ORTHOGONALITY IS NOT REALIZABILITY: SIGNED GENERATIVE RISKS AND PROBABILITY-CONSTRAINED COUNTERFACTUAL LEARNING

Anonymous authors
Paper under double-blind review

ABSTRACT

Orthogonal scores remove first-order nuisance bias, but their empirical targets need not be probability laws. We characterize this gap through a failure–recovery theory for proper generative risks. We construct smooth nuisance estimates whose L_4 errors vanish while their pseudo-density ratios remain negative on rare covariate regions. For categorical log risks, we give recovery conditions in minimum cell mass, observations per stratum, and nuisance-product error, with separate counterexamples isolating sampling and nuisance effects. Within a released counterfactual diffusion architecture, growing denoising-feature moments yield an exactly integrated objective that is unbounded below even with nuisance-product error $o(n^{-1/2})$. We formulate orthogonal embedding projection (OEP) as a minimum-distance probability repair of doubly robust feature estimates, preserving accuracy in the chosen distance and the nuisance-product error bound. Exact finite-feature guarantees and a Project STAR study connect the theory to computation.

1 INTRODUCTION

Counterfactual distribution learning estimates how outcomes would be distributed under a specified treatment. For example, a school may want to know how smaller classes would affect the proportion of pupils below a reading threshold. A fitted distribution can answer this question, give quantiles, and generate outcomes for a group of pupils defined by baseline characteristics.

A common approach estimates treatment probabilities and conditional outcome distributions, then uses these nuisance estimates to correct a generative model's training loss (Melnychuk & Feuerriegel, 2026; Luedtke & Fukumizu, 2025). This extends the orthogonal-score principle of double/debiased machine learning (Chernozhukov et al., 2018): the population objective is insensitive to first-order nuisance errors. In a finite sample, however, the correction can subtract more probability than the fitted outcome model assigns. The resulting negative mass can prevent the training loss from having a minimum.

Consider two outcomes, 0 and 1. Suppose treatment has probability $1/2$, and the correct fitted outcome model assigns probability $1/2$ to each outcome. One observation with $A = a$ and $Y = 0$ receives inverse-probability weight 2. Its corrected target is

$$q = \widehat{P} + 2\{(1,0) - \widehat{P}\} = (3/2, -1/2), \qquad \widehat{P} = (1/2, 1/2).$$

The masses sum to one, but the second is negative. For candidate probabilities p_0, p_1, the negative log loss becomes

$$L_q(p) = -\tfrac{3}{2}\log p_0 + \tfrac{1}{2}\log p_1 \longrightarrow -\infty \qquad \text{as } p_1 \downarrow 0.$$

The loss keeps decreasing as the model removes probability from outcome 1, even though the preliminary estimates are correct.

For asymptotic recovery, the question is whether these negative targets disappear as the sample grows. GDR-learners require a positive pseudo-density ratio; Remark 1 and Eq. (86) of Melnychuk & Feuerriegel (2026) infer this condition from Hölder smoothness and a vanishing L_4 nuisance-error product. On an unbounded covariate domain, these conditions allow fixed relative errors in increasingly rare regions. Our smooth construction gives a negative ratio in such regions, which an

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.

evaluation sample encounters with probability bounded away from zero (Appendix A.5). The missing control is uniform positivity at the resolution used for fitting.

We determine what supplies that control. Theorems 2–3 give a recovery condition in minimum cell mass, observations per stratum, and nuisance-product error, with separate constructions isolating the two error terms. Proposition 4 establishes unbounded loss within the released diffusion architecture after exact integration of all stochastic variables, even with a nuisance product $o(n^{-1/2})$. The construction identifies the role of growing denoising-feature moments.

When correction produces an invalid target, we replace it by the closest probability distribution in a chosen distance. For the binary example, Euclidean projection gives $(1, 0)$. For general outcomes, we apply this principle to averages of distributional features and call it orthogonal embedding projection (OEP). Theorems 5–6 state its minimum-distance criterion and the resulting error bound. Threshold features give a cumulative distribution function (CDF), with projection computed by weighted isotonic regression. The chosen distance determines the repair's error guarantee.

OEP uses the augmented-feature regression of conditional counterfactual mean embeddings (Anancharoenkij & Ponnoprat, 2026). Isotonic correction also has established error and asymptotic guarantees (Westling et al., 2020). Appendix E relates these methods to our recovery conditions and probability repair.

2 When is the corrected target a probability distribution?

2.1 Correction can produce negative probabilities

Let $O_i = (X_i, A_i, Y_i)$ be independent observations, and let the generator condition on $V = v(X)$. Throughout the theory we fix a treatment a and suppress its index. Write

$$e(x) = \mathbb{P}(A = a \mid X = x), \qquad P_x(B) = \mathbb{P}(Y \in B \mid X = x, A = a).$$

Under consistency, conditional exchangeability, and positivity, the target is

$$P_v(B) = \mathbb{P}(Y^a \in B \mid V = v) = \mathbb{E}\{P_X(B) \mid V = v\}.$$

Hats denote estimates. Expectations involving fitted nuisances condition on their training sample, and quantities may depend on n without an n subscript.

For finite conditioning strata, let $n_v = \sum_i \mathbf{1}(V_i = v)$. With nuisances fitted on independent training data, define the signed target

$$\widehat{Q}_v = \frac{1}{n_v} \sum_{i:V_i=v} [w_i \delta_{Y_i} + (1 - w_i)\widehat{P}_{X_i}], \qquad w_i = \frac{\mathbf{1}(A_i = a)}{\widehat{e}(X_i)}. \tag{1}$$

Each summand has mass one. For a row with $A_i = a$ and $\widehat{e}(X_i) < 1$, the coefficient $1 - w_i$ is negative. Thus $\widehat{Q}_v$ is a signed measure: its mass on a set may be negative. Averaging a generative loss $\ell_g(v, y)$ against this target gives

$$\bar{R}(g) = \sum_v \frac{n_v}{n} \int \ell_g(v, y)\, d\widehat{Q}_v(y). \tag{2}$$

The integral uses the fitted law itself. Replacing it by M nuisance draws gives the sampled risk $\widehat{R}_M(g)$; we keep this sampling step explicit in the experiments. Risk identities apply when the displayed terms are absolutely integrable.

For an outcome partition $B_1, \ldots, B_J$, the cell masses and log risk are

$$q_{vj} = \widehat{Q}_v(B_j), \qquad L_{q_v}(p_v) = -\sum_{j=1}^{J} q_{vj} \log p_{vj}, \quad p_v \in \Delta_J. \tag{3}$$

The following result applies when the model can fit any probability vector in each stratum.

Theorem 1 (Log loss with signed masses)**.** *Let $q \in \mathbb{R}^J$ satisfy $\sum_j q_j = 1$. If some $q_j < 0$, then $\inf_{p \in \Delta_J^\circ} L_q(p) = -\infty$. If $q_j \geq 0$ for all j, the infimum is finite and equals $-\sum_j q_j \log q_j$, with $0 \log 0 = 0$. It is attained at $p = q$ on the closed simplex under the extended convention; on the open simplex it is attained if and only if every $q_j > 0$.*

Appendix A proves the result and extends the argument to other losses.

2.2 When are the corrected probabilities positive?

Let S be the number of strata, $K = SJ$ the number of cells, and $n_{\min} = \min_v n_v$. The smallest true mass is

$$p_* = \min_{v,j} p_{vj}, \qquad p_{vj} = P_v(B_j).$$

The partition and these quantities may change with sample size. Write

$$\delta_e = \|\widehat{e} - e\|_\infty, \qquad \delta_P = \max_j \sup_x |\widehat{P}_x(B_j) - P_x(B_j)|.$$

The conditional nuisance bias is

$$b_{vj} = \mathbb{E}\left[\frac{\widehat{e}(X) - e(X)}{\widehat{e}(X)}\{\widehat{P}_X(B_j) - P_X(B_j)\} \mid V = v\right], \qquad b = \max_{v,j} |b_{vj}|. \tag{4}$$

The corrected masses are positive when sampling error and this bias are small relative to the smallest true probability p_*.

Theorem 2 (Recovery of positive masses)**.** *Suppose the evaluation rows are independent of the nuisance fits and $e, \widehat{e} \geq c > 0$. Conditional on the fits and stratum labels, if $n_{\min} \geq 1$, then for every $t > 0$,*

$$\mathbb{P}\{\max_{v,j} |q_{vj} - p_{vj}| > t + b\} \leq 2K \exp\left(-\frac{c^2 n_{\min} t^2}{2}\right). \tag{5}$$

Consequently,

$$\sqrt{\frac{\log(2K)}{n_{\min}}} + b = o_p(p_*) \tag{6}$$

implies that every $q_{vj} > 0$ with probability tending to one. Since $b \leq c^{-1}\delta_e\delta_P$, a sufficient condition is

$$\sqrt{\frac{\log(2K)}{n_{\min}}} + c^{-1}\delta_e\delta_P = o_p(p_*). \tag{7}$$

Under either condition, the saturated empirical log risk has the finite, unique interior optimizer $p_v = q_v$ in every stratum.

The argument applies fold by fold to a fixed-fold cross-fit.

Example: sample size and resolution. Suppose the J outcome cells in each of the S strata have comparable mass, so $p_* \asymp 1/J$, and the conditional product bias satisfies $b = o_p(1/J)$. Theorem 2's sufficient condition becomes

$$\frac{n_{\min}}{J^2 \log(2SJ)} \longrightarrow \infty.$$

For balanced strata, $n_{\min} \asymp n/S$, this is $n/\{SJ^2 \log(2SJ)\} \to \infty$. The per-stratum sample-size scale in this bound grows quadratically with J, apart from the logarithm. Thus finer outcome grids require more observations within every stratum; increasing the total sample alone need not improve the least-populated stratum.

The next result isolates why both terms in equation 7 are needed.

Theorem 3 (Examples showing both error terms are needed)**.** *The two terms in equation 7 cannot be uniformly improved up to constants, even with a binary partition and uniform overlap.*

(i) Sampling. *Let $S \to \infty$ strata each contain $m \to \infty$ independent evaluation rows, with $\log S = o(m^{1/3})$. For every $0 < \kappa < 1/\sqrt{2}$, there are proper fitted nuisances with $e = \widehat{e} = 1/2$, $K = 2S$, and*

$$p_* = \kappa\sqrt{\frac{\log S}{m}}$$

such that $\mathbb{P}\{\min_{v \leq S} q_{v1} < 0\} \to 1$.

(ii) Nuisance product. *With one stratum, two outcome cells, and $n_{\min} = n$, there are proper fitted nuisances satisfying*

$$\delta_e = \delta_P = n^{-1/8}, \qquad p_* = \delta_e\delta_P = n^{-1/4},$$

for which $\sqrt{\log(2K)/n} = o(p_*)$ and $\mathbb{P}(q_{11} < 0) \to 1$.

The first example isolates sampling error because its propensity estimate is exact. The second isolates the product of nuisance errors because sampling error is negligible. Appendix Theorem 8 extends the recovery argument to smoothed continuous densities.

3 THE SAME PROBLEM IN A DIFFUSION MODEL

Diffusion models use a squared denoising loss. When the prediction is affine in some parameters and the remaining parameters are fixed, this loss is a quadratic. Its boundedness can then be checked through a matrix. For signed coefficients α_r, targets z_r, offsets c_r, and features h_r, write

$$\begin{aligned} J(\beta) &= \sum_r \alpha_r (z_r - c_r - h_r^\top \beta)^2 = \beta^\top G \beta - 2b^\top \beta + C, \\ G &= \sum_r \alpha_r h_r h_r^\top, \qquad b = \sum_r \alpha_r (z_r - c_r) h_r. \end{aligned} \tag{8}$$

This objective is bounded below exactly when $G \succeq 0$ and $b \in \mathrm{Range}(G)$. A negative eigenvalue identifies a direction along which the loss decreases without limit.

The released GDR diffusion model for scalar outcomes contains such a quadratic. Its hypernetwork generates an affine denoiser output layer. Holding the preceding features fixed makes the prediction linear in the corresponding hypernetwork weights, giving equation 8. The identity holds for fixed draws and after averaging over the stochastic variables. Appendix A.4 gives the layer dimensions and identifies these weights in the implementation.

