2023 IEEE Symposium Series on Computational Intelligence (SSCI)
Mexico City, Mexico. December 5-8, 2023

Runtime Analysis of (1+1)-EA on a Biobjective Test Function in Unbounded Integer Search Space

Günter Rudolph
Department of Computer Science
TU Dortmund University
Dortmund, Germany
Guenter.Rudolph@tu-dortmund.de

***Abstract*—Runtime results of multiobjective evolutionary algorithms in unbounded integer spaces are scarce at present. In order to advance this research field we consider two versions of the (1+1)-EA and analyze their runtime to the Pareto front of a carefully designed biobjective test problem.**

***Index Terms*—runtime analysis, multiobjective evolutionary algorithm, unbounded integer space**

I. Introduction

The theory of multiobjective evolutionary algorithms (MOEA) on unbounded integer space can be regarded as a *terra incognita*, i.e., almost nothing is known about this case at present. There are theoretical results for EAs in integer space (e.g. [1], [2]), but the integer space is mostly bounded and only the singleobjective case is considered. Typically, the first step in the theoretical exploration of a new field is the runtime analysis of a version of the (1+1)-EA for a carefully designed test problem. This path is also followed here.

After introducing the mathematical background and terminology in section II we design an integer test problem with two objectives in section III which may be seen as an analogon to the bi-sphere problem in continuous space which has been amenable to an analysis in the past [5]. After deriving Pareto set and front for this problem we define a biobjective (1+1)-EA that rejects only dominated solutions, and a singleobjective (1+1)-EA that works on the scalarized objective function in the sense of the reference point method [6] in section IV. Both algorithms and their runtime for arbitrary starting points are theoretically analyzed in 2-dimensional decision space in section V. Section VI summarizes the results.

II. Background and Terminology

In the following we consider unconstrained multiobjective optimization problems (MOPs) of the form $\min\{f(x) : x \in \mathbb{Z}^n\}$ where $f(x) = (f_1(x), \ldots, f_d(x))'$ is a vector-valued mapping with $d \geq 2$ objective functions $f_i : \mathbb{Z}^n \to \mathbb{R}$ for $i = 1, \ldots, d$ that are to be minimized simultaneously. The optimality of a MOP is defined by the concept of *dominance*.

Let $u, v \in F \subseteq \mathbb{R}^d$ where F is equipped with the partial order $\preceq$ defined by $u \preceq v \Leftrightarrow \forall i = 1, \ldots d : u_i \leq v_i$. If $u \prec v \Leftrightarrow u \preceq v \wedge u \neq v$ then v is said to be *dominated by* u. An element u is termed *nondominated* relative to $V \subseteq F$ if there is no $v \in V$ that dominates u. The set $\mathsf{ND}(V, \preceq) = \{u \in V \mid \nexists\, v \in V : v \prec u\}$ is called the *nondominated set* relative to V.

If $F = f(X)$ is the objective space of some MOP with decision space $X \subseteq \mathbb{Z}^n$ and objective function $f(\cdot)$ then the set $F^* = \mathsf{ND}(f(X), \preceq)$ is called the *Pareto front* (PF). Elements $x \in X$ with $f(x) \in F^*$ are termed *Pareto-optimal* and the set X^* of all Pareto-optimal points is called the *Pareto set* (PS).

If we are not interested in finding an approximation of the entire PF a reference point method [6] can be used to find a solution that is closest to a so-called reference point gathering the user-given level of aspiration for each objective.

III. Test Problem

Consider the biobjective optimization problem $f : \mathbb{Z}^n \to \mathbb{N}_0^2$ with

$$f(x) = \begin{pmatrix} \|x - c\|_1 \\ \|x + c\|_1 \end{pmatrix} \to \min! \quad \text{and} \quad c \in \mathbb{Z}^n \setminus \{0\} \quad (1)$$

where $\|\cdot\|_1$ denotes the ℓ_1-norm. Setting $c = (c_1, 0, \ldots, 0)^\intercal \neq 0 \in \mathbb{Z}^n$ we obtain the special case $f : \mathbb{Z}^n \to \mathbb{N}_0^2$ with

$$f(x) = \begin{pmatrix} |x_1 - c_1| + |x_2| + \ldots |x_n| \\ |x_1 + c_1| + |x_2| + \ldots |x_n| \end{pmatrix} \to \min! \quad (2)$$

Furthermore, we insist on $c_1 =: a \in \mathbb{N}$ and $n = 2$, so that problem (1) finally specializes to $f : \mathbb{Z}^2 \to \mathbb{N}_0^2$ with

$$f(x) = \begin{pmatrix} |x_1 - a| + |x_2| \\ |x_1 + a| + |x_2| \end{pmatrix} \to \min! \quad \text{and} \quad a \in \mathbb{N}. \quad (3)$$

Theorem III.1
The Pareto set of problem (2) is

$$X^* = \{x \in \mathbb{Z}^n : x = (-|c_1| + k, 0, \ldots, 0)^\intercal\} \quad (4)$$

with Pareto front $F^* = f(X^*) =$

$$\{(v_1, v_2)^\intercal \in \mathbb{Z}^2 : (v_1, v_2)^\intercal = (k,\, 2 \cdot |c_1| - k)^\intercal\} \quad (5)$$

where $k = 0, 1, \ldots, 2 \cdot |c_1|$ in (4) and (5).

