Synthetic Counterfactual Networks

Chen-Wei Hsiang*

University College London

This version: November 7, 2026

(Click here for the latest version)

Abstract

We propose the *Synthetic Counterfactual Network* method for imputing how a network would respond to a treatment intervention observed only in other networks. Researchers observe a target network and a fixed number of donor networks before the intervention, and only the donors are observed after the intervention, with the estimand being the post-intervention counterfactual link probability of each pair in the target network. However, the challenges arise from the fact that networks differ in size, lack node correspondence, and consist of two entangled sources of heterogeneity: the difference in distributions of individual latent characteristics and the network-specific formation environments. The proposed method resolves these issues in the following steps: with the large network setting, we first denoise each adjacency matrix by kernel neighborhood smoothing; next, align each donor to the target through the Gromov–Wasserstein optimal coupling, which softly matches individuals by comparing within-network differences in their connection profiles; then choose convex weights so that the aligned donors reproduce the target before the intervention; finally, carry the learned soft alignment and network weights forward to impute the target's post-intervention counterfactual network. In the population problem, we provide conditions for the identification of the counterfactual network. We then establish uniform consistency of the finite-sample estimator and the max-norm rate of convergence.

*I thank Ben Deaner, Martin Weidner, Andrei Zeleneev, and the participants in the UCL Econometrics Brown Bag Seminar, the UCL-CeMMAP Ph.D./PostDoc Econometrics Research Day, and the CeMMAP Seminar for their valuable comments and suggestions. E-mail: chen-wei.hsiang.21@ucl.ac.uk.

1 Introduction

Empirical analysis involving network interactions is challenging, and such network structure matters for empirical analysis in multiple ways. First, the network itself may be the object of interest, with researchers studying who interacts with whom, the density of connections, or the formation of groups. Second, individuals' outcomes may depend on the characteristics, behavior, or outcomes of those to whom they are connected, as in the peer effects literature. Third, aggregate outcomes may depend on how individuals are connected; for example, an economy's aggregate production may depend on the organization of supplier–customer relationships across firms. Hence, understanding network structure is important both for studying interactions themselves and for analyzing outcomes that depend on them. However, when a treatment is introduced or a policy is implemented, it may alter the underlying network structure on which these outcomes depend. Therefore, studying the resulting post-intervention counterfactual network is crucial for causal analysis.

To further illustrate how the treatment intervention can reshape a network, Banerjee et al. (2024), for example, study the impact of the introduction of formal microfinance in 75 villages in Karnataka, India. They find that exposure to formal credit shrinks the informal lending and risk-sharing network: relative to villages without microfinance exposure, network density falls by about 11%. However, an average change in link probabilities may fail to capture a substantial rearrangement of the network. As discussed in their work, households least likely to borrow from microfinance experience larger losses in informal borrowing and risk sharing, and links between two households unlikely to borrow are at least as likely to disappear as links involving households likely to borrow. A policy implementation can create some links and destroy others. Auerbach and Cai (2025) further study such heterogeneity in treatment-induced network disruption. By comparing a specific pair of treated-control villages that happen to have similar village-level characteristics, and assuming weak isomorphism between these two villages' networks, they bound the negative effects on connection change of $[8.30\%, 11.93\%]$ and positive effects on connection change of $[7.69\%, 11.03\%]$, resulting in the total disruption of $[15.99\%, 22.96\%]$, despite the much smaller change suggested by the average effect. These heterogeneous changes in the post-intervention network structure immediately motivate our question:

"Can we go beyond the bounds on average network disruptions
and estimate the entire counterfactual network down to the pair level?"

To tackle this question of estimating counterfactual network, rather than relying on a single treated-control village pair with similar village-level characteristics and assuming weak isomorphism, we utilize different identification strategy and exploit more information from multiple

source (donor) networks. As illustrated in the above informal lending and risk-sharing network example, researchers have access to data from 75 networks (villages), of which 43 are treated and 32 are control. The proposed *Synthetic Counterfactual Network* method uses these additional source networks to help pin down the target counterfactual network.[1]

With a finite number of source networks, the conceptual idea is motivated by the synthetic control method. By first *comparing* networks in the pre-treatment period, we approximate the target network by the convex combination of source networks. That is, we construct synthetic control-like weights, such that the target can be best reproduced by the convex combination of sources:

$$\text{Network}_{\text{target}}^{(\text{pre})} \approx \sum_{g \in \text{sources}} \delta_g \cdot \text{Network}_g^{(\text{pre})}, \qquad \text{where } \delta_g \geq 0 \text{ and } \sum_{g \in \text{sources}} \delta_g = 1.$$

We then carry the same weights learned in the pre-treatment period forward to the post-treatment period and combine the sources' post-treatment networks to impute the target's post-treatment period counterfactual network:

$$\widehat{\text{Network}}_{\text{target}}^{(\text{post})} = \sum_{g \in \text{sources}} \delta_g \cdot \text{Network}_g^{(\text{post})}.$$

This concept immediately raises a fundamental difficulty. Unlike the scalar outcomes used in the traditional synthetic control method, which can be compared directly by simple subtraction, how should we *compare* networks?

Comparing networks, however, is substantially more difficult than comparing scalar outcomes. To illustrate the difficulty, consider a fixed number of networks indexed by $g = 1, ..., G$, where G denotes the target network and $g < G$ denotes source networks. For treatment status $w \in \{0, 1\}$, we model each network's adjacency matrix (link) by

$$A_{g,i,j}^{(w)} = P_{g,i,j}^{(w)} + e_{g,i,j}^{(w)} = f_g^{(w)}(\xi_{g,i}, \xi_{g,j}) + e_{g,i,j}^{(w)}, \qquad \text{for } w \in \{0, 1\}, \quad \text{for } i, j \in [N_g], \quad \text{for all } g \in [G],$$

where $P_{g,i,j}^{(w)}$ the denotes probability of individual i and j will form a link in network g, and it's determined by ξ the unobserved individual characteristics and $f_g^{(w)}$ the network formation environment. Throughout this paper, we refer to this link probability matrix $P_g^{(w)} = \left\{P_{g,i,j}^{(w)}\right\}_{i,j \in [N_g]}$ as our network.[2]

[1]For example, consider a specific untreated village (say village no. 44) as our target. We are interested in the network that would have formed among its households in the post-intervention period had the target village been exposed to microfinance. With this, the 43 villages exposed to microfinance serve as source (donor) networks. We use the information of target's pre-intervention network and the donors' pre- and post-intervention networks to impute the target's counterfactual post-intervention network.

[2]For an exchangeable network, where its joint distribution of links remains unchanged under any permutation

First, notice that networks may contain different numbers of individuals. When $N_g \neq N_G$ for $g \neq G$, their link or link probability matrices have different sizes and cannot be compared entry by entry. Even when $N_g = N_G$, there is generally no canonical node correspondence between individuals across networks because of the exchangeable nature of the network. Second, differences in link probabilities across networks may reflect two entangled sources of heterogeneity: ξ the individual characteristics and $f_g^{(w)}$ the network formation environment. The distributions of individual characteristics may differ across networks, while the network formation environments themselves may also differ. Hence, even if two individuals from two different networks have the same link probability connection profiles to the rest of their own networks, it does not imply that these two individuals are similar in terms of their unobserved individual characteristics, since the same link probability may arise under different network formation models. Together, these difficulties make it challenging to impute counterfactual link probabilities for all individual pairs in the target network.

The proposed *Synthetic Counterfactual Network* method addresses all these difficulties. Building on existing literature on denoising large networks (cf. Zhang et al., 2017; Zeleneev, 2020; Nowakowicz, 2024) and working with the estimates of the link probability matrices, the key idea is then *not to* compare the levels of link probabilities across networks directly, but instead to compare individuals through their *internal relational structure* across different networks. Within network g, for example, we measure how differently two individuals i and j connect to the rest of their own network through a dissimilarity of the form

$$\left\| P^{(0)}_{g,i,\cdot} - P^{(0)}_{g,j,\cdot} \right\|.$$

Because this comparison is made within the same network, the network-specific link formation environment is held fixed. The resulting dissimilarity therefore reflects differences in the *roles* that individuals play within network g, rather than differences in absolute link probability values across different environments.

We can similarly construct the within-network dissimilarities analogue for the target network G, $\left\| P^{(0)}_{G,k,\cdot} - P^{(0)}_{G,\ell,\cdot} \right\|$, and compare these two internal relational structures using Gromov-Wasserstein optimal transport, developed by Mémoli (2011). Rather than requiring the source and target networks to have the same size or a pre-specified node correspondence, Gromov-Wasserstein optimizes for a coupling that preserves their internal pairwise relationships as

of the individuals' labels, the Aldous-Hoover representation theorem provides a latent-variable foundation for the graphon model. This motivates representing link probabilities as $P^{(w)}_{g,i,j} = f^{(w)}_g(\xi_{g,i}, \xi_{g,j})$, where the individual latent characteristics $\xi_{g,i}$ are independently and identically distributed within each network and $f^{(w)}_g$ is a symmetric link formation function.

much as possible. The resulting optimal coupling provides a soft alignment between individuals across the two networks: one individual in a source network may correspond to several individuals in the target network with different weights. Under our cross-network conditions, this alignment allows us to represent each source network on the target population. The resulting aligned source-g network has a natural interpretation: *what would individuals i, j, originally from the target network G, look like if they were placed in source network environment g?*

Once all source networks are represented on this common target population space, we combine them using synthetic control-like weights. In the pre-treatment period, we choose weights on the simplex[3] so that the convex combination of the aligned source networks approximates the target network. These weights capture how the target's network formation environment can be represented by the source environments. We then carry the learned cross-network relation and the same convex weights forward to the post-treatment period. The resulting estimator recovers the full matrix of counterfactual link probabilities for the target network. By tracing each target pair's link probability across treatment statuses, the proposed method allows researchers to study the counterfactual network itself or to use it as an input in a subsequent analysis of another outcome.

We study identification in the population problem and establish large-sample guarantees for the proposed method. Two crucial assumptions underpin the identification argument. First, a convex hull condition requires that, under an appropriate cross-network alignment, the target's link formation function can be expressed as a convex combination of the source functions, with the same weights across individual pairs and treatment statuses. This condition allows weights identified from the pre-intervention network to remain valid for constructing the post-intervention counterfactual network. Second, a comonotonicity condition requires that the source and target networks order their pairwise dissimilarities consistently. These dissimilarities measure how differently two individuals connect to others within their own network. Their magnitudes may vary across networks, but the ordering of corresponding pairs must not reverse. This condition makes the cross-network alignment compatible with Gromov-Wasserstein optimal transport. Under these assumptions and additional regularity conditions, the oracle estimator exactly recovers the target's counterfactual network in the population problem. Building on this result, we further establish uniform consistency of the finite-sample estimator under additional regularity conditions, as network sizes grow while the number of networks remains fixed.

[3]Restricting the weights to the simplex provides a compact feasible set and ensures that convex combinations of link probabilities stay in $[0, 1]$. It also provides a natural interpretation: after alignment, the target's link formation environment is approximated by a convex combination of the source environments.

Literature Review

This paper contributes to the literature on counterfactual estimation using synthetic controls. The classical synthetic control method, introduced by Abadie and Gardeazabal (2003) and Abadie et al. (2010), constructs a convex weighted combination of donor units that approximates the treated unit before treatment implementation, and applies the same weights to the donors' post-intervention time period outcomes to estimate the treated unit's counterfactual. We refer the reader to Abadie (2021) for a more comprehensive review of the synthetic control method.

Recent work extends synthetic controls beyond scalar outcomes. Gunsilius (2023) constructs synthetic counterfactual distributions by combining donor quantile functions in Wasserstein space. Lin (2025) likewise fits pre-treatment quantile functions in a functional space and studies the connection to functional predictor regression. Kurisu et al. (2025) further generalize such idea and develop geodesic synthetic controls for random-object and functional data outcomes using weighted Fréchet means, while Okano and Kurisu (2026) develop functional synthetic controls for metric spaces admitting isometric embeddings into Hilbert spaces with closed, convex images. These approaches provide important precedents for estimating the counterfactual object. The network examples in the latter two papers include graph Laplacians of a fixed dimension, compared in common matrix coordinates.

Our setting instead concerns heterogeneous network populations with no observed node correspondence, potentially different sizes and latent distributions, and different network-formation environments. We estimate a coupling of latent node characteristics from internal connection-profile geometry before fitting synthetic weights. Bateni et al. (2025) develop a closely related transport approach that learns a cross-site transformation from control-group data using a regularized fused Gromov-Wasserstein criterion and applies it to source treated observations to synthesize the target treated distribution. Both settings require restrictions under which alignment learned from observed outcomes remains valid for counterfactual outcomes. Our estimand is a dyadic link probability, recovered by combining multiple aligned source environments.

The transfer across network environments also connects our problem to the extrapolation of treatment effects across heterogeneous sites. Menzel (2023) uses functional information in baseline outcomes to predict target-site treatment effects while allowing for unobserved site heterogeneity. Our setting concerns dyadic outcomes and requires alignment of latent node characteristics before information from different populations can be combined.

The studied network model is built on top of a latent factor structure, in which link probabilities depend on unobserved individual characteristics through a network-specific link for-

mation function. Due to the Aldous-Hoover representation (cf. Aldous, 1981; Hoover, 1979) of exchangeable networks, this latent factor structure provides a foundation for modeling dyadic interactions through general graphon functions. Such a formulation further connects to the literature on network formation with unobserved heterogeneity. Graham (2017) incorporates individual-specific degree heterogeneity into a model of link formation. Broader treatments of network formation and its econometric analysis are provided by De Paula (2020) and Graham and De Paula (2020). Estimating the underlying link probabilities nonparametrically without structural assumption also connects our approach to matrix estimation and graphon denoising. Chatterjee (2015) develops universal singular value thresholding for noisy and partially observed matrices, with guarantees covering both low-rank matrices and general graphon models. Gao et al. (2015) establish minimax rates for link probability matrix estimation under stochastic block and Hölder's graphon models. Zeleneev (2020) further studies identification and estimation of network models in which unobserved individual characteristics interact through an unknown function, allowing for latent homophily beyond additive fixed effects.

More specifically, the denoising step of our proposed method builds on estimating link probabilities using similarities between individuals' connection profiles. Zhang et al. (2017) develop a neighborhood smoothing method based on an estimable pseudo-distance, avoiding direct estimation of the non-identifiable latent characteristics. Related utilization of pseudo-distances to measure latent similarity includes controlling for unobserved heterogeneity in outcome regressions (cf. Auerbach, 2022), counterfactual imputation in nonlinear panel models (cf. Deaner et al., 2025; Feng, 2020), and identification and estimation of network models with nonparametric unobserved heterogeneity (cf. Zeleneev, 2020). Adapting the pseudo-distance-based kernel smoothing approach of Deaner et al. (2025) to network data, we establish an element-wise uniform convergence rate under our regularity and informativeness conditions.

The cross-network alignment step connects our proposed method to applications of optimal transport in causal inference. Lin et al. (2023) use Wasserstein geometry to define and estimate causal effect maps for distribution-valued outcomes. Van Dijcke (2025) develops regression discontinuity methods for such outcomes using local Fréchet regression in Wasserstein space, while Adjaho and Christensen (2022) study policy choice that is robust to distributional shifts within Wasserstein neighborhoods. Gunsilius (2025) reviews broader connections between optimal transport and causal identification, including instrumental variables, difference-in-differences, and synthetic controls. This paper contributes to the usage of Gromov-Wasserstein optimal transport and aligning across network populations. Gromov-Wasserstein couplings match their internal relational structure and transfer source link-probability functions to the

target population, where synthetic weights combine the transferred functions.

Our alignment step builds on the Gromov-Wasserstein framework developed by Mémoli (2011), which compares populations through their internal pairwise relationships without requiring a common coordinate system or an observed correspondence between individuals. Mémoli (2014) provides a brief overview, and Sturm (2023) studies the associated geometry of the space of metric measure spaces. Chowdhury and Mémoli (2019) extend Gromov-Wasserstein comparisons to weighted, potentially directed networks. Related work by Chowdhury and Mémoli (2023) develops a generalized Gromov-Hausdorff distance for comparing networks and shows how closeness under this distance controls differences in network statistics. Dumont et al. (2025) study when an optimal Gromov-Wasserstein coupling can be represented by a Monge map, assigning each source point to a single target point rather than spreading its mass across several points. However, matching internal pairwise relationships alone may leave several equally good correspondences. Demetci et al. (2024) introduce augmented Gromov-Wasserstein transport, which incorporates observed features and prior knowledge to guide the alignment and help distinguish between matches that internal relationships alone cannot separate. In our setting, alignment is based on connection-profile dissimilarities, whereas the objects transferred through the resulting couplings are link-probability functions. Our identification conditions ensure that alternative optimal alignments yield the same projected probability functions, allowing counterfactual recovery even when the optimal coupling is not unique.

The convex combination of aligned sources connects more specifically to Gromov-Wasserstein barycentric representations. Peyré et al. (2016) compare distance or kernel matrices of different sizes using a softassign criterion, which measures how well their pairwise relationships agree under a soft matching of the underlying points. They then use this criterion to construct Gromov-Wasserstein barycenters, providing a method for averaging these matrices without requiring a known correspondence. Xu et al. (2021) use Gromov-Wasserstein barycenters to estimate a shared link formation function from unaligned graphs generated by a common graphon. In contrast, our proposed method combines sources with potentially different link formation functions to reproduce a target network before intervention, then carries the learned alignment and network weights forward to impute its post-intervention counterfactual link probabilities. Chowdhury and Needham (2020) develop Gromov-Wasserstein methods for computing average networks. Their focus is on summarizing observed networks, whereas ours is on imputing a target network under an unobserved treatment status.

The work of Martín et al. (2026) is most closely related to our weight-fitting criterion, as they study the inverse, or analysis, problem of recovering the barycentric weights of an observed tar-

get network relative to known templates. When the target belongs to the Gromov-Wasserstein barycenter family, their fixed-point characterization yields a simplex-constrained quadratic program based on Frobenius norm. This formulation provides the immediate methodological precedent for our proposed method. In contrast, we study conditions under which the target network belongs to the Gromov-Wasserstein barycenter family generated by a finite collection of source networks with different latent characteristic distributions and link formation environments. We further establish statistical guarantees for the resulting counterfactual estimator.

Our statistical analysis also relates to finite-sample convergence results for empirical Gromov-Wasserstein estimation. Zhang et al. (2024) establish convergence rates for the $(2, 2)$ Gromov-Wasserstein cost between distributions on Euclidean spaces and its entropically regularized version. Kato and Wang (2025) further study the unregularized cost, establishing limits on estimation accuracy and bounds on the probability of large estimation errors. Our problem requires additional analysis because accurately estimating a transport cost does not by itself ensure accurate counterfactual link probabilities. We therefore examine how errors in denoising, cross-network alignment, and weight estimation affect the final imputation. Our assumptions ensure that alignments that approximately preserve the networks' internal relationships also yield accurate transferred link probabilities, while empirical Wasserstein bounds (cf. Fournier and Guillin, 2015) help account for variation from sampling individuals. Together, these arguments establish uniform consistency and an explicit convergence rate for the largest absolute error across target pairs, as network sizes grow while the number of source networks remains fixed.

The remainder of the paper is organized as follows. Section 2 introduces the proposed Synthetic Counterfactual Network method. Section 3 formulates the population problem and discusses the required assumptions for the identification. Section 4 establishes identification in the population problem. Section 5 develops the large sample theory. Section 6 examines the finite sample performance through Monte Carlo simulations. Section 7 applies the proposed method to empirical data. Finally, Appendix contains proofs, auxiliary results, discussions on the assumptions, and additional simulation results.

2 Synthetic Counterfactual Networks

2.1 Setting

Consider a setting in which there are a finite number of G different network environments. For simplicity, we will focus on two periods setting: a pre-treatment period and a post-treatment

period, though this can be easily generalized to multiple periods and staggered settings. In the pre-treatment period, researchers can observe the adjacency matrices of all G networks. However, in the post-treatment period, which is under different intervention schemes, only the first $G-1$ of these networks can be observed. We are interested in understanding what the G-th network would look like if it were subjected to the same intervention. With this context in mind, the model of the dynamic network adjacency matrix is defined as follows

$$\begin{aligned} A_{g,i,j}^{(w)} &= P_{g,i,j}^{(w)} + e_{g,i,j}^{(w)} \\ &= f_g^{(w)}(\xi_{g,i}, \xi_{g,j}) + e_{g,i,j}^{(w)} \\ &= f^{(w)}(\xi_{g,i}, \xi_{g,j}, \gamma_g) + e_{g,i,j}^{(w)} \end{aligned} \tag{1}$$

where i and j are the individual indices, g represents the network index, and w indicates the treatment status. Here, the link/edge $A_{g,i,j}^{(w)} \in \{0,1\}$ is characterized by the link/edge probability $P_{g,i,j}^{(w)}$ plus the Bernoulli error $e_{g,i,j}^{(w)}$. The function $f_g^{(w)}(\cdot,\cdot) = f^{(w)}(\cdot,\cdot,\gamma_g)$ is known as the graphon, a symmetric measurable function that maps a pair of latent characteristics to a link probability (cf. Lovász, 2012). Conditional on the latent characteristics, network g is a random graph with independent links. We refer to $f^{(w)}$ as the common graphon function, since it is shared across networks and differs only through γ_g, and to $P_{g,i,\cdot}^{(w)}$ as individual i's graphon slice. In this model, the link probability is determined by unobserved individual latent characteristics/factors ξ as well as the network-specific effect γ. As mentioned, for simplicity, we will only consider the case where $w \in \{0,1\}$.[4] Note that different networks have different individuals, and naturally, their latent characteristics are also distributed differently across various networks g.

We treat the network-specific effects $\{\gamma_g\}_{g=1}^G$ as fixed and specify the sampling process/data generating process conditional on these effects in Assumption 1.

Assumption 1 (Sampling design)**.**
The following three conditions hold:

(i) For each $g \in \{1, ..., G-1, G\}$, the latent factors $\xi_{g,1}, ..., \xi_{g,N_g}$ are independent draws from the law μ_g. On top of that, all latent factors across different networks are independent.

[4] In a conventional potential outcomes framework, network would be indexed by both treatment status and time. To simplify notation, we combine these into a single index: (0) denotes the pre-intervention baseline, and (1) denotes the post-intervention regime whose target link probabilities we seek to impute. In the setting described here, the missing target is the network that would arise if the target received the intervention observed in the sources. The method can also be used to impute the target's post-intervention network in the absence of treatment, as in conventional synthetic control applications, using sources observed without treatment in the post-intervention period, provided the corresponding identification conditions hold.

(ii) Conditional on the latent factors, the observed adjacency matrices follow model (1), with $\mathbb{E}\left[e^{(w)}_{g,i,j} \mid \xi_{g,1}, ..., \xi_{g,N_g}\right] = 0$, and satisfy symmetry $A^{(w)}_{g,i,j} = A^{(w)}_{g,j,i}$. For each w, conditional on all latent factors, the adjacency entries are mutually independent across distinct unordered pairs, including pairs that share an individual, and across networks.

(iii) With the number of networks G fixed, assume the large network setting, where

$$N_{\min} := \min_{g \in \{1,...,G-1,G\}} N_g \to \infty.$$

Assumption 1 allows dependence between links through shared latent characteristics, including differences in individuals' propensities to form links and preferences for connecting with similar individuals. However, it requires that distinct links be independent conditional on these characteristics. This restriction excludes link formation mechanisms that generate residual dependence between links, such as a link forming in response to other realized links, binding constraints on individuals' numbers of connections, or shocks that jointly affect several links and are not captured by the latent factors. The assumption does not rule out clustering. That is, individuals with similar latent characteristics may be more likely to connect, generating triangles even under conditional independence.

Under this setting, researchers are able to observe the pre-treatment period adjacency matrices for all networks $\left\{A^{(0)}_{g,i,j}\right\}_{i,j \in \{1,...,N_g\}}$ for $g \in \{1, ..., G-1, G\}$ and the post-treatment period adjacency matrices for the source networks $\left\{A^{(1)}_{g,i,j}\right\}_{i,j \in \{1,...,N_g\}}$ for $g \in \{1, ..., G-1\}$, but are not able to observe post-treatment target network $\left\{A^{(1)}_{G,i,j}\right\}_{i,j \in \{1,...,N_G\}}$. Given this, the estimand we are aiming for is the link probability matrix for the target network G

$$\left\{P^{(1)}_{G,i,j}\right\}_{i,j \in \{1,...,N_G\}}.$$

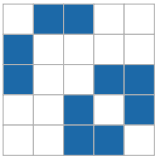
2.2 Challenges and Conceptual Solutions

Before laying down the details of the proposed method, we first discuss the challenges of this setting and the conceptual solution to them. Firstly, with the following two networks g and G

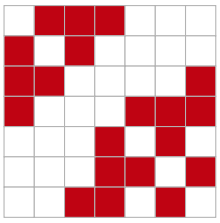


Network g Network G

researchers are able to observe the adjacency matrices $\left\{A^{(0)}_{g,i,j}\right\}_{i,j \in [N_g]}$ and $\left\{A^{(0)}_{G,k,\ell}\right\}_{k,\ell \in [N_G]}$

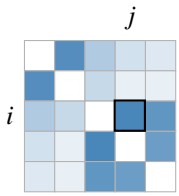


Adjacency matrix of g

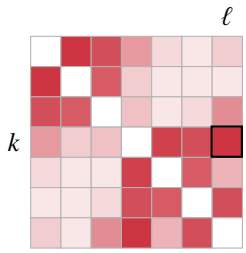


Adjacency matrix of G

which, under model (1), are the link probabilities plus the noise. Comparing the observed adjacency matrices directly is therefore complicated by the realization noise $e^{(0)}$. That is, the differences between two links reflect both the underlying link probabilities and the random link realizations. Furthermore, as studied in Cai (2026), link probabilities can sometimes better capture the economically relevant interaction propensities.[5] Hence, we would like to first denoise the observed adjacency matrices and work on the estimated link probabilities $\left\{\hat{P}^{(0)}_{g,i,j}\right\}_{i,j\in[N_g]}$ and $\left\{\hat{P}^{(0)}_{G,k,\ell}\right\}_{k,\ell\in[N_G]}$



Link probabilities of g



Link probabilities of G

which can be estimated well under the large network setting in Assumption 1 (iii). The details of the denoising step are deferred to the next subsection. Hence, the question becomes whether these denoised link probabilities are enough to answer what the target network G would look like under the intervention.

To answer this, we first look within a single network. From the network estimation literature, we learn that within the same network, say network g, we can compare the *graphon slices* of two individuals i and j, that is, their rows in the link probability matrix

$\hat{P}^{(0)}_{g,i,\cdot} = i$  vs $\hat{P}^{(0)}_{g,j,\cdot} = j$ 

[5]The network literature has focused on different objects depending on the application. The link prediction literature aims at recovering the realized links $A^{(w)}_{g,i,j}$ themselves, which is natural when the goal is, for instance, to study how an actual connection affects outcomes, as in the peer effects literature. In contrast, the graphon estimation literature (cf. Zhang et al., 2017) aims at the underlying link probabilities $P^{(w)}_{g,i,j}$. In this paper, we focus on the link probabilities. Under model (1), a realized link is the link probability plus a Bernoulli error draw $e^{(w)}_{g,i,j'}$ and for the unobserved post-treatment target network such a draw never takes place. Hence, the counterfactual is only identified up to its link probability. This is already sufficient for many applications, such as the counterfactual expected degree of an individual or the expected number of links between two groups, and any counterfactual binary network can be obtained from $\hat{P}^{(1)}_G$ afterwards by thresholding or by sampling, which we leave to the practitioner.

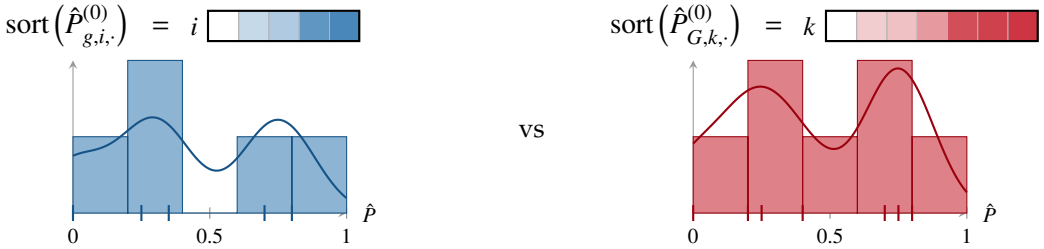
In this case, individuals i and j in network g are similar if they happen to have similar connection profiles, that is, if they tend to connect to the same individuals with similar probabilities. The natural next question is whether we can extend this concept across networks and *compare* different networks' link probability matrices directly. That is, can we, for each target individual $k \in \{1, ..., N_G\}$, find a counterpart $i \in \{1, ..., N_g\}$ in the source network g by comparing the following two graphon slices

$$\hat{P}^{(0)}_{g,i,\cdot} = i \quad \text{vs} \quad \hat{P}^{(0)}_{G,k,\cdot} = k$$

and thereby match i in network g to k in network G?

Unfortunately, this route does not work and we cannot directly compare $\hat{P}^{(0)}_{g,i,\cdot}$ and $\hat{P}^{(0)}_{G,k,\cdot}$ to match individuals across networks, and this is due to two reasons. First, the network sizes are most likely to be different. That is, $N_g \neq N_G$ in general, which makes $\hat{P}^{(0)}_g$ and $\hat{P}^{(0)}_G$ matrices of different sizes, and their graphon slices vectors of different lengths. On top of that, there is no cross-network entry-by-entry correspondence between them to begin with because of the exchangeable nature of the networks. Even in the extremely rare case where $N_g = N_G$, the indices carry no meaning across networks, since the individuals in network g are simply different people from those in network G. Second, and more importantly, even if the (estimated) link probabilities are the same, i.e., $\hat{P}^{(0)}_{g,i,\cdot} = \hat{P}^{(0)}_{G,k,\cdot}$, the individual characteristics behind them can be very different between the two.

To illustrate this second challenge, consider a simple toy example in which, after sorting the entries of their graphon slices, individual i in network g and individual k in network G happen to have similar empirical distributions of link probabilities



Now suppose that, behind these two graphon slices, the two networks look as follows:

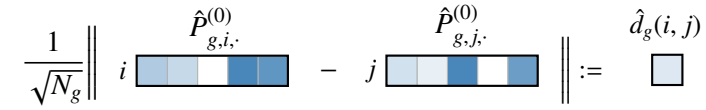
	Network g	**Network** G
Individual latent characteristics	Age: $\{20, \overset{i}{22}, 24, 25, 27\}$	Age: $\{21, 27, 40, 54, \overset{k}{57}, 62, 70\}$
Network-specific effect	Sparse living environment	Dense living environment
Link probability formation	$\hat{P}^{(0)}_g$ mostly driven by *physical distances*	$\hat{P}^{(0)}_G$ mostly driven by *age differences*

That is, the 22-year-old i in network g and the 57-year-old k in network G have nearly identical connection profiles, yet they are very different individuals, and their links form for very

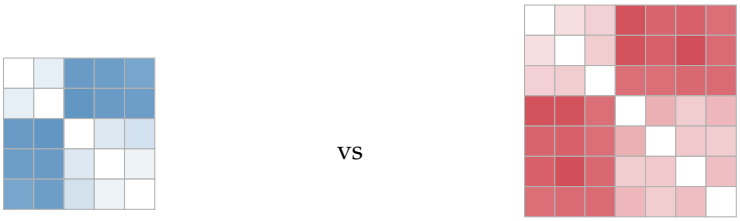
different reasons. More formally, under model (1), the source pair forms the link probability $f^{(0)}(\xi_{g,i}, \xi_{g,j}, \gamma_g)$ while the target pair forms link probability $f^{(0)}(\xi_{G,k}, \xi_{G,\ell}, \gamma_G)$. Because the network-specific effects γ_g and γ_G may be different, the same link probability can arise from very different latent characteristics $(\xi_{g,i}, \xi_{g,j})$ and $(\xi_{G,k}, \xi_{G,\ell})$. In other words, the heterogeneity in the individual characteristics ξ and the heterogeneity in the network-specific effects γ are entangled in the network formation, and comparing two link probabilities cannot separate them. To resolve this issue, the core idea of the proposed method is to *change the object of comparison and fix one source of heterogeneity*.

Instead of comparing link probabilities across networks, we compare the *internal relational structures* of the networks. Specifically, for a pair of individuals i, j in network g, we first take the difference between their graphon slices

$$\frac{1}{\sqrt{N_g}} \left\| \hat{P}^{(0)}_{g,i,\cdot} - \hat{P}^{(0)}_{g,j,\cdot} \right\| := \hat{d}_g(i,j)$$



where $\hat{d}_g(i,j)$ measures the (normalized) similarity between i and j by looking at if they happen to have similar connection profiles in network g. As studied in the graphon estimation literature (cf. Zhang et al., 2017), a small $\hat{d}_g(i,j)$ means that i and j have similar connection profiles to the rest of their own network, and, under some regularity conditions, it also means that they have similar latent characteristics $\xi_{g,i}$ and $\xi_{g,j}$. Collecting the similarity measure for all $i, j \in [N_g]$, we can form a similarity matrix $\left\{\hat{d}_g(i,j)\right\}_{i,j\in[N_g]}$ to measure the *internal geometric structure of network* g. Since this internal geometric structure is measured within a single network g, the network-specific effect γ_g is held the same. We can construct the internal geometric structure of network G, $\left\{\hat{d}_G(k,\ell)\right\}_{k,\ell\in[N_G]}$, in a similar fashion.



`Internal structure of g`
`measured by` $\hat{d}_g(i,j)$ $\forall i, j \in [N_g]$

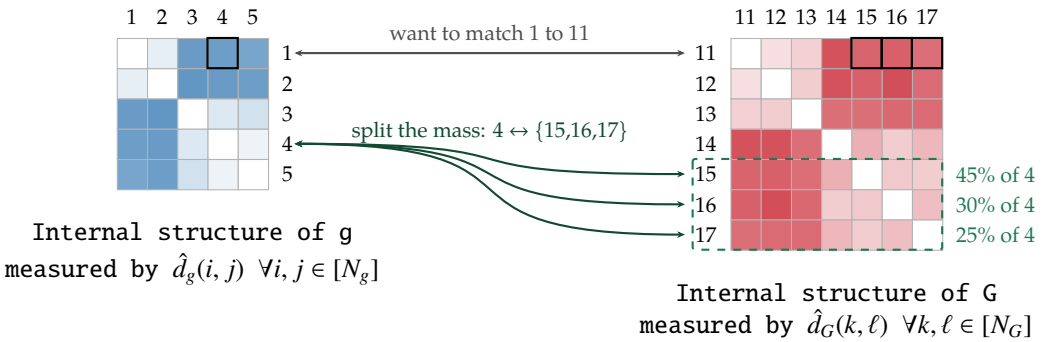
`Internal structure of G`
`measured by` $\hat{d}_G(k,\ell)$ $\forall k, \ell \in [N_G]$

Now, since one source of the heterogeneity, network-specific effect, is fixed, we can then compare two networks' internal structures. Such comparison has the following advantage. Under model (1), the two rows $\hat{P}^{(0)}_{g,i,\cdot}$ and $\hat{P}^{(0)}_{g,j,\cdot}$ are generated under the same network-specific effect γ_g, and hence, their difference reflects only how different $\xi_{g,i}$ and $\xi_{g,j}$ are, as seen through the common environment γ_g. That is, the internal distance measures how different two individuals are

relative to their own network, and γ_g is held fixed within each comparison rather than entangled with ξ. With these internal relational structures, how can we utilize them and align individuals across different networks?

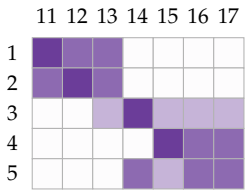
Although the network formation environment is held fixed using the internal relational structures, the other challenges mentioned earlier remain. There is still no cross-network entry-by-entry correspondence between $\left\{\hat{d}_g(i,j)\right\}_{i,j\in[N_g]}$ and $\left\{\hat{d}_G(k,\ell)\right\}_{k,\ell\in[N_G]}$, and the two matrices are still of different sizes. Hence, a one-to-one matching of individuals is unavailable. We tackle this issue by *softly aligning* the two networks. Specifically, the soft alignment is based on the following rule: a pair (i,j) in the source network g is a good counterpart for a pair (k,ℓ) in the target network G if the two pairs sit at similar *relative locations* within their respective networks.

To illustrate, label the individuals of network g by $1,...,5$ and the individuals of network G by $11,...,17$.



Suppose we want to match individual 1 in network g to individual 11 in network G. Notice that, in network g, individuals 1 and 4 are far apart, as $\hat{d}_g(1,4)$ is large, while in network G, individuals 11 and 15 are also far apart, as $\hat{d}_G(11,15)$ is large. If 1 is matched to 11, then 4 should go to someone equally far from 11 inside network G, for example, 15. However, $\hat{d}_G(11,16)$ and $\hat{d}_G(11,17)$ are of similar magnitude as well, which means that, in network G, individuals 16 and 17 sit just as far from 11 as 15 does. Hence, the alignment should be *soft*: individual 4 is shared out over 15, 16, and 17 in proportions. In general, each individual in one network is matched to a distribution over the individuals in the other network, rather than to a single one, with more weight placed on those that occupy a similar *relative position* within their own network.

Optimizing for such soft alignment over all pairs of pairs, we obtain the alignment $\hat{\pi}^{g,G}$, an $N_g \times N_G$ matrix whose rows are indexed by the individuals $i \in \{1,...,N_g\}$ of network g, whose columns are indexed by the individuals $k \in \{1,...,N_G\}$ of network G, and whose (i,k) entry records how much of individual i is matched to individual k.



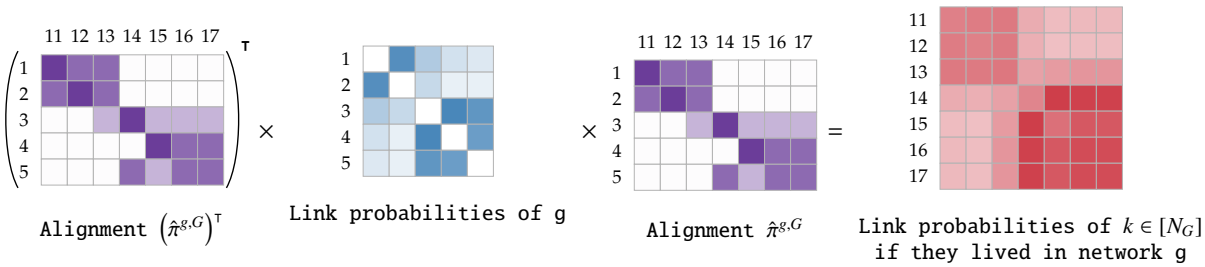
Alignment $\hat{\pi}^{g,G}$

This concept of comparing two spaces through their within-space distances, rather than through the points themselves, is known as the Gromov-Wasserstein optimal transport problem in the optimal transport literature (cf. Mémoli, 2011), which provides a well-established tool to carry out such matching in practice. The alignment $\hat{\pi}^{g,G}$ is precisely the Gromov-Wasserstein optimal coupling, and such matching neither requires $N_g = N_G$ nor requires any cross-network comparability of the latent characteristics.

With this soft alignment, captured by the Gromov-Wasserstein optimal coupling, we are able to ask and answer the following question:

"What would the individuals $k, \ell \in \{1, ..., N_G\}$ look like if they lived in network environment g?"

Using $\hat{\pi}^{g,G}$, the link probabilities of network g can be transported into the coordinates of the individuals of network G



Alignment $\left(\hat{\pi}^{g,G}\right)^{\intercal}$ | Link probabilities of g | Alignment $\hat{\pi}^{g,G}$ | Link probabilities of $k \in [N_G]$ if they lived in network g

where (up to a normalization) the resulting $N_G \times N_G$ matrix collects the link probabilities among the individuals $k \in \{1, ..., N_G\}$ of the target network as if they lived in network g.

By answering this question for every source network $g \in \{1, ..., G-1\}$, each source network is transported into the coordinates of the target network. Once all source networks are expressed in the same coordinates, they can be compared with the target network entry by entry and combined, in the same spirit as the synthetic control method of Abadie et al. (2010), to reproduce the target network. The next subsection lays down the details of each step.

2.3 Algorithm of the Proposed Method

We further concretize the aforementioned conceptual solution, where the proposed Synthetic Counterfactual Network method utilizes the technique in the optimal transport literature.

In order to align individuals across different networks, we couple each donor $g \in \{1, ..., G-1\}$ to the target G with the Gromov-Wasserstein optimal coupling $\pi^{g,G}$, which is a soft matching based on the internal structures of the networks. And with such alignments, we align and project individuals across different networks. After that, we optimize for convex weights δ via the Gromov-Wasserstein barycenter, so the source networks $g \in \{1, ..., G-1\}$ reproduce the target network G. Lastly, by carrying the same alignment and weights we learn in the pre-treatment period to the post-treatment period, we can pin down and estimate the desired $P_G^{(1)}$.

Step 0. Denoising the adjacency matrices. Every subsequent step of the proposed method operates on estimated link probabilities rather than on the raw links themselves. The reason is that if we use the realized link, which is a single Bernoulli realization, to study the alignment across different networks, the results will be complicated by the noise $e_{g,i,j}^{(w)}$ rather than only through the signal $P_{g,i,j}^{(w)}$. Thus, we begin by denoising each observed adjacency matrix with the estimate of its underlying link probability matrix.

To conceptualize the idea, suppose that the latent factors $\xi_{g,1}, \ldots, \xi_{g,N_g}$ were actually observable by the researcher. Under the condition of the link probability $P_{g,i,j}^{(w)} = f_g^{(w)}(\xi_{g,i}, \xi_{g,j})$ is a sufficiently smooth function of the latent factors, the link probability between i and j could be easily estimated by the Nadaraya-Watson type estimator

$$\hat{P}_{g,i,j}^{(w)} = \frac{\sum_{k=1}^{N_g} K\left(\left\|\xi_{g,i} - \xi_{g,k}\right\| / h\right) A_{g,k,j}^{(w)}}{\sum_{k=1}^{N_g} K\left(\left\|\xi_{g,i} - \xi_{g,k}\right\| / h\right)}$$

where $K(\cdot)$ denotes the kernel and h denotes the bandwidth.

However, in practice, the latent factors are not observable, and under model (1) they are also not identifiable without further restriction on the functional form of f. To tackle this question, following Zhang et al. (2017), Deaner et al. (2025), Zeleneev (2020), and Feng (2020), we work around this by replacing the unobservable ℓ_2 distance $\|\xi_{g,i} - \xi_{g,k}\|$ with an estimable pseudo distance that can be computed directly from the adjacency matrix,

$$\hat{d}_g^{\text{pseudo}}(i,k) := \max_{\ell \neq i,k} \frac{1}{N_g} \left| \left\langle A_{g,i,\cdot}^{(w)} - A_{g,k,\cdot}^{(w)}, A_{g,\ell,\cdot}^{(w)} \right\rangle \right| \tag{2}$$

and forming the neighborhood kernel smoothing estimator

$$\hat{P}_{g,i,j}^{(w)} = \frac{\sum_{k=1}^{N_g} K\left(\hat{d}_g^{\text{pseudo}}(i,k)/h\right) A_{g,k,j}^{(w)}}{\sum_{k=1}^{N_g} K\left(\hat{d}_g^{\text{pseudo}}(i,k)/h\right)}$$

Since we are studying the symmetric networks, we further rewrite the symmetric estimator as $\hat{P}_g^{(w)} \leftarrow \left(\hat{P}_g^{(w)} + \left(\hat{P}_g^{(w)}\right)^\intercal\right)/2$ as recommended by Zhang et al. (2017).

The pseudo distance $\hat{d}_g^{\text{pseudo}}(i,k)$ is small when individuals i and k connect to the rest of the network g in the same way.[6] Applying this step to every observed adjacency matrix yields $\hat{P}_g^{(0)}$ for $g \in \{1,...,G-1,G\}$ and $\hat{P}_g^{(1)}$ for $g \in \{1,...,G-1\}$, which are the inputs to the next alignment and projection step.

Step 1. Gromov-Wasserstein alignment and projection. As discussed earlier, networks may differ in size and lack a known correspondence between individuals, preventing direct entry-wise comparison of their link probability matrices. Moreover, link probabilities reflect both individual latent characteristics and network-specific effects, so similar link probabilities across networks need not imply similar latent characteristics. Hence, we align source and target networks' individuals using within-network distances that measure differences in their connection profiles. The Gromov-Wasserstein optimal transport, developed by Mémoli (2011), provides a criterion for this soft alignment by seeking a coupling between individuals across networks that minimizes discrepancies between the corresponding within-network distances.

Concretely, we equip each pre-treatment network with the squared internal distance

$$\hat{d}_g^2(i,j) = \frac{1}{N_g}\left\|\hat{P}^{(0)}_{g,i,\cdot} - \hat{P}^{(0)}_{g,j,\cdot}\right\|_2^2 \qquad \text{for } g \in \{1,...,G-1,G\}$$

so that, within the same network, two individuals are similar if their (denoised) link probability graphon slices are close.[7] The normalization by N_g keeps the internal distances on a common scale when $N_g \neq N_G$. Let $\Pi^{g,G} = \{\pi \in \mathbb{R}_+^{N_g \times N_G} : \pi \mathbf{1}_{N_G} = \frac{1}{N_g}\mathbf{1}_{N_g}, \pi^\intercal \mathbf{1}_{N_g} = \frac{1}{N_G}\mathbf{1}_{N_G}\}$ denote the set of couplings between the two node measures.

For each source network $g \in \{1,...,G-1\}$, we solve the Gromov-Wasserstein optimization problem between the pre-treatment networks g and G,

$$\hat{\pi}^{g,G} = \underset{\pi \in \Pi^{g,G}}{\arg\min} \sum_{i,j \in [N_g]} \sum_{k,\ell \in [N_G]} \left(\hat{d}_g^2(i,j) - \hat{d}_G^2(k,\ell)\right)^2 \pi_{i,k}\,\pi_{j,\ell}$$

Here, the Gromov-Wasserstein optimal coupling $\hat{\pi}^{g,G}$ defines a soft alignment of individuals. That is, $\hat{\pi}^{g,G}_{i,k}$ is the amount of mass placed on the alignment of individual i in source network g with individual k in target network G. The objective evaluates this alignment only using within-network distances. When i is aligned with k and j with ℓ, it penalizes the squared discrepancy

[6] Under our model assumptions, the latent factors do not vary with w. Although we estimate the pseudo-distances $\hat{d}^{\text{pseudo}}g(i,k)i,k$ using pre-intervention data, they may also be estimated using the post-intervention network. We leave this choice to the practitioner.

[7] Here we only consider the case of each node carries the same amount of information/weights. But the proposed method can be trivially extended to weighted network setting.

between the distance separating i and j in the source and the distance separating k and ℓ in the target. It thus favors alignments under which individuals that are close within the source correspond to individuals that are similarly close within the target, i.e., it favors alignments that preserve within-network distances.

Two features of the Gromov-Wasserstein optimal transport problem are worth emphasizing. First, the alignment does not require a distance between an individual in the source and an individual in the target. Instead, it uses distances computed separately within each network to assess how well a proposed alignment preserves their internal relationships. Computing this alignment therefore does not require directly comparable latent characteristics across networks, and in general, the networks may even have latent characteristics defined on different spaces. Second, the alignment criterion does not depend on arbitrary node labels. With this, relabeling the nodes just relabels the internal structure matrices and the corresponding set of optimal couplings, leaving the optimal objective value unchanged.

With such an optimal coupling, we can answer the aforementioned question: what the link between two target individuals $k, \ell \in \{1, ..., N_G\}$ would look like if they lived in network environment g. Normalizing the coupling by the target weights, $\frac{1}{N_G}\hat{\pi}^{g,G}_{k,i}$ is the conditional probability that the counterpart of target individual k in network G is unit i. Averaging the source link probabilities over the counterparts of k and ℓ gives the projected link probability matrix

$$\hat{Q}^{(0)}_g = \frac{1}{(1/N_G)^2}\left(\hat{\pi}^{g,G}\right)^\intercal \hat{P}^{(0)}_g \hat{\pi}^{g,G} \in \mathbb{R}^{N_G \times N_G},$$

or in the element-wise representation

$$\hat{Q}^{(0)}_{g,k,\ell} = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \frac{\hat{\pi}^{g,G}_{i,k}}{1/N_G}\frac{\hat{\pi}^{g,G}_{j,\ell}}{1/N_G}\hat{P}^{(0)}_{g,i,j}.$$

The link probability matrix $\hat{Q}^{(0)}_g$ projects source network g onto the target network G's individual space.[8] That is, each entry averages source link probabilities over the source counterparts assigned to the corresponding target individuals. Such projection is discussed in the Gromov-Wasserstein barycenters literature (cf. Peyré et al., 2016; Martín et al., 2026), which motivates the combination of projected sources in the next step.

Step 2. Gromov-Wasserstein barycentric weights. After projection, the source link probability matrices and the target link probability matrix share the same individual index set and belong to $\mathbb{R}^{N_G \times N_G}$, allowing entry-wise comparison. However, each projected source retains its

[8] Here, each target individual has equal mass $1/N_G$. The projection also allows unequal node weights. For a target node-weight vector q with $q_k > 0$ and $\sum_{k=1}^{N_G} q_k = 1$, the coupling has target marginal q, and the denominators $1/N_G$ in the elementwise formula are replaced by q_k and q_ℓ, respectively.

own link formation environment. Because network-specific effects can generate differences in link formation environments, a single source may fail to reproduce the target even when individuals are correctly aligned. To account for this remaining heterogeneity, we choose convex weights over the sources to best reproduce the target's pre-intervention link probabilities:

$$\left\{\hat{\delta}_g\right\}_{g=1}^{G-1} = \underset{\delta \in \Delta^{G-2}}{\arg\min} \left\| \hat{P}_G^{(0)} - \sum_{g=1}^{G-1} \delta_g \, \hat{Q}_g^{(0)} \right\|_{\mathrm{F}}^2,$$

where $\Delta^{G-2} = \{\delta \in \mathbb{R}^{G-1} : \delta_g \geq 0, \sum_g \delta_g = 1\}$ is the probability simplex and $\|\cdot\|_{\mathrm{F}}$ is the Frobenius norm. This optimization problem is a quadratic programming in δ and is inexpensive to solve.

The objective has a precise geometric meaning. For fixed couplings $\{\hat{\pi}^{g,G}\}_g$ and squared loss, the Gromov-Wasserstein barycenter of the source networks with weights δ is available in closed form and equals exactly $\sum_g \delta_g \hat{Q}_g^{(0)}$ (cf. Peyré et al., 2016). This step is therefore asking which point in the barycenter space spanned by the sources is closest to the target. This question is studied as the analysis problem for Gromov-Wasserstein barycenters in Martín et al. (2026), where their Theorem 3.6, Corollary 3.7, and Remark 3.8 justify the usage of the Frobenius norm criterion above.

This step is also the direct analogue of the pre-treatment fit in the synthetic control method as in Abadie et al. (2010). The projected source networks matrices $\hat{Q}_g^{(0)}$ play the role of the donors' pre-treatment outcomes and the target network matrix $\hat{P}_G^{(0)}$ plays the role of the treated unit's pre-treatment outcome. On top of that, the simplex constraint delivers the usual interpretability of the weights. The difference is that, because the units of different networks are not comparable, the donor "outcomes" first have to be transported into the target's coordinates through alignment and projection step via Gromov-Wasserstein optimal transport before they can be combined.

Step 3. Extension to the post-treatment period. The pre-treatment period was used to learn two objects: the Gromov-Wasserstein optimal couplings $\{\hat{\pi}^{g,G}\}_{g=1}^{G-1}$, which record how the individuals in each sources network correspond and align to the individuals in the target network, and the network weights $\{\hat{\delta}_g\}_{g=1}^{G-1}$, which record how the target network's environment can be composed by the source networks. To impute the post-treatment target, we carry them to the post-treatment networks.

Projecting each denoised post-treatment adjacency matrix with the same coupling as before gives the projected link probabilities among target individuals as if they lived in network g after the intervention,

$$\hat{Q}_g^{(1)} = \frac{1}{(1/N_G)^2} \left(\hat{\pi}^{g,G}\right)^{\intercal} \hat{P}_g^{(1)} \hat{\pi}^{g,G} \in \mathbb{R}^{N_G \times N_G},$$

and combining these with the pre-treatment network weights yields the synthetic counterfactual network

$$\hat{P}_G^{(1)} = \sum_{g=1}^{G-1} \hat{\delta}_g \hat{Q}_g^{(1)}.$$

The extension from the pre-treatment to post-treatment resemble the conditions of the traditional synthetic control method, where the donor weights are estimated on pre-treatment outcomes and then carried over unchanged to the post-treatment period.

In our setting, both the couplings and the network weights are carried from the pre-intervention to the post-intervention period, and model (1), together with the assumptions in Section 3, provides the justification for doing so. The couplings can be carried over because the relative positions of individuals, both within and across networks, are governed by the latent characteristics ξ, and these are time-invariant. The intervention changes the graphon function $f^{(w)}$ but not who the individuals are, i.e., the treatment intervention can only affect one's link probabilities with others only through the network formation change. The weights are carried over because they summarize how the network-specific effect γ_G relates to the effects $\gamma_1, \ldots, \gamma_{G-1}$ of the sources. On top of that, this relation, like the latent factors, is a fixed feature of the environments rather than of the treatment. With this, the two components of the estimator separate cleanly and perform different tasks: the couplings handle the individual level heterogeneity across networks, and the network weights tackle the network-level heterogeneity. Sections 3 and 4 will then discuss the conditions under which such transportation is valid.

Combining all these steps, the following Algorithm 1 summarizes the proposed method in detail.

3 Population Problem

Section 2 presented Algorithm 1, the operational form of the proposed Synthetic Counterfactual Network method for networks with finite sampled networks. This section turns to the population problem, the idealized counterpart in which both sources of sampling variability are switched off.

Recall from model (1) and Assumption 1, the observed network is generated in two stages: individuals $i = 1, \ldots, N_g$ are first drawn with latent characteristics $\xi_{g,i} \sim \mu_g$, and then each link $A_{g,i,j}^{(w)}$ is a Bernoulli draw with probability $P_{g,i,j}^{(w)} = f^{(w)}(\xi_{g,i}, \xi_{g,j}, \gamma_g)$. The population problem removes the second stage, replacing links by link probabilities, and sends $N_g \to \infty$ in the first, replacing the finitely many sampled individuals by the whole distribution μ_g. Consequently, the $N_g \times N_g$ link probability matrix $P_g^{(w)}$ of Section 2 is replaced by its population counterpart,

Algorithm 1 Synthetic Counterfactual Network (w/ Pre-denoising)

Inputs: $\left\{A^{(0)}_{g,i,j}\right\}_{i,j\in\{1,...,N_g\}}$ for $g \in \{1, ..., G-1, G\}$; $\left\{A^{(1)}_{g,i,j}\right\}_{i,j\in\{1,...,N_g\}}$ for $g \in \{1, ..., G-1\}$

Returns: $\left\{P^{(1)}_{G,i,j}\right\}_{i,j\in\{1,...,N_G\}}$

1: Aquire $\left\{\hat{P}^{(0)}_{g,i,j}\right\}_{i,j\in\{1,...,N_g\}}$ for $g \in \{1, ..., G-1, G\}$ and $\left\{\hat{P}^{(1)}_{g,i,j}\right\}_{i,j\in\{1,...,N_g\}}$ for $g \in \{1, ..., G-1\}$ by denoising all input adjacency matrices.

2: **for all** $g \in \{1, ..., G-1\}$ **do**

3: Solve the Gromov-Wasserstein problem between pre-treatment network g and G and aquire the optimal coupling:

$$\hat{\pi}^{g,G} \leftarrow \underset{\pi^{g,G}\in\Pi^{g,G}}{\arg\min} \sum_{k,\ell\in\{1,...,N_g\}} \sum_{i,j\in\{1,...,N_G\}} \left\|\hat{d}^2_g(k,\ell) - \hat{d}^2_G(i,j)\right\|^2 \cdot \pi^{g,G}_{k,i} \cdot \pi^{g,G}_{\ell,j}$$

where $\hat{d}_g(i, j)$ is the Euclidean distance and is defined as follow:

$$\hat{d}^2_g(i,j) = \frac{1}{N_g}\left\|\hat{P}^{(0)}_{g,i,\cdot} - \hat{P}^{(0)}_{g,j,\cdot}\right\|^2_2 \qquad \text{for } g \in \{1, ..., G-1, G\}$$

4: Estimate the edge probability between $i, j \in \{1, ..., N_G\}$ as if they live in network g (instead of G)

$$\left\{\hat{Q}^{(0)}_{g,i,j}\right\}_{i\in\{1,...,N_G\};j\in\{1,...,N_G\}} \leftarrow \frac{1}{qq^\intercal} \odot \left(\hat{\pi}^{g,G}\right)^\intercal \hat{P}^{(0)}_{g,\cdot,\cdot} \left(\hat{\pi}^{g,G}\right) \in \mathbb{R}^{N_G\times N_G}$$

where q is the vector of node weights, and by default it is set to be $q_i = \frac{1}{N_G} \forall i \in \{1, ..., N_G\}$.

5: **end for**

6: Solve the Gromov-Wasserstein barycenter analysis problem

$$\left\{\delta_g\right\}_{g=1}^{G-1} \leftarrow \underset{\delta\in\Delta}{\arg\min}\, d^2_{\text{analysis}}\left(\left\{\hat{P}^{(0)}_{G,i,j}\right\}_{i,j\in[N_G]}, \sum_{g=1}^{G-1} \delta_g \left\{\hat{Q}^{(0)}_{g,i,j}\right\}_{i,j\in[N_G]}\right)$$

where in Theorem 3.6 the Frobenius norm is being used:

$$d^2_{\text{analysis}}(a,b) = \|a-b\|^2_{\text{Frobenius}}$$

7: **for all** $g \in \{1, ..., G-1\}$ **do**

8: For the post-treatment time period, the same couplings are being used

$$\left\{\hat{Q}^{(1)}_{g,i,j}\right\}_{i\in\{1,...,N_G\};j\in\{1,...,N_G\}} \leftarrow \frac{1}{qq^\intercal} \odot \left(\hat{\pi}^{g,G}\right)^\intercal \hat{P}^{(1)}_{g,\cdot,\cdot} \left(\hat{\pi}^{g,G}\right) \in \mathbb{R}^{N_G\times N_G}$$

9: **end for**

10: Lastly, impute the target edge probability

$$\left\{\hat{P}^{(1)}_{G,i,j}\right\}_{i\in\{1,...,N_G\};j\in\{1,...,N_G\}} \leftarrow \sum_{g=1}^{G-1} \hat{\delta}_g \left\{\hat{Q}^{(1)}_{g,i,j}\right\}_{i\in\{1,...,N_G\};j\in\{1,...,N_G\}}$$

the objects that generate these matrices. For each network g, write

$$P_g^{(w)}(\xi, \xi') := f^{(w)}(\xi, \xi', \gamma_g), \qquad \xi, \xi' \in \text{supp}(\mu_g),$$

for the graphon of network g and call the pair $(P_g^{(w)}, \mu_g)$ the population network g. The link probability matrix of Algorithm 1 is this graphon evaluated at the sampled individuals, $P_{g,i,j}^{(w)} = P_g^{(w)}(\xi_{g,i}, \xi_{g,j})$, and hence, the population network is exactly the object that generates the matrix and determines its law. We use the same symbol for the matrix and the graphon, with the indices or arguments distinguishing them. Since the graphon is measurable and takes values in $[0, 1]$, it belongs to $L^2(\mu_g \otimes \mu_g)$, the space of square-integrable functions under the product measure $\mu_g \otimes \mu_g$.