Proposition 4 (An unbounded diffusion loss)**.** For any $1/4 < \gamma < 1/2$, the scalar-outcome diffusion architecture above admits a triangular array with the released layer dimensions held fixed, uniform propensity overlap, and proper smooth outcome densities $p, \widehat{p}$ such that

$$\|\widehat{e} - e\|_{L_4} = O(n^{-\gamma}), \qquad \|\widehat{p} - p\|_{L_4} = O(n^{-\gamma}),$$

but the loss restricted to one generated output-layer coordinate has an exactly integrated empirical Gram coefficient $G_n \to G_\infty < 0$ in probability. The integration includes the nuisance outcome, diffusion time, Gaussian denoising noise, and context jitter. Hence the corresponding affine-head objective is unbounded below with probability tending to one even though the nuisance product is $O(n^{-2\gamma}) = o(n^{-1/2})$.

For example, $\gamma = 1/3$ gives a product rate $O(n^{-2/3})$. The construction moves a vanishing amount of fitted probability far into the tail, where the denoising feature grows. Its fitted second moment is of order n^γ; this growing moment makes the quadratic coefficient negative despite the small density error. The density norms use the covariate law and a fixed dominating outcome measure. Appendix A.4 gives the construction and integrates all stochastic variables explicitly.

4 PROJECTION ONTO PROBABILITY DISTRIBUTIONS

OEP first estimates conditional averages of distributional features by augmented-feature regression (Anancharoenkij & Ponnoprat, 2026). It then finds the closest feature average attainable by a probability law. The feature map and norm determine this distance. For CDFs, the projection is an isotonic regression (Westling et al., 2020).

Let $\mathcal{Y}$ be compact metric, and let k be a bounded continuous kernel with RKHS $\mathcal{H}$ and feature $\phi(y) = k(y, \cdot)$. Write

$$m(x) = \mathbb{E}\{\phi(Y) \mid X = x, A = a\}, \qquad \mu(v) = \mathbb{E}\{m(X) \mid V = v\}.$$

Pointwise norms are Hilbert norms. For a function h of X or V, $\|h\|_{L_r} = (\mathbb{E}\|h\|^r)^{1/r}$, using the law of its argument. The cross-fitted augmented feature is

$$\psi_i = \widehat{m}(X_i) + \frac{\mathbf{1}(A_i = a)}{\widehat{e}(X_i)} \{\phi(Y_i) - \widehat{m}(X_i)\}. \tag{9}$$

Regress ψ_i on V_i to obtain $\widehat{\mu}(v)$, then project:

$$\widetilde{\mu}(v) = \Pi_{\mathcal{M}} \widehat{\mu}(v), \qquad \mathcal{M} = \left\{ \mu_P = \int \phi(y)\, dP(y) : P \in \mathcal{P}(\mathcal{Y}) \right\}. \tag{10}$$

The feature average μ_P is the mean embedding of P. For cell averages, projection also minimizes the signed kernel-score loss over probability laws (Appendix B.4). The Hilbert projection theorem (Bauschke & Combettes, 2017) gives the following guarantee for this repair.

Theorem 5 (Projection onto probability embeddings)**.** *The set $\mathcal{M}$ is nonempty, compact, and convex. Its projection is unique and 1-Lipschitz. For every $z \in \mathcal{H}$, $\Pi_{\mathcal{M}} z$ is the unique probability embedding closest to z, and for every $\mu_0 \in \mathcal{M}$,*

$$\|\Pi_{\mathcal{M}} z - \mu_0\|^2 \le \|z - \mu_0\|^2 - \|z - \Pi_{\mathcal{M}} z\|^2. \tag{11}$$

If k is characteristic, the projected embedding identifies a unique law.

Projection therefore reduces squared error by at least the squared size of the correction in equation 11. Combining this property with the doubly robust conditional-bias identity bounds the repaired error by regression error and the product of the two nuisance errors.

Theorem 6 (Error after projection)**.** *Suppose $\widehat{e} \ge c > 0$, and let $\bar{\mu}(v) = \mathbb{E}(\psi \mid V = v)$. The conditional bias is*

$$b(v) = \mathbb{E}\left[\frac{\widehat{e}(X) - e(X)}{\widehat{e}(X)}\{\widehat{m}(X) - m(X)\} \mid V = v\right]. \tag{12}$$

Then $\bar{\mu} - \mu = b$, with the pointwise identity

$$\mathbb{E}(\psi \mid X) - m(X) = \frac{\widehat{e}(X) - e(X)}{\widehat{e}(X)}\{\widehat{m}(X) - m(X)\}. \tag{13}$$

The repaired estimate satisfies

$$\begin{aligned} \|\widetilde{\mu} - \mu\|_{L_2} &\le \|\widehat{\mu} - \bar{\mu}\|_{L_2} + \|b\|_{L_2} \\ &\le \|\widehat{\mu} - \bar{\mu}\|_{L_2} + c^{-1}\|\widehat{e} - e\|_{L_4}\|\widehat{m} - m\|_{L_4}. \end{aligned} \tag{14}$$

The distance between probability embeddings is maximum mean discrepancy (MMD). The feature map determines which distributional differences it measures.

4.1 Computing the projection for a scalar outcome

For a scalar outcome, we can work directly with its CDF. Let Q have mass one and $\int |y|\, d|Q|(y) < \infty$, and write $H(t) = Q((-\infty, t])$. Minimizing its signed continuous ranked probability score (CRPS) is equivalent to minimizing $\int \{F_P(t) - H(t)\}^2 dt$, up to a constant independent of P. Choose support $s_1 < \cdots < s_m$, put $\Delta_j = s_{j+1} - s_j$, and average the target CDF on each interval, $u_j = \Delta_j^{-1} \int_{s_j}^{s_{j+1}} H(t)dt$. Solve

$$\widehat{c} = \operatorname*{arg\,min}_{0 \le c_1 \le \cdots \le c_{m-1} \le 1} \sum_{j=1}^{m-1} \Delta_j (c_j - u_j)^2. \tag{15}$$

Weighted isotonic regression supplies $\widehat{c}$. Consecutive differences give nonnegative weights $(\widehat{c}_1, \widehat{c}_2 - \widehat{c}_1, \ldots, 1 - \widehat{c}_{m-1})$ that sum to one, so the result is a sampleable probability law.

Random Fourier features (RFF) (Rahimi & Recht, 2007) give a finite-vector implementation of the same projection, using a quadratic program over probability weights. Restricting the support and solving the program approximately each add a term to the error bound. Appendix B.2 gives this feature implementation. Appendix B.4 derives the CDF program and its numerical error bounds for CDF values, tail probabilities, and quantiles.

Choosing the repair distance. Taking the positive part and renormalizing also returns a probability law. It minimizes total variation from the signed target (Appendix B.4). CDF projection instead preserves accuracy in the distance that determines expected CRPS. These choices can behave differently when nearby positive and negative masses cancel.

For example, let the true law be $P = \delta_1$, where δ_y is a point mass at y, and take $Q = \delta_1 + \varepsilon\delta_0 - \varepsilon\delta_h$, with $0 < \varepsilon, h < 1$. Its CDF error is confined to $[0, h)$. Positive part gives $P_+ = (\delta_1 + \varepsilon\delta_0)/(1 + \varepsilon)$; CDF projection gives $P_{\mathrm{CDF}} = (1 - \varepsilon h)\delta_1 + \varepsilon h \delta_0$. Their squared CDF errors are

$$\|H - F_P\|_2^2 = \varepsilon^2 h, \qquad \|F_{P_+} - F_P\|_2^2 = \frac{\varepsilon^2}{(1+\varepsilon)^2}, \qquad \|F_{P_{\mathrm{CDF}}} - F_P\|_2^2 = \varepsilon^2 h^2.$$

As h decreases, CDF projection retains the target's convergence to P; positive part leaves a fixed error. This is the reason to specify a distance: the repair should preserve the accuracy established for the target.

5 EXPERIMENTS

5.1 WHEN IS THE TRAINING OBJECTIVE UNBOUNDED?

We audit the released GDR objective on Project STAR kindergarten reading scores, with small class as action 1 and the two regular-class arms pooled as action 0 (Word et al., 1990; Krueger, 1999). This experiment uses the 5,785 complete cases from 79 schools, a 4,048/1,737 training/evaluation split, and cross-fitted Gaussian outcome nuisances. The conditioning cells use gender, coarsened ethnicity, and school type. Appendix C.1 specifies the data and audit protocol.

Table 1 compares one nuisance draw with exact integration of the fitted Gaussian law, using the same observations, nuisances, 32 outcome bins, batches, and groups. Both integrated action targets have negative mass. Lowering a negative-mass cell's probability floor from 10^{-3} to 10^{-12} increases the log score, hence lowers the negative log risk, as predicted by Theorem 1.

Table 1: Signed-mass audit on Project STAR. Medians are among failing units. Integrated rows average the same fitted Gaussian nuisance law analytically.

Objective unit	N	Negative	Neg. mass	Log-score increase
One-draw action-batch	126	93.65%	0.03604	0.76828
Integrated action-batch	126	84.13%	0.00554	0.12931
Integrated eligible V-action cell	20	80.00%	0.00098	0.03294
Integrated full-train action	2	100.00%	0.000064	0.00483

The matched diffusion audit reconstructs the released architecture on the same data and nuisances, fixes the trunk at a deterministic initialization, and sums both action losses. The one-draw full-data affine-head Gram has two negative eigenvalues. Averaging 512 draws over nuisance outcomes and internal diffusion randomness gives a positive full-data minimum eigenvalue, with block-jackknife interval $[3.90, 4.03] \times 10^{-6}$. Nine of 63 minibatches have intervals entirely below zero. Their negative eigenvalues identify directions of unbounded loss for the corresponding averaged objectives. Proposition 4 gives a construction with negative curvature after exact integration. Appendices C.1–C.2 report the effect of averaging more draws and of redrawing them during training.

5.2 HOW DOES THE CHOICE OF DISTANCE AFFECT THE FITTED DISTRIBUTION?

We first ask when CDF projection predicts better than positive part. Let $Y = \sqrt{r}X + \sqrt{1-r}\epsilon$, with independent standard-normal X, ϵ, so covariates explain fraction r of outcome variance while the target distribution stays $N(0,1)$. An independent sample fits the Gaussian outcome regression; both repairs use the same fitted nuisance, observations, and support. Table 2 shows that the ordering changes with r. At $r = 0.9$ and 0.99, CDF projection reduces mean CRPS excess by 10.2% and 31.6%. The advantage remains when both repairs start from the same mapped masses and at 65 or 257 support points. With weakly predictive covariates, positive part has lower error.

Table 2: Predictive error as covariates explain more outcome variation. Mean CRPS excess, multiplied by 10^3, over 1,000 independent repetitions; $n = 256$, $e = 0.1$, and 129 support points. The last column gives CDF minus positive part with a paired 95% interval. Mapped CDF uses the same signed masses as positive part.

r	Positive part	CDF	Mapped CDF	Plug-in	Difference [95% CI]
0	20.671	22.781	23.245	16.573	$2.110\ [1.675, 2.546]$
0.5	16.087	17.061	17.479	10.004	$0.974\ [0.627, 1.322]$
0.9	9.174	8.242	8.630	3.663	$-0.932\ [-1.157, -0.707]$
0.99	5.478	3.746	4.101	2.319	$-1.732\ [-1.907, -1.558]$

The correctly specified plug-in regression has the lowest error here. Shifting its fitted mean by 0.35 raises its error to 0.03691 at $r = 0.99$; CDF and positive part then give 0.01042 and 0.01121, respectively. Appendix C.4 gives the fitted-nuisance design, independent confirmation, and finite-draw comparison.

On the same STAR split, all repairs use the same Gaussian nuisances, 60-point support, and ten conditioning cells. We compare 64-feature RFF projection, CDF projection, and positive part, which clips and renormalizes the signed Voronoi masses. Correction distances include input mapping from the common continuous target.

Figure 1 shows that a smaller RFF correction can accompany worse CRPS. Compared with first mapping the target to the support, projecting its continuous feature mean reduces RFF correction from 0.005434 to 0.001227 for action 0 and from 0.007752 to 0.001856 for action 1. Its held-out CRPS is 16.953 and 17.181, respectively. CDF projection gives 15.830 and 16.184, close to positive part. Appendix C.3 examines support resolution, weight penalties, and tail probabilities.

