OPBO: ORDER-PRESERVING BAYESIAN OPTIMIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Bayesian optimization (BO) is widely used for optimizing expensive black-box problems under limited evaluation budgets. However, its performance often degrades as the problem dimension increases. Most high-dimensional BO methods address dimensionality while retaining value-based surrogate objectives. We argue that the surrogate learning objective itself may also contribute to this difficulty. We therefore study replacing *value*-based surrogate objectives with *order*-based objectives in high dimensions and characterize when order learning becomes preferable to value regression. Our analysis shows that the advantage of order learning increases when the numerical representation is harder to fit and when the dimension is large relative to the number of observations. These results motivate Order-Preserving Bayesian Optimization (OPBO), which replaces value regression with neural ranking surrogates. We compare value-based and order-based surrogates across 20 synthetic instances and four real-world tasks under Standard BO, TuRBO, and HeSBO. Under the same evaluation budget, ranking surrogates generally perform better, with larger gains at higher dimensions. The source code is available at `https://anonymous.4open.science/r/OPBO-89B4/`.

1 INTRODUCTION

Black-box optimization searches for good solutions through input–output observations, without requiring an analytic expression or gradients of the objective (Frazier, 2018; Shahriari et al., 2016). For example, in chip design-space exploration, costly cycle-accurate simulations limit the number of designs that can be evaluated (Lv et al., 2024). Bayesian optimization (BO) fits a surrogate to previous observations and uses it to guide subsequent queries under a limited evaluation budget (Snoek et al., 2012; Frazier, 2018; Dewancker et al., 2016). BO has been widely applied to hyperparameter tuning (Snoek et al., 2012) and policy search in reinforcement learning (Calandra et al., 2016). In high dimensions, however, the evaluation budget is often too small to explore the search space adequately (Wang et al., 2013; Nayebi et al., 2019; Eriksson et al., 2019).

High-dimensional Bayesian optimization (HDBO) methods typically reduce the effective search dimension, impose structural assumptions on the surrogate, or restrict optimization to local regions (Wang et al., 2013; Nayebi et al., 2019; Eriksson et al., 2019; Eriksson & Jankowiak, 2021). Recent studies also show that Gaussian process (GP) fitting can be sensitive to length-scale initialization, and that dimension-scaled initialization can improve high-dimensional BO performance (Xu et al., 2025; Papenmeier et al., 2025). Across these approaches, numerical value prediction generally remains the surrogate learning task.

Within a fixed candidate pool, the best point is determined by the ordering of objective values rather than their numerical magnitudes. Strictly increasing transformations preserve this ordering while changing the targets presented to value regression. Ranking-based global optimization directly learns the ordering induced by the objective (Malherbe et al., 2016), while preference-based BO operates on comparative feedback rather than exact function values (Chu & Ghahramani, 2005; González et al., 2017). More recent work has applied ranking objectives even when numerical evaluations are available, including in Bayesian optimization and offline model-based optimization (Wang et al., 2023; Tom et al., 2025; Tan et al., 2025; Lyu et al., 2026). These studies establish that order can be an effective learning signal, but its theoretical advantages over value regression in high dimensions remain poorly understood.

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.

We study this question by comparing value regression and order learning under order-preserving transformations of the same realized observations. An exact decomposition expresses expected candidate-selection regret as a sum of objective gaps weighted by the probabilities of missing the corresponding top-ranked candidate sets. Under a Gaussian single-index model, we characterize when initial ranking directions achieve lower asymptotic latent selection regret than affine least squares. A matched one-step value comparison separates representation dependence from the effect of empirical covariance correction. These results motivate Order-Preserving Bayesian Optimization (OPBO), which replaces value regression with pairwise or listwise neural ranking surrogates within existing BO frameworks. Experiments with neural surrogates show that nonlinear models exhibit the same pattern predicted by our theory. Changing the numerical representation affects value regression while leaving the ranking objectives unchanged. End-to-end evaluations under Standard BO, TuRBO, and HeSBO further show that pairwise and listwise ranking surrogates achieve the two best mean ranks among five surrogate choices under matched evaluation budgets.

2 METHOD

2.1 OPBO FRAMEWORK

We consider the black-box optimization problem

$$x^\star \in \arg\min_{x\in\Omega} f(x), \qquad \Omega = [0,1]^d.$$

At iteration r, BO maintains observations $\mathcal{D}_r = \{(x_i, y_i)\}_{i=1}^{n_r}$, fits a surrogate, and uses it to select new points for evaluation. OPBO replaces value regression with a ranking surrogate and performs candidate selection using the resulting score or rank information, while leaving candidate generation and framework-specific search updates unchanged. Algorithm 1 summarizes the procedure.

2.2 ORDER-PRESERVING SURROGATE OBJECTIVE

For a finite candidate set, selecting a promising point depends on relative quality rather than an exact reconstruction of objective values. Motivated by this observation and following prior work (Lv et al., 2024), we formalize numerical representations that preserve this order information.

Definition 1 (Order-Preserving Functions)**.** Two functions $f, g : \Omega \to \mathbb{R}$ are *order-preserving*, written as $f \overset{\text{OP}}{\Longleftrightarrow} g$, if $f(a) \geq f(b) \Longleftrightarrow g(a) \geq g(b)$ for all $a, b \in \Omega$.

Equivalently, $f \overset{\text{OP}}{\Longleftrightarrow} g$ if and only if $g = h \circ f$ for some strictly increasing $h : f(\Omega) \to \mathbb{R}$, called the *OP link*. Thus g is an OP representation of f, and the two functions share the same ordering, ties, and minimizers. For example, $F(x_1, x_2) = \exp\big(-(x_1^2 + x_2^2)/(2\tau^2)\big)$ and $H(x_1, x_2) = -(x_1^2 + x_2^2)$ satisfy $H = 2\tau^2 \log F$, and therefore $F \overset{\text{OP}}{\Longleftrightarrow} H$.

Given $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^n$ and a strictly increasing h, define $\mathcal{D}^h = \{(x_i, h(y_i))\}_{i=1}^n$. The transformation preserves every comparison among the observations but generally changes a mean squared error (MSE) objective. For example,

$$\mathcal{L}_{\text{MSE}}(\theta; \mathcal{D}^h) = \frac{1}{n}\sum_{i=1}^{n}\big(h(y_i) - \hat{y}_\theta(x_i)\big)^2$$

need not yield the same predictor as $\mathcal{L}_{\text{MSE}}(\theta; \mathcal{D})$. Target standardization removes positive affine recodings, but not general nonlinear increasing transformations.

Proposition 2 (Pathwise Order-Preserving Invariance)**.** *Let $\mathcal{D}$ be any realized data set and let h be strictly increasing. Suppose that $\mathcal{L}_{\text{ord}}$ depends on the targets only through their ordering. Then*

$$\mathcal{L}_{\text{ord}}(\theta; \mathcal{D}^h) = \mathcal{L}_{\text{ord}}(\theta; \mathcal{D})$$

for every θ. If any tie-breaking or sampling randomness is coupled across the two data sets and depends only on their ordering, then the same training randomness produces the same optimization trajectory, fitted scores, and rank-based candidate selection.

Algorithm 1 Order-Preserving Bayesian Optimization

Require: Initial dataset $\mathcal{D}_0 = \{(x_1, y_1), \ldots, (x_k, y_k)\}$; evaluation budget B; batch size b; ranking loss ℓ_{ord}; acquisition rule α; ensemble size M

$\mathcal{D} \leftarrow \mathcal{D}_0$
while $|\mathcal{D}| < B$ **do**
  Fit ranking surrogates $\{s_{\theta_m}\}_{m=1}^{M}$ on $\mathcal{D}$ using ℓ_{ord}
  Generate candidate set $\mathcal{C}_t$
  $b_t \leftarrow \min\{b, B - |\mathcal{D}|\}$
  $\mathcal{Q}_t \leftarrow \text{Top}_{b_t}\{\alpha(x) : x \in \mathcal{C}_t\}$
  Evaluate $\{y(x) : x \in \mathcal{Q}_t\}$
  $\mathcal{D} \leftarrow \mathcal{D} \cup \{(x, y(x)) : x \in \mathcal{Q}_t\}$
  Update the search state when required
end while
return $\widehat{x}^\star \in \arg\min_{(x_i, y_i) \in \mathcal{D}} y_i$

2.3 Surrogate Model

We use larger scores for better candidates, so for minimization a smaller y_i receives a larger score $s_i = s_\theta(x_i)$. We consider ListMLE (Xia et al., 2008) and Pairwise (Burges et al., 2005).

For ListMLE, let π order observations from best to worst, $y_{\pi(1)} \le \cdots \le y_{\pi(n)}$, with ties broken independently of their numerical magnitudes. The probability of placing $\pi(i)$ at position i given the previous positions is $P_\theta(\pi(i) \mid \pi(1), \ldots, \pi(i-1)) = \frac{\exp(s_{\pi(i)})}{\sum_{k=i}^{n} \exp(s_{\pi(k)})}$. The training loss is

$$\mathcal{L}_{\text{ListMLE}}(\theta) = -\sum_{i=1}^{n-1} \log \frac{\exp\big(s_{\pi(i)}\big)}{\sum_{k=i}^{n} \exp\big(s_{\pi(k)}\big)}. \tag{1}$$

For Pairwise, let $\mathcal{P} = \{(i, j) : 1 \le i < j \le n,\ y_i \ne y_j\}$ be the non-tied pairs and $q_{ij} = \text{sign}(y_j - y_i)$, so $q_{ij} = 1$ means observation i is better than j. The training loss is

$$\mathcal{L}_{\text{Pair}}(\theta) = \frac{1}{|\mathcal{P}|} \sum_{(i,j) \in \mathcal{P}} \log\left(1 + \exp\left[-q_{ij}(s_i - s_j)\right]\right). \tag{2}$$

Both losses depend only on the observed ordering and therefore satisfy Proposition 2.

2.4 Acquisition Function

The learned scores are ordinal rather than calibrated, so candidate selection uses only their induced ordering. Standard BO and TuRBO sample one ensemble member and select candidates with the largest scores, with exploration induced by variation across ensemble members. HeSBO instead measures improvement directly in rank space.

For HeSBO, let $\mathcal{X}_r^{\text{obs}} = \{x_i : (x_i, y_i) \in \mathcal{D}_r\}$ be the observed inputs and $\mathcal{C}_r$ the current candidate set. Each ensemble member jointly scores $\mathcal{A}_r = \mathcal{X}_r^{\text{obs}} \cup \mathcal{C}_r$, so that observations and candidates are ranked together. For member m, let $R_m(x)$ denote the rank of x when $\mathcal{A}_r$ is sorted by decreasing score, with $R_m(x) = 1$ denoting the best predicted point. Define the normalized rank

$$\bar{R}_m(x) = 1 - \frac{R_m(x) - 1}{|\mathcal{A}_r| - 1} \in [0, 1],$$

and the best observed rank as

$$\bar{R}_m^\star = \max_{x \in \mathcal{X}_r^{\text{obs}}} \bar{R}_m(x).$$

The rank-based acquisition value of $x \in \mathcal{C}_r$ is

$$\alpha_{\text{rank}}(x) = \frac{1}{M} \sum_{m=1}^{M} \left[\bar{R}_m(x) - \bar{R}_m^\star - \xi\right]_+, \tag{3}$$

where M is the ensemble size, $\xi \geq 0$ is an improvement margin, and $[z]_+ = \max\{z, 0\}$. HeSBO selects the candidate with the largest $\alpha_{\mathrm{rank}}(x)$. The acquisition measures the ensemble-averaged positive improvement over the best observed rank, providing a rank-space analogue of improvement-based selection without requiring calibrated objective values.

3 Theoretical Analysis

To understand when order learning helps, we study objectives with the same ordering but different numerical representations. The analysis proceeds from the cost of ranking errors, to order-preserving transformations that make MSE estimation harder, to candidate-selection regret.

3.1 What Candidate Selection Requires

Condition on $\mathcal{D}_t$ and a realized candidate set $\mathcal{C}_t$. Let $f_{t,(1)} \leq \cdots \leq f_{t,(N)}$ be the candidate values sorted from best to worst, and let K_t be the rank of the selected best candidate, with selection regret $r_t^{\mathrm{sel}} = f_{t,(K_t)} - f_{t,(1)}$.

Proposition 3 (Exact Selection Regret Decomposition)**.** *For every rule that selects a nonempty batch from $\mathcal{C}_t$,*

$$\mathbb{E}[r_t^{\mathrm{sel}} \mid \mathcal{D}_t, \mathcal{C}_t] = \sum_{g=1}^{N-1} \big(f_{t,(g+1)} - f_{t,(g)}\big)\mathbb{P}(K_t > g \mid \mathcal{D}_t, \mathcal{C}_t). \tag{4}$$

The gaps $f_{t,(g+1)} - f_{t,(g)}$ measure the cost of crossing rank boundary g, and the probabilities measure how often it is crossed. Following Ho et al. (2007), we describe the cost by an ordered performance curve.

