Hypervolume Sharpe-Ratio Indicator: Formalization and First Theoretical Results

Andreia P. Guerreiro(✉) and Carlos M. Fonseca

CISUC, Department of Informatics Engineering, University of Coimbra, Pólo II, Pinhal de Marrocos, 3030-290 Coimbra, Portugal
{apg,cmfonsec}@dei.uc.pt

Abstract. Set-quality indicators have been used in Evolutionary Multiobjective Optimization Algorithms (EMOAs) to guide the search process. A new class of set-quality indicators, the Sharpe-Ratio Indicator, combining the selection of solutions with fitness assignment has been recently proposed. This class is based on a formulation of fitness assignment as a Portfolio Selection Problem which sees solutions as assets whose returns are random variables, and fitness as the investment in such assets/solutions. An instance of this class based on the Hypervolume Indicator has shown promising results when integrated in an EMOA called POSEA. The aim of this paper is to formalize the class of Sharpe-Ratio Indicators and to demonstrate some of the properties of that particular Sharpe-Ratio Indicator instance concerning monotonicity, sensitivity to scaling and parameter independence.

Keywords: Sharpe Ratio · Portfolio selection · Evolutionary algorithms · Multiobjective optimization

1 Introduction

Indicator-based Evolutionary Multiobjective Optimization Algorithms (EMOAs) are currently among the state-of-the-art in Evolutionary Multiobjective Optimization. These EMOAs rely on quality indicators to guide the search, which map a point set into a scalar value, such as the Hypervolume Indicator [5,9]. Good quality indicators capture in a single value the proximity to the Pareto front and the sparsity/diversity of the set, which tends to enhance the capability of indicator-based EMOAs to find well-spread sets of good solutions.

Studies of quality-indicator properties have shown the abilities and limitations of indicator-based EMOAs. Such properties allow one to better understand, for example, whether an indicator-based EMOA aiming at the maximization of the indicator, is able to converge to the Pareto Front (monotonicity [10]) or understand which distribution each indicator favors (optimal μ-distributions [1]).

Yevseyeva *et al.* [8] established a link between the theory of Portfolio Selection and selection in Evolutionary Algorithms (EAs) by making an analogy between assets and individuals, expected return and individual quality, and return covariance and lack of diversity. They proposed that individuals be assessed through

© Springer International Publishing AG 2016
J. Handl et al. (Eds.): PPSN XIV 2016, LNCS 9921, pp. 814–823, 2016.
DOI: 10.1007/978-3-319-45823-6_76

the optimization of a Portfolio Selection Problem (PSP), formalized as the bi-objective problem of assigning investment to a set of assets so as to maximize expected return while minimizing return variance (associated to risk). This translates into the problem of assigning fitness to an EA population so as to maximize overall population quality while minimizing lack of diversity. Due to the bi-objective nature of the PSP, different optimal investment strategies balancing risk and expected return may be defined, such as the Sharpe Ratio, a risk-adjusted performance index well known in Finance [3]. A new indicator related to the Hypervolume Indicator, but based on the maximization of the Sharpe Ratio, was proposed and integrated in an EMOA with promising results. However, its theoretical properties have not been considered so far.

The goal of this paper is to formalize the class of Sharpe-Ratio Indicators and to study some of the properties of the indicator proposed by Yevseyeva *et al.* [8]. Section 2 provides the background. Section 3 details and formalizes the class of indicators based on the Sharpe Ratio and reintroduces the indicator proposed by Yevseyeva *et al.*, which will be called Hypervolume Sharpe Ratio (HSR) Indicator, as an instance of this class. Then, some properties of the HSR Indicator regarding monotonicity, reference points, and scaling independence, will be demonstrated in Sect. 4. Some conclusions are drawn in Sect. 5.

