Stanford University

ICME

Synthetic Principle Component Design

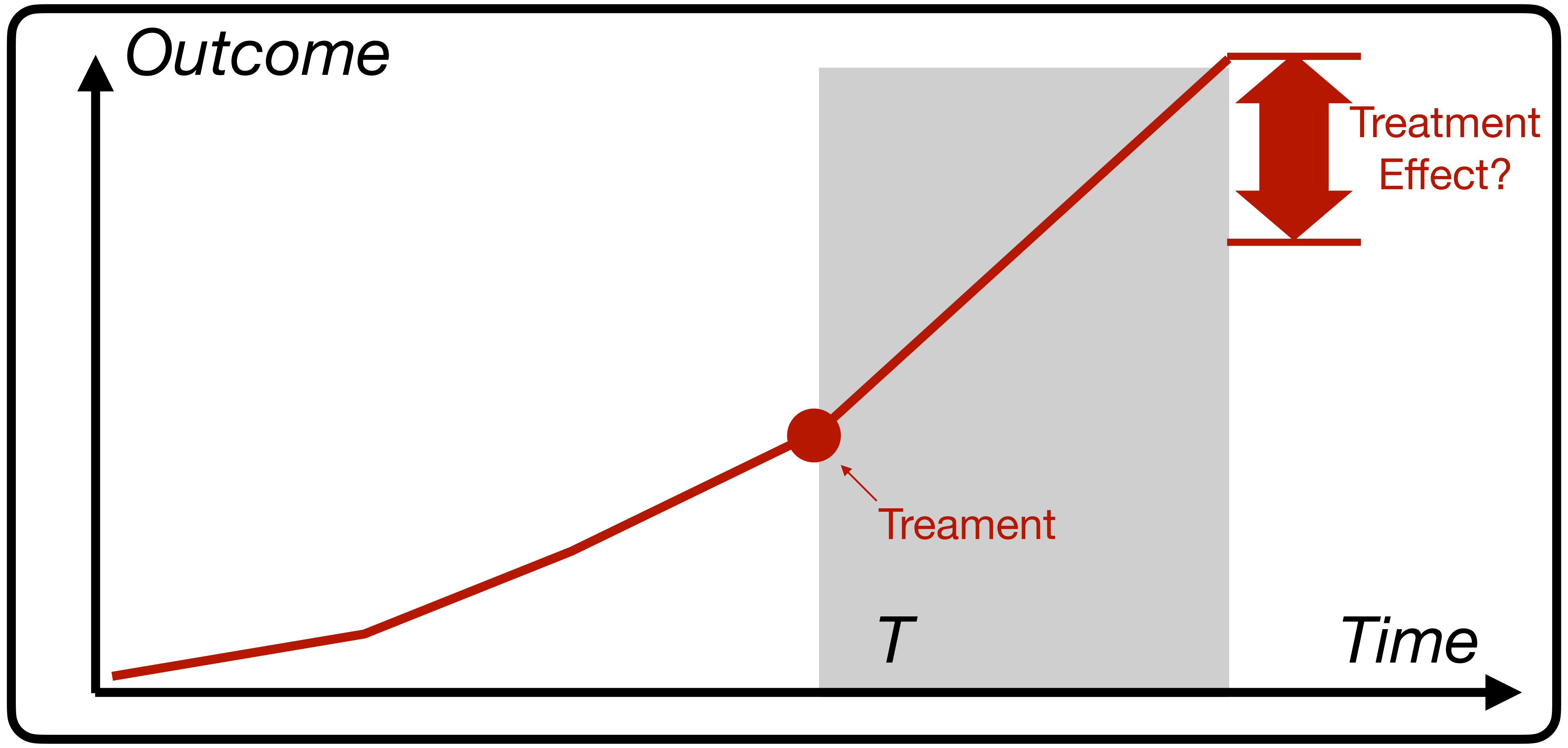
Phase synchronization and Experiment Design

Yiping Lu. Stanford University

Joint work with Jiajin Li, Lexing Ying, Jose Blanchet

Synthetic Control

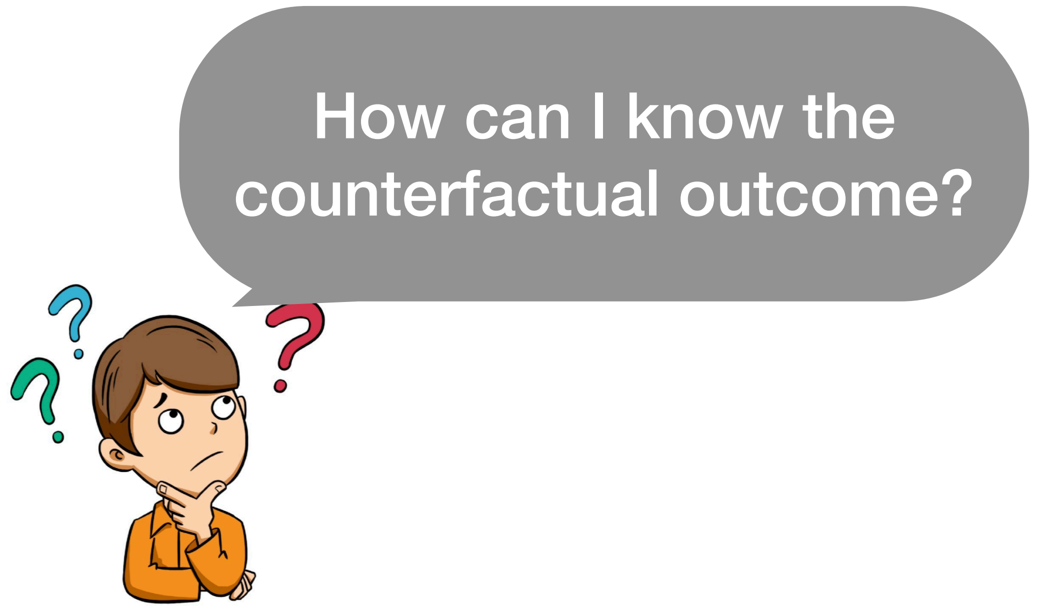
Causal Inference for Panel Data



Synthetic Control

Aim Estimate the effect of an applied policy

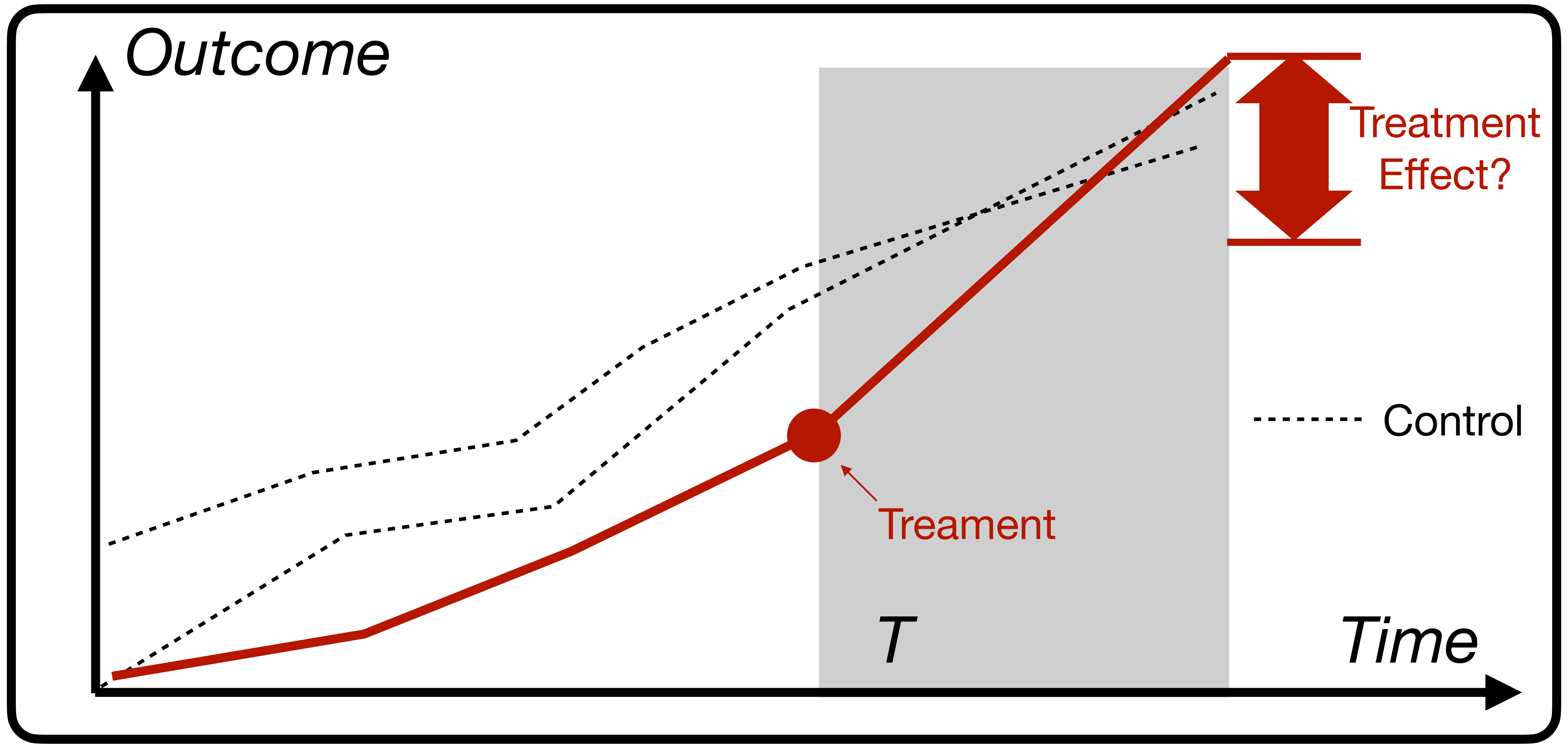
We need to know the **counterfactual** outcome!



Abadie A, Diamond A, Hainmueller J. Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program[J]. Journal of the American statistical Association, 2010, 105(490): 493-505.

Synthetic Control

Causal Inference for Panel Data



Synthetic Control

Aim Estimate the effect of an applied policy

We need to know the **counterfactual** outcome!

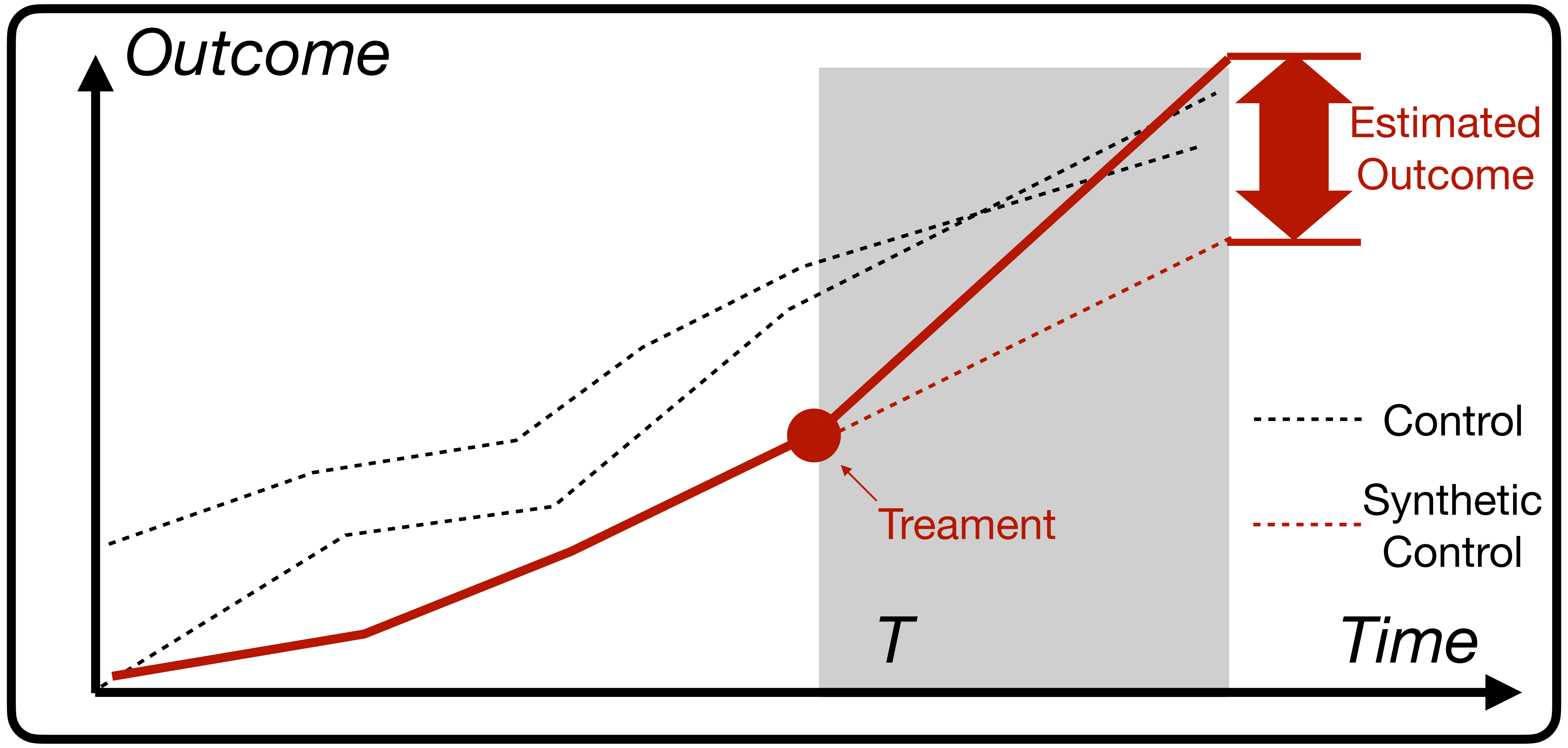
Synthetic Control

Step 1. Find out some control group.

Abadie A, Diamond A, Hainmueller J. Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program[J]. Journal of the American statistical Association, 2010, 105(490): 493-505.

Synthetic Control

Causal Inference for Panel Data



Synthetic Control

Aim Estimate the effect of an applied policy

We need to know the **counterfactual** outcome!

Synthetic Control

Step 1. Find out some control group.

Step 2. Regression on pre-treatment data.

Step 3. Synthetic the counterfactual outcome.

`California = 0.334* Utah+0.234*Nevada+0.164*Colorado+0.069*Connecticut`

Abadie A, Diamond A, Hainmueller J. Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program[J]. Journal of the American statistical Association, 2010, 105(490): 493-505.

Experiment Design

Covariate Balancing

NP-Hard

Once I have dataset, how can I design whom to treat?

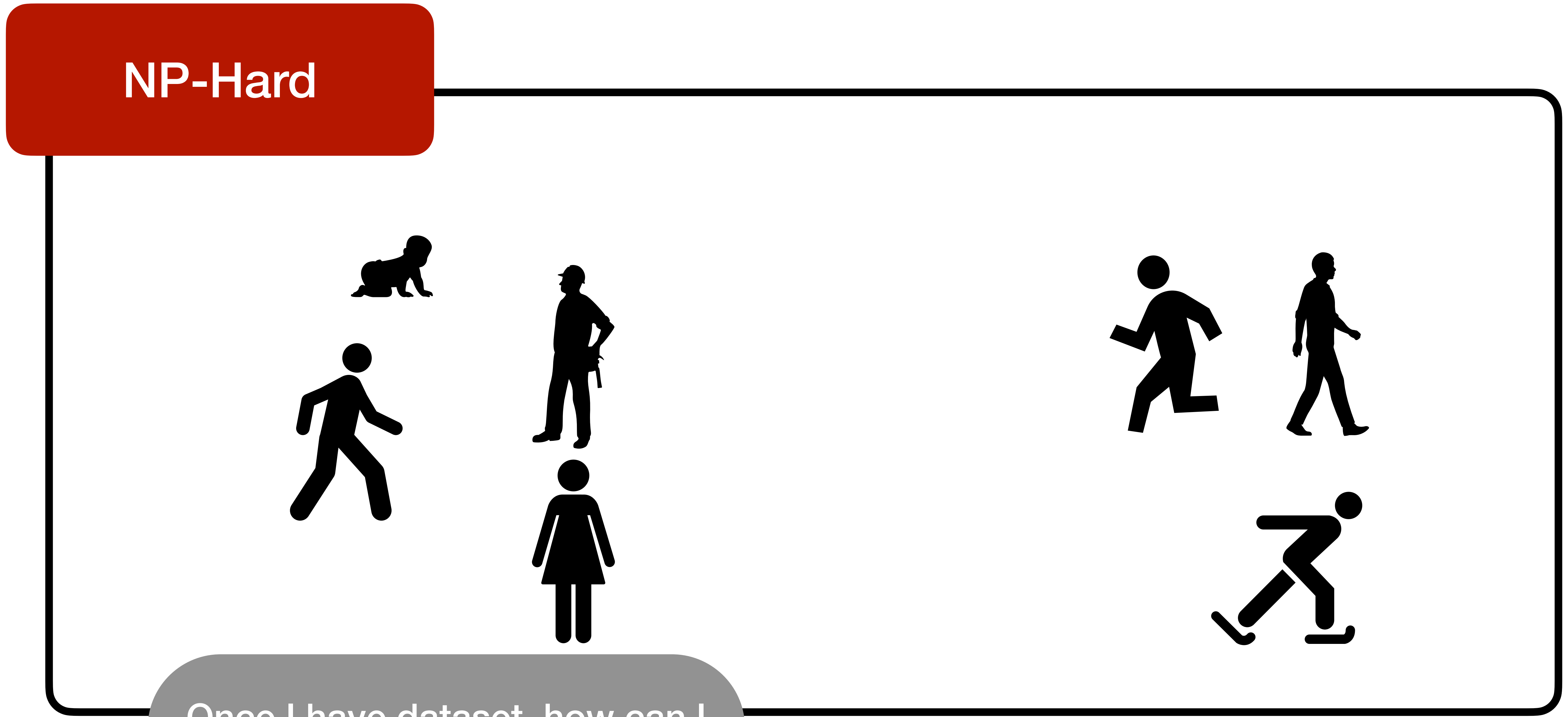
Treated data should similar to control data

Nonbipartite matching problem

divides a single group of 2n subjects into n pairs to minimize covariate differences within pairs

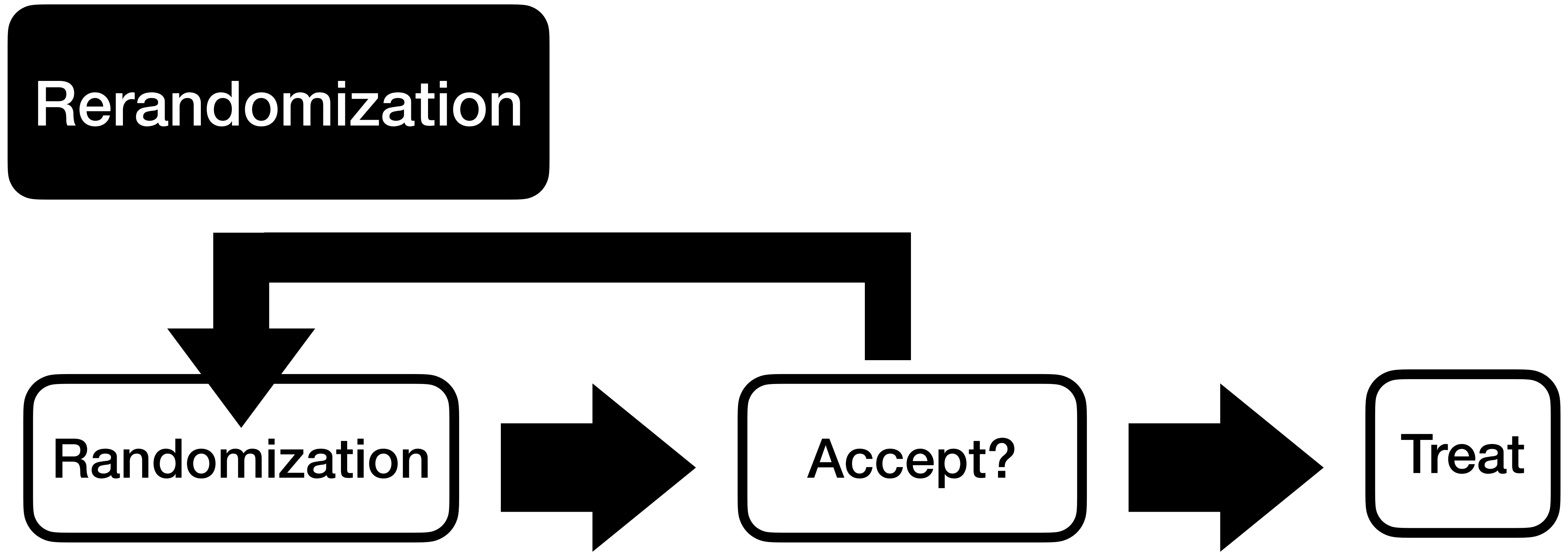
Experiment Design

Covariate Balancing



Once I have dataset, how can I design whom to treat?

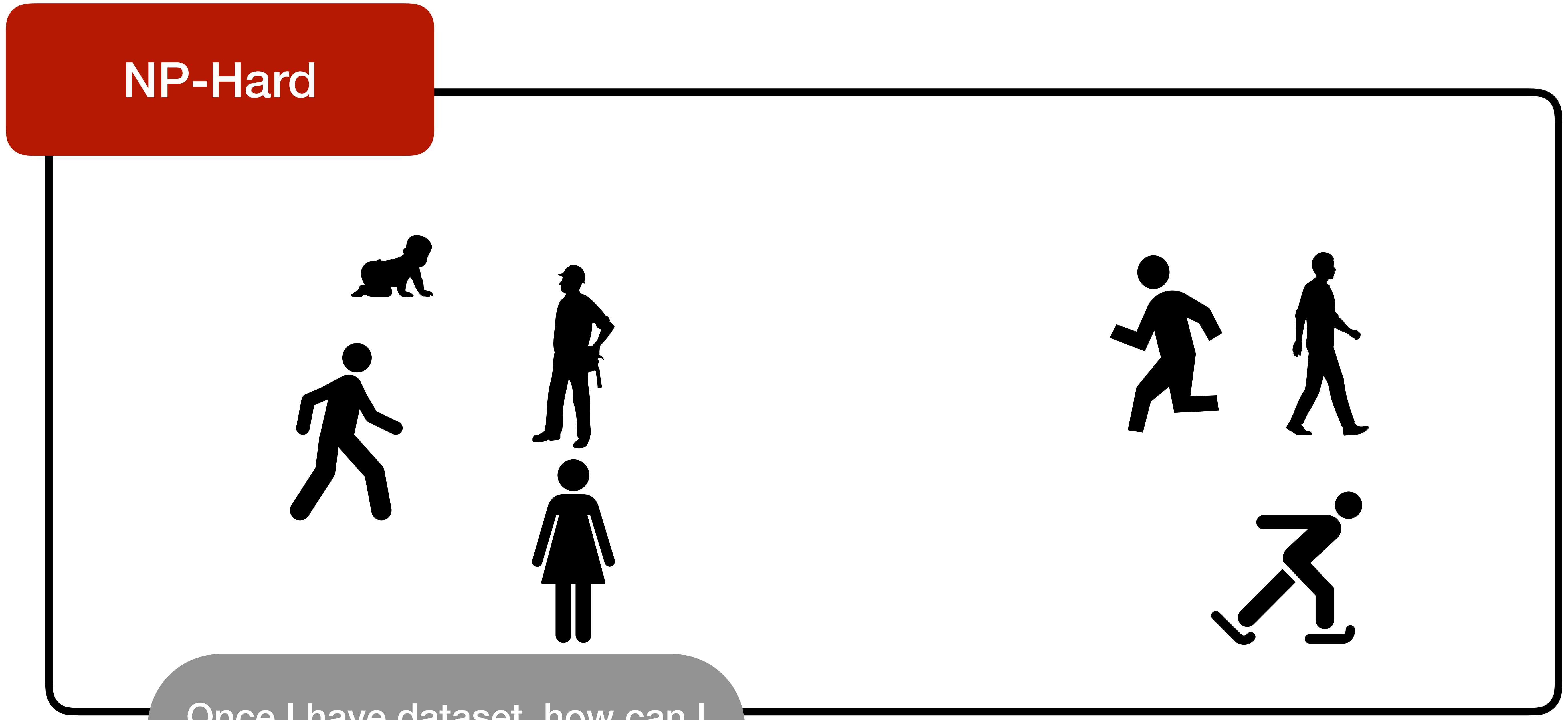
Treated data should similar to control data



Morgan K L, Rubin D B. Rerandomization to improve covariate balance in experiments. The Annals of Statistics, 2012, 40(2): 1263-1282.

Experiment Design

Covariate Balancing

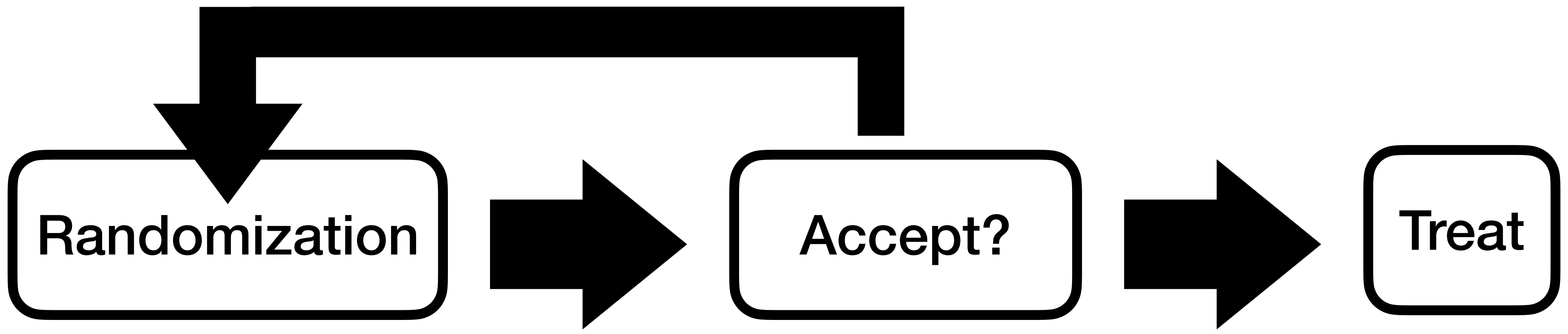


NP-Hard

Once I have dataset, how can I design whom to treat?

Treated data should similar to control data

Rerandomization



Morgan K L, Rubin D B. Rerandomization to improve covariate balance in experiments. The Annals of Statistics, 2012, 40(2): 1263-1282.

