LinkedIn

A/B Testing in Dense Large-Scale Networks: Design and Inference

NeurIPS 2020

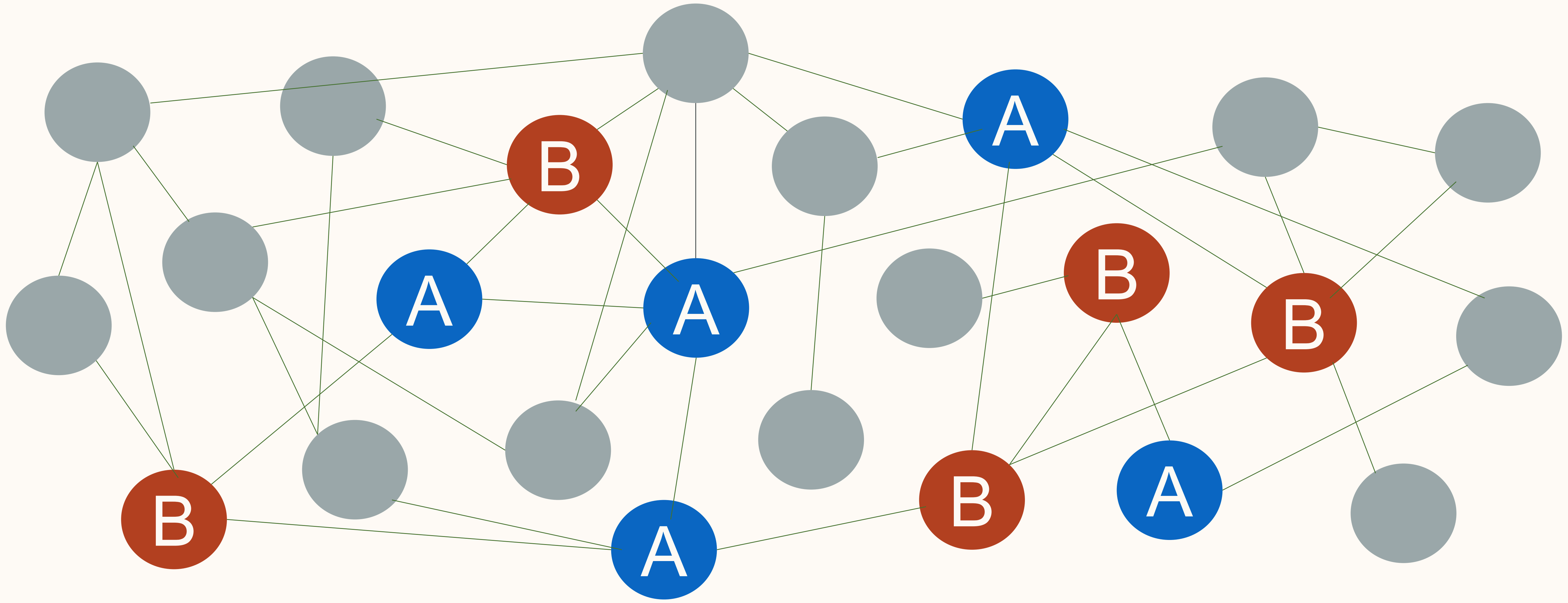
Preetam Nandy

Kinjal Basu

Shaunak Chatterjee

Ye Tu

Network Effect in A/B Testing



- A **violation of** Stable Unit Treatment Value Assumption (**SUTVA**), where
  Response depends on Own treatment + Neighbors' treatment (**network effect**)

- Average treatment effect (**ATE**):
  Average response when the whole population is receiving treatment B - Average response when the whole population is receiving treatment A

- **Incorrect ATE** estimation in **classical A/B testing**, since all neighbors of a treated node are not receiving the same treatment as the treated node

OASIS: Optimal Allocation Strategy and Importance Sampling

- Most literature on network A/B testing:
  - Graph clustering + cluster-level randomization for treatment assignment
  - Not suitable for dense graphs

- We present a complementary approach that does not require graph clustering by relies on certain assumption on treatment propagation.

