J Glob Optim (2009) 45:631–644
DOI 10.1007/s10898-008-9388-5

Asymptotic convergence of an inertial proximal method for unconstrained quasiconvex minimization

Paul-Emile Maingé

Received: 14 October 2007 / Accepted: 4 December 2008 / Published online: 27 December 2008

Abstract This paper deals with the convergence analysis of a second order proximal method for approaching critical points of a smooth and quasiconvex objective function defined on a real Hilbert space. The considered method, well-known in the convex case, unifies proximal method, relaxation and inertial-type extrapolation. The convergence theorems established in this new setting improve recent ones.

Keywords Quasiconvex minimization · Proximal algorithm · Inertial algorithm · Over-relaxation

Mathematics Subject Classification (2000) 49M45 · 65C25

1 Introduction

In this paper, we study the asymptotic behavior of some modified proximal algorithms for computing critical points of a real-valued function f, assumed to be quasiconvex (see Definition 2.1) and differentiable over a real Hilbert space $\mathcal{H}$. As a special case, we also pay some attention to the unconstrained minimization problem

$$\min_{x\in\mathcal{H}} f(x). \tag{1}$$

The space $\mathcal{H}$ will be endowed with a scalar product $\langle .,.\rangle$ and its induced norm $|.|$. The algorithms on which we focus, are mainly based upon an implicit discretization of the following second order evolution system [1,6,10]

$$\begin{cases} x''(t) + \gamma x'(t) + \nabla f(x(t)) = 0, \\ x(0) = x_0, \quad x'(0) = y_0, \end{cases} \tag{2}$$

P.-E. Maingé (✉)
Département Scientifique Interfacultaire, GRIMAAG, Université des Antilles et de la Guyane,
Campus de Schoelcher, 97230 Cedex, Martinique (F.W.I.), France
e-mail: Paul-Emile.Mainge@martinique.univ-ag.fr

where x_0, y_0 belong to $\mathcal{H}$, γ is a positive (damping or friction) term, and $\nabla f(.)$ stands for the gradient of f. Equation (2), also called "heavy ball with friction dynamical system", gave rise to various numerical methods (for monotone inclusions and fixed point problems) related to the "inertial" terminology (first introduced in [3]). They incorporate second order information to achieve nice convergence properties [1–3,18,19]. In particular, the so-called inertial proximal method was originally proposed in [1] as an algorithmic solution to (1), when f is convex and possibly noonsmooth. It consists of the iteration

$$x_{n+1} = \operatorname{argmin}_{x\in\mathcal{H}} \left\{ f(x) + (1/2)\lambda_n^{-1}|x - x_n - \theta_n(x_n - x_{n-1})|^2 \right\}, \tag{3}$$

with parameters $(\lambda_n) \subset (0, +\infty)$ and damping terms $(\theta_n) \subset [0, 1)$. In the special case when $\theta_n = 0$, the method (3) reduces to the classical proximal point algorithm initiated in [17] (also see [12,23,24]), namely

$$x_{n+1} = \operatorname{argmin}_{x\in\mathcal{H}} \left\{ f(x) + (1/2)\lambda_n^{-1}|x - x_n|^2 \right\}. \tag{4}$$

Let us briefly recall the motivation for the previously mentioned processes. Proximal algorithms were deeply studied as fundamental tools for solving ill-posed or ill-conditioned (convex or nonconvex) minimization problems. These methods consists of replacing the original objective by a sequence of better behaved functionals. This leads to solve more computationally stable problems. Furthermore, due to their mechanical interpretations, second order dynamical models such as (2) open interesting perspectives for designing new efficient algorithms. Indeed, in the special case when $\mathcal{H} = \mathbb{R}^2$, (2) is a simplified model which describes the motion of a heavy ball rolling over the graph of f and which keeps rolling under its own inertia until stopped by friction at a critical point of f [6]. Let us point out that (2) is not a descent method, but it is some global energy of the system that decreases, which allows to overcome some drawbacks of the steepest descent method

$$x'(t) + \nabla f(x(t)) = 0. \tag{5}$$

For instance, this latter system is unable to cross any non-minimum critical point of f, while (2) captures some exploring properties of the ball's motion. For further details related to the exploration of local minima and their practical applications, we refer to [4,10]. The second-order nature of (2) may be also exploited to obtain faster convergence [3,13]. In contrast with (3), the standard proximal point algorithm is based upon an implicit discretization of the first order method (5). The "inertia" induced by the term "$\theta_n(x_n - x_{n-1})$" in (3) considerably improves the speed of convergence comparing with (4) (see [13], for numerical experiments). This can be explained by the fact that the vector "$x_n - x_{n-1}$" acts as an impulsion term (mostly at the beginning of trajectories), while the coefficient "θ_n" plays the role of a speed regulator.

On the other hand, quasiconvex minimization problems arise in several important applications in economic theory, location theory, control theory and so on [4,9–11,26]. It is noteworthy that numerical approaches to some of these problems can be obtained by steepest descent methods [16] (also see [15] for more efficient subgradient methods), based upon non-implicit discretizations of the first order equation (5). Proximal methods, based upon implicit discretizations of (2) and (5), were only recently introduced as alternative (also complementary) approaches to quasiconvex minimization [10,22]. Nonquadratic proximal methods were also investigated, by replacing the quadratic distance in (4) with a Bregman distance or an entropy-like distance [5,21]. Other interesting works combining the proximal methodology and nonconvex programming can be found in [14] for some situations when the objective function in (4) becomes strongly convex, while f itself is assumed to be nonconvex (but not necessarily quasiconvex). Note that most of these methods deal with the quasiconvex situation only in the finite dimensional case $\mathcal{H} = \mathbb{R}^n$, except for that of [15] and the inertial proximal method in [10] which are both considered in setting of Hilbert spaces.

Remark 1.1 Obviously, to deal with practical purposes, the proximal methods proposed above have to be combined with an optimization algorithm for solving auxiliary problems of the form $\min\{f(x) + d(x) : \ x \in \mathcal{H}\}$, where $d : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$ is strongly convex. Roughly speaking, one cannot expect that these regularized subproblems are easier solvable than the original one. In particular, when f is quasiconvex (or nonconvex), the resulting proximity functions are not necessary single-valued (i.e. a subproblem may admit more than one solution), besides the function to be minimized in a subproblem may not be quasiconvex (because the class of quasiconvex functions is not closed by addition). Nevertheless, the proximal-type methodology turned out to be of fundamental importance for the development of successful numerical approaches such as operator splitting and bundle methods (see, e.g., [25]). In the same spirit, applying proximal methods to quasiconvex functions may pave the way for new optimization techniques, although such works could be regarded as purely theoretical. Another point of interest is to keep in mind that proximal methods have been successfully applied to classes of nonconvex functions occurring in concrete applications [14] (also see [21] for some quasiconvex cases). Similar investigations could be done in the present framework, but this is out of scope of our study.

In our work, we focus on the inertial proximal method in [10], obtained by an implicit discretization of (2). First, we recall some of the main convergence results given in [10] regarding the continuous system (2). These results improve that of [6] (for a convex objective) and can be summarized as follows.

