OPTIMIZATION ERUDITORUM
Volume 2 (2025), No. 3, 157–164
https://doi.org/10.69829/oper-025-0203-ta01
Tulipa Opera Scholarum

GRADIENT-FREE METHODS FOR NONSMOOTH NONCONVEX MULTIOBJECTIVE OPTIMIZATION

YONG ZHAO[1], CHUN-LIN YOU[1], AND WANG CHEN[2,*]

[1]*College of Mathematics and Statistics, Chongqing Jiaotong University, Chongqing, 400074, China*
[2]*National Center for Applied Mathematics in Chongqing, Chongqing Normal University, Chongqing, 401331, China*

Dedicated to Professor Hari Mohan Srivastava on the Occasion of His 85th Birthday

Abstract. This paper introduces a gradient-free method for solving a class of nonsmooth nonconvex multiobjective optimization problems. Under standard regularity conditions, we establish a non-asymptotic convergence rate of the proposed algorithm, where optimality is quantified in terms of (δ, ϵ)-Goldstein stationarity measure. Numerical experiments are presented to demonstrate the effectiveness of the proposed method.

Keywords. Nonsmooth nonconvex multiobjective optimization, Gradient-free method, Convergence rate.

1. Introduction

Multiobjective optimization is an optimization problem that involves the simultaneous optimization of multiple objective functions. It has wide applications in fields such as engineering, economic management, traffic planning, and machine learning [4, 9, 19]. These problems are called multiobjective optimization problems (MOPs) and the objective functions are conflicting, and since there is no single point that can optimize all objective functions at once, the concept of optimality is replaced by the concept of Pareto optimality. Due to this, the numerical computation of solutions to MOPs is more challenging. In addition, there are numerous applications involving nonsmooth objectives, such as financial risk control in asset portfolios or customer satisfaction optimization in service/supply chain systems, further complicating the problem.

Multiobjective optimization and nonsmooth optimization separately, there exist a large number of solution methods. Classical methods for solving MOPs include heuristic methods [7, 5] and scalarization methods [14]. The heuristic methods suffer from a lack of theoretical convergence guarantees and difficulties in scaling to large-scale optimization problems. The scalarization methods convert the MOPs into a parameterized scalar one, but it is difficult to determine the scalarization parameters of each function in advance. For nonsmooth single-objective optimization, standard methods include subgradient methods [17], bundle methods [11], and gradient sampling methods [3]. For nonsmooth multiobjective optimization, the literature is a lot more scarce. The subgradient method was generalized to the multiobjective optimization in [1, 6], but they acknowledged computational inefficiency limiting its real-world applicability. The proximal point method was generalized to convex vector optimization problems in [2], where differentiability of the objectives is not required. The multiobjective version of the proximal bundle method was proposed in [13]. In [16], the MultiSQP-GS method was proposed for

*Corresponding author.
E-mail address: zhaoyongty@126.com (Yong Zhao), 1332967558@qq.com (Chun-Lin You), chenwangff@163.com (Wang Chen)

2020 Mathematics Subject Classification: 90C26, 90C29.
Accepted: June 04, 2025.

constrained nonsmooth multiobjective optimization, integrating a Sequential Quadratic Programming (SQP) framework with the gradient sampling technique. In [10], an efficient descent method was introduced for unconstrained locally Lipschitz multiobjective optimization problems, combining theoretical computations of descent directions for nonsmooth objectives with practical subdifferential approximation techniques; [18] extends this framework from finite-dimensional to general Hilbert spaces.

It is worth noting that these methods either do not consider convergence rate analysis or have a slow convergence rate. Motivated by the works in [8, 12, 15], we propose the gradient-free method for solving a class of nonsmooth nonconvex multiobjective optimization problems. Under some mild conditions, we establish the convergence rate of the algorithm where the optimality gap is measured in terms of (δ, ϵ)-Goldstein stationary point. Numerical experiments are presented to demonstrate the effectiveness of the proposed method.

