DGEMO: a Case of New Wine in Old Bottle

Yongqi FENG

目录 Contents

Introduction of MOP

$$\textbf{Minimize} \quad \boldsymbol{F}(x) = (f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \dots, f_K(\boldsymbol{x}))^T$$

$$\textbf{Subject to} \begin{cases} \boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_N) \in \boldsymbol{X} \\ \boldsymbol{G}(x) = (g_1(\boldsymbol{x}), g_2(\boldsymbol{x}), \dots, g_M(\boldsymbol{x}))^T \le 0 \\ \boldsymbol{H}(x) = (h_1(\boldsymbol{x}), h_2(\boldsymbol{x}), \dots, h_L(\boldsymbol{x}))^T = 0 \end{cases}$$

Introduction of MOP

- Expensive MOP: MOP whose objectives are costly and time-consuming
- Related work:
  - ParEGO: Surrogate-based method by Joshua Knowles in 2006
  - SMS-EGO: Model-assisted S-Metric Section method by Wolfgang Ponweiser in 2008
  - MOEA/D-EGO: MOEA/D with Gaussian Process by Qingfu Zhang in 2010
  - MOBO/D, SBMO…

Generalized Homotopy Approach to MOP

Generalized homotopy approach to multiobjective optimization

作者: Hillermeier, C (Hillermeier, C)

JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS
卷: 110 **期:** 3 **页:** 557-583
DOI: 10.1023/A:1017536311488
出版年: SEP 2001
文献类型: Article

摘要

This paper proposes a new generalized homotopy algorithm for the solution of multiobjective optimization problems with equality constraints. We consider the set of Pareto candidates as a differentiable manifold and construct a local chart which is fitted to the local geometry of this Pareto manifold. New Pareto candidates are generated by evaluating the local chart numerically. The method is capable of solving multiobjective optimization problems with an arbitrary number k of objectives, makes it possible to generate all types of Pareto optimal solutions, and is able to produce a homogeneous discretization of the Pareto set. The paper gives a necessary and sufficient condition for the set of Pareto candidates to form a (k-l)-dimensional differentiable manifold, provides the numerical details of the proposed algorithm, and applies the method to two multiobjective sample problems.

Hillermeier, C. (2001). Generalized homotopy approach to multiobjective optimization. Journal of Optimization Theory and Applications, 110(3), 557-583.

Generalized Homotopy Approach to MOP

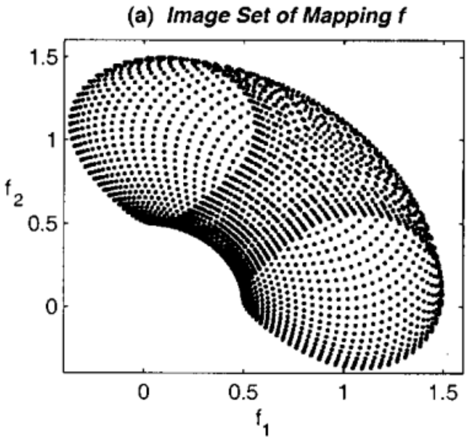
- a necessary and sufficient condition for Pareto candidates to form a $(K-1)$ -dimensional differentiable manifold.

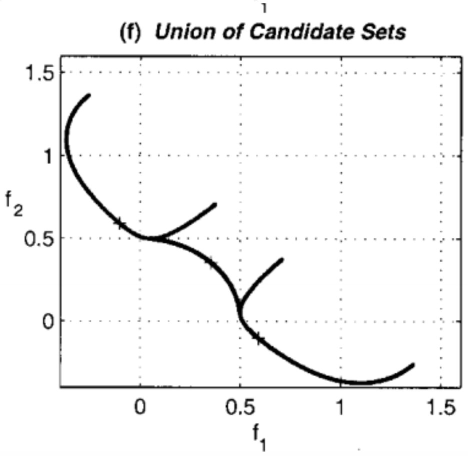
$$f:\begin{cases}\mathbb{R}^2 \to \mathbb{R}^2,\\ x \mapsto \begin{bmatrix}\cos(a(x))\cdot b(x)\\ \sin(a(x))\cdot b(x)\end{bmatrix},\end{cases}$$

