Cost-Aware Interpolation of Soft Interventions: Blend of Propensity, Target Law, and Product of Experts

Johan de Aguas
Department of Mathematics, University of Oslo
Integreat – Norwegian Centre for Knowledge-driven Machine Learning
johanmd@uio.no

Abstract

We introduce a family of stochastic interventions for discrete treatments that generalizes incremental propensity score interventions and bridges causal modeling with cost-sensitive domains. The formulation consists of a cost-penalized information projection problem of the independent product of the organic propensity and a user-specified target, yielding closed-form couplings. The induced marginals represent modified stochastic interventions and move smoothly, via a single tilt parameter, from the status quo or from the target distribution toward a product-of-experts limit when all destination costs are strictly positive. For inference, we derive efficient influence functions of their expected outcomes under a nonparametric model and construct one-step estimators with uniform confidence bands that exhibit stable performance and improved robustness relative to plug-in baselines. This framework can operationalizes graded scientific hypotheses under realistic constraints. Because the tilt is continuous, the costs and targets are modular, and expert-informed targets can integrate naturally with data-driven propensities, analysts can sweep feasible policy spaces, prototype candidates, and prioritize scarce experimental resources before committing them. This can help close the loop between observational evidence and resource-aware experimental design.

1 Introduction

From agricultural field trials to modern biomedicine, science advances by posing causal and counterfactual questions such as *what would happen under a different action, policy, treatment dose, or regimen?*, and then intervening in the system to gather evidence and test such queries [1, 2, 3]. Randomized controlled experiments are considered the scientific gold standard for isolating causal effects, but real-world constraints; including ethical, logistical, and economic considerations; often make deterministic population-wide interventions infeasible [4, 5, 6]. For some domains, two implications can follow: *(i)* experimental interventions may be impractical altogether, requiring inference from observational data, and *(ii) hard* interventions, in particular, may be unrealistic or lead to unstable inference, making *soft/stochastic* interventions a compelling analytical alternative. In such settings, studying stochastic interventions with observational data can become an engine for *hypothesis generation* and *policy prototyping*, provided the relevant causal queries are identifiable and can be estimated efficiently.

Stochastic interventions offer a practical alternative by allowing surgical modifications in treatment assignment rules using probabilistic or functional shifts [7]. Among them, *incremental propensity score interventions* (IPI) [8] represent a particularly useful subclass: by tilting the propensity score via a single tuning parameter, they generate a continuum of actions that interpolate between a non-intervention and a hard policy, without the stringent positivity requirements typically needed for the

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: CauScien: Uncovering Causality in Science.

latter [9]. Standard IPIs, however, ignore heterogeneous treatment costs and, by design, interpolate only toward a hard target. In many policy settings, planners instead have a structured soft target in mind. For example, logistical constraints may call for allocating treatment 1 to 80% of units, treatment 2 to 10%, and no treatment to the remaining 10%. Because treatment options can carry different costs, planners also seek to keep the overall deployment costs low.

Contributions. We introduce a formulation that yields two families of stochastic interventions governed by a single tilt parameter. The first family smoothly interpolates from a non-intervention to the *product of experts* (PoE) blend of the observed propensity score and a pre-specified target distribution; the second interpolates from the target distribution to the PoE. Such PoE limit arises when all treatment options carry strictly positive costs; if some actions are costless, the limiting behavior depends on the zero-cost set. These cost-aware interpolations are related to solutions to a *cost-penalized I-projection* problem. The first family directly generalizes IPIs and, as with them, does not require the positivity conditions demanded by hard interventions. Under standard identification assumptions for observational data, we derive efficient influence functions, under a nonparametric model, for the expected outcome after these interventions, develop doubly robust one-step estimators, and construct asymptotically valid confidence bands. We demonstrate consistent performance gains over naïve plug-in estimators in controlled simulations.

2 Preliminaries

Let $A \in \mathcal{A}$ be a discrete point-exposure variable, $Y \in \mathbb{R}$ a continuous outcome variable, $W \in \mathcal{W}$ a vector of covariates, and $\pi(a \,|\, w) := \mathbb{P}(A = a \,|\, W = w)$ the propensity score of treatment option $a \in \mathcal{A}$. Potential outcomes Y^a encode unit-level counterfactuals after a hard intervention $\mathrm{do}(A = a)$ [2]. Under counterfactual consistency, positivity $\pi(a \,|\, w) \in (0, 1), \forall w \in \mathcal{W}$, and conditional ignorability / backdoor admissibility of W, one can identify the expected outcome after intervention $\mathrm{do}(A = a)$ from observational data via the g-computation / backdoor formula, as $\mathbb{E}[Y^a] = \mathbb{E}[Y \mid \mathrm{do}(A = a)] = \mathbb{E}\{Q(W, a)\}$, with $Q(w, a) = \mathbb{E}[Y \,|\, W = w, A = a]$.

For a binary exposure, i.e. $\mathcal{A} = \{0, 1\}$, an IPI tilts the propensity score $\pi(1 \,|\, w)$ to be:

$$\widetilde{\pi}_\delta(1 \,|\, w) := \frac{e^\delta \pi(1 \,|\, w)}{e^\delta \pi(1 \,|\, w) + \pi(0 \,|\, w)}, \qquad \text{with } \delta \in \mathbb{R}. \tag{1}$$

The value $\delta = 0$ corresponds to a non-intervention, while $\delta \to \infty$ (resp. $\delta \to -\infty$) pushes toward the hard intervention $\mathrm{do}(A = 1)$ (resp. $\mathrm{do}(A = 0)$) [8]. Under conditional ignorability / backdoor admissibility, the expected outcome under such stochastic intervention remains identifiable without the positivity condition [8, 9], and it is given by:

$$\mathbb{E}[Y^{\widetilde{\pi}_\delta}] = \mathbb{E}\big\langle \widetilde{\pi}_\delta(\cdot \,|\, W),\, Q(W, \cdot) \big\rangle, \tag{2}$$

where the inner product $\langle \cdot, \cdot \rangle$ is taken over $\mathcal{A}$.

