Optimized Covariance Design for AB Test on Social Network under Interference
APPENDIX ONLY

Anonymous Author(s)
Affiliation
Address
`email`

A Proofs

A.1 Proof of Proposition 1

Firstly, we notice that

$$\mathbb{E}[z_i] = \frac{1}{2} \quad \mathrm{Var}[z_i] = E[z_i] - E[z_i]^2 = \frac{1}{4} \tag{1}$$

Then we substitute $Y_i(\boldsymbol{z})$ in HT estimator

$$\hat{\tau} = \frac{1}{n} \sum_{i \in [n]} \left(\left(\frac{z_i}{\mathbb{E}\left[z_i\right]} - \frac{(1-z_i)}{\mathbb{E}\left[1-z_i\right]} \right) Y_i(z) \right) \tag{2}$$

with our potential outcome model and take expectation w.r.t. treatment assignments,

$$\begin{aligned}
\mathbb{E}[\hat{\tau}] &= \frac{2}{n} \left(\sum_{i=1}^{n} \mathbb{E}[(z_i - (1-z_i))Y_i(z)] \right) \\
&= \frac{2}{n} \sum_{i=1}^{n} \left(\alpha_i \mathbb{E}[2z_i - 1] + \beta_i \mathbb{E}[z_i^2] + \gamma \sum_{j \in N_i} \mathbb{E}[(2z_i - 1)z_j] \right) \\
&= \frac{2}{n} \sum_{i=1}^{n} \left(\beta_i \mathbb{E}[z_i] + 2\gamma \, \mathrm{Cov}[z_i, \sum_{j \in N_i} z_j] \right) \\
&= \frac{1}{n} \sum_{i=1}^{n} \left(\beta_i + 4\gamma \, \mathrm{Cov}[z_i, \sum_{j \in N_i} z_j] \right)
\end{aligned} \tag{3}$$

Moreover, we can calculate the GATE under our potential outcome model

$$\tau = \frac{1}{n} \sum_{i=1}^{n} \left(Y_i(\mathbf{1}) - Y_i(\mathbf{0}) \right) = \frac{1}{n} \sum_{1=1}^{n} \left(\beta_i + \gamma d_i \right) \tag{4}$$

Hence, the bias of HT estimator is

$$\begin{aligned}
E[\hat{\tau}] - \tau &= \frac{1}{n} \sum_{i=1}^{n} \left(4\gamma \, \mathrm{Cov}[z_i, \sum_{j \in N_i} z_j] - \gamma d_i \right) \\
&= \frac{\gamma}{n} \sum_{i=1}^{n} \left(\frac{\mathrm{Cov}[z_i, \sum_{j \in N_i} z_j]}{\mathrm{Var}[z_i]} - d_i \right)
\end{aligned} \tag{5}$$

Submitted to 37th Conference on Neural Information Processing Systems (NeurIPS 2023). Do not distribute.

Then we derive the matrix form with cluster-level treatment vector $\boldsymbol{t}$.

$$
\begin{aligned}
\mathbb{E}[\hat{\tau}] - \tau &= \frac{\gamma}{n}\sum_{i=1}^{n}\left(\frac{\operatorname{Cov}[z_i, \sum_{k\in\mathcal{N}_i} z_k]}{\operatorname{Var}[z_i]} - d_i\right) \\
&= \frac{\gamma}{n}\left(4\sum_{i\neq j\in[K]}\operatorname{Cov}[t_i, t_j]\boldsymbol{C}_{ij} - \sum_{i\in[n]} d_i\right) \\
&= \frac{\gamma}{n}\left(4\sum_{i,j\in[K]}\operatorname{Cov}[t_i, t_j]\boldsymbol{C}_{ij} - \sum_{i,j\in[K]} \boldsymbol{C}_{ij}\right) \\
&= \frac{\gamma}{n}\left(4\operatorname{trace}(\boldsymbol{C}\operatorname{Cov}[\boldsymbol{t}]) - \sum_{i,j\in[K]} \boldsymbol{C}_{ij}\right)
\end{aligned} \tag{6}
$$

A.2 Proof of Proposition 2

Firstly we rewrite the estimator

$$
\begin{aligned}
\hat{\tau} &= \frac{2}{n}\sum_{i\in[n]}\left((\beta_i - \gamma d_i)z_i + 2\gamma\sum_{j\in N_i} z_i z_j\right) \\
&= \frac{2}{n}\left(\sum_{i\in[n]}(\beta_i - \gamma d_i)z_i + 2\gamma\sum_{(j,k)\in\mathcal{E}} z_i z_j\right)
\end{aligned} \tag{7}
$$

