Supplement: Proofs, Derivations and Extra Example

Qiong Zhang*[1] and Lulu Kang†[2]

[1]School of Mathematical and Statistical Sciences, Clemson University
[2]Department of Applied Mathematics, Illinois Institute of Technology

S1. Proof of Theorem 1

Proof. Because $T(\boldsymbol{x},\rho) = m - T_1(\boldsymbol{x},\rho) - T_2(\boldsymbol{x},\rho)$, we investigate the derivatives T_1 and T_2 with respect to ρ. For T_1,

$$T_1(\boldsymbol{x},\rho) = \rho\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{x},$$
$$\frac{\partial T_1(\boldsymbol{x},\rho)}{\partial\rho} = \boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{x}, \quad \frac{\partial^2 T_1(\boldsymbol{x},\rho)}{\partial\rho^2} = 0.$$

To derive the derivatives for T_2, we first introduce some notation to shorten the formulas. Let $\boldsymbol{A} := \boldsymbol{F}^\top(\boldsymbol{D}-\rho\boldsymbol{W})F$, $\boldsymbol{A}_1 := \frac{\partial\boldsymbol{A}^{-1}}{\partial\rho}$, and $\boldsymbol{A}_2 := \frac{\partial^2\boldsymbol{A}^{-1}}{\partial\rho^2}$. Following the calculus of matrix,

$$\begin{aligned}
\boldsymbol{A}_1 &= -\boldsymbol{A}^{-1}\frac{\partial\boldsymbol{A}}{\partial\rho}\boldsymbol{A}^{-1} = -\boldsymbol{A}^{-1}\frac{\partial\boldsymbol{F}^\top\boldsymbol{D}\boldsymbol{F}-\rho\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{F}}{\partial\rho}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}^{-1},\\
\boldsymbol{A}_2 &= \frac{\partial\boldsymbol{A}_1}{\partial\rho} = \frac{\partial\boldsymbol{A}^{-1}}{\partial\rho}\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}^{-1} + \boldsymbol{A}^{-1}\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{F}\frac{\partial\boldsymbol{A}^{-1}}{\partial\rho} = 2\boldsymbol{A}^{-1}\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}^{-1}.
\end{aligned}$$

Using the new notation,

$$\begin{aligned}
T_2(\boldsymbol{x},\rho) &= \boldsymbol{x}^\top(\boldsymbol{D}-\rho\boldsymbol{W})\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top(\boldsymbol{D}-\rho\boldsymbol{W})\boldsymbol{x}\\
&= \underbrace{\boldsymbol{x}^\top\boldsymbol{D}\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top\boldsymbol{D}\boldsymbol{x}}_{\text{Term 1}} - 2\underbrace{\rho\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top\boldsymbol{D}\boldsymbol{x}}_{\text{Term 2}} + \underbrace{\rho^2\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{x}}_{\text{Term 3}}.
\end{aligned}$$

The first order derivative of the three terms with respect to ρ are

$$\begin{aligned}
\frac{\partial\text{Term 1}}{\partial\rho} &= \boldsymbol{x}^\top\boldsymbol{D}\boldsymbol{F}\frac{\partial\boldsymbol{A}^{-1}}{\partial\rho}\boldsymbol{F}^\top\boldsymbol{D}\boldsymbol{x} = \boldsymbol{x}^\top\boldsymbol{D}\boldsymbol{F}\boldsymbol{A}_1F^\top\boldsymbol{D}\boldsymbol{x}\\
\frac{\partial\text{Term 2}}{\partial\rho} &= \boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top\boldsymbol{D}\boldsymbol{x} + \rho\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\frac{\partial\boldsymbol{A}^{-1}}{\partial\rho}\boldsymbol{F}^\top\boldsymbol{D}\boldsymbol{x}\\
&= \boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top\boldsymbol{D}\boldsymbol{x} + \rho\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}_1\boldsymbol{F}^\top\boldsymbol{D}\boldsymbol{x}\\
\frac{\partial\text{Term 3}}{\partial\rho} &= 2\rho\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{x} + \rho^2\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\frac{\partial\boldsymbol{A}^{-1}}{\partial\rho}\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{x}\\
&= 2\rho\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{x} + \rho^2\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}_1\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{x}.
\end{aligned}$$

*qiongz@clemson.edu
†lkang2@iit.edu

To combine the three derivatives,

$$\frac{\partial T_2(\boldsymbol{x},\rho)}{\partial \rho} = \boldsymbol{x}^\top(\boldsymbol{D}-\rho\boldsymbol{W})\boldsymbol{F}\boldsymbol{A}_1\boldsymbol{F}^\top(\boldsymbol{D}-\rho\boldsymbol{W})\boldsymbol{x} - 2\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top(\boldsymbol{D}-\rho\boldsymbol{W})\boldsymbol{x}$$

The derivative of $T(\boldsymbol{x},\rho)$ is,

$$\begin{aligned}\frac{\partial T(\boldsymbol{x},\rho)}{\partial \rho} &= -\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{x} - \boldsymbol{x}^\top(\boldsymbol{D}-\rho\boldsymbol{W})\boldsymbol{F}\boldsymbol{A}_1\boldsymbol{F}^\top(\boldsymbol{D}-\rho\boldsymbol{W})\boldsymbol{x} + 2\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top(\boldsymbol{D}-\rho W)\boldsymbol{x}\\ &= -\boldsymbol{x}^\top\boldsymbol{W}\left[\boldsymbol{I}_n - \boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top(\boldsymbol{D}-\rho\boldsymbol{W})\right]\boldsymbol{x}\\ &- \boldsymbol{x}^\top\left[(\boldsymbol{D}-\rho\boldsymbol{W})\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top - \boldsymbol{I}_n\right]\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top(\boldsymbol{D}-\rho\boldsymbol{W})\boldsymbol{x}\\ &= -\boldsymbol{x}^\top\left[\boldsymbol{I}_n - \boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top(\boldsymbol{D}-\rho\boldsymbol{W})\right]^\top\boldsymbol{W}\left[\boldsymbol{I}_n - \boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top(\boldsymbol{D}-\rho\boldsymbol{W})\right]\boldsymbol{x}.\end{aligned}$$

It is interesting to notice that $\boldsymbol{s} := \left[\boldsymbol{I}_n - \boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top(\boldsymbol{D}-\rho\boldsymbol{W})\right]\boldsymbol{x}$ can be considered as the residuals of regression model $\boldsymbol{x} = \boldsymbol{F}\boldsymbol{\beta} + \boldsymbol{v}$, where $\boldsymbol{v}$ is the vector with mean equal to $\boldsymbol{0}$ and covariance matrix $\boldsymbol{D}-\rho\boldsymbol{W}$. By the definition of the adjacency matrix,

$$\frac{\partial T(\boldsymbol{x},\rho)}{\partial \rho} = -\sum_{w_{i,j}=1} s_i s_j.$$

Thus, the sign of $\partial T(\boldsymbol{x},\rho)/\partial\rho$ is uncertain and is possible to be either positive or negative.