3 Cost-penalized I-projection and tilted marginals

Given two input probability measures π (*source*) and ν (*target*) over the set $\mathcal{A}$, a cost function on pairs $c : \mathcal{A}^2 \to [0, \infty)$, and a tuning parameter $\delta > 0$, we define the *cost-penalized I-projection* (CPIP) of the independent product $\pi \otimes \nu$ as the joint distribution $\gamma \in \mathcal{M}_+^1(\mathcal{A}^2)$ (in the class of distributions over the Cartesian product $\mathcal{A}^2$) that solves:

$$\inf_{\gamma \in \mathcal{M}_+^1(\mathcal{A}^2)} \mathbb{D}_{\mathrm{KL}}(\gamma \,|\, \pi \otimes \nu) + \delta\, \mathbb{E}_\gamma \left[c(A_1, A_2)\right]. \tag{3}$$

This formulation is closely related to unconstrained, relaxed, and limiting-case variants of entropic optimal transport and Schrödinger bridge problems [10, 11, 12, 13]. Yet, while entropic optimal transport problems typically require iterative solvers such as Sinkhorn's algorithm, the CPIP problem admits a closed-form solution thanks to the strong convexity and smoothness of its objective. Its unique minimizer is given by the Boltzmann–Gibbs kernel:

$$\gamma_\delta^*(a', a'') \propto \pi(a')\, \nu(a'')\, e^{-\delta c(a', a'')}. \tag{4}$$

3.1 Tilted marginal distributions with treatment-specific costs

Remark 1. *Let $A \in \mathcal{A} = \{\alpha_1, \ldots, \alpha_K\}$ be a categorical exposure with $K > 1$ treatment options, and ν be any valid probability distribution over $\mathcal{A}$. Let the reallocation cost from $A = \alpha_j$ to $A = \alpha_k \neq \alpha_j$ be a value that is specific for the received treatment α_k and constant over profiles W, i.e., $c(\alpha_j, \alpha_k) = c(\alpha_k)\,\mathbb{I}(\alpha_j \neq \alpha_k)$, with $0 \leq c(a) < \infty$ for all $a \in \mathcal{A}$. Then, for each $w \in \mathcal{W}$, the marginals of the CPIP solution with source $\pi(\cdot \mid w)$ and target ν are:*

$$\pi^*_\delta(a \mid w) := \frac{(\zeta_\delta + \xi_\delta(a))\,\pi(a \mid w)}{\sum_{a' \in \mathcal{A}}(\zeta_\delta + \xi_\delta(a'))\,\pi(a' \mid w)} \quad \text{and} \quad \nu^*_\delta(a \mid w) := \frac{\nu(a) - \xi_\delta(a)(1 - \pi(a \mid w))}{\sum_{a' \in \mathcal{A}}(\zeta_\delta + \xi_\delta(a'))\,\pi(a' \mid w)}, \tag{5}$$

where $\xi_\delta(a) := \nu(a)\left(1 - e^{-\delta c(a)}\right)$ and $\zeta_\delta := \sum_{a' \in \mathcal{A}} \nu(a')\, e^{-\delta c(a')}$.

We call π^*_δ and ν^*_δ the *tilted source / target marginal distributions*.

Remark 2. *When $K = 2$, the target ν is set to be the hard policy of giving treatment to everyone, i.e. $\nu(a) = \mathbb{I}\{a = 1\}$, and the cost structure is the Hamming cost, $c(a', a'') = \mathbb{I}(a' \neq a'')$, the tilted source π^*_δ reduces to an IPI with tilt parameter δ.*

We provide proofs of these remarks in the technical appendix A. Figure 1 presents the tilted marginal distributions for a binary exposure under varying configurations of the tilt parameter δ, cost functions c, and target distribution ν.

Note that at $\delta = 0$ and for all $a \in \mathcal{A}, w \in \mathcal{W}$, one gets $\pi^*_0(a \mid w) = \pi(a \mid w)$, a non-intervention, and $\nu^*_0(a \mid w) = \nu(a)$, a target intervention. In other words, setting $\delta = 0$ results in no modification of the input distributions. Furthermore, denote $\mathcal{A}_0 = \{a \in \mathcal{A} : c(a) = 0\}$, $\mathcal{A}_+ = \{a \in \mathcal{A} : c(a) > 0\}$, $\nu^\dagger(a) := \nu(a)\,\mathbb{I}(a \in \mathcal{A}_+) + \sum_{a \in \mathcal{A}_0} \nu(a)$ and $\pi^\dagger(a \mid w) := \pi(a \mid w) + (1 - \pi(a \mid w))\mathbb{I}(a \in \mathcal{A}_0)$. Then, in the limit $\delta \to \infty$, one obtains:

$$\pi^*_\infty(a \mid w) := \frac{\pi(a \mid w)\,\nu^\dagger(a)}{\sum_{a' \in \mathcal{A}} \pi(a' \mid w)\,\nu^\dagger(a')} \quad \text{and} \quad \nu^*_\infty(a \mid w) := \frac{\pi^\dagger(a \mid w)\,\nu(a)}{\sum_{a' \in \mathcal{A}} \pi^\dagger(a' \mid w)\,\nu(a')}. \tag{6}$$

When all treatment costs are positive, both reduce to the *product of experts* (PoE) distribution $\mathrm{PoE}(a) \propto \pi(a \mid w)\,\nu(a)$ [14].

The CPIP objective is strictly convex and well-posed for $\delta > 0$, yielding a unique solution. Interestingly, for $\delta < 0$, the joint distribution γ^*_δ in (4) is not the optimizer of the CPIP program, but it yields smooth and closed-form *parametric extensions* of π^*_δ and ν^*_δ, provided the cost is bounded and the action set $\mathcal{A}$ is finite. In this regime, the induced tilted marginals act as extrapolations that shift mass toward higher-cost and repulsive pairings. For continuous settings with lower unbounded costs, additional integrability conditions are needed for $\delta < 0$ to remain well-posed.

Since covariate set W is assumed to fulfill the backdoor criterion, the expected outcomes under the stochastic interventions π^*_δ and ν^*_δ are identified from observational data [7, 15], and given by:

$$\mathbb{E}[Y^{\pi^*_\delta}] \equiv \mu^S_\delta = \mathbb{E}\,\langle \pi^*_\delta(\cdot \mid W), Q(W, \cdot)\rangle \;\text{ and }\; \mathbb{E}[Y^{\nu^*_\delta}] \equiv \mu^T_\delta = \mathbb{E}\,\langle \nu^*_\delta(\cdot \mid W), Q(W, \cdot)\rangle\,. \tag{7}$$

Notably, while the identification of μ^T_δ requires the usual positivity / overlap condition, i.e. $\operatorname{ess\,sup}_{a \in \mathcal{A}} \nu^*_\delta(a \mid W)/\pi(a \mid W) < \infty$ (P_W-almost surely), the expected outcome under the tilted source, μ^S_δ, as a generalization of IPIs preserving the support of the organic law, does not impose any additional positivity beyond the stability condition $\mathbb{E} \sum_{a \in \mathcal{A}} [\pi^*_\delta(a \mid W)/\pi(a \mid W)]^2 < \infty$, also needed for the former.

