The Set-Based Hypervolume Newton Method for Bi-Objective Optimization

Víctor Adrián Sosa Hernández, Oliver Schütze, Hao Wang, André Deutz, and Michael Emmerich

***Abstract*—In this paper, we propagate the use of a set-based Newton method that enables computing a finite size approximation of the Pareto front (PF) of a given twice continuously differentiable bi-objective optimization problem (BOP). To this end, we first derive analytically the Hessian matrix of the hypervolume indicator, a widely used performance indicator for PF approximation sets. Based on this, we propose the hypervolume Newton method (HNM) for hypervolume maximization of a given set of candidate solutions. We first address unconstrained BOPs and focus further on first attempts for the treatment of inequality constrained problems. The resulting method may even converge quadratically to the optimal solution, however, this property is—as for all Newton methods—of local nature. We hence propose as a next step a hybrid of HNM and an evolutionary strategy in order to obtain a fast and reliable algorithm for the treatment of such problems. The strengths of both HNM and hybrid are tested on several benchmark problems and comparisons of the hybrid to state-of-the-art evolutionary algorithms for hypervolume maximization are presented.**

***Index Terms*—Hessian matrix, hypervolume indicator, memetic algorithms, Newton method, set-based local search.**

I. Introduction

In MANY applications, the problem arises where several objectives have to be optimized concurrently leading to a multiobjective optimization problem (MOP). One important characteristic of MOPs is that their solutions sets, the Pareto sets (PSs), typically (under certain conditions when the MOP is continuous and nondegenerate) form objects of dimension $(k-1)$, where k is the number of objectives involved in the problem. In many cases, the decision maker is interested in a suitable finite size approximation of the entire PS, namely, its image, and the Pareto front (PF) [1]. To address this problem, specialized evolutionary algorithms are predominantly used [2]–[8]. Such methods are capable of returning a finite size approximation of the entire set of interest in one single run of the algorithm, and as they are very robust and of global nature [9], [10]. One major drawback of all such methods, however, is that they are quite costly in terms of function evaluations. That is, they typically require quite a few function evaluations in order to obtain suitable approximations. Moreover, they do not converge toward the optimal solution in the mathematical sense, instead, their populations may get stuck in nonoptimal (but in most cases in nearly optimal) regions.

The inclusion of gradient information has long been considered as a remedy to imprecision [11], [12]. However, the initially proposed methods were only able to improve the closeness to the PF, but not the diversity of approximations. To also include the diversity in gradient-based search for PFs, directed search methods were proposed in [13] and [14], which allow not only to move toward the PF based on local gradient information but also along this set in order to increase diversity. Alternatively, it proposed the use of the hypervolume gradient defined on an entire population to find diverse set approximations to the PF (see [15], [16]).

In this paper, we first present the hypervolume Newton method (HNM) for hypervolume maximization. The hypervolume [17] is a widely used indicator to measure the approximation quality of a set. The method allows for a finite size approximation of the PF of a bi-objective optimization problem (BOP) in one run of the algorithm. To this end, we first derive an analytical expression of the Hessian matrix of the hypervolume indicator (which is being considered as a particular scalar optimization problem defined on sets). Based on this, we formulate the set-based HNM first for unconstrained and later on for inequality constrained BOP. The resulting method is defined on entire sets and yields up to quadratic convergence rates. The method, however, is as all Newton methods of local nature and requires good starting sets. That is why in the second step of this paper, we propose a hybrid evolutionary algorithm that utilizes HNM as a local search engine. Numerical results against state-of-the-art evolutionary algorithms for hypervolume indicator maximization demonstrate the benefit of the novel approach.

A first study of this paper can be found in [18], where an HNM is proposed for unconstrained BOPS, albeit using finite difference approximations of the Hessians. The idea to use the

Manuscript received July 11, 2018; revised October 8, 2018; accepted November 26, 2018. Date of publication December 25, 2018; date of current version April 15, 2020. This work was supported by CONACyT under Project 285599. The work of V. A. Sosa Hernández was supported by CONACyT. The work of H. Wang was supported by NWO PROMIMOOC under Project 650.002.001. This paper was recommended by Associate Editor Y. S. Ong. *(Corresponding author: Víctor Adrián Sosa Hernández.)*

V. A. Sosa Hernández is with the School of Engineering and Sciences, Tecnologico de Monterrey, Atizapan de Zaragoza 52926, Mexico (e-mail: vsosa@tec.mx).

O. Schütze is with the Department of Computer Science, CINVESTAV-IPN, Mexico City 07360, Mexico, and also with the Department of Applied Mathematics and Systems, UAM Cuajimalpa, Mexico City 05370, Mexico (e-mail: schuetze@cs.cinvestav.mx).

H. Wang, A. Deutz, and M. Emmerich are with the Leiden Institute of Advanced Computer Science, Leiden University, 2333 CA Leiden, The Netherlands (e-mail: h.wang@liacs.leidenuniv.nl; a.h.deutz@liacs.leidenuniv.nl; m.t.m.emmerich@liacs.leidenuniv.nl).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCYB.2018.2885974

Newton method for general convex sets has been proposed in [19], however, the work does not deal with multiobjective optimization.

This paper is structured as follows. In Section II, we introduce important mathematical definitions and preliminaries. In Section III, the Hessian matrix of the hypervolume indicator is derived and based on it, the HNM is described in Section IV. Section V deals with the hybrid evolutionary algorithm and Section VI shows and discusses the results of this algorithm on common bi-objective benchmark problems. Finally, Section VII concludes this paper with a brief discussion of the main results and possible future work.

II. Preliminaries and Background

A. Set-Based Multiobjective Optimization

In this paper, continuous MOPs are considered

$$\min_{\mathbf{x}\in Q} \mathbf{f}(\mathbf{x}), \quad \mathbf{f} : Q \subseteq \mathbb{R}^n \to \mathbb{R}^m \tag{1}$$

where the vector-valued function is defined as the concatenation of m objective functions: $\mathbf{f}(\mathbf{x}) := (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x}))^\top$. Throughout this paper, it is assumed that each objective function, $f_i : Q \subseteq \mathbb{R}^n \to \mathbb{R}$ $(i = 1, 2, \ldots, m)$, is *twice continuously differentiable* almost everywhere in Q.

Optimality is based on the concept of dominance.

1) Let $\mathbf{v} = (v_1, \ldots, v_m)^\top, \mathbf{w} = (w_1, \ldots, w_m)^\top \in \mathbb{R}^m$. Then the vector $\mathbf{v}$ is less than $\mathbf{w}$ ($\mathbf{v} <_p \mathbf{w}$), if and only if $v_i < w_i$ for all $i \in \{1, \ldots, m\}$. The relation $\leq_p$ is defined analogously.
2) *Dominance:* A vector $\mathbf{x} \in Q$ dominates another vector $\mathbf{x}' \in \mathbb{R}^n$ ($\mathbf{x} \prec \mathbf{x}'$) with respect to the vector-valued function $\mathbf{f}$ (1) if $\mathbf{f}(\mathbf{x}) \leq_p \mathbf{f}(\mathbf{x}')$ and $\mathbf{f}(\mathbf{x}) \neq \mathbf{f}(\mathbf{x}')$ [i.e., there exists a $j \in \{1, \ldots, m\}$ such that $f_j(\mathbf{x}) < f_j(\mathbf{x}')$].
3) *Pareto Optimality:* A point $\mathbf{x} \in Q$ is called a Pareto optimal or a Pareto point if there is no $\mathbf{x}' \in Q$ which dominates $\mathbf{x}$. The set of all Pareto optimal solutions

$$\mathcal{X} := \{\mathbf{x} \in Q \mid \mathbf{x} \text{ is Pareto optimal with respect to } \mathbf{f}\}$$

is called the PS and the PF is the image of it under $\mathbf{f}$: $P_{\mathcal{X}} = \{\mathbf{f}(\mathbf{x}) \mid \mathbf{x} \in \mathcal{X}\}$.

We adhere to the set-based approach for MOPs. Moreover, as discussed in [16], the finite approximation sets of size μ to the efficient set $\mathcal{X}$ are represented by vectors $\mathbf{X}$ in $\mathbb{R}^{\mu\cdot n}$. Each of the search points can be obtained as follows: the first point $\mathbf{x}^{(1)} \in Q$ is represented by the first n consecutive components of $\mathbf{X}$, the second one by the following n consecutive components, etc.,

$$\mathbf{X} = \left(\mathbf{x}^{(1)^\top}, \mathbf{x}^{(2)^\top}, \ldots, \mathbf{x}^{(\mu)^\top}\right)^\top, \ \mathbf{x}^{(i)} \in Q, \ i = 1, \ldots, \mu.$$

In addition, a mapping $\mathbf{F} : \mathbb{R}^{\mu\cdot n} \to \mathbb{R}^{\mu\cdot m}$ is obtained from the objective function $\mathbf{f}$

$$\mathbf{F}(\mathbf{X}) := \left(\mathbf{f}\left(\mathbf{x}^{(1)}\right)^\top, \mathbf{f}\left(\mathbf{x}^{(2)}\right)^\top, \ldots, \mathbf{f}\left(\mathbf{x}^{(\mu)}\right)^\top\right)^\top. \tag{2}$$

Approximation sets (which we represent by vectors $\mathbf{X}$) to the efficient set give rise to the vector-representations of approximation sets Y to the PF by defining $\mathbf{Y} := \mathbf{F}(\mathbf{X})$. Note that the PF approximation set Y can be obtained from the $\mu \cdot m$ vector $\mathbf{Y}$ by the same segmentation operation on $\mathbf{X}$ as above

$$\mathbf{Y} = \left(\mathbf{y}^{(1)^\top}, \mathbf{y}^{(2)^\top}, \ldots, \mathbf{y}^{(\mu)^\top}\right)^\top, \ \mathbf{y}^{(i)} \in \mathbb{R}^m, \ i = 1, \ldots, \mu.$$

The benefits of such a treatment will become clear in the sequel, where the hypervolume indicator is related to the decision space through the mapping $\mathbf{F}$ and it is differentiated with respect to the $\mu \cdot n$-vector $\mathbf{X}$.

