Experimental Design for Target-Population Treatment Effect Estimation under Heterogeneous Privacy Constraints

Anonymous authors
Paper under double-blind review

Abstract

Estimating a target-population average treatment effect (ATE) in multisite randomized trials requires allocating limited enrollment across sites that differ in target-population relevance, recruitment costs, and privacy constraints. Existing target-population trial designs do not directly account for site-specific privacy constraints, while private multisite ATE estimation generally takes site enrollment as given. In this paper, we study experimental design for multisite randomized trials under site-specific participant-level privacy. We characterize each site's privacy-constrained information contribution and derive a target-specific criterion that balances estimation variance and worst-case squared bias. The criterion characterizes the minimax mean squared error within constants independent of the budget, site count, and allocation. A site's information contribution grows quadratically with enrollment in the privacy-limited branch and linearly after a site-specific threshold, so heterogeneous privacy can change the allocation favored by classical design. The resulting design problem is strongly NP-hard even in the high-privacy regime. We develop sparse conic constructions with dimension-dependent approximation guarantees under explicit solver-accuracy conditions, budget-feasible integer recovery, and private estimator weights. Experiments show when accounting for heterogeneous privacy reduces target-population ATE error under a fixed enrollment budget and when the added privacy variance outweighs the reduction in squared bias.

1 Introduction

Randomized trials provide causal evidence for clinical and policy decisions, but the population in which a treatment will be used can differ from the participants enrolled in a trial. Estimating an average treatment effect (ATE) for a specified target population therefore depends on both the evidence collected and assumptions connecting trial participants to that population (Dahabreh et al., 2019). Multisite trials provide access to different patient populations under a shared study protocol. With a limited enrollment budget, an investigator must decide how many participants to recruit at each site before trial outcomes are observed. This allocation is an experimental design decision that determines the precision of subsequent treatment effect estimation (Raudenbush & Liu, 2000).

Classical optimal design allocates observations to estimate a specified parameter or linear functional efficiently (Elfving, 1952; Sagnol, 2011). Target-population causal inference connects experimental samples to populations of interest (Dahabreh et al., 2019), motivating experiment design, sample selection, and allocation that use information about the target population (Phan et al., 2021; Shi & Lin, 2022; Dang et al., 2026). These approaches make the target population part of the design problem. Their statistical value, however, depends on how much information the recruited participants can contribute to the final estimator.

Privacy constraints can substantially change that information. Private multisite estimation combines protected local evidence (Koga et al., 2024), while privacy-aware experimental and survey design accounts for release noise when allocating samples or privacy budgets (Chen et al., 2024; 2025). Information contraction results characterize how much information can remain in a privacy-constrained public transcript for a fixed experiment (Cai & Li, 2026). Consider an **example** with a multisite hospital trial: institutions follow a common treatment protocol but impose different participant-level

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.

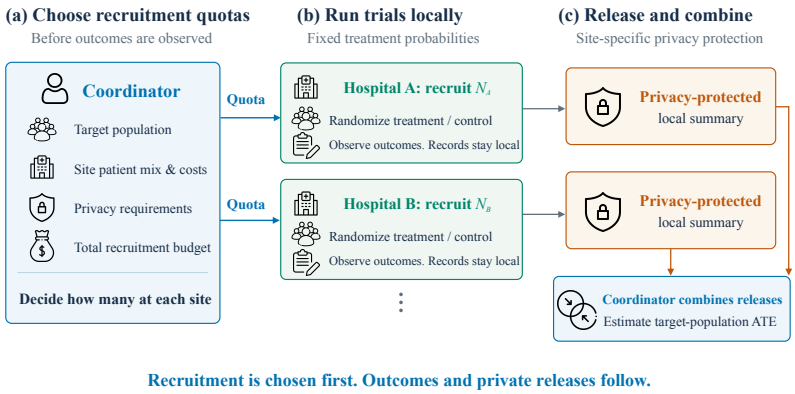


Figure 1: **Experimental design for a multisite randomized trial.** Before observing outcomes, the investigator allocates the enrollment budget across sites using the target population, site covariate distributions, costs, and privacy constraints. Treatment is randomized within each site. Local private summaries are combined to estimate the target-population ATE. The illustrated single-round release is an admissible protocol within the interactive class analyzed in the paper.

privacy requirements on their released summaries, and the investigator must allocate a common enrollment budget before outcomes are observed, while the information available from each site depends on its covariate distribution, enrollment, and privacy constraint. However, existing approaches do not directly determine how enrollment should be allocated for target-population ATE estimation when sites differ jointly in covariate distributions, recruitment costs, and privacy constraints.

The above problem introduces two statistical difficulties. First, under participant-level privacy, the information bound for a site is no longer linear in its sample size. Increasing enrollment both adds observations and reduces the sensitivity of the released summary. As a result, the privacy-constrained information bound grows quadratically with enrollment while privacy is limiting and linearly after a site-specific threshold. Second, a finite enrollment budget may not provide enough information to estimate every component relevant to the target-population ATE precisely. Allocating more enrollment to cover these components reduces estimation variance, while leaving part of the target direction unestimated incurs worst-case squared bias. The design must therefore determine how site sample sizes and privacy constraints jointly affect this variance-bias tradeoff. This leads to our central question:

> *How should a fixed enrollment budget be allocated across sites to minimize target-population ATE estimation error when sites differ in covariate distributions, recruitment costs, and participant-level privacy constraints?*

To answer this question, we study multisite randomized trials with a common linear treatment effect, known site covariate second moments and target mean, and site-specific participant-level zero-concentrated differential privacy (zCDP) constraints (Bun & Steinke, 2016). Enrollment is chosen before outcomes are observed, with treatment randomized within each site, as illustrated in Figure 1. We first characterize the privacy-constrained information available for the target ATE and derive a target-specific criterion that balances estimation variance with worst-case squared bias. We show that this criterion characterizes the minimax mean squared error over admissible private protocols within constants independent of the budget, site count, and allocation. Privacy makes the criterion nonlinear in site sample sizes, and its global optimization is strongly NP-hard even when every site lies in the high-privacy regime. We then develop sparse conic constructions with dimension-dependent approximation guarantees under explicit solver-accuracy conditions and budget-feasible integer recovery, together with private releases and estimator weights for the resulting allocations.

Our main **contributions** are as follows: (1) We formulate target-population experimental design under site-specific participant-level privacy and show that heterogeneous privacy can change the enrollment allocation relative to classical design. (2) We characterize the minimax MSE for the target-population ATE within constants and establish the statistical and computational properties of the resulting design problem, including strong NP-hardness and dimension-dependent guarantees for sparse conic constructions. (3) Extensive experiments demonstrate when accounting for hetero-

geneous privacy can improve target-population ATE estimation under a fixed budget and relate these gains to the tradeoff among squared bias, sampling variance, and privacy variance.

2 Related work

Experimental design for target-population effects. Classical c-optimal design allocates information to estimate a specified linear functional efficiently (Elfving, 1952), with finite multiresponse designs admitting second-order cone programming (SOCP) formulations (Sagnol, 2011). Multisite trial design similarly allocates a fixed budget across sites and participants (Raudenbush & Liu, 2000). For target-population causal effects, transporting trial evidence requires assumptions connecting the experimental and target populations (Dahabreh et al., 2019), and recent designs use target covariates for transportable experiment design, trial sample selection, and target-weighted allocation (Phan et al., 2021; Shi & Lin, 2022; Dang et al., 2026). Related minimax linear estimation methods balance variance against worst-case bias over a parameter class (Hirshberg & Wager, 2021), providing a risk-based perspective when limited information does not support precise estimation of every target-relevant direction.

Privacy-aware design and multisite estimation. Privacy introduces an additional constraint on how much information recruited participants can contribute. Federated experiment design studies private ATE estimation under distributed privacy and secure aggregation (Chen et al., 2024), private survey design optimizes stratified sample sizes for specified release mechanisms (Chen et al., 2025), and private multisite causal estimation combines protected local estimates (Koga et al., 2024). Information-contraction results characterize the information retained in privacy-constrained public transcripts (Cai & Li, 2026), while participant-level zCDP (Bun & Steinke, 2016) differs from observation-wise local privacy (Duchi et al., 2018) in its privacy model and information constraints. On the computational side, SOCP formulations and sparse representations provide standard tools for structured design problems (Alizadeh & Goldfarb, 2003), whereas related constrained and regularized L_p optimization problems are known to be strongly NP-hard (Ge et al., 2011; Chen et al., 2014).

3 Problem setup

We consider a multisite randomized trial for estimating an ATE in a prespecified target population. There are S candidate sites. Before outcomes are observed, the investigator chooses their sample sizes $N = (N_1, \dots, N_S)$ under an enrollment budget $\sum_s c_s N_s \le B$, where $B > 0$ and $c_s > 0$ is the cost per participant at site s. Treatment is randomized with probability $1/2$ within every site. The design controls enrollment across sites, and each site's complete public transcript must satisfy its site-specific participant-level privacy constraint.

Causal model and target population. Let J denote site membership. At site $s \in [S]$, covariates $X \in \mathbb{R}^d$ follow Q_s, and $A \mid X, J = s \sim \text{Bernoulli}(1/2)$ satisfies $A \perp \{Y(0), Y(1)\} \mid X, J = s$. For $a \in \{0, 1\}$,

$$Y(a) \mid X = x, J = s \sim \text{Bernoulli}\left(\frac{1 + (2a-1)x^\top \theta}{2}\right). \tag{1}$$

We assume consistency, $Y = Y(A)$, and no interference between participants. The conditional treatment effect is $x^\top \theta$, shared across sites and the target population. Writing Q_T for the target covariate distribution and $w = \mathbb{E}_{Q_T} X$ for its known mean gives the target ATE $\tau_T(\theta) = w^\top \theta$.

Let $\mathcal{X} = \text{supp}(Q_T) \cup_s \text{supp}(Q_s)$ and fix $\eta \in (0, 1)$. Define

$$\Theta_0 = \left\{\theta \in \mathbb{R}^d : \sup_{x \in \mathcal{X}} |x^\top \theta| \le 1 - \eta\right\}, \qquad H_s = \mathbb{E}_{Q_s}[XX^\top]. \tag{2}$$

Thus both response probabilities lie in $[\eta/2, 1 - \eta/2]$. Let $\mathcal{U} = \text{span}\{\text{range}(H_s) : s \in [S]\}$ and assume $w \in \mathcal{U}$. For $r > 0$, we evaluate risk on $\Theta(r) = \{\theta \in \mathcal{U} : \|\theta\|_2 \le r\} \subseteq \Theta_0$. The shared conditional treatment effect makes the same θ govern source and target effects, while $w \in \mathcal{U}$ ensures that the target linear functional is covered by the source-site covariate moments. The radius r specifies the effect magnitudes considered at design time.

At $\theta = 0$, the score $V = (2A-1)(2Y-1)X$ has covariance H_s, so N_s participants supply raw Fisher information $N_s H_s$. Uniformly over $\Theta(r)$,

$$H_s \preceq I_s(\theta) = \mathbb{E}_{Q_s}\left[\frac{XX^\top}{1-(X^\top\theta)^2}\right] \preceq [\eta(2-\eta)]^{-1} H_s. \tag{3}$$

Appendix A derives these identities. Appendix E extends the risk comparison to unknown site-specific baseline response functions.

Privacy and public design inputs. The inputs $(H_s, c_s, \rho_s, B, w, r)$ and a leverage bound L are fixed and public before enrollment. At site s, the complete public transcript must satisfy ρ_s-zCDP under replacement of one (X, A, Y) record, conditional on every fixed dataset at the other sites (Bun & Steinke, 2016). Communication may be adaptive, using independent private site randomness and public coordination. The model excludes hidden shared randomness, secure aggregation, and trusted-server private state. Appendix A specifies the protocol and regularity conditions. Privacy accounting covers the recruited records conditional on the public inputs. Acquiring those inputs from protected data requires additional accounting.

Feasible designs and estimability. Write N for integer sample sizes and n for their continuous relaxation. Their feasible sets are

$$\mathcal{D}_B^{\mathrm{int}} = \{N \in \mathbb{N}_0^S : \sum_s c_s N_s \le B\}, \qquad \mathcal{D}_B^{\mathrm{cont}} = \{n \in \mathbb{R}_+^S : \sum_s c_s n_s \le B\}. \tag{4}$$

Here $\mathbb{N}_0$ denotes the nonnegative integers. For $M \succeq 0$, the extended c-optimal form is $q_M(x) = x^\top M^\dagger x$ when $x \in \mathrm{range}(M)$ and $+\infty$ otherwise. Equivalently,

$$q_M(x) = \sup_v \{2x^\top v - v^\top M v\}. \tag{5}$$

It records whether the information in M covers the target direction.

4 Statistical characterization

4.1 Private information

To compare sample allocations, we bound each site's information using its sample size and privacy constraint. Define

$$\alpha_s(n_s) = \min\{n_s, \kappa_s n_s^2\}, \qquad \kappa_s = e^{2\rho_s} - 1, \qquad \overline{M}_\rho(n) = \sum_s \alpha_s(n_s) H_s. \tag{6}$$

The coefficient α_s has a linear sampling branch and a quadratic privacy-limited branch. Their crossover is $\kappa_s n_s = 1$. Thus a site's covariates, privacy constraint, and assigned sample size jointly determine its contribution to the information bound.

At $\theta = 0$, Theorem 3.4 of Cai & Li (2026) bounds the information in the complete public transcript using the raw site information $J_s = N_s H_s$. Rescaling its whitened contribution gives the following contraction class, extended algebraically to continuous sample sizes:

$$0 \preceq R_s \preceq n_s P_{H_s}, \qquad \mathrm{Tr}(R_s) \le \kappa_s n_s^2. \tag{7}$$

Here P_{H_s} projects onto $\mathrm{range}(H_s)$, and $P_{\mathcal{U}}$ projects onto $\mathcal{U}$. Both constraints imply $R_s \preceq \alpha_s(n_s) P_{H_s}$.

Theorem 1 (Scalar-target information envelope)**.** *For fixed $n \in \mathcal{D}_B^{\mathrm{cont}}$, let $\phi_\rho(n; w)$ be the infimum of $q_{\sum_s H_s^{1/2} R_s H_s^{1/2}}(w)$ over* (7)*. Then*

$$\phi_\rho(n; w) = q_{\overline{M}_\rho(n)}(w) = \min_{\substack{z_s \in \mathrm{range}(H_s) \\ \sum_s H_s^{1/2} z_s = w}} \sum_s \frac{\|z_s\|_2^2}{\alpha_s(n_s)}. \tag{8}$$

A zero denominator has cost zero for a zero block and $+\infty$ otherwise.

Theorem 1 converts the matrix contraction into an exact scalar-target criterion over its feasible class. The vector $H_s^{1/2} z_s$ is the portion of the target recovered at site s, and $\|z_s\|^2/\alpha_s(n_s)$ is its information cost. Rank-one contributions aligned with these blocks attain the variational equality. The statistical risk of an admissible private protocol is established next. Appendix B.1 gives the minimum-energy proof.

4.2 Minimax target-ATE risk

With a bounded parameter class, recovering every target component can incur more variance than its contribution warrants. For $M_N = \overline{M}_\rho(N)$, define

$$\psi_r(N; w) = w^\top [M_N + r^{-2} P_{\mathcal{U}}]^{-1} w, \tag{9}$$

$$\Psi^{\mathrm{int}}_{B,r}(w) = \inf_{N \in \mathcal{D}^{\mathrm{int}}_B} \psi_r(N; w), \qquad \Psi^{\mathrm{cont}}_{B,r}(w) = \inf_{n \in \mathcal{D}^{\mathrm{cont}}_B} \psi_r(n; w). \tag{10}$$

The inverse acts on $\mathcal{U}$, and the same formula defines $\psi_r(n; w)$ for continuous n. Adding a residual target component to the minimum-energy identity yields

$$\psi_r(N; w) = \min_{\substack{z_s \in \mathrm{range}(H_s),\ b \in \mathcal{U} \\ \sum_s H_s^{1/2} z_s + b = w}} \left\{ \sum_s \frac{\|z_s\|_2^2}{\alpha_s(N_s)} + r^2 \|b\|_2^2 \right\}. \tag{11}$$

The vector b is the unestimated part of the target. Its worst-case squared bias is $\sup_{\theta \in \Theta(r)} (b^\top \theta)^2 = r^2 \|b\|^2$. Equation (11) therefore balances the cost of estimating $w - b$ with the bias incurred by omitting b. Increasing r makes omission more costly and changes the sample allocation needed to estimate the target.

For a fixed integer design N, let $\mathcal{P}_N$ denote the admissible private protocols and G_Π the complete transcript of protocol Π. Define

$$R^\star_N = \inf_{\Pi \in \mathcal{P}_N} \inf_{\widehat{\tau}} \sup_{\theta \in \Theta(r)} \mathbb{E}_{\theta,\Pi} \left[\{\widehat{\tau}(G_\Pi) - w^\top \theta\}^2 \right], \qquad \mathfrak{R}^\star_{B,r}(w) = \inf_{N \in \mathcal{D}^{\mathrm{int}}_B} R^\star_N. \tag{12}$$

The first quantity is the best achievable risk at fixed site sample sizes. The second is the experimental design objective. For the constructive upper bound, assume $\|H_s^{\dagger/2} X\| \le L$ almost surely and $0 < \rho_s \le \bar{\rho} < \infty$.

Theorem 2 (Finite-radius minimax characterization)**.** *Under the model and transcript assumptions above, let* $C_- = \max\{[\eta(2-\eta)]^{-1}, \pi^2\}$ *and* $C_+ = 1 + 4L^2 e^{2\bar{\rho}}$*. For every feasible integer design,*

$$C_-^{-1} \psi_r(N; w) \le R^\star_N \le C_+ \psi_r(N; w). \tag{13}$$

Consequently, $\mathfrak{R}^\star_{B,r}(w) \asymp \Psi^{\mathrm{int}}_{B,r}(w)$*. With* L*,* $\bar{\rho}$*, and* η *fixed, the constants are independent of the budget, site count, allocation, and number of communication rounds.*

The lower bound combines transcript information contraction with a scalar van Trees argument. The upper bound releases a bounded target-directed scalar from each active site with independent Gaussian noise. Thus optimizing ψ_r has a statistical interpretation across admissible protocols. It also permits bias when a small sample budget cannot support precise estimation of all relevant directions. Appendix B.2 proves the result. Proposition 3 in Appendix E extends the same comparison to unknown site-specific baseline responses under equal-probability within-site randomization.

