Multi-Objective Molecular Design Through Learning Latent Pareto Set

Yiping Liu[1*], Jiahao Yang[1*], Xuanbai Ren[1], Xinyi Zhang[1], Yuansheng Liu[1], Bosheng Song[1], Xiangxiang Zeng[1†], Hisao Ishibuchi[2]

[1]College of Computer Science and Electronic Engineering, Hunan University, China
[2]Department of Computer Science and Engineering, Southern University of Science and Technology, China.
{yiping0liu, jiahaoyoung0520}@gmail.com, xzeng@foxmail.com

Abstract

Molecular design inherently involves the optimization of multiple conflicting objectives, such as enhancing bioactivity and ensuring synthesizability. Evaluating these objectives often requires resource-intensive computations or physical experiments. Current molecular design methodologies typically approximate the Pareto set using a limited number of molecules. In this paper, we present an innovative approach, called Multi-Objective Molecular design through Learning Latent Pareto Set (MLPS). MLPS initially utilizes an encoder-decoder model to seamlessly transform the discrete chemical space into a continuous latent space. We then employ local Bayesian optimization models to efficiently search for local optimal solutions (i.e., molecules) within predefined trust regions. Using surrogate objective values derived from these local models, we train a global Pareto set learning model to understand the mapping between direction vectors (called "preferences") in the objective space and the entire Pareto set in the continuous latent space. Both the global Pareto set learning model and local Bayesian optimization models collaborate to discover high-quality solutions and adapt the trust regions dynamically. Our work is an effective endeavor towards learning the Pareto set for multi-objective molecular design, providing decision-makers with the capability to fine-tune their preferences and thoroughly explore the Pareto set. Experimental results demonstrate that MLPS achieves state-of-the-art performance across various multi-objective scenarios, encompassing diverse objective types and varying numbers of objectives. The effectiveness of MLPS was further validated through real-world challenges in discovering antifungal peptides with low toxicity and high activity.

Code — https://github.com/JiahaoYoung0520/MLPS/source

Extended version — https://github.com/JiahaoYoung0520/MLPS/extend-version

Introduction

Molecular design plays a pivotal role in a multitude of applications, including drug discovery (Meyers, Fabian, and Brown 2021), material science (Butler et al. 2018), and catalyst development (Wan, Duan, and Huang 2020). Influenced by the rapid development of artificial intelligence, machine learning has shown its powerful ability to assist in efficient candidate discovery (Chen et al. 2021; Hoffman et al. 2022). However, in practice, molecules often need to fulfill multiple property requirements, e.g., drug molecules not only need to bind efficiently to the target, but also need to have suitable pharmacokinetic properties and low toxicity. In this context, multi-objective molecular design is challenging. It aims to discover the Pareto set of molecules, where improving one objective inevitably entails compromising others.

Traditional approaches to multi-objective molecular design often involve simplifying multi-objective problems by converting them into single-objective ones using specific weightings (Abels et al. 2019; SV et al. 2022). While effective to some extent, these methods rely on human experts for weighting. In general, weighting of multiple objectives is difficult, and inappropriate weighting leads to inappropriate solutions. Other techniques focus on identifying the Pareto set through a two-stage process: molecule sampling and non-dominated sorting (Yasonik 2020; Verhellen 2022). However, this approach can become prohibitively expensive and time-consuming, especially in cases with multiple objectives and a large pool of molecules to consider. To improve sampling efficiency in multi-objective molecular design, Bayesian optimization has emerged as a powerful and efficient method (Xie et al. 2021; Gao et al. 2022). Nevertheless, Bayesian optimization faces limitations when dealing with high-dimensional spaces and high computational complexity associated with Gaussian process inference. These challenges make it less suited for complex multi-objective molecular design tasks.

Conventional multi-objective molecular design approaches often yield a limited Pareto set (i.e., a small number of trade-off solutions), which may not align well with decision-makers preferences. multi-objective molecular design problems may have a complicated Pareto set with various trade-offs. Accessing this entire set can provide significant advantages, enabling decision-makers to select the most preferred solution and accelerating the molecular design process. Recent advancements in this field, particularly Pareto set learning methods, aim to approximate the entire Pareto set using learning models (Lin, Yang, and

*These authors contributed equally.
†Corresponding Author

Zhang 2021; Lin et al. 2022). However, most existing methods focus on low-dimensional Pareto sets, which are not suitable for addressing the complexity of high-dimensional and discrete multi-objective molecular design problems. These limitations underscore the need for a novel approach that can efficiently handle multi-objective molecular design scenarios while learning the comprehensive Pareto set.

In this paper, we introduce a groundbreaking approach to multi-objective molecular design, termed as Multi-Objective Molecular Design through Latent Pareto Set Learning (MLPS). MLPS fundamentally transforms the conventional paradigm by learning the entire Pareto set in a high-dimensional latent space. The core of MLPS begins with the utilization of an encoder-decoder model, which effectively maps the discrete chemical space into a continuous latent space. This transformation process lays the foundation for subsequent operations. MLPS partitions the latent space into multiple trust regions and uses a local Bayesian optimization model for efficient sampling within each region. Notably, MLPS incorporates a global Pareto set learning model, which establishes a crucial link between direction vectors (i.e., preferences) in the objective space and the comprehensive Pareto set in the continuous latent space. It is capable of aggregating and disseminating valuable information from each trust region, thereby enhancing both global and local optimization processes.

MLPS provides a vast array of solutions for decision-makers to choose from, enabling them to select molecules that best align with their preferences and specific requirements. This is clearly different from existing multi-objective molecular design methods where only a limited number of trade-off solutions are obtained. Importantly, this work marks the pioneering effort in molecular design by directly learning the entire Pareto set for the complex multi-objective molecular design problem. Our paper makes several significant contributions to the field of multi-objective molecular design.

