CISER Causal Inference Workshop

Session 4: Interference and Spillover Effects

Chencheng Cai

Department of Mathematics and Statistics

Washington State University

Apr. 2, 2025

SUTVA Assumption

The Stable Unit Treatment Value Assumption (SUTVA) is a fundamental assumption in causal inference.

- The potential outcomes for each unit do not depend on the treatment assignment of other units.

SUTVA Assumption

The Stable Unit Treatment Value Assumption (SUTVA) is a fundamental assumption in causal inference.

- The potential outcomes for each unit do not depend on the treatment assignment of other units.

SUTVA is violated when the potential outcomes of one unit depend on the treatment assignment of other units.
Such phoneomena are often referred to as **interference** or **spillover effects**.

- Social networks
- Transportation networks
- Field experiments

Potential Outcomes

- Let $\boldsymbol{Z} = (Z_1, \ldots, Z_N) \in \{0, 1\}^N$ be the treatment assignment vector for all N units.
- Let $Y_i(\boldsymbol{Z})$ be the potential outcome for unit i given treatment assignment vector $\boldsymbol{Z}$.

Potential Outcomes

- Let $\boldsymbol{Z} = (Z_1, \ldots, Z_N) \in \{0,1\}^N$ be the treatment assignment vector for all N units.
- Let $Y_i(\boldsymbol{Z})$ be the potential outcome for unit i given treatment assignment vector $\boldsymbol{Z}$.
- SUTVA assumes that for all i,

$$Y_i(\boldsymbol{Z}) = Y_i(\boldsymbol{Z}') \quad \text{whenever } Z_i = Z_i'$$

The consequence is we can simply write $Y_i(Z_i)$ instead of $Y_i(\boldsymbol{Z})$.

Potential Outcomes

- Let $\boldsymbol{Z} = (Z_1, \ldots, Z_N) \in \{0, 1\}^N$ be the treatment assignment vector for all N units.
- Let $Y_i(\boldsymbol{Z})$ be the potential outcome for unit i given treatment assignment vector $\boldsymbol{Z}$.
- SUTVA assumes that for all i,

  $$Y_i(\boldsymbol{Z}) = Y_i(\boldsymbol{Z}') \quad \text{whenever } Z_i = Z_i'$$

  The consequence is we can simply write $Y_i(Z_i)$ instead of $Y_i(\boldsymbol{Z})$.
- If SUTVA is violated, we cannot write $Y_i(\boldsymbol{Z})$ as $Y_i(Z_i)$ because Y_i depends on the treatment assignment of other units.

Interference

To represent the dependence of Y_i on $Z_{i'}$ for $i' \neq i$, we consider a directed graph $\mathfrak{G}$.

- Vertices: $V = \{1, \ldots, N\}$, representing the units.
- Edges: $E = \{(i, i') : Y_i \text{ depends on } Z_{i'}\}$.
- The indegree neighbor of unit i is defined as

$$\mathcal{N}_i = \{i' : (i', i) \in E\}.$$

- $\mathfrak{G}$ is often assumed to be observed and fixed for the dataset.

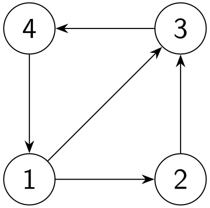
Interference

To represent the dependence of Y_i on $Z_{i'}$ for $i' \neq i$, we consider a directed graph $\mathfrak{G}$.

- Vertices: $V = \{1, \ldots, N\}$, representing the units.
- Edges: $E = \{(i, i') : Y_i \text{ depends on } Z_{i'}\}$.
- The indegree neighbor of unit i is defined as

$$\mathcal{N}_i = \{i' : (i', i) \in E\}.$$

- $\mathfrak{G}$ is often assumed to be observed and fixed for the dataset.



$$\mathcal{N}_1 = \{4\}$$
$$\mathcal{N}_2 = \{1\}$$
$$\mathcal{N}_3 = \{1, 2\}$$
$$\mathcal{N}_4 = \{3\}$$

SUTNVA Assumption

When the interference is present, we assume the **Stable Unit Treatment on Neighborhood Value Assumption** (SUTNVA):

1. For each i, for any two treatment assignments $\boldsymbol{Z} = (Z_i, Z_{\mathcal{N}_i}, Z_{\mathcal{N}_{-i}})$ and $\boldsymbol{Z}' = (Z'_i, Z'_{\mathcal{N}_i}, Z'_{\mathcal{N}_{-i}})$, we have

$$Y_i(\boldsymbol{Z}) = Y_i(\boldsymbol{Z}') \quad \text{whenever } Z_i = Z'_i \text{ and } Z_{\mathcal{N}_i} = Z'_{\mathcal{N}_i}$$