The population information consists of the pre-intervention network $(P_g^{(0)}, \mu_g)$ for every $g \in \{1, \ldots, G\}$ and the post-intervention network $(P_g^{(1)}, \mu_g)$ for every source $g \in \{1, \ldots, G-1\}$. Within each source, the two graphons are specified jointly on the same probability space, preserving the correspondence of individuals across periods. The object that the research aims to recover is the target's post-intervention graphon $P_G^{(1)}$.

We would like to further discuss two additional aspects of this population problem. First, in each population network, the latent characteristics ξ are never observed, and hence, $(P_g^{(w)}, \mu_g)$ is known up to a relabeling of Ξ that preserves μ_g, On top of that, neither the common graphon function $f^{(w)}$ nor the network effect γ_g is known separately.[9] This is the exact information that Algorithm 1 uses: the within-network distances and the Gromov-Wasserstein objective are invariant to such relabelings. Second, removing sampling variability eliminates the need for denoising but does not resolve the remaining identification problem. The population analogues of the alignment coupling and the network weights are defined through optimization problems that may have multiple solutions, and even unique solutions need not reproduce $P_G^{(1)}$. What remains is therefore a question of identification: under what conditions on the graphon function $f^{(w)}$, the network effects γ_g, and the individual latent characteristics distributions μ_g is $P_G^{(1)}$ uniquely determined by the population object through the proposed construction?

3.1 Population Model

We now formally define the population model introduced above. Let Ξ be a Polish space of individual latent characteristics. For each treatment status $w \in \{0, 1\}$, let

$$f^{(w)} : \Xi \times \Xi \times \mathbb{R}^{d_\gamma} \to [0, 1]$$

[9] Equivalently, what is known is the law of the exchangeable array $(P_{g,i,j}^{(w)})_{i,j \geq 1}$ generated by $\xi_{g,i} \overset{iid}{\sim} \mu_g$, which determines the graphon $P_g^{(w)}$ exactly up to measure-preserving relabeling (cf. Lovász, 2012). This is strictly more than the marginal distribution of the scalar $P_{g,i,j}^{(w)}$, which does not determine the relational structure.

be the link formation function symmetric in its first two arguments, with the network-specific effect as its third argument. Each network $g \in \{1, \ldots, G\}$ is characterized by a Borel probability measure μ_g on Ξ, describing the distribution of its individual latent characteristics, and a network-specific effect $\gamma_g \in \mathbb{R}^{d_\gamma}$. The distribution μ_g may be discrete, continuous, or mixed. Both μ_g and γ_g remain fixed across the two periods. Following the notation introduced above, the graphon of network g in period w is

$$P_g^{(w)}(\xi, \xi') := f^{(w)}(\xi, \xi', \gamma_g), \qquad \xi, \xi' \in \text{supp}(\mu_g).$$

which we view as an element of $L^2(\mu_g \otimes \mu_g)$. Together with μ_g, this graphon represents the corresponding population network.

Throughout, we maintain that $f^{(w)}(\cdot, \cdot, \gamma_g)$ is Borel measurable and bounded on $\text{supp}(\mu) \times \text{supp}(\mu)$ for every $g \in \{1, ..., G-1, G\}$ and $w \in \{0, 1\}$, where $\text{supp}(\mu) := \bigcup_{g \in \{1,...,G-1,G\}} \text{supp}(\mu_g) \subseteq \Xi$. This guarantees that all integrals below are finite and that every graphon lies in the corresponding L^2 space. Again, under the population notation, networks $g \in \{1, ..., G-1\}$ are the source networks and network G is the target network, with G being a fixed finite number. The object of interest is $P_G^{(1)}$.

With the population model defined, we then move on to the corresponding population version of the proposed method.

3.2 Population Form of the Proposed Algorithm

As implemented, Algorithm 1 operates on finite sampled networks, where, in Step 0, each observed adjacency matrix is denoised by the kernel-weighted neighborhood smoothing method (cf. Zhang et al., 2017; Deaner et al., 2025; Zeleneev, 2020), and all subsequent steps work on the resulting link probability matrix estimates. This subsection studies the population counterpart of Algorithm 1, in which each estimated link probability matrix is replaced by the true graphon and each finite coupling between nodes is replaced by a coupling between the laws μ_g and μ_G.

The main step that requires care is the Gromov-Wasserstein projection of Step 1 of Algorithm 1, which in the finite sample algorithm is the matrix operation $\frac{1}{qq^\intercal} \odot (\hat{\pi}^{g,G})^\intercal \hat{P}_g^{(0)} \hat{\pi}^{g,G}$. Recall from Section 2 that this operation averages the source link probabilities over the conditional distribution of the source counterparts of each target unit, with conditional probabilities $\hat{\pi}_{k,i}^{g,G}/q_i$. In the population counterpart, a coupling $\tilde{\pi}^{g,G} \in \Pi(\mu_g, \mu_G)$ is a joint law on $\Xi \times \Xi$ of a pair (ξ, ζ) with $\xi \sim \mu_g$ and $\zeta \sim \mu_G$, and the projection requires the conditional distribution of ξ given ζ under $\tilde{\pi}^{g,G}$. When both measures are discrete, this conditional distribution is obtained by dividing by the marginal mass, exactly as in the finite sample algorithm. But in the more general measure theoretical setting, it is delivered by the concept of *disintegration*.

Following the disintegration theorem (see, e.g., Ambrosio et al., 2008, Theorem 5.3.1), since Ξ is Polish, there exists a family $\left\{\tilde{\pi}_{\zeta}^{g,G}\right\}_{\zeta\in\Xi}$ of Borel probability measures on Ξ, unique up to a μ_G-null set of ζ, such that $\zeta \to \tilde{\pi}_{\zeta}^{g,G}(B)$ is Borel for every Borel B and

$$\mathrm{d}\tilde{\pi}^{g,G}(\xi,\zeta) = \mathrm{d}\tilde{\pi}_{\zeta}^{g,G}(\xi)\mathrm{d}\mu_G(\zeta)$$

This disintegration notation means that integrating a bounded Borel function against $\tilde{\pi}^{g,G}$ is the same as integrating it first against $\tilde{\pi}_{\zeta}^{g,G}$ and then against μ_G. The measure $\tilde{\pi}_{\zeta}^{g,G}$ is the conditional law of the source individual matched to a target individual at ζ. The precise statement is recorded as Lemma C.1 in Appendix C.

Note that the construction of a disintegration requires no optimality of $\tilde{\pi}^{g,G}$, as any coupling can be disintegrated. In Algorithm 2 below, $\tilde{\pi}^{g,G}$ will be the coupling returned by the (population version of the) Gromov-Wasserstein optimization step. With the conditional laws in hand, the population projection is defined by integrating the source graphon against them.

Definition 1 (Gromov-Wasserstein projection)**.**
For $w \in \{0,1\}$ and $\zeta, \zeta' \in \mathrm{supp}(\mu_G)$, define

$$\tilde{Q}_g^{(w)}(\zeta,\zeta') := \int_{\Xi}\int_{\Xi} P_g^{(w)}(\xi,\xi')\mathrm{d}\tilde{\pi}_{\zeta}^{g,G}(\xi)\mathrm{d}\tilde{\pi}_{\zeta'}^{g,G}(\xi').$$

The function $\tilde{Q}_g^{(w)}$ is the link probability among individuals with latent characteristics drawn from the target law μ_G, obtained by transporting $P_g^{(w)}$ through the coupling $\tilde{\pi}^{g,G}$. It serves as the population counterpart of $\hat{Q}_g^{(w)}$ in Algorithm 1. It also shares the same interpretation, i.e., if (ξ,ζ) and (ξ',ζ') are independent draws from $\tilde{\pi}^{g,G}$, then

$$\tilde{Q}_g^{(w)}(\zeta,\zeta') = \mathbb{E}\left[P_g^{(w)}(\xi,\xi') \,\middle|\, \zeta,\zeta'\right] \qquad \text{almost surely.}$$

Hence, $\tilde{Q}_g^{(w)}(\zeta,\zeta')$ is the expected link probability between the source counterparts of two target individuals ζ and ζ'.

Note that the population version of the Gromov-Wasserstein projection $\tilde{Q}_g^{(w)}$ in Definition 1 is constructed through the true link probability, and hence, it carries most of the properties we assume for $P_g^{(w)}$ as well. That is, the function $\tilde{Q}_g^{(w)}$ is Borel, symmetric, and takes values in $[0,1]$, as $P_g^{(w)}$ does. Since the disintegration is unique only up to a μ_G-null set of ζ, different versions of $\left\{\tilde{\pi}_{\zeta}^{g,G}\right\}$ could potentially provide different functions. However, they differ only on a $\mu_G \otimes \mu_G$-null set, and hence, $\tilde{Q}_g^{(w)}$ is a well-defined element of $L^2(\mu_G \otimes \mu_G)$. Lastly, as $\tilde{Q}_g^{(w)}$ is a conditional expectation, transporting a graphon through a coupling never amplifies it, i.e., $\|\tilde{Q}_g^{(w)}\|_{L^2(\mu_G\otimes\mu_G)} \le \|P_g^{(w)}\|_{L^2(\mu_g\otimes\mu_g)}$. These properties of the population Gromov-Wasserstein projection are more thoroughly discussed in Lemma C.2 of Appendix C.

With such a definition, Algorithm 2 states the population form of the proposed method. It follows Algorithm 1 step by step: the population version of the Gromov-Wasserstein optimization problem produces the couplings of μ_g and μ_G, the internal distances are $L^2(\mu_g)$ distances between graphon slices, the projection is that of Definition 1, and the weights minimize an $L^2(\mu_G \otimes \mu_G)$ distance rather than the Frobenius norm.

With the population algorithm defined, we aim to study whether the imputation output $\tilde{P}_G^{(1)}$ of Algorithm 2 coincides with the estimand $P_G^{(1)}$ in $L_2(\mu_G \otimes \mu_G)$. However, two issues can prevent identification within the population problem. First, the Gromov-Wasserstein step may return a coupling that aligns target individuals to inappropriate source individuals, in which case the projections $\tilde{Q}_g^{(0)}$ misrepresent the source. Second, even with the right coupling, the target graphon may fail to lie in the convex hull of the projected sources, or may lie in it with weights that change between the two periods. The assumptions listed in the next subsection rule out all these potential issues.

3.3 Reference Representation and Assumptions

This subsection introduces the structure and assumptions under which identification can be achieved in the population problem. However, discussing these conditions directly is challenging because the latent characteristics distributions μ_g differ across networks. Hence, we start by introducing a common auxiliary space, the *reference space*, on which all population networks can be represented.

Definition 2 (Reference space and reference measure)**.**
A reference space is a pair $(\check{\Xi}, \check{\mu})$ consisting of an uncountable Polish space $\check{\Xi}$ and a non-atomic Borel probability measure $\check{\mu}$ on it. A reference space always exists, and for example, the pair $([0,1], \text{Uniform})$ is one such space.

To illustrate this idea, we consider the following simple toy example. Suppose that the only latent characteristic is age and suppose network g is skewed toward young individuals while network G is skewed toward old individuals. In this case, μ_g and μ_G therefore describe different age compositions. We can represent both distributions using the reference space $[0,1]$ equipped with the uniform distribution. A reference point $\check{\xi} \in (0,1)$ represents an age percentile, which each network maps to the corresponding age in its own distribution. For example, $\check{\xi} = 0.5$ corresponds to the median age in each network, which might be 25 in network g and 60 in network G. Thus, the same reference point can represent different ages while retaining the same percentile interpretation. Mapping a uniformly distributed reference point to the corresponding age quantile in each network reproduces that network's age distribution. This provides a common space

Algorithm 2 Synthetic Counterfactual Network (Population Form)

Inputs: $\left\{P_g^{(0)}\right\}_{g\in\{1,\dots,G-1,G\}}$; $\left\{P_g^{(1)}\right\}_{g\in\{1,\dots,G-1\}}$

Returns: $\tilde{P}_G^{(1)}$

1: **for all** $g \in \{1, ..., G-1\}$ **do**

2: Solve the Gromov-Wasserstein problem between the pre-treatment networks g and G and acquire the optimal coupling:

$$\tilde{\pi}^{g,G} \leftarrow \underset{\pi^{g,G}\in\Pi(\mu_g,\mu_G)}{\arg\inf} \iint_{\Xi\times\Xi}\iint_{\Xi\times\Xi} \left(d_g^2(\xi,\xi') - d_G^2(\zeta,\zeta')\right)^2 \mathrm{d}\pi^{g,G}(\xi,\zeta)\mathrm{d}\pi^{g,G}(\xi',\zeta')$$

where $\Pi(\mu_g, \mu_G)$ is the set of Borel probability measures on $\Xi \times \Xi$ with marginals μ_g and μ_G, and d_g^2 is the squared $L^2(\mu_g)$ distance between link-probability profiles,

$$d_g^2(\xi,\xi') = \left\|P_g^{(0)}(\xi,\cdot) - P_g^{(0)}(\xi',\cdot)\right\|_{L^2(\mu_g)}^2 = \int_\Xi \left[f^{(0)}(\xi,u,\gamma_g) - f^{(0)}(\xi',u,\gamma_g)\right]^2 \mathrm{d}\mu_g(u) \qquad \text{for } g \in \{1,...,G-1,G\}.$$

3: Gromov-Wasserstein projection: disintegrate $\tilde{\pi}^{g,G}$ with respect to its second marginal μ_G, $\mathrm{d}\tilde{\pi}^{g,G}(\xi,\zeta) = \mathrm{d}\tilde{\pi}_\zeta^{g,G}(\xi)\mathrm{d}\mu_G(\zeta)$, and compute the edge probability between $\zeta, \zeta' \in \mathrm{supp}(\mu_G)$ as if they lived in network g (instead of G):

$$\tilde{Q}_g^{(0)}(\zeta,\zeta') \leftarrow \int_\Xi\int_\Xi P_g^{(0)}(\xi,\xi')\mathrm{d}\tilde{\pi}_\zeta^{g,G}(\xi)\mathrm{d}\tilde{\pi}_{\zeta'}^{g,G}(\xi') \;\in\; L^2(\mu_G\otimes\mu_G),$$

which is well-defined by Lemma C.2.

4: **end for**

5: Solve the Gromov-Wasserstein barycenter analysis problem:

$$\left\{\tilde{\delta}_g\right\}_{g=1}^{G-1} \leftarrow \underset{\delta\in\Delta_{G-2}}{\arg\min} \left\| P_G^{(0)} - \sum_{g=1}^{G-1}\delta_g\tilde{Q}_g^{(0)}\right\|_{L^2(\mu_G\otimes\mu_G)}^2$$

6: **for all** $g \in \{1, \dots, G-1\}$ **do**

7: For the post-treatment period, use the same couplings, hence the same disintegrations $\{\tilde{\pi}_\zeta^{g,G}\}$:

$$\tilde{Q}_g^{(1)}(\zeta,\zeta') \leftarrow \int_\Xi\int_\Xi P_g^{(1)}(\xi,\xi')\mathrm{d}\tilde{\pi}_\zeta^{g,G}(\xi)\mathrm{d}\tilde{\pi}_{\zeta'}^{g,G}(\xi')$$

8: **end for**

9: Lastly, impute the target edge probability:

$$\tilde{P}_G^{(1)}(\zeta,\zeta') \leftarrow \sum_{g=1}^{G-1}\tilde{\delta}_g\tilde{Q}_g^{(1)}(\zeta,\zeta'), \qquad \zeta,\zeta' \in \mathrm{supp}(\mu_G).$$

on which to represent networks with different latent characteristics distributions.

Note that, at this stage, the latent characteristics distribution μ_g of each network may be discrete, continuous, or mixed, and need not have full support on Ξ. Only the reference measure $\check{\mu}$ is required to be non-atomic. This allows each μ_g to be represented by mapping a point drawn from $\check{\mu}$ to the network's latent characteristics space, as established in Lemma 1.

Lemma 1 (Alignment maps from the reference space)
Let $(\check{\Xi}, \check{\mu})$ be a reference space and let μ be any Borel probability measure on a Polish space Ξ. Then there exists a Borel measurable map $\rho : \check{\Xi} \to \Xi$ such that:

(i) $\rho_{\#}\check{\mu} = \mu$, *that is,* $\check{\mu}\left(\rho^{-1}(A)\right) = \mu(A)$ *for every Borel* $A \subseteq \Xi$;

(ii) $\rho(\check{\Xi}) \subseteq \text{supp}(\mu)$.

In particular, applying the construction to each μ_g, $g \in \{1, ..., G-1, G\}$, yields a family of alignment maps $\rho_g : \check{\Xi} \to \Xi$ with $\left(\rho_g\right)_{\#}\check{\mu} = \mu_g$ and $\rho_g(\check{\Xi}) \subseteq \text{supp}(\mu_g)$ for every $g \in \{1, ..., G-1, G\}$.

Lemma 1 suggests that there exists an alignment map connecting a common auxiliary reference space to each network's individual latent characteristics distribution. Continuing with the previous toy example, we can think of a point in the reference space representing a *role* on the young-to-old spectrum. The alignment map serves as a connection between these roles/reference coordinates and corresponding age numbers. And with different alignment maps, the same role/reference coordinate corresponds to different age values in different networks.

Lemma 1 requires no additional assumption other than the maintained model framework. The reference space is an abstract auxiliary tool that allow us to analyze the cross network alignment and network formation on a common space. It is worth noting that the reference space itself is not a modeling restriction, and the family of alignment maps $\{\rho_g\}$ need not be unique. The results below do not require the alignment maps to be bijective. Thus, each network's latent space may remain unrestricted. If we are willing to assume the bijectivity, which requires the non-atomicity of μ_g, then ρ_g^{-1} can be read as a role/reference-extraction operator.

We further extend the aforementioned toy example to illustrates what the reference space and the alignment maps can represent.

Example 1 (Toy empirical example of reference space and alignment maps)**.**
Take the reference space $(\check{\Xi}, \check{\mu}) = ([0, 1], \text{Uniform})$[10] and two cities with latent characteristics $\xi = (\text{age}, \text{income}) \in \Xi \subseteq \mathbb{R}^2$. City g is young and poor, city G is old and rich. Suppose that, within

[10]The reference space can be multi-dimensional; nothing in Definition 2 restricts its dimension.

each city, both characteristics are driven by a single socioeconomic index $\check{\xi}$, and for concreteness let the alignment maps take the explicit functional forms

$$\rho_g(\check{\xi}) = \left(20 + 60\,\check{\xi}^2,\ 20 + 100\,\check{\xi}^2\right) \qquad \text{and} \qquad \rho_G(\check{\xi}) = \left(20 + 60\sqrt{\check{\xi}},\ 20 + 100\sqrt{\check{\xi}}\right)$$

The following observations describe how the reference structure interpret this scenario:

(i) **The pushforward maps produce two skewed populations**
With the uniform draw $\check{\xi} \sim$ Uniform, under ρ_g age piles mass around 20, and the median resident of city g is 35 years old earning 45. Under ρ_G, age piles mass around 80, and the median resident of city G is 62 years old earning 91.

(ii) **A common reference coordinate represents different individuals across cities**
For example, the role/reference coordinate $\check{\xi} = 0.1$ is occupied by a 21-year-old earning 21 in city g and by a 39-year-old earning 52 in city G. The role/reference coordinate answers the question of where an individual stands within their own network.

(iii) **There is no restriction on the dimension of the reference space**
In this toy example, the common ranking of age and income is a simplifying choice for illustration. More generally, the reference space may have any positive integer dimension. For instance, maps $\rho_g : [0,1]^4 \to \Xi \subseteq \mathbb{R}^2$ can generate age and income from a uniformly distributed reference point without requiring their rankings to coincide. Lemma 1 establishes the existence of such representations.

With a reference space and alignment maps defined, the graphon from each network can be represented on the common reference space domain $\check{\Xi}$ along with its alignment map.

Definition 3 (Graphon and internal distance in reference coordinates)**.**
For $g \in \{1, ..., G-1, G\}$ and $(\check{\xi}, \check{\xi}') \in \check{\Xi} \times \check{\Xi}$, define

$$\check{f}_g^{(w)}(\check{\xi}, \check{\xi}') := f^{(w)}\left(\rho_g(\check{\xi}), \rho_g(\check{\xi}'), \gamma_g\right) = P_g^{(w)}\left(\rho_g(\check{\xi}), \rho_g(\check{\xi}')\right)$$

and

$$\check{d}_g^2(\check{\xi}, \check{\xi}') := \int_{\check{\Xi}} \left[\check{f}_g^{(0)}(\check{\xi}, \check{u}) - \check{f}_g^{(0)}(\check{\xi}', \check{u})\right]^2 \mathrm{d}\check{\mu}(\check{u}). \tag{3}$$

Under this representation, $\check{f}_g^{(w)}(\check{\xi}, \check{\xi}')$ is the probability that the individuals occupying reference coordinates $\check{\xi}$ and $\check{\xi}'$ in network g are linked, and $\check{d}_g^2$ is the within network internal distance of Algorithm 2 pulled back to reference coordinates.

After defining the reference representation, we now ready to lay down the assumptions for the identification in the population problem. The first is the convex hull restriction resemble the traditional synthetic control method, the second restricts how the internal geometries of

different networks may differ, and the third ensures that the alignment step leaves no ambiguity that the imputation step could see.

Assumption 2 (Convex hull in reference measure)**.**
There exist a reference space $(\check{\Xi}, \check{\mu})$, alignment maps $\{\rho_g\}_{g=1}^{G}$ satisfying $(\rho_g)_{\#}\check{\mu} = \mu_g$, and weights $\delta^* \in \Delta_{G-2}$ such that, for both $w \in \{0, 1\}$,

$$f_G^{\check{(w)}} = \sum_{g=1}^{G-1} \delta_g^* f_g^{\check{(w)}} \quad \text{in} \quad L^2(\check{\mu} \otimes \check{\mu}),$$

with the same weights δ^* for both treatment statuses.[11]

Assumption 2 is the synthetic control type identification restriction of the proposed method, expressed in reference space representation. Reading on the common system of reference coordinates $\check{\xi}$, the target network's link probability is a convex combination of the source networks' link probabilities, with weights that do not depend on the treatment status. With the number of networks fixed, this assumption imposes an exact convex hull relation between the target and source graphons in the common reference representation. Additionally, the time invariance of the weights is what allows them to be learned from the pre-treatment time period and carried over to the post-treatment time period, exactly as in the traditional synthetic control method. Assumptions 3 and 4 are imposed on the same reference space and alignment maps.

Assumption 3 (Alignment of comonotone geometries across different networks)**.**
For every $g \in \{1, ..., G-1\}$,

$$\left[\check{d}_g^2(\check{\xi}, \check{\xi}') - \check{d}_g^2(\check{\zeta}, \check{\zeta}')\right]\left[\check{d}_G^2(\check{\xi}, \check{\xi}') - \check{d}_G^2(\check{\zeta}, \check{\zeta}')\right] \geq 0$$

for $(\check{\mu} \otimes \check{\mu}) \otimes (\check{\mu} \otimes \check{\mu})$-a.e. pair of pairs of reference coordinates $\left((\check{\xi}, \check{\xi}'), (\check{\zeta}, \check{\zeta}')\right)$.

Assumption 3 is the key identifying restriction. Recall that $\check{d}_g^2(\check{\xi}, \check{\xi}')$ measures how differently two reference coordinates $\check{\xi}$ and $\check{\xi}'$ connect to the rest of network g. This assumption states nothing about the level of these dissimilarities, but rather their ordering. *Whenever network g regards one pair of positions as at least as dissimilar as another pair, the target network G does too.* Equivalently, the two random variables $\check{d}_g^2$ and $\check{d}_G^2$, each defined on the pair space $(\check{\Xi} \times \check{\Xi}, \check{\mu} \otimes \check{\mu})$, are comonotone. In the scalar case with $(\check{\Xi}, \check{\mu}) = ([0, 1], \text{Uniform})$ and ρ_g the quantile function of μ_g, a reference coordinate $\check{\xi}$ is a within-network rank. Under this simplification, Assumption 3 is

[11] The alignment maps induce the joint law $\mathcal{P} := (\rho_1, \ldots, \rho_G)_{\#}\check{\mu}$ on Ξ^G, with marginals $\mu_1, \ldots, \mu_G$. Under this law, the coordinates $\xi_g = \rho_g(\check{\xi})$ correspond to the same reference point $\check{\xi} \sim \check{\mu}$. The convex hull restriction can therefore be written as $f^{(w)}(\xi_G, \xi_G', \gamma_G) = \sum_{g=1}^{G-1} \delta_g^* f^{(w)}(\xi_g, \xi_g', \gamma_g)$ for $\mathcal{P} \otimes \mathcal{P}$-almost every pair of vectors $(\xi_1, \ldots, \xi_G)$ and $(\xi_1', \ldots, \xi_G')$, with the same weights for both $w \in \{0, 1\}$.

equivalent to all networks agreeing on which pairs of ranks are more alike and which are less alike. Passing from the target to a source may stretch or compress the scale on which dissimilarity is measured, and may do so nonlinearly, but it may not reorder the pairs. Note that the scalar case with a quantile interpretation is just one of the special cases. Assumption 3 states a much more general cross-network relation and allows potentially higher-dimensional reference representations. Remark 1 illustrates the condition with one case that satisfies it and one that does not.

Remark 1 (What Assumption 3 allows and what it rules out). Suppose the only latent characteristic is age. For simplicity let all networks share the same latent distribution μ. In this case, reference coordinates can be read directly as ages and $\check{d}_g = d_g$. Consider three networks that form links by three different rules:

- network g sorts by *age group*: $P_g(\xi, \xi') = 0.8$ if both are young (≤ 30) or both are old (≥ 60), and 0.2 otherwise;
- network G sorts by *age gap*: $P_G(\xi, \xi') = 0.8 - |\xi - \xi'|/100$;
- network X sorts by *leisure time* (*distance to middle age*): $P_X(\xi, \xi') = 0.8 - \big||\xi - 50| - |\xi' - 50|\big|/50$.

Networks g and G satisfy Assumption 3. Network g sees only two levels of dissimilarity: d_g vanishes within an age group, where connection profiles are identical, and takes one positive value across groups. Network G grades pairs continuously by the age gap. Provided the two age groups are separated, every across-group pair has a larger gap than every within-group pair, and hence, G ranks it as more dissimilar too. With this, the ordering is preserved across g and G.

Networks g and X violate the assumption. Consider the example of having a student and a retiree who are equally far from middle age, one below it and one above. In network X, they have the same amount of free leisure time, and hence they connect to everyone else in the same way and are indistinguishable. Network g, by contrast, places them in different age groups and regards them as maximally dissimilar. The reverse happens for two young individuals at different distances from middle age: identical in g, distinct in X. The two networks disagree about which individuals are alike, so no monotone rescaling of d_g can recover d_X, and the assumption fails.

The example shows what the assumption does and does not restrict. It allows the latent distributions μ_g to differ, since the alignment maps absorb them. It allows the same sorting rule to operate with different intensities. It also allows one network to see blocks where another sees a continuum, as g and G do. What it rules out is disagreement about the ordering itself. Hence, the assumption does not require the latent characteristics to be scalar or even to be the same

across different networks, but rather requires the networks to follow similar rules expressed through the reference representation. Corollary 1 gives sufficient conditions on the network formation model under which this requirement holds.

Assumption 3 is the natural condition for the proposed method as the Gromov-Wasserstein optimal transport only sees each network only through its internal distances. If two networks disagree on which pairs are alike, no coupling can respect both geometries. In this case, the Gromov-Wasserstein optimum is a compromise that mixes positions across the disagreement, and the projection of Definition 1 then misrepresents the source.

As will be discussed in more detail in the later Section 4, under the assumption, the reference coordinates sort both geometries simultaneously, and Theorem 1 shows that there is a specific alignment coupling $\pi_\rho^{g,G}$ that is optimal, with an optimal value equal to the $L^2(\check{\mu} \otimes \check{\mu})$ distance between $\check{d}_g^2$ and $\check{d}_G^2$. This value is the second lower bound of Mémoli (2011), namely, the squared Wasserstein distance between the two one-dimensional laws of dissimilarity. Proposition B.1 shows that the assumption is exactly the condition under which the bound is attained: if the Gromov-Wasserstein optimum equals this bound for every source, a reference representation satisfying Assumption 3 exists.

What Assumption 3 leaves open is how we handle ties. In the toy example of Remark 1, network g cannot distinguish one across-group pair from another, leaving potential ambiguity in the optimal alignment. Assumption 4 addresses exactly this remaining gap and ensures, together with the other maintained conditions, that the individual-level link probabilities can be recovered.

Assumption 4 (Rigidity of the target network structure)**.**
The following two conditions hold:

(i) For every $g \in \{1, ..., G-1\}$, at least one of the following holds:

  (a) The law of $\check{d}_g^2$ under $\check{\mu} \otimes \check{\mu}$ is non-atomic.

  (b) There exists a Borel function $b_g : [0, \infty) \to [0, 1]$ with $\check{d}_G^2 = b_g \circ \check{d}_g^2$ $\check{\mu} \otimes \check{\mu}$-a.e.

(ii) Every $\check{\pi} \in \Pi(\check{\mu}, \check{\mu})$ with $\check{d}_G^2(\check{\zeta}, \check{\zeta}') = \check{d}_G^2(\check{\xi}, \check{\xi}')$ for $\check{\pi} \otimes \check{\pi}$-a.e. pairs satisfies $\check{f}_g^{(w)}(\check{\zeta}, \check{\zeta}') = \check{f}_g^{(w)}(\check{\xi}, \check{\xi}')$ for $\check{\pi} \otimes \check{\pi}$-a.e. pairs, for every $g \in \{1, ..., G-1, G\}$ and both $w \in \{0, 1\}$.

Assumption 4 closes the gap between what the alignment step optimizes and what the imputation step consumes. The Gromov-Wasserstein optimal transport problem only sees the internal distances. Hence, its optimum is pinned down only up to whatever those internal distances cannot distinguish.

Assumption 4(i) strengthens the comonotone relation of Assumption 3 into a deterministic link, and either of its two conditions suffices:

(a) the first condition states that no single dissimilarity level is shared by a positive fraction of reference coordinate pairs, so that the source network's internal geometry contains no ties, and

(b) the second condition states that the target's internal dissimilarity is a deterministic function of the source's.

Assumption 4(ii) requires that whatever the target's internal distance cannot distinguish, the link probabilities do not distinguish either. With this, every optimal coupling, which is not unique, from the Gromov-Wasserstein optimization produces the same projection. Note that Assumption 4 does not force the optimal coupling to be unique. It only forces all optimal couplings to be equivalent in the sense of imputation purpose.

To further concretize the need for Assumption 4, Remark 2 discusses why only Assumption 3 is not enough and how Assumption 4 is required to resolve the such problem, with a two-community toy example in which the assumption fails.

Remark 2 (Why Assumption 4 is needed). The projection uses whichever optimal coupling that Algorithm 2 Step 2 returns, and the Gromov-Wasserstein problem in general has several optimal couplings. The reason is that the objective of Step 2 sees the target only through its internal distance. If a measure-preserving map $T : \check{\Xi} \to \check{\Xi}$ satisfies $\check{d}_G^2(T\check{\zeta}, T\check{\zeta}') = \check{d}_G^2(\check{\zeta}, \check{\zeta}')$, then composing the target side of any coupling with T leaves the optimization objective unchanged. Hence, an optimal coupling is optimal after such a relabeling of target reference coordinates, and identification must be robust to which of these couplings the optimizer selects.

Here we present a simple example to demonstrate this concern. Consider a target network with two equally sized communities: with within-community link probabilities p and q, and cross-community link probability r. The dissimilarity $\check{d}_G^2$ between two reference coordinates in the same community is zero, and between two reference coordinates in different communities it is the constant $\frac{1}{2}(p-r)^2 + \frac{1}{2}(r-q)^2$, the same on both sides. The map that swaps the two communities thus preserves both $\check{\mu}$ and $\check{d}_G^2$, so any optimal coupling remains optimal after the target communities are interchanged. When $p \neq q$, the interchanged coupling projects the source graphons onto the target with its communities swapped, and the imputation fails even though the coupling is optimal.

Assumption 4 rules out this scenario while still allowing several optimal Gromov-Wasserstein couplings. Assumption 4(i) requires the target dissimilarity to be a function of the source dissimilarity. When the source dissimilarity distribution is non-atomic, each source dissimilarity

value determines a percentile in that distribution, and comonotonicity assigns the corresponding quantile of the target dissimilarity distribution to that percentile. Hence, the deterministic relationship follows from Assumption 3. When the source distribution has atoms, as in block models, tied source dissimilarities may correspond to different target dissimilarities. The second branch of condition (i) imposes this relationship explicitly.

Assumption 4(ii) then requires that any coupling of the reference with itself preserving the target dissimilarity also preserves every network's link probability. With this, then any relabeling where the optimal couplings can still perform is invisible to the projection. In the example, statement (ii) holds when $p = q$: the swap then preserves the link probabilities, and every optimal coupling delivers the same projection. It fails when $p \neq q$, exactly the case in which the imputation can fail.

To summarize, Assumption 4 does not force the optimal coupling to be unique. It forces every optimal coupling to produce the same projection.

3.4 Discussion on the Assumptions

3.4.1 Relation to the Network Formation Literature

As discussed, Assumptions 3 and 4 are high-level conditions on the internal geometries across different networks. Assumption 3 requires that each source and the target do not reverse the ordering of pairs by their within-network dissimilarities. Assumption 4 ensures that, although the optimal coupling may not be unique, different optimal couplings produce the same projected link probabilities. These are not restrictions on which network formation model is allowed. In fact, most of the common network formation models without strategic interaction are included in our framework. To make this precise, we give three classes of modeling assumptions, each of which specifies a network formation model on the reference representation $\check{f}_g^{(w)}$ and each of which implies Assumptions 3 and 4.

Corollary 1 (Modeling assumptions examples)
Each of the following modeling assumptions implies Assumptions 3 and 4.
(*i*) *Basis functions with co-monotonicity* *For $K \geq 1$, network-specific loadings $\psi_g^{(w)} \in \mathbb{R}^K$ and intercepts $\alpha_g^{(w)} \in \mathbb{R}$, and a common vector $\phi := (\phi_1, ..., \phi_K)^\intercal$ of symmetric bounded Borel functions on $\check{\Xi} \times \check{\Xi}$,*

$$\check{f}_g^{(w)}(\check{\xi}, \check{\xi}') = \psi_g^{(w)\intercal} \phi(\check{\xi}, \check{\xi}') + \alpha_g^{(w)} \qquad \check{\mu} \otimes \check{\mu}\text{-a.e.},$$

such that the following two conditions hold:

(*a*) ***All K dimensions agree with one another.*** *That is, one common notion of how far apart two reference coordinates sit drives every dimension, and no dimension orders pairs of reference coordinates against the others.*

Let

$$M_{k\ell}(\check{\xi},\check{\xi}') := \int_{\check{\Xi}} \Big[\phi_k(\check{\xi},u) - \phi_k(\check{\xi}',u)\Big]\Big[\phi_\ell(\check{\xi},u) - \phi_\ell(\check{\xi}',u)\Big]\,\mathrm{d}\check{\mu}(u)$$

measure how differently $\check{\xi}$ and $\check{\xi}'$ connect to others through dimensions k and ℓ, so that $\sum_{k=1}^{K} M_{kk}$ is the total discrepancy across all dimensions. Every ϕ_k and every $M_{k\ell}$ is, up to $\check{\mu}\otimes\check{\mu}$-null sets, a Borel function of the total discrepancy $\sum_{k=1}^{K} M_{kk}$, non-decreasing in the case of the $M_{k\ell}$ and strictly increasing in the case of the diagonal M_{kk}.

(*b*) ***Difference in network-specific effects.*** *$\psi_g^{(0)}$ is componentwise nonnegative with at least one strictly positive component for every $g \in \{1,...,G-1,G\}$.*

(*ii*) *Links driven by similarity alone* *For network-specific monotone functions $\psi_g^{(w)} : [0,\infty) \to [0,1]$ and a common symmetric Borel function $\phi : \check{\Xi}\times\check{\Xi} \to [0,\infty)$,*

$$f_g^{\check{}(w)}(\check{\xi},\check{\xi}') = \psi_g^{(w)}\Big(\phi(\check{\xi},\check{\xi}')\Big) \qquad \check{\mu}\otimes\check{\mu}\text{-a.e.},$$

such that the following two conditions hold:

(*a*) ***Valid pre-treatment.*** *That is, in $w = 0$, different similarity levels always produce different edge probabilities.*

Each $\psi_g^{(0)}$, $g \in \{1,...,G-1,G\}$, is strictly increasing or strictly decreasing; the direction may differ across networks.

(*b*) ***From every individual, the rest of the network looks the same.*** *That is, the environment looks the same from every $\check{\xi}$: no $\check{\xi}$ sits at an edge, in a cluster, or in a thin region. How much the neighborhoods of $\check{\xi}$ and $\check{\xi}'$ overlap is therefore determined by their similarity $\phi(\check{\xi},\check{\xi}')$ alone, and the overlap strictly shrinks as $\phi(\check{\xi},\check{\xi}')$ grows.*

For $\check{\mu}\otimes\check{\mu}$-a.e. $(\check{\xi},\check{\xi}')$ and all $r, r' > 0$, the overlap

$$\check{\mu}\Big(\big\{u : \phi(\check{\xi},u) < r\big\} \cap \big\{u : \phi(\check{\xi}',u) < r'\big\}\Big)$$

depends on $(\check{\xi},\check{\xi}')$ only through $\phi(\check{\xi},\check{\xi}')$ and is strictly decreasing in $\phi(\check{\xi},\check{\xi}')$.

(*iii*) *Block structure* *For a common Borel block assignment $\phi : \check{\Xi} \to \{1,...,B\}$ with block masses $\beta_b := \check{\mu}\left(\phi^{-1}(b)\right) > 0$ for every b, and network-specific symmetric matrices $B_g^{(w)} \in \mathbb{R}^{B\times B}$,*

$$f_g^{\check{}(w)}(\check{\xi},\check{\xi}') = B_g^{(w)}\Big(\phi(\check{\xi}),\phi(\check{\xi}')\Big) \qquad \check{\mu}\otimes\check{\mu}\text{-a.e.}$$

Denote

$$C_g(b,b') := \sum_{\ell=1}^{B} \beta_\ell \Big[B_g^{(0)}(b,\ell) - B_g^{(0)}(b',\ell)\Big]^2,$$

which measures how differently blocks b and b' connect to the rest of the blocks, and require that the following two conditions hold:

(*a*) ***All networks rank pairs of blocks in the same way.*** *That is, if a source regards one pair of blocks as no more different than another pair, so does the target, ties included.*

For every $g \in \{1, ..., G-1\}$ and all $b, b', \ell, \ell' \in \{1, ..., B\}$, $C_g(b, b') \leq C_g(\ell, \ell')$ implies $C_G(b, b') \leq C_G(\ell, \ell')$.

(*b*) ***Undetectable shuffles change nothing.*** *That is, any way of shuffling blocks that respects the block masses and preserves the target geometry also preserves every network's block matrix, before and after the treatment.*

Every shuffling matrix $S \in [0,1]^{B \times B}$ with both marginals equal to $(\beta_1, ..., \beta_B)$ that preserves the target block geometry, in the sense that $C_G(b, b') = C_G(\ell, \ell')$ whenever $S(b, \ell) S(b', \ell') > 0$, also satisfies $B_g^{(w)}(b, b') = B_g^{(w)}(\ell, \ell')$ whenever $S(b, \ell) S(b', \ell') > 0$, for every $g \in \{1, ..., G-1, G\}$ and both $w \in \{0, 1\}$.

The three modeling assumptions in Corollary 1 cover the majority of network formation models in the literature that do not involve strategic interaction. A more detailed discussion of those specific network formation models and the interpretation of Assumptions 3 and 4 within those models is provided in Appendix B.

3.4.2 Relation to the Optimal Transport Literature

On the other hand, Assumptions 3 and 4 also relate to the theory of Gromov-Wasserstein optimal transport. Mémoli (2011) organizes computable lower bounds for the Gromov-Wasserstein distance between two metric measure spaces according to how much of the internal structure they retain:

(i) the first lower bound compares eccentricities,

(ii) the second compares the distributions of pairwise distances, and

(iii) the third compares the local distributions of distances seen from each point.

The second lower bound is the one that is relevant to Assumption 3. For each network, let $\mathcal{L}_g$ be the law of $\check{d}_g^2$ under two independent draws from $\check{\mu}$. That is, the distribution of dissimilarities between a random pair of reference coordinates, which is a probability measure on $[0, 1]$. Then, for every coupling of μ_g and μ_G, the Gromov-Wasserstein objective is at least the squared Wasserstein distance $W_2^2(\mathcal{L}_g, \mathcal{L}_G)$, and Assumption 3 is exactly the condition under which this bound is attained. The reference space coordinates sort the two internal geometries simultaneously, so that a pair of reference coordinates that is far apart in a source is also far apart in the target. On top of that, the matching problem collapses onto the comparison of two distributions on the line

of dissimilarity. Conversely, if the bound is attained for every source, a reference representation satisfying Assumption 3 exists.

Assumption 4 addresses what the second lower bound cannot see. Its statement (i) asks the optimal coupling of the two one-dimensional dissimilarity laws to be a map rather than a genuine coupling, so that the target dissimilarity is a function of the source dissimilarity. Its statement (ii) responds to the fact that the distribution of distances is not a complete invariant of a metric measure space, as two genuinely different networks can share it (cf. Mémoli, 2011). As discussed in Chowdhury and Mémoli (2019), a coupling of the reference with itself that preserves $\check{d}_G^2$ is a weak isomorphism of the target measure network, and statement (ii) requires that every such isomorphism also preserve every network's link probabilities. Thus, the final imputation is never affected.

Formal statements of the above discussion about the optimal transport literature, together with their proofs, are given as Propositions B.1 and B.2 in Appendix B.

4 Identification in the Population Problem

With the population model and problem defined, in this section we study identification in the population problem. That is, as described in the previous section, we assume we have access to the laws of link probabilities $\left\{P_g^{(0)}\right\}_{g\in[G-1]}$ and $\left\{P_g^{(1)}\right\}_{g\in[G]}$, where $P_g^{(w)}$ are in $L_2(\mu_g \otimes \mu_g)$. We show that the population version of the proposed method, Algorithm 2, recovers the post-treatment link probabilities $P_G^{(1)}$ of the target network exactly under the assumptions of Section 3.

The difficulty in analyzing the population problem lies in the Gromov-Wasserstein optimal transport step of Algorithm 2. The Gromov-Wasserstein optimization returns an optimal coupling, but such an optimal coupling is not unique. Furthermore, no analytical form for it is available. Without further structure, one therefore does not know which optimal coupling the algorithm will return, and one cannot analyze the projection and the weights that are built from it as there is no analytical form.

To resolve this issue, we proceed in the following steps. First, we construct an explicit coupling, which we call the *alignment coupling*, from the reference space representation of Section 3.3, and its explicit formulation enables us to proceed. Second, we show that, under Assumption 3, this alignment coupling belongs to the set of Gromov-Wasserstein optimal couplings. Third, assuming that the Gromov-Wasserstein optimization actually returns this alignment coupling, we compute the resulting Gromov-Wasserstein projection and show that, under Assumption 2, the subsequent weight and imputation steps recover $P_G^{(1)}$ exactly. Lastly, we remove the supposition: under Assumptions 2-4, all the optimal couplings that the Gromov-Wasserstein optimiza-

tion produces work the same as the alignment coupling, and hence, the population problem is solved exactly.

Throughout this section, for notation simplicity, denote

$$F_{C,D}(\pi) := \iint \iint \left[C(\xi,\xi') - D(\zeta,\zeta')\right]^2 \mathrm{d}\pi(\xi,\zeta)\mathrm{d}\pi(\xi',\zeta')$$

4.1 Alignment Coupling and its Optimality

As mentioned earlier, the Gromov-Wasserstein optimal transport problem in the Step 2 of Algorithm 2 returns an optimal coupling, but such optimal coupling is not directly workable. Not only because the optimal coupling is in general not unique, but also the analytic form for it is not available, which is the problem stems from non-convex optimization over $\Pi(\mu_g, \mu_G)$. Hence, we proceed the analysis indirectly by first constructing an explicit coupling, named *alignment coupling* via the reference space representation.

The reference space of Definition 2 and the alignment maps of Lemma 1 supply this natural candidate for the coupling that the Gromov-Wasserstein optimal transport problem may return.

Definition 4 (Alignment coupling)**.**
Let $(\check{\Xi}, \check{\mu})$ be a reference space and $\left\{\rho_g\right\}_{g=1}^{G}$ be the family of alignment maps as in Lemma 1. For $g \in \{1, ..., G-1\}$, the alignment coupling is defined as

$$\pi_\rho^{g,G} := \left(\rho_g, \rho_G\right)_{\#} \check{\mu},$$

that is, the Borel probability measure on $\Xi \times \Xi$ with $\pi_\rho^{g,G}(A \times B) = \check{\mu}\left(\rho_g^{-1}(A) \cap \rho_G^{-1}(B)\right)$ for all Borel $A, B \subseteq \Xi$.

The alignment coupling $\pi_\rho^{g,G}$ draws a *reference coordinate* $\check{\xi}$ from the reference measure and pairs the source individual $\rho_g(\check{\xi})$ with the target individual $\rho_G(\check{\xi})$ who occupies the same reference coordinate. By Lemma 1 (i), it has the marginal distributions of $\left(\rho_g\right)_{\#}\check{\mu} = \mu_g$ and $(\rho_G)_{\#}\check{\mu} = \mu_G$. Thus, $\pi_\rho^{g,G} \in \Pi(\mu_g, \mu_G)$ is indeed a feasible candidate for the Gromov-Wasserstein optimization problem.

Since the alignment coupling is built from the reference space, the natural way to study it is to rewrite the whole Gromov-Wasserstein optimal transport problem to the reference space as well. In the reference space representation, the internal distances pull back to their reference coordinates $\check{\xi}$ through the alignment maps ρ_g. That is, for every network g and all $\check{\xi}, \check{\xi}' \in \check{\Xi}$, it follows that

$$d_g^2\left(\rho_g(\check{\xi}), \rho_g(\check{\xi}')\right) = \check{d}_g^2(\check{\xi}, \check{\xi}').$$

This is the identity behind Definition 3. Since $\left(\rho_g\right)_{\#}\check{\mu} = \mu_g$, integrating over $\xi \sim \mu_g$ in the definition of d_g^2 is the same as integrating over the reference coordinate $\check{\xi} \sim \check{\mu}$ in the definition of $\check{d}_g^2$. In other words, the dissimilarity measurement between two source individuals equals the dissimilarity between the reference coordinates they occupy. Same thing can be applied to the target network G.

With this, the alignment coupling $\check{\pi} \in \Pi(\check{\mu}, \check{\mu})$ of the reference measure with itself can be pushed forward through $\rho_g \times \rho_G$ to a coupling of μ_g and μ_G, and the Gromov-Wasserstein objective is unchanged. That is, $F_{d_g^2, d_G^2}\left(\left(\rho_g \times \rho_G\right)_{\#}\check{\pi}\right) = F_{\check{d}_g^2, \check{d}_G^2}(\check{\pi})$. Conversely, every $\pi \in \Pi(\mu_g, \mu_G)$ originates this way from some $\check{\pi}$, obtained by drawing $(\xi, \zeta) \sim \pi$ and then drawing the reference coordinates behind ξ and behind ζ from the conditional laws of $\check{\mu}$ given $\rho_g = \xi$ and given $\rho_G = \zeta$. Hence, the original Gromov-Wasserstein optimization problem is therefore equivalent to the problem represents on the reference space

$$\inf_{\check{\pi} \in \Pi(\check{\mu}, \check{\mu})} F_{\check{d}_g^2, \check{d}_G^2}(\check{\pi}), \tag{4}$$

in the sense that the two infima coincide and optimal couplings of one problem map to optimal couplings of the other. The reference problem asks that what is the best way to couple the reference coordinates of network g with the reference coordinates of network G.

With such relation between the original Gromov-Wasserstein problem and its reference problem, the alignment coupling has a natural counterpart on the reference space. For the reference problem, among all couplings of $\check{\mu}$ with itself, the simplest is the identity coupling $\check{\pi}^{\text{diag}} := (\text{id}, \text{id})_{\#}\check{\mu}$, which pairs every reference coordinate $\check{\xi}$ with itself and hence is supported on the diagonal of $\check{\Xi} \times \check{\Xi}$. Since pushing $\check{\pi}^{\text{diag}}$ forward through $\rho_g \times \rho_G$ sends the pair $(\check{\xi}, \check{\xi})$ to the pair $\left(\rho_g(\check{\xi}), \rho_G(\check{\xi})\right)$, which is exactly how the alignment coupling pairs individuals, we have $\pi_\rho^{g,G} = \left(\rho_g \times \rho_G\right)_{\#}\check{\pi}^{\text{diag}}$. That is, the alignment coupling is the identity coupling read through the alignment maps.

Consequently, $\pi_\rho^{g,G}$ is optimal if and only if $\check{\pi}^{\text{diag}}$ is optimal for the reference problem (4). The question of whether the alignment coupling is optimal then reduces to the question of whether matching each reference coordinate to itself is the best way to align the two reference geometries. This is also the motivation of why the assumptions of Section 3.3 were stated in reference coordinates: the matching step sees the two networks only through $\check{d}_g^2$ and $\check{d}_G^2$. The formal statements are collected in Lemma A.1 in the Appendix.

With all of these in mind, the next theorem states that the alignment coupling $\pi_\rho^{g,G}$ is indeed one of the Gromov-Wasserstein optimal coupling.

Theorem 1 (Optimality of the alignment coupling)

Let Assumption 3 hold and fix $g \in \{1, ..., G-1\}$. The following statements hold:

(*i*) *The alignment coupling* $\pi_\rho^{g,G}$ *is optimal for the Gromov-Wasserstein optimization over* $\Pi(\mu_g, \mu_G)$. *That is, the infimum is attained at* $\pi_\rho^{g,G}$, *and its value is*

$$\inf_{\pi \in \Pi(\mu_g, \mu_G)} F_{d_g^2, d_G^2}(\pi) = \left\| \check{d}_g^2 - \check{d}_G^2 \right\|_{L^2(\check{\mu} \otimes \check{\mu})}^2 .$$

(*ii*) *Every optimal* $\pi \in \Pi(\mu_g, \mu_G)$ *can be written as* $\pi = \left(\rho_g \times \rho_G\right)_\# \check{\pi}$, *where* $\check{\pi} \in \Pi(\check{\mu}, \check{\mu})$ *is an optimal coupling of the reference problem* (4). *Moreover, for every optimal* $\check{\pi}$ *of the reference problem and two independent draws* $(\check{\xi}, \check{\zeta})$ *and* $(\check{\xi}', \check{\zeta}')$ *from* $\check{\pi}$, *the pair* $\left(\check{d}_g^2(\check{\xi}, \check{\xi}'),\ \check{d}_G^2(\check{\zeta}, \check{\zeta}')\right)$ *is comonotone.*

Theorem 1 (i) establishes that the alignment coupling $\pi_\rho^{g,G}$ is optimal for Gromov-Wasserstein optimization. It does not state that the optimal coupling is unique, as in general it is no. Hence, $\pi_\rho^{g,G}$ is one optimal coupling among possibly many.

Theorem 1 (ii) constrains the alternatives. Every optimal coupling is a role-to-role matching read through the reference coordinates, and every one of them sorts the two internal structures across different networks in the same way. Hence, two optimal couplings can only differ in how they shuffle positions that the within network distances cannot distinguish. Whether such shuffles are harmless for the imputation is a separate question, which is why Assumption 4 is imposed and what Proposition 3 below is for. Note that no regularity condition on any μ_g enters Theorem 1. The reference space representation provides the coupling unconditionally, and the only condition used is the ordering restriction of Assumption 3, which is exactly what the matching step sees.

Remark 2 (The optimal Gromov-Wasserstein objective value is the discrepancy between aligned internal structures). As shown in Theorem 1 (i), the objective achieved at the alignment coupling is the L^2 distance between the internal structures of source and target, both read on the common reference space. The value vanishes if and only if $\check{d}_g^2 = \check{d}_G^2$ $\check{\mu} \otimes \check{\mu}$-a.e. That is, if and only if the two networks place every pair of reference coordinates at identical dissimilarity. The objective value represent the structural discrepancy across networks that survives the alignment, and the optimality of the alignment coupling is unaffected by the value being positive. Such discrepancy cause by network specific effects γ is then tackle by the later synthetic weights construction under the convex hull condition in Assumption 2.

4.2 The Gromov-Wasserstein Projection through the Alignment Coupling

In this subsection, we study the Gromov-Wasserstein projection under the supposition that the coupling returned by the Gromov-Wasserstein optimization in Step 2 of Algorithm 2 is the alignment coupling. That is, we suppose that $\tilde{\pi}^{g,G} = \pi_\rho^{g,G} = \left(\rho_g, \rho_G\right)_\# \check{\mu}$, and analyze the projection that follows.

We make this supposition for the following reason. The projection $\tilde{Q}_g^{(w)}$ of Definition 1 is built from the conditional laws of whichever optimal coupling $\tilde{\pi}^{g,G}$ the optimization returns. Theorem 1 establishes that, under Assumption 3, $\pi_\rho^{g,G}$ belongs to the optimal set, but it does not guarantee that the optimization returns this particular coupling. Without an explicit form for $\tilde{\pi}^{g,G}$, we do not have a direct expression for its conditional laws or the resulting projection. The alignment coupling provides an explicit optimal coupling by pairing individuals via the reference space. Hence, assuming $\tilde{\pi}^{g,G} = \pi_\rho^{g,G}$ allows us to derive an expression for the projection and then analyze the subsequent weight and imputation steps. We defer the analysis of the other optimal couplings to Proposition 3 in Section 4.3, where we show that, under Assumptions 3 and 4, every optimal coupling yields the same projection as the alignment coupling. The supposition made here can, therefore, be removed under these assumptions.

To compute the projection through the alignment coupling, we need the conditional distribution of source individuals matched to a given target individual ζ, as in Definition 1. This conditional law answers the question: *given that the target side of a matched pair is ζ, where does the source side land?* Under the alignment coupling, a matched pair takes the form $\left(\rho_g(\check{\xi}), \rho_G(\check{\xi})\right)$, where $\check{\xi} \sim \check{\mu}$. Thus, conditioning on the target individual being ζ is equivalent to conditioning the reference coordinate $\check{\xi}$ on $\rho_G(\check{\xi}) = \zeta$. Because ρ_G need not be injective, the target position ζ need not determine a unique reference coordinate. Hence, the question becomes one about reference coordinates: *given that $\rho_G(\check{\xi}) = \zeta$, how is $\check{\xi}$ distributed?* The disintegration theorem supplies this conditional distribution.

Denote this conditional distribution by $\check{\mu}_\zeta^G$. Since $(\rho_G)_\# \check{\mu} = \mu_G$, the disintegration theorem provides a μ_G-a.e. unique Borel family $\left\{\check{\mu}_\zeta^G\right\}_{\zeta \in \Xi}$ of probability measures on $\check{\Xi}$ satisfying

$$\check{\mu}_\zeta^G\left(\rho_G^{-1}(\{\zeta\})\right) = 1 \quad \text{for } \mu_G\text{-a.e. } \zeta, \qquad \check{\mu} = \int_\Xi \check{\mu}_\zeta^G \mathrm{d}\mu_G(\zeta).$$

Applying ρ_g to a reference coordinate drawn from $\check{\mu}_\zeta^G$ gives the corresponding source individual. Consequently, the projection $\tilde{Q}_g^{(w)}(\zeta, \zeta')$ can be computed by averaging the source reference representation over independent draws from $\check{\mu}_\zeta^G$ and $\check{\mu}_{\zeta'}^G$. This average gives the expected link probability between the source individuals matched to ζ and ζ'.

For $g \in \{1, ..., G-1, G\}$ and $w \in \{0, 1\}$, define the corresponding average of the reference representation $\check{f}_g^{(w)}$ by

$$\bar{f}_g^{(w)}(\zeta, \zeta') := \iint_{\check{\Xi}\times\check{\Xi}} \check{f}_g^{(w)}(\check{\xi}, \check{\xi}') \mathrm{d}\check{\mu}_\zeta^G(\check{\xi}) \mathrm{d}\check{\mu}_{\zeta'}^G(\check{\xi}'), \qquad (\zeta, \zeta') \in \Xi \times \Xi. \tag{5}$$

Under the alignment coupling, the conditional law in Definition 1 satisfies $\tilde{\pi}_\zeta^{g,G} = \left(\rho_g\right)_\# \check{\mu}_\zeta^G$ for μ_G-a.e. ζ. Substituting this identity into the projection and using $P_g^{(w)}\left(\rho_g(\check{\xi}), \rho_g(\check{\xi}')\right) = \check{f}_g^{(w)}(\check{\xi}, \check{\xi}')$ from Definition 3 turns the double integral over source individuals into the double integral

over reference coordinates in (5). Thus, for each source network g, the projection through the alignment coupling is $\bar{f}_g^{(w)}$ in $L^2(\mu_G \otimes \mu_G)$. Lemma A.2 in Appendix A collects the properties of $\bar{f}_g^{(w)}$. The following proposition summarizes its relation to the Gromov-Wasserstein projection through the alignment coupling.

With the reference space representation $\bar{f}_g^{(w)}$ defined, Proposition 1 below shows that it is exactly the Gromov-Wasserstein projection through the alignment coupling.

