SIAM J. CONTROL OPTIM.
Vol. 38, No. 2, pp. 538–565

STRONG CONVERGENCE OF BLOCK-ITERATIVE OUTER APPROXIMATION METHODS FOR CONVEX OPTIMIZATION*

PATRICK L. COMBETTES†

Abstract. The strong convergence of a broad class of outer approximation methods for minimizing a convex function over the intersection of an arbitrary number of convex sets in a reflexive Banach space is studied in a unified framework. The generic outer approximation algorithm under investigation proceeds by successive minimizations over the intersection of convex supersets of the feasibility set determined in terms of the current iterate and variable blocks of constraints. The convergence analysis involves flexible constraint approximation and aggregation techniques as well as relatively mild assumptions on the constituents of the problem. Various well-known schemes are recovered as special realizations of the generic algorithm and parallel block-iterative extensions of these schemes are devised within the proposed framework. The case of inconsistent constraints is also considered.

Key words. block-iterative, convex feasibility problem, convex programming, constrained minimization, cutting plane, fixed point, inconsistent constraints, outer approximation, projection onto an intersection of convex sets, reflexive Banach space, surrogate cut, uniformly convex function

AMS subject classifications. 49M27, 65J05, 65K05, 90C25

PII. S036301299732626X

1. Introduction.

Let $\mathcal{X}$ be a real reflexive Banach space, let $J\colon \mathcal{X} \to]-\infty, +\infty]$ be a proper function, and let $(S_i)_{i\in I}$ be an arbitrary family of closed convex subsets of $\mathcal{X}$. We investigate a broad class of block-iterative outer approximation methods for solving the program

$$\text{(P)}\qquad \text{find}\,\overline{x} \in S \triangleq \bigcap_{i\in I} S_i \quad \text{such that } J(\overline{x}) = \inf_{x\in S} J(x) \triangleq \overline{J}$$

under the following assumptions:

(A1) J is lower semicontinuous and convex.

(A2) For some closed convex set $E \supset S$, there exists a point $u \in S \cap \operatorname{dom} J$ such that the set $C \triangleq \{x \in E \mid J(x) \leq J(u)\}$ is bounded and J is uniformly convex with modulus of convexity c on C, i.e., [53], [54]

$$(1.1)\qquad \big(\forall (x,y) \in C^2\big)\quad J\left(\frac{x+y}{2}\right) \leq \frac{J(x)+J(y)}{2} - c(\|x-y\|),$$

where $c\colon \mathbb{R}_+ \to \mathbb{R}_+$ is nondecreasing and $(\forall \tau \in \mathbb{R}_+)\ c(\tau) = 0 \Leftrightarrow \tau = 0$.

(A3) For every $i \in I$, $S_i = \{x \in \mathcal{X} \mid g_i(x) \leq 0\}$, where g_i belongs to the class $\mathcal{G}$ of

*Received by the editors September 17, 1997; accepted for publication (in revised form) October 29, 1998; published electronically February 2, 2000. This work was supported by the National Science Foundation under grant MIP-9705504.

http://www.siam.org/journals/sicon/38-2/32626.html

†Laboratoire d'Analyse Numérique, Université Pierre et Marie Curie — Paris 6, 75005 Paris, France and City College and Graduate Center, City University of New York, New York, NY 10031 (plc@ann.jussieu.fr).

all functions $g\colon \mathcal{X} \to]-\infty,+\infty]$ such that

$$\text{(1.2)} \qquad \begin{cases} \text{(i) } \{x \in \mathcal{X} \mid g(x) \leq 0\} \text{ is nonempty and convex.} \\ \text{(ii) For every sequence } (y_n)_{n\geq 0} \subset \mathcal{X} \\ \qquad\qquad \begin{cases} y_n \overset{n}{\rightharpoonup} y \\ \overline{\lim}_n g(y_n) \leq 0 \end{cases} \quad \Rightarrow \quad g(y) \leq 0. \end{cases}$$

Assumptions (A1)–(A2) are rather standard and ensure in particular that (P) admits a unique solution $\overline{x}$ [1]. Assumption (A3) provides an explicit description of the constraint sets $(S_i)_{i\in I}$ as lower level sets of functions $(g_i)_{i\in I} \subset \mathcal{G}$. As will be seen in section 2, the class $\mathcal{G}$ is quite broad, and (A3) therefore covers a wide range of constraints encountered in convex optimization problems.

In the past four decades, various outer approximation methods for constrained minimization problems have been proposed, following their introduction by Cheney and Goldstein [8] and Kelley [31] in the form of cutting plane algorithms. The underlying principle is to replace (P) by a sequence of minimizations over simple closed convex supersets $(Q_n)_{n\geq 0}$ of the feasibility set S. Typically, the approximation at iteration n can be written as $Q_n = D_n \cap H_n$, where D_n and H_n are two closed convex supersets of S, the latter being termed a cut. Outer approximation methods can be divided into two main categories, namely, cutoff methods and constraints disintegration methods. In cutoff method [37], $D_{n+1} = Q_n$ and Q_{n+1} therefore results from the accumulation of all previous cuts. Several classical algorithms fit in this framework that differ in the way the cuts are defined, e.g., [8], [30], [31], [52], [55]. Naturally, a limitation of cutoff methods is that the minimization of J over the sets $(Q_n)_{n\geq 0}$ becomes increasingly demanding in terms of both computational load and storage requirements. This shortcoming prompted the development of filtered cutoff methods in which some of the old cuts can be discarded under various hypotheses, thereby keeping the complexity of the outer approximations manageable, e.g., [5], [16], [19], [49], [50]. These methods are cumulative in the sense that every cut must be retained until it is definitely dropped. By contrast, in the somewhat less well known constraints disintegration methods, D_n is a half-space depending solely on x_n and a subgradient of J at x_n. Such schemes were first proposed by Haugazeau in the 1960s for the minimization of quadratic forms in Hilbert spaces [26] and several variants have since been proposed for this particular problem [14], [27], [41], [44], [45]. The extension to convex functions was dealt with in [35] in Banach spaces and rediscovered in Euclidean spaces in [29] and [39].

The goal of the present work is to develop a general framework for outer approximation methods that captures and extends the above algorithms. Our investigation will not only provide a unified strong convergence analysis of existing outer approximation methods for solving (P) but also yield flexible generalizations of these methods in the form of parallel block-iterative algorithms.

The paper is built around the following generic outer approximation scheme. For brevity, $\mathfrak{m}(A)$ denotes the minimizer of J over a convex set A and $\mathfrak{C}(A)$ the family of all closed convex supersets of A.

ALGORITHM 1.1. *A sequence $(x_n)_{n\geq 0}$ is constructed as follows, where E is supplied by* (A2).

Step 0. *Set $D_0 = E$, $x_0 = \mathfrak{m}(D_0)$, and $n = 0$.*

Step 1. *Take a nonempty finite index set* $I_n \subset I$ *and generate* H_n *such that*

$$H_n \in \mathfrak{C}\left(\bigcap_{i\in I_n} S_i\right). \tag{1.3}$$

Step 2. *Set* $Q_n = E \cap D_n \cap H_n$ *and* $x_{n+1} = \mathfrak{m}(Q_n)$.
Step 3. *Generate* D_{n+1} *such that*

$$D_{n+1} \in \mathfrak{C}(S) \quad \textit{and} \quad x_{n+1} = \mathfrak{m}(D_{n+1}). \tag{1.4}$$

Step 4. *Set* $n = n+1$ *and go to Step* 1.

Associated with this algorithm is the following terminology.