This paper is organized as follows. Section 2 establishes foundational definitions. Section 3 develops the randomized smoothing method. Section 4 rigorously analyzes the convergence properties of our method. Finally, Section 5 presents numerical experiments.

2. Preliminaries

Consider the following multiobjective optimization problem

$$\min_{x\in\mathbb{R}^n} \; F(x) := (F_1(x), \cdots, F_m(x)), \tag{2.1}$$

where $F_i : \mathbb{R}^n \to \mathbb{R}, i = 1, \cdots, m$, are Lipschitz continuous functions.

The optimum for (2.1) is usually given by Pareto optimality: for a point $x^* \in \mathbb{R}^n$, x^* is a Pareto optimal solution of (2.1) if there exists no $x \in \mathbb{R}^n$ such that $F_i(x) \leq F_i(x^*)$ for all $i = 1, \cdots, m$ and $F_j(x) < F_j(x^*)$ for at least one index j. In practice, to check whether a given point is Pareto optimal, we need optimality conditions. In the smooth case, there are the well-known Karush–Kuhn–Tucker (KKT) conditions [14], which are based on the gradients of the objective functions. In case the objective functions are merely Lipschitz continuous, the KKT conditions can be generalized using the concept of subdifferentials. In the following, we recall the required definitions and results from nonsmooth analysis.

Definition 2.1. For $i = 1, \cdots, m$, let $\Omega_i \subset \mathbb{R}^n$ be the set of points where F_i is not differentiable. Then,

$$\partial F_i(x) := \operatorname{conv}\left(\{\xi \in \mathbb{R}^n : \exists \{x_j\}_j \in \mathbb{R}^n \setminus \Omega_i \text{ with } x_j \to x \text{ and } \nabla F_i(x_j) \to \xi \text{ for } j \to \infty\}\right)$$

is the (Clarke) subdifferential of F_i in x.

By employing the subdifferential, a necessary optimality condition can be formulated for Lipschitz-continuous MOPs.

Theorem 2.2. [10] Let $x \in \mathbb{R}^n$ be Pareto optimal. Then,

$$0 \in \operatorname{conv}\left(\bigcup_{i=1}^{k} \partial F_i(x)\right).$$

3. Randomized Smoothing

It is noted that the functions $F_i, i = 1, \cdots, m$, in problem (2.1) are nonsmooth. To address this, the randomized smoothing method is employed to approximate the original problem, thereby providing a foundation for developing gradient-free methods. Letting $\mathbb{P}$ denote a uniform distribution over the unit ball in the ℓ_2-norm, define $F_i^\delta(x) = \mathbb{E}_{u\sim\mathbb{P}}[F_i(x + \delta u)]$. The smoothed functions $F_i^\delta (i = 1, \cdots, m)$ then exhibit the following properties.

Lemma 3.1. [8] Suppose that for each $i = 1, \cdots, m$, F_i is L-Lipschitz. Then, we have
(i) $|F_i^\delta(x) - F_i(x)| \leq \delta L$;
(ii) F_i^δ is differentiable with the $\frac{cL\sqrt{n}}{\delta}$-Lipschitz gradient where $c > 0$ is a constant and $n \geq 1$ is the problem dimension.

In order to ensure convergence of the algorithm to an approximate stationary point in finite-time, we introduce the following definition.

Definition 3.2. Let $\delta \geq 0$, $x \in \mathbb{R}^n$, $\mathbb{B}_\delta(x) := \{y \in \mathbb{R}^n : \|y - x\| \leq \delta\}$ and $i \in \{1, \cdots, m\}$. Then,

$$\partial_\delta F_i(x) := \operatorname{conv}\left(\bigcup_{y \in \mathbb{B}_\delta(x)} \partial F_i(y)\right)$$

is the (Goldstein) δ-subdifferential of F_i in x.

