Causal Inference under Networked Interference and Intervention Policy Enhancement: Supplementary Materials

A HSIC

The empirical HSIC using a Gaussian RBF kernel is written as $H\hat{S}IC_{\mathcal{K}_\sigma}$. According to Gretton et al. (2005), given samples $\{\Phi(\mathbf{X}_i), T_i\}_{i=1}^n$, the empirical estimation of HSIC in Gaussian kernel $\mathcal{K}_\sigma$ reads

$$H\hat{S}IC_{\mathcal{K}_\sigma} = \frac{1}{n^2}\sum_{i,j=1}^{n} \mathcal{K}_\sigma(\Phi(\mathbf{X}_i), \Phi(\mathbf{X}_j))\mathcal{K}_\sigma(T_i, T_j)$$
$$+ \frac{1}{n^4}\sum_{i,j,k,l=1}^{n} \mathcal{K}_\sigma(\Phi(\mathbf{X}_i), \Phi(\mathbf{X}_j))\mathcal{K}_\sigma(T_k, T_l) - \frac{2}{n^3}\sum_{i,j,k=1}^{n} \mathcal{K}_\sigma(\Phi(\mathbf{X}_i), \Phi(\mathbf{X}_j))\mathcal{K}_\sigma(T_i, T_k).$$

B Nonparametric Identifiability of Causal Effect

The nonparametric identifiability of expected causal response is guaranteed following Ogburn et al. (2017); Forastiere et al. (2016). For the sake of simplicity, we assume that influences are only from the first-order neighbors. To prove the identifiability, we introduce a variable $V_i := S_{V,i}(\mathbf{X}_{\mathcal{N}_i}, \mathbf{T}_{\mathcal{N}_i})$, where

$$S_{V,i} : \{0,1\}^{|\mathcal{N}_i|} \otimes \boldsymbol{\chi}^{\otimes|\mathcal{N}_i|} \to \mathcal{V}_i,$$

for $i = 1, \ldots, n$, represents the aggregation of neighboring covariates and treatment assignments, e.g., the average of neighboring treatments and the output of a GNN. Following reasonable assumptions are necessary for the nonparametric identifiability.

Assumption 1.
(1) Given summary function $S_{V,i}$, for $i = 1, \ldots, n$, $\forall \mathbf{T}_{\mathcal{N}_i}, \mathbf{T}'_{\mathcal{N}_i}$, $\forall \mathbf{X}_{\mathcal{N}_i}, \mathbf{X}'_{\mathcal{N}_i}$, $\forall \mathbf{T}_{\mathcal{N}_{-i}}, \mathbf{T}'_{\mathcal{N}_{-i}}$, and $\forall \mathbf{X}_{\mathcal{N}_{-i}}, \mathbf{X}'_{\mathcal{N}_{-i}}$, with $S_{V,i}(\mathbf{T}_{\mathcal{N}_i}, \mathbf{X}_{\mathcal{N}_i}) = S_{V,i}(\mathbf{T}'_{\mathcal{N}_i}, \mathbf{X}'_{\mathcal{N}_i})$, then it holds

$$Y_i(T_i, \mathbf{T}_{\mathcal{N}_i}, \mathbf{X}_{\mathcal{N}_i}, \mathbf{T}_{\mathcal{N}_{-i}}, \mathbf{X}_{\mathcal{N}_{-i}}) = Y_i(T_i, \mathbf{T}'_{\mathcal{N}_i}, \mathbf{X}'_{\mathcal{N}_i}, \mathbf{T}'_{\mathcal{N}_{-i}}, \mathbf{X}'_{\mathcal{N}_{-i}}).$$

(2) Unconfoundedness assumption: $Y_i(t_i, v_i) \perp T_i, V_i | \mathbf{X}_i$, $\forall t_i \in \{0,1\}$ and $v_i \in \mathcal{V}_i$, for $i = 1, \ldots, n$.

Hence, the expected response of one unit under network inference can be identified as $\mathbb{E}[Y_i(t_i, v_i)] = \mathbb{E}[Y_i | T_i = t_i, V_i = v_i, \mathbf{X}_i]$, $\forall t_i \in \{0,1\}$, and $v_i \in \mathcal{V}_i$, for $i = 1, \ldots, n$. It is derived by

$$\mathbb{E}[Y_i | T_i = t_i, V_i = v_i, \mathbf{X}_i] \overset{Asm.(1)}{=} \mathbb{E}[Y_i(t_i, v_i) | T_i = t_i, V_i = v_i, \mathbf{X}_i]$$
$$\overset{Asm.(2)}{=} \mathbb{E}[Y_i(t_i, v_i) | \mathbf{X}_i].$$

H1GH52	Do you get enough sleep?	H1ED3	Have you skipped a grade?
H1ED5	Have you repeated a grade?	H1ED7	Have you received an suspension?
H1HS1	Have you had a routine physical examination?	H1HS3	Have you received psychological counseling?
H1WP17B	Played a sport in the past 4 weeks?	H1TO51	Is alcohol easily available in your home?
H1TO53	Is a gun easily available in your home?	H1NB5	Do you feel safe in your neighborhood?
H1EE3	Did you work for pay in the last 4 weeks?	PA57D	Food stamps?
H1DA5	How often do you play sport?	H1DA7	How do you hang out with friends?
H1ED11	Your grade in English or language arts?	H1ED12	Your grade in mathematics?
H1ED13	Your grade in history or social studies?	H1ED14	Your grade in science?
H1DS12	How often did you sell marijuana or other drugs	–	–

Table 1: Selected questions from the Wave1 data Harris and Udry (2018) that are used as feature vectors.

	$k=5$		$k=10$	
	$\sqrt{\mathcal{MSE}}$	ϵ_{PEHE}	$\sqrt{\mathcal{MSE}}$	ϵ_{PEHE}
GPS	0.279 ± 0.071	0.210 ± 0.043	0.281 ± 0.049	0.139 ± 0.052
GCN	0.212 ± 0.035	0.095 ± 0.055	0.211 ± 0.013	0.058 ± 0.036
GraphSAGE	$\mathbf{0.200 \pm 0.032}$	$\mathbf{0.088 \pm 0.054}$	$\mathbf{0.199 \pm 0.030}$	$\mathbf{0.057 \pm 0.039}$
1-GNN	0.214 ± 0.039	0.096 ± 0.062	0.203 ± 0.033	$\mathbf{0.057 \pm 0.040}$
Improve	28.3%	58.1%	29.2%	59.0%

Table 2: Evaluation metrics on under-treated synthetic data with $\mathbf{p} = \mathbf{0.1}$, $\alpha = 0.5$, and $k = 5, 10$. Improvements are obtained by comparing with the GPS baseline. Both representation balancing $\hat{HSIC}^{\Phi}$ and $\hat{HSIC}^{GNN}$ are deployed in the GNN-based estimators for searching for the best performance.

C Synthetic Randomized Experiments on Wave1

On the in-school friendship network derived from the Wave1 questionnaire data, we conduct randomized intervention experiments that simulate the improvement of performance after assigning a student to a tutoring or support program. Recall that $Y_i(T_i = 0, \mathcal{G} = \emptyset)$ indicates the overall performance of student i before assigning it to a tutoring program or being influenced by peers. We select specific questions from the questionnaire and regard the corresponding answers as the features of corresponding students. These feature vectors are further used to construct a symmetrized k-NN similarity graph as the in-school friendship network. Questions related to the potential performance of students are list in Table 1.

Using the answers of selected questions and their abbreviations, $Y_i(T_i = 0, \mathcal{G} = \emptyset)$ is generated as follows

$$\begin{aligned} Y_i(T_i = 0, \mathcal{G} = \emptyset) := & -X_{i,H1GH52} + 2X_{i,H1ED3} - X_{i,H1ED5} - 2X_{i,H1ED7} \\ & - 0.5(X_{i,H1ED11} + X_{i,H1ED12} + X_{i,H1ED13} + X_{i,H1ED14}) \\ & + 0.5(X_{i,H1DA5} + X_{i,H1DA7}) - 3X_{i,H1DS12} + f_{\mathcal{N}}(X_{i,H1HS1} \\ & + X_{i,H1HS3} + X_{i,H1WP17B} + X_{i,H1TO51} + X_{i,H1TO53} \\ & + X_{i,H1NB5} + X_{i,H1EE3} + X_{i,PA57D}), \end{aligned}$$

where $f_{\mathcal{N}}(\cdot)$ represents a 1-layer neural network with random coefficients.

The generating process of the individual treatment response also depends on the selected properties. For example, by assigning a student who has repeated grade will probably improve this student's performance. The treatment effect is simulated as follows:

$$\begin{aligned} \tau(\mathbf{X}_i) := & X_{i,H1ED3} + 0.5(X_{i,H1GH52} + X_{i,H1ED5} + X_{i,H1ED7}) \\ & + 0.5(X_{i,H1ED11} + X_{i,H1ED12} + X_{i,H1ED13} + X_{i,H1ED14}) \\ & + X_{i,H1DS12} + f_{\mathcal{N}}, \end{aligned}$$

where $f_{\mathcal{N}}$ represents a nonlinear random function depending on the rest of variables. Furthermore, peer effect in this synthetic experiment is generated by

$$\delta_i(\mathbf{X}, \mathbf{T}, \mathcal{G}) := \alpha \frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} T_j \tau(\mathbf{X}_j), \tag{1}$$

where the decay parameter α characterizes the decay of influence. Eq. 1 means that the peer effect applied to the node i is determined by individual treatment responses of its neighbors who are under treatment. Finally, the outcome, e.g., the linear response G_0, is simulated by

$$Y_i = Y_i(T_i = 0, \mathcal{G} = \emptyset) + T_i\tau(\mathbf{X}_i) + \delta_i(\mathbf{X}, \mathbf{T}, \mathcal{G}) + \epsilon_{Y_i}. \tag{2}$$

	$k=5$		$k=10$	
	$\sqrt{\mathcal{MSE}}$	ϵ_{PEHE}	$\sqrt{\mathcal{MSE}}$	ϵ_{PEHE}
GPS	0.318 ± 0.010	0.409 ± 0.008	0.363 ± 0.087	0.491 ± 0.200
GCN	0.277 ± 0.007	0.051 ± 0.007	0.288 ± 0.063	0.087 ± 0.053
GraphSAGE	0.276 ± 0.024	$\mathbf{0.050 \pm 0.007}$	0.301 ± 0.054	0.083 ± 0.033
1-GNN	$\mathbf{0.249 \pm 0.006}$	0.054 ± 0.015	$\mathbf{0.278 \pm 0.056}$	$\mathbf{0.076 \pm 0.034}$
Improve	21.7%	87.8%	23.4%	84.5%

Table 3: Evaluation metrics on over-treated synthetic data with $\mathbf{p = 0.7}$, $\alpha = 0.5$, and $k = 5, 10$. Improvements are obtained by comparing with the GPS baseline. Both representation balancing $H\hat{S}IC^{\Phi}$ and $H\hat{S}IC^{GNN}$ are deployed in the GNN-based estimators for searching for the best performance.

	$k=5$		$k=10$	
	$\sqrt{\mathcal{MSE}}$	ϵ_{PEHE}	$\sqrt{\mathcal{MSE}}$	ϵ_{PEHE}
GPS	0.329 ± 0.005	0.207 ± 0.015	0.294 ± 0.008	0.224 ± 0.071
GCN	0.269 ± 0.011	0.047 ± 0.006	0.215 ± 0.020	0.050 ± 0.012
GraphSAGE	0.279 ± 0.015	0.044 ± 0.003	0.223 ± 0.018	0.037 ± 0.011
1-GNN	$\mathbf{0.268 \pm 0.015}$	$\mathbf{0.042 \pm 0.005}$	$\mathbf{0.214 \pm 0.015}$	$\mathbf{0.032 \pm 0.007}$
Improve	18.5%	79.7%	27.2%	85.7%

Table 4: Evaluation metrics on balanced synthetic data with $\mathbf{p = 0.5}$, $\alpha = 0.5$, and $k = 5, 10$. Improvements are obtained by comparing with the GPS baseline. Both representation balancing $H\hat{S}IC^{\Phi}$ and $H\hat{S}IC^{GNN}$ are deployed in the GNN-based estimators for searching for the best performance.

