JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 114, No. 1, pp. 209–222, July 2002 (©2002)

# Stochastic Method for the Solution of Unconstrained Vector Optimization Problems

S. SCHÄFFLER,[1] R. SCHULTZ,[2] AND K. WEINZIERL[3]

Communicated by H. J. Pesch

**Abstract.** We propose a new stochastic algorithm for the solution of unconstrained vector optimization problems, which is based on a special class of stochastic differential equations. An efficient algorithm for the numerical solution of the stochastic differential equation is developed. Interesting properties of the algorithm enable the treatment of problems with a large number of variables. Numerical results are given.

**Key Words.** Vector optimization problems, curves of dominated points, Brownian motion, stochastic differential equations.

## 1. Introduction

In this paper, we investigate unconstrained vector optimization problems of the general form

$$\text{(VOP)} \quad \min_{x \in \mathbb{R}^n} \{ f(x) \},$$

where $f: \mathbb{R}^n \to \mathbb{R}^m$ is a twice continuously differentiable vector-valued function. Minimization of real-valued functions means that we look for a point $\hat{x} \in \mathbb{R}^n$ with

$$f(\hat{x}) \leq f(x), \qquad \text{for all } x \in \mathbb{R}^n.$$

Now, we have to clarify in which sense we want to minimize along vectors $f(x) \in \mathbb{R}^m$. Therefore, we define a partial order on $\mathbb{R}^m$.

[1]Professor of Mathematics, Universität der Bundeswehr, EIT 1, Neubiberg, Germany.
[2]Scientist, Corporate Technology, Siemens AG, München, Germany.
[3]Engineer, Automation and Control, Siemens AG, Erlangen, Germany.



0022-3239/02/0700-0209/0 

**Definition 1.1.** Let $u, v \in \mathbb{R}^m$. The vector $u$ is called less or equal to $v$, written $u \leq_p v$, if $u_i \leq v_i$ for all $i \in \{1, \ldots, n\}$. Equality holds if $u_i = v_i$ for all $i \in \{1, \ldots, n\}$.

For arbitrary $u, v, w, z \in \mathbb{R}^m$, we obtain the following properties:

(i) Reflexivity: $u \leq_p u$.
(ii) Transitivity: $u \leq_p v$, $v \leq_p w \Rightarrow u \leq_p w$.
(iii) Antisymmetry: $u \leq_p v$, $v \leq_p u \Rightarrow u = v$.
(iv) Compatibility with addition: $u \leq_p v$, $w \leq_p z \Rightarrow u + w \leq_p v + z$.
(v) Compatibility with scalar multiplication: $u \leq_p v$, $\alpha \in \mathbb{R}_0^+ \Rightarrow \alpha u \leq_p \alpha v$.

It is obvious that two arbitrary vectors $u, v \in \mathbb{R}^m$, $m > 1$, are not comparable with respect to $\leq_p$ in general. Using the partial order $\leq_p$ of $\mathbb{R}^m$, we are able to define an optimality concept for vector optimization problems.

**Definition 1.2.** Consider problem (VOP). A point $\hat{x} \in \mathbb{R}^n$ is called an efficient point or a Pareto-optimal solution if there exists no $w \in \mathbb{R}^n$ with $f(w) \neq f(\hat{x})$ and $f(w) \leq_p f(\hat{x})$.

If the above condition holds for only a neighborhood $U(\hat{x}) \subseteq \mathbb{R}^n$ of $\hat{x}$, then $\hat{x}$ is called locally efficient or a local Pareto-optimal solution. A point $x_1 \in \mathbb{R}^n$ is said to be dominated by a point $x_2 \in \mathbb{R}^n$ if

$$f(x_2) \neq f(x_1) \quad \text{and} \quad f(x_2) \leq_p f(x_1).$$

A large number of different approaches for the numerical solution of problem (VOP) is discussed in several papers. A special class of methods minimizes scalar-valued objective functions (weighted sums of objectives) in order to use classical optimization routines. Papers dealing with this approach and the problem of determining the scalar objective functions are Refs. 1–3 for instance. The method of the global criterion (see e.g., Refs. 4–5) tries to compute a solution as close as possible to a given reference point. In Ref. 6, a differential geometric analysis of the set of Pareto-optimal solutions is used for the development of a special homotopy method. Approaches based on evolutionary algorithms are summarized in Ref. 7. A population-based procedure with stochastic search is proposed in Ref. 8. Unfortunately, the excellent chances for the numerical solution of (VOP) using stochastic techniques are not exhausted at all in Ref. 8.