- A two-step approach (OASIS)
  1. An approximate experimental design to provide a “correct” counterfactual experience to each experiment unit (by solving a constrained optimization problem)
  2. Post-experiment adjustment (via importance sampling) to correct for any leftover bias in estimating the average treatment effect

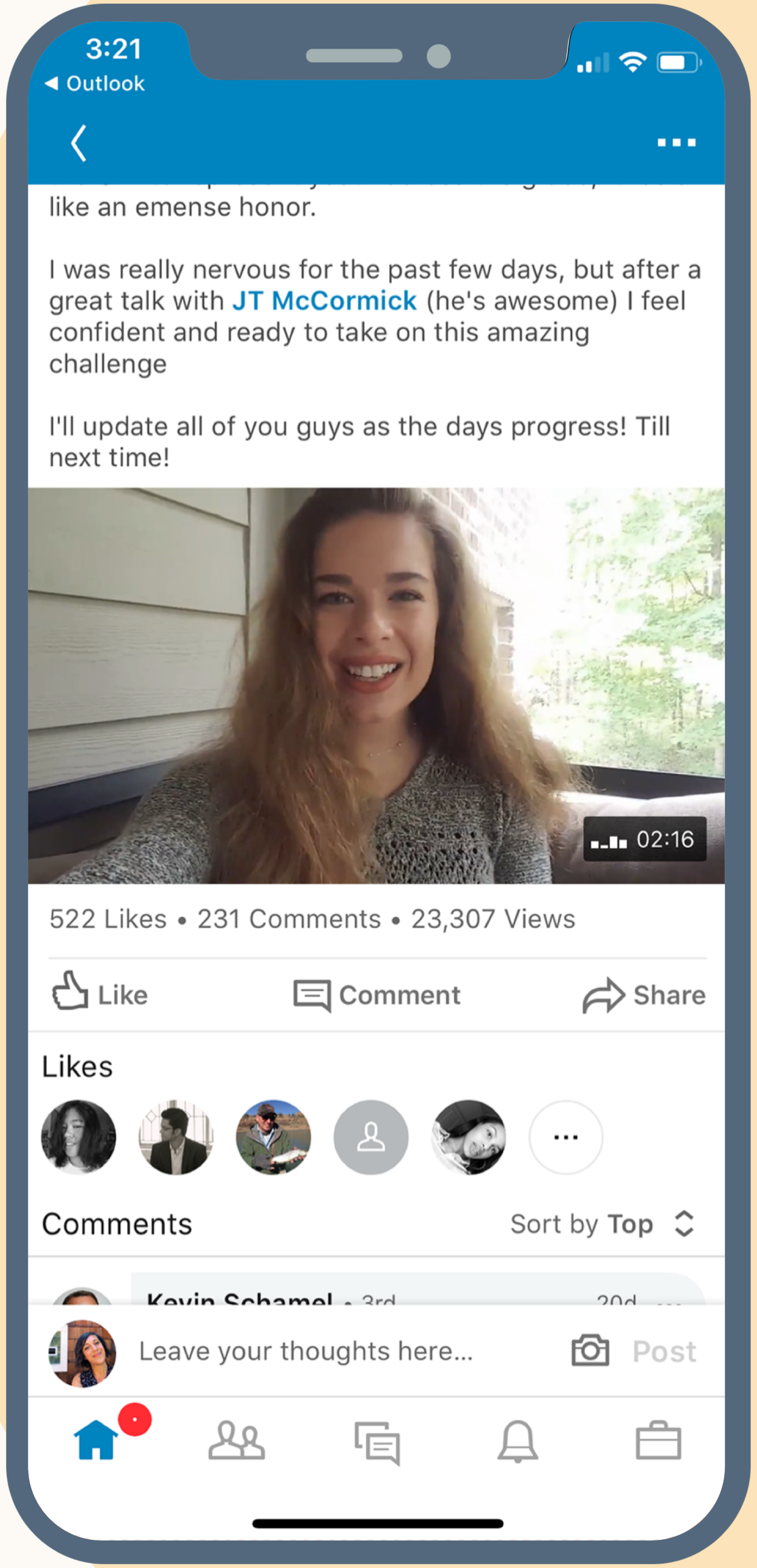
Preprint available on arXiv:
P. Nandy, K. Basu, S. Chatterjee, Y. Tu (2020). *A/B Testing in Dense Large-Scale Networks: Design and Inference*. arXiv:1901.10505.

