A Organization of the Appendix

The appendix includes the missing proofs, detailed discussions of some argument in the main body and more numerical experiments. We organize the appendix as follows:

- The proof of infeasibility condition (Theorem 3.2) is provided in Section B.
- Explanations on conditions derived in Theorem 3.2 are included in Section C.
- The proof of properties of the proposed model (r)LogSpecT (Proposition 3.4 & 3.6) is given in Section D and some additional properties are discussed.
- The truncated Hausdorff distance based proof details of Theorem 4.1 and Corollary 4.4 are given in Section E.
- Details of L-ADMM and its convergence analysis are in Section F.
- Additional experiments and discussions on synthetic data are included in Section G.

B Proof of Theorem 3.2

Since the linear system (4) has no solution, we know from Farkas' lemma that the following system has solutions:

$$\begin{cases} \begin{bmatrix} \boldsymbol{I}_{m-1} & \boldsymbol{0}_{\frac{(m-1)(m-2)}{2}} \end{bmatrix} \boldsymbol{B}^\top \boldsymbol{A}_n^\top \boldsymbol{x} < \boldsymbol{0}_{(m-1)\times 1}, \\ \begin{bmatrix} \boldsymbol{0}_{\frac{(m-1)(m-2)}{2}\times(m-1)} & \boldsymbol{I}_{\frac{(m-1)(m-2)}{2}} \end{bmatrix} \boldsymbol{B}^\top \boldsymbol{A}_n^\top \boldsymbol{x} \leq \boldsymbol{0}_{\frac{(m-1)(m-2)}{2}\times 1}. \end{cases} \tag{11}$$

Let $\boldsymbol{x}^* \in \mathbb{R}^{m^2}$ be a solution to (11). Denote $\boldsymbol{x}_+ := \max\{\boldsymbol{x}^*, \boldsymbol{0}\}$, $\boldsymbol{x}_- := \max\{-\boldsymbol{x}^*, \boldsymbol{0}\}$. Then, there exists $c \in (0, 1]$ such that

$$\boldsymbol{B}^\top \boldsymbol{A}_n^\top (\boldsymbol{x}_+ - \boldsymbol{x}_-) + c\boldsymbol{1}_{m^2}^\top (\boldsymbol{x}_+ + \boldsymbol{x}_-)[\boldsymbol{1}_{m-1}; \boldsymbol{0}_{\frac{(m-1)(m-2)}{2}}] \leq \boldsymbol{0}.$$

Define $y := -\boldsymbol{1}_{m^2}^\top(\boldsymbol{x}_+ + \boldsymbol{x}_-)$, $z := c\boldsymbol{1}_{m^2}^\top(\boldsymbol{x}_+ + \boldsymbol{x}_-)$ and set $\bar{\delta} = c$. For all $\delta \in [0, \bar{\delta})$, $(\boldsymbol{x}_+, \boldsymbol{x}_-, y, z)$ is a solution to the following linear system:

$$\begin{cases} \boldsymbol{B}^\top \boldsymbol{A}_n^\top (\boldsymbol{x}_+ - \boldsymbol{x}_-) + z[\boldsymbol{1}_{m-1}; \boldsymbol{0}_{\frac{(m-1)(m-2)}{2}}] \leq \boldsymbol{0}, \\ \boldsymbol{1}_{m^2}^\top (\boldsymbol{x}_+ + \boldsymbol{x}_-) + y \leq 0, \\ \delta y + z > 0, \\ \boldsymbol{x}_+, \boldsymbol{x}_-, -y \geq \boldsymbol{0}. \end{cases}$$

Again, from Farkas' lemma, this implies that the following linear system does not have a solution:

$$\begin{cases} \boldsymbol{A}_n \boldsymbol{B} \boldsymbol{s} + t\boldsymbol{1}_{m^2} \geq \boldsymbol{0}, \\ \boldsymbol{A}_n \boldsymbol{B} \boldsymbol{s} - t\boldsymbol{1}_{m^2} \leq \boldsymbol{0}, \\ t \leq \delta, \\ \begin{bmatrix} \boldsymbol{1}_{m-1} & \boldsymbol{0}_{\frac{(m-1)(m-2)}{2}} \end{bmatrix} \boldsymbol{s} = 1, \end{cases} \tag{12}$$

where $\boldsymbol{s} \in \mathbb{R}^{m(m-1)/2}$ and $t \in \mathbb{R}$. Since (12) is equivalent to:

$$\begin{cases} \|\boldsymbol{C}_n \boldsymbol{S} - \boldsymbol{S}\boldsymbol{C}_n\|_{\infty,\infty} \leq \delta, \\ (\boldsymbol{S}\boldsymbol{1})_1 = 1, \\ \boldsymbol{S} \in \mathcal{S}, \end{cases} \tag{13}$$

the above argument indicates that (13) does not have a solution. Suppose rSpecT has a feasible solution $\boldsymbol{S}'$, then

$$\|\boldsymbol{C}_n \boldsymbol{S}' - \boldsymbol{S}'\boldsymbol{C}_n\|_{\infty,\infty} \leq \|\boldsymbol{C}_n \boldsymbol{S}' - \boldsymbol{S}'\boldsymbol{C}_n\|_F \leq \delta.$$

Hence, $\boldsymbol{S}'$ is also a solution to (13). However, (13) does not have a solution. We can conclude that rSpecT is infeasible in this case.

C Explanations on Sufficient Conditions in Theorem 3.2

We elaborate more on the infeasibility condition that$\boldsymbol{A}_n\boldsymbol{B}$ has full column rank. An application of the condition is Example 3.1. Specifically, we know that in this case,

$$\boldsymbol{B} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{A}_n = \begin{pmatrix} 0 & h_{12} & -h_{12} & 0 \\ h_{12} & h_{22}-h_{11} & 0 & -h_{12} \\ -h_{12} & 0 & h_{11}-h_{22} & h_{12} \\ 0 & -h_{12} & h_{12} & 0 \end{pmatrix}.$$

This implies that

$$\boldsymbol{A}_n\boldsymbol{B} = \begin{pmatrix} 0 \\ h_{22}-h_{11} \\ h_{11}-h_{22} \\ 0 \end{pmatrix}.$$

Hence, when $h_{11} \neq h_{22}$, $\boldsymbol{A}_n\boldsymbol{B}$ has full column rank. This means that when δ is small enough (from Example 3.1 we know $\tilde{\delta} = \sqrt{2}|h_{11} - h_{22}|$), the model rSpecT is infeasible.