The benefit of using synthetic data is that we can modify the experiment settings. Three parameters control the experimental settings: number of neighbors k, which determines the graph structure and density; the probability p of assigning a node to treatment which controls the population imbalance between treatment and control groups; the decay parameter α, which determines the intensity of peer effect. For the evaluation results reported in the main text we generate the simulation data with parameters $k = 10$, $p = 0.1$, and $\alpha = 0.5$. We report more evaluations in Table 2, Table 3, and Table 4. One observation is that in the randomized experiment setting with linear response, the GraphSAGE-based estimator is a good candidate for causal inference in an under-treated population, while 1-GNN-based estimator is superior in a balanced- or over-treated population.

D Synthetic Randomized Experiments on Pokec

The motivation for using a real social network dataset is that the k-NN similarity graph can violate the power-law degree distribution, as shown in Fig. 1. Consider hypothetical intervention experiments to the users of the Pokec social network. After reading a personalized advertisement or getting influenced by social contacts, a user is encouraged to purchase a new medicine. To simulate the individual buying behavior, we use profile features that are related to the health condition of a user. Table 5 lists the related features used in semi-synthetic experiments.

features	values	features	values
gender	$[0, 1]$	age	$[15, 16, \cdots, 60]$
height	$[140, 141, \cdots, 200]$	weight	$[30, 31, \cdots, 200]$
completed level of education	$[0, 1, 2, 3]$	eyesight	$[0, 1]$
relation to smoking	$[0, 1, 2, 3]$	relation to alcohol	$[0, 1, 2, 3]$
relation to casual sex	$[0, 1, 2]$	–	–

Table 5: Characteristics of users and corresponding ranges of values selected from the Pokec social network data.

We assume that a healthy person with good habits is self-motivated to purchase health medicine even without

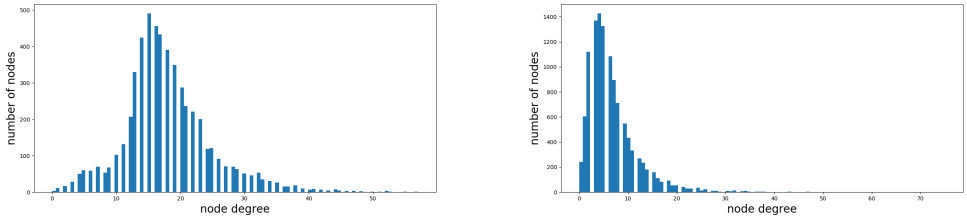

Figure 1: Number of nodes vs. node degree from the k-NN similarity graph of Wave1 with $k = 10$ (left), and from the Pokec social network (right).

external influences. Hence, $Y_i(T_i = 0, \mathcal{G} = \emptyset)$ is simulated as follows:

$$Y_i(T_i = 0, \mathcal{G} = \emptyset) := 0.2(1 - X_{i,gender}) + 0.5X_{i,age} - 0.2X_{i,weight} + 0.5X_{i,education} \\ - 0.6(3 - X_{i,smoke}) + 0.2X_{i,sex} - 0.6(3 - X_{i,alcohol}) + \epsilon,$$

where ϵ is a Gaussian random variable with mean 0.1. Suppose that new health medicine is advertised to offer miraculous effects on weight loss, quit smoking, abstinence, etc. Then the individual treatment response can be generated by

$$\tau(\mathbf{X}_i) := 0.8(1 - X_{i,gender}) + X_{i,age} + 0.3X_{i,weight} + 0.5(1 - X_{i,eyesight}) \\ 0.5(X_{i,education} + 0.5) + 0.6X_{i,smoke} + 0.5X_{i,alcohol} + \epsilon.$$

	$\alpha = 0.1$		$\alpha = 0.9$	
	$\sqrt{\mathcal{MSE}}$	ϵ_{PEHE}	$\sqrt{\mathcal{MSE}}$	ϵ_{PEHE}
GPS	0.263 ± 0.001	0.156 ± 0.017	0.595 ± 0.005	0.185 ± 0.005
GCN	0.230 ± 0.017	0.147 ± 0.031	0.573 ± 0.033	0.163 ± 0.005
GraphSAGE	$\mathbf{0.227 \pm 0.005}$	$\mathbf{0.128 \pm 0.015}$	$\mathbf{0.569 \pm 0.032}$	$\mathbf{0.151 \pm 0.011}$
1-GNN	0.231 ± 0.006	0.132 ± 0.014	0.571 ± 0.033	0.197 ± 0.020
Improve	13.5%	17.9%	4.4%	18.4%

Table 6: Evaluation metrics on under-treated Pokec social network with $\mathbf{p = 0.1}$, $\alpha = 0.1, 0.9$. Improvements are obtained by comparing with the GPS baseline. Both representation balancing $H\hat{S}IC^{\Phi}$ and $H\hat{S}IC^{GNN}$ are deployed in the GNN-based estimators for searching for the best performance.

	$\alpha = 0.1$		$\alpha = 0.9$	
	$\sqrt{\mathcal{MSE}}$	ϵ_{PEHE}	$\sqrt{\mathcal{MSE}}$	ϵ_{PEHE}
GPS	0.404 ± 0.007	0.126 ± 0.004	1.438 ± 0.000	$\mathbf{0.533 \pm 0.015}$
GCN	0.247 ± 0.008	0.044 ± 0.003	1.426 ± 0.030	0.594 ± 0.039
GraphSAGE	0.240 ± 0.006	0.041 ± 0.001	1.417 ± 0.021	0.662 ± 0.061
1-GNN	$\mathbf{0.233 \pm 0.001}$	$\mathbf{0.039 \pm 0.002}$	$\mathbf{1.390 \pm 0.033}$	1.076 ± 0.094
Improve	42.3%	69.0%	3.3%	-11.4%

Table 7: Evaluation metrics on over-treated Pokec social network with $\mathbf{p = 0.7}$, $\alpha = 0.1, 0.9$. Improvements are obtained by comparing with the GPS baseline. Both representation balancing $H\hat{S}IC^{\Phi}$ and $H\hat{S}IC^{GNN}$ are deployed in the GNN-based estimators for searching for the best performance.

Since Pokec is a social network, in the semi-synthetic experiments, we also take into account long-range influences to simulate opinion propagation in the social network. To be more specific, the spillover effect on one node not only depends on the nearest neighboring nodes but also next-nearest neighboring nodes. Formally, it is defined as

$$\delta_i(\mathbf{X}, \mathbf{T}, \mathcal{G}) := \alpha \frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} T_j \tau(\mathbf{X}_j) + \alpha^2 \frac{1}{|\mathcal{N}_i^{(2)}|} \sum_{k \in \mathcal{N}_i^{(2)}} T_k \tau(\mathbf{X}_k), \tag{3}$$

where α is the decay factor and $\mathcal{N}_i^{(2)}$ represents the next-nearest neighbors of i. Finally, the observed data in the randomized experiments can be derived from $Y_i(T_i = 0, \mathcal{G} = \emptyset)$, $\tau(\mathbf{X}_i)$, and social network structure $\mathcal{G}_{Pokec}$ using

Eq. 3 and Eq. 2 for the linear response or Eq. 4 and Eq. 5 for nonlinear responses. The experiments reported in the main text use the setting $\alpha = 0.5$ and $p = 0.1$.

Since the network structure $\mathcal{G}_{Pokec}$ is given, we provide more experiment results in Table 6 and Table 7 to understand the effect of decay parameter α. In particular, we consider regimes from negligible peer effects with $\alpha = 0.1$ to significant peer effects with $\alpha = 0.9$. Since the covariates of neighboring units in the Pokec dataset have strong cosine similarity, and the simulation generation process is relatively simple, GNN-based causal estimators might overfit the superimposed causal effects and poorly recover the individual treatment effect. It is becoming more evident if the peer effects are strong and the population is over-treated, where the GPS baseline can achieve comparable results as other GNN-based estimators using only the information of exposure level (see Table 7).

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

Table 8: Experimental results of randomized experiments on the Wave1 and Pokec dataset using nonlinear response generation functions G_1 and G_2 with $\kappa = \mathbf{0.2}$. For Wave1, we set (node degree) $k = 10$, (decay parameter) $\alpha = 0.5$, and (treatment probability) $p = 0.1$. For Pokec, we set $p = 0.1$. $\hat{HSIC}^{\Phi}$ and $\hat{HSIC}^{GNN}$ are deployed in the GNN-based estimators. $\hat{HSIC}^{\Phi}$ and $\hat{HSIC}^{GNN}$ are deployed in the GNN-based estimators. Improvements are obtained by comparing with the best baselines.

E Experimental Results of Nonlinear Causal Responses

To further investigate the superiority of the GNN-based causal estimators on nonlinear causal responses, we consider the following nonlinear data generation function inspired by Section 4.2 of Toulis and Kao (2013),

$$G_1 : Y_i = Y_i(T_i = 0, \mathcal{G} = \emptyset) + T_i\tau(\mathbf{X}_i) + \delta_i(\mathbf{X}, \mathbf{T}, \mathcal{G}) + \kappa\delta_i^2(\mathbf{X}, \mathbf{T}, \mathcal{G}) + \epsilon_{Y_i}, \tag{4}$$

where κ characterizes the strength of nonlinear effects. In addition, a more complicated nonlinear response generation function

$$G_2 : Y_i = Y_i(T_i = 0, \mathcal{G} = \emptyset) + T_i\tau(\mathbf{X}_i) + \delta_i(\mathbf{X}, \mathbf{T}, \mathcal{G}) + \frac{\kappa}{2}\delta_i^2(\mathbf{X}, \mathbf{T}, \mathcal{G}) + \frac{\kappa}{2}\tau(\mathbf{X}_i)\delta_i(\mathbf{X}, \mathbf{T}, \mathcal{G}) + \epsilon_{Y_i} \tag{5}$$

is considered, where the quadratic terms signify the spillover effect depending on the individual treatment effect.

Table 8 reports the performance of GNN-based causal estimators on nonlinear causal effects prediction tasks. Nonlinear responses are generated via G_1 and G_2 with $\kappa = 0.2$. 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.

