

BEACON: A Bayesian Optimization Inspired Strategy for Efficient Novelty Search

Tang, Wei-Ting; Chakrabarty, Ankush; Paulson, Joel A.

TR2026-151 October 07, 2026

Abstract

Novelty search (NS) refers to a class of exploration algorithms that automatically uncover diverse system behaviors through simulations or experiments. Uncovering diversity is a key aspect of engineering design problems with connections to material and drug discovery, neural architecture search, reinforcement learning, and robot navigation. Since the relationship between the inputs and behaviors (outputs) of modern engineering systems not always available or easily represented in closed analytical form, novelty search must be able to handle model opacity. For systems whose behaviors are expensive to simulate or evaluate, we propose a sample-efficient NS method inspired by Bayesian optimization principles. This involves modeling the input-to-behavior mapping with multi-output Gaussian processes (MOGP) and selecting inputs to evaluate that maximize a novelty metric while balancing the exploration-exploitation trade-off. By leveraging advances in efficient posterior sampling and high-dimensional Gaussian process modeling, we discuss how our approach can be made scalable with respect to both the amount of data and number of inputs. We demonstrate the potential of our approach on several well-studied benchmark problems and multiple real-world examples. We show that BEACON comprehensively outperforms existing baselines by finding substantially larger sets of diverse behaviors under limited sampling budgets.

Digital Discovery 2026

© 2026 MERL. This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of Mitsubishi Electric Research Laboratories, Inc.; an acknowledgment of the authors and individual contributions to the work; and all applicable portions of the copyright notice. Copying, reproduction, or republishing for any other purpose shall require a license with payment of fee to Mitsubishi Electric Research Laboratories, Inc. All rights reserved.

Cite this: DOI: 00.0000/xxxxxxxxxx

BEACON: A Bayesian Optimization Inspired Strategy for Efficient Novelty Search

Wei-Ting Tang,^a Ankush Chakrabarty,^b and Joel A. Paulson^{*a}

Received Date

Accepted Date

DOI: 00.0000/xxxxxxxxxx

Novelty search (NS) refers to a class of exploration algorithms that automatically uncover diverse system behaviors through simulations or experiments. Uncovering diversity is a key aspect of engineering design problems with connections to material and drug discovery, neural architecture search, reinforcement learning, and robot navigation. Since the relationship between the inputs and behaviors (outputs) of modern engineering systems not always available or easily represented in closed analytical form, novelty search must be able to handle model opacity. For systems whose behaviors are expensive to simulate or evaluate, we propose a sample-efficient NS method inspired by Bayesian optimization principles. This involves modeling the input-to-behavior mapping with multi-output Gaussian processes (MOGP) and selecting inputs to evaluate that maximize a novelty metric while balancing the exploration-exploitation trade-off. By leveraging advances in efficient posterior sampling and high-dimensional Gaussian process modeling, we discuss how our approach can be made scalable with respect to both the amount of data and number of inputs. We demonstrate the potential of our approach on several well-studied benchmark problems and multiple real-world examples. We show that BEACON comprehensively outperforms existing baselines by finding substantially larger sets of diverse behaviors under limited sampling budgets.

1 Introduction

With the increased modeling complexity of modern engineering systems, it is often easier to represent the system input-output behavior via black-box functions f whose arguments are user-defined system inputs $\mathbf{x} \in \mathcal{X}$ that yield outcomes (or behaviors) via the map $f: \mathcal{X} \mapsto \mathcal{O}$ in some outcome space $\mathcal{O}$; we will define notation more rigorously later. Derivative-free search methods ranging from simplices (e.g., Nelder-Mead), meta-heuristic population-based algorithms (e.g. particle swarms), and surrogate-based optimization (e.g. Bayesian optimization) provide powerful tools to *optimize* the quality of the outcomes based on a pre-defined criteria of optimality; c.f.^{1–4} of an overview of these classes of algorithms. If instead of seeking the highest quality outcomes we are interested in understanding the range of outcomes that can be rendered possible by the black-box system, then a different approach is required.

Let us consider the following motivating example. Suppose f denotes the solution map of a forced Duffing oscillator and $\mathbf{x}$ are the parameters of the oscillator and the external forcing function. We know that the solutions of the Duffing oscillator are either chaotic, periodic, quasi-periodic, or asymptotically convergent to

a fixed point depending on the values of $\mathbf{x}$ ⁵. If our objective was to numerically discover the range of behaviors of f with no prior knowledge of the fact that these behaviors are possible, we would struggle to put it into an optimization framework. First, because there is no clear objective function that will induce all of these behaviors simultaneously, e.g. minimizing or maximizing Lyapunov exponents will only yield convergence to a fixed point or chaotic oscillations, but not the intermediate periodic behaviors. Second, such an objective landscape would be fraught with local optima, and extremely hard to navigate without exorbitant function evaluations. A more natural solution to uncovering diverse behaviors is possible by novelty search, wherein the focus is shifted to pure exploration.

Instead of optimizing a pre-defined objective, novelty search aims to discover a new behavior of f given all past observed behaviors by modifying $\mathbf{x}$ and defining (actually, optimizing) a ‘novelty metric’^{6,7}. This is especially useful for automated scientific discovery such as for learning new protocell behaviors⁸, novel material design⁹, and scenario generation to inform policy search in reinforcement learning¹⁰. The NS perspective has also been found to be very useful in so-called “deceptive problems” wherein the optimization path taken towards a specific objective can often get stuck in suboptimal solutions. Since the mapping between inputs and system outcomes is generally unknown, established NS methods rely on meta-heuristics, such as evolutionary algorithms, to select new evaluation points. Genetic Algo-

^a University of Wisconsin-Madison, Department of Chemical and Biological Engineering, Madison, WI, USA.

^b Insulet Corporation, Acton, MA, USA.

^{*} Correspondence: joel.paulson@wisc.edu

gorithms (GAs) have frequently been employed in this context. For instance, the authors of¹¹ proposed diversity-promoting mechanisms to address deceptive optimization landscapes. Similarly, swarm intelligence techniques such as Particle Swarm Optimization (PSO) have been adapted for novelty search, as demonstrated by¹², who explored behavioral diversity for robotics applications. Furthermore, the MAP-Elites algorithm introduced by¹³ integrates population-based exploration to balance diversity and performance metrics. However, these methods are known to be *sample-inefficient*, requiring a very large number of evaluations, which limits their use on expensive-to-evaluate systems.

We take inspiration from the fundamental principles of Bayesian optimization (BO) to propose a Bayesian novelty search algorithm that exhibits high sample-efficiency. To this end, we construct Gaussian process (GP) surrogate models¹⁴ to learn the unknown system’s input-output mapping. Unlike BO, our surrogate models do not approximate optimization objectives but instead represent outcomes that a user wishes to explore. The outcome space can be freely defined by the user, meaning it can be a multi-element vector encompassing all the user’s desired features of exploration. For example, in drug discovery, this vector could comprise the efficacy, size, synthesizability, and stability of a candidate drug compound.

