Causal Inference under Networked Interference and Intervention Policy Enhancement

Yunpu Ma[*]
LMU
Siemens AG
Otto-Hahn-Ring 6
81739 Munich

Yuyi Wang
ETH Zurich

Volker Tresp
LMU
Siemens AG
Otto-Hahn-Ring 6
81739 Munich

Abstract

Estimating individual treatment effects from data of randomized experiments is a critical task in causal inference. The Stable Unit Treatment Value Assumption (SUTVA) is usually made in causal inference. However, interference can introduce bias when the assigned treatment on one unit affects the potential outcomes of the neighboring units. This interference phenomenon is known as spillover effect in economics or peer effect in social science. Usually, in randomized experiments or observational studies with interconnected units, one can only observe treatment responses under interference. Hence, how to estimate the superimposed causal effect and recover the individual treatment effect in the presence of interference becomes a challenging task in causal inference. In this work, we study causal effect estimation under general network interference using GNNs, which are powerful tools for capturing the dependency in the graph. After deriving causal effect estimators, we further study intervention policy improvement on the graph under capacity constraint. We give policy regret bounds under network interference and treatment capacity constraint. Furthermore, a heuristic graph structure-dependent error bound for GNN-based causal estimators is provided.

1 Introduction

A common assumption made in causal inference is the consistency and interference-free assumption, i.e., the Stable Unit Treatment Value Assumption (SUTVA) Rubin [1980], under which the individual treatment response is consistently defined and unaffected by variations in other individuals. However, this assumption is problematic under a social network setting since peers are not independent or "no man is an island," as written by the poet John Donne.

Interference occurs when the treatment response of an individual is influenced through the exposure to its social contacts' treatments or affected by its social neighbors' outcomes through peer effects Bowers et al. [2013], Toulis and Kao [2013]. For instance, the treatment effect of an individual under a vaccination against an infectious disease might influence the health conditions of its surrounding individuals; or a personalized online advertisement might affect other individuals' purchase of the advertised item through opinion propagation on social networks. Separating individual treatment effect and peer effect in causal inference becomes an intractable problem under interference since, in randomized experiments or observational studies, one can only observe the superposition of both effects. The issue of how to estimate causal responses and make optimal policies on the network is studied in this work.

One of the main objectives of treatment effect estimation is to make better treatment decision rules for individuals according to their characteristics. Population-averaged utility functions have been studied in Manski [2009], Athey and Wager [2017], Kallus [2018], Kallus and Zhou [2018]. In those publications, a policy learner can adapt and improve its decision rules through the utility function. However, interactions among units are always ignored. On the other hand, a policy learner usually faces a capacity or budget constraint, as studied in Kitagawa and Tetenov [2017]. Therefore, in this work, we develop a new type of utility function defined on interconnected units and investigate provable policy improvement with budget constraints.

[*]yunpu.ma@siemens.com

1.1 Related Work

Causal inference with interference was studied in Hudgens and Halloran [2008], Tchetgen and VanderWeele [2012], Liu and Hudgens [2014]. However, the assumption of group-level interference, having partial interference within the groups and independence across different groups, is often invalid. Hence, several works focus on unit-level causal effects under cross-unit interference and arbitrary treatment assignments, such as Aronow et al. [2017], Forastiere et al. [2016], Ogburn et al. [2017a,b], Viviano [2019]. Other approaches for estimating causal effects on networks use graphical models, which are studied in Arbour et al. [2016], Tchetgen et al. [2017], Ogburn et al. [2018], Sherman and Shpitser [2018], Bhattacharya et al. [2019].

1.2 Notations and Previous Approaches

Let $\mathcal{G} = (\mathcal{N}, \mathcal{E}, A)$ denote a directed or undirected graph with a node set $\mathcal{N}$ of size n, an edge set $\mathcal{E}$, and an adjacency matrix $A \in \{0,1\}^{n\times n}$. For a node, or unit, $i \in \mathcal{N}$, let $\mathcal{N}_i$ indicate the set of neighboring nodes with $A_{ij} = 1$ excluding the node i itself, and let $\mathbf{X}_i$ denote the covariate vector of node i which is defined in the space $\boldsymbol{\chi}$. We focus on the Neyman–Rubin causal inference model Rubin [1974], Splawa-Neyman et al. [1990] here temporally. Let T_i be a binary variable with $T_i = 1$ indicating that node i is in the treatment group, and $T_i = 0$ if i is in the control group. Moreover, let Y_i be the outcome variable with $Y_i(T_i = 1)$ indicating the potential outcome of i under treatment $T_i = 1$ and $Y_i(T_i = 0)$ the potential outcome under control $T_i = 0$. Moreover, we use $T_{\mathcal{N}_i}$ and $Y_{\mathcal{N}_i}$ to represent the treatment assignments and potential outcomes of neighboring nodes $\mathcal{N}_i$, and $\mathbf{T}$ the entire treatment assignments vector.

In the SUTVA assumption, the individual treatment effect on node i is defined as the difference between outcomes under treatment and under control, i.e., $\tau(\mathbf{X}_i) := \mathbb{E}[Y_i(T_i = 1) - Y_i(T_i = 0)|\mathbf{X}_i]$. To estimate treatment effects under network interference, an exposure variable G is proposed in Toulis and Kao [2013], Bowers et al. [2013], Aronow et al. [2017]. The exposure variable G_i is a function of neighboring treatments $T_{\mathcal{N}_i}$. For instance, G_i can be a variable indicating the level of exposure to the treated neighbors, i.e., $G_i := \frac{\sum_{j\in\mathcal{N}_i} T_j}{|\mathcal{N}_i|}$.

Under the assumption that the outcome only depends on the individual treatment and neighborhood treatments, Forastiere et al. [2016] defines an individual treatment effect under the exposure $G_i = g$ as

$$\tau(\mathbf{X}_i, G_i = g) := \mathbb{E}[Y_i(T_i = 1, G_i = g) - Y_i(T_i = 0, G_i = g)|\mathbf{X}_i]. \quad (1)$$

Moreover, the spillover effect under the treatment $T_i = t$ and the exposure $G_i = g$ is defined as $\delta(\mathbf{X}_i, T_i = t, G_i = g) := \mathbb{E}[Y_i(T_i = t, G_i = g) - Y_i(T_i = t, G_i = 0)|\mathbf{X}_i]$. Treatment and spillover effects are then estimated using generalized propensity score (GPS) weighted estimators.

In general, the outcome model can be more complicated, depending on network topology and covariates of neighboring units. Ogburn et al. [2017a] investigates more general causal structural equations under dimension-reducing assumption, and the potential outcome reads $Y_{i,t} := f_Y(\mathbf{X}_i, s_X(\{\mathbf{X}_j|j \in \mathcal{N}_i\}), T_i, s_T(\{T_j|j \in \mathcal{N}_i\}))$, where s_X and s_T are summary functions of neighborhood covariates and treatment, e.g., they could be the summation or average of neighboring treatment assignments and covariates, respectively. Motivated from the above causal structural equation model, we incorporate GNN-based causal estimators with appropriate covariates and treatment aggregation functions as inputs.

Contributions This work has four major contributions. First, we propose GNN-based causal estimators for causal effect prediction and to recover direct treatment effect under interference (Section 2). Second, we define a novel utility function for policy optimization on a network and derive a graph-dependent policy regret bound (Section 3). Third, we provide an error bound for the GNN-based causal estimators (Section 3 and Appendix). Last, we conduct extensive experiments to verify the superiority of GNN-based causal estimators and show that the accuracy of a causal estimator is crucial for finding the optimal policy (Section 4).

2 GNN-based Causal Estimators

In this section, we introduce our GNN-based causal effect estimators under general network interference.