We then expand the variance directly, and get

$$
\begin{aligned}
\operatorname{Var}[\hat{\tau}] = \frac{4}{n^2}(&\sum_{i,j\in[n]}(\beta_i - \gamma d_i)(\beta_j - \gamma d_j)\operatorname{Cov}[z_i, z_j] \\
&+ \sum_{i\in[n]}\sum_{(j,k)\in\mathcal{E}} 4\gamma(\beta_i - \gamma d_i)\operatorname{Cov}[z_i, z_j z_k] \\
&+ \sum_{(i,j)\in\mathcal{E}}\sum_{(k,l)\in\mathcal{E}} 4\gamma^2\operatorname{Cov}[z_i z_j, z_k z_l])
\end{aligned} \tag{8}
$$

then we write it as matrix form with cluster-level treatment vector $\boldsymbol{t}$,

$$
\hat{\tau} = \frac{2}{n}(\boldsymbol{h}^T t + 2\gamma t^T \boldsymbol{C} t) \tag{9}
$$

and expand the variance as covariance, we get the desired format

$$
\begin{aligned}
\operatorname{Var}[\hat{\tau}] =& \frac{4}{n^2}(\operatorname{trace}(\boldsymbol{h}\boldsymbol{h}^T\operatorname{Cov}[\boldsymbol{t}]) + 4\gamma\operatorname{Cov}[\boldsymbol{h}^T\boldsymbol{t}, \boldsymbol{t}^T\boldsymbol{C}\boldsymbol{t}] \\
&+ 4\gamma^2\operatorname{Var}[\boldsymbol{t}^T\boldsymbol{C}\boldsymbol{t}])
\end{aligned} \tag{10}
$$

A.3 Proof of Proposition 3

Firstly we consider the estimator of matrix form again

$$
\hat{\tau} = \frac{2}{n}(\boldsymbol{h}^T t + 2\gamma t^T \boldsymbol{C} t) \tag{11}
$$

Since our assumption can't guarantee $\mathbb{E}[\hat{\tau}] > 0$ always hold, we drop the square of expectation term, namely

$$
\operatorname{Var}[\hat{\tau}] = \mathbb{E}[\hat{\tau}^2] - \mathbb{E}[\hat{\tau}]^2 \leq \mathbb{E}[\hat{\tau}^2] \tag{12}
$$

Notice that all elements of matrix $E[tt^T]$ is non-negative, we have

$$
\boldsymbol{h}^T\mathbb{E}[tt^T]\boldsymbol{h} \leq \omega^2\boldsymbol{d}^T\mathbb{E}[tt^T]\boldsymbol{d} = \omega^2\operatorname{trace}(\boldsymbol{d}\boldsymbol{d}^T\mathbb{E}[tt^T]) \tag{13}
$$

Similarly, since all elements of matrix $\boldsymbol{C}$, namely, $\boldsymbol{C_{ij}}$ is also non-negative, we have

$$\begin{aligned}\mathbb{E}[(t^T\boldsymbol{C}t)^2] &= \mathrm{trace}(\mathbb{E}[\boldsymbol{C}tt^T\boldsymbol{C}tt^T]) \\ &\leq \mathrm{trace}(\mathbb{E}[\boldsymbol{C}\mathbf{1}\mathbf{1}^T\boldsymbol{C}tt^T] \\ &= \mathrm{trace}(\boldsymbol{C}\mathbf{1}\mathbf{1}^T\boldsymbol{C}\mathbb{E}[tt^T])\end{aligned} \tag{14}$$

In summary, we have

$$\begin{aligned}\mathrm{Var}[\hat{\tau}] &\leq \mathbb{E}[\hat{\tau}^2] \\ &\leq \frac{8}{n^2}(\boldsymbol{h}^T\mathbb{E}[tt^T]\boldsymbol{h} + 4\gamma^2\mathbb{E}[(t^T\boldsymbol{C}t)^2]) \\ &\leq \frac{8\gamma^2}{n^2}\left(\omega^2\,\mathrm{trace}(\boldsymbol{d}\boldsymbol{d}^T\mathbb{E}[tt^T]) + 4\,\mathrm{trace}(\boldsymbol{C}\mathbf{1}\mathbf{1}^T\boldsymbol{C}\mathbb{E}[tt^T])\right)\end{aligned} \tag{15}$$

By definition, we have

$$\boldsymbol{C}\mathbf{1} = \boldsymbol{d} \tag{16}$$

Thus the upper bound above is actually

$$\mathrm{Var}[\hat{\tau}] \leq \frac{8\gamma^2(\omega^2+4)}{n^2}\left(\mathrm{trace}(\boldsymbol{d}\boldsymbol{d}^T\mathbb{E}[tt^T])\right) \tag{17}$$

At last, plug in following equation.

$$\mathbb{E}[tt^T] = \mathrm{Cov}[t] + \frac{1}{4}\mathbf{1}\mathbf{1}^T \tag{18}$$

A.4 Proof of Lemma 1

Here we provide an intuitive proof. Utilizing Box-Muller transformation or pure algebraic analysis are also feasible.

We consider X and Y is generated from multivariate Gaussian distribution,

$$X = \langle x, g\rangle \quad Y = \langle y, g\rangle \tag{19}$$

where $g \sim \mathcal{N}(0, I_n)$ and x, y are two n-dim real vectors. Then we know that

$$\mathrm{Cov}[X, Y] = \langle x, y\rangle \tag{20}$$

Moreover, we have

$$\mathbb{E}[\mathrm{sgn}(X)] = 0 \quad \mathbb{E}[\mathrm{sgn}(Y)] = 0 \tag{21}$$

thus

$$\mathrm{Cov}[\mathrm{sgn}(X), \mathrm{sgn}(Y)] = \mathbb{E}[\mathrm{sgn}(X)\,\mathrm{sgn}(Y)] \tag{22}$$

Then we think geometrically that $\mathrm{sgn}(\langle x, g\rangle)\,\mathrm{sgn}(\langle y, g\rangle) > 0$ holds iff. g lies above or below both of the hyperplanes that is orthogonal to x and y respectively.