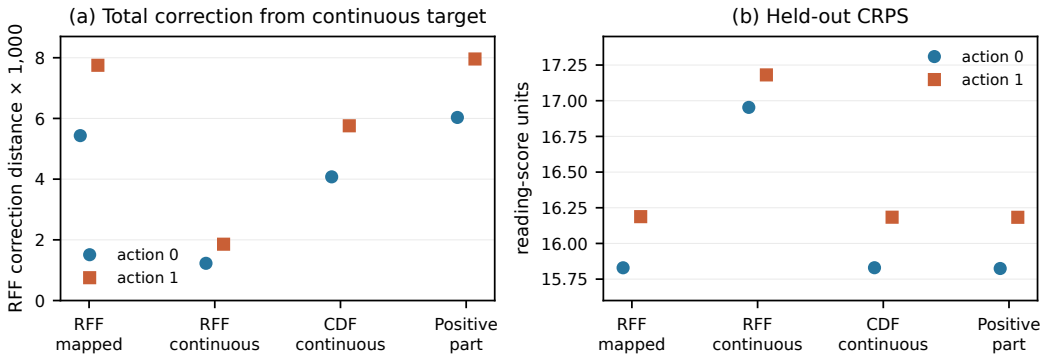

Figure 1: Repairs of a common continuous signed target. Left: total correction in RFF distance. Right: held-out CRPS. "RFF mapped" first maps the target to the support; "RFF continuous" uses its analytic feature expectation.

A separate causal evaluation trains in 40 schools and evaluates in 39 others. It targets baseline-complete pupils from the same school population, conditions on assignment counts, and assumes outcomes are missing at random given treatment and baseline covariates. Nuisance models omit school identifiers. Selection-adjusted CRPS differences between CDF and positive part are -0.0003 $[-0.0040, 0.0034]$ and 0.0021 $[-0.0078, 0.0121]$ for actions 0 and 1, respectively (95% intervals). The sampling target, nuisance-rate assumptions, cluster intervals, and retention validation are specified together in Appendix D.

5.3 How well do the trained models predict?

We next compare prediction after a fixed training budget. Every method uses the same three conditioning groups and two actions, giving six Gaussian target laws. The propensity is correct; the outcome nuisance mean is shifted by $0.35(1 - 2a)$. Neural branches train all layers for 200 epochs with paired initializations, batches, and stochastic draws. They share the calibrated 100-step schedule and standard-normal sampling initialization. The oracle-positive control trains on true-law draws.

Table 3: CRPS excess under the calibrated common-V comparison. Entries are mean (standard deviation) over eight paired seeds, averaging six conditional laws per seed. Neural fits use 200 epochs and a common terminal-calibrated schedule; oracle-positive diffusion trains on true-law draws. Appendix C.5 gives all eight methods and the sampling protocol.

n_v	Signed diffusion	Oracle-positive	CDF projection	Positive part
32	0.0280 (0.0194)	0.0129 (0.0059)	0.0308 (0.0123)	0.0298 (0.0122)
256	0.0051 (0.0037)	0.0009 (0.0003)	0.0050 (0.0017)	0.0050 (0.0014)

Table 3 shows comparable CRPS for CDF projection and positive part at both sample sizes. At 256 observations, signed diffusion is also close: its mean excess is 0.00509, against 0.00504 for CDF and 0.00503 for positive part.

6 Discussion

Corrected generative risks can be unbounded even with accurate nuisance estimates. Our counterexamples locate the failure in rare covariate regions and growing denoising-feature moments. The recovery bounds specify how sample size within each group and nuisance accuracy must scale with outcome resolution. When the target has negative mass, OEP returns its closest probability embedding and preserves the product-error bound.

The repair distance determines the error guarantee. CDF projection improves CRPS over positive part in our simulations with strongly predictive covariates; the two repairs have similar CRPS on STAR and in the calibrated generation experiment. Signed diffusion also attains comparable CRPS after the fixed training budget at the larger sample size.

AI use statement

In this work, we used generative AI tools for the following tasks with required disclosure: generating synthetic data sets, designing or providing feedback on research methodology or experiments, implementing methods, assisting with translation, cleaning and reformatting datasets, and interpreting results. We have not used generative AI tools to help develop theoretical models or conceptual frameworks, formulate mathematical claims, provide critical ingredients for proving mathematical claims, assist in the writing of proofs, or propose or refine hypotheses; the remaining required disclosure task, supporting qualitative and thematic data analysis, is not applicable to this work. Additionally, we used generative AI tools to refine the phrasing and presentation of mathematical statements to improve readability. All AI-assisted work was thoroughly reviewed and validated by the authors: mathematical statements were manually checked, and all AI-generated code was verified and rigorously tested for correctness. The authors take full responsibility for the final content of this work, including all text, claims, and code produced with the aid of generative AI.

Reproducibility statement

The appendices provide complete proofs, architecture correspondences, projection algorithms, and experimental protocols. The supplement contains the public STAR data, executable analyses, fixed random seeds, and all reference results. Its reproduction record identifies the executed code, numerical comparisons, and figure and table checks. Data selection and independent-school inference are specified in Appendix D; the common-conditioning generation comparison and terminal calibration are in Appendix C.5.

References

Thatchanon Anancharoenkij and Donlapark Ponnoprat. Conditional counterfactual mean embeddings: Doubly robust estimation and learning rates, 2026. arXiv:2602.04736.

Heinz H. Bauschke and Patrick L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, second edition, 2017. doi: 10.1007/978-3-319-48311-5.

Michela Bia, Martin Huber, and Lukáš Lafférs. Double machine learning for sample selection models. *Journal of Business & Economic Statistics*, 42(3):958–969, 2024. doi: 10.1080/07350015.2023.2271071.

Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68, 2018. doi: 10.1111/ectj.12097.

Tilmann Gneiting and Adrian E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007. doi: 10.1198/016214506000001437.

Daeyoung Ham, Ted Westling, and Charles R. Doss. Doubly robust estimation and inference for a log-concave counterfactual density, 2024. arXiv:2403.19917.

S. M. Hammer, D. A. Katzenstein, M. D. Hughes, H. Gundacker, R. T. Schooley, R. H. Haubrich, W. K. Henry, M. M. Lederman, J. P. Phair, M. Niu, M. S. Hirsch, and T. C. Merigan. A trial comparing nucleoside monotherapy with combination therapy in HIV-infected adults with CD4 cell counts from 200 to 500 per cubic millimeter. *New England Journal of Medicine*, 335(15): 1081–1090, 1996. doi: 10.1056/NEJM199610103351501.

Jennifer L. Hill. Bayesian nonparametric modeling for causal inference. *Journal of Computational and Graphical Statistics*, 20(1):217–240, 2011. doi: 10.1198/jcgs.2010.08162.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, volume 33, pp. 6840–6851, 2020.

Audrey Huang, Liu Leqi, Zachary C. Lipton, and Kamyar Azizzadenesheli. Off-policy risk assessment for Markov decision processes. In *Proceedings of the 25th International Conference on Artificial Intelligence and Statistics*, volume 151 of *Proceedings of Machine Learning Research*, pp. 5022–5050. PMLR, 2022.

Saksham Jain and Alex Luedtke. Conditional distributional treatment effects: Doubly robust estimation and testing. In *International Conference on Machine Learning*, 2026.

Nathan Kallus and Miruna Oprescu. Robust and agnostic learning of conditional distributional treatment effects. In *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, pp. 6037–6060. PMLR, 2023.

Alan B. Krueger. Experimental estimates of education production functions. *The Quarterly Journal of Economics*, 114(2):497–532, 1999. doi: 10.1162/003355399556052.

Alex Luedtke and Incheoul Chung. One-step estimation of differentiable Hilbert-valued parameters. *The Annals of Statistics*, 52(4):1534–1563, 2024. doi: 10.1214/24-AOS2403.

Alex Luedtke and Kenji Fukumizu. DoubleGen: Debiased generative modeling of counterfactuals, 2025. arXiv:2509.16842.

Diego Martinez-Taboada and Edward H. Kennedy. Counterfactual density estimation using kernel Stein discrepancies. In *International Conference on Learning Representations*, 2024.

Diego Martinez-Taboada, Aaditya Ramdas, and Edward H. Kennedy. An efficient doubly-robust test for the kernel treatment effect. In *Advances in Neural Information Processing Systems*, volume 36, pp. 59924–59952. Curran Associates, Inc., 2023. doi: 10.52202/075280-2619.

Valentyn Melnychuk and Stefan Feuerriegel. GDR-learners: Orthogonal learning of generative models for potential outcomes. In *International Conference on Learning Representations*, 2026.

Krikamol Muandet, Motonobu Kanagawa, Sorawit Saengkyongam, and Sanparith Marukatat. Counterfactual mean embeddings. *Journal of Machine Learning Research*, 22(162):1–71, 2021.

Alexander Quinn Nichol and Prafulla Dhariwal. Improved denoising diffusion probabilistic models. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 8162–8171, 2021.

Junhyung Park, Uri Shalit, Bernhard Schölkopf, and Krikamol Muandet. Conditional distributional treatment effect with kernel conditional mean embeddings and U-statistic regression. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 8401–8412. PMLR, 2021.

Ali Rahimi and Benjamin Recht. Random features for large-scale kernel machines. In *Advances in Neural Information Processing Systems*, volume 20, pp. 1177–1184. Curran Associates, Inc., 2007.

Bharath K. Sriperumbudur, Arthur Gretton, Kenji Fukumizu, Bernhard Schölkopf, and Gert R. G. Lanckriet. Hilbert space embeddings and metrics on probability measures. *Journal of Machine Learning Research*, 11(50):1517–1561, 2010.

Mark J. van der Laan and Daniel Rubin. Targeted maximum likelihood learning. *The International Journal of Biostatistics*, 2(1), 2006. doi: 10.2202/1557-4679.1043. Article 11.

Ted Westling, Mark J. van der Laan, and Marco Carone. Correcting an estimator of a multivariate monotone function with isotonic regression. *Electronic Journal of Statistics*, 14(2):3032–3069, 2020. doi: 10.1214/20-EJS1740.

Elizabeth R. Word, John Johnston, Helen Pate Bain, B. DeWayne Fulton, Jayne Boyd Zaharias, Charles M. Achilles, Martha Nannette Lintz, John Folger, and Carolyn Breda. The state of Tennessee's student/teacher achievement ratio (STAR) project: Technical report (1985–1990). Technical Report ED328356, Tennessee State University, Center of Excellence for Research in Basic Skills, and Tennessee State Department of Education, November 1990.

Houssam Zenati, Bariscan Bozkurt, and Arthur Gretton. Doubly-robust estimation of counterfactual policy mean embeddings. In *Advances in Neural Information Processing Systems*, volume 38, pp. 57044–57100. Curran Associates, Inc., 2025. doi: 10.52202/085713-1709.

APPENDIX CONTENTS

A Existence and recovery **12**
A.1 Log risk and loss separation . . . 12
A.2 Recovery and the two converse constructions . . . 12
A.3 Continuous densities . . . 13
A.4 Quadratic risks and the diffusion construction . . . 14
A.5 Smoothness and the positivity condition . . . 15

B Probability repair and computation **16**
B.1 Proofs of Theorems 5 and 6 . . . 16
B.2 Finite features and probability weights . . . 16
B.3 Approximation and optimization error . . . 17
B.4 Proper scores and CDF projection . . . 18

C Experimental design and additional results **19**
C.1 Data and signed-objective audit . . . 19
C.2 Averaging and resampling nuisance draws . . . 20
C.3 The choice of distance and finite features . . . 20
C.4 Covariates and predictive repair . . . 21
C.5 Generation with a calibrated terminal law . . . 22

D Independent-school risk evaluation **23**
D.1 Target and assumptions . . . 23
D.2 STAR implementation and results . . . 24
D.3 Coverage with fitted retention and risk regressions . . . 24

E Related work **25**

A EXISTENCE AND RECOVERY

A.1 LOG RISK AND LOSS SEPARATION

Proof of Theorem 1. If $q_j < 0$, set $p_j = \epsilon$ and distribute $1 - \epsilon$ over the remaining coordinates in fixed positive proportions. Then $L_q(p) = -q_j \log \epsilon + O(1) \to -\infty$. If $q \geq 0$, then $L_q(p) - L_q(q) = \mathrm{KL}(q\|p) \geq 0$. Equality requires $p = q$, including its zero coordinates. Thus the closed simplex minimum is the entropy of q, and an interior point attains it exactly when $q > 0$. □

Proposition 7 (Loss separation)**.** *Let $Q = Q^+ - Q^-$ be a finite signed measure. If a model class contains g_t such that*

$$\sup_t \int \ell_{g_t}\, dQ^+ < \infty, \qquad \int \ell_{g_t}\, dQ^- \to \infty,$$

then $\int \ell_{g_t}\, dQ \to -\infty$.

Proof. Subtract the second integral from the first. □

For squared loss on distinct sample inputs, an interpolating model can hold its predictions fixed at all positive-weight points and send its prediction at one negative-weight point to infinity. The risk is then $\alpha t^2 + O(t) \to -\infty$, with $\alpha < 0$. For a compact parameter space and a continuous loss, a minimum instead exists by compactness. The quadratic penalty threshold is given below.