Our *main contributions* are summarized as follows: (i) We introduce a new NS algorithm, referred to as **BEACON** (Bayesian Exploration Algorithm for outCOme Novelty), for noisy, expensive black-box systems that is designed to discover unseen behaviors of the system using a minimal number of evaluations; (ii) we propose a novel Thompson sampling-based acquisition function for NS applications that, to our knowledge, is the first to address the exploration-exploitation trade-off, handle stochastic observations noise, and is suitable for gradient-based acquisition function optimization; (iii) we improve the scalability of the core BEACON algorithm to high-dimensional problems including a general approach that uses fully Bayesian sparsity-inducing function priors and a specialized approach for computational chemistry applications; and (iv) we conduct extensive experiments on several synthetic and real-world problems that demonstrate the substantial benefits BEACON can achieve in terms of identified behaviors over relevant existing NS works. These include a first-time application of NS to discovery of metal organic frameworks that are useful materials in clean energy applications, and a challenging molecular discovery benchmark problems with 2133 input dimensions.

The rest of the paper is organized as follows. In Section 2, we define the black-box NS problem considered in this work. In Section 3, we present some necessary background, introducing BO terminology and multi-output GP regression. Notation is also fixed in this section. In Section 4, we present the proposed BEACON algorithm, along with pseudocode. We also provide additional customization for high-dimensional problems, and for problems where domain-relevant behavior constraints exist. Our validation of BEACON and baseline comparisons on benchmark problems is presented in Section 5, along with multiple real-world applications that reflect the potential of the base BEACON algorithm.

2 Problem Description

We consider a vector-valued black-box function $\mathbf{f}: \mathcal{X} \rightarrow \mathcal{O}$ with $\mathbf{f} = (f^{(1)}, \dots, f^{(n)})$ that maps an input space $\mathcal{X} \subset \mathbb{R}^d$ to a multi-dimensional outcome space $\mathcal{O} \subset \mathbb{R}^n$. Although $\mathcal{O}$ is continuous, we assume nearby outcomes correspond to similar system behaviors. We discretize this space using an ε -covering $\mathcal{B} \subset \mathbb{R}^n$, such that for every $\mathbf{y} \in \mathcal{O}$, there exists $\mathbf{y}' \in \mathcal{B}$ with $\|\mathbf{y} - \mathbf{y}'\| \leq \varepsilon$.

The goal is to discover inputs whose outcomes cover as many distinct behaviors in $\mathcal{B}$ as possible. Since $\mathbf{f}$ is expensive to evaluate, we operate under a finite query budget T , evaluating $\mathbf{f}$ sequentially over $t = 1, \dots, T$. Each query point $\mathbf{x}_t \in \mathcal{X}$ yields a noisy observation $\mathbf{y}_t = \mathbf{f}(\mathbf{x}_t) + \boldsymbol{\eta}_t$, where $\boldsymbol{\eta}_t \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_n)$ denotes independent Gaussian noise. Let $\boldsymbol{\phi}: \mathcal{O} \rightarrow \mathcal{B}$ denote the projection from the continuous outcome to its discrete behavior. We define the behavior gap at iteration t as

$$\text{BG}_t = 1 - (|\{\boldsymbol{\phi}(\mathbf{f}(\mathbf{x}_1)), \dots, \boldsymbol{\phi}(\mathbf{f}(\mathbf{x}_t))\}| / |\mathcal{B}|),$$

which quantifies the fraction of behaviors not yet observed. The objective is to minimize the cumulative behavior gap $\sum_{t=1}^T \text{BG}_t$, or equivalently, to maximize behavioral diversity across evaluations.

We model $\mathbf{f}$ using a multi-output Gaussian process (MOGP) prior¹⁵, assuming it lies in a smooth function class such as a Reproducing Kernel Hilbert Space (RKHS)¹⁶. The MOGP yields a posterior distribution $\mathbb{P}(\mathbf{f} | \mathcal{D}_t)$ over the function $\mathbf{f}$ given observations $\mathcal{D}_t = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^t$, with closed-form expressions for the posterior mean and covariance, as detailed in the next section.

3 Background

In this section, we provide background on multi-output Gaussian processes and Bayesian optimization, which form the core techniques underlying our proposed BEACON algorithm.

3.1 Multi-Output Gaussian Process Surrogate

Gaussian processes (GPs) offer a flexible, non-parametric way to model black-box functions¹⁴. A GP prior is fully specified by a prior mean function μ and covariance (or kernel) function κ . To (potentially) capture correlations among the n outputs of $\mathbf{f}$, we adopt the *latent-input* construction: define $h(\mathbf{x}, j) = f^{(j)}(\mathbf{x})$ for $j \in \mathcal{J} = \{1, \dots, n\}$ and place a scalar GP prior $h \sim \mathcal{GP}(\mu, \kappa)$ over the extended space $\mathcal{X} \times \mathcal{J}$, following, e.g.,¹⁷. Because observation noise is assumed i.i.d. Gaussian, the posterior after N observations $\mathcal{A} = \{(\mathbf{x}_i, j_i, y_i)\}_{i=1}^N$, denoted by $h|\mathcal{A} \sim \mathcal{GP}(\mu_{\mathcal{A}}, \kappa_{\mathcal{A}})$, remains a GP with analytically tractable mean $\mu_{\mathcal{A}}$ and covariance $\kappa_{\mathcal{A}}$ functions. These closed-form expressions are shown in below.

$$\begin{aligned} \mu_{\mathcal{A}}(\mathbf{x}, j) &= \mu(\mathbf{x}, j) + \boldsymbol{\kappa}_{\mathcal{A}}^{\top}(\mathbf{x}, j) \left(\mathbf{K}_{\mathcal{A}} + \sigma^2 \mathbf{I}_N \right)^{-1} (\mathbf{y} - \boldsymbol{\mu}_{\mathcal{A}}), \end{aligned} \quad (1a)$$

$$\begin{aligned} \kappa_{\mathcal{A}}((\mathbf{x}, j), (\mathbf{x}', j')) &= \kappa((\mathbf{x}, j), (\mathbf{x}', j')) - \boldsymbol{\kappa}_{\mathcal{A}}^{\top}(\mathbf{x}, j) \left(\mathbf{K}_{\mathcal{A}} + \sigma^2 \mathbf{I} \right)^{-1} \boldsymbol{\kappa}_{\mathcal{A}}(\mathbf{x}', j'), \end{aligned} \quad (1b)$$

where $\mathbf{y} = [y_1, \dots, y_N]^{\top}$, $\boldsymbol{\mu}_{\mathcal{A}} = [\mu(x_1), \dots, \mu(x_N)]^{\top}$, $\boldsymbol{\kappa}_{\mathcal{A}}(\mathbf{x}, j) \in \mathbb{R}^N$ is the vector of covariance values between the test input $(\mathbf{x}, j)$

and the observed inputs in $\mathcal{A}$, and $\mathbf{K}_{\mathcal{A}} \in \mathbb{R}^{N \times N}$ is the covariance matrix between all observed inputs in $\mathcal{A}$.

This construction has two advantages: (i) information shared across outputs improves sample efficiency and (ii) any scalable single-output GP approximation (e.g., inducing points or stochastic variational inference) can be easily applied without modification.