	Wave1			
	G_1		G_2	
	$\sqrt{\mathcal{MSE}}$	ϵ_{PEHE}	$\sqrt{\mathcal{MSE}}$	ϵ_{PEHE}
DA GB	$0.742 \pm .083$	$0.210 \pm .008$	$1.060 \pm .047$	$0.400 \pm .054$
DA RF	$1.007 \pm .027$	$0.527 \pm .141$	$1.243 \pm .089$	$1.056 \pm .222$
DR GB	$0.784 \pm .019$	$0.352 \pm .074$	$1.116 \pm .106$	$0.633 \pm .195$
DR EN	$0.882 \pm .053$	$0.575 \pm .015$	$1.258 \pm .176$	$0.841 \pm .293$
GPS	$0.280 \pm .017$	$0.142 \pm .032$	$0.289 \pm .012$	$0.244 \pm .066$
GCN	$0.224 \pm .008$	$\mathbf{0.038 \pm .003}$	$0.237 \pm .020$	$0.095 \pm .010$
GraphSAGE	$\mathbf{0.214 \pm .007}$	$0.045 \pm .002$	$\mathbf{0.231 \pm .014}$	$\mathbf{0.072 \pm .003}$
1-GNN	$0.216 \pm .003$	$0.040 \pm .001$	$0.250 \pm .020$	$0.103 \pm .015$
Improve	23.6%	73.2%	20.1%	70.5%
	Pokec			
DA GB	$1.342 \pm .070$	$0.551 \pm .026$	$2.095 \pm .070$	$0.828 \pm .282$
DA RF	$1.369 \pm .060$	$1.015 \pm .074$	$2.125 \pm .080$	$1.389 \pm .109$
DR GB	$1.324 \pm .081$	$0.306 \pm .011$	$2.038 \pm .090$	$0.438 \pm .005$
DR EN	$1.325 \pm .078$	$0.336 \pm .032$	$2.043 \pm .089$	$0.338 \pm .040$
GPS	$0.693 \pm .058$	$0.450 \pm .042$	$0.813 \pm .068$	$0.375 \pm .089$
GCN	$0.483 \pm .010$	$0.193 \pm .001$	$0.729 \pm .007$	$0.242 \pm .032$
GraphSAGE	$0.480 \pm .009$	$0.198 \pm .004$	$\mathbf{0.713 \pm .017}$	$\mathbf{0.217 \pm .025}$
1-GNN	$\mathbf{0.454 \pm .003}$	$\mathbf{0.159 \pm .005}$	$0.767 \pm .023$	$\mathbf{0.218 \pm .002}$
Improve	34.5%	48.0%	12.3%	35.8%

Table 9: Experimental results of randomized experiments on the Wave1 and Pokec datasets using nonlinear response generation functions G_1 and G_2 with $\kappa = \mathbf{0.5}$. For Wave1, other parameters are set as (node degree) $k = 10$, (decay parameter) $\alpha = 0.5$, and (treatment probability) $p = 0.1$. For Pokec, we set the treatment probability as $p = 0.1$ and the decay parameter as $\alpha = 0.5$. Both representation balancing $\hat{HSIC}^{\Phi}$ and $\hat{HSIC}^{GNN}$ are deployed in the GNN-based estimators for searching for the best performance. Improvements are obtained by comparing with the best baseline.

Table 9 reports the performance of GNN-based causal estimators on the Wave1 and Pokec datasets using nonlinear response models. Nonlinear responses are generated via G_1 and G_2 under $\kappa = 0.5$. For the $\sqrt{\mathcal{MSE}}$ metric, GNN-based estimators outperform the best baseline by 23.6%(G_1) and 20.1%(G_2) on Wave1, and by 34.5%(G_1) and 12.3%(G_2) on the Pokec dataset. Moreover, GNN-based causal estimators significantly outperform the best baseline in the individual treatment effect recovery task. Especially, a 73.2%(G_1) and a 70.5%(G_2) improvement on Wave1 are observed, and a 48.0%(G_1) and a 35.8%(G_2) improvement on Pokec. The significantly improved metric ϵ_{PEHE} indicates that even in the regime with higher nonlinear causal effects, GNN-based causal estimators can disentangle and extract individual treatment effects from strong interference.

F Additional Experiments for Intervention Policy Optimization

In addition to the policy optimization experiments on the Wave1 and Pokec simulation data under the treatment capacity constraint $p_t = 0.3$, in Table 10 we also report the intervention policy improvement under the treatment capacity constraint with $p_t = 0.5$.

Until now, we have only employed a simple neural network as the policy network with feature vectors as input. For GNN-based methods, the policy learner can adjust its treatment rules according to the neighboring nodes' features and responses through the GNN-based causal estimators. However, through baseline estimators, e.g., doubly-robust estimators, a simple policy network cannot access the neighboring features of a node. Therefore, for a fair comparison, we employ another 1-GNN as the policy network, and the evaluations on the Wave1 dataset are given in Table 11. The results further confirm that the accuracy of causal effect estimators is crucial for intervention policy optimization on interconnected units.

	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	0.636 ± 0.028	0.012 ± 0.025	0.479 ± 0.066	0.002 ± 0.055
DA RF	0.644 ± 0.027	0.016 ± 0.023	0.477 ± 0.049	0.008 ± 0.045
DR GB	0.761 ± 0.037	0.003 ± 0.031	0.712 ± 0.133	0.001 ± 0.089
DR EN	0.901 ± 0.150	0.006 ± 0.100	0.708 ± 0.093	0.001 ± 0.078
GPS	0.964 ± 0.091	0.018 ± 0.076	0.841 ± 0.072	0.007 ± 0.060
GCN	0.725 ± 0.015	0.544 ± 0.012	0.747 ± 0.041	0.566 ± 0.035
GraphSAGE	0.712 ± 0.031	0.532 ± 0.024	0.754 ± 0.099	0.559 ± 0.079
1-GNN	0.722 ± 0.052	$\mathbf{0.546 \pm 0.041}$	0.806 ± 0.031	$\mathbf{0.586 \pm 0.023}$

Table 10: Intervention policy improvements on the Wave1 and Pokec semi-synthetic datasets under treatment capacity constraint with $p_t = 0.5$. Note that *only* $\Delta S(\hat{\pi}_n^{p_t})$ reflects the genuine policy improvement.

	Wave1	
	$\Delta\hat{S}(\hat{\pi}_n^{p_t})$	$\Delta S(\hat{\pi}_n^{p_t})$
DA GB	0.291 ± 0.031	0.004 ± 0.026
DA RF	0.310 ± 0.041	0.003 ± 0.032
DR GB	0.102 ± 0.057	0.002 ± 0.048
DR EN	0.360 ± 0.044	0.002 ± 0.037
GPS	0.278 ± 0.061	0.006 ± 0.051
GCN	0.279 ± 0.029	0.179 ± 0.026
GraphSAGE	0.268 ± 0.023	0.169 ± 0.019
1-GNN	0.310 ± 0.022	$\mathbf{0.201 \pm 0.016}$

Table 11: Intervention policy improvements on the Wave1 semi-synthetic dataset under treatment capacity constraint with $p_t = 0.3$. The policy network employed is another 1-GNN. Note that *only* $\Delta S(\hat{\pi}_n^{p_t})$ reflects the real policy improvement.

G Experiment Settings

G.1 GNN-based Estimators in Causal Inference Experiments

For GNN-based estimators, we use Adam as a default optimizer with learning rate 0.001 and weight decay 0.0001. The number of total epochs is $20,000$; early stopping is employed by monitoring the loss on the validation set every 2000 epochs. Hyperparameter κ in $\mathcal{L}_{\text{est}}$ for penalizing the distribution discrepancy is searched from $\{0.001, 0.005, 0.1, 0.2\}$ for the Wave1 and Pokec datasets, and from $\{0.1, 0.2, 0.5, 1.\}$ for the Amazon dataset. The feature map neural network Φ has hidden dimensions $[64, 64]$ for the Wave1 and Pokec datasets, and $[256, 128, 128]$ for the Amazon dataset. GNNs have hidden dimensions $[128, 32]$ for the Wave1 and Pokec datasets, and $[256, 128, 64]$ for the Amazon dataset. Outcome prediction networks h_0 and h_1 have hidden dimensions $[64, 32]$ for the Wave1 and Pokec datasets, and $[256, 128, 64]$ for the Amazon dataset. ReLU is used as the activation function between hidden layers. Dropout is also employed between hidden layers with dropout rate a 0.5.

G.2 Baseline Estimators in Causal Inference Experiments

For baseline models, learning rate of the DR EN model is searched from $\{0.001, 0.01, 0.1\}$ with maximal iteration 10000. For the DA RF model, the number of estimators is searched from $\{5, 10, 20\}$, the maximal depth from $\{5, 10, 20\}$, and the minimum number of samples at a leaf node from $\{5, 10, 20\}$. For the DR GB and DA GB models, the number of estimators is searched from $\{10, 50\}$, and the maximal depth is searched from $\{5, 10\}$. In our experiments, the training procedure of Domain Adaption estimators for causal inference under interference is

given as below

$$\begin{aligned}
\hat{\mu}_0 &= M_1\left(Y_i^0 \sim [\mathbf{X}_i^0; G_i], \text{weights} = \frac{g(\mathbf{X}_i^0)}{1-g(\mathbf{X}_i^0)}\right),\\
\hat{\mu}_1 &= M_2\left(Y_i^1 \sim [\mathbf{X}_i^1; G_i], \text{weights} = \frac{1-g(\mathbf{X}_i^1)}{g(\mathbf{X}_i^1)}\right),\\
\hat{D}_i^1 &= Y_i^1 - \hat{\mu}_0([\mathbf{X}_i^1, G_i]),\\
\hat{D}_i^0 &= \hat{\mu}_1([\mathbf{X}_i^0; G_i]) - Y_i^0,\\
\hat{\tau} &= M_3(\hat{D}_i^0 | \hat{D}_i^1 \sim \mathbf{X}_i^0 | \mathbf{X}_i^1),
\end{aligned}$$

where M_1, M_2, M_3 are machine learning algorithms; $Y_i^0, \mathbf{X}_i^0$ represent the outputs and covariates of units under control in the training dataset, and $Y_i^1, \mathbf{X}_i^1$ under treatment. To capture the interference, the exposure variable G_i is concatenated to the covariates. $g(\mathbf{X}_i)$ is an estimation of $\Pr[T_i = 1|\mathbf{X}_i]$ in the observational study using the Amazon dataset, while it is the predefined treatment probability p in randomized experiments using the Wave1 and Pokec datasets. Similarly, the training procedure of Doubly Robust estimators for causal inference under interference is given as

$$\begin{aligned}
\hat{\mu}_0 &= M_1(Y_i^0 \sim [\mathbf{X}_i^0; G_i]),\\
\hat{\mu}_1 &= M_2(Y_i^1 \sim [\mathbf{X}_i^1; G_i]),\\
\hat{D}_i^1 &= \hat{\mu}_1([\mathbf{X}_i; G_i]) + \frac{Y_i - \hat{\mu}_1([\mathbf{X}_i; G_i])}{g(\mathbf{X}_i)}\mathbb{1}\{T_i = 1\},\\
\hat{D}_i^0 &= \hat{\mu}_0([\mathbf{X}_i; G_i]) + \frac{Y_i - \hat{\mu}_0([\mathbf{X}_i; G_i])}{1-g(\mathbf{X}_i)}\mathbb{1}\{T_i = 0\},\\
\hat{\tau} &= M_3((\hat{D}_i^1 - \hat{D}_i^0) \sim \mathbf{X}_i),
\end{aligned}$$

where M_1, M_2, M_3 are machine learning algorithms; $g(\mathbf{X}_i)$ is an estimation of $\Pr[T_i = 1|\mathbf{X}_i]$ in the observational study using the Amazon dataset, while it is the predefined treatment probability p in randomized experiments using the Wave1 and Pokec datasets.

G.3 Intervention Policy Experiments

Causal estimators with the best performance will be saved and fixed for the subsequent intervention policy improvement experiments on the same dataset. We use Adam as a default optimizer for the policy network with a learning rate of 0.001. The policy network has hidden dimensions $[64, 32]$ for the Wave1 and Pokec datasets, and $[128, 64, 64]$ for the Amazon dataset. ReLU is employed as the activation function between hidden layers, and a sigmoid function is applied to the output. Treatment is then sampled from a Bernoulli distribution using the output of the policy network as the probability. The Gumbel-softmax trick Jang et al. (2017) is employed such that errors can be back-propagated. Hyperparameter γ in $\mathcal{L}_{\text{pol}}$ for enforcing the constraint is chosen from $\{5, 50, 100, 200, 500\}$, such that the pre-defined constraint can be satisfied within the tolerance ± 0.01. Besides, we also penalize the distribution discrepancy under the new intervention policy given by the policy network, and the hyperparameter for penalizing this term is chosen from $\{0.0, 0.0001, 0.001, 0.01, 0.1, 1\}$. The number of training epochs is 2000, and each experiment is repeated 5 times.

H Error Bound of Causal Estimators

In this section we will give a heuristic explanation why the causal estimators are difficult to obtain under interference. We briefly summarize the theoretical result of this section in 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, or sample dependency, becomes unreachable when $d_{\max}(n)$ depends on

the number of units, namely the maximal node degree increases with the number of nodes. 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. This claim will be used as one of the important assumptions for proving the policy regret bounds.