2 Background

2.1 Definitions

In multiobjective optimization, each solution is mapped according to d objective functions onto a point in the objective space, $\mathbb{R}^d$. For simplicity, only those points in objective space will be considered throughout this paper. Note that a number in parentheses in superscript is used for enumeration (e.g. $a^{(1)}, a^{(2)}, a^{(3)} \in \mathbb{R}^d$) while a number in subscript is used to refer to a coordinate of a point/vector (e.g. v_i is the i^{th} coordinate of $v \in \mathbb{R}^d$). As the objective space is a partially ordered set, the Pareto dominance relation is introduced [4,11]:

Definition 1 ***(Dominance)****. A point $u \in \mathbb{R}^d$ is said to weakly dominate a point $v \in \mathbb{R}^d$, iff $u_i \leq v_i$ for all $1 \leq i \leq d$, and this is represented as $u \leq v$. If, in addition $u \neq v$, then u is said to dominate v and is represented as $u < v$. If $u_i < v_i$ for all $1 \leq i \leq d$, then u is said to strongly dominate v, and this is represented as $u \ll v$.*

Definition 2 ***(Set dominance)****. A set $\mathrm{A} \subset \mathbb{R}^d$ is said to weakly dominate a set $\mathrm{B} \subset \mathbb{R}^d$ iff $\forall_{b \in \mathrm{B}}, \exists_{a \in \mathrm{A}} : a \leq b$. This is represented as $\mathrm{A} \preceq \mathrm{B}$. A is said to dominate a set B iff $\mathrm{A} \preceq \mathrm{B}$ and $\mathrm{B} \npreceq \mathrm{A}$, and this is represented as $\mathrm{A} \prec \mathrm{B}$.*

2.2 Properties

A set-indicator is a function I that assigns a real value to a non-empty set of points in $\mathbb{R}^d$ [10]. Among the properties a set-indicator may possess [10], this paper will cover parameter independence, sensitivity to scaling and monotonicity.

Typically, an indicator is easier to use the lower is the number of parameters that must be set. A *scaling invariant* indicator (e.g. the cardinality indicator [10]) guarantees that the indicator value for any subset of the objective space remains unchanged when the objective space is scaled. A weaker form of invariance, called *scaling independence*, ensures that the order defined by an indicator among all subsets of the objective space is kept when the objective space is scaled.

Monotonicity is an important property as it formalizes the empirical notion of agreement between indicator values and set dominance. A monotonic indicator guarantees that a set of nondominated solutions is never considered to be worse than another set which it dominates. A definition of (weak) monotonicity of a set-quality indicator with respect to set dominance is given in [10]:

Definition 3 ***(Monotonicity).*** *A set-indicator I is weakly monotonic w.r.t set dominance iff, given two point sets* $\mathrm{A}, \mathrm{B} \subset \mathbb{R}^d$*,* $\mathrm{A} \prec \mathrm{B}$ *implies* $I(\mathrm{A}) \geq I(\mathrm{B})$*.*

The above properties have been studied for indicators such as the hypervolume indicator (strictly monotonic [10] for sets of points that strongly dominate the reference point, parameter-dependent [1], scaling independent [5,9]) and the additive ϵ-indicator (weakly monotonic [10], dependent on multiple parameters [10]), thereby motivating their use in EMOAs as well as in performance assessment. Not holding such properties may discourage the use of an indicator in EMOAs. For example, a non-monotonic indicator may prefer non-Pareto Front solutions over Pareto front solutions dominating them, as is the case with the Average Hausdorff distance [7] and cardinality [10].

2.3 Sharpe Ratio

A portfolio balancing return and risk, is obtained by optimizing Problem 1:

Problem 1 (**Sharpe-Ratio Maximization**). Let $\mathrm{A} = \{a^{(1)}, \ldots, a^{(n)}\}$ be a non-empty set of assets, let vector $r \in \mathbb{R}^n$ denote the expected return of these assets and matrix $Q \in \mathbb{R}^{n \times n}$ denote the return covariance between pairs of assets. Let $x \in [0,1]^n$ be the investment vector where x_i denotes the investment in asset $a^{(i)}$. The Sharpe-Ratio maximization problem is defined as:

$$\max_{x \in [0,1]^n} \quad h(x) = \frac{r^T x - r_f}{\sqrt{x^T Q x}} \quad \text{s. t.} \quad \sum_{i=1}^{n} x_i = 1 \tag{1}$$

where r_f represents the return of a riskless asset and $h(x)$ is the Sharpe Ratio [3].

Although Problem 1 is non-linear, $h(x)$ may be homogenized and thus, it may be restated as an equivalent convex quadratic programming (QP) problem [3]:

Problem 2 (**Sharpe-Ratio Maximization - QP Formulation**).

$$\min_{y \in \mathbb{R}^n} \quad g(y) = y^T Q y \tag{2a}$$

$$\text{s. t.} \quad \sum_{i=1}^{n} (r_i - r_f) y_i = 1 \tag{2b}$$

$$y_i \geq 0, \quad i = 1, \ldots, n \tag{2c}$$

The optimal investment x^* for Problem 1, i.e., the optimal risky portfolio, is given by $x^* = y^*/k$, where y^* is the optimal solution of Problem 2 and $k = \sum_{i=1}^{n} y_i^*$.