DEFINITION 1.1. *Let* H_n *and* D_{n+1} *be two subsets of* $\mathcal{X}$. *Then* H_n *(respectively,* D_{n+1}*) will be said to be a* cut *(respectively, a* base*) for Algorithm* 1.1 *at iteration* n *if* (1.3) *(respectively,* (1.4)*) holds.*

At iteration n, x_n and an outer approximation D_n to S are given such that x_n minimizes J over D_n. A finite block of sets $(S_i)_{i\in I_n}$ is then selected and H_n is constructed as an outer approximation to their intersection. The update x_{n+1} is the minimizer of J over $Q_n = E \cap D_n \cap H_n$. We observe that, since $u \in Q_n$, x_{n+1} is the minimizer of J over $Q_n \cap C$. Consequently, since J is weakly lower semicontinuous (by (A1)) and since $Q_n \cap C$ is nonempty and weakly compact (bounded, closed, and convex by (A1)–(A2) in the reflexive space $\mathcal{X}$), the existence of x_{n+1} follows from Weierstrass' theorem [1, Thm. 2.1.1]; its uniqueness follows from the strict convexity of J over $Q_n \cap C$, which is secured by (A2). The iteration is completed by generating a new outer approximation D_{n+1} to S over which J achieves its infimum at x_{n+1}.

The remainder of the paper is divided into six sections. In section 2 basic notation and definitions are introduced and assumptions (A1)–(A3) are illustrated through specific examples. In section 3 we establish the strong convergence of Algorithm 1.1 to the solution $\overline{x}$ of (P) for two types of control sequence $(I_n)_{n\geq 0}$ under certain "tightness" conditions. Four frameworks are then considered individually. In section 4, two general cut construction techniques are described, namely, exact-constraint cuts in section 4.1 and surrogate cuts in section 4.2. In the former case, the cuts are drawn directly from the pool of constraint sets $(S_i)_{i\in I}$, whereas in the latter they are constructed as surrogate half-spaces based on approximate projections of the current iterate onto the selected block of sets. Section 5 is devoted to the construction of bases. In section 5.1 the bases are cumulative, as in cutoff methods, whereas in section 5.2 the bases are instantaneous, as in constraints disintegration methods. By coupling a cut construction strategy from section 4.1 or 4.2 with a base construction strategy from section 5.1 or 5.2, we obtain in section 6 four general realizations of the abstract Algorithm 1.1. In each case, strong convergence theorems are given and existing methods are exhibited as special cases. As a by-product, a block-iterative algorithm for projecting onto an intersection of convex sets in a Hilbert space is presented in detail. Finally, problems with inconsistent constraints and feasibility problems are discussed in section 7.

2. Preliminaries.

2.1. Notation, definitions, and basic facts. The definitions and results stated hereafter can be found in [1].

$\mathbb{N}$ is the set of nonnegative integers, $\mathbb{N}^*$ the set of positive integers, $\mathbb{R}_+$ the set of nonnegative reals, $\mathbb{R}_+^*$ the set of positive reals, and $\mathbb{R}^N$ the standard N-dimensional

Euclidean space. $\mathcal{X}$ is a real reflexive Banach space, and Id its identity operator. $\operatorname{bd} A$ denotes the boundary of a set $A \subset \mathcal{X}$, A° its interior, $\mathfrak{m}(A)$ the minimizer of J over A (i.e., $\mathfrak{m}(A) \in A$ and $(\forall x \in A)\ J(\mathfrak{m}(A)) \le J(x)$) provided such a point exists and is unique, and $\mathfrak{C}(A)$ the family of all closed convex supersets of A. The norm of $\mathcal{X}$ and that of its topological dual $\mathcal{X}'$ is denoted by $\|\cdot\|$, the associated distance by d, and the canonical bilinear form on $\mathcal{X} \times \mathcal{X}'$ by $\langle \cdot, \cdot \rangle$. The expressions $x_n \overset{n}{\rightharpoonup} x$ and $x_n \overset{n}{\to} x$ denote, respectively, the weak and strong convergence to x of a sequence $(x_n)_{n\ge0}$ and $\mathfrak{W}(x_n)_{n\ge0}$ its set of weak cluster points. The closed ball of center x and radius γ in $\mathcal{X}$ or $\mathcal{X}'$ is denoted by $B(x,\gamma)$ and the normalized duality mapping of $\mathcal{X}$ by Δ, i.e.,

$$(2.1) \qquad (\forall x \in \mathcal{X}) \quad \Delta(x) = \big\{x' \in \mathcal{X}' \mid \|x\|^2 = \langle x, x'\rangle = \|x'\|^2\big\}.$$

It follows from the reflexivity of $\mathcal{X}$ that Δ is surjective ($\Delta^{-1}(x') \neq \varnothing$ for every $x' \in \mathcal{X}'$). Δ is single valued if $\mathcal{X}'$ is strictly convex.

Let $F\colon \mathcal{X} \to]-\infty,+\infty]$ be a proper function, i.e., $\operatorname{dom} F = \{x \in \mathcal{X} \mid F(x) < +\infty\} \neq \varnothing$. F is subdifferentiable at $x \in \operatorname{dom} F$ if its subdifferential at this point,

$$(2.2) \qquad \partial F(x) = \big\{t' \in \mathcal{X}' \mid \ (\forall y \in \mathcal{X})\ \langle y - x, t'\rangle + F(x) \le F(y)\big\},$$

is not empty. A subgradient of F at x is an element of $\partial F(x)$. The lower level set of F at height $\lambda \in \mathbb{R}$ is $\operatorname{lev}_{\le\lambda} F = \{x \in \mathcal{X} \mid F(x) \le \lambda\}$. F is quasi-convex if its lower level sets $(\operatorname{lev}_{\le\lambda} F)_{\lambda\in\mathbb{R}}$ are convex and it is (respectively, weakly) lower semicontinuous if they are (respectively, weakly) closed. Now suppose that $A \subset \mathcal{X}$ is a nonempty convex set and that F is convex and continuous at a point in $A \cap \operatorname{dom} F$, and let $p \in A$. Then

$$(2.3) \qquad F(p) = \inf_{y\in A} F(y) \quad \Leftrightarrow \quad \big(\exists t' \in \partial F(p)\big)(\forall y \in A) \quad \langle p - y, t'\rangle \le 0.$$

In particular, fix $x \in \mathcal{X}$ and let $F\colon y \mapsto \|x-y\|^2/2$. Then (2.3) yields

$$(2.4) \qquad \|x - p\| = d(x, A) \quad \Leftrightarrow \quad \big(\exists q' \in \Delta(x-p)\big)(\forall y \in A) \quad \langle y - p, q'\rangle \le 0$$

and p is called a projection of x onto A. Such a point exists if A is closed and it is unique if in addition $\mathcal{X}$ is strictly convex, as is the case when $\mathcal{X}$ is uniformly convex, i.e.,

$$(2.5) \quad (\forall \epsilon \in]0,2])(\exists \delta \in]0,2])\big(\forall (x,y) \in B(0,1)^2\big) \quad \|x-y\| \ge \epsilon \Rightarrow \|x+y\| \le 2-\delta,$$

and a fortiori when $\mathcal{X}$ is a Hilbert space.

If $\mathcal{X}$ is a Hilbert space, the identifications $\mathcal{X}' = \mathcal{X}$ and $\Delta = \mathrm{Id}$ will be made and the scalar product of $\mathcal{X}$ will also be denoted by $\langle \cdot, \cdot \rangle$. Thus, expressions such as $\langle x, y'\rangle$, where $(x,y) \in \mathcal{X}^2$ and $y' \in \Delta(y)$, will reduce to $\langle x, y\rangle$.