2.1 Structural Equation Model

Given the graph $\mathcal{G}$, the covariates of all units in the graph $\mathbf{X}$, and the entire treatment assignments vector $\mathbf{T}$, the structural equation model describing the considered data generation process is given as follows

$$\begin{aligned} T_i &= f_T(X_i) \\ Y_i &= f_Y(T_i, \mathbf{X}, \mathbf{T}, \mathcal{G}) + \epsilon_{Y_i}, \end{aligned} \quad (2)$$

for units $i = 1, \ldots, n$. This structural equation model encodes both the observational studies and the randomized experiments setting. In observational studies, e.g., on the Amazon dataset (see Section 4.1), the treatment $\mathbf{T}_i$ depends on the covariate $\mathbf{X}_i$ and the unknown specification of f_T, or even on the neighboring units under network interference. In the setting of the randomized experiment, e.g., experiments on Wave1 and Pokec

datasets, the treatment assignment function is specified as $f_T = \text{Bern}(p)$, where p represents predefined treatment probability. Function f_Y characterizes the causal response, which depends on, in addition to $\mathbf{X}_i$ and $\mathbf{T}_i$, the graph and neighboring covariates and treatment assignments. If only influences from first-order neighbors are considered, the response generation can be specified as $Y_i = f_Y(T_i, \mathbf{X}_{\mathcal{N}_i}, \mathbf{T}_{\mathcal{N}_i}, \mathcal{G}) + \epsilon_{Y_i}$. When the graph structure is given and fixed, we leave out $\mathcal{G}$ in the notation.

2.2 Distribution Discrepancy Penalty

Even without network interference, a covariate shift problem of counterfactual inference is commonly observed, namely the factual distribution $\Pr(\mathbf{X}, T)$ differs from the counterfactual distribution $\Pr(\mathbf{X}, 1-T)$. To avoid biased inference, Johansson et al. [2016], Shalit et al. [2017] propose a balancing counterfactual inference using domain-adapted representation learning. Covariate vectors are first mapped to a feature space via a feature map Φ. In the feature space, treated and control populations are balanced by penalizing the distribution discrepancy between $\Pr(\Phi(\mathbf{X})|T=0)$ and $\Pr(\Phi(\mathbf{X})|T=1)$ using the *Integral Probability Metric*. This approach is equivalent to finding a feature space such that the treatment assignment T and representation $\Phi(\mathbf{X})$ become approximately disentangled, namely $\Pr(\Phi(\mathbf{X}), T) \approx \Pr(\Phi(\mathbf{X}))P(T)$. We use the Hilbert-Schmidt Independence Criterion (HSIC) as the dependence test in the feature space. The empirical HSIC using a Gaussian RBF kernel is written as $\hat{HSIC}_{\mathcal{K}_\sigma}$ [1]. Note that incorporating the feature map and the representation balancing penalty is essential to tackle the imbalanced assignments in observational studies, e.g., on the Amazon dataset (see Section 4.1).

2.3 Graph Neural Networks

Graph neural networks can learn and aggregate feature information from distant neighbors, which makes it a right candidate for capturing the spillover effect given by the neighboring units. Different GNNs are employed and compared in our model, and we briefly provide a review.

Graph Convolutional Network (GCN) Kipf and Welling [2016] The graph convolutional layer in GCN is defined as $\mathbf{X}^{(l+1)} = \sigma\left(\hat{\mathbf{D}}^{-1/2}\hat{\mathbf{A}}\hat{\mathbf{D}}^{-1/2}\mathbf{X}^{(l)}\mathbf{W}^{(l)}\right)$, where $\mathbf{X}^{(l+1)}$ is the hidden output from the l-th layer with $\mathbf{X}^{(0)}$ being the input features matrix, and σ is the activation function, e.g., ReLU. The modified adjacency $\hat{\mathbf{A}}$ with inserted self-connections is defined as $\hat{\mathbf{A}} := \mathbf{A} + \mathbf{I}$, and $\hat{\mathbf{D}}$ denotes the node degree matrix of $\hat{\mathbf{A}}$.

GraphSAGE GraphSAGE Hamilton et al. [2017] is an inductive framework for calculating node embeddings and aggregating neighbor information. The mean aggregation operator in the GraphSAGE reads $\mathbf{X}_i^{(l+1)} = \sigma\left(\text{mean}_{j\in\mathcal{N}_i\cup\{i\}}\mathbf{X}_j^{(l)}\mathbf{W}^{(l)}\right)$. Traditional GCN algorithms perform spectral convolution via eigen-decomposition of the full graph Laplacian. In contrast, GraphSAGE computes a localized convolution by aggregating the neighborhood around a node, which resembles the simulation protocol of linear treatment response with spillover effect for semi-synthetic experiments (see Section 4.1). Due to the resemblance, a better causal estimator is expected when using GraphSAGE as the aggregation function (see the beginning of Appendix G.3 for more heuristic motivations.).

1-GNN 1-GNN Morris et al. [2018] is a variation of GraphSAGE, which performs separate transformations of node features and aggregated neighborhood features. Since the features of the considered unit and its neighbors contribute differently to the superimposed outcome, it is expected that the 1-GNN is more expressive than GraphSAGE. The convolutional operator of 1-GNN has the form $\mathbf{X}_i^{(l+1)} = \sigma\left(\mathbf{X}_i^{(l)}\mathbf{W}_1^{(l)} + \text{mean}_{j\in\mathcal{N}_i}\mathbf{X}_j\mathbf{W}_2^{(l)}\right)$.

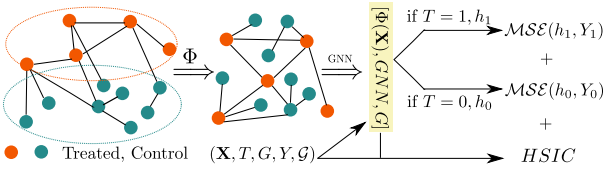

Figure 1: Treated and control populations have different distributions in the covariate vectors space. Through a map Φ and distribution discrepancy term HSIC, features and treatment assignments become disentangled in the feature space. Before and after the feature map Φ, the adjacency matrix remains the same. After applying GNNs, for each node i, the concatenation $[\Phi(\mathbf{X}_i), GNN(\Phi(\mathbf{X}), \mathbf{T})_i, G_i]$ is fed into outcome prediction network h_1 or h_0 depending on the treatment assignment. The loss function consists of outcome prediction error and the distribution discrepancy in the feature space.

2.4 GNN-based Causal Estimators

We use the percentage of treated neighboring nodes, i.e., the random variable G, as the treatment summary function, and the output of GNNs as the covariate aggregation function. The concatenation $[\Phi(\mathbf{X}_i), GNN(\Phi(\mathbf{X}), \mathbf{T})_i, G_i]$ of node i is then fed into the outcome prediction network h_1 or h_0, depending on T_i, where h_1 and h_0 are neural networks with a scalar

[1] Expression for $\hat{HSIC}_{\mathcal{K}_\sigma}$ is relegated to Appendix A.

output. Note that $GNN(\Phi(\mathbf{X}), \mathbf{T})_i$ indicates that the treatment vector $\mathbf{T}$ is also a GNNs' input. During the implementation, the treatment assignment vector masks the covariates, and GNN models use the masked covariates $T_i\mathbf{X}_i$, for $i = 1, \ldots, n$, as inputs. In summary, given $(\Phi(X)_i, T_i, G_i, Y_i)$ and graph $\mathcal{G}$, the loss function for GNN-based estimators is defined as

$$\mathcal{L}_{\text{est}} := \mathcal{MSE}\left(h_{T_i}([\Phi(\mathbf{X}_i), GNN(\Phi(\mathbf{X}), \mathbf{T})_i, G_i]), Y_i\right) + \kappa \hat{HSIC}_{\mathcal{K}_\sigma},$$

where κ and σ are tunable hyperparameters. Our model is illustrated in Fig. 1. During the implementation, we incorporate two types of empirical representation balancing: balancing the outputs of representation network Φ to tackle imbalanced assignments, denoted as $\hat{HSIC}^{\Phi}$, and balancing the outputs of the GNN representations to tackle imbalanced spillover exposure, denoted as $\hat{HSIC}^{GNN}$.

