Superior Runtime Guarantees for the MOEA/D Multi-Objective Optimizer via Weighted-Sum Decomposition

Danyang Zhang[1#], Zerong Zhong[1#], Weijie Zheng[1,2*], Benjamin Doerr[3]

[1]School of Computer Science and Technology, State Key Laboratory of Smart Farm Technologies and Systems, International Research Institute for Artificial Intelligence, Harbin Institute of Technology, Shenzhen, China
[2]Pengcheng Laboratory, Shenzhen, China
[3]Laboratoire d'Informatique (LIX), CNRS, École Polytechnique, Institut Polytechnique de Paris, Palaiseau, France
zhangdanyang@stu.hit.edu.cn, zhongzerong1@outlook.com, zhengweijie@hit.edu.cn, lastname@lix.polytechnique.fr

Abstract

The MOEA/D is the most popular decomposition-based evolutionary algorithm to solve multi-objective optimization problems. However, among the two common decomposition approaches, weighted-sum and Tchebycheff, the existing theoretical research almost exclusively focuses on the latter one. In this first complete mathematical runtime analysis for the MOEA/D using the original weighted-sum decomposition, we show that this variant of the algorithm solves the classic ONEMINMAX benchmark considerably faster than both the MOEA/D with Tchebycheff decomposition and many other classic algorithms such as the NSGA-II, NSGA-III, SMS-EMOA, and SPEA2. More precisely, we show that already a logarithmic number of subproblems suffices for the algorithm to be efficient, and then typically $O(n \log^2 n)$ function evaluations suffice to compute the full Pareto front. This beats the other algorithms by a factor of $\Theta(n/\log n)$. For a second benchmark, the ONEJUMPZEROJUMP problem, we show a speed-up by a factor of $\Theta(n)$. Overall, this work shows that a further development of the weighted-sum approach might be fruitful.

Introduction

Multi-objective evolutionary algorithms (MOEAs) have been widely used for solving optimization problems with multiple conflicting objectives. Among the decomposition-based MOEAs, the multi-objective evolutionary algorithm based on decomposition (MOEA/D) proposed by Zhang and Li (2007) is the most popular one. This algorithm decomposes the multi-objective problem into multiple single-objective subproblems, and then co-evolutionarily solves them. The weighted-sum and Tchebycheff approach are two common decomposition strategies.

The first runtime analysis of the MOEA/D was conducted by Li et al. (2016). They proved that with Tchebycheff decomposition, the algorithm (with standard bitwise mutation and no crossover) will cover the full Pareto front for the COCZ problem in expected $O(n^2 \log n)$ fitness evaluations. For the LPTNO problem, using the Tchebycheff decomposition with a modified set of weight vectors will result in an expected runtime of $O(n^3)$. Huang et al. (2019) proved that with the Tchebycheff decomposition, the usage of the crossover will reduce the needed size of the subproblems into 3, and thus will result in the better runtime of $O(n \log n)$ for COCZ and $O(n^2)$ for LPTNO. Still considering the Tchebycheff decomposition, Huang and Zhou (2020) considered to replace the standard bitwise mutation by the somatic contiguous hypermutation operators from the community of artificial immune systems, and proved that for this replacement (also without crossover), the same runtime $O(n^2 \log n)$ is reached for COCZ but a better runtime of $O(n^2 \log n)$ for LPTNO, compared with the standard bitwise mutation. Further, Corus, Oliveto, and Yazdani (2025) proposed a fast contiguous somatic hypermutation operator, and proved that this variant (with the Tchebycheff decomposition) further achieves a better runtime of $O(n \log n)$ for COCZ and maintains the runtime of $O(n^2 \log n)$ for LPTNO. Doerr, Krejca, and Weeks (2024) proved that under the assumption that the subproblems spread uniformly across the Pareto front, the MOEA/D with the Tchebycheff decomposition and $N+1$ subproblems will cover the full Pareto front for OMM in expected runtime of $O(nN \log n + n^{n/(2N)} N \log n)$ for standard bit mutation, and $O(nN \log n + n^{\beta} \log n)$ for power-law mutation with exponent $\beta \in (1, 2)$. Zheng (2024) proved the expected runtime of $O(\max\{k, 1\} n \log n)$ for cover the Pareto front of the OMM variant with conflict degree of k.

There are only two runtime results for the MOEA/D with other decomposition approaches. The weighted-sum, Tchebycheff, and penalty-based boundary intersection decomposition approaches are discussed in Huang et al. (2021) for solving many-objective optimization problem mCOCZ and mLOTZ. Instead of analyzing the runtime to cover the full Pareto front, they only obtained the expected runtime for the algorithm to reach the optimum of any given subproblem. Since it is not answered whether the Pareto front can be covered by the construction of the subproblems or by the optimization process in solving these subproblems, this runtime is not complete, and cannot be transferred to derive the rigorous runtime for the coverage of the full Pareto front. A special variant of the weighted-sum decomposition

[#]These authors contributed equally.
[*]Corresponding author.

approach was discussed in Do et al. (2023) and tailored for the multi-objective minimum weight base problem. They proved that it finds all extreme points within expected fixed-parameter polynomial time. In summary, the majority of the runtime results for the MOEA/D are for the Tchebycheff decomposition, and have given rise to fruitful results for the effect of crossover, other mutation strategies, etc. Although some theoretical works tried to analyze the weighted-sum approach, the runtime they obtained was incomplete or the weighted-sum approach was modified. Till now, it is still an open problem about the complete runtime for the MOEA/D using the original weighted-sum decomposition.

Our contributions: This paper will conduct the first complete runtime analysis of MOEA/D using the original weight-sum decomposition. Following the analysis for the first runtime results for the MOEA/D using the Tchebycheff decomposition (Li et al. 2016), and for other MOEAs (like NSGA-II (Zheng, Liu, and Doerr 2022; Zheng and Doerr 2023a), NSGA-III (Wietheger and Doerr 2023; Opris et al. 2024), SMS-EMOA (Bian et al. 2023; Zheng and Doerr 2024), SPEA2 (Ren et al. 2024; Alghouass et al. 2025)), we only consider the (standard bitwise) mutation as the only variation operator (the crossover is not employed) for the MOEA/D via the weighted-sum decomposition, and compare its runtime also with other MOEAs that also employ the standard bitwise mutation only. The effect of the crossover and the performance of other mutation operators will be our interesting future work. For the classic OneMinMax problem with size n, we will prove that although the weighted-sum decomposition cannot ensure every Pareto front point mapped to the optimum of a certain subproblem, the event that the whole Pareto front is covered when all subproblems reach their optimum for the first time, happens with probability at least $1 - n^{-\Omega(1)}$, if the number of the subproblems is $\Omega(\log n)$ (see Lemmas 3 and 5). Thus, we will obtain an expected runtime of $O(n \log^2 n)$ (see Theorem 6). We note that the runtime of the MOEA/D with the original Tchebycheff decomposition (Li et al. 2016) and of other original MOEAs with standard bitwise mutation only, like the GSEMO (Giel and Lehre 2010), NSGA-II (Zheng and Doerr 2023a), NSGA-III (setting $m = 2$ for mOMM in (Opris et al. 2024)), SMS-EMOA (Zheng and Doerr 2024), SPEA2 (setting $m = 2$ for mOMM in (Ren et al. 2024)), is $O(n^2 \log n)$. Hence the weighted-sum decomposition obtains a $\Theta(n/\log n)$ speed-up.