So far, the set of assets A has been considered to be fixed and so have r and Q. However, in this paper, r and Q are computed as function of a set of assets A that is not fixed and thus, $h^{\mathrm{A}}(x)$ and $g^{\mathrm{A}}(y)$ will be used instead of $h(x)$ and $g(y)$, respectively, to highlight this dependence where needed. Moreover, with a slight abuse of language, a solution y to Problem 2 will also be called an investment vector, as for a solution x for Problem 1.

3 Sharpe-Ratio Indicator

In this section, the class of Sharpe-Ratio Indicators is formalized, and the Hypervolume Sharpe-Ratio Indicator proposed by Yevseyeva *et al.* [8] is instantiated.

The return of each individual is related to the preferences of a Decision Maker (DM) and different methods can be used to model the uncertainty surrounding DM preferences. Yevseyeva *et al.*'s [8] interpretation of selection in EAs as a portfolio selection problem sees the return of each individual asset as a random variable whose expected values can be computed.

Problem 1 does not state what the expected return and covariance of assets/individuals are. Different preferences lead to different ways of modeling return (and vice-versa) which may lead to different investment strategies in EAs. Therefore, a broad class of indicators based on the Sharpe Ratio can be defined:

Definition 4 ***(Sharpe-Ratio Indicator)****. Given a non-empty set of assets* $\mathrm{A} = \{a^{(1)}, \ldots, a^{(n)}\}$*, the corresponding expected return,* r*, and covariance matrix,* Q*, the Sharpe-Ratio Indicator,* $I_{SR}(\mathrm{A})$*, is defined as follows:*

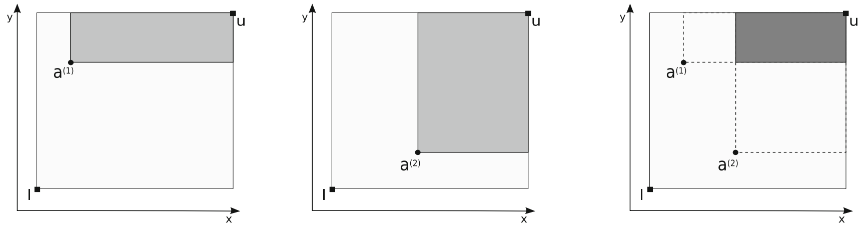
$$I_{SR}(\mathrm{A}) = \max_{x \in \Omega} \quad h^{\mathrm{A}}(x) \tag{3}$$

where $\Omega \subset [0,1]^n$ *is the set of solutions that satisfy the constraints of Problem 1.*

Note that the Sharpe-Ratio Indicator simultaneously evaluates the quality of the set A through a scalar, $I_{SR}(\cdot)$, and also the importance of each solution in that set through the optimal investment vector x^*.

The Hypervolume Sharpe-Ratio Indicator (HSR Indicator) is an instance of the Sharpe-Ratio Indicator where the expected return vector and the return covariance matrix are computed based on the Hypervolume Indicator as proposed by Yevseyeva *et al.* [8]. The expected return of a solution is the probability of that solution being satisfactory to the DM, assuming a uniform distribution of the DM's goal vector in an orthogonal range $[l, u]$, $l, u \in \mathbb{R}^d$. For the i^{th} individual in the population, this is represented by component p_i of a vector p, whereas the return covariance between the i^{th} and j^{th} individuals is represented by element q_{ij} of a matrix Q $(i, j = 1, \ldots, n)$. Let:

$$p_{ij}(l, u) = \frac{\Lambda([l, u] \cap [a^{(i)}, \infty[\cap [a^{(j)}, \infty[)}{\Lambda([l, u])} = \frac{\prod_{k=1}^{d}(u_k - \max(a_k^{(i)}, a_k^{(j)}))}{\prod_{k=1}^{d}(u_k - l_k)} \tag{4}$$



(a) $\Lambda([l,u] \cap [a^{(1)}, \infty[)$ (b) $\Lambda([l,u] \cap [a^{(2)}, \infty[)$ (c) $\Lambda([l,u] \cap [a^{(1)}, \infty[\cap [a^{(2)}, \infty[)$