D Proofs of Properties of (r)LogSpecT

D.1 Proof of Proposition 3.4

Since the constraint set $\mathcal{S}$ is a cone, it follows that for all $\gamma > 0$, $\gamma\mathcal{S} = \mathcal{S}$. Then, we know that

$$\begin{aligned}
\mathrm{Opt}(\boldsymbol{C}, \alpha) &= \underset{\boldsymbol{S}\in\mathcal{S}, \boldsymbol{C}\boldsymbol{S}=\boldsymbol{S}\boldsymbol{C}}{\operatorname{argmin}} \; \|\boldsymbol{S}\|_{1,1} - \alpha\boldsymbol{1}^\top \log(\boldsymbol{S}\boldsymbol{1}) \\
&= \gamma \cdot \underset{\gamma\boldsymbol{S}\in\mathcal{S}, \boldsymbol{C}\gamma\boldsymbol{S}=\gamma\boldsymbol{S}\boldsymbol{C}}{\operatorname{argmin}} \; \|\gamma\boldsymbol{S}\|_{1,1} - \alpha\boldsymbol{1}^\top \log(\gamma\boldsymbol{S}\boldsymbol{1}) \\
&= \gamma \cdot \underset{\boldsymbol{S}\in\frac{1}{\gamma}\mathcal{S}, \boldsymbol{C}\boldsymbol{S}=\boldsymbol{S}\boldsymbol{C}}{\operatorname{argmin}} \; \gamma\|\boldsymbol{S}\|_{1,1} - \alpha\boldsymbol{1}^\top \log(\boldsymbol{S}\boldsymbol{1}) \\
&= \gamma \cdot \underset{\boldsymbol{S}\in\mathcal{S}, \boldsymbol{C}\boldsymbol{S}=\boldsymbol{S}\boldsymbol{C}}{\operatorname{argmin}} \; \|\boldsymbol{S}\|_{1,1} - \frac{\alpha}{\gamma}\boldsymbol{1}^\top \log(\boldsymbol{S}\boldsymbol{1}) \\
&= \gamma\,\mathrm{Opt}(\boldsymbol{C}, \alpha/\gamma),
\end{aligned}$$

where the third equality is from the basic calculus rule of the logarithm function. Set $\gamma = \alpha$ and then $\mathrm{Opt}(\boldsymbol{C}, \alpha) = \alpha\,\mathrm{Opt}(\boldsymbol{C}, 1)$, which completes the proof.

D.2 Proof of Proposition 3.6

The proof will be conducted by constructing a feasible solution for rLogSpecT. Recall that $\boldsymbol{A}_n = \boldsymbol{I} \otimes \boldsymbol{C}_n - \boldsymbol{C}_n \otimes \boldsymbol{I}$ and the matrix $\boldsymbol{B} \in \mathbb{R}^{m^2\times m(m-1)/2}$ that maps a non-negative vector to the vectorization of a valid adjacency matrix. Let $\boldsymbol{S} = \min\{\frac{\delta}{\|\boldsymbol{A}_n\boldsymbol{B}\boldsymbol{s}\|_2}, 1\}\cdot\mathrm{mat}(\boldsymbol{B}\boldsymbol{s})$ with $\boldsymbol{s} \in \mathbb{R}^{(m-1)m/2}$ being a non-negative vector, where $\mathrm{mat}(\cdot)$ is the matricization operator. Note that

$$\mathrm{vec}(\boldsymbol{C}_n\boldsymbol{S} - \boldsymbol{S}\boldsymbol{C}_n) = (\boldsymbol{I} \otimes \boldsymbol{C}_n - \boldsymbol{C}_n \otimes \boldsymbol{I})\,\mathrm{vec}(\boldsymbol{S}) = \boldsymbol{A}_n\mathrm{vec}(\boldsymbol{S}).$$

Then, we know that

$$\|\boldsymbol{C}_n\boldsymbol{S} - \boldsymbol{S}\boldsymbol{C}_n\|_F = \|\mathrm{vec}(\boldsymbol{C}_n\boldsymbol{S} - \boldsymbol{S}\boldsymbol{C}_n)\|_2 = \min\left\{\frac{\delta}{\|\boldsymbol{A}_n\boldsymbol{B}\boldsymbol{s}\|_2}, 1\right\} \cdot \|\boldsymbol{A}_n\boldsymbol{B}\boldsymbol{s}\|_2 \leq \delta.$$

Thus, the given $\boldsymbol{S}$ is a feasible solution for rLogSpecT and it completes the proof.

D.3 Properties of optimal solutions and values of (r)LogSpecT

In this section, we further discuss some properties of the optimal solutions/value of the proposed models, which are useful for deriving the recovery guarantee. More specifically, we obtain an upper bound on the optimal solutions (which may not be unique) independent of the sample size n and the inaccuracy parameter δ_n. Also, a lower bound of optimal values follows.

Proposition D.1. *The following statements hold:*

- *For an optimal solution $\boldsymbol{S}^*$ (resp. $\boldsymbol{S}_n^*$) to LogSpecT (resp. rLogSpecT with any given sample size n), it follows that*

$$\|\boldsymbol{S}^*\|_{1,1} = \alpha m \ \text{ and } \ \|\boldsymbol{S}_n^*\|_{1,1} \leq \alpha m, \ \ \forall \delta_n > 0.$$

- *If $\delta_n \geq 2\alpha m \|\boldsymbol{C}_n - \boldsymbol{C}_\infty\|$, then*

$$\alpha m(1 - \log \alpha) \leq f_n^* \leq f^*, \ \ \forall n,$$

where f^* (resp. f_n^*) denotes the optimal value of LogSpecT (resp. rLogSpecT).

For the first statement, let us consider the Karush-Kuhn-Tucker (KKT) conditions of LogSpecT and rLogSpecT. Since the LogSpecT is a convex problem and Slater's condition holds, the KKT conditions are necessary and sufficient for the optimality, i.e., there exists $(\boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2) \in \mathbb{R}^{m \times m} \times \mathcal{N}_{\mathcal{S}}(\boldsymbol{S}^*)$ such that

$$\begin{cases} \nabla_{\boldsymbol{S}}(\|\boldsymbol{S}^*\|_{1,1} - \alpha \mathbf{1}^\top \log(\boldsymbol{S}^*\mathbf{1})) + \boldsymbol{C}_\infty \boldsymbol{\Lambda}_1 - \boldsymbol{\Lambda}_1 \boldsymbol{C}_\infty + \boldsymbol{\Lambda}_2 = \mathbf{0}, \\ \boldsymbol{C}_\infty \boldsymbol{S}^* = \boldsymbol{S}^* \boldsymbol{C}_\infty, \\ \boldsymbol{S}^* \in \mathcal{S}, \end{cases} \tag{14}$$

where $\mathcal{N}_{\mathcal{S}}(\boldsymbol{S}^*) \coloneqq \{\boldsymbol{N} \in \mathbb{R}^{m \times m} : \sup_{\boldsymbol{X} \in \mathcal{S}} \langle \boldsymbol{X} - \boldsymbol{S}^*, \boldsymbol{N} \rangle \leq 0\}$ is the normal cone of $\mathcal{S}$ at $\boldsymbol{S}^*$, and $\nabla \|\boldsymbol{S}^*\|_{1,1}$ is well-defined since $\|\cdot\|_{1,1} = \langle \cdot, \mathbf{1}\mathbf{1}^\top \rangle$ at $\boldsymbol{S}^* \geq 0$, which is differentiable. Taking further calculation gives that

$$\nabla \|\boldsymbol{S}^*\|_{1,1} = \mathbf{1}\mathbf{1}^\top, \quad (\nabla_{\boldsymbol{S}} \mathbf{1}^\top \log(\boldsymbol{S}^*\mathbf{1}))_{ij} = \frac{1}{(\boldsymbol{S}^*\mathbf{1})_i}.$$

Combining this with (14) by taking inner product of both sides with $\boldsymbol{S}^*$, we obtain that

$$\sum_{i,j} (\boldsymbol{S}^*)_{ij} - \alpha \sum_{i,j} \frac{(\boldsymbol{S}^*)_{ij}}{(\boldsymbol{S}^*\mathbf{1})_i} + \langle \boldsymbol{\Lambda}_1, \boldsymbol{C}_\infty \boldsymbol{S}^* - \boldsymbol{S}^* \boldsymbol{C}_\infty \rangle + \langle \boldsymbol{\Lambda}_2, \boldsymbol{S}^* \rangle = 0. \tag{15}$$

