Abstract—In this article, we present a novel evolutionary algorithm for approximating the efficient set of a multiobjective optimization problem (MOP) with continuous variables. The algorithm is based on populations of variable size and exploits new rules for selecting alternatives generated by mutation and recombination. A special feature of the algorithm is that it solves at the same time the original problem and a dual problem such that solutions converge towards the efficient border from two “sides”, the feasible set and a subset of the infeasible set. Together with additional assumptions on the considered MOP and further specifications on the algorithm, theoretical results on the approximation quality and the convergence of both subpopulations, the feasible and the infeasible one, are derived.

I. INTRODUCTION

In the area of multiobjective decision making (MODM) (see, e.g., [13], [31], [34], [35]), a significant number of algorithms based on evolutionary approaches have been proposed during the last two decades (see survey articles by [6], [19], and [32], and, for more recent results, the proceedings of the first three International Conferences on Evolutionary Multi-Criterion Optimization edited by Zitzler et al. [36], Fonseca et al. [7]), and Coello Coello et al. [3]). Especially, for analyzing hard-to-solve multiobjective optimization problems (MOPs) such multiobjective evolutionary algorithms (MOEAs) have become a flourishing area of research with successful applications to various real-life problems.

Typically, such problems, based on continuous and/or discrete variables, are characterized by numerous restrictions. These restrictions cause offspring entities generated during the run of a MOEA by mutation and/or recombination to be infeasible. Typical means to cope with infeasible offspring are, for instance, the anew generation of the offspring (see, e.g., [11]), the usage of a penalty function (see, e.g., [30]), or a repair mechanism for modifying an infeasible solution into a feasible one (see, e.g., [23]).

For the first strategy, the possibly high effort for repeatedly generating infeasible entities is lost. For the second approach, it is not clear in general whether the penalty function causes infeasible solutions to re-enter the feasible domain. For the third approach, it is often not obvious how to construct an efficient repair operator, i.e., an operator mapping infeasible solutions ‘uniformly’ onto the space of feasible alternatives. Therefore, it is of considerable importance in real-life applications (which require an economical consumption of computing time) to seek for alternative ideas in dealing with infeasibility.

In this article, we outline a MOEA which actively uses infeasible offspring as solutions to a ‘dual MOP’. Being based on a MOEA for approximating the efficient set (or Pareto frontier) of an MOP discussed in [15], the new algorithm approximates the Pareto frontier from “both sides”, the feasible and the infeasible one. In this way, information is used more economically and hard-to-reach regions of the efficient set might be approached more easily. Because of optimizing the original, primal and a dual MOP at the same time, the new algorithm is called primal-dual multiobjective evolutionary algorithm, or PDMOEA for short. For the tabu search metaheuristics, a similar approach of utilizing infeasible solutions has been explored in [10].

Off course, the preferability of a mechanism for dealing with infeasibility very much depends on the specific problem. For a highly constrained problem, it may be hard to obtain feasible alternatives at all. In such cases using a penalty approach might be better than using the current algorithm which assumes that there is a reasonable probability of finding feasible solutions in the neighborhood of already found feasible or infeasible solutions.

The article is organized as follows: In Section 2, the MOP and notations related to solution sets are introduced. A dual MOP is analyzed in Section 3. Section 4 presents a description of the new MOEA. Theoretical results for the algorithm are discussed in Section 5. In Section 6, a numerical example is presented. Section 7 gives some conclusions.

II. SOME NOTATIONS

Mathematically, multiobjective optimization problems (MOPs) can be described by (P)

\[
\begin{align*}
\text{"min" } f(a) \\
\text{s.t. } a \in A
\end{align*}
\]

(1)

(2)

where \( f : \mathbb{R}^n \rightarrow \mathbb{R}^g \) is a vector-valued objective function and “\text{min}” means that each of the objective functions (the components of \( f \)) should be minimized. Instead of minimization, a maximization can be assumed as well. \( A \subset \mathbb{R}^n \) is called the feasible set which is usually defined by restriction functions,

\[
A = \{ a \in \mathbb{R}^n : g_j(a) \leq 0, j \in \{1, ..., m\} \}.
\]

(3)
Usually for an MOP, there does not exist a unique solution which optimizes all objective functions at the same time. Therefore, mostly the set of efficient or Pareto-optimal alternatives is regarded as the solution set of the problem (1)–(2). For specifying this and related sets, the Pareto relation “≺” defined by

\[ x \triangleq y :\iff x_i \leq y_i \forall i \in \{1, ..., q\} \text{ and there exists } j \in \{1, ..., q\} \text{ with } x_j < y_j \]