2. For each i,

$$Y_i^{obs} = Y_i(\boldsymbol{Z})$$

SUTNVA Assumption

When the interference is present, we assume the **Stable Unit Treatment on Neighborhood Value Assumption** (SUTNVA):

1. For each i, for any two treatment assignments $\boldsymbol{Z} = (Z_i, Z_{\mathcal{N}_i}, Z_{\mathcal{N}_{-i}})$ and $\boldsymbol{Z}' = (Z'_i, Z'_{\mathcal{N}_i}, Z'_{\mathcal{N}_{-i}})$, we have

$$Y_i(\boldsymbol{Z}) = Y_i(\boldsymbol{Z}') \quad \text{whenever } Z_i = Z'_i \text{ and } Z_{\mathcal{N}_i} = Z'_{\mathcal{N}_i}$$

2. For each i,

$$Y_i^{obs} = Y_i(\boldsymbol{Z})$$

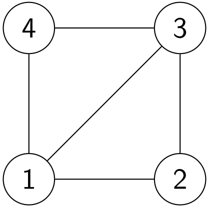
Under SUTNVA, we can write the potential outcome as

$$Y_i(\underbrace{Z_i}_{\text{direct treatment}}, \underbrace{Z_{\mathcal{N}_i}}_{\text{interference}})$$

Potential Outcomes

For demonstration purpose, we consider a simpler case the the interference graph $\mathcal{G}$ is undirected.

Inteference Graph



Neighbors

$$\mathcal{N}_1 = \{2, 3, 4\}$$
$$\mathcal{N}_2 = \{1, 3\}$$
$$\mathcal{N}_3 = \{1, 2, 4\}$$
$$\mathcal{N}_4 = \{1, 3\}$$

Potential Outcomes

$$Y_1^{obs} = Y_1(Z_1, Z_2, Z_3, Z_4)$$
$$Y_2^{obs} = Y_2(Z_2, Z_1, Z_3)$$
$$Y_3^{obs} = Y_3(Z_3, Z_1, Z_2, Z_4)$$
$$Y_4^{obs} = Y_4(Z_4, Z_1, Z_3)$$

Exposure Mapping

The **exposure mapping** is a function $g_i : \{0,1\}^{N_i} \to \mathcal{G}_i$ for all i, such that SUTNVA holds for $G_i = g_i(Z_{\mathcal{N}_i})$:

$$Y_i(\boldsymbol{Z}) = Y_i(\boldsymbol{Z}') \quad \text{whenever } Z_i = Z_i' \text{ and } g_i(Z_{\mathcal{N}_i}) = g_i(Z'_{\mathcal{N}_i})$$

Exposure Mapping

The **exposure mapping** is a function $g_i : \{0,1\}^{N_i} \to \mathcal{G}_i$ for all i, such that SUTNVA holds for $G_i = g_i(Z_{\mathcal{N}_i})$:

$$Y_i(\boldsymbol{Z}) = Y_i(\boldsymbol{Z}') \quad \text{whenever } Z_i = Z_i' \text{ and } g_i(Z_{\mathcal{N}_i}) = g_i(Z'_{\mathcal{N}_i})$$

We can write the potential outcomes under interference as

$$Y_i(Z_i, G_i) \quad \text{with} \quad G_i = g_i(Z_{\mathcal{N}_i}).$$

Exposure Mapping

Common choices of exposure mapping:

- number of treated neighbors:

$$G_i = \sum_{i' \in \mathcal{N}_i} Z_{i'}$$

- proportion of treated neighbors:

$$G_i = N_i^{-1} \sum_{i' \in \mathcal{N}_i} Z_{i'},$$

  where $N_i = |\mathcal{N}_i|$ is the number of neighbors of unit i.

- heterogeneous interference from neighbors:

$$G_i = \sum_{i' \in \mathcal{N}_i} w_{ii'} Z_{i'},$$

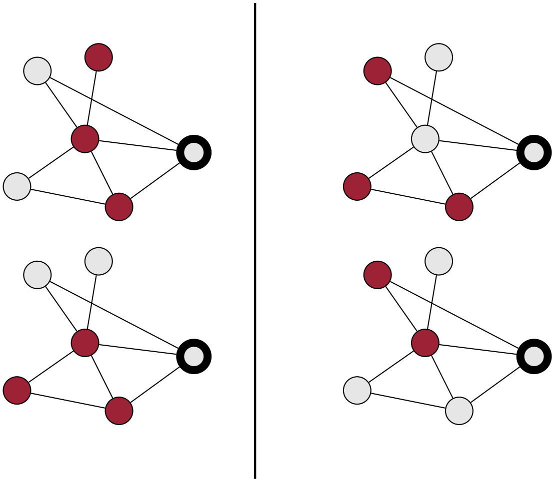
  where $w_{ii'}$ is usually determined by the distance between their covariates.