Propensity Score

$$\boxed{E_X} p(x)X - (1 - p(x))(X)$$

In expectation Balance

Imai K, Ratkovic M. Covariate balancing propensity score. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2014

Experiment Design

Covariate Balancing

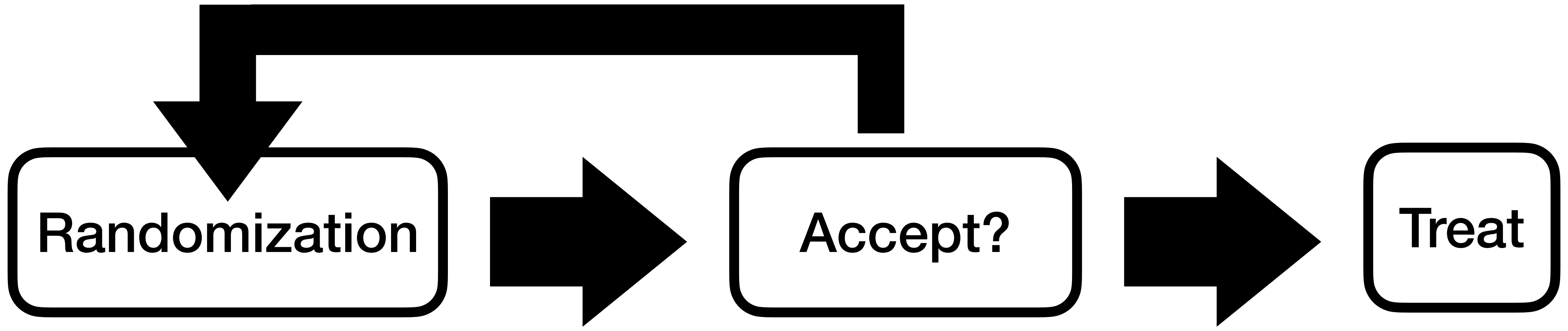
NP-Hard

Once I have dataset, how can I design whom to treat?

Treated data should similar to control data

What if #agent is small?

Rerandomization



Randomization

Accept?

Treat

Morgan K L, Rubin D B. Rerandomization to improve covariate balance in experiments. The Annals of Statistics, 2012, 40(2): 1263-1282.

Propensity Score

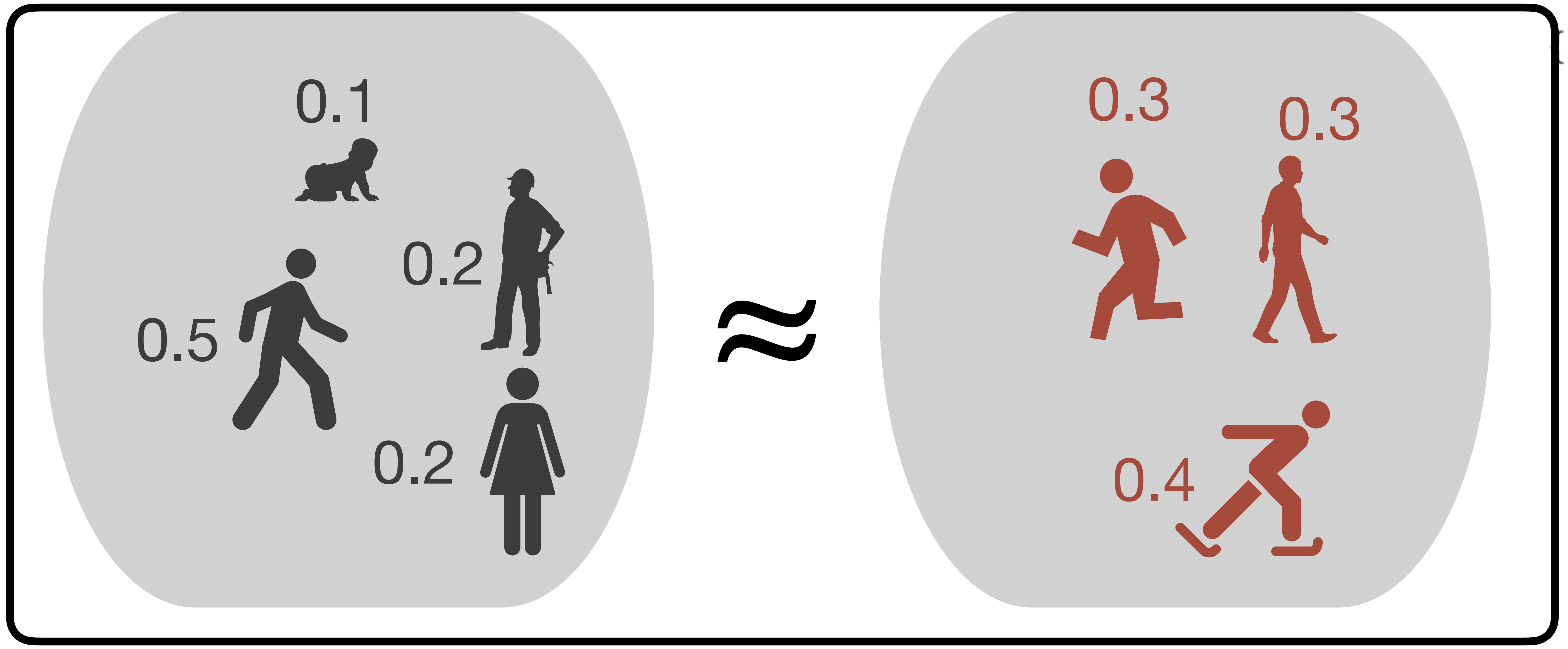
$$\boxed{E_X} p(x)X - (1 - p(x))(X)$$

In expectation Balance

Imai K, Ratkovic M. Covariate balancing propensity score. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2014

Synthetic Design

Weighted Covariate Balancing



Synthetic Design

Matching a weighted average

$$\min_{\{D_i,w_i\}_{i=1}^N} \frac{1}{T}\sum_{t=1}^{T}\left(\boxed{\sum_{i=1}^{N} w_i D_i Y_{it}} - \sum_{i=1}^{N}\boxed{w_i(1-D_i)Y_{it}}\right)^2 + \lambda\sum_{i=1}^{N} w_i^2$$

$$\text{s.t.}\quad w_i \geq 0,\quad D_i \in \{0,1\}\ \text{for}\ i = 1,\ldots,N,$$

$$\sum_{i=1}^{N} D_i = K,\quad \sum_{i=1}^{N} w_i D_i = 1,\quad \sum_{i=1}^{N} w_i(1-D_i) = 1$$

Treatment Effect =

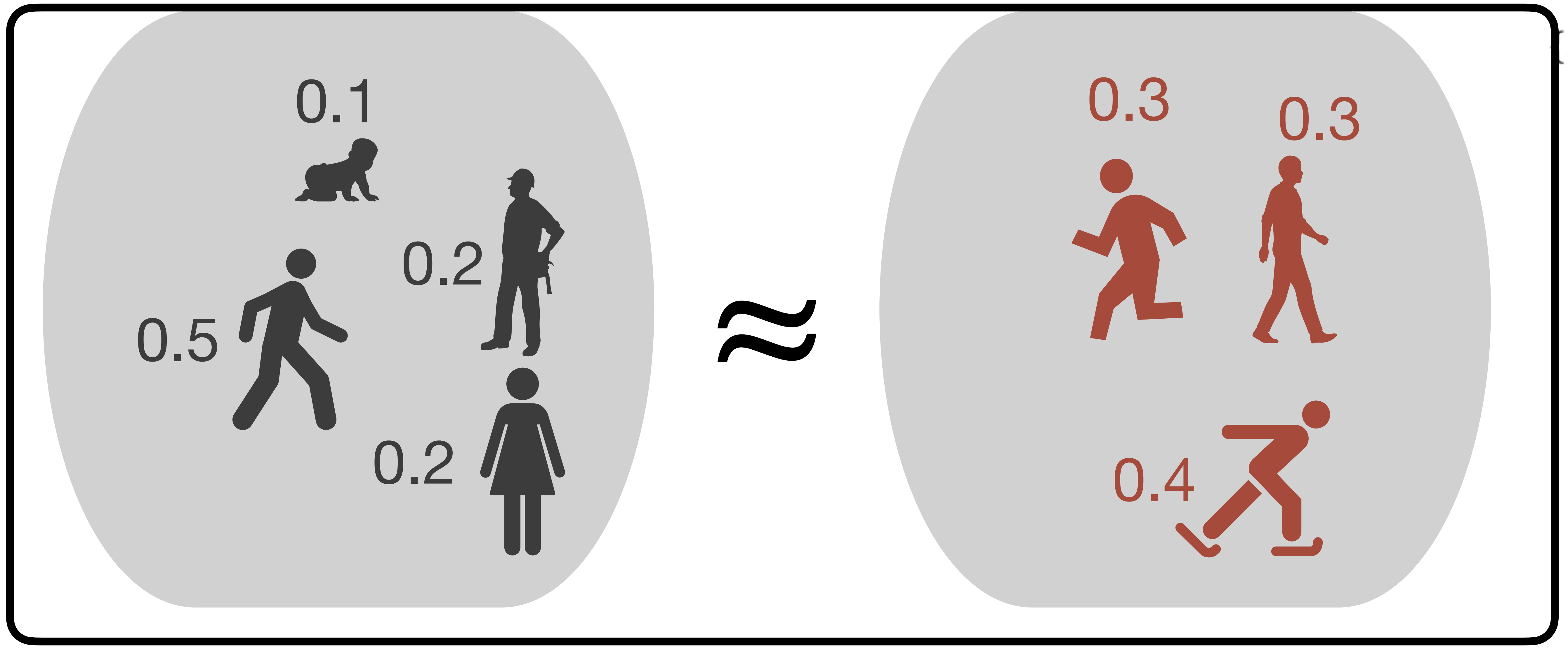
$$\boxed{\sum_{i:\ D_i=1} w_i Y_{i,T+1}} - \boxed{\sum_{i:\ D_i=0} w_i Y_{i,T+1}}.$$

Weighted mean of treatment group **Weighted mean of control group**

Doudchenko N, Khosravi K, Pouget-Abadie J, et al. Synthetic Design: An Optimization Approach to Experimental Design with Synthetic Controls. Advances in Neural Information Processing Systems, 2021, 34.

Synthetic Design

Weighted Covariate Balancing



Synthetic Design

$$\min_{\{D_i, w_i\}_{i=1}^N} \underline{\frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=1}^{N} w_i D_i Y_{it} - \sum_{i=1}^{N} w_i (1-D_i) Y_{it}\right)^2} + \lambda \sum_{i=1}^{N} w_i^2$$

$$\text{s.t.} \quad w_i \geq 0, \quad D_i \in \{0,1\} \ \text{ for } i = 1, \ldots, N,$$

$$\sum_{i=1}^{N} D_i = K, \quad \sum_{i=1}^{N} w_i D_i = 1, \quad \sum_{i=1}^{N} w_i (1-D_i) = 1$$

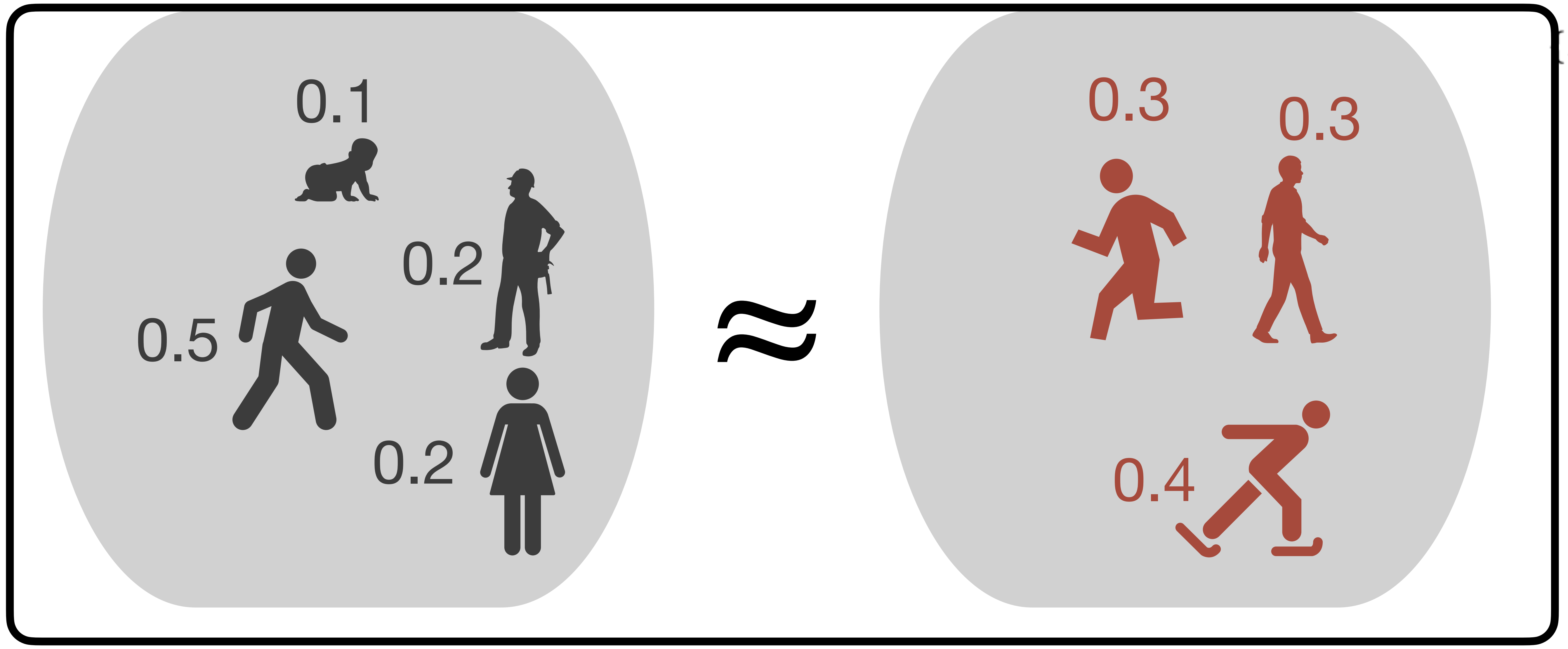
Change D_i to $\{1, -1\}$

$$\| Y(w \odot d) \|_2^1$$

Doudchenko N, Khosravi K, Pouget-Abadie J, et al. Synthetic Design: An Optimization Approach to Experimental Design with Synthetic Controls. Advances in Neural Information Processing Systems, 2021, 34.

Synthetic Design

Weighted Covariate Balancing



Synthetic Design

$$\min_{\{D_i,w_i\}_{i=1}^N} \frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=1}^{N} w_i D_i Y_{it} - \sum_{i=1}^{N} w_i(1-D_i)Y_{it}\right)^2 + \lambda\sum_{i=1}^{N} w_i^2$$

$$\text{s.t.} \quad w_i \geq 0, \quad D_i \in \{0,1\} \text{ for } i = 1,\ldots,N,$$

$$\sum_{i=1}^{N} D_i = K, \quad \boxed{\sum_{i=1}^{N} w_i D_i = 1, \quad \sum_{i=1}^{N} w_i(1-D_i) = 1}$$

Change D_i to $\{1,-1\}$

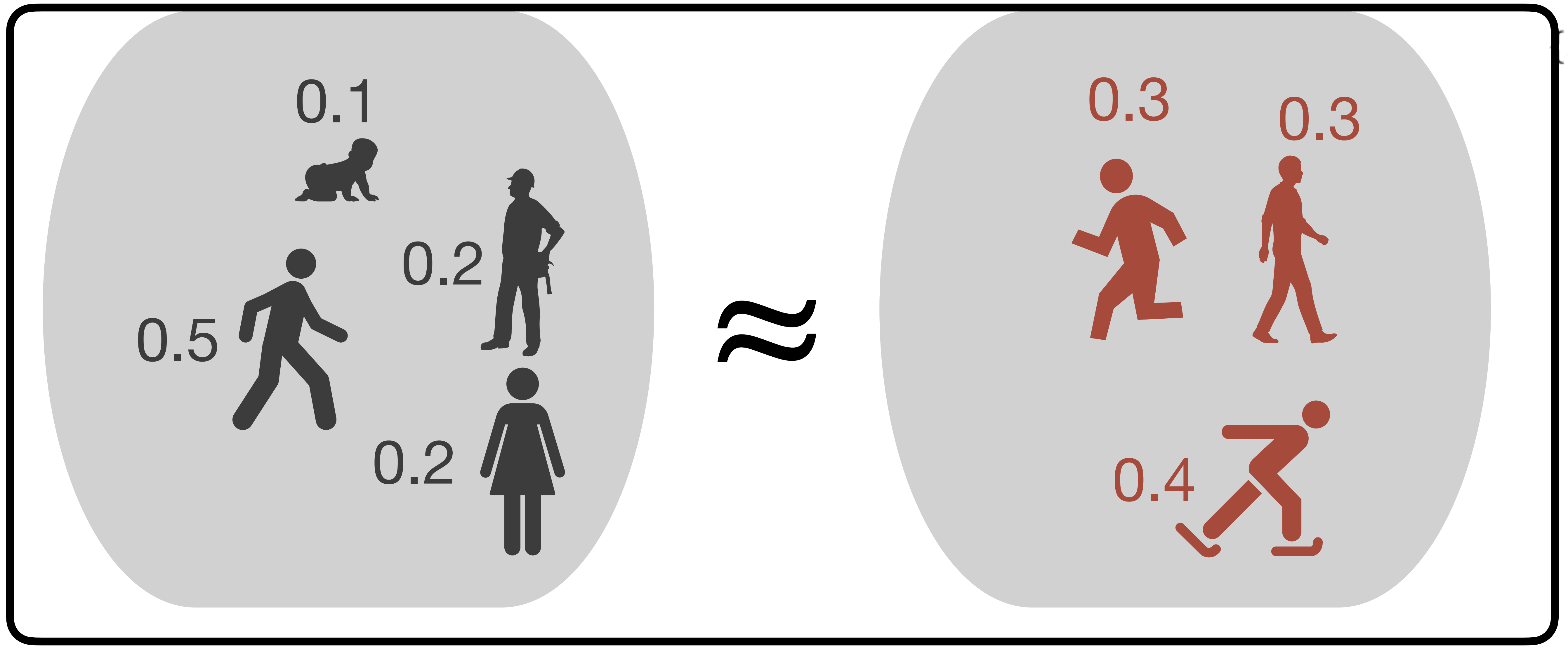
$$\|Y(w\odot d)\|_2^1$$

$$\|w\odot d\|_1 = 1, 1^\top(w\odot d) = 0$$

Doudchenko N, Khosravi K, Pouget-Abadie J, et al. Synthetic Design: An Optimization Approach to Experimental Design with Synthetic Controls. Advances in Neural Information Processing Systems, 2021, 34.

Synthetic Design

Weighted Covariate Balancing



Synthetic Design

$$\min_{\{D_i,w_i\}_{i=1}^N} \quad \frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=1}^{N} w_iD_iY_{it} - \sum_{i=1}^{N} w_i(1-D_i)Y_{it}\right)^2 + \lambda\sum_{i=1}^{N} w_i^2$$

$$\text{s.t.} \quad w_i \geq 0, \quad D_i \in \{0,1\} \text{ for } i = 1,\ldots,N,$$

$$\boxed{\sum_{i=1}^{N} D_i = K,} \quad \sum_{i=1}^{N} w_iD_i = 1, \quad \sum_{i=1}^{N} w_i(1-D_i) = 1$$

Drop

Change D_i to $\{1,-1\}$

Min $\|Y(w\odot d)\|_2^2$

Still non-convex!

s.t. $\|w\odot d\|_1 = 1, 1^{\top}(w\odot d) = 0$

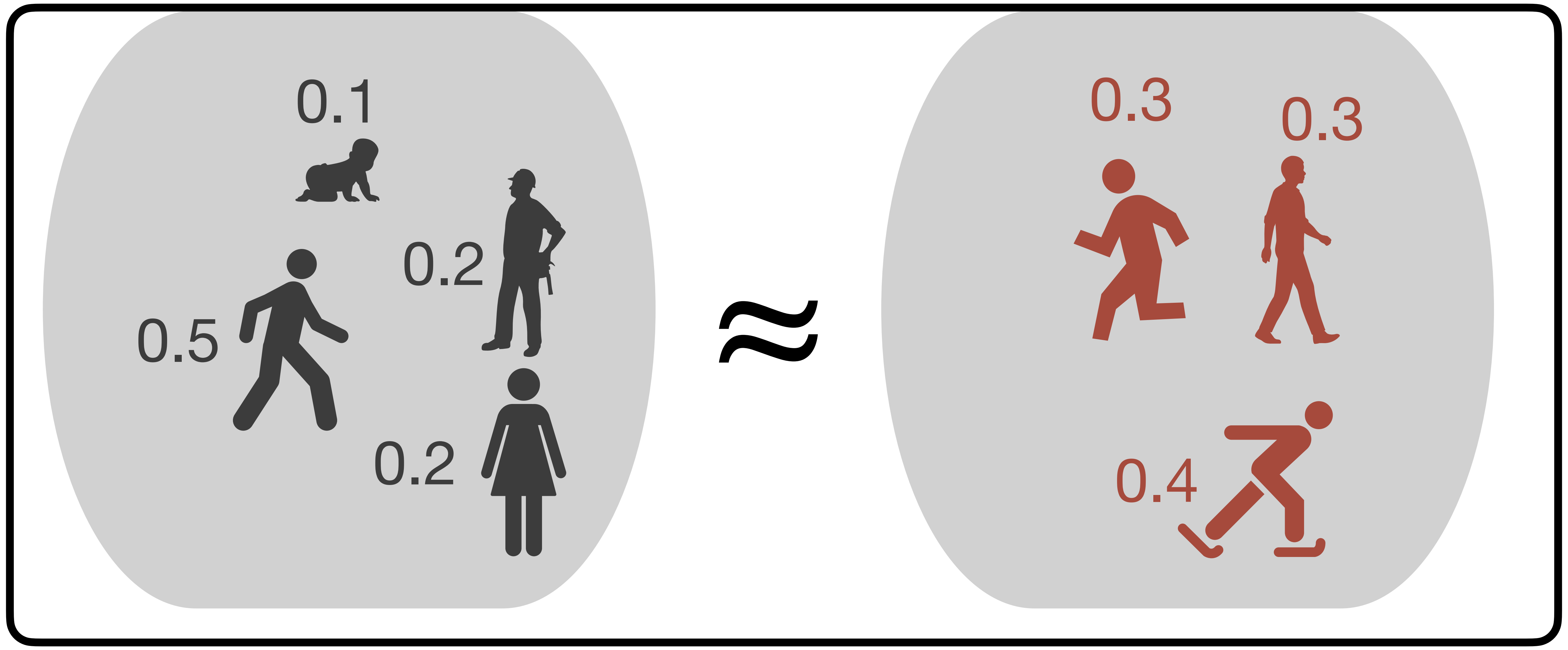
Doudchenko N, Khosravi K, Pouget-Abadie J, et al. Synthetic Design: An Optimization Approach to Experimental Design with Synthetic Controls. Advances in Neural Information Processing Systems, 2021, 34.

Synthetic Design

Weighted Covariate Balancing

Min $\|Y(w \odot d)\|_2^2$

s.t. $\|w \odot d\|_1 = 1, \boxed{1^\top (w \odot d) = 0}$

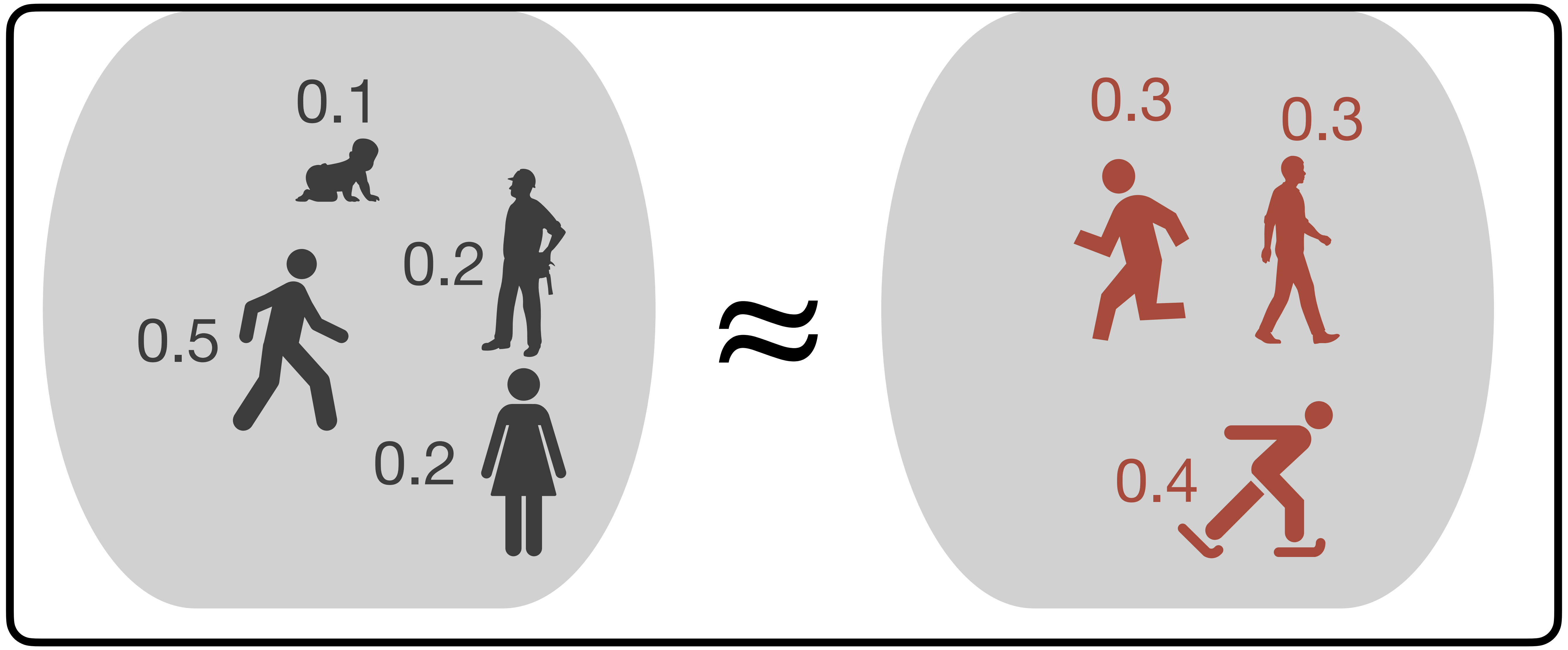


Synthetic Design

Doudchenko N, Khosravi K, Pouget-Abadie J, et al. Synthetic Design: An Optimization Approach to Experimental Design with Synthetic Controls. Advances in Neural Information Processing Systems, 2021, 34.