A.2 RECOVERY AND THE TWO CONVERSE CONSTRUCTIONS

Proof of Theorem 2. The row score for cell j is

$$\psi_{ij} = \widehat{P}_{X_i}(B_j) + \frac{\mathbf{1}(A_i = a)}{\widehat{e}(X_i)}\{\mathbf{1}(Y_i \in B_j) - \widehat{P}_{X_i}(B_j)\}.$$

Conditional expectation gives

$$\mathbb{E}(\psi_{ij} \mid X_i) - P_{X_i}(B_j) = \frac{\widehat{e}(X_i) - e(X_i)}{\widehat{e}(X_i)}\{\widehat{P}_{X_i}(B_j) - P_{X_i}(B_j)\}. \tag{16}$$

Hence $\mathbb{E}(q_{vj}) = p_{vj} + b_{vj}$, conditional on the fits and labels, and $b \leq c^{-1}\delta_e\delta_P$. Each score lies between $1 - 1/c$ and $1/c$, an interval of length at most $2/c$. Conditional independence, Hoeffding's inequality, and a union bound over the K cells give equation 5. Under equation 6, the maximum error is less than $p_*/2$ with probability tending to one. Theorem 1 gives the optimizer. For a fixed number of cross-fitting folds, apply this argument within each held-out fold and take a finite union bound. □

Proof of Theorem 3(i). Take S independent strata of m rows, with independent $A_{vi} \sim \mathrm{Bernoulli}(1/2)$ and $Z_{vi} = \mathbf{1}(Y^1_{vi} \in B_1) \sim \mathrm{Bernoulli}(p_*)$. Set $\widehat{e} = e = 1/2$ and $\widehat{P}_x(B_1) = 1/2$. The cell score is

$$\psi_{vi} = \tfrac{1}{2} + 2A_{vi}(Z_{vi} - \tfrac{1}{2}) = Z_{vi} + \tfrac{1}{2}R_{vi}, \qquad R_{vi} = (2A_{vi} - 1)(2Z_{vi} - 1).$$

The R_{vi} are independent Rademacher variables, also independent of all outcome indicators. Therefore

$$q_{v1} = \bar{Z}_v + \frac{1}{2m}\sum_{i=1}^{m} R_{vi}. \tag{17}$$

Set $p_* = \kappa\sqrt{\log S/m}$. Bernstein's inequality gives $\max_v |\bar{Z}_v - p_*| = o_p(p_*)$: for example, with $\epsilon = (\log S/m)^{1/8}$, the union bound vanishes because $mp_*\epsilon^2/\log S \to \infty$. On this event, a Rademacher fluctuation below $-2(1 + o(1))mp_*$ makes $q_{v1} < 0$. Stirling's expansion for the binomial tail applies since $2\sqrt{m}\, p_* = O(\sqrt{\log S}) = o(m^{1/6})$, and gives uniformly

$$\mathbb{P}(q_{v1} < 0 \mid Z_{v1}, \ldots, Z_{vm}) \geq S^{-2\kappa^2 + o(1)}.$$

Stratum independence and $\kappa < 1/\sqrt{2}$ imply

$$\mathbb{P}(\min_v q_{v1} \geq 0) \leq \exp\{-S^{1-2\kappa^2+o(1)}\} + o(1) \to 0.$$

The second true mass is $1 - p_*$, so p_* is the minimum mass and $K = 2S$. □

Proof of Theorem 3(ii). Use one stratum with independent $A_i \sim \text{Bernoulli}(1/2)$ and $Z_i \sim \text{Bernoulli}(p)$. Set

$$d = h = n^{-1/8}, \quad p = dh, \quad e = \tfrac{1}{2}, \quad \widehat{e} = \tfrac{1}{2} - d, \quad P(B_1) = p, \quad \widehat{P}(B_1) = p + h.$$

Complementary cell masses make both laws proper for large n, with uniform overlap, $\delta_e = d, \delta_P = h$, and $p_* = p$. For $\psi_i = p + h + A_i(Z_i - p - h)/(1/2 - d)$,

$$\mathbb{E}(q_{11}) = p - \frac{dh}{1/2 - d} = -(1 + o(1))p, \qquad \text{Var}(\psi_i) = \frac{p(1-p)/2 + h^2/4}{(1/2 - d)^2} = O(n^{-1/4}).$$

Thus $\text{Var}(q_{11}) = O(n^{-5/4})$ and

$$\mathbb{P}(q_{11} \geq 0) \leq \frac{\text{Var}(q_{11})}{\{\mathbb{E}(q_{11})\}^2} = O(n^{-3/4}) \to 0.$$

Here the sampling term is $O(n^{-1/2}) = o(p_*)$, while $\delta_e \delta_P = p_*$, isolating the product term. □

A.3 Continuous densities

Fix $0 < h \leq 1$ and a continuous nonnegative smoother K_h on compact $\mathcal{Y} \subset \mathbb{R}^d$, with $\int K_h(y, z)\, dy = 1$. Define

$$f_x(y) = \int K_h(y, z)\, dP_x(z), \qquad \widehat{f}_x(y) = \int K_h(y, z)\, d\widehat{P}_x(z),$$

and the smoothed signed density

$$\widehat{q}_v(y) = \frac{1}{n_v} \sum_{i:V_i = v} [\widehat{f}_{X_i}(y) + w_i\{K_h(y, Y_i) - \widehat{f}_{X_i}(y)\}]. \tag{18}$$

Let p_v be the true density in stratum v, let $p_h(v, y) = \mathbb{E}\{f_X(y) \mid V = v\}$, and set

$$p_* = \inf_{v,y} p_v(y), \qquad \beta = \|p_h - p\|_\infty, \qquad b = \sup_{v,y} \left| \mathbb{E}\left[\frac{\widehat{e} - e}{\widehat{e}} \{\widehat{f}_X(y) - f_X(y)\} \mid V = v \right] \right|.$$

Thus β is smoothing bias and b is conditional product bias.

Theorem 8 (Smoothed continuous log-score recovery)**.** *Suppose evaluation rows are independent of the fits, $e, \widehat{e} \geq c > 0$, and the conditional outcome densities and true and fitted smoothed densities are uniformly bounded in probability. Assume*

$$\|K_h\|_\infty \leq Ch^{-d}, \qquad |K_h(y, z) - K_h(y', z)| \leq Ch^{-d-1}\|y - y'\|.$$

Put $\ell = \log(Sn/h)$. If $\ell = o(n_{\min} h^d)$, then

$$\|\widehat{q} - p\|_\infty = O_p\left\{ \sqrt{\frac{\ell}{n_{\min} h^d}} + \frac{\ell}{n_{\min} h^d} + \beta + b \right\}. \tag{19}$$

If the expression in braces is $o_p(p_*)$, every $\widehat{q}_v$ is a strictly positive continuous density with probability tending to one. Its log risk has the finite, unique almost-everywhere minimizer $g = \widehat{q}_v$ over all densities. A sup-norm compact density sieve with a positive floor attains a minimum; if it contains $\widehat{q}_v$, this is its minimizer.

Proof. Apply equation 16 to the feature $K_h(y, Y)$. The conditional mean of $\widehat{q}_v(y)$ is $p_h(v, y)$ plus the bias defining b. For a row score $\psi_i(y)$ in equation 18, the bounded nuisance envelopes give

$$|\psi_i(y)| \leq Ch^{-d}, \quad \text{Var}\{\psi_i(y) \mid V_i\} \leq Ch^{-d}, \quad |\psi_i(y) - \psi_i(y')| \leq Ch^{-d-1}\|y - y'\|.$$

The variance bound uses $K_h^2 \leq \|K_h\|_\infty K_h$; integration against a probability law preserves the Lipschitz bound. A grid of mesh $h^{d+1}n^{-2}$ has logarithmic cardinality $O(\log(n/h))$. Bernstein's inequality and a union bound over this grid and S strata yield the first two terms of equation 19; interpolation adds $O(n^{-2})$. Tightness of the nuisance envelopes gives the unconditional O_p bound. Adding smoothing and product bias proves the result.

Each row score integrates to one. On the event that its uniform error is below $p_*/2$, the fitted density is continuous and positive. Finally, $L_{\widehat{q}_v}(g) - L_{\widehat{q}_v}(\widehat{q}_v) = \text{KL}(\widehat{q}_v \| g) \geq 0$, with equality only almost everywhere. Compactness and the positive floor give the sieve claim. □

If $\beta = O(h^s)$ and S, h^{-1} grow polynomially, choosing $h \asymp (\log n/n_{\min})^{1/(2s+d)}$ gives the recovery condition

$$(\log n/n_{\min})^{s/(2s+d)} + c^{-1}\delta_e\|\widehat{f} - f\|_\infty = o_p(p_*).$$

A.4 Quadratic risks and the diffusion construction

Theorem 9 (Boundedness of a quadratic loss)**.** *The quadratic in equation 8 is bounded below exactly when $G \succeq 0$ and $b \in \mathrm{Range}(G)$. Its minimizers solve $G\beta = b$, and the minimizer is unique exactly when $G \succ 0$.*

Proof. A negative eigenvector u gives $J(tu) = t^2u^\top Gu - 2tb^\top u + C \to -\infty$. If $G \succeq 0$ but $b \notin \mathrm{Range}(G)$, some $u \in \ker G$ has $b^\top u \neq 0$, so one sign of t again gives divergence. Otherwise, with $\beta_0 = G^\dagger b$,

$$J(\beta) = J(\beta_0) + (\beta - \beta_0)^\top G(\beta - \beta_0).$$

This proves boundedness and the minimizer statements. □

Proposition 10 (Quadratic regularization)**.** *For $\lambda \geq 0$, let $J_\lambda(\beta) = J(\beta) + \lambda\|\beta\|^2$ and $\lambda_* = \max\{0, -\lambda_{\min}(G)\}$. Below λ_* the objective is unbounded; above it the unique minimizer is $(G + \lambda I)^{-1}b$. At λ_*, boundedness holds exactly when $b \in \mathrm{Range}(G + \lambda_* I)$. As $\lambda \downarrow \lambda_*$, the minimizer diverges along a critical eigendirection exactly when b has a nonzero component in that direction.*

Proof. Apply Theorem 9 to $G + \lambda I$. Along an eigenvector with eigenvalue $-\lambda_*$, the minimizing coordinate is $b_j/(\lambda - \lambda_*)$. □

The affine parameter subspace. The released diffusion model uses a hypernetwork to generate the weights (W_1, b_1, W_2, b_2) of a two-layer denoiser. For a scalar outcome its denoiser is

$$s = \mathrm{ELU}(W_1 x + b_1), \qquad \widehat{\epsilon} = W_2 s + b_2.$$

The hypernetwork is implemented by Pyro 1.9.1's `DenseNN`, whose output layer is affine. The supplement gives the correspondence to the released model classes. With the last context-hidden vector augmented by an intercept $u \in \mathbb{R}^{11}$, write $W_{2j} = \theta_j^\top u$, $b_2 = \theta_0^\top u$. Fixing the trunk and the heads for W_1, b_1 gives

$$\widehat{\epsilon} = \sum_{j=1}^{10} s_j\theta_j^\top u + \theta_0^\top u = h^\top\beta, \qquad h = (s \otimes u, u) \in \mathbb{R}^{121}.$$

These are global parameter coordinates. For fixed draws, or after integration, their signed loss is exactly equation 8.