Exposure Mapping

- trivial exposure mapping:
$$G_i = \boldsymbol{Z}_{\mathcal{N}_i}$$
- Consequences of misspecification of exposure mapping:

  Aronow & Samii (2017), Estimating Average Causal Effects Under General Interference. AOAS

Exposure Mapping



Entanglement between Treatment and Interference

Consider the assignment mechanism:

$$P(\boldsymbol{Z}, \boldsymbol{G} \mid \boldsymbol{X}, \mathbb{Y}, \mathfrak{G})$$

- $\boldsymbol{Z} = (Z_1, \ldots, Z_N)$ is the treatment assignment vector.
- $\boldsymbol{G} = (G_1, \ldots, G_N)$ is the interference exposure vector.
- $\boldsymbol{X} = (X_1, \ldots, X_N)$ is the covariate vector.
- $\mathbb{Y} = \{Y_i(z, g), i = 1, \ldots, N : z \in \{0, 1\}, g \in \mathcal{G}_i\}$ is all potential outcomes.
- $\mathfrak{G}$ is the interference graph.

Entanglement between Treatment and Interference

Consider the assignment mechanism:

$$P(\boldsymbol{Z}, \boldsymbol{G} \mid \boldsymbol{X}, \mathbb{Y}, \mathfrak{G})$$

- $\boldsymbol{Z} = (Z_1, \ldots, Z_N)$ is the treatment assignment vector.
- $\boldsymbol{G} = (G_1, \ldots, G_N)$ is the interference exposure vector.
- $\boldsymbol{X} = (X_1, \ldots, X_N)$ is the covariate vector.
- $\mathbb{Y} = \{Y_i(z, g), i = 1, \ldots, N : z \in \{0, 1\}, g \in \mathcal{G}_i\}$ is all potential outcomes.
- $\mathfrak{G}$ is the interference graph.

The biggest problem in the inference framework is that $\boldsymbol{G}$ is a deterministic function of $\boldsymbol{Z}$ given all the conditions.

Unconfoundedness Condition

The unconfoundedness condition now becomes:

$$Z_i, G_i \perp\!\!\!\perp Y_i(z, g) \mid X_i$$

Unconfoundedness Condition

The unconfoundedness condition now becomes:

$$Z_i, G_i \perp\!\!\!\perp Y_i(z, g) \mid X_i$$

- Any randomization on the treatment that is independent of X_i satisfies the unconfoundedness condition.

Causal Effect

- Unit-level direct treatment effect:

$$\tau_i^{(d)}(g) = Y_i(1, g) - Y_i(0, g)$$

- Unit-level indirect/spillover treatment effect:

$$\tau_i^{(i)}(g, g'; z) = Y_i(z, g) - Y_i(z, g')$$

- Unit-level total treatment effect: (often the most interesting one)

$$\tau_i^{(t)} = Y_i(1, \overline{g}) - Y_i(0, \underline{g})$$

  where

$$\overline{g} = g_i(\mathbf{1}), \quad \underline{g} = g_i(\mathbf{0}).$$

- The population average treatment effects are defined as the average of the unit-level treatment effects.

Average Dose Response Function

We define the following populational average potential outcomes:

$$\mu(z, g) = E[Y_i(z, g) \mid i \in V_g], \quad \forall z \in \{0, 1\}, g \in \mathcal{G},$$

where $V_g = \{i : g \in \mathcal{G}_i\}$ is the set of units with possible exposure g and $\mathcal{G} = \bigcup_{i=1}^{N} \mathcal{G}_i$ is the set of all possible exposures.

It is also called the **average dose response function** (ADRF).

Average Dose Response Function

We define the following populational average potential outcomes:

$$\mu(z, g) = E[Y_i(z, g) \mid i \in V_g], \quad \forall z \in \{0, 1\}, g \in \mathcal{G},$$

where $V_g = \{i : g \in \mathcal{G}_i\}$ is the set of units with possible exposure g and $\mathcal{G} = \bigcup_{i=1}^{N} \mathcal{G}_i$ is the set of all possible exposures.

