Projective Preferential Bayesian Optimization

Petrus Mikkola [1]   **Milica Todorović** [2]   **Jari Järvi** [2]   **Patrick Rinke** [2]   **Samuel Kaski** [1 3]

Abstract

Bayesian optimization is an effective method for finding extrema of a black-box function. We propose a new type of Bayesian optimization for learning user preferences in high-dimensional spaces. The central assumption is that the underlying objective function cannot be evaluated directly, but instead a minimizer along a projection can be queried, which we call a projective preferential query. The form of the query allows for feedback that is natural for a human to give, and which enables interaction. This is demonstrated in a user experiment in which the user feedback comes in the form of optimal position and orientation of a molecule adsorbing to a surface. We demonstrate that our framework is able to find a global minimum of a high-dimensional black-box function, which is an infeasible task for existing preferential Bayesian optimization frameworks that are based on pairwise comparisons.

1. Introduction

Let $f : \mathcal{X} \rightarrow \mathbb{R}$ be a *black-box* function defined on a hypercube $\mathcal{X} = \prod_{d=1}^{D}[a_d, b_d]$ where $D \geq 2$. Without loss of generality we assume that $\mathbf{0} \in \mathcal{X}$. The objective is to find a global minimizer

$$\mathbf{x}^* = \operatorname*{argmin}_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}). \tag{1}$$

We assume, as in Preferential Bayesian Optimization (PBO, González et al., 2017), that f is not directly accessible. In PBO, queries to f can done in pairs of points $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$, and the binary feedback indicates whether $f(\mathbf{x}) > f(\mathbf{x}')$. In contrast, in our work we assume that queries to f are be done

[1] Helsinki Institute for Information Technology HIIT, Department of Computer Science, Aalto University, Espoo, Finland [2] Department of Applied Physics, Aalto University, Espoo, Finland [3] The University of Manchester, UK. Correspondence to: Petrus Mikkola <petrus.mikkola@aalto.fi>, Samuel Kaski <samuel.kaski@aalto.fi>.
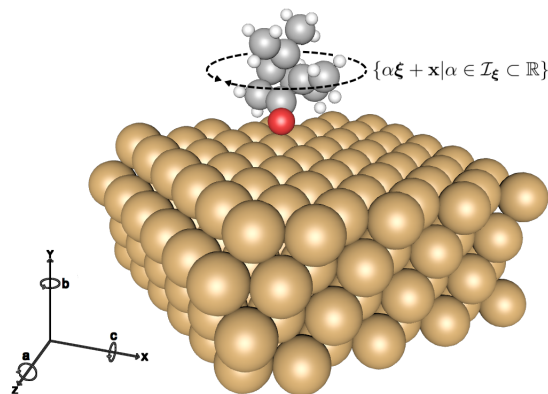
Figure 1. An illustration of a projective preferential query on molecular properties: in which rotation is the molecule most likely to bind to the surface. Here, $\mathbf{x} \in \mathcal{X}$ describes the location and orientation of a molecule as a vector $\mathbf{x} = (X, Y, Z, a, b, c)$. In the figure, the projective preferential query $(\boldsymbol{\xi}, \mathbf{x})$ finds the optimal rotation along the horizontal plane, defined by the coordinate b. This corresponds to setting $\xi_i = 0$ to all coordinates i except the one corresponding to b, and by rotating the molecule the expert then gives the optimal value α^* that corresponds to the optimal value for coordinate b. The other coordinates are kept fixed (these are determined by $\mathbf{x}$).

over the projection onto a *projection vector* $\boldsymbol{\xi} \in \Xi \subset \mathcal{X}$. The feedback is the optimal scalar projection, that is, the length α^* of the projection in the direction $\boldsymbol{\xi}$. We assume that there are zero coordinates in $\boldsymbol{\xi}$, and these coordinates are set to fixed values described by a *reference vector* $\mathbf{x} \in \mathcal{X}$. Formally, given a query $(\boldsymbol{\xi}, \mathbf{x})$, the feedback is obtained as *a minimizer over the possible scalar projections*,

$$\alpha^* = \operatorname*{argmin}_{\alpha \in \mathcal{I}_{\boldsymbol{\xi}}} f(\alpha\boldsymbol{\xi} + \mathbf{x}), \tag{2}$$

where $\mathcal{I}_{\boldsymbol{\xi}} \equiv \{\alpha \in \mathbb{R} | \alpha\boldsymbol{\xi} + \mathbf{x} \in \mathcal{X}\}$.

What are then natural use cases for such *projective preferential queries*? The main motivation comes from humans serving as the oracles. The form of the query enables efficient learning of user preferences over choice sets in which each choice has multiple attributes, and in particular over continuous choice sets. An important application is knowledge elicitation from trained professionals (doctors, physi-

cists, etc.). For example, we may learn a material scientists preferences, that is, insight based on prior knowledge and experience, over molecular translations and orientations as a molecule adsorbs to a surface. In this case, a projective preferential query could correspond to finding an optimal rotation (see Figure 1), which the scientists can easily give by rotating the molecule in a visual interface.

Probabilistic preference learning is a relatively new topic in machine learning research but has a longer history in econometrics and psychometrics (McFadden, 1981; 2001; Stern, 1990; Thurstone, 1927). A wide range of applications of these models exists, for instance in computer graphics (Brochu et al., 2010), expert knowledge elicitation (Soufiani et al., 2013), revenue management systems of airlines (Carrier & Weatherford, 2015), rating systems, and almost any application that contains users' preference modeling. An established probabilistic model is Thurstone-Mosteller (TM) model that measures a process of pairwise comparisons (Thurstone, 1927; Mosteller, 1951). In the preference learning context, the models based on the TM-model can be applied to learning preferences from pairwise comparison feedback (e.g. Chu & Ghahramani, 2005). An extension of this research into the interactive learning setting is studied by, among others, Brochu et al. (2008) and González et al. (2017). All these approaches resort to pairwise feedbacks. Koyama et al. (2017) proposed an extension of Bayesian optimization for learning optimal parameter values for visual design tasks by letting a user give feedback as a slider manipulation. They considered only two pairwise comparisons per slider since they tested to "included more sampling points" but they "did not observe any significant improvement in the optimization behavior".