Fig. 1. An example of the region measured to compute p_{ij}, given a point set $\mathrm{A} = \{a^{(1)}, a^{(2)}\} \subset \mathbb{R}^2$. The region measured to compute p_1 and p_2 and p_{12} is depicted in darker gray in Figures (a), (b) and (c), respectively.

where $l, u \in \mathbb{R}^d$ are two reference points and $\Lambda(\cdot)$ denotes the Lebesgue measure [2]. Note that p_{ij} is, therefore, the normalized hypervolume indicator of the region jointly dominated by $a^{(i)}$ and $a^{(j)}$ inside the region of interest, $[l, u]$. Moreover, from the formulation [8], $r_i(l, u) = p_i(l, u) = p_{ii}(l, u)$ and $q_{ij}(l, u) = p_{ij}(l, u) - p_i(l, u)p_j(l, u)$. For the sake of readability, $P = [p_{ij}]_{n\times n}$ and $Q = [q_{ij}]_{n\times n}$ will be assumed to have been previously calculated and, therefore, parameters l and u from expression (4) will be omitted as long as no ambiguity arises. Note that, from the definition of q_{ij}, $Q = P - pp^T$.

In Fig. 1, assuming w.l.o.g. that $l = (0, 0)$ and $u = (1, 1)$, and thus, $\Lambda([l, u]) = 1$, the area of the darker regions in Figs. 1(a) to (c) are, exactly, p_1, p_2 and p_{12}, respectively. Note that p_{ii} is related to the area dominated by $a^{(i)}$ inside the region $[l, u]$, while p_{ij} is related to the area simultaneously dominated by $a^{(i)}$ and $a^{(j)}$ inside the region $[l, u]$.

The Sharpe Ratio $h^{\mathrm{A}}(x)$ for the set of solutions A where r and Q are defined as in (4) will be represented by $h^{\mathrm{A}}_{\mathrm{HSR}}(x, l, u)$. Analogously to the Sharpe-Ratio Indicator, the HSR Indicator is formally defined as follows:

Definition 5 ***(Hypervolume Sharpe-Ratio Indicator).*** *Given a non-empty point set* $\mathrm{A} = \{a^{(1)}, \ldots, a^{(n)}\} \subset \mathbb{R}^d$*, the points* $l, u \in \mathbb{R}^d$*, the expected return* p *and the covariance* Q *computed as expressed in (4), the Hypervolume Sharpe-Ratio Indicator* $I_{\mathrm{HSR}}(\mathrm{A}, l, u)$ *is given by:*

$$I_{\mathrm{HSR}}(\mathrm{A}, l, u) = \max_{x \in \Omega} \; h^{\mathrm{A}}_{\mathrm{HSR}}(x, l, u) \tag{5}$$

where $\Omega \subset [0, 1]^n$ *is the set of solutions that satisfy the constraints of Problem 1.*

As Yevseyeva *et al.* [8] pointed out, it follows from the definition of q_{ij} that the riskless asset is such that $r_f = 0$. Consequently, Problem 2 may be simplified by noting that the constraint (2b) must always be satisfied. Therefore, the following is true for any solution y in the feasible space Ω:

$$y^T Q y = \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} y_i y_j - \sum_{i=1}^{n} p_i y_i \sum_{j=1}^{n} p_j y_j = \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} y_i y_j - 1 = y^T P y - 1 \tag{6}$$

Note that this simplification of Problem 2 is applicable to any DM preference model where r_f is zero.

4 Properties of the HSR Indicator

In the following, the optimal investment is shown to be invariant to the setting of l under certain conditions. Varying l can also be interpreted as applying linear transformations to the objective space, under which the indicator is scaling independent. Finally, the HSR Indicator is shown to be weakly monotonic.