3.2 Bayesian Optimization

BO is a sequential decision-making strategy to efficiently find a global optimum $\mathbf{x}^* = \operatorname{argmin}_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x})$ for an expensive-to-evaluate black-box function f , where $\mathcal{X}$ represents the admissible search space. Classical BO performs the following two steps *ad infinitum*, or until a predefined sampling budget is exhausted: (1) Constructing uncertainty-aware (probabilistic) surrogate models based on previously sampled data pairs $\{(\mathbf{x}, y)\}$ where y is a possibly noisy observation of $f(\mathbf{x})$ and (2) Optimizing an acquisition function (AF) that provides some measure of how valuable a future point is to sample by balancing the exploration-exploitation trade-off during the BO search process. Common acquisition functions include upper confidence bound (UCB)¹⁸, expected improvement (EI)¹⁹, and Thompson sampling (TS)²⁰. TS is a randomized approach for sequential decision making under uncertainty and has been shown to provide strong empirical performance in BO²¹. TS works by selecting $\operatorname{argmin}_{\mathbf{x} \in \mathcal{X}} g(\mathbf{x})$ as the next sample location where g is a posterior realization of the probabilistic surrogate model.

4 Proposed BEACON Method

In this section, we introduce the proposed BEACON algorithm. We begin by presenting the newly proposed NS acquisition function. We then provide an overview and pseudocode for the BEACON algorithm, followed by details of the optimization procedure for the proposed acquisition function. Finally, we extend BEACON to high-dimensional settings and problems with behavior constraints.

4.1 Thompson Sampling-Based Acquisition for Novelty Search

Our goal is to discover novel system behaviors by exploring unobserved regions of the outcome space $\mathcal{O}$. A common strategy in the NS literature is to evaluate how far a predicted outcome $\mathbf{f}(\mathbf{x})$ lies from previously observed outcomes. One such novelty metric is the average distance to the k nearest neighbors among past evaluations:

$$\rho(\mathbf{x}|\mathcal{D}) = \frac{1}{k} \sum_{i=1}^k \operatorname{dist}(\mathbf{f}(\mathbf{x}), \mathbf{y}_i^*), \quad (2)$$

where $\{\mathbf{y}_1^*, \dots, \mathbf{y}_k^*\} \subset \{\mathbf{y}_1, \dots, \mathbf{y}_N\}$ are the k nearest outcomes to $\mathbf{f}(\mathbf{x})$ in the dataset $\mathcal{D} = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^N$, and $\operatorname{dist}(\cdot)$ is a user-defined distance metric over $\mathcal{O}$.

Although intuitive, this metric (and other related metrics that are accumulation operators over a finite set of points) poses two challenges in our setting. First, $\mathbf{f}$ is a black-box function, so $\rho(\mathbf{x}|\mathcal{D})$ is not available until after an expensive query. Second,

Algorithm 1 BEACON

- 1: **Input:** Outcome function $\mathbf{f}$, domain $\mathcal{X}$, initial dataset $\mathcal{D}_0$, total budget T , MOGP prior $\mathcal{M} \mathcal{O} \mathcal{G} \mathcal{P}(\boldsymbol{\mu}, \boldsymbol{\kappa})$, neighbor count k , and outcome distance metric $\operatorname{dist}(\cdot)$
 - 2: **for** $t = 1$ to T **do**
 - 3: Fit MOGP posterior: $\mathcal{M} \mathcal{O} \mathcal{G} \mathcal{P}_{t-1} \leftarrow \mathbf{f} | \mathcal{D}_{t-1}$
 - 4: Sample $\mathbf{g} \sim \mathcal{M} \mathcal{O} \mathcal{G} \mathcal{P}_{t-1}$
 - 5: $\mathbf{x}_t \leftarrow \operatorname{argmax}_{\mathbf{x} \in \mathcal{X}} \alpha_{\text{NS}}(\mathbf{x} | \mathbf{g}, \mathcal{D}_{t-1})$
 - 6: Observe outcome $\mathbf{y}_t = \mathbf{f}(\mathbf{x}_t) + \boldsymbol{\eta}_t$
 - 7: Update $\mathcal{D}_t \leftarrow \mathcal{D}_{t-1} \cup \{(\mathbf{x}_t, \mathbf{y}_t)\}$
 - 8: **end for**
 - 9: **Return:** Final dataset $\mathcal{D}_T$
-

observational noise can degrade NS performance by misclassifying behaviors – for example, due to a realization of $\boldsymbol{\eta}$, it may hold that $\boldsymbol{\phi}(\mathbf{f}(\mathbf{x}) + \boldsymbol{\eta}) \neq \boldsymbol{\phi}(\mathbf{f}(\mathbf{x}))$.

To address these issues, we replace the true $\mathbf{f}$ with a MOGP posterior surrogate $\mathbf{f}|\mathcal{D} \sim \mathcal{M} \mathcal{O} \mathcal{G} \mathcal{P}(\boldsymbol{\mu}_{\mathcal{D}}, \boldsymbol{\kappa}_{\mathcal{D}})$. The MOGP serves two key purposes: (i) it enables us to filter observation noise by replacing measured data with a surrogate prediction and (ii) it provides a way for us to fantasize future outcome realizations while accounting for uncertainty. The latter step is accomplished through Thompson Sampling (TS), which is a classical strategy for decision-making under uncertainty²⁰. Our acquisition function can thus be expressed as follows:

$$\alpha_{\text{NS}}(\mathbf{x} | \mathbf{g}, \mathcal{D}) = \frac{1}{k} \sum_{i=1}^k \operatorname{dist}(\mathbf{g}(\mathbf{x}), \boldsymbol{\mu}_{\mathcal{D}}(\mathbf{x}_i^*)), \quad (3)$$

where $\mathbf{g}(\mathbf{x}) \sim \mathbf{f}|\mathcal{D}$ is a posterior function sample and $\{\mathbf{x}_1^*, \dots, \mathbf{x}_k^*\} \subset \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ are the set of k closest outcomes to $\mathbf{g}(\mathbf{x})$ in terms of distance to the posterior mean predictions $\{\boldsymbol{\mu}_{\mathcal{D}}(\mathbf{x}_1), \dots, \boldsymbol{\mu}_{\mathcal{D}}(\mathbf{x}_N)\}$. This acquisition function promotes exploration of unvisited outcome regions in the presence of model uncertainty. To generate $\mathbf{g}$ efficiently, we employ the decoupled sampling strategy of²², which combines weight- and function-space views to produce high-accuracy differentiable posterior samples. We describe how to efficiently and differentially compute the k -nearest neighbors needed for this acquisition in Section 4.3.

4.2 Overview of the BEACON algorithm

We now summarize the pseudocode for BEACON in Algorithm 1, which combines MOGP surrogate modeling with a TS-based acquisition function to promote exploration of novel behaviors. At each iteration, BEACON fits a MOGP model to all available data, draws a posterior sample, and selects a new query point that maximizes predicted novelty in the outcome space. This process is then repeated until the maximum number of iterations has been reached (or some other simple stopping criteria is employed). A visual illustration of this process is shown in Figure 1. The pseudocode presents BEACON in a standard sequential setting for clarity. However, it can be readily extended to parallel or asynchronous evaluation using posterior sampling strategies from prior work on parallel BO²¹.