Notice that the direction of g is uniform, it follows that

$$\mathbb{P}(\mathrm{sgn}(\langle x, g\rangle)\,\mathrm{sgn}(\langle y, g\rangle) > 0) = \frac{2}{2\pi}(\pi - \arccos(\langle x, y\rangle) \tag{23}$$

Now we put things together

$$\begin{aligned}\mathbb{E}[\mathrm{sgn}(X)\,\mathrm{sgn}(Y)] &= \mathbb{P}(\mathrm{sgn}(X)\,\mathrm{sgn}(Y) > 0) - \mathbb{P}(\mathrm{sgn}(X)\,\mathrm{sgn}(Y) < 0) \\ &= 2\mathbb{P}(\mathrm{sgn}(X)\,\mathrm{sgn}(Y) > 0) - 1 \\ &= 2(\frac{1}{\pi}(\pi - \arccos(\langle x, y\rangle)) - 1 \\ &= 1 - \frac{2}{\pi}\arccos(\langle x, y\rangle) \\ &= \frac{2}{\pi}\arcsin(\langle x, y\rangle)\end{aligned} \tag{24}$$

which gives the desired outcome.

B Simulation Details

B.1 Methods

We consider following linear potential outcome model,

$$Y_i(\boldsymbol{z}) = \alpha + \beta \cdot z_i + c \cdot \frac{d_i}{\bar{d}} + \sigma \cdot \epsilon_i + \gamma \frac{\sum_{j \in N_i} z_j}{d_i} \tag{25}$$

and multiplicative model, which is a simplified version of that in [6], with removing a covariate.

$$Y_i(\boldsymbol{z}) = (\alpha + \sigma \cdot \epsilon_i) \cdot \frac{d_i}{\bar{d}} \cdot (1 + \beta z_i + \gamma \frac{\sum_{j \in N_i} z_j}{d_i}) \tag{26}$$

We choose to fix all parameters except for interference density, γ. Namely, we set $(\alpha, \beta, c, \sigma) = (1, 1, 0.5, 0.1)$ for both models, and set $\gamma \in \{0.5, 1, 2\}$ to construct three regimes. $\epsilon_i \sim \mathcal{N}(0, 1)$.

We set the clusters as given by Louvain algorithm with fixed random seed and resolution parameter as $2, 5, 10$.

We consider social network FB-Stanford3[5],[1] and FB-Cornell5[5].[2]. These two social networks provide the network topology, and we generate potential outcome for each unit with mentioned potential outcome models.

Besides our optimized covariance design (OCD), we implement following randomization schemes. We first consider two baselines, independent Bernoulli randomization (Ber) and complete randomization (CR), both of which are cluster-level. We also implement two adaptive schemes, rerandomized-adaptive randomization (ReAR) and pairwise-sequential randomization (PSR) [3, 4], which balance heuristic covariates adaptively and act as competitive baselines, since the average degree is considered as a covariate in these two methods and exactly appear in both two of our models, explicitly. Then we implement the independent block randomization (IBR) [1] and heuristic version IBR-p that creates blocks with size 2.

In the methods mentioned above, ReAR is the only one concerned with hyperparameters setting. We set $(q, B, \alpha) = (0.85, 400, 0.1)$, which corresponds to the recommendation in the original paper.

We estimate GATE with standard HT estimator and difference-in-means (DIM) estimator, where the latter refers to

$$\hat{\tau}_{DIM} = \sum_{i \in [n]} \left(\left(\frac{z_i}{\sum_{j \in [n]} z_j} - \frac{(1 - z_i)}{\sum_{j \in [n]} (1 - z_j)} \right) Y_i(z) \right) \tag{27}$$

To summarize, we provide the bias, standard deviation and MSE of two estimators under two potential outcome models, three γ levels and two datasets. All of these three metrics are calculated by repetition of Monte Carlo simulation of randomization, 200 times. In the main paper we've presented the results of HT estimator on FB-Stanford3, and we'll present the detailed results in this section.

B.2 Discussion on Estimators

Here we also provide discussion on another popular estimator that's considered in existing literature, which is the HT estimator with exposure indicator. We consider the exposure condition that's fully treated or fully controlled in 1-hop neighborhood here, and the corresponding indicator is defined as $\delta_i(z_0) = \mathbb{I}\{\sum_{j \in N_i} z_j = d_i z_0\}$. The estimator is

$$\hat{\tau}' = \frac{1}{n} \sum_{i \in [n]} \left((\frac{\delta_i(1)}{\mathbb{E}[\delta_i(1)]} - \frac{\delta_i(0)}{\mathbb{E}[\delta_i(0)]}) Y_i(\boldsymbol{z}) \right) \tag{28}$$

We don't consider this estimator not only because of it's high variance resulted from low effective sample size, and but also its high calculation cost. For calculating it in our repeated simulations, we

[1]Network topology data can be found in `https://networkrepository.com/socfb-Stanford3.php`

