le \frac{C\|\mathbf{u}\|\|\mathbf{v}\|}{\min_i \tilde{d}_i}\Big(\frac{(\nu+b^2)\log^2 n}{\mathbf{s}_d} + \frac{\kappa^2(\mathbf{P})(\nu+b^2)^2 n^2\log^4 n}{\mathbf{s}_d^3}\Big). \tag{69}$$

Since the largest eigenvalue of the adjacency matrix is an upper bound on the minimum degree, we have

$$\frac{\kappa^2(\mathbf{P})(\nu+b^2)^2 n^2 \log^4 n}{\mathbf{s}_d^3 \min_i \tilde{d}_i} = \frac{\min_i \tilde{d}_i}{\mathbf{s}_d} \frac{\kappa^2(\mathbf{P})(\nu+b^2)^2 n^2 \log^4 n}{\mathbf{s}_d^2 \min_i \tilde{d}_i^2} \le \frac{\kappa^3(\mathbf{P})(\nu+b^2)^2 n^2 \log^4 n}{\mathbf{s}_d^2 \min_i \tilde{d}_i^2},$$

and our assumption in Equation (24) implies

$$\frac{\kappa^2(\mathbf{P})(\nu+b^2)^2 n^2 \log^4 n}{\mathbf{s}_d^3 \min_i \tilde{d}_i} = o_p\left(\frac{1}{n}\right). \tag{70}$$

Applying this to Equation (69), along with our growth assumption in Equation (24)

$$\left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1}\left(\widehat{\mathbf{X}}\mathbf{Q}-\mathbf{X}\right)\left(\mathbf{Q}\widehat{\mathbf{X}}-\mathbf{X}\right)^\top \mathbf{v}\right| = o_p\left(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\right).$$

Applying this to Equation (68),

$$\left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1}\left(\hat{\mathbf{P}}-\mathbf{P}\right)\mathbf{v}\right| = o_p\left(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\right).$$

Applying this to Equation (66) in turn,

$$\left|\mathbf{u}^\top\left(\widehat{\mathbf{G}}-\widetilde{\mathbf{G}}\right)\mathbf{v}\right| \le \left|\mathbf{u}^\top\left(\widehat{\mathbf{D}}^{-1}-\widetilde{\mathbf{D}}^{-1}\right)\hat{\mathbf{P}}\mathbf{v}\right| + o_p\left(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\right). \tag{71}$$

By the triangle inequality,

$$\left|\mathbf{u}^\top\left(\widehat{\mathbf{D}}^{-1}-\widetilde{\mathbf{D}}^{-1}\right)\hat{\mathbf{P}}\mathbf{v}\right| \le \left|\mathbf{u}^\top\left(\widehat{\mathbf{D}}^{-1}-\widetilde{\mathbf{D}}^{-1}\right)\left(\hat{\mathbf{P}}-\mathbf{P}\right)\mathbf{v}\right| + \left|\mathbf{u}^\top\left(\widehat{\mathbf{D}}^{-1}-\widetilde{\mathbf{D}}^{-1}\right)\mathbf{P}\mathbf{v}\right|. \tag{72}$$

By submultiplicativity,

$$\left|\mathbf{u}^\top\left(\widehat{\mathbf{D}}^{-1}-\widetilde{\mathbf{D}}^{-1}\right)\left(\hat{\mathbf{P}}-\mathbf{P}\right)\mathbf{v}\right| \le \|\mathbf{u}\|\|\mathbf{v}\|\left\|\widehat{\mathbf{D}}^{-1}-\widetilde{\mathbf{D}}^{-1}\right\|\left\|\hat{\mathbf{P}}-\mathbf{P}\right\|.$$

Upper bounding $\|\hat{\mathbf{P}}-\mathbf{P}\| \le 2\|\mathbf{A}-\mathbf{P}\|$, since $\hat{\mathbf{P}}$ is a truncation of $\mathbf{A}$, Lemma 11 yields

$$\left|\mathbf{u}^\top\left(\widehat{\mathbf{D}}^{-1}-\widetilde{\mathbf{D}}^{-1}\right)\left(\hat{\mathbf{P}}-\mathbf{P}\right)\mathbf{v}\right| \le C\|\mathbf{u}\|\|\mathbf{v}\|\left(\max_i \frac{|\hat{d}_i-\tilde{d}_i|}{\tilde{d}_i^2}\right)\sqrt{\nu+b^2}\sqrt{n}\log n.$$

Applying Lemma 16 and using the growth assumption in Equation (23),

$$\left|\mathbf{u}^\top\left(\widehat{\mathbf{D}}^{-1}-\widetilde{\mathbf{D}}^{-1}\right)\left(\hat{\mathbf{P}}-\mathbf{P}\right)\mathbf{v}\right| = o_p\left(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\right).$$

Applying this to Equation (72) and applying the result to Equation (71),

$$\left|\mathbf{u}^\top\left(\widehat{\mathbf{G}}-\widetilde{\mathbf{G}}\right)\mathbf{v}\right| \le \left|\mathbf{u}^\top\left(\widehat{\mathbf{D}}^{-1}-\widetilde{\mathbf{D}}^{-1}\right)\mathbf{P}\mathbf{v}\right| + o_p\left(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\right). \tag{73}$$

Recalling $\widetilde{\mathbf{G}} = \widetilde{\mathbf{D}}^{-1}\mathbf{P}$, factoring appropriately and applying the triangle inequality yields

$$\left|\mathbf{u}^\top \left(\widehat{\mathbf{D}}^{-1} - \widetilde{\mathbf{D}}^{-1}\right) \mathbf{P}\mathbf{v}\right| \leq \left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \left(\widehat{\mathbf{D}} - \widetilde{\mathbf{D}}\right) \widetilde{\mathbf{G}}\mathbf{v}\right| + \left|\mathbf{u}^\top \left(\widehat{\mathbf{D}}^{-1} - \widetilde{\mathbf{D}}^{-1}\right) \left(\widehat{\mathbf{D}} - \widetilde{\mathbf{D}}\right) \widetilde{\mathbf{G}}\mathbf{v}\right|. \tag{74}$$

By submultiplicativity and using our growth assumption in Equation (23),

$$\left|\mathbf{u}^\top \left(\widehat{\mathbf{D}}^{-1} - \widetilde{\mathbf{D}}^{-1}\right) \left(\widehat{\mathbf{D}} - \widetilde{\mathbf{D}}\right) \widetilde{\mathbf{G}}\mathbf{v}\right| \leq \frac{C\|\mathbf{u}\|\|\mathbf{v}\| \left\|\widehat{\mathbf{D}} - \widetilde{\mathbf{D}}\right\|^2}{\min_i \tilde{d}_i^2}.$$

where we have used the fact that $\|\widetilde{\mathbf{G}}\| = 1$. Applying Lemma 16 and using the growth assumption in Equation (23), it follows that

$$\left|\mathbf{u}^\top \left(\widehat{\mathbf{D}}^{-1} - \widetilde{\mathbf{D}}^{-1}\right) \left(\widehat{\mathbf{D}} - \widetilde{\mathbf{D}}\right) \widetilde{\mathbf{G}}\mathbf{v}\right| = o_p\left(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\right). \tag{75}$$

Defining $\mathbf{R} \in \mathbb{R}^{n \times n}$ as in Equation (63), recalling that the degrees $\hat{d}_i$ and $\tilde{d}_i$ are given by row-sums of $\widehat{\mathbf{P}}$ and $\mathbf{P}$, respectively,

$$\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \left(\widehat{\mathbf{D}} - \widetilde{\mathbf{D}}\right) \widetilde{\mathbf{G}}\mathbf{v} = 1^\top \left(\widehat{\mathbf{P}} - \mathbf{P}\right) \mathbf{R}\mathbf{v}.$$

By the triangle inequality,

$$\begin{aligned} \left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \left(\widehat{\mathbf{P}} - \mathbf{P}\right) \mathbf{v}\right| \leq \left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \left(\widehat{\mathbf{X}}\mathbf{Q} - \mathbf{X}\right) \mathbf{X}^\top \mathbf{v}\right| &+ \left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \mathbf{X} \left(\widehat{\mathbf{X}}\mathbf{Q} - \mathbf{X}\right)^\top \mathbf{v}\right| \\ &+ \left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \left(\widehat{\mathbf{X}}\mathbf{Q} - \mathbf{X}\right) \left(\widehat{\mathbf{X}}\mathbf{Q} - \mathbf{X}\right)^\top \mathbf{v}\right|. \end{aligned} \tag{76}$$

Applying submultiplicativity and Lemma 15 with $\mathbf{B} = \widetilde{\mathbf{D}}^{-1}\mathbf{u}$,

$$\left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \left(\widehat{\mathbf{X}}\mathbf{Q} - \mathbf{X}\right) \mathbf{X}^\top \mathbf{v}\right| \leq \frac{C\|\mathbf{u}\|\|\mathbf{v}\|}{\min_i \tilde{d}_i} \left(\sqrt{\kappa}\sqrt{\nu + b^2} \log n + \frac{\kappa^{3/2}(\nu + b^2) n \log^2 n}{\mathbf{s}_d}\right) = o_p\left(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\right),$$

where we have used our growth assumptions in Equations (23) and (24). A near-identical argument yields

$$\left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \mathbf{X} \left(\widehat{\mathbf{X}}\mathbf{Q} - \mathbf{X}\right)^\top \mathbf{v}\right| = o_p\left(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\right).$$

Applying the above two displays to Equation (76),

$$\left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \left(\widehat{\mathbf{P}} - \mathbf{P}\right) \mathbf{v}\right| \leq \left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \left(\widehat{\mathbf{X}}\mathbf{Q} - \mathbf{X}\right) \left(\widehat{\mathbf{X}}\mathbf{Q} - \mathbf{X}\right)^\top \mathbf{v}\right| + o_p\left(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\right). \tag{77}$$

Applying Cauchy-Schwarz and invoking Lemma 15 two more times,

$$\left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \left(\widehat{\mathbf{X}}\mathbf{Q} - \mathbf{X}\right) \left(\widehat{\mathbf{X}}\mathbf{Q} - \mathbf{X}\right)^\top \mathbf{v}\right| \leq \frac{C\|\mathbf{u}\|\|\mathbf{v}\|}{\min_i \tilde{d}_i} \left(\frac{(\nu + b^2) \log^2 n}{\mathbf{s}_d} + \frac{\kappa^2 (\nu + b^2)^2 n^2 \log^4 n}{\mathbf{s}_d^3}\right) = o_p\left(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\right),$$

where we have used our growth assumption in Equation (24) and the bound in Equation (70). Applying this to Equation (77),

$$\left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \left(\widehat{\mathbf{P}} - \mathbf{P}\right) \mathbf{v}\right| \leq \left|\mathbf{u}^\top \widetilde{\mathbf{D}}^{-1} \left(\widehat{\mathbf{X}}\mathbf{Q} - \mathbf{X}\right) \left(\widehat{\mathbf{X}}\mathbf{Q} - \mathbf{X}\right)^\top \mathbf{v}\right| + o_p\left(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\right).$$

Applying this to Equation (73) in turn completes the proof. □

Lemma 22. *Suppose that* $\mathbf{A}$ *follows a sub-gamma model as in Definition 1 and suppose that Assumptions 1 through 4 hold. Let* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ *be such that* $\mathbf{A} - \mathbf{P}$ *is independent of* $\mathbf{v}, \mathbf{u}$ *conditional on* $\mathbf{X}$*. Then under either Assumption 5 or 6,*

$$\left|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \mathbf{G}\right) \mathbf{v}\right| = o_p\left(\frac{\|\mathbf{u}\|\|\mathbf{v}\|}{\sqrt{n}}\right).$$

Proof. By the triangle inequality,

$$\left|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \mathbf{G}\right) \mathbf{v}\right| \leq \left|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right) \mathbf{v}\right| + \left|\mathbf{u}^\top \left(\mathbf{G} - \widetilde{\mathbf{G}}\right) \mathbf{v}\right|.$$

Applying Lemmas 21 and 20 completes the proof. □

D Controlling the Responses

Lemma 23. *Under the latent contagion model in Equation (6), suppose that Assumptions 1 through 3 and Assumption 6 hold. Then*

$$\|\mathbf{Y}\| = O_p(\sqrt{n}).$$

Proof. Recalling Equation (6) and applying the triangle inequality and submultiplicativity,

$$\|\mathbf{Y}\| \leq \left\|\left(\mathbf{I} - \theta_{\mathrm{y}}\widetilde{\mathbf{G}}\right)^{-1}\right\| \|\mathbf{1}_n\theta_0 + \mathbf{W}\boldsymbol{\theta}_{\mathrm{w}} + \mathbf{X}\boldsymbol{\theta}_{\mathrm{x}}\| + \left\|\left(\mathbf{I} - \theta_{\mathrm{y}}\widetilde{\mathbf{G}}\right)^{-1}\boldsymbol{\varepsilon}\right\|.$$

Controlling the first term using Lemma 1 and the triangle inequality, and controlling the second term using Lemma 8,

$$\|\mathbf{Y}\| \leq C\left(\frac{\theta_0\sqrt{n}}{|1 - \theta_{\mathrm{y}}|} + \|\mathbf{W}\boldsymbol{\theta}_{\mathrm{w}}\| + \|\mathbf{X}\|\|\boldsymbol{\theta}_{\mathrm{x}}\|\right)$$

The proof is complete after applying the growth rate in Equation (28) and using the fact that the model parameters are constant in n. □

Lemma 24. *Under the peer contagion model in Equation (4), suppose that Assumptions 1 through 3 and Assumption 5 hold. Then*

$$\mathbf{Y} = \widetilde{\mathbf{Y}} + \zeta_n,$$

where $\mathbf{A} - \mathbf{P}$ *and* $\widetilde{\mathbf{Y}}$ *are independent given* $\mathbf{X}$*,* $\|\widetilde{\mathbf{Y}}\| = O_p(\sqrt{n})$*, and* $\|\zeta_n\| = o_p(\sqrt{n})$*. Further,*

$$\|\mathbf{Y}\| = O_p(\sqrt{n}).$$

Proof. Recalling the definition of $\widetilde{\mathbf{G}}$ from Section 3.1, define

$$\widetilde{\mathbf{Y}} = \left(\mathbf{I} - \beta_{\mathrm{y}}\widetilde{\mathbf{G}}\right)^{-1}(\mathbf{1}_n\beta_0 + \mathbf{W}\boldsymbol{\beta}_{\mathrm{w}} + \mathbf{X}\boldsymbol{\beta}_{\mathrm{x}} + \boldsymbol{\varepsilon})$$

and

$$\zeta_n = \left[\left(\mathbf{I} - \beta_{\mathrm{y}}\mathbf{G}\right)^{-1} - \left(\mathbf{I} - \beta_{\mathrm{y}}\widetilde{\mathbf{G}}\right)^{-1}\right](\mathbf{1}_n\beta_0 + \mathbf{W}\boldsymbol{\beta}_{\mathrm{w}} + \mathbf{X}\boldsymbol{\beta}_{\mathrm{x}} + \boldsymbol{\varepsilon}). \tag{78}$$

Then adding and subtracting appropriate quantities in Equation (4),

$$\mathbf{Y} = \widetilde{\mathbf{Y}} + \zeta_n.$$

An argument parallel to that given in the proof of Lemma 23 above yields

$$\left\|\widetilde{\mathbf{Y}}\right\| = O_p(\sqrt{n}),$$

so our proof will be complete once we control ζ_n.

For ease of notation, define

$$\mathbf{L} = (\mathbf{1}_n \beta_0 + \mathbf{W}\boldsymbol{\beta}_{\mathrm{w}} + \mathbf{X}\boldsymbol{\beta}_{\mathrm{x}} + \boldsymbol{\varepsilon}). \tag{79}$$

Applying the Neumann expansion and the triangle inequality,

$$\|\zeta_n\| \leq \sum_{q=0}^{\infty} |\beta_{\mathrm{y}}|^q \left\|\mathbf{G}^q - \widetilde{\mathbf{G}}^q\right\| \|\mathbf{L}\| = \sum_{q=1}^{\infty} |\beta_{\mathrm{y}}|^q \left\|\mathbf{G}^q - \widetilde{\mathbf{G}}^q\right\| \|\mathbf{L}\|.$$

Expanding $\left\|\mathbf{G}^q - \widetilde{\mathbf{G}}^q\right\|$ as a telescoping sum in q

$$\left\|\mathbf{G}^q - \widetilde{\mathbf{G}}^q\right\| \leq \sum_{k=0}^{q-1} \|\mathbf{G}^{q-1-k}\| \left\|\mathbf{G} - \widetilde{\mathbf{G}}\right\| \left\|\widetilde{\mathbf{G}}^k\right\| \leq q \left\|\mathbf{G} - \widetilde{\mathbf{G}}\right\|. \tag{80}$$

Plugging this into the Neumann expansion and evaluating the series

$$\|\zeta_n\| \leq \|\mathbf{L}\| \left\|\mathbf{G} - \widetilde{\mathbf{G}}\right\| \sum_{q=1}^{\infty} |\beta_{\mathrm{y}}|^q q = \|\mathbf{L}\| \left\|\mathbf{G} - \widetilde{\mathbf{G}}\right\| \frac{|\beta_{\mathrm{y}}|}{(1 - \beta_{\mathrm{y}})^2}$$

Lemma 17 then yields

$$\|\zeta_n\| \leq C \|\mathbf{L}\| \left(1 + \frac{\mathbf{s}_1}{\min_{i \in [n]} \tilde{d}_i}\right) \frac{\sqrt{\nu + b^2}\sqrt{n} \log n}{\min_{i \in [n]} \tilde{d}_i}.$$

Applying our growth assumption in Equation (25) yields

$$\|\zeta_n\| = o_p(\|\mathbf{L}\|). \tag{81}$$

Applying the triangle inequality,

$$\|\mathbf{L}\| \leq |\beta_0| \|\mathbf{1}_n\| + \|\boldsymbol{\beta}_{\mathrm{w}}\| \|\mathbf{W}\| + \|\boldsymbol{\beta}_{\mathrm{x}}\| \|\mathbf{X}\| + \|\boldsymbol{\varepsilon}\|,$$

from which Lemma 7 and our growth rate in Equation (28) implies

$$\|\mathbf{L}\| = O_p(\sqrt{n}). \tag{82}$$

Applying this to Equation (81) completes the proof. □

Lemma 25. *Suppose that Assumptions 1 through 4 hold and that $\mathbf{u} \in \mathbb{R}^n$ is such that $\mathbf{A} - \mathbf{P}$ is independent of $\mathbf{u}$ given $\mathbf{X}$. Then under either of the models in Equations (6) and (4),*

$$\left|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \mathbf{G}\right) \mathbf{Y}\right| = o_p(\|\mathbf{u}\|).$$

Proof. We note that this bound is trivial under latent contagion given our growth assumptions: simply use Lemma 22 and the fact that $\|\mathbf{Y}\| = O_p(\sqrt{n})$. Under peer contagion as in Equation (4), the result requires more careful analysis.

Using Lemma 24, write $\mathbf{Y} = \widetilde{\mathbf{Y}} + \zeta_n$, where $\widetilde{\mathbf{Y}}$ and ζ_n obey the growth rates in Equation (135) and $\widetilde{\mathbf{Y}}$ is independent of $\mathbf{A} - \mathbf{P}$ conditional on $\mathbf{X}$. The triangle inequality then yields

$$\left|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \mathbf{G}\right) \mathbf{Y}\right| \le \left|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \mathbf{G}\right) \zeta_n\right| + \left|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \mathbf{G}\right) \widetilde{\mathbf{Y}}\right| \le \left|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \mathbf{G}\right) \zeta_n\right| + o_p(\|\mathbf{u}\|), \quad (83)$$

where the second bound follows from Lemma 22 and the fact that $\|\mathbf{Y}\| = O_p(\sqrt{n})$ by construction. By the definition of ζ_n in Equation (78),

$$\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \mathbf{G}\right) \zeta_n = \sum_{q=0}^{\infty} \beta_y^q \mathbf{u}^\top \left(\widehat{\mathbf{G}} - \mathbf{G}\right) \left(\mathbf{G}^q - \widetilde{\mathbf{G}}^q\right) \mathbf{L}, \quad (84)$$

where $\mathbf{L}$ is as in Equation (79). By Lemma 22 and Equation (82),

$$\left|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \mathbf{G}\right) \mathbf{L}\right| = o_p(\|\mathbf{u}\|).$$

By submultiplicativity followed by Lemmas 19 and 17,

$$\left|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \mathbf{G}\right) \left(\mathbf{G}^q - \widetilde{\mathbf{G}}^q\right) \mathbf{L}\right| \le C\|\mathbf{u}\|\|\mathbf{L}\| \left(1 + \frac{\mathbf{s}_1}{\min_{i\in[n]} \tilde{d}_i}\right)^2 \frac{(\nu + b^2) n \log^2 n}{\min_{i\in[n]} \tilde{d}_i^2} = o_p(\|\mathbf{u}\|),$$

where the equality follows from Equation (82) and our growth assumption in Equation (25). Applying the triangle inequality in Equation (84) and using the above two bounds,

$$\left|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \mathbf{G}\right) \zeta_n\right| \le \sum_{q=2}^{\infty} \left|\beta_y^q \mathbf{u}^\top \left(\widehat{\mathbf{G}} - \mathbf{G}\right) \left(\mathbf{G}^q - \widetilde{\mathbf{G}}^q\right) \mathbf{L}\right| + o_p(\|\mathbf{u}\|). \quad (85)$$

For any $q \ge 2$, the triangle inequality yields

$$\left|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \mathbf{G}\right) \left(\mathbf{G}^q - \widetilde{\mathbf{G}}^q\right) \mathbf{L}\right| \le \left|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \mathbf{G}\right) \widetilde{\mathbf{G}} \left(\mathbf{G}^{q-1} - \widetilde{\mathbf{G}}^{q-1}\right) \mathbf{L}\right| + \left|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \mathbf{G}\right) \left(\mathbf{G} - \widetilde{\mathbf{G}}\right) \mathbf{G}^{q-1}\mathbf{L}\right|.$$

Recursing, it holds for any $q \ge 2$ that

$$\left|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \mathbf{G}\right) \left(\mathbf{G}^q - \widetilde{\mathbf{G}}^q\right) \mathbf{L}\right| \le q\|\mathbf{u}\|\|\mathbf{L}\| \left\|\widehat{\mathbf{G}} - \mathbf{G}\right\| \left\|\mathbf{G} - \widetilde{\mathbf{G}}\right|.$$

Applying Lemmas 19 and 17 along with the bound in Equation (82),

$$\left|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \mathbf{G}\right) \left(\mathbf{G}^q - \widetilde{\mathbf{G}}^q\right) \mathbf{L}\right| \le Cq\|\mathbf{u}\|\sqrt{n} \left(1 + \frac{\mathbf{s}_1}{\min_{i\in[n]} \tilde{d}_i}\right)^2 \frac{(\nu + b^2) n \log^2 n}{\min_{i\in[n]} \tilde{d}_i^2}.$$

Applying this bound to each of the terms in Equation (85),

$$\left|\mathbf{u}^\top\left(\widehat{\mathbf{G}}-\mathbf{G}\right)\zeta_n\right| \le C\|\mathbf{u}\|\sqrt{n}\left(1+\frac{\mathbf{s}_1}{\min_{i\in[n]}\tilde{d}_i}\right)^2\frac{(\nu+b^2)n\log^2 n}{\min_{i\in[n]}\tilde{d}_i^2}\sum_{q=2}^{\infty}q\left|\beta_y\right|^q + o_p(\|\mathbf{u}\|).$$

Applying our growth assumption in Equation (25),

$$\left|\mathbf{u}^\top\left(\widehat{\mathbf{G}}-\mathbf{G}\right)\zeta_n\right| = o_p(\|\mathbf{u}\|).$$

Applying this to Equation (83) completes the proof. □

E Controlling the Instruments

We now turn to describing the behavior of our instruments (e.g., $\widehat{\mathbf{H}}, \breve{\mathbf{H}}, \widetilde{\mathbf{H}}$ and $\mathbf{H}$). As an oracle counterpart to the instrument $\breve{\mathbf{H}}$ in definition 4, we define

$$\widetilde{\mathbf{H}} = \left[\mathbf{W}\ \mathbf{X}\ \widetilde{\mathbf{G}}\mathbf{W}\ \widetilde{\mathbf{G}}\mathbf{X}\ \widetilde{\mathbf{G}}^2\mathbf{W}\ \widetilde{\mathbf{G}}^2\mathbf{X}\right] \in \mathbb{R}^{n\times(3p+3d)}. \tag{86}$$

Similarly, as an oracle counterpart to $\widehat{\mathbf{H}}$ in Definition 3, we define

$$\mathbf{H} = \left[\mathbf{W}\ \mathbf{X}\ \mathbf{G}\mathbf{W}\ \mathbf{G}\mathbf{X}\ \mathbf{G}^2\mathbf{W}\ \mathbf{G}^2\mathbf{X}\right] \in \mathbb{R}^{n\times(3p+3d)}. \tag{87}$$

The technical results below relate these four different versions of the instruments to one another. To account for the rotational non-identifiability of $\mathbf{X}$, we must consider appropriate orthogonal rotations of $\widehat{\mathbf{H}}$ and $\breve{\mathbf{H}}$, which are given by

$$\mathbf{Q}_{\mathbf{H}} = \begin{bmatrix} \mathbf{I}_{p\times p} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{Q} & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{I}_{p\times p} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{Q} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{I}_{p\times p} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{Q} \end{bmatrix} \in \mathbb{R}^{3(p+d)\times 3(p+d)}, \tag{88}$$

where $\mathbf{Q} \in \mathbb{R}^{d\times d}$ is the orthogonal matrix guaranteed by Lemma 15.

