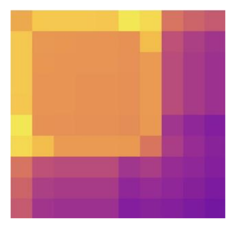
Transfer Learning on Edge Connecting Probability Estimation Under Graphon Model

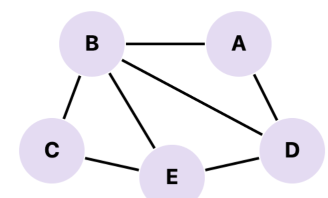
Yuyao Wang[1], Yu-Hung Cheng[2], Debarghya Mukherjee[1*], Huimin Cheng[2*]
[1]Department of Mathematics and Statistics, Boston University; [2]Department of Biostatistics, Boston University.
* Co-corresponding authors.

Motivation

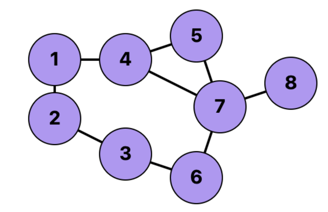
Graphon estimation is data-hungry: accuracy improves with network size, yet many real datasets have very few nodes.



Small graphs limit inference: poor estimation cascades into weak performance in downstream tasks.



Transfer opportunities exist: related domains provide larger graphs, motivating transfer of knowledge.



Current gap: prior transfer methods for graphons either require known node correspondences or subset relationships, which are impractical in real settings.

To answer these motivations, we propose **GTRANS,** a novel transfer learning framework for graphon estimation, with the following contributions:

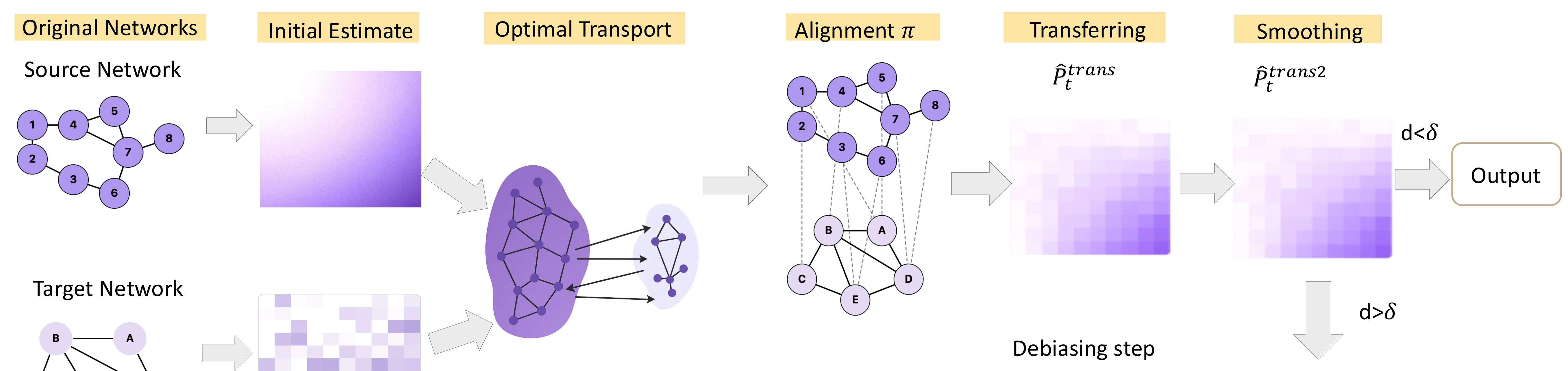
- **GTRANS framework**: first transfer method, GTRANS-GW and GTRANS-EGW (entropic), for graphon estimation without node correspondences.
- **Theory**: consistency guarantees for alignment via smoothed estimators.
- **Experiments**: lower MSE on synthetic networks; better graph classification & link prediction on real data.