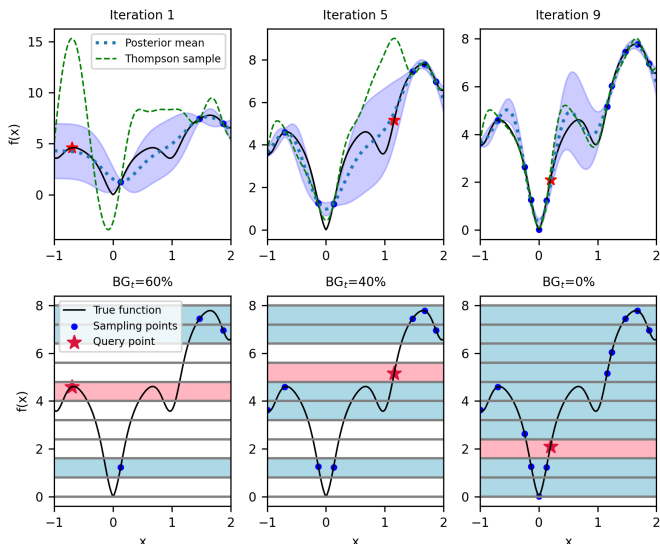


Fig. 1 Visual illustration of BEACON applied to a 1D test problem. Top: the true function (black), GP posterior mean (blue dashed), and a TS (green dashed) at iterations 1, 5, and 9. Shaded regions show posterior uncertainty, and the red star marks the selected query point. Bottom: behavioral bins defined over the outcome space. Blue bands denote previously observed outcomes; the red band shows the predicted outcome bin of the next query. As BEACON progresses, it efficiently reduces the behavior gap BG_t by identifying unexplored regions.

4.3 Gradient-Based Acquisition Function Optimization

Efficiently maximizing the BEACON acquisition function α_{NS} is essential to practical performance (Line 5 of Algorithm 1). To enable scalable search over $\mathbf{x}$, we leverage gradient-based optimizers such as L-BFGS-B²³ with multiple random restarts. Although nearest-neighbor novelty metrics are inherently discrete, we can reformulate α_{NS} in a way that enables differentiation:

$$\alpha_{NS}(\mathbf{x}|\mathbf{g}, \mathcal{D}) = \frac{1}{k} \mathbf{e}_k^\top \text{sort}(\{\text{dist}(\mathbf{g}(\mathbf{x}), \boldsymbol{\mu}_{\mathcal{D}}(\mathbf{x}))\}_{\mathbf{x} \in \mathcal{D}}), \quad (4)$$

where $\mathbf{e}_k$ is a vector whose first k entries are equal to 1 and the remaining entries are equal to 0 and $\text{sort}(\cdot)$ denotes the descending sort operator. As discussed in²⁴, the standard sort operator is continuous and almost everywhere differentiable (with non-zero gradients), allowing us to apply gradient-based optimization effectively.

The computational cost of this formulation scales with N , which can be large in practice. To mitigate this, we exploit a key property of BEACON: only distinct behavioral outcomes matter. Instead of computing distances to all outcomes in $\mathcal{D}$, we restrict to a deduplicated subset with unique behavior bins. This reduces the complexity of (4) from $O(N)$ to $O(|\mathcal{B}|)$, which is constant across iterations. In Appendix F, we show how combining this idea with sparse GP approximations enables BEACON to easily scale to evaluation budgets in the tens of thousands.

4.4 Scaling BEACON to High-Dimensional Problems

One of the advantages of BEACON’s design is its compatibility with modern BO advances, particularly for high-dimensional in-

put spaces. When the input dimension $d \gg 10$, surrogate models typically require structural priors to generalize well. In this section, we describe two complementary strategies that enable BEACON to operate effectively in such settings. Additional extensions, such as accounting for uncertainty in the hyperparameters²⁵ or black-box output constraints²⁶, can also be incorporated with minor changes.

4.4.1 SAAS Priors.

The sparse axis-aligned subspace (SAAS) prior²⁷ is a fully Bayesian approach that introduces adaptive sparsity into GP models. SAAS places a hierarchical prior over the lengthscales of each input dimension, favoring explanations that involve a small number of relevant inputs. This formulation naturally supports input relevance discovery, which is especially important when only a few dimensions influence the outcome behavior. Due to its generality and compatibility with automatic relevance determination, SAAS is well suited as a default surrogate prior in high-dimensional BO. We use SAAS to scale BEACON on a challenging molecular discovery task with 125 input dimensions in Section 5.3.

4.4.2 GAUCHE for Discrete Molecular Inputs.

While SAAS is effective when input relevance is sparse and independent, it may underperform in domains where inputs exhibit strong correlations (e.g., spatially structured data) or belong to non-Euclidean spaces. In these cases, domain-specific surrogate models are more appropriate. The GAUCHE framework²⁸ offers a principled GP modeling interface for structured chemical/biological inputs, including SMILES strings, proteins, and reaction graphs. Since GAUCHE is directly built on top of standard GP principles, BEACON can directly use GAUCHE models as drop-in surrogates. We demonstrate this extension in our solubility case study (Section 5.3).

4.5 Incorporating user information in BEACON

In many practical settings, users bring domain knowledge or preferences about the behaviors they seek to discover. For example, a scientist may wish to focus exploration on materials with specific target properties or avoid re-evaluating regions that have already yielded satisfactory solutions. To accommodate such guidance, we introduce “UG-BEACON”, which is a user-guided extension of BEACON that incorporates behavior-level constraints to focus exploration. UG-BEACON modifies the acquisition strategy to avoid selecting candidates whose predicted outcomes fall within behavior regions already explored – this can be straightforwardly accomplished by enforcing a constraint on the prediction during the acquisition optimization.

UG-BEACON builds directly on the BEACON algorithm with a simple but effective modification to incorporate user-defined behavior-level constraints. Specifically, recall that BEACON selects the next evaluation point by maximizing the novelty-based acquisition function:

$$\mathbf{x}^* = \underset{\mathbf{x} \in \mathcal{X}}{\operatorname{argmax}} \alpha_{NS}(\mathbf{x}|\mathbf{g}, \mathcal{D}),$$

where $\alpha_{\text{NS}}(\mathbf{x}|\mathbf{g}, \mathcal{D})$ denotes the novelty score computed using a posterior sample $\mathbf{g}$ of the outcome function and currently available data $\mathcal{D}$.