Lemma 26. *Suppose that Assumptions 1 through 4 hold. Then under either the peer contagion model in Equation (4) with Assumption 5 or the latent contagion model in Equation (6) with Assumption 6, with $\mathbf{Q}_{\mathbf{H}} \in \mathbb{R}^{(3p+3d)\times(3p+3d)}$ as given in Equation (88),*

$$\left\|\widehat{\mathbf{H}}\mathbf{Q}_{\mathbf{H}}^\top - \widetilde{\mathbf{H}}\right\| = o_p\left(\sqrt{\sigma_{\min}(\widetilde{\mathbf{H}})}\right).$$

Proof. Recalling the definitions of $\widehat{\mathbf{H}}$ and $\widetilde{\mathbf{H}}$ from Definition 3 and Equation (86), respectively, and applying the triangle inequality,

$$\begin{aligned}\left\|\widehat{\mathbf{H}}\mathbf{Q}_{\mathbf{H}}^\top - \widetilde{\mathbf{H}}\right\| \le \left\|\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right\| + \left\|\left(\widehat{\mathbf{G}}-\widetilde{\mathbf{G}}\right)\mathbf{W}\right\| + \left\|\widehat{\mathbf{G}}\widehat{\mathbf{X}}\mathbf{Q}^\top - \widetilde{\mathbf{G}}\mathbf{X}\right\| \\ + \left\|\left(\widehat{\mathbf{G}}^2-\widetilde{\mathbf{G}}^2\right)\mathbf{W}\right\| + \left\|\widehat{\mathbf{G}}^2\widehat{\mathbf{X}}\mathbf{Q}^\top - \widetilde{\mathbf{G}}^2\mathbf{X}\right\|.\end{aligned} \tag{89}$$

Applying Lemma 15 with $\mathbf{B} = \mathbf{I}$ followed by our growth assumptions in Equations (20), (28) and (21),

$$\left\|\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right\| \le \frac{C\sqrt{\nu + b^2}\sqrt{n}\log n}{\sqrt{\mathbf{s}_d}} + \frac{C\kappa(\mathbf{P})(\nu + b^2)n\log^2 n}{\mathbf{s}_d^{3/2}} = o_p\left(n^{1/4}\right).$$

Applying our assumption in Equation (30),

$$\left\|\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right\| = o_p\left(\sqrt{\sigma_{\min}(\widetilde{\mathbf{H}})}\right). \tag{90}$$

Applying submultiplicativity of the norm and Lemma 18,

$$\left\|\left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\mathbf{W}\right\| \le C\left(1 + \frac{\mathbf{s}_1}{\min_{i\in[n]}\tilde{d}_i}\right)\frac{\sqrt{\nu + b^2}\sqrt{n}\log n}{\min_{i\in[n]}\tilde{d}_i}\left\|\mathbf{W}\right\|.$$

Applying our growth assumptions in Equations (28) and (25),

$$\left\|\left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\mathbf{W}\right\| = o_p\left(n^{1/4}\right).$$

Applying the growth assumption in Equation (30),

$$\left\|\left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\mathbf{W}\right\| = o_p\left(\sqrt{\sigma_{\min}(\widetilde{\mathbf{H}})}\right). \tag{91}$$

Similarly, by the triangle inequality, submultiplicativity of the norm and the fact that $\widehat{\mathbf{G}}$ and $\widetilde{\mathbf{G}}$ both have unit spectral norm,

$$\left\|\left(\widehat{\mathbf{G}}^2 - \widetilde{\mathbf{G}}^2\right)\mathbf{W}\right\| \le \left\|\widehat{\mathbf{G}}\left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\mathbf{W}\right\| + \left\|\left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\widetilde{\mathbf{G}}\mathbf{W}\right\| \le 2\left\|\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right\|\left\|\mathbf{W}\right\|.$$

The same argument as that leading to Equation (91) then yields

$$\left\|\left(\widehat{\mathbf{G}}^2 - \widetilde{\mathbf{G}}^2\right)\mathbf{W}\right\| = o_p\left(\sqrt{\sigma_{\min}(\widetilde{\mathbf{H}})}\right). \tag{92}$$

By the triangle inequality, again using submultiplicativity of the norm and the fact that $\widehat{\mathbf{G}}$ has unit spectral norm,

$$\left\|\widehat{\mathbf{G}}\widehat{\mathbf{X}}\mathbf{Q}^\top - \widetilde{\mathbf{G}}\mathbf{X}\right\| \le \left\|\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right\|\left\|\mathbf{X}\right\| + \left\|\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right\|,$$

whence Lemma 18 and Equation (90) yield

$$\left\|\widehat{\mathbf{G}}\widehat{\mathbf{X}}\mathbf{Q}^\top - \widetilde{\mathbf{G}}\mathbf{X}\right\| \le C\sqrt{\mathbf{s}_d}\left(1 + \frac{\mathbf{s}_1}{\min_{i\in[n]}\tilde{d}_i}\right)\frac{\sqrt{\nu + b^2}\sqrt{n}\log n}{\min_{i\in[n]}\tilde{d}_i} + o_p\left(\sqrt{\sigma_{\min}(\widetilde{\mathbf{H}})}\right).$$

Once more applying the same argument as that leading to Equation (91), this time using the growth assumption in Equation (28),

$$\left\|\widehat{\mathbf{G}}\widehat{\mathbf{X}}\mathbf{Q}^\top - \widetilde{\mathbf{G}}\mathbf{X}\right\| = o_p\left(\sqrt{\sigma_{\min}(\widetilde{\mathbf{H}})}\right). \tag{93}$$

By an argument parallel to that yielding Equation (92),

$$\left\|\left(\widehat{\mathbf{G}}^2-\widetilde{\mathbf{G}}^2\right)\mathbf{X}\right\| = o_p\left(\sqrt{\sigma_{\min}(\widetilde{\mathbf{H}})}\right). \tag{94}$$

Applying Equations (90), (91), (92), (93), and (94) to Equation (89) completes the proof. □

Lemma 27. *Suppose that Assumptions 1 through 3 hold. Then under either the peer contagion model in Equation (4) with Assumption 5 or the latent contagion model in Equation (6) with Assumption 6, with $\mathbf{Q}_{\mathbf{H}} \in \mathbb{R}^{(3p+3d)\times(3p+3d)}$ as given in Equation (88),*

$$\left\|\breve{\mathbf{H}}\mathbf{Q}_{\mathbf{H}}^\top - \widetilde{\mathbf{H}}\right\| = o_p\left(\sqrt{\sigma_{\min}(\widetilde{\mathbf{H}})}\right).$$

Proof. Recalling the definitions of $\breve{\mathbf{H}}$, $\widetilde{\mathbf{H}}$ and $\mathbf{Q}_{\mathbf{H}}$, the triangle inequality yields

$$\begin{aligned}\left\|\breve{\mathbf{H}}\mathbf{Q}_{\mathbf{H}}^\top - \widetilde{\mathbf{H}}\right\| \le \left\|\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right\| + \left\|\left(\mathbf{G}-\widetilde{\mathbf{G}}\right)\mathbf{W}\right\| + \left\|\mathbf{G}\widehat{\mathbf{X}}\mathbf{Q}^\top - \widetilde{\mathbf{G}}\mathbf{X}\right\| \\ + \left\|\left(\mathbf{G}^2-\widetilde{\mathbf{G}}^2\right)\mathbf{W}\right\| + \left\|\mathbf{G}^2\widehat{\mathbf{X}}\mathbf{Q}^\top - \widetilde{\mathbf{G}}^2\mathbf{X}\right\|.\end{aligned}$$

The result then follows from the same argument as in Lemma 26, using Lemma 17 in place of Lemma 18. □

Lemma 28. *Suppose that Assumptions 1 through 3 hold. Then under either the peer contagion model in Equation (4) with Assumption 5 or the latent contagion model in Equation (6) with Assumption 6, with $\breve{\mathbf{H}}$ and $\mathbf{H}$ as given in Definition 3 and Equation (87), respectively, and with $\mathbf{Q}_{\mathbf{H}} \in \mathbb{R}^{3(p+d)\times 3(p+d)}$ as given in Equation (88),*

$$\left\|\breve{\mathbf{H}}\mathbf{Q}_{\mathbf{H}}^\top - \mathbf{H}\right\| = o_p\left(\sqrt{\sigma_{\min}(\widetilde{\mathbf{H}})}\right).$$

Proof. Recalling the definitions of $\breve{\mathbf{H}}$ and $\mathbf{H}$ and applying the triangle inequality,

$$\left\|\breve{\mathbf{H}}\mathbf{Q}_{\mathbf{H}}^\top - \mathbf{H}\right\| \le \left\|\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right\| + \left\|\mathbf{G}\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)\right\| + \left\|\mathbf{G}^2\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)\right\| \le 3\left\|\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right\|,$$

where the second inequality follows from submultiplicativity and the fact that $\|\mathbf{G}\| = 1$. Applying Lemma 15 with $\mathbf{B} = \mathbf{I}$ followed by our assumption in Equations (20), (28) and (21),

$$\left\|\breve{\mathbf{H}}\mathbf{Q}_{\mathbf{H}}^\top - \mathbf{H}\right\| \le \frac{C\sqrt{\nu+b^2}\sqrt{n}\log n}{\sqrt{\mathbf{s}_d}} + \frac{C\kappa(\mathbf{P})(\nu+b^2)n\log^2 n}{\mathbf{s}_d^{3/2}} = o_p(\sqrt{n}).$$

Applying our growth assumption in Equation (30) completes the proof. □

Lemma 29. *Suppose that Assumptions 1 through 3 hold. Then under either the peer contagion model in Equation (4) with Assumption 5 or the latent contagion model in Equation (6) with Assumption 6, with $\mathbf{H}$ and $\widetilde{\mathbf{H}}$ as given in Definitions 3 and 4,*

$$\left\|\mathbf{H} - \widetilde{\mathbf{H}}\right\| = o_p\left(\sqrt{\sigma_{\min}(\widetilde{\mathbf{H}})}\right).$$

Proof. By the triangle inequality, letting $\mathbf{Q}_{\mathbf{H}} \in \mathbb{R}^{(3p+3d)\times(3p+3d)}$ be as given in Equation (88),

$$\left\|\mathbf{H} - \widetilde{\mathbf{H}}\right\| \leq \left\|\breve{\mathbf{H}}\mathbf{Q}_{\mathbf{H}}^{\top} - \mathbf{H}\right\| + \left\|\breve{\mathbf{H}}\mathbf{Q}_{\mathbf{H}}^{\top} - \widetilde{\mathbf{H}}\right\|,$$

and Lemmas 28 and 27 complete the proof. □

Lemma 30. *Let $\mathbf{u} \in \mathbb{R}^n$ be independent of $\mathbf{A} - \mathbf{P}$ given $\mathbf{X}$ and suppose that Assumptions 1 through 4 hold. Then under either the peer contagion model in Equation (4) with Assumption 5 or the latent contagion model in Equation (6) with Assumption 6, with $\mathbf{Q}_{\mathbf{H}} \in \mathbb{R}^{(3p+3d)\times(3p+3d)}$ as given in Equation (88),*

$$\left\|\mathbf{u}^{\top}\left(\widehat{\mathbf{H}}\mathbf{Q}_{\mathbf{H}}^{\top} - \widetilde{\mathbf{H}}\right)\right\| = o_p(\|\mathbf{u}\|).$$

Proof. Recalling the definitions of $\widehat{\mathbf{H}}$ and $\widetilde{\mathbf{H}}$ and applying the triangle inequality,

$$\begin{aligned}\left\|\mathbf{u}^{\top}\left(\widehat{\mathbf{H}}\mathbf{Q}_{\mathbf{H}}^{\top} - \widetilde{\mathbf{H}}\right)\right\| \leq \left\|\mathbf{u}^{\top}\left(\widehat{\mathbf{X}}\mathbf{Q}^{\top} - \mathbf{X}\right)\right\| + \left\|\mathbf{u}^{\top}\left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\mathbf{W}\right\| + \left\|\mathbf{u}^{\top}\left(\widehat{\mathbf{G}}\widehat{\mathbf{X}}\mathbf{Q}^{\top} - \widetilde{\mathbf{G}}\mathbf{X}\right)\right\| \\ + \left\|\mathbf{u}^{\top}\left(\widehat{\mathbf{G}}^2 - \widetilde{\mathbf{G}}^2\right)\mathbf{W}\right\| + \left\|\mathbf{u}^{\top}\left(\widehat{\mathbf{G}}^2\widehat{\mathbf{X}}\mathbf{Q}^{\top} - \widetilde{\mathbf{G}}^2\mathbf{X}\right)\right\|.\end{aligned} \tag{95}$$

Applying Lemma 15 with $\mathbf{B} = \mathbf{u}$, followed by our growth assumptions in Equations (20), (28) and (21),

$$\left\|\mathbf{u}^{\top}\left(\widehat{\mathbf{X}}\mathbf{Q}^{\top} - \mathbf{X}\right)\right\| \leq \frac{C\|\mathbf{u}\|\sqrt{\nu + b^2}\log n}{\sqrt{\mathbf{s}_d}} + \frac{C\kappa(\mathbf{P})\|\mathbf{u}\|(\nu + b^2)n\log^2 n}{\mathbf{s}_d^{3/2}} = o_p(\|\mathbf{u}\|). \tag{96}$$

Applying the SVD to $\mathbf{W}$ and using Lemma 21,

$$\left\|\mathbf{u}^{\top}\left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\mathbf{W}\right\| = o_p\left(\frac{\|\mathbf{W}\|\|\mathbf{u}\|}{\sqrt{n}}\right) = o_p(\|\mathbf{u}\|), \tag{97}$$

where the second equality follows from the growth assumption in Equation (28).

Similarly, by the triangle inequal)ity,

$$\begin{aligned}\left\|\mathbf{u}^{\top}\left(\widehat{\mathbf{G}}^2 - \widetilde{\mathbf{G}}^2\right)\mathbf{W}\right\| &\leq \left\|\mathbf{u}^{\top}\widehat{\mathbf{G}}\left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\mathbf{W}\right\| + \left\|\mathbf{u}^{\top}\left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\widetilde{\mathbf{G}}\mathbf{W}\right\| \\ &\leq \left\|\mathbf{u}^{\top}\widetilde{\mathbf{G}}\left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\mathbf{W}\right\| + \left\|\mathbf{u}^{\top}\left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\mathbf{W}\right\| + \left\|\mathbf{u}^{\top}\left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\widetilde{\mathbf{G}}\mathbf{W}\right\|.\end{aligned}$$

Applying Lemma 21 to the first and third terms, recalling that $\|\widetilde{\mathbf{G}}\| = 1$,

$$\left\|\mathbf{u}^{\top}\left(\widehat{\mathbf{G}}^2 - \widetilde{\mathbf{G}}^2\right)\mathbf{W}\right\| \leq \left\|\mathbf{u}^{\top}\left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\mathbf{W}\right\| + o_p\left(\frac{\|\mathbf{W}\|\|\mathbf{u}\|}{\sqrt{n}}\right).$$

By submultiplicativity,

$$\begin{aligned}\left\|\mathbf{u}^{\top}\left(\widehat{\mathbf{G}}^2 - \widetilde{\mathbf{G}}^2\right)\mathbf{W}\right\| &\leq \|\mathbf{u}\|\|\mathbf{W}\|\left\|\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right\|^2 + o_p\left(\frac{\|\mathbf{W}\|\|\mathbf{u}\|}{\sqrt{n}}\right) \\ &\leq C\|\mathbf{u}\|\|\mathbf{W}\|\left(1 + \frac{\mathbf{s}_1}{\min_{i\in[n]}\tilde{d}_i}\right)^2 \frac{(\nu + b^2)n\log^2 n}{\min_{i\in[n]}\tilde{d}_i^2} + o_p\left(\frac{\|\mathbf{W}\|\|\mathbf{u}\|}{\sqrt{n}}\right),\end{aligned}$$

where the second inequality follows from Lemma 18. Applying our growth assumptions in Equations (28) and (25),

$$\left\|\mathbf{u}^\top \left(\widehat{\mathbf{G}}^2 - \widetilde{\mathbf{G}}^2\right) \mathbf{W}\right\| = o_p(\|\mathbf{u}\|). \tag{98}$$

Again using the triangle inequality and submultiplicativy of the norm,

$$\left\|\mathbf{u}^\top \left(\widehat{\mathbf{G}}\widehat{\mathbf{X}}\mathbf{Q}^\top - \widetilde{\mathbf{G}}\mathbf{X}\right)\right\| \le \left\|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right) \mathbf{X}\right\| + \left\|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)\right\| + \left\|\mathbf{u}^\top \widetilde{\mathbf{G}}\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)\right\|.$$

Applying Lemma 21 and recalling that $\|\mathbf{X}\| = O_p(\sqrt{n})$ by Equation (28),

$$\left\|\mathbf{u}^\top \left(\widehat{\mathbf{G}}\widehat{\mathbf{X}}\mathbf{Q}^\top - \widetilde{\mathbf{G}}\mathbf{X}\right)\right\| \le \left\|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)\right\| + \left\|\mathbf{u}^\top \widetilde{\mathbf{G}}\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)\right\| + o_p(\|\mathbf{u}\|).$$

By Lemma 15 with $\mathbf{B} = \widetilde{\mathbf{G}}^\top \mathbf{u}$ and our growth assumptions in Equations (20), (28) and (21) and recalling that $\|\widetilde{\mathbf{G}}\| = 1$,

$$\left\|\mathbf{u}^\top \left(\widehat{\mathbf{G}}\widehat{\mathbf{X}}\mathbf{Q}^\top - \widetilde{\mathbf{G}}\mathbf{X}\right)\right\| \le \left\|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)\right\| + o_p\left(\frac{\|\mathbf{X}\|\|\mathbf{u}\|}{\sqrt{n}}\right). \tag{99}$$

By submultiplicativity and Lemma 18,

$$\left\|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)\right\| \le C\|\mathbf{u}\|\left(1 + \frac{\mathbf{s}_1}{\min_{i\in[n]} \tilde{d}_i}\right)\frac{\sqrt{\nu + b^2}\sqrt{n}\log n}{\min_{i\in[n]} \tilde{d}_i}\left\|\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right\|.$$

Applying Lemma 15 with $\mathbf{B} = \mathbf{I}$ and collecting terms,

$$\begin{aligned}
&\left\|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)\right\| \\
&\le C\|\mathbf{u}\|\left(1 + \frac{\mathbf{s}_1}{\min_{i\in[n]} \tilde{d}_i}\right)\left[\frac{(\nu + b^2) n \log^2 n}{\sqrt{\mathbf{s}_d}\min_{i\in[n]} \tilde{d}_i} + \frac{\kappa(\mathbf{P})(\nu + b^2)^{3/2} n^{3/2}\log^3 n}{\mathbf{s}_d^{3/2}\min_{i\in[n]} \tilde{d}_i}\right] \\
&\le C\|\mathbf{u}\|\frac{\mathbf{s}_1}{\min_{i\in[n]} \tilde{d}_i}\left[\frac{(\nu + b^2) n \log^2 n}{\sqrt{\mathbf{s}_d}\min_{i\in[n]} \tilde{d}_i} + \frac{\kappa(\mathbf{P})(\nu + b^2)^{3/2} n^{3/2}\log^3 n}{\mathbf{s}_d^{3/2}\min_{i\in[n]} \tilde{d}_i}\right],
\end{aligned}$$

where the second inequality follows from the fact that $\mathbf{s}_1$ is an upper bound on the minimum degree. Applying our growth assumptions in Equations (23), (28), (23) and (24),

$$\left\|\mathbf{u}^\top \left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)\right\| = o_p(\|\mathbf{u}\|).$$

Applying this to Equation (99),

$$\left\|\mathbf{u}^\top \left(\widehat{\mathbf{G}}\widehat{\mathbf{X}}\mathbf{Q}^\top - \widetilde{\mathbf{G}}\mathbf{X}\right)\right\| = o_p(\|\mathbf{u}\|). \tag{100}$$

By an argument parallel to that yielding Equation (98), this time using the growth assumption in Equation (28),

$$\left\|\mathbf{u}^\top \left(\widehat{\mathbf{G}}^2 - \widetilde{\mathbf{G}}^2\right) \mathbf{X}\right\| = o_p(\|\mathbf{u}\|). \tag{101}$$

Applying Equations (96), (97), (98), (100) and (101) to Equation (95) completes the proof. □

Lemma 31. *Suppose that Assumptions 1 through 3 hold. Let* $\breve{\mathbf{H}}$ *and* $\widetilde{\breve{\mathbf{H}}}$ *be as given in Definition 3 and Equation (86), respectively, and let* $\mathbf{u} \in \mathbb{R}^n$ *be independent of* $\mathbf{A} - \mathbf{P}$ *given* $\mathbf{X}$*. Then under either the peer contagion model in Equation (4) with Assumption 5 or the latent contagion model in Equation (6) with Assumption 6, with* $\mathbf{Q}_\mathbf{H} \in \mathbb{R}^{(3p+3d)\times(3p+3d)}$ *as given in Equation (88),*

$$\left\|\mathbf{u}^\top\left(\breve{\mathbf{H}}\mathbf{Q}_\mathbf{H}^\top - \widetilde{\breve{\mathbf{H}}}\right)\right\| = o_p(\|\mathbf{u}\|).$$

Proof. Recalling the definitions of $\breve{\mathbf{H}}$ and $\widetilde{\breve{\mathbf{H}}}$ and applying the triangle inequality,

$$\begin{aligned}\left\|\mathbf{u}^\top\left(\breve{\mathbf{H}}\mathbf{Q}_\mathbf{H}^\top - \widetilde{\breve{\mathbf{H}}}\right)\right\| \le \left\|\mathbf{u}^\top\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)\right\| + \left\|\mathbf{u}^\top\left(\mathbf{G} - \widetilde{\mathbf{G}}\right)\mathbf{W}\right\| + \left\|\mathbf{u}^\top\left(\mathbf{G}\widehat{\mathbf{X}}\mathbf{Q}^\top - \widetilde{\mathbf{G}}\mathbf{X}\right)\right\| \\ + \left\|\mathbf{u}^\top\left(\mathbf{G}^2 - \widetilde{\mathbf{G}}^2\right)\mathbf{W}\right\| + \left\|\mathbf{u}^\top\left(\mathbf{G}^2\widehat{\mathbf{X}}\mathbf{Q}^\top - \widetilde{\mathbf{G}}^2\mathbf{X}\right)\right\|.\end{aligned}$$

The result then follows from an argument analogous to that in the proof of Lemma 30, using Lemma 17 in place of Lemma 18 and Lemma 20 in place of Lemma 21. □

Lemma 32. *Suppose that Assumptions 1 through 3 hold. Let* $\breve{\mathbf{H}}$ *and* $\mathbf{H}$ *be as given in Definition 3 and Equation (87) and let* $\mathbf{u} \in \mathbb{R}^n$ *be independent of* $\mathbf{A} - \mathbf{P}$ *given* $\mathbf{X}$*. Then under either the peer contagion model in Equation (4) with Assumption 5 or the latent contagion model in Equation (6) with Assumption 6, with* $\mathbf{Q}_\mathbf{H} \in \mathbb{R}^{(3p+3d)\times(3p+3d)}$ *as given in Equation (88),*

$$\left\|\mathbf{u}^\top\left(\breve{\mathbf{H}}\mathbf{Q}_\mathbf{H}^\top - \mathbf{H}\right)\right\| = o_p(\|\mathbf{u}\|).$$

Proof. Recalling the definitions of $\breve{\mathbf{H}}$ and $\mathbf{H}$ from Definition 3 and Equation (87), respectively, and applying the triangle inequality,

$$\left\|\mathbf{u}^\top\left(\breve{\mathbf{H}}\mathbf{Q}_\mathbf{H}^\top - \mathbf{H}\right)\right\| \le \left\|\mathbf{u}^\top\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)\right\| + \left\|\mathbf{u}^\top\mathbf{G}\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)\right\| + \left\|\mathbf{u}^\top\mathbf{G}^2\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)\right\|.$$

Three applications of Lemma 15 with $\mathbf{B} = \mathbf{u}$, $\mathbf{B} = \mathbf{G}^\top\mathbf{u}$ and $\mathbf{B} = (\mathbf{G}^2)^\top\mathbf{u}$ yields, after recalling that $\|(\mathbf{G}^2)^\top\mathbf{u}\| \le \|\mathbf{G}^\top\mathbf{u}\| \le \|\mathbf{u}\|$ since $\|\mathbf{G}\| = 1$,

$$\left\|\mathbf{u}^\top\left(\breve{\mathbf{H}}\mathbf{Q}_\mathbf{H}^\top - \mathbf{H}\right)\right\| \le \frac{C\|\mathbf{u}\|\sqrt{\nu + b^2}\log n}{\sqrt{\mathbf{s}_d}} + \frac{C\kappa(\mathbf{P})\|\mathbf{u}\|(\nu + b^2)n\log^2 n}{\mathbf{s}_d^{3/2}}.$$

Applying the growth assumptions in Equations (20), (28) and (21) completes the proof. □

Lemma 33. *Let* $\mathbf{u} \in \mathbb{R}^n$ *be independent of* $\mathbf{A} - \mathbf{P}$ *given* $\mathbf{X}$ *and suppose that Assumptions 1 through 3 hold. Then under either the peer contagion model in Equation (4) with Assumption 5 or the latent contagion model in Equation (6) with Assumption 6,*

$$\left\|\mathbf{u}^\top\left(\mathbf{H} - \widetilde{\breve{\mathbf{H}}}\right)\right\| = o_p\left(\sqrt{\|\mathbf{u}\|}\right).$$

Proof. Letting $\mathbf{Q}_\mathbf{H} \in \mathbb{R}^{(3p+3d)\times(3p+3d)}$ be as in Equation (88), the triangle inequality yields

$$\left\|\mathbf{u}^\top\left(\mathbf{H} - \widetilde{\breve{\mathbf{H}}}\right)\right\| \le \left\|\mathbf{u}^\top\left(\breve{\mathbf{H}}\mathbf{Q}_\mathbf{H}^\top - \mathbf{H}\right)\right\| + \left\|\mathbf{u}^\top\left(\breve{\mathbf{H}}\mathbf{Q}_\mathbf{H}^\top - \widetilde{\breve{\mathbf{H}}}\right)\right\|,$$

and Lemmas 32 and 31 complete the proof. □

F Controlling Projections

Here we collect results controlling the projection matrices used in our two-stage least squares estimators. Our primary task is to relate the projection $\widehat{\mathbf{M}}$, from Definition 4 and the projection $\widetilde{\mathbf{M}}$, from Definition 4, to their oracle versions,

$$\widetilde{\mathbf{M}} = \widetilde{\mathbf{H}}\left(\widetilde{\mathbf{H}}^\top \widetilde{\mathbf{H}}\right)\widetilde{\mathbf{H}}^\top, \tag{102}$$

and

$$\mathbf{M} = \mathbf{H}(\mathbf{H}^\top \mathbf{H})^{-1}\mathbf{H}^\top, \tag{103}$$

where $\widetilde{\mathbf{H}} \in \mathbb{R}^{n\times(3p+3d)}$ is as defined in Equation (86) and $\mathbf{H}$ is as defined in Equation (87).

