A Multiobjective Approach for Finding Equivalent Inverse Images of Pareto-optimal Objective Vectors

Günter Rudolph and Mike Preuss

Abstract—**Supply bottlenecks or sudden changes in legal regulations may lead to the situation that certain factor combinations for producing some commodity cannot be used any longer. In this case it is important to know alternative factor combinations leading to a product with identical characteristics represented by a Pareto-optimal objective vector of a multiobjective optimization problem. Here, we present a biobjective approach that finds equivalent inverse images of a given Pareto-optimal objective vector, provided they exist.**

I. INTRODUCTION

Typically, evolutionary multiobjective algorithms (EMOA) deliver a set of objective vectors representing an approximation of the Pareto front. Elements on the Pareto front are characterized by the fact that an improvement with regard to an arbitrary objective must deteriorate at least one of the remaining objective values. Those elements in objective space are important for the product designer as varied values with regard to the objectives typify different products. The inverse images of these objective vectors in decision space determine the factor combinations for producing the commodity. Consequently, they are of interest for the product engineer being responsible of the production process. Since the mapping from decision space to objective space is not injective in general, it may well happen that different factor combinations may be used to produce equivalent commidities with identical properties as selected by the product designer. If the production is running short of a specific factor due to supply bottlenecks or if new legal regulations prohibit the usage of some factor, the product engineer needs to know these alternative factor combinations to keep the production running.

Unfortunately, standard versions of popular EMOAs typically maintain only a single inverse image per objective vector. As a result, equivalent inverse images are not at the product engineer's disposal when they are needed. Therefore we [1], [2] and others [3], [4] have had developed special purpose EMOAs which do not only approximate the Pareto front in objective space; rather, they were designed for covering the Pareto set in decision space as completely as possible. In this case we only need to enumerate the solutions in decision space until we find an equivalent inverse image (if it exists).

But actually it is not necessary to approximate the entire Pareto set in decision space. As the product designer selects

a single objective vector we only need to find all inverse images of this particular objective vector instead of the inverse images of all objective vectors. This observation leads to an approach with drastically decreased computational effort.

Let $x^* \in \mathcal{X} \subseteq \mathbb{R}^n$ be a Pareto-optimal solution of a multiobjective optimization problem

$$f(x) \to \min! \qquad \text{for } x \in \mathcal{X} \subseteq \mathbb{R}^n .$$

We are interested in solutions from decision space $\mathcal{X}$ that are also Pareto-optimal but distinct from x^*, i.e., we are seeking elements of the set $\{x \in \mathcal{X} : f(x) = f(x^*) \wedge \|x - x^*\| > \varepsilon\}$ for some $\varepsilon > 0$. Evidently, it is hopeless in general to seek for different solutions $x, y \in \mathcal{X}$ with $f(x) = f(y)$ if X is a continuous subset of $\mathbb{R}^n$. Therefore, we define two solutions $x, y \in \mathcal{X}$ as *equivalent* if $\|f(x) - f(y)\| < \delta$ for some arbitrary norm $\| \cdot \|$ and a prescribed $\delta > 0$.

The basic concept of the method presented here relies on the idea of seeking a solution in decision space that minimizes the distance to a given objective vector in objective space and maximizes the distance to the known inverse image in decision space. Thus, the problem of finding equivalent inverse images will be addressed by a biobjective surrogate problem.

In [5] this task was addressed by means of singleobjective optimization using penalty functions: minimize the distance in objective space and penalize the nearness to the original inverse image. More formally,

$$F(x) = \|f(x) - f(x^*)\|^2 - S(\|x - x^*\|) \to \min!$$

for some strictly monotonic increasing $S : \mathbb{R}_+ \to \mathbb{R}_+$ with $S(0) = 0$. Once an alternative solution (inverse image) has been detected in this manner, it is necessary to penalize also the nearness to this inverse image just found. Thus, we have to change the penalty function for each equivalent inverse image that is found. All in all we get a sequence of optimization problems. Numerical experience has shown that this approach works passably if we succeed in making appropriate design decisions for the shape of the penalty function and in adjusting its parameters.

This approach of finding all local optima sequentially can be traced back to the mid-1970s [6]. The idea was picked up by the development of the *filled function* approach [7] and the *tunneling* method [8] in the mid-1980s. Later it was reinvented in the field of Evolutionary Computation [9].