Agenda

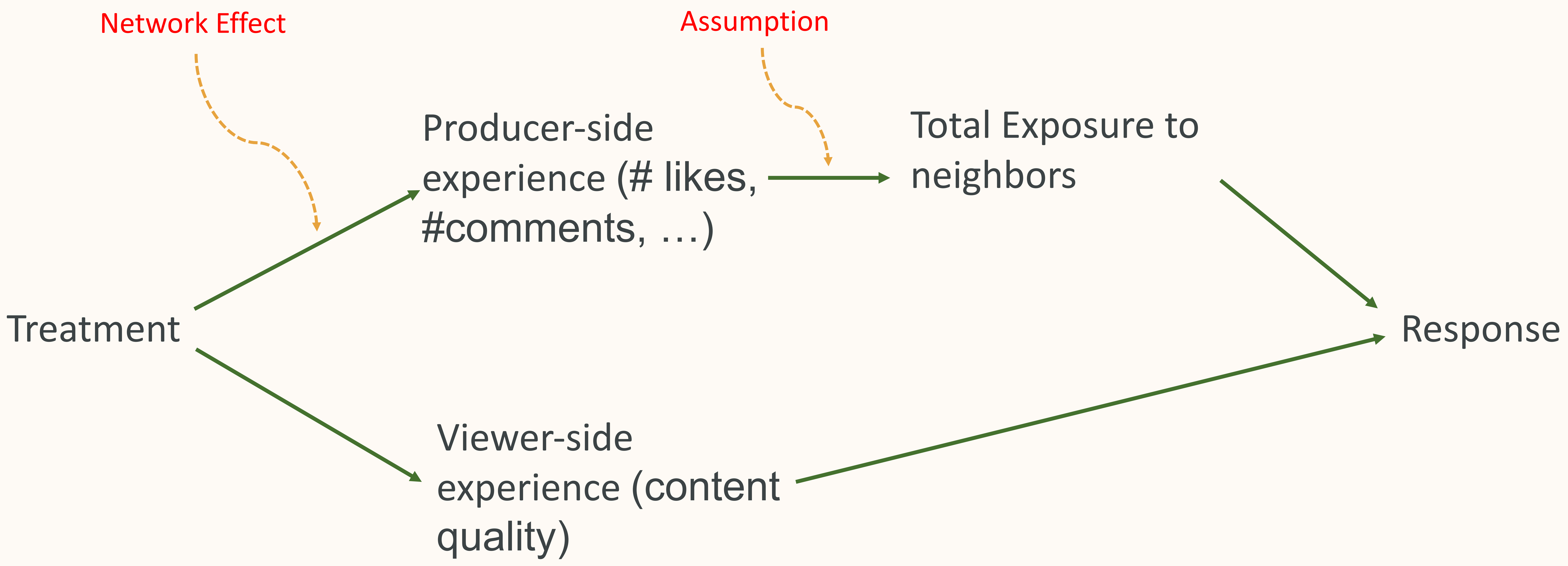
- Motivating Example
- Main Assumption
- OASIS
- Empirical Evaluation