Lemma 34. *Suppose that Assumptions 1, 2, 3 and 4 hold. With $\widehat{\mathbf{M}}$ as given in Definition 4 and $\widetilde{\mathbf{M}}$ as defined in Equation (102), under either the latent contagion model in Equation (6) with Assumption 6 or the peer contagion model in Equation (4) with Assumption 5,*

$$\left\|\widehat{\mathbf{M}} - \widetilde{\mathbf{M}}\right\| = o_p\left(\frac{1}{\sigma_{\min}(\widetilde{\mathbf{H}})}\right).$$

Proof. Let and $\mathbf{U}_{\widehat{\mathbf{H}}}$ and $\mathbf{U}_{\widetilde{\mathbf{H}}}$ be the left singular vectors of $\widehat{\mathbf{H}}$ and $\mathbf{H}$, respectively. Since $\widehat{\mathbf{M}}$ and $\widetilde{\mathbf{M}}$ are projection matrices,

$$\begin{aligned}\widehat{\mathbf{M}} &= \mathbf{U}_{\widehat{\mathbf{H}}}\mathbf{U}_{\widehat{\mathbf{H}}}^\top\\ \widetilde{\mathbf{M}} &= \mathbf{U}_{\widetilde{\mathbf{H}}}\mathbf{U}_{\widetilde{\mathbf{H}}}^\top.\end{aligned}$$

$\mathbf{U}_{\widehat{\mathbf{H}}}$ and $\mathbf{U}_{\widetilde{\mathbf{H}}}$ are the eigenvectors of $\widehat{\mathbf{H}}\widehat{\mathbf{H}}^\top$ and $\widetilde{\mathbf{H}}\widetilde{\mathbf{H}}^\top$ respectively, and thus by the Davis-Kahan $\sin\Theta$ theorem (Bhatia, 1997; Y. Yu, Wang, and Samworth, 2015),

$$\left\|\widehat{\mathbf{M}} - \widetilde{\mathbf{M}}\right\| = \left\|\mathbf{U}_{\widehat{\mathbf{H}}}\mathbf{U}_{\widehat{\mathbf{H}}}^\top - \mathbf{U}_{\widetilde{\mathbf{H}}}\mathbf{U}_{\widetilde{\mathbf{H}}}^\top\right\| \le \frac{C\left\|\widehat{\mathbf{H}}\widehat{\mathbf{H}}^\top - \widetilde{\mathbf{H}}\widetilde{\mathbf{H}}^\top\right\|}{\sigma_{\min}^2\left(\widetilde{\mathbf{H}}\right)}, \tag{104}$$

where $\widetilde{\mathbf{H}}$ is as defined in Equation (86) and $\widehat{\mathbf{H}}$ is as in Definition 4.

Let $\mathbf{Q}_{\mathbf{H}} \in \mathbb{R}^{3(p+d)\times 3(p+d)}$ be as in Equation (88), and recall that $\mathbf{Q}_{\mathbf{H}}$ is orthogonal by construction. By the triangle inequality,

$$\left\|\widehat{\mathbf{H}}\widehat{\mathbf{H}}^\top - \widetilde{\mathbf{H}}\widetilde{\mathbf{H}}^\top\right\| = \left\|\widehat{\mathbf{H}}\mathbf{Q}_{\mathbf{H}}^\top\mathbf{Q}_{\mathbf{H}}\widehat{\mathbf{H}}^\top - \widetilde{\mathbf{H}}\widetilde{\mathbf{H}}^\top\right\| \le 2\left\|\left(\widehat{\mathbf{H}}\mathbf{Q}_{\mathbf{H}}^\top - \widetilde{\mathbf{H}}\right)\widetilde{\mathbf{H}}^\top\right\| + \left\|\widehat{\mathbf{H}}\mathbf{Q}_{\mathbf{H}}^\top - \widetilde{\mathbf{H}}\right\|^2.$$

Controlling the first term with Lemma 30 and the second term with Lemma 26,

$$\left\|\widehat{\mathbf{H}}\widehat{\mathbf{H}}^\top - \widetilde{\mathbf{H}}\widetilde{\mathbf{H}}^\top\right\| \le o_p\left(\|\widetilde{\mathbf{H}}\|\right) + o_p\left(\sigma_{\min}(\widetilde{\mathbf{H}})\right) = o_p\left(\|\widetilde{\mathbf{H}}\|\right),$$

where we have used the growth rate in Equation (28). Applying this to Equation (104),

$$\left\|\widehat{\mathbf{M}} - \widetilde{\mathbf{M}}\right\| = o_p\left(\frac{\kappa(\widetilde{\mathbf{H}})}{\sigma_{\min}(\widetilde{\mathbf{H}})}\right).$$

Our growth assumption in Equation (30) completes the proof. □

Lemma 35. *Suppose that Assumptions 1, 2 and 3 hold. With* $\breve{\mathbf{M}}$ *as given in Definition 3 and* $\widetilde{\mathbf{M}}$ *as defined in Equation (102), under either the latent contagion model in Equation (6) with Assumption 6 or the peer contagion model in Equation (4) with Assumption 5,*

$$\left\|\breve{\mathbf{M}} - \widetilde{\mathbf{M}}\right\| = o_p\left(\frac{1}{\sigma_{\min}(\widetilde{\mathbf{H}})}\right).$$

Proof. The proof follows by the same argument as in the proof of Lemma 34, using Lemmas 31 and 27 in place of, respectively, Lemmas 30 and 26. Details are omitted. □

Lemma 36. *Suppose that Assumptions 1, 2 and 3 hold. With* $\breve{\mathbf{M}}$ *as given in Definition 3 and* $\mathbf{M}$ *as defined in Equation (103), under either the latent contagion model in Equation (6) with Assumption 6 or the peer contagion model in Equation (4) with Assumption 5,*

$$\left\|\breve{\mathbf{M}} - \mathbf{M}\right\| = o_p\left(\frac{1}{\sigma_{\min}(\widetilde{\mathbf{H}})}\right).$$

Proof. The proof follows from the same argument as Lemma 34, using Lemmas 32 and 28 in place of, respectively, Lemmas 30 and 26. Details are omitted. □

Lemma 37. *Suppose that Assumptions 1, 2 and 3 hold. With* $\mathbf{M}$ *as defined in Equation (103) and* $\widetilde{\mathbf{M}}$ *as defined in Equation (102), under either the latent contagion model in Equation (6) with Assumption 6 or the peer contagion model in Equation (4) with Assumption 5,*

$$\left\|\mathbf{M} - \widetilde{\mathbf{M}}\right\| = o_p\left(\frac{1}{\sigma_{\min}(\widetilde{\mathbf{H}})}\right).$$

Proof. The proof follows from the same argument as Lemma 34, using Lemmas 33 and 29 in place of, respectively, Lemmas 30 and 26. Details are omitted. □

Lemma 38. *Suppose that Assumptions 1, 2, 3 and 4 hold. With* $\widehat{\mathbf{M}}$ *as given by Definition 3 and* $\mathbf{M}$ *as defined in Equation (103), under either the latent contagion model in Equation (6) with Assumption 6 or the peer contagion model in Equation (4) with Assumption 5,*

$$\left\|\widehat{\mathbf{M}} - \mathbf{M}\right\| = o_p\left(\frac{1}{\sigma_{\min}(\widetilde{\mathbf{H}})}\right).$$

Proof. Applying the triangle inequality,

$$\left\|\widehat{\mathbf{M}} - \mathbf{M}\right\| \le \left\|\widehat{\mathbf{M}} - \widetilde{\mathbf{M}}\right\| + \left\|\mathbf{M} - \widetilde{\mathbf{M}}\right\|,$$

and Lemmas 34 and 37 complete the proof. □

G Controlling the Design Matrices

The following results are aimed toward controlling the behavior of our design matrices $\widehat{\mathbf{Z}}, \breve{\mathbf{Z}}, \mathbf{Z}$ and $\widetilde{\mathbf{Z}}$, as well as their interactions with the projections onto the span of the instrument (i.e., $\widehat{\mathbf{M}}, \breve{\mathbf{M}}, \mathbf{M}$ and $\widetilde{\mathbf{M}}$). As in our results on latent position estimation in Appendix B, we must account for the rotational non-identifiability of the latent positions.

This is achieved via an orthogonal matrix $\mathbf{Q_Z} \in \mathbb{R}^{(2+p+d)\times(2+p+d)}$, defined according to

$$\mathbf{Q_Z} = \begin{bmatrix} 1 & 0_{1\times p} & 0_{1\times d} & 0 \\ 0_{p\times 1} & \mathbf{I}_{p\times p} & 0_{p\times d} & 0 \\ 0_{d\times 1} & 0_{d\times p} & \mathbf{Q} & 0_{d\times 1} \\ 0 & 0_{1\times p} & 0_{1\times d} & 1 \end{bmatrix}, \tag{105}$$

where $\mathbf{Q} \in \mathbb{R}^{d\times d}$ is the orthogonal matrix guaranteed by Lemma 15.

Lemma 39. *Suppose that Assumptions 1, 2, 3 and 4 hold, Then under either the latent contagion model in Equation (6) with Assumption 6 or the peer contagion model in Equation (4) with Assumption 5, with $\mathbf{Q_Z} \in \mathbb{R}^{(2+p+d)\times(2+p+d)}$ as given by Equation (105),*

$$\left\|\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}}\right\| = o_p\left(\sigma_{\min}(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})\right).$$

Proof. Recalling the definitions of $\widehat{\mathbf{Z}}$ and $\widetilde{\mathbf{Z}}$ from Definition (4) and Equation (110), respectively, the triangle inequality and construction of $\mathbf{Q_Z}$ imply

$$\left\|\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}}\right\| \leq \left\|\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right\| + \left\|\left(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}\right)\mathbf{Y}\right\|. \tag{106}$$

Applying Lemma 15 with $\mathbf{B} = \mathbf{I}$,

$$\left\|\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right\| \leq \frac{C\sqrt{n}\sqrt{\nu + b^2}\log n}{\sqrt{\mathbf{s}_d}} + \frac{C\kappa(\mathbf{P})(\nu + b^2)n\log^2 n}{\mathbf{s}_d^{3/2}},$$

and our growth assumptions in Equations (20) (28) and (21) imply

$$\left\|\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right\| = o_p\left(\sqrt{n}\right) = o_p\left(\sigma_{\min}(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})\right), \tag{107}$$

where the second equality follows from our assumption in Equation (29) z Submultiplicativity and Lemma 18 yield

$$\left\|(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}})\mathbf{Y}\right\| \leq C\left(1 + \frac{\mathbf{s}_1}{\min_{i\in[n]}\tilde{d}_i}\right)\frac{\|\mathbf{Y}\|\sqrt{\nu + b^2}\sqrt{n}\log n}{\min_{i\in[n]}\tilde{d}_i} = o_p(\|\mathbf{Y}\|),$$

where the equality follows from our growth assumption in Equation (25). Using Lemma 23 or Lemma 24 to control $\|\mathbf{Y}\| = O_p\left(\sqrt{n}\right)$ depending on whether latent contagion or peer contagion, respectively, is assumed,

$$\left\|(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}})\mathbf{Y}\right\| = o_p\left(\sqrt{n}\right) = o_p\left(\sigma_{\min}(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})\right),$$

where the second equality follows from our growth assumption in Equation (29). Applying this and Equation (107) to Equation (106) completes the proof. □

Lemma 40. *Suppose that Assumptions 1, 2 and 3 hold. Then under either the latent contagion model in Equation (6) with Assumption 6 or the peer contagion model in Equation (4) with Assumption 5, with $\mathbf{Q_Z} \in \mathbb{R}^{(2+p+d)\times(2+p+d)}$ as in Equation (105),*

$$\left\|\breve{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}}\right\| = o_p\left(\sigma_{\min}(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})\right).$$

Proof. Recalling the definitions of $\breve{\mathbf{Z}}$ and $\widetilde{\mathbf{Z}}$ from Definition (3) and Equation (110), the triangle inequality implies

$$\left\|\breve{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^{\top} - \widetilde{\mathbf{Z}}\right\| \leq \left\|\widehat{\mathbf{X}}\mathbf{Q}^{\top} - \mathbf{X}\right\| + \left\|\left(\mathbf{G} - \widetilde{\mathbf{G}}\right)\mathbf{Y}\right\|.$$

The proof then follows by the same argument as in Lemma 39, using Lemma 17 in place of Lemma 18. ☐

Lemma 41. *Suppose that Assumptions 1, 2 and 3 hold. Then under either the latent contagion model in Equation (6) with Assumption 6 or the peer contagion model in Equation (4) with Assumption 5, with $\mathbf{Q}_{\mathbf{Z}} \in \mathbb{R}^{(2+p+d)\times(2+p+d)}$ as in Equation (105),*

$$\left\|\breve{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^{\top} - \mathbf{Z}\right\| = o_p\big(\sigma_{\min}(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})\big).$$

Proof. Recalling the definitions of $\breve{\mathbf{Z}}$ and $\mathbf{Z}$ from Definition (3) and Equation (129) and using the definition of $\mathbf{Q}_{\mathbf{Z}}$, the triangle inequality implies

$$\left\|\breve{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^{\top} - \mathbf{Z}\right\| \leq \left\|\widehat{\mathbf{X}}\mathbf{Q}^{\top} - \mathbf{X}\right\| + \left\|\left(\mathbf{G} - \mathbf{G}\right)\mathbf{Y}\right\| = \left\|\widehat{\mathbf{X}}\mathbf{Q}^{\top} - \mathbf{X}\right\|. \tag{108}$$

Applying Lemma 15 with $\mathbf{B} = \mathbf{I}$,

$$\left\|\breve{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^{\top} - \mathbf{Z}\right\| \leq \frac{C\sqrt{n}\sqrt{\nu + b^2}\log n}{\sqrt{\mathbf{s}_d}} + \frac{C\kappa(\mathbf{P})(\nu + b^2)n\log^2 n}{\mathbf{s}_d^{3/2}}.$$

Applying our growth assumptions in Equations (20), (28) and (21),

$$\left\|\breve{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^{\top} - \mathbf{Z}\right\| = o_p(\sqrt{n}),$$

and our growth assumption in Equation (29) completes the proof. ☐

Lemma 42. *Suppose that Assumptions 1, 2 and 3 hold. Then under either the latent contagion model in Equation (6) with Assumption 6 or the peer contagion model in Equation (4) with Assumption 5,*

$$\left\|\mathbf{Z} - \widetilde{\mathbf{Z}}\right\| = o_p\big(\|\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\|\big).$$

Proof. Recalling the definitions of $\mathbf{Z}$ and $\widetilde{\mathbf{Z}}$ from Equations (129) and (110), respectively, the triangle inequality implies

$$\left\|\mathbf{Z} - \widetilde{\mathbf{Z}}\right\| \leq \left\|\left(\mathbf{G} - \widetilde{\mathbf{G}}\right)\mathbf{Y}\right\|.$$

The proof then follows by the same argument as given in the second half of the proof of Lemma 39, using Lemma 17 in place of Lemma 18. ☐

Lemma 43. *Suppose that Assumptions 1, 2, 3 and 4 hold. Then under either the latent contagion model in Equation (6) with Assumption 6 or the peer contagion model in Equation (4) with Assumption 5, with $\mathbf{Q}_{\mathbf{Z}} \in \mathbb{R}^{(2+p+d)\times(2+p+d)}$ as in Equation (105),*

$$\left\|\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^{\top} - \mathbf{Z}\right\| = o_p\big(\|\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\|\big).$$

Proof. By the triangle inequality,

$$\left\|\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^{\top} - \mathbf{Z}\right\| \leq \left\|\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^{\top} - \widetilde{\mathbf{Z}}\right\| + \left\|\mathbf{Z} - \widetilde{\mathbf{Z}}\right\|,$$

and Lemmas 39 and 42 complete the proof. □

Lemma 44. *Suppose that Assumptions 1, 2, 3 and 4 hold. Then under either the latent contagion model in Equation (6) with Assumption 6 or the peer contagion model in Equation (4) with Assumption 5, with* $\mathbf{Q}_{\mathbf{Z}} \in \mathbb{R}^{(2+p+d)\times(2+p+d)}$ *as given by Equation (105),*

$$\left\|\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^{\top} - \widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right\| = o_p\big(\sigma_{\min}(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})\big).$$

Proof. By the triangle inequality,

$$\left\|\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^{\top} - \widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right\| \leq \left\|(\widehat{\mathbf{M}} - \widetilde{\mathbf{M}})\widetilde{\mathbf{Z}}\right\| + \left\|\widehat{\mathbf{M}}\left(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^{\top} - \widetilde{\mathbf{Z}}\right)\right\| \tag{109}$$

Since $\widehat{\mathbf{M}}$ is a projection, submultiplicativity and Lemma 39 yields

$$\left\|\widehat{\mathbf{M}}\left(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^{\top} - \widetilde{\mathbf{Z}}\right)\right\| = o_p\big(\sigma_{\min}(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})\big).$$

Similarly, submultiplicativity and Lemma 34 yield

$$\left\|(\widehat{\mathbf{M}} - \widetilde{\mathbf{M}})\widetilde{\mathbf{Z}}\right\| = o_p\left(\frac{\|\widetilde{\mathbf{Z}}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})}\right) = o_p\big(\sigma_{\min}(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})\big),$$

where the second equality follows from our growth assumptions in Equations (30) and (29). Applying the above two bounds to Equation (109) yields the desired result. □

Lemma 45. *Suppose that Assumptions 1, 2 and 3 hold. Then under either the latent contagion model in Equation (6) with Assumption 6 or the peer contagion model in Equation (4) with Assumption 5, with* $\mathbf{Q}_{\mathbf{Z}} \in \mathbb{R}^{(2+p+d)\times(2+p+d)}$ *as given by Equation (105),*

$$\left\|\widetilde{\mathbf{M}}\breve{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^{\top} - \widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right\| = o_p\big(\sigma_{\min}(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})\big).$$

Proof. The proof follows by the same argument as Lemma 44, using Lemma 40 in place of Lemma 39 and Lemma 35 in place of Lemma 34. □

Lemma 46. *Suppose that Assumptions 1, 2 and 3 hold. Then under either the latent contagion model in Equation (6) with Assumption 6 or the peer contagion model in Equation (4) with Assumption 5, with* $\mathbf{Q}_{\mathbf{Z}} \in \mathbb{R}^{(2+p+d)\times(2+p+d)}$ *as given by Equation (105),*

$$\left\|\mathbf{M}\mathbf{Z} - \widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right\| = o_p\big(\sigma_{\min}(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})\big).$$

Proof. The proof follows the same argument as the proof of Lemma 44, using Lemma 42 in place of Lemma 39 and Lemma 37 in place of Lemma 34. □

H Convergence of Oracle Estimators

Our proofs of Theorems 1, 2, 3 and 4 rely on showing that the estimates $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$ are close to "oracle" estimates based on using the true latent positions rather than estimates thereof (see Sections I and J below). The following theorem implies that these oracle estimates are asymptotically normal about their population targets, so that we need only to show that the four estimates listed above are close to the oracle estimates. This latter argument is carried out in Sections I and J below.

Theorem 6 (Theorem 3 of Harry H Kelejian and Ingmar R Prucha 1998)**.** *Under the peer contagion model of Equation* (3), *suppose that Assumption 5 holds, so that*

$$\mathbf{Y} = \left(\mathbf{I} - \beta_{\mathrm{y}}\mathbf{G}\right)^{-1}(\mathbf{1}_n\beta_0 + \mathbf{W}\boldsymbol{\beta}_{\mathrm{w}} + \mathbf{X}\boldsymbol{\beta}_{\mathrm{x}} + \boldsymbol{\varepsilon}).$$

Then

$$\sqrt{n}(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) \to \mathcal{N}\Big(0, \sigma_{\varepsilon}^2\left(\mathbf{Z}^\top\mathbf{H}\mathbf{Z}\right)^{-1}\Big).$$

An analogous result holds, replacing $\mathbf{G}$ *with* $\widetilde{\mathbf{G}}$, *under Assumption 6 instead of Assumption 5.*

I Latent Contagion Data Model

Here we provide proof details of Theorem 2 and 4, which describe the behavior of our estimators when the responses $\mathbf{Y}$ are driven by contagion on the graph structure encoded by the latent positions $\mathbf{X}$, i.e., what we have termed *latent contagion*, as given in Equation (6). If we had access to this latent structure, we could construct oracle analogues of the estimator, where $\widetilde{\mathbf{M}}$ is as defined in Equation (102) and, recalling the definition of $\widetilde{\mathbf{Z}}$ from Equation (11),

$$\widetilde{\mathbf{Z}} = \left[\mathbf{1}_n \; \mathbf{W} \; \mathbf{X} \; \widetilde{\mathbf{G}}\mathbf{Y}\right] \in \mathbb{R}^{n\times(p+d+2)}. \tag{110}$$

We begin by proving convergence of the latent contagion estimators under the latent contagion data model. We give analogous proofs for the peer contagion estimators in Section I.2 below.

I.1 Latent Contagion Estimators

Our proof of Theorem 2 relies on showing that when the model in Equation (6) holds, our estimator in Equation (10) is suitably close to an "oracle" version of the estimator that uses $\widetilde{\mathbf{Z}}$,

$$\tilde{\boldsymbol{\theta}} = (\widetilde{\mathbf{Z}}^\top\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})^{-1}\left(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right)^\top\mathbf{Y}. \tag{111}$$

This is established in Lemma 47 for the TSLS estimator.

Lemma 47. *Under the latent contagion model in Equation* (6), *suppose that Assumptions 1 through 4 hold. Then there exists a sequence of orthogonal matrices* $\mathbf{Q}_{\mathbf{Z}} \in \mathbb{R}^{(2+p+d)\times(2+p+d)}$ *such that*

$$\sqrt{n}\left(\mathbf{Q}_{\mathbf{Z}}\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}}\right) = o_p(1).$$

Proof. Take $\mathbf{Q_Z}$ to be as in Equation (105). Recalling the definitions in Equations (10) and (111),

$$\mathbf{Q_Z}\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}} = \left[\mathbf{Q_Z}(\widehat{\mathbf{Z}}^\top\widehat{\mathbf{M}}\widehat{\mathbf{Z}})^{-1}\left(\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\right)^\top - (\widetilde{\mathbf{Z}}^\top\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})^{-1}\left(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right)^\top\right]\mathbf{Y}.$$

Using the fact that $\mathbf{Q_Z}$ is orthogonal, the triangle inequality yields

$$\begin{aligned}\left\|\mathbf{Q_Z}\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}}\right\| \le &\left\|\mathbf{Q_Z}(\widehat{\mathbf{Z}}^\top\widehat{\mathbf{M}}\widehat{\mathbf{Z}})^{-1}\mathbf{Q}_\mathbf{Z}^\top\left(\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right)^\top\mathbf{Y}\right\| \\ &+ \left\|\left[\mathbf{Q_Z}(\widehat{\mathbf{Z}}^\top\widehat{\mathbf{M}}\widehat{\mathbf{Z}})^{-1}\mathbf{Q}_\mathbf{Z}^\top - (\widetilde{\mathbf{Z}}^\top\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})^{-1}\right]\left(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right)^\top\mathbf{Y}\right\|.\end{aligned} \tag{112}$$

We will control each of these right-hand terms separately.

