CONVERGE THEN DIVERSIFY: DECOUPLING CONVERGENCE AND DIVERSITY IN MULTI-OBJECTIVE BAYESIAN OPTIMISATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Multi-objective Bayesian optimisation (MOBO) is a sample-efficient approach for optimising expensive black-box functions with multiple objectives. In MOBO, the goal is to adequately approximate the Pareto front; that is, to obtain a high-quality solution set with 1) good convergence (closeness to the Pareto front) and 2) good diversity (spread across the Pareto front). Existing MOBO methods typically aim to accomplish these two tasks simultaneously, i.e., driving the search towards the Pareto front while maintaining a diverse set of nondominated solutions, such that the solutions, ideally, can gradually approach the entire front. When sufficient search budgets are available, this approach is effective. However, considering both convergence and diversity throughout the search is not easy and requires careful design. Under very tight budgets, there may not be enough solutions generated to be able to simultaneously approach the entire Pareto front. To address this issue, this paper proposes a *converge-then-diversify* (CTD) approach that decouples convergence and diversity into two stages. In the first stage, CTD focuses on convergence, aiming to quickly drive the search toward a single point on the Pareto front. In the second stage, CTD focuses on diversity, aiming to spread solutions across the front. We present two simple instantiations of CTD by using widely adopted acquisition functions in the area. Experimental results show that, across all 446 pairwise comparisons, CTD statistically outperforms state-of-the-art methods in 72.9% of the cases, performs equivalently in 21.1%, and is statistically worse in only 6.1%, with the advantage being particularly evident in settings with very tight evaluation budgets or in high-dimensional problems.

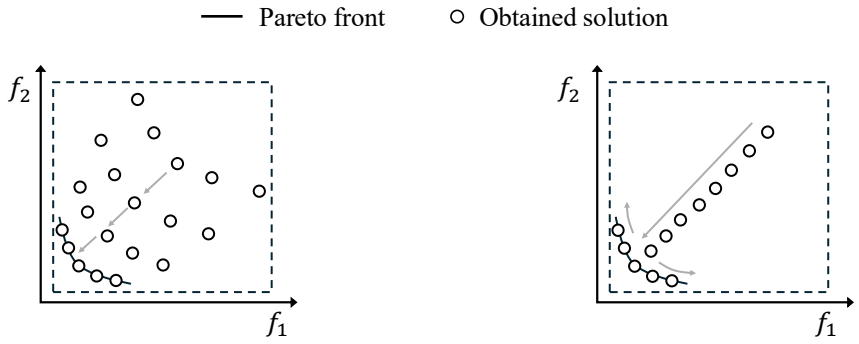
1 INTRODUCTION

Multi-objective Bayesian optimisation (MOBO) is a sample-efficient approach for optimising expensive black-box functions with multiple objectives (Garnett, 2023; Wang et al., 2023b; Karl et al., 2023). MOBO has been used in many areas, such as machine learning (Jang et al., 2024; Wang et al., 2025), physics (Irshad et al., 2021; 2023), and vehicle design (Daulton et al., 2021; Anosri et al., 2023). In MOBO, the goal is to adequately approximate the Pareto front, that is, to obtain a high-quality solution set with both good convergence (closeness to the Pareto front) and good diversity (spread across the Pareto front).

To achieve this goal, various MOBO methods have been proposed. They can be loosely divided into scalarisation-based and Pareto-based methods. Scalarisation-based methods (Knowles, 2006; Paria et al., 2020) convert a multi-objective problem into a number of single-objective subproblems by a set of weight vectors, and then use single-objective techniques to tackle these subproblems (Chugh, 2020). In these methods, different weight vectors are used to guide the search towards different regions of the Pareto front.

Pareto-based methods consider Pareto dominance relations over objectives and directly solve a multi-objective problem without converting it into a number of single-objective subproblems. Representative examples include hypervolume-based and information-based methods. Hypervolume-based methods (Emmerich et al., 2006; Daulton et al., 2020; 2021) evaluate the hypervolume gain produced by each candidate, favouring points that either push the current front towards better objective values

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.



(a) Conventional search in MOBO (b) Proposed convergence-then-diversity search

Figure 1: Illustration of the search processes of the conventional MOBO approach and the proposed approach. (a) The conventional approach that moves towards the Pareto front while simultaneously spreading nondominated solutions. (b) The proposed *convergence-then-diversity* strategy that first moves directly towards the Pareto front and then spreads along it.

or enlarge its coverage over previously nondominated regions. Information-based methods select points that reduce uncertainty about the Pareto front, favouring regions that are both promising and uncertain. These methods differ in the space over which information gain is measured: the input space of Pareto-optimal points (Garrido-Merchán et al., 2023; Garrido-Merchán & Hernández-Lobato, 2019; Hernandez-Lobato et al., 2016), the output space of the Pareto front (Belakaria et al., 2019; 2021; Suzuki et al., 2020; Ishikura & Karasuyama, 2025), or jointly over both (Tu et al., 2022). Apart from hypervolume-based and information-based methods, other Pareto-based methods select points using different criteria, such as maximum uncertainty (Belakaria et al., 2020b) and maximin distance (Renganathan & Carlson, 2025), from a Pareto front obtained by optimising multiple acquisition functions (e.g., expected improvement (Jones et al., 1998) and Thompson sampling (Thompson, 1933)).

All of the above methods, to obtain the Pareto front, consider convergence and diversity simultaneously throughout the search, aiming to approach multiple regions of the front. This strategy is effective when the evaluation budget is sufficient since maintaining a diverse set of solutions helps explore different areas. However, when the budget is very limited, balancing convergence and diversity becomes challenging. The number of evaluations is often insufficient to adequately approach the entire Pareto front. As a result, the search may yield a well-distributed set of nondominated solutions that are far from the true Pareto front.

To address this issue, this paper proposes an approach that decouples convergence from diversity, termed *converge-then-diversify* (CTD). CTD operates in two stages. The first stage focuses on convergence, aiming to quickly drive the search toward a single point on the Pareto front. Once the solutions converge or are close to the front, the search switches to exploring along the front, promoting diversity. We hope that this design enables more efficient use of the evaluation budget while still obtaining a good approximation of the Pareto front. Figure 1 illustrates the proposed two-stage process in comparison with the conventional approach that optimises convergence and diversity simultaneously.

The main contributions of this work are summarised as follows:

- We propose a novel search framework for MOBO under very limited evaluation budgets, which decouples convergence and diversity into separate stages, enabling faster progress toward finding an approximation of the Pareto front.
- We show that the framework is general and flexible: it can be instantiated with different acquisition functions. In particular, we demonstrate its effectiveness when directly working with two widely adopted acquisition functions in the area.
- We validate the proposed framework through an extensive experimental study, comparing it with state-of-the-art methods under noiseless, noisy, and high-dimensional settings on both benchmark and real-world problems.

2 BACKGROUND

Multi-Objective Optimisation. Multi-objective optimisation refers to the process of optimising two or more objectives simultaneously, with the aim of obtaining a set of trade-off solutions rather than a single optimum. Without loss of generality, this paper considers the problem of minimising a vector-valued function: $\boldsymbol{f}(\boldsymbol{x}) : \mathcal{X} \to \mathbb{R}^m$, where $\boldsymbol{x} \in \mathcal{X}$ ($\mathcal{X} \subset \mathbb{R}^d$) and m is the number of objectives. In multi-objective optimisation, a solution $\boldsymbol{x}^*$ is said to dominate $\boldsymbol{x}'$, denoted by $\boldsymbol{x}^* \prec \boldsymbol{x}'$, if $\forall i \in \{1, \dots, m\}$, $f_i(\boldsymbol{x}^*) \leq f_i(\boldsymbol{x}')$ and $\exists j \in \{1, \dots, m\}$, $f_j(\boldsymbol{x}^*) < f_j(\boldsymbol{x}')$. If a solution $\boldsymbol{x}^* \in \mathcal{X}$ is not dominated by any other solution, then $\boldsymbol{x}^*$ is said to be Pareto optimal. The collection of Pareto optimal solutions of a problem is called the Pareto set, and its mapping to the objective space is called Pareto front.

Bayesian Optimisation (BO). BO is a sample-efficient global optimisation method that builds a probabilistic surrogate, typically a Gaussian process (GP), and uses an acquisition function $\alpha : \mathcal{X} \to \mathbb{R}$ to decide which point to evaluate (Garnett, 2023). In this work, we model each objective with an independent Gaussian process $f_i \sim \mathcal{GP}(m_i(\boldsymbol{x}), k_i(\boldsymbol{x}, \boldsymbol{x}'))$, where $m_i : \mathcal{X} \to \mathbb{R}$ is the ith mean function, and $k_i(\cdot, \cdot) : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is the ith covariance function. Given n observed points $\mathcal{D}^{(n)} = \{(\boldsymbol{x}^{(t)}, \boldsymbol{y}^{(t)})\}_{t=1}^n$, where $\boldsymbol{y}^{(t)} = \boldsymbol{f}(\boldsymbol{x}^{(t)}) + \boldsymbol{\zeta}^{(t)}$ and the noise $\boldsymbol{\zeta}^{(t)} \sim \mathcal{N}(\boldsymbol{0}, \mathrm{diag}(\boldsymbol{\sigma}_\zeta^2))$, the posterior distribution of the ith objective at a new location $\boldsymbol{x}$ is Gaussian: $f_i(\boldsymbol{x}) \mid \mathcal{D}^{(n)} \sim \mathcal{N}(\mu_i(\boldsymbol{x}), \sigma_i^2(\boldsymbol{x}))$, where $\mu_i(\boldsymbol{x})$ and $\sigma_i^2(\boldsymbol{x})$ are the posterior mean and variance at $\boldsymbol{x}$, respectively. Detailed expressions of the mean and variance are given in Appendix A.

3 RELATED WORK

Over the past decades, various MOBO methods have been proposed (Zuluaga et al., 2013; Shah & Ghahramani, 2016; Daulton et al., 2022a; Tian et al., 2023; Yang et al., 2026; Rashidi et al., 2024; Wang et al., 2023a; Binois et al., 2020; Picheny et al., 2019; Zhu et al., 2023; Li et al., 2025a). They can be loosely divided into scalarisation-based and Pareto-based methods. In scalarisation-based methods (Knowles, 2006; Paria et al., 2020; Zhang, 2024), a multi-objective problem is converted into a number of single-objective problems (Chugh, 2020; Jiang et al., 2026a). This allows acquisition functions from single-objective BO to be used for selecting the next point to evaluate. In these studies, different acquisition functions are employed, such as probability of improvement (PI) (Kushner, 1964) in (De Ath et al., 2022), expected improvement (EI) (Jones et al., 1998) in (Knowles, 2006; Zhang et al., 2010; Namura et al., 2017; Pires et al., 2026), upper confidence bound (UCB) (Lai & Robbins, 1985) in Paria et al. (2020); Zhang & Golovin (2020); Li et al. (2024); Lin et al. (2022); Li et al. (2025b); Cheng et al. (2026), Thompson sampling (TS) (Thompson, 1933) in Paria et al. (2020); Zhang & Golovin (2020), and knowledge gradient (KG) (Frazier et al., 2008) in (Buckingham et al., 2025).

Pareto-based methods consider Pareto dominance relations over objectives and directly solve a multi-objective problem without converting it into a number of single-objective subproblems. Representative examples include hypervolume-based and information-based methods. Hypervolume-based methods (Ponweiser et al., 2008; Shang et al., 2020; Bhatija et al., 2025; Hung et al., 2025; Qing et al., 2023a; Zhang et al., 2025) are prevalent since maximising the hypervolume value is equivalent to finding the entire Pareto front (Zitzler et al., 2007; Li & Yao, 2019). Many hypervolume-based methods are based on expected hypervolume improvement (EHVI), which extends expected improvement (EI) to the multi-objective setting (Emmerich et al., 2006). EHVI has demonstrated strong performance across a range of settings (Yang et al., 2016; Feliot et al., 2018; Deng et al., 2025), including sequential optimisation (Emmerich et al., 2006; Rahat et al., 2022), batch optimisation (Yang et al., 2019c; Wada & Hino, 2019; Daulton et al., 2020; Lu et al., 2024), constrained optimisation (Abdolshah et al., 2018; de Winter et al., 2022), and noisy environments (Daulton et al., 2021). Substantial efforts have also been devoted to improving its computational efficiency (Couckuyt et al., 2014; Hupkens et al., 2015; Zhao et al., 2018; Yang et al., 2019a;b; Mei & Dawei, 2025). Apart from the EHVI family, hypervolume-based methods have also been developed based on other acquisition functions, including posterior mean (Konakovic Lukovic et al., 2020), probability of improvement (PI) (Wang et al., 2022; Couckuyt et al., 2014), upper confidence bound (UCB) (Roussel et al., 2021; Wang et al., 2026), Thompson sampling (TS) (Bradford et al., 2018; Daulton et al., 2022b), and knowledge gradient (KG) (Daulton et al., 2023).

Algorithm 1: Converge-Then-Diversify (CTD) Framework

Input: $\boldsymbol{f}$: Expensive black-box problem with m objectives; T: Maximum number of iterations;
α_{converge}: Acquisition function in the convergence-focused stage;
$\alpha_{\text{diversify}}$: Acquisition function in the diversity-focused stage;
$\mathcal{D}^{(n_0)} := \{(\boldsymbol{x}^{(t)}, \boldsymbol{y}^{(t)})\}_{t=1}^{n_0}$: Initial observed solution set.

1 **for** $t = n_0 + 1$ **to** T **do**
2 $\quad GPs \leftarrow$ Train $\mathcal{GP}s(\mathcal{D}^{(t-1)})$ `// Train m Gaussian process models`
3 $\quad$ **if** *convergence-focused stage* **then**
4 $\quad\quad \boldsymbol{x}^{(t)} \leftarrow \arg\max_{\boldsymbol{x}\in\mathcal{X}} \alpha_{\text{converge}}(\boldsymbol{x}, GPs)$ `// Convergence-focused search`
$\quad$ **else**
5 $\quad\quad \boldsymbol{x}^{(t)} \leftarrow \arg\max_{\boldsymbol{x}\in\mathcal{X}} \alpha_{\text{diversify}}(\boldsymbol{x}, GPs)$ `// Diversity-focused search`
6 $\quad \boldsymbol{y}^{(t)} \leftarrow \boldsymbol{f}(\boldsymbol{x}^{(t)}) + \boldsymbol{\zeta}^{(t)}$ `// Evaluate the selected solution`
7 $\quad \mathcal{D}^{(t)} \leftarrow \mathcal{D}^{(t-1)} \cup \{(\boldsymbol{x}^{(t)}, \boldsymbol{y}^{(t)})\}$ `// Augment the observed solution set`

Output: $\mathcal{D}^{(T)}$: Observed solution set.

Information-based methods leverage information theory to guide exploration toward regions likely to contribute to the Pareto front (Qing et al., 2023b; James & Thomas, 2024; Garrido-Merchán, 2025). Representative approaches are predictive entropy search (PES), which focuses on the optimal inputs (i.e., the approximated Pareto set) (Garrido-Merchán et al., 2023; Garrido-Merchán & Hernández-Lobato, 2019; Hernandez-Lobato et al., 2016), max-value entropy search (MES), on the optimal outputs (i.e., the approximated Pareto front) (Belakaria et al., 2019; 2021; Suzuki et al., 2020), and joint entropy search (JES), on both of them (Tu et al., 2022; Fernández-Sánchez & Hernández-Lobato, 2025).

Apart from hypervolume-based and information-based methods, other Pareto-based methods select points according to other criteria (Ahmadianshalchi et al., 2024; Park et al., 2024). For example, some methods use maximum uncertainty (Belakaria et al., 2020b), while others use maximin distance (Renganathan & Carlson, 2025) to select points from an approximate Pareto front obtained by optimising multiple acquisition functions for each objective, such as PI, EI, UCB, and TS.

All of the above methods consider convergence and diversity simultaneously across the search, i.e., guiding the search towards the Pareto front while maintaining a diverse set of nondominated solutions.

4 The Proposed Method

Unlike existing MOBO methods, the proposed method decouples the tasks of converging into the Pareto front and diversifying along the Pareto front into two separate stages. In the first stage, it focuses on driving the search towards the Pareto front, and in the second stage, it focuses on spreading the solutions over the Pareto front. Algorithm 1 gives the procedure of the proposed *converge-then-diversify* (CTD) framework. As can be seen in the algorithm, CTD follows the basic procedure of MOBO. For each iteration, it trains m Gaussian process models (Line 2), selects a solution for evaluation (Lines 3–5), and then evaluates the selected solution and augments the observed solution set (Lines 6 and 7). The key steps of CTD are the design in the convergence-focused stage (Line 4), the design in the diversity-focused stage (Line 5), and how to decide switching from convergence-focused search to diversity-focused search (Line 3). Next, we explain them individually in detail.

4.1 Convergence-focused Search

Designing the search strategy in this stage is straightforward. In principle, any acquisition function that can quickly guide the search towards one point of the Pareto front can be used. Here, we consider a well-known scalarisation-based acquisition function, EI based on the augmented Tchebycheff scalarisation (Miettinen, 1999). This scalarisation function has been frequently used in MOBO (Balandat et al., 2020; De Ath et al., 2022; Ozaki et al., 2024), including ParEGO (Knowles, 2006) and PSL (Lin et al., 2022). However, unlike these methods which consider different weight vectors to diversify the search directions for a well-distributed set of nondominated solutions, we consider a fixed weight vector, $(\frac{1}{m}, \ldots, \frac{1}{m}) \in \mathbb{R}^m$, where m denotes the number of objectives. We hope this

setting can enable the search quickly approach one point of the Pareto front, ideally a good trade-off solution among the objectives.

4.2 Diversity-focused Search

After the search converges to a region of the Pareto front, the focus shifts to diversifying the solutions. Note that diversification does not mean encouraging exploration in arbitrary directions - we do not want the search to move back towards regions that are dominated by solutions we found. Instead, the goal is to spread solutions along the Pareto front.

Some existing acquisition functions can serve this purpose. The EI family is suitable since it seeks solutions that are expected to improve upon the current best solution (i.e., in our case, the converged solution already found). Integrated with a criterion that encourages diversity, EI would lead the search to expand from the converged solution. Acquisition functions of this type include hypervolume-based EI, because spreading solutions along the Pareto front always leads to improvements in hypervolume and the hypervolume value is maximised when the entire front is found. Similarly, scalarisation-based EI with diverse weight vectors can also serve the purpose, as it can lead the search to different regions of the Pareto front. In this paper, we adopt a classic hypervolume-based EI, EHVI (Emmerich et al., 2006; Ament et al., 2023). In the experimental study, we additionally consider a classic scalarisation-based EI (i.e., ParEGO) to further validate the proposed framework.

4.3 Switching from Convergence to Diversity

A critical issue in the proposed method is determining when to switch from convergence to diversity, i.e., when to stop the convergence-focused search. Stopping too early may result in solutions that are still far from the Pareto front, whereas stopping too late may lead to unnecessary exploration after the search has already converged. In this regard, the properties of EI come in handy, as a small maximum EI value indicates that further evaluations are unlikely to improve much (Jones, 2001). Existing studies (Jones, 2001; Nguyen et al., 2017) show that when EI is smaller than a small positive constant, the search should be stopped:

Lemma 1 (Nguyen et al. (Nguyen et al., 2017))**.** *The value of the EI acquisition function at the selected point should be positive for a valid optimisation, i.e.,* $\forall \boldsymbol{x}^{(t)} \in \mathcal{D}^{(t)}, \alpha_{EI}^{(t)}(\boldsymbol{x}^{(t)}) \geq \kappa > 0$*, where* κ *is a small positive constant. If this condition is violated, the optimisation should be stopped.*

In our context, this means the search has reached, or is very close to, a point of the Pareto front, since for a weight vector only a Pareto optimal solution can obtain the optimum of the augmented Tchebycheff scalarisation (Steuer, 1986). Specifically, according to (Nguyen et al., 2017), we set $\kappa = 10^{-4}$ (i.e., the maximum EI value is smaller than 10^{-4}) as the switching criterion. In addition, it is worth mentioning that other stopping rules in BO (Lorenz et al., 2015; Frazier et al., 2008; Xie et al., 2025) (e.g., regret-based rules (Makarova et al., 2022; Ishibashi et al., 2023; Wilson, 2024)) may also be considered.