(4)

for \( x, y \in \mathbb{R}^q \) is used. The component-wise generalization of the scalar “≺” is then defined by

\[ x \preceq y :\iff x_i \leq y_i \forall i \in \{1, ..., q\} \]

(5)

Using the Pareto relation, the efficient set is defined by

\[ E(A, f) := \{ a \in A : \nexists b \in A : f(b) \prec f(a) \}. \]

(6)

The \( \varepsilon \)-efficient set can be defined by

\[ E_\varepsilon(A, f) := \{ a \in A : \nexists b \in A : f(b) \prec f(a) + (\varepsilon, ..., \varepsilon) \}. \]

(7)

Thus, any feasible solution with a distance not more than \( \varepsilon \) (according to the maximum metrics) to an efficient solution is called \( \varepsilon \)-efficient. See [8] for a survey of further concepts of efficient sets and other mathematical solution concepts for MOPs.

The set of dominating alternatives with respect to a given set \( B \subseteq \mathbb{R}^m \) is defined as

\[ \text{Dom}(B, f) := \{ b \in \mathbb{R}^n : \exists c \in B : f(b) < f(c) \land \nexists d \in B : f(d) \leq f(b) \}. \]

(8)

Basically, for \( B = A \) and \( f(A) \) being closed, the dominating set consists of all nonfeasible vectors being (strictly) better than solutions from the efficient set.

While the efficient set usually contains several alternatives, in practice decision makers desire a single solution to be selected or realized as a solution of a decision problem. Therefore, many MCDM methods utilize additional information such as weights, aspiration levels, or trade-off information (see [13]).

The most popular idea of applying EAs for MOPs is to use them as a robust method for approximating the efficient set. This is especially important for hard-to-solve problems for which ‘standard’ algorithms for determining the efficient set do not exist. A subsequent application of other methods might then be desirable for selecting one alternative from the determined set of (approximately) efficient ones.

### III. A DUAL PROBLEM

With respect to the MOP (P) defined by equations 1–3 we can specify a dual problem (D) by ‘inverting’ the objective functions (replace \( \min \) by \( \max \)) and by considering only those alternatives which are infeasible for the original problem and in its dominance cone. The formal definition is as follows:

\[ \text{"max"} f(b) \]

\[ \text{s.t. } b \in \text{Dom}(A, f). \]

(10)

Since \( \text{Dom}(A, f) \) is not known when, for instance, an evolutionary algorithm is applied, this set may be replaced by an approximation taken from the current population (see the selection rules in Section IV C for details). Note that for any \( b \in \text{Dom}(A, f) \) there exists \( a \in A \) with \( f(a) \prec f(b) \).

Furthermore, we have that for any \( a \in A \) there does not exist \( b \in \text{Dom}(A, f) \) with \( f(b) \prec f(a) \).

Note that (9) can be rewritten as a minimization problem as follows:

\[ \text{"min"} - f(b) \]

(11)

such that the above introduced concept of the efficient set can be applied. Defining the efficient set for the dual problem analogously to (6) we find the following result:

**Duality theorem:**

Let \( a \) be an efficient solution of (P). Then \( a \) is also a dominating solution of (D). Let \( b \) be an efficient solution of (D). Then \( b \) is also a dominating solution of (P).

The proof is straightforward and omitted here for brevity.

Note that according to the definition of the dual problem, there is no duality gap as often observed for dual pairs of MOPs defined according to different duality concepts [2]. For further and more general results on duality in MOP, see [27].

### IV. THE ALGORITHM PDMOEA

Originally, evolutionary algorithms (EAs) have been developed for scalar, or ordinary, optimization problems (see, e.g., [4]), i.e. problems with one objective function only. Traditional optimization methods such as the simplex algorithm for linear optimization or Fletcher and Powell’s [4] method for nonlinear optimization assume special properties for the considered optimization problems, e.g. a linearity of the objective function and the restriction functions, convexity of the feasible set, or differentiability. In contrast to these methods, evolutionary algorithms are applicable to a wider range of optimization problems and proved to be a robust and, at the same time, fast optimization method (cf. [29], [30]).

#### A. General Framework

In the following, we sketch a new multiobjective evolutionary algorithm suitable for approximating the efficient set of an MOP (defined by equations 1–2 or, respectively, 9–10).

The basic idea of EAs is that from a set of intermediary solutions (population) a subsequent set of solutions is generated by imitating concepts of natural evolution such as mutation, recombination, and selection. From a set of ‘parents’ in generation \( t \in N \), a set of ‘offspring’ is generated using
some kind of variation. From the offspring set and, possibly, the parent set, the subsequent solution set, i.e., the population of the next generation \( t+1 \), is selected.