Proof
Suppose the Pareto set X^* is as given in (4). Insertion of all $x^* \in X^*$ in the objective function (2) reveals that the Pareto front F^* is given by (5).

It remains to prove the validity of (4). First, we show that X^* is an antichain (i.e., all distinct elements are mutually

incomparable). Second, we show that the image of each $y \in \mathbb{Z}^n \setminus X^*$ is dominated by an element of $f(X^*)$. These facts then lead to the conclusion that X^* is complete and nondominated. Hereinafter, we assume w.l.o.g. that $c_1 =: a > 0$.

Let $x, y \in X^*$ with $x \neq y$ so that necessarily $x_i = y_i = 0$ for $i = 2, \ldots, n$. Two cases must be considered.

1. If $x_1 < y_1$ then $f_1(x) = |x_1 - a| > |y_1 - a| = f_1(y)$ and $f_2(x) = |x_1 + a| < |y_1 + a| = f_2(y)$.
2. If $x_1 > y_1$ then $f_1(x) = |x_1 - a| < |y_1 - a| = f_1(y)$ and $f_2(x) = |x_1 + a| > |y_1 + a| = f_2(y)$.

Thus, we have shown that $f(x) \parallel f(y)$ for each pair of distinct elements of X^*. This proves that X^* is an antichain.

Let $y \in \mathbb{Z}^n \setminus X^*$ and $x^* \in X^*$. We partition $\mathbb{Z}^n \setminus X^*$ into three sets:

1. $|y_1| > a$ and $\forall i = 2, \ldots, n : y_i = 0$.
   If $y_1 > a$ then $f_1(y) = |y_1 - a| = y_1 - a > 0$ and $f_2(y) = |y_1 + a| = y_1 + a > 2\,a$.
   Those points are dominated by $f(x^*) = f((a, 0, \ldots, 0)^\intercal) = (0, 2\,a)^\intercal \prec f(y)$.
   If $y_1 < -a$ then $f_1(y) = |y_1 - a| = a - y_1 > 2\,a$ and $f_2(y) = |y_1 + a| = -y_1 - a > 0$.
   Those points are dominated by $f(x^*) = f((-a, 0, \ldots, 0)^\intercal) = (2\,a, 0)^\intercal \prec f(y)$.
2. $|y_1| > a$ and $\exists i = 2, \ldots, n : |y_i| > 0$.
   If $y_1 > a$ then $f_1(y) = \underbrace{|y_1 - a|}_{>0} + \underbrace{|y_2| + \ldots + |y_n|}_{>0} > 0$ and $f_2(y) = \underbrace{|y_1 + a|}_{>2\,a} + \underbrace{|y_2| + \ldots + |y_n|}_{>0} > 2\,a$.
   Those points are dominated by $f(x^*) = f((a, 0, \ldots, 0)^\intercal) = (0, 2\,a)^\intercal \prec f(y)$.
   If $y_1 < -a$ then $f_1(y) = \underbrace{|y_1 - a|}_{>2\,a} + \underbrace{|y_2| + \ldots + |y_n|}_{>0} > 2\,a$ and $f_2(y) = \underbrace{|y_1 + a|}_{>0} + \underbrace{|y_2| + \ldots + |y_n|}_{>0} > 0$.
   Those points are dominated by $f(x^*) = f((-a, 0, \ldots, 0)^\intercal) = (2\,a, 0)^\intercal \prec f(y)$.
3. $|y_1| \leq a$ and $\exists i = 2, \ldots, n : |y_i| > 0$.
   Note that in this case $(y_1, 0, \ldots, 0)^\intercal \in X^*$. As a consequence,
   $f_1(y) = |y_1 - a| + \underbrace{|y_2| + \ldots + |y_n|}_{>0} > |y_1 - a| = f_1((y_1, 0, \ldots, 0)^\intercal)$ and
   $f_2(y) = |y_1 + a| + \underbrace{|y_2| + \ldots + |y_n|}_{>0} > |y_1 + a| = f_2((y_1, 0, \ldots, 0)^\intercal)$.