A drawback of most preference learning frameworks is their incapability to handle high-dimensional input spaces. The underlying reason is a *combinatorial explosion* in the number of possible comparisons with respect to the number of dimensions D, $O(K^{2D})$, given K grid-points per dimension. This implies that a single pairwise comparison has *low information content* in high-dimensional spaces. This problem was mitigated by González et al. (2017) by capturing the correlations among pairs of queries (duels). However, it is still difficult to scale that method to high-dimensional spaces, say higher than 2-dimensional (see Section 4). Furthermore, the numerical computations become infeasible in a high-dimensional setting, especially the optimization of an acquisition function or finding a Condorcet winner.

In this paper, we introduce a Bayesian framework, which we call *Projective Preferential Bayesian Optimization* (PPBO), that scales to high-dimensional input spaces. A main reason is that the information content of a projective preferential query is much higher than that of a pairwise preferential query. A projective preferential query is equiv-

alent to infinite pairwise comparisons along a projection. An important consequence is that with projective preferential queries, the user's workload in answering the queries will be considerably reduced. Source code is available at `https://github.com/AaltoPML/PPBO`.

2. Learning preferences from projective preferential feedback

In this section we introduce a Bayesian framework capable of dealing with projective preferential data. A central idea is to model the user's utility function, that is f, as a Gaussian process as first proposed by Chu & Ghahramani (2005). We extend this line of study to allow projective preferential queries, by deriving a tractable likelihood, proposing a method to approximate it, and introducing four acquisition criteria for enabling interactive learning in this setting.

In this paper, for convenience, we will formulate the method for maximization instead of minimization as in (2), without loss of generality.

2.1. Likelihood

Our probabilistic model of user preferences is built upon the Thurstone's law of comparative judgement (Thurstone, 1927). A straightforward way to formalize this would be to assume pairwise comparisons are corrupted by Gaussian noise: $x \succ x'$, if and only if $f(x) + \varepsilon > f(x') + \varepsilon'$, where the latent function f is *a utility function* that characterizes user preferences described by the preference relation $\succ$. The standard assumption is that ε and ε' are identically and independently distributed Gaussians. Here, we deviate slightly from this assumption: Given two alternatives $(\alpha\boldsymbol{\xi} + \mathbf{x}), (\beta\boldsymbol{\xi} + \mathbf{x}) \in \mathcal{X}$, we assume that $\alpha\boldsymbol{\xi} + \mathbf{x} \succ \beta\boldsymbol{\xi} + \mathbf{x}$, if and only if $f(\alpha\boldsymbol{\xi} + \mathbf{x}) + W(\alpha) > f(\beta\boldsymbol{\xi} + \mathbf{x}) + W(\beta)$, where W is a *Gaussian white noise process* with zeromean and autocorrelation $\mathbb{E}(W(t)W(t + \tau)) = \sigma^2$ if $\tau = 0$, and zero otherwise.

We would like to find the likelihood for an observation $(\alpha, (\boldsymbol{\xi}, \mathbf{x}))$ that corresponds to uncountably infinite pairwise comparisons: $\alpha\boldsymbol{\xi} + \mathbf{x} \succ \beta\boldsymbol{\xi} + \mathbf{x}$ for $\beta \neq \alpha$. For each comparison we condition on $W(\alpha)$ (more details in Supplementary material),

$$P(\alpha\boldsymbol{\xi} + \mathbf{x} \succ \beta\boldsymbol{\xi} + \mathbf{x} \mid W(\alpha) = w)$$
$$= 1 - \Phi\left(\frac{f(\beta\boldsymbol{\xi} + \mathbf{x}) - f(\alpha\boldsymbol{\xi} + \mathbf{x}) - w}{\sigma}\right),$$

where Φ is the cumulative distribution function of the standard normal distribution. For a single comparison we have

$$P(\alpha\boldsymbol{\xi} + \mathbf{x} \succ \beta\boldsymbol{\xi} + \mathbf{x})$$
$$= 1 - [\Phi * \phi]\left(\frac{f(\beta\boldsymbol{\xi} + \mathbf{x}) - f(\alpha\boldsymbol{\xi} + \mathbf{x})}{\sigma}\right),$$

where ϕ is the probability density of the standard normal distribution and $*$ is the convolution operator. For infinite comparisons, we first consider a finite number of comparisons m. By their independence, we have

$$\mathrm{P}\left(\alpha\boldsymbol{\xi} + \mathbf{x} \succ \beta_1\boldsymbol{\xi} + \mathbf{x}, ..., \alpha\boldsymbol{\xi} + \mathbf{x} \succ \beta_m\boldsymbol{\xi} + \mathbf{x}\right)$$
$$= \prod_{j=1}^{m}\left(1 - [\Phi * \phi]\left(\frac{f(\beta_j\boldsymbol{\xi} + \mathbf{x}) - f(\alpha\boldsymbol{\xi} + \mathbf{x})}{\sigma}\right)\right).$$

By letting the number of points m in an increasing sequence $\beta_1, ..., \beta_m$ of the partition of the interval $\mathcal{I}_{\boldsymbol{\xi}}\backslash\{\alpha\}$ to approach infinity, we can interpret this as a Volterra (product) integral

$$\exp\left(-\int_{\mathcal{I}_{\boldsymbol{\xi}}}[\Phi * \phi]\left(\frac{f(\beta\boldsymbol{\xi} + \mathbf{x}) - f(\alpha\boldsymbol{\xi} + \mathbf{x})}{\sigma}\right)d\beta\right).$$