Definition 4 (Ordered Performance Curve)**.** Assume $f_{t,(N)} > f_{t,(1)}$. For $i = 1, \ldots, N$, define the normalized rank $\rho_i = (i-1)/(N-1)$ and normalized value

$$v_{t,i} = \frac{f_{t,(i)} - f_{t,(1)}}{f_{t,(N)} - f_{t,(1)}}.$$

The *ordered performance curve* (OPC) is the piecewise linear curve through the points $(\rho_i, v_{t,i})$.

The quantity $v_{t,g+1} - v_{t,g}$ is the normalized cost of crossing the g-th boundary. To describe how often boundaries are crossed, we define the selection failure curve.

Definition 5 (Selection Failure Curve)**.** For $g = 1, \ldots, N-1$, define

$$\Psi_t(g) = \mathbb{P}(K_t > g \mid \mathcal{D}_t, \mathcal{C}_t).$$

The *selection failure curve* (SFC) gives the probability that the selected batch misses all of the best g candidates.

In normalized form, Proposition 3 becomes

$$\frac{\mathbb{E}[r_t^{\mathrm{sel}} \mid \mathcal{D}_t, \mathcal{C}_t]}{f_{t,(N)} - f_{t,(1)}} = \sum_{g=1}^{N-1} (v_{t,g+1} - v_{t,g})\Psi_t(g). \tag{5}$$

Thus the OPC describes where ranking errors are costly and the SFC how often they occur. Candidate selection does not require accurate value prediction, but rather avoiding ranking errors across costly boundaries. This raises whether different numerical representations of the same ordering change how difficult these errors are to avoid.

3.2 A High-Dimensional Value–Order Crossover

Let $X \sim \mathcal{N}(0, I_d)$, and let a unit vector $w_\star$ define the clean latent score $w_\star^\top X$. We observe $Y = h(S)$, where $S = w_\star^\top X + \varepsilon$ and $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ is independent of X. Here h is strictly increasing. We compare representations on the same n independent and identically distributed realizations of (X, S), so changing h alters the numerical targets but preserves every observed comparison.

For affine MSE regression, we quantify representation difficulty through the following residual-to-signal ratio.

Proposition 6 (Value Representation Difficulty)**.** *Suppose h is continuously differentiable and strictly increasing, with $\mathbb{E}[h(S)^2] < \infty$ and $0 < \mathbb{E}[h'(S)] < \infty$. Define*

$$\mathcal{D}_{\mathrm{val}}(h) = \frac{\mathrm{Var}[h(S)]}{\big(\mathbb{E}h'(S)\big)^2} - 1 = \frac{1+\sigma^2}{\mathrm{Corr}(S, h(S))^2} - 1. \tag{6}$$

Then $\mathcal{D}_{\mathrm{val}}(h) \geq \sigma^2$, with equality if and only if h is affine.

The population affine fit has slope $(\mathbb{E}h'(S))w_\star$, and $\mathcal{D}_{\mathrm{val}}(h)$ is its residual variance divided by $(\mathbb{E}h'(S))^2$.

We next study how this representation difficulty affects directional estimation when $d, n \to \infty$ with $d/n \to \gamma \in (0,1)$. All directions below are oriented as costs so that smaller $v^\top x$ indicates a better candidate. Let $\widehat{v}_{\mathrm{val}}$ be the normalized slope of the affine least-squares solution. For $k \in \{\mathrm{Pair}, \mathrm{List}\}$, let $\widehat{v}_{\mathrm{ord},k}$ be the normalized slope gradient of the corresponding ranking loss at the zero affine score. These ranking directions use the targets only through their ranks and are therefore invariant to h; their explicit forms are given in Appendix B.4. A matched one-step value comparison below isolates the effect of covariance correction.

Theorem 7 (High-Dimensional Value–Order Crossover)**.** *Under the Gaussian model and the conditions of Proposition 6, let h and σ be fixed. If $d, n \to \infty$ with $d/n \to \gamma \in (0,1)$, then*

$$\widehat{v}_{\mathrm{val}}^\top w_\star \xrightarrow{p} \left(1 + \mathcal{D}_{\mathrm{val}}(h)\frac{\gamma}{1-\gamma}\right)^{-1/2}, \tag{7}$$

$$\widehat{v}_{\mathrm{ord},k}^\top w_\star \xrightarrow{p} \big(1 + C_k(1+\sigma^2)\gamma\big)^{-1/2}, \qquad k \in \{\mathrm{Pair}, \mathrm{List}\}, \tag{8}$$

where $C_{\mathrm{Pair}} = \pi/3$, $C_{\mathrm{List}} \approx 1.226$. Consequently, the ranking direction k has strictly larger limiting alignment with $w_\star$ than $\widehat{v}_{\mathrm{val}}$ if and only if

$$\mathcal{D}_{\mathrm{val}}(h) > C_k(1+\sigma^2)(1-\gamma). \tag{9}$$

In the limiting alignments, the dimension-dependent factor is $\gamma/(1-\gamma)$ for least squares and γ for the ranking directions. Increasing either $\mathcal{D}_{\mathrm{val}}(h)$ or γ makes Eq. (9) easier to satisfy.

A matched one-step value control. For a matched comparison with the initial ranking directions, take one gradient step on the squared loss $\ell(\beta, b) = (2n)^{-1}\sum_{i=1}^n (Y_i - b - \beta^\top X_i)^2$. At $(\beta, b) = (0, \bar{Y})$, the negative slope gradient and its normalized direction are

$$\widehat{q}_{\mathrm{val},1} = \frac{1}{n}\sum_{i=1}^n X_i(Y_i - \bar{Y}), \qquad \widehat{v}_{\mathrm{val},1} = \frac{\widehat{q}_{\mathrm{val},1}}{\|\widehat{q}_{\mathrm{val},1}\|}. \tag{10}$$

Centering the targets makes this direction invariant to additive target shifts. The affine least-squares slope satisfies $\widehat{\beta}_{\mathrm{val}} = \widehat{\Sigma}_X^{-1}\widehat{q}_{\mathrm{val},1}$, where $\widehat{\Sigma}_X = n^{-1}\sum_i (X_i - \bar{X})(X_i - \bar{X})^\top$. Thus the two value estimators differ only in empirical covariance correction.

Proposition 8 (Matched One-Step Comparison)**.** *Under the assumptions of Theorem 7, the one-step value direction satisfies*

$$\widehat{v}_{\mathrm{val},1}^\top w_\star \xrightarrow{p} \left[1 + \gamma\{1 + \mathcal{D}_{\mathrm{val}}(h)\}\right]^{-1/2}. \tag{11}$$

For $k \in \{\mathrm{Pair}, \mathrm{List}\}$, the ranking direction has strictly larger limiting alignment than $\widehat{v}_{\mathrm{val},1}$ if and only if

$$1 + \mathcal{D}_{\mathrm{val}}(h) > C_k(1+\sigma^2). \tag{12}$$

Equation (11) shows that representation dependence persists without covariance inversion. Unlike Eq. (9), the matched-step threshold (12) is independent of γ, although the one-step alignments still depend on it.

3.3 From Directional Alignment to Candidate Selection

We now translate the directional limits in Section 3.2 into candidate-selection regret.

Order-preserving transformations can change the objective gaps and OPC in Section 3.1. We therefore evaluate selection regret on the common clean latent scale $w_\star^\top X$.

Let $X_1^{\mathrm{c}}, \dots, X_N^{\mathrm{c}}$ be independent $\mathcal{N}(0, I_d)$ candidates, independent of the training sample, with fixed $N \geq 2$. For a unit direction v independent of this pool, let $\widehat{j}(v) = \arg\min_{1 \leq j \leq N} v^\top X_j^{\mathrm{c}}$. Define its latent selection regret as

$$R_{\mathrm{lat}}(v) = \mathbb{E}\left[w_\star^\top X^{\mathrm{c}}_{\widehat{j}(v)} - \min_j w_\star^\top X_j^{\mathrm{c}} \,\middle|\, v \right]. \tag{13}$$

The expectation is over the candidate pool conditional on v.

Lemma 9 (Alignment Determines Selection Regret)**.** *Let $\rho = v^\top w_\star$. For a Gaussian candidate pool of fixed size N,*

$$R_{\mathrm{lat}}(v) = \kappa_N(1 - \rho), \tag{14}$$

where $\kappa_N = -\mathbb{E}[\min_{1 \leq j \leq N} Z_j] > 0$ for independent $Z_j \sim \mathcal{N}(0,1)$.

Combining Lemma 9 with Theorem 7 gives the following selection-regret limits.

Corollary 10 (High-Dimensional Candidate-Selection Crossover)**.** *Under the assumptions of Theorem 7, let the candidate pool be as above, with fixed $N \geq 2$. For $k \in \{\mathrm{Pair}, \mathrm{List}\}$, the latent selection regrets satisfy*

$$R_{\mathrm{lat}}(\widehat{v}_{\mathrm{val}}) \xrightarrow{p} \kappa_N \left[1 - \left(1 + \mathcal{D}_{\mathrm{val}}(h) \frac{\gamma}{1-\gamma} \right)^{-1/2} \right],$$
$$R_{\mathrm{lat}}(\widehat{v}_{\mathrm{ord},k}) \xrightarrow{p} \kappa_N \left[1 - \left(1 + C_k(1+\sigma^2)\gamma \right)^{-1/2} \right].$$

Consequently, the ranking direction k has strictly smaller limiting latent selection regret than $\widehat{v}_{\mathrm{val}}$ if and only if Eq. (9) *holds.*

When $\sigma > 0$, Proposition 6 gives $\mathcal{D}_{\mathrm{val}}(h) \geq \sigma^2 > 0$, so both ranking directions have strictly smaller limiting latent regret than affine least squares for every fixed admissible representation when γ is sufficiently close to one. For direction k, this limiting advantage holds for every $\gamma \in (0,1)$ if and only if $\mathcal{D}_{\mathrm{val}}(h) \geq C_k(1+\sigma^2)$. In contrast, for a noiseless affine representation, $\mathcal{D}_{\mathrm{val}}(h) = 0$, and least squares recovers the latent direction almost surely when $n > d$, giving zero latent selection regret.

For the matched one-step value direction, Lemma 9 and Proposition 8 give

$$R_{\mathrm{lat}}(\widehat{v}_{\mathrm{val},1}) \xrightarrow{p} \kappa_N \left[1 - \{1 + \gamma(1 + \mathcal{D}_{\mathrm{val}}(h))\}^{-1/2} \right]. \tag{15}$$

For each $k \in \{\mathrm{Pair}, \mathrm{List}\}$, the ranking direction has strictly smaller limiting latent selection regret than $\widehat{v}_{\mathrm{val},1}$ if and only if Eq. (12) holds.

3.4 Finite-Sample Verification of the Value–Order Crossover

We test whether the asymptotic value–order crossover from Section 3.2 appears at finite sample sizes. Under the Gaussian single-index model, we fix $n = 300$, $\sigma = 0.5$, and candidate-pool size $N = 200$, vary $\gamma = d/n \in \{0.1, \dots, 0.9\}$, and run 100 paired trials.

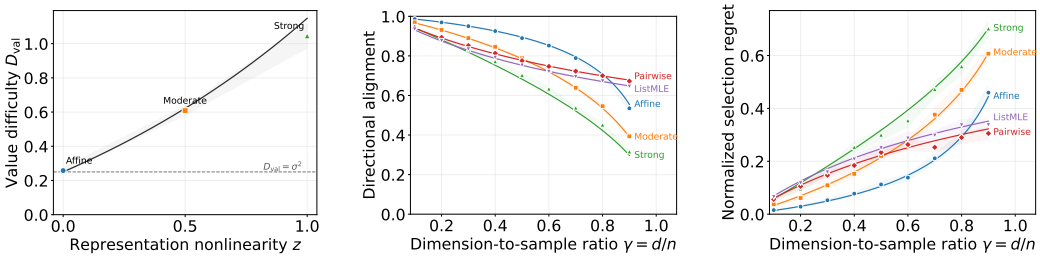
To vary numerical representation while preserving the observed ordering, let $\tau^2 = 1 + \sigma^2$ and define

$$h_z(s) = \frac{\exp(\sqrt{z}\, s/\tau - z/2) - 1}{\sqrt{e^z - 1}}, \qquad h_0(s) = s/\tau,$$

for $z \in \{0, 0.5, 1\}$. All three representations have zero population mean and unit population variance, while their representation difficulty is

$$\mathcal{D}_{\mathrm{val}}(h_z) = (1+\sigma^2)\frac{e^z - 1}{z} - 1,$$

with $\mathcal{D}_{\mathrm{val}}(h_0) = \sigma^2$. Larger $\mathcal{D}_{\mathrm{val}}$ corresponds to a harder affine value-regression problem. We evaluate directional alignment with $w_\star$ and normalized latent selection regret.