Proposition 1 (Projections through the alignment coupling)
Fix $g \in \{1, ..., G-1\}$ and suppose the Gromov-Wasserstein optimization problem returns $\tilde{\pi}^{g,G} = \pi_\rho^{g,G}$. The following statements hold:

(*i*) *For both $w \in \{0,1\}$, $\tilde{Q}_g^{(w)} = \bar{f}_g^{(w)}$ in $L^2(\mu_G \otimes \mu_G)$, where $\bar{f}_g^{(w)}$ is defined in (5).*

(*ii*) *For both $w \in \{0,1\}$, $\bar{f}_G^{(w)} = P_G^{(w)}$ in $L^2(\mu_G \otimes \mu_G)$.*

(*iii*) *Suppose $\rho_g = \theta_g \circ \rho_G$ $\check{\mu}$-a.e. for a Borel map $\theta_g : \Xi \to \mathrm{supp}(\mu_g)$, that is, the source position of a reference coordinate is determined by its target position. Then $\left(\theta_g\right)_{\#} \mu_G = \mu_g$ and $\tilde{Q}_g^{(w)} = P_g^{(w)} \circ \left(\theta_g \times \theta_g\right)$ in $L^2(\mu_G \otimes \mu_G)$.*

Proposition 1 (i) indicates that, through the alignment coupling $\pi_\rho^{g,G}$, the projection returns the source graphon read in reference coordinates and then averaged over the reference coordinates that ρ_G cannot distinguish. Statement (ii) shows that this averaging leaves the target's own graphon unchanged. More precisely, for independent draws $\check{\xi}, \check{\xi}' \sim \check{\mu}$,

$$\bar{f}_g^{(w)}\left(\rho_G(\check{\xi}), \rho_G(\check{\xi}')\right) = \mathbb{E}\left[\check{f}_g^{(w)}(\check{\xi}, \check{\xi}') \,\middle|\, \rho_G(\check{\xi}), \rho_G(\check{\xi}')\right] \quad \text{a.s.}$$

Being a conditional expectation, this average can only retain or lose information from $\check{f}_g^{(w)}$. That is, the projection retains the part of the source reference representation determined by the target positions and averages out any remaining variation. In particular,

$$\left\|\bar{f}_g^{(w)}\right\|_{L^2(\mu_G \otimes \mu_G)} \leq \left\|\check{f}_g^{(w)}\right\|_{L^2(\check{\mu} \otimes \check{\mu})}.$$

This potential loss of information arises because the reference space may distinguish reference coordinates more finely than the target's latent characteristics do, and any variation in the source graphon across those indistinguishable reference coordinates is averaged out.

On the other hand, Statement (iii) gives a sufficient condition for no information to be lost when the source position is determined by the target position. In this case, $\bar{f}_g^{(w)} \circ (\rho_G \times \rho_G) = \check{f}_g^{(w)}$ in $L^2(\check{\mu} \otimes \check{\mu})$. This includes the case in which ρ_G is injective. If μ_G has atoms, the non-atomicity of $\check{\mu}$ prevents ρ_G from being injective, and averaging across distinct reference coordinates is

unavoidable at those atoms. Such averaging need not, however, lose information about the source link probabilities, as statement (iii) illustrates.

Note that Proposition 1 requires none of Assumptions 2-4. Its conclusions hold for projection through the alignment coupling irrespective of the objective value $F_{d_g^2,d_G^2}\left(\pi_\rho^{g,G}\right)$, and rely only on the reference space structure and the disintegration of the alignment coupling.

After discussing the properties of the Gromov-Wasserstein projection through alignment coupling, the next subsection uses this expression for the projection to analyze the subsequent weight construction and counterfactual imputation.

4.3 Identification in the Population Problem

Throughout this subsection, we consider the case $G \geq 3$, as when $G = 2$, the simplex $\Delta_{2-2} = \{1\}$ is a singleton and the network weights optimization problem is trivial.

In the pre-treatment fit weights optimization, it can only pin down the weights to the extent that different weight vectors produce different synthetic networks. Since the weights live in the simplex Δ_{G-2}, the difference $v := \delta - \delta'$ between any two candidate weight vectors sums to zero. Hence, the relevant question is whether some zero-sum direction leaves the synthetic combination $\sum_g \delta_g \bar{f}_g^{(0)}$ unchanged.

To answer this, define

$$\underline{V} := \min_{v \in \mathbb{R}^{G-1}: \|v\|_2 = 1, \mathbf{1}^\intercal v = 0} \left\| \sum_{g=1}^{G-1} v_g \bar{f}_g^{(0)} \right\|_{L^2(\mu_G \otimes \mu_G)} \tag{6}$$

and, for $w \in \{0, 1\}$,

$$\mathcal{V}^{(w)} := \left\{ v \in \mathbb{R}^{G-1} : \mathbf{1}^\intercal v = 0 \text{ and } \sum_{g=1}^{G-1} v_g \bar{f}_g^{(w)} = 0 \text{ in } L^2(\mu_G \otimes \mu_G) \right\}, \tag{7}$$

which is a linear subspace of $\mathbb{R}^{G-1}$. The constant $\underline{V}$ is the smallest change in the pre-treatment synthetic network that an unit zero-sum change of the weights can produce. When moving the weights from δ to δ', it changes $\sum_g \delta_g \bar{f}_g^{(0)}$ by at least $\underline{V}\|\delta - \delta'\|_2$ in $L^2(\mu_G \otimes \mu_G)$. The subspace $\mathcal{V}^{(0)}$ collects the directions of weights change that are invisible to the pre-treatment fit. We can observe the relation between these two objects

$$\underline{V} > 0 \iff \mathcal{V}^{(0)} = \{0\} \iff \left\{\bar{f}_g^{(0)}\right\}_{g=1}^{G-1} \text{ is affinely independent in } L^2(\mu_G \otimes \mu_G).$$

Similar to tranditional synthetic control method, the condition $\underline{V} > 0$ is the network analogue of the requirement of no source be an affine combination of the others in the pre-treatment period. A small $\underline{V}$ means that some direction of weight change is nearly invisible to the fit, and along

that direction the weights are only weakly pinned down. Since in the weight optimization step, it can only see the projection, the condition is stated for the averaged objects $\bar{f}_g^{(0)}$.

A reference space counterpart $\underline{\check{V}}$ can be similarly defined by replacing $\bar{f}_g^{(0)}$ with $\check{f}_g^{(0)}$ and $L^2(\mu_G \otimes \mu_G)$ with $L^2(\check{\mu} \otimes \check{\mu})$. It satisfies $\underline{V} \leq \underline{\check{V}}$ by the nature of the average. Again, there is no loss when every ρ_g factors through ρ_G as in Proposition 1 (iii), in which case $\underline{V} = \underline{\check{V}}$. Detailed discuss of such weights variation are collected in Lemma A.3 in Appendix.

Having ensured that the source networks are affinely independent, so that any perturbation of the weights changes the synthetic network, Proposition 2 below shows that, once the projections through the alignment coupling $\pi_\rho^{g,G}$ are $\bar{f}_g^{(w)}$, the convex hull condition of Assumption 2 makes the weights optimization step return δ^* and the exactness of the imputation is acquired.

Proposition 2 (Downstream exactness through the alignment coupling)
Let Assumption 2 hold and suppose that, for every $g \in \{1, ..., G-1\}$, the Gromov-Wasserstein optimization returned the alignment coupling $\tilde{\pi}^{g,G} = \pi_\rho^{g,G}$, so that $\tilde{Q}_g^{(w)} = \bar{f}_g^{(w)}$ for both $w \in \{0, 1\}$ by Proposition 1 (i). Let

$$\mathcal{D} := \arg\min_{\delta \in \Delta_{G-2}} \left\| P_G^{(0)} - \sum_{g=1}^{G-1} \delta_g \tilde{Q}_g^{(0)} \right\|_{L^2(\mu_G \otimes \mu_G)}$$

be the minimizer set of the Gromov-Wasserstein barycenter analysis problem. The following statements hold:

(i) *The minimum value is zero and $\delta^* \in \mathcal{D}$.*

(ii) *$\mathcal{D} = \left\{\delta \in \Delta_{G-2} : \delta - \delta^* \in \mathcal{V}^{(0)}\right\}$. In particular, if $\underline{V} > 0$, then $\mathcal{D} = \{\delta^*\}$.*

(iii) *If the Gromov-Wasserstein barycenter analysis problem returns $\tilde{\delta} = \delta^*$, which is guaranteed whenever $\underline{V} > 0$ by statement (ii), then the counterfactual network imputation is exact, that is, $\tilde{P}_G^{(1)} = P_G^{(1)}$ in $L^2(\mu_G \otimes \mu_G)$.*

Proposition 2 discusses the population counterpart of the analysis problem of Martín et al. (2026). It states that, if the alignment coupling is indeed what the Gromov-Wasserstein optimization returns, then the population problem is solved exactly. As discussed in Remark 2, however, this cannot be guaranteed. The reason is that the optimal coupling is not unique, and the optimization algorithm may return any element of the optimal set. The next proposition resolves this issue by showing that, under additional conditions, the choice among the optimal couplings does not affect the projection.

Adding Assumptions 3 and 4, Proposition 3 states that, although the Gromov-Wasserstein optimal coupling is not unique, every coupling in the optimal set produces the same Gromov-Wasserstein projection, and it is the projection through the alignment coupling.

Proposition 3 (Projection over the optimal set)

Let Assumptions 3 and 4 hold, fix $g \in \{1, ..., G-1\}$, and let π be any optimal coupling of the Gromov-Wasserstein optimization. Then the Gromov-Wasserstein projection through π satisfies, for both $w \in \{0, 1\}$,

$$\tilde{Q}_g^{(w)} = \bar{f}_g^{(w)} \qquad \text{in } L^2(\mu_G \otimes \mu_G).$$

In particular, with Proposition 1 (i), every optimal coupling produces the same projection, and it is the projection produced by the alignment coupling.

Combining the result from the Propositions 2 and 3, Theorem 2 closes the argument of this section.

Theorem 2 (Identification in the population problem)

Let Assumptions 2, 3, and 4 hold, and let $\underline{V} > 0$. Then Algorithm 2 is exact for every family of optimal couplings that the Gromov-Wasserstein optimization may return. That is, whatever optimal couplings the Gromov-Wasserstein optimization outputs, weights optimization step has the unique solution $\tilde{\delta} = \delta^*$, and the final counterfactual network imputation satisfies

$$\tilde{P}_G^{(1)} = P_G^{(1)} \qquad \text{in } L^2(\mu_G \otimes \mu_G).$$

In Theorem 2, since the last step consumes the coupling only through the projection, the placeholder hypothesis of Propositions 1 and 2, that the Gromov-Wasserstein optimization returns the alignment coupling, is no longer needed. Whichever optimal coupling the algorithm returns, the imputation is exact.

Note that the theorem imposes no regularity condition on any μ_g, which may be atomic, and makes no condition to the value of the Gromov-Wasserstein objective, which is in general strictly positive (Remark 2). On top of that, the matching step is not asked to make the two networks look alike, but rather just align them, and the remaining discrepancy is then absorbed by the convex weights.

5 Large Sample Theory

Sections 3 and 4 established in identification results in the population problem. In this section, we return to Algorithm 1, which operates on finite sampled networks, and study how well the imputation $\hat{P}_G^{(1)}$ recovers the missing link probabilities as the networks grow. With the number of networks G fixed, we study the large network setting

$$N_{\min} := \min_{g \in \{1,\dots,G-1,G\}} N_g \to \infty.$$

Throughout this section, Assumptions 1-4 are maintained, and we work with the reference space representation as defined in Section 3. That is, the graphons $\check{f}_g^{(w)}$ and the internal distances $\check{d}_g^2$ are those of Definition 3, the objective of the reference problem is written as $F_{d_g^2,d_G^2}$, and for the sampled individuals we denote them as $\xi_{g,i} = \rho_g(\check{\xi}_{g,i})$, so that $\check{\xi}_{g,i}$ is the reference coordinate of individual i of network g. Note that the reference coordinate $\check{\xi}_{g,i}$ and the maps ρ_g are just the devices for the analysis and they never enter the actual algorithm.

5.1 Assumptions

Unlike the population analysis of Sections 3 and 4, the finite sample analysis has no access to the laws of the measure μ_g or the graphons $P_g^{(w)}$. The algorithm only sees N_g draws from each latent characteristics distribution and one Bernoulli realization of each link, and every population object has to be recovered from these. For such a multi-step estimation procedure, we impose two further assumptions. Assumption 5 discusses the reference representation, graphon regularity, and separation of reference coordinates. These conditions are used for uniform consistency. Assumption 6 adds conditions on the internal geometries for the rate of convergence.

Assumption 5 (Regularity on reference representation and smoothness conditions)**.**
There exists a reference representation satisfying the maintained assumptions such that, for every g, $\rho_g : \check{\Xi} \to \text{supp}(\mu_g)$ is a Borel bijection with Borel inverse and $(\rho_g)_{\#}\check{\mu} = \mu_g$. With this, the following conditions hold:

(i) $\check{\Xi} \subseteq \mathbb{R}^{d_\xi}$ is compact and convex.[12]

(ii) Assume $\text{supp}(\check{\mu}) = \check{\Xi}$. There exists $0 < \underline{c}_\mu \leq \bar{c}_\mu < \infty$ and $r_\mu > 0$ such that

$$\underline{c}_\mu r^{d_\xi} \leq \check{\mu}(B(x,r) \cap \check{\Xi}) \leq \bar{c}_\mu r^{d_\xi}, \qquad x \in \check{\Xi}, \quad 0 < r \leq r_\mu. \tag{8}$$

(iii) Every $\check{f}_g^{(w)}$ is symmetric, locates in $[0,1]$, and Lipschitz in each argument, with a common finite constant L_f:

$$|\check{f}_g^{(w)}(\check{\xi},\check{u}) - \check{f}_g^{(w)}(\check{\zeta},\check{u})| \leq L_f \|\check{\xi} - \check{\zeta}\|.$$

(iv) If $\check{d}_G(\check{\xi},\check{\xi}') = 0$, then $\check{\xi} = \check{\xi}'$ for all $\check{\xi},\check{\xi}' \in \check{\Xi}$.

[12] Convexity can be relaxed to a more general finite short path condition: *there is a constant $L_\Xi < \infty$ such that any two points $x, y \in \check{\Xi}$ can be joined by a curve inside $\check{\Xi}$ of length at most $L_\Xi\|x - y\|$.* This allows other nonconvex supports, with L_Ξ entering the constants of the proofs but not the rates. Note that $\check{\Xi}$ is must at least to be connected. For example, instead of the connected $\check{\Xi}$, suppose that $\check{\Xi}$ is the union of two disjoint balls. A point in one ball and a point in the other cannot be joined by any curve inside $\check{\Xi}$, so the misalignment analysis of Subsection 5.4, which walks from one reference coordinate to another through a path of intermediate reference coordinates, has no path to walk along.

Assumption 5 (i) states that any two reference coordinates in the reference space can be connected by a path lying within the reference space. Assumption 5 (ii) states that every open set of reference coordinates has positive mass, with the lower and upper bounds in (8) ensuring that every reference coordinate $\check{\xi}$ has enough, but not too many, neighbors for the denoising and alignment steps. The bijection condition on ρ_g guarantees that each sampled individual has a well-defined reference coordinate $\check{\xi}_{g,i} = \rho_g^{-1}(\xi_{g,i})$. Since $\check{\mu}$ is non-atomic, a bijective ρ_g exists exactly when μ_g is non-atomic. Thus, this condition restricts the latent laws to non-atomic ones. Assumption 5 (iii) says that individuals with similar reference coordinates have similar neighborhood in terms of link probability, which is what makes averaging over neighbors in the denoising step a valid estimate. Assumption 5 (iv) says that no two distinct reference coordinates look the same from the target's perspective.

Remark 3 (Assumption compatibility with stochastic block models). The condition of non-atomic μ_g for the large sample theory presented in this section excludes the block models. However, this is the limitation of the proof strategy used here rather than of the method. The argument of this section is built for a continuum of reference coordinate $\check{\xi}$. It measures how far the estimated coupling is from a correct alignment by the Euclidean distance between reference coordinates, and it corrects a misaligned pair by walking through a path of intermediate reference coordinates between them. When the reference coordinates only take finitely many values, reference coordinates are either identical or separated by a fixed gap, and hence, the aforementioned approach fails. To tackle such block models, the appropriate object is the block itself, rather than the irrelevant individual reference coordinates within that block. With blocks as the objects of interest, the same three-channel error decomposition argument in Section 5.2 goes through in a more simplified fashion. The smoother of the denoising step with a fixed bandwidth below the smallest inter-block pseudo-distance averages over own-block partners only and has no bias. Then the Gromov-Wasserstein optimal transport reduces to a finite problem over $B \times B$ block couplings, whose optimal set consists of the shuffles preserving the target block geometry, all of which produce the same projection by the shuffle check of Corollary 1 1. And lastly, the convex weights optimization step applies as it is. Since every quantity is then finite dimensional, the resulting rate of convergence should be faster than of Theorem 4. We do not pursue this approach here for generality purpose.

While Assumption 5 states how link probabilities vary with reference coordinates, Assumption 6 further discuss the internal geometries of the networks.

Assumption 6 (Target network structure and cross network relation).
Consider the following conditions

(i) For fixed $0 < \underline{c}_b \leq \bar{c}_b < \infty$, every source $g \in \{1, ..., G-1\}$, and all reference coordinate pairs $(\check{\xi}, \check{\xi}'), (\check{\zeta}, \check{\zeta}') \in \check{\Xi} \times \check{\Xi}$,

$$\underline{c}_b \left| \check{d}_g^2\left((\check{\xi}, \check{\xi}')\right) - \check{d}_g^2\left((\check{\zeta}, \check{\zeta}')\right) \right| \leq \left| \check{d}_G^2\left((\check{\xi}, \check{\xi}')\right) - \check{d}_G^2\left((\check{\zeta}, \check{\zeta}')\right) \right| \leq \bar{c}_b \left| \check{d}_g^2\left((\check{\xi}, \check{\xi}')\right) - \check{d}_g^2\left((\check{\zeta}, \check{\zeta}')\right) \right|.$$

(ii) Let $\mathcal{T}$ be the set of symmetries of the target, that is, Borel bijections $T : \check{\Xi} \to \check{\Xi}$ with $T_{\#}\check{\mu} = \check{\mu}$ and $\check{d}_G(T\check{\xi}, T\check{\xi}') = \check{d}_G(\check{\xi}, \check{\xi}')$ for all $\check{\xi}, \check{\xi}'$. Then there exists $\kappa > 0$, such that, for every coupling $\check{\pi} \in \Pi(\check{\mu}, \check{\mu})$ of the reference measure with itself,

$$F_{\check{d}_G^2, \check{d}_G^2}(\check{\pi}) \geq \kappa \inf_{T \in \mathcal{T}} \int_{\check{\Xi} \times \check{\Xi}} \|T\check{\xi} - \check{\xi}'\|^2 \mathrm{d}\check{\pi}(\check{\xi}, \check{\xi}').$$

Assumption 6 (i) quantifies the relation between source and target dissimilarities that Assumptions 3 and 4 already deliver. That is, whenever two pairs of reference coordinates are at different distances in a source, they are also at different distances in the target, and the two differences are of the same order. The comparison is between changes at arbitrary reference coordinate pairs, not only between their distances from zero.

Assumption 6 (ii) states that any reassignment of reference coordinates that substantially changes them must also substantially distort the target's geometry. It considers arbitrary couplings of the reference measure with itself, not only couplings returned by the Gromov–Wasserstein optimization. And if such a coupling produces a distortion of the target dissimilarities, then it must have some reference coordinate difference up to measure preserving symmetry. Note that, unlike the earlier rigidity condition in Assumption 4 that ensures that the exact symmetries preserve every formation function used in the imputation, Assumption 6 (ii) instead controls deviations from the symmetries and supplies a quantitative bound when the geometry is preserved only approximately. And unlike the upper bound in the Lipschitz condition from Assumption 5 that prevents small reference coordinate movements from producing disproportionately large target network geometric changes, Assumption 6 (ii), on the other hand, provides the lower bound. Similar to the curvature restriction, it excludes excessively flat or weakly identified reassignments while retaining exact symmetries. This assumption also resembles the bi-Lipschitz condition used in the neighborhood smoothing estimator in Zhang et al. (2017) for network denoising, and the informativeness condition imposed by Deaner et al. (2025) for counterfactual outcome imputation in nonlinear panel models.

5.2 Decomposition of the Imputation Error $\left\|\hat{P}_G^{(1)} - P_G^{(1)}\right\|_{\max}$

We start by defining some intermediate objects. For every source $g \in \{1, ..., G-1\}$ and both $w \in \{0, 1\}$, denote

$$\bar{Q}_{g,i,j}^{(w)} := N_G^2 \sum_{k,\ell} \hat{\pi}_{k,i}^{g,G} \hat{\pi}_{\ell,j}^{g,G} P_{g,k,\ell}^{(w)} \qquad \text{and} \qquad \check{Q}_{g,i,j}^{(w)} := \check{f}_g^{(w)}(\check{\xi}_{G,i}, \check{\xi}_{G,j}).$$

where they are the intermediate oracle and the population counterpart of the projection $\hat{Q}_g^{(w)}$. $\bar{Q}_{g,i,j}^{(w)}$ uses the estimated coupling but the true source probabilities. $\check{Q}_{g,i,j}^{(w)}$ is the link probability that source network g assigns to the two individuals occupying the reference coordinates of target individuals i and j. Under the bijective alignment maps in Assumption 5 and with Proposition 1 (iii), it is also the population projection $\vec{f}_g^{(w)}$ from Section 4 evaluated at the sampled target individuals.

With these intermediate objects defined, we then define the three error channels as

$$\begin{aligned} R_{\text{denoise}} &:= \max\left\{ \max_{g \le G} \max_{i,j \le N_g} |\hat{P}_{g,i,j}^{(0)} - P_{g,i,j}^{(0)}|,\ \max_{g<G} \max_{i,j \le N_g} |\hat{P}_{g,i,j}^{(1)} - P_{g,i,j}^{(1)}| \right\}, \\ R_{\text{align}} &:= \max_{g<G} \max_{w \in \{0,1\}} \max_{i,j \le N_G} |\bar{Q}_{g,i,j}^{(w)} - \check{Q}_{g,i,j}^{(w)}|, \\ R_{\text{weight}} &:= \|\hat{\delta} - \delta^*\|_1, \end{aligned} \tag{9}$$

where R_{denoise} represents the error induced by denoising the adjacency matrices, R_{align} indicates the error caused by using the estimated Gromov-Wasserstein optimal coupling for matching, and R_{weight} comes from the error in estimating network weights.

We are now ready to decompose the imputation error $\hat{P}_G^{(1)} - P_G^{(1)}$. By adding and subtracting the aforementioned two intermediate projections, we have

$$\hat{P}_G^{(1)} - P_G^{(1)} = \sum_{g<G} \hat{\delta}_g (\hat{Q}_g^{(1)} - \bar{Q}_g^{(1)}) + \sum_{g<G} \hat{\delta}_g (\bar{Q}_g^{(1)} - \check{Q}_g^{(1)}) + \sum_{g<G} (\hat{\delta}_g - \delta_g^*) \check{Q}_g^{(1)},$$

where we utilize the fact of $P_G^{(w)} = \sum_{g<G} \delta_g^* \check{Q}_g^{(w)}$ for $w \in \{0, 1\}$, which is guaranteed by Assumptions 2 and 5. For the first term, by the definitions of $\hat{Q}_g^{(1)}$ and $\bar{Q}_g^{(1)}$, we have

$$\hat{Q}_{g,i,j}^{(1)} - \bar{Q}_{g,i,j}^{(1)} = \sum_{k,\ell} (N_G \hat{\pi}_{k,i}^{g,G})(N_G \hat{\pi}_{\ell,j}^{g,G})(\hat{P}_{g,k,\ell}^{(1)} - P_{g,k,\ell}^{(1)}).$$

Since $\hat{\pi}^{g,G}$ has nonnegative entries and column sums $1/N_G$, the weights $N_G \hat{\pi}_{k,i}^{g,G} N_G \hat{\pi}_{\ell,j}^{g,G}$ are nonnegative and sum to one over (k, ℓ). Hence, for every feasible collection of couplings,

$$\max_{g<G} \max_{w \in \{0,1\}} \max_{i,j \le N_G} |\hat{Q}_{g,i,j}^{(w)} - \bar{Q}_{g,i,j}^{(w)}| \le R_{\text{denoise}}. \tag{10}$$

For the second term, since each $|\bar{Q}_{g,i,j}^{(1)} - \check{Q}_{g,i,j}^{(1)}| \le R_{\text{align}}$ and $\hat{\delta} \in \Delta_{G-2}$, it is bounded in absolute value by R_{align}. For the third term, since $\check{Q}_{g,i,j}^{(1)} \in [0, 1]$, it is bounded in absolute value by $\sum_{g=1}^{G} |\hat{\delta}_g - \delta_g^*| = R_{\text{weight}}$.

Hence, under Assumptions 2 and Assumption 5, for every sample and every feasible choice of couplings and weights,

$$\|\hat{P}_G^{(1)} - P_G^{(1)}\|_{\max} \le R_{\text{denoise}} + R_{\text{align}} + R_{\text{weight}}. \tag{11}$$

The remaining work is to bound each channel.

5.3 The Denoising Channel

Previously, the denoising step in Section 2 was stated in terms of the latent characteristics $\xi_{g,i}$. Here, we rewrite everything in the reference space representation. The link probability we aim to estimate is $P_{g,i,j}^{(w)} = \check{f}_g^{(w)}(\check{\xi}_{g,i}, \check{\xi}_{g,j})$, and the pseudo distance in (2) is the sample analogue of

$$d_g^{\text{pseudo}}(\check{\xi}_{g,i}, \check{\xi}_{g,k}) := \sup_{\check{\zeta} \in \check{\Xi}} \left| \int_{\check{\Xi}} \left[\check{f}_g^{(0)}(\check{\xi}_{g,i}, \check{\eta}) - \check{f}_g^{(0)}(\check{\xi}_{g,k}, \check{\eta}) \right] \check{f}_g^{(0)}(\check{\zeta}, \check{\eta}) \mathrm{d}\check{\mu}(\check{\eta}) \right|$$

How fast the denoising step converges depends on how informative this pseudo distance is about the actual distance between reference coordinates. Proposition 5 below controls the denoising step error.

Proposition 5 (Denoising step rate of covergence)
Assume Assumptions 1-5 hold. Let R_{denoise} be the denoising error in (9), where the neighborhood smoothing estimator of Step 0 uses a kernel K that is nonnegative, bounded, Lipschitz, supported on $[0,1]$, and bounded below by some $\underline{K} > 0$ on $[0, t_K]$ for some $t_K \in (0,1)$, and where the average defining $\hat{P}_{g,i,j}^{(w)}$ excludes the indices i and j.

(i) If $h_N \to 0$ and

$$\frac{\log(N_{\min})}{N_{\min} h_N^{d_\xi}} \to 0, \qquad \frac{(\log(N_{\min})/N_{\min})^{1/2} + (\log(N_{\min})/N_{\min})^{1/d_\xi}}{h_N^{d_\xi+1}} \to 0,$$

then $R_{\text{denoise}} = o_p(1)$. In particular, these conditions hold for

$$h_N \asymp \left[(\log(N_{\min})/N_{\min})^{1/2} + (\log(N_{\min})/N_{\min})^{1/d_\xi} \right]^{1/(d_\xi+2)}. \tag{12}$$

(ii) Additionally with Assumption 6, if $h_N \to 0$, $(\log(N_{\min})/N_{\min})^{1/2}/h_N \to 0$, and $\log(N_{\min})/(N_{\min} h_N^{d_\xi}) \to 0$, then

$$R_{\text{denoise}} = O_p\left(h_N + \frac{(\log(N_{\min})/N_{\min})^{1/2}}{h_N^{\max\{d_\xi/2,\, 1\}}} \right),$$

and by setting the bandwidth as

$$h_N \asymp \left(\frac{\log(N_{\min})}{N_{\min}} \right)^{1/\max\{d_\xi+2,\, 4\}} \tag{13}$$

we have $R_{\text{denoise}} = O_p((\log(N_{\min})/N_{\min})^{1/\max\{d_\xi+2,\,4\}})$.

Proposition 5 is the network counterpart of Theorem 1 of Deaner et al. (2025), where they have the kernel smoothed counterfactual mean estimator for the nonlinear panel model, utilizing T_0 amount of the pre-treatment history used to build the pseudo distance. As first implemented in Zhang et al. (2017), under the network context, the large network size is used for the denoising purpose, similar to the use of long pre-treatment history in the panel model.

Note that the inequality (10) shows that the denoising error would not be amplified by the Gromov-Wasserstein optimal coupling alignment and projection step. The next subsection studies the remaining error caused by matching different reference coordinates.

5.4 The Misalignment Channel

The alignment error R_{align} compares two projections of the true source probabilities: $\bar{Q}_g^{(w)}$, which projects them through the estimated coupling $\hat{\pi}^{g,G}$, and $\check{Q}_g^{(w)}$, which evaluates them at the reference coordinates of the target individuals, that is, projects them through the population alignment. Since both use the same source probabilities, R_{align} reflects only the error that comes from using the estimated coupling. If every source-target pair receiving positive mass under $\hat{\pi}^{g,G}$ had same reference coordinates, the two projections would coincide. And more generally, by the Lipschitz condition of Assumption 5 (iii), if those reference coordinates are close, so are the two projections. The remaining task is therefore to show that the estimated coupling only pairs source and target individuals whose reference coordinates are close.

However, the utilization of estimated Gromov-Wasserstein coupling makes this task nontrivial. First, since its inputs are estimated internal distances and sampled reference coordinates, $\hat{\pi}^{g,G}$ is only an approximate minimizer of the reference problem (4). Second, the non-convex Gromov-Wasserstein objective function offers no guarantee that an approximate minimizer lies close to an exact one. Third, even the exact minimizers are not unique. By Theorem 1 (ii), the Gromov-Wasserstein alignment is determined only up to symmetries of the target geometry. Such symmetries preserve the graphons under Assumption 4 and therefore do not affect the projected link probabilities. We consequently measure discrepancies between matched reference coordinates after allowing for these symmetries.

Let $\mathcal{T}$ be the symmetry group defined in Assumption 6 (ii), and define

$$\epsilon := \max_{g<G} \inf_{T\in\mathcal{T}} \max_{(k,i):\hat{\pi}_{k,i}^{g,G}>0} \|T\check{\xi}_{g,k} - \check{\xi}_{G,i}\|. \tag{14}$$

For each source network g, the condition $\hat{\pi}_{k,i}^{g,G} > 0$ indicates the source-target pairs to which the estimated coupling assigns positive mass. The inner maximum measures the largest reference coordinate discrepancy among these pairs after applying a common symmetry T to the source

reference coordinates. Taking the infimum over T allows for the choice of symmetry separately for each source network, and the outer maximum takes the largest discrepancy across sources. Thus, ϵ measures the worst matching discrepancy after allowing for target symmetries, regardless of how much positive mass a pair receives.

This matching discrepancy directly controls the alignment error. Under Assumptions 1-5, invariance of the graphons under $\mathcal{T}$ and their Lipschitz continuity imply that every feasible collection of couplings satisfies

$$R_{\text{align}} = \max_{g<G} \max_{w\in\{0,1\}} \max_{i,j\le N_G} |\bar{Q}^{(w)}_{g,i,j} - \check{Q}^{(w)}_{g,i,j}| \le 2L_f\epsilon \quad \text{and} \quad \max_{g<G} \max_{w\in\{0,1\}} \max_{i,j\le N_G} |\hat{Q}^{(w)}_{g,i,j} - \check{Q}^{(w)}_{g,i,j}| \le R_{\text{denoise}} + R_{\text{align}}. \tag{15}$$

For the first inequality, each source pair receiving positive weight in the projection differs from the corresponding target pair by at most ϵ in each reference coordinate, after allowing for a symmetry. Lipschitz continuity therefore bounds the link probability discrepancy by $2L_f\epsilon$, and averaging over source pairs preserves this bound. The second inequality follows by combining (10) with the definition of R_{align}.

It remains to show that ϵ vanishes for the couplings returned by the Gromov-Wasserstein optimization. This requires more than showing that the estimated internal distances and empirical reference coordinate distributions approach their population counterparts. The Gromov-Wasserstein objective averages discrepancies in pairwise internal distances over the coupling, whereas ϵ is the largest reference coordinate discrepancy among all pairs receiving positive mass. Hence, near-optimality of $\hat{\pi}^{g,G}$ can control the alignment only on average, and the analysis must convert this average control into uniform control over matched pairs.

Proposition 6 establishes this connection. Under Assumptions 1-5, consistency of the denoising step implies that ϵ, and hence R_{align}, converges to zero. Assumption 6 supplies the additional quantitative conditions needed to obtain a convergence rate.

Proposition 6 (The alignment channel)
Suppose Assumptions 1-5 hold, and let $\hat{\pi}^{g,G}$ be any global minimizers of the Gromov-Wasserstein problems in Step 1 of Algorithm 1.

(*i*) *If $R_{\text{denoise}} = o_p(1)$, then $\epsilon = o_p(1)$, and hence $R_{\text{align}} = o_p(1)$.*

(*ii*) *If in addition Assumption 6 holds and $R_{\text{denoise}} = O_p(r_N)$ for $r_N \to 0$, then*

$$\epsilon = O_p\left(\left[r_N^2 + (\log(N_{\min}))^{3/2} N_{\min}^{-1/\max\{d_\xi, 2\}}\right]^{1/(d_\xi+2)}\right). \tag{16}$$

Proposition 6 (i) states that, under the maintained assumptions and as the networks grow, consistency of the denoising step ensures that the estimated coupling approaches the correct alignment, up to a symmetry of the target.

Proposition 6 (ii) provides the convergence rate under Assumption 6. The exponent $1/(d_{\check{\xi}}+2)$ arises when passing from average to worst-case alignment. The two terms inside the bracket of (16) correspond to the two sources of error entering the matching: estimating the internal distances from denoised probabilities and representing the reference distribution with finitely many sampled individuals. The first contributes the squared denoising error. The second combines two effects: randomness in the proportion of sampled individuals falling in a given region of the reference space, and the distances over which population mass must be moved to reach the sampled reference coordinates. These give the scales $N_{\min}^{-1/2}$ and $N_{\min}^{-1/d_{\check{\xi}}}$, respectively. The bound used here is governed by the slower of the two, giving $N_{\min}^{-1/\max\{d_{\check{\xi}},2\}}$, with the two scales coinciding at $d_{\check{\xi}} = 2$. In higher dimensions, the sampled reference coordinates are more spread out, making the reference distribution harder to approximate. The logarithmic factor accommodates corrections in the empirical measure bound and the concentration needed for a high-probability statement.

With R_{denoise} and R_{align} controlled, the projection $\hat{Q}_g^{(w)}$ is close to its population counterpart $\check{Q}_g^{(w)}$ by (15). The remaining channel for the imputation error is the weights optimization.

5.5 The Weights Channel

The weight error R_{weight} measures how far the estimated weights $\hat{\delta}$ are from the true weights δ^*.

The weight optimization uses the squared Frobenius criterion. For the analysis, we use its non-squared version, which has the same minimizers:

$$\hat{L}(\delta) := \frac{1}{N_G}\left\|\hat{P}_G^{(0)} - \sum_{g<G}\delta_g \hat{Q}_g^{(0)}\right\|_{\text{F}}, \qquad \delta \in \Delta_{G-2}. \tag{17}$$

To tackle the weight error, we compare $\hat{L}$ with the criterion formed from the true probabilities,

$$\check{L}(\delta) := \frac{1}{N_G}\left\|P_G^{(0)} - \sum_{g<G}\delta_g \check{Q}_g^{(0)}\right\|_{\text{F}},$$

evaluated at the sampled target individuals. Under the maintained assumptions, the convex hull condition of Assumption 2 holds at every sampled pair, and hence $\check{L}(\delta^*) = 0$. The condition $\underline{V} > 0$ ensures that different weight vectors produce different population synthetic networks. As the target network size grows, the sample counterpart of $\underline{V}$ is bounded below by $\underline{V}/2$ with probability tending to one. And on this event, moving the weights away from δ^* increases $\check{L}$. Consequently, if $\hat{L}$ is uniformly close to $\check{L}$, its minimizers must be close to δ^*. Proposition 7

formalizes this argument and bounds the weight error in terms of the denoising and alignment errors.

Proposition 7 (The weights channel)
Suppose Assumptions 1-5 hold, let $G \geq 3$, and assume $\underline{V} > 0$. Then, with probability tending to one, every minimizer of (17) satisfies

$$R_{\text{weight}} \leq \frac{2\sqrt{G-1}}{\underline{V}} \left(4R_{\text{denoise}} + 2R_{\text{align}}\right).$$

The bound shows that weight estimation introduces no additional bottleneck. It transmits the denoising and alignment errors, with their effect amplified when $\underline{V}$ is small and different weight vectors produce similar synthetic networks.

The next subsection assembles these bounds through (11) into the main results of this section.

5.6 Main Results

We now combine the aforementioned three imputation error channels. The first result uses Assumptions 1-5 and delivers uniform consistency, without the quantitative restrictions of Assumption 6.

Theorem 3 (Uniform consistency)
Suppose Assumptions 1-5 hold, and $\underline{V} > 0$. Let $\hat{\pi}^{g,G}$ be global minimizers of the Gromov-Wasserstein problems of Step 1 of Algorithm 1, and let $\hat{\delta}$ be a minimizer of (17). If the first-stage estimates satisfy $R_{\text{denoise}} = o_p(1)$, then

$$\|\hat{P}_G^{(1)} - P_G^{(1)}\|_{\max} = o_p(1),$$

and likewise the projection error $\max_{g<G} \max_{w\in\{0,1\}} \max_{i,j\leq N_G} |\hat{Q}_{g,i,j}^{(w)} - \check{Q}_{g,i,j}^{(w)}|$ and the weight error R_{weight} are $o_p(1)$. In particular, the neighborhood smoothing estimator of Proposition 5, with bandwidth (12), satisfies the first-stage premise.

Theorem 3 does not require either clause of Assumption 6. The reference support, local mass, graphon regularity, and reference coordinate separation in Assumption 5 suffice for uniform consistency with genuinely multidimensional reference coordinates.

The second result adds Assumption 6 and delivers a max norm rate of convergence.

Theorem 4 (Max norm rate of convergence)
Suppose Assumptions 1-6 and $\underline{V} > 0$ hold, with the same optimization convention as in Theorem 3. If

$R_{\text{denoise}} = O_p(r_N)$ *for a* $r_N \to 0$, *then*

$$\|\hat{P}_G^{(1)} - P_G^{(1)}\|_{\max} = O_p\left(\left[r_N^2 + (\log(N_{\min}))^{3/2} N_{\min}^{-1/\max\{d_\xi, 2\}}\right]^{1/(d_\xi+2)}\right),$$

and the same bound holds for the projection error and the weight error. In particular, with the neighborhood smoothing kernel estimator indicated in Proposition 5 and the bandwidth of (13), it follows that

$$\|\hat{P}_G^{(1)} - P_G^{(1)}\|_{\max} = O_p\left((\log(N_{\min}))^{\frac{3}{2(d_\xi+2)}} N_{\min}^{-\frac{1}{\max\{d_\xi,2\}(d_\xi+2)}}\right).$$

6 Simulation

In this section, we demonstrate the finite sample performance of the proposed Synthetic Counterfactual Network method across different simulation designs. Here, we compare the proposed method with other benchmarks through various network statistics.

Common setup for sampling and implementation. Throughout this section, we consider $G = 10$ networks, where networks $g \in \{1, ..., G-1\}$ are source networks and network G is the target network. Network sizes are allowed to differ across networks, and each network size ranges uniformly from the $[N_{\min}, N_{\max}]$ in each setting.

A common feature across all simulation designs is how the target network environment is generated from the source environments. In every replication, we first draw the source network environments and independently draw

$$\delta^* \sim \text{Dirichlet}\left(\alpha_\delta \mathbf{1}_{G-1}\right), \qquad \alpha_\delta = 0.375. \tag{18}$$

The target environment is then constructed using the same generating weights δ^* in both treatment statuses. The choice $\alpha_\delta < 1$ results in a concentrated but non-degenerate mixture with source networks. That is, a relatively small number of sources usually carries most of the total weight, while all other source weights remain small but positive. This construction does not impose which sources are important before the draw.

For the practical implementation, the Epanechnikov kernel $K(x) = \frac{3}{4}(1 - x^2)\mathbf{1}\{|x| \le 1\}$ is used. Note that the Gromov-Wasserstein optimization is non-convex. For the numerical optimization detail, in the implementation, each source-target optimization problem is solved from several feasible initial couplings and the solution with the smallest achieved objective is retained. This numerical procedure does not guarantee global optimality.

We consider three network-size settings, $[N_{\min}, N_{\max}] \in \{[100, 110], [200, 220], [500, 550]\}$, and use 500 Monte Carlo replications for each setting. Rather than selecting the smoothing band-

width from the realized data[13], we report sensitivity across the fixed grid

$$\mathcal{H} = \{0.025, 0.05, 0.075, 0.10, 0.15, 0.20\}.$$

The infeasible oracle estimators do not depend on the bandwidth h.[14]

Benchmarks and estimation performance evaluation. For every design and replication, we compare the proposed Synthetic Counterfactual Network method with five benchmark procedures.

(i) *Nearest Gromov-Wasserstein source.* We select the source network with the smallest pre-treatment Gromov-Wasserstein distance,

$$\hat{g}_{\mathrm{GW}} = \arg\min_{g<G} \widehat{\mathrm{GW}}(g, G),$$

and use its projected post-treatment network, $\hat{P}^{(1)}_{G,\mathrm{GW}} = \hat{Q}^{(1)}_{\hat{g}_{\mathrm{GW}}}$.

(ii) *Best single projected source.* We select the source whose projected pre-treatment probability matrix best fits the target pre-treatment probability matrix,

$$\hat{g}_{\mathrm{single}} = \arg\min_{g<G} \left\| \hat{P}^{(0)}_{G} - \hat{Q}^{(0)}_{g} \right\|_F^2,$$

and use the corresponding projected post-treatment network, $\hat{P}^{(1)}_{G,\mathrm{single}} = \hat{Q}^{(1)}_{\hat{g}_{\mathrm{single}}}$.

(iii) *Uniform Gromov-Wasserstein barycenter.* We use the same Gromov-Wasserstein projections as the proposed method but replace the estimated network weights with equal weights, $\delta_g = \frac{1}{G-1}$ for $g < G$. The counterfactual network is therefore $\hat{P}^{(1)}_{G,\mathrm{uniform}} = \frac{1}{G-1}\sum_{g=1}^{G-1} \hat{Q}^{(1)}_{g}$.

(iv) *Degree alignment.* We replace the Gromov-Wasserstein coupling by a coupling $\hat{\pi}^{g,G}_{\mathrm{deg}}$ that matches individuals according to their pre-treatment degree ranks.[15] Let $\hat{Q}^{(w)}_{g,\mathrm{deg}}$ denote the corresponding projected source network. We then estimate the synthetic weights using

[13] For a more data-driven bandwidth selection procedure, the edge cross-validation bandwidth selection studied by Li et al. (2020) can be further implemented.

[14] When no individual receives positive kernel weight for a given smoothing average, we use the average observed link probability to the corresponding individual, excluding the two endpoints, as a fallback. This occurs mainly at very small bandwidths and becomes much less frequent as the network size or bandwidth increases.

[15] For the degree-alignment benchmark, we first compute each individual's normalized pre-treatment degree, $\hat{\mathrm{deg}}^{(0)}_{g,i} = \frac{1}{N_g-1}\sum_{j\neq i} \hat{P}^{(0)}_{g,i,j}$, and rank individuals within each network from lowest to highest degree. We then match source and target individuals by their degree quantiles. Note that if the network sizes are equal, individuals with the same rank are matched; otherwise, each individual's rank interval, $\left[\frac{r_g(i)-1}{N_g}, \frac{r_g(i)}{N_g}\right]$, is matched to the overlapping rank intervals in the target network. This produces a feasible coupling $\hat{\pi}^{g,G}_{\mathrm{deg}}$ with uniform marginals. We then replace the Gromov-Wasserstein coupling by $\hat{\pi}^{g,G}_{\mathrm{deg}}$, construct the projected source networks in the same way as in Step 1, estimate the synthetic weights using the same simplex optimization as in Step 2, and use these weights and projections to impute the target post-treatment network.

the same simplex optimization as the proposed method,

$$\hat{\delta}_{\text{deg}} \in \arg\min_{\delta \in \Delta^{G-2}} \left\| \hat{P}_G^{(0)} - \sum_{g=1}^{G-1} \delta_g \hat{Q}_{g,\text{deg}}^{(0)} \right\|_F^2,$$

and impute $\hat{P}_{G,\text{deg}}^{(1)} = \sum_{g=1}^{G-1} \hat{\delta}_{\text{deg},g} \hat{Q}_{g,\text{deg}}^{(1)}$.

(v) *No change.* We assume that the target network does not change after treatment and use its pre-treatment probability matrix directly, $\hat{P}_{G,\text{nochange}}^{(1)} = \hat{P}_G^{(0)}$.

For each procedure, we report both the feasible version based on the denoised adjacency matrices and an oracle version that replaces the denoised probability matrices by their true values. The oracle counterparts are used to isolate the loss contributed from the denoising step. Note that the feasible-oracle difference also incorporates the downstream effect of estimation error on the Gromov-Wasserstein couplings, projections, and estimated network weights.

Our primary object of interest is the pair-specific counterfactual link probability. Denote

$$E_{i,j} = \hat{P}_{G,i,j}^{(1)} - P_{G,i,j}^{(1)} \qquad \text{for } i \neq j.$$

We report the root mean squared error $\text{RMSE} = \left[1/\binom{N_G}{2} \sum_{i<j} E_{i,j}^2 \right]^{1/2}$, the median absolute error, and the 95-th percentile of $|E_{i,j}|$. These statistics summarize the overall accuracy, the error for a typical pair, and the upper tail of the pair-specific imputation errors, respectively.

We additionally examine whether the method recovers the treatment-induced change in individual relationships. Define $\Delta_{i,j} = P_{G,i,j}^{(1)} - P_{G,i,j}^{(0)}$ and $\hat{\Delta}_{i,j} = \hat{P}_{G,i,j}^{(1)} - \hat{P}_{G,i,j}^{(0)}$. We report the median absolute error of $\hat{\Delta}_{i,j}$ and, among pairs satisfying $|\Delta_{i,j}| \geq 0.02$, the fraction for which the direction of the change is correctly recovered.

Lastly, we study whether accurate pair-level imputation translates into accurate network summaries. For every target individual, define the normalized expected degree $\text{deg}_i(P) = \frac{1}{N_G - 1} \sum_{j \neq i} P_{i,j}$. We report the median absolute error of the individual expected degrees. We also report the absolute error in the normalized expected triangle count $\left(\text{tr}(P^3)/6\right) / \binom{N_G}{3}$, which summarizes whether the imputed counterfactual network recovers the extent of triadic structure in the target network.

6.1 Design 1: One-dimensional Homophily and Popularity Network Formation

The first design considers a one-dimensional latent reference coordinate $\check{\xi}$ and combines three common network formation mechanisms: homophily, popularity, and baseline density. In this simulation design, we aim to study how the proposed method does in a relatively simple setting with heterogeneous network formation environment across different network and after treatment.

Here, we consider simple one-dimensional quantile space as the reference. For every network, draw reference coordinate $\check{\xi}_{g,i} \overset{\text{i.i.d.}}{\sim} \text{Uniform}[0,1]$. With the reference coordinates drawn, we then map each $\check{\xi}$ to a different distribution of individual latent characteristics by $\xi_{g,i} = \rho_g(\check{\xi}_{g,i})$. Specifically, after mapping, we will have

$$\xi_{g,i} \overset{\text{i.i.d.}}{\sim} \mu_g = \text{Beta}\Big(1 + 4\gamma_g, 1 + 4(1-\gamma_g)\Big), \qquad \text{for } i \in [N_g] \qquad \text{for } g \in [G].$$

With this design, μ_g will have the mode of $\gamma_g \in [0.15, 0.85]$ and the concentration of $\kappa = 6$. Note that here the inverse of the mapping ρ_g^{-1} will be the CDF of a Beta$\Big(1 + \gamma_g(\kappa-2), 1 + (1-\gamma_g)(\kappa-2)\Big)$ distribution. Hence, different networks have different latent characteristic distributions even though their reference coordinates are drawn from the same reference measure.

For the source networks, draw the scalar network-specific effects

$$\gamma_g = 0.15 + 0.7 \cdot B_g, \qquad B_g \sim \text{Beta}(1.5, 1.5), \qquad \text{for } g < G,$$

and construct the target effect using the generating weights of (18), $\gamma_G = \sum_{g=1}^{G-1} \delta_g^* \gamma_g$.

The link probability function takes the form

$$f^{(w)}(\xi, \xi', \gamma) = \sum_{k=1}^{3} \psi_k^{(w)}(\gamma)\phi_k(\xi, \xi'), \qquad w \in \{0,1\}. \tag{19}$$

The three functions ϕ_k describe common formation mechanisms, while $\psi^{(w)}(\gamma)$ determines their relative importance in each network and period. The loadings are nonnegative and sum to one, and each ϕ_k takes values in $[0,1]$, so (19) gives a valid link probability. The formation mechanisms are

$$\phi_1(\xi, \xi') = \exp\left(-\frac{\left[\rho_g^{-1}(\xi) - \rho_g^{-1}(\xi')\right]^2}{2 \cdot (0.12)^2}\right), \qquad \phi_2(\xi, \xi') = \rho_g^{-1}(\xi)\rho_g^{-1}(\xi'), \qquad \phi_3(\xi, \xi') = 0.35,$$

and the loadings are

$$\psi^{(0)}(\gamma) = (1-\gamma) \cdot (0.75, 0.05, 0.20) + \gamma \cdot (0.30, 0.55, 0.15),$$
$$\psi^{(1)}(\gamma) = (1-\gamma) \cdot (0.15, 0.55, 0.30) + \gamma \cdot (0.05, 0.75, 0.20).$$

The first favors links between individuals with similar characteristics. The second gives higher link probabilities to pairs with larger values of ξ and ξ'. The third is a constant baseline. In every network, the intervention reduces the weight on homophily and increases the weight on popularity.

The treatment changes the relative importance of the formation mechanisms substantially, where the pre-treatment network puts more weight on homophily, while the post-treatment network puts more weight on popularity. At the same time, the ratio between the homophily

and popularity components varies across networks. Note that this design does not impose that the internal geometries are perfectly aligned across all networks, and hence, it provides a useful first robustness exercise for the alignment conditions.

Results. Figure 1 shows that the performance of the proposed method depends on the kernel smoothing bandwidth. With bandwidth being too small at $h = 0.025$, at least 90% of the nodes cannot find proper neighbor for kernel smoothing and use the fallback in each size setting, and mean RMSE remains high. Increasing the bandwidth initially improves prediction, but larger bandwidths eventually include too many neighbors and remove useful variation in connection profiles. The lowest mean SCN RMSE over the reported grid occurs at $h = 0.10$, 0.075, and 0.05 for $N_g \in [100, 110]$, $[200, 220]$, and $[500, 550]$, respectively.

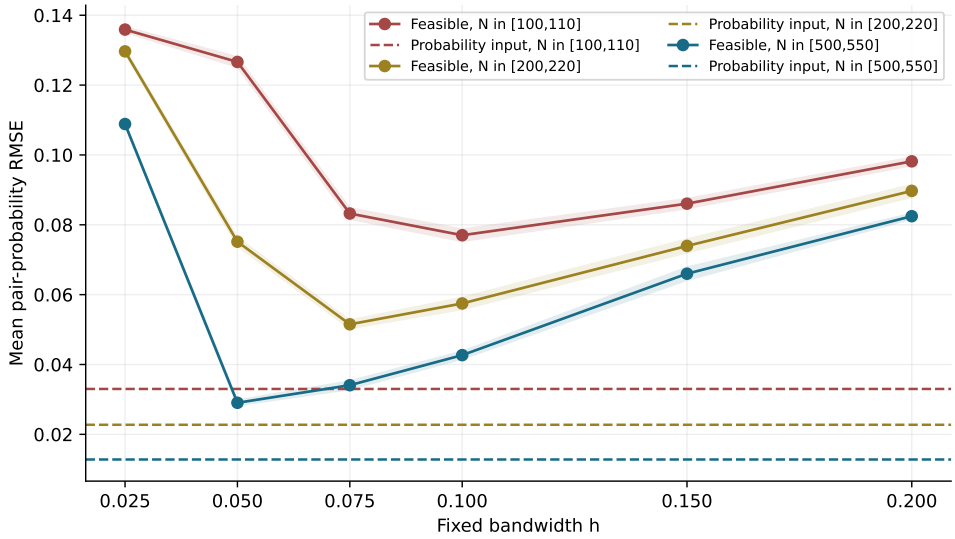


Figure 1: SCN pair-probability RMSE across fixed bandwidths for Design 1

Note: Solid curves report Monte Carlo means over 500 replications, with pointwise 95% Monte Carlo intervals. Dashed lines report the corresponding probability-input oracle means, which do not depend on the bandwidth.

Table 1 reports all seven evaluation measures for every feasible procedure and its oracle counterpart at the largest network size $N_g \in [500, 550]$.[16] Feasible SCN has the smallest RMSE 0.0290, compared with all other benchmark procedures: 0.0652 for the nearest Gromov-Wasserstein source, 0.0448 for the best single projected source, 0.0509 for uniform Gromov-Wasserstein weighting, 0.0798 for degree alignment, and 0.1506 for no change. SCN also has smaller median and upper-tail probability errors than each benchmark.

The treatment-change measures give a similar feasible ranking. SCN has median change error 0.0316 and direction accuracy 95.0%, compared with errors of 0.0515, 0.0380, 0.0437, and

[16]The complete results for every network size setting and bandwidth are reported in Appendix D.

Table 1: Simulation Results of Design 1 ($N_g \in [500, 550]$, $h = 0.05$)

Method	RMSE	MedAE	Q95	Δ MedAE	Direction	Degree	Triangle
Feasible procedures							
SCN	0.0290 (0.0004)	0.0185 (0.0003)	0.0571 (0.0007)	0.0316 (0.0002)	0.950 (0.0004)	0.0117 (0.0003)	0.0021 (0.0001)
Nearest GW source	0.0652 (0.0027)	0.0419 (0.0020)	0.1294 (0.0052)	0.0515 (0.0018)	0.895 (0.0038)	0.0274 (0.0018)	0.0028 (0.0001)
Best single source	0.0448 (0.0001)	0.0273 (0.0001)	0.0905 (0.0002)	0.0380 (0.0001)	0.923 (0.0006)	0.0138 (0.0001)	0.0024 (0.0001)
Uniform GW barycenter	0.0509 (0.0017)	0.0345 (0.0012)	0.0983 (0.0032)	0.0437 (0.0010)	0.931 (0.0018)	0.0274 (0.0011)	0.0045 (0.0001)
Degree alignment	0.0798 (0.0007)	0.0311 (0.0003)	0.1657 (0.0016)	0.0395 (0.0003)	0.937 (0.0004)	0.0221 (0.0002)	0.0027 (0.0001)
No change	0.1506 (0.0006)	0.0946 (0.0004)	0.3124 (0.0013)	0.0954 (0.0004)	0.000 (0.0000)	0.0435 (0.0003)	0.0021 (0.0001)
Oracle procedures							
SCN	0.0128 (0.0004)	0.0080 (0.0003)	0.0234 (0.0008)	0.0080 (0.0003)	0.997 (0.0003)	0.0073 (0.0003)	0.0016 (0.0001)
Nearest GW source	0.0232 (0.0019)	0.0157 (0.0014)	0.0433 (0.0037)	0.0157 (0.0014)	0.986 (0.0028)	0.0144 (0.0012)	0.0028 (0.0001)
Best single source	0.0133 (0.0002)	0.0083 (0.0002)	0.0244 (0.0004)	0.0083 (0.0002)	0.998 (0.0001)	0.0076 (0.0002)	0.0021 (0.0001)
Uniform GW barycenter	0.0279 (0.0013)	0.0193 (0.0010)	0.0528 (0.0026)	0.0193 (0.0010)	0.984 (0.0015)	0.0171 (0.0008)	0.0033 (0.0001)
Degree alignment	0.0640 (0.0010)	0.0174 (0.0004)	0.1218 (0.0022)	0.0174 (0.0004)	0.971 (0.0006)	0.0138 (0.0003)	0.0022 (0.0001)
No change	0.1519 (0.0006)	0.0954 (0.0004)	0.3250 (0.0012)	0.0954 (0.0004)	0.000 (0.0000)	0.0442 (0.0003)	0.0020 (0.0001)

Notes: Each entry is the mean of a replication-level statistic across 500 replications, with its Monte Carlo standard error in parentheses. RMSE, MedAE, and Q95 concern absolute errors in target post-treatment link probabilities. Δ MedAE concerns absolute errors in treatment-induced probability changes. Direction is the correctly predicted sign fraction among pairs with $|\Delta_{i,j}| \geq 0.02$. Degree is the median absolute error in normalized expected degrees. Triangle is the absolute error in normalized expected triangle density. Smaller values are better except for Direction.

0.0395 for the nearest source, best single source, uniform weighting, and degree alignment, respectively. Their corresponding direction accuracies are 89.5%, 92.3%, 93.1%, and 93.7%. No change has median change error 0.0954 and zero direction accuracy because it predicts a zero change for every pair. Thus, the pair-probability gains extend to recovery of treatment-induced changes in individual relationships.

The network-summary results require a separate interpretation. SCN's normalized-degree error is 0.0117, below the benchmark errors, which range from 0.0138 for the best single source to 0.0435 for no change. Its triangle-density error is 0.0021, compared with 0.0028 for the nearest source, 0.0024 for the best single source, 0.0045 for uniform weighting, and 0.0027 for degree alignment. However, no change also has triangle-density error close to 0.0021, despite its much larger pair-probability and degree errors. Accurate recovery of this aggregate statistic alone therefore does not establish accurate recovery of individual relationships.

The lower panel shows how the comparisons change when the true link probability matrices are supplied and the denoising step is not needed. Oracle SCN has RMSE 0.0128, median error 0.0080, and upper-tail error 0.0234, beating all other benchmark procedures. Similarly for other network statistics and evaluations. Hence, under the ideal data generating process of this simulation design, Oracle SCN accurately recovers the target network structure and down to individual pair information.