- We introduce a novel approach that learns a mapping from a preference to the corresponding Pareto optimal solution in multi-objective molecular design. It empowers decision-makers to fine-tune their preferences, facilitating exploration across the entire Pareto set. This is a reverse process compared to existing methods and offers a fresh perspective on solving the problem.
- We develop an efficient information-sharing mechanism wherein local Bayesian models communicate with each other through a global neural network. This combination of local and global optimization enhances the search capability for the Pareto set, allowing for quicker convergence to better solutions.
- Our approach outperforms existing state-of-the-art methods across a wide range of multi-objective scenarios and can be applied to real-world challenging tasks. This includes scenarios involving diverse types of objectives and varying numbers of objectives, showcasing its effectiveness and versatility.

Related Works

Multi-objective Molecular Design. Multi-objective molecular design aims to improve specific molecular several properties simultaneously. Recent advances have seen the emergence of various artificial intelligence-based approaches.These approaches can be broadly categorized into three groups: 1) Use Generative Models to sample : This category includes methods such as variational autoencoders (Lim et al. 2018; Liu et al. 2018; Jin, Barzilay, and Jaakkola 2018), generative adversarial networks (Guimaraes et al. 2017; Kadurin et al. 2017; Prykhodko et al. 2019), and diffusion models (Ho, Jain, and Abbeel 2020; Xu et al. 2021; Hoogeboom et al. 2022). They work with continuous latent representations and require substantial data for training. 2) Combinatorial Optimization Methods: These methods involve combinatorial search in a discrete chemical space. They include evolutionary algorithms (Nigam et al. 2020; Ahn et al. 2020), reinforcement learning (You et al. 2018; Jin, Barzilay, and Jaakkola 2020a; Fu et al. 2021a), and Monte Carlo methods (Xie et al. 2021; Fu et al. 2021b; Sun et al. 2022). 3) Multi-objective Bayesian Optimization: these approaches often involve learning a latent space from molecular data using generative models and then searching for molecules with desired properties in the latent space (Gómez-Bombarelli et al. 2018; Jin, Barzilay, and Jaakkola 2018; Siivola et al. 2021). Existing Multi-objective optimization methods in molecular design often face limitations. Many methods involve transforming a multi-objective problem into a set of single-objective ones by using scalarizing functions, allowing single-objective optimization to be applied. These transforming may not capture the full spectrum of solutions (i.e., multi-objective trade-offs). Additionally, representing complex chemical spaces in high-dimensional latent spaces can be challenging, and traditional multi-objective Bayesian optimization methods have limitations in handling high-dimensional spaces (Frazier 2018). Our MLPS provides a novel framework to learn the mapping from preferences to the Pareto set while efficiently navigating the search in high-dimensional spaces.

Pareto Set Learning. In the field of multi-objective optimization, there has been a growing interest in developing methods that learn to approximate the Pareto set. A prevalent trend in multi-objective optimization involves integrating preference information into neural networks. This approach has demonstrated success in various domains, including multi-task learning (Sener and Koltun 2018; Navon et al. 2020) and reinforcement learning (Abdolmaleki et al. 2020, 2021). By incorporating preferences, neural networks can adapt to specific preferences of decision-makers. Some notable works have aimed to learn and model the entire Pareto set. P-MOCO (Lin, Yang, and Zhang 2021) introduces an end-to-end reinforcement learning algorithm designed to train models capable of accommodating different preferences in multi-objective combinatorial optimization problems. In the context of expensive multi-objective optimization problems, PSL (Lin et al. 2022) has been developed. However, it is primarily designed for low-dimensional problems. Moreover, both P-MOCO and PSL rely on a single global model, which may not be well-suited for captur-

ing highly complicated Pareto sets encountered in molecular design. In contrast, our MLPS takes a hybrid approach that combines several local models with a global model, which is specifically designed to address the challenges posed by complicated high-dimensional Pareto sets in molecular design.

Method

In our study, we treat the multi-objective molecular design problem as a multi-objective maximization problem, formulated as follows:

$$\max_{\boldsymbol{x}\in\boldsymbol{\chi}} \mathbf{F}(\boldsymbol{x}) = [f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \cdots, f_M(\boldsymbol{x})]. \quad (1)$$

Here, $\boldsymbol{x}$ denotes a solution (i.e., molecule) in the N-dimensional latent space $\boldsymbol{\chi} \in \mathbb{R}^N$, M represents the number of objectives, and evaluating these objective functions can be expensive.

In this section, we first provide an overview of our MLPS framework. Then, we delve into the specifics of the local and global models within our MLPS respectively. To provide readers with a solid foundation in multi-objective optimization, Appendix A offers explanations on fundamental concepts such as Pareto dominance, Pareto set/front, the hypervolume (HV) indicator (f_{HV}), the hypervolume contribution (f_{HVC}), and the hypervolume improvement (f_{HVI}). For those seeking more in-depth information, Appendix B supplies additional details about our work.

Framework

The framework of our MLPS is illustrated in Fig 1. MLPS operates as a guided search model, where guidance and feedback from current candidates continuously inform and improve the search strategy in each iteration. This search is conducted within a continuous latent space established by a pre-trained encoder-decoder model. Our approach utilizes a latent space model with a dimensionality of 256, employing SELFIES as the molecular sequence representation. Details about the molecular representations and the encoder-decoder model are found in Appendix B.1.

The first step in the search process involves setting the initial molecular embeddings. MLPS offers two methods for accomplishing this: 1) One method samples a set of molecules from a database, and their embeddings are obtained using the encoder. 2) The other method directly samples initial embeddings from the latent space. In our work, we employ the Sobol sampler (Renardy et al. 2021), which enables uniform sampling in the high-dimensional latent space.

Once the initial molecular embeddings are in place, we proceed to initialize multiple trust regions within the latent space. Within each trust region, we build a local surrogate model. Subsequently, based on these local surrogate models, we train a global Pareto set learning model, which is implemented as a neural network. The global Pareto set learning model is designed to establish a mapping between direction vectors (i.e., preferences) in the objective space and the entire Pareto set across the latent space. To train the global Pareto set learning model, we randomly sample a set of preferences and utilize it to predict Pareto optimal solutions. The predicted solutions are then evaluated by the local surrogate models within their respective trust regions to generate loss values. The loss is utilized as the training signal for the global model, enabling it to improve its mapping accuracy.

