



## On Bi-Objective convex-quadratic problems

Cheikh Touré, Anne Auger, Dimo Brockhoff, Nikolaus Hansen

### ► To cite this version:

Cheikh Touré, Anne Auger, Dimo Brockhoff, Nikolaus Hansen. On Bi-Objective convex-quadratic problems. EMO 2019 - 10th International Conference on Evolutionary Multi-Criterion Optimization, Mar 2019, East Lansing, Michigan, United States. hal-01942159

**HAL Id: hal-01942159**

**<https://inria.hal.science/hal-01942159>**

Submitted on 3 Dec 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# On Bi-Objective convex-quadratic problems

Cheikh Toure<sup>1</sup>, Anne Auger<sup>1</sup>, Dimo Brockhoff<sup>1</sup>, and Nikolaus Hansen<sup>1</sup>

Inria and CMAP, Ecole Polytechnique, France  
firstname.lastname@inria.fr  
cheikh.toure@polytechnique.edu

**Abstract.** In this paper we analyze theoretical properties of bi-objective convex-quadratic problems. We give a complete description of their Pareto set and prove the convexity of their Pareto front. We show that the Pareto set is a line segment when both Hessian matrices are proportional. We then propose a novel set of convex-quadratic test problems, describe their theoretical properties and the algorithm abilities required by those test problems. This includes in particular testing the sensitivity with respect to separability, ill-conditioned problems, rotational invariance, and whether the Pareto set is aligned with the coordinate axis.

**Keywords:** Bi-objective optimization · Pareto set · Convex front · convex-quadratic problems.

## 1 Introduction

Convex-quadratic functions are among the simplest yet very useful test functions in optimization. Given a positive definite matrix  $Q$  of  $\mathbb{R}^{n \times n}$ , a convex quadratic function is defined as

$$f(x) = \frac{1}{2}(x - x^*)^\top Q(x - x^*)$$

where  $x^*$  is the unique optimum of the function. The Hessian of  $f$  coincides with the matrix  $Q$ . The level-sets of  $f$  defined as  $\{x \in \mathbb{R}^n : (x - x^*)^\top Q(x - x^*) = c, c \geq 0\}$  are hyper-ellipsoids whose main axes are the eigenvectors of the matrix  $Q$  with length proportional to the inverse of the eigenvalues of  $Q$ .

By changing the eigenvalues and eigenvectors of  $Q$ , one can model different essential difficulties in numerical optimization: if the eigenvectors are not aligned with the coordinate axes (if the matrix  $Q$  is not diagonal), then the associated function is non-separable: it cannot be efficiently optimized by coordinate-wise search. In practice, difficult optimization problems are non-separable. Having a large condition number for  $Q$ , that is a large ratio between the largest and smallest eigenvalue of  $Q$  models ill-conditioned problems where the characteristic scale along different directions is very different. Ill-conditioning is very frequent in real-world problems. They arise naturally as one often optimizes quantities that have different natures and different intrinsic scales (some variables can be

akin to time, others to weights, ...) such that a unit change along each variable can have a completely different impact on the function optimized. More generally, the eigenspectrum of  $Q$  entirely characterizes the scale among the different axes of the hyper-ellipsoidal level sets and parametrizes the difficulty of the function: from the arguably easiest function, the sphere function  $f(x) = \sum_{i=1}^n x_i^2$ , to very difficult ill-conditioned functions where condition numbers of  $Q$  of up to  $10^{10}$  have been observed in real-world problems, for example in [3].

Convex-quadratic functions have been central to the design of several important classes of optimization algorithms for single-objective optimization. Newton or quasi-Newton methods use or learn a second order approximation of the objective function optimized [11]. This second order approximation is done by convex-quadratic functions (assuming that the function is twice continuously differentiable and convex). Introduced more recently, the class of derivative-free-optimization (DFO) trust-region based algorithms builds a second-order approximation of the objective function by interpolation [12]. In the evolutionary computation (EC) context, convex-quadratic functions have also played a central role for the design of algorithms like CMA-ES: they have been intensively used for designing the algorithm and the performance of the method has been carefully quantified on different eigenspectra of the matrix  $Q$  for different condition numbers [6].

Given that a multiobjective problem is “simply” the simultaneous optimization of single-objective problems, the typical difficulties of each objective function are the same as the typical difficulties of single-objective problems. In particular non-separability and ill-conditioning are important difficulties that the single functions have. Therefore, combining convex-quadratic problems seems natural for testing and designing multiobjective algorithms. This has already been done in the past for instance for the design of multiobjective versions of CMA-ES [7] or as a subset of the biobjective BBOB test function suite [2, 13].

Yet, while the difficulties encoded and parametrized within a convex-quadratic problem are well-understood for single-objective optimization, the situation is different for multiobjective optimization, starting from bi-objective optimization. Simple properties like convexity of the Pareto front associated to bi-objective convex-quadratic problems as well as properties of the Pareto set have not been systematically investigated. Additionally, convex-quadratic bi-objective test problems used in the literature do not capture all important properties one could be testing with convex-quadratic problems. There is more degree of freedom than for single objective optimization that is not exploited: we can combine two functions having the same Hessian matrix, place the optima on the functions both on one axis of the search space, ... and this will affect how the Pareto set and Pareto front look like.

