Supplementary File of "Bayesian Inverse Transfer in Evolutionary Multiobjective Optimization"

ACM Reference format:
. 2024. Supplementary File of "Bayesian Inverse Transfer in Evolutionary Multiobjective Optimization". 1, 1, Article 1 (June 2024), 11 pages.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

S-I BENCHMARK PROBLEMS

S-I.1 Fomulation of the Benchmark Problems

Based on the widely known DTLZ and DTLZ^{-1} benchmark suite, eight base problems, i.e., mDTLZ1-(δ_1, δ_2), mDTLZ2-(δ_1, δ_2), mDTLZ3-(δ_1, δ_2), mDTLZ4-(δ_1, δ_2), mDTLZ1^{-1}-(δ_1, δ_2), mDTLZ2^{-1}-(δ_1, δ_2), mDTLZ3^{-1}-(δ_1, δ_2), and mDTLZ4^{-1}-(δ_1, δ_2) are constructed. These eight base problems are shown as follows:

- *mDTLZ1-(δ_1, δ_2)*:

$$\begin{aligned}
\min : f_1(\mathbf{x}) &= \frac{1}{2} x_1^{\delta_1} x_2^{\delta_1} \cdots x_{m-1}^{\delta_1}(1 + g(\mathbf{x})) \\
f_2(\mathbf{x}) &= \frac{1}{2} x_1^{\delta_1} x_2^{\delta_1} \cdots (1 - x_{m-1}^{\delta_1})(1 + g(\mathbf{x})) \\
&\cdots \\
f_{m-1}(\mathbf{x}) &= \frac{1}{2} x_1^{\delta_1}(1 - x_2^{\delta_1})(1 + g(\mathbf{x})) \\
f_m(\mathbf{x}) &= \frac{1}{2}(1 - x_1^{\delta_1})(1 + g(\mathbf{x})) \\
\text{s.t. } & 0 \le x_j \le 1, j = 1, \ldots, d
\end{aligned} \tag{S-1}$$

  where

$$g(\mathbf{x}) = (d - m + 1) + \sum_{j=m}^{d}((x_j - 0.5 - \delta_2)^2 - \cos(2\pi(x_j - 0.5 - \delta_2))) \tag{S-2}$$

XXXX-XXXX/2024/6-ART1 $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

- *mDTLZ2-(δ_1, δ_2)*:

$$\begin{aligned}
\min: f_1(\mathbf{x}) &= (1+g(\mathbf{x}))\cos(\frac{\pi}{2}x_1^{\delta_1})\cos(\frac{\pi}{2}x_2^{\delta_1})\cdots\cos(\frac{\pi}{2}x_{m-1}^{\delta_1}) \\
f_2(\mathbf{x}) &= (1+g(\mathbf{x}))\cos(\frac{\pi}{2}x_1^{\delta_1})\cos(\frac{\pi}{2}x_2^{\delta_1})\cdots\sin(\frac{\pi}{2}x_{m-1}^{\delta_1}) \\
&\cdots \\
f_{m-1}(\mathbf{x}) &= (1+g(\mathbf{x}))\cos(\frac{\pi}{2}x_1^{\delta_1})\sin(\frac{\pi}{2}x_2^{\delta_1}) \\
f_m(\mathbf{x}) &= (1+g(\mathbf{x}))\sin(\frac{\pi}{2}x_1^{\delta_1}) \\
\text{s.t. } & 0 \le x_j \le 1, j = 1, \ldots, d
\end{aligned} \tag{S-3}$$

where

$$g(\mathbf{x}) = \sum_{j=m}^{d}(x_j - 0.5 - \delta_2)^2 \tag{S-4}$$

- *mDTLZ3-(δ_1, δ_2)*:

$$\begin{aligned}
\min: f_1(\mathbf{x}) &= (1+g(\mathbf{x}))\cos(\frac{\pi}{2}x_1^{\delta_1})\cos(\frac{\pi}{2}x_2^{\delta_1})\cdots\cos(\frac{\pi}{2}x_{m-1}^{\delta_1}) \\
f_2(\mathbf{x}) &= (1+g(\mathbf{x}))\cos(\frac{\pi}{2}x_1^{\delta_1})\cos(\frac{\pi}{2}x_2^{\delta_1})\cdots\sin(\frac{\pi}{2}x_{m-1}^{\delta_1}) \\
&\cdots \\
f_{m-1}(\mathbf{x}) &= (1+g(\mathbf{x}))\cos(\frac{\pi}{2}x_1^{\delta_1})\sin(\frac{\pi}{2}x_2^{\delta_1}) \\
f_m(\mathbf{x}) &= (1+g(\mathbf{x}))\sin(\frac{\pi}{2}x_1^{\delta_1}) \\
\text{s.t. } & 0 \le x_j \le 1, j = 1, \ldots, d
\end{aligned} \tag{S-5}$$

where

$$g(\mathbf{x}) = \frac{1}{10}(d-m+1) + \sum_{j=m}^{d}((x_j - 0.5 - \delta_2)^2 - \frac{1}{10}\cos(2\pi(x_j - 0.5 - \delta_2))) \tag{S-6}$$

