

PROXIMAL POINT ALGORITHMS ON HADAMARD MANIFOLDS: LINEAR CONVERGENCE AND FINITE TERMINATION*

JINHUA WANG[†], CHONG LI[‡], GENARO LOPEZ[§], AND JEN-CHIH YAO[¶]

Abstract. In the present paper, we consider inexact proximal point algorithms for finding singular points of multivalued vector fields on Hadamard manifolds. The rate of convergence is shown to be linear under the mild assumption of metric subregularity. Furthermore, if the sequence of parameters associated with the iterative scheme converges to 0, then the convergence rate is superlinear. At the same time, the finite termination of the inexact proximal point algorithm is also provided under a weak sharp minima-like condition. Applications to optimization problems are provided. Some of our results are new even in Euclidean spaces, while others improve and/or extend some known results in Euclidean spaces. As a matter of fact, in the case of exact proximal point algorithm, our results improve the corresponding results in [G. C. Bento and J. X. Cruz Neto, *Optim.*, 63 (2014), pp. 1281–1288]. Finally, several examples are provided to illustrate that our results are applicable while the corresponding results in the Hilbert space setting are not.

Key words. inexact proximal point algorithms, Hadamard manifolds, monotone vector fields, finite termination, convergence rate

AMS subject classifications. Primary, 49J40; Secondary, 58D17

DOI. 10.1137/15M1051257

1. Introduction. Recent interests are focused on the study of various problems on manifolds. The reason is that many optimization problems arising in various applications are posed on manifolds and require a manifold structure (not necessarily with linear structure), such as geometric models for the human spine [4], some eigenvalue optimization problems [6, 61], and so on. Moreover, some constrained optimization problems may be seen as unconstrained ones on some manifolds. For example, some eigenvalue problems, invariant subspace computations, optimization problems with equality constraints, etc., can be reformulated as unconstrained optimization problems on some manifolds (cf. [4, 50, 56] and references therein). Then, these problems can be resolved more efficiently by using the underlying geometric structure of the manifold, which sometimes has a lower complexity and quite often has better numerical properties (as explained in [6]). On the other hand, some nonconvex/nonmonotone problems in linear spaces may become convex/monotone problems through the introduction of

*Received by the editors December 3, 2015; accepted for publication (in revised form) September 20, 2016; published electronically December 1, 2016.

<http://www.siam.org/journals/siopt/26-4/M105125.html>

Funding: The first author was partially supported by the National Natural Science Foundation of China (grant 11371325) and by Zhejiang Provincial Natural Science Foundation of China (grant LY17A010021). The second author was partially supported by the National Natural Science Foundation of China (grant 11571308). The third author was partially supported by DGES (grant MTM2015-65242-C2-1) and Junta de Andalucía (grant P08-FQM-03453). Research of the fourth author was partially supported by the National Science Council of Taiwan (grant MOST 105-2115-M-039-002-MY3).

[†]Department of Mathematics, Zhejiang University of Technology, Hangzhou 310032, People's Republic of China (wjh@zjut.edu.cn).

[‡]Corresponding author. School of Mathematical Sciences, Zhejiang University, Hangzhou 310027, People's Republic of China (cli@zju.edu.cn).

[§]Departamento de Análisis Matemático, Universidad de Sevilla, Apdo. 1160, 41080-Sevilla, Spain (glopez@us.es).

[¶]Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Sichuan, China (yaojc@math.nsysu.edu.tw).

an appropriate Riemannian metric (in [15] many such examples are provided). Thus, these problems can be studied by using some convex analysis results on manifolds and the intrinsic geometry property of the manifold endowed with some appropriate Riemannian metric.

Hence, extension of concepts, techniques, and algorithms from Euclidean spaces to Riemannian manifolds are natural and meaningful. Indeed, some important ones such as monotone operators, weak sharp minima, variational inequality, subdifferentials, Newton’s method, the conjugate gradient method, the trust-region method, the proximal point method, and their modifications for optimization problems on linear spaces are extended to the Riemannian manifolds setting in [4, 7, 10, 17, 22, 29, 32, 33, 36, 37, 38, 39, 44, 45, 53, 54, 55, 58] and the references therein. Furthermore, some recent works about optimization algorithms on Riemannian manifolds for nonsmooth functions can be found in [8, 25, 26] and so on.

Our interest in the present paper is focused on inexact proximal point algorithms for finding singular points of multivalued vector fields on Hadamard manifolds. We first simply recall the definition of this algorithm and its history in the Hilbert space setting. Let H be a real Hilbert space, and $T : H \rightarrow H$ be a multivalued maximal monotone operator. The problem of finding

$$(1.1) \quad z \in H \text{ such that } 0 \in T(z)$$

has interesting interpretations in various fields, and a lot of problems such as convex minimization problems, variational inequalities, saddle point problems, and complementarity problems, and so on can be recast into the form (1.1); see, e.g., [41, 51], etc. A large variety of methods for solving (1.1) have been proposed and investigated (see, e.g., [40, 51]). Among them, one of the most important is the well-known proximal point algorithm defined by

$$(1.2) \quad z^{k+1} = (\mathbf{I} + c_k T)^{-1}(z^k), \quad k = 0, 1, \dots,$$

where $z^0 \in H$ is an initial point and $\{c_k\} \subseteq (0, +\infty)$ is a sequence of the regularized parameters. The proximal point algorithm was earlier used for regularizing linear equations [31], and seems to have been applied for the first time to convex minimization by Martinet [41], and has been explored extensively; see, e.g., [5, 24, 40, 51] and the book [9], together with the reference therein.

In particular, in his seminal work [51], Rockafellar developed two inexact versions of the proximal point algorithm to generate the sequences $\{z^{k+1}\}$ by using two approximate rules $z^{k+1} \approx (\mathbf{I} + c_k T)^{-1}(z^k)$. One criterion adopted by Rockafellar is

$$(1.3) \quad \|z^{k+1} - (\mathbf{I} + c_k T)^{-1}(z^k)\| \leq \varepsilon_k, \quad k = 0, 1, \dots,$$

while the other is

$$(1.4) \quad d(0, S_k(z^{k+1})) \leq \frac{\varepsilon_k}{c_k}, \quad k = 0, 1, \dots,$$

where each $S_k(\cdot) := T(\cdot) + c_k^{-1}(\cdot - z^k)$ and the real sequence $\{\varepsilon_k\}$ satisfies $\sum_{k=0}^{\infty} \varepsilon_k < \infty$. As shown in [51], that criterion (1.4) implies criterion (1.3). Under the assumption that the solution set of (1.1) is nonempty and the regularized parameters $\{c_k\}$ are bounded away from zero, it is proved in [51] that the sequence $\{z^k\}$ generated by (1.3) (and so by (1.4)) with any initial point z^0 converges weakly to a solution of (1.1); furthermore the convergence rate is at least linear if T^{-1} is additionally assumed to be Lipschitz continuous at 0.

The proximal point algorithm was first extended in [21] to solve convex optimization problems on Hadamard manifolds and, further, in [33] to find singular points of multivalued vector fields on a Hadamard manifold. Recently, the convergence property of the proximal point algorithm for solving variational inequality problems for multivalued mappings on general Riemannian manifolds has been studied in [35]. The setting in the Hadamard manifold is as follows. Let M be a Hadamard manifold, and let $A : M \rightarrow 2^{TM}$ be a multivalued vector field such that $A(x) \subseteq T_x M$ for each $x \in M$ and the domain $\mathcal{D}(A)$ of A is closed and convex. Given a sequence of parameters $\{\lambda_n\} \subseteq (0, +\infty)$, then the proximal point algorithm on Hadamard manifolds for finding a singularity of A (i.e., $0 \in A(x)$) is defined as follows.

Algorithm P. Give initial point $x_0 \in \mathcal{D}(A)$. Letting $n = 0, 1, 2, \dots$ and having x_n , determine x_{n+1} such that

$$(1.5) \quad 0 \in A(x_{n+1}) - \lambda_n \exp_{x_{n+1}}^{-1} x_n.$$

As expected, the extension to the Hadamard manifold setting of the inexact versions of the proximal point algorithm in Hilbert spaces with two approximate rules (1.3) and (1.4) has been done in [59], where their convergence results are established. Given a sequence of parameters $\{\lambda_n\} \subseteq (0, +\infty)$ and a sequence of error controls $\{\varepsilon_n\} \subseteq (0, +\infty)$, the two inexact proximal point algorithms are formulated, respectively, as the following Algorithm IP₁ and Algorithm IP₂.

Algorithm IP₁. Give initial point $x_0 \in \mathcal{D}(A)$. Letting $n = 0, 1, 2, \dots$ and having x_n , determine x_{n+1} such that

$$(1.6) \quad d(x_{n+1}, y_n) \leq \varepsilon_n$$

with the associated $y_n \in M$ satisfying

$$(1.7) \quad 0 \in A(y_n) - \lambda_n \exp_{y_n}^{-1} x_n.$$

(The sequence $\{y_n\}$ is called an associated sequence generated by Algorithm IP₁).

Algorithm IP₂. Give initial point $x_0 \in \mathcal{D}(A)$. Letting $n = 0, 1, 2, \dots$ and having x_n , determine x_{n+1} such that

$$(1.8) \quad d(0, A_{\lambda_n, x_n}(x_{n+1})) \leq \lambda_n \varepsilon_n,$$

where, for each $n = 0, 1, \dots$, the multivalued vector field A_{λ_n, x_n} is defined by

$$A_{\lambda_n, x_n}(x) := A(x) - \lambda_n \exp_x^{-1} x_n \quad \text{for each } x \in M.$$

As pointed out in [59], it is unknown whether a sequence generated by Algorithm IP₂ is also a sequence generated by Algorithm IP₁ even in the case when A is maximal monotone (because whether (1.7) is solvable is unknown) in general. From the viewpoint of application, Algorithm IP₂ is more implementable than Algorithm IP₁.

The aim of the present paper is to continue to study the convergence properties of Algorithms IP₁ and IP₂. The interesting topics considered here are two: one is the convergence rate of these algorithms and the other is the finite termination of these algorithms. More precisely, we will establish in sections 3 and 4 that linear convergence and finite termination for these two inexact algorithms under the assumptions of metric subregularity and of weak sharp minima-like conditions, respectively. Furthermore, if the involved parameters $\lambda_n \rightarrow 0$, then the convergence rate is superlinear. In particular, some of our results are new (e.g., most of the results in section 3) even

in Euclidean spaces, while some corresponding result in [11] for the proximal point algorithm on Hadamard manifolds is improved/extended (as explained before Corollary 5.11). To the best of our knowledge, the present paper seems to explore the linear convergence property of the proximal point algorithms on Riemannian manifolds. As shown by an example provided in the last section (i.e., Example 6.4), the results in the present paper cannot be extended directly to the general Riemannian manifold setting.

The remainder of the paper is organized as follows. In section 2, some fundamental definitions, properties, and notations of Riemannian manifolds are provided. In section 3, convergence rates of Algorithms IP₁ and IP₂ are presented. Finite terminations of Algorithms IP₁ and IP₂ are explored in section 4. In section 5, applications to optimization problems are provided. In the last section, several examples are provided to illustrate our theorems obtained in the present paper.

2. Preliminaries. In this section we introduce some fundamental definitions, properties, and notations of Riemannian manifolds, which can be found in any textbook on Riemannian geometry, for example, [18, 48, 52].

Let M be a complete and connected m -dimensional manifold. We assume that M can be endowed with a Riemannian metric $\langle \cdot, \cdot \rangle$, with the corresponding norm denoted by $\| \cdot \|$, to become a *Riemannian* manifold. Let $p \in M$. The tangent space of M at p is denoted by T_pM and the tangent bundle of M by $TM = \bigcup_{p \in M} T_pM$, which is naturally a manifold. Given a piecewise smooth curve $\gamma : [a, b] \rightarrow M$ joining p to q (i.e., $\gamma(a) = p$ and $\gamma(b) = q$), we can define the length of γ by using the metric as $L(\gamma) = \int_a^b \|\gamma'(t)\| dt$. Then the Riemannian distance $d(p, q)$ is defined by minimizing this length over the set of all such curves joining p to q , which induces the original topology on M .

For a Banach space or a Riemannian manifold Z , we use $\mathbf{B}_Z(p, r)$ and $\overline{\mathbf{B}}_Z(p, r)$ to denote, respectively, the open metric ball and the closed metric ball at p with radius r , that is,

$$\mathbf{B}_Z(p, r) = \{q \in Z : d(p, q) < r\} \quad \text{and} \quad \overline{\mathbf{B}}_Z(p, r) = \{q \in Z : d(p, q) \leq r\}.$$

We often omit the subscript Z if no confusion arises. In particular, we use \mathbb{B}_p to denote the closed unit ball of T_pM , i.e.,

$$\mathbb{B}_p := \{v \in T_pM \mid \|v\| \leq 1\}.$$

Let ∇ be the Levi-Civita connection associated with $(M, \langle \cdot, \cdot \rangle)$. Let γ be a smooth curve in M . A vector field X is said to be parallel along γ if $\nabla_{\gamma'} X = 0$. If γ' itself is parallel along γ , we say that γ is a *geodesic*, and in this case $\|\gamma'\|$ is constant. When $\|\gamma'\| = 1$, γ is called normalized. A geodesic joining p to q in M is said to be minimal if its length equals $d(p, q)$. By the Hopf–Rinow theorem we know that if M is complete then any pair of points in M can be joined by a minimal geodesic. Moreover, (M, d) is a complete metric space and bounded closed subsets are compact.

We use $P_{\gamma, \cdot, \cdot}$ to denote the parallel transport on the tangent bundle TM along γ with respect to ∇ , which is defined by

$$P_{\gamma, \gamma(b), \gamma(a)}(v) = V(\gamma(b)) \quad \forall a, b \in \mathbb{R} \text{ and } v \in T_{\gamma(a)}M,$$

where V is the unique vector field satisfying $\nabla_{\gamma'(t)} V = 0$ for all t and $V(\gamma(a)) = v$. Then, for any $a, b \in \mathbb{R}$, $P_{\gamma, \gamma(b), \gamma(a)}$ is an isometry from $T_{\gamma(a)}M$ to $T_{\gamma(b)}M$. Note that,

for any $a, b, b_1, b_2 \in \mathbb{R}$,

$$P_{\gamma, \gamma(b_2), \gamma(b_1)} \circ P_{\gamma, \gamma(b_1), \gamma(a)} = P_{\gamma, \gamma(b_2), \gamma(a)} \quad \text{and} \quad P_{\gamma, \gamma(b), \gamma(a)}^{-1} = P_{\gamma, \gamma(a), \gamma(b)}.$$

We will write $P_{q,p}$ instead of $P_{\gamma, q, p}$ in the case when γ is a minimal geodesic joining p to q and no confusion arises.

Recall that the exponential map $\exp : TM \rightarrow M$ is defined as follows: for all $q \in M, v \in T_q M$,

$$\exp_q v = \gamma(1, q, v) = \gamma\left(\|v\|, q, \frac{v}{\|v\|}\right),$$

that is, $\exp_q v$ is obtained by going out the length equal to $\|v\|$, starting from q , along a geodesic γ which passes through q with velocity equal to $\frac{v}{\|v\|}$. Then, for any value of t , $\exp_q tv = \gamma(t, q, v)$. Note that the map \exp_q is differentiable on $T_q M$ for any $q \in M$.

A complete simply connected Riemannian manifold of nonpositive sectional curvature is called a *Hadamard manifold*. Throughout the remainder of this paper, we will always assume that M is an m -dimensional Hadamard manifold. Let $p \in M$. It follows from [52, p. 221, Theorem 4.1] that $\exp_p : T_p M \rightarrow M$ is a diffeomorphism, and for any two points $p, q \in M$ there exists a unique normalized geodesic joining p to q , which is, in fact, a minimal geodesic. Furthermore, for the Hadamard manifold, one of the most important properties is the following comparison theorem (cf. [52, p. 223, Proposition 4.5]), which is useful in our study. Let $\Delta(p_1 p_2 p_3)$ be a geodesic triangle of a Riemannian manifold, that is, $\Delta(p_1 p_2 p_3)$ is a set consisting of three points p_1, p_2, p_3 , and three minimal geodesics joining these points. Then, for each $i = 1, 2, 3 \pmod{3}$,

$$(2.1) \quad d^2(p_i, p_{i+1}) + d^2(p_{i+1}, p_{i+2}) - 2 \left\langle \exp_{p_{i+1}}^{-1} p_i, \exp_{p_{i+1}}^{-1} p_{i+2} \right\rangle \leq d^2(p_{i-1}, p_i).$$

A subset $K \subseteq M$ is said to be convex if for any two points p and q in K , the geodesic joining p to q is contained in K , that is, if $\gamma : [a, b] \rightarrow M$ is a geodesic such that $p = \gamma(a)$ and $q = \gamma(b)$, then $\gamma((1-t)a + tb) \in K$ for all $t \in [0, 1]$.