In this paper, we investigate a stochastic method for the numerical computation of Pareto-optimal solutions of the proposed unconstrained vector optimization problem. Since the Pareto-optimal solutions of (VOP) are not unique in general, we focus on the computation of either all or a

large number of these points. In Section 2, we consider the construction of a special curve of dominated points using a differential equation approach. This curve forms the drift part of a special stochastic differential equation which is introduced and analyzed in Section 3. The numerical treatment of this stochastic differential equation leads to an algorithm for the computation of Pareto-optimal solutions of (VOP). Finally, in Section 3, numerical examples are given.

## 2. Curve of Dominated Points

A large class of numerical methods for the unconstrained minimization of a twice continuously differentiable real-valued objective function $f: \mathbb{R}^n \to \mathbb{R}$ can be interpreted as numerical solutions of the following initial-value problem:

$$\text{(IVP)} \quad \dot{x}(t) = -\nabla f(x(t)), \qquad x(0) = x_0,$$

where $\nabla f(x)$ denotes the gradient of $f$ at $x$ and $x_0$ is chosen arbitrarily. The solution $x: [0, \infty[ \to \mathbb{R}^n$ of this initial-value problem consists of points with decreasing function values, which means that, if $\nabla f(x_0) \neq 0$, then

$$f(x(s)) > f(x(t)), \qquad \text{for all } 0 \leq s < t < \infty.$$

In this section, we want to generalize this approach to unconstrained problems (VOP) with objective function $f: \mathbb{R}^n \to \mathbb{R}^m$. Therefore, we are going to formulate on initial-value problem, whose unique solution is a curve $x: [0, \infty[ \to \mathbb{R}^n$ consisting of dominated points; i.e.,

$$f(x(s)) \geq_p f(x(t)) \text{ and } f(x(s)) \neq f(x(t)), \qquad \text{for all } 0 \leq s < t < \infty.$$

For the construction of our initial-value problem, we consider the following quadratic optimization problem for each $x \in \mathbb{R}^n$:

$$\text{(QOP}(x)) \quad \min_{\alpha \in \mathbb{R}^m} \left\{ \left\| \sum_{i=1}^{m} \alpha_i \nabla f_i(x) \right\|_2^2, \alpha_i \geq 0, i = 1, \ldots, m, \sum_{i=1}^{m} \alpha_i = 1 \right\},$$

where $\nabla f_i(x)$ denotes the gradient of the $i$th component of our objective function

$$f = (f_1, \ldots, f_m)^T, \qquad f_i: \mathbb{R}^n \to \mathbb{R}, \qquad i = 1, \ldots, m.$$

Now, we summarize some properties of problem (QOP($x$)) resulting from convex analysis:

(i) For each $x \in \mathbb{R}^n$, there exists a global minimizer $\hat{\alpha}$ of (QOP($x$)), which is not unique in general. Each local minimizer of (QOP($x$)) is a global minimizer.

(ii) Let $\hat{\alpha}$ and $\tilde{\alpha}$ be two global minimizers of (QOP($x$)) for fixed $x \in \mathbb{R}^n$. Then,

$$\sum_{i=1}^{m} \hat{\alpha}_i \nabla f_i(x) = \sum_{i=1}^{m} \tilde{\alpha}_i \nabla f_i(x).$$

Based on these properties, we define the function

$$q: \mathbb{R}^n \to \mathbb{R}^n, \qquad x \mapsto \sum_{i=1}^{m} \hat{\alpha}_i \nabla f_i(x),$$

where $\hat{\alpha}$ is a global minimizer of problem (QOP($x$)). We investigate the function $q$ and its relationship to problem (VOP) in the following theorem.

**Theorem 2.1.** Consider problem (QOP($x$)), and let $q: \mathbb{R}^n \to \mathbb{R}^n$ be the function given by

$$x \mapsto \sum_{i=1}^{m} \hat{\alpha}_i \nabla f_i(x),$$

where $\hat{\alpha}$ is a global minimizer of (QOP($x$)). Then, we obtain:

(i) Either $q(x) = 0$ or $-q(x)$ is a descent direction for all functions $f_1, \ldots, f_m$ at $x$; hence, each $x$ with $q(x) = 0$ fulfills the first-order necessary conditions for Pareto optimality.