At this point, it is necessary to emphasize that only the causal responses of a part of the units in $\mathcal{N}$ are relevant to the models. The GNN-based models use this part of causal responses, the network structure $\mathcal{G}$, and covariates $\mathbf{X}$ as input, and can predict the superimposed causal effects of the remaining units. Note that for GNN-based nonparametric models, the identifiability of causal response is guaranteed under reasonable assumptions similar to those given in Section 3.2 of Ogburn et al. [2017a]. The proof is relegated to Appendix B.

Notice that the outcome prediction networks h_0 and h_1 are trained to estimate the superposition of individual treatment effect and spillover effect. Still, after fitting the observed outcomes, we expect to extract the non-interfered individual treatment effect from the causal estimators by assuming that the considered unit is isolated. An individual treatment effect estimator can be defined similarly to Eq. 1. To be more specific, the individual treatment effect of unit i is expected to be extracted from GNN-based estimators by setting its exposure to $G_i = 0$ and its neighbors' covariates to $\mathbf{0}$, namely [2]

$$\hat{\tau}(\mathbf{X}_i) = h_1([\Phi(\mathbf{X}_i), \mathbf{0}, 0]) - h_0([\Phi(\mathbf{X}_i), \mathbf{0}, 0]). \quad (3)$$

3 Intervention Policy on Graph

After obtaining the treatment effect estimator, we develop an algorithm for learning intervention assignments to maximize the utility on the entire graph, and the learned rule for assignment is called a policy. As suggested in Athey and Wager [2017], without interference a utility function is defined as $A(\pi) = \mathbb{E}[(2\pi(\mathbf{X}_i) - 1)(Y_i(T_i = 1) - Y_i(T_i = 0))] = \mathbb{E}[(2\pi(\mathbf{X}_i) - 1)\tau(\mathbf{X}_i)]$. An optimal policy $\hat{\pi}_n$ is obtained by maximizing the n-sample empirical utility function $\hat{A}^{\tau}_n(\pi) := \frac{1}{n}\sum_{i=1}^{n}(2\pi(\mathbf{X}_i)-1)\hat{\tau}(\mathbf{X}_i)$ given the individual treatment response estimator $\hat{\tau}$, i.e., $\hat{\pi}_n \in \operatorname{argmax}_{\pi\in\Pi}\hat{A}^{\tau}_n(\pi)$, where Π indicates the policy function class. Notably, $\hat{\pi}_n$ tends to assign treatment to units with positive treatment effect and control to units with negative responses.

Now, consider the outcome variable Y_i under network interference. For notational simplicity and clarity of the later proof, we assume first-order interference from nearest neighboring units, hence the outcome variable can be written as $Y_i(T_i, \mathbf{X}_{\mathcal{N}_i}, T_{\mathcal{N}_i})$. Inspired by the definition of $A(\pi)$, the utility function of a policy π under interference is defined as

$$S(\pi) := \mathbb{E}[(2\pi(\mathbf{X}_i) - 1)(Y_i(T_i = 1, \mathbf{X}_{\mathcal{N}_i}, T_{\mathcal{N}_i} = \pi(\mathbf{X}_{\mathcal{N}_i})) - Y_i(T_i = 0, \mathcal{G} = \emptyset))], \quad (4)$$

where $Y_i(T_i = 0, \mathcal{G} = \emptyset)$ with an empty graph represents the individual outcome under control without any network influence [3]. After some manipulations, $S(\pi)$ equals the sum of individual treatment effect and spillover effect, i.e., $S(\pi) = \mathbb{E}[(2\pi(\mathbf{X}_i) - 1)(\tau_i + \delta_i(\pi))]$, where

$$\begin{aligned}\tau_i &:= \mathbb{E}[Y_i(T_i = 1, \mathcal{G} = \emptyset) - Y_i(T_i = 0, \mathcal{G} = \emptyset)|\mathbf{X}_i] \\ \delta_i(\pi) &:= \mathbb{E}[Y_i(T_i = 1, \mathbf{X}_{\mathcal{N}_i}, T_{\mathcal{N}_i} = \pi(\mathbf{X}_{\mathcal{N}_i})) \\ &\qquad - Y_i(T_i = 1, \mathcal{G} = \emptyset)|\mathbf{X}_i, \mathbf{X}_{\mathcal{N}_i}].\end{aligned}$$

To be more specific, τ_i is the conventional individual treatment effect, while $\delta_i(\pi)$ represents the spillover effect under the policy π and when $T_i = 1$. Due to the network-dependency in the spillover effect, an optimal policy will not merely treat units with positive responses but also adjust its intervention on the entire graph to maximize the spillover effects.

Next, we establish guarantees for the regret of learned intervention policy. Let $\hat{\tau}_i$ and $\hat{\delta}_i(\pi)$ denote the estimator of τ_i and $\delta_i(\pi)$, respectively. Given the true models τ_i and $\delta_i(\pi)$, let $S^{\pi,\delta}_n(\pi) := \frac{1}{n}\sum_{i=1}^{n}(2\pi(\mathbf{X}_i) - 1)(\tau_i + \delta_i(\pi))$ be the empirical analogue of $S(\pi)$, and let $\hat{S}^{\pi,\delta}_n(\pi) := \frac{1}{n}\sum_{i=1}^{n}(2\pi(\mathbf{X}_i) - 1)(\hat{\tau}_i + \hat{\delta}_i(\pi))$ be the empirical utility with estimators plugged in. Using learned causal estimators, an optimal intervention policy from the empirical utility perspective can be obtained from $\hat{\pi}_n \in \operatorname{argmax}_{\pi\in\Pi}\hat{S}^{\pi,\delta}_n(\pi)$. Moreover, the best possible intervention policy from the functional class Π with respect to the utility $S(\pi)$ is written as $\pi^\star := \operatorname{argmax}_{\pi\in\Pi}S(\pi)$, and the policy regret between $\pi^\star$ and $\hat{\pi}_n$ is defined as $\mathcal{R}(\hat{\pi}_n) := S(\pi^\star) - S(\hat{\pi}_n)$. Throughout the estimation of policy regret, we maintain the following assumptions.

[2] Spillover effect can be extracted similarly.

[3] Hence $\mathbf{X}_{\mathcal{N}_i}$ and $T_{\mathcal{N}_i}$ are omitted in the expression.

