

GRAPH NEURAL NETWORKS FOR CAUSAL INFERENCE UNDER NETWORK CONFOUNDING

Marc ae P Leun[†] Pante s Loupos[‡]

Marc

Abstract. s p p s s | n^e n^e o s on | n^e s

A n n s p o^e s n s e p o s s | o^e n^e f^e n^e n

o p o^e n | o o^e s n s | on n o e n o e p l^e e o p^e

o s n^e e s e e e o^e on s e nonp e s p n s

o s e s e e e o^e s o^e s | n^e o s e on s o | s s | s n

e n s on | n^e s on o n n s e e n^e s n^e o v e s o^e ||

n s on s e s o e s o^e s | on s e | e p^e o n n | s e s

on o n n n e s e nonp | o v e n s on | n on o^e

e s o e s n s n | s s | on o | s s o on o e s

o^e n on s W s o v e p n^e | n^e s s e e s || s e o

s o W e n s on | n^e s on o n n e e s | s n^e s n

o o^e p p o e s p s n^e p s on on s on n^e f^e n^e e s

e on s e e o e | s | o v e n s on | s e e s s

1 Introduction

U VA

x m y x

A X A

GNNs for Network Confounding

Leung and Loupos

'

'

A |

A | "

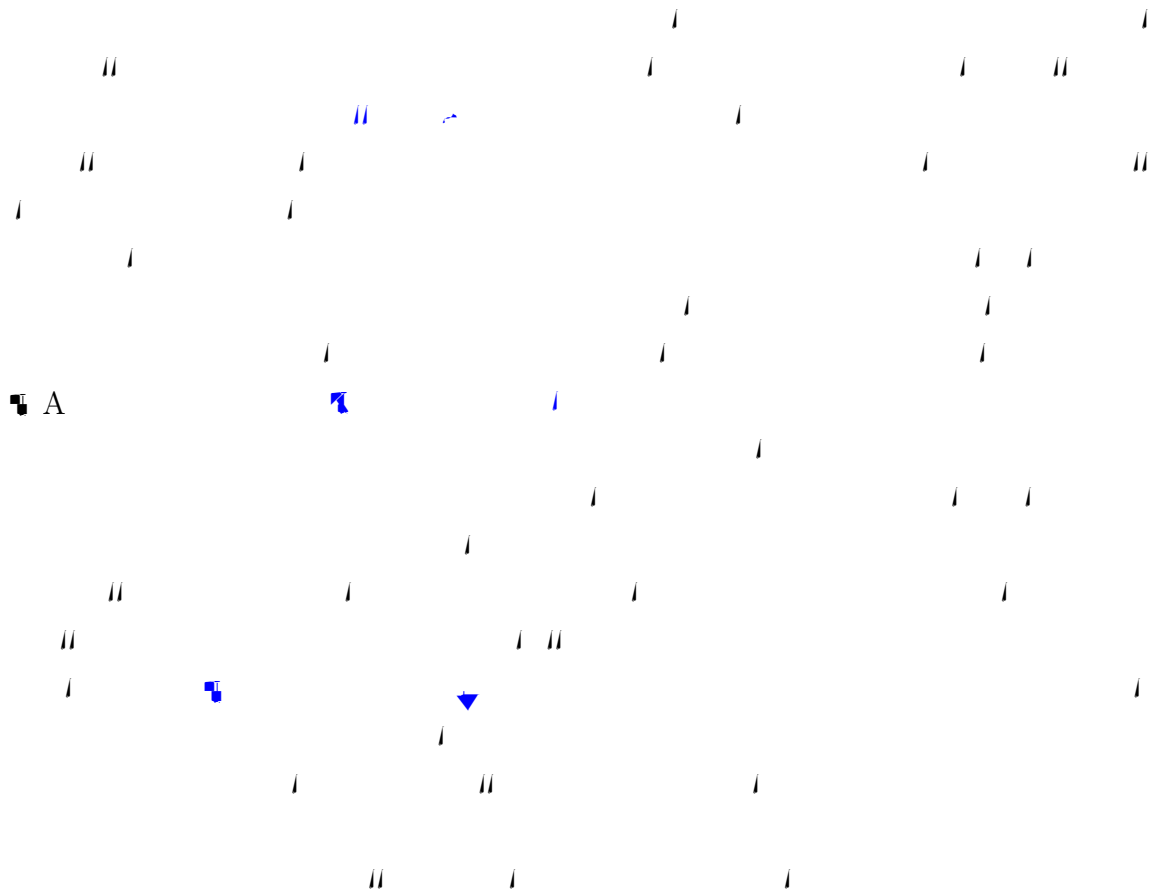
|

'

'

'

GNNs for Network Confounding



1.2 Related Literature



GNNs for Network Confounding

$i \in \mathcal{N}_n = \{1, \dots, n\}$ $K = 1$ $i \in \mathcal{N}_n$
 $A_{ij} \in \{0, 1\}$ $\mathbf{n}(i, 1)$

2 Setup

$\mathcal{N}_n = \{1, \dots, n\}$ $A \in \mathbb{R}^{n \times n}$
 $i \in \mathcal{N}_n$ $(\mathbf{x}_i, \mathbf{z}_i) \in \mathbb{R}^{d_x} \times \mathbb{R}^{d_z}$ $\mathbf{x}_i \in \mathbb{R}^{d_x}$

$$Y_i = g_n(i, D, X, A, \varepsilon) \quad D_i = h_n(i, X, A, \nu)$$

$\nu \in \mathbb{R}$ $X = (\mathbf{x}_i)_{i=1}^n$ $\mathbf{x}_i \in \mathbb{R}^{d_x}$ $\mathbf{z}_i \in \mathbb{R}^{d_z}$ Y, D, ε
 $(g_n, h_n)_{n \in \mathbb{N}}$
 $g_n: \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ $h_n: \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$
 $(Y, D, X, A) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{d_x \times n} \times \mathbb{R}^{n \times n}$ $(A, X, \varepsilon, \nu) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{d_x \times n} \times \mathbb{R} \times \mathbb{R}$
 $\mathbf{x}_i, \mathbf{z}_i \in \mathbb{R}^{d_x}, \mathbb{R}^{d_z}$

$(A, X, \varepsilon, \nu) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{d_x \times n} \times \mathbb{R} \times \mathbb{R}$
 $h_n: \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ $g_n: \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$
 $h_n: \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ $g_n: \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$
 Y_i, D_i

$E = \mathbb{P}_1$

$$Y_i = \frac{\sum_{j=1}^n A_{ij} Y_j}{\sum_{j=1}^n A_{ij}} + \frac{\sum_{j=1}^n A_{ij} Z'_j}{\sum_{j=1}^n A_{ij}} + Z'_i + \varepsilon_i$$

$$Z_i = (D_i, X'_i)'$$

Leung and Loupos

A

$$Y_i = \frac{1}{1+Z} + \sum_{k=0}^{\infty} \tilde{A}^{k+1} Z + \sum_{k=0}^{\infty} \tilde{A}^k \varepsilon_i$$

$$Y_i = g_n(i, D, X, A, \varepsilon)$$

E p 2

$$D_i = 1 + \frac{\sum_{j=1}^n A_{ij} D_j}{\sum_{j=1}^n A_{ij}} + \frac{\sum_{j=1}^n A_{ij} Z'_j}{\sum_{j=1}^n A_{ij}} + Z'_i + \varepsilon_i$$