The next section II presents our biobjective approach along with a first empirical evaluation. In the last section we summarize our findings and close with our prospects regarding the extension of our method.

II. Biobjective Approach with Constraints

A. Basic Concept of "Multi-Objectivization"

The notion of "multi-objectivization" of singleobjective problems was coined by Knowles et al. [10]. The hypothesis of this approach is that singleobjective problems can be solved easier by adding additional objective functions. These additional objective function can be interpreted as *helper functions* [11]. Needless to say, it depends on the particular choice of the helper function if it is actually helpful for a multiobjective optimization method. It seems to be an open question if such a helper function always exists. There are a number of carefully constructed examples revealing in terms of formal proofs that additional objective functions can be helpful as well as harmful [12], [13], [14]. But these negative examples are of course not an argument against the existence and potential utility of helper functions: You simply have to find them!

We argue that helper functions are useful if they encode additional problem knowledge that can be exploited by by the optimization method. In case of a real-world problem in the field of thermodynamics [15] we were able to support this point of view: our helper function enlarged the basins of the local optima and it encoded implicitly which area of the search space should be avoided. This course of action has facilitated the detection of the desired optima significantly. As a consequence, we shall follow the approach of encoding additional knowledge in helper functions also for the problem considered here.

B. Formulation of the Problem

In the scenario considered here we assume that some approximation of the Pareto front has been determined already and that the product designer has chosen some objective vector from this approximation. The optimization method used for finding the approximation immediately delivers the corresponding inverse image x^* of the selected objective vector $f(x^*)$.

The product engineer now likes to know if there are additional inverse images $x \in \mathcal{X}$ that are different from x^* but map to the same objective vector $f(x^*)$. A necessary condition for such an inverse image $x \in \mathcal{X}$ is $\|f(x) - f(x^*)\| = 0$ for some norm $\| \cdot \|$. This observation leads to the first objective function

$$F_1(x) = \|f(x) - f(x^*)\|^2 \rightarrow \min! \qquad (1)$$

If we deploy a multistart approach it might be possible to identify additional equivalent inverse images by minimizing objective (1). But the original solution will be found also when proceeding in this manner. The original solution will be found the more frequent the larger is the region of attraction of the local optimum for the optimization method used. In order to avoid repeated heading to the known optimum the optimization method has to be "informed" appropriately that it ought not to search in the vicinity of the known optimum. This can be achieved by assigning worse objective function values to solutions within the undesired region. We realized

this idea by distinguishing inverse images with identical objective value in F_1 in terms of distance to the known solution x^*. Solutions with larger distance are preferred. This can be modelled by maximizing $\|x - x^*\|^2$, or equivalently

$$F_2(x) = -\|x - x^*\|^2 \rightarrow \min! \qquad (2)$$

Evidently, Pareto front as well as Pareto set of unbounded problems can be unbounded. Fortunately, there exists a natural constraint that enforces the boundedness of the Pareto front: As we are interested in inverse images whose images should be as similar as possible it is reasonable to introduce a maximum threshold $\theta > 0$ with respect to objective function (1):

$$g(x) = F_1(x) - \theta < 0 \qquad (3)$$

Finally, the biobjective surrogate problem has been specified by (1), (2) and (3), where (2) takes the role of the helper function: It encodes the knowledge that one should seek solutions distant to x^*.

C. Algorithm

The method proposed here is a hybrid metaheuristic: At first, we apply an evolutionary multiobjective optimization algorithm (EMOA) to the surrogate problem (1) - (3). The EMOA stops after a prescribed number of objective function evaluations and provides an approximation of the Pareto front. This set is sorted in ascending order with respect to objective (1) and we select a prescribed number of best solutions. Next, we run a singleobjective (local) optimizer for each of the best solutions using the inverse images as starting points and (1) as objective function. This second step is necessary for

1) localizing the optimum more accurately and
2) diagnosing if we are heading to an optimum already known.

For evaluating our approach we have used NSGA-II as EMOA and Hooke & Jeeves as (local) direct search method. The strategy parameters of the NSGA-II were $\eta_c = 15$, $\eta_m = 10$, $p_c = 0.6$ and $p_m = 1.0$.