The joint log-likelihood of a dataset $\mathcal{D}$, denoted as $\mathcal{L}(\mathcal{D}|\mathbf{f})$, takes the form

$$-\sum_{i=1}^{N}\int_{\mathcal{I}_{\boldsymbol{\xi}^i}}[\Phi * \phi]\left(\frac{f(\beta\boldsymbol{\xi}^i + \mathbf{x}^i) - f(\alpha^i\boldsymbol{\xi}^i + \mathbf{x}^i)}{\sigma}\right)d\beta.$$

2.2. Prior

First, we introduce notation. Assume that N projective preferential queries have been performed and gathered into a dataset $\mathcal{D} = \{(\alpha^i, (\boldsymbol{\xi}^i, \mathbf{x}^i))\}_{i=1}^{N}$. For every data instance $(\alpha^i, (\boldsymbol{\xi}^i, \mathbf{x}^i))$, we also consider a sequence of pseudo-observations $\{(\beta_j^i\boldsymbol{\xi}^i, \mathbf{x}^i)\}_{j=1}^{m}$. Technically, the pseudo-observations are Monte-Carlo samples needed for integrating the likelihood. The latent function values evaluated on those points are gathered into a vector,

$$\mathbf{f}^{(i)} \equiv \left(f(\alpha^i\boldsymbol{\xi}^i + \mathbf{x}^i), \{f(\beta_j^i\boldsymbol{\xi}^i + \mathbf{x}^i)\}_{j=1}^{m}\right).$$

The latent function vector over all points is formed by concatenating over $\mathbf{f} \equiv (\mathbf{f}^{(i)})_{i=1}^{N}$.

The user's utility function f is modelled as a Gaussian process (Rasmussen & Williams, 2005). GP model fits ideally to this objective, since it is flexible (non-parametric) and can conveniently handle uncertainty (predictive distributions can be derived analytically). In particular, it allows us to have insight into those regions of the space $\mathcal{X}$ in which either we are uncertain about user preferences due to lack of data, or because the user gives inconsistent feedback. A possible reason for the latter is that one of the preference axioms, *transitivity* or *completeness*, is violated in those regions. A weak preference relation $\succeq$ is *complete* if for all $\mathbf{x}, \mathbf{y} \in \mathcal{X}$, either $\mathbf{x} \succeq \mathbf{y}$ or $\mathbf{y} \succeq \mathbf{x}$ holds. That is, a user is able to reveal their preferences over all possible pairwise comparisons. Similarly, $\succeq$ is *transitive* if for any $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{X}$ the following holds: ($\mathbf{x} \succeq \mathbf{y}$ and $\mathbf{y} \succeq \mathbf{z}$) implies that $\mathbf{x} \succeq \mathbf{z}$. That is, a user has consistent preferences. This together

with the continuity of $\succeq$ guarantees the existence of a real-valued continuous utility function that represents $\succeq$ (Debreu, 1954).

Thus, we assume as Chu & Ghahramani (2005), that the prior of the utility function follows a zero-mean Gaussian process,

$$p(\mathbf{f}) = \frac{1}{(2\pi)^{\frac{N}{2}}|\Sigma|^{\frac{1}{2}}}\exp(-\frac{1}{2}\mathbf{f}^{\top}\Sigma^{-1}\mathbf{f}),$$

where the ij^{th}-element of the covariance matrix is determined by a kernel k as $\Sigma_{ij} = k(\mathbf{x}^i, \mathbf{x}^j)$. Throughout the paper, we assume the *squared exponential* kernel $k(\mathbf{x}, \mathbf{x}) = \sigma_f^2\exp(\frac{1}{-2l}\|\mathbf{x} - \mathbf{x}'\|^2)$, where the σ_f and l are hyperparameters.

2.3. Posterior

For the sake of simplicity, we use the Laplace approximation for the posterior distribution. A maximum a posteriori (MAP) estimate is needed for that,

$$\underset{\mathbf{f}}{\operatorname{argmax}}\,\mathrm{P}(\mathbf{f}|\mathcal{D}) = \underset{\mathbf{f}}{\operatorname{argmax}}(\mathrm{P}(\mathbf{f})\mathcal{L}(\mathcal{D}|\mathbf{f})) = \underset{\mathbf{f}}{\operatorname{argmax}}\,T(\mathbf{f}),$$

where we denote the functional (log-scaled posterior)

$$T(\mathbf{f}) \equiv -\frac{1}{2}\mathbf{f}^{\top}\Sigma^{-1}\mathbf{f}$$
$$-\sum_{i=1}^{N}\int_{\mathcal{I}_{\boldsymbol{\xi}^i}}[\Phi * \phi]\left(\frac{f(\beta\boldsymbol{\xi}^i + \mathbf{x}^i) - f(\alpha^i\boldsymbol{\xi}^i + \mathbf{x}^i)}{\sigma}\right)d\beta.$$

The convolution $\Phi * \phi$ can be efficiently approximated by Gauss-Hermite quadrature. The outer integral is approximated as a Monte-Carlo integral,

$$\int_{\mathcal{I}_{\boldsymbol{\xi}}}[\Phi * \phi]\left(\frac{f(\beta\boldsymbol{\xi} + \mathbf{x}) - f(\alpha\boldsymbol{\xi} + \mathbf{x})}{\sigma}\right)d\beta$$
$$\approx \frac{\ell(\mathcal{I}_{\boldsymbol{\xi}})}{m}\sum_{j=1}^{m}[\Phi * \phi]\left(\frac{f(\beta_j\boldsymbol{\xi} + \mathbf{x}) - f(\alpha\boldsymbol{\xi} + \mathbf{x})}{\sigma}\right),$$

where the pseudo-observations $(\beta_j\boldsymbol{\xi})_j^m$ for $j = 1, ..., m$ are sampled from a suitable distribution. Our choice is to use a family of truncated generalized normal (TGN) distributions, since it provides a continuous transformation from the uniform distribution to the truncated normal distribution, such that the locations of distributions can be specified. The idea is to concentrate pseudo-observations more densely around the optimal value $\alpha\boldsymbol{\xi}$ as the number of queries increases. For more details, see Supplementary material.