B. Multiobjective Evolutionary Algorithms

These days, apart from mathematical techniques, multiobjective evolutionary algorithms (MOEAs) have been successful in solving MOPs. Algorithms of this type are able to obtain good (finite) approximations of the efficient set. In the last decade, several MOEAs have integrated performance indicators as selection criteria. For instance, a framework for designing indicator-based evolutionary algorithms (IBEAs) is proposed in [20] by using any arbitrary indicator. Another example is the $\mathcal{S}$ metric selection EMOA (SMS-EMOA) [2] which takes the contribution of the hypervolume indicator into the selection operator. A one point steady-selection strategy is used, in order to guarantee a sequence of populations that is nondecreasing with respect to hypervolume and at the same time keep the computational cost of selection small. One more example can be found in [3]. There are also some other algorithms that have been designed using different performance indicators, for instance Δ_p [21] and $R2$ [22], for more information see [23], [24].

C. Hypervolume Indicator

The hypervolume indicator [17] is a common indicator to assess the quality of an approximation to the PF. In terms of minimization, the hypervolume indicator of the PF approximation Y is defined as the Lebesgue measure of the subspace dominated by Y, that is cut from above by a reference point $\mathbf{r} \in \mathbb{R}^m$

$$H(Y) := \lambda\left(\{\mathbf{p} \in \mathbb{R}^m | \mathbf{p} <_p \mathbf{r} \wedge \exists \mathbf{y} \in Y : \mathbf{y} <_p \mathbf{p}\}\right)$$

with λ being the Lebesgue measure in $\mathbb{R}^m$. Moreover, such a definition can be extended to the $\mu\cdot m$-vector $\mathbf{Y}$ by introducing the notation $\mathcal{H}(\mathbf{Y}) := H(Y)$. Based on the mapping stated in (2), the hypervolume indicator of $\mathcal{H}(\mathbf{Y})$ can also be further related to the $\mu \cdot n$-vector $\mathbf{X}$ by function composition

$$\mathcal{H}_{\mathbf{F}}(\mathbf{X}) := \mathcal{H}(\mathbf{F}(\mathbf{X})) = \mathcal{H}(\mathbf{Y}).$$

Note that the reformulation of the hypervolume indicator defines a mapping: $\mathbb{R}^{\mu\cdot n} \to \mathbb{R}$. Furthermore, such a performance indicator allows us to transform an MOP into the following scalar optimization problem in higher dimension:

$$\max_{\mathbf{X}\in\mathbb{R}^{\mu\cdot n}} \ \mathcal{H}_{\mathbf{F}}(\mathbf{X}) \quad \text{s.t.} \ \ \mathbf{x}^{(i)} \in Q, \quad i = 1, \ldots, \mu. \tag{3}$$

The hypervolume indicator has been successfully exploited also as a selection criterion in evolutionary multiobjective optimization algorithms (see [2], [3]).

D. Hypervolume Gradient Field

The hypervolume gradient is defined in [15], where it is stated that for the function $\mathcal{H}_{\mathbf{F}}$, the gradient field $\nabla\mathcal{H}_{\mathbf{F}}$ is defined as (the denominator layout is adopted for the calculation of the derivatives in this paper)

$$\begin{aligned}&\nabla\mathcal{H}_{\mathbf{F}}(\mathbf{X})\\&=\left(\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial x_1^{(1)}},\ldots,\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial x_n^{(1)}},\ldots,\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial x_1^{(\mu)}},\ldots,\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial x_n^{(\mu)}}\right)^{\top}\\&=\left(\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial\mathbf{x}^{(1)}}^{\top},\ldots,\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial\mathbf{x}^{(\mu)}}^{\top}\right)^{\top}.\end{aligned}\tag{4}$$

Note that each term in the right-hand side of the above equation is called *subgradient*, which is the local hypervolume change rate by moving only one decision vector infinitesimally. Moreover, the subgradients can be computed by applying the chain rule [16], [25]

$$\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial\mathbf{x}^{(i)}}=\frac{\partial\mathbf{Y}}{\partial\mathbf{x}^{(i)}}\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial\mathbf{Y}}=\sum_{k=1}^{m}\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial f_k\left(\mathbf{x}^{(i)}\right)}\nabla f_k\left(\mathbf{x}^{(i)}\right).\tag{5}$$

E. Hypervolume-Based Local Search Techniques Within MOEAs

In recent years, different local search techniques have been designed for improving the performance of some evolutionary algorithms by using first- or second-order derivative methods (see [13], [26]–[29]). However, it is not clearly specified how these improvements can (or should) be measured. The integration of an indicator can steer the local search in better directions and also be more effective for improving points. In the following, we describe some hypervolume-based local search algorithms that have been developed to construct memetic strategies. The hypervolume gradient was used as a gradient-ascent method for maximizing the hypervolume. Such a local search procedure was applied after performing one run of the SMS-EMOA [2]. The results showed how gradient-based indicator methods can improve the convergence rate of evolutionary algorithms. In [16], the definition of the hypervolume gradient was extended to $m > 2$ and also an efficient algorithm for its computation for $m \leq 4$ was provided. Sosa Hernández *et al.* [18] proposed the multiobjective HVDS. This local search procedure is based on the directed search method (see [14]) to improve the hypervolume contribution for a selected point or points in the current population. HVDS divides the objective space into three regions and for each region another local search strategy is proposed. The effect of the integration of the local searcher pushes the population to converge faster, which reduces the number of function evaluations to obtain a desirable approximation of the efficient set. Recently, the use of Newton method for sets based on the hypervolume indicator has been proposed [18]. This method uses the hypervolume gradient and an approximation of the Hessian matrix by using finite differences to compute a search direction. First, the results showed the potential of this method even by using an approximation, however, the exact computation of the hypervolume Hessian matrix and the extension of the usability of this method are called for.

III. Hypervolume Hessian Matrix

A. Derivation of the Hessian Matrix

We first derive the Hessian matrix of the hypervolume indicator for the general *multiobjective* optimization scenario. The Hessian matrix in *bi-objective* cases is treated in Section III-D. The hypervolume Hessian matrix is the Jacobian of the hypervolume gradient defined as follows:

$$\begin{aligned}\nabla^2\mathcal{H}_{\mathbf{F}}(\mathbf{X})&=\frac{\partial}{\partial\mathbf{X}}\left(\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial\mathbf{X}}\right)\\&=\left(\underbrace{\frac{\partial}{\partial\mathbf{X}}\left(\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial\mathbf{x}^{(1)}}\right)}_{\mu\cdot n\times n},\ldots,\frac{\partial}{\partial\mathbf{X}}\left(\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial\mathbf{x}^{(\mu)}}\right)\right)\\&=\begin{pmatrix}\frac{\partial}{\partial\mathbf{x}^{(1)}}\left(\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial\mathbf{x}^{(1)}}\right)&\cdots&\frac{\partial}{\partial\mathbf{x}^{(1)}}\left(\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial\mathbf{x}^{(\mu)}}\right)\\\vdots&\ddots&\vdots\\\frac{\partial}{\partial\mathbf{x}^{(\mu)}}\left(\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial\mathbf{x}^{(1)}}\right)&\cdots&\frac{\partial}{\partial\mathbf{x}^{(\mu)}}\left(\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial\mathbf{x}^{(\mu)}}\right)\end{pmatrix}\end{aligned}\tag{6}$$

where each *subgradient* is differentiated with respect to $\mathbf{X}$. This results in μ^2 block partitions ($n\times n$) of the Hessian matrix. The (i,j)-block matrix can be further expressed as follows:

$$\begin{aligned}\frac{\partial}{\partial\mathbf{x}^{(i)}}\left(\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial\mathbf{x}^{(j)}}\right)&=\frac{\partial}{\partial\mathbf{x}^{(i)}}\left(\frac{\partial\mathbf{y}^{(j)}}{\partial\mathbf{x}^{(j)}}\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial\mathbf{y}^{(j)}}\right)\\&=\sum_{k=1}^{m}\frac{\partial}{\partial\mathbf{x}^{(i)}}\left(\frac{\partial f_k\left(\mathbf{x}^{(j)}\right)}{\partial\mathbf{x}^{(j)}}\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial f_k\left(\mathbf{x}^{(j)}\right)}\right)\\&=\underbrace{\sum_{k=1}^{m}\frac{\partial}{\partial\mathbf{x}^{(i)}}\left(\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial f_k\left(\mathbf{x}^{(j)}\right)}\right)\nabla f_k\left(\mathbf{x}^{(j)}\right)^{\top}}_{\mathbf{A}_{ij}}\\&\quad+\underbrace{\sum_{k=1}^{m}\frac{\partial^2 f_k\left(\mathbf{x}^{(j)}\right)}{\partial\mathbf{x}^{(i)}\partial\mathbf{x}^{(j)}}\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial f_k\left(\mathbf{x}^{(j)}\right)}}_{\mathbf{B}_{ij}}.\end{aligned}$$

According to the differentiation above, each (i,j)-block matrix is a combination of two components: $\mathbf{A}_{ij}$ and $\mathbf{B}_{ij}$. Note that in matrix $\mathbf{A}_{ij}$, $[\partial/(\partial\mathbf{x}^{(i)})][([\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})]/(\partial f_k(\mathbf{x}^{(j)}))])$ is a column vector of size n and stands for the subgradient of $[(\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X}))/(\partial f_k(\mathbf{x}^{(j)}))]$ at $\mathbf{x}^{(j)}$. In the following, we abbreviate $f_k(\mathbf{x}^{(i)})$ as $f_k^{(i)}$ and its gradient $\nabla f_k(\mathbf{x}^{(i)})$ as $\nabla f_k^{(i)}$.