Assumption 1.
(BO) Bounded treatment and spillover effects: There exist $0 < M_1, M_2 < \infty$ *such that the individual treatment effect satisfies* $|\tau_i| \leq M_1$ *and the spillover effect satisfies* $\forall \pi \in \Pi, |\delta_i(\pi)| \leq M_2$.
(WI) Weak independence assumption: For any node indices i *and* j*, the weak independence assumption assumes that* $\mathbf{X}_i \perp \mathbf{X}_j$ *if* $A_{ij} = 0$*, or* $\nexists k$ *with* $A_{ik} = A_{kj} = 1$.
(LIP) Lipschitz continuity of the spillover effect w.r.t. policy: Given two treatment policies π_1 *and* π_2*, for any node* i *the spillover effect satisfies* $|\delta_i(\pi_1) - \delta_i(\pi_2)| \leq L||\pi_1 - \pi_2||_\infty$*, where the Lipschitz constant satisfies* $L > 0$ *and* $||\pi_1 - \pi_2||_\infty := \sup_{\mathbf{X} \in \boldsymbol{\chi}} |\pi_1(\mathbf{X}) - \pi_2(\mathbf{X})|$.
(ES) Uniformly consistency: after fitting experimental or observational data on $\mathcal{G}$*, individual treatment effect estimator satisfies* $\frac{1}{n}\sum_{i=1}^{n} |\tau_i - \hat{\tau}_i| < \frac{\alpha_\tau}{\sqrt{n}}$*, and spillover estimator satisfies* $\forall \pi \in \Pi$, $\frac{1}{n}\sum_{i=1}^{n} |\delta_i(\pi) - \hat{\delta}_i(\pi)| < \frac{\alpha_\delta}{\sqrt{n}}$*, where* $\alpha_\tau > 0$ *and* $\alpha_\delta > 0$ *are scaling factors that characterize the errors of estimators.*

Notice that the (ES) assumption requires consistent estimators of the individual treatment effect and the spillover effect, which is the fundamental problem of causal inference with interference. In our GNN-based model, these empirical errors are particularly difficult to estimate due to the lack of proper theoretical tools for understanding GNNs. To grasp how these GNN-based causal estimators are influenced by the network structure and network effect, in Appendix G.3, we study a particular class of GNNs, which is inspired by the *surrogate model* of nonlinear graph neural networks and have the following claim.

Claim 1. *GNN-based causal estimators restricted to a particular class for predicting the superimposed causal effects have an error bound* $\mathcal{O}(\sqrt{\frac{D_{max}^3 \ln D_{max}}{n}})$*, where* $D_{max} := 1 + d_{max} + d_{max}^2$ *and* d_{max} *is the maximal node degree in the graph.*

The above claim indicates that an accurate and consistent causal estimator is difficult with large network effects. Worse case is that the $\frac{1}{\sqrt{n}}$ convergence rate in the (ES) assumption becomes unreachable when $d_{\max}(n)$ depends on the number of units. The exact convergence rate of causal estimators is impossible to derive since it depends on the topology of the network, and it beyond the theoretical scope of this work.

Besides, (LIP) assumes that the change of received spillover effect is bounded after modifying the treatment assignments of one unit's neighbors. We will use hypergraph techniques, instead of chromatic number arguments, to give a tighter bound of policy regrets. Another advantage is that the weak independence (WI) assumption can be relaxed to support longer dependencies on the network. However, by relaxing (WI), the power of $d_{\max}$ in Theorem 1 and 2 needs to be modified correspondingly. For example, if we assume a next-nearest neighbors dependency of covariates, i.e., $\mathbf{X}_i \perp \mathbf{X}_j$ for $j \notin i \cup \mathcal{N}_i \cup \mathcal{N}_i^{(2)}$, then the term $d_{\max}^2$ in Theorem 1 and 2 needs to be modified to $d_{\max}^4$.

Under Assumption 1, we can have the following bound.

Theorem 1. *By Assumption 1, for any small* $\epsilon > 0$*, the policy regret is bounded by* $\mathcal{R}(\hat{\pi}_n) \leq \frac{2(\alpha_\tau + \alpha_\delta)}{\sqrt{n}} + 2\epsilon$ *with probability at least* $1 - \mathcal{N}\left(\Pi, \frac{\epsilon}{4(2M_1+2M_2+L)}\right) \exp\left(-\frac{n\epsilon^2}{32(d_{\max}^2+1)(M_1+M_2)^2}\right)$*, where* $\mathcal{N}\left(\Pi, \frac{\epsilon}{4(2M_1+2M_2+L)}\right)$ *indicates the covering number* [4] *on the functional class* Π *with radius* $\frac{\epsilon}{4(2M_1+2M_2+L)}$*, and* $d_{\max}$ *is the maximal node degree in the graph* $\mathcal{G}$.

Proof. Under (WI) and (BO), we can use concentration inequalities of networked random variables defined on a hypergraph, which is derived from graph $\mathcal{G}$ to bound the convergence rate. Moreover, the Lipschitz assumption (LIP) allows an estimation of the covering number of the policy functional class Π. More discussions on the plausibility of Assumption 1 and the full proof are relegated to Appendix G. ∎

Suppose that the policy functional class Π is finite and its capacity is bounded by $|\Pi|$. According to Theorem 1, with probability at least $1 - \delta$, the policy regret is bounded by $\mathcal{R}(\hat{\pi}_n) \leq \frac{2(\alpha_\tau+\alpha_\delta)}{\sqrt{n}} + 8(M_1+M_2)\sqrt{\frac{2(d_{\max}^2+1)}{n} \log \frac{|\Pi|}{\delta}} \approx \frac{2(\alpha_\tau+\alpha_\delta)}{\sqrt{n}} + 8d_{\max}(M_1 + M_2)\sqrt{\frac{2}{n} \log \frac{|\Pi|}{\delta}}$. It indicates that optimal policies are more difficult to find in a dense graph even under weak interactions between neighboring nodes.

In a real-world setting, treatments could be expensive. So the policymaker usually encounters a budget or capacity constraints, e.g., the proportion of patients receiving treatment is limited, and to decide who should be treated under constraints is a challenging problem Kitagawa and Tetenov [2017]. Through the interference-free welfare function $A(\pi)$, a policy is trained to make treatment choices using only each individual's features. In contrast, under interference, a smart policy should maximize the utility function Eq. (4) by deciding whether to treat an individual or expose it under neighboring treatment effects such that a required constraint can be satisfied. Therefore, in the second part of the experiments, after fitting causal estimators, we investigate policy networks

[4] The covering number characterizes the capacity of a functional class. Definition is provided in the Appendix G.

that maximize the utility function $S(\pi)$ on the graph and satisfy a treatment proportion constraint.

To be more specific, we consider the constraint where only p_t percentage of the population can be assigned to treatment [5]. The corresponding sample-averaged loss function for a policy network π under capacity constraint is defined as $\mathcal{L}_{\text{pol}}(\pi) := -\hat{S}_n^{\tau,\delta}(\pi) + \gamma(\frac{1}{n}\sum_{i=1}^n \pi(\mathbf{X}_i) - p_t)$, where γ is a hyperparameter for the constraint. Optimal policy under capacity constraint is obtained by $\hat{\pi}_n^{p_t} \in \min_{\pi\in\Pi} \mathcal{L}_{\text{pol}}(\pi)$. A capacity-constrained policy regret bound is provided in Theorem 2, which is proved in Appendix G.2. It indicates that if in the constraint p_t is small, then the optimal capacity-constrained policy will be challenging to find. Increasing the treatment probability can not guarantee the improvement of the group's interest due to the non-linear network effect. Therefore, finding the balance between optimal treatment probability, treatment assignment, and group's welfare is a provocative question in social science.