Thus, those points are dominated by $f(x^*) = f((y_1, 0, \ldots, 0)^\intercal) \prec f(y)$. □

IV. Algorithms

A. (1+1)-Pareto-EA

This dominance-based EA accepts the offspring if it is not dominated by its parent. Thus, an offspring being incomparable to its parent is also accepted. The offspring is generated by a local mutation operation: draw any dimension k uniformly at random and increment or decrement x_k with equal probability. Effectively, we add a mutation vector with a specific distribution to the parent.

The notation $U_n(k, V; v)$ with $n \in \mathbb{N}$, $k \in \{1, \ldots, n\}$, $V \subset \mathbb{Z}$ with $|V| < \infty$, and $v \in \mathbb{Z}$ denotes a discrete n-dimensional distribution, whose realizations are generated by drawing k distinct indices uniformly distributed from $\{1, \ldots, n\}$, before drawing the values for these k index positions independently and uniformly distributed from V; the remaining values are set to v. Thus, mutation vector $Z \sim U_n(1, \{-1, +1\}; 0)$ is a zero vector where a single uniformly distributed position in the vector is set either to -1 or to $+1$.

Algorithm 1 $(1 + 1)$-Pareto-EA

```
choose X^(0) ∈ Z^n at random; set t = 0
repeat
    draw Z ~ U_n(1, {−1, +1}; 0)
    Y = X^(t) + Z
    if f(X^(t)) ≺ f(Y) then
        X^(t+1) = X^(t)
    else
        X^(t+1) = Y
    end if
    t = t + 1
until stopping criterion fulfilled
```

B. (1+1)-RefPoint-EA

The reference point method is related to the goal programming approach and they share the psychologically appealing idea that the decision maker should set a reference point (or goal) in objective space and then some algorithm tries to come close to it. Coming close to a goal is typically realized by minimizing a distance measure between an attainable objective vector and the goal vector.

Algorithm 2 $(1 + 1)$-RefPoint-EA

```
choose X^(0) ∈ Z^n at random; set t = 0
choose a reference point r ∈ Z^d
repeat
    draw Z ~ U_n(1, {−1, +1}; 0)
    Y = X^(t) + Z
    if f̃_r(X^(t)) < f̃_r(Y) then
        X^(t+1) = X^(t)
    else
        X^(t+1) = Y
    end if
    t = t + 1
until stopping criterion fulfilled
```

Here, the scalarized objective function to be minimized is given by

$$\tilde{f}_r(x) = \rho(f(x), r)$$

where $r \in \mathbb{Z}^d$ is the reference point and $\rho : \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{R}_0^+$ is any metric on d-dimensional integers.

The mutation is identical to that of the dominance-based EA, but the selection method now accepts every offspring that is not more distant to the reference point than its parent.

V. Theory

A. (1+1)-Pareto-ES

1) Local behavior: Suppose $x \notin X^*$. We like to know how the objective function values alter for each possible mutation. Here, we consider the special case $n = 2$. In this case four different mutations are possible:

1. $Z = (0, +1)^\intercal$

$$f(x+Z) = \begin{cases} f(x) + (+1,+1)^\intercal & \text{if } x_2 \geq 0 \\ f(x) + (-1,-1)^\intercal & \text{if } x_2 < 0 \end{cases}$$

2. $Z = (0, -1)^\intercal$

$$f(x+Z) = \begin{cases} f(x) + (-1,-1)^\intercal & \text{if } x_2 > 0 \\ f(x) + (+1,+1)^\intercal & \text{if } x_2 \leq 0 \end{cases}$$

3. $Z = (+1, 0)^\intercal$

$$f(x+Z) = \begin{cases} f(x) + (+1,+1)^\intercal & \text{if } x_1 \geq a \\ f(x) + (-1,+1)^\intercal & \text{if } -a \leq x_1 < a \\ f(x) + (-1,-1)^\intercal & \text{if } x_1 < -a \end{cases}$$

4. $Z = (-1, 0)^\intercal$

$$f(x+Z) = \begin{cases} f(x) + (-1,-1)^\intercal & \text{if } x_1 > a \\ f(x) + (+1,-1)^\intercal & \text{if } -a < x_1 \leq a \\ f(x) + (+1,+1)^\intercal & \text{if } x_1 \leq -a \end{cases}$$

Due to symmetries in the problem (see fig. 1) it suffices to consider starting points in the region $\{x \in \mathbb{Z}^2 : x_1 > a \wedge x_2 > 0\}$ to analyze the behavior also for other regions.

$x_1 > +a$	$\Leftrightarrow$	$x_1 > +a$
$x_2 > 0$		$x_2 < 0$
$\Updownarrow$		$\Updownarrow$
$x_1 < -a$		$x_1 < -a$
$x_2 > 0$	$\Leftrightarrow$	$x_2 < 0$

Fig. 1. Dominance symmetries in the test problem if $n = 2$.