We will also extend our analysis to the popular multi-modal OneJumpZeroJump benchmark, and will show that conditional on a similar event to the one for OneMinMax, which happens with probability $1 - n^{-\Omega(1)}$, the weighted-sum decomposition has the expected runtime of $O(n^k)$ to cover the full Pareto front (See Theorem 11). In contrast to the runtime of $O(n^{k+1})$ for the Tchebycheff decomposition (proved in Corollary 15), and for other original MOEAs with standard bitwise mutation only, like the GSEMO (Zheng and Doerr 2023b), NSGA-II (Doerr and Qu 2023), SMS-EMOA (Bian et al. 2023), NSGA-III (Opris 2025), SPEA2 (setting $m = 2$ for mOJZJ in (Ren et al. 2024)), we see a speed-up by a factor of $\Theta(n)$.

Preliminaries

Bi-Objective Optimization Problem

Formally, a pseudo-Boolean bi-objective optimization problem is to maximize $f = (f_1, f_2) : \{0,1\}^n \to \mathbb{R}^2$, where $n \in \mathbb{N}$ is the problem size. For any $x, y \in \{0,1\}^n$, we say that x weakly dominates y, denoted as $x \succeq y$, if $f_1(x) \geq f_1(y)$ and $f_2(x) \geq f_2(y)$. If at least one of inequalities is strict, that is, $f_1(x) > f_1(y)$ or $f_2(x) > f_2(y)$, then we say that x dominates y, denoted as $x \succ y$. A solution $x \in \{0,1\}^n$ is called Pareto optimal if there exists no solution $y \in \{0,1\}^n$ such that $y \succ x$. The set of all Pareto optimal solutions is called the Pareto set, and the corresponding set of objective function values is Pareto front.

Following the theory community of MOEAs (like (Laumanns et al. 2002; Laumanns, Thiele, and Zitzler 2004; Giel and Lehre 2010; Zheng and Doerr 2023a; Bian et al. 2023; Wietheger and Doerr 2023; Opris et al. 2024)), the runtime in this paper refers to the number of function evaluations required to cover the full Pareto front. Additionally, in this paper, we use $[a, b]$ to denote the set $\{a, a+1, \ldots, b\}$ for $a \leq b$ and $a, b \in \mathbb{N}$, and for any $x \in \{0,1\}^n$, we use $|x|_1 = \sum_{i=1}^n x_i$ to denote the number of ones in x, and $|x|_0 = n - |x|_1$ for the number of zeros in x.

The MOEA/D

The MOEA/D (Zhang and Li 2007) first decomposes a multi-objective problem into $H + 1$ single-objective optimization subproblems. The weighted-sum and Tchebycheff are two common decomposition approaches. Let $\lambda^i = (\lambda_1^i, \lambda_2^i), i = 0, ..., H$ be weight vectors with $\lambda_1^i, \lambda_2^i \geq 0$ and $\lambda_1^i + \lambda_2^i = 1$. As in (Zhang and Li 2007), usually $\lambda_1^i = \frac{i}{H}$ and $\lambda_2^i = 1 - \frac{i}{H}$. Then for the maximization of $f = (f_1, f_2)$, the i-th subproblem ($i = 0, ..., H$) via weighted-sum decomposition is to maximize

$$g_i(x) = \lambda_1^i f_1(x) + \lambda_2^i f_2(x), \quad (1)$$

and the i-th subproblem via Tchebycheff decomposition is to minimize

$$g_i(x|z^*) = \max\{\lambda_1^i |f_1(x) - z_1^*|, \lambda_2^i |f_2(x) - z_2^*|\}, \quad (2)$$

where $z^* = (z_1^*, z_2^*)$ denotes the reference point, which is calculated as $z_1^* = \max\{f_1(x) \mid x \in S\}$ and $z_2^* = \max\{f_2(x) \mid x \in S\}$ for S the set of all solutions generated so far.

After the decomposition, the MOEA/D will solve the subproblems in a co-evolutionary manner, where the solutions for T closest neighbor subproblems will join in the update of the solution w.r.t. the current subproblem (Zhang and Li 2007). Following all existing theoretical analyses of the MOEA/D (e.g. (Li et al. 2016; Doerr, Krejca, and Weeks 2024)), in this work, we also set $T = 1$. In each iteration, each subproblem $i \in [0..H]$ will generate a new solution y from the current solution x_i. For the Tchebycheff decomposition, the reference point z^* is updated accordingly, and the update of z^* is skipped for the weighted-sum decomposition. For the Tchebycheff decomposition, the new solution y will replace x_i as the solution for the i-th subproblem only when y has a smaller or equal value of g_i, and for

Algorithm 1: The MOEA/D maximizing a bi-objective problem $f : \{0,1\}^n \to \mathbb{R}^2$.

Input: A parameter H, weight vectors $\{\lambda^0, \lambda^1, ..., \lambda^H\}$, subproblems $\{g_i\}_{i\in[0..H]}$, a mutation operator: $\{0,1\}^n \to \{0,1\}^n$, a stopping criterion.
Output: A set P usually containing non-dominated solutions.

