Amortized Active Generation of Pareto Sets

Daniel M. Steinberg, Asiri Wijesinghe, Rafael Oliveira, Piotr Koniusz, Cheng Soon Ong & Edwin V. Bonilla
CSIRO's Data61, Australia
{dan.steinberg, rafael.dossantosdeoliveira, piotr.koniusz}@data61.csiro.au

arXiv: 2510.21052

TL; DR

Task: online multi-objective optimization (MOO), with discrete/mixed design spaces, $\mathbf{x} \in \mathcal{X}$, e.g. *proteins*. Our method (A-GPS) generalizes VSD [2],

- learns a generative model of the Pareto set,
- conditions this model on subjective preferences,
- uses a non-dominance class probability estimator (CPE) to estimate the probability of hypervolume improvement (PHVI) — avoiding expensive hypervolume computation.

We use A-GPS for high dimensional sequence optimization using causal and masked transformers.

Preliminaries

Multi-Objective Optimization: of black-box functions,

$$\max_{\mathbf{x}\in\mathcal{X}} [f_\bullet^1(\mathbf{x}), \ldots, f_\bullet^l(\mathbf{x}), \ldots, f_\bullet^L(\mathbf{x})].$$

But there is no total ordering, or gradients $\nabla_\mathbf{x} f_\bullet^l(\mathbf{x})$. Instead, we wish to find the global Pareto set,

$$\mathcal{S}^*_{\text{Pareto}} = \{\mathbf{x} : \mathbf{x}' \not\succ \mathbf{x},\ \forall \mathbf{x}' \in \mathcal{X}\}, \quad \text{where}$$
$$\mathbf{x}' \succ \mathbf{x} \text{ iff } f_\bullet^l(\mathbf{x}') \geq f_\bullet^l(\mathbf{x}) \ \ \forall l \in \{1,\ldots,L\}, \quad \text{and}$$
$$\exists l \in \{1,\ldots,L\} \text{ such that } f_\bullet^l(\mathbf{x}') > f_\bullet^l(\mathbf{x}).$$

This is the set of $\mathbf{x}$ where we cannot increase one $f_\bullet^l$ without compromising others. Pareto front: $\mathcal{F}^*_{\text{Pareto}} := \{\mathbf{f}_\bullet(\mathbf{x}) : \forall \mathbf{x} \in \mathcal{S}^*_{\text{Pareto}}\}$.

Active Generation: reframes online black-box optimization as sequential learning of a generative model, $q_\phi(\mathbf{x}) \approx p(\mathbf{x}|z)$, conditioned by a CPE, as in *variational search distributions* (VSD) [2],

$$\pi^z_\theta(\mathbf{x}) \approx p(z = 1|\mathbf{x}),$$

where $z = \mathbb{1}[\mathbf{x} \in \mathcal{S}]$ indicates membership in some desired set, $\mathcal{S}$. We use the evidence lower bound to learn the generative model,

$$\mathcal{L}_{\text{ELBO}}(\phi, \theta) = \mathbb{E}_{q_\phi(\mathbf{x})}[\log \pi^z_\theta(\mathbf{x})] - \mathbb{D}_{\text{KL}}[q_\phi(\mathbf{x}) \| p(\mathbf{x}|\mathcal{D}_0)],$$

where $p(\mathbf{x}|\mathcal{D}_0)$ is a prior over the design space. Then each round $t \in \{1, \ldots, T\}$ using $\mathcal{D}^z_N = \{(\mathbf{x}_n, z_n)\}^N_{n=1}$ we,

$$\theta^*_t \leftarrow \underset{\theta}{\text{argmin}}\, \mathcal{L}_{\text{CPE}}(\theta, \mathcal{D}^z_N), \quad \phi^*_t \leftarrow \underset{\phi}{\text{argmax}}\, \mathcal{L}_{\text{ELBO}}(\phi, \theta^*_t),$$

and sample new $\mathbf{x}^{(b)} \sim q_{\phi^*_t}(\mathbf{x})$ for evaluation by $f_\bullet(\mathbf{x})$ and labelling.