[2]`https://networkrepository.com/socfb-Cornell5.php`

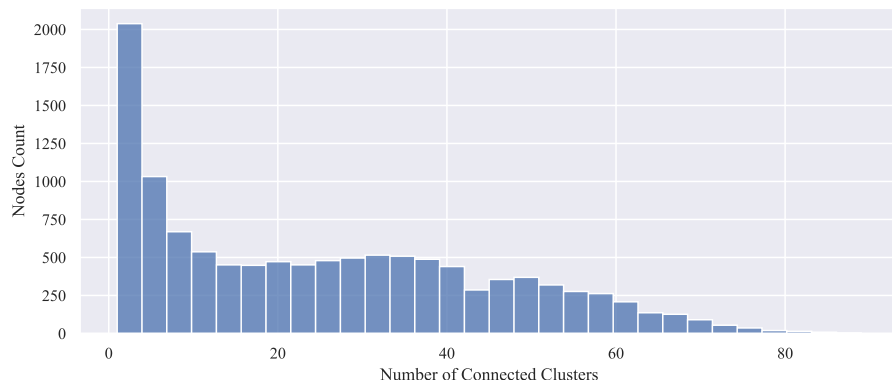

Figure 1: Distribution of c_i with 95 clusters on FB-Stanford3

need to estimate $\mathbb{E}[\delta_i(1)]$, the probability of fully treated, for each unit i. We denote the number of clusters a unit i connects to by c_i. For a unit i in the exterior of its cluster, namely, $c_i > 1$, such a quantity can very high in a social network, which is also decided by the resolution of clustering, we present an instance in figure 1.

Unfortunately, we must estimate such generalized propensity score [2] by Monte Carlo simulation for most of randomization schemes, even if it's just a little bit more complex than independent Bernoulli randomization, where such quantity can be calculated directly, $(1/2)^{c_i}$. For every simulation, we should visit every node and query the treatment assignment of its 1-hop neighborhood, whose time complexity is $O(|E|)$. Roughly, we need 2^{c_i} repetitions of randomization to achieve effective estimation on $\delta_i(1), \delta_i(0)$, which is too time-consuming to be acceptable for many units.

We stress the computational issue here since the resulted variance can be reduced by self-normalization, which corresponds to the Hájek estimator.

$$\hat{\tau}' = \sum_{i \in [n]} \left(\left(\frac{\frac{\delta_i(1)}{\mathbb{E}[\delta_i(1)]}}{\sum_{j \in [n]} \frac{\delta_j(1)}{\mathbb{E}[\delta_j(1)]}} \right) - \left(\frac{\frac{\delta_i(0)}{\mathbb{E}[\delta_i(0)]}}{\sum_{j \in [n]} \frac{\delta_j(0)}{\mathbb{E}[\delta_j(0)]}} \right) \right) Y_i(\boldsymbol{z}) \tag{29}$$

However, we argue that the computational cost can't be bypassed without further modification and restricts the practicality of such HT estimator with exposure indicator. [3] proposes the so-called cluster-adjusted estimator that assign the units on the exterior of same cluster the same modified propensity score, which reduces the computation complexity at the cost of unpredictable distortion of estimator, which can be viewed as heuristics.

Therefore, we choose to consider the standard HT estimator and DIM estimator in our simulation.

B.3 Detailed Simulation Results

We present the results according to following sequence: HT-DIM estimator, linear-multiplicative model, 2-5-10 clustering resolution. In every table, we report the average bias, standard deviation (SD) and MSE of randomization schemes with three γ levels.

Table 1: Simulation results of **HT** estimator under **linear** model with resolution **2** on FB-Stanford3

gamma		0.5			1.0			2.0	
metric	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
method									
Ber	-0.232	0.121	0.069	-0.506	0.126	0.272	-0.999	0.131	1.016
CR	-0.242	0.118	0.073	-0.481	0.120	0.247	-0.998	0.127	1.013
ReAR	-0.259	0.062	0.071	-0.491	0.059	0.245	-0.971	0.062	0.949
PSR	-0.240	0.058	0.061	-0.483	0.060	0.237	-0.973	0.058	0.951
IBR	-0.253	0.112	0.077	-0.496	0.103	0.257	-0.987	0.114	0.988
IBR-p	-0.248	0.086	0.069	-0.490	0.096	0.250	-0.986	0.093	0.981
OCD	-0.212	0.053	0.048	-0.423	0.056	0.182	-0.838	0.061	0.706

Table 2: Simulation results of **DIM** estimator under **linear** model with resolution **2** on FB-Stanford3

gamma		0.5			1.0			2.0	
metric	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
method									
Ber	-0.352	1.038	1.202	-0.451	1.089	1.388	-0.793	1.462	2.765
CR	-0.235	0.643	0.469	-0.444	0.722	0.719	-1.057	0.839	1.822
ReAR	-0.247	0.159	0.086	-0.544	0.147	0.318	-0.951	0.185	0.940
PSR	-0.244	0.236	0.115	-0.458	0.246	0.271	-0.972	0.286	1.028
IBR	-0.233	0.230	0.108	-0.460	0.246	0.273	-0.972	0.296	1.033
IBR-p	-0.239	0.158	0.082	-0.484	0.166	0.262	-0.971	0.193	0.980
OCD	-0.198	0.227	0.091	-0.440	0.280	0.272	-0.848	0.327	0.827

Table 3: Simulation results of **HT** estimator under **multiplicative** model with resolution **2** on FB-Stanford3

gamma		0.5			1.0			2.0	
metric	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
method									
Ber	-0.328	0.398	0.267	-0.620	0.481	0.617	-1.281	0.657	2.075
CR	-0.247	0.440	0.254	-0.585	0.495	0.588	-1.163	0.617	1.734
ReAR	-0.290	0.209	0.128	-0.542	0.219	0.343	-1.308	0.312	1.810
PSR	-0.318	0.207	0.144	-0.617	0.241	0.440	-1.208	0.309	1.556
IBR	-0.311	0.362	0.228	-0.640	0.401	0.570	-1.230	0.548	1.816
IBR-p	-0.324	0.303	0.197	-0.613	0.355	0.503	-1.175	0.436	1.572
OCD	-0.270	0.178	0.105	-0.549	0.210	0.346	-1.060	0.270	1.198