U

D

$$D_i = h_n(i, X, A, \nu)$$

$$D_j = h_n(j, X, A, \nu)$$

$$D_i = h_n(i, X, A, \nu)$$

E p 2

$$D_i = i$$

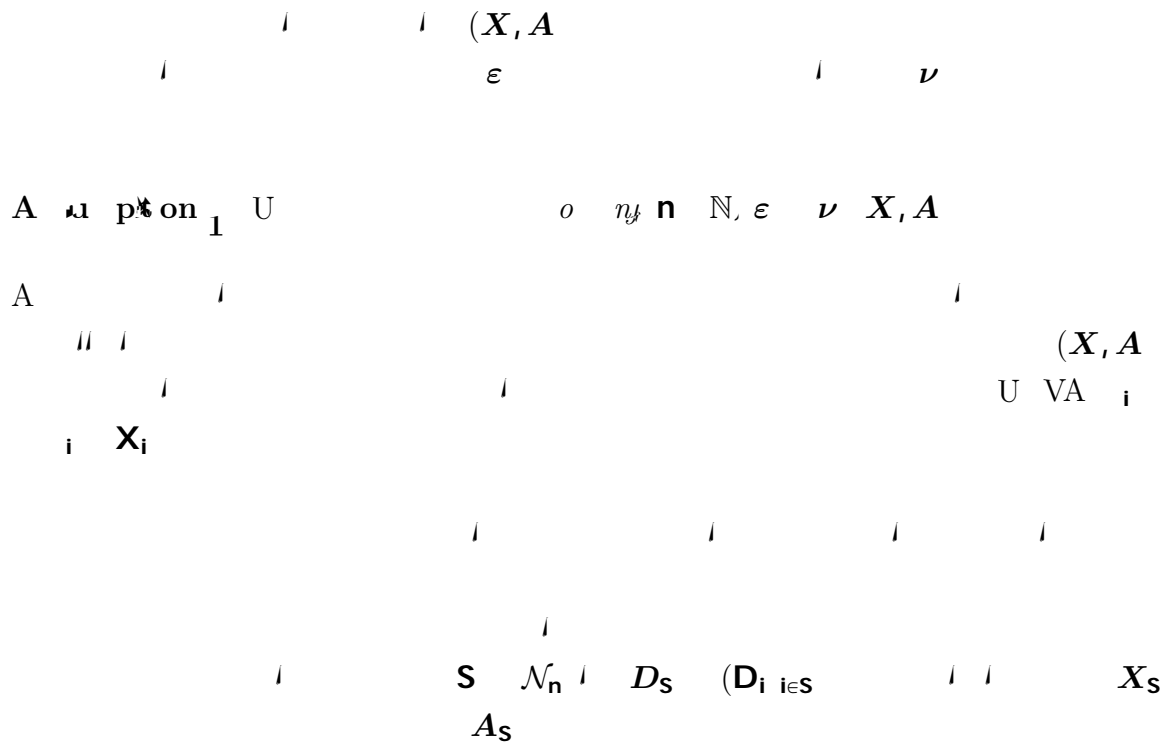
$$Y_i = g_n(D_{\mathcal{N}(i,K)}, i, W'_i)$$