Synthetic Design

Weighted Covariate Balancing



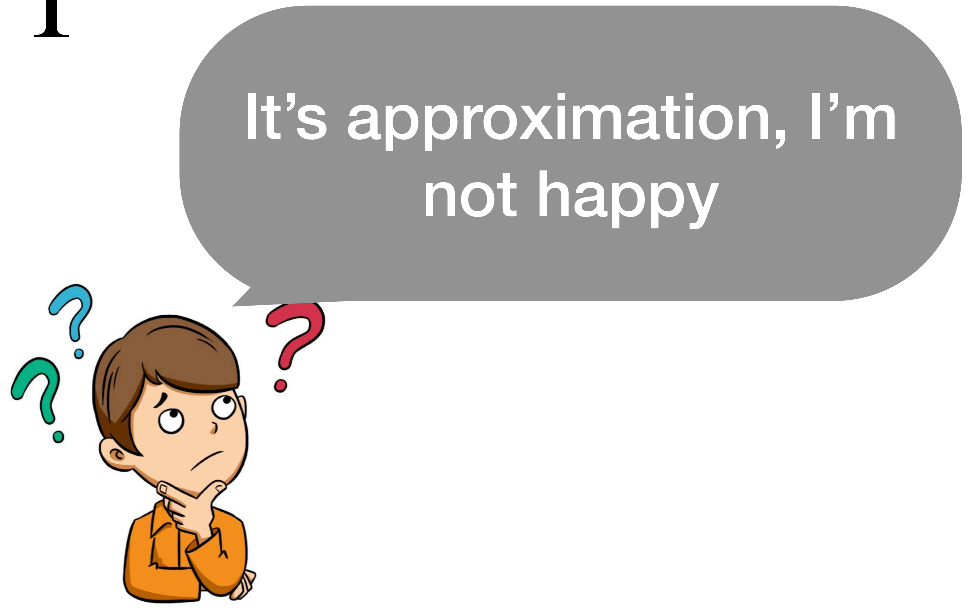
Synthetic Design

Min $\|Y(w \odot d)\|_2^2$

s.t. $\|w \odot d\|_1 = 1, \boxed{1^\top (w \odot d) = 0}$

Min $(w \odot d)^\top (Y^\top Y + \boxed{\lambda 1 1^\top})(w \odot d)$

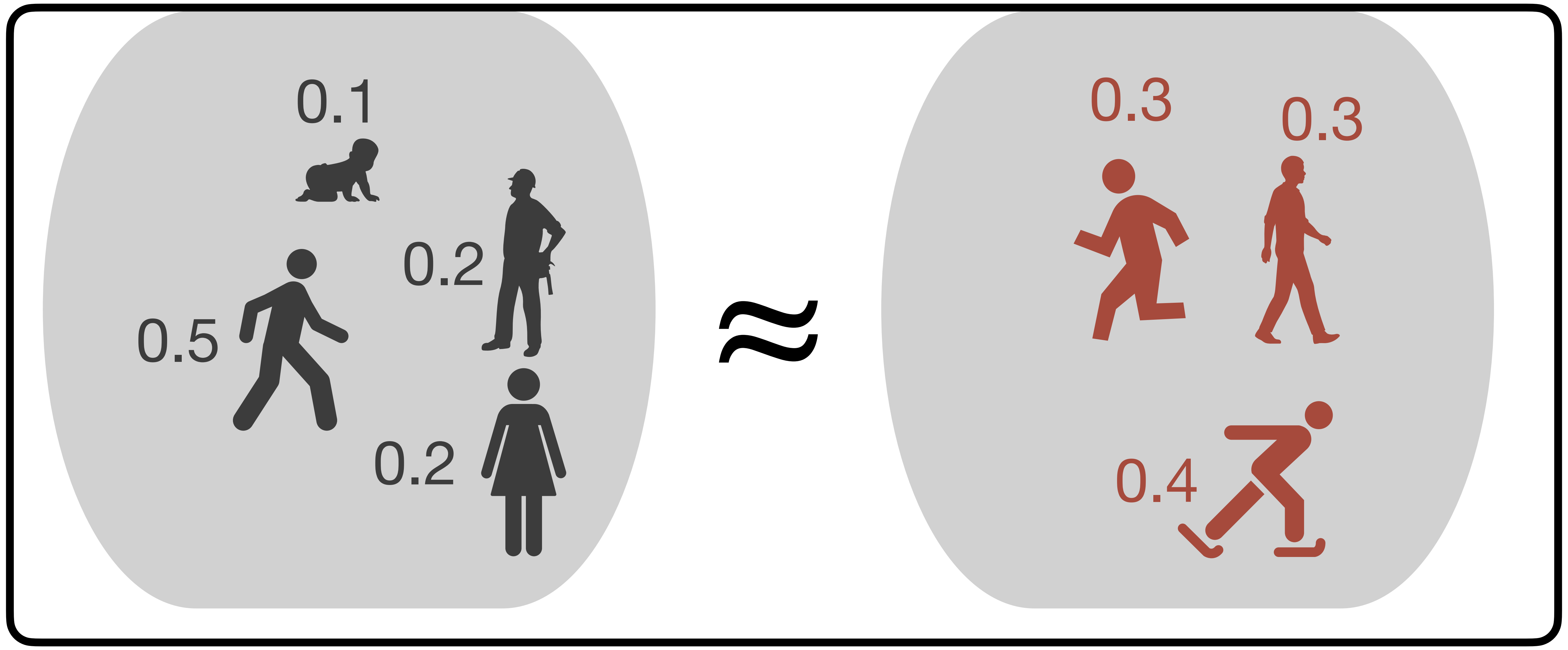
s.t. $\|w \odot d\|_1 = 1$



Doudchenko N, Khosravi K, Pouget-Abadie J, et al. Synthetic Design: An Optimization Approach to Experimental Design with Synthetic Controls. Advances in Neural Information Processing Systems, 2021, 34.

Synthetic Design

Weighted Covariate Balancing



Synthetic Design

If knows the sign, it’s convex!

Min $\|Y(w \odot d)\|_2^2$

s.t. $\|w \odot d\|_1 = 1, \boxed{1^\top (w \odot d) = 0}$

Min $(w \odot d)^\top (Y^\top Y + \boxed{\lambda 1 1^\top})(w \odot d)$

s.t. $\|w \odot d\|_1 = 1$

Theorem

If λ is large enough, the sign of the two solution are the same

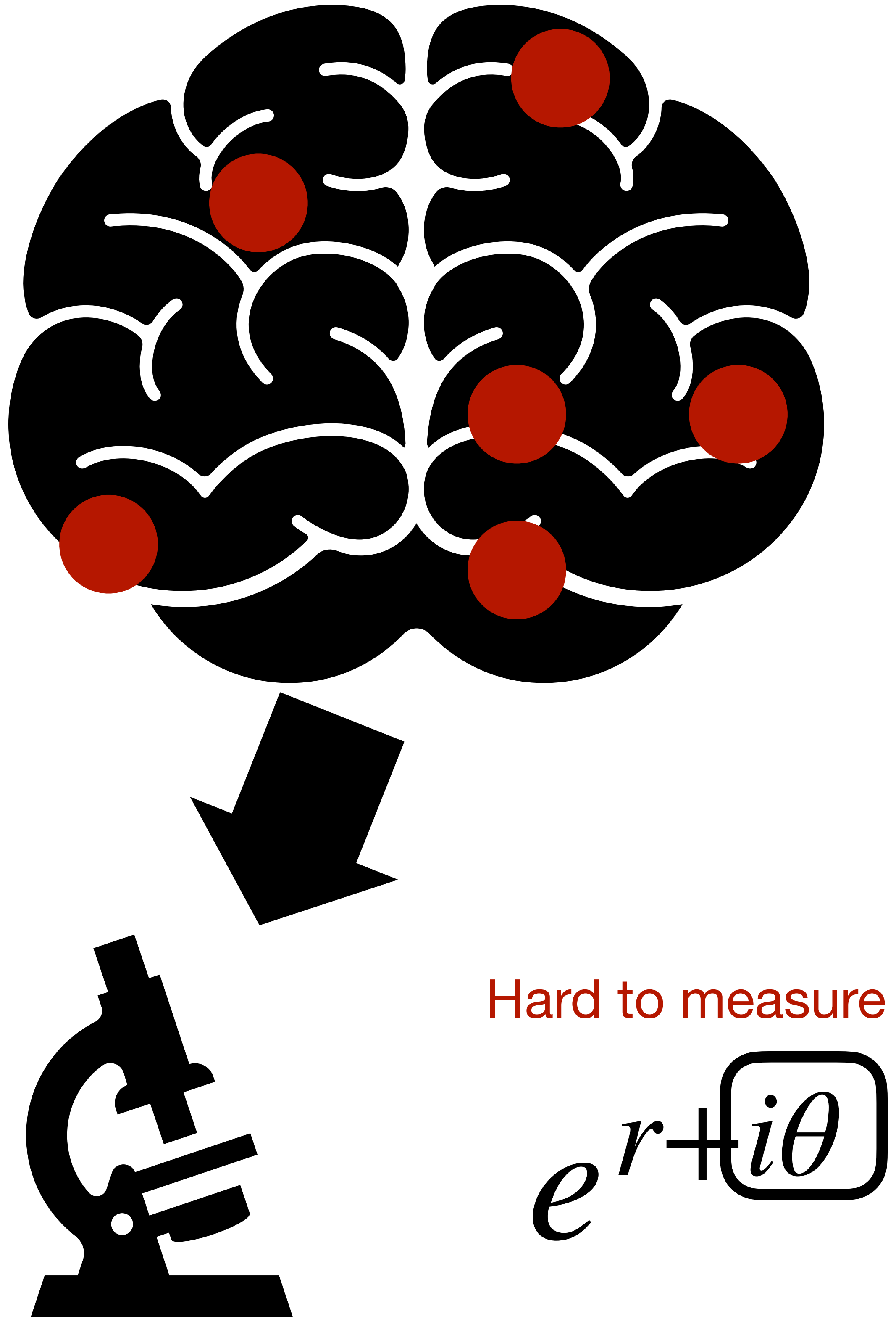
Doudchenko N, Khosravi K, Pouget-Abadie J, et al. Synthetic Design: An Optimization Approach to Experimental Design with Synthetic Controls. Advances in Neural Information Processing Systems, 2021, 34.

The second reformulation

Equal to Phase Synchronization

Phase Synchronization

$$\max_{\|x\|_2=1} \|Ax\|_1 = \max_{\|x\|_2=1, y\in\{-1,+1\}} y^\top Ax = \boxed{\max_{y\in\{-1,+1\}} \|A^\top y\|_2}$$



Hard to measure

$$e^{r+\boxed{i\theta}}$$

Singer A. Angular synchronization by eigenvectors and semidefinite programming. Applied and computational harmonic analysis, 2011, 30(1): 20-36.

The second reformulation

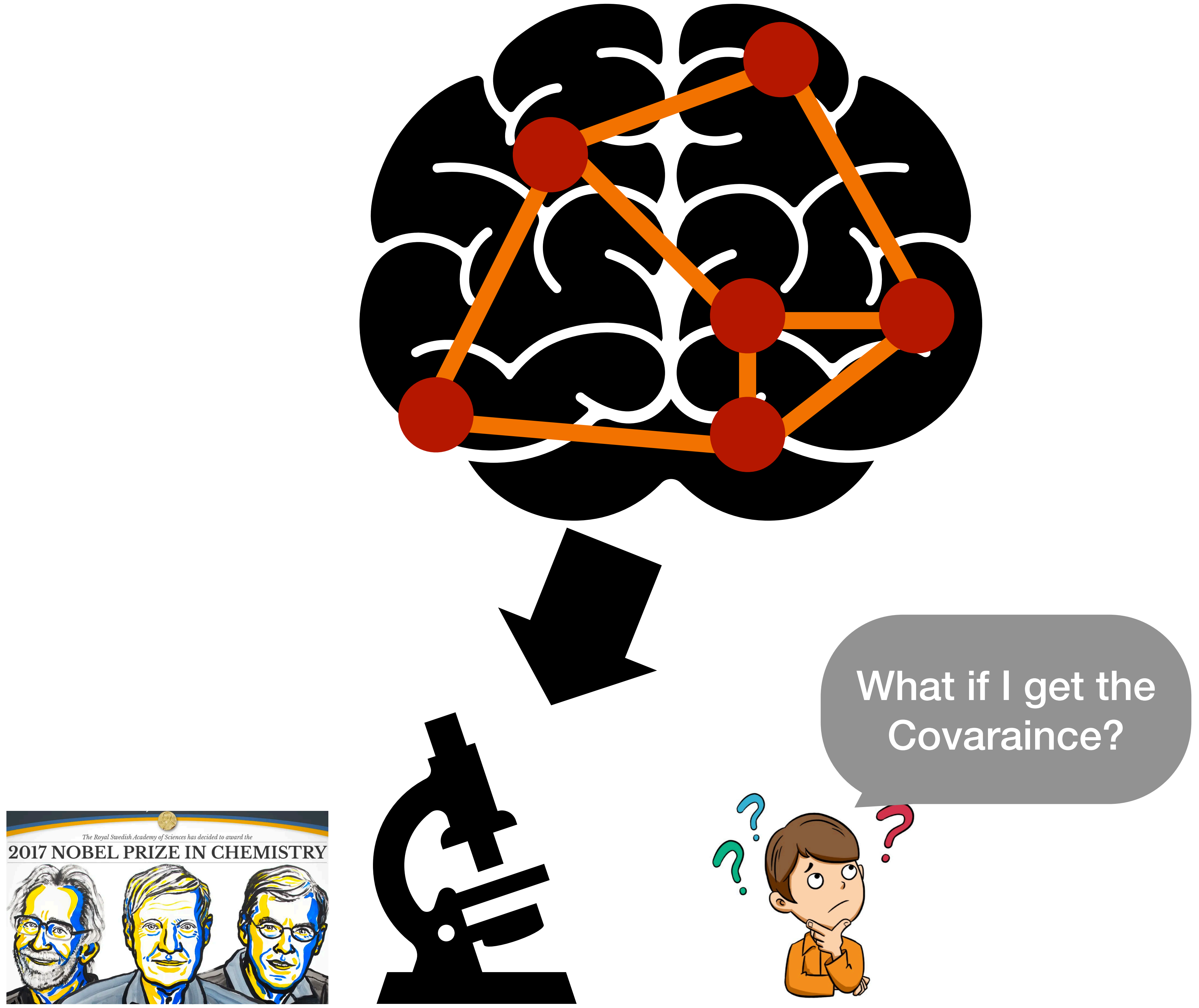
Equal to Phase Synchronization

Phase Synchronization

$$\max_{\|x\|_2=1} \|Ax\|_1 = \max_{\|x\|_2=1, y\in\{-1,+1\}} y^\top Ax = \max_{\boxed{y\in\{-1,+1\}}} \boxed{\|A^\top y\|_2}$$

Match covariance

Find phase



Basic idea behind Cro-EM (Nobel Prize 2017)

Singer A. Angular synchronization by eigenvectors and semidefinite programming. Applied and computational harmonic analysis, 2011, 30(1): 20-36.

The second reformulation

Equal to Phase Synchronization

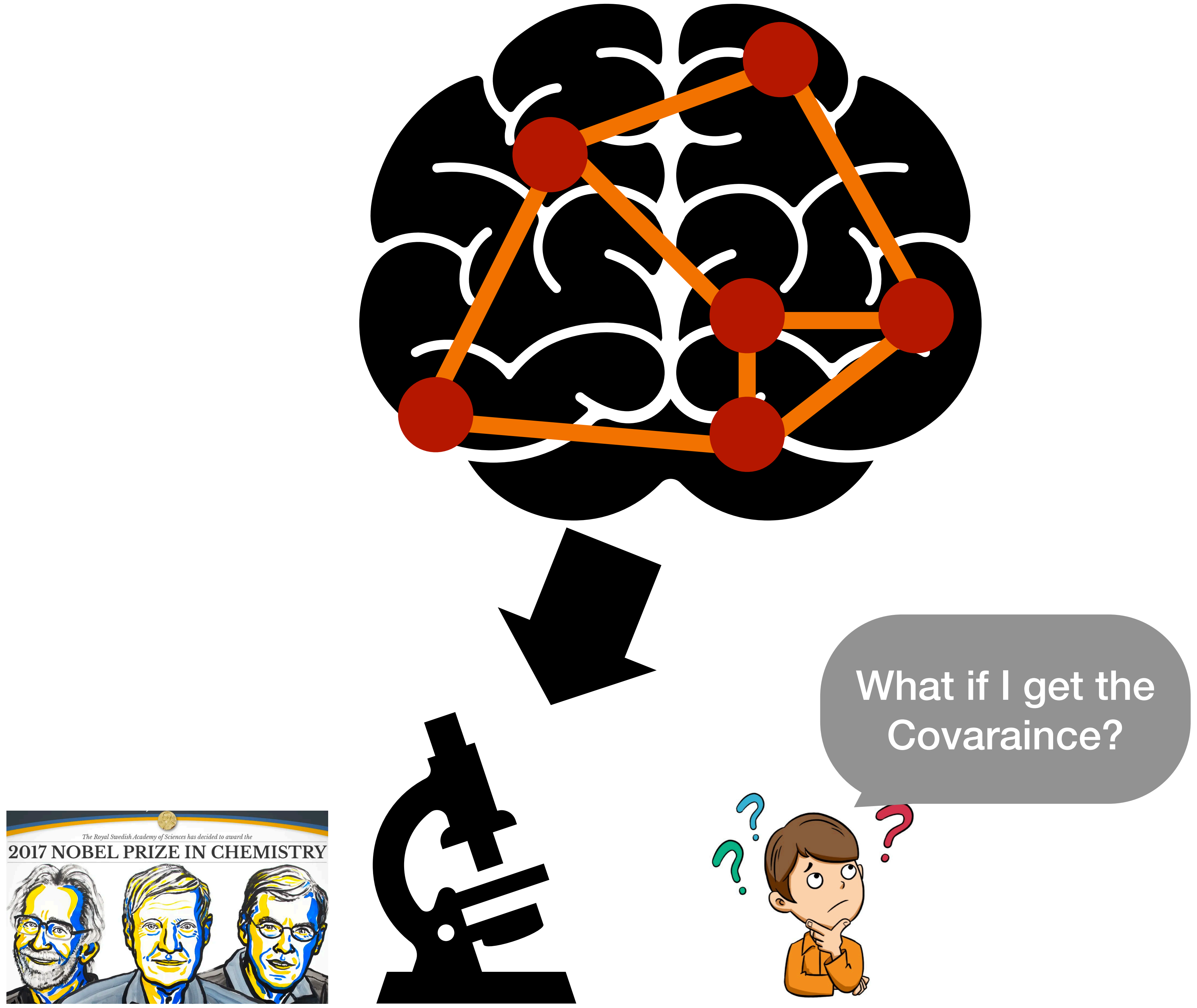
Phase Synchronization

$$\max_{\|x\|_2=1} \|Ax\|_1 = \max_{\|x\|_2=1,y\in\{-1,+1\}} y^\top Ax = \max_{\boxed{y\in\{-1,+1\}}} \boxed{\|A^\top y\|_2}$$

Match covariance

Find phase

Still provable NP-hard



Basic idea behind Cro-EM (Nobel Prize 2017)

Boumal N. Nonconvex phase synchronization. SIAM Journal on Optimization, 2016, 26(4): 2355-2377.

The second reformulation

Equal to Phase Synchronization

Phase Synchronization

$$\max_{\|x\|_2=1} \|Ax\|_1 = \max_{\|x\|_2=1, y\in\{-1,+1\}} y^\top Ax = \max_{\boxed{y\in\{-1,+1\}}} \boxed{\|A^\top y\|_2}$$

Match covariance

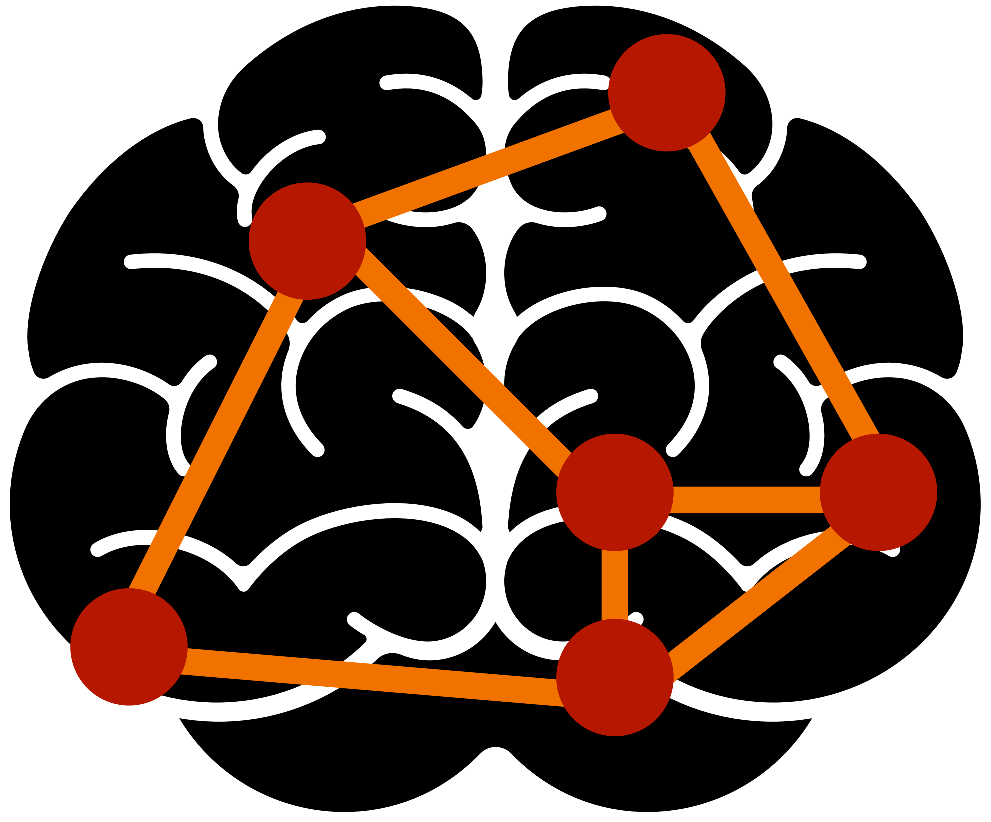
Find phase

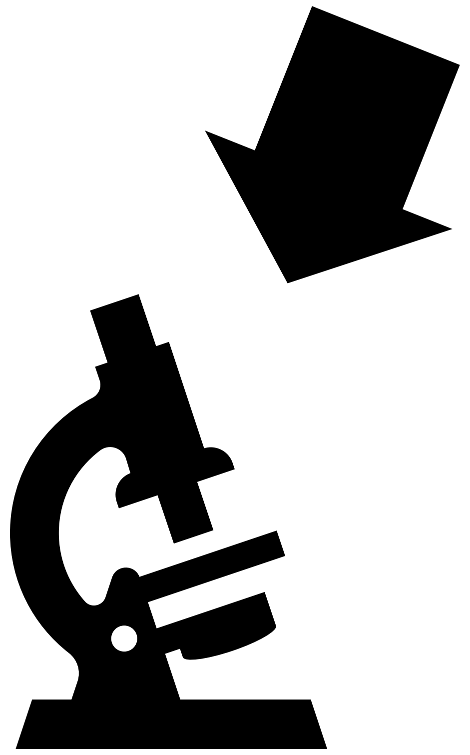
Still provable NP-hard

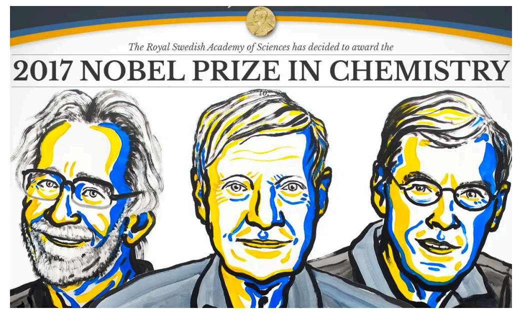
Algorithm

Step 1. Relax $y \in \{-1,1\}$ to $\|y\|_2^2 = n$ and change it to Eigenvalue problem.

Econ intuition: Experiment through Smallest Principle Component







Basic idea behind Cro-EM (Nobel Prize 2017)

Boumal N. Nonconvex phase synchronization. SIAM Journal on Optimization, 2016, 26(4): 2355-2377.