5 Experimental Design

Compared Methods. To evaluate the proposed CTD, we compare it with nine well-established methods: five Pareto-based methods, three scalarisation-based methods, and one baseline, Sobol (Sobol, 1967). The five Pareto-based methods include three HV-based methods, i.e., EHVI (Daulton et al., 2020; Ament et al., 2023) (along with its noisy variant NEHVI (Daulton et al., 2021)), MORBO (Daulton et al., 2022b), and MOBO-OSD (Ngo et al., 2025); one information-based method, JES (Tu et al., 2022); and one maximin-distance-based method, qPOTS (Renganathan & Carlson, 2025). The three scalarisation-based methods are ParEGO (Knowles, 2006; Ament et al., 2023) (along with its noisy variant NParEGO (Daulton et al., 2021)), TS-TCH (Paria et al., 2020), and PSL (Lin et al., 2022). For all EI-based methods, namely EHVI, NEHVI, ParEGO, NParEGO, and CTD, we use their log variants for numerical stability, as suggested in (Ament et al., 2023). Detailed implementations of our method and the peer methods can be found in Appendix B.1.

Benchmarks and Real-World Problems. For benchmark problems, we choose the most widely used scalable functions, DTLZ1–DTLZ7 (Deb et al., 2005). Each problem is considered with 2 and 3 objectives, following the practice in (Deb & Jain, 2013; Jain & Deb, 2013; Li et al., 2014). We also

Table 1: The HV results (mean and standard deviation) of the ten methods on the 20 benchmark and real-world problems. The method with the best mean HV is highlighted in bold. The symbols "−", "∼" and "+" indicate that our method CTD is statistically worse than, equivalent to, and better than the competitor, respectively.

Method	DTLZ1 ($m=2$)		DTLZ2 ($m=2$)		DTLZ3 ($m=2$)		DTLZ4 ($m=2$)		DTLZ5 ($m=2$)		DTLZ6 ($m=2$)		DTLZ7 ($m=2$)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Sobol	0.00e+00	0.00e+00 (∼)	1.23e-02	1.82e-02 (+)	0.00e+00	0.00e+00 (∼)	1.34e-02	2.21e-02 (+)	1.20e-02	1.33e-02 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)
ParEGO	0.00e+00	0.00e+00 (∼)	2.97e-01	7.15e-02 (+)	1.94e+02	5.61e+02 (∼)	2.84e-01	6.52e-02 (+)	2.39e-01	1.06e-01 (+)	2.67e-01	1.34e-01 (∼)	3.86e-01	2.38e-01 (−)
TS-TCH	0.00e+00	0.00e+00 (∼)	3.68e-02	3.99e-02 (+)	0.00e+00	0.00e+00 (∼)	3.47e-02	3.43e-02 (+)	4.32e-02	4.18e-02 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)
PSL	**7.92e+00**	**2.82e+01** (∼)	8.61e-02	5.44e-02 (+)	2.66e+02	7.12e+02 (∼)	9.88e-02	3.64e-02 (+)	8.61e-02	5.44e-02 (+)	0.00e+00	0.00e+00 (+)	4.64e-01	1.71e-01 (−)
JES	0.00e+00	0.00e+00 (∼)	2.43e-01	6.00e-02 (+)	5.37e+02	6.34e+02 (∼)	2.65e-01	5.99e-02 (+)	2.74e-01	4.40e-02 (+)	2.22e-01	7.26e-02 (+)	3.72e-01	1.86e-01 (−)
EHVI	0.00e+00	0.00e+00 (∼)	1.63e-01	1.15e-01 (+)	5.54e+02	9.03e+02 (∼)	1.25e-01	9.14e-02 (+)	1.32e-01	1.15e-01 (+)	2.26e-01	3.17e-02 (+)	1.69e-01	2.30e-02 (∼)
MORBO	0.00e+00	0.00e+00 (∼)	4.07e-02	3.46e-02 (+)	0.00e+00	0.00e+00 (∼)	3.91e-02	3.80e-02 (+)	4.98e-02	3.98e-02 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)
MOBO-OSD	0.00e+00	0.00e+00 (∼)	1.60e-01	6.90e-02 (+)	**8.62e+02**	**8.90e+02** (∼)	1.10e-01	1.06e-01 (+)	1.54e-01	5.76e-02 (+)	1.57e-01	8.31e-02 (+)	**7.19e-01**	**2.86e-02** (−)
qPOTS	0.00e+00	0.00e+00 (∼)	6.94e-03	1.19e-02 (+)	0.00e+00	0.00e+00 (∼)	3.99e-03	9.05e-03 (+)	4.09e-03	8.35e-03 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)
CTD (ours)	6.68e+00	2.43e+01	**3.46e-01**	**2.09e-02**	4.17e+02	8.64e+02	**3.52e-01**	**1.46e-02**	**3.49e-01**	**1.95e-02**	**3.19e-01**	**4.51e-02**	1.71e-01	2.43e-02

Method	DTLZ1 ($m=3$)		DTLZ2 ($m=3$)		DTLZ3 ($m=3$)		DTLZ4 ($m=3$)		DTLZ5 ($m=3$)		DTLZ6 ($m=3$)		DTLZ7 ($m=3$)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Sobol	1.19e+03	3.27e+03 (∼)	2.90e-02	2.10e-02 (+)	0.00e+00	0.00e+00 (+)	3.69e-02	2.42e-02 (+)	1.12e-03	1.87e-03 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)
ParEGO	4.31e+03	6.79e+03 (∼)	1.40e-01	1.20e-01 (+)	9.97e+05	8.31e+05 (∼)	9.85e-02	1.12e-01 (+)	6.85e-02	3.30e-02 (+)	8.18e-02	3.32e-02 (+)	4.87e-01	2.72e-01 (−)
TS-TCH	9.97e+02	4.40e+03 (+)	3.23e-02	2.83e-02 (+)	0.00e+00	0.00e+00 (+)	2.56e-02	2.31e-02 (+)	3.30e-03	6.49e-03 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)
PSL	5.58e+03	6.89e+03 (∼)	6.66e-02	6.60e-03 (+)	1.80e+06	6.38e+05 (∼)	5.71e-02	5.03e-02 (+)	5.81e-02	2.42e-02 (+)	0.00e+00	0.00e+00 (+)	2.23e-01	1.25e-01 (−)
JES	8.38e+03	5.49e+03 (−)	1.40e-01	1.04e-01 (+)	1.77e+06	6.85e+05 (∼)	1.78e-01	1.23e-01 (+)	5.04e-02	3.76e-02 (+)	6.77e-02	2.40e-02 (+)	3.93e-01	1.66e-01 (−)
EHVI	9.41e+03	5.68e+03 (−)	4.56e-02	5.17e-02 (+)	1.88e+06	7.78e+05 (∼)	4.70e-02	5.37e-02 (+)	3.20e-02	3.18e-02 (+)	6.67e-02	4.78e-03 (+)	2.23e-01	1.73e-02 (−)
MORBO	9.50e+02	4.52e+03 (+)	1.95e-02	2.32e-02 (+)	0.00e+00	0.00e+00 (+)	3.06e-02	3.52e-02 (+)	1.22e-03	2.35e-03 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)
MOBO-OSD	**1.25e+04**	**4.79e+03** (−)	1.77e-02	1.98e-02 (+)	**2.68e+06**	**1.01e+06** (−)	1.52e-02	6.79e-02 (+)	1.86e-02	2.78e-02 (+)	2.68e-02	2.68e-02 (+)	**8.11e-01**	**7.21e-02** (−)
qPOTS	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)
CTD (ours)	4.29e+03	5.30e+03	**3.73e-01**	**7.95e-02**	1.58e+06	8.91e+05	**3.60e-01**	**9.01e-02**	**1.02e-01**	**1.30e-02**	**1.01e-01**	**1.73e-02**	2.22e-01	1.90e-02

Method	Four bar truss design		Pressure vessel design		Hatch cover design		Vehicle safety		Car side impact		LPA		Summary
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	−/∼/+
Sobol	4.45e+01	6.61e-01 (+)	3.90e+09	1.87e+09 (+)	2.01e+04	7.43e+02 (+)	1.48e+01	9.38e-01 (+)	2.25e+02	5.17e+00 (+)	2.53e+07	2.65e+06 (+)	0/3/17
ParEGO	5.19e+01	1.34e+00 (+)	1.10e+10	5.81e+07 (∼)	2.18e+04	3.04e+01 (+)	2.62e+01	5.52e-01 (∼)	2.96e+02	1.01e+02 (+)	4.12e+07	3.08e+06 (+)	4/5/11
TS-TCH	5.07e+01	1.32e+00 (+)	1.03e+10	2.79e+08 (+)	2.16e+04	8.35e+01 (+)	2.26e+01	3.76e-01 (+)	2.54e+02	8.64e+01 (+)	3.75e+07	3.11e+06 (+)	0/2/18
PSL	5.39e+01	3.87e-02 (+)	1.09e+10	1.72e+08 (∼)	2.17e+04	1.42e+01 (+)	2.37e+01	1.50e+00 (+)	2.77e+02	1.09e+01 (+)	**4.81e+07**	**2.13e+06** (∼)	2/6/12
JES	5.15e+01	1.62e+00 (+)	1.08e+10	1.39e+08 (+)	2.17e+04	4.07e+01 (+)	**2.66e+01**	**4.17e-01** (−)	**3.27e+02**	**9.44e+00** (−)	4.03e+07	4.05e+06 (+)	5/3/12
EHVI	5.40e+01	9.20e-02 (+)	1.09e+10	1.51e+08 (∼)	**2.18e+04**	**1.83e+01** (∼)	2.54e+01	8.40e-01 (+)	2.79e+02	9.52e+01 (∼)	4.75e+07	3.02e+06 (∼)	2/8/10
MORBO	5.06e+01	1.23e+00 (+)	1.03e+10	3.62e+08 (+)	2.16e+04	9.12e+01 (+)	2.24e+01	4.56e-01 (+)	2.83e+02	5.65e+00 (+)	3.73e+07	2.77e+06 (+)	0/2/18
MOBO-OSD	5.20e+01	8.56e-01 (+)	**1.10e+10**	**2.77e+07** (∼)	2.13e+04	3.62e+02 (+)	2.35e+01	1.14e+00 (+)	2.78e+02	9.90e+00 (+)	3.65e+07	3.96e+06 (+)	4/3/13
qPOTS	3.97e+01	1.31e+00 (+)	1.08e+10	1.93e+08 (+)	1.85e+04	4.03e+03 (+)	1.30e+01	2.37e+00 (+)	2.43e+02	1.06e+01 (+)	3.85e+07	2.16e+06 (+)	0/2/18
CTD (ours)	**5.41e+01**	**2.82e-02**	1.09e+10	1.79e+08	2.18e+04	2.05e+01	2.60e+01	6.38e-01	3.05e+02	1.55e+01	4.62e+07	3.71e+06	−/−/−

consider six well-studied real-world problems (Tanabe & Ishibuchi, 2020; Irshad et al., 2023), i.e., four bar truss design (Cheng & Li, 1999), pressure vessel design (Kannan & Kramer, 1994), hatch cover design (Amir & Hasegawa, 1989), vehicle safety design (Liao et al., 2008), car side impact design (Jain & Deb, 2013), and laser plasma acceleration (LPA) (Irshad et al., 2021; 2023). These problems are widely used in the MOBO literature (Bradford et al., 2018; Daulton et al., 2020; 2021; Lin et al., 2022; Ament et al., 2023; Renganathan & Carlson, 2025; Yang et al., 2026). The details of the problem formulations are given in Appendix B.2. We test the above problems under both noiseless and noisy cases. To make their noise cases, we include additive zero-mean Gaussian noise with a standard deviation set to approximately 10% of the objective ranges, as suggested in Daulton et al. (2021); Tu et al. (2022).

6 Experimental Results

6.1 How Does CTD Perform Compared with Well-Established Methods?

Noiseless Cases. We begin our evaluation in the noiseless setting. We set the dimensionality to $d = m + 4$ for DTLZ1, $d = m + 9$ for DTLZ2–DTLZ6, and $d = m + 19$ for DTLZ7, according to the original paper (Deb et al., 2005). For all the methods, we consider a very limited budget, allowing a maximum of 100 evaluations, which is not an uncommon setting in the area (e.g., Daulton et al. (2020); Ip et al. (2025)).

Table 1 shows the HV results (mean and standard deviation) of our method and the nine peer methods on the 20 benchmark and real-world problems. As shown in the table, our method demonstrates a clear performance advantage, performing statistically better, equivalently, and worse in 129, 34, and 17 out of the 180 pairwise comparisons, respectively. Despite its strong overall performance, CTD performs relatively poorly on DTLZ7, which has a disconnected Pareto front. This is mainly because EHVI is used during the diversity-focused search, as EHVI itself struggles to explore the entire Pareto front of DTLZ7 (Chugh, 2022). That said, CTD can, in principle, handle different Pareto-front structures because it does not restrict the search space during the diversity-focused stage. Evidence of this can be found in the pressure-vessel design problem, which also has a disconnected Pareto front (Tanabe & Ishibuchi, 2020). On this problem, where EHVI performs well, CTD is among the best-performing methods.

To help understand their search behaviour, Figure 2 shows the hypervolume trajectories with respect to the number of evaluations of the ten methods on the problem DTLZ2 with $m = 2$. As shown in

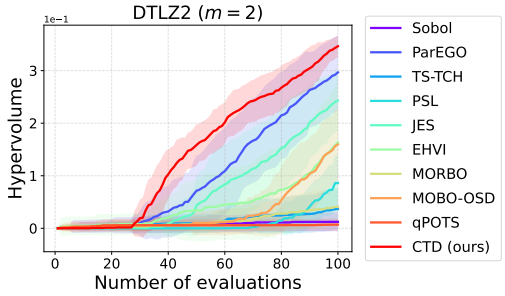
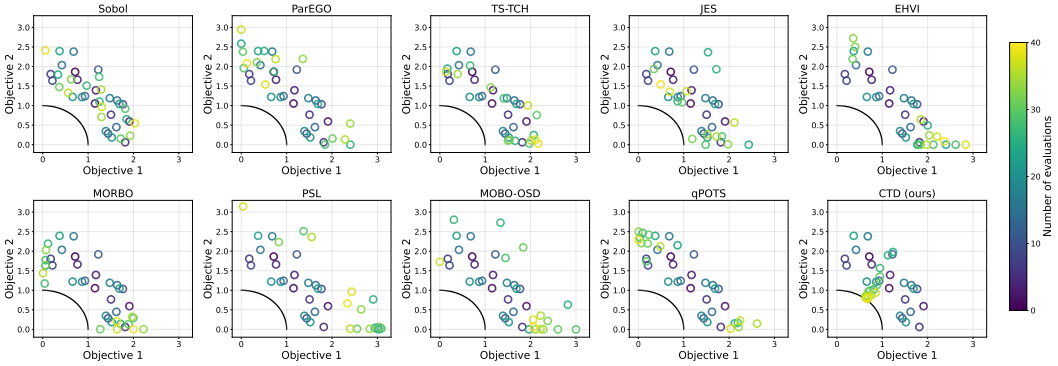
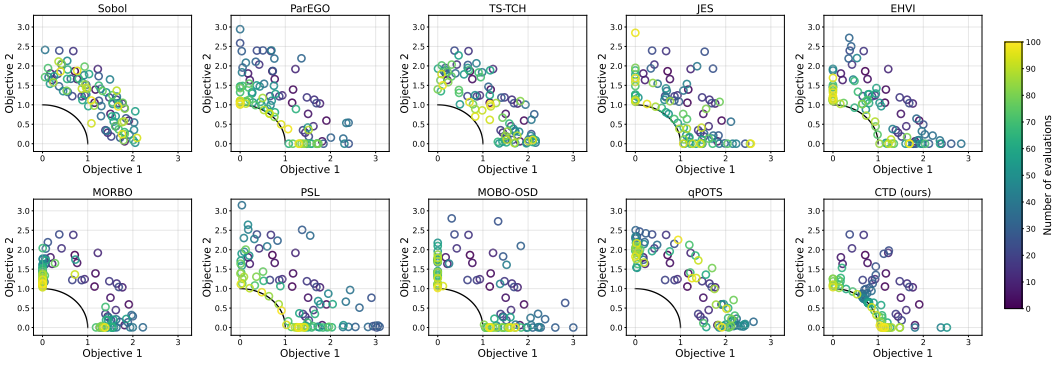


Figure 2: Mean hypervolume trajectories of the ten methods on the representative DTLZ2 problem ($m = 2$, $d = 11$) over 30 independent runs. The shaded regions indicate one standard deviation around the mean.



(a) Solutions obtained by the proposed CTD and the nine peer methods under 40 evaluations.



(b) Solutions obtained by the proposed CTD and the nine peer methods under 100 evaluations.

Figure 3: Solutions obtained by the proposed CTD and the nine peer methods on DTLZ2 ($m = 2$, $d = 11$) under budgets of 40 and 100 function evaluations. All methods start from the same $2(d + 1)$ initial points generated by a scrambled Sobol sequence. The black line denotes the Pareto front, and brighter colours indicate solutions obtained at later evaluations.

the figure, since around 30 evaluations (notably, the first 24 evaluations are initial samples), CTD exhibits a clear advantage, obtaining substantially higher hypervolume values than its competitors.

To visually understand the results, Figure 3 shows the solutions obtained by each method on DTLZ2 ($m = 2$, $d = 11$), where for facilitating comparison, all methods start from the same $2(d + 1)$ initial points generated by a scrambled Sobol sequence. As shown in the plot (a), the convergence-focused search enables CTD to approach the Pareto front under a very limited budget, while the other methods remain distant from the Pareto front. Furthermore, after approaching the Pareto front, CTD is able to spread solutions along the front and achieve good diversity (plot (b) in Figure 3).

Noisy Cases. Then, we consider noisy settings. Here, we do not include PSL (Lin et al., 2022) and MOBO-OSD (Ngo et al., 2025), as their original studies focus on noiseless settings. For EI-based methods, i.e., ParEGO, EHVI, and CTD, we use their noisy variants (Ament et al., 2023; Daulton

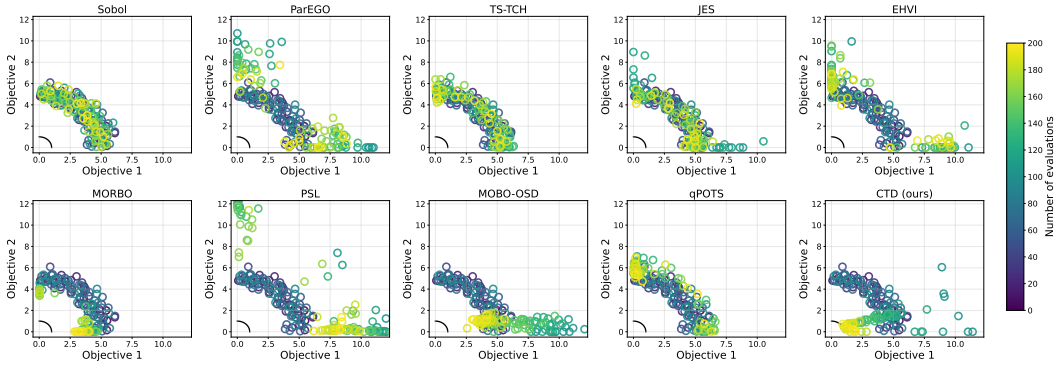


Figure 4: Solutions obtained by the proposed CTD and the nine peer methods on DTLZ2 ($m = 2$, $d = 50$) under 200 evaluations. All methods start from the same $2(d + 1)$ initial points generated by a scrambled Sobol sequence. The black line denotes the Pareto front, and brighter colours indicate solutions obtained at later evaluations.

et al., 2021). We evaluate the proposed CTD on the same 20 benchmark and real-world problems as in the noiseless setting but with observation noise.