5 Design and estimation

Theorem 2 reduces experimental design to minimizing ψ_r over the budget set. We solve a conic problem for the site contributions, compress their representation, and recover feasible sample sizes. The number of active sites controls the approximation loss. Write $d_{\mathrm{eff}} = \dim \mathcal{U}$ and use supplied factors $H_s = G_s G_s^\top$. In this section, z_s denotes a coefficient in the coordinates of G_s. Appendix C.1 gives its correspondence with the symmetric-root representation.

5.1 Computing the design

Sampling benchmark. The sampling benchmark replaces $\alpha_s(n_s)$ by n_s, agreeing with the private criterion when $\kappa_s n_s \ge 1$ at every active site. Its optimum is

$$\Psi^{\mathrm{samp}}_{B,r}(w) = \min_{\sum_s G_s z_s + b = w} \left\{ B^{-1} \Big(\sum_s \sqrt{c_s} \|z_s\| \Big)^2 + r^2 \|b\|^2 \right\}, \tag{14}$$

an exact SOCP. This is a regularized c-optimal design criterion, with $r^2\|b\|^2$ controlling worst-case bias. Requiring $b = 0$ recovers the classical multiresponse formulation (Sagnol, 2011). Optimal expenditure for fixed site contributions is proportional to $\sqrt{c_s}\|z_s\|$. If the recovered allocation lies on every active site's sampling branch, it also minimizes the private criterion.

High-privacy regime. If $\kappa_s B/c_s \le 1$ for every site, the entire feasible region has $\alpha_s(n_s) = \kappa_s n_s^2$. With $u_s = c_s z_s/\sqrt{\kappa_s}$ and $\widetilde{G}_s = \sqrt{\kappa_s} G_s/c_s$,

$$\Psi^{\mathrm{hp}}_{B,r}(w) = \min_{\sum_s \widetilde{G}_s u_s + b = w} \left\{ B^{-2}\Big(\sum_s \|u_s\|^{2/3} \Big)^3 + r^2\|b\|^2 \right\}. \tag{15}$$

The block-$2/3$ term is nonconvex and favors concentrating enrollment on fewer sites. Its convex lower bound replaces $(\sum_s \|u_s\|^{2/3})^3$ by $(\sum_s \|u_s\|)^2$. We solve this SOCP and compress the representation to at most $d_{\mathrm{eff}} + 1$ active blocks. Compression preserves $w - b$, b, and $T = \sum_s \|u_s\|$. On this support, the original privacy term is at most $d_{\mathrm{eff}} + 1$ times its convex replacement. The recovered expenditure is proportional to $\|u_s\|^{2/3}$.

Mixed privacy. For general privacy constraints, different sites can lie on different information branches. The conic comparison program accounts for both sampling and privacy costs. Its conic objective uses the norm sums $L_S = \sum_s \sqrt{c_s}\|z_s\|$ and $L_P = \sum_s c_s\|z_s\|/\sqrt{\kappa_s}$. Compression preserves both sums and the represented target using at most $d_{\mathrm{eff}} + 2$ blocks. We assign half the budget according to the sampling allocation and half according to the privacy allocation, then add the site sample sizes. Equations (35) and (37) give the program and recovery rule.

Theorem 3 (Computational characterization)**.** *The sampling benchmark* (14) *has an exact SOCP formulation. Global optimization of* (15) *is strongly NP-hard for rational inputs* $(G_s, c_s, \kappa_s, B, r, w)$*, even with rank-one* H_s*, equal costs and privacy coefficients, fixed* $r = 1/4$*, and a fixed positivity margin.*

For $\varepsilon > 0$*, suppose the branch-specific conic return is feasible and its objective, recomputed from the returned block norms, is at most* $(1 + \varepsilon)$ *times the conic optimum. Let* $\Psi^{\mathrm{hp}}_{\mathrm{alg}}$ *and* $\Psi^{\mathrm{mix}}_{\mathrm{alg}}$ *denote the returned criterion values. Apply the compression above with b unchanged and recover feasible sample allocations by Equation* (34) *or Equation* (37)*. With exact compression and recovery, the returned continuous criterion values satisfy*

$$\Psi^{\mathrm{hp}}_{\mathrm{alg}} \le (d_{\mathrm{eff}} + 1)(1 + \varepsilon)\Psi^{\mathrm{hp}}_{B,r}(w), \tag{16}$$

$$\Psi^{\mathrm{mix}}_{\mathrm{alg}} \le (4d_{\mathrm{eff}} + 10)(1 + \varepsilon)\Psi^{\mathrm{cont}}_{B,r}(w). \tag{17}$$

The high-privacy branch requires $\kappa_s B/c_s \le 1$ *for all sites. The mixed branch allows arbitrary positive* κ_s*. Finite-precision returns satisfy these bounds when the support, feasibility, and combined error conditions in Appendix C.5 hold with* $\Gamma \le 1 + \varepsilon$*.*

The support size determines the dimension dependence in Theorem 3. Compression preserves the represented target and the branch-specific norm sums, so the recovery bound depends on d_{eff} even when the number of available sites is large. Appendix B.3 proves hardness. Appendix C proves the construction bounds, with numerical error conditions in Appendix C.5.

Integer sample allocation. Flooring each continuous site sample size preserves the budget but can discard small active sites. A second rule reserves the support cost $C_K = \sum_{s:n_s>0} c_s$ and rounds the scaled sample allocations upward when $C_K < B$. Comparing the two feasible candidates under ψ_r gives a hybrid rule. Appendix D bounds the loss of each rule. Algorithm 1 in Appendix C.1 summarizes design and private estimation.

5.2 Private estimation after allocation

For estimation, we use symmetric factors $H_s^{1/2}$. At fixed site sample sizes, the representation determines the private release and its weights. The criterion-based (CB) representation minimizes Equation (11). The Gaussian-bound (GB) representation replaces $1/\alpha_s(N_s)$ by $\beta_s = N_s^{-1} + 2L^2/(\rho_s N_s^2)$

and minimizes

$$g_r(N; w) = \min_{\sum_s H_s^{1/2} z_s + b = w} \left\{ \sum_{s:N_s>0} \beta_s \|z_s\|^2 + r^2 \|b\|^2 \right\}. \tag{18}$$

The optimized bound has the closed form

$$g_r(N; w) = w^\top \left[\sum_{s:N_s>0} \beta_s^{-1} H_s + r^{-2} P_{\mathcal{U}} \right]_{\mathcal{U}}^{-1} w. \tag{19}$$

Let q be the bracketed inverse applied to w. Then $z_s^{\mathrm{GB}} = \beta_s^{-1} H_s^{1/2} q$ at active sites and $b^{\mathrm{GB}} = r^{-2} q$. Computing GB requires one linear solve after allocation. Appendix J derives Equation (19) and the corresponding Gaussian-risk identities. Appendix I.2 gives the eICU matrix-split implementation, with criterion sensitivity analyzed in Appendix F.

Inactive sites have $z_s = 0$. For $z_s \neq 0$, let $u_s = z_s/\|z_s\|$. We average the clipped score $\mathrm{clip}_{[-L,L]}(u_s^\top H_s^{\dagger/2}(2A-1)(2Y-1)X)$ and add Gaussian noise with variance $2L^2/(\rho_s N_s^2)$. The coordinator sums the releases with weights $\|z_s\|$. The coefficient β_s includes sampling variance and Gaussian privacy variance. GB therefore chooses weights for the specified release at fixed sample sizes. Its variance crossover is $N_s = 2L^2/\rho_s$, which generally differs from the information-envelope crossover $1/\kappa_s$.

For the Gaussian estimator $\widehat{\tau}_z$, define its actual worst-case risk by

$$R_G^{\mathrm{wc}}(N, z) = \sup_{\theta \in \Theta(r)} \mathbb{E}_\theta[(\widehat{\tau}_z - w^\top \theta)^2]. \tag{20}$$

Thus ψ_r characterizes protocol-class minimax risk, g_r optimizes a bound for the specified release, and R_G^{wc} evaluates a selected estimator. With consistent moment factors, $\psi_r \leq g_r \leq C_G \psi_r$ for $C_G \leq C_+$, and GB gives $R_G^{\mathrm{wc}} \leq g_r$. Appendix J.1 proves these comparisons. These quantities can select different allocations, which the experiments compare.

6 Experiments

In this section, we investigate three questions: Q1. Can heterogeneous site-specific participant-level privacy constraints change the enrollment allocation selected before outcomes are observed? Q2. When does incorporating these privacy constraints into enrollment reduce target-population ATE MSE, and how do squared bias, sampling variance, and Gaussian privacy variance determine the change? Q3. How do numerical optimization, objective choice, and estimator weighting affect the risk of the implemented Gaussian estimator?

Experimental setup. Synthetic studies use Equation (1) with controlled site populations, costs, privacy constraints, and budgets. The semi-synthetic eICU study uses 162,282 stays from 158 hospitals, with their stay-weighted mixture as the target population. Enrollment, treatment, outcomes, and release noise are simulated on the empirical covariate supports. We compare Uniform, Classical c-optimal design, the sampling benchmark (Sampling) in Equation (14), and the conic construction for ψ_r, denoted ψ-Conic. Classical sets $b = 0$ in the sampling objective. Uniform uses $n_s = B/\sum_t c_t$ before flooring. On small problems, ψ-Reference and g-Reference enumerate the integer feasible set to minimize ψ_r and g_r, respectively. CB and GB denote the representations in Section 5.2. All eICU allocation comparisons use GB. Design quality is measured by ψ_r relative to numerical references and by active-site count. Estimator performance is measured by target-population ATE MSE over prespecified parameters or by actual worst-case risk R_G^{wc}. Paired simulations share random draws. MCSE denotes Monte Carlo standard error. Aggregate MCSEs are computed for the mean paired difference across conditions. W/U/L denotes wins, inconclusive conditions, and losses under the paired procedure in Appendix I. Appendices G, H, and I give the full protocols.

Privacy changes enrollment. A two-site synthetic construction isolates the enrollment decision. Site 1 supplies more sampling information per participant and is preferred by Classical c-optimal design, while site 2 has the weaker privacy constraint and is preferred by ψ_r. At $B = 256$ and $r = 1/4$, the Classical c-optimal allocation has a ψ_r value 5.00 times the optimum. Proposition 5 and

Table 1: Target-population ATE MSE in the semi-synthetic eICU study. Results over 108 conditions. Entries are MSE differences, ψ-Conic minus comparator. Negative values favor ψ-Conic. W/U/L denotes wins, inconclusive conditions, and losses under the paired procedure in Appendix I.

Comparator	Mean MSE difference	MCSE	W/U/L
Uniform	-0.0585	4.74×10^{-5}	86/2/20
Classical c-optimal design	-0.0109	3.28×10^{-5}	87/10/11
Sampling benchmark	-0.0107	3.28×10^{-5}	46/11/51

Table 2: Bias and variance decomposition of the eICU MSE difference. Each row summarizes 36 conditions after averaging the three randomization replicates. Differences are ψ-Conic minus the sampling benchmark. Negative values favor ψ-Conic. Analytic components and Monte Carlo means are evaluated separately.

Cost	Δ squared bias	Δ sampling variance	Δ Gaussian privacy variance	Mean MSE difference	W/U/L
Equal	-0.0392	$+1.05 \times 10^{-4}$	$+0.00787$	-0.0313	27/0/9
Inverse-size	-2.20×10^{-5}	$+4.70 \times 10^{-7}$	$+6.30 \times 10^{-5}$	$+4.40 \times 10^{-5}$	4/4/28
Lognormal	-6.16×10^{-4}	$+1.00 \times 10^{-6}$	$+7.80 \times 10^{-5}$	-5.43×10^{-4}	14/1/21

Corollary 2 in Appendix G.1 establish the reversal and show that the corresponding minimax-risk ratio is unbounded along the budget-indexed family. Heterogeneous privacy can therefore change which sites receive enrollment before any private estimator is formed.

Target-population ATE error. The semi-synthetic eICU study evaluates whether this change in enrollment improves the final target-population ATE estimate on empirical hospital covariate distributions. Table 1 holds GB fixed and reports MSE differences over 108 conditions. ψ-Conic has lower mean MSE than Uniform, Classical c-optimal design, and the sampling benchmark. Relative to the sampling benchmark, the mean difference is -0.0107 with 46/11/51 W/U/L. Pooling all three randomization replicates gives a mean difference of -0.0106 with 45/5/58 W/U/L (Table 5, Appendix I.4). The negative mean therefore coexists with substantial condition-level heterogeneity.

Table 2 identifies the source of the MSE changes. Under equal costs, the squared-bias reduction of 0.0392 outweighs the added variance, giving a mean difference of -0.0313. Under inverse-size costs, the bias reduction is smaller than the added Gaussian privacy variance, and MSE increases. Lognormal costs favor ψ-Conic on average. Appendix I.5 reports the corresponding decomposition by treatment effect direction. Privacy-aware enrollment helps when coverage of target-relevant directions reduces squared bias enough to offset the release variance. Because all eICU allocation comparisons use GB, Tables 1 and 2 hold estimator weighting fixed across designs.

Realized estimator risk. We isolate estimator weighting at fixed site sample sizes. Appendix H, Figure 4(b), compares the method selected by ψ_r with the method minimizing evaluated estimator risk among the candidates. Across 32,400 problems, replacing CB with GB raises the selection agreement rate from 71.9% to 81.3%, while mean Spearman correlation rises from 0.642 to 0.890. A separate set of 20 small problems uses the same site sample sizes and random draws under CB and GB. GB gives selection agreement in 19/20 problems for maximum empirical MSE over the prespecified parameters and 20/20 for R_G^{wc}. GB includes both sampling variance and Gaussian privacy variance in its fixed-allocation objective.

Using GB, we enumerate all 3,016 feasible integer allocations in these 20 problems. The two reference designs agree on all ten high-privacy allocations and differ on all ten mixed allocations. In Table 3, ψ-Reference lowers mean ψ_r from 0.296 to 0.198 relative to ψ-Conic. Relative to ψ-Reference, g-Reference lowers mean g_r from 0.459 to 0.433. ψ-Conic has higher R_G^{wc} than Sampling in nine of the ten mixed problems.

Table 3: Design-objective comparison on ten mixed small problems. Mean objective values. Lower is better. All allocations use GB.

Design	ψ_r	g_r	R_G^{wc}
Sampling benchmark	0.229	0.494	0.484
ψ-Conic	0.296	0.519	0.512
ψ-Reference	0.198	0.459	0.447
g-Reference	0.238	0.433	0.425

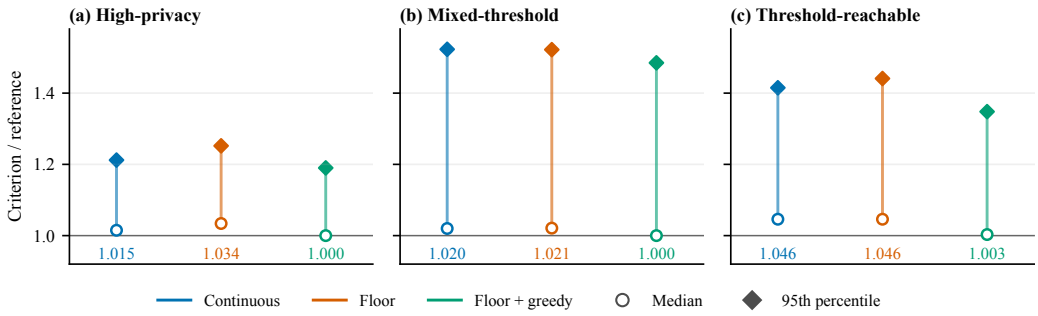


Figure 2: Optimization quality and integer recovery. Each input group contains 15,360 instances per candidate. Circles and diamonds show median and 95th-percentile criterion ratios. Continuous designs use numerical lower references and integer designs use exhaustive floating-point minima. Floor + greedy fills residual budget after flooring.

Figure 2 evaluates the conic construction and integer recovery on 46,080 small instances in three privacy configurations. The continuous median/95th-percentile ratios of ψ_r to the numerical reference are 1.015/1.212, 1.020/1.523, and 1.046/1.415. On 115,200 larger instances, compression leaves a median of three or four active sites (Appendix G.2, Figure 3(e)). Appendix D.2 reports floor-to-continuous criterion ratios of 1.025/1.121 and 3,097 improvements from the floor/ceiling hybrid. On 192 held-out validation instances, a local exchange search that reduces ψ_r lowers R_G^{wc} relative to ψ-Conic in 153 cases and raises it in 39.

Further minimizing g_r gives 82 risk improvements and 110 ties (Table 7, Appendix J.2). The 39 increases after the ψ_r exchange search reflect the distinction between the two risk criteria. ψ_r characterizes protocol-class minimax risk within constants, while R_G^{wc} is the risk of the specified Gaussian estimator. The fixed-allocation CB/GB comparison changes only estimator weights. The ψ-Reference and g-Reference comparison searches the same integer feasible set with GB fixed, so it changes only the design objective. Comparing ψ-Conic with ψ-Reference then measures the loss from the conic construction and integer recovery on these small problems. On the ten mixed small problems, exhaustive optimization of ψ_r reduces mean R_G^{wc} from 0.512 to 0.447 relative to ψ-Conic. Optimizing g_r reduces it further to 0.425, with a median reduction of 4.83% relative to ψ-Reference (Appendix J, Figure 5). On the ten high-privacy small problems, exhaustive ψ_r and g_r optimization selects the same allocation in every case (Appendix J, Table 6).