Table 4: Simulation results of **DIM** estimator under **multiplicative** model with resolution **2** on FB-Stanford3

gamma		0.5			1.0			2.0	
metric	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
method									
Ber	-0.262	0.799	0.707	-0.585	0.893	1.141	-1.169	1.233	2.887
CR	-0.293	0.493	0.329	-0.592	0.607	0.719	-1.117	0.699	1.737
ReAR	-0.276	0.232	0.130	-0.519	0.307	0.363	-1.308	0.403	1.874
PSR	-0.312	0.309	0.193	-0.612	0.364	0.508	-1.199	0.477	1.666
IBR	-0.315	0.348	0.221	-0.638	0.370	0.544	-1.240	0.501	1.789
IBR-p	-0.330	0.287	0.192	-0.595	0.336	0.468	-1.207	0.412	1.629
OCD	-0.293	0.166	0.114	-0.533	0.178	0.316	-1.086	0.236	1.236

Table 5: Simulation results of **HT** estimator under **linear** model with resolution **5** on FB-Stanford3

gamma metric method	Bias	0.5 SD	MSE	Bias	1.0 SD	MSE	Bias	2.0 SD	MSE
Ber	-0.278	0.092	0.086	-0.569	0.089	0.332	-1.111	0.097	1.244
CR	-0.274	0.089	0.083	-0.559	0.093	0.321	-1.124	0.100	1.274
ReAR	-0.307	0.031	0.095	-0.581	0.045	0.340	-1.136	0.057	1.294
PSR	-0.275	0.049	0.078	-0.554	0.053	0.310	-1.113	0.058	1.243
IBR	-0.274	0.087	0.083	-0.550	0.090	0.311	-1.111	0.090	1.243
IBR-p	-0.200	0.073	0.045	-0.492	0.073	0.248	-1.053	0.075	1.115
OCD	-0.199	0.108	0.051	-0.409	0.106	0.179	-0.788	0.106	0.632

Table 6: Simulation results of **DIM** estimator under **linear** model with resolution **5** on FB-Stanford3

gamma metric method	Bias	0.5 SD	MSE	Bias	1.0 SD	MSE	Bias	2.0 SD	MSE
Ber	-0.254	0.672	0.516	-0.420	0.694	0.659	-1.037	0.796	1.710
CR	-0.308	0.453	0.300	-0.643	0.543	0.709	-1.161	0.674	1.802
ReAR	-0.045	0.093	0.011	-0.361	0.182	0.164	-0.904	0.248	0.880
PSR	-0.262	0.226	0.120	-0.554	0.252	0.371	-1.122	0.306	1.354
IBR	-0.302	0.269	0.164	-0.531	0.312	0.380	-1.094	0.380	1.341
IBR-p	-0.696	0.063	0.489	-1.028	0.075	1.064	-1.685	0.096	2.852
OCD	-0.214	0.371	0.183	-0.376	0.411	0.311	-0.857	0.510	0.996

Table 7: Simulation results of **HT** estimator under **multiplicative** model with resolution **5** on FB-Stanford3

gamma metric method	Bias	0.5 SD	MSE	Bias	1.0 SD	MSE	Bias	2.0 SD	MSE
Ber	-0.366	0.333	0.245	-0.684	0.391	0.621	-1.321	0.486	1.984
CR	-0.329	0.348	0.230	-0.717	0.390	0.667	-1.378	0.483	2.134
ReAR	-0.440	0.130	0.211	-0.822	0.177	0.707	-1.614	0.181	2.639
PSR	-0.344	0.186	0.153	-0.667	0.219	0.493	-1.346	0.273	1.887
IBR	-0.322	0.300	0.194	-0.693	0.332	0.590	-1.396	0.464	2.165
IBR-p	-0.080	0.265	0.077	-0.407	0.315	0.265	-1.017	0.423	1.214
OCD	-0.258	0.354	0.193	-0.506	0.418	0.432	-1.059	0.505	1.378

Table 8: Simulation results of **DIM** estimator under **multiplicative** model with resolution **5** on FB-Stanford3

gamma metric method	Bias	0.5 SD	MSE	Bias	1.0 SD	MSE	Bias	2.0 SD	MSE
Ber	-0.357	0.462	0.342	-0.683	0.514	0.732	-1.349	0.637	2.227
CR	-0.383	0.288	0.230	-0.706	0.319	0.601	-1.429	0.429	2.228
ReAR	-0.261	0.120	0.083	-0.655	0.135	0.447	-1.377	0.175	1.927
PSR	-0.339	0.132	0.133	-0.674	0.145	0.476	-1.369	0.172	1.906
IBR	-0.335	0.172	0.142	-0.713	0.190	0.545	-1.401	0.259	2.031
IBR-p	-0.473	0.184	0.258	-0.847	0.219	0.766	-1.571	0.294	2.556
OCD	-0.255	0.052	0.068	-0.519	0.076	0.276	-1.043	0.101	1.100