(ii) For each $\hat{x} \in \mathbb{R}^n$, there exists a neighborhood $U(\hat{x})$ and a constant $L_{\hat{x}} \in \mathbb{R}_0^+$ such that

$$\|q(x) - q(y)\|_2 \leq L_{\hat{x}} \|x - y\|_2, \qquad \text{for all } x, y \in U(\hat{x}).$$

**Proof.**

(i) Define

$$K(x) = \left\{ \sum_{i=1}^{m} \alpha_i \nabla f_i(x), \alpha_i \geq 0, i = 1, \ldots, n, \sum_{i=1}^{m} \alpha_i = 1 \right\},$$

and assume that $0 \notin K(x)$ for any fixed $x \in \mathbb{R}^n$. Furthermore, assume that there exists a vector $v(x) \in K(x)$ with $q(x)^T v(x) \leq 0$. Then, we obtain the following properties of the vectors $\lambda(q(x) - v(x))$, $0 \leq \lambda \leq 1$:

(a) $(v(x) + \lambda(q(x) - v(x))) \in K(x)$, for all $0 \leq \lambda \leq 1$;

(b) $q(x)^T(\lambda(q(x) - v(x))) > 0$, for all $0 < \lambda \leq 1$.

Let $\hat{\lambda}$ be the global minimizer of the quadratic programming problem

$$\min_{0 \leq \lambda \leq 1} \{\|v(x) + \lambda(q(x) - v(x))\|_2^2\};$$

then, it is obvious that

$$\|v(x) + \hat{\lambda}(q(x) - v(x))\|_2^2 < \|q(x)\|_2^2,$$

because

$$\hat{\lambda} = 1 \Leftrightarrow q(x)^T(\lambda(q(x) - v(x))) \leq 0, \qquad \text{for all } 0 < \lambda \leq 1.$$

Since

$$v(x) + \hat{\lambda}(q(x) - v(x)) \in K(x),$$

we obtain a contradiction to the definition of $q$. Hence,

$$v(x)^T q(x) > 0, \qquad \text{for all } v(x) \in K(x).$$

(ii) Consider the following system of nonlinear equations with inequalities in $(\alpha(x), \xi(x), \lambda(x), \mu(x)) \in \mathbb{R}^{m+n+1+m}$, where $e_i$ denotes the $i$th unit vector:

$$(\nabla f_1(x), \ldots, \nabla f_m(x))\alpha(x) - \xi(x) = 0, \tag{1a}$$

$$(\nabla f_1(x), \ldots, \nabla f_m(x))^T \xi(x) - \lambda(x) \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} + \sum_{i=1}^{m} \mu_i(x) e_i = 0, \tag{1b}$$

$$\sum_{i=1}^{m} \alpha_i(x) - 1 = 0, \tag{1c}$$

$$\mu_i(x)\alpha_i(x) = 0, \qquad i = 1, \ldots, m, \tag{1d}$$

$$\mu_i(x) \leq 0, \qquad i = 1, \ldots, m, \tag{1e}$$

$$\lambda(x) \geq 0. \tag{1f}$$

The system (1) represents the first-order necessary and sufficient optimality conditions for global minimizers of (QOP($x$)). Let $(\hat{\alpha}(x), q(x), \hat{\lambda}(x), \hat{\mu}(x))$ be a solution of (1) for fixed $x \in \mathbb{R}^n$; then, we obtain:

(a) $q(x)$ is unique [cf. Theorem 2.1(i)];
(b) $\hat{\lambda}(x)$ and $\hat{\mu}(x)$ are unique.

Let $\{x_i\}_{i \in \mathbb{N}}$ be a sequence of vectors $x_i \in \mathbb{R}^n$ which converge to a point $\bar{x} \in \mathbb{R}^n$. Then, the sequences $\{q(x_i)\}_{i \in \mathbb{N}}$ and $\{\hat{\alpha}(x_i)\}_{i \in \mathbb{N}}$ are bounded and there exist convergent subsequences $\{q(x_j)\}$ with limit $\tilde{q}$ and $\{\hat{\alpha}(x_j)\}$ with limit $\tilde{\alpha}$. Therefore, we obtain a vector $(\tilde{\alpha}, \tilde{q}, \tilde{\lambda}, \tilde{\mu})$ that solves (1) at $x = \bar{x}$. Because of Theorem 2.1(i), $\tilde{q}$ is equal to $q(\bar{x})$, and $q$, $\tilde{\lambda}$, $\tilde{\mu}$ are continuous functions. If $\hat{\alpha}_i(\bar{x})$ is unique and greater then zero for all $i \in \{1, \ldots, m\}$ with $\hat{\mu}_i(\bar{x}) = 0$, then $q$ is continuously differentiable in a neighborhood of $\bar{x}$.

Otherwise, there exists a finite number of points $x_1, \ldots, x_k$ and closed neighborhoods $U(x_1), \ldots, U(x_k)$ of these points such that

(A) $x_i$ is an inner point of $U(x_i)$ for $i = 1, \ldots, k$;
(B) $\bar{x} \in U(x_i)$ for $i = 1, \ldots, k$;
(C) $\bar{x}$ is an inner point of $U(x_1) \cup \cdots \cup U(x_k)$;
(D) the function $q$ restricted on $U(x_i)$ is for all $i = 1, \ldots, k$ a continuously differentiable rational function in some components of the first-order derivatives of the objective function $f: \mathbb{R}^n \to \mathbb{R}^m$ of (VOP); see the system (1).