- *mDTLZ4-(δ_1, δ_2)*:

$$\begin{aligned}
\min: f_1(\mathbf{x}) &= (1+g(\mathbf{x}))\cos(\frac{\pi}{2}x_1^{2\cdot\delta_1})\cos(\frac{\pi}{2}x_2^{2\cdot\delta_1})\cdots\cos(\frac{\pi}{2}x_{m-1}^{2\cdot\delta_1}) \\
f_2(\mathbf{x}) &= (1+g(\mathbf{x}))\cos(\frac{\pi}{2}x_1^{2\cdot\delta_1})\cos(\frac{\pi}{2}x_2^{2\cdot\delta_1})\cdots\sin(\frac{\pi}{2}x_{m-1}^{2\cdot\delta_1}) \\
&\cdots \\
f_{m-1}(\mathbf{x}) &= (1+g(\mathbf{x}))\cos(\frac{\pi}{2}x_1^{2\cdot\delta_1})\sin(\frac{\pi}{2}x_2^{2\cdot\delta_1}) \\
f_m(\mathbf{x}) &= (1+g(\mathbf{x}))\sin(\frac{\pi}{2}x_1^{2\cdot\delta_1}) \\
\text{s.t. } & 0 \le x_j \le 1, j = 1, \ldots, d
\end{aligned} \tag{S-7}$$

where

$$g(\mathbf{x}) = \sum_{j=m}^{d}(x_j - 0.5 - \delta_2)^2 \tag{S-8}$$

- *mDTLZ1*$^{-1}$-(δ_1, δ_2):

$$
\begin{aligned}
\min: f_1(\mathbf{x}) &= -\frac{1}{2}x_1^{\delta_1}x_2^{\delta_1}\cdots x_{m-1}^{\delta_1}(1-g(\mathbf{x})) \\
f_2(\mathbf{x}) &= -\frac{1}{2}x_1^{\delta_1}x_2^{\delta_1}\cdots(1-x_{m-1}^{\delta_1})(1-g(\mathbf{x})) \\
&\cdots \\
f_{m-1}(\mathbf{x}) &= -\frac{1}{2}x_1^{\delta_1}(1-x_2^{\delta_1})(1-g(\mathbf{x})) \\
f_m(\mathbf{x}) &= -\frac{1}{2}(1-x_1^{\delta_1})(1-g(\mathbf{x})) \\
\text{s.t. } 0 \leq x_j &\leq 1, j = 1, \ldots, d
\end{aligned} \tag{S-9}
$$

where

$$
g(\mathbf{x}) = (d-m+1) + \sum_{j=m}^{d}((x_j - 0.5 - \delta_2)^2 - \cos(2\pi(x_j - 0.5 - \delta_2))) \tag{S-10}
$$

- *mDTLZ2*$^{-1}$-(δ_1, δ_2):

$$
\begin{aligned}
\min: f_1(\mathbf{x}) &= -(1-g(\mathbf{x}))\cos(\frac{\pi}{2}x_1^{\delta_1})\cos(\frac{\pi}{2}x_2^{\delta_1})\cdots\cos(\frac{\pi}{2}x_{m-1}^{\delta_1}) \\
f_2(\mathbf{x}) &= -(1-g(\mathbf{x}))\cos(\frac{\pi}{2}x_1^{\delta_1})\cos(\frac{\pi}{2}x_2^{\delta_1})\cdots\sin(\frac{\pi}{2}x_{m-1}^{\delta_1}) \\
&\cdots \\
f_{m-1}(\mathbf{x}) &= -(1-g(\mathbf{x}))\cos(\frac{\pi}{2}x_1^{\delta_1})\sin(\frac{\pi}{2}x_2^{\delta_1}) \\
f_m(\mathbf{x}) &= -(1-g(\mathbf{x}))\sin(\frac{\pi}{2}x_1^{\delta_1}) \\
\text{s.t. } 0 \leq x_j &\leq 1, j = 1, \ldots, d
\end{aligned} \tag{S-11}
$$