For notational convenience, define

$$\Delta_{i,j}(\mathbf{f}) \equiv \frac{f(\beta_j^i\boldsymbol{\xi}^i + \mathbf{x}^i) - f(\alpha^i\boldsymbol{\xi}^i + \mathbf{x}^i)}{\sigma}.$$

If the domain is normalized to $\mathcal{X} = \prod_{d=1}^{D}[0, 1]$, and the projections are normalized to $\xi = \frac{\xi}{\|\xi\|_\infty}$, then $\ell(\mathcal{I}_\xi) = 1$. Hence, under this normalization, the functional T can be approximated as

$$T(\mathbf{f}) \approx -\frac{1}{2}\mathbf{f}^\top \Sigma^{-1} \mathbf{f} - \frac{1}{m}\sum_{i=1}^{N}\sum_{j=1}^{m}[\Phi * \phi]\big(\Delta_{i,j}(\mathbf{f})\big).$$

The MAP estimate can be efficiently solved by a second-order iterative optimization algorithm, since the gradient and the Hessian can be easily derived for T.

The Laplace approximation of the posterior amounts to the second-order Taylor approximation of the log posterior around the MAP estimate. In the ordinary (non-log) scale, this reads

$$P(\mathbf{f}|\mathcal{D}) \approx P(\mathbf{f}_{\text{MAP}}|\mathcal{D})\exp\big(-\frac{1}{2}(\mathbf{f} - \mathbf{f}_{\text{MAP}})^\top H(\mathbf{f} - \mathbf{f}_{\text{MAP}})\big),$$

where the matrix H is the negative Hessian of the log-posterior at the MAP estimate, $H \equiv -\nabla\nabla\log P(\mathbf{f}|\mathcal{D})|_{\mathbf{f}=\mathbf{f}_{\text{MAP}}} = \Sigma^{-1} + \Lambda$.[1] In other words, the posterior distribution is approximated as a multivariate normal distribution with mean $\mathbf{f}_{\text{MAP}}$ and the covariance matrix $(\Sigma^{-1} + \Lambda)^{-1}$.

2.4. Predictive distribution

Based on the well-known properties of the multivariate Gaussian distribution, the predictive distribution of $\mathbf{f}$ is also Gaussian. Given test locations $(\mathbf{y}^{(1)}, ..., \mathbf{y}^{(M)})$, consider the N by M matrix $K \equiv [k(\mathbf{y}^{(j)}, \mathbf{x}^{(i)})]_{ij}$. The predictive mean and the predictive covariance at test locations are (for more details see (Rasmussen & Williams, 2005) or (Chu & Ghahramani, 2005))

$$\boldsymbol{\mu}_{\text{pred}} = K^\top \Sigma^{-1}\mathbf{f}_{\text{MAP}}$$
$$\Sigma_{\text{pred}} = \Sigma' - K^\top(\Sigma + \Lambda^{-1})^{-1}K,$$

where Σ' is the covariance matrix of the test locations.

3. Sequential learning by projective preferential query

In this section, we discuss how to select the next projective preferential query $(\xi, \mathbf{x})$. We will choose the next query as a maximizer of an *acquisition function* $\alpha(\xi, \mathbf{x})$, for instance, we will consider a modified version of the *expected*

---

[1] We denote the partial derivatives matrix evaluated at MAP estimate as

$$\Lambda \equiv \frac{\partial^2}{\partial\mathbf{f}\partial\mathbf{f}^\top}\frac{1}{m}\sum_{i=1}^{N}\sum_{j=1}^{m}[\Phi * \phi]\big(\Delta_{i,j}(\mathbf{f})\big)\bigg|_{\mathbf{f}=\mathbf{f}_{\text{MAP}}}.$$

improvement acquisition function (Jones et al., 1998). The optimization $(\xi, \mathbf{x})_{next} = \text{argmax}_{(\xi, \mathbf{x})}\,\alpha(\xi, \mathbf{x})$ is carried out by using Bayesian optimization (more details in Supplementary material).

If the oracle is a human, this allows us to learn user preferences in an iterative loop, making PPBO *interactive*. However, this interesting special case, where f is a utility function of a human, also brings forth issues due to *bounded rationality*. We apply here the following narrow definition of this more general concept (see Simon, 1990): Bounded rationality is the idea that users give feedback that reflects their preferences, but within the limits of the information available to them and their mental capabilities.

3.1. The effects of bounded rationality on the optimal next query

If the oracle is a human, it is important to realize that the optimal next $(\xi, \mathbf{x})$ is not solely the one which optimally balances the *exploration-exploitation trade-off* – as it is for a perfect oracle – but the optimal $(\xi, \mathbf{x})$ takes also into account human cognitive capabilities and limitations. For instance, the more there are non-zero coordinates in ξ, the greater the "cognitive burden" to a human user, and the harder it becomes to give useful feedback. Thus, if there is a human in the loop, the choice of ξ should take into account both the optimization needs and what types of queries are convenient for the user.

The projective preferential feedback (2) may not be single-valued or even well-defined for all $\xi \in \Xi$, if the oracle is a human. For instance, the user may not be able to explicate their preferences with respect to the dth attribute, that is, the preferences do not satisfy the *completeness axiom*. Formally, this means that if $\xi = \mathbf{e}_d$ (the dth-standard unit vector), then for some $\mathbf{x} \in \mathcal{X}$ it holds that $\text{argmax}_{\alpha \in \mathcal{I}_\xi}\,f(\alpha\xi + \mathbf{x})$ is multi-valued, a random variable or not well-defined – depending on how we interpret the scenario in which the user should say "I do not know" but instead gives an arbitrary feedback. Fortunately, this incompleteness can be easily handled when a GP is used for modelling f; it just implies that the posterior variance is high along the dimension d.