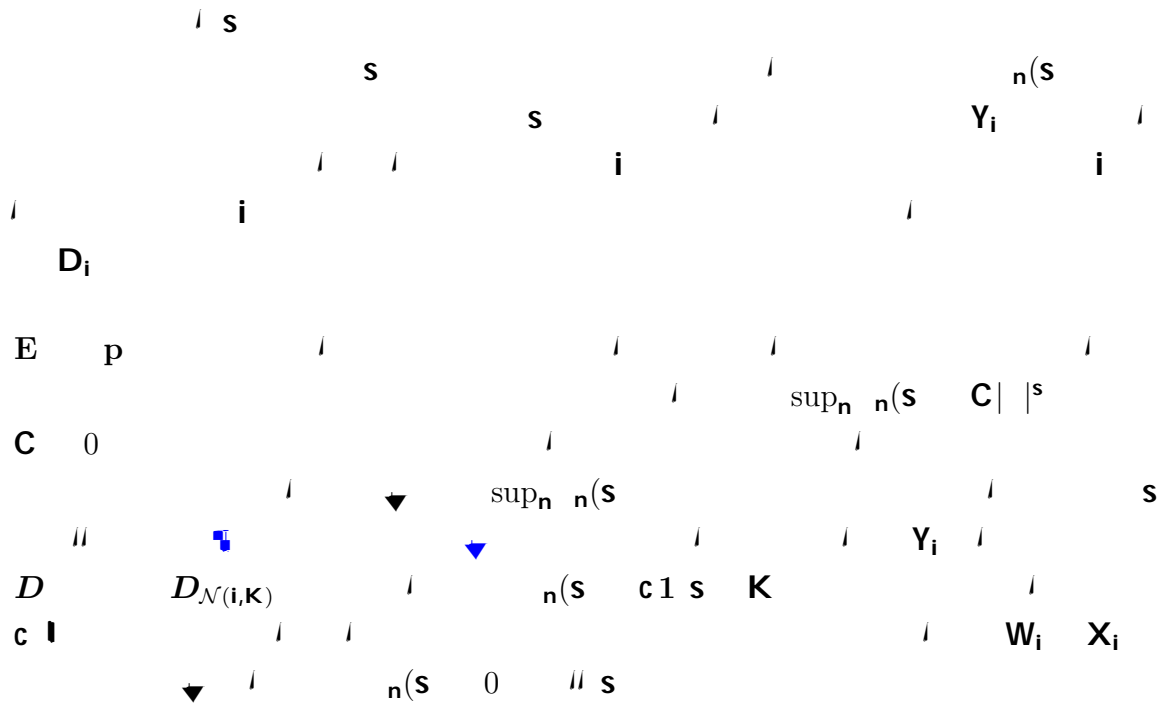
$$W_i = (X, A, K)$$

$$Y_i(d) = g_n(i, d, X, A, \varepsilon)$$

$$Y_i(d) = \dots D_i$$

GNNs for Network Confounding





2.1 Related Literature

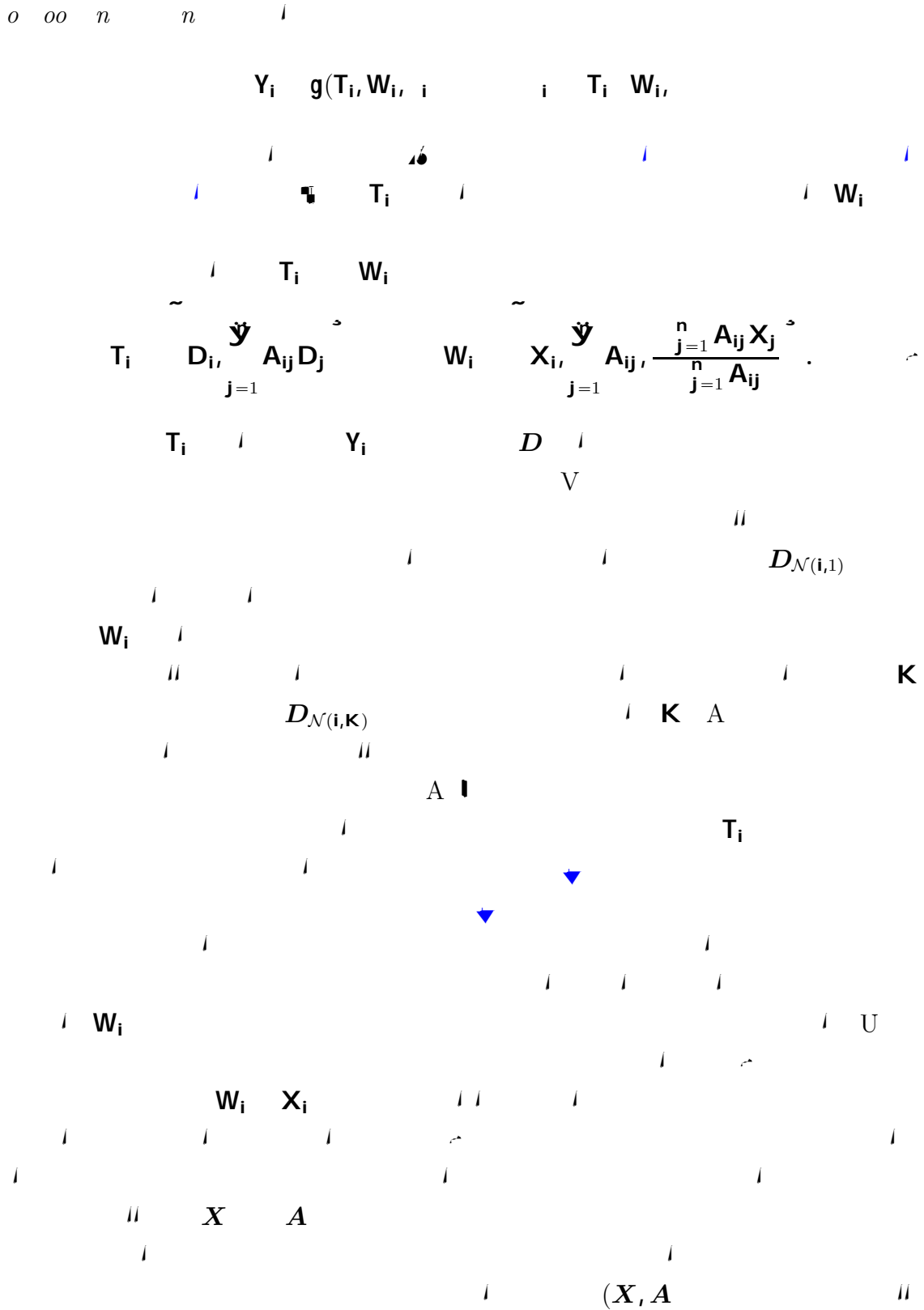
U VA

$$Y_i = g(D_i, X_i, i) \quad i = D_i, X_i.$$

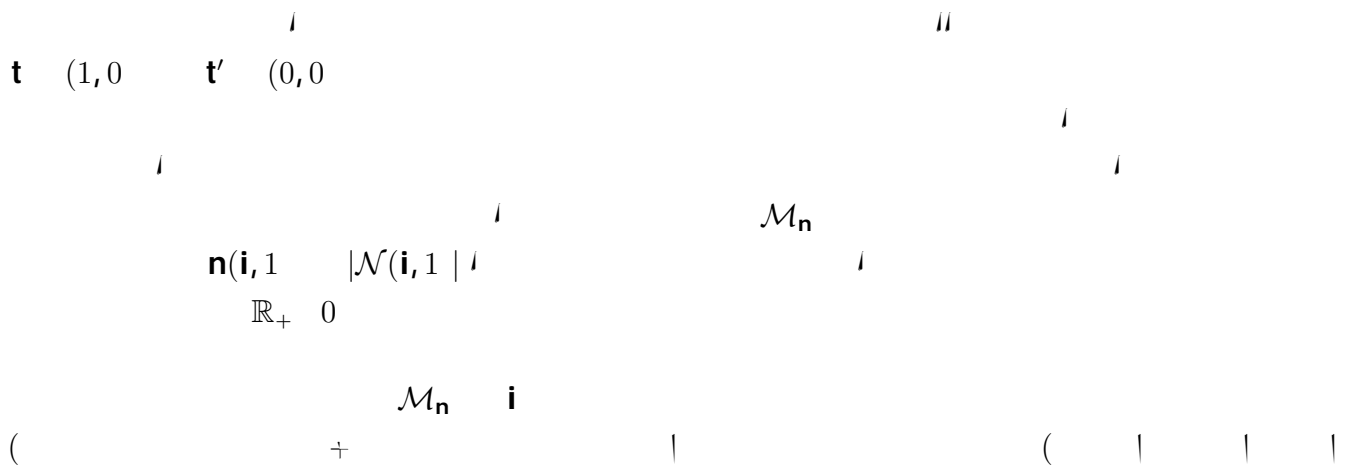
$$T_i = f_n(i, D, A) \quad W_i = q_n(i, X, A)$$

$$f_n(\dots) \quad q_n(\dots) \quad n \quad A$$

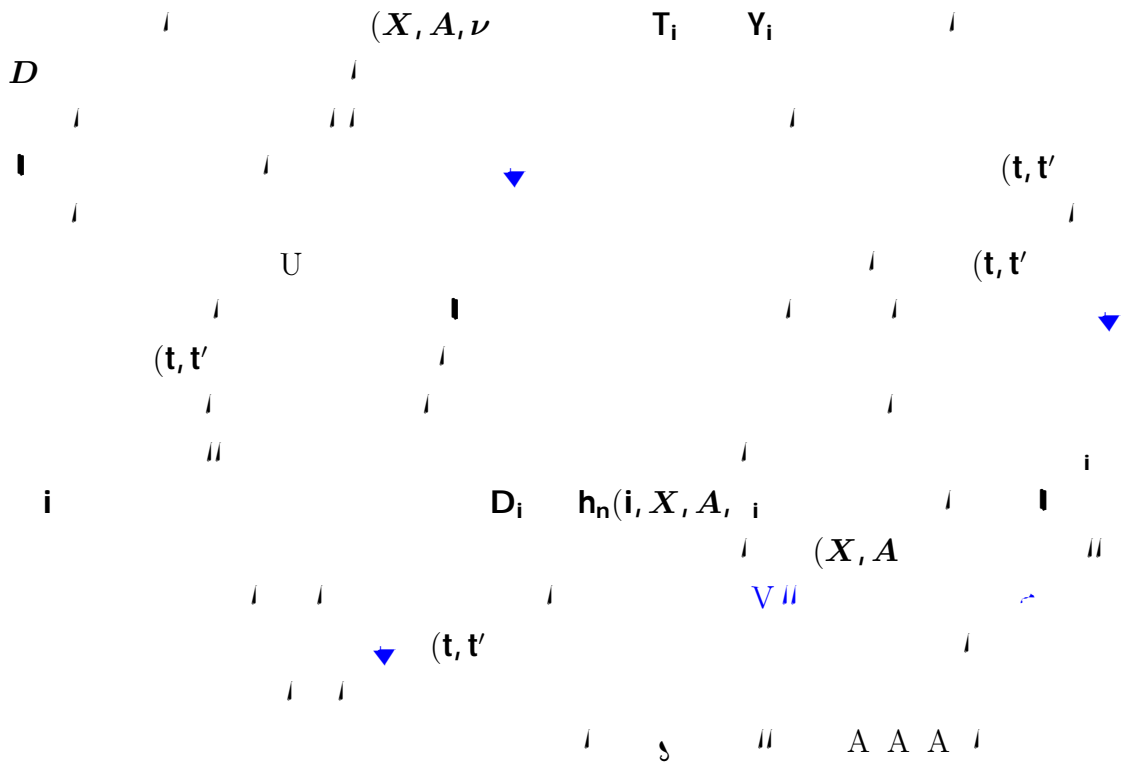
GNNs for Network Confounding



GNNs for Network Confounding



Leung and Loupos



GNNs for Network Confounding

$$\hat{\mu}_i(\mathbf{t}, \mathbf{t}') = \frac{1}{T_i} \frac{\sum_{\mathbf{t}} T_{i, \mathbf{t}} (Y_i - \hat{\mu}_{\mathbf{t}}(\mathbf{i}, \mathbf{X}, \mathbf{A}))}{\hat{\rho}_{\mathbf{t}}(\mathbf{i}, \mathbf{X}, \mathbf{A})} + \hat{\mu}_{\mathbf{t}}(\mathbf{i}, \mathbf{X}, \mathbf{A}) - \frac{1}{T_i} \frac{\sum_{\mathbf{t}'} T_{i, \mathbf{t}'} (Y_i - \hat{\mu}_{\mathbf{t}'}(\mathbf{i}, \mathbf{X}, \mathbf{A}))}{\hat{\rho}_{\mathbf{t}'}(\mathbf{i}, \mathbf{X}, \mathbf{A})} + \hat{\mu}_{\mathbf{t}'}(\mathbf{i}, \mathbf{X}, \mathbf{A}) .$$

3.1 Architecture

GNNs for Network Confounding

Leung and Loupos

$$\begin{aligned}
 & \mathbf{X}_i \\
 & \Gamma(\boldsymbol{\mu}(\cdot, \cdot), \Sigma(\cdot, \cdot), \min(\cdot, \cdot), \max(\cdot, \cdot)) \\
 & \Gamma(\cdot) \Gamma_1(\cdot) \\
 & \Gamma_1(\boldsymbol{\mu}(\cdot, \cdot), \Sigma(\cdot, \cdot), \min(\cdot, \cdot), \max(\cdot, \cdot)) \\
 & \mathbf{n}(\mathbf{i}, 1) \\
 & \Gamma_1(\cdot) \\
 & \mathbf{S}(\cdot, \frac{\log(|\cdot| + 1)}{\cdot}, \frac{1}{\mathbf{n}} \sum_{i=1}^{\mathbf{y}} \log \sum_{j=1}^{\mathbf{y}} \mathbf{A}_{ij} + 1, \cdot) \quad | \cdot, 1, 1. \\
 & 1 \\
 & 0
 \end{aligned}$$