In the following, we will first motivate GNN-based causal estimators and then prove Claim 1 step by step. First, with abuse of notation, we consider the following linear model with deterministic outcome

$$\mu_\star(\mathbf{X}_i, \mathbf{X}, \mathbf{T}, \mathcal{G}) = T_i\tau_\star(\mathbf{X}_i) + \alpha_1 \sum_{j\in\mathcal{N}_i} T_j\tau_\star(\mathbf{X}_j) + \alpha_2 \sum_{k\in\mathcal{N}_i^{(2)}} T_k\tau_\star(\mathbf{X}_k) \tag{6}$$

by setting $Y_i(T_i = 0) = 0$, $\alpha = 1$ and letting $\alpha_1 = \frac{1}{|\mathcal{N}_i|}$, $\alpha_2 = \frac{1}{|\mathcal{N}_i^{(2)}|}$, where $\tau_\star$ stands for the ground truth individual treatment response which is bounded by $||\tau_\star||_\infty \leq M$.

One motivation for employing localized graph convolution network, such as GraphSAGE, is that the surrogate model of a 2-layer GraphSAGE can recover the linear model, especially, when $\mathbf{T} = \mathbf{1}$. To be more specific, consider the following form of a 2-layer GraphSAGE

$$\begin{aligned}
\mathbf{X}_i^{(1)} &= \text{ ReLU } (\mathbf{X}_i + \sum_{j\in\mathcal{N}_i} \mathbf{X}_j\mathbf{W}^{(1)}) \\
\mathbf{X}_i^{(2)} &= \text{ ReLU } (\mathbf{X}_i^{(1)} + \sum_{j\in\mathcal{N}_i} \mathbf{X}_j^{(1)}\mathbf{W}^{(2)}) \\
&= \text{ ReLU } [\text{ ReLU } (\mathbf{X}_i + \sum_{j\in\mathcal{N}_i} \mathbf{X}_j\mathbf{W}^{(1)}) + \sum_{j\in\mathcal{N}_i} \text{ ReLU } (\mathbf{X}_j + \sum_{k\in\mathcal{N}_i^{(2)}} \mathbf{X}_k\mathbf{W}^{(1)})\mathbf{W}^{(2)}].
\end{aligned}$$

A prediction from it reads $o(\mathbf{X}_i) = \mathbf{X}_i^{(2)\intercal}\mathbf{v}$, where $\mathbf{v}$ is a vector mapping the second hidden layer to the outcome prediction. In a *surrogate model* [1], where an identity mapping replaces the ReLU activation function, the model returns the outcome prediction

$$o_{surrogate}(\mathbf{X}_i) = \mathbf{X}_i^\intercal\mathbf{v} + \sum_{j\in\mathcal{N}_i} (\mathbf{X}_j\mathbf{W}^{(1)} + \mathbf{X}_j\mathbf{W}^{(2)})^\intercal\mathbf{v} + \sum_{k\in\mathcal{N}_i^{(2)}} (\mathbf{X}_k\mathbf{W}^{(1)}\mathbf{W}^{(2)})^\intercal\mathbf{v},$$

which correctly recovers the linear model and the simulation protocol of spillover effects when all units are assigned to treatment. Moreover, according to the universal approximation properties of GNNs Scarselli et al. (2008), $\mu_\star$ can be approximated. However, this claim cannot reflect an explicit dependence of estimation error on the graph structure. Hence, motivated by the *surrogate model* and the universal approximation property, we study the following class of functions derived from the universal GNN. Let $\mathcal{T}$ be a class of bounded functions with envelop $M < \infty$ and finite VC-dimension $VC(\mathcal{T}) < \infty$, and let

$$\mathcal{M}_{GNN} := \{\tau_1 + \cdots + \tau_{D_{max}},\ \tau_i \in \mathcal{T} \cup \{0\},\ i = 1, \ldots, D_{max},\ ||\tau_1 + \cdots + \tau_{D_{max}}||_\infty \leq 3M\}, \tag{7}$$

where D_{max} is related to the maximal degree of the graph, for a 2-layer GNN $D_{max} := 1 + d_{max} + d_{max}^2$. Function from $\mathcal{M}_{GNN}$ takes $(\mathbf{X}_i, \mathbf{X}_{j\in\mathcal{N}_i}, \mathbf{X}_{k\in\mathcal{N}_i^{(2)}})_{i=1}^n$ as input [2] and returns outcome prediction. The maximal subscript D_{max} serves as a padding, to fit it, the function class $\mathcal{T}$ is extended to $\mathcal{T} \cup \{0\}$. As an example, one can find a function $\mu_{GNN} \in \mathcal{M}_{GNN}$ which approximates $\mu_\star(\mathbf{X}_i, \mathbf{X}, \mathbf{T}, \mathcal{G})$ as

$$\mu_{GNN}(\mathbf{X}_i, \mathbf{X}, \mathbf{T}, \mathcal{G}) = \tau_0(\mathbf{X}_i) + \sum_{j\in\mathcal{N}_i} \tau_j(\mathbf{X}_j) + \sum_{k\in\mathcal{N}_i^{(2)}} \tau_k(\mathbf{X}_k),$$

where $\tau_0, \tau_j, \tau_k \in \mathcal{T}$, for $j \in \mathcal{N}_i$, $k \in \mathcal{N}_i^{(2)}$. In other words, there exists a function in the class $\mathcal{M}_{GNN}$ which, for every node in the network, only uses the representations of this node, this node's neighbors, and this node's 2-hop neighbors, similar to the surrogate model. Assumptions used in this section are summarized in Assumption 2.

[1] The *surrogate models* of graph convolutional networks are first studied in Zügner et al. (2018) for designing adversarial attacks on GNNs and finding robust nodes.

[2] Note that, treatment assignments can be combined with the covariates and fed into the function. In the experiments, we fed $T_i\mathbf{X}_i$ into the GNNs, meaning that only covariates of treated units are non-zero.

Assumption 2.
(A1) Outcome simulation under interference follows the protocol given in Eq. 6 with $||\mu_\star||_\infty \leq 3M$ *due to the requirement* $||\tau_\star||_\infty \leq M$.
(A2) Outcome prediction model is drawn from $\mathcal{M}_{GNN}$ *defined in Eq. 7.*
(A3) There are no isolated nodes in the network [3].

Define the best approximation realized by the class $\mathcal{M}_{GNN}$ as

$$\tilde{\mu}_{GNN} := \text{argmin}_{\mu \in \mathcal{M}_{GNN}} ||\mu - \mu_\star||_\infty,$$

and the approximation error

$$\epsilon_{GNN} := ||\tilde{\mu}_{GNN} - \mu_\star||_\infty. \tag{8}$$

Moreover, define the optimal empirical estimator as

$$\hat{\mu}_{GNN} := \text{argmin}_{\mu \in \mathcal{M}_{GNN}} \sum_{i=1}^{n} \ell(\mu(\mathbf{X}_i, \mathbf{X}, \mathbf{T}, \mathcal{G}), Y_i).$$

Since both $\tilde{\mu}_{GNN}$ and $\hat{\mu}_{GNN}$ belong to the same class $\mathcal{M}_{GNN}$, it is easy to see

$$\mathbb{E}_n[\ell(\tilde{\mu}_{GNN}(\mathbf{X}_i), Y_i)] \geq \mathbb{E}_n[\ell(\hat{\mu}_{GNN}(\mathbf{X}_i), Y_i)],$$

where we write $\tilde{\mu}_{GNN}(\mathbf{X}_i)$ and $\hat{\mu}_{GNN}(\mathbf{X}_i)$ for the sake of simplicity.

We can decompose the approximation error of the empirical causal estimator using the following fact

$$\begin{aligned}
&\mathbb{E}[\ell(\hat{\mu}_{GNN}(\mathbf{X}_i), Y_i) - \ell(\mu_\star(\mathbf{X}_i), Y_i)] \\
&\quad = \mathbb{E}\mathbb{E}_{Y_i}[\hat{\mu}^2_{GNN}(\mathbf{X}_i) - 2Y_i\hat{\mu}_{GNN}(\mathbf{X}_i) + 2Y_i\mu_\star(\mathbf{X}_i) - \mu^2_\star(\mathbf{X}_i)] \\
&\quad = \mathbb{E}[\hat{\mu}^2_{GNN}(\mathbf{X}_i) - 2\mu_\star(\mathbf{X}_i)\hat{\mu}_{GNN}(\mathbf{X}_i) + \mu^2_\star(\mathbf{X}_i)] \\
&\quad = \mathbb{E}[(\hat{\mu}_{GNN}(\mathbf{X}_i) - \mu_\star(\mathbf{X}_i))^2].
\end{aligned}$$

It then yields

$$\begin{aligned}
\mathbb{E}[(\hat{\mu}_{GNN}(\mathbf{X}_i) - \mu_\star(\mathbf{X}_i))^2] &= \mathbb{E}[\ell(\hat{\mu}_{GNN}(\mathbf{X}_i), Y_i) - \ell(\mu_\star(\mathbf{X}_i), Y_i)] \\
&\leq \mathbb{E}[\ell(\hat{\mu}_{GNN}(\mathbf{X}_i), Y_i) - \ell(\mu_\star(\mathbf{X}_i), Y_i)] \\
&\quad - \mathbb{E}_n[\ell(\hat{\mu}_{GNN}(\mathbf{X}_i), Y_i)] + \mathbb{E}_n[\ell(\tilde{\mu}_{GNN}(\mathbf{X}_i), Y_i)] \\
&= \underbrace{(\mathbb{E} - \mathbb{E}_n)[\ell(\hat{\mu}_{GNN}(\mathbf{X}_i), Y_i) - \ell(\mu_\star(\mathbf{X}_i), Y_i)]}_{\text{(I)}} \\
&\quad + \underbrace{\mathbb{E}_n[\ell(\tilde{\mu}_{GNN}(\mathbf{X}_i), Y_i) - \ell(\mu_\star(\mathbf{X}_i), Y_i)]}_{\text{(II)}}.
\end{aligned}$$

The second term (II) can be bounded by applying the Bernstein inequality. The following inequality holds with probability at least $1 - \mathrm{e}^{-\gamma}$

$$\begin{aligned}
\text{(II)} &\leq \mathbb{E}[\ell(\tilde{\mu}_{GNN}(\mathbf{X}_i), Y_i) - \ell(\mu_\star(\mathbf{X}_i), Y_i)] + \sqrt{\frac{2C^2_\ell ||\tilde{\mu}_{GNN} - \mu_\star||^2_\infty \gamma}{n}} \\
&\quad + \frac{2C_\ell ||\tilde{\mu}_{GNN} - \mu_\star||_\infty \gamma}{3n} \\
&= \mathbb{E}[(\tilde{\mu}_{GNN}(\mathbf{X}_i) - \mu_\star(\mathbf{X}_i))^2] + \sqrt{\frac{2C^2_\ell \epsilon^2_{GNN} \gamma}{n}} + \frac{2C_\ell \epsilon_{GNN} \gamma}{3n} \\
&\leq \epsilon^2_{GNN} + \epsilon_{GNN}\sqrt{\frac{2C^2_\ell \gamma}{n}} + \frac{4C_\ell M \gamma}{n}
\end{aligned} \tag{9}$$

[3] This assumption will be used later

using the facts $||\ell(\tilde{\mu}_{GNN}(\mathbf{X}_i), Y_i) - \ell(\mu_\star(\mathbf{X}_i), Y_i)||_\infty \leq C_\ell ||\tilde{\mu}_{GNN} - \mu_\star||_\infty$ and $\epsilon_{GNN} := ||\tilde{\mu}_{GNN} - \mu_\star||_\infty \leq 6M$ where C_ℓ represents the finite Lipschitz constant of loss function.