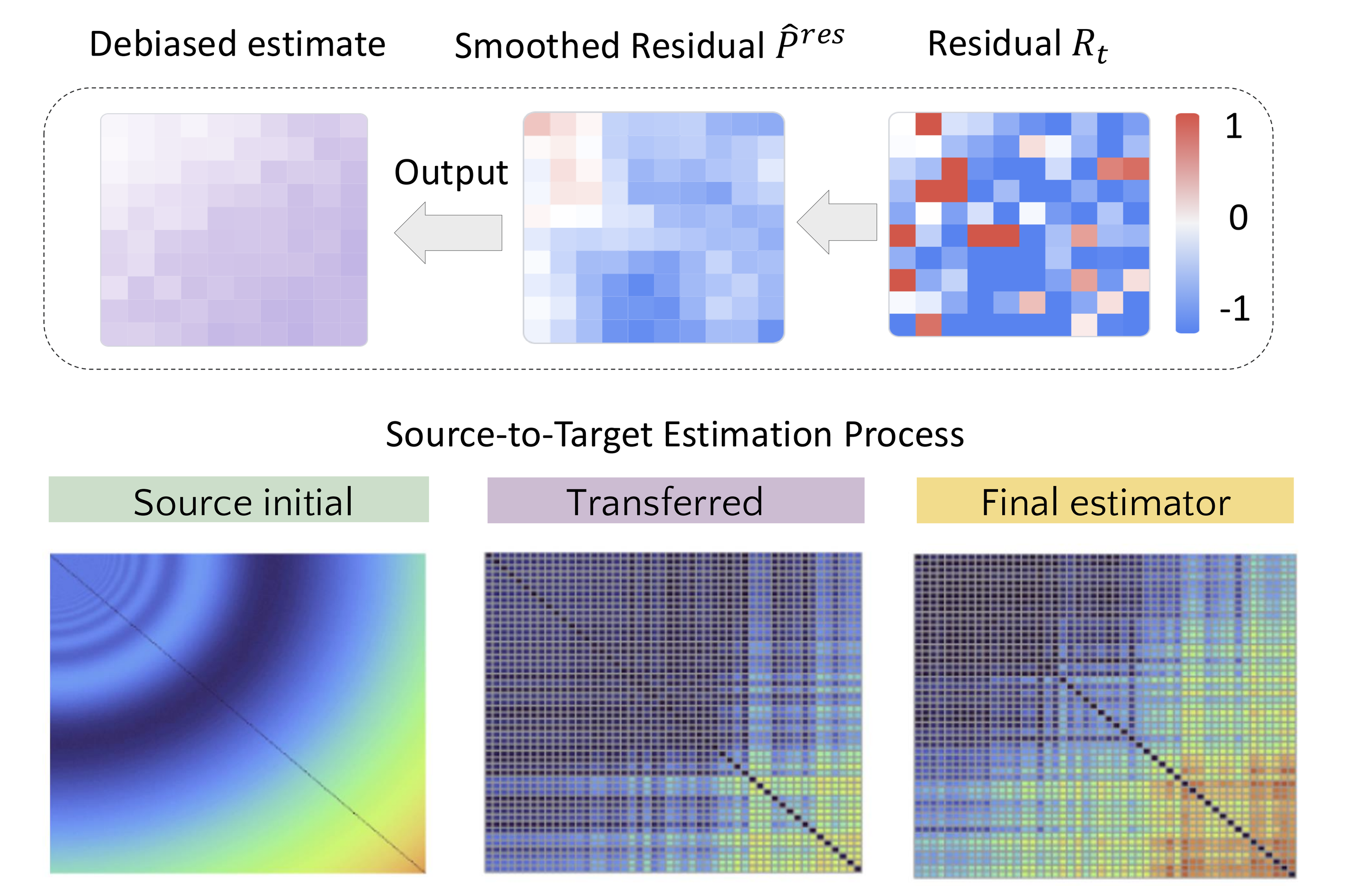
Theoretical Theorem

Theorem 1: Let $\hat{\pi}$ and π^* be the solutions of the EGW optimization problem with $(C,D) = \left(\widehat{P_s^{ini}}, \widehat{P_t^{ini}}\right)$ and $(C,D) = (P_s, P_t)$ respectively. If $|\pi^*|_\infty \leq \frac{\epsilon}{C_1 |P_s \otimes P_t|_{op}}, C_1 > 2$, then there exists a universal constant C_2 such that