GNNs for Network Confounding

(

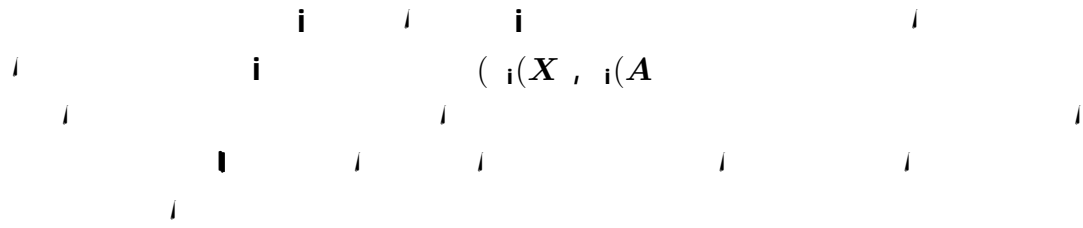
$$\hat{\mathbf{f}}_{\text{GNN}} = \underset{\mathbf{f} \in \mathcal{F}_{\text{GNN}}(\mathbf{L})}{\text{argmin}} \|\mathbf{y} - \mathbf{f}\|_2$$

A

δ

$\mathbf{f} \quad \mathcal{F}_{\text{GNN}}(\mathbf{L}$

GNNs for Network Confounding



Proposition 1. Let $\mathcal{N} = (\mathcal{V}, \mathcal{E})$ be a directed acyclic graph with n nodes and m edges.

$$\mathbf{f}_n(\mathbf{i}, D, A) = \mathbf{f}_n(\mathbf{i}, (D, (A, \dots$$

$$\mathbf{g}_n(\mathbf{i}, D, X, A, \epsilon) = \mathbf{g}_n(\mathbf{i}, (D, (X, (A, (\epsilon, \dots$$

Leung and Loupos

$$\begin{aligned}
 & \mu_{t,t'}(\mathbf{i}) = \frac{1}{\mathbf{p}_t(\mathbf{i}, \mathbf{X}, \mathbf{A})} \sum_{\mathbf{d} \in \mathcal{D}} \mathbf{Y}_i(\mathbf{d} | \mathbf{X}, \mathbf{A}) \mu_{t,t'}(\mathbf{i}, \mathbf{X}, \mathbf{A}, \mathbf{d}) \\
 & \mu_{t,t'}(\mathbf{i}, \mathbf{X}, \mathbf{A}, \mathbf{d}) = \frac{1}{\mathbf{p}_{t'}(\mathbf{i}, \mathbf{X}, \mathbf{A})} \sum_{\mathbf{d}' \in \mathcal{D}} \mathbf{Y}_i(\mathbf{d}' | \mathbf{X}, \mathbf{A}) \mu_{t,t'}(\mathbf{i}, \mathbf{X}, \mathbf{A}, \mathbf{d}, \mathbf{d}') \\
 & \mathbf{i} \in \mathcal{M}_n \\
 & \mathbf{V} = \frac{1}{m_n} \sum_{\mathbf{i} \in \mathcal{M}_n} \mathbf{Y}_i(\mathbf{d} | \mathbf{X}, \mathbf{A})^2
 \end{aligned}$$

A $n \times p$ matrix \mathbf{M} of n rows and p columns of 0 's and 1 's
 $\mathbf{n} \in \mathbb{N}$, $\mathbf{i} \in \mathcal{M}_n$, n dimensional $\mathbf{d} \in \{0, 1\}^n$, $\mathbf{Y}_i(\mathbf{d} | \mathbf{X}, \mathbf{A}) \in \mathbb{M}$
 $\mathbf{p}_t, \mathbf{p}_{t'} \in (0, 1)^n$, $\hat{\mathbf{p}}_t(\mathbf{i}, \mathbf{X}, \mathbf{A}), \hat{\mathbf{p}}_{t'}(\mathbf{i}, \mathbf{X}, \mathbf{A}) \in (0, 1)^n$

GNNs for Network Confounding

$$t, t' \left(\prod_{i=1}^n \right)$$



Leung and Loupos

$$\Lambda_n(s) = \int_{-2M}^{-M} \int_{-M}^M \int_{-M}^M \prod_{i=1}^n t_i^{t_i} \dots$$

GNNs for Network Confounding



or $\mathcal{D}_{n, \mathbf{t}, \mathbf{t}'} \sim_{\mathbf{t}, \mathbf{t}'}(\mathbf{i}, \mathcal{I})$ $n(\mathbf{t}, \mathbf{t}')$ n on $\mathcal{D}_{n, \mathbf{t}, \mathbf{t}'}(\mathbf{i}, \mathcal{I})$ $\mathbf{i}(\mathbf{t}, \mathbf{t}')$
 $\{Y_i, T_i, \mathbf{t}, X_i\}$

$$\| \hat{\beta} - \beta \|_2 \leq \frac{\sigma}{\lambda} \| \mathbf{A} \|_F$$

5 Approximate Sparsity

$$\mathbf{A} \beta \approx \mathbf{y} \quad \mathbf{L}$$

$(\mathbf{X}_{\mathcal{N}(i,L)})$

GNNs for Network Confounding

$$(W_i)_{i=1}^n \sim \nu \quad (i)_{i=1}^n$$

$$V_i(W, \nu; \epsilon) = \frac{\sum_{j=1}^n A_{ij} W_j}{\sum_{j=1}^n A_{ij}} + \frac{\sum_{j=1}^n A_{ij} X_j}{\sum_{j=1}^n A_{ij}} + X_j + \epsilon + \frac{\sum_{j=1}^n A_{ij} j}{\sum_{j=1}^n A_{ij}}$$

$$Y_i \sim V_i(Y, \epsilon; y) \quad y \in (0.5, 0.8, 10, 1) \quad D_i \sim \nu \quad (0.5, 1.5, 1, 1)$$

$$D_i^0 \sim \nu \quad D_i^1 \sim \nu \quad V_i(0, \nu; d) = 0$$

$$A_{ij} = \frac{\sum_{j=1}^n A_{ij} j}{\sum_{j=1}^n A_{ij}} \quad A_{ij}$$

$$A_{ij}$$

$$T_i \sim t \quad (Y_i, T_i) \sim t, X, A \quad n = 1000, 2000, 4000$$

6.2 Nonparametric Estimators

$$L = 1, \quad \Gamma_2(\dots)$$

Leung and Loupos

		$L = 1$			$L = 2$			$L = 3$		
n	$e \quad e$									
H		<hr/>								
$\hat{\tau}(1,0)$										
Cl										
$e \quad Cl$										
$W \hat{\tau}(1,0)$										
$W Cl$										
W										
$\mu \quad Cl$										
μ										

• In situations, the estimator is $\hat{\tau}(1,0) = 0$, treated \approx

Leung and Loupos

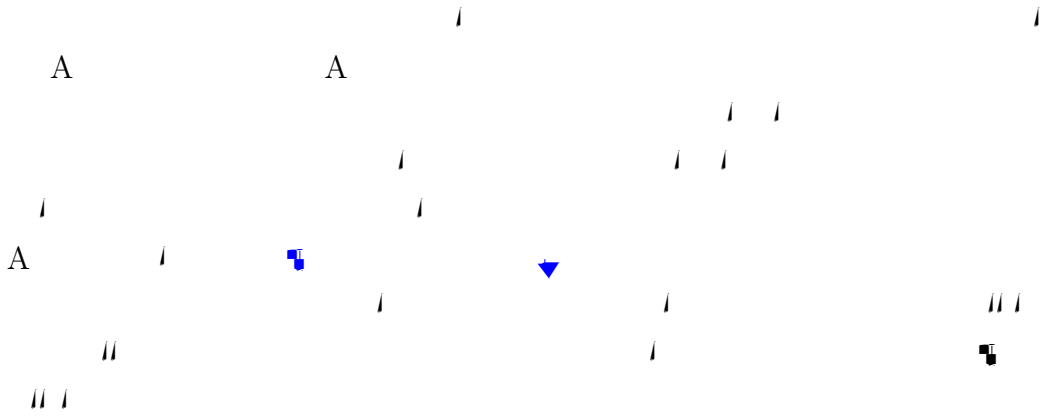


!