The HV results (mean and standard deviation) of each method can be found in Appendix C.1.2. As can be seen in Table 3, similar to the noiseless setting, CTD demonstrates a clear performance advantage over its competitors. In total, it performs statistically better, equivalently and worse in 85, 52 and 3 out of the 140 pairwise comparisons, respectively.

High-Dimensional Problems. We now investigate how the proposed CTD performs on high-dimensional problems. In this experiment, we include a new MOBO algorithm dedicated to high-dimensional problems, MORBO (Daulton et al., 2022b). We consider the scalable problem suite DTLZ, which are frequently used in the high-dimensional MOBO literature (Daulton et al., 2022b; Rashidi et al., 2024). We set the dimensionality $d = 50$ for DTLZ1–DTLZ7 with $m = 2$ and $m = 3$ objectives. Note that since our initial design contains $2(d + 1)$ points, i.e., 102 points, which already exceeds the budget of 100 evaluations used in the previous experiments, we allow a total budget of 200 evaluations, following the practice in (Daulton et al., 2020; 2021; Konakovic Lukovic et al., 2020).

The HV results (mean and standard deviation) of each method can be found in Appendix C.1.3. As shown, the advantage of the proposed CTD is even clearer in high-dimensional cases. It performs statistically better, equivalently and worse in 111, 8 and 7 out of the 126 pairwise comparisons, respectively. This suggests that prioritising convergence before diversity is effective when the search space becomes substantially larger. We observe that CTD statistically outperforms the high-dimensional method MORBO on all the problems. One possible reason for the weaker performance of MORBO is that the method may be better suited to settings with generous evaluation budgets, e.g., 2,000 evaluations (Daulton et al., 2022b).

To visually understand the results, Figure 4 shows the solutions obtained by each method on DTLZ2 ($m = 2$, $d = 50$), where all methods start from the same $2(d + 1)$ initial points generated by a scrambled Sobol sequence. As shown in the figure, the proposed CTD is able to approach the Pareto front, while the other methods remain distant from the Pareto front. This indicates that the proposed convergence-first strategy is particularly effective in high-dimensional problems under the limited evaluation budget.

It is worth mentioning that the results obtained by the proposed method can echo recent findings in single-objective Bayesian optimisation, where standard Gaussian processes with scaled priors can perform well in high-dimensional spaces (Hvarfner et al., 2024; Papenmeier et al., 2025; Xu et al., 2025) even without using high-dimensional handling techniques (Binois & Wycoff, 2022; González-Duque et al., 2024; Papenmeier et al., 2023; Santoni et al., 2024; Wang et al., 2016; Hoang et al., 2025; Doumont et al., 2025) (e.g., variable selection (Eriksson & Jankowiak, 2021), additive models (Delbridge et al., 2020; Han et al., 2021; Wang et al., 2018; Ziomek & Ammar, 2023),

embeddings (Antonov et al., 2022; Letham et al., 2020; Raponi et al., 2020), and trust regions (Daulton et al., 2022b; Diouane et al., 2023; Eriksson et al., 2019)).

6.2 Sensitivity Analysis

Within the proposed method CTD, the threshold κ is used to control when the search switches to the diversity-focused stage by comparing it with the maximum EI value. A large κ may result in solutions that are still far from the Pareto front before the switch, whereas a small κ may lead to unnecessary exploration after the search has already converged.

In this work, we set $\kappa = 10^{-4}$, as suggested by (Nguyen et al., 2017). The results show that this setting allows the search to obtain well-converged solutions and can switch to diversity-focused search without unnecessary exploration (see Figure 3). We further now evaluate the sensitivity of κ and set $\kappa \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-5}, 10^{-6}\}$. The HV results (mean and standard deviation) of each method can be found in Appendix C.2. As can be seen in Table 7, CTD is relatively insensitive to κ within the range from 10^{-3} to 10^{-5}. Compared with the default setting $\kappa = 10^{-4}$, both $\kappa = 10^{-3}$ and $\kappa = 10^{-5}$ achieve statistically equivalent results on 19 out of the 20 problems. By contrast, the more extreme settings generally lead to poorer performance. Setting κ to 10^{-1} or 10^{-2} may trigger the diversity-focused stage while the solutions are still far from the Pareto front, resulting in statistically worse performance on six and five problems, respectively. Meanwhile, setting $\kappa = 10^{-6}$ may consume unnecessary evaluations after the search has already converged, resulting in statistically worse performance on eight problems.

6.3 Using an Alternative Acquisition Function

In the proposed CTD framework, we employ EHVI as the acquisition function during the diversity-focused search. In this section, we want to investigate whether other acquisition functions also work. We now consider EI with random augmented Tchebycheff scalarisations (used in ParEGO (Knowles, 2006)) in the diversity-focused search, denoted as CTD$_{tch}$.

The HV results (mean and standard deviation) of each method can be found in Appendix C.3. Table 9 shows the results (mean and standard deviation) of the HV obtained by the proposed CTD with different acquisition functions (EI based on HV and random scalarisations) in the diversity-focused search across 20 benchmark and real-world problems under the noiseless setting. Interestingly, CTD with random scalarisations (CTD$_{tch}$) is highly competitive, performing statistically better, equivalently, and worse than the original CTD on 6, 11, and 3 out of the 20 problems, respectively. Notably, CTD$_{tch}$ improves the performance on some problems (e.g., DTLZ7) where the EHVI-based CTD performs relatively poorly (see Table 1). This suggests that CTD itself does not restrict global exploration, while a more suitable acquisition function can significantly improve its performance. The results for the noisy and high-dimensional settings, shown in Tables 10 and 11, respectively, are consistent with those observed in the noiseless setting. Specifically, CTD$_{tch}$ is statistically better than, equivalent to, and worse than the original CTD on 4, 11, and 5 out of the 20 noisy problems, respectively, and on 7, 4, and 3 out of the 14 high-dimensional problems, respectively.

7 Conclusion

This work presented a new MOBO framework, CTD, that prioritises convergence before diversity, rather than balancing both throughout the search. We evaluated CTD on a range of benchmark and real-world problems, covering noiseless, noisy, and high-dimensional settings. The results show that CTD performs effectively across these scenarios. We used EHVI as the default acquisition function in the diversity-focused stage, and found that using EI based on random augmented Tchebycheff scalarisations (i.e., ParEGO) can achieve even better performance. Other acquisition functions for the diversity-focused stage could be further explored in future work. For the convergence-focused stage and the switching criterion, we used conventional methods: EI with the augmented Tchebycheff scalarisation and an EI-based stopping rule, respectively. Future work could explore other methods, including alternative acquisition functions (e.g., UCB (Lai & Robbins, 1985)) and scalarisations (e.g., smooth Tchebycheff scalarisation (Lin et al., 2024)) for the convergence-focused search, as well as other BO stopping rules (Lorenz et al., 2015; Frazier et al., 2008; Xie et al., 2025; Makarova et al., 2022; Ishibashi et al., 2023; Wilson, 2024) for determining the switching time.

AI USE STATEMENT

Generative AI tools were used only for language editing, e.g., grammar, spelling, word choice.

REFERENCES

Majid Abdolshah, Alistair Shilton, Santu Rana, Sunil Gupta, and Svetha Venkatesh. Expected hypervolume improvement with constraints. In *2018 24th International Conference on Pattern Recognition (ICPR)*, pp. 3238–3243. IEEE, 2018.

Alaleh Ahmadianshalchi, Syrine Belakaria, and Janardhan Rao Doppa. Pareto front-diverse batch multi-objective Bayesian optimization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 10784–10794, 2024.

Sebastian Ament, Samuel Daulton, David Eriksson, Maximilian Balandat, and Eytan Bakshy. Unexpected improvements to expected improvement for Bayesian optimization. In *Advances in Neural Information Processing Systems*, volume 36, pp. 20577–20612. Curran Associates, Inc., 2023.

Sasan Amini and Inneke Van Nieuwenhuyse. A Bayesian optimization algorithm for constrained simulation optimization problems with heteroscedastic noise. In *International Conference on Learning and Intelligent Optimization*, pp. 78–91. Springer, 2023.

Sasan Amini, Inneke Vannieuwenhuyse, and Alejandro Morales-Hernández. Constrained Bayesian optimization: A review. *IEEE Access*, 13:1581–1593, 2024.

Hossain M Amir and Takashi Hasegawa. Nonlinear mixed-discrete structural optimization. *Journal of Structural Engineering*, 115(3):626–646, 1989.

Siwakorn Anosri, Natee Panagant, Pakin Champasak, Sujin Bureerat, Chinnapat Thipyopas, Sumit Kumar, Nantiwat Pholdee, Betül Sultan Yıldız, and Ali Riza Yildiz. A comparative study of state-of-the-art metaheuristics for solving many-objective optimization problems of fixed wing unmanned aerial vehicle conceptual design. *Archives of Computational Methods in Engineering*, 30(6):3657–3671, 2023.

Kirill Antonov, Elena Raponi, Hao Wang, and Carola Doerr. High dimensional Bayesian optimization with kernel principal component analysis. In *International Conference on Parallel Problem Solving from Nature*, pp. 118–131. Springer, 2022.

Setareh Ariafar, Jaume Coll-Font, Dana Brooks, and Jennifer Dy. Admmbo: Bayesian optimization with unknown constraints using admm. *Journal of Machine Learning Research*, 20(123):1–26, 2019.

Maximilian Balandat, Brian Karrer, Daniel Jiang, Samuel Daulton, Ben Letham, Andrew G Wilson, and Eytan Bakshy. Botorch: A framework for efficient monte-carlo Bayesian optimization. In *Advances in Neural Information Processing Systems*, volume 33, pp. 21524–21538. Curran Associates, Inc., 2020.

Syrine Belakaria, Aryan Deshwal, and Janardhan Rao Doppa. Max-value entropy search for multi-objective Bayesian optimization. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

Syrine Belakaria, Aryan Deshwal, and Janardhan Rao Doppa. Multi-fidelity multi-objective Bayesian optimization: An output space entropy search approach. *Proceedings of the AAAI Conference on artificial intelligence*, 34(06):10035–10043, 2020a.

Syrine Belakaria, Aryan Deshwal, Nitthilan Kannappan Jayakodi, and Janardhan Rao Doppa. Uncertainty-aware search framework for multi-objective Bayesian optimization. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(06):10044–10052, 2020b.

Syrine Belakaria, Aryan Deshwal, and Janardhan Rao Doppa. Output space entropy search framework for multi-objective Bayesian optimization. *Journal of artificial intelligence research*, 72:667–715, 2021.

Shriya Bhatija, Paul-David Zuercher, Jakob Thumm, and Thomas Bohné. Multi-objective causal Bayesian optimization. *arXiv preprint arXiv:2502.14755*, 2025.

Mickael Binois and Nathan Wycoff. A survey on high-dimensional Gaussian process modeling with application to Bayesian optimization. *ACM Transactions on Evolutionary Learning and Optimization*, 2(2):1–26, 2022.

Mickael Binois, Victor Picheny, Patrick Taillandier, and Abderrahmane Habbal. The kalai-smorodinsky solution for many-objective Bayesian optimization. *Journal of Machine Learning Research*, 21(150):1–42, 2020.

Eric Bradford, Artur M Schweidtmann, and Alexei Lapkin. Efficient multiobjective optimization employing gaussian processes, spectral sampling and a genetic algorithm. *Journal of global optimization*, 71(2):407–438, 2018.

Jack M Buckingham, Sebastian Rojas-Gonzalez, and Juergen Branke. Knowledge gradient for multi-objective Bayesian optimization with decoupled evaluations. In *International Conference on Evolutionary Multi-Criterion Optimization*, pp. 117–132. Springer, 2025.

Franklin Y Cheng and Xiang-Song Li. Generalized center method for multiobjective engineering optimization. *Engineering Optimization*, 31(5):641–661, 1999.

Ji Cheng, Bo Xue, and Qingfu Zhang. Parametric pareto set learning for expensive multi-objective optimization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 36829–36837, 2026.

Tinkle Chugh. Scalarizing functions in Bayesian multiobjective optimization. In *2020 IEEE Congress on Evolutionary Computation (CEC)*, pp. 1–8. IEEE, 2020.

Tinkle Chugh. Mono-surrogate vs multi-surrogate in multi-objective Bayesian optimisation. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, pp. 2143–2151, 2022.

Ivo Couckuyt, Dirk Deschrijver, and Tom Dhaene. Fast calculation of multiobjective probability of improvement and expected improvement criteria for Pareto optimization. *Journal of Global Optimization*, 60(3):575–594, 2014.

Sam Daulton, Maximilian Balandat, and Eytan Bakshy. Hypervolume knowledge gradient: A lookahead approach for multi-objective Bayesian optimization with partial information. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 7167–7204. PMLR, 2023.

Samuel Daulton, Maximilian Balandat, and Eytan Bakshy. Differentiable expected hypervolume improvement for parallel multi-objective Bayesian optimization. In *Advances in Neural Information Processing Systems*, volume 33, pp. 9851–9864. Curran Associates, Inc., 2020.

Samuel Daulton, Maximilian Balandat, and Eytan Bakshy. Parallel Bayesian optimization of multiple noisy objectives with expected hypervolume improvement. In *Advances in Neural Information Processing Systems*, volume 34. Curran Associates, Inc., 2021.

Samuel Daulton, Sait Cakmak, Maximilian Balandat, Michael A Osborne, Enlu Zhou, and Eytan Bakshy. Robust multi-objective Bayesian optimization under input noise. In *International Conference on Machine Learning*, pp. 4831–4866. PMLR, 2022a.

Samuel Daulton, David Eriksson, Maximilian Balandat, and Eytan Bakshy. Multi-objective Bayesian optimization over high-dimensional search spaces. In *Uncertainty in Artificial Intelligence*, pp. 507–517. PMLR, 2022b.

George De Ath, Tinkle Chugh, and Alma AM Rahat. MBORE: multi-objective Bayesian optimisation by density-ratio estimation. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 776–785, 2022.

Roy de Winter, Bas van Stein, and Thomas Bäck. Multi-point acquisition function for constraint parallel efficient multi-objective optimization. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 511–519, 2022.

Kalyanmoy Deb and Himanshu Jain. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: solving problems with box constraints. *IEEE transactions on Evolutionary Computation*, 18(4):577–601, 2013.

Kalyanmoy Deb, Lothar Thiele, Marco Laumanns, and Eckart Zitzler. Scalable test problems for evolutionary multiobjective optimization. In *Evolutionary multiobjective optimization: theoretical advances and applications*, pp. 105–145. Springer, 2005.

Ian Delbridge, David Bindel, and Andrew Gordon Wilson. Randomly projected additive Gaussian processes for regression. In *International Conference on Machine Learning*, pp. 2453–2463. PMLR, 2020.

Jingda Deng, Jianyong Sun, Qingfu Zhang, and Hui Li. Expected hypervolume improvement is a particular hypervolume improvement. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(15):16217–16225, 2025.

Youssef Diouane, Victor Picheny, Rodolophe Le Riche, and Alexandre Scotto Di Perrotolo. Trego: a trust-region framework for efficient global optimization. *Journal of Global Optimization*, 86(1): 1–23, 2023.

Colin Doumont, Donney Fan, Natalie Maus, Jacob R Gardner, Henry Moss, and Geoff Pleiss. We still don't understand high-dimensional Bayesian optimization. *arXiv preprint arXiv:2512.00170*, 2025.

Michael TM Emmerich, Kyriakos C Giannakoglou, and Boris Naujoks. Single-and multiobjective evolutionary optimization assisted by Gaussian random field metamodels. *IEEE Transactions on Evolutionary Computation*, 10(4):421–439, 2006.

David Eriksson and Martin Jankowiak. High-dimensional Bayesian optimization with sparse axis-aligned subspaces. In *Uncertainty in Artificial Intelligence*, pp. 493–503. PMLR, 2021.

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

Paul Feliot, Julien Bect, and Emmanuel Vazquez. User preferences in bayesian multi-objective optimization: The expected weighted hypervolume improvement criterion. In *Proceedings of the 4th International Conference on Machine Learning, Optimization, and Data Science (LOD 2018)*, pp. 533–544. Springer, 2018. doi: 10.1007/978-3-030-13709-0_45.

Daniel Fernández-Sánchez and Daniel Hernández-Lobato. Joint entropy search for multi-objective Bayesian optimization with constraints and multiple fidelities. *Neurocomputing*, pp. 131674, 2025.

Peter I Frazier, Warren B Powell, and Savas Dayanik. A knowledge-gradient policy for sequential information collection. *SIAM Journal on Control and Optimization*, 47(5):2410–2439, 2008.

Jacob Gardner, Geoff Pleiss, Kilian Q Weinberger, David Bindel, and Andrew G Wilson. Gpytorch: Blackbox matrix-matrix Gaussian process inference with GPU acceleration. In *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.

Roman Garnett. *Bayesian optimization*. Cambridge University Press, 2023.

Eduardo C Garrido-Merchán. Information-theoretic Bayesian optimization: survey and tutorial. *arXiv preprint arXiv:2502.06789*, 2025.

Eduardo C Garrido-Merchán, Daniel Fernández-Sánchez, and Daniel Hernández-Lobato. Parallel predictive entropy search for multi-objective Bayesian optimization with constraints applied to the tuning of machine learning algorithms. *Expert Systems with Applications*, 215:119328, 2023.

Eduardo C. Garrido-Merchán and Daniel Hernández-Lobato. Predictive entropy search for multi-objective Bayesian optimization with constraints. *Neurocomputing*, 361:50–68, 2019.

Michael A Gelbart, Jasper Snoek, and Ryan P Adams. Bayesian optimization with unknown constraints. *arXiv preprint arXiv:1403.5607*, 2014.

Miguel González-Duque, Richard Michael, Simon Bartels, Yevgen Zainchkovskyy, Sø ren Hauberg, and Wouter Boomsma. A survey and benchmark of high-dimensional Bayesian optimization of discrete sequences. In *Advances in Neural Information Processing Systems*, volume 37, pp. 140478–140508. Curran Associates, Inc., 2024.

Eric Han, Ishank Arora, and Jonathan Scarlett. High-dimensional Bayesian optimization via tree-structured additive models. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(9): 7630–7638, 2021.

Daniel Hernandez-Lobato, Jose Hernandez-Lobato, Amar Shah, and Ryan Adams. Predictive entropy search for multi-objective Bayesian optimization. In *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pp. 1492–1501, New York, USA, 2016. PMLR.

Vu Viet Hoang, Hung The Tran, Sunil Gupta, and Vu Nguyen. High dimensional Bayesian optimization using lasso variable selection. *arXiv preprint arXiv:2504.01743*, 2025.

Sture Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, pp. 65–70, 1979.

Yu-Heng Hung, Kai-Jie Lin, Yu-Heng Lin, Chien-Yi Wang, Cheng Sun, and Ping-Chun Hsieh. Boformer: Learning to solve multi-objective Bayesian optimization via non-markovian RL. *arXiv preprint arXiv:2505.21974*, 2025.

Iris Hupkens, André Deutz, Kaifeng Yang, and Michael Emmerich. Faster exact algorithms for computing expected hypervolume improvement. In *international conference on evolutionary multi-criterion optimization*, pp. 65–79. Springer, 2015.

Carl Hvarfner, Erik Orm Hellsten, and Luigi Nardi. Vanilla Bayesian optimization performs great in high dimensions. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235, pp. 20793–20817. PMLR, 21–27 Jul 2024.