From the structure of $\mathcal{S}$ and the fact that $\Lambda_2 \in \mathcal{N}_{\mathcal{S}}(\boldsymbol{S}^*)$, one has that $\langle \boldsymbol{\Lambda}_2, \boldsymbol{S}^* \rangle = 0$. Also, note that $\boldsymbol{C}_\infty \boldsymbol{S}^* = \boldsymbol{S}^* \boldsymbol{C}_\infty$. Hence, the equation (15) can be simplified as the desired result:

$$\|\boldsymbol{S}^*\|_{1,1} = \sum_{i,j} (\boldsymbol{S}^*)_{ij} = \alpha \sum_{i,j} \frac{(\boldsymbol{S}^*)_{ij}}{(\boldsymbol{S}^*\mathbf{1})_i} = \alpha \sum_{i=1}^m \sum_{j=1}^m \frac{(\boldsymbol{S}^*)_{ij}}{(\boldsymbol{S}^*\mathbf{1})_i} = \alpha m.$$

The KKT conditions of rLogSpecT indicate that there exist $\lambda_1 \geq 0$, $\boldsymbol{\Lambda}_2 \in \mathcal{N}_{\mathcal{S}}(\boldsymbol{S}_n^*)$ and $\boldsymbol{Q} \in \partial \|\boldsymbol{C}_n \boldsymbol{S}_n^* - \boldsymbol{S}_n^* \boldsymbol{C}_n\|_F$ (i.e., the subgradient of the function $\boldsymbol{S} \mapsto \|\boldsymbol{C}_n \boldsymbol{S} - \boldsymbol{S} \boldsymbol{C}_n\|_F$ at $\boldsymbol{S}_n^*$) such that

$$\begin{cases} \nabla_{\boldsymbol{S}}(\|\boldsymbol{S}_n^*\|_{1,1} - \alpha \mathbf{1}^\top \log(\boldsymbol{S}_n^*\mathbf{1})) + \lambda_1 \boldsymbol{Q} + \boldsymbol{\Lambda}_2 = \mathbf{0}, \\ \lambda_1(\|\boldsymbol{C}_n \boldsymbol{S}_n^* - \boldsymbol{S}_n^* \boldsymbol{C}_n\|_F - \delta_n) = 0, \\ \boldsymbol{S}_n^* \in \mathcal{S}. \end{cases} \tag{16}$$

Moreover, from the definition of the convex subdifferential we know that $0 \geq \|\boldsymbol{C}_n \boldsymbol{S}_n^* - \boldsymbol{S}_n^* \boldsymbol{C}_n\|_F - \langle \boldsymbol{Q}, \boldsymbol{S}_n^* \rangle$. Thus, after taking inner product of both sides of the equation (16) with $\boldsymbol{S}_n^*$, it follows that:

$$\begin{aligned} 0 &= \sum_{i,j} (\boldsymbol{S}_n^*)_{ij} - \alpha m + \lambda_1 \langle \boldsymbol{Q}, \boldsymbol{S}_n^* \rangle + \langle \boldsymbol{\Lambda}_2, \boldsymbol{S}_n^* \rangle \\ &\geq \sum_{i,j} (\boldsymbol{S}_n^*)_{ij} - \alpha m + \lambda_1 \|\boldsymbol{C}_n \boldsymbol{S}_n^* - \boldsymbol{S}_n^* \boldsymbol{C}_n\|_F + \langle \boldsymbol{\Lambda}_2, \boldsymbol{S}_n^* \rangle \\ &= \sum_{i,j} (\boldsymbol{S}_n^*)_{ij} - \alpha m + \lambda_1 \delta_n, \end{aligned}$$

which implies that $\sum_{i,j} (\boldsymbol{S}_n^*)_{ij} \leq \alpha m - \lambda_1 \delta_n \leq \alpha m$. This completes the proof of the first statement.

For the second statement, we first prove that v_n^* and v^* are larger than $\alpha m(1-\log\alpha)$. Define the auxiliary function $g:\mathbb{R}\to\mathbb{R}$ such that $g(x) \coloneqq x-\alpha\log x$ for any $x\in\mathbb{R}_+$, whose minimum is attained at α. Since for any $\boldsymbol{S}\in\mathcal{S}$,

$$f(\boldsymbol{S}) = \sum_{i=1}^{m} g\left(\sum_{j=1}^{m} S_{ij}\right),$$

where f is the objective in LogSpecT, it follows that

$$f(\boldsymbol{S}) \geq \sum_{i=1}^{m} g(\alpha) = \alpha m(1-\log\alpha).$$

This implies that v_n^* and v^* are larger than $\alpha m(1-\log\alpha)$. Next, we will show $v_n^* \leq v^*$. Consider any optimal solution $\boldsymbol{S}^*$ to LogSpecT. We show that it is feasible for rLogSpecT.

$$\begin{aligned}\|\boldsymbol{C}_n\boldsymbol{S}^* - \boldsymbol{S}^*\boldsymbol{C}_n\| &= \|\boldsymbol{C}_n\boldsymbol{S}^* - \boldsymbol{C}_\infty\boldsymbol{S}^* + \boldsymbol{S}^*\boldsymbol{C}_\infty - \boldsymbol{S}^*\boldsymbol{C}_n\| \\ &\leq 2\|\boldsymbol{S}^*\|_{1,1}\|\boldsymbol{C}_n - \boldsymbol{C}_\infty\| \leq 2\alpha m\|\boldsymbol{C}_n - \boldsymbol{C}_\infty\| \leq \delta_n,\end{aligned}$$

where the equality comes from $\boldsymbol{C}_\infty\boldsymbol{S}^* = \boldsymbol{S}^*\boldsymbol{C}_\infty$, the first inequality comes from the fact that $\|\boldsymbol{X}\boldsymbol{Y}\| \leq \|\boldsymbol{X}\|_F\|\boldsymbol{Y}\| \leq \|\boldsymbol{X}\|_{1,1}\|\boldsymbol{Y}\|$, the second one comes from the first statement and the last one is due to $\delta_n \geq 2\alpha m\|\boldsymbol{C}_n - \boldsymbol{C}_\infty\|$. Hence, $\boldsymbol{S}^*$ is feasible for rLogSpecT, which indicates that $v_n^* \leq v^*$. The proof is completed.

E Proof of Theorem 4.1 & Corollary 4.4

E.1 Truncated Hausdorff distance

In this section, we introduce an advanced technique in optimization that is efficient in analyzing the recovery guarantee of robust formulations. Before that, we introduce the concept of truncated Hausdorff distance between two sets.

Definition E.1 (Truncated Hausdorff Distance [28, 6.J])**.** *For any $\rho \geq 0$, the truncated Hausdorff distance between two sets $\mathcal{C}$ and $\mathcal{D}$ is defined as:*

$$\hat{\mathrm{d}}_\rho(\mathcal{C},\mathcal{D}) = \max\{\mathrm{dist}(\mathcal{C}\cap\mathbb{B}(\mathbf{0},\rho),\mathcal{D}), \mathrm{dist}(\mathcal{D}\cap\mathbb{B}(\mathbf{0},\rho),\mathcal{C})\}.$$

It turns out that the distance between the optimum of two minimization problems can be bounded with the truncated Hausdorff distance of the epigraphs under some conditions. The result is captured in the following lemma.