Theorem 1.1 ([10]) *Assume that f is $C^{1,1}$ and quasiconvex on $\mathcal{H}$, with $S = \operatorname{argmin}_{\mathcal{H}} f \neq \emptyset$. Then, for every trajectories $x(.)$ of system (2), the following properties hold:*

(i) $\lim_{t\to\infty} x'(t) = \lim_{t\to\infty} x''(t) = \lim_{t\to\infty} \nabla f(x(t)) = 0$ *and* $\lim_{t\to\infty} f(x(t))$ *exists.*
(ii) *$x(t)$ weakly converges in $\mathcal{H}$ towards some x_∞.*
(iii) $\nabla f(x_\infty) = 0$.
(iv) *If $x_\infty \notin \operatorname{argmin}_{\mathcal{H}} f$, then the convergence is strong.*

Theorem 1.1 puts out the remarkable properties of system (2) regarding the exploration of non-minimum critical point of f. A discrete version of this theorem was then established for the inertial proximal method in [10]. This latter method, considered in the setting of Hilbert spaces, is written as

$$x_{n+1} \in \operatorname{argmin}_{x\in\mathcal{H}} \left\{ f(x) + (1/2)\lambda_n^{-1} |x - x_n - \theta_n(x_n - x_{n-1})|^2 \right\}, \tag{6}$$

with particular parameters (λ_n) and (θ_n) set to be such that

$$\theta_n = \frac{h_n}{h_{n-1}} \frac{1}{(1+\nu h_n)}, \qquad \lambda_n = \frac{h_n^2}{(1+\nu h_n)}, \tag{7}$$

where $\nu \in (0,\infty)$ and (h_n) is a positive sequence verifying $\lim_{n\to\infty} h_n = \infty$.

Remark 1.2 Let us emphasize that the method proposed in [10] only consider a special case of (6), also in the special range of parameter $\lambda_n \to \infty$. Interesting convergent results are given in [10] regarding either the exploration of local minima, namely the case when the iterates (x_n) converge to a non-minimum critical point, or the exploration of global minima. However, it seems that there is a gap in the proofs for optimality of weak accumulation points of (x_n). Indeed, the condition $h_n \to \infty$ entails $(1/\lambda_n) \to 0$, which cannot allow the exploration of minima else than global ones. This can be clarified in light of Theorems 3.1 and 3.2. Furthermore, studies only in the range $h_n \to \infty$ remain far from being satisfactory, since the method (4) is well-known to become computationally unstable as $(1/\lambda_n) \to 0$.

It is our purpose here to propose a correct analysis of the proximal method (6), for a broader range of the involved parameters, even in a more general setting which combines inertial-type extrapolation with suitable relaxation factors. The latter strategy is intended to act as a speeding up process. The spirit of our work is the same as that in [2] related to monotone operators. More precisely, given parameters $(\lambda_n) \subset (0, \infty)$, $(\theta_n) \subset [0, 1)$ and $(w_n) \subset (0, 2)$ (including the over-relaxation case $(w_n) \subset (1, 2)$), we investigate the following procedure:

Inertial and Relaxed Proximal Algorithm:

- Step 0 (Initialization). Choose any $x_0, \quad x_1 \in \mathcal{H}$.
- Step 1 (for $n \geq 1$). Set $v_n = x_n + \theta_n(x_n - x_{n-1})$, and compute

$$y_n \in \operatorname{argmin}_{\mathcal{H}} \left\{ f(x) + (1/2)\lambda_n^{-1}|x - v_n|^2 \right\}. \tag{8a}$$

- Step 2. Compute

$$x_{n+1} = (1 - w_n)v_n + w_n y_n; \tag{8b}$$

go to Step 1.

Let us precise that the method (8) makes sense, provided that f is bounded from below (see Proposition 2.1). This paper establishes the asymptotic convergence of the method (8) under conditions related to the involved parameters (damping term (θ_n), relaxation factor (w_n) and proximal stepsize (λ_n)). Two main convergence theorems are given (Theorems 3.1 and 3.2). It turns out that the main results claimed in [10] are rigorously stated in a new setting, while some of them are improved in either form or requirements of parameters.

Remark 1.3 To the best of our knowledge, results such as in Theorem 3.1 and 3.2 does not appear in the existing literature, even in the special case of (8) when $w_n \equiv 1$, $\theta_n \equiv 0$ and $\mathcal{H}$ is finite dimensional [22] (which corresponds to the classical proximal method). Then our work is also complementary to that of [22] (related to a possibly nonsmooth objective function).

Remark 1.4 Algorithm (8) would be considerably enhanced by introducing inexact computations in Step 1 (see Remark 3.2). However for legibility's sake we do not include such a procedure.

2 Basic results and preliminaries

We begin with definition and other characterization of quasiconvexity [7,8].

Definition 2.1 A real-valued function $f : \mathcal{H} \to \mathbb{R}$ is called quasiconvex on some convex subset C of $\mathcal{H}$ if, for every $t \in (0, 1)$ and for all x, y in C, there holds $f(tx + (1 - t)y) \leq \max\{f(x), f(y)\}$.

Remark 2.1 Definition 2.1 can be alternatively characterized by the fact that, for every $\alpha \in \mathbb{R}$, the level set $L_f(\alpha) := \{x \in C; \quad f(x) \leq \alpha\}$ is convex. In the case when C is an open convex set and f is continuously differentiable over C, Definition 2.1 is equivalent to ([8]):

$$(\forall x, y \in C), \quad f(y) \leq f(x) \Rightarrow \langle \nabla f(x), y - x \rangle \leq 0. \tag{9}$$

Clearly, the notion of quasiconvexity includes that of convexity. However, in contrast with convex functions, quasiconvex ones may present critical points which are not minima. Through our study, we will be essentially concerned with the characterization (9) of quasiconvexity.

Now we show that the proximal method (8) makes sense.

Proposition 2.1 *If f is quasiconvex, continous and bounded from below over $\mathcal{H}$, then the sequence (x_n) generated by (8) is well-defined with $(\lambda_n) \subset (0, \infty)$.*