Hence, $q$ is locally Lipschitzian, because $f \in C^2$. □

The properties of the function $q$ enable us to generalize the initial-value problem (IVP) for scalar optimization problems to the following initial-value problem for vector optimization problems:

$$\text{(IVPV)} \quad \dot{x}(t) = -q(x(t)), \qquad x(0) = x_0,$$

where $q: \mathbb{R}^n \to \mathbb{R}^n$ is defined above.

**Theorem 2.2.** Consider the vector optimization problem (VOP) and the corresponding initial value problem (IVPV) with $q(x_0) \neq 0$. Define

$$R_{\leq_p} := \{x \in \mathbb{R}^n; f(x) \leq_p f(x_0)\},$$

and assume that $R_{\leq_p}$ is bounded. Then, we obtain the existence and uniqueness of a solution $x: [0, \infty[ \to \mathbb{R}^n$ of (IVPV) with

$$f(x(s)) \geq_p f(x(t)) \text{ and } f(x(s)) \neq f(x(t)), \qquad \text{for all } 0 \leq s < t < \infty.$$

**Proof.** Since $q$ is locally Lipschitzian (see Theorem 2.1), there exist a real number $T > 0$ and a unique solution $x: [0, T[ \to \mathbb{R}^n$ of (IVPV). Using Theorem 2.1, for all $i \in \{1, \ldots, m\}$ we get

$$\dot{f}_i(x(t)) = \nabla f_i(x(t))^T \dot{x}(t) = -\nabla f_i(x(t))^T q(x(t)) < 0,$$

for all $t \in [0, T[$. Therefore, $f_i(x(\cdot)): [0, T[ \to \mathbb{R}$ is a strictly monotone decreasing function for each $i \in \{1, \ldots, m\}$; hence,

$$f(x(s)) \geq_p f(x(t)) \text{ and } f(x(s)) \neq f(x(t)), \qquad \text{for all } 0 \leq s < t < T.$$

Now, we assume that $T$ is the largest real number such that $x: [0, T[ \to \mathbb{R}^n$ is a solution of (IVPV). Since $R_{\leq_p}$ is bounded and since $x(t) \in R_{\leq_p}$ for all $t \in [0, T[$, the solution $x$ of (IVPV) can be extended continuously to $x(T)$ at $t = T$ with $q(x(T)) = 0$. For the initial-value problem

$$\dot{y}(t) = q(y(t)), \qquad y(0) = x(T),$$

we know two solutions:

$$y(t) \equiv x(T) \text{ and } y(t) = x(T-t), \qquad \text{for all } t \in [0, T[.$$

This is a contradiction to Theorem 2.1(ii); therefore the solution $x$ of (IVPV) is defined on $[0, \infty[$. □

A numerical treatment of (IVPV) using explicit numerical schemes (because $q$ is not continuously differentiable) leads to points $x^* \in \mathbb{R}^n$, where the first-order optimality conditions for Pareto-optimal solutions are fulfilled. The fact that

$$f(x(s)) \geq_p f(x(t)) \text{ and } f(x(s)) \neq f(x(t)), \qquad \text{for all } 0 \leq s < t < \infty$$

has to be considered for a suitable stepsize control. The initial-value problem (IVPV) can be used for the computation of a single candidate for a Pareto-optimal solution. The application of a special stochastic perturbation to (IVPV), which we consider in Section 3, leads to a method for the numerical computation of either all or a large number of Pareto-optimal solutions.

## 3. Stochastic Approach to Vector Optimization

Let $\Omega$ be the set of all continuous functions $v: [0, \infty[ \to \mathbb{R}^n$, $n \in \mathbb{N}$. We define a metric $d$ on $\Omega$ by

$$d(\omega_1, \omega_2) := \sum_{m=1}^{\infty} (1/2^m) \min\left[ \max_{0 \leq t \leq m} \|\omega_1(t) - \omega_2(t)\|_2, 1 \right].$$

By $\mathcal{B}(\Omega)$, we mean the smallest $\sigma$-field containing all the open sets of $\Omega$ in the topology defined by the metric $d$. The open sets in $\mathbb{R}^p$, $p \in \mathbb{N}$, are given by the Euclidean norm. Throughout this paper, we use only the probability spaces $(\Omega, \mathcal{B}(\Omega), \mathcal{W})$, where $\mathcal{W}$ denotes an $n$-dimensional Wiener measure, which we consider later on. Let $\bar{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ be the compactification of $\mathbb{R}$ with $0 \cdot (\infty) = 0 \cdot (-\infty) = (\infty) \cdot 0 = (-\infty) \cdot 0 := 0$, and let $\mathcal{B}(\bar{\mathbb{R}})$ be the Borel $\sigma$-field of $\bar{\mathbb{R}}$, given by

$$B \in \mathcal{B}(\bar{\mathbb{R}}) \Leftrightarrow (B \cap \mathbb{R}) \in \mathcal{B}(\mathbb{R}).$$