B. First Component Matrix: $\mathbf{A}_{ij}$

Matrix $\mathbf{A}_{ij}$ has size $n\times n$ and can be expressed as a sum of outer products

$$\mathbf{A}_{ij}=\sum_{k=1}^{m}\underbrace{\frac{\partial}{\partial\mathbf{x}^{(i)}}\left(\frac{\partial\mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial f_k^{(j)}}\right)\nabla f_k^{(j)}{}^{\top}}_{\text{rank 1}}.\tag{7}$$

Due to the fact that it is a sum of m outer products, this term has *at most* rank m. It is possible to make $\mathbf{A}_{ij}$ to have full rank only if $m \geq n$. In other cases, $\mathbf{A}_{ij}$ is always rank deficient ($\text{rank}(\mathbf{A}_{ij}) \leq m < n$). This indicates that in the "usual" multiobjective optimization case, where the number of objective functions is smaller than the number of decision variables, such a matrix $\mathbf{A}_{ij}$ is *always singular*.

In the following lemma, a detailed expression of $\mathbf{A}_{ij}$ is given for the bi-objective case ($m = 2$). Without loss of generality, we assume that the objective vectors (and corresponding decision vectors) are arranged according to the ascending order of the first objective values.

Lemma 1: Let $m = 2$ and all vectors $\mathbf{x}^{(i)}$, $i = 1 \ldots, \mu$, be mutually nondominated, the first component $\mathbf{A}_{ij}$ is nonzero only if the block matrix is located on the main diagonal ($i = j$) or the first diagonal above/below the main diagonal ($|i - j| = 1$), and it can be written as

$$\mathbf{A}_{ij} = \begin{cases} \nabla f_2^{(j)} \nabla f_1^{(j)\top} + \nabla f_1^{(j)} \nabla f_2^{(j)\top} & \text{if } i = j \\ -\nabla f_1^{(j+1)} \nabla f_2^{(j)\top} & \text{if } i = j+1 \\ -\nabla f_2^{(j-1)} \nabla f_1^{(j)\top} & \text{if } i = j-1 \\ \mathbf{0} & \text{otherwise.} \end{cases} \tag{8}$$

Proof: Assume a fixed reference point $\mathbf{r} = (r_1, r_2)^\top$. To simplify the formulation, we denote $f_1^{(\mu+1)} \coloneqq r_1$ and $f_2^{(0)} \coloneqq r_2$. The partial derivative of the hypervolume indicator with respect to the objective value is derived in [16], which corresponds to the length of the steps of the attainment curve

$$\frac{\partial \mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial f_1^{(j)}} = f_2^{(j)} - f_2^{(j-1)}, \quad \frac{\partial \mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial f_2^{(j)}} = f_1^{(j)} - f_1^{(j+1)}. \tag{9}$$

It is clear that $[(\partial \mathcal{H}_{\mathbf{F}}(\mathbf{X}))/(\partial f_1^{(j)})]$ is a function of only $\mathbf{x}^{(j)}$ and $\mathbf{x}^{(j-1)}$ (similar argument holds for $[(\partial \mathcal{H}_{\mathbf{F}}(\mathbf{X}))/(\partial f_2^{(j)})]$). The gradient of the partial derivatives can be given, for example: $[\partial/(\partial \mathbf{x}^{(j)})]([(\partial \mathcal{H}_{\mathbf{F}}(\mathbf{X}))/(\partial f_k^{(j)})]) = \nabla f_2^{(j)}$. Such a gradient is nonzero for at least one objective function, when $i = j$, $i = j+1$, or $i = j-1$. By substituting the required gradients into (7), the expression of $\mathbf{A}_{ij}$ can be obtained. ∎

C. Second Component Matrix: $\mathbf{B}_{ij}$

$\mathbf{B}_{ij}$ is a weighted sum of second-order derivatives of the objective functions, where the weights are partial derivatives of the hypervolume indicator at each objective value [see (5)]. Note that the second-order derivative $[(\partial^2 f_k^{(j)})/(\partial \mathbf{x}^{(i)} \partial \mathbf{x}^{(j)})]$ is not zero if and only if $i = j$

$$\mathbf{H}_k^{(j)} \coloneqq \frac{\partial^2 f_k^{(j)}}{\partial \mathbf{x}^{(j)2}}$$

is the Hessian matrix of objective function f_k at point $\mathbf{x}^{(j)}$. Consequently, matrix $\mathbf{B}_{ij}$ can be written as

$$\mathbf{B}_{ij} = \begin{cases} \sum_{k=1}^{m} \frac{\partial \mathcal{H}_{\mathbf{F}}(\mathbf{X})}{\partial f_k^{(j)}} \mathbf{H}_k^{(j)} & \text{if } i = j \\ \mathbf{0} & \text{if } i \neq j. \end{cases} \tag{10}$$

Note that $[(\partial \mathcal{H}_{\mathbf{F}}(\mathbf{X}))/(\partial f_k^{(j)})]$ can be obtained from (9). The singularity of matrix $\mathbf{B}_{ij}$ depends on the properties of the Hessian matrices of the objective functions. Under the assumption that all objective functions are convex (the objective-wise Hessian matrices are positive definite), matrix $\mathbf{B}_{ij}$ is also positive definite, under the condition that all the objective functions are subject to maximization (for minimization, $\mathbf{B}_{ij}$ is negative definite). In general, if each objective function has nonsingular Hessian matrix almost everywhere, it is obvious that the matrix $\mathbf{B}_{ij}$ is nonsingular.

D. Hypervolume Hessian in Bi-Objective Case

For a BOP, the hypervolume Hessian matrix has the following structure:

$$\nabla^2 \mathcal{H}_{\mathbf{F}}(\mathbf{X}) = \begin{pmatrix} \mathbf{D}_1 & \tilde{\mathbf{A}}_1 & & & \\ \tilde{\mathbf{A}}_1^\top & \mathbf{D}_2 & \tilde{\mathbf{A}}_2 & & \\ & \tilde{\mathbf{A}}_2^\top & \ddots & \ddots & \\ & & \ddots & \ddots & \tilde{\mathbf{A}}_{\mu-1} \\ & & & \tilde{\mathbf{A}}_{\mu-1}^\top & \mathbf{D}_\mu \end{pmatrix}$$

where $\mathbf{D}_i = \mathbf{A}_{ii} + \mathbf{B}_{ii}$ and $\tilde{\mathbf{A}}_i = \mathbf{A}_{i(i+1)} = -\nabla f_2^{(i)} \nabla f_1^{(i+1)\top}$ according to (8). Note that the Hessian matrix $\nabla^2 \mathcal{H}_{\mathbf{F}}(\mathbf{X})$ is a *tridiagonal block* matrix.

IV. Hypervolume Newton Method

After having stated the hypervolume gradient and Hessian matrix for a $\mu \cdot n$-vector $\mathbf{X}$ for a given MOP, we are now in the position to address the set-based Newton method for hypervolume maximization. For this, we will first consider the unconstrained case and later on discuss first attempts to treat constrained problems.

Given an unconstrained MOP and a population of μ individuals, the Newton step (or Newton function) is defined as follows:

$$\begin{aligned} &\mathbf{N} : \mathbb{R}^{\mu \cdot n} \to \mathbb{R}^{\mu \cdot n} \\ &\quad \mathbf{N}(\mathbf{X}) \coloneqq \mathbf{X} - \nabla^2 \mathcal{H}_F(\mathbf{X})^{-1} \nabla \mathcal{H}_F(\mathbf{X}). \end{aligned} \tag{11}$$

The Newton direction for the entire population is given by

$$\nu_{\mathbf{N}} \coloneqq -\nabla^2 \mathcal{H}_F(\mathbf{X})^{-1} \nabla \mathcal{H}_F(\mathbf{X}) \in \mathbb{R}^{\mu \cdot n} \tag{12}$$

and the according directions for each individual $\mathbf{x}^{(i)}$ of $\mathbf{X}$ are denoted by $\nu_{\mathbf{N}}^{(i)} \in \mathbb{R}^n$, $i = 1, \ldots, \mu$. Since the hypervolume indicator subgradient (4) for the strictly dominated subvector $\mathbf{x}^{(i)}$ of $\mathbf{X}$ is zero and also its corresponding Newton direction will be zero. Consequently, such a point will remain stationary when applying the set-based Newton method. In the following, we restrict the approximation set $\mathbf{X}$ to the set of mutually nondominated elements. For this purpose, we define the set $\tilde{\mathbf{X}}$ as the subset of $\mathbf{X}$ that contains all its nondominated elements. The HNM is thus defined as

$$\begin{aligned} &\mathbf{X}_0 \in \mathbb{R}^{\mu \cdot n} \\ &\mathbf{X}_{i+1} = \mathbf{N}\left(\tilde{\mathbf{X}}_i\right), \text{ for } \quad i = 0, 1, 2, \ldots \end{aligned} \tag{13}$$

The pseudocode for HNM is shown in Algorithm 1. For the step size control, we suggest to choose the initial step $\tilde{t}_0 = 1$ together with quadratic backtracking to satisfy the

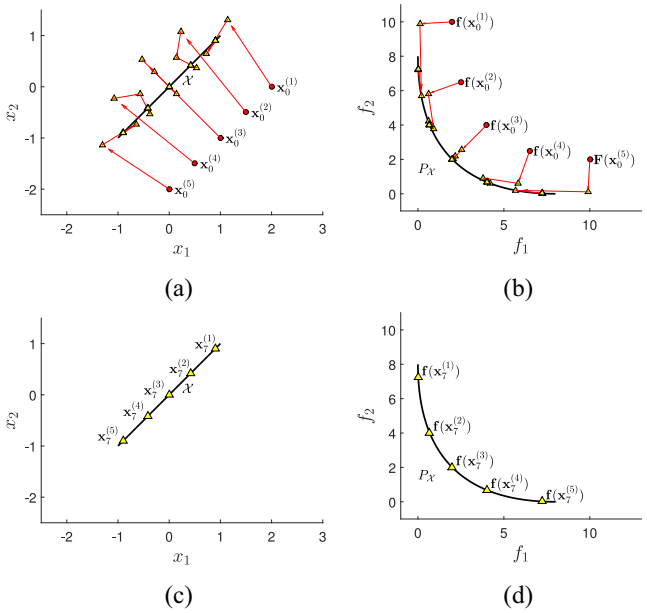


Fig. 1. Numerical result of HNM on MOP1 using (14) as initial population. (a) and (b) Iterations in decision and objective space. (c) and (d) Optimal solution and its image for $\mu = 5$ and $\mathbf{r} = (20, 20)^\top$.