2.2. On assumptions (A1)–(A3). We first describe basic scenarios covered by assumptions (A1)–(A2). It should be noted at this point that the boundedness of C in (A2) is mentioned only for the sake of clarity and that it is actually implicit. Indeed, if $F\colon \mathcal{B} \to]-\infty,+\infty]$ is lower semicontinuous and uniformly convex on a closed convex set $A \subset \mathcal{B}$, where $\mathcal{B}$ is a reflexive Banach space, then $A \cap \operatorname{lev}_{\le F(w)} F$ is bounded for every $w \in A \cap \operatorname{dom} F$ [53, Thm. 1(1)].

PROPOSITION 2.1. *Assumptions* (A1) *and* (A2) *are satisfied in each of the following cases.*

(i) *J is lower semicontinuous and convex and, for some $E \in \mathfrak{C}(S)$, there exists $u \in S \cap \operatorname{dom} J$ such that $C = E \cap \operatorname{lev}_{\le J(u)} J$ is compact and J is strictly convex and continuous on C.*

(ii) $\mathcal{X} = \mathbb{R}^N$, *$J$ is finite and strictly convex, and either of the following conditions is fulfilled:*
 (a) $E = \mathcal{X}$ *and, for some* $\overline{\lambda} \in \mathbb{R}$, $\operatorname{lev}_{\leq \overline{\lambda}} J$ *is nonempty and bounded.*
 (b) $E \in \mathfrak{C}(S)$ *is bounded.*

(iii) *$\mathcal{X}$ is a Hilbert space, $E = \mathcal{X}$, and J is a coercive quadratic form, i.e.,* $J\colon x \mapsto a(x,x)/2 - \langle x, b\rangle$, *where $b \in \mathcal{X}$ and a is a symmetric bounded bilinear form on $\mathcal{X}^2$ that satisfies*

$$\big(\exists \gamma \in \mathbb{R}_+^*\big)(\forall x \in \mathcal{X}) \quad a(x,x) \geq \gamma \|x\|^2. \tag{2.6}$$

(iv) *$\mathcal{X}$ is uniformly convex, $E = \mathcal{X}$, and* $J\colon x \mapsto \int_0^{\|x-w\|} \varphi(t)dt$, *where $w \in \mathcal{X}$ and* $\varphi\colon \mathbb{R}_+ \to \mathbb{R}_+$ *is increasing.*

(v) *Let $(\Omega, \mathcal{F}, \mu)$ be a complete finite measure space, let $p \in]1, +\infty[$, and let X be a separable, real reflexive Banach space with norm $\|\cdot\|_{\mathsf{X}}$ and Borel σ-algebra $\mathcal{B}$. $\mathcal{X} = \mathbf{L}^p_{\mathsf{X}}$ is the Lebesgue space of (equivalence classes of μ — almost everywhere (a.e.) equal) measurable functions $x\colon (\Omega, \mathcal{F}) \to (\mathsf{X}, \mathcal{B})$ such that $\int_\Omega \|x(\omega)\|^p_{\mathsf{X}} \mu(d\omega) < +\infty$ and $J\colon x \mapsto \int_\Omega \varphi(\omega, x(\omega))\mu(d\omega)$, where the integrand $\varphi\colon \Omega \times \mathsf{X} \to]-\infty, +\infty]$ fulfills the following conditions:*
 (a) *φ is measurable relative to the product σ-algebra $\mathcal{F} \times \mathcal{B}$.*
 (b) $(\forall x \in \mathcal{X})$ $\int_\Omega |\varphi(\omega, x(\omega))|\mu(d\omega) < +\infty$.
 (c) *The functions $(\varphi(\omega, \cdot))_{\omega \in \Omega}$ are lower semicontinuous, proper, and uniformly convex on X with common modulus of convexity c_0, hence*

$$(\forall \omega \in \Omega)(\forall (\mathsf{x}, \mathsf{y}) \in \big(\operatorname{dom} \varphi(\omega, \cdot)\big)^2) \quad \varphi\left(\omega, \frac{\mathsf{x}+\mathsf{y}}{2}\right) \leq \frac{\varphi(\omega, \mathsf{x}) + \varphi(\omega, \mathsf{y})}{2} - c_0(\|\mathsf{x} - \mathsf{y}\|_{\mathsf{X}}). \tag{2.7}$$

Moreover, c_0 is continuous and $\varliminf_{\tau \to +\infty} c_0(\tau)/\tau^p > 0$.

Proof. (i) is a consequence of [36, Thm. 4.1.8.(1)]. (ii) $\Rightarrow$ (i): J is convex and, by [46, Cor. 10.1.1], continuous. Moreover, for any $u \in S$, $E \cap \operatorname{lev}_{\leq J(u)} J$ is compact. This follows from the compactness of $\operatorname{lev}_{\leq \lambda} J$ for any $\lambda \in \mathbb{R}$ in (a) [46, Cor. 8.7.1] and from that of E in (b). (iii): $a(\cdot,\cdot)$ is a scalar product on $\mathcal{X}$ with associated norm $|||\cdot|||\colon x \mapsto \sqrt{a(x,x)}$. The parallelogram identity applied to $|||\cdot|||$ and (2.6) then shows its uniform convexity on $\mathcal{X}$ with modulus of convexity $\alpha \mapsto \gamma\alpha^2/4$. Hence, $|||\cdot|||$ satisfies (A1)–(A2) and so does J. (iv): Without loss of generality, let $w = 0$. The function $\psi\colon \alpha \mapsto \int_0^\alpha \varphi(t)dt$ is well defined, finite, increasing, convex, and continuous on $\mathbb{R}_+$ [46, Thm. 24.2]. Hence, $J = \psi \circ \|\cdot\|$ is convex and continuous, and (A1) is satisfied. Finally, (A2) is satisfied due to the uniform convexity of J on any closed ball [54, Thm. 4.1(ii)] and therefore on $\operatorname{lev}_{\leq J(u)} J$ for any $u \in S$. (v): $\mathcal{X}$ is a reflexive Banach space with norm $\|\cdot\|\colon x \mapsto (\int_\Omega \|x(\omega)\|^p_{\mathsf{X}} \mu(d\omega))^{1/p}$ [20, Thm. 8.20.5]. Moreover, J is finite, continuous, and convex on $\mathcal{X}$ [47, Thm. 22(a)], which gives (A1). As regards (A2), we claim that J is uniformly convex on $\mathcal{X}$. Indeed, take arbitrarily $(x,y) \in \mathcal{X}^2$. Then it follows from (b) that $\varphi(\cdot, x(\cdot)) < +\infty$ and $\varphi(\cdot, y(\cdot)) < +\infty$ μ-a.e. Consequently, by virtue of (2.7), for μ almost every $\omega \in \Omega$, it holds that

$$\varphi\left(\omega, \frac{x(\omega)+y(\omega)}{2}\right) \leq \frac{\varphi(\omega, x(\omega)) + \varphi(\omega, y(\omega))}{2} - c_0\left(\|x(\omega) - y(\omega)\|_{\mathsf{X}}\right), \tag{2.8}$$

where, under our assumptions, the function $\omega \mapsto c_0(\|x(\omega) - y(\omega)\|_{\mathsf{X}})$ is measurable. Upon integrating (2.8), we obtain

$$J\left(\frac{x+y}{2}\right) \leq \frac{J(x)+J(y)}{2} - \int_\Omega c_0\left(\|x(\omega) - y(\omega)\|_{\mathsf{X}}\right)\mu(d\omega). \tag{2.9}$$

Now fix $\varepsilon \in \mathbb{R}_+^*$ arbitrarily. Then, since $\mu(\Omega) < +\infty$ and $\underline{\lim}_{\,\tau\to+\infty} c_0(\tau)/\tau^p > 0$, it follows from [54, Lem. 4.4] that there exists $c \in \mathbb{R}_+^*$ depending only on ε such that $\int_\Omega c_0(\|x(\omega) - y(\omega)\|_{\mathsf{X}})\mu(d\omega) \geq c$ whenever $\|x - y\| \geq \varepsilon$. This proves the claim. $\square$