Furthermore, the first term (I), the maximal deviation between empirical and true means, can be bounded using the standard symmetrization method (see Theorem 2.1 in Bartlett et al. (2005)). Consider a class of functions $\mathcal{F}$, for any $f \in \mathcal{F}$, assume that $||f||_\infty \leq F$ and $\mathbb{V}[f] \leq V$. Then for every $\gamma > 0$, with probability at least $1 - \mathrm{e}^{-\gamma}$

$$\sup_{f \in \mathcal{F}} (\mathbb{E}[f] - \mathbb{E}_n[f]) \leq \inf_{\alpha > 0} \left(2(1+\alpha)\mathcal{R}_n\mathcal{F} + \sqrt{\frac{2V\gamma}{n}} + 2F(\frac{1}{3} + \frac{1}{\alpha})\frac{\gamma}{n} \right),$$

where $\mathcal{R}_n\mathcal{F}$ indicates the Rademacher complexity of $\mathcal{F}$. Hence, it gives

$$\text{(I)} \leq 4\mathcal{R}_n\{\ell(\mu) - \ell(\mu_\star) : \mu \in \mathcal{M}_{GNN}\} + 6\sqrt{\frac{2C_\ell^2 M^2 \gamma}{n}} + \frac{16 C_\ell M \gamma}{n} \tag{10}$$

by setting $\alpha = 1$ and using $||\ell(\mu(\mathbf{X}_i), Y_i) - \ell(\mu_\star(\mathbf{X}_i), Y_i)||_\infty \leq C_\ell ||\mu - \mu_\star||_\infty \leq 6C_\ell M$, $\mathbb{V}[\ell(\mu(\mathbf{X}_i), Y_i) - \ell(\mu_\star(\mathbf{X}_i), Y_i)] \leq \mathbb{E}[(\ell(\mu(\mathbf{X}_i), Y_i) - \ell(\mu_\star(\mathbf{X}_i), Y_i))^2] \leq 36 C_\ell^2 M^2$ for any $\mu \in \mathcal{M}_{GNN}$. Moreover, the Rademacher complexity term is defined as

$$\begin{aligned}
&\mathcal{R}_n\{\ell(\mu) - \ell(\mu_\star) : \mu \in \mathcal{M}_{GNN}\} \\
&\quad := \mathbb{E}_\sigma \left[\sup_{\mu \in \mathcal{M}_{GNN}} |\frac{1}{n} \sum_{i=1}^n \sigma_i (\ell(\mu(\mathbf{X}_i), Y_i) - \ell(\mu_\star(\mathbf{X}_i), Y_i))| \,\middle|\, \mathbf{X}, \mathbf{T}, \mathcal{G} \right] \\
&\quad \leq \underbrace{C_\ell \mathbb{E}_\sigma \left[\sup_{\mu \in \mathcal{M}_{GNN}} |\frac{1}{n} \sum_{i=1}^n \sigma_i (\mu(\mathbf{X}_i) - \mu_\star(\mathbf{X}_i))| \,\middle|\, \mathbf{X}, \mathbf{T}, \mathcal{G} \right]}_{(\#)},
\end{aligned} \tag{11}$$

where $\{\sigma_i\}_{i=1}^n$ are Rademacher random variables. Before using the covering number arguments to further bound the Rademacher complexity term we introduce the following lemmas.

Lemma 1 (Theorem 29.6 in Devroye et al. (2013))**.** *Let $\mathcal{F}_1, \ldots, \mathcal{F}_k$ be classes of real functions on $\mathbb{R}^d$. For n arbitrary fixed points $z_1^n = (z_1, \ldots, z_n)$ in $\mathbb{R}^d$, define the sets $\mathcal{F}_1(z_1^n), \ldots, \mathcal{F}_k(z_1^n)$ by $\mathcal{F}_j(z_1^n) = \{f_j(z_1), \ldots, f_j(z_n) : f_j \in \mathcal{F}_j\}$, $j = 1 \ldots, k$. Also introduce $\mathcal{F} = \{f_1 + \cdots + f_k : f_j \in \mathcal{F}_j,\ j = 1, \ldots, k\}$. Then for every $\epsilon > 0$ and z_1^n,*

$$\mathcal{N}_1(\epsilon, \mathcal{F}(z_1^n)) \leq \prod_{j=1}^k \mathcal{N}_1(\epsilon/k, \mathcal{F}_j(z_1^n)). \tag{12}$$

Lemma 2. *Let $\mathcal{F}_1, \ldots, \mathcal{F}_k$ be classes of bounded real functions on $\mathbb{R}^d$ with envelop F and finite VC-dimension $v < \infty$, for $3 \leq k \leq K$. Also introduce $\mathcal{F} = \{f_1 + \cdots + f_k, f_j \in \mathcal{F}_j,\ j = 1, \ldots, k\}$ and let $\mathcal{F}(z_1^n) = \{f(z_1), \ldots, f(z_n), f \in \mathcal{F}\}$ for arbitrary fixed points z_1^n in $\mathbb{R}^d$. Then we have the following bound*

$$\mathbb{E}_\sigma[\sup_{f \in \mathcal{F}} |\frac{1}{n} \sum_{i=1}^n \sigma_i f(z_i)|] \leq C_F \sqrt{\frac{kv \ln k}{n}}, \tag{13}$$

where C_F is a constant which depends only on the envelop.

Proof. According to the Theorem 5.22 in Wainwright (2019), the Rademacher complexity term is bounded as

$$\mathbb{E}_\sigma[\sup_{f \in \mathcal{F}} |\frac{1}{n} \sum_{i=1}^n \sigma_i f(z_i)|] \leq \underbrace{\frac{32}{\sqrt{n}} \int_0^{2F} \sqrt{\ln \mathcal{N}_1(\epsilon, \mathcal{F}(z_1^n))}\, d\epsilon}_{(\star)}.$$

Using Lemma 1 and $\mathcal{N}_1(\epsilon, \mathcal{F}) \leq \mathcal{N}_2(\epsilon, \mathcal{F})$, it gives

$$(\star) \leq \frac{32}{\sqrt{n}} \int_0^{2F} \sqrt{\sum_{j=1}^k \ln \mathcal{N}_2(\epsilon/k, \mathcal{F}_j(z_1^n))}.$$

Moreover, a uniform entropy bound for the covering number is given by the Theorem 2.6.7 in Van Der Vaart and Wellner (1996). A small modification gives

$$\mathcal{N}_2(\epsilon, \mathcal{F}_j(z_1^n)) \leq C(v+1)(16e)^{(v+1)}(k/\epsilon)^{2v}, j = 1, \ldots, k,$$

where C is a universal constant. Furthermore, following the same technique used by Eq. A.6 in Kitagawa and Tetenov (2018), we obtain

$$\begin{aligned}
(\star) &\leq \frac{32}{\sqrt{n}}\sqrt{k}\int_0^{2F}\sqrt{\ln C + \ln(v+1) + (v+1)\ln(16e) + 2v\ln k - 2v\ln\epsilon}\, d\epsilon \\
&\overset{(1)}{\leq} \frac{32}{\sqrt{n}}\sqrt{kv}\int_0^{2F}\sqrt{\ln C + \ln 2 + \ln(16e) + 2\ln k - 2\ln\epsilon}\, d\epsilon \\
&\overset{(2)}{\leq} \frac{32}{\sqrt{n}}\sqrt{kv\ln k}\int_0^{2F}\sqrt{\ln C + \ln 2 + \ln(16e) + 2 - 2\ln\epsilon/\ln K}\, d\epsilon := C_F\sqrt{\frac{kv\ln k}{n}},
\end{aligned}$$

where (1) uses the fact that usually v is large enough and (2) is due to the condition $3 \leq k \leq K$. ∎

Now, we can further bound the term $\mathcal{R}_n\{\ell(\mu) - \ell(\mu_\star) : \mu \in \mathcal{M}_{GNN}\}$ after Eq. 11. Note that

$$\begin{aligned}
(\#) = C_\ell \mathbb{E}_\sigma \Bigg[\sup_{f\in\mathcal{M}_{GNN}} \Bigg| \frac{1}{n}\sum_{i=1}^n \sigma_i [(\tau_0(\mathbf{X}_i) - T_i\tau_\star(\mathbf{X}_i)) + \sum_{j\in\mathcal{N}_i}(\tau_j(\mathbf{X}_j) - T_j\tau_\star(\mathbf{X}_j)) \\
+ \sum_{k\in\mathcal{N}_i^{(2)}}(\tau_k(\mathbf{X}_k) - T_k\tau_\star(\mathbf{X}_k))] \Bigg| \Bigg]. \qquad (14)
\end{aligned}$$

Define a new constant for each node $D_i := 1 + |\mathcal{N}_i| + |\mathcal{N}_i^{(2)}|$, $i = 1, \ldots, n$. According to (A3) in Assumption 2, we have $D_i \geq 3$. Also introduce a new class of function $\Omega := \{\mathcal{T} \pm \tau_\star\}$. Note that class Ω has the same VC-dimension as $\mathcal{T}$, i.e., $VC(\Omega) = VC(\mathcal{T})$, and $||\omega||_\infty \leq 2M$ for any $\omega \in \Omega$. Recall the definition of $D_{max} := 1 + d_{max} + d_{max}^2$. By decomposing the node subscript i into groups with the same D_i, Eq. 14 can be further written as

$$\begin{aligned}
(\#) &= C_\ell \mathbb{E}_\sigma \left[\sup_{f\in\mathcal{M}_{GNN}} \left| \frac{1}{n}\sum_{i=1}^n \sigma_i \sum_{l=1}^{D_i} \omega_l(\mathbf{X}_i, \mathbf{X}, \mathbf{T}, \mathcal{G}) \right| \right] \quad \omega_l \in \Omega \\
&= C_\ell \mathbb{E}_\sigma \left[\sup_{f\in\mathcal{M}_{GNN}} \left| \sum_{k=3}^{D_{max}} \frac{1}{n} \sum_{i:D_i=k} \sigma_i \sum_{l=1}^{k} \omega_l(\mathbf{X}_i, \mathbf{X}, \mathbf{T}, \mathcal{G}) \right| \right] \\
&\overset{(1)}{\leq} C_\ell \sum_{k=3}^{D_{max}} \mathbb{E}_\sigma \left[\sup_{f\in\mathcal{M}_{GNN}} \left| \frac{1}{n} \sum_{i:D_i=k} \sigma_i \sum_{l=1}^{k} \omega_l(\mathbf{X}_i, \mathbf{X}, \mathbf{T}, \mathcal{G}) \right| \right] \\
&\overset{(2)}{\leq} C_\ell C_F \sum_{k=3}^{D_{max}} \frac{1}{n}\sqrt{|i : D_i = k| k VC(\mathcal{T})\ln k} \leq C_\ell C_F \sum_{k=3}^{D_{max}} \sqrt{\frac{kVC(\mathcal{T})\ln k}{n}} \\
&\leq C_\ell C_F \sqrt{\frac{D_{max}^3 VC(\mathcal{T})\ln D_{max}}{n}},
\end{aligned}$$

where (1) uses the triangle inequality and (2) uses Lemma 2. Hence, the (I) term is bounded by

$$\text{(I)} \leq 4C_\ell C_F \sqrt{\frac{D_{max}^3 VC(\mathcal{T})\ln D_{max}}{n}} + 6\sqrt{\frac{2C_\ell^2 M^2\gamma}{n}} + \frac{16C_\ell M\gamma}{n}.$$

By combining (I) and (II) we have the following theorem.