Wolfe conditions [30] on the hypervolume. If automatic differentiation [30] is used to evaluate the (exact) gradient and the Hessian matrix at the iterate $\mathbf{X}_i$, the cost for each Newton step is given by $5\mu + (4 + 6n)\mu$ function evaluations.

Hence, if the hypervolume Hessian of the optimal set $\mathbf{X}^*$ is regular and the initial set $\mathbf{X}_0$ is close enough to $\mathbf{X}^*$, we can expect local quadratic convergence of the sequence of sets $\mathbf{X}_i$, $i \in \mathbf{N}$, toward $\mathbf{X}^*$ [30]. The investigation of the regularity of the Hessian at the optimal solution is subject of ongoing work.

Example 1: In order to demonstrate the performance of the HNM we consider the following BOP:

$$\begin{aligned} f_1 &= (x_1 - 1)^2 + (x_2 - 1)^2 \\ f_2 &= (x_1 + 1)^2 + (x_2 + 1)^2 \end{aligned} \qquad \text{(MOP1)}$$

where we choose a reference point $\mathbf{r} = (20, 20)^\top$.

1) We choose $\mu = 5$ and the initial population $\mathbf{X}_0$ as

$$\left\{\mathbf{x}_0^{(1)}, \mathbf{x}_0^{(2)}, \mathbf{x}_0^{(3)}, \mathbf{x}_0^{(4)}, \mathbf{x}_0^{(5)}\right\} = \left\{\begin{pmatrix}0\\-2\end{pmatrix}, \begin{pmatrix}0.5\\-1.5\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}, \begin{pmatrix}1.5\\-0.5\end{pmatrix}, \begin{pmatrix}2\\2\end{pmatrix}\right\}. \tag{14}$$

Fig. 1 shows the performance of HNM both in decision and objective space. As it can be seen, the iterations quickly approach the optimal solution for $\mu = 5$ and a given reference point. This observation is confirmed in Table I, where the hypervolume values, the norm of the gradients, and the error—measured in terms of the Hausdorff distance [21] of $\mathbf{X}_i$ and the optimal solution—are displayed for each iteration. The values indicate quadratic convergence.

2) Next, we consider the same setting but using as initial population

$$\left\{\mathbf{x}_0^{(1)}, \mathbf{x}_0^{(2)}, \mathbf{x}_0^{(3)}, \mathbf{x}_0^{(4)}, \mathbf{x}_0^{(5)}\right\}$$

TABLE I
HYPERVOLUME VALUES, ERROR, AND HYPERVOLUME GRADIENTS FOR THE APPLICATION OF HNM ON MOP1 USING (14) (SEE FIG. 1)

Iter	$\mathcal{H}$	Error (HNM)	$\|\|\nabla\mathcal{H}_{\mathbf{F}}\|\|$
0	306.500000000000	76.569536787325	48.826222462934
1	369.562245664015	13.507291123309	21.062836757428
2	379.065240846390	4.004295940935	13.997313947531
3	382.734095975048	0.335440812277	2.879962272550
4	383.068028174229	0.001508613095	0.199992009700
5	383.069536709027	0.000000078297	0.001265863184
6	383.069536787325	0.000000000000	0.000000101250
7	383.069536787325	0.000000000000	0.000000000000

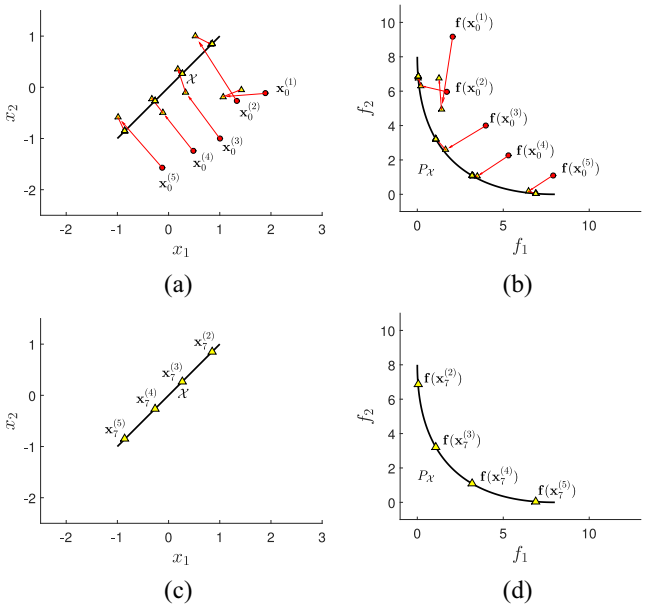


Fig. 2. Numerical result of HNM on MOP1 using (15) as initial population. (a) and (b) Iterations in decision and objective space. (c) and (d) Optimal solution and its image for $\mu = 4$ and $\mathbf{r} = (20, 20)^\top$.

$$= \left\{\begin{pmatrix}-0.12\\-1.57\end{pmatrix}, \begin{pmatrix}0.48\\-1.24\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}, \begin{pmatrix}1.32\\-0.26\end{pmatrix}, \begin{pmatrix}1.89\\-0.11\end{pmatrix}\right\}. \tag{15}$$

Fig. 2 and Table II show the numerical results of HNM. In step 2, $\mathbf{x}_2^{(1)}$ gets dominated by $\mathbf{x}_2^{(3)}$. The iteration thus continues with $\tilde{\mathbf{X}}_2 \in \mathbb{R}^{4\cdot 2}$, that is, with a set of 4 2-D vectors. HNM converges (again quadratically) toward the optimal hypervolume population, albeit for population size $\mu = 4$.

Next, we consider inequality constrained MOPs of the form

$$\begin{aligned} &\min_{\mathbf{x}\in\mathbb{R}^n} \ \mathbf{f}(\mathbf{x}) \\ &\text{s.t.} \ \ g_j(\mathbf{x}) \leq 0, \quad j = 1, \ldots, q. \end{aligned} \tag{16}$$

For the treatment of such problems we propose to utilize a penalization approach that transforms the original (constrained) MOP into an auxiliary unconstrained one. In the current context, the related unconstrained problem reads as

$$\min_{\mathbf{x}\in\mathbb{R}^n} \mathbf{f}(\mathbf{x}) + p(\mathbf{x})\mathbf{c} \tag{17}$$

TABLE II
HYPERVOLUME VALUES, ERROR, AND HYPERVOLUME GRADIENTS FOR THE APPLICATION OF HNM ON MOP1 USING (15) FOR $\mathbf{X}_0$ (COMPARE TO FIG. 2)

Iter	$\mathcal{H}$	Error (HNM)	$\|\|\nabla\mathcal{H}_\mathbf{F}\|\|$
$\mu = 5$			
0	321.548361341053	61.521175446272	52.900629002755
1	376.616155150423	6.453381636902	14.685530752574
$\mu = 4$			
2	373.544616989121	9.524919798204	2.013226644050
3	380.698229481420	2.371307305905	0.110423150875
4	380.698531798899	2.371004988425	0.000213267251
5	380.698531800725	2.371004986600	0.000000002021
6	380.698531800725	2.371004986600	0.000000000000
7	380.698531800725	2.371004986600	0.000000000000

Algorithm 1: HNM($\mathbf{X}_0, I_{\max}, \text{tol}_x$)

Data: initial point $\mathbf{X}_0 \in \mathbb{R}^{\mu\cdot n}$, maximal number of iterations I_{max}, tolerance $tol_x \in \mathbf{R}_+$.
Result: best found Newton iteration $\mathbf{X}_{(N)}$ according to $\mathcal{H}$.

```
for i = 0 to I_max do
    Compute X̃_i, ∇H_F(X̃_i), ∇²H_F(X̃_i)
    Compute step size t_i ∈ R_+
    X_{i+1} := X̃_i − t_i ∇²H_F(X̃_i)^{-1} ∇H_F(X̃_i)
    if ||∇H_F(X_{i+1})|| < tol_x then
        Return X_(N) := X_{i+1}
    end
end
Return X_(N) := X_{i+1}
```

where $\mathbf{c} = (c, \ldots, c)^\top \in \mathbb{R}^m$, $c > 0$ a given (large) constant, and

$$p(\mathbf{x}) := \sum_{j=1}^{q} \max\big(0, g_j(\mathbf{x})\big)^2 \tag{18}$$

the penalization function. To solve inequality constrained MOPs of the form (16), HNM is thus applied on the unconstrained problem (17) as described above. To avoid convergence toward spurious solutions, the value of c cannot remain fixed during the computations. Instead, we need a sequence $c_i > 0$ with $\lim_{i\to\infty} c_i = \infty$. In our computations, we have chosen $c_0 = 10$ and in each Newton step the value is increased by a factor of 10, i.e., $c_i = 10^{i+1}$.