It is also called the **average dose response function** (ADRF).

The population average treatment effects are defined as the contrast of ADRFs.

Observational Study

Now consider the observational study problem.

- Observed, fixed interference graph $\mathfrak{G}$.
- Observed confounders $\boldsymbol{X}$.
- Observed treatment assignment $\boldsymbol{Z}$.
- Observed interference exposure $\boldsymbol{G}$ — usually computed from $\boldsymbol{Z}$ and $\mathfrak{G}$.

Observational Study

Now consider the observational study problem.

- Observed, fixed interference graph $\mathfrak{G}$.
- Observed confounders $\boldsymbol{X}$.
- Observed treatment assignment $\boldsymbol{Z}$.
- Observed interference exposure $\boldsymbol{G}$ — usually computed from $\boldsymbol{Z}$ and $\mathfrak{G}$.

Goal: estimate the casual effects, as well as $\mu(z, g)$.

Propensity Score

The **joint propensity score** of (z, g) for unit i is

$$\psi(z; g; x) = P(Z_i = z, G_i = g \mid X_i = x)$$

Assumptions used here:

- Unconfoundedness.
- The probability depends on its own covariates X_i only.

Propensity Score

The **joint propensity score** of (z, g) for unit i is

$$\psi(z; g; x) = P(Z_i = z, G_i = g \mid X_i = x)$$

Assumptions used here:

- Unconfoundedness.
- The probability depends on its own covariates X_i only.

The joint propensity score works as a balancing score:

$$Y_i(z, g) \perp\!\!\!\perp Z_i, G_i \mid \psi(z; g; X_i) \quad \forall z, g$$

Propensity Score

The propensity score can be expanded as

$$\psi(z; g; x) = P(G_i = g \mid Z_i = z, X_i^z = x^z)P(Z_i = z \mid X_i^g = x^g)$$

where X_i^g and X_i^z are the covariates that are used to predict Z_i and G_i, respectively.

Propensity Score

The propensity score can be expanded as

$$\psi(z; g; x) = P(G_i = g \mid Z_i = z, X_i^z = x^z)P(Z_i = z \mid X_i^g = x^g)$$

where X_i^g and X_i^z are the covariates that are used to predict Z_i and G_i, respectively.

- The first term is the **neighborhood propensity score**:

$$P(G_i = g \mid Z_i = z, X_i^z = x^z) = \lambda(g; z; x^g)$$

- The second term is the **individual propensity score**:

$$P(Z_i = z \mid X_i^g = x^g) = \phi(z; x^z)$$

Propensity Score

The propensity score can be expanded as

$$\psi(z; g; x) = P(G_i = g \mid Z_i = z, X_i^z = x^z)P(Z_i = z \mid X_i^g = x^g)$$

where X_i^g and X_i^z are the covariates that are used to predict Z_i and G_i, respectively.

- The first term is the **neighborhood propensity score**:

$$P(G_i = g \mid Z_i = z, X_i^z = x^z) = \lambda(g; z; x^g)$$

- The second term is the **individual propensity score**:

$$P(Z_i = z \mid X_i^g = x^g) = \phi(z; x^z)$$

They jointly satisfy the unconfoundedness condition:

$$Z_i, G_i \perp\!\!\!\perp Y_i(z, g) \mid \lambda(g; z; X_i^g), \phi(z; X_i^z) \quad \forall z, g$$

Stratification Strategy

1. Stratification on the individual propensity score $\phi(z; X_i^z)$:
   1.1 Fit $\phi(1; X_i^z)$ using logistic regression.
   1.2 Divide the units into J strata, $B_1, \ldots, B_J$, based on the estimated propensity score.
   1.3 We approximately have $Z_i \perp\!\!\!\perp X_i^z \mid i \in B_j$ for any j.

Stratification Strategy

1. Stratification on the individual propensity score $\phi(z; X_i^z)$:
   1.1 Fit $\phi(1; X_i^z)$ using logistic regression.
   1.2 Divide the units into J strata, $B_1, \ldots, B_J$, based on the estimated propensity score.
   1.3 We approximately have $Z_i \perp\!\!\!\perp X_i^z \mid i \in B_j$ for any j.
2. Estimation within each stratum:
   2.1 Fit $\lambda(g; z; X_i^g)$ using logistic regression.
   2.2 Fit a parametric model $Y_i(z, g) \sim Z_i + G_i + \hat{\lambda}_i$.
   2.3 For the pair (z, g), for each eligible unit, make a prediction of $Y_i(z, g)$ using the fitted model.
   2.4 The estimator is