Next, we compute the second order derivative of $T_2(\boldsymbol{x},\rho)$ with respect to ρ.

$$\begin{aligned}\frac{\partial^2 \text{Term 1}}{\partial \rho^2} &= \boldsymbol{x}^\top\boldsymbol{D}\boldsymbol{F}\frac{\partial \boldsymbol{A}_1}{\partial \rho}\boldsymbol{F}^\top\boldsymbol{D}\boldsymbol{x} = \boldsymbol{x}^\top\boldsymbol{D}\boldsymbol{F}\boldsymbol{A}_2\boldsymbol{F}^\top\boldsymbol{D}\boldsymbol{x},\\ \frac{\partial^2 \text{Term 2}}{\partial \rho^2} &= \boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\frac{\partial \boldsymbol{A}^{-1}}{\partial \rho}\boldsymbol{F}^\top\boldsymbol{D}\boldsymbol{x} + \boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}_1 F^\top\boldsymbol{D}\boldsymbol{x} + \rho\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\frac{\partial \boldsymbol{A}_1}{\partial \rho}\boldsymbol{F}^\top\boldsymbol{D}\boldsymbol{x}\\ &= 2\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}_1\boldsymbol{F}^\top\boldsymbol{D}\boldsymbol{x} + \rho\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}_2\boldsymbol{F}^\top\boldsymbol{D}\boldsymbol{x},\\ \frac{\partial^2 \text{Term 3}}{\partial \rho^2} &= 2\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{x} + 2\rho\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\frac{\partial \boldsymbol{A}^{-1}}{\partial \rho}\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{x} + 2\rho\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}_1\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{x}\\ &+ \rho^2\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\frac{\partial \boldsymbol{A}_1}{\partial \rho}\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{x}\\ &= 2\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{x} + 4\rho\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}_1\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{x} + \rho^2\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}_2\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{x}.\end{aligned}$$

Let $\boldsymbol{C} := \boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{F}$. Then

$$\begin{aligned}\frac{\partial^2 T_2(\boldsymbol{x},\rho)}{\partial \rho^2} &= 2\boldsymbol{x}^\top(\boldsymbol{D}-\rho\boldsymbol{W})\boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{C}^\top(\boldsymbol{D}-\rho\boldsymbol{W})\boldsymbol{x} - 4\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{C}^\top(\boldsymbol{D}-\rho\boldsymbol{W})\boldsymbol{x}\\ &+ 2\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{x}\\ &= 2\boldsymbol{x}^\top\left[(\boldsymbol{D}-\rho\boldsymbol{W})\boldsymbol{C} - \boldsymbol{W}\boldsymbol{F}\right]\boldsymbol{A}^{-1}\left[\boldsymbol{C}^\top(\boldsymbol{D}-\rho\boldsymbol{W}) - \boldsymbol{F}^\top\boldsymbol{W}\right]\boldsymbol{x}.\end{aligned}$$

For any $\rho \in (0,1)$, it is apparent that $\boldsymbol{D}-\rho\boldsymbol{W}$ is the Laplacian matrix of the weighted undirected graph with the constant weight ρ for each edge, and it is also clear that $\boldsymbol{D}-\rho\boldsymbol{W}$ is a positive definite matrix. We assume $\boldsymbol{F}$ is a full rank matrix so that the regression model is valid. So $\boldsymbol{A}$ and $\boldsymbol{A}^{-1}$ are both positive definite. Thus, $\frac{\partial^2 T_2(\boldsymbol{x},\rho)}{\partial\rho^2} \geq 0$ and $\frac{\partial^2 T(\boldsymbol{x},\rho)}{\partial\rho^2} \leq 0$ for any $\rho\in(0,1)$. The design criterion $T(\boldsymbol{x},\rho)$, which is to be maximized, is concave. $\square$

S2. The Gap between $\mathbb{E}[T(\boldsymbol{x},\rho)]$ and $T(\boldsymbol{x},\rho_0)$

We randomly generate a network of size $n = 50$. For each pair of nodes, an edge will connect the two with a probability of 1/4 and the existence of the edge is independent of any other random variables. The covariate z_i is generated from a one-dimensional normal distribution $N(0, 10^2)$ and z_i's are independent of each other and the network structure. The prior distribution of ρ is uniform distribution in $[0, 1]$ and $\rho_0 = \mathbb{E}(\rho) = 1/2$. We randomly generate 400 completely randomized designs $\boldsymbol{x}_l$ for $l = 1, \ldots, 400$ and calculate $T(\boldsymbol{x}_l, \rho_0)$, whose histogram is plotted in the left panel of Figure S1. For any given design $\boldsymbol{x}_l$, we randomly samples ρ_i for $i = 1, \ldots, 200$ and calculate $T(\boldsymbol{x}_l, \rho_i)$. The mean $\mathbb{E}[T(\boldsymbol{x}_l, \rho)]$ is approximated by the sample mean of $T(\boldsymbol{x}_l, \rho_i)$'s. The histogram of the gap $T(\boldsymbol{x}_l, \rho_0) - \mathbb{E}[T(\boldsymbol{x}_l, \rho)]$ for all the random designs is plotted in the right panel of Figure S1. Based on the two histograms, the gap $T(\boldsymbol{x}, \rho_0) - \mathbb{E}[T(\boldsymbol{x}, \rho)]$ is relatively small compared to the range of $T(\boldsymbol{x}, \rho_0)$. Thus, it is reasonable to use the surrogate local design criterion $T(\boldsymbol{x}_l, \rho_0)$ to replace $\mathbb{E}[T(\boldsymbol{x}, \rho)]$ for this simple example.