4.1 Reference Points and Linear Scaling

Given a non-empty point set $\mathrm{A} \subset \mathbb{R}^d$ and the reference points $l, u \in \mathbb{R}^d$, such that for all $a \in \mathrm{A}$, $l \leq a \ll u$ holds, the location of l can be shown to have no effect on the optimal investment in A as long as $\{l\} \leq \mathrm{A}$ and u remains fixed. This is equivalent to applying a linear transformation to the objective space, with u as the center of the transformation. Thus, in practice, only one parameter of the HSR Indicator needs to be set (the upper reference point, u). Formally:

Theorem 1. *Let* $\mathrm{A} \subset \mathbb{R}^d$ *be a non-empty point set, let* $l, u \in \mathbb{R}^d$ *be two reference points such that* $\forall_{a\in\mathrm{A}}$, $l \leq a \ll u$, *and let* $x^* \in [0,1]^n$ *be such that* $I_{\mathrm{HSR}}(\mathrm{A}, l, u) = h^{\mathrm{A}}_{\mathrm{HSR}}(x^*, l, u)$. *If* $l' \in \mathbb{R}^d$ *is such that* $\{l'\} \leq \mathrm{A}$, *then* x^* *also satisfies* $I_{\mathrm{HSR}}(\mathrm{A}, l', u) = h^{\mathrm{A}}_{\mathrm{HSR}}(x^*, l', u)$.

Proof. Recall expression (4), of p_{ij}, for a given point set $\mathrm{A} = \{a^{(1)}, \ldots, a^{(n)}\} \subset \mathbb{R}^d$, where $p = [p_{ii}]_{n\times 1}$ and $P = [p_{ij}]_{n\times n}$ $(i, j = 1, ..., n)$. $P(l', u)$ and $p(l', u)$ may be defined as functions of $P(l, u)$ and $p(l, u)$, respectively, in the following way:

$$P(l', u) = \frac{v}{v'} P(l, u) \tag{7a}$$

$$p(l', u) = \frac{v}{v'} p(l, u) \tag{7b}$$

where $v = \Lambda([l, u])$ and $v' = \Lambda([l', u])$.

Assume that $y \in \mathbb{R}^n$ is the vector of variables of Problem 2 (minimizing $g^{\mathrm{A}}_{\mathrm{HSR}}(y, l, u) = y^T P y - 1$), when l is set as the lower reference point and that, analogously, $y' \in \mathbb{R}^n$ is the corresponding vector of variables when l' is used instead. Taking into account expressions (7b) and the equality constraint of Problem 2, the following is derived:

$$p(l, u)^T y = p(l', u)^T y' \quad \Leftrightarrow \quad p(l, u)^T y = \frac{v}{v'} p(l, u)^T y' \Leftrightarrow \quad y = \frac{v}{v'} y' \tag{8}$$

which implies that when y is such that $y = \frac{v}{v'} y'$, if $y' > 0$ then $y > 0$ and therefore, if y' is feasible so is y and vice-versa. Hence, the following holds:

$$g^{\mathrm{A}}_{\mathrm{HSR}}(y', l', u) = y'^T P(l', u) y' - 1 = \frac{v'}{v} y^T P(l, u) y - 1 = \frac{v'}{v} g^{\mathrm{A}}_{\mathrm{HSR}}(y, l, u) - 1 + \frac{v'}{v}$$

Therefore, the optimal solution y'^* for Problem 2, given l', can be obtained from the optimal solution y^*, given l, i.e., $y'^* = \frac{v'}{v} y^*$. Consequently, the optimal solution x^* for Problem 1:

$$x^* = \frac{y^*}{\sum_{i=1}^n y_i^*} = \frac{\frac{v}{v'} y'^*}{\frac{v}{v'} \sum_{i=1}^n y_i'^*} = \frac{y'^*}{\sum_{i=1}^n y_i'^*}. \tag{9}$$

Hence, $I_{\text{HSR}}(\text{A}, l, u) = h_{\text{HSR}}(x^*, l, u)$ implies that $I_{\text{HSR}}(\text{A}, l', u) = h_{\text{HSR}}(x^*, l', u)$ thus, Theorem 1 is proved.

Note that moving the lower reference point, l, for example, to a lower value of one of the objectives while the others are kept the same, is equivalent to scaling down that objective with respect to the other objectives. Thus, the placement of l can also be seen as a way of linearly scaling the objective functions (as long as this reference point continues to dominate A). Therefore, by Theorem 1, scaling the objective space under such conditions does not affect the optimal investment.

Scaling through l comes down to multiplying p_i and p_{ij} by a positive constant as in the proof of Theorem 1. Observing the Sharpe Ratio expression $h(x)$ in Problem 1, the HSR-indicator is not scaling invariant, i.e., scaling the objective space will affect the indicator value. However, the HSR-indicator is scaling independent under these linear transformations, as shown next.