This paper aims at filling the gaps from the literature on multiobjective optimization with respect to convex-quadratic problems. More precisely the objectives are twofold: clarify theoretical Pareto properties of bi-objective problems where each function is convex-quadratic and define sets of bi-objective convex-quadratic problems that allow to test different (well-understood) difficulties of

bi-objective problems. The paper is organized as follows: in Section 2 we present theoretical properties of convex-quadratic problems and discuss new test functions in Section 3.

## 2 Theoretical Properties of Bi-Objective Convex-Quadratic Problems

### 2.1 Preliminaries

We consider bi-objective problems  $(f_1, f_2)$  defined on the search space  $\mathbb{R}^n$ .

The Pareto set of  $(f_1, f_2)$  is defined as the set of all non-dominated (or efficient) solutions  $\{x \in \mathbb{R}^n \mid \nexists y \in \mathbb{R}^n \text{ such that } f_1(y) \leq f_1(x) \text{ and } f_2(y) \leq f_2(x) \text{ and at least one inequality is strict}\}$ . The image of the Pareto set (in the objective space  $\mathbb{R}^2$ ) is called the Pareto front of  $(f_1, f_2)$ . We first remark that the Pareto set remains unchanged if we compose the objective functions with a strictly increasing function. More precisely the following lemma holds.

**Lemma 1 (Invariance of the Pareto set to strictly increasing transformations of the objectives).** *Given a bi-objective problem  $x \mapsto (f_1(x), f_2(x))$  and  $g_1 : \text{Im}(f_1) \mapsto \mathbb{R}$ ,  $g_2 : \text{Im}(f_2) \mapsto \mathbb{R}$  two strictly increasing functions, then  $(f_1, f_2)$  and  $(g_1 \circ f_1, g_2 \circ f_2)$  have the same Pareto set.*

*Proof.* If  $x$  is not in the Pareto set of  $(g_1 \circ f_1, g_2 \circ f_2)$ , then there exists  $y$  such that  $g_1 \circ f_1(y) \leq g_1 \circ f_1(x)$  and  $g_2 \circ f_2(y) \leq g_2 \circ f_2(x)$  with one inequality being strict, which is equivalent to the fact that  $f_1(y) \leq f_1(x)$  and  $f_2(y) \leq f_2(x)$ , with one inequality being strict. And vice versa. Hence  $x$  is not in the Pareto set of  $(g_1 \circ f_1, g_2 \circ f_2)$  if and only if it is not in the Pareto set of  $(f_1, f_2)$ , which shows that both problems have the same Pareto set.  $\square$

From now on  $(f_1, f_2)$  denote a bi-objective convex-quadratic problem. More precisely, let  $x_1, x_2$  be two *different* vectors in  $\mathbb{R}^n$ , and  $\alpha, \beta > 0$ . Let  $Q_1$  and  $Q_2$  (in  $\mathbb{R}^{n \times n}$ ) be two positive definite matrices and consider the bi-objective *minimization* problem  $(f_1, f_2)$  defined for  $x \in \mathbb{R}^n$  as

$$f_1(x) = \frac{1}{\alpha} (x - x_1)^\top Q_1 (x - x_1), f_2(x) = \frac{1}{\beta} (x - x_2)^\top Q_2 (x - x_2). \quad (1)$$

We denote this general bi-objective convex-quadratic problem by  $\mathcal{P}$ , and assume that the optimization goal is to find (an approximation of) the Pareto set of  $\mathcal{P}$ .

### 2.2 Pareto set

We characterize in this section the Pareto set of  $\mathcal{P}$ . We use the linear scalarization method to obtain the whole Pareto set. This is doable, whenever  $f_1$  and  $f_2$  are strict convex functions (see [8]). Then the Pareto set of  $\mathcal{P}$  is described by the solutions of

$$\min_{x \in \mathbb{R}^n} (1 - t) f_1(x) + t f_2(x), \text{ for } t \in [0, 1].$$

We prove in the next proposition that the Pareto set of  $\mathcal{P}$  is a continuous and differentiable parametric curve of  $\mathbb{R}^n$  whose extremes are  $x_1$  and  $x_2$ .

**Proposition 1.** *The Pareto set of  $\mathcal{P}$  is the image of the function  $\varphi$  defined as*

$$\varphi : t \in [0, 1] \mapsto [(1-t)Q_1 + tQ_2]^{-1} [(1-t)Q_1x_1 + tQ_2x_2] . \quad (2)$$

*The function  $\varphi$  is differentiable and verifies for any  $t$  in  $[0, 1]$*

$$(1-t)Q_1 (\varphi(t) - x_1) = tQ_2 (x_2 - \varphi(t)) , \quad (3)$$

$$t[(1-t)Q_1 + tQ_2] \varphi'(t) = Q_1 (\varphi(t) - x_1) . \quad (4)$$

*Hence, the Pareto set is a continuous (differentiable) curve of  $\mathbb{R}^n$  whose extremes are  $x_1 = \varphi(0)$  and  $x_2 = \varphi(1)$ .*