(a) Representation difficulty (b) Directional alignment (c) Latent selection regret

Figure 1: Finite-sample verification of the high-dimensional value–order crossover over 100 paired trials. **(a)** Representation difficulty. **(b)** Directional alignment with $w_\star$. **(c)** Normalized latent selection regret. Dashed curves show theoretical predictions, solid lines connect finite-sample means, and error bars show pointwise 95% confidence intervals across trials.

Figure 1a shows that representation difficulty increases with z. Figure 1b shows that affine least squares loses alignment more rapidly as both γ and representation nonlinearity increase, so the crossover with Pairwise and ListMLE occurs at smaller γ. The ranking directions are unchanged across recodings, consistent with Proposition 2. Figure 1c shows the corresponding effect on candidate selection, with the finite-sample curves closely following the theoretical predictions.

Matched estimator control. The comparison above also differs in estimator structure. Affine least squares applies the empirical covariance correction $\widehat{\Sigma}_X^{-1}$, whereas the initial Pairwise and ListMLE directions do not. To remove this difference, we compare them with the one-step MSE direction, so none of the three estimators inverts the empirical covariance.

Using the same Gaussian model, with 100 training trials per dimension and five candidate pools of size 200 per trial, Proposition 8 predicts a ranking advantage when

$$\frac{e^z - 1}{z} > C_k, \qquad k \in \{\mathrm{Pair}, \mathrm{List}\}.$$

Thus $z = 0$ predicts no ranking advantage, whereas $z = 0.5$ and $z = 1$ satisfy the crossover condition.

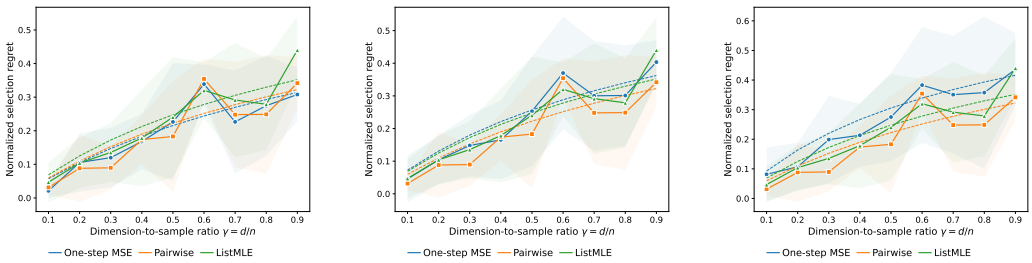
Figure 2 follows this pattern. One-step MSE is competitive under the affine representation in Figure 2a, while the ranking directions become more favorable in Figures 2b and 2c. Hence covariance inversion can amplify the high-dimensional deterioration of least squares, but is not required for representation sensitivity to arise.

Figure 2 is consistent with these predictions. Under the affine representation in Figure 2a, one-step MSE generally has the lowest regret. As the numerical representation becomes nonlinear in Figures 2b and 2c, the ranking directions become more favorable, with Pairwise showing a particularly clear advantage. The finite-sample curves are noisier than the asymptotic predictions, but the overall change across z follows the predicted crossover. Thus the representation effect does not arise solely from covariance inversion in ordinary least squares. Covariance correction strengthens the high-dimensional deterioration of OLS, but representation sensitivity is already present at the level of the one-step value direction.

3.5 Neural Verification of Representation Sensitivity

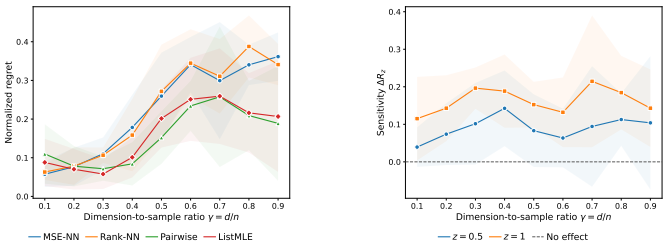
The matched control shows that representation sensitivity is not created by covariance inversion. It remains to check whether the effect is specific to affine estimators. A sufficiently expressive nonlinear surrogate could in principle absorb part of a monotone distortion of the numerical targets. We therefore repeat the controlled experiment with nonlinear neural surrogates.

We use the same neural architecture and training budget as in the end-to-end evaluation, and compare MSE-NN, Rank-NN, Pairwise, and ListMLE under the same controlled Gaussian setup. Within each paired trial, the methods share the training inputs, latent scores, candidate pools, and network initializations, while only the numerical recoding h_z is changed.



(a) Affine: $z = 0$. (b) Moderate: $z = 0.5$. (c) Strong: $z = 1$.

Figure 2: Matched one-step comparison on the common latent scale. None of the three estimators uses empirical covariance inversion. Solid curves show finite-sample mean normalized latent selection regret, dashed curves show high-dimensional predictions, and shaded regions show pointwise 95% confidence intervals. Lower is better.



(a) Affine representation ($z = 0$). (b) Paired representation effect.

Figure 3: Neural verification of representation sensitivity. **(a)** Normalized latent selection regret for $z = 0$. **(b)** Paired regret change of MSE-NN under nonlinear recodings. Rank-based surrogates remain invariant. Shaded regions show pointwise 95% confidence intervals over 30 paired trials.

Figure 3a shows performance under the affine representation $z = 0$. At small dimension-to-sample ratios, MSE-NN remains competitive with the rank-based surrogates, so direct ranking is not uniformly superior. Figure 3b measures the increase in MSE-NN regret relative to the affine representation, $\Delta R_z(\gamma) = R_z(\gamma) - R_0(\gamma)$. The increase is positive across most settings and is generally larger for $z = 1$ than for $z = 0.5$, showing that MSE-NN becomes more sensitive as the numerical representation becomes more nonlinear. The zero line corresponds to no representation effect, as expected for the rank-based surrogates.

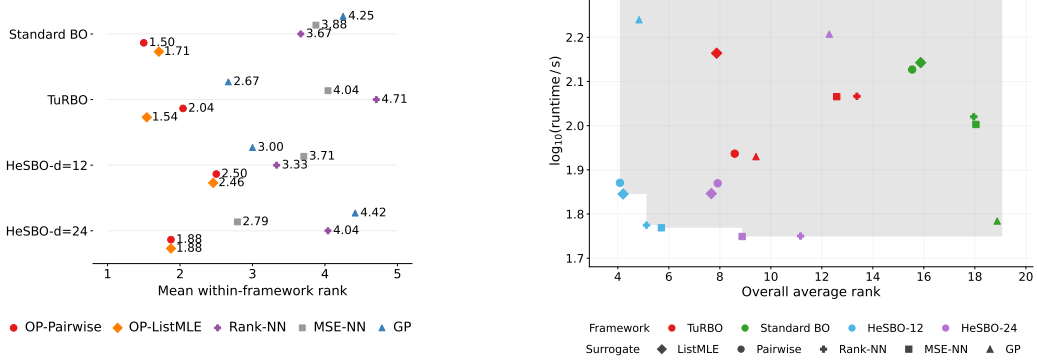
4 Comprehensive Evaluation

4.1 Experimental Setting

Benchmarks. We evaluate on 24 high-dimensional optimization problems. The synthetic suite contains Ackley, Levy, Rosenbrock, and Dixon–Price, each evaluated at $d \in \{200, 400, 600, 800, 1000\}$, giving 20 synthetic instances. The four real-world tasks are Rover (Wang et al., 2018), MOPTA08 (Jones & Anjos, 2025), Walker (Towers et al., 2024; Todorov et al., 2012), and DNA from LassoBench (Šehić et al., 2022). All tasks are formulated as minimization problems.

Surrogates and protocol. The purpose of this experiment is to isolate the effect of the surrogate learning objective rather than to compare different outer BO frameworks. We therefore replace the surrogate while keeping the remaining components of each framework fixed.

We compare five surrogate choices: GP, MSE-NN, Rank-NN, Pairwise, and ListMLE. MSE-NN performs numerical value regression, Rank-NN regresses rank-derived targets, and Pairwise and ListMLE learn the observed ordering directly. The same five choices are evaluated under Standard BO, TuRBO, and HeSBO with $d_{\text{emb}} \in \{12, 24\}$.



(a) Mean surrogate rank within each outer BO framework over 24 instances. Lower is better.

(b) End-to-end runtime versus average optimization rank across algorithm configurations. Lower is better on both axes.

Figure 4: End-to-end comparison of five surrogate choices across 24 benchmark instances and four outer BO frameworks.

Within each outer framework, all variants share the same initial design, candidate-generation rule, evaluation budget, and framework-specific updates. Each configuration uses 10 seeds, 40 shared Latin hypercube initial points (McKay et al., 1979), and a total budget of 500 evaluations.

4.2 Optimization Performance

Because objective scales differ across the 24 tasks, we summarize performance using within-framework ranks. For each task, the five surrogate variants are ranked by their median final incumbent over 10 trials, and these ranks are averaged across tasks. Figure 4a shows that Pairwise and ListMLE obtain the two best mean ranks in all four outer frameworks, while Rank-NN performs worse than both direct ranking losses. This suggests that replacing values with rank-derived regression targets alone does not reproduce the benefit of directly learning the ordering. Neither Pairwise nor ListMLE consistently dominates the other. Figure 4b shows that these gains do not come with a systematic runtime increase relative to the other neural surrogates. Full optimization trajectories are provided in Appendix E.

5 Conclusion

Bayesian optimization ultimately requires identifying promising candidates rather than reproducing their objective values exactly. We show that different numerical representations of the same ordering can substantially change the statistical efficiency of value regression, even when their first two moments are matched. Under the Gaussian single-index model, this yields a value–order crossover governed jointly by representation difficulty and the dimension-to-sample ratio. Harder numerical representations and larger d/n increasingly favor order-based estimation. Matched one-step comparisons show that this effect is not created by empirical covariance inversion, and neural experiments show that it persists beyond affine estimators. In end-to-end optimization, Pairwise and ListMLE achieve the two best mean ranks under matched evaluation budgets across all four outer BO settings. Together, these results suggest that direct order learning is particularly useful when fitting objective magnitudes introduces representation complexity that is unnecessary for candidate selection.

AI use statement

We used generative AI tools to assist with theoretical analysis and derivation checks, method development and experimental design, code implementation and debugging, interpretation of diagnostic results, manuscript drafting and editing, literature identification, and reference formatting and cross-

checking. AI tools were not used to generate synthetic data, translate the manuscript, clean or reformat data, or perform qualitative or thematic analysis; the study did not involve surveys, interviews, or transcription. The authors independently verified all mathematical statements and proofs, reviewed and tested AI-assisted code, and manually checked AI-suggested references. All reported numerical results were obtained from code executed by the authors, and the authors take full responsibility for the manuscript, its mathematical claims, experimental results, and accompanying artifacts.

REFERENCES

Chris Burges, Tal Shaked, Erin Renshaw, Ari Lazier, Matt Deeds, Nicole Hamilton, and Greg Hullender. Learning to rank using gradient descent. In *Proceedings of the 22nd International Conference on Machine Learning*, pp. 89–96. ACM, 2005.

Roberto Calandra, André Seyfarth, Jan Peters, and Marc Peter Deisenroth. Bayesian optimization for learning gaits under uncertainty. *Annals of Mathematics and Artificial Intelligence*, 76(1–2):5–23, 2016.

Wei Chu and Zoubin Ghahramani. Preference learning with Gaussian processes. In *Proceedings of the 22nd International Conference on Machine Learning*, volume 119 of *ACM International Conference Proceeding Series*, pp. 137–144. ACM, 2005.

Ian Dewancker, Michael McCourt, and Scott Clark. Bayesian optimization for machine learning: A practical guidebook. *arXiv preprint arXiv:1612.04858*, 2016.

David Eriksson and Martin Jankowiak. High-dimensional Bayesian optimization with sparse axis-aligned subspaces. In *Proceedings of the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence*, volume 161 of *Proceedings of Machine Learning Research*, pp. 493–503. PMLR, 2021.

David Eriksson, Michael Pearce, Jacob R. Gardner, Ryan D. Turner, and Matthias Poloczek. Scalable global optimization via local Bayesian optimization. In *Advances in Neural Information Processing Systems*, volume 32, pp. 5496–5507. Curran Associates, Inc., 2019.

Peter I. Frazier. A tutorial on Bayesian optimization. *arXiv preprint arXiv:1807.02811*, 2018.

Javier González, Zhenwen Dai, Andreas Damianou, and Neil D. Lawrence. Preferential Bayesian optimization. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 1282–1291. PMLR, 2017.

Yu-Chi Ho, Qian-Chuan Zhao, and Qing Shan Jia. *Ordinal Optimization: Soft Optimization for Hard Problems*. Springer, New York, NY, USA, 2007.

Donald Jones and Miguel Anjos. MOPTA 2008 jones benchmark. Dataset, Edinburgh DataShare, 2025. URL `https://doi.org/10.7488/ds/7894`.

Donald R. Jones, Matthias Schonlau, and William J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4):455–492, 1998.