7 Conclusion

In this paper, we study enrollment allocation across multisite randomized trials when sites differ in target-population relevance, recruitment cost, and participant-level privacy. We develop a statistical and computational framework for privacy-aware trial design targeted to a specific population. We characterize the information available under heterogeneous privacy constraints and derive a design criterion balancing estimation variance and worst-case squared bias. The criterion achieves protocol-class minimax MSE within constants independent of the budget and number of sites. We also establish the computational difficulty of global optimization and develop sparse conic constructions, budget-feasible integer recovery, and compatible private estimator weights. Experiments show that privacy heterogeneity can materially change optimal enrollment and that accounting for it can improve target-population estimation when bias reduction outweighs the additional privacy variance.

Limitations and future work. Our analysis focuses on a common linear treatment effect with public site covariate moments and target mean, which allows an explicit characterization of the privacy-aware design problem. An interesting next step is to estimate these design inputs from protected data, which would couple enrollment decisions with statistical uncertainty and additional privacy expenditure. Extending the theory to nonlinear treatment effect models or uncertain target populations is also nontrivial because the resulting bias need not admit the finite-dimensional characterization used here. Adaptive enrollment raises a distinct challenge, as recruitment, estimation, and privacy accounting become sequentially dependent as trial data accumulate.

AI Use Statement

Generative AI assisted with manuscript organization and language editing.

References

Farid Alizadeh and Donald Goldfarb. Second-order cone programming. *Mathematical programming*, 95(1):3–51, 2003.

Mark Bun and Thomas Steinke. Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *Theory of cryptography conference*, pp. 635–658. Springer, 2016.

T. Tony Cai and Yicheng Li. A van Trees lower bound for fully interactive differentially private federated learning. *arXiv preprint arXiv:2605.19813*, 2026.

Wei-Ning Chen, Graham Cormode, Akash Bharadwaj, Peter Romov, and Ayfer Özgür. Federated experiment design under distributed differential privacy. In *Proceedings of the 27th International Conference on Artificial Intelligence and Statistics*, pp. 2458–2466. PMLR, 2024.

Xiaojun Chen, Dongdong Ge, Zizhuo Wang, and Yinyu Ye. Complexity of unconstrained L_2–L_p minimization. *Mathematical Programming*, 143(1):371–383, 2014.

Yu-Wei Chen, Raghu Pasupathy, and Jordan Awan. Optimal survey design for private mean estimation. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 8856–8870, 2025. URL `https://proceedings.mlr.press/v267/chen25ay.html`.

Issa J Dahabreh, Sarah E Robertson, Eric J Tchetgen, Elizabeth A Stuart, and Miguel A Hernán. Generalizing causal inferences from individuals in randomized trials to all trial-eligible individuals. *Biometrics*, 75(2):685–694, 2019.

Hoang Dang, Luan Pham, and Minh Nguyen. Target-weighted Neyman allocation: Experimental design for heterogeneous treatment effects under population shift. *arXiv preprint arXiv:2608.06512*, 2026.

John C Duchi, Michael I Jordan, and Martin J Wainwright. Minimax optimal procedures for locally private estimation. *Journal of the American Statistical Association*, 113(521):182–201, 2018.

Gustav Elfving. Optimum allocation in linear regression theory. *The Annals of Mathematical Statistics*, pp. 255–262, 1952.

Michael R Garey and David S Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*, volume 29. wh freeman New York, 2002.

Dongdong Ge, Xiaoye Jiang, and Yinyu Ye. A note on the complexity of L_p minimization. *Mathematical programming*, 129(2):285–299, 2011.

Richard D. Gill and Boris Y. Levit. Applications of the van Trees inequality: a Bayesian Cramér-Rao bound. *Bernoulli*, 1(3):59–79, 1995. URL `http://dml.mathdoc.fr/item/1186078362`.

David A Hirshberg and Stefan Wager. Augmented minimax linear estimation. *The Annals of Statistics*, 49(6):3206–3227, 2021.

Tatsuki Koga, Kamalika Chaudhuri, and David Page. Differentially private multi-site treatment effect estimation. In *2024 IEEE Conference on Secure and Trustworthy Machine Learning (SaTML)*, pp. 472–489. IEEE, 2024.

My Phan, David Arbour, Drew Dimmery, and Anup Rao. Designing transportable experiments under S-admissability. In *International Conference on Artificial Intelligence and Statistics*, pp. 2539–2547. PMLR, 2021.

Tom Pollard, Alistair Johnson, Jesse Raffa, Leo Anthony Celi, Omar Badawi, and Roger Mark. eICU Collaborative Research Database. *PhysioNet*, April 2019. doi: 10.13026/C2WM1R. URL `https://doi.org/10.13026/C2WM1R`. Version 2.0.

Tom J. Pollard, Alistair E. W. Johnson, Jesse D. Raffa, Leo A. Celi, Roger G. Mark, and Omar Badawi. The eICU collaborative research database, a freely available multi-center database for critical care research. *Scientific data*, 5(1):180178, 2018.

Stephen W. Raudenbush and Xiaofeng Liu. Statistical power and optimal design for multisite randomized trials. *Psychological Methods*, 5(2):199–213, 2000.

Guillaume Sagnol. Computing optimal designs of multiresponse experiments reduces to second-order cone programming. *Journal of Statistical Planning and Inference*, 141(5):1684–1708, 2011.

Wenqi Shi and Xi Lin. Strategy to select most efficient RCT samples based on observational data. *arXiv preprint arXiv:2211.04672*, 2022.

A Complete assumptions and causal identities

Sampling and causal model. Participants have independent covariates $X \sim Q_s$ within each site and treatment $A \mid X, J = s \sim \text{Bernoulli}(1/2)$ assigned by conditional randomization. Potential outcomes follow Equation (1) with a shared θ across source sites and the target. The parameter class satisfies $\Theta(r) \subseteq \Theta_0$ with positivity margin η, and the target mean satisfies $w \in \mathcal{U}$.

Design inputs and release bounds. The coordinator knows $(H_s, c_s, \rho_s, B, w, r)$ and the public leverage bound L before recruitment. Each complete public transcript satisfies participant-level ρ_s-zCDP at site s. The constructive upper bound additionally assumes $\|H_s^{\dagger/2} X\| \le L$ and $0 < \rho_s \le \bar{\rho}$. The lower bound does not require either the leverage bound or this upper bound on the privacy parameters.

Design inputs and privacy scope. The privacy analysis treats the site moment matrices H_s, target vector w, and inputs used to calibrate the private release as fixed public inputs. Estimating or releasing these quantities from protected records requires additional privacy accounting and is outside the present analysis. The semi-synthetic study evaluates the subsequent trial mechanism conditional on its empirical covariate distributions and moments.

Protocol and regularity correspondence. The local dataset at a hospital consists of independent observed triples (X, A, Y). Datasets are independent across hospitals at fixed θ. Recruitment and design inputs are fixed before outcomes. The protocol uses θ-independent kernels on standard Borel spaces, with private local states and independent local randomness, and public coordination based only on the recorded history. We allow finitely many ordered transmissions and almost surely finite stopping recorded in the transcript. For an almost surely finite stopping time, apply the finite-round contraction bound to stopped-transcript prefixes and pass to the complete stopped transcript by L^2 convergence of conditional expectations of the bounded full-data score. The dataset size is fixed, and only the communication length is random. Privacy concerns the joint transcript conditional on every fixed value of the other hospitals' data. Hidden shared randomness, secure aggregation, and trusted-server private state are excluded. These conditions match Definition 2.2 and the transcript conventions of Cai & Li (2026). The privacy unit is one participant record.

Likelihood regularity. For this Bernoulli experiment, with $D = (2A-1)(2Y-1)$, the single-record likelihood ratio relative to $\theta = 0$ is $1 + DX^\top\theta$. On the parameter ball it is bounded between η and $2-\eta$. Let $\mathcal{X}_S = \bigcup_s \text{supp}(Q_s)$ be the source-covariate range and define $K_X = \sup_{x \in \mathcal{X}_S} \|x\|_2$. Source covariates lie in $\mathcal{U}$ almost surely, and positivity on the whole ball gives $K_X \le (1-\eta)/r < \infty$. Consequently the projected score $DX/(1 + DX^\top\theta)$ is uniformly bounded at each fixed r. Finite product likelihoods have integrable derivatives. Applying any admissible parameter-independent transcript kernel preserves domination (by its law at zero) and differentiation under the integral, giving the transcript score as the conditional expectation of the full-data score. Thus the differentiability, integrability and score-projection requirements of the external contraction theorem hold, including singular site information. The one-dimensional prior used below supplies the boundary and prior-information requirements of van Trees.

The upper-bound constant $C_+ = 1 + 4L^2 e^{2\bar{\rho}}$ increases quadratically with L. The leverage condition enters through $\mathbb{E}\|H_s^{\dagger/2} X\|^2 = \text{rank}(H_s) \le L^2$. Uniform risk constants across a sequence of instances require fixed bounds on L, $\bar{\rho}$, and η. The lower bound uses only the positivity margin.

Moments of the signed outcome. The conditional causal moment used by the estimator is exact:

$$V_{si} = (2A_{si} - 1)(2Y_{si} - 1)X_{si},$$
$$\mathbb{E}_\theta[V_{si} \mid X_{si}] = X_{si}X_{si}^\top\theta, \qquad \mathbb{E}_\theta[V_{si}] = H_s\theta. \tag{21}$$

Since $D^2 = 1$, $\mathbb{E}_\theta[V_{si}V_{si}^\top] = H_s$ for every valid θ. Consequently, $\text{Cov}_\theta(V_{si}) = H_s - H_s\theta\theta^\top H_s$, which reduces to H_s at $\theta = 0$. In general, differentiating the Bernoulli log likelihood gives (3). For any instance with $K_X r < 1$, including singular H_s, the pointwise scalar bound on the integrand gives the uniform Loewner comparison

$$H_s \preceq I_s(\theta) \preceq \{1 - (K_X r)^2\}^{-1} H_s, \qquad \theta \in \Theta(r). \tag{22}$$

Thus along a sequence with $K_X r \to 0$, the scalar upper multiplier is $1 + O(K_X^2 r^2)$.

B Proofs of the information, risk, and hardness results

B.1 Minimum-energy identity and Theorem 1

For positive semidefinite H_s and $\alpha_s \geq 0$, let $G = [\sqrt{\alpha_1}H_1^{1/2}, \ldots, \sqrt{\alpha_S}H_S^{1/2}]$. Since $GG^\top = \sum_s \alpha_s H_s$, the Moore-Penrose minimum-norm identity gives

$$q_{\sum_s \alpha_s H_s}(w) = \min_{\substack{z_s \in \mathrm{range}(H_s) \\ \sum_s H_s^{1/2} z_s = w}} \sum_s \frac{\|z_s\|^2}{\alpha_s}, \tag{23}$$

with $\|z\|^2/0 = 0$ for $z = 0$ and $+\infty$ otherwise. From (7), $R_s \preceq \min\{n_s, \kappa_s n_s^2\} P_{H_s} = \alpha_s P_{H_s}$, so every feasible aggregate is Loewner-dominated by $\overline{M}_\rho(n)$ and its q value is no smaller. If $w \notin \mathrm{range}(\overline{M}_\rho(n))$, both sides are infinite. Otherwise, for a minimizer of (23), choose $R_s^\star = \alpha_s z_s^\star z_s^{\star\top}/\|z_s^\star\|^2$ for nonzero blocks. It satisfies both constraints, and applying the block minimum-energy identity to $R_s^\star$ yields the reverse inequality. This proves (8) within the contraction class.

Applying (23) after adding the block $r^{-2}P_{\mathcal{U}}$ gives (11). Its new representation variable is b, and its quadratic cost is $r^2\|b\|^2$.

B.2 Theorem 2: lower and upper bounds

The lower bound controls the information in any private transcript, then chooses a one-dimensional submodel aligned with the target. The upper bound constructs a scalar Gaussian release for each site and optimizes its representation.

Lower bound. For a fixed integer design, put $J_s(\theta) = N_s I_s(\theta)$. Theorem 3.4 of Cai & Li (2026) gives contributions $A_s = J_s^{1/2} B_s J_s^{1/2}$ with $0 \preceq B_s \preceq P_{J_s}$ and $\mathrm{Tr}\, B_s \leq \kappa_s N_s$. Since a PSD matrix's largest eigenvalue is at most its trace,

$$A_s \preceq \min(1, \kappa_s N_s) J_s = \alpha_s(N_s) I_s(\theta) \preceq [\eta(2-\eta)]^{-1} \alpha_s(N_s) H_s.$$

Zero-count sites contribute zero. Summing gives, uniformly on $\Theta(r)$,

$$I_G(\theta) \preceq C_\eta M_N, \qquad C_\eta = [\eta(2-\eta)]^{-1}. \tag{24}$$

For nonzero $v \in \mathcal{U}$, set $a = r/\|v\|$ and restrict to $\theta = tv$, $|t| \leq a$. Use the density $p_a(t) = a^{-1}\cos^2(\pi t/(2a))$ on this interval and zero outside. It integrates to one, is continuously differentiable with zero boundary values, and has Fisher information π^2/a^2. The target derivative is $w^\top v$. Scalar van Trees (Gill & Levit, 1995) gives

$$R_N^\star \geq \frac{(w^\top v)^2}{C_\eta v^\top M_N v + \pi^2 r^{-2}\|v\|^2}.$$

The supremum of this Rayleigh quotient equals $w^\top (C_\eta M_N + \pi^2 r^{-2} P_{\mathcal{U}})^{-1} w$. Loewner comparison with $C_-(M_N + r^{-2}P_{\mathcal{U}})$ yields $R_N^\star \geq C_-^{-1}\psi_r(N; w)$ for $C_- = \max(C_\eta, \pi^2)$. A different least-favorable line may be chosen for each N because the bound holds for every fixed design and protocol before taking their infima.

Upper bound. Omit $N_s = 0$ sites (their blocks are necessarily zero), and fix a feasible (z, b) in (11). When $z_s \neq 0$, let $u_s = z_s/\|z_s\|$ and define the globally bounded statistic

$$T_{si} = \mathrm{clip}_{[-L,L]}\left(u_s^\top H_s^{\dagger/2} V_{si}\right), \qquad \mathrm{clip}_{[-L,L]}(t) = \max\{-L, \min\{L, t\}\}.$$

The leverage assumption makes clipping inactive almost surely under the model, so $\mathbb{E}T_{si} = u_s^\top H_s^{1/2}\theta$ and $\mathbb{E}T_{si}^2 = 1$ remain unchanged. Site s releases

$$\widetilde{T}_s = N_s^{-1} \sum_{i=1}^{N_s} T_{si} + Z_s, \qquad Z_s \sim \mathcal{N}\left(0, \frac{2L^2}{\rho_s N_s^2}\right). \tag{25}$$

Boundedness on the entire record domain gives replace-one sensitivity at most $2L/N_s$, so the Gaussian zCDP identity verifies ρ_s-zCDP. The server estimator $\widehat{\tau}_z = \sum_s \|z_s\| \widetilde{T}_s$ has expectation $(w-b)^\top \theta$ and hence squared bias at most $r^2\|b\|^2$. Its variance is at most

$$\sum_s \|z_s\|^2 \left(N_s^{-1} + \frac{2L^2}{\rho_s N_s^2} \right) \le C_+(L, \bar{\rho}) \sum_s \frac{\|z_s\|^2}{N_s\{1 \wedge \kappa_s N_s\}}. \tag{26}$$

One may take $C_+ = 1 + 4L^2 e^{2\bar{\rho}}$, since $1/\rho \le 2e^{2\bar{\rho}}/(e^{2\rho} - 1)$ on $(0, \bar{\rho}]$. Optimizing (z, b) proves the upper inequality and then optimizing N proves design-level equivalence.

B.3 Direct strong NP-hardness with signed variables

We use the strongly NP-complete restricted 3-PARTITION family (Garey & Johnson, 2002, Section 4.2 and problem SP15): $n_0 = 3m \ge 3$, positive integers a_i, T with $T/4 < a_i < T/2$, $\sum_i a_i = mT$, and T polynomially bounded in n_0. The assignment vector is *unrestricted in sign*, $\gamma \in \mathbb{R}^{n_0 m}$. Define $C \in \mathbb{Z}^{(n_0+m)\times n_0 m}$ by

$$C_{\ell,(i,j)} = \mathbf{1}\{\ell = i\} + a_i \mathbf{1}\{\ell = n_0 + j\}, \qquad h = (\mathbf{1}_{n_0}, T\mathbf{1}_m)^\top.$$

Thus $C\gamma = h$ imposes item sums one and weighted bin sums T. Set $p = 2/3$, $q(\gamma) = \sum_{i,j} |\gamma_{ij}|^p$, and

$$P = 256Tn_0^3, \qquad \delta = P^{-1}, \qquad \Lambda = 2n_0^3 P^2, \qquad F_\Lambda(\gamma) = q(\gamma)^3 + \Lambda \|C\gamma - h\|_2^2. \tag{27}$$

Lemma 1 (Quantitative signed near-integrality)**.** *If $\|C\gamma - h\|_2 \le \delta$ and $q(\gamma) \le n_0 + \delta$, then every row admits an index j_i with $\gamma_{ij_i} > 0$ and*

$$\|\gamma_i - e_{j_i}\|_1 < 31 n_0^{3/2} \delta.$$

Rounding row i to e_{j_i} gives a valid 3-partition.

Proof. Put $s_i = \sum_j \gamma_{ij}$, $q_i = \sum_j |\gamma_{ij}|^p$, and $e_i = q_i - |s_i|^p$. The residual bound gives $|s_i - 1| \le \delta < 1/2$. Subadditivity for $0 < p < 1$ gives $e_i \ge 0$ and

$$\sum_i e_i \le n_0 + \delta - n_0(1-\delta)^p \le 2n_0\delta^p, \qquad e_i \le 2n_0\delta^p,$$

where $1 - (1-\delta)^p \le \delta^p$ and $\delta \le \delta^p$. For a row, let $M_i = \max_j |\gamma_{ij}|$ and $h_i = \sum_{j \ne j_i} |\gamma_{ij}|$, choosing any maximizer j_i. Regardless of signs,

$$e_i \ge M_i^p + h_i^p - (M_i + h_i)^p \ge (1-p)\min(M_i, h_i)^p.$$

For the last inequality, if $a \ge b > 0$, concavity yields $(a+b)^p \le a^p + pa^{p-1}b$, hence $a^p + b^p - (a+b)^p \ge b^p[1 - p(b/a)^{1-p}] \ge (1-p)b^p$. Zero cases follow directly.