1: Initialization: $P = \emptyset$. Decompose f into $H+1$ scalar subproblems according to the weight vectors $\{\lambda^0, \lambda^1, ..., \lambda^H\}$. Generate solution $x_i \in \{0,1\}^n$ uniformly at random for each g_i and set $z_1^* = \max_{i\in[0..H]} f_1(x_i)$, $z_2^* = \max_{i\in[0..H]} f_2(x_i)$. For all $i \in [0..H]$, if x_i is not dominated by any element in P, remove all elements in P that are weakly dominated by x_i from P and add x_i into P.
2: **while** stopping criterion is not met **do**
3:   **for** each subproblem $i \in [0..H]$ **do**
4:     Mutation: $y \leftarrow \text{mut}\,(x_i)$;
5:     Update z^*: set $z_1^* \leftarrow \max(z_1^*, f_1(y))$, $z_2^* \leftarrow \max(z_2^*, f_2(y))$;
6:     Update x_i: if $g_i(y|z^*) \le g_i(x_i|z^*)$, then $x_i \leftarrow y$;
7:     Update P: If y is not dominated by any element in P, remove all elements in P that are weakly dominated by y from P and add y into P;
8:   **end for**
9: **end while**

the weighted-sum decomposition, the new solution y will replace x_i only when y has a larger or equal value of g_i. No matter y replaces x_i or not, it will be considered for the update of P. If no element in P dominates y, then all weakly dominated solutions (by y) in P will be removed and y will be added into P. See Algorithm 1 for the whole procedure.

Following the previous theoretical (Li et al. 2016; Doerr, Krejca, and Weeks 2024), we do not set the termination criterion, and calculate the first time when P covers the full Pareto front w.r.t. f. Also, we only consider standard bitwise mutation, where each bit of any solution $x \in \{0,1\}^n$ is independently flipped with a probability of $1/n$, as in (Li et al. 2016; Doerr, Krejca, and Weeks 2024).

$\Theta(n/\log n)$ Runtime Speed-up on OneMinMax

This section will analyze the runtime of the MOEA/D with the weighted-sum decomposition on OneMinMax (OMM for short), and prove a runtime speed-up by a factor of $\Theta(n/\log n)$, compared with the existing results of the MOEA/D with the Tchebycheff decomposition, and other MOEAs. Note that the above comparisons are made for all algorithms with standard bitwise mutation as the only variation operator, as we discussed in section Introduction. The effect of crossover and other mutation strategies will be our interesting future work.

OMM and Decomposed Subproblems

The OMM (Giel and Lehre 2010) is a classic bi-objective counterpart of the OneMax benchmark, where the first objective is to maximize the number of zeros, while the second one is to maximize the number of ones. In this paper, we swap the first and second objectives to align with the OneJumpZeroJump problem discussed in the next section (see Definition 1). This change is purely notational and does not affect the essence of the problem.

Definition 1 ((Giel and Lehre 2010))**.** *For any $x \in \{0,1\}^n$, the bi-objective* OMM $f = (f_1, f_2) : \{0,1\}^n \to \mathbb{R}^2$ *is defined by:*

$$f(x) = (f_1(x), f_2(x)) = (|x|_1, |x|_0).$$

For the maximization of the OMM problem, every solution is Pareto optimal (that is, the Pareto set is $\{x \mid x \in \{0,1\}^n\}$), and the Pareto front is $\{(j, n-j) \mid j \in [0..n]\}$.

Recalling the weighted-sum decomposition discussed before (see (1) in the previous section), we obtain the following results for the decomposed subproblems and the corresponding optimal sets. Due to the limited space, all our proofs are omitted.

Lemma 2. *Via the weighted-sum decomposition with $H+1$ subproblems, the i-th ($i \in [0..H]$) subproblem for the maximization of* OMM *is to maximize*

$$g_i(x) = \left(\frac{2i}{H} - 1\right)|x|_1 + \left(1 - \frac{i}{H}\right)n.$$

The optimal set S_i of the i-th subproblem is

$$S_i = \begin{cases} \{x \mid |x|_1 = 0\}, & \text{if } i < \frac{H}{2}, \\ \{x \mid |x|_1 = [0..n]\}, & \text{if } H \text{ is even and } i = \frac{H}{2}, \\ \{x \mid |x|_1 = n\}, & \text{if } i > \frac{H}{2}. \end{cases}$$

Recalling the Pareto front of $\{(j, n-j) \mid j \in [0..n]\}$, we see that not any Pareto front point (like $(j, n-j)$ with $j \in [1..n-1]$) can be mapped to the optimum for a certain subproblem (like when H is odd). It is widely regarded as a drawback of the weighted-sum decomposition, like in (Zhang and Li 2007). However, in the following subsection, we will show that this "drawback" does not prevent the MOEA/D from covering the full Pareto front (that is, for P to satisfy $\{f(P)\} \supseteq \{(j, n-j) \mid j \in [0..n]\}$).

$O(n\log^2 n)$ Runtime on OMM

Before we establish the runtime of the MOEA/D with weighted-sum decomposition, we first give the following lemma that discusses the population in the initial iteration, which holds for all multi-objective problems defined in the binary space $\{0,1\}^n$, and is independent of the used decomposition approach. We believe that it will be useful for future analysis on the MOEA/D as well, when the initial state of P is required.

Lemma 3. *Let $n \ge 8$. Consider the MOEA/D decomposes the multiobjective f into $H+1$ subproblems. For the initial iteration (step 1 in Algorithm 1), let $N_0 = |\{x_i \mid |x_i|_1 \in [\frac{1}{2}n, \frac{3}{4}n], i < \frac{1}{2}H\}|$ and $N_1 = |\{x_i \mid |x_i|_1 \in [\frac{1}{4}n, \frac{1}{2}n], i > \frac{1}{2}H\}|$. Then $\Pr[\min\{N_0, N_1\} > \frac{1}{32}H] \ge 1 - 2\exp(-\frac{1}{128}H)$.*

We note that $H+1$ subproblems are solved in the procedure of the MOEA/D, and technically, we need to estimate the first time when all subproblems reach their optimum. The expectation of this random variable is naturally upper bounded by summing up $H+1$ random variables, each of which denotes the first time when the optimum of a specific subproblem is reached. The following lemma establishes a general tool for a tighter estimate for such expectation. This result was implicitly used in some previous work, and we now put it as a general tool, as we believe that it will facilitate future analysis as well.

Lemma 4. *Let $m, n \in \mathbb{N}$, $a > 1$ and $m \leq n^{a-1}$. Let $X_1, ..., X_n$, $X_{n+1}, ..., X_{2n}, ..., X_{(m-1)n+1}, ..., X_{mn}$ be independent geometric random variables with success probabilities $p_1, ..., p_{mn}$. Assume that there is a number $C \leq 1$ such that $p_{(j-1)n+i} \geq C\frac{i}{n}$ for all $i \in [1..n]$ and $j \in [1..m]$. Let $Y_j = \sum_{i=(j-1)n+1}^{jn} X_i, j = 1, ..., m$. Then $E[\max\{Y_1, ..., Y_m\}] \leq \frac{1}{C}(1 + a\ln n)n$.*

Now we start to analyze the number of function evaluations (which is obtained by multiplying the number of iterations by the number of subproblems $H+1$) needed by the MOEA/D with the weighted-sum decomposition to cover the full Pareto front for OMM. Technically, recalling the “drawback” mentioned in the previous subsection, we know that one cannot simply use the time when all subproblems reach their optimum to represent the total runtime to cover the full Pareto front, as the full coverage is not unknown. Note that it is quite different from the analysis of the Tchebycheff decomposition (with the standard bitwise mutation only as the variation operator) (Li et al. 2016).