Let $\mathcal{X}(M)$ denote the set of all multivalued vector fields $A : M \rightrightarrows TM$ such that $A(x) \subseteq T_x M$ for each $x \in M$ and the effective domain $\mathcal{D}(A)$ of A is closed and convex, where the domain $\mathcal{D}(A)$ of A is defined by

$$\mathcal{D}(A) = \{x \in M : A(x) \neq \emptyset\}.$$

In the spirit of the corresponding notions in Hilbert spaces (cf. [14, 42] and [60] for example), Definition 2.1 below taken from [59] extends some notions of the monotonicity to multivalued vector fields on Hadamard manifolds. In particular, concepts (a), (b), (c) were introduced and studied in [44] for the single-valued case and in [16] for the multivalued case.

DEFINITION 2.1. *Let $A \in \mathcal{X}(M)$. A is said to be*

(a) *monotone if the following condition holds for any $x, y \in \mathcal{D}(A)$:*

$$(2.2) \quad \langle u, \exp_x^{-1} y \rangle \leq \langle v, -\exp_y^{-1} x \rangle \quad \forall u \in A(x) \text{ and } \forall v \in A(y);$$

(b) *strictly monotone if (2.2) holds with strict inequality for any $x, y \in \mathcal{D}(A)$ with $x \neq y$, that is,*

$$(2.3) \quad \langle u, \exp_x^{-1} y \rangle < \langle v, -\exp_y^{-1} x \rangle \quad \forall u \in A(x) \text{ and } \forall v \in A(y);$$

(c) ρ -strongly monotone if there exists $\rho > 0$ such that, for any $x, y \in \mathcal{D}(A)$,

$$(2.4) \quad \langle u, \exp_x^{-1} y \rangle - \langle v, -\exp_y^{-1} x \rangle \leq -\rho d^2(x, y) \quad \forall u \in A(x) \text{ and } \forall v \in A(y);$$

(d) maximal monotone if it is monotone and the following implication holds for any $x \in M$ and $u \in T_x M$:

$$(2.5) \quad \langle u, \exp_x^{-1} y \rangle \leq \langle v, -\exp_y^{-1} x \rangle \quad \forall y \in \mathcal{D}(A) \text{ and } v \in A(y) \implies u \in A(x).$$

Example 2.1. Let $y \in M$. Define a vector field $V_y : M \rightarrow TM$ by

$$V_y(x) := -\exp_x^{-1} y.$$

Then, by [58], V_y is 1-strongly maximal monotone. Furthermore, for any $\lambda > 0$ and any monotone vector field A on M , the multivalued vector field $A_{\lambda,y}$ defined by

$$A_{\lambda,y}(x) := A(x) - \lambda \exp_x^{-1} y \quad \text{for each } x \in M$$

is λ -strongly monotone; see [33] for example.

Let A be a multivalued vector field. We use $A^{-1}(0)$ to denote the set of all singularities of A , that is,

$$A^{-1}(0) := \{x \in \mathcal{D}(A) : 0 \in A(x)\}.$$

The following proposition is known in [33, Theorem 4.3] and provides a sufficient condition ensuring the existence of singularities of A .

PROPOSITION 2.2. *Let $A \in \mathcal{X}(M)$ be a maximal strongly monotone vector field with the domain $\mathcal{D}(A) = M$. Then there exists a unique singularity of A .*

Remark 2.1. Note that Algorithm P is an implicit method. Thus one basic problem is when this algorithm is well-defined. For each n , define $A_{\lambda_n, x_n} \in \mathcal{X}(M)$ by

$$A_{\lambda_n, x_n}(x) := A(x) - \lambda_n \exp_x^{-1} x_n \quad \forall x \in \mathcal{D}(A).$$

Then, in the case when $A \in \mathcal{X}(M)$ is monotone, each A_{λ_n, x_n} is strongly monotone. Thus, in view of Proposition 2.2, the following assertions hold when A is monotone.

- (i) Algorithm P is well-defined if and only if $A_{\lambda_n, x_n}^{-1}(0) \neq \emptyset$ for each $n = 0, 1, 2, \dots$
- (ii) If $\mathcal{D}(A) = M$ and A is maximal monotone, then Algorithm P is well-defined.

The following proposition taken from [59] shows that if (1.7) is solvable for each n , then any sequence generated by Algorithm IP₂ with initial point x_0 is also a sequence generated by Algorithm IP₁ with the same initial point.

PROPOSITION 2.3. *Suppose that $A \in \mathcal{X}(M)$ is a maximal monotone multivalued vector field. Let $\{\lambda_n\}$ and $\{\varepsilon_n\}$ be two sequences of positive numbers. Let $x_0 \in M$ and let $\{x_n\}$ be a sequence generated by Algorithm IP₂. Suppose that, for each n , the problem*

$$(2.6) \quad 0 \in A(x) - \lambda_n \exp_x^{-1} x_n$$

is solvable. Then $\{x_n\}$ can also be regarded as a sequence generated by Algorithm IP₁ with same initial point x_0 .

Consider a sequence $\{x_n\}$ in M . For convenience, we write

$$(2.7) \quad \delta_n := \frac{\varepsilon_n}{d(x_{n+1}, x_n)} \quad \text{for each } n \in \mathbb{N},$$

where we adopt the convention that $\frac{0}{0} = 0$. Throughout the paper, we always assume that $A \in \mathcal{X}(M)$ is maximal monotone with $A^{-1}(0) \neq \emptyset$, and let $x_0 \in \mathcal{D}(A)$ be a given initial point. We first need the following lemma about the properties of sequences generated by Algorithm IP₁.

LEMMA 2.4. *Let $\{x_n\}$ be a sequence generated by Algorithm IP₁ (together with the associated sequence $\{y_n\}$) and let $\bar{x} \in A^{-1}(0)$. Then the following assertions hold:*

(i) *For any $n = 0, 1, \dots$, the following inequality holds:*

$$(2.8) \quad d^2(y_n, \bar{x}) + d^2(y_n, x_n) \leq d^2(x_n, \bar{x}).$$

(ii) *The following implication holds:*

$$(2.9) \quad \sum_{n \geq 0} \delta_n < +\infty \implies \sum_{n \geq 0} \varepsilon_n < +\infty.$$

(iii) *If $\{\varepsilon_n\} \subset [0, +\infty)$ satisfies*

$$(2.10) \quad \sum_{n \geq 0} \varepsilon_n < +\infty,$$

then the sequences $\{d(x_n, \bar{x})\}$ and $\{d(y_n, \bar{x})\}$ are convergent for any $\bar{x} \in A^{-1}(0)$.

Proof. Assertions (i) and (iii) are known in [59, Lemma 3.7].

(ii). Suppose that $\sum_n \delta_n < +\infty$. Fix n . Then, by the definition, one has that

$$(2.11) \quad \varepsilon_n = \delta_n d(x_{n+1}, x_n).$$

Let $\bar{x} \in A^{-1}(0)$. We observe by (2.8) that

$$(2.12) \quad d(y_n, \bar{x}) \leq d(x_n, \bar{x}) \quad \text{for each } n.$$

Hence

$$(2.13) \quad d(x_{n+1}, \bar{x}) \leq d(x_{n+1}, y_n) + d(y_n, \bar{x}) \leq d(x_n, \bar{x}) + \varepsilon_n \quad \text{for each } n.$$

Combining (2.11) and the second inequality of (2.13), we conclude that

$$d(x_{n+1}, \bar{x}) \leq \delta_n d(x_{n+1}, x_n) + d(x_n, \bar{x}) \leq \delta_n (d(x_{n+1}, \bar{x}) + d(x_n, \bar{x})) + d(x_n, \bar{x}).$$

It follows that

$$(2.14) \quad d(x_{n+1}, \bar{x}) \leq \frac{1 + \delta_n}{1 - \delta_n} d(x_n, \bar{x}) = \left(1 + \frac{2\delta_n}{1 - \delta_n}\right) d(x_n, \bar{x}).$$

Since $\sum_{n \geq 0} \frac{2\delta_n}{1 - \delta_n} < \infty$ by assumption, we apply [49, Lemma 2.2.2] to obtain that the sequence $\{d(x_n, \bar{x})\}$ is convergent and so $\{d(x_{n+1}, x_n)\}$ is bounded. It follows from (2.11) that

$$\sum_{n \geq 0} \varepsilon_n \leq \sup_n d(x_{n+1}, x_n) \sum_{n \geq 0} \delta_n < +\infty$$

and so $\sum_{n \geq 0} \varepsilon_n \leq +\infty$. The proof is complete. \square

Note that, in general, a sequence generated by Algorithm IP₂ cannot be regarded as a sequence generated by Algorithm IP₁. The following lemma is an analogue to Lemma 2.4 for Algorithm IP₂.

LEMMA 2.5. *Let $\{x_n\}$ be a sequence generated by Algorithm IP₂ and let $\bar{x} \in A^{-1}(0)$. Then the following assertions hold:*

(i) *For any $n \in \mathbb{N}$, the following inequality holds:*

$$(2.15) \quad d^2(x_{n+1}, \bar{x}) + \frac{1}{2}d^2(x_{n+1}, x_n) \leq \frac{1}{1 - 2\delta_n^2}d^2(x_n, \bar{x}).$$

(ii) *The following implication holds:*

$$(2.16) \quad \sum_{n \geq 0} \delta_n^2 < +\infty \implies \sum_{n \geq 0} \varepsilon_n < +\infty.$$

(iii) *If $\{\varepsilon_n\} \subset [0, +\infty)$ satisfies (2.10), then the sequence $\{d(x_n, \bar{x})\}$ is convergent for each $\bar{x} \in A^{-1}(0)$.*

Proof. Assertions (i) and (iii) are known in [59, Lemma 3.9]. Below, we show that assertion (ii) holds. Suppose that $\sum_{n \geq 0} \delta_n^2 < +\infty$. Note by definition that $\delta_n = \frac{\varepsilon_n}{d(x_{n+1}, x_n)}$. Then

$$(2.17) \quad 2\varepsilon_n \leq \delta_n^2 + d^2(x_{n+1}, x_n).$$

Let N be a positive integer such that $2\delta_n^2 < 1$ for all $n \geq N$. Fix $n \geq N$ and let $\bar{x} \in A^{-1}(0)$. Then, applying (2.15) to $2\delta_n^2$ in place of t , we have the following assertion:

$$(2.18) \quad d^2(x_{n+1}, \bar{x}) \leq \left(1 + \frac{2\delta_n^2}{1 - 2\delta_n^2}\right) d^2(x_n, \bar{x}) - \frac{1}{2}d^2(x_{n+1}, x_n).$$

This implies that

$$\frac{1}{2}d^2(x_{n+1}, x_n) \leq (d^2(x_n, \bar{x}) - d^2(x_{n+1}, \bar{x})) + \frac{2\delta_n^2 d^2(x_n, \bar{x})}{1 - 2\delta_n^2};$$

hence,

$$(2.19) \quad \sum_{n \geq N} \frac{1}{2}d^2(x_{n+1}, x_n) \leq \sum_{n \geq N} (d^2(x_n, \bar{x}) - d^2(x_{n+1}, \bar{x})) + \sum_{n \geq N} \frac{2\delta_n^2 d^2(x_n, \bar{x})}{1 - 2\delta_n^2}.$$

It follows from (2.17) and (2.18) that [49, Lemma 2.2.2] is applicable and so $\{d^2(x_n, \bar{x})\}$ is convergent (and so bounded). Consequently, (2.19) implies that $\sum d^2(x_{n+1}, x_n) < +\infty$ thanks to the fact that $\sum_n \delta_n^2 < +\infty$. Therefore, $\sum_n \varepsilon_n < +\infty$ by (2.17), and the proof is complete. \square

3. Metrical subregularity and rate of convergence. Let $A : M \rightarrow 2^{TM}$ be a multivalued vector field, and consider the following generalized singularity problem:

$$(3.1) \quad 0 \in A(x).$$

This section is devoted to the study of the convergence rate of the inexact proximal point algorithm for solving (3.1) under metric regular condition. First, we recall some

notions and notations. Let X be a metric space and let $x \in X$. Let $S_1 \subseteq X$. The distance from x to S_1 is defined by

$$d(x, S_1) := \inf\{d(x, z) \mid z \in S_1\}.$$

Let $v \in T_z M$ for some $z \in M$. Then, we define

$$A^{-1}(v) := \{x \in M : P_{x,z}v \in A(x)\},$$

while $\text{gph}A$ and $\text{dom}A$ are defined similarly.

The metric subregularity (the terminology ‘‘error bond’’ is sometimes adopted instead of metric subregularity; see [19] for history of terminology) and its equivalent calmness counterpart for inverse mappings have been studied extensively in linear spaces. Various results in this direction and their applications can be found in [1, 2, 3, 19, 23, 27, 28, 30, 62, 63] and references therein. Below, we extend the notion of metric subregularity (cf. [20]) to Riemannian manifolds.

DEFINITION 3.1. *Let $A : M \rightarrow 2^{TM}$ be a multivalued vector field. Let $\kappa \geq 0$ and $U \subset M$. A is said to be*

- (a) *metrically subregular on U with modulus κ if*

$$(3.2) \quad d(x, A^{-1}(0)) \leq \kappa d(0, A(x)) \quad \text{for all } x \in U;$$

- (b) *metrically subregular if, for any $\bar{x} \in A^{-1}(0)$, there exist a neighborhood $U_{\bar{x}} \subset M$ of \bar{x} and $\kappa_{\bar{x}} > 0$ such that A is metrically subregular on $U_{\bar{x}}$ with modulus $\kappa_{\bar{x}}$;*

- (c) *global metrically subregular with modulus κ if A is metrically subregular on M with modulus κ , that is,*

$$(3.3) \quad d(x, A^{-1}(0)) \leq \kappa d(0, A(x)) \quad \text{for all } x \in M.$$

Remark 3.1. Let $\bar{x} \in A^{-1}(0)$, and define

$$\text{sreg}_{\bar{x}}A := \lim_{\rho \rightarrow 0} \sup_{x \in \mathbf{B}(\bar{x}, \rho) \setminus A^{-1}(0)} \frac{d(x, A^{-1}(0))}{d(0, A(x))}.$$

Then A is metrically subregular if and only if $\text{sreg}_{\bar{x}}A < +\infty$ for each $\bar{x} \in A^{-1}(0)$.

Below, we get a lemma about the convergence rate of sequences generated by Algorithms IP₁ and IP₂, which will be useful. Recall that $\{\delta_n\}$ is defined by (2.7):

$$\delta_n = \frac{\varepsilon_n}{d(x_n, x_{n+1})} \quad \text{for each } n.$$

Let $\kappa > 0$ and set

$$(3.4) \quad \mu_n := \kappa \lambda_n (1 + \delta_n) \quad \text{for each } n.$$

LEMMA 3.2. *Let $N \in \mathbb{N}$, $\kappa > 0$, and let $U \subseteq M$. Suppose that A is metrically subregular on U with modulus κ .*

- (i) *Let $\{x_n\}$ be a sequence generated by Algorithm IP₁ (together with the associated sequence $\{y_n\}$) such that $\{y_n : n \geq N\} \subset U$ and*

$$(3.5) \quad \sup_{n \geq N} \delta_n < 1.$$

Then, one has that, for each $n \geq N$,

$$(3.6) \quad d(x_{n+1}, A^{-1}(0)) \leq \left(\frac{\delta_n}{1 - \delta_n} + \sqrt{\frac{\lambda_n^2 \kappa^2}{1 + \lambda_n^2 \kappa^2}} \right) d(x_n, A^{-1}(0)).$$

(ii) Let $\{x_n\}$ be a sequence generated by Algorithm IP₂ such that $\{x_n : n \geq N\} \subset U$ and (3.5) holds. Then one has that, for each $n \geq N$,

$$(3.7) \quad d(x_{n+1}, A^{-1}(0)) \leq \frac{\sqrt{2}\mu_n}{\sqrt{(1 + 2\mu_n^2)(1 - 2\delta_n^2)}} d(x_n, A^{-1}(0)),$$

where μ_n is given by (3.4).