If $h_i > M_i$, then $M_i \ge |s_i|/m \ge 1/(2m)$ and

$$e_i \ge \tfrac{1}{3}(2m)^{-2/3} > \frac{1}{8n_0} \ge 2n_0\delta^{2/3}.$$

Here $P \ge 64n_0^3$ implies the last inequality. The strict middle inequality follows from $m = n_0/3$ and $8n_0^{1/3} > 3(2/3)^{2/3}$ for $n_0 \ge 3$. This contradicts the slack bound, so $h_i \le M_i$. Consequently,

$$h_i \le (3e_i)^{3/2} \le (6n_0)^{3/2}\delta < 15 n_0^{3/2}\delta.$$

The dominant coefficient must be positive: otherwise $s_i \le -M_i + h_i \le 0$, whereas $s_i \ge 1 - \delta > 0$. Therefore $|\gamma_{ij_i} - 1| \le \delta + h_i$ and $\|\gamma_i - e_{j_i}\|_1 \le \delta + 2h_i < 31 n_0^{3/2}\delta$.

Write $\widehat{\gamma}_i = e_{j_i}$ and $\widehat{B}_j = \sum_i a_i \widehat{\gamma}_{ij} \in \mathbb{Z}$. Each bin residual obeys

$$\begin{aligned} |\widehat{B}_j - T| &\le \delta + \sum_i a_i |\widehat{\gamma}_{ij} - \gamma_{ij}| \\ &< \delta + \tfrac{T}{2} \sum_i \|\widehat{\gamma}_i - \gamma_i\|_1 < \frac{1}{256Tn_0^3} + \frac{31}{512\sqrt{n_0}} < \frac{1}{2}. \end{aligned}$$

For example the last expression is at most $1/256 + 31/512 < 1/2$ for $T, n_0 \ge 1$. Thus every integer bin total equals T. Since $T/4 < a_i < T/2$, every bin has exactly three items. □

Gap. A yes-instance has a one-hot feasible assignment, so $\min F_\Lambda \le n_0^3$. On a no-instance, Lemma 1 implies, for every γ, either $\|C\gamma - h\|_2 > \delta$ or $q(\gamma) > n_0 + \delta$. In the first case $F_\Lambda > 2n_0^3 > n_0^3 + 3n_0^2\delta$ since $n_0 > 3\delta$. In the second, $F_\Lambda > (n_0 + \delta)^3 > n_0^3 + 3n_0^2\delta$. The continuous coercive function F_Λ attains its minimum, so the strict gap also holds at the optimum:

$$\text{yes: } \min_\gamma F_\Lambda \le n_0^3, \qquad \text{no: } \min_\gamma F_\Lambda > n_0^3 + 3n_0^2/P. \tag{28}$$

Embedding, target and privacy feasibility. The causal feature dimension is $d_{\text{feat}} = n_0 + m$ and there are $S = n_0 m$ sites. Put $D = 1 + n_0 + mT$, $a = (2D)^{-1}$, $x_{ij} = aC_{\cdot,(i,j)}$, $Q_{ij} = \delta_{x_{ij}}$, $G_{ij} = x_{ij}$ and $H_{ij} = x_{ij}x_{ij}^\top$. Choose $w = ah$ and $Q_T = \delta_w$. Since the fractional assignment $\gamma_{ij} = 1/m$ satisfies $C\gamma = h$, $w \in \text{range}(C) = \mathcal{U}$. Both $\|x_{ij}\|$ and $\|w\|$ are at most $1/2$. With fixed $r_0 = 1/4$, all source and target inner products have magnitude at most $1/8$, giving positivity margin $\eta = 7/8$ on the entire ball. Point-mass sites have leverage one.

Set $c_s = 1$, $B = 64D^2\Lambda$ and $\kappa_s = B^{-1}$. Every feasible n satisfies $n_s \le B$ and $\kappa_s n_s \le 1$, so the *entire* continuous domain is high-privacy. Scalar factor coordinates are signed: $G_s\gamma_s = x_s\gamma_s$. Eliminating the sample allocations in the minimum-energy formula gives

$$\Psi^{\text{hp}}_{B,r_0}(w) = \min_{\gamma \in \mathbb{R}^{n_0 m}} \left\{ \frac{q(\gamma)^3}{\kappa B^2} + r_0^2 a^2 \|C\gamma - h\|_2^2 \right\}.$$

Equivalently, the scaled coordinates of (15) are $u_s = \gamma_s/\sqrt{\kappa}$. Since $\kappa B^2 = B$ and $Br_0^2a^2 = \Lambda$, the optimal-value identity is

$$B\Psi^{\text{hp}}_{B,r_0}(w) = \min_{\gamma \in \mathbb{R}^{n_0 m}} F_\Lambda(\gamma). \tag{29}$$

Computational representation. The hardness input consists of rational factors G_s, costs, κ_s, B, r_0 and w. The statistical privacy level is represented symbolically as $\rho_s = \log(1 + \kappa_s)/2$. On the restricted 3-PARTITION family all numerators and denominators are polynomially bounded in n_0: $P = O(Tn_0^3)$, $\Lambda = O(T^2n_0^9)$, $D = O(n_0T)$ and $B = O(T^4n_0^{11})$. There are polynomially many coordinates and each has $O(\log n_0)$ bits when T is polynomial in n_0. Squaring rational factors to obtain H_s preserves these bounds. The decision threshold $(n_0^3 + 3n_0^2/(2P))/B$ separates the cases, with inverse-polynomial value gap $3n_0^2/(PB)$. Approximating the optimum to additive error less than one quarter of that gap would decide the restricted family. This proves strong NP-hardness for rational κ_s inputs.

C Conic construction of the experimental design

C.1 Construction stages and coordinate conventions

Algorithm 1 Experimental design construction and private estimation

Require: $(H_s, c_s, \rho_s, w, r, B)$, factors $H_s = G_sG_s^\top$, public L, record domain
Ensure: Integer sample sizes N, estimator representation (z, b), private estimate $\widehat{\tau}$

```
if κ_s B/c_s ≤ 1 for every site then
    Solve Equation (32). Compress preserving b, w − b, T
    Recover continuous sample sizes by Equation (34)
else
    Solve Equation (35). Compress preserving b, w − b, L_S, L_P
    Recover continuous sample sizes by Equation (37)
end if
Check branch support, representation accuracy, and budget feasibility
Recover N by flooring or an applicable rule in Appendix D
Compute CB or GB at N using Equation (11) or Equation (18)
Calibrate the site releases using L on the declared record domain
Combine private site summaries to obtain τ̂
```

The continuous guarantee uses the accuracy conditions of Theorem 3. Integer comparisons use the chosen recovery rule. The semi-synthetic eICU study uses pure flooring and the GB implementation in Appendix I.2.

For a supplied factor $G_s \in \mathbb{R}^{d\times m_s}$, the conic block z_s belongs to $\mathbb{R}^{m_s}$. Removing its component in $\ker(G_s)$ preserves $G_s z_s$ and cannot increase its norm. On $\mathrm{range}(G_s^\top)$ the map $z_s \mapsto H_s^{\dagger/2} G_s z_s$ is an isometry into $\mathrm{range}(H_s)$ and gives the symmetric-square-root coordinates used in Equation (11). In particular, $H_s^{1/2} H_s^{\dagger/2} G_s z_s = G_s z_s$. The standard solvers use $d\times d$ symmetric roots and d-component blocks. For rank-deficient cases, we instead solve in effective coordinates and lift the result back. These transformations preserve the represented target and block norms. Appendix I describes the truncated factors used for eICU.

For nonnegative t_s on k active blocks,

$$(\sum_s t_s)^2 \le (\sum_s t_s^{2/3})^3 \le k(\sum_s t_s)^2. \tag{30}$$

The high-privacy construction applies this inequality after compression. For mixed privacy, write $V(z;n) = \sum_s \|z_s\|^2/\alpha_s(n_s)$. Its minimum over sample allocations is

$$V_B(z) = \min_{\sum_s c_s n_s \le B} \sum_s \|z_s\|^2 \max\{n_s^{-1}, (\kappa_s n_s^2)^{-1}\}. \tag{31}$$

The branch-specific programs and recovery bounds follow below.

Sampling-optimum applicability. $\alpha_s(n_s) \le n_s$ makes the sampling optimum a lower bound on the private continuous optimum. If an optimal sample-size vector associated with (14) satisfies $\kappa_s n_s \ge 1$ at every active site, equality at that vector proves it is also private-optimal.

C.2 Sampling recovery and conic conventions

For the sampling program (14), return $c_s n_s = B\sqrt{c_s}\|z_s\|/L_S$ if $L_S = \sum_s \sqrt{c_s}\|z_s\| > 0$. Otherwise set all site sample sizes to zero. All block norms use nonnegative epigraph variables, and squared terms use rotated second-order cones. The high-privacy and mixed programs below use the same input factors. Zero blocks receive zero samples, and an all-zero representation returns the zero allocation. Compression uses a linear-programming basic feasible representation with block directions fixed. The exact-real version uses a feasible SOCP return with relative objective error $\varepsilon_{\mathrm{solve}}$ and exact compression and recovery. With numerical post-processing, $\varepsilon_{\mathrm{solve}}$ is only part of the total ε in Theorem 3. Appendix C.5 gives the combined condition.

C.3 High-privacy SOCP and compression

Obtain a feasible return with objective at most $1 + \varepsilon_{\mathrm{solve}}$ times the optimum of

$$\min_{\sum_s \widetilde{G}_s u_s + b = w} \left\{ B^{-2}(\sum_s \|u_s\|)^2 + r^2\|b\|^2 \right\}. \tag{32}$$

Let $T = \sum_s \|u_s\|$ and $v_s = u_s/\|u_s\|$ for active blocks. The vector $(w-b, T)$ is a conic combination of $(\widetilde{G}_s v_s, 1)$ in dimension $d_{\mathrm{eff}} + 1$. Fixing v_s, a basic feasible solution of

$$\sum_s \lambda_s(\widetilde{G}_s v_s, 1) = (w-b, T), \qquad \lambda_s \ge 0 \tag{33}$$

uses at most $d_{\mathrm{eff}} + 1$ atoms and preserves the relaxed objective. Replace u_s by the compressed block $u_s = \lambda_s v_s$, then recover $z_s = \sqrt{\kappa_s} u_s/c_s$ and use the exact high-privacy budget allocation

$$c_s n_s = B\frac{c_s^{2/3}\kappa_s^{-1/3}\|z_s\|^{2/3}}{\sum_j c_j^{2/3}\kappa_j^{-1/3}\|z_j\|^{2/3}}. \tag{34}$$

C.4 Arbitrary-mixed SOCP, augmented compression, and allocation

For a representation z, define the sampling and privacy envelopes

$$S_B = B^{-1}(\sum_s \sqrt{c_s}\|z_s\|)^2, \qquad P_B = B^{-2}(\sum_s c_s^{2/3}\kappa_s^{-1/3}\|z_s\|^{2/3})^3,$$

and the convex privacy envelope $\widetilde{P}_B = B^{-2}(\sum_s c_s\|z_s\|/\sqrt{\kappa_s})^2$.

Let $k_{\mathrm{mix}} = d_{\mathrm{eff}} + 2$ and obtain the same feasible relative-accuracy guarantee for

$$\min_{\sum_s G_s z_s + b = w} \left\{2S_B(z) + 4k_{\mathrm{mix}}\widetilde{P}_B(z) + r^2\|b\|^2\right\}. \tag{35}$$

At any feasible representation, the objective in Equation (35) bounds the criterion of the design recovered after exact compression and the allocation below. For active z_s, fix $v_s = z_s/\|z_s\|$ and compress the augmented atoms

$$\mathfrak{a}_s = (G_s v_s, \sqrt{c_s}, c_s/\sqrt{\kappa_s}) \in \mathcal{U} \times \mathbb{R}^2. \tag{36}$$

A basic feasible representation of $(w-b, L_S, L_P)$ uses at most $d_{\mathrm{eff}}+2$ atoms while exactly preserving $L_S = \sum_s \sqrt{c_s}\|z_s\|$ and $L_P = \sum_s c_s\|z_s\|/\sqrt{\kappa_s}$. Set the compressed blocks to $z_s = \lambda_s v_s$. On this representation, compute a sampling-only allocation with budget $B/2$ and a privacy-only allocation with budget $B/2$:

$$\begin{aligned} c_s n_s^{(S)} &= \frac{B}{2}\frac{\sqrt{c_s}\|z_s\|}{\sum_j \sqrt{c_j}\|z_j\|}, \\ c_s n_s^{(P)} &= \frac{B}{2}\frac{c_s^{2/3}\kappa_s^{-1/3}\|z_s\|^{2/3}}{\sum_j c_j^{2/3}\kappa_j^{-1/3}\|z_j\|^{2/3}}, \qquad n^{\mathrm{mix}} = n^{(S)} + n^{(P)}. \end{aligned} \tag{37}$$

Monotonicity gives $V(z; n^{\mathrm{mix}}) \le 2S_B + 4P_B$ because the summed sample sizes dominate both half-budget allocations componentwise. Support size and (30) give $P_B \le k_{\mathrm{mix}}\widetilde{P}_B$ after compression. Thus exact compression and recovery give $\Psi^{\mathrm{cont}}_{B,r}(w) \le J^\star_M$, where $J^\star_M$ is the comparison-program optimum. For any exact mixed optimizer $(z^\star, b^\star, n^\star)$, both $S_B(z^\star)$ and $\widetilde{P}_B(z^\star) \le P_B(z^\star)$ are at most its variance $V^\star$. Hence $J^\star_M \le (2 + 4k_{\mathrm{mix}})V^\star + r^2\|b^\star\|^2 \le (4d_{\mathrm{eff}} + 10)\Psi^{\mathrm{cont}}_{B,r}(w)$. A feasible relative $1 + \varepsilon_{\mathrm{solve}}$ return, exact compression and the displayed allocation prove (17) with $\varepsilon = \varepsilon_{\mathrm{solve}}$. The high-privacy proof compares its lower-envelope relaxation with the exact quasi-norm objective, then applies the support inequality including the unchanged bias term.

C.5 Accuracy, feasibility and the experimental solver interface

To apply Theorem 3 to a numerical return, we account for solve accuracy, representation residuals, compression, and sample-size recovery in one factor Γ. The required condition is $\Gamma \le 1 + \varepsilon$ on the supplied instance.

Computational model. The hardness input uses rational factors and rational (c_s, κ_s, B, r, w). SOCP coefficients may involve algebraic square roots. The approximation proof assumes a feasible relative-accuracy SOCP oracle followed by exact real arithmetic, comparisons, roots, and elimination. Compression and recovery terminate in this model. The guarantee is conditional on this conic-solver interface and accounts for objective accuracy and arithmetic errors.

All stages use one supplied instance $H_s = G_s G_s^\top$. A truncated factor defines a different instance. Appendix F gives a conditional criterion comparison. For the conic solve, Alizadeh & Goldfarb (2003, Section 7) give $O(\sqrt{q})$ primal-dual iterations per constant-factor gap reduction under full-row-rank, strict-feasibility, and central-neighborhood conditions, where q counts cone blocks. The following accounting specifies the additional accuracy needed for the full construction.

Objective scale and feasibility repair. Let J denote the high-privacy relaxation objective or the mixed convex comparison objective, evaluated from actual block norms. For the high-privacy relaxation define $a_H = \max_s \|\widetilde{G}_s\|_{\mathrm{op}}$, and for mixed privacy let $a_S = \max_s \|G_s\|_{\mathrm{op}}/\sqrt{c_s}$. For $w \ne 0$ the respective positive lower bounds are

$$\ell_H = \frac{\|w\|^2}{B^2 a_H^2 + r^{-2}}, \qquad \ell_M = \frac{\|w\|^2}{(B/2)a_S^2 + r^{-2}}.$$

Indeed, $\|w\| \le a_H T + \|b\|$ and weighted Cauchy-Schwarz give $\|w\|^2 \le (B^2 a_H^2 + r^{-2})(B^{-2}T^2 + r^2\|b\|^2)$. The mixed argument uses $\|w\| \le a_S L_S + \|b\|$ and drops the nonnegative privacy term from J. Certified upper bounds on a_H, a_S suffice. The feasible point with zero blocks and $b = w$ provides

the upper bound $r^2\|w\|^2$ and bounds the relevant sublevel set. An additive objective certificate on the scale $\varepsilon_{\text{solve}}\ell_H$ or $\varepsilon_{\text{solve}}\ell_M$ can therefore imply relative accuracy. An absolute tolerance must be calibrated to this lower-bound scale to imply the stated relative guarantee.

Project b onto $\mathcal{U}$ first if necessary. For a candidate with representation residual $e = w - \sum_s G_s z_s - b$ (use $\widetilde{G}_s u_s$ in high privacy), replace b by $b + e$. This gives exact representation feasibility in the mathematical problem. Since only the squared-bias term changes,

$$\sqrt{J_{\text{repaired}}} \le \sqrt{J} + r\|e\|. \tag{38}$$

Thus a pre-repair objective certificate $J \le (1+\varepsilon_{\text{solve}})J^\star$ yields $J_{\text{repaired}} \le [\sqrt{1+\varepsilon_{\text{solve}}}+r\|e\|/\sqrt{\ell}]^2 J^\star$.

Support and preserved quantities. Compression has a constructive exact-arithmetic interpretation. If more atoms than augmented dimensions have positive weights, find a nonzero dependence $\sum_s h_s \mathfrak{a}_s = 0$. The positive cost coordinate forces h to have both signs. Subtract th from the weights, where $t = \min_{h_s>0} \lambda_s/h_s$. At least one weight becomes zero, none becomes negative, and the augmented sum is preserved. Repeating at most $\max\{0, S-k\}$ times gives at most k atoms. Unit block directions make the preserved sums equal to the actual block norms. If numerical compression gives a support-bounded representation with recomputed objective $J_c \le J_{\text{repaired}}+\delta_c$, the additional relative cost is at most δ_c/ℓ, followed by another application of (38) for its equality residual.