Lemma E.2 ([28, Theorem 6.56])**.** *Let $\rho\in[0,\infty)$. Suppose that the extended-real-valued functions $f,g:\mathbb{R}^n\to\overline{\mathbb{R}}$ satisfy*

- $\inf f, \inf g \in [-\rho,\rho]$,
- $\operatorname{argmin} f, \operatorname{argmin} g \subseteq \mathbb{B}(\mathbf{0},\rho)$.

Then, it follows that

$$|\inf f - \inf g| \leq \hat{\mathrm{d}}_\rho(\operatorname{epi} f, \operatorname{epi} g).^{3} \tag{17}$$

Suppose further that $\varepsilon > 2\,\hat{\mathrm{d}}_\rho(\operatorname{epi} f,\operatorname{epi} g)$, then one has

$$\mathrm{dist}(\boldsymbol{x}_g^*, \varepsilon\text{-}\operatorname{argmin} f) \leq \hat{\mathrm{d}}_\rho(\operatorname{epi} f, \operatorname{epi} g), \tag{18}$$

where ε-argmin f is the ε-suboptimal solution set of f that is defined as ε-argmin $f \coloneqq \{\boldsymbol{x}\in\mathbb{R}^n : f(\boldsymbol{x}) \leq \inf f + \varepsilon\}$, and $\boldsymbol{x}_g^*$ is a minimizer of g.

From the above lemma, we know that if two optimization problems are close enough (in the sense of truncated Hausdorff distance), then the optimum of them should be close to each other. Hence, in order to apply this result, we need to bound the truncated Hausdorff distance in an explicit way, which is solved by the following Kenmochi condition.

[3] For a function $f:\mathbb{R}^n\to\overline{\mathbb{R}}$, its epigraph is defined as $\operatorname{epi} f \coloneqq \{(\boldsymbol{x},y) \mid y \geq f(\boldsymbol{x})\}$.

Lemma E.3 (Kenmochi Condition [28, Proposition 6.58]). *Let $\rho \in [0,\infty)$. Then, for $f,g:\mathbb{R}^n \to \overline{\mathbb{R}}$ with nonempty epigraphs, one has that*

$$\hat{\mathrm{d}}_\rho(\operatorname{epi} f, \operatorname{epi} g) = \inf \left\{ \eta > 0 : \begin{array}{l} \inf\limits_{\mathbb{B}(\boldsymbol{x},\eta)} g \le \max\{f(\boldsymbol{x}), -\rho\} + \eta,\ \forall \boldsymbol{x} \in [f \le \rho] \cap \mathbb{B}(\boldsymbol{0},\rho) \\ \inf\limits_{\mathbb{B}(\boldsymbol{x},\eta)} f \le \max\{g(\boldsymbol{x}), -\rho\} + \eta,\ \forall \boldsymbol{x} \in [g \le \rho] \cap \mathbb{B}(\boldsymbol{0},\rho) \end{array} \right\},$$

where $[f \le \rho] \coloneqq \{\boldsymbol{x} \in \mathbb{R}^n : f(\boldsymbol{x}) \le \rho\}$.

E.2 Proof of Theorem 4.1

Before presenting the proof, we first introduce the following lemma.

Lemma E.4 (Hoffman's Error Bound [11]). *Consider the set $\mathcal{S} \coloneqq \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}\}$. There exists $C > 0$ such that for any $\boldsymbol{x} \in \mathbb{R}^n$, one has*

$$\operatorname{dist}(\boldsymbol{x}, \mathcal{S}) \le C \cdot \|(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b})_+\|_2.$$

For the sake of brevity, we denote

$$\bar{f}_n(\boldsymbol{S}) \coloneqq \|\boldsymbol{S}\|_{1,1} - \alpha \boldsymbol{1}^\top \log(\boldsymbol{S}\boldsymbol{1}) + \iota_{\mathbb{R}_-}(\|\boldsymbol{C}_n\boldsymbol{S} - \boldsymbol{S}\boldsymbol{C}_n\|_F - \delta_n) + \iota_{\mathcal{S}}(\boldsymbol{S}),$$
$$\bar{f}(\boldsymbol{S}) \coloneqq \|\boldsymbol{S}\|_{1,1} - \alpha \boldsymbol{1}^\top \log(\boldsymbol{S}\boldsymbol{1}) + \iota_{\{0\}}(\|\boldsymbol{C}_\infty\boldsymbol{S} - \boldsymbol{S}\boldsymbol{C}_\infty\|_F) + \iota_{\mathcal{S}}(\boldsymbol{S}).$$

Hence, the optimization problem LogSpecT (resp. rLogSpecT) is equivalent to $\inf \bar{f}$ (resp. $\inf \bar{f}_n$).

Now, we aim to use Lemma E.3 to bound $\hat{\mathrm{d}}_\rho(\operatorname{epi}\bar{f}, \operatorname{epi}\bar{f}_n)$. Let $\boldsymbol{S} \in \mathcal{S} \cap \mathbb{B}(\boldsymbol{0},\rho)$ satisfy

$$\bar{f}(\boldsymbol{S}) \le \rho \quad \text{and} \quad \boldsymbol{S}\boldsymbol{C}_\infty = \boldsymbol{C}_\infty\boldsymbol{S}.$$

Then, we know that

$$\|\boldsymbol{S}\boldsymbol{C}_n - \boldsymbol{C}_n\boldsymbol{S}\|_F \le 2\|\boldsymbol{S}\|_F\|\boldsymbol{C}_n - \boldsymbol{C}_\infty\| \le 2\rho\|\boldsymbol{C}_n - \boldsymbol{C}_\infty\| \le \delta_n,$$

and consequently $\boldsymbol{S}$ is in the domain of $\bar{f}_n$. Then, it follows that for any $\eta > 0$, we have

$$\inf_{\mathbb{B}(\boldsymbol{S},\eta)} \bar{f}_n \le \bar{f}_n(\boldsymbol{S}) = \bar{f}(\boldsymbol{S}) \le \max\{\bar{f}(\boldsymbol{S}), -\rho\}, \quad \forall \boldsymbol{S} \in [\bar{f} \le \rho] \cap \mathbb{B}(\boldsymbol{0},\rho). \tag{19}$$

Before verifying the reverse side of the Kenmochi condition, we first consider the non-emptiness of $[\bar{f}_n \le \rho] \cap \mathbb{B}(\boldsymbol{0},\rho)$. Since

$$\delta_n \ge 2\rho\|\boldsymbol{C}_n - \boldsymbol{C}_\infty\| \ge 2\alpha m\|\boldsymbol{C}_n - \boldsymbol{C}_\infty\|,$$

it follows from Proposition D.1 that $\|\boldsymbol{S}_n^*\|_{1,1} \le \alpha m \le \rho$ and $f_n^* \le f^* \le \rho$, which implies that $[\bar{f}_n \le \rho] \cap \mathbb{B}(\boldsymbol{0},\rho)$ is nonempty. Let $\boldsymbol{S}_n \in [\bar{f}_n \le \rho] \cap \mathbb{B}(\boldsymbol{0},\rho)$. Then, one has that

$$\boldsymbol{S}_n \in \mathcal{S} \quad \text{and} \quad \|\boldsymbol{C}_n\boldsymbol{S}_n - \boldsymbol{S}_n\boldsymbol{C}_n\|_F \le \delta_n.$$