4 Estimators and inference

Plug-in estimators for μ^S_δ and μ^T_δ can be readily obtained by estimating π and Q, modifying the former via transformations in equations (5), substituting them into the inner product expressions in equations (7), and averaging over IID samples. Building upon foundational work in doubly robust estimation of classical causal effects [16, 17, 18, 19, 20, 21], we also present estimators grounded in semiparametric theory that incorporate data-adaptive statistical methods while achieving optimal asymptotic performance under flexible and realistic assumptions about the data-generating process (DGP) [22]. These *Newton–Raphson one-step corrected estimators* ensure consistency even under partial misspecification of some nuisance components [23], and leverage the *efficient influence function* (EIF) of the smooth functional of the distribution that defines the parameter of interest [24].

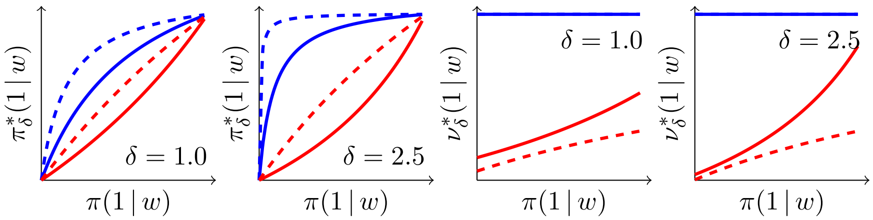

Figure 1: Tilted source distribution π^*_δ (left) and tilted target distribution ν^*_δ (right) for a binary exposure at $W = w$, shown as pointwise transformation of the propensity score $\pi(1 \mid w)$, for cases $\delta = 1.0$ and $\delta = 2.5$. Line color indicates the target configuration: $\nu = (0, 1)$ (blue) and $\nu = (0.7, 0.3)$ (red). Line style denotes the cost structure: $c = (1, 1)$ (solid), $c = (0, 2)$ (dashed). The first component is for $A = 0$ and second for $A = 1$.

Remark 3. *Let $\mathcal{S}_\delta[P]$ denote the functional that evaluates expression $\mathbb{E}_P[Y^{\pi^*_\delta}]$ at an arbitrary distribution P. Similarly, let $\mathcal{T}_\delta[P]$ represent the corresponding P-functional from expression $\mathbb{E}_P[Y^{\nu^*_\delta}]$. Let $O_i = (W_i, A_i, Y_i)$ denote a sample drawn from the true DGP P^*. Under a nonparametric model, suitable smoothness and regularity conditions, the* uncentered *EIF of $\mathcal{S}_\delta[P]$ at P^*, evaluated at point O_i, exists and given by:*

$$D^S_\delta(O_i) = \frac{\pi^*_\delta(A_i \mid W_i)}{\pi(A_i \mid W_i)} \left[Y_i - \sum_{a \in \mathcal{A}} \pi^*_\delta(a \mid W_i)\, Q(W_i, a) \right] + \sum_{a \in \mathcal{A}} \pi^*_\delta(a \mid W_i)\, Q(W_i, a). \tag{8}$$

Analogously, the uncentered *EIF of $\mathcal{T}_\delta[P]$ at P^*, evaluated at point O_i, exists and is given by:*

$$\begin{aligned} D^T_\delta(O_i) = {} & \frac{\nu^*_\delta(A_i \mid W_i)}{\pi(A_i \mid W_i)} [Y_i - Q(W_i, A_i)] + \left[2 - \frac{\pi^*_\delta(A_i \mid W_i)}{\pi(A_i \mid W_i)} \right] \sum_{a \in \mathcal{A}} \nu^*_\delta(a \mid W_i)\, Q(W_i, a) \\ & + \frac{\pi^*_\delta(A_i \mid W_i)}{\pi(A_i \mid W_i)} \varrho_\delta(A_i)\, Q(W_i, A_i) - \sum_{a \in \mathcal{A}} \pi^*_\delta(a \mid W_i)\, \varrho_\delta(a)\, Q(W_i, a). \end{aligned} \tag{9}$$

where $\varrho_\delta(a) := \frac{\xi_\delta(a)}{\zeta_\delta + \xi_\delta(a)}$. Derivation is provided in technical appendix A.

Observe that $\varrho_0(a) = 0$ and $\pi^*_0(a \mid w)/\pi(a \mid w) = 1$ for all $a \in \mathcal{A}, w \in \mathcal{W}$. Hence, when $\delta = 0$ and $\nu(a) = \mathbb{I}(a = a')$, we recover the standard uncentered EIF for the expected outcome of a hard intervention $\mathrm{do}(A = a')$, identified with $\mathbb{E}[Q(W, a')]$ [17, 18].

Consequently, one-step estimators of the expected outcomes in expressions (7), along with asymptotical confidence intervals can be constructed as follows:

1. Split data into K folds. On each training fold k, fit $\widehat{\pi}_k$ and $\widehat{Q}_k$ with data-adaptive learners (ML).
2. For a grid G of δ-values, compute $\widehat{\pi}^*_{\delta,k}, \widehat{\nu}^*_{\delta,k}$ via transformations (5).
3. Evaluate and average $\widehat{D}^S_{\delta,k}, \widehat{D}^T_{\delta,k}$ on held-out folds and units to form one-step estimates $\widehat{\mu}^S_\delta, \widehat{\mu}^T_\delta$.
4. **Uniform confidence bands**: For each $b \in \{1, \dots, B\}$, draw $\{\chi_i^{(b)}\}_{i=1}^n \overset{iid}{\sim} N(0, 1)$, compute $\zeta^{(b)} = \sup_{\delta \in G} \left| n^{-1/2} \sum_i \chi_i^{(b)} (\widehat{D}^S_\delta(O_i) - \widehat{\mu}^S_\delta) / \widehat{\mathrm{var}}\left(D^S_\delta(O)\right)^{1/2} \right|$, and save the 95% quantile of $\{\zeta^{(b)}\}_{b=1}^B$ for band construction around $\widehat{\mu}^S_\delta$. Do the same for $\widehat{\mu}^T_\delta$.
5. Compute final estimates and intervals:

$$\widehat{\mu}^S_\delta \mp z^S \, \frac{\widehat{\mathrm{var}}\left(D^S_\delta(O)\right)^{1/2}}{\sqrt{n}} \quad \text{and} \quad \widehat{\mu}^T_\delta \mp z^T \, \frac{\widehat{\mathrm{var}}\left(D^T_\delta(O)\right)^{1/2}}{\sqrt{n}}, \tag{10}$$

where z^S, z^T equal 1.96 for a 95% *pointwise* Wald-type confidence interval at a fixed value of δ, or correspond to the saved 95% quantile of the respective samples $\{\zeta^{(b)}\}_{b=1}^B$ for *uniform* confidence bands [25].

