Inexact scalarization proximal methods for multiobjective quasiconvex minimization on Hadamard manifolds.

E.A. Papa Quiroz, N. Baygorrea, and N. Maculan

Federal University of Rio de Janeiro
PESC - COPPE - UFRJ
PO Box 68511, Rio de Janeiro, CEP 21941-972, Brazil.
erikpapa@gmail.com, nbaygorrea@gmail.com, maculan@cos.ufrj.br

Abstract

In this paper we extend naturally the scalarization proximal point method to solve multiobjective unconstrained minimization problems, proposed by Apolinario et al.[1], from Euclidean spaces to Hadamard manifolds for locally Lipschitz and quasiconvex vector objective functions. Moreover, we present a convergence analysis, under some mild assumptions on the multiobjective function, for two inexact variants of the scalarization proximal point algorithm for this kind of functions. In this sense, strong convergence of all the sequences produced by the methods are obtained. Indeed, each accumulation point, of any sequence generated by these algorithm, is a Pareto critical point for the multiobjective function.

Keywords: Proximal point method, quasiconvex function, Hadamard manifolds, multiobjetive optimization, pareto optimality.

1 Introduction

In this paper, we consider the class of problems known as multiobjective quasiconvex optimization on Hadamard manifolds, which minimize a set of quasiconvex objective functions on Hadamard manifolds. The multiobjective optimization problem, also called multicriteria optimization, can be defined as the problem of finding a vector of decision variables which satisfies constraints and optimizes a vector function who elements represent the objective functions. Multiobjective optimization models have many significant applications in decision-making problems such as economic theory, management science and engineering design. On the other hand, the class of quasiconvex functions has many applications for real life problems. For instance, usually, in Economy, utility functions are quasiconcave functions; see [21, 22].

Nevertheless, several methods were proposed to solve multiobjective optimization problems. We can cite, for instance, the steepest descent method for multiobjective optimization was dealt with in [18], an extension of the projective gradient method to the case of convex constrained vector optimization can be found in [20], a interior-point method was proposed for solving convex multiobjective problem in [18], the proximal point scalarization methods for multiobjective optimization can be found in [23, 32], a novel multiobjective derivative-free methodology which does not aggregate any of the objective functions was proposed by Custodio et al[13], Bonnel et al.[10] generalized the classical results of Rockafellar[33] from the scalar case to the vector case, Ceng and Yao [30] generalized the results of Bonnel et al.[10] to the approximate case

For the case multiobjective quasiconvex minimization, the proximal point scalarization methods in [1], a subgradient type method for nonsmooth unconstrained multiobjective optimization where the objective function is componentwise essentially quasiconvex and Lipschitz continuous is proposed by da Cruz Neto et al.[14], the projected gradient method for solving the problem of finding a Pareto optimum in Bello Cruz et al.[5], a method for resolving multiple objective goal programming problems with quasi-convex linear penalty functions is presented by Li and Yu[30]. In the last years researchers began the study of the multiobjective optimization problems on Riemannian manifolds, see [7, 8, 9, 25].

In this paper, we are interesed in solving the following multiobjective optimization problem:

$$\min\{F(x) \mid x \in M\}, \tag{1.1}$$

where M is a Hadamard manifold, which is a complete simply connected Riemannian manifold with non-positive sectional curvature, and $F : M \to I\!R^m$ with each $F_i : M \to I\!R \cup \{+\infty\}$, for $i = 1, \cdots, m$, is a locally Lipschitz and quasiconvex (non-necessarily differentiable) vector function.

A popular strategy for solving (1.1) is the scalarization approach. So, in this paper we propose an extension and generalization of the scalarization proximal point method, proposed by Apolinario et al.[1], for solving multiobjective unconstrained minimization problems with locally Lipschitz and quasiconvex vector functions. The two differences between the algorithm studied in this paper and existing scalarization proximal methods in [1] is the use of more general topological spaces (namely Hadamard manifolds), requiring the use of tools of riemannian geometry, and also an extension, in order to make the algorithm more implementable, of the analysis for the cases where inexact solution holds. In fact, we present and analyze the convergence of two inexact variants of the proximal point algorithm for locally Lipschitz and quasiconvex multiobjective functions, which were already studied by Baygorrea et al.[3] for proper, lower semicontinuous and scalar-valued quasiconvex functions on Hadamard manifolds whose convergence rate of these algorithms was analized in [4] The purpose of this paper is to prove that the sequence generated by these algorithms, for solving (1.1) on Hadamard manifolds, is well-defined and that global convergence to Pareto-Clarke critical points is also obtained. We recall that an algorithm is said to be globally convergent if the algorithm is guaranteed to generate a sequence of points converging to a critical point for arbitrary starting points.

The outline of this paper is as follows: Section 2, we recall some definitions and results about Riemannian geometry, quasiconvex analysis, Fréchet subdifferential and Clarke's generalized ones. In Section 3, we study some sufficient conditions to guarantee existence of global minimizer for single-valued functions on Riemannian manifolds. In Section 4, we state the quasiconvex multiobjective optimization problem and the inexact scalarization proximal point method for solving the problem (1.1). In Section 5, we analyze two inexact variants of the algorithm mentioned and convergence results are studied under certain given hypothesis on error criteria. Finally, in Section 6, some final remarks are given.

2 Basic results and properties

In this section we recall some fundamental properties and notations on Riemannian manifolds. Those basic facts can be seen, for example, in do Carmo[16], Sakai[34], Udriste[38], Rapcsák[31] and references therein.

Let M be a differential manifold with finite dimension n. We denote by T_xM the tangent space of M at x and $TM = \bigcup_{x\in M} T_xM$. T_xM is a linear space and has the same dimension of M. Because we restrict ourselves to real manifolds, T_xM is isomorphic to $I\!R^n$. If M is endowed with a Riemannian metric g, then M is a Riemannian manifold and we denote it by (M, G) or only by M when no confusion can arise, where G denotes the matrix representation of the metric g. The inner product of two vectors $u, v \in T_xM$ is written as $\langle u, v\rangle_x := g_x(u, v)$, where g_x is the metrics at point x. The norm of a vector $v \in T_xM$ is set by $||v||_x := \langle v, v\rangle_x^{1/2}$. If there is no confusion we denote $\langle , \rangle = \langle , \rangle_x$ and $||.|| = ||.||_x$. The metrics can be used to define the length of a piecewise smooth curve $\alpha : [t_0, t_1] \to M$ joining $\alpha(t_0) = p'$ to $\alpha(t_1) = p$ through $L(\alpha) = \int_{t_0}^{t_1} \|\alpha'(t)\|_{\alpha(t)}dt$. Minimizing this length functional over the set of all curves we obtain a Riemannian distance $d(p', p)$ which induces the original topology on M.

Given any open set U in a smooth manifold M. We denote by $C^\infty(U)$ the algebra of all smooth real-valued functions on U. A vector field $X : C^\infty(U) \to C^\infty(U)$ is C^∞ on U provided so is $X(g)$ for any $g \in C^\infty(U)$. The collection of all C^∞ vector fields on M is denoted by $V^\infty(M)$. Given two vector fields X and Y on M, the covariant derivative of Y in the direction X is denoted by $\nabla_X Y$. In this paper ∇ is the Levi-Civita connection associated to (M, G). This connection defines an unique covariant derivative D/dt, where, for each vector field X, along a smooth curve $\alpha : [t_0, t_1] \to M$, another vector field is obtained, denoted by DX/dt. The parallel transport along α from $\alpha(t_0)$ to $\alpha(t_1)$, denoted by P_{α,t_0,t_1}, is an application $P_{\alpha,t_0,t_1} : T_{\alpha(t_0)}M \to T_{\alpha(t_1)}M$ defined by $P_{\alpha,t_0,t_1}(v) = X(t_1)$ where X is the unique vector field along α so that $DX/dt = 0$ and $X(t_0) = v$. Since ∇ is a Riemannian connection, P_{α,t_0,t_1} is a linear isometry, furthermore $P^{-1}_{\alpha,t_0,t_1} = P_{\alpha,t_1,t_0}$ and $P_{\alpha,t_0,t_1} = P_{\alpha,t,t_1}P_{\alpha,t_0,t}$, for all $t \in [t_0, t_1]$. A curve $\gamma : I \to M$ is called a geodesic if $D\gamma'/dt = 0$.

A Riemannian manifold is complete if its geodesics are defined for any value of $t \in I\!R$. Let $x \in M$, the exponential map $\exp_x : T_xM \to M$ is defined $\exp_x(v) = \gamma(1, x, v)$, for each $x \in M$.

If M is complete, then $\exp_x$ is defined for all $v \in T_xM$. Besides, there is a minimal geodesic (its length is equal to the distance between the extremes). Take $x \in M$. Let $\exp_x^{-1} : M \to T_xM$ be the inverse of the exponential maps which is also C^∞. Note that $d(x, y) = \|\exp_y^{-1} x\|$.

Complete simply-connected Riemannian manifolds with nonpositive curvature are called *Hadamard manifolds*. Some examples of Hadamard manifolds may be found in Section 4 of Papa Quiroz and Oliveira [25]. In the sequel of this paper, we consider (M, g) as a Hadamard manifolds.