Theorem 2 ***(Linear-Scaling Independence of*** I_{HSR}*). Consider two point sets* $\text{A}, \text{B} \subset \mathbb{R}^d$ *and two reference points* $l, u \in \mathbb{R}^d$ *such that* $\forall_{a \in \text{A}, b \in \text{B}}, l \leq a, b \ll u$*. Assume w.l.o.g. that* A *and* B *are such that* $I_{\text{HSR}}(\text{A}, l, u) \leq I_{\text{HSR}}(\text{B}, l, u)$*. Then,* $I_{\text{HSR}}(\text{A}, l', u) \leq I_{\text{HSR}}(\text{B}, l', u)$ *holds for any* $l' \in \mathbb{R}^d$ *such that* $\{l'\} \leq \text{A}, \text{B}$*.*

Proof. Let p_A, P_A and Q_A denote, respectively, the expected return vector, the matrix of expected return and the return covariance matrix with respect to point set A. Scaling is applied to A and B in expression $h(x)$ in Problem 1 by multiplying a constant $t > 0$ by each p_i and p_{ij} and, therefore, $p'_\text{A} = tp_\text{A}$ and $P'_\text{A} = tP_\text{A}$, where $t = \frac{\Lambda([l,u])}{\Lambda([l',u])}$. Consequently,

$$\begin{aligned} I_{\text{HSR}}(\text{A}, l', u) \leq I_{\text{HSR}}(\text{B}, l', u) \qquad & \Leftrightarrow \\ \frac{tp_\text{A}^T x_\text{A}}{\sqrt{tx_\text{A}^T P_\text{A} x_\text{A} - t^2 x_\text{A}^T p_\text{A} p_\text{A}^T x_\text{A}}} \leq \frac{tp_\text{B}^T x_\text{B}}{\sqrt{tx_\text{B}^T P_\text{B} x_\text{B} - t^2 x_\text{B}^T p_\text{B} p_\text{B}^T x_\text{B}}} \qquad & \Leftrightarrow \\ \frac{1}{t}(x_\text{A}^T p_\text{A} p_\text{A}^T x_\text{A})(x_\text{B}^T P_\text{B} x_\text{B}) \leq \frac{1}{t}(x_\text{B}^T p_\text{B} p_\text{B}^T x_\text{B})(x_\text{A}^T P_\text{A} x_\text{A}) \qquad & \end{aligned} \tag{10}$$

Since the constant t vanishes from the inequality, which includes the case where the lower reference point is not changed ($t = 1$), Theorem 2 is proved.

4.2 Monotonicity

The property of monotonicity may now be stated for the HSR Indicator:

Theorem 3 ***(Weak Monotonicity of the Hypervolume Sharpe-Ratio Indicator).*** *Consider two reference points $l, u \in \mathbb{R}^d$ and two point sets $\mathrm{A}, \mathrm{B} \subset [l, u[$ such that $\mathrm{A} \prec \mathrm{B}$. Then $I_{\mathrm{HSR}}(\mathrm{A}, l, u) \geq I_{\mathrm{HSR}}(\mathrm{B}, l, u)$.*

In order to prove this theorem, two auxiliary results are stated first. Lemma 1 is used to prove Lemma 2, which is then used in the proof of the theorem. Similarly to expression (4), for any two points $a, b \in [l, u[$, let p_{ab} denote the measure of the region bounded above by $u \in \mathbb{R}^d$ that is dominated simultaneously by a and b, and let $p_a = p_{aa}$. Note that $p_c > 0$ for any point $c \in [l, u[$.

Lemma 1. *Consider two points $a, b \in [l, u[$ such that $a < b$. Then, for all $c \in [l, u[\subset \mathbb{R}^d$, $p_b p_{ac} \leq p_{bc} p_a$ holds.*

Proof. Consider w.l.o.g. that $l = (0, ..., 0)$ and $u = (1, \ldots, 1)$ and therefore, $\Lambda([l, u]) = 1$. Lemma 1 will be proved by contradiction. Hence, suppose that, for some choice of $c \in [l, u[$:

$$\begin{gathered} p_b p_{ac} > p_{bc} p_a \Leftrightarrow \\ \prod_{i=1}^{d} (1 - b_i)(1 - \max(a_i, c_i)) > \prod_{i=1}^{d} (1 - \max(b_i, c_i))(1 - a_i) \end{gathered} \tag{11}$$

Thus, there should be, at least, a dimension i for which the following holds:

$$(1 - b_i)(1 - \max(a_i, c_i)) > (1 - \max(b_i, c_i))(1 - a_i) \tag{12}$$

However, by manipulating expression (12), it is possible to verify that $b_i \geq c_i$ implies $a_i > \max(a_i, c_i)$, and that $b_i < c_i$ implies $a_i > b_i$, which are both untrue. Consequently, expression (11) does not hold either, and Lemma 1 is proved.