6.2 Design 2: Complex Mixing with No Explicit Reference Representation

In the next simulation design, we aim to examine the robustness of the proposed method under more complex and not ideal setting. We retain the individual latent characteristic generation mechanism of Design 1 and mix the four graphon functions considered by Zhang et al. (2017). Unlike the previous simulation design having explicit reference coordinate $\check{\xi}$ and map ρ_g, here each graphon is evaluated at the raw characteristics ξ, rather than at the reference coordinates $\rho_g^{-1}(\xi)$. Consequently, variation in the network environment changes both the graphon loadings and the distribution of the arguments entering link formation.

Again, for each source network, draw

$$\gamma_g = 0.15 + 0.70B_g, \qquad B_g \sim \text{Beta}(1.5, 1.5), \qquad g < G,$$

and construct $\gamma_G = \sum_{g<G} \delta_g^* \gamma_g$ using the weights in (18). Also, we have $\xi_{g,i} \overset{\text{i.i.d.}}{\sim} \text{Beta}(1 + 4\gamma_g, 1 + 4(1 - \gamma_g))$. The characteristics remain fixed across treatment statuses.

Following Zhang et al. (2017), define

$$\phi_1(\xi,\xi') = \begin{cases} k/(7+1), & \xi,\xi' \in I_k = \begin{cases} [(k-1)/7, k/7), & \text{for } k < 7 \\ [(7-1)/7, 1], & \text{for } k = 7 \end{cases} \\ 0.3/(7+1), & \text{otherwise,} \end{cases}$$

$$\phi_2(\xi,\xi') = 0.5 + 0.5\sin\{5\pi(\xi+\xi'-1)+1\},$$

$$\phi_3(\xi,\xi') = 1 - \left[1 + \exp\{15(0.8|\xi-\xi'|)^{4/5} - 0.1\}\right]^{-1},$$

$$\phi_4(\xi,\xi') = 0.15 + \frac{\xi^2+\xi'^2}{3}\cos\{1/(\xi^2+\xi'^2)\},$$

where $\phi_4(0,0) = 0.15$ by continuity. These templates combine a block structure, an oscillation in the sum of characteristics, a nonlinear distance effect, and an oscillation in their squared radius. The pre- and post-treatment loading vectors are

$$\psi^{(0)}(\gamma) = (1-\gamma)(0.45, 0.20, 0.20, 0.15) + \gamma(0.15, 0.45, 0.15, 0.25),$$
$$\psi^{(1)}(\gamma) = (1-\gamma)(0.15, 0.45, 0.20, 0.20) + \gamma(0.35, 0.15, 0.15, 0.35).$$

Each loading is positive and the loadings sum to one. All four templates therefore enter both periods, and their mixture gives valid probabilities. The loading on ϕ_3 is unchanged across periods, so the intervention does not operate through a change in the weight on this relatively dense component. For $i \neq j$, set

$$P^{(w)}_{g,i,j} = f^{(w)}(\xi_{g,i}, \xi_{g,j}, \gamma_g), \qquad f^{(w)}(\xi,\xi',\gamma) = \sum_{k=1}^{4} \psi_k^{(w)}(\gamma)\phi_k(\xi,\xi').$$

And the affineness of the loadings yields the exact identity

$$f^{(w)}(\xi,\xi',\gamma_G) = \sum_{g<G} \delta_g^* f^{(w)}(\xi,\xi',\gamma_g), \qquad w \in \{0,1\},$$

at fixed raw characteristics (ξ,ξ').

However, the reference representation is not as explicit as the previous simulation design. Unlike design 1, because quantile maps vary across networks, the identity at fixed raw values does not imply the same mixture identity in reference coordinates. Matching equal raw values also does not give a feasible population coupling when the source and target characteristic distributions differ. In addition, internal profile distances integrate against network-specific characteristic distributions, and common ordering of the pre-treatment geometries is not imposed. Therefore, we consider this a robustness numerical stress test without imposing an explicit reference space convex hull and alignment conditions, despite there possibly being some underlying reference or some sort of approximation.

Results To demonstrate the complexity of the data-generating process in Simulation Design 2, as well as the performance of the proposed method, Figure 2 presents the true post-treatment target link probability matrix $P_G^{(1)}$, the oracle estimated link probability matrix $\tilde{P}_G^{(1)}$, and the feasible estimates of the link probability matrix obtained using the proposed method $\hat{P}_G^{(1)}$ from one of the Monte Carlo replications. In the true probability matrix, we can see that the block structure is fused with other network formation models, making it very difficult to denoise. Despite that, with a large enough network size and a suitable bandwidth for neighborhood kernel smoothing, the proposed method recovers the counterfactual network really well. If we remove the error channel caused by the denoising step and look at the oracle estimates, the proposed Gromov-Wasserstein alignment and projection, and the following convex weight optimization reproduce the post-treatment target network event better, despite the fact that there is no explicit reference space defined in this simulation design.

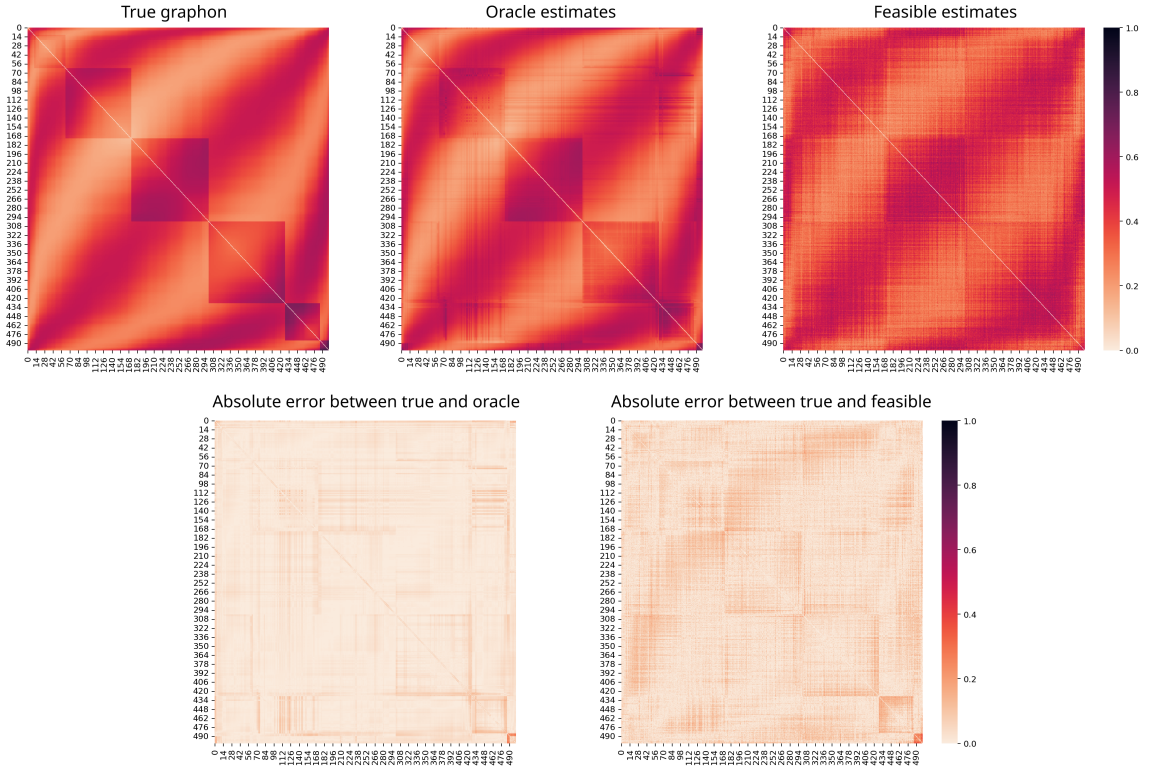


Figure 2: Comparison between true graphon, oracle, and feasible estimates for Design 2

Note: These plots report the first replication among 500 Monte Carlo replications for $N_g \in [500, 550]$ and $h = 0.05$.

Table 2 reports all seven evaluation measures for the largest networks, using both feasible and oracle inputs. Complete results for all six procedures, all three size settings, and every reported bandwidth, again, appear in Appendix D.

The proposed method continues to perform the best compared to other benchmarks. At this setting, SCN has smaller RMSE, pair MedAE, pair Q95, and expected-degree MedAE than all

Table 2: Simulation Results of Design 2 ($N_g \in [500, 550]$, $h = 0.05$)

Method	RMSE	MedAE	Q95	Δ MedAE	Direction	Degree	Triangle
Feasible procedures							
SCN	0.0634 (0.0002)	0.0422 (0.0001)	0.1239 (0.0003)	0.0561 (0.0004)	0.644 (0.0024)	0.0135 (0.0002)	0.0022 (0.0001)
Nearest GW source	0.0972 (0.0003)	0.0609 (0.0002)	0.1940 (0.0006)	0.0730 (0.0003)	0.630 (0.0024)	0.0172 (0.0001)	0.0017 (0.0001)
Best single source	0.0946 (0.0003)	0.0592 (0.0002)	0.1891 (0.0006)	0.0721 (0.0003)	0.618 (0.0023)	0.0176 (0.0001)	0.0024 (0.0001)
Uniform GW barycenter	0.0662 (0.0002)	0.0452 (0.0002)	0.1280 (0.0004)	0.0578 (0.0004)	0.622 (0.0027)	0.0160 (0.0002)	0.0030 (0.0001)
Degree alignment	0.1223 (0.0004)	0.1012 (0.0005)	0.2135 (0.0005)	0.1000 (0.0002)	0.680 (0.0086)	0.0148 (0.0001)	0.0032 (0.0001)
No change	0.0905 (0.0005)	0.0575 (0.0004)	0.1793 (0.0011)	0.0197 (0.0004)	0.000 (0.0000)	0.0196 (0.0003)	0.0072 (0.0002)
Oracle procedures							
SCN	0.0277 (0.0002)	0.0137 (0.0002)	0.0572 (0.0005)	0.0137 (0.0002)	0.897 (0.0023)	0.0058 (0.0001)	0.0012 (0.0000)
Nearest GW source	0.0380 (0.0004)	0.0157 (0.0003)	0.0809 (0.0009)	0.0157 (0.0003)	0.873 (0.0029)	0.0064 (0.0002)	0.0014 (0.0001)
Best single source	0.0367 (0.0003)	0.0149 (0.0002)	0.0778 (0.0009)	0.0149 (0.0002)	0.877 (0.0029)	0.0062 (0.0001)	0.0014 (0.0001)
Uniform GW barycenter	0.0393 (0.0004)	0.0264 (0.0003)	0.0757 (0.0007)	0.0264 (0.0003)	0.774 (0.0039)	0.0129 (0.0002)	0.0031 (0.0001)
Degree alignment	0.1159 (0.0005)	0.0919 (0.0006)	0.2088 (0.0007)	0.0919 (0.0006)	0.728 (0.0109)	0.0088 (0.0002)	0.0030 (0.0001)
No change	0.0280 (0.0006)	0.0197 (0.0004)	0.0518 (0.0010)	0.0197 (0.0004)	0.000 (0.0000)	0.0163 (0.0005)	0.0070 (0.0002)

Notes: Each entry is the mean of a replication-level statistic across 500 replications, with its Monte Carlo standard error in parentheses. RMSE, MedAE, and Q95 concern absolute errors in target post-treatment link probabilities. Δ MedAE concerns absolute errors in treatment-induced probability changes. Direction is the correctly predicted sign fraction among pairs with $|\Delta_{i,j}| \geq 0.02$. Degree is the median absolute error in normalized expected degrees. Triangle is the absolute error in normalized expected triangle density. Smaller values are better except for Direction.

five benchmarks. Its RMSE is 0.0634, compared with 0.0662 for uniform Gromov-Wasserstein, 0.0946 for the best single projected source, and 0.0972 for nearest Gromov-Wasserstein.

Recovering probability levels and recovering treatment changes give different rankings. At $N_g \in [500, 550]$ and $h = 0.05$, no change has smaller change MedAE than SCN, 0.0197 versus 0.0561, and degree alignment has higher change-direction accuracy, 68.0% versus 64.4%. No change predicts zero change, which limits its magnitude error when true changes are small but gives zero direction accuracy on the evaluated pairs. Increasing SCN's bandwidth to $h = 0.075$ improves change MedAE to 0.0272 and direction accuracy to 74.8%, and lowers expected-degree and triangle-density errors to 0.0124 and 0.00163, respectively, while increasing pair-probability RMSE to 0.0880. Network summaries also need separate interpretation where nearest Gromov-Wasserstein has the smallest triangle-density error at $h = 0.05$, despite its substantially larger pair-probability errors.

With oracle inputs, SCN's mean RMSE decreases from 0.0277 for $N_g \in [500, 550]$. It has smaller mean RMSE than all four source-based benchmarks at each size. Degree alignment remains near 0.116 across sizes, indicating that degree ranks provide insufficient information about the pair-specific structure in this mixture.

7 Empirical Illustration

In this section, we illustrate the proposed method using the informal risk-sharing network data of Banerjee et al. (2024) and conduct a leave-one-out placebo study. Each village's observed post-treatment network is withheld and imputed from its pre-treatment network and the pre- and post-treatment networks of other villages. Note that, unlike in the simulation, the true link probabilities are not observed. Hence, here we present the comparisons with a denoised validation network from prediction of realized links.

7.1 Data and Placebo Study Design

The data cover 75 villages in Karnataka, India, of which 43 were exposed to microfinance and 32 were not. We use the 43 treated villages for the placebo exercise. Their baseline household counts range from 114 to 356, with a mean of approximately 223 and a standard deviation of approximately 56. The different network sizes and the absence of a natural correspondence between households across villages make this a useful setting for the alignment step.

In this context, nodes represent households, and links are the binary household connections recorded in the replication data. We construct the pre- and post-treatment adjacency matrices from the link variables `A_ij_1` and `A_ij_2`, using household identifiers to retain a common node

ordering across the two periods within each village. The construction follows the household-indexing and symmetrization procedure associated with Auerbach and Cai (2025). Households in different villages are not matched by their identifiers; their alignment is instead estimated from the pre-treatment network structures via the proposed method.

In the leave-one-out placebo study each of the 43 treated villages takes turn as the target G and use the other 42 treated villages as sources. The target's observed post-treatment adjacency matrix is used only for evaluation. This exercise evaluates prediction under the same treatment exposure as the sources. It does not construct a no-microfinance outcome for a treated village and therefore does not estimate the causal effect of microfinance. In particular, the before-after changes examined below should not be interpreted as treatment effects, but rather the counterfactual imputation exercise.

We use the neighborhood smoother and Epanechnikov kernel described in Section 6, excluding both endpoints of the pair from its smoothing average. We then compare SCN with the same five benchmarks as in Section 6: the nearest Gromov-Wasserstein donor, the best single projected source, the uniform Gromov-Wasserstein barycenter, degree alignment, and no change.

Unlike in the simulation, the true post-treatment link probabilities are not observed. We therefore use the same seven evaluation measures as in Section 6, replacing the true post-treatment probability matrix by the denoised held-out post-treatment network, denoted by $\hat{P}_G^{(1),\mathrm{obs}}$. Specifically, we report pair-level RMSE, median absolute error, and the 95-th percentile of the absolute error; median absolute error[17] and direction accuracy for the denoised before-after changes; median absolute error in normalized expected degrees; and absolute error in normalized expected triangle density. Direction accuracy is evaluated among pairs with absolute denoised before-after change at least 0.02, matching the cutoff used in the simulation.

We first calculate each statistic within each target village and then take an unweighted average across the 43 villages. Thus, the reported medians and 95-th percentiles are averages of village-level statistics rather than statistics calculated after pooling household pairs. These comparisons measure agreement with an estimated probability matrix and should not be interpreted as recovery of the true link probabilities or as estimates of causal treatment effects.

[17]For the empirical change-magnitude comparison, both the predicted and observed changes use the same denoised pre-treatment network: $\hat{\Delta}_{i,j} = \hat{P}^{(1)}_{G,i,j} - \hat{P}^{(0),\mathrm{obs}}_{G,i,j}$ and $\Delta^{\mathrm{obs}}_{i,j} = \hat{P}^{(1),\mathrm{obs}}_{G,i,j} - \hat{P}^{(0),\mathrm{obs}}_{G,i,j}$. Hence, $\hat{\Delta}_{i,j} - \Delta^{\mathrm{obs}}_{i,j} = \hat{P}^{(1)}_{G,i,j} - \hat{P}^{(1),\mathrm{obs}}_{G,i,j}$ and the empirical Δ MedAE is identical to the pair-level MedAE.

7.2 Results

Table 3 first reports the pair-probability RMSE over the bandwidth grid. As in the simulation, the performance of the feasible procedures depends on the bandwidth used in the denoising step. At $h = 0.05$, SCN has a mean RMSE of 0.0534, similar to uniform Gromov-Wasserstein weighting, degree alignment, and no change. Its error falls sharply as the bandwidth increases and is relatively stable from $h = 0.15$ onward. The lowest mean SCN RMSE over the reported grid is 0.0238 at $h = 0.20$.

Table 3: Leave-one-out Placebo RMSE across Bandwidths

Method	$h = 0.05$	$h = 0.10$	$h = 0.15$	$h = 0.20$	$h = 0.30$	$h = 0.50$
SCN	0.0534	0.0276	0.0244	0.0238	0.0239	0.0240
Nearest GW source	0.0653	0.0343	0.0305	0.0300	0.0298	0.0300
Best single source	0.0617	0.0355	0.0316	0.0310	0.0311	0.0314
Uniform GW barycenter	0.0524	0.0287	0.0259	0.0255	0.0254	0.0255
Degree alignment	0.0533	0.0276	0.0244	0.0239	0.0238	0.0238
No change	0.0531	0.0284	0.0260	0.0257	0.0258	0.0259

Notes: Each entry is the unweighted mean of the village-level off-diagonal RMSEs across the 43 leave-one-out targets. Each target is imputed using the other 42 treated villages. For each bandwidth, both the input networks and the held-out validation network are denoised using the same bandwidth. Hence, comparisons across bandwidths constitute a sensitivity exercise rather than a bandwidth-selection rule.

The comparison across bandwidths differs from the simulation because the true post-treatment probability matrix is not observed. Changing h changes both the imputation and the denoised validation network. Therefore, the lower RMSE at $h = 0.20$ should not by itself be interpreted as evidence that this bandwidth recovers the unobserved true probability matrix most accurately. We use $h = 0.20$ below as a representative bandwidth from the relatively stable low-error region.

Table 4 reports the same seven evaluation measures as in the simulation. At $h = 0.20$, SCN has the smallest mean RMSE, 0.0238, compared with 0.0300 for the nearest Gromov-Wasserstein source, 0.0310 for the best single projected source, 0.0255 for uniform Gromov-Wasserstein weighting, 0.0239 for degree alignment, and 0.0257 for no change. SCN also has smaller upper-tail pair error than all five benchmarks. Its median pair error is smaller than those of the nearest source, best single source, uniform weighting, and no change, while degree alignment performs almost identically.

Table 4: Leave-one-out Placebo Results ($h = 0.20$)

Method	RMSE	MedAE	Q95	Direction	Degree	Triangle
SCN	0.0238	0.0139	0.0483	0.806	0.0102	1.155
Nearest GW source	0.0300	0.0178	0.0607	0.701	0.0137	1.486
Best single source	0.0310	0.0181	0.0635	0.725	0.0140	1.679
Uniform GW barycenter	0.0255	0.0163	0.0502	0.778	0.0140	1.588
Degree alignment	0.0239	0.0138	0.0485	0.807	0.0101	1.156
No change	0.0257	0.0160	0.0511	0.000	0.0135	1.416

Notes: Each entry is the unweighted mean of a village-level statistic across the 43 leave-one-out targets. RMSE, MedAE, and Q95 concern discrepancies between the imputed and denoised held-out post-treatment networks. Direction is the correctly predicted sign fraction among pairs whose absolute denoised before–after change is at least 0.02. Degree is the median absolute error in normalized expected degrees. Triangle is the absolute error in normalized expected triangle density, multiplied by 10^4. Smaller values are better except for Direction. Unlike in the simulation, no oracle procedures are available because the true probability matrices are unobserved.

The comparisons with the two single-source procedures provide the clearest distinction. SCN reduces mean RMSE by approximately 20.5% relative to the nearest Gromov-Wasserstein source and by 23.1% relative to the best single projected source. Uniform weighting performs considerably better than either single-source procedure but remains less accurate than SCN, with mean RMSE 0.0255. This comparison, which holds the Gromov-Wasserstein projections fixed and changes only how the projected source networks are combined, suggests that estimating the source weights can improve held-out reconstruction in this application.

Degree alignment is substantially more competitive. Its mean RMSE, pair MedAE, Q95, direction accuracy, normalized-degree error, and triangle-density error are all very close to those of SCN. SCN has slightly smaller RMSE, upper-tail pair error, and triangle-density error, while degree alignment has slightly smaller median pair error, slightly higher direction accuracy, and slightly smaller degree error. Thus, unlike in the simulation designs where the Gromov-Wasserstein and degree-alignment procedures can differ substantially, this empirical exercise does not provide strong evidence that Gromov-Wasserstein alignment improves prediction relative to degree alignment.

The before–after change measures lead to a similar conclusion. SCN correctly recovers the direction of 80.6% of the evaluated denoised changes, compared with 70.1% for the nearest source, 72.5% for the best single source, and 77.8% for uniform weighting. Degree alignment

has essentially the same direction accuracy at 80.7%. No change has zero direction accuracy by construction because it predicts a zero change for every pair.

The network-summary measures are also broadly consistent with the pair-level comparisons. SCN has normalized-degree MedAE 0.0102, below the nearest source, best single source, uniform weighting, and no change, and close to degree alignment at 0.0101. Its normalized triangle-density error is 1.155×10^{-4}, again below the corresponding errors for the four non-degree benchmarks and nearly identical to degree alignment. These results indicate that the pair-level reconstruction gains also translate into closer agreement in these node-level and triadic summaries of the denoised held-out networks.

Overall, the leave-one-out exercise provides evidence that combining multiple source networks can improve reconstruction of a held-out post-treatment network relative to relying on a single source, equal weighting, or assuming no change. The improvement over uniform weighting is more modest than the improvement over single-source procedures, and degree alignment performs almost identically to SCN in this application. The exercise remains descriptive: the validation object is itself estimated, the leave-one-out folds share source villages, and the comparison does not identify or validate an unobserved no-microfinance counterfactual.

References

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, 59(2):391–425.

Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American statistical Association*, 105(490):493–505.

Abadie, A. and Gardeazabal, J. (2003). The economic costs of conflict: A case study of the basque country. *American Economic Review*, 93(1):113–132.

Adjaho, C. and Christensen, T. (2022). Externally valid policy choice. *arXiv preprint arXiv:2205.05561*.

Aldous, D. J. (1981). Representations for partially exchangeable arrays of random variables. *Journal of Multivariate Analysis*, 11(4):581–598.

Ambrosio, L., Gigli, N., and Savaré, G. (2008). *Gradient flows: in metric spaces and in the space of probability measures*. Springer.

Auerbach, E. (2022). Identification and estimation of a partially linear regression model using network data. *Econometrica*, 90(1):347–365.

Auerbach, E. and Cai, Y. (2025). Identifying socially disruptive policies. Working paper, arXiv version 3, July 29, 2025; first circulated in 2023 under this title.

Banerjee, A., Breza, E., Chandrasekhar, A. G., Duflo, E., Jackson, M. O., and Kinnan, C. (2024). Changes in social network structure in response to exposure to formal credit markets. *Review of Economic Studies*, 91(3):1331–1372.

Bateni, B., Yuan, Y., Xu, Q., and Qu, A. (2025). Distributional treatment effect estimation across heterogeneous sites via optimal transport. *arXiv preprint arXiv:2511.09759*.

Cai, Y. (2026). Linear regression with centrality measures.

Chatterjee, S. (2015). Matrix estimation by universal singular value thresholding. *The annals of statistics*, pages 177–214.

Chowdhury, S. and Mémoli, F. (2019). The Gromov-Wasserstein distance between networks and stable network invariants. *Information and Inference: A Journal of the IMA*, 8(4):757–787.

Chowdhury, S. and Mémoli, F. (2023). Distances and isomorphism between networks: stability and convergence of network invariants. *Journal of Applied and Computational Topology*, 7(2):243–361.

Chowdhury, S. and Needham, T. (2020). Gromov-wasserstein averaging in a riemannian framework. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 842–843.

De Paula, Á. (2020). Econometric models of network formation. *Annual Review of Economics*, 12(1):775–799.

Deaner, B., Hsiang, C.-W., and Zeleneev, A. (2025). Simple specification-free causal inference in large panels. *Working paper*.

Demetci, P., Tran, Q. H., Redko, I., and Singh, R. (2024). Breaking isometric ties and introducing priors in gromov-wasserstein distances. In *International Conference on Artificial Intelligence and Statistics*, pages 298–306. PMLR.

Dumont, T., Lacombe, T., and Vialard, F.-X. (2025). On the existence of monge maps for the gromov–wasserstein problem. *Foundations of Computational Mathematics*, 25(2):463–510.

Feng, Y. (2020). Causal inference in possibly nonlinear factor models. *arXiv preprint arXiv:2008.13651*.

Fournier, N. and Guillin, A. (2015). On the rate of convergence in wasserstein distance of the empirical measure. *Probability theory and related fields*, 162(3):707–738.

Gao, C., Lu, Y., and Zhou, H. H. (2015). Rate-optimal graphon estimation. *The Annals of Statistics*, 43(6):2624–2652.

Graham, B. and De Paula, Á. (2020). *The econometric analysis of network data*. Academic Press.

Graham, B. S. (2017). An econometric model of network formation with degree heterogeneity. *Econometrica*, 85(4):1033–1063.

Gunsilius, F. F. (2023). Distributional synthetic controls. *Econometrica*, 91(3):1105–1117.

Gunsilius, F. F. (2025). A primer on optimal transport for causal inference with observational data. *arXiv preprint arXiv:2503.07811*.

Hoover, D. N. (1979). Relations on probability spaces and arrays of random variables. *Preprint, Institute for Advanced Study*, 15.

Kato, K. and Wang, B. (2025). Convergence of empirical gromov-wasserstein distance. *arXiv preprint arXiv:2508.03985*.

Kurisu, D., Zhou, Y., Otsu, T., and Müller, H.-G. (2025). Geodesic synthetic control methods for random objects and functional data. *arXiv preprint arXiv:2505.00331*.

Li, T., Levina, E., and Zhu, J. (2020). Network cross-validation by edge sampling. *Biometrika*, 107(2):257–276.

Lin, C. (2025). Functional synthetic control. *Working Paper*.

Lin, Z., Kong, D., and Wang, L. (2023). Causal inference on distribution functions. *Journal of the Royal Statistical Society: Series B*, 85(2):378–398.

Lovász, L. (2012). *Large networks and graph limits*, volume 60. American Mathematical Soc.

Martín, R. D., Medri, I. V., and Murphy, J. M. (2026). Gromov-wasserstein barycenters: The analysis problem. *arXiv preprint arXiv:2507.09865*.

Mémoli, F. (2011). Gromov–wasserstein distances and the metric approach to object matching. *Foundations of computational mathematics*, 11(4):417–487.

Mémoli, F. (2014). The gromov–wasserstein distance: A brief overview. *Axioms*, 3(3):335–341.

Menzel, K. (2023). Transfer estimates for causal effects across heterogeneous sites. *arXiv preprint arXiv:2305.01435*.

Nowakowicz, K. (2024). Nonparametric network bootstrap. *arXiv preprint arXiv:2401.12345*.

Okano, R. and Kurisu, D. (2026). Functional synthetic control methods for metric space-valued outcomes. *arXiv preprint arXiv:2601.07539*.

Peyré, G., Cuturi, M., and Solomon, J. (2016). Gromov-wasserstein averaging of kernel and distance matrices. In *International conference on machine learning*, pages 2664–2672. PMLR.

Sturm, K.-T. (2023). *The space of spaces: curvature bounds and gradient flows on the space of metric measure spaces*, volume 290. American Mathematical Society.

Van Dijcke, D. (2025). Regression discontinuity design with distribution-valued outcomes. *arXiv preprint arXiv:2504.03992*.

Xu, H., Luo, D., Carin, L., and Zha, H. (2021). Learning graphons via structured Gromov–Wasserstein barycenters. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 10505–10513.

Zeleneev, A. (2020). Identification and estimation of network models with nonparametric unobserved heterogeneity. *arXiv preprint arXiv:2602.06885*.

Zhang, Y., Levina, E., and Zhu, J. (2017). Estimating network edge probabilities by neighbourhood smoothing. *Biometrika*, 104(4):771–783.

Zhang, Z., Goldfeld, Z., Mroueh, Y., and Sriperumbudur, B. K. (2024). Gromov–wasserstein distances: Entropic regularization, duality and sample complexity. *The Annals of Statistics*, 52(4):1616–1645.

Appendix

A Proofs

A.1 Proofs of the results of Section 4

A.1.1 Proof of Theorem 1

To prove the optimality of the alignment coupling, we first rewrite the Gromov-Wasserstein optimization in the reference space representation with the following lemma.

Lemma A.1 demonstrates that the internal distances d_g^2 and d_G^2 pull back to $\check{d}_g^2$ and $\check{d}_G^2$ through the alignment maps. With this, the Gromov-Wasserstein problem is equivalent to a self-coupling problem on the reference space, in which the alignment coupling corresponds to the identity coupling.

Lemma A.1 (Reference representation of Gromov-Wasserstein optimization problem)
The internal distances d_g^2 and d_G^2 are bounded on their respective supports, and their reference counterparts $\check{d}_g^2$ and $\check{d}_G^2$ are bounded by Lemma 1 (ii). Thus, all objectives below are finite, and the following statements hold:

(*i*) *For every $g \in \{1, ..., G-1, G\}$ and all $(\check{\xi}, \check{\xi}') \in \check{\Xi} \times \check{\Xi}$,*

$$d_g^2\left(\rho_g(\check{\xi}), \rho_g(\check{\xi}')\right) = \check{d}_g^2(\check{\xi}, \check{\xi}').$$

(*ii*) *Fix $g \in \{1, ..., G-1\}$. For every $\check{\pi} \in \Pi(\check{\mu}, \check{\mu})$, the pushforward $\left(\rho_g \times \rho_G\right)_\# \check{\pi}$ belongs to the set of valid couplings $\Pi(\mu_g, \mu_G)$, and it satisfies $F_{d_g^2, d_G^2}\left(\left(\rho_g \times \rho_G\right)_\# \check{\pi}\right) = F_{\check{d}_g^2, \check{d}_G^2}(\check{\pi})$. Conversely, every coupling $\pi \in \Pi(\mu_g, \mu_G)$ is such a pushforward. Hence,*

$$\inf_{\pi \in \Pi(\mu_g, \mu_G)} F_{d_g^2, d_G^2}(\pi) = \inf_{\check{\pi} \in \Pi(\check{\mu}, \check{\mu})} F_{\check{d}_g^2, \check{d}_G^2}(\check{\pi}). \tag{20}$$

(*iii*) *Denote $\check{\pi}^{\mathrm{diag}} := (\mathrm{id}, \mathrm{id})_\# \check{\mu} \in \Pi(\check{\mu}, \check{\mu})$ as the identity coupling, which is supported on the diagonal of $\check{\Xi} \times \check{\Xi}$. Then with the alignment coupling $\pi_\rho^{g,G} = \left(\rho_g \times \rho_G\right)_\# \check{\pi}^{\mathrm{diag}}$, it attains the left infimum in (20) if and only if $\check{\pi}^{\mathrm{diag}}$ attains the right infimum.*

Proof.

Proof of statement (i). Fix g and $(\check{\xi}, \check{\xi}') \in \check{\Xi} \times \check{\Xi}$, for notation simplicity, we define

$$h(\zeta) := \left[f^{(0)}\left(\rho_g(\check{\xi}), \zeta, \gamma_g\right) - f^{(0)}\left(\rho_g(\check{\xi}'), \zeta, \gamma_g\right)\right]^2 \qquad \text{for } \zeta \in \mathrm{supp}(\mu_g),$$

where we extend h by zero outside $\mathrm{supp}(\mu_g)$. By Lemma 1 (ii) and the maintained model described in Section 3, h is bounded and Borel. And since $\mu_g\left(\mathrm{supp}(\mu_g)\right) = 1$, the integral below is well defined. By the definition of d_g^2 in the population algorithm 2 and the change of variables for pushforward measures, which states that $\int_\Xi h \mathrm{d}\left((\rho_g)_\# \check{\mu}\right) = \int_{\check{\Xi}} \left(h \circ \rho_g\right) \mathrm{d}\check{\mu}$ for bounded Borel h, it follows that

$$d_g^2\left(\rho_g(\check{\xi}), \rho_g(\check{\xi}')\right) = \int_\Xi h \mathrm{d}\mu_g = \int_{\check{\Xi}} h\left(\rho_g(\zeta)\right) \mathrm{d}\check{\mu}(\zeta) = \int_{\check{\Xi}} \left[\check{f}_g^{(0)}(\check{\xi}, \zeta) - \check{f}_g^{(0)}(\check{\xi}', \zeta)\right]^2 \mathrm{d}\check{\mu}(\zeta) = \check{d}_g^2(\check{\xi}, \check{\xi}'),$$

where the second equality uses $\left(\rho_g\right)_\# \check{\mu} = \mu_g$ from Lemma 1 (i), and the third and fourth follow from Definition 3.

Proof of statement (ii). We start by showing that the pushforward $\left(\rho_g \times \rho_G\right)_\# \check{\pi}$ is a valid coupling. For Borel sets $A, B \subseteq \Xi$, by Lemma 1 (i), we have

$$\left(\rho_g \times \rho_G\right)_\# \check{\pi}(A \times \Xi) = \check{\pi}\left(\rho_g^{-1}(A) \times \check{\Xi}\right) = \check{\mu}\left(\rho_g^{-1}(A)\right) = \mu_g(A),$$
$$\left(\rho_g \times \rho_G\right)_\# \check{\pi}(\Xi \times B) = \check{\pi}\left(\check{\Xi} \times \rho_G^{-1}(B)\right) = \check{\mu}\left(\rho_G^{-1}(B)\right) = \mu_G(B).$$

Thus, $\left(\rho_g \times \rho_G\right)_\# \check{\pi} \in \Pi(\mu_g, \mu_G)$.

We then show the objectives transfers via alignment maps. Notice that two independent draws from the pushforward coupling $\left(\rho_g \times \rho_G\right)_\# \check{\pi}$ have the same joint distribution as two independent draws from $\check{\pi}$ pushed forward coordinate-wise, that is,

$$\left(\left(\rho_g \times \rho_G\right)_\# \check{\pi}\right) \otimes \left(\left(\rho_g \times \rho_G\right)_\# \check{\pi}\right) = \left(\left(\rho_g \times \rho_G\right) \times \left(\rho_g \times \rho_G\right)\right)_\# (\check{\pi} \otimes \check{\pi}).$$

Hence, by the change of variables for pushforward measures and then statement (i) applied to both networks,

$$F_{d_g^2, d_G^2}\left(\left(\rho_g \times \rho_G\right)_\# \check{\pi}\right) = \iint_{\check{\Xi}\times\check{\Xi}} \iint_{\check{\Xi}\times\check{\Xi}} \left[d_g^2\left(\rho_g(\check{\xi}), \rho_g(\check{\xi}')\right) - d_G^2\left(\rho_G(\check{\zeta}), \rho_G(\check{\zeta}')\right)\right]^2 \mathrm{d}\check{\pi}(\check{\xi}, \check{\zeta}) \mathrm{d}\check{\pi}(\check{\xi}', \check{\zeta}') = F_{\check{d}_g^2, \check{d}_G^2}(\check{\pi}).$$

Next, we prove the converse. Given a pair $(\xi, \zeta) \sim \pi$, let $\check{\xi}$ and $\check{\zeta}$ be conditionally independent with conditional distributions being $\check{\mu}$ given $\rho_g = \xi$ and given $\rho_G = \zeta$, respectively. Both of these are provided by the disintegration theorem, and write $\check{\pi}$ for the joint law of $(\check{\xi}, \check{\zeta})$. Since the first marginal of π is μ_g and averaging the conditional distributions over μ_g reconstructs $\check{\mu}$, and symmetrically for the target, both marginals of $\check{\pi}$ are $\check{\mu}$. Moreover, since the conditional law of $\check{\mu}$ given $\rho_g = \xi$ is supported on the reference coordinates that ρ_g sends to ξ, and symmetrically for ρ_G and ζ, it follows that $\rho_g(\check{\xi}) = \xi$ and $\rho_G(\check{\zeta}) = \zeta$ almost surely. Hence, $\check{\pi} \in \Pi(\check{\mu}, \check{\mu})$ and $\left(\rho_g \times \rho_G\right)_\# \check{\pi} = \pi$.

Finally, we prove (20). Since every $\check{\pi} \in \Pi(\check{\mu}, \check{\mu})$ produces the valid $\left(\rho_g \times \rho_G\right)_\# \check{\pi}$ with the same objective value, the left infimum is at most the right one. And on the other hand, since every $\pi \in \Pi(\mu_g, \mu_G)$ is such a pushforward by the converse, again with the same objective value, the right infimum is at most the left one.

Proof of statement (iii). Since $\left(\rho_g, \rho_G\right) = \left(\rho_g \times \rho_G\right) \circ (\mathrm{id}, \mathrm{id})$, and pushing a measure forward through a composition of two maps is the same as pushing it forward through the first map and then through the second, Definition 4 gives $\pi_\rho^{g,G} = \left(\rho_g \times \rho_G\right)_\# \check{\pi}^{\mathrm{diag}}$. By statement (ii), $F_{d_g^2, d_G^2}\left(\pi_\rho^{g,G}\right) = F_{\check{d}_g^2, \check{d}_G^2}\left(\check{\pi}^{\mathrm{diag}}\right)$, and the two infima in (20) are equal. Therefore, $\pi_\rho^{g,G}$ attains the left infimum if and only if its objective value equals this value, which holds if and only if $\check{\pi}^{\mathrm{diag}}$ attains the right infimum.

□

Note that Lemma A.1 (ii) is the reason the assumptions of Section 3 were stated in reference representation. This is because the matching step sees the two networks only through $\check{d}_g^2$ and $\check{d}_G^2$, and Assumption 3 is a statement about exactly these two objects.

With Lemma A.1 on hand, we then proceed to prove Theorem 1.

Proof.

Proof of statement (i). By Lemma A.1 (ii), the Gromov-Wasserstein optimization problem has the same infimum as the reference problem $\inf_{\check{\pi} \in \Pi(\check{\mu}, \check{\mu})} F_{\check{d}_g^2, \check{d}_G^2}(\check{\pi})$. Moreover, by Lemma A.1 (iii), $\pi_\rho^{g,G}$ attains the former infimum if and only if $\check{\pi}^{\mathrm{diag}}$ attains the latter. Therefore, it suffices to show that under Assumption 3, $\check{\pi}^{\mathrm{diag}}$ minimizes $F_{\check{d}_g^2, \check{d}_G^2}$ over $\Pi(\check{\mu}, \check{\mu})$ with minimum value $\left\|\check{d}_g^2 - \check{d}_G^2\right\|_{L^2(\check{\mu} \otimes \check{\mu})}^2$.

Under $\check{\pi} \otimes \check{\pi}$, the pair $(\check{\xi}, \check{\xi}') \sim \check{\mu} \otimes \check{\mu}$ and $(\check{\zeta}, \check{\zeta}') \sim \check{\mu} \otimes \check{\mu}$. In particular,

$$\mathbb{P}\left(\check{d}_g^2(\check{\xi}, \check{\xi}') > u\right) = (\check{\mu} \otimes \check{\mu})\left(\check{d}_g^2 > u\right),$$
$$\mathbb{P}\left(\check{d}_G^2(\check{\zeta}, \check{\zeta}') > v\right) = (\check{\mu} \otimes \check{\mu})\left(\check{d}_G^2 > v\right),$$

for every $\check{\pi} \in \Pi(\check{\mu}, \check{\mu})$. That is, the marginal distributions of $\check{d}_g^2(\check{\xi}, \check{\xi}')$ and $\check{d}_G^2(\check{\zeta}, \check{\zeta}')$ do not depend on the coupling. Expanding the squared term in $F_{\check{d}_g^2, \check{d}_G^2}(\check{\pi})$ and integrating the three bounded terms separately by Fubini's theorem gives the square loss decomposition,

$$F_{\check{d}_g^2, \check{d}_G^2}(\check{\pi}) = \left\|\check{d}_g^2\right\|_{L^2(\check{\mu} \otimes \check{\mu})}^2 + \left\|\check{d}_G^2\right\|_{L^2(\check{\mu} \otimes \check{\mu})}^2 - 2\mathbb{E}\left\{\left[\check{d}_g^2(\check{\xi}, \check{\xi}')\right]\left[\check{d}_G^2(\check{\zeta}, \check{\zeta}')\right]\right\}.$$

Similar decomposition also appears in the optimal transport literature, e.g., Proposition 1 of Peyré et al. (2016) for the square loss, where they separates the objective into two terms that only depend on the marginals and one cross product term that depends on the coupling. Therefore, minimizing $F_{\check{d}_g^2, \check{d}_G^2}$ is equivalent to maximizing the cross product $\mathbb{E}\left\{\left[\check{d}_g^2(\check{\xi}, \check{\xi}')\right]\left[\check{d}_G^2(\check{\zeta}, \check{\zeta}')\right]\right\}$, as the first two norms are coupling-free.

For notational simplicity, denote $U := \check{d}_g^2(\check{\xi}, \check{\xi}')$ and $V := \check{d}_G^2(\check{\zeta}, \check{\zeta}')$. With the nonnegative integrand, it follows that

$$\begin{aligned} UV &= \int_0^\infty \int_0^\infty \mathbf{1}\{U > u\}\mathbf{1}\{V > v\}\mathrm{d}u\mathrm{d}v, \\ \mathbb{E}\left\{\left[\check{d}_g^2(\check{\xi}, \check{\xi}')\right]\left[\check{d}_G^2(\check{\zeta}, \check{\zeta}')\right]\right\} = \mathbb{E}(UV) &= \int_0^\infty \int_0^\infty \mathbb{P}\left(U > u, V > v\right) \mathrm{d}u\mathrm{d}v \\ &\leq \int_0^\infty \int_0^\infty \min\left\{\mathbb{P}(U > u), \mathbb{P}(V > v)\right\} \mathrm{d}u\mathrm{d}v, \end{aligned}$$

where the bound is finite and coupling-free. Note that the integrand vanishes for $u > \sup \check{d}_g^2$ or $v > \sup \check{d}_G^2$, it's at most one, and it only depends on the two marginal distribution and not on the coupling.

We then record a fact about comonotonicity. If the pair (U, V) is comonotone in the sense of statement (ii), then

$$\mathbb{P}(U > u, V > v) = \min\left\{\mathbb{P}(U > u), \mathbb{P}(V > v)\right\} \qquad \text{for every } (u, v).$$

To see this, fix (u, v) and consider the two events $\{U > u, V \leq v\}$ and $\{U \leq u, V > v\}$. They cannot both have positive probability. Indeed, if they did, then the product of the two events would have positive product measure, and at any pair (ω, ω') in that product we would have $U(\omega) > u \geq U(\omega')$ and $V(\omega) \leq v < V(\omega')$. Hence the resulting $[U(\omega) - U(\omega')]\,[V(\omega) - V(\omega')] < 0$‘ contradicts the comonotonicity. Thus, at least one of the two events is null. Suppose $\mathbb{P}(U > u, V \leq v) = 0$. Then $\mathbb{P}(U > u, V > v) = \mathbb{P}(U > u)$. Since the joint probability is also at most $\mathbb{P}(V > v)$, this gives $\mathbb{P}(U > u) \leq \mathbb{P}(V > v)$. Hence, the minimum is $\mathbb{P}(U > u)$ and the equality holds. If instead $\mathbb{P}(U \leq u, V > v) = 0$, the same argument with the reference coordinates of U and V exchanged leads to $\mathbb{P}(U > u, V > v) = \mathbb{P}(V > v)$, which again is the minimum.

Now take $\check{\pi} = \check{\pi}^{\mathrm{diag}}$, and it follows that $\check{\zeta} = \check{\xi}$ and $\check{\zeta}' = \check{\xi}'$ almost surely. With this, $(U, V) = \left(\check{d}_g^2(\check{\xi}, \check{\xi}'), \check{d}_G^2(\check{\xi}, \check{\xi}')\right)$ with $(\check{\xi}, \check{\xi}') \sim \check{\mu} \otimes \check{\mu}$. And by Assumption 3, this pair is comonotone. Combining the fact about comonotonicity that we previously stated, we have

$$\mathbb{E}_{\check{\pi}^{\mathrm{diag}} \otimes \check{\pi}^{\mathrm{diag}}}\left\{\left[\check{d}_g^2(\check{\xi}, \check{\xi}')\right]\left[\check{d}_G^2(\check{\zeta}, \check{\zeta}')\right]\right\} = \int_0^\infty \int_0^\infty \min\left\{\mathbb{P}(U > u), \mathbb{P}(V > v)\right\} \mathrm{d}u\mathrm{d}v,$$

and for every $\check{\pi} \in \Pi(\check{\mu}, \check{\mu})$,

$$F_{\check{d}_g^2, \check{d}_G^2}(\check{\pi}) \geq \left\|\check{d}_g^2\right\|_{L^2(\check{\mu}\otimes\check{\mu})}^2 + \left\|\check{d}_G^2\right\|_{L^2(\check{\mu}\otimes\check{\mu})}^2 - 2\iint_{\check{\Xi}\times\check{\Xi}} \check{d}_g^2 \check{d}_G^2 \,\mathrm{d}(\check{\mu} \otimes \check{\mu}) = \left\|\check{d}_g^2 - \check{d}_G^2\right\|_{L^2(\check{\mu}\otimes\check{\mu})}^2,$$

with equality at $\check{\pi}^{\mathrm{diag}}$.

Proof of statement (ii). Let $\pi \in \Pi(\mu_g, \mu_G)$ be optimal for the Gromov-Wasserstein optimization. By Lemma A.1 (ii), there exists $\check{\pi} \in \Pi(\check{\mu}, \check{\mu})$ with $\left(\rho_g \times \rho_G\right)_\# \check{\pi} = \pi$ and $F_{\check{d}_g^2, \check{d}_G^2}(\check{\pi}) = F_{d_g^2, d_G^2}(\pi)$, and

the two optimizations share the same infimum. Hence, $\check{\pi}$ is optimal for the reference problem, and the first claim is proved.

For the second claim, let $\check{\pi} \in \Pi(\check{\mu}, \check{\mu})$ be an arbitrary optimal coupling of the reference problem, and denote $U := \check{d}_g^2(\check{\xi}, \check{\xi}')$ and $V := \check{d}_G^2(\check{\zeta}, \check{\zeta}')$ just as in the proof of statement (i). We want to show that the pair (U, V) is comonotone.

Since the upper bound on $\mathbb{E}(UV)$ is attained in statement (i), optimality of $\check{\pi}$ forces $\mathbb{E}(UV)$ to attain the same bound. That is,

$$\mathbb{E}(UV) = \int_0^\infty \int_0^\infty \mathbb{P}\left(U > u, V > v\right) \mathrm{d}u\mathrm{d}v = \int_0^\infty \int_0^\infty \min\left\{\mathbb{P}(U > u), \mathbb{P}(V > v)\right\} \mathrm{d}u\mathrm{d}v$$

Since $\mathbb{P}(U > u, V > v) \leq \min\{\mathbb{P}(U > u), \mathbb{P}(V > v)\}$ for every (u, v), the left-hand side is never larger than the right-hand side. Two integrands with one dominating the other can only have the same finite integral if they agree except on a set of zero area, and hence, these two are equal for all $(u, v) \in (0, \infty)^2$ outside a set of zero area.

This equality extends to every $(u, v) \in (0, \infty)^2$ by continuity from the right. Indeed, as $u_n \downarrow u$ and $v_n \downarrow v$, the events $\{U > u_n, V > v_n\}$, $\{U > u_n\}$, and $\{V > v_n\}$ increase to their respective limiting events, and hence their probabilities converge. Given any (u, v), since the set where they differ has zero area and cannot contain every point of a neighborhood of (u, v) from the upper right, we can then pick a sequence $(u_n, v_n) \downarrow (u, v)$ along which the two sides are equal. Passing to the limit, equality then holds at (u, v) itself. Hence, for every (u, v), since $\mathbb{P}(U > u, V > v) = \min\{\mathbb{P}(U > u), \mathbb{P}(V > v)\}$, it follows that

$$\mathbb{P}\left(U > u, V \leq v\right) \cdot \mathbb{P}\left(U \leq u, V > v\right) = \left[\mathbb{P}(U > u) - \mathbb{P}\left(U > u, V > v\right)\right] \left[\mathbb{P}(V > v) - \mathbb{P}\left(U > u, V > v\right)\right] = 0.$$

Therefore, at least one of the events $\{U > u, V \leq v\}$ and $\{U \leq u, V > v\}$ is null.

Finally, we are now ready to show that the pair (U, V) is comonotone. Let (U, V) and (U', V') be independent pairs with the same joint distribution, and suppose that $U > U'$ and $V < V'$. Then there exist positive rational numbers u, v with $U > u \geq U'$ and $V \leq v < V'$. With this, the first pair lies in the event $\{U > u, V \leq v\}$ and the second pair lies in the event $\{U' \leq u, V' > v\}$. By the previous discussion, one of these two events is null, and hence the product event has probability zero for each such (u, v). Similarly, the case $U < U'$ and $V > V'$ can be handled by swapping the two. Since there are only countably many pairs of rational numbers, the event $(U - U')(V - V') < 0$ is contained in a countable union of null events, and hence it is also null. That is, $(U - U')(V - V') \geq 0$ almost surely, which is the comonotonicity of $\left(\check{d}_g^2(\check{\xi}, \check{\xi}'), \check{d}_G^2(\check{\zeta}, \check{\zeta}')\right)$.

□

A.1.2 Proof of Proposition 1

To study the properties of Gromov-Wasserstein projections through the alignment coupling, we first discuss the properties of $\bar{f}_g^{(w)}$ in the following Lemma A.2.

Lemma A.2 (Properties of $\bar{f}_g^{(w)}$)
Let $\{\breve{\mu}_\zeta^G\}_{\zeta\in\Xi}$ be the disintegration of $\breve{\mu}$ with respect to ρ_G fixed in Section 4.2, and let $\bar{f}_g^{(w)}$ be the average defined in (5), for $g \in \{1, ..., G-1, G\}$ and $w \in \{0, 1\}$. The following statements hold:

(*i*) *$\bar{f}_g^{(w)}$ is Borel on $\Xi \times \Xi$ with values in $[0, 1]$. $\bar{f}_g^{(w)}$ also does not depend on the choice of the disintegration up to $\mu_G \otimes \mu_G$-null sets.*

(*ii*) *The average is linear. For every $a \in \mathbb{R}^G$ and every pair (ζ, ζ'),*

$$\sum_{g=1}^{G} a_g \bar{f}_g^{(w)}(\zeta, \zeta') = \iint_{\breve{\Xi}\times\breve{\Xi}} \left[\sum_{g=1}^{G} a_g \breve{f}_g^{(w)}(\breve{\xi}, \breve{\xi}')\right] \mathrm{d}\breve{\mu}_\zeta^G(\breve{\xi})\mathrm{d}\breve{\mu}_{\zeta'}^G(\breve{\xi}').$$

Moreover, if $\sum_{g=1}^G a_g \breve{f}_g^{(w)} = 0$ $\breve{\mu}\otimes\breve{\mu}$-a.e., then $\sum_{g=1}^G a_g \bar{f}_g^{(w)} = 0$ $\mu_G \otimes \mu_G$-a.e.

(*iii*) *For every $a \in \mathbb{R}^G$,*

$$\left\|\sum_{g=1}^{G} a_g \bar{f}_g^{(w)}\right\|_{L^2(\mu_G\otimes\mu_G)} \leq \left\|\sum_{g=1}^{G} a_g \breve{f}_g^{(w)}\right\|_{L^2(\breve{\mu}\otimes\breve{\mu})}.$$

(*iv*) *If $\breve{f}_g^{(w)} = H \circ (\rho_G \times \rho_G)$ $\breve{\mu}\otimes\breve{\mu}$-a.e. for a Borel function H on $\Xi \times \Xi$, where $\breve{f}_g^{(w)}$ is bounded on $\mathrm{supp}(\mu_G) \times \mathrm{supp}(\mu_G)$, then $\bar{f}_g^{(w)} = H$ in $L^2(\mu_G \otimes \mu_G)$.*

Proof.

For notation simplicity, denote the bounded Borel function h on $\breve{\Xi}\times\breve{\Xi}$ and its average

$$\bar{h}(\zeta, \zeta') := \iint_{\breve{\Xi}\times\breve{\Xi}} h(\breve{\xi}, \breve{\xi}')\mathrm{d}\breve{\mu}_\zeta^G(\breve{\xi})\mathrm{d}\breve{\mu}_{\zeta'}^G(\breve{\xi}').$$

Hence, $h = \breve{f}_g^{(w)}$ and $\bar{h} = \bar{f}_g^{(w)}$. We first lay down some useful properties of $\bar{h}$.

Property (a): $\bar{h}$ is a bounded Borel function on $\Xi\times\Xi$. Since $\breve{\mu}_\zeta^G \otimes \breve{\mu}_{\zeta'}^G$ is a probability measure, the boundedness holds. For measurability, since $\zeta \to \breve{\mu}_\zeta^G$ is a measurable family of probability measures by the disintegration theorem, it follows that $(\zeta, \zeta') \to \breve{\mu}_\zeta^G \otimes \breve{\mu}_{\zeta'}^G$ is also measurable. Integrating a bounded Borel function over a measurable family of probability measures will result a Borel function of the parameter.

Property (b): integrating the average recovers the original integral. For every bounded Borel h,

$$\iint_{\Xi\times\Xi} \bar{h}\,\mathrm{d}(\mu_G\otimes\mu_G) = \iint_{\check{\Xi}\times\check{\Xi}} h\,\mathrm{d}(\check{\mu}\otimes\check{\mu})\,. \tag{21}$$

Here, we utilize the property $\check{\mu} = \int_{\Xi} \check{\mu}_\zeta^G \mathrm{d}\mu_G(\zeta)$ from the disintegration.

Property (c): the average ignores null sets. If $h = 0$ $\check{\mu}\otimes\check{\mu}$-a.e., then $\bar{h} = 0$ $\mu_G\otimes\mu_G$-a.e. Indeed, by the triangle inequality, $|\bar{h}| \le \overline{|h|}$ at every pair inputs. Then by (21), we have $\iint \overline{|h|}\,\mathrm{d}(\mu_G\otimes\mu_G) = \iint |h|\,\mathrm{d}(\check{\mu}\otimes\check{\mu}) = 0$. Hence, the nonnegative function $\overline{|h|}$ vanishes $\mu_G\otimes\mu_G$-a.e.

Proof of statement (i). Since $\check{f}_g^{(w)}$ lies in $[0,1]$, $\bar{f}_g^{(w)}$ also lies in $[0,1]$, and its Borel by abovementioned property (a). Let $\left\{\check{\mu}_\zeta^G\right\}$ and $\left\{\left(\check{\mu}_\zeta^G\right)'\right\}$ be two disintegrations, and denote $\varnothing$ as the μ_G-null set. By the uniqueness statement of the disintegration theorem, $\check{\mu}_\zeta^G = \left(\check{\mu}_\zeta^G\right)'$ for every ζ outside $\varnothing$. Then two $\bar{f}_g^{(w)}$ with different disintegrations are equal at every pair (ζ,ζ') with $\zeta\notin\varnothing$ and $\zeta'\notin\varnothing$, and the excluded pairs form a $\mu_G\otimes\mu_G$-null set.

Proof of statement (ii). To prove the equality relation, we first fix a pair (ζ,ζ'). Since each $\bar{f}_g^{(w)}(\zeta,\zeta')$ is an integral against the same probability measure $\check{\mu}_\zeta^G\otimes\check{\mu}_{\zeta'}^G$, and the integral of a linear combination is the linear combination of the integrals, this completes the proof. For the second claim, let $h = \sum_{g=1}^G a_g \check{f}_g^{(w)}$ and $\bar{h} = \sum_{g=1}^G a_g \bar{f}_g^{(w)}$. If $h = 0$ $\check{\mu}\otimes\check{\mu}$-a.e., then $\bar{h} = 0$ $\mu_G\otimes\mu_G$-a.e. by property (c), and this completes the proof.

Proof of statement (iii). Let $h = \sum_{g=1}^G a_g \check{f}_g^{(w)}$. At every pair (ζ,ζ'), by the Cauchy-Schwarz inequality, it follows that $\bar{h}(\zeta,\zeta')^2 \le \overline{h^2}(\zeta,\zeta')$. Integrating over $\mu_G\otimes\mu_G$ and applying property (b) to h^2, we have

$$\left\|\sum_{g=1}^G a_g \bar{f}_g^{(w)}\right\|_{L^2(\mu_G\otimes\mu_G)}^2 = \iint \bar{h}^2\,\mathrm{d}(\mu_G\otimes\mu_G) \le \iint h^2\,\mathrm{d}(\check{\mu}\otimes\check{\mu}) = \left\|\sum_{g=1}^G a_g \check{f}_g^{(w)}\right\|_{L^2(\check{\mu}\otimes\check{\mu})}^2 .$$

Proof of statement (iv). Since $\check{f}_g^{(w)}$ and $H\circ(\rho_G\times\rho_G)$ only differ on a $\check{\mu}\otimes\check{\mu}$-null set, property (c) implies that their averages differ only on a $\mu_G\otimes\mu_G$-null set. Hence, it suffices to show that

$$\iint_{\check{\Xi}\times\check{\Xi}} H(\rho_G(\check{\xi}),\rho_G(\check{\xi}'))\mathrm{d}\check{\mu}_\zeta^G(\check{\xi})\mathrm{d}\check{\mu}_{\zeta'}^G(\check{\xi}') = H(\zeta,\zeta')$$

for $\mu_G\otimes\mu_G$-a.e. (ζ,ζ'). We fix such a pair. By the disintegration theorem, $\check{\mu}_\zeta^G$ gives all its mass to reference coordinates $\check{\xi}$ with $\rho_G(\check{\xi}) = \zeta$, and $\check{\mu}_{\zeta'}^G$ gives all its mass to reference coordinates $\check{\xi}'$ with $\rho_G(\check{\xi}') = \zeta'$. For every such $(\check{\xi},\check{\xi}')$, the integrand takes the value $H(\zeta,\zeta')$, which does not depend on $(\check{\xi},\check{\xi}')$. Hence, integral equals $H(\zeta,\zeta')$, and therefore, $\bar{f}_g^{(w)} = H$ $\mu_G\otimes\mu_G$-a.e.

□

With the properties of $\bar{f}_g^{(w)}$, we then proceed to prove Proposition 1.

Proof.