where

$$
g(\mathbf{x}) = \sum_{j=m}^{d}(x_j - 0.5 - \delta_2)^2 \tag{S-12}
$$

- *mDTLZ3*$^{-1}$-(δ_1, δ_2):

$$
\begin{aligned}
\min: f_1(\mathbf{x}) &= -(1-g(\mathbf{x}))\cos(\frac{\pi}{2}x_1^{\delta_1})\cos(\frac{\pi}{2}x_2^{\delta_1})\cdots\cos(\frac{\pi}{2}x_{m-1}^{\delta_1}) \\
f_2(\mathbf{x}) &= -(1-g(\mathbf{x}))\cos(\frac{\pi}{2}x_1^{\delta_1})\cos(\frac{\pi}{2}x_2^{\delta_1})\cdots\sin(\frac{\pi}{2}x_{m-1}^{\delta_1}) \\
&\cdots \\
f_{m-1}(\mathbf{x}) &= -(1-g(\mathbf{x}))\cos(\frac{\pi}{2}x_1^{\delta_1})\sin(\frac{\pi}{2}x_2^{\delta_1}) \\
f_m(\mathbf{x}) &= -(1-g(\mathbf{x}))\sin(\frac{\pi}{2}x_1^{\delta_1}) \\
\text{s.t. } 0 \leq x_j &\leq 1, j = 1, \ldots, d
\end{aligned} \tag{S-13}
$$

where

$$
g(\mathbf{x}) = \frac{1}{10}(d-m+1) + \sum_{j=m}^{d}((x_j - 0.5 - \delta_2)^2 - \frac{1}{10}\cos(2\pi(x_j - 0.5 - \delta_2))) \tag{S-14}
$$

- *mDTLZ4^{-1}-(δ_1, δ_2)*:

$$
\begin{aligned}
\min: \ & f_1(\mathbf{x}) = -(1-g(\mathbf{x}))\cos(\frac{\pi}{2}x_1^{2\cdot\delta_1})\cos(\frac{\pi}{2}x_2^{2\cdot\delta_1})\cdots\cos(\frac{\pi}{2}x_{m-1}^{2\cdot\delta_1}) \\
& f_2(\mathbf{x}) = -(1-g(\mathbf{x}))\cos(\frac{\pi}{2}x_1^{2\cdot\delta_1})\cos(\frac{\pi}{2}x_2^{2\cdot\delta_1})\cdots\sin(\frac{\pi}{2}x_{m-1}^{2\cdot\delta_1}) \\
& \quad \cdots \\
& f_{m-1}(\mathbf{x}) = -(1-g(\mathbf{x}))\cos(\frac{\pi}{2}x_1^{2\cdot\delta_1})\sin(\frac{\pi}{2}x_2^{2\cdot\delta_1}) \\
& f_m(\mathbf{x}) = -(1-g(\mathbf{x}))\sin(\frac{\pi}{2}x_1^{2\cdot\delta_1}) \\
\text{s.t. } & 0 \le x_j \le 1, j = 1, \ldots, d
\end{aligned}
\tag{S-15}
$$

where

$$
g(\mathbf{x}) = \sum_{j=m}^{d}(x_j - 0.5 - \delta_2)^2 \tag{S-16}
$$

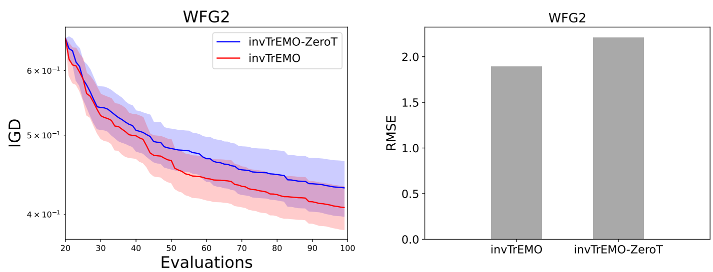
S-I.2 Further Discussions on the Benchmark Problems

S-I.2.1 ***Influence of*** δ_1 ***and*** δ_2. Let's begin by discussing the characteristics of the mDTLZ and mDTLZ^{-1} benchmark problems. Across all these problems, a notable pattern emerges: the first $(m-1)$ decision variables (referred to as $\mathbf{x}_I$) define the PF, while the remaining decision variables (denoted as $\mathbf{x}_{II}$) determine the position of the PS within the decision space. All of the nondominated solutions in the PS share the same $\mathbf{x}_{II}$ values. Then, according to the formulations, the first parameter, δ_1, primarily influences $\mathbf{x}_I$, thereby affecting the mapping relationship from a preference vector to a Pareto optimal solution. Specifically, for two mDTLZ problems with different δ_1 values and the same δ_2 values, the same preference vector $\mathbf{w}$ will correspond to different $\mathbf{x}_I$ values. Notably, however, this parameter does not alter the shape of the PF. Conversely, the second parameter, δ_2, predominantly influences $\mathbf{x}_{II}$. It distinguishes the PS corresponding to mDTLZ or mDTLZ^{-1} problems with varying δ_2 values by shifting the PS within the decision space.