The second reformulation

Equal to Phase Synchronization

Phase Synchronization

$$\max_{\|x\|_2=1} \|Ax\|_1 = \max_{\|x\|_2=1, y\in\{-1,+1\}} y^\top Ax = \max_{\boxed{y\in\{-1,+1\}}} \boxed{\|A^\top y\|_2}$$

Match covariance

Find phase

Still provable NP-hard

Algorithm

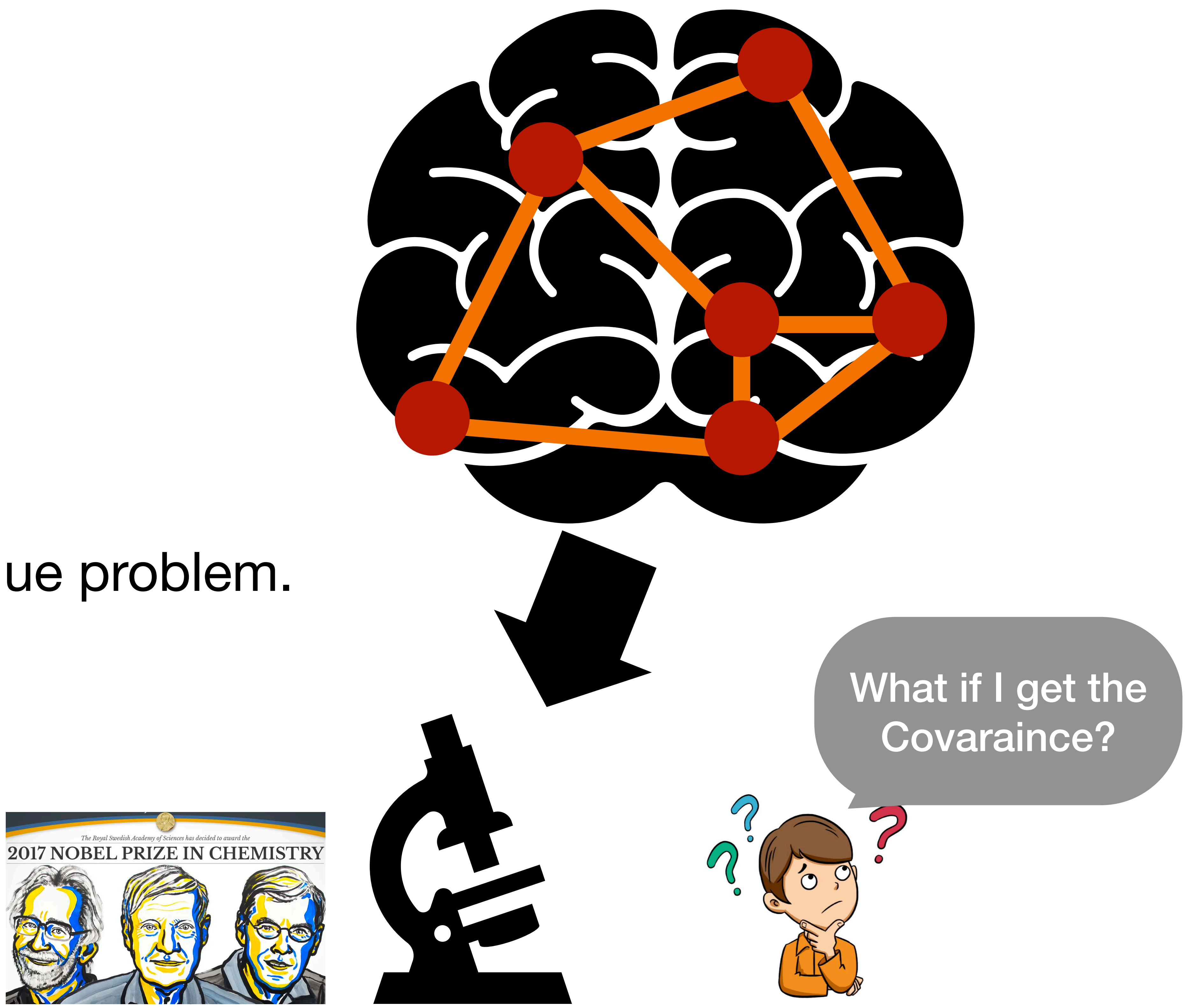
Step 1. Relax $y \in \{-1,1\}$ to $\|y\|_2^2 = n$ and change it to Eigenvalue problem.

Step 2. Local Refinement via Power Method.

$$y^k = \boxed{\mathrm{sgn}(}(\boxed{AA^\top + \alpha I})y^{k-1})$$

Projection Back

Power Method



Basic idea behind Cro-EM (Nobel Prize 2017)

Boumal N. Nonconvex phase synchronization. SIAM Journal on Optimization, 2016, 26(4): 2355-2377.

The second reformulation

Equal to Phase Synchronization

Phase Synchronization

$$\max_{\|x\|_2=1} \|Ax\|_1 = \max_{\|x\|_2=1, y\in\{-1,+1\}} y^\top Ax = \max_{\boxed{y\in\{-1,+1\}}} \boxed{\|A^\top y\|_2}$$

Match covariance

Find phase

Still provable NP-hard

Algorithm

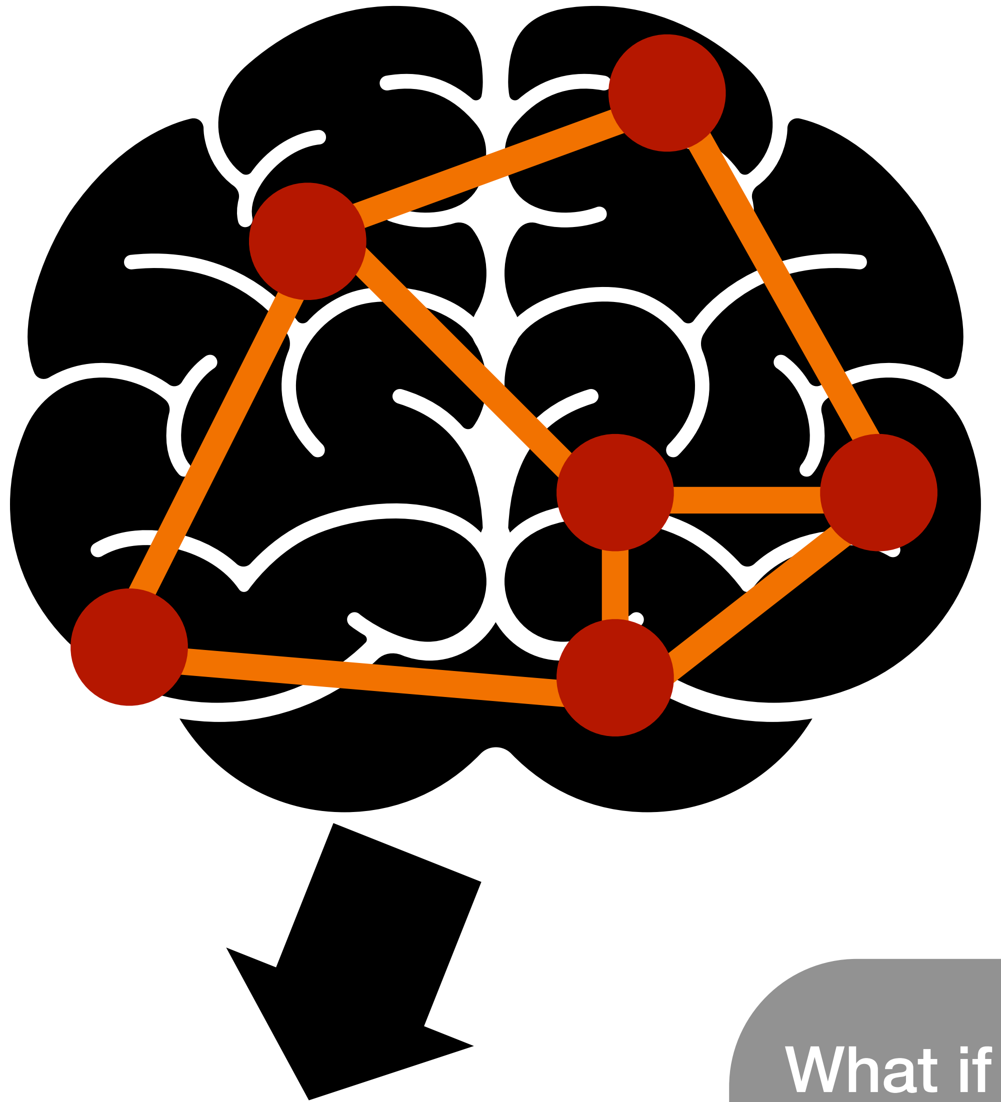
Step 1. Relax $y \in \{-1,1\}$ to $\|y\|_2^2 = n$ and change it to Eigenvalue problem.

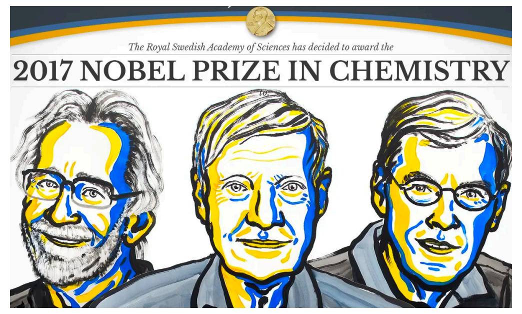
Step 2. Local Refinement via Power Method.

$$y^k = \operatorname{sgn}((\boxed{AA^\top} + \alpha I)y^{k-1})$$

Inverse of the covariance matrix

Generalized Inverse Power Method !





Basic idea behind Cro-EM (Nobel Prize 2017)

Boumal N. Nonconvex phase synchronization. SIAM Journal on Optimization, 2016, 26(4): 2355-2377.

The second reformulation

Equal to Phase Synchronization

Phase Synchronization

$$\max_{\|x\|_2=1} \|Ax\|_1 = \max_{\|x\|_2=1, y\in\{-1,+1\}} y^\top Ax = \max_{\boxed{y\in\{-1,+1\}}} \boxed{\|A^\top y\|_2}$$

Match covariance

Find phase

Still provable NP-hard

Best experiment: Smallest "Eigen" !

Algorithm

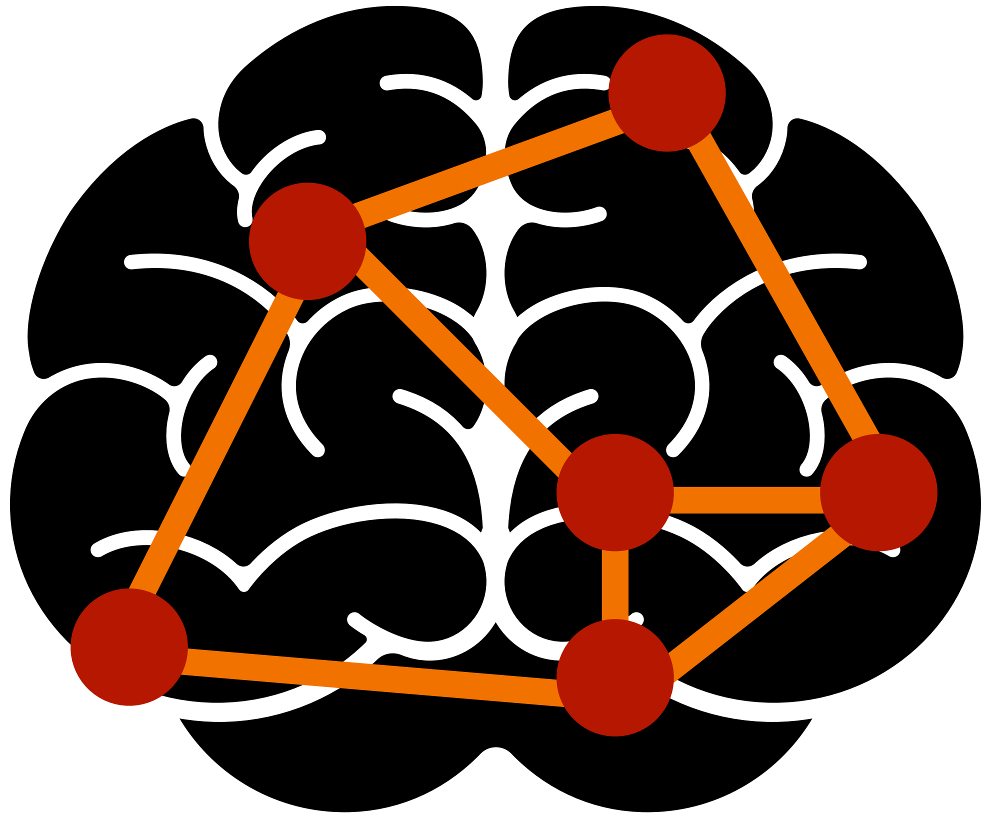
Step 1. Relax $y \in \{-1,1\}$ to $\|y\|_2^2 = n$ and change it to Eigenvalue problem.

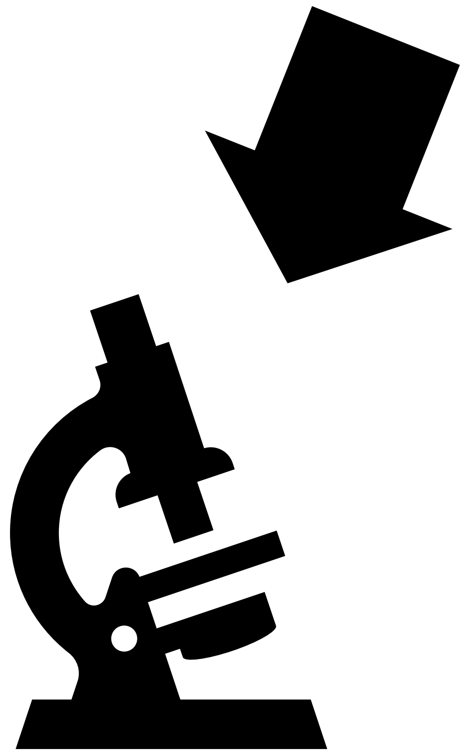
Step 2. Local Refinement via Power Method.

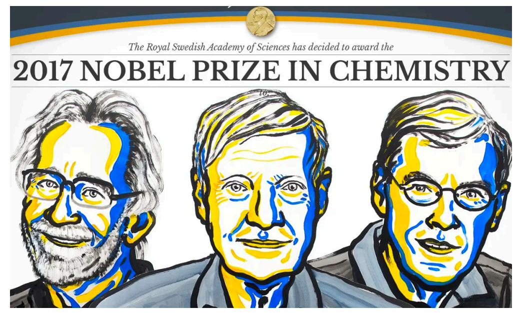
$$y^k = \mathrm{sgn}((\boxed{AA^\top} + \alpha I)y^{k-1})$$

Inverse of the covariance matrix

Generalized Inverse Power Method !





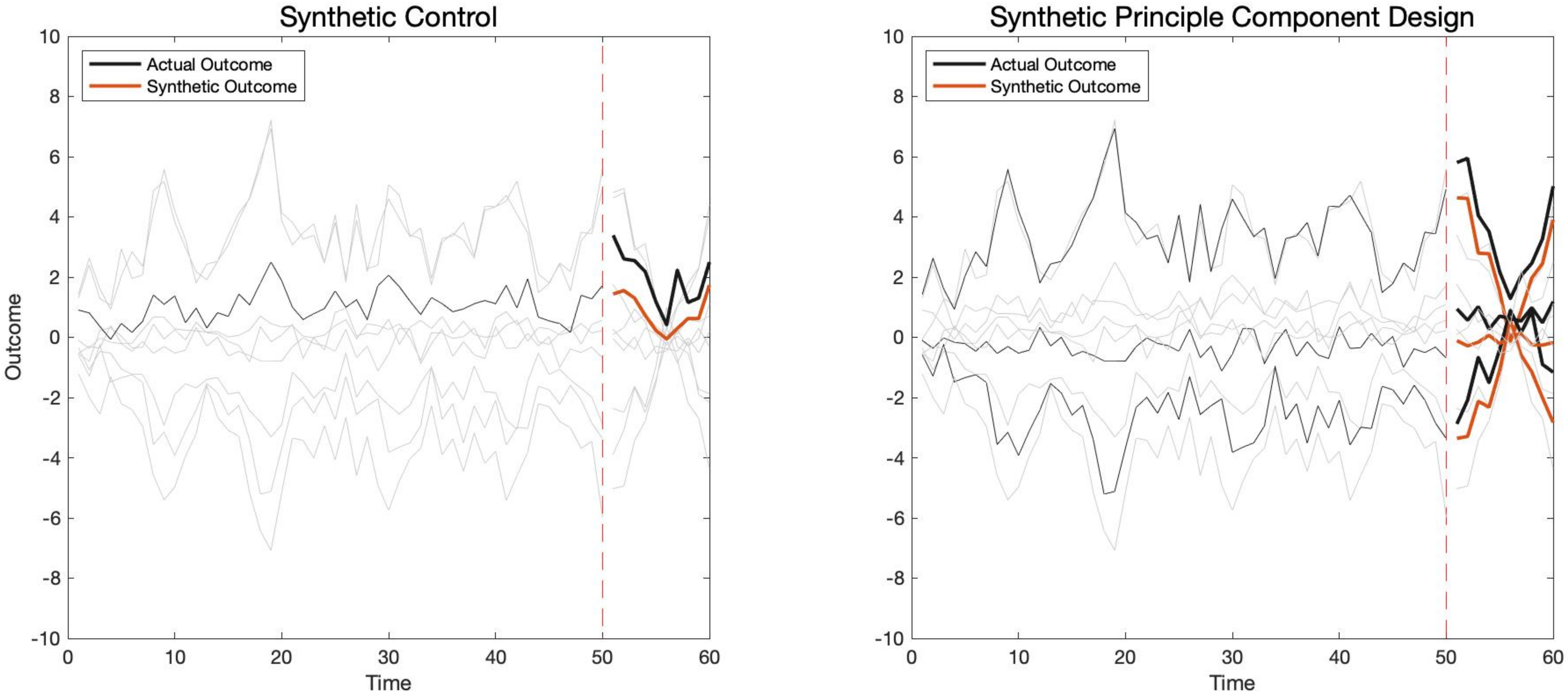


Basic idea behind Cro-EM (Nobel Prize 2017)

Boumal N. Nonconvex phase synchronization. SIAM Journal on Optimization, 2016, 26(4): 2355-2377.

Designed Experiment

“**representative**” agents in market



AR(1) Process

The second reformulation

Equal to Phase Synchronization

$$\max_{\|x\|_2=1} \|Ax\|_1 = \boxed{\max_{\|x\|_2=1, y\in\{-1,+1\}} y^\top Ax} = \max_{y\in\{-1,+1\}} \|A^\top y\|_2$$

$\underline{x^* = A^\top y}$

The second reformulation

Equal to Phase Synchronization

$$\max_{\|x\|_2=1} \|Ax\|_1 = \boxed{\max_{\|x\|_2=1, y\in\{-1,+1\}} y^\top Ax} = \max_{y\in\{-1,+1\}} \|A^\top y\|_2$$

$\underline{x^* = A^\top y}$

Input Optimal experiment profile y

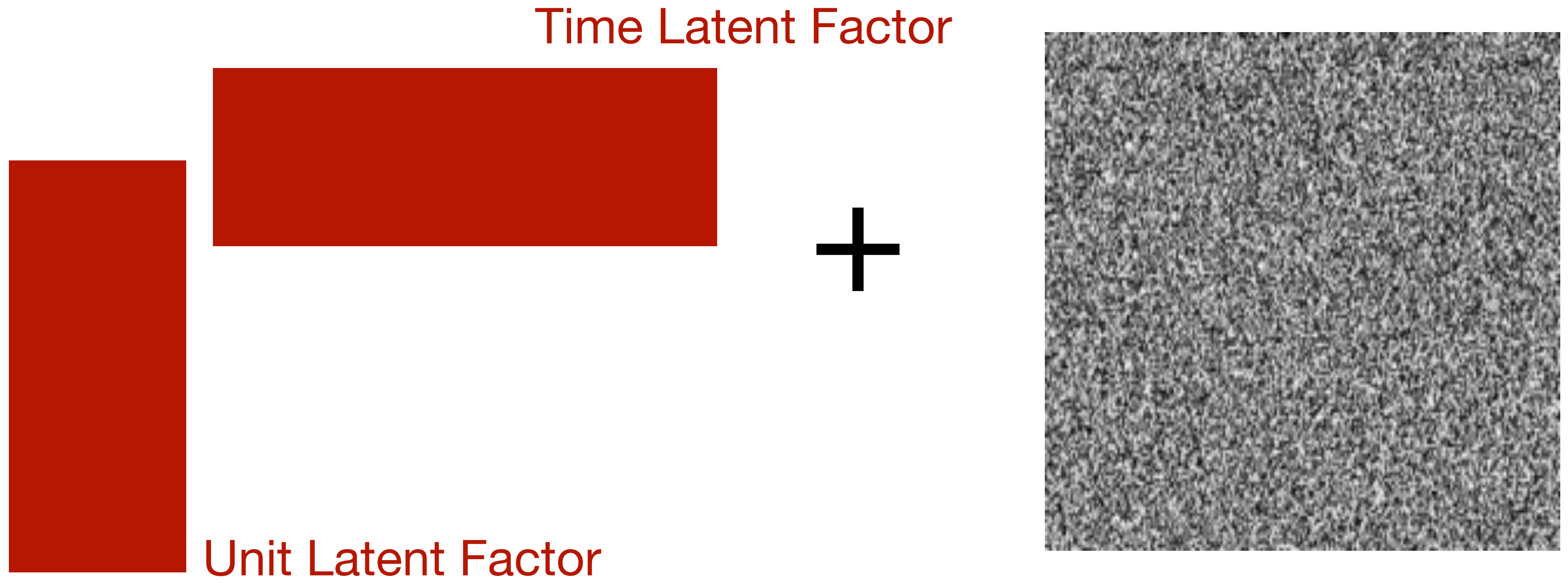
Estimator Weight $w = \Sigma^{-1} y$ Optimality condition leads to sgn$(w) = y$

Final Estimation $\tau = w \times \boxed{\text{(post-treamtnet outcome)}}$

Principle Component Design

Simulated Data

$$Y_{\text{Unit,Time}} =$$



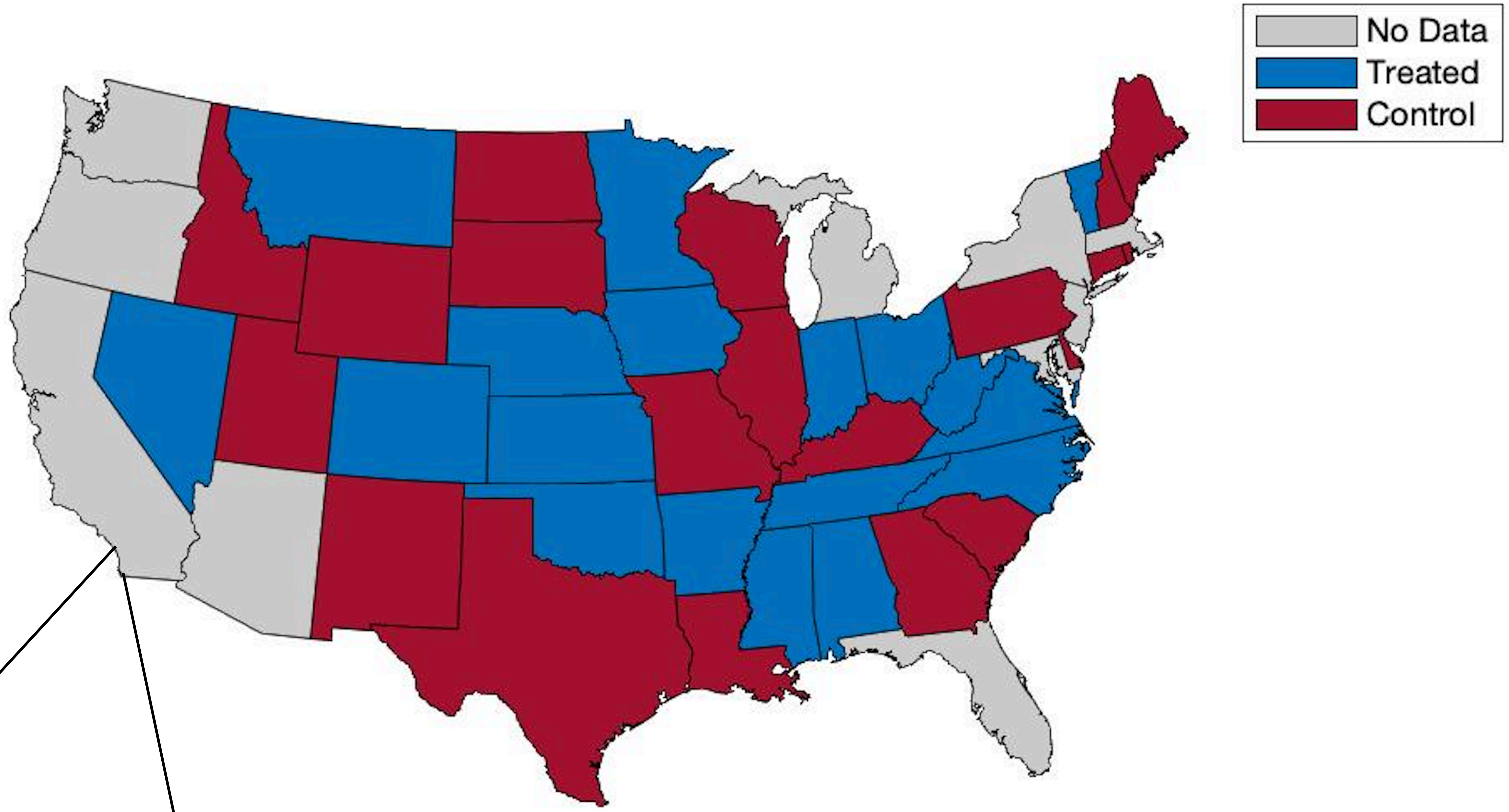
Athey S, Bayati M, Doudchenko N, et al. Matrix completion methods for causal panel data models. Journal of the American Statistical Association, 2021, 116(536): 1716-1730.