Proof Clearly, it suffices to prove that each iterate y_n exists (even non-uniquely) for (8) to make sense. To put out this latter fact, we introduce the mapping $g : \mathcal{H} \to \mathbb{R}$ defined, for every $x \in \mathcal{H}$, by $g(x) = f(x) + \beta|x - v|^2$, where $\beta > 0$ and $v \in \mathcal{H}$. Then we just need to prove that the set of minimizers of g is nonempty, under the conditions that f is quasiconvex, continous and bounded from below on $\mathcal{H}$. Indeed, by the latter condition, we observe that g is bounded below (as $f \leq g$), and we set $m_* = \inf_{\mathcal{H}} g$ (hence m_* is finite). Thus, for any x in the level set $L_g(g(a)) := \{x \in C; \quad g(x) \leq g(a)\}$, a being some element in $\mathcal{H}$, we have $g(a) \geq m_* + \beta|x - v|^2$, so that $L_g(g(a))$ is bounded. Now consider a minimizing sequence (z_n) of g, namely (z_n) satisfies $m_* = \lim_{n\to\infty} g(z_n)$. It is then immediate that $g(z_n) \leq g(a)$ for $n \geq n_0$ (n_0 being some large enough integer). Consequently, we have $(z_n)_{n\geq n_0} \subset L_g(g(a))$, which proves that $(z_n)_{n\geq n_0}$ is a bounded sequence, hence we can consider a subsequence (z_{n_k}) of (z_n) such that (z_{n_k}) converges weakly to some element $q \in \mathcal{H}$ as $k \to \infty$. By convergence of the whole sequence $(g(z_n))$, we also have $\lim_{k\to\infty} g(z_{n_k}) = \lim_{n\to\infty} g(z_n) = m_*$. Furthermore, it is well-known that continuous quasi-convex functions are weakly lower semicontinuous. As f and the mapping $d : x \to \beta|x - v|^2$ are continuous and quasiconvex (thus they are weakly lower semicontinuous), we then obtain $m_* \leq g(q) = f(q) + d(q) \leq \liminf_{k\to\infty} f(z_{n_k}) + \liminf_{k\to\infty} d(z_{n_k}) \leq \liminf_{k\to\infty}(f(z_{n_k}) + d(z_{n_k})) = \liminf_{k\to\infty} g(z_{n_k}) = m_*$. Therefore, we deduce $g(q) = m_*$, which leads to $\operatorname{argmin}_{\mathcal{H}} g \neq \emptyset$. ◻

In the next section, we establish a discrete version of Theorem 1.1 for Algorithm (8). For simplicity's sake, we introduce here three preliminary propositions needed for our convergence analysis. The first one is the well-known discrete version of the Opial result ([20]).

Proposition 2.2 ([20]) *Let $\mathcal{H}$ be a Hilbert space and (x_n) a sequence in $\mathcal{H}$ such that there exists a nonempty set $\Omega \subset \mathcal{H}$ satisfying:*

(i) For every $u \in \Omega$, $\lim_{n\to\infty}|x_n - u|$ exists.
(ii) Any weak cluster point of (x_n) belongs to Ω.

Then, there exists $\bar{x} \in \Omega$ such that (x_n) weakly converges to $\bar{x}$.

The second proposition is a quite standard result (see, e.g., [3]).

Proposition 2.3 *Let (a_n), (δ_n), (θ_n) be nonnegative sequences verifying:*

$$a_{n+1} \leq \theta_n a_n + \delta_n, \quad \forall n \geq 0. \tag{10}$$

If $(\theta_n) \subset [0, \theta]$ (for some $\theta \in [0, 1)$) and $\sum_n \delta_n < \infty$, then $\sum_n a_n < \infty$.

The third proposition appears implicitly in [2,3], and its proof is given for the sake of completeness.

Proposition 2.4 *Let* (ϕ_n)*,* (δ_n) *and* (θ_n) *in* $[0,\infty)$ *be such that :*

$$\phi_{n+1}-\phi_n \le \theta_n(\phi_n-\phi_{n-1})+\delta_n. \tag{11}$$

If $(\theta_n)\subset[0,\theta]$ *(for some* $\theta\in[0,1)$*) and* $\sum_n \delta_n<\infty$*, then* (ϕ_n) *converges and* $\sum_n[\phi_{n+1}-\phi_n]_+<\infty$*, where* $[t]_+:=\max\{t,0\}$ *for any* $t\in(-\infty,\infty)$*.*

Proof Setting $u_n:=\phi_n-\phi_{n-1}$ and using (11), we have $[u_{n+1}]_+\le\theta^n[u_n]_++\delta_n$, which by Proposition 2.3 amounts to $\sum_{n\ge0}[u_{n+1}]_+<\infty$. Then, setting $w_n:=\phi_n-\sum_{j=1}^n[u_j]_+$, we deduce that the sequence (w_n) is bounded from below. Furthermore, by a simple calculation, we obtain $w_{n+1}=\phi_{n+1}-[u_{n+1}]_+-\sum_{j=1}^n[u_j]_+\le w_n$, hence (w_n) is nonincreasing. Consequently, we deduce that (w_n) is convergent, and so is (ϕ_n). □

3 Convergence analysis

In the remainder of the paper, we assume that the objective function f occurring in (1) is such that:

$$f \text{ is quasiconvex and continuously differentiable on } \mathcal{H}, \tag{12}$$

$$\inf_{x\in\mathcal{H}} f(x)>-\infty \text{ (i.e. } f \text{ is bounded from below).} \tag{13}$$

Two main theorems will be given regarding the asymptotic convergence of (8):

(1) The first one (Theorem 3.1) can be viewed as a discrete variant of Theorem 1.1, and it is stated under the additional conditions:

$$0<\inf_n w_n\le\sup_n w_n<2. \tag{14}$$

$$\exists\,\lambda>0 \text{ such that } (\lambda_n)\subset(\lambda,+\infty). \tag{15}$$

$$\exists\theta\in[0,1) \text{ such that } (\theta_n)\subset[0,\theta). \tag{16}$$

$$\sum_{n\ge1}\theta_n|x_n-x_{n-1}|^2<\infty. \tag{17}$$

(2) The second result (Theorem 3.2) establishes the same conclusions as Theorem 3.1 when conditions (16)–(17) are replaced by the following one:

$$\text{There exists } \delta_0\in\left(0,\frac{\delta}{\delta+\mu}\right), \text{ where } \delta=\inf_n\frac{2-w_n}{2w_n} \text{ and } \mu=\max\{1,2\delta\},$$
$$\text{such that } (\theta_n) \text{ is a nondecreasing sequence in } [0,\delta_0]. \tag{18}$$

Remark 3.1 Observe that conditions (16)–(17) in Theorem 3.1 are not restrictive at all, as they are easily implemented in numerical computations. To see this, choose $\theta\in[0,1)$, a sequence $(\epsilon_n)\subset[0,\infty)$ such that $\sum_n\epsilon_n<\infty$, and consider the sequence (e_n) defined by

$$e_n:=\min\left\{\frac{\epsilon_n}{|x_n-x_{n-1}|^2},\quad\theta\right\}\quad\text{if}\quad x_n\ne x_{n-1},\quad e_n:=\theta\quad\text{otherwise.} \tag{19}$$

Clearly, (16)–(17) are satisfied if, at each step (x_{n-1} and x_n being given), the update x_{n+1} is computed by (8) and by choosing θ_n such that $0\le\theta_n\le e_n$. Recall that one of the main convergence theorems in [3] (regarding the convex case) was treated under conditions (15)–(17).

Before proving our main results, we give two preliminary lemmas.