The key idea to establish the runtime is to prove that with high probability, the whole Pareto front is covered by P for the first time when the optima are obtained for certain (almost all) subproblems. Specifically, the Pareto front points $(j, n-j)$ for all $j \in [0, \frac{1}{2}n]$ are obtained when the first half of subproblems (that is, the i-th subproblem with $i \in [0, \frac{1}{2}H)$) (starting from initial solutions x_i with $|x_i|_1 \geq \frac{1}{2}n$) are all successfully solved. Analogously, the Pareto front points $(j, n-j)$ for all $j \in [\frac{1}{2}n, n]$ are expected to be discovered in the second half subproblems (that is, the i-th subproblem with $i \in (\frac{1}{2}H, H]$) (starting from initial solutions x_i with $|x_i|_1 \leq \frac{1}{2}n$). From Lemma 3, we know that with high probability, the algorithm for $\Theta(H)$ subproblems will start with the desired solutions. The following lemma will show that conditional on this initial state, when the number of subproblems $H+1$ is at least of logarithmic order, with high probability $1-n^{\Omega(-1)}$, P will cover the full Pareto front when specific subproblems have obtained their optimum.

Lemma 5. *Let $n \geq 8, c > \frac{500}{3}$ and $H \geq c\ln n$. Consider using MOEA/D with weighted-sum decomposition and standard bitwise mutation to maximize* OMM. *Let N_0 and N_1 be defined as in Lemma 3 and assume $\min\{N_0, N_1\} > \frac{1}{32}H$. Let T_0 denote the first time when $x_0, ..., x_H$ satisfy $x_i = 1^n$ for $i > \frac{1}{2}H$ and $x_i = 0^n$ for $i < \frac{1}{2}H$. Then with probability at least $1 - n^{-(\frac{3}{500}c-1)}$, the whole Pareto front, that is, $\{(j, n-j) \mid j \in [0..n]\}$, is covered by P within T_0 iterations.*

Combining Lemmas 3 and 5, we obtain the following theorem for the runtime of the MOEA/D via weighted-sum decomposition on OMM.

Theorem 6. *Let $n \geq 8, c > \frac{500}{3}$, $a > 1$ and $c\ln n \leq H \leq n^{a-1}-1$. Consider using MOEA/D with weighted-sum decomposition and standard bitwise mutation to maximize* OMM. *Conditional on an event that happens with probability at least $1-3n^{-(\frac{3}{500}c-1)}$, P covers the whole Pareto front of* OMM *within $e(1+a\ln n)n$ expected iterations, that is, within $e(H+1)(1+a\ln n)n$ expected function evaluations.*

From Theorem 6, we know that when using the weighted-sum decomposition with $\Theta(\log n)$ subproblems, the set P covers the full Pareto front typically (conditional on an event that happens with high probability $1-n^{-\Omega(1)}$) in expected $O(n\log^2 n)$ number of function evaluations. We note here the expected runtime is $O(n^2\log n)$ for the MOEA/D with the Tchebycheff decomposition (using the standard bitwise mutation as the only variation operator) (Li et al. 2016), and for other classical MOEAs that adopt standard bitwise mutation, such as GSEMO (Giel and Lehre 2010), NSGA-II (Zheng, Liu, and Doerr 2022; Zheng and Doerr 2023a), NSGA-II (setting $m=2$ for the mOMM in (Opris et al. 2024)), SMS-EMOA (Zheng and Doerr 2024), and SPEA2 (setting $m=2$ for the mOMM in (Ren et al. 2024)). Therefore, the weighted-sum decomposition demonstrates strong performance and achieves an improvement by a factor of $\Theta(n/\log n)$.

$\Theta(n)$ Runtime Speed-up on OneJumpZeroJump

In the previous section, we have shown that the weighted-sum decomposition achieves at least a factor of $\Theta(n/\log n)$ improvement in optimizing the OMM problem. Here, we further investigate its performance on another popular OneJumpZeroJump (OJZJ for short) problem, to see how the weighted-sum decomposition performs for the multi-modal problem. Since there is no analysis on the MOEA/D with Tchbycheff decomposition either, we will in this section also consider its runtime for a good comparison with the weighted-sum decomposition.

OJZJ and Decomposed Subproblems

The OJZJ benchmark (Doerr and Zheng 2021) is a bi-objective counterpart of the classical multimodal Jump problem with a gap size k (Droste, Jansen, and Wegener 2002), where the first objective is the standard Jump function (w.r.t. the number of ones in the bitstring), and the second one is structurally identical, but w.r.t. the number of zeros in the bitstring (see Definition 7).

Definition 7 ((Doerr and Zheng 2021))**.** *Let $k \in [1..n/2]$. For any $x \in \{0,1\}^n$, the bi-objective problem* OJZJ $f =$

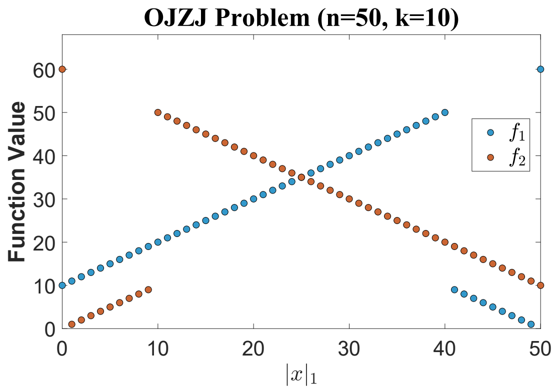

Figure 1: The values of two objectives in OJZJ with respect to $|x|_1$.