An important aspect to consider is that, although δ_1 doesn't alter the shape of the PF, it remains pertinent to examine its influence. This is due to a distinctive requirement in our context compared to traditional multiobjective optimization endeavors. Here, our concern extends beyond merely obtaining the PS of the problem; *we also aim to ensure that, upon receiving a preference vector from the user, the inverse model accurately retrieves the corresponding nondominated solution.* Thus, in addition to the commonly discussed convergence and diversity metrics, it becomes imperative to delve deeper into the mapping relationship from preference vectors to nondominated solutions. Fortunately, parameter δ_1 fulfills this need.

S-I.2.2 ***Limitations of mDTLZ and mDTLZ^{-1} Benchmark Problems***. As previously mentioned, since no parameters are specifically designed to adjust the shape of the PF in the benchmark problems, it remains unexplored whether the inverse transfer technique can be effectively employed when the source and target exhibit different PF shapes. Furthermore, the benchmark problems, particularly mDTLZ benchmarks, often feature regular-shaped PFs (e.g., hyper-plane or hyper-sphere), which might not readily generalize to more realistic scenarios.

To further explore the effectiveness of invTrEMO when the source and target problems feature different PF shapes, we apply it to solve WFG2 within the WFG benchmark suite. We employ mDTLZ2-(1,0) as the source task to generate the source dataset. For comparison, we also utilize invTrEMO-ZeroT to solve WFG2, aiming to assess the efficacy of inverse transfer. The convergence trend in terms of IGD and RMSE results are depicted in Fig. S-1. The results indicate that invTrEMO, leveraging the source data, exhibits faster convergence and achieves superior RMSE performance compared to invTrEMO-ZeroT. This observation suggests that even in scenarios where the source and target problems possess different PF shapes, the inverse transfer technique can still bolster convergence rates and enhance the accuracy of inverse models.



(a) (b)

Fig. S-1. The IGD convergence trends and the RMSE results provided by invTrEMO and invTrEMO-ZeroT on WFG2. The shaded areas represent half a standard deviation in performance on either side of the mean. (a) IGD convergence trends. (b) RMSE results.

Indeed, from the perspective of the inverse modeling, there remains a necessity for further exploration of the characteristics and construction of benchmark problems. While current benchmarks predominantly emphasize the shapes of the PF, they often overlook the mapping from preference vectors to nondominated solutionsâĂŤan essential aspect for inverse modeling. This area offers ample research opportunities, but the focus of this study does not lie here, so we will not delve into it in depth.

S-II ADDITIONAL RESULTS

S-II.1 Further Investigations on the Transfer Mechanism

In Section 3.4, we outlined how invTGPs harness information from source data to enhance optimization. This section delves deeper into the experimental investigation of this transfer mechanism. Specifically, we compare the UCB value of the selected candidate solution $\mathbf{x}_T$ with the maximum UCB value in each iteration. Figure S-2 displays the corresponding data from a single run utilizing invTrEMO to tackle mDTLZ-$(1, 0)$. Notably, throughout each iteration, the UCB value of $\mathbf{x}_T$ consistently remains lower than the maximum UCB value. This observation suggests that invTGPs confine the search within the region covered by their generated offspring solutions rather than pursuing the solution with the highest UCB across the entire search space. Such an approach yields optimization benefits, as discussed in Section 3.4. By concentrating the search primarily on regions learned by the invTGPs, the algorithm effectively avoids over-exploration and redundant evaluations in unpromising regions, thereby reducing the algorithm's evaluation budget.