Lremma 3.1 *Let* (x_n) *and* (y_n) *be given by* (8) *under conditions* (12)–(17) *and assume the set* $U = \{z \in \mathcal{H}, \quad f(z) \leq \inf_n f(y_n)\}$ *is nonempty. Then the following statements hold:*

(i1) $(|x_n - q|)$ *converges for every* $q \in U$.
(i2) $\sum_{n\geq 0} |x_n - x_{n-1}|^2 < \infty$ *and* $\sum_{n\geq 0} |x_n - y_n|^2 < \infty$.
(i3) $\sum_{n\geq 0} \lambda_n w_n \langle \nabla f(y_n), y_n - q\rangle < \infty$ *for every* $q \in U$.

Proof Clearly, by (8a), we classically have $0 = (y_n - v_n)/\lambda_n + \nabla f(y_n)$, which in addition to (8b) leads to

$$x_{n+1} = v_n - \lambda_n w_n \nabla f(y_n), \tag{20}$$

hence, taking any $q \in U$, we obtain

$$-\lambda_n w_n \langle \nabla f(y_n), y_n - q\rangle = \langle x_{n+1} - v_n, y_n - q\rangle. \tag{21}$$

Moreover, again by (8b) we have $x_{n+1} - q = (1 - w_n)(v_n - q) + w_n(y_n - q)$, namely $y_n - q = \frac{1}{w_n}(x_{n+1} - q) - \frac{1-w_n}{w_n}(v_n - q)$, which by (21) gives

$$\begin{aligned} -\lambda_n w_n \langle \nabla f(y_n), y_n - q\rangle &= \left\langle x_{n+1} - v_n, \frac{1}{w_n}(x_{n+1} - q) - \frac{1 - w_n}{w_n}(v_n - q)\right\rangle \\ &= \frac{1}{w_n} \langle x_{n+1} - v_n, x_{n+1} - q\rangle + \frac{1 - w_n}{w_n} \langle v_n - x_{n+1}, v_n - q\rangle . \end{aligned} \tag{22}$$

Furthermore, for any $a, b \in \mathcal{H}$, it is easily checked that

$$\langle a, b\rangle = -(1/2)|a - b|^2 + (1/2)|a|^2 + (1/2)|b|^2, \tag{23}$$

so that

$$\langle x_{n+1} - v_n, x_{n+1} - q\rangle = -\frac{1}{2}|v_n - q|^2 + \frac{1}{2}|x_{n+1} - v_n|^2 + \frac{1}{2}|x_{n+1} - q|^2,$$
$$\langle v_n - x_{n+1}, v_n - q\rangle = -\frac{1}{2}|x_{n+1} - q|^2 + \frac{1}{2}|x_{n+1} - v_n|^2 + \frac{1}{2}|v_n - q|^2.$$

Combining (22) with the previous two inequalities, we obtain

$$\begin{aligned} -2\lambda_n w_n^2 \langle \nabla f(y_n), y_n - q\rangle &= -|v_n - q|^2 + |x_{n+1} - v_n|^2 + |x_{n+1} - q|^2 \\ &\quad + (1 - w_n)(-|x_{n+1} - q|^2 + |x_{n+1} - v_n|^2 + |v_n - q|^2) \\ &= -w_n|v_n - q|^2 + (2 - w_n)|x_{n+1} - v_n|^2 + w_n|x_{n+1} - q|^2, \end{aligned} \tag{24}$$

that is

$$|x_{n+1} - q|^2 + 2\lambda_n w_n \langle \nabla f(y_n), y_n - q\rangle = |v_n - q|^2 - \frac{2 - w_n}{w_n}|x_{n+1} - v_n|^2. \tag{25}$$

Moreover, by an easy computation, we obtain

$$\begin{aligned} |v_n - q|^2 &= |(x_n - q) + \theta_n(x_n - x_{n-1})|^2 \\ &= |x_n - q|^2 + 2\theta_n \langle x_n - q, x_n - x_{n-1}\rangle + \theta_n^2 |x_{n-1} - x_n|^2, \end{aligned}$$

hence, by (25), we obtain

$$\begin{aligned}|x_{n+1}-q|^2-|x_n-q|^2+2\lambda_n w_n\langle\nabla f(y_n),y_n-q\rangle\le\\ 2\theta_n\langle x_n-q,x_n-x_{n-1}\rangle+\theta_n^2|x_{n-1}-x_n|^2-\frac{2-w_n}{w_n}|x_{n+1}-v_n|^2.\end{aligned}\tag{26}$$

From (23) we also have

$$\langle x_n-q,x_n-x_{n-1}\rangle=-\frac{1}{2}|x_{n-1}-q|^2+\frac{1}{2}|x_n-q|^2+\frac{1}{2}|x_n-x_{n-1}|^2,$$

which together with the previous inequality amounts to

$$\begin{aligned}&|x_{n+1}-q|^2-|x_n-q|^2+2\lambda_n w_n\langle\nabla f(y_n),y_n-q\rangle\le\theta_n(|x_n-q|^2-|x_{n-1}-q|^2)\\&+(\theta_n^2+\theta_n)|x_{n-1}-x_n|^2-\frac{2-w_n}{w_n}|x_{n+1}-v_n|^2.\end{aligned}\tag{27}$$

Furthermore, by quasiconvexity of f and by $f(q)\le f(y_n)$, we have

$$\langle\nabla f(y_n),y_n-q\rangle\ge 0\quad\forall n\ge 0,\tag{28}$$

then, in light of (27), we obtain

$$\begin{aligned}|x_{n+1}-q|^2-|x_n-q|^2\le\theta_n(|x_n-q|^2-|x_{n-1}-q|^2)\\+(\theta_n+\theta_n^2)|x_{n-1}-x_n|^2-\tfrac{2-w_n}{w_n}|x_{n+1}-v_n|^2,\end{aligned}\tag{29}$$

hence, for $w_n\in(0,2]$ and $\theta_n\in[0,1]$, we deduce

$$|x_{n+1}-q|^2-|x_n-q|^2\le\theta_n(|x_n-q|^2-|x_{n-1}-q|^2)+2\theta_n|x_{n-1}-x_n|^2.\tag{30}$$

Suppose now $\sum_{n\ge 0}\theta_n|x_n-x_{n-1}|^2<\infty$, where $(\theta_n)\subset[0,\theta]$ and $\theta\in[0,1)$. Then, by (30) and applying Proposition 2.4, we deduce that $(|x_n-q|)$ is convergent (hence (x_n) is bounded). Again from (30) and Proposition 2.4, we obtain $\sum_{n\ge 0}[|x_n-q|^2-|x_{n-1}-q|^2]_+<\infty$, while from (27) we have

$$\begin{aligned}&\frac{2-w_n}{w_n}|x_{n+1}-v_n|^2+2\lambda_n w_n\langle\nabla f(y_n),y_n-q\rangle\\&\quad\le|x_n-q|^2-|x_{n+1}-q|^2+\theta_n[|x_n-q|^2-|x_{n-1}-q|^2]_++2\theta_n|x_{n-1}-x_n|^2,\end{aligned}$$

hence we obviously obtain

$$\sum_{n\ge 0}\frac{2-w_n}{w_n}|x_{n+1}-v_n|^2<\infty\quad\text{and}\quad\sum_{n\ge 0}\lambda_n w_n\langle\nabla f(y_n),y_n-q\rangle<\infty.$$