After training the global model in the current iteration, we employ it to generate predicted Pareto optimal solutions by randomly inputting preferences into the model. Subsequently, we select the best global batch of solutions $\boldsymbol{X}_g$ based on their hypervolume improvement values. Additionally, we sample a set of solutions within the trust regions, and we select the best local batch of solutions $\boldsymbol{X}_l$ according to their hypervolume improvement values. The details of the global and local batch selection can be found in Appendix B.2.

The selected solutions $\boldsymbol{X} = \boldsymbol{X}_g \cup \boldsymbol{X}_l$ are then decoded to obtain their molecular representations. These molecules are subsequently evaluated for their true objective vectors $\boldsymbol{Y}$. The pair $\{\boldsymbol{X}, \boldsymbol{Y}\}$ is used to update the trust regions, and the current iteration is over. The next iteration starts with the updated trust regions, allowing the optimization process to iteratively improve and refine the Pareto set. This approach combines both global and local information, leveraging the strength of each to enhance the efficiency and effectiveness of the multi-objective molecular design process. We summarize the core components of MLPS in Algorithm 1.

Algorithm 1: The framework of MLPS

Require: Pre-trained encoder and decoder, initial embeddings and their objective vectors $\{\boldsymbol{X}_{init}, \boldsymbol{Y}_{init}\}$, the number of trust regions n_{tr}, the minimum edge length of trust regions L_{min}

Initialize n_{tr} trust regions $\boldsymbol{T} = \{T_1, T_2, \cdots, T_{n_{tr}}\}$ with $\{\boldsymbol{X}_{init}, \boldsymbol{Y}_{init}\}$;
Build a local surrogate model in each trust region;
while budget not exhausted **do**
    Train the global Pareto set learning model $h_{\boldsymbol{\theta}}$ based on the local surrogate models (Alg.2);
    Generate solutions from random preferences by $h_{\boldsymbol{\theta}}$ to form the global batch $\boldsymbol{X}_g$ (Appendix Alg.2);
    Sample solutions in $\boldsymbol{T}$ to form the local batch $\boldsymbol{X}_l$ (Appendix Alg.1);
    $\boldsymbol{X} \leftarrow \boldsymbol{X}_g \cup \boldsymbol{X}_l$;
    Decode $\boldsymbol{X}$ and obtain their true objective vectors $\boldsymbol{Y}$;
    for $j = 1, 2, \cdots, n_{tr}$ **do**
        Update the edge length of T_j based on $\{\boldsymbol{X}, \boldsymbol{Y}\}$;
        if the edge length of T_j is less than L_{min} **then**
            Reinitialize T_j (Appendix Alg.3);
        end if
        Update the center of T_j based on f_{HVC} (Subsection Local Models);
    end for
    Update the local surrogate model in each trust region;
end while

Ensure: global Pareto set learning model $h_{\boldsymbol{\theta}}$

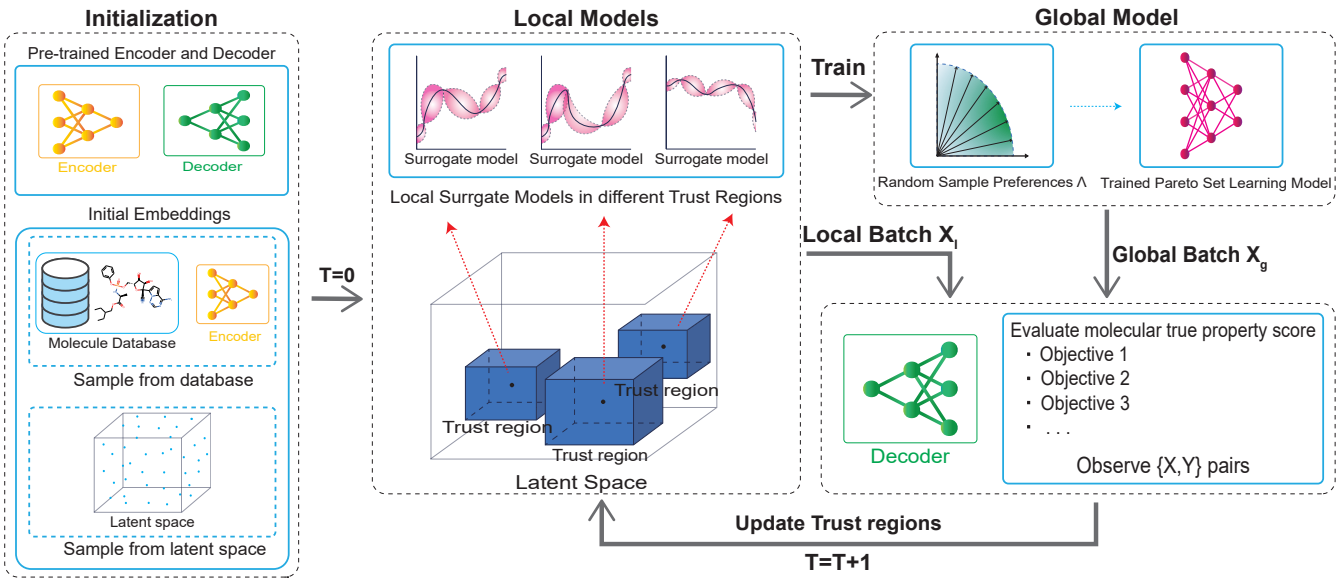

Figure 1: Illustration of the framework of the proposed MLPS

Local Models

In our MLPS, each trust region is equipped with a local surrogate model, which makes MLPS different from other global Bayesian optimization methods. Specifically, each local surrogate model is a Gaussian process implemented using BoTorch (Balandat et al. 2020) and GPyTorch (Gardner et al. 2018). Our approach focuses on maintaining the accuracy of each local surrogate model within its respective trust region. Similar to Turbo (Eriksson et al. 2019), a trust region is a hypercube region defined within the latent space in this work. There are two key issues for each trust region in our work: the setting of its center and the reinitialization strategy, which we discuss below:

Center Setting. In single-objective optimization, existing approaches often place the center of a trust region at the best observed point. However, this approach is no longer suitable for multi-objective optimization since there exists no single best solution. In MLPS, we select the point with the maximum hypervolume contribution as the trust region center. Given a non-dominated solution set within the current trust region, we calculate the hypervolume contribution for each solution and choose the point with the highest hypervolume contribution to be the center of the trust region, while excluding points that have already been selected as the centers of other trust regions. Placing the center of the trust region at the point with the highest hypervolume contribution enhances diversity. This is because more crowded solutions tend to have smaller hypervolume contributions.