Xiaoliang Lv, Qiaozhu Zhai, Jianchen Hu, Yuhang Zhu, Jinhui Liu, and Xiaohong Guan. An efficient binary programming method for black-box optimization and its application in processor design. *Science China Information Sciences*, 67(12):222102, 2024.

Shen-Huan Lyu, Rong-Xi Tan, Ke Xue, Yi-Xiao He, Yu Huang, Qingfu Zhang, and Chao Qian. On the learnability of offline model-based optimization: A ranking perspective. *arXiv preprint arXiv:2603.04000*, 2026.

Cédric Malherbe, Emile Contal, and Nicolas Vayatis. A ranking approach to global optimization. In *Proceedings of the 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pp. 1539–1547. PMLR, 2016.

Michael D. McKay, Richard J. Beckman, and William J. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245, 1979.

Samuel Müller, Matthias Feurer, Noah Hollmann, and Frank Hutter. PFNs4BO: In-context learning for Bayesian optimization. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 25444–25470. PMLR, 2023.

Amin Nayebi, Alexander Munteanu, and Matthias Poloczek. A framework for Bayesian optimization in embedded subspaces. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 4752–4761. PMLR, 2019.

Yann Ollivier, Ludovic Arnold, Anne Auger, and Nikolaus Hansen. Information-geometric optimization algorithms: A unifying picture via invariance principles. *Journal of Machine Learning Research*, 18(18):1–65, 2017.

Leonard Papenmeier, Matthias Poloczek, and Luigi Nardi. Understanding high-dimensional Bayesian optimization. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 47902–47923. PMLR, 2025.

Dat Phan-Trong, Hung Tran-The, and Sunil Gupta. NeuralBO: A black-box optimization algorithm using deep neural networks. *Neurocomputing*, 559:126776, 2023.

Kenan Šehić, Alexandre Gramfort, Joseph Salmon, and Luigi Nardi. LassoBench: A high-dimensional hyperparameter optimization benchmark suite for Lasso. In *Proceedings of the First International Conference on Automated Machine Learning*, volume 188 of *Proceedings of Machine Learning Research*, pp. 1–24. PMLR, 2022.

Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P. Adams, and Nando de Freitas. Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE*, 104(1): 148–175, 2016.

Jasper Snoek, Hugo Larochelle, and Ryan P. Adams. Practical Bayesian optimization of machine learning algorithms. In *Advances in Neural Information Processing Systems*, volume 25, pp. 2951–2959. Curran Associates, Inc., 2012.

Jasper Snoek, Oren Rippel, Kevin Swersky, Ryan Kiros, Nadathur Satish, Narayanan Sundaram, Md. Mostofa Ali Patwary, Prabhat, and Ryan P. Adams. Scalable Bayesian optimization using deep neural networks. In *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pp. 2171–2180. PMLR, 2015.

Niranjan Srinivas, Andreas Krause, Sham M. Kakade, and Matthias W. Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. In *Proceedings of the 27th International Conference on Machine Learning*, pp. 1015–1022. Omnipress, 2010.

Rong-Xi Tan, Ke Xue, Shen-Huan Lyu, Haopu Shang, Yao Wang, Yaoyuan Wang, Sheng Fu, and Chao Qian. Offline model-based optimization by learning to rank. In *International Conference on Learning Representations*, 2025.

Louis C. Tiao, Aaron Klein, Matthias W. Seeger, Edwin V. Bonilla, Cedric Archambeau, and Fabio Ramos. BORE: Bayesian optimization by density-ratio estimation. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 10289–10300. PMLR, 2021.

Emanuel Todorov, Tom Erez, and Yuval Tassa. MuJoCo: A physics engine for model-based control. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5026–5033. IEEE, 2012.

Gary Tom, Stanley Lo, Samantha Corapi, Alán Aspuru-Guzik, and Benjamin Sanchez-Lengeling. Ranking over regression for Bayesian optimization and molecule selection. *APL Machine Learning*, 3(3):036113, 2025.

Mark Towers, Ariel Kwiatkowski, Jordan Terry, John U. Balis, Gianluca De Cola, Tristan Deleu, Manuel Goulão, Andreas Kallinteris, Markus Krimmel, Arjun KG, Rodrigo Perez-Vicente, Andrea Pierré, Sander Schulhoff, Jun Jet Tai, Hannah Tan, and Omar G. Younis. Gymnasium: A standard interface for reinforcement learning environments. *arXiv preprint arXiv:2407.17032*, 2024.

Joel A. Tropp. A comparison principle for functions of a uniformly random subspace. *Probability Theory and Related Fields*, 153(3–4):759–769, 2012.

Xiaoxing Wang, Jiaxing Li, Chao Xue, Wei Liu, Weifeng Liu, Xiaokang Yang, Junchi Yan, and Dacheng Tao. Poisson process for Bayesian optimization. In *Proceedings of the Second International Conference on Automated Machine Learning*, volume 224 of *Proceedings of Machine Learning Research*, pp. 1–20. PMLR, 2023.

Zi Wang, Clement Gehring, Pushmeet Kohli, and Stefanie Jegelka. Batched large-scale Bayesian optimization in high-dimensional spaces. In *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*, volume 84 of *Proceedings of Machine Learning Research*, pp. 745–754. PMLR, 2018.

Ziyu Wang, Masrour Zoghi, Frank Hutter, David Matheson, and Nando de Freitas. Bayesian optimization in high dimensions via random embeddings. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence*, pp. 1778–1784, 2013.

Fen Xia, Tie-Yan Liu, Jue Wang, Wensheng Zhang, and Hang Li. Listwise approach to learning to rank: Theory and algorithm. In *Proceedings of the 25th International Conference on Machine Learning*, pp. 1192–1199. ACM, 2008.

Zhitong Xu, Haitao Wang, Jeff M. Phillips, and Shandian Zhe. Standard Gaussian process is all you need for high-dimensional Bayesian optimization. In *International Conference on Learning Representations*, 2025.

Yisong Yue, Josef Broder, Robert Kleinberg, and Thorsten Joachims. The K-armed dueling bandits problem. *Journal of Computer and System Sciences*, 78(5):1538–1556, 2012.

A Related Work

Classical BO models numerical objective values with probabilistic surrogates, most commonly GPs, and constructs acquisition functions from their predictive means and uncertainties (Jones et al., 1998; Srinivas et al., 2010). Neural and pretrained surrogates provide alternatives to GPs while generally retaining numerical value prediction (Snoek et al., 2015; Phan-Trong et al., 2023; Müller et al., 2023). HDBO methods address scalability through random embeddings (Wang et al., 2013; Nayebi et al., 2019), sparse or structured surrogates (Eriksson & Jankowiak, 2021), or adaptive local trust regions (Eriksson et al., 2019). These methods differ substantially in how they represent or search the input space, while generally retaining value regression as the surrogate learning objective.

Ordinal Optimization exploits relative performance to identify sufficiently good alternatives (Ho et al., 2007), while RankOpt directly learns the ranking induced by the objective (Malherbe et al., 2016). Information Geometric Optimization constructs optimization dynamics using invariance to increasing objective transformations (Ollivier et al., 2017). OPBO uses standard pairwise and listwise ranking losses based on RankNet and ListMLE (Burges et al., 2005; Xia et al., 2008).

Preference-based BO and dueling methods take comparisons, rather than numerical objective values, as feedback (Chu & Ghahramani, 2005; González et al., 2017; Yue et al., 2012). When scalar evaluations are available, BORE replaces value regression with quantile-based classification (Tiao et al., 2021), while PoPBO models candidate rankings probabilistically (Wang et al., 2023). Ranking objectives have also been studied in molecular optimization (Tom et al., 2025), offline model-based optimization (Tan et al., 2025; Lyu et al., 2026), and processor design-space exploration (Lv et al., 2024). Building on this line of work, we study how order-preserving transformations of the same realized observations affect the high-dimensional selection efficiency of affine least squares and initial ranking directions. A matched one-step value control distinguishes representation dependence from empirical covariance correction.

B PROOFS

B.1 PROOF OF PROPOSITION 2

Proof. Because h is strictly increasing,

$$y_i < y_j \iff h(y_i) < h(y_j), \qquad y_i = y_j \iff h(y_i) = h(y_j).$$

Hence $\mathcal{D}$ and $\mathcal{D}^h$ induce exactly the same pairwise labels, list permutations, and tie decisions.

Any ordinal loss that accesses the targets only through these quantities therefore has exactly the same empirical objective on the two data sets. In particular, this applies to the Pairwise and ListMLE losses used in this work. If the optimization and selection randomness are coupled across the two runs, the fitted scores and the selected candidates are identical. □

B.2 PROOF OF PROPOSITION 3

Proof. Let

$$\Delta_{t,g} = f_{t,(g+1)} - f_{t,(g)}, \qquad g = 1, \dots, N-1.$$

For every realized selected batch,

$$\begin{aligned} r_t^{\mathrm{sel}} &= f_{t,(K_t)} - f_{t,(1)} \\ &= \sum_{g=1}^{K_t-1} \Delta_{t,g} \\ &= \sum_{g=1}^{N-1} \Delta_{t,g} \mathbf{1}\{K_t > g\}. \end{aligned}$$

Taking conditional expectation with respect to the randomness of the selection rule gives

$$\mathbb{E}[r_t^{\mathrm{sel}} \mid \mathcal{D}_t, \mathcal{C}_t] = \sum_{g=1}^{N-1} \Delta_{t,g} \mathbb{P}(K_t > g \mid \mathcal{D}_t, \mathcal{C}_t),$$

which is Eq. (4). □

B.3 PROOF OF PROPOSITION 6

Proof. Let

$$U = w_\star^\top X \sim \mathcal{N}(0,1), \qquad S = U + \varepsilon,$$

and write $Y = h(S)$. Since X is isotropic Gaussian and Y depends on X only through U, the population affine regression slope is parallel to $w_\star$. Its magnitude is

$$a_h = \mathrm{Cov}(U, Y) = \mathbb{E}[U(Y - \mathbb{E}Y)].$$

Gaussian integration by parts gives

$$a_h = \mathbb{E}[h'(S)].$$

Define the population residual

$$r = h(S) - \mathbb{E}h(S) - a_h U.$$

By construction, $\mathbb{E}[Ur] = 0$, and therefore

$$\mathbb{E}r^2 = \mathrm{Var}[h(S)] - a_h^2.$$

Hence

$$\frac{\mathbb{E}r^2}{a_h^2} = \frac{\mathrm{Var}[h(S)]}{(\mathbb{E}h'(S))^2} - 1 = \mathcal{D}_{\mathrm{val}}(h).$$

Since $S \sim \mathcal{N}(0, 1+\sigma^2)$, Stein's identity yields

$$\mathrm{Cov}(S, h(S)) = (1+\sigma^2)\mathbb{E}[h'(S)].$$

Thus

$$\operatorname{Corr}(S, h(S))^2 = \frac{(1+\sigma^2)(\mathbb{E}h'(S))^2}{\operatorname{Var}[h(S)]},$$

and consequently

$$\mathcal{D}_{\mathrm{val}}(h) = \frac{1+\sigma^2}{\operatorname{Corr}(S, h(S))^2} - 1.$$

Because $|\operatorname{Corr}(S, h(S))| \le 1$,

$$\mathcal{D}_{\mathrm{val}}(h) \ge \sigma^2.$$

Equality holds if and only if $|\operatorname{Corr}(S, h(S))| = 1$, which is equivalent to $h(S)$ being affine in S almost surely. Since S has full support on $\mathbb{R}$ and h is continuous and strictly increasing, this is equivalent to h being affine. □

B.4 Rank-Score Asymptotics

Work under the Gaussian single-index model

$$X_i \sim \mathcal{N}(0, I_d), \qquad U_i = w_\star^\top X_i, \qquad S_i = U_i + \varepsilon_i,$$

where $\|w_\star\| = 1$ and the noises $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$ are independent of the inputs and of one another. Let R_i be the increasing rank of S_i among $S_1, \dots, S_n$. These ranks are almost surely unique and coincide with the ranks of $h(S_i)$ for every strictly increasing h.

For deterministic weights $b_n(1), \dots, b_n(n)$, define

$$\widehat{q}_b = \frac{1}{n} \sum_{i=1}^{n} b_n(R_i) X_i.$$

To specify the required convergence of the weights, introduce the step interpolation

$$B_n(u) = b_n(r), \qquad \frac{r-1}{n} < u \le \frac{r}{n}, \qquad 0 < u < 1.$$

Assume that there exists $b \in L^2(0,1)$ such that

$$\|B_n - b\|_{L^2(0,1)} \longrightarrow 0. \tag{16}$$

Define

$$\nu_b = \int_0^1 b(u)^2 \, du, \qquad \kappa_b = \int_0^1 \Phi^{-1}(u) b(u) \, du.$$

The latter integral is finite by Cauchy–Schwarz, since $\int_0^1 \Phi^{-1}(u)^2 \, du = 1$.