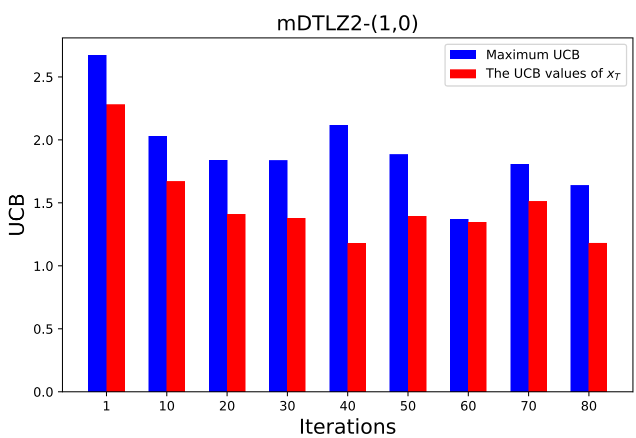
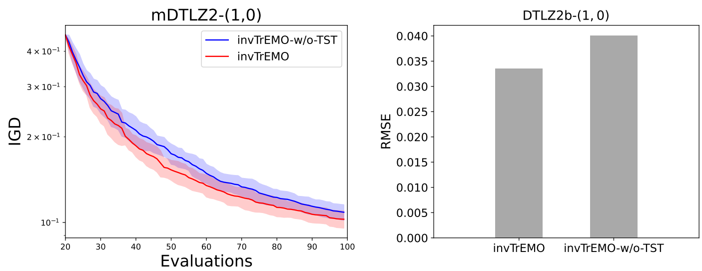


Fig. S-2. The UCB values of $\mathbf{x}_T$ and the maximum UCB values provided by the invTrEMO on solving mDTLZ-$(1, 0)$ in the $1st$-$80th$ iterations. The results are shown every 10 iterations.

S-II.2 Effectiveness of the Two Step Training Process

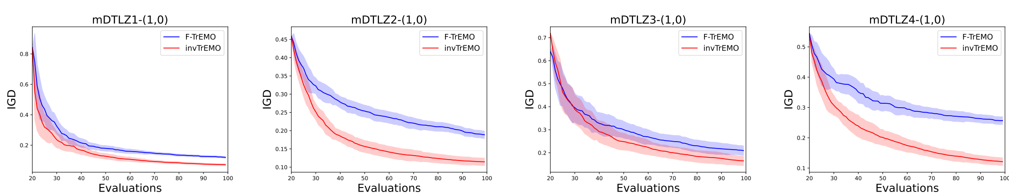
In Section 3.2, we employ a two step training process for optimizing the hyperparameters of the invTGPs. This strategy is designed to mitigate the risk of the inverse models exhibiting bias towards the source task rather than the target task, especially when the number of samples in the target dataset is significantly smaller than that in the source dataset. By doing so, the negative transfer can be alleviated to some extent. To assess the efficacy of this two step training process, we compare the performance of invTrEMO with a variant referred to as invTrEMO-w/o-TST on solving mDTLZ2-$(1, 0)$, leveraging a MS source dataset for assistance. In invTrEMO-w/o-TST, we train the hyperparameters of the invTGPs jointly in one pass. The convergence trend of IGD results and the RMSE results are illustrated in Fig. S-3. Notably, invTrEMO demonstrates superior IGD convergence trends and RMSE results than invTrEMO-w/o-TST. These outcomes can be attributed to the removal of the two step training process, which potentially leads to biased predictions of the invTGPs-w/o-TST towards source tasks. Consequently, negative transfer occurs, resulting in misguided optimization.



(a) (b)

Fig. S-3. The IGD convergence trends and the RMSE results provided by invTrEMO and invTrEMO-w/o-TST on mDTLZ-(1, 0). The shaded areas represent one standard deviation in performance on either side of the mean. (a) IGD convergence trends. (b) RMSE results.

S-II.3 Comparison with Forward Transfer



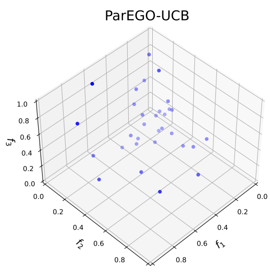
(a) (b) (c) (d)

Fig. S-4. Comparison of IGD convergence trends averaged over 20 independent runs of F-TrEMO and invTrEMO. Shaded areas represent one standard deviation on either side of the mean. (a) mDTLZ1-(1, 0). (b) mDTLZ2-(1, 0). (c) mDTLZ3-(1, 0). (d) mDTLZ4-(1, 0).