Reinitialization Strategy. An unpromising trust region is penalized by halving the edge length. This reduction in edge length is not unlimited. If the length of an edge becomes smaller than a predefined threshold L_{min}, the current trust region is terminated, and a new trust region is generated. For generating a new trust region, we determine the position of the new center and build a new local surrogate model. Let $\boldsymbol{D}_p = (\boldsymbol{X}_p, \boldsymbol{Y}_p)$ be the set of previously reinitialized center points $\boldsymbol{X}_p$ and their corresponding observations $\boldsymbol{Y}_p$. $\hat{\boldsymbol{f}}_r$ is the surrogate model corresponding to the trust region to be reinitialized T_r. Then, we build a new local surrogate model to $\boldsymbol{D}_p$: $\hat{\boldsymbol{f}}_{new} \sim P(\hat{\boldsymbol{f}}_r \mid \boldsymbol{D}_p)$. Based on $\hat{\boldsymbol{f}}_{new}$, we identify the center point $\boldsymbol{x}_c$ of the new trust region by maximizing a scalarizing function specified by a random preference. The reinitialization strategy is outlined in Appendix Alg.3. This approach ensures that trust regions can be reinitialized in promising parts of the overall space.

Global Pareto Set Learning Model

As previously highlighted, the Pareto set of a multi-objective problem may contain an infinite or large number of solutions, each offering a different trade-off among the objectives. In contrast to traditional multi-objective molecular design approaches, MLPS employs a global Pareto set learning model to map a user defined trade-off preference to the corresponding Pareto optimal solution. To achieve this, we employ scalarizing functions to calculate gradients using local surrogate models, and these gradients are utilized to update the global Pareto set learning model. The global Pareto set learning model serves as a bridge connecting different trust regions. It can generate improved global optimal solutions, encouraging local search within trust regions to explore and discover even better solutions. Conversely, The enhanced local solutions found in trust regions can provide guidance to update the global model. This feedback loop ensures that the global model continuously improves its understanding of the Pareto set and adapts to changing preferences.

Model Formulation. The function of the global Pareto set learning model is expressed as follows:

$$\boldsymbol{x} = h_{\boldsymbol{\theta}}(\boldsymbol{\lambda}). \tag{2}$$

Here, $\boldsymbol{\lambda}$ represents any valid preference selected from $\boldsymbol{\Lambda} = \{\boldsymbol{\lambda} \in \mathbb{R}^M \mid \sum \lambda_i = 1\}$, with $i = 1, 2, \ldots, M$. $\boldsymbol{x}$ is

the corresponding solution to the preference $\boldsymbol{\lambda}$ in the N-dimensional latent space. Typically, N is much larger than M in multi-objective molecular design. $h_{\boldsymbol{\theta}}$ denotes a neural network with parameters $\boldsymbol{\theta}$. We utilize thermometer encoding (Buckman et al. 2018) for preferences and apply a cross-attention mechanism for different preferences. Specifically, we divide the preference values [0,1] into 128-dimensional scales to enhance the model's sensitivity to minor fluctuations. We apply two layers of cross-attention to balance the focus on different targets, thereby improving the effectiveness and diversity of the Pareto set. We introduce residual connections within the network to ensure that the model learns these distinctions effectively. These connections help preserve the original preference information throughout the transformation process. By combining preference data with its features, our model can generate corresponding solutions that are more distinguishable and faithful to the preferences. To ensure that the model learns these distinctions effectively, we introduce residual connections within the network. These connections help preserve the original preference information throughout the transformation process. By combining preference data with its features, our model can generate corresponding solutions that are more distinguishable and faithful to the preferences.

Model Training. The training process aims to update the model's parameters $\boldsymbol{\theta}$ such that the generated solutions align with the optimal solutions by minimizing the augmented Tchebycheff scalarization. This can be expressed as:

$$\boldsymbol{x}^* = h_{\boldsymbol{\theta}^*}(\boldsymbol{\lambda}) = \arg\min_{\boldsymbol{x}\in\boldsymbol{\chi}} g_{tch\_aug}(\boldsymbol{x} \mid \boldsymbol{\lambda}). \quad (3)$$

The detaild of the augmented Tchebycheff scalarization g_{tch_aug} can be found in Appendix B.3. This function establishes a connection between a set of preferences $\boldsymbol{\Lambda} = \{\boldsymbol{\lambda} \in \mathbb{R}^M \mid \sum \lambda_i = 1\}$ and their corresponding solutions within the Pareto set. It guides the global Pareto set learning model to generate solutions that are close to the Pareto front.

To find the optimal parameter $\boldsymbol{\theta}^*$, we propose an efficient algorithm. Since the optimal solution of the augmented Tchebycheff scalarizing function is unknown, we need to optimize all solutions generated by our model with the corresponding augmented Tchebycheff scalarizing functions for all valid preferences:

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} E_{\boldsymbol{\theta}\sim\boldsymbol{\Lambda}} g_{tch\_aug}(\boldsymbol{x} = h_{\boldsymbol{\theta}}(\boldsymbol{\lambda}) \mid \boldsymbol{\lambda}). \quad (4)$$

Solving Eq. (4) directly is challenging due to the expectation over an infinite set of preferences. Therefore, we employ a Monte Carlo sampling and gradient descent approach to iteratively update the model with different surrogate models in different trust regions. The update equation is as follows:

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t - \eta \sum_{k=1}^{K} \nabla_{\boldsymbol{\theta}} \hat{g}_{tch\_aug}(\boldsymbol{x} = h_{\boldsymbol{\theta}}(\boldsymbol{\lambda}_k) \mid \boldsymbol{\lambda}_k). \quad (5)$$