Controlling the first term in Equation (112)

By submultiplicativity and unitary invariance of the norm,

$$\left\|\mathbf{Q_Z}(\widehat{\mathbf{Z}}^\top\widehat{\mathbf{M}}\widehat{\mathbf{Z}})^{-1}\mathbf{Q}_\mathbf{Z}^\top(\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})^\top\mathbf{Y}\right\| \le \frac{\left\|(\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})^\top\mathbf{Y}\right\|}{\sigma_{\min}^2(\widehat{\mathbf{M}}\widehat{\mathbf{Z}})} \le \frac{C\left\|(\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})^\top\mathbf{Y}\right\|}{\sigma_{\min}^2(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})},$$

where the second inequality follows from Weyl's inequality and Lemma 44. Using our assumption in Equation (29), it follows that

$$\left\|\mathbf{Q_Z}(\widehat{\mathbf{Z}}^\top\widehat{\mathbf{M}}\widehat{\mathbf{Z}})^{-1}\mathbf{Q}_\mathbf{Z}^\top\left(\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right)^\top\mathbf{Y}\right\| \le \frac{C\left\|\left(\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right)^\top\mathbf{Y}\right\|}{n}. \tag{113}$$

By the triangle inequality,

$$\left\|\left(\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right)^\top\mathbf{Y}\right\| \le \left\|\mathbf{Q_Z}\widehat{\mathbf{Z}}^\top\left(\widehat{\mathbf{M}} - \widetilde{\mathbf{M}}\right)\mathbf{Y}\right\| + \left\|\left(\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}}\right)^\top\widetilde{\mathbf{M}}\mathbf{Y}\right\|. \tag{114}$$

Again using the triangle inequality,

$$\left\|\mathbf{Q_Z}\widehat{\mathbf{Z}}^\top\left(\widehat{\mathbf{M}} - \widetilde{\mathbf{M}}\right)\mathbf{Y}\right\| \le \left\|\left(\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}}\right)^\top\left(\widehat{\mathbf{M}} - \widetilde{\mathbf{M}}\right)\mathbf{Y}\right\| + \left\|\widetilde{\mathbf{Z}}^\top\left(\widehat{\mathbf{M}} - \widetilde{\mathbf{M}}\right)\mathbf{Y}\right\|. \tag{115}$$

By submultiplicativity, Lemma 34 and the assumptions in Equations (30) and (29),

$$\left\|\widetilde{\mathbf{Z}}^\top\left(\widehat{\mathbf{M}} - \widetilde{\mathbf{M}}\right)\mathbf{Y}\right\| = o_p\left(\frac{\|\widetilde{\mathbf{Z}}\|\|\mathbf{Y}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})}\right) = o_p(\sqrt{n}),$$

where we have used Lemma 23 to ensure $\|\mathbf{Y}\| = O_p(\sqrt{n})$, Using submultiplicativity followed by Lemmas 34 and 39,

$$\left\|\left(\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}}\right)^\top\left(\widehat{\mathbf{M}} - \widetilde{\mathbf{M}}\right)\mathbf{Y}\right\| \le \left\|\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}}\right\|\left\|\widehat{\mathbf{M}} - \widetilde{\mathbf{M}}\right\|\|\mathbf{Y}\| = o_p\left(\frac{\sigma_{\min}(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})\|\mathbf{Y}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})}\right) = o_p(\sqrt{n}),$$

where the last equality follows from the growth assumption in Equation (30) and using Lemma 23 to ensure $\|\mathbf{Y}\| = O_p(\sqrt{n})$.

Applying the above two displays to Equation (115),

$$\left\| \mathbf{Q_Z} \widehat{\mathbf{Z}}^\top \left(\widehat{\mathbf{M}} - \widetilde{\mathbf{M}} \right) \mathbf{Y} \right\| = o_p(\sqrt{n}).$$

Applying this to Equation (114),

$$\left\| \left(\widehat{\mathbf{M}} \widehat{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{M}} \widetilde{\mathbf{Z}} \right)^\top \mathbf{Y} \right\| \le \left\| \left(\widehat{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}} \right)^\top \widetilde{\mathbf{M}} \mathbf{Y} \right\| + o_p(\sqrt{n}). \tag{116}$$

Recalling the structure of $\widehat{\mathbf{Z}}$ and $\widetilde{\mathbf{Z}}$ from Definition 4 and Equation (110), respectively, and applying the triangle inequality,

$$\left\| (\widehat{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}})^\top \widetilde{\mathbf{M}} \mathbf{Y} \right\| \le \left\| (\widehat{\mathbf{X}} \mathbf{Q}^\top - \mathbf{X})^\top \widetilde{\mathbf{M}} \mathbf{Y} \right\| + \left| \mathbf{Y}^\top \widetilde{\mathbf{M}} (\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}) \mathbf{Y} \right|. \tag{117}$$

Applying Lemma 15 with $\mathbf{B} = \widetilde{\mathbf{M}}\mathbf{Y}$, noting that $\|\widetilde{\mathbf{M}}\mathbf{Y}\| \le \|\mathbf{Y}\|$ since $\widetilde{\mathbf{M}}$ is a projection,

$$\|(\widehat{\mathbf{X}} \mathbf{Q}^\top - \mathbf{X})^\top \widetilde{\mathbf{M}} \mathbf{Y}\| \le C \|\mathbf{Y}\| \left(\frac{\sqrt{\nu + b^2} \log n}{\sqrt{\mathbf{s}_d}} + \frac{\kappa(\mathbf{P})(\nu + b^2) n \log^2 n}{\mathbf{s}_d^{3/2}} \right) = o_p(\sqrt{n}),$$

where the second equality follows from the growth rates in Equations (20), (28) and (21) and using Lemma 23 to ensure $\|\mathbf{Y}\| = O_p(\sqrt{n})$.

Noting that $\mathbf{Y}$ is independent of $\mathbf{A} - \mathbf{P}$ given $\mathbf{X}$ under the latent contagion model, Lemma 21 yields

$$\left| \mathbf{Y}^\top \widetilde{\mathbf{M}} (\widehat{\mathbf{G}} - \widetilde{\mathbf{G}}) \mathbf{Y} \right| = o_p\left(\frac{\|\mathbf{Y}\|^2}{\sqrt{n}} \right) = o_p(\sqrt{n}),$$

where we have used the fact that $\|\widetilde{\mathbf{M}}\mathbf{Y}\| \le \|\mathbf{Y}\|$ trivially and again used Lemma 23 to ensure $\|\mathbf{Y}\| = O_p(\sqrt{n})$. Applying the above two displays to Equation (117),

$$\left\| (\widehat{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}})^\top \widetilde{\mathbf{M}} \mathbf{Y} \right\| = o_p(\sqrt{n}).$$

Applying this to Equation (116), and applying the resulting bound to Equation (113),

$$\left\| \mathbf{Q_Z} (\widehat{\mathbf{Z}}^\top \widehat{\mathbf{M}} \widehat{\mathbf{Z}})^{-1} \mathbf{Q}_\mathbf{Z}^\top \left(\widehat{\mathbf{M}} \widehat{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{M}} \widetilde{\mathbf{Z}} \right)^\top \mathbf{Y} \right\| = o_p\left(\frac{1}{\sqrt{n}} \right). \tag{118}$$

Controlling the second term in Equation (112)

By Lemma 44 and our growth assumption in Equation (29), both $\widehat{\mathbf{Z}}^\top \widehat{\mathbf{M}} \widehat{\mathbf{Z}}$ and $\widetilde{\mathbf{Z}}^\top \widetilde{\mathbf{M}} \widetilde{\mathbf{Z}}$ are invertible with high probability for all n suitably large. Factoring appropriately, applying submultiplicativity and recalling that $\widehat{\mathbf{M}}$ and $\widetilde{\mathbf{M}}$ are projections,

$$\left\| \left[\mathbf{Q_Z} (\widehat{\mathbf{Z}}^\top \widehat{\mathbf{M}} \widehat{\mathbf{Z}})^{-1} \mathbf{Q}_\mathbf{Z}^\top - (\widetilde{\mathbf{Z}}^\top \widetilde{\mathbf{M}} \widetilde{\mathbf{Z}})^{-1} \right] \left(\widetilde{\mathbf{M}} \widetilde{\mathbf{Z}} \right)^\top \mathbf{Y} \right\| \le \frac{\left\| \mathbf{Q_Z} \widehat{\mathbf{Z}}^\top \widehat{\mathbf{M}} \widehat{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}}^\top \widetilde{\mathbf{M}} \widetilde{\mathbf{Z}} \right\| \|\widetilde{\mathbf{Z}}\| \|\mathbf{Y}\|}{\sigma_{\min}^2(\widehat{\mathbf{M}} \widehat{\mathbf{Z}}) \sigma_{\min}^2(\widetilde{\mathbf{M}} \widetilde{\mathbf{Z}})}.$$

Using Lemma 44 again along with Weyl's inequality, we can ensure that $\sigma_{\min}(\widehat{\mathbf{M}} \widehat{\mathbf{Z}}) = \Omega_p(\sigma_{\min}(\widetilde{\mathbf{M}} \widetilde{\mathbf{Z}}))$. Our growth assumptions in Equations (29) and (29), as well as using Lemma 23 to ensure $\|\mathbf{Y}\| = O_p(\sqrt{n})$, then yield

$$\left\| \mathbf{Q_Z} \left[(\widehat{\mathbf{Z}}^\top \widehat{\mathbf{M}} \widehat{\mathbf{Z}})^{-1} \mathbf{Q}_\mathbf{Z}^\top - (\widetilde{\mathbf{Z}}^\top \widetilde{\mathbf{M}} \widetilde{\mathbf{Z}})^{-1} \right] \left(\widetilde{\mathbf{M}} \widetilde{\mathbf{Z}} \right)^\top \mathbf{Y} \right\| \le \frac{C \left\| \mathbf{Q_Z} \widehat{\mathbf{Z}}^\top \widehat{\mathbf{M}} \widehat{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}}^\top \widetilde{\mathbf{M}} \widetilde{\mathbf{Z}} \right\|}{n}. \tag{119}$$

Adding and subtracting appropriate quantities and applying the triangle inequality,

$$\left\|\mathbf{Q_Z}\widehat{\mathbf{Z}}^\top\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}}^\top\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right\| \le 2\left\|(\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}})^\top\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right\| + \left\|\mathbf{Q_Z}\widehat{\mathbf{Z}}^\top(\widehat{\mathbf{M}} - \widetilde{\mathbf{M}})\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top\right\| + \left\|(\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}})^\top\widetilde{\mathbf{M}}(\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}})\right\|. \tag{120}$$

Decomposing $\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}}$ as in Equation (117),

$$\left\|(\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}})^\top\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right\| \le \left\|(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X})^\top\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right\| + \left|\widetilde{\mathbf{Z}}^\top\widetilde{\mathbf{M}}(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}})\mathbf{Y}\right|. \tag{121}$$

Applying Lemma 15 with $\mathbf{B} = \widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}$ and using the growth rates in Equations (20), (28) and (21),

$$\left\|(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X})^\top\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right\| = o_p\left(\|\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\|\right) = o_p\left(\sqrt{n}\right),$$

where the second equality follows from the fact that $\widetilde{\mathbf{M}}$ is a projection and the growth assumption in Equation (29). Using similar growth assumptions, this time with Lemma 21 and using Lemma 23 to ensure $\|\mathbf{Y}\| = O_p\left(\sqrt{n}\right)$,

$$\left|\widetilde{\mathbf{Z}}^\top\widetilde{\mathbf{M}}(\widehat{\mathbf{G}} - \widetilde{\mathbf{G}})\mathbf{Y}\right| = o_p\left(\frac{\|\mathbf{Y}\|\|\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\|}{\sqrt{n}}\right) = o_p\left(\sqrt{n}\right).$$

Applying the above two displays to Equation (121),

$$\left\|(\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}})^\top\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right\| = o_p\left(\sqrt{n}\right). \tag{122}$$

By submultiplicativity,

$$\left\|\mathbf{Q_Z}\widehat{\mathbf{Z}}^\top(\widehat{\mathbf{M}} - \widetilde{\mathbf{M}})\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top\right\| \le \left\|\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top\right\|^2\left\|\widehat{\mathbf{M}} - \widetilde{\mathbf{M}}\right\| \le C\left\|\widetilde{\mathbf{Z}}\right\|^2\left\|\widehat{\mathbf{M}} - \widetilde{\mathbf{M}}\right\|,$$

where the second inequality follows from Lemma 39, Weyl's inequality, and the fact that $\sigma_{\min}(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}) \le \sigma_{\min}(\widetilde{\mathbf{Z}})$, since $\widetilde{\mathbf{M}}$ is a projection. Applying Lemma 34 and our growth assumptions in Equations (30) and (29),

$$\left\|\mathbf{Q_Z}\widehat{\mathbf{Z}}^\top(\widehat{\mathbf{M}} - \widetilde{\mathbf{M}})\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top\right\| = o_p\left(\sqrt{n}\right). \tag{123}$$

Using idempotence of $\widetilde{\mathbf{M}}$,

$$\left\|(\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}})^\top\widetilde{\mathbf{M}}(\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}})\right\| = \left\|\widetilde{\mathbf{M}}(\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}})\right\|^2. \tag{124}$$

Recalling the definition of $\widetilde{\mathbf{M}}$ from Equation (102) and applying submultiplicativity,

$$\left\|\widetilde{\mathbf{M}}(\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}})\right\| \le \frac{\left\|\widetilde{\mathbf{H}}^\top(\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}})\right\|}{\sigma_{\min}(\widetilde{\mathbf{H}})}.$$

By an argument analogous to those controlling Equations (117) and (121) above,

$$\left\|\widetilde{\mathbf{M}}(\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}})\right\| = o_p\left(\frac{\|\widetilde{\mathbf{H}}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})}\right).$$

Applying this to Equation (124),

$$\left\|(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^{\top}-\widetilde{\mathbf{Z}})^{\top}\widetilde{\mathbf{M}}(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^{\top}-\widetilde{\mathbf{Z}})\right\|=o_p\big(\kappa^2(\widetilde{\mathbf{H}})\big),$$

and our growth assumption in Equation (30) yields

$$\left\|(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^{\top}-\widetilde{\mathbf{Z}})^{\top}\widetilde{\mathbf{M}}(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^{\top}-\widetilde{\mathbf{Z}})\right\|=o_p\big(\sqrt{n}\big). \tag{125}$$

Applying Equations (122), (123) and (125) to Equation (120),

$$\left\|\mathbf{Q}_{\mathbf{Z}}\widehat{\mathbf{Z}}^{\top}\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^{\top}-\widetilde{\mathbf{Z}}^{\top}\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right\|=o_p\big(\sqrt{n}\big).$$

Applying this to Equation (119),

$$\left\|\left[\mathbf{Q}_{\mathbf{Z}}(\widehat{\mathbf{Z}}^{\top}\widehat{\mathbf{M}}\widehat{\mathbf{Z}})^{-1}\mathbf{Q}_{\mathbf{Z}}^{\top}-(\widetilde{\mathbf{Z}}^{\top}\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})^{-1}\right]\left(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right)^{\top}\mathbf{Y}\right\|=o_p\left(\frac{1}{\sqrt{n}}\right). \tag{126}$$

Applying this and Equation (118) to Equation (112) and multiplying through by $\sqrt{n}$ completes the proof. □

Proof of Theorem 2. By Theorem 6,

$$\sqrt{n}\big(\tilde{\boldsymbol{\theta}}-\boldsymbol{\theta}\big)\to\mathcal{N}\big(\mathbf{0},\Sigma\big(\hat{\boldsymbol{\theta}}\big)\big)$$

Thus, it will suffice for us to show that

$$\sqrt{n}\big(\mathbf{Q}_{\mathbf{Z}}\hat{\boldsymbol{\theta}}-\tilde{\boldsymbol{\theta}}\big)=o_p(1).$$

This is precisely the content of Lemmas 47. □

I.2 Peer Contagion Estimators

Lemma 48. *Under the latent contagion model in Equation (6), suppose that Assumptions 1 through 4 hold. Then there exists a sequence of orthogonal matrices $\mathbf{Q}_{\mathbf{Z}}\in\mathbb{R}^{(2+p+d)\times(2+p+d)}$ such that*

$$\sqrt{n}\Big(\mathbf{Q}_{\mathbf{Z}}\hat{\boldsymbol{\beta}}-\tilde{\boldsymbol{\theta}}\Big)=o_p(1).$$

Proof. Take $\mathbf{Q}_{\mathbf{Z}}$ to be as in Equation (105). Recalling the definitions from Equations (8) and (111),

$$\mathbf{Q}_{\mathbf{Z}}\hat{\boldsymbol{\beta}}-\tilde{\boldsymbol{\theta}}=\mathbf{Q}_{\mathbf{Z}}\left(\breve{\mathbf{Z}}^{\top}\widetilde{\mathbf{M}}\breve{\mathbf{Z}}\right)^{-1}\breve{\mathbf{Z}}^{\top}\widetilde{\mathbf{M}}\mathbf{y}-\left(\widetilde{\mathbf{Z}}^{\top}\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right)^{-1}\widetilde{\mathbf{Z}}^{\top}\widetilde{\mathbf{M}}\mathbf{Y}.$$

Applying the triangle inequality,

$$\begin{aligned}\left\|\mathbf{Q}_{\mathbf{Z}}\hat{\boldsymbol{\beta}}-\tilde{\boldsymbol{\theta}}\right\|\le&\left\|\mathbf{Q}_{\mathbf{Z}}\left(\breve{\mathbf{Z}}^{\top}\widetilde{\mathbf{M}}\breve{\mathbf{Z}}\right)^{-1}\mathbf{Q}_{\mathbf{Z}}^{\top}\left(\widetilde{\mathbf{M}}\breve{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^{\top}-\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right)^{\top}\mathbf{Y}\right\|\\&+\left\|\left[\mathbf{Q}_{\mathbf{Z}}\left(\breve{\mathbf{Z}}^{\top}\widetilde{\mathbf{M}}\breve{\mathbf{Z}}\right)^{-1}\mathbf{Q}_{\mathbf{Z}}^{\top}-\left(\widetilde{\mathbf{Z}}^{\top}\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right)^{-1}\right]\left(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\right)^{\top}\mathbf{Y}\right\|.\end{aligned} \tag{127}$$

We will control each of these right-hand terms separately.

Following an argument parallel to that given leading up to Equation (118) in the proof of Lemma 47, but using Lemmas 45, 35 and 40 in place of Lemmas 44, 34 and 39, respectively, we obtain

$$\left\| \mathbf{Q_Z} (\breve{\mathbf{Z}}^\top \widetilde{\mathbf{M}} \breve{\mathbf{Z}})^{-1} \mathbf{Q}_{\mathbf{Z}}^\top \left(\widetilde{\mathbf{M}} \breve{\mathbf{Z}} \mathbf{Q}_{\mathbf{Z}}^\top - \widetilde{\mathbf{M}} \widetilde{\mathbf{Z}} \right)^\top \mathbf{Y} \right\| = o_p\left(\frac{1}{\sqrt{n}} \right).$$

Following an argument parallel to that leading to Equation (126) in the proof of Lemma 47, but using Lemmas 45, 20, 40 and 35 in place of Lemmas 44, 21, 39 and 34, respectively, yields

$$\left\| \mathbf{Q_Z} \left[(\breve{\mathbf{Z}}^\top \widetilde{\mathbf{M}} \breve{\mathbf{Z}})^{-1} \mathbf{Q}_{\mathbf{Z}}^\top - (\widetilde{\mathbf{Z}}^\top \widetilde{\mathbf{M}} \widetilde{\mathbf{Z}})^{-1} \right] \left(\widetilde{\mathbf{M}} \widetilde{\mathbf{Z}} \right)^\top \mathbf{Y} \right\| = o_p\left(\frac{1}{\sqrt{n}} \right).$$

Applying the above two displays to Equation (127) and multiplying through by $\sqrt{n}$ completes the proof. □

J Peer Contagion Data Model

Here we provide proof details for Theorems 1 and 3, which describe the behavior of our estimators when the responses $\mathbf{Y}$ are driven by contagion on the graph structure encoded by the observed network $\mathbf{A}$, i.e., what we have termed *peer contagion*, as given in Equation (4). The key challenge under the peer contagion model, as in the latent contagion model considered in Appendix I, is that we do not have access to the latent positions $\mathbf{X}$. If we had access to this latent structure, we could construct oracle analogues of the estimators in Equation (8),

$$\tilde{\beta} = (\mathbf{Z}^\top \mathbf{M} \mathbf{Z})^{-1} \mathbf{Z}^\top \mathbf{M} \mathbf{Y}, \tag{128}$$

where $\mathbf{Z}$ is given by

$$\mathbf{Z} = \left[\mathbf{1}_n \ \mathbf{W} \ \mathbf{X} \ \mathbf{GY} \right] \in \mathbb{R}^{n \times (p+d+2)} \quad \text{and} \quad \mathbf{H} = \left[\mathbf{W} \ \mathbf{X} \ \mathbf{GW} \ \mathbf{GX} \ \mathbf{G}^2\mathbf{W} \ \mathbf{G}^2\mathbf{X} \right] \in \mathbb{R}^{n \times (3p+3d)} \tag{129}$$

and $\mathbf{M}$ is as defined in Equation (103).

We begin by proving convergence of the peer contagion estimators under the peer contagion data model to establish Theorem 1. We give analogous proofs for the latent contagion estimators in Section J.2 below.

J.1 Peer Contagion Estimators

Our proof of Theorem 1 relies on showing that when the model in Equation (3) holds, our estimators are close to those in Equation (128). Lemmas 49 establishes this.

Lemma 49. *Under the peer contagion model in Equation (4), suppose that Assumptions 1 through 3 and Assumption 5 hold. Then there exists a sequence of orthogonal matrices $\mathbf{Q_Z} \in \mathbb{R}^{(2+p+d)\times(2+p+d)}$ such that*

$$\sqrt{n}\left(\mathbf{Q_Z} \hat{\beta} - \tilde{\beta} \right) = o_p(1).$$

Proof. Take $\mathbf{Q}_{\mathbf{Z}}$ to be as in Equation (105). Recalling the definitions from Equations (8) and (128) and using the fact that $\mathbf{Q}_{\mathbf{Z}}$ is orthogonal,

$$\mathbf{Q}_{\mathbf{Z}}\hat{\beta} - \tilde{\beta} = \mathbf{Q}_{\mathbf{Z}} \left(\breve{\mathbf{Z}}^\top \widetilde{\mathbf{M}} \breve{\mathbf{Z}}\right)^{-1} \mathbf{Q}_{\mathbf{Z}}^\top \mathbf{Q}_{\mathbf{Z}} \breve{\mathbf{Z}}^\top \widetilde{\mathbf{M}} \mathbf{Y} - (\mathbf{Z}^\top \mathbf{M} \mathbf{Z})^{-1} \mathbf{Z}^\top \mathbf{M} \mathbf{Y}.$$

Applying the triangle inequality,

$$\begin{aligned} \left\| \mathbf{Q}_{\mathbf{Z}}\hat{\beta} - \tilde{\beta} \right\| \le & \left\| \mathbf{Q}_{\mathbf{Z}} \left(\breve{\mathbf{Z}}^\top \widetilde{\mathbf{M}} \breve{\mathbf{Z}}\right)^{-1} \mathbf{Q}_{\mathbf{Z}}^\top \left(\widetilde{\mathbf{M}} \breve{\mathbf{Z}} \mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{M}\mathbf{Z}\right)^\top \mathbf{Y} \right\| \\ & + \left\| \left[\mathbf{Q}_{\mathbf{Z}} \left(\breve{\mathbf{Z}}^\top \widetilde{\mathbf{M}} \breve{\mathbf{Z}}\right)^{-1} \mathbf{Q}_{\mathbf{Z}}^\top - \left(\mathbf{Z}^\top \mathbf{M} \mathbf{Z}\right)^{-1} \right] (\mathbf{M}\mathbf{Z})^\top \mathbf{Y} \right\|. \end{aligned} \tag{130}$$

We will control each of these right-hand terms separately.

Controlling the first term in Equation (130)

By submultiplicativity and unitary invariance of the norm,

$$\left\| \mathbf{Q}_{\mathbf{Z}} (\breve{\mathbf{Z}}^\top \widetilde{\mathbf{M}} \breve{\mathbf{Z}})^{-1} \mathbf{Q}_{\mathbf{Z}}^\top \left(\widetilde{\mathbf{M}} \breve{\mathbf{Z}} \mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{M}\mathbf{Z}\right)^\top \mathbf{Y} \right\| \le \frac{\left\| \left(\widetilde{\mathbf{M}} \breve{\mathbf{Z}} \mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{M}\mathbf{Z}\right)^\top \mathbf{Y} \right\|}{\sigma_{\min}^2 (\widetilde{\mathbf{M}} \breve{\mathbf{Z}})} \le \frac{C \left\| \left(\widetilde{\mathbf{M}} \breve{\mathbf{Z}} \mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{M}\mathbf{Z}\right)^\top \mathbf{Y} \right\|}{\sigma_{\min}^2 (\widetilde{\mathbf{M}} \widetilde{\mathbf{Z}})},$$

where the second inequality follows from Weyl's inequality and Lemma 45. Using our assumption in Equation (29), it follows that

$$\left\| \mathbf{Q}_{\mathbf{Z}} (\breve{\mathbf{Z}}^\top \widetilde{\mathbf{M}} \breve{\mathbf{Z}})^{-1} \mathbf{Q}_{\mathbf{Z}}^\top \left(\widetilde{\mathbf{M}} \breve{\mathbf{Z}} \mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{M}\mathbf{Z}\right)^\top \mathbf{Y} \right\| \le \frac{C \left\| \left(\widetilde{\mathbf{M}} \breve{\mathbf{Z}} \mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{M}\mathbf{Z}\right)^\top \mathbf{Y} \right\|}{n}. \tag{131}$$

By the triangle inequality,

$$\left\| \left(\widetilde{\mathbf{M}} \breve{\mathbf{Z}} \mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{M}\mathbf{Z}\right)^\top \mathbf{Y} \right\| \le \left\| \mathbf{Q}_{\mathbf{Z}} \breve{\mathbf{Z}}^\top \left(\widetilde{\mathbf{M}} - \mathbf{M}\right) \mathbf{Y} \right\| + \left\| \left(\breve{\mathbf{Z}} \mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}\right)^\top \mathbf{M}\mathbf{Y} \right\|. \tag{132}$$

Again using the triangle inequality,

$$\left\| \mathbf{Q}_{\mathbf{Z}} \breve{\mathbf{Z}}^\top \left(\widetilde{\mathbf{M}} - \mathbf{M}\right) \mathbf{Y} \right\| \le \left\| \left(\breve{\mathbf{Z}} \mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}\right)^\top \left(\widetilde{\mathbf{M}} - \mathbf{M}\right) \mathbf{Y} \right\| + \left\| \mathbf{Z}^\top \left(\widetilde{\mathbf{M}} - \mathbf{M}\right) \mathbf{Y} \right\|. \tag{133}$$

Using submultiplicativity followed by Lemmas 41 and 36,

$$\left\| \left(\breve{\mathbf{Z}} \mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}\right)^\top : \left(\widetilde{\mathbf{M}} - \mathbf{M}\right) \mathbf{Y} \right\| \le \left\| \breve{\mathbf{Z}} \mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z} \right\| \left\| \widetilde{\mathbf{M}} - \mathbf{M} \right\| \|\mathbf{Y}\| = o_p\left(\frac{\sigma_{\min}(\widetilde{\mathbf{M}} \widetilde{\mathbf{Z}}) \|\mathbf{Y}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})} \right).$$

Using the fact that $\widetilde{\mathbf{M}}$ is a projection and applying our growth assumption in Equation (29), we have $\sigma_{\min}(\widetilde{\mathbf{M}} \widetilde{\mathbf{Z}}) \le \|\widetilde{\mathbf{Z}}\| = O_p(\sqrt{n})$, and thus

$$\left\| \left(\breve{\mathbf{Z}} \mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}\right)^\top \left(\widetilde{\mathbf{M}} - \mathbf{M}\right) \mathbf{Y} \right\| = o_p\left(\frac{\|\mathbf{Y}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})} \right) = o_p(\sqrt{n}), \tag{134}$$

where the second equality follows from our growth assumption in Equation (30) and using Lemma 24 to ensure $\|\mathbf{Y}\| = O_p(\sqrt{n})$.