Note that for each $i = 1, \cdots, m$, $\partial_\delta F_i(x) = \partial_0 F_i(x)$ and $\partial F_i(x) \subset \partial_\delta F_i(x)$. For $\delta \geq 0$, we define for the multiobjective setting

$$\partial_\delta F(x) := \operatorname{conv}\left(\bigcup_{i=1}^{m} \partial_\delta F_i(x)\right).$$

The Goldstein subdifferential of F_i at x is defined as the convex hull of the union of all generalized gradients at points within a δ-ball centered at x. Therefore, we define the (δ, ϵ)-Goldstein stationary points for (2.1) as follows, which serve as an optimality condition for general nonsmooth nonconvex multiobjective optimization problems.

Definition 3.3. [10] $x \in \mathbb{R}^n$ is a (δ, ϵ)-Goldstein stationary point if the following statement holds:

$$\min\{\|g\| : g \in \partial_\delta F(x)\} \leq \epsilon.$$

Lemma 3.4. [12] Suppose that for each $i = 1, \cdots, m$, F_i is L-Lipschitz. Then, we have

$$\nabla F_i^\delta(x) \in \partial_\delta F_i(x).$$

4. Convergence Analysis

In this section, we first present the gradient-free method for solving (2.1), which is given in Algorithm 1. We subsequently analyze the non-asymptotic convergence properties of the proposed method.

Lemma 4.1. Suppose that for each $i = 1, \cdots, m$, F_i is L-Lipschitz and let $\{g_i^k\}_{k=0}^{K-1}$ and $\{x^k\}_{k=0}^{K-1}$ be generated by Algorithm 1. Then, we have

$$\mathbb{E}[g_i^k | x^k] = \nabla F_i^\delta(x^k) \text{ and } \mathbb{E}[\|g_i^k - \nabla F_i^\delta(x^k)\|^2 | x^k] \leq \frac{16\sqrt{2\pi} n L^2}{N_k}.$$

Proof. By the definition of g_i^k and the symmetry of the distribution of w_j^k, we have

$$\begin{aligned}
\mathbb{E}[g_i^k | x^k] &= \mathbb{E}\left[\frac{n}{2N_k\delta}\sum_{j=1}^{N_k}\left(F_i(x^k + \delta w_j^k) - F_i(x^k - \delta w_j^k)\right) w_j^k | x^k\right] \\
&= \frac{1}{2}\left(\mathbb{E}\left[\frac{n}{N_k\delta}\sum_{j=1}^{N_k} F_i(x^k + \delta w_j^k) w_j^k | x^k\right] + \mathbb{E}\left[\frac{n}{N_k\delta}\sum_{j=1}^{N_k} F_i(x^k + \delta(-w_j^k))(-w_j^k) | x^k\right]\right) \\
&= \frac{1}{2}(\nabla F_i^\delta(x^k) + \nabla F_i^\delta(x^k)) = \nabla F_i^\delta(x^k).
\end{aligned}$$

Algorithm 1: Gradient-Free Method for (2.1)

Input: Initial point $x^0 \in \mathbb{R}^n$, stepsize $\eta > 0$, problem dimension $n \geq 1$, smoothing parameter δ, batch size N_k and iteration number $K \geq 1$.

for $k = 0, \cdots, K-1$ **do**

Sample $w^k \in \mathbb{R}^n$ uniformly from a unit sphere in $\mathbb{R}^n$.

Compute $g_i^k = \frac{n}{2N_k\delta}\sum_{j=1}^{N_k}\left(F_i(x^k+\delta w_j^k) - F_i(x^k-\delta w_j^k)\right) w_j^k$.

Compute $\lambda^k \in \mathbb{R}^m$ by solving

$$\min_{\lambda\in\mathbb{R}^m} \left\|\sum_{i=1}^m \lambda_i g_i^k\right\|^2$$

$$\text{s.t. } \lambda \geq 0 \text{ and } \sum_{i=1}^m \lambda_i = 1.$$

Denote $d_k = \sum_{i=1}^m \lambda_i^k g_i^k$ and compute $x^{k+1} = x^k - \eta d_k$.

end

Output: x^R, where $R \in \{0, \cdots, K-1\}$ is uniformly sampled.