Joshua Hang Sai Ip, Ankush Chakrabarty, Ali Mesbah, and Diego Romeres. User preference meets Pareto-optimality in multi-objective Bayesian optimization. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(19):20246–20254, 2025.

Faran Irshad, Stefan Karsch, and Andreas Döpp. Expected hypervolume improvement for simultaneous multi-objective and multi-fidelity optimization. *arXiv preprint arXiv:2112.13901*, 10, 2021.

Faran Irshad, Stefan Karsch, and Andreas Döpp. Multi-objective and multi-fidelity Bayesian optimization of laser-plasma acceleration. *Physical Review Research*, 5(1):013063, 2023.

Hideaki Ishibashi, Masayuki Karasuyama, Ichiro Takeuchi, and Hideitsu Hino. A stopping criterion for Bayesian optimization by the gap of expected minimum simple regrets. In *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, pp. 6463–6497. PMLR, 25–27 Apr 2023.

Masanori Ishikura and Masayuki Karasuyama. Pareto-frontier entropy search with variational lower bound maximization. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=9UxEydnM1N.

Himanshu Jain and Kalyanmoy Deb. An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach, part ii: Handling constraints and extending to an adaptive approach. *IEEE Transactions on evolutionary computation*, 18(4):602–622, 2013.

Thomas James and Sinnu Thomas. Multi objective regionalized Bayesian optimization via entropy search. In *OPT 2024: Optimization for Machine Learning*, 2024.

Chaeyun Jang, Hyungi Lee, Jungtaek Kim, and Juho Lee. Model fusion through Bayesian optimization in language model fine-tuning. *Advances in Neural Information Processing Systems*, 37: 29878–29912, 2024.

Chao Jiang and Miqing Li. Multi-objectivising acquisition functions in Bayesian optimisation. *ACM Transactions on Evolutionary Learning and Optimization*, 5(2), 2025a.

Chao Jiang and Miqing Li. Trading off quality and uncertainty through multi-objective optimisation in batch Bayesian optimisation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39 (25):27027–27035, 2025b.

Chao Jiang, Jingyu Huang, and Miqing Li. Do we really need to approach the entire pareto front in many-objective Bayesian optimisation? *arXiv preprint arXiv:2604.09417*, 2026a.

Chao Jiang, Yulong Ye, Tao Chen, and Miqing Li. When to use which? benchmarking optimisers for configurable systems under varying budgets. *arXiv preprint arXiv:2607.16476*, 2026b.

Donald R Jones. A taxonomy of global optimization methods based on response surfaces. *Journal of global optimization*, 21(4):345–383, 2001.

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

Kirthevasan Kandasamy, Gautam Dasarathy, Jeff Schneider, and Barnabás Póczos. Multi-fidelity Bayesian optimisation with continuous approximations. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 1799–1808. PMLR, 2017.

BK Kannan and Steven N Kramer. An augmented lagrange multiplier based method for mixed integer discrete continuous optimization and its applications to mechanical design. *Journal of mechanical design*, 116(2):405–411, 1994.

Florian Karl, Tobias Pielok, Julia Moosbauer, Florian Pfisterer, Stefan Coors, Martin Binder, Lennart Schneider, Janek Thomas, Jakob Richter, Michel Lang, et al. Multi-objective hyperparameter optimization in machine learning—an overview. *ACM Transactions on Evolutionary Learning and Optimization*, 3(4):1–50, 2023.

Joshua Knowles. ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE transactions on Evolutionary Computation*, 10(1): 50–66, 2006.

Mina Konakovic Lukovic, Yunsheng Tian, and Wojciech Matusik. Diversity-guided multi-objective Bayesian optimization with batch evaluations. In *Advances in Neural Information Processing Systems*, volume 33, pp. 17708–17720. Curran Associates, Inc., 2020.

Harold J Kushner. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. *Journal Basic Engineering*, 86(1):97–106, 1964.

Tze Leung Lai and Herbert Robbins. Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics*, 6(1):4–22, 1985.

Ben Letham, Roberto Calandra, Akshara Rai, and Eytan Bakshy. Re-examining linear embeddings for high-dimensional Bayesian optimization. In *Advances in Neural Information Processing Systems*, volume 33, pp. 1546–1558. Curran Associates, Inc., 2020.

Benjamin Letham, Brian Karrer, Guilherme Ottoni, and Eytan Bakshy. Constrained Bayesian optimization with noisy experiments. *Bayesian Analysis*, 14(2):495–519, 2019.

Bingdong Li, Zixiang Di, Yongfan Lu, Hong Qian, Feng Wang, Peng Yang, Ke Tang, and Aimin Zhou. Expensive multi-objective Bayesian optimization based on diffusion models. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(25):27063–27071, 2025a.

Diantong Li, Fengxue Zhang, Chong Liu, and Yuxin Chen. Constrained multi-objective Bayesian optimization through optimistic constraints estimation. *arXiv preprint arXiv:2411.03641*, 2024.

Diantong Li, Fengxue Zhang, Chong Liu, and Yuxin Chen. Constrained multi-objective Bayesian optimization through optimistic constraints estimation. In *International Conference on Artificial Intelligence and Statistics*, pp. 370–378. PMLR, 2025b.

Ke Li, Kalyanmoy Deb, Qingfu Zhang, and Sam Kwong. An evolutionary many-objective optimization algorithm based on dominance and decomposition. *IEEE transactions on Evolutionary Computation*, 19(5):694–716, 2014.

Miqing Li and Xin Yao. Quality evaluation of solution sets in multiobjective optimisation: A survey. *ACM Computing Surveys (CSUR)*, 52(2):1–38, 2019.

Shibo Li, Wei Xing, Robert Kirby, and Shandian Zhe. Multi-fidelity Bayesian optimization via deep neural networks. In *Advances in Neural Information Processing Systems*, volume 33, pp. 8521–8531. Curran Associates, Inc., 2020.

Xingtao Liao, Qing Li, Xujing Yang, Weigang Zhang, and Wei Li. Multiobjective optimization for crash safety design of vehicles using stepwise regression model. *Structural and multidisciplinary optimization*, 35(6):561–569, 2008.

Xi Lin, Zhiyuan Yang, Xiaoyuan Zhang, and Qingfu Zhang. Pareto set learning for expensive multi-objective optimization. In *Advances in Neural Information Processing Systems*, volume 35, pp. 19231–19247. Curran Associates, Inc., 2022.

Xi Lin, Xiaoyuan Zhang, Zhiyuan Yang, Fei Liu, Zhenkun Wang, and Qingfu Zhang. Smooth tchebycheff scalarization for multi-objective optimization. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 30479–30509. PMLR, 2024.

Romy Lorenz, Ricardo P Monti, Ines R Violante, Aldo A Faisal, Christoforos Anagnostopoulos, Robert Leech, and Giovanni Montana. Stopping criteria for boosting automatic experimental design using real-time fmri with Bayesian optimization. *arXiv preprint arXiv:1511.07827*, 2015.

Yongfan Lu, Bingdong Li, and Aimin Zhou. Are you concerned about limited function evaluations: Data-augmented pareto set learning for expensive multi-objective optimization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 14202–14210, 2024.

Anastasia Makarova, Huibin Shen, Valerio Perrone, Aaron Klein, Jean Baptiste Faddoul, Andreas Krause, Matthias Seeger, and Cedric Archambeau. Automatic termination for hyperparameter optimization. In *International Conference on Automated Machine Learning*, pp. 7–1. PMLR, 2022.

Li Mei and Zhan Dawei. Pointwise expected hypervolume improvement for expensive multi-objective optimization. *Journal of Global Optimization*, 91(1):171–197, 2025.

Kaisa Miettinen. *Nonlinear Multiobjective Optimization*, volume 12. Springer Science & Business Media, 1999.

Jonas Mockus, Vytautas Tiesis, and Antanas Zilinskas. The application of Bayesian methods for seeking the extremum. *Towards Global Optimization*, 2(2):117–129, 1978.

Henry B. Moss, David S. Leslie, Javier Gonzalez, and Paul Rayson. Gibbon: General-purpose information-based Bayesian optimisation. *Journal of Machine Learning Research*, 22(235):1–49, 2021.

Nobuo Namura, Koji Shimoyama, and Shigeru Obayashi. Expected improvement of penalty-based boundary intersection for expensive multiobjective optimization. *IEEE Transactions on Evolutionary Computation*, 21(6):898–913, 2017.

Lam Ngo, Huong Ha, Jeffrey Chan, and Hongyu Zhang. Mobo-osd: Batch multi-objective Bayesian optimization via orthogonal search directions. *Advances in neural information processing systems*, 2025.

Vu Nguyen, Sunil Gupta, Santu Rana, Cheng Li, and Svetha Venkatesh. Regret for expected improvement over the best-observed value and stopping condition. In *Asian Conference on Machine Learning*, pp. 279–294. PMLR, 2017.

Ryota Ozaki, Kazuki Ishikawa, Youhei Kanzaki, Shion Takeno, Ichiro Takeuchi, and Masayuki Karasuyama. Multi-objective Bayesian optimization with active preference learning. *Proceedings of the AAAI conference on artificial intelligence*, 38(13):14490–14498, 2024.

Leonard Papenmeier, Luigi Nardi, and Matthias Poloczek. Bounce: Reliable high-dimensional Bayesian optimization for combinatorial and mixed spaces. In *Advances in Neural Information Processing Systems*, volume 36, pp. 1764–1793. Curran Associates, Inc., 2023.

Leonard Papenmeier, Matthias Poloczek, and Luigi Nardi. Understanding high-dimensional Bayesian optimization. *arXiv preprint arXiv:2502.09198*, 2025.

Biswajit Paria, Kirthevasan Kandasamy, and Barnabás Póczos. A flexible framework for multi-objective Bayesian optimization using random scalarizations. In Ryan P. Adams and Vibhav Gogate (eds.), *Proceedings of The 35th Uncertainty in Artificial Intelligence Conference*, volume 115 of *Proceedings of Machine Learning Research*, pp. 766–776. PMLR, 2020.

Ji Won Park, Natasa Tagasovska, Michael Maser, Stephen Ra, and Kyunghyun Cho. BOtied: Multi-objective Bayesian optimization with tied multivariate ranks. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 39813–39833. PMLR, 2024.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

Victor Picheny, Robert B Gramacy, Stefan Wild, and Sebastien Le Digabel. Bayesian optimization under mixed constraints with a slack-variable augmented lagrangian. *Advances in neural information processing systems*, 29, 2016.

Victor Picheny, Mickael Binois, and Abderrahmane Habbal. A Bayesian optimization approach to find nash equilibria. *Journal of Global Optimization*, 73:171–192, 2019.

Nogueira Pires, RP Cardoso Coelho, and FM Andrade Pires. Composite Bayesian optimisation for multi-objective problems with smooth Tchebycheff scalarisation. *Structural and Multidisciplinary Optimization*, 69(2), 2026.

Wolfgang Ponweiser, Tobias Wagner, Dirk Biermann, and Markus Vincze. Multiobjective optimization on a limited budget of evaluations using model-assisted $\mathcal{S}$-metric selection. In *Parallel Problem Solving from Nature – PPSN X*, pp. 784–794. Springer, 2008.

Tony Pourmohamad and Herbert KH Lee. Bayesian optimization via barrier functions. *Journal of Computational and Graphical Statistics*, 31(1):74–83, 2022.

Jixiang Qing, Ivo Couckuyt, and Tom Dhaene. A robust multi-objective Bayesian optimization framework considering input uncertainty. *Journal of Global Optimization*, 86(3):693–711, 2023a.

Jixiang Qing, Henry B. Moss, Tom Dhaene, and Ivo Couckuyt. {PF}^{2}ES: Parallel feasible pareto frontier entropy search for multi-objective Bayesian optimization. In *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, pp. 2565–2588. PMLR, 2023b.

Alma Rahat, Tinkle Chugh, Jonathan Fieldsend, Richard Allmendinger, and Kaisa Miettinen. Efficient approximation of expected hypervolume improvement using gauss-hermite quadrature. In *International Conference on Parallel Problem Solving from Nature*, pp. 90–103. Springer, 2022.

Elena Raponi, Hao Wang, Mariusz Bujny, Simonetta Boria, and Carola Doerr. High dimensional Bayesian optimization assisted by principal component analysis. In *Parallel Problem Solving from Nature–PPSN XVI: 16th International Conference, PPSN 2020, Leiden, The Netherlands, September 5-9, 2020, Proceedings, Part I 16*, pp. 169–183. Springer, 2020.

Bahador Rashidi, Kerrick Johnstonbaugh, and Chao Gao. Cylindrical thompson sampling for high-dimensional Bayesian optimization. In *International Conference on Artificial Intelligence and Statistics*, pp. 3502–3510. PMLR, 2024.

Ashwin Renganathan and Kade Carlson. qPOTS: Efficient batch multiobjective Bayesian optimization via pareto optimal Thompson sampling. In *Proceedings of The 28th International Conference on Artificial Intelligence and Statistics*, volume 258 of *Proceedings of Machine Learning Research*, pp. 4051–4059. PMLR, 03–05 May 2025.

Ryan Roussel, Adi Hanuka, and Auralee Edelen. Multiobjective Bayesian optimization for online accelerator tuning. *Physical Review Accelerators and Beams*, 24(6):062801, 2021.

Maria Laura Santoni, Elena Raponi, Renato De Leone, and Carola Doerr. Comparison of high-dimensional Bayesian optimization algorithms on BBOB. *ACM Transactions on Evolutionary Learning and Optimisation*, 4(3):1–33, 2024.

Michael J Sasena, Panos Papalambros, and Pierre Goovaerts. Exploration of metamodeling sampling criteria for constrained global optimization. *Engineering optimization*, 34(3):263–278, 2002.

Amar Shah and Zoubin Ghahramani. Pareto frontier learning with expensive correlated objectives. In *International conference on machine learning*, pp. 1919–1927. PMLR, 2016.

Ke Shang, Hisao Ishibuchi, Linjun He, and Lie Meng Pang. A survey on the hypervolume indicator in evolutionary multiobjective optimization. *IEEE Transactions on Evolutionary Computation*, 25 (1):1–20, 2020.

Ilya M Sobol. The distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational Mathematics and Mathematical Physics*, 7(4):86–112, 1967.

Jialin Song, Yuxin Chen, and Yisong Yue. A general framework for multi-fidelity Bayesian optimization with Gaussian processes. In *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, volume 89 of *Proceedings of Machine Learning Research*, pp. 3158–3167. PMLR, 2019.

Ralph E Steuer. Multiple criteria optimization: Theory, computation, and application, 1986.

Shinya Suzuki, Shion Takeno, Tomoyuki Tamura, Kazuki Shitara, and Masayuki Karasuyama. Multi-objective Bayesian optimization using Pareto-frontier entropy. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 9279–9288. PMLR, 2020.

Shion Takeno, Hitoshi Fukuoka, Yuhki Tsukada, Toshiyuki Koyama, Motoki Shiga, Ichiro Takeuchi, and Masayuki Karasuyama. Multi-fidelity Bayesian optimization with max-value entropy search and its parallelization. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 9334–9345. PMLR, 2020.

Ryoji Tanabe and Hisao Ishibuchi. An easy-to-use real-world multi-objective optimization problem suite. *Applied Soft Computing*, 89:106078, 2020.

William R Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3/4):285–294, 1933.

Yunsheng Tian, Pavle Vanja Konakovic, Beichen Li, Ane Zuniga, Michael Foshey, Timothy Erps, Wojciech Matusik, and Mina Konakovic Lukovic. Autodex: Automated optimal design of experiments platform with data-and time-efficient multi-objective optimization. In *NeurIPS 2023 Workshop on Adaptive Experimental Design and Active Learning in the Real World*, 2023.

Ben Tu, Axel Gandy, Nikolas Kantas, and Behrang Shafei. Joint entropy search for multi-objective Bayesian optimization. In *Advances in Neural Information Processing Systems*, volume 35, pp. 9922–9938. Curran Associates, Inc., 2022.

Takashi Wada and Hideitsu Hino. Bayesian optimization for multi-objective optimization and multi-point search. *arXiv preprint arXiv:1905.02370*, 2019.

Hao Wang, Kaifeng Yang, and Michael Affenzeller. Probability distribution of hypervolume improvement in bi-objective Bayesian optimization. *arXiv preprint arXiv:2205.05505*, 2022.

Hongyan Wang, Hua Xu, and Zeqiu Zhang. High-dimensional multi-objective Bayesian optimization with block coordinate updates: Case studies in intelligent transportation system. *IEEE Transactions on Intelligent Transportation Systems*, 25(1):884–895, 2023a.

Wenyu Wang, Zheyi Fan, and Szu Hui Ng. A trajectory-based Bayesian approach to multi-objective hyperparameter optimization with epoch-aware trade-offs. In *Conference on Uncertainty in Artificial Intelligence*, pp. 4370–4394. PMLR, 2025.

Xilu Wang, Yaochu Jin, Sebastian Schmitt, and Markus Olhofer. Recent advances in Bayesian optimization. *ACM Computing Surveys*, 55(13s):1–36, 2023b.

Yanhong Wang, Yungai Tang, Cong-Feng Wu, and Guangyao Feng. Multiobjective Bayesian optimization for the shape design of rf cavity in particle accelerators. *Physical Review Accelerators and Beams*, 29(3):034601, 2026.

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

Ziyu Wang, Frank Hutter, Masrour Zoghi, David Matheson, and Nando De Feitas. Bayesian optimization in a billion dimensions via random embeddings. *Journal of Artificial Intelligence Research*, 55:361–387, 2016.

Frank Wilcoxon. Individual comparisons by ranking methods. In *Breakthroughs in Statistics: Methodology and Distribution*, pp. 196–202. Springer, 1992.

James T Wilson. Stopping Bayesian optimization with probabilistic regret bounds. *Advances in Neural Information Processing Systems*, 37:98264–98296, 2024.

Jian Wu, Saul Toscano-Palmerin, Peter I. Frazier, and Andrew Gordon Wilson. Practical multi-fidelity Bayesian optimization for hyperparameter tuning. In *Proceedings of The 35th Uncertainty in Artificial Intelligence Conference*, volume 115 of *Proceedings of Machine Learning Research*, pp. 788–798. PMLR, 22–25 Jul 2020.

Qian Xie, Linda Cai, Alexander Terenin, Peter I Frazier, and Ziv Scully. Cost-aware stopping for Bayesian optimization. *arXiv preprint arXiv:2507.12453*, 2025.

Wenjie Xu, Yuning Jiang, Bratislav Svetozarevic, and Colin Jones. Constrained efficient global optimization of expensive black-box functions. In *International Conference on Machine Learning*, pp. 38485–38498. PMLR, 2023.

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

Kaifeng Yang, Andre Deutz, Zhiwei Yang, Thomas Back, and Michael Emmerich. Truncated expected hypervolume improvement: Exact computation and application. In *2016 IEEE Congress on Evolutionary Computation (CEC)*, pp. 4350–4357. IEEE, 2016.

Kaifeng Yang, Michael Emmerich, André Deutz, and Thomas Bäck. Efficient computation of expected hypervolume improvement using box decomposition algorithms. *Journal of Global Optimization*, 75(1):3–34, 2019a.

Kaifeng Yang, Michael Emmerich, André Deutz, and Thomas Bäck. Multi-objective Bayesian global optimization using expected hypervolume improvement gradient. *Swarm and Evolutionary Computation*, 44:945–956, 2019b.

Kaifeng Yang, Pramudita Satria Palar, Michael Emmerich, Koji Shimoyama, and Thomas Bäck. A multi-point mechanism of expected hypervolume improvement for parallel multi-objective Bayesian global optimization. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 656–663, 2019c.

Yaohong Yang, Sammie Katt, and Samuel Kaski. Multi-objective Bayesian optimization via adaptive ε-constraints decomposition. *arXiv preprint arXiv:2604.15959*, 2026.