Combined error budget and sample-size recovery. Write $J_0 \le (1+\varepsilon_{\text{solve}})J^\star$ for the pre-repair objective certificate and let $0 < \ell \le J^\star$. Let e_1 be the first representation residual, and let numerical compression return at most the branch's prescribed number of blocks with actual objective $J_c \le J_1 + \delta_c$, where J_1 is the first repaired objective and $\delta_c \ge 0$. Repair the compressed representation residual e_2 in b and call its objective J_2. Put $t_i = r\|e_i\|/\sqrt{\ell}$. Applying (38) twice gives

$$J_2 \le QJ^\star, \qquad Q = \left[\sqrt{(\sqrt{1+\varepsilon_{\text{solve}}}+t_1)^2 + \delta_c/\ell} + t_2\right]^2.$$

The support restriction remains exact: the repair changes b only. Exact preservation is the special case $\delta_c = 0$, $e_2 = 0$. A numerical implementation must certify the displayed inequalities from actual blocks and the same input instance.

Set $d_H = d_{\text{eff}} + 1$ and $d_M = 1$. For the final feasible representation, define its criterion value at computed nonnegative sample allocations by $F(\widehat{n}) = \sum_s \|z_s\|^2/\alpha_s(\widehat{n}_s) + r^2\|b\|^2$, using $z_s = \sqrt{\kappa_s}u_s/c_s$ in high privacy. Require a recovery certificate $F(\widehat{n}) \le (1+\varepsilon_{\text{rec}})d_j J_2$, $j \in \{H, M\}$, with $\varepsilon_{\text{rec}} \ge 0$. The exact allocation formulas (34) and (37) give this inequality with $\varepsilon_{\text{rec}} = 0$ and cost B for nonzero representations. Zero blocks receive no samples.

If $C = \sum_s c_s\widehat{n}_s > B$, put $a = B/C$. Otherwise put $a = 1$, including $C = 0$. Return $n = a\widehat{n}$. Since $\alpha_s(a\widehat{n}_s) \ge a^2\alpha_s(\widehat{n}_s)$, the fixed-representation bound increases by at most a^{-2} under this budget repair. The required total factor is

$$\Gamma = a^{-2}(1+\varepsilon_{\text{rec}})\left[\sqrt{(\sqrt{1+\varepsilon_{\text{solve}}}+t_1)^2 + \delta_c/\ell} + t_2\right]^2 \le 1+\varepsilon. \tag{39}$$

Indeed, $\psi_r(n) \le a^{-2}F(\widehat{n}) \le d_j\Gamma J^\star$. The relevant comparisons are $J_H^\star \le \Psi_{B,r}^{\text{hp}}(w)$ for the high-privacy lower-envelope relaxation and $\Psi_{B,r}^{\text{cont}}(w) \le J_M^\star \le (4d_{\text{eff}}+10)\Psi_{B,r}^{\text{cont}}(w)$ for the mixed comparison program, giving exactly the two factors in Theorem 3 with $\Gamma \le 1+\varepsilon$. If $a = 1$, the residuals and compression error vanish, and recovery is exact, then $\Gamma = 1+\varepsilon_{\text{solve}}$. Using the entire ε for the solve leaves no allowance for positive errors at later stages.

Coefficient rounding and rank truncation change the supplied instance. Appendix F bounds the effect of a matrix perturbation on the criterion.

Numerical implementation. The numerical studies use CLARABEL feasibility and gap tolerances 10^{-8}, compression equality tolerance 2×10^{-7}, support tolerance 10^{-12}, and numerical rank tolerance 10^{-10}. An alternative norm-epigraph formulation minimizes $\sqrt{J}$ and squares the resulting value. Tolerances on that scale differ from tolerances on J. The optimization-quality study supplies

numerical branch-and-bound references (Appendix G.2.1). Solver tolerances and residuals summarize the numerical runs. Applying the approximation factor also requires the combined bound in Equation (39).

The semi-synthetic eICU study discards positive eigenvalues up to 1.27×10^{-16}. The maximum operator-norm difference between the empirical and truncated moment matrices is 1.21×10^{-15}. Numerical checks cover all 8,640 method-design combinations, and the maximum GB representation residual is 2.93×10^{-15}. Appendix I.2 specifies the release domain and representation correction, while Appendix F quantifies criterion sensitivity to the matrix approximation.

C.6 Integer rounding

For $n_{\min} = \min_{s:n_s>0} n_s > 1$, let $a = 1 - 1/n_{\min} > 0$. Every active site sample size satisfies $\lfloor n_s \rfloor \geq a n_s$, so $\alpha_s(\lfloor n_s \rfloor) \geq a^2 \alpha_s(n_s)$ on both branches. Flooring preserves the budget, and

$$\overline{M}_\rho(\lfloor n \rfloor) + r^{-2} P_{\mathcal{U}} \succeq a^2 (\overline{M}_\rho(n) + r^{-2} P_{\mathcal{U}}), \qquad \psi_r(\lfloor n \rfloor) \leq a^{-2} \psi_r(n).$$

For an A-approximate continuous return,

$$\psi_r(\lfloor n \rfloor) \leq a^{-2} \psi_r(n) \leq a^{-2} A \Psi^{\text{cont}}_{B,r}(w) \leq a^{-2} A \Psi^{\text{int}}_{B,r}(w). \tag{40}$$

The condition $n_{\min} > 1$ is checked on every positive recovered site sample size. Zero recruitment is unchanged by flooring. Appendix D gives comparisons that also cover sample allocations at most one.

D Budget-feasible integer sample allocation

Continuous optimization determines fractional site sample sizes. Integer recovery must preserve the enrollment budget while controlling the information lost by rounding. The minimum-sample-size bound in Appendix C applies when $n_{\min} > 1$. The following comparisons also account for arbitrarily small positive allocations. Fix one feasible continuous allocation n, finite $r > 0$, positive costs and privacy coefficients, and positive-semidefinite H_s. All inverses are on the unchanged effective subspace $\mathcal{U}$ containing w.

Proposition 1 (Observed-loss flooring comparison)**.** *Let $F = \lfloor n \rfloor$, $A_n = \overline{M}_\rho(n) + r^{-2} I_{\mathcal{U}}$, $D = \overline{M}_\rho(n) - \overline{M}_\rho(F)$, and $\eta_{\text{loss}} = \|A_n^{-1/2} D A_n^{-1/2}\|_{\text{op}}$. Then $0 \leq \eta_{\text{loss}} < 1$ in exact arithmetic, F is budget feasible, and*

$$\psi_r(F; w) \leq (1 - \eta_{\text{loss}})^{-1} \psi_r(n; w).$$

If this n has an A-factor continuous guarantee, then $\psi_r(F; w) \leq A(1 - \eta_{\text{loss}})^{-1} \Psi^{\text{int}}_{B,r}(w)$.

Proof. Since α_s is nondecreasing and $H_s \succeq 0$, $0 \preceq D \preceq \overline{M}_\rho(n)$. The ridge makes $A_n \succ 0$, and congruence gives

$$0 \preceq A_n^{-1/2} D A_n^{-1/2} \preceq I - r^{-2} A_n^{-1} \prec I.$$

Thus $\eta_{\text{loss}} < 1$ and $D \preceq \eta_{\text{loss}} A_n$. Hence $A_F = A_n - D \succeq (1 - \eta_{\text{loss}}) A_n$. Inverse Loewner order and evaluation at w prove the criterion inequality. Flooring proves budget feasibility. Finally use $\psi_r(n; w) \leq A \Psi^{\text{cont}}_{B,r}(w) \leq A \Psi^{\text{int}}_{B,r}(w)$. □

The bound controls the largest information loss across all directions. It can therefore exceed the observed ratio for a specific target w. A numerical certificate requires an error bound for the generalized eigensolve.

Proposition 2 (Support-cost ceiling recovery)**.** *Let $K = \{s : n_s > 0\}$ and $C_K = \sum_{s \in K} c_s$. If $C_K < B$, put $a = 1 - C_K/B$ and set $C_s = \lceil a n_s \rceil$ for $s \in K$, with $C_s = 0$ elsewhere. Then C is integer budget feasible and*

$$\alpha_s(C_s) \geq a^2 \alpha_s(n_s), \qquad \psi_r(C; w) \leq a^{-2} \psi_r(n; w).$$

Consequently a continuous A-factor guarantee gives $\psi_r(C; w) \leq A a^{-2} \Psi^{\text{int}}_{B,r}(w)$.

Proof. For $s \in K$, $\lceil an_s \rceil \le an_s + 1$, and therefore

$$\sum_s c_s C_s \le a \sum_s c_s n_s + C_K \le aB + C_K = B.$$

Also $C_s \ge an_s$. Monotonicity of $\alpha_s(x) = \min\{x, \kappa_s x^2\}$ and $0 < a \le 1$ imply

$$\alpha_s(C_s) \ge \min\{an_s, \kappa_s a^2 n_s^2\} \ge a^2 \min\{n_s, \kappa_s n_s^2\}.$$

Summing against $H_s \succeq 0$ gives $\overline{M}_\rho(C) \succeq a^2 \overline{M}_\rho(n)$. The ridge gives $A_C \succeq a^2 A_n$. Inverse order proves the criterion comparison, and $\Psi^{\mathrm{cont}}_{B,r}(w) \le \Psi^{\mathrm{int}}_{B,r}(w)$ completes the conditional composition. □

D.1 Conditional two-candidate hybrid

Corollary 1 (Conditional floor/ceiling hybrid)**.** *Let the available-candidate set be*

$$C(n) = \begin{cases} \{F, C\}, & C_K < B, \\ \{F\}, & C_K \ge B. \end{cases}$$

Choose in exact arithmetic

$$N^{\mathrm{hyb}} \in \arg\min_{N \in C(n)} \psi_r(N; w).$$

Every available candidate is integer budget feasible. When $C_K < B$,

$$\psi_r(N^{\mathrm{hyb}}; w) \le \min\{\psi_r(F; w), \psi_r(C; w)\}.$$

Otherwise $N^{\mathrm{hyb}} = F$ up to ties. If the same input has a valid flooring factor b_F and the support-cost condition holds, then

$$\psi_r(N^{\mathrm{hyb}}; w) \le \min\{b_F, a^{-2}\} \psi_r(n; w).$$

A continuous A-factor guarantee then gives $\psi_r(N^{\mathrm{hyb}}; w) \le A \min\{b_F, a^{-2}\} \Psi^{\mathrm{int}}_{B,r}(w)$.

Proof. The two propositions make the available candidates feasible and supply their criterion bounds. Exact minimization over them proves both displayed inequalities. The final composition applies the continuous approximation guarantee to the chosen integer candidate. □

When $C_K \ge B$, the hybrid uses only F and inherits only an available flooring comparison. If $n = 0$, both candidates are zero. If $w = 0$, all criteria are zero although empirical multiplicative ratios are undefined. As $C_K \uparrow B$, a^{-2} correctly diverges. Exact rational arithmetic or directed rounding plus a final integer-cost check is required near a ceiling boundary. With certified criterion intervals $[l_i, u_i]$, selecting candidate j gives only $\psi_r(N_j; w) \le \min_i \psi_r(N_i; w) + u_j - \min_i l_i$. A certified comparison accounts for objective-evaluation error.

D.2 Predefined integer-recovery diagnostics

The unit of analysis is one regime-appropriate continuous return for each of 115,200 scale inputs. The minimum-sample-size flooring condition $n_{\min} > 1$ holds for 99 returns (85 high-privacy and 14 mixed). Every strictly positive site sample size enters this test: the minimum positive site sample size has median 3.15×10^{-4}, 83,476 returns have $n_{\min} < 10^{-3}$, and 69 have $n_{\min} < 10^{-6}$. Flooring deletes a median of two sites and continuous expenditure 0.00149. Removing only these sites raises the floating criterion by more than 1% in 558 cases and by more than 5% in 165. These losses motivate the observed-loss comparison, which measures the affected information directions. Table 4 compares the observed ratios with the available bounds.

The numerical observed-loss factor is finite on all inputs. Exact-rational checks are reported below.

The minimum-sample-size flooring factor reaches 7266.98, and the support-cost factor reaches 2.49×10^6 near its budget boundary. The support-cost gate holds in 115,159 inputs. Ceiling improves on floor by more than 10^{-10} in 3,097, ties in seven, and is worse in 112,055. The median budget utilization is 96.875% for both. The hybrid therefore switches 3,097 times and retains floor in 112,103 cases. Among switches its relative improvement has median/95th percentile/maximum

Table 4: Deterministic integer post-processing on the predefined scale grid. Observed ratios and generalized-eigenvalue factors are computed in binary64. The symbol $\widehat{\eta}_{\mathrm{loss}}$ denotes the numerically evaluated loss factor. Ceiling and hybrid are evaluated on the same continuous returns as flooring.

Quantity	Applicable inputs	Median	95th percentile
Floor / continuous criterion	115,200	1.025	1.121
Observed-loss factor $(1-\widehat{\eta}_{\mathrm{loss}})^{-1}$	115,200	1.031	1.149
Minimum-sample-size flooring factor	99	2.257	50.935
Support-cost ceiling factor a^{-2}	115,159	1.378	5.482
Ceiling / continuous criterion	115,159	1.092	1.408
Hybrid / continuous criterion	115,200	1.024	1.120

2.864%/12.683%/22.103%. Its all-grid observed-ratio maximum is 1.38957. Thus conditional ceiling recovery broadens theoretical coverage but is usually inferior on these instances.

The floating-point allocations exceed the budget by at most 2.85×10^{-14} before inward rational scaling. The scaling restores exact budget feasibility without changing the flooring or applicable ceiling allocations. An $LDL^{\top}$-based verification of the rationalized inputs confirms all 115,200 flooring and hybrid comparisons and all 115,159 ceiling comparisons. The hybrid-factor median/95th-percentile/maximum values are 1.031/1.149/1.830, and the corresponding observed ratios are 1.024/1.120/1.390. Combining these integer comparisons with a continuous approximation factor gives the corresponding integer guarantee.

E Unknown baseline response functions

The symmetric Bernoulli model supports the information analysis. The following extension allows unknown site-specific baseline responses while retaining the same design inputs, equal-probability within-site randomization, and parameter ball $\Theta(r) \subseteq \Theta_0$. For each θ, let $\mathcal{M}_\eta(\theta)$ consist of measurable site-specific functions m_s such that, on each declared covariate support and for $a \in \{0,1\}$,

$$\eta/2 \le m_s(x) + (a - 1/2)x^{\top}\theta \le 1 - \eta/2.$$

The observed Bernoulli response probabilities are

$$p_a(x,s) = m_s(x) + (a - 1/2)x^{\top}\theta.$$

The functions m_s are unknown and are not supplied to the sample-allocation rule or release mechanism. Each m_s is the midpoint of the two response probabilities. The target effect remains $w^{\top}\theta$. Retain the stated source-to-target common-effect assumption and target-support restrictions.

Define the extended risk by

$$R^{\star}_{N,\mathrm{ext}} = \inf_{\Pi \in \mathcal{P}_N} \inf_{\widehat{\tau}} \sup_{\theta \in \Theta(r)} \sup_{m \in \mathcal{M}_\eta(\theta)} \mathbb{E}_{\theta,m,\Pi}(\widehat{\tau} - w^{\top}\theta)^2. \tag{41}$$

Protocol kernels cannot depend on the unknown θ or m. Privacy holds on the original declared record domain for the same complete-public-transcript class. The design and representation are fixed before trial outputs are observed.

Proposition 3 (Risk comparison with unknown baseline response functions)**.** *Under the remaining assumptions of Theorem 2, the extended model satisfies*

$$C_-^{-1}\psi_r(N;w) \le R^{\star}_{N,\mathrm{ext}} \le C_+\psi_r(N;w)$$

for every feasible integer design, with the same constants C_- and C_+ as in Theorem 2. The stated Gaussian estimator, with fixed design and representation, has a distribution that is independent of the admissible baseline functions.

Proof. Under equal-probability within-site randomization, for $D = (2A-1)(2Y-1)$,

$$\Pr_{\theta,m}(D = 1 \mid X = x, J = s) = \tfrac{1}{2}p_1(x,s) + \tfrac{1}{2}\{1 - p_0(x,s)\} = \tfrac{1}{2}(1 + x^{\top}\theta).$$

Since $D \in \{-1,1\}$, this identifies its entire conditional distribution, which is independent of m_s. The distribution of (X, DX), and therefore of the stated clipped scalar releases and their independent

Gaussian perturbations, agrees with that in the symmetric submodel. The Gaussian construction consequently has the same risk at each θ, uniformly over the nuisance class. Clipping controls sensitivity on the full declared record domain independently of the baseline functions. For the lower bound, $m_s \equiv 1/2$ is admissible for every $\theta \in \Theta(r)$, so restricting the extended supremum to this symmetric submodel gives $R^{\star}_{N,\mathrm{ext}} \geq R^{\star}_{N}$. Apply Theorem 2. The extension follows from the distribution of the released statistic and the symmetric submodel lower bound. □

Singular H_s retain the range and Moore-Penrose conventions. Sites with zero sample size or zero representation contribute no release. A zero target admits the identically zero estimator, including when the effective span is zero. These cases require no normalized direction for a zero block.

F Sensitivity to site moment approximation

Finite-radius criterion under a matrix split. Approximating site moments changes the information criterion. We first bound this change at fixed sample sizes, then derive a uniform bound over the budget set.