Proof. (i) Let $z \in A^{-1}(0)$ and fix n . By (2.8), we have $d(y_n, x_n) \leq d(x_n, z)$. As $z \in A^{-1}(0)$ is arbitrary, it follows that

$$(3.8) \quad d(y_n, x_n) \leq d(x_n, A^{-1}(0)).$$

Observe further that

$$d(x_{n+1}, y_n) \leq \varepsilon_n = \delta_n d(x_{n+1}, x_n) \leq \delta_n (d(x_{n+1}, y_n) + d(y_n, x_n)).$$

This gives that

$$d(x_{n+1}, y_n) \leq \frac{\delta_n}{1 - \delta_n} d(x_n, y_n).$$

Combining this with (3.8) yields

$$(3.9) \quad d(x_{n+1}, y_n) \leq \frac{\delta_n}{1 - \delta_n} d(x_n, A^{-1}(0)).$$

This implies that

$$(3.10) \quad d(x_{n+1}, A^{-1}(0)) \leq d(x_{n+1}, y_n) + d(y_n, A^{-1}(0)) \leq \frac{\delta_n}{1 - \delta_n} d(x_n, A^{-1}(0)) + d(y_n, A^{-1}(0)).$$

Since $\lambda_n \exp_{y_n}^{-1} x_n \in A(y_n)$ by the definition of Algorithm IP₁, it follows that

$$(3.11) \quad d(0, A(y_n)) \leq \lambda_n \|\exp_{y_n}^{-1} x_n\|.$$

By assumption, we get that for each $n \geq N_1$,

$$d(y_n, A^{-1}(0)) \leq \kappa d(0, A(y_n)).$$

Combining this with (3.11) gives that

$$(3.12) \quad \|\exp_{y_n}^{-1} x_n\| \geq \frac{1}{\kappa \lambda_n} d(y_n, A^{-1}(0)).$$

Write $z = P_{A^{-1}(0)}(x_n)$. Then

$$\begin{aligned} d^2(y_n, A^{-1}(0)) &\leq d^2(y_n, z) \\ &\leq d^2(x_n, z) - d^2(x_n, y_n) \\ &\leq d^2(x_n, A^{-1}(0)) - \frac{1}{\kappa^2 \lambda_n^2} d^2(y_n, A^{-1}(0)), \end{aligned}$$

where the second inequality holds because of (2.8), while the last inequality holds because of (3.12). This implies that

$$d(y_n, A^{-1}(0)) \leq \sqrt{\frac{\lambda_n^2 \kappa^2}{1 + \lambda_n^2 \kappa^2}} d(x_n, A^{-1}(0)).$$

Combining this with (3.10) yields that for all $n \geq N_1$,

$$d(x_{n+1}, A^{-1}(0)) \leq \left(\frac{\delta_n}{1 - \delta_n} + \sqrt{\frac{\lambda_n^2 \kappa^2}{1 + \lambda_n^2 \kappa^2}} \right) d(x_n, A^{-1}(0)).$$

Hence, (3.6) is seen to hold.

(ii) For each n , by the definition of Algorithm IP₂, one can choose $v_{n+1} \in T_{x_{n+1}}M$ such that

$$(3.13) \quad \lambda_n v_{n+1} \in \Lambda_n(x_{n+1}) = A(x_{n+1}) - \lambda_n \exp_{x_{n+1}}^{-1} x_n$$

and

$$\lambda_n \|v_{n+1}\| \leq d(0, \Lambda_n(x_{n+1})) \leq \lambda_n \varepsilon_n.$$

This, together with the definition of δ_n , gives that

$$\lambda_n \|v_{n+1}\| \leq \lambda_n \delta_n d(x_{n+1}, x_n).$$

We get by assumption that, for each $n \geq N$,

$$(3.14) \quad d(x_n, A^{-1}(0)) \leq \kappa d(0, A(x_n)).$$

Observe further that $\lambda_n v_{n+1} + \lambda_n \exp_{x_{n+1}}^{-1} x_n \in A(x_{n+1})$ by (3.13). Thus, it follows from (3.14) that

$$(3.15) \quad d(x_{n+1}, A^{-1}(0)) \leq \kappa \|\lambda_n v_{n+1} + \lambda_n \exp_{x_{n+1}}^{-1} x_n\| \leq \kappa \lambda_n (1 + \delta_n) d(x_n, x_{n+1}).$$

Applying (2.15), we conclude that

$$d^2(x_{n+1}, A^{-1}(0)) + \frac{1}{2} d^2(x_{n+1}, x_n) \leq \frac{1}{1 - 2\delta_n^2} d^2(x_n, A^{-1}(0)).$$

This, together with (3.15), gives that

$$\begin{aligned} & (\kappa \lambda_n (1 + \delta_n))^2 d^2(x_{n+1}, A^{-1}(0)) + \frac{1}{2} d^2(x_{n+1}, A^{-1}(0)) \\ & \leq (\kappa \lambda_n (1 + \delta_n))^2 \left[d^2(x_{n+1}, A^{-1}(0)) + \frac{1}{2} d^2(x_{n+1}, x_n) \right] \\ & \leq \frac{(\kappa \lambda_n (1 + \delta_n))^2}{1 - 2\delta_n^2} d^2(x_n, A^{-1}(0)). \end{aligned}$$

Therefore,

$$d^2(x_{n+1}, A^{-1}(0)) \leq \frac{(\kappa \lambda_n (1 + \delta_n))^2}{\left(\frac{1}{2} + (\kappa \lambda_n (1 + \delta_n))^2\right)(1 - 2\delta_n^2)} d^2(x_n, A^{-1}(0)),$$

and (3.7) is seen to hold. The proof is complete. \square

LEMMA 3.3. Let $\{x_n\}$ be a sequence generated by Algorithm IP₁ (together with the associated sequence $\{y_n\}$) or Algorithm IP₂ such that

$$(3.16) \quad \lim_{n \rightarrow \infty} \frac{\varepsilon_n}{d(x_n, x_{n+1})} = 0 \quad \text{and} \quad \sup_n \lambda_n < \infty.$$

Suppose that there exist $\kappa \in [0, +\infty)$, $N \in \mathbb{N}$, and $U \subseteq M$ such that

- (a) A is metrically subregular on U with modulus κ , and
- (b) $\{x_n : n \geq N\} \subset U$ and $\{y_n : n \geq N\} \subset U$.

Then there exist $\bar{x} \in A^{-1}(0)$, $q \in [\frac{1}{2}, 1)$, and a bounded sequence $\{a_n\} \subset (0, +\infty)$ satisfying

$$(3.17) \quad \overline{\lim}_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} \leq 2\sqrt{2}\kappa \overline{\lim}_n \lambda_n$$

such that

$$(3.18) \quad d(x_n, \bar{x}) \leq a_n q^{n-N} d(x_N, A^{-1}(0)) \quad \text{for any } n > N;$$

hence $\{x_n\}$ converges linearly to \bar{x} .

Proof. By assumption (3.16), there exists $q \in [\frac{1}{2}, 1)$ such that the following inequalities hold for any $n \geq N$ (N can be taken to be larger if necessary):

$$(3.19) \quad \delta_n \leq \frac{1}{2}, \quad \frac{\delta_n}{1 - \delta_n} + \sqrt{\frac{\lambda_n^2 \kappa^2}{1 + \lambda_n^2 \kappa^2}} < q, \quad \text{and} \quad \frac{\sqrt{2}\mu_n}{\sqrt{(1 + 2\mu_n^2)(1 - 2\delta_n^2)}} < q,$$

where μ_n is given by (3.4). Note further by assumption that Lemma 3.2 is applicable. Define the sequence $\{a_n\}$ by

$$(3.20) \quad a_n := \frac{4}{(1 - q)q^{n-N}} \prod_{l=N}^{n-1} \left(\frac{\delta_l}{1 - \delta_l} + \sqrt{\frac{\lambda_l^2 \kappa^2}{1 + \lambda_l^2 \kappa^2}} \right) \quad \text{for any } n > N$$

if $\{x_n\}$ is generated by Algorithm IP₁; otherwise,

$$(3.21) \quad a_n := \frac{4}{(1 - q)q^{n-N}} \prod_{l=N}^{n-1} \frac{\sqrt{2}\mu_l}{\sqrt{(1 + 2\mu_l^2)(1 - 2\delta_l^2)}} \quad \text{for any } n > N.$$

Then, $\{a_n\}$ is decreasing and so bounded thanks to the second inequality in (3.19), and (3.17) holds because, by (3.16),

$$\overline{\lim}_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = \frac{1}{q} \overline{\lim}_{n \rightarrow \infty} \left(\frac{\delta_n}{1 - \delta_n} + \sqrt{\frac{\lambda_n^2 \kappa^2}{1 + \lambda_n^2 \kappa^2}} \right) = \frac{\kappa \overline{\lim}_n \lambda_n}{q \sqrt{1 + \kappa^2 \overline{\lim}_n \lambda_n^2}} \leq \frac{\kappa \overline{\lim}_n \lambda_n}{q}$$

if $\{a_n\}$ is defined by (3.20); otherwise,

$$\overline{\lim}_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = \frac{1}{q} \overline{\lim}_{n \rightarrow \infty} \left(\frac{\sqrt{2}\mu_n}{\sqrt{(1 + 2\mu_n^2)(1 - 2\delta_n^2)}} \right) = \frac{\sqrt{2}\kappa \overline{\lim}_n \lambda_n}{q \sqrt{1 + 2\kappa^2 \overline{\lim}_n \lambda_n^2}} \leq \frac{\sqrt{2}\kappa \overline{\lim}_n \lambda_n}{q}.$$

Below we shall show that

$$(3.22) \quad d(x_{n+1}, x_n) \leq a_n (1 - q) q^{n-N} d(x_N, A^{-1}(0)) \quad \text{for each } n > N.$$

Granting this, we have that, for any $m \in \mathbb{N}$ and $n > N$,

$$(3.23) \quad d(x_{n+m}, x_n) \leq a_n(1-q)q^{n-N}(1+q+\cdots+q^{m-1})d(x_N, A^{-1}(0))$$

as $\{a_n, n \geq N+1\}$ is monotone decreasing. Passing to the limit as $m \rightarrow \infty$, one sees that (3.18) holds.

Thus, to complete the proof, it remains to show (3.22). We only prove this for the case when $\{x_n\}$ is generated by Algorithm IP₁ as the proof is similar for the other case. Then, by Lemma 3.2, for each $n > N$, (3.6) holds, and so

$$(3.24) \quad d(x_n, A^{-1}(0)) \leq \prod_{l=N}^{n-1} \left(\frac{\delta_l}{1-\delta_l} + \sqrt{\frac{\lambda_l^2 \kappa^2}{1+\lambda_l^2 \kappa^2}} \right) d(x_N, A^{-1}(0)).$$

In terms of the definition of $\{a_n\}$, one has that

$$(3.25) \quad d(x_n, A^{-1}(0)) \leq \frac{(1-q)a_n}{4} q^{n-N} d(x_N, A^{-1}(0)) \quad \text{for each } n > N.$$

Fix $n \geq N$ and let $z := P_{A^{-1}(0)}(x_n)$. Then $d(y_n, z) \leq d(x_n, z) = d(x_n, A^{-1}(0))$. Thus, using (1.6) and recalling (2.7), we conclude that

$$d(x_{n+1}, z) \leq d(x_{n+1}, y_n) + d(y_n, z) \leq \delta_n d(x_{n+1}, x_n) + d(x_n, z).$$

Hence,

$$d(x_{n+1}, x_n) \leq d(x_{n+1}, z) + d(x_n, z) \leq \delta_n d(x_{n+1}, x_n) + 2d(x_n, z).$$

This implies that

$$d(x_{n+1}, x_n) \leq \frac{2d(x_n, z)}{1-\delta_n} = \frac{2}{1-\delta_n} d(x_n, A^{-1}(0)) \leq 4d(x_n, A^{-1}(0))$$

(noting that $\delta_n \leq \frac{1}{2}$ and $d(x_n, z) = d(x_n, A^{-1}(0))$). Combining this with (3.25) yields that (3.22) holds, and the proof is complete. \square

THEOREM 3.4. *Suppose that*

- (a) *A is globally metrically subregular, or*
- (b) *A is metrically subregular and $\sum_n \varepsilon_n < +\infty$.*

Let $\{x_n\}$ be a sequence generated by Algorithms IP₁ or IP₂ such that (3.16) holds. Then $\{x_n\}$ converges linearly to a point $\bar{x} \in A^{-1}(0)$: there exist $q \in [\frac{1}{2}, 1)$ and a bounded sequence $\{a_n\} \subset (0, +\infty)$ satisfying

$$(3.26) \quad \overline{\lim}_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} \leq 2\sqrt{2} \text{sreg}_{\bar{x}} A \overline{\lim}_n \lambda_n$$

such that

$$(3.27) \quad d(x_n, \bar{x}) \leq a_n q^n d(x_0, A^{-1}(0)) \quad \text{for any } n \geq 1.$$

In particular, if $\lim_{n \rightarrow \infty} \lambda_n = 0$ or $\text{sreg}_{\bar{x}} A = 0$, then $\{x_n\}$ is superlinearly convergent.

Proof. We first note $x_n \rightarrow \bar{x} \in A^{-1}(0)$, which is clear by Lemma 3.3 in case (a) (applied to some $\kappa > 0$ and $U := M$), and by [59, Theorems 3.8 and 3.11] in case (b) (noting that $\sum_n \frac{1}{\lambda_n^2} = +\infty$ and $\lim_{n \rightarrow \infty} \lambda_n \varepsilon_n = 0$ hold by (3.16) and (b) assumed here).

To complete the proof, it suffices to show that there exist $q \in [\frac{1}{2}, 1)$, $N \in \mathbb{N}$, and a bounded sequence $\{a_n\} \subset (0, +\infty)$ such that (3.18) and (3.26) hold. To this purpose, let $\varepsilon > 0$. By the definition of $\text{sreg}_{\bar{x}}A$, there exists $\delta > 0$ such that

$$\sup_{x \in \mathbf{B}(\bar{x}, \delta) \setminus A^{-1}(0)} \frac{d(x, A^{-1}(0))}{d(0, A(x))} \leq \text{sreg}_{\bar{x}}A + \varepsilon;$$

hence, assumption (a) of Lemma 3.3 holds with $\mathbf{B}(\bar{x}, \delta)$ and $\text{sreg}_{\bar{x}}A + \varepsilon$ in place of U and κ . Moreover, one can choose a positive integer N such that $\{x_n : n \geq N\} \subset U$ and $\{y_n : n \geq N\} \subset U$ because $x_n \rightarrow \bar{x}$ and $\|y_n - \bar{x}\| \leq \|x_n - \bar{x}\|$ for each n . Thus Lemma 3.3 is applicable to concluding that there exist $q \in [\frac{1}{2}, 1)$ and a bounded sequence $\{a_n\} \subset (0, +\infty)$ satisfying (3.17) such that (3.18) holds. Passing to the limit, one sees that (3.26) follows from (3.17), and the proof is complete. \square

Corollary 3.5 follows directly from Lemmas 2.4 and 2.5 and Theorem 3.4.

COROLLARY 3.5. *Let $\{x_n\}$ be a sequence generated by Algorithm IP_1 with $\sum_n \delta_n < \infty$, or by Algorithm IP_2 with $\sum_n \delta_n^2 < \infty$. Suppose that A is metrically subregular and $\sup_n \lambda_n < \infty$. Then, $\{x_n\}$ is linearly convergent to a point $\bar{x} \in A^{-1}(0)$. Moreover, if $\lim_{n \rightarrow \infty} \lambda_n = 0$ or $\text{sreg}_{\bar{x}}A = 0$, then $\{x_n\}$ is superlinearly convergent.*

If $\varepsilon_n = 0$ for each $n = 0, 1, \dots$, then $\delta_n = 0$. Thus, we get the following corollary directly from Theorem 3.4 for the convergence rate of the classical proximal point Algorithm P.

COROLLARY 3.6. *Let $\{x_n\}$ be a sequence generated by Algorithm P. Suppose that A is metrically subregular and $\sup_n \lambda_n < \infty$. Then, $\{x_n\}$ is linearly convergent to a point $\bar{x} \in A^{-1}(0)$. Moreover, if $\lim_{n \rightarrow \infty} \lambda_n = 0$ or $\text{sreg}_{\bar{x}}A = 0$, then $\{x_n\}$ is superlinearly convergent.*

The notion of Lipschitz continuous for a multivalued mapping in Banach spaces was given in [51]. Below, we extend this notion to the Riemannian manifold setting.