D. Results

1) Test Problem TWO-ON-ONE: TWO-ON-ONE stands for the structurally simplest case of a multiobjective problem with a Pareto front that is multiply covered by disjoint parts of the Pareto set (see fig. 1). The multiple covering is caused by symmetries and the multimodality of the objective functions: a polynomial of 4th order and a simple paraboloid:

$$\begin{aligned} f_1(x) &= x_1^4 + x_2^4 - x_1^2 + x_2^2 - 10\,x_1\,x_2 + 20 \\ f_2(x) &= x_1^2 + x_2^2 \,. \end{aligned}$$

An empirical study in [1] revealed that all EMOAs tested (NSGA-II, SPEA2, SMS-EMOA) were not capable of covering both distinct subsets of the Pareto set for which each subset alone can represent the entire Pareto front in objective space. Even though both subsets were detected initially there was a rapid tendency to loose one of the subsets completely.
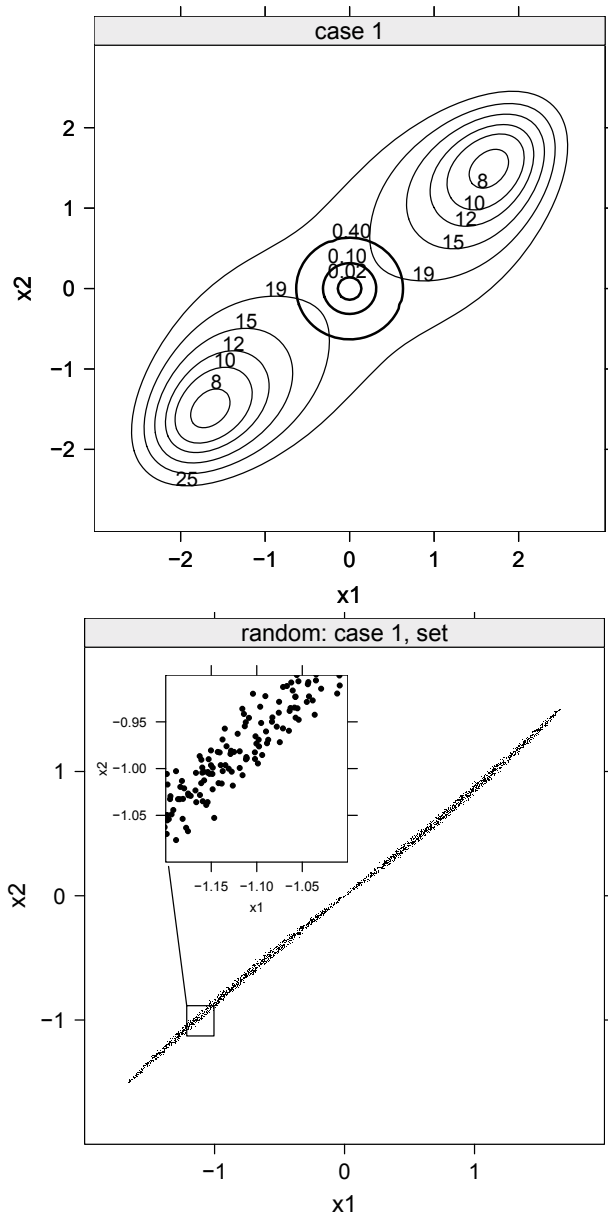
Fig. 1. Top: Contours of both objective functions of TWO-ON-ONE in decision space; the first function generates both outer peaks, the second function the peak in the center. Bottom: The Pareto set consists of a left and right "wing" which is actually a line. The areal impression is caused by (algorithmic relevant) imprecision in the sampling process. Each "wing" covers the Pareto front completely. The Pareto front is convex and resembles the shape of the pareto front for problem SYM-PART shown in fig. 4.

Fig. 2 illustrates the outcome of a typical run of NSGA-II with 100 individuals after 250 generations: The Pareto front appears to be well covered so that the product designer has sufficient information at his/her disposal for selecting the objective vector. Suppose the product designer chooses the objective vector $f(x^*) = (-17.4765, 4.97609)'$ with inverse image $x^* = (-1.595774, -1.558715)'$ from the approximation shown here. As can be seen from fig. 1 and the symmetry of the problem, the alternative equivalent inverse image is not in the approximation of the Pareto set.