Theorem 2.1 *Given a geodesic triangle $[x, y, z]$ in a Hadamard manifold, it holds that:*

$$d^2(x, z) + d^2(z, y) - 2\langle \exp_z^{-1} x, \exp_z^{-1} y\rangle \leq d^2(x, y),$$

Proof. See Sakai [34, Proposition 4.5]. ∎

Theorem 2.2 *Let M be a Hadamard manifold and $y \in M$ be a fixed point. Then, the function $g(x) = d^2(x, y)$ is strictly convex and its gradient at $x \in M$ is $\operatorname{grad} g(x) = -2\exp_x^{-1} y$.*

Proof. See Sakai[34] and Shiga[35]. ∎

Let $D \subset M$ be a convex set, and $x \in D$. The normal cone to D at $x \in D$, denoted by $\mathcal{N}_D(x)$, is defined by

$$\mathcal{N}_D(x) := \{v \in T_xM \mid \langle v, \exp_x^{-1} y\rangle \leq 0, \quad \forall\, y \in D\}.$$

Assume that $D = \{x\}$, then $\mathcal{N}_D(x) = T_xM$.

Given an extended real valued function $f : M \to I\!R \cup \{+\infty\}$, we denote its domain by $\operatorname{dom}(f) := \{x \in M \mid f(x) < +\infty\}$. It is said to be a proper function if $\operatorname{dom}(f) \neq \emptyset$ and for all $x \in \operatorname{dom}(f)$ we have $f(x) > -\infty$.

Let $f : M \to I\!R \cup \{+\infty\}$ be a proper function, it is called quasiconvex if for all $x, y \in dom(f)$ and $t \in [0, 1]$, it holds that

$$f(\gamma(t)) \leq \max\{f(x), f(y)\},$$

for the geodesic $\gamma : [0, 1] \to I\!R$, so that $\gamma(0) = x$ and $\gamma(1) = y$.

A function $f : M \longrightarrow I\!R \cup \{+\infty\}$ is said to be lower semicontinuous at a point $\bar{x} \in M$ if, for any sequence $\{x^k\}$ convergent to $\bar{x}$, we obtain

$$\liminf_{k\to\infty} f(x^k) \geq f(\bar{x}).$$

The function f is lower semicontinuous on M if it is lower semicontinuous at each point in M.

A function $f : M \to I\!R \cup \{+\infty\}$ is said to be locally Lipschitz if for every $z \in M$ there exists a constant $L := L_z > 0$ such that for every x, y in a neigbourhood of z we have

$$|f(x) - f(y)| \leq Ld(x, y). \tag{2.2}$$

Now we start with some definitions of nonsmooth analysis on Riemannian manifolds. See, for instance, Hosseini and Pouryayevali[26], Azagra et al.[2], Ledyaev and Zhu[29].

Let $f : M \to I\!R \cup \{+\infty\}$ be a lower semicontinuous function with $dom(f) \neq \emptyset$. The Fréchet subdifferential of f at a point $x \in dom(f)$ is defined as the set $\partial^F f(x)$ of all $s \in T_x M$ with the property that

$$\liminf_{\substack{u \to x \\ u \neq x}} \frac{1}{d(x, u)} \left[f(u) - f(x) - \langle s, \exp_x^{-1} u \rangle_x \right] \geq 0.$$

Clearly, $s \in \partial^F f(x)$ iff for each $\eta > 0$, there is $\epsilon > 0$ such that

$$\langle s, \exp_x^{-1} u \rangle_x \leq f(u) - f(x) + \eta d(x, u), \quad \text{for all } u \in B(x, \epsilon).$$

If $x \notin dom(f)$ then we set $\partial^F f(x) = \emptyset$.

Theorem 2.3 *[29, Theorem 4.4] Let $f_1, \ldots, f_L$ be lower semicontinuous functions on a manifold M and let $s \in \partial^F (f_1 + \cdots + f_L)(\bar{x})$. Then for any $\epsilon > 0$, any neigbourhood of $\bar{x}$ and any $v \in V^\infty(M)$ there exist $x_i \in B(\bar{x}, \epsilon)$ and $s_i \in \partial^F f_i(x_i)$, $i = 1, 2, \cdots, L$, satisfying $|f_i(x_i) - f_i(\bar{x})| < \epsilon$ and*

$$|\langle s, v \rangle_{\bar{x}} - \sum_{i=1}^{L} \langle s_i, v \rangle_{x_i}| < \epsilon. \tag{2.3}$$

Lemma 2.1 *Let $f, g : M \to I\!R \cup \{+\infty\}$, be proper functions such that f is lower semicontinuous at $\bar{x} \in dom(f) \cap dom(g)$ and g is locally Lipschitz at $\bar{x} \in dom(f) \cap dom(g)$. Then*

$$\partial^F (f + g)(\bar{x}) = \partial^F f(\bar{x}) + \partial^F g(\bar{x}).$$

Proof. The inclusion " $\supset$" is easily obtained as consequence of limit inferior property. To prove the converse, pick any $s \in \partial^F (f + g)(\bar{x})$ and employ Theorem 2.3 of sum rule property for lower semicontinuous functions. For any $\epsilon > 0$, there exist x_1, $x_2 \in B(\bar{x}, \epsilon)$, $s_1 \in \partial^F f(x_1)$ such that $|f(x_1) - f(\bar{x})| < \epsilon$ and $s_2 \in \partial^F g(x_2)$ such that $|g(x_2) - g(\bar{x})| < \epsilon$ satisfying relation (2.3) for any $v \in V^\infty(M)$.

Furthermore, since $s_2 \in \partial^F g(x_2)$ and for any $\epsilon > 0$ find $\tau = \epsilon - d(x_2, \bar{x})$ such that

$$\langle s_2, \exp_{x_2}^{-1} u \rangle_{x_2} \leq g(u) - g(x_2) + \epsilon d(u, x_2), \tag{2.4}$$

whenever $d(x_2, u) < \tau$.

On the other hand, we can take $v := \exp_{x_2}^{-1}u \in T_{x_2}M$. So, relation (2.3) holds

$$\begin{aligned}\epsilon + \langle s_2, v\rangle_{x_2} &> \langle s, \mathcal{P}_{x_2,\bar{x}}v\rangle_{\bar{x}} - \langle s_1, \mathcal{P}_{x_2,x_1}v\rangle_{x_1} \\ &= \langle s, \mathcal{P}_{x_2,\bar{x}}v\rangle_{\bar{x}} - \langle \mathcal{P}_{x_1,\bar{x}}s_1, \mathcal{P}_{x_1,\bar{x}}\mathcal{P}_{x_2,x_1}v\rangle_{x_1} \\ &= \langle s - \mathcal{P}_{x_1,\bar{x}}s_1, \mathcal{P}_{x_2,\bar{x}}v\rangle_{\bar{x}}. \qquad (2.5)\end{aligned}$$

Denote $z := \exp_{\bar{x}}\mathcal{P}_{x_2,\bar{x}}v$. So, it follows from (2.4) and (2.5), for all $z \in B(\bar{x}, \tau)$, that

$$\begin{aligned}\langle s - \mathcal{P}_{x_1,\bar{x}}\, s_1, \exp_{\bar{x}}^{-1}z\rangle_{\bar{x}} &< g(u) - g(x_2) + \epsilon d(u, x_2) + \epsilon \\ &= g(z) - g(\bar{x}) + g(\bar{x}) - g(z) + g(u) - g(x_2) + \epsilon d(u, x_2) + \epsilon \quad (2.6)\end{aligned}$$

Since g is a locally Lipschitz function, there exist positive constants $L_{\bar{x}}$ and L_{x_2} such that relation (2.6) yields

$$\begin{aligned}\langle s - \mathcal{P}_{x_1,\bar{x}}\, s_1, \exp_{\bar{x}}^{-1}z\rangle_{\bar{x}} &< g(z) - g(\bar{x}) + L_{\bar{x}}d(\bar{x}, z) + (L_{x_2} + \epsilon)d(u, x_2) + \epsilon \\ &= g(z) - g(\bar{x}) + (L_{\bar{x}} + L_{x_2} + \epsilon)d(\bar{x}, z) + \epsilon. \qquad (2.7)\end{aligned}$$

The last equality is obtained since $d(u, x_2) = d(\bar{x}, z)$.

Therefore, for $\epsilon > 0$ sufficiently small, relation (2.7) implies

$$\langle s - \mathcal{P}_{x_1,\bar{x}}s_1, z - \bar{x}\rangle \leq g(z) - g(\bar{x}) + (L_{\bar{x}} + L_{x_2})d(\bar{x}, z),$$

whenever $z \in B(\bar{x}, \tau)$. So, $s - \mathcal{P}_{x_1,\bar{x}}s_1 \in \partial^F g(\bar{x})$. This implies that $s \in \mathcal{P}_{x_1,\bar{x}}s_1 + \partial^F g(\bar{x})$. Moreover, in this conditions, it can be proved that if $s_1 \in \partial^F f(x_1)$ then $\mathcal{P}_{x_1,\bar{x}}s_1 \in \partial^F f(\bar{x})$. Thus, $s \in \partial^F f(\bar{x}) + \partial^F g(\bar{x})$ and ends the proof of this result. ■

Let $f : M \to I\!R \cup \{+\infty\}$ be a locally Lipschitz proper function, with $dom(f) \neq \emptyset$, at $x \in dom(f)$ and $d \in T_xM$. The Clarke's generalized directional derivate of f at $x \in M$ in the direction $d \in T_xM$, denoted by $f^\circ(x, d)$, is defined as

$$f^\circ(x, d) := \limsup_{\substack{u \to x \\ t \searrow 0}} \frac{f(\exp_u t(D\exp_x)_{\exp_x^{-1}u}d) - f(u)}{t},$$

where $(D\exp_x)_{\exp_x^{-1}u} : T_{\exp_x^{-1}u}(T_xM) \simeq T_xM \to T_uM$ is the differential of exponential mapping at $\exp_x^{-1}u$. The Clarke's generalized subdifferential of f at $x \in M$, is the subset $\partial^\circ f(x)$ of $T_xM^* \simeq T_xM$ defined by

$$\partial^\circ f(x) := \{s \in T_xM \mid \langle s, d\rangle_x \leq f^\circ(x, d),\ \forall\, d \in T_xM\}.$$

Note that for a locally Lipschitz function $f : M \to I\!R \cup \{+\infty\}$, the set $\partial^\circ f(x)$ is a nonempty closed convex subset of T_xM for every $x \in M$.