Example 2: We reconsider the MOP from Example 1 but additionally impose the box constraints

$$x_i \in [-0.5, -0.25], \quad i = 1, 2. \tag{19}$$

We have chosen $\mathbf{r} = (20, 20)^T$ and $\mu = 5$ as before, and the initial population as

$$\left\{\mathbf{x}_0^{(1)}, \mathbf{x}_0^{(2)}, \mathbf{x}_0^{(3)}, \mathbf{x}_0^{(4)}, \mathbf{x}_0^{(5)}\right\} = \left\{\begin{pmatrix}-0.46\\-0.43\end{pmatrix}, \begin{pmatrix}-0.41\\-0.38\end{pmatrix}, \begin{pmatrix}-0.36\\-0.33\end{pmatrix}, \begin{pmatrix}-0.31\\-0.28\end{pmatrix}, \begin{pmatrix}-0.26\\-0.23\end{pmatrix}\right\}. \tag{20}$$

Fig. 3 shows a numerical result of HNM for this setting. Table III indicates that the iterations converges, but that the convergence speed is only linear, which is due to the penalization approach.

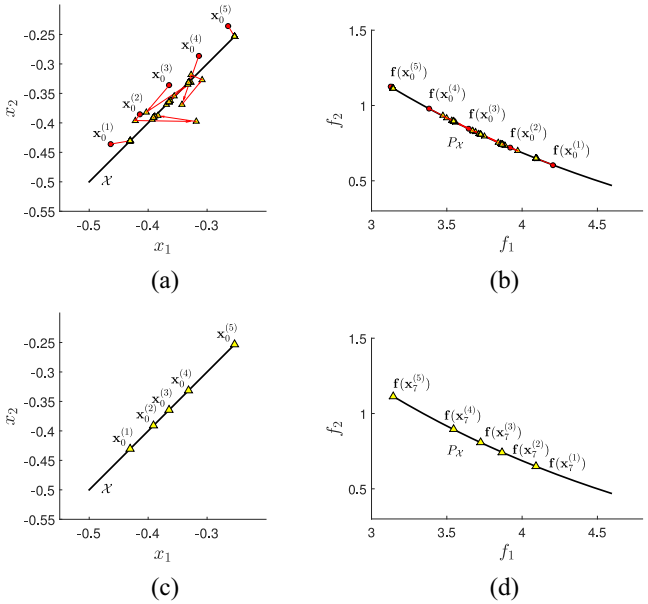


Fig. 3. Numerical result of HNM on MOP1 together with the box constraints (19) using (20) as initial population. (a) and (b) Iterations in decision and objective space. (c) and (d) Optimal solution and its image for $\mu = 5$ and $\mathbf{r} = (20, 20)^\top$.

TABLE III
HYPERVOLUME VALUES, ERROR, AND HYPERVOLUME GRADIENTS FOR THE APPLICATION OF HNM ON MOP1 WITH CONSTRAINT (19) USING (20) FOR $\mathbf{X}_0$ (COMPARE TO FIG. 3)

Iter	$\mathcal{H}$	Error (HNM)	$\|\|\nabla\mathcal{H}_\mathbf{F}\|\|$
1	325.901249212877	0.020663552304	1.212734648416
2	325.920707182928	0.001205582253	0.051243303186
3	325.921835652869	0.000077112312	0.004631644572
4	325.921906470408	0.000006294772	0.000422880099
5	325.921912146183	0.000000618997	0.000039564993
6	325.921912707233	0.000000057948	0.000003700931
7	325.921912759765	0.000000005415	0.000000346099

For equality constrained MOPs, one can proceed analogously to "classical" equality constrained SOPs via adapting the Lagrange multiplier method [30], the explicit description is left out here due to space limitations. More advanced constraint handling techniques are certainly interesting but beyond the scope of this paper.

Here, another important aspect to consider when using the HNM is the choice of a reference point $\mathbf{r}$. For the previous examples, we have chosen $\mathbf{r}$ much larger than the nadir point to consider all the possible elements by the HNM. Further, this is for adjusting the extreme points of the approximation. This strategy represents no problem for BOPs (see [31]), since $\mathbf{r}$ can be infinitely large, however, in the case of $m > 2$, this will not be appropriate anymore [32]. Different strategies have to be explored for the choosing of the best $\mathbf{r}$ when m goes up, but, for the moment, the choosing of a reference point greater than each greatest f_k in the final approximation is enough.

V. Memetic Strategy

By the above discussion, we can expect linear convergence for inequality constrained problems and is even quadratic convergence for unconstrained problems. The convergence, however, is only of local nature as it is the case for all Newton variants. Thus, a carefully chosen starting population is required or else the Newton iteration may fail to converge. We propose thus in the next step to hybridize the HNM with an evolutionary algorithm in order to obtain a fast and reliable memetic strategy. As an evolutionary algorithm we will use the SMS-EMOA which is a state-of-the-art EMOA that aims for hypervolume maximization [2]. Crucial for the effective realization of the hybrid is to decide when to switch from the evolutionary strategy to HNM. We suggest to apply the Newton method once: 1) the current population of SMS-EMOA only consists of mutually nondominated elements and 2) the difference of the hypervolume values of two consecutive populations is less than a given threshold. More precisely, the hybrid method we propose here, SMS-EMOA-HNM, works as follows (compare to Algorithm 2). First, an initial population with μ elements is chosen at random. In each iteration step (steps 4–11), a new offspring is generated and added to the current archive. Then, the new population is categorized into s subfronts, $G_1, \ldots, G_s$, according to the domination grade. Next, the hypervolume contribution is computed for each individual of the worst front G_s, and the individual with the least contribution is deleted from the archive (denoted as updated archive). This process is continued until the population consists of only one subfront and the difference of the hypervolume value of the current and the updated archive is less than or equal to a threshold tol_{HNM}. Finally, HNM is performed on the best-found population and the improved result is returned by the algorithm.

Algorithm 2: SMS-EMOA-HNM

Data: A reference point $\mathbf{r} \in \mathbb{R}^m$, a given tolerance $tol_{HNM} \in \mathbb{R}_+$ for SMS-EMOA, and a maximal number of iterations I_{max} and a tolerance $tol_x \in \mathbb{R}_+$ for HNM.
Result: Final population $X_{(F)}$.

1 Initialize a population X with μ elements at random.
2 $\mathcal{H}_u := \mathcal{H}(\mathbf{f}(X))$
3 **do**
4 $\quad \mathcal{H}_c := \mathcal{H}_u$ `// HV value of current archive`
5 $\quad$ Generate an offspring $\mathbf{x}' \in \mathbf{S}$ from X.
6 $\quad$ Set $X := X \cup \{\mathbf{x}'\}$.
7 $\quad$ Build a ranking $X = G_1 \cup G_2 \cup \ldots \cup G_s$ according the grade of dominance, where G_s denotes the worst sub-front.
8 $\quad$ Compute the hypervolume contribution for each $\mathbf{x} \in G_s$.
9 $\quad$ Denote by $\mathbf{x}^*$ the element with least hypervolume contribution in G_s.
10 $\quad X := X \setminus \{\mathbf{x}^*\}$
11 $\quad \mathcal{H}_u := \mathcal{H}(\mathbf{f}(X))$ `// HV value of updated archive`
12 **while** $|\mathcal{H}_u - \mathcal{H}_c| > tol_{HNM}$ *or* $s > 1$;
13 Construct a $\mu \cdot n$-vector $\mathbf{X}$ from X
14 $\tilde{\mathbf{X}} := HNM(\mathbf{X}, I_{max}, tol_x)$
15 Return $X_{(F)}$ constructed from $\tilde{\mathbf{X}}$

VI. Numerical Results

All of the experiments conducted through this section were performed on a computer with an Intel Core i7-2670QM 2.20 GHz CPU, 8 GB RAM, and Microsoft Windows 10 x64. As a first step, we will empirically evaluate the performance of the novel Newton method as well as the hybrid. To this end, we investigate HNM as a stand-alone algorithm. One natural question is the performance against the version presented in [18], where the Hessian has been approximated via finite differences (denoted by HNMA). Performance-wise, we have not observed a noticeable difference in our computations. This changes, however, when considering the computational times for both methods due to the bandwidth structure of the Hessian. As an example, Table IV shows the CPU times of the HNM and HNMA on MOP1 for two different population sizes. As anticipated, the CPU times are significantly less when computing the exact Hessian, for increasing values of μ and n. In our example, the CPU times differ by 2 orders of magnitude for $\mu = 100$, where we only have $n = 2$ decision variables.

TABLE IV
CPU Times (in Seconds) of HNMA and HNM for Different Population Sizes Using Seven Iterations Over 20 Experiments

μ	Description	Time (HNMA)	Time (HNM)
20	Mean (Std)	52.1808 (0.5563)	**4.2234** (0.0274)
100	Mean (Std)	3710.7411 (22.1876)	**13.7274** (0.1342)

Next, we evaluate the potential of the novel memetic strategy. As test problems, we selected unconstrained and constrained MOPs that have different properties, such as the shapes of the PF and the modalities of the objective functions. We select to test our algorithm over DTLZ [33], modified ZDT [34], and UF [35] test functions among others. The dimensions of the decision spaces vary from 2 to 30. For this comparison, we choose to use SMS-EMOA [2], the base algorithm of our memetic strategy; IBEA, where the selection mechanism is governed by the performance indicator [20]; NSGA-II, an algorithm based on dominance [4]; and finally, we compare against MOEA/D-DRA, an approach based on decomposition [6].