It remains to show that $\mathbb{E}[\|g_i^k - \nabla F_i^\delta(x^k)\|^2|x^k] \leq \frac{16\sqrt{2\pi}nL^2}{N_k}$. By the definition of g_i^k, we have zyc

$$\begin{aligned}
&\mathbb{E}[\|g_i^k - \nabla F_i^\delta(x^k)\|^2|x^k] \\
&= \mathbb{E}\left[\left\|\frac{n}{2N_k\delta}\sum_{j=1}^{N_k}\left(F_i(x^k+\delta w_j^k) - F_i(x^k-\delta w_j^k)\right) w_j^k - \nabla F_i^\delta(x^k)\right\|^2 |x^k\right] \\
&= \frac{1}{N_k^2}\mathbb{E}\left[\left\|\sum_{j=1}^{N_k}\left(\frac{n}{2\delta}(F_i(x^k+\delta w_j^k) - F_i(x^k-\delta w_j^k))w_j^k) - \nabla F_i^\delta(x^k)\right)\right\|^2 |x^k\right] \\
&= \frac{1}{N_k^2}\sum_{j=1}^{N_k}\mathbb{E}\left[\left\|\frac{n}{2\delta}(F_i(x^k+\delta w_j^k) - F_i(x^k-\delta w_j^k))w_j^k) - \nabla F_i^\delta(x^k)\right\|^2 |x^k\right] \\
&\leq \frac{1}{N_k^2}\sum_{j=1}^{N_k} 16\sqrt{2\pi}nL^2 = \frac{16\sqrt{2\pi}nL^2}{N_k},
\end{aligned}$$

where the inequality follows from Lemma D.1 in [12]. □

Theorem 4.2. Suppose that for each $i = 1, \cdots, m$, F_i is L-Lipschitz and let $\{x^k\}_{k=0}^{K-1}$ be generated by Algorithm 1. Then, we have

$$\frac{1}{K}\sum_{k=0}^{K-1}\mathbb{E}[\|\sum_{i=1}^m \lambda_i^k \nabla F_i^\delta(x^k)\|^2] \leq \left(\frac{32\sqrt{2\pi}nL^2}{K} + \frac{64\sqrt{2\pi}nL^2}{K\eta}\right)\sum_{k=0}^{K-1}\frac{1}{N_k} + \frac{8(F_i^\delta(x^0)) - \mathbb{E}[F_i^\delta(x^K)]}{K\eta}. \tag{4.1}$$

Proof. By Lemma 3.1, we have F_i^δ is differentiable and $\frac{cL\sqrt{n}}{\delta}$-Lipschitz gradient, where $c > 0$ is a constant. This implies

$$F_i^\delta(x^{k+1}) \leq F_i^\delta(x^k) - \eta\langle\nabla F_i^\delta(x^k), d_k\rangle + \frac{cL\sqrt{d}}{2\delta}\eta^2\|d_k\|^2.$$

Taking the expectation of both sides conditioned on x^k, we have

$$\begin{aligned}\mathbb{E}[F_i^\delta(x^{k+1})|x^k] \leq & F_i^\delta(x^k) - \eta\mathbb{E}[\langle\nabla F_i^\delta(x^k), d_k\rangle|x^k] + \frac{cL\sqrt{n}}{2\delta}\eta^2\mathbb{E}[\|d_k\|^2|x^k] \\ = & F_i^\delta(x^k) - \eta\mathbb{E}[\langle\nabla F_i^\delta(x^k) - g_i^k, d_k\rangle|x^k] - \eta\mathbb{E}[\langle g_i^k, d_k\rangle|x^k] + \frac{cL\sqrt{n}}{2\delta}\eta^2\mathbb{E}[\|d_k\|^2|x^k] \\ \leq & F_i^\delta(x^k) + \frac{1}{2}\mathbb{E}[\|\nabla F_i^\delta(x^k) - g_i^k\|^2|x^k] + \frac{1}{2}\eta^2\mathbb{E}[\|d_k\|^2|x^k] - \eta\mathbb{E}[\|d_k\|^2|x^k] \\ & + \frac{cL\sqrt{n}}{2\delta}\eta^2\mathbb{E}[\|d_k\|^2|x^k] \\ \leq & F_i^\delta(x^k) + \frac{8\sqrt{2\pi}nL^2}{N_k} - (\eta - \frac{1}{2}\eta^2 - \frac{cL\sqrt{n}\eta^2}{2\delta})\mathbb{E}[\|d_k\|^2|x^k],\end{aligned}$$