Hence, it follows that

$$\|\boldsymbol{C}_\infty\boldsymbol{S}_n - \boldsymbol{S}_n\boldsymbol{C}_\infty\| \le 2\|\boldsymbol{S}_n\|_F\|\boldsymbol{C}_\infty - \boldsymbol{C}_n\| + \|\boldsymbol{C}_n\boldsymbol{S}_n - \boldsymbol{S}_n\boldsymbol{C}_n\|_F \le 2\rho\|\boldsymbol{C}_\infty - \boldsymbol{C}_n\| + \delta_n.$$

Also, note that there exists $\beta > 0$ such that $(\boldsymbol{S}_n\boldsymbol{1})_i \ge \beta$ for all $i \in [m]$ as $\bar{f}_n \le \rho$ and $\|\boldsymbol{S}_n\|_{1,1} - \alpha\boldsymbol{1}^\top \log(\boldsymbol{S}_n\boldsymbol{1}) \to \infty$ when $\boldsymbol{S}_n \to \boldsymbol{0}$. Thus, applying Lemma E.4 to the linear system

$$\tilde{\mathcal{S}} \coloneqq \{\boldsymbol{S} \in \mathbb{R}^{m\times m} : \boldsymbol{S}\boldsymbol{C}_\infty = \boldsymbol{C}_\infty\boldsymbol{S},\ \boldsymbol{S} \in \mathcal{S},\ (\boldsymbol{S}\boldsymbol{1})_i \ge \beta,\ \forall i \in [m]\}$$

yields that there exists $\tilde{c} > 0$ such that

$$\operatorname{dist}(\boldsymbol{S}_n, \tilde{\mathcal{S}}) \le \tilde{c} \cdot (2\rho\|\boldsymbol{C}_\infty - \boldsymbol{C}_n\| + \delta_n).$$

Hence, there exists $\tilde{\boldsymbol{S}}$ in the domain of $\bar{f}$ such that

$$\|\boldsymbol{S}_n - \tilde{\boldsymbol{S}}\|_F \le \tilde{c} \cdot (2\rho\|\boldsymbol{C}_\infty - \boldsymbol{C}_n\| + \delta_n) \quad \text{and} \quad (\tilde{\boldsymbol{S}}\boldsymbol{1})_i \ge \beta,\ \forall i \in [m].$$

Since the function $\boldsymbol{S} \mapsto \|\boldsymbol{S}\|_{1,1} - \alpha \mathbf{1}^\top \log(\boldsymbol{S}\mathbf{1})$ is locally Lipschitz continuous when $(\boldsymbol{S}\mathbf{1})_i \geq \beta$, there exists $L > 0$ such that

$$\begin{aligned}\bar{f}(\tilde{\boldsymbol{S}}) = \|\tilde{\boldsymbol{S}}\|_{1,1} - \alpha \mathbf{1}^\top \log(\tilde{\boldsymbol{S}}\mathbf{1}) &\leq \|\boldsymbol{S}_n\|_{1,1} - \alpha \mathbf{1}^\top \log(\boldsymbol{S}_n\mathbf{1}) + L\|\boldsymbol{S}_n - \tilde{\boldsymbol{S}}\|_F \\ &= \bar{f}_n(\boldsymbol{S}_n) + L\|\boldsymbol{S}_n - \tilde{\boldsymbol{S}}\|_F \\ &\leq \bar{f}_n(\boldsymbol{S}_n) + L\tilde{c} \cdot (2\rho\|\boldsymbol{C}_\infty - \boldsymbol{C}_n\| + \delta_n).\end{aligned}$$

Setting $c_1 \geq \max\{1, L\} \cdot \tilde{c}$, one can obtain that for any $\boldsymbol{S}_n \in [\bar{f}_n \leq \rho] \cap \mathbb{B}(\mathbf{0}, \rho)$

$$\inf_{\mathbb{B}(\boldsymbol{S}_n,\eta)} \bar{f} \leq \bar{f}(\tilde{\boldsymbol{S}}) \leq \bar{f}_n(\boldsymbol{S}_n) + c_1 \cdot (2\rho\|\boldsymbol{C}_\infty - \boldsymbol{C}_n\| + \delta_n) \leq \max\{\bar{f}_n(\boldsymbol{S}_n), -\rho\} + \eta, \tag{20}$$

where $\eta \coloneqq c_1 \cdot (2\rho\|\boldsymbol{C}_\infty - \boldsymbol{C}_n\| + \delta_n)$. Combining inequality (19) and (20), we can conclude that

$$\hat{\mathrm{d}}_\rho(\mathrm{epi}\, \bar{f}, \mathrm{epi}\, \bar{f}_n) \leq c_1 \cdot (2\rho\|\boldsymbol{C}_\infty - \boldsymbol{C}_n\| + \delta_n). \tag{21}$$

In order to derive the conclusion (i) and (ii), it remains to check the requirements in Lemma E.2. Since $\rho \geq \alpha m$, the first statement of Proposition D.1 shows that the optimal solutions to $\inf \bar{f}$ and $\inf \bar{f}_n$ lie in $\mathbb{B}(\mathbf{0}, \rho)$. Since $\rho \geq f^*$ and $-\rho \leq \alpha m(1 - \log \alpha)$, the second statement of the proposition shows that $\inf \bar{f}, \inf \bar{f}_n \in [-\rho, \rho]$. Hence, applying Lemma E.2 completes the proof of the first two statements.

To prove conclusion (iii), we first make the following two claims:

(a) $\mathcal{S}_0^*\mathbf{1}$ is a singleton, whose element is denoted by $\boldsymbol{S}^*\mathbf{1}$,

(b) For any $\bar{\varepsilon} \in [0, \infty)$, there exists a $\delta(\bar{\varepsilon}) > 0$ such that for all $0 \leq \varepsilon \leq \bar{\varepsilon}$ and $\boldsymbol{S}_\varepsilon \in \mathcal{S}_\varepsilon^*$, one has that

$$\|\boldsymbol{S}_\varepsilon\mathbf{1} - \boldsymbol{S}^*\mathbf{1}\|_2 \leq \delta(\bar{\varepsilon}) \cdot \sqrt{\varepsilon}. \tag{22}$$

Granting these and with the help of Theorem 4.1, we can derive that for all $\boldsymbol{S}_n^* \in \mathcal{S}^{n,*}$

$$\begin{aligned}\mathrm{dist}(\boldsymbol{S}_n^*\mathbf{1}, \mathcal{S}_0^*\mathbf{1}) = \|\boldsymbol{S}_n^*\mathbf{1} - \boldsymbol{S}^*\mathbf{1}\|_2 &\leq \|\boldsymbol{S}_n^*\mathbf{1} - \boldsymbol{S}_{2\varepsilon_n}\mathbf{1}\|_2 + \|\boldsymbol{S}_{2\varepsilon_n}\mathbf{1} - \boldsymbol{S}^*\mathbf{1}\|_2 \\ &\leq \sqrt{m}\, \mathrm{dist}(\boldsymbol{S}_n^*, \mathcal{S}_{2\varepsilon_n}^*) + \|\boldsymbol{S}_{2\varepsilon_n}\mathbf{1} - \boldsymbol{S}^*\mathbf{1}\|_2 \\ &\leq \tilde{c}_1\varepsilon_n + \tilde{c}_2\sqrt{\varepsilon_n},\end{aligned}$$

where $\tilde{c}_1$, $\tilde{c}_2$ are positive constants, and $\boldsymbol{S}_{2\varepsilon_n} \in \mathcal{S}_{2\varepsilon_n}^*$ satisfies $\|\boldsymbol{S}_n^* - \boldsymbol{S}_{2\varepsilon_n}\|_F = \mathrm{dist}(\boldsymbol{S}_n^*, \mathcal{S}_{2\varepsilon_n}^*)$ (whose existence is guaranteed since $\mathcal{S}_\varepsilon^*$ is convex and compact). Hence,

$$\mathrm{dist}(\mathcal{S}^{n,*}\mathbf{1}, \mathcal{S}_0^*\mathbf{1}) \leq \tilde{c}_1\varepsilon_n + \tilde{c}_2\sqrt{\varepsilon_n}.$$

