WHEN MORE OUTCOMES STOP HELPING: MINIMAX INFORMATION FRONTIERS FOR NETWORK EXPERIMENTS

Anonymous authors
Paper under double-blind review

ABSTRACT

Network experiments can assign treatment broadly while measuring only $B \ll n$ outcomes. Under interference, these measurements need not yield mutually compatible causal observations. We formulate this constraint as a budgeted experiment on a labeled exposure-conflict graph. Including the analysis support in the design record preserves the physical experiment's minimax risk. For Direct and Global effects with bounded second moments on bounded-degeneracy networks, we establish the common decomposition $R_{B,c}(G; M) \asymp_k M^2\{B^{-1} + \Psi_c(G)\}$. It separates the B^{-1} cost of outcome measurement from a contrast-specific, network-induced risk floor. We develop Design-Supported Conflict-Graph Design (`DS-CGD`): an offline search builds a finite library of compatible supports, and convex mixture optimization allocates probability using pretreatment residual scores. The resulting exact inclusion probabilities yield a design-unbiased generalized-difference estimator. Across evaluated conditions, geometric-mean Direct-effect RMSE is 12% lower than the best matched baseline in each condition on synthetic graphs, and 16% and 22% lower on the two observed network datasets (CAI and ca-GrQc), respectively. The corresponding Global-effect reductions are 40%, 4%, and 11%.

1 INTRODUCTION

Consider the weather-insurance information experiment of Cai et al. (2015). Information sessions can be offered throughout a friendship network, but post-season household surveys are costly. If the field team can revisit only B farmers, a revisit is informative only when its exposure is realized by the treatment assignment. A Direct effect compares a treated farmer with untreated friends to an untreated farmer with untreated friends; a Global effect requires an entire local neighborhood to be treated or untreated. These conditions may conflict on shared neighbors, so B measured outcomes need not yield B observations relevant to the causal contrast.

This problem joins two strands of causal inference. Exposure mappings define effects under interference (Aronow & Samii, 2017; Leung, 2022), and graph-aware designs improve the realization of informative exposures (Ugander et al., 2013; Ugander & Yin, 2023; Cai et al., 2024; Kandiros et al., 2024). A separate literature studies costly outcome collection, including finite-population sampling and active acquisition (Aronow & Lopatto, 2026; Nwankwo et al., 2026). In a network experiment, follow-up selection and treatment assignment must be chosen jointly because the selected unit–exposure pairs must coexist under one assignment.

Choosing follow-up first may select incompatible exposures, while choosing treatment first may spend measurements on outcomes irrelevant to the target contrast. The experiment must coordinate treatment and outcome measurement before outcomes are observed. This leads to the question studied here:

> *What is the minimax precision with only B outcome measurements under network interference, and when do additional measurements cease to improve the optimal rate?*

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

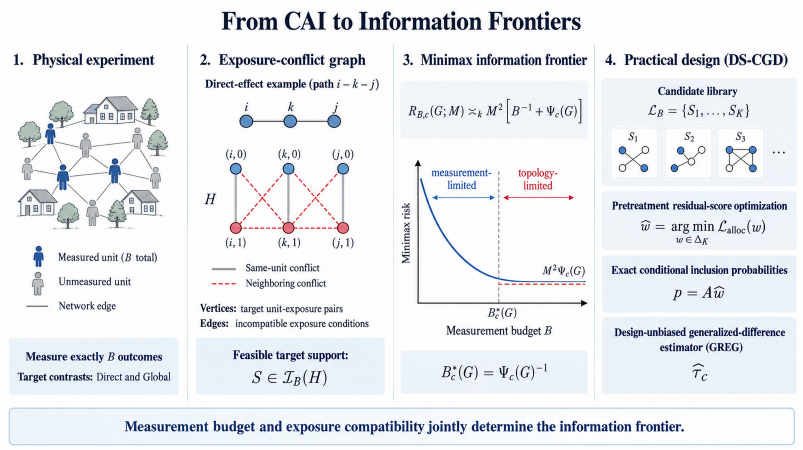


Figure 1: Overview of budgeted outcome measurement, the exposure-conflict graph, the minimax information frontier, and a finite support library for DS−CGD.

We represent each estimand-relevant unit–exposure pair by a vertex of an exposure-conflict graph H, joining two vertices when their conditions impose contradictory treatments. Jointly realizable unit–exposure pairs form independent sets of H. This representation preserves minimax risk: treatment is assigned on the physical network, outcomes are measured on physical units, and H records which potential outcomes can be observed together.

The resulting risk has a simple form. For second-moment-bounded outcomes and a choice of contrast and estimator class c, our fixed-network results give, with constants depending only on the degeneracy bound k,

$$R_{B,c}(G;M) \asymp_k M^2\left\{\frac{1}{B}+\Psi_c(G)\right\}. \tag{1}$$

Thus the corresponding worst-case RMSE is $\sqrt{R_{B,c}(G;M)} \asymp_k M\sqrt{B^{-1}+\Psi_c(G)}$. The term B^{-1} is the cost of measuring only B outcomes. The quantity $\Psi_c(G)$ captures the remaining difficulty of realizing and observing the target exposure conditions. For unrestricted Direct, design-unbiased Direct, and Global effects, it is respectively $n^{-1}\Phi(G,2)$, $n^{-1}U(G,2)$, and $\Theta(G,2)$; see Theorem 1. The scale $B_c^\star(G) := \Psi_c(G)^{-1}$ marks matching full-budget precision up to a constant factor. Below this scale, additional follow-up mainly reduces estimation uncertainty from limited observation. Beyond it, network interference determines the dominant error.

The form of $\Psi_c(G)$ depends on the causal contrast and estimator class. For Direct effects, allowing bias leads to a trimmed degree functional, because a small number of difficult units can be downweighted. Exact design unbiasedness retains the full degree functional and can therefore be substantially more costly. Contrasts that change neighbors' treatment status, including spillover and Global effects, inherit a two-hop exposure constraint and belong to a different information class. Thus the relevant notion of network complexity is determined by the exposure contrast, not only by the physical degree sequence.

For practical design, we restrict the randomization distribution to a finite collection of compatible exposure sets. DS−CGD uses pretreatment predictions to optimize the mixture and its exact inclusion probabilities in a design-unbiased generalized-difference estimator. The library ensures feasibility under the measurement budget; its risk depends on the resulting joint inclusion structure.

Contributions. Our contributions are threefold.

1. **Outcome measurement under interference.** We formulate treatment assignment and costly outcome measurement as one design-based problem. For concrete exposure conditions, the physical network experiment and its exposure-conflict representation have the same minimax value.
2. **Minimax information frontiers on an observed network.** We establish matching upper and lower bounds of the form (1) for Direct and Global effects on sparse networks. The Direct

frontier separates the unrestricted and exactly design-unbiased estimator classes, while spillover and Global contrasts share a two-hop exposure constraint.

3. **Network complexity, saturation, and design.** We characterize growing-degeneracy regimes, including when one measured outcome attains the full-budget minimax risk, and develop `DS-CGD` for compatible exposure sets. We evaluate it on synthetic, CAI friendship-network, and ca-GrQc experiments.

The paper follows the scientific chain from motivation to method and validation. Section 3 specifies the one-hop interference model, exposure contrasts, and conflict graph. Section 4 derives the minimax measurement cost and network-induced frontiers. Section 5 introduces `DS-CGD` as a finite compatible-support design, and Section 6 evaluates it on synthetic and observed networks; the appendices give related work, algorithmic details, additional results, and proofs.

2 Literature Review

Exposure-based frameworks define causal contrasts under interference (Hudgens & Halloran, 2008; Aronow & Samii, 2017; Leung, 2022); graph-aware designs coordinate assignments through clusters, quasi-colorings, focal independent sets, and worst-case-MSE clusterings (Ugander et al., 2013; Ugander & Yin, 2023; Jagadeesan et al., 2020; Cai et al., 2024; Viviano et al., 2023). Conflict-Graph Design makes pairwise exposure incompatibility explicit (Kandiros et al., 2024). Most closely, Kandiros et al. (2026) optimize all designs and estimators under unrestricted outcome observation and show that the conflict graph controls minimax risk. Their analysis also treats general moment restrictions, bias–variance tradeoffs, and priority-based conflict-graph designs. Here the physical protocol measures exactly B outcomes. We derive matching budgeted rates for Direct and Global effects on sparse networks, distinguish unrestricted from exactly design-unbiased Direct risk, and obtain bounds uniform over growing network complexity.

The second neighboring literature allocates scarce observations. Under an expected sample-size budget, independent unequal-probability sampling is minimax for design-unbiased estimation of a bounded finite-population total (Aronow & Lopatto, 2026). Batch-adaptive annotation can reduce the asymptotic variance of a doubly robust estimator under no interference (Nwankwo et al., 2026). Zhang et al. (2025) study active treatment-effect estimation with a conflict-graph extension under network interference, while Yamin et al. (2025) study dependent rounding for budget-constrained treatment assignments. Here graph edges constrain which target exposure conditions can be observed together. The estimand is finite-population, the design is fixed before current outcomes, and validity follows from its randomization law. Pretreatment adjustment is complementary (Lu et al., 2026b). A complete discussion of the work most closely related to our setting appears in Appendix A.

3 Problem Setup

We consider a finite population of n units connected by a known undirected network $G = (V, E)$, where $V = [n]$, and treatment assignment $Z = (Z_1, \dots, Z_n) \in \{0,1\}^n$. For unit i, write N_i for its set of neighbors and $N_i^+ = N_i \cup \{i\}$ for its closed neighborhood. The potential outcome of unit i under assignment z is denoted by $Y_i(z)$.

We first state the one-hop form of arbitrary neighborhood interference (ANI), which fixes the range of interference used throughout the finite-population analysis.

Assumption 1 (Arbitrary Neighborhood Interference (ANI))**.** *For every $i \in [n]$ and assignments $z, z' \in \{0,1\}^n$,*

$$z_{N_i^+} = z'_{N_i^+} \quad \Longrightarrow \quad Y_i(z) = Y_i(z').$$

Assumption 1 leaves the response surface unrestricted over the closed neighborhood N_i^+, while making that neighborhood the boundary of interference. Heterogeneous and nonlinear interactions among the unit and its immediate neighbors are allowed, but treatments outside N_i^+ cannot alter Y_i. The general exposure-mapping perspective follows Aronow & Samii (2017), with the earlier design-based interference framework of Hudgens & Halloran (2008) as background. The local

conflict-graph setting is developed by Kandiros et al. (2024). Approximate neighborhood interference instead allows effects from more distant units to decay without being exactly zero (Leung, 2022). The supplied graph therefore determines the exposure range, whereas all potential-outcome values within that range remain free. This separation is the modeling step on which the rest of the paper rests: ANI fixes which treatment coordinates can affect an outcome, but it does not impose a parametric response surface within that neighborhood. The causal contrast can therefore be defined on local exposure conditions before any design or estimator is chosen.

Target estimand. We next define the finite-population contrast induced by two realizable local exposure conditions. Under Assumption 1, write $Y_i(e)$ for the potential outcome under local exposure e. For each unit, fix two distinct realizable conditions $e_{i,0}, e_{i,1} \in \{0,1\}^{N_i^+}$. The target is

$$\tau = \frac{1}{n}\sum_{i=1}^{n}\{Y_i(e_{i,1}) - Y_i(e_{i,0})\}. \tag{2}$$

Homogeneous exposure contrasts. For the rates below, we specialize to homogeneous neighborhood exposures. Let $e_i^{ab} := \{Z_i = a,\ Z_j = b \text{ for all } j \in N_i\}$, where $a, b \in \{0,1\}$. The Direct effect, spillover effect among treated units, and Global effect use respectively (e_i^{00}, e_i^{10}), (e_i^{10}, e_i^{11}), and (e_i^{00}, e_i^{11}). The conflict-graph reduction applies to any distinct pair of realizable conditions; the degree-based rates below concern the stated homogeneous contrasts. For spillover contrasts, the two conditions are distinct only on nonisolated units. The effective population is $V_+ = \{i : d_i > 0\}$; isolated units have identical spillover exposures and contribute zero, so the corresponding rescaling is given in Appendix D.6.

Exposure compatibility. We now encode which target exposure conditions can be realized together. A single treatment assignment need not realize every exposure condition in (2) simultaneously. We represent this restriction by an exposure-level conflict graph $H = (\mathcal{V}_H, \mathcal{E}_H)$, with $\mathcal{V}_H = [n] \times \{0,1\}$. Vertex (i,a) represents the potential outcome $Y_i(e_{i,a})$. Two vertices (i,a) and (j,b) are adjacent when their local exposure conditions disagree on at least one treatment coordinate shared by their closed neighborhoods:

$$(i,a) \sim_H (j,b) \quad \Longleftrightarrow \quad \exists\, k \in N_i^+ \cap N_j^+ \text{ such that } (e_{i,a})_k \neq (e_{j,b})_k. \tag{3}$$

In particular, $(i,0) \sim_H (i,1)$. We call H a conflict graph because an edge records incompatibility, while an independent set is precisely a collection of unit–exposure pairs that can be realized under one assignment. With measurement budget B, let $\mathcal{I}_B(H)$ be the independent sets of H of size at most B.

The physical experiment measures exactly B distinct units. Its recorded support of estimand-relevant unit–exposure pairs has at most B elements and lies in $\mathcal{I}_B(H)$; measured outcomes outside this support need not enter the estimator. Conversely, every $S \in \mathcal{I}_B(H)$ can be realized and completed to exactly B physical measurements. Section 4 formalizes this equivalence.

Design-based risk. We next place the problem in the design-based framework: the network and the full schedule of potential outcomes are fixed, and probability is induced only by the experimental design. Let $y_i^a := Y_i(e_{i,a})$, $y^a = (y_1^a, \ldots, y_n^a)$, $a \in \{0,1\}$. Then $\tau(y) = n^{-1}\sum_{i=1}^n (y_i^1 - y_i^0)$.

For $q \in [2, \infty]$ and $M > 0$, the finite-population moment class is

$$\mathcal{Y}_q(M) := \left\{ y = (y^0, y^1) : \max_{a\in\{0,1\}} n^{-1/q}\|y^a\|_q \le M \right\}, \tag{4}$$

For $2 \le q < \infty$, $\|u\|_q = (\sum_{i=1}^n |u_i|^q)^{1/q}$, and $\|u\|_\infty = \max_i |u_i|$. Thus q controls the allowed concentration of the finite-population outcomes, while M is their scale.

A design is a pre-outcome distribution $w = (w_S)_{S\in\mathcal{I}_B(H)}$ over feasible sets; after $S \sim w$, it observes y_S. An estimator is a collection of measurable functions $T_S = T_S(y_S)$.

For a fixed schedule y, its design-based risk is $r_{w,T}(y) := \mathbb{E}_w[\{T_S(y_S) - \tau(y)\}^2]$ and its bias is $b_{w,T}(y) := \mathbb{E}_w[T_S(y_S)] - \tau(y)$. The unrestricted minimax risk is $R(B, H; q, M) := \inf_{w,T} \sup_{y\in\mathcal{Y}_q(M)} r_{w,T}(y)$; for $\delta \ge 0$, the bias-constrained counterpart is $R(B, H, \delta; q, M) :=$

$\inf_{w,T:\sup_{y\in\mathcal{Y}_q(M)}|b_{w,T}(y)|\le\delta}\sup_{y\in\mathcal{Y}_q(M)} r_{w,T}(y)$. Here $\delta = 0$ gives the exactly design-unbiased class; $R(B, H; q, M)$ places no bias restriction. The full-budget benchmark is $R(n, H; q, M)$.

The setup has now separated the causal contrast from the set of exposure conditions that can be observed together. Section 4 uses this reduction to quantify the measurement cost first and then the additional cost created by network incompatibility.

4 Minimax Risk and Outcome Measurement

This section answers four design questions. What cost is imposed by measuring only B outcomes? What network-induced error remains when all outcomes are measured? At what budget does the risk reach the full-measurement order? How do these quantities change as graph degeneracy grows with n? We address the questions in this order and defer proofs to Appendix D. The first result establishes that the physical experiment and its exposure graph have the same minimax value; the next results then separate the universal measurement term from the network term and identify the budget at which the full-measurement order becomes attainable.

4.1 Minimax equivalence

Fix the network and exposure contrast in Section 3. Before bounding risk, we remove a possible ambiguity in the representation: the physical protocol measures units on G, whereas the minimax problem is written on H. Let $R_{\mathrm{phys}}(B, \delta; q, M)$ be the minimax risk of the physical experiment, whose design record contains the unit–exposure pairs used for estimation. Potential outcomes at all other exposures remain unrestricted.

We first show that the physical experiment and its exposure-conflict graph have the same minimax value, so the graph representation loses no information (Proposition 1).

Proposition 1 (Minimax equivalence)**.** *For every $1 \le B \le n$, $q \in [2, \infty]$, $M > 0$, and $\delta \ge 0$,*

$$\begin{aligned} R_{\mathrm{phys}}(B, \delta; q, M) &= R(B, H, \delta; q, M), \\ R(B, H, \delta; q, M) &= M^2 R(B, H, \delta/M; q, 1). \end{aligned} \tag{5}$$

Both equalities also hold without a bias constraint.