where the second inequality follows from Young inequality and $\langle g_i^k, d_k\rangle \geq \|d_k\|^2$. By setting $1 - \frac{1}{2}\eta - \frac{cL\sqrt{d}\eta}{2\delta} \geq \frac{1}{4}$, i.e., $\eta \leq \frac{3\delta}{2(\delta + cL\sqrt{d})}$. Therefore, we have

$$\frac{\eta}{4}\mathbb{E}[\|d_k\|^2|x^k] \leq -\mathbb{E}[F_i^\delta(x^{k+1})|x^k] + F_i^\delta(x^k) + \frac{8\sqrt{2\pi}nL^2}{N_k}.$$

Taking the expectation of both sides and summing up the above inequality over $k = 0, \cdots, K-1$ yields that

$$\frac{1}{K}\sum_{k=0}^{K-1}\mathbb{E}[\|d_k\|^2] \leq \frac{4(F_i^\delta(x^0)) - \mathbb{E}[F_i^\delta(x^K)]}{K\eta} + \frac{32\sqrt{2\pi}nL^2}{K\eta}\sum_{k=0}^{K-1}\frac{1}{N_k}.$$

Since

$$\begin{aligned}\|\sum_{i=1}^m \lambda_i^k \nabla F_i^\delta(x^k)\|^2 &\leq 2\|\sum_{i=1}^m \lambda_i^k\nabla F_i^\delta(x^k) - d_k\|^2 + 2\|d_k\|^2 \\ &\leq 2\sum_{i=1}^m \lambda_i^k\|\nabla F_i^\delta(x^k) - g_i^k\|^2 + 2\|d_k\|^2,\end{aligned}$$

thus, we have

$$\begin{aligned}\frac{1}{K}\sum_{k=0}^{K-1}\mathbb{E}[\|\sum_{i=1}^m\lambda_i^k\nabla F_i^\delta(x^k)\|^2] \leq & \frac{32\sqrt{2\pi}nL^2}{K}\sum_{k=0}^{K-1}\frac{1}{N_k} + \frac{8(F_i^\delta(x^0)) - \mathbb{E}[F_i^\delta(x^K)]}{K\eta} \\ & + \frac{64\sqrt{2\pi}nL^2}{K\eta}\sum_{k=0}^{K-1}\frac{1}{N_k},\end{aligned}$$

which concludes the proof. □

In the following theorem, we establish the convergence rate of the proposed algorithm.

Theorem 4.3. Suppose that for each $i = 1, \cdots, m$, F_i is L-Lipschitz and let $\{x^k\}_{k=0}^{K-1}$ be generated by Algorithm 1. For each $i = 1, \cdots, m$, assume that F_i has a lower bound $F_i^{\min}$ and $\sum_{k=0}^{\infty}\frac{1}{N_k} \leq M$ with a constant $M > 0$. Let $F^{\min} = \min_{i=1,\cdots,m} F_i^{\min}$ and $F^{\max}(x_0) = \max_{i=1,\cdots,m} F_i(x_0)$, where x_0 is a given initial point. Then, we have

$$\mathbb{E}\left[\min\{\|g\| : g \in \partial_\delta F(x^R)\}\right] \leq \frac{\left((81nL^2 + 161nL^2/\eta)M + 8(F^{\max}(x_0) - F^{\min} + \delta L)/\eta\right)^{\frac{1}{2}}}{\sqrt{K}}.$$