UG-BEACON introduces a filtering step to exclude candidates whose predicted outcomes fall within previously explored or user-restricted regions of the behavior space. Let $\mathcal{G} = \{G_1, \dots, G_M\}$ denote a user-defined partition of the outcome space into M non-overlapping regions. This partition can reflect application-specific preferences (e.g., fine-grained in important regions or coarse in irrelevant ones). Let $\mathcal{G}_{\text{visited}} \subseteq \mathcal{G}$ denote the set of regions containing outcomes previously observed during the search (identified by the user). The feasible set of candidates is then given by:

$$\mathcal{X}_{\text{feas}} = \left\{ \mathbf{x} \in \mathcal{X} \mid \mathbf{g}(\mathbf{x}) \notin \bigcup_{G \in \mathcal{G}_{\text{visited}}} G \right\},$$

and the next evaluation point is chosen as:

$$\mathbf{x}^* = \underset{\mathbf{x}}{\operatorname{argmax}} \alpha_{\text{NS}}(\mathbf{x}|\mathbf{g}, \mathcal{D}) \text{ subject to } \mathbf{x} \in \mathcal{X}_{\text{feas}}.$$

This mechanism enables UG-BEACON to avoid redundant sampling and to steer exploration toward novel or user-relevant behavior regions. All other components of the BEACON algorithm, including posterior sampling, surrogate modeling, and acquisition computation, remain unchanged.

5 Numerical Experiments

In this section, we compare BEACON against established NS algorithms and their variants including the standard evolutionary-based method (NS-EA)⁷, a distance-enhanced evolutionary method (NS-DEA)²⁹, and a feature space-based method (NS-FS). We also compare with the performance of three other algorithms: one that chooses points uniformly at random over $\mathcal{X}$ (RS); one that chooses points using low-discrepancy quasi-random Sobol samples over $\mathcal{X}$ (Sobol); and a maximum variance active learning strategy (MaxVar). Full implementation details for BEACON and the benchmark algorithms are provided in Section 1 of the supplementary information (SI).

We evaluate performance on 10 synthetic and 8 real-world problems, which are described below or in the Appendices. In all problems, we generate an initial set of 10 points uniformly at random over $\mathcal{X}$ as a first stage to seed the algorithms. Then, we use the algorithms to select an additional 80 to 300 points for evaluation, which is depicted in the figures below. All experiments are replicated 20 times. We use reachability as our performance metric, which is defined as $\text{Reach}_t = 1 - \text{BG}_t$ (the fraction of total behaviors observed). In all figures, we report the mean of Reach_t plus and minus one standard deviation computed over the independent replicates.

Lastly, to demonstrate UG-BEACON’s flexibility to handle user-defined preferences, we also present case studies on a MNIST digit discovery problem and an oil sorbent material discovery problem.

5.1 Synthetic Functions

We create three synthetic test functions from popular global optimization benchmarks – **Ackley**, **Rosenbrock**, and **Styblinski-Tang** – by adapting them to the NS problem setting. We treat the output of these functions as our outcomes and partition the range of outcomes into 25 equally-spaced intervals to define the behaviors. Further details are provided in Section 2 of the SI. The reachability performance of all algorithms across 9 problems (3 synthetic functions, each with 3 input dimensions $d \in \{4, 8, 12\}$) is shown in Figure 3. BEACON consistently outperforms the other algorithms in all cases, achieving near perfect reachability values of 1 in several cases. We further see that the performance gap increases as problem dimension increases.

To test BEACON in the multi-output setting, we also created a challenging **Multi-Output Plus** function with a long-tail joint distribution. The function has 6 inputs and 2 outputs (full description given in Section 2 of the SI). The distribution of outcomes for 10,000 inputs drawn uniformly at random from the space $\mathcal{X} = [-5, 5]^D$ is shown in Figure 2 (top). Reachability is computed with respect to intervals defined on a 10×10 grid over the two outcomes. NS algorithms find such jointly long-tail distributions challenging, as they tend to overexplore the high density center zone and ignore the tails. This is evident in the reachability results in Figure 2 (bottom). We see BEACON consistently has larger reachability than the other algorithms, uncovering nearly 80% while others only find 40% or less.

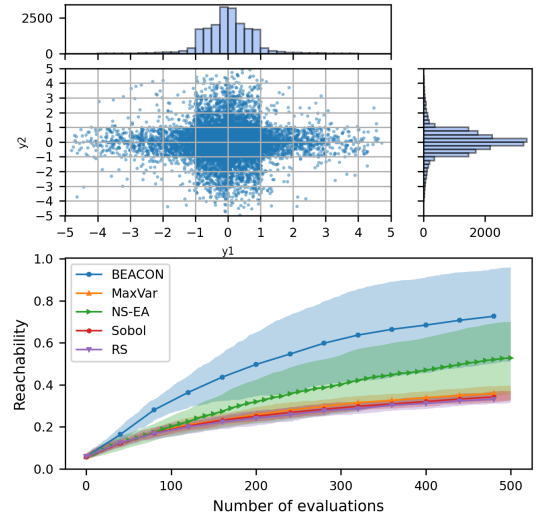


Fig. 2 Results for multi-output cross plus function. (Top) Scatter plot of outcomes for 10,000 random input values. (Bottom) Reachability performance for all algorithms.

5.2 Material Discovery Applications

Metal-organic frameworks (MOFs) are a highly diverse class of porous crystalline materials constructed from metal ions and organic linkers. Due to the vast design space, which spans linkers, metal centers, topologies, and defects, there is currently no comprehensive theoretical framework to assess the full chemical or property diversity that MOFs can exhibit³⁰. Recent studies have