$$\hat{\mu}_j(z, g) = |B_j^g|^{-1} \sum_{i \in B_j^g} \hat{Y}_i(z, g)$$

Stratification Strategy

1. Stratification on the individual propensity score $\phi(z; X_i^z)$:
   1.1 Fit $\phi(1; X_i^z)$ using logistic regression.
   1.2 Divide the units into J strata, $B_1, \ldots, B_J$, based on the estimated propensity score.
   1.3 We approximately have $Z_i \perp\!\!\!\perp X_i^z \mid i \in B_j$ for any j.
2. Estimation within each stratum:
   2.1 Fit $\lambda(g; z; X_i^g)$ using logistic regression.
   2.2 Fit a parametric model $Y_i(z, g) \sim Z_i + G_i + \hat{\lambda}_i$.
   2.3 For the pair (z, g), for each eligible unit, make a prediction of $Y_i(z, g)$ using the fitted model.
   2.4 The estimator is

$$\hat{\mu}_j(z, g) = |B_j^g|^{-1} \sum_{i \in B_j^g} \hat{Y}_i(z, g)$$

3. The final estimator is

$$\hat{\mu}(z, g) = \sum_{j=1}^{J} \hat{\mu}_j(z, g) \pi_j^g$$

where $\pi_j^g = |B_j^g| / |B_j|$ is the proportion of units in stratum j with exposure g.

Stratification Strategy

More details in
Forastiere, Airoldi, & Mealli (2020). Identication and estimation of treatment and interference effects in observational studies on networks. JASA.

Stratification Strategy

More details in
Forastiere, Airoldi, & Mealli (2020). Identication and estimation of treatment and interference effects in observational studies on networks. JASA.

Why parametric model in the second step: for example,

- A network of $N = 1000$ units.
- Each unit has $|\mathcal{N}_i| = 4$ neighbors.
- $N_t = 500$ units are randomly assigned to treatment.
- We want to estimate $\mu(1, 2)$ for exposure mapping of number of treated neighbors.

Stratification Strategy

More details in
Forastiere, Airoldi, & Mealli (2020). Identication and estimation of treatment and interference effects in observational studies on networks. JASA.

Why parametric model in the second step: for example,

- A network of $N = 1000$ units.
- Each unit has $|\mathcal{N}_i| = 4$ neighbors.
- $N_t = 500$ units are randomly assigned to treatment.
- We want to estimate $\mu(1, 2)$ for exposure mapping of number of treated neighbors.
- The expected number of units that are exposed to $(z, g) = (1, 2)$ is

$$\approx 1000 \times \frac{1}{2} \times \binom{4}{2} \times \frac{1}{2^4} \approx 188 \ll 1000.$$

Stratification Strategy

More details in
Forastiere, Airoldi, & Mealli (2020). Identication and estimation of treatment and interference effects in observational studies on networks. JASA.

Why parametric model in the second step: for example,

- A network of $N = 1000$ units.
- Each unit has $|\mathcal{N}_i| = 4$ neighbors.
- $N_t = 500$ units are randomly assigned to treatment.
- We want to estimate $\mu(1, 2)$ for exposure mapping of number of treated neighbors.
- The expected number of units that are exposed to $(z, g) = (1, 2)$ is

$$\approx 1000 \times \frac{1}{2} \times \binom{4}{2} \times \frac{1}{2^4} \approx 188 \ll 1000.$$

- Use parametric model to "impute" the potential outcomes for other units.

Empirical Matching

Now we consider a model-free approach.
Suppose we want to estimate the following effect:

$$\hat{\mu}(z, g) - \hat{\mu}(z', g')$$

Empirical Matching

Now we consider a model-free approach.
Suppose we want to estimate the following effect:

$$\hat{\mu}(z, g) - \hat{\mu}(z', g')$$

- A natural way is to find two subsets of the samples:

$$\mathcal{A} = \{i : Z_i = z, G_i = g\}, \quad \mathcal{B} = \{i : Z_i = z', G_i = g'\}.$$

- The average treatment effect can be estimated based on these two sets of units using, e.g., matching methods.

Empirical Matching

Now we consider a model-free approach.
Suppose we want to estimate the following effect:

$$\hat{\mu}(z, g) - \hat{\mu}(z', g')$$

- A natural way is to find two subsets of the samples:

$$\mathcal{A} = \{i : Z_i = z, G_i = g\}, \quad \mathcal{B} = \{i : Z_i = z', G_i = g'\}.$$

- The average treatment effect can be estimated based on these two sets of units using, e.g., matching methods.
- Drawback 1: Small sample size.
- Drawback 2: Correlation.