Remark 2.1 *We recall that if $f : M \to I\!R \cup \{+\infty\}$ is a locally Lipschitzian proper function then, $\partial^F f(x) \subseteq \partial^\circ f(x)$ for all $x \in dom(f)$. It is due to the fact that $\partial^F f(x) = \partial^F (f \circ \exp_x)(0_x)$, $\partial^\circ f(x) = \partial^\circ (f \circ \exp_x)(0_x)$ and by Correa et al.[11, property 2.7].*

Lemma 2.2 *Let $f_i : M \to I\!R \cup \{+\infty\}$, $i \in \mathcal{I} \subset I\!N$, be locally Lipschitz proper functions. Then, for all $d \in T_xM$,*

$$
\begin{aligned}
(\sum_{i\in\mathcal{I}} f_i)^\circ(x,d) &\leq \sum_{i\in\mathcal{I}} f_i^\circ(x,d) && (2.8)\\
(\alpha f_i)^\circ(x,d) &= \alpha(f_i^\circ(x,d)), \quad \forall\, \alpha \geq 0. && (2.9)\\
f_i^\circ(x,\alpha d) &= \alpha f_i^\circ(x,d), \quad \forall\, \alpha \geq 0. && (2.10)
\end{aligned}
$$

Proof. Relation (2.8) follows by using simple upper limit properties together with the definition of Clarke's directional derivative. Relation (2.9) is immediate. Finally, relation (2.10) means that function $d \mapsto f^\circ(x,d)$ is positively homogeneous. Its proof is trivial for every $d \in T_xM$ and $\alpha \geq 0$. Indeed, for $\alpha > 0$ we get

$$
\begin{aligned}
\alpha f^\circ(x,d) &= \limsup_{\substack{u\to x\\ t\searrow 0}} \frac{\alpha f(\exp_u t(D\exp_x)_{\exp_x^{-1}u} d) - \alpha f(u)}{t}\\
&= \limsup_{\substack{u\to x\\ (t/\alpha)\searrow 0}} \frac{f(\exp_u (t/\alpha)(D\exp_x)_{\exp_x^{-1}u} \alpha d) - f(u)}{t/\alpha}\\
&= f^\circ(x,\alpha d),
\end{aligned}
$$

and it is immediate as $\alpha = 0$. ■

Lemma 2.3 *Let $f_i : M \to I\!R \cup \{+\infty\}, i \in \mathcal{I} \subset I\!N$, be locally Lipschitz functions. Then*

$$
\begin{aligned}
\partial^\circ(\sum_{i\in\mathcal{I}} f_i)(x) &\subset \sum_{i\in\mathcal{I}} \partial^\circ f_i(x), && (2.11)\\
\partial^\circ(\alpha f_i)(x) &= \alpha \partial^\circ f_i(x) && (2.12)
\end{aligned}
$$

Proof. Using Lemma 2.2, relation (2.11) follows immediately. To show relation (2.12), consider $s \in \partial^\circ(\alpha f_i)(x)$. By relation (2.9), for all $d \in T_xM$ we have $\langle s, d\rangle \leq \alpha f_i^\circ(x,d)$. So $s/\alpha \in \partial^\circ f(x)$ is given by the nonnegative number α. Therefore, $s \in \alpha\, \partial^\circ f(x)$. The converse can be proved in the same way. ■

Lemma 2.4 *Let $\Omega \subset M$ be an open set. If $f : M \to I\!R \cup \{+\infty\}$ is locally Lipschitz and $g : M \to I\!R$ is convex and differentiable on Ω, then*

$$
\partial^\circ(f+g)(x) = \partial^\circ f(x) + grad\, g(x), \qquad \forall\, x \in \Omega
$$

Proof. See Bento et al.[6], Lemma 3.1 ■

Proposition 2.1 *Let $f : M \to I\!R \cup \{+\infty\}$ be a locally Lipschitz function on M and $g : I\!R \to I\!R$ be a continuously differentiable function on $I\!R$. Then, for any $\bar{x} \in dom(f)$, it holds*

$$(g \circ f)^{\circ}(\bar{x}, d) = g'(f(\bar{x}))f^{\circ}(\bar{x}, d), \qquad \forall\ d \in T_{\bar{x}}M$$

Proof. Let $x \in dom(f)$ such that x converges to $\bar{x}$. Let $\tilde{d} := (D\exp_{\bar{x}})_{\exp_{\bar{x}}^{-1}x} d \in T_x M$, for any $d \in T_x M$. Consider $\gamma : I\!R \to M$ be a geodesic such that $\gamma(x, 1, t\tilde{d}) := \exp_x t\tilde{d} \in M$ with $t > 0$. As $t \searrow 0$, $\gamma(x, 1, t\tilde{d})$ converges to x. As x converges to $\bar{x}$, $\lim_{x \to \bar{x}} (D\exp_{\bar{x}})_{\exp_{\bar{x}}^{-1}x} d = (D\exp_{\bar{x}})_0 d = d$, so $\gamma(x, 1, t\tilde{d})$ converges to $\bar{x}$.

Define

$$r(y) := \begin{cases} \frac{g(y) - g(f(x))}{y - f(\bar{x})} - g'(f(x)), & y \neq f(x) \\ 0, & y = f(x) \end{cases}$$

Since g is continuously differentiable, it follows that

$$\lim_{y \to f(x)} r(y) = 0.$$

Thus, we have

$$g(y) - g(f(x)) = g'(f(x))(y - f(x)) + r(y)(y - f(x)). \tag{2.13}$$

Take $y := f(\gamma(x, 1, t\tilde{d}))$. Since f is a locally Lipschitz function, y converges to $f(\bar{x})$ as $x \to \bar{x}$ and $t \searrow 0$. Moreover, $\lim r(y) = 0$ as $y \to f(\bar{x})$.

Applying now limsup to relation (2.13), we get

$$\begin{aligned} (g \circ f)^{\circ}(\bar{x}, d) &= \limsup_{\substack{x \to \bar{x} \\ t \searrow 0}} \frac{(g \circ f)(\gamma(x, 1, t\tilde{d})) - (g \circ f)(x)}{t} && (2.14) \\ &= g'(f(\bar{x})) \limsup_{\substack{x \to \bar{x} \\ t \searrow 0}} \frac{f(\gamma(x, 1, t\tilde{d})) - f(x)}{t} \\ &= g'(f(\bar{x}))f^{\circ}(\bar{x}, d). && (2.15) \end{aligned}$$

The result follows immediately from the definition of the Clarke's generalized subdifferential to the left-hand side of the relation (2.15). ∎

Lemma 2.5 *Let M be a Hadamard manifold and let $f : M \to I\!R \cup \{+\infty\}$ be a proper quasiconvex locally Lipschitz function on M. If $g \in \partial^{\circ} f(x)$ and $\langle g, \exp_x^{-1} z \rangle > 0$, then $f(x) \leq f(z)$.*

Proof. See Papa Quiroz and Oliveira [24] ∎

The following notations are used throughout our presentation. We denote $\mathcal{I} = \{1, \cdots, m\}$. Consider the (multiobjective) function $F := (F_1, \ldots, F_m)$, for every $F_i : M \to I\!R \cup \{+\infty\}$ with $i \in \mathcal{I}$. It is said to be componentwise quasiconvex, or simply quasiconvex, if all component functions $F_i, i \in \mathcal{I}$, are quasiconvex. Similarly, it is locally Lipschitz if all component functions are ones. Moreover, the inequality sign $\preceq$ between vectors is to be understood in a componentwise sense, that is, for $x, y \in M$, $F(x) \preceq F(y)$ iff $F_i(x) \leq F_i(y)$ for all $i \in \mathcal{I}$.

We say $x \in M$ to be a Pareto-Clarke critical point of a multiobjective function F if for all $d \in T_xM$ there exists $i_0 := i_0(d) \in \mathcal{I}$ such that $F^{\circ}_{i_0}(x, d) \geq 0$.

3 Existence of minimizers for single-valued functions on Riemannian manifolds

The following theorem ensures the existence of global minima for a proper, lower semicontinuous, extended and single-valued function with compact domain. We recall the problem of determining the global minimum solution set is called the global optimization problems.