While $x_1 > a$ and $x_2 > 0$ the EA makes successful mutations with probability $1/2$, because it may go either to the left with probability $1/4$ or down with probability $1/4$. In this case both fitness functions will be decreased by 1. Otherwise the EA stays at the current position.

As soon as either $x_1 = a$ or $x_2 = 0$ the behavior changes.

If $x_1 = a$ only moves down are successful which happens with probability $1/4$. It is also possible to move to the left (with probability $1/4$), but this leads to an incomparable solution which is accepted. If this has happened ($x_1 < a$) the EA may go to left or right to incomparable solutions but it will never leave $\{-a \leq x_2 \leq a\}$. Repeated downward mutations will finally end at the Pareto set with $x_2 = 0$.

If $x_2 = 0$ only moves to the left are successful which happend with probability $1/4$. Other mutations lead to dominated solutions which are rejected. Repeated leftward mutations will finally end at the Pareto set with $x_1 = a$.

2) Runtime: Suppose the EA is started in $x \in \mathbb{Z}^2$ with $x_1 > a$ and $x_2 > 0$. We first distinguish how many FEs are necessary to reach a positions with either $x_1 = a$ or $x_2 = 0$. Next, we consider then runtime to the Pareto set once either of both positions is reached.

1. Arrival at $(k, 0)^\intercal$ with $a < k \leq x_1$
   This is only possible via $(k, 1)^\intercal$. Regardless which path has been realized to end up in $(k, 1)^\intercal$, the EA must have taken $x_2 - 1$ downward and $x_1 - k$ leftward steps. On average the EA leaves its current position in the 2nd trial. Therefore, $2\,(x_1 + x_2 - k - 1)$ trials are required on average to reach $(k, 1)^\intercal$. Finally, it must do the downward move which probability $1/2$ which adds 2 trials on average.
2. Arrival at $(a, k)^\intercal$ with $k > 0$
   This is only possible via $(a + 1, k)^\intercal$. Again, regardless of the realized path, the EA must have taken $x_1 - (a + 1)$ leftward and $x_2 - k$ downward steps. As the EA leaves its current position by a successful mutation in the 2nd trial on average as well, now $2\,(x_1 + x_2 - k - a - 1)$ trials are required on average to reach $(a + 1, k)^\intercal$. Finally, it must do the leftward move which probability $1/2$ which adds 2 trials on average.
3. Path from $(k, 0)^\intercal$
   Now only leftward mutations are possible which happen with probability $1/4$. A successful move therefore happens after 4 trials on average and the EA needs $k - a$ of them to reach the Pareto set at $(a, 0)^\intercal$. Therefore, $4\,(k - a)$ trials are necessary on average.
4. Path from $(a, k)^\intercal$
   Now only downward mutations are successful which happen with probability $1/4$. Moves to the left and right are accepted as long as the constraint $|x_1| \leq a$ is valid. Repeated left- and rightward moves lead to a random walk (with reflecting barriers) of variable x_1 and they do not change the runtime (because they count effectively like unsuccessful moves without progress). Since the EA needs k downward steps to reach the Pareto set, $4\,k$ trials are necessary on average.

So far we have derived

$$\mathsf{E}[\,T \,|\, \text{path through } (k,0)^\intercal\,] = 2\,(x_1 + x_2 - k - 1) + 4\,(k - a) + 2 = 2\,(x_1 + x_2 + k) - 4\,a \text{ and} \quad (6)$$

$$\mathsf{E}[\,T \,|\, \text{path through } (a,k)^\intercal\,] = 2\,(x_1 + x_2 - k - a - 1) + 4\,k + 2 = 2\,(x_1 + x_2 + k) - 2\,a \quad (7)$$

It remains to specify with which probability either of the both cases occur. To that end we regard both points as absorbing

states of a Markov chain. Please notice that the disregard of the waiting times on transient states (i.e., positions that are left after an accepted mutation) do not affect the absorption probabilities of a Markov chain if we reformulate the transition probabilities for a modified, new process:

$$\mathsf{P}\{\text{ move downward in new process }\} = \frac{\mathsf{P}\{\text{ move downward in original process }\}}{\mathsf{P}\{\text{ stay at position in original process }\}} = \frac{1}{4} : \frac{1}{2} = \frac{1}{2}$$

and analogous for the leftward move. Therefore we obtain:

1. Absorption to $(k,0)^\intercal$
   We know that the EA must do $x_2 - 1$ downward and $x_1 - k$ leftward steps each with (new) probability $1/2$ and finally a downward move with probability $1/2$. This happens with probability
   $$\mathsf{P}\{\,(k,0)^\intercal\,\} = \frac{1}{2}\binom{x_1 + x_2 - k - 1}{x_1 - k} 2^{-(x_1+x_2-k-1)} \quad (8)$$
   where the binomial term describes the number of possible paths.
2. Absorption to $(a,k)^\intercal$
   We know that the EA must do $x_2 - k$ downward and $x_1 - (a+1)$ leftward steps each with (new) probability $1/2$ and finally a leftward move with probability $1/2$. This happens with probability
   $$\mathsf{P}\{\,(a,k)^\intercal\,\} = \frac{1}{2}\binom{x_1 + x_2 - k - a - 1}{x_2 - k} 2^{-(x_1+x_2-k-a-1)} \quad (9)$$
   where the binomial term describes the number of possible paths.