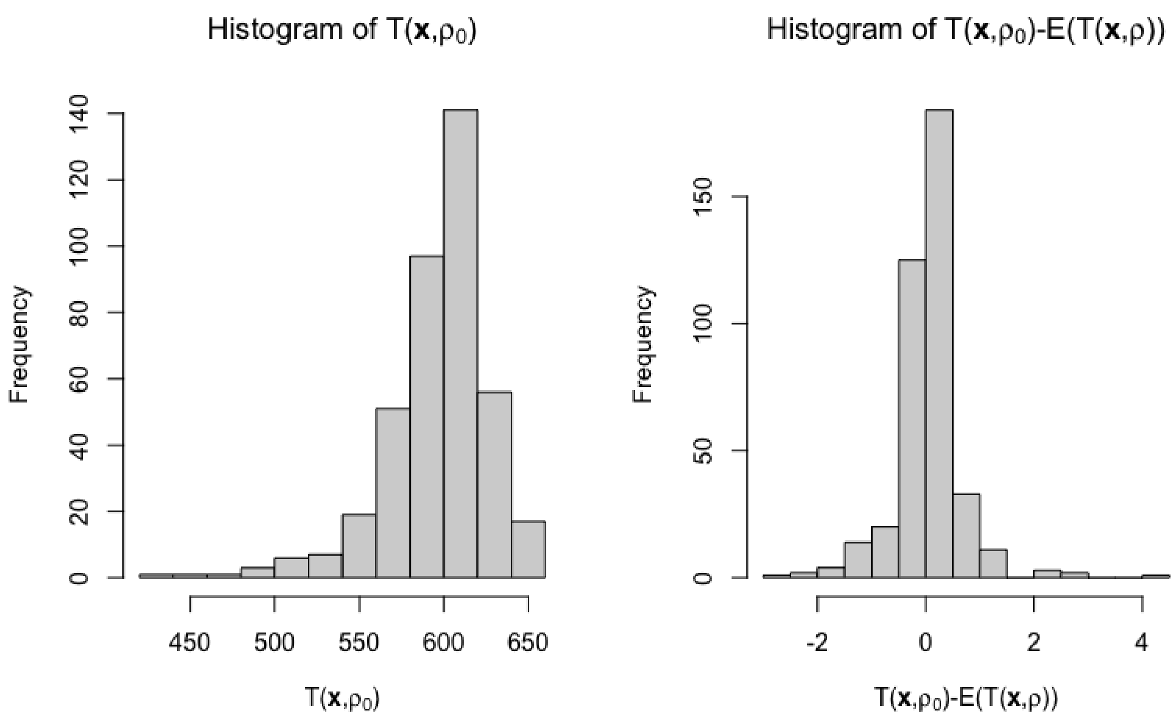

Figure S1: Histogram of $T(\boldsymbol{x}, \rho_0)$ and the gap $T(\boldsymbol{x}, \rho_0) - \mathbb{E}[T(\boldsymbol{x}, \rho)]$

In more general case, Proposition S1 provides the analytic gap between $T(\boldsymbol{x}, \rho_0)$ and $\mathbb{E}[T(\boldsymbol{x}, \rho)]$. Its proof is provided in the Supplement. Proposition S1 also provides two different upper bounds of the gap. Which one of the two upper bounds is larger depends on the adjacency matrix $\boldsymbol{W}$ and ρ_0. Regrettably, since both the upper bounds are independent of the design $\boldsymbol{x}$, they are too large to have any practical guidance, even though they might still be attainable for certain extreme design $\boldsymbol{x}$. For the above simulation example, since the skewness of uniform distribution is 0, the two upper bounds of (4) and (5) are calculated as 902.4 and 650.1, respectively. They are much larger than the range shown in the histogram in Figure S1. On the other hand, the two upper bounds increase as the size and density of the network become larger. Therefore, for large and dense networks we should be more careful applying the locally optimal design.

Proposition S1. *The difference between* $T(\boldsymbol{x}, \rho_0)$ *and* $\mathbb{E}(T(\boldsymbol{x}, \rho))$ *is*

$$T(\boldsymbol{x}, \rho_0) - \mathbb{E}(T(\boldsymbol{x}, \rho)) = \frac{1}{2} \left.\frac{\partial^2 T_2(\boldsymbol{x}, \rho)}{\partial \rho^2}\right|_{\rho=\rho_0} \mathrm{var}(\rho) - \mathbb{E}(O(\rho - \rho_0)^3), \tag{1}$$

where

$$\frac{1}{2} \left.\frac{\partial^2 T_2(\boldsymbol{x}, \rho)}{\partial \rho^2}\right|_{\rho=\rho_0} = \boldsymbol{s}^\top \boldsymbol{W} \boldsymbol{F} \left[\boldsymbol{F}^\top (\boldsymbol{D} - \rho_0 \boldsymbol{W}) \boldsymbol{F}\right]^{-1} \boldsymbol{F}^\top \boldsymbol{W} \boldsymbol{s}, \tag{2}$$

$$\text{and} \quad \boldsymbol{s} := \left[\boldsymbol{I}_n - \boldsymbol{F}(\boldsymbol{F}^\top (\boldsymbol{D} - \rho_0 \boldsymbol{W}) \boldsymbol{F})^{-1} \boldsymbol{F}^\top (\boldsymbol{D} - \rho_0 \boldsymbol{W})\right] \boldsymbol{x}. \tag{3}$$