Experimental Design

Due to the complicated nature of the interference problems, experimental design approaches are often more prefered over the observational study approaches.

Experimental Design

Due to the complicated nature of the interference problems, experimental design approaches are often more prefered over the observational study approaches.

For the further demonstrations, we consider the following exposure mapping:

$$g_i(\boldsymbol{Z}_{\mathcal{N}_i}) = |\mathcal{N}_i|^{-1} \sum_{i' \in \mathcal{N}_i} Z_{i'}$$

The proportion of treated neighbors.

Experimental Design

Due to the complicated nature of the interference problems, experimental design approaches are often more prefered over the observational study approaches.

For the further demonstrations, we consider the following exposure mapping:

$$g_i(\boldsymbol{Z}_{\mathcal{N}_i}) = |\mathcal{N}_i|^{-1} \sum_{i' \in \mathcal{N}_i} Z_{i'}$$

The proportion of treated neighbors.

The total treatment effect becomes

$$\tau^{(t)} = \mu(1, 1) - \mu(0, 0).$$

Ego-Centric Design

An **ego-cluster** consists of a unit (ego center) and all its neighbors (alters).

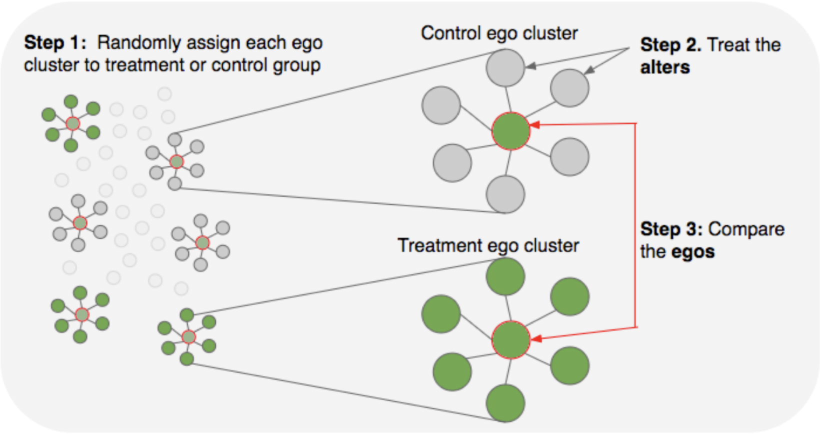
Ego-Centric Design

An **ego-cluster** consists of a unit (ego center) and all its neighbors (alters).

In an **ego-centric design**.

- Find maximal disjoint ego-clusters $C_1, \ldots, C_K$ from the interference graph.
- Randomly assign half of the ego-clusters to treatment and half to control.
- The total treatment effect is estimated by the difference-in-means estimator on the ego-center's outcomes.

Ego-Centric Design



Saint-Jacques, Varshney, Simpson, & Xu (2019). Using Ego-Clusters to Measure Network Effects at LinkedIn.

Ego-Centric Design

Advntages:

- Easy control of the interference.
- Similar procedure as the RCT.

Ego-Centric Design

Advntages:

- Easy control of the interference.
- Similar procedure as the RCT.

Disadvantages:

- Limited sample size.
- Require sparse interference graph.

Independent-Set Design

An **independent set** is a subset of the vertices in a graph such that no two vertices in the subset are adjacent.

Independent-Set Design

An **independent set** is a subset of the vertices in a graph such that no two vertices in the subset are adjacent.

In an **independent-set design**,

- Find a maximal independent set $\mathcal{I}$ from the interference graph. Call the rest of the vertices the auxiliary set $\mathcal{A}$.

Independent-Set Design

An **independent set** is a subset of the vertices in a graph such that no two vertices in the subset are adjacent.

In an **independent-set design**,

- Find a maximal independent set $\mathcal{I}$ from the interference graph. Call the rest of the vertices the auxiliary set $\mathcal{A}$.
- We focus on the units in $\mathcal{I}$. Their interference exposure is $\boldsymbol{G}_{\mathcal{I}} = \boldsymbol{\Gamma} \boldsymbol{Z}_A$, where $\boldsymbol{\Gamma}$ is the (normalized) adjacency matrix between $\mathcal{I}$ and $\mathcal{A}$.
- $\boldsymbol{Z}_A$ is chosen to maximize the variance of the interference exposure:

$$\boldsymbol{Z}_A^T \boldsymbol{\Gamma}^T \left[\boldsymbol{I} - n_I^{-1} \mathbf{1} \mathbf{1}^T \right] \boldsymbol{\Gamma} \boldsymbol{Z}_A .$$

Independent-Set Design

An **independent set** is a subset of the vertices in a graph such that no two vertices in the subset are adjacent.