Proposition 4 (Fixed-design finite-radius perturbation)**.** *Let $H_s, \widetilde{H}_s \succeq 0$, $r > 0$, and $\alpha_s(n_s) = n_s \min\{1, \kappa_s n_s\}$ with $\kappa_s > 0$. On the full feature space put $M_H(n) = \sum_s \alpha_s(n_s) H_s$, $M_{\widetilde{H}}(n) = \sum_s \alpha_s(n_s)\widetilde{H}_s$ and $\bar{\psi}_H(n; w) = w^\top [M_H(n) + r^{-2} I]^{-1} w$, with the analogous $\bar{\psi}_{\widetilde{H}}$. Then*

$$\left|\bar{\psi}_H(n; w) - \bar{\psi}_{\widetilde{H}}(n; w)\right| \leq r^4 \|w\|_2^2 \left\| \sum_s \alpha_s(n_s)(H_s - \widetilde{H}_s) \right\|_{\mathrm{op}} \leq r^4 \|w\|_2^2 \sum_s \alpha_s(n_s) \|H_s - \widetilde{H}_s\|_{\mathrm{op}}. \tag{42}$$

This full-space comparison allows different ranks and effective subspaces. Define

$$\widetilde{\mathcal{U}} = \mathrm{span}\{\mathrm{range}(\widetilde{H}_s) : s \in [S]\}.$$

If w belongs to both effective spans and each collection's moments are supported on its own span, its full-space expression equals the paper's inverse-on-effective-span criterion for that collection. If $w \notin \widetilde{\mathcal{U}}$, the full-space expression remains defined, but projecting w onto $\widetilde{\mathcal{U}}$ changes the target. Theorem 2 applies to a truncated instance when its effective span contains w.

Proof. Write $A = M_H(n) + r^{-2} I$ and $\widetilde{A} = M_{\widetilde{H}}(n) + r^{-2} I$. Both are positive definite and $A, \widetilde{A} \succeq r^{-2} I$, hence $\|A^{-1}\|_{\mathrm{op}}, \|\widetilde{A}^{-1}\|_{\mathrm{op}} \leq r^2$. The resolvent identity $A^{-1} - \widetilde{A}^{-1} = A^{-1}(\widetilde{A} - A)\widetilde{A}^{-1}$ and $|w^\top C w| \leq \|w\|_2^2 \|C\|_{\mathrm{op}}$ prove the first inequality. The triangle inequality proves the second. On an effective span $\mathcal{U}$, M_H has no cross block with $\mathcal{U}^\perp$. In the decomposition $\mathcal{U} \oplus \mathcal{U}^\perp$, A is block diagonal with blocks $M_H|_{\mathcal{U}} + r^{-2} I_{\mathcal{U}}$ and $r^{-2} I_{\mathcal{U}^\perp}$. For $w \in \mathcal{U}$ its quadratic form is therefore exactly the restricted inverse. The same argument applies separately to $\widetilde{\mathcal{U}}$. □

Optimal values and conditional factor transfer. For a common continuous or integer feasible set $\mathcal{F} = \{n \geq 0 : \sum_s c_s n_s \leq B\}$ (or its integer subset), with $c_s, B > 0$, let $v_H = \inf_{n \in \mathcal{F}} \bar{\psi}_H(n; w)$ and similarly $v_{\widetilde{H}}$. Since $\alpha_s(n_s) \leq n_s$, Equation (42) holds uniformly with

$$E_B = r^4 \|w\|_2^2 B \max_s \frac{\|H_s - \widetilde{H}_s\|_{\mathrm{op}}}{c_s}. \tag{43}$$

Thus $|v_H - v_{\widetilde{H}}| \leq E_B$. Indeed, every fixed-design criterion differs by at most E_B, and taking infima preserves both inequalities. If an allocation $n_{\mathrm{alg}} \in \mathcal{F}$ has a truncated-instance factor, for $A \geq 1$, $\bar{\psi}_{\widetilde{H}}(n_{\mathrm{alg}}; w) \leq A v_{\widetilde{H}}$, then

$$\bar{\psi}_H(n_{\mathrm{alg}}; w) \leq A v_H + (A + 1) E_B. \tag{44}$$

For $w \neq 0$, a positive common lower bound may be supplied by

$$v_H \geq \ell_H := \frac{\|w\|_2^2}{r^{-2} + B \max_s \|H_s\|_{\mathrm{op}} / c_s} > 0, \tag{45}$$

because $\|M_H(n)\|_{\mathrm{op}} \leq B \max_s \|H_s\|_{\mathrm{op}} / c_s$ for every feasible n. Using this lower bound converts Equation (44) to the multiplicative factor $A + (A+1) E_B / \ell_H$ for the empirical criterion. When $w = 0$, both criteria and optima are zero, and this relative expression is unnecessary.

Numerical sensitivity on eICU designs. The deterministic calculation uses the empirical moment matrices and all 8,640 integer method-design combinations, weighting each site difference by that design's $\alpha_s(N_s)$. For heterogeneous privacy, only the public upper range for ρ_s is available, so the corresponding weighted perturbation bounds use the available privacy upper limits. Across all combinations, the numerical right-hand side of Equation (42) has median 3.87×10^{-15}, 95th percentile 7.26×10^{-13}, and maximum 2.49×10^{-12}. On homogeneous-policy configurations the directly weighted operator difference is at most 2.58×10^{-11}, the bound is at most 2.04×10^{-12}, and direct criteria differ by at most 2.78×10^{-17}. Among nonzero displayed differences the bound is much looser than the direct discrepancy (median ratio 3.79×10^{4}), but its largest ratio to the empirical criterion over all checked designs is 9.72×10^{-8}. Numerical ranks agree at 4, 8 and 17 and target projection residuals are of order 10^{-16}.

These binary64 diagnostics assess criterion sensitivity. Estimator risk uses the direct identities in Appendix I.2.

G Synthetic design generation and numerical checks

G.1 Allocation reversal

Proposition 5 (Privacy-induced allocation reversal)**.** *Let $d = 1, w = 1$, $H_1 = 1, H_2 = 1/4$, equal unit costs, integer $B \geq 4$, fixed $0 < r_0 < 1/2$, and $\kappa_{1,B} = B^{-3}, \kappa_{2,B} = B^{-1}$. The classical-criterion optimum is $(B, 0)$, the private-criterion optimum is $(0, B)$, and*

$$\frac{\psi_{r_0}((B,0);1)}{\psi_{r_0}((0,B);1)} = \frac{B/4 + r_0^{-2}}{B^{-1} + r_0^{-2}} \sim \frac{r_0^2}{4} B \to \infty. \tag{46}$$

Proof. On $n_1 + n_2 = B$, sampling information is $n_1 + n_2/4$ and is maximized at $n_1 = B$. Private information is $n_1^2/B^3 + n_2^2/(4B)$ throughout the feasible region. Its convex quadratic reaches its maximum at an endpoint. $B/4 > B^{-1}$ for $B > 2$, so the private-criterion optimum is $n_2 = B$. The criterion is the reciprocal of information plus r_0^{-2}, giving the displayed ratio and asymptotic. Both endpoints are integer feasible. □

Corollary 2 (Minimax-risk consequence of Proposition 5)**.** *Realize the construction with $Q_1 = \delta_1$, $Q_2 = \delta_{1/2}$ and $Q_T = \delta_1$. Let $N^{\mathrm{cl}} = (B, 0)$ and $N^{\mathrm{ref}} = (0, B)$. For fixed $0 < r_0 < 1/2$ and integer $B \geq 4$, with all risks referring to this budget-indexed instance,*

$$\frac{R^\star_{N^{\mathrm{cl}}}}{\mathfrak{R}^\star_{B,r_0}(1)} \geq \frac{1}{6\pi^2} \frac{B/4 + r_0^{-2}}{B^{-1} + r_0^{-2}} \longrightarrow \infty. \tag{47}$$

The constants are uniform in B.

Proof. The point-mass model gives $H_1 = 1$, $H_2 = 1/4$ and $w = 1$. Since $r_0 < 1/2$, $\Theta(r_0) \subseteq \Theta_0$ with $\eta = 1/2$, and both sites have whitened leverage one. Moreover $\rho_{s,B} = \frac{1}{2}\log(1 + \kappa_{s,B}) \leq \bar{\rho} = \frac{1}{2}\log(5/4)$ for $B \geq 4$. Thus Theorem 2 applies uniformly with $C_- = \pi^2$ and $C_+ = 6$. Since N^{ref} is feasible,

$$\frac{R^\star_{N^{\mathrm{cl}}}}{\mathfrak{R}^\star_{B,r_0}(1)} \geq \frac{R^\star_{N^{\mathrm{cl}}}}{R^\star_{N^{\mathrm{ref}}}} \geq \frac{\psi_{r_0}(N^{\mathrm{cl}};1)}{6\pi^2 \psi_{r_0}(N^{\mathrm{ref}};1)}.$$

Equation (46) gives (47) and its divergence. □

G.2 Optimization quality

The small-instance reference grid uses 20 randomization replicates, $S \in \{2, 3\}$, $d \in \{1, 2, 3, 5\}$, four constructions of H_s, two cost modes, three privacy-input groups, $B \in \{8, 16, 32, 64\}$, and $r \in \{0.25, 0.50, 0.90\}$, yielding 46,080 instances, with six evaluation procedures considered for each instance. The scale grid replaces S by $\{5, 10, 20, 50, 100\}$ and omits the global numerical reference procedures, yielding 115,200 instances with four applicable methods per instance.

The four constructions of H_s are aligned, complementary, redundant, and heterogeneous-rank. Aligned sites use $H_s = (0.80 + 0.20s/S)vv^\top$, where v is a normalized standard Gaussian vector. Complementary sites use the same scale and cycle through a random orthonormal basis. Redundant sites use scale $0.84 + 0.12s/S$, with the first two sites sharing one basis direction and the remaining sites sharing a second direction. When $S = 2$ or $d = 1$, all sites share one direction. Heterogeneous-rank sites have independent random orthonormal bases and rank $k_s = \min(d, s)$. Their positive eigenvalues decrease linearly from 1 to 0.35, with eigenvalue 1 when $k_s = 1$. Random bases are obtained by orthonormalizing standard Gaussian matrices.

For heterogeneous costs, we draw costs from Uniform$(0.7, 1.4)$. Privacy is generated through exact rational values of $\kappa_s B/c_s$, drawn uniformly on a grid with spacing 0.001: $[0.15, 0.85]$ for globally high-privacy inputs, $[1.5, 4.0]$ for threshold-reachable inputs, and alternating values from $[0.25, 0.75]$ and $[1.4, 3.2]$ for mixed-threshold inputs. In the threshold-reachable group, each site can enter its sampling branch if assigned enough of the budget. Group membership is determined by the inputs. The target combines all d_{eff} effective orthonormal coordinates with weights linearly decreasing from 1 to 0.55, then is normalized to norm 0.75.

All 46,080 numerical intervals meet the stopping rule in Appendix G.2.1, with maximum absolute gap 4.00×10^{-4}. Figure 3 separates continuous optimization, integer recovery, and scalability. Conic runtime increases with site count, while the active-site fraction decreases after compression. This connects the construction's computational cost to its use of sparse site allocations.

Timing scope. Reported conic runtimes include optimization, compression, sample-size recovery, and criterion evaluation. Integer processing is timed separately, and instance generation and I/O are excluded. The implementation uses CPython 3.14.6, NumPy 2.5.2, CVXPY 1.9.2, Clarabel 0.11.1, and macOS 26.4 arm64.

G.2.1 Numerical reference method

The continuous reference uses monotonic branch-and-bound. On a box $[l, u]$, monotonicity gives $\psi_r(n; w) \geq \psi_r(u; w)$, even when u exceeds the budget. Boxes with $c^\top l > B + 10^{-14}$ are removed. The longest scaled side is bisected and a feasible incumbent updated. Bounds of tolerance-pruned boxes remain in the global lower value ℓ. With incumbent U, the stopping rule is

$$\max\{0, U - \ell\} \leq 2 \times 10^{-4} + 2 \times 10^{-4} \max\{1, |U|\},$$

with at most 250,000 child-box evaluations. The incumbent U may be improved with feasible Sampling or conic returns. Calculations use binary64 linear algebra and rank tolerance 10^{-10} without outward rounding.

For exact budget-feasibility checks in the enumeration, heterogeneous costs are represented to six decimal places as $c_s = m_s/10^6$. For $S = 2, 3$, feasibility is evaluated as $\sum_s m_s N_s \leq 10^6 B$ using integer arithmetic, with floating-point objective comparisons and ties resolved by cost and enumeration order. Recomputed minima agree within $2 \times 10^{-10} \max\{1, |\psi_r|\}$. Figure 2 reports each conic method and recovery separately: 15,360 valid instances per group and candidate, excluding reference self-ratios. Quantiles use linear interpolation (Hyndman-Fan type 7). Both reference procedures cover all 46,080 small instances.

H Estimator calibration and weights at fixed designs

The minimax design criterion and the Gaussian-release bound can order designs differently. We compare their rankings with estimator-risk rankings on an analytic grid and check the risk calculation by paired simulation. A separate fixed-allocation study holds site sample sizes fixed to isolate the effect of estimator weights.

Estimator-ranking protocol and calibration. The analytic grid uses 20 randomization replicates, $d \in \{2, 5, 10\}$, $S \in \{5, 10, 20\}$, five constructions of H_s, two cost regimes, three privacy regimes, $B \in \{50, 200, 800\}$, and radii $0.20/K_X$ or $0.70/K_X$. The five constructions are aligned, complementary, redundant, distractor, and heterogeneous-rank. Aligned sites share one random unit direction. Complementary sites cycle through permuted coordinate directions. Redundant sites duplicate a

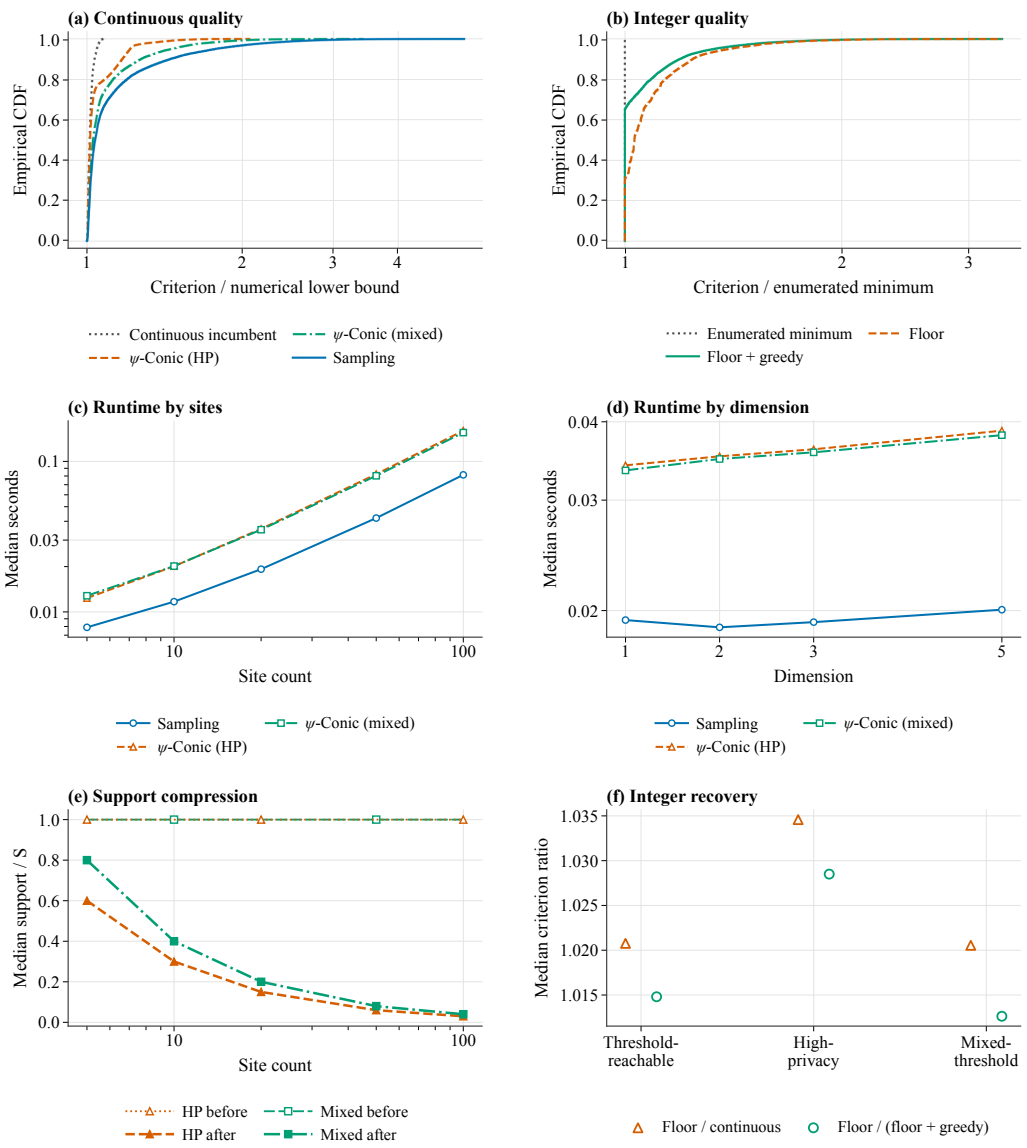


Figure 3: Criterion quality and scale behavior. (a) and (b) compare continuous candidates with numerical lower bounds and integer candidates with exhaustive floating-point minima. (c) and (d) report scale-grid runtime by site count and dimension. Appendix G.2.1 specifies the floating-point reference procedure. (e) reports support before and after compression, and (f) reports integer recovery. Floor + greedy denotes residual-budget filling after flooring, separate from the floor/ceiling hybrid. The continuous incumbent is a feasible reference design, and the enumerated minimum is the integer reference. HP and mixed denote the two conic constructions. Panel (f) groups inputs by privacy regime, as in Figure 2.

small random direction pool. Distractor sites reserve signal coordinate directions and draw remaining directions with signal coordinates downweighted by 0.08. Heterogeneous-rank sites use random orthogonal bases, ranks cycling through $1, \ldots, d$, and positive eigenvalues decreasing linearly from 1 to 0.25. Methods are Uniform, Classical c-optimal design, Sampling, and the high-privacy (HP) and mixed conic constructions. HP is evaluated where applicable. Each applicable allocation is evaluated with CB and GB. For each allocation, both representations use the same $15 + 2d$ parameter settings comprising zero, signed target, effective-coordinate and bias directions on the radius, eight random effective-sphere directions, and independent interior directions at $r/4$ and $r/2$. The $2d$ signed coordinate settings cycle through an orthonormal basis of $\mathcal{U}$ when $d_{\text{eff}} < d$. Random directions are

shared across candidates. Each allocation uses its CB bias direction for both representations. Draws are normalized in the effective span and checked on the complete source and target supports.