Finally, we have to sum over all possible values of k

$$\mathsf{E}[\,T\,] = \sum_{k=a+1}^{x_1} \mathsf{E}[\,T \,|\, \text{path through } (k,0)^\intercal\,] \cdot \mathsf{P}\{\,(k,0)^\intercal\,\} + \sum_{k=1}^{x_2} \mathsf{E}[\,T \,|\, \text{path through } (a,k)^\intercal\,] \cdot \mathsf{P}\{\,(a,k)^\intercal\,\}.$$

inserting the expressions given in (6) - (9). Table I compares our findings with empirical results (100 runs) for some starting points for the test problem with $a = 50$. Figure 2 provides a graphical illustration and an impression about the runtime variability.

TABLE I
COMPARISION OF EXACT, EMPIRICAL AND APPROXIMATED EXPECTED RUNTIMES FOR DIFFERENT STARTING POINTS $X^{(0)}$ IN CASE OF $a = 50$. EMPIRICAL RESULTS ARE BASED ON 100 RUNS.

experiment	$X^{(0)}$	$\mathsf{E}[\,T\,]$	median	mean	meanfield
E1	$(0,1000)^\intercal$	4000	3994	4001	4000
E2	$(100,900)^\intercal$	3600	3577	3591	3600
E3	$(300,700)^\intercal$	2800	2807	2802	2800
E4	$(500,500)^\intercal$	2003	2003	2011	2000
E5	$(700,300)^\intercal$	2600	2589	2586	2600
E6	$(900,100)^\intercal$	3400	3400	3405	3400
E7	$(1000,0)^\intercal$	3800	3795	3805	3800

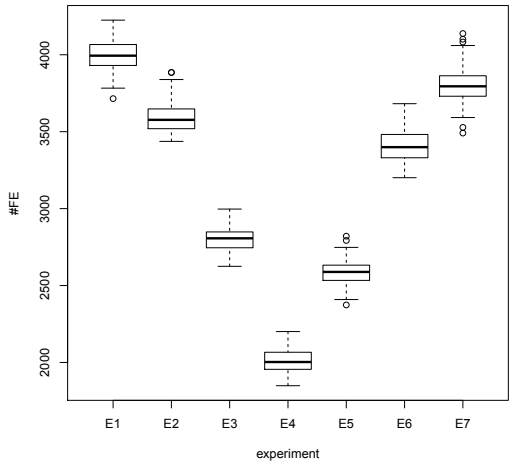

Fig. 2. Number of function evaluations until Pareto front reached.

Taking a meanfield point of view, i.e., we average over the stochastic fluctuations and follow the "expected path", the runtime can be estimated as follows: When starting from $x_1 > a$ and $x_2 > 0$ the EA either goes a step down or a step to the left in the second trial.

1. $x_1 - a \geq x_2$:
   After 4 trials it has made a move down and to the left. This is repeated until it hits $(x_1 - x_2, 0)^\intercal$ after $4\,x_2$ trials. Then the EA can only move to the left every 4th trial, so that it ends in $(0,a)^\intercal$ after $4\,(x_1 - x_2 - a)$ trials. In total it needs $4\,x_2 + 4\,(x_1 - x_2 - a) = 4\,(x_1 - a)$ trials on average.
2. $x_1 - a < x_2$:
   After 4 trials it has made a move down and to the left. This is repeated until it hits $(a, x_2 - (x_1 - a), 0)^\intercal$ after $4\,(x_1 - a)$ trials. Then the EA can only move downwards every 4th trial, so that it ends in $(0,a)^\intercal$ after $4\,(x_2 - x_1 + a)$ trials. In total it needs $4\,(x_1 - a) + 4\,(x_2 - x_1 + a) = 4\,x_2$ trials on average.

The above expressions for the trials on average have been used to calculate the values given in the rightmost column of table I. The agreement with the exact values for $\mathsf{E}[\,T\,]$ and the experimental results in table I is intriguing. The expressions lead to the conjecture that the expected runtime is roughly $4\,\|X^{(0)} - c\|_\infty$ or $O(\|X^{(0)}\|_\infty)$, where $c = (a,0)^\intercal$.