Clearly, if $0<\inf_n w_n\le\sup_n w_n<2$, we deduce $\sum_{n\ge 0}|x_{n+1}-v_n|^2<\infty$. Concerning this series, we immediately have

$$\begin{aligned}|v_n-x_{n+1}|^2&=|(x_n-x_{n+1})+\theta_n(x_n-x_{n-1})|^2\\&=|x_n-x_{n+1}|^2+2\theta_n\langle x_n-x_{n+1},x_n-x_{n-1}\rangle+\theta_n^2|x_n-x_{n-1}|^2,\end{aligned}\tag{31}$$

while by Young's inequality we have

$$\begin{aligned}2\theta_n\langle x_n-x_{n+1},x_n-x_{n-1}\rangle&=\langle x_n-x_{n+1},(2\theta_n)(x_n-x_{n-1})\rangle\\&\ge-\tfrac{1}{2}(2\theta_n)^2|x_n-x_{n-1}|^2-\tfrac{1}{2}|x_n-x_{n+1}|^2\\&=-2\theta_n^2|x_n-x_{n-1}|^2-\tfrac{1}{2}|x_n-x_{n+1}|^2.\end{aligned}$$

Consequently, combining this last inequality with (31), we get

$$|v_n - x_{n+1}|^2 \geq \frac{1}{2}|x_n - x_{n+1}|^2 - \theta_n^2|x_n - x_{n-1}|^2, \tag{32}$$

hence, as $\theta_n^2 \leq \theta_n$, we obtain

$$|x_n - x_{n+1}|^2 \leq 2|v_n - x_{n+1}|^2 + 2\theta_n|x_n - x_{n-1}|^2,$$

which amounts to $\sum_{n\geq 0}|x_n - x_{n+1}|^2 < \infty$. Moreover, by (8b), we recall that $x_{n+1} = (1-w_n)v_n + w_n y_n$, so that $x_{n+1} - x_n = (1-w_n)(v_n - x_n) + w_n(y_n - x_n)$, hence

$$\begin{aligned}
w_n^2|y_n - x_n|^2 &= |(x_{n+1} - x_n) - (1-w_n)(v_n - x_n)|^2 \\
&= |(x_{n+1} - x_n) - \theta_n(1-w_n)(x_n - x_{n-1})|^2 \\
&= |x_{n+1} - x_n|^2 + 2\theta_n(1-w_n)\langle x_{n+1} - x_n, x_{n-1} - x_n\rangle \\
&\quad + \theta_n^2(1-w_n)^2|x_n - x_{n-1}|^2,
\end{aligned}$$

then, again from Young's inequality, we obtain

$$w_n^2|y_n - x_n|^2 \leq 2|x_{n+1} - x_n|^2 + 2\theta_n^2(1-w_n)^2|x_n - x_{n-1}|^2.$$

Consequently, reminding that $\sum_{n\geq 0}|x_{n+1} - x_n|^2 < \infty$ and $\inf_n w_n > 0$, we deduce $\sum_{n\geq 0}|y_n - x_n|^2 < \infty$, which completes the proof. □

In the next lemma, we state some preliminary convergence results, given in a general setting in view of forthcoming developments.

Lremma 3.2 *Let (x_n) and (y_n) given by (8) under (12)–(17) and assume $U = \{z \in \mathcal{H},\quad f(z) \leq \inf_n f(y_n)\} \neq \emptyset$. Then the following results hold:*

- *(j1) The sequence $(f(y_n))$ is convergent.*
- *(j2) If $\lim_{n\to\infty} f(y_n) = \inf_{q\in U} f(q)$, then (x_n) converges weakly to some element x_* in $\mathcal{H}$ verifying $f(x_*) = \inf_{q\in U} f(q)$.*
- *(j3) If $\lim_{n\to\infty} f(y_n) > \inf_{q\in U} f(q)$, then (x_n) converges strongly to some element x_* in $\mathcal{H}$.*
- *(j4) If (x_n) converges weakly to some element x_* and if $\limsup_{n\to\infty} \lambda_n = \infty$, then $f(x_*) = \inf_{q\in U} f(q)$.*

Proof Let us prove (j1). From (8a), we obviously have

$$(\forall n \geq 0), \quad (\forall x \in \mathcal{H}), \quad f(y_n) + \frac{1}{2\lambda_n}|y_n - v_n|^2 \leq f(x) + \frac{1}{2\lambda_n}|x - v_n|^2. \tag{33}$$

Then, setting $x = y_{n-1}$ in (33) and by (15) (i.e, $\lambda_n \geq \lambda > 0$), we get

$$f(y_n) - f(y_{n-1}) \leq \frac{1}{2\lambda}|y_{n-1} - v_n|^2. \tag{34}$$

Moreover, regarding the term in the right-hand side of the above inequality, we have

$$\begin{aligned}
|v_n - y_{n-1}|^2 &= |(x_n - y_{n-1}) + \theta_n(x_n - x_{n-1})|^2 \\
&= |(1+\theta_n)(x_n - x_{n-1}) + (x_{n-1} - y_{n-1})|^2 \\
&= (1+\theta_n)^2|x_n - x_{n-1}|^2 + |x_{n-1} - y_{n-1}|^2 \\
&\quad + 2(1+\theta_n)\langle x_n - x_{n-1}, x_{n-1} - y_{n-1}\rangle,
\end{aligned}$$

hence, by Young's inequality, we get

$$|v_n - y_{n-1}|^2 \leq 2(1+\theta_n)^2|x_n - x_{n-1}|^2 + 2|x_{n-1} - y_{n-1}|^2, \tag{35}$$

which by $\sum_{n\geq 0}|x_n - x_{n-1}|^2 < \infty$ and $\sum_{n\geq 0}|y_n - x_n|^2 < \infty$ (from Lemma 3.1 (i2)) yields

$$\sum_{n\geq 0}|v_n - y_{n-1}|^2 < \infty. \tag{36}$$

In light of (34) and (36), we easily deduce that $(f(y_n))$ is convergent, because this latter sequence is also bounded from below (as U is assumed to be nonempty).

Now we prove (j2). Assume there holds

$$\lim_{n\to\infty} f(y_n) = \inf_{z\in U} f(z). \tag{37}$$

Given any element q in U, we know by Lemma 3.1 that $(|x_n - q|)$ is convergent, so that (x_n) is a bounded sequence. Let us prove that any weak accumulation point of (x_n) lies in U. To this aim, we consider a subsequence (x_{n_k}) of (x_n) such that (x_{n_k}) converges weakly to some u in $\mathcal{H}$ (as $k \to \infty$). It is then immediate that (y_{n_k}) also converges weakly to u, because $|y_n - x_n| \to 0$ (as $\sum_{n\geq 0}|x_n - y_n|^2 < \infty$ by Lemma 3.1). In addition, f being assumed to be quasiconvex and continuous, we recall that f is weakly lower semi-continuous, which by (37) amounts to $f(u) \leq \liminf_{k\to\infty} f(y_{n_k}) = \liminf_{n\to\infty} f(y_n) = \inf_{z\in U} f(z)$. Therefore, using the definition of U, we immediately obtain

$$f(u) \leq \inf_{z\in U} f(z) \leq \inf_{n} f(y_n), \tag{38}$$

so that $u \in U$. Also note that $u \in U$ and (38) shows that $f(u) = \inf_{z\in U} f(z)$. Now, applying Proposition 2.2, we conclude that (x_n) converges weakly to some element x_* in U, which proves (j2).