Table 9: Simulation results of **HT** estimator under **linear** model with resolution **10** on FB-Stanford3

gamma		0.5			1.0			2.0	
metric	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
method									
Ber	-0.300	0.078	0.096	-0.585	0.076	0.348	-1.190	0.078	1.424
CR	-0.300	0.077	0.096	-0.596	0.077	0.362	-1.192	0.081	1.429
ReAR	-0.296	0.022	0.088	-0.592	0.025	0.351	-1.201	0.022	1.444
PSR	-0.293	0.036	0.087	-0.595	0.038	0.356	-1.192	0.038	1.422
IBR	-0.294	0.059	0.090	-0.593	0.052	0.354	-1.191	0.058	1.424
IBR-p	-0.295	0.047	0.090	-0.590	0.051	0.351	-1.191	0.052	1.423
OCD	-0.189	0.099	0.046	-0.396	0.102	0.168	-0.796	0.107	0.646

Table 10: Simulation results of **DIM** estimator under **linear** model with resolution **10** on FB-Stanford3

gamma		0.5			1.0			2.0	
metric	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
method									
Ber	-0.330	0.527	0.387	-0.567	0.586	0.665	-1.200	0.714	1.950
CR	-0.265	0.423	0.250	-0.548	0.450	0.504	-1.157	0.551	1.643
ReAR	-0.364	0.235	0.188	-0.688	0.262	0.543	-1.281	0.318	1.744
PSR	-0.321	0.234	0.158	-0.605	0.263	0.436	-1.191	0.325	1.526
IBR	-0.299	0.266	0.160	-0.571	0.306	0.420	-1.171	0.370	1.510
IBR-p	-0.305	0.232	0.147	-0.594	0.264	0.423	-1.156	0.316	1.437
OCD	-0.240	0.354	0.183	-0.395	0.417	0.330	-0.771	0.539	0.885

Table 11: Simulation results of **HT** estimator under **multiplicative** model with resolution **10** on FB-Stanford3

gamma		0.5			1.0			2.0	
metric	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
method									
Ber	-0.348	0.258	0.188	-0.766	0.307	0.682	-1.446	0.390	2.244
CR	-0.375	0.264	0.211	-0.729	0.322	0.636	-1.504	0.369	2.400
ReAR	-0.353	0.074	0.130	-0.725	0.094	0.536	-1.426	0.117	2.050
PSR	-0.384	0.120	0.162	-0.755	0.129	0.588	-1.489	0.179	2.249
IBR	-0.402	0.172	0.192	-0.742	0.227	0.603	-1.466	0.274	2.226
IBR-p	-0.360	0.154	0.154	-0.739	0.186	0.581	-1.468	0.238	2.212
OCD	-0.253	0.355	0.190	-0.546	0.396	0.455	-1.060	0.499	1.374

Table 12: Simulation results of **DIM** estimator under **multiplicative** model with resolution **10** on FB-Stanford3

gamma		0.5			1.0			2.0	
metric	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
method									
Ber	-0.407	0.326	0.272	-0.750	0.383	0.710	-1.463	0.480	2.373
CR	-0.336	0.232	0.167	-0.719	0.269	0.590	-1.474	0.348	2.297
ReAR	-0.416	0.152	0.197	-0.763	0.174	0.613	-1.546	0.234	2.446
PSR	-0.356	0.137	0.146	-0.747	0.162	0.585	-1.446	0.188	2.129
IBR	-0.361	0.160	0.157	-0.726	0.174	0.558	-1.497	0.228	2.295
IBR-p	-0.375	0.166	0.169	-0.750	0.189	0.600	-1.480	0.238	2.248
OCD	-0.264	0.059	0.074	-0.527	0.060	0.282	-1.051	0.083	1.113

Table 13: Simulation results of **HT** estimator under **linear** model with resolution **2** on FB-Cornell5

gamma		0.5			1.0			2.0	
metric	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
method									
Ber	-0.251	0.129	0.080	-0.506	0.122	0.271	-0.996	0.124	1.008
CR	-0.244	0.111	0.072	-0.502	0.114	0.265	-0.992	0.129	1.002
ReAR	-0.233	0.057	0.058	-0.484	0.065	0.239	-0.981	0.065	0.968
PSR	-0.242	0.058	0.062	-0.483	0.059	0.237	-0.977	0.061	0.960
IBR	-0.251	0.103	0.074	-0.493	0.104	0.255	-1.001	0.117	1.016
IBR-p	-0.235	0.087	0.063	-0.488	0.090	0.246	-0.986	0.095	0.982
OCD	-0.184	0.085	0.041	-0.371	0.083	0.145	-0.769	0.087	0.600

Table 14: Simulation results of **DIM** estimator under **linear** model with resolution **2** on FB-Cornell5

gamma		0.5			1.0			2.0	
metric	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
method									
Ber	-0.166	1.048	1.126	-0.214	1.101	1.258	-1.003	1.335	2.790
CR	-0.228	0.611	0.425	-0.465	0.679	0.678	-0.938	0.838	1.582
ReAR	-0.255	0.172	0.095	-0.462	0.159	0.239	-0.975	0.190	0.987
PSR	-0.239	0.234	0.112	-0.492	0.261	0.311	-1.006	0.298	1.102
IBR	-0.236	0.227	0.107	-0.501	0.253	0.315	-0.983	0.291	1.051
IBR-p	-0.242	0.149	0.081	-0.504	0.164	0.282	-0.985	0.209	1.015
OCD	-0.205	0.318	0.143	-0.422	0.348	0.300	-0.723	0.476	0.750