Theorem 2. *By Assumption 1, for any small $\epsilon > 0$, the policy regret under the capacity constraint p_t is bounded by $\mathcal{R}(\hat{\pi}_n^{p_t}) \leq \frac{2(\alpha_\tau+\alpha_\delta)}{\sqrt{n}} + 2\epsilon$ with probability at least $1 - \mathcal{N}\exp\left(-\frac{n\epsilon^2}{32(d_{\max}^2+1)(M_1+M_2)^2}\right)$, where $\mathcal{N} := \mathcal{N}\left(\Pi, \frac{\epsilon}{8[(M_1+M_2+L)+\frac{1}{p_t}(M_1+M_2)]}\right)$ indicates the covering number on the functional class Π with radius $\frac{\epsilon}{8[(M_1+M_2+L)+\frac{1}{p_t}(M_1+M_2)]}$, and $d_{\max}$ is the maximal node degree in the graph $\mathcal{G}$.*

4 Experiments

4.1 Datasets

The difficulties of evaluating the performance of the proposed estimators lie in the broad set of missing outcomes under counterfactual inference. Therefore, we conduct randomized experiments on two semi-synthetic datasets with *ground-truth* response generation functions, and observational studies on one real dataset with *unknown* treatment assignment and response generation functions. Notably, in the randomized experiment setting, we consider a linear response generation function inspired by Eq. 5 of Toulis and Kao [2013], $G_0 : Y_i = Y_i(T_i = 0, \mathcal{G} = \emptyset) + T_i\tau(\mathbf{X}_i) + \delta_i(\mathbf{X}, \mathbf{T}, \mathcal{G}) + \epsilon_{Y_i}$, where $Y_i(T_i = 0, \mathcal{G} = \emptyset)$ is the outcome under control and without network interference, and ϵ_{Y_i} represents Gaussian noise. $\tau(\mathbf{X}_i)$ and $\delta_i(\mathbf{X}, \mathbf{T}, \mathcal{G})$ represent individual treatment effect and spillover effect, respectively, whose forms are dataset-dependent and discussed below.

[5] Note that here p_t differs from the treatment probability p from causal structural equations in the randomized experiment setting.

To further investigate the superiority of the GNN-based causal estimators on nonlinear causal responses, we consider the following data generation function inspired by Section 4.2 of Toulis and Kao [2013], $G_1 : Y_i = Y_i(T_i = 0, \mathcal{G} = \emptyset) + T_i\tau(\mathbf{X}_i) + \delta_i(\mathbf{X}, \mathbf{T}, \mathcal{G}) + \kappa\delta_i^2(\mathbf{X}, \mathbf{T}, \mathcal{G}) + \epsilon_{Y_i}$, where κ characterizes the strength of nonlinear effects. In addition, a more complicated nonlinear response generation function $G_2 : Y_i = Y_i(T_i = 0, \mathcal{G} = \emptyset) + T_i\tau(\mathbf{X}_i) + \delta_i(\mathbf{X}, \mathbf{T}, \mathcal{G}) + \frac{\kappa}{2}\delta_i^2(\mathbf{X}, \mathbf{T}, \mathcal{G}) + \frac{\kappa}{2}\tau(\mathbf{X}_i)\delta_i(\mathbf{X}, \mathbf{T}, \mathcal{G}) + \epsilon_{Y_i}$ is considered, where the quadratic terms signify the spillover effect depending on the individual treatment effect.

Wave1 Wave1 is an in-school questionnaire data collected through the National Longitudinal Study of Adolescent Health project Chantala and Tabor [1999]. The questionnaire contains questions such as age, grade, health insurance, etc. Due to the anonymity of Wave1, we use the symmetrized k-NN graph derived from the questionnaire data as the friendship network. In our experiments, we choose $k = 10$, and the resulting friendship network has $5,578$ nodes and $100,158$ links. We assume a randomized experiment conducted on the friendship network which describes students' improvements of performance through assigning to a tutoring program or through the peer effect. Hence $Y_i(T_i = 0, \mathcal{G} = \emptyset)$ represents the overall performance of student i before assignment to a tutoring program and before being exposed to peer influences, $\tau(\mathbf{X}_i)$ the simulated performance difference after an assignment, and $\delta_i(\mathbf{X}, \mathbf{T}, \mathcal{G})$ the synthetic peer effect. Exact forms of $Y_i(T_i = 0, \mathcal{G} = \emptyset)$ and $\tau(\mathbf{X}_i)$ depend nonlinearly on the features of each student. Moreover, the first-order peer effect is simulated as $\delta_i(\mathbf{X}, \mathbf{T}, \mathcal{G}) := \alpha\frac{1}{|\mathcal{N}_i|}\sum_{j\in\mathcal{N}_i} T_j\tau(\mathbf{X}_j)$, where the decay parameter α characterizes the decay of influence. In randomized experiments reported in the main text, we randomly assign 10% of the population to the treatment. Details of the generating process and more experiment results with different settings are relegated to Appendix C and F.

Pokec The friendship network derived from the Wave1 questionnaire data may violate the power-law degree distribution of real networks. Hence, we further conduct experiments on the real social network Pokec Takac and Zabovsky [2012] with generated responses. Pokec is an online social network in Slovakia with profile data, including age, gender, education, etc. We consider randomized experiments on the Pokec social network, in which personalized advertisements of a new health medicine are pushed to some users. We assume that the response of exposed users to the advertisement only depends on a few properties, such as age, weight, smoking status, etc. We keep profiles with complete infor-

Table 1: Experimental results of randomized experiments on the Wave1 and Pokec datasets using linear response generation function G_0. For Wave1, we set (node degree) $k = 10$, (decay parameter)$\alpha = 0.5$, and (treatment probability) $p = 0.1$, and for Pokec $p = 0.1$. Improvements are obtained by comparing with the best baselines.

	Wave1		Pokec	
	$\sqrt{\mathcal{MSE}}$	ϵ_{PEHE}	$\sqrt{\mathcal{MSE}}$	ϵ_{PEHE}
DA GB	0.721 ± 0.054	0.289 ± 0.061	0.713 ± 0.016	0.321 ± 0.057
DA RF	1.037 ± 0.122	0.790 ± 0.215	0.749 ± 0.023	0.840 ± 0.087
DR GB	0.831 ± 0.109	0.499 ± 0.185	0.686 ± 0.020	0.275 ± 0.051
DR EN	0.929 ± 0.091	0.733 ± 0.135	0.695 ± 0.019	0.247 ± 0.060
GPS	0.238 ± 0.012	0.150 ± 0.047	0.329 ± 0.010	0.147 ± 0.010
GCN + $H\hat{S}IC^{\Phi/GNN}$	0.192 ± 0.019	0.047 ± 0.018	0.305 ± 0.011	0.136 ± 0.009
GraphSAGE + $H\hat{S}IC^{\Phi/GNN}$	0.181 ± 0.016	0.042 ± 0.020	0.303 ± 0.008	$\mathbf{0.123 \pm 0.003}$
1-GNN + $H\hat{S}IC^{\Phi/GNN}$	$\mathbf{0.176 \pm 0.011}$	$\mathbf{0.035 \pm 0.011}$	$\mathbf{0.302 \pm 0.004}$	0.130 ± 0.006
Improve	26.1%	76.7%	8.2%	16.3%

Table 2: Experimental result on the pos Amazon dataset without representation balancing and under different imbalance penalties. Improvements are obtained by comparing with the best baselines.