Let us prove (j3). Assume there holds

$$\lim_{n\to\infty} f(y_n) > \inf_{z\in U} f(z). \tag{39}$$

In that case, it is a simple matter to see that there exists $q \in U$ verifying $\lim_{n\to\infty} f(y_n) > f(q)$. Hence, from continuity of f (at q), it is immediate that exist a nonempty closed ball $\Omega := B(q, \rho)$ (closed ball of center q and radius ρ) and some integer n_0 such that

$$(\forall n \geq n_0), \quad (\forall z \in \Omega), \quad f(z) \leq f(y_n). \tag{40}$$

Let $z_n = q + (\rho/|\nabla f(y_n)|)\nabla f(y_n)$ (provided that $|\nabla f(y_n)| \neq 0$). Clearly, z_n belongs to Ω (since $|z_n - q| = \rho$), while by an easy computation we have

$$\begin{aligned} |\nabla f(y_n)| &= \langle \nabla f(y_n), \tfrac{\rho}{|\nabla f(y_n)|}\nabla f(y_n)\rangle \\ &= \langle \nabla f(y_n), z_n - q\rangle \\ &= \langle \nabla f(y_n), z_n - y_n\rangle + \langle \nabla f(y_n), y_n - q\rangle, \end{aligned}$$

then, as $\langle \nabla f(y_n), z_n - y_n\rangle \leq 0$ (by quasiconvexity of f and by $f(z_n) \leq f(y_n)$), we get $|\nabla f(y_n)| \leq \langle \nabla f(y_n), y_n - q\rangle$. It is obvious that the latter inequality remains valid even if $\nabla f(y_n) = 0$, which by Lemma 3.1 yields

$$\sum_{n\geq 0}\lambda_n w_n |\nabla f(y_n)| < \infty. \tag{41}$$

Moreover, by (20), we obtain

$$x_{n+1} - x_n = \theta_n(x_n - x_{n-1}) - \lambda_n w_n \nabla f(y_n), \tag{42}$$

hence, setting $\delta_n = \lambda_n w_n |\nabla f(y_n)|$, we obviously get

$$|x_{n+1} - x_n| \leq \theta_n |x_n - x_{n-1}| + \delta_n. \tag{43}$$

Then, applying Proposition 2.3, we deduce $\sum_{n\geq 0} |x_{n+1} - x_n| < \infty$, so that (x_n) converges strongly to some element x_* in $\mathcal{H}$.

It remains to prove (j4). Suppose (x_n) converges weakly to some element x_* in $\mathcal{H}$. Again by Lemma 3.1 we have $|x_n - y_n| \to 0$ (as $\sum_{n\geq 1} |x_n - y_n|^2 < \infty$), hence (y_n) converges weakly to x_*. Moreover, by (33) and taking any $q \in U$, we obviously get $f(y_n) \leq f(q) + \frac{1}{2\lambda_n}|q - v_n|^2$. Passing to the limit in this last inequality, using the weak lower semicontinuity of f, and observing that $\liminf_{n\to\infty} \frac{1}{2\lambda_n}|q - v_n|^2 = 0$ (as $\limsup_{n\to\infty} \lambda_n = \infty$ and (v_n) is bounded), we deduce $f(x_*) \leq \liminf_{n\to\infty} f(y_n) \leq f(q)$, hence, by definition of U, we get

$$f(x_*) \leq \inf_{q\in U} f(q) \leq \inf_n f(y_n). \tag{44}$$

Then $x_* \in S$, which together with (44) entails $f(x_*) = \inf_{q\in U} f(q)$, and the proof is completed. □

Now we are in position to state the first main result of this section.

Theorem 3.1 *Let (x_n) and (y_n) be the sequences given by (8) under conditions (12)–(17). Then it holds that:*

- *(r1) If $S = \operatorname{argmin}_{\mathcal{H}} f \neq \emptyset$, then (x_n) converges weakly to some element x_* in $\mathcal{H}$ satisfying $\nabla f(x_*) = 0$. Moreover, $(f(y_n))$ is convergent, besides $\sum_n |x_n - x_{n-1}|^2 < \infty$ and $\sum_n |x_n - y_n|^2 < \infty$.*
  If in addition $x_* \notin S$, then (x_n) converges strongly to x_*.
- *(r2) If $S = \operatorname{argmin}_{\mathcal{H}} f \neq \emptyset$ and $\limsup_{n\to\infty} \lambda_n = \infty$, then the weak limit x_* attained by (x_n) belongs to S.*
- *(r3) If the set $\Gamma = \{x \in \mathcal{H};\ \ \nabla f(x) = 0\}$ of critical points of f is empty, then (x_n) is unbounded.*

Proof Let us prove (r1). Assume $S \neq \emptyset$, hence we immediately have $U := \{z \in \mathcal{H};\ \ f(z) \leq \inf_n f(y_n)\} \neq \emptyset$, because $S \subset U$. Then the sequence $(f(y_n))$ is convergent (by Lemma 3.2) and the two given estimates hold (from Lemmas 3.1 and 3.2). Furthermore, taking $\overline{x} \in S$, we note that $\inf_{q\in U} f(q) = f(\overline{x})$. It is also clear that the limit of $(f(y_n))$ satisfies one of the following two items:

(i) (Case 1) $\lim_{n\to\infty} f(y_n) = f(\overline{x})$; (ii) (Case 2) $\lim_{n\to\infty} f(y_n) > f(\overline{x})$.