Proof of Proposition 4. Set one context-hidden coordinate to $\mathrm{ELU}(A) = A$ and all others to zero. Freeze a first-layer denoiser row to select only the noised outcome. Varying one generated-W_2 coefficient then contributes $\beta\,\mathrm{ELU}(Y_t)$ under action 1 and zero under action 0, where

$$Y_t = \sqrt{\bar{\alpha}_t}Y + \sqrt{1 - \bar{\alpha}_t}\epsilon.$$

This coordinate ignores the representation, so context jitter has no effect. For the released $T = 100$ schedule, define the integrated feature moment

$$r(y) = \frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}_\epsilon[\mathrm{ELU}(\sqrt{\bar{\alpha}_t}y + \sqrt{1 - \bar{\alpha}_t}\epsilon)^2].$$

It is finite and continuous, and $r(y) \to \infty$ as $y \to \infty$: every $\bar{\alpha}_t > 0$, and ELU grows linearly on the positive half-line. Let p be a bounded smooth compactly supported density, and write

$$p_L(y) = p(y - L), \qquad R_0 = \int r(y)p(y)\,dy, \qquad R(L) = \int r(y)p_L(y)\,dy.$$

The function R is continuous and diverges. Fix $e \in (0, 1)$, $\kappa > eR_0$, and $1/4 < \gamma < 1/2$. For all sufficiently large n, choose

$$d = n^{-\gamma}, \qquad \widehat{e} = e - d, \qquad \widehat{p} = (1 - d)p + dp_L, \qquad R(L) = \kappa n^{2\gamma}.$$

Then $\widehat{e} \geq e/2$. The density is proper and uniformly smooth, with $\|\widehat{p} - p\|_{L_4} \leq 2d\|p\|_{L_4}$, while $|\widehat{e} - e| = d$, so the error product is $O(n^{-2\gamma}) = o(n^{-1/2})$. Put $\widehat{R} = (1-d)R_0 + \kappa n^{\gamma}$. After integrating nuisance outcomes, time, noise, and jitter, the expected Gram coefficient is

$$\bar{G} = R_0 + \frac{d}{e-d}(R_0 - \widehat{R}) \longrightarrow R_0 - \kappa/e < 0. \tag{20}$$

Its empirical counterpart is

$$G_n = \widehat{R} + \frac{1}{n(e-d)} \sum_{i=1}^{n} A_i\{r(Y_i) - \widehat{R}\}.$$

More explicitly,

$$G_n - \bar{G} = \frac{n^{-1}\sum_i A_i r(Y_i) - eR_0}{e-d} - \frac{\widehat{R}}{e-d}\left(n^{-1}\sum_i A_i - e\right).$$

Compact support makes the first term $O_p(n^{-1/2})$. Since $\widehat{R} = O(n^{\gamma})$, the second is $O_p(n^{\gamma-1/2}) = o_p(1)$. Thus $G_n - \bar{G} = o_p(1)$, and $G_n < 0$ with probability tending to one. The linear and constant terms are finite, so Theorem 9 completes the proof. □

A.5 Smoothness and the positivity condition

Remark 1 and Eq. (86) of Melnychuk & Feuerriegel (2026) use the lower bound

$$1 + \left\{\frac{\mathbf{1}(A=a)}{\widehat{e}(x)} - 1\right\}\left\{1 - \frac{\widehat{p}_x(y)}{p_x(y)}\right\} \geq \delta > 0 \qquad \text{where } p_x(y) > 0.$$

The remark derives asymptotic positivity from Hölder smoothness and a vanishing L_4 nuisance product. The next construction places smooth estimation errors in a covariate region of mass $O(1/n)$. The ratio is negative there, and a sample of size n encounters an observed negative ratio with probability bounded away from zero.

Theorem 11 (Unbounded log loss with fitted densities)**.** *There is a triangular array with an unbounded covariate space, uniformly smooth bounded nuisance densities, and $\|\widehat{e} - e\|_{L_4}\|\widehat{p} - p\|_{L_4} = O(n^{-1/2})$, for which a sample-separating categorical model with $V = X$ has unbounded integrated log risk with probability tending to one. On covariate regions of probability λ/n, its treated-row pseudo-density ratio is negative over an outcome set of fixed positive probability. An evaluation sample encounters an observed negative ratio with nonvanishing probability.*

Proof. Give X a positive smooth density with right tail Ce^{-x}. Translate a smooth bump $b : \mathbb{R} \to [0,1]$, supported on $[0,2]$ and equal to one on $[1/2, 3/2]$, so that its plateau E_n has probability λ/n. Its full support U_n has probability $O(n^{-1})$; translation preserves derivative bounds. Denote the translated bump by b_n.

Let p_B, p_C be bounded smooth densities on disjoint compact sets B, C. With $r = 0.2, \rho = 0.4$, set

$$e = \tfrac{1}{2}, \qquad \widehat{e} = \tfrac{1}{2} - \tfrac{1}{4}b_n, \qquad p_x = rp_B + (1-r)p_C, \qquad \widehat{p}_x = p_x + (\rho - r)b_n(x)(p_B - p_C).$$

Both laws are proper and all derivatives are uniformly bounded. Each nuisance error has L_4 norm $O(n^{-1/4})$, since it is bounded and supported where $X \in U_n$. The density norm uses the covariate law times a fixed dominating outcome measure.

Outside U_n, a row with $A = a, Y \in C$ assigns signed mass $-r$ to B. Its probability tends to $(1-r)/2 > 0$, so such a row occurs with probability tending to one. Observed covariates are distinct almost surely; sample separation and Theorem 1 give unbounded log risk. On $E_n \times B$, the inverse fitted propensity for a treated row is four and $\widehat{p}/p = \rho/r = 2$, so the displayed ratio equals $1 + (4-1)(1-2) = -2$. A row has $X \in E_n, A = a, Y \in B$ with probability $\lambda r/(2n)$. The probability of observing at least one such row is

$$1 - \{1 - \lambda r/(2n)\}^n \longrightarrow 1 - e^{-\lambda r/2} > 0.$$

□

Uniform relative accuracy supplies the positivity condition directly. Under $\widehat{e} \geq c > 0$, the absolute difference between the displayed ratio and one is at most

$$c^{-1} \sup_{x,y:p_x(y)>0} \left| \frac{\widehat{p}_x(y)}{p_x(y)} - 1 \right|.$$

If this bound is $o_p(1)$, the ratio exceeds any fixed $\delta < 1$ with probability tending to one, uniformly on the true support. For the empirical target, Theorem 2 additionally controls sampling error relative to cell mass; Theorem 8 gives the corresponding density condition with smoothing bias and bandwidth.

B Probability repair and computation

B.1 Proofs of Theorems 5 and 6

Weak compactness of $\mathcal{P}(\mathcal{Y})$ and continuity of the bounded kernel make $P \mapsto \mu_P$ continuous in Hilbert norm: expand $\|\mu_{P_n} - \mu_P\|^2 = \iint k \, d(P_n - P) \, d(P_n - P)$ and use weak convergence of the product measures. Thus $\mathcal{M}$ is compact; linearity makes it convex. The Hilbert projection theorem (Bauschke & Combettes, 2017) gives a unique, 1-Lipschitz, hence measurable projection. Its variational inequality is

$$\langle z - \Pi_{\mathcal{M}} z, \mu_0 - \Pi_{\mathcal{M}} z \rangle \leq 0, \qquad \mu_0 \in \mathcal{M}.$$

Expanding $\|z - \mu_0\|^2$ proves equation 11. A characteristic kernel makes $P \mapsto \mu_P$ injective (Sriperumbudur et al., 2010), proving the law-uniqueness claim. The correction $\|\widehat{\mu} - \widetilde{\mu}\| = \operatorname{dist}(\widehat{\mu}, \mathcal{M})$ is 1-Lipschitz in the input and vanishes exactly on $\mathcal{M}$.

For Theorem 6, conditional expectation in equation 9 gives

$$\mathbb{E}(\psi \mid X) = \widehat{m}(X) + \frac{e(X)}{\widehat{e}(X)} \{m(X) - \widehat{m}(X)\}.$$

Subtracting $m(X)$ proves equation 13; conditioning on V gives $\bar{\mu} - \mu = b$. Conditional Jensen and Hölder yield

$$\|b\|_{L_2} \leq \left\| \frac{\widehat{e} - e}{\widehat{e}} (\widehat{m} - m) \right\|_{L_2} \leq c^{-1} \|\widehat{e} - e\|_{L_4} \|\widehat{m} - m\|_{L_4}.$$

Since $\mu(v) \in \mathcal{M}$, nonexpansiveness gives $\|\widetilde{\mu}(v) - \mu(v)\| \leq \|\widehat{\mu}(v) - \mu(v)\|$. The L_2 triangle inequality completes the proof.

B.2 Finite features and probability weights

In this subsection take a continuous feature map $\phi : \mathcal{Y} \to \mathbb{R}^D$. Use the same definitions of $m, \mu, \psi, \widehat{\mu}, \bar{\mu}, b$ as in Section 4, now with Euclidean norms. For support $s_1, \ldots, s_m$, write $F = [\phi(s_1), \ldots, \phi(s_m)]$ and $C = F\Delta_m$. The computation is

$$\widehat{w}(v) \in \operatorname*{arg\,min}_{w \in \Delta_m} \|Fw - \widehat{\mu}(v)\|^2, \qquad \widehat{P}_v = \sum_j \widehat{w}_j(v) \delta_{s_j}, \qquad \widetilde{\mu}(v) = F\widehat{w}(v). \tag{21}$$

Theorem 12 (Projection with finite features)**.** *If $\mathcal{Y}$ is compact, then $\mathcal{M} = \operatorname{conv}\{\phi(y) : y \in \mathcal{Y}\}$ is compact and convex. Theorems 5–6 hold in $\mathbb{R}^D$. For the support-constrained fit equation 21, let $\delta = \|\Pi_C \mu - \mu\|_{L_2}$. Then*

$$\|\widetilde{\mu} - \mu\|_{L_2} \leq \|\widehat{\mu} - \bar{\mu}\|_{L_2} + \|b\|_{L_2} + \delta \leq \|\widehat{\mu} - \bar{\mu}\|_{L_2} + c^{-1} \|\widehat{e} - e\|_{L_4} \|\widehat{m} - m\|_{L_4} + \delta. \tag{22}$$

Every optimizing weight vector defines a probability law.

Proof. Carathéodory's theorem makes the convex hull of the compact image $\phi(\mathcal{Y})$ compact. Every point of this hull is a discrete probability mean, and every probability mean lies in the hull. The preceding projection and conditional-expectation arguments apply unchanged. For the restricted support, $\|\Pi_C \widehat{\mu} - \Pi_C \mu\|_{L_2} \leq \|\widehat{\mu} - \mu\|_{L_2}$; add δ and apply Theorem 6. □

Proposition 13 (Weight selection and identified functionals). *Let $A = [\mathbf{1}^\top; F]$ and $\mathcal{W}(z) = \{w \in \Delta_m : Fw = z\}$, for $z \in C$. The selection*

$$w^\dagger(z) = \operatorname*{arg\,min}_{w \in \mathcal{W}(z)} \|w\|^2 \tag{23}$$

is unique and Borel measurable. Every $h \in \mathrm{Range}(A^\top)$ has constant expectation $h^\top w$ over $\mathcal{W}(z)$. This condition is also necessary when the set contains a strictly positive vector. Every $z \in C$ has a unique weight representation if and only if $\ker A = \{0\}$.

Proof. The feasible set is nonempty, compact, and convex. Strict convexity gives uniqueness, and the measurable maximum theorem gives measurability. If $h = A^\top u$, then $h^\top w = u^\top(1, z^\top)^\top$. Conversely, at a strictly positive w_0, every $d \in \ker A$ permits both perturbations $w_0 \pm td$ for small $t > 0$. Constant expectation therefore requires $h \perp \ker A$. The same perturbation of the uniform weight vector proves the last claim. □

At tolerance ϵ, define $\mathcal{W}_\epsilon(z) = \{w \in \Delta_m : \|Fw - z\|_\infty \le \epsilon\}$. Two linear programs give the smallest and largest possible $h^\top w$. For any α, β, their difference is at most

$$2\|h - \alpha\mathbf{1} - F^\top\beta\|_\infty + 2\epsilon\|\beta\|_1. \tag{24}$$

Indeed, subtract $\alpha\mathbf{1} + F^\top\beta$ from h and use $\|w - w'\|_1 \le 2$ and $\|F(w - w')\|_\infty \le 2\epsilon$. Appendix C.3 applies this calculation to a tail probability.

B.3 Approximation and optimization error

For $f(w) = \|Fw - z\|^2$, a feasible numerical solution has the certificate

$$g = 2F^\top(Fw - z), \qquad \gamma = g^\top w - \min_j g_j, \qquad \max\{0, f(w) - \gamma\} \le f^* \le f(w). \tag{25}$$

Convexity bounds f below by its tangent, whose minimum over the simplex is $f(w) - \gamma$. The projection inequality also gives $\|Fw - \Pi_C z\| \le \sqrt{\gamma}$. For $f(w) + \lambda\|w\|^2$, $\lambda > 0$, use $g + 2\lambda w$ in the same certificate to obtain γ_λ. Its unique minimizer w_λ satisfies

$$\|w - w_\lambda\| \le \sqrt{\gamma_\lambda/\lambda}, \qquad f(w_\lambda) - f^* \le \lambda.$$

The first bound follows from quadratic growth; the second compares with an unpenalized minimizer and uses $\|w\|^2 \le 1$.