Tobacco Control Dataset

Real world dataset

SC	Random	SPCD
7.89	3.13 ±0.19	**0.98**

Random select treated and control group

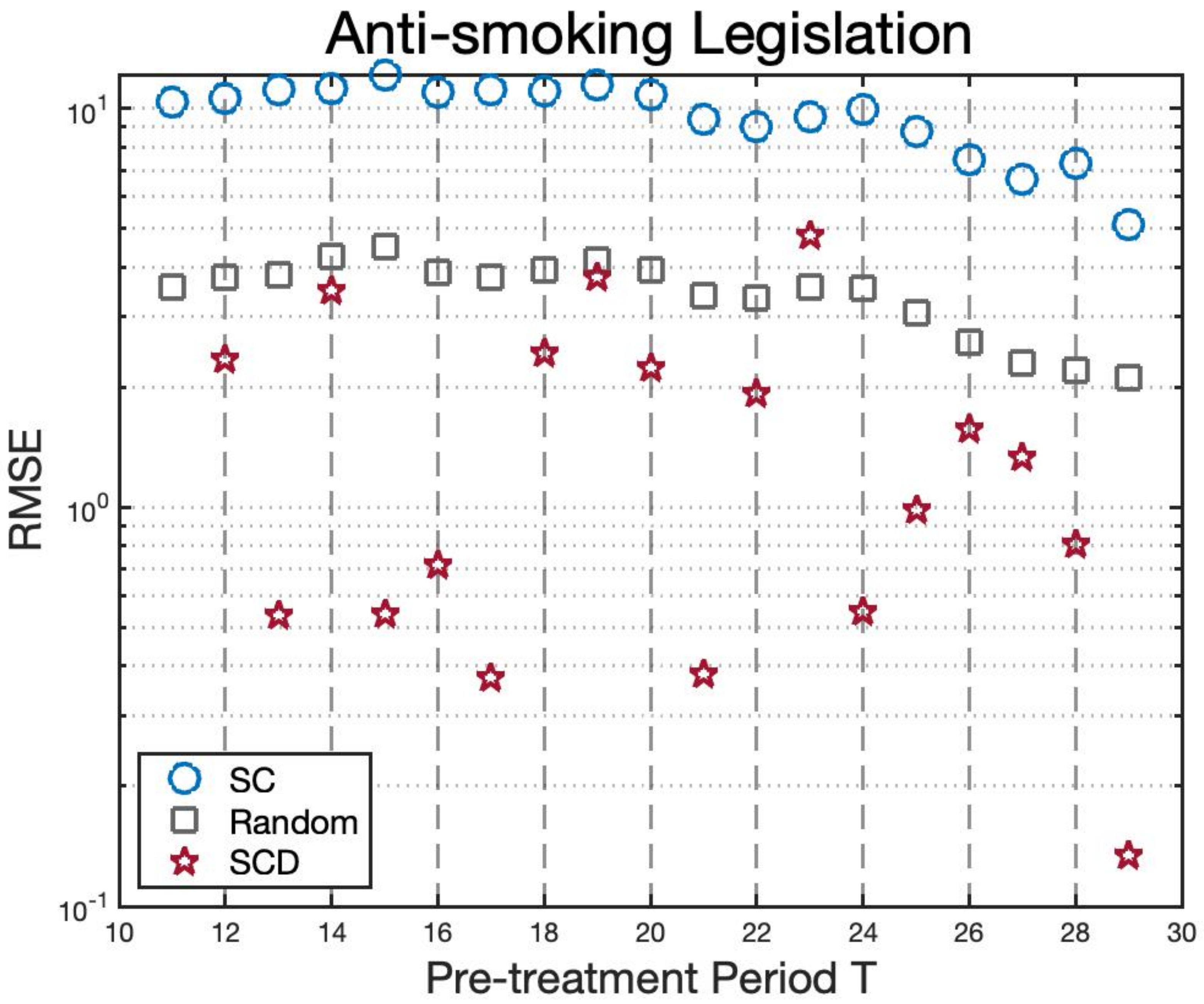


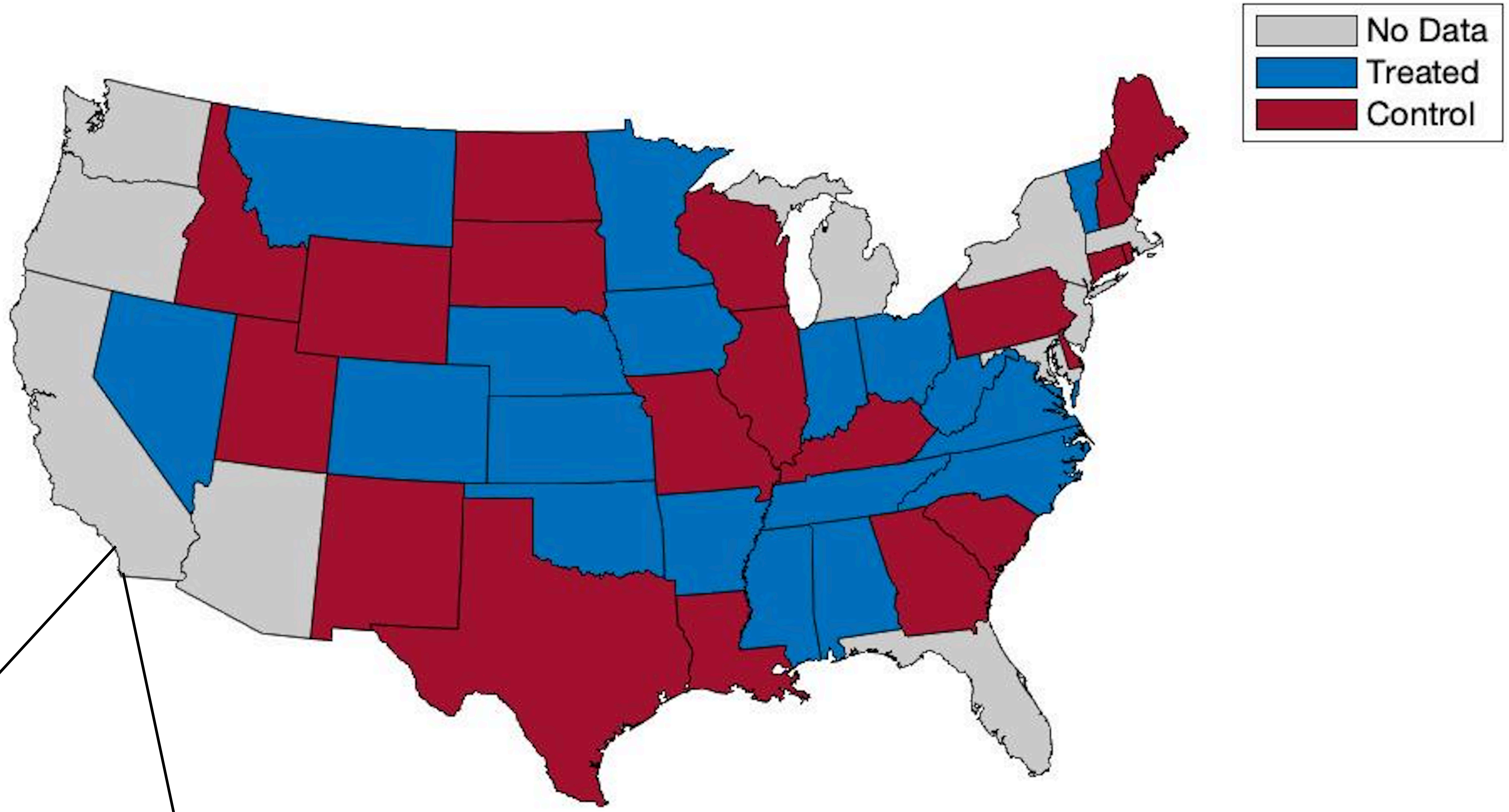
California = 0.334* Utah+0.234*Nevada+0.164*Colorado+0.069*Connecticut

Abadie A, Diamond A, Hainmueller J. Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program. Journal of the American statistical Association, 2010, 105(490): 493-505.

Tobacco Control Dataset

Real world dataset



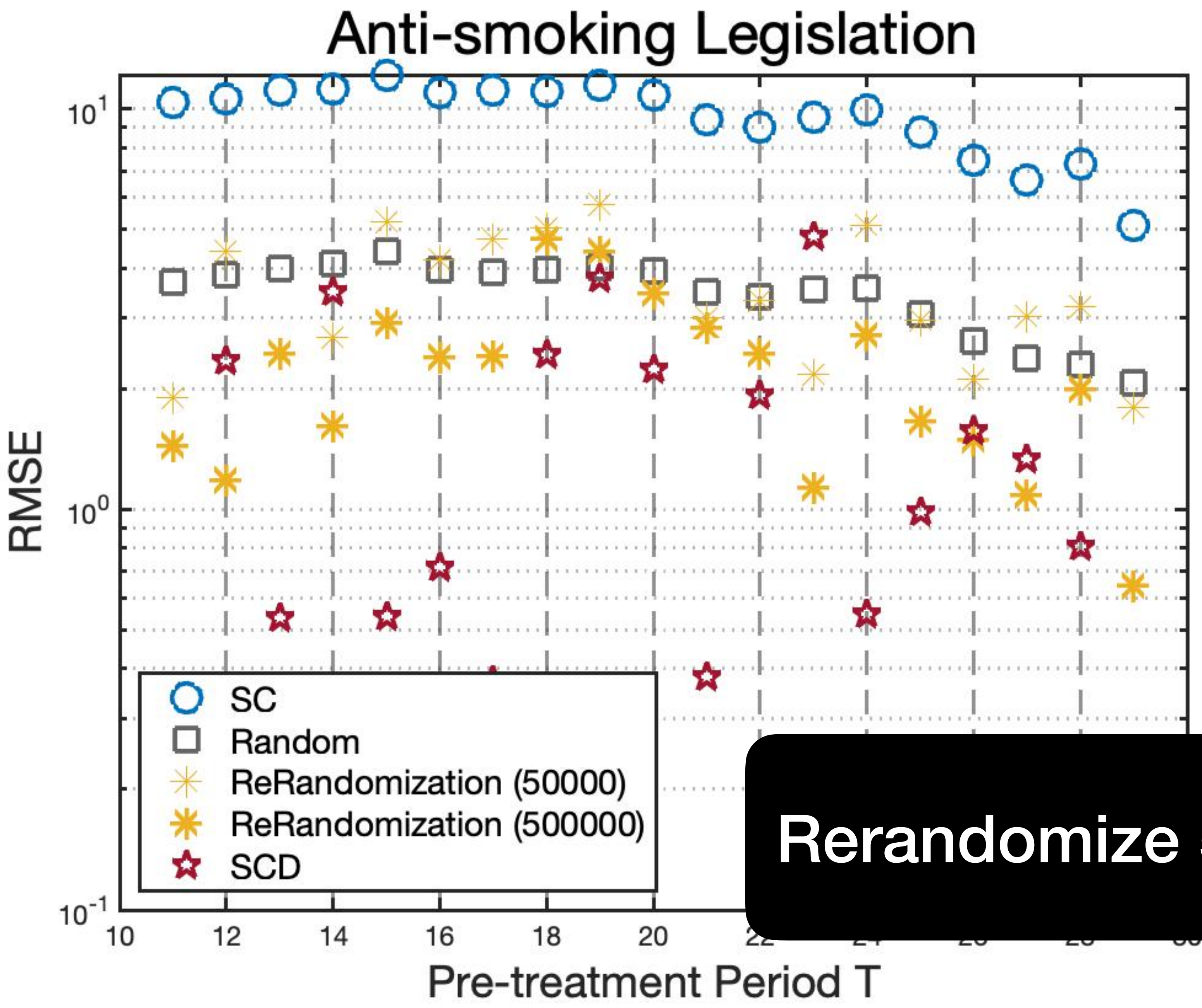


California = 0.334* Utah+0.234*Nevada+0.164*Colorado+0.069*Connecticut

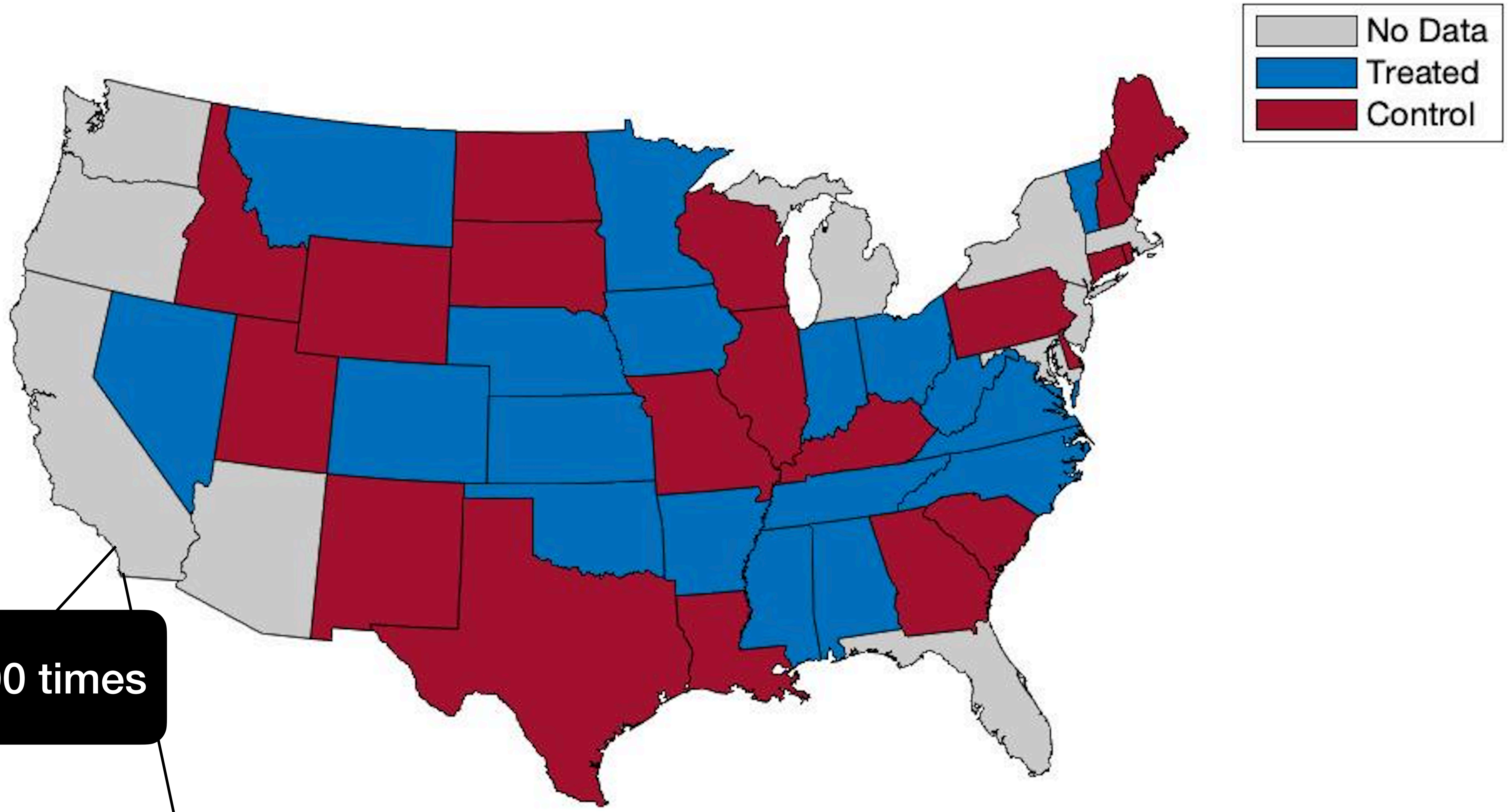
Abadie A, Diamond A, Hainmueller J. Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program. Journal of the American statistical Association, 2010, 105(490): 493-505.

Tobacco Control Dataset

Real world dataset



Rerandomize 500000 times

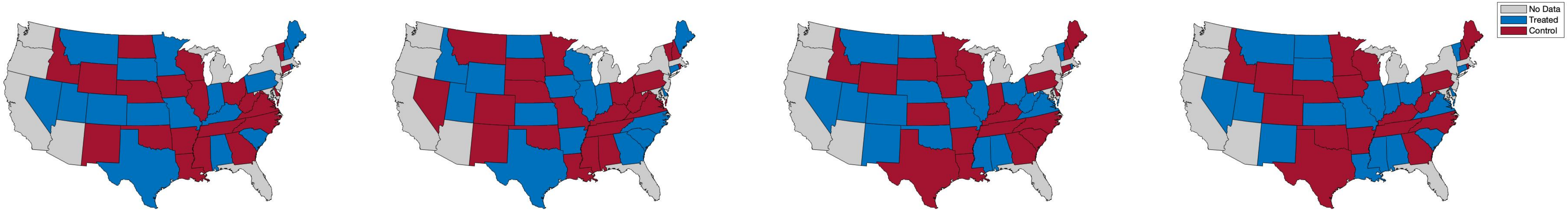


California = 0.334* Utah+0.234*Nevada+0.164*Colorado+0.069*Connecticut

Abadie A, Diamond A, Hainmueller J. Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program. Journal of the American statistical Association, 2010, 105(490): 493-505.

Tobacco Control Dataset

Real world dataset



T=15 T=20 T=25 T=30

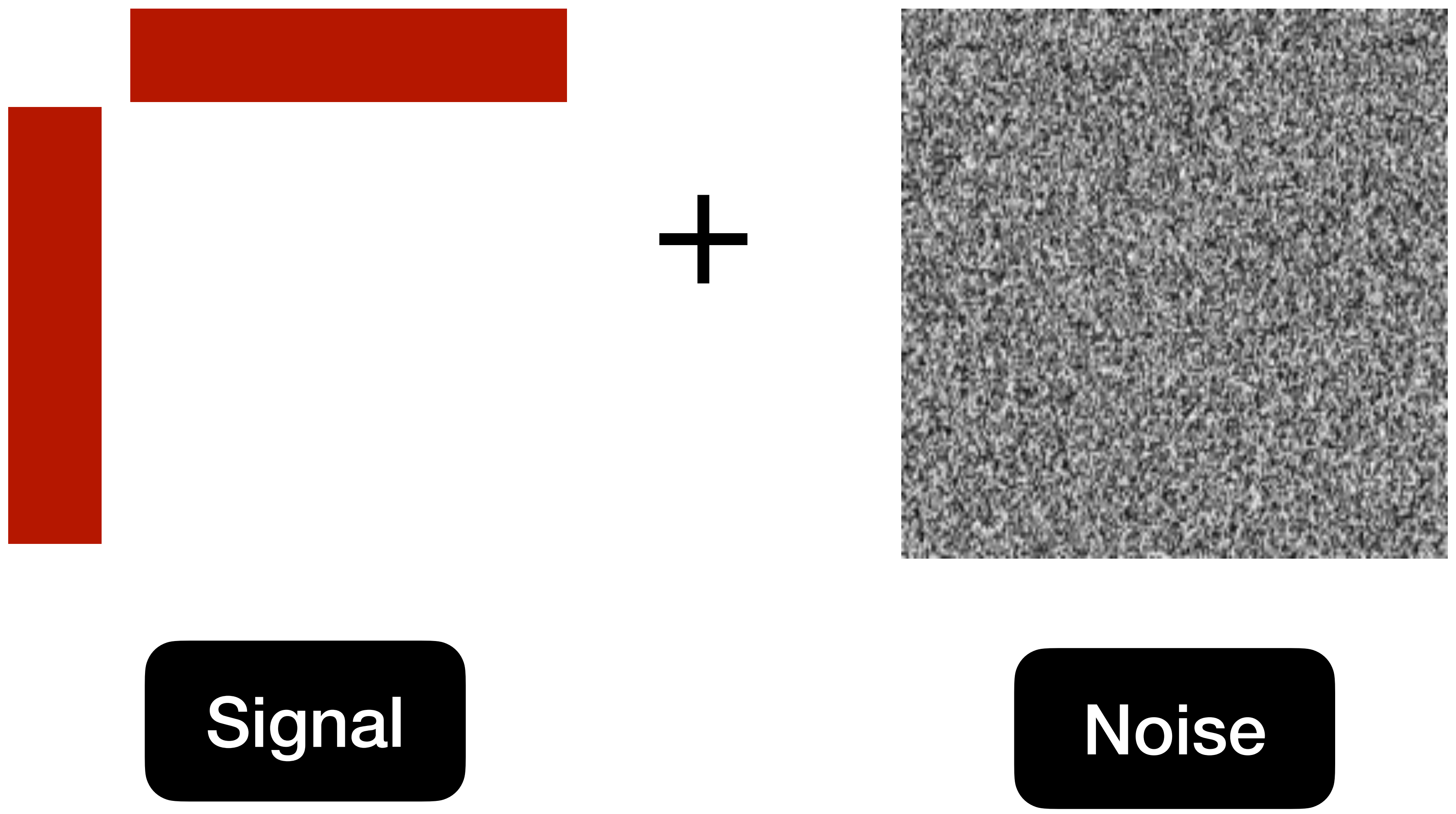
Abadie A, Diamond A, Hainmueller J. Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. Journal of the American statistical Association, 2010, 105(490): 493-505.

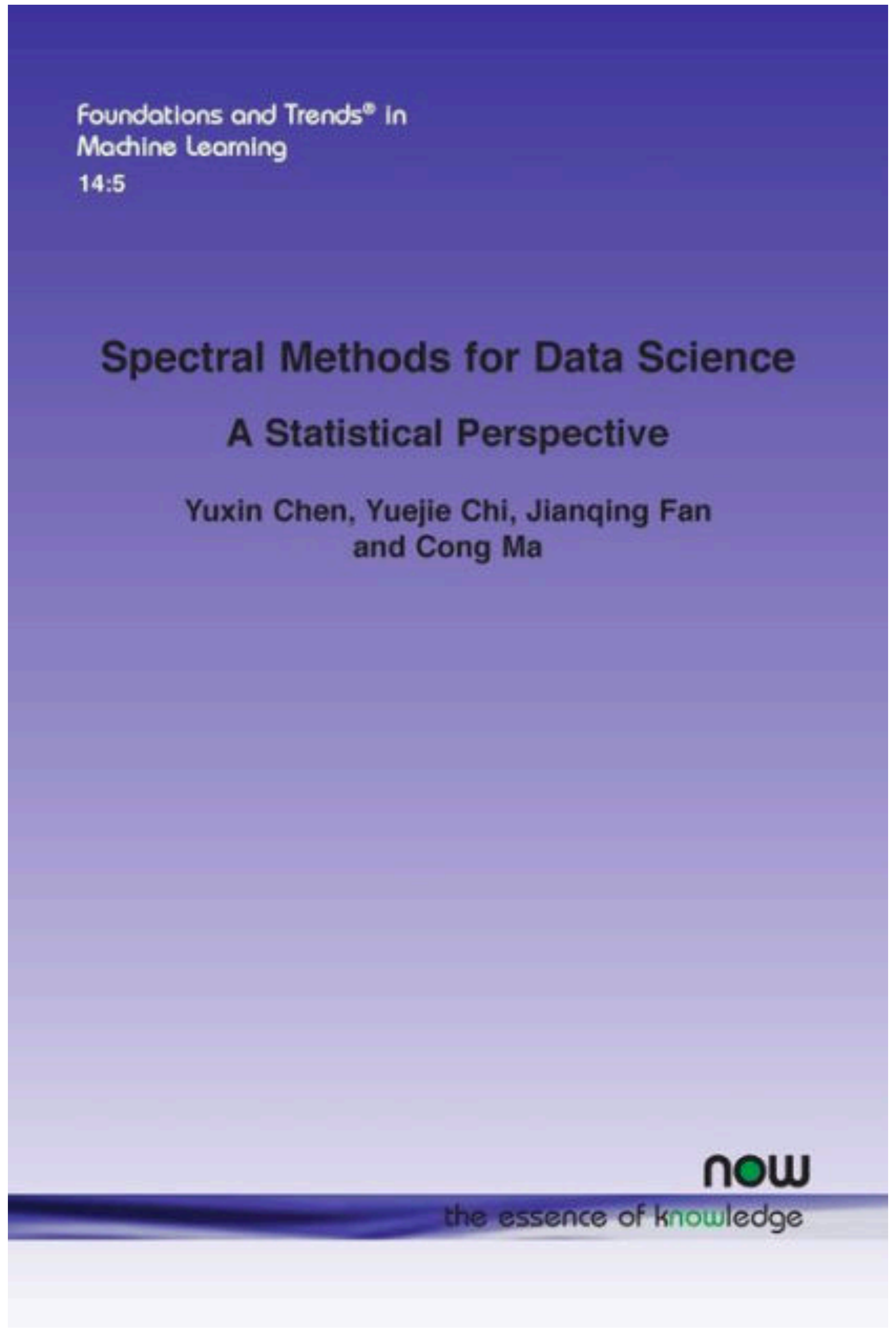
Global for Certain DGP

Spectral Method + Local Improve Meant

Although NP-hard, it’s solvable under certain data generating process (DGP)

Idea




Foundations and Trends® in
Machine Learning
14:5

Spectral Methods for Data Science

A Statistical Perspective

Yuxin Chen, Yuejie Chi, Jianqing Fan
and Cong Ma

now
the essence of knowledge

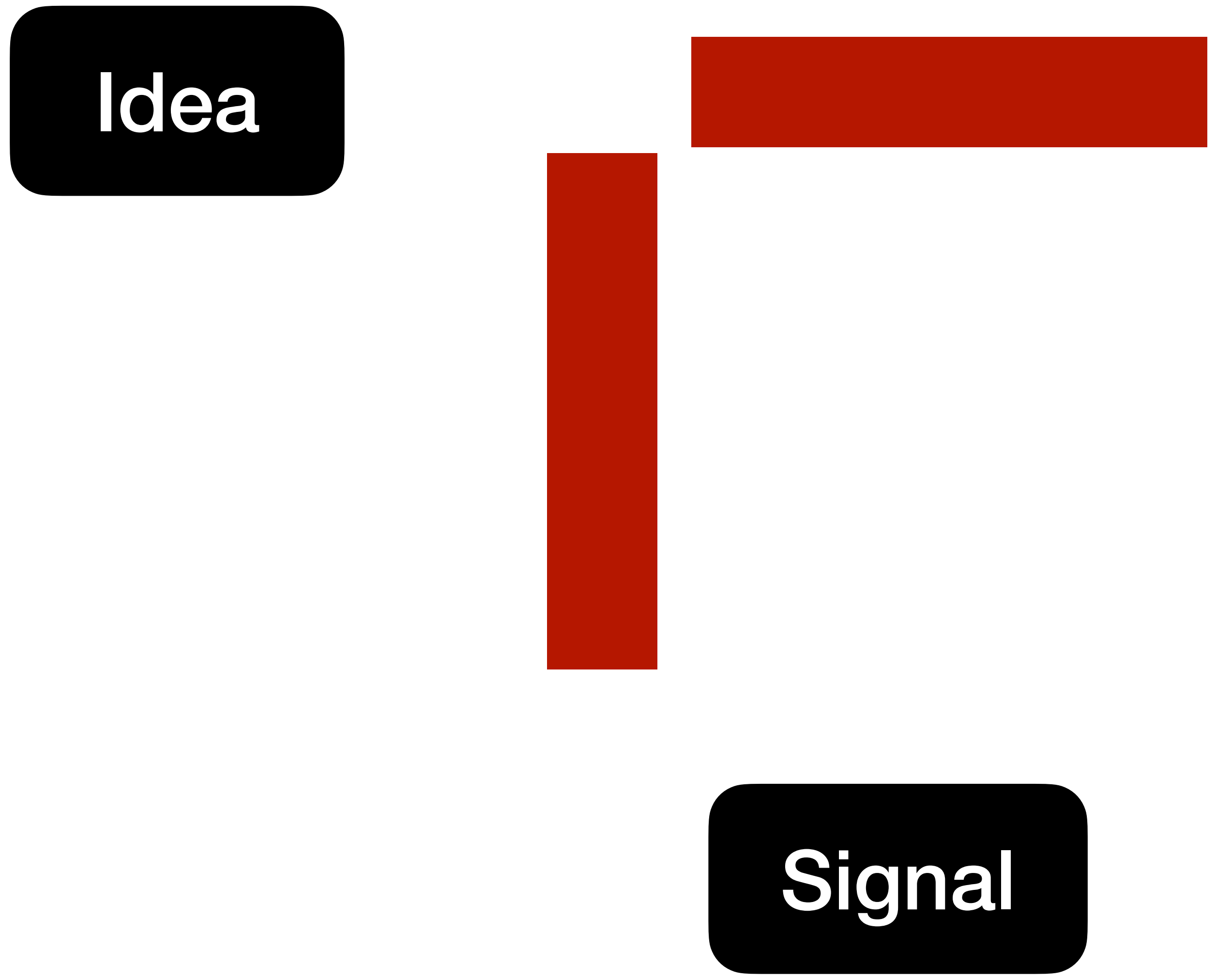
Examples: Phase Synchronization/Retrieval, Matrix Completion, Random Block Model

Chen Y, Chi Y, Fan J, et al. Spectral methods for data science: A statistical perspective. Foundations and Trends® in Machine Learning, 2021, 14(5): 566-806.

