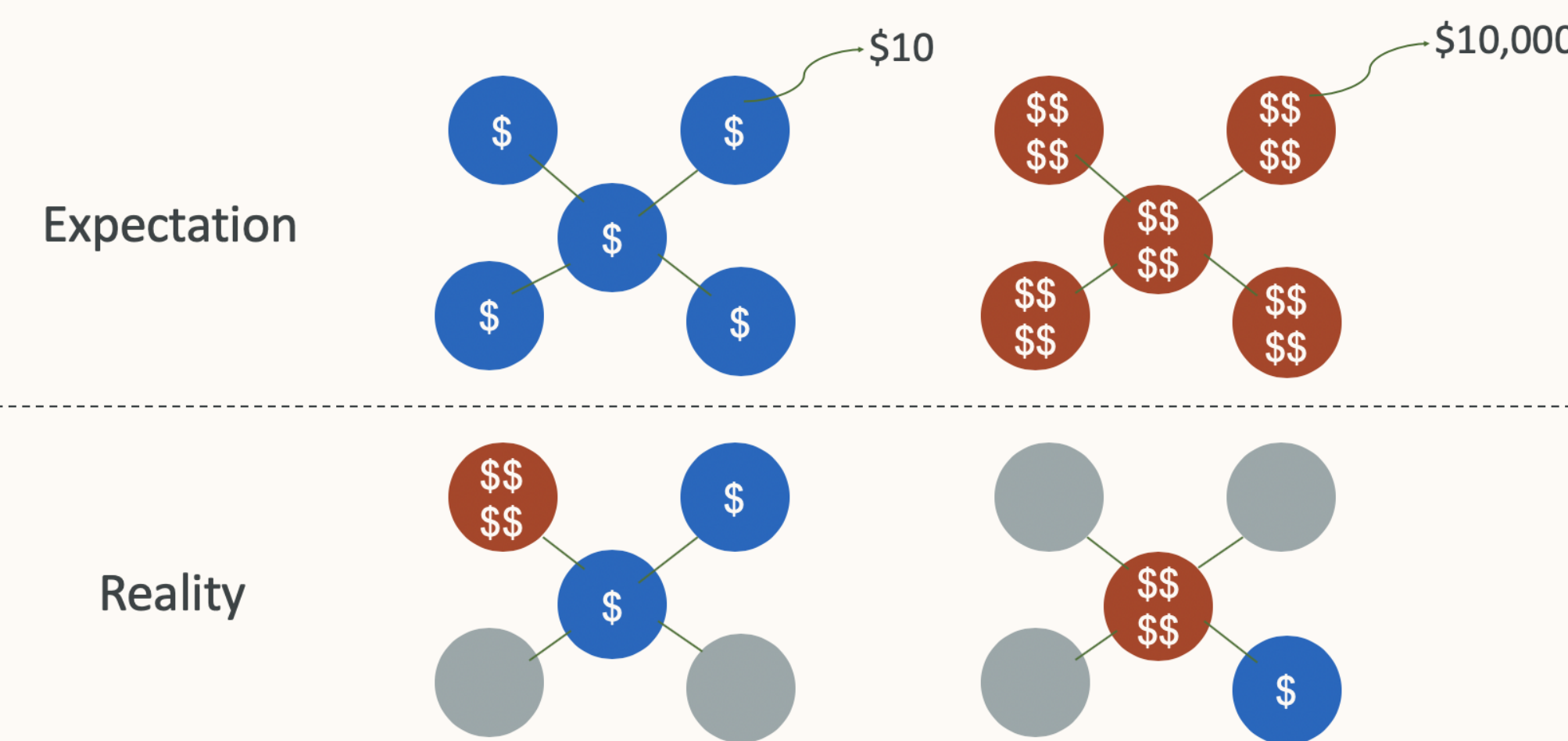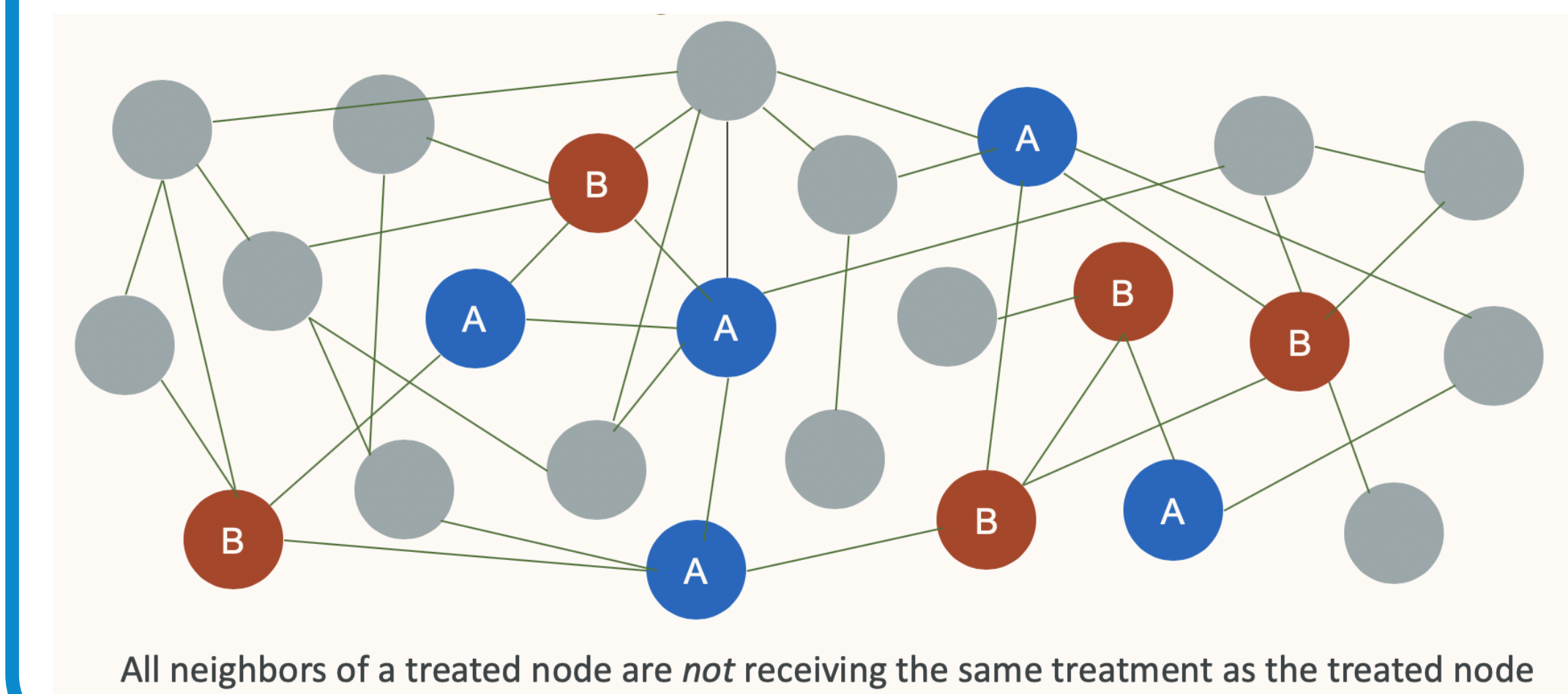# A/B Testing in Dense Large-Scale Networks: Design and Inference