DEFINITION 3.7. *Let $A : M \rightarrow 2^{TM}$ be a multivalued vector field. Let $\kappa \geq 0$. Then A^{-1} is said to be Lipschitz continuous at 0 with modulus κ if $A^{-1}(0) = \{\bar{x}\}$ and there exists $r > 0$ such that*

$$d(x, \bar{x}) \leq \kappa \|w\| \quad \text{for each } x \in M \text{ and } w \in \overline{\mathbf{B}(0, r)} \cap A(x).$$

Proposition 3.8 below provides some sufficient conditions ensuring the Lipschitz continuity at 0 of A^{-1} .

PROPOSITION 3.8. *Suppose that $A \in \mathcal{X}(M)$ is monotone and $\bar{x} \in A^{-1}(0)$. Then the following assertions hold:*

- (i) *If A is ρ -strongly monotone at \bar{x} with $\rho > 0$, that is,*

$$\langle P_{\bar{x}, x} w, \exp_{\bar{x}}^{-1} x \rangle \geq \rho d^2(x, \bar{x}) \quad \text{for each } x \in \mathcal{D}(A) \text{ and } w \in A(x),$$

then A^{-1} is Lipschitz continuous at 0 with modulus $\kappa := \rho^{-1}$.

- (ii) *If $0 \in \text{int}A(\bar{x})$, then A^{-1} is Lipschitz continuous at 0 with modulus $\kappa := 0$.*

Proof. (i). The proof is trivial.

(ii). By definition, to complete the proof, it's sufficient to show that there exists $\varepsilon > 0$ such that the following implication holds:

$$(3.28) \quad \text{for any } x \in M, \text{ and } w \in A(x) \text{ with } \|w\| < \varepsilon \implies x = \bar{x}.$$

Since $0 \in \text{int}A(\bar{x})$, there exists $\varepsilon > 0$ such that for each $\bar{w} \in T_{\bar{x}}M$ with $\|\bar{w}\| < \varepsilon$, one has that $\bar{w} \in \text{int}A(\bar{x})$. Write

$$\mathbf{B}(0, \varepsilon) := \{\bar{w} \in T_{\bar{x}}M : \|\bar{w}\| < \varepsilon\}.$$

Then $\mathbf{B}(0, \varepsilon) \subset A(\bar{x})$. Let $x \in M$ be such that $x \neq \bar{x}$. Let $w \in A(x)$. Let $\bar{w} \in \mathbf{B}(0, \varepsilon) \subset A(\bar{x})$. Let $\gamma \in \Gamma_{\bar{x}x}$ be such that $\|\gamma'(0)\| = \|\gamma'(1)\| = d(\bar{x}, x)$. Noting that A is monotone, we obtain

$$(3.29) \quad \langle \bar{w}, \gamma'(0) \rangle \leq \langle w, \gamma'(1) \rangle.$$

Since $\bar{w} \in \mathbf{B}(0, \varepsilon) \subset A(\bar{x})$ is arbitrary, it follows from (3.29) that

$$\sup_{\bar{w} \in \mathbf{B}(0, \varepsilon)} \langle \bar{w}, \gamma'(0) \rangle \leq \langle w, \gamma'(1) \rangle,$$

which implies that

$$\varepsilon \|\gamma'(0)\| \leq \|w\| \|\gamma'(1)\|$$

and so $\|w\| \geq \varepsilon$. Hence, we get the following implication:

$$\text{for any } x \in M \text{ with } x \neq \bar{x}, \text{ and } w \in A(x) \implies \|w\| \geq \varepsilon,$$

which is equivalent to (3.28). The proof is complete. \square

Remark 3.2. Let A be any monotone, and let $\bar{x} \in A^{-1}(0)$. For any $\lambda > 0$, let $A_{\lambda, \bar{x}}$ be the multivalued vector field given in Example 2.1 with \bar{x} in place of y . Then $A_{\lambda, \bar{x}}$ is λ -strongly monotone at \bar{x} , and so $A_{\lambda, \bar{x}}^{-1}$ is Lipschitz continuous at 0 with modulus $\kappa := \lambda$ thanks to Proposition 3.8(i).

Remark 3.3. Suppose that A^{-1} is Lipschitz continuous at 0. Then A is metrically subregular and $A^{-1}(0) = \{\bar{x}\}$. Thus, by Lemma 3.2, there exists a positive N such that, for each $n \geq N$,

$$d(x_{n+1}, \bar{x}) \leq \left(\frac{\delta_n}{1 - \delta_n} + \sqrt{\frac{\lambda_n^2 \kappa^2}{1 + \lambda_n^2 \kappa^2}} \right) d(x_n, \bar{x})$$

in the case when $\{x_n\}$ is a sequence generated by Algorithm IP₁ with $\sum_n \delta_n < \infty$, and

$$d(x_{n+1}, \bar{x}) \leq \frac{\sqrt{2}\mu_n}{\sqrt{(1 + 2\mu_n^2)(1 - 2\delta_n^2)}} d(x_n, \bar{x})$$

in the case when $\{x_n\}$ is a sequence generated by Algorithm IP₂ with $\sum_n \delta_n^2 < \infty$.

Then the following corollary follows directly.

COROLLARY 3.9. *Let $\{x_n\}$ be a sequence generated by Algorithm IP₁ with $\sum_n \delta_n < \infty$, or by Algorithm IP₂ with $\sum_n \delta_n^2 < \infty$. Suppose that A^{-1} is Lipschitz continuous at 0 with modulus $\kappa \geq 0$ and $\sup_n \lambda_n < \infty$. Then, $\{x_n\}$ is (Q-)linearly convergent to the unique point $\bar{x} \in A^{-1}(0)$:*

$$\overline{\lim}_{n \rightarrow \infty} \frac{d(x_{n+1}, \bar{x})}{d(x_n, \bar{x})} \leq 1.$$

Furthermore, if $\lim_{n \rightarrow \infty} \lambda_n = 0$ or $\kappa = 0$ (e.g., $0 \in \text{int}A(\bar{x})$), then $\{x_n\}$ is (Q-)superlinearly convergent:

$$\overline{\lim}_{n \rightarrow \infty} \frac{d(x_{n+1}, \bar{x})}{d(x_n, \bar{x})} = 0.$$

Remark 3.4. To the best of our knowledge, Theorem 3.4 and Corollaries 3.5, 3.6, and 3.9 are new even in Euclidean spaces.

4. Finite termination of the algorithm. This section is devoted to establishing results on finite termination of the inexact proximal point algorithm. We say that a sequence $\{x_n\}$ terminates in a finite number of iterations if there exists a positive integer N such that $x_n = x_{n+1}$ and $x_n \in A^{-1}(0)$ for all $n \geq N$. We begin with the following useful lemma. Let S be a convex subset in M , and let $x \in S$. Recall that a vector $v \in T_x M$ is tangent to S if there is a smooth curve $\gamma : [0, \varepsilon) \rightarrow S$ such that $\gamma(0) = x$ and $\gamma'(0) = v$. Then the collection $T_x S$ of all tangent vectors to S at x is a convex cone in the space $T_x M$; see [54, p. 71]. Thus the normal cone $N_S(\cdot)$ is defined by

$$(4.1) \quad N_S(x) = \begin{cases} \{w \in T_x M \mid \langle w, v \rangle \leq 0 \text{ for all } v \in T_x S\} & \text{for each } x \in S, \\ \emptyset & \text{otherwise.} \end{cases}$$

Below, we always assume that $S = A^{-1}(0)$.

DEFINITION 4.1. Let $\bar{x} \in S$. Let $r, \alpha > 0$. Problem (3.1) is said to satisfy

(a) *weak sharp minima-like condition on $\mathbf{B}(\bar{x}, r)$ (with modulus α) if*

$$(4.2) \quad \overline{\mathbf{B}(0, \alpha)} \subset A(z) + T_z S \quad \text{for each } z \in \mathbf{B}(\bar{x}, r) \cap S;$$

(b) *local weak sharp minima-like condition if, for any $\bar{z} \in A^{-1}(0)$, there exists $r_{\bar{z}} > 0$ such that problem (3.1) satisfies weak sharp minima-like condition on $\mathbf{B}(\bar{z}, r_{\bar{z}})$;*

(c) *global weak sharp minima-like condition (with modulus α) if problem (3.1) satisfies weak sharp minima-like condition on M , that is,*

$$\overline{\mathbf{B}(0, \alpha)} \subset A(z) + T_z S \quad \text{for each } z \in S.$$

Note by [57, Proposition 3.1] that the following equivalence holds for each $z \in S$:

$$(4.3) \quad \overline{\mathbf{B}(0, \alpha)} \subset A(z) + T_z S \iff \overline{\mathbf{B}(0, \alpha)} \cap N_S(z) \subset A(z).$$

The following remark shows that the weak sharp minima-like condition implies the metrical subregularity.

Remark 4.1. Let $S = A^{-1}(0)$. If problem (3.1) satisfies a weak sharp minima-like condition on $\mathbf{B}(\bar{x}, r)$, then A is metrically subregular on $\mathbf{B}(\bar{x}, r_{\bar{x}})$ with $r_{\bar{x}} < \min\{1, r\}$, that is,

$$(4.4) \quad d(x, A^{-1}(0)) \leq \frac{1}{\alpha} d(0, A(x)) \quad \text{for all } x \in \mathbf{B}(\bar{x}, r_{\bar{x}}).$$

In fact, let $x \in \mathbf{B}(\bar{x}, r_{\bar{x}})$. Without loss of generality, we assume that $d(x, S) > 0$. Set $z = P_S(x)$. Then $\exp_z^{-1} x \in N_S(z)$ and $\|\exp_z^{-1} x\| = d(x, z) < 1$ which implies that

$$\alpha \exp_z^{-1} x \in \overline{\mathbf{B}(0, \alpha)} \cap N_S(z).$$

Then, by (4.3), we have that $\alpha \exp_z^{-1} x \in A(z)$. Since A is monotone, it follows that

$$(4.5) \quad \langle w, -\exp_x^{-1} z \rangle \geq \langle \alpha \exp_z^{-1} x, \exp_z^{-1} x \rangle \quad \text{for any } w \in A(x).$$

Hence, $\alpha \|\exp_z^{-1} x\| \leq d(0, A(x))$, and (4.4) is shown.

LEMMA 4.2. Let $\bar{x} \in S$ and let $\alpha > 0$. Let $r \in (0, +\infty]$. Suppose that problem (3.1) satisfies a weak sharp minima-like condition on $\mathbf{B}(\bar{x}, r)$ (with modulus α). Let $x, y \in \mathbf{B}(\bar{x}, r)$. Let $\lambda > 0$ and $u \in T_y M$ be such that

$$(4.6) \quad \lambda \exp_y^{-1} x + u \in A(y) \quad \text{and} \quad \|\lambda \exp_y^{-1} x + u\| < \alpha.$$

Then $y \in S$.

Proof. Assume on the contrary that $y \notin S$. Let $\bar{y} = P_S(y)$ be the projection of y on S . Then, by [34, Proposition 3.4], one has that

$$(4.7) \quad \langle v, \exp_{\bar{y}}^{-1} y \rangle \leq 0 \quad \text{for any } v \in T_{\bar{y}}S.$$

Furthermore, by (2.1), one has

$$d^2(y, \bar{y}) + d^2(\bar{y}, \bar{x}) - 2 \langle \exp_{\bar{y}}^{-1} y, \exp_{\bar{y}}^{-1} \bar{x} \rangle \leq d^2(y, \bar{x}).$$

This, together with (4.7), gives that $d(\bar{y}, \bar{x}) \leq d(y, \bar{x}) < r$. Thus, $\bar{y} \in \mathbf{B}(\bar{x}, r) \cap S$. Write $w := \alpha \frac{\exp_{\bar{y}}^{-1} y}{d(\bar{y}, y)}$. Then, w is well-defined (as $y \notin S$ by assumption), and $\|w\| = \alpha$. This, together with (4.2), implies that $w \in A(\bar{y}) + T_{\bar{y}}S$. Hence there exists $v \in T_{\bar{y}}S$ such that $w - v \in A(\bar{y})$. Since $\lambda \exp_y^{-1} x + u \in A(y)$ by (4.6), it follows from the monotonicity of A that

$$(4.8) \quad \langle \bar{w} - v, \exp_{\bar{y}}^{-1} y \rangle \leq \langle -\lambda \exp_y^{-1} x - u, \exp_y^{-1} \bar{y} \rangle.$$

Using (4.7), we see that $\langle w, \exp_{\bar{y}}^{-1} y \rangle \leq \langle \bar{w} - v, \exp_{\bar{y}}^{-1} y \rangle$. This, together with (4.8), implies that

$$\langle w, \exp_{\bar{y}}^{-1} y \rangle \leq \langle -\lambda \exp_y^{-1} x - u, \exp_y^{-1} \bar{y} \rangle.$$

Hence, we have $\|\lambda \exp_y^{-1} x + u\| \geq \|w\| = \alpha$, which contradicts (4.6), and so, we have that $y \in S$. \square

Recall that $A \in \mathcal{X}(M)$ is monotone and upper Kuratowski semicontinuous with $A^{-1}(0) \neq \emptyset$, and $x_0 \in \mathcal{D}(A)$ is a given initial point. Throughout the remainder of this section, we always assume that $S := A^{-1}(0)$ is convex.

Remark 4.2. Noting that A is maximal monotone, we see that S is closed. If $\mathcal{D}(A) = M$, then $S = A^{-1}(0)$ is convex (see [35, Corollary 4.7]).

LEMMA 4.3. *Let $\bar{x} \in S$ and let $\alpha > 0$. Let $r \in (0, +\infty]$. Suppose that problem (3.1) satisfies a weak sharp minima-like condition on $\mathbf{B}(\bar{x}, r)$ with modulus α . Let $\{x_n\}$ be a sequence generated by Algorithm IP₁ (together with the associated sequence $\{y_n\}$) or by Algorithm IP₂ such that*

$$(4.9) \quad \overline{\lim}_{n \rightarrow \infty} \delta_n < 1 \quad \text{and} \quad \underline{\lim}_{n \rightarrow \infty} \lambda_n d(x_n, x_{n+1}) = 0.$$

Suppose further that there exists $N \in \mathbb{N}$ such that $\{x_n, y_n, n \geq N\} \subseteq \mathbf{B}(\bar{x}, r)$. If $A(x)$ is a singleton for each $x \in S$, or $\{x_n\}$ is generated by Algorithm IP₂, then, $\{x_n\}$ terminates in a finite number of iterations.

Proof. First, let $\{x_n\}$ be generated by Algorithm IP₁ with the associated sequence $\{y_n\}$. Then

$$d(x_{n+1}, y_n) \leq \varepsilon_n = \delta_n d(x_{n+1}, x_n).$$

Thus, it follows that

$$\begin{aligned} (1 - \delta_n)d(x_{n+1}, x_n) &\leq d(x_{n+1}, x_n) - d(x_{n+1}, y_n) \\ &\leq d(x_n, y_n) \\ &\leq d(x_{n+1}, x_n) + d(x_{n+1}, y_n) \\ &\leq (1 + \delta_n)d(x_{n+1}, x_n). \end{aligned}$$

This, together with (4.9) implies that

$$\liminf_{n \rightarrow \infty} \lambda_n d(x_n, y_n) = 0.$$

Then, there exists a subsequence $\{n_k\}$ of $\{n\}$ such that

$$\lim_{k \rightarrow \infty} \lambda_{n_k} d(x_{n_k}, y_{n_k}) = 0.$$

Thus, there is a positive integer $K \geq N$ such that, for all $k \geq K$,

$$(4.10) \quad x_{n_k}, y_{n_k} \in \mathbf{B}(\bar{x}, r) \quad \text{and} \quad \lambda_{n_k} \left\| \exp_{y_{n_k}}^{-1} x_{n_k} \right\| < \alpha.$$

Noting

$$(4.11) \quad \lambda_{n_k} \exp_{y_{n_k}}^{-1} x_{n_k} \in A(y_{n_k})$$

for each k , Lemma 4.2 is applicable (with $u = 0$) to concluding that $y_{n_k} \in S$ for all $k \geq K$. Thus, there exists $\tilde{N} \geq n_K$ such that $y_{\tilde{N}} \in S$. By assumption, $A(y_{\tilde{N}})$ is a singleton. Therefore, thanks to (4.11), one sees that $0 \in A(y_{\tilde{N}}) = \lambda_{\tilde{N}} \exp_{y_{\tilde{N}}}^{-1} x_{\tilde{N}}$, that is, $x_{\tilde{N}} = y_{\tilde{N}}$ (as $\lambda_{\tilde{N}} > 0$). Thus, $x_{\tilde{N}} \in S$. Below, we show that $x_{n+1} = x_n$ for all $n \geq \tilde{N}$ and then complete the proof. To proceed, we fix $n \geq \tilde{N}$. Without loss of generality, assume that $x_n \in S = A^{-1}(0)$ and $\delta_n < 1$. Below, we prove that $x_{n+1} = x_n$. Note by Algorithm IP₁ that there exists y_n such that $0 \in A(y_n) - \lambda_n \exp_{y_n}^{-1} x_n$. By assumption, $0 \in A(x_n)$ and so $0 \in A(x_n) - \lambda_n \exp_{x_n}^{-1} x_n$. As $A(\cdot) - \lambda_n \exp_{x_n}^{-1} x_n$ is strongly monotone as noted in Example 2.1, it follows that $x_n = y_n$. Thus, we have

$$d(x_{n+1}, x_n) = d(x_{n+1}, y_n) \leq \varepsilon_n.$$

Then, it follows that

$$d(x_{n+1}, x_n) \leq \frac{\varepsilon_n}{d(x_{n+1}, x_n)} d(x_{n+1}, x_n) = \delta_n d(x_{n+1}, x_n).$$

This, together with the assumption that $\delta_n < 1$, implies that $x_n = x_{n+1}$.