Here, $\hat{g}_{tch_aug}(\cdot)$ is the augmented Tchebycheff scalarizing function with objective vectors predicted by local surrogate models (Please refer to Eq. (6) in Appendix B.3). To account for the uncertainty of the surrogate model, we use the lower confidence bound to obtain the surrogate objective vector:

$$\hat{\boldsymbol{f}}(\boldsymbol{x}) = \hat{\boldsymbol{\mu}}(\boldsymbol{x}) - \beta\hat{\boldsymbol{\sigma}}(\boldsymbol{x}), \quad (6)$$

Algorithm 2: Global Pareto Set Learning Model Training

Require: global Pareto set learning model $h_{\boldsymbol{\theta}}$, the number of iterations for training the global model T_g, trust regions $\boldsymbol{T} = \{T_1, T_2, \cdots, T_{n_{tr}}\}$, corresponding local surrogate models $\{\hat{\boldsymbol{f}}_1, \hat{\boldsymbol{f}}_2, \cdots, \hat{\boldsymbol{f}}_{n_{tr}}\}$, the number of sampled preference n

for $t_g = 1, 2, \cdots, T_g$ **do**
  Randomly sample n preferences $\{\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \ldots, \boldsymbol{\lambda}_n\}$;
  Use the current model $h_{\boldsymbol{\theta}}$ to generate solutions $\boldsymbol{X}_{\boldsymbol{\theta}} = \{\boldsymbol{x} \mid \boldsymbol{x}_i = h_{\boldsymbol{\theta}}(\boldsymbol{\lambda}_i), i = 1, 2, \cdots, n\}$;
  Initialize an empty set $loss$;
  for $i = 1, 2, \cdots, n$ **do**
    Find the closest trust region to $\boldsymbol{x}_i$, denoted as T_j;
    Calculate the scalarizing function $\hat{g}_{tch\_aug}(\boldsymbol{x}_i)$ based on the surrogate model $\hat{\boldsymbol{f}}_j$ (Appendix B.3);
    Calculate $\nabla_{\boldsymbol{\theta}}\hat{g}_{tch\_aug}(\boldsymbol{x}_i)$ and add it to the $loss$ set;
  end for
  Update $\boldsymbol{\theta}$ with $loss$;
  $t_g = t_g + 1$;
end for

Ensure: updated Pareto set learning model $h_{\boldsymbol{\theta}}$

where $\hat{\boldsymbol{\mu}}$ is the mean value, $\hat{\boldsymbol{\sigma}}$ is the variance value, and β is a parameter that balances the weight between the mean and variance. In this work, we set β to 0.1.

Algorithm 2 describes the training of the global Pareto set learning model. This process iteratively updates the global Pareto set learning model, allowing it to learn the mapping from preferences to corresponding solutions in the Pareto set.

Experiments

In our experiments, we aim to explore the effectiveness of the proposed MLPS in various multi-objective molecular design scenarios. We divided the experimental setup into multi-objective molecular design tasks and real-world challenging tasks.

Multi-Objective Molecular Design Tasks

We consider four objectives related to molecular properties, following (Jin, Barzilay, and Jaakkola 2020a; Xie et al. 2021; Sun et al. 2022). To comprehensively evaluate the performance of our approach across different types and numbers of objectives, we consider several objective combinations within our experiments:

1) QED + SA (non-biological objectives): Drug-likeness (QED) and synthetic accessibility (SA) assess the drug-likeness and synthesizability of molecules, predicted using RDKit. 2) GSK3β + JNK3 (biological objectives): Inhibition of GSK3β and JNK3 (related to Alzheimer's) is predicted using a random forest model. 3) QED + SA + GSK3β/JNK3: Inhibition of GSK3β or JNK3 with constraints on drug-like properties (QED) and synthesizability (SA). 4) QED + SA + GSK3β + JNK3: Joint inhibition of GSK3β and JNK3 with constraints on QED and SA.

Objective	QED+SA	GSK3β+JNK3	QED+SA+GSK3β	QED+SA+JNK3	QED+SA+GSK3β+JNK3
GA+D	0.598	0.350	0.243	0.251	0.137
JT-VAE	0.832	0.460	0.276	0.287	0.254
GCPN	0.850	0.830	0.186	0.191	0.100
RationalRL	0.750	0.762	0.722	0.567	0.539
MARS	0.916	0.898	0.763	0.778	0.679
MolSearch	/	0.723	0.183	0.217	0.571
PSL	0.763	0.512	0.481	0.472	0.432
RetMol	0.847	**0.910**	0.771	0.781	0.701
HN-GFN	0.907	0.798	0.702	0.713	0.578
MLPS	**0.922**	0.902	**0.781**	**0.788**	**0.714**

Table 1: Comparison of different methods on HV

Baseline. Here is a brief description of each baseline method: 1) **GA+D** (Nigam et al. 2020) combines a genetic algorithm with a generative model. 2) **JT-VAE**(Jin, Barzilay, and Jaakkola 2018) generates molecules by constructing a tree-structured scaffold over chemical substructures and then assembling them into complete molecules, leveraging variational autoencoders. 3) **GCPN** (You et al. 2018) employs reinforcement learning to generate molecules atom by atom, utilizing a graph neural network as the basis. 4) **RationaleRL** (Jin, Barzilay, and Jaakkola 2020a) extends molecule rationales into complete molecules using reinforcement learning. 5) **MARS** (Xie et al. 2021) utilizes Markov sampling to generate molecules using a combination of GNNs and molecule fragments. 6) **MolSearch** (Sun et al. 2022) employs a Monte Carlo tree search algorithm to discover molecular design moves and generate molecules. 7) **PSL** (Lin et al. 2022) utilizes a neural network to learn the entire Pareto set for expensive multi-objective optimization. 8) **RetMol** (Wang et al. 2023) introduces a search-based framework that guides controlled molecule generation using a small set of molecules, iteratively refining the process to meet desired properties. 9) **HN-GFN** (Zhu et al. 2023) introduces a conditional variant of GFlowNet that can efficiently sample candidates. Please refer to Appendix C.2 for more details of these baseline methods.