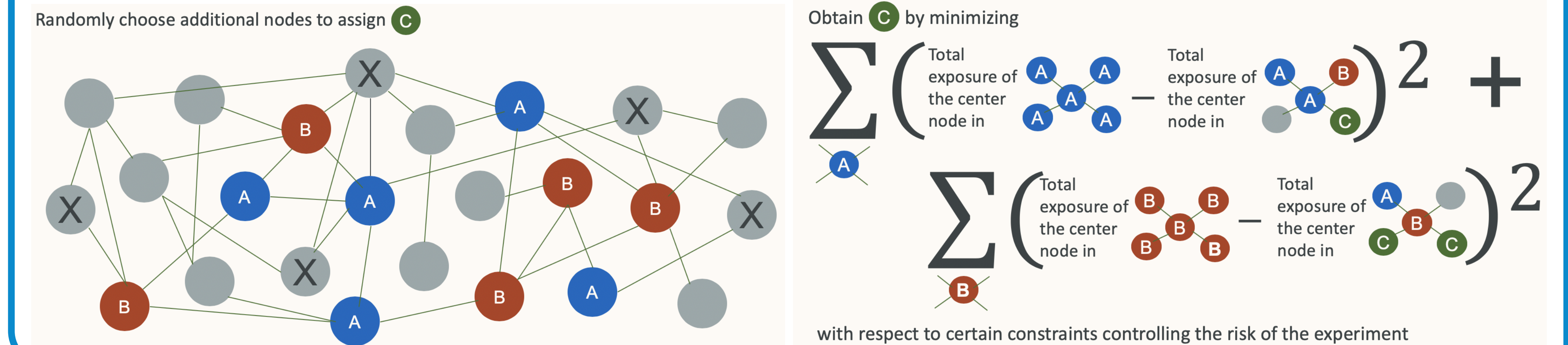
Preetam Nandy, Kinjal Basu, Shaunak Chatterjee, Ye Tu
LinkedIn Corporation