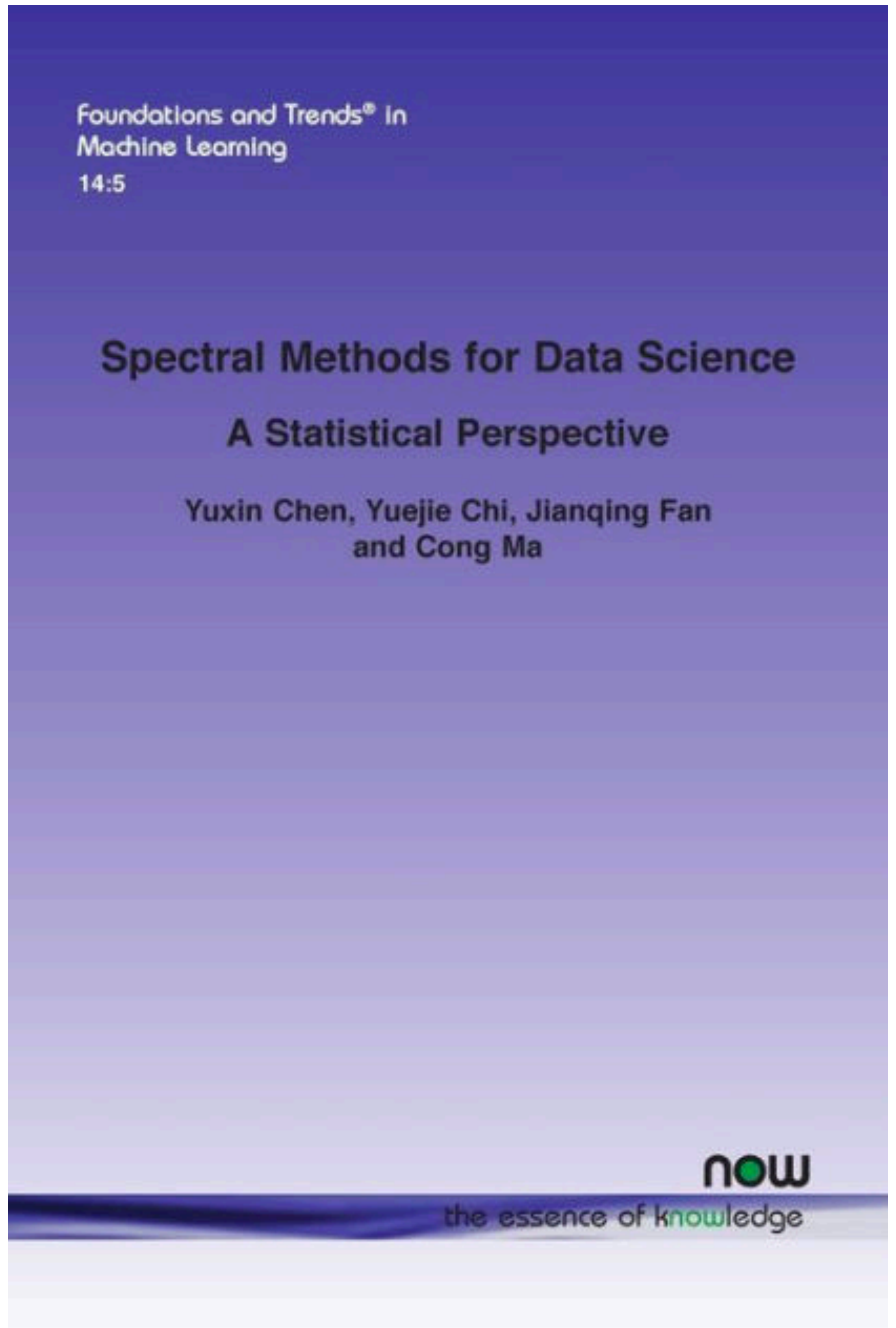
Global for Certain DGP

Spectral Method + Local Improve Meant

Although NP-hard, it’s solvable under certain data generating process (DGP)



Can be solved via spectral method!



Foundations and Trends® in
Machine Learning
14:5

Spectral Methods for Data Science

A Statistical Perspective

Yuxin Chen, Yuejie Chi, Jianqing Fan
and Cong Ma

now
the essence of knowledge

Examples: Phase Synchronization/Retrieval, Matrix Completion, Random Block Model

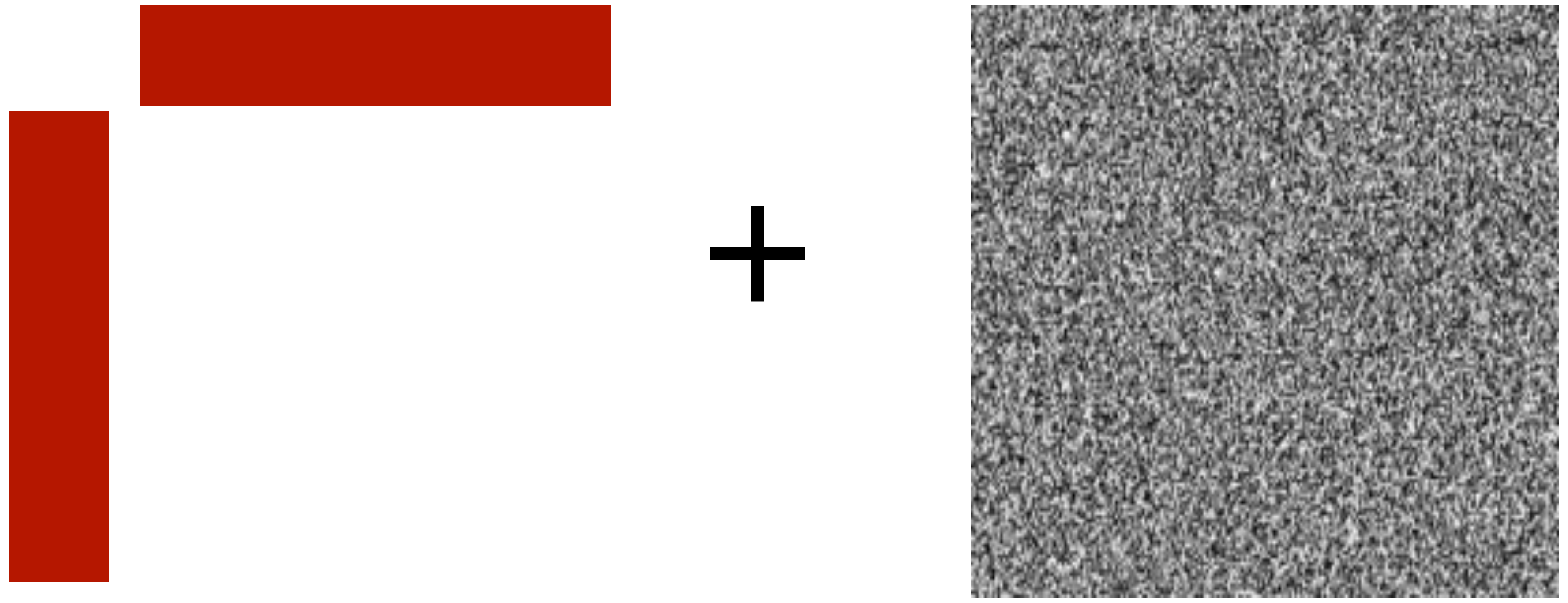
Chen Y, Chi Y, Fan J, et al. Spectral methods for data science: A statistical perspective. Foundations and Trends® in Machine Learning, 2021, 14(5): 566-806.

Global for Certain DGP

Spectral Method + Local Improve Meant

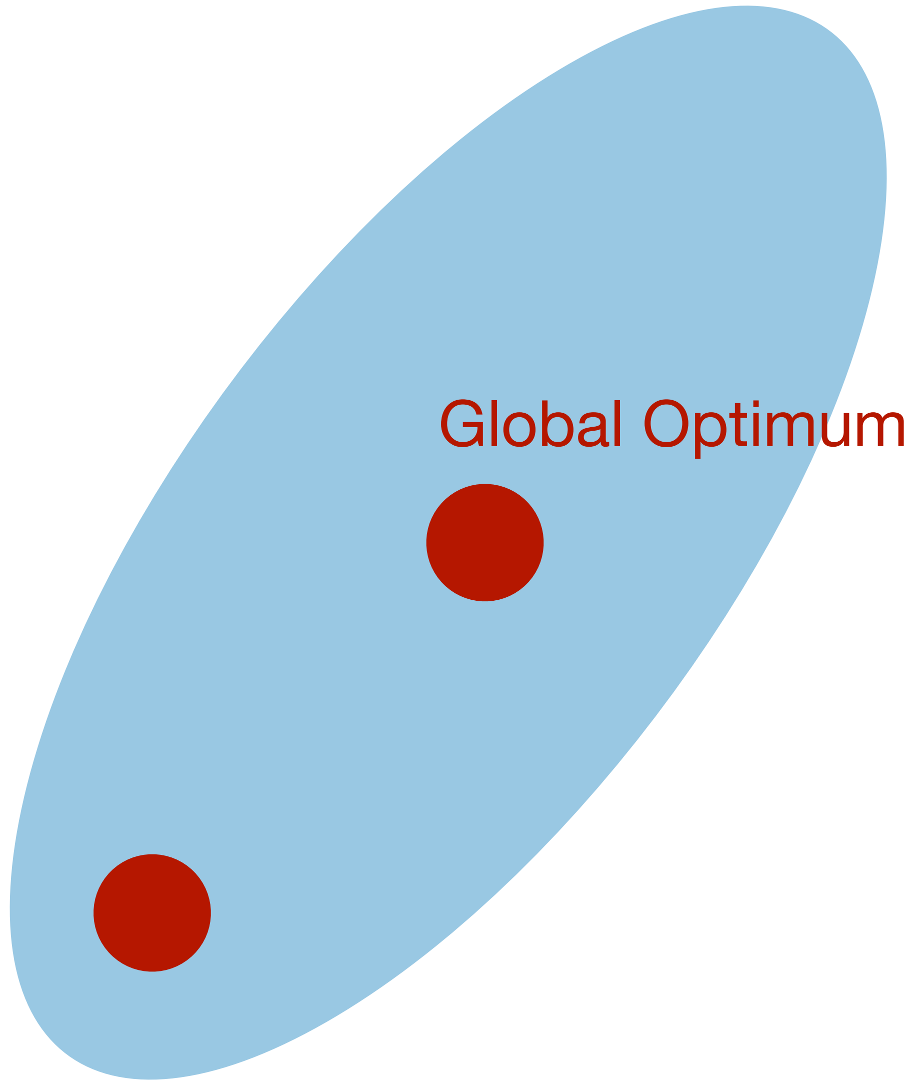
Although NP-hard, it's solvable under certain data generating process (DGP)

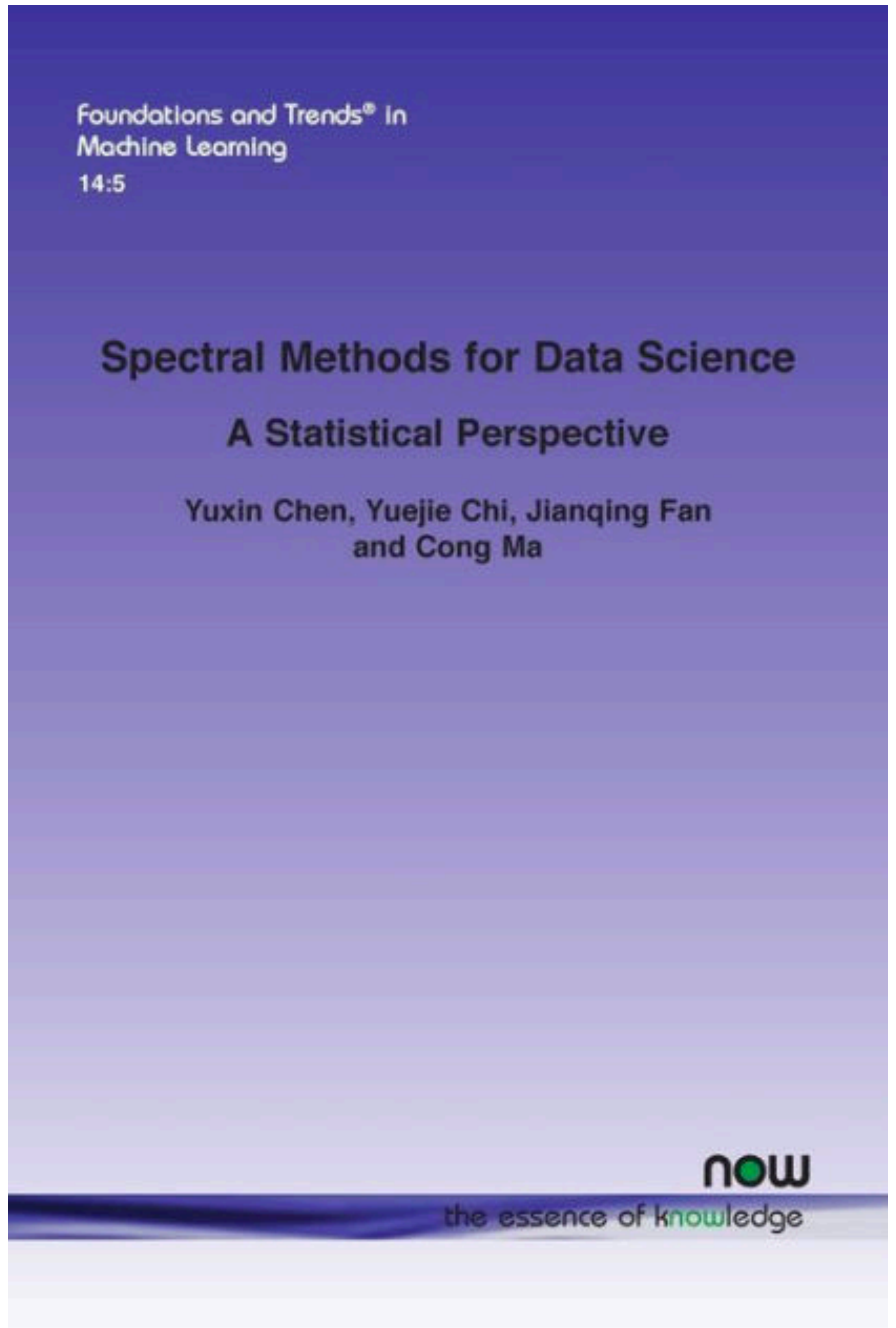
Idea



Signal **Noise**

Spectral initialization!





Foundations and Trends® in
Machine Learning
14:5

Spectral Methods for Data Science

A Statistical Perspective

Yuxin Chen, Yuejie Chi, Jianqing Fan
and Cong Ma

now
the essence of knowledge

Examples: Phase Synchronization/Retrieval, Matrix Completion, Random Block Model

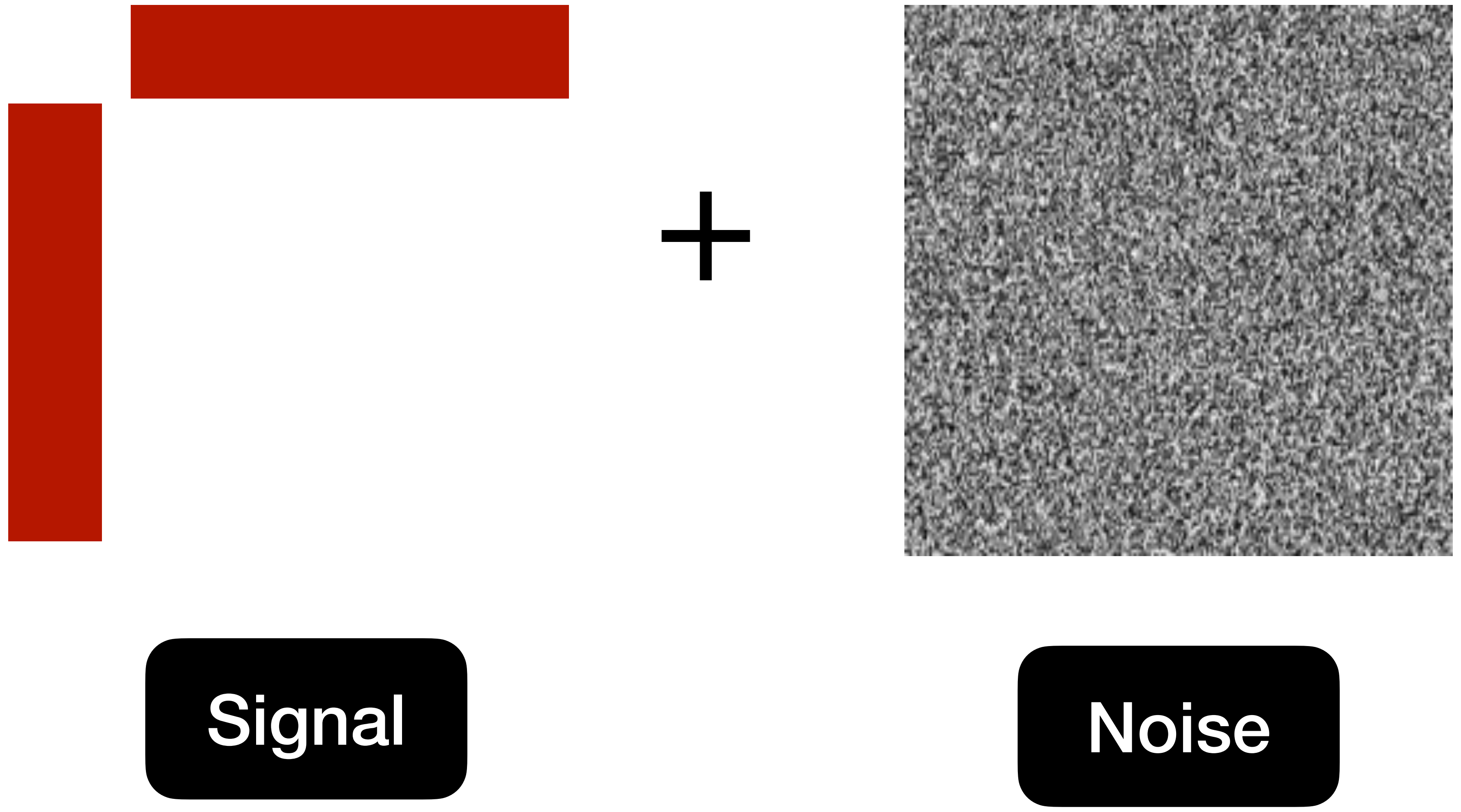
Chen Y, Chi Y, Fan J, et al. Spectral methods for data science: A statistical perspective. Foundations and Trends® in Machine Learning, 2021, 14(5): 566-806.

Global for Certain DGP

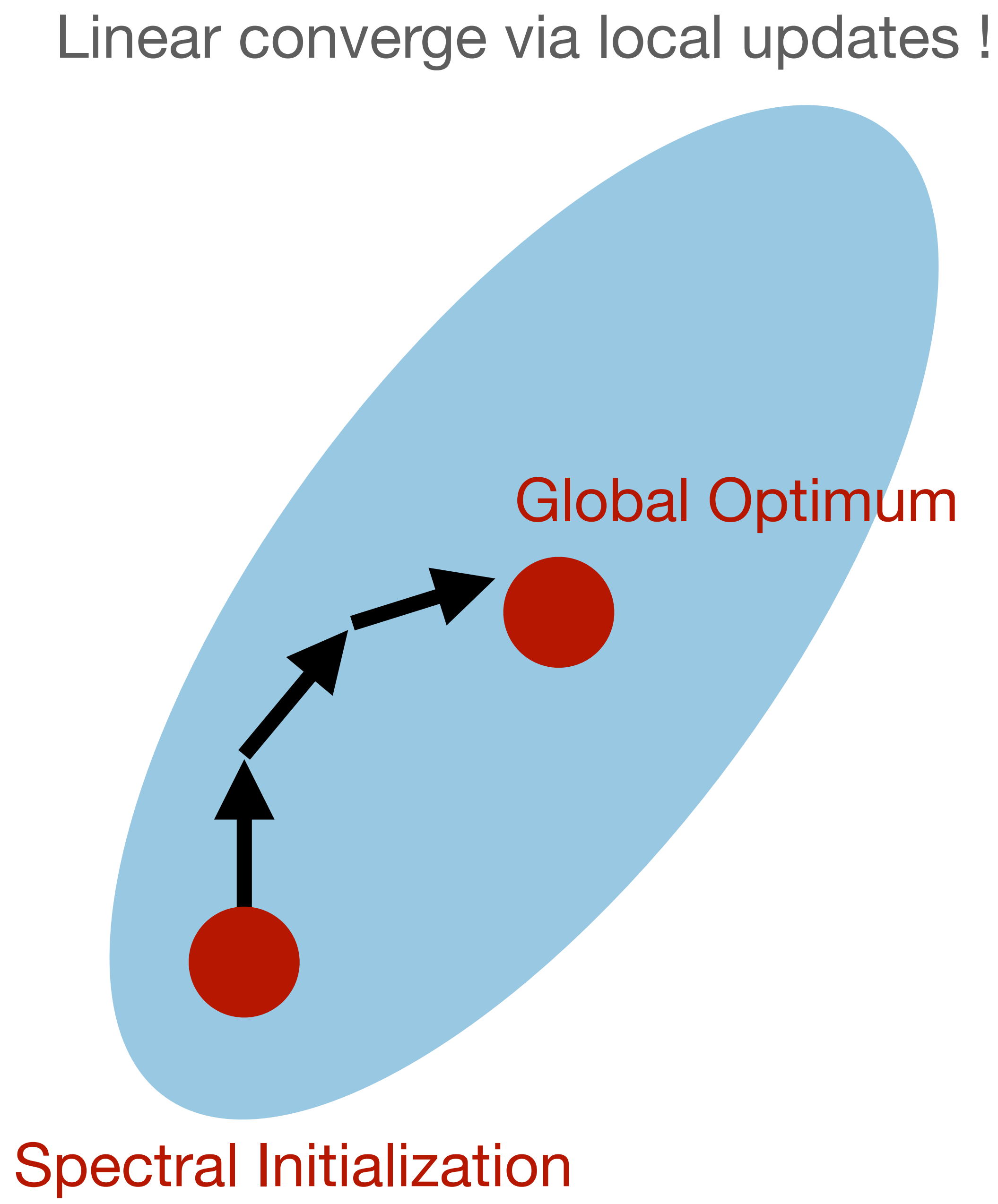
Spectral Method + Local Improve Meant

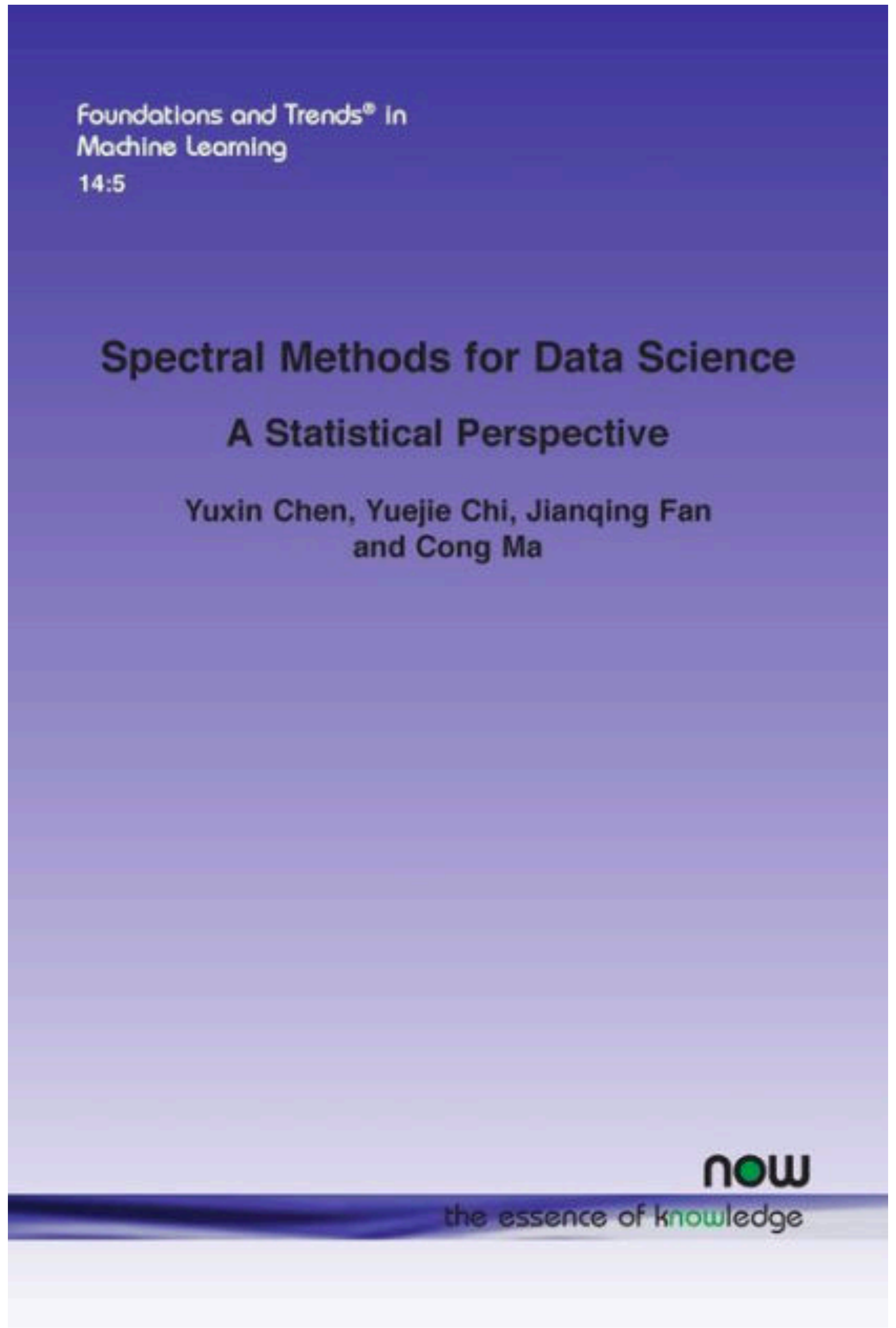
Although NP-hard, it’s solvable under certain data generating process (DGP)

Idea



Spectral initialization!