To account for preprocessing, let z be the Hilbert-space target actually supplied to the projection and put $\tau = \|z - \widehat{\mu}\|_{L_2}$. If a map T moves outcomes by at most h, a Lipschitz feature map gives

$$\left\| \int \{\phi(T(y)) - \phi(y)\}\, dQ_v(y) \right\| \le Lh\|Q_v\|_{\mathrm{TV}}, \qquad \tau \le Lh\|\|Q_V\|_{\mathrm{TV}}\|_{L_2}.$$

Analytic integration of continuous nuisance features gives $\tau = 0$.

Corollary 14 (Rate in the original kernel metric). *Let $\mathcal{Y} = [0,1]^d$, with $\|\phi(y) - \phi(y')\| \le L\|y - y'\|$, and suppose*

$$\|\widehat{\mu} - \bar{\mu}\|_{L_2} + \|b\|_{L_2} + \tau = O_p(r), \qquad r \to 0.$$

Draw an independent m-point support from a density bounded below. Use D random features whose kernel satisfies $\sup_{y,y'} |k_D(y,y') - k(y,y')| = O_p\{\sqrt{\log(n+m)/D}\}$. Assume the signed representations of the supplied targets have uniformly bounded total variation and that their finite features are evaluated from those same representations. Let $\epsilon(v)$ bound the numerical squared-objective gap, and let ϵ_{gen} be the L_2 embedding error of an optional sampler relative to the computed support law $\widehat{P}$. Then the returned law has

$$\|\mu_{\mathrm{out}} - \mu\|_{L_2} = O_p\left\{ r + \left(\frac{\log m}{m}\right)^{1/d} + \left(\frac{\log(n+m)}{D}\right)^{1/4} + \|\sqrt{\epsilon}\|_{L_2} + \epsilon_{\mathrm{gen}} \right\}. \tag{26}$$

Thus $m \gtrsim r^{-d}\log(1/r)$, $D \gtrsim r^{-4}\log(n+m)$, $\|\epsilon\|_{L_1} = O_p(r^2)$, and $\epsilon_{\mathrm{gen}} = O_p(r)$ preserve rate r. For $r = n^{-1/4}$, one may take $m \asymp n^{d/4}\log n$, $D \asymp n\log n$, and $\|\epsilon\|_{L_1} = O_p(n^{-1/2})$. In one dimension, quantiles of an independent proposal sample of size n permit $m \asymp n^{1/4}$. Direct sampling from $\widehat{P}$ gives $\epsilon_{\mathrm{gen}} = 0$.

Proof. For the support fill distance $H = \sup_y \min_j \|y - s_j\|$, a grid cover gives $\mathbb{P}(H > u) \le C_0 u^{-d} \exp(-c_0 m u^d)$. Thus $H = O_p\{(\log m/m)^{1/d}\}$. Moving a probability law to its nearest support points changes its embedding by at most LH.

Put $\zeta = \sup |k_D - k|$. For two signed measures of total variation at most B, their squared embedding distances differ by at most $4B^2\zeta$; their distances therefore differ by at most $2B\sqrt{\zeta}$. Apply Theorem 12 in the feature metric, then transfer the output error, input error, and support approximation to the original kernel metric. Together with the numerical and sampling errors this gives

$$\|\mu_{\text{out}} - \mu\|_{L_2} \le \|\widehat{\mu} - \bar{\mu}\|_{L_2} + \|b\|_{L_2} + \tau + LH + 6B\sqrt{\zeta} + \|\sqrt{\epsilon}\|_{L_2} + \epsilon_{\text{gen}}. \tag{27}$$

Here $B \ge 1$ also bounds probability measures. Substitution proves the rate and the stated choices. For scalar quantile support, the uniform empirical-CDF bound and the lower density bound give $H = O_p(m^{-1} + n^{-1/2})$. The assumed random-feature approximation is the usual uniform-kernel condition (Rahimi & Recht, 2007). □

B.4 Proper scores and CDF projection

For any unit-mass finite signed measure Q, the kernel score satisfies

$$\int \|\mu_P - \phi(y)\|^2 \, dQ(y) = \|\mu_P - \mu_Q\|^2 + \int k(y,y) \, dQ(y) - \|\mu_Q\|^2. \tag{28}$$

Thus projection of cell averages minimizes this signed proper score. For scalar outcomes, let $H(t) = Q((-\infty, t])$ and assume $\int |y| \, d|Q|(y) < \infty$. Using the loss convention for CRPS (Gneiting & Raftery, 2007) and expanding its squared indicator difference gives, for every P with finite first moment,

$$\int \text{CRPS}(P, y) \, dQ(y) = \int \{F_P(t) - H(t)\}^2 \, dt + \int \{H(t) - H(t)^2\} \, dt. \tag{29}$$

The moment assumption justifies interchanging the integrals. The feature $\phi_y(t) = \mathbf{1}(y \le t) - \mathbf{1}(0 \le t)$ lies in $L_2(\mathbb{R})$; on compact outcome domains, Theorems 5–6 apply to this CDF distance.

Proposition 15 (CDF projection and numerical bounds)**.** *Let $s_1 < \cdots < s_m$, $m \ge 2$, $\Delta_j = s_{j+1} - s_j$, and $u_j = \Delta_j^{-1} \int_{s_j}^{s_{j+1}} H(t) \, dt$. The unique minimizer of signed CRPS on this support has cumulative weights*

$$c^* = \operatorname*{arg\,min}_{0 \le c_1 \le \cdots \le c_{m-1} \le 1} J(c), \qquad J(c) = \sum_j \Delta_j (c_j - u_j)^2. \tag{30}$$

For feasible w, set $c_j = \sum_{l \le j} w_l$, $g_l = 2 \sum_{j \ge l} \Delta_j (c_j - u_j)$, $g_m = 0$, and $\gamma = g^\top w - \min_l g_l$. Then

$$\sum_j \Delta_j (c_j - c_j^*)^2 \le J(c) - J(c^*) \le \gamma, \qquad |c_j - c_j^*| \le \sqrt{\gamma / \Delta_j}. \tag{31}$$

The resulting laws obey

$$d_{\text{TV}}(P_w, P_{w^*}) \le \sqrt{\gamma \sum_{j=1}^{m-1} \Delta_j^{-1}}.$$

For $s_j < \tau \le s_{j+1}$, the numerical error in $P(Y \ge \tau)$ is at most $\sqrt{\gamma/\Delta_j}$. Put $l_j = \max(0, c_j - \sqrt{\gamma/\Delta_j})$, $r_j = \min(1, c_j + \sqrt{\gamma/\Delta_j})$, and $l_m = r_m = 1$. The exact fitted p-quantile, $0 < p < 1$, lies in

$$\left[s_{\min\{j : r_j \ge p\}}, \; s_{\min\{j : l_j \ge p\}}\right].$$

Proof. The CDF equals c_j between consecutive support points, so completing the square gives $J(c)$ plus a constant. Positive Δ_j imply strict convexity. Weighted isotonic regression followed by truncation to $[0, 1]$ gives the minimizer. The projection inequality and the simplex tangent bound prove equation 31. For $e_j = c_j - c_j^*$, with $e_0 = e_m = 0$, $\frac{1}{2} \sum_l |e_l - e_{l-1}| \le \sum_j |e_j|$. Weighted Cauchy–Schwarz proves the total-variation bound. The remaining bounds follow by reading and inverting the CDF envelopes. □

Gaussian targets are integrated analytically using

$$\int \Phi((t-\mu)/\sigma)\,dt = (t-\mu)\Phi((t-\mu)/\sigma) + \sigma\varphi((t-\mu)/\sigma).$$

A factual point mass contributes $(t-y)_+$. The reported numerical bounds include a double-precision arithmetic allowance and concern the exact minimizer of the computed program.

Positive part and the choice of distance. Write the Jordan decomposition of a mass-one signed measure as $Q = Q^+ - Q^-$, with $m = Q^-(\mathcal{Y})$, and put $P_+ = Q^+/(1+m)$. For any probability law P, the negative part of $Q - P$ has mass at least m. Since $Q - P$ has mass zero,

$$d_{\mathrm{TV}}(Q,P) := \tfrac{1}{2}|Q-P|(\mathcal{Y}) \geq m = d_{\mathrm{TV}}(Q,P_+).$$

Thus positive part is a total-variation minimizer. If Q and P are supported on $[a,b]$, its CDF error also satisfies

$$\|F_{P_+} - F_P\|_2 \leq \|H - F_P\|_2 + m\sqrt{b-a}.$$

Indeed, for $m > 0$, $Q - P_+ = m(P_+ - Q^-/m)$, whose CDF is bounded by m in absolute value. Projection in CDF distance removes this additional negative-mass term through Theorem 5.

For the three-point example in the main text, H equals ε on $[0,h)$ and zero on $[h,1)$. Within each interval, replacing any nondecreasing CDF by its average lowers squared error and preserves order. The projection therefore reduces to two coordinates with weights h and $1-h$. Isotonic regression pools them at εh, proving the stated law and all three error formulas. These formulas hold for projection over all laws on $[0,1]$, as well as over the common support $\{0,h,1\}$. With $\varepsilon = 0.1$ and $h = 0.01$, the target, CDF projection, and positive part have squared CDF errors 10^{-4}, 10^{-6}, and approximately 8.264×10^{-3}, respectively. For $h > 1/(1+\varepsilon)$, the same formulas give positive part the smaller error. The supplement checks both orderings, using $\varepsilon \in \{0.05, 0.1, 0.5\}$ and $h \in \{0.99, 0.9, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$.

C Experimental design and additional results

C.1 Data and signed-objective audit

We use the AER/Rdatasets version of Project STAR, whose randomized design is described by Word et al. (1990); Krueger (1999). Of 6,325 kindergarten-assigned pupils, 5,785 have the retained baseline fields and reading outcome. Action 1 is small class; action 0 combines regular and regular-with-aide classes in their design proportions. The objective and repair audits use these complete cases in 79 schools. A 30% holdout within school and arm (seed 5601) gives 4,048 training and 1,737 held-out pupils. The training mean and standard deviation of reading scores are 437.12 and 32.27. The covariate design has 105 columns including school identifiers. Two stratified training folds fit arm-specific Gaussian ridge regressions with penalty 10 and fold-specific residual scales. The assignment weights use school fractions computed before outcome selection; their range, 0.1613–0.4426, lies inside the clipping interval $[0.05, 0.95]$. The conditioning variable V combines gender, coarsened ethnicity, and school type. Ten cells have at least 60 training pupils and, in each arm, 12 held-out pupils. Held-out losses in this audit use complete-case cell frequencies. The causal evaluation with outcome-selection adjustment is specified separately in Appendix D.

The GDR implementation of Melnychuk & Feuerriegel (2026)[1] uses one nuisance draw per action and epoch to approximate equation 2, retaining the coefficients w_i and $1 - w_i$. We compare this sampled objective with analytic Gaussian integration on 32 outcome bins, using identical rows, nuisances, batches, and cells. Table 1 reports the negative-mass frequency and the log-score change when a negative cell's probability floor decreases from 10^{-3} to 10^{-12}. Both full-training integrated action targets have negative mass.

For the affine-head audit, the released architecture has representation, context-hidden, and denoiser-hidden widths 10, time embedding dimension 20, 100 diffusion steps, context jitter 0.1, and batch size 64. Fixing the trunk at initialization seed 5617 gives the 121 coordinates in Appendix A.4. We

[1]The GDR-learners source code is linked at the version used in our analysis. The supplement records the full version identifier and its correspondence to the mathematical objective.

sum both action losses and average M independent draws of nuisance outcomes, diffusion time, Gaussian noise, and context jitter. Table 4 reports all five integration resolutions. At 512 draws, an eight-block jackknife gives the full-data minimum eigenvalue interval $[3.90, 4.03] \times 10^{-6}$; nine of 63 batches and one of ten cells have intervals strictly below zero. The largest batch penalty threshold is $\lambda_* = 0.015744$ under the stated loss scaling.

Table 4: Affine-head signed curvature. The last two columns give the fraction of matrices with a negative minimum eigenvalue.

Draws M	Full-data $\lambda_{\min}$	64-row batches	V-cells
1	-6.30×10^{-4}	100.0%	100.0%
8	3.95×10^{-6}	100.0%	60.0%
32	3.83×10^{-6}	85.7%	20.0%
128	3.98×10^{-6}	38.1%	20.0%
512	3.96×10^{-6}	34.9%	20.0%

C.2 Averaging and resampling nuisance draws

We train the same affine head for 200 epochs, with 63 updates per epoch, AdamW step size 0.005, weight decay 0.01, and exponential moving average (EMA) decay 0.995. The four settings combine one or 32 nuisance draws with either fixed draws or new draws each epoch. Eight paired seeds share initialization perturbations and minibatch order; each single draw is the first of the paired set of 32. We evaluate excess loss relative to the exact minimizer of an independent quadratic formed from 256 draws.