*Proof.* For any  $s$  in  $[0, 1]$ , define  $g_s \stackrel{\text{def}}{=} (1-s)f_1 + sf_2$ . We observe that  $g_s$ , like  $f_1$  and  $f_2$ , is strictly convex, differentiable, and diverges to  $\infty$  when  $\|x\|$  goes to  $\infty$  (where  $\|x\|$  denotes the Euclidean norm). Then its critical point minimizes  $g_s$ . Let us now compute the gradient of  $g_s$  times  $\alpha\beta$  for  $x$  in  $\mathbb{R}^n$ :

$$\alpha\beta \nabla g_s(x) = (1-s)\alpha\beta \nabla f_1(x) + s\alpha\beta \nabla f_2(x) = 2(1-s)\beta Q_1(x-x_1) + 2s\alpha Q_2(x-x_2)$$

$$\text{Thus, } \alpha\beta \nabla g_s(x) = 2[(1-s)\beta Q_1 + s\alpha Q_2]x - 2(1-s)\beta Q_1x_1 - 2s\alpha Q_2x_2.$$

Then it follows that for any  $s$  in  $[0, 1]$ , the point that minimizes  $g_s$  (its critical point), denoted by  $\tilde{x}_s$  verifies  $\frac{(1-s)\beta Q_1 + s\alpha Q_2}{(1-s)\beta + s\alpha} \tilde{x}_s = \frac{(1-s)\beta Q_1x_1 + s\alpha Q_2x_2}{(1-s)\beta + s\alpha}$ .

Since  $[0, 1] \ni s \mapsto \frac{s\alpha}{(1-s)\beta + s\alpha} \in [0, 1]$  is bijective (its derivative is  $s \mapsto \frac{\alpha\beta}{((1-s)\beta + s\alpha)^2}$ ), then it is equivalent to parametrize the Pareto set with  $t \stackrel{\text{def}}{=} \frac{s\alpha}{(1-s)\beta + s\alpha}$ . Hence, the Pareto set is fully described by  $(\varphi(t))_{t \in [0, 1]}$  such that:

$$[(1-t)Q_1 + tQ_2] \varphi(t) = (1-t)Q_1x_1 + tQ_2x_2, \quad (5)$$

$$(1-t)Q_1 (\varphi(t) - x_1) = tQ_2 (x_2 - \varphi(t)) . \quad (6)$$

The function  $t \mapsto [(1-t)Q_1 + tQ_2]^{-1}$  is differentiable as inverse of a differentiable and invertible matrix function. Then  $\varphi$  is differentiable.

We differentiate (5) and multiply by  $t$  to obtain  $t[(1-t)Q_1 + tQ_2] \varphi'(t) = tQ_2x_2 - tQ_1x_1 + tQ_1\varphi(t) - tQ_2\varphi(t)$ . Injecting in (6) gives  $t[(1-t)Q_1 + tQ_2] \varphi'(t) = Q_1(\varphi(t) - x_1)$ , for any  $t \in [0, 1]$ .  $\square$

We obtain as corollary that when  $f_1$  and  $f_2$  have proportional Hessian matrices, then the Pareto set is the line segment between the optima of the functions  $f_1$  and  $f_2$ .

**Corollary 1.** *In the case where  $f_1$  and  $f_2$  have proportional Hessian matrices, the Pareto set of  $\mathcal{P}$  is the line segment between  $x_1$  and  $x_2$ .*

*Proof.* In that case, there exists a real  $\gamma$  such that  $\frac{Q_1}{\alpha} = \gamma \frac{Q_2}{\beta}$ . Then, Proposition 1 implies that for any  $t \in [0, 1]$ ,

$$\varphi(t) = \left[ (1-t)\gamma \frac{\alpha}{\beta} Q_2 + tQ_2 \right]^{-1} \left[ (1-t)\gamma \frac{\alpha}{\beta} Q_2 x_1 + tQ_2 x_2 \right] = \frac{\gamma\alpha(1-t)x_1 + t\beta x_2}{(1-t)\alpha\gamma + t\beta},$$

which is  $[x_1, x_2]$ , since  $[0, 1] \ni t \mapsto \frac{t\beta}{(1-t)\alpha\gamma + t\beta} \in [0, 1]$  is a bijection.  $\square$

Using Lemma 1, we directly deduce the following corollary.

**Corollary 2.** *If  $f_1$  and  $f_2$  have proportional Hessian matrices,  $g_1 : \text{Im}(f_1) \mapsto \mathbb{R}$ ,  $g_2 : \text{Im}(f_2) \mapsto \mathbb{R}$  are two strictly increasing functions, then the Pareto set of the problem  $(g_1 \circ f_1, g_2 \circ f_2)$  is the line segment between  $x_1$  and  $x_2$ .*

As an example, the double-norm problem defined as:

$(x \rightarrow \|x - x_1\|_2, x \rightarrow \|x - x_2\|_2)$  can be seen as:  $(g \circ f_1, g \circ f_2)$  where  $g(x) = \sqrt{x}$ ,  $f_1(x) = \|x - x_1\|_2^2$  and  $f_2(x) = \|x - x_2\|_2^2$ .

Then  $(g \circ f_1, g \circ f_2)$  has the same Pareto set than the double-sphere problem  $(f_1, f_2)$ , which is the line segment between  $x_1$  and  $x_2$ . Therefore the Pareto front of the double-norm problem is described by  $(t\|x_2 - x_1\|_2, (1-t)\|x_2 - x_1\|_2)_{t \in [0, 1]}$ . Thereby, the front is described by the function  $u \mapsto \|x_2 - x_1\|_2 - u$ . We recover the well-known result that the double-norm problem has a linear front.