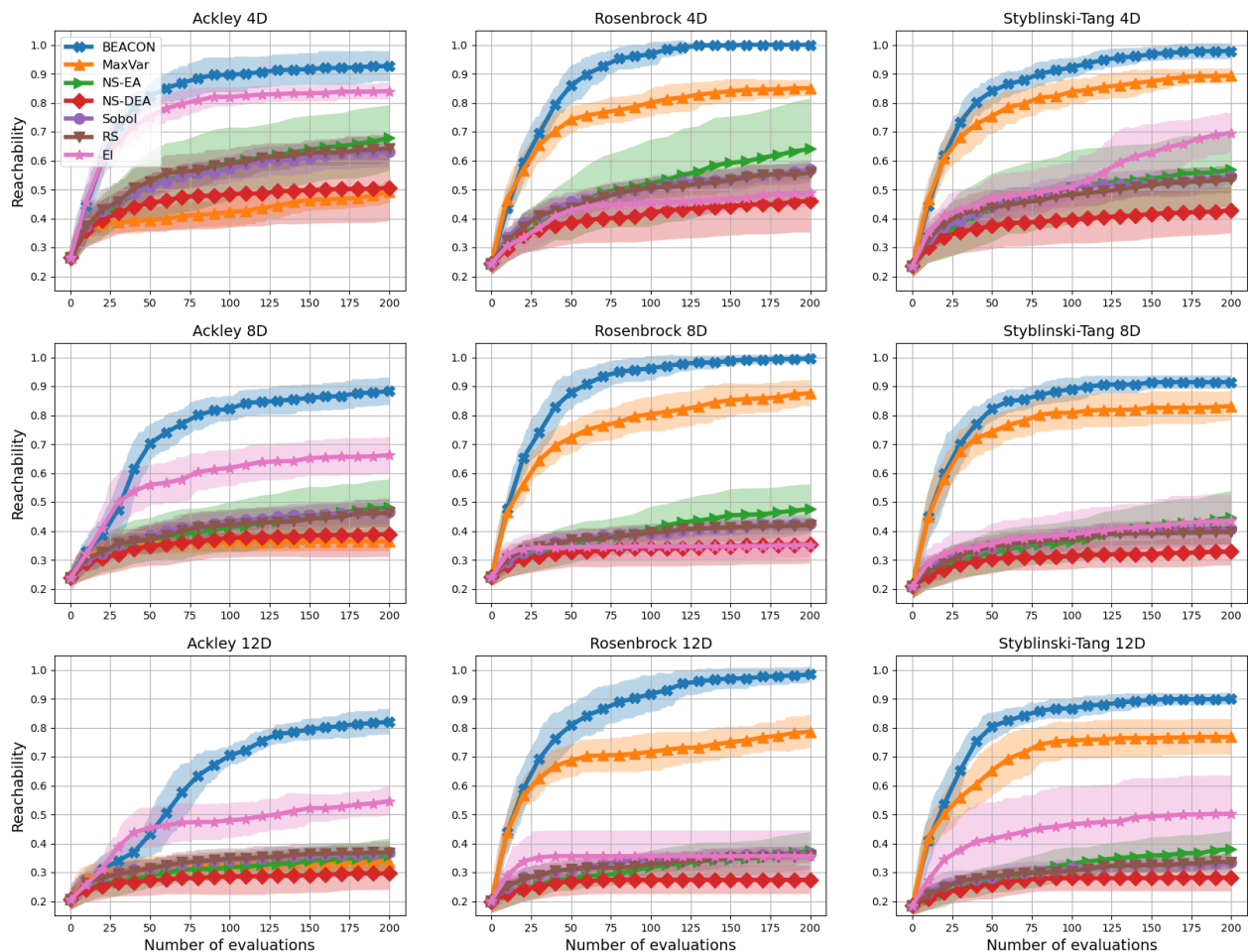


Fig. 3 Results on synthetic test problems for Ackley (Left column), Rosenbrock (Middle column), and Styblinski-Tang (Right column) with 4 (Top row), 8 (Middle row), and 12 (Bottom row) dimensions. BEACON provides substantially better outcome reachability over the benchmark methods, with larger improvements for higher-dimensional problems.

shown that existing MOF libraries are often biased, which can distort conclusions drawn from machine learning screening studies³¹. These challenges make novelty-driven search especially appealing. However, because MOF property evaluations are often costly (experimentally or computationally), sample-efficient approaches like BEACON are well-suited for the task.

In our setting, each candidate MOF is represented by a fixed-dimensional vector of molecular descriptors $\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^d$. The associated properties of interest form the outcome $\mathbf{y} = \mathbf{f}(\mathbf{x}) \in \mathcal{O} \subset \mathbb{R}^n$, where $\mathbf{f}$ is modeled as a black-box function. The choice of descriptors is adapted from domain-specific prior work (see section 2.2 in SI for details). We consider three MOF discovery tasks that reflect different search complexities and outcome dimensionalities.

5.2.1 Hydrogen uptake capacity.

Hydrogen storage is central to many clean energy technologies. We use a dataset from³² with 98,000 unique MOFs. Each MOF is encoded using 7 real-valued features, yielding $\mathcal{X} \subset \mathbb{R}^7$ and $\mathcal{O} \subset \mathbb{R}^1$.

5.2.2 Nitrogen uptake capacity.

Efficient nitrogen removal is important for natural gas purification. We use a dataset from³³ consisting of 5,224 MOFs, each described by 20 features, i.e., $\mathcal{X} \subset \mathbb{R}^{20}$ and $\mathcal{O} \subset \mathbb{R}^1$.

5.2.3 Joint CO₂ and CH₄ uptake.

We aim to discover MOFs with diverse combinations of carbon dioxide (y_1) and methane (y_2) uptake capacities. The dataset from³¹ includes 7,000 MOFs, represented by 25 features. Thus, $\mathcal{X} \subset \mathbb{R}^{25}$ and $\mathcal{O} \subset \mathbb{R}^2$. The joint outcome distribution is highly skewed and visualized in section 2.2.3 in the SI.

Figure 4 compares the reachability performance of BEACON and competing methods on these tasks. As the MOF design problems involve discrete input sets, standard evolutionary NS methods (NS-EA, NS-DEA, NS-FS) are not applicable. Across all three tasks, BEACON consistently achieves full reachability, indicating its ability to uncover maximal property diversity under limited evaluations.

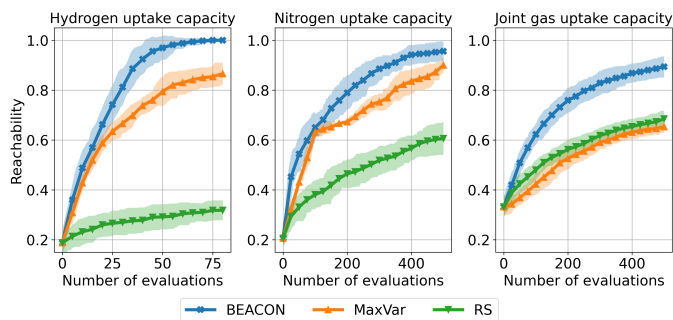


Fig. 4 Performance of BEACON and baseline methods on MOF discovery problems. Shown are results for hydrogen uptake (left), nitrogen uptake (middle), and joint CO_2/CH_4 uptake (right).

5.3 Drug Discovery Applications

Exploring the chemical space to identify small molecules with diverse properties is a foundational task in pharmaceutical research. BEACON provides a sample-efficient strategy to promote novelty across molecular properties while accommodating the unique modeling challenges of chemical data. For all tasks, the input space $\mathcal{X} \subset \mathbb{R}^d$ encodes molecular structure using either descriptor-based or high-dimensional representations, while the outcome space $\mathcal{O} \subset \mathbb{R}$ corresponds to molecular properties of interest. We consider three benchmark datasets with varying dimensionality and modeling complexity.

5.3.1 Water solubility.

We use the dataset from³⁴ containing 900 small organic molecules. Each molecule is encoded using a 14-dimensional molecular descriptor vector.

5.3.2 ESOL.

We use the dataset from³⁵ containing 1,128 molecules with experimentally measured aqueous solubility. Instead of low-dimensional descriptors, we adopt a 2,133-dimensional binary fragprint representation, which captures detailed structural information. BEACON is paired with the Tanimoto-fragprint kernel²⁸.

5.3.3 LogD.

The octanol-water distribution coefficient (LogD) is a key property in drug discovery. We use a dataset of 2,070 molecules from³⁶, each encoded with 125 features. To address the high dimensionality, we incorporate the sparse axis-aligned subspace (SAAS) prior²⁷, which promotes automatic feature selection within BEACON.

Figure 5 shows that BEACON consistently outperforms other methods in all tasks. Notably, in high-dimensional settings (ESOL and LogD), the use of chemically informed priors/kernels significantly improves performance. Additional experiment details are provided in section 2.3 in the supplementary information (SI).