Scenario (ii) is an important practical instance of (i) in which (P) takes the form of a semi-infinite convex program, as commonly found in numerical applications. In scenario (iii), since there exists $w \in \mathcal{X}$ such that $(\forall x \in \mathcal{X})\langle x, b\rangle = a(x, w)$ [7, Chap. V], we can write $J\colon x \mapsto a(x-w, x-w)/2 - a(w, w)/2$. (P) can therefore be looked upon as the problem of finding the projection of w onto the intersection of the closed convex sets $(S_i)_{i\in I}$ relative to the norm $|||\cdot|||\colon x \mapsto \sqrt{a(x,x)}$. Alternatively, since for every $y \in \mathcal{X}$ $\langle y, \nabla J(\overline{x})\rangle = a(y, \overline{x}) - \langle y, b\rangle$ [15, Chap. VII], (2.3) shows that (P) is equivalent to solving the variational inequality

$$\text{find}\, \overline{x} \in S = \bigcap_{i\in I} S_i \quad \text{such that } (\forall x \in S) \quad a(x - \overline{x}, \overline{x}) \geq \langle x - \overline{x}, b\rangle, \tag{2.10}$$

which arises in numerous areas of mathematical sciences [1], [7], [15]. Next, scenario (iv) describes the problem of projecting w onto the intersection of the closed convex sets $(S_i)_{i\in I}$ in a uniformly convex Banach space. It is noted that if $J\colon x \mapsto \|x-w\|^2/2$, then $\partial J\colon x \mapsto \Delta(x - w)$ [1]. Finally, scenario (v) is of interdisciplinary interest and covers problems in areas such as stochastic programming, economics, and control theory; see, e.g., [1], [43], [47]. It should be added that $t' \in \partial J(x) \Leftrightarrow t'(\cdot) \in \partial\varphi(\cdot, x(\cdot))$ μ — a.e. [47, Thm. 22(c)] and that $\mathcal{X}$ is a Hilbert space if X is a Hilbert space and $p = 2$.

We now turn to assumption (A3). The motivation for introducing the class of functions $\mathcal{G}$ stems from its ability to capture in the convenient form of functional inequalities a wide range of convex constraints arising in theoretical and practical optimization problems. As illustrated below, constraint sets in the form of lower level sets of quasi-convex functions or of fixed point sets of quasi-nonexpansive operators, as found for instance in [4], [10], [11], [12], [29], [34], [35], and [51], are included. Let us also call attention to the fact that (1.2)(ii) implies that $\mathrm{lev}_{\leq 0}\, g$ is weakly closed for every $g \in \mathcal{G}$.

Proposition 2.2. *Let $g\colon \mathcal{X} \to]-\infty, +\infty]$ be a function such that, for some $w \in \mathcal{X}$, $g(w) \leq 0$. Then $g \in \mathcal{G}$ if one of the conditions below is fulfilled.*

(i) $\mathrm{lev}_{\leq 0}\, g$ *is convex and g is weakly lower semicontinuous.*

(ii) *g is lower semicontinuous and quasi-convex.*

(iii) $\mathrm{lev}_{\leq 0}\, g$ *is closed and convex and the constraint "$g(x) \leq 0$" is correct* [37]*:*

$$\big(\forall (y_n)_{n\geq 0} \subset \mathcal{X}\big)\ \overline{\lim}_{\,n}\, g(y_n) \leq 0 \Rightarrow \underline{\lim}_{\,n}\, d(y_n, \mathrm{lev}_{\leq 0}\, g) = 0. \tag{2.11}$$

(iv) *$g\colon x \mapsto \|Tx - x\|$ is the displacement function of an operator $T\colon \mathcal{X} \to \mathcal{X}$ whose fixed point set* $\mathrm{Fix}\, T \triangleq \{x \in \mathcal{X} \mid Tx = x\}$ *is convex and such that* $T - \mathrm{Id}$ *is demiclosed at the origin:*

$$\big(\forall (y_n)_{n\geq 0} \subset \mathcal{X}\big) \quad \begin{cases} y_n \overset{n}{\rightharpoonup} y \\ Ty_n - y_n \overset{n}{\to} 0 \end{cases} \quad \Rightarrow \quad y \in \mathrm{Fix}\, T. \tag{2.12}$$

These conditions are fulfilled in each of the following cases.

(a) $\mathcal{X}$ *is uniformly convex and* T *is nonexpansive:* $(\forall(x,y)\in\mathcal{X}^2)$ $\|Tx-Ty\| \le \|x-y\|$.

(b) $\operatorname{Fix} T$ *is closed and convex and, for every sequence* $(y_n)_{n\ge 0}\subset\mathcal{X}$, $Ty_n - y_n \xrightarrow{n} 0 \Rightarrow \underline{\lim}_n d(y_n,\operatorname{Fix} T)=0$.

(c) $T-\mathrm{Id}$ *is demiclosed at the origin and there exists* $\eta\in\mathbb{R}_+^*$ *such that*

$$(\forall x\in\mathcal{X})\big(\exists z'\in\Delta(x-Tx)\big)(\forall y\in\operatorname{Fix} T)\quad \langle x-y,z'\rangle\ge\eta\|Tx-x\|^2. \tag{2.13}$$

(d) $\mathcal{X}$ *is a Hilbert space,* $T-\mathrm{Id}$ *is demiclosed at the origin, and* T *is quasi-nonexpansive:*

$$\big(\forall(x,y)\in\mathcal{X}\times\operatorname{Fix} T\big)\quad \|Tx-y\|\le\|x-y\|. \tag{2.14}$$

(e) $\mathcal{X}$ *is a Hilbert space and* T *is firmly nonexpansive:*

$$\big(\forall(x,y)\in\mathcal{X}^2\big)\quad \|Tx-Ty\|^2\le\|x-y\|^2-\|(T-\mathrm{Id})x-(T-\mathrm{Id})y\|^2. \tag{2.15}$$

Proof. (i) Since g is weakly lower semicontinuous, $y_n \overset{n}{\rightharpoonup} y \Rightarrow g(y)\le\overline{\lim}_n g(y_n)$. Hence (1.2) (ii) holds. (ii) $\Rightarrow$ (i) is immediate. (iii) Since $\operatorname{lev}_{\le 0} g$ is convex, $d(\cdot,\operatorname{lev}_{\le 0} g)$ is convex and Lipschitzian and therefore weakly lower semicontinuous. Accordingly,

$$y_n \overset{n}{\rightharpoonup} y \quad\Rightarrow\quad d(y,\operatorname{lev}_{\le 0} g)\le\underline{\lim}_n d(y_n,\operatorname{lev}_{\le 0} g). \tag{2.16}$$

Hence, if we further assume $\overline{\lim}_n g(y_n)\le 0$, (2.11) gives $d(y,\operatorname{lev}_{\le 0} g)=0$; i.e., $g(y)\le 0$ since $\operatorname{lev}_{\le 0} g$ is closed. (iv) is immediate. (a) is proved in [24, Lem. 3.4 and Thm. 8.4]. (b) follows from (iii). (c): Let

$$(\forall x\in\mathcal{X})\quad Q_x=\big\{y\in\mathcal{X}\mid\langle x-y,z'\rangle\ge\eta\|Tx-x\|^2\big\} \tag{2.17}$$

(z' being as in (2.13)) and $Q=\bigcap_{x\in\mathcal{X}}Q_x$. Then Q is convex as an intersection of half-spaces. Let us show $\operatorname{Fix} T=Q$. $\operatorname{Fix} T\subset Q$ results at once from (2.13). Conversely, let $x\in Q$. Then $x\in Q_x$ and therefore $0\ge\eta\|Tx-x\|^2$. Thus, $Tx=x$ and, in turn, $Q\subset\operatorname{Fix} T$. (d) $\Rightarrow$ (c): In Hilbert spaces, (2.13) becomes

$$\big(\forall(x,y)\in\mathcal{X}\times\operatorname{Fix} T\big)\quad\langle x-y,x-Tx\rangle\ge\eta\|Tx-x\|^2. \tag{2.18}$$