Lemma 2. *Consider a point set $\mathrm{A} = \{a^{(1)}, \ldots, a^{(n)}\} \subset [l, u[$, where $n \geq 2$, and, without loss of generality, assume that $a^{(2)} < a^{(1)}$. Then, the investment vector $x^* \in [0, 1]^n$ that maximizes the Sharpe Ratio for the set A is such that the investment in $a^{(1)}$, denoted by x_1^*, is zero.*

Proof. Note that, for constraint (2b) to be satisfied, there has to be a strictly positive investment in, at least one asset and thus, all constraints are linearly independent for any feasible solution to Problem 2. Thus, the prerequisites of the first-order necessary optimality conditions (KKT conditions) [6] are satisfied.

Following the notation and definitions in Nocedal and Wright [6], the KKT conditions state that if a feasible solution y^* is optimal, then there is a Lagrange multiplier vector λ^* for which all components associated to an inequality constraint are nonnegative and the product of each component of λ^* and the corresponding constraint at y^* is zero. Moreover, the gradient of the Lagrangian function w.r.t y^* is zero ($\nabla_y \mathcal{L}(y^*, \lambda^*) = 0$). The Lagrangian function, for the HSR Indicator (in Problem 2) is:

$$\mathcal{L}(y, \lambda) = y^T P y - 1 - \lambda_1 p^T y - \sum_{i=2}^{n+1} \lambda_i y_{i-1} \tag{13}$$

and the corresponding partial derivative w.r.t. y_k at (y^*, λ^*) for $k = 1, ..., n$ is:

$$\frac{\partial \mathcal{L}(y^*, \lambda^*)}{\partial y_k} = 2\sum_{i=1}^{n} p_{ik} y_i^* - p_k \lambda_1^* - \lambda_{k+1}^* = 0 \tag{14}$$

Lemma 2 is proved by contradiction. Let y_1^* and y_2^* represent the investments in $a^{(1)}$ and $a^{(2)}$, respectively. Since $a^{(2)}$ dominates $a^{(1)}$, the following holds:

$$p_1 = p_{12}, \quad p_1 < p_2 \quad \text{and} \quad p_{1i} \leq p_{2i}, \quad i = 3, \ldots, n \tag{15}$$

Suppose that the optimal investment y^* is such that $y_1^* > 0$. Then, the KKT conditions imply that $\lambda_2^* = 0$. By manipulating Eq. (14) for $k = 1, 2$ using the conditions in (15), the following condition on λ_3^* is obtained:

$$p_1(p_{12} - p_2)y_1^* + \sum_{i=3}^{n}(p_1 p_{2i} - p_{1i} p_2)y_i^* = \tfrac{p_1 \lambda_3^*}{2} \geq 0 \tag{16}$$

$\lambda_3^* \geq 0$ must be true so that it is a valid Lagrange multiplier. Therefore, since $p_1 > 0$, the left-hand side of expression (16) must be zero or positive. However, the first term is clearly negative since $p_{12} = p_1 < p_2$, and the sum is non-positive by Lemma 1.

Therefore, no optimal Lagrange multiplier vector λ^* exists for which the KKT conditions hold true when y_1^* is strictly positive, and consequently, y^* cannot be optimal. Therefore, $y_1^* = 0$ which implies that $x_1^* = 0$ and proves Lemma 2.