As design parameters, we have chosen $\text{tol}_{\text{HNM}} = 1 \times 10^{-3}$, $I_{\max} = 6$, $\text{tol}_x = 1 \times 10^{-10}$, and $\mu = 50$. As reference point, we have chosen $r = (11, 11)^\top$ for all problems and have set a budget of 50 000 function evaluations, however, the SMS-EMOA-HNM can stop before. All experiments have been repeated for 20 independent runs. The following comparison is made at the point that the SMS-EMOA-HNM stops. Table V shows the hypervolume values for all the five methods. In 17 of 21 test functions, SMS-EMOA-HNM achieves the best values. Moreover, in 11 of 17 test functions, where SMS-EMOA-HNM wins, it exists significant statistical differences [marked with an asterisk (*) in Table V]. Our approach

TABLE V
STATISTICAL RESULTS (MEAN AND STANDARD DEVIATION) ACCORDING TO HYPERVOLUME FOR THE FIVE ALGORITHMS. THE MEASURES WERE TAKEN WHEN THE SMS-EMOA-HNM STOPS TO USE THE SAME BUDGET OF FUNCTION EVALUATIONS

Problem	NSGA-II	MOEA/D-DRA	IBEA	SMS-EMOA	SMS+HNM
MOP 2 *	1519.8573	1521.4253	1489.2452	1524.5748	**1530.5184**
	0.8424	0.3942	5.2421	0.4520	0.4928
Fons. 1 *	15.0014	15.0064	14.9865	15.0113	**15.0616**
	0.0003	0.0002	0.0050	0.0034	0.0001
LSS *	15.3524	15.6347	15.2394	15.6420	**15.9472**
$\alpha = 1.5$	0.0034	0.0011	0.0845	0.0001	0.0001
ZDT1M *	109.5616	109.5591	109.5431	109.5650	**109.6572**
	0.0010	0.0028	0.0201	0.0021	0.0001
ZDT2M	109.2270	109.2245	109.2306	109.2264	**109.2340**
	0.0018	0.0056	0.0081	0.0002	0.0004
ZDT3M *	117.1980	117.6477	113.9760	117.6778	**117.7758**
	1.1432	0.0157	3.5800	0.0117	0.0042
ZDT4M *	109.3948	109.2322	100.4140	109.4266	**109.5530**
	0.0879	0.47442	1.1027	0.7433	0.01100
ZDT6M *	102.1341	103.0401	102.0334	102.1739	**103.8428**
	0.8281	0.5224	0.6328	0.6189	0.5643
DTLZ1 *	120.8701	120.8722	120.7640	120.8712	**120.8735**
	0.0004	0.0001	0.0146	0.00020	0.0.00011
DTLZ2	120.2037	120.2054	120.2035	120.2074	**120.2086**
	0.0006	0.00001	0.0003	0.0001	0.0001
DTLZ3 *	120.0483	119.1455	119.9853	119.8617	**120.2004**
	0.5180	1.8524	0.0186	0.7603	0.00422
DTLZ4	120.2038	120.2047	120.2035	120.2074	**120.2074**
	0.0005	0.0001	0.0002	0.0001	0.0001
DTLZ7	92.6981	93.34604	85.4129	94.6510	**94.6514**
	4.6433	3.9004	6.0292	0.3523	0.0023
LSS *	118.3043	118.3485	118.3042	118.3495	**118.4901**
$\alpha = 0.5$	0.0034	0.0042	0.0001	0.0068	0.0040
UF1 *	120.0554	120.2095	119.5118	119.8055	**120.2965**
	0.6714	0.8838	0.6193	0.6947	0.1866
UF2	120.4717	120.5766	120.1257	120.4244	**120.5767**
	0.2024	0.0696	0.4277	0.1922	0.0260
UF3 *	115.5531	**120.4099**	115.2628	115.3350	116.0815
	1.9915	0.1094	1.6808	1.5367	1.4376
UF4	119.7310	119.8712	119.7509	119.6860	**119.8728**
	0.1478	0.1658	0.1727	0.1252	0.1039
UF5 *	116.0595	116.0345	113.8333	**116.5355**	116.3029
	2.4378	4.1337	2.8913	1.8056	1.8536
UF6 *	**116.1578**	114.5945	114.3792	115.3433	115.1701
	2.1682	3.8867	2.6050	2.6149	2.7687
UF7 *	119.9853	**120.4549**	119.2574	120.1600	120.2817
	1.6355	0.0586	2.3080	0.2954	0.2555

seems to be effective for linear PSs even in higher dimension (DTLZ, ZDT, LSS, Fonseca 1, and MOP 2 test functions). However, its performance is diminished when it tackles more difficult PSs. See UF test functions results in Table V; our algorithm only wins in 3 of 7 cases. We notice that HNM depends on the starting population due to its local nature. Then, the performance of HNM is affected when the population produced by the global approach, in this particular case, the SMS-EMOA is not good like in UF3. In UF3, the SMS-EMOA-HNM outperforms SMS-EMOA, however, cannot reach the value obtained by MOEA/D-DRA. For nonlinear and disconnected PSs (see UF5 and UF6) the method is clearly not effective. The refinement of the solutions in these areas of the PS can be very difficult for our approach, since, a slight movement can affect the hypervolume value. Then, the method prefers to not move what causes that HNM gets stuck. By considering the geometry of the PFs, the HNM has worked well over convex (MOP2, ZDT1M, LSS, ZDT4M, and UF1), concave (Fonseca 1, DTLZ3, LSS, ZDT6M, and UF4), linear (DTLZ1), and disconnected fronts (ZDT3M). The performance of the HNM is invariant to the geometry of the PF, nevertheless, the method has difficulties in dealing with nonlinear and disconnected PSs. In most of the test functions our memetic approach, the SMS-EMOA-HNM, seems to be competitive and superior against the other approaches. It is important to mention that the HNM has not been designed for replacing evolutionary algorithms instead this paper shows that the method can be used for obtaining better approximations of the PF with higher hypervolume values almost near to the optimal. Finally, for other kind of problems, such as the imbalance MOPs proposed in [36] and [37], the method as it is will not perform well. This is since the final approximations produced by commonly used evolutionary algorithms (NSGA-II, MOEA/D, and SMS-EMOA) generally do not cover the whole PF. Due to these, algorithms aim more for convergence than diversity. Therefore, by using the HNM with a starting point produced by SMS-EMOA will not be good since the local nature of the Newton method only will allow to refine the solution set locally. However, Liu *et al.* [36] developed an approach called multiobjective to multiobjective that can be coupled with any MOEA. This strategy can be helpful to produce a good starting point for the HNM for tackling these kind of problems.

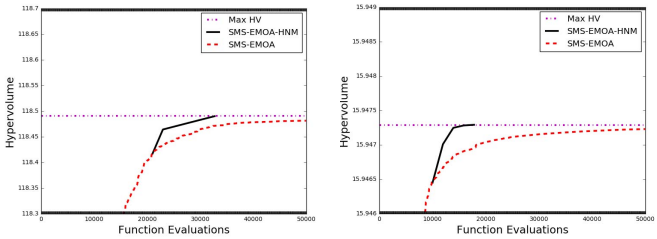


Fig. 4. Convergence plot comparisons over the whole execution of SMS-EMOA and SMS-EMOA-HNM for some test functions. (a) LSS $\alpha = 1.5$. (b) LSS $\alpha = 0.5$.

Considering the computational time expended to compute the local search, our algorithm use less time in most of the cases than the SMS-EMOA (see Table VI). This means that the switch between the global and the local approach decreases the computational time spent using the same number of function evaluations. However, NSGA-II, MOEA/D-DRA, and IBEA are the algorithms that perform much faster than the SMS-EMOA. Nevertheless, by selecting the one with the best performance, we could apply the HNM also to refine the solution and decrease the computational time. To make the comparison fair, we use the implementation of these algorithms provided by jMetal [38] using the default configuration.

Fig. 4 shows the convergence plots for SMS-EMOA against SMS-EMOA-HNM in some test problems. The black lines represent the means for every 100 function evaluations of SMS-EMOA-HNM, it is important to emphasize that our proposal stops before spending all the function evaluations in the budget. The red dotted lines represent the means for the competitor (SMS-EMOA) and the magenta dots represent

TABLE VI
STATISTICAL RESULTS (MEAN AND STANDARD DEVIATION) ACCORDING TO TIME (SECONDS) FOR ALL ALGORITHMS

Problem	NSGA-II	MOEA/D-DRA	IBEA	SMS-EMOA	SMS+HNM
MOP2	0.1848	**0.1383**	0.1762	0.7285	0.4492
	0.0615	0.0256	0.0520	0.0573	0.1353
Fons. 1	0.2132	**0.1337**	0.1877	0.6513	0.3527
	0.0021	0.0186	0.0555	0.0646	0.0432
LSS	0.1977	**0.1467**	0.1827	0.7707	0.4799
$\alpha = 1.5$	0.0400	0.0343	0.0412	0.0073	0.0342
ZDT1M	**0.1538**	0.1612	0.2103	0.9525	0.6665
	0.0656	0.0510	0.0675	0.0176	0.0244
ZDT2M	0.1967	**0.1589**	0.2077	0.9084	0.6089
	0.0051	0.0585	0.0081	0.0573	0.0134
ZDT3M	0.1848	**0.1769**	0.2080	0.8535	0.5728
	0.0391	0.0151	0.0643	0.0131	0.0525
ZDT4M	0.2132	**0.1800**	0.2170	0.8935	0.5941
	0.0153	0.0294	0.0052	0.0549	0.0353
ZDT6M	**0.1977**	0.2051	0.2194	0.8801	0.5822
	0.0155	0.0082	0.0633	0.0142	0.0242
DTLZ1	0.1538	**0.1484**	0.1908	0.8189	0.5211
	0.0147	0.0110	0.0199	0.0399	0.1243
DTLZ2	0.1692	**0.1127**	0.1771	0.9715	0.6946
	0.0335	0.0327	0.0630	0.0409	0.0243
DTLZ3	0.1733	**0.1337**	0.2059	0.8243	0.5349
	0.0168	0.0234	0.0366	0.0036	0.0243
DTLZ4	0.1661	**0.1477**	0.1824	0.7815	0.4836
	0.0498	0.0442	0.0033	0.0624	0.0353
DTLZ7	0.1814	**0.1485**	0.1881	0.9746	0.6855
	0.0645	0.0450	0.0413	0.0593	0.0564
LSS	0.1871	**0.1421**	0.1523	0.6509	0.3561
$\alpha = 0.5$	0.0335	0.0071	0.0017	0.0574	0.0434
UF1	0.2056	**0.1288**	0.2347	0.8443	0.5488
	0.0374	0.0118	0.0644	0.0382	0.0423
UF2	0.1862	**0.1774**	0.2043	0.9041	0.6084
	0.0393	0.0087	0.0643	0.0380	0.0733
UF3	0.1730	**0.1577**	0.1825	0.8341	0.5525
	0.0072	0.0147	0.0615	0.0047	0.1323
UF4	0.1733	**0.1401**	0.2039	0.9055	0.6151
	0.0092	0.0609	0.0672	0.0115	0.0543
UF5	**0.1427**	0.1605	0.1979	0.7890	0.4940
	0.0153	0.0504	0.0576	0.0465	0.0342
UF6	**0.1360**	0.1861	0.2043	0.7992	0.5105
	0.0660	0.0645	0.0303	0.0325	0.0234
UF7	0.1382	**0.1252**	0.2072	0.8993	0.6111
	0.0044	0.0528	0.0590	0.0247	0.0324