5 Empirical evaluation

We conducted an evaluation task for the proposed estimators using repeated simulations with synthetic data. The employed DGP is adapted from Kang and Schafer [26] and Kennedy [8], with bespoke

Table 1: Integrated bias (iBias) and root mean squared error (iRMSE) for the estimated expected outcome under stochastic interventions π^*_δ and ν^*_δ. Results are averaged over 200 simulations and shown for two estimators (plug-in and one-step), across three model specifications: *(i)* correctly specified, *(ii)* misspecified outcome regression Q, and *(iii)* misspecified propensity score π; and under three combinations of cost c and target ν.

		Setup 1: $c = (2.0, 1.0, 1.0)$ $\nu = (0.4, 0.4, 0.2)$				**Setup 2:** $c = (1.0, 0.5, 2.0)$ $\nu = (0.5, 0.3, 0.2)$				**Setup 3:** $c = (1.0, 1.0, 2.0)$ $\nu = (0.0, 0.2, 0.8)$			
		$\widehat{\mu}^S_\delta$		$\widehat{\mu}^T_\delta$		$\widehat{\mu}^S_\delta$		$\widehat{\mu}^T_\delta$		$\widehat{\mu}^S_\delta$		$\widehat{\mu}^T_\delta$	
Estim.	**Misspec.**	**iBias**	**iRMSE**	**iBias**	**iRMSE**	**iBias**	**iRMSE**	**iBias**	**iRMSE**	**iBias**	**iRMSE**	**iBias**	**iRMSE**
plug-in	–	0.40	2.36	3.03	5.69	0.21	2.20	3.43	5.66	1.20	3.04	9.46	11.82
plug-in	Q	1.99	3.21	15.22	15.43	0.54	2.25	17.61	17.80	3.82	4.77	27.41	27.76
plug-in	π	14.00	14.57	7.92	9.48	14.67	15.14	3.08	5.59	13.32	14.12	10.28	12.68
one-step	–	**0.02**	2.20	**0.57**	8.44	**0.02**	2.14	**0.38**	7.38	**0.04**	**2.40**	**0.75**	**8.86**
one-step	Q	**0.03**	**2.20**	**0.35**	**13.89**	**0.02**	2.14	**0.50**	**11.24**	**0.06**	**2.39**	**0.49**	**16.07**
one-step	π	**0.39**	**2.32**	**3.98**	**5.86**	**0.36**	**2.23**	**2.66**	**4.82**	**1.15**	**2.83**	**5.17**	**7.50**

modifications to introduce a three-leveled categorical exposure and to increase the noise in the system, thereby aligning the signal-to-noise ratio with amounts commonly observed in social science and observational clinical data. The DGP is given in appendix B.

Table 1 presents the results across three setups, with varying costs and targets, with $n = 1000$ and after 200 repetitions. We compare plug-in vs. one-step estimators for the expected outcome under both stochastic policies π^*_δ and ν^*_δ, based on the integrated bias (iBias) and integrated RMSE (iRMSE) averaged over $\delta \in [-2, 2]$. The one-step method dominates overall: under correct specification it delivers near-zero bias and lower or comparable RMSE, especially for the tilted source intervention. For the tilted target intervention, a more diffuse target ν trades small bias for a modest RMSE increase, but the one-step estimator remains competitive. Under misspecification, gains are decisive: with a misspecified outcome model $\widehat{Q}$, one-step maintains low iBias / iRMSE across all setups while plug-in degrades sharply (most notably for the tilted target case); with a misspecified propensity $\widehat{\pi}$, one-step still substantially reduces both metrics relative to plug-in, though performance drops more than in the Q-misspecified case. Overall, the one-step correction consistently attenuates misspecification sensitivity, yielding the most reliable performance.

6 Conclusion

This work introduces a cost-aware family of stochastic interventions for discrete treatments that generalizes incremental propensity score interventions (IPI) and provides an explicit bridge between causal modeling and domains with cost-sensitive interventions. By formulating a cost-penalized I-projection (CPIP) of the independent product of organic and target distributions, we obtain closed-form Boltzmann–Gibbs couplings whose induced marginals interpolate, via a single tilting parameter δ, from the propensity scores or from the target intervention, respectively, toward a product-of-experts (PoE) limit when all destination costs are strictly positive. On the inferential side, we derive efficient influence functions (EIF), under a nonparametric model, for the expected outcomes under these policies and construct one-step estimators that deliver stable performance and improved misspecification robustness compared to plug-in baselines, with uniform bands over δ.

These policies can operationalize graded scientific hypotheses under realistic constraints. Because δ is continuous and the costs c and targets ν are modular, analysts can sweep feasible spaces to prototype and evaluate policies for prospective study (e.g., stepped-wedge trials), turning rich observational registries into a pre-experimental policy prototyping engine. The explicit appearance of costs of actions makes prioritization more transparent: planners can align candidate interventions with budgets and logistical burdens, and quantify trade-offs before committing experimental resources. In clinical and other applied settings, clinician-informed targets (e.g. favoring a low dose of medication) can integrate naturally in this framework, ensuring that proposed policies reflect both expert priors and empirical regularities. Taken together, these features could help close the loop between causal identification and estimation from observational data and resource-aware experimental design when hard interventions are impractical, as is common in social, economic and clinical sciences.

References

[1] R. Fisher. *The Design of Experiments*. Oliver and Boyd, Edinburgh, 1935.

[2] Judea Pearl. *Causality*. Cambridge University Press, Cambridge, UK, 2 edition, 2009. ISBN 978-0-521-89560-6. doi: 10.1017/cbo9780511803161.

[3] Guido Imbens and Donald Rubin. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press, USA, 2015. ISBN 0521885884.

[4] Iván Díaz and Mark van der Laan. Assessing the causal effect of policies: an example using stochastic interventions. *The International Journal of Biostatistics*, 9(2):149–165, 2013. doi: 10.1515/ijb-2013-0014.

[5] Sebastien Haneuse and Andrea Rotnitzky. Estimation of the effect of interventions that modify the received treatment. *Statistics in Medicine*, 32(30):5260–5277, 2013. doi: 10.1002/sim.5905.

[6] Aaron Sarvet, Kerollos Wanis, Jessica Young, Roberto Hernandez-Alejandro, and Mats Stensrud. Longitudinal incremental propensity score interventions for limited resource settings. *Biometrics*, 79(4): 3418–3430, Dec 2023. doi: 10.1111/biom.13859.

[7] Juan Correa and Elias Bareinboim. A calculus for stochastic interventions: Causal effect identification and surrogate experiments. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(06):10093–10100, 4 2020. doi: 10.1609/aaai.v34i06.6567. URL https://ojs.aaai.org/index.php/AAAI/article/view/6567.

[8] Edward Kennedy. Nonparametric causal effects based on incremental propensity score interventions. *Journal of the American Statistical Association*, 113(522):645–656, 2018. doi: 10.1080/01621459.2017.1422737. Published online: 07 Aug 2018.