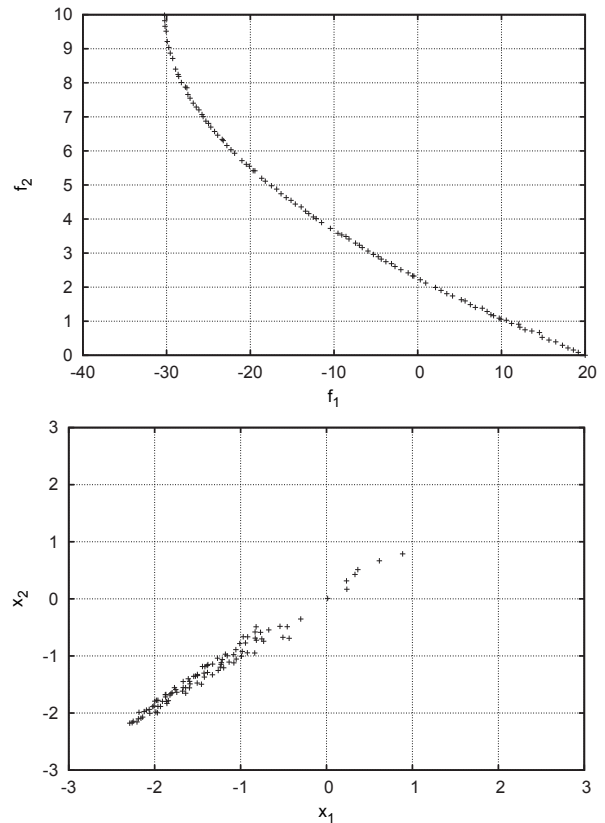In order to find an equivalent inverse image we initiate

Fig. 2. Result of a typical run of NSGA-II with 100 individuals after 250 generations. Top: Well covered Pareto front. Bottom: Incompletely covered Pareto set.

a run of our hybrid method: At first, the NSGA-II with 20 individuals is run for 100 generations in the search set $\mathcal{X} = [-5, 5]^2$ using maximum threshold of $\theta = 100$. The result of a typical run is given in fig. 3.

Next, we take a look at those 5 inverse images whose images are best with respect to F_1 (their objective vectors are the leftmost points in the left picture in fig. 3. Table I summarizes the results. Finally, we start HOOKE & JEEVES for a more accurate localization and we recognize that all starting points of table I converge to the same solution (different to the know solution). Evidently, only one equivalent inverse image does exist for this problem. The test problem in the subsequent subsection offers several equivalent inverse images.

rank	F_1	F_2	x_1	x_2
1	0.003433	-19.98722	1.705102	1.456486
2	1.326537	-21.57204	1.985328	1.398945
3	1.752043	-22.29680	1.502349	2.004770
4	3.937764	-23.40602	1.532254	2.132012
5	7.266334	-24.73223	1.594493	2.256312

TABLE I

THE 5 BEST SOLUTIONS OF THE SURROGATE PROBLEM FOR TWO-ON-ONE. WHEN USING THE INVERSE IMAGES (x_1, x_2) FROM THE TABLE AS STARTING POINTS OF A LOCAL OPTIMIZER OF F_1, WE ALWAYS FIND THE SAME EQUIVALENT INVERSE IMAGE.
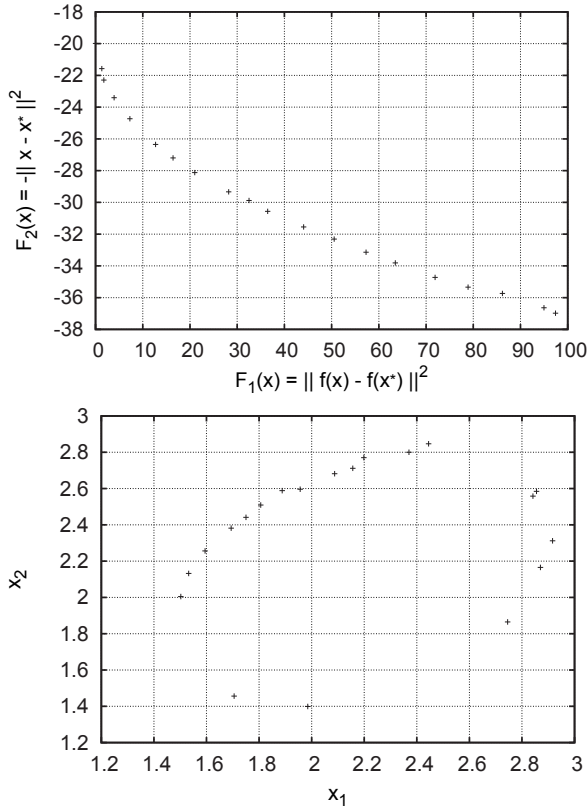
Fig. 3. Typical run of the NSGA-II for the surrogate problem based on TWO-ON-ONE. Top: Only the leftmost solutions in the figure are of interest since the distance to the desired objective vector is too large for the remaining solutions. Bottom: Corresponding inverse images of the surrogate problem. The solution with lowest F_1-value is $x = (1.705102, 1, 456496)'$.