Proof of statement (i). The proof consists two steps. We first identify the conditional distributions of the alignment coupling given the target position, which $\tilde{Q}_g^{(w)}$ in Definition 1 needs. We then substitute them into the definition.

For the first step, the alignment coupling is the joint distribution of $\left(\rho_g(\check{\xi}), \rho_G(\check{\xi})\right)$ for $\check{\xi} \sim \check{\mu}$. Conditional on $\rho_G(\check{\xi}) = \zeta$, the reference coordinate $\check{\xi}$ is distributed as $\check{\mu}_\zeta^G$. Hence, the source individual $\rho_g(\check{\xi})$ is distributed as $\pi_\zeta := \left(\rho_g\right)_\# \check{\mu}_\zeta^G$. To verify that $\left\{\pi_\zeta\right\}$ is indeed a disintegration of $\pi_\rho^{g,G}$ with respect to μ_G, note that $\zeta \to \pi_\zeta(B) = \check{\mu}_\zeta^G\left(\rho_g^{-1}(B)\right)$ is Borel for every Borel B. And for every bounded Borel h on $\Xi \times \Xi$, it follows that

$$\begin{aligned}\iint_{\Xi\times\Xi} h(\xi,\zeta)\mathrm{d}\pi_\rho^{g,G}(\xi,\zeta) &= \int_{\check{\Xi}} h\left(\rho_g(\check{\xi}), \rho_G(\check{\xi})\right)\mathrm{d}\check{\mu}(\check{\xi}) \\ &= \int_\Xi \left[\int_{\check{\Xi}} h\left(\rho_g(\check{\xi}),\zeta\right)\mathrm{d}\check{\mu}_\zeta^G(\check{\xi})\right]\mathrm{d}\mu_G(\zeta) = \int_\Xi \left[\int_\Xi h(\xi,\zeta)\mathrm{d}\pi_\zeta(\xi)\right]\mathrm{d}\mu_G(\zeta),\end{aligned}$$

where the derivation utilizes the fact that $\check{\mu}_\zeta^G$ gives all its mass to reference coordinates with $\rho_G(\check{\xi}) = \zeta$, for μ_G-a.e. ζ. Hence, $\left\{\pi_\zeta\right\}$ is indeed a disintegration of $\pi_\rho^{g,G}$ with respect to μ_G, and by Lemma C.2 (iii) it may be used for $\tilde{Q}_g^{(w)}$.

For the second step, by substituting $\left\{\pi_\zeta\right\}$ into $\tilde{Q}_g^{(w)}$ in Definition 1 and changing variables through ρ_g in each integral, we have

$$\tilde{Q}_g^{(w)}(\zeta,\zeta') = \int_\Xi\int_\Xi P_g^{(w)}(\xi,\xi')\mathrm{d}\pi_\zeta(\xi)\mathrm{d}\pi_{\zeta'}(\xi') = \iint_{\check{\Xi}\times\check{\Xi}} P_g^{(w)}\left(\rho_g(\check{\xi}),\rho_g(\check{\xi}')\right)\mathrm{d}\check{\mu}_\zeta^G(\check{\xi})\mathrm{d}\check{\mu}_{\zeta'}^G(\check{\xi}') = \bar{f}_g^{(w)}(\zeta,\zeta'),$$

where the last equality uses $P_g^{(w)}\left(\rho_g(\check{\xi}),\rho_g(\check{\xi}')\right) = \check{f}_g^{(w)}(\check{\xi},\check{\xi}')$ from Definition 3 and the definition of $\bar{f}_g^{(w)}$ in equation (5).

Proof of statement (ii). By Definition 3, $\check{f}_G^{(w)} = P_G^{(w)} \circ (\rho_G \times \rho_G)$ at every pair of reference coordinates, where $P_G^{(w)} = f^{(w)}(\cdot,\cdot,\gamma_G)$ is Borel on $\Xi \times \Xi$ and bounded on $\mathrm{supp}(\mu_G) \times \mathrm{supp}(\mu_G)$. Hence, $\check{f}_G^{(w)}$ has the form required in Lemma A.2 (iv) by setting $H = P_G^{(w)}$, and it follows that $\bar{f}_G^{(w)} = P_G^{(w)}$ in $L^2(\mu_G \otimes \mu_G)$.

Proof of statement (iii). We first show $\left(\theta_g\right)_\# \mu_G = \mu_g$. Since $\rho_g = \theta_g \circ \rho_G$ except on a $\check{\mu}$-null set, the two maps push $\check{\mu}$ forward to the same measure, and hence

$$\left(\theta_g\right)_\#\mu_G = \left(\theta_g\right)_\#\left(\rho_G\right)_\#\check{\mu} = \left(\theta_g\circ\rho_G\right)_\#\check{\mu} = \left(\rho_g\right)_\#\check{\mu} = \mu_g,$$

where the first and last equalities utilize $(\rho_G)_{\#}\breve{\mu} = \mu_G$ and $\left(\rho_g\right)_{\#}\breve{\mu} = \mu_g$ from Lemma 1 (i).

We then want to show $\tilde{Q}_g^{(w)} = P_g^{(w)} \circ \left(\theta_g \times \theta_g\right)$. Let $H := P_g^{(w)} \circ \left(\theta_g \times \theta_g\right)$. As $P_g^{(w)}$ is bounded and θ_g takes values in $\text{supp}(\mu_g)$, H is Borel on $\Xi \times \Xi$. Now consider a pair $(\check{\xi}, \check{\xi}')$ such that $\rho_g = \theta_g \circ \rho_G$ holds at both $\check{\xi}$ and $\check{\xi}'$. This kind of pairs form a set of full $\breve{\mu} \otimes \breve{\mu}$-measure, and for each of them, we have

$$\breve{f}_g^{(w)}(\check{\xi}, \check{\xi}') = P_g^{(w)}\left(\rho_g(\check{\xi}), \rho_g(\check{\xi}')\right) = P_g^{(w)}\left(\theta_g\left(\rho_G(\check{\xi})\right), \theta_g\left(\rho_G(\check{\xi}')\right)\right) = H\left(\rho_G(\check{\xi}), \rho_G(\check{\xi}')\right).$$

That is, $\breve{f}_g^{(w)} = H \circ (\rho_G \times \rho_G)$ $\breve{\mu} \otimes \breve{\mu}$-a.e. And by Lemma A.2 (iv), it follows that $\bar{f}_g^{(w)} = H$ in $L^2(\mu_G \otimes \mu_G)$. Combining with statement (i) completes the proof.

□

A.1.3 Proof of Proposition 2

To analyze the downstream imputation properties presented in the Proposition 2, we first need to discuss the weight optimization step and how the weight perturbation will affect the results.

Lemma A.3 (Weight perturbation)
Let $G \geq 3$, and let $\underline{V}$ and $\mathcal{V}^{(w)}$ be as in (6) and (7). Furthermore, let $\underline{\breve{V}}$ be defined as $\underline{V}$ with $\breve{f}_g^{(0)}$ and $L^2(\breve{\mu} \otimes \breve{\mu})$ in place of $\bar{f}_g^{(0)}$ and $L^2(\mu_G \otimes \mu_G)$. The following statements hold:

(*i*) *The minimum defining $\underline{V}$ is attained, and for every $v \in \mathbb{R}^{G-1}$ with $\mathbf{1}^\intercal v = 0$,*

$$\left\| \sum_{g=1}^{G-1} v_g \bar{f}_g^{(0)} \right\|_{L^2(\mu_G \otimes \mu_G)} \geq \underline{V} \|v\|_2 .$$

(*ii*) *$\underline{V} > 0$ if and only if $\mathcal{V}^{(0)} = \{0\}$, that is, if and only if the family $\left\{\bar{f}_g^{(0)}\right\}_{g=1}^{G-1}$ is affinely independent in $L^2(\mu_G \otimes \mu_G)$.*

(*iii*) *For every $v \in \mathbb{R}^{G-1}$,*

$$\left\| \sum_{g=1}^{G-1} v_g \bar{f}_g^{(0)} \right\|_{L^2(\mu_G \otimes \mu_G)} \leq \left\| \sum_{g=1}^{G-1} v_g \breve{f}_g^{(0)} \right\|_{L^2(\breve{\mu} \otimes \breve{\mu})},$$

and consequently $\underline{V} \leq \underline{\breve{V}}$. If moreover $\rho_g = \theta_g \circ \rho_G$ $\breve{\mu}$-a.e. for a Borel map $\theta_g : \Xi \to \text{supp}(\mu_g)$ for every $g \in \{1, ..., G-1\}$, as in Proposition 1 (iii), then the displayed inequality is an equality for every v, and $\underline{V} = \underline{\breve{V}}$.

Proof.

Proof of statement (i). We first show that the minimum is attained. Since the objective $v \to \left\|\sum_{g=1}^{G-1} v_g \bar{f}_g^{(0)}\right\|_{L^2(\mu_G \otimes \mu_G)}$ since it is a linear map into $L^2(\mu_G \otimes \mu_G)$ followed by the norm, and each $\bar{f}_g^{(0)}$ has finite norm by Lemma A.2 (i), it follows that the objective is also continuous on $\mathbb{R}^{G-1}$. The constraint set $\left\{v \in \mathbb{R}^{G-1} : \|v\|_2 = 1, \mathbf{1}^\top v = 0\right\}$ is closed and bounded, and it is nonempty as $G \geq 3$. Hence, the minimum defining $\underline{V}$ is attained.

We then show the inequality. Fix $v \in \mathbb{R}^{G-1}$ with $\mathbf{1}^\top v = 0$. If $v = 0$, both $\left\|\sum_{g=1}^{G-1} v_g \bar{f}_g^{(0)}\right\|_{L^2(\mu_G \otimes \mu_G)}$ and $\underline{V}\|v\|_2$ equal zero. If $v \neq 0$, then $v/\|v\|_2$ has unit norm and sums to zero, which is in the constraint set. By the definition of $\underline{V}$ being the minimum over the constraint set, we ahve

$$\left\|\sum_{g=1}^{G-1} \frac{v_g}{\|v\|_2} \bar{f}_g^{(0)}\right\|_{L^2(\mu_G \otimes \mu_G)} \geq \underline{V}.$$

Multiplying both sides by $\|v\|_2$ completes the proof.

Proof of statement (ii). We first show that $\underline{V} > 0$ implies $\mathcal{V}^{(0)} = \{0\}$. Suppose $\underline{V} > 0$ and let $v \in \mathcal{V}^{(0)}$. By the definition of $\mathcal{V}^{(0)}$, we know that $\mathbf{1}^\top v = 0$ and $\sum_{g=1}^{G-1} v_g \bar{f}_g^{(0)} = 0$ in $L^2(\mu_G \otimes \mu_G)$. Applying the inequality in statement (i) to this v leads to

$$0 = \left\|\sum_{g=1}^{G-1} v_g \bar{f}_g^{(0)}\right\|_{L^2(\mu_G \otimes \mu_G)} \geq \underline{V}\|v\|_2 .$$

Since $\underline{V} > 0$, it forces $\|v\|_2 = 0$, which leads to $v = 0$. Hence $\mathcal{V}^{(0)} = \{0\}$.

We then show that $\underline{V} = 0$ implies $\mathcal{V}^{(0)} \neq \{0\}$. Suppose $\underline{V} = 0$. Again, by statement (i), the minimum is attained at some v with $\|v\|_2 = 1$ and $\mathbf{1}^\top v = 0$, and $\underline{V} = 0$ means $\sum_{g=1}^{G-1} v_g \bar{f}_g^{(0)} = 0$ in $L^2(\mu_G \otimes \mu_G)$. This v is a nonzero element of $\mathcal{V}^{(0)}$.

Finally, by definition, the family $\left\{\bar{f}_g^{(0)}\right\}_{g=1}^{G-1}$ is affinely independent in $L^2(\mu_G \otimes \mu_G)$ if the only $v \in \mathbb{R}^{G-1}$ with $\mathbf{1}^\top v = 0$ and $\sum_{g=1}^{G-1} v_g \bar{f}_g^{(0)} = 0$ in $L^2(\mu_G \otimes \mu_G)$ is $v = 0$. This is exactly the statement $\mathcal{V}^{(0)} = \{0\}$.

Proof of statement (iii). We first prove the inequality. Apply Lemma A.2 (iii) with the coefficients set to $a = (v_1, ..., v_{G-1}, 0)$, where the last 0 is for the sums in that lemma run over the source networks only. With this, inequality is shown.

We then show $\underline{V} \leq \check{V}$. Let $\check{v}$ attain the minimum defining $\check{V}$, which exists by the same argument as in statement (i). Since $\check{v}$ has unit norm and sums to zero, it is also feasible for the minimization defining $\underline{V}$. Hence

$$\underline{V} \leq \left\|\sum_{g=1}^{G-1} \check{v}_g \bar{f}_g^{(0)}\right\|_{L^2(\mu_G \otimes \mu_G)} \leq \left\|\sum_{g=1}^{G-1} \check{v}_g \check{f}_g^{(0)}\right\|_{L^2(\check{\mu} \otimes \check{\mu})} = \check{V}.$$

Lastly, we show the case of $\underline{V} \leq \check{\underline{V}}$ under additional conditions. Suppose we have $\rho_g = \theta_g \circ \rho_G$ $\check{\mu}$-a.e. for every g. Notice that both sides of the inequality are then norms of the same function, but one on the reference space and another on the target space. With the change of variables through ρ_G, we can make the two norms equal.

To see this, fix $v \in \mathbb{R}^{G-1}$ and define

$$H(\zeta, \zeta') := \sum_{g=1}^{G-1} v_g P_g^{(0)}\left(\theta_g(\zeta), \theta_g(\zeta')\right), \qquad (\zeta, \zeta') \in \Xi \times \Xi,$$

which is Borel and bounded on $\text{supp}(\mu_G) \times \text{supp}(\mu_G)$. In the proof of Proposition 1 (iii), we learned that, on the reference space, $\check{f}_g^{(0)}(\check{\xi}, \check{\xi}') = P_g^{(0)}\left(\theta_g(\rho_G(\check{\xi})), \theta_g(\rho_G(\check{\xi}'))\right)$ for $\check{\mu} \otimes \check{\mu}$-a.e. $(\check{\xi}, \check{\xi}')$; on the target space, $\bar{f}_g^{(0)}(\zeta, \zeta') = P_g^{(0)}\left(\theta_g(\zeta), \theta_g(\zeta')\right)$ for $\mu_G \otimes \mu_G$-a.e. (ζ, ζ'). Multiplying both by v_g and summing over g, we have

$$\sum_{g=1}^{G-1} v_g \check{f}_g^{(0)}(\check{\xi}, \check{\xi}') = H\left(\rho_G(\check{\xi}), \rho_G(\check{\xi}')\right) \qquad \text{and} \qquad \sum_{g=1}^{G-1} v_g \bar{f}_g^{(0)}(\zeta, \zeta') = H(\zeta, \zeta'),$$

$\check{\mu} \otimes \check{\mu}$-a.e. and $\mu_G \otimes \mu_G$-a.e., respectively. We then compute the squared norm on the reference space. Since ρ_G pushes $\check{\mu}$ forward to μ_G, integrating a function of $\left(\rho_G(\check{\xi}), \rho_G(\check{\xi}')\right)$ over $\check{\mu} \otimes \check{\mu}$ is the same as integrating that function of (ζ, ζ') over $\mu_G \otimes \mu_G$. Hence

$$\left\| \sum_{g=1}^{G-1} v_g \check{f}_g^{(0)} \right\|_{L^2(\check{\mu} \otimes \check{\mu})}^2 = \iint_{\check{\Xi} \times \check{\Xi}} H\left(\rho_G(\check{\xi}), \rho_G(\check{\xi}')\right)^2 \mathrm{d}(\check{\mu} \otimes \check{\mu}) = \iint_{\Xi \times \Xi} H(\zeta, \zeta')^2 \, \mathrm{d}(\mu_G \otimes \mu_G) = \left\| \sum_{g=1}^{G-1} v_g \bar{f}_g^{(0)} \right\|_{L^2(\mu_G \otimes \mu_G)}^2.$$

Since this equality for every v, it follows that $\underline{V} = \check{\underline{V}}$.

□

We then proceed to prove Proposition 2.

Proof.

Throughout the proof, by the hypothesis and Proposition 1 (i), we know that $\tilde{Q}_g^{(w)} = \bar{f}_g^{(w)}$ in $L^2(\mu_G \otimes \mu_G)$ for every g and both w. With this, we know that for $w \in \{0, 1\}$,

$$\sum_{g=1}^{G-1} \delta_g^* \bar{f}_g^{(w)} = P_G^{(w)} \qquad \text{in } L^2(\mu_G \otimes \mu_G). \tag{22}$$

This is because, by Assumption 2, we have $\sum_{g=1}^{G-1} \delta_g^* \check{f}_g^{(w)} - \check{f}_G^{(w)} = 0$ $\check{\mu} \otimes \check{\mu}$-a.e. Applying Lemma A.2 (ii) with the coefficients $\left(\delta_1^*, ..., \delta_{G-1}^*, -1\right)$, we have $\sum_{g=1}^{G-1} \delta_g^* \bar{f}_g^{(w)} - \bar{f}_G^{(w)} = 0$ $\mu_G \otimes \mu_G$-a.e. That is, $\sum_{g=1}^{G-1} \delta_g^* \bar{f}_g^{(w)} = \bar{f}_G^{(w)}$ in $L^2(\mu_G \otimes \mu_G)$. By Proposition 1 (ii), we then replace $\bar{f}_G^{(w)}$ by $P_G^{(w)}$, and the result of equation (22) is acquired.

Proof of statement (i). We first show that the minimum is attained. Since $\delta \to \left\| P_G^{(0)} - \sum_{g=1}^{G-1} \delta_g \tilde{Q}_g^{(0)} \right\|_{L^2(\mu_G \otimes \mu_G)}$ is an affine map into $L^2(\mu_G \otimes \mu_G)$ followed by the norm, and since each $\tilde{Q}_g^{(0)}$ has finite norm by Lemma C.2 (i), the objective is continuous on $\mathbb{R}^{G-1}$. On top of that, because the simplex Δ_{G-2} is closed, bounded, and nonempty, it follows that the minimum is attained and $\mathcal{D}$ is nonempty.

We then show that the minimum value is zero and $\delta^* \in \mathcal{D}$. Computing the objective at $\delta = \delta^*$ and using $\tilde{Q}_g^{(0)} = \bar{f}_g^{(0)}$, we have

$$\left\| P_G^{(0)} - \sum_{g=1}^{G-1} \delta_g^* \tilde{Q}_g^{(0)} \right\|_{L^2(\mu_G \otimes \mu_G)} = \left\| P_G^{(0)} - \sum_{g=1}^{G-1} \delta_g^* \bar{f}_g^{(0)} \right\|_{L^2(\mu_G \otimes \mu_G)} = 0,$$

where the last equality comes from above derived equation (22). Since the objective is nonnegative, δ^* is attained with the minimum value of objective being zero.

Proof of statement (ii). Since the minimum value is zero by statement (i), $\delta \in \mathcal{D}$ if and only if

$$\sum_{g=1}^{G-1} \delta_g \bar{f}_g^{(0)} = P_G^{(0)} \quad \text{in } L^2(\mu_G \otimes \mu_G).$$

And since $P_G^{(0)} = \sum_{g=1}^{G-1} \delta_g^* \bar{f}_g^{(0)}$ in $L^2(\mu_G \otimes \mu_G)$ by (22), this holds if and only if

$$\sum_{g=1}^{G-1} \left(\delta_g - \delta_g^* \right) \bar{f}_g^{(0)} = 0 \quad \text{in } L^2(\mu_G \otimes \mu_G).$$

As δ and δ^* both sum to one, $\mathbf{1}^\intercal(\delta - \delta^*) = 0$. Hence, by the definition of $\mathcal{V}^{(0)}$ in (7), $\delta \in \mathcal{D}$ if and only if $\delta - \delta^* \in \mathcal{V}^{(0)}$. Further more, if $\underline{V} > 0$, Lemma A.3 (ii) indicates that $\mathcal{V}^{(0)} = \{0\}$. Therefore, combining these arguments leads to $\mathcal{D} = \{\delta \in \Delta_{G-2} : \delta = \delta^*\} = \{\delta^*\}$.

Proof of statement (iii). Notice that with $\underline{V} > 0$, statement (ii) shows that $\mathcal{D} = \{\delta^*\}$. As the Gromov-Wasserstein barycenter analysis problem returns an element of $\mathcal{D}$, it will then return such $\tilde{\delta} = \delta^*$.

With $\tilde{\delta} = \delta^*$ and $\tilde{Q}_g^{(1)} = \bar{f}_g^{(1)}$ by Proposition 1 (i), the final imputation is then

$$\tilde{P}_G^{(1)} = \sum_{g=1}^{G-1} \tilde{\delta}_g \tilde{Q}_g^{(1)} = \sum_{g=1}^{G-1} \delta_g^* \bar{f}_g^{(1)}.$$

With the right-hand side equals $P_G^{(1)}$ in $L^2(\mu_G \otimes \mu_G)$ by (22), we have $\tilde{P}_G^{(1)} = P_G^{(1)}$ in $L^2(\mu_G \otimes \mu_G)$.

□

A.1.4 Proof of Proposition 3

We then start to prove the Proposition 3.

Proof.

Throughout the proof, let $\check{\pi} \in \Pi(\check{\mu}, \check{\mu})$ be an optimal coupling of the reference problem (4) with $\pi = \left(\rho_g \times \rho_G\right)_{\#} \check{\pi}$, which exists by Theorem 1 (ii). We work on $\check{\pi} \otimes \check{\pi}$ with generic $\left((\check{\xi}, \check{\zeta}), (\check{\xi}', \check{\zeta}')\right)$. In each pair $(\check{\xi}, \check{\zeta})$ drawn from $\check{\pi}$, the first coordinate $\check{\xi}$ is the reference coordinate read through ρ_g as a source position and the second coordinate $\check{\zeta}$ is the reference coordinate read through ρ_G as a target position, so that $\check{\pi}$ matches source reference coordinates to target reference coordinates. Since both marginals of $\check{\pi}$ are $\check{\mu}$, the pairs $(\check{\xi}, \check{\xi}')$ and $(\check{\zeta}, \check{\zeta}')$ both have the distribution of $\check{\mu} \otimes \check{\mu}$.

Showing that $\check{\pi}$ preserves the target dissimilarity. Since the optimal coupling from the reference problem $\check{\pi}$ only states that it ranks source and target dissimilarities consistently, first, we would like to show that $\check{\pi}$ do preserve the target dissimilarity, that is,

$$\check{d}_G^2(\check{\zeta}, \check{\zeta}') = \check{d}_G^2(\check{\xi}, \check{\xi}') \qquad \text{for } \check{\pi} \otimes \check{\pi}\text{-a.e. } \left((\check{\xi}, \check{\zeta}), (\check{\xi}', \check{\zeta}')\right). \tag{23}$$

Denote $U := \check{d}_g^2(\check{\xi}, \check{\xi}')$, $V := \check{d}_G^2(\check{\zeta}, \check{\zeta}')$, and $W := \check{d}_G^2(\check{\xi}, \check{\xi}')$. That is, U is the source dissimilarity of the two source reference coordinates, V is the target dissimilarity of the two target reference coordinates matched to them by $\check{\pi}$, and W is the target dissimilarity of the same two source reference coordinates. With this notation, equation (23) reads as $V = W$ almost surely. By Assumption 4 (i) and Proposition B.2, there exist a Borel function $b_g : [0, \infty) \to [0, 1]$ with $\check{d}_G^2 = b_g \circ \check{d}_g^2$ $\check{\mu} \otimes \check{\mu}$-a.e. Since $(\check{\xi}, \check{\xi}')$ has distribution $\check{\mu} \otimes \check{\mu}$, it follows that $W = b_g(U)$ almost surely.

Since $W = b_g(U)$ relates the two dissimilarities of the same pair of source reference coordinates, while V is evaluated at the target reference coordinates matched to them, we need to transfer the relation from (U, W) to (U, V). This can be done by showing that the two pairs have the same joint distribution. Since $(\check{\xi}, \check{\xi}')$ has distribution $\check{\mu} \otimes \check{\mu}$, the pair (U, W) is comonotone by Assumption 3. Also, since $\check{\pi}$ is optimal for the reference problem, the pair (U, V) is comonotone by Theorem 1 (ii). As both V and W are $\check{d}_G^2$ evaluated at a pair with distribution $\check{\mu} \otimes \check{\mu}$, they have the same distribution. Lastly, because two comonotone pairs with the same marginal distributions have the same joint distribution, as shown in the proof of Proposition B.2, (U, V) and (U, W) also have the same joint distribution.

We now transfer the relation. Since the two pairs have the same joint law, the bounded Borel function $(u, v) \to \left(v - b_g(u)\right)^2$, which is zero when $v = b_g(u)$, has the same expectation under both

pairs,

$$\mathbb{E}\left[\left(V - b_g(U)\right)^2\right] = \mathbb{E}\left[\left(W - b_g(U)\right)^2\right].$$

With $W = b_g(U)$ almost surely, the right-hand side is zero, so the left-hand side is zero as well. Hence $V = b_g(U)$ almost surely. Since also $W = b_g(U)$ almost surely, we obtain $V = W$ almost surely, which is (23).

Showing that $\check{\pi}$ preserves the source graphon. We then show that $\check{\pi}$ preserves the graphon of network g. That is, for $w \in \{0, 1\}$,

$$\check{f}_g^{(w)}(\check{\xi}, \check{\xi}') = \check{f}_g^{(w)}(\check{\zeta}, \check{\zeta}') \qquad \text{for } \check{\pi} \otimes \check{\pi}\text{-a.e. } \left((\check{\xi}, \check{\zeta}), (\check{\xi}', \check{\zeta}')\right). \tag{24}$$

This is what allows the projection, which evaluates $\check{f}_g^{(w)}$ at the source reference coordinates matched by $\check{\pi}$, to be rewritten as an evaluation at the target reference coordinates, which is how $\bar{f}_g^{(w)}$ is defined. Since $\check{\pi}$ preserves the target dissimilarity by (23), Assumption 4 (ii) leads to the desired (24).

Showing that $\tilde{Q}_g^{(w)} = \bar{f}_g^{(w)}$. Now, we're ready to show that $\tilde{Q}_g^{(w)} = \bar{f}_g^{(w)}$ in $L^2(\mu_G \otimes \mu_G)$. Since Definition 1 defines $\tilde{Q}_g^{(w)}$ through a disintegration of π on $\Xi \times \Xi$, but (24) is a statement about $\check{\pi}$ on $\check{\Xi} \times \check{\Xi}$, we first express that disintegration through $\check{\pi}$.

Let $\left\{\check{\pi}_\zeta\right\}_{\zeta \in \Xi}$ be the disintegration of $\check{\pi}$ with respect to the map $(\check{\xi}, \check{\zeta}) \to \rho_G(\check{\zeta})$, obtained as $\left\{\check{\mu}_\zeta^G\right\}_{\zeta \in \Xi}$ was in Section 4.2. That is, $\check{\pi}_\zeta$ is the conditional distribution of $(\check{\xi}, \check{\zeta}) \sim \check{\pi}$ given $\rho_G(\check{\zeta}) = \zeta$, and $\check{\pi} = \int_\Xi \check{\pi}_\zeta \mathrm{d}\mu_G(\zeta)$. Let π_ζ be the distribution of $\rho_g(\check{\xi})$ under $\check{\pi}_\zeta$. Since $\pi = \left(\rho_g \times \rho_G\right)_\# \check{\pi}$, π_ζ is the conditional distribution of $\xi = \rho_g(\check{\xi})$ given $\zeta = \rho_G(\check{\zeta})$ under π, so $\left\{\pi_\zeta\right\}_{\zeta \in \Xi}$ is the disintegration of π with respect to its second marginal μ_G in the sense of Lemma C.1.

With Definition 1 using this disintegration, for $\mu_G \otimes \mu_G$-a.e. (ζ, ζ'), we have

$$\begin{aligned}
\tilde{Q}_g^{(w)}(\zeta, \zeta') &= \iint f^{(w)}(\xi, \xi', \gamma_g) \mathrm{d}\pi_\zeta(\xi) \mathrm{d}\pi_{\zeta'}(\xi') \\
&= \iint \check{f}_g^{(w)}(\check{\xi}, \check{\xi}') \mathrm{d}\check{\pi}_\zeta(\check{\xi}, \check{\zeta}) \mathrm{d}\check{\pi}_{\zeta'}(\check{\xi}', \check{\zeta}') \\
&= \iint \check{f}_g^{(w)}(\check{\zeta}, \check{\zeta}') \mathrm{d}\check{\pi}_\zeta(\check{\xi}, \check{\zeta}) \mathrm{d}\check{\pi}_{\zeta'}(\check{\xi}', \check{\zeta}') \\
&= \iint_{\check{\Xi} \times \check{\Xi}} \check{f}_g^{(w)}(\check{\zeta}, \check{\zeta}') \mathrm{d}\check{\mu}_\zeta^G(\check{\zeta}) \mathrm{d}\check{\mu}_{\zeta'}^G(\check{\zeta}') \\
&= \bar{f}_g^{(w)}(\zeta, \zeta').
\end{aligned}$$

The second equality is the change of variables $\xi = \rho_g(\check{\xi})$ and $\xi' = \rho_g(\check{\xi}')$. The third equality switches the integrand from source reference coordinates to target reference coordinates. Since

$\check{\pi} \otimes \check{\pi} = \iint \check{\pi}_\zeta \otimes \check{\pi}_{\zeta'} \mathrm{d}\mu_G(\zeta)\mathrm{d}\mu_G(\zeta')$ and (24) holds $\check{\pi} \otimes \check{\pi}$-a.e., the two integrands are equal $\check{\pi}_\zeta \otimes \check{\pi}_{\zeta'}$-a.e. for $\mu_G \otimes \mu_G$-a.e. (ζ, ζ'). The fourth equality integrates out the source reference coordinates, on which the integrand no longer depends. Since the distributions of $\check{\xi}$ under the $\check{\pi}_\zeta$ form a disintegration of $\check{\mu}$ with respect to ρ_G, they equal $\check{\mu}_\zeta^G$ for μ_G-a.e. ζ by the a.e. uniqueness of the disintegration, and likewise for $\check{\zeta}'$.

Since $\tilde{Q}_g^{(w)}$ belongs to $L^2(\mu_G \otimes \mu_G)$ by Lemma C.2 (i), and $\bar{f}_g^{(w)}$ belongs to $L^2(\mu_G \otimes \mu_G)$ by Lemma A.2 (i), this $\mu_G \otimes \mu_G$-a.e. equality is the desired equality in $L^2(\mu_G \otimes \mu_G)$.

□

A.1.5 Proof of Theorem 2

Finally, we prove the Theorem 2.

Proof.

We first show that the Gromov-Wasserstein projections satisfy $\tilde{Q}_g^{(w)} = \bar{f}_g^{(w)}$ in $L^2(\mu_G \otimes \mu_G)$ for every $g \in \{1, ..., G-1\}$ and both $w \in \{0, 1\}$. Fix g and let π be the optimal coupling returned by the Gromov-Wasserstein optimization for this g. Since Assumptions 3 and 4 hold, Proposition 3 gives $\tilde{Q}_g^{(w)} = \bar{f}_g^{(w)}$ in $L^2(\mu_G \otimes \mu_G)$ for both w, whichever optimal coupling π is.

Since Assumption 2 holds and the projections satisfy $\tilde{Q}_g^{(w)} = \bar{f}_g^{(w)}$, Proposition 2 applies. Since $\underline{V} > 0$, Proposition 2 (ii) gives $\mathcal{D} = \{\delta^*\}$, so the weights optimization has the unique solution $\tilde{\delta} = \delta^*$. Proposition 2 (iii) then gives $\tilde{P}_G^{(1)} = P_G^{(1)}$ in $L^2(\mu_G \otimes \mu_G)$.

□

A.2 Proof for the Large Sample Theory in Section 5: Proposition 5

We separate the proof for Proposition 5 into following steps:

- We first try to answer the question of if two reference coordinates $\check{\xi}$ and $\check{\xi}'$ have small population pseudo distance. Will these two reference coordinates also be close in $\check{\Xi}$. *Lemma A.4 shows that the small d_g^{pseudo} will leads to nearby reference coordinates.*
- Now, since the population pseudo distance can correctly recover the neighborhood in reference coordinates $\check{\xi}$, we then investigate the intermediate oracle pseudo distance's property. *And indeed, in Lemma A.5, we will find that the intermediate oracle pseudo distance can also preserve the neighborhood structure in terms of the reference coordinates.*
- We then analyze the infeasible oracle estimator. *With the oracle neighborhood and the oracle kernel estimator, Lemma A.6 derives its error.*

- Lastly, we transfer the oracle result to the actual proposed estimator. *We show that the feasible estimator is close to the oracle and complete the proof for Proposition 5.*

Lemma A.4 (Informativeness of the population pseudo-distance)

Suppose Assumptions 3-5 hold. For every $g \le G$, recall the population pseudo distance of Section 5:

$$d_g^{\text{pseudo}}(\check{\xi}, \check{\xi}') := \sup_{\check{\zeta} \in \check{\Xi}} \left| \int_{\check{\Xi}} \left[\check{f}_g^{(0)}(\check{\xi}, \check{\eta}) - \check{f}_g^{(0)}(\check{\xi}', \check{\eta}) \right] \check{f}_g^{(0)}(\check{\zeta}, \check{\eta}) \mathrm{d}\check{\mu}(\check{\eta}) \right|.$$

We have

(*i*) *For every $g < G$, there exists a continuous nondecreasing function $b_g : \check{d}_g^2(\check{\Xi}^2) \to \check{d}_G^2(\check{\Xi}^2)$ such that*

$$\check{d}_G^2 = b_g \circ \check{d}_g^2 \qquad \text{with} \qquad b_g(0) = 0.$$

Also, for every $g \le G$, $\check{d}_g$ is a metric on $\check{\Xi}$ and

$$\check{d}_g(\check{\xi}, \check{\xi}') = 0 \iff \check{\xi} = \check{\xi}' \qquad \text{and} \qquad \check{d}_g(\check{\xi}, \check{\xi}') \le L_f \|\check{\xi} - \check{\xi}'\|.$$

(*ii*) *The pseudo-distance is jointly continuous. For every $g \le G$ and $\check{\xi}, \check{\xi}' \in \check{\Xi}$,*

$$\frac{1}{2} \check{d}_g(\check{\xi}, \check{\xi}')^2 \le d_g^{\text{pseudo}}(\check{\xi}, \check{\xi}') \le \check{d}_g(\check{\xi}, \check{\xi}') \le L_f \|\check{\xi} - \check{\xi}'\|$$

and

$$\|\check{d}_g^2(\check{\xi}, \cdot) - \check{d}_g^2(\check{\xi}', \cdot)\|_{L^2(\check{\mu})} \le 4 d_g^{\text{pseudo}}(\check{\xi}, \check{\xi}').$$

Consequently, for every $\varepsilon > 0$,

$$\min_{g \le G} \min_{\check{\xi}, \check{\xi}' \in \check{\Xi},\ \|\check{\xi} - \check{\xi}'\| \ge \varepsilon} d_g^{\text{pseudo}}(\check{\xi}, \check{\xi}') > 0.$$

(*iii*) *If in addition Assumption 6 holds, then, with $c_1 = \frac{1}{2}\sqrt{\kappa} \min\{1, 1/\bar{c}_b\} > 0$,*

$$\|\check{d}_g^2(\check{\xi}, \cdot) - \check{d}_g^2(\check{\xi}', \cdot)\|_{L^2(\check{\mu})} \ge c_1 \|\check{\xi} - \check{\xi}'\|, \qquad g \le G,$$

and hence

$$d_g^{\text{pseudo}}(\check{\xi}, \check{\xi}') \ge \frac{1}{4} c_1 \|\check{\xi} - \check{\xi}'\|.$$

Proof.

Proof of statement (i). Since $\check{f}_g^{(0)}$ is L_f-Lipschitz by Assumption 5 (iii), we have $\check{d}_g(\check{\xi},\check{\xi}') \le L_f\|\check{\xi}-\check{\xi}'\|$. And since $\check{d}_g \le 1$, it follows that $\check{d}_g^2$ is continuous and $2L_f$-Lipschitz in each argument. With this and the full support of $\check{\mu}$, the comonotonicity in Assumption 3 holds at every pair of points $\check{\xi},\check{\xi}',\check{\zeta},\check{\zeta}' \in \check{\Xi}$:

$$\left[\check{d}_g^2(\check{\xi},\check{\xi}') - \check{d}_g^2(\check{\zeta},\check{\zeta}')\right]\left[\check{d}_G^2(\check{\xi},\check{\xi}') - \check{d}_G^2(\check{\zeta},\check{\zeta}')\right] \ge 0. \tag{25}$$

We first show that pairs with equal distance in network g have equal distance in network G, that is,

$$\check{d}_g^2(\check{\xi},\check{\xi}') = \check{d}_g^2(\check{\zeta},\check{\zeta}') \qquad \text{implies} \qquad \check{d}_G^2(\check{\xi},\check{\xi}') = \check{d}_G^2(\check{\zeta},\check{\zeta}'),$$

and we prove it by contradiction. Suppose not: $\check{d}_g^2(\check{\xi},\check{\xi}') = \check{d}_g^2(\check{\zeta},\check{\zeta}')$ but $\check{d}_G^2(\check{\xi},\check{\xi}') < \check{d}_G^2(\check{\zeta},\check{\zeta}')$ for some $\check{\xi},\check{\xi}',\check{\zeta},\check{\zeta}'$. With this, we know that the set

$$\left\{(\check{\eta},\check{\eta}') \in \check{\Xi}\times\check{\Xi} : \check{d}_G^2(\check{\xi},\check{\xi}') < \check{d}_G^2(\check{\eta},\check{\eta}') < \check{d}_G^2(\check{\zeta},\check{\zeta}')\right\}$$

is open and nonempty. And in this set, by (25), we know that $\check{d}_g^2(\check{\eta},\check{\eta}') \ge \check{d}_g^2(\check{\xi},\check{\xi}')$ and $\check{d}_g^2(\check{\eta},\check{\eta}') \le \check{d}_g^2(\check{\zeta},\check{\zeta}') = \check{d}_g^2(\check{\xi},\check{\xi}')$, which means that $\check{d}_g^2$ equals $\check{d}_g^2(\check{\xi},\check{\xi}')$ on the whole set. As the open and nonempty set has positive mass, $\check{d}_g^2$ takes the single value $\check{d}_g^2(\check{\xi},\check{\xi}')$ with positive probability. This contradicts Assumption 4 (i).

Since $\check{d}_g^2(\check{\xi},\check{\xi}') = \check{d}_g^2(\check{\zeta},\check{\zeta}')$ indeed implies $\check{d}_G^2(\check{\xi},\check{\xi}') = \check{d}_G^2(\check{\zeta},\check{\zeta}')$, there exists a function b_g with $\check{d}_G^2 = b_g \circ \check{d}_g^2$. We also know that $b_g(0) = 0$ as $\check{d}_g^2(\check{\xi},\check{\xi}) = \check{d}_G^2(\check{\xi},\check{\xi}) = 0$. On top of that, (25) makes b_g nondecreasing. Since $\check{d}_g^2$ is a continuous map from the compact set $\check{\Xi}\times\check{\Xi}$ onto its range, and $\check{d}_G^2$ is continuous, hence the map b_g linking them is continuous as well.

Being an $L^2(\check{\mu})$ distance, each $\check{d}_g$ is symmetric and satisfies the triangle inequality. If $\check{d}_g(\check{\xi},\check{\xi}') = 0$, then $\check{d}_G^2(\check{\xi},\check{\xi}') = b_g(0) = 0$. With this, Assumption 5 (iv) guarantees $\check{\xi} = \check{\xi}'$. Therefore, we conclude that $\check{d}_g$ is a metric.

Proof of statement (ii). Here we follow the graphon slice calculation as in Zhang et al. (2017) and Deaner et al. (2025).

For the first lower bound, write

$$\check{d}_g(\check{\xi},\check{\xi}')^2 = \int_{\check{\Xi}}\left[\check{f}_g^{(0)}(\check{\xi},\check{\eta}) - \check{f}_g^{(0)}(\check{\xi}',\check{\eta})\right]\check{f}_g^{(0)}(\check{\xi},\check{\eta})\mathrm{d}\check{\mu}(\check{\eta}) - \int_{\check{\Xi}}\left[\check{f}_g^{(0)}(\check{\xi},\check{\eta}) - \check{f}_g^{(0)}(\check{\xi}',\check{\eta})\right]\check{f}_g^{(0)}(\check{\xi}',\check{\eta})\mathrm{d}\check{\mu}(\check{\eta}).$$

Each integral is at most $d_g^{\text{pseudo}}(\check{\xi},\check{\xi}')$ in absolute value if we take $\check{\zeta} = \check{\xi}$ and $\check{\zeta} = \check{\xi}'$, hence $\check{d}_g^2 \le 2d_g^{\text{pseudo}}$. For the upper bound, by Cauchy-Schwarz inequality and the fact that $0 \le \check{f}_g^{(0)} \le 1$, we can bound the integral in d_g^{pseudo} by $\check{d}_g(\check{\xi},\check{\xi}')$ for every $\check{\zeta}$.

For the second inequality, with a specific $\check{\zeta} \in \check{\Xi}$, we expand the square operations and get

$$\check{d}_g^2(\check{\xi},\check{\zeta}) - \check{d}_g^2(\check{\xi}',\check{\zeta}) = \int_{\check{\Xi}} \left[\check{f}_g^{(0)}(\check{\xi},\check{\eta}) - \check{f}_g^{(0)}(\check{\xi}',\check{\eta})\right]\left[\check{f}_g^{(0)}(\check{\xi},\check{\eta}) + \check{f}_g^{(0)}(\check{\xi}',\check{\eta})\right] \mathrm{d}\check{\mu}(\check{\eta})$$
$$- 2\int_{\check{\Xi}} \left[\check{f}_g^{(0)}(\check{\xi},\check{\eta}) - \check{f}_g^{(0)}(\check{\xi}',\check{\eta})\right] \check{f}_g^{(0)}(\check{\zeta},\check{\eta}) \mathrm{d}\check{\mu}(\check{\eta}).$$

Since each line is at most $2d_g^{\mathrm{pseudo}}(\check{\xi},\check{\xi}')$ in absolute value, the difference is then at most $4d_g^{\mathrm{pseudo}}(\check{\xi},\check{\xi}')$ for every $\check{\zeta}$. Same bound extend to its $L^2(\check{\mu})$ norm.

Lastly, since graphons are continuous on the compact $\check{\Xi}$, the integral in d_g^{pseudo} is uniformly continuous in $(\check{\xi},\check{\xi}',\check{\zeta})$. Take supremum over $\check{\zeta}$ of such functions will retain it continuity, and it follows that d_g^{pseudo} is jointly continuous. Furthermore. the set of pairs $(\check{\xi},\check{\xi}')$ with $\|\check{\xi}-\check{\xi}'\| \geq \varepsilon$ is compact. For every such pair, $\check{\xi} \neq \check{\xi}'$, hence $d_g^{\mathrm{pseudo}}(\check{\xi},\check{\xi}') \geq \frac{1}{2}\check{d}_g(\check{\xi},\check{\xi}')^2 > 0$ by the first lower bound and statement (i). A continuous positive function on a compact set has a positive minimum, and the minimum over finitely many g remains positive.

Proof of statement (iii). The result is trival if $\check{\xi} = \check{\xi}'$, hence, we consider distinct $\check{\xi},\check{\xi}'$. It suffices to prove

$$\|\check{d}_G^2(\check{\xi},\cdot) - \check{d}_G^2(\check{\xi}',\cdot)\|_{L^2(\check{\mu})} \geq \frac{\sqrt{\kappa}}{2}\|\check{\xi}-\check{\xi}'\|.$$

This is because, with the same reasoning, Assumption 6 (i) then provides the same bound for every source with constant $\sqrt{\kappa}/(2\bar{c}_b)$. The second inequality is the direct result of utilizing statement (ii).

We first note that, for every $\check{\pi} \in \Pi(\check{\mu},\check{\mu})$,

$$F_{\check{d}_G^2,\check{d}_G^2}(\check{\pi}) \leq 4\int \|\check{d}_G^2(\check{\xi},\cdot) - \check{d}_G^2(\check{\zeta},\cdot)\|_{L^2(\check{\mu})}^2 \mathrm{d}\check{\pi}(\check{\xi},\check{\zeta}). \tag{26}$$

To see this, write $\check{d}_G^2(\check{\xi},\check{\xi}') - \check{d}_G^2(\check{\zeta},\check{\zeta}')$ as $[\check{d}_G^2(\check{\xi},\check{\xi}') - \check{d}_G^2(\check{\zeta},\check{\xi}')] + [\check{d}_G^2(\check{\zeta},\check{\xi}') - \check{d}_G^2(\check{\zeta},\check{\zeta}')]$ and use $(a+b)^2 \leq 2a^2 + 2b^2$. Since both marginals are $\check{\mu}$ and $\check{d}_G^2$ is symmetric, each of the two resulting integrals equals the integral on the right.

Then we want to state a fact about the coupling that may swap nearby reference coordinates. Pick radius $r > 0$ small enough such that

$$S_r = B(\check{\xi},r) \cap \check{\Xi} \qquad \text{and} \qquad S_r' = B(\check{\xi}',r) \cap \check{\Xi}$$

are disjoint. By Assumption 5 (ii), both sets have positive mass and $\check{\mu}(S_r \cup S_r') \to 0$ as $r \downarrow 0$. Draw independent reference coordinates $\check{u},\check{z}$ from $\check{\mu}$. If one of them lies in S_r and the other in S_r', set $\check{v} = \check{z}$; if they locate in the same ball, set $\check{v} = \check{u}$. Let $\check{\pi}_r$ be the distribution of $(\check{u},\check{v})$. The aforementioned rule treats $\check{u}$ and $\check{z}$ symmetrically, and hence both $\check{v}$ and $\check{u}$ have the same distribution $\check{\mu}$ with

$\check{\pi}_r \in \Pi(\check{\mu}, \check{\mu})$. In other words, this paragraphs shows that, with small probability, if a reference coordinate in one set moves to another random reference coordinate in the other set, every other reference coordinate will still stay in place.

Continue the discussion the coupling that may swap nearby reference coordinates. Fix $\check{u}$, and let p be its probability of moving, so $p \le \check{\mu}(S_r \cup S'_r)$. Then for any points $\check{t}$ and $\check{z}$, it follows that

$$(1-p)\|\check{t}-\check{u}\|^2 + p\|\check{t}-\check{z}\|^2 \ge p(1-p)\|\check{u}-\check{z}\|^2.$$

Select $\check{t} = T\check{u}$ and average $\check{z}$ over the destination of a move. The inequality becomes

$$\mathbb{E}\left[\|T\check{u}-\check{v}\|^2|\check{u}\right] \ge (1-p)\mathbb{E}\left[\|\check{u}-\check{v}\|^2|\check{u}\right]$$

Then for every Borel map T, by integrating over $\check{u}$, we have

$$\int \|T\check{u}-\check{v}\|^2 \mathrm{d}\check{\pi}_r(\check{u},\check{v}) \ge [1-\check{\mu}(S_r \cup S'_r)] \int \|\check{u}-\check{v}\|^2 \mathrm{d}\check{\pi}_r(\check{u},\check{v}). \tag{27}$$

By Assumption 6 (ii), and combining (27) and (26), it follows that

$$\kappa[1-\check{\mu}(S_r \cup S'_r)] \int \|\check{u}-\check{v}\|^2 \mathrm{d}\check{\pi}_r \le 4 \int \|\check{d}_G^2(\check{u},\cdot) - \check{d}_G^2(\check{v},\cdot)\|^2_{L^2(\check{\mu})} \mathrm{d}\check{\pi}_r.$$

Both integrands vanish when there is no move. When a move happens, $\check{u}$ and $\check{v}$ lie in opposite sets, and hence $\|\check{u}-\check{v}\| \ge \|\check{\xi}-\check{\xi}'\| - 2r$. Since, by statement (i), $\check{d}_G^2$ is bounded by one and $2L_f$-Lipschitz, we also have

$$\|\check{d}_G^2(\check{u},\cdot) - \check{d}_G^2(\check{v},\cdot)\|^2_{L^2(\check{\mu})} \le \|\check{d}_G^2(\check{\xi},\cdot) - \check{d}_G^2(\check{\xi}',\cdot)\|^2_{L^2(\check{\mu})} + Cr,$$

with C depending only on L_f. A move happens with probability $2\check{\mu}(S_r)\check{\mu}(S'_r) > 0$. Bounding both integrands when a move happens and dividing by this probability gives

$$\kappa[1-\check{\mu}(S_r \cup S'_r)](\|\check{\xi}-\check{\xi}'\| - 2r)^2 \le 4\left[\|\check{d}_G^2(\check{\xi},\cdot) - \check{d}_G^2(\check{\xi}',\cdot)\|^2_{L^2(\check{\mu})} + Cr\right].$$

Sending $r \downarrow 0$ leads to the desired

$$\kappa\|\check{\xi}-\check{\xi}'\|^2 \le 4\|\check{d}_G^2(\check{\xi},\cdot) - \check{d}_G^2(\check{\xi}',\cdot)\|^2_{L^2(\check{\mu})}.$$

□

As the population pseudo distance is not available and hard to directly analyze, hence, we introduce an intermediate oracle using the true probability matrix at the same sampled individuals. The next lemma shows both that this oracle retains the population reference coordinate information, and replacing its probabilities by observed adjacency matrices only changes the pseudo distance slightly.

Lemma A.5 (Oracle pseudo distance)

Suppose Assumptions 1-5 hold. For every network $g \le G$, define the oracle pseudo distance

$$\bar{d}_g^{\text{pseudo}}(i,k) := \max_{\ell \notin \{i,k\}} \frac{1}{N_g} \left| \sum_{a=1}^{N_g} (P_{g,i,a}^{(0)} - P_{g,k,a}^{(0)}) P_{g,\ell,a}^{(0)} \right|, \qquad i \neq k.$$

Let $\hat{d}_g^{\text{pseudo}}$ be the pseudo-distance computed from the baseline adjacency matrix in Step 0. Then:

(*i*) *For every sampled pair,*

$$\bar{d}_g^{\text{pseudo}}(i,k) \le L_f \|\check{\xi}_{g,i} - \check{\xi}_{g,k}\|.$$

For every fixed $\lambda > 0$, there is a constant $C_\lambda < \infty$ such that, for all sufficiently large network sizes, with probability at least $1 - C_\lambda \sum_{g \le G} N_g^{-\lambda}$,

$$\max_{i \neq k} \left| \hat{d}_g^{\text{pseudo}}(i,k) - \bar{d}_g^{\text{pseudo}}(i,k) \right| \le C_\lambda \left(\frac{\log(N_g)}{N_g} \right)^{1/2}, \qquad g \le G.$$

(*ii*) *For every deterministic sequence $h_N \to 0$ with $h_N > 0$,*

$$\max_{g \le G} \max_{i \neq k, \bar{d}_g^{\text{pseudo}}(i,k) \le h_N} \|\check{\xi}_{g,i} - \check{\xi}_{g,k}\| = o_p(1), \tag{28}$$

where a maximum over an empty set is taken to be zero.

(*iii*) *If in addition Assumption 6 holds, then, for every fixed $\lambda > 0$ and all sufficiently large network sizes, the bound in statement (i) and*

$$\bar{d}_g^{\text{pseudo}}(i,k) \ge \frac{c_1}{8} \|\check{\xi}_{g,i} - \check{\xi}_{g,k}\|, \qquad g \le G, \quad i \neq k,$$

hold together with probability at least $1 - C_\lambda \sum_{g \le G} N_g^{-\lambda}$, where c_1 is the constant in Lemma A.4 (iii).

Proof.

Proof of statement (i). With that fact that $0 \le P_{g,\ell,a}^{(0)} \le 1$ and $\check{f}_g^{(0)}$ being Lipschitz, for every ℓ, it follows that

$$\frac{1}{N_g} \left| \sum_{j=1}^{N_g} (P_{g,i,j}^{(0)} - P_{g,k,j}^{(0)}) P_{g,\ell,j}^{(0)} \right| \le \frac{1}{N_g} \sum_{j=1}^{N_g} \left| P_{g,i,j}^{(0)} - P_{g,k,j}^{(0)} \right| \le L_f \|\check{\xi}_{g,i} - \check{\xi}_{g,k}\|.$$

For the second bound, $\hat{d}_g^{\text{pseudo}}(i,k)$ is defined with $A_g^{(0)}$ in place of $P_g^{(0)}$. By the relation of $|\max_\ell |x_\ell| - \max_\ell |y_\ell|| \le \max_\ell |x_\ell - y_\ell|$, the difference $|\hat{d}_g^{\text{pseudo}}(i,k) - \bar{d}_g^{\text{pseudo}}(i,k)|$ is at most the largest

difference between the two sums inside the maxima. And since each difference is a row-(i,ℓ) term minus a row-(k,ℓ) term, it suffices to show that, uniformly over distinct rows i,ℓ,

$$\frac{1}{N_g}\left|\sum_{j=1}^{N_g}\left(A^{(0)}_{g,i,j}A^{(0)}_{g,\ell,j}-P^{(0)}_{g,i,j}P^{(0)}_{g,\ell,j}\right)\right|\le C_\lambda\left(\frac{\log\left(N_g\right)}{N_g}\right)^{1/2}, \tag{29}$$

and then apply it to the pairs (i,ℓ) and (k,ℓ). Same argument can be found in (Zhang et al., 2017, Proof of Lemma 2).

The two terms with $j\in\{i,\ell\}$ contribute at most $\frac{2}{N_g}$. And for each remaining $j\notin\{i,\ell\}$, these products $A^{(0)}_{g,i,j}A^{(0)}_{g,\ell,j}$ are independent, lie in $[0,1]$, and have means $P^{(0)}_{g,i,j}P^{(0)}_{g,\ell,j}$. By Bernstein's inequality and a union bound over the at most N_g^2 pairs, with probability at least $1-C_\lambda N_g^{-\lambda}$, the equation (29) hold.

Auxiliary result used for statements (ii) and (iii). Pick distinct i,k,ℓ and condition on their reference coordinates. For $j\notin\{i,k,\ell\}$, the j-th summand in $\bar{d}_g^{\text{pseudo}}(i,k)$ is

$$\left(P^{(0)}_{g,i,j}-P^{(0)}_{g,k,j}\right)P^{(0)}_{g,\ell,j}=\left[\check{f}^{(0)}_g(\check{\xi}_{g,i},\check{\xi}_{g,j})-\check{f}^{(0)}_g(\check{\xi}_{g,k},\check{\xi}_{g,j})\right]\check{f}^{(0)}_g(\check{\xi}_{g,\ell},\check{\xi}_{g,j}),$$

which depends only on the reference coordinate $\check{\xi}_{g,j}$. Since these reference coordinates are independent draws from $\check{\mu}$, the summands are independent, and each has mean

$$\int_{\check{\Xi}}[\check{f}^{(0)}_g(\check{\xi}_{g,i},\check{\eta})-\check{f}^{(0)}_g(\check{\xi}_{g,k},\check{\eta})]\check{f}^{(0)}_g(\check{\xi}_{g,\ell},\check{\eta})\mathrm{d}\check{\mu}(\check{\eta}).$$

By the Lipschitz property of $\check{f}^{(0)}_g$ and $0\le\check{f}^{(0)}_g\le 1$, each summand is at most $L_f\|\check{\xi}_{g,i}-\check{\xi}_{g,k}\|$ in absolute value. If $\check{\xi}_{g,i}=\check{\xi}_{g,k}$, the summands and their mean are all zero, so there is nothing to prove.

Otherwise, divide the summands and their mean by $\|\check{\xi}_{g,i}-\check{\xi}_{g,k}\|$. The divided summands lie in $[-L_f,L_f]$, so their deviations from the mean are at most $2L_f$ and their variances are at most L_f^2. Take $t=C_\lambda(\log\left(N_g\right)/N_g)^{1/2}$, with N_g large enough that $t\le 1$ and $N_g-3\ge N_g/2$. Bernstein's inequality then bounds the probability that their average over the N_g-3 indices $j\notin\{i,k,\ell\}$ deviates from the mean by at least t by

$$2\exp\left(-\frac{(N_g-3)\,t^2}{2(L_f^2+2L_ft/3)}\right)\le 2\exp\left(-\frac{N_g\,t^2}{4(L_f^2+L_f)}\right)=2N_g^{-C_\lambda^2/(4(L_f^2+L_f))}\le N_g^{-\lambda-3},$$

where in the last step we set $C_\lambda^2=4(L_f^2+L_f)(\lambda+4)$ and $N_g\ge 2$. Multiplying back by $\|\check{\xi}_{g,i}-\check{\xi}_{g,k}\|$, the average of the original summands is within $C_\lambda(\log\left(N_g\right)/N_g)^{1/2}\|\check{\xi}_{g,i}-\check{\xi}_{g,k}\|$ of their mean.

The full sum in $\bar{d}_g^{\text{pseudo}}(i,k)$ also contains the three summands with $j\in\{i,k,\ell\}$, and it divides by N_g rather than N_g-3. Both changes move the average by at most $C\|\check{\xi}_{g,i}-\check{\xi}_{g,k}\|/N_g$, which is

smaller than the bound above. A union bound over the at most N_g^3 triples then gives, uniformly over distinct i, k, ℓ,

$$\left|\frac{1}{N_g}\sum_{j=1}^{N_g}(P^{(0)}_{g,i,j} - P^{(0)}_{g,k,j})P^{(0)}_{g,\ell,j} - \int_{\check{\Xi}}\left[\check{f}^{(0)}_g(\check{\xi}_{g,i},\check{\eta}) - \check{f}^{(0)}_g(\check{\xi}_{g,k},\check{\eta})\right]\check{f}^{(0)}_g(\check{\xi}_{g,\ell},\check{\eta})\mathrm{d}\check{\mu}(\check{\eta})\right|$$
$$\leq C_\lambda\left(\frac{\log(N_g)}{N_g}\right)^{1/2}\|\check{\xi}_{g,i} - \check{\xi}_{g,k}\|, \tag{30}$$

with probability at least $1 - N_g^{-\lambda}$.