[9] Matteo Bonvini, Alec McClean, Zach Branson, and Edward Kennedy. Incremental causal effects: An introduction and review. In Michael Elliott, Elizabeth Stuart, and Daniel Westreich, editors, *Handbook of Matching and Weighting Adjustments for Causal Inference*, pages 1–24. Chapman and Hall/CRC, 1st edition, 2023. ISBN 9781003102670. doi: 10.1201/9781003102670.

[10] Christian Léonard. A survey of the schrödinger problem and some of its connections with optimal transport. *Discrete and Continuous Dynamical Systems*, 34(4):1533–1574, 2014. doi: 10.3934/dcds.2014.34.1533. Mathematics Subject Classification: 46N10, 60J25, 60F10.

[11] Charlie Frogner, Chiyuan Zhang, Hossein Mobahi, Mauricio Araya, and Tomaso Poggio. Learning with a wasserstein loss. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc., 2015. URL https://proceedings.neurips.cc/paper_files/paper/2015/file/a9eb812238f753132652ae09963a05e9-Paper.pdf.

[12] Lénaïc Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. Scaling algorithms for unbalanced optimal transport problems. *Mathematics of Computation*, 87(314):2563–2609, 2018. doi: 10.1090/mcom/3303. URL https://doi.org/10.1090/mcom/3303. Published electronically on February 6, 2018.

[13] Gabriel Peyré and Marco Cuturi. Computational optimal transport: With applications to data science. *Foundations and Trends® in Machine Learning*, 11(5-6):355–607, 2019. ISSN 1935-8237. doi: 10.1561/2200000073. URL http://dx.doi.org/10.1561/2200000073.

[14] Geoffrey Hinton. Products of experts. In *9th International Conference on Artificial Neural Networks (ICANN'99)*, pages 1–6. IEE, 1999. ISBN 978-0-85296-721-8. doi: 10.1049/cp:19991075.

[15] Juan Correa and Elias Bareinboim. General transportability of soft interventions: completeness results. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS '20, Red Hook, NY, USA, 2020. Curran Associates Inc. ISBN 9781713829546.

[16] P. Bickel, C. Klaassen, Y. Ritov, and J. Wellner. *Efficient and adaptive estimation for semiparametric models*, volume 4. Springer New York, 1993. ISBN 9780387984735.

[17] James Robins, Andrea Rotnitzky, and Lue Zhao. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association*, 90 (429):106–121, 1995. doi: 10.1080/01621459.1995.10476493.

[18] James Robins and Andrea Rotnitzky. Comment on the bickel and kwon article, 'inference for semiparametric models: Some questions and an answer'. *Statistica Sinica*, 7(1):167–199, 1997.

[19] A. van der Vaart. *Asymptotic Statistics*. Asymptotic Statistics. Cambridge University Press, 2000. ISBN 9780521784504. URL https://books.google.com/books?id=UEuQEM5RjWgC.

[20] Iván Díaz and Mark van der Laan. Population intervention causal effects based on stochastic interventions. *Biometrics*, 68(2):541–549, Jun 2012. doi: 10.1111/j.1541-0420.2011.01685.x.

[21] Tri-Dung Duong, Qian Li, and Guandong Xu. Stochastic intervention for causal effect estimation. In *2021 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2021.

[22] Iván Díaz. Machine learning in the estimation of causal effects: targeted minimum loss-based estimation and double/debiased machine learning. *Biostatistics*, 21(2):353–358, 11 2019. ISSN 1465-4644. doi: 10.1093/biostatistics/kxz042. URL `https://doi.org/10.1093/biostatistics/kxz042`.

[23] Rhian Daniel. *Double Robustness*, pages 1–14. John Wiley and Sons, Ltd, 2018. ISBN 9781118445112. doi: https://doi.org/10.1002/9781118445112.stat08068. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118445112.stat08068`.

[24] Oliver Hines, Oliver Dukes, Karla Díaz-Ordaz, and Stijn Vansteelandt. Demystifying statistical learning based on efficient influence functions. *The American Statistician*, 76(3):292–304, 2022.

[25] Victor Chernozhukov, Denis Chetverikov, and Kohei Kato. Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *The Annals of Statistics*, 41(6): 2786–2819, 2013. doi: 10.1214/13-AOS1170.

[26] Joseph Kang and Joseph Schafer. Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data. *Statistical Science*, 22(4):523–539, 2007. doi: 10.1214/07-STS227.

A Technical appendix

A.1 Remark 1: tilted marginal distributions

Consider cost-penalized I-projection (CPIP) of the independent product of input distributions π and ν over discrete sample space $\mathcal{A}$, formulated as:

$$\inf_{\gamma \in \mathcal{M}_+^1(\mathcal{A}^2)} \sum_{a', a'' \in \mathcal{A}} c(a', a'')\, \gamma(a', a'') + \frac{1}{\delta} \mathbb{D}_{\mathrm{KL}}(\gamma \,|\, \pi \otimes \nu),$$

where $\delta > 0$, $\mathcal{M}_+^1(\mathcal{A}^2)$ denotes the set of probability measures over sample space $\mathcal{A}^2$, and $c : \mathcal{A}^2 \to \mathbb{R}$ is a nonnegative and integrable cost function. Owing to the strict convexity of the objective, the unique minimizer satisfies the KKT conditions obtained from the Lagrangian:

$$\mathcal{L}(\gamma, \lambda) = \sum_{a', a'' \in \mathcal{A}} \left[c(a', a'') + \frac{1}{\delta} \log\left(\frac{\gamma(a', a'')}{\pi(a')\, \nu(a'')} \right) \right] \gamma(a', a'') - \lambda \left[\sum_{a', a'' \in \mathcal{A}} \gamma(a', a'') - 1 \right],$$

which admits a unique closed-form solution. The first-order condition is given by:

$$\begin{aligned}
&\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}\gamma(a', a'')} = c(a', a'') + \frac{1}{\delta} \left[\log\left(\frac{\gamma(a', a'')}{\pi(a')\, \nu(a'')} \right) + 1 \right] - \lambda = 0, \\
&\Rightarrow \log\left(\frac{\gamma(a', a'')}{\pi(a')\, \nu(a'')} \right) = \delta[\lambda - c(a', a'')] - 1, \\
&\Rightarrow \gamma(a', a'') = \pi(a')\, \nu(a'')\, e^{-\delta c(a', a'')}\, e^{\delta \lambda - 1} \propto \pi(a')\, \nu(a'')\, e^{-\delta c(a', a'')}.
\end{aligned}$$

Now suppose $\pi \equiv \pi(\cdot \,|\, w)$ is a conditional distribution, and the cost function c does not depend on the covariate profile $W = w$. Then, for each profile w, the solution plan becomes:

$$\gamma^*(a', a'' \,|\, w) = \frac{\pi(a' \,|\, w)\, \nu(a'')\, e^{-\delta c(a', a'')}}{\sum_{a', a'' \in \mathcal{A}} \pi(a' \,|\, w)\, \nu(a'')\, e^{-\delta c(a', a'')}}.$$