Proof. By Lemma 3.1, we have $F_i(x_0) \leq F_i^\delta(x_0) \leq F_i(x_0) + \delta L$. In addition, we see from the definition of F_i^δ that $F_i^\delta(x) \geq \inf_{x\in\mathbb{R}^n} F_i(x)$ for any $x \in \mathbb{R}^n$ and thus $\mathbb{E}[F_i^\delta(x^K)] \geq \inf_{x\in\mathbb{R}^n} F_i(x)$, which implies that

$$F_i^\delta(x^0) - \mathbb{E}[F_i^\delta(x^K)] \leq F_i(x_0) - F^{\min} + \delta L \leq F^{\max}(x_0) - F^{\min} + \delta L.$$

Therefore, it follows from (4.1) that we have

$$\begin{aligned}\frac{1}{K}\sum_{k=0}^{K-1}\mathbb{E}[\|\sum_{i=1}^m \lambda_i^k \nabla F_i^\delta(x^k)\|^2] &\leq \left(\frac{32\sqrt{2\pi}nL^2}{K} + \frac{64\sqrt{2\pi}nL^2}{K\eta}\right)\sum_{k=0}^{K-1}\frac{1}{N_k} + \frac{8(F^{\max}(x_0) - F^{\min} + \delta L)}{K\eta} \\ &\leq \frac{(81nL^2 + 161nL^2/\eta)M}{K} + \frac{8(F^{\max}(x_0) - F^{\min} + \delta L)}{K\eta}.\end{aligned}$$

Since

$$\frac{1}{K}\sum_{k=0}^{K-1}\mathbb{E}[\min_{\lambda\in\triangle}\|\sum_{i=1}^m \lambda_i \nabla F_i^\delta(x^k)\|^2] \leq \frac{1}{K}\sum_{k=0}^{K-1}\mathbb{E}[\|\sum_{i=1}^m \lambda_i^k \nabla F_i^\delta(x^k)\|^2]$$

and the random count $R \in \{0, 1, \cdots, T-1\}$ is uniformly sampled, we have

$$\begin{aligned}\mathbb{E}[\min_{\lambda\in\triangle}\|\sum_{i=1}^m \lambda_i \nabla F_i^\delta(x^R)\|^2] &= \frac{1}{K}\sum_{k=0}^{K-1}\mathbb{E}[\min_{\lambda\in\triangle}\|\sum_{i=1}^m \lambda_i \nabla F_i^\delta(x^k)\|^2] &(4.2)\\ &\leq \frac{(81dL^2 + 161dL^2/\eta)M}{K} + \frac{8(F^{\max}(x_0) - F^{\min} + \delta L)}{K\eta}. &(4.3)\end{aligned}$$

By Lemma 3.4, we have $\nabla F_i^\delta(x^R) \in \partial_\delta F_i(x^R)$ for each $i = 1, \cdots, m$, which means that

$$\operatorname{conv}\left(\bigcup_{i=1}^m \nabla F_i^\delta(x^R)\right) \subset \partial_\delta F(x^R).$$

This together with (4.2) implies that

$$\begin{aligned}\mathbb{E}\left[\min\{\|g\| : g \in \partial_\delta F(x^R)\}\right] &\leq \mathbb{E}[\min_{\lambda\in\triangle}\|\sum_{i=1}^m \lambda_i \nabla F_i^\delta(x^R)\|] \\ &\leq \frac{\left((81nL^2 + 161nL^2/\eta)M + 8(F^{\max}(x_0) - F^{\min} + \delta L)/\eta\right)^{\frac{1}{2}}}{\sqrt{K}}.\end{aligned}$$

This completes the proof. □

Theorem 4.3 establishes a convergence rate of $\mathcal{O}(n^{\frac{1}{2}}\epsilon^{-1})$ for a randomized gradient-based method, measured in terms of (δ, ϵ)-Goldstein stationarity for (2.1).