To proceed, it remains to prove the claims. Define an auxiliary function $h : \mathbb{R}^m \to \mathbb{R}$ as $h(\boldsymbol{x}) = \sum_{i=1}^m x_i - \alpha \sum_{i=1}^m \log x_i$ for each $\boldsymbol{x} \in \mathbb{R}_+^m$. Consider the following optimization problem:

$$\begin{aligned}&\min_{\boldsymbol{x}}\ h(\boldsymbol{x}) \\ &\text{s.t.}\ \ \boldsymbol{x} \in \{\boldsymbol{S}\mathbf{1} \in \mathbb{R}^m \mid \boldsymbol{S} \text{ that is feasible for LogSpecT}\}.\end{aligned} \tag{23}$$

For the sake of brevity, denote the ε-suboptimal solution set of (23) as $\mathcal{H}_\varepsilon^*$. In the remaining part, we will first show that $\mathcal{S}_\varepsilon^*\mathbf{1} = \mathcal{H}_\varepsilon^*$ and then, by the strict convexity of h, the desired two claims hold.

The first step is to show that the optimal function value of the problem (23) satisfies $h^* = f^*$. Since it is obvious that $\tilde{\boldsymbol{x}} = \boldsymbol{S}^*\mathbf{1}$ is feasible for (23), $h^* \leq h(\tilde{\boldsymbol{x}}) = f(\boldsymbol{S}^*) = f^*$. Suppose to the contrary that $h^* < f^*$, from the fact that the objective function is coercive and continuous and the feasible set is closed, there exists $\tilde{\boldsymbol{S}}$ such that it is feasible for LogSpecT and $\boldsymbol{x}^* = \tilde{\boldsymbol{S}}\mathbf{1}$, where $\boldsymbol{x}^*$ is an optimal solution to (23). Since $h^* = h(\boldsymbol{x}^*) = h(\tilde{\boldsymbol{S}}\mathbf{1}) = f(\tilde{\boldsymbol{S}})$, this contradicts the fact that $f(\tilde{\boldsymbol{S}}) \geq f^*$. Hence, $h^* = f^*$. Next, we will show that $\mathcal{S}_\varepsilon^*\mathbf{1} = \mathcal{H}_\varepsilon^*$. Consider any ε-suboptimal solution $\boldsymbol{S} \in \mathcal{S}_\varepsilon^*$, i.e.,

$$h(\boldsymbol{S}\mathbf{1}) = f(\boldsymbol{S}) \leq f^* + \varepsilon = h^* + \varepsilon.$$

Hence, $\boldsymbol{S}\mathbf{1} \in \mathcal{H}_\varepsilon^*$ and it implies that $\mathcal{S}_\varepsilon^*\mathbf{1} \subseteq \mathcal{H}_\varepsilon^*$. On the other hand, for any ε-suboptimal solution $\boldsymbol{x} \in \mathcal{H}_\varepsilon^*$, there exists $\boldsymbol{S}$ that is feasible for LogSpecT such that $\boldsymbol{x} = \boldsymbol{S}\mathbf{1}$. Thus,

$$f(\boldsymbol{S}) = h(\boldsymbol{x}) \leq h^* + \varepsilon = f^* + \varepsilon.$$

This implies that $\boldsymbol{S} \in \mathcal{S}_\varepsilon^*$ and consequently $\mathcal{H}_\varepsilon^* \subseteq \mathcal{S}_\varepsilon^* \mathbf{1}$. Hence, $\mathcal{H}_\varepsilon^* = \mathcal{S}_\varepsilon^* \mathbf{1}$.

Since h is strictly convex, its optimal solution set $\mathcal{H}_0^*$ is a singleton. Then, $\mathcal{S}_0^*\mathbf{1} = \mathcal{H}_0^*$ is a singleton, which proves the first claim. For the second claim, we know that for any $\boldsymbol{S}_\varepsilon \in \mathcal{S}_\varepsilon^*$ there exists $\boldsymbol{x}_\varepsilon \in \mathcal{H}_\varepsilon^*$ such that

$$\|\boldsymbol{S}_\varepsilon \mathbf{1} - \boldsymbol{S}^*\mathbf{1}\|_2 = \|\boldsymbol{x}_\varepsilon - \boldsymbol{x}^*\|_2, \tag{24}$$

where $\boldsymbol{x}^* \in \mathcal{H}_0^*$. The coerciveness of h asserts that $\boldsymbol{x}_\varepsilon$ and $\boldsymbol{x}^*$ are bounded. This together with the fact that h is strongly convex on any bounded set, illustrates that there exists $\mu > 0$ such that

$$h(\boldsymbol{x}_\varepsilon) \geq h(\boldsymbol{x}^*) + \langle \nabla h(\boldsymbol{x}^*), \boldsymbol{x}_\varepsilon - \boldsymbol{x}^* \rangle + \frac{1}{\mu}\|\boldsymbol{x}_\varepsilon - \boldsymbol{x}^*\|_2^2 \geq h(\boldsymbol{x}^*) + \frac{1}{\mu}\|\boldsymbol{x}_\varepsilon - \boldsymbol{x}^*\|_2^2, \tag{25}$$

where the second inequality comes from the global optimality of $\boldsymbol{x}^*$. Combining (24) and (25) gives that

$$\|\boldsymbol{S}_\varepsilon \mathbf{1} - \boldsymbol{S}^*\mathbf{1}\|_2 = \|\boldsymbol{x}_\varepsilon - \boldsymbol{x}^*\|_2 \leq \sqrt{\mu(h(\boldsymbol{x}_\varepsilon) - h(\boldsymbol{x}^*))} \leq \sqrt{\mu\varepsilon}.$$

This completes the proof of the claims.

E.3 Proof of Corollary 4.4

Suppose to the contrary that there exists a sequence $\{\boldsymbol{S}_n^*\}_n$, where the nth element is an optimal solution to rLogSpecT with sample size n, such that

$$\operatorname{dist}(\boldsymbol{S}_n^*, \mathcal{S}_0^*) \not\to 0.$$

From Proposition D.1, we know that $\{\boldsymbol{S}_n^*\}_n$ is bounded, and consequently, has a convergent subsequence. Without loss of generality, we may assume that the sequence itself is convergent and the limiting point is $\boldsymbol{S}^*$. Note that

$$\|\boldsymbol{C}_n \boldsymbol{S}_n^* - \boldsymbol{S}_n^* \boldsymbol{C}_n\|_F \leq \delta_n, \;\; \boldsymbol{C}_n \to \boldsymbol{C}_\infty \;\text{ and }\; \delta_n \to 0.$$

Hence, $\boldsymbol{C}_\infty \boldsymbol{S}^* = \boldsymbol{S}^* \boldsymbol{C}_\infty$. This indicates that $\boldsymbol{S}^*$ is feasible for LogSpecT. Then, from Theorem 4.1, we know that $f(\boldsymbol{S}_n^*) = f_n^* \to f^*$, which leads to $f(\boldsymbol{S}^*) = f^*$ since $f(\cdot) = \|\cdot\|_{1,1} - \alpha \mathbf{1}^\top \log(\cdot\mathbf{1})$ is continuous. Together with the fact that $\boldsymbol{S}^*$ is feasible, we conclude that $\boldsymbol{S}^*$ is an optimal solution to LogSpecT. This further implies that $\operatorname{dist}(\boldsymbol{S}_n^*, \mathcal{S}_0^*) \to 0$, which is a contradiction.