The novel concept of the PDMOEA is to solve both problems, the original one and the dual problem, simultaneously. Therefore, we distinguish the parent and the offspring solution sets into feasible and non-feasible solutions. Let us denote the populations of the original and the dual problem in generation \( t \) as follows:

| TABLE I NOTATIONS USED FOR POPULATIONS IN THE PDMOEA. |
|---------------------------------|-------------------|
| parents                         | \( M^t \)          |
|                  | \( P^t \)          |
| offspring                       | \( N^t \)          |
|                  | \( Q^t \)          |

In that terminology, the complete parent population, \( M^t \cup P^t \), in generation \( t \in N \), serves to generate the offspring, being either in \( N^t \) or in \( Q^t \). From \( N^t \cup M^t \) and \( Q^t \cup P^t \) the subsequent solution sets, \( M^{t+1} \) and \( P^{t+1} \), are selected. In below, we consider the population sets, \( M^t, N^t, P^t, Q^t, t \in N \), to be multisets such that identical elements may be included several times in each of them.

In Fig. 1, the general framework of the evolutionary algorithm is shown. Specifications necessary for multiobjective optimization and for guaranteeing the approximation of the efficient set are discussed in the following sections.

B. The Usual Genetic Operators

In the discussion of MOEAs, the specification of the operations for mutation and recombination are of smaller interest than the selection. This is because in the first instance, the multiobjective nature of the objective function only affects the fitness evaluation of a solution, and thus the selection step. Of course, it may be useful to modify the mutation and recombination operators as well in order to solve multiobjective optimizations problems faster or with a better accuracy. However, for brevity these genetic operators and steps of the PDMOEA are discussed rather quickly in the following. More details on the design of these operators can be found in [15].

1) The starting population \( M^0 \) is assumed to consist of feasible alternatives only, i.e., \( M^0 \subseteq A \). Since the algorithm is based on the concept of a growing population, having only a single starting solution would be sufficient. Alternatively, starting with \( \mu > 1 \) copies of the same starting solution would be feasible as well. The starting solution(s) may be generated randomly or given by the user. For the dual problem, starting solutions are not required, i.e., the algorithm starts with \( P^0 = \emptyset \).

2) Basically, an offspring population \( N^t \cup Q^t \) consisting of \( \lambda \geq 1 \) entities is created by replicating the data of the current ‘parent’ populations \( M^t \) and \( P^t \) (for \( t = 0, 1, \ldots \)). The most simple approach for doing so is to select parent alternatives \( a^{ti} \in \{ a^{t1}, \ldots, a^{tn} \} = M^t \cup P^t \) randomly with equal probability for each offspring alternative \( b^{tj} \) for \( j = 1, \ldots, \lambda \) and to copy data, i.e., \( b^{tj} = a^{ti} \) for some \( i \in \{ 1, \ldots, \mu^t \} \).

3) For having some variation in the offspring data taken from the parents it is usually advantageous to use data from more than one parent to construct an offspring entity [29], [30]. The following concept of bisexual reproduction can be used for this recombination: For each \( j \in \{ 1, \ldots, \lambda \} \) and with a specific probability, \( p_{\text{reco}} \leq 0 \leq p_{\text{reco}} < 1 \), the offspring alternative is reproduced by using two randomly chosen parents, \( a^{ts}, a^{tk} \in M^t \cup P^t \). The data of offspring entity \( b^{tj} \) is then, for each vector component \( b^{ts}, s = 1, \ldots, n \), randomly chosen (with equal probabilities) from one of the parents. Note that the parents may come from \( M^t \) or \( P^t \).

In this way, recombining a feasible and an infeasible parent might especially be useful to find an ‘intermediate’ solution.
being closer to the efficient front. An example of such a feasible-infeasible recombination is illustrated by Figure 2.

![Recombination with infeasible solutions](image-url)

**Fig. 2.** Recombination with infeasible solutions may be advantageous.

4) A further mechanism of introducing completely new data into the population is realized by mutation. This is accomplished by adding some ‘noise’ defined by the vector-valued random variables \( z^j, t \geq 0, j \in \{1, ..., \lambda \} \), to the data of the offspring alternatives. \( z^j \) are assumed to be normally distributed with expected value \( 0 \in \mathbb{R}^n \) and standard deviation \( \sigma^j \in \mathbb{R}^n \) with \( \sigma^j_i > 0 \) for any \( i \in \{1, ..., n \} \):

\[
b^{ij} := b^{ij} + z^j, j = 1, ..., \lambda. \tag{12}
\]

In this way, any alternative in the solution space may be reached and locally efficient regions may be permuted [12].