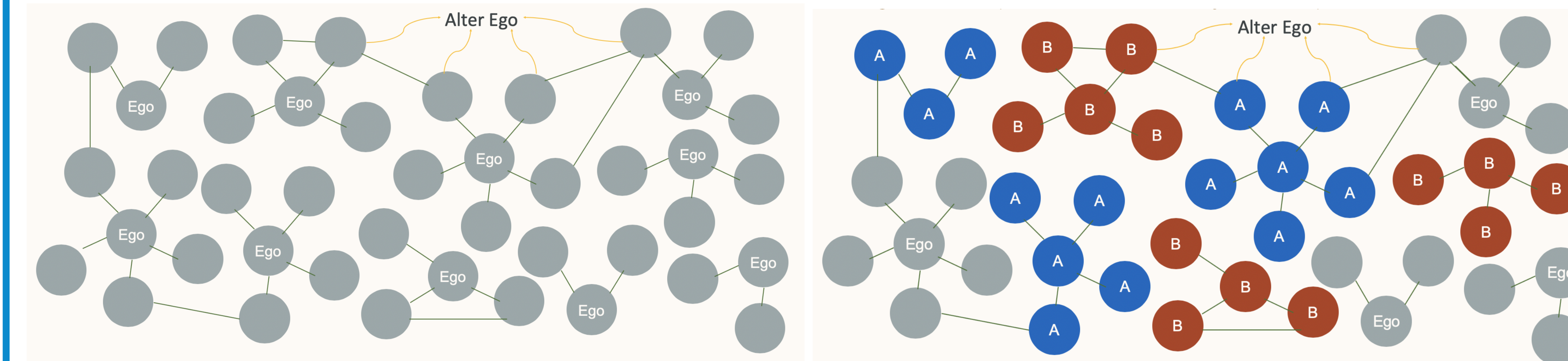## Network Effect in A/B Testing

Expectation

Reality

$10 $10,000
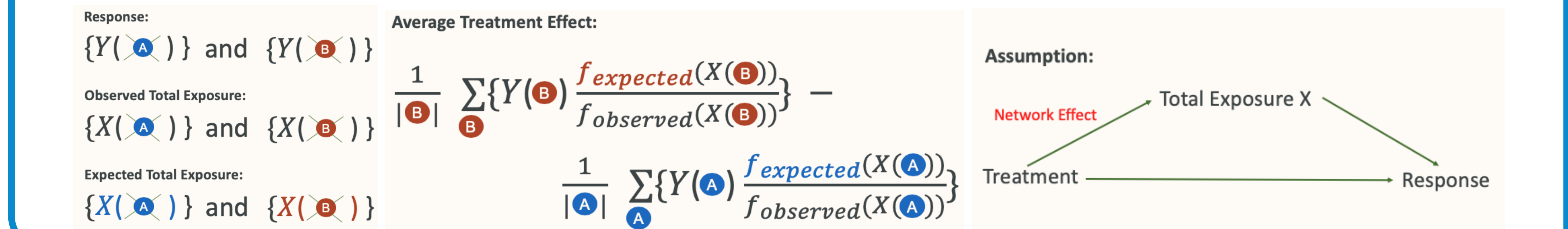
All neighbors of a treated node are *not* receiving the same treatment as the treated node

Ego-cluster Experiment

Alter Ego

+ Works under minimal assumptions.