A possible solution would be to allow the answer "I do not know", and to design an acquisition function that is capable of discovering and avoiding those regions in the space $\mathcal{X}$ where the user gives inconsistent feedback due to any source of bounded rationality. This challenge is left for future research. It is noteworthy that the acquisition function we introduce next, performed well in the user experiment covered in Section 5.

3.2. Expected improvement by projective preferential query

We define *the expected improvement by projective preferential query* at the nth-iteration by

$$\text{EI}_n(\boldsymbol{\xi}, \mathbf{x}) \equiv \mathbb{E}_n\Big(\max\big\{ \max_{\alpha \in \mathcal{I}_{\boldsymbol{\xi}}} f(\alpha\boldsymbol{\xi} + \mathbf{x}) - \mu_n^*, 0 \big\} \Big), \quad (3)$$

where μ_n^* denotes the highest value of the predictive posterior mean, and the expectation is conditioned on the data up to the nth-iteration. The maximum over α models the anticipated feedback.

EI_n can be approximated as a Monte-Carlo integral (up to a multiplicative constant that does not depend on $(\boldsymbol{\xi}, \mathbf{x})$),

$$\frac{1}{K} \sum_{k=1}^K \max\big\{ \max_{\alpha \in \mathcal{I}_{\boldsymbol{\xi}}} \tilde{f}_k(\alpha\boldsymbol{\xi} + \mathbf{x}) - \mu_n^*, 0 \big\}, \quad (4)$$

where $\max_{\alpha \in \mathcal{I}_{\boldsymbol{\xi}}} \tilde{f}_k(\alpha\boldsymbol{\xi} + \mathbf{x})$ is approximated by using *discrete[2] Thompson sampling* as described in (Hernández-Lobato et al., 2014). Discrete Thompson sampling draws a finite sample from the GP posterior distribution, and then returns the maximum over the sample. The steps needed to approximate EI_n are summarized in Algorithm 1.

Algorithm 1 Approximate $\text{EI}_n(\boldsymbol{\xi}, \mathbf{x})$

---
input $(\boldsymbol{\xi}, \mathbf{x})$ and $K \geq 1$, $J \geq 1$
   1. Compute $(\Sigma + \Lambda^{-1})^{-1}$
  for $k = 1, 2, ..., K$ **do**
    2. Draw $(\beta_j)_{j=1}^J$
    3. Draw $\big(\tilde{f}(\beta_j\boldsymbol{\xi} + \mathbf{x})\big)_{j=1}^J$ from the predictive distribution $\mathcal{N}(\boldsymbol{\mu}_{\text{pred}}, \Sigma_{\text{pred}})$
    4. $z_k \leftarrow \max_j \big\{ \big(\tilde{f}(\beta_j\boldsymbol{\xi} + \mathbf{x})\big)_{j=1}^J \big\}$
  end for
output $\frac{1}{K} \sum_{k=1}^K \max\big\{ z_k - \mu_n^*, 0 \big\}$

The bottlenecks are the first and the third steps. In the third step, a predictive covariance matrix of size $J \times J$ needs to be computed, and then a sample from the multivariate normal distribution needs to be drawn. Hence, the time complexity of Algorithm 1 is $O(N^3m^3 + KN^2m^2J + KNmJ^2 + KJ^3)$, where the terms come from a matrix inversion (the first step), two matrix multiplications, and a Cholesky decomposition, respectively. Recall that N, m, K and J refer to the number of observations, pseudo-observations, Monte-Carlo samples, and grid points, respectively.

---

[2]Another alternative is to consider *continuous Thompson sampling* to draw a continuous sample path from the GP model, and then maximize it. The method is based on Bochner's theorem and the equivalence between a Bayesian linear model with random features and a Gaussian process. For more details see (Hernández-Lobato et al., 2014).

3.3. Pure exploitation and exploration

In the experiments we use pure exploration and exploitation as baselines. A natural interpretation of pure exploitation in our context is to select the next query $(\boldsymbol{\xi}, \mathbf{x})$ such that $\boldsymbol{\xi} + \mathbf{x} = \mathbf{x}^* \equiv \text{argmax}_{\mathbf{x}' \in \mathcal{X}} \mu_n(\mathbf{x}')$, where $\mu_n(\mathbf{x}')$ is the posterior mean of the GP model at location $\mathbf{x}'$, given all the data so far $\mathcal{D}_n$.

We interpret pure exploration as maximization of the GP variance on a query $(\boldsymbol{\xi}, \mathbf{x})$ given the anticipated feedback. That is, the pure explorative acquisition strategy maximizes the following acquisition function

$$\text{Explore}(\boldsymbol{\xi}, \mathbf{x}) \equiv \mathbb{V}_n\Big(\max_{\alpha \in \mathcal{I}_{\boldsymbol{\xi}}} f(\alpha\boldsymbol{\xi} + \mathbf{x}) \Big). \quad (5)$$

In practice, (5) is approximated by Monte-Carlo integration and discrete Thompson sampling in the same vein as in Algorithm 1.

3.4. Preferential coordinate descent (PCD)

The fourth acquisition strategy corresponds to the interesting special case where $\boldsymbol{\xi} = \mathbf{e}_d$ (the dth-standard unit vector), and the coordinates d are rotated in a cyclical order for each query. The reference vector $\mathbf{x} = (x_1, ..., x_{d-1}, 0, x_{d+1}, ..., x_D)$ can be chosen in several ways, but it is natural to consider an exploitative strategy in which $\mathbf{x}$ is set to $\mathbf{x}^*$ except for the dth-coordinate which is set to zero. We call this acquisition strategy *Preferential Coordinate Descent* (PCD), since PPBO with PCD acquisition is closely related to a *coordinate descent algorithm* that successively minimizes an objective function along coordinate directions. The PPBO method with PCD acquisition (PPBO-PCD) differs from the classical coordinate descent in two ways: First, PPBO-PCD assumes that direct function evaluations are not possible but instead projective preferential queries are. Second, it models the black-box function f (as a GP) whereas the classical coordinate descent does not. This makes PPBO-PCD able to take advantage of past queries from every one-dimensional optimization.