Foundations and Trends® in
Machine Learning
14:5

Spectral Methods for Data Science

A Statistical Perspective

Yuxin Chen, Yuejie Chi, Jianqing Fan
and Cong Ma

now
the essence of knowledge

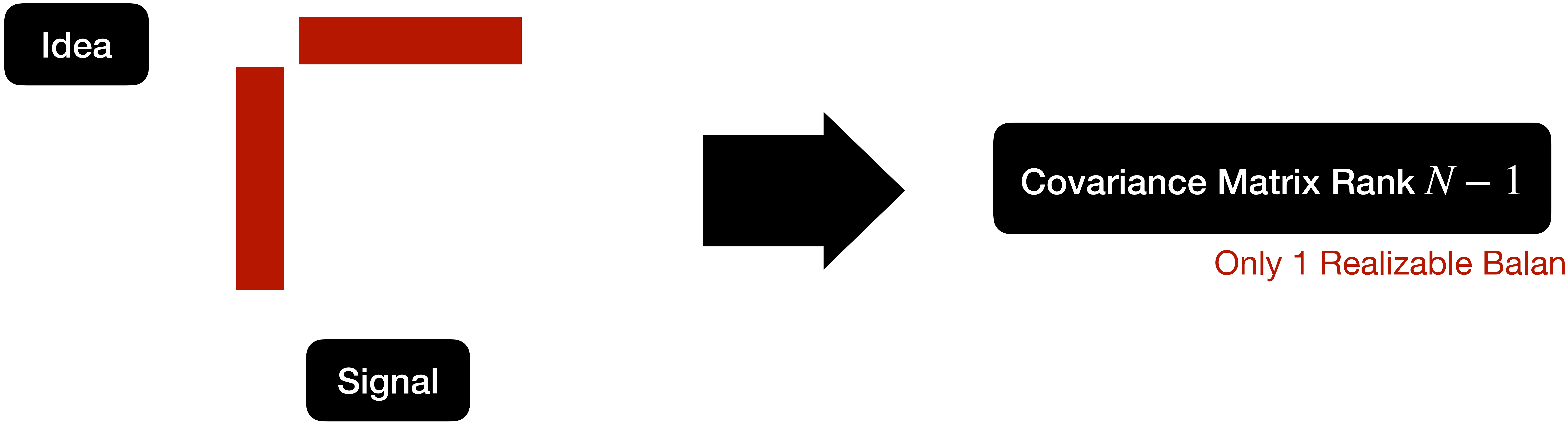
Examples: Phase Synchronization/Retrieval, Matrix Completion, Random Block Model

Chen Y, Chi Y, Fan J, et al. Spectral methods for data science: A statistical perspective. Foundations and Trends® in Machine Learning, 2021, 14(5): 566-806.

Global for Certain DGP

Spectral Method + Local Improve Meant

Although NP-hard, it’s solvable under certain data generating process (DGP)



Idea

Signal

Covariance Matrix Rank $N - 1$

Only 1 Realizable Balance Profile!

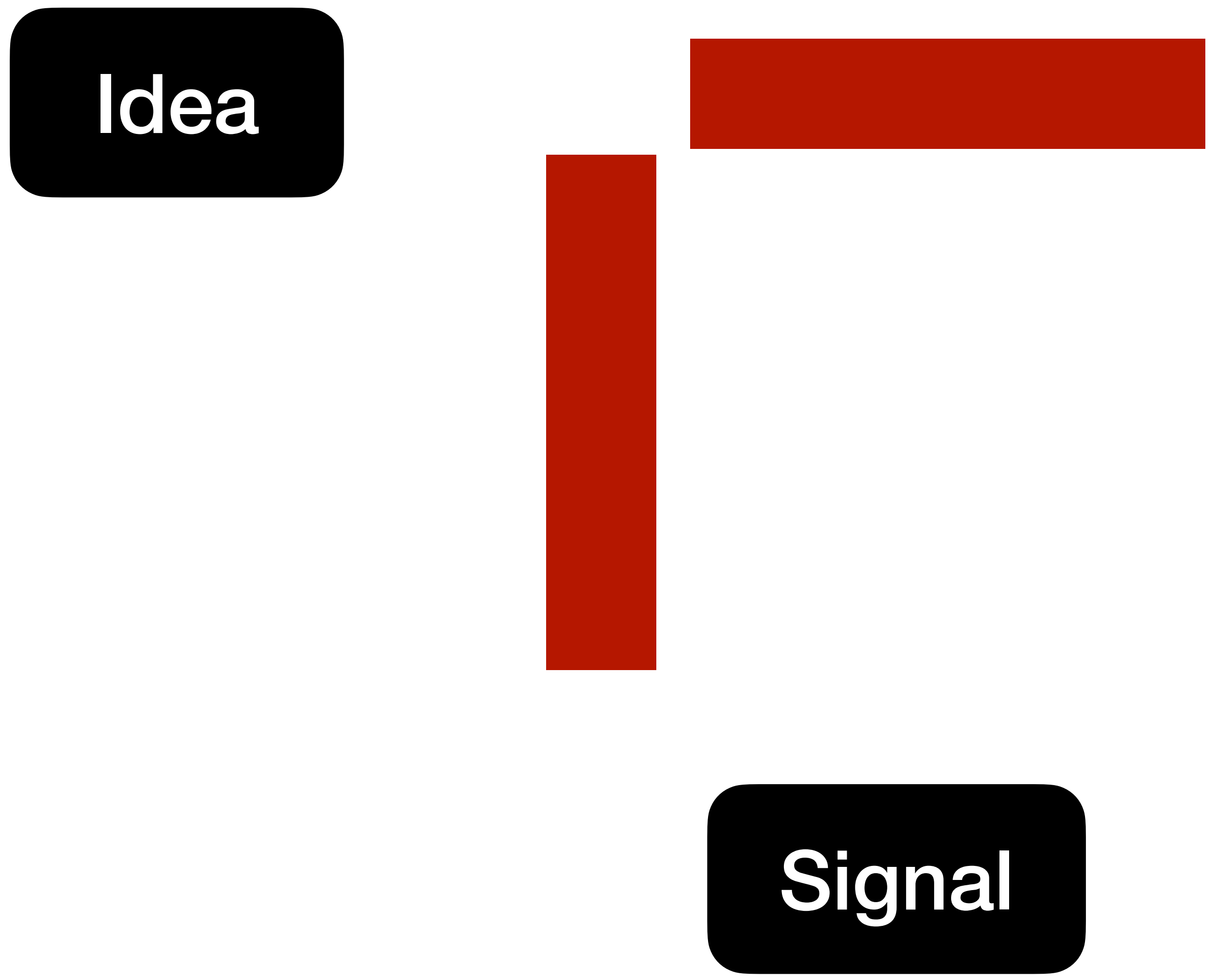
Can be solved via spectral method!

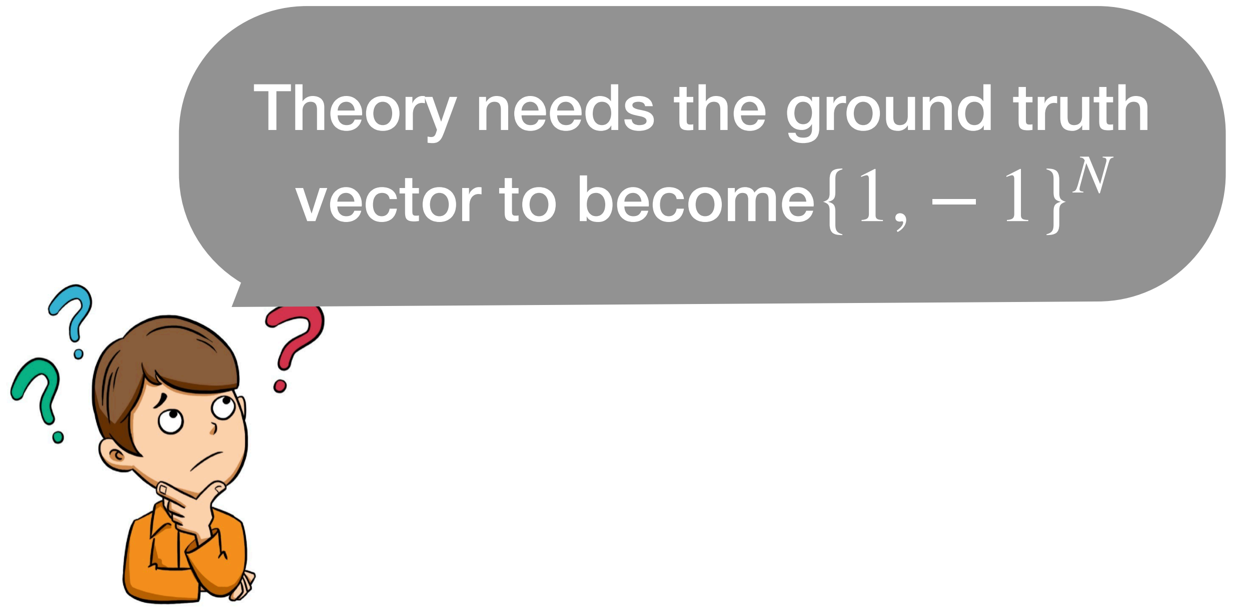
Chen Y, Chi Y, Fan J, et al. Spectral methods for data science: A statistical perspective. Foundations and Trends® in Machine Learning, 2021, 14(5): 566-806.

Global for Certain DGP

Spectral Method + Local Improve Meant

Although NP-hard, it’s solvable under certain data generating process (DGP)





Can be solved via spectral method!

Relates to Degree Corrected Block Model

Jin J. Fast community detection by SCORE. The Annals of Statistics, 2015, 43(1): 57-89.

Global for Certain DGP

Spectral Method + Local Improve Meant

Although NP-hard, it’s solvable under certain data generating process (DGP)

Idea

Signal

Can be solved via spectral method!

Theory needs the ground truth vector to become $\{1,-1\}^N$

Global Result needs $|z_i| > 1 - \frac{\sqrt{3}}{2}$

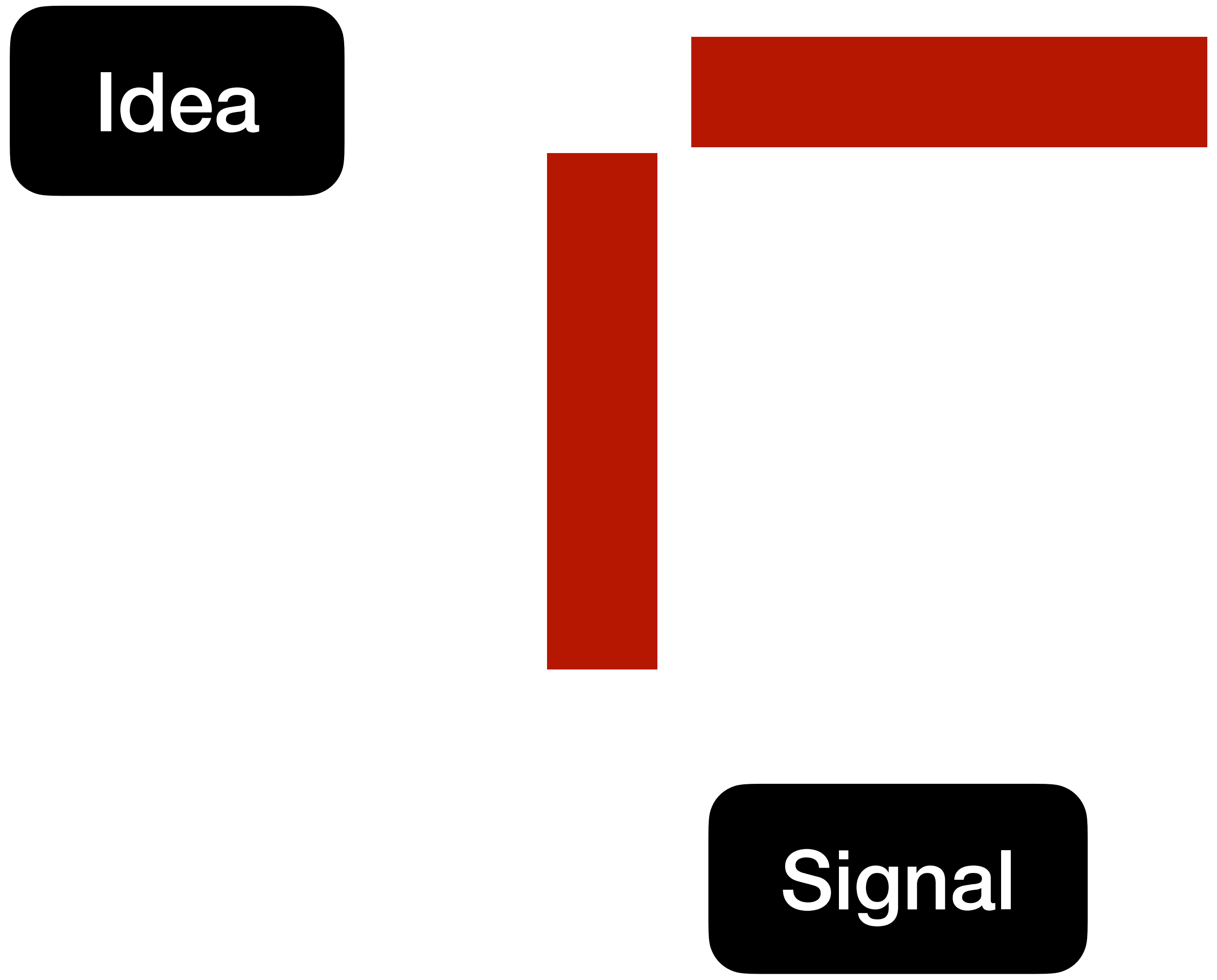
Relates to Degree Corrected Block Model

Jin J. Fast community detection by SCORE. The Annals of Statistics, 2015, 43(1): 57-89.

Global for Certain DGP

Spectral Method + Local Improve Meant

Although NP-hard, it’s solvable under certain data generating process (DGP)



Can be solved via spectral method!

Theory needs the ground truth vector to become $\{1,-1\}^N$

Normalize !

Change the Riemannian Metric

Normalize by the diagonal term of inverse covariance matrix

$$y^k = \text{sgn}((AA^\top + \alpha I)y^{k-1}\,./d)$$

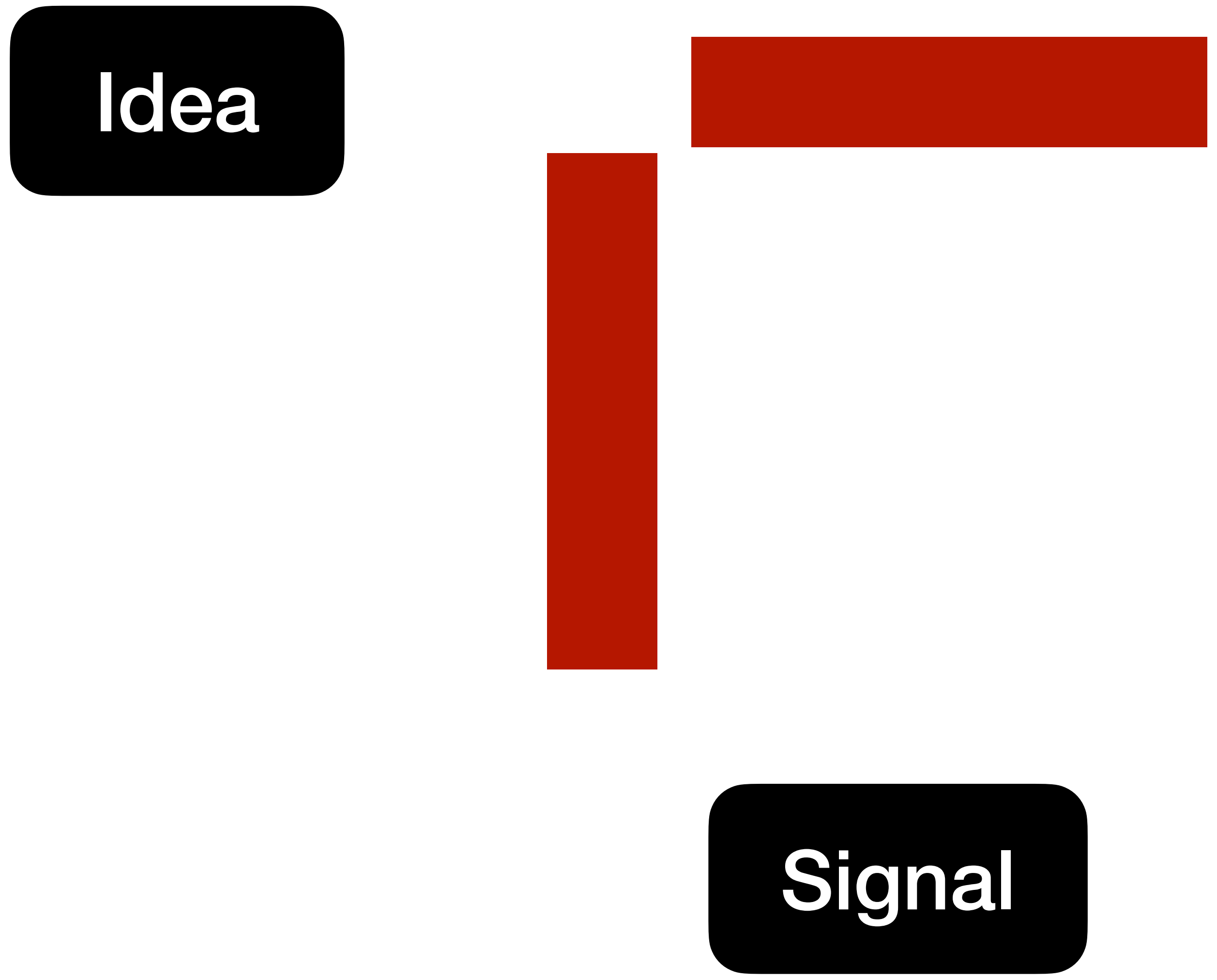
Relates to Degree Corrected Block Model

Jin J. Fast community detection by SCORE. The Annals of Statistics, 2015, 43(1): 57-89.

Global for Certain DGP

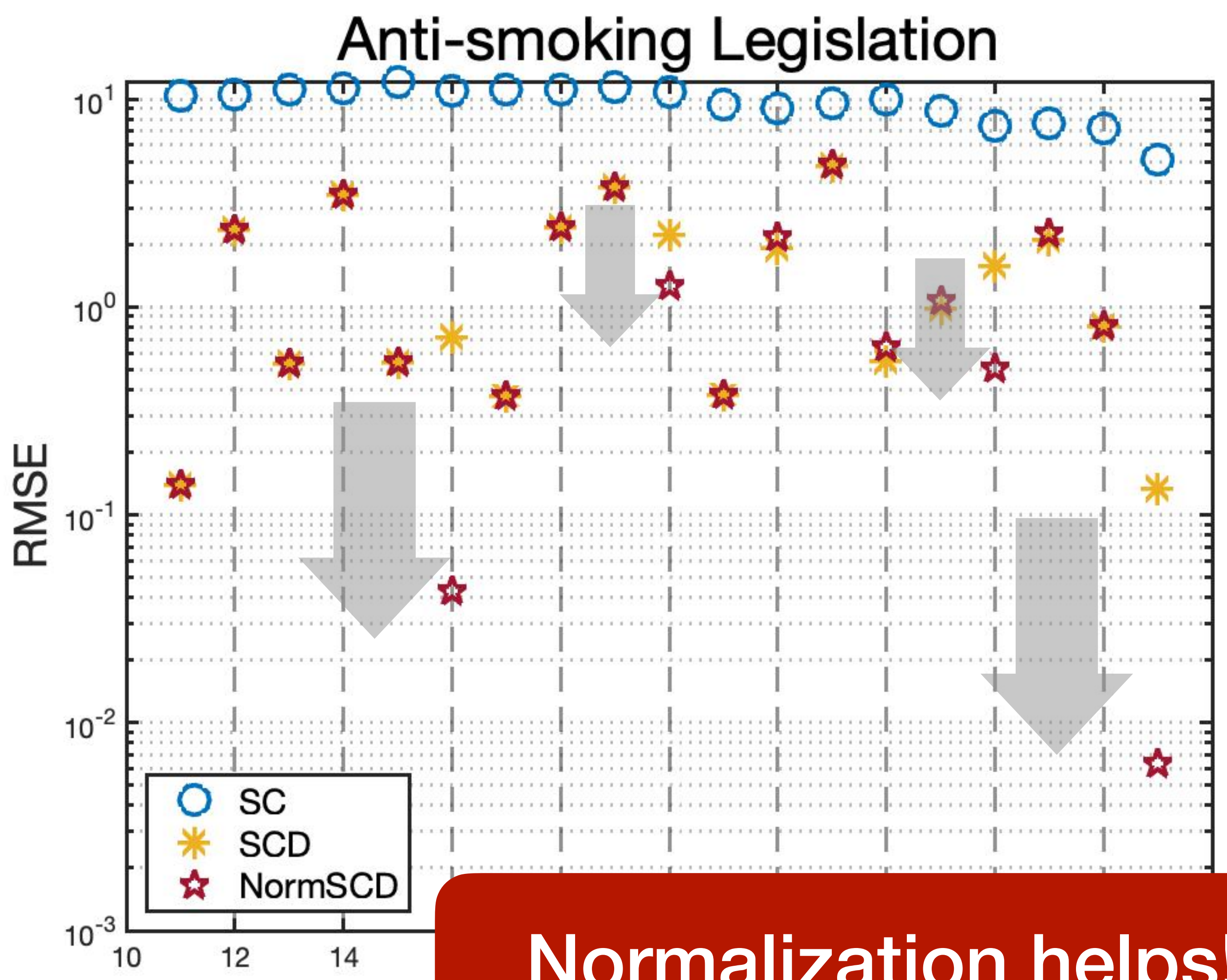
Spectral Method + Local Improve Meant

Although NP-hard, it’s solvable under certain data generating process (DGP)



Theory needs the ground truth vector to become $\{1,-1\}^N$

Normalize !



Normalization helps!

Normalize by the diagonal term of inverse covariance matrix

$$y^k = \mathrm{sgn}((AA^\top + \alpha I)y^{k-1}\boxed{./d})$$

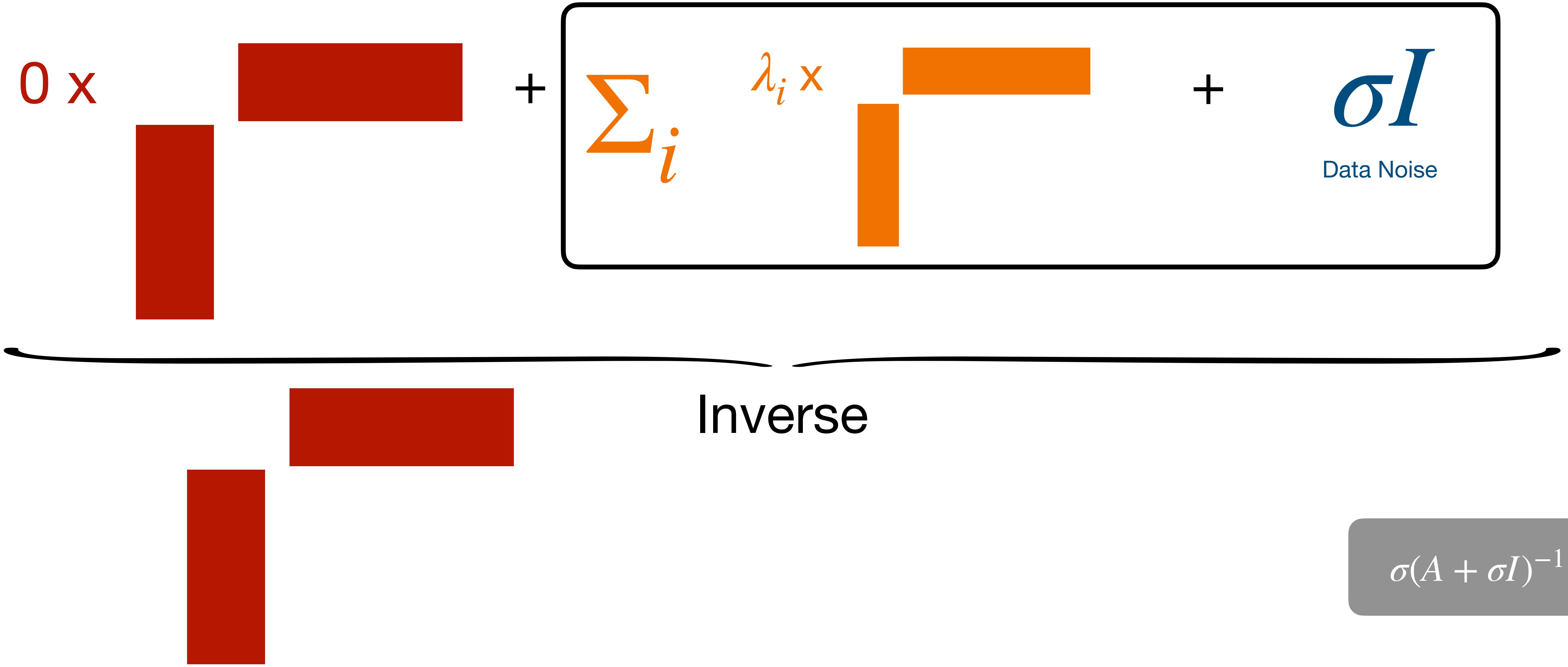
Can be solved via spectral method!

Relates to Degree Corrected Block Model

Jin J. Fast community detection by SCORE. The Annals of Statistics, 2015, 43(1): 57-89.