In an **independent-set design**,

- Find a maximal independent set $\mathcal{I}$ from the interference graph. Call the rest of the vertices the auxiliary set $\mathcal{A}$.
- We focus on the units in $\mathcal{I}$. Their interference exposure is $\boldsymbol{G}_{\mathcal{I}} = \boldsymbol{\Gamma} \boldsymbol{Z}_A$, where $\boldsymbol{\Gamma}$ is the (normalized) adjacency matrix between $\mathcal{I}$ and $\mathcal{A}$.
- $\boldsymbol{Z}_A$ is chosen to maximize the variance of the interference exposure:

$$\boldsymbol{Z}_A^T \boldsymbol{\Gamma}^T \left[\boldsymbol{I} - n_I^{-1} \mathbf{1}\mathbf{1}^T \right] \boldsymbol{\Gamma} \boldsymbol{Z}_A.$$

- Units in $\mathcal{I}$ are assigned according to the interference exposure:

$$Z_i = \begin{cases} 1 & \text{if } G_i \geq 0.5 \\ 0 & \text{otherwise} \end{cases}$$

Independent-Set Design

An **independent set** is a subset of the vertices in a graph such that no two vertices in the subset are adjacent.

In an **independent-set design**,

- Find a maximal independent set $\mathcal{I}$ from the interference graph. Call the rest of the vertices the auxiliary set $\mathcal{A}$.
- We focus on the units in $\mathcal{I}$. Their interference exposure is $\boldsymbol{G}_{\mathcal{I}} = \boldsymbol{\Gamma} \boldsymbol{Z}_A$, where $\boldsymbol{\Gamma}$ is the (normalized) adjacency matrix between $\mathcal{I}$ and $\mathcal{A}$.
- $\boldsymbol{Z}_A$ is chosen to maximize the variance of the interference exposure:

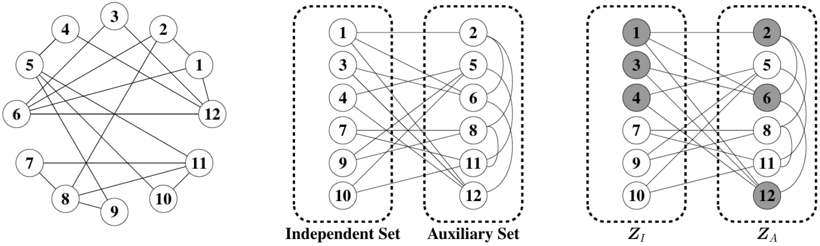
$$\boldsymbol{Z}_A^T \boldsymbol{\Gamma}^T \left[\boldsymbol{I} - n_I^{-1} \mathbf{1} \mathbf{1}^T \right] \boldsymbol{\Gamma} \boldsymbol{Z}_A.$$

- Units in $\mathcal{I}$ are assigned according to the interference exposure:

$$Z_i = \begin{cases} 1 & \text{if } G_i \geq 0.5 \\ 0 & \text{otherwise} \end{cases}$$

- The total treatment effect is estimated by the difference-in-means estimator on the units in $\mathcal{I}$.

Independent-Set Design



Cai, Zhang, & Airoldi (2025). Independent-Set Design of Experiments for Estimating Treatment and Spillover Effects under Network Interference. ICLR.

Independent-Set Design

Advantages:

- Independent control on the treatment and interference.
- Large independent set size (compared to ego-centeric design) with high probability.

Independent-Set Design

Advantages:

- Independent control on the treatment and interference.
- Large independent set size (compared to ego-centeric design) with high probability.

Disadvantages:

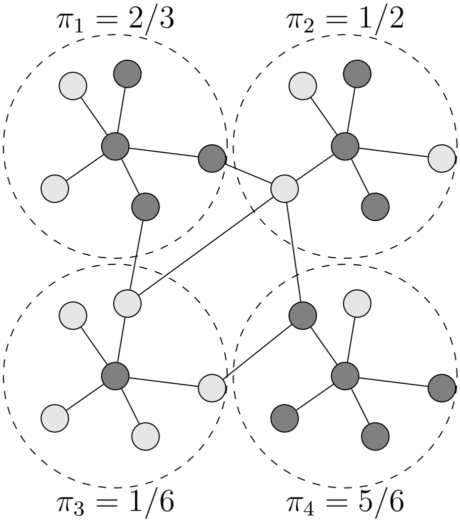
- Computation of the maximal independent set is NP-hard.
- Could have bias.
- Require sparse interference graph.