Corollary 2 allows also to recover the Pareto set description for the one-peak scenario in the Mixed-Peak Bi-Objective Problem (see [9] and [10]).<sup>1</sup>

In general, the Pareto set of a bi-objective convex-quadratic problem is not necessarily a line segment. Consider for instance for  $n = 2$  the case where  $x_1 = (0, 0)^\top$ ,  $x_2 = (1, 1)^\top$  and where we generate two different matrices  $Q_1$  and  $Q_2$  by randomly rotating a diagonal matrix with eigenvalues 1 and 10. Two resulting Pareto fronts associated to different random rotations are depicted in Figure 1.

For  $n = 10$ , we also define  $\mathcal{P}_{10}$  setting  $x_1 = (0, \dots, 0)^\top$ ,  $x_2 = (1, \dots, 1)^\top$  and  $Q_1$  and  $Q_2$  as diagonal matrices such that for  $i = 1, \dots, 10$

$$Q_1(i, i) = 100^{\frac{i-1}{9}}, \text{ and } Q_2(i, i) = 10^{\frac{i-1}{9}}. \quad (7)$$

The different coordinates of the Pareto set given in (3) are depicted in Figure 1.

### 2.3 Convexity of the Pareto front

Corollary 1 proves that in the case where we have proportional Hessian matrices in problem  $\mathcal{P}$ , the Pareto set is a line segment. Then it is reasonable to expect a simple analytic expression for the corresponding Pareto front. In what follows, we will express the Pareto front of a bi-objective problem as a one-dimensional

<sup>1</sup> In that scenario, we set  $f_1(x) = (x - c)^\top \Sigma (x - c)$ ,  $f_2(x) = (x - c')^\top \Sigma' (x - c')$  ( $f_1$  and  $f_2$  are seen as squares of the Mahalanobis distance to the optima, with respect to the Hessian matrices),  $g_1(u) = 1 - \frac{h_1}{1 + \frac{\sqrt{u}}{r_1}}$ ,  $g_2(u) = 1 - \frac{h_2}{1 + \frac{\sqrt{u}}{r_2}}$ .

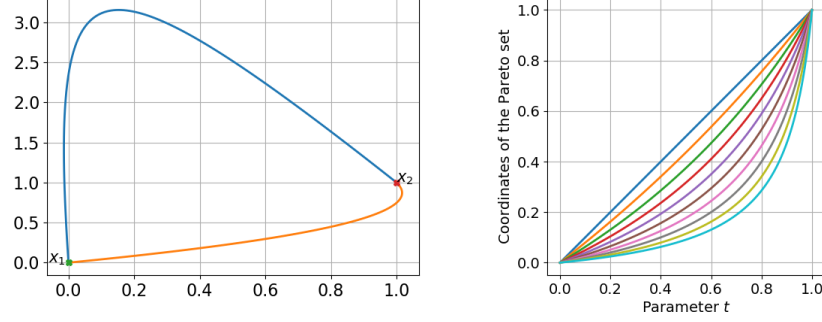


Fig. 1: Left: Two Pareto sets for  $n = 2$  represented in  $\mathbb{R}^2$  with  $Q_1$  and  $Q_2$  randomly sampled and different. Right: Pareto set for  $n = 10$  with matrices given in (7) represented as the function of the parameter  $t$  given in (3). The coordinates are ordered, the first one is on top and last one below.

function  $u \in \mathbb{R} \mapsto g(u)$ . Formally, if  $t \in \mathbb{R} \mapsto \varphi(t) \in \mathbb{R}^n$  is a parametrization of the Pareto set, then the function  $g$  satisfies  $f_2(\varphi(t)) = g(f_1(\varphi(t)))$ . It is well-known that when  $(f_1, f_2)$  is the double-sphere, that is  $f_1(x) = \frac{1}{n} \sum_{i=1}^n x_i^2$  and  $f_2(x) = \frac{1}{n} \sum_{i=1}^n (x_i - 1)^2$ , then the Pareto front expression is given by  $g(u) = (1 - \sqrt{u})^{\frac{n}{2}}$  [4]. In the next proposition, we show that this expression of the Pareto front holds (up to a normalization) for all bi-objective convex-quadratic problems, provided the Hessians of  $f_1$  and  $f_2$  are proportional.

**Proposition 2.** *When we have proportional Hessian matrices in the problem  $\mathcal{P}$ , the Pareto front is described by the following continuous and convex function:*

$$u \in [0, \kappa_\alpha] \mapsto \kappa_\beta \left(1 - \sqrt{\frac{u}{\kappa_\alpha}}\right)^2, \text{ where } \begin{cases} \kappa_\alpha = \frac{(x_2 - x_1)^\top Q_1 (x_2 - x_1)}{\alpha} \\ \kappa_\beta = \frac{(x_2 - x_1)^\top Q_2 (x_2 - x_1)}{\beta} \end{cases} \quad (8)$$

*Proof.* Denote  $u \stackrel{\text{def}}{=} f_1 \circ \varphi$  and  $v \stackrel{\text{def}}{=} f_2 \circ \varphi$ , where  $\varphi : [0, 1] \ni t \mapsto (1-t)x_1 + tx_2 \in [x_1, x_2]$  is the line segment between  $x_1$  and  $x_2$ .