Table 15: Simulation results of **HT** estimator under **multiplicative** model with resolution **2** on FB-Cornell5

gamma		0.5			1.0			2.0	
metric	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
method									
Ber	-0.360	0.451	0.333	-0.616	0.513	0.643	-1.241	0.591	1.890
CR	-0.303	0.418	0.267	-0.653	0.453	0.633	-1.274	0.575	1.955
ReAR	-0.317	0.201	0.141	-0.685	0.237	0.526	-1.231	0.308	1.612
PSR	-0.323	0.199	0.144	-0.608	0.235	0.426	-1.217	0.291	1.568
IBR	-0.292	0.368	0.221	-0.589	0.410	0.516	-1.211	0.509	1.728
IBR-p	-0.315	0.318	0.201	-0.606	0.385	0.516	-1.190	0.454	1.623
OCD	-0.266	0.285	0.152	-0.529	0.327	0.388	-0.984	0.410	1.137

Table 16: Simulation results of **DIM** estimator under **multiplicative** model with resolution **2** on FB-Cornell5

gamma		0.5			1.0			2.0	
metric	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
method									
Ber	-0.381	0.792	0.773	-0.602	0.919	1.208	-1.171	1.130	2.651
CR	-0.344	0.462	0.332	-0.634	0.578	0.737	-1.206	0.722	1.977
ReAR	-0.334	0.221	0.161	-0.692	0.294	0.566	-1.235	0.329	1.634
PSR	-0.311	0.325	0.203	-0.614	0.353	0.502	-1.201	0.465	1.659
IBR	-0.279	0.336	0.191	-0.608	0.367	0.505	-1.225	0.471	1.723
IBR-p	-0.328	0.288	0.191	-0.612	0.357	0.502	-1.184	0.410	1.570
OCD	-0.250	0.070	0.068	-0.496	0.037	0.248	-0.998	0.068	1.002

Table 17: Simulation results of **HT** estimator under **linear** model with resolution **5** on FB-Cornell5

gamma metric method	Bias	0.5 SD	MSE	Bias	1.0 SD	MSE	Bias	2.0 SD	MSE
Ber	-0.263	0.098	0.079	-0.567	0.095	0.331	-1.121	0.105	1.270
CR	-0.268	0.092	0.080	-0.553	0.093	0.315	-1.111	0.090	1.244
ReAR	-0.307	0.036	0.096	-0.593	0.049	0.354	-1.138	0.038	1.298
PSR	-0.279	0.056	0.082	-0.551	0.055	0.307	-1.110	0.056	1.237
IBR	-0.273	0.088	0.083	-0.544	0.080	0.303	-1.113	0.091	1.248
IBR-p	-0.210	0.073	0.050	-0.493	0.074	0.249	-1.048	0.079	1.105
OCD	-0.203	0.103	0.052	-0.401	0.109	0.173	-0.792	0.103	0.639

Table 18: Simulation results of **DIM** estimator under **linear** model with resolution **5** on FB-Cornell5

gamma metric method	Bias	0.5 SD	MSE	Bias	1.0 SD	MSE	Bias	2.0 SD	MSE
Ber	-0.380	0.649	0.567	-0.504	0.783	0.867	-1.090	0.923	2.042
CR	-0.375	0.502	0.393	-0.656	0.533	0.716	-1.159	0.663	1.784
ReAR	-0.075	0.127	0.022	-0.373	0.175	0.171	-0.855	0.216	0.779
PSR	-0.248	0.220	0.110	-0.551	0.250	0.367	-1.082	0.313	1.269
IBR	-0.277	0.284	0.158	-0.572	0.319	0.429	-1.094	0.365	1.331
IBR-p	-0.689	0.066	0.479	-1.025	0.077	1.058	-1.682	0.091	2.839
OCD	-0.178	0.368	0.168	-0.384	0.425	0.328	-0.820	0.525	0.949

Table 19: Simulation results of **HT** estimator under **multiplicative** model with resolution **5** on FB-Cornell5

gamma metric method	Bias	0.5 SD	MSE	Bias	1.0 SD	MSE	Bias	2.0 SD	MSE
Ber	-0.317	0.333	0.212	-0.691	0.357	0.605	-1.432	0.461	2.263
CR	-0.332	0.329	0.219	-0.688	0.396	0.632	-1.386	0.480	2.153
ReAR	-0.456	0.137	0.228	-0.816	0.142	0.686	-1.550	0.231	2.457
PSR	-0.344	0.193	0.156	-0.669	0.222	0.497	-1.361	0.287	1.937
IBR	-0.345	0.304	0.212	-0.697	0.349	0.608	-1.368	0.432	2.061
IBR-p	-0.115	0.287	0.096	-0.425	0.300	0.271	-1.038	0.421	1.256
OCD	-0.218	0.375	0.188	-0.488	0.427	0.421	-1.042	0.542	1.380