with

$$a(x) := (2\pi/360)[a_c + a_1\sin(2\pi x_1) + a_2\sin(2\pi x_2)],$$

$$b(x) := 1 + d\cos(2\pi x_1).$$





Generalized Homotopy Approach to MOP

- Let x^* be Pareto optimal with respect to (MOP). Let $h(x)$ comply with the following constraint qualification:
- (CQ) the set of vectors $\{\nabla h_i(x^*) | i = 1, \ldots, L\}$ are linearly independent.
- Karush–Kuhn–Tucker condition(first-order necessary condition):
  - $\sum_{i=1}^{K} \alpha_i \nabla f_i(x^*) + \sum_{j=1}^{L} \lambda_j \nabla h_j(x^*) = 0$
  - $h_i(x^*) = 0, i = 1, \ldots, L$
  - $\sum_{i=1}^{K} \alpha_i = 1, \alpha_i \geq 0, i = 1, \ldots, K$

Hillermeier, C. (2001). Generalized homotopy approach to multiobjective optimization. Journal of Optimization Theory and Applications, 110(3), 557-583.

Generalized Homotopy Approach to MOP

- From KKT condition, let us figure out why the dimension of PF is $(K-1)$
  - $F(x,\lambda,\alpha) := \begin{bmatrix} \sum_{i=1}^{K} \alpha_i \nabla f_i(x^*) + \sum_{j=1}^{L} \lambda_j \nabla h_j(x^*) \\ h(x) \\ \sum_{i=1}^{K} \alpha_i - 1 \end{bmatrix} \in \mathbb{R}^{N+L+1}$
- And
  - $M \coloneqq \{(x,\lambda,\alpha) \in \mathbb{R}^{N+L+K} | F(x,\lambda,\alpha) = 0 and \alpha_i \geq 0, i = 1, \ldots, K\}$
- Theorem 1: let $(x^*, \lambda^*, \alpha^*) \in M$, if $rank\ F'(x^*, \lambda^*, \alpha^*) = N + L + 1$. Then , there exists an open neighborhood $U \in \mathbb{R}^{N+L+K}$, such that $M \cap U$ is a $(k-1)$ -dimensional C^1 -submanifold of $\mathbb{R}^{N+L+K}$, and therefore a $(k-1)$ -dimensional C^1 - manifold.

Hillermeier, C. (2001). Generalized homotopy approach to multiobjective optimization. Journal of Optimization Theory and Applications, 110(3), 557-583.

Generalized Homotopy Approach to MOP

- If we get a new possible point from the neighborhood U, whether is it the local optimal we need?
  - Check whether it is from M
  - Check whether it is local minimum
- We need second-order sufficient condition
- Let $L_{\alpha^*}(x^*) = \sum_{i=1}^{K} \alpha_i f_i(x^*) + \sum_{j=1}^{L} \lambda_j h_j(x^*)$
  - if $\nabla^2 L_{\alpha^*}(x^*)$ is positive definite, $x^*$is a isolated minimum point
  - if $\nabla^2 L_{\alpha^*}(x^*)$ is negative definite, $x^*$is a isolated maximum point
  - If if $\nabla^2 L_{\alpha^*}(x^*)$ has both positive and negative eigenvalues, $x^*$is a saddle point

0?→theorem 2.3

Hillermeier, C. (2001). Generalized homotopy approach to multiobjective optimization. Journal of Optimization Theory and Applications, 110(3), 557-583.

Generalized Homotopy Approach to MOP

- How to find new points?
  - Find the starting points x^*
  - Discover other points in U
  - Check new points

- Find the starting points x^*
  - Choose a weight vector α^* and transform it into a single objective problem
  - Optimize the problem with the state-of-the-art technique (L-BFGS)
  - Check the rank condition

$$rank\ F'(x^*, \lambda^*, \alpha^*) = N + L + 1$$

Hillermeier, C. (2001). Generalized homotopy approach to multiobjective optimization. Journal of Optimization Theory and Applications, 110(3), 557-583.