**Definition 3.1.** A function $g: \Omega \to \bar{\mathbb{R}}$ is called a numerical function.

**Definition 3.2.** A collection $\{X_t\}$ of $n$-dimensional real random variables $X_t$, $t \geq 0$, defined on $\Omega$ is called a stochastic process.

**Definition 3.3.** Let $\{X_t\}$ be a stochastic process. For each fixed $\omega \in \Omega$, the function $X_\omega\colon [0, \infty[ \to \mathbb{R}^n$, $t \mapsto X_t(\omega)$ is called a path of $\{X_t\}$.

**Definition 3.4.** Let $\{X_t\}$ be a stochastic process. $\{X_t\}$ is called continuous if each path of $\{X_t\}$ is a continuous function.

For each $n \in \mathbb{N}$, our Wiener measure $\mathscr{W}$ is determined uniquely by the following conditions using the stochastic process $\{B_t\}$ defined by $B_t(\omega) := \omega(t)$ for all $t > 0$ (a proof is given in Ref. 9):

(i) $B_0(\omega) = 0$, $\mathscr{W}$-almost surely.

(ii) For every $0 = t_0 < t_1 < \cdots < t_m$, $m \in \mathbb{N}$, the random variables $B_0$, $B_{t_1} - B_{t_0}, \ldots, B_{t_m} - B_{t_{m-1}}$ are stochastically independent.

(iii) For every $0 \leq s < t$, the random variable $B_t - B_s$ is $N(0, (t-s)I_n)$ Gaussian distributed, where $I_n$ denotes the $n$-dimensional identity matrix.

Now, we are able to define the Brownian motion process.

**Definition 3.5.** The stochastic process $\{B_t\}$ defined by $B_t(\omega) := \omega(t)$ is called $n$-dimensional Brownian motion.

The following definition is very important for the investigation of some properties of stochastic processes.

**Definition 3.6.** A $\mathscr{B}(\Omega) - \mathscr{B}(\bar{\mathbb{R}})$-measurable numerical function $g$ is called a random time.

An important class of random times is give by the following theorem, which we restate without proof (see e.g. Ref. 10).

**Theorem 3.1.** Let $\{X_t\}$ be a continuous stochastic process. Then, the function $g\colon \Omega \to \bar{\mathbb{R}}$,

$$\omega \to \begin{cases} \inf\{t \geq 0; X_\omega(t) \in A\}, & \text{if } \bigcup_{t \geq 0}(X_\omega(t) \cap A) \neq \varnothing, \\ \infty, & \text{otherwise,} \end{cases}$$

is a random time for each open or closed set $A \in \mathscr{B}(\mathbb{R}^n)$, $n \in \mathbb{N}$.

The function value $g(\omega)$ of $g$ at $\omega$ indicates the shortest time at which the path $X_\omega$ of $\{X_t\}$ hits the Borel set $A$.

Now, we consider a stochastic method for the solution of the vector optimization problem (VOP). The following assumption concerning the function $q$ is fundamental for further investigations.

**Assumption A.** There exists an $\epsilon > 0$ such that

$$x^\top q(x) \geq [(1 + n\epsilon^2)/2] \max(1, \|q(x)\|_2),$$

for all $x \in \mathbb{R}^n \setminus \{x \in \mathbb{R}^n; \|x\|_2 \leq r\}$ with some $r > 0$, where $q$ is defined in Section 2.

Assumption A describes a special behavior of $q$ and therefore of $f$ outside a ball with radius $r$ for which only the existence is postulated. It is obvious that each function $q$ for which Assumption A is fulfilled results from a vector optimization problem which has at least one Pareto-optimal solution.

For the solution of (VOP), we investigate the following class of stochastic Itô differential equations, where $\{B_t\}$ is an $n$-dimensional Brownian motion, $\omega \in \Omega$, and $q$ is the above defined function:

$$\text{(SDE)} \quad dX_t = -q(X_t)\, dt + \epsilon\, dB_t, \qquad X_0 = x_0,$$

with $\epsilon > 0$, $x_0 \in \mathbb{R}^n$. We study the existence, uniqueness, and regularity of the solution of (SDE).