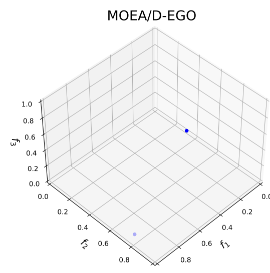
As discussed in Section 1, a prominent characteristic of invTrEMO is its ability to effectively transfer knowledge across heterogeneous source-target pairs. To further validate the effectiveness of our approach in handling such pairs, we conduct a comparative study between invTrEMO and a forward transfer-based algorithm, denoted as F-TrEMO, for solving mDTLZ problems. Unlike invTGPs used in invTrEMO, F-TrEMO utilizes a forward TGP for knowledge transfer. Specifically, the forward TGP approximates the scalarized function by utilizing both the source and target data, guiding optimization by identifying the solution with the maximum UCB value based on its predictions in each iteration. To adapt the forward TGP for heterogeneous decision spaces, we align the dimensions of the decision vectors of the source data with those of the target data by concatenating a set of random numbers within the range of [0,1]. The convergence trends of IGD for invTrEMO and F-TrEMO are illustrated in Fig. S-4. Notably, invTrEMO exhibits significantly faster convergence compared to F-TrEMO. These results underscore the effectiveness of invTrEMO in facilitating efficient knowledge transfer across heterogeneous scenarios.

S-II.4 Pareto Front Approximations

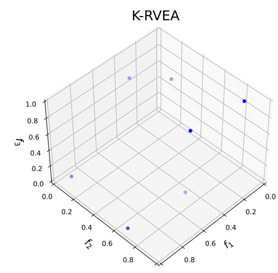
Figs. S-1 to S-4 shows the Pareto front approximation provided by ParEGO-UCB, MOEA/D-EGO, K-RVEA, CSEA, PSL-MOBO, qNEHVI, and invTrEMO on mDTLZ-(1, 0) benchmark problems after 100 evaluations. It is observed that compared with ParEGO-UCB, MOEA/D-EGO, K-RVEA, CSEA, PSL-MOBO, qNEHVI, invTrEMO obtains solutions that offer better convergence and spread.



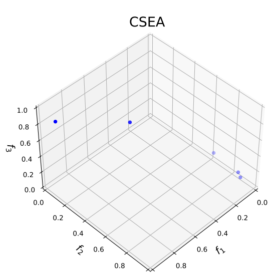
(a)



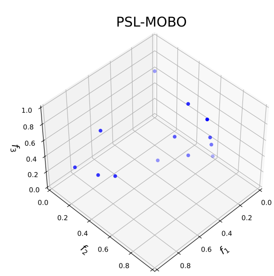
(b)



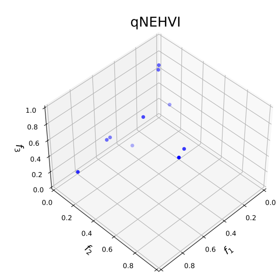
(c)



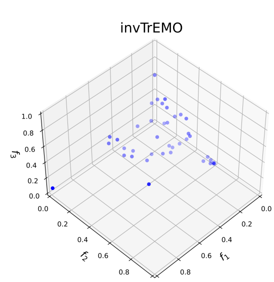
(d)



(e)



(f)



(g)

Fig. S-5. Pareto front approximation by ParEGO-UCB, MOEA/D-EGO, K-RVEA, CSEA, qNEHVI, PSL-MOBO, and invTrEMO on mDTLZ1-(1, 0) over 100 evaluations. Note that, only the solutions with the objective function values in the region $[0, 1] \times [0, 1] \times [0, 1]$ are shown in the figure. (a) ParEGO-UCB. (b) MOEA/D-EGO. (c) K-RVEA. (d) CSEA. (e) qNEHVI. (f) PSL-MOBO. (g) invTrEMO.