Using Lemma 24 again, write $\mathbf{Y} = \widetilde{\mathbf{Y}} + \zeta$, where $\widetilde{\mathbf{Y}}$ is independent of $\mathbf{A} - \mathbf{P}$ conditional on $\mathbf{X}$ with

$$\left\|\widetilde{\mathbf{Y}}\right\| = O_p(\sqrt{n}) \text{ and } \|\zeta\| = o_p(\sqrt{n}). \tag{135}$$

Applying the triangle inequality,

$$\left\|\mathbf{Z}^\top\left(\widetilde{\mathbf{M}} - \mathbf{M}\right)\mathbf{Y}\right\| \leq \left\|\mathbf{Z}^\top\left(\widetilde{\mathbf{M}} - \mathbf{M}\right)\widetilde{\mathbf{Y}}\right\| + \left\|\mathbf{Z}^\top\left(\widetilde{\mathbf{M}} - \mathbf{M}\right)\zeta\right\|. \tag{136}$$

By submultiplicativity, Lemma 36 and the assumptions in Equations (30) and (29),

$$\left\|\mathbf{Z}^\top\left(\widetilde{\mathbf{M}} - \mathbf{M}\right)\widetilde{\mathbf{Y}}\right\| = o_p\left(\frac{\|\widetilde{\mathbf{Z}}\|\|\widetilde{\mathbf{Y}}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})}\right) = o_p(\|\widetilde{\mathbf{Y}}\|) = o_p(\sqrt{n}),$$

where the last equality follows from $\|\widetilde{\mathbf{Y}}\| = O_p(\sqrt{n})$. Applying submultiplicativity, Lemma 36 and the fact that $\zeta = o_p(\sqrt{n})$,

$$\left\|\mathbf{Z}^\top\left(\widetilde{\mathbf{M}} - \mathbf{M}\right)\zeta\right\| = o_p\left(\frac{\|\mathbf{Z}\|\sqrt{n}}{\sigma_{\min}(\widetilde{\mathbf{H}})}\right) = o_p(\sqrt{n}),$$

where the second equality follows from our growth assumptions in Equations (30) and (29). Applying the above two display equations to Equation (136),

$$\left\|\mathbf{Z}^\top\left(\widetilde{\mathbf{M}} - \mathbf{M}\right)\mathbf{Y}\right\| = o_p(\sqrt{n}).$$

Applying this and Equation (134) to Equation (133),

$$\left\|\mathbf{Q}_{\mathbf{Z}}\breve{\mathbf{Z}}^\top\left(\widetilde{\mathbf{M}} - \mathbf{M}\right)\mathbf{Y}\right\| = o_p(\sqrt{n}).$$

Finally, applying this to Equation (132),

$$\left\|\left(\widetilde{\mathbf{M}}\breve{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{M}\mathbf{Z}\right)^\top\mathbf{Y}\right\| \leq \left\|\left(\breve{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}\right)^\top\mathbf{M}\mathbf{Y}\right\| + o_p(\sqrt{n}). \tag{137}$$

Recalling the structure of $\mathbf{Z}$ from Equation (129) and applying the triangle inequality,

$$\left\|\left(\breve{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}\right)^\top\mathbf{M}\mathbf{Y}\right\| \leq \|\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)^\top\mathbf{M}\mathbf{Y}\| + |\mathbf{Y}^\top\mathbf{M}\left(\mathbf{G} - \mathbf{G}\right)\mathbf{Y}| = \|\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)^\top\mathbf{M}\mathbf{Y}\|.$$

Applying the triangle inequality and recalling $\widetilde{\mathbf{M}}$ from Equation (102),

$$\left\|\left(\breve{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}\right)^\top\mathbf{M}\mathbf{Y}\right\| \leq \|\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)^\top\left(\mathbf{M} - \widetilde{\mathbf{M}}\right)\mathbf{Y}\| + \|\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)^\top\widetilde{\mathbf{M}}\mathbf{Y}\|. \tag{138}$$

By submultiplicativity followed by Lemmas 15 and 37,

$$\begin{aligned}\left\|\left(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right)^\top\left(\mathbf{M} - \widetilde{\mathbf{M}}\right)\mathbf{Y}\right\| &\leq \left\|\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X}\right\|\left\|\mathbf{M} - \widetilde{\mathbf{M}}\right\|\|\mathbf{Y}\| \\ &= o_p\left(\frac{\|\mathbf{Y}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})}\left[\frac{\sqrt{\nu + b^2}\sqrt{n}\log n}{\sqrt{\mathbf{s}_d}} + \frac{\kappa(\mathbf{P})(\nu + b^2)n^{3/2}\log^2 n}{\mathbf{s}_d^{3/2}}\right]\right) \\ &= o_p(\sqrt{n}),\end{aligned} \tag{139}$$

where we have used Lemma 24 to ensure that $\|\mathbf{Y}\| = O_p(\sqrt{n})$ and the growth rates in Equations (20), (28), (21) and (30). Applying this to Equation (138),

$$\left\|(\breve{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z})^\top \mathbf{M}\mathbf{Y}\right\| \le \|(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X})^\top \widetilde{\mathbf{M}}\mathbf{Y}\| + o_p(\sqrt{n}). \tag{140}$$

Recalling $\widetilde{\mathbf{M}} = \widetilde{\mathbf{H}}\left(\widetilde{\mathbf{H}}^\top\widetilde{\mathbf{H}}\right)^{-1}\widetilde{\mathbf{H}}^\top$ and using submultiplicativity,

$$\|(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X})^\top \widetilde{\mathbf{M}}\mathbf{Y}\| \le \frac{\left\|(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X})^\top \widetilde{\mathbf{H}}\right\| \|\widetilde{\mathbf{H}}\| \|\mathbf{Y}\|}{\sigma_{\min}^2(\widetilde{\mathbf{H}})} \le \frac{\kappa(\widetilde{\mathbf{H}}) \left\|(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X})^\top \widetilde{\mathbf{H}}\right\| \|\mathbf{Y}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})}.$$

Applying Lemma 15 with $\mathbf{B} = \widetilde{\mathbf{H}}$ and using Lemma 24 to ensure that $\|\mathbf{Y}\| = O_p(\sqrt{n})$,

$$\|(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X})^\top \widetilde{\mathbf{M}}\mathbf{Y}\| \le \frac{\kappa(\widetilde{\mathbf{H}})\sqrt{n}\|\widetilde{\mathbf{H}}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})}\left[\frac{\sqrt{\nu + b^2}\log n}{\sqrt{\mathbf{s}_d}} + \frac{\kappa(\mathbf{P})(\nu + b^2) n \log^2 n}{\mathbf{s}_d^{3/2}}\right].$$

Applying our growth rates in Equations (20), (28), (21) and (30),

$$\|(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X})^\top \widetilde{\mathbf{M}}\mathbf{Y}\| = o_p(\sqrt{n}).$$

Applying this to Equation (140),

$$\left\|(\breve{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z})^\top \mathbf{M}\mathbf{Y}\right\| = o_p(\sqrt{n}).$$

Applying this to Equation (137) in turn,

$$\left\|\left(\widetilde{\mathbf{M}}\breve{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{M}\mathbf{Z}\right)^\top \mathbf{Y}\right\| = o_p(\sqrt{n}).$$

Finally, applying this bound to Equation (131),

$$\left\|\mathbf{Q}_{\mathbf{Z}}(\breve{\mathbf{Z}}^\top\widetilde{\mathbf{M}}\breve{\mathbf{Z}})^{-1}\mathbf{Q}_{\mathbf{Z}}^\top \left(\widetilde{\mathbf{M}}\breve{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{M}\mathbf{Z}\right)^\top \mathbf{Y}\right\| = o_p\left(\frac{1}{\sqrt{n}}\right). \tag{141}$$

Controlling the second term in Equation (130)

By Lemmas 45 and 46 along with our growth assumption in Equation (29), both $\breve{\mathbf{Z}}^\top\widetilde{\mathbf{M}}\breve{\mathbf{Z}}$ and $\mathbf{Z}^\top\mathbf{M}\mathbf{Z}$ are invertible with high probability for all n suitably large. Thus, factoring appropriately, applying submultiplicativity and recalling that $\widetilde{\mathbf{M}}$ and $\mathbf{M}$ are projections,

$$\left\|\left[\mathbf{Q}_{\mathbf{Z}}\left(\breve{\mathbf{Z}}^\top\widetilde{\mathbf{M}}\breve{\mathbf{Z}}\right)^{-1}\mathbf{Q}_{\mathbf{Z}}^\top - (\mathbf{Z}^\top\mathbf{M}\mathbf{Z})^{-1}\right](\mathbf{M}\mathbf{Z})^\top \mathbf{Y}\right\| \le \frac{\left\|\mathbf{Q}_{\mathbf{Z}}\breve{\mathbf{Z}}^\top\widetilde{\mathbf{M}}\breve{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}^\top\mathbf{M}\mathbf{Z}\right\| \|\mathbf{Z}\| \|\mathbf{Y}\|}{\sigma_{\min}^2(\widetilde{\mathbf{M}}\breve{\mathbf{Z}})\sigma_{\min}^2(\mathbf{M}\mathbf{Z})}.$$

Using Lemmas 45 and 46 again along with Weyl's inequality, we can ensure that

$$\sigma_{\min}(\widetilde{\mathbf{M}}\breve{\mathbf{Z}})\sigma_{\min}(\mathbf{M}\mathbf{Z}) = \Omega_p\left(\sigma_{\min}^2(\widetilde{\mathbf{M}}\breve{\mathbf{Z}})\right),$$

so that our growth assumption in Equation (29), along with Lemma 24 yield

$$\left\|\left[\mathbf{Q}_{\mathbf{Z}}(\breve{\mathbf{Z}}^\top\widetilde{\mathbf{M}}\breve{\mathbf{Z}})^{-1}\mathbf{Q}_{\mathbf{Z}}^\top - (\mathbf{Z}^\top\mathbf{M}\mathbf{Z})^{-1}\right](\mathbf{M}\mathbf{Z})^\top \mathbf{Y}\right\| \le \frac{C\left\|\mathbf{Q}_{\mathbf{Z}}\breve{\mathbf{Z}}^\top\widetilde{\mathbf{M}}\breve{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}^\top\mathbf{M}\mathbf{Z}\right\| \|\mathbf{Z}\|}{n^{3/2}}.$$

Applying Lemma 42, Weyl's inequality and our growth assumption in Equation (29),

$$\left\| \left[\mathbf{Q}_\mathbf{Z} (\breve{\mathbf{Z}}^\top \widetilde{\mathbf{M}} \breve{\mathbf{Z}})^{-1} \mathbf{Q}_\mathbf{Z}^\top - (\mathbf{Z}^\top \mathbf{M} \mathbf{Z})^{-1} \right] (\mathbf{M}\mathbf{Z})^\top \mathbf{Y} \right\| \le \frac{C \left\| \mathbf{Q}_\mathbf{Z} \breve{\mathbf{Z}}^\top \widetilde{\mathbf{M}} \breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z}^\top \mathbf{M} \mathbf{Z} \right\|}{n}. \tag{142}$$

Adding and subtracting appropriate quantities and applying the triangle inequality,

$$\left\| \mathbf{Q}_\mathbf{Z} \breve{\mathbf{Z}}^\top \widetilde{\mathbf{M}} \breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z}^\top \mathbf{M} \mathbf{Z} \right\| \le 2 \left\| (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z})^\top \mathbf{M} \mathbf{Z} \right\| + \left\| \mathbf{Q}_\mathbf{Z} \breve{\mathbf{Z}}^\top (\widetilde{\mathbf{M}} - \mathbf{M}) \breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top \right\| + \left\| (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z})^\top \mathbf{M} (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z}) \right\|. \tag{143}$$

Decomposing $\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z}$ as in Equation (138),

$$\left\| (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z})^\top \mathbf{M} \mathbf{Z} \right\| \le \left\| (\widehat{\mathbf{X}} \mathbf{Q}^\top - \mathbf{X})^\top \mathbf{M} \mathbf{Z} \right\| + \left| \mathbf{Z}^\top \mathbf{M} (\mathbf{G} - \mathbf{G}) \mathbf{Y} \right| = \left\| (\widehat{\mathbf{X}} \mathbf{Q}^\top - \mathbf{X})^\top \mathbf{M} \mathbf{Z} \right\|.$$

Applying the triangle inequality,

$$\left\| (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z})^\top \mathbf{M} \mathbf{Z} \right\| \le \left\| (\widehat{\mathbf{X}} \mathbf{Q}^\top - \mathbf{X})^\top \left(\mathbf{M} - \widetilde{\mathbf{M}} \right) \mathbf{Z} \right\| + \left\| (\widehat{\mathbf{X}} \mathbf{Q}^\top - \mathbf{X})^\top \widetilde{\mathbf{M}} \mathbf{Z} \right\|. \tag{144}$$

By submultiplicativity followed by Lemmas 15 and 37,

$$\left\| (\widehat{\mathbf{X}} \mathbf{Q}^\top - \mathbf{X})^\top \left(\mathbf{M} - \widetilde{\mathbf{M}} \right) \mathbf{Z} \right\| \le o_p\left(\frac{\sqrt{n} \|\mathbf{Z}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})} \frac{\sqrt{\nu + b^2} \log n}{\sqrt{\mathbf{s}_d}} \right) + o_p\left(\frac{\|\mathbf{Z}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})} \frac{\kappa(\mathbf{P})(\nu + b^2) n \log^2 n}{\mathbf{s}_d^{3/2}} \right),$$

and the growth rates in Equations (20), (28) and (21) yield

$$\left\| (\widehat{\mathbf{X}} \mathbf{Q}^\top - \mathbf{X})^\top \left(\mathbf{M} - \widetilde{\mathbf{M}} \right) \mathbf{Z} \right\| = o_p\left(\frac{\sqrt{n} \|\mathbf{Z}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})} \right) + o_p\left(\frac{\|\mathbf{Z}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})} \right) = o_p(\sqrt{n}),$$

where the second equality follows from Lemma 42 and our growth assumptions in Equations (29) and (30). Applying this to Equation (144),

$$\left\| (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z})^\top \mathbf{M} \mathbf{Z} \right\| \le \left\| (\widehat{\mathbf{X}} \mathbf{Q}^\top - \mathbf{X})^\top \widetilde{\mathbf{M}} \mathbf{Z} \right\| + o_p(\sqrt{n}).$$

Recalling the definition of $\widetilde{\mathbf{M}}$ from Equation (102) and applying submultiplicativity,

$$\left\| (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z})^\top \mathbf{M} \mathbf{Z} \right\| \le \frac{\left\| (\widehat{\mathbf{X}} \mathbf{Q}^\top - \mathbf{X})^\top \widetilde{\mathbf{H}} \right\| \kappa(\widetilde{\mathbf{H}}) \|\mathbf{Z}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})} + o_p(\sqrt{n}).$$

Applying Lemma 15 with $\mathbf{B} = \widetilde{\mathbf{H}}$ and using Equations (20), (28), (21) and (30),

$$\left\| (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z})^\top \mathbf{M} \mathbf{Z} \right\| \le o_p(\|\mathbf{Z}\|) + o_p(\sqrt{n}).$$

Applying Lemma 43 and our growth assumption in Equation (29),

$$\left\| (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z})^\top \mathbf{M} \mathbf{Z} \right\| = o_p(\sqrt{n}), \tag{145}$$

By submultiplicativity followed by Lemma 40, Weyl's inequality, and the fact that $\sigma_{\min}(\widetilde{\mathbf{M}} \widetilde{\mathbf{Z}}) \le \sigma_{\min}(\widetilde{\mathbf{Z}})$,

$$\left\| \mathbf{Q}_\mathbf{Z} \breve{\mathbf{Z}}^\top (\widetilde{\mathbf{M}} - \mathbf{M}) \breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top \right\| \le \left\| \breve{\mathbf{Z}} \right\|^2 \left\| \widetilde{\mathbf{M}} - \mathbf{M} \right\| \le C \left\| \widetilde{\mathbf{Z}} \right\|^2 \left\| \widetilde{\mathbf{M}} - \mathbf{M} \right\|,$$

Applying Lemma 36 and our growth assumptions in Equations (30) and (29),

$$\left\| \mathbf{Q_Z} \breve{\mathbf{Z}}^\top (\widetilde{\mathbf{M}} - \mathbf{M}) \breve{\mathbf{Z}} \right\| = o_p(\sqrt{n}). \tag{146}$$

Using idempotence of $\mathbf{M}$,

$$\left\| (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z})^\top \mathbf{M} (\breve{\mathbf{Z}} - \mathbf{Z}) \right\| = \left\| \mathbf{M} (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z}) \right\|^2. \tag{147}$$

Applying the triangle inequality and using submultiplicativity,

$$\left\| \mathbf{M} (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z}) \right\| \le \left\| \widetilde{\mathbf{M}} (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z}) \right\| + \left\| \mathbf{M} - \widetilde{\mathbf{M}} \right\| \left\| \breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z} \right\| \le \left\| \widetilde{\mathbf{M}} (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z}) \right\| + o_p\left(\frac{\|\widetilde{\mathbf{M}} \widetilde{\mathbf{Z}}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})} \right),$$

where the second inequality follows from Lemmas 41 and 34. Recalling that $\widetilde{\mathbf{M}}$ is a projection and applying our growth assumptions in Equations (29) and (30),

$$\left\| \mathbf{M} (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z}) \right\| \le \left\| \widetilde{\mathbf{M}} (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z}) \right\| + o_p(1).$$

Recalling the definition of $\widetilde{\mathbf{M}}$ from Equation (102) and applying submultiplicativity,

$$\left\| \mathbf{M} (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z}) \right\| \le \frac{\kappa(\widetilde{\mathbf{H}})}{\sigma_{\min}(\widetilde{\mathbf{H}})} \left\| \widetilde{\mathbf{H}}^\top (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z}) \right\| + o_p(1). \tag{148}$$

Recalling the structure of $\breve{\mathbf{Z}}$ and $\mathbf{Z}$ from Definition 3 and Equation (129), respectively,

$$\left\| \widetilde{\mathbf{H}}^\top (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z}) \right\| \le \left\| \widetilde{\mathbf{H}}^\top \left(\widehat{\mathbf{X}} \mathbf{Q}^\top - \mathbf{X} \right) \right\| + \left\| \widetilde{\mathbf{H}}^\top (\mathbf{G} - \mathbf{G}) \mathbf{Y} \right\| = \left\| \widetilde{\mathbf{H}}^\top \left(\widehat{\mathbf{X}} \mathbf{Q}^\top - \mathbf{X} \right) \right\|,$$

and Lemma 15 implies, using our growth assumptions in Equations (20), (28) and (21),

$$\left\| \widetilde{\mathbf{H}}^\top (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z}) \right\| = o_p(\|\widetilde{\mathbf{H}}\|).$$

Applying this to Equation (148) and applying our growth assumption in Equation (30),

$$\left\| \mathbf{M} (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z}) \right\| = o_p(1).$$

Applying this in turn to Equation (147),

$$\left\| (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z})^\top \mathbf{M} (\breve{\mathbf{Z}} \mathbf{Q}_\mathbf{Z}^\top - \mathbf{Z}) \right\| = o_p(1). \tag{149}$$

Applying Equations (145), (146) and (149) to Equation (142),

$$\left\| \left[\mathbf{Q_Z} (\breve{\mathbf{Z}}^\top \widetilde{\mathbf{M}} \breve{\mathbf{Z}})^{-1} \mathbf{Q}_\mathbf{Z}^\top - (\mathbf{Z}^\top \mathbf{M} \mathbf{Z})^{-1} \right] (\mathbf{M} \mathbf{Z})^\top \mathbf{Y} \right\| = o_p\left(\frac{1}{\sqrt{n}} \right). \tag{150}$$

Applying this and Equation (141) to Equation (130) and multiplying through by $\sqrt{n}$ completes the proof. □

J.2 Latent Contagion Estimators

We now establish corresponding results for the behavior of the latent contagion estimators when the peer contagion model is true.

Lemma 50. *Under the peer contagion model in Equation (6), suppose that Assumptions 1 through 4 hold. Then there exists a sequence of orthogonal matrices $\mathbf{Q_Z} \in \mathbb{R}^{(2+p+d)\times(2+p+d)}$ such that*

$$\sqrt{n}(\mathbf{Q_Z}\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\beta}}) = o_p(1).$$

Proof. Take $\mathbf{Q_Z}$ to be as in Equation (105). Recalling the definitions in Equations (10) and (128) and using the fact that $\mathbf{Q_Z}$ is orthogonal,

$$\mathbf{Q_Z}\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\beta}} = \left[\mathbf{Q_Z}(\widehat{\mathbf{Z}}^\top\widehat{\mathbf{M}}\widehat{\mathbf{Z}})^{-1}\mathbf{Q}_\mathbf{Z}^\top\left(\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top\right)^\top - (\mathbf{Z}^\top\mathbf{M}\mathbf{Z})^{-1}\left(\mathbf{M}\mathbf{Z}\right)^\top\right]\mathbf{Y}.$$

Simplifying and applying the triangle inequality,

$$\begin{aligned}\left\|\mathbf{Q_Z}\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\beta}}\right\| \le{}& \left\|\mathbf{Q_Z}(\widehat{\mathbf{Z}}^\top\widehat{\mathbf{M}}\widehat{\mathbf{Z}})^{-1}\mathbf{Q}_\mathbf{Z}^\top\left(\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \mathbf{M}\mathbf{Z}\right)^\top\mathbf{Y}\right\| \\ &+ \left\|\left[\mathbf{Q_Z}(\widehat{\mathbf{Z}}^\top\widehat{\mathbf{M}}\widehat{\mathbf{Z}})^{-1}\mathbf{Q}_\mathbf{Z}^\top - (\mathbf{Z}^\top\mathbf{M}\mathbf{Z})^{-1}\right]\left(\mathbf{M}\mathbf{Z}\right)^\top\mathbf{Y}\right\|.\end{aligned} \tag{151}$$

We will control each of these right-hand terms separately.