B. (1+1)-RefPoint-EA

Let the reference point $r = (a,a)^\intercal$ be located on the Pareto front with $f^{-1}(r) = (0,\ldots,0)^\intercal$. In principle, every metric or norm may be used to express the distance of a solution to the reference point. In integer decision and objective space the ℓ_1-norm seems most natural. In this case, the scalarized objective function is $\tilde{f}_r(x) = \|f(x) - r\|_1$. If $n = 2$ then

$$\begin{aligned}\tilde{f}_r(x) &= \|f(x) - r\|_1 = |f_1(x) - a| + |f_2(x) - a| \\ &= \Big||x_1 - a| + |x_2| - a\Big| + \Big||x_1 + a| + |x_2| - a\Big|. \quad (10)\end{aligned}$$

For the runtime analysis one needs to know which mutations are improving, neutral or worsening. This information is available from the level sets (= elements in decision space with equal fitness value). The absolute values require a case-by-case analysis.

1. Let $x_1 \geq a$ and $x_2 \geq 0$. Then (10) simplifies to
$$\begin{aligned}\tilde{f}_r(x) &= |(x_1 - a) + x_2 - a| + |(x_1 + a) + x_2 + a| \\ &= |x_1 + x_2 - 2\,a| + x_1 + x_2. \end{aligned} \tag{11}$$
If additionally $x_1 + x_2 \geq 2\,a$ equation (11) reduces to
$$\tilde{f}_r(x) = x_1 + x_2 - 2\,a + x_1 + x_2 = 2\,(\,[x_1 - a] + x_2\,)$$
and otherwise to
$$\tilde{f}_r(x) = 2\,a - x_1 - x_2 + x_1 + x_2 = 2\,a.$$
2. Let $x_1 \leq -a$ and $x_2 \geq 0$. Due to symmetry we obtain (analogous to case 1 above) that (10) finally simplifies to
$$\tilde{f}_r(x) = \begin{cases} 2\,(\,[-x_1 - a] + x_2\,) & \text{if } x_2 \geq x_1 + 2\,a \\ 2\,a & \text{otherwise.} \end{cases}$$
3. Let $-a < x_1 < a$ and $x_2 \geq 0$. In this case equation (10) simplifies to
$$\begin{aligned}\tilde{f}_r(x) &= \Big||x_1 - a| + x_2 - a|\Big| + |x_1 + x_2| \\ &= |(a - x_1) + x_2 - a| + |x_1 + x_2| \\ &= |x_2 - x_1| + |x_1 + x_2|. \end{aligned} \tag{12}$$
If additionally $x_1 \geq 0$ equation (12) reduces to
$$\tilde{f}_r(x) = |x_2 - x_1| + x_1 + x_2 = \begin{cases} 2\,x_2 & \text{if } x_2 \geq x_1 \\ 2\,x_1 & \text{if } x_2 < x_1. \end{cases}$$
The case with additional constraint $x_1 < 0$ leads to
$$\tilde{f}_r(x) = x_2 - x_1 + |x_1 + x_2| = \begin{cases} 2\,x_2 & \text{if } x_1 + x_2 \geq 0 \\ -2\,x_1 & \text{if } x_1 + x_2 < 0. \end{cases}$$
4. Due to symmetries of the problem one obtains analogue expressions for the preceding cases if $x_2 < 0$.

This analysis has shown that there exists plateaus in the fitness landscape with $f(x) = 2\,a$. Although the size of the plateaus is finite this fact may slow down the algorithm's approach to the optimum as it behaves like a random walk on a lattice until it reaches the border of the plateau with better fitness values. If there are worse fitness values at the rim of the plateau the random walk is reflected by this barrier. Figure 3 illustrates the resulting fitness landscape for $r = (a, a)^\intercal$ with $a = 4$.

From a theoretician's point of view the situation becomes even worse, since the size and values of the plateaus vary depending on the location of the reference point—even if it is placed on the Pareto front.

If the reference point is placed in the utopian region 'behind' the Pareto front, then there are still plateaus and, additionally, several points have the same minimal attainable distance to the reference point. For example, setting $r = (2, 2)^\intercal$ in case $a = 4$,