Lemma 11 (High-Dimensional Rank-Score Alignment)**.** *Suppose that Eq.* (16) *holds, $\kappa_b > 0$, and σ is fixed. If $d, n \to \infty$ with $d/n \to \gamma \in (0, \infty)$, then*

$$\frac{\widehat{q}_b^\top w_\star}{\|\widehat{q}_b\|} \xrightarrow{p} \left(1 + (1+\sigma^2)\gamma \frac{\nu_b}{\kappa_b^2}\right)^{-1/2}.$$

The normalized vector may be assigned an arbitrary unit value on the event $\widehat{q}_b = 0$, whose probability tends to zero.

Proof. Set $\tau = \sqrt{1+\sigma^2}$. The Gaussian decomposition

$$X_i = U_i w_\star + Z_i$$

has

$$Z_i \sim \mathcal{N}(0, I_d - w_\star w_\star^\top),$$

with the Z_i's independent of the complete collection $(U_i, S_i)_{i=1}^n$. Moreover,

$$\nu_{b,n} := \frac{1}{n} \sum_{r=1}^{n} b_n(r)^2 = \|B_n\|_{L^2(0,1)}^2 \longrightarrow \nu_b. \tag{17}$$

First consider the weighted ordered scores. Let F_n^{-1} be the empirical quantile function of $S_1, \dots, S_n$, and write

$$q(u) = \tau \Phi^{-1}(u).$$

Under an i.i.d. coupling of the scalar samples, $F_n^{-1}(u) \to q(u)$ almost surely for every $u \in (0,1)$. In fact,

$$\|F_n^{-1} - q\|_{L^2(0,1)} \longrightarrow 0 \qquad \text{almost surely.} \tag{18}$$

To verify this stronger convergence, let $T_M(t) = \max\{-M, \min\{t, M\}\}$. For each fixed M, dominated convergence gives

$$\|T_M \circ F_n^{-1} - T_M \circ q\|_{L^2(0,1)} \longrightarrow 0.$$

The remaining tails satisfy

$$\begin{aligned} \|F_n^{-1} - T_M \circ F_n^{-1}\|_{L^2(0,1)}^2 &= \frac{1}{n}\sum_{i=1}^{n} (|S_i| - M)_+^2 \\ &\longrightarrow \mathbb{E}[(|S| - M)_+^2] \\ &= \|q - T_M \circ q\|_{L^2(0,1)}^2 \end{aligned}$$

by the strong law. Since $\mathbb{E}S^2 < \infty$, the last quantity tends to zero as $M \to \infty$. The triangle inequality proves Eq. (18).

Consequently,

$$\frac{1}{n}\sum_{r=1}^{n} b_n(r) S_{(r)} = \int_0^1 B_n(u) F_n^{-1}(u)\, du \xrightarrow{p} \int_0^1 b(u) q(u)\, du = \tau \kappa_b.$$

Indeed, the difference between the two integrals is bounded by

$$\|B_n - b\|_2 \|F_n^{-1}\|_2 + \|b\|_2 \|F_n^{-1} - q\|_2,$$

which tends to zero.

For the component parallel to $w_\star$, joint Gaussianity gives

$$U_i = \frac{S_i}{\tau^2} + \eta_i, \qquad \eta_i \sim \mathcal{N}\left(0, \frac{\sigma^2}{\tau^2}\right),$$

where the η_i's are independent of one another and of $(S_1, \dots, S_n)$. Conditional on the scores, the weighted residual average has mean zero and variance

$$\operatorname{Var}\left(\frac{1}{n}\sum_{i=1}^{n} b_n(R_i)\eta_i \,\middle|\, S_1, \dots, S_n \right) = \frac{\sigma^2 \nu_{b,n}}{\tau^2 n} \longrightarrow 0.$$

It follows that

$$w_\star^\top \widehat{q}_b = \frac{1}{\tau^2 n}\sum_{r=1}^{n} b_n(r) S_{(r)} + \frac{1}{n}\sum_{i=1}^{n} b_n(R_i)\eta_i \xrightarrow{p} \frac{\kappa_b}{\tau}. \tag{19}$$

For the orthogonal component, let $P_\perp = I_d - w_\star w_\star^\top$. Conditional on the scores,

$$P_\perp \widehat{q}_b = \frac{1}{n}\sum_{i=1}^{n} b_n(R_i) Z_i \sim \mathcal{N}\left(0, \frac{\nu_{b,n}}{n} P_\perp\right).$$

Thus its squared norm has conditional distribution

$$\|P_\perp \widehat{q}_b\|^2 \overset{\mathrm{d}}{=} \frac{\nu_{b,n}}{n} \chi^2_{d-1}.$$

Since $\chi^2_{d-1}/(d-1) \xrightarrow{p} 1$, Eq. (17) yields

$$\|P_\perp \widehat{q}_b\|^2 \xrightarrow{p} \gamma \nu_b. \tag{20}$$

Combining Eqs. (19) and (20) gives

$$\frac{\widehat{q}_b^\top w_\star}{\|\widehat{q}_b\|} \xrightarrow{p} \frac{\kappa_b/\tau}{\sqrt{\kappa_b^2/\tau^2 + \gamma \nu_b}} = \left(1 + \tau^2 \gamma \frac{\nu_b}{\kappa_b^2}\right)^{-1/2}.$$

Because $\kappa_b > 0$, the parallel component is bounded away from zero with probability tending to one, which also justifies the normalization. □

Pairwise score. Let $\pi(r)$ denote the observation with increasing rank r. For the Pairwise loss in Eq. (2), all $\binom{n}{2}$ pairs are almost surely non-tied. Differentiating at the zero score gives

$$\left.\frac{\partial \mathcal{L}_{\mathrm{Pair}}}{\partial s_{\pi(r)}}\right|_{s=0} = \frac{(r-1)-(n-r)}{n(n-1)} = \frac{2r-n-1}{n(n-1)}.$$

Therefore, for the affine score $s_\theta(x) = \theta^\top x + c$, the slope gradient satisfies

$$\nabla_\theta \mathcal{L}_{\mathrm{Pair}}|_{(\theta,c)=0} = \frac{n}{n-1}\widehat{q}_{\mathrm{Pair}}, \qquad b_{\mathrm{Pair,n}}(r) = \frac{2r-n-1}{n}.$$

Under the larger-is-better score convention, gradient descent moves in a utility direction, so the positive slope gradient is the corresponding cost direction.

The limiting score is $b_{\mathrm{Pair}}(u) = 2u - 1$. Its step interpolation satisfies

$$|B_{\mathrm{Pair,n}}(u) - (2u-1)| \le \frac{1}{n}, \qquad 0 < u < 1,$$

so Eq. (16) holds. Moreover,

$$\nu_{\mathrm{Pair}} = \int_0^1 (2u-1)^2\, du = \frac{1}{3}.$$

For $Z \sim \mathcal{N}(0,1)$, Gaussian integration by parts gives

$$\begin{aligned}\kappa_{\mathrm{Pair}} &= \mathbb{E}[Z(2\Phi(Z)-1)] \\ &= 2\mathbb{E}[\varphi(Z)] = 2\int_{\mathbb{R}} \varphi(z)^2\, dz = \frac{1}{\sqrt{\pi}}.\end{aligned}$$

Hence

$$C_{\mathrm{Pair}} = \frac{\nu_{\mathrm{Pair}}}{\kappa_{\mathrm{Pair}}^2} = \frac{\pi}{3}.$$

ListMLE score. Define $H_0 = 0$ and $H_j = \sum_{\ell=1}^{j} \ell^{-1}$ for $j \ge 1$. The final ListMLE factor is identically one, so appending its zero log-loss allows Eq. (1) to be written as a sum over all n positions. At the zero score, differentiation therefore gives, for every $r = 1, \dots, n$,

$$\begin{aligned}\left.\frac{\partial \mathcal{L}_{\mathrm{ListMLE}}}{\partial s_{\pi(r)}}\right|_{s=0} &= -1 + \sum_{j=1}^{r} \frac{1}{n-j+1} \\ &= H_n - H_{n-r} - 1 = b_{\mathrm{List,n}}(r).\end{aligned}$$

Thus

$$\nabla_\theta \mathcal{L}_{\mathrm{ListMLE}}|_{(\theta,c)=0} = n\widehat{q}_{\mathrm{List}},$$

so $\widehat{q}_{\mathrm{List}}$ has the required cost orientation.

For each fixed $u \in (0,1)$, with $r = \lceil nu \rceil$, the harmonic-number asymptotics yield

$$B_{\mathrm{List,n}}(u) = H_n - H_{n-r} - 1 \longrightarrow b_{\mathrm{List}}(u) = -1 - \log(1-u).$$

The logarithmic endpoint does not prevent L^2 convergence. Indeed, for $(r-1)/n < u \le r/n$,

$$0 \le H_n - H_{n-r} \le 1 + \log\left(\frac{n}{n-r+1}\right) \le 1 - \log(1-u).$$

Consequently,

$$|B_{\mathrm{List,n}}(u)| \le 1 + |\log(1-u)|.$$

The square of this bound is integrable on $(0,1)$, so dominated convergence proves

$$\|B_{\mathrm{List,n}} - b_{\mathrm{List}}\|_{L^2(0,1)} \longrightarrow 0.$$

For $V \sim \mathrm{Unif}(0,1)$, $-\log(1-V)$ has the unit-rate exponential distribution. Hence

$$\mathbb{E}[b_{\mathrm{List}}(V)] = 0, \qquad \nu_{\mathrm{List}} = \mathbb{E}[b_{\mathrm{List}}(V)^2] = 1.$$

To compute κ_{List}, set

$$g(z) = b_{\mathrm{List}}(\Phi(z)) = -1 - \log(1 - \Phi(z)).$$

Then $g'(z) = \varphi(z)/(1 - \Phi(z))$. Integration by parts gives

$$\begin{aligned}\kappa_{\mathrm{List}} &= \int_{\mathbb{R}} z\varphi(z)g(z)\,dz \\ &= \int_{\mathbb{R}} \varphi(z)g'(z)\,dz = \int_{\mathbb{R}} \frac{\varphi(z)^2}{1-\Phi(z)}\,dz \approx 0.903197.\end{aligned}$$

The boundary term vanishes because $g(z) = O(1 + z^2)$ as $|z| \to \infty$. The first integral is absolutely convergent by Cauchy–Schwarz; applying integration by parts on bounded intervals and passing to the limit also establishes finiteness of the last integral. Therefore,

$$C_{\mathrm{List}} = \frac{\nu_{\mathrm{List}}}{\kappa_{\mathrm{List}}^2} = \kappa_{\mathrm{List}}^{-2} \approx 1.22584.$$

B.5 Proof of Theorem 7

Proof. By rotational invariance, take $w_\star = e_1$ and write

$$X_i = (U_i, Z_i), \qquad Z_i \sim \mathcal{N}(0, I_{d-1}),$$

where the Z_i's are independent of the complete collection $(U_i, \varepsilon_i)_{i=1}^n$. Set

$$\mu_h = \mathbb{E}[h(S)], \qquad a_h = \mathbb{E}[h'(S)] > 0, \qquad r_i = h(S_i) - \mu_h - a_h U_i.$$

Proposition 6 gives

$$\mathbb{E}r_i = 0, \qquad \mathbb{E}[U_i r_i] = 0, \qquad v_r := \mathbb{E}r_i^2 = a_h^2 \mathcal{D}_{\mathrm{val}}(h) < \infty.$$

The residual r_i need not be independent of U_i.

Write

$$u = (U_1, \ldots, U_n)^\top, \qquad r = (r_1, \ldots, r_n)^\top,$$

and let $Z \in \mathbb{R}^{n\times p}$, with $p = d - 1$, have rows $Z_i^\top$. Let $\mathbf{1} \in \mathbb{R}^n$ be the all-ones vector and define

$$J = I_n - \frac{\mathbf{1}\mathbf{1}^\top}{n}, \qquad u_c = Ju, \qquad r_c = Jr.$$

Since $\mathbb{E}U_i^2 = 1$, $\mathbb{E}r_i^2 < \infty$, and $\mathbb{E}|U_i r_i| < \infty$, the law of large numbers yields

$$\frac{\|u_c\|^2}{n} \xrightarrow{p} 1, \qquad \frac{u_c^\top r_c}{n} \xrightarrow{p} 0, \qquad \frac{\|r_c\|^2}{n} \xrightarrow{p} v_r. \tag{21}$$

Consider affine least squares with an intercept, and decompose its fitted slope as $\widehat{\beta} = (\widehat{\beta}_\parallel, \widehat{\beta}_\perp)$. All design matrices below have full column rank almost surely for sufficiently large n, since $d/n \to \gamma < 1$.