Controlling the first term in Equation (151)

By submultiplicativity of the norm,

$$\left\|\mathbf{Q_Z}(\widehat{\mathbf{Z}}^\top\widehat{\mathbf{M}}\widehat{\mathbf{Z}})^{-1}\mathbf{Q}_\mathbf{Z}^\top\left(\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \mathbf{M}\mathbf{Z}\right)^\top\mathbf{Y}\right\| \le \frac{\left\|\left(\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \mathbf{M}\mathbf{Z}\right)^\top\mathbf{Y}\right\|}{\sigma^2_{\min}(\widehat{\mathbf{M}}\widehat{\mathbf{Z}})} \le \frac{C\left\|\left(\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \mathbf{M}\mathbf{Z}\right)^\top\mathbf{Y}\right\|}{\sigma^2_{\min}(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})},$$

where the second inequality follows from Weyl's inequality and Lemma 44. Using the growth assumption in Equation (29), it follows that

$$\left\|\mathbf{Q_Z}(\widehat{\mathbf{Z}}^\top\widehat{\mathbf{M}}\widehat{\mathbf{Z}})^{-1}\mathbf{Q}_\mathbf{Z}^\top\left(\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \mathbf{M}\mathbf{Z}\right)^\top\mathbf{Y}\right\| \le \frac{C\left\|\left(\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \mathbf{M}\mathbf{Z}\right)^\top\mathbf{Y}\right\|}{n}. \tag{152}$$

By the triangle inequality,

$$\left\|\left(\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \mathbf{M}\mathbf{Z}\right)^\top\mathbf{Y}\right\| \le \left\|\mathbf{Q_Z}\widehat{\mathbf{Z}}^\top\left(\widehat{\mathbf{M}} - \mathbf{M}\right)\mathbf{Y}\right\| + \left\|\left(\widehat{\mathbf{Z}}\mathbf{Q_Z} - \mathbf{Z}\right)^\top\mathbf{M}\mathbf{Y}\right\|. \tag{153}$$

Again using the triangle inequality,

$$\left\|\mathbf{Q_Z}\widehat{\mathbf{Z}}^\top\left(\widehat{\mathbf{M}} - \mathbf{M}\right)\mathbf{Y}\right\| \le \left\|\left(\widehat{\mathbf{Z}}\mathbf{Q}_\mathbf{Z}^\top - \widetilde{\mathbf{Z}}\right)^\top\left(\widehat{\mathbf{M}} - \mathbf{M}\right)\mathbf{Y}\right\| + \left\|\widetilde{\mathbf{Z}}^\top\left(\widehat{\mathbf{M}} - \mathbf{M}\right)\mathbf{Y}\right\|. \tag{154}$$

Using Lemma 24 to write $\mathbf{Y} = \widetilde{\mathbf{Y}} + \zeta$, where $\widetilde{\mathbf{Y}}$ and ζ obey the growth rates in Equation (135) and $\widetilde{\mathbf{Y}}$ is independent of $\mathbf{A} - \mathbf{P}$ conditional on $\mathbf{X}$,

$$\left\|\widetilde{\mathbf{Z}}^\top\left(\widehat{\mathbf{M}} - \mathbf{M}\right)\mathbf{Y}\right\| \le \left\|\widetilde{\mathbf{Z}}^\top\left(\widehat{\mathbf{M}} - \mathbf{M}\right)\widetilde{\mathbf{Y}}\right\| + \left\|\widetilde{\mathbf{Z}}^\top\left(\widehat{\mathbf{M}} - \mathbf{M}\right)\zeta\right\|.$$

Controlling the first term using Lemma 38 followed by the growth assumptions in Equations (30) and (29), and using Lemma 24 to ensure that $\|\mathbf{Y}\| = O_p(\sqrt{n})$,

$$\left\|\widetilde{\mathbf{Z}}^\top \left(\widehat{\mathbf{M}} - \mathbf{M}\right) \mathbf{Y}\right\| \le \left\|\widetilde{\mathbf{Z}}^\top \left(\widehat{\mathbf{M}} - \mathbf{M}\right) \zeta\right\| + o_p\left(\frac{\|\widetilde{\mathbf{Z}}\|\|\widetilde{\mathbf{Y}}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})}\right) \le \left\|\widetilde{\mathbf{Z}}^\top \left(\widehat{\mathbf{M}} - \mathbf{M}\right) \zeta\right\| + o_p(\sqrt{n}).$$

Applying submultiplicativity to the first term followed by Lemma 38, recalling that

$$\left\|\widetilde{\mathbf{Z}}^\top \left(\widehat{\mathbf{M}} - \mathbf{M}\right) \mathbf{Y}\right\| \le o_p\left(\frac{\|\widetilde{\mathbf{Z}}\| \sqrt{n}}{\sigma_{\min}(\widetilde{\mathbf{H}})}\right) + o_p(\sqrt{n}) = o_p(\sqrt{n}),$$

where the equality follows from the growth assumptions in Equations (29) and (30).

Using submultiplicativity followed by Lemmas 43 and 38,

$$\left\|\left(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \widetilde{\mathbf{Z}}\right)^\top \left(\widehat{\mathbf{M}} - \mathbf{M}\right) \mathbf{Y}\right\| \le \left\|\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \widetilde{\mathbf{Z}}\right\| \left\|\widehat{\mathbf{M}} - \mathbf{M}\right\| \|\mathbf{Y}\| = o_p\left(\frac{\sigma_{\min}(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})\|\mathbf{Y}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})}\right) = o_p(\sqrt{n}),$$

where the last equality follows using Lemma 24 to ensure $\|\mathbf{Y}\| = O_p(\sqrt{n})$ and the growth assumption in Equation (30). Applying the above two displays to Equation (154),

$$\left\|\mathbf{Q}_{\mathbf{Z}}\widehat{\mathbf{Z}}^\top \left(\widehat{\mathbf{M}} - \mathbf{M}\right) \mathbf{Y}\right\| = o_p(\sqrt{n}).$$

Applying this to Equation (153),

$$\left\|\left(\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{M}\mathbf{Z}\right)^\top \mathbf{Y}\right\| \le \left\|\left(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}\right)^\top \mathbf{M}\mathbf{Y}\right\| + o_p(\sqrt{n}). \tag{155}$$

Applying the triangle inequality,

$$\left\|\left(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}\right)^\top \mathbf{M}\mathbf{Y}\right\| \le \left\|\left(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}\right)^\top \widetilde{\mathbf{M}}\mathbf{Y}\right\| + \left\|\left(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}\right)^\top \left(\mathbf{M} - \widetilde{\mathbf{M}}\right) \mathbf{Y}\right\|.$$

Applying submultiplicativity to the second term followed by Lemmas 43, 37 and using Lemma 24 to ensure $\|\mathbf{Y}\| = O_p(\sqrt{n})$,

$$\left\|\left(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}\right)^\top \mathbf{M}\mathbf{Y}\right\| \le \left\|\left(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}\right)^\top \widetilde{\mathbf{M}}\mathbf{Y}\right\| + o_p\left(\frac{\sqrt{n}\,\|\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})}\right) = \left\|\left(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}\right)^\top \widetilde{\mathbf{M}}\mathbf{Y}\right\| + o_p(\sqrt{n}),$$

where the second equality follows from our assumptions in Equations (29) and (30). Applying this to Equation (155),

$$\left\|\left(\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{M}\mathbf{Z}\right)^\top \mathbf{Y}\right\| \le \left\|\left(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}\right)^\top \widetilde{\mathbf{M}}\mathbf{Y}\right\| + o_p(\sqrt{n}). \tag{156}$$

Recalling the structure of $\mathbf{Z}$ from Equation (129) and applying the triangle inequality,

$$\left\|(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z})^\top\widetilde{\mathbf{M}}\mathbf{Y}\right\| \le \|(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X})^\top\widetilde{\mathbf{M}}\mathbf{Y}\| + \left|\mathbf{Y}^\top\widetilde{\mathbf{M}}(\widehat{\mathbf{G}} - \mathbf{G})\mathbf{Y}\right|. \tag{157}$$

Recalling the definition of $\widetilde{\mathbf{M}}$ from Equation (102), and applying submultiplicativity,

$$\|(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X})^\top\widetilde{\mathbf{M}}\mathbf{Y}\| \le \frac{\left\|(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X})^\top\widetilde{\mathbf{H}}\right\| \|\widetilde{\mathbf{H}}\| \|\mathbf{Y}\|}{\sigma_{\min}^2(\widetilde{\mathbf{H}})} \le C\left\|(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X})^\top\widetilde{\mathbf{H}}\right\|,$$

where we have used our growth assumption in Equation (30) and used Lemma 24 to bound $\|\mathbf{Y}\| = O_p(\sqrt{n})$. Applying Lemma 15 with $\mathbf{B} = \widetilde{\mathbf{H}}$ and using our growth assumptions in Equations (20), (28) and (21),

$$\|(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X})^\top \widetilde{\mathbf{M}}\mathbf{Y}\| = o_p(\|\widetilde{\mathbf{H}}\|) = o_p(\sqrt{n}), \tag{158}$$

where the second equality follows from our assumptions in Equation (30). Applying this to Equation (157),

$$\left\|(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z})^\top \widetilde{\mathbf{M}}\mathbf{Y}\right\| \le \left|\mathbf{Y}^\top \widetilde{\mathbf{M}}(\widehat{\mathbf{G}} - \mathbf{G})\mathbf{Y}\right| + o_p(\sqrt{n}).$$

Applying this to Equation (156),

$$\left\|\left(\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{M}\mathbf{Z}\right)^\top \mathbf{Y}\right\| \le \left|\mathbf{Y}^\top \widetilde{\mathbf{M}}(\widehat{\mathbf{G}} - \mathbf{G})\mathbf{Y}\right| + o_p(\sqrt{n}). \tag{159}$$

Recalling the definition of $\widetilde{\mathbf{M}}$ from Equation (102) and applying submultiplicativity,

$$\left|\mathbf{Y}^\top \widetilde{\mathbf{M}}(\widehat{\mathbf{G}} - \mathbf{G})\mathbf{Y}\right| \le \frac{\|\mathbf{Y}\|\kappa(\widetilde{\mathbf{H}})}{\sigma_{\min}(\widetilde{\mathbf{H}})} \left\|\widetilde{\mathbf{H}}^\top (\widehat{\mathbf{G}} - \mathbf{G})\mathbf{Y}\right\| \le C\left\|\widetilde{\mathbf{H}}^\top (\widehat{\mathbf{G}} - \mathbf{G})\mathbf{Y}\right\|,$$

where the second inequality follows from our assumptions in Equation (30) and using Lemma 24 to ensure $\|\mathbf{Y}\| = O_p(\sqrt{n})$. Applying Lemma 25 with $\mathbf{u} = \widetilde{\mathbf{H}}$ and using our growth assumption in Equation (30),

$$\left|\mathbf{Y}^\top \widetilde{\mathbf{M}}(\widehat{\mathbf{G}} - \mathbf{G})\mathbf{Y}\right| = o_p(\|\widetilde{\mathbf{H}}\|) = o_p(\sqrt{n}).$$

Applying this to Equation (159) and applying the resulting bound to Equation (152),

$$\left\|\mathbf{Q}_{\mathbf{Z}}(\widehat{\mathbf{Z}}^\top \widehat{\mathbf{M}}\widehat{\mathbf{Z}})^{-1}\mathbf{Q}_{\mathbf{Z}}^\top \left(\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{M}\mathbf{Z}\right)^\top \mathbf{Y}\right\| = o_p\left(\frac{1}{\sqrt{n}}\right). \tag{160}$$

Controlling the second term in Equation (151)

By Lemmas 44 and 46 and our growth assumption in Equation (29), both $\widehat{\mathbf{Z}}^\top \widehat{\mathbf{M}}\widehat{\mathbf{Z}}$ and $\mathbf{Z}^\top \mathbf{M}\mathbf{Z}$ are invertible with high probability for all n suitably large. Thus, factoring appropriately, applying submultiplicativity and recalling that $\widehat{\mathbf{M}}$ and $\mathbf{M}$ are projections,

$$\left\|\left[\mathbf{Q}_{\mathbf{Z}}(\widehat{\mathbf{Z}}^\top \widehat{\mathbf{M}}\widehat{\mathbf{Z}})^{-1}\mathbf{Q}_{\mathbf{Z}}^\top - (\mathbf{Z}^\top \mathbf{M}\mathbf{Z})^{-1}\right](\mathbf{M}\mathbf{Z})^\top \mathbf{Y}\right\| \le \frac{\left\|\mathbf{Q}_{\mathbf{Z}}\widehat{\mathbf{Z}}^\top \widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}^\top \mathbf{M}\mathbf{Z}\right\| \|\mathbf{Z}\| \|\mathbf{Y}\|}{\sigma_{\min}^2(\widehat{\mathbf{M}}\widehat{\mathbf{Z}})\sigma_{\min}^2(\mathbf{M}\mathbf{Z})}.$$

Using Lemmas 44 and 46 again along with Weyl's inequality, we can ensure that

$$\sigma_{\min}(\widehat{\mathbf{M}}\widehat{\mathbf{Z}})\sigma_{\min}(\mathbf{M}\mathbf{Z}) = \Omega_p\left(\sigma_{\min}^2(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}})\right).$$

Our growth assumptions in Equations (29) and (29), as well as Lemma 24 then yield

$$\left\|\left[\mathbf{Q}_{\mathbf{Z}}(\widehat{\mathbf{Z}}^\top \widehat{\mathbf{M}}\widehat{\mathbf{Z}})^{-1}\mathbf{Q}_{\mathbf{Z}}^\top - (\mathbf{Z}^\top \mathbf{M}\mathbf{Z})^{-1}\right](\mathbf{M}\mathbf{Z})^\top \mathbf{Y}\right\| \le \frac{C\left\|\mathbf{Q}_{\mathbf{Z}}\widehat{\mathbf{Z}}^\top \widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}^\top \mathbf{M}\mathbf{Z}\right\|}{n}. \tag{161}$$

Adding and subtracting appropriate quantities and applying the triangle inequality,

$$\left\|\mathbf{Q}_{\mathbf{Z}}\widehat{\mathbf{Z}}^\top\widehat{\mathbf{M}}\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}^\top\mathbf{M}\mathbf{Z}\right\| \le 2\left\|(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z})^\top\mathbf{M}\mathbf{Z}\right\| + \left\|\widehat{\mathbf{Z}}^\top(\widehat{\mathbf{M}} - \mathbf{M})\widehat{\mathbf{Z}}\right\| + \left\|(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z})^\top\mathbf{M}(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z})\right\|. \tag{162}$$

Applying the triangle inequality followed by submultiplicativity,

$$\left\|(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z})^\top\mathbf{M}\mathbf{Z}\right\| \le \left\|(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z})^\top\widetilde{\mathbf{M}}\mathbf{Z}\right\| + \left\|\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z}\right\|\left\|\mathbf{M} - \widetilde{\mathbf{M}}\right\|\|\mathbf{Z}\| = \left\|(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z})^\top\widetilde{\mathbf{M}}\mathbf{Z}\right\| + o_p\left(\frac{\|\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}\|\,\|\mathbf{Z}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})}\right).$$

Controlling $\|\mathbf{Z}\|$ with Lemma 42 and Weyl's inequality and using using our growth assumptions in Equations (29) and (30),

$$\left\|(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z})^\top\mathbf{M}\mathbf{Z}\right\| \le \left\|(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z})^\top\widetilde{\mathbf{M}}\mathbf{Z}\right\| + o_p(\sqrt{n}).$$

Decomposing $\widehat{\mathbf{Z}} - \mathbf{Z}$ as in Equation (157),

$$\left\|(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z})^\top\mathbf{M}\mathbf{Z}\right\| \le \left\|(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X})^\top\widetilde{\mathbf{M}}\mathbf{Z}\right\| + \left\|\mathbf{Z}^\top\widetilde{\mathbf{M}}(\widehat{\mathbf{G}} - \mathbf{G})\mathbf{Y}\right\| + o_p(\sqrt{n}). \tag{163}$$

Recalling the definition of $\widetilde{\mathbf{M}}$ from Equation (102) and using submultiplicativity,

$$\left\|(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X})^\top\widetilde{\mathbf{M}}\mathbf{Z}\right\| \le \frac{\kappa(\widetilde{\mathbf{H}})\left\|(\widehat{\mathbf{X}} - \mathbf{X})^\top\widetilde{\mathbf{H}}\right\|\|\mathbf{Z}\|}{\sigma_{\min}(\widetilde{\mathbf{H}})} = o_p(\|\mathbf{Z}\|),$$

where we have used Lemma 15 with $\mathbf{B} = \widetilde{\mathbf{H}}$ along with the growth rates in Equation (20) (28), (21), and (30). Applying Lemma 42, Weyl's inequality and our growth assumption in Equation (29),

$$\left\|(\widehat{\mathbf{X}}\mathbf{Q}^\top - \mathbf{X})^\top\widetilde{\mathbf{M}}\mathbf{Z}\right\| = o_p(\sqrt{n}).$$

Applying this to Equation (163),

$$\left\|(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z})^\top\mathbf{M}\mathbf{Z}\right\| \le \left\|\mathbf{Z}^\top\widetilde{\mathbf{M}}(\widehat{\mathbf{G}} - \mathbf{G})\mathbf{Y}\right\| + o_p(\sqrt{n}).$$

Again using the definition of $\widetilde{\mathbf{M}}$ and submultiplicativity,

$$\left\|(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z})^\top\mathbf{M}\mathbf{Z}\right\| \le \frac{\|\mathbf{Z}\|\,\kappa(\widetilde{\mathbf{H}})}{\sigma_{\min}(\widetilde{\mathbf{H}})}\left\|\widetilde{\mathbf{H}}^\top(\widehat{\mathbf{G}} - \mathbf{G})\mathbf{Y}\right\| + o_p(\sqrt{n}) \le C\left\|\widetilde{\mathbf{H}}^\top(\widehat{\mathbf{G}} - \mathbf{G})\mathbf{Y}\right\| + o_p(\sqrt{n}),$$

Applying Lemma 25 with $\mathbf{u}$ equal to each of the columns of $\widetilde{\mathbf{H}}$ and using our growth assumption in Equation (30),

$$\left\|(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top - \mathbf{Z})^\top\mathbf{M}\mathbf{Z}\right\| = o_p(\sqrt{n}). \tag{164}$$

By submultiplicativity,

$$\left\|\mathbf{Q}_{\mathbf{Z}}\widehat{\mathbf{Z}}^\top(\widehat{\mathbf{M}} - \mathbf{M})\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top\right\| \le \left\|\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top\right\|^2\left\|\widehat{\mathbf{M}} - \mathbf{M}\right\| \le C\left\|\widehat{\mathbf{Z}}\right\|^2\left\|\widehat{\mathbf{M}} - \mathbf{M}\right\|,$$

where the second inequality follows from Lemma 39, Weyl's inequality, and the fact that $\sigma_{\min}(\widetilde{\mathbf{M}}\widetilde{\mathbf{Z}}) \leq \sigma_{\min}(\widetilde{\mathbf{Z}})$, since $\widetilde{\mathbf{M}}$ is a projection. Applying Lemma 38 and our growth assumptions in Equations (30) and (29),

$$\left\|\mathbf{Q}_{\mathbf{Z}}\widehat{\mathbf{Z}}^\top(\widehat{\mathbf{M}}-\mathbf{M})\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top\right\| = o_p(\sqrt{n}). \tag{165}$$

Using idempotence of $\mathbf{M}$,

$$\left\|(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top-\mathbf{Z})^\top\mathbf{M}(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top-\mathbf{Z})\right\| = \left\|\mathbf{M}(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top-\mathbf{Z})\right\|^2. \tag{166}$$

Applying the triangle inequality and submultiplicativity,

$$\left\|\mathbf{M}(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top-\mathbf{Z})\right\| \leq \left\|\widetilde{\mathbf{M}}\left(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top-\mathbf{Z}\right)\right\| + \left\|\mathbf{M}-\widetilde{\mathbf{M}}\right\|\left\|\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top-\mathbf{Z}\right\| \leq \left\|\widetilde{\mathbf{M}}\left(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top-\mathbf{Z}\right)\right\| + o_p(1),$$

where the second inequality follows from Lemmas 37 and 43 along with our growth assumptions in Equations (30) and (29). Recalling the definition of $\widetilde{\mathbf{M}}$ from Equation (102) and applying submultiplicativity,

$$\left\|\mathbf{M}(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top-\mathbf{Z})\right\| \leq \frac{\kappa(\widetilde{\mathbf{H}})\left\|\widetilde{\mathbf{H}}^\top(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top-\mathbf{Z})\right\|}{\sigma_{\min}(\mathbf{H})} + o_p(1).$$

An argument parallel to that leading up to Equation (164) yields

$$\left\|\widetilde{\mathbf{H}}^\top(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top-\mathbf{Z})\right\| = o_p(\sqrt{n}),$$

from which our growth assumption in Equation (30) yield

$$\left\|\mathbf{M}(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top-\mathbf{Z})\right\| = o_p(1).$$

Applying this to Equation (166),

$$\left\|(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top-\mathbf{Z})^\top\mathbf{M}(\widehat{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top-\mathbf{Z})\right\| = o_p(1). \tag{167}$$

Applying Equations (164), (165) and (167) to Equation (162),

$$\left\|\mathbf{Q}_{\mathbf{Z}}\breve{\mathbf{Z}}^\top\widetilde{\mathbf{M}}\breve{\mathbf{Z}}\mathbf{Q}_{\mathbf{Z}}^\top-\mathbf{Z}^\top\mathbf{M}\mathbf{Z}\right\| = o_p(\sqrt{n}).$$

Applying this to Equation (161),

$$\left\|\left[\mathbf{Q}_{\mathbf{Z}}(\widehat{\mathbf{Z}}^\top\widehat{\mathbf{M}}\widehat{\mathbf{Z}})^{-1}\mathbf{Q}_{\mathbf{Z}}^\top-(\mathbf{Z}^\top\mathbf{M}\mathbf{Z})^{-1}\right](\mathbf{M}\mathbf{Z})^\top\mathbf{Y}\right\| = o_p\left(\frac{1}{\sqrt{n}}\right). \tag{168}$$

Applying this and Equation (160) to Equation (151) and multiplying through by $\sqrt{n}$ completes the proof. □

K Projection Equivalence

Here, we prove Theorem 5, which concerns asymptotic equivalence of the projection parameters defined in Equation (14). We handle the two different model settings of Theorem 5 in two different lemmas below. Recall that $\mathbb{E}_\beta$ indicates that an expectation is with respect to a true underlying peer contagion process (3) and $\mathbb{E}_\theta$ indicates that the expectation is with respect to a true underlying latent contagion process (5). We note that our results on (approximate) projection equivalence do not require the same assumptions as our first four Theorems, but holds under slightly weaker assumptions. For example, instead of the subgamma edge behavior of Definition 1, we require only that the edge noise has bounded second moment.

K.1 Projection equivalence under latent contagion

Here we prove projection equivalence when the true data generating model is latent contagion, as described in Equation (6). For ease of notation, define the sample covariance matrices

$$\widetilde{\boldsymbol{\Sigma}} = \frac{1}{n}\widetilde{\mathbf{Z}}^\top\widetilde{\mathbf{Z}} \text{ and } \boldsymbol{\Sigma} = \frac{1}{n}\mathbf{Z}^\top\mathbf{Z}. \tag{169}$$

as well as

$$\widetilde{\boldsymbol{\Xi}} = \left(\mathbf{I} - \theta_y\widetilde{\mathbf{G}}\right)^{-1} \in \mathbb{R}^{n\times n} \text{ and } \tilde{\mathbf{L}} = \mathbf{1}_n\theta_0 + \mathbf{W}\boldsymbol{\theta}_{\mathrm{w}} + \mathbf{X}\boldsymbol{\theta}_{\mathrm{x}}. \tag{170}$$

Lemma 51. *Under the latent contagion model in Equation (6), suppose that Assumptions 6, 7, 8 and 9 hold. Then*

$$\left\|\mathbb{E}_\theta\big[\widetilde{\mathbf{Z}}^\top\widetilde{\mathbf{Z}}\big]^{-1}\mathbb{E}_\theta\big[\widetilde{\mathbf{Z}}^\top\mathbf{Y}\big] - \mathbb{E}_\theta\big[\mathbf{Z}^\top\mathbf{Z}\big]^{-1}\mathbb{E}_\theta\big[\mathbf{Z}^\top\mathbf{Y}\big]\right\| = o(n^{-1/2}).$$

In preparation for a proof of Lemma 51, we establish a handful of technical results.

Lemma 52. *Under the conditions of Lemma 51,*

$$\frac{1}{n}\left\|\mathbb{E}_\theta\widetilde{\mathbf{Z}}^\top\mathbf{Y}\right\| \le C.$$

Proof. Recalling the structure of $\widetilde{\mathbf{Z}}$ from Equation (110),

$$\left\|\frac{1}{n}\mathbb{E}_\theta\big[\widetilde{\mathbf{Z}}^\top\mathbf{Y}\big]\right\| \le \frac{1}{n}\left\|\mathbb{E}_\theta\mathbf{1}_n^\top\mathbf{Y}\right\| + \frac{1}{n}\left\|\mathbb{E}_\theta\mathbf{W}^\top\mathbf{Y}\right\| + \frac{1}{n}\left\|\mathbb{E}_\theta\mathbf{X}^\top\mathbf{Y}\right\| + \frac{1}{n}\left\|\mathbb{E}_\theta\mathbf{Y}^\top\widetilde{\mathbf{G}}^\top\mathbf{Y}\right\|.$$

Writing $\mathbf{Y} = \widetilde{\boldsymbol{\Xi}}\left(\tilde{\mathbf{L}} + \boldsymbol{\varepsilon}\right)$ as in Equation (170) and using the fact that $\boldsymbol{\varepsilon}$ is conditionally independent of $\mathbf{A}$ conditional on $\mathbf{X}$ under the latent contagion model, the moment bounds in Assumption 6 imply

$$\left\|\frac{1}{n}\mathbb{E}_\theta\big[\widetilde{\mathbf{Z}}^\top\mathbf{Y}\big]\right\| \le C + \frac{1}{n}\left\|\mathbb{E}_\theta\mathbf{Y}^\top\widetilde{\mathbf{G}}^\top\mathbf{Y}\right\|.$$

Again recalling the decomposition of $\mathbf{Y}$ from Equation (170) and using the independence structure of the latent contagion model,

$$\begin{aligned}\left\|\frac{1}{n}\mathbb{E}_\theta[\widetilde{\mathbf{Z}}^\top \mathbf{Y}]\right\| &\leq C + \frac{1}{n}\mathbb{E}_\theta\left\|\tilde{\mathbf{L}}^\top \widetilde{\Xi}^\top \widetilde{\mathbf{G}}^\top \widetilde{\Xi}\tilde{\mathbf{L}}\right\| + \frac{1}{n}\mathbb{E}_\theta\left\|\varepsilon^\top \widetilde{\Xi}^\top \widetilde{\mathbf{G}}^\top \widetilde{\Xi}\varepsilon\right\| \\ &\leq C + \frac{C}{n}\mathbb{E}_\theta\left\|\tilde{\mathbf{L}}\right\|^2 + \frac{C}{n}\mathbb{E}_\theta\left\|\varepsilon\right\|^2,\end{aligned}$$

where the second inequality follows from submultiplicativity and Lemma 1. Items 3 and 4 in Assumption 6 then imply

$$\left\|\frac{1}{n}\mathbb{E}_\theta[\widetilde{\mathbf{Z}}^\top \mathbf{Y}]\right\| \leq C,$$

as we set out to show. □

Lemma 53. *Under the conditions of Lemma 51,*

$$\left\|\mathbb{E}_\theta\left(\mathbf{Z} - \widetilde{\mathbf{Z}}\right)^\top \widetilde{\mathbf{Z}}\right\| = o(\sqrt{n}).$$

Proof. Recalling the structure of $\mathbf{Z}$ and $\widetilde{\mathbf{Z}}$ from Equations (129) and (110), respectively,

$$\left\|\mathbb{E}_\theta\left(\mathbf{Z} - \widetilde{\mathbf{Z}}\right)^\top \widetilde{\mathbf{Z}}\right\| = \left|\mathbb{E}_\theta \mathbf{Y}^\top\left(\mathbf{G} - \widetilde{\mathbf{G}}\right)\widetilde{\mathbf{G}}\mathbf{Y}\right|. \tag{171}$$

Recalling $\mathbf{G} = \mathbf{D}^{-1}\mathbf{A}$ and $\widetilde{\mathbf{G}} = \widetilde{\mathbf{D}}^{-1}\mathbf{P}$, adding and subtracting appropriate quantities yields

$$\begin{aligned}\mathbb{E}_\theta \mathbf{Y}^\top\left(\mathbf{G} - \widetilde{\mathbf{G}}\right)\widetilde{\mathbf{G}}\mathbf{Y} &= \mathbb{E}_\theta \mathbf{Y}^\top\left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1}\right)\mathbf{A}\widetilde{\mathbf{G}}\mathbf{Y} + \mathbb{E}_\theta \mathbf{Y}^\top \widetilde{\mathbf{D}}^{-1}\left(\mathbf{A} - \mathbf{P}\right)\widetilde{\mathbf{G}}\mathbf{Y} \\ &= \mathbb{E}_\theta \mathbf{Y}^\top\left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1}\right)\mathbf{A}\widetilde{\mathbf{G}}\mathbf{Y},\end{aligned}$$

where the second equality follows from the fact that $\mathbf{A} - \mathbf{P}$ is mean zero and independent of ε conditional on $\mathbf{X}$. Rearranging slightly,

$$\begin{aligned}\mathbb{E}_\theta \mathbf{Y}^\top\left(\mathbf{G} - \widetilde{\mathbf{G}}\right)\widetilde{\mathbf{G}}\mathbf{Y} &= \mathbb{E}_\theta \mathbf{Y}^\top\left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right)\mathbf{G}\widetilde{\mathbf{G}}\mathbf{Y} \\ &= \mathbb{E}_\theta \mathbf{Y}^\top\left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right)\left(\mathbf{G} - \widetilde{\mathbf{G}}\right)\widetilde{\mathbf{G}}\mathbf{Y},\end{aligned} \tag{172}$$

where the second equality follows from the fact that

$$\mathbb{E}_\theta \mathbf{Y}^\top\left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right)\widetilde{\mathbf{G}}^2\mathbf{Y} = 0,$$

again because $\mathbf{D} - \widetilde{\mathbf{D}}$ is mean zero and independent of ε conditional on $\mathbf{X}$.