The marginals of $\gamma^*(a', a'' \,|\, w)$ are then straightforward to compute:

$$\pi_\delta^*(a \,|\, w) = \frac{\pi(a \,|\, w) \sum_{a'' \in \mathcal{A}} \nu(a'')\, e^{-\delta c(a, a'')}}{\sum_{a', a'' \in \mathcal{A}} \pi(a' \,|\, w)\, \nu(a'')\, e^{-\delta c(a', a'')}} \quad \text{and} \quad \nu_\delta^*(a \,|\, w) = \frac{\nu(a) \sum_{a' \in \mathcal{A}} \pi(a' \,|\, w)\, e^{-\delta c(a', a)}}{\sum_{a', a'' \in \mathcal{A}} \pi(a' \,|\, w)\, \nu(a'')\, e^{-\delta c(a', a'')}}.$$

Let:

- $A \in \mathcal{A} = \{\alpha_1, \ldots, \alpha_K\}$ be a categorical point-exposure variable with K treatment options,
- The target marginal ν be any valid probability distribution over $\mathcal{A}$,
- The reallocation cost from $A = \alpha_j$ to $A = \alpha_k \neq \alpha_j$ be a value that is specific for the received treatment α_k and constant over profiles $W = w$, i.e., $c(\alpha_j, \alpha_k) = c(\alpha_k)\, \mathbb{I}(\alpha_j \neq \alpha_k)$, with $0 \leq c(a) < \infty$ for all $a \in \mathcal{A}$.

Then, the tilted source marginal of the CPIP solution with parameter δ corresponds to:

$$\pi^*_\delta(a \,|\, w) = \frac{\pi(a \,|\, w) \sum_{a'' \in \mathcal{A}} \nu(a'') \, e^{-\delta c(a'') \, \mathbb{I}(a \neq a'')}}{\sum_{a', a'' \in \mathcal{A}} \pi(a' \,|\, w) \, \nu(a'') \, e^{-\delta c(a'') \, \mathbb{I}(a' \neq a'')}} = \frac{\pi(a \,|\, w) \left[\nu(a) + \sum_{a'' \neq a} \nu(a'') \, e^{-\delta c(a'')}\right]}{\sum_{a' \in \mathcal{A}} \pi(a' \,|\, w) \left[\nu(a') + \sum_{a'' \neq a'} \nu(a'') \, e^{-\delta c(a'')}\right]},$$

$$= \frac{\pi(a \,|\, w) \left[\nu(a) + \sum_{a'' \in \mathcal{A}} \nu(a'') \, e^{-\delta c(a'')} - \nu(a) \, e^{-\delta c(a)}\right]}{\sum_{a' \in \mathcal{A}} \pi(a' \,|\, w) \left[\nu(a') + \sum_{a'' \in \mathcal{A}} \nu(a'') \, e^{-\delta c(a'')} - \nu(a') \, e^{-\delta c(a')}\right]} = \frac{\pi(a \,|\, w) \, (\zeta_\delta + \xi_\delta(a))}{\sum_{a' \in \mathcal{A}} \pi(a' \,|\, w) \, (\zeta_\delta + \xi_\delta(a'))},$$

where $\xi_\delta(a) := \nu(a) \left(1 - e^{-\delta c(a)}\right)$ and $\zeta_\delta := \sum_{a' \in \mathcal{A}} \nu(a') \, e^{-\delta c(a')}$.

Similarly, the tilted target marginal of the CPIP solution with parameter δ corresponds to:

$$\nu^*_\delta(a \,|\, w) = \frac{\nu(a) \sum_{a' \in \mathcal{A}} \pi(a' \,|\, w) \, e^{-\delta c(a) \, \mathbb{I}(a' \neq a)}}{\sum_{a', a'' \in \mathcal{A}} \pi(a' \,|\, w) \, \nu(a'') \, e^{-\delta c(a'') \, \mathbb{I}(a' \neq a'')}} = \frac{\nu(a) \left[\pi(a \,|\, w) + \sum_{a' \neq a} \pi(a' \,|\, w) \, e^{-\delta c(a)}\right]}{\sum_{a' \in \mathcal{A}} \pi(a' \,|\, w) \left[\nu(a') + \sum_{a'' \neq a'} \nu(a'') \, e^{-\delta c(a'')}\right]},$$

$$= \frac{\nu(a) \left[\pi(a \,|\, w) + e^{-\delta c(a)} (1 - \pi(a \,|\, w))\right]}{\sum_{a' \in \mathcal{A}} \pi(a' \,|\, w) \left[\nu(a') + \sum_{a'' \in \mathcal{A}} \nu(a'') \, e^{-\delta c(a'')} - \nu(a') \, e^{-\delta c(a')}\right]},$$

$$= \frac{\nu(a) \left[1 - (1 - e^{-\delta c(a)})(1 - \pi(a \,|\, w))\right]}{\sum_{a' \in \mathcal{A}} \pi(a' \,|\, w) \, (\zeta_\delta + \xi_\delta(a'))} = \frac{\nu(a) - \xi_\delta(a)(1 - \pi(a \,|\, w))}{\sum_{a' \in \mathcal{A}} \pi(a' \,|\, w) \, (\zeta_\delta + \xi_\delta(a'))}.$$

A.2 Remark 2: IPIs as a special case

Let $A \in \{0, 1\}$ be a binary point-exposure. Let the target marginal ν be the degenerate distribution that always assigns treatment, $\nu(a) = \mathbb{I}(a = 1)$, and let $c(a', a'') = \mathbb{I}(a' \neq a'')$ be the Hamming cost. Then, the tilted source marginal of the CPIP solution with regularization parameter δ is:

$$\pi^*_\delta(a \,|\, w) = \frac{\pi(a \,|\, w) \sum_{a'' \in \mathcal{A}} \mathbb{I}(a'' = 1) \, e^{-\delta \mathbb{I}(a \neq a'')}}{\sum_{a', a'' \in \mathcal{A}} \pi(a' \,|\, w) \, \mathbb{I}(a'' = 1) \, e^{-\delta \mathbb{I}(a' \neq a'')}} = \frac{\pi(a \,|\, w) \, e^{-\delta \mathbb{I}(a \neq 1)}}{\sum_{a' \in \mathcal{A}} \pi(a' \,|\, w) \, e^{-\delta \mathbb{I}(a' \neq 1)}},$$

$$\Rightarrow \pi^*_\delta(1 \,|\, w) = \frac{\pi(1 \,|\, w) \, e^{-\delta \cdot 0}}{\pi(1 \,|\, w) \, e^{-\delta \cdot 0} + \pi(0 \,|\, w) \, e^{-\delta \cdot 1}} = \frac{\pi(1 \,|\, w) \, e^{\delta}}{\pi(1 \,|\, w) \, e^{\delta} + \pi(0 \,|\, w)},$$

which coincides with an IPI with tilt parameter δ, and thus $\pi^*_\delta(1 \,|\, w) = \widetilde{\pi}_\delta(1 \,|\, w)$ for all $w \in \mathcal{W}$.