Parallel coefficient. Let M_n be the orthogonal projector onto the complement of the column space of $[\mathbf{1}, Z]$. Eliminating the intercept and the Z-coefficients from the normal equations gives

$$\widehat{\beta}_\parallel - a_h = \frac{u^\top M_n r}{u^\top M_n u} = \frac{u_c^\top M_n r_c}{u_c^\top M_n u_c}. \tag{22}$$

The matrix M_n is independent of (u, r). Restricted to $\mathbf{1}^\perp$, it is a uniformly distributed orthogonal projector of rank $n - d$ in dimension $n - 1$. Therefore, for every fixed nonzero $v \in \mathbf{1}^\perp$,

$$\frac{v^\top M_n v}{\|v\|^2} \overset{\mathrm{d}}{=} \frac{\sum_{j=1}^{n-d} g_j^2}{\sum_{j=1}^{n-1} g_j^2} \sim \mathrm{Beta}\left(\frac{n-d}{2}, \frac{d-1}{2}\right),$$

where the g_j's are independent standard Gaussian variables. In particular, writing $\alpha_n = (n-d)/(n-1)$,

$$\mathbb{E}\left[\frac{v^\top M_n v}{\|v\|^2}\right] = \alpha_n, \qquad \mathrm{Var}\left(\frac{v^\top M_n v}{\|v\|^2}\right) = \frac{2(n-d)(d-1)}{(n-1)^2(n+1)} = O(n^{-1}).$$

These bounds do not depend on the direction of v.

Conditioning on (u, r), the same bounds apply to vectors measurable with respect to (u, r). Thus, whenever such a vector $v \in \mathbf{1}^\perp$ satisfies $\|v\|^2/n = O_p(1)$, conditional Chebyshev gives

$$\frac{v^\top M_n v - \alpha_n \|v\|^2}{n} \xrightarrow{p} 0.$$

This conclusion follows by first restricting to $\|v\|^2/n \le K$ and then letting $K \to \infty$, so it does not require an unconditional fourth-moment bound.

Apply this result to u_c and to $u_c + r_c$ and $u_c - r_c$. Using polarization and Eq. (21), we obtain

$$\frac{u_c^\top M_n u_c}{n} = \alpha_n \frac{\|u_c\|^2}{n} + o_p(1) \xrightarrow{p} 1 - \gamma,$$
$$\frac{u_c^\top M_n r_c}{n} = \alpha_n \frac{u_c^\top r_c}{n} + o_p(1) \xrightarrow{p} 0.$$

Since $1 - \gamma > 0$, Eq. (22) proves

$$\widehat{\beta}_\parallel \xrightarrow{p} a_h. \tag{23}$$

Orthogonal coefficients. Let $m = n - 2$, and choose an orthonormal basis $Q \in \mathbb{R}^{n \times m}$ for $\mathrm{span}\{\mathbf{1}, u\}^\perp$, measurable with respect to u. Define

$$G = Q^\top Z, \qquad e = Q^\top r.$$

Conditional on (u, r), the matrix G has independent standard Gaussian entries and e is fixed. Eliminating the intercept and the u-coefficient gives the exact identity

$$\widehat{\beta}_\perp = (G^\top G)^{-1} G^\top e. \tag{24}$$

Moreover,

$$\|e\|^2 = \|r_c\|^2 - \frac{(u_c^\top r_c)^2}{\|u_c\|^2},$$

so Eq. (21) implies

$$\frac{\|e\|^2}{m} \xrightarrow{p} v_r. \tag{25}$$

We establish the distributional identity needed to control Eq. (24). For an $m \times p$ standard Gaussian matrix with $p < m$, use its polar factorization

$$G = OW^{1/2}, \qquad O^\top O = I_p, \qquad W = G^\top G.$$

The matrix O is uniformly distributed on the Stiefel manifold of orthonormal p-frames in $\mathbb{R}^m$ and is independent of W (Tropp, 2012, Proposition 9). These properties also follow from left orthogonal invariance: conditional on W, the distribution of O is invariant under every left orthogonal transformation, and hence is the same uniform distribution for every W.

For fixed $e \neq 0$, the vector $O^\top e/\|e\|$ has the distribution of the first p coordinates of a uniform unit vector in $\mathbb{R}^m$. Consequently, it can be written as

$$O^\top e = \|e\| \sqrt{B_{m,p}}\, \omega, \qquad B_{m,p} \sim \mathrm{Beta}\left(\frac{p}{2}, \frac{m-p}{2}\right),$$

where ω is uniform on the unit sphere in $\mathbb{R}^p$. The variables $B_{m,p}$ and ω are independent, and both are independent of W. It follows that

$$\|(G^\top G)^{-1} G^\top e\|^2 = \|e\|^2 B_{m,p}\, \omega^\top W^{-1} \omega.$$

By orthogonal invariance of W, conditional on ω, the quadratic form $\omega^\top W^{-1} \omega$ has the same law as $(W^{-1})_{11}$. To identify this law, represent W as

$$W = (g, H)^\top (g, H),$$

where $g \in \mathbb{R}^m$ and $H \in \mathbb{R}^{m\times(p-1)}$ are independent standard Gaussian arrays. Block inversion gives

$$(W^{-1})_{11} = \left[g^\top \left\{I_m - H(H^\top H)^{-1}H^\top\right\} g\right]^{-1}.$$

Conditional on H, the projector in braces has rank $m-p+1$. Its quadratic form in g therefore has distribution χ^2_{m-p+1}. Hence

$$\omega^\top W^{-1}\omega \overset{\mathrm{d}}{=} D_{m,p}^{-1}, \qquad D_{m,p} \sim \chi^2_{m-p+1}.$$

Because $B_{m,p}$ is independent of (ω, W), $D_{m,p}$ can be taken independent of $B_{m,p}$.

Applying this argument conditionally on (u, r) proves

$$\|\widehat{\beta}_\perp\|^2 \mid (u,r) \overset{\mathrm{d}}{=} \|e\|^2 \frac{B_{m,p}}{D_{m,p}}, \tag{26}$$

where the auxiliary variables $B_{m,p}$ and $D_{m,p}$ may also be taken independent of (u, r). The identity holds trivially when $e = 0$.

Since $p/m \to \gamma \in (0,1)$, the beta and chi-square representations give

$$B_{m,p} \xrightarrow{p} \gamma, \qquad \frac{D_{m,p}}{m} \xrightarrow{p} 1-\gamma.$$

Combining Eqs. (25) and (26) yields

$$\|\widehat{\beta}_\perp\|^2 \xrightarrow{p} v_r \frac{\gamma}{1-\gamma} = a_h^2 \mathcal{D}_{\mathrm{val}}(h) \frac{\gamma}{1-\gamma}. \tag{27}$$

Directional limits and crossover. By Eqs. (23) and (27),

$$\begin{aligned}
\widehat{v}_{\mathrm{val}}^\top w_\star &= \frac{\widehat{\beta}_\parallel}{\sqrt{\widehat{\beta}_\parallel^2 + \|\widehat{\beta}_\perp\|^2}} \\
&\xrightarrow{p} \frac{a_h}{\sqrt{a_h^2 + a_h^2 \mathcal{D}_{\mathrm{val}}(h)\gamma/(1-\gamma)}} \\
&= \left(1 + \mathcal{D}_{\mathrm{val}}(h)\frac{\gamma}{1-\gamma}\right)^{-1/2}.
\end{aligned}$$

Since $a_h > 0$, the fitted slope is nonzero with probability tending to one; its normalized direction may be assigned an arbitrary unit value on the zero event.

Appendix B.4 verifies the L^2 weight condition and the cost orientation for both ranking losses. Applying Lemma 11 therefore gives

$$\widehat{v}_{\mathrm{ord},k}^\top w_\star \xrightarrow{p} \left(1 + C_k(1+\sigma^2)\gamma\right)^{-1/2}, \qquad k \in \{\mathrm{Pair}, \mathrm{List}\}.$$

Finally, $t \mapsto (1+t)^{-1/2}$ is strictly decreasing. Thus direction k has strictly larger limiting alignment than affine least squares if and only if

$$C_k(1+\sigma^2)\gamma < \mathcal{D}_{\mathrm{val}}(h)\frac{\gamma}{1-\gamma}.$$

Since $\gamma > 0$ and $1-\gamma > 0$, this is equivalent to

$$\mathcal{D}_{\mathrm{val}}(h) > C_k(1+\sigma^2)(1-\gamma),$$

which proves the theorem. □

B.6 Proof of Proposition 8

Proof. Write $U_i = w_\star^\top X_i$ and $X_i = U_i w_\star + Z_i$, where the centered Gaussian vector Z_i has covariance $P_\perp = I_d - w_\star w_\star^\top$ and is independent of (U_i, Y_i). Set $a_h = \mathbb{E}[h'(S)] > 0$ and $\mu_h = \mathbb{E}Y$. The identity $\mathbb{E}[UY] = a_h$ follows from the Gaussian integration-by-parts calculation in Appendix B.3.

Since $\mathbb{E}Y^2 < \infty$ and $\mathbb{E}U^2 = 1$, Cauchy–Schwarz gives $\mathbb{E}|UY| < \infty$. The law of large numbers therefore yields

$$A_n := w_\star^\top \widehat{q}_{\mathrm{val},1} = \frac{1}{n}\sum_i U_i Y_i - \bar{U}\bar{Y} \xrightarrow{p} a_h. \tag{28}$$

For the orthogonal component, $B_n := P_\perp \widehat{q}_{\mathrm{val},1} = n^{-1}\sum_i Z_i(Y_i - \bar{Y})$. Conditional on the complete collection $(U_i, Y_i)_{i=1}^n$, the vectors Z_i remain independent Gaussian vectors. Consequently,

$$B_n \mid (U_i, Y_i)_{i=1}^n \sim \mathcal{N}\left(0, \frac{V_n}{n} P_\perp\right), \qquad V_n = \frac{1}{n}\sum_i (Y_i - \bar{Y})^2.$$

The scalar sample variance satisfies $V_n \xrightarrow{p} \mathrm{Var}(Y)$. In an orthonormal basis of $w_\star^\perp$, the conditional distribution of $\|B_n\|^2$ is $(V_n/n)\chi^2_{d-1}$. Since $\chi^2_{d-1}/(d-1) \xrightarrow{p} 1$ and $(d-1)/n \to \gamma$,

$$\|B_n\|^2 \xrightarrow{p} \gamma\,\mathrm{Var}(Y). \tag{29}$$

Combining the orthogonal and parallel limits gives

$$\frac{\widehat{q}_{\mathrm{val},1}^\top w_\star}{\|\widehat{q}_{\mathrm{val},1}\|} = \frac{A_n}{\sqrt{A_n^2 + \|B_n\|^2}} \xrightarrow{p} \frac{a_h}{\sqrt{a_h^2 + \gamma\,\mathrm{Var}(Y)}} = [1 + \gamma\{1 + \mathcal{D}_{\mathrm{val}}(h)\}]^{-1/2}.$$

Because $a_h > 0$, the probability that the sample direction is nonzero tends to one; it can be assigned an arbitrary unit value on its zero event.

The map $t \mapsto (1+t)^{-1/2}$ is strictly decreasing. Comparing with Eq. (8) and cancelling $\gamma > 0$ gives Eq. (12). Comparing instead with Eq. (7) gives

$$\frac{\gamma\mathcal{D}_{\mathrm{val}}(h)}{1-\gamma} < \gamma\{1 + \mathcal{D}_{\mathrm{val}}(h)\} \quad \Longleftrightarrow \quad \mathcal{D}_{\mathrm{val}}(h) < \frac{1-\gamma}{\gamma}.$$

Finally, Lemma 9, applied conditionally on the training data, gives the stated one-step selection-regret limit. □

B.7 Proof of Lemma 9

Proof. Let

$$A_j = v^\top X_j^{\mathrm{c}}, \qquad B_j = w_\star^\top X_j^{\mathrm{c}}.$$

Then (A_j, B_j) is standard bivariate Gaussian with correlation $\rho = v^\top w_\star$, so we may write

$$B_j = \rho A_j + \sqrt{1-\rho^2}\, E_j,$$

where $E_j \sim \mathcal{N}(0,1)$ are independent of the A_j's. Since $\widehat{j}(v) = \arg\min_j A_j$ depends only on the A_j's,

$$\mathbb{E}[B_{\widehat{j}(v)}] = \rho\,\mathbb{E}[\min_j A_j] = -\rho\kappa_N.$$

Also,

$$\mathbb{E}[\min_j B_j] = -\kappa_N.$$

Therefore

$$R_{\mathrm{lat}}(v) = -\rho\kappa_N + \kappa_N = \kappa_N(1-\rho).$$

□

B.8 Proof of Corollary 10

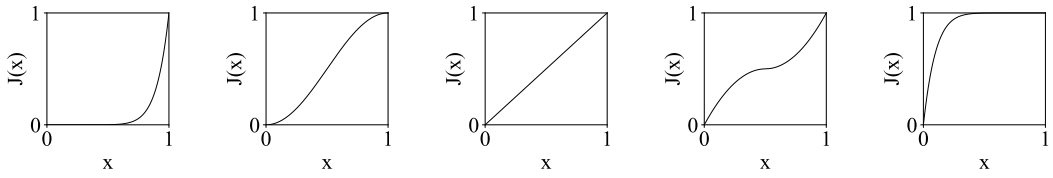
Proof. By Lemma 9, $R_{\mathrm{lat}}(v) = \kappa_N(1 - v^\top w_\star)$, so latent selection regret is strictly decreasing in directional alignment. Therefore, for $k \in \{\mathrm{Pair}, \mathrm{List}\}$, the order-induced cost direction has smaller asymptotic latent selection regret exactly when its limiting alignment in Theorem 7 exceeds that of the value-regression direction. This occurs exactly when

$$\mathcal{D}_{\mathrm{val}}(h) > C_k(1+\sigma^2)(1-\gamma),$$

equivalently,

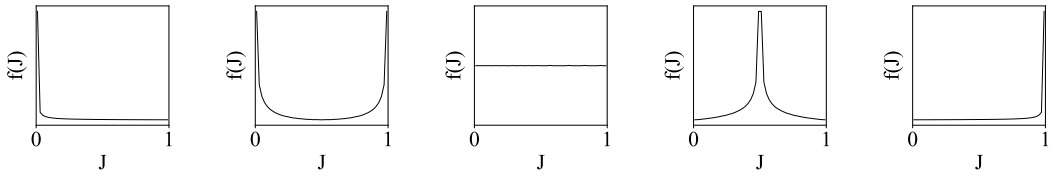
$$\gamma > 1 - \frac{\mathcal{D}_{\mathrm{val}}(h)}{C_k(1+\sigma^2)}.$$

□



(a) Flat (b) U-Shaped (c) Neutral (d) Bell (e) Steep

Figure 5: Normalized OPCs.