Recalling $\mathbf{G} = \mathbf{D}^{-1}\mathbf{A}$, adding and subtracting appropriate quantities in Equation (172) yields

$$\mathbb{E}_\theta \mathbf{Y}^\top(\mathbf{G} - \widetilde{\mathbf{G}})\widetilde{\mathbf{G}}\mathbf{Y} = \mathbb{E}_\theta \mathbf{Y}^\top\left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right)\widetilde{\mathbf{D}}^{-1}(\mathbf{A} - \mathbf{P})\widetilde{\mathbf{G}}\mathbf{Y} + \mathbb{E}_\theta \mathbf{Y}^\top\left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right)^2\mathbf{G}\widetilde{\mathbf{G}}\mathbf{Y}. \tag{173}$$

Recalling $\mathbf{Y} = \widetilde{\Xi}(\tilde{\mathbf{L}} + \varepsilon)$ from Equation (170), using the fact that ε is mean zero and independent of $\mathbf{A}$ conditional on $\mathbf{X}$,

$$\begin{aligned}\mathbb{E}_\theta \mathbf{Y}^\top\left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right)^2\mathbf{G}\widetilde{\mathbf{G}}\mathbf{Y} &= \mathbb{E}_\theta\left(\widetilde{\Xi}\tilde{\mathbf{L}}\right)^\top\left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right)^2\mathbf{G}\widetilde{\mathbf{G}}\widetilde{\Xi}\tilde{\mathbf{L}} \\ &\quad + \mathbb{E}_\theta\left(\widetilde{\Xi}\varepsilon\right)^\top\left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right)^2\mathbf{G}\widetilde{\mathbf{G}}\widetilde{\Xi}\varepsilon.\end{aligned} \tag{174}$$

Expanding the matrix-vector products,

$$\mathbb{E}_\theta \left(\widetilde{\boldsymbol{\Xi}}\tilde{\mathbf{L}}\right)^\top \left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right)^2 \mathbf{G}\widetilde{\mathbf{G}}\widetilde{\boldsymbol{\Xi}}\tilde{\mathbf{L}} = \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}_\theta \left(\widetilde{\boldsymbol{\Xi}}\tilde{\mathbf{L}}\right)_i \frac{(d_i - \tilde{d}_i)^2}{\tilde{d}_i^2} \left[\mathbf{G}\widetilde{\mathbf{G}}\widetilde{\boldsymbol{\Xi}}\right]_{ij} L_j$$
$$\leq C \sum_{i=1}^n \mathbb{E}_\theta \|\tilde{\mathbf{L}}\|_\infty^2 \frac{(d_i - \tilde{d}_i)^2}{\tilde{d}_i^2} \leq C n \nu_n \sum_{i=1}^n \mathbb{E}_\theta \frac{\|\tilde{\mathbf{L}}\|_\infty^2}{\tilde{d}_i^2}.$$

where the first inequality follows from the fact that $\mathbf{G}\widetilde{\mathbf{G}}\widetilde{\boldsymbol{\Xi}}$ is a transition matrix up to a constant and the second inequality follows from Assumption 8. Using Assumption 6 to ensure that all entries of $\mathbf{X}$ and $\mathbf{W}$ are bounded almost surely, Equation (32) implies

$$\mathbb{E}_\theta \tilde{\mathbf{L}}^\top \left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right)^2 \mathbf{G}\widetilde{\mathbf{G}}\tilde{\mathbf{L}} = o(\sqrt{n}). \tag{175}$$

Applying the triangle inequality to Equation (174) and using the above bound,

$$\left|\mathbb{E}_\theta \mathbf{Y}^\top \left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right)^2 \mathbf{G}\widetilde{\mathbf{G}}\mathbf{Y}\right| \leq \left|\mathbb{E}_\theta \boldsymbol{\varepsilon}^\top \left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right)^2 \mathbf{G}\widetilde{\mathbf{G}}\boldsymbol{\varepsilon}\right| + o(\sqrt{n}).$$

An argument analogous to that leading to Equation (175), this time using Assumption 6 and the fact that $\boldsymbol{\varepsilon}$ has uncorrelated entries, yields

$$\left|\mathbb{E}_\theta \boldsymbol{\varepsilon}^\top \left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right)^2 \mathbf{G}\widetilde{\mathbf{G}}\boldsymbol{\varepsilon}\right| = o(\sqrt{n}),$$

and therefore

$$\left|\mathbb{E}_\theta \mathbf{Y}^\top \left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right)^2 \mathbf{G}\widetilde{\mathbf{G}}\mathbf{Y}\right| = o(\sqrt{n}). \tag{176}$$

Recalling Equation (170) and using the fact that $\boldsymbol{\varepsilon}$ is mean zero and independent of $\mathbf{A} - \mathbf{P}$ conditional on $\mathbf{X}$,

$$\begin{aligned}\mathbb{E}_\theta \mathbf{Y}^\top \left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right) \widetilde{\mathbf{D}}^{-1} (\mathbf{A} - \mathbf{P}) \widetilde{\mathbf{G}}\mathbf{Y} = &\ \mathbb{E}_\theta \left(\widetilde{\boldsymbol{\Xi}}\tilde{\mathbf{L}}\right)^\top \left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right) \widetilde{\mathbf{D}}^{-1} (\mathbf{A} - \mathbf{P}) \widetilde{\mathbf{G}} \left(\widetilde{\boldsymbol{\Xi}}\tilde{\mathbf{L}}\right) \\ &+ \mathbb{E}_\theta \left(\widetilde{\boldsymbol{\Xi}}\boldsymbol{\varepsilon}\right)^\top \left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right) \widetilde{\mathbf{D}}^{-1} (\mathbf{A} - \mathbf{P}) \widetilde{\mathbf{G}}\widetilde{\boldsymbol{\Xi}}\boldsymbol{\varepsilon}.\end{aligned} \tag{177}$$

Expanding the matrix-vector products,

$$\mathbb{E}_\theta \left(\widetilde{\boldsymbol{\Xi}}\tilde{\mathbf{L}}\right)^\top \left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right) \widetilde{\mathbf{D}}^{-1} (\mathbf{A} - \mathbf{P}) \widetilde{\mathbf{G}}\widetilde{\boldsymbol{\Xi}}\tilde{\mathbf{L}} = \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}_\theta \frac{(d_i - \tilde{d}_i) \left(\widetilde{\boldsymbol{\Xi}}\tilde{\mathbf{L}}\right)_i}{\tilde{d}_i^2} (\mathbf{A} - \mathbf{P})_{ij} \left(\widetilde{\mathbf{G}}\widetilde{\boldsymbol{\Xi}}\tilde{\mathbf{L}}\right)_j.$$

By our conditional edge independence assumptions, we have

$$\mathbb{E}_\theta \left[(d_i - \tilde{d}_i)(\mathbf{A} - \mathbf{P})_{ij} \mid \mathbf{X}\right] = \mathbb{E}_\theta \left[(\mathbf{A} - \mathbf{P})_{ij}^2 \mid \mathbf{X}\right],$$

and thus,

$$\mathbb{E}_\theta \left(\widetilde{\boldsymbol{\Xi}}\tilde{\mathbf{L}}\right)^\top \left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right) \widetilde{\mathbf{D}}^{-1} (\mathbf{A} - \mathbf{P}) \widetilde{\mathbf{G}}\widetilde{\boldsymbol{\Xi}}\tilde{\mathbf{L}} = \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}_\theta \frac{\left(\widetilde{\boldsymbol{\Xi}}\tilde{\mathbf{L}}\right)_i}{\tilde{d}_i^2} (\mathbf{A} - \mathbf{P})_{ij}^2 \left(\widetilde{\mathbf{G}}\widetilde{\boldsymbol{\Xi}}\tilde{\mathbf{L}}\right)_j$$
$$\leq C n \nu_n \sum_{i=1}^n \mathbb{E}_\theta \frac{\|\tilde{\mathbf{L}}\|_\infty^2}{\tilde{d}_i^2},$$

where the inequality follows from our assumption in Equation (31) and the fact that $\widetilde{\Xi}$ and $\widetilde{\mathbf{G}}\widetilde{\Xi}$ are both transition matrices, up to multiplicative factors. Applying Item 3 from Assumption 6 to bound $\|\tilde{\mathbf{L}}\| \le C$ almost surely, our assumption in Equation (32) implies

$$\left|\mathbb{E}_\theta \left(\widetilde{\Xi}\tilde{\mathbf{L}}\right)^\top \left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right) \widetilde{\mathbf{D}}^{-1} (\mathbf{A} - \mathbf{P}) \widetilde{\mathbf{G}}\widetilde{\Xi}\tilde{\mathbf{L}}\right| = o(\sqrt{n}).$$

Applying this to Equation (177),

$$\mathbb{E}_\theta \mathbf{Y}^\top \left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right) \widetilde{\mathbf{D}}^{-1} (\mathbf{A} - \mathbf{P}) \widetilde{\mathbf{G}} \mathbf{Y} = \mathbb{E}_\theta \left(\widetilde{\Xi}\boldsymbol{\varepsilon}\right)^\top \left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right) \widetilde{\mathbf{D}}^{-1} (\mathbf{A} - \mathbf{P}) \widetilde{\mathbf{G}}\widetilde{\Xi}\boldsymbol{\varepsilon} + o(\sqrt{n}). \tag{178}$$

Since the entries of $\boldsymbol{\varepsilon}$ are uncorrelated by assumption,

$$\mathbb{E}_\theta \left(\widetilde{\Xi}\boldsymbol{\varepsilon}\right)^\top \left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right) \widetilde{\mathbf{D}}^{-1} (\mathbf{A} - \mathbf{P}) \widetilde{\mathbf{G}}\widetilde{\Xi}\boldsymbol{\varepsilon} = \sigma_\epsilon^2 \sum_{i=1}^n \mathbb{E}_\theta \left[\widetilde{\Xi}^\top \left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right) \widetilde{\mathbf{D}}^{-1} (\mathbf{A} - \mathbf{P}) \widetilde{\mathbf{G}}\widetilde{\Xi}\right]_{ii}.$$

Expanding the matrix products,

$$\begin{aligned}
\mathbb{E}_\theta \left(\widetilde{\Xi}\boldsymbol{\varepsilon}\right)^\top \left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right) \widetilde{\mathbf{D}}^{-1} (\mathbf{A} - \mathbf{P}) \widetilde{\mathbf{G}}\widetilde{\Xi}\boldsymbol{\varepsilon} &= \sigma_\epsilon^2 \sum_{i=1}^n \mathbb{E}_\theta \sum_{j=1}^n [\widetilde{\Xi}]_{ji} \frac{(\tilde{d}_j - d_j)}{\tilde{d}_j^2} \left[(\mathbf{A} - \mathbf{P}) \widetilde{\mathbf{G}}\widetilde{\Xi}\right]_{ji} \\
&= \sigma_\epsilon^2 \mathbb{E}_\theta \sum_{i=1}^n \sum_{j=1}^n [\widetilde{\Xi}]_{ji} \frac{(\tilde{d}_j - d_j)}{\tilde{d}_j^2} \sum_{k=1}^n (\mathbf{A} - \mathbf{P})_{jk} \left[\widetilde{\mathbf{G}}\widetilde{\Xi}\right]_{ki}.
\end{aligned}$$

Owing to the conditional independence structure of $\mathbf{A} - \mathbf{P}$ and our assumption in Equation (31), we have

$$\mathbb{E}\left[(\tilde{d}_j - d_j)\,(\mathbf{A} - \mathbf{P})_{jk} \middle| \mathbf{X}\right] = \mathbb{E}\left[(\mathbf{A} - \mathbf{P})_{jk}^2 \mid \mathbf{X}\right] \le C\nu,$$

from which

$$\begin{aligned}
\left|\mathbb{E}_\theta \left(\widetilde{\Xi}\boldsymbol{\varepsilon}\right)^\top \left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right) \widetilde{\mathbf{D}}^{-1} (\mathbf{A} - \mathbf{P}) \widetilde{\mathbf{G}}\widetilde{\Xi}\boldsymbol{\varepsilon}\right| &\le C\sigma_\epsilon^2 \nu \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \mathbb{E}_\theta \frac{[\widetilde{\Xi}]_{ji} \left[\widetilde{\mathbf{G}}\widetilde{\Xi}\right]_{ki}}{\tilde{d}_j^2} \\
&\le C\sigma_\epsilon^2 \nu_n n \sum_{j=1}^n \mathbb{E}_\theta \frac{1}{\tilde{d}_j^2},
\end{aligned}$$

where the second inequality follows from the fact that $\widetilde{\Xi}$ and $\widetilde{\mathbf{G}}\widetilde{\Xi}$ are both transition matrices, up to multiplicative factors. Applying Equation (32),

$$\left|\mathbb{E}_\theta \left(\widetilde{\Xi}\boldsymbol{\varepsilon}\right)^\top \left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right) \widetilde{\mathbf{D}}^{-1} (\mathbf{A} - \mathbf{P}) \widetilde{\mathbf{G}}\widetilde{\Xi}\boldsymbol{\varepsilon}\right| = o(\sqrt{n}).$$

Applying this to Equation (178),

$$\left|\mathbb{E}_\theta \mathbf{Y}^\top \left(\mathbf{I} - \mathbf{D}\widetilde{\mathbf{D}}^{-1}\right) \widetilde{\mathbf{D}}^{-1} (\mathbf{A} - \mathbf{P}) \widetilde{\mathbf{G}} \mathbf{Y}\right| = o(\sqrt{n}). \tag{179}$$

Applying the triangle inequality in Equation (173) followed by Equations (176) and (179) completes the proof. □

Lemma 54. *Under the conditions of Lemma 51,*

$$\left\| \mathbb{E}_\theta \left(\mathbf{Z} - \widetilde{\mathbf{Z}}\right)^\top \left(\mathbf{Z} - \widetilde{\mathbf{Z}}\right) \right\| = o(\sqrt{n}).$$

Proof. Recalling the structure of $\mathbf{Z}$ and $\widetilde{\mathbf{Z}}$ from Equations (129) and (110), respectively,

$$\left\| \mathbb{E}_\theta \left(\mathbf{Z} - \widetilde{\mathbf{Z}}\right)^\top \left(\mathbf{Z} - \widetilde{\mathbf{Z}}\right) \right\| = \left| \mathbb{E}_\theta \mathbf{Y}^\top \left(\mathbf{G} - \widetilde{\mathbf{G}}\right)^\top \left(\mathbf{G} - \widetilde{\mathbf{G}}\right) \mathbf{Y} \right|.$$

Recalling the definitions of $\widetilde{\Xi}$ and $\tilde{\mathbf{L}}$ from Equation (170), we have $\mathbf{Y} = \widetilde{\Xi}\left(\tilde{\mathbf{L}} + \boldsymbol{\varepsilon}\right)$. Since $\boldsymbol{\varepsilon}$ is independent of $\mathbf{A}$ conditional on $\mathbf{X}$ under the latent contagion model in Equation (5), we have

$$\mathbb{E}_\theta \mathbf{Y}^\top \left(\mathbf{G} - \widetilde{\mathbf{G}}\right)^\top \left(\mathbf{G} - \widetilde{\mathbf{G}}\right) \mathbf{Y} = \mathbb{E}_\theta \left\| \left(\mathbf{G} - \widetilde{\mathbf{G}}\right) \widetilde{\Xi} \tilde{\mathbf{L}} \right\|^2 + \mathbb{E}_\theta \left\| \left(\mathbf{G} - \widetilde{\mathbf{G}}\right) \widetilde{\Xi} \boldsymbol{\varepsilon} \right\|^2,$$

so that

$$\left\| \mathbb{E}_\theta \left(\mathbf{Z} - \widetilde{\mathbf{Z}}\right)^\top \left(\mathbf{Z} - \widetilde{\mathbf{Z}}\right) \right\| = \mathbb{E}_\theta \left\| \left(\mathbf{G} - \widetilde{\mathbf{G}}\right) \widetilde{\Xi} \tilde{\mathbf{L}} \right\|^2 + \mathbb{E}_\theta \left\| \left(\mathbf{G} - \widetilde{\mathbf{G}}\right) \widetilde{\Xi} \boldsymbol{\varepsilon} \right\|^2. \tag{180}$$

Applying the triangle inequality,

$$\mathbb{E}_\theta \left\| \left(\mathbf{G} - \widetilde{\mathbf{G}}\right) \widetilde{\Xi} \tilde{\mathbf{L}} \right\|^2 \le C \mathbb{E}_\theta \left\| \widetilde{\mathbf{D}}^{-1} (\mathbf{A} - \mathbf{P}) \widetilde{\Xi} \tilde{\mathbf{L}} \right\|^2 + C \mathbb{E}_\theta \left\| \left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1}\right) \mathbf{A} \widetilde{\Xi} \tilde{\mathbf{L}} \right\|^2. \tag{181}$$

Expanding the norm,

$$\mathbb{E}_\theta \left\| \widetilde{\mathbf{D}}^{-1} (\mathbf{A} - \mathbf{P}) \widetilde{\Xi} \tilde{\mathbf{L}} \right\|^2 = \sum_{i=1}^n \mathbb{E}_\theta \left[\frac{1}{\tilde{d}_i} \sum_{j=1}^n (\mathbf{A} - \mathbf{P})_{ij} \left(\widetilde{\Xi} \tilde{\mathbf{L}}\right)_j \right]^2 = \sum_{i=1}^n \mathbb{E}_\theta \frac{1}{\tilde{d}_i^2} \left[\sum_{j=1}^n (\mathbf{A} - \mathbf{P})_{ij} \left(\widetilde{\Xi} \tilde{\mathbf{L}}\right)_j \right]^2.$$

Since the entries of $\mathbf{A} - \mathbf{P}$ are independent mean zero conditional on the latent positions, expanding the square yields

$$\mathbb{E}_\theta \left\| \widetilde{\mathbf{D}}^{-1} (\mathbf{A} - \mathbf{P}) \widetilde{\Xi} \tilde{\mathbf{L}} \right\|^2 = \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}_\theta \frac{1}{\tilde{d}_i^2} (\mathbf{A} - \mathbf{P})_{ij}^2 \left(\widetilde{\Xi} \tilde{\mathbf{L}}\right)_j^2 \le C \nu_n \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}_\theta \frac{\left(\widetilde{\Xi} \tilde{\mathbf{L}}\right)_j^2}{\tilde{d}_i^2},$$

where the second inequality follows from Equation (31). Since $\widetilde{\Xi}$ is a transition matrix up to a scaling factor,

$$\mathbb{E}_\theta \left\| \widetilde{\mathbf{D}}^{-1} (\mathbf{A} - \mathbf{P}) \widetilde{\Xi} \tilde{\mathbf{L}} \right\|^2 \le C \nu_n \sum_{i=1}^n \mathbb{E}_\theta \frac{\|\tilde{\mathbf{L}}\|_\infty^2}{\tilde{d}_i^2},$$

and Item 3 from Assumption 6 along with Equation (32) yields

$$\mathbb{E}_\theta \left\| \widetilde{\mathbf{D}}^{-1} (\mathbf{A} - \mathbf{P}) \widetilde{\Xi} \tilde{\mathbf{L}} \right\|^2 = o(\sqrt{n}). \tag{182}$$

Again using the fact that $\mathbf{G}\widetilde{\Xi}$ is a transition matrix up to scaling,

$$\mathbb{E}_\theta \left\| \left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1}\right) \mathbf{A} \widetilde{\Xi} \tilde{\mathbf{L}} \right\|^2 = \mathbb{E}_\theta \left\| \widetilde{\mathbf{D}}^{-1} \left(\widetilde{\mathbf{D}} - \mathbf{D}\right) \mathbf{G} \widetilde{\Xi} \tilde{\mathbf{L}} \right\|^2 = \sum_{i=1}^n \mathbb{E}_\theta \left(\frac{d_i - \tilde{d}_i}{\tilde{d}_i} \right)^2 \left(\mathbf{G} \widetilde{\Xi} \tilde{\mathbf{L}}\right)_i^2$$
$$\le C \sum_{i=1}^n \mathbb{E}_\theta \left(\frac{d_i - \tilde{d}_i}{\tilde{d}_i} \right)^2 \|\tilde{\mathbf{L}}\|_\infty^2.$$

Using conditional independence of the entries of $\mathbf{A} - \mathbf{P}$ along with Equation (31),

$$\mathbb{E}_\theta \left\| \left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1} \right) \mathbf{A} \widetilde{\Xi} \tilde{\mathbf{L}} \right\|^2 \le C n \nu_n \sum_{i=1}^n \mathbb{E}_\theta \frac{\|\tilde{\mathbf{L}}\|_\infty^2}{\tilde{d}_i^2},$$

and Item 3 from Assumption 6 along with Equation (32) yields

$$\mathbb{E}_\theta \left\| \left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1} \right) \mathbf{A} \widetilde{\Xi} \tilde{\mathbf{L}} \right\|^2 = o(\sqrt{n}).$$

Applying this and Equation (182) to Equation (181),

$$\mathbb{E}_\theta \left\| \left(\mathbf{G} - \widetilde{\mathbf{G}} \right) \widetilde{\Xi} \tilde{\mathbf{L}} \right\|^2 = o(\sqrt{n}).$$

Applying this to Equation (180) in turn,

$$\left\| \mathbb{E}_\theta \left(\mathbf{Z} - \widetilde{\mathbf{Z}} \right)^\top \left(\mathbf{Z} - \widetilde{\mathbf{Z}} \right) \right\| \le \mathbb{E}_\theta \left\| \left(\mathbf{G} - \widetilde{\mathbf{G}} \right) \widetilde{\Xi} \boldsymbol{\varepsilon} \right\|^2 + o(\sqrt{n}). \tag{183}$$

Expanding the norm, using the independence assumptions of the latent contagion model, and applying our assumption that the entries of $\boldsymbol{\varepsilon}$ are uncorrelated,

$$\mathbb{E}_\theta \left\| \left(\mathbf{G} - \widetilde{\mathbf{G}} \right) \widetilde{\Xi} \varepsilon \right\|^2 = \sum_{i=1}^n \mathbb{E}_\theta \varepsilon_i^2 \left[\widetilde{\Xi}^\top \left(\mathbf{G} - \widetilde{\mathbf{G}} \right)^\top \left(\mathbf{G} - \widetilde{\mathbf{G}} \right) \widetilde{\Xi} \right]_{i,i} \le \sigma_\epsilon^2 \mathbb{E}_\theta \left\| \left(\mathbf{G} - \widetilde{\mathbf{G}} \right) \widetilde{\Xi} \right\|_F^2.$$

Using basic properties of the Frobenius norm and Lemma 1,

$$\mathbb{E}_\theta \left\| \left(\mathbf{G} - \widetilde{\mathbf{G}} \right) \widetilde{\Xi} \varepsilon \right\|^2 \le C \sigma_\epsilon^2 \mathbb{E}_\theta \left\| \mathbf{G} - \widetilde{\mathbf{G}} \right\|_F^2.$$