5) When the offspring is finally generated, it has to be evaluated. This means that for each \( b^{ij} \) it is checked whether the alternative is feasible \((b^{ij} \in A)\) or not, i.e. whether it belongs to feasible offspring population, \( N^f \), or to the infeasible offspring population, \( Q^f \). Moreover, the objective values for the alternatives are calculated.

6) For the stopping condition of the generation loop being checked in each iteration, usually a maximum number of generations, \( t_{max} \), is specified. Other or additional stopping conditions could be given by requiring a minimum average progress per generation in objective space or in alternative space.

### C. Selection

With respect to the multiobjective nature of the considered optimization problem, the design of selection is the most interesting step in specifying the MOEA. This is because the concepts discussed in short (mutation, recombination, ...) have originally been developed for scalar evolution strategies and do not require an adaptation for multiobjective problems. For the selection step, the concept used in scalar evolutionary algorithms cannot be used without modification for the multiobjective case since scalar EAs use the linear order implicated by the single objective function. Using several criteria (or objective functions) instead of one, only the Pareto order for comparing alternatives is available. Especially, it is possible that two alternatives are incomparable with respect to several criteria, i.e. neither \( f(a) < f(b) \) nor \( f(b) < f(a) \) nor \( f(a) = f(b) \) holds.

For avoiding such frequently occurring situations [14] many MOEAs use some kind of substitute scalarization concept, e.g. an additive weighting, mapping the \( q \) objective functions to a single one. By doing so, the selection concepts used in scalar optimization EAs can be applied without modification. The most important disadvantage of this concept is, however, that, in general, not all efficient solutions can be reached or that the obtained solutions concentrate in a specific region of the efficient set [5], [6]. Such problems also occur with more advanced concepts of scalarization, like modifying the scalarization function in each step of the MOEA [20], [21], [5], [6].

To overcome these difficulties and to allow for a truly multiobjective analysis of the problem, other approaches have been developed which get by on the Pareto relation.

One of these concepts denoted as dominance grade is defined by counting the number of other alternatives which dominate the considered solution \( a \in A \), i.e.

\[
domgrade(a) := |\{b \in M^f \cup N^f : f(b) < f(a)\}|. \tag{13}
\]

A significant disadvantage of this and similar concepts is their low discriminative power. After a small number of generations, usually all or almost all parent entities are efficient with respect to the population, i.e. each parent entity has dominance grade 0. In such an “all efficient situation” [14] it is difficult for offspring entities to become new parents especially when the algorithm is conservative in the sense that parents are preferred to offspring for building the subsequent generation in the case that they are evaluated equally.

The main advantage of such elitist conservation principle is that they allow for the analysis of theoretical properties of MOEAs such as convergence. For instance, in [11] we have discussed a selection concept which allows to replace parent entities by offspring entities only when being dominated by one of them. For this kind of selection the convergence of the population to solutions in \( E(A, f) \) is proven.

The algorithm discussed in [11] as well as many others discussed in the literature is based on a constant population size. With such an assumption it is hardly possible to approximate or represent the whole efficient set \( E(A, f) \) well enough since the number of required solutions to approximate \( E(A, f) \) with a given exactness \( \varepsilon \) cannot, in general, be estimated in advance. Therefore, MOEAs based on a variable population size or utilizing some kind of archive have been developed. In [15] we have analyzed such an approach utilizing an increasable population for which a property of approximating the efficient set could be shown.

For the PDMOEA we choose a similar concept allowing to increase the population by new alternatives in order to support approximation. This is especially important since we assume that generating and evaluating new solutions might
be computationally expensive (see [17] for an example) such that even infeasible alternatives might be worth keeping for approximating the efficient set from the ‘infeasible side’.

Based on these considerations, we propose the following selection rules:

1) If \( a \in M^t \cap E_c(M^t \cup N^t, f) \) then \( a \in M^{t+1} \). (Keep feasible parents which remain \( \varepsilon \)-efficient.)

2) If \( a \in N^t \cap Dom(M^t) \) then \( a \in M^{t+1} \). (Add feasible offspring which dominates efficient parents.)

3) If \( a \in N^t \cap E(M^t \cup N^t, f) \) and \( \text{dist}(a, b) \geq \varepsilon \forall b \in M^t \cap E(M^t \cup N^t, f) \cup N^t \cap Dom(M^t, f) \) then \( a \in M^{t+1} \). (Add feasible efficient offspring which has a minimum distance of \( \varepsilon \) to all already selected parent and offspring alternatives.)