Evaluation Metrics. We generate 5000 molecules by each method and use **hypervolume (HV)** to compare these methods. HV assesses how well the generated molecules cover the Pareto front in the objective space. It indicates how closely the generated molecules approximate the entire Pareto set and provides insights into their distribution in the objective space across different objectives (see details in Appendix A). In addition to HV, we consider several traditional metrics that are commonly used in molecular design, even though they were not explicitly designed for multi-objective molecular design: **Success Rate**, **Novelty**, **Diversity**, and **Product Metric**. (see details in Appendix C.2). We also evaluate the values of single objective in multi-objective task setting. Please refer to Appendix C.4 for details.

Benchmark Results

Table 1 presents the average HV values over 10 runs obtained by the compared methods. From the results, we can observe that: 1) In all five tasks, MLPS consistently achieves the highest HV scores among the compared methods. This indicates that MLPS excels in approximating the entire Pareto set, showcasing its ability to provide diverse and high-quality solutions across different multi-objective scenarios. 2) MARS and RationalRL typically secure the second and third places in terms of HV scores. These methods exhibit strong capabilities in multi-objective molecular design, but MLPS consistently outperforms them across different scenarios. 3) Several methods, such as GA+D, JT-VAE, and GCPN, perform acceptably in the two-objective scenarios. However, their performance degrades when faced with problems involving three or four objectives. 4) MolSearch performs relatively better in the four-objective scenario compared to the three-objective scenarios. This behavior might be attributed to the optimization strategy that MolSearch employs: it starts with a specialized molecule set tailored for the four-objective scenario. Note that MolSearch is not applicable to the QED+SA scenario due to its two-stage nature, where the optimization of biological objectives is prioritized in the first stage.

Table 2 presents the PM values (SR $\times$ Nov $\times$ Div) which can describe the comprehensive quanlity of molecules. From the results, we can observe the following: 1) MLPS consistently achieves the best performance across all five tasks. 2) Compared to other methods, such as JT-VAE and GCPN, MLPS excels in 3-objective and 4-objective tasks, demonstrating its capability to handle more objectives and more complex scenarios effectively. The more detailed results are shown in Appendix C.4.

Comparison with other Pareto set learning methods. PSL and P-MOCO are two other methods used for learning the Pareto set. PSL is designed to solve general multi-objective tasks, but it cannot directly optimize molecular design. To address this, we have combined our encoder-decoder architecture with PSL to make it more suitable for molecular design tasks. Additionally, while P-MOCO takes graph data as input, our encoder-decoder is based on sequence data. For the sake of a fair comparison, we only compare MLPS with PSL. The results, shown in Table 1 and 2 , highlight the following observations: 1) MLPS outperforms PSL in both tasks regarding HV. As the number of objectives increases, the performance gap between MLPS and PSL widens. As noted in the PSL study, PSL struggles with high-dimensional tasks, whereas MLPS achieves a better molecular Pareto set. 2) In terms of PM, the gap is also significant. PSL's inability to enhance PM effectively stems from its lack of local search capabilities, which makes it difficult to identify high-quality molecules in complex scenarios. In contrast, MLPS's global-local structure enables it to discover molecules with higher PM.

Objective	QED+SA	GSK3β+JNK3	QED+SA+GSK3β	QED+SA+JNK3	QED+SA+GSK3β+JNK3
GA+D	0.428	0.360	0.610	0.430	0.310
JT-VAE	0.002	0.002	0.063	0.131	0.015
GCPN	0.003	0.002	0.000	0.000	0.000
RationalRL	0.716	0.686	0.270	0.131	0.294
MARS	0.538	0.518	0.680	0.674	0.547
MolSearch	/	0.650	0.702	0.651	0.664
PSL	0.002	0.002	0.001	0.000	0.000
RetMol	0.462	0.436	0.521	0.578	0.611
HN-GFN	0.782	0.712	0.702	0.710	0.646
MLPS	**0.789**	**0.734**	**0.737**	**0.722**	**0.669**

Table 2: Comparison of different methods on PM

Ablation Study. In the ablation study, we aim to assess the contributions of the designed global and local models in MLPS. We compare MLPS with two variants: one that lacks the global model and another that lacks the local models. The results are presented in Figure 2, which illustrates the Hypervolume (HV) values as a function of the number of solutions that are evaluated by their true objective vectors during the MLPS learning process for these three methods in the context of the QED+SA+GSK3β+JNK3 task. MLPS performs better than MLPS without the global model and MLPS without the global model.This suggests that leveraging multiple local optimizations is more effective than relying solely on a single global optimization in the high-dimensional latent space. More details of ablation study are shown in Appendix C.6.

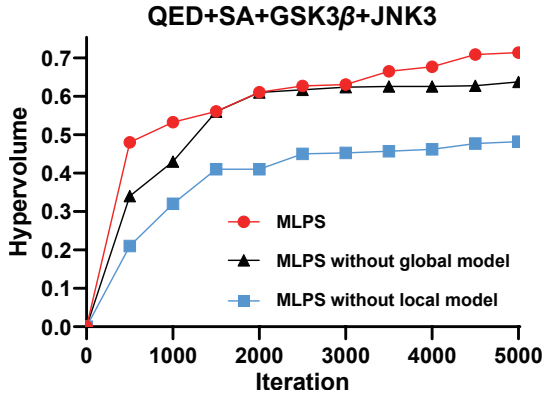

Figure 2: Performance comparison among MLPS and its variants on QED+SA+GSK3β+JNK3.

Real-world Challenging Tasks

We use a challenging drug discovery tasks in the real world to demonstrate the generalization and multi-objective optimization capabilities of MLPS: Antifungal activity + toxicity. Antifungal active peptides with low toxicity, measure the active score against fungal and non-toxic to normal cells.

Evaluation Metrics. For the antifungal peptide optimization task, we chose 5000 antimicrobial peptides as the initial population to optimize them into antifungal peptides with low toxicity. The candidates were scored for activity by the Chemprop model (Jin, Barzilay, and Jaakkola 2020b) and toxicity by the ToxinPred3.0 model (Rathore et al. 2023).