An upper bound of the gap $T(\boldsymbol{x}, \rho_0) - \mathbb{E}(T(\boldsymbol{x}, \rho))$ *is*

$$T(\boldsymbol{x}, \rho_0) - \mathbb{E}(T(\boldsymbol{x}, \rho)) \le \min \{n \lambda_{\max}(\boldsymbol{D} - \rho_0 \boldsymbol{W}), (1 + \rho_0) m\} \frac{|\lambda(\boldsymbol{W})|^2_{\max} \mathrm{var}(\rho)}{\lambda^2_{\min}(\boldsymbol{D} - \rho_0 \boldsymbol{W})} - \mathbb{E}\left[O(\rho - \rho_0)^3\right], \tag{4}$$

where $\lambda_{\min}(\boldsymbol{D} - \rho_0 \boldsymbol{W})$ *and* $\lambda_{\max}(\boldsymbol{D} - \rho_0 \boldsymbol{W})$ *are the minimum and maximum eigenvalues of the Laplacian matrix* $\boldsymbol{D} - \rho_0 \boldsymbol{W}$*, which is positive definite for* $\rho_0 \in (0, 1)$*,* $|\lambda(\boldsymbol{W})|_{\max}$ *is the spectrum radius of* $\boldsymbol{W}$*, and* $m = \sum_{i=1}^n m_i$*. Based on Theorem 2, an alternative upper bound (5) holds asymptotically with probability of* $100(1 - \alpha)\%$ *and* $\alpha \in (0, 1)$*,*

$$T(\boldsymbol{x}, \rho_0) - \mathbb{E}(T(\boldsymbol{x}, \rho)) \le (m + z_\alpha \sqrt{m}) \frac{|\lambda(\boldsymbol{W})|^2_{\max} \mathrm{var}(\rho)}{\lambda^2_{\min}(\boldsymbol{D} - \rho_0 \boldsymbol{W})} - \mathbb{E}\left[O(\rho - \rho_0)^3\right], \tag{5}$$

where $z_\alpha = \Phi^{-1}(\alpha)$ *is the upper* α *quantile of the standard normal distribution.*

Lemma S1. *Let* $\boldsymbol{A}$ *be an* $n \times n$ *real symmetric positive definite matrix. For any vector* $\boldsymbol{x} \in \mathbb{R}^n$*,* $\lambda_{\min}(\boldsymbol{A}) ||\boldsymbol{x}||_2^2 \le \boldsymbol{x}^\top \boldsymbol{A} \boldsymbol{x} \le \lambda_{\max}(\boldsymbol{A}) ||\boldsymbol{x}||_2^2$*. The equality holds if* $\boldsymbol{x} = \boldsymbol{0}$ *or* $\boldsymbol{A} = a \boldsymbol{I}_n$ *for* $a \ge 0$*.*

Proof. Because $\boldsymbol{A}$ is a real symmetric positive definite matrix, via eigendecomposition, $\boldsymbol{A} = \boldsymbol{Q} \boldsymbol{\Lambda} \boldsymbol{Q}^{-1}$, where $\boldsymbol{\Lambda} = \mathrm{diag}\{\lambda_1, \dots, \lambda_n\}$ is a diagonal matrix of the eigenvalues of $\boldsymbol{A}$, $\boldsymbol{Q}$ is the square $n \times n$ matrix whose ith column is the eigenvector corresponding to eigenvalue λ_i. Also, $\boldsymbol{Q}^\top = \boldsymbol{Q}^{-1}$. Denote $\boldsymbol{l} := \boldsymbol{Q}^\top \boldsymbol{x}$.

$$\boldsymbol{x}^\top \boldsymbol{A} \boldsymbol{x} = \boldsymbol{x}^\top \boldsymbol{Q} \boldsymbol{\Lambda} \boldsymbol{Q}^\top \boldsymbol{x} = \boldsymbol{l}^\top \boldsymbol{\Lambda} \boldsymbol{l} = \sum_{i=1}^n \lambda_i l_i^2,$$

$$\lambda_{\min}(\boldsymbol{A}) ||\boldsymbol{l}||_2^2 = \lambda_{\min}(\boldsymbol{A}) \sum_{i=1}^n l_i^2 \le \sum_{i=1}^n \lambda_i l_i^2 \le \lambda_{\max}(\boldsymbol{A}) \sum_{i=1}^n l_i^2 = \lambda_{\max}(\boldsymbol{A}) ||\boldsymbol{l}||_2^2.$$

Here $\lambda_{\max}(\boldsymbol{A})$ and $\lambda_{\min}(\boldsymbol{A})$ are the maximum and minimum eigenvalues of $\boldsymbol{A}$, and since $\boldsymbol{A}$ is positive definite, $\lambda_{\min}(\boldsymbol{A}) > 0$. The norm $|| \cdot ||_2$ is the l_2-norm of a vector, and $||\boldsymbol{l}||_2^2 = \boldsymbol{l}^\top \boldsymbol{l} = \boldsymbol{x}^\top \boldsymbol{Q} \boldsymbol{Q}^\top \boldsymbol{x} = ||\boldsymbol{x}||_2^2$. Thus the lemma is proved. □

Lemma S2. *Let* $\boldsymbol{A}$ *be an* $n \times n$ *real symmetric matrix. For any vector* $\boldsymbol{x} \in \mathbb{R}^n$*,* $|\boldsymbol{x}^\top \boldsymbol{A} \boldsymbol{x}| \le |\lambda(\boldsymbol{A})|_{\max} ||\boldsymbol{x}||_2^2$*.*

Proof. For any real symmetric matrix, based on eigenvalue decomposition, $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^\top$, where $\boldsymbol{\Lambda} = \text{diag}\{\lambda_1, \dots, \lambda_n\}$ is a diagonal matrix of the eigenvalues of $\boldsymbol{A}$, and $\boldsymbol{Q}$ is the $n \times n$ orthogonal matrix as above. Denote $\boldsymbol{l} := \boldsymbol{Q}^\top \boldsymbol{x}$.

$$|\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}| = |\boldsymbol{x}^\top \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^\top \boldsymbol{x}| = |\boldsymbol{l}^\top \boldsymbol{\Lambda} \boldsymbol{l}| = |\sum_{i=1}^n \lambda_i l_i^2| \leq \sum_{i=1}^n |\lambda_i| l_i^2 \leq |\lambda(\boldsymbol{A})|_{\max} ||\boldsymbol{l}||_2^2 = |\lambda(\boldsymbol{A})|_{\max} ||\boldsymbol{x}||_2^2.$$