For any  $t \in [0, 1]$ ,  $u(t) = f_1(\varphi(t)) = \frac{1}{\alpha} (x_2 - x_1)^\top Q_1 (x_2 - x_1) t^2$ ,  $v(t) = f_2(\varphi(t)) = \frac{1}{\beta} (x_2 - x_1)^\top Q_2 (x_2 - x_1) (1-t)^2$ . It follows that for any  $t \in [0, 1]$ :

$$v(t) = \frac{(x_2 - x_1)^\top Q_2 (x_2 - x_1)}{\beta} \left(1 - \sqrt{\frac{\alpha u(t)}{(x_2 - x_1)^\top Q_1 (x_2 - x_1)}}\right)^2. \quad \square$$

From Proposition 2, we deduce that if we set  $\kappa_\alpha = \kappa_\beta = 1$ , then the Pareto front will be independent from the Hessian matrix and will be described by the front of the double-sphere problem:  $u \mapsto (1 - \sqrt{u})^2$ .

We investigate now the general case where the Hessians of the functions  $f_1$  and  $f_2$  are not necessarily proportional. Yet, before digging into the general convex-quadratic problems, we show a result on the shape of the Pareto front of a larger class of bi-objective problems.

**Theorem 1.** *Let  $f_1 : \mathbb{R}^n \mapsto \mathbb{R}$  and  $f_2 : \mathbb{R}^n \mapsto \mathbb{R}$  be strict convex differentiable functions such that the problem  $(f_1, f_2)$  has, as Pareto set, the image of a differentiable function  $\varphi : [0, 1] \mapsto \mathbb{R}^n$ .*

*Assume that: (i)  $f_1 \circ \varphi$  is strictly monotone, (ii)  $\lim_{t \rightarrow 0} \frac{(f_1 \circ \varphi)'(t)}{t} \neq 0$  and (iii)  $\lim_{t \rightarrow 1} \frac{(f_2 \circ \varphi)'(t)}{1-t} \neq 0$ . Then, the Pareto front is a **convex** curve, with **vertical** tangent at  $t = 0$  and horizontal tangent at  $t = 1$ .*

*Proof.* Denote by  $u \stackrel{\text{def}}{=} f_1 \circ \varphi$  and  $v \stackrel{\text{def}}{=} f_2 \circ \varphi$ . Then the Pareto front is described by the parametric equation  $(u(t), v(t))$ , for  $t \in [0, 1]$ . We will show that  $u'v'' - u''v' > 0$  which implies the convexity of the curve.

By linear scalarization (see [8], or weighted sum method in [5]), as in the proof of Proposition 1, we have  $(1-t)\nabla f_1(\varphi(t)) + t\nabla f_2(\varphi(t)) = 0$ . If we take the scalar product of the former equation with  $\varphi'(t)$ , we obtain that

$$(1-t)\langle \nabla f_1(\varphi(t)), \varphi'(t) \rangle + t\langle \nabla f_2(\varphi(t)), \varphi'(t) \rangle = 0. \quad (9)$$

Moreover, for any differentiable function  $f$  with suitable domains,

$$(f \circ \varphi)'(t) = d(f \circ \varphi)_t(1) = df_{\varphi(t)}(d\varphi_t(1)) = \langle \nabla f(\varphi(t)), \varphi'(t) \rangle. \quad (10)$$

Inserting this in (9) shows  $(1-t)(f_1 \circ \varphi)'(t) + t(f_2 \circ \varphi)'(t) = 0$ , which is the same as:

$$(1-t)u'(t) + tv'(t) = 0, \text{ for any } t \in [0, 1]. \quad (11)$$

Since  $\lim_{t \rightarrow 0} \frac{(f_1 \circ \varphi)'(t)}{t}$  exists, (11) implies that:

$$v'(t) = \left(1 - \frac{1}{t}\right)u'(t), \text{ for any } t \in [0, 1]. \quad (12)$$

By deriving (12) and multiplying by  $u'(t)$  in a suitable way, we obtain

$$u'(t)v''(t) = \frac{1}{t^2}u'(t)^2 + \left(1 - \frac{1}{t}\right)u'(t)u''(t), \text{ for any } t \in [0, 1]. \quad (13)$$

Using (12) in (13) gives  $u'(t)v''(t) = \frac{1}{t^2}u'(t)^2 + v'(t)u''(t)$ . Thanks to the assertions on  $f_1 \circ \varphi$ , we have that  $u'(t)v''(t) - u''(t)v'(t) > \frac{1}{t^2}u'(t)^2 > 0$ , for any  $t \in [0, 1]$ . Thus, the Pareto front is a **convex** curve.

Evaluating (11) at  $t = 0$  and at  $t = 1$  implies that  $u'(0) = 0, v'(1) = 0$ . And if we divide (11) by  $t$  (resp.  $1-t$ ) and take the limit to 0 (resp. 1), it follows that  $v'(0) \neq 0$  (resp.  $u'(1) \neq 0$ ). Thereby we also obtain the derivative assumptions on the extremal points.  $\square$



*Remark 1.* Note that the above result about the tangents in the extremal points have additional consequences: according to [1], the assumptions of Theorem 1 imply that the extremal points are never included in any optimal  $\mu$ -distributions of the Hypervolume indicator.

We now deduce the convexity of the Pareto front for convex-quadratic bi-objective problems and characterize the derivatives at the extremes of the front.