Non-Dominance Classification ⇔ PHVI

Given an observed (empirical) Pareto set, $\mathcal{S}^t_{\text{Pareto}}$, we define a labeling function $z(\mathbf{x}) := \mathbb{1}[\mathbf{x} \in \text{Pareto}(\mathcal{S}^t_{\text{Pareto}} \cup \{\mathbf{x}\})]$, where $\text{Pareto}(\mathcal{S})$ is the Pareto subset of an arbitrary set $\mathcal{S} \subset \mathcal{X}$. Then,

Theorem 1 (Equivalence of Indicators). *For every $\mathbf{x} \notin \mathcal{S}^t_{\text{Pareto}}$, the hypervolume improvement (HVI) indicator is equivalent to a non-dominance indicator,*

$$\mathbb{1}[\text{HVI}(\mathbf{x}) > 0] = z(\mathbf{x}).$$

Corollary 2 (Non-Dominance CPE estimates PHVI). *Following from Theorem 1,*

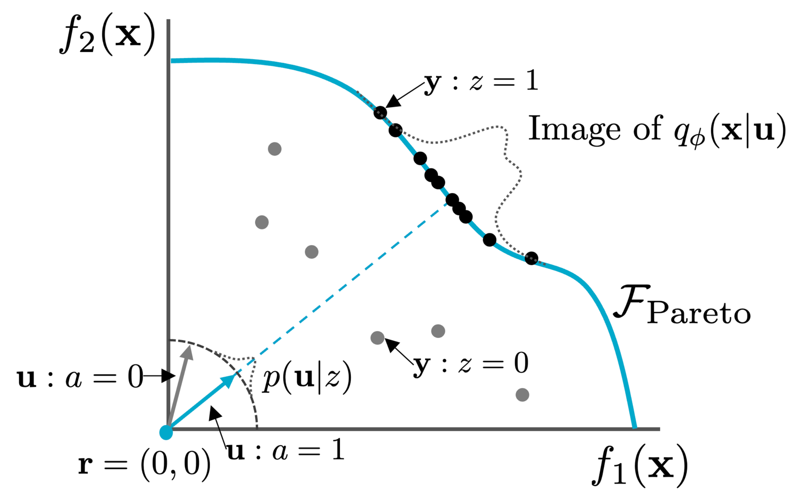
$$\mathbb{P}(z(\mathbf{x}) = 1|\mathbf{x}) = \mathbb{P}(\text{HVI}(\mathbf{x}) > 0|\mathbf{x}) := \text{PHVI}(\mathbf{x}), \quad \forall \mathbf{x} \notin \mathcal{S}^t_{\text{Pareto}},$$

as the events are equivalent. Thus, a CPE trained on z, using a proper loss, is predicting PHVI. □

Preference Direction Vectors, $\mathbf{u}$

Our objective is to learn a conditional generative model, $q_\phi(\mathbf{x}|\mathbf{u})$, where $\mathbf{u}$ are *preference direction vectors*, $\mathbf{u} \in \{\mathbf{u} \in \mathbb{R}^L : \|\mathbf{u}\|_2 = 1\}$.

We use these vectors, $\mathbf{u}$, instead of scalarization weights to indicate the region of the Pareto front to generate designs for. We base their training data on the noisy observations of the black-box objectives, $\mathbf{y}$,



$$\mathbf{u}_n = g(\mathbf{y}_n) := \frac{\mathbf{y}_n - \mathbf{r}}{\|\mathbf{y}_n - \mathbf{r}\|_2}.$$

An alignment CPE is trained using contrastive data with labels, a_n, to reward learning the conditional relationship between $(\mathbf{x}_n, \mathbf{u}_n)$.