Table 20: Simulation results of **DIM** estimator under **multiplicative** model with resolution **5** on FB-Cornell5

gamma metric method	Bias	0.5 SD	MSE	Bias	1.0 SD	MSE	Bias	2.0 SD	MSE
Ber	-0.305	0.464	0.308	-0.670	0.504	0.704	-1.337	0.622	2.175
CR	-0.354	0.277	0.203	-0.719	0.348	0.639	-1.448	0.428	2.282
ReAR	-0.291	0.091	0.093	-0.619	0.125	0.399	-1.295	0.199	1.718
PSR	-0.320	0.127	0.118	-0.672	0.141	0.472	-1.360	0.180	1.883
IBR	-0.349	0.187	0.157	-0.677	0.216	0.506	-1.369	0.270	1.948
IBR-p	-0.500	0.206	0.293	-0.863	0.208	0.790	-1.590	0.294	2.617
OCD	-0.258	0.031	0.068	-0.526	0.081	0.284	-1.033	0.065	1.071

Table 21: Simulation results of **HT** estimator under **linear** model with resolution **10** on FB-Cornell5

gamma		0.5			1.0			2.0	
metric	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
method									
Ber	-0.301	0.080	0.098	-0.600	0.072	0.365	-1.175	0.078	1.387
CR	-0.294	0.076	0.093	-0.603	0.077	0.370	-1.192	0.083	1.428
ReAR	-0.294	0.023	0.087	-0.600	0.021	0.361	-1.196	0.027	1.432
PSR	-0.299	0.034	0.091	-0.597	0.036	0.358	-1.195	0.036	1.431
IBR	-0.294	0.054	0.090	-0.596	0.056	0.359	-1.186	0.059	1.411
IBR-p	-0.302	0.046	0.094	-0.598	0.044	0.360	-1.186	0.056	1.410
OCD	-0.192	0.118	0.051	-0.377	0.118	0.157	-0.762	0.123	0.597

Table 22: Simulation results of **DIM** estimator under **linear** model with resolution **10** on FB-Cornell5

gamma		0.5			1.0			2.0	
metric	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
method									
Ber	-0.292	0.541	0.378	-0.536	0.552	0.593	-1.291	0.740	2.214
CR	-0.286	0.399	0.242	-0.547	0.450	0.503	-1.194	0.580	1.764
ReAR	-0.328	0.243	0.167	-0.626	0.265	0.463	-1.288	0.310	1.757
PSR	-0.271	0.234	0.129	-0.566	0.263	0.389	-1.168	0.326	1.471
IBR	-0.302	0.278	0.169	-0.582	0.309	0.435	-1.169	0.377	1.509
IBR-p	-0.288	0.233	0.137	-0.598	0.264	0.428	-1.216	0.308	1.575
OCD	-0.193	0.407	0.203	-0.416	0.464	0.388	-0.834	0.577	1.029

Table 23: Simulation results of **HT** estimator under **multiplicative** model with resolution **10** on FB-Cornell5

gamma		0.5			1.0			2.0	
metric	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
method									
Ber	-0.348	0.275	0.197	-0.710	0.299	0.595	-1.477	0.365	2.317
CR	-0.348	0.249	0.184	-0.775	0.307	0.696	-1.482	0.380	2.342
ReAR	-0.367	0.074	0.141	-0.747	0.096	0.569	-1.468	0.106	2.166
PSR	-0.363	0.125	0.148	-0.748	0.137	0.579	-1.492	0.175	2.257
IBR	-0.399	0.182	0.193	-0.741	0.215	0.597	-1.514	0.277	2.370
IBR-p	-0.389	0.173	0.181	-0.755	0.188	0.606	-1.489	0.231	2.273
OCD	-0.248	0.404	0.225	-0.452	0.461	0.416	-0.969	0.589	1.286

Table 24: Simulation results of **DIM** estimator under **multiplicative** model with resolution **10** on FB-Cornell5

gamma		0.5			1.0			2.0	
metric	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
method									
Ber	-0.336	0.320	0.216	-0.758	0.399	0.734	-1.468	0.482	2.389
CR	-0.399	0.212	0.205	-0.709	0.265	0.574	-1.448	0.348	2.218
ReAR	-0.429	0.147	0.206	-0.798	0.191	0.674	-1.516	0.245	2.361
PSR	-0.368	0.132	0.153	-0.743	0.157	0.578	-1.489	0.187	2.255
IBR	-0.383	0.157	0.171	-0.711	0.191	0.542	-1.483	0.214	2.245
IBR-p	-0.372	0.157	0.164	-0.751	0.201	0.605	-1.468	0.227	2.207
OCD	-0.260	0.036	0.069	-0.519	0.035	0.271	-1.031	0.034	1.065

References

[1] Ozan Candogan, Chen Chen, and Rad Niazadeh. Correlated cluster-based randomized experiments: Robust variance minimization. *Chicago Booth Research Paper (21-17)*, 2021.

[2] Dean Eckles, Brian Karrer, and Johan Ugander. Design and analysis of experiments in networks: Reducing bias from interference. *Journal of Causal Inference*, 5(1):20150021, 2016.

[3] Yang Liu, Yifan Zhou, Ping Li, and Feifang Hu. Adaptive a/b test on networks with cluster structures. In *International Conference on Artificial Intelligence and Statistics*, pages 10836–10851. PMLR, 2022.

[4] Yichen Qin, Yang Li, Wei Ma, and Feifang Hu. Pairwise sequential randomization and its properties. *arXiv preprint arXiv:1611.02802*, 2016.

[5] Ryan A. Rossi and Nesreen K. Ahmed. The network data repository with interactive graph analytics and visualization. In *AAAI*, 2015.

[6] Johan Ugander and Hao Yin. Randomized graph cluster randomization. *arXiv preprint arXiv:2009.02297*, 2020.