The identity $2\langle x-y,x-Tx\rangle=\|Tx-x\|^2+\|x-y\|^2-\|Tx-y\|^2$ shows that (2.18) is equivalent to

$$\big(\forall(x,y)\in\mathcal{X}\times\operatorname{Fix} T\big)\quad\|Tx-y\|^2\le\|x-y\|^2-(2\eta-1)\|Tx-x\|^2, \tag{2.19}$$

which reduces to (2.14) for $\eta=1/2$. (e) $\Rightarrow$ (c): T is nonexpansive and $T-\mathrm{Id}$ is therefore demiclosed by (a). In addition, (2.15) $\Rightarrow$ (2.19) with $\eta=1$. ∎

3. Convergence analysis. This section is devoted to establishing the strong convergence of Algorithm 1.1 under suitable conditions. Our starting point is the following proposition, which collects some basic properties of the algorithm.

PROPOSITION 3.1. *Let* $(x_n)_{n\ge 0}$ *be an arbitrary orbit of Algorithm* 1.1. *Then:*

(i) $(\forall n\in\mathbb{N})$ $J(x_n)\le J(x_{n+1})\le\overline{J}$.

(ii) $(x_n)_{n\ge 0}\subset C$.

(iii) $\mathfrak{W}(x_n)_{n\ge 0}\neq\varnothing$.

(iv) $(J(x_n))_{n\geq 0}$ *converges and* $\lim_n J(x_n) \leq \overline{J}$.
(v) $(\exists n \in \mathbb{N})\ x_n \in S \quad \Rightarrow \quad (\forall k \in \mathbb{N})\ x_{n+k} = \overline{x}$.
(vi) $\mathfrak{W}(x_n)_{n\geq 0} \subset S \quad \Rightarrow \quad x_n \overset{n}{\to} \overline{x}$.
(vii) $x_{n+1} - x_n \overset{n}{\to} 0$.
(viii) $d(x_n, H_n) \overset{n}{\to} 0$.

Proof. (i) results from the inclusions $(\forall n \in \mathbb{N}) D_n \supset Q_n \supset S$. (ii), (iv), and (v) follow from (i). (ii) $\Rightarrow$ (iii): It follows from (A1)–(A2) and the reflexivity of $\mathcal{X}$ that C is weakly compact. (vi): Assume $\mathfrak{W}(x_n)_{n\geq 0} \subset S$ and take $x \in \mathfrak{W}(x_n)_{n\geq 0}$, say $x_{n_k} \overset{k}{\rightharpoonup} x$. By virtue of (A1), J is weakly lower semicontinuous and it follows from (iv) that $J(x) \leq \underline{\lim}_k J(x_{n_k}) = \lim_n J(x_n) \leq \overline{J}$. However, $x \in S$ and $\overline{x}$ is the unique solution to (P). Hence, $x = \overline{x}$, $\mathfrak{W}(x_n)_{n\geq 0} = \{\overline{x}\}$, and, since C is weakly compact, (ii) yields $x_n \overset{n}{\rightharpoonup} \overline{x}$. Repeating the above argument, we obtain $\overline{J} = J(\overline{x}) \leq \underline{\lim}_n J(x_n)$ and, by (iv), $J(x_n) \overset{n}{\to} \overline{J}$. Since $(x_n + \overline{x})/2 \overset{n}{\rightharpoonup} \overline{x}$, the weak lower semicontinuity of J and (1.1) yield

$$
\begin{aligned}
\overline{J} &\leq \underline{\lim}_n J\left(\frac{x_n + \overline{x}}{2}\right) \\
&\leq \overline{\lim}_n \frac{J(x_n) + \overline{J}}{2} - \overline{\lim}_n c(\|x_n - \overline{x}\|) \\
&= \overline{J} - \overline{\lim}_n c(\|x_n - \overline{x}\|). \qquad (3.1)
\end{aligned}
$$

Hence, $c(\|x_n - \overline{x}\|) \overset{n}{\to} 0$ and, by (A2), $x_n \overset{n}{\to} \overline{x}$. (vii): For every $n \in \mathbb{N}$, $(x_n, x_{n+1}) \in D_n^2$ and therefore $y_n = (x_n + x_{n+1})/2 \in D_n$. Since $x_n = \mathfrak{m}(D_n)$, (1.1) then yields

$$
J(x_n) \leq J(y_n) \leq \frac{J(x_n) + J(x_{n+1})}{2} - c(\|x_{n+1} - x_n\|). \qquad (3.2)
$$

Hence, (iv) implies $c(\|x_{n+1} - x_n\|) \overset{n}{\to} 0$ and, in turn, $x_{n+1} - x_n \overset{n}{\to} 0$. (vii) $\Rightarrow$ (viii): $(\forall n \in \mathbb{N}) x_{n+1} \in H_n \Rightarrow \|x_{n+1} - x_n\| \geq d(x_n, H_n)$. $\square$

Item (i) above shows that Algorithm 1.1 is an ascent method. On the other hand, item (vi) guarantees the strong convergence of any orbit to the solution of (P) as long as each of its weak cluster points satisfies all the constraints. In view of (1.3), for this condition to hold, the control sequence $(I_n)_{n\geq 0}$ determining the blocks of constraints activated over the course of the iterations must sweep through the index set I in a coherent fashion; three suitable control modes will be considered in Definition 3.1. In addition, the constraint sets $(S_i)_{i\in I}$ must be tightly approximated by the cuts $(H_n)_{n\geq 0}$ in a sense that will be made precise in Definition 3.2.

DEFINITION 3.1. *Algorithm* 1.1 *operates under*

- admissible *control if I is countable and there exist positive integers $(M_i)_{i\in I}$ such that*

$$
\big(\forall (i,n) \in I \times \mathbb{N}\big) \quad i \in \bigcup_{k=n}^{n+M_i-1} I_k; \qquad (3.3)
$$

- chaotic *control if I is countable and*

$$
I = \overline{\lim}_n I_n \triangleq \bigcap_{n\geq 0} \bigcup_{k\geq n} I_k; \qquad (3.4)
$$

- coercive *control if*

$$\left(\exists (\mathrm{i}(n))_{n\geq 0} \in \bigtimes_{n\geq 0} I_n\right) \quad \overline{\lim}_n g_{\mathrm{i}(n)}(x_n) \leq 0 \Rightarrow \overline{\lim}_n \sup_{i\in I} g_i(x_n) \leq 0. \tag{3.5}$$

In addition, Algorithm 1.1 *is* serial *if* $(I_n)_{n\geq 0}$ *reduces to a sequence of singletons* $(\{i(n)\})_{n\geq 0}$. *The above admissibility and coercivity conditions then read*

$$(\forall i \in I)(\exists M_i \in \mathbb{N}^*)(\forall n \in \mathbb{N}) \quad i \in \{\mathrm{i}(n), \ldots, \mathrm{i}(n+M_i-1)\} \tag{3.6}$$

and

$$\overline{\lim}_n g_{\mathrm{i}(n)}(x_n) \leq 0 \Rightarrow \overline{\lim}_n \sup_{i\in I} g_i(x_n) \leq 0, \tag{3.7}$$

respectively.

The coercive control mode is found in [13] with $(\forall i \in I)\ g_i \colon x \mapsto d(x, S_i)$. The admissible and chaotic control modes have already been used at various levels of generality in convex feasibility problems [4], [10], [13], [34], [42].