Amortized ELBO

A-GPS minimizes the *expected* KL divergence each round, t,

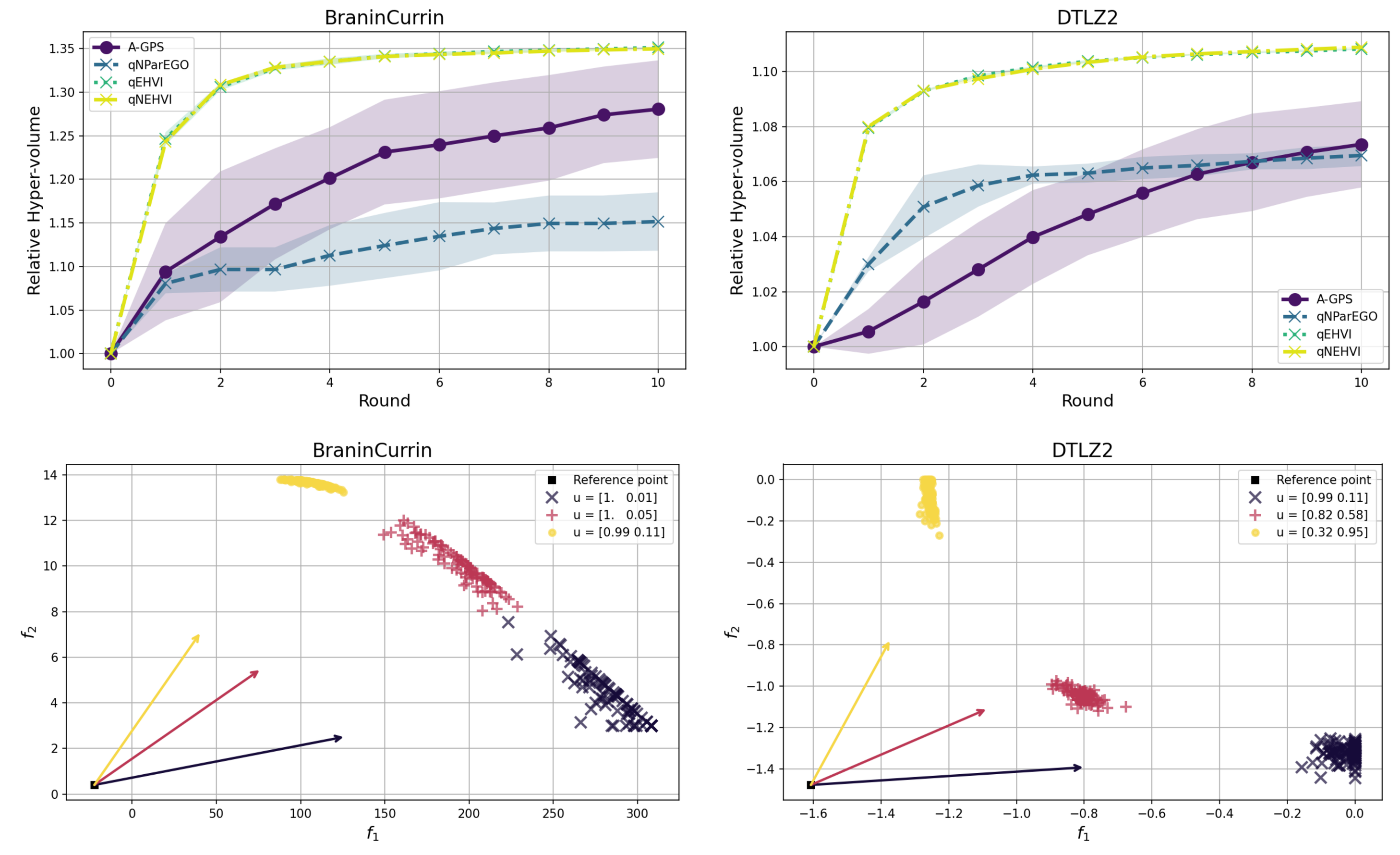
$$\phi^*_t = \underset{\phi}{\text{argmin}}\, \mathbb{E}_{p(\mathbf{u}|z)}[\mathbb{D}_{\text{KL}}[q_\phi(\mathbf{x}|\mathbf{u}) \| p(\mathbf{x}|\mathbf{u}, z, a)]] = \underset{\phi}{\text{argmax}}\, \mathcal{L}_{\text{A-ELBO}}(\phi).$$