Qingfu Zhang, Wudong Liu, Edward Tsang, and Botond Virginas. Expensive multiobjective optimization by MOEA/D with Gaussian process model. *IEEE Transactions on Evolutionary Computation*, 14(3):456–474, 2010.

Qiuyi (Richard) Zhang. Optimal scalarizations for sublinear hypervolume regret. In *Proceedings of the 38th International Conference on Neural Information Processing Systems*, NIPS '24, Red Hook, NY, USA, 2024. Curran Associates Inc. ISBN 9798331314385.

Richard Zhang and Daniel Golovin. Random hypervolume scalarizations for provable multi-objective black box optimization. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 11096–11105. PMLR, 2020.

Xinyu Zhang, Conor Hassan, Julien Martinelli, Daolang Huang, and Samuel Kaski. In-context multi-objective optimization. *arXiv preprint arXiv:2512.11114*, 2025.

Yehong Zhang, Trong Nghia Hoang, Bryan Kian Hsiang Low, and Mohan Kankanhalli. Information-based multi-fidelity Bayesian optimization. In *NIPS workshop on Bayesian optimization*, volume 49. Journal of Machine Learning Research JMLR. org Cambridge, MA, 2017.

Guang Zhao, Raymundo Arroyave, and Xiaoning Qian. Fast exact computation of expected hypervolume improvement. *arXiv preprint arXiv:1812.07692*, 2018.

Yiheng Zhu, Jialu Wu, Chaowen Hu, Jiahuan Yan, Tingjun Hou, Jian Wu, et al. Sample-efficient multi-objective molecular optimization with gflownets. *Advances in Neural Information Processing Systems*, 36:79667–79684, 2023.

Juliusz Krzysztof Ziomek and Haitham Bou Ammar. Are random decompositions all we need in high dimensional Bayesian optimisation? In *International Conference on Machine Learning*, pp. 43347–43368. PMLR, 2023.

Eckart Zitzler and Lothar Thiele. Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach. *IEEE Transactions on Evolutionary Computation*, 3(4):257–271, 1999.

Eckart Zitzler, Dimo Brockhoff, and Lothar Thiele. The hypervolume indicator revisited: On the design of pareto-compliant indicators via weighted integration. In *International conference on evolutionary multi-criterion optimization*, pp. 862–876. Springer, 2007.

Marcela Zuluaga, Guillaume Sergent, Andreas Krause, and Markus Püschel. Active learning for multi-objective optimization. In *International conference on Machine Learning*, pp. 462–470. PMLR, 2013.

Appendix to:

Converge Then Diversify: Decoupling Convergence and Diversity in Multi-Objective Bayesian Optimisation

A Multi-Objective Bayesian Optimisation (MOBO)

MOBO consists of two main steps, i.e., training m Gaussian process models based on the observed solutions and optimising an acquisition function $\alpha : \mathcal{X} \to \mathbb{R}$ to select a solution for evaluation. In this work, we model each objective with an independent Gaussian process $f_i \sim \mathcal{GP}(m_i(\boldsymbol{x}), k_i(\boldsymbol{x}, \boldsymbol{x}'))$, where $m_i : \mathcal{X} \to \mathbb{R}$ is the ith mean function, and $k_i(\cdot, \cdot) : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is the ith covariance function. We use the notation $K(\mathrm{A}, \mathrm{B})$ to represent the covariance matrix at all pairs of solutions in set A and in set B. Given n observed points $\mathcal{D}^{(n)} = \{(\boldsymbol{x}^{(t)}, \boldsymbol{y}^{(t)})\}_{t=1}^{n}$, where $\boldsymbol{y}^{(t)} = \boldsymbol{f}(\boldsymbol{x}^{(t)}) + \boldsymbol{\zeta}^{(t)}$ and the noise $\boldsymbol{\zeta}^{(t)} \sim \mathcal{N}(\mathbf{0}, \mathrm{diag}(\boldsymbol{\sigma}_\zeta^2))$, the posterior distribution of the ith objective at a new location $\boldsymbol{x}$ is Gaussian:

$$f_i(\boldsymbol{x}) \mid \mathcal{D}^{(n)} \sim \mathcal{N}(\mu_i(\boldsymbol{x}), \sigma_i^2(\boldsymbol{x})) \tag{1}$$

$$\mu_i(\boldsymbol{x}) = K(\boldsymbol{x}, X^{(n)})(K(X^{(n)}, X^{(n)}) + \sigma_{\zeta_i}^2 \mathbf{I})^{-1} Y_i^{(n)} \tag{2}$$

$$\sigma_i^2(\boldsymbol{x}) = K(\boldsymbol{x}, \boldsymbol{x}) - K(\boldsymbol{x}, X^{(n)})(K(X^{(n)}, X^{(n)}) + \sigma_{\zeta_i}^2 \mathbf{I})^{-1} K(X^{(n)}, \boldsymbol{x}) \tag{3}$$

where $\mu_i(\boldsymbol{x})$ and $\sigma_i^2(\boldsymbol{x})$ are the mean and variance at $\boldsymbol{x}$, respectively; $X^{(n)} = (\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(n)}) \in \mathbb{R}^{n \times d}$ and $Y_i^{(n)} = (y_i^{(1)}, \ldots, y_i^{(n)}) \in \mathbb{R}^n$ are the matrix of evaluated solutions and the corresponding vector of y values, respectively; $\sigma_{\zeta_i}^2$ is the variance of the observation noise $\zeta_i \sim \mathcal{N}(0, \sigma_{\zeta_i}^2)$, and corresponds to the i-th diagonal entry of the noise covariance matrix $\mathrm{diag}(\boldsymbol{\sigma}_\zeta^2)$;

Based on the posterior means and variances provided by the GP models, an acquisition function is optimised to determine the next solution to evaluate. Expected Improvement (EI) (Mockus et al., 1978; Jones et al., 1998; Ament et al., 2023; Jiang & Li, 2025a;b; Jiang et al., 2026b) is one of the most prevalent acquisition functions. Considering a single-objective case, given the best observed value y^*, the improvement over y^* is defined as $I(x) = \max(y^* - \mu(x), 0)$. Then the EI can be expressed as: $\alpha_{EI}(x) = \mathbb{E}[I(x)] = \sigma(x)(\lambda\Phi(\lambda) + \phi(\lambda))$, where $\lambda = \frac{y^* - \mu(x)}{\sigma(x)}$; Φ and ϕ are the standard normal cumulative distribution function and the probability density function, respectively. EI can be readily extended to multi-objective cases. Representative examples include scalarisation-based EI (Knowles, 2006; Zhang et al., 2010) and hypervolume-based EI (Emmerich et al., 2006; Daulton et al., 2020).

B Experiment Settings

B.1 Implementation details

The computational studies were performed on a Red Hat Enterprise Linux 8.8 system, operating on a 64-bit x86 CPU architecture. The computing cluster utilised Intel Xeon Platinum 8360Y processors running at 2.40 GHz. The implementation details are as follows.

CTD. We implement CTD based on the open-source Python framework BoTorch (Balandat et al., 2020), which builds on GPyTorch (Gardner et al., 2018) for Gaussian process modelling and PyTorch (Paszke et al., 2019) for automatic differentiation. For the convergence-focused search, we use LogEI (Ament et al., 2023) based on the augmented Tchebycheff scalarisation (Miettinen, 1999) with a fixed weight vector $(\frac{1}{m}, \ldots, \frac{1}{m}) \in \mathbb{R}^m$. For the diversity-focused search, we use LogEHVI (Ament

et al., 2023) as the main implementation and LogEI (Ament et al., 2023) with random scalarisations as an alternative diversity-focused variant. In noisy cases, we replace LogEI with LogNEI and LogEHVI with LogNEHVI (Ament et al., 2023; Daulton et al., 2021). For the switching criterion, we set $\kappa = 10^{-4}$ (i.e., the maximum EI value is smaller than 10^{-4}), according to (Nguyen et al., 2017). Since LogEI (Ament et al., 2023) is used in the implementation, the corresponding threshold is applied on the log scale, i.e., $\log(\kappa)$. The data and code are available at an anonymised repository: `https://anonymous.4open.science/r/CTD-7C25/README.md`.

ParEGO (Knowles, 2006) and NParEGO (Daulton et al., 2021). We implement ParEGO and NParEGO based on the implementation provided by BoTorch.[1] Both methods use the augmented Tchebycheff scalarisation function and their log version, following the recommendation in (Ament et al., 2023).

TS-TCH (Paria et al., 2020). We implement TS-TCH based on the implementation provided by BoTorch.[2] The augmented Tchebycheff scalarisation function is used. After converting to single-objective optimisation problem, Thompson sampling is used as the acquisition function. We draw a sample from the joint posterior over a discrete set of $1000d$ points sampled from a scrambled Sobol sequence, suggested by (Daulton et al., 2020).

JES (Tu et al., 2022). We implement JES using the open-source implementation provided by the authors[3].

EHVI (Daulton et al., 2020) and NEHVI (Daulton et al., 2021). We implement EHVI and NEHVI based on the implementation provided by BoTorch.[4]. For both methods, we use their log versions, following the recommendation of (Ament et al., 2023).

MORBO (Daulton et al., 2022b). We implement MORBO using the open-source implementation provided by the authors[5] and set the batch size to one.

PSL (Lin et al., 2022). We implement PSL using the open-source implementation provided by the authors[6] and set the batch size to one.

MOBO-OSD (Ngo et al., 2025). We implement MOBO-OSD using the open-source implementation provided by the authors[7] and set the batch size to one.

qPOTS (Renganathan & Carlson, 2025). We implement qPOTS using the open-source implementation provided by the authors[8] and set the batch size to one.

For all the methods, we sample $2(d+1)$ initial points from a scrambled Sobol sequence, following the practice in (Daulton et al., 2020; 2021). For methods using HV, i.e., EHVI, NEHVI, MORBO, PSL, MOBO-OSD, and CTD, we consider a setting frequently considered in the literature (Daulton et al., 2020; Tu et al., 2022). That is, the reference point is $\boldsymbol{r} = \boldsymbol{y}_{\text{nadir}} + 0.1\,|\boldsymbol{y}_{\text{nadir}} - \boldsymbol{y}_{\text{ideal}}|$, where $\boldsymbol{y}_{\text{nadir}} = (\max_{y_1 \in \mathcal{D}_1^{(t)}} y_1, \ldots, \max_{y_m \in \mathcal{D}_m^{(t)}} y_m)$, $\boldsymbol{y}_{\text{ideal}} = (\min_{y_1 \in \mathcal{D}_1^{(t)}} y_1, \ldots, \min_{y_m \in \mathcal{D}_m^{(t)}} y_m)$, and $\mathcal{D}_i^{(t)} = \{y_i \mid \boldsymbol{y} \in \mathcal{D}^{(t)}\}$ denotes the set of observed values for the ith objective in the current observed dataset $\mathcal{D}^{(t)}$.

For all the problems, we normalise the input variables and standardise the objective values before training Gaussian processes. We assume an independent surrogate model for each objective, using a

[1] EI, NEI, and the BoTorch multi-objective tutorial.
[2] TS and BoTorch multi-objective tutorial
[3] JES implementation
[4] EHVI, NEHVI, and BoTorch multi-objective tutorial
[5] MORBO implementation
[6] PSL implementation
[7] MOBO-OSD implementation
[8] qPOTS implementation

constant mean function and a Matérn 5/2 ARD kernel with a dimension-scaled prior (see https://github.com/meta-pytorch/botorch/releases/tag/v0.12.0).

To enable statistical comparisons, each optimisation was repeated 30 times. We use the Wilcoxon rank-sum test (Wilcoxon, 1992) at a significance level of 0.05 and Holm-Bonferroni correction (Holm, 1979) to see if our method differs significantly from each peer method.

HV (Zitzler & Thiele, 1999) is used to compare all methods. For the evaluation, we use the commonly used setting in quality evaluation of multi-objective optimisation (Li & Yao, 2019). That is, we collect all solutions obtained by all methods over 30 independent runs and set the reference point $\boldsymbol{r} = \boldsymbol{y}_{\text{nadir}} + 0.1\,|\boldsymbol{y}_{\text{nadir}} - \boldsymbol{y}_{\text{ideal}}|$.

B.2 Problem Details

The details of the benchmark problems and real-world problems are given in the following.

DTLZ1–DTLZ7. DTLZ1–DTLZ7 are scalable multi-objective benchmark problems. Their mathematical formulations are provided in Deb et al. (2005).

Four Bar Truss Design. The four-bar truss design problem aims to optimise a lightweight truss structure while controlling its structural displacement (Cheng & Li, 1999; Tanabe & Ishibuchi, 2020). It involves $m = 2$ objectives with $d = 4$ continuous design variables. The first objective is related to the structural volume, while the second objective measures the displacement of the truss under load. This problem reflects the trade-off between reducing material usage and maintaining structural stiffness. The problem is formulated as:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & f_1(\boldsymbol{x}) = L\left(2x_1 + \sqrt{2}x_2 + \sqrt{x_3} + x_4\right), \\ & f_2(\boldsymbol{x}) = \frac{FL}{E}\left(\frac{2}{x_1} + \frac{2\sqrt{2}}{x_2} - \frac{2\sqrt{2}}{x_3} + \frac{2}{x_4}\right), \end{aligned}$$

where $F = 10$, $E = 2 \times 10^5$, and $L = 200$. The decision variables are bounded by

$$x_1, x_4 \in [1, 3], \qquad x_2, x_3 \in [\sqrt{2}, 3].$$

Pressure Vessel Design. The pressure vessel design problem aims to optimise the design of a cylindrical pressure vessel subject to engineering requirements on thickness, radius, length, and stress-related constraints (Kannan & Kramer, 1994; Tanabe & Ishibuchi, 2020). It involves $m = 2$ objectives with $d = 4$ design variables. The objectives are associated with minimising the manufacturing cost and the overall constraint violation. This problem reflects the trade-off between economical design and feasibility under pressure-vessel safety requirements. The objective functions are:

$$\begin{aligned} f_1(\boldsymbol{x}) &= 0.6224\tilde{x}_1 x_3 x_4 + 1.7781\tilde{x}_2 x_3^2 + 3.1661\tilde{x}_1^2 x_4 + 19.84\tilde{x}_1^2 x_3, \\ f_2(\boldsymbol{x}) &= \max(0, -g_1(\boldsymbol{x})) + \max(0, -g_2(\boldsymbol{x})) + \max(0, -g_3(\boldsymbol{x})), \end{aligned}$$

where

$$\tilde{x}_1 = 0.0625\,\text{round}(x_1), \qquad \tilde{x}_2 = 0.0625\,\text{round}(x_2),$$

and

$$\begin{aligned} g_1(\boldsymbol{x}) &= \tilde{x}_1 - 0.0193x_3, \\ g_2(\boldsymbol{x}) &= \tilde{x}_2 - 0.00954x_3, \\ g_3(\boldsymbol{x}) &= \pi x_3^2 x_4 + \frac{4}{3}\pi x_3^3 - 1296000. \end{aligned}$$

The decision variables are bounded by

$$x_1, x_2 \in [1, 100], \qquad x_3 \in [10, 200], \qquad x_4 \in [10, 240].$$

Hatch Cover Design. The hatch cover design problem considers the structural design of a hatch cover under loading conditions (Amir & Hasegawa, 1989; Tanabe & Ishibuchi, 2020). It involves $m = 2$ objectives with $d = 2$ continuous design variables. The objectives are related to minimising the structural weight and controlling the deflection of the hatch cover. This problem captures the trade-off between lightweight structural design and mechanical reliability. The objective functions are:

$$f_1(\boldsymbol{x}) = x_1 + 120x_2,$$
$$f_2(\boldsymbol{x}) = \sum_{i=1}^{4} \max(0, -g_i(\boldsymbol{x})),$$

where

$$g_1(\boldsymbol{x}) = 1 - \frac{\sigma_b}{\sigma_{b,\max}},$$
$$g_2(\boldsymbol{x}) = 1 - \frac{\tau}{\tau_{\max}},$$
$$g_3(\boldsymbol{x}) = 1 - \frac{\delta}{\delta_{\max}},$$
$$g_4(\boldsymbol{x}) = 1 - \frac{\sigma_b}{\sigma_k},$$

with

$$\sigma_k = \frac{Ex_1^2}{100},$$
$$\sigma_b = \frac{4500}{x_1 x_2},$$
$$\tau = \frac{1800}{x_2},$$
$$\delta = \frac{56.2 \times 10000}{Ex_1 x_2^2}.$$

Here, $E = 700000$, $\sigma_{b,\max} = 700$, $\tau_{\max} = 450$, and $\delta_{\max} = 1.5$. The decision variables are bounded by

$$x_1 \in [0.5, 4], \qquad x_2 \in [0.5, 50].$$

Vehicle Safety Design. The vehicle safety design problem aims to improve crashworthiness performance while reducing the vehicle weight (Liao et al., 2008; Tanabe & Ishibuchi, 2020). It involves $m = 3$ objectives with $d = 5$ continuous design variables. The objectives include vehicle weight and safety-related crashworthiness responses, such as acceleration and structural intrusion measures. This problem reflects the trade-off between lightweight design and passenger safety in vehicle crashworthiness optimisation. The objective functions are:

$$\begin{aligned}
f_1(\boldsymbol{x}) &= 1640.2823 + 2.3573285x_1 + 2.3220035x_2 + 4.5688768x_3 + 7.7213633x_4 + 4.4559504x_5, \\
f_2(\boldsymbol{x}) &= 6.5856 + 1.15x_1 - 1.0427x_2 + 0.9738x_3 + 0.8364x_4 - 0.3695x_1x_4 \\
&\quad + 0.0861x_1x_5 + 0.3628x_2x_4 - 0.1106x_1^2 - 0.3437x_3^2 + 0.1764x_4^2, \\
f_3(\boldsymbol{x}) &= -0.0551 + 0.0181x_1 + 0.1024x_2 + 0.0421x_3 - 0.0073x_1x_2 + 0.024x_2x_3 \\
&\quad - 0.0118x_2x_4 - 0.0204x_3x_4 - 0.008x_3x_5 - 0.0241x_2^2 + 0.0109x_4^2.
\end{aligned}$$

The decision variables are bounded by

$$x_i \in [1, 3], \qquad i = 1, \ldots, 5.$$

Car Side Impact Design. The car side-impact problem aims to minimise vehicle weight while satisfying safety constraints related to occupant injury and structural response Jain & Deb (2013). It involves $m = 4$ objectives with $d = 7$ variables. The problem reflects the trade-off between safety performance and weight under uncertainty. The mathematical formulations are given as follow:

$$
\begin{aligned}
f_1(\boldsymbol{x}) &= 1.98 + 4.9x_1 + 6.67x_2 + 6.98x_3 + 4.01x_4 + 1.78x_5 + 10^{-5}x_6 + 2.73x_7 \\
f_2(\boldsymbol{x}) &= 4.72 - 0.5x_4 - 0.19x_2x_3 \\
f_3(\boldsymbol{x}) &= 0.5\left(V_{\text{MBP}}(\boldsymbol{x}) + V_{\text{FD}}(\boldsymbol{x})\right) \\
f_4(\boldsymbol{x}) &= -\sum_{i=1}^{10} \max\left(g_i(\boldsymbol{x}), 0\right)
\end{aligned}
$$

where the constraint functions $g_i(\boldsymbol{x})$ are defined as:

$$
\begin{aligned}
g_1(\boldsymbol{x}) &= 1 - 1.16 + 0.3717x_2x_4 + 0.0092928x_3 \\
g_2(\boldsymbol{x}) &= 0.32 - 0.261 + 0.0159x_1x_2 + 0.06486x_1 + 0.019x_2x_7 - 0.0144x_3x_5 - 0.0154464x_6 \\
g_3(\boldsymbol{x}) &= 0.32 - 0.214 - 0.00817x_5 + 0.045195x_1 + 0.0135168x_1 - 0.03099x_2x_6 \\
&\quad + 0.018x_2x_7 - 0.007176x_3 - 0.023223x_3 + 0.00364x_5x_6 + 0.018x_2^2 \\
g_4(\boldsymbol{x}) &= 0.32 - 0.74 + 0.61x_2 + 0.031296x_3 + 0.031872x_7 - 0.227x_2^2 \\
g_5(\boldsymbol{x}) &= 32 - 28.98 - 3.818x_3 + 4.2x_1x_2 - 1.27296x_6 + 2.68065x_7 \\
g_6(\boldsymbol{x}) &= 32 - 33.86 - 2.95x_3 + 5.057x_1x_2 + 3.795x_2 + 3.4431x_7 - 1.45728 \\
g_7(\boldsymbol{x}) &= 32 - 46.36 + 9.9x_2 + 4.4505x_1 \\
g_8(\boldsymbol{x}) &= 4 - f_2(\boldsymbol{x}) \\
g_9(\boldsymbol{x}) &= 9.9 - V_{\text{MBP}}(\boldsymbol{x}) \\
g_{10}(\boldsymbol{x}) &= 15.7 - V_{\text{FD}}(\boldsymbol{x})
\end{aligned}
$$

with volume terms defined as:

$$
\begin{aligned}
V_{\text{MBP}}(\boldsymbol{x}) &= 10.58 - 0.674x_1x_2 - 0.67275x_2 \\
V_{\text{FD}}(\boldsymbol{x}) &= 16.45 - 0.489x_3x_7 - 0.8435x_6x_7.
\end{aligned}
$$

The search space is defined as:

$$
\begin{aligned}
x_1 &\in [0.5, 1.5], \\
x_2 &\in [0.45, 1.35], \\
x_3, x_4 &\in [0.5, 1.5], \\
x_5 &\in [0.875, 2.625], \\
x_6, x_7 &\in [0.4, 1.2].
\end{aligned}
$$

Laser Plasma Acceleration. The laser plasma acceleration (LPA) problem considers the optimisation of beam quality in a laser-plasma accelerator (Irshad et al., 2021; 2023). It involves $m = 3$ objectives with $d = 4$ design variables. The objectives describe competing beam-quality metrics, such as improving the accelerated electron beam performance while controlling undesirable properties such as energy spread or beam instability. Following the practice in (Ament et al., 2023; Daulton et al., 2023), we use the public dataset from (Irshad et al., 2023) to construct a surrogate model in our experiments.