Theorem 3.1 *Let $f : M \longrightarrow I\!R \cup \{+\infty\}$ be a proper lower semicontinuous function. If* $\operatorname{dom}(f)$ *is compact, then f achieves its global minimum over M.*

Proof. We first show that f is bounded from below. To this end, assume for the contradiction that there exists a sequence $\{x^k\} \subset \operatorname{dom}(f)$ such that

$$\lim_{k\to+\infty} f(x^k) = -\infty. \tag{3.16}$$

Since $\operatorname{dom}(f)$ is compact, there exist $\hat{x} \in dom(f)$ and a subsequence $\{x^{k_j}\} \subset \{x^k\}$ such that $\lim_{j\to+\infty} x^{k_j} = \widehat{x}$. Because f is lower semicontinuous, we get

$$\liminf_{j\to+\infty} f(x^{k_j}) \geq f(\widehat{x}),$$

which contradicts relation (3.16). Thus, f is bounded from below on M. It implies that there exists $f^* \in I\!R$ such that $f^* := inf\{f(x) \mid x \in M\}$. From the infimum property, there exists a subsequence $\{x^k\} \subset \operatorname{dom}(f)$ such that $\lim_{k\to\infty} f(x^k) = f^*$.

From the compactness of $\operatorname{dom}(f)$, there exist $\bar{x} \in dom(f)$ and $\{x^{k_j}\} \subseteq \{x^k\}$ such that $\lim_{j\to\infty} x^{k_j} = \bar{x}$. From the lower semicontinuity of f, it holds

$$\liminf_{j\to\infty} f(x^{k_j}) \geq f(\bar{x}). \tag{3.17}$$

Since $\{f(x^k)\}$ converge to f^*, then the subsequence $\{f(x^{k_j})\}$ converges to f^*, and from (3.17) we have $f^* \geq f(\bar{x})$, Therefore, $\bar{x}$ is a global minimum point of f on M. ■

Remark 3.1 *Note that compactness is a very strong condition to ensure existence of a global minimum. Indeed, consider the following problem*

$$\min\{x^2 \mid x \in I\!R\} \tag{3.18}$$

In this case, $M = I\!R$ is a 1-dimensional Euclidean space. Note that $\operatorname{dom}(f) = I\!R$ *is no compact but the problem* (3.18) *has a global minimum at* 0.

In order to give a weak form of the Theorem 3.1, we define, for $\alpha \in I\!R$, the following set

$$L_f(\alpha) := \{x \in M \mid f(x) \leq \alpha\},$$

which is called *lower level set* of f.

Corollary 3.1 *Let $f : M \to I\!R \cup \{+\infty\}$ where M is a Riemannian manifolds. If $L_f(\alpha)$ is a nonempty compact set, for any $\alpha \in I\!R$, and f is lower semicontinuous on $L_f(\alpha)$, then there exists a global minimum point of f on M.*

Proof. Let an arbitrary point $x \in M$. If $x \in L_f(\alpha)$, by Theorem 3.1, it follows that there exists a global minimizer of f on $L_f(\alpha)$. It can be taken as $\bar{x} \in L_f(\alpha)$ such that $f(\bar{x}) \leq f(x)$, for all $x \in L_f(\alpha)$. Otherwise, it implies that $f(x) > \alpha \geq f(\bar{x})$. So, $f(\bar{x}) < f(x)$. Therefore, in both cases, we conclude that $\bar{x} \in dom(f)$ is a global minimizer of f on M. ■

Definition 3.1 *A function $f : M \to I\!R \cup \{+\infty\}$ is said to be coercive if for every sequence $\{x^k\} \subset M$ for which $\lim_{k\to+\infty} d(x^k, x) = +\infty$, for some $x \in M$, it must be the case that*

$$\limsup_{k\to+\infty} f(x^k) = +\infty$$

as well.

Example 3.1 *Let $x \in M$ be a fixed point. We show that $f(\cdot) = d^2(\cdot, x)$ is a coercive function. In fact, let $\{y^k\} \subset M$ such that $\lim_{k\to+\infty} d(y^k, w) = +\infty$ for some $w \in M$. By applying* Theorem 2.1 *to the geodesic triangle $[y^k, x, w]$, it yields*

$$d^2(y^k, w) + d^2(w, x) - 2\langle \exp_w^{-1} y^k, \exp_w^{-1} x\rangle \leq d^2(y^k, x).$$

Since $-d(y^k, w)d(x, w) \leq -\langle \exp_w^{-1} y^k, \exp_w^{-1} x\rangle$, it holds

$$(d(y^k, w) - d(x, w))^2 \leq d^2(y^k, x) := f(y^k).$$

Now, we suppose that $\limsup_{k\to+\infty} f(y^k) < +\infty$. That is, there exists $M > 0$ such that $|d(y^k, w) - d(x, w)| \leq M$. So, $\lim_{k\to+\infty} d(y^k, w) \leq M + d(x, w)$. Since $d(x, w)$ is finite, it follows that $\lim_{k\to+\infty} d(y^k, w) < +\infty$, which is a contradiction. ■

Corollary 3.2 *Let M be a complete Riemannian manifold and let $f : M \to I\!R \cup \{+\infty\}$ be a coercive and lower semicontinuous function, then there exists a global minimizer of f*

Proof. Since $M \neq \emptyset$, there exists $\tilde{x} \in M$ such that $L_f(\alpha) \neq \emptyset$ with $\alpha = f(\tilde{x})$. We will show that $L_f(\alpha)$ is bounded. By contradiction, suppose that $L_f(\alpha)$ is no bounded. Then, there exists a sequence $\{x^k\} \subset L_f(\alpha)$ such that

$$\lim_{k\to+\infty} d(x^k, \bar{x}) = +\infty,$$

for any $\bar{x} \in L_f(\alpha)$. Since $x^k \in L_f(\alpha)$, we have $f(x^k) \leq \alpha$. Taking the upper limit, as k goes to $+\infty$, and using coercive property of f we obtain $+\infty = \limsup_{k \to +\infty} f(x^k) \leq \alpha$, which is a contradiction. Therefore, the set $L_f(\alpha)$ is bounded.
Let us now prove that $L_f(\alpha)$ is a closed set. Let $\{x^k\} \subset L_f(\alpha)$ be a sequence such that $\{x^k\}$ converges to $\bar{x}$ and suppose that $\bar{x} \notin L_f(\alpha)$, that is, $f(\bar{x}) > \alpha$. By the lower semicontinuity, it follows that

$$\alpha < f(\bar{x}) \leq \liminf_{k \to +\infty} f(x^k) \leq \alpha,$$

which is a contradiction. Therefore, $\bar{x} \in L_f(\alpha)$ and so $L_f(\alpha)$ is a closed set.
Finally, since M is a complete Riemannian manifold and using the well known Hopf-Rinow Theorem, see for example Do Carmo[16, Theorem 2.8], we get $L_f(\alpha)$ is a compact set. The result follows by Corolary 3.1. ■

4 Definition of the problem and the algorithm.

Let M be a Hadamard manifold. We are interested in solving the problem:

$$\min_{x \in M} F(x) \tag{4.19}$$

where $F := (F_1, \ldots, F_m)$, for every $F_i : M \to I\!R \cup \{+\infty\}$, is a vector function which satisfies the following assumptions:

(H1) F is a locally Lipschitz function.

(H2) F is a quasiconvex function.

Let us define the set

$$\Omega_k := \{x \in M \mid F(x) \preceq F(x^k)\}. \tag{4.20}$$

We now extend the inexact scalarization proximal point algorithm, based on the work of Apolinario et al.[1] in the context of Euclidean spaces, to solve the problem (4.19). It will be denoted by HISPP algorithm and is described as follows:

HISPP algorithm.

Initialization: Take an arbitrary point $x^0 \in M$. Set $k = 0$.

Main step: Given $x^k \in M$, find $x^{k+1} \in \Omega_k$ such that

$$\epsilon^{k+1} \in \partial^{\circ}(\langle F(\cdot), z^k \rangle)(x^{k+1}) - \lambda_k \exp^{-1}_{x^{k+1}} x^k + \mathcal{N}_{\Omega_k}(x^{k+1}), \tag{4.21}$$

where λ_k is an exogenous positive parameter and $\{z^k\} \subset I\!R^m_+ \setminus \{0\}$ with $\|z^k\| = 1$

Stop criterion: If $x^{k+1} = x^k$ or x^{k+1} is a Pareto-Clarke critical point, stop. Otherwise, $k+1 \leftarrow k$ and go to *Main step*.

Remark 4.1 *1. If $M = I\!R^n$, then equation* (4.21) *becomes*

$$\epsilon^{k+1} \in \partial^\circ(\langle F(\cdot), z^k\rangle)(x^{k+1}) - \lambda_k(x^k - x^{k+1}) + \mathcal{N}_{\Omega_k}(x^{k+1}).$$

From Lemma 2.4*, it holds*

$$\epsilon^{k+1} \in \partial^\circ \left(\langle F(\cdot), z^k\rangle + \frac{\lambda_k}{2}\|\cdot - x^k\|^2\right)(x^{k+1}) + \mathcal{N}_{\Omega_k}(x^{k+1}).$$

Thus, the HISPP algorithm is an extension of the scalarization proximal point method proposted by Apolinario et al.[1].

2. If F is a single-valued function, $z^k = 1$ and $x^{k+1} \in int(\Omega_k)$, then (4.21) *yields*

$$\epsilon^{k+1} \in \partial^\circ F(x^k) - \lambda_k \exp^{-1}_{x^{k+1}} x^k, \tag{4.22}$$

which is a particular case of the Inexact proximal point methods for quasiconvex minimization studied by Baygorrea et al.[3], in sense that abstract subdifferential discussed here is reduced to Clarke's generalized subdifferential.