Generalized Homotopy Approach to MOP

- Discover other points in U

Calculate a QR-factorization of the matrix $(F'(x^*, \lambda^*, \alpha^*))^T$

$$(F'(x^*, \lambda^*, \alpha^*))^T = QR$$

Where $Q = (q_1, q_2, \ldots, q_{N+L+K}) \in \mathbb{R}^{(N+L+K)*(N+L+K)}$ is an orthonormal matrix and $R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \in \mathbb{R}^{(N+L+K)*(N+L+1)}$

R_1 is an upper triangular matrix

We use the tangent space Q to generate the local coordinate chart from $M \cap U$ such that

$$Span(q_1, q_2, q_3, \ldots, q_{K-1}) = T_{(x^*, \lambda^*, \alpha^*)}(M \cap U)$$

And

$$Span(q_K, q_{K+1}, \ldots, q_{N+L+K}) = T_{(x^*, \lambda^*, \alpha^*)}(M \cap U)^{\perp}$$

Hillermeier, C. (2001). Generalized homotopy approach to multiobjective optimization. Journal of Optimization Theory and Applications, 110(3), 557-583.

Generalized Homotopy Approach to MOP

- Discover other points in U

We take $\xi \in \mathbb{R}^{K-1}$ as the parameter to discover new points by the chart $(M \cap U, \varphi)$

Where

$$\varphi: \begin{cases} M \cap U \to U \\ (x^*, \lambda^*, \alpha^*) + Q \begin{bmatrix} \xi \\ \eta(\xi) \end{bmatrix} \end{cases}$$

And

$$\eta: \mathbb{R}^{K-1} \supseteq T \to \mathbb{R}^{N+L+1} \text{ with } \eta(0) = 0 \; and \; \frac{\delta\eta}{\delta\xi}(0) = 0$$

Now we need to find η, by solve the kkt condition

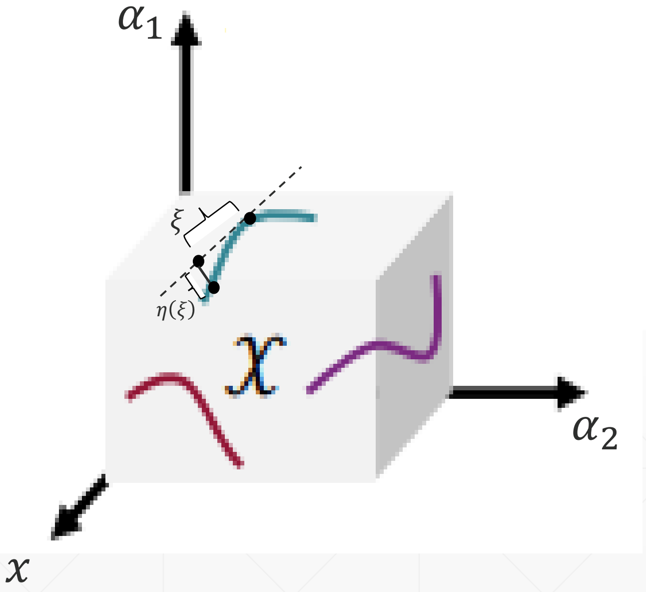
$$F(\varphi(\xi)) = F\left[(x^*, \lambda^*, \alpha^*) + Q \begin{bmatrix} \xi \\ \eta(\xi) \end{bmatrix}\right] = 0$$

With $N + L + 1$ equations and $N + L + 1$ unknowns

Generalized Homotopy Approach to MOP

- An example to illustrate the mechanism

  Suppose that $N = 1, K = 2, L = 0$



Hillermeier, C. (2001). Generalized homotopy approach to multiobjective optimization. Journal of Optimization Theory and Applications, 110(3), 557-583.

DGEMO

- NIPS2020：Diversity-Guided Multi-Objective Bayesian Optimization With Batch Evaluations

Diversity-Guided Multi-Objective Bayesian Optimization With Batch Evaluations

Part of Advances in Neural Information Processing Systems 33 (NeurIPS 2020)

AuthorFeedback » Bibtex » MetaReview » Paper » Review » Supplemental »