7.1 Comparison with He and Song (2024)

Leung and Loupos

(t, t')



GNNs for Network Confounding

n 4413

|0.01, 0.99

A

δ

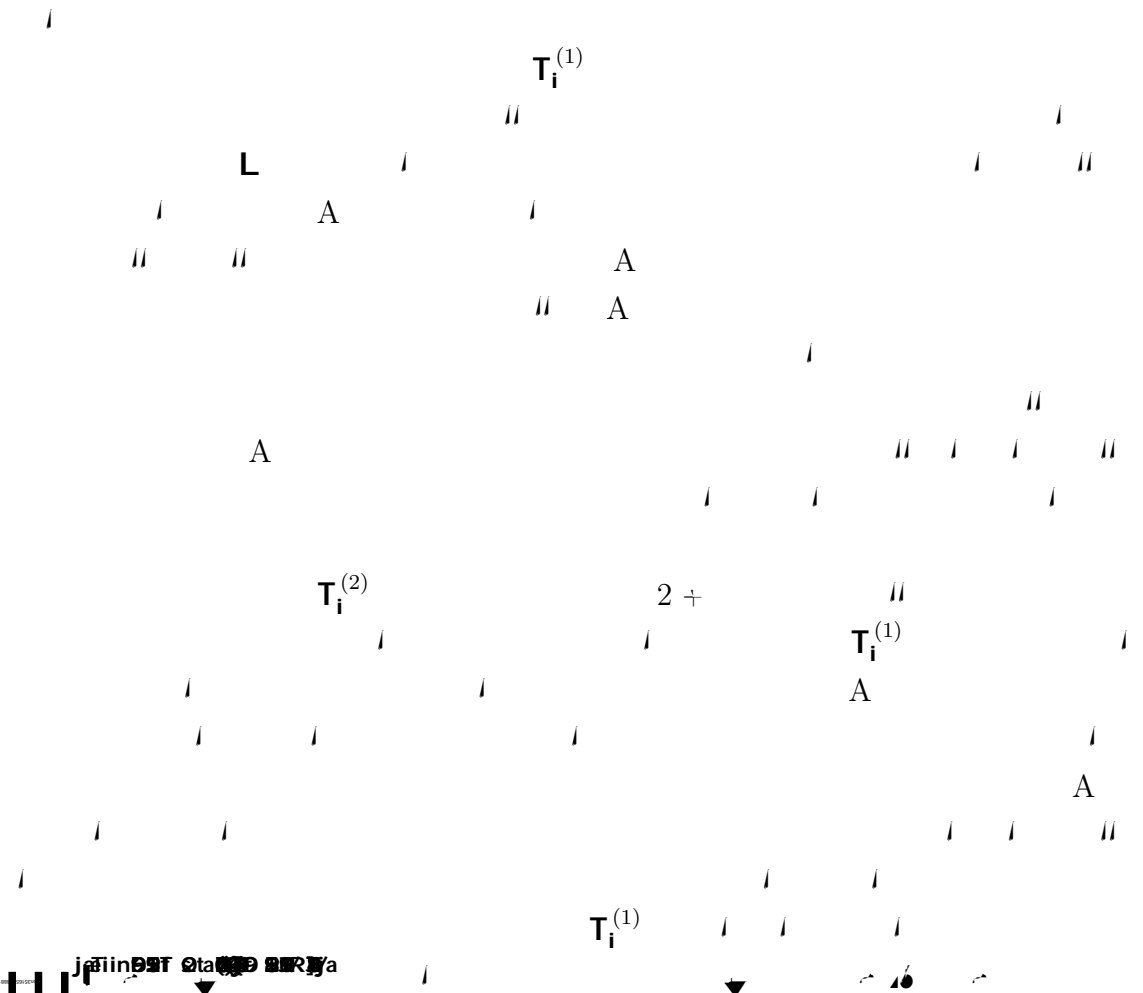
i

$T_i^{(1)}$

	ADM	GNN			GLM		
		Layer	Layer	Layer	Order	Order	Order
Leader case							
G_{ee}	003	000	000	000	000	000	000
G_{sc}	00	00	00	00	00	00	00
G_{all}	003	00	00	00	000	003	000
Leader adopter case							
G_{ee}	0'3	00	00	003	00	00	0'3
G_{sc}	0'	00	003	00	00	00	0'00
G_{all}	0'3	00	00	00	00	0'00	0'300
Adopter case							
G_{ee}	0'	00	00	003	003	003	0'00
G_{sc}	0'	00	003	00	00	003	0'00
G_{all}	0'3	00	00	00	00	003	0'00

n = 4413

Leung and Loupos



misal sobe

jein BT Stage 2021

A Additional Results on GNNs

$$\begin{aligned}
 & \frac{1}{m_n} \sum_{i \in \mathcal{M}_n} \|\hat{\mathbf{p}}_t(\mathbf{i}, \mathbf{X}, \mathbf{A}) - \mathbf{p}_t(\mathbf{i}, \mathbf{X}, \mathbf{A})\|^2 = o_p(n^{-1/2}). \\
 & \frac{1}{m_n} \sum_{i \in \mathcal{M}_n} \|\hat{\mathbf{p}}_t(\mathbf{i}, \mathbf{X}, \mathbf{A}) - \mathbf{p}_t(\mathbf{i}, \mathbf{X}_{\mathcal{N}(\mathbf{i}, L)}, \mathbf{A}_{\mathcal{N}(\mathbf{i}, L)})\|^2 = o_p(n^{-1/2}). \\
 & \|\hat{\mathbf{p}}_t(\mathbf{i}, \mathbf{X}, \mathbf{A}) - \mathbf{p}_t(\mathbf{i}, \mathbf{X}_{\mathcal{N}(\mathbf{i}, L)}, \mathbf{A}_{\mathcal{N}(\mathbf{i}, L)})\| \leq C \frac{WL \log R}{n} \log n + \frac{\log \log n}{n} + \frac{1}{n^2}. \\
 & \frac{1}{n} \sum_{i=1}^n \|\hat{\mathbf{p}}_t(\mathbf{i}, \mathbf{X}, \mathbf{A}) - \mathbf{p}_t(\mathbf{i}, \mathbf{X}_{\mathcal{N}(\mathbf{i}, L)}, \mathbf{A}_{\mathcal{N}(\mathbf{i}, L)})\|^2 \leq C \frac{WL \log R}{n} \log n + \frac{\log \log n}{n} + \frac{1}{n^2}.
 \end{aligned}$$



GNNs for Network Confounding

A

A

δ

W, R, L, n

δ

A.1 WL Function Class

$(X, A), (X', A')$

f

$f(X, A) \neq f(X', A')$

W

n

(X, A)