A/B Testing Content Recommendation on LinkedIn News Feed

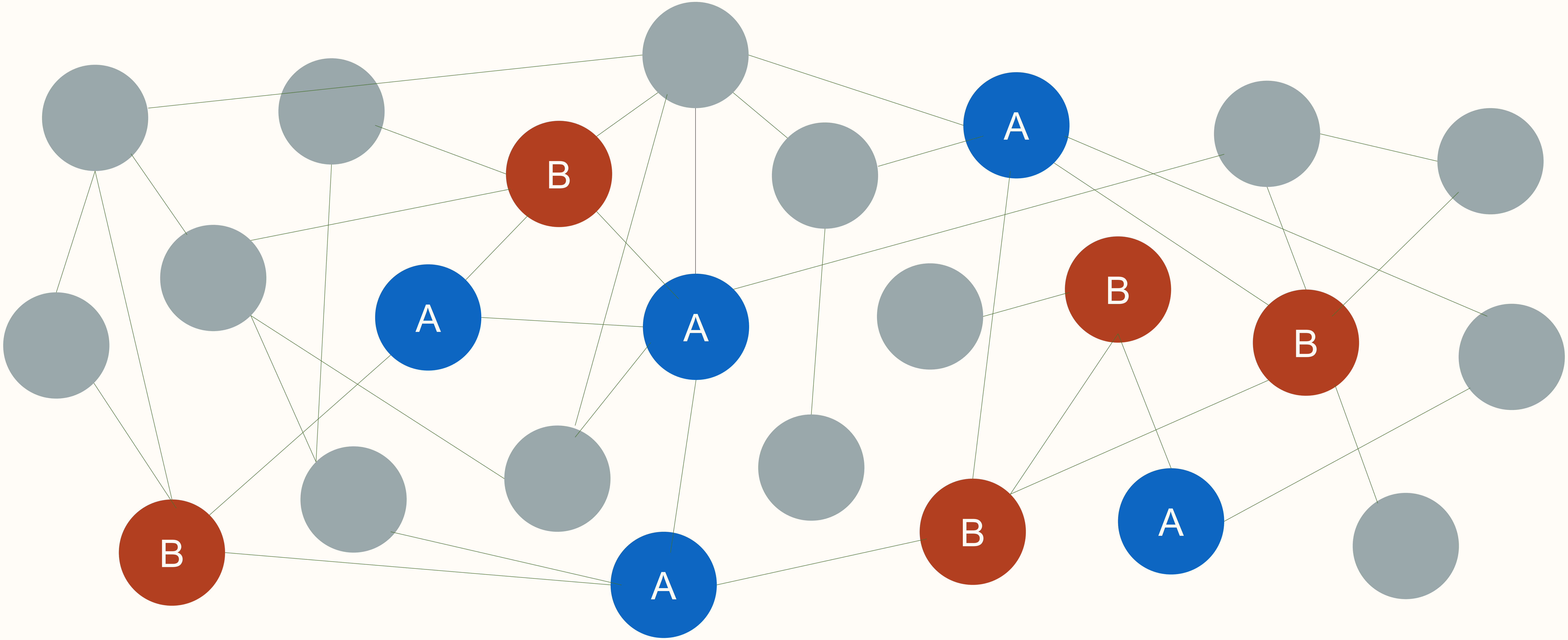
- Every member in the network is a content consumer and a content producer
- Treatment = Content recommendation model
- Response = Member engagement
- Response of a member depends on
  consumer-side experience (content quality) +
  producer-side experience (# likes, #comments, …)
- Producer-side experience (network effect) depends on neighbors' recommendation model (treatment)