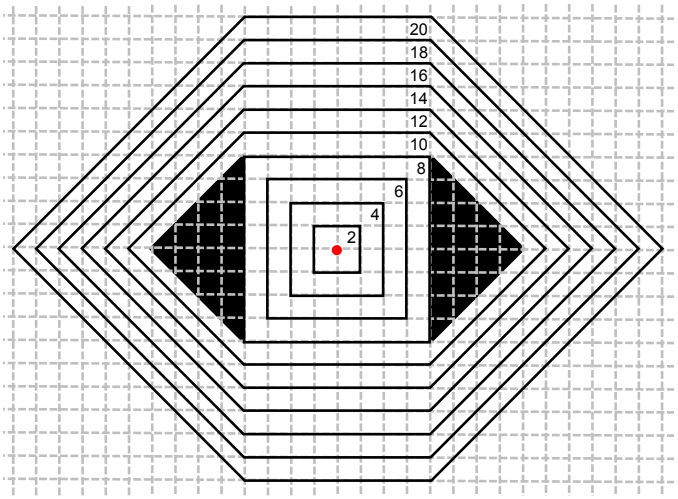

Fig. 3. Kind of contour plot of the scalarized objective function with reference point $r = (a, a)^\intercal = (4, 4)^\intercal$ in $n = 2$. Lattice points lying on the a connected solid line or on a black area have the same objective function value. Black areas indicate plateaus. Numbers next to the contour lines refer to the objective function value. The red dot in the center marks the position of the optimum at $x^* = (0, 0)^\intercal$ with $\tilde{f}(x^*) = 0$.

the Pareto-optimal solutions $(6, 2)^\intercal, (5, 3)^\intercal, (4, 4)^\intercal, (3, 5)^\intercal$ and $(2, 6)^\intercal$ have the same ℓ_1-distance to r. Even worse, if $r = (0, 0)^\intercal$ then the entire Pareto front has the same distance to the reference point.

Another extreme case is caused when using the discrete metric with $\rho(x, y) = 0$ if $x = y$, and $\rho(x, y) = 1$ if $x \neq y$. In this case only the reference point has distance zero whereas all other points are on a plateau with value 1. Evidently, a useful distance measure must provide some guidance to better solutions.

Recommendation:
The metric used for expressing 'closeness' to a solution in objective space should be chosen with prudent consideration. The ostensibly obvious ℓ_1-metric in integer objective spaces can lead to optimality results that are quite counter-intuitive to the DM's perception. In addition, this metric can cause the emergence of plateaus which may hamper the algorithm's approach of the optimum and the theoretician's efforts of deriving runtime bounds. □

Most likely, the concept of 'closeness' for humans is more naturally represented by the Euclidean ℓ_2-norm. As table II reveals there are no plateaus any longer. After this finding, we restrict the analysis to distances in ℓ_2-norm. Using the squared norm does not change the problem in ordinal sense and the objective function values remain in $\mathbb{N}_0$. Thus,
$$f_r(x) = \|f(x) - r\|_2^2 = (f_1(x) - r_1)^2 + (f_2(x) - r_2)^2$$
in general, and with $r = (a, a)^\intercal$ and $a = 4$ in the following. With this choice we can revert to table II as a starting point for our runtime analysis.

As can be seen from table II the problem is symmetric to x_1- and x_2-axis. Since the EA only makes steps either in x_1-direction or in x_2-direction and cannot cross either of both

TABLE II
OBJECTIVE FUNCTION VALUES $\tilde{f}_r(x)$ USING ℓ_2-DISTANCE WITH $r = (a,a)^\mathsf{T}$ AND $a = 4$ FOR $x \in [-10,10]^2 \cap \mathbb{Z}^2$.

10	544	482	424	370	320	274	232	218	208	202	200	202	208	218	232	274	320	370	424	482	544
9	482	424	370	320	274	232	194	180	170	164	162	164	170	180	194	232	274	320	370	424	482
8	424	370	320	274	232	194	160	146	136	130	128	130	136	146	160	194	232	274	320	370	424
7	370	320	274	232	194	160	130	116	106	100	98	100	106	116	130	160	194	232	274	320	370
6	320	274	232	194	160	130	104	90	80	74	72	74	80	90	104	130	160	194	232	274	320
5	274	232	194	160	130	104	82	68	58	52	50	52	58	68	82	104	130	160	194	232	274
4	232	194	160	130	104	82	64	50	40	34	32	34	40	50	64	82	104	130	160	194	232
3	194	160	130	104	82	64	50	36	26	20	18	20	26	36	50	64	82	104	130	160	194
2	160	130	104	82	64	50	40	26	16	10	8	10	16	26	40	50	64	82	104	130	160
1	130	104	82	64	50	40	34	20	10	4	2	4	10	20	34	40	50	64	82	104	130
0	104	82	64	50	40	34	32	18	8	2	0	2	8	18	32	34	40	50	64	82	104
-1	130	104	82	64	50	40	34	20	10	4	2	4	10	20	34	40	50	64	82	104	130
-2	160	130	104	82	64	50	40	26	16	10	8	10	16	26	40	50	64	82	104	130	160
-3	194	160	130	104	82	64	50	36	26	20	18	20	26	36	50	64	82	104	130	160	194
-4	232	194	160	130	104	82	64	50	40	34	32	34	40	50	64	82	104	130	160	194	232
-5	274	232	194	160	130	104	82	68	58	52	50	52	58	68	82	104	130	160	194	232	274
-6	320	274	232	194	160	130	104	90	80	74	72	74	80	90	104	130	160	194	232	274	320
-7	370	320	274	232	194	160	130	116	106	100	98	100	106	116	130	160	194	232	274	320	370
-8	424	370	320	274	232	194	160	146	136	130	128	130	136	146	160	194	232	274	320	370	424
-9	482	424	370	320	274	232	194	180	170	164	162	164	170	180	194	232	274	320	370	424	482
-10	544	482	424	370	320	274	232	218	208	202	200	202	208	218	232	274	320	370	424	482	544
x	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10

axes, it suffices to consider the nonnegative quadrant with $x_1 \geq 0$ and $x_2 \geq 0$.