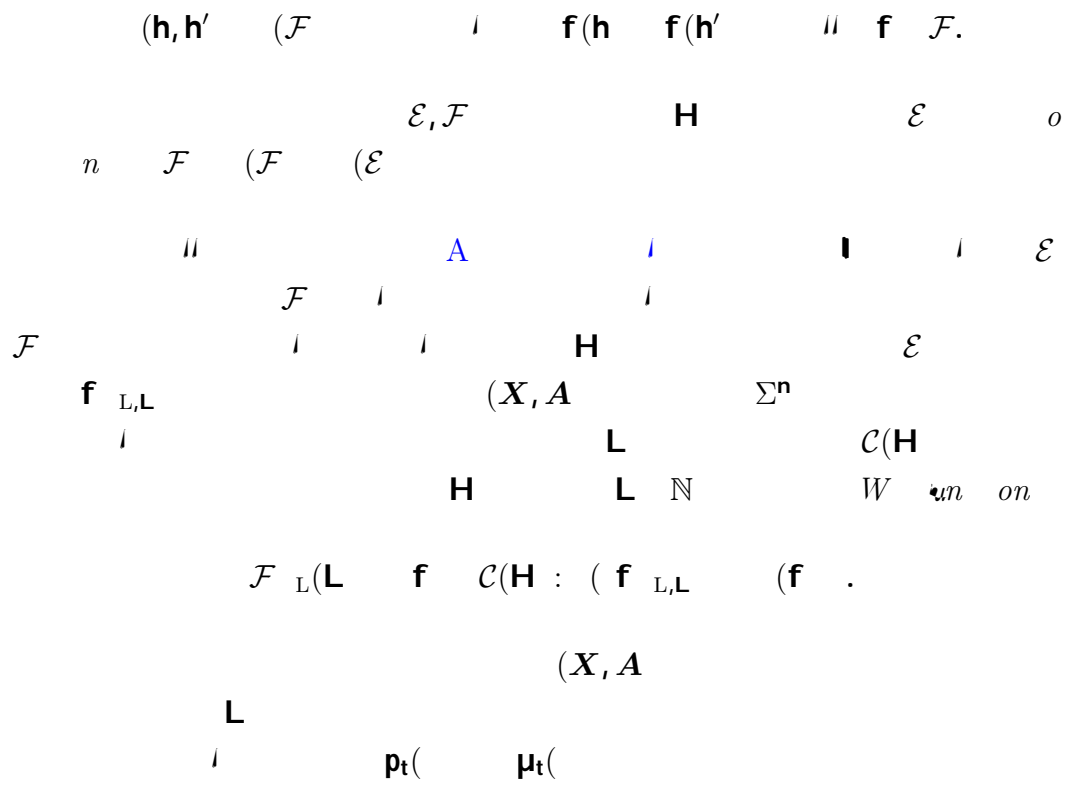
Leung and Loupos

||

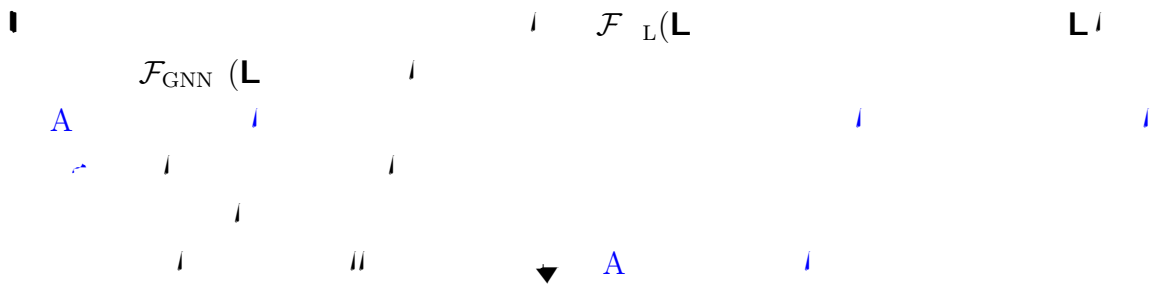
||

GNNs for Network Confounding

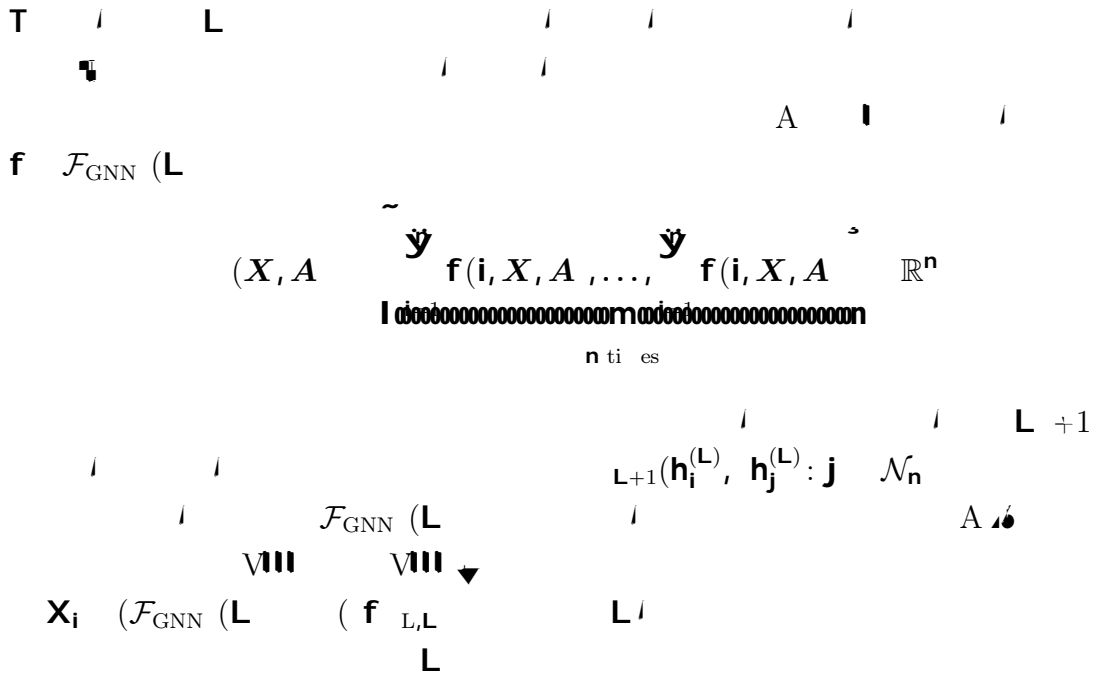
H^2



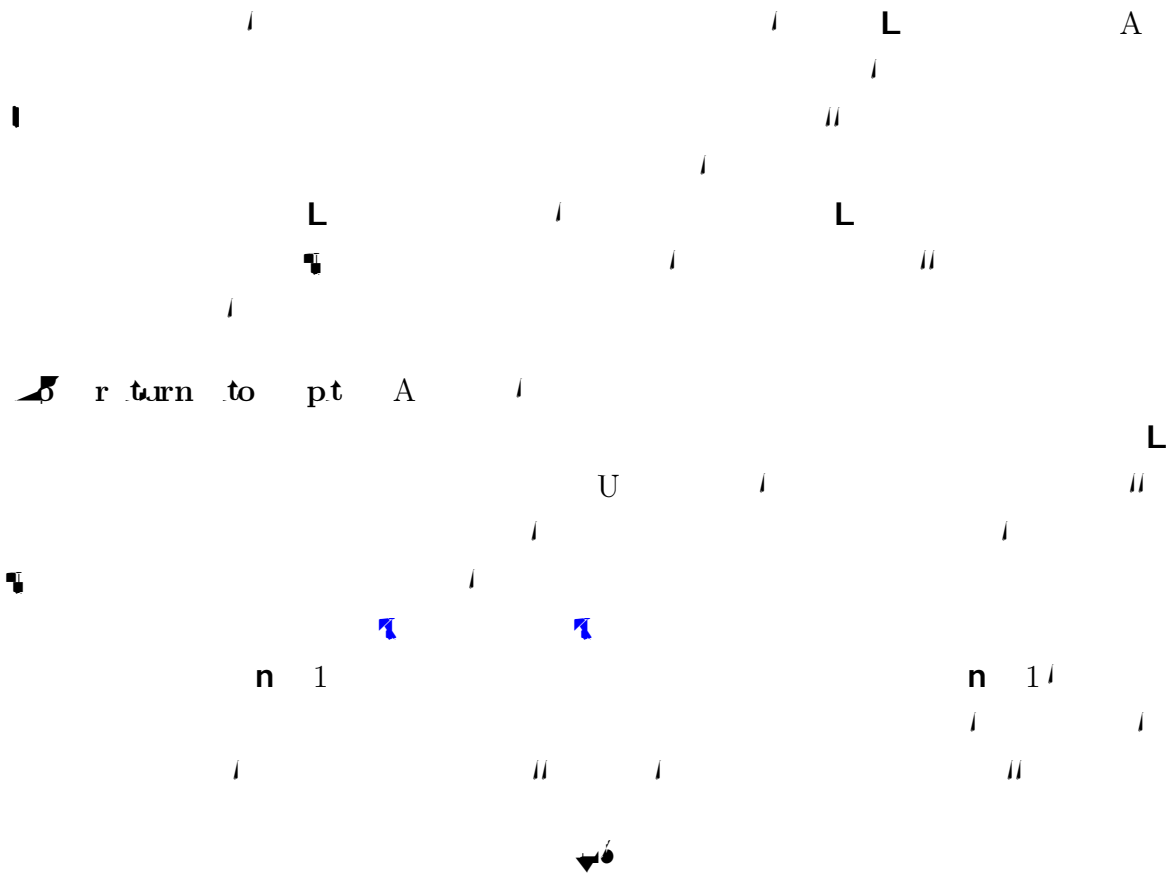
Leung and Loupos



GNNs for Network Confounding



A.2 Disadvantages of Depth



GNNs for Network Confounding

B Verifying §8 Assumptions

$$\begin{aligned}
 & \sup_{\mathbf{s} \in \mathcal{S}} \max_{\mathbf{i} \in \mathcal{N}_n} |\mathcal{N}_A(\mathbf{i}, \mathbf{s})| \leq C s^d \\
 & \text{where } \mathcal{N}_A(\mathbf{i}, \mathbf{s}) = \{j \in \mathcal{N}_n : \|\mathbf{i} - \mathbf{j}\| \leq s\}
 \end{aligned}$$