Fix $\eta > 0$ and cover $\check{\Xi}$ by finitely many balls of radius η centred in $\check{\Xi}$. By Assumption 5 (ii), each ball has positive mass. By binomial, with probability at least $1 - C_\eta \exp(-c_\eta N_g)$, each ball then contains at least three sampled reference coordinates. On this event, fix $i \neq k$ and any $\check{\zeta} \in \check{\Xi}$. The point $\check{\zeta}$ lies within η of some centre, and the ball around that centre contains a sampled reference coordinate $\check{\xi}_{g,\ell}$ with $\ell \notin \{i,k\}$. Hence $\|\check{\xi}_{g,\ell} - \check{\zeta}\| \leq 2\eta$. For any $\check{\zeta}, \check{\zeta}' \in \check{\Xi}$, the Lipschitz property of $\check{f}^{(0)}_g$ gives

$$\left|\int_{\check{\Xi}}\left[\check{f}^{(0)}_g(\check{\xi}_{g,i},\check{\eta}) - \check{f}^{(0)}_g(\check{\xi}_{g,k},\check{\eta})\right]\left[\check{f}^{(0)}_g(\check{\zeta},\check{\eta}) - \check{f}^{(0)}_g(\check{\zeta}',\check{\eta})\right]\mathrm{d}\check{\mu}(\check{\eta})\right| \leq L_f^2\|\check{\xi}_{g,i} - \check{\xi}_{g,k}\|\,\|\check{\zeta} - \check{\zeta}'\|.$$

The integral in d_g^{pseudo} is continuous in $\check{\zeta}$ and $\check{\Xi}$ is compact, so the supremum is attained at some $\check{\zeta}$. Take $\ell \notin \{i,k\}$ with $\|\check{\xi}_{g,\ell} - \check{\zeta}\| \leq 2\eta$. Using the definition of $\bar{d}_g^{\text{pseudo}}$, then (30), then the Lipschitz bound above, we have

$$\begin{aligned}
\bar{d}_g^{\text{pseudo}}(i,k) &\geq \left|\frac{1}{N_g}\sum_{j=1}^{N_g}(P^{(0)}_{g,i,j} - P^{(0)}_{g,k,j})P^{(0)}_{g,\ell,j}\right| \\
&\geq \left|\int_{\check{\Xi}}[\check{f}^{(0)}_g(\check{\xi}_{g,i},\check{\eta}) - \check{f}^{(0)}_g(\check{\xi}_{g,k},\check{\eta})]\check{f}^{(0)}_g(\check{\xi}_{g,\ell},\check{\eta})\mathrm{d}\check{\mu}(\check{\eta})\right| - C_\lambda\left(\frac{\log(N_g)}{N_g}\right)^{1/2}\|\check{\xi}_{g,i} - \check{\xi}_{g,k}\| \\
&\geq d_g^{\text{pseudo}}(\check{\xi}_{g,i},\check{\xi}_{g,k}) - \left[2L_f^2\eta + C_\lambda\left(\frac{\log(N_g)}{N_g}\right)^{1/2}\right]\|\check{\xi}_{g,i} - \check{\xi}_{g,k}\|.
\end{aligned} \tag{31}$$

Proof of statement (ii). Fix $\varepsilon > 0$. If $\mathrm{diam}(\check{\Xi}) < \varepsilon$, there is nothing to prove. Otherwise, let $c_\varepsilon > 0$ be the minimum in Lemma A.4 (ii), and choose η with $2L_f^2\eta\,\mathrm{diam}(\check{\Xi}) \leq c_\varepsilon/4$. By (31), with probability tending to one, every pair with $\|\check{\xi}_{g,i} - \check{\xi}_{g,k}\| \geq \varepsilon$ satisfies

$$\bar{d}_g^{\text{pseudo}}(i,k) \geq c_\varepsilon - \frac{c_\varepsilon}{4} - C_\lambda\left(\frac{\log(N_g)}{N_g}\right)^{1/2}\mathrm{diam}(\check{\Xi}) \geq \frac{c_\varepsilon}{2}$$

for large N_g. Since $h_N \to 0$, eventually we will have $h_N < c_\varepsilon/2$. Thus, no such pair enters the maximum in (28). Hence, this maximum is below ε with probability tending to one. As $\varepsilon > 0$ is arbitrary, statement (ii) is shown.

Proof of statement (iii). By Lemma A.4 (iii), we have $d_g^{\text{pseudo}}(\check{\xi}_{g,i}, \check{\xi}_{g,k}) \geq (c_1/4)\|\check{\xi}_{g,i} - \check{\xi}_{g,k}\|$. Choose η with $2L_f^2\eta \leq c_1/16$, and have N_g large enough that $C_\lambda(\log(N_g)/N_g)^{1/2} \leq c_1/16$. Then with (31), we have

$$\bar{d}_g^{\text{pseudo}}(i,k) \geq \left(\frac{c_1}{4} - \frac{c_1}{16} - \frac{c_1}{16}\right)\|\check{\xi}_{g,i} - \check{\xi}_{g,k}\| = \frac{c_1}{8}\|\check{\xi}_{g,i} - \check{\xi}_{g,k}\|.$$

With η fixed, every ball of the cover contains at least three sampled reference coordinates with probability at least $1 - C_\eta \exp\left(-c_\eta N_g\right) \geq 1 - C_\lambda N_g^{-\lambda}$ for all large N_g. The bounds (30) and those in statement (i) each hold with probability at least $1 - C_\lambda N_g^{-\lambda}$. So all of them hold together, for every $g \leq G$, with probability at least $1 - C_\lambda \sum_{g\leq G} N_g^{-\lambda}$, which proves the joint probability bound.

□

Lemma A.5 indicates which oracle neighbors can be used and how their pseudo distances change when probabilities are replaced by observed adjacency matrices.

The next Lemma A.6 further pass this information to bound the actual proposed estimator. It separates the error into oracle bias, oracle averaging noise, and the difference between feasible and oracle weights.

Lemma A.6 (Bound of denoising channel)

Suppose Assumptions 1-5 hold. Use the kernel and the setting specified in Proposition 5, and let R_{denoise} be the error of its symmetrized estimator. Suppose

$$h_N \to 0, \qquad \frac{\log(N_{\min})}{N_{\min} h_N^{d_\xi}} \to 0, \qquad \frac{(\log(N_{\min})/N_{\min})^{1/2}}{h_N} \to 0. \tag{32}$$

For every fixed $\lambda > 0$ and all sufficiently large network sizes, the oracle and feasible kernel denominators are positive with probability at least $1 - C_\lambda \sum_{g\leq G} N_g^{-\lambda}$. On the event with the same probability, the following statements hold:

(i) *The denoising channel error satisfies*

$$R_{\text{denoise}} \leq L_f \max_{g\leq G} \max_{i\neq k, \bar{d}_g^{\text{pseudo}}(i,k)\leq h_N} \|\check{\xi}_{g,i} - \check{\xi}_{g,k}\| + C_\lambda \left[\sqrt{\frac{\log(N_{\min})}{N_{\min} h_N^{d_\xi}}} + \frac{(\log(N_{\min})/N_{\min})^{1/2}}{h_N^{d_\xi+1}}\right],$$

where the first term tends to zero in probability by Lemma A.5 (ii).

(ii) *If in addition Assumption 6 holds, then*

$$R_{\text{denoise}} \leq C_\lambda \left[h_N + \sqrt{\frac{\log(N_{\min})}{N_{\min} h_N^{d_\xi}}} + \frac{(\log(N_{\min})/N_{\min})^{1/2}}{h_N}\right].$$

Proof.

Empirical probability of nearby reference coordinates. Fix $c > 0$, following the arguments in (Deaner et al., 2025, Lemma A.4) and (Zeleneev, 2020, Lemma A.5), we first compare the empirical proportion of reference coordinates near $\check{\xi}_{g,i}$ with its conditional probability. By Assumption 1, conditional on $\check{\xi}_{g,i}$, the indicators $\mathbf{1}\{\|\check{\xi}_{g,i} - \check{\xi}_{g,k}\| \le ch_N\}$ for all $k \neq i$ are independent and have the same conditional mean. With Assumption 5 (ii), for all sufficiently large network sizes, we know that

$$\mathbb{E}\left[\mathbf{1}\{\|\check{\xi}_{g,i} - \check{\xi}_{g,k}\| \le ch_N\} \,\middle|\, \check{\xi}_{g,i}\right] = \check{\mu}(B(\check{\xi}_{g,i}, ch_N) \cap \check{\Xi}) \quad \text{and} \quad \underline{c}_\mu c^{d_\xi} h_N^{d_\xi} \le \check{\mu}(B(\check{\xi}_{g,i}, ch_N) \cap \check{\Xi}) \le \bar{c}_\mu c^{d_\xi} h_N^{d_\xi}.$$

The conditional variance of each indicator is no greater than its conditional mean, and hence, it is bounded by $\bar{c}_\mu c^{d_\xi} h_N^{d_\xi}$. Applying Bernstein's inequality conditionally on $\check{\xi}_{g,i}$, integrating the conditional probability bound, and then taking a union bound over i, that for some $c'' > 0$, it follows that

$$\mathbb{P}\left(\max_{i \le N_g} \left| \frac{1}{N_g - 1} \sum_{k \neq i} \mathbf{1}\{\|\check{\xi}_{g,i} - \check{\xi}_{g,k}\| \le ch_N\} - \check{\mu}(B(\check{\xi}_{g,i}, ch_N) \cap \check{\Xi}) \right| > \frac{1}{2}\underline{c}_\mu c^{d_\xi} h_N^{d_\xi} \right) \le 2N_g \exp(-c'' N_g h_N^{d_\xi}) \le N_g^{-\lambda},$$

for all sufficiently large network sizes. Note that with in (32), and since $N/\log N$ is increasing for large N and $N_g \ge N_{\min}$, we have $N_g h_N^{d_\xi}/\log N_g \to \infty$ and leads to the last inequality.

Replacing the normalization $N_g - 1$ with N_g and excluding $k = j$ changes this proportion by at most $2/N_g = o(h_N^{d_\xi})$, uniformly over i, j. Therefore, there are constants $c', C' > 0$ such that, with probability at least $1 - N_g^{-\lambda}$,

$$c' h_N^{d_\xi} \le \frac{1}{N_g} \sum_{k \notin \{i,j\}} \mathbf{1}\{\|\check{\xi}_{g,i} - \check{\xi}_{g,k}\| \le ch_N\} \le C' h_N^{d_\xi}, \qquad i, j \le N_g. \tag{33}$$

Denominators in the kernel estimator. By Lemma A.5 (i), the oracle kernel weight is at least $\underline{K}$ whenever $\|\check{\xi}_{g,i} - \check{\xi}_{g,k}\| \le t_K h_N / L_f$, and Lipschitz condition of K further provides

$$\max_{i \neq k} \left| K\left(\hat{d}_g^{\text{pseudo}}(i,k)/h_N\right) - K\left(\bar{d}_g^{\text{pseudo}}(i,k)/h_N\right) \right| \le C_\lambda \frac{(\log N_g/N_g)^{1/2}}{h_N}. \tag{34}$$

The right-hand side tends to zero by the third condition in (32). Hence, the feasible kernel weight is at least $\underline{K}/2$ on the same neighborhood for all sufficiently large network sizes. With (33) and set $c = t_K/L_f$, it follows that

$$\begin{aligned} &\frac{1}{N_g} \sum_{k \notin \{i,j\}} K\left(\bar{d}_g^{\text{pseudo}}(i,k)/h_N\right) \ge c h_N^{d_\xi}, \\ &\frac{1}{N_g} \sum_{k \notin \{i,j\}} K\left(\hat{d}_g^{\text{pseudo}}(i,k)/h_N\right) \ge c' h_N^{d_\xi} > 0, \end{aligned} \tag{35}$$

uniformly over i, j.

Decomposition of the denoising error. For $k \notin \{i, j\}$, denote the weights

$$\bar{\alpha}_{g,i,k}(j) := \frac{K\left(\bar{d}_g^{\text{pseudo}}(i,k)/h_N\right)}{\sum_{\ell \notin \{i,j\}} K\left(\bar{d}_g^{\text{pseudo}}(i,\ell)/h_N\right)} \qquad \text{and} \qquad \hat{\alpha}_{g,i,k}(j) := \frac{K\left(\hat{d}_g^{\text{pseudo}}(i,k)/h_N\right)}{\sum_{\ell \notin \{i,j\}} K\left(\hat{d}_g^{\text{pseudo}}(i,\ell)/h_N\right)}.$$

Under the maintained setting, both sets of weights are well defined, nonnegative, and sum to one. By adding and subtracting, we have

$$\begin{aligned}\sum_{k \notin \{i,j\}} \hat{\alpha}_{g,i,k}(j) A^{(w)}_{g,k,j} - P^{(w)}_{g,i,j} = &\sum_{k \notin \{i,j\}} \bar{\alpha}_{g,i,k}(j)(P^{(w)}_{g,k,j} - P^{(w)}_{g,i,j}) \\ &+ \sum_{k \notin \{i,j\}} \bar{\alpha}_{g,i,k}(j)(A^{(w)}_{g,k,j} - P^{(w)}_{g,k,j}) \\ &+ \sum_{k \notin \{i,j\}} (\hat{\alpha}_{g,i,k}(j) - \bar{\alpha}_{g,i,k}(j)) A^{(w)}_{g,k,j}, \end{aligned} \tag{36}$$

where these are the oracle bias, the oracle noise, and the error from replacing the oracle weights.

For the oracle bias, a positive oracle weight requires $\bar{d}_g^{\text{pseudo}}(i,k) \le h_N$. Since the weights sum to one, Lipschitz $\check{f}_g^{(w)}$ provides

$$\left| \sum_{k \notin \{i,j\}} \bar{\alpha}_{g,i,k}(j)(P^{(w)}_{g,k,j} - P^{(w)}_{g,i,j}) \right| \le L_f \max_{\substack{k \neq i \\ \bar{d}_g^{\text{pseudo}}(i,k) \le h_N}} \|\check{\xi}_{g,i} - \check{\xi}_{g,k}\|. \tag{37}$$

For the oracle noise, conditional on the reference coordinates, the oracle weights are fixed. Furthermore, the errors $A^{(w)}_{g,k,j} - P^{(w)}_{g,k,j}$ for all $k \notin \{i, j\}$ are independent, centered, and bounded by one in absolute value. And following (35), boundedness of K implies

$$\max_{k \notin \{i,j\}} \alpha_{g,i,k}(j) \le \frac{C}{N_g h_N^{d_\xi}}, \qquad \sum_{k \notin \{i,j\}} (\alpha_{g,i,k}(j))^2 \le \frac{C}{N_g h_N^{d_\xi}}.$$

Apply Bernstein's inequality conditionally on these reference coordinate realizations, take a union bound over i, j, and integrate over the reference coordinates, then with probability at least $1 - C_\lambda N_g^{-\lambda}$, we have

$$\max_{i,j \le N_g} \left| \sum_{k \notin \{i,j\}} \bar{\alpha}_{g,i,k}(j)(A^{(w)}_{g,k,j} - P^{(w)}_{g,k,j}) \right| \le C_\lambda \sqrt{\frac{\log N_g}{N_g h_N^{d_\xi}}},$$

where the linear term $\log N_g/(N_g h_N^{d_\xi})$ is dominated by the square root term under the second condition in (32).

For the weight difference, adding and subtracting the feasible kernel weights divided by the oracle denominator provides

$$\sum_{k \notin \{i,j\}} \left|\hat{\alpha}_{g,i,k}(j) - \bar{\alpha}_{g,i,k}(j)\right| \le \frac{2N_g^{-1} \sum_{k \notin \{i,j\}} \left|K\left(\hat{d}_g^{\text{pseudo}}(i,k)/h_N\right) - K\left(\bar{d}_g^{\text{pseudo}}(i,k)/h_N\right)\right|}{N_g^{-1} \sum_{k \notin \{i,j\}} K\left(\bar{d}_g^{\text{pseudo}}(i,k)/h_N\right)}. \tag{38}$$

Since the elements in adjacency matrices lie in $[0, 1]$, this also bounds the last term of (36) in absolute value.

Proof of statement (i). The empirical average of the absolute kernel differences in (38) is bounded by their maximum in (34). Dividing by the oracle lower bound in (35) gives

$$\sum_{k\notin\{i,j\}} \left|\hat{\alpha}_{g,i,k}(j) - \bar{\alpha}_{g,i,k}(j)\right| \le C_\lambda \frac{(\log N_g/N_g)^{1/2}}{h_N^{d_\xi+1}}.$$

Combine this with the bias and noise bounds in (36). Taking maxima over the finitely many networks and observed statuses, using $N_g \ge N_{\min}$, and intersecting the events gives the bound in statement (i) with probability at least $1 - C_\lambda \sum_{g\le G} N_g^{-\lambda}$. Lemma A.5 (ii) gives the stated convergence of its bias term.

Proof of statement (ii). Under Assumption 6, the kernel differences are restricted to nearby reference coordinates. A nonzero difference requires either $\bar{d}_g^{\text{pseudo}}(i,k) \le h_N$ or $\hat{d}_g^{\text{pseudo}}(i,k) \le h_N$. With Lemma A.5 (iii), its bounds and the third condition in (32) imply

$$\bar{d}_g^{\text{pseudo}}(i,k) \le 2h_N \qquad \text{and} \qquad \|\check{\xi}_{g,i} - \check{\xi}_{g,k}\| \le 16h_N/c_1.$$

Hence, the empirical average of the absolute kernel differences is bounded by their maximum multiplied by the empirical probability of this neighborhood. Plug (34) and (35) into (38), followed by the upper bound in (33) with $c = 16/c_1$, we have

$$\begin{aligned}\sum_{k\notin\{i,j\}} \left|\hat{\alpha}_{g,i,k}(j) - \bar{\alpha}_{g,i,k}(j)\right| &\le C_\lambda \frac{(\log N_g/N_g)^{1/2}}{h_N^{d_\xi+1}} \frac{1}{N_g} \sum_{k\notin\{i,j\}} \mathbf{1}\{\|\check{\xi}_{g,i} - \check{\xi}_{g,k}\| \le 16h_N/c_1\} \\ &\le C_\lambda \frac{(\log N_g/N_g)^{1/2}}{h_N}.\end{aligned}$$

For the oracle bias, since the positive weight indicates $\bar{d}_g^{\text{pseudo}}(i,k) \le h_N$, we can bounds the reference coordinate difference by $8h_N/c_1$ as well. Thus (37) is at most $8L_f h_N/c_1$. Adding the noise bound, and taking the maxima proves statement (ii).

□

We are now ready to prove Proposition 5.

Proof.

Proof of statement (i). Note that the bandwidth conditions in Proposition 5 imply the three common bandwidth conditions of Lemma A.6. On top of that, with Lemma A.5 (ii), it follows that the oracle bias term in Lemma A.6 (i) vanish in probability. Also, the remaining two terms vanish by the stated bandwidth conditions. Hence $R_{\text{denoise}} = o_p(1)$.

Proof of statement (ii). Apply Lemma A.6 (ii). With $h_N \leq 1$ eventually, will have

$$\sqrt{\frac{\log(N_{\min})}{N_{\min} h_N^{d_\xi}}} + \frac{(\log(N_{\min})/N_{\min})^{1/2}}{h_N} \leq 2\frac{(\log(N_{\min})/N_{\min})^{1/2}}{h_N^{\max\{d_\xi/2,1\}}}.$$

Together with the bias term h_N, this proves the first bound in statement (ii). Balancing with h_N gives the bandwidth.

□

A.3 Proof for the Large Sample Theory in Section 5: Proposition 6

We separate the proof for Proposition 6 into following steps:

- We first try to answer the question of, if we have the population objects, what would an optimal Gromov-Wasserstein coupling look like and what can the coupling recover? *In this step, we will see that it matches reference coordinates correctly up to a target symmetry.*
- However, since the estimated within-network graphon distance matrices differ from their population counterparts, and empirical reference coordinate distributions differ from $\check{\mu}$, can the finite sample Gromov-Wasserstein optimization problem be close to the population problem? *As long as the differences between finite sample and population counterparts are small, the estimated object can still be close to it population counterparts.*
- Now, we know that an estimated coupling has almost the correct Gromov-Wasserstein objective. The next question is that does it actually match similar reference coordinates $\check{\xi}$? *In this step, we will find that the small Gromov-Wasserstein distortion will lead to small mean-square reference coordinate mismatch. These three steps will be discussed in Lemma A.7.*
- We now have the good average alignment with small mean-square reference coordinate mismatch, but how can we translate this small average mismatch to the argument of the worst case mismatch being small. *In which, this will be tackled in Lemma A.8.*
- Lastly, we translate the result of the worst case reference coordinate misalignment to projection alignment. *With small reference coordinate misalignment, the projected networks are accurate.*

Throughout the derivation, we use the reference representation of Section 5. Let W_1 and W_2 denote the 1-Wasserstein and quadratic Wasserstein distances on $\check{\Xi}$, and define

$$\begin{aligned}
\nu_g &:= \frac{1}{N_g}\sum_{i=1}^{N_g} \text{Dirac}_{\check{\xi}_{g,i}}, \\
s_N &:= \max_{g\leq G} W_2^2(\nu_g, \check{\mu}), \\
e_N &:= \max_{g\leq G}\max_{i,j\leq N_g} \left|\hat{d}_g^2(i,j) - \check{d}_g^2(\check{\xi}_{g,i}, \check{\xi}_{g,j})\right|.
\end{aligned} \tag{39}$$

Here, ν_g is the empirical distribution of sampled reference coordinates in network g, s_N measures the largest discrepancy between the empirical reference coordinate distributions and the unobserved population reference measure $\check{\mu}$, and e_N measures the largest error in the estimated internal distances used by the Gromov-Wasserstein matching step. Whenever a finite sample estimated coupling appears in an integral with population costs, we also regard it as the discrete probability measure

$$\hat{\pi}^{g,G} = \sum_{k \le N_g} \sum_{i \le N_G} \hat{\pi}^{g,G}_{k,i} \mathrm{Dirac}_{(\check{\xi}_{g,k}, \check{\xi}_{G,i})} \in \Pi(\nu_g, \nu_G).$$

The uniform marginals of its index matrix remain $1/N_g$ and $1/N_G$.

With this, Lemma A.7 shows that when the empirical reference coordinate distributions and estimated internal distances are close to their population counterparts, any global Gromov-Wasserstein minimizer must align source and target reference coordinates closely on average, up to a symmetry of the target. And under Assumption 6, it also provides an explicit bound for this average alignment error in terms of e_N and s_N.

Lemma A.7 (From Gromov-Wasserstein to average reference coordinate alignment)
Suppose Assumptions 3-5 hold. Fix a source $g < G$ and let $\hat{\pi}^{g,G}$ be any global minimizer of finite sample Gromov-Wasserstein objective $F_{\hat{d}_g^2, \hat{d}_G^2}$. Then, the followings hold:

(*i*) *As $e_N + s_N \to 0$,*

$$\inf_{T \in \mathcal{T}} \left\{ \int \|T\check{\xi} - \check{\zeta}\|^2 \mathrm{d}\hat{\pi}^{g,G}(\check{\xi}, \check{\zeta}) + W_2^2(T_\# \nu_g, \check{\mu}) \right\} \to 0.$$

(*ii*) *If in addition Assumption 6 holds, there exists $C < \infty$ and $t_0 > 0$, such that, whenever $e_N + s_N \le t_0$, a symmetry $T \in \mathcal{T}$ can be selected and leads to*

$$\int \|T\check{\xi} - \check{\zeta}\|^2 \mathrm{d}\hat{\pi}^{g,G}(\check{\xi}, \check{\zeta}) \le C(e_N^2 + s_N) \qquad \textit{and} \qquad W_2^2(T_\# \nu_g, \check{\mu}) \le C s_N.$$

Proof.

The population optimum matches reference coordinates correctly, up to a symmetry. Proposition 3 already shows that every population optimal reference coupling preserves $\check{d}_G^2$ almost surely. Under Assumption 5 (iv), $\check{d}_G$ is a metric on $\check{\Xi}$. Continuity and full support therefore let us apply Mémoli (2011, Lemma 2.2, Lemma 10.4 and proof of Theorem 5.1(a)). That is, every population optimal coupling is supported on the graph of a measure-preserving isometry $T \in \mathcal{T}$. Conversely, every $(\mathrm{id}, T)_\# \check{\mu}$ has the same objective value as the diagonal coupling. Hence

$$\arg\min_{\check{\pi} \in \Pi(\check{\mu}, \check{\mu})} F_{\check{d}_g^2, \check{d}_G^2}(\check{\pi}) = \{(\mathrm{id}, T)_\# \check{\mu} : T \in \mathcal{T}\}. \tag{40}$$

Proposition 3 and Assumption 4 (ii) also imply that these symmetries preserve every reference graphon almost everywhere. Continuity and full support make this pointwise, and therefore, for every $g' \le G$ and $T \in \mathcal{T}$, we have

$$\check{f}^{(w)}_{g'}(T\check{\xi}, T\check{\zeta}) = \check{f}^{(w)}_{g'}(\check{\xi}, \check{\zeta}) \qquad \text{and} \qquad \check{d}^2_{g'}(T\check{\xi}, T\check{\zeta}) = \check{d}^2_{g'}(\check{\xi}, \check{\zeta}). \tag{41}$$

Finally, we record that $\mathcal{T}$ is compact for uniform convergence: every sequence in $\mathcal{T}$ has a subsequence converging uniformly to a map in $\mathcal{T}$. By Lemma A.4 (i), $\check{d}_G$ is a metric on $\check{\Xi}$ giving the Euclidean topology, so $(\check{\Xi}, \check{d}_G)$ is a compact metric space, and its isometries form a compact set for uniform convergence. Within it, the maps with $T_\# \check{\mu} = \check{\mu}$ form a closed subset, since $T_n \to T$ uniformly gives

$$\int \phi(T\check{\xi}) \mathrm{d}\check{\mu}(\check{\xi}) = \lim_{n\to\infty} \int \phi(T_n\check{\xi}) \mathrm{d}\check{\mu}(\check{\xi}) = \int \phi \mathrm{d}\check{\mu}$$

for every continuous ϕ. A closed subset of a compact set is compact.

Comparing the finite and population problems. What we first need is an upper bound on the population objective evaluated at the estimated coupling. Take a W_2-optimal coupling between ν_g and ν_G. By the triangle inequality for W_2, we have

$$W_2(\nu_g, \nu_G) \le W_2(\nu_g, \check{\mu}) + W_2(\nu_G, \check{\mu}) \le 2\sqrt{s_N}.$$

This coupling has mean squared reference coordinate mismatch at most $4s_N$. Since the internal distances are Lipschitz, replacing its target reference coordinates by its source reference coordinates changes the objective by at most $C\sqrt{s_N}$. The resulting objective is an empirical average of

$$\left[\check{d}^2_g(\check{\xi}, \check{\xi}') - \check{d}^2_G(\check{\xi}, \check{\xi}')\right]^2,$$

whose difference from its population average under $\check{\mu} \otimes \check{\mu}$ is also at most $C\sqrt{s_N}$, as $W_1(\nu_g, \check{\mu}) \le W_2(\nu_g, \check{\mu}) \le \sqrt{s_N}$. Therefore there exists a feasible finite coupling such that

$$F_{\check{d}^2_g, \check{d}^2_G}(\pi) \le F_{\check{d}^2_g, \check{d}^2_G}(\check{\pi}^{\mathrm{diag}}) + C\sqrt{s_N}.$$

Replacing true internal distances by the estimated version changes the objective of every finite coupling by at most Ce_N when $e_N \le 1$. Since $\hat{\pi}^{g,G}$ minimizes the estimated objective, it follows that

$$F_{\check{d}^2_g, \check{d}^2_G}(\hat{\pi}^{g,G}) \le F_{\check{d}^2_g, \check{d}^2_G}(\check{\pi}^{\mathrm{diag}}) + C(e_N + \sqrt{s_N}). \tag{42}$$

Proof of statement (i). We prove this statement by contradiction. Suppose the conclusion fails with $e_N + s_N \to 0$. Since $\check{\Xi} \times \check{\Xi}$ is compact, a subsequence of $\hat{\pi}^{g,G}$ converges weakly to some probability measure π^*. Note that since the marginals of $\hat{\pi}^{g,G}$ are ν_g and ν_G, while $s_N \to 0$ implies $\nu_g \Rightarrow \check{\mu}$ and $\nu_G \Rightarrow \check{\mu}$, the marginals also pass to the limit. Hence, $\pi^* \in \Pi(\check{\mu}, \check{\mu})$.

By the continuity of the population objective and (42), we have

$$F_{\check{d}_g^2, \check{d}_G^2}(\pi^*) \le F_{\check{d}_g^2, \check{d}_G^2}(\check{\pi}^{\mathrm{diag}}).$$

Since the diagonal coupling is population optimal, hence the equality holds and π^* is optimal. Then by (40), for some $T \in \mathcal{T}$, we have

$$\pi^* = (\mathrm{id}, T)_{\#}\check{\mu}.$$

Since $\|T\check{\xi} - \check{\zeta}\|^2$ is bounded and continuous, it follows that

$$\int \|T\check{\xi} - \check{\zeta}\|^2 \mathrm{d}\hat{\pi}^{g,G} \to 0.$$

Also, since $\nu_g \to \check{\mu}$, we have $T_{\#}\nu_g \to T_{\#}\check{\mu} = \check{\mu}$. with the compactness of $\check{\Xi}$, it then leads to

$$W_2^2(T_{\#}\nu_g, \check{\mu}) \to 0.$$

This contradicts the supposed failure and completes the proof of statement (i).

Rewriting the objective as a quadratic in the source costs. We then want to turn the optimality of $\hat{\pi}^{g,G}$ into a quadratic bound on the mismatch between source and target internal distances., and to do so, we start by proving an auxiliary result. Let b_g be the continuous nondecreasing function in Lemma A.4 (i), and define the Bregman divergence by

$$B_g(a, v) := \int_v^a [b_g(t) - b_g(v)]\mathrm{d}t.$$

It is nonnegative and vanishes exactly when $b_g(a) = b_g(v)$. With some derivation, we have

$$[a - b_g(v)]^2 = 2B_g(a, v) + \left[a^2 - 2\int_0^a b_g(t)\mathrm{d}t\right] + \left[b_g(v)^2 + 2\int_0^v b_g(t)\mathrm{d}t - 2vb_g(v)\right]. \tag{43}$$

The second term depends only on the source pair and the third only on the target pair. Since every feasible coupling has the same source and target marginals, these two terms cancel when two coupling objectives are compared. Hence, only the Bregman term depends on how source and target reference coordinates are matched.

Proof of statement (ii). Note that the function b_g is originally defined on the range $\check{d}_g^2(\check{\Xi}^2)$. Since $\hat{d}_G^2$ may fall slightly outside the corresponding target range, extend b_g linearly beyond both endpoints of its domain with slope $\underline{c}_b$. The extended function remains increasing with slopes between $\underline{c}_b$ and $\bar{c}_b$, and therefore, this leaves the population relation unchanged and makes b_g^{-1} well defined on $\mathbb{R}$. Assumption 6 (i) then gives

$$\left|b_g^{-1}(\hat{d}_G^2(i,j)) - \check{d}_g^2(\check{\xi}_{G,i}, \check{\xi}_{G,j})\right| \le \frac{e_N}{\underline{c}_b} \qquad \text{and} \qquad \frac{\underline{c}_b}{2}(a-v)^2 \le B_g(a,v) \le \frac{\bar{c}_b}{2}(a-v)^2. \tag{44}$$

Next, let π be a W_2-optimal coupling between ν_g and ν_G. By the triangle inequality, we have

$$\int \|\check{\xi} - \check{\zeta}\|^2 \mathrm{d}\pi(\check{\xi}, \check{\zeta}) = W_2^2(\nu_g, \nu_G) \le \left[W_2(\nu_g, \check{\mu}) + W_2(\nu_G, \check{\mu})\right]^2 \le 4 s_N. \tag{45}$$

We then want to compare this π with the estimated coupling. Apply (43) with $a = \hat{d}_g^2(k,\ell)$ and $v = b_g^{-1}(\hat{d}_G^2(i,j))$. Since $b_g(v) = \hat{d}_G^2(i,j)$, the left-hand side is the estimated Gromov-Wasserstein loss. On the right-hand side, notice that the second term depends only on the source marginals and the third term depends only on the target marginals. Indeed, for every feasible coupling π',

$$\sum_{i,j} \pi'_{k,i}\pi'_{\ell,j} = \frac{1}{N_g^2}, \qquad \sum_{k,\ell} \pi'_{k,i}\pi'_{\ell,j} = \frac{1}{N_G^2}.$$

Hence these terms can be canceled when comparing two feasible couplings using (43). Optimality of $\hat{\pi}^{g,G}$ implies that the B_g term under $\hat{\pi}^{g,G}$ is no larger than under π. Using the lower bound in (44) on the $\hat{\pi}^{g,G}$ and the upper bound on the π leads to

$$F_{\hat{d}_g^2, b_g^{-1}\circ\hat{d}_G^2}(\hat{\pi}^{g,G}) \le \frac{\bar{c}_b}{\underline{c}_b} F_{\hat{d}_g^2, b_g^{-1}\circ\hat{d}_G^2}(\pi).$$

On the other hand, note that since π is a W_2-optimal coupling between ν_g and ν_G, the right-hand side is small because π matches nearby reference coordinates. Then for two independent pairs $(\check{\xi}, \check{\zeta})$ and $(\check{\xi}', \check{\zeta}')$ drawn from π, the definition of e_N, (44), and Lipschitz continuity of $\check{d}_g^2$ give

$$\left|\hat{d}_g^2(\check{\xi}, \check{\xi}') - b_g^{-1}\left(\hat{d}_G^2(\check{\zeta}, \check{\zeta}')\right)\right| \le C e_N + C\left(\|\check{\xi} - \check{\zeta}\| + \|\check{\xi}' - \check{\zeta}'\|\right).$$

Therefore, by (45),

$$F_{\hat{d}_g^2, b_g^{-1}\circ\hat{d}_G^2}(\pi) \le C(e_N^2 + s_N),$$

and hence

$$F_{\hat{d}_g^2, b_g^{-1}\circ\hat{d}_G^2}(\hat{\pi}^{g,G}) \le C(e_N^2 + s_N). \tag{46}$$

The next step is to express this bound in the target geometry. By the definition of e_N and (44), we have

$$\left|\check{d}_g^2(\check{\xi}_{g,k}, \check{\xi}_{g,\ell}) - \check{d}_g^2(\check{\xi}_{G,i}, \check{\xi}_{G,j})\right| \le \left|\hat{d}_g^2(k,\ell) - b_g^{-1}(\hat{d}_G^2(i,j))\right| + C e_N.$$

Squaring and averaging under $\hat{\pi}^{g,G} \otimes \hat{\pi}^{g,G}$, together with (46), gives

$$F_{\check{d}_g^2, \check{d}_g^2}(\hat{\pi}^{g,G}) \le C(e_N^2 + s_N).$$

Then with $\check{d}_G^2 = b_g \circ \check{d}_g^2$ and Assumption 6 (i), it follows that

$$F_{\check{d}_G^2, \check{d}_G^2}(\hat{\pi}^{g,G}) \le \bar{c}_b^2 F_{\check{d}_g^2, \check{d}_g^2}(\hat{\pi}^{g,G}) \le C(e_N^2 + s_N). \tag{47}$$

What remains is that Assumption 6 (ii) applies only to couplings in $\Pi(\check{\mu}, \check{\mu})$, while $\hat{\pi}^{g,G}$ has marginals ν_g and ν_G. Thus, we replace the marginals once. Take W_2-optimal couplings of ν_g and ν_G with $\check{\mu}$. Conditional on $(\check{\xi}, \check{\zeta}) \sim \hat{\pi}^{g,G}$, draw the corresponding population reference coordinates $(\check{u}, \check{v})$. Their joint law $\check{\pi}_N$ belongs to $\Pi(\check{\mu}, \check{\mu})$ and satisfies

$$\mathbb{E}\left[\|\check{\xi} - \check{u}\|^2 + \|\check{\zeta} - \check{v}\|^2\right] \le 2 s_N. \tag{48}$$

This follows from disintegration of the two Wasserstein couplings (Ambrosio et al., 2008, Theorem 5.3.1); see also Chowdhury and Mémoli (2019, proof of Theorem 2.3). Since $\check{d}_G^2$ is Lipschitz, by Minkowski's inequality, we have

$$F_{\check{d}_G^2, \check{d}_G^2}(\check{\pi}_N)^{1/2} \le F_{\check{d}_G^2, \check{d}_G^2}(\hat{\pi}^{g,G})^{1/2} + C\sqrt{s_N}.$$

Together with (47), we have

$$F_{\check{d}_G^2, \check{d}_G^2}(\check{\pi}_N) \le C(e_N^2 + s_N).$$

We can now apply Assumption 6 (ii). Assumption 6 (ii) and compactness of $\mathcal{T}$ provide a symmetry $T \in \mathcal{T}$ such that

$$\int \|T\check{u} - \check{v}\|^2 \mathrm{d}\check{\pi}_N(\check{u}, \check{v}) \le C(e_N^2 + s_N).$$

Finally, we carry this symmetry back to the sampled reference coordinates. Lemma A.4 (iii) gives the first inequality below; for the second, the triangle inequality for $\check{d}_G$ and $\check{d}_G \le 1$ leads to

$$c_1 \|\check{\xi} - \check{\zeta}\| \le \|\check{d}_G^2(\check{\xi}, \cdot) - \check{d}_G^2(\check{\zeta}, \cdot)\|_{L^2(\check{\mu})} \le 2\check{d}_G(\check{\xi}, \check{\zeta}).$$

Using invariance of $\check{d}_G$ and the upper bound in Lemma A.4 (i),

$$\|T\check{\xi} - T\check{\zeta}\| \le \frac{2L_f}{c_1}\|\check{\xi} - \check{\zeta}\|, \qquad T \in \mathcal{T}.$$

Using (48) and the squared triangle inequality,

$$\int \|T\check{\xi} - \check{\zeta}\|^2 \mathrm{d}\hat{\pi}^{g,G} \le C\left[\mathbb{E}\|\check{\xi} - \check{u}\|^2 + \int \|T\check{u} - \check{v}\|^2 \mathrm{d}\check{\pi}_N + \mathbb{E}\|\check{v} - \check{\zeta}\|^2\right]$$
$$\le C(e_N^2 + s_N).$$

Applying (T, T) to a W_2-optimal coupling of ν_g and $\check{\mu}$ and using $T_{\#}\check{\mu} = \check{\mu}$ gives

$$W_2^2(T_{\#}\nu_g, \check{\mu}) \le \left(\frac{2L_f}{c_1}\right)^2 W_2^2(\nu_g, \check{\mu}) \le C s_N.$$

Therefore, the same symmetry T satisfies both bounds in statement (ii).

□

Lemma A.8 (From average to uniform reference coordinate alignment)
Suppose Assumptions 3-5 hold. Fix a source $g < G$ and let $\hat{\pi}^{g,G}$ be any global optimizer of $F_{\hat{d}_g^2, \hat{d}_G^2}$. For any fixed $T \in \mathcal{T}$, define

$$\epsilon_g(T) := \max_{(k,i):\, \hat{\pi}^{g,G}_{k,i} > 0} \|T\check{\xi}_{g,k} - \check{\xi}_{G,i}\|,$$

$$\theta_{g,N} := \int \|T\check{\xi} - \check{\zeta}\|^2 \mathrm{d}\hat{\pi}^{g,G}(\check{\xi}, \check{\zeta}) + W_2^2(T_{\#}\nu_g, \check{\mu}) + W_2^2(\nu_G, \check{\mu}).$$

With this, it follows that

(i) *As $e_N + \theta_{g,N} \to 0$,*

$$\epsilon_g(T) \to 0.$$

(ii) *If in addition Assumption 6 holds, there exists $C < \infty$ and $t_0 > 0$, such that, whenever $e_N + \theta_{g,N} \le t_0$,*

$$\epsilon_g(T) \le C\left[e_N + \theta_{g,N}^{1/(d_\xi+2)}\right].$$

Proof.

Proof of statement (i). The idea of this proof is as following: We prove it by contradiction. Suppose there were still one badly matched positive pair, optimality would let us compare it with a chain of many nearly correct pairs. The bad match has a strictly positive population cost, while the total cost of the increasingly short chain goes to zero. Contradiction.

We first define a population cost for being wrongly matched. Using B_g from the proof of Lemma A.7, define

$$\check{r}_g(\check{\xi}, \check{\zeta}) := 2\int_{\check{\Xi}} B_g\Big(\check{d}_g^2(\check{\xi}, \check{u}), \check{d}_g^2(\check{\zeta}, \check{u})\Big)\, \mathrm{d}\check{\mu}(\check{u}).$$

By (43), this differs from the population linearized cost $\int[\check{d}_g^2(\check{\xi}, \check{u}) - \check{d}_G^2(\check{\zeta}, \check{u})]^2 \mathrm{d}\check{\mu}(\check{u})$ only by terms depending separately on $\check{\xi}$ and $\check{\zeta}$. These terms cancel when target reference coordinates are permuted.

Note that the $\check{r}_g(\check{\xi}, \check{\zeta})$ is continuous, nonnegative, and zero exactly on the diagonal. This is because, for distinct reference coordinates, its integrand at $\check{u} = \check{\xi}$ is positive as $b_g(\check{d}_g^2(\check{\zeta}, \check{\xi})) = \check{d}_G^2(\check{\zeta}, \check{\xi}) > b_g(0)$, and continuity and full support make the integral positive. The integral defining B_g, uniform continuity of b_g, and Lipschitz continuity of $\check{d}_g^2$ also leads to

$$\check{r}_g(\check{\xi}, \check{\zeta}) = o(\|\check{\xi} - \check{\zeta}\|), \qquad \check{\xi}, \check{\zeta} \in \check{\Xi}. \tag{49}$$

We now prove statement (i) by contradiction. Suppose $e_N + \theta_{g,N} \to 0$ while $\epsilon_g(T)$ stays bounded away from zero, that is, we assume average alignment becomes perfect, but some positive coupling entry remains a fixed distance away. The definition of $\theta_{g,N}$ implies

$$(T, \mathrm{id})_{\#}\hat{\pi}^{g,G} \to \check{\pi}^{\mathrm{diag}}$$

where the target marginal converges to $\check{\mu}$, and the mean squared distance between the two transformed coordinates tends to zero. Passing to a subsequence, there exists $\varepsilon > 0$ such that $\epsilon_g(T) \geq \varepsilon$. For each configuration, choose a positive entry (k_N, i_N) attaining $\epsilon_g(T)$. Since

$$(T\check{\xi}_{g,k_N}, \check{\xi}_{G,i_N}) \in \check{\Xi} \times \check{\Xi}$$

and $\check{\Xi} \times \check{\Xi}$ is compact, those bad pairs have a subsequence satisfies

$$T\check{\xi}_{g,k_N} \to \check{\xi}, \qquad \check{\xi}_{G,i_N} \to \check{\zeta}.$$

Moreover, since

$$\|\check{\xi} - \check{\zeta}\| = \lim_N \epsilon_g(T) \geq \varepsilon,$$

it follows that $\check{\xi} \neq \check{\zeta}$.

Fix some $M \geq 2$ and divide the segment from $\check{\xi}$ to $\check{\zeta}$ into equal pieces, with endpoints $\check{z}_0 = \check{\xi}, \ldots, \check{z}_M = \check{\zeta}$. Fix an interior point $\check{z}_j$. Since $\check{\mu}$ has full support,

$$\check{\pi}^{\mathrm{diag}}\Big(B(\check{z}_j, \delta)^2\Big) = \check{\mu}(B(\check{z}_j, \delta)) > 0$$

for every $\delta > 0$. Because $(T, \mathrm{id})_{\#}\hat{\pi}^{g,G} \to \check{\pi}^{\mathrm{diag}}$, for all large N, this neighborhood contains a positive atom of $(T, \mathrm{id})_{\#}\hat{\pi}^{g,G}$. Hence, we can choose positive entries $(k_{j,N}, i_{j,N})$ such that

$$T\check{\xi}_{g,k_{j,N}} \to \check{z}_j, \qquad \check{\xi}_{G,i_{j,N}} \to \check{z}_j.$$

Let (k_0, i_0) be the distant positive pair and (k_j, i_j), $j = 1, \ldots, M-1$, the positive pairs near the interior points. Reassign

$$(k_0, i_0), (k_1, i_1), \ldots, (k_{M-1}, i_{M-1}) \quad \to \quad (k_0, i_1), (k_1, i_2), \ldots, (k_{M-1}, i_0).$$

This preserves both marginals and replaces one long match by M short ones. By (41), $e_N \to 0$, and continuity of the population costs, the finite linearized cycle inequality passes to the limit. Canceling the single-coordinate terms gives

$$0 < \check{r}_g(\check{\xi}, \check{\zeta}) \le \sum_{j=0}^{M-1} \check{r}_g(\check{z}_j, \check{z}_{j+1}) \le M \max_{\check{u},\check{v}\in\check{\Xi};\ \|\check{u}-\check{v}\|\le\|\check{\xi}-\check{\zeta}\|/M} \check{r}_g(\check{u}, \check{v}).$$

Letting $M \to \infty$ contradicts (49). This proves statement (i), uniformly over configurations and the chosen symmetry.

Proof of statement (ii). Here we use the same idea as statement (i): compare the worst positive match with a chain of shorter positive matches. Under Assumption 6, the Bregman cost now gives a quadratic measure of the reference coordinate mismatch, which makes this comparison quantitative.

Using the extension of b_g and the quadratic bounds on B_g from the proof of Lemma A.7 (ii), define

$$\hat{r}^{g,G}_{k,i} := 2 \sum_{a\le N_g} \sum_{b\le N_G} B_g\Big(\hat{d}^2_g(k,a), b_g^{-1}(\hat{d}^2_G(i,b))\Big)\hat{\pi}^{g,G}_{a,b}.$$

By (43), the remaining terms depend only on the marginals. Hence the same target reassignment used in statement (i) and first-order optimality give, for positive entries (k_j, i_j) and $i_M = i_0$,

$$\sum_{j=0}^{M-1} \hat{r}^{g,G}_{k_j,i_j} \le \sum_{j=0}^{M-1} \hat{r}^{g,G}_{k_j,i_{j+1}}. \tag{50}$$

We next relate $\hat{r}^{g,G}_{k,i}$ to the reference coordinate displacement. By the reverse triangle inequality in L^2, the definition of e_N, (44), invariance (41), and the definition of $\theta_{g,N}$,

$$\left| \left(\sum_{a,b} \left|\hat{d}^2_g(k,a) - b_g^{-1}(\hat{d}^2_G(i,b))\right|^2 \hat{\pi}^{g,G}_{a,b} \right)^{1/2} - \left\|\check{d}^2_g(T\check{\xi}_{g,k}, \cdot) - \check{d}^2_g(\check{\xi}_{G,i}, \cdot)\right\|_{L^2(\check{\mu})} \right| \le C\left(e_N + \sqrt{\theta_{g,N}}\right).$$

Indeed, removing the two estimation errors costs Ce_N, replacing the target partner by the transformed source partner costs $C\sqrt{\theta_{g,N}}$, and replacing $T_\#\nu_g$ by $\check{\mu}$ costs another $C\sqrt{\theta_{g,N}}$.

The quadratic bounds on B_g, Lemma A.4 (iii), and the Lipschitz upper bound therefore imply

$$c_0\left[\|T\check{\xi}_{g,k} - \check{\xi}_{G,i}\| - C\delta_N\right]_+^2 \le \hat{r}^{g,G}_{k,i} \le C_0\left[\|T\check{\xi}_{g,k} - \check{\xi}_{G,i}\| + C\delta_N\right]^2, \tag{51}$$

where

$$\delta_N := e_N + \sqrt{\theta_{g,N}}.$$

What remains is to find the short positive matches quantitatively. If $\theta_{g,N} = 0$, the conclusion is immediate. Otherwise set

$$r = \tilde{c}\theta_{g,N}^{1/(d_{\check{\xi}}+2)}.$$

For any $\check{z} \in \check{\Xi}$, the target marginal and Markov's inequality give

$$\hat{\pi}^{g,G}\left\{(\check{\xi},\check{\zeta}) : \|\check{\zeta}-\check{z}\| \le r, \|T\check{\xi}-\check{\zeta}\| \le r\right\} \ge \check{\mu}(B(\check{z},r/2)\cap\check{\Xi}) - \frac{5\theta_{g,N}}{r^2} \ge cr^{d_{\check{\xi}}} - \frac{5\theta_{g,N}}{r^2}.$$

Choose $\tilde{c}$ such that $c\tilde{c}^{d_{\check{\xi}}+2} > 5$. For sufficiently small $\theta_{g,N}$, the right-hand side is positive. Thus every $\check{z} \in \check{\Xi}$ has a positive pair whose target lies within r of $\check{z}$ and whose displacement is at most r.

Now choose a positive pair attaining $\epsilon_g(T)$. If $\epsilon_g(T) \le r$, the desired bound already holds. Otherwise divide the segment between its transformed source and target into M equal pieces and, exactly as in the proof of statement (i), choose positive pairs near the $M-1$ interior points and cyclically reassign their targets. Each new match then has distance at most

$$\frac{\epsilon_g(T)}{M} + 3r.$$

Since $\hat{r}^{g,G} \ge 0$, (50) and (51) give

$$c_0\left[\epsilon_g(T) - C\delta_N\right]_+^2 \le C_0 M\left[\frac{\epsilon_g(T)}{M} + 3r + C\delta_N\right]^2.$$

Choose the fixed integer M sufficiently large. Taking square roots and absorbing the term proportional to $\epsilon_g(T)/\sqrt{M}$ gives

$$\epsilon_g(T) \le C(r + \delta_N).$$

Finally, when $\theta_{g,N} \le 1$,

$$\sqrt{\theta_{g,N}} \le \theta_{g,N}^{1/(d_{\check{\xi}}+2)},$$

so

$$\epsilon_g(T) \le C\left[e_N + \theta_{g,N}^{1/(d_{\check{\xi}}+2)}\right].$$

Under the short-path alternative in Assumption 5, use the same construction along a path of length at most L_Ξ times the endpoint distance; only the constants change. This proves statement (ii). □

We now prove Proposition 6 by applying Lemmas A.7 and A.8 to the estimated couplings.

Proof.

From reference coordinate error to projection error. We first show how ϵ in Lemma A.8 controls the projected probabilities. Fix a source $g < G$ and $T \in \mathcal{T}$. With the the invariance established in (41) and the Lipschitz, we have

$$\begin{aligned}|\check{f}_g^{(w)}(\check{\xi}_{g,k}, \check{\xi}_{g,\ell}) - \check{f}_g^{(w)}(\check{\xi}_{G,i}, \check{\xi}_{G,j})| &\le L_f(\|T\check{\xi}_{g,k} - \check{\xi}_{G,i}\| + \|T\check{\xi}_{g,\ell} - \check{\xi}_{G,j}\|) \\ &\le 2L_f\epsilon_g(T)\end{aligned}$$

whenever $\hat{\pi}_{k,i}^{g,G}\hat{\pi}_{\ell,j}^{g,G} > 0$. By (Peyré et al., 2016, Proposition 3), we know that the projection weights are nonnegative and sum to one. Then averaging this bound, taking the infimum over T, and then the maximum over sources leads to $R_{\text{align}} \le 2L_f\epsilon$.

The two sampling errors. We next bound the inputs to Lemmas A.7 and A.8. For every fixed $\lambda > 0$ and all sufficiently large network sizes, with probability at least $1 - C_\lambda \sum_{g' \le G} N_{g'}^{-\lambda}$, we want to show

$$s_N \le C_\lambda(\log(N_{\min}))^{3/2} N_{\min}^{-1/\max\{d_{\check{\xi}},2\}} \qquad \text{and} \qquad e_N \le 4R_{\text{denoise}} + 4R_{\text{denoise}}^2 + C_\lambda\left(\frac{\log(N_{\min})}{N_{\min}}\right)^{1/2}. \tag{52}$$

To obtain the first bound, fix a network g'. Since every pair in $\check{\Xi}$ satisfies $\|\check{\xi} - \check{\zeta}\|^2 \le \operatorname{diam}(\check{\Xi})\|\check{\xi} - \check{\zeta}\|$, we have $W_2^2(\nu_{g'}, \check{\mu}) \le \operatorname{diam}(\check{\Xi})\, W_1(\nu_{g'}, \check{\mu})$, and it suffices to bound W_1. Since the reference coordinates are $N_{g'}$ independent draws from $\check{\mu}$, which has compact support in $\mathbb{R}^{d_{\check{\xi}}}$, with Fournier and Guillin (2015, Theorem 1), we have

$$\mathbb{E}\, W_1(\nu_{g'}, \check{\mu}) \le C \log\left(N_{g'}\right) N_{g'}^{-1/\max\{d_{\check{\xi}},2\}}.$$

Replacing one sampled reference coordinate $\check{\xi}$ by $\check{\xi}'$ changes $\nu_{g'}$ by moving mass $1/N_{g'}$ from $\check{\xi}$ to $\check{\xi}'$, so the two empirical measures are at W_1-distance at most $\operatorname{diam}(\check{\Xi})/N_{g'}$. Then by the triangle inequality $W_1(\nu_{g'}, \check{\mu})$ changes by at most the same amount. With McDiarmid's inequality, we have, with probability at least $1 - N_{g'}^{-\lambda}$,

$$W_1(\nu_{g'}, \check{\mu}) \le \mathbb{E}\, W_1(\nu_{g'}, \check{\mu}) + \operatorname{diam}(\check{\Xi})\sqrt{\frac{\lambda \log N_{g'}}{2N_{g'}}},$$

where the second term is of no larger order than the first. A union bound over the G networks and $N_{g'} \ge N_{\min}$ give the first bound in (52).

To obtain the second bound, let $\tilde{d}^2(i,j)$ be $\hat{d}_{g'}^2(i,j)$ with the true probabilities $P_{g'}^{(0)}$ in place of $\hat{P}_{g'}^{(0)}$, and split the error into $|\hat{d}_{g'}^2 - \tilde{d}^2| + |\tilde{d}^2 - \check{d}_{g'}^2|$. For the first part, since $0 \le P_{g',i,\ell}^{(0)} \le 1$ and $|\hat{P}_{g',i,\ell}^{(0)} - P_{g',i,\ell}^{(0)}| \le R_{\text{denoise}}$, writing $a^2 - b^2 = (a-b)(a+b)$ gives, for each ℓ,

$$\left|(\hat{P}_{g',i,\ell}^{(0)} - \hat{P}_{g',j,\ell}^{(0)})^2 - (P_{g',i,\ell}^{(0)} - P_{g',j,\ell}^{(0)})^2\right| \le 4R_{\text{denoise}} + 4R_{\text{denoise}}^2,$$

and averaging over ℓ keeps this bound. For the second part, fix $i \neq j$ and condition on their reference coordinates. The summands of $\tilde{d}^2(i,j)$ with $\ell \notin \{i,j\}$ are independent functions of $\check{\xi}_{g',\ell} \sim \check{\mu}$, lie in $[0,1]$, and have mean $\check{d}^2_{g'}(\check{\xi}_{g',i}, \check{\xi}_{g',j})$, so Bernstein's inequality and a union bound over the at most $N^2_{g'}$ pairs bound their average's deviation by $C_\lambda(\log N_{g'}/N_{g'})^{1/2}$, as in the proof of Lemma A.5. The two omitted summands change the average by at most $2/N_{g'}$, and $i = j$ gives zero on both sides. Taking the maximum over networks proves the second bound in (52).

Proof of statement (i). We now turn small sampling errors into uniform reference coordinate alignment. If $R_{\text{denoise}} = o_p(1)$, then $e_N + s_N = o_p(1)$ by (52). For each source, Lemma A.7 (i) allows a symmetry for which the first two terms of $\theta_{g,N}$ tend to zero as $e_N + s_N \to 0$. Its last term is at most s_N. Hence, with Lemma A.8 (i), we have

$$\inf_{T \in \mathcal{T}} \epsilon_g(T) \longrightarrow 0 \qquad \text{as} \qquad e_N + s_N \to 0.$$

Lemmas A.7 and A.8 provide this implication uniformly over the configurations and chosen minimizers. Applying it to the random inputs and taking the maximum over the fixed number of sources yields $\epsilon = o_p(1)$. The projection bound proved above then gives $R_{\text{align}} = o_p(1)$.

Proof of statement (ii). We now use the bounds for the same argument. For each source, choose the symmetry supplied by Lemma A.7 (ii). Its two bounds and $W_2^2(\nu_G, \check{\mu}) \le s_N$ lead to

$$\theta_{g,N} \le C(e_N^2 + s_N).$$

Then with Lemma A.8 (ii), followed by the maximum over sources, we know that, whenever $e_N + s_N$ is sufficiently small,

$$\epsilon \le C\left[e_N + (e_N^2 + s_N)^{1/(d_\xi+2)}\right] \le C'(e_N^2 + s_N)^{1/(d_\xi+2)}. \tag{53}$$

The last step uses $e_N \le 1$ and $d_\xi \ge 1$.

We finally insert the first-stage rate. If $R_{\text{denoise}} = O_p(r_N)$ with $r_N \to 0$, the required small-error conditions hold with probability tending to one. Equation (52) gives

$$\begin{aligned} e_N^2 + s_N &= O_p\left(r_N^2 + \frac{\log(N_{\min})}{N_{\min}} + (\log(N_{\min}))^{3/2} N_{\min}^{-1/\max\{d_\xi,2\}}\right) \\ &= O_p\left(r_N^2 + (\log(N_{\min}))^{3/2} N_{\min}^{-1/\max\{d_\xi,2\}}\right). \end{aligned}$$

The second line absorbs the middle term into the last. Substituting into (53) proves statement (ii). □

A.4 Proof for the Large Sample Theory in Section 5: Proposition 7, Theorem 3, and Theorem 4

A.4.1 Proof of Proposition 7

Proof.