Table 5: Affine-head outcomes after 12,600 updates. Dispersion is the standard deviation across eight seeds. Multi-draw branches use 32 times as many sampled contributions per update.

Mechanism	Median excess	Excess dispersion	Median EMA norm
Fixed, 1 draw	0.076524	0.051530	5.605
Fixed, 32 draws	0.003943	0.000414	3.618
Refreshed, 1 draw	0.006022	0.000582	2.692
Refreshed, 32 draws	0.001759	0.000048	3.321

Both averaging and refresh reduce the median excess in this comparison. The supplement contains the seed-level trajectories, spectra, and integration diagnostics.

C.3 The choice of distance and finite features

All repairs use the same signed target, which combines fitted Gaussian laws with point masses at observed outcomes, and 60 training-quantile support points. RFF projection uses 64 Fourier features. For Gaussian outcomes on the standardized scale, each feature mean is evaluated as

$$\mathbb{E}\{\sqrt{2/D}\cos(\omega Y + b)\} = \sqrt{2/D}\exp(-\omega^2\sigma^2/2)\cos(\omega\mu + b).$$

Mapped RFF first assigns every outcome to its nearest support point. Continuous RFF uses the displayed expectation; continuous CDF uses equation 30. Positive part clips and normalizes the signed Voronoi masses. Table 6 measures every correction from the continuous input, including the movement caused by mapping. The positive-part to mapped-RFF correction ratios are 1.110 and 1.027. Continuous RFF reduces that correction further, while its CRPS and threshold error depend strongly on the weight penalty.

For the fixed feature matrix $F \in \mathbb{R}^{64\times 60}$, the augmented matrix $A = [\mathbf{1}^\top; F]$ has numerical rank 17 at singular-value threshold 1.42×10^{-13}, and condition number 8.6×10^{17}. The rank threshold is $65\epsilon_{\text{mach}}\sigma_1(A)$, using double precision. In the male, Caucasian, suburban cell under action 1, two nonnegative weight vectors summing to one produce embeddings within coordinate tolerance 10^{-8} of the same fitted embedding. Their probabilities of a reading score of at least 460 are 0.11146 and 0.50979. Their embedding distance is 6.91×10^{-9} (Figure 2). At tolerance 10^{-10}, a feasible pair

Table 6: Common continuous STAR target and 60-point support. Each entry gives action 0 / action 1. Correction includes input mapping; tail error uses the held-out complete-case probability at 460. Continuous RFF uses weight penalty 10^{-12} unless shown otherwise.

Fit	10^3 RFF correction	CRPS	Tail error (%)
Mapped RFF	5.434 / 7.752	15.830 / 16.188	4.00 / 6.98
Continuous RFF	1.227 / 1.856	16.953 / 17.181	7.93 / 8.21
RFF, $\lambda = 10^{-4}$	1.353 / 1.958	15.876 / 16.221	2.87 / 5.77
Continuous CDF	4.076 / 5.758	15.830 / 16.184	3.99 / 7.15
Positive part	6.033 / 7.959	15.825 / 16.183	3.95 / 7.16
Plug-in	89.077 / 89.607	16.018 / 16.285	5.67 / 8.07

still gives 0.12412 and 0.48228. These calculations evaluate the functional sets in Proposition 13 and equation 24.

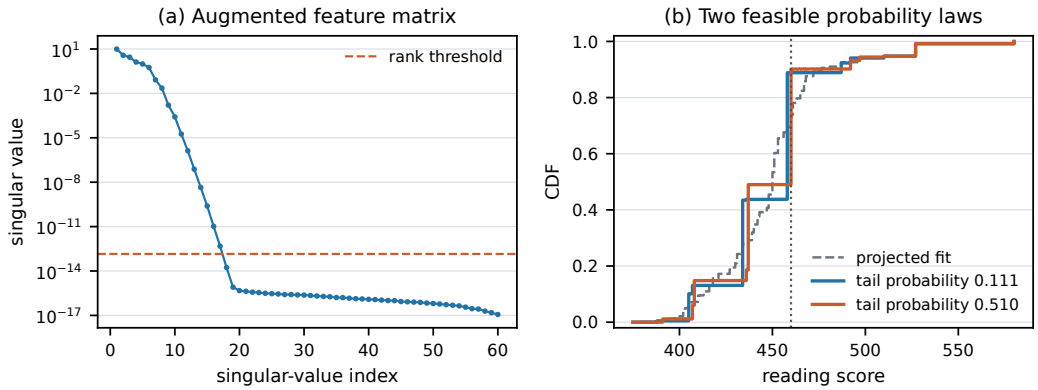

Figure 2: Finite-feature resolution. Left: singular values of $[\mathbf{1}^\top; F]$ and the numerical-rank threshold. Right: CDFs of the fitted law and two feasible laws at coordinate tolerance 10^{-8}; the vertical line marks reading score 460.

Evaluation with the complete Gaussian kernel at bandwidths $0.5s_Y, s_Y, 2s_Y$, $s_Y = 32.27$, places the mapped repair below plug-in and within 0.00047 of positive part in weighted root MMD. For the displayed feasible pair, complete-kernel MMD at s_Y is 2.25×10^{-4}, and an independent 4,096-feature estimate is 3.08×10^{-5}. For this cell, six combinations of SLSQP initialization and tolerance give fitted effective support sizes ranging from 39.50 to 48.31. Certified refits of the direct-feature and common-Gaussian targets have maximum unpenalized objective gaps 4.88×10^{-13} and 1.25×10^{-12}. The supplement reports the penalized gaps as well.

Two rounds of midpoint insertion give nested supports of 60, 119, and 237 points with fixed endpoints. Across the 60 CDF programs, the largest certified numerical errors are 6.23×10^{-6} in CDF, 4.41×10^{-6} in the tail at 460, and 7.41×10^{-5} in total variation; all 180 quantile brackets at levels 0.1, 0.5, and 0.9 contain one support point. Full support-refinement and solver results accompany the fitted weights in the supplement.

C.4 Covariates and Predictive Repair

The target is the marginal law of Y^1 in a single conditioning group. Generate independent $X, \epsilon \sim N(0,1)$, $A \sim \text{Bernoulli}(e)$, and $Y^a = \sqrt{r}X + \sqrt{1-r}\epsilon$ for both actions. An auxiliary sample of size n fits a linear Gaussian regression using its treated observations; residual variance uses the residual degrees of freedom. A second independent sample of size n forms the signed target with known propensity e. Covariates enter the nuisance regression; every returned distribution targets the same marginal $N(0,1)$ law. The mean-shift extension adds 0.35 to the fitted regression.

Continuous CDF projection integrates the fitted Gaussian nuisance exactly. Positive part clips and normalizes its signed Voronoi masses on the same equally spaced support in $[-6, 6]$. Mapped CDF projects these same masses, isolating the repair from the preceding input mapping. Expected CRPS excess is analytic against $N(0, 1)$; direct integration of squared CDF error agrees within 1.2×10^{-15}. The plug-in comparison averages the fitted conditional Gaussian laws on the common support.

Development uses 32 seeds over $n \in \{64, 256, 1024\}$, $e \in \{0.1, 0.25, 0.5\}$, and $r \in \{0, 0.5, 0.9, 0.99\}$. Independent confirmation retains all four r's at $n = 256$, both $e = 0.1, 0.25$, and 65, 129, or 257 support points, using 1,000 new seeds per setting. The mean-shift extension has a separate eight-setting development run and 1,000 new seeds per setting, with support checks at $e = 0.25$, $r = 0.9, 0.99$. The supplement retains every attempted configuration and seed-level score. Both high-r advantages in Table 2, including the mapped-CDF contrasts, survive a Bonferroni correction across all 144 confirmation contrasts in this experiment and the extensions below.

For a finite-draw target, the fitted nuisance draws are shared by both repairs. We cross $n = 256, 1024$, $r = 0, 0.5, 0.9$, and 1, 8, or 32 draws per covariate at $e = 0.25$, with the same 0.35 mean shift and 1,000 new seeds. At $n = 1024, r = 0.9$ with one draw, mean CRPS excess is 0.002007 for continuous CDF, 0.002097 for mapped CDF, and 0.002304 for positive part on the 129-point grid. The CDF–positive-part difference is -0.000298 $[-0.000344, -0.000251]$. On the common support of all sampled atoms, CDF projection and positive part have errors 0.001868 and 0.007265, respectively. Mapping nearby positive and negative atoms to a common grid allows them to cancel before clipping, substantially reducing positive-part error.

We also evaluate two public datasets. The `speff2trial` release of ACTG 175 (Hammer et al., 1996) contains 2,139 participants. We reserve 642 for evaluation and use the other 1,497 for nuisance selection. With 129 support points, the selected CD4 comparison gives a CRPS difference of -0.0182 $[-0.0481, 0.0118]$, in original CD4 units. For the public release of the IHDP semi-synthetic benchmark (Hill, 2011), the first 20 replications select the nuisance model and the next 200 evaluate it. At 129 support points, CDF and positive part give mean CRPS 0.33680 and 0.33640 in training-standardized outcome units. Positive part has lower error at all three support resolutions. Data sources, baseline variables, cross-fitting, reserved splits, and all development results are specified in the supplementary predictive-comparison protocol.

C.5 Generation with a calibrated terminal law

All methods condition on $V \in \{0, 1, 2\}$, with $n_v = 32$ or 256 observations per level, $A \sim \text{Bernoulli}(0.25)$, and

$$Y^a \mid V = v \sim N(\mu_v + 0.4a, \sigma_v^2), \quad (\mu_0, \mu_1, \mu_2) = (-0.8, 0, 0.8), \quad \sigma_v = 0.6 + 0.15\mathbf{1}(v = 1).$$

The common outcome nuisance shifts the mean by $0.35(1 - 2a)$, retaining the true scale and propensity. The neural branches train all 2,811 network parameters with AdamW step size 0.005, weight decay 0.01, EMA decay 0.995, and batch size 64, including the last short batch. Training lasts 200 epochs: 400 or 2,400 updates at the two sample sizes. The branches use signed loss, signed loss plus $0.001\|\theta\|^2$, positive loss on true-law draws, or positive loss on nuisance draws. Eight paired seeds share data, initialization, batches, outcome base noise, diffusion randomness, and jitter. Each EMA generator returns 2,048 outcomes per cell. The reconstruction matches released forward values and parameter gradients within 1.12×10^{-16} and 8.89×10^{-16}. Probability repairs use 121 equally spaced support points on $[-6, 6]$, 64 features, and analytic continuous targets.

All neural branches use 100 diffusion steps with β_t linear from 0.001 to 0.2, following the step-count scaling in the implementation accompanying Nichol & Dhariwal (2021). Here $\bar{\alpha}_{100} = \prod_t (1 - \beta_t) = 2.0390 \times 10^{-5}$, and the largest terminal $D_{\text{KL}}(q_T \| N(0, 1))$ across cells is 1.47×10^{-5}. Every neural sampler starts from $N(0, 1)$. Table 7 gives all eight methods' CRPS excess, calculated as integrated squared CDF error against truth. The supplement also reports complete-kernel MMD at bandwidth 1, tail error at 0.5, and quantile error averaged over levels 0.1, 0.5, and 0.9. Each metric first averages six cells within a seed.

Table 7: Calibrated generation: CRPS excess, mean (seed standard deviation) over eight paired repetitions. Neural branches share the 100-step schedule and standard-normal initialization. Oracle-positive diffusion uses true-law training draws. Every branch completed with finite output.