The simulated subset was fixed before outcomes and contains 1,080 problems. The five evaluation parameters are 0, $\pm rw/\|w\|$, ru_1, and ru^{rand}, where u_1 is the first effective-basis vector and u^{rand} is a random unit direction in that span. There are 2,000 paired repeats. Across 46,800 applicable method-parameter evaluations, the Bonferroni-adjusted Student-t family contains 140,400 expectation, MSE, and privacy-variance diagnostics. Their maximum absolute standardized discrepancies are 3.8903, 4.3094, and 5.0722, below the predefined threshold 5.1082. The Student-t intervals have approximate nominal coverage conditional on the fixed inputs and selected designs. The ordinary 95% interval contains the analytic value in 44,558, 44,430, and 44,534 of 46,800 evaluations.

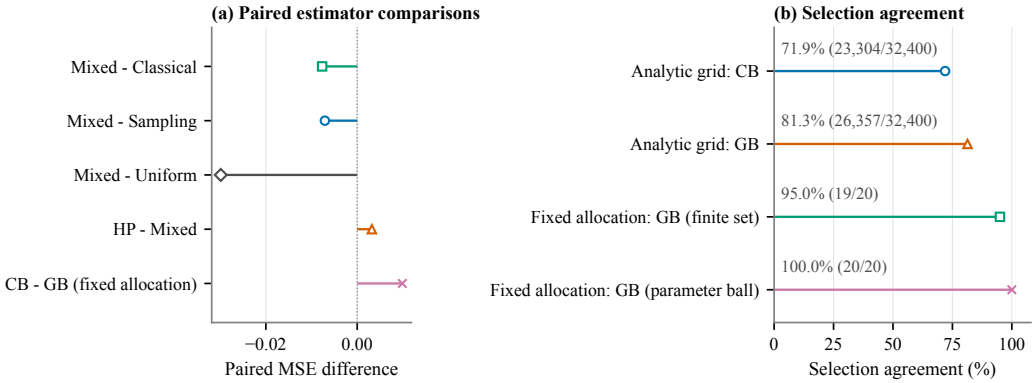


Figure 4: Estimator comparisons. (a) Mean paired MSE differences over 5,400 problem and parameter combinations, except HP versus Mixed, which uses 1,800 applicable combinations. Negative favors the first named method. Allocation comparisons use GB. CB versus GB holds mixed sample allocations fixed. (b) Agreement between the criterion-selected and minimum-risk methods. The analytic grid uses R_G^{wc}. The fixed-allocation study uses maximum empirical MSE over the prespecified parameters (finite set) or R_G^{wc} (parameter ball). HP and Mixed denote the high-privacy and mixed conic constructions.

Figure 4(b) reports selection agreement. The analytic-grid mean Spearman correlation rises from 0.6416 under CB to 0.8898 under GB, while 1,747 selection disagreements in the mixed-privacy group remain. In Figure 4(a), the high-privacy minus mixed mean MSE difference is 0.003214 across 360 applicable problems and five parameters per problem. The mixed construction has lower MSE on average. GB improves rank agreement while this allocation comparison remains unfavorable to the high-privacy construction.

Fixed-allocation comparison. Twenty $d = 2, S = 3$ problems evaluate Uniform, Sampling, ψ-Conic, and ψ-Reference allocations under both CB and GB. The prespecified set of 13 parameter values contains zero, signed target and ψ-Reference CB-bias directions, signed directions along two effective coordinates, and signed interior draws at $r/4$ and $r/2$. The 4,000 paired Monte Carlo replications per parameter value share sample allocations, parameters, and random draws across representations. The additional 32 controlled problems compare ψ-Conic and ψ-Reference allocations with both representations using $17 + 2d_{\text{eff}}$ treatment effect parameter settings, including four random effective-span directions and two interior radii.

Under CB weighting, the Spearman correlation between ψ_r and maximum empirical MSE over the 13 prespecified parameters is negative in all ten high-privacy instances. Figure 4(b) shows GB selection agreement in 19/20 problems for maximum empirical MSE over the prespecified finite parameter set. The only remaining disagreement under GB occurs for instance P10, where the corresponding Spearman correlation is 0.4. Analytic risk on the parameter ball gives agreement in 20/20. Across the 52 controlled problems, conic-to-integer-ψ_r criterion ratios range from 1 to 1.7045, while CB-to-GB R_G^{wc} ratios at the same conic allocations range from 1.0330 to 2.5683. At fixed sample sizes, GB minimizes the Gaussian risk upper bound in Equation (18).

I Semi-synthetic eICU covariate study: cohort, estimator, and inference

I.1 Cohort and design conditions

The eICU-CRD v2.0 hospital, patient, and periodic-vital tables supply the covariates (Pollard et al., 2018; 2019). We select one unit stay per admission, apply adult eligibility, then exclude hospitals with fewer than 200 eligible stays. Of 165,795 eligible stays, 162,282 from 135,985 patients at 158 hospitals remain. Repeated admissions are retained.

Features comprise age, admission anthropometrics and categories, and the minimum, median, and maximum of seven vital signs within 1,440 minutes. Implausible values become missing. We exclude variables above 30% pooled missingness, impute pooled medians, add missingness indicators, one-hot encode, center, and scale. The $d-1$ highest-variance columns are normalized by $z \leftarrow z/\max(1, \|z\|)$ and augmented as $x = (1, z)/\sqrt{2}$. Thus $\|x\| \le 1$ for $d \in \{5, 10, 20\}$, with effective ranks 4, 8, and 17. Let m_s be the number of eligible stays at hospital s. Each Q_s is uniform on its empirical covariates, and $Q_T = \sum_s m_s Q_s / \sum_t m_t$. Treatment and outcome data are simulated.

The 108 conditions cross three cost policies, three privacy policies, $\rho_0 \in \{0.005, 0.02, 0.1, 0.5\}$, and three treatment effect directions. Costs are proportional to 1, $1/m_s$, or $\exp(0.5 Z_s^{\text{cost}})$, then normalized to mean one. Privacy is homogeneous, independently heterogeneous, or size-correlated. For the heterogeneous policies, define

$$u_s = \frac{a_s - \bar{a}}{\max_t |a_t - \bar{a}|}, \qquad \bar{a} = S^{-1} \sum_t a_t, \qquad \rho_s = \rho_0(1 + 0.5 u_s),$$

with $u_s = 0$ if the denominator vanishes, followed by normalization of ρ_s to mean ρ_0. Size-correlated privacy uses $a_s = \log m_s$, and independent heterogeneity uses $a_s = \tanh Z_s^{\text{privacy}}$. All Z_s^{cost} and Z_s^{privacy} are independent standard Gaussian draws.

The 20 parameter combinations are fixed before outcomes. They use the grid $d \in \{5, 10, 20\}$, $B \in \{500, 1000, 2500, 5000\}$ and $r \in \{0.20/K_X, 0.70/K_X\}$. We omit $(d, B, rK_X) = (5, 1000, 0.20)$, $(10, 1000, 0.70)$, $(10, 2500, 0.20)$ and $(20, 2500, 0.70)$.

Let $\mu_s = \mathbb{E}_{Q_s} X$ and $H_{\text{pool}} = \sum_s m_s H_s / \sum_t m_t$. The target-mean direction is the normalized projection of w onto $\mathcal{U}$. The leading pooled second-moment direction is the normalized projection of w onto the top eigenspace of H_{pool}. If this projection vanishes, we normalize the first nonzero projection of a coordinate axis onto that eigenspace. The between-hospital direction is proportional to

$$P_{\mathcal{U}} \sum_s \omega_s(\mu_s - w), \qquad \omega_s = \frac{\xi_s - \bar{\xi}}{\sqrt{\sum_t (\xi_t - \bar{\xi})^2}}, \qquad \xi_s \overset{\text{iid}}{\sim} \mathcal{N}(0, 1),$$

where $\bar{\xi} = S^{-1} \sum_s \xi_s$. We retain nonzero draws with at least two nonzero projected hospital contributions. Each direction is normalized to a unit vector v and oriented to have $w^\top v \ge 0$. If $w^\top v = 0$, its largest absolute coordinate is made positive. The treatment effect parameter is

$$\theta = \gamma v, \qquad \gamma = \min\left\{ r, \frac{0.8 - 10^{-10}}{\max_x |x^\top v|} \right\}.$$

The maximum is over the pooled empirical covariate support. The trial then samples covariates with replacement from each empirical support, $A \sim \text{Bernoulli}(1/2)$, and Y from Equation (1).

All methods use componentwise flooring without residual-budget reallocation and then the implemented GB representation. ψ-Conic uses the mixed construction. The other sample-allocation rules are defined in Section 6.

I.2 Matrix split, scalar release, and representation residual

The release operates on trial records sampled from the fixed empirical supports described above. Design optimization uses truncated moment factors, while estimator evaluation uses empirical moments. We specify both matrices, derive the release sensitivity, and retain the resulting representation residual in the risk calculation. Privacy accounting conditions on these empirical moment matrices and

covariate supports. Constructing or releasing them from protected source records requires additional accounting.

If $H_s^{\mathrm{emp}} = U_s \operatorname{diag}(\lambda_{sj}) U_s^\top$, numerical factorization uses

$$\tau_s = 10^{-8} \max\{1, \max_j |\lambda_{sj}|\}, \qquad \widetilde{H}_s = U_s \operatorname{diag}(\lambda_{sj} \mathbf{1}\{\lambda_{sj} > \tau_s\}) U_s^\top, \tag{48}$$

with $G_s^{\mathrm{tr}} = \widetilde{H}_s^{1/2}$ and $J_s^{\mathrm{tr}} = \widetilde{H}_s^{\dagger/2}$. Eigenvalues below $-\tau_s$ are rejected. Design optimization, compression, and recovery use G_s^{tr}. The returned-design criterion and GB system matrix use H_s^{emp}. The representation multiplies the resulting q by G_s^{tr}, while release sensitivity uses J_s^{tr} on the empirical support. The analytic risk approximation assumes zero representation residual. Below we quantify that residual and give the direct risk identities.

The implementation computes $L_s = \max_{x \in \mathrm{supp}(Q_s)} \|J_s^{\mathrm{tr}} x\|_2$ on every covariate observation and uses $L = \max_s L_s$. The unclipped scalar release operates on the fixed simulated-trial record domain $\mathcal{R}_s^{\mathrm{trial}} = \mathrm{supp}(Q_s) \times \{0,1\} \times \{0,1\}$.

For nominal dimensions $d = 5, 10, 20$, the effective ranks are 4, 8, 17 and the common public leverage bounds are $L = 61.9, 68.5, 68.5$, respectively. Each uses all 158 hospital supports and 162,282 eligible stays.

The information-envelope threshold is $1/\kappa_s$. Sampling and privacy variance balance at $2L^2/\rho_s$. These differ by a factor $2L^2\kappa_s/\rho_s$. The consistent-matrix comparison in Proposition 6 is evaluated on the synthetic instances in Appendix J.

In an adjacent-data comparison, H_s^{emp}, $\widetilde{H}_s$, J_s^{tr}, L, ρ_s, N_s, and z_s are held fixed. Two size-N_s datasets replace one (X, A, Y) record within $\mathcal{R}_s^{\mathrm{trial}}$. For $D_{si} = (2A_{si} - 1)(2Y_{si} - 1)$, $V_{si} = D_{si} X_{si}$ and, when $z_s \neq 0$, $u_s = z_s / \|z_s\|_2$, let $T_{si} = u_s^\top J_s^{\mathrm{tr}} V_{si}$. Since $|T_{si}| \leq L$ on this domain, replacement sensitivity of the site mean is at most $2L/N_s$. The release and estimator are

$$\widetilde{T}_s = N_s^{-1} \sum_{i=1}^{N_s} T_{si} + Z_s, \quad Z_s \sim \mathcal{N}\left(0, \frac{2L^2}{\rho_s N_s^2}\right), \qquad \widehat{\tau}_z = \sum_{s: z_s \neq 0} \|z_s\|_2 \widetilde{T}_s. \tag{49}$$

Thus the executed support-domain release is ρ_s-zCDP and has privacy variance $2L^2\|z_s\|_2^2/(\rho_s N_s^2)$. For a larger record domain, the clipped extension

$$T_{si}^{\mathrm{clip}} = \mathrm{clip}_{[-L,L]}(u_s^\top J_s^{\mathrm{tr}} V_{si})$$

has the same sensitivity bound and agrees with T_{si} on the stated empirical support. The simulations use the unclipped support-domain release. Under the simulated outcome law, $\mathbb{E}[V_{si}] = H_s^{\mathrm{emp}} \theta$ and $\mathbb{E}[V_{si} V_{si}^\top] = H_s^{\mathrm{emp}}$.

For active sites define

$$\begin{aligned} \beta_s &= N_s^{-1} + \frac{2L^2}{\rho_s N_s^2}, \quad A_N^{\mathrm{emp}} = r^{-2} I + \sum_{s: N_s > 0} \beta_s^{-1} H_s^{\mathrm{emp}}, \\ q &= (A_N^{\mathrm{emp}})^{-1} w, \qquad z_s = \beta_s^{-1} G_s^{\mathrm{tr}} q, \qquad b = r^{-2} q. \end{aligned} \tag{50}$$

Inactive sites have $\beta_s = +\infty$ and $z_s = 0$. Their scalar contribution is defined as zero. The matrix split leaves residual

$$\delta = b + \sum_s G_s^{\mathrm{tr}} z_s - w = \sum_{s: N_s > 0} \beta_s^{-1} (\widetilde{H}_s - H_s^{\mathrm{emp}}) q. \tag{51}$$

This quantity measures the representation error caused by the matrix split.

Fixed-parameter and worst-case risk. Condition on the selected design and empirical covariate supports and moments, holding $(H_s^{\mathrm{emp}}, J_s^{\mathrm{tr}}, L, N, z)$ fixed. For $N_s > 0$ put

$$v_s = (J_s^{\mathrm{tr}})^\top z_s, \qquad h_s = H_s^{\mathrm{emp}} v_s, \qquad b_{\mathrm{act}} = \sum_{s: N_s > 0} h_s - w,$$

$$C_0 = \sum_{s: N_s > 0} \left\{ \frac{v_s^\top H_s^{\mathrm{emp}} v_s}{N_s} + \frac{2L^2 \|z_s\|_2^2}{\rho_s N_s^2} \right\}, \quad K = b_{\mathrm{act}} b_{\mathrm{act}}^\top - \sum_{s: N_s > 0} \frac{h_s h_s^\top}{N_s}. \tag{52}$$

On the empirical support, where the clipped extension leaves the statistic unchanged, the implemented mechanism satisfies

$$\mathrm{MSE}_\theta(\widehat{\tau}_z) = C_0 + \theta^\top K \theta. \tag{53}$$

Indeed, one site's unnoised contribution has mean $h_s^\top \theta$ and variance $\{v_s^\top H_s^{\mathrm{emp}} v_s - (h_s^\top \theta)^2\}/N_s$. Independence across sites and of the Gaussian noises, followed by adding the squared bias $(b_{\mathrm{act}}^\top \theta)^2$, proves Equation (53). If Q has orthonormal columns spanning $\mathcal{U}$, maximizing this quadratic over $\Theta(r)$ gives

$$R_G^{\mathrm{wc}}(N, z) = C_0 + r^2 \max\{0, \lambda_{\max}(Q^\top K Q)\}. \tag{54}$$

For $\mathcal{U} = \{0\}$ the value is C_0. The maximum with zero is needed because the parameter ball contains $\theta = 0$.

In the ideal common spectral factorization, $h_s = G_s^{\mathrm{tr}} z_s$, $v_s^\top H_s^{\mathrm{emp}} v_s = \|z_s\|^2$, and $b_{\mathrm{act}} = \delta - b$. The analytic approximation further sets $\delta = 0$. For separately formed floating matrices, the direct definitions above retain

$$b_{\mathrm{act}} = (\delta - b) + \sum_s \{H_s^{\mathrm{emp}} (J_s^{\mathrm{tr}})^\top - G_s^{\mathrm{tr}}\} z_s.$$

Across 8,640 method-design combinations, $r\|\delta\| \le 1.72 \times 10^{-15}$ and the squared-bias correction is at most 4.66×10^{-16}. Criterion sensitivity is analyzed in Appendix F.

I.3 Pairing and inference

For randomization replicate ℓ, selected parameter combination k, condition c, Monte Carlo replicate j, and comparator h, let $E^h_{\ell k c j}$ denote its squared estimation error. Inference uses paired differences

$$D^h_{\ell c j} = \frac{1}{20} \sum_{k=1}^{20} (E^{\psi\text{-Conic}}_{\ell k c j} - E^h_{\ell k c j}), \qquad D^h_{\mathrm{pool},cj} = \frac{1}{3} \sum_{\ell=0}^{2} D^h_{\ell c j},$$
$$\widehat{\Delta}^h_c = \frac{1}{1000} \sum_{j=1}^{1000} D^h_{cj}, \qquad \mathrm{MCSE}(\widehat{\Delta}^h_c) = \frac{\widehat{\sigma}^h_c}{\sqrt{1000}}. \tag{55}$$

Here $\widehat{\sigma}^h_c$ is the sample standard deviation of the 1,000 paired differences, with denominator 999. For view $v \in \{0, 1, 2, \mathrm{pool}\}$, set $D^h_{cj} = D^h_{vcj}$. Bonferroni-adjusted Student-t critical values are 3.79857 for the 324 main comparisons and 4.13437 for the 1,296 supplemental comparisons, with 999 degrees of freedom. Intervals have approximate nominal coverage conditional on the fixed empirical covariate distributions, moment matrices, and designs. Set $\tau^h_c = 10^{-10} \max\{1, \widehat{\mathrm{MSE}}^h_c\}$, where $\widehat{\mathrm{MSE}}^h_c$ is the comparator's estimated MSE. A win requires the corrected interval's upper endpoint below $-\tau^h_c$ and the estimated MSE ratio of ψ-Conic to the comparator below $1 - 10^{-10}$. A loss requires its lower endpoint above τ^h_c and the ratio above $1 + 10^{-10}$. Other cases are inconclusive. Summary means average the 108 conditions. Summary MCSEs are the sample standard deviations of $108^{-1} \sum_c D^h_{cj}$ across the 1,000 Monte Carlo replicates, divided by $\sqrt{1000}$. Summary interval endpoints average conditionwise simultaneous endpoints.