C Additional Experimental Results

C.1 How Well Does CTD Perform Compared with Well-Established Methods?

C.1.1 Noiseless Cases.

Table 2: The HV results (mean and standard deviation) of the ten methods on the 20 benchmark and real-world problems under 200 evaluations. The method with the best mean HV is highlighted in bold. The symbols "$-$", "$\sim$" and "$+$" indicate that our method CTD is statistically worse than, equivalent to, and better than the competitor, respectively.

Method	DTLZ1 ($m=2$)		DTLZ2 ($m=2$)		DTLZ3 ($m=2$)		DTLZ4 ($m=2$)		DTLZ5 ($m=2$)		DTLZ6 ($m=2$)		DTLZ7 ($m=2$)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Sobol	5.13e-02	2.81e-01 ($\sim$)	2.45e-02	2.33e-02 ($+$)	3.26e-01	1.79e+00 ($\sim$)	2.47e-02	2.69e-02 ($+$)	1.76e-02	1.61e-02 ($+$)	0.00e+00	0.00e+00 ($+$)	0.00e+00	0.00e+00 ($+$)
ParEGO	2.56e+00	8.00e+00 ($\sim$)	4.05e-01	2.23e-03 ($+$)	3.91e+02	6.99e+02 ($\sim$)	4.04e-01	2.04e-03 ($+$)	4.03e-01	2.73e-03 ($+$)	3.01e-01	1.13e-01 ($\sim$)	6.32e-01	2.12e-01 ($-$)
TS-TCH	0.00e+00	0.00e+00 ($\sim$)	1.49e-01	2.42e-02 ($+$)	0.00e+00	0.00e+00 ($\sim$)	1.51e-01	2.30e-02 ($+$)	1.51e-01	2.23e-02 ($+$)	0.00e+00	0.00e+00 ($+$)	0.00e+00	0.00e+00 ($+$)
PSL	**9.15e+01**	**1.44e+02** ($-$)	3.94e-01	1.10e-02 ($+$)	1.05e+03	1.62e+03 ($\sim$)	1.17e-01	2.07e-02 ($+$)	3.94e-01	1.10e-02 ($+$)	0.00e+00	0.00e+00 ($+$)	4.91e-01	1.63e-01 ($-$)
JES	0.00e+00	0.00e+00 ($\sim$)	3.99e-01	2.77e-03 ($+$)	9.96e+02	8.75e+02 ($\sim$)	3.99e-01	3.57e-03 ($+$)	4.00e-01	2.76e-03 ($+$)	2.51e-01	7.30e-02 ($+$)	6.45e-01	1.86e-01 ($-$)
EHVI	6.21e-01	3.40e+00 ($\sim$)	4.08e-01	1.42e-03 ($+$)	1.10e+03	1.35e+03 ($\sim$)	4.08e-01	1.26e-03 ($+$)	4.08e-01	1.75e-03 ($+$)	2.63e-01	4.59e-02 ($+$)	1.84e-01	3.21e-02 ($\sim$)
MORBO	0.00e+00	0.00e+00 ($\sim$)	1.50e-01	2.21e-02 ($+$)	0.00e+00	0.00e+00 ($\sim$)	1.48e-01	1.94e-02 ($+$)	1.66e-01	1.88e-02 ($+$)	0.00e+00	0.00e+00 ($+$)	0.00e+00	0.00e+00 ($+$)
MOBO-OSD	0.00e+00	0.00e+00 ($\sim$)	3.73e-01	1.13e-02 ($+$)	**1.47e+03**	**9.24e+02** ($-$)	2.51e-01	1.07e-01 ($+$)	3.76e-01	8.43e-03 ($+$)	2.03e-01	7.66e-02 ($+$)	**7.40e-01**	**1.35e-02** ($-$)
qPOTS	0.00e+00	0.00e+00 ($\sim$)	3.37e-02	3.01e-02 ($+$)	0.00e+00	0.00e+00 ($\sim$)	3.62e-02	3.49e-02 ($+$)	3.67e-02	3.93e-02 ($+$)	0.00e+00	0.00e+00 ($+$)	0.00e+00	0.00e+00 ($+$)
CTD (ours)	1.31e+01	3.48e+01	**4.09e-01**	**1.35e-03**	5.86e+02	9.79e+02	**4.10e-01**	**1.24e-03**	**4.10e-01**	**1.11e-03**	**3.29e-01**	**3.49e-02**	1.78e-01	3.00e-02

Method	DTLZ1 ($m=3$)		DTLZ2 ($m=3$)		DTLZ3 ($m=3$)		DTLZ4 ($m=3$)		DTLZ5 ($m=3$)		DTLZ6 ($m=3$)		DTLZ7 ($m=3$)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Sobol	1.86e+03	4.49e+03 ($\sim$)	5.43e-02	2.41e-02 ($+$)	0.00e+00	0.00e+00 ($+$)	5.71e-02	2.64e-02 ($+$)	1.67e-03	2.00e-03 ($+$)	0.00e+00	0.00e+00 ($+$)	0.00e+00	0.00e+00 ($+$)
ParEGO	6.76e+03	9.88e+03 ($\sim$)	5.86e-01	7.34e-02 ($+$)	1.41e+06	9.03e+05 ($\sim$)	5.75e-01	7.23e-02 ($+$)	1.20e-01	3.74e-03 ($+$)	9.69e-02	2.11e-02 ($+$)	8.61e-01	1.29e-01 ($-$)
TS-TCH	9.99e+02	4.40e+03 ($+$)	1.04e-01	4.29e-02 ($+$)	0.00e+00	0.00e+00 ($+$)	9.29e-02	4.47e-02 ($+$)	2.07e-02	1.11e-02 ($+$)	0.00e+00	0.00e+00 ($+$)	0.00e+00	0.00e+00 ($+$)
PSL	1.01e+04	9.05e+03 ($\sim$)	2.75e-01	1.23e-01 ($+$)	2.41e+06	5.14e+05 ($\sim$)	1.37e-01	5.12e-02 ($+$)	1.26e-01	5.36e-03 ($\sim$)	0.00e+00	0.00e+00 ($+$)	2.64e-01	2.07e-01 ($-$)
JES	1.47e+04	4.10e+03 ($-$)	5.46e-01	8.12e-02 ($+$)	2.39e+06	6.57e+05 ($\sim$)	5.62e-01	5.97e-02 ($+$)	1.15e-01	6.52e-03 ($+$)	7.35e-02	2.19e-02 ($+$)	7.71e-01	1.91e-01 ($-$)
EHVI	1.39e+04	6.10e+03 ($-$)	3.23e-01	1.20e-01 ($+$)	2.45e+06	6.89e+05 ($\sim$)	3.31e-01	1.12e-01 ($+$)	1.11e-01	1.79e-02 ($+$)	7.86e-02	1.49e-02 ($+$)	2.27e-01	7.35e-03 ($+$)
MORBO	2.35e+03	6.92e+03 ($\sim$)	1.14e-01	5.52e-02 ($+$)	0.00e+00	0.00e+00 ($+$)	1.02e-01	5.37e-02 ($+$)	1.65e-02	8.94e-03 ($+$)	0.00e+00	0.00e+00 ($+$)	0.00e+00	0.00e+00 ($+$)
MOBO-OSD	**1.66e+04**	**3.47e+03** ($-$)	2.51e-01	7.23e-02 ($+$)	**3.63e+06**	**6.96e+05** ($-$)	2.58e-02	8.77e-02 ($+$)	6.60e-02	2.11e-02 ($+$)	3.71e-02	2.67e-02 ($+$)	**9.31e-01**	**2.49e-02** ($-$)
qPOTS	0.00e+00	0.00e+00 ($+$)	0.00e+00	0.00e+00 ($+$)	0.00e+00	0.00e+00 ($+$)	0.00e+00	0.00e+00 ($+$)	0.00e+00	0.00e+00 ($+$)	0.00e+00	0.00e+00 ($+$)	0.00e+00	0.00e+00 ($+$)
CTD (ours)	6.15e+03	7.35e+03	**6.43e-01**	**3.14e-02**	1.91e+06	1.00e+06	**6.45e-01**	**2.70e-02**	**1.29e-01**	**2.18e-03**	**1.07e-01**	**1.57e-02**	2.28e-01	6.42e-04

Method	Four bar truss design		Pressure vessel design		Hatch cover design		Vehicle safety		Car side impact		LPA		Summary
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	$-$/$\sim$/$+$
Sobol	4.65e+01	5.45e-01 ($+$)	5.45e+09	1.70e+09 ($+$)	2.10e+04	2.38e+02 ($+$)	1.61e+01	8.15e-01 ($+$)	2.42e+02	3.96e+00 ($+$)	3.08e+07	2.43e+06 ($+$)	0/3/17
ParEGO	5.34e+01	9.61e-01 ($+$)	1.11e+10	2.86e+07 ($-$)	2.18e+04	1.09e+01 ($+$)	2.68e+01	5.05e-01 ($-$)	3.15e+02	1.07e+02 ($+$)	4.71e+07	4.94e+06 ($+$)	4/5/11
TS-TCH	5.20e+01	8.68e-01 ($+$)	1.05e+10	2.11e+08 ($+$)	2.17e+04	5.49e+01 ($+$)	2.35e+01	4.11e-01 ($+$)	2.71e+02	9.21e+01 ($+$)	4.38e+07	2.96e+06 ($+$)	0/2/18
PSL	5.40e+01	3.69e-02 ($+$)	1.09e+10	1.67e+08 ($\sim$)	2.18e+04	7.26e+00 ($+$)	2.38e+01	1.48e+00 ($+$)	3.31e+02	5.83e+00 ($\sim$)	5.00e+07	2.91e+06 ($+$)	3/6/11
JES	5.34e+01	1.01e+00 ($+$)	1.09e+10	8.06e+07 ($+$)	2.18e+04	2.92e+01 ($+$)	**2.71e+01**	**3.63e-01** ($-$)	**3.46e+02**	**8.98e+00** ($-$)	4.69e+07	4.69e+06 ($+$)	5/3/12
EHVI	5.43e+01	7.44e-02 ($+$)	1.10e+10	9.64e+07 ($\sim$)	**2.18e+04**	**7.04e+00** ($\sim$)	2.59e+01	7.30e-01 ($+$)	2.94e+02	1.00e+02 ($\sim$)	**5.86e+07**	**2.70e+06** ($\sim$)	1/8/11
MORBO	5.21e+01	7.71e-01 ($+$)	1.05e+10	1.99e+08 ($+$)	2.17e+04	6.28e+01 ($+$)	2.33e+01	4.64e-01 ($+$)	3.01e+02	3.37e+00 ($+$)	4.39e+07	2.63e+06 ($+$)	0/3/17
MOBO-OSD	5.21e+01	8.54e-01 ($+$)	**1.11e+10**	**1.40e+07** ($-$)	2.15e+04	2.76e+02 ($+$)	2.38e+01	1.10e+00 ($+$)	2.84e+02	1.11e+01 ($+$)	4.07e+07	3.76e+06 ($+$)	6/1/13
qPOTS	3.98e+01	1.31e+00 ($+$)	1.10e+10	4.88e+07 ($\sim$)	1.88e+04	4.14e+03 ($+$)	1.32e+01	2.38e+00 ($+$)	2.53e+02	1.01e+01 ($+$)	4.38e+07	2.55e+06 ($+$)	0/3/17
CTD (ours)	**5.43e+01**	**2.80e-02**	1.09e+10	1.91e+08	2.18e+04	1.52e+01	2.66e+01	6.20e-01	3.26e+02	1.07e+01	5.82e+07	3.97e+06	–/–/–

C.1.2 NOISY CASES.

Table 3: The HV results (mean and standard deviation) of the eight methods on the 20 benchmark and real-world problems under noisy settings and 100 evaluations. The method with the best mean HV is highlighted in bold. The symbols "−", "∼" and "+" indicate that the proposed CTD is statistically worse than, equivalent to, and better than the competitor, respectively.

Method	DTLZ1 ($m=2$)		DTLZ2 ($m=2$)		DTLZ3 ($m=2$)		DTLZ4 ($m=2$)		DTLZ5 ($m=2$)		DTLZ6 ($m=2$)		DTLZ7 ($m=2$)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Sobol	0.00e+00	0.00e+00 (∼)	1.54e-02	1.39e-02 (+)	0.00e+00	0.00e+00 (∼)	1.37e-02	1.66e-02 (+)	1.17e-02	1.61e-02 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (∼)
NParEGO	0.00e+00	0.00e+00 (∼)	6.94e-02	4.58e-02 (∼)	3.98e+02	4.96e+02 (∼)	5.68e-02	6.14e-02 (∼)	**7.16e-02**	**6.93e-02** (∼)	1.36e-01	9.29e-02 (∼)	1.39e-02	7.63e-02 (∼)
TS-TCH	0.00e+00	0.00e+00 (∼)	2.76e-03	4.60e-03 (+)	0.00e+00	0.00e+00 (∼)	2.92e-03	8.28e-03 (+)	5.82e-03	1.47e-02 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (∼)
JES	0.00e+00	0.00e+00 (∼)	5.95e-02	5.53e-02 (∼)	3.24e+02	7.44e+02 (∼)	3.93e-02	4.22e-02 (∼)	4.20e-02	4.51e-02 (∼)	1.12e-01	9.06e-02 (∼)	0.00e+00	0.00e+00 (∼)
NEHVI	0.00e+00	0.00e+00 (∼)	2.38e-02	4.35e-02 (+)	**4.40e+02**	**6.41e+02** (∼)	2.82e-02	4.75e-02 (+)	3.20e-02	5.52e-02 (+)	**1.71e-01**	**6.73e-02** (∼)	**1.50e-02**	**7.13e-02** (∼)
MORBO	0.00e+00	0.00e+00 (∼)	5.55e-04	1.62e-03 (+)	0.00e+00	0.00e+00 (∼)	2.07e-03	7.19e-03 (+)	3.47e-03	8.53e-03 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (∼)
qPOTS	0.00e+00	0.00e+00 (∼)	4.22e-03	8.03e-03 (+)	0.00e+00	0.00e+00 (∼)	5.04e-03	1.39e-02 (+)	5.77e-03	1.68e-02 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (∼)
CTD (ours)	**8.72e-03**	**4.78e-02**	**7.59e-02**	**5.95e-02**	3.69e+02	6.54e+02	**5.68e-02**	**4.74e-02**	6.70e-02	4.69e-02	1.33e-01	7.48e-02	1.48e-02	5.13e-02

Method	DTLZ1 ($m=3$)		DTLZ2 ($m=3$)		DTLZ3 ($m=3$)		DTLZ4 ($m=3$)		DTLZ5 ($m=3$)		DTLZ6 ($m=3$)		DTLZ7 ($m=3$)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Sobol	3.08e+02	1.35e+03 (+)	3.16e-02	1.96e-02 (+)	0.00e+00	0.00e+00 (+)	3.56e-02	2.71e-02 (+)	1.48e-03	3.22e-03 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (∼)
NParEGO	5.77e+03	5.09e+03 (∼)	6.26e-02	6.14e-02 (+)	1.05e+06	7.39e+05 (∼)	5.32e-02	4.45e-02 (+)	7.85e-03	1.08e-02 (+)	2.53e-02	3.25e-02 (+)	0.00e+00	0.00e+00 (∼)
TS-TCH	1.18e+03	2.40e+03 (+)	8.98e-03	1.37e-02 (+)	0.00e+00	0.00e+00 (+)	1.29e-02	1.40e-02 (+)	2.89e-04	1.04e-03 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (∼)
JES	6.80e+03	5.15e+03 (∼)	5.61e-02	4.62e-02 (+)	7.39e+05	6.05e+05 (+)	5.40e-02	4.67e-02 (+)	4.93e-03	9.71e-03 (+)	1.27e-02	2.50e-02 (+)	0.00e+00	0.00e+00 (∼)
NEHVI	**9.89e+03**	**5.04e+03** (∼)	1.61e-02	2.48e-02 (+)	**1.87e+06**	**6.86e+05** (−)	1.96e-02	3.03e-02 (+)	9.73e-04	2.28e-03 (+)	3.72e-02	3.40e-02 (∼)	0.00e+00	0.00e+00 (∼)
MORBO	7.29e+02	3.38e+03 (+)	7.14e-03	9.57e-03 (+)	0.00e+00	0.00e+00 (+)	1.67e-02	2.77e-02 (+)	2.68e-04	7.75e-04 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (∼)
qPOTS	4.85e+02	1.53e+03 (+)	1.26e-02	1.53e-02 (+)	0.00e+00	0.00e+00 (+)	8.72e-03	1.11e-02 (+)	2.56e-04	7.47e-04 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (∼)
CTD (ours)	8.33e+03	6.55e+03	**9.15e-02**	**4.19e-02**	1.43e+06	6.62e+05	**8.17e-02**	**4.46e-02**	**1.47e-02**	**1.20e-02**	**5.16e-02**	**3.09e-02**	**0.00e+00**	**0.00e+00**

Method	Four bar truss design		Pressure vessel design		Hatch cover design		Vehicle safety		Car side impact		LPA		Summary
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	−/∼/+
Sobol	4.48e+01	8.24e-01 (+)	3.93e+09	1.91e+09 (+)	1.98e+04	8.07e+02 (+)	1.49e+01	1.02e+00 (+)	2.25e+02	5.93e+00 (+)	2.51e+07	2.46e+06 (+)	0/4/16
NParEGO	5.06e+01	1.65e+00 (∼)	1.03e+10	4.58e+08 (+)	2.14e+04	2.55e+02 (∼)	2.58e+01	9.75e-01 (∼)	2.45e+02	1.73e+01 (+)	4.14e+07	4.25e+06 (+)	0/13/7
TS-TCH	4.87e+01	1.53e+00 (+)	7.83e+09	1.55e+09 (+)	2.13e+04	3.01e+02 (∼)	2.15e+01	7.63e-01 (+)	2.37e+02	9.85e+00 (+)	3.68e+07	2.93e+06 (+)	0/5/15
JES	4.57e+01	1.97e+00 (+)	1.03e+10	6.69e+08 (+)	2.11e+04	5.55e+02 (+)	2.36e+01	1.48e+00 (+)	2.32e+02	1.86e+01 (+)	4.05e+07	3.37e+06 (+)	0/9/11
NEHVI	5.09e+01	1.99e+00 (∼)	**1.07e+10**	**3.32e+08** (∼)	2.14e+04	4.33e+02 (∼)	**2.68e+01**	**3.74e-01** (−)	**2.89e+02**	**1.90e+01** (−)	4.52e+07	5.15e+06 (∼)	3/11/6
MORBO	4.81e+01	1.89e+00 (+)	8.23e+09	1.49e+09 (+)	2.13e+04	4.11e+02 (∼)	2.08e+01	9.21e-01 (+)	2.37e+02	1.21e+01 (+)	3.78e+07	2.20e+06 (+)	0/5/15
qPOTS	4.85e+01	1.57e+00 (+)	8.27e+09	1.19e+09 (+)	2.14e+04	2.67e+02 (∼)	2.09e+01	8.46e-01 (+)	2.39e+02	1.03e+01 (+)	3.87e+07	2.54e+06 (+)	0/5/15
CTD (ours)	**5.11e+01**	**1.08e+00**	1.06e+10	2.83e+08	**2.15e+04**	**1.70e+02**	2.57e+01	6.86e-01	2.57e+02	1.95e+01	**4.52e+07**	**5.62e+06**	–/–/–

Table 4: The HV results (mean and standard deviation) of the eight methods on the 20 benchmark and real-world problems under noisy settings and 200 evaluations. The method with the best mean HV is highlighted in bold. The symbols "−", "∼" and "+" indicate that the proposed CTD is statistically worse than, equivalent to, and better than the competitor, respectively.