E.4 Proof of Lemma 4.7

Recall the generative model (1). Since $\boldsymbol{w}$ follows a sub-Gaussian distribution, it can be shown that for every $t > 0$,

$$\mathbb{P}(\|\boldsymbol{x}\|_2 > t) \leq \mathbb{P}\left(\|\boldsymbol{w}\|_2 > \frac{t}{\|\mathcal{H}(\boldsymbol{S})\|}\right) \leq Ce^{-v' t^2},$$

for some positive constant v', which means that $\boldsymbol{x}$ also follows a sub-Gaussian distribution. Thus, due to the sub-Gaussian property, $\|\boldsymbol{C}_n - \boldsymbol{C}_\infty\|$ can be explictly bounded by the following lemma.

Lemma E.5 ([39, Proposition 2.1])**.** *Consider sub-Gaussian, identical, independent random vectors* $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n \in \mathbb{R}^m$ *with* $n > m$*. Then for all* $\varepsilon > 0$*, it follows that*

$$\mathbb{P}\left(\left\|\frac{1}{n}\sum_{i=1}^n \boldsymbol{x}_i \boldsymbol{x}_i^\top - \mathbb{E}[\boldsymbol{x}\boldsymbol{x}^\top]\right\|_2 \leq \varepsilon\right) \geq 1 - 2e^{2m - l\varepsilon^2 n},$$

for some constant $l > 0$*.*

Setting $\varepsilon^2 = (4/l)\log(2n)m/n$, Lemma E.5 indicates that with high probability (lower bounded by $1 - n^{-1}$),

$$\|\boldsymbol{C}_n - \boldsymbol{C}_\infty\| \leq \mathcal{O}\left(\sqrt{\frac{\log n}{n}}\right).$$

F Derivations of L-ADMM and Convergence Analysis

This section includes the details of L-ADMM for rLogSpecT.

F.1 Proof of Proposition 5.1

Note that the minimization problem (7) is separable for $\boldsymbol{Z}$ and $\boldsymbol{q}$, and can be split into two subproblems:

$$\min_{\boldsymbol{Z}\in\mathbb{B}(\boldsymbol{0},\delta_n)} \|\boldsymbol{C}_n\boldsymbol{S}^{(k)} - \boldsymbol{S}^{(k)}\boldsymbol{C}_n + \boldsymbol{\Lambda}^{(k)}/\rho - \boldsymbol{Z}\|_F^2, \tag{26}$$

$$\min_{\boldsymbol{q}} \ -\alpha\boldsymbol{1}^\top \log \boldsymbol{q} + \boldsymbol{\lambda}_2^{(k)\top}(\boldsymbol{q} - \boldsymbol{S}^{(k)}\boldsymbol{1}) + \frac{\rho}{2}\|\boldsymbol{q} - \boldsymbol{S}^{(k)}\boldsymbol{1}\|_2^2. \tag{27}$$

For problem (26), the optimal solution is the projection of $\boldsymbol{C}_n\boldsymbol{S}^{(k)} - \boldsymbol{S}^{(k)}\boldsymbol{C}_n + \boldsymbol{\Lambda}^{(k)}/\rho$ onto $\mathbb{B}(\boldsymbol{0},\delta_n)$, which is given by

$$\boldsymbol{Z}^{(k+1)} = \min\left\{1, \frac{\delta_n}{\|\tilde{\boldsymbol{Z}}\|_F}\right\}\tilde{\boldsymbol{Z}} \ \text{ with } \ \tilde{\boldsymbol{Z}} = \boldsymbol{C}_n\boldsymbol{S}^{(k)} - \boldsymbol{S}^{(k)}\boldsymbol{C}_n + \boldsymbol{\Lambda}^{(k)}/\rho.$$

For problem (27), the first-order optimality condition gives

$$-\alpha 1/\boldsymbol{q} + \boldsymbol{\lambda}_2^{(k)} + \rho(\boldsymbol{q} - \boldsymbol{S}^{(k)}\boldsymbol{1}) = 0.$$

This together with the fact that the objective function is convex implies that

$$\boldsymbol{q}^{(k+1)} = \frac{\tilde{\boldsymbol{q}} + \sqrt{\tilde{\boldsymbol{q}}^2 + 4\alpha/\rho\boldsymbol{1}}}{2} \ \text{ with } \ \tilde{\boldsymbol{q}} = \frac{1}{\rho}(\rho\boldsymbol{S}^{(k)}\boldsymbol{1} - \boldsymbol{\lambda}_2^{(k)}).$$

F.2 Calculation of $\Pi_{\mathcal{S}}(\cdot)$

The projection of $\boldsymbol{X}$ to $\mathcal{S}$ can be calculated via an optimization problem:

$$\begin{aligned} \min_{\boldsymbol{S}} \ & \|\boldsymbol{X} - \boldsymbol{S}\|_F^2 \\ \text{s.t.} \ & \boldsymbol{S}^\top = \boldsymbol{S}, \\ & S_{ii} = 0, \ i = 1, 2, \ldots, m, \\ & S_{ij} \geq 0, \ \forall i, j, \end{aligned}$$

which is equivalent to

$$\begin{aligned} \min \ & \sum_{i<j} \left((X_{ij} - S_{ij})^2 + (X_{ji} - S_{ij})^2\right) \\ \text{s.t.} \ & S_{ij} \geq 0, \ \forall i < j, \\ & S_{ii} = 0, \ \forall i. \end{aligned}$$

Hence

$$(\Pi_{\mathcal{S}}(\boldsymbol{X}))_{ij} = \begin{cases} \frac{1}{2}\max\{0, X_{ij} + X_{ji}\}, & i \neq j, \\ 0, & i = j. \end{cases}$$

F.3 Stopping criterion and updating rule of ρ

We follow the procedures in [3] to update ρ in each iteration. Similarly, we define the primal residual and dual residual as follows:

$$\begin{aligned} p_{\text{res}}^{(k+1)} &= \sqrt{\|\boldsymbol{Z}^{(k+1)} - \boldsymbol{C}_n\boldsymbol{S}^{(k+1)} + \boldsymbol{S}^{(k+1)}\boldsymbol{C}_n\|_F^2 + \|\boldsymbol{q}^{(k+1)} - \boldsymbol{S}^{(k+1)}\boldsymbol{1}\|_2^2}, \\ d_{\text{res}}^{(k+1)} &= \rho^{(k)}\left(\boldsymbol{C}_n(\boldsymbol{S}^{(k+1)} - \boldsymbol{S}^{(k)}) - (\boldsymbol{S}^{(k+1)} - \boldsymbol{S}^{(k)})\boldsymbol{C}_n + \boldsymbol{1}^\top(\boldsymbol{S}^{(k+1)} - \boldsymbol{S}^{(k)})\boldsymbol{1}\right). \end{aligned}$$

The aim of updating ρ is to control the decaying speed of p_{res} and d_{res} such that their difference is not too large. To this end, we update ρ adaptively following the scheme:

$$\rho^{(k+1)} := \begin{cases} 2\rho^{(k)}, & \text{if } p_{\text{res}}^{(k+1)} > 5d_{\text{res}}^{(k+1)}, \\ \rho^{(k)}/2, & \text{if } d_{\text{res}}^{(k+1)} > 5p_{\text{res}}^{(k+1)}, \\ \rho^{(k)}, & \text{otherwise.} \end{cases}$$

When p_{res} and d_{res} are both smaller than the threshold $\varepsilon = 10^{-5}$, we stop the algorithm.