DEFINITION 3.2. *Let* $(x_n)_{n\geq 0}$ *be an arbitrary orbit of Algorithm* 1.1. *Then the algorithm will be said to be*

- tight *if, for every* $i \in I$ *and every increasing sequence* $(n_k)_{k\geq 0} \subset \mathbb{N}$ *such that* $i \in \bigcap_{k\geq 0} I_{n_k}$, *we have* $\overline{\lim}_k g_i(x_{n_k}) \leq 0$;
- strongly tight *if* $\overline{\lim}_n \max_{i\in I_n} g_i(x_n) \leq 0$.

It is clear that strong tightness implies tightness. We show below that, when I is finite, the distinction between the two notions disappears.

PROPOSITION 3.2. *Suppose that* I *is finite. Then Algorithm* 1.1 *is tight if and only if it is strongly tight.*

Proof. To show necessity, take an arbitrary orbit $(x_n)_{n\geq 0}$ and suppose that the algorithm is not strongly tight, i.e., that $\epsilon \triangleq \overline{\lim}_n \max_{i\in I_n} g_i(x_n) > 0$. Define a sequence $(\mathrm{i}(n))_{n\geq 0} \subset I$ by $(\forall n \in \mathbb{N})\ g_{\mathrm{i}(n)}(x_n) = \max_{i\in I_n} g_i(x_n)$. Then, since I is finite, there exists an index $i \in I$ and an increasing sequence $(n_k)_{k\geq 0} \subset \mathbb{N}$ such that $(\forall k \in \mathbb{N})\ \mathrm{i}(n_k) = i$ and $g_i(x_{n_k}) \xrightarrow{k} \epsilon$, in contradiction of the tightness assumption. ◻

We are now ready to state and prove the following strong convergence result.

THEOREM 3.1. *Let* $(x_n)_{n\geq 0}$ *be an arbitrary orbit of Algorithm* 1.1 *generated under either of the following conditions:*

(i) *tightness,* I *countable, and admissible control;*

(ii) *strong tightness and coercive control.*

Then $x_n \xrightarrow{n} \overline{x}$.

Proof. By virtue of Proposition 3.1(vi), it suffices to show $\mathfrak{W}(x_n)_{n\geq 0} \subset S$. Fix arbitrarily $i \in I$ and $x \in \mathfrak{W}(x_n)_{n\geq 0}$, say $x_{n_k} \overset{k}{\rightharpoonup} x$. Then it is enough to show $x \in S_i$, i.e., that $g_i(x) \leq 0$. (i): By (3.3), there exist $M_i \in \mathbb{N}^*$ and an increasing sequence $(p_k)_{k\geq 0} \subset \mathbb{N}$ such that

$$(\forall k \in \mathbb{N}) \quad n_k \leq p_k \leq n_k + M_i - 1 \qquad \text{and} \qquad i \in I_{p_k}. \tag{3.8}$$

Hence

$$(\forall k \in \mathbb{N}) \quad \|x_{p_k} - x_{n_k}\| \leq \sum_{l=n_k}^{n_k+M_i-2} \|x_{l+1} - x_l\| \tag{3.9}$$

and Proposition 3.1(vii) yields $x_{p_k} - x_{n_k} \xrightarrow{k} 0$. Consequently, $x_{p_k} \xrightarrow{k} x$. On the other hand, the tightness condition gives $\overline{\lim}_k g_i(x_{p_k}) \leq 0$ and (A3) then yields $g_i(x) \leq 0$, as desired. (ii): The strong tightness condition gives $\overline{\lim}_n \max_{j\in I_n} g_j(x_n) \leq 0$. However, since the control is coercive, we obtain

$$\begin{aligned} \overline{\lim}_n \max_{j\in I_n} g_j(x_n) \leq 0 &\Rightarrow \overline{\lim}_n g_{\mathrm{i}(n)}(x_n) \leq 0 \\ &\Rightarrow \overline{\lim}_n \sup_{j\in I} g_j(x_n) \leq 0 \\ &\Rightarrow \overline{\lim}_k g_i(x_{n_k}) \leq 0, \end{aligned} \tag{3.10}$$

where the sequence $(\mathrm{i}(n))_{n\geq 0}$ is as in (3.5). It then follows from (A3) that $g_i(x) \leq 0$, which completes the proof. ☐

We conclude this section by supplying a theoretical condition under which Theorem 3.1(i) can be extended to the chaotic control mode (3.4).

PROPOSITION 3.3. *Suppose that Algorithm* 1.1 *is tight and that I is countable, and let $(x_n)_{n\geq 0}$ be any of its orbits generated under chaotic control. Then, if $(x_n)_{n\geq 0}$ admits at most one weak cluster point, $x_n \xrightarrow{n} \overline{x}$.*

Proof. It follows from Proposition 3.1(ii) and the weak compactness of C that if $(x_n)_{n\geq 0}$ admits at most one weak cluster point, then it converges weakly, say $x_n \overset{n}{\rightharpoonup} x$. Now, fix $i \in I$ arbitrarily. According to Proposition 3.1(vi), it remains to show $g_i(x) \leq 0$. By condition (3.4), there exists an increasing sequence $(n_k)_{k\geq 0} \subset \mathbb{N}$ such that $i \in \bigcap_{k\geq 0} I_{n_k}$. In turn, tightness implies $\overline{\lim}_k g_i(x_{n_k}) \leq 0$ and, since $x_{n_k} \overset{k}{\rightharpoonup} x$, (A3) yields $g_i(x) \leq 0$. ☐

The execution of iteration n of Algorithm 1.1 necessitates the construction of a cut H_n at Step 1 and of a base D_{n+1} at Step 3 (see Definition 1.1). This question is addressed in the next two sections.

4. Cut construction schemes. In this section, we describe two techniques to construct cuts for Algorithm 1.1 and provide examples of families of constraint functions $(g_i)_{i\in I}$ that yield tight and strongly tight algorithms in each case.

4.1. Exact-constraint cuts. Here, Algorithm 1.1 is assumed to operate under serial control, say $(\forall n \in \mathbb{N})\ I_n = \{\mathrm{i}(n)\}$. In view of Definition 1.1, the following observation is self-evident.

PROPOSITION 4.1. *The set $H_n = S_{i(n)}$ is a cut for Algorithm* 1.1 *at iteration n.*

When it operates under serial control with cuts generated as above, Algorithm 1.1 will be said to be implemented with *exact-constraint cuts*. We now proceed with some examples of families $(g_i)_{i\in I}$ that yield tight and strongly tight algorithms (see also Proposition 3.2). In Propositions 4.2 and 4.3, γ is the diameter of C in (A2) and $Q = B(u, 2\gamma)$.