When comparing to other acquisition strategies, we show that PCD performs well in numerical experiments (when f is not a utility function but a numerical test function). This agrees with the results in the optimization literature; for instance: if f is pseudoconvex with continuous gradient, and $\mathcal{X}$ is compact and convex with "nice boundary", then the coordinate descent algorithm converges to a global minimum (Spall, 2012, Corollary 3.1). However, PCD may not perform so well in high-dimensional spaces, since it cannot query in between the dimensions. For instance, the expected improvement by projective preferential query outperformed PCD on a 20D test function (see Section 4), since it allows to query arbitrary projections.

4. Numerical experiments

In this section we demonstrate the efficiency of the PPBO method in high-dimensional spaces, and experiment with various acquisition strategies in numerical experiments on simulated functions.

The goal is to find a global minimum of a black-box function f by querying it either through (i) pairwise comparisons or (ii) projective preferential queries. For (i) we use the PBO method of González et al. (2017), which is state of the art among Gaussian process preference learning frameworks that are based on pairwise comparisons. For (ii) we use the PPBO method as introduced in this paper. The four different acquisition strategies introduced in Section 3 are compared against the baseline that samples a random $(\boldsymbol{\xi}, \mathbf{x})$. For the PBO method, we consider a random and a dueling-Thompson sampling acquisition strategies. In total seven different methods are compared: the expected improvement by projective preferential query (PPBO-EI), pure exploitation (PPBO-EXT), pure exploration (PPBO-EXR), preferential coordinate descent (PPBO-PCD), random (PPBO-RAND), and for the PBO; random (PBO-RAND) and a variant of dueling-Thompson sampling (PBO-DTS). For more details, see Supplementary material.

For f we consider four different test functions: Six-hump-camel2D, Hartmann6D, Levy10D and Ackley20D.[3] We add a small Gaussian error term to the test function outputs. There are as many initial queries as there are dimensions in a test function. The ith-initial query corresponds to $\boldsymbol{\xi} = \mathbf{e}_i$, that is, to the ith-coordinate projection, and the reference vector $\mathbf{x}$ is uniformly random. We consider a total budget of 100 queries. The results are depicted in Figure 2.[4]

PPBO-PCD obtained the best performance on three of the four test functions. On the high-dimensional test function Ackley20D, PPBO-EI performed best. Unsurprisingly, all PPBO variants clearly outperformed all PBO variants. Since the performance gap between PPBO-RAND and PBO-RAND is so high, we conclude that from the optimization perspective; whenever a projective preferential query is possible, a PPBO-type of approach should be preferred to an approach that is based on pairwise comparisons. However, we note that it is better to think of PPBO as a complement, not as a substitute for PBO. In the applications, pairwise comparisons may be preferred, for instance, if they are more convenient to a user, and the underlying choice space is low-dimensional.

To illustrate the low information content of pairwise comparisons, we ran a test on the Six-hump-camel2D function.

---

[3]https://www.sfu.ca/~ssurjano/optimization.html

[4]All experiments of each test function were run on a computing infrastructure of 24x Xeon Gold 6148 2.40GHz cores and 72GB RAM. The longest experiment (Ackley20D) took in total 24h.

We trained a GP classifier with 2000 random queries (duels), and found a Condorcet winner (see González et al., 2017) by maximizing the soft-Copeland score (33×33 MC-samples used for the integration) by using Bayesian optimization (500 iterations with 10 optimization restarts). This took 41 minutes on the 8th-gen Intel i5-CPU, and the distance to a true global minimizer was $\|\mathbf{x}_c - \mathbf{x}_{true}\| = \|(0.1770, -0.0488) - (0.0898, -0.7126)\| \approx 0.67,$ and the corresponding function value was 0.1052 compared to a true global minimum value -1.0316. In contrast, PPBO-RAND reached this level of accuracy at the first queries, as seen from Figure 2.

5. User experiment

In this section we demonstrate the capability of PPBO to correctly and efficiently encode user preferences from projective preferential feedback.

We consider a material science problem of a single organic molecule adsorbing to an inorganic surface. This is a key step in understanding the structure at the interface between organic and inorganic films inside electronic devices, coatings, solar cells and other materials of technological relevance. The molecule can bind in different adsorption configurations, altering the electronic properties at the interface and affecting device performance. Exploring the structure and property phase space of materials with accurate but costly computer simulations is a difficult task. Our objective is to find the most stable surface adsorption configuration through human intuition and subsequent computer simulations. The optimal configuration is the one that minimises the computed adsorption energy.

Our test case is the adsorption of a non-symmetric, bulky molecule camphor on the flat surface of (111)-plane terminated Cu slab. Some understanding of chemical bonding is required to infer correct adsorption configurations. The user is asked to consider the adsorption structure as a function of molecular orientation and translation near the surface. These are represented with 6 physical variables: angles α, β, γ of molecular rotation around the X, Y, Z Cartesian axes (in the range [0, 360] deg.), and distances x, y, z of translation above the surface (with lattice vectors following the translational symmetry of the surface). The internal structures of the molecule and surface were kept fixed since little structural deformation is expected with adsorption. A similar organic/inorganic model system and experiment scenario was previously employed to detect the most stable surface structures with autonomous BO, given the energies of sampled configurations (Todorović et al., 2019).