Moreover, if $\epsilon^{k+1} = 0$, so by Lemma 2.4, equation (4.22) becomes

$$0 \in \partial^\circ(F(\cdot) + \frac{\lambda_k}{2} d^2(\cdot, x^k))(x^{k+1}).$$

Thus, HISPP algorithm also extends the proximal point algorithm proposed by Papa Quiroz and Oliveira[24].

On the other hand, according to the Remark 3.0.1 made by Apolinario et al[1] (see also Huang and Yang[28]) we also can to consider, without loss of generaty, the following condition:

(H3) $\quad 0 \preceq F$, that is, $F_i(x) \geq 0$ for all $i \in \mathcal{I}$ and $x \in dom(F_i)$.

Theorem 4.1 *Under assumptions* (H1) *and* (H2)*, the sequence $\{x^k\}$ generated by the HISPP algorithm is well defined.*

Proof. We proceed by induction. Due to Initialization step of the HISPP algorithm. It holds for $k = 0$. Assume that x^k exists and define

$$\psi_k(x) = \langle F(x), z^k\rangle + \frac{\lambda_k}{2} d^2(x, x^k) + \delta_{\Omega_k}(x).$$

Note that $0 \preceq F$ and $\{z^k\} \subset I\!R^m_+\backslash\{0\}$, so $\psi_k : M \to I\!R \cup \{+\infty\}$ is bounded below. We recall that Ω_k is closed iff $epi(\Omega_k) = \Omega_k \times I\!R_+$ is closed iff δ_{Ω_k} is lower semicontinuous. So, since F is locally Lipschitz, it holds that $\psi_k(\cdot)$ is also lower semicontinuous. Moreover, since $d^2(\cdot, x^k)$ is coercive, see Example 3.1, it follows that $\psi_k(\cdot)$ is also coercive. Therefore, by Corollary 3.2, there exists a global minimum and it can be taken as x^{k+1}. Thus, by Lemma 2.1 we get $0 \in \partial^F(\langle F(\cdot), z^k\rangle)(x^{k+1}) + \frac{\lambda_k}{2}\operatorname{grad} d^2(x^{k+1}, x^k) + \mathcal{N}_{\Omega_k}(x^{k+1})$. The result follows by Theorem 2.2. $\blacksquare$

The following condition, given in various works in the literature of proximal algorithms (e.g., Ceng and Yao[12], Bonnel et al.[10] and Villacorta and Oliveira[39]), is both necessary and sufficient for the iteration to have a limit point, which guarantees strong convergence of the HISPP algorithm: the set

$$\text{(H4)} \qquad (F(x^0) - I\!R^m_+) \cup F(M)$$

is $I\!R^m_+$-complete, which means that for all sequences $\{a^k\} \subset M$ with $a^0 = x^0$, such that $F(a^{k+1}) \preceq F(a^k)$ for all $k \in I\!N$, there exists $a \in M$ such that $F(a) \preceq F(a^k)$ for all $k \in I\!N$.

Note that when stop criterion applies at some iteration, the sequence remains constant thereafter, and so it is convergent to the stopping iterate. Therefore, we will assume from now on that the stop criterion rule never applies.

Denote the following set:

$$E := \{x \in M \mid F(x) \preceq F(x^k), \forall\, k \in I\!N\} \tag{4.23}$$

Note that $E \subseteq \Omega_k$, for any $k \in I\!N$.

Remark 4.2 *By the definition of the HISPP algorithm, $\{x^k\}$ satisfies $F(x^{k+1}) \preceq F(x^k)$. Then, by hypothesis* (H4)*, there exists $a \in M$ such that $F(a) \preceq F(x^k)$, for all $k \in I\!N$. Therefore, the set E, given by (4.23), is nonempty. Moreover, by assumptions* (H1) *and* (H2)*, it is clear to see that E is also closed and convex, respectively.*

The following results will play a central role in the sequel. More precisely, it will be useful in the convergence analysis of the sequence generated by the HISPP algorithm.

Lemma 4.1 *Let $F := (F_1, \ldots, F_m)$, for every $F_i : M \to I\!R \cup \{+\infty\}$. Let $\{x^l\} \subset M$ be a sequence such that $x^{l+1} \in \Omega_l$. Suposse that assumptions* (H1)*,* (H3) *and* (H4) *are satisfied. If $\hat{x} \in M$ is an accumulation point of $\{x^l\}$, then $\hat{x} \in E$.*

Proof. Let $\hat{x} \in M$ be an accumulation point, then there exists $\{x^{l_j}\} \subset \{x^l\}$ such that $x^{l_j} \to \hat{x}$. Let any arbitrary $p \in M$. Since F is locally Lipschitz, there exist $L_p > 0$ and $\delta_p > 0$ such that for any x, $y \in B(p, \delta_p)$, we have $|F_i(x) - F_i(y)| \leq L_p\, d(x, y)$, for all $i \in \mathcal{I}$.

Let $w \in I\!R^m_+ \backslash \{0\}$ be an arbitrary vector. Since F_i is locally Lipschitz and as consequence of the Cauchy-Schwarz inequality, it holds $|\langle F(x), w\rangle - \langle F(y), w\rangle| \leq L_p\|w\|d(x,y)$. Thus, $\langle F(\cdot), w\rangle$ is locally Lipschitz and so, continuous on M. It follows that $\lim_{j\to+\infty}\langle F(x^{l_j}), w\rangle = \langle F(\hat{x}), w\rangle$.

On the other hand, since $x^{l+1} \in \Omega_l$, it follows by relation (4.20) that $F_i(x^{l+1}) \leq F_i(x^l)$ for all $i \in \mathcal{I}$. It is clear that $\langle F(x^{l+1}), w\rangle \leq \langle F(x^l), w\rangle$, for any $w \in I\!R^m_+\backslash\{0\}$. Thus, $\langle F(\cdot), w\rangle$ is a decreasing function.

Moreover, by assumption (H3) and for any $w \in I\!R^m_+\backslash\{0\}$, $0 \leq \langle F(x^l), w\rangle$. So, since $\langle F(\hat{x}), w\rangle$ is an accumulation point of $\{\langle F(\cdot), w\rangle\}$, it follows that $\langle F(\cdot), w\rangle$ is convergent. Then, it yields

$$\langle F(\hat{x}), w\rangle = \lim_{l\to+\infty}\langle F(x^l), w\rangle = \inf_{l\in I\!N}\{\langle F(x^l), w\rangle\} \leq \langle F(x^l), w\rangle, \tag{4.24}$$

for all $l \in I\!N$.

Relation (4.24) implies that

$$0 \leq \langle F(x^l) - F(\hat{x}), w\rangle, \qquad \forall\, l \in I\!N.$$

and for any $w \in I\!R^m_+\backslash\{0\}$. Take $w = \{e_1, \ldots, e_m\}$. It implies that $F(\hat{x}) \preceq F(x^l)$, for all $l \in I\!N$. By relation (4.23), it implies that $\hat{x} \in E$. ■

Proposition 4.1 *Let* $\{x^k\} \subset M$ *be a sequence generated by the* HISPP *algorithm and* $\{\epsilon^k\} \subset T_{x^k}M$ *be an error sequence. If assumptions* (H1), (H2), (H3) *and* (H4) *are satisfied with* $\tilde{\lambda}$ *such that* $\tilde{\lambda} < \lambda_k$, *then for any* $x \in E$ *and all* $k \in I\!N$ *we have*

$$d^2(x^{k+1}, x) \leq d^2(x^k, x) - d^2(x^{k+1}, x^k) - \frac{2}{\tilde{\lambda}}\langle \epsilon^{k+1}, \exp^{-1}_{x^{k+1}} x\rangle.$$

Proof. By Lemma 2.3, relation (4.21) becomes

$$\epsilon^{k+1} = \sum_{i\in\mathcal{I}} z_i^k g_i^k - \lambda_k \exp^{-1}_{x^{k+1}} x^k + v^k,$$

where $v^k \in \mathcal{N}_{\Omega_k}(x^{k+1})$ and $g_i^k \in \partial^\circ F_i(x^{k+1})$. Thus,

$$\exp^{-1}_{x^{k+1}} x^k = \frac{1}{\lambda_k}(\sum_{i\in\mathcal{I}} z_i^k g_i^k + v^k - \epsilon^{k+1}). \tag{4.25}$$

Let $x \in E \subseteq \Omega_{k+1}$ be an arbitrary point for any $k \in I\!N$. Since E is a closed set, see Remark 4.2, there exists $\{x^l\} \subset int(\Omega_{k+1})$ such that $\{x^l\}$ converges to x.

Since $x^l \in int(\Omega_{k+1})$, it follows that $F_i(x^{k+1}) > F_i(x^l)$ for all $i \in \mathcal{I}$. Hence, for $g_i^k \in \partial^\circ F_i(x^{k+1})$ and according to Lemma 2.5, we obtain $\langle g_i^k, \exp^{-1}_{x^{k+1}} x^l\rangle \leq 0$, for all $i \in \mathcal{I}$. It implies that

$$\langle \sum_{i\in\mathcal{I}} z_i^k g_i^k, \exp^{-1}_{x^{k+1}} x^l\rangle \leq 0 \tag{4.26}$$

with $z_i^k \in I\!R_+$.