First, define the criterion formed from the true probabilities at the sampled target individuals:

$$\check{L}(\delta) := \frac{1}{N_G}\left\| P_G^{(0)} - \sum_{g<G} \delta_g \check{Q}_g^{(0)} \right\|_{\mathrm{F}}, \qquad \delta \in \Delta_{G-2}.$$

The reverse triangle inequality, the simplex constraint, and (15) give, uniformly over $\delta \in \Delta_{G-2}$,

$$\begin{aligned} |\hat{L}(\delta) - \check{L}(\delta)| &\le \frac{1}{N_G}\|\hat{P}_G^{(0)} - P_G^{(0)}\|_{\mathrm{F}} + \sum_{g<G} \frac{\delta_g}{N_G}\|\hat{Q}_g^{(0)} - \check{Q}_g^{(0)}\|_{\mathrm{F}} \\ &\le 2R_{\text{denoise}} + R_{\text{align}}. \end{aligned}$$

The normalized Frobenius norm is bounded by the maximum norm. Introduce the sampled counterpart of the affine identification constant:

$$\underline{V}_N := \min_{v\in\mathbb{R}^{G-1},\ \|v\|_2=1,\ \sum_{g<G} v_g=0} \frac{1}{N_G}\left\| \sum_{g<G} v_g \check{Q}_g^{(0)} \right\|_{\mathrm{F}}.$$

By Lemma A.3 (iii) and the bijective reference maps, $\underline{V}$ is the same minimum with the sampled norm replaced by the $L^2(\check{\mu}\otimes\check{\mu})$ norm of $\sum_{g<G} v_g \check{f}_g^{(0)}$. The empirical measure ν_G in (39) converges weakly to $\check{\mu}$ almost surely, hence $\nu_G^{\otimes 2} \Rightarrow \check{\mu}^{\otimes 2}$. The finitely many products $\check{f}_g^{(0)}\check{f}_h^{(0)}$ are bounded and continuous. Expanding the squared norms gives

$$|\underline{V}_N^2 - \underline{V}^2| \le (G-1)\max_{g,h<G}\left| \int_{\check{\Xi}^2} \check{f}_g^{(0)}(\check{\xi},\check{\zeta})\check{f}_h^{(0)}(\check{\xi},\check{\zeta})\,\mathrm{d}\left(\nu_G^{\otimes 2} - \check{\mu}^{\otimes 2}\right)(\check{\xi},\check{\zeta}) \right| \to 0 \qquad \text{almost surely.}$$

The factor $G-1$ bounds $(\sum_{g<G}|v_g|)^2$ on the zero-sum unit sphere. Thus $\underline{V}_N \ge \underline{V}/2$ with probability tending to one. The pointwise convex hull gives $\check{L}(\delta^*) = 0$, and every minimizing $\hat{\delta}$ satisfies, on that event,

$$\begin{aligned} \underline{V}_N\|\hat{\delta} - \delta^*\|_2 &\le \check{L}(\hat{\delta}) \\ &\le \hat{L}(\hat{\delta}) + 2R_{\text{denoise}} + R_{\text{align}} \\ &\le \hat{L}(\delta^*) + 2R_{\text{denoise}} + R_{\text{align}} \\ &\le 4R_{\text{denoise}} + 2R_{\text{align}}. \end{aligned}$$

Finally, $\|\hat{\delta} - \delta^*\|_1 \le \sqrt{G-1}\|\hat{\delta} - \delta^*\|_2$ proves the bound in Proposition 7. □

A.4.2 Proof of Theorem 3

Proof.

Proposition 6 (i) gives $R_{\text{align}} = o_p(1)$, and (15) gives maximum-norm projection error $o_p(1)$. Proposition 7 gives $R_{\text{weight}} = o_p(1)$ when $G \geq 3$. For $G = 2$, the weight simplex is a singleton and the weight error is identically zero. Together with $R_{\text{denoise}} = o_p(1)$, (11) proves Theorem 3. The bandwidth in Proposition 5 (i) supplies the first-stage premise for the specified smoother. □

A.4.3 Proof of Theorem 4

Proof.

Proposition 6 (ii) and (15) give the alignment and projection errors the order stated in Theorem 4. Here r_N is absorbed because $r_N \leq (r_N^2)^{1/(d_\xi+2)}$ eventually. Proposition 7 gives the same order for R_{weight} when $G \geq 3$, while it is zero for $G = 2$. Equation (11) proves the general rate in Theorem 4.

For the specified smoother, Proposition 5 (ii) permits $r_N = (\log(N_{\min})/N_{\min})^{1/\max\{d_\xi+2,4\}}$. When $d_\xi \leq 2$, $r_N^2 = (\log(N_{\min})/N_{\min})^{1/2}$. When $d_\xi \geq 3$, $2/(d_\xi + 2) > 1/d_\xi$. In either case,

$$r_N^2 \leq C(\log(N_{\min}))^{3/2} N_{\min}^{-1/\max\{d_\xi,2\}}.$$

Substituting into the general rate proves Theorem 4. The projection and weight errors have the same rate by the preceding bounds. □

B Further Discussion on the Assumptions

Assumptions 3 and 4 are high-level conditions on the internal geometries of across networks. In this subsection we further discuss these assumptions in two directions. From the optimal transport side, we show that they are statements related to the lower bounds for the Gromov-Wasserstein distance. From the network formation perspective, we show that they are satisfied by most of the standard non-strategic network formation models.

B.1 Relation to Optimal Transport Literature

Before we relate Assumptions 3 and 4 to the optimal transport literature, we start by restating these two assumptions in different ways in Propositions B.1 and B.2.

For $g \in \{1, ..., G-1, G\}$, define the distribution of internal structure of network g as the law of its within-network disimilarity (cost) under two independent draws from its own latent distribution,

$$\mathcal{L}_g := \left(d_g^2\right)_{\#} \left(\mu_g \otimes \mu_g\right)$$

which is a Borel probability measure on $[0, 1]$. With the reference space representation, it is also the law of $\check{d}_g^2$ under $\check{\mu} \otimes \check{\mu}$. Denote W_2 as the quadratic Wasserstein distance between Borel probability measures on $\mathbb{R}$. The next proposition restates Assumption 3: it is exactly the condition under which the Step 2 of Algorithm 2 collapses onto the one-dimensional comparison of these distributions.

Proposition B.1 (Restatement of Assumption 3)
Denote:

$$F_{C,D}(\pi) := \iint_{\mathcal{X}\times\mathcal{X}} \iint_{\mathcal{X}\times\mathcal{X}} \left[C(\xi, \xi') - D(\zeta, \zeta')\right]^2 \mathrm{d}\pi(\xi, \zeta)\mathrm{d}\pi(\xi', \zeta')$$

With this, the following statements hold:

(*i*) *For every $g \in \{1, ..., G-1\}$ and every $\pi \in \Pi(\mu_g, \mu_G)$,*

$$F_{d_g^2, d_G^2}(\pi) \geq W_2^2\left(\mathcal{L}_g, \mathcal{L}_G\right). \tag{54}$$

(*ii*) *Under Assumption 3, for every $g \in \{1, ..., G-1\}$,*

$$F_{d_g^2, d_G^2}\left(\pi^{g,G}\right) = \left\|\check{d}_g^2 - \check{d}_G^2\right\|_{L^2(\check{\mu}\otimes\check{\mu})}^2 = W_2^2\left(\mathcal{L}_g, \mathcal{L}_G\right), \tag{55}$$

so the achieved objective value of Step 2 of Algorithm 2 is the squared Wasserstein distance between the two networks' dissimilarity distributions.

(*iii*) *Conversely, suppose that for every $g \in \{1, ..., G-1\}$ there exists $\pi^{g,G} \in \Pi(\mu_g, \mu_G)$ with $F_{d_g^2, d_G^2}(\pi^{g,G}) = W_2^2\left(\mathcal{L}_g, \mathcal{L}_G\right)$. Then there is an admissible family of alignment maps on the same reference space $(\check{\Xi}, \check{\mu})$, with $\left(\rho_g, \rho_G\right)_{\#} \check{\mu} = \pi^{g,G}$ for every g. And their pulled-back dissimilarities satisfy the comonotonicity of Assumption 3.*

Proof.

Proof of statement (i). Fix $\pi \in \Pi(\mu_g, \mu_G)$ and consider the associated $d_g^2(\xi, \xi')$ and $d_G^2(\zeta, \zeta')$. By definition, $F_{d_g^2, d_G^2}(\pi) = \mathbb{E}\left\{\left[d_g^2(\xi, \xi') - d_G^2(\zeta, \zeta')\right]^2\right\}$. The pair (ξ, ξ') collects the first coordinates of two independent draws from π, so it has law $\mu_g \otimes \mu_g$, and $d_g^2(\xi, \xi')$ has law $\mathcal{L}_g$. Symmetrically, $d_G^2(\zeta, \zeta')$ has law $\mathcal{L}_G$. Hence, the joint law of $\left(d_g^2(\xi, \xi'), d_G^2(\zeta, \zeta')\right)$ is one coupling of $\left(\mathcal{L}_g, \mathcal{L}_G\right)$, while $W_2^2\left(\mathcal{L}_g, \mathcal{L}_G\right)$ is the smallest value of $\mathbb{E}\left\{\left[d_g^2(\xi, \xi') - d_G^2(\zeta, \zeta')\right]^2\right\}$ over all such couplings. This completes the proof.

Proof of statement (ii). The first equality of (55) is Theorem 1 (i). For the second equality, denote $X := \check{d}_g^2(\check{\xi}, \check{\xi}')$ and $Y := \check{d}_G^2(\check{\xi}, \check{\xi}')$ for $(\check{\xi}, \check{\xi}') \sim \check{\mu} \otimes \check{\mu}$. The middle term is then $\mathbb{E}\left[(X - Y)^2\right]$. The laws of X and Y are $\mathcal{L}_g$ and $\mathcal{L}_G$, and by Assumption 3 the pair (X, Y) is comonotone. It remains to show that a comonotone pair attains the smallest value of $\mathbb{E}\left[(X - Y)^2\right]$ among all couplings of the two laws.

Expand $(X - Y)^2 = X^2 + Y^2 - 2XY$. The expectations of X^2 and Y^2 depend only on the two laws, not on the coupling. For $\mathbb{E}[XY]$, using the same argument as in the proof of Theorem 1 (i), $\mathbb{E}[XY] = \int_0^\infty \int_0^\infty \mathbb{P}(X > u, Y > v)\, \mathrm{d}u\mathrm{d}v$, and every coupling satisfies $\mathbb{P}(X > u, Y > v) \leq \min\{\mathbb{P}(X > u), \mathbb{P}(Y > v)\}$. By the property of comonotonicity used in the proof of Theorem 1 (i), the comonotone pair turns this inequality into an equality at every (u, v). Thus, it maximizes $\mathbb{E}[XY]$ and hence minimizes $\mathbb{E}\left[(X - Y)^2\right]$. Therefore, the middle term of (55) equals $W_2^2\left(\mathcal{L}_g, \mathcal{L}_G\right)$.

Proof of statement (iii). First, fix g and $\pi^{g,G} \otimes \pi^{g,G}$ with $((\xi, \zeta), (\xi', \zeta'))$. Expanding the square as in the proof of statement (ii), the condition forces $\mathbb{E}\left[d_g^2(\xi, \xi')\, d_G^2(\zeta, \zeta')\right]$ to equal its coupling-free upper bound. The upper bound replaces the joint probability $\mathbb{P}\left(d_g^2(\xi, \xi') > u,\ d_G^2(\zeta, \zeta') > v\right)$ by the smaller of the two marginal probabilities. Hence, the equality means that the two dissimilarities cross every level together. That is, the pair $\left(d_g^2(\xi, \xi'),\ d_G^2(\zeta, \zeta')\right)$ is comonotone: two independent evaluations of this pair never rank the source dissimilarity and the target dissimilarity in opposite directions.

Second, define a measure $\mathcal{P}$ on Ξ^G by the following construction. Draw one target position $\zeta \sim \mu_G$. Then draw each source position from its conditional law given ζ, the conditional laws being the disintegrations of the $\pi^{g,G}$ as in Lemma C.1. By construction, the coordinates (g, G) under $\mathcal{P}$ have joint law $\pi^{g,G}$, so in particular the g-th marginal of $\mathcal{P}$ is μ_g.

Third, since Ξ^G is Polish, Lemma 1 gives a Borel map $B : \check{\Xi} \to \Xi^G$ with $B_\# \check{\mu} = \mathcal{P}$. Let ρ'_g be the g-th coordinate of B. Then $\left(\rho'_g\right)_\# \check{\mu} = \mu_g$ with values in $\mathrm{supp}(\mu_g)$. Thus, $\left\{\rho'_g\right\}$ is an admissible alignment family, and $\left(\rho'_g, \rho'_G\right)_\# \check{\mu} = \pi^{g,G}$. Lastly, drawing a pair of reference coordinates from $\check{\mu} \otimes \check{\mu}$ and reading it through $\left(\rho'_g, \rho'_G\right)$ reproduces a draw from $\pi^{g,G}$. Hence, the pulled-back dissimilarities of the new family inherit the comonotonicity of the first step, which is exactly Assumption 3.

□

Under Assumption 3, the non-atomic route of Assumption 4 (i) implies the existence of the deterministic maps. Hence, two conditions of Assumption 4 (i) are not unrelated alternatives, where the first is a primitive sufficient condition for the second.

Proposition B.2 (Restatement of Assumption 4 (i))

Let Assumption 3 hold and fix $g \in \{1, ..., G-1\}$. If the law of $\check{d}_g^2$ under $\check{\mu} \otimes \check{\mu}$ is non-atomic, then there exists a Borel function $b_g : [0, \infty) \to [0, 1]$ with $\check{d}_G^2 = b_g \circ \check{d}_g^2$ $\check{\mu} \otimes \check{\mu}$-a.e.

Proof.

Denote $X := \check{d}_g^2(\check{\xi}, \check{\xi}')$ and $Y := \check{d}_G^2(\check{\xi}, \check{\xi}')$ for $(\check{\xi}, \check{\xi}') \sim \check{\mu} \otimes \check{\mu}$, the pair from the proof of Proposition B.1 (ii). As recorded there, X has law $\mathcal{L}_g$, Y has law $\mathcal{L}_G$, both take values in $[0, 1]$, and the pair is comonotone by Assumption 3. Let $\mathcal{P}_g$ be the distribution function of $\mathcal{L}_g$, continuous because $\mathcal{L}_g$ is non-atomic. Let $\mathcal{P}_g^{-1}$ and $\mathcal{P}_G^{-1}$ be the quantile functions of $\mathcal{L}_g$ and $\mathcal{L}_G$, nondecreasing and hence Borel, with values in $[0, 1]$.

Let $T \sim \text{Uniform}(0, 1)$ and consider the pair $\left(\mathcal{P}_g^{-1}(T), \mathcal{P}_G^{-1}(T)\right)$. Both coordinates are nondecreasing in T, so the pair is comonotone. By the classical quantile transform, as in the proof of Lemma 1, its marginal laws are $\mathcal{L}_g$ and $\mathcal{L}_G$. We now have two comonotone pairs with the same marginal laws. By the property of comonotonicity used in the proof of Proposition B.1 (ii), both pairs have $\mathbb{P}(\cdot > u, \cdot > v) = \min\{\mathbb{P}(X > u), \mathbb{P}(Y > v)\}$ at every (u, v). This joint probability of exceeding levels, together with the two marginal laws, determines the joint distribution function by inclusion-exclusion. Hence, the two pairs have the same joint law.

Define $b_g(x) := \mathcal{P}_G^{-1}\left(\mathcal{P}_g(x)\right)$ where $\mathcal{P}_g(x) \in (0, 1)$ and $b_g(x) := 0$ otherwise, a Borel function with values in $[0, 1]$. Since $\mathcal{P}_g$ is continuous, we have $\mathcal{P}_g\left(\mathcal{P}_g^{-1}(t)\right) = t$ for every $t \in (0, 1)$, the value t being attained by $\mathcal{P}_g$ and $\mathcal{P}_g^{-1}(t)$ being the smallest point attaining at least t. Therefore, $b_g\left(\mathcal{P}_g^{-1}(T)\right) = \mathcal{P}_G^{-1}\left(\mathcal{P}_g\left(\mathcal{P}_g^{-1}(T)\right)\right) = \mathcal{P}_G^{-1}(T)$ almost surely. That is, under the quantile pair, the second coordinate is exactly b_g of the first. Now, the set of pairs (x, y) with $y \neq b_g(x)$ is Borel, and its probability is determined by the joint law alone. It is null under the quantile pair, and the two joint laws are equal, so it is null under (X, Y) as well. Hence, $Y = b_g(X)$ almost surely, which is the identity $\check{d}_G^2 = b_g \circ \check{d}_g^2$ $\check{\mu} \otimes \check{\mu}$-a.e.

□

Propositions B.1 and B.2 locate Assumptions 3 and 4 within the optimal transport literature. The isomorphism invariants of a metric measure space were first studied by Mémoli (2011) and later extended to the network setting by Chowdhury and Mémoli (2019). Mémoli (2011) organizes computable lower bounds for the Gromov-Wasserstein distance as follows:

(i) The first lower bound (FLB) is based on eccentricities: how far a position sits from the rest of the population on average.

(ii) The second lower bound (SLB) is based on the overall distribution of pairwise distances: how far positions tend to be from one another across the population, without consider to

which positions generate those distances.

(iii) The third lower bound (TLB) is based on local distance distributions: the full distribution of distances from each position to the rest of the population, comparing positions according to how similar these distributions are.

These lower bounds satisfy

$$\mathrm{GW}(X, Y) \geq \mathrm{TLB}(X, Y) \geq \mathrm{FLB}(X, Y) \qquad \text{and} \qquad \mathrm{GW}(X, Y) \geq \mathrm{SLB}(X, Y).$$

Within this hierarchy, our assumptions can be read as follows.

- **Assumption 3 makes the SLB tight.**
  - *Interpretation:* The reference coordinates sort all the geometries simultaneously: a pair of reference coordinates that is far apart in a source is far apart in the target, and near in one is near in the other. One ranking of reference coordinate pairs by dissimilarity serves every network at once, so the matching problem has no conflict to resolve.
  - *Relation to the SLB:* By Proposition B.1 (ii), Assumption 3 collapses the inequality (54) to an equality, the comonotone coupling of two one-dimensional laws being the optimal one. The Gromov-Wasserstein optimization (Step 2 of Algorithm 2) is then solved at an objective value that is precisely the SLB quantity $W_2^2\left(\mathcal{L}_g, \mathcal{L}_G\right)$, computable in closed form by sorting the two dissimilarity distributions.
- **Assumption 4 (i) makes the SLB deterministic.**
  - *Interpretation:* Assumption 3 says the two dissimilarities move together, and Assumption 4 (i) says the co-movement leaves no slack: the source dissimilarity pins down the target dissimilarity.
  - *Relation to the SLB:* It ensures that the optimal coupling of the two one-dimensional laws underlying the SLB is induced by a map rather than being a genuine coupling.
- **Assumption 4 (ii) makes the SLB complete enough.**
  - *Interpretation:* Whatever the geometry cannot distinguish, the graphons must not distinguish either. Optimality can only pin the coupling down up to the symmetries of the target dissimilarity, and Assumption 4 (ii) demands that those symmetries be harmless for the object we actually impute.
  - *Relation to the invariants:* The distribution of distances is not a complete invariant, as two genuinely different networks can share it (cf. Mémoli (2011), Section 5.1). The identification notion of Chowdhury and Mémoli (2019) is weak isomorphism, and a coupling of the reference with itself that preserves the target dissimilarity is exactly a weak self-isomorphism of the target measure network. Statement (ii) requires that

every such weak self-isomorphism be simultaneously a weak self-isomorphism of every network's edge probability, the target's own included, so that the incompleteness of the distance invariant never matters for the imputation.

B.2 Relation to Commonly Used Network Formation Models

Corollary 1 provides three sets of modeling examples that satisfy Assumptions 3 and 4. We now relate these conditions to particular specifications of commonly used network formation models. The relevant restrictions concern not only how links form within each network, but also how the resulting connection patterns compare across networks after alignment. We use friendship networks across cities to explain this distinction.

Example B.1 (Friendship networks across cities)**.**
Consider friendship networks observed in G cities:

- Cities $g \in \{1, ..., G-1\}$ are sources, and city G is the target. All networks are observed pre-treatment $(w = 0)$, and only the source networks are observed post-treatment $(w = 1)$.
- A resident's latent characteristics ξ may describe age, income, social group, sociability, or position in a social space. Their distribution in city g is μ_g. Conditional on individual characteristics and the city-specific effect γ_g, friendships form independently across pairs, with probabilities determined by $f^{(w)}$.
- Take the reference space to be $([0,1], \text{Leb})$. A reference coordinate $\check{\xi}$ indexes an individual in a common reference population, and $\rho_g(\check{\xi})$ gives the characteristics assigned to that individual in city g. The requirement $\left(\rho_g\right)_{\#} \text{Leb} = \mu_g$ ensures that these assignments reproduce the city's distribution of characteristics.
- When the characteristic is scalar, we may choose ρ_g as the quantile function of μ_g. Then $\check{\xi}$ has a within-city percentile interpretation: for example, $\rho_g(0.5)$ and $\rho_G(0.5)$ are the median characteristics in the two cities, although their values may differ. With discrete characteristics, several reference coordinates may correspond to the same value. For multivariate characteristics, $\check{\xi}$ is a common index rather than an intrinsic ranking of the characteristics.
- The reference representation $\check{f}_g^{(w)}(\check{\xi}, \check{\xi}') = f^{(w)}\left(\rho_g(\check{\xi}), \rho_g(\check{\xi}'), \gamma_g\right)$ gives the friendship probability between the residents corresponding to $\check{\xi}$ and $\check{\xi}'$ in city g. Holding the reference coordinates fixed therefore allows us to compare corresponding pairs across cities.
- The internal dissimilarity $\check{d}_g^2(\check{\xi}, \check{\xi}') = \int_0^1 \left[\check{f}_g^{(0)}(\check{\xi}, \check{u}) - \check{f}_g^{(0)}(\check{\xi}', \check{u})\right]^2 \mathrm{d}\check{u}$ measures how differently these two residents connect to the rest of the reference population. It is not their probability of being friends with each other. Assumption 3 requires that, when comparing two pairs of reference individuals, the ordering of these dissimilarities cannot reverse between

a source city and the target city.

Different cities may have different distributions μ_g and different friendship rules. The existence of the reference representation does not itself restrict these differences. Assumptions 3 and 4, however, impose restrictions on the resulting reference kernels $\check{f}_g^{(w)}$. The specifications below illustrate what those restrictions mean.

Remark B.1 (Common network formation models in reference coordinates). We consider four specifications within the friendship setting of Example B.1. Each uses the same reference-coordinate notation, while giving a different interpretation to individual characteristics. The discussion concerns Assumptions 3 and 4; the convex-hull restriction of Assumption 2 must be imposed separately.

(i) Stochastic block model illustrates statement (iii).

- *Model in reference coordinates*: suppose residents belong to finitely many social groups. Let $z_g(\xi)$ denote a resident's group in city g, and suppose the alignment maps satisfy $z_g\left(\rho_g(\check{\xi})\right) = \phi(\check{\xi})$ for a common assignment $\phi : [0,1] \to \{1,...,B\}$. Thus, corresponding reference individuals belong to corresponding groups. The reference kernel is $\check{f}_g^{(w)}(\check{\xi},\check{\xi}') = B_g^{(w)}\left(\phi(\check{\xi}),\phi(\check{\xi}')\right)$, where the symmetric matrix $B_g^{(w)}$ contains the friendship probabilities between groups. The common reference group masses are $\beta_b = \text{Leb}\left(\phi^{-1}(b)\right)$. The Erdős–Rényi model is the special case with one group and a constant friendship probability.
- *Conditions of the corollary*: the internal dissimilarity is $\check{d}_g^2(\check{\xi},\check{\xi}') = C_g\left(\phi(\check{\xi}),\phi(\check{\xi}')\right)$, where $C_g(b,b') = \sum_{\ell=1}^{B}\beta_\ell\left[B_g^{(0)}(b,\ell) - B_g^{(0)}(b',\ell)\right]^2$. Statement (iii) requires the ordering condition on these block dissimilarities and the condition that matchings preserving the target dissimilarities also preserve every network's link probabilities.

  A simple specification satisfies both conditions. Take two reference groups, with $\phi(\check{\xi}) = 1$ for $\check{\xi} \in [0, 0.3)$ and $\phi(\check{\xi}) = 2$ for $\check{\xi} \in [0.3, 1]$, and suppose the two rows of $B_g^{(0)}$ differ in every city. Each city then has exactly two dissimilarity levels: zero within groups and a positive value $C_g(1,2)$ between groups. Their ordering is therefore the same in every city. Moreover, a matching that preserves the target dissimilarities must map each group entirely into one group. Because the group masses are unequal, it cannot exchange the two groups. It must preserve group membership and hence all the block kernels, before and after treatment. Both conditions of statement (iii) hold.
- *Interpretation in the city example*: the residents corresponding to $\check{\xi} = 0.1$ and $\check{\xi}' = 0.2$ belong to the same group in every city and have identical friendship probabilities with every reference group. Their internal dissimilarity is zero. The residents corresponding to $\check{\xi} = 0.1$ and $\check{\xi}' = 0.8$ belong to different groups and have different friendship patterns. The

size of that difference may vary across cities, but the distinction between within-group and between-group pairs remains unchanged. With equally sized groups, exchanging the groups can create the ambiguity described in Remark 2.

(ii) **Chung–Lu expected-degree model** illustrates the restrictions on differences in sociability.

- *Model in reference coordinates*: let $s_g^{(w)}(\check{\xi})$ be the sociability of the resident with characteristics $\rho_g(\check{\xi})$ in city g, and consider $f_g^{(w)}(\check{\xi},\check{\xi}') = s_g^{(w)}(\check{\xi})\, s_g^{(w)}(\check{\xi}')$, with nonnegative sociability levels restricted so that all products lie in $[0,1]$. If sociability takes finitely many values on a common reference partition, write $s_g^{(w)}(\check{\xi}) = s^{(w)}_{g,\phi(\check{\xi})}$. The model then has the block form in statement (iii), with $B_g^{(w)}(b,b') = s_{g,b}^{(w)}\, s_{g,b'}^{(w)}$.
- *Conditions of the corollary*: the pre-treatment internal dissimilarity is formulated as $\check{d}_g^2(\check{\xi},\check{\xi}') = \left[s_g^{(0)}(\check{\xi}) - s_g^{(0)}(\check{\xi}')\right]^2 \int_0^1 \left[s_g^{(0)}(\check{u})\right]^2 \mathrm{d}\check{u}$. Thus, Assumption 3 restricts the ordering of sociability gaps between pairs of reference individuals. Preserving the ranking of individual sociability is not sufficient.

  For example, if $s_g^{(0)}(\check{\xi}) = a_g s(\check{\xi})$ with $a_g > 0$, then $\check{d}_g^2(\check{\xi},\check{\xi}') = a_g^4 \left[s(\check{\xi}) - s(\check{\xi}')\right]^2 \int_0^1 s(\check{u})^2 \mathrm{d}\check{u}$. All cities therefore order reference pairs in the same way, establishing Assumption 3. This calculation does not by itself establish Assumption 4. For the finite-category version, both conditions of statement (iii) must be checked. In particular, the two-group specification of item (1), with unequal group masses and distinct pre-treatment sociability levels in every city, satisfies both conditions, provided sociability remains constant within those groups for both treatment statuses.
- *Interpretation in the city example*: $\check{\xi}$ identifies a corresponding resident across cities, while $s_g^{(0)}(\check{\xi})$ describes that resident's sociability in city g. Under proportional sociability, a city may make everyone more sociable by the same factor without changing which pairs have more different friendship patterns. By contrast, suppose three corresponding groups have sociability levels $(0.1, 0.2, 0.4)$ in a source city and $(0.1, 0.4, 0.5)$ in the target. The ranking of the groups is unchanged, but the gap between the first two groups changes from smaller to larger than the gap between the last two. The ordering of internal dissimilarities reverses, violating Assumption 3.

(iii) **Latent-space model with circular positions** gives an explicit example of statement (i).

- *Model in reference coordinates*: suppose residents have positions on a circle representing a social characteristic. Let $x(\xi)$ denote a resident's angular position, and assume the alignment maps satisfy $x\left(\rho_g(\check{\xi})\right) = 2\pi\check{\xi}$ for every city, almost everywhere. This specification gives all cities the same uniform distribution of angular positions and aligns residents at the same position. Consider $\check{f}_g^{(w)}(\check{\xi},\check{\xi}') = a_g^{(w)} + b_g^{(w)} \cos\left(2\pi(\check{\xi} - \check{\xi}')\right)$, where $0 \le b_g^{(w)} \le$

$\min\left\{a_g^{(w)}, 1 - a_g^{(w)}\right\}$ and $b_g^{(0)} > 0$. These restrictions ensure valid link probabilities and a non-constant pre-treatment dependence on circular distance.

- *Conditions of the corollary*: take $K = 1$ and $\phi_1(\check{\xi}, \check{\xi}') = \cos\left(2\pi(\check{\xi} - \check{\xi}')\right)$. Direct integration gives $M_{11}(\check{\xi}, \check{\xi}') = 1 - \phi_1(\check{\xi}, \check{\xi}')$. Hence, ϕ_1 is a function of M_{11}, and the positive pre-treatment coefficient $b_g^{(0)}$ satisfies the loading restriction in statement (i). Both conditions of that statement hold.

  In particular, $\check{d}_g^2(\check{\xi}, \check{\xi}') = \left(b_g^{(0)}\right)^2 \left[1 - \cos\left(2\pi(\check{\xi} - \check{\xi}')\right)\right]$. All cities therefore order reference pairs by the same circular distance. Furthermore, preserving the target dissimilarity preserves the cosine term and hence every network's link probabilities for both treatment statuses. This example establishes Assumptions 3 and 4 without invoking the neighborhood-overlap condition in statement (ii).
- *Interpretation in the city example*: here $\check{\xi}$ indexes a position on a social circle, rather than a ranking from low to high sociability. Residents at nearby positions have similar friendship patterns. The parameter $a_g^{(w)}$ changes the overall probability of friendship, while $b_g^{(w)}$ changes the strength of the preference for nearby residents. Changing these parameters can alter friendship probabilities substantially while preserving the ordering of differences in connection patterns across cities.

(iv) Dyadic link formation with degree heterogeneity (**Graham, 2017**) has an exact finite-type specialization under statement (iii).

- *Model in reference coordinates*: let $X(\xi)$ denote observed characteristics and $A(\xi)$ an individual sociability effect. Define $x_g(\check{\xi}) = X\left(\rho_g(\check{\xi})\right)$ and $a_g(\check{\xi}) = A\left(\rho_g(\check{\xi})\right)$. Consider the reference kernel $\check{f}_g^{(w)}(\check{\xi}, \check{\xi}') = \Lambda\left(W\left(x_g(\check{\xi}), x_g(\check{\xi}')\right)' c_g^{(w)} + a_g(\check{\xi}) + a_g(\check{\xi}')\right)$, where Λ is the logistic distribution function, W is symmetric in its arguments, and $c_g^{(w)}$ is a vector of coefficients.

  Suppose the pair $\left(x_g(\check{\xi}), a_g(\check{\xi})\right)$ is constant on each block of a common finite reference partition, taking values $\left(x_{g,b}, a_{g,b}\right)$ on block b. Then the model has block probabilities $B_g^{(w)}(b, b') = \Lambda\left(W\left(x_{g,b}, x_{g,b'}\right)' c_g^{(w)} + a_{g,b} + a_{g,b'}\right)$.
- *Conditions of the corollary*: statement (iii) applies when these induced block matrices satisfy its two conditions. For example, the two-group specification in item (1) applies if the group masses are unequal and the induced pre-treatment block rows differ in every city. This is an exact application of the corollary and does not require a sparse approximation to the logistic function. With continuously varying characteristics or individual effects, the finite-block argument no longer applies directly, and Assumptions 3 and 4 require separate verification.
- *Interpretation in the city example*: the resident corresponding to $\check{\xi}$ has observed characteris-

tics $x_g(\check{\xi})$ and sociability $a_g(\check{\xi})$ in city g. The pair covariates describe which residents tend to befriend one another, while the individual effects describe their general tendency to form friendships. These characteristics and effects may differ across cities. The corollary restricts the resulting differences in friendship patterns after residents have been aligned through the common reference partition.

Models with finitely many individual types, including suitable degree-corrected block models and dyad-independent exponential random graph models, can also be written as block models, subject to the two conditions of statement (iii). More general latent-space models must satisfy the common-distance and neighborhood-overlap conditions of statement (ii), while spline or series models must satisfy the restrictions on basis functions and coefficients in statement (i).

Other commonly used models such as the exact configuration model, Watts–Strogatz rewiring model, dyad-dependent exponential random graph models, and subgraph-generated models with joint edge formation generally violate our conditional independence requirement. This requirement means that, after conditioning on all individuals' latent characteristics and the network-specific effect, whether one pair forms a link is independent of whether other pairs form links. Their independent-edge alternatives may fit the framework, but they are different models. Strategic-interaction and dynamic models are not considered here.

C Auxiliary Lemmas and Proofs

C.1 Auxiliary Lemmas

Lemma C.1 (Disintegration; see, e.g., Ambrosio et al. (2008) Theorem 5.3.1)
Fix $g \in \{1, ..., G-1\}$ and let $\tilde{\pi}^{g,G} \in \Pi(\mu_g, \mu_G)$ be any coupling of μ_g and μ_G. Since Ξ is Polish, every Borel probability measure on the Polish space $\Xi \times \Xi$ is Radon. Disintegrating $\tilde{\pi}^{g,G}$ with respect to the second coordinate projection $(\xi, \zeta) \to \zeta$, whose pushforward of $\tilde{\pi}^{g,G}$ is the second marginal μ_G, there exists a μ_G-a.e. unique family $\left\{\tilde{\pi}_\zeta^{g,G}\right\}_{\zeta \in \Xi}$ of Borel probability measures on Ξ, the conditional laws of the first coordinate given that the second coordinate equals ζ, such that

(*i*) *$\zeta \to \tilde{\pi}_\zeta^{g,G}(B)$ is Borel measurable for every $B \in \mathcal{B}(\Xi)$, and*

(*ii*) *for every bounded Borel $h : \Xi \times \Xi \to \mathbb{R}$,*

$$\iint_{\Xi \times \Xi} h(\xi, \zeta) \mathrm{d}\tilde{\pi}^{g,G}(\xi, \zeta) = \int_\Xi \left[\int_\Xi h(\xi, \zeta) \mathrm{d}\tilde{\pi}_\zeta^{g,G}(\xi) \right] \mathrm{d}\mu_G(\zeta).$$

We abbreviate this as $\mathrm{d}\tilde{\pi}^{g,G}(\xi, \zeta) = \mathrm{d}\tilde{\pi}_\zeta^{g,G}(\xi)\mathrm{d}\mu_G(\zeta)$.

Lemma C.2 (Properties of the Gromov-Wasserstein projection)
Let $\left\{\tilde{\pi}_{\zeta}^{g,G}\right\}$ *be any disintegration of* $\tilde{\pi}^{g,G}$ *with respect to* μ_G*. Then:*

(*i*) $\tilde{Q}_g^{(w)}$ *is Borel measurable on* $\Xi \times \Xi$ *and*

$$\sup\left|\tilde{Q}_g^{(w)}\right| \leq \sup_{\mathrm{supp}(\mu)\times\mathrm{supp}(\mu)} \left|f^{(w)}(\cdot,\cdot,\gamma_g)\right| < \infty.$$

In particular, as $P_g^{(w)}$ *takes values in* $[0,1]$*, so does* $\tilde{Q}_g^{(w)}$*.*

(*ii*) $\tilde{Q}_g^{(w)}(\zeta,\zeta') = \tilde{Q}_g^{(w)}(\zeta',\zeta)$ *for all* ζ,ζ'*.*

(*iii*) $\tilde{Q}_g^{(w)}$ *is well-defined as an element of* $L^2(\mu_G \otimes \mu_G)$*. That is, any two disintegrations of* $\tilde{\pi}^{g,G}$ *produce versions of* $\tilde{Q}_g^{(w)}$ *that coincide at* $\mu_G \otimes \mu_G$*-a.e. pair* (ζ,ζ')*, so the choice made in Definition 1 does not matter.*

(*iv*) *Let* (ξ,ζ) *and* (ξ',ζ') *be independent with joint distribution* $\tilde{\pi}^{g,G}$*. Then*

$$\tilde{Q}_g^{(w)}(\zeta,\zeta') = \mathbb{E}\left[P_g^{(w)}(\xi,\xi') \,\middle|\, \zeta,\zeta'\right] \quad \textit{almost surely},$$

and consequently

$$\left\|\tilde{Q}_g^{(w)}\right\|_{L^2(\mu_G\otimes\mu_G)} \leq \left\|P_g^{(w)}\right\|_{L^2(\mu_g\otimes\mu_g)}.$$

Proof.

Proof of statement (i). When the integrand is the indicator of a measurable rectangle $A \times B$ with $A, B \in \mathcal{B}(\Xi)$, the double integral of Definition 1 evaluates to

$$\int_\Xi \int_\Xi \mathbf{1}_A(\xi)\,\mathbf{1}_B(\xi')\mathrm{d}\tilde{\pi}_{\zeta}^{g,G}(\xi)\mathrm{d}\tilde{\pi}_{\zeta'}^{g,G}(\xi') = \tilde{\pi}_{\zeta}^{g,G}(A)\,\tilde{\pi}_{\zeta'}^{g,G}(B),$$

which is a product of Borel functions of ζ and ζ' by Lemma C.1 (i). The class of bounded Borel functions on $\Xi \times \Xi$ for which this double integral is a Borel function of (ζ,ζ') contains these indicators, is a vector space, and is closed under bounded monotone limits by the monotone convergence theorem applied inside the double integral. Hence, by the functional monotone class theorem, the rectangles forming a π-system generating $\mathcal{B}(\Xi \times \Xi)$ as Ξ is second countable, the class contains all bounded Borel functions, and in particular $P_g^{(w)}$, which is Borel and bounded by the maintained assumptions of Section 3.1. This proves the measurability. The bound holds because each $\tilde{\pi}_{\zeta}^{g,G} \otimes \tilde{\pi}_{\zeta'}^{g,G}$ is a probability measure, so the integral never exceeds the supremum of the integrand, and since the first marginal of $\tilde{\pi}^{g,G}$ is μ_g, we have $\tilde{\pi}_{\zeta}^{g,G}\left(\mathrm{supp}(\mu_g)\right) = 1$ for μ_G-a.e. ζ, so $P_g^{(w)}$ is evaluated on $\mathrm{supp}(\mu_g) \times \mathrm{supp}(\mu_g)$ for $\mu_G \otimes \mu_G$-a.e. (ζ,ζ'). In particular, as $P_g^{(w)}$ takes values in $[0,1]$ there, so does $\tilde{Q}_g^{(w)}$.

Proof of statement (ii). Fix (ζ, ζ'). Then

$$\tilde{Q}_g^{(w)}(\zeta', \zeta) = \iint P_g^{(w)}(\xi, \xi') \mathrm{d}\tilde{\pi}_{\zeta'}^{g,G}(\xi) \mathrm{d}\tilde{\pi}_{\zeta}^{g,G}(\xi') = \iint P_g^{(w)}(\xi', \xi) \mathrm{d}\tilde{\pi}_{\zeta'}^{g,G}(\xi') \mathrm{d}\tilde{\pi}_{\zeta}^{g,G}(\xi) = \tilde{Q}_g^{(w)}(\zeta, \zeta'),$$

where the second equality renames the two integration variables and exchanges the order of integration, licit by Fubini's theorem for the finite product measure, and the last equality uses the symmetry $P_g^{(w)}(\xi', \xi) = P_g^{(w)}(\xi, \xi')$.

Proof of statement (iii). Let $\left\{\left(\tilde{\pi}_{\zeta}^{g,G}\right)'\right\}_{\zeta \in \Xi}$ be another disintegration of $\tilde{\pi}^{g,G}$ with respect to μ_G. By the μ_G-a.e. uniqueness in Lemma C.1, there is a μ_G-null set O with $\left(\tilde{\pi}_{\zeta}^{g,G}\right)' = \tilde{\pi}_{\zeta}^{g,G}$ for all $\zeta \notin O$. Hence, the two versions of $\tilde{Q}_g^{(w)}$ coincide at every pair (ζ, ζ') off the set $(O \times \Xi) \cup (\Xi \times O)$, which is $\mu_G \otimes \mu_G$-null as $\mu_G(O) = 0$. Since $\tilde{Q}_g^{(w)}$ is bounded by statement (i), it is square integrable, and all versions define one element of $L^2(\mu_G \otimes \mu_G)$.

Proof of statement (iv). Let $((\xi, \zeta), (\xi', \zeta')) \sim \tilde{\pi}^{g,G} \otimes \tilde{\pi}^{g,G}$, so that $(\zeta, \zeta') \sim \mu_G \otimes \mu_G$ and $(\xi, \xi') \sim \mu_g \otimes \mu_g$. By statement (i), $\tilde{Q}_g^{(w)}(\zeta, \zeta')$ is a bounded Borel function of (ζ, ζ'). It remains to show that $\tilde{Q}_g^{(w)}(\zeta, \zeta')$ and $P_g^{(w)}(\xi, \xi')$ have the same integral against $B(\zeta, \zeta')$ for every bounded Borel B on $\Xi \times \Xi$. Applying Lemma C.1 (ii) once in each component of the product measure and Fubini's theorem, we have

$$\begin{aligned}
\mathbb{E}\left[P_g^{(w)}(\xi, \xi') B(\zeta, \zeta')\right] &= \int_{\Xi} \int_{\Xi} \left[\iint P_g^{(w)}(\xi, \xi') \mathrm{d}\tilde{\pi}_{\zeta}^{g,G}(\xi) \mathrm{d}\tilde{\pi}_{\zeta'}^{g,G}(\xi')\right] B(\zeta, \zeta') \mathrm{d}\mu_G(\zeta) \mathrm{d}\mu_G(\zeta') \\
&= \mathbb{E}\left[\tilde{Q}_g^{(w)}(\zeta, \zeta') B(\zeta, \zeta')\right],
\end{aligned}$$

and the a.e. uniqueness of the conditional expectation proves the representation. By Jensen's inequality, the square of a conditional expectation is dominated by the conditional expectation of the square, so $\tilde{Q}_g^{(w)}(\zeta, \zeta')^2 \leq \mathbb{E}\left[P_g^{(w)}(\xi, \xi')^2 \mid \zeta, \zeta'\right]$ almost surely. Taking expectations on both sides,

$$\left\|\tilde{Q}_g^{(w)}\right\|_{L^2(\mu_G \otimes \mu_G)}^2 = \mathbb{E}\left[\tilde{Q}_g^{(w)}(\zeta, \zeta')^2\right] \leq \mathbb{E}\left[P_g^{(w)}(\xi, \xi')^2\right] = \left\|P_g^{(w)}\right\|_{L^2(\mu_g \otimes \mu_g)}^2.$$

□

C.2 Additional Proofs for Other Auxiliary Results

C.2.1 Proof of Lemma 1

Proof.

Proof of statement (i). We first prove the statement for the reference $(\check{\Xi}, \check{\mu}) = ([0,1], \text{Leb})$, and then generalize to any other reference space.

Since $(\Xi, \mathcal{B}(\Xi))$ is a standard Borel space, it is Borel isomorphic to a Borel subset $B \subseteq [0,1]$. Let $b : \Xi \to B$ be such an isomorphism, let H be the distribution function of b under μ, and let $H^{-1}(t) := \inf\{x \in [0,1] : H(x) \geq t\}$, $t \in (0,1)$, be its quantile function. By the quantile transform, which places no restriction on μ, the variable $H^{-1}(\check{\xi})$ with $\check{\xi}$ uniform on $(0,1)$ has the same law as $b(\xi)$ with $\xi \sim \mu$. Define $\rho(\check{\xi}) := b^{-1}\left(H^{-1}(\check{\xi})\right)$, where H^{-1} turns the uniform draw $\check{\xi}$ into a draw with the law of $b(\xi)$, and b^{-1} carries this draw back into Ξ. This definition only fails on a Leb-null Borel set, namely the two endpoints and the $\check{\xi}$ with $H^{-1}(\check{\xi}) \notin B$. There, we set $\rho(\check{\xi}) := \zeta$ for a fixed point $\zeta \in \text{supp}(\mu)$, the support being nonempty. The resulting map is Borel. Then, for every Borel $A \subseteq \Xi$,

$$\begin{aligned}
\rho_{\#}\text{Leb}(A) &= \text{Leb}\left(\left\{\check{\xi} \in (0,1) : H^{-1}(\check{\xi}) \in b(A)\right\}\right) \\
&= \mu\left(\{\xi : b(\xi) \in b(A)\}\right) \\
&= \mu(A),
\end{aligned}$$

where the first equality drops the null exceptional set and unwinds b^{-1}, the second uses the quantile transform, and the third uses the injectivity of b. This proves statement (i) for the reference $([0,1], \text{Leb})$.

For a general reference $(\check{\Xi}, \check{\mu})$, one extra step is required. Since $\check{\Xi}$ is an uncountable Polish space, the Borel isomorphism theorem provides a Borel bijection $\check{b} : \check{\Xi} \to [0,1]$. Let F be the distribution function of $\check{b}$ under $\check{\mu}$. It is continuous, because a jump would give a single point positive mass, which the non-atomicity of $\check{\mu}$ forbids. By the probability integral transform, $F\left(\check{b}(\check{\xi})\right)$ is uniform on $[0,1]$ when $\check{\xi} \sim \check{\mu}$. That is, the map $u := F \circ \check{b}$ transforms a draw from $\check{\mu}$ into a uniform draw. The map ρ constructed above then transforms the uniform draw into a draw from μ. Hence, the composition $\rho \circ u$ pushes $\check{\mu}$ forward to μ, which proves statement (i) for any reference space. Notice that the continuity of F is the place where the non-atomicity of the reference is used.

Proof of statement (ii). The set $\rho^{-1}(\Xi \setminus \text{supp}(\mu))$ is Borel, the support being closed. It is also $\check{\mu}$-null, because statement (i) gives it the mass $\mu(\Xi \setminus \text{supp}(\mu)) = 0$. Redefine $\rho := \zeta$ on this set, with $\zeta \in \text{supp}(\mu)$ the same support point as in the proof of statement (i). The map stays Borel. The pushforward is unchanged, as the modification occurs on a $\check{\mu}$-null set. Hence, every value of ρ now lies in $\text{supp}(\mu)$, which is statement (ii). Finally, applying the construction to each μ_g, $g \in \{1, ..., G-1, G\}$, yields the family of alignment maps in the last statement.

□

C.2.2 Proof of Corollary 1

Statements (i) and (ii) share one structure: every within-network dissimilarity and every reference graphon is a function of a single scalar index of the pair of reference coordinates. We first show that this structure delivers Assumptions 3 and 4, then verify it in statements (i) and (ii), and finally treat the block structure of statement (iii) directly.

Proof.

All identities below hold up to $\check{\mu} \otimes \check{\mu}$-null sets, and Assumptions 3 and 4 are statements up to null sets as well. By Fubini's theorem, the model identities may be substituted into the integral defining $\check{d}_g^2$ in (3). We fix a Borel set $E \subseteq \check{\Xi} \times \check{\Xi}$ of full $\check{\mu} \otimes \check{\mu}$-measure, symmetric in its two coordinates, on which the finitely many identities used below hold. Moreover, for every $\check{\pi} \in \Pi(\check{\mu}, \check{\mu})$, both $(\check{\xi}, \check{\xi}')$ and $(\check{\zeta}, \check{\zeta}')$ have law $\check{\mu} \otimes \check{\mu}$ under $\check{\pi} \otimes \check{\pi}$, so both pairs lie in E almost surely.

A sufficient condition. Suppose that there exist a Borel function $T : \check{\Xi} \times \check{\Xi} \to [0, \infty)$ and, for every $g \in \{1, ..., G-1, G\}$ and $w \in \{0, 1\}$, functions H_g and $F_g^{(w)}$ on $[0, \infty)$ such that

$$\check{d}_g^2 = H_g \circ T \qquad \text{and} \qquad \check{f}_g^{(w)} = F_g^{(w)} \circ T \qquad \text{on } E, \tag{56}$$

with every H_g strictly increasing on $S := T(E)$. We show that Assumptions 3 and 4 follow.

For Assumption 3, take two pairs in E with index values t and s. The product of differences in Assumption 3 equals $\left[H_g(t) - H_g(s)\right][H_G(t) - H_G(s)] \geq 0$, as both factors share the sign of $t - s$.

For Assumption 4 (i), since H_g is strictly increasing on S, the map $b_g := H_G \circ H_g^{-1}$ is well defined and non-decreasing on $H_g(S) \subseteq [0, 1]$, with values in $[0, 1]$. Extending it to $[0, \infty)$ by $b_g(x) := \sup\left\{b_g(y) : y \in H_g(S),\ y \leq x\right\}$, with the supremum of the empty set taken to be zero, keeps it non-decreasing, hence Borel, and $\check{d}_G^2 = H_G \circ T = b_g \circ H_g \circ T = b_g \circ \check{d}_g^2$ on E. This is branch (b) of Assumption 4 (i).

For Assumption 4 (ii), let $\check{\pi} \in \Pi(\check{\mu}, \check{\mu})$ preserve $\check{d}_G^2$. By the observation on null sets, almost surely under $\check{\pi} \otimes \check{\pi}$, both pairs lie in E and $H_G\left(T(\check{\zeta}, \check{\zeta}')\right) = H_G\left(T(\check{\xi}, \check{\xi}')\right)$ with both arguments in S. Injectivity of H_G on S gives $T(\check{\zeta}, \check{\zeta}') = T(\check{\xi}, \check{\xi}')$, and then (56) gives $\check{f}_g^{(w)}(\check{\zeta}, \check{\zeta}') = \check{f}_g^{(w)}(\check{\xi}, \check{\xi}')$ for every g and both w.

Proof of statement (i). Let $T := \sum_{k=1}^K M_{kk}$, which is Borel by Tonelli's theorem. Substituting the model into (3), the intercept cancels, and expanding the square gives

$$\check{d}_g^2(\check{\xi}, \check{\xi}') = \int_{\check{\Xi}} \left[\sum_{k=1}^K \psi_{k,g}^{(0)} \left(\phi_k(\check{\xi}, \check{u}) - \phi_k(\check{\xi}', \check{u})\right)\right]^2 \mathrm{d}\check{\mu}(\check{u}) = \sum_{k=1}^K \sum_{\ell=1}^K \psi_{k,g}^{(0)} \psi_{\ell,g}^{(0)} M_{k\ell}(\check{\xi}, \check{\xi}'),$$

where $\psi_{k,g}^{(w)}$ denotes the k-th component of $\psi_g^{(w)}$. By condition (a), $M_{k\ell} = m_{k\ell} \circ T$ and $\phi_k = a_k \circ T$ for Borel functions $m_{k\ell}$ and a_k on $[0, \infty)$, with every $m_{k\ell}$ non-decreasing and every diagonal m_{kk} strictly increasing. Hence (56) holds with

$$H_g(t) := \sum_{k=1}^{K} \sum_{\ell=1}^{K} \psi_{k,g}^{(0)} \psi_{\ell,g}^{(0)} m_{k\ell}(t) \qquad \text{and} \qquad F_g^{(w)}(t) := \sum_{k=1}^{K} \psi_{k,g}^{(w)} a_k(t) + \alpha_g^{(w)}.$$

By condition (b), every coefficient $\psi_{k,g}^{(0)} \psi_{\ell,g}^{(0)}$ is nonnegative, so H_g is non-decreasing, and at least one $\psi_{k,g}^{(0)}$ is strictly positive, so the strictly increasing m_{kk} enters with the strictly positive coefficient $\left(\psi_{k,g}^{(0)}\right)^2$. Therefore, H_g is strictly increasing for every g, the target included, and the sufficient condition applies.

Proof of statement (ii). Let $T := \phi$. The model gives $\check{f}_g^{(w)} = \psi_g^{(w)} \circ T$, so it remains to show that $\check{d}_g^2 = H_g \circ T$ on E for a function H_g that is strictly increasing on $S = \phi(E)$, for every g.

By condition (b), the overlap of two neighborhoods depends on the pair only through T. We write it as

$$A_{r,r'}\left(T(\check{\xi}, \check{\xi}')\right) := \check{\mu}\left(\left\{\check{u} : \phi(\check{\xi}, \check{u}) < r\right\} \cap \left\{\check{u} : \phi(\check{\xi}', \check{u}) < r'\right\}\right) \qquad \text{for } r, r' > 0,$$

a strictly decreasing function of T. Letting $r' \to \infty$ gives the mass of the first neighborhood, and the same limit after exchanging $\check{\xi}$ and $\check{\xi}'$ gives the mass of the second, since T is symmetric. Hence $\check{\mu}\left(\{\check{u} : \phi(\check{\xi}, \check{u}) < r\}\right) = \check{\mu}\left(\{\check{u} : \phi(\check{\xi}', \check{u}) < r\}\right)$ for every $r > 0$ and every pair in E. Taking all positive rational r and using Fubini's theorem and monotone limits shows that the law of $\phi(\check{\xi}, \check{u})$ for $\check{u} \sim \check{\mu}$ is the same for $\check{\mu}$-a.e. $\check{\xi}$. Denote this common law on $[0, \infty)$ by λ. Restricting E to pairs whose two coordinates have this law, and redefining $S := \phi(E)$, we may assume that both profile distributions equal λ at every pair in E.

Fix g and set $h_g := \psi_g^{(0)}$ if $\psi_g^{(0)}$ is strictly increasing and $h_g := 1 - \psi_g^{(0)}$ otherwise. Then h_g is strictly increasing with values in $[0, 1]$ and leaves the squared differences in (3) unchanged. For $s, s' \in (0, 1)$, define

$$J_g\left(s, s'; T(\check{\xi}, \check{\xi}')\right) := \check{\mu}\left(\left\{\check{u} : h_g\left(\phi(\check{\xi}, \check{u})\right) < s\right\} \cap \left\{\check{u} : h_g\left(\phi(\check{\xi}', \check{u})\right) < s'\right\}\right).$$

Since h_g is increasing, each sublevel set $\{v : h_g(v) < s\}$ is a lower interval $[0, q)$ or $[0, q]$, possibly empty or the whole domain. Thus, J_g is an $A_{q,q'}$, or a monotone limit of such overlaps, or is determined by λ alone. In every case it depends on the pair only through T, as the notation indicates, and is non-increasing in T. This argument allows jumps in h_g and never modifies its values.

Using $1 - h_g(v) = \int_0^1 \mathbf{1}\{h_g(v) < s\} \mathrm{d}s$ to expand $\left[h_g(a) - h_g(b)\right]^2$ as a double integral over $(s, s') \in (0, 1)^2$, and integrating over $\check{u}$ by Tonelli's theorem, the two own terms depend only on λ and

the two cross terms are J_g. Hence, on E,

$$\check{d}_g^2 = H_g \circ T, \qquad H_g(t) := 2\int_{[0,\infty)} \left[1 - h_g(v)\right]^2 \mathrm{d}\lambda(v) - 2\int_0^1 \int_0^1 J_g(s, s'; t)\mathrm{d}s\mathrm{d}s',$$

and H_g is non-decreasing on S because J_g is non-increasing in t.

For the strict increase, take $t_1 < t_2$ in S and choose $r_0 > 0$ with $\lambda(\{r_0\}) = 0$. Condition (b) gives $A_{r_0,r_0}(t_1) > A_{r_0,r_0}(t_2)$, and continuity of measure at the atomless point r_0 allows us to choose $0 < a < r_0 < b$ with $A_{a,a}(t_1) > A_{b,b}(t_2)$. For $s, s' \in \left(h_g(a), h_g(b)\right)$, the sublevel intervals of h_g at s and s' contain $[0, a)$ and are contained in $[0, b)$, so

$$J_g(s, s'; t_1) - J_g(s, s'; t_2) \geq A_{a,a}(t_1) - A_{b,b}(t_2) > 0.$$

This rectangle of (s, s') has positive area by the strict monotonicity of h_g, and the integrand difference is nonnegative elsewhere, so $H_g(t_2) > H_g(t_1)$. The sufficient condition therefore applies with $F_g^{(w)} = \psi_g^{(w)}$.

Proof of statement (iii). Substituting the block model into (3) and integrating block by block gives

$$\check{d}_g^2(\check{\xi}, \check{\xi}') = \sum_{\ell=1}^{B} \beta_\ell \left[B_g^{(0)}\left(\phi(\check{\xi}), \ell\right) - B_g^{(0)}\left(\phi(\check{\xi}'), \ell\right)\right]^2 = C_g\left(\phi(\check{\xi}), \phi(\check{\xi}')\right),$$

where every entry of $B_g^{(w)}$ lies in $[0, 1]$ because each block pair has positive mass, so C_g takes values in $[0, 1]$. By condition (a), the ordering of any two source block dissimilarities agrees with that of the corresponding target block dissimilarities, so the product of differences in Assumption 3 is nonnegative.

For Assumption 4 (i), the law of $\check{d}_g^2$ takes finitely many values, so we verify branch (b). Applying condition (a) in both directions, $C_g(b, b') = C_g(\ell, \ell')$ implies $C_G(b, b') = C_G(\ell, \ell')$. Hence, $b_g\left(C_g(b, b')\right) := C_G(b, b')$ is well defined on the finite set of source block dissimilarities; setting $b_g := 0$ elsewhere gives a Borel function with values in $[0, 1]$ and $\check{d}_G^2 = b_g \circ \check{d}_g^2$.

For Assumption 4 (ii), let $\check{\pi} \in \Pi(\check{\mu}, \check{\mu})$ preserve $\check{d}_G^2$, and define its induced block coupling $S(b, \ell) := \check{\pi}\left(\phi^{-1}(b) \times \phi^{-1}(\ell)\right)$, whose marginals are both $(\beta_1, ..., \beta_B)$. If $S(b, \ell)\, S(b', \ell') > 0$, the block configuration (b, ℓ, b', ℓ') has positive $\check{\pi} \otimes \check{\pi}$-mass, and since $\check{d}_G^2$ is constant on each block pair, its preservation forces $C_G(b, b') = C_G(\ell, \ell')$. Thus, S is a shuffling matrix preserving the target block geometry, and condition (b) gives $B_g^{(w)}(b, b') = B_g^{(w)}(\ell, \ell')$ whenever $S(b, \ell)\, S(b', \ell') > 0$. Finally, the finitely many configurations with zero product mass form a $\check{\pi} \otimes \check{\pi}$-null set, so

$$\check{f}_g^{(w)}(\check{\zeta}, \check{\zeta}') = B_g^{(w)}\left(\phi(\check{\zeta}), \phi(\check{\zeta}')\right) = B_g^{(w)}\left(\phi(\check{\xi}), \phi(\check{\xi}')\right) = \check{f}_g^{(w)}(\check{\xi}, \check{\xi}') \qquad \check{\pi} \otimes \check{\pi}\text{-a.e.}$$

for every g and both w, which is Assumption 4 (ii).