2) Test Problem SYMPART: This test problem is an advancement of TWO-ON-ONE, as the number of equivalent inverse images of elements of the Pareto front is increased to 9 (see fig. 4). Moreover, both test functions are quadratic and identical except for translation. The behavior of commonly used EMOA for this type of problems was studied in [2] for the first time. As expected, all standard EMOAs failed to find and keep all 9 subsets.

This observation is not surprising as these EMOAs do not have built-in mechanisms for addressing this special problem. This problem has been included in the CEC 2007 contest [16] with high-dimensional decision space. Neither of the EMOAs participating in the contest was able to find more than one subset on average.

Finally, we like to examine how the hybrid method performs in case of multiple equivalent inverse images. Fig. 5 shows a typical run of the NSGA-II with 100 individuals after 250 generations in the region $\mathcal{X} = [-20, 20]^2$. The Pareto front seems to be well represented; no wonder as each of the 9 Pareto subsets covers the entire Pareto front completely. The approximation of the Pareto set, however, is incomplete. If the NSGA-II is run for some additional generations then the approximation of the 9 subsets becomes worse and worse due to drift effects.

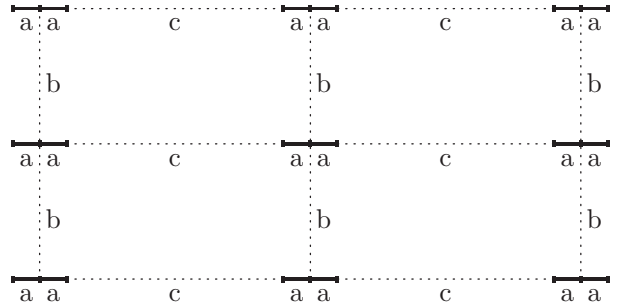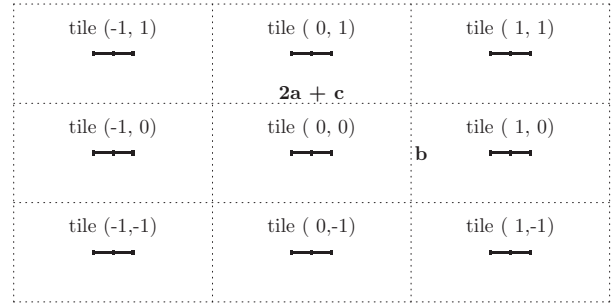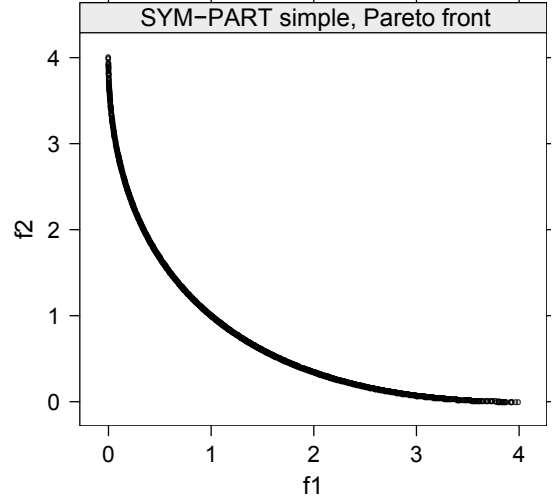Suppose the product designer would select objective vector

Fig. 4. Top: Pareto front of the SYMPART problem. Middle and bottom: Schematic setup in decision space consisting of 9 symmetric "tiles". Each tile contains a part of the Pareto set that covers the Pareto front completely. The Pareto front is generated by 2 simple quadratic objective functions. The Pareto subsets have length $2a$ and the vertical of horizontal distance to the next subset is characterized by the parameters b resp. c.