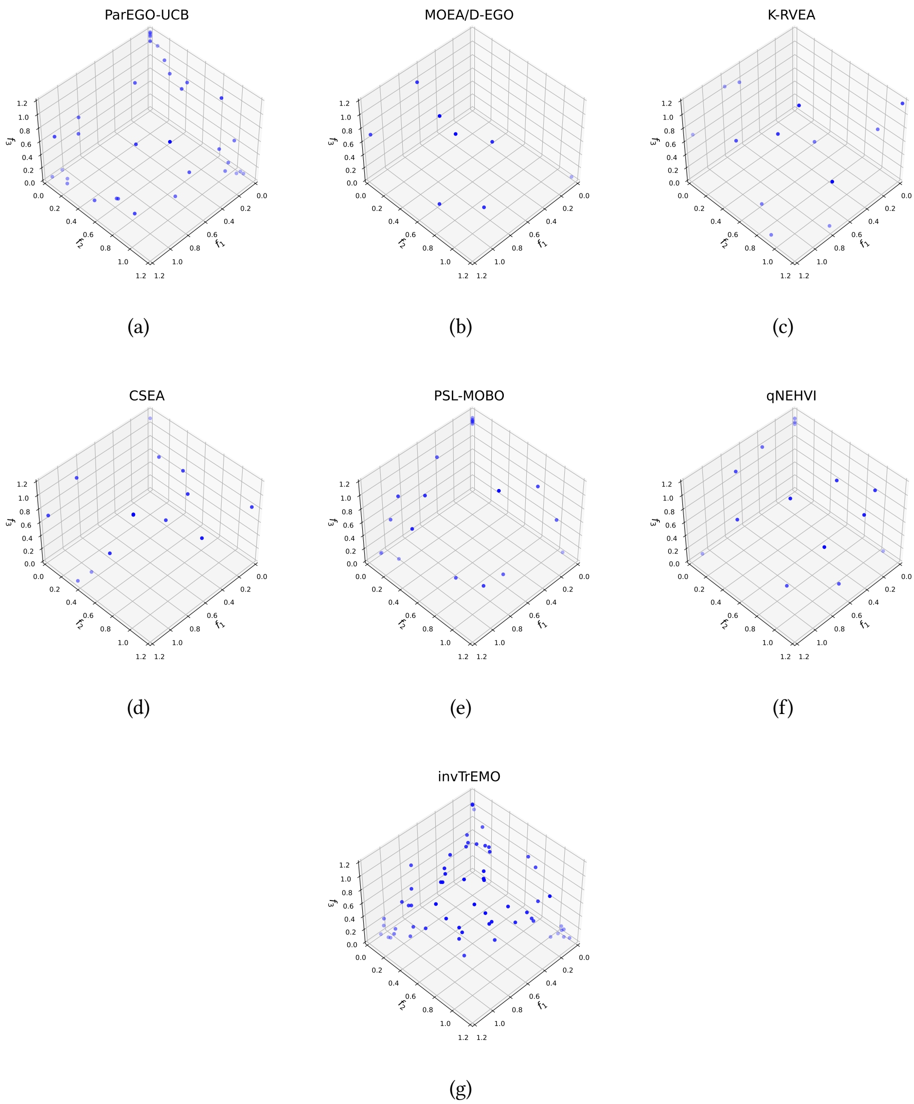


Fig. S-6. Pareto front approximation by ParEGO-UCB, MOEA/D-EGO, K-RVEA, CSEA, qNEHVI, PSL-MOBO, and invTrEMO on mDTLZ2-(1, 0) over 100 evaluations. Note that, only the solutions with the objective function values in the region $[0, 1.2] \times [0, 1.2] \times [0, 1.2]$ are shown in the figure. (a) ParEGO-UCB. (b) MOEA/D-EGO. (c) K-RVEA. (d) CSEA. (e) qNEHVI. (f) PSL-MOBO. (g) invTrEMO.