$(f_1, f_2) : \{0,1\}^n \to \mathbb{R}^2$ *is defined by*

$$f_1(x) = \begin{cases} k + |x|_1, & \text{if } |x|_1 \le n-k \text{ or } x = 1^n, \\ n - |x|_1, & \text{else,} \end{cases}$$

$$f_2(x) = \begin{cases} k + |x|_0, & \text{if } |x|_0 \le n-k \text{ or } x = 0^n, \\ n - |x|_0, & \text{else.} \end{cases}$$

In the OJZJ problem, the Pareto set is $\{x \mid |x|_1 \in [k..n-k] \cup \{0,n\}\}$, and the Pareto front is $\{(k+j, n+k-j) \mid j \in [k..n-k] \cup \{0,n\}\}$. Figure 1 gives an illustrative plot of this function with problem size $n = 50$ and gap size $k = 10$. We clearly see that the Pareto front contains an inner part and two extreme points of $(k, n+k)$ and $(n+k, k)$. Now we calculate the subproblems and the optimal set of these subproblems under weighted-sum decomposition and Tchebycheff decomposition, respectively. With (1) and (2), we have the following lemmas.

Lemma 8. *Via the weighted-sum decomposition with $H+1$ subproblems, the i-th ($i \in [0..H]$) subproblem for the maximization of* OJZJ *is to maximize*

$$g_i(x) = \begin{cases} (2\frac{i}{H} - 1)|x|_1 + k + (1 - \frac{i}{H})n, \\ \qquad \text{if } |x|_1 \in [k..n-k] \cup \{0,n\}, \\ |x|_1 + \frac{i}{H}k, \\ \qquad \text{if } |x|_1 \in [1..k-1], \\ -|x|_1 + n + (1 - \frac{i}{H})k, \\ \qquad \text{if } |x|_1 \in [n-k+1..n-1]. \end{cases}$$

The optimal set S_i of the i-th subproblem is

$$S_i = \begin{cases} \{x \mid |x|_1 = 0\}, & \text{if } i < \frac{H}{2}, \\ \{x \mid |x|_1 = [k..n-k] \cup \{0,n\}\}, & \text{if } i = \frac{H}{2}, \\ \{x \mid |x|_1 = n\}, & \text{if } i > \frac{H}{2}. \end{cases}$$

Lemma 9. *Let the reference point $z^* = (n+k, n+k)$ and $c_i = \frac{i}{H}n - \lfloor \frac{i}{H}n \rfloor \in [0,1)$ for $i \in [0..H]$. Via the Tchebycheff decomposition with $H+1$ subproblems, the i-th ($i \in [0..H]$) subproblem for the maximization of* OJZJ *is to minimize*

$$g_i(x|z^*) = \begin{cases} \max\{\frac{i}{H}(n - |x|_1), (1 - \frac{i}{H})|x|_1\}, \\ \qquad \text{if } |x|_1 \in [k..n-k] \cup \{0,n\}, \\ \max\{\frac{i}{H}(n - |x|_1), (1 - \frac{i}{H})(n + k - |x|_1)\}, \\ \qquad \text{if } |x|_1 \in [1..k-1], \\ \max\{\frac{i}{H}(k + |x|_1), (1 - \frac{i}{H})|x|_1\}, \\ \qquad \text{if } |x|_1 \in [n-k+1..n-1]. \end{cases}$$

The optimal set S_i of the i-th subproblem is

$$S_i = \begin{cases} \{x \mid |x|_1 = 0\}, & \text{if } 0 \le i < \frac{kH}{n+k}, \\ \{x \mid |x|_1 \in \{0,k\}\}, & \text{if } \frac{kH}{n+k} \in \mathbb{N},\ i = \frac{kH}{n+k}, \\ \{x \mid |x|_1 = k\}, & \text{if } \frac{kH}{n+k} < i < \frac{kH}{n}, \\ \{x \mid |x|_1 = \lfloor \frac{i}{H}n \rfloor\}, & \text{if } \frac{kH}{n} \le i \le \frac{(n-k)H}{n} \\ & \text{and } i < (1 - c_i)H, \\ \{x \mid |x|_1 \in \{\lfloor \frac{i}{H}n \rfloor, \lceil \frac{i}{H}n \rceil\}\}, & \text{if } \frac{kH}{n} \le i \le \frac{(n-k)H}{n} \\ & \text{and } i = (1 - c_i)H, \\ \{x \mid |x|_1 = \lceil \frac{i}{H}n \rceil\}, & \text{if } \frac{kH}{n} \le i \le \frac{(n-k)H}{n} \\ & \text{and } i > (1 - c_i)H, \\ \{x \mid |x|_1 = n-k\}, & \text{if } \frac{(n-k)H}{n} < i < \frac{nH}{n+k}, \\ \{x \mid |x|_1 \in \{n-k, n\}\}, & \text{if } \frac{nH}{n+k} \in \mathbb{N},\ i = \frac{nH}{n+k}, \\ \{x \mid |x|_1 = n\}, & \text{if } \frac{nH}{n+k} < i \le H. \end{cases}$$

Recalling the Pareto front of $\{(k+j, n+k-j) \mid j \in [k..n-k] \cup \{0,n\}\}$, we see that through the weighted-sum decomposition of the OJZJ problem, similar to OMM, some Pareto front points (like $(k+j, n+k-j)$ with $j \in [k..n-k]$) cannot be mapped to the optimum of any subproblem (like when H is odd). However, for the Tchebycheff decomposition with $H = n$, we see that each Pareto front point can be mapped to the optimum of a certain subproblem. Although not all Pareto front points can be mapped in this way for $H < n$, the middle subproblems with $i \in [\frac{kH}{n}, \frac{(n-k)H}{n}]$ have their optima uniformly distributed along the Pareto front. To have an illustrative feeling, Figure 2 shows the example plots between $g_i(x)$ (via the weighted-sum decomposition) and $|x|_1$ for two representative cases $i < \frac{1}{2}H$ and $i > \frac{1}{2}H$, and Figure 3 for the representative case for the Tchebycheff decomposition.

In the following two subsections, we will analyze the runtime of the MOEA/D with the weighted-sum decomposition and the Tchebycheff decomposition (for a wide range of subproblem sizes), respectively. We will also show that despite this similar "drawback" of the subproblem construction as seen for OMM, the weighted-sum decomposition beats the Tchebycheff via a factor of $\Theta(n)$ on OJZJ.