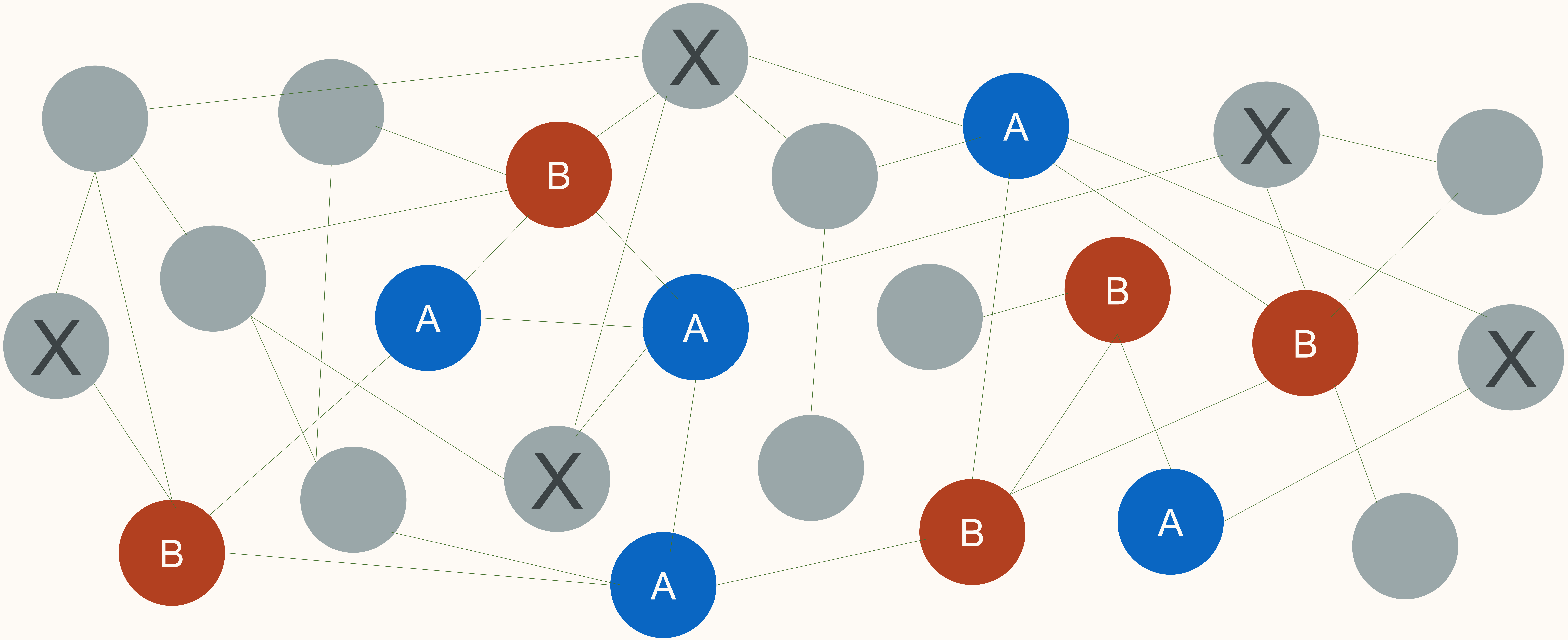
The Main Assumption



Step 1: Randomly Assign A and B

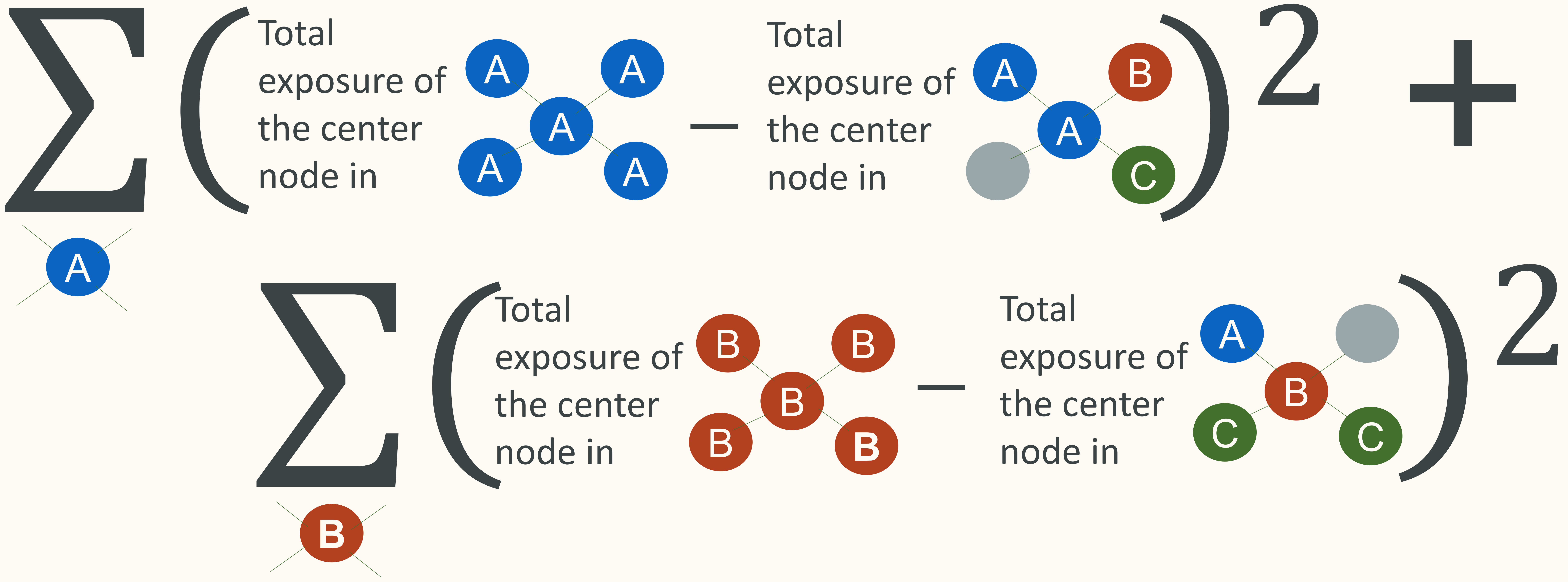


Step 2: Randomly Choose Nodes to Assign C



Step 3: Solve a Constrained Optimization Problem

Obtain C by minimizing



with respect to certain constraints controlling the risk of the experiment

Step 4: Run Experiment and Collect Data

Response:

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

Observed Total Exposure:

$$\{X(A)\} \text{ and } \{X(B)\}$$

Expected Total Exposure:

$$\{X(A)\} \text{ and } \{X(B)\}$$

Step 5: Importance Sampling Correction

Average Treatment Effect:

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

OASIS: Optimal Allocation Strategy and Importance Sampling

1. Randomly assign A and B

2. Randomly choose additional nodes X