Consequently, again from Lemma 3.2, we deduce that (x_n) converges weakly in Case 1 and strongly in Case 2. This establishes the weak convergence of (x_n) to some x_* in $\mathcal{H}$. Let us now focus on the weak limit point of (x_n). Note that, in Case 1, the limit x_* of (x_n) belongs to S (by Lemma 3.2), hence we additionally have $\nabla f(x_*) = 0$. Concerning the strong limit x_* of (x_n) in Case 2, by Lemma 3.1, we have $|x_n - y_n| \to 0$ (since $\sum_n |x_n - y_n|^2 < \infty$), hence (y_n) converges strongly to x_*. Also recall that $|x_{n+1} - x_n| \to 0$, while $(\lambda_n w_n)$ is bounded away from zero (thanks to (14) and (15)). Then, passing to the limit in (42) and using the continuity of ∇f, we deduce that $\nabla f(x_*) = 0$. Thus, in Case 1 and Case 2, the limit (x_*) of (x_n) satisfies $\nabla f(x_*) = 0$, namely x_* is a critical point of f. Also recall that, in Case 1, we obtain $x_* \in S$. Therefore, in any situation when $x_* \notin S$, we are concerned with Case 2, in which strong convergence of (x_n) to x_* holds. This completes the proof of (r1). (r2) is a straightforward consequence of Lemma 3.2 (j4). Now let us prove (r3), namely (x_n) is unbounded under the two conditions that Γ is empty and $\limsup_{n\to\infty} \lambda_n = \infty$. Otherwise, (x_n) is bounded (hence so is (v_n) occurring in (8)), so that (y_n) is bounded, because $y_n = (1/w_n)(x_{n+1} - (1 - w_n)v_n)$. Set $m_* = \liminf_{n\to\infty} f(y_n)$ (possibly $m_* = -\infty$) and let (y_{n_k}) be a subsequence of (y_n) such that $\liminf_{k\to\infty} f(y_{n_k}) = m_*$. Clearly, as (y_n) is bounded, there also exists a subsequence (y_{m_k}) of (y_n) such that $\liminf_{k\to\infty} f(y_{m_k}) = m_*$ and such that (y_{m_k}) converges weakly (as $k \to \infty$) to some element u in $\mathcal{H}$. By weak lower semicontinuity of f (assumed to be quasiconvex and continuous), we then have $f(u) \le \lim_{k\to\infty} f(y_{m_k}) = \liminf_{n\to\infty} f(y_n)$. It is then immediate that there exists q in $\mathcal{H}$ such that $f(q) \le \inf_n f(y_n)$, so that $U := \{z \in \mathcal{H};\quad f(z) \le \inf_n f(y_n)\} \neq \emptyset$. In light of Lemma 3.2 (j1), we then deduce that the sequence $(f(y_n))$ is convergent, hence, we easily obtain $f(u) \le \lim_{n\to\infty} f(y_n) = m_*$. Now, take any element $q \in U$. Since q is not a global minimizer of f (as the set Γ is assumed to be nonempty), we can also exhibit a minimizing sequence $(q_n) \subset \mathcal{H}$ of f, namely $\lim_{n\to\infty} f(q_n) = \inf_{\mathcal{H}} f$. Clearly, for $n \ge n_0$ (where n_0 is some large enough integer), we have $f(q_n) \le f(q)$, hence $(q_n)_{n\ge n_0} \subset U$, so that $\inf_{\mathcal{H}} f \le \inf_{q\in U} f(q) \le \lim_{n\to\infty} f(q_n) = \inf_{\mathcal{H}} f$, which leads to $\inf_{q\in U} f(q) = \inf_{\mathcal{H}} f$. Hence, as u is not a global minimizer of f, we easily obtain

$$\inf_{q\in U} f(q) = \inf_{\mathcal{H}} f < f(u) \le m_* = \lim_{n\to\infty} f(y_n) \tag{45}$$

(hence $\inf_{q\in U} f(q) < \lim_{n\to\infty} f(y_n)$). Consequently, invoking Lemma 3.2 (j3), we deduce that (x_n) converges strongly to some element x_*, and so does (y_n), as $|x_n - y_n| \to 0$ (by Lemma 3.1). Also recall that $|x_{n+1} - x_n| \to 0$ (by Lemma 3.1), while $(\lambda_n w_n)$ is bounded away from zero (by (14) and (15)). Therefore, using (42) together with the continuity of ∇f, we get $\nabla f(x_*) = 0$, namely $x_* \in \Gamma$, which is absurd. This completes the proof. □

Now we claim the second main result of this section.

Theorem 3.2 *Let (x_n) and (y_n) be given by (8) with conditions (12)–(15) and (18). Then (x_n) satisfies the same conclusions as in Theorem 3.1.*

Proof First, we prove that assumptions (16)–(17), when they are replaced by condition (18), are automatically satisfied in the framework of Lemmas 3.1 and 3.2. More precisely, we show that (17) holds under conditions (12)–(15), (18) provided that $U = \{q \in \mathcal{H},\quad f(q) \le \inf_n f(y_n)\} \neq \emptyset$. Set $\delta = \inf_n \frac{2-w_n}{2w_n}$. Clearly, by (14), we have $\delta \in (0, \infty)$, hence, taking any $q \in U$ and using (29), we immediately have

$$|x_{n+1} - q|^2 - |x_n - q|^2 \le \theta_n(|x_n - q|^2 - |x_{n-1} - q|^2) + (\theta_n + \theta_n^2)|x_{n-1} - x_n|^2 \\ - 2\delta|x_{n+1} - v_n|^2.$$

Furthermore (31) is written as

$$|v_n - x_{n+1}|^2 = |x_n - x_{n+1}|^2 + \theta_n^2|x_n - x_{n-1}|^2 + 2\theta_n\langle x_n - x_{n+1}, x_n - x_{n-1}\rangle,$$

which by Young's inequality amounts to

$$|v_n - x_{n+1}|^2 \ge (1 - \theta_n)|x_n - x_{n+1}|^2 + (\theta_n^2 - \theta_n)|x_n - x_{n-1}|^2.$$

Therefore, combining the above inequalities, we obtain

$$|x_{n+1} - q|^2 - |x_n - q|^2 \le \theta_n(|x_n - q|^2 - |x_{n-1} - q|^2) - 2\delta(1 - \theta_n)|x_{n+1} - x_n|^2 \\ +[(\theta_n + \theta_n^2) + 2\delta(\theta_n - \theta_n^2)]|x_{n-1} - x_n|^2.$$

while it is also a simple matter to see that

$$(1/2)(1 + \theta_n) + \delta(1 - \theta_n) \le 2\max\{1/2, \delta\},$$

hence, setting $\mu = 2\max\{1/2, \delta\}$, we deduce

$$|x_{n+1} - q|^2 - |x_n - q|^2 \le \theta_n(|x_n - q|^2 - |x_{n-1} - q|^2) - 2\delta(1 - \theta_n)|x_{n+1} - x_n|^2 \\ + \mu\theta_n|x_{n-1} - x_n|^2.$$

Then the desired estimate (17) can be obtained by following the same lines as in [3]. However, we give full details so that the proof is self-contained. For simplicity's sake, we set $\phi_n = (1/2)|x_n - q|^2$ and $d_n = |x_n - x_{n-1}|^2$, which by the previous inequality leads to

$$\begin{aligned}\phi_{n+1} - \phi_n - \theta_n(\phi_n - \phi_{n-1}) &\le -\delta(1-\theta_n)d_{n+1} + \mu\theta_n d_n \\ &= -\mu\theta_{n+1}d_{n+1} + \mu\theta_n d_n + [\mu\theta_{n+1} - \delta(1-\theta_n)]d_{n+1}.\end{aligned}$$

Then, assuming (θ_n) is nondecreasing, we obtain

$$\phi_{n+1} - \phi_n - (\theta_n\phi_n - \theta_{n-1}\phi_{n-1}) \le -\mu\theta_{n+1}d_{n+1} + \mu\theta_n d_n - [\delta - (\mu+\delta)\theta_{n+1}]d_{n+1}.$$