Method	$n_v = 32$	$n_v = 256$
Signed diffusion	0.02805 (0.01941)	0.00509 (0.00371)
Signed diffusion + ridge	0.02969 (0.02238)	0.00562 (0.00396)
Oracle-positive diffusion	0.01287 (0.00587)	0.00094 (0.00027)
Plug-in diffusion	0.06712 (0.00748)	0.05379 (0.00224)
Continuous RFF	0.03861 (0.01337)	0.00785 (0.00293)
Continuous CDF	0.03079 (0.01229)	0.00504 (0.00167)
Positive part	0.02984 (0.01216)	0.00503 (0.00143)
Plug-in	0.05342 (0.00000)	0.05342 (0.00000)

The terminal-law check is analytic. Applying the DDPM forward process (Ho et al., 2020) to Gaussian truth and taking a Gaussian conditional expectation gives the forward law and exact noise predictor:

$$Y_t \sim N(\sqrt{\bar\alpha_t}\mu, \bar\alpha_t\sigma^2 + 1 - \bar\alpha_t), \qquad \epsilon_t^*(x) = \frac{\sqrt{1-\bar\alpha_t}(x - \sqrt{\bar\alpha_t}\mu)}{\bar\alpha_t\sigma^2 + 1 - \bar\alpha_t}.$$

The reverse chain is Gaussian at every step, allowing exact propagation. With the original 100-step endpoints 10^{-4} and 0.02, $\bar\alpha_{100} = 0.363563$: at $(\mu, \sigma) = (1.2, 0.6)$, the true terminal mean and variance are 0.72355 and 0.76732. Under the code's DDPM reverse variance, changing a standard-normal start to the true terminal law reduces the mean exact-denoiser CRPS excess from 0.007117 to 0.0000306. With the calibrated schedule both values are approximately 2.647×10^{-4}, differing by less than 2×10^{-11}. Using both the true terminal law and exact Gaussian reverse variance recovers truth to numerical precision. The supplement retains both schedules' complete trained results and paired differences.

D Independent-school risk evaluation

D.1 Target and assumptions

This analysis targets baseline-complete pupils in independent schools from the same population as the training schools. Of the 6,325 assigned STAR pupils, 6,317 have the retained baseline fields; 5,785 have observed outcomes and 532 have missing outcomes. Let R indicate outcome observation. We use all baseline-complete rows, assume consistency and conditional random assignment, and require

$$Y \perp R \mid A, X, \qquad \pi(x) = \mathbb{P}(R = 1 \mid A = a, X = x) > 0.$$

Here X includes baseline and assignment-design information; a is fixed. We condition on school arm counts and assume exchangeability of assignment labels. The known e_s is the fraction assigned to a in school s before outcome selection; for pupil i in that school, $e_i = e_s$. The retention and outcome regressions use the transferable baseline variables specified below. Their accuracy requirements enter Proposition 16.

All expectations below condition on the independent training sample, including the fitted probability predictors P, P_0, support, features, and eligibility rule. For a proper loss S, define

$$d(y, v) = S(P(v), y) - S(P_0(v), y), \qquad m(x) = \mathbb{E}\{d(Y, V) \mid A = a, X = x\}.$$

Let $C_1, \ldots, C_G$ be iid evaluation schools and N_s their eligible pupil counts. Dependence within a school is unrestricted. The conditional, student-weighted risk difference is

$$\Delta = \frac{\mathbb{E}\{\sum_{i \in C_s} d(Y_i^a, V_i)\}}{\mathbb{E}N_s}.$$

Applying the sample-selection score of Bia et al. (2024) to the fixed loss difference $d(Y, V)$ gives

$$\psi_i = \widehat m(X_i) + \frac{\mathbf{1}(A_i = a)R_i}{e_i\widehat\pi(X_i)}\{d(Y_i, V_i) - \widehat m(X_i)\}, \qquad \widehat Z_s = \sum_{i \in C_s}\psi_i, \qquad \widehat\Delta = \frac{\sum_s \widehat Z_s}{\sum_s N_s}. \quad (32)$$

Proposition 16 (Independent-school evaluation)**.** *Let Z_s be the school sum in equation 32 using the true m, π. Suppose $e\pi, e\widehat{\pi}$ are bounded away from zero, N_s, Z_s have finite second moments, $\mathbb{E}N_s > 0$, and a nondegenerate cluster CLT holds for $Z_s - \Delta N_s$. If*

$$\mathbb{E}(\widehat{Z}_s - Z_s)^2 = o_p(1), \qquad \frac{\mathbb{E}\sum_{i\in C_s} |\widehat{\pi}(X_i) - \pi(X_i)|\,|\widehat{m}(X_i) - m(X_i)|}{\mathbb{E}N_s} = o_p(G^{-1/2}),$$

then

$$\widehat{\Delta} - \Delta = \frac{1}{G\mathbb{E}N_s} \sum_{s=1}^{G} (Z_s - \Delta N_s) + o_p(G^{-1/2}). \tag{33}$$

Under the corresponding conditional law of large numbers,

$$\widehat{\mathrm{se}}^2 = \frac{G}{G-1} \frac{\sum_s (\widehat{Z}_s - \widehat{\Delta} N_s)^2}{(\sum_s N_s)^2} \tag{34}$$

consistently estimates the first-order variance. When retention is known, the same conclusion holds with $Z_s = \widehat{Z}_s$ for any fixed square-integrable risk regression satisfying the stated cluster moment and CLT conditions.

Proof. Conditional expectation gives

$$\frac{\mathbb{E}(\widehat{Z}_s - Z_s)}{\mathbb{E}N_s} = \frac{\mathbb{E}\sum_{i\in C_s} \{1 - \pi(X_i)/\widehat{\pi}(X_i)\}\{\widehat{m}(X_i) - m(X_i)\}}{\mathbb{E}N_s}.$$

Positivity and the product condition bound this by $o_p(G^{-1/2})$. Independence of schools and the score L_2 condition give the same order for the centered nuisance remainder. Expanding the denominator proves equation 33; the conditional CLT and law of large numbers give equation 34. For known retention the displayed bias is zero for every fixed regression. □

Normal critical values give asymptotic coverage; we use t_{G-1} critical values for the reported intervals. The target conditions on the fitted predictors, so this expansion involves evaluation scores alone.

D.2 STAR IMPLEMENTATION AND RESULTS

A seeded permutation assigns 40 schools to training and 39 to evaluation. Their baseline-complete sample sizes are 3,123 and 3,194, including 2,891 and 2,894 observed outcomes. Two folds of 20 training schools cross-fit Gaussian ridge regressions with penalty 10 and arm-specific logistic retention regressions with $C = 1$. Predictors use gender, coarsened ethnicity, birth quarter, and school type. School identifiers are omitted. Evaluation nuisances are refitted on all training rows; their retention predictions range from 0.634 to 0.984, inside the clipping interval $[0.05, 0.995]$.

Training observed-outcome quantiles give 59 distinct support points. The training-defined eligibility rule requires 60 baseline rows and 12 observed outcomes per arm in a V-cell. Its ten cells contain 2,992 evaluation pupils across all 39 schools. Every repair uses the same nuisances and support. The largest certified numerical CDF error is 1.16×10^{-6}. For a support law, CRPS is evaluated by

$$\mathrm{CRPS}(P_w, y) = \sum_j w_j |s_j - y| - \tfrac{1}{2} \sum_{j,l} w_j w_l |s_j - s_l|.$$

Gaussian expectations of absolute distance, kernel values, and the threshold indicator at 460 are analytic. Table 8 gives the CRPS contrasts. The supplement contains kernel-score and Brier-score contrasts under the same sampling target.

D.3 COVERAGE WITH FITTED RETENTION AND RISK REGRESSIONS

For each of 1,000 paired repetitions, draw 40 new training schools and 39 independent evaluation schools, with 48 pupils per school. Each school randomly assigns 14 pupils to action 1; the ten eligible V-labels occur uniformly. For label index $v = 0, \ldots, 9$, outcomes have mean $420 + 6v + 12a$, marginal standard deviation 24, and a shared Gaussian school component giving intracluster correlation 0.35.

Table 8: CRPS risk differences on 39 independent evaluation schools, conditional on training in 40 other schools. Entries give estimates and cluster t_{38} intervals from Proposition 16; negative values favor the row method.

Action	Fit	Minus positive part	Minus plug-in
0	Continuous RFF	$0.8519\ [0.7881, 0.9156]$	$0.7230\ [0.4109, 1.0351]$
0	Continuous CDF	$-0.0003\ [-0.0040, 0.0034]$	$-0.1291\ [-0.4516, 0.1933]$
1	Continuous RFF	$1.1458\ [1.0403, 1.2512]$	$1.4472\ [0.9998, 1.8945]$
1	Continuous CDF	$0.0021\ [-0.0078, 0.0121]$	$0.3035\ [-0.1302, 0.7372]$

Retention probabilities are 0.90 and 0.85 in actions 0 and 1. The evaluated STAR probability predictors remain fixed.

In every repetition, the training sample refits logistic retention regressions with $C = 1$ and Gaussian ridge regressions with penalty 10, using an intercept and ten V-indicators. Outcome scaling and residual variance use training observed outcomes only. These fits use the same routines as the STAR evaluation. Table 9 crosses known versus fitted retention with known versus fitted risk regression on the same evaluation samples. With both fitted, the twelve coverage rates are 93.3–95.7%, and the ratio of mean reported standard error to empirical standard deviation is 0.951–1.033. The saved repetition-level results include both nuisance errors and the exact conditional bias

$$B = \frac{1}{10} \sum_{v=0}^{9} \left(1 - \frac{\pi}{\widehat{\pi}(v)}\right) \{\widehat{m}(v) - m(v)\}.$$

Table 9: Coverage (%) of nominal 95% intervals in 1,000 paired repetitions, each with 40 new training schools and 39 independent evaluation schools. Columns identify the fitted evaluation nuisances; the remaining nuisances equal their known values. All contrasts are against positive part. Monte Carlo standard errors are 0.6–0.8 percentage points.

Action	Fit	Loss	Neither	Risk only	Retention only	Both
0	Continuous CDF	kernel	93.9	93.9	94.1	94.0
0	Continuous CDF	crps	94.6	94.7	94.7	94.6
0	Continuous CDF	brier	95.6	95.5	95.4	95.4
0	Continuous RFF	kernel	94.1	94.0	94.3	93.7
0	Continuous RFF	crps	93.2	93.2	93.2	93.3
0	Continuous RFF	brier	94.9	95.1	95.0	95.1
1	Continuous CDF	kernel	94.7	94.3	94.5	94.6
1	Continuous CDF	crps	94.8	94.8	94.7	94.9
1	Continuous CDF	brier	94.9	95.3	95.0	95.2
1	Continuous RFF	kernel	93.8	94.2	94.5	94.3
1	Continuous RFF	crps	96.1	95.8	95.9	95.7
1	Continuous RFF	brier	96.0	95.8	96.0	95.2

A known-retention control additionally shifts the Gaussian risk-regression mean by 0.3×24. Its 24 prescribed interval coverages over 1,000 repetitions range from 93.3% to 95.9%. Known retention makes these scores unbiased, while the fitted-retention branches exercise the product remainder in Proposition 16.

E Related work

GDR-learners implement orthogonal generative risks (Melnychuk & Feuerriegel, 2026). Their Remark 1 and Eq. (86) link positivity of a pseudo-density ratio to Hölder smoothness and a vanishing L_4 nuisance product. Our smooth counterexample in Appendix A.5 shows how this implication fails on an unbounded covariate domain. The recovery conditions control errors relative to cell mass, and the diffusion construction identifies the role of denoising-feature moments. DoubleGen establishes robustness and oracle-rate guarantees for loss-based generators (Luedtke & Fukumizu, 2025).

Counterfactual mean embeddings (Muandet et al., 2021) and Hilbert-valued one-step estimation (Luedtke & Chung, 2024) supply the representation and semiparametric foundations of the estimation stage. Conditional counterfactual mean embeddings (CCME) (Anancharoenkij & Ponnoprat, 2026) use the same augmented feature and regression on V, and establish learning rates for that estimation stage. OEP takes this estimate as input and specifies its nearest probability embedding in a chosen Hilbert distance. Related distributional targets include policy embeddings (Zenati et al., 2025), smoothed discrepancies between conditional outcome distributions (Jain & Luedtke, 2026), conditional quantile and risk-measure effects (Kallus & Oprescu, 2023), conditional kernel tests (Park et al., 2021), marginal doubly robust kernel tests (Martinez-Taboada et al., 2023), and counterfactual density learning through Stein discrepancies (Martinez-Taboada & Kennedy, 2024).

Isotonic correction has a direct precedent in Westling et al. (2020). They establish finite-sample improvement in supremum error on a grid and, under regularity conditions, asymptotic equivalence to the initial estimate, with an application to G-computed distribution functions. Theorems 5–6 apply Hilbert projection geometry to the chosen distributional distance and carry the augmented estimate's product-error bound through the repair. The signed proper-score identities then specify which risk is minimized; the finite-support results account for representation and numerical error.

Probability-constrained causal estimation also includes targeted substitution (van der Laan & Rubin, 2006) and running-maximum and clipping correction of doubly robust CDFs (Huang et al., 2022). Ham et al. (2024) repair a one-step CDF before projection onto log-concave densities. These approaches provide probability-valued estimates through constraints adapted to their target distributions.