**Theorem 3.2.** Consider (SDE). For all $x_0 \in \mathbb{R}^n$ and for all $\epsilon$ for which Assumption A is fulfilled, we obtain the following:

(i) there exists a unique stochastic process $\{X_t\}$ that solves (SDE);
(ii) all paths of $\{X_t\}$ are continuous;
(iii) $X_0 \equiv x_0$.

Basically, the proof of this theorem is identical to what has been proven before (see Ref. 11) in context of classical optimization problems.

Let $\bar{x} \in \mathbb{R}^n$, and consider

$$S(\bar{x}, p) := \{x \in \mathbb{R}^n \mid \|x - \bar{x}\|_2 \leq p\}, p > 0.$$

For the investigation of the relations between the solutions of (SDE) and (VOP), we need the random times $s_{\bar{x},p}: \Omega \to \mathbb{R}^n$,

$$\omega \mapsto \begin{cases} \inf\{t \geq 0 \mid \|X_t(\omega) - \bar{x}\|_2 \leq p\}, & \text{if } \{t \geq 0 \mid \|X_t(\omega) - \bar{x}\|_2 \leq p\} \neq \emptyset, \\ \infty, & \text{otherwise.} \end{cases}$$

The following theorem is a direct application of some important results from the Lyapunov stability calculus for stochastic differential equations (see Ref. 12).

**Theorem 3.3.** Consider (SDE) with $\epsilon > 0$ such that Assumption A is fulfilled. Then, we obtain the following results for each $x_0 \in \mathbb{R}^n$ and for each

Pareto optimal solution $\bar{x} \in \mathbb{R}^n$, $p > 0$:

(i) $\mathscr{W}(\{\omega \in \Omega |_{S_{\bar{x},p}}(\omega) < \infty\}) = 1$.
(ii) $\mathbb{E}(s_{\bar{x},p}) < \infty$, where $\mathbb{E}(\cdot)$ denotes expectation.
(iii) The stochastic process $\{X_t\}$ defined in Theorem 3.2 converges in distribution to a random variable $X: \Omega \to \mathbb{R}^n$ with $\mathbb{E}(q(X)) = 0$.

From Theorem 3.3, we know that, for each Pareto-optimal solution $\bar{x}$, $\mathscr{W}$-almost all paths of $\{X_t\}$ hit any ball $S(\bar{x}, p)$ centered at $\bar{x}$ for arbitrary chosen $p > 0$ after a finite time for all $x_0 \in \mathbb{R}^n$. Moreover, the expectation of $s_{\bar{x},p}$ is finite. This is an important analytic property of $\{X_t\}$. Now, we investigate the question of how to choose $\epsilon$. For that, we consider once more

$$\text{(SDE)} \quad dX_t = -q(X_t)\,dt + \epsilon\,dB_t,\ X_0 = x_0.$$

The parameter $\epsilon$ is a measure of the balance between the curve of dominated points,

$$X_t^{x_0} = x_0 - \int_0^t q(X_\tau^{x_0})\,d\tau, \tag{2}$$

and a random search with realizations of Gaussian distributed random vectors with increasing variance,

$$X_t^{x_0}(\omega) = x_0 + B_t(\omega) - B_0(\omega). \tag{3}$$

If we choose $\epsilon$ for a fixed $\omega$ such that (2) dominates, then the chosen path of $\{X_t^{x_0}\}$ spends a long time close to any (local) Pareto-optimal solution of (VOP). If we choose $\epsilon$ such that the random search (3) dominates, then the Pareto-optimal solutions of (VOP) play no significant role along this path of $\{X_t^{x_0}\}$.

The optimal balance between (2) and (3), therefore the choice of $\epsilon$, depends on the objective function $f$ and the scale used. If one observes during the numerical computations of a path of $\{X_t\}$ that this path spends a very long time close to any (local) Pareto-optimal solution of (VOP), then $\epsilon$ is too small. If the (local) Pareto-optimal solutions of (VOP) play no significant role along this path, then $\epsilon$ is too large.

For the numerical computation of a path of $\{X_t\}$, we consider the following iteration scheme, which results from a standard approach in the numerical analysis of ordinary differential equations (Euler method). For a fixed stepsize $\sigma$, set

$$x_{j+1}^1 := x_j - \sigma q(x_j) - \epsilon n_3(\sigma/2)^{1/2}, \tag{4a}$$

$$x(\sigma/2) := x_j - (\sigma/2) q(x_j) - \epsilon n_1(\sigma/2)^{1/2}, \tag{4b}$$

$$x_{j+1}^2 := x(\sigma/2) - (\sigma/2) q(x(\sigma/2)) - \epsilon n_2(\sigma/2)^{1/2}, \tag{4c}$$

where $n_1$ and $n_2$ are realizations of the independent $N(0, I_n)$ normally distributed random vectors, which are computed by pseudorandom numbers, and where

$$n_3 = n_1 + n_2.$$

For a fixed positive $\delta$, we take

$$x_{j+1} = x^2_{j+1}, \qquad \text{if } \|x^1_{j+1} - x^2_{j+1}\|_2 \leq \delta;$$

otherwise, Steps (4a) and (4c) have to be repeated with $\sigma/2$ instead of $\sigma$.