In this interactive experiment, the users encode their preferred adsorption geometry as a location in the 6-dimensional phase space. We employ the quantum-
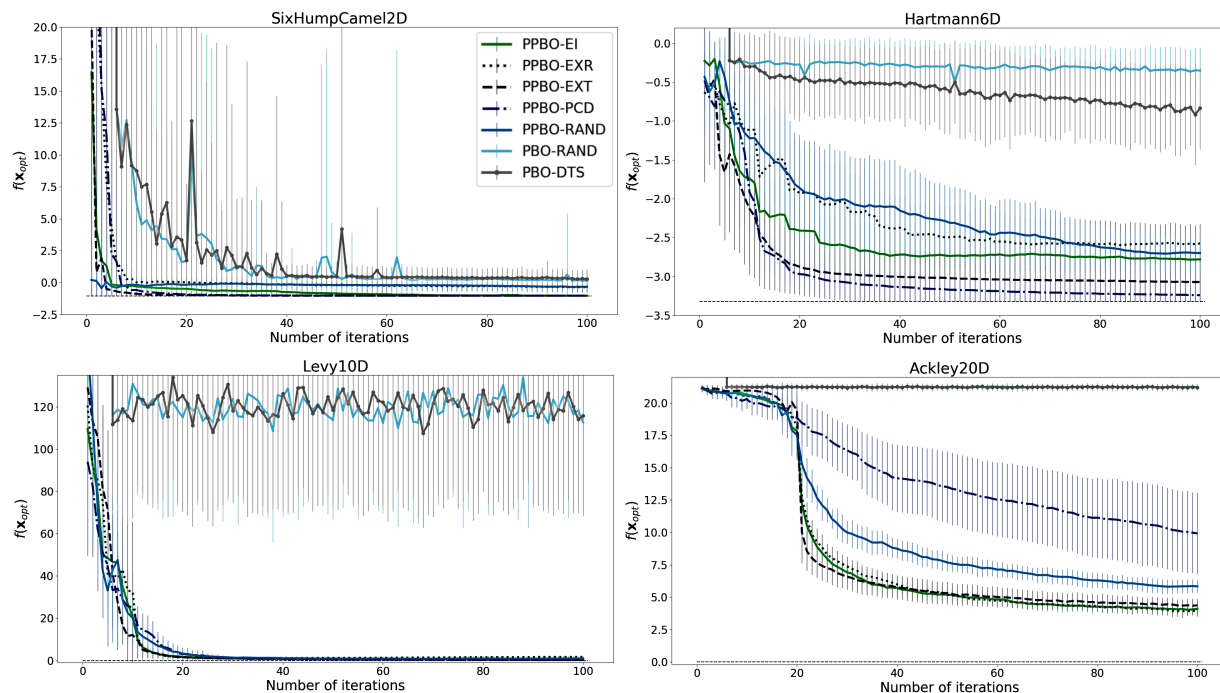
Figure 2. Convergence of different methods across 4 objective functions. The results are averaged over 25 different random initializations (for PBO we run 100 seeds). The vertical axis represent the value of the true objective function at the best guess: $\mathbf{x}_{opt} = \operatorname{argmax}_{\mathbf{x}} \mu_n(\mathbf{x})$ in PPBO, and $\mathbf{x}_{opt} = "Condorcet\ winner"$ in PBO. The horizontal axis represents the number of projective preferential queries in PPBO, and the number of pairwise comparisons in PBO. The black horizontal dashed line indicates the global minimum of the objective function (a small Gaussian noise is added to the function values). The standard deviations are depicted by the vertical lines.

mechanical atomistic simulation code FHI-aims (Blum et al., 2009) to i) compute the adsorption energy **E** of this preferred choice, and ii) optimise the structure from this initial position to find the nearest local energy minimum in phase space, **E***. We also consider the number of optimization steps **N** needed to reach the nearest minimum as a measure of quality of the initial location.

There are four different test users: two materials science experts (human: a PhD student and an experienced researcher, both of them know the optimal solution), a non-expert (human), and a random bot (computer). The hypothesis is that: *the materials science experts should obtain structures associated with lower energy minimum points*. We consider only coordinate projections, that is $\boldsymbol{\xi} \in \{\mathbf{e}_1, ..., \mathbf{e}_6\}$. In other words, we let the user choose the optimal value for one dimension at a time.

The total number of queries was 24, of which 6 were initial queries. The ith-initial query corresponded to $\boldsymbol{\xi} = \mathbf{e}_i$, that is, to the ith-coordinate projection. The initial values for the reference coordinate vector **x** were fixed to the same value across all user sessions. For acquisition, we used the expected improvement by projective preferential query. Since we allowed only coordinate projections for $\boldsymbol{\xi}$, we first selected $\boldsymbol{\xi}_{n+1} = \operatorname{argmax}_{\boldsymbol{\xi} \in \{\mathbf{e}_1, ..., \mathbf{e}_6\}} \int \mathrm{EI}_n(\boldsymbol{\xi}, \mathbf{x}) d\mathbf{x}$,

and then, either $\mathbf{x}_{n+1} = \operatorname{argmax}_{\mathbf{x}} \mu_n(\mathbf{x})$ (EI-EXT), or the next $\mathbf{x}_{n+1}$ was drawn uniformly at random (EI-RAND). The computer bot gave random values to the queries; to provide some consistency to the bot, $\mathbf{x}_{n+1}$ was selected by maximizing a standard expected improvement function (EI-EI). The results are summarized in Table 1.

Our first observation is that PPBO can distinguish between the choices made by a human and a computer bot. Human choices pinpoint atomic arrangements that are close to nearby local minima (small **N**), while the random bot's choices are far less reasonable and require much subsequent computation to optimise structures. For all human users, the preferred molecular structures were placed somewhat high above the surface, which led to relatively high **E** values. With this query arrangement, it appears the z variable was the most difficult one to estimate visually. Human-preferred molecular orientations were favourable, so the structures were optimised quickly (few **N** steps).

The quality of user preference is best judged by the depth of the nearest energy basin, denoted by **E***. It describes the energy of the structure preferred by a user. Here, there is a marked divide by expertise. The structures refined from the choices of the bot and non-expert are local minima of adsorption, characterised by weak dispersive interactions.