PROPOSITION 4.2. *Algorithm* 1.1 *with exact-constraint cuts is tight if, for every $i \in I$, one of the following conditions holds.*

(i) *g_i is uniformly continuous on Q.*
(ii) *g_i is weakly continuous on Q.*
(iii) *$\mathcal{X} = \mathbb{R}^N$ and g_i is finite and convex.*
(iv) *g_i is the displacement function of an operator $T_i\colon \mathcal{X} \to \mathcal{X}$ which satisfies condition* (c) *(in particular* (d) *or* (e)*) in Proposition* 2.2(iv) *with constant $\eta_i \in \mathbb{R}_+^*$.*

Proof. Given an arbitrary orbit $(x_n)_{n\geq 0}$, Propositions 3.1(viii) and 4.1 give $d(x_n, S_{\mathrm{i}(n)}) \xrightarrow{n} 0$. Now take an index $i \in I$ and an increasing sequence $(n_k)_{k\geq 0} \subset \mathbb{N}$

such that, for every $k \in \mathbb{N}$, $i = \underline{\mathrm{i}(n_k)}$. Then $d(x_{n_k}, S_i) \xrightarrow{k} 0$ and, in view of Definition 3.2, it must be proved that $\overline{\lim}_k g_i(x_{n_k}) \leq 0$. (i) is similar to Proposition 4.3(i) and thus is omitted. (ii) $\Rightarrow$ (i) follows from the weak compactness of Q. (iii) $\Rightarrow$ (i): g_i is Lipschitzian on Q by [46, Thm. 10.4]. (iv): For every $k \in \mathbb{N}$, let $p_{i,k}$ be a projection of x_{n_k} onto S_i and suppose that (c) in Proposition 2.2(iv) holds with constant $\eta_i \in \mathbb{R}_+^*$. Then, there exists $z'_k \in \Delta(x_{n_k} - T_i x_{n_k})$ such that $\|T_i x_{n_k} - x_{n_k}\|^2 \leq \eta_i^{-1} \langle x_{n_k} - p_{i,k}, z'_k \rangle$. Consequently, $\|T_i x_{n_k} - x_{n_k}\|^2 \leq \eta_i^{-1} \|x_{n_k} - p_{i,k}\| \cdot \|z'_k\| = \eta_i^{-1} d(x_{n_k}, S_i) \cdot \|T_i x_{n_k} - x_{n_k}\|$ and we conclude $g_i(x_{n_k}) \leq \eta_i^{-1} d(x_{n_k}, S_i)$. □

PROPOSITION 4.3. *Algorithm* 1.1 *with exact-constraint cuts is strongly tight if one of the following conditions holds.*

(i) $(g_i)_{i \in I}$ *is uniformly equicontinuous on* Q.

(ii) $(g_i)_{i \in I}$ *is weakly equicontinuous on* Q*: For every* $(x, \epsilon) \in Q \times \mathbb{R}_+^*$*, there exists a weak neighborhood* V *of* x *such that* $(\forall y \in V)(\forall i \in I)\ |g_i(x) - g_i(y)| \leq \epsilon$.

(iii) $(g_i)_{i \in I}$ *is a family of affine functions associated with a family of pointwise bounded continuous linear functions.*

(iv) $\mathcal{X} = \mathbb{R}^N$ *and* $(g_i)_{i \in I}$ *is a family of pointwise bounded convex functions.*

(v) $(g_i)_{i \in I}$ *is a family of displacement functions of operators* $(T_i)_{i \in I}$ *as in Proposition* 4.2(iv) *with* $\eta \triangleq \inf_{i \in I} \eta_i > 0$ *(in particular, each* T_i *satisfies condition* (d) *or* (e) *in Proposition* 2.2(iv)*).*

Proof. Take an arbitrary orbit $(x_n)_{n \geq 0}$. Then, as above, $d(x_n, S_{\mathrm{i}(n)}) \xrightarrow{n} 0$ and, in view of Definition 3.2, it must be proved that $\overline{\lim}_n g_{\mathrm{i}(n)}(x_n) \leq 0$. (i): Fix $\epsilon \in \mathbb{R}_+^*$, extract a subsequence $(x_{n_k})_{k \geq 0}$ such that $(g_{\mathrm{i}(n_k)}(x_{n_k}))_{k \geq 0} \subset]0, +\infty]$ (if no such subsequence exists, the proof is complete), and let p_k be a projection of x_{n_k} onto $S_{\mathrm{i}(n_k)}$. Since $u \in S_{\mathrm{i}(n_k)}$ and $(x_{n_k}, u) \in C^2$, we have

$$\|p_k - u\| \leq \|x_{n_k} - u\| + \|x_{n_k} - p_k\| \leq 2\|x_{n_k} - u\| \leq 2\gamma \tag{4.1}$$

and, in turn, $p_k \in B(u, 2\gamma) = Q$. Next, as $x_{n_k} - p_k \xrightarrow{k} 0$ and $((x_{n_k}, p_k))_{k \geq 0} \subset Q^2$, the uniform equicontinuity of $(g_i)_{i \in I}$ on Q gives, for k sufficiently large, $\sup_{i \in I} |g_i(x_{n_k}) - g_i(p_k)| \leq \epsilon$ and, therefore, $0 < g_{\mathrm{i}(n_k)}(x_{n_k}) \leq \epsilon$. Since ϵ can be arbitrarily small, strong tightness ensues. (ii) $\Rightarrow$ (i) follows from the weak compactness of Q. (iii) $\Rightarrow$ (i): $(\forall i \in I) g_i \colon x \mapsto \langle x, z'_i \rangle + \alpha_i$, where $(z'_i, \alpha_i) \in \mathcal{X}' \times \mathbb{R}$ and $(\forall x \in \mathcal{X}) \sup_{i \in I} |\langle x, z'_i \rangle| < +\infty$. The uniform boundedness principle [1, Thm. 1.1.4] asserts that $\zeta \triangleq \sup_{i \in I} \|z'_i\| < +\infty$ and, therefore, that $(g_i)_{i \in I}$ is equi-Lipschitzian with constant ζ. (iv) $\Rightarrow$ (i): $(g_i)_{i \in I}$ is equi-Lipschitzian on Q by [46, Th. 10.6]. (v): Following the proof of Proposition 4.2(iv), we obtain $(\forall n \in \mathbb{N})\ g_{\mathrm{i}(n)}(x_n) \leq \eta^{-1} d(x_n, S_{\mathrm{i}(n)})$. □

It should be remarked that in Hilbert spaces, projectors are nonexpansive [24, Chap. 12]. Accordingly, the inequalities $\|p_k - u\| \leq \|x_{n_k} - u\| \leq \gamma$ can be used in lieu of (4.1), and one can take $Q = B(u, \gamma)$ in Propositions 4.2 and 4.3.

Next, we recover the framework proposed by Laurent and Martinet in [35].

EXAMPLE 4.1. *Under the strong tightness condition, Algorithm* 1.1 *implemented with coercive control and exact-constraint cuts contains the setting of* [35]. *There,* (P) *is investigated under assumptions* (A1)–(A2) *with* $E = \mathcal{X}$ *and the special instance of* (A3) *when the functions* $(g_i)_{i \in I}$ *are lower semicontinuous, convex, and satisfy the condition*

$$(\exists \Omega \in \mathbb{R}_+^*)\big(\forall (x, i) \in C \times I\big) \quad g_i(x) \leq \Omega d(x, S_i). \tag{4.2}$$

Furthermore, the serial control rule

$$(4.3)\qquad (\forall n \in \mathbb{N})\ g_{\mathrm{i}(n)}(x_n) \geq \theta \sup_{i\in I} g_i(x_n) - \rho_n, \quad \text{where } 0 < \theta \leq 1 \text{ and } 0 \leq \rho_n \overset{n}{\to} 0$$

is in force. Since $d(x_n, S_{\mathrm{i}(n)}) \overset{n}{\to} 0$, (4.2) $\Rightarrow \overline{\lim}_n g_{\mathrm{i}(n)}(x_n) \leq 0$, *which shows strong tightness. On the other hand, since* (4.3) $\Rightarrow$ (3.7)*, the control is coercive. Hence,* [35, Thm. 1] *is a corollary of Theorem* 3.1(ii)*.*

4.2. Surrogate cuts. In this section, the cut H_n at iteration n is constructed as a surrogate half-space (this terminology appears in [23]). The basic idea is, for every $i \in I_n$, to "linearize" g_i by approximating it by a continuous affine function $g_{i,n}$ (determined here geometrically via a projection onto a simple superset of S_i). A surrogate function $\widetilde{g}_n$ is then formed as a convex combination of the family $(g_{i,n})_{i\in I_n}$, and the cut is defined as $H_n = \mathrm{lev}_{\leq \gamma_n} \widetilde{g}_n$ for some $\gamma_n \in \mathbb{R}_+$. We formally define *surrogate cuts* as follows.