$O(n^k)$ Runtime for Weighted-Sum

In this subsection, we will analyze the number of function evaluations for the MOEA/D using weighted-sum decomposition on OJZJ. We divide the process into two phases. In the first phase, we consider the situation when certain (almost all) subproblems reach one of the desired local optima. Similar to the analysis for OMM, the key idea to establish the runtime is to prove that, with high probability, the entire inner part of the Pareto front is covered by the set P for

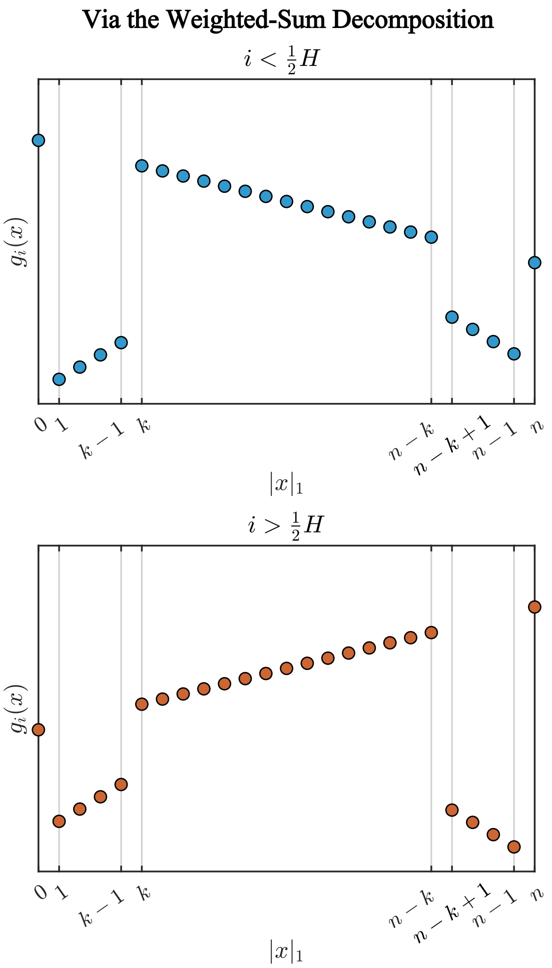

Figure 2: The function values of $g_i(x)$ (via the weighted-sum decomposition) with respect to $|x|_1$ when H is odd (Top: $i < \frac{1}{2}H$; Bottom: $i > \frac{1}{2}H$).

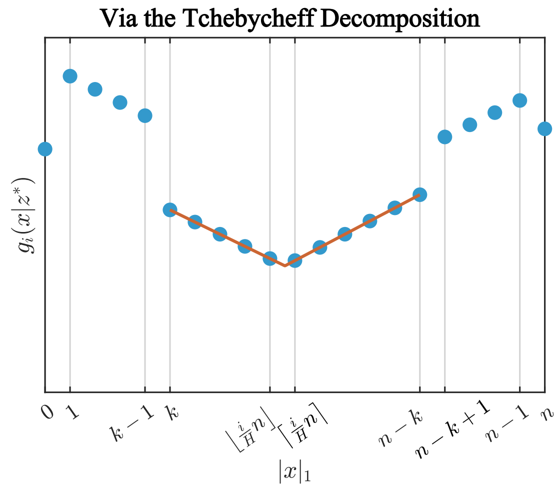

Figure 3: The function values of $g_i(x|z^*)$ (via the Tchebycheff decomposition) with respect to $|x|_1$. The orange curve represents the function $h(w) = \max\{\frac{i}{H}(n-w), (1-\frac{i}{H})w\}$, where $w \in [k, n-k]$.

the first time when the local optima of all these certain subproblems are obtained. Specifically, the Pareto front points $(k+j, n+k-j)$ for all $j \in [k, \frac{1}{2}n]$ are obtained when the first half of subproblems (that is, the i-th subproblem with $i \in [0, \frac{1}{2}H)$) (starting from initial solutions x_i with $|x_i|_1 \in [\frac{1}{2}n, \frac{3}{4}n]$) are successfully solved. Analogously, the Pareto points $(k+j, n+k-j)$ for all $j \in [\frac{1}{2}n, n-k]$ are expected to be discovered in the second half subproblems (that is, the i-th subproblem with $i \in (\frac{1}{2}H, H]$) (starting from initial solutions x_i with $|x_i|_1 \in [\frac{1}{4}n, \frac{1}{2}n]$). From Lemma 3, with high probability, $\Theta(H)$ subproblems start with desired solutions, and conditioned on this, the following lemma shows that if $H+1$ is at least logarithmic, then with probability $1-n^{\Omega(-1)}$, P covers all inner Pareto points $(k+j, n+k-j)$ for $j \in [k..n-k]$ when all certain subproblems reach their desired local optimum.

Lemma 10. *Let $n \geq 8, c > \frac{500}{3}$ and $H \geq c \ln n$. Consider using MOEA/D with weighted-sum decomposition and standard bitwise mutation to maximize* OJZJ. *Let N_0 and N_1 be defined in Lemma 3 and assume $\min\{N_0, N_1\} > \frac{1}{32}H$. Let T_0 denote the first time when $x_0, ..., x_H$ satisfy $x_i \in \{x \mid |x|_1 \in \{n-k, n\}\}$ for $i > \frac{1}{2}H$ and $x_i \in \{x \mid |x|_1 \in \{0, k\}\}$ for $i < \frac{1}{2}H$. Then with probability at least $1 - n^{-(\frac{3}{500}c-1)}$, the inner part of the Pareto front, that is, $\{(k+j, n+k-j) \mid j \in [k..n-k]\}$, is covered by P within T_0 iterations.*

Combining Lemmas 3 and 10, and considering the possible additional runtime to cover two extreme points, we obtain the following theorem for the runtime of the MOEA/D via weighted-sum decomposition on OJZJ.

Theorem 11. *Let $n \geq 8, c > \frac{500}{3}$, $a > 1$ and $c \ln n \leq H \leq n^{a-1} - 1$. Consider using MOEA/D with weighted-sum decomposition and standard bit mutation to maximize* OJZJ. *Conditional on an event that happens with probability at least $1 - 3n^{-(\frac{3}{500}c-1)}$, P covers the whole Pareto front of* OJZJ *within $e(1 + a \ln n)n + \frac{4en^k}{H} + 2$ expected iterations, that is, within $(H+1)\Big(e(1 + a \ln n)n + \frac{4en^k}{H} + 2\Big)$ expected function evaluations.*

$O(n^{k+1})$ Runtime for Tchebycheff

Since there is no runtime of the MOEA/D via the Tchebycheff decomposition on OJZJ, this subsection will establish it, also for the comparison with the weighted-sum decomposition. Note that as discussed before, the analyzed MOEA/D with the Tchebycheff decomposition considers the standard bitwise mutation as the only variation operator, also for the fair comparison with the weighted-sum one. Other mutation strategies and the usage of the crossover will be our interesting future work.