Appendix D.1 gives the formal reduction and scaling argument. This proposition is the bridge for the rest of the theory: it lets us separate how many outcomes are measured from which target exposures can be realized jointly.

We next isolate the universal cost of measuring only B outcomes; it is independent of the conflict pattern:

$$R(B, H, \delta; q, M) \ge \max\left\{ R(n, H, \delta; q, M), \frac{M^2}{B} \right\}. \tag{6}$$

The first term is the full-measurement benchmark, and the second is the unavoidable M^2/B cost of observing only B outcomes. It persists even when all exposure conflicts are removed, so measurement scarcity and interference impose distinct restrictions on attainable precision.

4.2 Minimax risk on an observed network

We next identify the network-induced term that remains after measurement cost is isolated. For the direct effect with untreated neighbors, an unrestricted procedure may omit or downweight difficult high-degree targets. For a set $D \subseteq V$ assigned to control, the resulting omission bias is bounded by

$$\sup_{y\in\mathcal{Y}_q(M)} \left| \frac{1}{n}\sum_{i\in D}(y_i^1 - y_i^0) \right| \le 2M \left(\frac{|D|}{n} \right)^s, \qquad s = 1 - \frac{1}{q}. \tag{7}$$

Removing a few highly connected units can make the remaining target events easier to observe, but this variance reduction must be balanced against the omission bias.

Write $d_i = |N_i|$, $v_i = d_i + 1$, and $t = q/(q-2)$, with $t = \infty$ at $q = 2$ and $t = 1$ at $q = \infty$. The following network difficulty functionals distinguish the Direct-effect tradeoff from exact

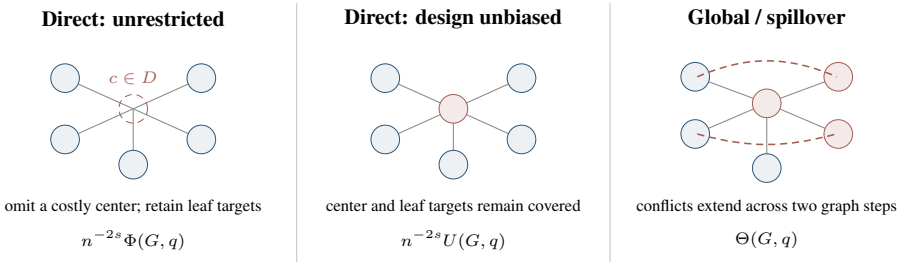


Figure 2: Schematic of the unrestricted Direct, design-unbiased Direct, and Global/spillover network difficulty functionals.

design unbiasedness and from Global or spillover contrasts. Exact unbiasedness cannot trim target contributions, whereas opposite-arm exposures in Global-type contrasts interact through two-hop neighborhoods.

$$\begin{aligned} U(G,q) &:= \left\|(v_i^s)_{i\in V}\right\|_t, \\ \Phi(G,q) &:= \min_{D\subseteq V} \left\{|D|^{2s} + \left\|(v_i^s)_{i\notin D}\right\|_t\right\}, \\ \Theta(G,q) &:= \min\left\{1,\, n^{-2s}\left\|(v_i^{2s})_{i\in V}\right\|_t\right\}. \end{aligned} \tag{8}$$

The empty-vector norm is zero. For a fixed size of D, the minimum in Φ removes the largest degrees. Thus Φ is the trimmed Direct functional, U its untrimmed counterpart, and Θ the two-hop functional for Global-type contrasts.

Figure 2 summarizes the three exposure-compatibility patterns that generate these functionals.

To obtain rates that are uniform over a graph class, we next impose a degeneracy bound. This condition controls the sparsity of every induced subgraph while still allowing individual degrees to grow with n: a graph has degeneracy at most k if every nonempty induced subgraph contains a vertex of degree at most k. We write $\asymp_k$ for comparisons with positive constants depending only on k.

The network difficulty functionals now yield the sharp Direct and Global frontiers on a fixed sparse network (Theorem 1).

Theorem 1 (Direct and global effects)**.** *For every network G of degeneracy at most $k \geq 1$, every $1 \leq B \leq n$, $q \in [2,\infty]$, and $M > 0$,*

$$R(B, H_{\mathrm{D}}(G); q, M) \asymp_k M^2\left\{B^{-1} + n^{-2s}\Phi(G,q)\right\}, \tag{9}$$

$$R(B, H_{\mathrm{D}}(G), 0; q, M) \asymp_k M^2\left\{B^{-1} + n^{-2s}U(G,q)\right\}, \tag{10}$$

$$R(B, H_{\mathrm{G}}(G), \delta; q, M) \asymp_k M^2\left\{B^{-1} + \Theta(G,q)\right\}, \qquad \delta \geq 0. \tag{11}$$

The global comparison also holds without a bias constraint.

The first term in each display is the topology-independent B^{-1} measurement cost; the second is the network term that remains at full measurement. For Direct effects, allowing bias replaces U by the trimmed functional Φ; for Global effects, the relevant functional is Θ. At $B = n$, the ratio of the design-unbiased Direct risk to the unrestricted risk has order $U(G,q)/\Phi(G,q)$, whose maximum over n-vertex graphs with a fixed degeneracy bound has order $n^{1/q}$. For Global effects, a design-unbiased estimator already attains the unrestricted rate.

Measurement scale. We next locate the budget at which additional measurements cease to change the minimax order. This decomposition gives a natural answer to when additional measurements stop changing risk by more than a constant factor. Set $\Theta_{\mathrm{D}}(G,q) = n^{-2s}\Phi(G,q)$ and $\Theta_{\mathrm{G}}(G,q) = \Theta(G,q)$. For $c \in \{\mathrm{D}, \mathrm{G}\}$,

$$B_\star(G,c,q) := \Theta_c(G,q)^{-1}, \qquad \frac{R(B, H_c(G); q, M)}{R(n, H_c(G); q, M)} \asymp_k 1 + \frac{B_\star(G,c,q)}{B}. \tag{12}$$

A budget of order $B_\star$ is therefore sufficient and necessary to match the risk of measuring all n outcomes within a constant factor. Since $R(n, H_c(G); q, M) \asymp_k M^2/B_\star$, an earlier transition can

also reflect a larger risk that persists at full measurement. The scale $B_\star$ characterizes a transition in minimax order, with comparison constants depending on k.

Other exposure contrasts. The preceding distinction extends to the four neighborhood assignments $00, 01, 10, 11$. On networks without isolated units, the two direct contrasts 00 versus 10 and 01 versus 11 have the rates in (9)–(10). The remaining four contrasts change neighbors' treatment, including both spillover effects. Each has the global rate in (11), at every bias allowance and without a bias constraint. The relevant difference is whether the contrast keeps neighbors' treatment fixed or changes it.

Exact saturation. A final question concerns whether one additional measurement can improve the risk at all. The scale $B_\star$ concerns constant-factor precision. A separate question is whether an additional measurement can improve the minimax value at all. There is an exact statement about the first additional measurement. For any conflict graph in the setup with $n \geq 2$, suppressing the fixed arguments of R,

$$R(1, \delta) = R(2, \delta) \quad \Longleftrightarrow \quad R(1, \delta) = R(n, \delta). \tag{13}$$

The same equivalence holds without a bias constraint. Thus, whenever a larger budget improves on one measurement, two measurements already give a strict improvement. Without a bias constraint, saturation occurs exactly when no connected component of the compatibility graph contains more than n target coordinates; the compatibility graph joins two coordinates when they can be observed together.

4.3 Network complexity and uniform consistency

Finally, we extend the fixed-network frontiers to graph classes whose degeneracy may grow with population size (Theorem 2). The preceding results give graph-specific rates for a fixed degeneracy bound. We next seek bounds uniform as that bound grows with n, while allowing a prescribed amount of design bias. Let the degeneracy bound grow with population size and define

$$\begin{aligned} \mathcal{G}(n, k) &:= \{G : |V(G)| = n,\ \mathrm{degen}(G) \leq k\}, \\ W(B, n, k, \delta; q, M) &:= \sup_{G \in \mathcal{G}(n,k)} R(B, H_{\mathrm{D}}(G), \delta; q, M). \end{aligned} \tag{14}$$

Each design and estimator may depend on the observed graph. The supremum then describes the most difficult network in the class.

Theorem 2 (Bias, measurement, and network complexity)**.** *Let $n \geq 2$, $1 \leq k \leq n-1$, $1 \leq B \leq n$, $q \in [2, \infty]$, $M > 0$, and $\delta \geq 0$. With $\rho = k/n$, $\eta = \delta/M$, and $s = 1 - 1/q$,*

$$W(B, n, k, \delta; q, M) \asymp M^2 \left\{ \frac{1}{B} + \rho^{2s/(1+s)} + \frac{\rho}{\left(\rho + \eta^{1/s}\right)^{1-s}} \right\}, \tag{15}$$

where the comparison constants are universal.

The three terms respectively describe limited measurement, the network restriction under unrestricted estimation, and the additional cost of controlling bias. In particular, design unbiasedness gives the rate $M^2\{B^{-1} + \rho^s\}$, while unrestricted estimation gives $M^2\{B^{-1} + \rho^{2s/(1+s)}\}$. Under second moments, the corresponding network terms are $\rho^{1/2}$ and $\rho^{2/3}$.

The same characterization yields the condition for uniform consistency. For fixed $M > 0$ and any admissible sequences $q_n \geq 2$, $\delta_n \geq 0$, B_n, and k_n,

$$W(B_n, n, k_n, \delta_n; q_n, M) \longrightarrow 0 \quad \Longleftrightarrow \quad B_n \longrightarrow \infty \text{ and } k_n = o(n). \tag{16}$$

Thus uniform consistency of direct-effect estimation requires an increasing measurement budget and permits network degeneracy to grow at any sublinear rate.

The theory now has a complete progression. Proposition 1 puts the physical experiment on the exposure graph; Theorem 1 separates measurement cost from contrast-specific incompatibility; and Theorem 2 extends the frontier to growing network complexity. Section 5 uses this decomposition to decide how the available budget should be allocated within the compatible exposure sets.

5 BUDGET-AWARE SAMPLING AND ESTIMATION

Theorem 1 separates the universal B^{-1} cost from the network term. Once the budget scale has been identified in Section 4, the remaining design question is how to place those measurements among exposure conditions that can be realized together. This leads naturally to DS−CGD: it first allocates the measurement budget and then accounts for exposure topology through a finite compatible-support library, using frozen pretreatment scores; see Appendix B.

Let $\mathcal{L}_B = \{S_1, \dots, S_K\} \subseteq \mathcal{I}_B(H)$, and define the incidence matrix and mixture marginals by $A_{vk} = \mathbf{1}\{v \in S_k\}$, $w \in \Delta_K$, and $p(w) = Aw$. Thus $p_v(w)$ is the exact inclusion probability of target event v under the finite mixture. Given a pretreatment estimate $\widehat{a}_v$ of its squared residual magnitude, DS−CGD chooses

$$\widehat{w} \in \arg\min_{w \in \Delta_K} \left\{ \frac{1}{n^2} \sum_{v \in \mathcal{V}_H} \frac{\widehat{a}_v + \delta_a}{(Aw)_v} + \lambda_{\mathrm{reg}} \|w - w^{\mathrm{ref}}\|_2^2 \right\}, \qquad (Aw)_v \geq \underline{p}_v. \tag{17}$$

The leading term in (17) is motivated by the diagonal contribution to the generalized-difference variance, whose residual component scales as r_v^2/p_v. Because the current residuals are unavailable when the design is formed, r_v^2 is replaced by the frozen pretreatment score $\widehat{a}_v$, which yields the stated marginal allocation criterion. The library and mixture induce joint inclusion probabilities, but the objective in (17) optimizes only a marginal variance surrogate. The coverage constraints ensure positive inclusion probabilities; they do not by themselves bound the off-diagonal covariance contribution. For a fixed library, this is a convex program. The first term allocates measurement probability according to predicted residual variation, the lower bounds maintain coverage of target events, and the second term stabilizes the mixture around a reference design. Here $\delta_a > 0$ is a positive residual-score offset and $\lambda_{\mathrm{reg}} \geq 0$ is a regularization coefficient; neither is the bias allowance δ in Section 3.

Finally, let $\widehat{m}$ denote the frozen pretreatment prediction and $r_{i,a} = \sigma_a\{Y_i(e_{i,a}) - \widehat{m}_i^a\}$ its signed residual, where $\sigma_0 = -1$ and $\sigma_1 = 1$. With $J \sim \mathrm{Categorical}(\widehat{w})$, $I_v = \mathbf{1}\{v \in S_J\}$, and $p_v = (A\widehat{w})_v$, we use the generalized-difference estimator

$$\widehat{\tau}_{\mathrm{DS}} = \tau(\widehat{m}) + \frac{1}{n} \sum_{v \in \mathcal{V}_H} \frac{I_v r_v}{p_v}. \tag{18}$$

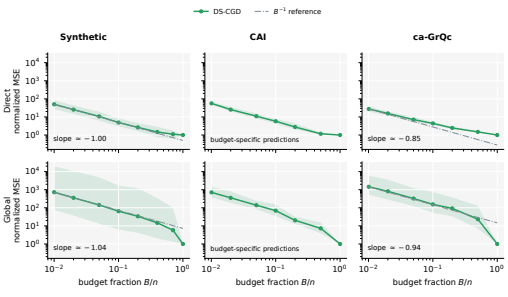
Conditional on the frozen library, mixture, and predictions, (18) is design-unbiased whenever $p_v > 0$. Its allocation criterion uses predicted marginal residual variation; the exact design variance also depends on joint inclusion probabilities. Thus the minimax bounds provide a benchmark, not a risk guarantee for this finite support algorithm.

6 EXPERIMENTS

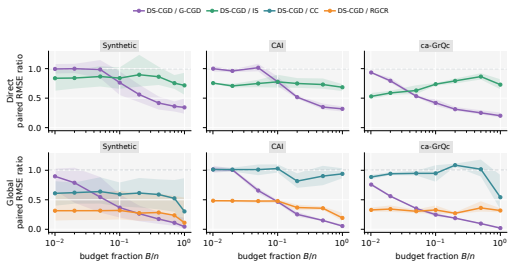
Setup. We evaluate DS−CGD on ER, preferential-attachment (PA), stochastic-block (SBM), and small-world (SW) graphs with $n \in \{2{,}000, 5{,}000\}$, and on CAI and ca-GrQc topologies. We study Direct and Global contrasts under arm-asymmetric, covariate-predictable, and degree-aligned outcome mechanisms. All potential outcomes are simulated. Every estimator uses GREG adjustment from independently generated historical outcomes. Methods share the graph, prediction, current population, and budget within each matched comparison; synthetic and ca-GrQc RGCR comparisons use a separate matched prediction and current population; CAI RGCR shares the common pair (Appendix C). Direct baselines are Generalized and Original CGD (Kandiros et al., 2024) and independent-set design (Cai et al., 2024); Global comparisons also include spectral causal clustering (CC) (Viviano et al., 2023) and randomized graph-cluster randomization (RGCR) (Ugander & Yin, 2023). Appendix C specifies the data, outcome mechanisms, historical information, and comparison protocol. We report RMSE and geometric means of paired DS−CGD/baseline RMSE ratios; values below one favor DS−CGD.

Finite-budget scaling. Figure 3a retains normalized MSE to compare with the B^{-1} term in Theorem 1. The fitted Direct slopes are -1.00 on synthetic graphs and -0.85 on ca-GrQc; Global slopes are -1.04 and -0.94, respectively. The CAI panels show performance across budgets with budget-specific frozen predictions, so their changes also reflect those prediction fits. The curves describe design MSE for the simulated finite populations, rather than worst-case minimax risk.

The two summaries separate budget effects from method comparisons. The left panel normalizes each condition by its full-budget error, so its slope can be read directly against the B^{-1} benchmark. The right panel keeps each method comparison within the same graph, outcome mechanism, prediction, and finite population; geometric aggregation then summarizes relative performance without allowing a single scale to dominate. This separation connects the rate predicted by Theorem 1 to the finite-sample comparisons reported below.



(a) Finite-budget scaling. Normalized `DS-CGD` MSE versus B/n with a B^{-1} reference.



(b) Paired RMSE ratios. Lines are geometric means and bands are the 10th–90th percentiles.

Figure 3: **Experimental scaling and relative performance.** The two panels summarize budget dependence and paired comparisons; the matching protocol is given in Appendix C.

Performance across budgets. Figure 3b shows paired gains at moderate and large budgets, with small-budget exceptions; CC remains competitive on observed networks and RGCR matching is detailed in Appendix C.

Table 1 gives the absolute scale: `DS-CGD` is best in five of six Direct and all six Global PA cells. At 20%, it leads 17/24 synthetic Direct cells, all six observed Direct cells, all 24 common synthetic Global cells, and all three CAI cells; CC leads on ca-GrQc. Appendix C.4 reports the remaining conditions and budgets.