5. Numerical Experiments

In this section, we report the results of computational experiments, showing the performance of the proposed algorithm. The code is written in Matlab software and run on a computer with the following characteristics: Intel(R) Core(TM) i7-11390H 3.40GHz, 16 GB RAM. In order to solve the subproblem in the step 4 of Algorithm 1, the Matlab solver `quadprog` is employed.

We consider two test problems in the experiments, which can be found in [10]. The first problem is (2.1) with $m = 2$, $n = 2$, and

$$F(x) = \begin{pmatrix}\max\{x_1^2 + (x_2-1)^2 + x_2 - 1, -x_1^2 - (x_2-1)^2 + x_2 + 1\} \\ -x_1 + 2(x_1^2 + x_2^2 - 1) + 1.75|x_1^2 + x_2^2 - 1|\end{pmatrix}^\top.$$

The second problem is the sparse optimization problem with an ℓ_1 penalty term, which can be described as

$$\min_{x\in\mathbb{R}^n} G(x) + \lambda\|x\|_1, \tag{5.1}$$

where $G : \mathbb{R}^n \to \mathbb{R}$ is a locally Lipschitz continuous function and $\lambda \geq 0$ is a regular parameter. Gebken and Peitz [10] reformulated this problem (5.1) as a bi-objective optimization problem with

$$F_1(x) = G(x)$$

and

$$F_2(x) = \|x\|_1,$$

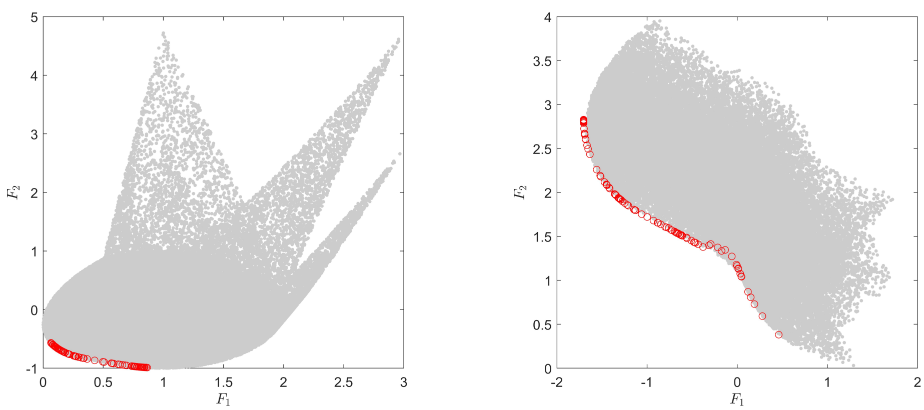
where

$$G(x) = \left(x_1 - \frac{1}{4}\right)^2 + \left(x_2 - \frac{1}{2}\right)^2 + (x_3 - 1)^4 - \frac{1}{2}\left(x_3 - \frac{1}{4}\right)^3.$$

Owing to the nonsmooth nature of F_2, the resulting MOP is inherently nonsmooth. For brevity, we refer to these two problems as TP1 and TP2, respectively.

Considering that Algorithm 1 is a single-point iterative method, we adopt a multi-start strategy to obtain an effective approximation of the true Pareto front. Specifically, Algorithm 1 uses 100 initial points on each test problem, which are selected uniformly from the given box. For TP1, the box is set as $[-1, 1]^2$, while for TP2, it is $[0, 0, 0] \times [1, 1, 2]$. The parameters for Algorithm 1 are set as follows: $\eta = 0.008$, $\delta = 0.005$, $K = 1000$, and $N_k = \lceil 0.9984^{-(k+1)} \rceil$.

For each problem, we discretize the above-mentioned box into fine grid points and plot all image points. Thus, it provides a good representation of the image space of F. Figure 1 gives the image points and the final solutions obtained by Algorithm 1. In this figure, the gray points mean the image points, while the red circle points represent the final solutions. As can be seen, the proposed algorithm is capable of satisfactorily estimating the Pareto front of the considered problems.