**Corollary 3.** *For the problem  $\mathcal{P}$ , the Pareto front is a **convex** curve, with **vertical** tangent at  $(0, f_2(x_1))$  and **horizontal** tangent at  $(f_1(x_2), 0)$ .*

*Proof.* We will show that  $f_1 \circ \varphi$  verifies the assumptions of Theorem 1. From (10) we know that

$$(f_1 \circ \varphi)'(t) = \langle \nabla f_1(\varphi(t)), \varphi'(t) \rangle. \quad (14)$$

In addition,  $\nabla f_1(\varphi(t)) = \frac{2}{\alpha} Q_1(\varphi(t) - x_1)$  and Eq. (4) of Proposition 1 gives  $t[(1-t)Q_1 + tQ_2]\varphi'(t) = Q_1(\varphi(t) - x_1)$ . Multiplying (14) by  $t \in [0, 1]$  shows

$$t(f_1 \circ \varphi)'(t) = \frac{2}{\alpha} \left\langle [(1-t)Q_1 + tQ_2]^{-1} Q_1(\varphi(t) - x_1), Q_1(\varphi(t) - x_1) \right\rangle. \quad (15)$$

Since  $[(1-t)Q_1 + tQ_2]^{-1}$  is a positive definite matrix, then  $t(f_1 \circ \varphi)'(t) \geq 0$ . Let us prove that  $\varphi(t) \neq x_1$ , for  $t \in (0, 1]$ . By contradiction, assume that there exists  $t \in (0, 1]$  such that  $\varphi(t) = x_1$ . Then Equation (3) in Proposition 1 shows that:  $tQ_2(x_2 - \varphi(t)) = (1-t)Q_1(\varphi(t) - x_1) = 0$ , which implies that  $x_2 = \varphi(t) = x_1$ : that is impossible since  $x_1 \neq x_2$ . Hence, by *reductio ad absurdum*,  $\varphi(t) \neq x_1$ , for  $t \in (0, 1]$ . From (15), it follows that

$$(f_1 \circ \varphi)'(t) > 0, \text{ for any } t \in (0, 1]. \quad (16)$$

If we use again the relation from Proposition 1, we obtain  $\lim_{t \rightarrow 0} \frac{Q_1(\varphi(t) - x_1)}{t} = Q_2(x_2 - \varphi(0)) = Q_2(x_2 - x_1)$ . Injecting this result in (15), it follows that:

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{(f_1 \circ \varphi)'(t)}{t} &= \frac{2}{\alpha} \langle Q_1^{-1} Q_2(x_2 - x_1), Q_2(x_2 - x_1) \rangle \\ &> 0, \text{ since } (Q_1^{-1} \text{ is a positive definite matrix}) \end{aligned} \quad (17)$$

In the same way as above, we obtain that

$$\lim_{t \rightarrow 1} \frac{(f_2 \circ \varphi)'(t)}{1-t} = -\frac{2}{\beta} \langle Q_2^{-1} Q_1(x_1 - x_2), Q_1(x_1 - x_2) \rangle < 0. \quad (18)$$

Equations (16), (17), and (18) allow us to apply Theorem 1.  $\square$

We illustrate the previous corollary by taking three random instances of our general problem  $\mathcal{P}$ , with the scalings always chosen as  $\alpha = \beta = \max(f_1(x_2), f_2(x_1))$ . The Pareto fronts are presented in Figure 2. We observe that the Pareto fronts are convex and their derivatives are infinite on the left and zero on the right.

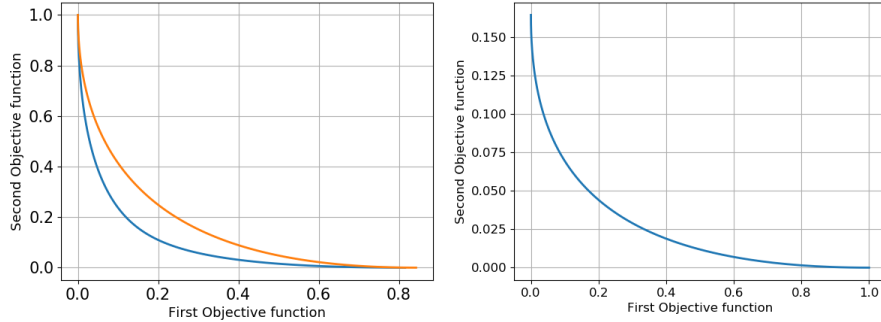


Fig. 2: Left: Two Pareto fronts for  $n = 2$  represented in  $\mathbb{R}^2$  with  $Q_1, Q_2$  randomly sampled and different. Right: Pareto front for  $n = 10$  with matrices given in (7).

### 3 New Classes of bi-objective test functions

Bi-objective problems using convex-quadratic functions have been used to test MO algorithms (see for example [7]). Problems where both Hessian matrices have the same eigenvalues have been used in particular. Yet, test problems considered so far do not explore the full possibilities of properties that can be tested. We therefore extend the test problems from the literature to be able to capture more properties. To do so we present seven classes of bi-objective convex-quadratic problems where the eigenspectra of both Hessian matrices are equal. A natural extension of these classes is to use in each objective different eigenspectra,  $\Delta$ , which leads in general to a nonlinear Pareto set.

The proposed construction parametrizes, apart from search space translations, *all* bi-objective convex-quadratic functions with identical Hessian eigenspectrum in seven classes with increasing difficulty. The particular focus is on problems with a linear Pareto set in five of the seven classes. Some classes represent essentially different problems, hence we do not expect uniform performance over all problems within each class. Independently of the given construction, invariance to search space rotation can be tested by applying an orthogonal transformation to the input argument.