Case Study Results

Optimizing existing antimicrobial peptides towards improved antifungal activity and reduced toxicity. As a motivating use case, fungal infections (e.g., Candida infections, Aspergillus infections, etc.) are a serious threat to human health globally, especially in immunocompromised patients. Antifungal peptides (AFPs), a new antifungal agent with broad-spectrum, high efficiency, and low toxicity, provide a new option for treating fungal infections. Optimal AFPs design requires balancing between multiple, tightly interacting attribute objectives, such as high activity and low toxicity. We used MLPS to optimize antimicrobial peptides for high antifungal activity and reduced toxicity to address this challenge. With 5000 samples for both the original and improved sequences, it is evident that the optimized candidate has significantly increased in activity and significantly decreased in toxicity (Fig.3). This result demonstrates that our approach is efficient and able to optimize the desired properties on real-world challenging tasks as well.

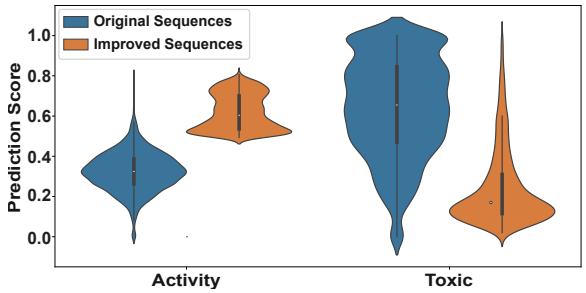

Figure 3: Comparison of the difference in properties between the original sequence and the optimized sequence.

Conclusion

This paper proposed an innovative algorithm called MLPS to address various challenges of multi-objective molecular design. MLPS leverages a combination of global and local optimization models, which empowers decision-makers to efficiently explore the Pareto Set. Our extensive experiments across diverse multi-objective molecular design and peptides scenarios have consistently demonstrated the superiority of MLPS. In future studies, we will try to further enhance the scalability and efficiency of MLPS to tackle multi-objective molecular design problems with a higher number of objectives.

Acknowledgments

This work was supported by the National Science and Technology Major Project (2023ZD0120902), the National Natural Science Foundation of China (U22A2037; 62472152; 62272151; 62106073; 62425204; 62122025; 62450002; 62432011; 62376115; 62372159), Hunan Provincial Natural Science Foundation of China (2024JJ4015; 2022JJ20016), the science and technology innovation Program of Hunan Province (2022RC1099), the Beijing Natural Science Foundation (L248013) and Guangdong Provincial Key Laboratory (Grant No. 2020B121201001).

References

Abdolmaleki, A.; Huang, S.; Hasenclever, L.; Neunert, M.; Song, F.; Zambelli, M.; Martins, M.; Heess, N.; Hadsell, R.; and Riedmiller, M. 2020. A distributional view on multi-objective policy optimization. In *International conference on machine learning*, 11–22. PMLR.

Abdolmaleki, A.; Huang, S. H.; Vezzani, G.; Shahriari, B.; Springenberg, J. T.; Mishra, S.; TB, D.; Byravan, A.; Bousmalis, K.; Gyorgy, A.; et al. 2021. On multi-objective policy optimization as a tool for reinforcement learning. *arXiv preprint arXiv:2106.08199*.

Abels, A.; Roijers, D.; Lenaerts, T.; Nowé, A.; and Steckelmacher, D. 2019. Dynamic weights in multi-objective deep reinforcement learning. In *International conference on machine learning*, 11–20. PMLR.

Ahn, S.; Kim, J.; Lee, H.; and Shin, J. 2020. Guiding deep molecular optimization with genetic exploration. *Advances in neural information processing systems*, 33: 12008–12021.

Balandat, M.; Karrer, B.; Jiang, D.; Daulton, S.; Letham, B.; Wilson, A. G.; and Bakshy, E. 2020. BoTorch: A framework for efficient Monte-Carlo Bayesian optimization. *Advances in neural information processing systems*, 33: 21524–21538.

Buckman, J.; Roy, A.; Raffel, C.; and Goodfellow, I. 2018. Thermometer Encoding: One Hot Way To Resist Adversarial Examples. In *International Conference on Learning Representations*.

Butler, K. T.; Davies, D. W.; Cartwright, H.; Isayev, O.; and Walsh, A. 2018. Machine learning for molecular and materials science. *Nature*, 559(7715): 547–555.

Chen, Z.; Min, M. R.; Parthasarathy, S.; and Ning, X. 2021. A deep generative model for molecule optimization via one fragment modification. *Nature machine intelligence*, 3(12): 1040–1049.

Eriksson, D.; Pearce, M.; Gardner, J.; Turner, R. D.; and Poloczek, M. 2019. Scalable global optimization via local Bayesian optimization. *Advances in neural information processing systems*, 32.

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

Fu, T.; Xiao, C.; Glass, L. M.; and Sun, J. 2021a. MOLER: incorporate molecule-level reward to enhance deep generative model for molecule optimization. *IEEE transactions on knowledge and data engineering*, 34(11): 5459–5471.

Fu, T.; Xiao, C.; Li, X.; Glass, L. M.; and Sun, J. 2021b. Mimosa: Multi-constraint molecule sampling for molecule optimization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, 125–133.

Gao, W.; Fu, T.; Sun, J.; and Coley, C. 2022. Sample efficiency matters: a benchmark for practical molecular optimization. *Advances in Neural Information Processing Systems*, 35: 21342–21357.

Gardner, J.; Pleiss, G.; Weinberger, K. Q.; Bindel, D.; and Wilson, A. G. 2018. Gpytorch: Blackbox matrix-matrix gaussian process inference with gpu acceleration. *Advances in neural information processing systems*, 31.

Gómez-Bombarelli, R.; Wei, J. N.; Duvenaud, D.; Hernández-Lobato, J. M.; Sánchez-Lengeling, B.; Sheberla, D.; Aguilera-Iparraguirre, J.; Hirzel, T. D.; Adams, R. P.; and Aspuru-Guzik, A. 2018. Automatic chemical design using a data-driven continuous representation of molecules. *ACS central science*, 4(2): 268–276.

Guimaraes, G. L.; Sanchez-Lengeling, B.; Outeiral, C.; Farias, P. L. C.; and Aspuru-Guzik, A. 2017. Objective-reinforced generative adversarial networks (organ) for sequence generation models. *arXiv preprint arXiv:1705.10843*.

Ho, J.; Jain, A.; and Abbeel, P. 2020. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33: 6840–6851.