C $0 < d < 1$

Leung and Loupos

$$D'_B = (D'_j)_{j \in B} \quad B \in \mathcal{N}_n \cup \{ \emptyset \}$$

$$p_t(i, X, A) = (D'_i + (D_i - D'_i) \mathbb{1}_{\{a, b\}}, V'_i + (V_i - V'_i) \mathbb{1}_{\{a, b\}}) \cdot (X, A) \\ + \mathbb{1}_{\{a, b\}} \cdot (D_i - D'_i) \cdot (X, A) + (V_i - V'_i) \cdot (X, A) \cdot R_0$$

/

$$(D'_i \mathbb{1}_{\{a, b\}})$$

GNNs for Network Confounding

1

$$p_t(\mathbf{i}, \mathbf{X}, \mathbf{A}) = p_t(\mathbf{i}, \mathbf{X}_{\mathcal{N}(\mathbf{i}, r_\lambda(s+1))}, \mathbf{A}_{\mathcal{N}(\mathbf{i}, r_\lambda(s+1))}) \mid \mathbf{n}(\mathbf{s} + 1) \sim \mathbf{R}_0.$$

$$\mu_t(\mathbf{i}, \mathbf{X}, \mathbf{A}) = \mathbb{E}[Y_i \mid \mathbf{i}, \mathbf{X}, \mathbf{A}] = \mathbb{E}[Y_i \mid \mathbf{i}, \mathbf{X}_{\mathcal{N}(\mathbf{i}, s)}, \mathbf{A}_{\mathcal{N}(\mathbf{i}, s)}, \mathbf{Y}'_{\mathcal{N}(\mathbf{i}, s)}, \mathbf{g}_{\mathbf{n}(\mathbf{i}, s)}(\mathbf{i}, \mathbf{D}'_{\mathbf{B}}, \mathbf{X}_{\mathbf{B}}, \mathbf{A}_{\mathbf{B}}, \epsilon_{\mathbf{B}})]$$

$$\mathbb{E}[Y_i \mid \mathbf{i}, \mathbf{X}, \mathbf{A}] = \mathbb{E}[Y_i \mid \mathbf{i}, \mathbf{X}_{\mathcal{N}(\mathbf{i}, s)}, \mathbf{A}_{\mathcal{N}(\mathbf{i}, s)}, \mathbf{Y}'_{\mathcal{N}(\mathbf{i}, s)}, \mathbf{g}_{\mathbf{n}(\mathbf{i}, s)}(\mathbf{i}, \mathbf{D}'_{\mathbf{B}}, \mathbf{X}_{\mathbf{B}}, \mathbf{A}_{\mathbf{B}}, \epsilon_{\mathbf{B}})]$$

Leung and Loupos

A



$|\mathbf{R}_1|$ n

GNNs for Network Confounding

$$|R_1| = \sum_{i \in \mathcal{N}_n} \Lambda_n(i, s) \mathbb{1}_i(s) \quad \blacksquare$$

\blacktriangleright $\mathcal{C}^c = \sum_{i \in \mathcal{N}_n} Y_i', D_i' \quad \sum_{i \in \mathcal{N}_n} \mathbb{1}_i(t' = 1) D_i' = d, \quad \sum_{j=1}^n A_{ij} D_j'$
 $n \times n$ matrix A with n rows and n columns, $C = 0$ with n rows and n columns.
 $n \in \mathbb{N}, i \in \mathcal{N}_n, n \times s > 0,$

$$|Y_i| \mathbb{1}_i(t = 1) \mathbb{1}_i(t' = 1) |X, A| = C(1 + \sum_{i \in \mathcal{N}_n} \mathbb{1}_i(s)).$$

Proof. \parallel $a, b, \dots, V_i = \sum_{j=1}^n A_{ij} D_j$
 $V_i' = \sum_{j=1}^n A_{ij} D_j' \quad C = |D_i - D_i'|, |V_i - V_i'|$

$$|Y_i| \mathbb{1}_i(t = 1) \mathbb{1}_i(t' = 1) |X = x, A = a$$

$$|Y_i| \mathbb{1}_i(t = 1) \mathbb{1}_i(t' = 1) |C, X = x, A = a + C = (C^c |X = x, A = a$$

$$|C = 0, A = \dots \quad \blacktriangledown \quad A$$

$$\mathbb{1}_i(t = 1) |D_i = |a, b, V_i|, \quad \mathbb{1}_i(t' = 1) |D_i' = |a, b, V_i'|, \dots$$

U C

$$\mathbb{1}_i(t = 1) |D_i = |a, b, V_i|, \quad C$$

$$\mathbb{1}_i(t' = 1) |D_i' = |a, b, V_i'| + (V_i - V_i') |D_i'$$

$$\mathbb{1}_i(t' = 1) |D_i' = |a, b, V_i'| + (V_i - V_i') |D_i' \quad \text{DUR 92f DOTd (BURE 92f$$

Leung and Loupos

$$\mathbb{R}^d \quad (\mathbf{f} \quad \mathbf{f} \quad \mathcal{L}_d$$

$$\mathcal{P}_n(\mathbf{h}, \mathbf{h}'; \mathbf{s}) = (\mathbf{H}, \mathbf{H}' : \mathbf{H}, \mathbf{H}' \sim \mathcal{N}_n, \mathbf{H} = \mathbf{h}, \mathbf{H}' = \mathbf{h}', \mathbb{P}(\mathbf{H}, \mathbf{H}' > \mathbf{s}) .$$

$$\mathbb{D}_n(\mathbf{h}, \mathbf{h}') = \mathbb{E} \left[\int_{\mathcal{C}} \left(\mathbf{f}(\mathbf{Z}_H) - \mathbf{f}'(\mathbf{Z}_{H'}) \right)^2 \mathcal{C}(\mathbf{h}, \mathbf{h}') \mathcal{P}_n(\mathbf{h}, \mathbf{h}'; \mathbf{s}) \mathbb{D}_n(\mathbf{h}, \mathbf{h}') \right]$$

$$\leq \mathbb{E} \left[\int_{\mathcal{C}} \left(\|\mathbf{f}(\mathbf{Z}_H) - \mathbf{f}'(\mathbf{Z}_{H'})\|_\infty \right)^2 \mathcal{C}(\mathbf{h}, \mathbf{h}') \mathcal{P}_n(\mathbf{h}, \mathbf{h}'; \mathbf{s}) \mathbb{D}_n(\mathbf{h}, \mathbf{h}') \right]$$

$$\leq \mathbb{E} \left[\int_{\mathcal{C}} \left(\|\mathbf{f} - \mathbf{f}'\|_\infty \right)^2 \mathcal{C}(\mathbf{h}, \mathbf{h}') \mathcal{P}_n(\mathbf{h}, \mathbf{h}'; \mathbf{s}) \mathbb{D}_n(\mathbf{h}, \mathbf{h}') \right]$$

$$\leq \mathbb{E} \left[\int_{\mathcal{C}} \left(\|\mathbf{f} - \mathbf{f}'\|_\infty \right)^2 \mathcal{C}(\mathbf{h}, \mathbf{h}') \mathcal{P}_n(\mathbf{h}, \mathbf{h}'; \mathbf{s}) \mathbb{D}_n(\mathbf{h}, \mathbf{h}') \right]$$