Here we consider a general setting of H, instead of assuming the same size of the subproblem as the Pareto front or as the maximal size of incomparable solutions, like as in (Li et al. 2016). From Lemma 9, we know that the set of optimal solutions obtained by all subproblems may not fully cover the entire Pareto front (like $H < n$). Hence, we need an additional analysis after all subproblems find their optimal solutions. Together with the evolving of the reference points, we divide the whole process into three phases, the first time for the reference point to be $(n+k, n+k)$, the first time when the optima of all subproblems are reached, and the remaining time to cover the remaining Pareto front points.

Firstly, we present the runtime required to reach the reference point $z^* = (n+k, n+k)$.

Lemma 12. *Consider using MOEA/D with the Tchebycheff decomposition and $H+1$ subproblems to maximize* OJZJ. *Then $z^* = (n+k, n+k)$ is reached within $2en\ln n + 2en^k$ expected iterations, that is, within $(H+1)(2en\ln n + 2en^k)$ expected function evaluations.*

It is not difficult to see that once the reference point z^* value of $(n+k, n+k)$ is reached, z^* will stay at this value. The following lemma shows the runtime required to find the optimal solution for each middle subproblem with $i \in [\frac{kH}{n+k}, \frac{nH}{n+k}]$ when $z^* = (n+k, n+k)$.

Lemma 13. *Let $a > 1$, $H \leq n^{a-1} - 1$ and suppose $z^* = (n+k, n+k)$. Consider using MOEA/D with the Tchebycheff decomposition and $H+1$ subproblems to maximize* OJZJ. *For all $i \in [\frac{kH}{n+k}, \frac{nH}{n+k}]$, i-th subproblem reaches its optimal solution within $e(H+1)(n\ln n + 2k^k)$ expected iterations, that is, within $e(H+1)^2(n\ln n + 2k^k)$ expected function evaluations.*

Finally, we analyze the runtime required to cover the remaining Pareto front points of OJZJ.

Lemma 14. *Let $a > 1$, $H \leq n^{a-1} - 1$ and suppose $z^* = (n+k, n+k)$. Let A denote $(\frac{kH}{n+k}, \frac{kH}{n}] \cap \mathbb{N} \neq \emptyset$, $\overline{A}$ for $(\frac{kH}{n+k}, \frac{kH}{n}] \cap \mathbb{N} = \emptyset$, and*

$$b = \begin{cases} \lceil \frac{n}{2H} + \frac{1}{2} \rceil, & \text{if } A \text{ or } (\overline{A}, H \geq n), \\ \lceil \frac{n}{H} \rceil, & \text{if } (\overline{A}, H < n \leq (k+1)H), \\ \lceil \frac{n}{2H} + \frac{k+1}{2} \rceil, & \text{if } (\overline{A}, n > (k+1)H). \end{cases}$$

Consider using MOEA/D with the Tchebycheff decomposition and $H+1$ subproblems to maximize OJZJ. *Assume that in the current iteration, for all $i \in [\frac{kH}{n+k}, \frac{nH}{n+k}] \cup \{0, H\}$, i-th subproblem reaches its optimal solution. Then P covers the remaining Pareto front of* OJZJ *within additional*

$$2en^b \ln b + \frac{4en^b}{\ln(H+1)}$$

expected iterations, that is, within additional

$$e(H+1)\left(2n^b \ln b + \frac{4n^b}{\ln(H+1)}\right)$$

expected function evaluations.

Combining the runtime of the three lemmas above, we obtain the following corollary for the runtime of the MOEA/D via Tchebycheff decomposition on OJZJ.

Corollary 15. *Let $a > 1$, $H \leq n^{a-1} - 1$. Let A denote $(\frac{kH}{n+k}, \frac{kH}{n}] \cap \mathbb{N} \neq \emptyset$, $\overline{A}$ for $(\frac{kH}{n+k}, \frac{kH}{n}] \cap \mathbb{N} = \emptyset$, and*

$$b = \begin{cases} \lceil \frac{n}{2H} + \frac{1}{2} \rceil, & \text{if } A \text{ or } (\overline{A}, H \geq n), \\ \lceil \frac{n}{H} \rceil, & \text{if } (\overline{A}, H < n \leq (k+1)H), \\ \lceil \frac{n}{2H} + \frac{k+1}{2} \rceil, & \text{if } (\overline{A}, n > (k+1)H). \end{cases}$$

Consider using MOEA/D with the Tchebycheff decomposition and $H+1$ subproblems to maximize OJZJ. *Then P covers the whole Pareto front of* OJZJ *within*

$$2en\ln n + 2en^k + e(H+1)(n\ln n + 2k^k) + 2en^b \ln b + \frac{4en^b}{\ln(H+1)}$$

expected iterations, that is, within

$$e(H+1)\left(2n\ln n + 2n^k + (H+1)(n\ln n + 2k^k) + 2n^b \ln b + \frac{4n^b}{\ln(H+1)}\right)$$

expected function evaluations.

From Theorem 11, we know that with the weighted-sum decomposition, the expected runtime (the number of function evaluations) remains $O(n^k)$ for all considerable numbers of the subproblems. For the MOEA/D via Tchebycheff decomposition, from Corollary 15, we know that the expected runtime is $O(Hn^k)$ for $H \geq n$, and roughly $O(H\max\{n^{n/H}\ln(n/H), n^k\})$ (which will become quite large for small value of H) for $H < n$. Thus, for the specific setting of $H = n$ as in (Li et al. 2016), we obtain an expected runtime of $O(n^{k+1})$ for the Tchebycheff decomposition. We also note that this expected runtime of $O(n^{k+1})$ was also proved for other classical MOEAs using only standard bit-wise mutation, such as GSEMO (Zheng and Doerr 2023b), NSGA-II (Doerr and Qu 2023), SMS-EMOA (Bian et al. 2023), NSGA-III (Opris 2025), and SPEA2 (setting $m = 2$ for mOJZJ in (Ren et al. 2024)). Therefore, we see that the MOEA/D via the weighted-sum decomposition beats these algorithms by a factor of $\Theta(n)$ on OJZJ.

Conclusion

Different from the existing theoretical research of the MOEA/D that almost exclusively focused on the Tchebycheff decomposition, this paper presented the first complete runtime analysis for the original weighted-sum decomposition. We proved that the MOEA/D using the weighted-sum decomposition typically (conditional on an event that happens with $1 - n^{-\Omega(1)}$ probability) solves the classic OMM benchmark within $O(n\log^2 n)$ expected function evaluations, which beats other original MOEAs (like the MOEA/D with the Tchebycheff decomposition, GSEMO, NSGA-II, NSGA-III, SMS-EMOA, SPEA2) by a factor of $\Theta(n/\log n)$. The superior runtime of the MOEA/D using the weighted-sum decomposition also holds for the popular OJZJ benchmark, and a speed-up by a factor of $\Theta(n)$ was proved. Our results indicate a further development of the weighted-sum approach might be fruitful.