If the EA is somewhere in the positive quadrant with $x_1 > 0$ and $x_2 > 0$ then it must do x_1 moves to the left and x_2 downward moves to reach the optimum in the origin. It can move to the left or downwards in each case with equal probability $\frac{1}{4}$, and it stays at the current position with probability $\frac{1}{2}$. If the EA has reached one of the axes then improvements can be achieved only in a single direction along the axis with probability $\frac{1}{4}$.

Since we consider effectively a singleobjective problem, we can apply all the techniques that have been developed so far for this case. Let random variable D_t denote the number of steps required to reach the optimum for the EA at position $X^{(t)}$ at step $t \geq 0$. The above argumentation reveals that $\mathsf{E}[\, D_t - D_{t+1}\,] \geq \frac{1}{4}$ so that the additive drift theorem [3] asserts $\mathsf{E}[\, T \,|\, D_0\,] \leq 4\, D_0$ where $D_0 = \|X^{(0)}\|_1$. Thus, we have proven:

Theorem V.1
The (1+1)-Refpoint-EA with local mutations and starting point $X^{(0)}$ requires $O(\|X^{(0)}\|_1)$ many steps to reach the reference point $r = (a,a)^\mathsf{T}$ with $a = 4$ of test problem (3) when using the squared ℓ_2-metric as distance measure. □

We dare to claim that this result also holds for arbitrary choices of $a > 0$ and that another reference point on the Pareto front will not change the order of the runtime.

VI. CONCLUSIONS

We have introduced two versions of (1+1)-EAs and applied them to a simple biobjective optimization problem in unbounded integer search space. Both EAs employ local mutations (i.e., adding ± 1 with equal probability to the parent vector at an arbitrary index). The first version accepts the offspring if it is not dominated by its parent, whereas the second version accepts the offspring if it is not more distant to some reference point than its parent. The formal analysis for the latter case revealed that the choice of the distance measure can have significant impact on the structure of the landscape (existence and size of fitness plateaus). The runtime analysis yields the result that both EA versions reach the Pareto set or front (1st version) and the reference point (2nd version) in $O(\|X^{(0)}\|_\infty)$ expected number of steps in dimension $n = 2$.

These results can be extended in various ways. The next steps should be a generalization for $n > 2$ and more realistic mutation operators like the maximum entropy distribution with unbounded support in $\mathbb{Z}^n$, that can also be coupled with self-adaptive step size control [4].

ACKNOWLEDGMENT

Parts of this work was done while the author visited Frank Neumann's lab at the University of Adelaide, Australia, in the course of a *short term scientific mission* (STSM) in April 2019 funded by COST Action CA 15140 of the European Union. The discussions with Frank about that topic were highly appreciated.

REFERENCES

[1] B. Doerr, C. Doerr, and T. Kötzing, "Static and Self-Adjusting Mutation Strengths for Multi-valued Decision Variables", *Algorithmica 80:1732—1768*, 2018. (doi: 10.1007/s00453-017-0341-1)

[2] J. Gadea Harder, T. Kötzing, X. Li, and A. Radhakrishnan, "Run Time Bounds for Integer-Valued OneMax Functions", CoRR abs/2307.11855, 2023 (doi: 10.48550/arXiv.2307.11855).

[3] J. He and X. Yao, 2004, "A study of drift analysis for estimating computation time of evolutionary algorithms", *Natural Computing 3(1):21-35*, 2004.

[4] G. Rudolph, "An Evolutionary Algorithm for Integer Programming", pp. 139-148 in Y. Davidor et al. (eds.): *Parallel Problem Solving From Nature (PPSN III)*. Berlin: Springer 1994.

[5] G. Rudolph, "On a multi-objective evolutionary algorithm and its convergence to the Pareto set", pp. 511-516 in: *Proceedings of the 5th IEEE Conference on Evolutionary Computation*, IEEE Press, Piscataway (NJ) 1998.

[6] A.P. Wierzbicki, "The use of reference objectives in multiobjective optimization", pp. 468-486 in: G. Fandel & T. Gal (eds.): *Multiple Criteria Decision Making; Theory and Applications*, Springer: Berlin-Heidelberg. 1980.