Authors

Mina Konakovic Lukovic, Yunsheng Tian, Wojciech Matusik

Abstract

<p>Many science, engineering, and design optimization problems require balancing the trade-offs between several conflicting objectives. The objectives are often black-box functions whose evaluations are time-consuming and costly. Multi-objective Bayesian optimization can be used to automate the process of discovering the set of optimal solutions, called Pareto-optimal, while minimizing the number of performed evaluations. To further reduce the evaluation time in the optimization process, testing of several samples in parallel can be deployed. We propose a novel multi-objective Bayesian optimization algorithm that iteratively selects the best batch of samples to be evaluated in parallel. Our algorithm approximates and analyzes a piecewise-continuous Pareto set representation. This representation allows us to introduce a batch selection strategy that optimizes for both hypervolume improvement and diversity of selected samples in order to efficiently advance promising regions of the Pareto front. Experiments on both synthetic test functions and real-world benchmark problems show that our algorithm predominantly outperforms relevant state-of-the-art methods. Code is available at https://github.com/yunshengtian/DGEMO.</p>

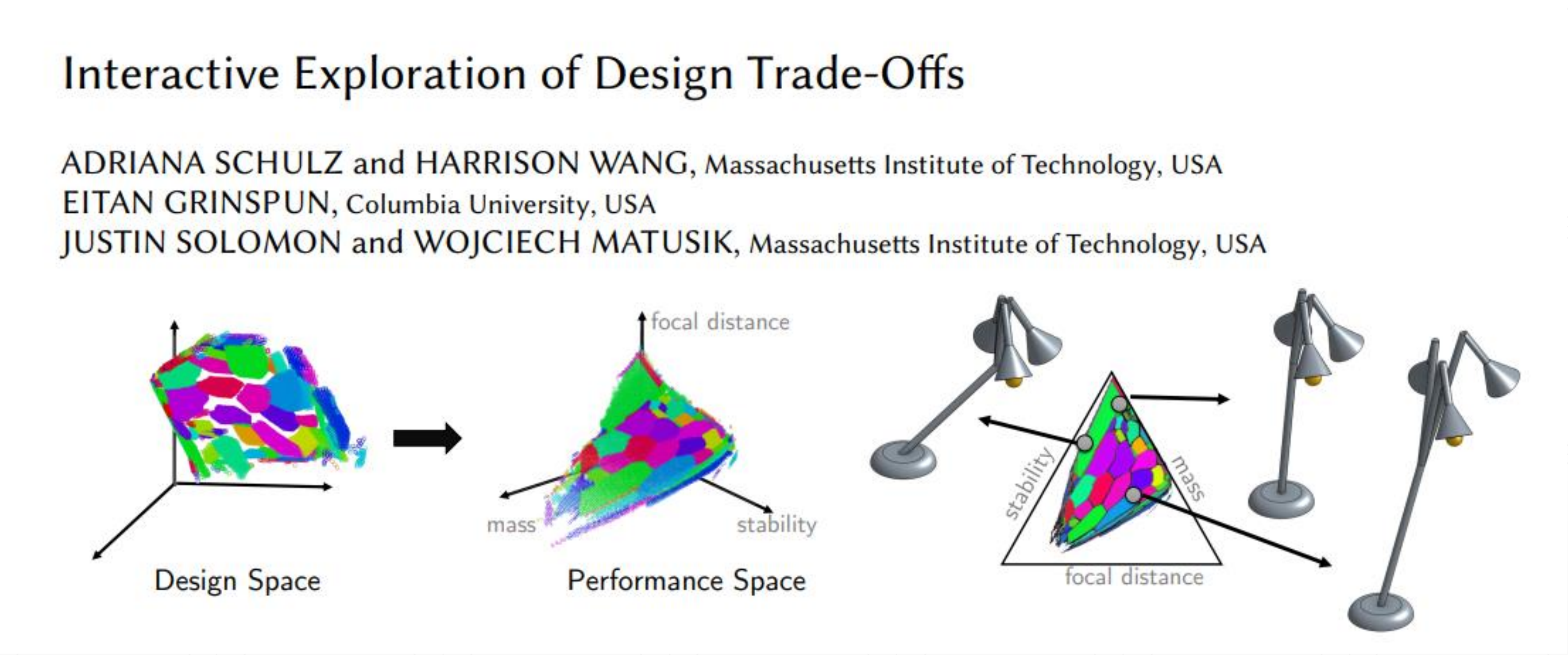
Lukovic, M. K., Tian, Y., & Matusik, W. Diversity-Guided Multi-Objective Bayesian Optimization With Batch Evaluations.