Table 1: **Absolute RMSE on PA graphs.** Degree-aligned outcomes in the common matched group; entries are square roots of mean MSE over two graph realizations.

(a) Direct effect

n	B	IS	G-CGD	`DS-CGD`
2000	20	.7019	**.4432**	.4666
2000	400	.1677	.2323	**.1261**
2000	2000	.1097	.2370	**.0584**
5000	50	.4654	.2824	**.2759**
5000	1000	.1053	.1555	**.0775**
5000	5000	.0673	.1552	**.0483**

(b) Global effect

n	B	G-CGD	CC	`DS-CGD`
2000	20	.5906	1.8322	**.4250**
2000	400	.6275	.4244	**.1108**
2000	2000	.6256	.1198	**.0248**
5000	50	.4488	.5731	**.2934**
5000	1000	.4522	.1358	**.0542**
5000	5000	.4508	.0433	**.0226**

Note. Boldface denotes the row-minimum RMSE. IS denotes independent-set design; G-CGD denotes Generalized CGD; CC denotes spectral causal clustering. RGCR is not part of the common matched comparison group.

7 Discussion and Conclusion

The minimax analysis separates the universal B^{-1} measurement cost from contrast-specific compatibility: Direct effects admit a bias–conflict tradeoff, whereas Global and spillover effects require two-hop compatibility, and $B_\star$ marks the full-budget transition. Guided by this decomposition, `DS-CGD` optimizes a compatible-support mixture from pretreatment scores and uses a design-unbiased generalized-difference estimator. Its low-budget MSE is approximately inverse in B on the evaluated networks, while relative performance depends on contrast and graph. The results assume a known graph, fixed local exposures, and pre-outcome design, so precision depends jointly on the measurement budget and exposure compatibility.

AI USE STATEMENT

We used generative AI tools for literature-retrieval support, language editing, code and figure development, and proof formalization. The authors verified the cited sources, mathematical statements, and generated artifacts and take responsibility for the final content.

ETHICS STATEMENT

The experiments compare randomized designs using synthetic potential outcomes. The CAI benchmark retains an existing social network and baseline covariates; its results concern design performance and do not re-estimate the original insurance intervention's effects. Social-network and household data warrant careful privacy protection. Applications of the proposed designs should consider consent, unequal exposure to interventions, and the consequences of selecting which outcomes to measure.

REPRODUCIBILITY STATEMENT

The exposure model and design assumptions are specified in Section 3. Algorithms 1 and 2, together with Equations (17) and (18), describe the sampling and estimation procedures. Appendix D contains the proofs of the theoretical results.

REFERENCES

Peter M. Aronow and Patrick Lopatto. Minimax unbiased estimation for finite populations with bounded outcomes. *arXiv preprint arXiv:2605.20572*, 2026.

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

Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286 (5439):509–512, 1999.

Chencheng Cai, Xu Zhang, and Edoardo M. Airoldi. Independent-set design of experiments for estimating treatment and spillover effects under network interference. In *The Twelfth International Conference on Learning Representations*, volume 2024, pp. 37362–37377, 2024.

Jing Cai, Alain de Janvry, and Elisabeth Sadoulet. Social networks and the decision to insure. *American Economic Journal: Applied Economics*, 7(2):81–108, 2015.

Jing Cai, Alain de Janvry, and Elisabeth Sadoulet. Replication data for: Social networks and the decision to insure. American Economic Association; ICPSR repository, 2019.

Claes M. Cassel, Carl-Erik Särndal, and Jan H. Wretman. Some results on generalized difference estimation and generalized regression estimation for finite populations. *Biometrika*, 63(3):615–620, 1976.

Paul Erdős and Alfréd Rényi. On random graphs I. *Publicationes Mathematicae Debrecen*, 6: 290–297, 1959.

Paul W. Holland, Kathryn Blackmond Laskey, and Samuel Leinhardt. Stochastic blockmodels: First steps. *Social Networks*, 5(2):109–137, 1983.

Michael G. Hudgens and M. Elizabeth Halloran. Toward causal inference with interference. *Journal of the American Statistical Association*, 103(482):832–842, 2008.

Ravi Jagadeesan, Natesh S. Pillai, and Alexander Volfovsky. Designs for estimating the treatment effect in networks with interference. *The Annals of Statistics*, 48(2):679–712, 2020.

Vardis Kandiros, Charilaos Pipis, Constantinos Daskalakis, and Christopher Harshaw. The conflict graph design: Estimating causal effects under arbitrary neighborhood interference. *arXiv preprint arXiv:2411.10908*, 2024.

Vardis Kandiros, Christopher Harshaw, and Fredrik Sävje. A design-based minimax theory for network experiments. *arXiv preprint arXiv:2608.04909*, 2026.

Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. Graph evolution: Densification and shrinking diameters. *ACM Transactions on Knowledge Discovery from Data*, 1(1):2, 2007. doi: 10.1145/1217299.1217301.

Michael P. Leung. Causal inference under approximate neighborhood interference. *Econometrica*, 90(1):267–293, 2022.

Xin Lu, Hongzi Li, and Hanzhong Liu. Causal inference under uniformly bounded neighbourhood interference. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, art. qkag077, 2026a. doi: 10.1093/jrsssb/qkag077.

Xin Lu, Yuhao Wang, and Zhiheng Zhang. Adjusting auxiliary variables under approximate neighborhood interference. *Journal of the American Statistical Association*, 2026b. doi: 10.1080/01621459.2026.2696737.

Ezinne Nwankwo, Lauri Goldkind, and Angela Zhou. Batch-adaptive causal annotations. In *Proceedings of The 29th International Conference on Artificial Intelligence and Statistics*, volume 300 of *Proceedings of Machine Learning Research*, pp. 4573–4581. PMLR, 2026.

Carl-Erik Särndal, Bengt Swensson, and Jan H. Wretman. *Model Assisted Survey Sampling*. Springer-Verlag, New York, 1992.

Stanford Network Analysis Project. General relativity and quantum cosmology collaboration network, n.d. Dataset description, accessed 24 September 2026.

Johan Ugander and Hao Yin. Randomized graph cluster randomization. *Journal of Causal Inference*, 11(1):20220014, 2023.

Johan Ugander, Brian Karrer, Lars Backstrom, and Jon M. Kleinberg. Graph cluster randomization: Network exposure to multiple universes. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 329–337, 2013.

Davide Viviano, Lihua Lei, Guido Imbens, Brian Karrer, Okke Schrijvers, and Liang Shi. Causal clustering: Design of cluster experiments under network interference. *arXiv preprint arXiv:2310.14983*, 2023.

Duncan J. Watts and Steven H. Strogatz. Collective dynamics of small-world networks. *Nature*, 393 (6684):440–442, 1998.

Bar Weinstein and Daniel Nevo. Causal inference with misspecified network interference structure. *Biometrics*, 82(1):ujag023, 2026. doi: 10.1093/biomtc/ujag023.

Khurram Yamin, Edward Kennedy, and Bryan Wilder. Dependent randomized rounding for budget constrained experimental design. In *Proceedings of the Forty-first Conference on Uncertainty in Artificial Intelligence*, volume 286 of *Proceedings of Machine Learning Research*, pp. 4681–4700. PMLR, 2025.

Zhiheng Zhang, Haoxiang Wang, Haoxuan Li, and Zhouchen Lin. Active treatment effect estimation via limited samples. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 76225–76245. PMLR, 2025.

Appendix Contents

CONTENTS

A Related Work 13

Notation 14

B `DS-CGD`: Algorithmic Principles 14

C Experimental Details 16

C.1 Data and Pretreatment Information 16

C.2 Potential Outcomes 17

C.3 Historical Information and Evaluation Protocol 17

C.4 Additional Experimental Results 19

D Proofs of the Main Results 21

D.1 Proof of Proposition 1 21

D.2 The Measurement Lower Bound 22

D.3 Compatible Sampling and Budget Reduction 23

D.4 Proof of Theorem 1 24

D.5 Measurement Scale and the Cost of Unbiasedness 27

D.6 Other Homogeneous Exposure Contrasts 27

D.7 Exact Saturation at the First Additional Measurement 28

D.8 Bias, Measurement, and Network Complexity 32

D.9 Endpoint Rates and Uniform Consistency 34

A Related Work

Relevant work follows four connected lines: design-based minimax theory for network experiments; network-aware experimental design; finite-population sampling under limited outcome observation; and model-assisted estimation.

Design-based minimax theory for network experiments. Exposure mappings define causal contrasts under interference (Hudgens & Halloran, 2008; Aronow & Samii, 2017). Approximate-neighborhood interference permits effects outside a selected neighborhood to decay (Leung, 2022), while uniformly bounded-neighborhood results study causal contrasts under a bounded exposure range (Lu et al., 2026a); misspecified-network analyses examine the consequences of supplying the wrong interference structure (Weinstein & Nevo, 2026). The closest theoretical comparison is Kandiros et al. (2026), who optimize design and estimator jointly on an exposure-level conflict graph, allow general moment classes, and characterize bias–variance tradeoffs when all outcomes are available. We inherit that exposure-level representation, but impose a finite outcome-measurement budget before current outcomes are observed. The resulting question is not only whether the graph permits a low-risk design, but whether a compatible target support can be measured within the available budget. Our lower bounds separate the full-measurement benchmark from the measurement term, while the upper bounds require a construction that attains both terms simultaneously. On bounded-degeneracy graphs, the quantities Φ, U, and Θ distinguish Direct, design-unbiased Direct, and Global contrasts and give matching rates over the budget range; the threshold $B_\star$ is a constant-factor transition to the full-measurement order, not the point at which risk becomes exactly constant. Theorem 1 gives the fixed-graph frontiers, and Theorem 2 extends them uniformly to growing degeneracy.

Network-aware experimental design. Network-aware designs control which exposure conditions can be realized by choosing clusters, independent sets, or other assignment structures. Graph cluster randomization creates repeated exposure opportunities (Ugander et al., 2013; Ugander & Yin, 2023); independent-set designs select focal units and assign the remaining treatments to control their exposures (Cai et al., 2024); and causal clustering optimizes a worst-case MSE criterion over a specified design class (Viviano et al., 2023). The Conflict-Graph Design framework makes pairwise exposure incompatibility explicit (Kandiros et al., 2024). Zhang et al. (2025) provide a closer procedural comparison: their conflict-graph active-sampling extension combines active selection with regression adjustment under network interference, but its network extension does not establish design unbiasedness. Our `DS-CGD` starts from a prespecified library $\mathcal{L}_B \subseteq \mathcal{I}_B(H)$ of compatible supports, freezes the pretreatment prediction scores, and optimizes a mixture w whose inclusion probabilities are $p = Aw$. Whenever every target event has positive inclusion probability, the generalized-difference estimator is conditionally design unbiased. This is a finite-library construction and a sampling guarantee; it should not be conflated with the all-design minimax frontiers above.

Outcome measurement and finite-population sampling. Aronow & Lopatto (2026) provide the closest finite-population statistical benchmark: under an expected sample-size constraint, they characterize minimax design-unbiased estimation of bounded totals with unequal-probability sampling. Our constraint is a hard budget on measured outcomes, and the target events must also be jointly realizable under the exposure-conflict graph. Nwankwo et al. (2026) allocate expensive annotations in batches to reduce the asymptotic variance of a doubly robust estimator, whereas Yamin et al. (2025) constrain treatment slots and preserve marginal treatment probabilities through dependent rounding. These works clarify different meanings of a budget—labels, treatment assignments, or expected sample size—and therefore provide useful statistical comparisons without supplying the minimax characterization under a fixed network and a binding outcome budget.

Model-assisted estimation. The generalized-difference and generalized-regression construction underlying our estimator is part of the model-assisted sampling tradition (Cassel et al., 1976; Särndal et al., 1992). In our notation, the frozen prediction $\widehat{m}$ improves allocation and can reduce residual variation, while the inclusion probabilities and residual correction determine the conditional design-unbiasedness of $\widehat{\tau}_{\mathrm{DS}}$. Related auxiliary adjustment under approximate interference is studied by Lu et al. (2026b); here the prediction is frozen before the current randomization and enters a design-based residual correction. Prediction quality therefore affects finite-sample precision, whereas the minimax results provide the statistical benchmark for all admissible designs and estimators. The optimization in `DS-CGD` is over a finite candidate library and uses a residual criterion; it is not itself a claim of attaining the unrestricted minimax rate.

NOTATION

Table A.1 collects the notation used repeatedly in the main text and appendices. Formal definitions remain in the section where each object first enters the argument.

Table A.1: Core notation.

Object	Meaning
$G = (V, E)$, N_i, N_i^+	Network, neighbors of unit i, and closed neighborhood.
D, V_+, m	Omitted Direct-effect targets, nonisolated units, and effective-population size.
Z, z, $Y_i(z)$	Treatment assignment, a fixed assignment, and unit i's potential outcome.
$e_{i,a}$, y_i^a, τ	Target exposure, corresponding arm value, and finite-population estimand.
H, H_{D}, H_{G}, $\mathcal{I}_B(H)$	Exposure-conflict graph, Direct and Global reductions, and independent sets of size at most B.
B, w, A, $p = Aw$	Outcome budget, support-mixture weights, incidence matrix, and target-event inclusion probabilities.
q, $\|\cdot\|_q$, M	Moment parameter, q-norm, and outcome scale.
$\mathcal{Y}_q(M)$, δ, R, R_{phys}	Moment class, bias allowance, conflict-graph minimax risk, and physical-protocol minimax risk.
$r_{w,T}$, $b_{w,T}$, T_S	Design-based risk, bias, and estimator used on support S.
d_i, v_i, s, t	Degree, $d_i + 1$, $1 - 1/q$, and $q/(q-2)$ with endpoint conventions.
U, Φ, Θ	Untrimmed, trimmed, and Global/spillover network difficulty functionals.
$\mathcal{G}(n, k)$, W, ρ, η	Degeneracy-bounded graph class, worst-case risk, normalized degeneracy, and normalized bias.
$\mathcal{L}_B$, Δ_K, $\widehat{w}$, $\widehat{\tau}_{\mathrm{DS}}$	Compatible-support library, simplex, optimized mixture, and generalized-difference estimator used by `DS-CGD`.
$\widehat{m}$, r_v, I_v, δ_a, λ_{reg}	Pretreatment prediction, signed residual, inclusion indicator, score offset, and regularization coefficient.

B DS-CGD: ALGORITHMIC PRINCIPLES

This appendix gives the mathematical construction behind `DS-CGD`. We first define the compatible-support library and its inclusion probabilities, then derive the mixture objective, the generalized-difference estimator, and its design-unbiasedness property.

Let $\mathcal{L}_B = \{S_1, \ldots, S_K\} \subseteq \mathcal{I}_B(H)$ and $\Delta_K = \{w \in \mathbb{R}_+^K : \sum_{k=1}^K w_k = 1\}$. Define $A_{vk} = \mathbf{1}\{v \in S_k\}$, $w \in \Delta_K$, $p(w) = Aw$, and $p_v(w) = \sum_{k: v \in S_k} w_k$. The mixture selected by `DS-CGD` is the solution of the convex program already stated in (17); its constraints enforce $p_v(w) \geq \underline{p}_v$ for every target event. The reference weights satisfy $w_k^{\mathrm{ref}} = (2n)^{-1}|\{v : S(v) = S_k\}|$ and $\underline{p} = \epsilon A w^{\mathrm{ref}}$. Choose $0 < \epsilon \leq 1$. Because every target event belongs to a nonempty coverage block, $(Aw^{\mathrm{ref}})_v > 0$

for every v, and hence $0 < \underline{p} \leq Aw^{\mathrm{ref}}$. Thus the reference mixture is feasible for the coverage constraints and assigns positive inclusion probability to every target event.

For a frozen prediction $\widehat{m}$, define the signed residual $r_v = \sigma_a\{Y_i(e_{i,a}) - \widehat{m}_i^a\}$ for $v = (i, a)$, with $\sigma_0 = -1$ and $\sigma_1 = 1$. If $J \sim \mathrm{Categorical}(\widehat{w})$, $S = S_J$, and $I_v = \mathbf{1}\{v \in S\}$, then

$$\widehat{\tau}_{\mathrm{DS}} = \tau(\widehat{m}) + \frac{1}{n} \sum_{v \in \mathcal{V}_H} \frac{I_v r_v}{p_v(\widehat{w})}. \tag{B.1}$$

Conditional on the frozen library, predictions, and weights,

$$\mathbb{E}[\widehat{\tau}_{\mathrm{DS}} \mid \widehat{m}, \mathcal{L}_B, \widehat{w}] = \tau(y) \quad \text{whenever } p_v(\widehat{w}) > 0 \text{ for all } v. \tag{B.2}$$

Algorithm 1 constructs the library, reference weights, and inclusion probabilities described above; Algorithm 2 draws the support and evaluates (B.1).