Table 1. Results of the user experiment. The absorption energies were computed by using density functional theory (DFT) methods. The energies represent the absorption energies of the relaxed structures corresponding to the most preferred configurations.

User	Acq. of (ξ, x)	E (eV)	E* (eV)	N
Expert 1	EI-RAND	-0.454	-1.023	62
Expert 1	EI-EXT	-0.371	-1.029	39
Expert 2	EI-RAND	-0.611	-1.007	37
Expert 2	EI-EXT	-0.564	-1.030	45
Non-expert	EI-RAND	-0.481	-0.771	32
Non-expert	EI-EXT	-0.511	-0.762	56
Bot	EI-EI	-0.365	-0.643	70
Bot	EI-EI	-0.612	-0.753	63
Bot	EI-EI	2.231	-0.959	127
Bot	EI-EI	-0.216	-0.783	88

The expert's choice led to two low-energy structure types that compete for the global minimum, and feature strong chemical bonding of the O atom to the Cu surface. Thus, the data (Table 1, column **E***: rows 1-4 versus rows 5-10) supports our hypothesis: the materials science experts do obtain structures associated with lower energy minimum points.

The findings above demonstrate that the PPBO framework was able to encode the expert knowledge described via preferences. However, since there are only 10 samples, further work will be needed to validate the results.

6. Conclusions

In this paper we have introduced a new Bayesian framework, PPBO, for learning user preferences from a special kind of feedback, which we call projective preferential feedback. The feedback is equivalent to a minimizer along a projection. Its form is especially applicable in a *human-in-the-loop* context. We demonstrated this in a user experiment in which the user gives the feedback as an optimal position or orientation of a molecule adsorbing to a surface. PPBO was capable of encoding user preferences in this case.

We demonstrated that PPBO can deal with high-dimensional spaces where existing preferential Bayesian optimization frameworks that are based on pairwise comparisons, such as IBO (Brochu, 2010) or PBO (González et al., 2017), have difficulties to operate. In the numerical experiments, the performance gap between PPBO and PBO was so high that we conclude: whenever a projective preferential query is possible, a PPBO-type of approach is preferable from the optimization perspective. However, we note that it is better to think of PPBO as a complement, not as a substitute for

PBO. In the applications, pairwise comparisons may be preferred, for instance, if they are more convenient to a user.

In summary, if it is possible to query a projective preferential query, then PPBO provides an efficient way for preference learning in high-dimensional problems. In particular, PPBO can be used for efficient expert knowledge elicitation in high-dimensional settings which are important in many fields.

Acknowledgements

References

Blum, V., Gehrke, R., Hanke, F., Havu, P., Havu, V., Ren, X., Reuter, K., and Scheffler, M. Ab initio molecular simulations with numeric atom-centered orbitals. *Computer Physics Communications*, 180(11):2175–2196, 2009.

Brochu, E. *Interactive Bayesian optimization: learning user preferences for graphics and animation*. PhD thesis, University of British Columbia, 2010.

Brochu, E., Brochu, N. D. F., and Ghosh, A. Active preference learning with discrete choice data. In Platt, J. C., Koller, D., Singer, Y., and Roweis, S. T. (eds.), *Advances in Neural Information Processing Systems 20*, pp. 409–416. Curran Associates, Inc., 2008.

Brochu, E., Brochu, T., and de Freitas, N. A Bayesian interactive optimization approach to procedural animation design. In *Proceedings of the 2010 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA 10, pp. 103–112. Eurographics Association, 2010.

Carrier, E. and Weatherford, L. Implementation of MNL choice models in the passenger origin-destination simulator (PODS). *Journal of Revenue and Pricing Management*, 14, 2015.

Chu, W. and Ghahramani, Z. Preference learning with gaussian processes. In *Proceedings of the 22nd International Conference on Machine Learning*, ICML'05, pp. 137–144. ACM, 2005.

Debreu, G. Representation of a preference ordering by a numerical function. *Decision Process*, pp. 159–165, 1954.

González, J., Dai, Z., Damianou, A., and Lawrence, N. D. Preferential Bayesian Optimization. In *Proceedings of*

the 34th International Conference on Machine Learning - Volume 70, ICML'17, pp. 1282–1291, 2017.

Hernández-Lobato, J. M., Hoffman, M. W., and Ghahramani, Z. Predictive entropy search for efficient global optimization of black-box functions. In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'14, pp. 918–926. MIT Press, 2014.

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

Koyama, Y., Sato, I., Sakamoto, D., and Igarashi, T. Sequential line search for efficient visual design optimization by crowds. *ACM Transactions on Graphics*, 36(4), 2017.

McFadden, D. Econometric models of probabilistic choice. *Structural analysis of discrete data with econometric applications*, pp. 198–272, 1981.

McFadden, D. Economic choices. *American Economic Review*, 91(3):351–378, 2001.

Mosteller, F. Remarks on the method of paired comparisons: The least squares solution assuming equal standard deviations and equal correlations. *Psychometrika*, 16(1):3–9, 1951.

Rasmussen, C. E. and Williams, C. K. I. *Gaussian Processes for Machine Learning*. The MIT Press, 2005.

Simon, H. *Reason in human affairs*. Stanford University Press, 1990.

Soufiani, H. A., Parkes, D. C., and Xia, L. Preference elicitation for general random utility models. In *Proceedings of the 29th Conference on Uncertainty in Artificial Intelligence*, 2013.

Spall, J. C. Cyclic seesaw process for optimization and identification. *Journal of Optimization Theory and Applications*, 154(1):187–208, 2012.

Stern, H. A continuum of paired comparisons models. *Biometrika*, 77(2):265–273, 1990.

Thurstone, L. L. A law of comparative judgment. *Psychological Review*, 34(4):273–286, 1927.

Todorović, M., Gutmann, M., Corander, J., and Rinke, P. Bayesian inference of atomistic structure in functional materials. In *npj Computational Materials*, volume 5, pp. 1–7. 2019.