$f(x^*) = (0.757702, 1.27587)'$ with corresponding inverse image $x^* = (9.870459, -10.00175)'$. We run our hybrid method with 100 individuals for a maximum number of 300 generations and $\theta = 0.5$. The subsequent graphics in fig. (6) yield an impression about the situation after 100, 200, and 300 iterations.

Initially the individuals are far away from the Pareto front, but there are clear tendencies towards eqivalent inverse images. If we would stop after 100 generation and start local searches w.r.t. F_1 the 8 equivalent inverse images would be identified. The longer the NSGA-II runs the more equivalent inverse images are getting lost. This behavior might be explained to some extent by drift effects but more importantly by the fact that in case of accurate approximations w.r.t. F_1
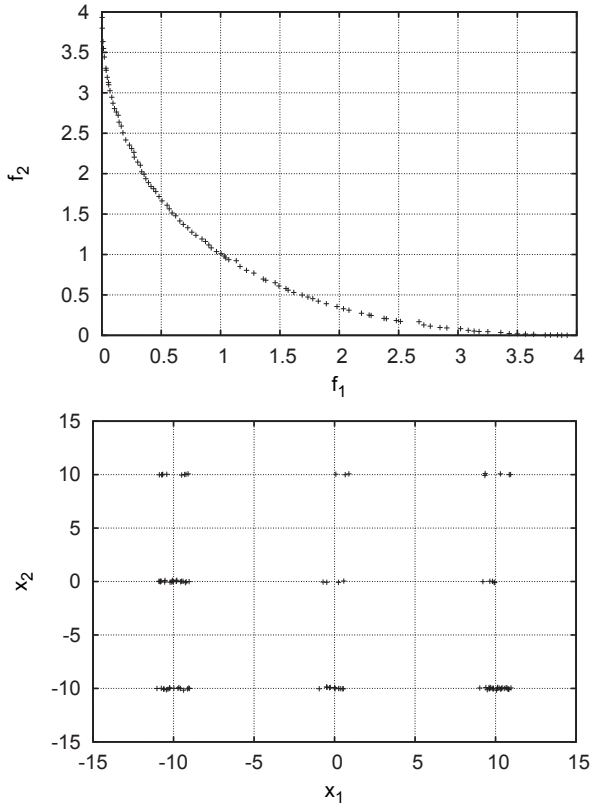
Fig. 5. Typical run of a NSGA-II with 100 individuals after 250 generations in region $\mathcal{X} = [-20, 20]^2$. Top: Well covered Pareto front. Bottom: Incompletely covered Pareto set.

individuals with smaller F_2 dominate solutions that are closer to the original inverse image x^*. This is a methodological problem that we plan to attack with an archive that collects individuals with good F_1 value and sufficiently large F_2 value. The size of the archive can be kept small by maintaining only representative solutions (e.g. cluster centroids).

III. CONCLUSIONS

We have introduced an evolutionary hybrid method for finding equivalent inverse images of Pareto-optimal objective vectors. The problem is addressed by creating a surrogate problem in terms of a constrained biobjective optimization problem. It was empirically shown that this approach works well for a single additional inverse image. If we need to find all inverse images simultaneously, then we have a severe problem that we plan to solve by an archiving strategy in future. Of course we are aware that we have only described a proof-of-principle, here. But we are convinced that the approach, if completely elaborated, will save many resources compared to special-purpose MOEAs that are designed to approximate also the entire Pareto set: If the original problem has $d \geq 2$ objectives and n decision variables then the dimensionality of the Pareto front can be as large as $\min\{d-1, n\}$. If d is large then a huge number of individuals is necessary to approximate the Pareto front and one needs a multiple of this number to approximate the entire Pareto set if there are multiple Pareto subsets with images completely covering

the Pareto front. In contrast, our approach always leads to a biobjective surrogate problem regardless of the number of objectives of the original problem. As a consequence, the complexity of the problem is reduced and much fewer function evaluations are necessary to identify equivalent inverse images of Pareto-optimal solutions.