	$\sqrt{\mathcal{MSE}}$	ϵ_{PEHE}
DA GB	0.601 ± 0.007	1.370 ± 0.016
DA RF	0.604 ± 0.019	1.398 ± 0.013
DR GB	0.615 ± 0.022	1.222 ± 0.020
DR EN	1.104 ± 0.001	1.929 ± 0.003
GPS	0.399 ± 0.003	1.968 ± 0.025
GCN	0.312 ± 0.002	2.400 ± 0.201
GCN + $H\hat{S}IC^{GNN}$	0.303 ± 0.006	1.881 ± 0.076
GCN + $H\hat{S}IC^{\Phi}$	0.301 ± 0.002	1.531 ± 0.024
GraphSAGE	0.305 ± 0.001	1.984 ± 0.026
GraphSAGE + $H\hat{S}IC^{GNN}$	0.296 ± 0.002	1.567 ± 0.051
GraphSAGE + $H\hat{S}IC^{\Phi}$	0.300 ± 0.002	1.358 ± 0.025
1-GNN	0.279 ± 0.000	1.512 ± 0.111
1-GNN + $H\hat{S}IC^{GNN}$	$\mathbf{0.276 \pm 0.002}$	1.434 ± 0.030
1-GNN + $H\hat{S}IC^{\Phi}$	$\mathbf{0.277 \pm 0.002}$	$\mathbf{1.098 \pm 0.031}$
Improve	30.8%	10.1%

mation on these properties, and the resulting Pokec social network contains $11,623$ nodes and $76,752$ links. Let $Y_i(T_i = 0, \mathcal{G} = \emptyset)$ represent the purchase of this new health medicine without external influence on the decision, $\tau(\mathbf{X}_i)$ the purchase difference after seeing the advertisement, $\delta_i(\mathbf{X}, \mathbf{T}, \mathcal{G})$ the purchase difference due to social influences. For randomized experiments on the Pokec social network, we also consider peer effects from next-nearest neighbors by defining $\delta_i(\mathbf{X}, \mathbf{T}, \mathcal{G}) := \alpha \frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} T_j \tau(\mathbf{X}_j) + \alpha^2 \frac{1}{|\mathcal{N}_i^{(2)}|} \sum_{k \in \mathcal{N}_i^{(2)}} T_k \tau(\mathbf{X}_k)$, where the decay parameter α characterizes the decay of influence. Details and more experimental results with different hyperparameter settings are given in Appendix D and F.

Amazon The co-purchase dataset from Amazon contains product details, review information, and a list of similar products. Therefore, there is a directed network of products that describes whether a substitutable or complementary product is getting co-purchased with another product Leskovec et al. [2007]. To study the causal effect of reviews on the sales of products, Rakesh et al. [2018] generates a dataset containing products with only positive reviews from the Amazon co-purchase dataset, named as pos Amazon, and Amazon for short. In this dataset, all items have positive reviews, i.e., the average rating is larger than 3, and one item is considered to be treated if there are more than three reviews under this item; otherwise, an item is in the control group. In this setting, pos Amazon is an over-treated dataset with more than 70% of products being in the treatment group. Word2vec embedding of an item's review serves as the feature vector of this item. Moreover, the individual treatment effect of an item is approximated by matching it to other items having similar features and under minimal exposure to neighboring nodes' treatments.

4.2 Results of Causal Estimators

Evaluation Metrics One evaluation metric is the square root of MSE for the prediction of the observed outcomes on the test dataset $\mathcal{U}_T$, which is defined as $\sqrt{\mathcal{MSE}} := \sqrt{\frac{1}{|\mathcal{U}_T|} \sum_{i \in \mathcal{U}_T} (Y_i - h_{T_i})^2}$, where h_{T_i} denotes the output of the outcome prediction network (see h_0 and h_1 in Fig. 1). This metric reflects how well an estimator can predict the superimposed individual treatment and spillover effects on a network. Another evaluation metric that quantifies the quality of extracted individual treatment effect is the Precision in Estimation of Heterogeneous Effect studied in Hill [2011], which is defined as $\epsilon_{PEHE} := \frac{1}{|\mathcal{U}_T|} \sum_{i \in \mathcal{U}_T} (\tau(\mathbf{X}_i) - \hat{\tau}(\mathbf{X}_i))^2$, where $\hat{\tau}(\mathbf{X}_i)$ is defined in Eq. (3).

Baselines Baseline models are domain adaption method Künzel et al. [2019] with gradient boosting regression (**DA GB**), with random forest regression (**DA RF**), doubly-robust estimator Funk et al. [2011] with gradient boosting regression (**DR GB**), and elastic net regression (**DR EN**). They are implemented via EconML Research [2019] with grid-searched hyperparameters. These baselines incorporate the feature vectors as inputs and exposure as the control variable into the model. For random-

Table 3: Experimental results of randomized experiments on the Wave1 and Pokec datasets using nonlinear response generation functions G_1 and G_2 with $\kappa = 0.2$. For Wave1, we set (node degree) $k = 10$, (decay parameter) $\alpha = 0.5$, and (treatment probability) $p = 0.1$, and for Pokec, we set $p = 0.1$. $\hat{HSIC}^{\Phi}$ and $\hat{HSIC}^{GNN}$ are deployed in the GNN-based estimators. Improvements are obtained by comparing with the best baselines.

	Wave1				Pokec			
	G_1		G_2		G_1		G_2	
	$\sqrt{\mathcal{MSE}}$	ϵ_{PEHE}	$\sqrt{\mathcal{MSE}}$	ϵ_{PEHE}	$\sqrt{\mathcal{MSE}}$	ϵ_{PEHE}	$\sqrt{\mathcal{MSE}}$	ϵ_{PEHE}
DA GB	$0.770 \pm .017$	$0.379 \pm .126$	$0.763 \pm .047$	$0.248 \pm .121$	$0.988 \pm .005$	$0.419 \pm .046$	$1.189 \pm .017$	$0.376 \pm .033$
DA RF	$1.047 \pm .046$	$0.701 \pm .029$	$0.977 \pm .021$	$0.599 \pm .193$	$1.016 \pm .024$	$1.075 \pm .031$	$1.225 \pm .009$	$1.016 \pm .037$
DR GB	$0.814 \pm .058$	$0.392 \pm .029$	$0.771 \pm .014$	$0.401 \pm .028$	$0.943 \pm .024$	$0.297 \pm .057$	$1.173 \pm .012$	$0.314 \pm .020$
DR EN	$1.063 \pm .037$	$0.843 \pm .005$	$0.886 \pm .010$	$0.636 \pm .173$	$0.947 \pm .023$	$0.181 \pm .031$	$1.172 \pm .013$	$0.282 \pm .041$
GPS	$0.236 \pm .001$	$0.158 \pm .031$	$0.262 \pm .071$	$0.163 \pm .063$	$0.420 \pm .006$	$0.212 \pm .070$	$0.475 \pm .004$	$0.220 \pm .013$
GCN	$0.192 \pm .003$	$0.050 \pm .007$	$0.201 \pm .034$	$0.044 \pm .026$	$0.367 \pm .005$	$0.162 \pm .004$	$0.423 \pm .017$	$0.183 \pm .010$
GraphSAGE	$\mathbf{0.191 \pm .004}$	$\mathbf{0.049 \pm .003}$	$0.198 \pm .022$	$\mathbf{0.039 \pm .018}$	$\mathbf{0.360 \pm .000}$	$\mathbf{0.146 \pm .001}$	$0.425 \pm .018$	$0.167 \pm .005$
1-GNN	$0.207 \pm .003$	$0.058 \pm .006$	$\mathbf{0.188 \pm .020}$	$0.043 \pm .024$	$0.366 \pm .013$	$0.151 \pm .006$	$\mathbf{0.408 \pm .009}$	$\mathbf{0.158 \pm .004}$
Improve	19.1%	19.0%	28.2%	76.1%	14.3%	19.3%	14.1%	28.2%

ized experiments on Wave1 and Pokec, the predefined treatment probability p is provided, while for the observational studies on the Amazon dataset, the covariate-dependent treatment probability is estimated. Moreover, the generalized propensity score (GPS) method is reproduced and enhanced for a fair comparison, equipped with the same feature map Φ function. More details of baselines, the sketch of the training procedure, and hyperparameters are relegated to Appendix F.