(a) Flat (b) U-Shaped (c) Neutral (d) Bell (e) Steep

Figure 6: Normalized performance density functions.

C Representative Ordered Performance Curves

The ordered performance curve (OPC) in Definition 4 records how normalized objective gaps are distributed along candidate rank. Following Ho et al. (2007), Figure 5 shows five representative shapes. They differ only in where the large rank-to-rank increments occur: near the worst end for Flat, mainly in the middle for U-Shaped, approximately uniformly for Neutral, near both ends for Bell, and near the best end for Steep. Hence two candidate sets with the same ranking errors can incur very different selection regret.

A continuous interpretation makes this relation explicit. Let F be the distribution function of the normalized objective value and let

$$q(r) = F^{-1}(r)$$

be its quantile function. The OPC is the finite-sample analogue of q. Whenever the corresponding density p is positive,

$$q'(r) = \frac{1}{p(q(r))}.$$

Thus low-density regions produce steep OPC segments and large objective gaps between nearby ranks. The corresponding representative densities are shown in Figure 6.

For a finite candidate set, the increment $v_{t,g+1} - v_{t,g}$ is precisely the normalized cost of crossing the g-th rank boundary, while $\Psi_t(g)$ is the probability of crossing it. Therefore all of the representative OPC shapes enter the same decomposition,

$$\frac{\mathbb{E}[r_t^{\mathrm{sel}} \mid \mathcal{D}_t, \mathcal{C}_t]}{f_{t,(N)} - f_{t,(1)}} = \sum_{g=1}^{N-1} (v_{t,g+1} - v_{t,g}) \Psi_t(g),$$

as established in Proposition 3.

D More Experiment Details

D.1 Definition of Synthetic Functions

The synthetic suite contains four standard minimization functions at $d \in \{200, 400, 600, 800, 1000\}$. Candidate-generation methods operate on $[0, 1]^d$; inputs are mapped coordinatewise to the native domain before the objective is evaluated. The functions are as follows.

Ackley. For $x \in [-32.768, 32.768]^d$,

$$f(x) = -20 \exp\left(-0.2\sqrt{\frac{1}{d}\sum_{i=1}^{d} x_i^2}\right) - \exp\left(\frac{1}{d}\sum_{i=1}^{d}\cos(2\pi x_i)\right) + 20 + e.$$

Levy. For $x \in [-10, 10]^d$, let $w_i = 1 + (x_i - 1)/4$. Then

$$f(x) = \sin^2(\pi w_1) + \sum_{i=1}^{d-1}(w_i - 1)^2\{1 + 10\sin^2(\pi w_i + 1)\} + (w_d - 1)^2\{1 + \sin^2(2\pi w_d)\}.$$

Rosenbrock. For $x \in [-5, 10]^d$,

$$f(x) = \sum_{i=1}^{d-1}\left[100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2\right].$$

Dixon–Price. For $x \in [-10, 10]^d$,

$$f(x) = (x_1 - 1)^2 + \sum_{i=2}^{d} i\left(2x_i^2 - x_{i-1}\right)^2.$$

D.2 Real-World Benchmarks

The four real-world tasks retain the domains and scalar objectives of their reference implementations. Rover is a 60-dimensional trajectory-design problem; MOPTA08 is a 124-dimensional engineering-design problem; Walker uses the continuous-control parameterization of the Gymnasium MuJoCo task; and DNA is the 180-dimensional regression instance from LassoBench. Reward-type outputs are negated so that every reported task is a minimization problem. No task values or bounds are modified separately for a surrogate variant.

D.3 Implementation

Table 1 collects the settings shared across the end-to-end experiments. Within an outer BO framework and seed, all surrogate variants receive the same initial design, embedding or trust-region state, candidate draws, and evaluation budget. The objective evaluations are deterministic, and the total budget includes the initial design.

Setting	Value
Instances	20 synthetic and 4 real-world
Independent trials	10 per configuration
Initial design	40 shared Latin-hypercube points
Total evaluation budget	500, including the initial design
Surrogates	GP, MSE-NN, Rank-NN, Pairwise, ListMLE
Outer frameworks	Standard BO, TuRBO, HeSBO-12, HeSBO-24
Neural ensemble size	3
Standard BO/TuRBO candidate pool	$\min\{100d, 5000\}$
HeSBO candidate pool	2000

Table 1: Shared protocol for the optimization benchmarks. Framework-specific training and selection rules are given in Appendix D.3.

Algorithm 1 describes the Pairwise and ListMLE variants of OPBO. Within each outer framework, all surrogate variants use the same initial design, candidate generator, evaluation budget, and region updates. The learning objectives are defined in Section 2.3, the acquisition rules in Section 2.4, and the benchmark protocol in Section 4.1.

All tasks are stated as minimization problems. We use the utility $u_i = -y_i$, so that a larger utility or predicted score always denotes a better candidate.

Neural surrogates. MSE-NN, Rank-NN, Pairwise, and ListMLE use the same scalar-output ReLU network with two 128-unit hidden layers. Each ensemble contains $M = 3$ independently initialized networks. Every scheduled fit starts from a new Xavier initialization and runs for 50 full-batch epochs using Adam with learning rate 10^{-2}.

For numerical regression, utilities are centered at their empirical median and scaled by their empirical standard deviation, with unit scale used when the standard deviation is below 10^{-6}. MSE-NN regresses these standardized utilities. Rank-NN instead uses the normal-score targets $z_i = \Phi^{-1}((r_i - \frac{1}{2})/n)$, where r_i is the average empirical rank of u_i from worst to best. Pairwise uses all pairs whose utilities differ by more than 10^{-8} and minimizes Eq. (2). ListMLE applies Eq. (1) to the complete current data as one list, with ties randomized before sorting in each epoch. The list size grows with the data, consistent with the rank-score asymptotics in Appendix B.4.

Standard BO and TuRBO. Both frameworks use ensemble sampling as an approximate Thompson rule. At each iteration, one ensemble member is sampled uniformly and scores all candidates; the b candidates with the largest scores are selected. Exact score ties are broken randomly. The candidate pool contains $N_{\text{cand}} = \min\{100d, 5000\}$ scrambled Sobol points. Standard BO samples over the full normalized domain, while TuRBO samples inside its current trust region and perturbs each coordinate with probability $\min\{20/d, 1\}$.

For TuRBO, the trust-region length starts at 0.8, is capped at 1.6, and triggers a restart below 2^{-7}. Three consecutive successes double the length, while $\lceil \max\{4/b, d/b\} \rceil$ consecutive failures halve it. The neural ensemble is refitted at every Standard BO or TuRBO iteration.

HeSBO. Each run uses a fixed sparse embedding that assigns every original coordinate to a uniformly sampled embedded coordinate and random sign. The embedding is shared across surrogate variants within each seed. At every step, $N = 2000$ Latin-hypercube candidates are drawn in the embedded domain and scored jointly with the observed points. The neural variants use Eq. (3) with $\xi = 0$, selecting the b candidates with the largest acquisition values. Ties are resolved first by mean ensemble rank and then randomly.

The HeSBO experiments begin with 40 observations. Neural ensembles are fitted initially and after every 20 new evaluations. Between refits, the networks remain fixed, but newly evaluated inputs are included when observed and candidate points are ranked jointly.

GP references. The Standard BO and TuRBO GP baselines retain the posterior and acquisition rules of their reference implementations. The HeSBO baseline uses an exact constant-mean GP with a scaled Matérn-$5/2$ kernel and Gaussian likelihood. The likelihood noise is initialized at 10^{-3} and constrained to $[10^{-6}, 0.2]$. Hyperparameters are optimized for 50 Adam steps at learning rate 0.1 on the same 20-evaluation schedule. Between these steps, the GP conditions on all new observations without reoptimizing its hyperparameters. Expected improvement is evaluated on the same 2000 candidates as the neural variants.

The four neural variants provide a controlled comparison of learning objectives, while GP serves as the framework-specific reference. Runtime includes surrogate fitting, candidate generation, acquisition, and objective evaluation.

D.4 Controlled Neural Bridge Protocol

The controlled bridge retains Gaussian inputs without clipping or mapping them to $[0, 1]^d$. This isolates the change from affine estimators to neural surrogates while leaving the single-index data model unchanged. It is a candidate-selection experiment, not an additional adaptive BO benchmark. Training and candidate data are independent, and the clean latent direction and scores are used only for evaluation. No hyperparameter, epoch, or model checkpoint is selected using candidate regret.

Each member uses a d–128–128–1 network, ReLU activations in the hidden layers, Xavier-uniform weights with gain one, and zero biases. Members have different initializations; the initialization for a particular member is shared across losses and representations. Adam uses learning rate 0.01, $(\beta_1, \beta_2) = (0.9, 0.999)$, $\epsilon = 10^{-8}$, and no weight decay. There is no dropout, early stopping, learning-rate scheduling, or gradient clipping. All trainable layers receive the 50 full-batch updates. The utility

target is $u_i = -Y_i$. MSE-NN subtracts its empirical median and divides by the sample standard deviation (denominator $n - 1$), using unit scale below 10^{-6}. Rank-NN uses $\Phi^{-1}((r_i - 1/2)/n)$ with average utility ranks. Pairwise uses the mean logistic loss in Eq. (2); ListMLE uses the sum in Eq. (1), evaluated by a numerically stable reverse log-cumulative sum of exponentials.

All comparisons and rank targets are computed from the actual transformed observations in double precision before the neural tensors are cast to single precision. Pairwise includes every exactly non-tied pair, without an absolute objective-gap tolerance. This differs from the numerical filtering reported for the existing end-to-end Pairwise implementation; no end-to-end result is reinterpreted as having used the new comparison rule. Sorting is stable, and Gaussian observations are almost surely untied. The code aborts if finite-precision recoding changes an observed tie or rank. Candidate-score ties are resolved by shared random priorities that do not depend on values.

For each trial, the code checks that all three order-based training inputs are identical across recodings. It then trains each order-based surrogate once and reuses this identical optimization problem's fit; these reused results are not counted as independent observations. In the first trial of each dimension, both nonlinear recodings are additionally trained from scratch for all three order-based methods. The audit compares parameter hashes, candidate predictions, and selections against the affine recoding. The option to retrain all representations is also provided. Deterministic PyTorch execution is enabled, and unsupported deterministic operations cause an error rather than silently weakening the audit.

The reported bridge run uses CPU execution with single-precision neural tensors and double-precision data generation. Its environment and complete configuration are recorded with the raw results. Confidence intervals treat the training trial, rather than the ensemble member or candidate pool, as the independent unit. They are pointwise approximate intervals, not simultaneous claims over all dimension and representation settings.

D.5 Computing Environment

The end-to-end experiments were conducted on Linux servers equipped with NVIDIA L40 and RTX 6000 PRO GPUs. All methods compared within the same benchmark use the same software environment and hardware allocation. Reported runtimes are end-to-end wall-clock measurements and include surrogate fitting, candidate generation, acquisition, and objective evaluation.

E Additional Results

Figures 7–10 show all synthetic benchmark trajectories. The TuRBO real-world results are in Figure 12; Figures 11–14 give the remaining frameworks. All panels retain the five surrogate variants.