Algorithm 1 `DS-CGD`: prespecified candidate library and mixture

Require: Conflict graph H on $\mathcal{V}_H$, budget B, frozen scores $\widehat{a}_v$, predictions $\widehat{m}$, coverage fraction $\epsilon \in (0, 1]$, and prespecified search parameters (including $\eta_{\mathrm{search}} > 0$) and solver parameters
Ensure: Compatible library $\mathcal{L}_B$, optimized weights $\widehat{w}$, and exact target-event marginals p

Set proposal weights $q_v \leftarrow \sqrt{\widehat{a}_v + \eta_{\mathrm{search}}}$ with a prespecified positive search floor η_{search}.
Allocation-first candidates
Draw B units by fixed-size rejective sampling with unit weights $q_{i,0} + q_{i,1}$; draw each unit's arm with probability proportional to $q_{i,a}$.
Pilot the compatibility rate of these proposals. If its prespecified binomial lower bound meets the threshold, retain compatible whole-set draws; otherwise run a fixed-B single-site heat-bath search with prespecified restarts from all-control supports (either arm for Global) and retain its distinct visited supports.
Topology-first candidates
For $\beta \in \{0, \frac{1}{2}, 1\}$, order vertices by exponential keys with rate $(1 + d_H(v))^{-\beta}$ and greedily construct independent sets; truncate sets larger than B by weighted sampling without replacement using weights q_v.
Use Laplace-smoothed pilot inclusion counts to select the mode with smallest $\sum_v \widehat{a}_v(1/\widetilde{p}_v - 1)$; generate topology supports with that mode. Pilot counts are not final inclusion probabilities.
Coverage and final mixture
Partition each arm's events into disjoint compatible blocks of size at most B: use balanced blocks for control (and both Global arms) and conflict-aware greedy packing for Direct treated events. Merge these coverage blocks with the candidate supports.
Let $S(v)$ be the unique coverage block containing v.
Form $A_{vk} = \mathbf{1}\{v \in S_k\}$ and $w_k^{\mathrm{ref}} = (2n)^{-1}|\{v : S(v) = S_k\}|$.
Set $\underline{p} \leftarrow \epsilon A w^{\mathrm{ref}}$ and solve (17) on the frozen library; record feasibility and the numerical optimality gap.
Set $p \leftarrow A\widehat{w}$ and **return** $\mathcal{L}_B, \widehat{w}, p, \widehat{m}$.

Algorithm 2 `DS-CGD`: randomization, measurement, and estimation

Require: Frozen $\mathcal{L}_B, \widehat{w}, p, \widehat{m}$ and budget B
Ensure: B measured units and the generalized-difference estimate

Draw $J \sim \mathrm{Categorical}(\widehat{w})$ and set $S \leftarrow S_J$.
Combine the exposure requirements of S into a compatible treatment assignment; complete unspecified coordinates by the prespecified rule.
Measure the units indexed by S and $B - |S|$ additional distinct units.
For each $v = (i, a) \in S$, compute $r_v \leftarrow \sigma_a(Y_i - \widehat{m}_i^a)$, where $\sigma_0 = -1$ and $\sigma_1 = 1$.
Compute $\widehat{\tau}_{\mathrm{DS}} \leftarrow \tau(\widehat{m}) + n^{-1} \sum_{v \in S} r_v / p_v$ and record J, S, the assignment, and the measured-unit set.
return $\widehat{\tau}_{\mathrm{DS}}$.

C EXPERIMENTAL DETAILS

This appendix specifies the simulation graphs, outcome mechanisms, baselines, and evaluation protocol used to validate the theoretical frontiers.

C.1 DATA AND PRETREATMENT INFORMATION

Insurance information in rural China. Cai et al. (2015) study how information sessions and social ties affect weather insurance adoption among rice farming households. We use the friendship ties and survey fields in their replication archive (Cai et al., 2019). Treatment represents an information session and outcome measurement represents a subsequent household survey. The numerical results use the observed network and pretreatment covariates with simulated potential outcomes. The processed network has 7,649 vertices and 20,304 distinct undirected edges. Survey records match 4,848 vertices; the remaining 2,801 vertices lack matched survey fields. Each missing field is imputed by its observed column mean, and the resulting columns are standardized across the 7,649 vertices. Village identifiers define prediction groups where available; missing identifiers use five deterministic graph partitions.

Scientific collaboration. The ca-GrQc data describe coauthorship among authors of arXiv papers in general relativity and quantum cosmology from January 1993 to April 2003 (Stanford Network Analysis Project, n.d.; Leskovec et al., 2007). In a research support interpretation, support is assigned to researchers, collaboration links may transmit information, and subsequent research output requires individual follow-up. We use the observed collaboration network with simulated pretreatment covariates and potential outcomes. SNAP lists 5,242 authors and 14,496 edges; the analysis graph has 5,242 vertices and 14,484 distinct undirected edges after removing self loops and repeated edges.

Table C.1: Processed observed networks. Counts describe the analysis graphs.

Network	Vertices	Edges	Components	Isolated	Covariates
CAI friendship	7,649	20,304	30	0	10 survey fields
ca-GrQc collaboration	5,242	14,484	355	1	2 simulated fields

Table C.2: CAI survey fields used as pretreatment information.

Field	Meaning	Field	Meaning
`age`	Age	`disaster_prob`	Perceived disaster probability
`agpop`	Household size	`general_trust`	Trust in government
`educ`	Education level	`male`	Male household head
`rice_inc`	Rice income share	`risk_averse`	Risk aversion
`ricearea_2010`	Rice production area	`literacy`	Literacy indicator

Synthetic networks. We use Erdős–Rényi (ER), preferential attachment (PA), stochastic block model (SBM), and Watts–Strogatz small world (SW) constructions (Erdős & Rényi, 1959; Barabási & Albert, 1999; Holland et al., 1983; Watts & Strogatz, 1998). Each has $n \in \{2{,}000, 5{,}000\}$. For the four equal SBM blocks, the within block probability is $4/(n/4-1)$ and the between block probability is $1/(3n/4)$. Two independent standard normal covariates are generated at every synthetic vertex and ca-GrQc author.

Table C.3: Synthetic network generators.

Family	Generator	Parameter
ER	$G(n, 3n)$	Exactly $3n$ edges
PA	Preferential attachment	3 links per new vertex
SBM	Four equal blocks	Within $4/(n/4-1)$; between $1/(3n/4)$
SW	Watts–Strogatz	6 neighbors; rewiring probability .10

The loader first imputes and standardizes original covariates over all n vertices. It then adds raw degree d_i, $\log(1+d_i)$, and the neighbor mean of every column in this augmented vector. Empty neighborhoods receive mean zero, implemented with denominator $\max\{d_i, 1\}$. The resulting eight columns for synthetic graphs and ca-GrQc, or 24 columns for CAI, are standardized again over all vertices. Below, x_{i1} and x_{i2} are the first two columns after this second standardization. Synthetic and ca-GrQc groups are assigned by vertex index modulo five. CAI groups use villages and, for vertices without a village identifier, five partitions of a deterministic breadth first traversal.

C.2 Potential outcomes

For Direct effects, $(Y_i(0), Y_i(1)) = (Y_i(00), Y_i(10))$; for Global effects, $(Y_i(0), Y_i(1)) = (Y_i(00), Y_i(11))$. Each contrast has its own stored current potential outcome arrays. Within a contrast, independent Gaussian innovations generate $Y_i(a) = \mu_{ia} + \sigma_{ia}\epsilon_{ia}$, where $\epsilon_{ia} \overset{\text{ind}}{\sim} N(0,1)$. Write $\bar{d} = n^{-1}\sum_i d_i$ and $\widetilde{d}_i = (d_i - \bar{d})/\operatorname{sd}(d)$. All three mechanisms share the systematic means

$$\mu_{i0} = .55x_{i1} - .30x_{i2} + .18\widetilde{d}_i,$$
$$\mu_{i1} = \mu_{i0} + .35 + .12x_{i1}.$$

For the degree-related terms, let

$$g_i = \operatorname{clip}\left(\sqrt{\frac{d_i+1}{\bar{d}+1}}, .5, 2.5\right), \qquad \bar{g} = n^{-1}\sum_i g_i, \quad c(z) = \operatorname{clip}(z, -1.2, 1.2).$$

The neighbor mean of the first standardized covariate is $\bar{x}_{N(i),1}$, set to zero for an empty neighborhood.

Arm asymmetric. The two target exposures have distinct residual scales:

$$\sigma_{i0} = \exp\{c(.70\sin(1.3x_{i1}) + .25\widetilde{d}_i)\},$$
$$\sigma_{i1} = \exp\{c(.75\mathbf{1}\{x_{i2} > 0\} - .30\bar{x}_{N(i),1})\}.$$

Covariate predictable. Both target exposures have the same covariate dependent residual scale, $\sigma_{i0} = \sigma_{i1} = \exp\{c(.60x_{i1})\}$.

Degree aligned. The treated target has degree-related effect heterogeneity and residual scale:

$$\mu_{i1} = \mu_{i0} + .35 + .12x_{i1} + .50(g_i - \bar{g}),$$
$$\sigma_{i0} = 1, \qquad \sigma_{i1} = g_i.$$

The finite-population estimand is $\tau = n^{-1}\sum_i\{Y_i(1) - Y_i(0)\}$. The centered systematic terms give mean effect .35; each realized finite-population effect also contains its drawn Gaussian innovations and is computed from the stored outcome array.

C.3 Historical information and evaluation protocol

Prediction and design inputs. For each target outcome at every node, historical information consists of $R = 200$ independent Gaussian rounds. The implementation stores their exact sufficient statistics $\bar{Y}_{ia}^{\text{hist}}$ and $S_{ia}^{2,\text{hist}}$. A degree two polynomial feature map contains an intercept, squares, and pairwise interactions. Separate grouped five-fold ridge fits predict the historical means in the

two target arms. Each training fold minimizes the *sum* of squared residuals plus $(1/R)\sum_{j\neq 0}\beta_j^2$; the intercept is unpenalized. Cross-fitted predictions $\widehat{m}_{ia}$ and historical variances give

$$\widehat{a}_{ia} = \max\left\{10^{-8}, S_{ia}^{2,\mathrm{hist}}(1-R^{-1}) + (\bar{Y}_{ia}^{\mathrm{hist}} - \widehat{m}_{ia})^2\right\}.$$

The entire $2n$-event risk vector is divided by its mean before optimization.

Target support and estimator. The network, covariates, historical outcomes, and predictions are fixed before the current potential outcomes are drawn. Candidate supports contain at most B compatible target exposure events. Their mixture probabilities are $p_v = (Aw)_v$. Physical measurement records exactly B distinct units, with additional units filling any unused target support positions. The design uses the candidate search and optimized mixture of Appendix B. Every reported method uses the full residual correction GREG estimator; for a sampled support S it is

$$\widehat{\tau}(S) = \frac{1}{n}\sum_i(\widehat{m}_{i1} - \widehat{m}_{i0}) + \frac{1}{n}\sum_{v=1}^{2n}\frac{\mathbf{1}\{v\in S\}}{p_v}s_v(Y_v - \widehat{m}_v), \qquad s_v = \begin{cases}-1, & v\leq n,\\ 1, & v>n.\end{cases}$$

Evaluation and pairing. The synthetic grid has two graph realizations for each network, size, and outcome condition; each observed network condition has one current finite population. `DS-CGD` risk is evaluated exactly over its finite mixture at the stored population. Baseline risks use their recorded inclusion probability procedures and randomization evaluations, with numerical MSE uncertainty retained in the archive. Generalized and Original CGD use analytic probabilities when available and native retention probabilities for the other budgeted designs. CC uses analytic probabilities, independent-set design uses its Rao–Blackwellized inclusion probabilities, and RGCR uses its recorded stratified probability calculation. All methods use the same physical budget and GREG estimator. Within each matched comparison, graph, features, prediction, current population, estimand, and budget agree. Synthetic budget fractions are $\{.01, .02, .05, .10, .20, .40, .70, 1\}$; observed networks use $\{.01, .02, .05, .10, .20, .50, 1\}$, with integer budgets rounded up. CAI has budget-specific prediction fits. Within each budget, the prediction is frozen for every matched method comparison. Absolute table entries equal the square root of arithmetic mean MSE over the matched graph realizations. The main relative plot takes the geometric mean of paired RMSE ratios within each budget. These two summaries are computed separately from the archived MSE records. The abstract compares `DS-CGD` with the lowest-RMSE baseline in each common-group condition and aggregates the paired ratios geometrically over conditions and budgets.

Baseline implementation. Each baseline follows its own assignment law, followed by the common physical measurement budget. The selected target events must be realized under that assignment and involve at most B distinct units. When fewer than B target events enter the estimator, additional distinct units complete the physical measurement set; their outcomes do not enter the GREG residual correction. For the conflict graph baselines, the archived probability source records analytic event probabilities in some conditions and native retention probabilities in the others. Spectral causal clustering uses an analytic probability calculation, while the independent-set design uses Rao–Blackwellization over its focal-unit selection. The RGCR variant in these results uses three-net graph clustering with a focal-first budget. Its inclusion probabilities are calibrated by a stratified procedure based on the first vertex in a uniformly randomized ordering. The mixture of `DS-CGD` is evaluated by its exact finite list of supports, yielding $p_v = (Aw)_v$ directly.

C.4 Additional Experimental Results

We report supplementary absolute-error curves and tabulated comparisons that complement the main experimental results.

Figure C.1 shows Direct and Global absolute RMSE over every budget for $n = 5{,}000$ and the degree-aligned mechanism. Each column fixes one synthetic network family. The shaded bands span the two graph realizations. The RGCR row uses its own matched DS−CGD records.

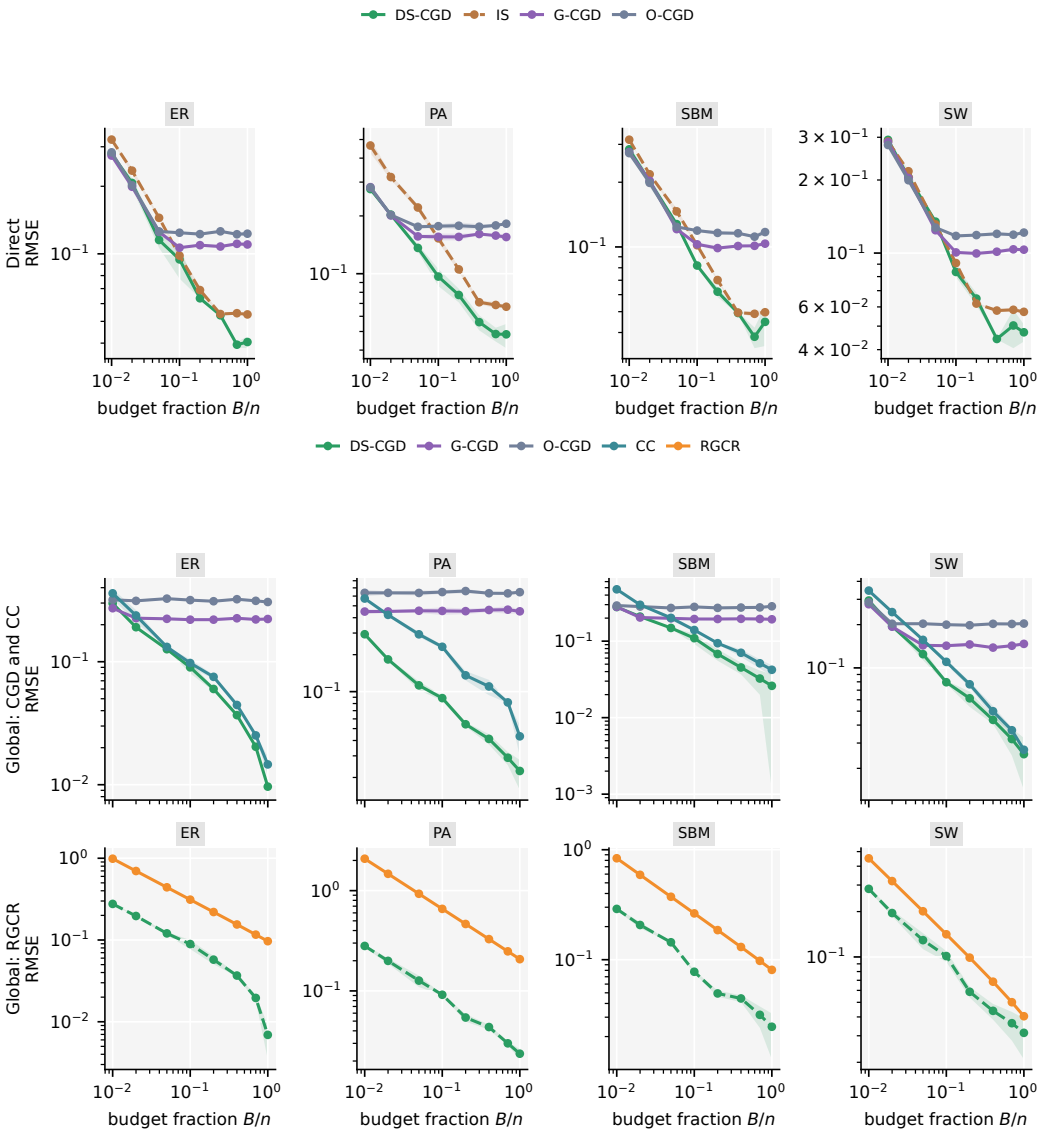


Figure C.1: Absolute RMSE across budgets for synthetic networks with $n = 5{,}000$ and degree-aligned outcomes. The top chart shows Direct effects; the lower chart shows Global effects in a common CGD/CC group and a separate RGCR matched group. Lines are square roots of mean MSE across two graph realizations; shading spans their individual RMSE values. IS denotes independent-set design; G-CGD and O-CGD are Generalized and Original CGD.