Randomized Saturation Design

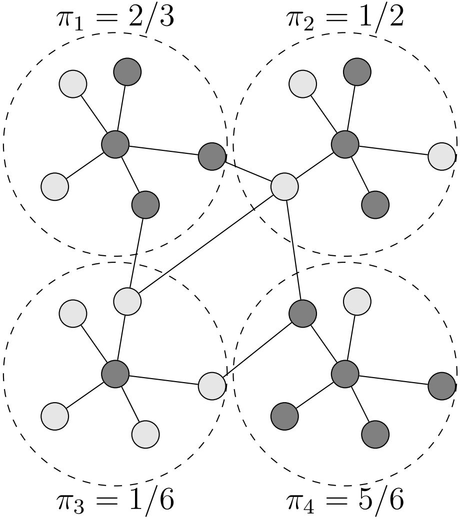
In a **randomized saturation design**,

- Units are divided into K clusters.
- A set of proportions $\boldsymbol{\pi} = \{\pi_1, \ldots, \pi_K\}$ is randomly assigned to the clusters. W.L.O.G., assume cluster k is assigned π_k.
- For the N_k units in cluster k, $\pi_k N_k$ units are randomly assigned to treatment and the rest are assigned to control.
- The causal effects are estimated by the difference-in-means.

Randomized Saturation Design



Randomized Saturation Design



Population: A collection of J clusters of units.

Two-step Randomization:

1. Randomly generate a *proportion vector* $\boldsymbol{\pi} = [\pi_1, \ldots, \pi_J]$ from $\mathbf{\Pi}$.
2. Randomly assign $n_j = \lfloor \pi_j N_j \rfloor$ units in cluster j to treatment.

Example: A realization of treatment assignment generated by a randomized saturation design where the realized proportion vector is $\boldsymbol{\pi} = \left[\frac{2}{3}, \frac{1}{2}, \frac{1}{6}, \frac{5}{6}\right]$.

●: treated units ○: control units

Randomized Saturation Design

Advantages:

- Easy to implement.
- Rough control of the interference.

Randomized Saturation Design

Advantages:

- Easy to implement.
- Rough control of the interference.

Disadvantages:

- Bias from inter-cluster interference.
- Require partial interference assumption (in oppose to our local interference assumption).

$$G_i = \text{proportion of treated unit in its cluster}$$

Randomized Saturation Design

- Proposed by Hudgens and Halloran (2008). Toward causal inference with interference. JASA.
- Theoretical properties: Jiang, Imai & Malani (2022). Statistical inference and power analysis for direct and spillover effects in two-stage randomized experiments. Biometrics.

Randomized Saturation Design

- Proposed by Hudgens and Halloran (2008). Toward causal inference with interference. JASA.
- Theoretical properties: Jiang, Imai & Malani (2022). Statistical inference and power analysis for direct and spillover effects in two-stage randomized experiments. Biometrics.
- Common clustering strategies:
  - Community detection algorithm. (tons of reference here)
  - Randomly assign units to clusters. Ugandar & Yin (2020). Randomized Graph Cluster Randomization.
  - Sample disjoint clusters from the population.
- Critization on poor clustering structures:
  Cai, Pouget-Abadie, & Airoldi (2022). Optimizing Randomized and Deterministic Saturation Designs under Interferenc.

Randomized Saturation Design

- Proposed by Hudgens and Halloran (2008). Toward causal inference with interference. JASA.
- Theoretical properties: Jiang, Imai & Malani (2022). Statistical inference and power analysis for direct and spillover effects in two-stage randomized experiments. Biometrics.
- Common clustering strategies:
  - Community detection algorithm. (tons of reference here)
  - Randomly assign units to clusters. Ugandar & Yin (2020). Randomized Graph Cluster Randomization.
  - Sample disjoint clusters from the population.
- Critization on poor clustering structures:
  Cai, Pouget-Abadie, & Airoldi (2022). Optimizing Randomized and Deterministic Saturation Designs under Interferenc.
- In practice, to estimate the total treatment effect, the proportion vector has $K/2$ 1's and $K/2$ 0's.

Thank you for joining the workshop!

Contact:
Me chencheng.cai@wsu.edu
CISER ciser.info@wsu.edu