Proof. $\mathcal{F}_n = \mathbb{E} \left[\int_{\mathcal{C}} \left(\mathbf{f}(\mathbf{Z}_H) - \mathbf{f}'(\mathbf{Z}_{H'}) \right)^2 \mathcal{C}(\mathbf{h}, \mathbf{h}') \mathcal{P}_n(\mathbf{h}, \mathbf{h}'; \mathbf{s}) \mathbb{D}_n(\mathbf{h}, \mathbf{h}') \right]$

$$\mathbf{D}_i^{(s)} = \mathbf{h}_{n(i,s)}(\mathbf{i}, \mathbf{X}_{\mathcal{N}(i,s)}, \mathbf{A}_{\mathcal{N}(i,s)}, \nu_{\mathcal{N}(i,s)}) .$$

$$\mathbf{D}_{\mathcal{N}(i,s)}^{(s)} = \left(\mathbf{D}_j^{(s)} \right)_{j \in \mathcal{N}(i,s)}$$

$$\mathbf{1}_i^{(s)}(\mathbf{t}) = \mathbf{1}_{\mathbf{f}_{n(i,s/2)}(\mathbf{i}, \mathbf{D}_{\mathcal{N}(i,s/2)}^{(s/2)}, \mathbf{A}_{\mathcal{N}(i,s/2)}) = \mathbf{t}}$$

$$\mathbf{Y}_i^{(s)} = \mathbf{g}_{n(i,s/2)}(\mathbf{i}, \mathbf{D}_{\mathcal{N}(i,s/2)}^{(s/2)}, \mathbf{X}_{\mathcal{N}(i,s/2)}, \mathbf{A}_{\mathcal{N}(i,s/2)}, \epsilon_{\mathcal{N}(i,s/2)})$$

$$\mathbf{Z}_i^{(s)} = \mathbf{1}_i^{(s)}(\mathbf{t}) \left(\mathbf{Y}_i^{(s)} - \mu_{\mathbf{t}}(\mathbf{i}, \mathbf{X}, \mathbf{A}) \right)$$

GNNs for Network Confounding

$$\begin{aligned}
 & A \quad \left(\mathbf{Z}_i^{(s/2, \cdot)} \right)_{i \in H} \quad \left(\mathbf{Z}_j^{(s/2, \cdot)} \right)_{j \in H^1} \quad \mathcal{F}_n \\
 & \left| \left(\cdot, \mathcal{F}_n \right) \right| \quad \left| \left(\cdot^{(s/2)}, \mathcal{F}_n \right) \right| + \left| \left(\cdot^{(s/2)}, \cdot^{(s/2)} \mathcal{F}_n \right) \right| \\
 & 2 \|\mathbf{f}'\|_\infty \left| \left(\cdot^{(s/2)}, \mathcal{F}_n \right) \right| + 2 \|\mathbf{f}\|_\infty \left| \left(\cdot^{(s/2)}, \cdot^{(s/2)} \mathcal{F}_n \right) \right| \\
 & 2 \mathbf{h} \|\mathbf{f}'\|_\infty \left(\mathbf{f} + \mathbf{h}' \|\mathbf{f}\|_\infty \right) \left(\mathbf{f}' \max_{i \in \mathcal{N}} \right)
 \end{aligned}$$

Leung and Loupos

$$\mathcal{L}_h = \mathcal{L}_{h'} \quad \mathbf{s} = 0 \quad (\mathbf{H}, \mathbf{H}') \quad \mathcal{P}_n(\mathbf{h}, \mathbf{h}'; \mathbf{s})$$

$$\mathbf{Y}_i^{(s)} = \mathbf{g}_{n(i,s)}(\mathbf{i}, \mathbf{D}_{\mathcal{N}(i,s)}, \mathbf{X}_{\mathcal{N}(i,s)}, \mathbf{A}_{\mathcal{N}(i,s)}, \boldsymbol{\varepsilon}_{\mathcal{N}(i,s)})$$

$$\mathbf{f}(\mathbf{Y}_i)_{i \in \mathcal{H}} = \mathbf{f}'(\mathbf{Y}_i)_{i \in \mathcal{H}'} \quad (\mathbf{s}) \quad \mathbf{f}(\mathbf{Y}_i^{(s)})_{i \in \mathcal{H}} = \mathbf{f}'(\mathbf{Y}_i^{(s)})_{i \in \mathcal{H}'}$$

A

$$\begin{aligned} & \left| \left(\mathbf{f}, \mathcal{F}'_n \right) \right| \left| \left(\mathbf{f}^{(s/2)}, \mathcal{F}'_n \right) \right| + \left| \left(\mathbf{f}^{(s/2)}, \mathcal{F}'_n \right) \right| \left| \left(\mathbf{f}, \mathcal{F}'_n \right) \right| \\ & 2 \|\mathbf{f}'\|_\infty \|\mathbf{f}\|_\infty \|\mathbf{f}^{(s/2)}\|_\infty + 2 \|\mathbf{f}\|_\infty \|\mathbf{f}'\|_\infty \|\mathbf{f}^{(s/2)}\|_\infty \\ & 2 \|\mathbf{h}\| \|\mathbf{f}'\|_\infty (\|\mathbf{f}\| + \|\mathbf{h}'\| \|\mathbf{f}\|_\infty) (\mathbf{f}') \max_{i \in \mathcal{N}_n} \|\mathbf{Y}_i - \mathbf{Y}_i^{(s/2)}\| \mathcal{F}'_n \\ & 2 \|\mathbf{h}\| \|\mathbf{f}'\|_\infty (\|\mathbf{f}\| + \|\mathbf{h}'\| \|\mathbf{f}\|_\infty) (\mathbf{f}') \mathbf{n}(\mathbf{s}^2) \end{aligned}$$

$$\left| \left(\mathbf{f}, \mathcal{F}'_n \right) \right| \left| \left(\mathbf{f}^{(s/2)}, \mathcal{F}'_n \right) \right| \quad \left| \left(\mathbf{f}^{(s/2)}, \mathcal{F}'_n \right) \right| \left| \left(\mathbf{f}, \mathcal{F}'_n \right) \right|$$

A

Leung and Loupos

$$\hat{\rho}_t(\mathbf{i}, X, A) \quad \text{where } \mathbf{C}, \mathbf{C}' = 0 \quad \mathbf{R}_{it}^2$$

$$\frac{1}{m_n} \sum_{i \in \mathcal{M}_n} \sum_{j \in \mathcal{M}_n} \left(Y_i - \mu_i \right) \left(Y_j - \mu_j \right) D_{ij} X_{ij} A_{ij} \mathbf{1}_i(t) \mathbf{1}_j(t)$$

GNNs for Network Confounding

$\hat{\rho}_t(\mathbf{i}, \mathbf{X}, \mathbf{A})$

$$\Delta_{\mathbf{i}}(\mathbf{t}) = (\hat{\rho}_t(\mathbf{i}) - \mu_t(\mathbf{i})) \mathbf{p}_t(\mathbf{i}) \mathbf{1}_{\mathbf{i}}(\mathbf{t})$$

Leung and Loupos

GNNs for Network Confounding

Bronstein

<https://towardsdatascience.com/do-we-need-deep-graph-neural-networks>
be 2dec A

→ Bran Co n n X n X 0

C n Z n C n n Bran A A
V

C rno D C t D r r E D u a C n n
ono ou n

Cor a C D B n n n
A A V

D root on o A n A o

D r F C r n q . Dono n n A n
B r r ou n o ono 5

D C o r n t Y B n q n X Br on
X n X 00 00

E n C po n n B nn
U A X n X 0 5

F r r A X n X 0

→ n n r
ono

Leung and Loupos

F \rightarrow M n \rightarrow n n
X n X 0 0

For \rightarrow r \rightarrow E A ro \rightarrow n F M \rightarrow l
A n A o on \rightarrow \rightarrow \rightarrow ou n o

ro \rightarrow M \rightarrow A l

\rightarrow X n on \rightarrow o o
no \rightarrow u o o n \rightarrow

orn \rightarrow \rightarrow M \rightarrow n o \rightarrow n \rightarrow t l l
U l A \rightarrow u o \rightarrow

GNNs for Network Confounding

⌘ n B M A ▼

GNNs for Network Confounding

o m M I I X n
 X 0 00 Σ

Leung and Loupos

Zop₁ μ

X 0 0 5

! A! A! Y

X

n