Table C.4 covers both sizes and all three outcome mechanisms at the 20% budget. Among its 24 graph, size, and outcome conditions, DS−CGD has the lowest Direct RMSE in 17 and the lowest Global RMSE in all 24 common CGD/CC comparisons. The counts refer to displayed conditions, and the budget charts show how the degree-aligned slice changes with budget.

Table C.4: **Synthetic network RMSE at the 20% budget.** Each entry is the square root of mean MSE over two graph realizations.

		Direct				Global: CGD and CC				Global: RGCR	
Graph	n	IS	G-CGD	O-CGD	DS	G-CGD	O-CGD	CC	DS	RGCR	DS
Arm asymmetric											
ER	2,000	**.1617**	.2619	.2951	.1964	.5236	.7396	.2407	**.1192**	.5351	**.1396**
ER	5,000	.1045	.1619	.1859	**.1025**	.3379	.4881	.1218	**.1039**	.3399	**.0861**
PA	2,000	.2146	.3515	.4070	**.1576**	.9341	1.3545	.3935	**.1171**	.9418	**.1513**
PA	5,000	.1534	.2388	.2629	**.1112**	.6756	.9366	.1886	**.0961**	.6332	**.0959**
SBM	2,000	.1608	.2396	.2777	**.1436**	.4580	.6565	.2198	**.1441**	.4535	**.0980**
SBM	5,000	.1062	.1536	.1749	**.0952**	.2996	.4234	.1481	**.1036**	.2872	**.0789**
SW	2,000	.1473	.2412	.2784	**.1367**	.3609	.4903	.1988	**.1207**	.2443	**.1161**
SW	5,000	**.0905**	.1493	.1691	.0939	.2155	.3174	.1150	**.0827**	.1532	**.0989**
Covariate predictable											
ER	2,000	**.1432**	.2373	.2692	.1586	.4881	.6604	.1904	**.1577**	.4406	**.1137**
ER	5,000	.0890	.1512	.1672	**.0755**	.3091	.4336	.1022	**.0836**	.2833	**.0906**
PA	2,000	.1696	.3225	.3643	**.1173**	.8469	1.2534	.3197	**.1349**	.5924	**.1290**
PA	5,000	**.1083**	.2143	.2455	.1131	.6082	.8415	.1345	**.0990**	.4179	**.0816**
SBM	2,000	.1428	.2217	.2544	**.1397**	.4123	.5697	.2019	**.1159**	.3790	**.1116**
SBM	5,000	.0913	.1393	.1563	**.0747**	.2719	.3815	.1242	**.0985**	.2375	**.0601**
SW	2,000	.1356	.2196	.2556	**.1177**	.3038	.4425	.1701	**.1131**	.2193	**.1274**
SW	5,000	**.0823**	.1440	.1615	.0949	.1983	.2751	.1016	**.0749**	.1341	**.0865**
Degree aligned											
ER	2,000	.1066	.1747	.1978	**.1017**	.3578	.5127	.1489	**.0763**	.3429	**.0545**
ER	5,000	.0689	.1093	.1224	**.0634**	.2198	.3103	.0753	**.0600**	.2199	**.0575**
PA	2,000	.1677	.2323	.2685	**.1261**	.6275	.8728	.4244	**.1108**	.6415	**.1061**
PA	5,000	.1053	.1555	.1778	**.0775**	.4522	.6594	.1358	**.0542**	.4650	**.0543**
SBM	2,000	.1109	.1580	.1846	**.0786**	.2999	.4381	.1519	**.0957**	.2923	**.0731**
SBM	5,000	.0701	.0988	.1163	**.0619**	.1946	.2721	.0940	**.0679**	.1858	**.0493**
SW	2,000	**.0981**	.1638	.1858	.1212	.2345	.3204	.1309	**.1307**	.1556	**.0870**
SW	5,000	**.0619**	.0996	.1186	.0649	.1456	.1981	.0769	**.0614**	.0990	**.0587**

Note. Boldface denotes the minimum RMSE within each matched group. IS: independent-set design; G-CGD: Generalized CGD; O-CGD: Original CGD; CC: spectral causal clustering; DS: `DS-CGD`. RGCR and its DS comparison form a separate matched group.

Table C.5: **Observed network RMSE.** Three budgets per mechanism; each entry uses one current population.

		Direct				Global: CGD and CC				Global: RGCR	
Graph	B	IS	G-CGD	O-CGD	DS	G-CGD	O-CGD	CC	DS	RGCR	DS
Arm asymmetric											
CAI	77	.4372	**.3170**	.3220	.3330	**.3148**	.3216	.3178	.3185	.6885	**.3185**
CAI	1,530	.0998	.1363	.1480	**.0659**	.2130	.3025	.0707	**.0587**	.1528	**.0587**
CAI	7,649	.0632	.1324	.1484	**.0466**	.2178	.3115	.0229	**.0195**	.0653	**.0195**
ca-GrQc	53	.7436	.4188	.4229	**.3896**	.5097	.7282	.4450	**.3746**	1.2521	**.3810**
ca-GrQc	1,049	.1482	.3808	.4040	**.1083**	.5220	.7107	**.0916**	.1026	.2808	**.0831**
ca-GrQc	5,242	.1035	.3764	.4161	**.0652**	.5116	.7424	**.0225**	.0226	.1245	**.0353**
Covariate predictable											
CAI	77	.4084	.3249	.3227	**.3032**	.3230	.3147	.3155	**.3049**	.6298	**.3049**
CAI	1,530	.0895	.1402	.1487	**.0776**	.2189	.3056	.0637	**.0427**	.1394	**.0427**
CAI	7,649	.0595	.1390	.1529	**.0352**	.2201	.2979	.0082	**.0075**	.0591	**.0075**
ca-GrQc	53	.5806	.3736	.3782	**.3499**	.4664	.6607	.3937	**.3490**	.9679	**.3728**
ca-GrQc	1,049	.1157	.3286	.3664	**.0966**	.4697	.6724	**.0813**	.0883	.2166	**.0562**
ca-GrQc	5,242	.0774	.3275	.3710	**.0565**	.4579	.6849	.0151	**.0042**	.0950	**.0351**
Degree aligned											
CAI	77	.2976	**.2221**	.2255	.2261	**.2211**	.2215	.2260	.2420	.4776	**.2420**
CAI	1,530	.0671	.0952	.1109	**.0491**	.1493	.2278	.0461	**.0451**	.1059	**.0451**
CAI	7,649	.0450	.0912	.1118	**.0330**	.1555	.2217	**.0080**	.0085	.0449	**.0085**
ca-GrQc	53	.5450	.2740	.2819	**.2564**	.3458	.4847	.2934	**.2736**	.9312	**.2807**
ca-GrQc	1,049	.1094	.2444	.2697	**.0893**	.3513	.4796	**.0605**	.0638	.2088	**.0541**
ca-GrQc	5,242	.0775	.2361	.2817	**.0653**	.3403	.4767	.0119	**.0069**	.0923	**.0283**

Note. Boldface denotes the minimum RMSE within each matched group. Abbreviations are in Table C.4. ca-GrQc RGCR uses a separate matched prediction and population.

D PROOFS OF THE MAIN RESULTS

Notation. Potential outcomes and the physical network are fixed. Expectations are over the experimental design; probability distributions on potential outcomes enter only the lower-bound arguments. Write $\sigma_0 = -1$, $\sigma_1 = 1$, and $x_{i,a} = \sigma_a y_i^a$. This change of signs preserves the armwise moment constraints and gives $\tau(x) = n^{-1}\sum_{i,a} x_{i,a}$. The corresponding parameter space is

$$\mathcal{X}_q(M) = \{x = (x_0, x_1) : \|x_a\|_q \le n^{1/q} M,\ a = 0, 1\}, \qquad 2 \le q \le \infty.$$

Here $x_a = (x_{i,a})_{i=1}^n$ and $n^{1/\infty} = 1$. We use $K = \mathcal{X}_q(M)$ when the arguments are fixed. In particular, $|\tau(x)| \le 2M$ and $\|x\|_2^2 \le 2nM^2$.

Symbol	Meaning
$H, \mathcal{I}_B(H)$	Exposure-conflict graph and its independent sets of size at most B.
$P = (w, T)$	Pre-outcome design $S \sim w$ and a Borel-measurable estimator $T_S(x_S)$.
$r_P(x), b_P(x)$	Design mean-squared error and design bias at the fixed schedule x.
$R(B, H; q, M)$	Minimax risk over all designs and estimators.
$R(B, H, \delta; q, M)$	Minimax risk under the uniform bias constraint $\sup_{x\in K} \lvert b_P(x)\rvert \le \delta$.
d_i, Δ, k	Physical degree, maximum physical degree, and a degeneracy bound.
s, t, v_i	$s = 1 - 1/q$, $t = q/(q-2)$, and $v_i = d_i + 1$; $t = \infty$ at $q = 2$ and $t = 1$ at $q = \infty$.
D_u, A_G, A_H	The diagonal matrix with diagonal u, and the physical and conflict adjacency matrices.

All vector norms are unnormalized, and vector powers are coordinatewise. Matrix norms are Euclidean operator norms. The order $\preceq$ is the positive-semidefinite order. The comparisons $\lesssim_k$ and $\asymp_k$ allow constants depending only on k; $\lesssim$ and $\asymp$ have universal constants, including uniformly over $q \in [2, \infty]$.

We retain the functionals from the main text:

$$\begin{aligned} U(G, q) &= \|(v_i^s)_{i\in[n]}\|_t, \\ \Phi(G, q) &= \min_{D\subseteq[n]} \left\{|D|^{2s} + \|(v_i^s)_{i\notin D}\|_t\right\}, \\ \Theta(G, q) &= \min\left\{1, n^{-2s}\|(v_i^{2s})_{i\in[n]}\|_t\right\}. \end{aligned}$$

The norm of an empty vector is zero. The physical protocol measures exactly B distinct units. Its selected target support has at most B elements; additional measured outcomes can be omitted from the estimator.

D.1 PROOF OF PROPOSITION 1

Proof. For the physical experiment, let

$$U \sim P, \qquad Z = Z(U), \qquad \mathcal{M} = \mathcal{M}(U), \qquad |\mathcal{M}| = B,$$

where the law P is fixed independently of the current potential outcomes. The target exposure events realized among the measured units are

$$S_{\mathrm{obs}}(U) = \{(i, a) : i \in \mathcal{M}(U),\ Z_{N_i^+} = e_{i,a}\},$$

and the estimator may use a recorded sub-support $S_{\mathrm{ana}}(U) \subseteq S_{\mathrm{obs}}(U)$, with $|S_{\mathrm{ana}}(U)| \le B$. Thus the reduction retains the exposure events used for analysis while allowing the physical protocol to measure additional units to complete its exactly-B budget.

For $S \in \mathcal{I}_B(H)$, the partial treatment assignments of its events agree on all overlaps. Their union extends to a full assignment $z(S)$. Same-unit conflicts ensure that S represents distinct units.

On the submodel with all nontarget outcomes zero, let $S = S_{\mathrm{obs}}(U)$ be the measured target set, and write the physical data as $O(U, x_S)$. Because the two local target assignments are distinct, setting all other local potential outcomes to zero defines a valid submodel under neighborhood interference.

The support $S_{\mathrm{obs}}(U)$ and its conditional randomization kernel are the same for every x. Conditional averaging defines a single Borel rule $\widetilde{T}_S(y) = \mathbb{E}[T_{\mathrm{phys}}\{O(U,y)\} \mid S_{\mathrm{obs}}(U) = S]$ and gives

$$\mathbb{E}\widetilde{T}_S(x_S) = \mathbb{E}T_{\mathrm{phys}},$$
$$\mathbb{E}\{\widetilde{T}_S(x_S) - \tau(x)\}^2 = \mathbb{E}\big[\{\mathbb{E}[T_{\mathrm{phys}} - \tau(x) \mid S]\}^2\big] \le \mathbb{E}\{T_{\mathrm{phys}} - \tau(x)\}^2.$$

For finite-risk procedures these conditional expectations are integrable; the integrand is Borel measurable in both the record and the observed coordinates. Zero-probability supports can be assigned arbitrary outputs. Taking the supremum over $x \in K$ and then the infimum over physical procedures proves $R(B, H, \delta; q, M) \le R_{\mathrm{phys}}(B, \delta; q, M)$.

Conversely, draw and record $S \sim w$, assign $z(S)$, and measure its units. If $|S| < B$, measure $B - |S|$ further distinct units and ignore their outcomes. The output $T_S(x_S)$ has exactly the graph procedure's loss and bias on every potential-outcome schedule, proving the reverse inequality.

Finally, the bijection $x = Mx'$ and $T_S(x_S) = MT'_S(x_S/M)$ gives

$$b_T(x) = Mb_{T'}(x'), \qquad r_T(x) = M^2 r_{T'}(x').$$

The parameter class becomes $\mathcal{X}_q(1)$ and the bias allowance becomes δ/M, proving the scaling identity. □

D.2 The measurement lower bound

Measurement bound. *Under the preceding setup,*

$$R(B, H, \delta; q, M) \ge \max\{R(n, H, \delta; q, M), M^2/B\}.$$

The same bound holds without a bias constraint.

Proof. Every budget-B graph procedure is feasible at budget n. Taking the infimum over the larger class gives the full-measurement term.

For $B = 1$, give each signed arm an independent common sign of magnitude M. A single observed coordinate reveals one sign, while the other sign remains independent and contributes M^2 to the posterior variance of τ. These tables belong to the moment class for every $q \ge 2$. Thus the measurement term holds when $B = 1$.

For $B \ge 2$, put $N = 2n$ and allow B distinct coordinate queries without any conflicts. This enlarges the experiment. On the bounded submodel $\{-M, M\}^N$, take $P \sim \mathrm{Beta}(a, a)$ and, conditional on P, independent $Z_v \sim \mathrm{Bernoulli}(P)$, with $X_v = M(2Z_v - 1)$. Let Q count successes among the queried coordinates and put $m = N - B$. Then $P \mid X_S \sim \mathrm{Beta}(a + Q, a + B - Q)$. Conditional variance and the law of total variance give

$$\mathbb{E}\,\mathrm{Var}(\tau \mid X_S) = \frac{4M^2}{n^2}\left\{m\mathbb{E}[P(1-P)] + m^2\mathbb{E}\,\mathrm{Var}(P \mid X_S)\right\},$$
$$\mathbb{E}[P(1-P)] = \frac{a}{2(2a+1)}, \qquad \mathbb{E}\,\mathrm{Var}(P \mid X_S) = \frac{a}{2(2a+1)(B+2a)}.$$

Consequently,

$$\mathbb{E}\,\mathrm{Var}(\tau \mid X_S) = \frac{2a(N-B)(N+2a)}{(2a+1)(B+2a)n^2}M^2.$$

For $B \ge 2$, choose $a = B/2$. Since $N \ge 2B$,

$$\mathbb{E}\,\mathrm{Var}(\tau \mid X_S) = \frac{2M^2}{B+1}\left(1 - \frac{B^2}{N^2}\right) \ge \frac{3M^2}{2(B+1)} \ge \frac{M^2}{B}.$$

Exchangeability makes this Bayes risk independent of the selected coordinate labels. It lower-bounds the risk of every estimator in the enlarged experiment, and hence the original minimax risk at every bias allowance. □

D.3 COMPATIBLE SAMPLING AND BUDGET REDUCTION

The next lemma supplies the sampling construction used in both main theorems. All coefficient vectors in its proof have finite support in probability.

Lemma D.1 (Compatible sampling). *If $\omega > 0$ and $D_\omega \succeq A_H$, then*

$$R(B, H, 0; q, M) \lesssim M^2 \left\{ B^{-1} + n^{-2s}(\|\omega_0\|_t + \|\omega_1\|_t) \right\}. \tag{D.1}$$

For $0 < \gamma_i \le 1$, put $\Gamma = D_\gamma$ and $\Lambda(Q, q) = \max_{z \ge 0,\, \|z\|_q \le 1} z^\top Q z$ for symmetric, entrywise nonnegative Q. There is a Direct-effect procedure satisfying

$$\begin{aligned} &\sup_{x \in K} |\mathbb{E}\widehat{\tau} - \tau(x)| = 2Mn^{-s}\|\mathbf{1} - \gamma\|_{1/s}, \\ &\sup_{x \in K} \mathbb{E}(\widehat{\tau} - \tau)^2 \lesssim M^2 n^{-2s} \left\{ \frac{\|\gamma\|_{1/s}^2}{B} + \|\gamma^2\|_t + \Lambda(\Gamma A_G \Gamma, q) + \|\mathbf{1} - \gamma\|_{1/s}^2 \right\}. \end{aligned} \tag{D.2}$$

Both designs use at most B selected targets and exactly B physical measurements. Bounds on retained targets keep the original denominator n.