4) If \( a \in P^t \cap E_c(P^t \cup Q^t, -f) \cap Dom(M^{t+1}, f) \) then \( a \in Q^{t+1} \). (Keep parents which remain feasible and \( \varepsilon \)-efficient for the dual problem.)

5) If \( a \in Q^t \cap Dom(P^t, -f) \cap Dom(M^{t+1}, f) \) then \( a \in P^{t+1} \). (Add offspring being feasible for the dual problem which dominates efficient parents of the dual problem.)

6) If \( a \in Q^t \cap E(P^t \cup Q^t, -f) \) and \( \text{dist}(a, b) \geq \varepsilon \forall b \in P^t \cap E(P^t \cup Q^t, -f) \cup Q^t \cap Dom(P^t, -f) \) then \( a \in P^{t+1} \). (Add dual efficient offspring which has a minimum distance of \( \varepsilon \) to all already selected dual feasible parent and offspring alternatives.)

These selection rules guarantee that we do not lose any solutions which are ‘good’ with respect to the original or the dual problem. An alternative in \( M^t \) or \( P^t \) is replaced by an offspring alternative only if it is not \( \varepsilon \)-efficient (dual \( \varepsilon \)-efficient) within \( M^t \) or \( P^t \) or if it is dominated by an alternative in \( N^t \) or \( Q^t \). In both cases, efficient alternatives (with respect to (P) or (D) which are within an \( \varepsilon \)-neighborhood of the another alternative in \( M^t \) or \( P^t \) are still within the \( \varepsilon \)-neighborhood of an alternative of the next population. Note that the above selection rules are softened compared to [15] based on a remark by [28]. For the distance function \( \text{dist}, \) a maximum metrics is chosen for allowing compatibility with the definition of \( \varepsilon \)-efficient solutions.

For applying selection rules 5) and 6) it is necessary to have found at least one feasible solution (contained in \( M^t \)). This may, of course, be difficult for some kind of problem but is frequent assumption for test problems as well as in practical problem solving. In highly constrained optimization problems, it would also be possible to combine the current approach with a penalty function approach to reduce the number of constraints. However, during the run of the algorithm, new solutions generated by mutating feasible solutions are with a rather high probability ‘close’ to them (in particular, when the mutation rates are decreased along with a better approximation and when recombination is not applied). In such cases, it can be expected that only a few (if any) constraints are violated. Usually, in the neighborhood of an efficient solution only one constraint (if any) is active. Otherwise, there would be some redundancy in the modelling of the problem (see [9]).

V. THEORETICAL RESULTS

In [11], [15] we have given various theoretical results on the convergence of multiobjective evolutionary algorithms. In [11] new concept of \( \varepsilon \)-efficient solutions has been introduced and used for analysis of a MOEA. This concept is based on \( \varepsilon \)-neighborhoods and does not require scalarization as some other definitions of \( \varepsilon \)-efficient solutions (cf. [18]). The approximately efficient set of an MOP is defined as \( E_c(A, f) := \{ a \in A : (f(A) + K) \cap f(A) \subseteq f(U_\varepsilon(a)) \} \) where \( K \) is the dominance cone and \( U_\varepsilon \) is an \( \varepsilon \)-neighborhood around the considered point.

Using this concept, the convergence of the MOEA to some efficient solutions is shown, in [11]. In [15] we have extended this result for a MOEA with a nonconstant (usually increasing) population. It is shown that each efficient solution will asymptotically be approximated by the algorithm, i.e. that the population converges to the complete efficient set.

Since the generation of the parent populations \( M^t \) is very similar to their generation in the AMOEBA discussed in [15], the convergence results given in that article can be used for the new algorithm. For the PDMOEA without stopping criterion and for \( \sigma_i^0 > 0 \) for all \( t \geq 0, i \in \{1, \ldots, n\} \) the following theorems are proved under regularity assumptions similar as below (but including the closedness of \( A \)):

1) Convergence in probability for \( M^t \) towards \( E_c(A, f) \) for some \( \varepsilon > 0 \).

2) Almost sure convergence of \( M^t \) to \( E(A, f) \).

For the PDMOEA we can moreover show convergence results for the infeasible population \( P^t \). The special feature of our algorithm is that it also works for non-closed sets (if the complementary set is closed) and for sets (regions) with an empty interior.