TABLE VII
Δ_2 VALUES BETWEEN THE FINAL APPROXIMATIONS AND THE PSS

Problem	NSGA-II	MOEA/D-DRA	IBEA	SMS-EMOA	SMS+HNM
MOP 2	0.0453	0.0436	0.0963	0.0436	**0.0027**
	0.0023	0.0034	0.0004	0.0035	0.0008
Fons. 1	0.0642	0.0599	0.0853	0.0524	**0.0503**
	0.0875	0.0064	0.0034	0.0098	0.0069
LSS	0.0911	0.0867	0.0982	0.0835	**0.0633**
$\alpha = 1.5$	0.0111	0.0087	0.0096	0.0099	0.0091
ZDT1M	0.1203	0.0995	0.1234	0.1131	**0.0923**
	0.0023	0.0089	0.0022	0.0036	0.0002
ZDT2M	0.2342	0.2124	0.2113	0.1934	**0.1289**
	0.0015	0.0032	0.0037	0.0025	0.0019
ZDT3M	0.1124	0.1024	0.1134	0.0998	**0.0978**
	0.0071	0.0049	0.0084	0.0044	0.0043
ZDT4M	0.5353	0.5633	1.4235	0.5734	**0.3101**
	0.0033	0.0073	0.0058	0.0044	0.0234
ZDT6M	5.32523	3.8567	4.0234	3.9424	**3.2423**
	0.0028	0.0046	0.0016	0.0038	0.06214
DTLZ1	0.0946	0.0834	0.0643	0.0345	**0.0103**
	0.0030	0.0015	0.0046	0.0071	0.0078
DTLZ2	0.0146	0.0104	0.0099	0.0092	**0.0047**
	0.0039	0.0042	0.0024	0.0073	0.0014
DTLZ3	0.1325	0.4356	0.6435	0.5242	**0.0987**
	0.0066	0.0031	0.0065	0.0088	0.2415
DTLZ4	0.8135	0.7342	0.8635	**0.7235**	0.8474
	0.0022	0.0018	0.0067	0.0081	0.4514
DTLZ7	1.0353	0.8635	5.3523	0.5463	**0.4071**
	0.0049	0.0023	0.0048	0.0041	0.0098
LSS	0.9355	0.7342	0.7354	0.83521	**0.2027**
$\alpha = 0.5$	0.0039	0.0010	0.0072	0.0067	0.0110
UF1	0.1024	0.1005	0.9422	0.1103	**0.0924**
	0.0030	0.0074	0.0028	0.0052	0.0097
UF2	0.0824	0.0722	0.9242	0.0835	**0.0624**
	0.0097	0.0095	0.0080	0.0023	0.0060
UF3	3.2532	**0.0653**	3.633	3.2453	2.3532
	0.0038	0.0029	0.0061	0.0094	0.0055
UF4	0.9474	0.8462	0.9943	0.9463	**0.7653**
	0.0043	0.0090	0.0041	0.0045	0.0042
UF5	1.3643	1.3533	5.2342	**1.1242**	1.2424
	0.0025	0.0080	0.0053	0.0091	0.0086
UF6	**0.8344**	1.3542	1.6453	1.2425	1.3053
	0.0067	0.0040	0.0074	0.0023	0.0018
UF7	1.2422	**0.8342**	1.1143	1.0242	0.9242
	0.0054	0.0032	0.0080	0.0032	0.0025

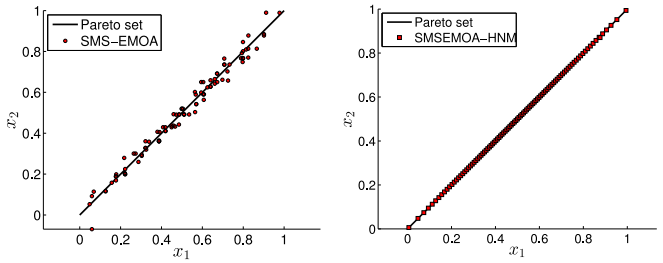


Fig. 5. Final approximation in parameter space obtained by (a) SMS-EMOA and (b) SMS-EMOA-HNM in the Lame Super Sphere problem with $n = 2$ and $\mu = 100$ after 20 000 function evaluations.

the maximal hypervolume values for each problem. We can observe that the new hybrid reaches the maximal hypervolume values. Further, even if all algorithms converge toward the maximum, the convergence is significantly faster for SMS-EMOA-HNM, saving about 20 000 function evaluations. Fig. 4 shows the advantage of using the HNM for the refinement of the final approximations produced by a global approach by saving function evaluations and reaching better hypervolume values. The above-mentioned superiority does not only hold when considering the hypervolume value but also the approximation quality of the PS. Fig. 5 shows a numerical result of SMS-EMOA-HNM and its base MOEA on the Lame Super Sphere problem with $n = 2$ for a budget of 20 000 function evaluations. The PS is the line segment connecting the points $(0, 0)^\top$ and $(1, 1)^\top$. Apparently, the approximation quality of SMS-EMOA-HNM is much better than the one from SMS-EMOA. Remarkably, the approximation quality of the latter cannot reach the other one even if the function evaluation budget is increased. This observation gets confirmed in Table VII, where the approximation qualities (measured via the averaged Hausdorff distance Δ_2 [21]) of the final

populations for all problems and algorithms are displayed. Then, SMS-EMOA-HNM significantly outperforms five other MOEAs on our chosen benchmark functions.

The proposed HNM only works for BOPs. Such problems have to be continuous and second-order differentiable. A problem for extending the applicability of the method for $m > 2$ is the time complexity of the computation of the hypervolume Hessian matrix. However, by using the same framework and by adapting a quasi-Newton method, we could tackle problems with $m > 2$. The adaptation of the HNM is left for future work.

VII. Conclusion

In this paper, we have propagated the use of a population-based Newton method for hypervolume maximization within specialized memetic strategies for the treatment of sufficiently smooth continuous MOPs. First of all, we have derived the hypervolume Hessian matrix for a given MOP and population analytically and have fully expressed the matrix for the case of a BOP. Further, we have investigated under which condition the matrix has full rank. Next, we have proposed the population-based HNM for hypervolume maximization for unconstrained MOPs and have made first attempts for the adequate treatment of constraints. For the proposed method, we could expect the following features:

1) local quadratic convergence in the unconstrained case;
2) set-based local search;
3) save function evaluations;
4) reach in most of the cases a near optimal $\mathcal{H}$ value;
5) invariant to different geometries of PFs;
6) not well performed over nonlinear and disconnect PSs;
7) necessity of a good starting point to avoid getting stuck;
8) necessity of the computation of the Hessian matrix.

Since the convergence holds only locally, we have proposed in the sequel SMS-EMOA-HNM, a hybrid of HNM and SMS-EMOA to obtain a fast and reliable solver for continuous BOPs. Numerical results and comparisons have finally shown the benefit of the novel memetic algorithm.

For future work, it would be desirable to develop and test more advanced constrained handling techniques which would increase the applicability of the method. Also, an adaptation of HNM for $m > 2$ (quasi-Newton method) and a study of the best choice of $\mathbf{r}$ will be pursued. Moreover, the singularity of the Hessian matrix will be inspected around the stationary point of the hypervolume indicator and the singularity conditions obtained in this paper should be generalized to the scenario when $\mu > 2$. Finally, the application of the novel hybrid to real-world problems would be of great interest.

References

[1] E. Dilettoso, S. A. Rizzo, and N. Salerno, "A weakly Pareto compliant quality indicator," *Math. Comput. Appl.*, vol. 22, no. 1, p. 25, 2017.

[2] N. Beume, B. Naujoks, and M. Emmerich, "SMS-EMOA: Multiobjective selection based on dominated hypervolume," *Eur. J. Oper. Res.*, vol. 181, no. 3, pp. 1653–1669, 2007.

[3] J. Bader and E. Zitzler, "HypE: An algorithm for fast hypervolume-based many-objective optimization," *Evol. Comput.*, vol. 19, no. 1, pp. 45–76, 2011.

[4] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002.

[5] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Trans. Evol. Comput.*, vol. 11, no. 6, pp. 712–731, Dec. 2007.

[6] W. Khan and Q. Zhang, "MOEA/D-DRA with two crossover operators," in *Proc. UKCI*, Sep. 2010, pp. 1–6.