F Limitation

The theory isolates representation dependence in a Gaussian single-index model with affine scores. It explains why value regression and order learning can have different directional efficiency, but it does not provide a regret theorem for the full adaptive BO loop. The matched one-step control separates representation dependence from covariance correction but does not characterize nonlinear neural-network training. The high-dimensional selection limits assume an independent training sample and candidate pool; extending them to an adaptively collected history requires sequential concentration and explicit control of the candidate generator.

Order learning also discards cardinal information. This is advantageous when objective magnitudes are a misleading representation of the decision problem, but it can be inefficient when calibrated value differences are informative. The comparison noise must remain positively aligned with the true order; large low-margin regions, heteroscedastic reversals, or systematic preference errors can weaken the guarantee and the empirical advantage. The crossover described after Theorem 7 makes this tradeoff explicit: near a linear representation, MSE can have the smaller leading error.

Finally, the end-to-end benchmarks cover continuous, deterministic, high-dimensional tasks under a fixed budget of 500 evaluations. They do not establish the same ranking for discrete, constrained, multiobjective, or strongly noisy problems, nor for substantially larger budgets. These settings require separate study.

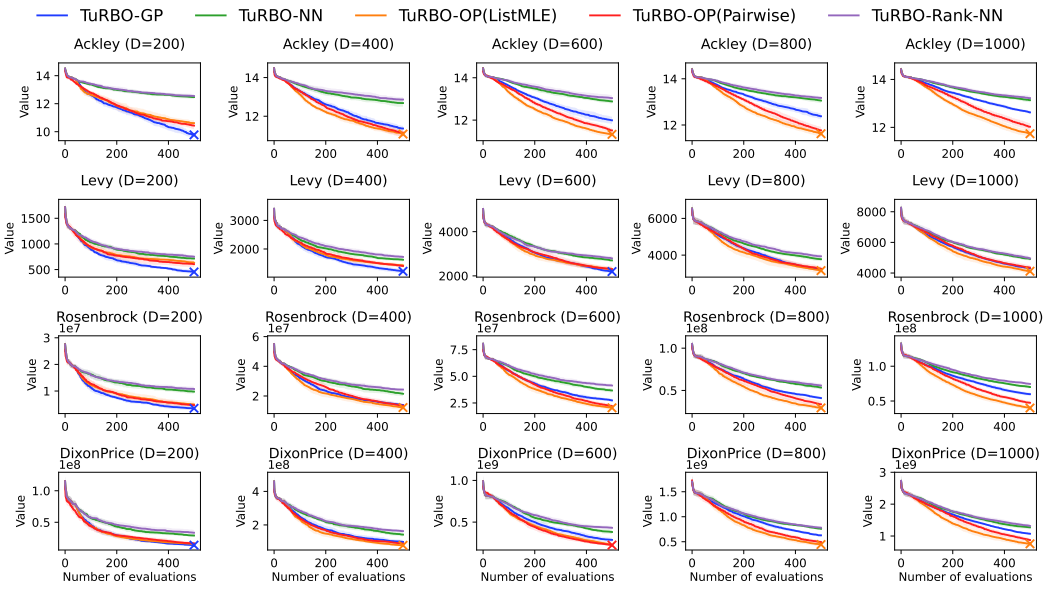


Figure 7: Synthetic benchmark trajectories under TuRBO. All five surrogates share the evaluation budget and Thompson-sampling acquisition.

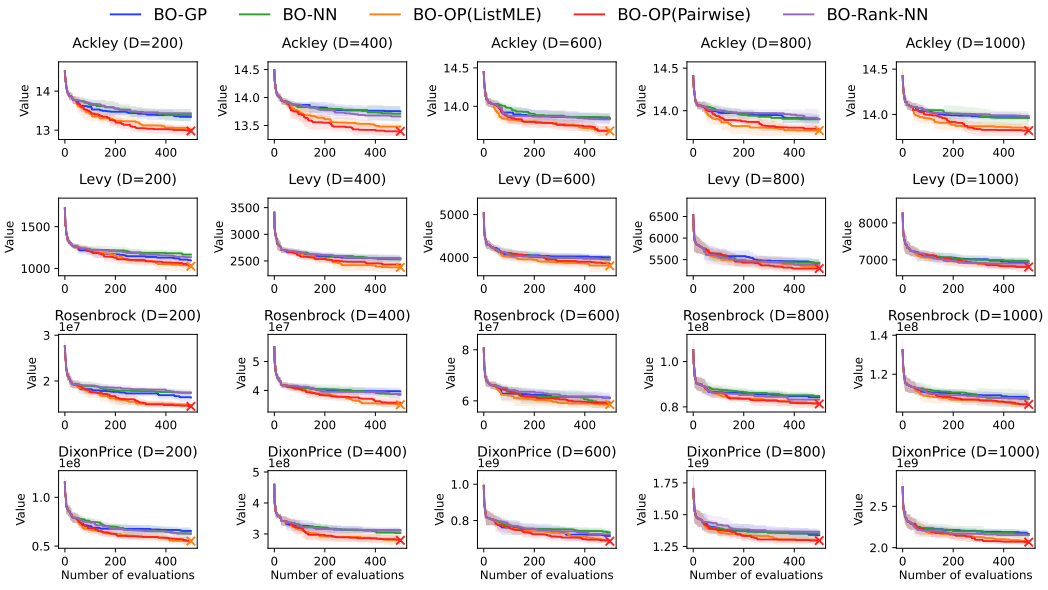


Figure 8: Synthetic benchmark trajectories under Standard BO, with the same evaluation budget and Thompson-sampling acquisition for all surrogates.

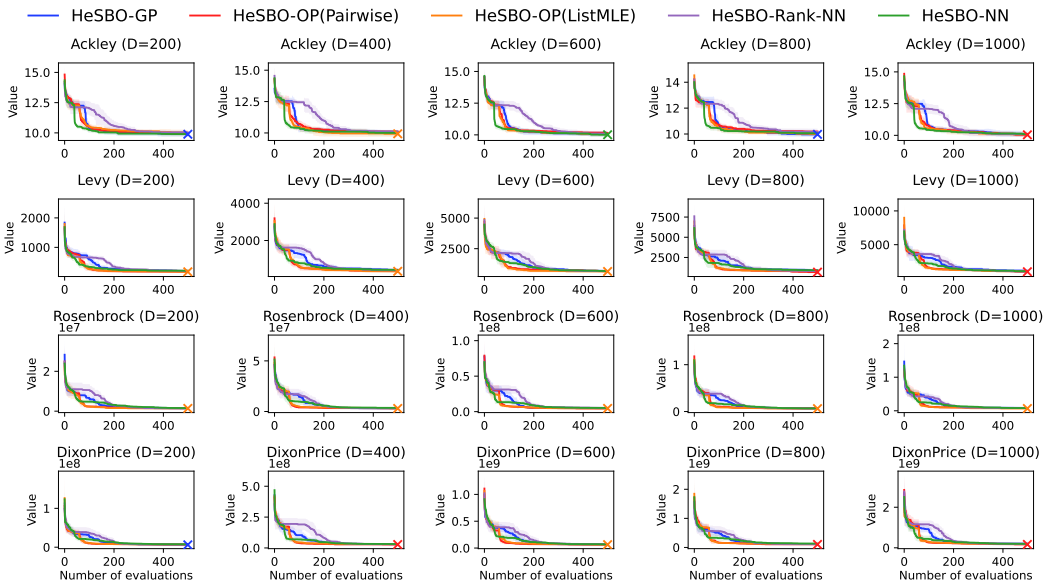


Figure 9: Synthetic benchmark trajectories under HeSBO with $d_{\text{emb}} = 12$. Neural variants share rank improvement; GP uses expected improvement.

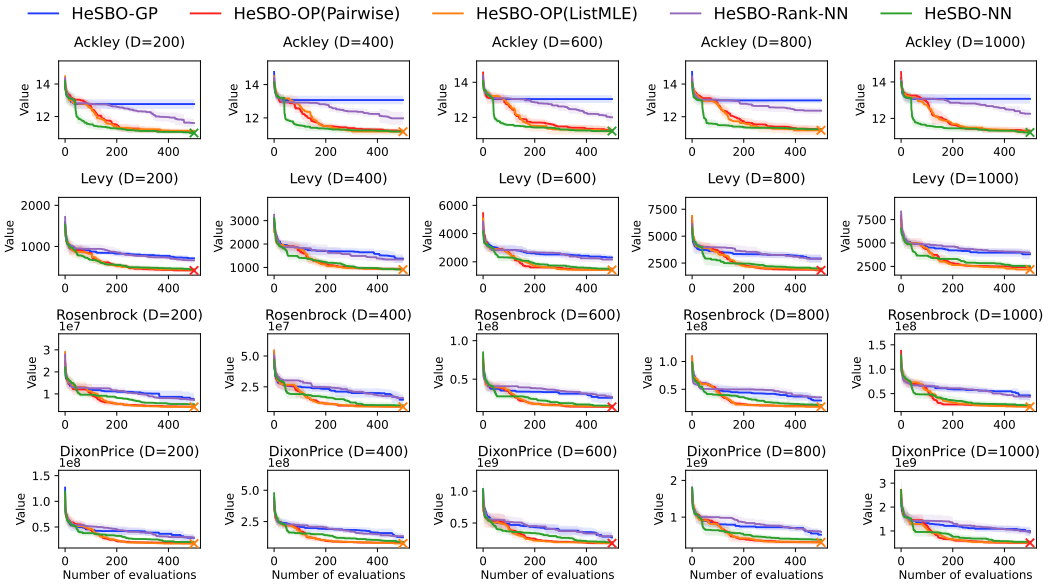


Figure 10: Synthetic benchmark trajectories under HeSBO with $d_{\text{emb}} = 24$. Acquisition rules are as in Section 2.4.

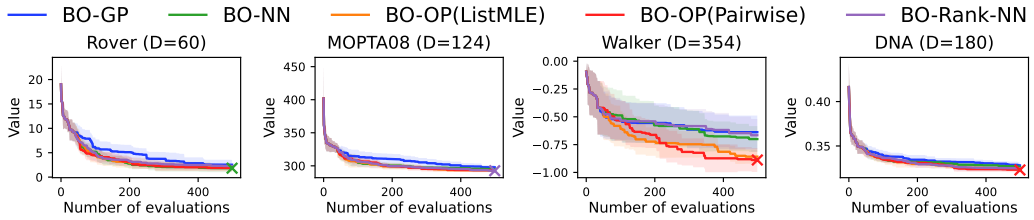


Figure 11: Real-world benchmark trajectories under Standard BO, with the same evaluation budget and Thompson-sampling acquisition for all surrogates.

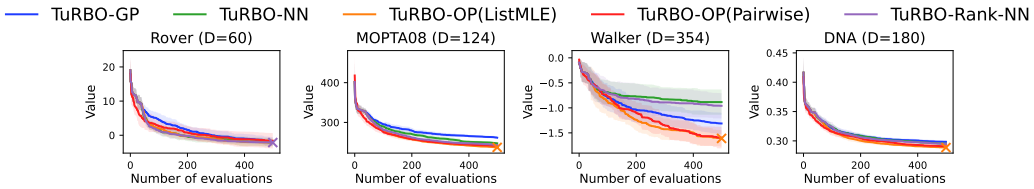


Figure 12: Real-world benchmark trajectories under TuRBO, with the same evaluation budget and Thompson-sampling acquisition for all surrogates.

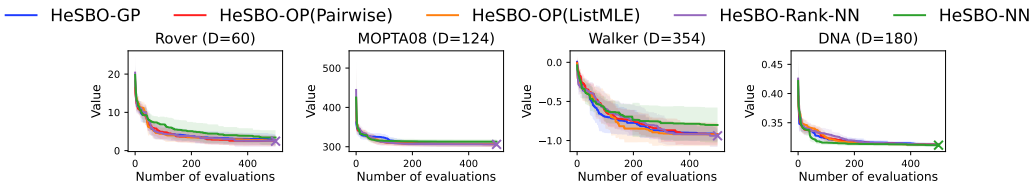


Figure 13: Real-world benchmark trajectories under HeSBO with $d_{\text{emb}} = 12$. Neural variants share rank improvement; GP uses expected improvement.

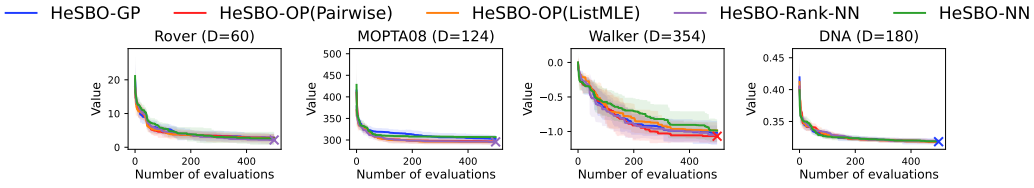


Figure 14: Real-world benchmark trajectories under HeSBO with $d_{\text{emb}} = 24$. Acquisition rules are as in Section 2.4.