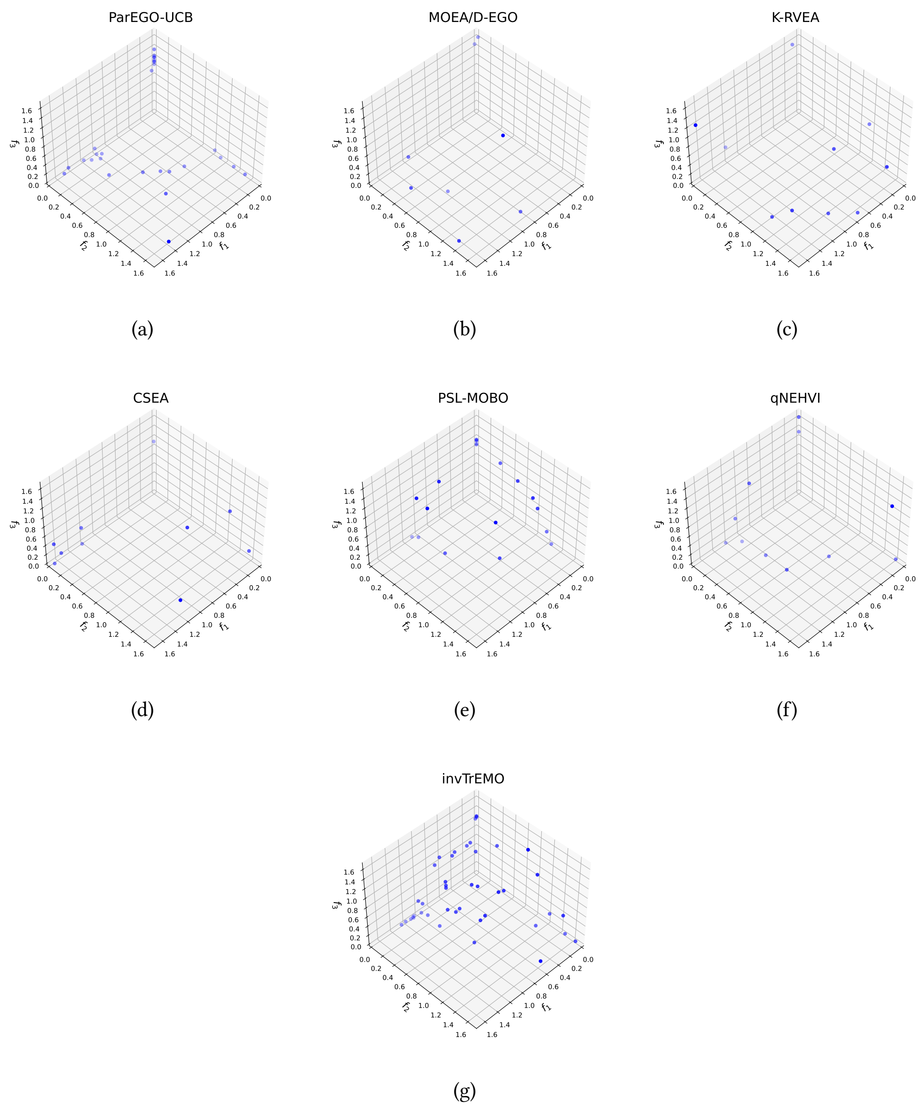
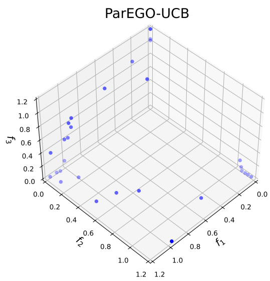
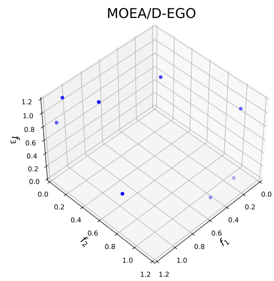


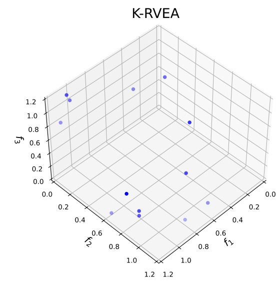
Fig. S-7. Pareto front approximation by ParEGO-UCB, MOEA/D-EGO, K-RVEA, CSEA, qNEHVI, PSL-MOBO, and invTrEMO on mDTLZ3-(1, 0) over 100 evaluations. Note that, only the solutions with the objective function values in the region $[0, 1.7] \times [0, 1.7] \times [0, 1.7]$ are shown in the figure. (a) ParEGO-UCB. (b) MOEA/D-EGO. (c) K-RVEA. (d) CSEA. (e) qNEHVI. (f) PSL-MOBO. (g) invTrEMO.



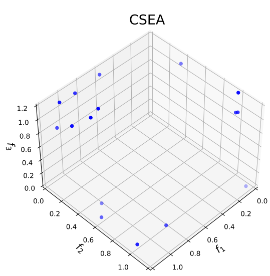
(a)



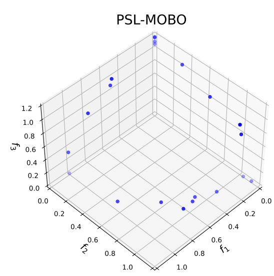
(b)



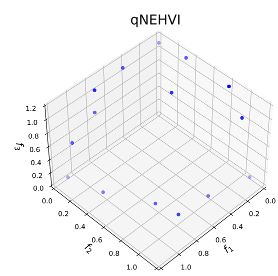
(c)



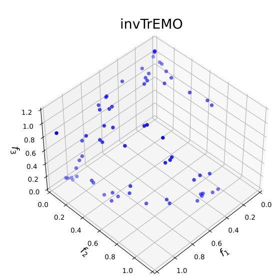
(d)



(e)



(f)



(g)

Fig. S-8. Pareto front approximation by ParEGO-UCB, MOEA/D-EGO, K-RVEA, CSEA, qNEHVI, PSL-MOBO, and invTrEMO on mDTLZ4-(1, 0) over 100 evaluations. Note that, only the solutions with the objective function values in the region $[0, 1.2] \times [0, 1.2] \times [0, 1.2]$ are shown in the figure. (a) ParEGO-UCB. (b) MOEA/D-EGO. (c) K-RVEA. (d) CSEA. (e) qNEHVI. (f) PSL-MOBO. (g) invTrEMO.