The expectation over $p(\mathbf{u}|z)$ leads to learning an *amortized* conditional generative model, $q_\phi(\mathbf{x}|\mathbf{u})$, over seen preference directions,

$$\mathcal{L}_{\text{A-ELBO}}(\phi) =$$
$$\mathbb{E}_{\underbrace{p(\mathbf{u}|z)}_{\text{Direction dist.}}} \left[\mathbb{E}_{q_\phi(\mathbf{x}|\mathbf{u})} \left[\underbrace{\log \pi^z_\theta(\mathbf{x}, \mathbf{u})}_{\text{Pareto CPE}} + \underbrace{\log \pi^a_\psi(\mathbf{x}, \mathbf{u})}_{\text{Align. CPE}} \right] - \beta \mathbb{D}_{\text{KL}} \left[q_\phi(\mathbf{x}|\mathbf{u}) \| \underbrace{p(\mathbf{x}|\mathcal{D}_0)}_{\text{prior}} \right] \right].$$

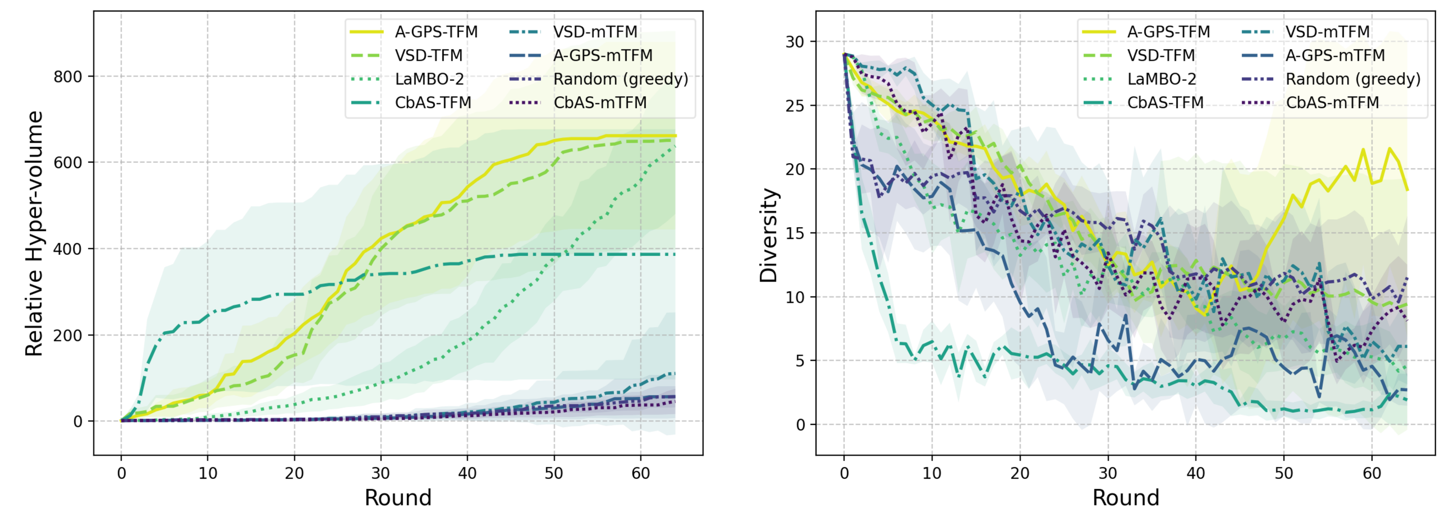
Experiments

Synthetic test functions ($\mathcal{X} = \mathbb{R}^D$): a demonstration of the amortized generative model.