3. Solve a constrained optimization to assign C to X

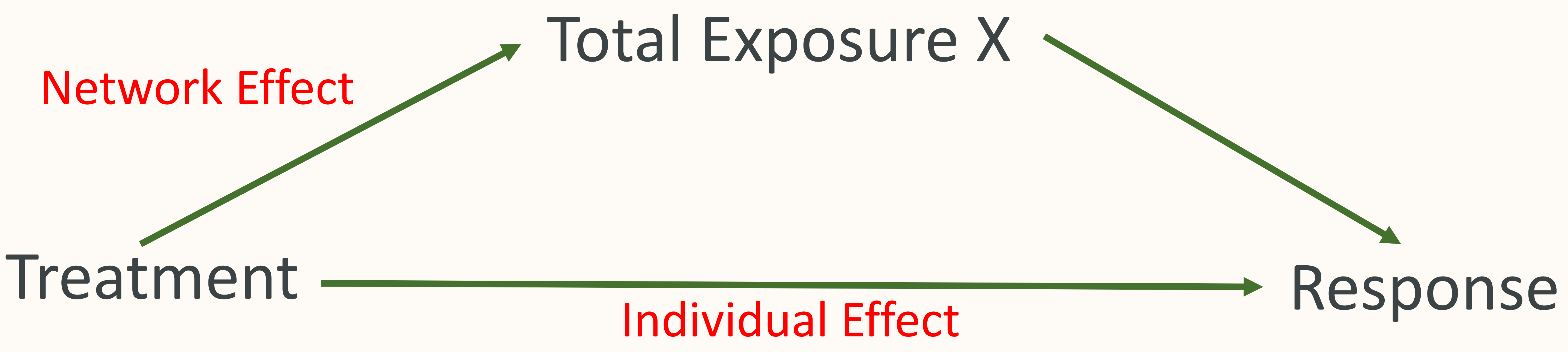
4. Run experiment and collect data

5. Importance sampling correction

OASIS: Optimal Allocation Strategy and Importance Sampling

- 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 a large-scale 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
- Correct bias with importance sampling
- See simulation and real-world experiment results in the paper

OASIS: Optimal Allocation Strategy and Importance Sampling



	Individual Effect	Network Effect	Total Effect
Classical A/B Testing	✓	×	×
OAS + Average Difference	✓	≈	≈
OASIS	✓	✓	✓

OASIS: Optimal Allocation Strategy and Importance Sampling

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

- 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.

Thank you