Now, since $v^k \in \mathcal{N}_{\Omega_k}(x^{k+1}) \subseteq \mathcal{N}_{\Omega_{k+1}}(x^{k+1})$ and $x^l \in \Omega_{k+1}$, it holds that $\langle v^k, \exp^{-1}_{x^{k+1}} x^l\rangle \leq 0$. So, relation (4.26) yields

$$\begin{aligned} 0 &\geq \langle \sum_{i\in\mathcal{I}} z_i^k g_i^k, \exp^{-1}_{x^{k+1}} x^l\rangle + \langle v^k, \exp^{-1}_{x^{k+1}} x^l\rangle \\ &= \langle \lambda_k \exp^{-1}_{x^{k+1}} x^k + \epsilon^{k+1}, \exp^{-1}_{x^{k+1}} x^l\rangle. \end{aligned} \tag{4.27}$$

This last result is obtained from relation (4.25). Thus, since $\tilde{\lambda} < \lambda_k$, (4.27) yields

$$\langle \exp^{-1}_{x^{k+1}} x^k, \exp^{-1}_{x^{k+1}} x^l\rangle + \frac{1}{\tilde{\lambda}}\langle \epsilon^{k+1}, \exp^{-1}_{x^{k+1}} x^l\rangle < 0.$$

It follows from Theorem 2.1, for the given geodesic triangle $[x^k, x^l, x^{k+1}]$, that

$$d^2(x^k, x^l) \geq d^2(x^k, x^{k+1}) + d^2(x^{k+1}, x^l) + \frac{2}{\tilde{\lambda}}\langle \epsilon^{k+1}, \exp^{-1}_{x^{k+1}} x^l\rangle, \tag{4.28}$$

and taking limit, as $l \to +\infty$, the result follows. ■

5 Inexact variants of the HISPP algorithm for multiobjective minimization

In this section, following the same approach done by Baygorrea et al.[3] for single-valued functions, we discuss convergence results of two inexact proximal algorithms for multiobjective minimization. Firts, about a classical version, as error estimative criterion is considered under mild and natural conditions. Second, it is considered the Tang and Huang's error criterion, see [37].

5.1 HISPP1 algorithm and convergence results.

In this subsection, we discuss the HISPP algorithm considering the classical version on error estimation criterion, which will called the HISPP1 algorithm, and it is as follows:

Let $\{x^k\} \in M$ be a sequence generated by the HISPP algorithm. Consider the following assumption:

(A1) $\sum_{k=0}^{+\infty} \|\epsilon^k\| < +\infty$.

The convergence analysis of the HISPP1 algorithm is based on the theory of p-Quasi-Fejér convergence, which was introduced by Ermol′ev [17] in the context of sequences of random variables. We recall that a sequence $\{x^k\} \subset M$ is said to be p-Quasi-Fejér convergent to a set $U \subset M$, $U \neq \emptyset$ if, for each $u \in U$ there exists a non-negative, summable sequence $\{\rho_k\}$ such that

$$d^p(x^{k+1}, u) \leq d^p(x^k, u) + \rho_k \qquad k = 0, 1 \ldots$$

The following result on p-Quasi-Fejér convergence is well known.

Lemma 5.1 *Let (M,d) be a complete metric space, $p \in (0,+\infty)$ and $\{x^k\} \subset M$ be a sequence which is p-Quasi-Fejér convergent to a nonempty set $U \subset M$, then $\{x^k\}$ is bounded. Furthermore, if a cluster point $x \in M$ of $\{x^k\}$ belongs to U, then it converges to such a point.*

Proof. See Svaiter[36], Proposition 2.1 ■

Lemma 5.2 *Let $\{x^k\} \subset M$ and $\{\epsilon^k\} \subset T_{x^k}M$ be sequences generated by HISPP1 algorithm. Consider assumptions* (H1), (H2), (H3) *and* (H4). *Then, it holds that*

(i) *The sequence $\{x^k\}$ is bounded;*

(ii) $\lim_{k\to+\infty} d(x^k, x^{k+1}) = 0$;

(iii) *If $\lim_{j\to+\infty} x^{k_j} = \hat{x}$ then $\lim_{j\to+\infty} x^{k_{j+1}} = \hat{x}$*

Proof. (i) Let $z \in E$ be a fixed point. From Proposition 4.1, we have

$$\begin{aligned} d^2(x^{k+1}, z) &\leq d^2(x^k, z) - \frac{2}{\tilde{\lambda}}\langle \epsilon^{k+1}, \exp^{-1}_{x^{k+1}} z\rangle \\ &\leq d^2(x^k, z) + \frac{2}{\tilde{\lambda}}\|\epsilon^{k+1}\| d(x^{k+1}, z). \end{aligned}$$

It implies that

$$\left(d(x^{k+1}, z) - \frac{1}{\tilde{\lambda}}\|\epsilon^{k+1}\|\right)^2 \leq d^2(x^k, z) + \frac{1}{\tilde{\lambda}^2}\|\epsilon^{k+1}\|^2 \tag{5.29}$$

It is clear that if $d(x^{k+1}, z) - \frac{1}{\tilde{\lambda}}\|\epsilon^{k+1}\| \leq 0$, we get $d(x^{k+1}, z) \leq \frac{1}{\tilde{\lambda}}\|\epsilon^{k+1}\| < \delta$, for some $\delta > 0$. It follows that $\{x^k\}$ is bounded. In this case, note that for k suffiently large, we get that $\{x^k\}$ converges to z. In this case, Since z was chosen arbitrarily, it holds that $E = \{z\}$. Otherwise, by (5.29) we have

$$\begin{aligned} d(x^{k+1}, z) - \frac{1}{\tilde{\lambda}}\|\epsilon^{k+1}\| &\leq \sqrt{d^2(x^k, z) + \frac{1}{\tilde{\lambda}^2}\|\epsilon^{k+1}\|^2} \\ &\leq d(x^k, z) + \frac{1}{\tilde{\lambda}}\|\epsilon^{k+1}\| \end{aligned}$$

Here, $d(x^{k+1}, z) \leq d(x^k, z) + \frac{2}{\tilde{\lambda}}\|\epsilon^{k+1}\|$. Thus, because of the summability of $\|\epsilon^{k+1}\|$ we have that $\{x^k\}$ is a 1-Quasi-Fejér convergente to E. By Lemma 5.1, it follows that $\{x^k\}$ is bounded.

(ii) By using again Proposition 4.1, it holds

$$\begin{aligned} d^2(x^{k+1}, x^k) &\leq d^2(x^k, z) - d^2(x^{k+1}, z) - \frac{2}{\tilde{\lambda}}\langle \epsilon^{k+1}, \exp^{-1}_{x^{k+1}} z\rangle \\ &\leq d^2(x^k, z) - d^2(x^{k+1}, z) + \frac{2}{\tilde{\lambda}}\|\epsilon^{k+1}\| d(x^{k+1}, z). \end{aligned}$$

By summing up, it yields

$$\sum_{k=0}^{n} d^2(x^{k+1}, x^k) \le d^2(x^0, z) - d^2(x^{k+1}, z) + \frac{2}{\tilde{\lambda}} \max\{d(x^{k+1}, z)\} \sum_{k=0}^{n} \|\epsilon^{k+1}\|.$$

By assumption (A1), it holds $\sum_{k=0}^{n} d^2(x^{k+1}, x^k) < +\infty$, and the result follows.

(iii) Since $d(x^{k_j+1}, \hat{x}) \le d(x^{k_j+1}, x^{k_j}) + d(x^{k_j}, \hat{x})$ and $\hat{x}$ is a limit point of $\{x^{k_j}\}$, result follows clearly by item (ii). ■

Proposition 5.1 *Let $\{x^k\} \subset M$ and $\{\epsilon^k\} \subset T_{x^k}M$ be sequences generated by HISPP1 algorithm. If the assumptions* (H1), (H2), (H3) *and* (H4) *are satisfied with $\tilde{\lambda} > 0$ such that $0 < \lambda_k < \tilde{\lambda}$, then the sequence $\{x^k\}$ converges to a point of E.*

Proof. By Lemma 5.2(i), it follows that there exist $\{x^{k_j}\} \subset \{x^k\}$ and $\hat{x} \in M$ such that x^{k_j} converges to $\hat{x}$. Moreover, by Lemma 4.1, it follows that $\hat{x} \in E$. Therefore, according to Lemma 5.1, we can conclude that the given sequence is convergent to $\hat{x} \in E$. ■

Theorem 5.1 *Let $F := (F_1, \ldots, F_m)$, for every $F_i : M \to I\!R \cup \{+\infty\}$. Let $\{x^k\} \subset M$ and $\{\epsilon^k\} \subset T_{x^k}$ be sequences generated by HISPP1 algorithm. If the assumptions* (H1), (H2), (H3) *and* (H4) *are satisfied with $\tilde{\lambda} > 0$ such that $0 < \lambda_k < \tilde{\lambda}$, then $\{x^k\}$ converges to a Pareto-Clarke critical point in M.*

Proof. By Proposition 5.1, there exists $\hat{x} \in E$ such that $\{x^k\}$ converges to $\hat{x}$. Now, we prove that $\hat{x}$ is Pareto-Clarke critical point. By contradiction, suppose that there exists $\tilde{d} \in T_{\hat{x}}M$ satisfying

$$F_i^\circ(\hat{x}, \tilde{d}) < 0, \tag{5.30}$$

for all $i \in \mathcal{I}$. That is, there exists $\delta > 0$ such that $F(\exp_{\hat{x}}(t\tilde{d})) \prec F(\hat{x})$, for all $t \in (0, \delta]$. Denote $\gamma(t) := \exp_{\hat{x}}(t\tilde{d})$. So $\gamma(t) \in E$

On the other hand, we recall that $\{x^{k+1}\} \in \Omega_k$ is a sequence generated by relation (4.21),

$$\epsilon^{k+1} + \lambda_k \exp^{-1}_{x^{k+1}} x^k - v^k \in \partial^\circ(\langle F(\cdot), z^k \rangle)(x^{k+1}),$$

where $v^k \in \mathcal{N}_{\Omega_k}(x^{k+1})$. By definition of Clarke's generalized subdifferential, for all $d \in T_{x^{k+1}}M$, we get

$$\langle \epsilon^{k+1}, d \rangle + \lambda_k \langle \exp^{-1}_{x^{k+1}} x^k, d \rangle - \langle v^k, d \rangle \le \langle F(\cdot), z^k \rangle^\circ(x^{k+1}, d). \tag{5.31}$$

Take $d := \exp^{-1}_{x^{k+1}} \gamma(t)$. Since $\gamma(t) \in \Omega_k$, it follows that $\langle v^k, \exp^{-1}_{x^{k+1}} \gamma(t) \rangle \le 0$, for all $t \in (0, \delta]$.