## 4. Numerical Results

Let us consider a product which consists of several components. The $j$th component has a certain failure rate which depends on the additional reliability cost $x_j$. The failure rate of this component may be described by a function

$$p_j: \mathbb{R}^+_0 \to \mathbb{R}^+.$$

The product fails whenever any component fails. For our numerical example, we consider 20 components with

$$p_1(x_1) = 0.01 \exp(-(x_1/20)^{2.5}),$$
$$p_2(x_2) = 0.01 \exp(-(x_2/20)^{2.5}),$$
$$p_j(x_j) = 0.01 \exp(-x_j/15), \qquad j = 3, \ldots, 20.$$

Our objective function $f: (\mathbb{R}^+_0)^{20} \to \mathbb{R}^2$ of (VOP) is given by the sum of the additional costs

$$f_1: (\mathbb{R}^+_0)^{20} \to \mathbb{R}, \qquad x \mapsto \sum_{j=1}^{20} x_j$$

and the total failure rate

$$f_2: (\mathbb{R}^+_0)^{20} \to \mathbb{R}, \qquad x \mapsto 1 - \sum_{j=1}^{20} (1 - p_j(x_j)).$$

With the starting point $x_0 = (20, \ldots, 20)^T$, we have calculated 5000 steps using the above algorithm for the numerical solution of (SDE) with $\epsilon = 0.5$ and $\delta = 0.1$. For the sake of transparency, we have focused our numerical analysis on the interval $x \in [0, 40]^{20}$. This can be done using projection techniques (see Figs. 1–4) or a penalty approach.

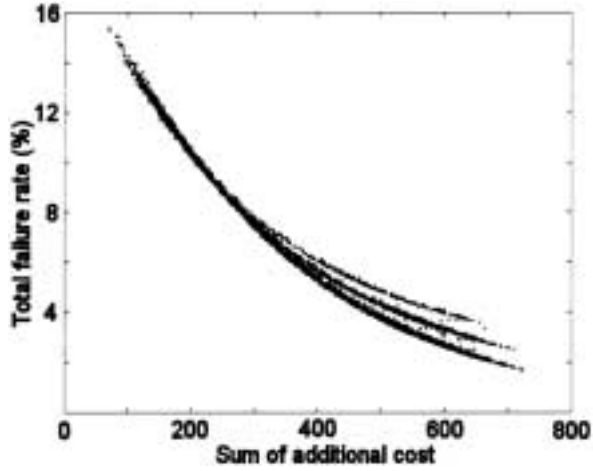


Fig. 1. Function values.

In Fig. 1, the function values of the components of the objective function evaluated at the 5000 calculated points are visualized.

In Fig. 2, the correlation between the component costs $x_1$ and $x_2$ according to the 5000 calculated points is visualized and these points are

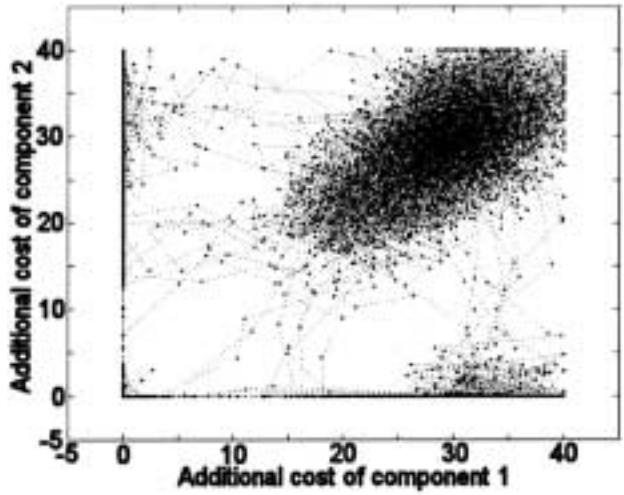


Fig. 2. Correlation between the cost components $x_1$ and $x_2$.