Second, let $\{x_n\}$ be generated by Algorithm IP₂. Note by assumption that

$$\liminf_{n \rightarrow \infty} \lambda_n d(x_n, x_{n+1}) = \liminf_{n \rightarrow \infty} \lambda_n \left\| \exp_{x_{n+1}}^{-1} x_n \right\| = 0.$$

Fix n . Then making use of the definition of Algorithm IP₂, we can choose $v_{n+1} \in T_{x_{n+1}} M$ such that

$$(4.12) \quad v_{n+1} \in A_{\lambda_n, x_n}(x_{n+1}) = A(x_{n+1}) - \lambda_n \exp_{x_{n+1}}^{-1} x_n$$

and

$$(4.13) \quad \|v_{n+1}\| = d(0, A_{\lambda_n, x_n}(x_{n+1})) \leq \lambda_n \varepsilon_n = \lambda_n \delta_n d(x_n, x_{n+1}),$$

where v_{n+1} is well-defined because $A_{\lambda_n, x_n}(x_{n+1})$ is closed and convex. (4.13) implies that

$$\left\| \lambda_n \exp_{x_{n+1}}^{-1} x_n + v_{n+1} \right\| \leq \lambda_n d(x_n, x_{n+1})(1 + \delta_n).$$

Combining this with (4.9) yields that

$$\liminf_{n \rightarrow \infty} \left(\lambda_n \exp_{x_{n+1}}^{-1} x_n + v_{n+1} \right) = 0.$$

Hence, there exists a positive integer $K \geq N$ such that the following assertions hold:

$$x_K, x_{K+1} \in \mathbf{B}(\bar{x}, r), \quad \lambda_K \exp_{x_{K+1}}^{-1} x_K + v_{K+1} \in A(x_{K+1}),$$

and

$$\left\| \lambda_K \exp_{x_{K+1}}^{-1} x_K + v_{K+1} \right\| < \alpha.$$

Then, Lemma 4.2 is applicable to concluding that $x_{K+1} \in S$. Below we show that $x_{n+1} = x_n$ for each $n \geq K+1$, and then complete the proof. By assumption, without loss of generality, we may assume that for each $n \geq K$,

$$(4.14) \quad \frac{\varepsilon_n}{d(x_n, x_{n+1})} < 1.$$

Let $n = K+1$ and let v_{n+1} be chosen to satisfy (4.12) and (4.13). Since $0 \in A(x_n)$ because $x_n \in S$, and A is monotone, it follows that

$$\left\langle v_{n+1} + \lambda_n \exp_{x_{n+1}}^{-1} x_n, \exp_{x_{n+1}}^{-1} x_n \right\rangle \leq \left\langle 0, -\exp_{x_n}^{-1} x_{n+1} \right\rangle = 0.$$

That is,

$$\left\langle \lambda_n \exp_{x_{n+1}}^{-1} x_n, \exp_{x_{n+1}}^{-1} x_n \right\rangle \leq \left\langle -v_{n+1}, \exp_{x_{n+1}}^{-1} x_n \right\rangle,$$

and so $\lambda_n d(x_n, x_{n+1}) = \lambda_n \|\exp_{x_{n+1}}^{-1} x_n\| \leq \|v_{n+1}\|$. Combining this with (4.13) yields that

$$d(x_{n+1}, x_n) \leq \varepsilon_{n+1}.$$

Furthermore, using (4.14), we obtain that

$$d(x_{n+1}, x_n) \leq \frac{\varepsilon_{n+1}}{d(x_{n+1}, x_n)} d(x_{n+1}, x_n) < d(x_{n+1}, x_n).$$

Hence $d(x_{n+1}, x_n) = 0$, that is $x_{n+1} = x_n$ as desired. Inductively, we get $x_{n+1} = x_n$ for each $n \geq K+1$ and the proof is complete. \square

THEOREM 4.4. *Suppose that problem (3.1) satisfies*

(a) *global weak sharp minima-like condition with modulus $\alpha > 0$*

or

(b) *local weak sharp minima-like condition and $\sum_n \varepsilon_n < +\infty$.*

Let $\{x_n\}$ be a sequence generated by Algorithm IP₁ such that (4.9) holds. If $A(x)$ is a singleton for each $x \in S$, then $\{x_n\}$ terminates in a finite number of iterations.

Proof. (a) If problem (3.1) satisfies a global weak sharp minima-like condition with modulus $\alpha > 0$, then for any $\bar{x} \in S$, problem (3.1) satisfies a weak sharp minima-like condition on $\mathbf{B}(\bar{x}, r)$ (for any $r > 0$) with uniform modulus α . Then, the conclusion follows from Lemma 4.3.

(b) Suppose that problem (3.1) satisfies a weak sharp minima-like condition and $\sum_n \varepsilon_n < +\infty$. Let $\{x_n\}$ and $\{y_n\}$ be sequences generated by Algorithm IP₁. Then, it follows from [59, Proposition 3.3] that $\{x_n\}$ converges to a singular point $\bar{x} \in A^{-1}(0)$. Hence, there exist $\alpha_{\bar{x}}, r_{\bar{x}} > 0$ such that problem (3.1) satisfies a weak sharp minima-like condition on $\mathbf{B}(\bar{x}, r_{\bar{x}})$ with modulus $\alpha_{\bar{x}}$. Thus, there exists a positive integer N such that $\{x_n, y_n, n \geq N\} \subset \mathbf{B}(\bar{x}, r_{\bar{x}})$. Therefore, Lemma 4.3 is applicable and the conclusion follows. \square

COROLLARY 4.5. *Suppose problem (3.1) satisfies a local weak sharp minima-like condition and suppose that*

$$(4.15) \quad \sum_n \frac{1}{\lambda_n^2} = +\infty.$$

Let $\{x_n\}$ be a sequence generated by Algorithm IP₁ with $\sum_n \delta_n < +\infty$. If $A(x)$ is a singleton for each $x \in S$, then $\{x_n\}$ terminates in a finite number of iterations.

Proof. By Theorem 4.4, to complete the proof, it's sufficient to show that

$$(4.16) \quad \sum_n \varepsilon_n < +\infty$$

and (4.9) hold. In fact, $\overline{\lim}_{n \rightarrow \infty} \delta_n < 1$ and (4.16) follows directly from the assumption $\sum_n \delta_n < +\infty$ and Lemma 2.4. Thus, we get from [59, Theorem 3.8] that (4.16) and (4.15) imply that the equality in (4.9) holds. \square

Theorem 4.6 below shows that if $\{x_n\}$ is a sequence generated by Algorithm IP₂, then the condition that $A(x)$ is a singleton for each $x \in S$ can be removed.

THEOREM 4.6. *Suppose that problem (3.1) satisfies condition (a) or (b) of Theorem 4.4. Let $\{x_n\}$ be a sequence generated by Algorithm IP₂ such that (4.9) holds. Then $\{x_n\}$ terminates in a finite number of iterations.*

Proof. Let $\{x_n\}$ be a sequence generated by Algorithm IP₂. Since $\lambda_n \varepsilon_n = \lambda_n \delta_n d(x_n, x_{n+1})$, it follows from (4.9) that

$$\liminf_{n \rightarrow \infty} \lambda_n (d(x_n, x_{n+1}) + \varepsilon_n) = 0.$$

Hence, [59, Proposition 3.3] is applicable to conclude that $\{x_n\}$ converges to a singular point $\bar{x} \in A^{-1}(0)$. Thus, with a similar technique mentioned in the proof of Theorem 4.4, we get that Lemma 4.3 is applicable and the conclusion follows. \square

COROLLARY 4.7. *Let $\{x_n\}$ be a sequence generated by Algorithm IP₂ with $\sum_n \delta_n^2 < +\infty$. Suppose problem (3.1) satisfies a local weak sharp minima-like condition and (4.15) holds. Then $\{x_n\}$ terminates in a finite number of iterations.*

Proof. By Theorem (4.6), to complete the proof, it's sufficient to show that (4.16) and (4.9) hold. Indeed, $\sum_n \delta_n^2 < +\infty$ implies that $\overline{\lim}_{n \rightarrow \infty} \delta_n < 1$ and (4.16) holds because of Lemma 2.5. Thus, we apply [59, Lemma 3.10] to conclude that

$$\underline{\lim}_{n \rightarrow \infty} \lambda_n \left\| \exp_{x_{n+1}}^{-1} x_n \right\| = 0.$$

Hence, (4.9) is seen to hold. \square

COROLLARY 4.8. *Let $\{x_n\}$ be a sequence generated by Algorithm IP₂ with $\sum_n \delta_n^2 < +\infty$. Suppose that (4.15) holds, and that there exists $\bar{x} \in M$ such that $0 \in \text{int}A(\bar{x})$. Then, $\{x_n\}$ terminates in a finite number of iterations.*

Proof. By Proposition 3.8, the assumption $0 \in \text{int}A(\bar{x})$ implies that $S = A^{-1}(0) = \{\bar{x}\}$. This implies that $T_{\bar{x}}S = 0$ and there exists $\alpha > 0$ such that (4.2) holds (with any $r > 0$). Hence, problem (3.1) satisfies a global weak sharp minima-like condition. Thus, Corollary 4.7 is applicable and the conclusion follows. \square

Applying Corollary 4.7 to the special case when $\varepsilon_n = 0$ for each $n \geq 0$, we get the following finite termination result for Algorithm P.

COROLLARY 4.9. *Let $\{x_n\}$ be a sequence generated by Algorithm P. Suppose that problem (3.1) satisfies a local weak sharp minima-like condition and (4.15) holds. Then, Algorithm P terminates in a finite number of iterations.*

The following corollary follows from Corollary 4.8 directly.

COROLLARY 4.10. *Let $\{x_n\}$ be a sequence generated by Algorithm P. Suppose that (4.15) holds, and there exists $\bar{x} \in M$ such that $0 \in \text{int}A(\bar{x})$. Then, Algorithm P terminates in a finite number of iterations.*

5. Applications. This section is devoted to applications of the results in the previous sections to minimization problems on Hadamard manifolds. Let $f : M \rightarrow (-\infty, +\infty]$ be a function with its effective domain denoted by $\mathcal{D}(f)$ and defined by

$$\mathcal{D}(f) := \{x \in M : f(x) \neq +\infty\}.$$

Recall that f is proper if its effective domain $\mathcal{D}(f)$ is nonempty. Recall also that f is convex if for any geodesic γ in M , the composition function $f \circ \gamma : \mathbb{R} \rightarrow (-\infty, +\infty]$ is convex, that is,

$$(f \circ \gamma)(ta + (1-t)b) \leq t(f \circ \gamma)(a) + (1-t)(f \circ \gamma)(b)$$

for any $a, b \in \mathbb{R}$ and $0 \leq t \leq 1$. Throughout this subsection, we assume that $f : M \rightarrow (-\infty, +\infty]$ is a proper lower semicontinuous convex function. The *subdifferential* $\partial f(x)$ of f at $x \in \mathcal{D}(f)$ is defined by

$$\partial f(x) := \{u \in T_x M : \langle u, \exp_x^{-1} y \rangle \leq f(y) - f(x) \text{ for each } y \in M\}.$$

The following proposition collects some useful properties of ∂f proved in [33].

PROPOSITION 5.1. *Let f be a proper lower semicontinuous convex function on M . Then the following assertions hold:*

- (i) *The subdifferential ∂f is a monotone and upper Kuratowski semicontinuous multivalued vector field.*
- (ii) *If $\mathcal{D}(f) = M$, then ∂f is maximal monotone.*

The constrained minimization problem considered here is described as follows:

$$(5.1) \quad \min_{x \in C} f(x),$$

where, $C \subseteq M$ is a closed convex subset. Let δ_C be the indicator function defined by $\delta_C(x) = 0$ if $x \in C$ and $\delta_C(x) = +\infty$ otherwise, and let $f_C := f + \delta_C$. Then the constrained minimization problem (5.1) is equivalent to the unconstrained minimization problem: $\min_{x \in M} f_C(x)$. Therefore, without loss of generality, we assume for the whole subsection that C is the manifold M . Thus, problem (5.1) is reduced to the following minimization problem:

$$(5.2) \quad \min_{x \in M} f(x).$$

We use S to denote the solution set of (5.2), that is,

$$S := \operatorname{argmin}\{f(x) : x \in M\}.$$

It is easy to check that

$$(5.3) \quad x \in S \iff 0 \in \partial f(x).$$

Throughout this section, we always assume that S is nonempty.

Let $\lambda > 0$, $z \in M$, and let $f_{\lambda,z}$ denote the regularized function defined by

$$f_{\lambda,z}(x) := f(x) + \frac{\lambda}{2}d^2(z, x) \quad \text{for each } x \in M.$$

The inexact proximal point algorithm that we are interested in for finding an approximate solution of problem (5.2), is defined as follows.

Algorithm IP_M. Give initial point $x_0 \in \mathcal{D}(f)$. Letting $n = 0, 1, \dots$ and having x_0, x_1, \dots, x_n , determine x_{n+1} as follows: choosing a parameter $\lambda_n > 0$ and an error $\varepsilon_n \geq 0$, then take x_{n+1} such that

$$f_{\lambda_n, x_n}(x_{n+1}) \leq \inf_{x \in M} f_{\lambda_n, x_n}(x) + \frac{1}{2}\lambda_n \varepsilon_n^2.$$

Clearly, in the special case when each $\varepsilon_n = 0$, Algorithm IP_M is reduced to the following classical proximal point algorithm which has been studied in [33] and [21].

Algorithm P_M. Give initial point $x_0 \in \mathcal{D}(f)$. Letting $n = 0, 1, \dots$ and having x_0, x_1, \dots, x_n , determine x_{n+1} as follows: choosing a parameter $\lambda_n > 0$, take x_{n+1} such that

$$x_{n+1} = \operatorname{argmin}_{x \in M} f_{\lambda_n, x_n}(x).$$

The following proposition describes the relationship between Algorithms IP_M and IP₁, which is taken from [59, Proposition 4.3].

PROPOSITION 5.2. *Let $\{x_n\}$ be a sequence generated by Algorithm IP_M with initial point x_0 . Then $\{x_n\}$ also is a sequence generated by Algorithm IP₁ for $A(\cdot) = \partial f(\cdot)$ with the same error sequence $\{\varepsilon_n\}$.*

5.1. Linear convergence. This subsection is devoted to the study of linear convergence of Algorithm IP_M. Below, we extend the notion of weak sharp minimizer of order p for problem (5.2) to a manifold setting. Here, we are interested in the case when $p = 2$.

DEFINITION 5.3. *Let S be the solution set of problem (5.2). Then, we say that*

- (a) *S is the set of local weak sharp minima of order 2 for problem (5.2) if for each $\bar{x} \in S$, there exist $\alpha, r > 0$ such that for all $x \in \mathbf{B}(\bar{x}, r)$,*

$$(5.4) \quad f(x) \geq f(\bar{x}) + \alpha d^2(x, S);$$

- (b) *S is the set of global weak sharp minima of order 2 for problem (5.2) with the uniform modulus $\alpha > 0$ if estimate (5.4) holds for all $\bar{x} \in S$ and $x \in M$.*

The following proposition is about metrical subregularity of ∂f .