$$|\hat{\pi} - \pi^*|_F \leq C_2 \frac{|(P_s \otimes P_t) - \left(\widehat{P_s^{ini}} \otimes \widehat{P_t^{ini}}\right)|_{op}}{|P_s \otimes P_t|_{op}}$$

Methodology





GTRANS framework:

Initial estimation

- Get initial estimates $\widehat{P_s^{ini}} \in [0,1]^{n_s \times n_s}, \widehat{P_t^{ini}} \in [0,1]^{n_t \times n_t}$ by neighborhood smoothing (NS)

Transfer step

- Use optimal transport to compute alignment matrix $\tilde{\pi}$
- Transfer source estimate $\rightarrow \widehat{P_t^{trans}} = \tilde{\pi}^T \widehat{P_s^{ini}} \tilde{\pi}$
- Apply neighborhood smoothing(NS) $\rightarrow \widehat{P_t^{trans2}}$
- If transport distance $d \leq \delta$: final $\widehat{P^t} = \widehat{P_t^{trans2}}$

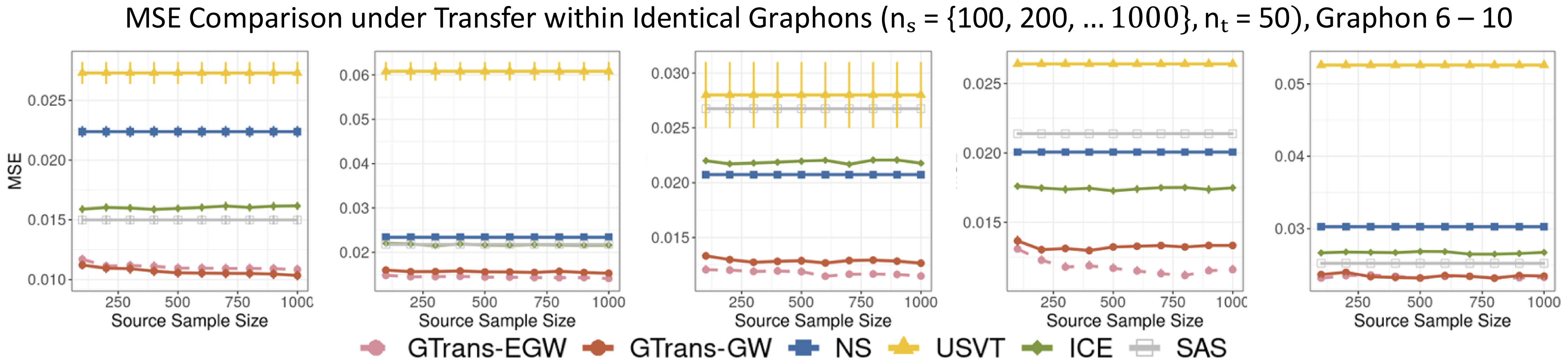
Debiasing step (when $d > \delta$)

- Residual:$R_t = \widehat{P_t^{ini}} - \widehat{P_t^{trans2}}$ & Apply NS $\rightarrow \widehat{P^{res}}$
- Final output: $\widehat{P^t} = \widehat{P_t^{trans2}} + \widehat{P^{res}}$

Experiment Results

Simulation: 10 graphons $(f_1, \ldots, f_{10})$ Networks: $f_t(u,v) = f_s(u,v) + \xi, \quad \xi \sim U(-0.01, 0.01)$