A.3 Remark 3

Let ν and c be given, $Q(W, A) = \mathbb{E}[Y \,|\, W, A]$, and:

$$\mathcal{S}_\delta[P] \equiv \mu^S_\delta = \mathbb{E} \left\langle \pi^*_\delta(\cdot \,|\, W), Q(W, \cdot) \right\rangle = \sum_{a \in \mathcal{A}} \mathbb{E} \left[\pi^*_\delta(a \,|\, W) \, Q(W, a)\right],$$

$$\mathcal{T}_\delta[P] \equiv \mu^s_\delta = \mathbb{E} \left\langle \nu^*_\delta(\cdot \,|\, W), Q(W, \cdot) \right\rangle = \sum_{a \in \mathcal{A}} \mathbb{E} \left[\nu^*_\delta(a \,|\, W) \, Q(W, a)\right],$$

Consider parametric submodel $P_\epsilon \in \mathfrak{P}$ indexed by a small fluctuation parameter $\epsilon \in \mathbb{R}$, and a point-mass contamination $O_i = (W_i, A_i, Y_i) \sim P^*$, such that, $P_\epsilon(O) = \epsilon \, \mathbb{I}(O = O_i) + (1 - \epsilon) \, P^*(O)$, where $P^* \in \mathfrak{P}$ is the true DGP distribution. Under some technical conditions involving *(i)* fully nonparametric or saturated model $\mathfrak{P}$, *(ii)* smoothness for the paths within the model, and *(iii)* boundedness of the outcome mean, the Gâteaux derivative and their variances, one has that $\mathcal{S}_\delta[P]$ and $\mathcal{T}_\delta[P]$ are pathwise differentiable at P^*.

The *uncentered* efficient influence function (EIF) of $\mathcal{S}_\delta[P]$ at P^* evaluated at O_i is given by $D^S_\delta(O_i) := \frac{\mathrm{d}\mathcal{S}_\delta[P_\epsilon]}{\mathrm{d}\epsilon}\big|_{\epsilon=0} + \mathcal{S}_\delta[P^*]$, and can be computed using the using the chain rule and gradient algebra for the Gâteaux derivative, as follows:

$$D^S_\delta(O_i) = \sum_{a \in \mathcal{A}} \frac{1}{Z_\delta(W_i)^2} \left[Z_\delta(W_i) \left(s_\delta(W_i, a) \, Q'(O_i, a) + s'_\delta(W_i, a) \, Q(W_i, a)\right)\right.$$
$$\left. - s_\delta(W_i, a) \, Q(W_i, a) \, Z'_\delta(O_i)\right],$$

$$= \sum_{a \in \mathcal{A}} \left\{ \frac{s_\delta(W_i, a) \, Q'(O_i, a)}{Z_\delta(W_i)} + \frac{s'_\delta(W_i, a) \, Q(W_i, a)}{Z_\delta(W_i)} - \frac{s_\delta(W_i, a) \, Q(W_i, a) \, Z'_\delta(O_i)}{Z_\delta(W_i)^2} \right\},$$

$$Q'(O_i,a) = \frac{\mathbb{I}(a=A_i)}{\pi(a\,|\,W_i)}\,[Y_i - Q(W_i,a)] + Q(W_i,a),$$

$$s_\delta(W_i,a) = (\zeta_\delta + \xi_\delta(a))\,\pi(a\,|\,W_i), \qquad s'_\delta(W_i,a) = (\zeta_\delta + \xi_\delta(a))\,\mathbb{I}(a=A_i) - s_\delta(W_i,a),$$

$$Z_\delta(W_i) = \sum_{a'\in\mathcal{A}} \left(\zeta_\delta + \xi_\delta(a')\right)\,\pi(a'\,|\,W_i), \qquad Z'_\delta(O_i) = (\zeta_\delta + \xi_\delta(A_i)) - Z_\delta(W_i).$$

These expressions satisfy the following equivalences:

$$s_\delta(W_i,a')/Z_\delta(W_i) = \pi^*_\delta(a'\,|\,W_i) \quad \text{and} \quad (\zeta_\delta + \xi_\delta(a'))/Z_\delta(W_i) = \pi^*_\delta(a'\,|\,W_i)/\pi(a'\,|\,W_i).$$

Therefore,

$$\begin{aligned}
D^S_\delta(O_i) &= \sum_{a\in\mathcal{A}} \Bigg\{ \pi^*_\delta(a\,|\,W_i)\left[\frac{\mathbb{I}(a=A_i)}{\pi(a\,|\,W_i)}\,[Y_i - Q(W_i,a)] + Q(W_i,a)\right] \\
&\quad + \left[\pi^*_\delta(a\,|\,W_i)\frac{\mathbb{I}(a=A_i)}{\pi(a\,|\,W_i)} - \pi^*_\delta(a\,|\,W_i)\right] Q(W_i,a) - \pi^*_\delta(a\,|\,W_i)\,Q(W_i,a)\left[\frac{\pi^*_\delta(A_i\,|\,W_i)}{\pi(A_i\,|\,W_i)} - 1\right]\Bigg\}, \\
&= \frac{\pi^*_\delta(A_i\,|\,W_i)}{\pi(A_i\,|\,W_i)}\,[Y_i - Q(W_i,A_i)] + \sum_{a\in\mathcal{A}} \pi^*_\delta(a\,|\,W_i)\,Q(W_i,a) + \frac{\pi^*_\delta(A_i\,|\,W_i)}{\pi(A_i\,|\,W_i)}\,Q(W_i,A_i) \\
&\quad - \frac{\pi^*_\delta(A_i\,|\,W_i)}{\pi(A_i\,|\,W_i)} \sum_{a\in\mathcal{A}} \pi^*_\delta(a\,|\,W_i)\,Q(W_i,a), \\
&= \underbrace{\frac{\pi^*_\delta(A_i\,|\,W_i)}{\pi(A_i\,|\,W_i)}\left[Y_i - \sum_{a\in\mathcal{A}} \pi^*_\delta(a\,|\,W_i)\,Q(W_i,a)\right]}_{D^{S,1}_\delta(O_i)} + \underbrace{\sum_{a\in\mathcal{A}} \pi^*_\delta(a\,|\,W_i)\,Q(W_i,a)}_{D^{S,2}_\delta(O_i)}.
\end{aligned}$$