PROPOSITION 5.4. *If S is the set of local (resp., global) weak sharp minima of order 2 for problem (5.2), then ∂f is metrical subregular (resp., global metrical subregular).*

Proof. Let $\bar{x} \in \partial f^{-1}(0)$. Then $\bar{x} \in S$. Thus, by assumption, there exist $\alpha, r > 0$ such that (5.4) holds for all $x \in \mathbf{B}(\bar{x}, r)$. Let $x \in \mathbf{B}(\bar{x}, \varepsilon)$. As $S = \partial f^{-1}(0)$, to complete the proof, it's sufficient to show that

$$(5.5) \quad d(x, S) \leq \frac{1}{\alpha}d(0, \partial f(x)).$$

In the case when $d(x, S) = 0$, it's trivial. Below, we assume that $d(x, S) > 0$. Since S is closed and convex, there exists $x_0 \in S$ such that

$$(5.6) \quad d(x, S) = d(x, x_0).$$

Let $v \in \partial f(x)$. Then

$$\langle v, -\exp_x^{-1} x_0 \rangle \geq f(x) - f(x_0).$$

This, together with (5.6) and (5.4), gives that

$$\|v\| \geq \frac{\langle v, -\exp_x^{-1} x_0 \rangle}{\|\exp_x^{-1} x_0\|} \geq \frac{f(x) - f(x_0)}{\|\exp_x^{-1} x_0\|} \geq \frac{\alpha d^2(x, S)}{d(x, S)} = \alpha d(x, S).$$

As $v \in \partial f(x)$ is arbitrary, (5.5) is seen to hold. \square

The linear convergence result of Algorithm IP_M is given below.

COROLLARY 5.5. *Let $\{x_n\}$ be a sequence generated by Algorithm IP_M . Suppose that (3.16) holds. Suppose further that*

- (a) S is the set of global weak sharp minima of order 2, or
- (b) S is the set of local weak sharp minima of order 2 and $\sum_n \varepsilon_n < +\infty$.

Then, $\{x_n\}$ is linearly convergent to a point $\bar{x} \in S$. Moreover, if $\lim_{n \rightarrow \infty} \lambda_n = 0$, then $\{x_n\}$ is superlinearly convergent.

Proof. Let $\{x_n\}$ be a sequence generated by Algorithm IP_M . Then, it follows from Proposition 5.2 that $\{x_n\}$ is also a sequence generated by Algorithm IP_1 for $A(\cdot) = \partial f(\cdot)$ with the same error sequence $\{\varepsilon_n\}$. Thus, the conclusion follows from Proposition 5.4 and Theorem 3.4. \square

For the classical proximal point Algorithm P_M , we get the following corollary about the linear convergence rate directly from Corollary 5.5.

COROLLARY 5.6. *Let $\{x_n\}$ be a sequence generated by Algorithm P_M . Suppose that $\sup_n \lambda_n < +\infty$. Suppose further that S is the set of local weak sharp minima of order 2. Then, $\{x_n\}$ is linearly convergent to a point $\bar{x} \in S$. Furthermore, if $\lim_{n \rightarrow \infty} \lambda_n = 0$, then $\{x_n\}$ is superlinearly convergent.*

5.2. Finite termination. This subsection is devoted to the study of finite termination of Algorithm IP_M . The notion of weak sharp minima has been explored thoroughly in linear spaces (cf. [12, 13] and so on), which has also been extended to a manifold setting (cf. [34]).

DEFINITION 5.7. *Let S be the solution set of problem (5.2). S is said to be the set of local weak sharp minima for problem (5.2) if for each $\bar{x} \in S$, there exist $\alpha, r > 0$ such that for all $x \in \mathbf{B}(\bar{x}, r)$,*

$$(5.7) \quad f(x) \geq f(\bar{x}) + \alpha d(x, S).$$

By [34, Corollary 4.10], S is the set of local weak sharp minima for problem (5.2) if and only if there are $\alpha, r > 0$ such that we have the inclusion

$$(5.8) \quad \overline{\mathbf{B}(0, \alpha)} \subset \partial f(z) + T_z S \quad \text{for all } z \in S \cap \mathbf{B}(\bar{x}, r).$$

Hence, we have the following proposition.

PROPOSITION 5.8. *S is the set of local weak sharp minima for problem (5.2) if and only if problem 0 $\in \partial f(x)$ satisfies a local weak sharp minima-like condition.*

The following theorem shows that under certain condition and assumption of local weak sharp minima, a sequence generated by Algorithm IP_M terminates in a finite number of iterations.

COROLLARY 5.9. *Suppose that S is the set of local weak sharp minima for problem (5.2), and that (4.15) holds. Let $\{x_n\}$ be a sequence generated by Algorithm IP_M. Then, $\{x_n\}$ terminates in a finite number of iterations provided one of the following two conditions holds:*

- (a) f is Gateaux differentiable on S_f and $\sum \delta_n < +\infty$;
- (b) $\sum \delta_n^2 < +\infty$ and

$$(5.9) \quad d(0, \partial f_{\lambda_n, x_n}(x_{n+1})) \leq \lambda_n \varepsilon_n \quad \text{for each } n \in \mathbb{N}.$$

Proof. By Proposition 5.8, a local weak sharp minima implies a local weak sharp minima-like condition. On the other hand, note that for each $x \in S_f$, f is Gateaux differentiable at x implies that $\partial f(x)$ is a singleton. Moreover, if assumption (5.9) holds, then $\{x_n\}$ is also a sequence generated by Algorithm IP₂ for $A(\cdot) = \partial f(\cdot)$ with the same error sequence $\{\varepsilon_n\}$. Hence, Corollary 4.5 is applicable and the conclusion follows. □

The following theorem shows that if $\{f(x_n)\}$ is bounded (e.g., $\mathcal{D}(f) = M$ or, more generally, any cluster point of $\{x_n\}$ is in $\text{int}_R(\mathcal{D}(f))$ (cf. [59, Remark 4.2])), then the condition $\sum_n \frac{1}{\lambda_n^2} = +\infty$ can be weakened to $\sum_n \frac{1}{\lambda_n} = +\infty$.

THEOREM 5.10. *Suppose that S is the set of local weak sharp minima for problem (5.2). Suppose further that*

$$(5.10) \quad \sum_n \frac{1}{\lambda_n} = +\infty.$$

Let $\{x_n\}$ be a sequence generated by Algorithm IP_M such that $\{f(x_n)\}$ is bounded. Then, $\{x_n\}$ terminates in a finite number of iterations provided one of the conditions (a) and (b) in Corollary 5.9 holds.

Proof. To complete the proof, we only need to show that

$$(5.11) \quad \liminf_{n \rightarrow \infty} \lambda_n d(x_n, x_{n+1}) = 0.$$

Granting this, we can check, as in the proof we did for Corollary 5.9, that Proposition 5.8 and Theorem 4.4 are applicable and then the conclusion follows.

By the definition of Algorithm IP_M, for each n ,

$$(5.12) \quad f(x_{n+1}) + \frac{\lambda_n}{2} d(x_n, x_{n+1})^2 \leq f(x_n) + \frac{1}{2} \lambda_n \varepsilon_n^2.$$

Recall that $\delta_n = \frac{\varepsilon_n}{d(x_n, x_{n+1})}$. Then, $\varepsilon_n = \delta_n d(x_n, x_{n+1})$. This, together with (5.12), yields that

$$(5.13) \quad \frac{\lambda_n}{2} (1 - \delta_n^2) d(x_n, x_{n+1})^2 \leq f(x_n) - f(x_{n+1}).$$

Since $p := \limsup_{n \rightarrow \infty} \delta_n < 1$, there exists a positive integer N such that for all $n \geq N$, $\delta_n < p$. Combing this with (5.13) yields that

$$(5.14) \quad \frac{1}{2} (1 - p^2) \sum_{n \geq N} \frac{1}{\lambda_n} (\lambda_n d(x_n, x_{n+1}))^2 \leq \sum_{n \geq N} (f(x_n) - f(x_{n+1})).$$

Since $\{f(x_n)\}$ is bounded, it follows from (5.14) that

$$(5.15) \quad \sum_{n \geq N} \frac{1}{\lambda_n} (\lambda_n d(x_n, x_{n+1}))^2 < +\infty.$$

As $\sum_{n \geq 0} \frac{1}{\lambda_n} = +\infty$, one sees that (5.11) follows from (5.15), and the proof is complete. \square

Corollary 5.11 below follows directly from Theorem 5.10, and shows that the classical proximal point algorithm P_M for problem (5.2) terminates in a finite number of iterations, which extends the corresponding results in [11], where it was proved under the assumptions that S_f is a global weak sharp minima for problem (5.2) and that $\{\lambda_n\} \subseteq [\lambda_-, \lambda_+]$ for some $0 < \lambda_- < \lambda_+$.

COROLLARY 5.11. *Suppose that S is the set of local weak sharp minima for problem (5.2). Suppose further that $\mathcal{D}(f) = M$ and (5.10) holds. Then, Algorithm P_M terminates in a finite number of iterations.*

6. Numerical examples. This section is devoted to several examples: Examples 6.1–6.3 are used to illustrate our theorems in the previous sections for inexact proximal point algorithms on Hadamard manifolds; Example 6.4 is used to show that our results obtained in the present paper cannot be extended directly to the general Riemannian manifold setting. It should be remarked that the functions or vector fields in Examples 6.1–6.3 are not convex or monotone in the Euclidean space and so the corresponding results in a Hilbert space setting (e.g., [51, 57]) cannot apply.

Let $M = H := \{(y_1, y_2) \in \mathbb{R}^2 \mid y_2 > 0\}$ be the Poincaré plane endowed with the Riemannian metric given by

$$g_{11} = g_{22} = \frac{1}{y_2^2}, \quad g_{12} = 0 \quad \text{for each point } (y_1, y_2) \in H$$

(cf. [54, p. 86]). Then, the sectional curvature of M is -1 (cf. [18]). Let $\bar{x} = (y_1, y_2) \in H$. Then $T_{\bar{x}}H = \mathbb{R}^2$ and

$$(6.1) \quad \langle u, v \rangle_{\bar{x}} = \frac{1}{y_2^2} \langle u, v \rangle \quad \text{for any pair } (u, v) \in T_{\bar{x}}H \times T_{\bar{x}}H,$$

where $\langle \cdot, \cdot \rangle_{\bar{x}}$ and $\langle \cdot, \cdot \rangle$ denote the inner product in $T_{\bar{x}}H$ and \mathbb{R}^2 , respectively. The geodesics of the Poincaré plane are the semilines $\gamma_a : y_1 = a, y_2 > 0$, and the semicircles $\gamma_{b,r} : (y_1 - b)^2 + y_2^2 = r^2, y_2 > 0$. Consider the set C given by

$$C := \{(y_1, y_2) \in \mathbb{R}^2 \mid (y_1 - 1)^2 + y_2^2 \leq 5, (y_1 + 1)^2 + y_2^2 \leq 5, y_2 \geq 1\}.$$

Then $C \subseteq H$ is convex. Example 6.1 gives an example of a multivalued vector field on the Poincaré plane satisfying a local weak sharp minima-like condition.

Example 6.1. Let $f_0 : M \rightarrow \mathbb{R}$ be the function defined by

$$f_0((y_1, y_2)) := \frac{1}{y_2} \quad \text{for each } (y_1, y_2) \in M.$$

By [54, p. 86], f_0 is convex on M . Define the function $f_1 : M \rightarrow \mathbb{R}$ by

$$f_1((y_1, y_2)) := \max \left\{ f_0(y_1, y_2), \frac{2}{3} \right\} \quad \text{for each } (y_1, y_2) \in M.$$

Then, f_1 is convex on M . Moreover, define the function $f : M \rightarrow \mathbb{R}$ by

$$(6.2) \quad f((y_1, y_2)) := f_1(y_1, y_2) + \delta_C(y_1, y_2) \quad \text{for each } (y_1, y_2) \in M.$$

Then, f is convex and the solution set of (5.2) is

$$(6.3) \quad S = C \cap \left\{ (y_1, y_2) \in M \mid y_2 \geq \frac{3}{2} \right\}.$$

Furthermore, by [34, Example 6.2], S is the set of local weak sharp minima for problem (5.2). Hence, Corollary 5.9 is applicable to concluding that any sequence $\{x_n\}$, generated by Algorithm IP_M with $\{\delta_n\}$ satisfying $\sum \delta_n < +\infty$ and the parameters $\{\lambda_n\}$ satisfying (4.15), terminates in a finite number of iterations.

Let $A : M \rightrightarrows TM$ be a multivalued vector field defined by

$$A(y_1, y_2) := \begin{cases} (0, -1) + N_C(y_1, y_2), & (y_1, y_2) \in C \setminus S, \\ \{t(0, -1) : t \in [0, 1]\} + N_C(y_1, y_2), & (y_1, y_2) \in S, \\ \emptyset, & \text{otherwise.} \end{cases}$$

By [54, p. 297], one checks that A is equal to the subdifferential of f given by (6.2). Since S is the set of local weak sharp minima for problem (5.2), it follows from Proposition 5.8 that the problem $0 \in A(x)$ satisfies a local weak sharp minima-like condition. Thus, Corollary 4.7 is applicable to concluding that any sequence $\{x_n\}$, generated by Algorithm IP_2 with $\{\delta_n\}$ satisfying $\sum \delta_n^2 < +\infty$ and the parameters $\{\lambda_n\}$ satisfying (4.15), terminates in a finite number of iterations.

Example 6.2 below provides a multivalued vector field on the Poincaré plane satisfying the metrical subregularity condition.

Example 6.2. Let $W, S \subseteq H$ be defined, respectively, by

$$W := \{(y_1, y_2) \in H : y_1 \in [1, 3], 2 \leq (y_1 - 2)^2 + y_2^2 \leq 10\}$$

and

$$S := \{(y_1, y_2) \in W : y_1^2 + (y_2 - 2)^2 \leq 4\} = \left\{ (y_1, y_2) \in W : \frac{y_1^2 + y_2^2}{y_2} \leq 4 \right\}.$$

Then, W and S are closed and convex as, by [54, p. 301], the function $(y_1, y_2) \rightarrow \frac{y_1^2 + y_2^2}{y_2}$ is convex on H . Let $A_1, A : M \rightrightarrows TM$ be multivalued vector fields defined, respectively, by

$$A_1(y_1, y_2) := \begin{cases} (2y_1y_2, y_2^2 - y_1^2), & y_1^2 + (y_2 - 2)^2 > 4, \\ \{t(2y_1y_2, y_2^2 - y_1^2) : t \in [0, 1]\}, & y_1^2 + (y_2 - 2)^2 = 4, \\ (0, 0), & y_1^2 + (y_2 - 2)^2 < 4, \end{cases}$$

and

$$A(y_1, y_2) := A_1(y_1, y_2) - \exp_{(y_1, y_2)}^{-1} P_S(y_1, y_2) + N_W(y_1, y_2),$$

where P_S denotes the projection on S . Clearly, $N_W(\cdot)$ is monotone. Furthermore, by [46], the vector field $(y_1, y_2) \rightarrow -\exp_{(y_1, y_2)}^{-1} P_S(y_1, y_2)$ is monotone and continuous on H . Below we show that A_1 also is monotone. To do this, define a function $f : H \rightarrow \mathbb{R}$ by

$$f(y_1, y_2) := \max \left\{ \frac{y_1^2 + y_2^2}{y_2}, 4 \right\} \quad \text{for each } (y_1, y_2) \in H.$$

Then, as mentioned above, f is convex. Furthermore, by [54, p. 297], one checks that A_1 is equal to the subdifferential of f . Hence, by [33, Theorem 5.1], A_1 is monotone

and upper Kuratowski semicontinuous on H . Therefore, A is monotone. Observe further that $S \subseteq A^{-1}(0)$. Below, we verify that

$$(6.4) \quad S = A^{-1}(0) \quad \text{and} \quad d(x, S) \leq d(0, A(x)) \quad \text{for all } x \in W,$$

that is, A is metrically subregular on W with modulus $\kappa = 1$. Clearly, the first assertion in (6.4) follows trivially from the second one of (6.4). Hence, it remains to show that the second assertion in (6.4) holds. To proceed, let $x \in W$. In the case when $d(x, S) = 0$, it's trivial. Below, we assume that $d(x, S) > 0$. Since S is closed and convex, there exists $x_0 \in S$ such that $d(x, S) = d(x, x_0)$, that is, $x_0 = P_S(x)$. Let $v \in A(x)$. Then, there are $v_1 \in A_1(x), v_2 \in N_W(x)$ such that

$$v = v_1 - \exp_x^{-1} P_S(x) + v_2 = v_1 - \exp_x^{-1} x_0 + v_2.$$

Clearly, by the definition of a normal cone $N_W(x)$, one has

$$(6.5) \quad \langle v_2, -\exp_x^{-1} x_0 \rangle \geq 0.$$

Note that f is convex and so

$$(6.6) \quad \langle v_1, -\exp_x^{-1} x_0 \rangle \geq f(x) - f(x_0) \geq 0.$$

Observe further that

$$(6.7) \quad \langle -\exp_x^{-1} x_0, -\exp_x^{-1} x_0 \rangle = d^2(x, S).$$

Hence, it follows from (6.7), (6.6), and (6.5) that

$$\langle v, -\exp_x^{-1} x_0 \rangle \geq d^2(x, S).$$

This implies that

$$\|v\| \geq \frac{\langle v, -\exp_x^{-1} x_0 \rangle}{\|-\exp_x^{-1} x_0\|} \geq \frac{d^2(x, S)}{d(x, S)} = d(x, S).$$

As $v \in A(x)$ is arbitrary, (6.4) is seen to hold. Thus, Corollary 3.5 is applicable to concluding that, with $\{\lambda_n\}$ satisfying $\sup_n \lambda_n < +\infty$, any sequence $\{x_n\}$, generated by IP₁ with $\sum_n \delta_n < \infty$, or by Algorithm IP₂ with $\sum_n \delta_n^2 < \infty$, converges linearly to a point $\bar{x} \in A^{-1}(0)$. Furthermore, if $\lim_{n \rightarrow \infty} \lambda_n = 0$, then $\{x_n\}$ is superlinearly convergent.