5.4 Exploring Deceptive Landscapes in Reinforcement Learning

We evaluate BEACON on a challenging reinforcement learning (RL) task involving navigation through a large maze. The ob-

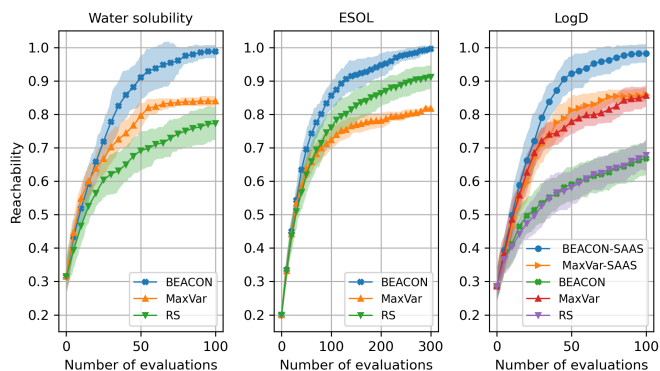


Fig. 5 Performance of BEACON and baseline methods on drug discovery problems. Shown are results for water solubility (left), ESOL (middle), and LogD (right). BEACON achieves superior reachability across all cases, particularly in high-dimensional settings.

jective is to control a ball from a start to a goal position within 300 steps using a linear policy defined by 8 tunable parameters. Instead of optimizing final distance alone, we use a normalized reward that measures relative improvement toward the target. See 2.4 in the supplementary information (SI) for a figure of the maze and additional details on the problem setup.

This setting is highly deceptive: reward-based acquisition can mislead search toward local optima. Figure 6 shows that BEACON consistently outperforms both standard NS variants and reward-driven approaches, including classical expected improvement (EI) acquisition in BO. The top panel shows the average best observed reward across 20 replicates, where BEACON outperforms all other methods, including the objective-focused EI approach, which ranks as the second-worst performing algorithm. The bottom panel presents a violin plot of the reward distribution at the final iteration for all algorithms. BEACON is the only method that consistently achieves successful escapes the maze, with all 20 replicates yielding a final reward of 1. These results clearly demonstrate the benefit of sample-efficient novelty-driven search in complex RL domains.

5.5 Incorporating User Information

We evaluate UG-BEACON on two problems with behavior constraints. The first case study uses the well-known MNIST dataset³⁷. The task is to discover as many distinct handwritten digits (0 through 9) as possible by sampling from the latent space of a trained variational autoencoder (VAE). UG-BEACON uses a convolutional neural network (CNN) classifier to identify whether a newly generated digit matches a previously observed class and prioritizes exploration of novel digit types. As shown in Figure 7, UG-BEACON achieves faster discovery of all digit classes compared to BEACON and baseline methods.

As a second demonstration of UG-BEACON, we consider the discovery of materials for oil adsorption applications. To reflect realistic application preferences, we define a non-uniform partition of the 2D outcome space based on target utility. As shown in Figure 8 (left), outcome regions with high strength and adsorption capacity are assigned finer grid resolution to encourage cov-

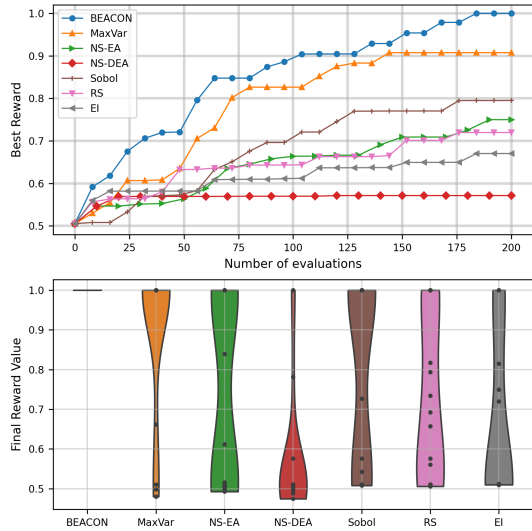


Fig. 6 Results for RL-based maze navigation case study. (Top) Average best observed reward for all algorithms across 20 replicates. (Bottom) Violin plot showing the distribution of best reward values at the final iteration for all algorithms.

erage in practically useful zones. In contrast, low-performance regions are coarsely partitioned, reducing sampling pressure in less relevant areas. UG-BEACON uses this user-defined grid to guide novelty evaluation and prevent redundant sampling by avoiding outcome cells already explored. Figure 8 (right) reports reachability performance across algorithms. UG-BEACON rapidly attains full (100%) coverage of the preferred outcome space within just 50 iterations, outperforming BEACON and other state-of-the-art novelty and reward-driven baselines. These results demonstrate the value of incorporating user behavior preferences into the NS process to accelerate discovery in practical material design tasks.

6 Conclusions

We presented BEACON, a sample-efficient algorithm for novelty search (NS) in settings where outcomes are defined by expensive, noisy black-box functions. BEACON integrates ideas from Bayesian optimization to guide exploration toward novel behaviors while accounting for uncertainty, enabling strong performance even in low-data regimes. Our experiments across a wide variety of synthetic, scientific, and reinforcement learning tasks demonstrate that BEACON consistently outperforms existing NS methods, particularly in high-dimensional or deceptive environments.

While BEACON offers substantial gains in sample efficiency, it does incur higher computational cost than simpler alternatives, which may limit its usefulness in low-cost evaluation settings. Theoretical convergence guarantees also remain an open question. Nonetheless, BEACON opens new opportunities for NS in real-world domains, and future work may extend its scope to even broader decision-making and design problems.

Author contributions

Wei-Ting Tang: Conceptualization, Methodology, Software, Formal analysis, Investigation, Visualization, Writing – original draft.

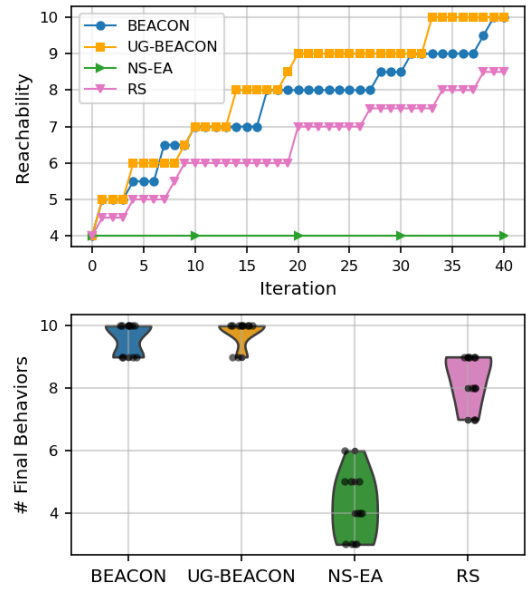


Fig. 7 Results for MNIST case study. (Top) Average number of discovered digits for all algorithms across 10 replicates. (Bottom) Violin plot showing distribution of discovered digits at the final iteration for all algorithms. UG-BEACON achieves faster convergence by incorporating user-defined behavior constraints on observed digits.

Ankush Chakrabarty: Methodology, Software, Formal analysis, Writing – review & editing. Joel A. Paulson: Conceptualization, Supervision, Funding acquisition, Project administration, Writing – review & editing.