[7] W. Fang, L. Zhang, S. Yang, J. Sun, and X. Wu, "A multiobjective evolutionary algorithm based on coordinate transformation," *IEEE Trans. Cybern.*, to be published.

[8] H. Xu, W. Zeng, D. Zhang, and X. Zeng, "MOEA/HD: A multiobjective evolutionary algorithm based on hierarchical decomposition," *IEEE Trans. Cybern.*, to be published.

[9] K. Deb, "Multi-objective optimisation using evolutionary algorithms: An introduction," in *Multi-Objective Evolutionary Optimisation for Product Design and Manufacturing*. London, U.K.: Springer, 2011, pp. 3–34.

[10] C. A. Coello Coello, G. B. Lamont, D. A. Van Veldhuizen, *Evolutionary Algorithms for Solving Multi-Objective Problems*, vol. 5. New York, NY, USA: Springer, 2007.

[11] K. Deb, S. Lele, and R. Datta, "A hybrid evolutionary multi-objective and SQP based procedure for constrained optimization," in *Proc. Symp. Intell. Comput. Appl.*, 2007, pp. 36–45.

[12] P. A. N. Bosman, "On gradients and hybrid evolutionary algorithms for real-valued multiobjective optimization," *IEEE Trans. Evol. Comput.*, vol. 16, no. 1, pp. 51–69, Feb. 2012.

[13] A. Lara, G. Sanchez, C. A. Coello Coello, and O. Schütze, "HCS: A new local search strategy for memetic multiobjective evolutionary algorithms," *IEEE Trans. Evol. Comput.*, vol. 14, no. 1, pp. 112–132, Feb. 2010.

[14] O. Schütze *et al.*, "The directed search method for multi-objective memetic algorithms," *Comput. Optim. Appl.*, vol. 63, no. 2, pp. 305–332, 2016.

[15] M. Emmerich, A. Deutz, and N. Beume, "Gradient-based/evolutionary relay hybrid for computing Pareto front approximations maximizing the S-metric," in *Hybrid Metaheuristics*, vol. 4771. Heidelberg, Germany: Springer, 2007, pp. 140–156.

[16] M. Emmerich and A. Deutz, "Time complexity and zeros of the hypervolume indicator gradient field," in *EVOLVE III*, O. Schütze *et al.*, Eds. Heidelberg, Germany: Springer Int., 2014, pp. 169–193.

[17] E. Zitzler and L. Thiele, *Multiobjective Optimization Using Evolutionary Algorithms—A Comparative Case Study*. Heidelberg, Germany: Springer, 1998, pp. 292–301.

[18] V. A. S. Hernández, O. Schütze, and M. Emmerich, "Hypervolume maximization via set based Newton's method," in *EVOLVE V* (Advances in Intelligent Systems and Computing), vol. 288, A. A. Tantar *et al.*, Eds. Cham, Switzerland: Springer, 2014, pp. 15–28.

[19] R. Baier, M. Dellnitz, M. H. von Molo, S. Sertl, and I. G. Kevrekidis, "The computation of convex invariant sets via Newton's method," *J. Comput. Dyn.*, vol. 1, no. 1, pp. 39–69, 2014.

[20] E. Zitzler and S. Künzli, "Indicator-based selection in multiobjective search," in *PPSN VIII*. Heidelberg, Germany: Springer, 2004, pp. 832–842.

[21] O. Schütze, X. Esquivel, A. Lara, and C. A. Coello Coello, "Using the averaged Hausdorff distance as a performance measure in evolutionary multiobjective optimization," *IEEE Trans. Evol. Comput.*, vol. 16, no. 4, pp. 504–522, Aug. 2012.

[22] M. P. Hanse and A. Jaszkiewicz, "Evaluating the quality of approximations to the non-dominated set," IMM, Dept. Math. Model., Tech. Univ. Denmark, Lyngby, Denmark, Tech. Rep., 1998.

[23] K. Gerstl, G. Rudolph, O. Schütze, and H. Trautmann, "Finding evenly spaced fronts for multiobjective control via averaging Hausdorff-measure," in *Proc. CCE*, 2011, pp. 1–6.

[24] H. Trautmann, T. Wagner, and D. Brockhoff, "R2-EMOA: Focused multiobjective search using R2-indicator-based selection," in *Proc. Learn. Intell. Optim. Conf. (LION)*, 2013, pp. 70–74.

[25] H. Wang, Y. Ren, A. Deutz, and M. Emmerich, "On steering dominated points in hypervolume indicator gradient ascent for bi-objective optimization," in *NEO*, O. Schütze *et al.*, Eds. Cham, Switzerland: Springer Int., 2017, pp. 175–203.

[26] J. Fliege and B. F. Svaiter, "Steepest descent methods for multicriteria optimization," *Math. Methods Oper. Res.*, vol. 51, no. 3, pp. 479–494, 2000.

[27] S. Schäffler, R. Schultz, and K. Weinzierl, "Stochastic method for the solution of unconstrained vector optimization problems," *J. Optim. Theory Appl.*, vol. 114, no. 1, pp. 209–222, 2002.

[28] P. Koch, O. Krämer, G. Rudolph, and N. Beume, "On the hybridization of SMS-EMOA and local search for continuous multiobjective optimization," in *Proc. GECCO*, 2009, pp. 603–610.
[29] H. Han, W. Lu, L. Zhang, and J. Qiao, "Adaptive gradient multiobjective particle swarm optimization," *IEEE Trans. Cybern.*, vol. 48, no. 11, pp. 3067–3079, Nov. 2018.
[30] S. Wright and J. Nocedal, *Numerical Optimization*. New York, NY, USA: Springer, 2006.
[31] A. Auger, J. Bader, D. Brockhoff, and E. Zitzler, "Hypervolume-based multiobjective optimization: Theoretical foundations and practical implications," *Theor. Comput. Sci.*, vol. 425, pp. 75–103, Mar. 2012.
[32] H. Ishibuchi, R. Imada, Y. Setoguchi, and Y. Nojima, "How to specify a reference point in hypervolume calculation for fair performance comparison," *Evol. Comput.*, vol. 26, no. 3, pp. 411–440, 2018.
[33] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, "Scalable test problems for evolutionary multiobjective optimization," in *Evolutionary Multiobjective Optimization*. London, U.K.: Springer, 2005, pp. 105–145.
[34] P. K. Shukla, "Gradient based stochastic mutation operators in evolutionary multi-objective optimization," in *Proc. ICANNGA*, 2007, pp. 58–66.
[35] Q. Zhang *et al.*, "Multiobjective optimization test instances for the CEC 2009 special session and competition," School Comput. Sci. Elect. Eng., Univ. Essex, Colchester, U.K., and School Elect. Electron. Eng., Nanyang Technol. Univ., Singapore, Rep. CES-487, 2008.
[36] H. L. Liu, L. Chen, K. Deb, and E. D. Goodman, "Investigating the effect of imbalance between convergence and diversity in evolutionary multiobjective algorithms," *IEEE Trans. Evol. Comput.*, vol. 21, no. 3, pp. 408–425, Jun. 2017.
[37] H. L. Liu, F. Gu, and Q. Zhang, "Decomposition of a multiobjective optimization problem into a number of simple multiobjective subproblems," *IEEE Trans. Evol. Comput.*, vol. 18, no. 3, pp. 450–455, Jun. 2014.
[38] J. J. Durillo, A. J. Nebro, and E. Alba, "The jMetal framework for multi-objective optimization: Design and architecture," in *Proc. Evol. Comput. (CEC)*, 2010, pp. 1–8.

Víctor Adrián Sosa Hernández received the B.Sc. degree in computer systems from the Instituto Politécnico Nacional, Mexico City, Mexico, in 2011 and the M.Sc. degree in computer science and the Ph.D. degree from CINVESTAV-IPN, Mexico City, in 2013 and 2017, respectively.

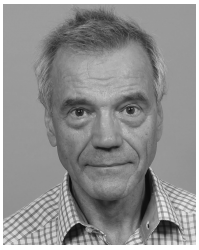
His current research interests include indicator-based local search techniques for multiobjective optimization, memetic strategies, and parameter-dependent multiobjective optimization.

Oliver Schütze received the Diploma degree in mathematics from the University of Bayreuth, Bayreuth, Germany, in 1999 and the Ph.D. degree in natural sciences from the University of Paderborn, Paderborn, Germany, in 2004.

He is currently a Professor with CINVESTAV-IPN, Mexico City, Mexico. His current research interest includes numerical and evolutionary optimization.

Hao Wang was born in Baoji, China, in 1989. He received the bachelor's degree in computer science from the Beijing Language and Culture University, Beijing, China, in 2011. He is currently pursuing the Ph.D. degree in computer science with the Leiden Institute of Advanced Computer Science, Leiden University, Leiden, The Netherlands.

His current research interests include algorithm improvement for evolutionary strategy as well as statistical machine learning for real-world data classification, regression, and prediction.



André Deutz received the first M.Sc. degree (*cum laude*) in pure mathematics from the University of Amsterdam, Amsterdam, The Netherlands, and the second M.Sc. degree in computer science from Cornell University, Ithaca, NY, USA.

He wrote a doctoral thesis on Algebraic Topology at Wayne State University, Detroit, MI, USA, and the University of California at Los Angeles, Los Angeles, CA, USA. Then, he was invited by the Mathematical Association of America to participate in a study of computer science for mathematicians. His current research interests include multiobjective optimization, and geometric algebra and its applications to multiobjective optimization and quantum computation.

Michael Emmerich received the Ph.D. degree in computer science from TU Dortmund, Dortmund, Germany, in 2005.

He is currently an Associate Professor with Leiden University, Leiden, The Netherlands, leading the Multidisciplinary Optimization and Decision Analysis Research Group. His current research interests include foundations of indicator-based multiobjective optimization and complex systems analysis, with applications in process engineering, drug discovery, and the biomedical sciences.