Proof of Lemma D.1. **Step 1. A compatible coefficient vector.** The scaled matrix is symmetric and entrywise nonnegative. Its largest absolute eigenvalue equals its largest eigenvalue, so the diagonal majorant implies $\|D_\omega^{-1/2} A_H D_\omega^{-1/2}\| \le 1$. On every remaining vertex set U, this inequality gives

$$\sum_{i \in U} \frac{1}{\omega_i} \sum_{j \in U} \frac{(A_H)_{ij}}{\omega_j} \le \sum_{i \in U} \frac{1}{\omega_i}.$$

Thus some vertex has inverse-weighted degree at most one. Successively removing such vertices gives a priority order with $\ell_i = \sum_{j \in P(i)} \omega_j^{-1} \le 1$, where $P(i)$ denotes the higher-priority neighbors of i.

Fix $u = 2/3$. Propose vertices independently with probabilities $q_i = u/(\omega_i + u)$, and accept i when it is proposed and no vertex in $P(i)$ is proposed. Write I_i for acceptance, $\pi_i = q_i \prod_{j \in P(i)}(1 - q_j)$, $Z_i = I_i/\pi_i$, and $D_{ji} = \mathbf{1}\{j \in P(i)\}$. Then $\mathbb{E}Z = \mathbf{1}$ and $\operatorname{supp} Z \in \mathcal{I}(H)$. For distinct nonadjacent i, j, and for the diagonal, respectively,

$$\begin{aligned} \operatorname{Cov}(Z_i, Z_j) &= \prod_{h \in P(i) \cap P(j)} (1 + u/\omega_h) - 1 \le (e^u - 1) \sum_{h \in P(i) \cap P(j)} \omega_h^{-1}, \\ \operatorname{Var}(Z_i) &\le (e^u/u)\omega_i + (e^u - 1)\ell_i. \end{aligned}$$

Adjacent entries equal -1. Since $D + D^\top = A_H$ and $\|D_\omega^{-1/2} D D_\omega^{-1/2}\| \le 1$, these bounds give, for every y,

$$\begin{aligned} y^\top \operatorname{Cov}(Z) y &\le (e^u/u) y^\top D_\omega y + |y|^\top A_H |y| + (e^u - 1)|y|^\top D^\top D_\omega^{-1} D |y| \\ &\lesssim y^\top D_\omega y. \end{aligned}$$

Thus $\operatorname{Cov}(Z) \preceq C_0 D_\omega$ for a universal C_0.

Step 2. Reduction to at most B observed coordinates. Let $C \ge 0$ be a finite random coefficient vector with compatible support, mean μ, and $\operatorname{Cov}(C) \preceq D_v$. Choose $D_d \succeq \mu\mu^\top$. For $J = \mathbb{E}CC^\top \preceq D_{v+d}$, the nonnegative symmetric matrix $D_{v+d}^{-1/2} J D_{v+d}^{-1/2}$ has spectral radius at most one. A positive Perron vector on each nonzero irreducible block, followed by rescaling, gives $z > 0$ with $Jz \le D_{v+d} z$. Coordinates with zero second moment can be omitted. Given $C = c \ne 0$, sample B indices independently with probabilities $p_i(c) = z_i c_i / (z^\top c)$ and set $\widehat{C}_i = c_i N_i / \{B p_i(c)\}$, where N_i is the count of i. Zero coordinates, including $c = 0$, contribute zero. We have $\mathbb{E}(\widehat{C} \mid C) = C$, $\operatorname{supp} \widehat{C} \subseteq \operatorname{supp} C$, and $|\operatorname{supp} \widehat{C}| \le B$. The covariance calculation is

$$\begin{aligned} \ell_i' := \mathbb{E}\frac{C_i^2}{p_i(C)} &= \frac{(Jz)_i}{z_i} \le v_i + d_i, \\ \operatorname{Cov}(\widehat{C}) &= (1 - B^{-1}) \operatorname{Cov}(C) + B^{-1}(D_{\ell'} - \mu\mu^\top) \\ &\preceq D_v + B^{-1} D_d. \end{aligned}$$

To control the moment cost, put $p = 1/s$, $h_a = \|\mu_a\|_p$ and $H_\mu = h_0 + h_1$. On nonzero coordinates take $d_{i,a} = H_\mu h_a^{p-1} |\mu_{i,a}|^{2-p}$. The identities

$$\sum_{i,a} \frac{\mu_{i,a}^2}{d_{i,a}} = 1, \qquad \|d_0\|_t + \|d_1\|_t = H_\mu^2$$

follow from $(2-p)t = p$. At $p = 2$ use a constant diagonal on each nonzero arm; at $p = 1$ use $d = H_\mu |\mu|$. Therefore

$$\sup_{x \in K} \mathrm{Var}(n^{-1} \widehat{C}^\top x) \le M^2 n^{-2s} \left\{ \|v_0\|_t + \|v_1\|_t + \frac{H_\mu^2}{B} \right\}.$$

Apply this to $C = Z$, $\mu = \mathbf{1}$. Conditional averaging of $\widehat{C}$ given its support preserves the mean and reduces risk, so the estimator has the required form $T_S(x_S)$. This proves (D.1).

Step 3. Coefficients with prescribed means. For symmetric nonnegative Q, there is $h \ge 0$ such that $D_h \succeq Q$ and $\|h\|_t = \Lambda(Q, q)$. For $2 < q < \infty$, a maximizing z is positive on every nonzero connected component and satisfies $Qz = \Lambda(Q, q) z^{q-1}$. Set $h_i = (Qz)_i / z_i$; for every real u,

$$u^\top (D_h - Q) u = \frac{1}{2} \sum_{i,j} Q_{ij} z_i z_j \left(\frac{u_i}{z_i} - \frac{u_j}{z_j} \right)^2 \ge 0.$$

Moreover, $(q-2)t = q$ gives the norm identity. At $q = 2$ take $h = \|Q\|\mathbf{1}$; at $q = \infty$ take $h = Q\mathbf{1}$. Zero components have zero matrix cost. A vanishing positive diagonal perturbation supplies strictly positive proposal parameters on such components.

For $Q = \Gamma A_G \Gamma$, the Direct conflict matrix satisfies

$$D_{\gamma \oplus \gamma} A_{H_{\mathrm{D}}} D_{\gamma \oplus \gamma} = \begin{pmatrix} 0 & D_{\gamma^2} + Q \\ D_{\gamma^2} + Q & Q \end{pmatrix} \preceq \mathrm{diag}(D_{\gamma^2 + h}, D_{\gamma^2 + 2h}).$$

The block inequality follows by separating the diagonal contribution and the two-block form with diagonal D_h and off-diagonal $-Q$. Unscale, apply the compatible construction, and then multiply its coefficients by $\gamma \oplus \gamma$. Their mean is $\mu = \gamma \oplus \gamma$, and the budget calculation above applies. Finally, armwise duality gives

$$\sup_{x \in K} \left| \frac{1}{n} \sum_{a=0}^{1} (\gamma - \mathbf{1})^\top x_a \right| = 2M n^{-s} \|\mathbf{1} - \gamma\|_{1/s}.$$

Adding squared bias proves (D.2). Measuring additional distinct units completes the physical budget. □

D.4 Proof of Theorem 1

Proof. **Step 1. Direct upper bounds.** Fix $D \subseteq [n]$, assign its units to control, and retain targets on $J = G[[n] \setminus D]$. Let $u_i = d_J(i) + 1$ and choose an acyclic outdegree-k orientation of J. Define

$$\omega_i = 1 + k^{1/q} \, \mathrm{indeg}_J(i) u_i^{-1/q} + k^{-1/q} \sum_{j : i \to j} u_j^{1/q}.$$

Weighted Young's inequality on each directed edge gives $D_\omega \succeq A_J$. For finite t,

$$\sum_i \left(\sum_{j : i \to j} u_j^{1/q} \right)^t \le k^{t-1} \sum_j \mathrm{indeg}_J(j) u_j^{t/q} \le k^{t-1} \sum_j u_j^{st},$$
$$\|\omega\|_t \lesssim_k \|(u_i^s)_i\|_t,$$

using $1 + t/q = st$. Maxima at $q = 2$ and sums at $q = \infty$ give the same bound. The Direct conflict matrix has the diagonal majorant

$$A_{H_{\mathrm{D}}(J)} = \begin{pmatrix} 0 & I + A_J \\ I + A_J & A_J \end{pmatrix} \preceq \mathrm{diag}(I + D_\omega, I + 2D_\omega).$$

Lemma D.1 therefore gives a retained-target unbiased estimator. Keeping the original denominator n, its bias and MSE satisfy

$$\sup_{x\in K}|\mathbb{E}\widehat{\tau}-\tau(x)| \le \frac{|D|^{1-1/q}}{n}(\|x_0\|_q+\|x_1\|_q) \le 2M(|D|/n)^s,$$
$$\sup_{x\in K}\mathbb{E}(\widehat{\tau}-\tau)^2 \lesssim_k M^2\left\{\frac{1}{B}+n^{-2s}\|(u_i^s)_{i\notin D}\|_t+n^{-2s}|D|^{2s}\right\}$$
$$\le C_k M^2\left\{\frac{1}{B}+n^{-2s}\left[\|(v_i^s)_{i\notin D}\|_t+|D|^{2s}\right]\right\}.$$

Minimizing over D proves the unrestricted upper bound. Taking $D=\varnothing$ proves the exactly unbiased upper bound. For $D=[n]$, use the zero estimator.

Step 2. Direct lower bound over all estimators. First take a star forest F on the original n vertices. Let $\ell_j \ge 1$ be the leaf counts of its nontrivial stars. Fix any procedure with worst risk $\mathcal{E}$, and write $V=n^{2s}\mathcal{E}/M^2$. Independent signs of magnitude M on all target coordinates give $V \ge n^{1/t}$. Let p_j be the probability of observing star j's treated center.

Figure D.1 depicts the spanning star forests and the two separate finite-population sign constructions used below.

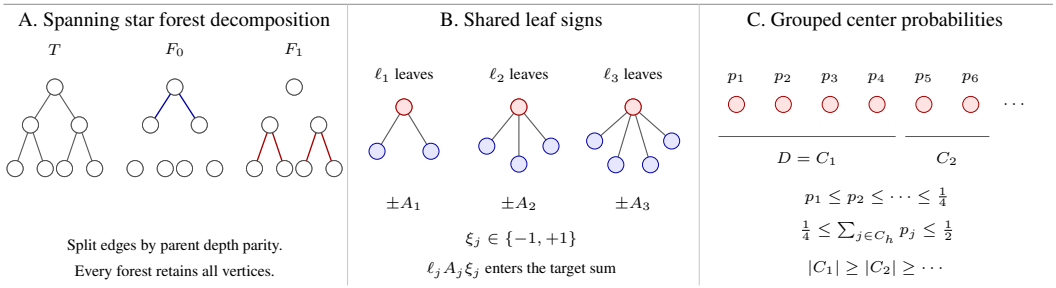


Figure D.1: Construction underlying the unrestricted Direct lower bound. Panel A illustrates the depth parity split on a tree; a graph of degeneracy at most k yields at most $2k$ spanning star forests. Panel B uses independent signs across stars and one common control leaf sign within each star: observing its treated center leaves both $+A_j$ and $-A_j$ possible. Its contribution $\ell_j A_j \xi_j$ yields the quadratic leaf count in the risk bound. Panel C uses a separate potential outcome family with a common treated-center sign in each group. Centers with $p_j \le 1/4$ are ordered; each complete group closes when its cumulative probability first reaches $1/4$, and the final group may be incomplete. The first group is D. Both potential outcome families test the same design distribution.

Give its control leaves a common sign of amplitude A_j, independently across stars, and set all other targets to zero. Observing a treated center hides its leaf sign. Optimizing posterior variance over the amplitudes gives

$$\mathcal{E} \ge \frac{1}{n^2}\sum_j p_j \ell_j^2 A_j^2, \quad \|(\ell_j^{1/q}A_j)_j\|_q \le n^{1/q}M \quad\Longrightarrow\quad \|(p_j\ell_j^{2s})_j\|_t \le V.$$

This is $\ell_{q/2}$–ℓ_t duality after rescaling A_j by $\ell_j^{1/q}$, with the maximum and sum interpretations at the endpoints. For disjoint groups of centers C_h with $\sum_{j\in C_h} p_j \le 1/2$, give the treated centers in a group one common sign. Writing $m_h=|C_h|$, the chance that this sign is hidden is at least $1/2$, so the same calculation gives $\|(m_h^{2s})_h\|_t \le 2V$.

Sort centers with $p_j \le 1/4$ by increasing p_j and form consecutive groups, closing a group when its probability sum reaches $1/4$. Every sum is at most $1/2$, and the group sizes are nonincreasing.

Delete the first group D. Each center of group $h+1$ has $p_j \geq 1/(4m_h)$; hence

$$\sum_{j\notin D:\, p_j\leq 1/4} p_j^{-t} \leq 4^t \sum_h m_{h+1} m_h^t \leq 4^t \sum_h m_h^{t+1} = 4^t \sum_h (m_h^{2s})^t,$$
$$|D|^{2s} \lesssim V, \qquad \|(p_j^{-1})_{j\notin D}\|_t \lesssim V,$$
$$\|(\ell_j^s)_{j\notin D}\|_t \leq \|(p_j \ell_j^{2s})_j\|_t^{1/2} \|(p_j^{-1})_{j\notin D}\|_t^{1/2} \lesssim V.$$

For centers with $p_j > 1/4$, the ℓ_t norm of (p_j^{-1}) is at most $4n^{1/t} \leq 4V$. Zero-probability centers belong to the first group D. The displayed interpolation bound therefore includes every retained center. At $t = \infty$ use maxima in the same inequalities. Centers satisfy $d_F(j) + 1 \leq 2\ell_j$, and all other degrees are at most one. Consequently $\Phi(F, q) \lesssim V$.

To pass to G, label outgoing edges distinctly in an acyclic outdegree-k orientation. Each label class is a forest; splitting each rooted tree by depth parity yields at most $2k$ spanning star forests F_j whose edge union is G. If D_j minimizes the functional of F_j, take $D = \bigcup_j D_j$. Subadditivity of u^s and Minkowski give

$$v(G) \leq \sum_j v(F_j),$$
$$\Phi(G, q) \leq (2k)^{2s-1} \sum_j |D_j|^{2s} + \sum_j \|(v_i(F_j)^s)_{i\notin D_j}\|_t \lesssim_k \max_j \Phi(F_j, q).$$

Physical edge deletion only removes Direct conflicts. Therefore $\mathcal{E} \gtrsim_k M^2 n^{-2s} \Phi(G, q)$. Combining this with the measurement bound proves the first lower bound.

Step 3. The lower bound under design unbiasedness. The needed lower bound applies to non-linear estimators as well. Let independent fair signs ξ_j have target $\sum_j h_j \xi_j$. A design record S reveals the signs in $O(S)$; put $\pi_j = \mathbb{P}\{j \in O(S)\}$. For an estimator unbiased at every sign vector, set $a_j(S) = \mathbb{E}_\xi[T(S, \xi_{O(S)})\xi_j]$. Then $a_j = 0$ off the revealing event and $\mathbb{E}_S a_j = h_j$. Orthogonal projection and Cauchy–Schwarz yield

$$\mathbb{E}_{S,\xi}(T - \tau)^2 \geq \sum_j \mathbb{E}_S(a_j - h_j)^2 \geq \sum_j h_j^2(\pi_j^{-1} - 1).$$

A nonzero h_j with $\pi_j = 0$ makes unbiasedness impossible.

Apply this first to independent treated-center signs on F, then to one common control-leaf sign per star. The same design probabilities p_j satisfy

$$\|((1-p_j)/p_j)_j\|_t \leq V, \qquad \|(\ell_j^{2s} p_j/(1-p_j))_j\|_t \leq V,$$
$$\|(\ell_j^s)_j\|_t \leq \|((1-p_j)/p_j)_j\|_t^{1/2} \|(\ell_j^{2s} p_j/(1-p_j))_j\|_t^{1/2} \leq V.$$

Here $0 < p_j < 1$, and the probability of seeing a control leaf is at most $1 - p_j$. Thus $U(F, q) \lesssim V$. The same star-forest decomposition gives $U(G, q) \leq \sum_j U(F_j, q) \lesssim_k \max_j U(F_j, q)$, proving $\mathcal{E} \gtrsim_k M^2 n^{-2s} U(G, q)$. Add the measurement lower bound.