Theorem 1. *Suppose Assumption 2 holds. Let $\hat{\mu}_{GNN}$ be the optimal causal estimator obtained by minimizing an empirical loss function using the data $\{\mathbf{X}_i, \mathbf{X}, \mathbf{T}, \mathcal{G}\}_{i=1}^n$. Suppose that the loss function has a finite Lipschitz*

constant C_ℓ and $\hat{\mu}_{GNN}$ is restricted to $\mathcal{M}_{GNN}$, Then with probability at least $1-2\mathrm{e}^{-\gamma}$, the causal estimator under interference has an error bound

$$\mathbb{E}[(\hat{\mu}_{GNN}(\mathbf{X}_i)-\mu_\star(\mathbf{X}_i))^2] \leq 4C_\ell C_F\sqrt{\frac{D_{max}^3 VC(\mathcal{T})\ln D_{max}}{n}}+6\sqrt{\frac{2C_\ell^2M^2\gamma}{n}} + \epsilon_{GNN}\sqrt{\frac{2C_\ell^2\gamma}{n}}+\frac{20C_\ell M\gamma}{n}+\epsilon_{GNN}^2, \tag{15}$$

where ϵ_{GNN} is defined in Eq. 8.

Keeping only the leading term with D_{max}, under network interference, the causal estimator has an error bound $\mathcal{O}(\sqrt{\frac{D_{max}^3\ln D_{max}}{n}})$. It indicates that an accurate causal estimator is difficult to obtain under large network interference. Recall that the prediction outcome from the GNN causal estimator is actually the superposition of individual treatment effect and spillover effect. Hence, it is expected that, similarly, the individual treatment effect becomes more and more difficult to recover under more substantial network interference. This intuitive expectation can be observed in the following experimental results in Table 12. We observe that the error of individual treatment effect estimator increases from $k=1$ to $k=4$.

	$k=1$	$k=2$	$k=4$
GraphSAGE	0.048	0.129	0.152

Table 12: ϵ_{PEHE} on the semi-synthetic Wave1 data with $p=0.1$, $\alpha=0.5$, and $k=1,2,4$. To fit the theoretical analysis, exposure level is not fed into the model.

I Policy Regret Bound

In this section, we provide a regret bound for the intervention policy that employs GNN-based causal estimators. The policy regret bound is first summarized in the following theorem.

Theorem 2. *By Assumption 3, for any small $\epsilon>0$, the policy regret is bounded by $\mathcal{R}(\hat{\pi}_n)\leq 2\left(\frac{\alpha_\tau}{n^{\zeta_\tau}}+\frac{\alpha_\delta}{n^{\zeta_\delta}}\right)+2\epsilon$ 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 $\mathcal{G}$.

Suppose that the policy functional class Π is finite and its capacity is bounded by $|\Pi|$. According to Theorem 2, with probability at least $1-\delta$, the policy regret is bounded by

$$\begin{aligned}\mathcal{R}(\hat{\pi}_n) &\leq 2\left(\frac{\alpha_\tau}{n^{\zeta_\tau}}+\frac{\alpha_\delta}{n^{\zeta_\delta}}\right)+8(M_1+M_2)\sqrt{\frac{2(d_{\max}^2+1)}{n}\log\frac{|\Pi|}{\delta}}\\ &\approx 2\left(\frac{\alpha_\tau}{n^{\zeta_\tau}}+\frac{\alpha_\delta}{n^{\zeta_\delta}}\right)+8d_{\max}(M_1+M_2)\sqrt{\frac{2}{n}\log\frac{|\Pi|}{\delta}}\end{aligned}$$

It indicates that optimal policies are more difficult to find in a dense graph even under weak interactions between neighboring nodes.

Throughout the estimation of policy regret, we maintain the following assumptions.

Assumption 3.
(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

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

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* $||\pi_1 - \pi_2||_\infty := \sup_{\mathbf{X}\in\boldsymbol{\mathcal{X}}} |\pi_1(\mathbf{X}) - \pi_2(\mathbf{X})|$.
(ES) Uniformly consistency: after fitting experimental or observational data on $\mathcal{G}$*, individual treatment effect estimator satisfies*

$$\frac{1}{n}\sum_{i=1}^{n} |\tau_i - \hat{\tau}_i| < \frac{\alpha_\tau}{n^{\zeta_\tau}},$$

and spillover estimator satisfies

$$\forall \pi \in \Pi,\ \frac{1}{n}\sum_{i=1}^{n} |\delta_i(\pi) - \hat{\delta}_i(\pi)| < \frac{\alpha_\delta}{n^{\zeta_\delta}} \tag{16}$$

where $\alpha_\tau > 0$ *and* $\alpha_\delta > 0$ *are scaling factors that characterize the errors of estimators.* ζ_τ *and* ζ_δ *control the convergence rate of estimators for individual treatment effect and spillover effect, respectively, which satisfy* $0 < \zeta_\tau, \zeta_\delta < 1$.

Before proving Theorem 2 step by step, we first discuss the plausibility of Assumption 3. 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 H, we have studied a particular class of GNNs, which is inspired by the *surrogate model* of nonlinear graph neural networks and derived Claim 1. Claim 1 indicates that the $\frac{1}{\sqrt{n}}$ error bound of GNN-based causal estimators might be unreachable when $d_{\max}(n)$ depends on the number of units. Therefore, in the (ES) assumption, we assume the coefficients ζ_τ and ζ_δ to characterize the convergence rates, which is line with the assumption made in Athey and Wager (2017) (see Assumption 2 of Athey and Wager (2017)).

Besides, (LIP) assumes that the change of received spillover effect is bounded after modifying the treatment assignments of one unit's neighbors. This assumption is plausible, at least, in the synthetic experiments. For instance, consider the spillover effect in the simulated experiments generated by $\delta_i(\pi) = \alpha\frac{1}{|\mathcal{N}_i|}\sum_{j\in\mathcal{N}_i} \pi(\mathbf{X}_j)\tau(\mathbf{X}_j)$ (see Eq. 1), then we can see

$$|\delta_i(\pi_1) - \delta_i(\pi_2)| \leq \alpha\frac{1}{|\mathcal{N}_i|}\sum_{j\in\mathcal{N}_i} M_1|\pi_1(\mathbf{X}_j) - \pi_2(\mathbf{X}_j)| \leq \alpha M_1||\pi_1 - \pi_2||_\infty.$$

Hence, in this example $L = \alpha M_1$.

The underlying difficulty of estimating the intervention policy regret is the networked setting. Weak independence assumption (WI) allows us to use hypergraph-based method and derive concentration inequalities for the networked random variables. 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 the regret bound 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 2 needs to be modified to $d_{\max}^4$. This change remains the same for the policy regret bound under capacity constraint, which will be provided in Theorem 4.

The flow of the proof for Theorem 2 can be summarized as: Under (WI) and (BO), we use concentration inequalities of networked random variables defined on a hypergraph, which is derived from graph $\mathcal{G}$ to bound the convergence rate. Besides, using the Lipschitz assumption (LIP) enables us to estimate the covering number of the policy functional class Π.

Concentration inequalities on partly dependent random variables are first given in Janson (2004). Later, Wang et al. (2017) provides tighter concentration inequalities using hypergraph and weak dependence assumption. A hypergraph is a generalization of graph in which a hyperedge groups a number of vertices in the graph. For instance, consider a graph with n vertices, and let $\mathcal{N} = \{v_1, v_2, \ldots, v_n\}$ represent the set of vertices. Hyperedges set $\mathcal{E}_h = \{e_{h,1}, e_{h,2} \cdots, e_{h,m}\}$ represents instances joining a number of vertices. In the following, let $\mathcal{G}_h = (\mathcal{N}, \mathcal{E}_h)$ denote a hypergraph.

Definition 1 (Definition 1 in Wang et al. (2017)). *Given a hypergraph $\mathcal{G}_h$, we call $\{\xi_i\}_{i=1}^n$ $\mathcal{G}_h$-networked random variables if there exist functions $f_i : \boldsymbol{\chi}^{\otimes|e_{h,i}|} \to \mathbb{R}$ such that $\xi_i = f_i(\{\mathbf{X}_v | v \in e_{h,i}\})$, where $\{\mathbf{X}_v | v \in e_{h,i}\}$ represents the set of covariates of the vertices in the hyperedge $e_{h,i}$.*

Furthermore, we have the following concentration inequality.

Theorem 3 (Corollary 7 in Wang et al. (2017)). *Let $\{\xi_i\}_{i=1}^n$ be $\mathcal{G}_h$-networked random variables with mean $\mathbb{E}[\xi_i] = \mu$, and satisfying $a < \xi_i < b$, $\forall i \in \{1, 2, \ldots, n\}$. Then for all $\epsilon > 0$,*

$$\Pr\left(\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i - \mu\right| \geq \epsilon\right) \leq \exp\left(-\frac{n\epsilon^2}{2\omega_{\mathcal{G}_h}(b-a)^2}\right), \tag{17}$$

where $\omega_{\mathcal{G}_h} := \max_{v\in\mathcal{N}} |\{e_h : v \in e_h\}|$ represents the maximal degree of $\mathcal{G}_h$.

Recall the following definitions of utility functions $S_n^{\tau,\delta}(\pi)$, $\hat{S}_n^{\tau,\delta}(\pi)$, and $S(\pi)$

$$\begin{aligned}
S(\pi) &:= \mathbb{E}[(2\pi(\mathbf{X}_i) - 1)(\tau_i + \delta_i(\pi))] \\
S_n^{\tau,\delta}(\pi) &:= \frac{1}{n}\sum_{i=1}^{n}(2\pi(\mathbf{X}_i) - 1)(\tau_i + \delta_i(\pi)) \\
\hat{S}_n^{\tau,\delta}(\pi) &:= \frac{1}{n}\sum_{i=1}^{n}(2\pi(\mathbf{X}_i) - 1)(\hat{\tau}_i + \hat{\delta}_i(\pi)),
\end{aligned}$$

where the policy π function has output in $[0, 1]$. An optimal empirical policy is obtained via $\hat{\pi}_n \in \operatorname{argmax}_{\pi\in\Pi}\hat{S}_n^{\tau,\delta}(\pi)$. Note that in the definition of $S(\pi)$ we still keep the subindex i to emphasize the dependence of spillover effect on neighboring nodes. Next we provide several lemmas related to the utility functions.

Lemma 3. *Let $\mathcal{S}(\pi) := S_n^{\tau,\delta}(\pi) - S(\pi)$, for any $\pi_1, \pi_2 \in \Pi$, where the policy class in contained in $[0, 1]$, according to the assumptions (BO) and (LIP) we have*

$$|\mathcal{S}(\pi_1) - \mathcal{S}(\pi_2)| \leq 2(2M_1 + 2M_2 + L)||\pi_1 - \pi_2||_\infty$$

Proof. First note that $|\mathcal{S}(\pi_1) - \mathcal{S}(\pi_2)| \leq |S(\pi_1) - S(\pi_2)| + |S_n^{\tau,\delta}(\pi_1) - S_n^{\tau,\delta}(\pi_2)|$, and we have

$$\begin{aligned}
&|S(\pi_1) - S(\pi_2)| = |\int_{\boldsymbol{\chi}} (2\pi_1(\mathbf{X}_i) - 1)(\tau_i + \delta_i(\pi_1)) - (2\pi_2(\mathbf{X}_i) - 1)(\tau_i + \delta_i(\pi_2))\ d\mathbf{X}_i| \\
&\leq \int_{\boldsymbol{\chi}} 2|\tau_i|||\pi_1 - \pi_2||_\infty + |(2\pi_1(\mathbf{X}_i) - 1)(\delta_i(\pi_2) + L||\pi_1 - \pi_2||_\infty) - (2\pi_2(\mathbf{X}_i) - 1)\delta_i(\pi_2)|\ d\mathbf{X}_i \\
&= \int_{\boldsymbol{\chi}} 2|\tau_i|||\pi_1 - \pi_2||_\infty + |2(\pi_1(\mathbf{X}_i) - \pi_2(\mathbf{X}_i))\delta_i(\pi_2) + L(2\pi_1(\mathbf{X}_i) - 1)||\pi_1 - \pi_2||_\infty|\ d\mathbf{X}_i \\
&\leq (2|\tau_i| + 2|\delta_i(\pi_2)| + L)||\pi_1 - \pi_2||_\infty \\
&\leq (2M_1 + 2M_2 + L)||\pi_1 - \pi_2||_\infty.
\end{aligned}$$

Similarly, we have $|S_n^{\tau,\delta}(\pi_1) - S_n^{\tau,\delta}(\pi_2)| \leq (2M_1 + 2M_2 + L)||\pi_1 - \pi_2||_\infty$. ∎

Using the concentration inequality in Theorem 3 we can obtain the convergence rate of the worst-case utility regret. We also use a capacity measure of the policy functional class Π, namely the covering number, to prove the convergence rate, which is defined in the following.