□

D Additional Simulation Results

D.1 Complete Results for Design 1

Table 5: Design 1: complete oracle results

Method	RMSE	MedAE	Q95	Δ MedAE	Direction (%)	Degree MedAE	Triangle AE
$N_g \in [100, 110]$							
SCN	0.0330 (0.0009)	0.0221 (0.0006)	0.0631 (0.0017)	0.0221 (0.0006)	98.2 (0.1)	0.0199 (0.0006)	0.0038 (0.00015)
Nearest GW source	0.0924 (0.0047)	0.0649 (0.0034)	0.1752 (0.0089)	0.0649 (0.0034)	88.8 (0.7)	0.0574 (0.0029)	0.0056 (0.00021)
Best single source	0.0295 (0.0008)	0.0192 (0.0007)	0.0552 (0.0016)	0.0192 (0.0007)	98.7 (0.1)	0.0179 (0.0007)	0.0048 (0.00017)
Uniform GW barycenter	0.0942 (0.0022)	0.0672 (0.0016)	0.1798 (0.0042)	0.0672 (0.0016)	90.9 (0.3)	0.0581 (0.0014)	0.0082 (0.00026)
Degree alignment	0.0768 (0.0013)	0.0290 (0.0007)	0.1546 (0.0028)	0.0290 (0.0007)	95.8 (0.1)	0.0241 (0.0006)	0.0043 (0.00017)
No change	0.1516 (0.0006)	0.0957 (0.0004)	0.3222 (0.0013)	0.0957 (0.0004)	0.0 (0.0)	0.0430 (0.0005)	0.0031 (0.00010)
$N_g \in [200, 220]$							
SCN	0.0227 (0.0006)	0.0150 (0.0005)	0.0425 (0.0013)	0.0150 (0.0005)	99.1 (0.1)	0.0135 (0.0005)	0.0026 (0.00011)
Nearest GW source	0.0571 (0.0037)	0.0399 (0.0027)	0.1077 (0.0072)	0.0399 (0.0027)	94.0 (0.6)	0.0361 (0.0023)	0.0046 (0.00016)
Best single source	0.0208 (0.0003)	0.0133 (0.0003)	0.0388 (0.0006)	0.0133 (0.0003)	99.4 (0.0)	0.0127 (0.0004)	0.0036 (0.00012)
Uniform GW barycenter	0.0634 (0.0021)	0.0450 (0.0015)	0.1207 (0.0041)	0.0450 (0.0015)	94.6 (0.3)	0.0392 (0.0013)	0.0060 (0.00020)
Degree alignment	0.0688 (0.0011)	0.0227 (0.0006)	0.1337 (0.0025)	0.0227 (0.0006)	96.6 (0.1)	0.0184 (0.0005)	0.0031 (0.00012)
No change	0.1518 (0.0006)	0.0952 (0.0004)	0.3244 (0.0013)	0.0952 (0.0004)	0.0 (0.0)	0.0437 (0.0004)	0.0026 (0.00008)
$N_g \in [500, 550]$							
SCN	0.0128 (0.0004)	0.0080 (0.0003)	0.0234 (0.0008)	0.0080 (0.0003)	99.7 (0.0)	0.0073 (0.0003)	0.0016 (0.00006)
Nearest GW source	0.0232 (0.0019)	0.0157 (0.0014)	0.0433 (0.0037)	0.0157 (0.0014)	98.6 (0.3)	0.0144 (0.0012)	0.0028 (0.00009)
Best single source	0.0133 (0.0002)	0.0083 (0.0002)	0.0244 (0.0004)	0.0083 (0.0002)	99.8 (0.0)	0.0076 (0.0002)	0.0021 (0.00007)
Uniform GW barycenter	0.0279 (0.0013)	0.0193 (0.0010)	0.0528 (0.0026)	0.0193 (0.0010)	98.4 (0.1)	0.0171 (0.0008)	0.0033 (0.00012)
Degree alignment	0.0640 (0.0010)	0.0174 (0.0004)	0.1218 (0.0022)	0.0174 (0.0004)	97.1 (0.1)	0.0138 (0.0003)	0.0022 (0.00007)
No change	0.1519 (0.0006)	0.0954 (0.0004)	0.3250 (0.0012)	0.0954 (0.0004)	0.0 (0.0)	0.0442 (0.0003)	0.0020 (0.00006)

Notes: Entries are means across 500 replications, with Monte Carlo standard errors in parentheses.

Table 6: Design 1: feasible results, $N_g \in [100, 110]$, $h \in \{0.025, 0.05, 0.075\}$

h	Method	RMSE	MedAE	Q95	Δ MedAE	Direction (%)	Degree MedAE	Triangle AE
0.025	**SCN**	0.1359 (0.0004)	0.0891 (0.0004)	0.2837 (0.0014)	0.1016 (0.0005)	47.5 (0.2)	0.0637 (0.0003)	0.0089 (0.00020)
	Nearest GW source	0.1447 (0.0007)	0.0927 (0.0007)	0.2962 (0.0017)	0.1090 (0.0007)	47.9 (0.2)	0.0654 (0.0006)	0.0087 (0.00024)
	Best single source	0.1425 (0.0004)	0.0911 (0.0004)	0.2909 (0.0014)	0.1069 (0.0006)	48.2 (0.2)	0.0640 (0.0004)	0.0082 (0.00022)
	Uniform GW barycenter	0.1399 (0.0006)	0.0946 (0.0006)	0.2935 (0.0017)	0.1016 (0.0006)	46.0 (0.3)	0.0695 (0.0005)	0.0098 (0.00027)
	Degree alignment	0.1366 (0.0004)	0.0892 (0.0004)	0.2852 (0.0014)	0.1021 (0.0005)	47.3 (0.2)	0.0640 (0.0003)	0.0091 (0.00021)
	No change	0.1339 (0.0003)	0.0903 (0.0003)	0.2792 (0.0012)	0.0957 (0.0004)	0.0 (0.0)	0.0634 (0.0003)	0.0089 (0.00019)
0.05	**SCN**	0.1266 (0.0007)	0.0663 (0.0006)	0.2784 (0.0018)	0.1216 (0.0005)	61.7 (0.3)	0.0518 (0.0006)	0.0148 (0.00026)
	Nearest GW source	0.1736 (0.0014)	0.1066 (0.0011)	0.3631 (0.0027)	0.1443 (0.0010)	57.6 (0.2)	0.0607 (0.0010)	0.0125 (0.00029)
	Best single source	0.1625 (0.0011)	0.0995 (0.0009)	0.3406 (0.0024)	0.1364 (0.0008)	59.0 (0.2)	0.0584 (0.0009)	0.0120 (0.00029)
	Uniform GW barycenter	0.1377 (0.0011)	0.0752 (0.0009)	0.3032 (0.0023)	0.1271 (0.0006)	60.6 (0.3)	0.0621 (0.0009)	0.0156 (0.00028)
	Degree alignment	0.1403 (0.0005)	0.0780 (0.0005)	0.3021 (0.0013)	0.1178 (0.0004)	63.2 (0.3)	0.0630 (0.0005)	0.0145 (0.00027)
	No change	0.1898 (0.0010)	0.1012 (0.0007)	0.4081 (0.0017)	0.0957 (0.0004)	0.0 (0.0)	0.0580 (0.0004)	0.0101 (0.00022)
0.075	**SCN**	0.0832 (0.0008)	0.0520 (0.0006)	0.1644 (0.0016)	0.0886 (0.0004)	79.9 (0.1)	0.0342 (0.0006)	0.0065 (0.00021)
	Nearest GW source	0.1766 (0.0033)	0.1121 (0.0025)	0.3552 (0.0061)	0.1336 (0.0021)	70.8 (0.3)	0.0767 (0.0028)	0.0067 (0.00024)
	Best single source	0.1261 (0.0011)	0.0743 (0.0008)	0.2602 (0.0021)	0.1021 (0.0006)	75.1 (0.2)	0.0378 (0.0009)	0.0060 (0.00022)
	Uniform GW barycenter	0.1179 (0.0016)	0.0796 (0.0012)	0.2269 (0.0030)	0.1063 (0.0009)	77.4 (0.2)	0.0639 (0.0013)	0.0093 (0.00027)
	Degree alignment	0.1181 (0.0006)	0.0613 (0.0005)	0.2485 (0.0013)	0.0942 (0.0004)	82.0 (0.1)	0.0427 (0.0004)	0.0066 (0.00021)
	No change	0.1846 (0.0007)	0.1171 (0.0005)	0.3725 (0.0016)	0.0957 (0.0004)	0.0 (0.0)	0.0451 (0.0004)	0.0033 (0.00011)

Notes: Entries are means across 500 replications, with Monte Carlo standard errors in parentheses.

Table 7: Design 1: feasible results, $N_g \in [100, 110]$, $h \in \{0.1, 0.15, 0.2\}$

h	Method	RMSE	MedAE	Q95	Δ MedAE	Direction (%)	Degree MedAE	Triangle AE
0.1	**SCN**	0.0770 (0.0009)	0.0474 (0.0006)	0.1505 (0.0018)	0.0640 (0.0005)	84.8 (0.1)	0.0330 (0.0006)	0.0057 (0.00020)
	Nearest GW source	0.1569 (0.0037)	0.1023 (0.0029)	0.3098 (0.0069)	0.1133 (0.0025)	73.2 (0.5)	0.0824 (0.0029)	0.0070 (0.00024)
	Best single source	0.0967 (0.0009)	0.0567 (0.0007)	0.1978 (0.0019)	0.0740 (0.0006)	80.0 (0.2)	0.0359 (0.0007)	0.0060 (0.00020)
	Uniform GW barycenter	0.1281 (0.0017)	0.0873 (0.0013)	0.2437 (0.0031)	0.0929 (0.0010)	79.2 (0.3)	0.0739 (0.0013)	0.0096 (0.00027)
	Degree alignment	0.1104 (0.0007)	0.0574 (0.0005)	0.2318 (0.0015)	0.0702 (0.0004)	85.5 (0.1)	0.0407 (0.0004)	0.0061 (0.00020)
	No change	0.1388 (0.0006)	0.0913 (0.0005)	0.2762 (0.0015)	0.0957 (0.0004)	0.0 (0.0)	0.0453 (0.0004)	0.0038 (0.00012)
0.15	**SCN**	0.0861 (0.0007)	0.0505 (0.0005)	0.1718 (0.0017)	0.0711 (0.0004)	77.8 (0.2)	0.0345 (0.0005)	0.0063 (0.00020)
	Nearest GW source	0.1572 (0.0033)	0.1024 (0.0026)	0.3110 (0.0062)	0.1151 (0.0020)	65.8 (0.5)	0.0849 (0.0027)	0.0064 (0.00023)
	Best single source	0.0971 (0.0008)	0.0554 (0.0005)	0.1998 (0.0019)	0.0800 (0.0004)	73.0 (0.2)	0.0374 (0.0006)	0.0062 (0.00021)
	Uniform GW barycenter	0.1350 (0.0015)	0.0902 (0.0012)	0.2670 (0.0030)	0.0908 (0.0009)	71.4 (0.4)	0.0766 (0.0012)	0.0099 (0.00027)
	Degree alignment	0.1110 (0.0006)	0.0598 (0.0004)	0.2356 (0.0016)	0.0705 (0.0003)	78.4 (0.2)	0.0430 (0.0003)	0.0071 (0.00020)
	No change	0.1057 (0.0004)	0.0686 (0.0004)	0.2030 (0.0009)	0.0957 (0.0004)	0.0 (0.0)	0.0492 (0.0004)	0.0060 (0.00015)
0.2	**SCN**	0.0982 (0.0006)	0.0582 (0.0004)	0.2026 (0.0016)	0.0843 (0.0005)	67.9 (0.3)	0.0386 (0.0004)	0.0073 (0.00020)
	Nearest GW source	0.1418 (0.0026)	0.0901 (0.0021)	0.2867 (0.0049)	0.1087 (0.0014)	59.5 (0.4)	0.0696 (0.0021)	0.0071 (0.00022)
	Best single source	0.1060 (0.0007)	0.0614 (0.0004)	0.2191 (0.0018)	0.0908 (0.0006)	63.7 (0.3)	0.0412 (0.0005)	0.0070 (0.00021)
	Uniform GW barycenter	0.1328 (0.0012)	0.0874 (0.0010)	0.2766 (0.0027)	0.0916 (0.0007)	63.4 (0.3)	0.0706 (0.0010)	0.0099 (0.00027)
	Degree alignment	0.1198 (0.0005)	0.0695 (0.0004)	0.2570 (0.0015)	0.0794 (0.0005)	68.6 (0.3)	0.0503 (0.0004)	0.0084 (0.00019)
	No change	0.1070 (0.0004)	0.0681 (0.0004)	0.2099 (0.0011)	0.0957 (0.0004)	0.0 (0.0)	0.0552 (0.0004)	0.0076 (0.00017)

Notes: Entries are means across 500 replications, with Monte Carlo standard errors in parentheses.

Table 8: Design 1: feasible results, $N_g \in [200, 220]$, $h \in \{0.025, 0.05, 0.075\}$

h	Method	RMSE	MedAE	Q95	Δ MedAE	Direction (%)	Degree MedAE	Triangle AE
0.025	**SCN**	0.1297 (0.0003)	0.0841 (0.0003)	0.2708 (0.0010)	0.1001 (0.0004)	48.7 (0.2)	0.0587 (0.0003)	0.0088 (0.00015)
	Nearest GW source	0.1352 (0.0005)	0.0859 (0.0005)	0.2773 (0.0014)	0.1049 (0.0006)	49.1 (0.2)	0.0596 (0.0004)	0.0084 (0.00020)
	Best single source	0.1343 (0.0004)	0.0853 (0.0003)	0.2754 (0.0011)	0.1040 (0.0005)	49.4 (0.2)	0.0592 (0.0003)	0.0082 (0.00017)
	Uniform GW barycenter	0.1294 (0.0003)	0.0859 (0.0004)	0.2709 (0.0012)	0.0975 (0.0005)	47.9 (0.3)	0.0603 (0.0003)	0.0089 (0.00020)
	Degree alignment	0.1297 (0.0003)	0.0843 (0.0003)	0.2713 (0.0010)	0.0998 (0.0004)	48.6 (0.2)	0.0591 (0.0003)	0.0089 (0.00015)
	No change	0.1325 (0.0002)	0.0917 (0.0003)	0.2758 (0.0009)	0.0952 (0.0004)	0.0 (0.0)	0.0632 (0.0002)	0.0088 (0.00014)
0.05	**SCN**	0.0752 (0.0006)	0.0452 (0.0005)	0.1518 (0.0012)	0.1004 (0.0004)	77.1 (0.1)	0.0258 (0.0005)	0.0065 (0.00018)
	Nearest GW source	0.1726 (0.0018)	0.0996 (0.0015)	0.3622 (0.0031)	0.1321 (0.0012)	70.2 (0.2)	0.0373 (0.0017)	0.0048 (0.00017)
	Best single source	0.1503 (0.0007)	0.0846 (0.0004)	0.3187 (0.0015)	0.1201 (0.0004)	71.4 (0.1)	0.0273 (0.0004)	0.0054 (0.00017)
	Uniform GW barycenter	0.0827 (0.0011)	0.0515 (0.0008)	0.1650 (0.0020)	0.1042 (0.0005)	76.7 (0.1)	0.0328 (0.0009)	0.0074 (0.00021)
	Degree alignment	0.1114 (0.0006)	0.0546 (0.0004)	0.2351 (0.0012)	0.1070 (0.0003)	78.9 (0.1)	0.0358 (0.0003)	0.0068 (0.00018)
	No change	0.2135 (0.0005)	0.1329 (0.0004)	0.4325 (0.0011)	0.0952 (0.0004)	0.0 (0.0)	0.0441 (0.0003)	0.0030 (0.00010)
0.075	**SCN**	0.0515 (0.0006)	0.0325 (0.0004)	0.1004 (0.0012)	0.0464 (0.0003)	90.8 (0.1)	0.0221 (0.0005)	0.0039 (0.00013)
	Nearest GW source	0.1252 (0.0040)	0.0831 (0.0030)	0.2450 (0.0076)	0.0920 (0.0027)	79.5 (0.5)	0.0656 (0.0030)	0.0048 (0.00016)
	Best single source	0.0674 (0.0006)	0.0405 (0.0004)	0.1365 (0.0012)	0.0538 (0.0004)	86.9 (0.1)	0.0236 (0.0005)	0.0042 (0.00014)
	Uniform GW barycenter	0.1043 (0.0019)	0.0724 (0.0014)	0.1975 (0.0036)	0.0768 (0.0012)	85.3 (0.3)	0.0610 (0.0014)	0.0081 (0.00021)
	Degree alignment	0.0948 (0.0007)	0.0441 (0.0004)	0.1989 (0.0015)	0.0538 (0.0004)	90.4 (0.1)	0.0310 (0.0003)	0.0042 (0.00015)
	No change	0.1443 (0.0006)	0.0929 (0.0004)	0.2929 (0.0014)	0.0952 (0.0004)	0.0 (0.0)	0.0437 (0.0004)	0.0029 (0.00009)

Notes: Entries are means across 500 replications, with Monte Carlo standard errors in parentheses.

Table 9: Design 1: feasible results, $N_g \in [200, 220]$, $h \in \{0.1, 0.15, 0.2\}$

h	Method	RMSE	MedAE	Q95	Δ MedAE	Direction (%)	Degree MedAE	Triangle AE
0.1	**SCN**	0.0575 (0.0009)	0.0356 (0.0006)	0.1094 (0.0017)	0.0470 (0.0005)	89.8 (0.1)	0.0242 (0.0006)	0.0042 (0.00014)
	Nearest GW source	0.1240 (0.0041)	0.0822 (0.0030)	0.2407 (0.0076)	0.0903 (0.0026)	78.8 (0.6)	0.0684 (0.0031)	0.0046 (0.00017)
	Best single source	0.0656 (0.0008)	0.0390 (0.0006)	0.1316 (0.0015)	0.0530 (0.0005)	86.4 (0.2)	0.0245 (0.0006)	0.0041 (0.00015)
	Uniform GW barycenter	0.1135 (0.0021)	0.0782 (0.0016)	0.2127 (0.0039)	0.0770 (0.0013)	83.3 (0.4)	0.0670 (0.0015)	0.0083 (0.00021)
	Degree alignment	0.0917 (0.0007)	0.0439 (0.0003)	0.1919 (0.0015)	0.0529 (0.0003)	89.0 (0.1)	0.0307 (0.0003)	0.0044 (0.00015)
	No change	0.1213 (0.0006)	0.0818 (0.0004)	0.2411 (0.0013)	0.0952 (0.0004)	0.0 (0.0)	0.0452 (0.0004)	0.0038 (0.00010)
0.15	**SCN**	0.0739 (0.0010)	0.0446 (0.0007)	0.1411 (0.0021)	0.0702 (0.0006)	78.7 (0.2)	0.0283 (0.0008)	0.0056 (0.00014)
	Nearest GW source	0.1312 (0.0036)	0.0867 (0.0028)	0.2536 (0.0068)	0.1032 (0.0020)	68.5 (0.5)	0.0705 (0.0029)	0.0046 (0.00016)
	Best single source	0.0779 (0.0010)	0.0459 (0.0007)	0.1534 (0.0020)	0.0751 (0.0005)	75.1 (0.2)	0.0286 (0.0007)	0.0050 (0.00016)
	Uniform GW barycenter	0.1212 (0.0019)	0.0813 (0.0015)	0.2355 (0.0037)	0.0827 (0.0010)	73.4 (0.4)	0.0681 (0.0015)	0.0086 (0.00021)
	Degree alignment	0.0991 (0.0006)	0.0521 (0.0003)	0.2095 (0.0013)	0.0660 (0.0003)	78.2 (0.2)	0.0363 (0.0003)	0.0060 (0.00014)
	No change	0.0975 (0.0003)	0.0622 (0.0004)	0.1834 (0.0006)	0.0952 (0.0004)	0.0 (0.0)	0.0510 (0.0004)	0.0062 (0.00011)
0.2	**SCN**	0.0897 (0.0009)	0.0553 (0.0007)	0.1800 (0.0021)	0.0857 (0.0005)	68.0 (0.2)	0.0329 (0.0007)	0.0069 (0.00014)
	Nearest GW source	0.1141 (0.0023)	0.0725 (0.0018)	0.2273 (0.0047)	0.0984 (0.0012)	61.7 (0.4)	0.0516 (0.0020)	0.0064 (0.00016)
	Best single source	0.0933 (0.0009)	0.0566 (0.0006)	0.1868 (0.0021)	0.0883 (0.0005)	65.1 (0.2)	0.0347 (0.0007)	0.0064 (0.00016)
	Uniform GW barycenter	0.1094 (0.0015)	0.0716 (0.0012)	0.2233 (0.0034)	0.0876 (0.0007)	66.7 (0.3)	0.0520 (0.0013)	0.0081 (0.00021)
	Degree alignment	0.1110 (0.0004)	0.0653 (0.0003)	0.2367 (0.0011)	0.0790 (0.0004)	67.6 (0.2)	0.0451 (0.0003)	0.0077 (0.00014)
	No change	0.1060 (0.0003)	0.0691 (0.0004)	0.2069 (0.0009)	0.0952 (0.0004)	0.0 (0.0)	0.0572 (0.0003)	0.0078 (0.00012)

Notes: Entries are means across 500 replications, with Monte Carlo standard errors in parentheses.

Table 10: Design 1: feasible results, $N_g \in [500, 550]$, $h \in \{0.025, 0.05, 0.075\}$

h	Method	RMSE	MedAE	Q95	Δ MedAE	Direction (%)	Degree MedAE	Triangle AE
0.025	**SCN**	0.1089 (0.0002)	0.0552 (0.0002)	0.2461 (0.0007)	0.1104 (0.0004)	54.8 (0.2)	0.0390 (0.0002)	0.0107 (0.00012)
	Nearest GW source	0.1236 (0.0003)	0.0687 (0.0003)	0.2701 (0.0009)	0.1134 (0.0005)	52.4 (0.2)	0.0433 (0.0003)	0.0101 (0.00016)
	Best single source	0.1218 (0.0003)	0.0682 (0.0003)	0.2652 (0.0009)	0.1106 (0.0005)	53.1 (0.2)	0.0433 (0.0003)	0.0099 (0.00015)
	Uniform GW barycenter	0.1086 (0.0002)	0.0546 (0.0002)	0.2470 (0.0007)	0.1115 (0.0005)	54.7 (0.2)	0.0388 (0.0002)	0.0109 (0.00012)
	Degree alignment	0.1204 (0.0003)	0.0652 (0.0003)	0.2608 (0.0007)	0.1044 (0.0004)	54.8 (0.2)	0.0468 (0.0002)	0.0111 (0.00012)
	No change	0.1416 (0.0003)	0.0758 (0.0003)	0.3228 (0.0009)	0.0954 (0.0004)	0.0 (0.0)	0.0613 (0.0002)	0.0104 (0.00010)
0.05	**SCN**	0.0290 (0.0004)	0.0185 (0.0003)	0.0571 (0.0007)	0.0316 (0.0002)	95.0 (0.0)	0.0117 (0.0003)	0.0021 (0.00007)
	Nearest GW source	0.0652 (0.0027)	0.0419 (0.0020)	0.1294 (0.0052)	0.0515 (0.0018)	89.5 (0.4)	0.0274 (0.0018)	0.0028 (0.00009)
	Best single source	0.0448 (0.0001)	0.0273 (0.0001)	0.0905 (0.0002)	0.0380 (0.0001)	92.3 (0.1)	0.0138 (0.0001)	0.0024 (0.00008)
	Uniform GW barycenter	0.0509 (0.0017)	0.0345 (0.0012)	0.0983 (0.0032)	0.0437 (0.0010)	93.1 (0.2)	0.0274 (0.0011)	0.0045 (0.00014)
	Degree alignment	0.0798 (0.0007)	0.0311 (0.0003)	0.1657 (0.0016)	0.0395 (0.0003)	93.7 (0.0)	0.0221 (0.0002)	0.0027 (0.00009)
	No change	0.1506 (0.0006)	0.0946 (0.0004)	0.3124 (0.0013)	0.0954 (0.0004)	0.0 (0.0)	0.0435 (0.0003)	0.0021 (0.00006)
0.075	**SCN**	0.0341 (0.0006)	0.0215 (0.0004)	0.0645 (0.0012)	0.0294 (0.0003)	95.1 (0.1)	0.0135 (0.0004)	0.0024 (0.00007)
	Nearest GW source	0.0684 (0.0032)	0.0449 (0.0023)	0.1334 (0.0060)	0.0520 (0.0021)	89.5 (0.5)	0.0338 (0.0022)	0.0029 (0.00009)
	Best single source	0.0414 (0.0006)	0.0253 (0.0004)	0.0823 (0.0011)	0.0341 (0.0004)	93.3 (0.1)	0.0148 (0.0004)	0.0025 (0.00008)
	Uniform GW barycenter	0.0646 (0.0019)	0.0441 (0.0014)	0.1216 (0.0036)	0.0460 (0.0012)	92.1 (0.3)	0.0354 (0.0013)	0.0054 (0.00015)
	Degree alignment	0.0743 (0.0007)	0.0306 (0.0003)	0.1526 (0.0016)	0.0353 (0.0002)	93.8 (0.0)	0.0209 (0.0002)	0.0028 (0.00009)
	No change	0.1300 (0.0006)	0.0859 (0.0004)	0.2632 (0.0013)	0.0954 (0.0004)	0.0 (0.0)	0.0449 (0.0003)	0.0028 (0.00007)

Notes: Entries are means across 500 replications, with Monte Carlo standard errors in parentheses.

Table 11: Design 1: feasible results, $N_g \in [500, 550]$, $h \in \{0.1, 0.15, 0.2\}$

h	Method	RMSE	MedAE	Q95	Δ MedAE	Direction (%)	Degree MedAE	Triangle AE
0.1	**SCN**	0.0427 (0.0006)	0.0267 (0.0004)	0.0770 (0.0011)	0.0391 (0.0003)	91.6 (0.1)	0.0154 (0.0004)	0.0031 (0.00008)
	Nearest GW source	0.0797 (0.0034)	0.0527 (0.0025)	0.1518 (0.0065)	0.0636 (0.0021)	85.0 (0.5)	0.0411 (0.0026)	0.0032 (0.00010)
	Best single source	0.0465 (0.0006)	0.0286 (0.0004)	0.0890 (0.0011)	0.0436 (0.0003)	89.7 (0.1)	0.0161 (0.0004)	0.0028 (0.00009)
	Uniform GW barycenter	0.0815 (0.0020)	0.0553 (0.0015)	0.1510 (0.0038)	0.0554 (0.0011)	87.6 (0.3)	0.0450 (0.0015)	0.0064 (0.00015)
	Degree alignment	0.0757 (0.0007)	0.0346 (0.0002)	0.1545 (0.0015)	0.0426 (0.0002)	89.7 (0.1)	0.0227 (0.0002)	0.0033 (0.00009)
	No change	0.1110 (0.0005)	0.0759 (0.0004)	0.2216 (0.0011)	0.0954 (0.0004)	0.0 (0.0)	0.0467 (0.0003)	0.0040 (0.00007)
0.15	**SCN**	0.0660 (0.0009)	0.0420 (0.0007)	0.1224 (0.0020)	0.0748 (0.0005)	78.9 (0.2)	0.0225 (0.0007)	0.0051 (0.00008)
	Nearest GW source	0.0894 (0.0027)	0.0580 (0.0020)	0.1695 (0.0052)	0.0866 (0.0013)	73.4 (0.4)	0.0402 (0.0022)	0.0045 (0.00010)
	Best single source	0.0673 (0.0009)	0.0419 (0.0007)	0.1270 (0.0019)	0.0767 (0.0005)	76.6 (0.2)	0.0226 (0.0007)	0.0046 (0.00010)
	Uniform GW barycenter	0.0871 (0.0018)	0.0573 (0.0014)	0.1663 (0.0037)	0.0766 (0.0008)	77.5 (0.3)	0.0414 (0.0015)	0.0065 (0.00014)
	Degree alignment	0.0909 (0.0005)	0.0484 (0.0003)	0.1919 (0.0013)	0.0684 (0.0002)	76.8 (0.1)	0.0317 (0.0002)	0.0057 (0.00008)
	No change	0.0937 (0.0002)	0.0577 (0.0003)	0.1712 (0.0005)	0.0954 (0.0004)	0.0 (0.0)	0.0537 (0.0003)	0.0066 (0.00007)
0.2	**SCN**	0.0825 (0.0004)	0.0537 (0.0003)	0.1618 (0.0010)	0.0859 (0.0003)	69.7 (0.1)	0.0273 (0.0004)	0.0065 (0.00007)
	Nearest GW source	0.0865 (0.0007)	0.0549 (0.0005)	0.1706 (0.0015)	0.0884 (0.0005)	66.4 (0.2)	0.0302 (0.0006)	0.0066 (0.00010)
	Best single source	0.0845 (0.0004)	0.0540 (0.0003)	0.1658 (0.0011)	0.0871 (0.0004)	67.3 (0.2)	0.0290 (0.0004)	0.0063 (0.00010)
	Uniform GW barycenter	0.0840 (0.0005)	0.0551 (0.0004)	0.1646 (0.0013)	0.0849 (0.0004)	70.2 (0.1)	0.0295 (0.0004)	0.0066 (0.00013)
	Degree alignment	0.1036 (0.0004)	0.0625 (0.0003)	0.2218 (0.0010)	0.0804 (0.0003)	66.2 (0.2)	0.0410 (0.0002)	0.0072 (0.00008)
	No change	0.1079 (0.0002)	0.0718 (0.0002)	0.2147 (0.0006)	0.0954 (0.0004)	0.0 (0.0)	0.0588 (0.0002)	0.0080 (0.00008)

Notes: Entries are means across 500 replications, with Monte Carlo standard errors in parentheses.

D.2 Complete Results for Design 2

Table 12: Design 4: complete oracle results

Method	RMSE	MedAE	Q95	Δ MedAE	Direction (%)	Degree MedAE	Triangle AE
$N_g \in [100, 110]$							
SCN	0.0461 (0.0005)	0.0285 (0.0004)	0.0913 (0.0010)	0.0285 (0.0004)	78.6 (0.4)	0.0135 (0.0002)	0.00260 (0.00009)
Nearest GW source	0.0682 (0.0013)	0.0409 (0.0013)	0.1379 (0.0021)	0.0409 (0.0013)	74.3 (0.5)	0.0147 (0.0003)	0.00279 (0.00011)
Best single source	0.0591 (0.0006)	0.0325 (0.0005)	0.1234 (0.0011)	0.0325 (0.0005)	75.8 (0.5)	0.0137 (0.0003)	0.00290 (0.00011)
Uniform GW barycenter	0.0680 (0.0011)	0.0508 (0.0012)	0.1248 (0.0018)	0.0508 (0.0012)	73.3 (0.7)	0.0181 (0.0003)	0.00318 (0.00010)
Degree alignment	0.1161 (0.0005)	0.0925 (0.0006)	0.2081 (0.0007)	0.0925 (0.0006)	72.3 (1.1)	0.0106 (0.0002)	0.00343 (0.00012)
No change	0.0279 (0.0006)	0.0197 (0.0005)	0.0515 (0.0010)	0.0197 (0.0005)	0.0 (0.0)	0.0159 (0.0005)	0.00692 (0.00021)
$N_g \in [200, 220]$							
SCN	0.0362 (0.0003)	0.0204 (0.0002)	0.0734 (0.0007)	0.0204 (0.0002)	84.0 (0.3)	0.0095 (0.0002)	0.00188 (0.00006)
Nearest GW source	0.0506 (0.0005)	0.0249 (0.0004)	0.1077 (0.0012)	0.0249 (0.0004)	80.6 (0.4)	0.0100 (0.0002)	0.00220 (0.00009)
Best single source	0.0475 (0.0004)	0.0227 (0.0003)	0.1014 (0.0010)	0.0227 (0.0003)	81.8 (0.4)	0.0095 (0.0002)	0.00216 (0.00008)
Uniform GW barycenter	0.0481 (0.0005)	0.0326 (0.0004)	0.0926 (0.0009)	0.0326 (0.0004)	75.5 (0.4)	0.0154 (0.0003)	0.00307 (0.00009)
Degree alignment	0.1164 (0.0005)	0.0932 (0.0006)	0.2083 (0.0007)	0.0932 (0.0006)	72.4 (1.1)	0.0096 (0.0002)	0.00317 (0.00010)
No change	0.0279 (0.0006)	0.0196 (0.0004)	0.0516 (0.0010)	0.0196 (0.0004)	0.0 (0.0)	0.0161 (0.0004)	0.00696 (0.00021)
$N_g \in [500, 550]$							
SCN	0.0277 (0.0002)	0.0137 (0.0002)	0.0572 (0.0005)	0.0137 (0.0002)	89.7 (0.2)	0.0058 (0.0001)	0.00119 (0.00004)
Nearest GW source	0.0380 (0.0004)	0.0157 (0.0003)	0.0809 (0.0009)	0.0157 (0.0003)	87.3 (0.3)	0.0064 (0.0002)	0.00142 (0.00006)
Best single source	0.0367 (0.0003)	0.0149 (0.0002)	0.0778 (0.0009)	0.0149 (0.0002)	87.7 (0.3)	0.0062 (0.0001)	0.00143 (0.00006)
Uniform GW barycenter	0.0393 (0.0004)	0.0264 (0.0003)	0.0757 (0.0007)	0.0264 (0.0003)	77.4 (0.4)	0.0129 (0.0002)	0.00312 (0.00008)
Degree alignment	0.1159 (0.0005)	0.0919 (0.0006)	0.2088 (0.0007)	0.0919 (0.0006)	72.8 (1.1)	0.0088 (0.0002)	0.00301 (0.00009)
No change	0.0280 (0.0006)	0.0197 (0.0004)	0.0518 (0.0010)	0.0197 (0.0004)	0.0 (0.0)	0.0163 (0.0005)	0.00696 (0.00021)

Notes: Entries are means across 500 replications, with Monte Carlo standard errors in parentheses.

Table 13: Design 4: feasible results, $N_g \in [100, 110]$, $h \in \{0.025, 0.05, 0.075\}$

h	Method	RMSE	MedAE	Q95	Δ MedAE	Direction (%)	Degree MedAE	Triangle AE
0.025	**SCN**	0.1205 (0.0003)	0.1035 (0.0005)	0.2042 (0.0003)	0.0301 (0.0003)	55.4 (0.6)	0.0216 (0.0003)	0.00357 (0.00012)
	Nearest GW source	0.1246 (0.0003)	0.1031 (0.0005)	0.2173 (0.0004)	0.0372 (0.0002)	56.2 (0.5)	0.0252 (0.0002)	0.00461 (0.00016)
	Best single source	0.1244 (0.0003)	0.1034 (0.0005)	0.2164 (0.0003)	0.0370 (0.0002)	56.9 (0.5)	0.0246 (0.0002)	0.00394 (0.00014)
	Uniform GW barycenter	0.1200 (0.0003)	0.1037 (0.0005)	0.2023 (0.0003)	0.0289 (0.0003)	55.4 (0.7)	0.0212 (0.0003)	0.00338 (0.00011)
	Degree alignment	0.1215 (0.0003)	0.1033 (0.0005)	0.2075 (0.0004)	0.0317 (0.0003)	55.3 (0.6)	0.0222 (0.0003)	0.00388 (0.00014)
	No change	0.1230 (0.0003)	0.1020 (0.0005)	0.2137 (0.0004)	0.0197 (0.0005)	0.0 (0.0)	0.0235 (0.0003)	0.00718 (0.00023)
0.05	**SCN**	0.1209 (0.0003)	0.1034 (0.0005)	0.2056 (0.0004)	0.0310 (0.0003)	55.6 (0.6)	0.0217 (0.0003)	0.00356 (0.00012)
	Nearest GW source	0.1254 (0.0003)	0.1032 (0.0005)	0.2190 (0.0004)	0.0377 (0.0003)	56.2 (0.5)	0.0253 (0.0002)	0.00474 (0.00017)
	Best single source	0.1254 (0.0003)	0.1035 (0.0005)	0.2184 (0.0004)	0.0376 (0.0002)	56.7 (0.5)	0.0248 (0.0002)	0.00414 (0.00015)
	Uniform GW barycenter	0.1201 (0.0003)	0.1037 (0.0005)	0.2027 (0.0003)	0.0295 (0.0003)	55.2 (0.6)	0.0213 (0.0003)	0.00336 (0.00011)
	Degree alignment	0.1217 (0.0003)	0.1034 (0.0005)	0.2077 (0.0004)	0.0319 (0.0003)	55.7 (0.6)	0.0220 (0.0003)	0.00359 (0.00013)
	No change	0.1237 (0.0004)	0.1021 (0.0005)	0.2148 (0.0004)	0.0197 (0.0005)	0.0 (0.0)	0.0235 (0.0003)	0.00718 (0.00023)
0.075	**SCN**	0.1439 (0.0004)	0.1048 (0.0004)	0.2714 (0.0006)	0.1839 (0.0007)	54.5 (0.2)	0.0233 (0.0003)	0.00297 (0.00010)
	Nearest GW source	0.2521 (0.0008)	0.1781 (0.0007)	0.4838 (0.0014)	0.2259 (0.0006)	54.3 (0.2)	0.0374 (0.0003)	0.00517 (0.00020)
	Best single source	0.2446 (0.0009)	0.1721 (0.0007)	0.4704 (0.0016)	0.2215 (0.0006)	53.7 (0.2)	0.0370 (0.0003)	0.00592 (0.00023)
	Uniform GW barycenter	0.1400 (0.0003)	0.1052 (0.0004)	0.2596 (0.0005)	0.1842 (0.0007)	54.9 (0.2)	0.0227 (0.0003)	0.00283 (0.00010)
	Degree alignment	0.1332 (0.0003)	0.1038 (0.0004)	0.2418 (0.0004)	0.1810 (0.0007)	55.6 (0.2)	0.0204 (0.0002)	0.00282 (0.00009)
	No change	0.2394 (0.0007)	0.1710 (0.0007)	0.4654 (0.0013)	0.0197 (0.0005)	0.0 (0.0)	0.0324 (0.0002)	0.00713 (0.00023)

Notes: Entries are means across 500 replications, with Monte Carlo standard errors in parentheses.

Table 14: Design 4: feasible results, $N_g \in [100, 110]$, $h \in \{0.1, 0.15, 0.2\}$

h	Method	RMSE	MedAE	Q95	Δ MedAE	Direction (%)	Degree MedAE	Triangle AE
0.1	**SCN**	0.1236 (0.0004)	0.0960 (0.0005)	0.2246 (0.0007)	0.0794 (0.0004)	58.7 (0.4)	0.0233 (0.0002)	0.00345 (0.00013)
	Nearest GW source	0.1560 (0.0006)	0.1113 (0.0005)	0.2962 (0.0012)	0.1016 (0.0005)	58.0 (0.4)	0.0352 (0.0003)	0.00472 (0.00018)
	Best single source	0.1521 (0.0006)	0.1065 (0.0005)	0.2916 (0.0013)	0.0983 (0.0004)	57.6 (0.3)	0.0336 (0.0003)	0.00551 (0.00020)
	Uniform GW barycenter	0.1225 (0.0004)	0.1007 (0.0005)	0.2146 (0.0004)	0.0799 (0.0004)	59.3 (0.4)	0.0225 (0.0003)	0.00317 (0.00010)
	Degree alignment	0.1237 (0.0003)	0.1023 (0.0005)	0.2160 (0.0004)	0.0777 (0.0004)	61.1 (0.4)	0.0206 (0.0002)	0.00318 (0.00011)
	No change	0.1259 (0.0004)	0.0850 (0.0003)	0.2450 (0.0007)	0.0197 (0.0005)	0.0 (0.0)	0.0310 (0.0002)	0.00792 (0.00025)
0.15	**SCN**	0.1209 (0.0003)	0.1021 (0.0005)	0.2078 (0.0004)	0.0382 (0.0003)	59.8 (0.5)	0.0213 (0.0002)	0.00353 (0.00011)
	Nearest GW source	0.1300 (0.0003)	0.1035 (0.0004)	0.2327 (0.0006)	0.0493 (0.0003)	58.9 (0.4)	0.0285 (0.0003)	0.00487 (0.00019)
	Best single source	0.1286 (0.0003)	0.1028 (0.0004)	0.2297 (0.0005)	0.0485 (0.0003)	59.4 (0.4)	0.0276 (0.0002)	0.00440 (0.00016)
	Uniform GW barycenter	0.1203 (0.0003)	0.1030 (0.0005)	0.2042 (0.0003)	0.0371 (0.0003)	59.6 (0.5)	0.0208 (0.0003)	0.00321 (0.00010)
	Degree alignment	0.1213 (0.0003)	0.1030 (0.0005)	0.2074 (0.0004)	0.0374 (0.0002)	60.5 (0.5)	0.0210 (0.0002)	0.00334 (0.00011)
	No change	0.1137 (0.0004)	0.0898 (0.0005)	0.2044 (0.0005)	0.0197 (0.0005)	0.0 (0.0)	0.0259 (0.0003)	0.00743 (0.00024)
0.2	**SCN**	0.1205 (0.0003)	0.1034 (0.0005)	0.2046 (0.0003)	0.0317 (0.0002)	58.2 (0.5)	0.0210 (0.0002)	0.00335 (0.00011)
	Nearest GW source	0.1262 (0.0003)	0.1032 (0.0005)	0.2220 (0.0004)	0.0404 (0.0002)	57.7 (0.5)	0.0263 (0.0002)	0.00471 (0.00018)
	Best single source	0.1258 (0.0003)	0.1033 (0.0005)	0.2205 (0.0004)	0.0400 (0.0002)	58.7 (0.4)	0.0256 (0.0002)	0.00387 (0.00015)
	Uniform GW barycenter	0.1200 (0.0003)	0.1036 (0.0005)	0.2027 (0.0003)	0.0306 (0.0002)	57.8 (0.6)	0.0208 (0.0003)	0.00328 (0.00011)
	Degree alignment	0.1214 (0.0003)	0.1032 (0.0005)	0.2075 (0.0004)	0.0329 (0.0002)	58.0 (0.5)	0.0218 (0.0003)	0.00360 (0.00013)
	No change	0.1183 (0.0004)	0.0966 (0.0005)	0.2079 (0.0004)	0.0197 (0.0005)	0.0 (0.0)	0.0243 (0.0003)	0.00724 (0.00023)

Notes: Entries are means across 500 replications, with Monte Carlo standard errors in parentheses.

Table 15: Design 4: feasible results, $N_g \in [200, 220]$, $h \in \{0.025, 0.05, 0.075\}$

h	Method	RMSE	MedAE	Q95	Δ MedAE	Direction (%)	Degree MedAE	Triangle AE
0.025	**SCN**	0.1195 (0.0003)	0.1036 (0.0006)	0.2010 (0.0003)	0.0247 (0.0003)	58.1 (0.8)	0.0195 (0.0002)	0.00304 (0.00010)
	Nearest GW source	0.1217 (0.0003)	0.1035 (0.0005)	0.2081 (0.0003)	0.0292 (0.0003)	59.4 (0.6)	0.0207 (0.0002)	0.00314 (0.00013)
	Best single source	0.1215 (0.0003)	0.1038 (0.0005)	0.2070 (0.0003)	0.0288 (0.0003)	60.7 (0.5)	0.0198 (0.0002)	0.00223 (0.00008)
	Uniform GW barycenter	0.1193 (0.0003)	0.1037 (0.0006)	0.2000 (0.0003)	0.0239 (0.0003)	57.4 (0.8)	0.0196 (0.0003)	0.00305 (0.00009)
	Degree alignment	0.1199 (0.0003)	0.1034 (0.0006)	0.2023 (0.0003)	0.0254 (0.0003)	57.9 (0.8)	0.0198 (0.0002)	0.00326 (0.00011)
	No change	0.1205 (0.0003)	0.1019 (0.0006)	0.2055 (0.0004)	0.0196 (0.0004)	0.0 (0.0)	0.0205 (0.0003)	0.00669 (0.00021)
0.05	**SCN**	0.1254 (0.0003)	0.1037 (0.0005)	0.2187 (0.0004)	0.0481 (0.0008)	55.2 (0.3)	0.0200 (0.0003)	0.00235 (0.00008)
	Nearest GW source	0.1545 (0.0008)	0.1127 (0.0006)	0.2971 (0.0022)	0.0608 (0.0013)	54.6 (0.4)	0.0244 (0.0003)	0.00451 (0.00018)
	Best single source	0.1424 (0.0006)	0.1087 (0.0005)	0.2614 (0.0017)	0.0537 (0.0009)	54.5 (0.3)	0.0236 (0.0002)	0.00363 (0.00013)
	Uniform GW barycenter	0.1244 (0.0003)	0.1034 (0.0005)	0.2162 (0.0003)	0.0482 (0.0007)	54.5 (0.3)	0.0212 (0.0003)	0.00272 (0.00009)
	Degree alignment	0.1220 (0.0003)	0.1033 (0.0005)	0.2092 (0.0003)	0.0401 (0.0008)	56.7 (0.4)	0.0180 (0.0003)	0.00225 (0.00008)
	No change	0.1530 (0.0010)	0.1116 (0.0007)	0.2923 (0.0026)	0.0196 (0.0004)	0.0 (0.0)	0.0213 (0.0003)	0.00668 (0.00022)
0.075	**SCN**	0.0974 (0.0006)	0.0720 (0.0006)	0.1819 (0.0011)	0.0658 (0.0005)	61.7 (0.4)	0.0186 (0.0002)	0.00256 (0.00009)
	Nearest GW source	0.1306 (0.0009)	0.0895 (0.0008)	0.2526 (0.0015)	0.0877 (0.0006)	60.4 (0.4)	0.0273 (0.0002)	0.00286 (0.00011)
	Best single source	0.1240 (0.0007)	0.0827 (0.0005)	0.2431 (0.0012)	0.0837 (0.0004)	59.7 (0.3)	0.0256 (0.0002)	0.00336 (0.00014)
	Uniform GW barycenter	0.1013 (0.0007)	0.0788 (0.0007)	0.1835 (0.0009)	0.0691 (0.0005)	61.6 (0.4)	0.0202 (0.0003)	0.00291 (0.00009)
	Degree alignment	0.1229 (0.0004)	0.1017 (0.0005)	0.2142 (0.0005)	0.0822 (0.0003)	66.2 (0.6)	0.0180 (0.0002)	0.00317 (0.00010)
	No change	0.1042 (0.0004)	0.0688 (0.0003)	0.2044 (0.0008)	0.0196 (0.0004)	0.0 (0.0)	0.0244 (0.0003)	0.00736 (0.00023)

Notes: Entries are means across 500 replications, with Monte Carlo standard errors in parentheses.

Table 16: Design 4: feasible results, $N_g \in [200, 220]$, $h \in \{0.1, 0.15, 0.2\}$

h	Method	RMSE	MedAE	Q95	Δ MedAE	Direction (%)	Degree MedAE	Triangle AE
0.1	**SCN**	0.1103 (0.0004)	0.0903 (0.0005)	0.1933 (0.0006)	0.0381 (0.0003)	64.9 (0.5)	0.0184 (0.0002)	0.00254 (0.00009)
	Nearest GW source	0.1223 (0.0005)	0.0949 (0.0005)	0.2220 (0.0009)	0.0510 (0.0004)	62.6 (0.5)	0.0246 (0.0003)	0.00315 (0.00011)
	Best single source	0.1169 (0.0004)	0.0902 (0.0004)	0.2130 (0.0007)	0.0470 (0.0003)	63.1 (0.4)	0.0230 (0.0002)	0.00308 (0.00013)
	Uniform GW barycenter	0.1140 (0.0004)	0.0962 (0.0006)	0.1957 (0.0004)	0.0406 (0.0003)	63.9 (0.5)	0.0206 (0.0003)	0.00286 (0.00009)
	Degree alignment	0.1201 (0.0004)	0.1024 (0.0005)	0.2048 (0.0004)	0.0402 (0.0002)	69.2 (0.6)	0.0178 (0.0002)	0.00294 (0.00010)
	No change	0.0953 (0.0005)	0.0733 (0.0006)	0.1740 (0.0007)	0.0196 (0.0004)	0.0 (0.0)	0.0222 (0.0003)	0.00703 (0.00023)
0.15	**SCN**	0.1187 (0.0003)	0.1025 (0.0006)	0.2006 (0.0003)	0.0255 (0.0002)	64.1 (0.6)	0.0185 (0.0002)	0.00285 (0.00010)
	Nearest GW source	0.1224 (0.0003)	0.1027 (0.0005)	0.2115 (0.0004)	0.0316 (0.0003)	63.1 (0.5)	0.0212 (0.0002)	0.00320 (0.00013)
	Best single source	0.1215 (0.0003)	0.1024 (0.0005)	0.2092 (0.0003)	0.0308 (0.0002)	63.6 (0.5)	0.0200 (0.0002)	0.00238 (0.00009)
	Uniform GW barycenter	0.1190 (0.0003)	0.1030 (0.0006)	0.2002 (0.0003)	0.0252 (0.0003)	63.0 (0.6)	0.0193 (0.0003)	0.00295 (0.00009)
	Degree alignment	0.1195 (0.0003)	0.1034 (0.0006)	0.2014 (0.0003)	0.0249 (0.0002)	65.5 (0.6)	0.0186 (0.0002)	0.00269 (0.00009)
	No change	0.1121 (0.0004)	0.0935 (0.0006)	0.1935 (0.0005)	0.0196 (0.0004)	0.0 (0.0)	0.0207 (0.0003)	0.00675 (0.00022)
0.2	**SCN**	0.1193 (0.0003)	0.1036 (0.0006)	0.2006 (0.0003)	0.0243 (0.0003)	62.1 (0.7)	0.0189 (0.0002)	0.00280 (0.00009)
	Nearest GW source	0.1220 (0.0003)	0.1033 (0.0005)	0.2093 (0.0003)	0.0295 (0.0002)	61.7 (0.5)	0.0208 (0.0002)	0.00313 (0.00013)
	Best single source	0.1218 (0.0003)	0.1036 (0.0005)	0.2084 (0.0003)	0.0291 (0.0002)	62.5 (0.5)	0.0200 (0.0002)	0.00227 (0.00008)
	Uniform GW barycenter	0.1193 (0.0003)	0.1036 (0.0006)	0.2002 (0.0003)	0.0238 (0.0003)	60.9 (0.7)	0.0192 (0.0003)	0.00300 (0.00009)
	Degree alignment	0.1197 (0.0003)	0.1034 (0.0006)	0.2019 (0.0003)	0.0248 (0.0003)	61.7 (0.7)	0.0193 (0.0002)	0.00297 (0.00010)
	No change	0.1166 (0.0004)	0.0982 (0.0006)	0.1998 (0.0004)	0.0196 (0.0004)	0.0 (0.0)	0.0206 (0.0003)	0.00671 (0.00021)

Notes: Entries are means across 500 replications, with Monte Carlo standard errors in parentheses.

Table 17: Design 4: feasible results, $N_g \in [500, 550]$, $h \in \{0.025, 0.05, 0.075\}$

h	Method	RMSE	MedAE	Q95	Δ MedAE	Direction (%)	Degree MedAE	Triangle AE
0.025	**SCN**	0.1185 (0.0003)	0.1040 (0.0006)	0.1975 (0.0003)	0.0208 (0.0004)	64.7 (0.9)	0.0167 (0.0002)	0.00230 (0.00008)
	Nearest GW source	0.1195 (0.0003)	0.1040 (0.0005)	0.2011 (0.0003)	0.0235 (0.0003)	66.1 (0.6)	0.0166 (0.0002)	0.00217 (0.00011)
	Best single source	0.1192 (0.0003)	0.1042 (0.0005)	0.2000 (0.0003)	0.0230 (0.0003)	66.8 (0.6)	0.0154 (0.0002)	0.00144 (0.00006)
	Uniform GW barycenter	0.1190 (0.0003)	0.1039 (0.0006)	0.1988 (0.0002)	0.0216 (0.0004)	58.7 (1.0)	0.0190 (0.0003)	0.00286 (0.00008)
	Degree alignment	0.1185 (0.0003)	0.1038 (0.0006)	0.1978 (0.0003)	0.0209 (0.0004)	63.6 (1.0)	0.0170 (0.0002)	0.00261 (0.00010)
	No change	0.1191 (0.0003)	0.1022 (0.0006)	0.2004 (0.0003)	0.0197 (0.0004)	0.0 (0.0)	0.0190 (0.0004)	0.00664 (0.00021)
0.05	**SCN**	0.0634 (0.0002)	0.0422 (0.0001)	0.1239 (0.0003)	0.0561 (0.0004)	64.4 (0.2)	0.0135 (0.0002)	0.00224 (0.00006)
	Nearest GW source	0.0972 (0.0003)	0.0609 (0.0002)	0.1940 (0.0006)	0.0730 (0.0003)	63.0 (0.2)	0.0172 (0.0001)	0.00166 (0.00007)
	Best single source	0.0946 (0.0003)	0.0592 (0.0002)	0.1891 (0.0006)	0.0721 (0.0003)	61.8 (0.2)	0.0176 (0.0001)	0.00235 (0.00010)
	Uniform GW barycenter	0.0662 (0.0002)	0.0452 (0.0002)	0.1280 (0.0004)	0.0578 (0.0004)	62.2 (0.3)	0.0160 (0.0002)	0.00301 (0.00008)
	Degree alignment	0.1223 (0.0004)	0.1012 (0.0005)	0.2135 (0.0005)	0.1000 (0.0002)	68.0 (0.9)	0.0148 (0.0001)	0.00322 (0.00009)
	No change	0.0905 (0.0005)	0.0575 (0.0004)	0.1793 (0.0011)	0.0197 (0.0004)	0.0 (0.0)	0.0196 (0.0003)	0.00716 (0.00022)
0.075	**SCN**	0.0880 (0.0003)	0.0707 (0.0003)	0.1559 (0.0004)	0.0272 (0.0002)	74.8 (0.5)	0.0124 (0.0001)	0.00163 (0.00006)
	Nearest GW source	0.0964 (0.0004)	0.0737 (0.0004)	0.1761 (0.0008)	0.0378 (0.0004)	69.9 (0.4)	0.0170 (0.0002)	0.00174 (0.00007)
	Best single source	0.0923 (0.0003)	0.0710 (0.0003)	0.1682 (0.0005)	0.0343 (0.0002)	72.4 (0.4)	0.0145 (0.0001)	0.00171 (0.00007)
	Uniform GW barycenter	0.0937 (0.0004)	0.0759 (0.0005)	0.1646 (0.0005)	0.0316 (0.0003)	69.2 (0.5)	0.0179 (0.0003)	0.00274 (0.00008)
	Degree alignment	0.1191 (0.0004)	0.1015 (0.0005)	0.2034 (0.0005)	0.0460 (0.0002)	74.7 (0.8)	0.0148 (0.0001)	0.00309 (0.00009)
	No change	0.0762 (0.0006)	0.0585 (0.0007)	0.1390 (0.0008)	0.0197 (0.0004)	0.0 (0.0)	0.0187 (0.0004)	0.00699 (0.00022)

Notes: Entries are means across 500 replications, with Monte Carlo standard errors in parentheses.

Table 18: Design 4: feasible results, $N_g \in [500, 550]$, $h \in \{0.1, 0.15, 0.2\}$

h	Method	RMSE	MedAE	Q95	Δ MedAE	Direction (%)	Degree MedAE	Triangle AE
0.1	**SCN**	0.1083 (0.0003)	0.0932 (0.0005)	0.1835 (0.0003)	0.0203 (0.0002)	72.9 (0.6)	0.0138 (0.0002)	0.00163 (0.00006)
	Nearest GW source	0.1114 (0.0003)	0.0938 (0.0004)	0.1917 (0.0004)	0.0259 (0.0003)	69.1 (0.5)	0.0160 (0.0003)	0.00181 (0.00008)
	Best single source	0.1097 (0.0003)	0.0929 (0.0004)	0.1882 (0.0003)	0.0238 (0.0002)	71.1 (0.5)	0.0143 (0.0001)	0.00170 (0.00006)
	Uniform GW barycenter	0.1120 (0.0003)	0.0959 (0.0005)	0.1896 (0.0003)	0.0244 (0.0003)	66.9 (0.6)	0.0188 (0.0003)	0.00274 (0.00008)
	Degree alignment	0.1180 (0.0004)	0.1032 (0.0006)	0.1976 (0.0004)	0.0217 (0.0003)	77.7 (0.6)	0.0150 (0.0001)	0.00224 (0.00007)
	No change	0.0998 (0.0006)	0.0835 (0.0007)	0.1714 (0.0006)	0.0197 (0.0004)	0.0 (0.0)	0.0185 (0.0004)	0.00675 (0.00021)
0.15	**SCN**	0.1167 (0.0003)	0.1025 (0.0006)	0.1945 (0.0003)	0.0197 (0.0003)	70.6 (0.8)	0.0151 (0.0002)	0.00188 (0.00007)
	Nearest GW source	0.1180 (0.0004)	0.1023 (0.0005)	0.1993 (0.0004)	0.0229 (0.0003)	68.8 (0.6)	0.0155 (0.0002)	0.00207 (0.00010)
	Best single source	0.1175 (0.0003)	0.1023 (0.0005)	0.1977 (0.0003)	0.0222 (0.0003)	69.4 (0.6)	0.0145 (0.0001)	0.00145 (0.00005)
	Uniform GW barycenter	0.1178 (0.0003)	0.1025 (0.0006)	0.1972 (0.0002)	0.0214 (0.0003)	63.8 (0.8)	0.0187 (0.0003)	0.00281 (0.00008)
	Degree alignment	0.1182 (0.0003)	0.1039 (0.0006)	0.1969 (0.0003)	0.0192 (0.0003)	72.8 (0.8)	0.0159 (0.0002)	0.00215 (0.00007)
	No change	0.1133 (0.0004)	0.0967 (0.0006)	0.1913 (0.0004)	0.0197 (0.0004)	0.0 (0.0)	0.0188 (0.0004)	0.00666 (0.00021)
0.2	**SCN**	0.1182 (0.0003)	0.1038 (0.0006)	0.1971 (0.0003)	0.0200 (0.0003)	68.7 (0.8)	0.0163 (0.0002)	0.00227 (0.00008)
	Nearest GW source	0.1195 (0.0003)	0.1038 (0.0005)	0.2013 (0.0003)	0.0230 (0.0003)	68.4 (0.6)	0.0163 (0.0002)	0.00218 (0.00011)
	Best single source	0.1192 (0.0003)	0.1040 (0.0005)	0.2001 (0.0003)	0.0224 (0.0003)	69.0 (0.5)	0.0152 (0.0002)	0.00145 (0.00006)
	Uniform GW barycenter	0.1190 (0.0003)	0.1037 (0.0006)	0.1989 (0.0002)	0.0214 (0.0003)	62.5 (0.9)	0.0189 (0.0003)	0.00283 (0.00008)
	Degree alignment	0.1184 (0.0003)	0.1039 (0.0006)	0.1973 (0.0003)	0.0199 (0.0003)	68.7 (0.9)	0.0164 (0.0002)	0.00236 (0.00009)
	No change	0.1162 (0.0004)	0.0994 (0.0006)	0.1958 (0.0004)	0.0197 (0.0004)	0.0 (0.0)	0.0189 (0.0004)	0.00665 (0.00021)

Notes: Entries are means across 500 replications, with Monte Carlo standard errors in parentheses.