Analogously, the *uncentered* EIF of $\mathcal{T}_\delta[P]$ at P^* evaluated at point O_i is given by $D^T_\delta(O_i) := \frac{\mathrm{d}\mathcal{T}_\delta[P_\epsilon]}{\mathrm{d}\epsilon}\big|_{\epsilon=0} + \mathcal{T}_\delta[P^*]$, and can be computed as:

$$D^S_\delta(O_i) = \sum_{a\in\mathcal{A}} \left\{ \frac{t_\delta(W_i,a)\,Q'(O_i,a)}{Z_\delta(W_i)} + \frac{t'_\delta(W_i,a)\,Q(W_i,a)}{Z_\delta(W_i)} - \frac{t_\delta(W_i,a)\,Q(W_i,a)\,Z'_\delta(O_i)}{Z_\delta(W_i)^2} \right\},$$

where $t_\delta(W_i,a) = \nu(a) - \xi_\delta(a)(1 - \pi(a\,|\,W_i))$ and $t'_\delta(W_i,a) = \xi_\delta(a)\,[\mathbb{I}(a=A_i) - \pi(a\,|\,W_i)]$. These expressions satisfy the following equivalences:

$$\begin{aligned}
t_\delta(W_i,a')/Z_\delta(W_i) &= \nu^*_\delta(a'\,|\,W_i), \\
\xi_\delta(a')/Z_\delta(W_i) &= \varrho_\delta(a')\,\pi^*_\delta(a'\,|\,W_i)/\pi(a'\,|\,W_i), \text{ with} \\
\varrho_\delta(a') &= \xi_\delta(a')/(\zeta_\delta + \xi_\delta(a')).
\end{aligned}$$

Therefore,

$$\begin{aligned}
D^T_\delta(O_i) &= \sum_{a\in\mathcal{A}} \Bigg\{ \nu^*_\delta(a\,|\,W_i)\left[\frac{\mathbb{I}(a=A_i)}{\pi(a\,|\,W_i)}\,[Y_i - Q(W_i,a)] + Q(W_i,a)\right] \\
&\quad + \left[\varrho_\delta(a)\,\pi^*_\delta(a\,|\,W_i)\frac{\mathbb{I}(a=A_i)}{\pi(a\,|\,W_i)} - \varrho_\delta(a)\,\pi^*_\delta(a\,|\,W_i)\right] Q(W_i,a) - \nu^*_\delta(a\,|\,W_i)\,Q(W_i,a)\left[\frac{\pi^*_\delta(A_i\,|\,W_i)}{\pi(A_i\,|\,W_i)} - 1\right]\Bigg\}, \\
&= \frac{\nu^*_\delta(A_i\,|\,W_i)}{\pi(A_i\,|\,W_i)}\,[Y_i - Q(W_i,A_i)] + 2\sum_{a\in\mathcal{A}} \nu^*_\delta(a\,|\,W_i)\,Q(W_i,a) + \frac{\pi^*_\delta(A_i\,|\,W_i)}{\pi(A_i\,|\,W_i)}\,\varrho_\delta(A_i)\,Q(W_i,A_i) \\
&\quad - \sum_{a\in\mathcal{A}} \pi^*_\delta(a\,|\,W_i)\,\varrho_\delta(a)\,Q(W_i,a) - \frac{\pi^*_\delta(A_i\,|\,W_i)}{\pi(A_i\,|\,W_i)} \sum_{a\in\mathcal{A}} \nu^*_\delta(a\,|\,W_i)\,Q(W_i,a), \\
&= \underbrace{\frac{\nu^*_\delta(A_i\,|\,W_i)}{\pi(A_i\,|\,W_i)}\,[Y_i - Q(W_i,A_i)]}_{D^{T,1}_\delta(O_i)} + \underbrace{\left[2 - \frac{\pi^*_\delta(A_i\,|\,W_i)}{\pi(A_i\,|\,W_i)}\right] \sum_{a\in\mathcal{A}} \nu^*_\delta(a\,|\,W_i)\,Q(W_i,a)}_{D^{T,2}_\delta(O_i)} \\
&\quad + \underbrace{\frac{\pi^*_\delta(A_i\,|\,W_i)}{\pi(A_i\,|\,W_i)}\,\varrho_\delta(A_i)\,Q(W_i,A_i) - \sum_{a\in\mathcal{A}} \pi^*_\delta(a\,|\,W_i)\,\varrho_\delta(a)\,Q(W_i,a)}_{D^{T,3}_\delta(O_i)}.
\end{aligned}$$

B Empirical evaluation details

The data-generating process (DGP) used in the evaluation task with synthetic data is given by:

$$
\begin{aligned}
W &\overset{iid}{\sim} N(\vec{0}, I_4),\\
\eta'(W) &= \exp(-2W_1 + W_2 - 0.5W_3 - 0.25W_4),\\
\eta''(W) &= \exp(-W_1 + 0.25W_2 + 2W_3 + 0.5W_4),\\
\pi(\alpha_k \,|\, W) &= \eta_k(W) / \left[\eta'(W) + \eta''(W) + 1\right],\ k \in \{1, 2\},\\
A \,|\, W &\overset{iid}{\sim} \mathrm{Cat}_3\left(\pi(\alpha' \,|\, W), \pi(\alpha'' \,|\, W), 1 - \pi(\alpha' \,|\, W) - \pi(\alpha'' \,|\, W)\right),\\
q(W) &= 2W_1 + W_2 + W_3 + W_4,\\
Q(W, A) &= \begin{cases} 10 - 8.7\, q(W) & \text{if } A = a' \\ 40 + 17.4\, q(W) & \text{if } A = a'' \\ 50 + 26.1\, q(W) & \text{if } A = a_3 \end{cases},\\
Y &= Q(W, A) + \varepsilon,\ \text{where } \varepsilon \overset{iid}{\sim} N(0, 50).
\end{aligned}
$$

For the exposure model class, we use multinomial logistic regression with a linear predictor, and for the outcome we employ multivariate adaptive regression splines (MARS) with extra linear predictors W. These model classes are correctly specified in the sense that they contain the true propensity score π and outcome regression function Q, respectively. Following the approach of Kang and Schafer [26], we introduce ad hoc misspecification in π and Q by using the same model classes but replacing the original covariates $W \in \mathbb{R}^4$ with a nonlinear transformation $Z(W) \in \mathbb{R}^3$, which also constitutes a valid adjustment set, defined as:

$$
\begin{aligned}
Z_1 &= 10 + W_2/(1 + \exp(W_1)),\\
Z_2 &= (0.6 + W_1 W_3/25)^3,\\
Z_3 &= (W_2 + W_4 + 20)^2.
\end{aligned}
$$