DGEMO

Algorithm 1 DGEMO

Inputs: Design space $\mathcal{X}$; black-box objectives $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), ..., f_m(\mathbf{x}))$; number of iterations n; number of initial samples k; batch size b.
Output: Pareto set $\mathcal{P}_s$ and Pareto front $\mathcal{P}_f$.
$X_0 \leftarrow \{\mathbf{x}_1, \ldots, \mathbf{x}_k\}, Y_0 \leftarrow \{\mathbf{f}(\mathbf{x}_1), \ldots, \mathbf{f}(\mathbf{x}_k)\}$ // initial samples drawn from LHS
for $i \leftarrow 0$ **to** n **do**
  Train surrogate models $G_j^{(i)}$ on X_i, Y_i for each objective $f_j, j \in \{1, \ldots, m\}$
  Define acquisition function $\tilde{f}_j^{(i)}$ from each $G_j^{(i)}$, $\tilde{\mathbf{f}}^{(i)}(\mathbf{x}) \leftarrow (\tilde{f}_1^{(i)}(\mathbf{x}), \ldots, \tilde{f}_m^{(i)}(\mathbf{x}))$
  Approximate Pareto set $\mathcal{P}_s^{(i)}$ and Pareto front $\mathcal{P}_f^{(i)}$ over $\tilde{\mathbf{f}}^{(i)}$
  Split points from $\mathcal{P}_s^{(i)}$ into diversity regions $\mathcal{D}_1^{(i)}, \ldots, \mathcal{D}_r^{(i)}$
  Select points $\mathbf{x}_1^{(i)}, \ldots, \mathbf{x}_b^{(i)}$ to evaluate from $\mathcal{D}_1^{(i)}, \ldots, \mathcal{D}_r^{(i)}$
  Evaluate and update $Y_{i+1} \leftarrow Y_i \cup \{\mathbf{f}^{(i)}(\mathbf{x}_1^{(i)}), \ldots, \mathbf{f}^{(i)}(\mathbf{x}_b^{(i)})\}$, $X_{i+1} \leftarrow X_i \cup \{\mathbf{x}_1^{(i)}, \ldots, \mathbf{x}_b^{(i)}\}$
Compute Pareto front $\mathcal{P}_f$ from points in Y_n and corresponding Pareto set $\mathcal{P}_s$

DGEMO


Interactive Exploration of Design Trade-Offs

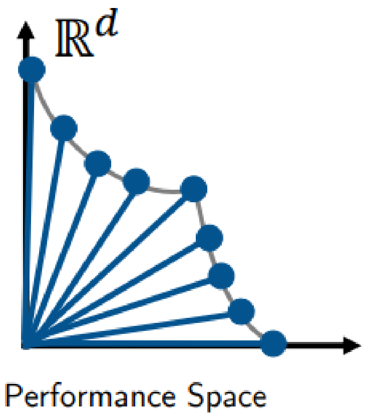
ADRIANA SCHULZ and HARRISON WANG, Massachusetts Institute of Technology, USA
EITAN GRINSPUN, Columbia University, USA
JUSTIN SOLOMON and WOJCIECH MATUSIK, Massachusetts Institute of Technology, USA

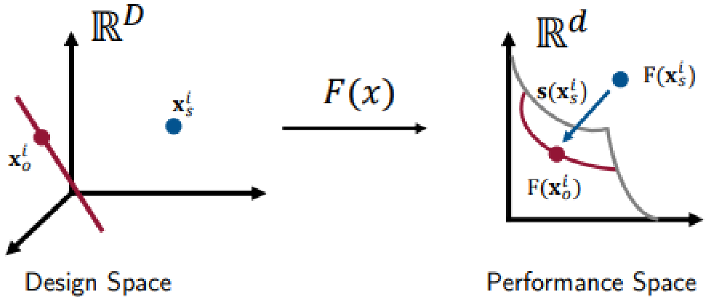
Lukovic, M. K., Tian, Y., & Matusik, W. Diversity-Guided Multi-Objective Bayesian Optimization With Batch Evaluations.
Schulz, A., Wang, H., Grinspun, E., Solomon, J., & Matusik, W. (2018). Interactive exploration of design trade-offs. ACM Transactions on Graphics (TOG), 37(4), 1-14.