Other results on convergence in MOEAs for continuous set of alternatives are given in [24] and [33]. Most recent results are discussed in [28]. Convergence results for finite search spaces can be found in [26]. A convergence result for finite problems with a partial ordering defining the favorability of alternatives is provided in [25]. Another convergence result is provided in [22].

A. Regularity Conditions of the Original and Dual MOP Problems

For deriving some theoretical results on our AMOEBA several further assumptions, some on the considered problem, others on the algorithm, are necessary:

1) \( A \neq \emptyset \).

This condition is also necessary for the algorithm to work since a feasible starting solution is required.

2) \( A \) is bounded.

With respect to our article [15] we do not require any more that \( A \) is closed.

3) \( A \) can be described by (3) using restriction functions \( g_j, j \in \{1, \ldots, m\} \), which are continuous.

4) The objective functions \( f_k, k \in \{1, \ldots, q\} \), are continuous.

5) \( \text{Dom}(A, f) \neq \emptyset \).
Otherwise the dual problem would not have any feasible solution. In that case, the PDMOE would still work but approximation of the efficient set could only be accomplished from the ‘feasible side’.

B. Further Assumptions for the PDMOE
5) No stopping criterion is defined.
   This condition is necessary to allow for asymptotic convergence.
6) $\sigma^2_i > 0$ for all $t \geq 0, i \in \{1, ..., n\}$.
   This condition means that mutations are possible in any direction in alternative space $R^n$.

C. Approximation Results for $P^t$
Let us first introduce some approximation concepts to be used for the analysis of our algorithm:

Let $dist$ be a metrics on $R^n$. A set $H \subseteq R^n$ is called an $\varepsilon$-approximation of a set $S \subseteq R^n$ if for each $s \in S$ there exists $h \in H$ such that $dist(s, h) \leq \varepsilon$.

Theorem 1 (convergence in probability):
Let $(A, f)$ be a regular MOP with properties (1)-(4) and the PDMOE with properties (5)-(6). Let $\varepsilon > 0$ and let $cl(A)$ denote the closure of $A$. Then there exists $\varepsilon > 0$ such that the probability of $P^t$ not being an $\varepsilon$-approximation of $E(cl(A), f)$ is $\leq e^{-ct}$ for all $t \geq 0$.

Proof:
As shown in [15] it is possible to approximate $E(A, f)$ with exactness $\varepsilon$ with a set of $[d/\delta]^n$ points, each of them being located in a unique cube $C_{i_1, ..., i_n}$. In that formula, $d := \max_{a, b \in E(A, f)} |a - b|$ is the maximal distance between two efficient points and $\delta := \varepsilon / \sqrt{n}$. Since $A$ is bounded, we can generalize this construction to $E(cl(A), f)$.

Because of the regularity conditions (3) and (4) we know that for each $a \in E(cl(A), f)$ there exists an $\varepsilon$-neighborhood $U_{\varepsilon}(a)$ such that $int(U_{\varepsilon}(a)) \neq \emptyset$. Let us assume that $a$ is close to a feasible solution of the dual problem (such a solution exists according to regularity assumption (5)), then $int(U_{\varepsilon}(a) \cap Dom(A, f)) \neq \emptyset$. Furthermore, there exists $a' \in U_{\varepsilon}(a) \cap Dom(A, f)$, $\varepsilon' > 0$ such that $U_{\varepsilon'}(a') \subseteq U_{\varepsilon}(a) \cap Dom(A, f)$.

For each $C_{i_1, ..., i_n}$ with $C_{i_1, ..., i_n} \cap Dom(A, f) \neq \emptyset$ let $a_{i_1, ..., i_n}$ be some element in $E(cl(A), f)$ and $a'_{i_1, ..., i_n}$ some point in $U_{\varepsilon}(a) \cap Dom(A, f)$ and $\varepsilon_{i_1, ..., i_n} > 0$ such that $U_{\varepsilon_{i_1, ..., i_n}}(a'_{i_1, ..., i_n}) \subseteq U_{\varepsilon}(a) \cap Dom(A, f)$.

Note that for any point $b \in U_{\varepsilon_{i_1, ..., i_n}}(a'_{i_1, ..., i_n})$ we have $dist(a, b) \leq \varepsilon$ for any $a \in C_{i_1, ..., i_n}$.

Let $\varepsilon^* = \min \{\varepsilon_{i_1, ..., i_n} \mid \sigma_i > 0 \}$ for $i \in \{1, ..., n\}$; $s := \min \{i \mid \sigma_i > 0\}$. Let $a^* \varepsilon^*$ be defined as a point in the closure of the decision space $cl(A)$ being most distant from any point in $E(cl(A), f)$ with respect to a Euclidean norm $| \cdot | : R^n \rightarrow R$, i.e. $a^* \varepsilon^* \in \arg \max \{a \mid dist(a, b) \}$.