Applying the triangle inequality,

$$\mathbb{E}_\theta \left\| \left(\mathbf{G} - \widetilde{\mathbf{G}} \right) \widetilde{\Xi} \boldsymbol{\varepsilon} \right\|^2 \le C \sigma_\epsilon^2 \left[\mathbb{E}_\theta \left\| \widetilde{\mathbf{D}}^{-1} \left(\mathbf{A} - \mathbf{P} \right) \right\|_F^2 + \mathbb{E}_\theta \left\| \left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1} \right) \mathbf{A} \right\|_F^2 \right]. \tag{184}$$

Expanding the norm, using conditional independence of the edges and Equation (31),

$$\mathbb{E}_\theta \left\| \widetilde{\mathbf{D}}^{-1} \left(\mathbf{A} - \mathbf{P} \right) \right\|_F^2 = \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}_\theta \frac{1}{\tilde{d}_i^2} \mathbb{E} \left[\left(\mathbf{A} - \mathbf{P} \right)_{i,j}^2 \mid \mathbf{X} \right] \le C n \nu_n \sum_{i=1}^n \mathbb{E}_\theta \frac{1}{\tilde{d}_i^2}.$$

Applying Equation (32),

$$\mathbb{E}_\theta \left\| \widetilde{\mathbf{D}}^{-1} \left(\mathbf{A} - \mathbf{P} \right) \right\|_F^2 = o(\sqrt{n}). \tag{185}$$

Using the fact that the rows of $\mathbf{G} = \mathbf{D}^{-1}\mathbf{A}$ sum to 1 and all entries are between 0 and 1,

$$\mathbb{E}_\theta \left\| \left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1} \right) \mathbf{A} \right\|_F^2 = \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}_\theta \frac{(\tilde{d}_i - d_i)^2}{\tilde{d}_i^2} G_{ij}^2 \le \sum_{i=1}^n \mathbb{E}_\theta \frac{(\tilde{d}_i - d_i)^2}{\tilde{d}_i^2}.$$

By conditional independence of the edges along with Equations (31) and (32),

$$\mathbb{E}_\theta \left\| \left(\mathbf{D}^{-1} - \widetilde{\mathbf{D}}^{-1} \right) \mathbf{A} \right\|_F^2 \le C n \nu_n \sum_{i=1}^n \mathbb{E}_\theta \frac{1}{\tilde{d}_i^2} = o(\sqrt{n}).$$

Applying this and Equation (185) to Equation (184),

$$\mathbb{E}_\theta \left\| \left(\mathbf{G} - \widetilde{\mathbf{G}}\right) \widetilde{\Xi}\varepsilon \right\|^2 = o(\sqrt{n}).$$

Applying this to Equation (183) completes the proof. □

Lemma 55. *Under the conditions of Lemma 51, let $\mathbf{\Sigma}$ and $\widetilde{\mathbf{\Sigma}}$ be as in Equation (169). Then*

$$\left\| \mathbb{E}_\theta \mathbf{\Sigma} - \mathbb{E}_\theta \widetilde{\mathbf{\Sigma}} \right\| = o(n^{-1/2}).$$

Proof. Adding and subtracting appropriate quantities and applying the triangle inequality,

$$\left\| \mathbb{E}_\theta \mathbf{\Sigma} - \mathbb{E}_\theta \widetilde{\mathbf{\Sigma}} \right\| \le \frac{2}{n} \left\| \mathbb{E}_\theta \left(\mathbf{Z} - \widetilde{\mathbf{Z}}\right)^\top \widetilde{\mathbf{Z}} \right\| + \frac{1}{n} \left\| \mathbb{E}_\theta \left(\mathbf{Z} - \widetilde{\mathbf{Z}}\right)^\top \left(\mathbf{Z} - \widetilde{\mathbf{Z}}\right) \right\|.$$

Applying Lemmas 53 and 54 yields the result. □

Lemma 56. *Under the conditions of Lemma 51, $\mathbb{E}_\theta \mathbf{\Sigma}$ is invertible and*

$$\left\| [\mathbb{E}_\theta \mathbf{\Sigma}]^{-1} \right\| \le C.$$

Proof. By Lemma 55,

$$\left\| \mathbb{E}_\theta \mathbf{\Sigma} - \mathbb{E}_\theta \widetilde{\mathbf{\Sigma}} \right\| = o(n^{-1/2}).$$

Combining this with our assumption that $\mathbb{E}_\theta \widetilde{\mathbf{\Sigma}}$ converges to an invertible matrix completes the proof. □

Lemma 57. *Under the conditions of Lemma 51, $\widetilde{\mathbf{\Sigma}}$ and $\mathbf{\Sigma}$, as defined in Equation (169), are both invertible and obey*

$$\left\| (\mathbb{E}_\theta \mathbf{\Sigma})^{-1} - \left(\mathbb{E}_\theta \widetilde{\mathbf{\Sigma}}\right)^{-1} \right\| = o(n^{-1/2}).$$

Proof. By Lemma 56 and Assumption 7, both $\mathbb{E}_\theta \mathbf{\Sigma}$ and $\mathbb{E}_\theta \widetilde{\mathbf{\Sigma}}$ are invertible for n suitably large. Applying submultiplicativity,

$$\left\| (\mathbb{E}_\theta \mathbf{\Sigma})^{-1} - \left(\mathbb{E}_\theta \widetilde{\mathbf{\Sigma}}\right)^{-1} \right\| \le \left\| (\mathbb{E}_\theta \mathbf{\Sigma})^{-1} \right\| \left\| \left(\mathbb{E}_\theta \widetilde{\mathbf{\Sigma}}\right)^{-1} \right\| \left\| \mathbb{E}_\theta \mathbf{\Sigma} - \mathbb{E}_\theta \widetilde{\mathbf{\Sigma}} \right\|.$$

Applying Lemma 56 and Assumption 7 again, it follows that

$$\left\| (\mathbb{E}_\theta \mathbf{\Sigma})^{-1} - \left(\mathbb{E}_\theta \widetilde{\mathbf{\Sigma}}\right)^{-1} \right\| \le C \left\| \mathbb{E}_\theta \mathbf{\Sigma} - \mathbb{E}_\theta \widetilde{\mathbf{\Sigma}} \right\|.$$

Applying Lemma 55 completes the proof. □

Proof of Lemma 51. Recalling the definitions from Equation (169),

$$\left\| \mathbb{E}_\theta [\widetilde{\mathbf{Z}}^\top \widetilde{\mathbf{Z}}]^{-1} \mathbb{E}_\theta [\widetilde{\mathbf{Z}}^\top \mathbf{Y}] - \mathbb{E}_\theta [\mathbf{Z}^\top \mathbf{Z}]^{-1} \mathbb{E}_\theta [\mathbf{Z}^\top \mathbf{Y}] \right\| = \left\| [\mathbb{E}_\theta \widetilde{\mathbf{\Sigma}}]^{-1} \mathbb{E}_\theta \frac{\widetilde{\mathbf{Z}}^\top \mathbf{Y}}{n} - [\mathbb{E}_\theta \mathbf{\Sigma}]^{-1} \mathbb{E}_\theta \frac{\mathbf{Z}^\top \mathbf{Y}}{n} \right\|.$$

By the triangle inequality, it will suffice for us to show that

$$\left\| [\mathbb{E}_\theta \boldsymbol{\Sigma}]^{-1} \, \mathbb{E}_\theta \frac{(\mathbf{Z} - \widetilde{\mathbf{Z}})^\top \mathbf{Y}}{n} \right\| = o(n^{-1/2}) \tag{186}$$

and

$$\left\| \left[(\mathbb{E}_\theta \boldsymbol{\Sigma})^{-1} - \left(\mathbb{E}_\theta \widetilde{\boldsymbol{\Sigma}}\right)^{-1} \right] \mathbb{E}_\theta \frac{\widetilde{\mathbf{Z}}^\top \mathbf{Y}}{n} \right\| = o(n^{-1/2}). \tag{187}$$

To see Equation (187), observe that by submultiplicativity,

$$\left\| \left[(\mathbb{E}_\theta \boldsymbol{\Sigma})^{-1} - \left(\mathbb{E}_\theta \widetilde{\boldsymbol{\Sigma}}\right)^{-1} \right] \mathbb{E}_\theta \frac{\widetilde{\mathbf{Z}}^\top \mathbf{Y}}{n} \right\| \leq \left\| (\mathbb{E}_\theta \boldsymbol{\Sigma})^{-1} - \left(\mathbb{E}_\theta \widetilde{\boldsymbol{\Sigma}}\right)^{-1} \right\| \left\| \mathbb{E}_\theta \frac{\widetilde{\mathbf{Z}}^\top \mathbf{Y}}{n} \right\|,$$

from which Lemmas 52 and 57 yield Equation (187).

To see Equation (186), note that submultiplicativity and Lemma 56 imply

$$\left\| \mathbb{E}_\theta [\mathbf{Z}^\top \mathbf{Z}]^{-1} \mathbb{E}_\theta \left(\mathbf{Z} - \widetilde{\mathbf{Z}}\right)^\top \mathbf{Y} \right\| \leq C \left\| \frac{1}{n} \mathbb{E}_\theta \left(\mathbf{Z} - \widetilde{\mathbf{Z}}\right)^\top \mathbf{Y} \right\|,$$

and Lemma 53 yields Equation (186), completing the proof. □

K.2 Projection equivalence under peer contagion

A result analogous to Lemma 51 holds under the peer contagion model. We state the result here, but omit the details for the sake of space, as the proof largely follows the same argument as for Lemma 51, with more complicated bookkeeping owing to the dependence of ε on $\mathbf{G}$.

Lemma 58. *Under the peer contagion model in Equation (4), suppose that Assumptions 5, 7, 8 and 9 hold. Then*

$$\left\| \mathbb{E}_\beta [\mathbf{Z}^\top \mathbf{Z}]^{-1} \mathbb{E}_\beta [\mathbf{Z}^\top \mathbf{Y}] - \mathbb{E}_\beta [\widetilde{\mathbf{Z}}^\top \widetilde{\mathbf{Z}}]^{-1} \mathbb{E}_\beta [\widetilde{\mathbf{Z}}^\top \mathbf{Y}] \right\| = o(n^{-1/2}).$$

References

Athreya, Avanti, Donniell E Fishkind, Minh Tang, Carey E Priebe, Youngser Park, Joshua T Vogelstein, Keith Levin, Vince Lyzinski, Yichen Qin, and Daniel L Sussman (2018). "Statistical Inference on Random Dot Product Graphs: A Survey". In: *Journal of Machine Learning Research* 18, pp. 1–92.

Besag, Julian (Jan. 1974). "Spatial Interaction and the Statistical Analysis of Lattice Systems". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 36.2, pp. 192–225. DOI: 10.1111/j.2517-6161.1974.tb00999.x.

Bhatia, Rajendra (1997). *Matrix Analysis*. Springer.

Bidart, Claire and Johanne Charbonneau (Aug. 2011). "How to Generate Personal Networks: Issues and Tools for a Sociological Perspective". In: *Field Methods* 23.3, pp. 266–286. DOI: 10.1177/1525822X11408513.

Bivand, Roger, Giovanni Millo, and Gianfranco Piras (June 2021). "A Review of Software for Spatial Econometrics in R". In: *Mathematics* 9.11, p. 1276. DOI: 10.3390/math9111276.

Boucheron, Stéphane, Gábor Lugosi, and Pascal Massart (Feb. 2013). *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press. DOI: 10.1093/acprof:oso/9780199535255.001.0001.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin (May 2009). "Identification of Peer Effects through Social Networks". In: *Journal of Econometrics* 150.1, pp. 41–55. DOI: 10.1016/j.jeconom.2008.12.021.

— (2020). "Peer Effects in Networks: A Survey". In: *Annual Review of Economics* 12.1, pp. 603–629.

Chan, Ga Ming Angus and Tianxi Li (2023). "Fitting Low-rank Models on Egocentrically Sampled Partial Networks". In: *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics*. Vol. 206. Valencia, Spain: PMLR.

Chang, Jae Ho and Subhadeep Paul (June 2024). *Embedding Network Autoregression for Time Series Analysis and Causal Peer Effect Inference*. arXiv: 2406.05944 [math, stat].

— (Mar. 2025). *Embedding Network Autoregression for Time Series Analysis and Causal Peer Effect Inference*. DOI: 10.48550/arXiv.2406.05944. arXiv: 2406.05944 [stat].

Chen, Elynn Y., Jianqing Fan, and Xuening Zhu (Aug. 2023). "Community Network Auto-Regression for High-Dimensional Time Series". In: *Journal of Econometrics* 235.2, pp. 1239–1256. DOI: 10.1016/j.jeconom.2022.10.005.

Chin, Alex (Sept. 2019). "Regression Adjustments for Estimating the Global Treatment Effect in Experiments with Interference". In: *Journal of Causal Inference* 7.2, p. 20180026. DOI: 10.1515/jci-2018-0026.

Cho, Juhee, Donggyu Kim, and Karl Rohe (Apr. 2019). "Intelligent Initialization and Adaptive Thresholding for Iterative Matrix Completion: Some Statistical and Algorithmic Theory for Adaptive-Impute". In: *Journal of Computational and Graphical Statistics* 28.2, pp. 323–333. DOI: 10.1080/10618600.2018.1518238.

De Bacco, Caterina, Martina Contisciani, Jonathan Cardoso-Silva, Hadiseh Safdari, Diego Baptista, Tracy Sweet, Jean-Gabriel Young, Jeremy Koster, Cody T. Ross, Richard McElreath, Daniel Redhead, and Eleanor A. Power (Dec. 2021). "Latent Network Models to Account for Noisy, Multiply-Reported Social Network Data". In: *arXiv:2112.11396 [cs]*. arXiv: 2112.11396 [cs].

Di Maria, Chiara, Antonino Abbruzzo, and Gianfranco Lovison (Feb. 2022). "Networks as Mediating Variables: A Bayesian Latent Space Approach". In: *Statistical Methods & Applications* 31, pp. 1015–1035. DOI: 10.1007/s10260-022-00621-w.

Drukker, David M., Peter Egger, and Ingmar R. Prucha (May 2013). "On Two-Step Estimation of a Spatial Autoregressive Model with Autoregressive Disturbances and Endogenous Regressors". In: *Econometric Reviews* 32.5-6, pp. 686–733. DOI: 10.1080/07474938.2013.741020.

Egami, Naoki and Eric J. Tchetgen Tchetgen (Sept. 2021). "Identification and Estimation of Causal Peer Effects Using Double Negative Controls for Unmeasured Network Confounding". In: *arXiv:2109.01933 [stat]*. arXiv: 2109.01933 [stat].

Griffith, Alan (Oct. 2022). "Name Your Friends, but Only Five? The Importance of Censoring in Peer Effects Estimates Using Social Network Data". In: *Journal of Labor Economics* 40.4, pp. 779–805. DOI: 10.1086/717935.

Griffith, Alan and Jungyoun Kim (July 2024). "The Impact of Missing Links on Linear Reduced-form Network-Based Peer Effects Estimates". In.

Hardy, Morgan, Rachel M. Heath, Wesley Lee, and Tyler H. McCormick (Mar. 2024). *Estimating Spillovers Using Imprecisely Measured Networks*. DOI: 10.48550/arXiv.1904.00136. arXiv: 1904.00136 [stat].

Hayes, Alex, Mark M Fredrickson, and Keith Levin (2025). "Estimating Network-Mediated Causal Effects via Principal Components Network Regression". In: *Journal of Machine Learning Research* 26, pp. 1–99.

Hayes, Alex and Keith Levin (Nov. 2025). *Minimax Rates for the Linear-in-Means Model Reveal an Identifiability-Estimability Gap*. DOI: 10.48550/arXiv.2410.10772. arXiv: 2410.10772 [stat].

Johnsson, Ida and Hyungsik Roger Moon (July 2019). *Estimation of Peer Effects in Endogenous Social Networks: Control Function Approach*. arXiv: 1709.10024 [econ].

Karrer, Brian and M. E. J. Newman (Jan. 2011). "Stochastic Blockmodels and Community Structure in Networks". In: *Physical Review E* 83.1, p. 016107. DOI: 10.1103/PhysRevE.83.016107. arXiv: 1008.3926.

Kelejian, Harry H and Ingmar R Prucha (1998). "A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances". In: *Journal of Real Estate Finance and Economics* Vol. 17.1, pp. 99–121. DOI: 10.1023/A:1007707430416.

— (Sept. 2007). "HAC Estimation in a Spatial Framework". In: *Journal of Econometrics* 140.1, pp. 131–154. DOI: 10.1016/j.jeconom.2006.09.005.

Lee, Lung-Fei (Apr. 2002). "Consistency and Efficiency of Least Squares Estimation for Mixed Regressive, Spatial Autoregressive Models". In: *Econometric Theory* 18.2, pp. 252–277. DOI: 10.1017/S0266466602182028.

— (Jan. 2003). "Best Spatial Two-Stage Least Squares Estimators for a Spatial Autoregressive Model with Autoregressive Disturbances". In: *Econometric Reviews* 22.4, pp. 307–335. DOI: 10.1081/ETC-120025891.

— (Nov. 2004). "Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Autoregressive Models". In: *Econometrica* 72.6, pp. 1899–1925. DOI: 10.1111/j.1468-0262.2004.00558.x.

Lee, Lung-Fei, Xiaodong Liu, and Xu Lin (July 2010). "Specification and Estimation of Social Interaction Models with Network Structures: Estimation of Social Network Models". In: *Econometrics Journal* 13.2, pp. 145–176. DOI: 10.1111/j.1368-423X.2010.00310.x.

LeSage, James P. and R. Kelley Pace (2009). *Introduction to Spatial Econometrics*. Statistics, Textbooks and Monographs. Boca Raton: CRC Press.

Leung, Michael P. (2022). "Causal Inference Under Approximate Neighborhood Interference". In: *Econometrica* 90.1, pp. 267–293. DOI: 10.3982/ECTA17841.

Levin, Keith, Avanti Athreya, Minh Tang, Vince Lyzinski, Youngser Park, and Carey E. Priebe (June 2019). "A Central Limit Theorem for an Omnibus Embedding of Multiple Random Graphs and Implications for Multiscale Network Inference". In: *arXiv:1705.09355 [stat]*. arXiv: 1705.09355 [stat].

Levin, Keith, Asad Lodhia, and Elizaveta Levina (2022). "Recovering Shared Structure from Multiple Networks with Unknown Edge Distributions". In: *Journal of Machine Learning Research* 23, pp. 1–48.
Lewbel, Arthur, Xi Qu, and Xun Tang (Apr. 2023). "Social Networks with Unobserved Links". In: *Journal of Political Economy* 131.4, pp. 898–946. DOI: 10.1086/722090.
— (June 2024a). *Estimating Social Network Models with Link Misclassification*.
— (June 2024b). "Ignoring Measurement Errors in Social Networks". In: *The Econometrics Journal* 27.2, pp. 171–187. DOI: 10.1093/ectj/utad028.
Li, Wei, Nilanjan Chakraborty, and Robert Lunde (Sept. 2025). *Assumption-Lean Inference for Network-linked Data*. DOI: 10.48550/arXiv.2510.00287. arXiv: 2510.00287 [stat].
Li, Wenrui, Daniel L. Sussman, and Eric D. Kolaczyk (Aug. 2022). *Causal Inference under Network Interference with Noise*. DOI: 10.48550/arXiv.2105.04518. arXiv: 2105.04518 [stat].
Lin, Xu and Lung-Fei Lee (July 2010). "GMM Estimation of Spatial Autoregressive Models with Unknown Heteroskedasticity". In: *Journal of Econometrics* 157.1, pp. 34–52. DOI: 10.1016/j.jeconom.2009.10.035.
Lyzinski, Vince, Minh Tang, Avanti Athreya, Youngser Park, and Carey E. Priebe (Jan. 2017). "Community Detection and Classification in Hierarchical Stochastic Blockmodels". In: *IEEE Transactions on Network Science and Engineering* 4.1, pp. 13–26. DOI: 10.1109/TNSE.2016.2634322.
Martellosio, Federico (June 2022). *Non-Identifiability in Network Autoregressions*. arXiv: 2011.11084 [econ, math, stat].
McFowland, Edward and Cosma Rohilla Shalizi (July 2021). "Estimating Causal Peer Influence in Homophilous Social Networks by Inferring Latent Locations". In: *Journal of the American Statistical Association* 0.0, pp. 1–12. DOI: 10.1080/01621459.2021.1953506.
Michell, L. and P. West (1996). "Peer Pressure to Smoke: The Meaning Depends on the Method". In: *Health Education Research* 11.1, pp. 39–49. DOI: 10.1093/her/11.1.39.
Michell, Lynn (1997). "Loud, Sad or Bad: Young People's Perceptions of Peer Groups and Smoking". In: *Health Education Research* 12.1, pp. 1–14. DOI: 10.1093/her/12.1.1-a.
Michell, Lynn and Amanda Amos (June 1997). "Girls, Pecking Order and Smoking". In: *Social Science & Medicine* 44.12, pp. 1861–1869. DOI: 10.1016/S0277-9536(96)00295-X.
Michell Lynn, Michael Pearson (Jan. 2000). "Smoke Rings: Social Network Analysis of Friendship Groups, Smoking and Drug-Taking". In: *Drugs: Education, Prevention and Policy* 7.1, pp. 21–37. DOI: 10.1080/dep.7.1.21.37.
Ord, Keith (Mar. 1975). "Estimation Methods for Models of Spatial Interaction". In: *Journal of the American Statistical Association* 70.349, pp. 120–126. DOI: 10.1080/01621459.1975.10480272.
Paul, Subhadeep, Shanjukta Nath, and Keith Warren (Aug. 2022). *Network Influence with Latent Homophily and Measurement Error*. arXiv: 2203.14223 [stat].
Rohe, Karl, Jun Tao, Xintian Han, and Norbert Binkiewicz (2018). "A Note on Quickly Sampling a Sparse Matrix with Low Rank Expectation". In: *Journal of Machine Learning Research* 19, pp. 1–13.
Rubin-Delanchy, Patrick, Joshua Cape, Minh Tang, and Carey E. Priebe (2022). "A Statistical Interpretation of Spectral Embedding: The Generalised Random Dot Product

Graph". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 84.4, pp. 1446–1473. DOI: 10.1111/rssb.12509.
Rue, Håvard (2005). *Gaussian Markov Random Fields: Theory and Applications*. Monographs on Statistics and Applied Probability 104. Boca Raton: Chapman & Hall/CRC.
Sävje, F (Feb. 2024). "Causal Inference with Misspecified Exposure Mappings: Separating Definitions and Assumptions". In: *Biometrika* 111.1, pp. 1–15. DOI: 10.1093/biomet/asad019.
Shakya, Holly B., Nicholas A. Christakis, and James H. Fowler (Jan. 2017). "An Exploratory Comparison of Name Generator Content: Data from Rural India". In: *Social Networks* 48, pp. 157–168. DOI: 10.1016/j.socnet.2016.08.008.
Shalizi, Cosma Rohilla and Andrew C. Thomas (May 2011). "Homophily and Contagion Are Generically Confounded in Observational Social Network Studies". In: *Sociological Methods & Research* 40.2, pp. 211–239. DOI: 10.1177/0049124111404820.
Smieszek, T., E. U. Burri, R. Scherzinger, and R. W. Scholz (Apr. 2012). "Collecting Close-Contact Social Mixing Data with Contact Diaries: Reporting Errors and Biases". In: *Epidemiology & Infection* 140.4, pp. 744–752. DOI: 10.1017/S0950268811001130.
Spohn, Meta-Lina, Leonard Henckel, and Marloes H. Maathuis (Dec. 2023). *A Graphical Approach to Treatment Effect Estimation with Observational Network Data*. arXiv: 2312.02717 [stat].
Su, Liangjun (Apr. 2012). "Semiparametric GMM Estimation of Spatial Autoregressive Models". In: *Journal of Econometrics* 167.2, pp. 543–560. DOI: 10.1016/j.jeconom.2011.09.034.
Sussman, Daniel L., Minh Tang, and Carey E. Priebe (Jan. 2014). "Consistent Latent Position Estimation and Vertex Classification for Random Dot Product Graphs". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 36.1, pp. 48–57. DOI: 10.1109/TPAMI.2013.135.
Sweet, Tracy and Samrachana Adhikari (June 2020). "A Latent Space Network Model for Social Influence". In: *Psychometrika* 85.2, pp. 251–274. DOI: 10.1007/s11336-020-09700-x.
Vazquez-Bare, Gonzalo (Nov. 2023). "Identification and Estimation of Spillover Effects in Randomized Experiments". In: *Journal of Econometrics* 237.1, p. 105237. DOI: 10.1016/j.jeconom.2021.10.014.
Vershynin, Roman (June 2020). *High-Dimensional Probability: An Introduction with Applications in Data Science*. 1st ed. Cambridge University Press. DOI: 10.1017/9781108231596.
Yu, Christina Lee, Edoardo M. Airoldi, Christian Borgs, and Jennifer T. Chayes (Nov. 2022). "Estimating the Total Treatment Effect in Randomized Experiments with Unknown Network Structure". In: *Proceedings of the National Academy of Sciences* 119.44, e2208975119. DOI: 10.1073/pnas.2208975119.
Yu, Y., T. Wang, and R. J. Samworth (June 2015). "A Useful Variant of the Davis–Kahan Theorem for Statisticians". In: *Biometrika* 102.2, pp. 315–323. DOI: 10.1093/biomet/asv008.
Zhang, Lina (Aug. 2024). *Spillovers of Program Benefits with Missing Network Links*. DOI: 10.48550/arXiv.2009.09614. arXiv: 2009.09614 [econ].