Proof (Theorem 3). Consider two point sets $\mathrm{A}, \mathrm{B} \subset [l, u[\subset \mathbb{R}^d$, such that $|\mathrm{A}|, |\mathrm{B}| \geq 1$ and $\mathrm{A} \prec \mathrm{B}$. Since any points in $\mathrm{B} - \mathrm{A}$ are dominated points in $\mathrm{A} \cup \mathrm{B}$, by Lemma 2 they are assigned zero investment, and $I_{\mathrm{HSR}}(\mathrm{A} \cup \mathrm{B}) = I_{\mathrm{HSR}}(\mathrm{A})$ must hold true. Suppose that $I_{\mathrm{HSR}}(\mathrm{B}) > I_{\mathrm{HSR}}(\mathrm{A})$. Then, an investment strategy in $\mathrm{A} \cup \mathrm{B}$ with Sharpe Ratio greater than $I_{\mathrm{HSR}}(\mathrm{A} \cup \mathrm{B})$ where zero investment is given to the points in $\mathrm{A} - \mathrm{B}$ would exist, which leads to a contradiction and proves the theorem.

5 Concluding Remarks

The Sharpe-Ratio Indicator class has been formalized, and theoretical results on the particular HSR Indicator have been presented regarding the independence of one of the reference points, scaling independence and the monotonicity property. Although the formulation of the HSR Indicator involves two reference points, only one needs to be set in practice. The second reference point is just a technical parameter that is required by the formulation. Indeed, the optimal investment is not affected by the linear objective rescaling implied by changes to this second reference point, and the indicator is scaling independent under such transformations. Thus, the HSR Indicator does not require more parameters to be set than, for example, the Hypervolume Indicator. The HSR Indicator is also weakly monotonic w.r.t. set dominance.

The study of other properties of interest, including optimal μ-distributions for the HSR Indicator, will be the subject of future work.

Acknowledgments. This work was supported by national funds through the Portuguese Foundation for Science and Technology (FCT), by the European Regional Development Fund (FEDER) through COMPETE 2020 – Operational Program for Competitiveness and Internationalization (POCI). A. P. Guerreiro acknowledges FCT for Ph.D. studentship SFHR/BD/77725/2011, co-funded by the European Social Fund and by the State Budget of the Portuguese Ministry of Education and Science in the scope of NSRF–HPOP–Type 4.1–Advanced Training.

References

1. Auger, A., Bader, J., Brockhoff, D., Zitzler, E.: Theory of the hypervolume indicator: optimal μ-distributions and the choice of the reference point. In: Foundations of Genetic Algorithms (FOGA 2009), pp. 87–102. ACM (2009)
2. Beume, N., Naujoks, B., Emmerich, M.: SMS-EMOA: multiobjective selection based on dominated hypervolume. EJOR **181**, 1653–1669 (2007)
3. Cornuejols, G., Tuntuncu, R.: Optimization Methods in Finance. Cambridge University Press, Cambridge (2007)
4. Ehrgott, M.: Multicriteria Optimization, 2nd edn. Springer, Heidelberg (2005)
5. Knowles, J.D.: Local-search and hybrid evolutionary algorithms for Pareto optimization. Ph.D. thesis, Department of Computer Science, University of Reading (2002)
6. Nocedal, J., Wright, S.J.: Numerical Optimization. Springer Series in Operations Research and Financial Engineering, 2nd edn. Springer, New York (2006)
7. Rudolph, G., Schütze, O., Trautmann, H.: On the closest averaged Hausdorff archive for a circularly convex Pareto front. In: Squillero, G., Burelli, P. (eds.) EvoApplications 2016. LNCS, vol. 9598, pp. 42–55. Springer, Heidelberg (2016). doi:10.1007/978-3-319-31153-1_4
8. Yevseyeva, I., Guerreiro, A.P., Emmerich, M.T.M., Fonseca, C.M.: A portfolio optimization approach to selection in multiobjective evolutionary algorithms. In: Bartz-Beielstein, T., Branke, J., Filipič, B., Smith, J. (eds.) PPSN 2014. LNCS, vol. 8672, pp. 672–681. Springer, Heidelberg (2014)
9. Zitzler, E.: Evolutionary algorithms for multiobjective optimization: methods and applications. Ph.D. thesis, ETH Zurich, Switzerland (1999)
10. Zitzler, E., Knowles, J.D., Thiele, L.: Quality assessment of pareto set approximations. In: Branke, J., Deb, K., Miettinen, K., Słowiński, R. (eds.) Multiobjective Optimization. LNCS, vol. 5252, pp. 373–404. Springer, Heidelberg (2008)
11. Zitzler, E., Thiele, L., Laumanns, M., Fonseca, C.M., da Fonseca, V.G.: Performance assessment of multiobjective optimizers: an analysis and review. IEEE Trans. Evol. Comput. **7**(2), 117–132 (2003)