Hoffman, S. C.; Chenthamarakshan, V.; Wadhawan, K.; Chen, P.-Y.; and Das, P. 2022. Optimizing molecules using efficient queries from property evaluations. *Nature Machine Intelligence*, 4(1): 21–31.

Hoogeboom, E.; Satorras, V. G.; Vignac, C.; and Welling, M. 2022. Equivariant diffusion for molecule generation in 3d. In *International conference on machine learning*, 8867–8887. PMLR.

Jin, W.; Barzilay, R.; and Jaakkola, T. 2018. Junction tree variational autoencoder for molecular graph generation. In *International conference on machine learning*, 2323–2332. PMLR.

Jin, W.; Barzilay, R.; and Jaakkola, T. 2020a. Multi-objective molecule generation using interpretable substructures. In *International conference on machine learning*, 4849–4859. PMLR.

Jin, W.; Barzilay, R.; and Jaakkola, T. S. 2020b. Composing Molecules with Multiple Property Constraints. *CoRR*, abs/2002.03244.

Kadurin, A.; Nikolenko, S.; Khrabrov, K.; Aliper, A.; and Zhavoronkov, A. 2017. druGAN: an advanced generative adversarial autoencoder model for de novo generation of new molecules with desired molecular properties in silico. *Molecular pharmaceutics*, 14(9): 3098–3104.

Lim, J.; Ryu, S.; Kim, J. W.; and Kim, W. Y. 2018. Molecular generative model based on conditional variational autoencoder for de novo molecular design. *Journal of cheminformatics*, 10(1): 1–9.

Lin, X.; Yang, Z.; and Zhang, Q. 2021. Pareto Set Learning for Neural Multi-Objective Combinatorial Optimization. In *International Conference on Learning Representations*.

Lin, X.; Yang, Z.; Zhang, X.; and Zhang, Q. 2022. Pareto set learning for expensive multi-objective optimization. *Advances in Neural Information Processing Systems*, 35: 19231–19247.

Liu, Q.; Allamanis, M.; Brockschmidt, M.; and Gaunt, A. 2018. Constrained graph variational autoencoders for molecule design. *Advances in neural information processing systems*, 31.

Meyers, J.; Fabian, B.; and Brown, N. 2021. De novo molecular design and generative models. *Drug Discovery Today*, 26(11): 2707–2715.

Navon, A.; Shamsian, A.; Fetaya, E.; and Chechik, G. 2020. Learning the Pareto Front with Hypernetworks. In *International Conference on Learning Representations*.

Nigam, A.; Friederich, P.; Krenn, M.; and Aspuru-Guzik, A. 2020. Augmenting Genetic Algorithms with Deep Neural Networks for Exploring the Chemical Space. *ICLR 2020*.

Prykhodko, O.; Johansson, S. V.; Kotsias, P.-C.; Arús-Pous, J.; Bjerrum, E. J.; Engkvist, O.; and Chen, H. 2019. A de novo molecular generation method using latent vector based generative adversarial network. *Journal of Cheminformatics*, 11(1): 1–13.

Rathore, A. S.; Arora, A.; Choudhury, S.; Tijare, P.; and Raghava, G. P. S. 2023. ToxinPred 3.0: An improved method for predicting the toxicity of peptides. *bioRxiv*.

Renardy, M.; Joslyn, L. R.; Millar, J. A.; and Kirschner, D. E. 2021. To Sobol or not to Sobol? The effects of sampling schemes in systems biology applications. *Mathematical biosciences*, 337: 108593.

Sener, O.; and Koltun, V. 2018. Multi-task learning as multi-objective optimization. *Advances in neural information processing systems*, 31.

Siivola, E.; Paleyes, A.; González, J.; and Vehtari, A. 2021. Good practices for Bayesian optimization of high dimensional structured spaces. *Applied AI Letters*, 2(2): e24.

Sun, M.; Xing, J.; Meng, H.; Wang, H.; Chen, B.; and Zhou, J. 2022. Molsearch: search-based multi-objective molecular generation and property optimization. In *Proceedings of the 28th ACM SIGKDD conference on knowledge discovery and data mining*, 4724–4732.

SV, S. S.; Law, J. N.; Tripp, C. E.; Duplyakin, D.; Skordilis, E.; Biagioni, D.; Paton, R. S.; and St. John, P. C. 2022. Multi-objective goal-directed optimization of de novo stable organic radicals for aqueous redox flow batteries. *Nature Machine Intelligence*, 4(8): 720–730.

Verhellen, J. 2022. Graph-based molecular Pareto optimisation. *Chemical Science*, 13(25): 7526–7535.

Wan, C.; Duan, X.; and Huang, Y. 2020. Molecular design of single-atom catalysts for oxygen reduction reaction. *Advanced Energy Materials*, 10(14): 1903815.

Wang, Z.; Nie, W.; Qiao, Z.; Xiao, C.; Baraniuk, R.; and Anandkumar, A. 2023. Retrieval-based controllable molecule generation. In *International Conference on Learning Representations*.

Xie, Y.; Shi, C.; Zhou, H.; Yang, Y.; Zhang, W.; Yu, Y.; and Li, L. 2021. MARS: Markov Molecular Sampling for Multi-objective Drug Discovery. In *International Conference on Learning Representations*.

Xu, M.; Yu, L.; Song, Y.; Shi, C.; Ermon, S.; and Tang, J. 2021. GeoDiff: A Geometric Diffusion Model for Molecular Conformation Generation. In *International Conference on Learning Representations*.

Yasonik, J. 2020. Multiobjective de novo drug design with recurrent neural networks and nondominated sorting. *Journal of Cheminformatics*, 12(1): 14.

You, J.; Liu, B.; Ying, Z.; Pande, V.; and Leskovec, J. 2018. Graph convolutional policy network for goal-directed molecular graph generation. *Advances in neural information processing systems*, 31.

Zhu, Y.; Wu, J.; Hu, C.; Yan, J.; Hsieh, C.-Y.; Hou, T.; and Wu, J. 2023. Sample-efficient Multi-objective Molecular Optimization with GFlowNets. In *Thirty-seventh Conference on Neural Information Processing Systems*.