Here $|\lambda(\boldsymbol{A})|_{\max} = \max_{i=1,\dots,n} |\lambda|_i$. □

Proof of Proposition S1

Proof. Using Taylor expansion, we have

$$T(\boldsymbol{x}, \rho) = T(\boldsymbol{x}, \rho_0) + \left.\frac{\partial T(\boldsymbol{x}, \rho)}{\partial \rho}\right|_{\rho=\rho_0} (\rho - \rho_0) + \frac{1}{2} \left.\frac{\partial^2 T(\boldsymbol{x}, \rho)}{\partial \rho^2}\right|_{\rho=\rho_0} (\rho - \rho_0)^2 + O((\rho - \rho_0)^3).$$

Apply expectation on both side of the equaiton with respet the priori $p(\rho)$, we have

$$\begin{aligned}
\mathbb{E}\left[T(\boldsymbol{x}, \rho)\right] &= T(\boldsymbol{x}, \rho_0) + \left.\frac{\partial T(\boldsymbol{x}, \rho)}{\partial \rho}\right|_{\rho=\rho_0} \mathbb{E}\left[\rho - \rho_0\right] + \frac{1}{2} \left.\frac{\partial^2 T(\boldsymbol{x}, \rho)}{\partial \rho^2}\right|_{\rho=\rho_0} \mathbb{E}\left[(\rho - \rho_0)^2\right] + \mathbb{E}\left[O((\rho - \rho_0)^3)\right] \\
&= T(\boldsymbol{x}, \rho_0) + \frac{1}{2} \left.\frac{\partial^2 T(\boldsymbol{x}, \rho)}{\partial \rho^2}\right|_{\rho=\rho_0} \text{var}(\rho) + \mathbb{E}\left[O((\rho - \rho_0)^3)\right].
\end{aligned}$$

From the proof of Theorem 1, we have that

$$\frac{\partial^2 T(\boldsymbol{x}, \rho)}{\partial \rho^2} = -\frac{\partial^2 T_2(\boldsymbol{x}, \rho)}{\partial \rho^2}.$$

Thus we obtain the gap between $T(\boldsymbol{x}, \rho_0)$ and $\mathbb{E}\left[T(\boldsymbol{x}, \rho)\right]$ in (1). Also in proof of Theorem 1,

$$\frac{1}{2} \left.\frac{\partial^2 T_2(\boldsymbol{x}, \rho)}{\partial \rho^2}\right|_{\rho=\rho_0} = \boldsymbol{s}^\top \boldsymbol{W}\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top \boldsymbol{W}\boldsymbol{s},$$

where

$$\begin{aligned}
\boldsymbol{A} &= \boldsymbol{F}^\top (\boldsymbol{D} - \rho_0 \boldsymbol{W})\boldsymbol{F}, \\
\boldsymbol{s} &= \left[\boldsymbol{I}_n - \boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top (\boldsymbol{D} - \rho_0 \boldsymbol{W})\right] \boldsymbol{x}.
\end{aligned}$$

From the definition of $\boldsymbol{s}$, we can see that

$$\begin{aligned}
\boldsymbol{s}^\top (\boldsymbol{D} - \rho_0 \boldsymbol{W})\boldsymbol{s} &= \boldsymbol{x}^\top \left[(\boldsymbol{D} - \rho_0 \boldsymbol{W}) - (\boldsymbol{D} - \rho_0 \boldsymbol{W})\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top (\boldsymbol{D} - \rho_0 \boldsymbol{W})\right] \boldsymbol{x} \\
&\leq \boldsymbol{x}^\top (\boldsymbol{D} - \rho_0 \boldsymbol{W})\boldsymbol{x}.
\end{aligned}$$

From Lemma S1, since $\boldsymbol{D} - \rho_0 \boldsymbol{W}$ is a real symmetric positive definite matrix as $\rho_0 \in (0, 1)$,

$$\lambda_{\min}(\boldsymbol{D} - \rho_0 \boldsymbol{W}) ||\boldsymbol{s}||_2^2 \leq \lambda_{\max}(\boldsymbol{D} - \rho_0 \boldsymbol{W}) ||\boldsymbol{x}||_2^2 = \lambda_{\max}(\boldsymbol{D} - \rho_0 \boldsymbol{W}) n.$$

On the other hand, $\boldsymbol{x}^\top (\boldsymbol{D} - \rho_0 \boldsymbol{W})\boldsymbol{x} \leq (1 + \rho_0)m$. Thus,

$$||\boldsymbol{s}||_2^2 \leq \frac{1}{\lambda_{\min}(\boldsymbol{D} - \rho_0 \boldsymbol{W})} \min\{n\lambda_{\max}(\boldsymbol{D} - \rho_0 \boldsymbol{W}), (1 + \rho_0)m\}.$$

According to Theorem 2, $\boldsymbol{x}^\top \boldsymbol{W}\boldsymbol{x}/\sqrt{m}$ converges in distribution to the standard normal distribution. Therefore, with probability of $100(1-\alpha)\%$, $\boldsymbol{x}^\top \boldsymbol{W}\boldsymbol{x} \geq -z_\alpha\sqrt{m}$, asymptotically. Here z_α is the upper α quantile of the standard normal distribution, i.e., $z_\alpha = \Phi^{-1}(1-\alpha)$. So we can obtain an asymptotic upper bound,

$$\boldsymbol{s}^\top(\boldsymbol{D}-\rho_0\boldsymbol{W})\boldsymbol{s} \leq \boldsymbol{x}^\top(\boldsymbol{D}-\rho_0\boldsymbol{W})\boldsymbol{x} = \boldsymbol{x}^\top\boldsymbol{D}\boldsymbol{x} - \rho_0\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{x} = m - \rho_0\boldsymbol{x}^\top\boldsymbol{W}\boldsymbol{x} \leq m + z_\alpha\sqrt{m},$$

which holds with probability of $100(1-\alpha)\%$. Consequently, an asymptotic upper bound for $||s||_2^2$ is

$$||\boldsymbol{s}||_2^2 \leq \frac{1}{\lambda_{\min}(\boldsymbol{D}-\rho_0\boldsymbol{W})}(m+z_\alpha\sqrt{m})$$

with probability of $100(1-\alpha)\%$.