− Does not work well for dense networks.

Network Effect in Exposure Redistribution Experiments

Like/Share/Comment

Content of Member 1
Content of Member 2
Content of Member 3
Content of Member N

Representing LinkedIn Feed Model by an Exposure Distribution Matrix

$$P = \begin{pmatrix} p_{11} & \cdots & p_{1N} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix}$$ such that

Column Sum = 1
Row Sum = Total Exposure

Feed of Member j

Exposure Distribution A

0	0.30	0.20	0	0.60	0	= 1.10
0	*	*	*	*	0	
0.50	*	*	. *	*	0.20	
0.25	*	*	*	*	0.60	
0	*	*	*	*	0.20	
0.25	0.10	0	0.10	0.10	0	

Exposure Distribution B

0	0.40	0.10	0	0.25	0	
0	*	*	*	*	0.30	
0.20	*	*	*	*	0.40	
0.45	*	*	*	*	0	
0	*	*	*	*	0.30	
0.35	0	0	0.25	0.25	0	= 0.85

A/B Testing Design

0	0.20	0.10	0	0.60	0	= 0.90 ≠ 1.10
0	*	*	*	*	0.30	
0.50	*	*	*	*	0.40	
0.25	*	*	*	*	0	
0	*	*	*	*	0.30	
0.25	0.20	0	0.40	0.10	0	= 0.95 ≠ 0.85

The total exposure mismatch can be cured by appropriately modifying the columns corresponding to some unassigned nodes

0	0.20	0.10	0.20	0.60	0	= 1.10
0	*	*	*	*	0.30	
0.50	*	*	*	*	0.40	
0.25	*	*	*	*	0	
0	*	*	*	*	0.30	
0.25	0.20	0	0.30	0.10	0	= 0.85

Network Effect

Assumption

Treatment

Row Values in the Exposure Distribution Matrix → Row Sum in the Exposure Distribution Matrix

Column Values in the Exposure Distribution Matrix

Response

Optimal Allocation Strategy (OAS)

Randomly choose additional nodes to assign C

Obtain C by minimizing

$$\sum_A \left(\text{Total exposure of the center node in} \quad - \quad \text{Total exposure of the center node in} \right)^2 +$$

$$\sum_B \left(\text{Total exposure of the center node in} \quad - \quad \text{Total exposure of the center node in} \right)^2$$

with respect to certain constraints controlling the risk of the experiment

Importance Sampling (IS)

Response:
$$\{Y(\text{A})\} \text{ and } \{Y(\text{B})\}$$

Observed Total Exposure:
$$\{X(\text{A})\} \text{ and } \{X(\text{B})\}$$

Expected Total Exposure:
$$\{X(\text{A})\} \text{ and } \{X(\text{B})\}$$

Average Treatment Effect:

$$\frac{1}{|\text{B}|} \sum_B \left\{ Y(\text{B}) \frac{f_{expected}(X(\text{B}))}{f_{observed}(X(\text{B}))} \right\} - \frac{1}{|\text{A}|} \sum_A \left\{ Y(\text{A}) \frac{f_{expected}(X(\text{A}))}{f_{observed}(X(\text{A}))} \right\}$$

Assumption:

Network Effect   Total Exposure X

Treatment   Response

OASIS Experiment

1. Randomly assign A and B

2. Randomly choose additional nodes X

3. Solve a constrained optimization to assign C to X

4. Run experiment and collect data

5. Importance sampling correction

Validation:

- Theoretical results for robustness and verified in simulation

- No statistically significant result in A/A test

- Uniform p-values in A/A test

- Comparing with ego-cluster results (where an ego-cluster experiment is possible)

- Other special experimental designs for online validation

Implementation:

- Implemented for LinkedIn Feed experiments and using it for experiments targeted toward creator experience enhancement

- C = A * boost factors (normalized to have each column sum equals 1)

- Solve optimization to get boost factors, where we control risk by setting a lower and an upper bound for boost factors