Consequently, assuming $(\theta_n) \subset [0, \delta_0]$ (where $\delta_0 \in (0, \frac{\delta}{\delta+\mu})$), also setting $\gamma = \delta - \delta_0(\mu+\delta)$ (hence $\gamma > 0$) and $\Gamma_n := \phi_n - \theta_{n-1}\phi_{n-1} + \mu\theta_n d_n$, we get

$$\Gamma_{n+1} - \Gamma_n \le -\gamma d_{n+1}. \tag{46}$$

Thus, (Γ_n) is nonincreasing, so that $\phi_n - \theta\phi_{n-1} \le \Gamma_n \le \Gamma_1$, which entails $\phi_n \le \theta^n\phi_0 + \frac{\Gamma_1}{1-\theta}$. Again with (46), we get $\gamma\sum_{k=1}^n d_{k+1} \le \Gamma_1 - \Gamma_{n+1}$, while $-\Gamma_n \le \theta\phi_{n-1}$, hence $\gamma\sum_{k=1}^n d_{k+1} \le \Gamma_1 + \theta\phi_{n-1} \le \Gamma_1 + \theta^n\phi_0 + \theta\frac{\Gamma_1}{1-\theta}$, which leads to $\sum_{k\ge1} d_{k+1} < \infty$, that is the desired estimate (17). It is then immediate that the conclusions of Lemmas 3.1 and 3.2 hold under conditions (12)–(15), $U \neq \emptyset$ and (18). The conclusions of Theorem 3.2 are therefore obtained as for the proof of Theorem 3.1 $\square$

Remark 3.2 In view of next developments regarding practical implementations of (8), it would be interesting to introduce an additional procedure allowing inexact computations in Step 1:

(1) A first example would be to consider inexact global minimization

$$y_n \in \epsilon_n - \text{argmin}_{\mathcal{H}}\left\{f(x) + (1/2)\lambda_n^{-1}|x - v_n|^2\right\}, \quad \text{where}(\epsilon_n) \subset [0, \infty),$$

under suitable conditions on the error tolerance (ϵ_n).

(2) A second example, even better suited, would be to consider inexact local stationary condition of the form

$$|\nabla f(y_n) + (1/\lambda_n)(y_n - v_n)| \le \epsilon_n, \quad \text{where } (\epsilon_n) \subset [0, \infty). \tag{47}$$

However (47) ensures only that (y_n) is close to a local minimum of $f(.) + (1/2)\lambda_n^{-1}|. - v_n|^2$, but not necessarily near a global minimum. Condition (47) would then be applicable for instance to situation when the local minima of $f(.) + (1/2)\lambda_n^{-1}|. - v_n|^2$ are global ones (see, e.g., [14] for classes of nonconvex objective, or [9] for strictly quasiconvex objective).

Acknowledgments The author would like to thank the anonymous referees for their helpful comments and suggestions on this work.

References

1. Alvarez, F.: On the minimizing property of a second order dissipative dynamical system in Hilbert spaces. SIAM J. Control Optim. **39**, 1102–1119 (2000)
2. Alvarez, F.: Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space. SIAM J. Optim. **14**(3), 773–782 (2004)
3. Alvarez, F., Attouch, H.: An inertial proximal method for monotone operators via discretization of a nonlinear oscillator with damping. Set Valued Anal. **9**, 3–11 (2001)

4. Attouch, H., Soubeyran, A.: Inertia and reactivity in decision making as cognitive variational inequalities. J. Convex Anal. **13**(2), 207–224 (2006)
5. Attouch, H., Teboulle, M.: Regularized Lotka-Volterra dynamical system as continuous proximal-like method in optimization. J. Optim. Theory Appl. **121**, 541–580 (2004)
6. Attouch, H., Goudon, X., Redont, P.: The heavy ball with friction I The continuous dynamical system. Commun. Contemp. Math. **2**(N1), 1–34 (2000)
7. Aussel, D.: Subdifferential properties of quasiconvex and pseudoconvex function, unified approach. J. Optim. Theor. Appl. **97**, 29–45 (1998)
8. Aussel, D.: Contributions en analyse multivoque et en optimization. Ph.D. thesis, Université Monpellier 2 (2005)
9. Crouzeix, J.P.: Characterizations of generalized convexity and generalized monotonicity: a survey. Nonconvex Optim. Appl. **27**, 237–256 (1998)
10. Goudou, X., Munier, J.: The gradient and heavy ball with friction dynamical systems: the quasiconvex case. Math. Progam. Ser. B. doi:10.1007/s10107-007-0109-5 (2007)
11. Gromicho, J.: Quasiconvex optimization and location theory. Kluwer Academic Publishers, Dordrecht, Netherlands (1998)
12. Guler, O.: On the convergence of the proximal point algorithm for convex minimization. SIAM J. Control **29**(2), 403–419 (1991)
13. Jules, F., Maingé, P.E.: Numerical approaches to a stationnary solution of a second order dissipative dynamical system. Optimization **51**(2), 235–255 (2002)
14. Kaplan, A., Tichatschke, R.: Proximal point methods and nonconvex optimization. J. Global Optim. **13**, 389–406 (1998)
15. Kiwiel, K.C.: Convergence and efficiency of subgradient methods for quasiconvex minimization. Math. Prog. Ser. A. **90**, 1–25 (2001)
16. Kiwiel, K.C., Murty, K.: Convergence of the steepest descent method for minimization of quasiconvex functions. J. Optim. Theory Appl. **89**(1), 221–226 (1996)
17. Martinet, B.: Perturbation des méthodes d'optimisation. Appl. R.A.I.R.O Numer. Anal. **12**, 154–171 (1978)
18. Maingé, P.E.: Inertial iterative process for fixed points of certain quasi-nonexpansive mappings. Set Valued Anal. **15**(1), 67–79 (2007)
19. Moudafi, A., Oliny, M.: Convergence of a splitting inertial proximal method for monotone operators. J. Comput. Appl. Math. **155**, 447–454 (2003)
20. Opial, Z.: Weak convergence of the sequence of successive approximations for nonexpansive mappings. Bull. Am. Math. Soc. **73**, 591–597 (1967)
21. Pan, S, Chen, J.S.: Entropy-like proximal algorithms based on a second order homogeneous distance function for quasi-convex programming. J. Glob. Optim. doi:10.1007/s10898-007 (2007)
22. Papa Quiroz, E.A., Roberto Oliviera, P.: Classic and Logarithmic-Quadratic proximal point method for quasiconvex minimization. Optimization Online (2006)
23. Rockafellar, R.T.: Augmented Lagrangians and applications of proximal point algorithm in convex programming. Math. Oper. Res. **1**, 97–116 (1976)
24. Rockafellar, R.T.: Monotone operators and the proximal point algorithm. SIAM J. Control Optim. **14**, 877–898 (1976)
25. Solodov, M.V., Svaiter, B.F.: A unified framework for some inexact proximal point algorithms. Numer. Funct. Anal. Optim. **22**, 1013–1035 (2001)
26. Takayama, A.: Mathematical Economics, 2nd edn. Cambridge University Press (1995)