We start from a diagonal matrix  $\Delta$  with positive entries that define a separable convex-quadratic function  $f(x) = \frac{1}{\alpha} x^\top \Delta x$ . For instance,  $\Delta$  can be equal to the identity and we recover the sphere function. If  $\Delta(1, 1) = 1$ ,  $\Delta(n, n) = 10^8$  and  $\Delta(i, i) = 10^4$ , we recover the separable cig-tab function and if  $\Delta(i, i) = 10^{6 \frac{i-1}{n-1}}$ , we recover the separable ellipsoid function.

In the sequel,  $O$  and  $O_2$  denote orthogonal matrices.  $O_1$  is either a permutation matrix, or an orthogonal matrix, depending on the context. The classes of problems proposed are summarized in Table 1 and Table 2.

**The Sep problem classes** We define the **Sep- $k$**  class by considering two separable functions and place the optimum of  $f_1$  in 0 and of  $f_2$  in the  $k^{\text{th}}$  unit vector:  $f_{1,\Delta}^{\text{sep-}k}(x) = \frac{1}{\alpha} (x - x_1)^\top \Delta (x - x_1)$  and  $f_{2,\Delta}^{\text{sep-}k}(x) = \frac{1}{\beta} (x - x_2)^\top \Delta (x - x_2)$ , where  $x_1 = (0, \dots, 0)^\top$  and  $x_2 = (0, \dots, 0, \sqrt{n}, 0, \dots, 0)^\top$  where  $\sqrt{n}$  is at coordinate  $k$ . According to Corollary 1, the Pareto set of this class of problems is the line segment between the optima of the single-objective problems. These problems allow to test the performance on separable problems with a Pareto set aligned with the coordinate axis and check the sensibility with respect to different axes (by varying  $k$ ).

For the **Sep-O** class, we only change the location of the optimum of the second objective by taking  $x_2 = O(1, \dots, 1)^\top$ . If  $O$  has elements  $\notin \{-1, 0, 1\}$ , the Pareto set is not anymore aligned with the coordinate system, but the objectives  $f_1$  and  $f_2$  themselves remain separable. Comparing with class **Sep- $k$** , we can test whether having the Pareto set not aligned with the coordinate axis has an influence on the performance of the algorithm.

For the **Sep-Two-O** class, we define  $f_{1,\Delta}^{\text{sep-Two-O}}(x) = \frac{1}{\alpha} (x - x_1)^\top \Delta (x - x_1)$  and  $f_{2,\Delta}^{\text{sep-Two-O}}(x) = \frac{1}{\beta} (x - x_2)^\top O_1^\top \Delta O_1 (x - x_2)$  where  $O_1$  is a permutation matrix,  $x_1 = (0, \dots, 0)^\top$  and  $x_2 = O(1, \dots, 1)^\top$ . The matrix  $O_1^\top \Delta O_1$  is also diagonal, and thereby each function is separable. Yet the Pareto set is generally not a line segment anymore since we have different Hessian matrices. We can test here the difficulty of having a nonlinear Pareto set on separable functions.

**The One and the One-O problem classes.** We now consider non-separable problems with a line segment as Pareto set. We define  $f_{1,\Delta}^{\text{one}}(x) = \frac{1}{\alpha} (x - x_1)^\top O_1^\top \Delta O_1 (x - x_1)$  and  $f_{2,\Delta}^{\text{one}}(x) = \frac{1}{\alpha} (x - x_2)^\top O_1^\top \Delta O_1 (x - x_2)$ , where  $O_1$  is an orthogonal matrix,  $x_1 = (0, \dots, 0)^\top$  and  $x_2 = (1, \dots, 1)^\top$ . We replace  $x_2$  by  $Ox_2$  to obtain the **One-O** problems.

These two problem classes allow to test the performance on non-separable problems that have a line segment as Pareto set comparing in particular to class **Sep-O**. Up to a reformulation, the problems ELLI1 and CIGTAB1 from [7] are from the **One-O** problem class. Generally, we do not expect different performance over all problems of the **One** vs the **One-O** class.

**The Two and the Two-O problem classes.** For these classes, we rotate each function independently; then the Pareto set is generally not a line segment anymore. We define  $f_{1,\Delta}^{\text{two}}(x) = \frac{1}{\alpha} (x - x_1)^\top O_1^\top \Delta O_1 (x - x_1)$  and  $f_{2,\Delta}^{\text{two}}(x) = \frac{1}{\alpha} (x - x_2)^\top O_2^\top \Delta O_2 (x - x_2)$ , with  $O_1$  orthogoanal,  $x_1 = (0, \dots, 0)^\top$  and  $x_2 = (1, \dots, 1)^\top$ . The corresponding **O** problems are obtained with  $Ox_2$  replacing  $x_2$ . All presented classes are subsets of the **Two-O** class. ELLI2 and CIGTAB2 from [7] fall within the **Two-O** class. Compared to the respective **One** classes, we can test the impact of having a nonlinear Pareto set.

Table 1: Unconstrained quadratic bi-objective test problems:  $\Delta$  is a positive diagonal matrix,  $O$  is an orthogonal matrix,  $O_1$  is a permutation matrix.