DGEMO

Algorithm 1 Pareto set discovery given performance metrics F and design constraints that define $\mathcal{X}$.

procedure PARETOFRONTDISCOVERY($\mathcal{X}$, F)
  B: performance buffer array
  $B(i) \leftarrow \emptyset, \forall i$
  do
    $\mathbf{x}_s^0, \ldots, \mathbf{x}_s^{N_s} \leftarrow$ stochasticSampling($B, F, \mathcal{X}$)
    for each $\mathbf{x}_s^i$ **do**
      $D(\mathbf{x}_s^i) \leftarrow$ selectDirection($B, \mathbf{x}_s^i$)
      $\mathbf{x}_o^i \leftarrow$ localOptimization($D(\mathbf{x}_s^i), F, \mathcal{X}$)
      $M^i \leftarrow$ firstOrderApproximation($\mathbf{x}_o^i, F, \mathcal{X}$)
      updateBuffer($B, F(M^i)$)
    if buffer not updated on past N_i iterations **then**
      break
  while within computation budget
  return B

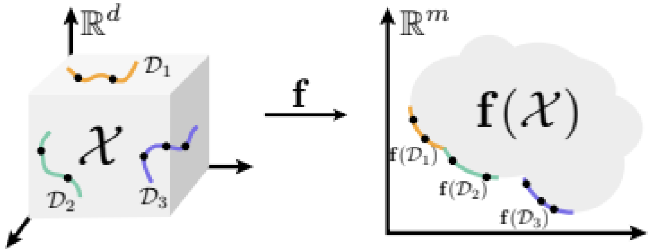




Lukovic, M. K., Tian, Y., & Matusik, W. Diversity-Guided Multi-Objective Bayesian Optimization With Batch Evaluations.
Schulz, A., Wang, H., Grinspun, E., Solomon, J., & Matusik, W. (2018). Interactive exploration of design trade-offs. ACM Transactions on Graphics (TOG), 37(4), 1-14.

DGEMO

- Diversity region



- Selection strategy

- $\arg\max_{X_B} \mathrm{HVI}(Y_B, \mathcal{P}_f) \quad \text{s.t.} \max_{1 \le i \le |\mathcal{D}|} \delta_i(X_B) - \min_{1 \le i \le |\mathcal{D}|} \delta_i(X_B) \le 1$