REFERENCES

[1] M. Preuss, B. Naujoks, and G. Rudolph, "Pareto set and EMOA behavior for simple multimodal multiobjective functions," in *Parallel Problem Solving from Nature (PPSN IX)*, T. Runarsson *et al.*, Eds. Berlin: Springer, 2006, pp. 513–522.

[2] G. Rudolph, B. Naujoks, and M. Preuss, "Capabilities of EMOA to detect and preserve equivalent pareto subsets," in *Proceedings of the 4th International Conference on Evolutionary Multi-Criterion Optimization (EMO 2007)*, S. Obayashi *et al.*, Eds. Berlin: Springer, 2007, pp. 36–50.

[3] K. Chan and T. Ray, "An evolutionary algorithm to maintain diversity in the parametric and the objective space," in *Proceedings of IEEE International Conference on Computational Robotics and Autonomous Systems (CIRAS 2005)*. National University of Singapore: Centre for Intelligent Control, 2005.

[4] K. Deb and S. Tiwari, "Omni-optimizer: A generic evolutionary algorithm for single and multi-objective optimization," *European Journal of Operational Research*, vol. 185, no. 3, pp. 1062–1087, 2008.

[5] T. Dickhöfer, "Entwicklung und Vergleich von evolutionären Algorithmen zur Approximation äquivalenter Urbilder von Pareto-optimalen Zielvektoren," Master's thesis, Technische Universität Dortmund, Fakultät für Informatik, April 2007.

[6] A. Vilkov, N. Zhidkov, and B. Shchedrin, "A method of search for the global minimum of a function of one variable," *Journal of Computational Mathematics and Mathematical Physics*, vol. 15, pp. 1040–1042, 1975.

[7] R. P. Ge and Y. F. Qin, "A class of filled functions for finding global minimizers of several variables," *Journal of Optimization Theory and Applications*, vol. 54, no. 2, pp. 241–252, 1987.

[8] A. V. Levy and A. Montalvo, "The tunneling algorithm for the global minimization of functions," *SIAM Journal on Scientific and Statistical Computing*, vol. 6, no. 1, pp. 15–29, 1985.

[9] D. Beasley, D. Bull, and R. Martin, "A sequential niche technique for multimodal function optimization," *Evolutionary Computation*, vol. 1, no. 2, pp. 101–125, 1993.

[10] J. Knowles, R. Watson, and D. Corne, "Reducing local optima in single-objective problems by multi-objectivization," in *Proceedings of the First International Conference on Evolutionary Multi-Criterion Optimization (EMO 2001)*. Berlin: Springer, 2001, pp. 269–283.

[11] M. Jensen, "Helper-objectives: Using multi-objective evolutionary algorithms for single-objective optimisation," *Journal of Mathematical Modelling and Algorithms*, vol. 3, no. 4, pp. 323–347, 2004.

[12] F. Neumann and I. Wegener, "Minimum spanning trees made easier via multi-objective optimization," *Natural Computing*, vol. 5, no. 3, pp. 305–319, 2006.

[13] D. Brockhoff, T. Friedrich, N. Hebbinghaus, C. Klein, F. Neumann, and E. Zitzler, "Do additional objectives make a problem harder?" in *Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation (GECCO 2007)*. New York: ACM Press, 2007, pp. 765–772.

[14] J. Handl, S. Lovell, and J. Knowles, "Multiobjectivization by decomposition of scalar cost functions," in *Parallel Problem Solving from Nature (PPSN X)*, G. Rudolph *et al.*, Eds. Berlin: Springer, 2008, pp. 31–40.

[15] M. Preuss, G. Rudolph, and F. Tumakaka, "Solving multimodal problems via multiobjective techniques with application to phase equilibrium detection," in *Proceedings of the 2007 IEEE International Congress on Evolutionary Computation (CEC2007)*, K. Tan *et al.*, Eds. Piscataway (NJ): IEEE Press, 2007, pp. 2703–2710.