Then, relation (5.31) yields

$$\langle \epsilon^{k+1}, \exp^{-1}_{x^{k+1}} \gamma(t) \rangle + \lambda_k \langle \exp^{-1}_{x^{k+1}} x^k, \exp^{-1}_{x^{k+1}} \gamma(t) \rangle \le \langle F(\cdot), z^k \rangle^\circ(x^{k+1}, \exp^{-1}_{x^{k+1}} \gamma(t)) \tag{5.32}$$

By triangle inequality property, and taking into account that $\|\exp_w^{-1}s\| = d(s,w)$, for all $w, s \in M$, relation (5.32) yields

$$-\|\epsilon^{k+1}\|d(\gamma(t), x^{k+1}) - \lambda_k d(x^k, x^{k+1})d(\gamma(t), x^{k+1}) \le \langle F(\cdot), z^k\rangle^{\circ}(x^{k+1}, \exp^{-1}_{x^{k+1}}\gamma(t)). \quad (5.33)$$

Since $\{z^k\}$ is an unit vector sequence, there exists $\{z^{k_j}\} \subseteq \{z^k\}$ such that $z^{k_j} \to \bar{z}$ with $\bar{z} \in I\!R^m \backslash \{0\}$. Considere a subsequence, if necessary, we may assume, without loss of generality, that $\{x^{k_j}\} \subseteq \{x^k\}$. By Lemma 2.2, and since $\lambda_k < \tilde{\lambda}$, (5.33) yields

$$-\|\epsilon^{k_j+1}\|d(\gamma(t), x^{k_j+1}) - \tilde{\lambda} d(x^{k_j}, x^{k_j+1})d(\gamma(t), x^{k_j+1}) \le \sum_{i\in\mathcal{I}} z_i^{k_j} F_i^{\circ}(x^{k_j+1}, \exp^{-1}_{x^{k_j+1}}\gamma(t)). \quad (5.34)$$

Moreover, we have that $\lim_{j\to+\infty} \exp^{-1}_{x^{k_j+1}}\gamma(t) = \exp^{-1}_{\hat{x}}(\exp_{\hat{x}} t\tilde{d}) = t\tilde{d}$, with $t \in (0, \delta]$. Note that by Lemma 5.2 (ii)-(iii), the left-hand side of (5.34) tends to 0, as $j \to +\infty$.

So, applying lim sup to relation (5.34) and since $\lim_{j\to+\infty} z_i^{k_j} = \bar{z}_i$, (5.34) yields

$$\begin{aligned} 0 &\le \limsup_{j\to+\infty}\Big(\sum_{i\in\mathcal{I}} z_i^{k_j} F_i^{\circ}(x^{k_j+1}, \exp^{-1}_{x^{k_j+1}}\gamma(t))\Big) \\ &\le \sum_{i\in\mathcal{I}} \limsup_{j\to+\infty} z_i^{k_j} F_i^{\circ}(x^{k_j+1}, \exp^{-1}_{x^{k_j+1}}\gamma(t)) \\ &\le \sum_{i\in\mathcal{I}} \bar{z}_i t F_i^{\circ}(\hat{x}, \tilde{d}), \end{aligned}$$

with $\bar{z}_i \in I\!R_+$ and $t \in (0, \delta]$. It follows that there exists, at least, some $i_0 \in \mathcal{I}$ such that $0 \le F_{i_0}(\hat{x}, \tilde{d})$, which contradicts relation (5.30). Therefore, $\hat{x}$ is a Pareto Clarke critical point. ■

5.2 HISPP2 algorithm and convergence results.

In this subsection, we discuss the HISPP algorithm and the error criteria studied by Baygorrea et al.[3] for solving quasiconvex minimization for single-valued function on Hadamard manifold, which will be called the HISPP2 algorithm and it is given as follows:

Let $\{x^k\} \subset M$ be a sequence generated by the HISPP algorithm. Consider the following assumption:

(B1) $\|\epsilon^k\| \le \eta_k d(x^k, x^{k-1})$;

(B2) $\sum_{k=1}^{+\infty} \eta_k^2 < +\infty$.

Lemma 5.3 *Let $\{x^k\} \subset M$ and $\{\epsilon^k\} \subset T_{x^k}M$ be sequences generated by* HISPP2 *algorithm. Then there exists an integer $k_0 \ge 0$ such that for all $k \ge k_0$ we have*

$$d^2(x^k, x) \le \left(1 + \frac{2\eta_k^2}{1 - 2\eta_k^2}\right) d^2(x^{k-1}, x) - \frac{1}{2} d^2(x^k, x^{k-1}), \forall\, x \in E.$$

Furthermore, $\{x^k\}$ is a bounded sequence and $\lim_{k\to+\infty} d(x^k, x^{k-1}) = 0$.

Proof. See Baygorrea et al[3]. ■

Proposition 5.2 *Let $F := (F_1, \dots, F_m)$, for every $F_i : M \to I\!R \cup \{+\infty\}$. Let $\{x^k\} \subset M$ and $\{\epsilon^k\} \subset T_{x^k}M$ be sequences generated by* HISPP2 *algorithm. If the assumptions* (H1), (H2), (H3) *and* (H4) *are satisfied with $\tilde{\lambda} > 0$ such that $0 < \lambda_k < \tilde{\lambda}$, then the sequence $\{x^k\}$ converges to a point of E.*

Proof. Lemma 5.3 implies that there exists $\{x^{k_j}\} \subset \{x^k\}$ and $\hat{x} \in M$ such that x^{k_j} converges to $\hat{x}$. In analogy to the last part in the proof of Theorem 3.1 by Tang and Huang[37], convergence of the whole sequence $\{x^k\}$ to a point $\hat{x} \in M$ is obtained. In adittion, From Lemma 4.1, $\hat{x} \in E$ ■

Theorem 5.2 *Let $F := (F_1, \dots, F_m)$, for every $F_i : M \to I\!R \cup \{+\infty\}$. Let $\{x^k\} \subset M$ and $\{\epsilon^k\} \subset T_{x^k}M$ be sequences generated by the* HISPP2 *algorithm. If the assumptions* (H1), (H2), (H3) *and* (H4) *are satisfied with $\tilde{\lambda} > 0$ such that $0 < \lambda_k < \tilde{\lambda}$, then $\{x^k\}$ converges to a Pareto-Clarke critical point.*

Proof. Consider the first paragraph of the proof of Theorem 5.1. Let $\{x^k\}$ be a sequence generated by the HISPP2 algorithm and $\hat{x} \in M$ such that x^k converges to $\hat{x}$. Suppose that $\hat{x}$ is no Pareto-Clarke critical point. Then, by assumptions (B1) and (B2), (5.34) yields

$$\begin{aligned} \sum_{i \in \mathcal{I}} z_i^{k_j} F_i^\circ(x^{k_j+1}, \exp^{-1}_{x^{k_j+1}} \gamma(t)) &\geq -\eta_{k_j} d(x^{k_j+1}, x^{k_j}) d(\gamma(t), x^{k_j+1}) - \tilde{\lambda} d(x^{k_j}, x^{k_j+1}) d(\gamma(t), x^{k_j+1}) \\ &\geq -(M + \tilde{\lambda}) d(x^{k_j+1}, x^{k_j}) d(\gamma(t), x^{k_j+1}), \end{aligned} \tag{5.35}$$

for some $M > 0$. Note that $d(\gamma(t), x^{k_j+1}) = \|\exp^{-1}_{x^{k_j+1}} \gamma(t)\|$. So, $\lim_{j \to +\infty} d(\gamma(t), x^{k_j+1}) = t\|\tilde{d}\|$ for all $t \in (0, \delta]$. Thus, since $\{x^k\}$ and $\{z^k\}$ are convergent, the right-hand side of (5.35) tends to 0, as $j \to +\infty$. Thus, relation (5.35) yields

$$\liminf_{j \to +\infty} \sum_{i \in \mathcal{I}} \bar{z}_i F_i^\circ(\hat{x}, t\tilde{d}) \geq 0,$$

where $\bar{z}_i > 0$ for all $i \in \mathcal{I}$, $t \in (0, \delta]$ and some $\tilde{d} \in T_{\hat{x}}M$. It implies that $F_i^\circ(\hat{x}, \tilde{d}) \geq 0$ for some $i \in \mathcal{I}$, which leads to a contradiction to relation (5.30). ■

Note that, in the case that $\bar{\lambda} d(x^{k+1}, x^k) \geq \|\epsilon^{k+1}\|$, if $\eta_k \leq \bar{\lambda}$ then ISPPA2 is a better approximation than SPPA1; otherwise, ISPPA2 will be an ISPPA1-like relaxation algorithm.a sesin, el embajador ruso ante la ONU, Vitali Churkin, seal que la votacin del Consejo de Seguridad deba ser aplazada hasta despus de celebrarse una reunin entre los expertos rusos y estadounidenses con el fin de elaborar un plan para arreglar la situacin en Alepo.