- Update boost factors regularly to handle dynamic network/treatment

Pros:

- Theoretically sound under certain assumptions
- Works well for dense networks
- Can handle multiple treatments simultaneously
- Can handle dynamic networks and dynamic treatments
- Can control the risk of the experiment explicitly by adding constraints in the optimization

Cons:

- Relies on a number of assumptions
- Works only for a certain type of experiments
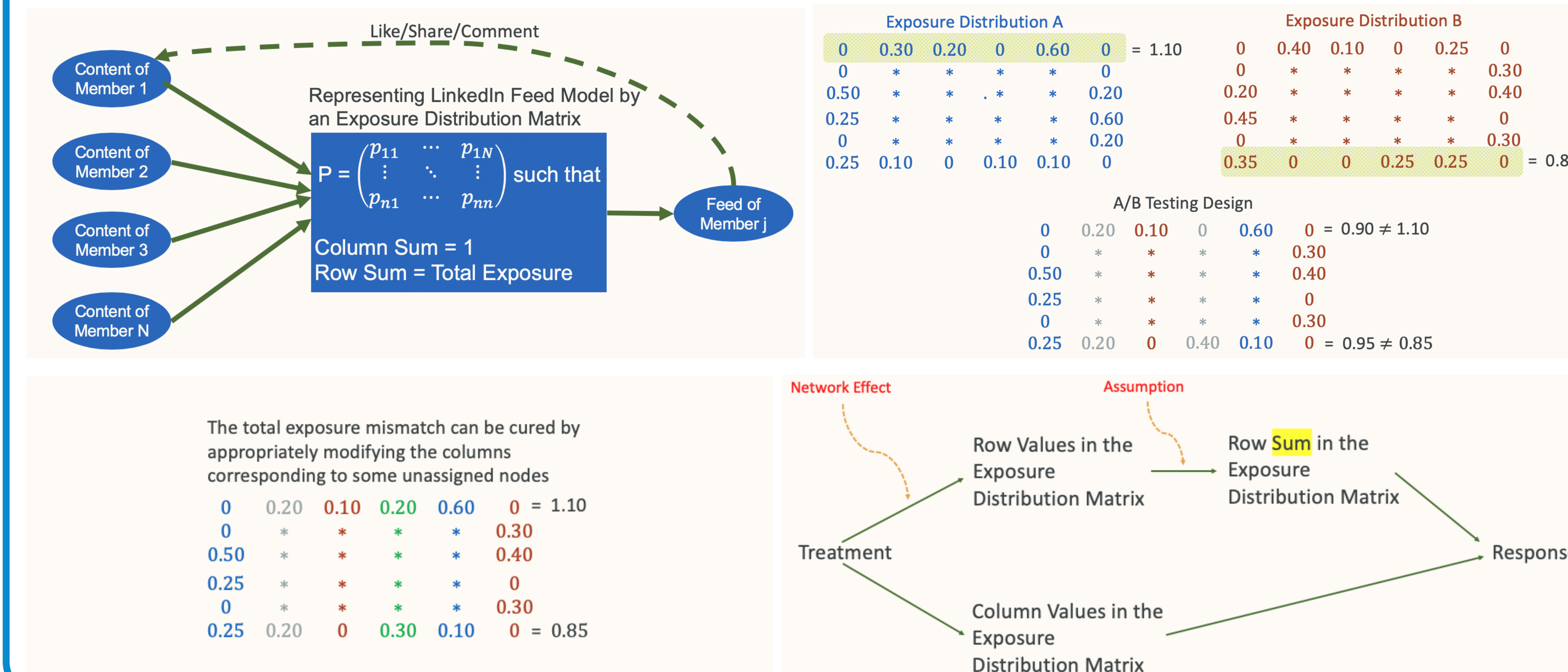
References

- P. Nandy, K. Basu, S. Chatterjee, Y. Tu. A/B Testing in Dense Large-Scale Networks: Design and Inference. arXiv preprint arXiv:1901.10505, 2019.

- G. Saint-Jacques, M. Varshney, J. Simpson, and Y. Xu. Using ego-clusters to measure network effects at LinkedIn. arXiv preprint arXiv:1903.08755, 2019.