(a) TP1 (b) TP2

FIGURE 1. The image points and the final solutions obtained by Algorithm 1

6. CONCLUSION

In this paper, we proposed a gradient-free method for solving multiobjective optimization problems with locally Lipschitz continuous objective functions. We established the convergence rate of the proposed algorithm and demonstrated its performance through numerical experiments on two benchmark problems.

STATEMENTS AND DECLARATIONS

The authors declare that they have no conflict of interest, and the manuscript has no associated data.

Acknowledgements

This work was supported by the Science and Technology Research Program of Chongqing Municipal Education Commission (No. KJQN202400760), the Chongqing Postdoctoral Research Project Special Grant (No. 2023CQBSHTB2026), and the Science and Technology Research Program of Chongqing Education Commission of China (No. KJQN202400516).

References

[1] J. Bello Cruz. A subgradient method for vector optimization problems. *SIAM Journal on Optimization*, 23(4):2169–2182, 2013.

[2] H. Bonnel, A. N. Iusem, and B. F. Svaiter. Proximal methods in vector optimization. *SIAM Journal on Optimization*, 15(4):953–970, 2005.

[3] J. V. Burke, A. S. Lewis, and M. L. Overton. A robust gradient sampling algorithm for nonsmooth, nonconvex optimization. *SIAM Journal on Optimization*, 15(3):751–779, 2005.

[4] J.-H. Cho, Y. Wang, R. Chen, K. S. Chan, and A. Swami. A survey on modeling and optimizing multi-objective systems. *IEEE Communications Surveys & Tutorials*, 19(3):1867–1901, 2017.

[5] C. A. C. Coello. *Evolutionary Algorithms for Solving Multi-Objective Problems*. Springer, New York, 2007.

[6] J. X. da Cruz Neto, G. Da Silva, O. P. Ferreira, and J. O. Lopes. A subgradient method for multiobjective optimization. *Computational Optimization and Applications*, 54(3):461–472, 2013.

[7] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002.

[8] J. C. Duchi, P. L. Bartlett, and M. J. Wainwright. Randomized smoothing for stochastic optimization. *SIAM Journal on Optimization*, 22(2):674–701, 2012.

[9] M. Ehrgott. *Multicriteria Optimization*. Springer Verlag, Berlin, 2005.

[10] B. Gebken and S. Peitz. An efficient descent method for locally lipschitz multiobjective optimization problems. *Journal of Optimization Theory and Applications*, 188(3):696–723, 2021.

[11] K. C. Kiwiel. Proximity control in bundle methods for convex nondifferentiable minimization. *Mathematical Programming*, 46(1):105–122, 1990.

[12] T. Lin, Z. Zheng, and M. Jordan. Gradient-free methods for deterministic and stochastic nonsmooth nonconvex optimization. *Advances in Neural Information Processing Systems*, 35:26160–26175, 2022.

[13] M. Mäkelä, N. Karmitsa, and O. Wilppu. Multiobjective proximal bundle method for nonsmooth optimization. Technical report, TUCS, 2014.

[14] K. Miettinen. *Nonlinear Multiobjective Optimization*. Springer, New York, 1998.

[15] Y. Nesterov. Smooth minimization of non-smooth functions. *Mathematical Programming*, 103:127–152, 2005.

[16] M. Rashidi and M. Soleimani-damaneh. MultiSQP-GS: A sequential quadratic programming algorithm via gradient sampling for nonsmooth constrained multiobjective optimization. *Computational Optimization and Applications*, 89(3):729–767, 2024.

[17] N. Z. Shor. *Minimization Methods for Non-Differentiable Functions*. Springer Berlin, Heidelberg, 2012.

[18] K. Sonntag, B. Gebken, G. Müller, S. Peitz, and S. Volkwein. A descent method for nonsmooth multiobjective optimization in hilbert spaces. *Journal of Optimization Theory and Applications*, 203(1):455–487, 2024.

[19] M. M. Wiecek, M. Ehrgott, and A. Engau. Continuous multiobjective programming. In S. Greco, M. Ehrgott, and J. Figueira, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 739–815. Springer, New York, 2016.