Definition 2 (Definition 3.1 in Cucker and Zhou (2007)). *Let Π be a metric space and $\epsilon > 0$, the covering number $\mathcal{N}(\Pi, \epsilon)$ is defined as the minimal $l \in \mathbb{N}$ such that there exist l disks in Π with radius ϵ covering Π.*

Lemma 4. *Under Assumption 3, for any $\{\mathbf{X}_i\}_{i=1}^n \in \boldsymbol{\chi}^{\otimes n}$ and $\epsilon > 0$, it satisfies*

$$\begin{aligned}
&\Pr\left(\sup_{\pi\in\Pi}|S_n^{\tau,\delta}(\pi) - S(\pi)| \leq \epsilon\right) \\
&\quad\geq 1 - \mathcal{N}\left(\Pi, \frac{\epsilon}{4(2M_1 + 2M_2 + L)}\right)\exp\left(-\frac{n\epsilon^2}{32(d_{\max}^2 + 1)(M_1 + M_2)^2}\right),
\end{aligned} \tag{18}$$

where $\mathcal{N}\left(\Pi, \frac{\epsilon}{4(2M_1+2M_2+L)}\right)$ *represents the covering number on the policy functional class* Π *with radius* $\frac{\epsilon}{4(2M_1+2M_2+L)}$.

Proof. According to the assumption (BO), the summands are bounded as $|(2\pi(\mathbf{X}_i)-1)(\tau_i+\delta_i(\pi))| \leq M_1+M_2$, $\forall i \in \{1,\ldots,n\}$. Given the graph $\mathcal{G}=(\mathcal{N},\mathcal{E})$ and its corresponding adjacency matrix A, using the weak independence assumption (WI) a dependence hypergraph can be defined as $\mathcal{G}_h=(\mathcal{N},\mathcal{E}_h)$, where a hyperedge $e_{h,i} \in \mathcal{E}_h$ is defined as $e_{h,i} := \{v_i\} \cup \{v_j | j \in \mathcal{N}_i\} \cup \{v_k | \exists j : A_{ij}=1 \wedge A_{jk}=1\}$. Therefore, the maximal degree of the hypergraph $\mathcal{G}_h$ satisfies $\omega_{\mathcal{G}_h} \leq d_{\max}^2+1$, where $d_{\max}$ indicates the maximal vertex degree of the graph $\mathcal{G}$. Via Theorem 3, we have

$$\Pr\left(|S_n^{\tau,\delta}(\pi)-S(\pi)| \geq \epsilon\right) \leq \exp\left(-\frac{n\epsilon^2}{8(d_{\max}^2+1)(M_1+M_2)^2}\right), \ \forall \pi \in \Pi. \tag{19}$$

Let $l=\mathcal{N}\left(\Pi, \frac{\epsilon}{2(2M_1+2M_2+L)}\right)$ denote the covering number. Consider policies π_j, with $j \in \{1,\ldots,l\}$ located in the center of disks D_j with radius $\frac{\epsilon}{2(2M_1+2M_2+L)}$ which cover the policy functional class Π. Recall the definition $\mathcal{S}(\pi) := S_n^{\tau,\delta}(\pi)-S(\pi)$, by Lemma 3, for any π_j and $\pi \in D_j$, we have

$$|\mathcal{S}(\pi)-\mathcal{S}(\pi_j)| \leq 2(2M_1+2M_2+L)\frac{\epsilon}{2(2M_1+2M_2+L)} = \epsilon.$$

Then $\forall \pi \in D_j$, $\sup_{\pi \in D_j} \mathcal{S}(\pi) \geq 2\epsilon \Rightarrow \mathcal{S}(\pi_j) \geq \epsilon$, which indicates

$$\Pr(\sup_{\pi \in D_j} \mathcal{S}(\pi) \geq 2\epsilon) \leq \Pr(\mathcal{S}(\pi_j) \geq \epsilon) \leq \exp\left(-\frac{n\epsilon^2}{8(d_{\max}^2+1)(M_1+M_2)^2}\right).$$

Since $\Pi = D_1 \cup \cdots \cup D_l$, it is easy to see

$$\begin{aligned}\Pr\left(\sup_{\pi\in\Pi}\mathcal{S}(\pi) \geq 2\epsilon\right) &\leq \sum_{j=1}^{l} \Pr\left(\sup_{\pi\in D_j}\mathcal{S}(\pi) \geq 2\epsilon\right) \\ &\leq \mathcal{N}\left(\Pi, \frac{\epsilon}{2(2M_1+2M_2+L)}\right)\exp\left(-\frac{n\epsilon^2}{8(d_{\max}^2+1)(M_1+M_2)^2}\right).\end{aligned}$$

Upper bound for the probability $\Pr\left(\sup_{\pi\in\Pi}\mathcal{S}(\pi) \leq -2\epsilon\right)$ can be derived in the same way. The statement becomes valid by replacing ϵ by $\frac{\epsilon}{2}$. ∎

Finally, we prove the policy regret bound in Theorem 2 as follows.

Proof. Consider an arbitrary policy $\tilde{\pi} \in \Pi$, we have the following utility difference

$$\begin{aligned}S(\tilde{\pi})-S(\hat{\pi}_n) &= S_n^{\tau,\delta}(\tilde{\pi})-S_n^{\tau,\delta}(\tilde{\pi})+S_n^{\tau,\delta}(\hat{\pi}_n)-S_n^{\tau,\delta}(\hat{\pi}_n) \\ &\quad + S(\tilde{\pi})-S(\hat{\pi}_n)+\hat{S}_n^{\tau,\delta}(\hat{\pi}_n)-\hat{S}_n^{\tau,\delta}(\hat{\pi}_n) \\ &\leq \underbrace{S_n^{\tau,\delta}(\tilde{\pi})-\hat{S}_n^{\tau,\delta}(\tilde{\pi})-S_n^{\tau,\delta}(\hat{\pi}_n)+\hat{S}_n^{\tau,\delta}(\hat{\pi}_n)}_{(1)} \\ &\quad + \underbrace{S(\tilde{\pi})-S_n^{\tau,\delta}(\tilde{\pi})+S_n^{\tau,\delta}(\hat{\pi}_n)-S(\hat{\pi}_n)}_{(2)}.\end{aligned}$$

Using $\forall \pi \in \Pi$, $\pi \in [0,1]$ and assumption (ES) the term $(\star)$ can be bounded as

$$\begin{aligned}
(1) &= \frac{1}{n}\sum_{i=1}^{n} 2(\tau_i - \hat{\tau}_i)(\tilde{\pi}(\mathbf{X}_i) - \hat{\pi}_n(\mathbf{X}_i)) \\
&+ \frac{1}{n}\sum_{i=1}^{n}(2\tilde{\pi}(\mathbf{X}_i) - 1)(\delta_i(\tilde{\pi}) - \hat{\delta}_i(\tilde{\pi})) - \frac{1}{n}\sum_{i=1}^{n}(2\hat{\pi}_n(\mathbf{X}_i) - 1)(\delta_i(\hat{\pi}_n) - \hat{\delta}_i(\hat{\pi}_n)) \\
&\leq \frac{1}{n}\sum_{i=1}^{n} 2|\tau_i - \hat{\tau}_i| + \frac{1}{n}\sum_{i=1}^{n}|\delta_i(\tilde{\pi}) - \hat{\delta}_i(\tilde{\pi})| + \frac{1}{n}\sum_{i=1}^{n}|\delta_i(\hat{\pi}_n) - \hat{\delta}_i(\hat{\pi}_n)| \\
&\leq 2\left(\frac{\alpha_\tau}{n^{\zeta_\tau}} + \frac{\alpha_\delta}{n^{\zeta_\delta}}\right).
\end{aligned}$$

Furthermore, $(2) \leq |S_n^{\tau,\delta}(\tilde{\pi}) - S(\tilde{\pi})| + |S_n^{\tau,\delta}(\hat{\pi}_n) - S(\hat{\pi}_n)| \leq 2\sup_{\pi\in\Pi}|S_n^{\tau,\delta}(\pi) - S(\pi)|$. In summary,

$$\mathcal{R}(\hat{\pi}_n) := \sup_{\tilde{\pi}\in\Pi}(S(\tilde{\pi}) - S(\hat{\pi}_n)) \leq 2\left(\frac{\alpha_\tau}{n^{\zeta_\tau}} + \frac{\alpha_\delta}{n^{\zeta_\delta}}\right) + 2\epsilon,$$

with probability at least $1 - \mathcal{N}\left(\Pi, \frac{\epsilon}{4(2M_1+2M_2+L)}\right)\exp\left(-\frac{n\epsilon^2}{32(d_{\max}^2+1)(M_1+M_2)^2}\right)$ via Lemma 4. ∎

J Capacity-constrained Policy Regret

In this section, we provide an additional policy regret bound under capacity constraint.

Theorem 4. *By Assumption 3, for any small $\epsilon > 0$, the policy regret under the capacity constraint p_t is bounded by $\mathcal{R}(\hat{\pi}_n^{p_t}) \leq 2\left(\frac{\alpha_\tau}{n^{\zeta_\tau}} + \frac{\alpha_\delta}{n^{\zeta_\delta}}\right) + 2\epsilon$ 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 $\mathcal{G}$.*

This capacity-constrained policy regret bound 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.

Before proving Theorem 4, let us first review the definition of utility function $A(\pi)$ following Section 2 of Athey and Wager (2017). The benefit of deploying the intervention policy π compared to assigning everyone in control group is defined as

$$V(\pi) := \mathbb{E}[Y_i(T_i = 1)\pi(\mathbf{X}_i) + Y_i(T_i = 0)(1 - \pi(\mathbf{X}_i))] - \mathbb{E}[Y_i(T_i = 0)] = \mathbb{E}[\pi(\mathbf{X}_i)\tau(\mathbf{X}_i)],$$

and the utility function equals

$$A(\pi) := 2V(\pi) - \mathbb{E}[\tau(\mathbf{X}_i)] = \mathbb{E}[(2\pi(\mathbf{X}_i) - 1)\tau(\mathbf{X}_i)].$$

In the following, let us consider policy learning under treatment constraint p_t, and we will introduce a capacity-constrained utility function under network interference. If the distribution of covariates $\mathcal{P}_{\boldsymbol{\mathcal{X}}}$ is known, and let $\mathcal{P}_{\boldsymbol{\mathcal{X}}}(\pi)$ denote the treatment rule on the covariates space, then a capacity-constrained welfare gain relative to treating no one is defined as (see also Section 4.1 of Kitagawa and Tetenov (2017))

$$\begin{aligned}
V_{p_t}(\pi) &:= \mathbb{E}[[Y_i(T_i = 1)\min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\mathcal{X}}}(\pi)}\} + Y_i(T_i = 0)(1 - \min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\mathcal{X}}}(\pi)}\})]\pi(\mathbf{X}_i) \\
&+ Y_i(T_i = 0)(1 - \pi(\mathbf{X}_i))] - \mathbb{E}[Y_i(T_i = 0)] \\
&= \min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\mathcal{X}}}(\pi)}\}\mathbb{E}[\pi(\mathbf{X}_i)\tau(\mathbf{X}_i))],
\end{aligned}$$

and the corresponding capacity-constrained utility function equals

$$A_{p_t}(\pi) := 2V_{p_t}(\pi) - \mathbb{E}[\tau(\mathbf{X}_i)] = \mathbb{E}[(2\min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi)}\}\pi(\mathbf{X}_i) - 1)\tau(\mathbf{X}_i))].$$