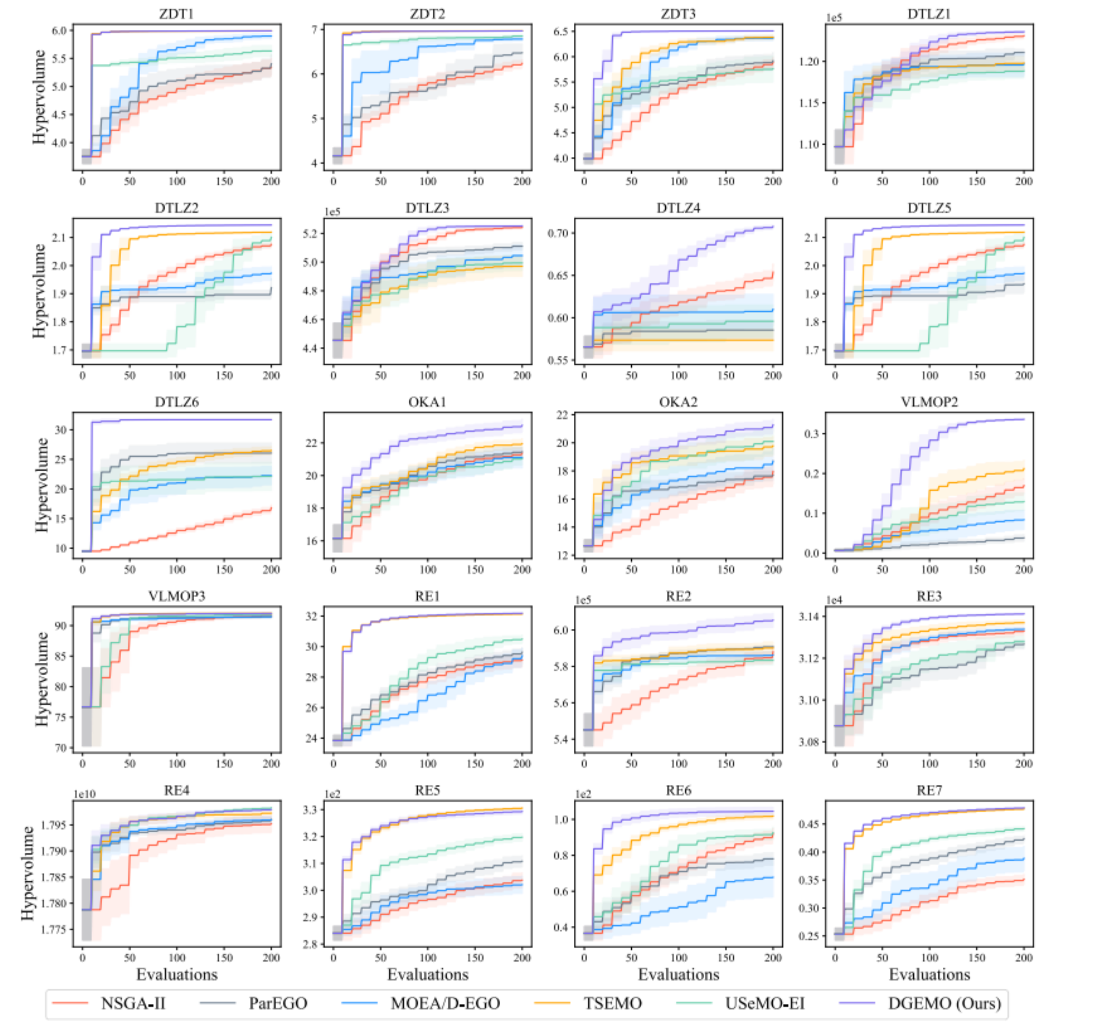
Algorithm 2 Batch Selection Algorithm

Inputs: Current Pareto front $\mathcal{P}_f$; batch size b; candidate points split into regions $\mathcal{D}_1, \ldots, \mathcal{D}_r$; surrogate objectives $\tilde{\mathbf{f}}(\mathbf{x}) = (\tilde{f}_1(\mathbf{x}), \ldots, \tilde{f}_m(\mathbf{x}))$.
Output: Batch of selected samples X_B.

$S \leftarrow \{1, 2, \ldots, r\}$ // set of available regions to choose
for $i \leftarrow 1$ **to** b **do**
  $h_{max} \leftarrow -\infty, s_{max} \leftarrow 0, \mathbf{x}_{max} \leftarrow null$ // record the sample with maximal HVI
  for each $s \in S$ **do**
    for each $\mathbf{x} \in \mathcal{D}_s$ **do**
      $h \leftarrow \mathrm{HVI}(\tilde{\mathbf{f}}(\mathbf{x}), \mathcal{P}_f)$
      if $h > h_{max}$ **then** $h_{max} \leftarrow h, s_{max} \leftarrow s, \mathbf{x}_{max} \leftarrow \mathbf{x}$
  $X_B \leftarrow X_B \cup \{\mathbf{x}_{max}\}, \mathcal{P}_f \leftarrow \mathcal{P}_f \cup \{\tilde{\mathbf{f}}(\mathbf{x}_{max})\}$ // aggregate solutions
  $\mathcal{D}_{s_{max}} \leftarrow \mathcal{D}_{s_{max}} - \{\mathbf{x}_{max}\}, S \leftarrow S - \{s_{max}\}$
  if $S = \emptyset$ **then** $S \leftarrow \{1, 2, \ldots, r\}$ // reset the counter of visited regions

Lukovic, M. K., Tian, Y., & Matusik, W. Diversity-Guided Multi-Objective Bayesian Optimization With Batch Evaluations.

Results


ZDT1
ZDT2
ZDT3
DTLZ1
DTLZ2
DTLZ3
DTLZ4
DTLZ5
DTLZ6
OKA1
OKA2
VLMOP2
VLMOP3
RE1
RE2
RE3
RE4
RE5
RE6
RE7
Hypervolume
Evaluations
NSGA-II
ParEGO
MOEA/D-EGO
TSEMO
USeMO-EI
DGEMO (Ours)