El diplomtico recalc que los autores del documento haban sido sometidos a una "presin flagrante" por parte de EE.UU., el Reino Unido y Francia, que exigan "llevar a votacin un proyecto condenado al fracaso". Churkin calific tales acciones de provocativas y agreg que socavan los intentos internacionales de poner fin al conflicto sirio.

Segn el analista internacional Andrs Pierantoni, los pases occidentales siempre recorren a la diplomacia cuando hay una derrota de los mercenarios en el terreno militar en Siria. El experto ha destacado que instan a un alto el fuego inmediato para mejorar el abastecimiento y reequipar a los combatientes.

6 Final Remarks.

- We extend and generalize the scalarization proximal point algorithm for (non-necessarily smooth) unconstrained multiobjective optimization for quasiconvex and locally Lipschitz objective functions from Euclidean spaces, studied by Apolinario et al.[1], to Hadamard manifolds.
- For this method, using the same error criteria studied by Baygorrea et al.[3], two inexact versions are presented: First, it was considered a classical and natural error estimation criterion. Second, the error estimative assumptions is related with the distance of adjacent iterations as an upper bound for this.
- Furthermore, under reasonable and mild assumptions, we proved that any accumulation point is itself a Pareto-Clarke critical ones for any of these two versions. Moreover, for any initial point, under the completion property, the sequence generated by this algorithm strongly converges to a Pareto-Clarke critical point.
- Nevertheless, to compute the iterates of the algorithms HISPP, we need to determine vectors of normal cones, which can be hard and non-effortleessly. Thus, overcoming this obstacle requires some technicalities to be studied.

References

[1] Apolinario, H.C.F., Papa Quiroz, E.A. and Oliveira, P.R., A scalarization proximal point method for quasiconvex multiobjective minimization, J. of Global Optim. 64(1), 79-96 (2016)

[2] Azagra D., Ferrera J. and López-Mesas F., Nonsmooth analysis and Hamilton-Jacobi equations on Riemannian manifolds, Journal of Functional Analysis 220, 304-361 (2005)

[3] Baygorrea N., Papa Quiroz E.A. and Maculan N., Inexact proximal point methods for quasiconvex minimization on Hadamard manifolds, Journal of the Operations Research Society of China, 4(4), 397-424 (2016)

[4] Baygorrea N., Papa Quiroz E.A. and Maculan N., On the convergence rate of an inexact proximal point algorithm for quasiconvex minimization on Hadamard manifolds, Journal of the Operations Research Society of China, 5(1), 1-11 (2016)

[5] Bello Cruz J.Y., Lucambio Pérez L.R. and Melo J.G., Convergence of the projected gradient method for quasiconvex multiobjective optimization, Nonlinear Analysis:Theory, Methods and Applications 74(16), 5268-5273, (2011)

[6] Bento G.C., Ferreira O.P. and Oliveira P.R., Local convergence of the proximal point method for a special class of nonconvex functions on Hadamard manifolds, Nonlinear Analysis, Theory Methods and App., 73(2), 564-572 (2010)

[7] Bento G.C., Ferreira O.P. and Oliveira P.R., Unconstrained steepest descent method for multicriteria optimization on Riemannian manifolds, Journal of Optimization Theory and Applications, 154(1), 88-107 (2012)

[8] Bento G.C., Cruz Neto J.X., A subgradient method for multiobjective optimization on Riemannian manifolds, J. of Optimization Theory and Applications, 159(1), 125-137 (2013)

[9] Bento G.C., Cruz Neto J.X. and Santos P.S.M, An inexact steepest descent method for multicriteria optimization on Riemannian manifolds, J. Optim. Theory Appl. 159, 108-124 (2013).

[10] Bonnel, H., Iusem, A.N. and Svaiter, B.F., Proximal methods in vector optimization, SIAM J. Optim. 15(4), 953-970 (2005)

[11] Correa R., Jofre A. and Thibault L., Characterization of lower semicontinuous convex functions, Proc. Amer. Math. Soc., 116, 67-72 (1992)

[12] Ceng, L.C. and Yao, J.C., Approximate proximal methods in vector optimization. European Journal of Operational Research. 183,1-19 (2007)

[13] Custdio A.L., Aguilar Madeira J.F., Vaz A.I.F. and Vicente L.N., Direct multisearch for multiobjective optimization, SIAM J. Optim., 21(3), 1109-1140 (2011)

[14] da Cruz Neto, J.X., da Silva G.J.P, Ferreira O.P. and Lopes J.O., A subgradient method for multiobjective optimization, Computational Optimization and Applications, 54(3), 461-472 (2013)

[15] da Cruz Neto, J.X., Bento G.C., A subgradient method for multiobjective optimization on Riemannian manifolds, Journal of Optimization Theory and Applications, 159(1), 125-137 (2013)

[16] do Carmo, M.P., Riemannian geometry, Bikhausen, Boston (1992)

[17] Ermol′ev J.M., The method of generalized stochastic gradients and quasi-Fejér sequences, Cybernetics and Systems Analysis, 5(2), 208-220 (1969)

[18] Fliege, J., An efficient interior-point method for convex multicriteria optimization problems, Math. Oper. Res., 31, 825845 (2006)

[19] Fliege, J. and Svaiter, B.F., Steepest descent methods for multicriteria optimization, Math. Methods Oper. Res. 51, 479494 (2000)

[20] Graa Drummond, L.M. and Iusem, A. N., A projected gradient method for vector optimization problems, Comput. Optim. Appl. 28, 5-30 (2008)

[21] Graziano, M.G., Fuzzy cooperative behavior in response to market imperfections, Int. J. Intell. Syst. 27, 108-131 (2012)

[22] Graziano, M.G., Romaniello, M., Linear cost share equilibria and the veto power of the grand coalition, Soc. Choice Welf. 38, 269-303 (2012)

[23] Gregrio, R., Oliveira, P.R., A logarithmic-quadratic proximal point scalarization method for multiobjective programming, J. Glob. Optim. 49(2), 281-291 (2011)

[24] Papa Quiroz, E.A. and Oliveira, P.R., Proximal point methods for minimizing quasiconvex locally lipschitz functions on Hadamard manifolds, Nonlinnear Analysis. 75(15), 5924-5932, (2012)

[25] Papa Quiroz, E.A. and Oliveira, P.R., Full convergence of the proximal point method for quasiconvex function on Hadamard manifolds, ESAIM: Control, Optimisation and Calculus of Variations, 18(2), 483-500 (2012)

[26] Hosseini, S. and Pouryayevali, M.R., Generalized gradient and characterization of epi-Lipschitz sets in Riemannian manifolds, Nonlinear Analysis: Theory, Methods and Applications, 74(12), 38843895 (2011)

[27] Hosseini, S. and Pouryayevali, M.R., Nonsmooth optimization techniques on Riemannian manifolds. J. Optim. Theory Appl. 158(2), 328342 (2013)

[28] Huang, X.X. and Yang, X.Q., Duality for multiobjective optimization via nonlinear Lagrangian functions, Journal of Optimization, theory and App., 120(1), 111-127 (2004)

[29] Ledyaev Y. and Zhu Q., Nonsmooth analysis on smooth manifolds, Transactions of the American Mathematical Society, 359(8), 3687-3732 (2007)

[30] Li, H.L. and Yu, C.S., Solving multiple objective quasiconvex goal programming problems by linear programming, International Transactions in operational Research, 7(3), 265-284 (2000)

[31] Rapcsák, T., Smooth nonlinear optimization in $\mathbb{R}^n$, Kluwer Academic Publishers (1997)

[32] Rocha, R.A., Oliveira, P.R., Gregrio, R.M., Souza, M., Logarithmic quasi-distance proximal point scalarization method for multi-objective programming, Appl. Math. Comput. 273, 856-867 (2016)

[33] Rockafellar, R.T., Monotone operators and the proximal point algorithm, SIAM J. Contr. Optim. 14(5), 877-898 (1976)

[34] Sakai, T., Riemannian geometry, American Mathematical Society, Providence, RI. (1996)

[35] Shiga, K., Hadamard manifolds, advanced studies in pure mathematics 3, Geometry of Geodesic and Related Topic, 239-281 (1984)

[36] Svaiter, B.F., A class of Fejér convergent algorithms, approximate resolvents and the Hybrid Proximal-Extragradient method, J. Optim. Theory App., 162(1), 133-153 (2014)

[37] Tang G.J. and Huang N.J., An inexact proximal point algorithm for maximal monotone vector fields on Hadamard manifolds. Operations Research Letters, (41)6, 586-591 (2013)

[38] Udriste, C., Convex function and optimization methods on Riemannian manifolds, Kluwer Academic Publishers, 1994.

[39] Villacorta, K.D.V and Oliveira, P.R., An interior proximal method in vector optimization. European Journal of Operational Research, 214, 485-492 (2011)