Experiments We use partial outcomes, both in the randomized experiments and observational settings, to train the GNN-based causal estimators. We investigate the effect of penalizing representation imbalance in the observational studies on the Amazon dataset. The entire data points $(\mathbf{X}_i, T_i, G_i, Y_i)$ are randomly divided into training (80%), validation (5%), and test (15%) sets. Note that the entire network $\mathcal{G}$ and the covariates of all units $\mathbf{X}$ are given during the training and test, while only the causal responses of units in the training set are provided in the training phase. For the randomized experiments using the Wave1 and Pokec datasets, we repeat the experiments 3 times and use different random parameters in the response generation process each time.

Experimental results on the Wave1 and Pokec data generated via linear model G_0 are presented in Table 1. Both representation balancing $\hat{HSIC}^{\Phi}$ and $\hat{HSIC}^{GNN}$ are deployed in the GNN-based estimators for searching for the best performance. GNN-based estimators, especially the 1-GNN estimator, are superior for superimposed causal effects prediction. One can observe a 26.1% improvement of the $\sqrt{\mathcal{MSE}}$ metric on the Wave1 dataset when comparing the 1-GNN estimator with the enhanced GPS method and a 8.2% improvement on the Pokec dataset. The covariates of neighboring units in the Pokec dataset actually have strong cosine similarity, hence the improvement on the Pokec dataset is not significant, and the network effect can be approximately captured from the exposure variable. Table 2 shows the experimental results on the pos Amazon dataset in the observational study. In particular, we demonstrate the effects of without representation penalty, and with different penalties. It shows that representation penalties can significantly improve the individual treatment effect recovery, serving as a regularization to avoid over-fitting the network interference. Furthermore, GNN-based estimators using $\hat{HSIC}^{GNN}$ penalty are slightly better than those using $\hat{HSIC}^{\Phi}$ penalty; however, by sacrificing the metric ϵ_{PEHE}.

Table 3 reports the performance of GNN-based causal estimators on nonlinear response models. Nonlinear responses are generated via G_1 and G_2 under $\kappa = 0.2$. For the $\sqrt{\mathcal{MSE}}$ metric, GNN-based estimators outperform the best baseline GPS dramatically, showing the effectiveness of predicting nonlinear causal responses. Moreover, a $19.0\%(G_1)$ and $76.1\%(G_2)$ performance improvement on the ϵ_{PEHE} metric with the Wave1 dataset shows that setting an empty graph, i.e., $\mathcal{G} = \emptyset$, in the GNN-based estimators is an appropriate approach for extracting individual causal effect. Results of nonlinear responses with larger strength parameter $\kappa = 0.5$ are reported in Appendix C and D.

4.3 Results on Improved Intervention Policy

Experiment Settings After obtaining the optimal causal effect estimators and feature map Φ (see Fig. 1), we subsequently optimize intervention policy on the same graph. A simple 2-layer neural network, with ReLU activation between hidden layers and sigmoid activation at the end, is employed as the policy network. The output of the policy network lies in $[0, 1]$, and it is interpreted as the probability of treating a node. The real intervention choice is then sampled from this probability via the Gumbel-softmax trick Jang et al. [2016] such that gradients can be back-propagated. Sampled treatment choices along with corresponding node features are then fed into the feature map Φ and subsequent causal estimators to

Table 4: Intervention policy improvements on the Wave1 and Pokec semi-synthetic datasets under treatment capacity constraint with $p_t = 0.3$. $\Delta\hat{S}(\hat{\pi}_n^{p_t})$ and $\Delta S(\hat{\pi}_n^{p_t})$ represent utility differences evaluated from learned estimators and ground truth, respectively. Note that *only* $\Delta S(\hat{\pi}_n^{p_t})$ reflects the real policy improvement.

	Wave1		Pokec	
	$\Delta\hat{S}(\hat{\pi}_n^{p_t})$	$\Delta S(\hat{\pi}_n^{p_t})$	$\Delta\hat{S}(\hat{\pi}_n^{p_t})$	$\Delta S(\hat{\pi}_n^{p_t})$
DA GB	0.276 ± 0.033	0.002 ± 0.025	0.231 ± 0.051	0.001 ± 0.036
DA RF	0.302 ± 0.029	0.003 ± 0.021	0.198 ± 0.080	0.001 ± 0.057
DR GB	0.322 ± 0.023	0.002 ± 0.019	0.338 ± 0.060	0.002 ± 0.046
DR EN	0.311 ± 0.019	0.001 ± 0.018	0.329 ± 0.028	0.001 ± 0.026
GPS	0.235 ± 0.042	0.004 ± 0.032	0.362 ± 0.069	0.001 ± 0.053
GCN	0.260 ± 0.024	0.163 ± 0.020	0.270 ± 0.007	0.190 ± 0.012
GraphSAGE	0.283 ± 0.031	0.176 ± 0.025	0.376 ± 0.049	0.211 ± 0.034
1-GNN	0.327 ± 0.038	$\mathbf{0.208 \pm 0.026}$	0.377 ± 0.041	$\mathbf{0.225 \pm 0.031}$

Table 5: Intervention policy improvements on the pos Amazon dataset under treatment capacity constraint with $p_t = 0.5$. Only domain adaption methods and GPS are compared since they are the best baseline estimators according to Table 2.

	DA GB	DA RF	GPS	GCN	GraphSAGE	1-GNN
$\Delta\hat{S}(\hat{\pi}_n^{p_t})$	38.9 ± 1.1	84.1 ± 2.3	$\underline{98.6 \pm 10.8}$	80.7 ± 0.9	$\mathbf{86.0 \pm 0.9}$	84.1 ± 1.3

evaluate the utility function under network interference defined in Eq. (4). Each experiment setting is repeated 5 times until convergence. The hyperparameter γ in $\mathcal{L}_{\text{pol}}$ is tuned such that the constraint for the percentage p_t is satisfied within the tolerance ± 0.01. More details of experiment settings and hyperparameters are relegated to Appendix D and E.

To quantify the optimized policy $\hat{\pi}_n^{p_t}$, we evaluate the difference $\Delta\hat{S}(\hat{\pi}_n^{p_t}) := \hat{S}_n^{\tau,\delta}(\hat{\pi}_n^{p_t}) - \hat{S}_n^{\tau,\delta}(\pi_R^{p_t})$, where $\pi_R^{p_t}$ represents a randomized intervention underlying the same capacity constraint. The difference $\Delta\hat{S}(\hat{\pi}_n^{p_t})$ indicates how a learned policy can outperform a randomized policy with the same constraint evaluated via learned causal effect estimators. However, from its definition, it is concerned that the policy improvement $\hat{\pi}_n^{p_t}$ may be very biased, such that any "expected improvement" may come from the inaccurate causal estimators. Hence, for the Wave1 and Pokec datasets, knowing the generating process of treatment and spillover effects, we also compare the actual utility difference $\Delta S(\hat{\pi}_n^{p_t}) := S_n^{\tau,\delta}(\hat{\pi}_n^{p_t}) - S_n^{\tau,\delta}(\pi_R^{p_t})$.

Table 4 displays policy optimization results on the under-treated Wave1 and Pokec simulation datasets, where initially only 10% of nodes are randomly assigned to treatment. It shows that an optimized policy network cannot even outperform a randomized policy in ground truth when the causal estimators perform poorly. Hence, policy networks learned from the utility function with plugged in doubly-robust or domain adaption estimators are not reliable. By contrast, the small difference between genuine utility improvement $\Delta S(\hat{\pi}_n^{p_t})$ and estimated improvement $\Delta\hat{S}(\hat{\pi}_n^{p_t})$ for the GNN-based causal estimators indicates the reliability of the optimized policy. Moreover, comparing the ground-truth utility improvement on GPS and GCN-based estimator shows that the policy network sensitively relies on the accuracy of the employed causal estimator. Furthermore, one might argue that through baseline estimators, a simple policy network cannot adjust its treatment choice according to neighboring nodes' features and responses, unlike through GNN-based estimators. For a fair comparison, in Appendix D, we also provide experimental results using a GNN-based policy network. However, we still cannot observe genuine utility improvements on $\Delta S(\hat{\pi}_n^{p_t})$ when using baseline models as causal estimators.