Step 4. Global lower and upper bounds. For a star forest with component sizes b_j (isolates have size one), give each component two independent common arm signs of amplitude A_j. No feasible observation reveals both signs in a component. Hence

$$\|(b_j^{1/q} A_j)_j\|_q \leq n^{1/q} M \quad \Longrightarrow \quad \mathcal{E} \geq M^2 n^{-2s} \|(b_j^{2s})_j\|_t \gtrsim M^2 n^{-2s} \|v(F)^{2s}\|_t,$$
$$\|v(G)^{2s}\|_t \leq (2k)^{2s} \max_j \|v(F_j)^{2s}\|_t.$$

The last inequality is Minkowski in ℓ_{2st}; at $t = \infty$ it is a maximum inequality. Global conflicts also increase under physical edge addition. This proves the unrestricted topology lower bound, and therefore the lower bound at every bias allowance.

For attainment, put $h_i = \max_{j \in N_i^+}(d_j + 1)$ and $C_{ij} = \mathbf{1}\{\mathrm{dist}_G(i,j) \le 2\}$. Write $A_G = P + P^\top$ with outdegree at most k. Rowwise Cauchy–Schwarz gives

$$\|Pz\|_2^2 \le k \sum_j \mathrm{indeg}(j) z_j^2 \le k z^\top D_h z,$$
$$\|P^\top z\|_2^2 \le \sum_i z_i^2 \sum_{j: i \to j} \mathrm{indeg}(j) \le k z^\top D_h z,$$
$$|z^\top C z| \le |z|^\top (I + A_G + A_G^2)|z| \le z^\top D_h z + \|(P + P^\top)|z|\|_2^2 \lesssim_k z^\top D_h z.$$

Thus $A_{H_G} = \left(\begin{smallmatrix} 0 & C \\ C & 0 \end{smallmatrix}\right)$ has a diagonal majorant proportional to $D_{h \oplus h}$. Also

$$\sum_i h_i^t \le \sum_i \sum_{j \in N_i^+} v_j^t = \sum_j v_j^{t+1} = \sum_j v_j^{2st}, \qquad \|h\|_t \le \|v^{2s}\|_t,$$

with the maximum interpretation at $q = 2$. Lemma D.1 gives an unbiased procedure of risk $\lesssim_k M^2\{B^{-1} + n^{-2s}\|v^{2s}\|_t\}$. If the raw functional exceeds one, instead sample a uniform target coordinate I and report $2x_I$, with risk at most $4M^2$. This gives the clipped functional Θ. The unbiased upper bound and unrestricted lower bound sandwich every bias-constrained value. □

D.5 Measurement scale and the cost of unbiasedness

Budget scale and cost of unbiasedness. *Let $\Theta_{\mathrm{D}} = n^{-2s}\Phi$, $\Theta_{\mathrm{G}} = \Theta$, and $B^\star(G, c, q) = \Theta_c^{-1}$. Then*

$$\frac{R(B, H_c(G); q, M)}{R(n, H_c(G); q, M)} \asymp_k 1 + \frac{B^\star(G, c, q)}{B}, \qquad R(n, H_c(G); q, M) \asymp_k M^2 \Theta_c.$$

For the Direct effect, the full-budget cost of exact unbiasedness on the same graph is of order U/Φ; its maximum over graphs with fixed $k \ge 1$ is of order $n^{1/q}$.

Proof. Since $v_i \ge 1$, $2s - 1 = 1/t$, and $\Phi \le U$,

$$\Phi \ge \min_{m \in \{0,1,\dots,n\}} \{m^{2s} + (n-m)^{1/t}\} \ge n^{1/t}, \qquad \Theta \ge n^{-1}.$$

Use the empty-tail convention when $t = \infty$. Hence $\Theta_c \ge n^{-1}$. Applying Theorem 1 at B and at n gives

$$\frac{R(B, H_c(G))}{R(n, H_c(G))} \asymp_k \frac{B^{-1} + \Theta_c}{n^{-1} + \Theta_c} \asymp_k 1 + \frac{1}{B\Theta_c}.$$

Likewise the Direct ratio is $(n^{2s}/B + U)/(n^{2s}/B + \Phi)$ up to k-dependent constants. At $B = n$, using $st \ge 1$ and $\sum_i v_i \lesssim_k n$ gives

$$\frac{U}{\Phi} \le \frac{\|v\|_{st}^s}{n^{2s-1}} \le \frac{(\sum_i v_i)^s}{n^{2s-1}} \lesssim_k n^{1-s} = n^{1/q}.$$

A single star has $U \asymp n^s$, while deleting its center gives $\Phi \asymp n^{2s-1}$, attaining the reverse order. □

At $q = 2$, $U = \sqrt{\Delta + 1}$, $\Phi = \min_m\{m + \sqrt{d_{(m+1)} + 1}\}$, and $\Theta = (\Delta + 1)/n$, where the square-root term is zero for the empty tail. At $q = \infty$, $\Phi \asymp_k n$ and $B^\star(G, \mathrm{G}, \infty) = \max\{1, n/[(\bar{d} + 1)^2 + \mathrm{Var}_n(d)]\}$, where $\bar{d}$ and $\mathrm{Var}_n(d)$ are the mean and variance of the degree sequence.

D.6 Other homogeneous exposure contrasts

Proposition (Homogeneous exposure contrasts). *On a network without isolated units, write ab for own treatment a and common neighbor treatment b. The contrasts 00 versus 10 and 01 versus 11 have the Direct rates of Theorem 1. The other four contrasts have the Global rate at every bias allowance and without a bias restriction.*

Proof. Put $A = A_G$, $C_{ij} = \mathbf{1}\{\mathrm{dist}_G(i,j) \le 2\}$, and $Q_{ij} = \mathbf{1}\{i = j \text{ or } (A^2)_{ij} > 0\}$. In the target orders $(00, 10)$, $(00, 11)$, $(11, 10)$, and $(10, 01)$, respectively,

$$A_{H_{\mathrm{D}}} = \begin{pmatrix} 0 & I + A \\ I + A & A \end{pmatrix}, \quad A_{H_{\mathrm{G}}} = \begin{pmatrix} 0 & C \\ C & 0 \end{pmatrix}, \quad A_{H_{\mathrm{S}}} = \begin{pmatrix} 0 & C \\ C & A \end{pmatrix}, \quad A_{H_{\mathrm{X}}} = \begin{pmatrix} A & Q \\ Q & A \end{pmatrix}.$$

These matrices follow by comparing the treatment requirements on overlaps. Treatment complementation preserves the experiment; reversing a contrast only negates its target and estimator. These operations identify the two Direct contrasts and the two spillover contrasts.

For spillover, $H_{\mathrm{G}} \subseteq H_{\mathrm{S}}$ gives the Global lower bound. For the upper bound, put

$$h_i = \max_{j \in N_i^+} (d_j + 1).$$

Step 4 of the proof of Theorem 1 gives

$$|z^\top C z| \lesssim_k z^\top D_h z, \qquad |z^\top A z| \le z^\top D_h z, \qquad \|h\|_t \le \|v^{2s}\|_t.$$

Hence $A_{H_{\mathrm{S}}}$ has a diagonal majorant proportional to $D_{h \oplus h}$. Lemma D.1, with the singleton construction when the raw functional exceeds one, proves the unbiased Global-order upper bound. Since $Q \le C$, the same argument gives the upper bound for H_{X}.

For its lower bound, on each star place independent common signs on $\{\text{center } 10,\ \text{leaves } 01\}$ and $\{\text{center } 01,\ \text{leaves } 10\}$. Every target in one class conflicts with every target in the other. Each original arm has exactly the component-size number of outcomes of the chosen amplitude. The Global star-forest calculation therefore applies:

$$\mathcal{E} \gtrsim M^2 n^{-2s} \|(b_j^{2s})_j\|_t \gtrsim M^2 n^{-2s} \|v(F)^{2s}\|_t.$$

Both A and Q increase under physical-edge addition, so the same star-forest decomposition transfers the bound to G. Together with treatment complementation, this covers all six contrasts. □

Spillover normalization with isolated units. *Let $V_+ = \{i : d_i > 0\}$, $m = |V_+|$, $G_+ = G[V_+]$, and $b = \min(B, m)$. If $m > 0$, set $\sigma = M(m/n)^s$. The spillover risk on the original population is $\sigma^2 R(b, H_{\mathrm{S}}(G_+), \delta/\sigma; q, 1)$. It is zero if $m = 0$. The same identity holds without a bias restriction.*

Proof. Isolated units have identical spillover exposures and zero contribution. On V_+, rescale $x = M_+ \widetilde{x}$, where $M_+ = M(n/m)^{1/q}$. Then

$$\|x_a\|_q \le n^{1/q} M = m^{1/q} M_+, \qquad \tau(x) = \frac{M_+}{n} \sum_{i \in V_+, a} \widetilde{x}_{i,a} = \sigma \widetilde{\tau}(\widetilde{x}).$$

Proposition 1 implements the active-population procedure with additional measurements. In the reverse direction, set isolated and nontarget outcomes to zero and condition away their measurements. Risk scales by σ^2 and bias by σ, retaining the original denominator n. □

D.7 Exact saturation at the first additional measurement

Let $\mathcal{F}$ be a finite downward-closed family of target sets containing all singletons and with maximum size at most n; put $\mathcal{F}_B = \{S \in \mathcal{F} : |S| \le B\}$. For the paper's experiment, $\mathcal{F} = \mathcal{I}(H)$. The compatibility graph $J(\mathcal{F})$ joins coordinates that can be observed together. Denote its components by C_j and their sizes by c_j. Every nonempty feasible set lies in one component.

Proposition (Exact initial saturation). *For $n \ge 2$ and every $\delta \ge 0$,*

$$R(1, \delta) = R(2, \delta) \quad \Longleftrightarrow \quad R(1, \delta) = R(n, \delta).$$

Without a bias restriction, this holds precisely when $\max_j c_j \le n$. With a bias restriction, the criterion is

$$\begin{cases} \max_j c_j \le n, & \delta \ge M, \text{ or } \delta = 0,\ q = 2, \\ \sum_{j \in A} c_j = n \text{ for some } A, & 0 < \delta < M, \text{ or } \delta = 0,\ q > 2. \end{cases}$$

Fixed arguments H, q, M are suppressed only within this proof.

Proof. **Step 1. The one-measurement value.** Let I be uniform on $[2n]$, put $\alpha = \max\{1, 2-\delta/M\}$, and report αx_I. Its bias is $(\alpha/2-1)\tau$ and

$$r(x) = \frac{\alpha^2}{2n}\|x\|_2^2 + (1-\alpha)\tau(x)^2 \le \alpha^2 M^2.$$

For a lower bound give each signed arm a common sign times M. Arm exchange, global sign reversal, and conditional averaging reduce this four-table problem to reporting z times the observed sign. Thus

$$R(1,\delta) = \min_{|z-2M|\le\delta} \max\{z^2, (z-2M)^2\} = [\max\{M, 2M-\delta\}]^2, \qquad R(1) = M^2.$$

Step 2. Decision representations used in the equality cases. For a probability measure μ on K, conditional expectation under μ is denoted by $\mathbb{E}_\mu$. The unrestricted value has the representation

$$R(B) = \max_{\mu\in\mathcal{P}(K)} \min_{S\in\mathcal{F}_B} \mathbb{E}_\mu \operatorname{Var}_\mu\{\tau(X) \mid X_S\}. \tag{D.3}$$

For $\delta > 0$ the corresponding bias-constrained representation is

$$R(B,\delta) = \max_{\mu,\nu} \left\{ \int \tau^2 \, d\mu - \int \tau \, d\nu - \delta\|\nu\|_{\mathrm{TV}} - \max_{S\in\mathcal{F}_B} \mathcal{J}(S;\mu,\nu) \right\}. \tag{D.4}$$

Here ν is a finite signed measure, μ_S is the projection of μ onto x_S, and λ_S is the projection of $\tau\mu - \nu/2$. The quadratic observation term is

$$\mathcal{J}(S;\mu,\nu) = \sup_{f\in C(\pi_S K)} \left\{ 2\int f \, d\lambda_S - \int f^2 \, d\mu_S \right\} = \int \left(\frac{d\lambda_S}{d\mu_S}\right)^2 d\mu_S,$$

with value $+\infty$ if no square-integrable density exists.

We justify both representations for all Borel estimators. On a finite parameter set K_0, set $u_{S,o} = w_S T_S(o)$ and use the closed perspective $\varphi(w,z) = z^2/w$. The finite optimization is

$$\begin{aligned} \min_{r,w,u} \quad & r \qquad \text{subject to } w \in \Delta(\mathcal{F}_B), \\ & \sum_S \varphi(w_S, u_{S,x_S} - w_S\tau(x)) \le r, \\ & \left| \sum_S u_{S,x_S} - \tau(x) \right| \le \delta, \qquad x \in K_0. \end{aligned}$$

Take $\varphi(0,0) = 0$ and $\varphi(0,z) = +\infty$ for $z \neq 0$. Positive support weights and an inverse-inclusion unbiased rule give relative strict feasibility after enlarging r. At $\delta = 0$ the bias constraints are equalities. Finite conic duality, integration of risk against μ, and completion of the square give the finite versions of (D.3)–(D.4).

For the limit, let $0 < h < M$, $a_h = 1 - h/M$, and $K_h = K \cap h\mathbb{Z}^{2n}$. Independently round each coordinate of $a_h x$ to its two neighboring grid points, preserving its mean. Then $\Xi \in K_h$, $\mathbb{E}_x\Xi = a_h x$, and the law of Ξ_S depends only on x_S. A grid rule lifts to the Borel rule $\widetilde{T}_S(x_S) = a_h^{-1}\mathbb{E}_x t_S(\Xi_S)$, for which

$$\sup_K r_{\widetilde{T}} \le a_h^{-2} \sup_{K_h} r_t, \qquad \sup_K |b_{\widetilde{T}}| \le a_h^{-1} \sup_{K_h} |b_t|.$$

For $\delta > 0$, first mix the grid rule with the unbiased singleton rule, with weights a_h and $1 - a_h$, to preserve the original bias allowance on lifting. If R_h denotes the finite value, this gives

$$\begin{aligned} & R_h(B,\delta) \le R(B,\delta) \le a_h^{-1} R_h(B,\delta) + 4M^2(1-a_h)a_h^{-2}, \\ & R_h(B,0) \le R(B,0) \le a_h^{-2} R_h(B,0), \qquad R_h(B) \le R(B) \le a_h^{-2} R_h(B). \end{aligned}$$

Thus finite values converge to the original values. Finite dual measures remain valid on K, proving equality of the displayed primal and dual values.

For $\delta > 0$, the dual objective is at most $4M^2 - \delta\|\nu\|_{\mathrm{TV}}$, by evaluation at the unbiased singleton rule. A maximizing sequence therefore has bounded total variation. Weak-star compactness and lower semicontinuity of $\mathcal{J}$ give an attaining pair in (D.4). For (D.3), each conditional variance is the infimum of continuous quadratic-loss functionals of μ; upper semicontinuity on compact $\mathcal{P}(K)$ gives an attaining prior.

Step 3. Unrestricted saturation and the component-size criterion. Suppose $R(2) = M^2$. Let μ attain (D.3), and put $P_S f = \mathbb{E}_\mu[f(X) \mid X_S]$ and $V_S = \|\tau - P_S\tau\|^2_{L^2(\mu)}$. Every V_S for $S \in \mathcal{F}_2$ is at least M^2. Orthogonal projection gives

$$\mathbb{E}_\mu(X_i - \tau)^2 = V_i + \|X_i - P_i\tau\|^2_{L^2(\mu)},$$
$$M^2 \le \frac{1}{2n}\sum_i V_i \le \frac{1}{2n}\sum_i \mathbb{E}_\mu(X_i - \tau)^2 = \frac{\mathbb{E}_\mu\|X\|_2^2}{2n} \le M^2.$$

All inequalities are equalities: $V_i = M^2$ and $P_i\tau = X_i$. Since $V_i = \mathbb{E}\tau^2 - \mathbb{E}X_i^2$, all coordinate second moments are equal. Their sum is $2nM^2$, so $\mathbb{E}X_i^2 = M^2$ and $\mathbb{E}\tau^2 = 2M^2$. For an available pair,

$$V_i - V_{ij} = \|P_{ij}\tau - P_i\tau\|^2_{L^2(\mu)} = 0, \qquad V_j - V_{ij} = 0.$$

Hence $X_i = X_j$ on every compatibility edge. The prior is constant on each component. Every nonempty feasible set then reveals the same sigma-field as one coordinate, so $V_S = M^2$ at all budgets. Thus $R(n) = M^2$. The converse follows from budget monotonicity.

For the criterion, let $X_i = Z_j$ on C_j, put $a_j = c_j/n$, and let $U = \tau(X) = \sum_j a_j Z_j$. Since $\sum_j a_j = 2$, the equality conditions imply

$$\mathbb{E}(U \mid Z_j) = Z_j, \qquad \mathbb{E}Z_j^2 = M^2, \qquad \mathbb{E}U^2 = 2M^2.$$

The variables $Y_j = Z_j - U/2$ obey $\sum_j a_j Y_j = 0$ and $\|Y_j\|_{L^2} = M/\sqrt{2}$. The triangle inequality gives $a_j \le \sum_{\ell \ne j} a_\ell = 2 - a_j$, so $c_j \le n$.