Acknowledgments

This work was supported by National Natural Science Foundation of China (Grant No. 62306086, 62350710797), Guangdong Basic and Applied Basic Research Foundation (Grant No. 2025A1515011936), and Xinjiang Tianshan Innovative Research Team (2025D14009). This research benefited from the support of the FMJH Program PGMO. This work has profited from many scientific discussions at the Dagstuhl Seminars 23361 "Multiobjective Optimization on a Budget" and 24271 "Theory of Randomized Optimization Heuristics".

References

Alghouass, Y.; Doerr, B.; Krejca, M. S.; and Lagmah, M. 2025. Proven approximation guarantees in multi-objective optimization: SPEA2 beats NSGA-II. In *International Joint Conference on Artificial Intelligence, IJCAI 2025*, 8833–8841. ijcai.org.

Bian, C.; Zhou, Y.; Li, M.; and Qian, C. 2023. Stochastic population update can provably be helpful in multi-objective evolutionary algorithms. In *International Joint Conference on Artificial Intelligence, IJCAI 2023*, 5513–5521. ijcai.org.

Corus, D.; Oliveto, P. S.; and Yazdani, D. 2025. Fast Contiguous Somatic Hypermutations for Single-Objective Optimisation and Multi-Objective Optimisation Via Decomposition. In *Conference on Artificial Intelligence, AAAI 2025*, volume 39, 26922–26930.

Do, A. V.; Neumann, A.; Neumann, F.; and Sutton, A. 2023. Rigorous runtime analysis of MOEA/D for solving multi-objective minimum weight base problems. *Advances in Neural Information Processing Systems*, 36: 36434–36448.

Doerr, B.; Krejca, M. S.; and Weeks, N. 2024. Proven runtime guarantees for how the MOEA/D computes the pareto front from the subproblem solutions. In *International Conference on Parallel Problem Solving from Nature, PPSN 2024*, 197–212. Springer.

Doerr, B.; and Qu, Z. 2023. A first runtime analysis of the NSGA-II on a multimodal problem. *IEEE Transactions on Evolutionary Computation*, 27: 1288–1297.

Doerr, B.; and Zheng, W. 2021. Theoretical analyses of multi-objective evolutionary algorithms on multi-modal objectives. In *Conference on Artificial Intelligence, AAAI 2021*, 12293–12301. AAAI Press.

Droste, S.; Jansen, T.; and Wegener, I. 2002. On the analysis of the (1+1) evolutionary algorithm. *Theoretical Computer Science*, 276: 51–81.

Giel, O.; and Lehre, P. K. 2010. On the effect of populations in evolutionary multi-objective optimisation. *Evolutionary Computation*, 18: 335–356.

Huang, Z.; and Zhou, Y. 2020. Runtime analysis of somatic contiguous hypermutation operators in MOEA/D framework. In *Conference on Artificial Intelligence, AAAI 2020*, 2359–2366. AAAI Press.

Huang, Z.; Zhou, Y.; Chen, Z.; and He, X. 2019. Running time analysis of MOEA/D with crossover on discrete optimization problem. In *Conference on Artificial Intelligence, AAAI 2019*, 2296–2303. AAAI Press.

Huang, Z.; Zhou, Y.; Luo, C.; and Lin, Q. 2021. A runtime analysis of typical decomposition approaches in MOEA/D framework for many-objective optimization problems. In *International Joint Conference on Artificial Intelligence, IJCAI 2021*, 1682–1688.

Laumanns, M.; Thiele, L.; and Zitzler, E. 2004. Running time analysis of multiobjective evolutionary algorithms on pseudo-Boolean functions. *IEEE Transactions on Evolutionary Computation*, 8: 170–182.

Laumanns, M.; Thiele, L.; Zitzler, E.; Welzl, E.; and Deb, K. 2002. Running time analysis of multi-objective evolutionary algorithms on a simple discrete optimization problem. In *Parallel Problem Solving from Nature, PPSN 2002*, 44–53. Springer.

Li, Y.-L.; Zhou, Y.-R.; Zhan, Z.-H.; and Zhang, J. 2016. A primary theoretical study on decomposition-based multiobjective evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 20: 563–576.

Opris, A. 2025. A first runtime analysis of NSGA-III on a many-objective multimodal problem: provable exponential speedup via stochastic population update. In *International Joint Conference on Artificial Intelligence, IJCAI 2025*, 8903–8911. ijcai.org.

Opris, A.; Dang, D.-C.; Neumann, F.; and Sudholt, D. 2024. Runtime analyses of NSGA-III on many-objective problems. In *Proceedings of the Genetic and Evolutionary Computation Conference*, 1596–1604.

Ren, S.; Bian, C.; Li, M.; and Qian, C. 2024. A first running time analysis of the Strength Pareto Evolutionary Algorithm 2 (SPEA2). In *Parallel Problem Solving from Nature, PPSN 2024*, 295–312. Springer.

Wietheger, S.; and Doerr, B. 2023. A mathematical runtime analysis of the Non-dominated Sorting Genetic Algorithm III (NSGA-III). In *International Joint Conference on Artificial Intelligence, IJCAI 2023*, 5657–5665. ijcai.org.

Zhang, Q.; and Li, H. 2007. MOEA/D: A multiobjective evolutionary algorithm based on decomposition. *IEEE Transactions on Evolutionary Computation*, 11: 712–731.

Zheng, W. 2024. Theoretical advantage of multiobjective evolutionary algorithms for problems with different degrees of conflict. *CoRR*, abs/2408.04207.

Zheng, W.; and Doerr, B. 2023a. Mathematical runtime analysis for the non-dominated sorting genetic algorithm II (NSGA-II). *Artificial Intelligence*, 325: 104016.

Zheng, W.; and Doerr, B. 2023b. Theoretical analyses of multiobjective evolutionary algorithms on multimodal objectives. *Evolutionary Computation*, 31: 337–373.

Zheng, W.; and Doerr, B. 2024. Runtime analysis of the SMS-EMOA for many-objective optimization. In *Conference on Artificial Intelligence, AAAI 2024*, 20874–20882. AAAI Press.

Zheng, W.; Liu, Y.; and Doerr, B. 2022. A first mathematical runtime analysis of the Non-Dominated Sorting Genetic Algorithm II (NSGA-II). In *Conference on Artificial Intelligence, AAAI 2022*, 10408–10416. AAAI Press.