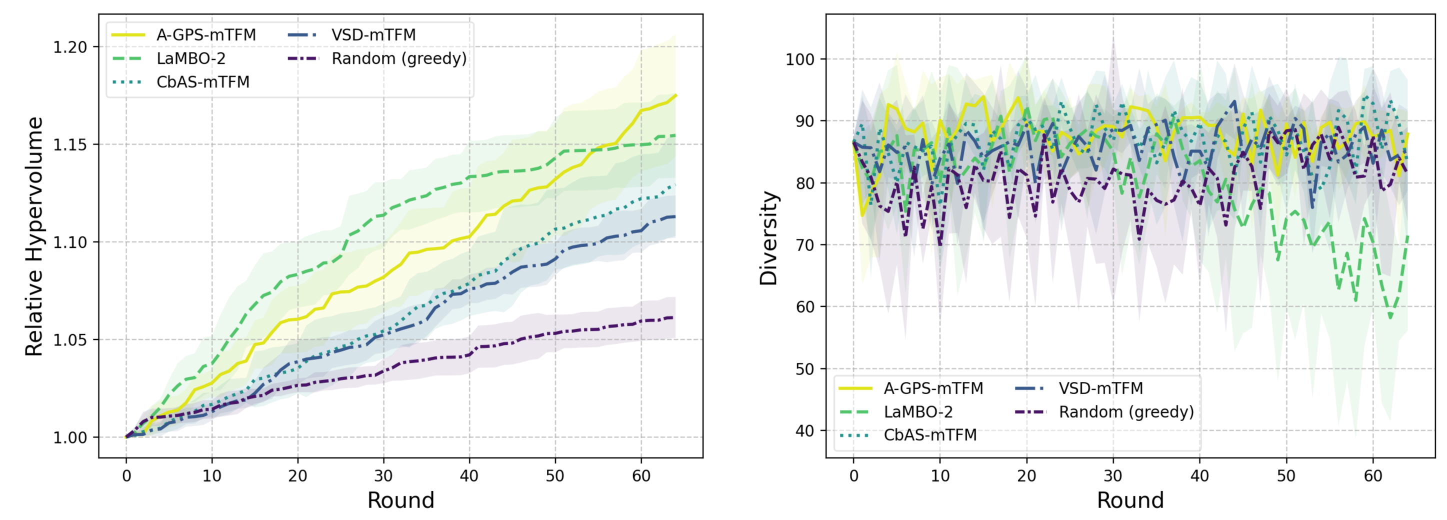
(a) Negative Branin-Currin (b) DTLZ2

Bi-gram counts ($\mathcal{X} = \mathcal{V}^{32}$): maximize the occurrence of `AC', `VC' and `CA' bigrams in a sequence [1].



(a) Relative hypervolume improvement (b) Sequence diversity

Stability vs. SASA ($\mathcal{X} = \mathcal{V}^{200+}$): maximise the folding stability and solvent accessible surface area of 6 red fluorescent proteins [1].



(a) Relative hypervolume improvement (b) Sequence diversity

References

[1] Samuel Stanton, Wesley Maddox, Nate Gruver, Phillip Maffettone, Emily Delaney, Peyton Greenside, and Andrew Gordon Wilson. Accelerating Bayesian optimization for biological sequence design with denoising autoencoders. In *International Conference on Machine Learning*, pages 20459--20478. PMLR, 2022.

[2] Daniel M Steinberg, Rafael Oliveira, Cheng Soon Ong, and Edwin V Bonilla. Variational search distributions. In *The Thirteenth International Conference on Learning Representations (ICLR)*, 2025.