Closer look at Theory

Difference to Phase synchronization



Perturbation

0 x + $\Sigma_i \ \lambda_i$ x + σI

Data Noise

Inverse

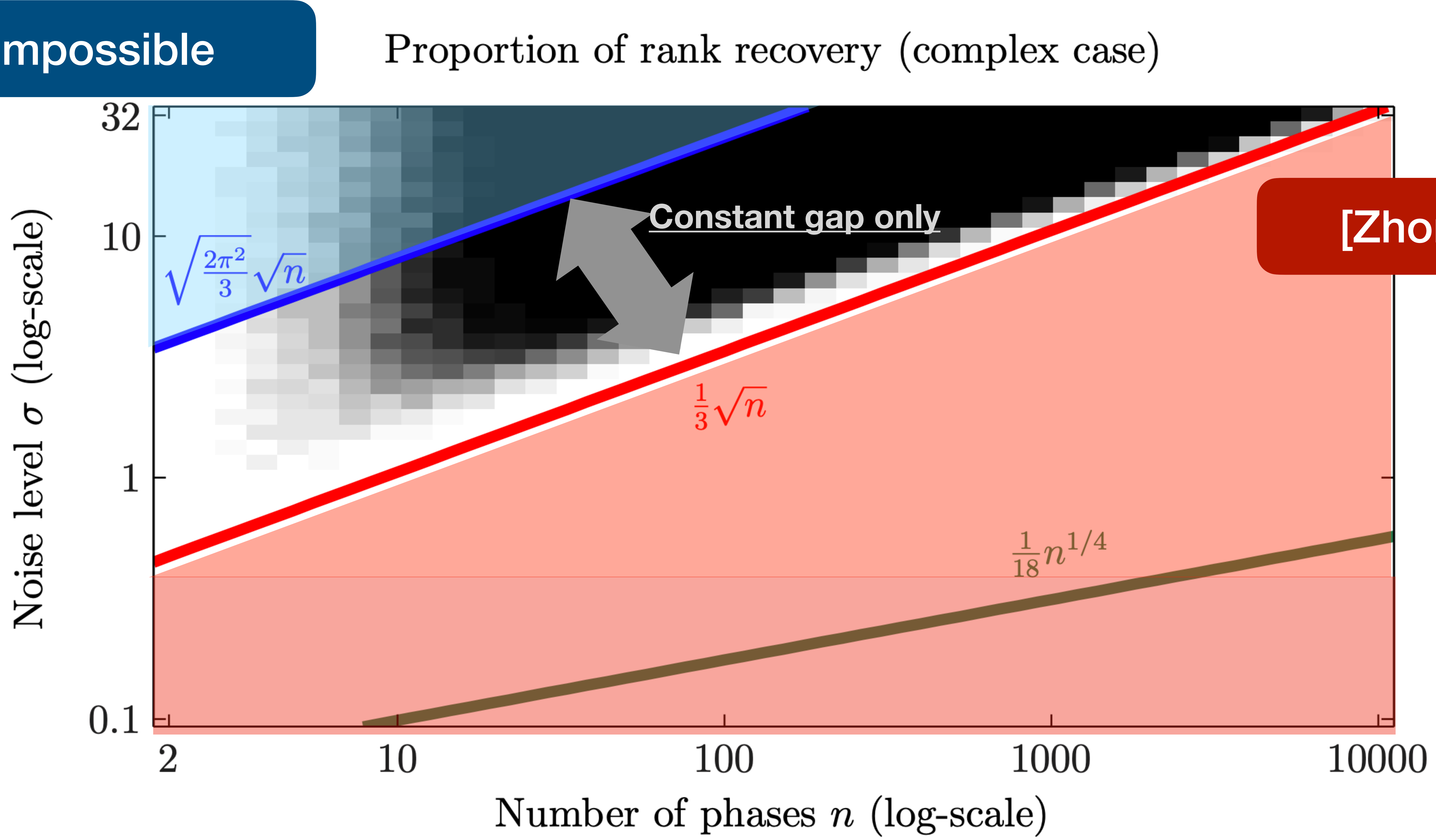
$\sigma(A + \sigma I)^{-1} \rightarrow vv^T$

Drawback of Theory

Connection to Phase synchronization

Stregth of signal = strength of the noise

Impossible



Proportion of rank recovery (complex case)

Constant gap only

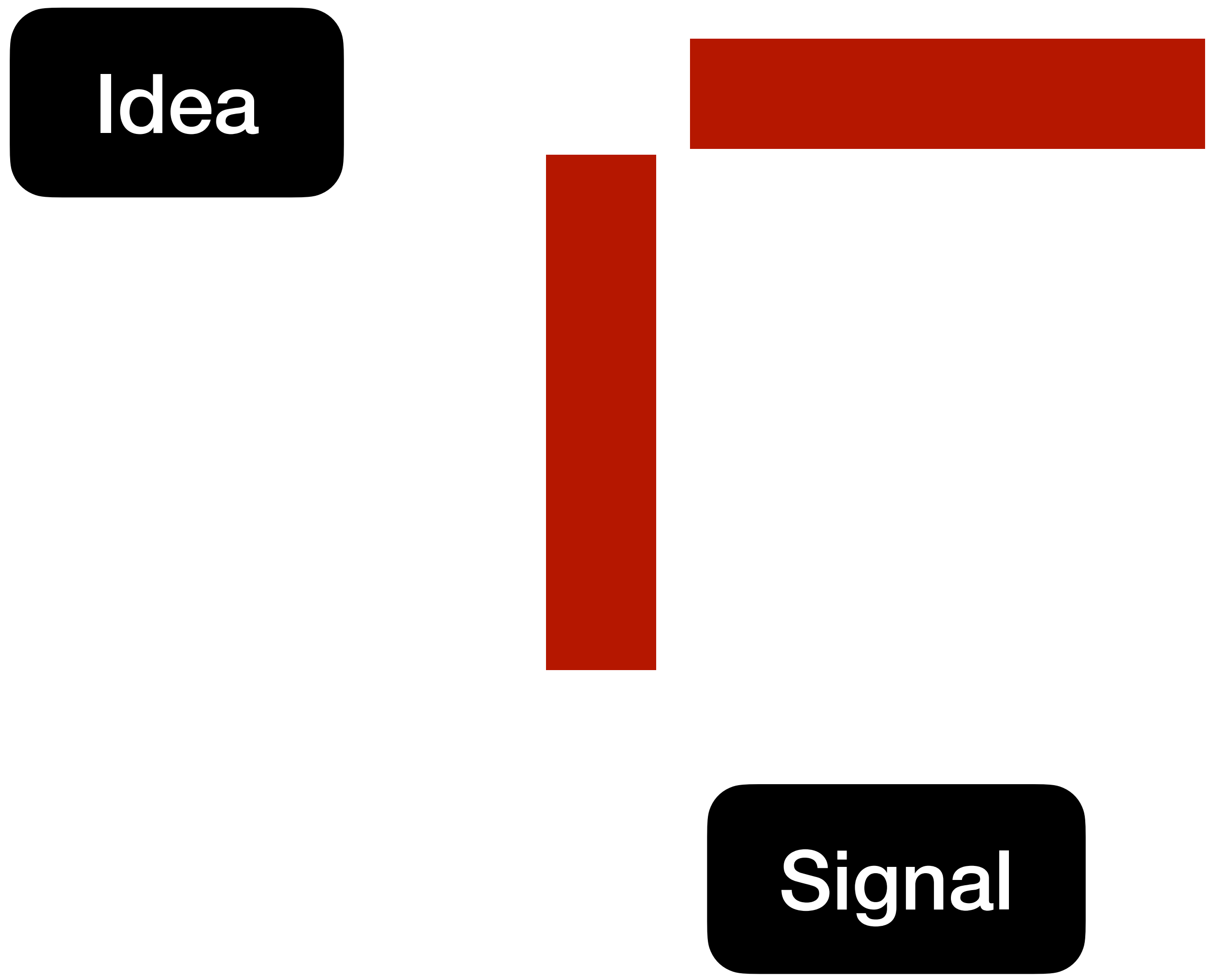
σ samll enough but $O(1)$

[Zhong et al SIOPT]

Global for Certain DGP

Spectral Method + Local Improve Meant

Although NP-hard, it’s solvable under certain data generating process (DGP)



Idea

Signal

Covariance Matrix Rank $N-1$

Only 1 Realizable Balance Profile!

Can be solved via spectral method!

What if I have multiple realizable balance profile?

Select Between Experiments

Chen Y, Chi Y, Fan J, et al. Spectral methods for data science: A statistical perspective. Foundations and Trends® in Machine Learning, 2021, 14(5): 566-806.

The second reformulation

Equal to ℓ_1 PCA

Phase Synchronization

$$\boxed{\max_{\|x\|_2=1} \|Ax\|_1 = \max_{\|x\|_2=1, y\in\{-1,+1\}} y^\top A x} = \max_{y\in\{-1,+1\}} \|A^\top y\|_2$$

ℓ_1-PCA

The second reformulation

Equal to ℓ_1 PCA

Phase Synchronization

$$\boxed{\max_{\|x\|_2=1} \|Ax\|_1 = \max_{\|x\|_2=1, y\in\{-1,+1\}} y^\top Ax} = \max_{y\in\{-1,+1\}} \|A^\top y\|_2$$

ℓ_1-PCA

Still provable NP-hard

Low Rank: N^{rank}

Markopoulos P P, Karystinos G N, Pados D A. Optimal algorithms for $ L_ {1} $-subspace signal processing. IEEE Transactions on Signal Processing, 2014, 62(19): 5046-5058.

The second reformulation

Equal to ℓ_1 PCA

Phase Synchronization

$$\boxed{\max_{\|x\|_2=1} \|Ax\|_1 = \max_{\|x\|_2=1, y\in\{-1,+1\}} y^\top Ax} = \max_{y\in\{-1,+1\}} \|A^\top y\|_2$$

ℓ_1-PCA

Still provable NP-hard

Low Rank: N^{rank}

Markopoulos P P, Karystinos G N, Pados D A. Optimal algorithms for $ L_ {1} $-subspace signal processing. IEEE Transactions on Signal Processing, 2014, 62(19): 5046-5058.

Algorithm

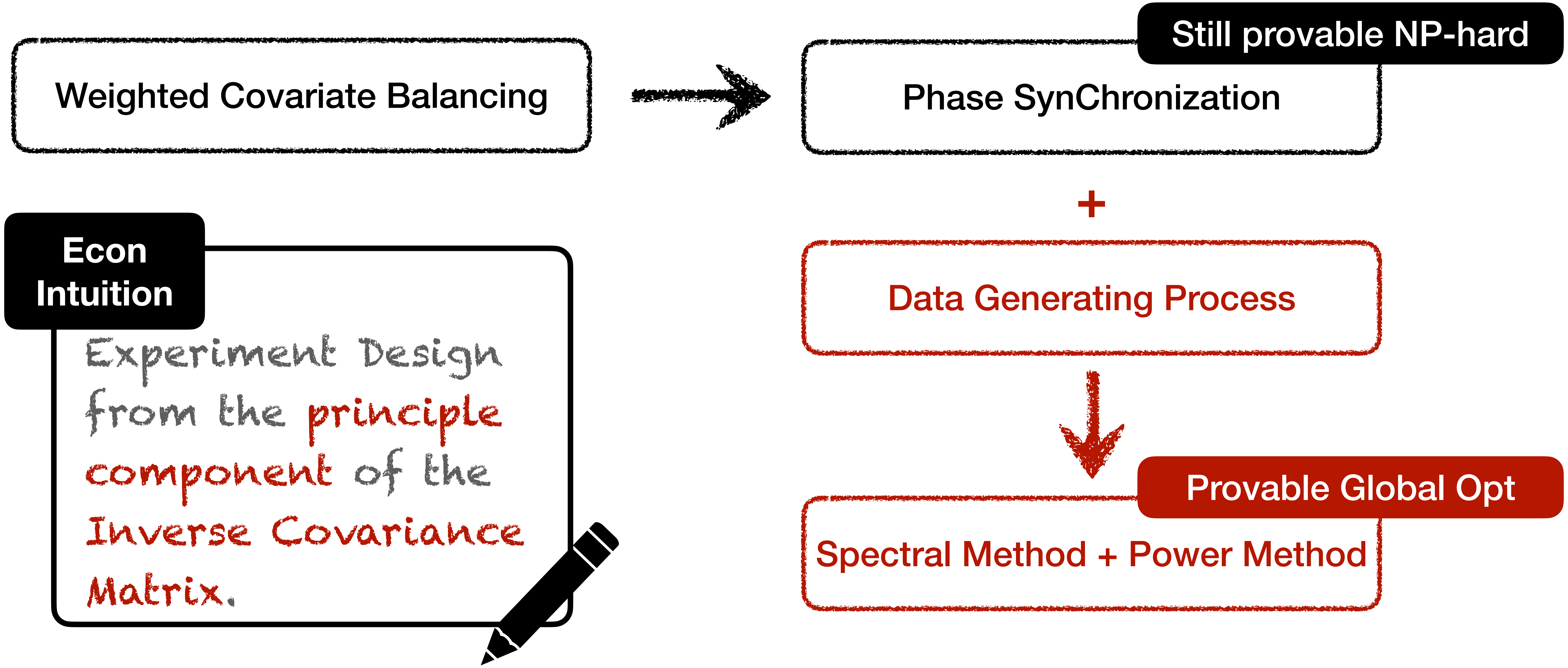
Step 1. Low rank approximate to the inverse covariance matrix.

Step 2. Using Algorithms for ℓ_1-PCA

Step 3. Local Refinement via Power Method $y^k = \text{sgn}((AA^\top + \alpha I)y^{k-1})$.

Take Home Message

Fast Covariate Balancing



Weighted Covariate Balancing → Phase SynChronization

Still provable NP-hard

\+

Data Generating Process

↓

Spectral Method + Power Method

Provable Global Opt

Econ Intuition

Experiment Design from the principle component of the Inverse Covariance Matrix.

Take Home Message

Fast Covariate Balancing

Weighted Covariate Balancing → Phase SynChronization

+

Data Generating Process

↓

Spectral Method + Power Method

Econ Intuition

Experiment Design from the principle component of the Inverse Covariance Matrix.

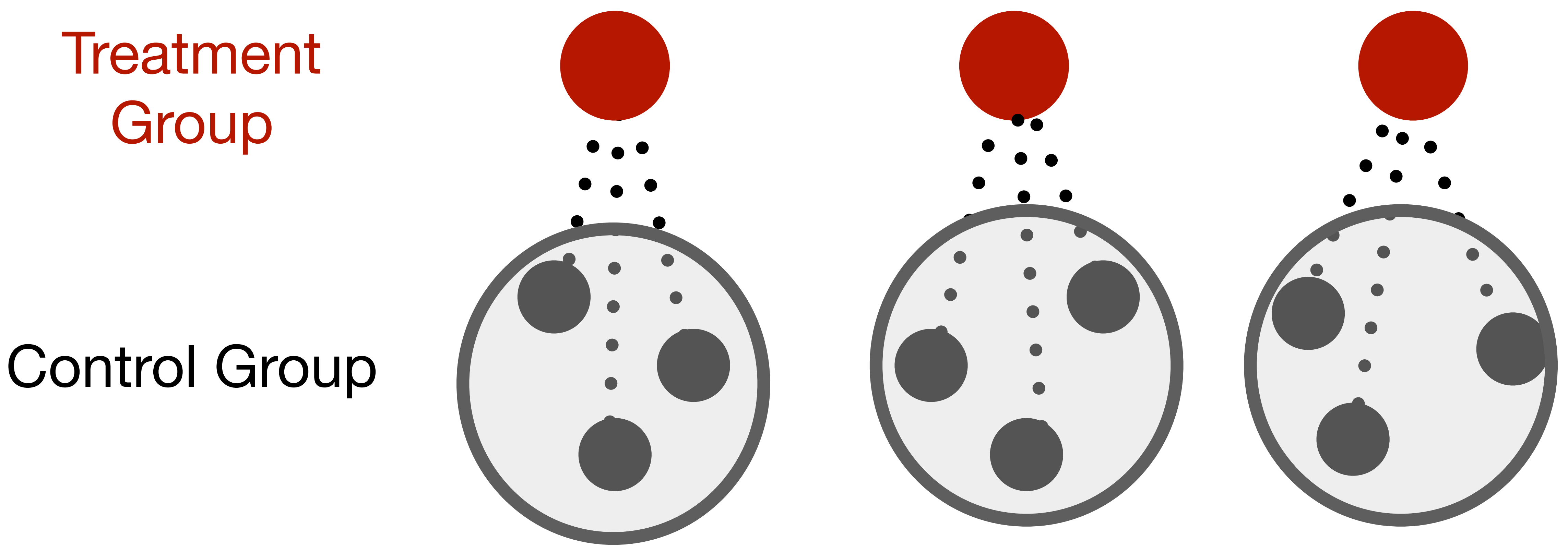
collect data

SVD+Power Method

Experiment

Still open questions

Happy to talk



Separate the data into two groups to minimize the <u>optimal transport distance</u> between a weighted version to the two group

$$\min_{\{D_i,w_i\}_{i=1}^N} \quad \frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=1}^{N} w_i D_i Y_{it} - \sum_{i=1}^{N} w_i(1-D_i)Y_{it}\right)^2 + \lambda\sum_{i=1}^{N} w_i^2$$

$$\text{s.t.} \quad w_i \geq 0, \quad D_i \in \{0,1\} \text{ for } i = 1,\ldots,N,$$

$$\boxed{\sum_{i=1}^{N} D_i = K,} \quad \sum_{i=1}^{N} w_i D_i = 1, \quad \sum_{i=1}^{N} w_i(1-D_i) = 1$$

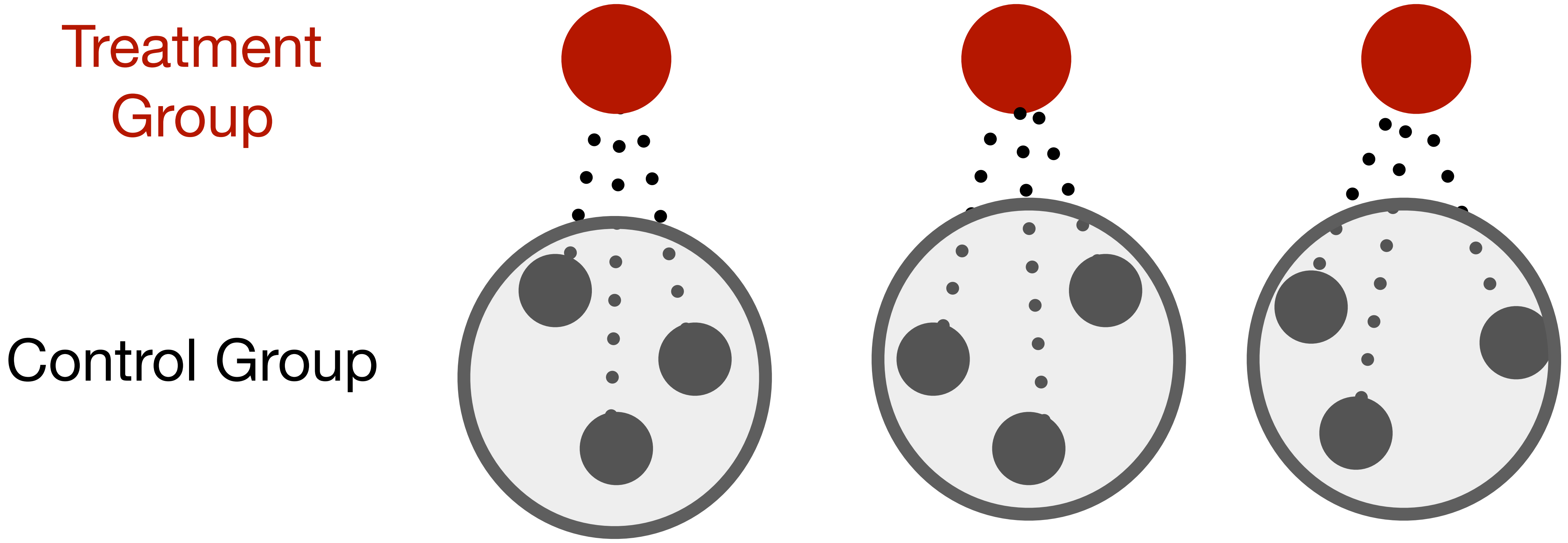
Constraint the cost of experiment

negatively correlated PCA

negatively correlated sparse PCA

Still open questions

Happy to talk



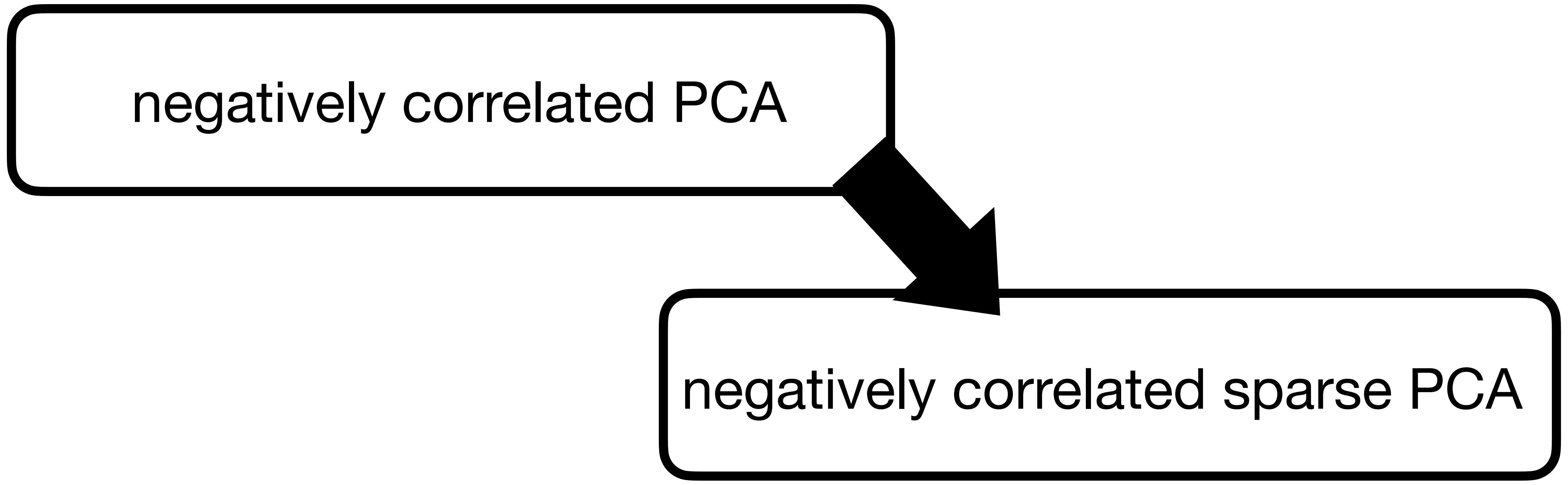
Separate the data into two groups to minimize the optimal transport distance between a weighted version to the two group

$$\min_{\{D_i,w_i\}_{i=1}^N} \frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=1}^{N} w_iD_iY_{it} - \sum_{i=1}^{N} w_i(1-D_i)Y_{it}\right)^2 + \lambda\sum_{i=1}^{N} w_i^2$$

$$\text{s.t.}\quad w_i \geq 0,\quad D_i \in \{0,1\}\ \text{for}\ i = 1,\ldots,N,$$

$$\boxed{\sum_{i=1}^{N} D_i = K,}\quad \sum_{i=1}^{N} w_iD_i = 1,\quad \sum_{i=1}^{N} w_i(1-D_i) = 1$$

Constraint the cost of experiment

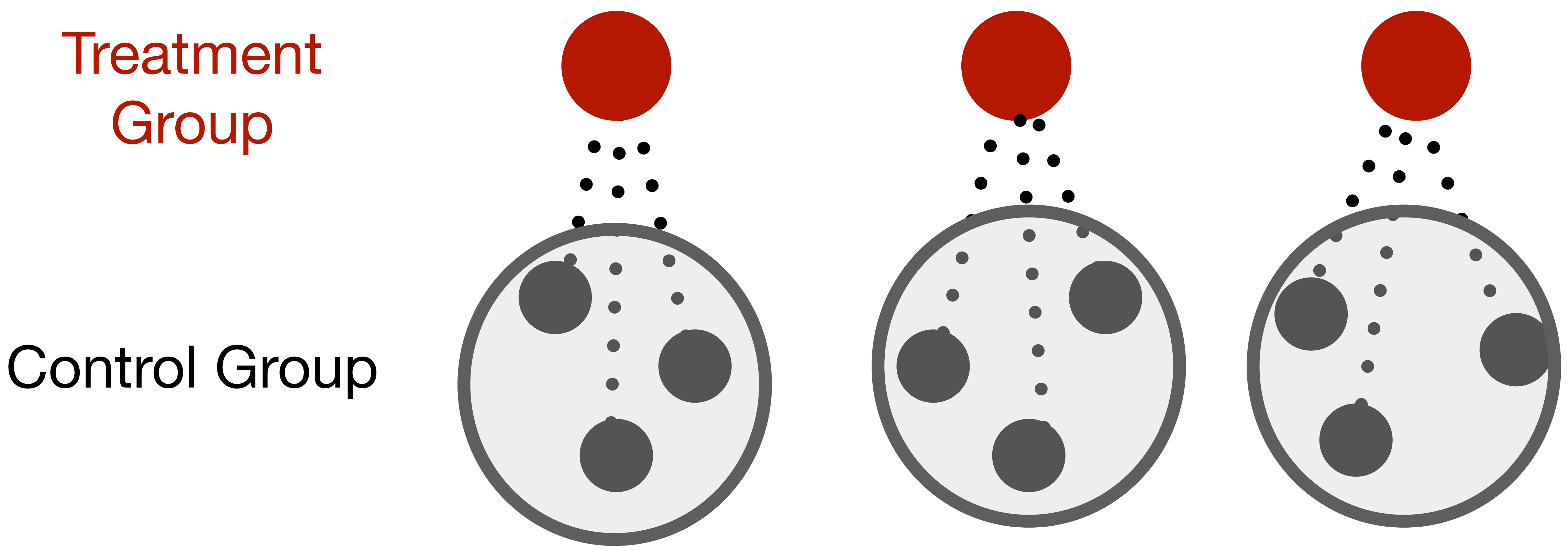


negatively correlated PCA

negatively correlated sparse PCA

Still open questions

Happy to talk



Separate the data into two groups to minimize the optimal transport distance between a weighted version to the two group

$$\min_{\{D_i,w_i\}_{i=1}^N} \quad \frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=1}^{N} w_i D_i Y_{it} - \sum_{i=1}^{N} w_i(1-D_i)Y_{it}\right)^2 + \lambda\sum_{i=1}^{N} w_i^2$$

$$\text{s.t.} \quad w_i \geq 0, \quad D_i \in \{0,1\} \text{ for } i = 1,\ldots,N,$$

$$\boxed{\sum_{i=1}^{N} D_i = K,} \quad \sum_{i=1}^{N} w_i D_i = 1, \quad \sum_{i=1}^{N} w_i(1-D_i) = 1$$

$$\min_{\substack{w_1,\ldots,w_J,\\ v_1,\ldots,v_J}} \left\|\boxed{\overline{\boldsymbol{X}}} - \sum_{j=1}^{J} w_j \boldsymbol{X}_j\right\|^2 + \left\|\boxed{\overline{\boldsymbol{X}}} - \sum_{j=1}^{J} v_j \boldsymbol{X}_j\right\|^2 \qquad \overline{\boldsymbol{X}} = \sum_{j=1}^{J} f_j \boldsymbol{X}_j.$$

Add a prior to the market

$$\text{s.t.} \quad \sum_{j=1}^{J} w_j = 1, \qquad w_j, v_j \geq 0, \quad j = 1,\ldots,J,$$

$$w_j v_j = 0, \quad j = 1,\ldots,J,$$

$$\sum_{j=1}^{J} v_j = 1, \qquad \underline{m} \leq \|\boldsymbol{w}\|_0 \leq \overline{m}.$$

Stanford
University

ICME

Thank You and Questions?

Contact: yplu@stanford.edu