F.4 Convergence analysis

Define $\boldsymbol{D} := \mathrm{Diag}(\mathbf{1}_m^\top, \ldots, \mathbf{1}_m^\top) \in \mathbb{R}^{m \times m^2}$. Then, $\boldsymbol{D}$ satisfies $\boldsymbol{D}\mathrm{vec}(\boldsymbol{S}) = \boldsymbol{S}\mathbf{1}$ and $\|\boldsymbol{D}^\top \boldsymbol{D}\| = m$. Denote

$$\boldsymbol{Q} := \tau \boldsymbol{I} - \boldsymbol{D}^\top \boldsymbol{D} - \boldsymbol{A}_n^\top \boldsymbol{A}_n.$$

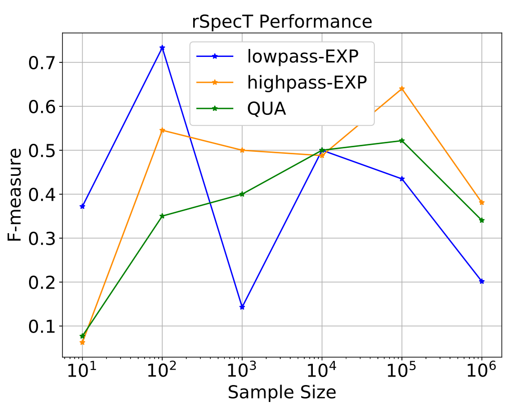
Then the linearized ADMM update (8) of $\boldsymbol{S}$ can be written as:

$$\min_{\boldsymbol{S}} \; L(\boldsymbol{S}) + \frac{\rho}{2} \|\mathrm{vec}(\boldsymbol{S}) - \mathrm{vec}(\boldsymbol{S}^{(k)})\|_{\boldsymbol{Q}},$$

where $\|\boldsymbol{x}\|_{\boldsymbol{Q}} := \boldsymbol{x}^\top \boldsymbol{Q} \boldsymbol{x}$. Since $\tau > m + \|\boldsymbol{A}_n\|^2$, we know that $\boldsymbol{Q}$ is positively definite. Consequently, by treating $(\boldsymbol{Z}, \boldsymbol{q})$ as one variable, we can apply Theorem 4.2 in [43] and directly obtain the result.

G More Experiments and Discussions on Synthetic Data

To make a fair comparison between rSpecT and rLogSpecT, we test rSpecT on BA graphs with the same graph filters and the results are reported in Figure 5. It is obvious that rSpecT fails in these cases and cannot benefit from the increase in sample size. This is reasonable since SpecT fails on BA graphs as indicated in Figure 1, let alone the approximation formulation rSpecT.



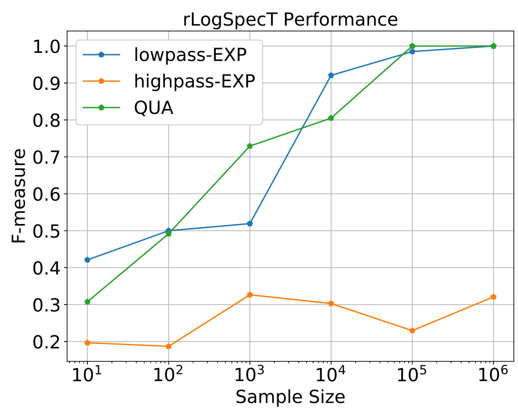


Figure 5: Performance of rSpecT on BA graphs. Figure 6: rLogSpecT on ER graphs with $\delta_n = 20\sqrt{\log n / n}$.

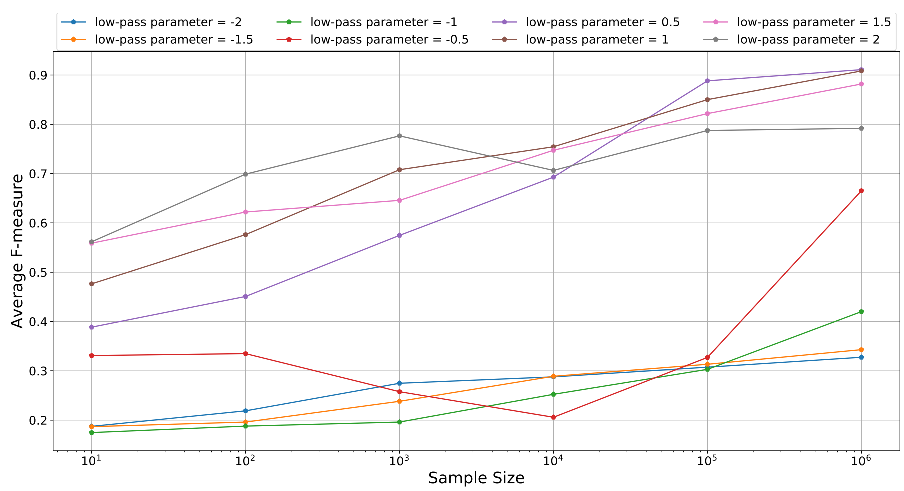


Figure 7: Effect of Low-Pass Parameter: different performance of graph filters $\exp(t\boldsymbol{S})$ with t ranging from -2 to 2.

We further test rLogSpecT on ER graphs with different numbers of signals observed. The parameter δ_n is set as $20\sqrt{\log n / n}$ and the results are reported in Figure 6. The figure shows that for graph filters that are not high-pass, rLogSpecT can achieve nearly perfect recovery when the sample size is large enough. Also, compared with the performance on BA graphs, rLogSpecT works better on ER graphs. This observation is in accordance with the conclusion from Figure 1 that LogSpecT performs better on ER graphs than BA ones. We further notice that the difference between the low-pass graph filter and the high-pass one is huge. To check the conjecture that rLogSpecT generally performs better on low-pass graph filters, we choose different graph filters $\exp(t\boldsymbol{S})$ with t ranging from -2 to 2 and conduct the experiments on ER graphs. When the graph shifting operator is the adjacency matrix, the positive low-pass parameter t corresponds to low-pass graph filters and the negative t corresponds to the high-pass ones [25, 10]. We omit the case when $t = 0$ since this filter does not contain any graph information (note that $\exp(0\boldsymbol{S}) = \boldsymbol{I}$).

We then repeat the experiments for 50 times and report the average results in Figure 7. The comparison between the performance of low-pass graph filters and high-pass graph filters indicates that the low-pass graph filters generally outperforms the high-pass ones. A closer look at the results shows that the performance grows faster when the absolute value of t is smaller. And eventually, the graph filter with smaller absolute value of t prevails. This observation is interesting since Figure 1 indicates that the choice of graph filters has few impacts on the model performance. One explanation is that both low-pass graph filters and high-pass graph filters attenuate some frequencies of the graph and the larger absolute value of t leads to the more loss of information carried by finite signals.