MSE Comparison under Transfer within Identical Graphons (n_s = {100, 200, … 1000}, n_t = 50), Graphon 6 – 10



MSE Comparison under Transfer between Different Graphons (n_s = 500, n_t = 50)

Similarity	Scenario	GTRANS-GW	GTRANS-EGW	NS	USVT	ICE	SAS
Similar	7 → 6	**0.9 ± 0.2**	1.0 ± 0.3	2.1 ± 0.3	3.0 ± 0.9	1.7 ± 0.3	1.6 ± 0.4
	6 → 7	**1.8 ± 0.3**	**1.8 ± 0.3**	2.3 ± 0.3	5.9 ± 2.7	2.2 ± 0.3	2.4 ± 0.6
	8 → 7	**1.3 ± 0.3**	1.4 ± 0.3	2.3 ± 0.3	5.9 ± 2.7	2.2 ± 0.3	2.4 ± 0.6
	3 → 1	**1.6 ± 0.5**	1.7 ± 0.2	2.5 ± 0.3	5.1 ± 4.6	2.7 ± 0.3	2.1 ± 0.3
Different	10 → 5	1.6 ± 0.2	**1.5 ± 0.2**	1.7 ± 0.1	3.5 ± 1.5	1.7 ± 0.2	5.5 ± 1.0
	5 → 10	2.6 ± 0.3	**2.5 ± 0.3**	3.1 ± 0.6	5.3 ± 0.6	2.8 ± 0.3	2.6 ± 0.3

Conclusion: GTRANS-GW / GTRANS-EGW consistently achieve lowest MSE.

Real Data – Graph Classification with G-Mixup Augmentation: GTRANS-GW / GTRANS-EGW consistently outperform baselines

Table: Graph classification accuracy (%): Real Networks.

Source	Target	GTRANS-GW	GTRANS-EGW	NS	USVT	SAS	ICE
Reddit-B	IMDB-B	76.30 ± 2.35	**76.80 ± 1.52**	72.90 ± 2.10	73.85 ± 2.40	74.25 ± 1.93	74.30 ± 2.16
COLLAB	IMDB-B	76.25 ± 2.06	**77.50 ± 2.13**	72.90 ± 2.10	73.85 ± 2.40	74.25 ± 1.93	74.30 ± 2.16
Reddit-B	IMDB-M	49.10 ± 1.33	**51.27 ± 1.98**	43.80 ± 2.59	48.00 ± 2.93	44.10 ± 2.05	43.90 ± 1.27
COLLAB	IMDB-M	**50.47 ± 1.42**	50.23 ± 0.92	43.80 ± 2.59	48.00 ± 2.93	44.10 ± 2.05	43.90 ± 1.27
D&D	PROTEINS	**69.33 ± 2.55**	68.52 ± 1.59	63.18 ± 1.94	65.11 ± 2.21	65.25 ± 1.85	65.38 ± 1.84

Real Data – Link Prediction with source wiki-vote (n_s = 889): GTRANS-GW / GTRANS-EGW consistently outperform baselines

Table: Link prediction AUC (%): Real Networks

Dataset	GTRANS-GW	GTRANS-EGW	NS	USVT	SAS	ICE
Dolphins (n_t = 62)	75.96 ± 8.53	**76.26 ± 8.54**	70.60 ± 9.01	72.36 ± 10.16	50.66 ± 6.35	75.36 ± 8.40
Firm (n_t = 33)	**71.31 ± 12.18**	71.26 ± 12.06	66.32 ± 12.27	65.56 ± 12.25	54.90 ± 7.59	64.26 ± 12.20
Football (n_t = 115)	86.64 ± 3.66	86.74 ± 3.72	**86.75 ± 3.49**	85.32 ± 3.56	44.56 ± 7.38	82.46 ± 4.59
Karate (n_t = 34)	82.47 ± 10.36	**82.53 ± 10.46**	76.74 ± 12.01	71.86 ± 15.20	63.88 ± 11.15	80.43 ± 11.22