In the following example, vector fields with global metrical subregularity or a global weak sharp minima-like condition on the Hadamard manifold of symmetric positive definite matrixes are presented.

Example 6.3. Let

$$\mathbb{S}_{++}^m := \{X \in \mathbb{R}^{m \times m} \mid X \text{ is a symmetric positive definite matrixe}\}$$

be the differentiable manifold with the Riemannian distance between $X, Y \in \mathbb{S}_{++}^m$ given by

$$d^2(X, Y) := \sum_{i=1}^m \ln^2 \lambda_i \left(X^{-\frac{1}{2}} Y X^{-\frac{1}{2}} \right),$$

where $\lambda_i(X^{-\frac{1}{2}}YX^{-\frac{1}{2}})$ is the i th eigenvalue of $X^{-\frac{1}{2}}YX^{-\frac{1}{2}}$ (cf. [47, p. 355], [43]). Let $X \in \mathbb{S}_{++}^m$. Then $T_X\mathbb{S}_{++}^m = \mathbb{S}^m$, where $\mathbb{S}^m = \{X \in \mathbb{R}^{m \times m} \mid X \text{ is a symmetric matrix}\}$. Furthermore, the inner product on $T_X\mathbb{S}_{++}^m$ is defined by

$$\langle U, V \rangle_X := \text{Tr}(X^{-1}UX^{-1}V) \quad \text{for each } U, V \in T_X\mathbb{S}_{++}^m.$$

Hence, \mathbb{S}_{++}^m is a Hadamard manifold with nonpositive curvature everywhere (cf. [47, p. 355], [43]). Let $X_0, Y_0 \in \mathbb{S}_{++}^m$ be such that $\det X_0 = \det Y_0 = 1$. Let γ be the geodesic connecting X_0 and Y_0 , that is, $\gamma(0) = X_0$, $\gamma(1) = Y_0$, and

$$\gamma(t) = X^{\frac{1}{2}} \left(X^{-\frac{1}{2}}YX^{-\frac{1}{2}} \right)^t X^{\frac{1}{2}} \quad \text{for each } t \in [0, 1].$$

Write

$$S := \{\gamma(t) \mid t \in [0, 1]\}.$$

Note that $\bigcup_{X \in \mathbb{S}_{++}^m} T_X\mathbb{S}_{++}^m = \mathbb{S}^m$. Define the vector field $A_1, A_2 : \mathbb{S}_{++}^m \rightrightarrows \mathbb{S}^m$, respectively, by

$$A_1(X) := 2(\ln \det X)X - \exp_X^{-1} P_S X \quad \text{for each } X \in \mathbb{S}_{++}^m,$$

and

$$A_2(X) := \begin{cases} 2(\ln \det X)X - \frac{\exp_X^{-1} P_S X}{d_S(X)} & \text{for each } X \in \mathbb{S}_{++}^m \setminus S, \\ 2(\ln \det X)X + N_S(X) \cap \mathbb{B}_X & \text{for each } X \in S, \end{cases}$$

where P_S denotes the projection on S . Recall from [15] that the vector field $X \rightarrow 2(\ln \det X)X$ is monotone (but this vector field isn't monotone in the classical sense), and recall also from [46] that the vector field $X \rightarrow -\exp_X^{-1} P_S X$ is monotone. Hence, A_1 and A_2 are monotone. Observe further that $A_1^{-1}(0) = A_2^{-1}(0) = S$. Then, with a similar technique to Example 6.2, one can verify that A_1 is globally metrically subregular with modulus $\kappa = 1$. Furthermore, by (4.3), we see that A_2 satisfies a global weak sharp minima-like condition. Then, Corollary 3.5 is applied to A_1 to conclude that, with $\{\lambda_n\}$ satisfying $\sup_n \lambda_n < +\infty$, any sequence $\{x_n\}$, generated by IP₁ with $\sum_n \delta_n < \infty$, or by Algorithm IP₂ with $\sum_n \delta_n^2 < \infty$, converges linearly to a point $\bar{x} \in A_1^{-1}(0)$, and $\{x_n\}$ is superlinearly convergent if $\lim_{n \rightarrow \infty} \lambda_n = 0$, while Corollary 4.7 is applied to A_2 to conclude that any sequence $\{x_n\}$, generated by Algorithm IP₂ with $\{\delta_n\}$ satisfying $\sum \delta_n^2 < +\infty$ and the parameters $\{\lambda_n\}$ satisfying (4.15), terminates in a finite number of iterations.

The numerical simulation aims to verify the linear convergence and finite convergence of Algorithms P, IP₁, and IP₂, established in the present paper. These algorithms are implemented for vector fields A_1 and A_2 on the symmetry positive matrix manifold \mathbb{S}_{++}^{10} in Example 6.3, respectively. In both cases, we randomly generate $X_0, Y_0 \in \mathbb{S}_{++}^{10}$ and initial point $x_0 \in \mathbb{S}_{++}^{10}$. Figure 1 plots the distances to the solution set along the number of iterations in a random trial: (a) is for algorithms P and IP₁, and (b) is for algorithms P and IP₂. It is illustrated by Figure 1 that the algorithms P, IP₁, and IP₂ are linearly convergent in the case when each $\lambda_n = \frac{1}{3}$ and superlinearly convergent when each $\lambda_n = \frac{1}{n+1}$. The finite convergence results for vector field A_2 are demonstrated in Tables 1, 2, and 3 for algorithms P, IP₁, and IP₂, respectively.

We end this paper with an example on the unit sphere in \mathbb{R}^3 , which shows that our results obtained in the present paper cannot be extended directly to a general Riemannian manifold setting.

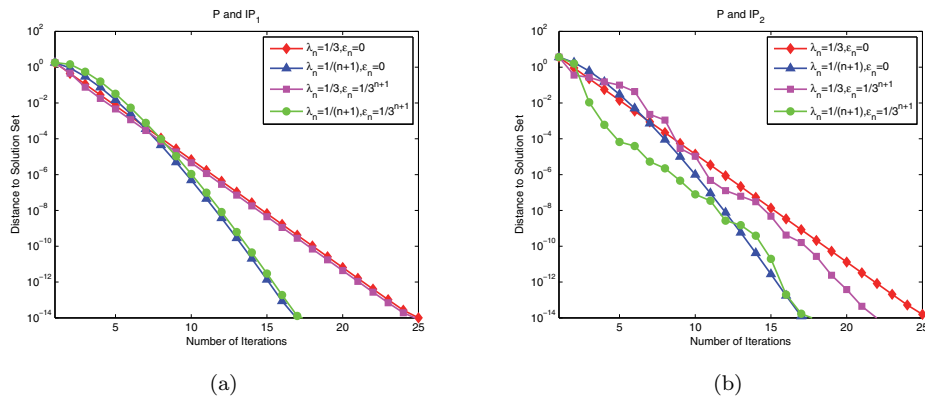


FIG. 1. Example 6.3 for A_1 .

TABLE 1
Example 6.3: \mathbf{P} for A_2 .

Iteration number n	$d_S(x_n)$	
	$\lambda_n := 2$	$\lambda_n := \sqrt{n+4}$
0	2.233920999084185	2.233920999084185
1	1.733920999084184	1.733920999084184
2	1.233920999084183	1.286707403584226
3	0.733920999084186	0.878459113120364
4	0.233920999084186	0.500494640111139
5	$x_5 \in S$	0.146941249517865
6		$x_6 \in S$

TABLE 2
Example 6.3: \mathbf{IP}_1 for A_2 with $\varepsilon_n := \frac{1}{3^n}$.

Iteration number n	$d_S(x_n)$	
	$\lambda_n := 2$	$\lambda_n := \sqrt{n+4}$
0	2.233920999084185	2.233920999084185
1	1.821521307724848	2.217547128688538
2	1.193546460337486	1.885254581098946
3	0.693365113332030	1.515619721840838
4	0.193670632261136	1.160862584759840
5	$x_5 \in S$	0.828914576382055
6		0.512671258870780
7		0.211117999117864
8		$x_8 \in S$

TABLE 3
Example 6.3: \mathbf{IP}_2 for A_2 with $\varepsilon_n := \frac{1}{3^n}$.

Iteration number n	$d_S(x_n)$	
	$\lambda_n := 2$	$\lambda_n := \sqrt{n+4}$
0	2.233920999084185	2.233920999084185
1	1.502241607420368	2.766956223934800
2	1.310317523279516	2.416771773578742
3	0.887048452410537	3.129591719836815
4	0.614692318479917	1.548693180993704
5	0.419260719126130	0.500470820802963
6	$x_6 \in S$	0.069219843701401
7		$x_7 \in S$

Example 6.4. Let

$$M = \mathbb{S}^2 := \{(y_1, y_2, y_3) \in \mathbb{R}^3 \mid y_1^2 + y_2^2 + y_3^2 = 1\}$$

be the 2-dimensional unit sphere. Write $\mathbf{x} := (0, 0, 1)$ and $\mathbf{y} := (0, 0, -1)$. Then $\mathbb{S}^2 \setminus \{\mathbf{x}, \mathbf{y}\}$ can be parametrized by $\Phi : (0, \pi) \times [0, 2\pi] \subset \mathbb{R}^2 \rightarrow \mathbb{S}^2 \setminus \{\mathbf{x}, \mathbf{y}\}$ defined by $\Phi(\theta, \varphi) := (y_1, y_2, y_3)^T$ for each $\theta \in (0, \pi)$ and $\varphi \in [0, 2\pi]$, where

$$\begin{cases} y_1 := \sin \theta \cos \varphi, \\ y_2 := \sin \theta \sin \varphi, \\ y_3 := \cos \theta. \end{cases}$$

Clearly, for each $x \in \mathbb{S}^2 \setminus \{\mathbf{x}, \mathbf{y}\}$, $(\mathbb{S}^2 \setminus \{\mathbf{x}, \mathbf{y}\}, \Phi^{-1})$ is a system of coordinates around x . Then the Riemannian metric on $\mathbb{S}^2 \setminus \{\mathbf{x}, \mathbf{y}\}$ is given by

$$g_{11} = 1, \quad g_{12} = 0, \quad g_{22} = \sin^2 \theta \quad \text{for each } \theta \in (0, \pi), \varphi \in [0, 2\pi].$$

The geodesics of $\mathbb{S}^2 \setminus \{\mathbf{x}, \mathbf{y}\}$ are great circles or semicircles. By the definition of the Riemannian metric on \mathbb{S}^2 , one can check

$$(6.8) \quad \langle u, v \rangle_{\bar{x}} = \langle u, v \rangle \quad \text{for any pair } (u, v) \in T_{\bar{x}}\mathbb{S}^2 \times T_{\bar{x}}\mathbb{S}^2,$$

where $\langle \cdot, \cdot \rangle_{\bar{x}}$ and $\langle \cdot, \cdot \rangle$ denote the inner product in $T_{\bar{x}}\mathbb{S}^2$ and \mathbb{R}^3 , respectively; see, for example, [54, p. 84].

Let

$$(6.9) \quad z_1 := \left(0, \frac{\sqrt{3}}{3}, \frac{\sqrt{6}}{3}\right), \quad z_2 := \left(0, \frac{\sqrt{6}}{3}, \frac{\sqrt{3}}{3}\right),$$

and consider the set C given by

$$C := \{(y_1, y_2, y_3) \in \mathbb{S}^2 \mid y_1 = 0, y_2 > 0, y_3 > 0\}.$$

Clearly C is strongly convex. Define the multivalued $A : M \rightrightarrows TM$ by

$$(6.10) \quad A(x) := \begin{cases} -\frac{\exp_x^{-1} z_1}{d(x, z_1)} - \frac{\exp_x^{-1} z_2}{d(x, z_2)} + N_C(x) & \text{for each } x \in C \text{ and } x \neq z_1, z_2, \\ \mathbb{B}_{z_1} - \frac{\exp_{z_1}^{-1} z_2}{d(z_1, z_2)} + N_C(z_1), & x = z_1, \\ \mathbb{B}_{z_2} - \frac{\exp_{z_2}^{-1} z_1}{d(z_1, z_2)} + N_C(z_2), & x = z_2, \\ \emptyset, & \text{otherwise,} \end{cases}$$

where $\mathbb{B}_{z_i} := \{v \in T_{z_i}M \mid \|v\| \leq 1\}$ for each $i = 1, 2$. Then, one checks that A is monotone and

$$A^{-1}(0) = S := \left\{ (y_1, y_2, y_3) \in \mathbb{S}^2 \mid y_1 = 0, y_2, y_3 \in \left[\frac{\sqrt{3}}{3}, \frac{\sqrt{6}}{3} \right] \right\}.$$

We first show that Problem (3.1) with A given by (6.10) satisfies the global weak sharp minima-like condition (with modulus $\alpha := 1$)

$$(6.11) \quad \mathbb{B}_{\bar{x}} \subset A(\bar{x}) + T_{\bar{x}}S \quad \text{for any } \bar{x} \in S$$

(and so A is metrically subregular by Remark 4.1). To do this, let $\bar{x} \in S$. If $\bar{x} \neq z_1$ and $\bar{x} \neq z_2$, then $T_{\bar{x}}S = T_{\bar{x}}C$ and so $N_C(\bar{x}) + T_{\bar{x}}S = T_{\bar{x}}M$. Therefore, the inclusion

in (6.11) holds for any $\bar{x} \in S \setminus \{z_1, z_2\}$. Thus, it remains to consider the case when $\bar{x} = z_1$ or z_2 . Without loss of generality, we assume that $\bar{x} = z_1$. Then, we conclude that $T_{\bar{x}}C = \{(0, y_2, y_3) \in \mathbb{R}^3 \mid y_2 + \sqrt{2}y_3 = 0\}$ and then

$$N_C(\bar{x}) = \{(y_1, 0, 0) \mid y_1 \in \mathbb{R}\}.$$

Moreover, one checks that

$$T_{\bar{x}}S = \left\{ (0, y_2, y_3) \in \mathbb{R}^3 \mid y_2 + \sqrt{2}y_3 = 0, y_3 \leq 0 \right\}.$$

Hence,

$$N_C(\bar{x}) + T_{\bar{x}}S = \left\{ (y_1, y_2, y_3) \in \mathbb{R}^3 \mid y_2 + \sqrt{2}y_3 = 0, y_3 \leq 0 \right\}.$$

By the definition of A , we have that

$$A(\bar{x}) = \mathbb{B}_{\bar{x}} + \left\{ -\frac{\exp_{\bar{x}}^{-1} z_2}{d(\bar{x}, z_2)} \right\} + N_C(\bar{x}) = \mathbb{B}_{\bar{x}} + \frac{1}{\sqrt{3}} (0, -\sqrt{2}, 1) + N_C(\bar{x}).$$

Noting that $\frac{1}{\sqrt{3}}(0, \sqrt{2}, -1) \in N_C(\bar{x}) + T_{\bar{x}}S$, we see that

$$A(\bar{x}) + T_{\bar{x}}S = \mathbb{B}_{\bar{x}} + \frac{1}{\sqrt{3}} (0, -\sqrt{2}, 1) + N_C(\bar{x}) + T_{\bar{x}}S \supseteq \mathbb{B}_{\bar{x}}.$$

Hence, (6.11) is shown.