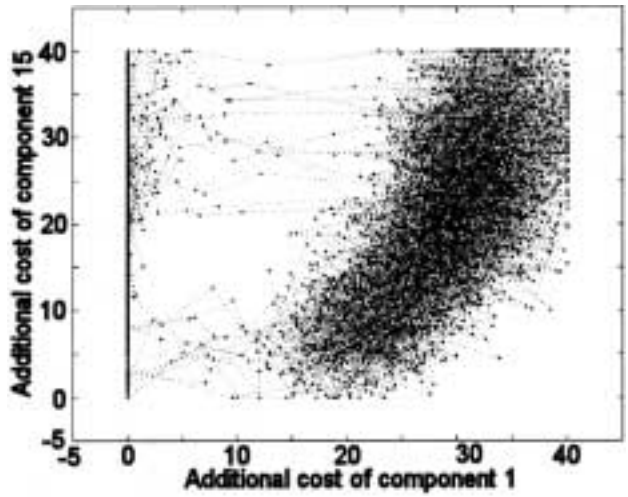


Fig. 3. Correlation between the cost components $x_1$ and $x_{15}$.

connected by dotted lines. Resulting from the nonconvex shape of $p_1$ and $p_2$, there exists nonconnected subsets of Pareto-optimal solutions.

In Fig. 3, the correlation between the component costs $x_1$ and $x_{15}$ according to the 5000 calculated points is visualized and these points are connected by dotted lines.

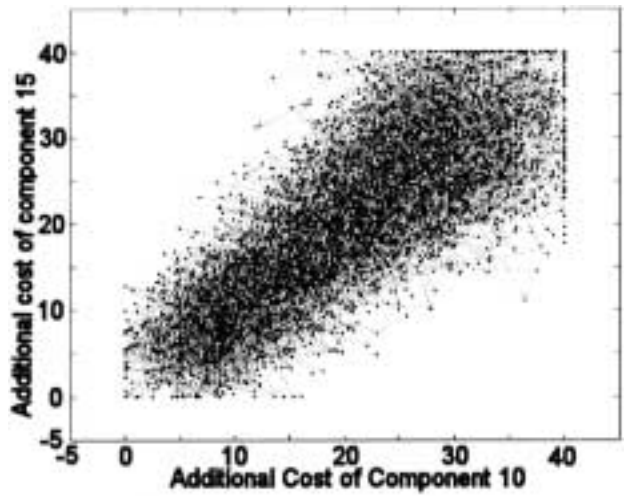


Fig. 4. Correlation between the cost components $x_{10}$ and $x_{15}$.

In Fig. 4, the correlation between the component costs $x_{10}$ and $x_{15}$ according to the 5000 calculated points is visualized and these points are connected by dotted lines.

The nonconnected subsets of Pareto-optimal solutions can be recognized in the above figures whenever component 1 or component 2 is included.

These results show that the numerical solution of (SDE) leads to a numerical cover of the Pareto-optimal solutions. Deterministic approaches to the numerical computation of global Pareto-optimal points lead to classical global optimization problems, but most of these approaches are not able to handle more than ten variables or to compute a cover of solutions in general.

## References


1. Fishburn, P. C., *Lexicographic Orders, Utilities, and Decision Rules*, Management Science, Vol. 20, pp. 1442–1471, 1973.
2. Huber, G. P., *Multiattribute Utility Models: A Review of Field and Fieldlike Studies*, Management Science, Vol. 20, pp. 1393–1402, 1973.
3. Keeney, R. L., *Utility Functions for Multiattributed Consequences*, Management Science, Vol. 18, pp. 276–287, 1972.
4. Yu, P., *A Class of Decisions for Group Decision Problems*, Management Science, Vol. 19, pp. 936–946, 1973.
5. Zeleny, M., *A Concept of Compromise Solutions and the Method of the Displaced Ideal*, Computational Operations Research, Vol. 7, pp. 479–489, 1974.
6. Hillermeier, C., *Eine Homotopiemethode zur Vektoroptimierung*, Habilitation Thesis, Technical University of Munich, 1999.
7. Fonseca, C. M., and Fleming, P. J., *An Overview of Evolutionary Algorithms in Multiobjective Optimization*, Evolutionary Computation, Vol. 3, pp. 416–423, 1995.
8. Timmel, G., *Ein stochastisches Suchverfahren zur Bestimmung der optimalen Kompromisslösung bei statistischen polykriteriellen Optimierungsaufgaben*, TH Illmenau, Report 5, 1980.
9. Bauer, H., *Wahrscheinlichkeitstheorie*, Walter de Gruyter, Berlin, Germany, 1991.
10. Protter, P., *Stochastic Integration and Differential Equations*, Springer, New York, NY, 1990.
11. Schäffler, S., *Unconstrained Global Optimization Using Stochastic Integral Equations*, Optimization, Vol. 35, pp. 43–60, 1995.
12. Hasminskij, R. Z., *Stochastic Stability of Differential Equations*, Sijthoff and Noordhoff, Amsterdam, Netherlands, 1980.