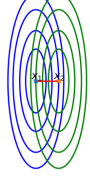
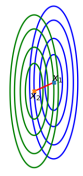
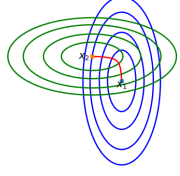
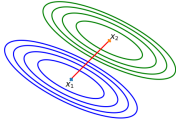
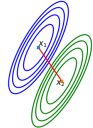
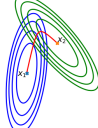
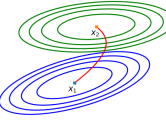
	Sep- $k$	Sep-O	Sep-Two-O
$x_1$	$(0, \dots, 0)^\top$	$(0, \dots, 0)^\top$	$(0, \dots, 0)^\top$
$x_2$	$\underbrace{(0, \dots, \sqrt{n}, \dots, 0)^\top}_{\sqrt{n} \text{ is at row } k}$	$O(1, \dots, 1)^\top$	$O(1, \dots, 1)^\top$
$Q_1, Q_2$	$\Delta, \Delta$	$\Delta, \Delta$	$\Delta, O_1^\top \Delta O_1$
Level sets			

Table 2: Unconstrained quadratic bi-objective test problems:  $\Delta$  is a positive diagonal matrix,  $O, O_1$  and  $O_2$  are three independent orthogonal matrices.

	One	One-O	Two	Two-O
$x_1$	$(0, \dots, 0)^\top$	$(0, \dots, 0)^\top$	$(0, \dots, 0)^\top$	$(0, \dots, 0)^\top$
$x_2$	$(1, \dots, 1)^\top$	$O(1, \dots, 1)^\top$	$(1, \dots, 1)^\top$	$O(1, \dots, 1)^\top$
$Q_1, Q_2$	$O_1^\top \Delta O_1, O_1^\top \Delta O_1$	$O_1^\top \Delta O_1, O_1^\top \Delta O_1$	$O_1^\top \Delta O_1, O_2^\top \Delta O_2$	$O_1^\top \Delta O_1, O_2^\top \Delta O_2$
Level sets				

## 4 Summary

We have presented an analytic description of the Pareto set for quadratic bi-objective problems. We have shown that the Pareto set is a line segment when both objectives have proportional Hessian matrices and deduced a complete description of the Pareto front in that case. We have also proven that some properties of the double-sphere are conserved in a wider framework that includes the general quadratic bi-objective problem: the Pareto front remains convex and its vertical and horizontal tangents remain at the extremal points of the front. Such assumptions on the derivatives imply that when looking at the optimal  $\mu$ -distributions of the Hypervolume indicator, the extremal points are always excluded [1]. We have also presented several classes of problems, where each one tests a specific capability of the multiobjective algorithm.

**Acknowledgments** The Ph.D. of Cheikh Touré is funded by Inria and Storengy. We particularly thank F. Huguet and A. Lange from Storengy for their strong support, practical ideas and expertise.

## References

1. Auger, A., Bader, J., Brockhoff, D., Zitzler, E.: Theory of the hypervolume indicator: optimal  $\mu$ -distributions and the choice of the reference point. In: Foundations of Genetic Algorithms (FOGA 2009). pp. 87–102. ACM (2009)
2. Brockhoff, D., Tran, T.D., Hansen, N.: Benchmarking Numerical Multiobjective Optimizers Revisited. In: Genetic and Evolutionary Computation Conference (GECCO 2015). pp. 639–646. ACM (2015). <https://doi.org/10.1145/2739480.2754777>
3. Collange, G., Delattre, N., Hansen, N., Quinquis, I., Schoenauer, M.: Multidisciplinary Optimisation in the Design of Future Space Launchers. In: Multidisciplinary Design Optimization in Computational Mechanics, pp. 487–496. Wiley (2010)
4. Emmerich, M.T., Deutz, A.H.: Test problems based on Lamé superspheres. In: Evolutionary Multi-Criterion Optimization (EMO 2007). pp. 922–936. Springer (2007)
5. Grodzewich, O., Romanko, O.: Normalization and other topics in multi-objective optimization. In: Proceedings of the Fields-MITACS Industrial Problems Workshop, 2006. pp. 89–101. Fields-MITACS (2006)
6. Hansen, N., Ostermeier, A.: Completely derandomized self-adaptation in evolution strategies. *Evolutionary Computation* **9**(2), 159–195 (2001)
7. Igel, C., Hansen, N., Roth, S.: Covariance matrix adaptation for multi-objective optimization. *Evolutionary Computation* **15**(1), 1–28 (2007)
8. Jahn, J.: Vector optimization: Theory, applications and extensions. 2004
9. Kerschke, P., Wang, H., Preuss, M., Grimme, C., Deutz, A., Trautmann, H., Emmerich, M.: Search dynamics on multimodal multi-objective problems. *Evolutionary computation* pp. 1–30 (2018)
10. Kerschke, P., Wang, H., Preuss, M., Grimme, C., Deutz, A., Trautmann, H., Emmerich, M.: Towards analyzing multimodality of continuous multiobjective landscapes. In: Parallel Problem Solving from Nature (PPSN 2016). pp. 962–972. Springer (2016)
11. Nocedal, J., Wright, S.J.: Numerical optimization (2006)
12. Powell, M.J.: The NEWUOA software for unconstrained optimization without derivatives. Tech. Rep. DAMTP 2004/NA05, CMS, University of Cambridge, Cambridge CB3 0WA, UK (November 2004)
13. Tušar, T., Brockhoff, D., Hansen, N., Auger, A.: COCO: the bi-objective black box optimization benchmarking (bbob-biobj) test suite. *CoRR* **abs/1604.00359** (2016), <http://arxiv.org/abs/1604.00359>