Against Uniform, the primary-replicate interval is $[-0.0603, -0.0567]$. The corresponding intervals are $[-0.0118, -0.0100]$ for Classical and $[-0.0115, -0.00982]$ for Sampling.

I.4 Sensitivity to design randomization

Sensitivity is assessed with two additional randomization replicates that vary the random cost draws, heterogeneous privacy draws, and random between-hospital treatment effect direction while holding all deterministic inputs fixed. Across the three randomization replicates, the descriptive standard deviations of the grand mean MSE differences are 1.7893×10^{-4}, 9.3156×10^{-5}, and 1.1108×10^{-4} relative to Uniform, Classical c-optimal design, and Sampling, respectively. The sign of the estimated MSE difference changes in 1, 10, and 17 of the 108 conditions, respectively. The Bonferroni-adjusted W/U/L classification changes in 0, 15, and 23 conditions. Thus the aggregate means are stable across randomizations, while condition-level conclusions remain heterogeneous.

Table 5: Pooled sensitivity analysis across three randomization replicates. Intervals are formed by averaging conditionwise endpoints from the 1,296-comparison Bonferroni-adjusted Student-t procedure. Differences are ψ-Conic minus comparator. Negative favors ψ-Conic.

Comparator	Mean difference	MCSE	Approx. nominal 95% interval	W/U/L
Uniform	−0.0583	2.78×10^{-5}	[−0.0594, −0.0572]	87/1/20
Classical c-optimal	−0.0108	1.94×10^{-5}	[−0.0114, −0.0103]	89/5/14
Sampling	−0.0106	1.94×10^{-5}	[−0.0111, −0.0101]	45/5/58

Against Sampling, 48 negative and 60 positive pooled estimates contribute −0.0116 and +0.000963 to the equally weighted mean, respectively. Large gains therefore coexist with more unfavorable conditions.

I.5 Bias and variance decomposition and unfavorable conditions

For the 108 equally weighted conditions let $\Delta_i = \mathrm{MSE}_{\psi\text{-Conic},i} - \mathrm{MSE}_{\mathrm{Sampling},i}$. Its grand mean decomposes as

$$\overline{\Delta} = 108^{-1} \sum_i \min(\Delta_i, 0) + 108^{-1} \sum_i \max(\Delta_i, 0).$$

These signed gain and loss contributions use point estimates. W/U/L uses the corrected-interval and MSE-ratio rule above. Conditionwise relative improvement $-\Delta_i/\mathrm{MSE}_{\mathrm{Sampling},i}$ and the aggregate ratio $-\sum_i \Delta_i / \sum_i \mathrm{MSE}_{\mathrm{Sampling},i}$ have different weighting. No comparator MSE is at or below the exclusion threshold 10^{-10}.

Table 2 reports the cost-stratified analytic components and Monte Carlo means. The Monte Carlo means include simulation error.

Under equal costs, bias reduction exceeds the added privacy variance. Under inverse-size costs, the added privacy variance exceeds the bias reduction. Equal-cost conditions contribute 98.4% of the signed net improvement under equal weighting of the 108 conditions. For the target-mean and leading pooled second-moment directions, mean differences are −0.017204 and −0.017194. The between-hospital direction is unfavorable at +0.002606 (5/3/28): its squared-bias reduction −0.0000759 is outweighed by privacy and sampling increases +0.0026709 and +0.0000412.

The full condition set also retains large relative deteriorations. At $\rho_0 = 0.005$, three equal-cost between-hospital conditions increase absolute MSE by 0.0117 to 0.0162 against nonzero comparator MSE 0.00148 to 0.00156, producing relative improvements between −1,037% and −792%. Conversely, six equal-cost target-mean and leading second-moment conditions at $\rho_0 = 0.5$ produce 98.32% of gross gains within that privacy slice. These decompositions identify bias and variance offsets on this semi-synthetic cohort.

J Design-objective comparisons and local refinement

This study separates numerical design construction from the objective being optimized. On the 20 small fixed-allocation problems, all candidate allocations use the same GB representation and analytic estimator-risk evaluation.

Design objectives and minimax comparison. We compare two design objectives using the same Gaussian release. The design criterion ψ_r connects site sample allocation to protocol-class minimax risk. The optimized Gaussian bound g_r accounts for the specified release variance. Proposition 6 compares the objectives under consistent matrices and gives the design-level consequence for a global integer minimizer of g_r. The actual worst-case risk R_G^{wc} evaluates the selected estimator over the parameter ball and retains its bias and sampling-variance correction.

The Gaussian-bound objective. For $r > 0$, $w \in \mathcal{U}$, and fixed integer N, let $G_s = H_s^{1/2}$ and $\beta_s = N_s^{-1} + 2L^2/(\rho_s N_s^2)$ at active sites. With $z_s = 0$ at inactive sites, Equation (19) follows by completing squares.

To prove this identity, denote the bracketed operator by A and set $q = A_{\mathcal{U}}^{-1} w$, $z_s^* = \beta_s^{-1} G_s q$, and $b^* = r^{-2} q$. These variables satisfy the representation constraint. For every other feasible representation, completing squares gives

$$\sum_s \beta_s \|z_s\|^2 + r^2 \|b\|^2 - w^\top q = \sum_s \beta_s \|z_s - z_s^*\|^2 + r^2 \|b - b^*\|^2 \geq 0.$$

The sums run over active sites. Projection onto range(H_s) and $\mathcal{U}$ justifies restricting z_s and b to those spaces without increasing the objective. Zero recruitment gives $g_r = r^2 \|w\|^2$. A zero target gives zero, including when $\mathcal{U} = \{0\}$. This is the optimized risk upper bound for the Gaussian construction.

J.1 Criterion comparison and its design-level consequence

Proposition 6 (Comparison of the design objectives)**.** *On nonzero site ranges, bounded leverage implies $L^2 \geq \operatorname{rank}(H_s) \geq 1$. Zero-information sites can be omitted from the representation. Let*

$$C_G = 1 + 2L^2 \max_{s: H_s \neq 0} \frac{\kappa_s}{\rho_s} \leq 1 + 4L^2 e^{2\bar{\rho}} = C_+.$$

The maximum is defined as zero if every $H_s = 0$. Then, for every feasible integer design,

$$\psi_r(N; w) \leq g_r(N; w) \leq C_G \psi_r(N; w).$$

Moreover, for the consistent-matrix Gaussian construction and its bound-optimal representation,

$$R_N^\star \leq R_G^{\mathrm{wc}}(N, z^{\mathrm{GB}}) \leq g_r(N; w).$$

Proof. Put $a_s = 2L^2/\rho_s$. For $n > 0$, $\beta_s(n) = 1/n + a_s/n^2$. Since $\kappa_s \geq 2\rho_s$ and $L^2 \geq 1$, $a_s \kappa_s \geq 1$, giving $\beta_s(n) \geq 1/\alpha_s(n)$. If $\kappa_s n \leq 1$, then $\beta_s(n)\alpha_s(n) = \kappa_s n + a_s \kappa_s \leq 1 + a_s \kappa_s$. If $\kappa_s n \geq 1$, the product is $1 + a_s/n \leq 1 + a_s \kappa_s$. Thus $1/\alpha_s(n) \leq \beta_s(n) \leq C_G/\alpha_s(n)$. Minimize over the same representations, retaining the nonnegative squared-bias term, to obtain the criterion comparison. The Gaussian risk is bounded by the optimized bound, while protocol minimax risk is no larger than the risk of this admissible construction. The bound on C_G follows from $e^{2\rho} - 1 \leq 2\rho e^{2\bar{\rho}}$. If every site has zero information, $w \in \mathcal{U}$ implies $w = 0$, and all criteria vanish. □

If N_g is a global minimizer of g_r on the same integer feasible domain, then

$$R_G^{\mathrm{wc}}(N_g, z^{\mathrm{GB}}) \leq C_G C_- \, \mathfrak{R}_{B,r}^\star(w). \tag{56}$$

Indeed, for every feasible N^*,

$$R_G^{\mathrm{wc}}(N_g, z^{\mathrm{GB}}) \leq g_r(N_g; w) \leq g_r(N^*; w) \leq C_G \psi_r(N^*; w) \leq C_G C_- R_{N^*}^\star.$$

Minimizing over N^* gives the claim. Positive costs and finite budget make the domain finite. This comparison uses globally optimized sample allocations and consistent matrices. The semi-synthetic eICU study uses the split described in Appendix I.2.

Actual worst-case risk and the ideal consistent-matrix simplification. In the consistent model with inactive clipping, $V = (2A - 1)(2Y - 1)X$ satisfies $\mathbb{E}_\theta V = H_s \theta$ and $\mathbb{E}_\theta V V^\top = H_s$ for every valid θ. Thus the weighted site release has mean $h_s^\top \theta$, with $h_s = G_s z_s$, and variance $\beta_s \|z_s\|^2 - (h_s^\top \theta)^2 / N_s$. Independence and the representation constraint give

$$C_0 = \sum_{s: N_s > 0} \beta_s \|z_s\|^2, \qquad K = b b^\top - \sum_{s: N_s > 0} h_s h_s^\top / N_s,$$

$$\mathrm{MSE}_\theta = C_0 + \theta^\top K \theta, \qquad R_G^{\mathrm{wc}} = C_0 + r^2 \max\{0, \lambda_{\max}(Q^\top K Q)\}, \tag{57}$$

where Q spans $\mathcal{U}$ orthonormally. The last equality is the Rayleigh-quotient maximum over the ball, including its origin. At $\theta = 0$ the risk is C_0. For the GB representation, $R_G^{\mathrm{wc}} \leq g_r$. A repeated top eigenvalue can make the component decomposition depend on the selected maximizer. We evaluate all components at the same maximizing θ.

For the actual floating matrices we use the general identities in Equations (52) to (54), with H_s^{emp} replaced by the finite-support second moment H_s^Q and J_s^{tr} by the executed inverse-square-root transform J_s. Specifically, $v_s = J_s^\top z_s$, $h_s = H_s^Q v_s$ and $b_{\mathrm{act}} = \sum_s h_s - w$ replace the ideal mean and bias, and the sampling constant is $\sum_s v_s^\top H_s^Q v_s / N_s$. These expressions include factorization and representation residuals in the evaluated risk.

Table 6: Design-objective comparison on the ten high-privacy small-instance problems. Entries are arithmetic means. All allocations use the same Gaussian-bound representation. g_r is an upper-bound objective and actual worst-case risk R_G^{wc} is evaluated separately. The two references use exhaustive floating-point enumeration over the same feasible integer domain. The mixed-group summary appears in Table 3.

Design	Mean ψ_r	Mean g_r	Mean R_G^{wc}
Sampling	0.460	0.534	0.533
ψ-Conic	0.409	0.514	0.513
ψ-Reference	0.402	0.511	0.510
g-Reference	0.402	0.511	0.510

Inputs and selection. Each site has two opposite support points with equal probability and public leverage bound $L = 1$. The maximum $r\|x\|$ is 0.850412, below $1 - \eta = 0.9$. The release uses these fixed supports. Sampling and ψ-Conic retain their componentwise-floor allocations. Enumeration covers all 3,016 feasible integer allocations across the 20 instances and selects each objective minimum using decimal-cost feasibility and binary64 linear algebra. Ties are ordered by objective, cost, then lexicographic allocation. All methods use the same GB representation. No pair of leading objective values lies within $10^{-12}\max\{1, |\text{minimum}|\}$.

Separating algorithm, objective and estimator effects. Table 6 reports the high-privacy means, where both objective references choose the same allocations. Figure 5 gives the instancewise comparison. The g-Reference improves actual worst-case risk R_G^{wc} over Sampling in all 20 instances and over the ψ-Conic allocation in 14, with identical sample allocations in the other six. In the mixed group, the ψ-Conic-to-ψ-Reference criterion ratio has median 1.498 and 95th percentile 1.681. This comparison measures the combined loss from continuous construction and integer recovery.

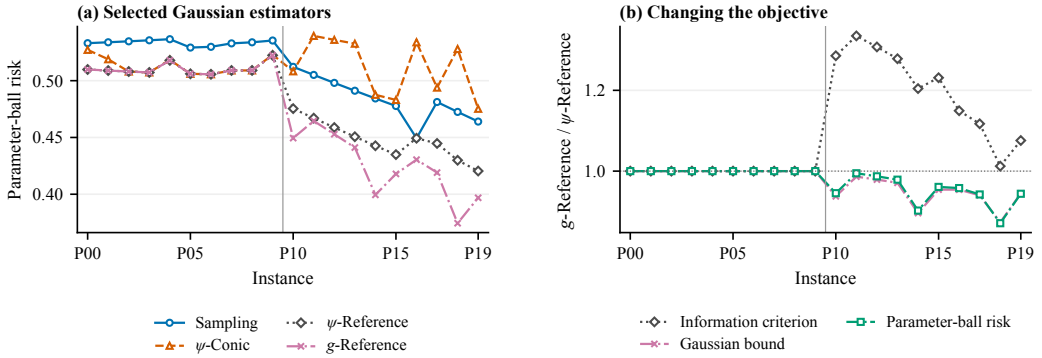


Figure 5: Supplementary small-instance design comparison. (a) reports the actual worst-case risk R_G^{wc} of the common Gaussian estimator at each selected allocation. (b) divides the g-Reference's criterion, bound and risk by the corresponding value at the ψ-Reference. All 20 original instances are shown. Curves connect instance IDs only. The two references coincide in the high-privacy group. In the mixed group, a higher information criterion accompanies a lower Gaussian bound and lower R_G^{wc}. Values are obtained by analytic risk evaluation.

J.2 Local refinement and validation

The fill and exchange stages optimize ψ_r, and the final local refinement optimizes g_r. A denotes ψ-Conic, B adds the fill stage, C adds the exchange stage, and D adds the g_r refinement. On the 20 development instances, C and D match their exhaustive objective references, with 100,000 evaluations and 1,000 accepted updates allowed per stage. These refinements are separate from ψ-Conic.

For validation, let f be the objective used by the current stage. Fill evaluates every budget-feasible $N+e_j$ and accepts an objective reduction exceeding $10^{-12}+10^{-12}\max\{|f(N)|, |f(N+e_j)|\}$. It selects the largest gain per cost, breaking ties by total decimal cost and lexicographic allocation. Exchange considers every donor, receiver, $1 \le k \le N_i$, and budget-feasible positive ℓ in $N - ke_i + \ell e_j$. It

Table 7: Separate 192-instance validation. Risk changes are $100(R_{\text{numerator}}/R_{\text{denominator}} - 1)$. Negative favors the numerator. R denotes R_G^{wc}. W/T/L denotes wins, ties, and losses under the rule in Equation (58). Criterion ratios are numerator over denominator. Incomplete searches return the last accepted allocation.

Comparison	n	ψ_r ratio		Risk change (%)			W/T/L
		Median	95th pct.	Median	95th pct.	Maximum	
B/A	192	0.973	0.997	−2.188	−0.4550	−0.07769	192/0/0
C/A	192	0.952	0.996	−1.530	14.70	25.26	153/0/39
C/Sampling	192	0.843	1.032	−10.74	−0.7828	2.174	188/0/4
C/C-S	192	1.000	1.038	0	0	3.011	52/135/5
D/C	192	1.000	1.603	0	0	0	82/110/0
D-S/C-S	192	1.000	1.603	0	0	0	81/111/0

selects the lowest improving objective, using the same tie rule, then refills. This allows activation of unused sites and many-for-many moves.

B and C each have a budget of 20,000 evaluations and 200 accepted updates, including the initial objective and every candidate evaluation. C starts from ψ-Conic. C-S starts from Sampling. D starts from C, and D-S starts from C-S. Both use the same neighborhood with g_r and a separate equal budget. Incomplete searches return the last accepted feasible allocation.

The 192 validation instances cross $S \in \{3, 5, 10, 20\}$, effective dimension 1 or 3, the three privacy-input groups in Appendix G.2, equal or heterogeneous integer costs, low or high budgets, and two replicates. Validation uses a separate set of random draws from development. Equal costs are one. Heterogeneous costs cycle a random permutation of $\{1, 2, 4\}$. Low and high budgets are $2S \max_s c_s$ and $8S \max_s c_s$. The finite-support construction uses $r = 0.72$ and public $L = 1$. Group membership is checked through the $\kappa_s B/c_s$ intervals. The privacy coefficients vary with the budget in this constructed family.

For risks $R_{\text{denominator}}$ and $R_{\text{numerator}}$, the comparison tolerance is

$$\tau = 10^{-9} + 10^{-8} \max(|R_{\text{denominator}}|, |R_{\text{numerator}}|). \tag{58}$$

A win requires $R_{\text{denominator}} - R_{\text{numerator}} > \tau$, a loss requires the reverse difference to exceed τ, and all other cases are ties. The zero median, 95th percentile, and maximum changes for D/C and D-S/C-S reflect their 110 and 111 ties, respectively. Improvements occur below the median.

Median sample-allocation times for B, C, C-S, D, and D-S are 0.0142, 0.0507, 0.0461, 0.129, and 0.100 seconds, respectively. Sample-allocation times include initialization and search.

Final GB computation and risk evaluation are timed separately. The D and D-S times include construction of their C and C-S starting sample allocations. C, C-S, D, and D-S exhaust the prespecified search budget in 74, 75, 79, and 79 of the 192 instances, respectively. The fill search for B terminates exhaustively in all 192 instances.

B improves both the criterion and R_G^{wc} on all 192 instances. C improves the criterion throughout, but worsens R_G^{wc} in 39 instances versus A and four versus Sampling. The g_r refinement improves or ties R_G^{wc} in every instance while sometimes increasing ψ_r (Table 7).

The validation runs use one Apple M5 Pro process with numerical-library threads set to one. Software versions match those reported in Appendix G.2.