Similarly, the capacity-constrained utility function under interference for interconnected units reads

$$S_{p_t}(\pi) := \mathbb{E}[(2\min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi)}\}\pi(\mathbf{X}_i) - 1)(\tau_i + \delta_i(\pi))].$$

Moreover, the empirical version of $S_{p_t}(\pi)$ reads

$$S^{\tau,\delta}_{n,p_t}(\pi) := \frac{1}{n}\sum_{i=1}^{n}(2\min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi)}\}\pi(\mathbf{X}_i) - 1)(\tau_i + \delta_i(\pi)).$$

The empirical estimation of $S_{p_t}(\pi)$ with causal estimators being plugged in reads

$$\hat{S}^{\tau,\delta}_{n,p_t}(\pi) := \frac{1}{n}\sum_{i=1}^{n}(2\min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi)}\}\pi(\mathbf{X}_i) - 1)(\hat{\tau}_i + \hat{\delta}_i(\pi)),$$

an corresponding optimal capacity-constrained policy is obtained via [5]

$$\hat{\pi}^{p_t}_n \in \operatorname{argmax}_{\pi\in\Pi}\hat{S}^{\tau,\delta}_{n,p_t}(\pi).$$

Moreover, let $\pi^{p_t\star}$ denote the best possible intervention policy from the functional class Π with respect to the utility $S_{p_t}(\pi)$, namely $\pi^{p_t\star} \in \operatorname{argmax}_{\pi\in\Pi}S_{p_t}(\pi)$. The capacity-constrained policy regret is defined as $\mathcal{R}(\hat{\pi}^{p_t}_n) := S_{p_t}(\pi^{p_t\star}) - S_{p_t}(\hat{\pi}^{p_t}_n)$. Before estimating the capacity-constrained intervention policy regret we derive the following inequality similar to Lemma 3.

Lemma 5. *Let $\mathcal{S}_{p_t}(\pi) := S^{\tau,\delta}_{n,p_t}(\pi) - S_{p_t}(\pi)$, for any $\pi_1, \pi_2 \in \Pi$, where the policy class in contained in $[0, 1]$, according to the assumptions (BO) and (LIP) we have*

$$|\mathcal{S}_{p_t}(\pi_1) - \mathcal{S}_{p_t}(\pi_2)| \leq 4[(M_1 + M_2 + L) + \frac{1}{p_t}(M_1 + M_2)]||\pi_1 - \pi_2||_\infty. \tag{20}$$

Proof. Note that $|\mathcal{S}_{p_t}(\pi_1) - \mathcal{S}_{p_t}(\pi_2)| \leq |S_{p_t}(\pi_1) - S_{p_t}(\pi_2)| + |S^{\tau,\delta}_{n,p_t}(\pi_1) - S^{\tau,\delta}_{n,p_t}(\pi_2)|$. We first rewrite $S_{p_t}(\pi)$ as

$$S_{p_t}(\pi) = \min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi)}\}\mathbb{E}[(2\pi(\mathbf{X}_i) - 1)(\tau_i + \delta_i(\pi))] + (\min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi)}\} - 1)\mathbb{E}[\tau_i + \delta_i(\pi)].$$

Recall the definition of $S(\pi)$, and define $T(\pi) := \mathbb{E}[\tau_i + \delta_i(\pi)]$, we have

$$\begin{aligned}
|S_{p_t}(\pi_1) - S_{p_t}(\pi_2)| = &|\min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi_1)}\}S(\pi_1) - \min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi_2)}\}S(\pi_2) \\
&+ (\min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi_1)}\} - 1)T(\pi_1) - (\min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi_2)}\} - 1)T(\pi_2)| \\
\leq &|\min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi_1)}\}||S(\pi_1) - S(\pi_2)| \\
&+ |S(\pi_2)||\min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi_1)}\} - \min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi_2)}\}| \\
&+ |\min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi_1)}\} - 1||T(\pi_1) - T(\pi_2)| \\
&+ |T(\pi_2)||\min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi_1)}\} - \min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi_2)}\}| \\
\leq &|S(\pi_1) - S(\pi_2)| + |T(\pi_1) - T(\pi_2)| \\
&+ (|S(\pi_2)| + |T(\pi_2)|)|\min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi_1)}\} - \min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi_2)}\}|.
\end{aligned}$$

[5] This optimal capacity-constrained policy is, in principle, equivalent to the one obtained by minimizing the loss function $\mathcal{L}_{\text{pol}}(\pi) := -\hat{S}^{\tau,\delta}_n(\pi) + \gamma(\frac{1}{n}\sum_{i=1}^{n}\pi(\mathbf{X}_i) - p_t)$, since, in practice, treatment capacity constraint can be satisfied via Lagrangian multiplier.

Using the following bounds

$$
\begin{aligned}
&|S(\pi_1) - S(\pi_2)| \leq (2M_1 + 2M_2 + L)||\pi_1 - \pi_2||_\infty, \\
&|T(\pi_1) - T(\pi_2)| \leq L||\pi_1 - \pi_2||_\infty, \\
&|S(\pi_2)| \leq M_1 + M_2, \\
&|T(\pi_2)| \leq M_1 + M_2, \\
&|\min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi_1)}\} - \min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi_2)}\}| = |\frac{p_t}{\max\{p_t, \mathcal{P}_{\boldsymbol{\chi}}(\pi_1)\}} - \frac{p_t}{\max\{p_t, \mathcal{P}_{\boldsymbol{\chi}}(\pi_2)\}}| \\
&\qquad\qquad \leq \frac{1}{p_t}|\mathcal{P}_{\boldsymbol{\chi}}(\pi_1) - \mathcal{P}_{\boldsymbol{\chi}}(\pi_2)| \leq \frac{1}{p_t}||\pi_1 - \pi_2||_\infty,
\end{aligned}
$$

yields $|S_{p_t}(\pi_1) - S_{p_t}(\pi_2)| \leq 2[(M_1 + M_2 + L) + \frac{1}{p_t}(M_1 + M_2)]||\pi_1 - \pi_2||_\infty$. Similarly, we also have $|S^{\tau,\delta}_{n,p_t}(\pi_1) - S^{\tau,\delta}_{n,p_t}(\pi_2)| \leq 2[(M_1 + M_2 + L) + \frac{1}{p_t}(M_1 + M_2)]||\pi_1 - \pi_2||_\infty$. ∎

In the same sense as Lemma 4, using Lemma 5 we obtain the following bound for the policy functional class under a capacity constraint p_t.

Lemma 6. *Under Assumption 3, for any $\{\mathbf{X}_i\}_{i=1}^n \in \boldsymbol{\chi}^{\otimes n}$ and $\epsilon > 0$, it satisfies*

$$
\Pr\left(\sup_{\pi\in\Pi} |S^{\tau,\delta}_{n,p_t}(\pi) - S_{p_t}(\pi)| \leq \epsilon\right) \geq 1 - \mathcal{N}\exp\left(-\frac{n\epsilon^2}{32(d^2_{\max} + 1)(M_1 + M_2)^2}\right),
$$

where $\mathcal{N} := \mathcal{N}\left(\Pi, \frac{\epsilon}{8[(M_1+M_2+L)+\frac{1}{p_t}(M_1+M_2)]}\right)$ represents the covering number on the policy functional class Π with radius $\frac{\epsilon}{8[(M_1+M_2+L)+\frac{1}{p_t}(M_1+M_2)]}$.

Finally, we can derive the capacity-constrained policy regret bound as follows.

Proof. Consider an arbitrary policy $\tilde{\pi} \in \Pi$, we have the following utility difference

$$
\begin{aligned}
S_{p_t}(\tilde{\pi}) - S_{p_t}(\hat{\pi}^{p_t}_n) \leq &\underbrace{S^{\tau,\delta}_{n,p_t}(\tilde{\pi}) - \hat{S}^{\tau,\delta}_{n,p_t}(\tilde{\pi}) - S^{\tau,\delta}_{n,p_t}(\hat{\pi}^{p_t}_n) + \hat{S}^{\tau,\delta}_{n,p_t}(\hat{\pi}^{p_t}_n)}_{(1)} \\
&\underbrace{S_{p_t}(\tilde{\pi}) - S^{\tau,\delta}_{n,p_t}(\tilde{\pi}) + S^{\tau,\delta}_{n,p_t}(\hat{\pi}^{p_t}_n) - S_{p_t}(\hat{\pi}^{p_t}_n)}_{(2)}.
\end{aligned}
$$

Using the fact that $\forall\pi \in \Pi$, $|2\pi(\mathbf{X}_i)\min\{1, \frac{p_t}{\mathcal{P}_{\boldsymbol{\chi}}(\pi)}\} - 1| \leq 1$, it is easy to see $(1) \leq 2\left(\frac{\alpha_\tau}{n^{\zeta_\tau}} + \frac{\alpha_\delta}{n^{\zeta_\delta}}\right)$. Furthermore,

$$
(2) \leq |S^{\tau,\delta}_{n,p_t}(\tilde{\pi}) - S_{p_t}(\tilde{\pi})| + |S^{\tau,\delta}_{n,p_t}(\hat{\pi}^{p_t}_n) - S_{p_t}(\hat{\pi}^{p_t}_n)| \leq 2\sup_{\pi\in\Pi} |S^{\tau,\delta}_{n,p_t}(\pi) - S_{p_t}(\pi)|.
$$

In summary, via Lemma 6 it yields the statement. ∎

References

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

P. L. Bartlett, O. Bousquet, S. Mendelson, et al. Local rademacher complexities. *The Annals of Statistics*, 33(4): 1497–1537, 2005.

F. Cucker and D. X. Zhou. *Learning theory: an approximation theory viewpoint*, volume 24. Cambridge University Press, 2007.

L. Devroye, L. Györfi, and G. Lugosi. *A probabilistic theory of pattern recognition*, volume 31. Springer Science & Business Media, 2013.

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

A. Gretton, O. Bousquet, A. Smola, and B. Schölkopf. Measuring statistical dependence with hilbert-schmidt norms. In *International conference on algorithmic learning theory*, pages 63–77. Springer, 2005.

K. M. Harris and J. R. Udry. National longitudinal study of adolescent to adult health (add health), 1994-2008 [public use]. *Ann Arbor, MI: Carolina Population Center, University of North Carolina-Chapel Hill [distributor], Inter-university Consortium for Political and Social Research [distributor]*, pages 08–06, 2018.

E. Jang, S. Gu, and B. Poole. Categorical reparameterization with gumbel-softmax. *International Conference on Learning Representations (ICLR)*, 2017.

S. Janson. Large deviations for sums of partly dependent random variables. *Random Structures & Algorithms*, 24 (3):234–248, 2004.

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

T. Kitagawa and A. Tetenov. Who should be treated? empirical welfare maximization methods for treatment choice. *Econometrica*, 86(2):591–616, 2018.

E. L. Ogburn, O. Sofrygin, I. Diaz, and M. J. van der Laan. Causal inference for social network data. *arXiv preprint arXiv:1705.08527*, 2017.

F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini. Computational capabilities of graph neural networks. *IEEE Transactions on Neural Networks*, 20(1):81–102, 2008.

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

A. W. Van Der Vaart and J. A. Wellner. Weak convergence. In *Weak convergence and empirical processes*, pages 16–28. Springer, 1996.

M. J. Wainwright. *High-dimensional statistics: A non-asymptotic viewpoint*, volume 48. Cambridge University Press, 2019.

Y. Wang, Z.-C. Guo, and J. Ramon. Learning from Networked Examples. In *28th International Conference on Algorithmic Learning Theory (ALT), Kyoto, Japan*, October 2017.

D. Zügner, A. Akbarnejad, and S. Günnemann. Adversarial attacks on neural networks for graph data. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2847–2856. ACM, 2018.