Conversely, suppose $a_j \le 1$. Merge two groups whenever their combined weight is at most one. The terminal weights are either $(1, 1)$ or $b_1, b_2, b_3 \in (0, 1)$ with sum two. In the first case use independent fair group signs. In the second put $A = \sum_{h=1}^3 (1 - b_h)^{-1}$, $p_h = \{2(A-1)(1-b_h)\}^{-1}$, and $p_0 = (A-2)/\{2(A-1)\}$. Choose the all-equal sign vector with probability p_0, or reverse only sign h with probability p_h, and multiply by an independent fair sign. Then

$$\mathbb{E}\left[\varepsilon_h \sum_\ell b_\ell \varepsilon_\ell\right] = 1.$$

Assign M times the resulting group sign to each target in the group. These tables belong to K. Every feasible observation reveals only one group sign, its conditional target mean is M times that sign, and its average posterior variance is M^2. Thus $R(n) = M^2$.

For $\delta \ge M$, the constrained and unrestricted values coincide: a rule with worst risk below M^2 has worst absolute bias below M by Jensen; when the risk equals M^2, the singleton rule attains it with bias at most M. Approximate minimizers suffice when the infimum is not attained.

Step 4. Positive bias allowances smaller than M. Scale to $M = 1$, put $\alpha = 2 - \delta \in (1, 2)$, and suppose $R(2, \delta) = \alpha^2$. The singleton rule αx_I is optimal. Let (μ, ν) attain (D.4). Equality in the risk and bias terms of its primal–dual comparison gives

$$\operatorname{supp}\mu \subseteq \{x : \tau(x) = 0,\ \|x\|_2^2 = 2n\}, \qquad \operatorname{supp}\nu \subseteq \{\mathbf{1}, -\mathbf{1}\},$$

with nonpositive mass at $\mathbf{1}$ and nonnegative mass at $-\mathbf{1}$. Indeed, its risk is $\alpha^2\|x\|_2^2/(2n) + (1 - \alpha)\tau(x)^2$, and its bias is $-\delta\tau(x)/2$; equality $|\tau| = 2$ occurs only at $\pm\mathbf{1}$.

Stationarity of every active singleton output yields $\nu_i = -2\alpha x_i \mu_i$. Thus μ_i is supported on $\{-1, 0, 1\}$; $\|X\|_2^2 = 2n$ excludes zero. The coordinate means agree and sum to zero, so all marginals are fair, with $\nu(\{\mathbf{1}\}) = -\alpha$ and $\nu(\{-\mathbf{1}\}) = \alpha$. For a feasible pair set $u = \mu(X_i = X_j = 1) = \mu(X_i = X_j = -1) \le 1/2$. Then

$$\mathcal{J}(\{i\}; \mu, \nu) = \alpha^2, \qquad \mathcal{J}(\{i, j\}; \mu, \nu) = \frac{\alpha^2}{2u} \le \alpha^2 \quad \Longrightarrow \quad u = \tfrac{1}{2}.$$

At $u = 0$ the pair term is infinite. It follows that $X_i = X_j$ on every compatibility edge. A component-constant sign table with $\tau = 0$ partitions the components into two groups of exactly n coordinates.

Conversely, given such a partition, assign independent common signs to the two groups. Every feasible set reveals at most one sign. The four-table lower bound in Step 1 applies at every budget and matches the singleton upper bound. Restoring M proves this case.

Step 5. Exact unbiasedness. Choose a symmetric stochastic matrix Π supported on the diagonal and the edges of the compatibility graph, with positive diagonal entries and positive entries on every compatibility edge. For example, if L_J is the Laplacian of $J(\mathcal{F})$ and Δ_J its maximum degree, take $\Pi = I - \xi L_J$ with $0 < \xi < (1 + \Delta_J)^{-1}$. Draw I uniformly on $N = 2n$ coordinates and I' from row I of Π, and report $X_I + X_{I'}$. This feasible pair rule is unbiased, and

$$r_0(x) = \frac{2}{N}\{\|x\|_2^2 + x^\top \Pi x\} - \tau(x)^2 \le 4M^2. \tag{D.5}$$

Equality requires maximal arm second moments, constancy on each component, and $\tau = 0$. For $q > 2$, equality in the moment comparison further requires $|x_i| = M$ for every coordinate. Hence equality requires a union of components of size n. Without such a union, compactness gives a strict risk bound below $4M^2$; with it, the four-table lower bound gives $R(n, 0) = 4M^2$.

For $q = 2$, scale to $M = 1$. Consider an oracle that selects component j with probability $p_j > 0$ and observes it completely. Let $\tau_j(x) = n^{-1} \sum_{i \in C_j} x_i$. Replace a rule by $T_j^{\mathrm{odd}}(y) = \{T_j(y) - T_j(-y)\}/2$; symmetry of K preserves unbiasedness and cannot increase worst risk. On tables supported in component j, outputs on other components vanish, so unbiasedness forces $T_j(x) = \tau_j(x)/p_j$. The oracle value is

$$U_{\mathrm{or}} = \min_{p \in \Delta_r,\, p > 0} f(p), \qquad f(p) = \max_{x \in K} \left\{ \sum_j \frac{\tau_j(x)^2}{p_j} - \tau(x)^2 \right\}.$$

The minimum is attained in the positive simplex, since f diverges at its boundary. The oracle can simulate any original feasible procedure. Put $a_j = c_j/n$ and $p_j^0 = a_j/2$. Componentwise Cauchy–Schwarz gives $f(p^0) \le 4$. Equality holds precisely on component-constant vectors $x_i = z_j$ satisfying

$$a^\top z = 0, \qquad \sum_j u_j z_j^2 = \sum_j v_j z_j^2 = 1, \quad u_j = |C_j \cap \text{arm } 0|/n, \quad v_j = |C_j \cap \text{arm } 1|/n. \tag{D.6}$$

Here $u + v = a$. We claim

$$U_{\mathrm{or}} = 4 \iff \exists\, Q \succeq 0 :\ Q_{jj} = 1,\ Qa = 0 \iff \max_j a_j \le 1.$$

If $U_{\mathrm{or}} = 4$, the interior point p^0 minimizes the convex function f. Active gradients equal $-4z_j^2$, so their convex combination is constant across coordinates. Since $\sum_j a_j z_j^2 = 2$, the resulting finite distribution has $\mathbb{E} z_j^2 = 1$; set $Q = \mathbb{E} z z^\top$. Conversely, given Q, the compact convex set

$$\{X \succeq 0 : Xa = 0,\ \mathrm{tr}(D_u X) = \mathrm{tr}(D_v X) = 1\}$$

contains it. Every extreme point has rank one. To see this, write a rank-r point as $LL^\top$, with $L^\top a = 0$. If $r \ge 2$, the space of symmetric $r \times r$ matrices has dimension at least three. There is a nonzero symmetric D satisfying the two equations $\mathrm{tr}(L^\top D_u L D) = \mathrm{tr}(L^\top D_v L D) = 0$. For sufficiently small $\varepsilon > 0$, the distinct matrices $L(I \pm \varepsilon D)L^\top$ are feasible and have midpoint $LL^\top$. Thus a rank-$r \ge 2$ point is not extreme. It follows that Q is a convex combination of $zz^\top$ satisfying (D.6), and for every p,

$$\mathbb{E} \sum_j \frac{a_j^2 z_j^2}{p_j} = \sum_j \frac{a_j^2}{p_j} \ge (\sum_j a_j)^2 = 4.$$

This proves the first equivalence. A Gram representation $Q_{ij} = e_i^\top e_j$ turns $Qa = 0$ into $\sum_j a_j e_j = 0$ with unit e_j. Such directions exist exactly when the largest length a_j does not exceed the sum of the others, by the closed-polygon criterion. Since $\sum_j a_j = 2$, this is $\max_j a_j \le 1$.

If this condition holds, the oracle lower bound proves saturation at every budget. Otherwise choose $p^* > 0$ with $f(p^*) < 4$. As an alternative pair rule select component j with probability p_j^*, choose I uniformly within it and I' from the corresponding stochastic kernel, and report $c_j(X_I + X_{I'})/(2np_j^*)$. This rule is unbiased and agrees with the oracle on the equality set (D.6) of the base rule (D.5). Its risk is strictly below four on a neighborhood of that set; the base risk is strictly below four on the compact complement. A sufficiently small mixture therefore has risk strictly below four everywhere on K. Thus failure of the component condition already permits improvement at budget two, completing all cases. □

D.8 Bias, measurement, and network complexity

For the class $\mathcal{G}(n,k) = \{G : |V(G)| = n,\ \mathrm{degen}(G) \le k\}$, write

$$W(B,n,k,\delta;q,M) = \sup_{G\in\mathcal{G}(n,k)} R(B, H_{\mathrm{D}}(G), \delta; q, M).$$

The design and estimator may depend on the observed graph.

Theorem (Bias, measurement, and network complexity). *Let $n \ge 2$, $1 \le k \le n-1$, $1 \le B \le n$, $q \in [2,\infty]$, $M > 0$, and $\delta \ge 0$. With $\rho = k/n$, $\eta = \delta/M$, and $\beta = 2s/(1+s)$,*

$$W(B,n,k,\delta;q,M) \asymp M^2 \left\{ \frac{1}{B} + \rho^{\beta} + \frac{\rho}{(\rho + \eta^{1/s})^{1-s}} \right\}.$$

The comparison constants are universal, including with respect to k and q.

Proof. **Step 1. Degree-dependent shrinkage.** For $z > 0$, let $\gamma_i(z) = \min\{1, \sqrt{z/d_i}\}$ when $d_i > 0$, and $\gamma_i(z) = 1$ otherwise. Put $\Gamma = D_\gamma$ and $L = \sum_i d_i \le 2kn$. Only degrees above z are shrunk, so

$$\sum_i (1-\gamma_i)^{1/s} \le \frac{L}{z}, \qquad \sup_{x\in K} |\mathbb{E}\widehat{\tau} - \tau(x)| \le 2M(2k/z)^s. \tag{D.7}$$

The second inequality is the exact bias formula of Lemma D.1. For an outdegree-k orientation P,

$$\|\Gamma P \Gamma x\|_2^2 \le k \sum_j \gamma_j^2 \left(\sum_{i:i\to j} \gamma_i^2 \right) x_j^2 \le kz\|x\|_2^2,$$
$$\|\Gamma A_G \Gamma\| \le 2\sqrt{kz}, \qquad \mathbf{1}^\top \Gamma A_G \Gamma \mathbf{1} \le L.$$

For a symmetric entrywise nonnegative Q, $2 < q < \infty$, and $u \ge 0$ with $\|u\|_q \le 1$, weighted Hölder gives

$$u^\top Q u \le (\mathbf{1}^\top Q \mathbf{1})^{1-2/q} \left(\sum_{i,j} Q_{ij} u_i^{q/2} u_j^{q/2} \right)^{2/q}$$
$$\le (\mathbf{1}^\top Q \mathbf{1})^{1-2/q} \|Q\|^{2/q}.$$

Taking the supremum over u proves the asserted bound for $\Lambda(Q,q)$. At $q = 2$ this is the spectral bound; at $q = \infty$ it is the entry sum. Using also $\Gamma A_G \Gamma \le A_G$ entrywise and $\|A_G\| \le \sqrt{L}$, we obtain

$$\Lambda(\Gamma A_G \Gamma, q) \lesssim \min\{k^s z^{1-s} n^{2s-1}, (kn)^s\}. \tag{D.8}$$

Step 2. An upper bound at the prescribed bias allowance. For $s < 1$ and $\eta > 0$, set

$$z_0 = 2k\rho^{-1/(1+s)}, \qquad z_\delta = 2k(2/\eta)^{1/s}, \qquad z = \min\{\Delta, \max(z_0, z_\delta)\}.$$

The first scale balances variance and squared bias; the second enforces $\sup |b_P| \le \delta$. When the cap Δ binds there is no shrinkage. Otherwise (D.7) gives

$$\sup_{x\in K} |b_P(x)|^2 \le \min\{\delta^2, 4M^2\rho^\beta\}.$$

The same bound is automatic when the cap binds. Since $\|\gamma\|_{1/s} \le n^s$ and $\|\gamma^2\|_t \le n^{2s-1}$, Lemma D.1 and (D.8) yield

$$\begin{aligned} M^{-2} \sup_{x \in K} r_P(x) &\lesssim B^{-1} + \rho^\beta + \min\{\rho(z/k)^{1-s}, \rho^s\} \\ &\lesssim B^{-1} + \rho^\beta + \min\{\rho\eta^{-(1-s)/s}, \rho^s\} \\ &\asymp B^{-1} + \rho^\beta + \frac{\rho}{(\rho + \eta^{1/s})^{1-s}}. \end{aligned}$$

These comparisons are uniform over $s \in [1/2, 1)$. At $\eta = 0$, take $\gamma = \mathbf{1}$; the unshrunk energy bound gives $M^2(B^{-1} + \rho^s)$. At $s = 1$, the same choice gives $M^2(B^{-1} + \rho)$. An edgeless graph is handled by $\gamma = \mathbf{1}$ directly. This proves the upper bound in every endpoint case.

Step 3. Matching lower bounds. The measurement term M^2/B holds for every graph. First suppose $\rho < 1/8$, and put $r_* = \rho^{-s/(1+s)}$. For an integer $1 \le r \le r_*/2$, take r disjoint copies of $K_{k,\ell}$, where $\ell = \lfloor n/r \rfloor - k \asymp n/r \ge k$, and add isolates. This is a k-degenerate graph. Call the k-side the centers. For a fixed procedure with worst risk $\mathcal{E}$, let p_j be the probability of observing either target of any center in component j, and set $P = \sum_j p_j$.

Give each component's treated leaves one independent common sign with amplitude $M\{n/(r\ell)\}^{1/q}$, and set all other targets to zero. Observation of a center forces those leaves untreated, hiding their sign. Their target contribution per component has magnitude of order M/r; hence

$$\mathcal{E} \gtrsim M^2 P/r^2. \tag{D.9}$$

For a second family, give both signed targets of all rk centers one common unknown sign of amplitude $M\{n/(rk)\}^{1/q}$. Its target is $\pm u$, where $u = 2M(r\rho)^s$; the probability of revealing the sign is at most P. If $P \le 1/2$, posterior variance gives $\mathcal{E} \ge u^2/2$ without any bias restriction.

When $\delta \le u/2$, couple the two outputs T_+, T_- using the same design randomization. They agree off the revealing event and $\mathbb{E}(T_+ - T_-) \ge 2(u - \delta)$. Thus, for $0 < P \le 1/2$,

$$\begin{aligned} 4\mathcal{E} &\ge \mathrm{Var}(T_+ - T_-) \ge \{\mathbb{E}(T_+ - T_-)\}^2 (P^{-1} - 1), \\ \mathcal{E} &\ge (u - \delta)^2 (P^{-1} - 1) \gtrsim \frac{M^2 (r\rho)^{2s}}{P}. \end{aligned}$$

At $P = 0$, bias less than u is impossible. Both families satisfy the original armwise moment constraints.

Choose $r = \lfloor r_*/2 \rfloor \asymp r_*$. Use the unrestricted center bound when $P \le 1/2$ and (D.9) otherwise. Balancing the two scales gives

$$W \gtrsim M^2 \min\{(r_*\rho)^{2s}, r_*^{-2}\} = M^2 \rho^{2s/(1+s)}.$$

Next put $h = 1 + \eta^{1/s}/\rho$. If $h \le r_*/2$, choose $r = \max\{1, \lceil \eta^{1/s}/\rho \rceil\} \asymp h$, so $\delta \le u/2$. The two cases give

$$\begin{aligned} \mathcal{E}/M^2 &\gtrsim \left\{ \frac{P}{r^2} \frac{(r\rho)^{2s}}{P} \right\}^{1/2} = \rho^s r^{s-1}, && 0 < P \le \tfrac{1}{2}, \\ \mathcal{E}/M^2 &\gtrsim r^{-2} \ge \rho^s r^{s-1}, && P > \tfrac{1}{2}. \end{aligned}$$

The last inequality uses $r \le r_*$. Consequently

$$W \gtrsim M^2 \rho^s h^{s-1} = \frac{M^2 \rho}{(\rho + \eta^{1/s})^{1-s}}.$$

If $h > r_*/2$, this final term is already bounded by a universal multiple of $M^2\rho^\beta$.

For $\rho \ge 1/8$, a clique of size $k + 1$ plus isolates, with independent common arm signs, gives $W \gtrsim M^2 \rho^{2s} \gtrsim M^2$. All three claimed terms are at most a universal multiple of M^2. Combining the measurement, unrestricted-network, and bias-constrained bounds proves the theorem. □

D.9 Endpoint rates and uniform consistency

Proof. At $\eta = 0$, the additional term is ρ^s, which dominates ρ^β. Without a bias restriction, use z_0 in the upper construction and the unrestricted lower bound from Step 3. Finally, $1/2 \le s \le 1$ and $2/3 \le \beta \le 1$ imply

$$M^2(B^{-1} + \rho) \lesssim W \lesssim M^2(B^{-1} + \rho^{2/3} + \sqrt{\rho}).$$

The lower and upper inequalities prove necessity and sufficiency. □