[16] V. L. Huang, A. K. Qin, K. Deb, E. Zitzler, P. N. Suganthan, J. J. Liang, M. Preuss, and S. Huband, "Problem definitions for performance assessment of multi-objective optimization algorithms," Nanyang Technological University, Singapore, Tech. Rep., 2007. [Online]. Available: CEC-07-TR-emoa-performance.pdf
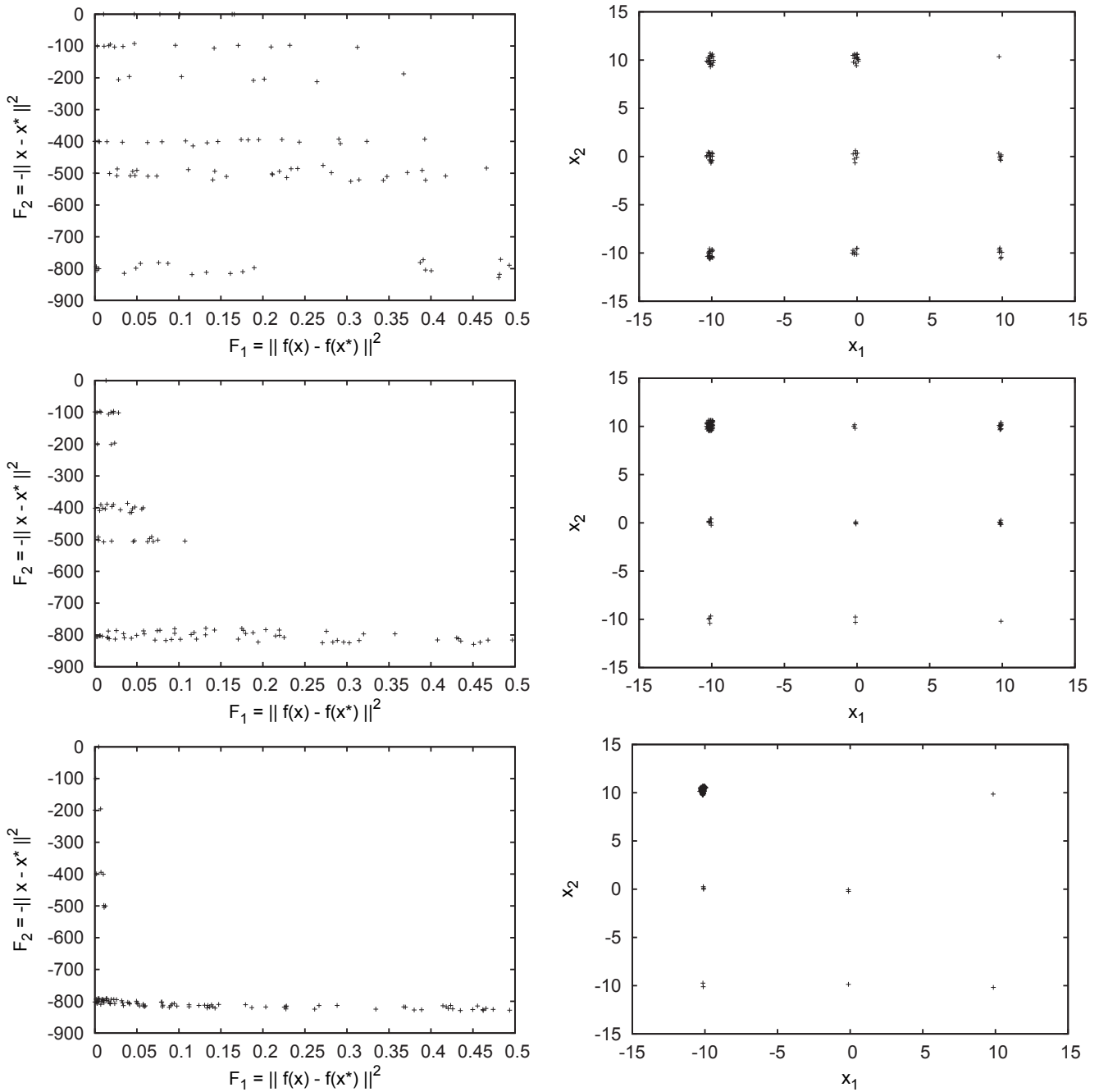
Fig. 6. Typical run of NSGA-II with 100 individuals after 100 (top), 200 (center) and 300 (bottom) generations in region $\mathcal{X} = [-20, 20]^2$. Left column: Approximations of Pareto front of surrogate problem. Right column: Corresponding inverse image.