Next, we conduct experiments for intervention policy learning on the over-treated pos Amazon dataset under treatment capacity constraint. Since we do not have access to the ground truth of the pos Amazon dataset, Table 5 shows the utility difference under treatment capacity constraint with $p_t = 0.5$ evaluated only from learned causal estimators. Although the optimized utility improvement $\Delta\hat{S}(\hat{\pi}_n^{p_t})$ achieves the best result via the GPS causal estimator, it might be unreliable compared to the ground truth. A reliable policy improvement having comparable utility improvement via a GNN-based causal estimator is expected.

5 Conclusion

In this work, we first introduced the task of causal inference under general network interference and proposed causal effect estimators using GNNs of various types. We also defined a novel utility function for policy optimization on interconnected nodes, of which a graph-dependent policy regret bound can be derived theoretically. We conduct experiments on semi-synthetic simulation and real datasets. Experiment results show that GNN-based causal effect estimators, especially GraphSAGE and 1-GNN, with an HSIC distribution discrepancy penalty are supe-

rior in superimposed causal effects prediction, and the individual treatment effect can be recovered reasonably well. Subsequent experiments of intervention policy optimization under capacity constraint further confirms the importance of employing an optimal and reliable causal estimator for policy improvement. In future work, we consider the scenario in which the network structure is only partially observed, or dynamic.

References

Donald B Rubin. Randomization analysis of experimental data: The fisher randomization test comment. *Journal of the American Statistical Association*, 75(371):591–593, 1980.

Jake Bowers, Mark M. Fredrickson, and Costas Panagopoulos. Reasoning about interference between units: A general framework. *Political Analysis*, 21: 97–124, 2013. doi: 10.1093/pan/mps038.

Panos Toulis and Edward Kao. Estimation of causal peer influence effects. In *International conference on machine learning*, pages 1489–1497, 2013.

Charles F Manski. *Identification for prediction and decision*. Harvard University Press, 2009.

Susan Athey and Stefan Wager. Efficient policy learning. *arXiv preprint arXiv:1702.02896*, 2017.

Nathan Kallus. Balanced policy evaluation and learning. In *Advances in Neural Information Processing Systems*, pages 8895–8906, 2018.

Nathan Kallus and Angela Zhou. Confounding-robust policy improvement. In *Advances in Neural Information Processing Systems*, pages 9269–9279, 2018.

Toru Kitagawa and Aleksey Tetenov. Who should be treated? empirical welfare maximization methods for treatment choice. Technical report, Cemmap working paper, 2017.

Michael G. Hudgens and M. Elizabeth Halloran. Toward causal inference with interference. 103(482), 2008.

Eric J Tchetgen Tchetgen and Tyler J VanderWeele. On causal inference in the presence of interference. *Statistical methods in medical research*, 21(1):55–75, 2012.

Lan Liu and Michael G Hudgens. Large sample randomization inference of causal effects in the presence of interference. *Journal of the american statistical association*, 109(505):288–301, 2014.

Peter M Aronow, Cyrus Samii, et al. Estimating average causal effects under general interference, with application to a social network experiment. *The Annals of Applied Statistics*, 11(4):1912–1947, 2017.

Laura Forastiere, Edoardo M Airoldi, and Fabrizia Mealli. Identification and estimation of treatment and interference effects in observational studies on networks. *arXiv preprint arXiv:1609.06245*, 2016.

Elizabeth L Ogburn, Oleg Sofrygin, Ivan Diaz, and Mark J van der Laan. Causal inference for social network data. *arXiv preprint arXiv:1705.08527*, 2017a.

Elizabeth L Ogburn, Tyler J VanderWeele, et al. Vaccines, contagion, and social networks. *The Annals of Applied Statistics*, 11(2):919–948, 2017b.

Davide Viviano. Policy targeting under network interference. *arXiv preprint arXiv:1906.10258*, 2019.

David Arbour, Dan Garant, and David Jensen. Inferring network effects from observational data. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 715–724. ACM, 2016.

Eric J Tchetgen Tchetgen, Isabel Fulcher, and Ilya Shpitser. Auto-g-computation of causal effects on a network. *arXiv preprint arXiv:1709.01577*, 2017.

Elizabeth L Ogburn, Ilya Shpitser, and Youjin Lee. Causal inference, social networks, and chain graphs. *arXiv preprint arXiv:1812.04990*, 2018.

Eli Sherman and Ilya Shpitser. Identification and estimation of causal effects from dependent data. In *Advances in neural information processing systems*, pages 9424–9435, 2018.

Rohit Bhattacharya, Daniel Malinsky, and Ilya Shpitser. Causal inference under interference and network uncertainty. *arXiv preprint arXiv:1907.00221*, 2019.

Donald B Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688, 1974.

Jerzy Splawa-Neyman, Dorota M Dabrowska, and TP Speed. On the application of probability theory to agricultural experiments. essay on principles. section 9. *Statistical Science*, pages 465–472, 1990.

Fredrik Johansson, Uri Shalit, and David Sontag. Learning representations for counterfactual inference. In *International conference on machine learning*, pages 3020–3029, 2016.

Uri Shalit, Fredrik D Johansson, and David Sontag. Estimating individual treatment effect: generalization bounds and algorithms. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 3076–3085. JMLR. org, 2017.

Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.

Will Hamilton, Zhitao Ying, and Jure Leskovec. Inductive representation learning on large graphs. In *Advances in Neural Information Processing Systems*, pages 1024–1034, 2017.

Christopher Morris, Martin Ritzert, Matthias Fey, William L Hamilton, Jan Eric Lenssen, Gaurav Rattan, and Martin Grohe. Weisfeiler and leman go neural: Higher-order graph neural networks. *arXiv preprint arXiv:1810.02244*, 2018.

Kim Chantala and Joyce Tabor. National longitudinal study of adolescent health: Strategies to perform a design-based analysis using the add health data. 1999.

Lubos Takac and Michal Zabovsky. Data analysis in public social networks. In *International Scientific Conference and International Workshop Present Day Trends of Innovations*, volume 1, 2012.

Jure Leskovec, Lada A Adamic, and Bernardo A Huberman. The dynamics of viral marketing. *ACM Transactions on the Web (TWEB)*, 1(1):5, 2007.

Vineeth Rakesh, Ruocheng Guo, Raha Moraffah, Nitin Agarwal, and Huan Liu. Linked causal variational autoencoder for inferring paired spillover effects. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pages 1679–1682. ACM, 2018.

Jennifer L Hill. Bayesian nonparametric modeling for causal inference. *Journal of Computational and Graphical Statistics*, 20(1):217–240, 2011.

Sören R Künzel, Jasjeet S Sekhon, Peter J Bickel, and Bin Yu. Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the National Academy of Sciences*, 116(10):4156–4165, 2019.

Michele Jonsson Funk, Daniel Westreich, Chris Wiesen, Til Stürmer, M Alan Brookhart, and Marie Davidian. Doubly robust estimation of causal effects. *American journal of epidemiology*, 173(7):761–767, 2011.

Microsoft Research. EconML: A Python Package for ML-Based Heterogeneous Treatment Effects Estimation. https://github.com/microsoft/EconML, 2019. Version 0.x.

Eric Jang, Shixiang Gu, and Ben Poole. Categorical reparameterization with gumbel-softmax. *arXiv preprint arXiv:1611.01144*, 2016.