It is easy to see that the matrix

$$\boldsymbol{I}_n - (\boldsymbol{D}-\rho_0\boldsymbol{W})^{1/2}\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top(\boldsymbol{D}-\rho_0\boldsymbol{W})^{1/2}$$

is a projection matrix, and thus

$$\begin{aligned}
&\boldsymbol{s}^\top\boldsymbol{W}\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top\boldsymbol{W}\boldsymbol{s} \\
=&\boldsymbol{s}^\top\boldsymbol{W}(\boldsymbol{D}-\rho_0\boldsymbol{W})^{-1/2}(\boldsymbol{D}-\rho_0\boldsymbol{W})^{1/2}\boldsymbol{F}\boldsymbol{A}^{-1}\boldsymbol{F}^\top(\boldsymbol{D}-\rho_0\boldsymbol{W})^{-1/2}(\boldsymbol{D}-\rho_0\boldsymbol{W})^{1/2}\boldsymbol{W}\boldsymbol{s} \\
\leq&\boldsymbol{s}^\top\boldsymbol{W}(\boldsymbol{D}-\rho_0\boldsymbol{W})^{-1}\boldsymbol{W}\boldsymbol{s} \leq \lambda_{\min}^{-1}(\boldsymbol{D}-\rho_0\boldsymbol{W})||\boldsymbol{W}\boldsymbol{s}||_2^2 \\
\leq&\lambda_{\min}^{-1}(\boldsymbol{D}-\rho_0\boldsymbol{W})||\boldsymbol{W}||_2^2||\boldsymbol{s}||_2^2 = \lambda_{\min}^{-1}(\boldsymbol{D}-\rho_0\boldsymbol{W})|\lambda(\boldsymbol{W})|_{\max}^2||\boldsymbol{s}||_2^2
\end{aligned}$$

The first inequality is due to Lemma S2. Here $|\lambda(\boldsymbol{W})|_{\max} = ||\boldsymbol{W}||_2$ is the spetrum radius of $\boldsymbol{W}$. Combining the previous steps we obtain the upper bound of the gap in (4). $\square$

S3. Proposition S2 and Its Proof

Proposition S2. *Let $x_1, \ldots, x_n$ of $\boldsymbol{x}$ are independent and identically distributed random variables from the discrete distribution with $\Pr(x_i = 1) = \Pr(x_i = -1) = 0.5$. For any two symmetric and non-zero $n \times n$ matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, we have that*

$$cor_{\boldsymbol{x}}(\boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}^\top\boldsymbol{B}\boldsymbol{x}) = \frac{\sum_{i<j} a_{ij}b_{ij}}{\sqrt{\sum_{i<j} a_{ij}^2}\sqrt{\sum_{i<j} b_{ij}^2}}, \tag{6}$$

where a_{ij} and b_{ij} are the (i,j)-th entries of matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ respectively.

Consider two $n\times n$ symmetric matrices $\boldsymbol{A}$ and $\boldsymbol{B}$. For random designs, we have that $\mathbb{E}(x_i) = 0$, $\mathrm{var}(x_i) = 1$, and $\mathrm{cov}(x_i, x_j) = 0$ for $i \neq j$. Therefore, $\mathrm{cov}(\boldsymbol{x}) = \boldsymbol{I}_n$ and

$$\begin{aligned}
\mathrm{cov}(\boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}^\top\boldsymbol{B}\boldsymbol{x}) &= \mathbb{E}(\boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x}\boldsymbol{x}^\top\boldsymbol{B}\boldsymbol{x}) - \mathbb{E}(\boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x})\mathbb{E}(\boldsymbol{x}^\top\boldsymbol{B}\boldsymbol{x}) \\
&= \mathbb{E}\left(\boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x}\boldsymbol{x}^\top\boldsymbol{B}\boldsymbol{x}\right) - \mathrm{tr}(\boldsymbol{A})\mathrm{tr}(\boldsymbol{B})
\end{aligned}$$

Note that

$$\boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x}\boldsymbol{x}^\top\boldsymbol{B}\boldsymbol{x} = (\boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x}) \otimes (\boldsymbol{x}^\top\boldsymbol{B}\boldsymbol{x}) = (\boldsymbol{x}^\top \otimes \boldsymbol{x}^\top)(\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{x}\otimes\boldsymbol{x}).$$

Then

$$\begin{aligned}
\boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x}\boldsymbol{x}^\top\boldsymbol{B}\boldsymbol{x} &= \mathrm{tr}(\boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x}\boldsymbol{x}^\top\boldsymbol{B}\boldsymbol{x}) = \mathrm{tr}((\boldsymbol{x}^\top \otimes \boldsymbol{x}^\top)(\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{x}\otimes\boldsymbol{x})) \\
&= \mathrm{tr}((\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{x}\otimes\boldsymbol{x})(\boldsymbol{x}^\top \otimes \boldsymbol{x}^\top)),
\end{aligned}$$

and thus

$$\mathbb{E}(\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}\boldsymbol{x}^\top \boldsymbol{B}\boldsymbol{x}) = \mathbb{E}(\mathrm{tr}(\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}\boldsymbol{x}^\top \boldsymbol{B}\boldsymbol{x})) = \mathrm{tr}((\boldsymbol{A}\otimes\boldsymbol{B})\mathbb{E}((\boldsymbol{x}\otimes\boldsymbol{x})(\boldsymbol{x}^\top\otimes\boldsymbol{x}^\top)))$$

We need to derive $\mathbb{E}((\boldsymbol{x}\otimes\boldsymbol{x})(\boldsymbol{x}^\top\otimes\boldsymbol{x}^\top))$. Note that $(\boldsymbol{x}\otimes\boldsymbol{x})(\boldsymbol{x}^\top\otimes\boldsymbol{x}^\top) = (\boldsymbol{x}\boldsymbol{x}^\top)\otimes(\boldsymbol{x}\boldsymbol{x}^\top)$ is an $n\times n$ block matrix, and the i,j-th block is $x_i x_j \boldsymbol{x}\boldsymbol{x}^\top$. The diagonal blocks are $\mathbb{E}(x_i^2\boldsymbol{x}\boldsymbol{x}^\top) = \boldsymbol{I}_n$ ($\mathbb{E}(x_i^4) = 1$). If $i \neq j$, $\mathbb{E}(x_i x_j \boldsymbol{x}\boldsymbol{x}^\top) = \boldsymbol{e}_i\boldsymbol{e}_j^\top + \boldsymbol{e}_j\boldsymbol{e}_i^\top$, where $\boldsymbol{e}_i$ is the element vector with i-th entry equal to 1 others 0 and $\boldsymbol{e}_i\boldsymbol{e}_j^\top + \boldsymbol{e}_j\boldsymbol{e}_i^\top$ is a matrix with (i,j)th and (j,i)th entries equal to 1 and the rest entries 0. Therefore, the resulting $n\times n$ block matrix should have diagonal blocks be an $n\times n$ identity matrix, and the (i,j)-th off-diagonal block be $\boldsymbol{e}_i\boldsymbol{e}_j^\top + \boldsymbol{e}_j\boldsymbol{e}_i^\top$. So we can decompose the block matrix to be