Conflicts of interest

There are no conflicts to declare

Data availability

The BEACON code is publicly available at <https://github.com/PaulsonLab/BEACON>. We provide code to reproduce all empirical experiments done in this paper.

Acknowledgements

This work was supported by the National Science Foundation (NSF) CAREER Award 2237616.

Notes and references

- 1 L. M. Rios and N. V. Sahinidis, *Journal of Global Optimization*, 2013, **56**, 1247–1293.
- 2 B. Shahriari, K. Swersky, Z. Wang, R. P. Adams and N. De Freitas, *Proceedings of the IEEE*, 2015, **104**, 148–175.
- 3 S. J. Russell and P. Norvig, *Artificial intelligence: A modern approach*, Pearson, 2016.
- 4 P. I. Frazier, *arXiv preprint arXiv:1807.02811*, 2018.
- 5 A. R. Zeni and J. A. Gallas, *Physica D: Nonlinear Phenomena*, 1995, **89**, 71–82.
- 6 J. Lehman and K. O. Stanley, *Evolutionary Computation*, 2011, **19**, 189–223.

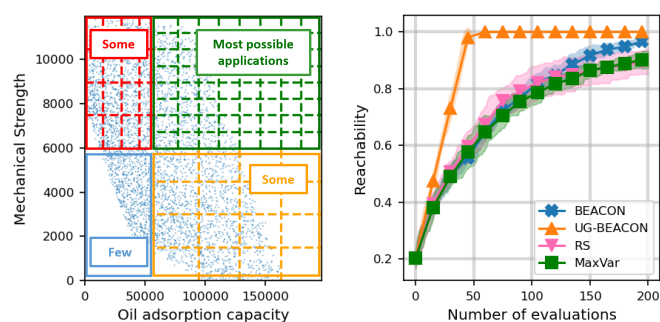


Fig. 8 Oil sorbent material case study. (Left) Partitioned 2D outcome space with user-defined behavior preferences. (Right) Reachability performance for UG-BEACON and other algorithms. UG-BEACON achieves 100% reachability within significantly fewer iterations than other methods.

- 7 J. Lehman and K. O. Stanley, *Genetic Programming Theory and Practice IX*, 2011, 37–56.
- 8 J. Grizou, L. J. Points, A. Sharma and L. Cronin, *Science Advances*, 2020, **6**, eaay4237.
- 9 K. Terayama, M. Sumita, R. Tamura, D. T. Payne, M. K. Chahal, S. Ishihara and K. Tsuda, *Chemical Science*, 2020, **11**, 5959–5968.
- 10 E. C. Jackson and M. Daley, *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, 2019, pp. 173–174.
- 11 J. Gomes, P. Mariano and A. L. Christensen, *Knowledge-Based Systems*, 2015, **82**, 1–20.
- 12 J.-B. Mouret and S. Doncieux, *IEEE Transactions on Evolutionary Computation*, 2012, **16**, 1–15.
- 13 J.-B. Mouret and J. Clune, *Artificial Life*, 2015, **22**, 245–268.
- 14 C. K. Williams and C. E. Rasmussen, *Gaussian Processes for Machine Learning*, MIT Press Cambridge, MA, 2006, vol. 2.
- 15 H. Liu, J. Cai and Y.-S. Ong, *Knowledge-Based Systems*, 2018, **144**, 102–121.
- 16 S. R. Chowdhury and A. Gopalan, *International Conference on Machine Learning*, 2017, pp. 844–853.
- 17 A. Kudva, W.-T. Tang and J. A. Paulson, *Computers & Chemical Engineering*, 2024, **181**, 108515.
- 18 N. Srinivas, A. Krause, S. M. Kakade and M. Seeger, *arXiv preprint arXiv:0912.3995*, 2009.
- 19 S. Ament, S. Daulton, D. Eriksson, M. Balandat and E. Bakshy, *Advances in Neural Information Processing Systems*, 2023, **36**, 20577–20612.

- 20 W. R. Thompson, *Biometrika*, 1933, **25**, 285–294.
- 21 K. Kandasamy, A. Krishnamurthy, J. Schneider and B. Póczos, *International Conference on Artificial Intelligence and Statistics*, 2018, pp. 133–142.
- 22 J. Wilson, V. Borovitskiy, A. Terenin, P. Mostowsky and M. Deisenroth, *International Conference on Machine Learning*, 2020, pp. 10292–10302.
- 23 R. H. Byrd, P. Lu, J. Nocedal and C. Zhu, *SIAM Journal on Scientific Computing*, 1995, **16**, 1190–1208.
- 24 S. Prillo and J. Eisenschlos, *International Conference on Machine Learning*, 2020, pp. 7793–7802.
- 25 F. Berkenkamp, A. P. Schoellig and A. Krause, *Journal of Machine Learning Research*, 2019, **20**, 1–24.
- 26 J. R. Gardner, M. J. Kusner, Z. E. Xu, K. Q. Weinberger and J. P. Cunningham, *International Conference on Machine Learning*, 2014, pp. 937–945.
- 27 D. Eriksson and M. Jankowiak, *Uncertainty in Artificial Intelligence*, 2021, pp. 493–503.
- 28 R.-R. Griffiths, L. Klarner, H. Moss, A. Ravuri, S. Truong, Y. Du, S. Stanton, G. Tom, B. Rankovic, A. Jamasb, A. Deshwal, J. Schwartz, A. Tripp, G. Kell, S. Frieder, A. Bourached, A. Chan, J. Moss, C. Guo, J. P. Dürholt, S. Chaurasia, J. W. Park, F. Strieth-Kalthoff, A. Lee, B. Cheng, A. Aspuru-Guzik, P. Schwaller and J. Tang, *Advances in Neural Information Processing Systems*, 2023, pp. 76923–76946.
- 29 S. Doncieux, G. Paolo, A. Laflaquière and A. Coninx, *Proceedings of the 2020 Genetic and Evolutionary Computation Conference*, 2020, pp. 85–93.
- 30 S. Lee, B. Kim, H. Cho, H. Lee, S. Y. Lee, E. S. Cho and J. Kim, *ACS Applied Materials & Interfaces*, 2021, **13**, 23647–23654.
- 31 S. M. Moosavi, A. Nandy, K. M. Jablonka, D. Ongari, J. P. Janet, P. G. Boyd, Y. Lee, B. Smit and H. J. Kulik, *Nature Communications*, 2020, **11**, 1–10.
- 32 S. Ghude and C. Chowdhury, *Chemistry–A European Journal*, 2023, **29**, e202301840.
- 33 H. Daglar and S. Keskin, *ACS Applied Materials & Interfaces*, 2022, **14**, 32134–32148.
- 34 S. Boobier, D. R. Hose, A. J. Blacker and B. N. Nguyen, *Nature Communications*, 2020, **11**, 5753.
- 35 J. S. Delaney, *Journal of Chemical Information and Computer Sciences*, 2004, **44**, 1000–1005.
- 36 Z.-M. Win, A. M. Cheong and W. S. Hopkins, *Journal of Chemical Information and Modeling*, 2023, **63**, 1906–1913.
- 37 L. Deng, *IEEE Signal Processing Magazine*, 2012, **29**, 141–142.