Finally, let $\sigma$ being the vector $\sigma_{i_1, ..., i_n} \in E(cl(A), f)$, $i_j$ as above, with maximal distance to $a^* \varepsilon^*$. Considering the density function used for the mutations, the probability of finding a point (an offspring alternative) within any of the $C_{i_1, ..., i_n}$ (if $int(C_{i_1, ..., i_n}) \cap Dom(A, f)$) is larger than or equal to

$$p := (1 - p_{\text{pro}})(1/s\sqrt{2\pi})^n \int_{\cdots} \exp(-1/2s^2 \sum_{i=1}^{n} (a_i - a_i^*)^2)da$$

Within $\gamma := [k/\lambda]$ generations at least $k$ offspring alternatives are generated, and the probability that each of the $C_{i_1, ..., i_n}$ (with $int(C_{i_1, ..., i_n}) \cap Dom(A, f)$) contains one of them is $\geq p^k$. Using the general framework of the algorithm, we can decompose the probability as follows:

$$p(\max_{a \in E(cl(A), f)} \min_{b \in P^t} dist(a, b) > \varepsilon) \leq p(\max_{a \in E(cl(A), f)} \min_{b \in P^t} dist(a, b) > \varepsilon) \cdot (1 - p^k)$$

and

$$p(\max_{a \in E(cl(A), f)} \min_{b \in P^t} dist(a, b) > \varepsilon) \leq \Pi_{j=1}^{\gamma} (1 - p^j)$$

With

$$p(\max_{a \in E(cl(A), f)} \min_{b \in P^t} dist(a, b) > \varepsilon) \leq 1$$

for $t = 0, 1, ...$ and $\varepsilon := p^k / \gamma$, we get

$$p(\max_{a \in E(cl(A), f)} \min_{b \in P^t} dist(a, b) > \varepsilon) \leq e^{-\varepsilon t}$$

(14)

From this, we can conclude immediately:

$$p(\max_{a \in E(cl(A), f)} \min_{b \in P^t} dist(a, b) > \varepsilon) \leq e^{-\varepsilon t}$$

for all $t \geq 0$.

Remark: From Theorem 1, we can conclude that with probability $p \geq 1 - e^{-\varepsilon t}$, $P^t$ is an $\varepsilon$-approximation for $E(cl(A), f)$.

Theorem 2 (almost sure convergence):
Let $(A, f)$ be a regular MOP and the PDMOE be fulfilling the properties (5)-(6) and the $\varepsilon$-values in each generation, $\varepsilon^t > 0$, be a decreasing sequence with $\varepsilon^t \to 0$ for $t \to \infty$.

Then

$$p(\lim_{t \to \infty} \max_{a \in E(cl(A), f)} \min_{b \in P^t} dist(a, b) = 0) = 1.$$

Proof:
We have a monotonicity of the events

$$\{\max_{a \in E(cl(A), f)} \min_{b \in P^t} dist(a, b) > \varepsilon \} \supseteq \cdots \supseteq \{\max_{a \in E(cl(A), f)} \min_{b \in P^t} dist(a, b) > \varepsilon \}$$

With Theorem 1 and because $e^{-\varepsilon t} \to 0$ for $t \to \infty$ and $\varepsilon > 0$ we can conclude

$$p(\lim_{t \to \infty} \max_{a \in E(cl(A), f)} \min_{b \in P^t} dist(a, b) > \varepsilon) = \lim_{t \to \infty} p(\lim_{t \to \infty} \max_{a \in E(cl(A), f)} \min_{b \in P^t} dist(a, b) > \varepsilon) = 0.$$  

(15)
almost surely for the population
(6) applied to an MOP with properties (1)-(4) converges
are generated with increasing
\[ \lim_{i \to \infty} \varepsilon_i = 0 \text{ and } 0 < \varepsilon_{i+1} \leq \varepsilon_i \text{ for all } i \geq 0. \]
Then we have a monotonicity
\[
\{ \cup_{t=0}^{\infty} \max_{a \in E(cl(A), f)} \min_{b \in P^t} \; \text{dist}(a, b) \leq \varepsilon_i \}
\subseteq
\{ \cup_{t=0}^{\infty} \max_{a \in E(cl(A), f)} \min_{b \in P^t} \; \text{dist}(a, b) \leq \varepsilon_{i+1} \}
\]
for each \( i \in N \). From this it follows that
\[
p\left( \cap_{i=1}^{\infty} \cup_{t=0}^{\infty} \max_{a \in E(cl(A), f)} \min_{b \in P^t} \; \text{dist}(a, b) \leq \varepsilon_i \right) = \lim_{i \to \infty} p\left( \cup_{t=0}^{\infty} \max_{a \in E(cl(A), f)} \min_{b \in P^t} \; \text{dist}(a, b) \leq \varepsilon_i \right) = 1.
\]
From this, Theorem 2 follows immediately.