$$\begin{aligned}
&\mathbb{E}((\boldsymbol{x}\otimes\boldsymbol{x})(\boldsymbol{x}^\top\otimes\boldsymbol{x}^\top)) = \boldsymbol{I}_n\otimes\boldsymbol{I}_n + \sum_{i\neq j}(\boldsymbol{e}_i\boldsymbol{e}_j^\top)\otimes(\boldsymbol{e}_i\boldsymbol{e}_j^\top + \boldsymbol{e}_j\boldsymbol{e}_i^\top)\\
=&\boldsymbol{I}_n\otimes\boldsymbol{I}_n + \sum_{i\neq j}(\boldsymbol{e}_i\boldsymbol{e}_j^\top)\otimes(\boldsymbol{e}_i\boldsymbol{e}_j^\top) + \sum_{i\neq j}(\boldsymbol{e}_i\boldsymbol{e}_j^\top)\otimes(\boldsymbol{e}_j\boldsymbol{e}_i^\top)
\end{aligned}$$

Then

$$\begin{aligned}
&\mathrm{tr}\left[(\boldsymbol{A}\otimes\boldsymbol{B})\mathbb{E}[(\boldsymbol{x}\otimes\boldsymbol{x})(\boldsymbol{x}^\top\otimes\boldsymbol{x}^\top)]\right]\\
=&\mathrm{tr}\left[(\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{I}_n\otimes\boldsymbol{I}_n)\right] + \mathrm{tr}\left[\sum_{i\neq j}(\boldsymbol{A}\otimes\boldsymbol{B})[(\boldsymbol{e}_i\boldsymbol{e}_j^\top)\otimes(\boldsymbol{e}_i\boldsymbol{e}_j^\top)]\right] + \mathrm{tr}\left[\sum_{i\neq j}(\boldsymbol{A}\otimes\boldsymbol{B})[(\boldsymbol{e}_i\boldsymbol{e}_j^\top)\otimes(\boldsymbol{e}_j\boldsymbol{e}_i^\top)]\right]\\
=&\mathrm{tr}\left[\boldsymbol{A}\otimes\boldsymbol{B}\right] + \sum_{i\neq j}\mathrm{tr}\left[(\boldsymbol{A}\otimes\boldsymbol{B})[(\boldsymbol{e}_i\boldsymbol{e}_j^\top)\otimes(\boldsymbol{e}_i\boldsymbol{e}_j^\top)]\right] + \sum_{i\neq j}\mathrm{tr}\left[(\boldsymbol{A}\otimes\boldsymbol{B})[(\boldsymbol{e}_i\boldsymbol{e}_j^\top)\otimes(\boldsymbol{e}_j\boldsymbol{e}_i^\top)]\right]\\
=&\mathrm{tr}(\boldsymbol{A})\mathrm{tr}(\boldsymbol{B}) + \sum_{i\neq j}\mathrm{tr}\left[(\boldsymbol{A}\boldsymbol{e}_i\boldsymbol{e}_j^\top)\otimes(\boldsymbol{B}\boldsymbol{e}_i\boldsymbol{e}_j^\top)\right] + \sum_{i\neq j}\mathrm{tr}\left[(\boldsymbol{A}\boldsymbol{e}_i\boldsymbol{e}_j^\top)\otimes(\boldsymbol{B}\boldsymbol{e}_j\boldsymbol{e}_i^\top)\right]\\
=&\mathrm{tr}(\boldsymbol{A})\mathrm{tr}(\boldsymbol{B}) + 2\sum_{i\neq j}\mathrm{tr}\left[(\boldsymbol{A}\boldsymbol{e}_i\boldsymbol{e}_j^\top)\right]\mathrm{tr}\left[(\boldsymbol{B}\boldsymbol{e}_i\boldsymbol{e}_j^\top)\right]\\
=&\mathrm{tr}(\boldsymbol{A})\mathrm{tr}(\boldsymbol{B}) + 4\sum_{i<j}\boldsymbol{A}_{ij}\boldsymbol{B}_{ij},
\end{aligned}$$

where $\boldsymbol{A}_{ij}$ is the ij-th entry of matrix $\boldsymbol{A}$. Then

$$\mathrm{cov}(\boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}^\top\boldsymbol{B}\boldsymbol{x}) = 4\sum_{i<j}\boldsymbol{A}_{ij}\boldsymbol{B}_{ij}$$

Accordingly,

$$\mathrm{cor}(\boldsymbol{x}^\top\boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}^\top\boldsymbol{B}\boldsymbol{x}) = \frac{\sum_{i<j}\boldsymbol{A}_{ij}\boldsymbol{B}_{ij}}{\sqrt{\sum_{i<j}\boldsymbol{A}_{ij}^2}\sqrt{\sum_{i<j}\boldsymbol{B}_{ij}^2}}$$

S4. Proof of Theorem 2

We first provide a useful Lemma.

Lemma S3. *Let X and Y be two random variables taking values from $\{-1, 1\}$. If $\mathrm{cov}(X, Y) = 0$, then X and Y are independent.*

Proof. Let U and V be two Bernoulli random variables. We first show that if $\mathrm{cov}(U, V) = 0$, then U and V are independent.