Next take $w_0, w_1 \in C$ as follows:

$$w_0 := \left(0, \frac{\sqrt{35}}{6}, \frac{1}{6} \right), \quad w_1 := \left(0, \frac{3\sqrt{7}}{8}, \frac{1}{8} \right).$$

Define sequences $\{x_n\} \subset M$ and $\{\lambda_n\} \subseteq \mathbb{R}_+$, respectively, by (6.12)

$$x_n := \begin{cases} w_0, & n = 2k, \\ w_1, & n = 2k + 1, \end{cases} \quad \text{and} \quad \lambda_n := \begin{cases} \frac{2}{2\pi - (\arctan \frac{1}{\sqrt{35}} - \arctan \frac{1}{3\sqrt{7}})}, & n = 2k, \\ \frac{2}{\arctan \frac{1}{\sqrt{35}} - \arctan \frac{1}{3\sqrt{7}}}, & n = 2k + 1 \end{cases}$$

for any $k := 0, 1, 2, \dots$. Below, we show that

$$(6.13) \quad 0 \in A(x_{n+1}) - \lambda_n \exp_{x_{n+1}}^{-1} x_n \quad \text{for all } n = 0, 1, \dots,$$

granting this, $\{x_n\}$ is the sequence generated by Algorithm P with initial point x_0 and the parameters $\{\lambda_n\}$ (noting that $\sup_n \lambda_n < \infty$). Clearly $\{x_n\}$ doesn't converge; hence Corollaries 3.6 and 4.9 (and so Theorems 3.4 and 4.6) fail.

To show (6.13), recall that z_1 and z_2 are given by (6.9). Then

$$u_1 := \left[2\pi - \left(\arctan \frac{1}{\sqrt{35}} - \arctan \frac{1}{3\sqrt{7}} \right) \right] \left(0, \frac{1}{8}, -\frac{3\sqrt{7}}{8} \right) \in \exp_{w_1}^{-1} w_0,$$

$$u_2 := \left(\arctan \sqrt{2} - \arctan \frac{1}{3\sqrt{7}} \right) \left(0, -\frac{1}{8}, \frac{3\sqrt{7}}{8} \right) \in \exp_{w_1}^{-1} z_1,$$

$$u_3 := \left(\arctan \frac{1}{\sqrt{2}} - \arctan \frac{1}{3\sqrt{7}} \right) \left(0, -\frac{1}{8}, \frac{3\sqrt{7}}{8} \right) \in \exp_{w_1}^{-1} z_2.$$

It follows from (6.10) that

$$A(w_1) = -\frac{\exp_{w_1}^{-1} z_1}{d(w_1, z_1)} - \frac{\exp_{w_1}^{-1} z_2}{d(w_1, z_2)} + N_C(w_1)$$

and

$$(6.14) \quad 0 = -\frac{u_2}{\|u_2\|} - \frac{u_3}{\|u_3\|} + 0 - \lambda_0 u_1 \in A(w_1) - \lambda_0 \exp_{w_1}^{-1} w_0.$$

Similarly,

$$\begin{aligned} v_1 &:= \left(\arctan \frac{1}{\sqrt{35}} - \arctan \frac{1}{3\sqrt{7}} \right) \left(0, \frac{1}{6}, -\frac{\sqrt{35}}{6} \right) \in \exp_{w_0}^{-1} w_1, \\ v_2 &:= \left(\arctan \sqrt{2} - \arctan \frac{1}{\sqrt{35}} \right) \left(0, -\frac{1}{6}, \frac{\sqrt{35}}{6} \right) \in \exp_{w_0}^{-1} z_1, \\ v_3 &:= \left(\arctan \frac{1}{\sqrt{2}} - \arctan \frac{1}{\sqrt{35}} \right) \left(0, -\frac{1}{6}, \frac{\sqrt{35}}{6} \right) \in \exp_{w_0}^{-1} z_2. \end{aligned}$$

Then one can check that

$$(6.15) \quad 0 = -\frac{v_2}{\|v_2\|} - \frac{v_3}{\|v_3\|} + 0 - \lambda_1 v_1 \in A(w_0) - \lambda_1 \exp_{w_0}^{-1} w_1.$$

In view of the definitions of $\{x_n\}$ and $\{\lambda_n\}$ in (6.12), (6.13) follows from (6.14) and (6.15).

$$\begin{aligned} x_{2n} = w_0, \quad w_{2n+1} = w_1, \quad \lambda_{2n} &= \frac{2}{2\pi - \left(\arctan \frac{1}{\sqrt{35}} - \arctan \frac{1}{3\sqrt{7}} \right)}, \\ \lambda_{2n+1} &= \frac{2}{\arctan \frac{1}{\sqrt{35}} - \arctan \frac{1}{3\sqrt{7}}}. \end{aligned}$$

REFERENCES

- [1] F. J. ARAGON ARTACHO AND M. H. GEOFFROY, *Characterization of metric regularity of subdifferentials*, J. Convex Anal., 15 (2008), pp. 365–380.
- [2] F. J. ARAGON ARTACHO AND B. S. MORDUKHOVICH, *Metric regularity and Lipschitzian stability of parametric variational systems*, Nonlinear Anal., 72 (2010), pp. 1149–1170.
- [3] F. J. ARAGON ARTACHO AND B.S. MORDUKHOVICH, *Enhanced metric regularity and Lipschitzian properties of variational systems*, J. Global Optim. 50 (2011), pp. 145–167.
- [4] R. ADLER, J. P. DEDIEU, J. MARGULIES, M. MARTENS, AND M. SHUB, *Newton’s method on Riemannian manifolds and a geometric model for human spine*, IMA J. Numer. Anal., 22 (2002), pp. 359–390.
- [5] H. ATTOUCH AND J. BOLTE, *On the convergence of the proximal algorithm for nonsmooth functions involving analytic features*, Math. Program. Ser. B, 116 (2009), pp. 5–16.
- [6] P. A. ABSIL, R. MAHONY, AND R. SEPULCHRE, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ, 2008.
- [7] D. AZAGRA, J. FERRERA, AND F. LÓPEZ-MESAS, *Nonsmooth analysis and Hamilton-Jacobi equations on Riemannian manifolds*, J. Funct. Anal., 220 (2005), pp. 304–361.
- [8] M. BAČÁK, R. BERGMANN, G. STEIDL, AND A. WEINMANN, *A second order nonsmooth variational model for restoring manifold-valued images*, SIAM J. Sci. Comput., 38 (2016), pp. A567–A597.
- [9] H. H. BAUSCHKE AND P. L. COMBETTES, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, 2011.

- [10] G. C. BENTO AND J. G. MELO, *Subgradient method for convex feasibility on Riemannian manifolds*, J. Optim. Theory Appl., 152 (2012), pp. 773–785.
- [11] G. C. BENTO AND J. X. CRUZ NETO, *Finite termination of the proximal point method for convex functions on Hadamard manifolds*, Optim., 63 (2014), pp. 1281–1288.
- [12] J. V. BURKE AND S. DENG, *Weak sharp minima revisited, I: Basic theory*, Control Cybernet. 31 (2002), pp. 439–469.
- [13] J. V. BURKE AND M. C. FERRIS, *Weak sharp minima in mathematical programming*, SIAM J. Control Optim., 31 (1993), pp. 1340–1359.
- [14] F. E. BROWDER, *Multi-valued monotone nonlinear mappings and duality mappings in Banach spaces*, Trans. Amer. Math. Soc., 118 (1965), pp. 338–351.
- [15] J. X. DA CRUZ NETO, O. P. FERREIRA, L. R. LUCAMBIO PÉREZ, AND S. Z. NÉMETH, *Convex and monotone-transformable mathematical programming problems and a proximal-like point method*, J. Global Optim., 35 (2006), pp. 53–69.
- [16] J. X. DA CRUZ NETO, O. P. FERREIRA, AND L. R. LUCAMBIO PÉREZ, *Monotone point-to-set vector fields*, Balkan J. Geom. Appl., 5 (2000), pp. 69–79.
- [17] J. P. DEDIEU, P. PRIOURET, AND G. MALAJOVICH, *Newton’s method on Riemannian manifolds: Covariant alpha theory*, IMA J. Numer. Anal., 23 (2003), pp. 395–419.
- [18] M. P. DO CARMO, *Riemannian Geometry*, Birkhäuser, Boston, 1992.
- [19] A. L. DONTCHEV AND R. T. ROCKAFELLAR, *Implicit Functions and Solution Mappings: A View from Variational Analysis*, Springer, New York, 2009.
- [20] A. L. DONTCHEV AND R. T. ROCKAFELLAR, *Regularity and conditioning of solution mappings in variational analysis*, Set-Valued Anal., 12 (2004), pp. 79–109.
- [21] O. P. FERREIRA AND P. R. OLIVEIRA, *Proximal point algorithm on Riemannian manifolds*, Optim., 51 (2002), pp. 257–270.
- [22] O. P. FERREIRA AND B. F. SVAITER, *Kantorovich’s theorem on Newton’s method in Riemannian manifolds*, J. Complexity, 18 (2002), pp. 304–329.
- [23] H. GFRERER, *First order and second order characterizations of metric subregularity and calmness of constraint set mappings*, SIAM J. Optim., 21 (2011), pp. 1439–1474.
- [24] O. GÜLER, *New proximal point algorithms for convex minimization*, SIAM J. Optim., 2 (1992), pp. 649–664.
- [25] P. GROHS AND S. HOSSEINI, *Nonsmooth trust region algorithms for locally Lipschitz functions on Riemannian manifolds*, IMA J. Numer. Anal., 36 (2016), pp. 1167–1192.
- [26] P. GROHS AND S. HOSSEINI, ε -subgradient algorithms for locally Lipschitz functions on Riemannian manifolds, Adv. Comput. Math., 42 (2016), pp. 333–360.
- [27] R. HENRION AND J. V. OUTFRATA, *Calmness of constraint systems with applications*, Math. Program., 104 (2005), pp. 437–464.
- [28] R. HENRION AND T. SUROWIEC, *On calmness conditions in convex bilevel programming*, Appl. Anal., 90 (2011), pp. 951–970.
- [29] W. HUANG, K. A. GALLIVAN, AND P.-A. ABSIL, *A Broyden class of quasi-Newton methods for Riemannian optimization*, SIAM J. Optim., 25 (2015), pp. 1660–1685.
- [30] A. D. IOFFE AND J. V. OUTFRATA, *On metric and calmness qualification conditions in subdifferential calculus*, Set-Valued Anal., 16 (2008), pp. 199–277.
- [31] A. V. KRYANEV, *The solution of incorrectly posed problems by methods of successive approximations*, Soviet Math., 14 (1973), pp. 673–676.
- [32] Y. S. LEDYAEV AND Q. J. ZHU, *Nonsmooth analysis on smooth manifolds*, Trans. Amer. Math. Soc., 359 (2007), pp. 3687–3732.
- [33] C. LI, G. LÓPEZ, AND V. MARTÍN-MÁRQUEZ, *Monotone vector fields and the proximal point algorithm on Hadamard manifolds*, J. Lond. Math. Soc. (2), 79 (2009), pp. 663–683.
- [34] C. LI, B. S. MORDUKHOVICH, J. WANG, AND J.-C. YAO, *Weak sharp minima on Riemannian manifolds*, SIAM J. Optim., 21 (2011), pp. 1523–1560.
- [35] C. LI AND J.-C. YAO, *Variational inequalities for set-valued vector fields on Riemannian manifolds: Convexity of the solution set and the proximal point algorithm*, SIAM J. Control Optim., 50 (2012), pp. 2486–2514.
- [36] C. LI AND J. H. WANG, *Convergence of the Newton method and uniqueness of zeros of vector fields on Riemannian manifolds*, Sci. China Ser. A., 48 (2005), pp. 1465–1478.
- [37] C. LI AND J. H. WANG, *Newton’s method on Riemannian manifolds: Smale’s point estimate theory under the γ -condition*, IMA J. Numer. Anal., 26 (2006), pp. 228–251.
- [38] C. LI AND J. H. WANG, *Newton’s method for sections on Riemannian manifolds: Generalized covariant α -theory*, J. Complexity, 24 (2008), pp. 423–451.
- [39] S. L. LI, C. LI, Y. C. LIU, AND J.-C. YAO, *Existence of solutions for variational inequalities on Riemannian manifolds*, Nonlinear Anal., 71 (2009), pp. 5695–5706.
- [40] F. J. LUQUE, *Asymptotic convergence analysis of the proximal point algorithm*, SIAM J. Control Optim., 22 (1984), pp. 277–293.

- [41] B. MARTINET, *Régularisation d'inéquations variationnelles par approximations successives*, Rev. Française Inform. Rech. Oper., 4 (1970), pp. 154–159.
- [42] G. J. MINTY, *On the monotonicity of the gradient of a convex function*, Pacific J. Math., 14 (1964), pp. 243–247.
- [43] M. MOAKHER, *A differential geometric approach to the geometric mean of symmetric positive-definite matrices*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 735–747.
- [44] S. Z. NÉMETH, *Monotone vector fields*, Publ. Math. Debrecen, 54 (1999), pp. 437–449.
- [45] S. Z. NÉMETH, *Variational inequalities on Hadamard manifolds*, Nonlinear Anal., 52 (2003), pp. 1491–1498.
- [46] S. Z. NÉMETH, *Monotonicity of the complementary vector field of a non-expansive map*, Acta Math. Hungar., 84 (1999), pp. 189–197.
- [47] Y. E. NESTEROV AND M. J. TODD, *On the Riemannian geometry defined by self-concordant barriers and interior-point methods*, Found. Comput. Math., 2 (2002), pp. 333–361.
- [48] P. PETERSEN, *Riemannian Geometry*, Grad. Texts in Math. 171, 2nd ed., Springer, New York, 2006.
- [49] B. T. POLYAK, *Introduction to Optimization*, Optimization Software, New York, 1987.
- [50] T. RAPCSK, *Smooth nonlinear optimization in \mathbb{R}^n* , Nonconvex Optim. Appl., 19, Kluwer Academic, Dordrecht, 1997.
- [51] R. T. ROCKAFELLAR, *Monotone operators and the proximal point algorithm*, SIAM J. Control Optim., 14 (1976), pp. 877–898.
- [52] T. SAKAI, *Riemannian Geometry*, Transl. Math. Monogr. 149, AMS, Providence, RI, 1996.
- [53] S. T. SMITH, *Optimization techniques on Riemannian manifolds*, in *Fields Institute Communications*, Vol. 3, AMS, Providence, RI, 1994, pp. 113–146.
- [54] C. UDRISTE, *Convex Functions and Optimization Methods on Riemannian Manifolds*, Math. Appl. 297, Kluwer Academic, Dordrecht, 1994.
- [55] J. H. WANG AND C. LI, *Uniqueness of the singular points of vector fields on Riemannian manifolds under the γ -condition*, J. Complexity, 22 (2006), pp. 533–548.
- [56] J. H. WANG AND C. LI, *Kantorovich's theorems for Newton's method for mappings and optimization problems on Lie groups*, IMA J. Numer. Anal., 31 (2011), pp. 322–347.
- [57] J. H. WANG, C. LI, AND J.-C. YAO, *Finite termination of inexact proximal point algorithms in Hilbert spaces*, J. Optim. Theory Appl., 166 (2015), pp. 188–212.
- [58] J. H. WANG, G. LÓPEZ, V. MARTÍN-MÁRQUEZ, AND C. LI, *Monotone and accretive operators on Riemannian manifolds*, J. Optim. Theory Appl., 146 (2010), pp. 691–708.
- [59] J. H. WANG, C. LI, G. LOPEZ, AND J.-C. YAO, *Convergence analysis of inexact proximal point algorithms on Hadamard manifolds*, J Global Optim., 61 (2015), pp. 553–573.
- [60] E. ZEIDLER, *Nonlinear Functional Analysis and Applications. II B, Nonlinear Monotone Operators*, Springer, New York, 1990.
- [61] L.-H. ZHANG, *Riemannian Newton method for the multivariate eigenvalue problem*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 2972–2996.
- [62] X. Y. ZHENG AND K. F. NG, *Metric subregularity and constraint qualifications for convex generalized equations in Banach spaces*, SIAM. J. Optim., 18 (2007), pp. 437–460.
- [63] X. Y. ZHENG AND K. F. NG, *Metric subregularity and calmness for nonconvex generalized equations in Banach spaces*, SIAM J. Optim., 20 (2010), pp. 2119–2136.