**Corollary:** The PDMOEAs with additional properties (5)-(6) applied to an MOP with properties (1)-(4) converges almost surely for the population \( P^t \) to the efficient set
\( E(cl(A), f) \) with an approximation accuracy \( \varepsilon \) for any \( \varepsilon > 0 \).

Since usually the number of efficient solutions is not finite, the PDMOEAs, therefore, do not converge to populations \( M^t \) and \( P^t \) of finite size. Instead, more and more alternatives are generated with increasing \( t \) (and decreasing \( \varepsilon \)) to “converge” to an arbitrarily exact representation of the efficient set. Without a stopping criterion, the algorithm does never terminate and requires an infinite capacity of storage.

Observe that the PDMOEAs may converge to a solution in \( Dom(A, f) \). Thus, if \( A \) is a nonclosed set, the algorithm may produce solutions arbitrarily close to solutions in \( A \) but dominating them. In a similar way, the algorithm may work for sets \( A \) with an empty interior (i.e. probability of finding a feasible solution is 0) or for feasible sets with small probabilities of finding alternatives randomly in them (i.e. for problems with very many restrictions).

**VI. A NUMERICAL EXAMPLE**

For demonstrating the new PDMOEAs we apply it to the following multiobjective optimization problem described in [12], p. 356:

\[
(P) \quad \text{min}_x f, \quad f: R^2 \to R^2 \text{ with } f_1 : x \mapsto x_i \text{ for } i \in \{1, 2\}
\]
and
\[
x \in X = \{ x \in R^2 : x_1 \geq 0, x_2 \geq 0, \\
x_2 - 5 + 0.5 x_1 \sin(4 x_1) \geq 0 \}.
\]

The problem \( P \) is characterized by 7 curve segments of efficient solutions which are separated by infeasible areas. (In Fig. 3, the curve which results from the third restriction and defines parts of the Pareto front is shown.) In such a case, it very much depends on the strength of the variation parameters (mutation rates, in particular) whether all these parts of the efficient set can be reached by the MOEA with a limited amount of time.

Fig. 3 shows the results of the algorithm after 1 second running time (1189 generations). All disconnected efficient areas are reached by the final population. Fig. 4 shows how infeasible solutions traverse an infeasible area: Infeasible solutions are accepted for the population in those parts of the objective space which connect the separated parts of the efficient set. This behavior can be characterized as a “tunneling effect”.

For comparison, problem \( P \) is analyzed using two other strategies for dealing with infeasible solutions: using recalculations of infeasible solutions and using a penalty function [16]. The rest of the applied MOEAs including the running time limit are identical to the PDMOEAs. In the case of using recalculations, only two of the 7 separated parts of the efficient set are reached. In case of using a penalty function, solutions in three of the separated parts are obtained. Thus, the PDMOEAs outperforms the other two variants in the considered example.

**VII. CONCLUSIONS**

In this article we have presented a new multiobjective evolutionary algorithm for approximating the efficient set which exploits information from generated infeasible solutions. This algorithm which considers also a problem being
due to the original MOP is called primal-dual multiobjective evolutionary algorithm, or PDMOEA for short.

Compared with other MOEAs, the PDMOEA does not require a time expensive re-calculation of infeasible solutions, the usage of a penalty function nor the employing of some repair mechanism for infeasible alternatives. Instead, infeasible solutions are used for approximating the efficient set (Pareto frontier) from the ‘other side’. Thus, the algorithm works economically with generated solutions. Especially, for hard-to-solve problems or problems with many restrictions such features should be of special importance.

For the algorithm, a double convergence can be shown. Under appropriate conditions, both the feasible and the infeasible subpopulations converge to the efficient set. The convergence results are, for some parts, based on our earlier results in [11] and [15].

For the future, we plan to apply the algorithm to typical hard-to-solve real-life problems, for instance problems in scheduling [17]. From a practical point of view, also the performing of further comparative studies with respect to other algorithms, especially other MOEAs, will be important.

REFERENCES