Notice that

$$\Pr(\{U = 1\}\text{and}\{V = 1\}) = \Pr(UV = 1) = \mathbb{E}(UV)$$

$$\mathbb{E}(U) = \Pr(U = 1)$$

and

$$\mathbb{E}(V) = \Pr(V = 1).$$

If $\mathrm{cov}(U, V) = 0$,

$$\Pr(\{U = 1\}\text{and}\{V = 1\}) - \Pr(U = 1)\Pr(V = 1) = \mathbb{E}(UV) - \mathbb{E}(U)\mathbb{E}(V) = 0.$$

Similarly, we can show that

$$\Pr(\{U = 0\}\text{and}\{V = 1\}) - \Pr(U = 0)\Pr(V = 1) = 0,$$

$$\Pr(\{U = 0\}\text{and}\{V = 0\}) - \Pr(U = 0)\Pr(V = 0) = 0,$$

and

$$\Pr(\{U = 1\}\text{and}\{V = 0\}) - \Pr(U = 1)\Pr(V = 0) = 0,$$

which demonstrate that U and V are independent.

For X and Y, we have that $X = 2U - 1$ and $Y = 2V - 1$. The independence of U and V indicates the independence of X and Y. Also,

$$\mathrm{cov}(X, Y) = 4\mathrm{cov}(U, V).$$

Thus, the conclusion holds. □

Proof. Recall that $w_{ii} = 0$ for $i = 1, \ldots, n$. Therefore, we only need to consider the terms $w_{ij}x_ix_j$ with $i \neq j$. Notice that

$$\mathrm{cov}(x_ix_j, x_{i'}x_{j'}) = \mathbb{E}(x_ix_jx_{i'}x_{j'}) - \mathbb{E}(x_ix_j)\mathbb{E}(x_{i'}x_{j'}) = 0$$

for $i \neq i'$ and $j \neq j'$. Also,

$$\mathrm{cov}(x_ix_j, x_ix_{j'}) = \mathbb{E}(x_i^2x_jx_{j'}) - \mathbb{E}(x_ix_j)\mathbb{E}(x_ix_{j'}) = 0$$

for $j \neq j'$. According to Lemma S3, we have that x_ix_j and $x_ix_{j'}$ are independent, and x_ix_j and $x_{i'}x_{j'}$ are independent. Thus, $w_{ij}x_ix_j$'s with $w_{ij} \neq 0$ are i.i.d random variables with mean

$$\mathbb{E}(w_{ij}x_ix_j) = \mathbb{E}(x_i)\mathbb{E}(x_j) = 0,$$

and variance

$$\mathrm{var}(w_{ij}x_ix_j) = \mathbb{E}(x_i^2x_j^2) - (\mathbb{E}(x_ix_j))^2 = 1.$$

According to the central limit theorem, the conclusion holds. □

S5. Proof of Proposition 1

Proof. Notice that

$$\mathbb{E}\left(\boldsymbol{x}^\top \boldsymbol{K}\boldsymbol{x}\right) = \operatorname{tr}\left[\mathbb{E}\left(\boldsymbol{x}^\top \boldsymbol{K}\boldsymbol{x}\right)\right] = \mathbb{E}\left[\operatorname{tr}(\boldsymbol{x}^\top \boldsymbol{K}\boldsymbol{x})\right] = \mathbb{E}\left[\operatorname{tr}(\boldsymbol{K}\boldsymbol{x}\boldsymbol{x}^\top)\right] = \operatorname{tr}\left[\boldsymbol{K}\mathbb{E}(\boldsymbol{x}\boldsymbol{x}^\top)\right].$$

For completely random design, under the same assumption as in Theorem 2, we have that

$$\mathbb{E}(x_i x_j) = \mathbb{E}(x_i)\mathbb{E}(x_j) = 0 \quad \text{for} \quad i \neq j$$

and $\mathbb{E}(x_i^2) = 1$ for $i = 1, \ldots, n$. Thus, $\mathbb{E}(\boldsymbol{x}\boldsymbol{x}^\top) = \boldsymbol{I}_n$.

Now we consider the case where $\boldsymbol{x}$ is a random balanced design. If n is even, we have that

$$\mathbb{E}\left(x_i \sum_{j=1}^n x_j\right) = 0$$

since the balanced constraint gives $\sum_{j=1}^n x_j = 0$ directly. If n is odd, $n = 2h+1$ with h be a positive integer. Due to the balance constraint, $\sum_{i=1}^n x_i = 1$ or -1. We have that

$$\begin{aligned}
\mathbb{E}\left(x_i \sum_{j=1}^n x_j\right) &= \Pr\left(\sum_{i=1}^n x_i = 1\right)\mathbb{E}\left(x_i \sum_{j=1}^n x_j \middle| \sum_{i=1}^n x_i = 1\right) + \Pr\left(\sum_{i=1}^n x_i = -1\right)\mathbb{E}\left(x_i \sum_{j=1}^n x_j \middle| \sum_{i=1}^n x_i = -1\right) \\
&= \frac{1}{2}\mathbb{E}\left(x_i \middle| \sum_{j=1}^n x_j = 1\right) - \frac{1}{2}\mathbb{E}\left(x_i \middle| \sum_{j=1}^n x_j = -1\right).
\end{aligned}$$

Note that

$$\mathbb{E}\left(x_i \middle| \sum_{i=1}^n x_i = 1\right) = \Pr\left(x_i = 1 \middle| \sum_{i=1}^n x_i = 1\right) - \Pr\left(x_i = -1 \middle| \sum_{i=1}^n x_i = 1\right) = \frac{h+1}{2h+1} - \frac{h}{2h+1} = \frac{1}{n},$$

$$\mathbb{E}\left(x_i \middle| \sum_{j=1}^n x_j = -1\right) = \Pr\left(x_i = 1 \middle| \sum_{j=1}^n x_j = -1\right) - \Pr\left(x_i = -1 \middle| \sum_{j=1}^n x_j = -1\right) = \frac{h}{2h+1} - \frac{h+1}{2h+1} = -\frac{1}{n}.$$

Thus, $\mathbb{E}\left(x_i \sum_{j=1}^n x_j\right) = 1/n$.

Therefore,

$$\mathbb{E}\left(x_1 \sum_{j=1}^n x_j\right) = 1 + (n-1)\mathbb{E}(x_1 x_2)$$

which gives that

$$\mathbb{E}(x_1 x_2) = \begin{cases} -\frac{1}{n-1} & \text{if} \quad n \text{ is even} \\ -\frac{1}{n} \quad if & \text{if} \quad n \text{ is odd} \end{cases}.$$

This conclusion holds for $\mathbb{E}(x_i x_j)$ with any $i \neq j$. $\square$