Method	DTLZ1 ($m=2$)		DTLZ2 ($m=2$)		DTLZ3 ($m=2$)		DTLZ4 ($m=2$)		DTLZ5 ($m=2$)		DTLZ6 ($m=2$)		DTLZ7 ($m=2$)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Sobol	3.20e-01	1.75e+00 (∼)	2.08e-02	1.60e-02 (+)	1.10e+02	6.03e+02 (+)	2.34e-02	2.02e-02 (+)	2.18e-02	2.17e-02 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (∼)
ParEGO	0.00e+00	0.00e+00 (∼)	1.42e-01	4.65e-02 (∼)	6.00e+02	7.25e+02 (∼)	**1.39e-01**	**7.34e-02** (∼)	**1.40e-01**	**8.26e-02** (∼)	**2.27e-01**	**7.34e-02** (∼)	**1.04e-01**	**1.55e-01** (∼)
TS-TCH	0.00e+00	0.00e+00 (∼)	5.39e-03	1.06e-02 (+)	0.00e+00	0.00e+00 (+)	8.57e-03	1.40e-02 (+)	1.35e-02	2.17e-02 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (∼)
JES	0.00e+00	0.00e+00 (∼)	**1.43e-01**	**6.86e-02** (∼)	8.24e+02	9.69e+02 (∼)	1.06e-01	5.03e-02 (∼)	9.97e-02	6.62e-02 (∼)	2.17e-01	6.15e-02 (∼)	1.37e-02	3.91e-02 (∼)
EHVI	0.00e+00	0.00e+00 (∼)	6.86e-02	8.47e-02 (+)	**8.92e+02**	**8.95e+02** (∼)	5.31e-02	5.55e-02 (+)	8.97e-02	7.69e-02 (∼)	1.94e-01	5.54e-02 (∼)	8.76e-02	1.61e-01 (∼)
MORBO	2.88e-01	1.58e+00 (∼)	4.13e-03	8.07e-03 (+)	0.00e+00	0.00e+00 (+)	5.99e-03	1.05e-02 (+)	7.80e-03	1.23e-02 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (∼)
qPOTS	**1.88e+00**	**1.03e+01** (∼)	7.95e-03	1.17e-02 (+)	0.00e+00	0.00e+00 (+)	7.90e-03	1.63e-02 (+)	8.16e-03	1.72e-02 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (∼)
CTD (ours)	8.72e-03	4.78e-02	1.37e-01	6.87e-02	6.03e+02	8.58e+02	1.33e-01	4.83e-02	1.33e-01	4.38e-02	2.07e-01	8.12e-02	3.02e-02	1.06e-01

Method	DTLZ1 ($m=3$)		DTLZ2 ($m=3$)		DTLZ3 ($m=3$)		DTLZ4 ($m=3$)		DTLZ5 ($m=3$)		DTLZ6 ($m=3$)		DTLZ7 ($m=3$)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Sobol	2.20e+03	4.59e+03 (+)	5.55e-02	2.42e-02 (+)	0.00e+00	0.00e+00 (+)	5.90e-02	2.20e-02 (+)	2.07e-03	3.37e-03 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (∼)
ParEGO	1.04e+04	5.79e+03 (∼)	1.49e-01	8.05e-02 (+)	1.86e+06	6.94e+05 (∼)	1.39e-01	9.21e-02 (+)	2.20e-02	1.66e-02 (+)	5.81e-02	2.93e-02 (∼)	**4.90e-02**	**1.48e-01** (∼)
TS-TCH	1.59e+03	2.67e+03 (+)	1.43e-02	1.59e-02 (+)	0.00e+00	0.00e+00 (+)	1.78e-02	1.71e-02 (+)	6.97e-04	2.19e-03 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (∼)
JES	**1.38e+04**	**4.54e+03** (∼)	1.12e-01	8.10e-02 (+)	1.91e+06	8.83e+05 (∼)	1.25e-01	7.43e-02 (+)	1.54e-02	1.27e-02 (+)	5.22e-02	2.97e-02 (∼)	2.13e-02	7.27e-02 (∼)
EHVI	1.37e+04	5.96e+03 (∼)	2.18e-02	4.04e-02 (+)	**2.44e+06**	**5.12e+05** (∼)	2.52e-02	3.46e-02 (+)	3.45e-03	7.29e-03 (+)	6.52e-02	2.47e-02 (∼)	4.62e-02	1.07e-01 (∼)
MORBO	2.12e+03	5.59e+03 (+)	1.24e-02	1.13e-02 (+)	0.00e+00	0.00e+00 (+)	2.05e-02	2.79e-02 (+)	3.30e-04	7.80e-04 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (∼)
qPOTS	2.62e+03	6.00e+03 (+)	1.56e-02	1.81e-02 (+)	2.81e+04	1.54e+05 (+)	1.32e-02	1.62e-02 (+)	3.61e-04	9.12e-04 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (∼)
CTD (ours)	1.07e+04	7.98e+03	**1.98e-01**	**6.82e-02**	2.15e+06	4.99e+05	**1.98e-01**	**6.62e-02**	**3.28e-02**	**1.38e-02**	**6.82e-02**	**2.66e-02**	4.97e-03	2.72e-02

Method	Four bar truss design		Pressure vessel design		Hatch cover design		Vehicle safety		Car side impact		LPA		Summary
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	−/∼/+
Sobol	4.65e+01	6.07e-01 (+)	5.28e+09	1.70e+09 (+)	2.05e+04	4.43e+02 (+)	1.62e+01	6.87e-01 (+)	2.42e+02	5.97e+00 (+)	2.90e+07	2.82e+06 (+)	0/3/17
ParEGO	5.23e+01	1.13e+00 (∼)	1.04e+10	4.70e+08 (+)	2.16e+04	8.50e+01 (∼)	2.67e+01	3.60e-01 (∼)	2.63e+02	1.55e+01 (+)	4.77e+07	4.94e+06 (+)	0/14/6
TS-TCH	5.08e+01	8.96e-01 (+)	9.00e+09	1.01e+09 (+)	2.14e+04	2.45e+02 (+)	2.26e+01	5.55e-01 (+)	2.60e+02	1.13e+01 (+)	4.33e+07	2.52e+06 (+)	0/3/17
JES	4.61e+01	2.13e+00 (+)	1.06e+10	2.84e+08 (∼)	2.12e+04	5.83e+02 (+)	2.36e+01	1.49e+00 (+)	2.74e+02	1.84e+01 (+)	4.65e+07	4.95e+06 (+)	0/12/8
EHVI	5.18e+01	1.86e+00 (∼)	**1.07e+10**	**2.96e+08** (∼)	2.15e+04	3.72e+02 (∼)	**2.72e+01**	**3.44e-01** (−)	**3.28e+02**	**1.11e+01** (−)	5.78e+07	3.75e+06 (∼)	2/13/5
MORBO	5.04e+01	1.26e+00 (+)	8.99e+09	9.19e+08 (+)	2.15e+04	3.67e+02 (+)	2.24e+01	6.05e-01 (+)	2.59e+02	1.04e+01 (+)	4.37e+07	2.98e+06 (+)	0/3/17
qPOTS	5.05e+01	9.04e-01 (+)	9.28e+09	8.85e+08 (+)	2.15e+04	2.09e+02 (+)	2.25e+01	5.79e-01 (+)	2.63e+02	9.98e+00 (+)	4.38e+07	3.23e+06 (+)	0/3/17
CTD (ours)	**5.26e+01**	**7.99e-01**	1.07e+10	2.52e+08	**2.17e+04**	**7.08e+01**	2.66e+01	3.16e-01	2.93e+02	1.59e+01	**5.80e+07**	**4.10e+06**	–/–/–

C.1.3 HIGH-DIMENSIONAL PROBLEMS.

Table 5: The HV results (mean and standard deviation) of the ten methods on 14 problems with $d = 50$. The method with the best mean HV is highlighted in bold. The symbols "$-$", "$\sim$" and "$+$" indicate that the proposed CTD is statistically worse than, equivalent to, and better than the competitor, respectively.

Method	DTLZ1 ($m = 2$)		DTLZ2 ($m = 2$)		DTLZ3 ($m = 2$)		DTLZ4 ($m = 2$)		DTLZ5 ($m = 2$)		DTLZ6 ($m = 2$)		DTLZ7 ($m = 2$)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Sobol	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)
ParEGO	1.34e+05	2.41e+04 (+)	6.48e-02	9.40e-02 (+)	1.91e+05	5.93e+04 (+)	2.08e-02	4.07e-02 (+)	2.92e-01	3.36e-01 (+)	8.85e-02	5.30e-02 (+)	4.01e-01	2.11e-01 ($-$)
TS-TCH	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)
PSL	8.86e+04	2.98e+04 (+)	0.00e+00	0.00e+00 (+)	3.99e+04	4.16e+04 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	3.50e-01	2.02e-01 ($\sim$)
JES	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)
EHVI	**2.04e+05**	**1.48e+04** ($\sim$)	0.00e+00	0.00e+00 (+)	2.78e+05	7.70e+04 ($\sim$)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	1.05e-01	6.48e-02 (+)	2.10e-01	2.64e-02 ($\sim$)
MORBO	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)
MOBO-OSD	1.77e+05	2.29e+04 (+)	0.00e+00	0.00e+00 (+)	**3.19e+05**	**6.53e+04** ($\sim$)	3.80e-02	7.14e-02 (+)	9.29e-02	1.97e-01 (+)	2.00e-03	1.10e-02 (+)	**7.39e-01**	**1.06e-02** ($-$)
qPOTS	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)
CTD (ours)	1.94e+05	2.83e+04	**1.84e-01**	**1.59e-01**	2.97e+05	7.55e+04	**1.29e-01**	**1.17e-01**	**7.42e-01**	**4.11e-01**	**1.56e-01**	**9.46e-02**	2.11e-01	2.45e-02

Method	DTLZ1 ($m = 3$)		DTLZ2 ($m = 3$)		DTLZ3 ($m = 3$)		DTLZ4 ($m = 3$)		DTLZ5 ($m = 3$)		DTLZ6 ($m = 3$)		DTLZ7 ($m = 3$)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Sobol	0.00e+00	0.00e+00 (+)	2.27e+00	7.62e-01 (+)	0.00e+00	0.00e+00 (+)	1.67e-01	2.06e-01 (+)	9.90e-01	5.03e-01 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)
ParEGO	9.98e+07	1.71e+07 (+)	3.19e+00	2.38e+00 (+)	4.21e+08	1.17e+08 (+)	7.42e-01	1.28e+00 (+)	4.42e+00	4.48e+00 (+)	1.83e-02	2.85e-02 (+)	6.84e-01	2.18e-01 ($-$)
TS-TCH	0.00e+00	0.00e+00 (+)	1.66e+00	6.99e-01 (+)	0.00e+00	0.00e+00 (+)	1.64e-01	1.93e-01 (+)	6.29e-01	3.69e-01 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)
PSL	8.34e+07	2.16e+07 (+)	3.56e+00	1.83e-02 (+)	2.47e+08	1.36e+08 (+)	3.85e-01	8.41e-01 (+)	1.15e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	4.66e-01	1.78e-01 ($-$)
JES	0.00e+00	0.00e+00 (+)	1.61e+00	8.66e-01 (+)	0.00e+00	0.00e+00 (+)	2.26e-01	2.79e-01 (+)	6.22e-01	4.00e-01 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)
EHVI	1.81e+08	1.40e+07 ($\sim$)	1.98e+00	1.35e+00 (+)	5.33e+08	1.00e+08 ($\sim$)	2.06e-01	3.63e-01 (+)	8.15e-01	8.83e-01 (+)	5.04e-02	2.29e-02 (+)	2.64e-01	1.21e-01 ($\sim$)
MORBO	0.00e+00	0.00e+00 (+)	1.75e+00	9.17e-01 (+)	0.00e+00	0.00e+00 (+)	1.81e-01	4.48e-01 (+)	6.41e-01	4.20e-01 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)
MOBO-OSD	**1.88e+08**	**4.93e+06** ($-$)	1.83e+00	1.29e+00 (+)	**7.87e+08**	**4.76e+07** ($-$)	3.16e+00	3.35e+00 (+)	1.49e+00	2.71e+00 (+)	3.02e-03	1.06e-02 (+)	**8.72e-01**	**7.41e-02** ($-$)
qPOTS	0.00e+00	0.00e+00 (+)	2.20e+00	4.52e-16 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	1.04e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)	0.00e+00	0.00e+00 (+)
CTD (ours)	1.80e+08	1.11e+07	**1.44e+01**	**2.97e+00**	5.09e+08	9.43e+07	**6.47e+00**	**2.16e+00**	**1.49e+01**	**4.24e+00**	**6.41e-02**	**1.56e-02**	2.50e-01	6.98e-02

Table 6: The summary of statistical results for the noisy experiments in Table 5. Here, the left, median, and right numbers are the counts of test problems where the CTD was statistically worse, equivalent, or better to the peer method, respectively.

	Sobol	**ParEGO**	**TS-TCH**	**PSL**	**JES**	**EHVI**	**MORBO**	**MOBO-OSD**	**qPOTS**
CTD (ours)	0/0/14	2/0/12	0/0/14	1/1/12	0/0/14	0/6/8	0/0/14	4/1/9	0/0/14

C.2 SENSITIVITY ANALYSIS

Table 7: The HV results (mean and standard deviation) of our method with six different threshold on the 20 benchmark and real-world problems under 100 evaluations. The method with the best mean HV is highlighted in bold. The symbols "−", "∼" and "+" indicate that our method CTD is statistically worse than, equivalent to, and better than the competitor, respectively.

Method	DTLZ1 ($m=2$)		DTLZ2 ($m=2$)		DTLZ3 ($m=2$)		DTLZ4 ($m=2$)		DTLZ5 ($m=2$)		DTLZ6 ($m=2$)		DTLZ7 ($m=2$)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
CTD (10^{-1})	0.00e+00	0.00e+00 (∼)	1.93e-01	1.07e-01 (+)	5.42e+02	7.71e+02 (∼)	2.21e-01	8.52e-02 (+)	2.20e-01	9.21e-02 (+)	**3.26e-01**	**4.89e-02** (∼)	1.73e-01	2.61e-02 (∼)
CTD (10^{-2})	1.92e+00	1.05e+01 (∼)	3.17e-01	5.62e-02 (∼)	4.93e+02	1.06e+03 (∼)	3.16e-01	4.19e-02 (+)	3.17e-01	4.51e-02 (+)	3.07e-01	5.16e-02 (∼)	1.71e-01	2.43e-02 (∼)
CTD (10^{-3})	1.87e-01	8.49e-01 (∼)	3.44e-01	2.32e-02 (∼)	6.01e+02	1.38e+03 (∼)	**3.55e-01**	**1.46e-02** (∼)	3.40e-01	2.22e-02 (∼)	3.00e-01	5.33e-02 (∼)	1.70e-01	2.37e-02 (∼)
CTD (10^{-5})	5.18e-01	2.84e+00 (∼)	3.33e-01	3.64e-02 (∼)	**7.00e+02**	**9.85e+02** (∼)	3.46e-01	2.06e-02 (∼)	3.43e-01	2.33e-02 (∼)	3.17e-01	6.28e-02 (∼)	**1.75e-01**	**2.73e-02** (∼)
CTD (10^{-6})	0.00e+00	0.00e+00 (∼)	2.90e-01	4.78e-02 (+)	6.96e+02	8.26e+02 (∼)	3.02e-01	3.33e-02 (+)	2.90e-01	5.75e-02 (+)	3.15e-01	5.20e-02 (∼)	1.62e-01	1.20e-02 (∼)
CTD (10^{-4})	**6.68e+00**	**2.43e+01**	**3.46e-01**	**2.09e-02**	4.17e+02	8.64e+02	3.52e-01	1.46e-02	**3.49e-01**	**1.95e-02**	3.19e-01	4.51e-02	1.71e-01	2.43e-02

Method	DTLZ1 ($m=3$)		DTLZ2 ($m=3$)		DTLZ3 ($m=3$)		DTLZ4 ($m=3$)		DTLZ5 ($m=3$)		DTLZ6 ($m=3$)		DTLZ7 ($m=3$)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
CTD (10^{-1})	**9.86e+03**	**5.85e+03** (−)	5.05e-02	7.32e-02 (+)	**2.12e+06**	**7.00e+05** (−)	5.98e-02	5.00e-02 (+)	2.30e-02	2.47e-02 (+)	9.28e-02	1.86e-02 (∼)	**2.26e-01**	**1.14e-02** (∼)
CTD (10^{-2})	6.58e+03	6.06e+03 (∼)	2.21e-01	7.10e-02 (+)	1.63e+06	7.83e+05 (∼)	2.37e-01	9.99e-02 (+)	8.08e-02	1.95e-02 (+)	9.51e-02	1.69e-02 (∼)	2.25e-01	1.16e-02 (∼)
CTD (10^{-3})	6.24e+03	7.01e+03 (∼)	3.31e-01	9.33e-02 (∼)	1.59e+06	9.17e+05 (∼)	3.03e-01	7.64e-02 (+)	9.42e-02	1.64e-02 (∼)	9.94e-02	1.76e-02 (∼)	2.26e-01	1.14e-02 (∼)
CTD (10^{-5})	1.67e+03	2.81e+03 (∼)	3.73e-01	7.25e-02 (∼)	1.39e+06	8.29e+05 (∼)	3.50e-01	6.30e-02 (∼)	9.59e-02	2.03e-02 (∼)	8.90e-02	2.17e-02 (∼)	2.24e-01	1.58e-02 (∼)
CTD (10^{-6})	3.93e+03	6.34e+03 (∼)	3.67e-01	6.17e-02 (∼)	1.42e+06	7.98e+05 (∼)	**3.68e-01**	**7.43e-02** (∼)	8.71e-02	2.27e-02 (+)	8.71e-02	2.52e-02 (+)	1.97e-01	3.17e-02 (+)
CTD (10^{-4})	4.29e+03	5.30e+03	**3.73e-01**	**7.95e-02**	1.58e+06	8.91e+05	3.60e-01	9.01e-02	**1.02e-01**	**1.30e-02**	**1.01e-01**	**1.73e-02**	2.22e-01	1.90e-02

Method	Four bar truss design		Pressure vessel design		Hatch cover design		Vehicle safety		Car side impact		LPA		Summary
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	−/∼/+
CTD (10^{-1})	5.41e+01	3.30e-02 (∼)	1.08e+10	1.90e+08 (∼)	2.18e+04	2.55e+01 (∼)	2.59e+01	7.70e-01 (∼)	3.09e+02	1.16e+01 (∼)	4.54e+07	4.07e+06 (∼)	2/12/6
CTD (10^{-2})	5.41e+01	4.12e-02 (∼)	1.08e+10	1.86e+08 (∼)	**2.18e+04**	**2.05e+01** (∼)	2.58e+01	9.54e-01 (∼)	**3.12e+02**	**1.25e+01** (∼)	4.61e+07	3.80e+06 (∼)	0/15/5
CTD (10^{-3})	**5.41e+01**	**3.02e-02** (∼)	1.09e+10	1.24e+08 (∼)	2.18e+04	2.18e+01 (∼)	2.55e+01	1.64e+00 (∼)	2.80e+02	9.56e+01 (∼)	4.51e+07	4.29e+06 (∼)	0/19/1
CTD (10^{-5})	5.40e+01	6.27e-02 (+)	1.09e+10	1.77e+08 (∼)	2.18e+04	1.88e+01 (∼)	**2.60e+01**	**7.75e-01** (∼)	3.10e+02	1.49e+01 (∼)	**4.64e+07**	**3.99e+06** (∼)	0/19/1
CTD (10^{-6})	5.30e+01	2.77e+00 (+)	**1.09e+10**	**1.40e+08** (∼)	2.18e+04	4.72e+01 (∼)	2.46e+01	2.18e+00 (+)	2.98e+02	2.02e+01 (∼)	4.46e+07	4.44e+06 (∼)	0/12/8
CTD (10^{-4})	5.41e+01	2.82e-02	1.09e+10	1.79e+08	2.18e+04	2.05e+01	2.60e+01	6.38e-01	3.05e+02	1.55e+01	4.62e+07	3.71e+06	–/–/–

Table 8: The HV results (mean and standard deviation) of our method with six different threshold on the 20 benchmark and real-world problems under 200 evaluations. The method with the best mean HV is highlighted in bold. The symbols "−", "∼" and "+" indicate that our method CTD is statistically worse than, equivalent to, and better than the competitor, respectively.

Method	DTLZ1 ($m=2$)		DTLZ2 ($m=2$)		DTLZ3 ($m=2$)		DTLZ4 ($m=2$)		DTLZ5 ($m=2$)		DTLZ6 ($m=2$)		DTLZ7 ($m=2$)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
CTD (10^{-1})	0.00e+00	0.00e+00 (∼)	4.08e-01	1.49e-03 (∼)	9.70e+02	1.01e+03 (∼)	4.08e-01	1.19e-03 (+)	4.08e-01	1.36e-03 (+)	**3.32e-01**	**3.53e-02** (∼)	**1.86e-01**	**3.22e-02** (∼)
CTD (10^{-2})	1.92e+00	1.05e+01 (∼)	4.09e-01	1.45e-03 (∼)	9.23e+02	1.30e+03 (∼)	4.09e-01	1.08e-03 (∼)	4.09e-01	1.53e-03 (∼)	3.14e-01	4.42e-02 (∼)	1.79e-01	3.03e-02 (∼)
CTD (10^{-3})	1.84e+00	6.59e+00 (∼)	**4.09e-01**	**1.32e-03** (∼)	**1.17e+03**	**1.60e+03** (∼)	4.10e-01	1.49e-03 (∼)	4.09e-01	1.05e-03 (∼)	3.08e-01	5.12e-02 (∼)	1.81e-01	3.07e-02 (∼)
CTD (10^{-5})	5.13e+00	1.36e+01 (∼)	4.09e-01	1.04e-03 (∼)	9.84e+02	1.13e+03 (∼)	4.10e-01	1.43e-03 (∼)	4.10e-01	1.41e-03 (∼)	3.31e-01	3.64e-02 (∼)	1.83e-01	3.09e-02 (∼)
CTD (10^{-6})	6.05e+00	2.74e+01 (∼)	4.08e-01	1.10e-03 (∼)	1.13e+03	1.14e+03 (∼)	4.09e-01	1.46e-03 (+)	4.09e-01	1.61e-03 (+)	3.23e-01	4.11e-02 (∼)	1.64e-01	1.66e-02 (∼)
CTD (10^{-4})	**1.31e+01**	**3.48e+01**	4.09e-01	1.35e-03	5.86e+02	9.79e+02	**4.10e-01**	**1.24e-03**	**4.10e-01**	**1.11e-03**	3.29e-01	3.49e-02	1.78e-01	3.00e-02

Method	DTLZ1 ($m=3$)		DTLZ2 ($m=3$)		DTLZ3 ($m=3$)		DTLZ4 ($m=3$)		DTLZ5 ($m=3$)		DTLZ6 ($m=3$)		DTLZ7 ($m=3$)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
CTD (10^{-1})	**1.28e+04**	**5.36e+03** (−)	3.55e-01	1.37e-01 (+)	**2.71e+06**	**6.94e+05** (−)	3.64e-01	1.29e-01 (+)	1.19e-01	5.24e-03 (+)	9.94e-02	1.85e-02 (∼)	2.28e-01	8.53e-04 (∼)
CTD (10^{-2})	1.03e+04	6.17e+03 (−)	5.82e-01	6.48e-02 (+)	2.67e+06	6.73e+05 (−)	5.78e-01	7.14e-02 (+)	1.25e-01	3.61e-03 (+)	1.06e-01	1.38e-02 (∼)	**2.41e-01**	**7.72e-02** (∼)
CTD (10^{-3})	1.13e+04	8.90e+03 (−)	6.23e-01	6.28e-02 (∼)	2.20e+06	9.50e+05 (∼)	6.25e-01	3.72e-02 (∼)	1.28e-01	1.87e-03 (∼)	1.07e-01	1.45e-02 (∼)	2.28e-01	6.05e-04 (∼)
CTD (10^{-5})	3.40e+03	5.01e+03 (∼)	**6.56e-01**	**2.20e-02** (∼)	1.73e+06	8.41e+05 (∼)	6.37e-01	4.38e-02 (∼)	1.29e-01	2.87e-03 (∼)	1.02e-01	1.71e-02 (∼)	2.28e-01	7.83e-04 (∼)
CTD (10^{-6})	8.67e+03	9.52e+03 (∼)	6.52e-01	2.80e-02 (∼)	1.75e+06	8.87e+05 (∼)	**6.55e-01**	**2.41e-02** (∼)	**1.30e-01**	**2.37e-03** (∼)	1.02e-01	1.62e-02 (∼)	2.28e-01	1.50e-03 (∼)
CTD (10^{-4})	6.15e+03	7.35e+03	6.43e-01	3.14e-02	1.91e+06	1.00e+06	6.45e-01	2.70e-02	1.29e-01	2.18e-03	**1.07e-01**	**1.57e-02**	2.28e-01	6.42e-04

Method	Four bar truss design		Pressure vessel design		Hatch cover design		Vehicle safety		Car side impact		LPA		Summary
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	−/∼/+
CTD (10^{-1})	5.43e+01	2.76e-02 (∼)	1.09e+10	1.91e+08 (∼)	2.18e+04	1.31e+01 (∼)	2.65e+01	6.18e-01 (∼)	3.23e+02	1.17e+01 (∼)	5.77e+07	3.53e+06 (∼)	2/13/5
CTD (10^{-2})	5.43e+01	3.06e-02 (∼)	1.10e+10	8.35e+07 (∼)	2.18e+04	1.28e+01 (∼)	2.64e+01	5.84e-01 (∼)	**3.30e+02**	**8.27e+00** (∼)	5.81e+07	3.39e+06 (∼)	2/15/3
CTD (10^{-3})	**5.43e+01**	**2.45e-02** (∼)	1.09e+10	1.16e+08 (∼)	2.18e+04	1.19e+01 (∼)	2.62e+01	1.22e+00 (∼)	2.94e+02	1.00e+02 (∼)	5.79e+07	3.78e+06 (∼)	1/19/0
CTD (10^{-5})	5.43e+01	2.72e-02 (∼)	1.09e+10	1.95e+08 (∼)	**2.18e+04**	**8.30e+00** (∼)	2.66e+01	5.84e-01 (∼)	3.25e+02	1.40e+01 (∼)	5.77e+07	3.25e+06 (∼)	0/20/0
CTD (10^{-6})	5.43e+01	5.14e-02 (+)	**1.10e+10**	**1.40e+08** (−)	2.18e+04	1.16e+01 (∼)	2.58e+01	1.65e+00 (∼)	3.17e+02	1.68e+01 (∼)	5.69e+07	4.39e+06 (∼)	1/16/3
CTD (10^{-4})	5.43e+01	2.80e-02	1.09e+10	1.91e+08	2.18e+04	1.52e+01	**2.66e+01**	**6.20e-01**	3.26e+02	1.07e+01	**5.82e+07**	**3.97e+06**	–/–/–

C.3 USING AN ALTERNATIVE ACQUISITION FUNCTION

Table 9: The HV results (mean and standard deviation) of the two methods on the 20 benchmark and real-world problems on 30 independent runs. The method with the best mean HV is highlighted in bold. The symbols "$-$", "$\sim$" and "$+$" indicate that CTD$_{tch}$ is statistically better than, equivalent to, and worse than EHVI-based CTD, respectively.

Method	**DTLZ1** ($m=2$)		**DTLZ2** ($m=2$)		**DTLZ3** ($m=2$)		**DTLZ4** ($m=2$)		**DTLZ5** ($m=2$)		**DTLZ6** ($m=2$)		**DTLZ7** ($m=2$)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
CTD$_{tch}$	0.00e+00	0.00e+00 ($\sim$)	**3.55e-01**	**1.22e-02** ($\sim$)	**5.07e+02**	**6.87e+02** ($\sim$)	**3.56e-01**	**1.20e-02** ($\sim$)	**3.51e-01**	**1.52e-02** ($\sim$)	**3.53e-01**	**6.73e-02** ($-$)	**3.75e-01**	**2.16e-01** ($-$)
CTD	**6.68e+00**	**2.43e+01**	3.46e-01	2.09e-02	4.17e+02	8.64e+02	3.52e-01	1.46e-02	3.49e-01	1.95e-02	3.19e-01	4.51e-02	1.71e-01	2.43e-02

Method	**DTLZ1** ($m=3$)		**DTLZ2** ($m=3$)		**DTLZ3** ($m=3$)		**DTLZ4** ($m=3$)		**DTLZ5** ($m=3$)		**DTLZ6** ($m=3$)		**DTLZ7** ($m=3$)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
CTD$_{tch}$	**4.30e+03**	**7.83e+03** ($\sim$)	**4.68e-01**	**5.62e-02** ($-$)	**1.60e+06**	**8.57e+05** ($\sim$)	**4.78e-01**	**5.39e-02** ($-$)	**1.05e-01**	**1.43e-02** ($\sim$)	9.86e-02	2.38e-02 ($\sim$)	**4.58e-01**	**2.68e-01** ($\sim$)
CTD	4.29e+03	5.30e+03	3.73e-01	7.95e-02	1.58e+06	8.91e+05	3.60e-01	9.01e-02	1.02e-01	1.30e-02	**1.01e-01**	**1.73e-02**	2.22e-01	1.90e-02

Method	**Four bar truss design**		**Pressure vessel design**		**Hatch cover design**		**Vehicle safety**		**Car side impact**		**LPA**		**Summary**
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	$-$/$\sim$/$+$
CTD$_{tch}$	5.13e+01	1.51e+00 ($+$)	**1.10e+10**	**3.73e+07** ($-$)	2.18e+04	3.11e+01 ($+$)	2.59e+01	1.61e+00 ($\sim$)	**3.26e+02**	**1.09e+01** ($-$)	4.13e+07	3.45e+06 ($+$)	6/11/3
CTD	**5.41e+01**	**2.82e-02**	1.09e+10	1.79e+08	**2.18e+04**	**2.05e+01**	**2.60e+01**	**6.38e-01**	3.05e+02	1.55e+01	**4.62e+07**	**3.71e+06**	–/–/–

Table 10: The HV results (mean and standard deviation) of the eight methods on the 20 benchmark and real-world problems under noisy settings. The method with the best mean HV is highlighted in bold. The symbols "$-$", "$\sim$" and "$+$" indicate that CTD$_{tch}$ is statistically better than, equivalent to, and worse than EHVI-based CTD, respectively.

Method	**DTLZ1** ($m=2$)		**DTLZ2** ($m=2$)		**DTLZ3** ($m=2$)		**DTLZ4** ($m=2$)		**DTLZ5** ($m=2$)		**DTLZ6** ($m=2$)		**DTLZ7** ($m=2$)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
CTD$_{tch}$	0.00e+00	0.00e+00 ($\sim$)	6.38e-02	5.88e-02 ($\sim$)	**4.43e+02**	**8.41e+02** ($\sim$)	4.19e-02	4.82e-02 ($\sim$)	4.88e-02	5.12e-02 ($\sim$)	1.02e-01	6.21e-02 ($\sim$)	**5.90e-02**	**1.15e-01** ($-$)
CTD	**8.72e-03**	**4.78e-02**	**7.59e-02**	**5.95e-02**	3.69e+02	6.54e+02	**5.68e-02**	**4.74e-02**	**6.70e-02**	**4.69e-02**	**1.33e-01**	**7.48e-02**	1.48e-02	5.13e-02

Method	**DTLZ1** ($m=3$)		**DTLZ2** ($m=3$)		**DTLZ3** ($m=3$)		**DTLZ4** ($m=3$)		**DTLZ5** ($m=3$)		**DTLZ6** ($m=3$)		**DTLZ7** ($m=3$)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
CTD$_{tch}$	6.61e+03	6.16e+03 ($\sim$)	5.20e-02	5.54e-02 ($+$)	**1.68e+06**	**6.09e+05** ($\sim$)	5.40e-02	5.44e-02 ($+$)	5.21e-03	8.78e-03 ($+$)	2.18e-02	2.95e-02 ($+$)	0.00e+00	0.00e+00 ($\sim$)
CTD	**8.33e+03**	**6.55e+03**	**9.15e-02**	**4.19e-02**	1.43e+06	6.62e+05	**8.17e-02**	**4.46e-02**	**1.47e-02**	**1.20e-02**	**5.16e-02**	**3.09e-02**	**0.00e+00**	**0.00e+00**

Method	**Four bar truss design**		**Pressure vessel design**		**Hatch cover design**		**Vehicle safety**		**Car side impact**		**LPA**		**Summary**
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	$-$/$\sim$/$+$
CTD$_{tch}$	**5.19e+01**	**9.48e-01** ($-$)	1.05e+10	5.25e+08 ($\sim$)	**2.15e+04**	**2.54e+02** ($\sim$)	**2.64e+01**	**5.92e-01** ($-$)	**2.70e+02**	**1.67e+01** ($-$)	4.11e+07	4.67e+06 ($+$)	4/11/5
CTD	5.11e+01	1.08e+00	**1.06e+10**	**2.83e+08**	2.15e+04	1.70e+02	2.57e+01	6.86e-01	2.57e+02	1.95e+01	**4.52e+07**	**5.62e+06**	–/–/–

Table 11: The HV results (mean and standard deviation) of the two methods on 14 problems with $d = 50$. The method with the best mean HV is highlighted in bold. The symbols "$-$", "$\sim$" and "$+$" indicate that the method is statistically better than, equivalent to, and worse than CTD, respectively.

Method	**DTLZ1** ($m=2$)		**DTLZ2** ($m=2$)		**DTLZ3** ($m=2$)		**DTLZ4** ($m=2$)		**DTLZ5** ($m=2$)		**DTLZ6** ($m=2$)		**DTLZ7** ($m=2$)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
CTD$_{tch}$	1.53e+05	1.97e+04 ($+$)	**3.16e-01**	**1.68e-01** ($-$)	2.71e+05	7.69e+04 ($\sim$)	**1.69e-01**	**1.09e-01** ($\sim$)	**1.12e+00**	**2.39e-01** ($-$)	**1.64e-01**	**1.00e-01** ($\sim$)	**3.63e-01**	**1.83e-01** ($-$)
CTD	**1.94e+05**	**2.83e+04**	1.84e-01	1.59e-01	**2.97e+05**	**7.55e+04**	1.29e-01	1.17e-01	7.42e-01	4.11e-01	1.56e-01	9.46e-02	2.11e-01	2.45e-02

Method	**DTLZ1** ($m=3$)		**DTLZ2** ($m=3$)		**DTLZ3** ($m=3$)		**DTLZ4** ($m=3$)		**DTLZ5** ($m=3$)		**DTLZ6** ($m=3$)		**DTLZ7** ($m=3$)	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
CTD$_{tch}$	1.10e+08	1.95e+07 ($+$)	**1.72e+01**	**2.47e+00** ($-$)	4.61e+08	1.09e+08 ($\sim$)	**7.68e+00**	**1.61e+00** ($-$)	**1.72e+01**	**2.01e+00** ($-$)	1.64e-02	2.91e-02 ($+$)	**6.34e-01**	**2.80e-01** ($-$)
CTD	**1.80e+08**	**1.11e+07**	1.44e+01	2.97e+00	**5.09e+08**	**9.43e+07**	6.47e+00	2.16e+00	1.49e+01	4.24e+00	**6.41e-02**	**1.56e-02**	2.50e-01	6.98e-02

Table 12: The summary of statistical results for the high-dimensional settings in Table 11. Here, the left, median, and right numbers are the counts of test problems where the CTD was statistically worse, equivalent, or better to CTD$_{tch}$, respectively.

	CTD$_{tch}$
CTD	7/4/3

D EXTENSIONS

The preceding discussion has addressed multi-objective optimisation problems in which evaluations are performed under noiseless, noisy, and high-dimensional settings. However, not all multi-objective settings conform to these scenarios. To accommodate a broader class of problems, we propose several extensions that enable the methodology to handle more optimisation scenarios.

Constrained Bayesian Optimisation (CBO). Existing CBO methods can be divided into explicit and implicit methods (Amini et al., 2024). Explicit methods first estimate the feasible region from constraint surrogate models and then optimise a standard acquisition function within this region (Sasena et al., 2002; Xu et al., 2023). In contrast, implicit methods avoid explicitly solving a constrained problem by incorporating feasibility into the acquisition function. For example, they may multiply improvement by the probability of feasibility (Gelbart et al., 2014; Letham et al., 2019), or use penalty (Ariafar et al., 2019; Pourmohamad & Lee, 2022; Amini & Van Nieuwenhuyse, 2023; Picheny et al., 2016). Our CTD can be naturally extended to the constrained setting through the above techniques.

Multi-Fidelity Bayesian Optimisation (MFBO). In many real-world optimisation scenarios, the evaluation is often available at multiple fidelity levels, where increasing fidelity typically leads to improved accuracy at the expense of higher computational cost. Many MFBO methods have been proposed to tackle such optimisation problems (Belakaria et al., 2020a; Kandasamy et al., 2017; Li et al., 2020; Moss et al., 2021; Song et al., 2019; Takeno et al., 2020; Wu et al., 2020; Zhang et al., 2017). Our proposed CTD can be potentially extended to the multi-fidelity setting by integrating prior techniques, e.g., building multiple surrogate models of different levels of fidelity.