PROPOSITION 4.4. *Fix* $(\delta, \varepsilon) \in]0,1[^2$ *and let*

$$(4.4)\qquad H_n = \left\{ x \in \mathcal{X} \mid \sum_{i\in I_n} w_{i,n} \langle x - p_{i,n}, q'_{i,n}\rangle \leq \gamma_n \right\},$$

where the following conditions hold.

(C1) *For every* $i \in I_n$, $p_{i,n}$ *is a projection of* x_n *onto a set* $S_{i,n} \in \mathfrak{C}(S_i)$ *and* $q'_{i,n} \in \Delta(x_n - p_{i,n})$ *satisfies*

$$(4.5)\qquad (\forall x \in S_{i,n}) \quad \langle x - p_{i,n}, q'_{i,n}\rangle \leq 0.$$

(C2) $(w_{i,n})_{i\in I_n} \subset [0,1]$, $\sum_{i\in I_n} w_{i,n} = 1$, *and*

$$(\exists j \in I_n) \quad \begin{cases} d(x_n, S_{j,n}) = \max_{i\in I_n} d(x_n, S_{i,n}), \\ w_{j,n} \geq \delta. \end{cases}$$

(C3) $0 \leq \gamma_n \leq (1-\varepsilon) \sum_{i\in I_n} w_{i,n} d(x_n, S_{i,n})^2$.

Then H_n *is a cut for Algorithm* 1.1 *at iteration* n*.*

Proof. Let us show that (1.3) holds. First, it is clear that H_n is closed and convex. Second, (C1) yields $S_i \subset S_{i,n} \subset \{x \in \mathcal{X} \mid \langle x - p_{i,n}, q'_{i,n}\rangle \leq 0\}$ for every $i \in I_n$. Hence, by virtue of (C2) and (C3), $H_n \in \mathfrak{C}(\bigcap_{i\in I_n} S_i)$. $\square$

The existence of $(q'_{i,n})_{i\in I_n}$ in (C1) is guaranteed by (2.4), while (2.1) yields

$$(4.6)\qquad (\forall i \in I_n) \quad \|q'_{i,n}\|^2 = d(x_n, S_{i,n})^2 = \langle x_n - p_{i,n}, q'_{i,n}\rangle.$$

On the other hand, $(p_{i,n})_{i\in I_n}$ and $(q'_{i,n})_{i\in I_n}$ are uniquely defined if $\mathcal{X}$ and $\mathcal{X}'$ are strictly convex, respectively [1]. In particular, if $\mathcal{X}$ is a Hilbert space, one can identify $q'_{i,n} = x_n - p_{i,n}$ hereafter, and (4.4) becomes

$$(4.7)\qquad H_n = \left\{ x \in \mathcal{X} \mid \sum_{i\in I_n} w_{i,n} \langle x - p_{i,n}, x_n - p_{i,n}\rangle \leq \gamma_n \right\}.$$

Surrogate half-spaces have already been used—explicitly or implicitly—for solving convex feasibility problems in Hilbert spaces. Thus, in the methods of [10], [12], [13], [22], [32], [34], [40], [42], [45], the update x_{n+1} is obtained by (under/over) projecting the current iterate x_n onto a half-space whose general form is (4.7). This point will be reexamined in section 7.2.

An important feature of Algorithm 1.1 with surrogate cuts is that it does not require the ability to enforce exactly a constraint "$g_i(x) \leq 0$" selected at Step 1 but merely the ability to move the current iterate x_n toward S_i by means of a projection onto a superset $S_{i,n}$. A wide range of approximating supersets are acceptable, and the construction of $S_{i,n}$ can be adapted to the nature of the function g_i. In the two examples below, $S_{i,n}$ is constructed as an affine half-space and the expressions for $p_{i,n}$, $q'_{i,n}$, and $d(x_n, S_{i,n})$ are derived from the following facts.

LEMMA 4.1 (see [48, Lem. I.1.2]). *Given a nonzero functional $z' \in \mathcal{X}'$ and $\alpha \in \mathbb{R}$, consider the closed affine half-space $A = \{y \in \mathcal{X} \mid \langle y, z'\rangle \leq \alpha\}$. Take $x \notin A$ and let $p = x + (\alpha - \langle x, z'\rangle)z/\|z'\|^2$, where $z \in \Delta^{-1}(z')$. Then $d(x, A) = (\langle x, z'\rangle - \alpha)/\|z'\|$ and p is a projection of x onto A.*

EXAMPLE 4.2. *The function g_i is convex and lower semicontinuous, and subdifferentiable on C. Then, given $t'_{i,n} \in \partial g_i(x_n)$, the function $x \mapsto \langle x - x_n, t'_{i,n}\rangle + g_i(x_n)$ minorizes g_i by* (2.2). *Thus*

$$S_{i,n} = \{x \in \mathcal{X} \mid \langle x_n - x, t'_{i,n}\rangle \geq g_i(x_n)\} \tag{4.8}$$

lies in $\mathfrak{C}(S_i)$. If $x_n \notin S_i$, then $p_{i,n} = x_n - g_i(x_n)t_{i,n}/\|t'_{i,n}\|^2$, $q'_{i,n} = g_i(x_n)t'_{i,n}/\|t'_{i,n}\|^2$, and $d(x_n, S_{i,n}) = g_i(x_n)/\|t'_{i,n}\|$, where $t_{i,n} \in \Delta^{-1}(t'_{i,n})$.

Approximations of type (4.8) go back to [31] and have been used extensively; see, e.g., [4], [12], [29], [32], [34].

EXAMPLE 4.3. *The function g_i is the displacement function of an operator T_i as in Proposition* 2.2(iv)(c) *with constant $\eta_i \in \mathbb{R}_+^*$. Hence $S_i = \operatorname{Fix} T_i$ and, for some $z'_{i,n} \in \Delta(x_n - T_i x_n)$, we have $\langle x_n - x, z'_{i,n}\rangle \geq \eta_i\|T_i x_n - x_n\|^2$ for every $x \in S_i$. Therefore*

$$S_{i,n} = \big\{x \in \mathcal{X} \mid \langle x - x_n - \eta_i(T_i x_n - x_n), z'_{i,n}\rangle \leq 0\big\} \tag{4.9}$$

lies in $\mathfrak{C}(S_i)$. Furthermore, $p_{i,n} = x_n + \eta_i(T_i x_n - x_n)$, $q'_{i,n} = \eta_i z'_{i,n}$, and $d(x_n, S_{i,n}) = \eta_i\|T_i x_n - x_n\|$.

Further examples can be derived from Example 4.3 by considering the special cases (d) or (e) of (c) in Proposition 2.2(iv). For instance, if $\mathcal{X}$ is a Hilbert space and T_i is firmly nonexpansive, (4.9) reads as

$$S_{i,n} = \{x \in \mathcal{X} \mid \langle x - T_i x_n, x_n - T_i x_n\rangle \leq 0\}. \tag{4.10}$$

This particular approximation appears implicitly in [4] and [10], and explicitly in [34].

We preface our study of the tightness of Algorithm 1.1 with surrogate cuts with two basic facts.

PROPOSITION 4.5. $(\forall n \in \mathbb{N})$ $\sum_{i\in I_n} w_{i,n} q'_{i,n} = 0 \Leftrightarrow x_n \in \bigcap_{i\in I_n} S_{i,n} \Leftrightarrow x_n \in H_n$.

Proof. Fix $(n, x) \in \mathbb{N} \times S$. Then (C2), (4.6), and (4.5) imply

$$\begin{aligned}\delta \max_{i\in I_n} d(x_n, S_{i,n})^2 &\le \sum_{i\in I_n} w_{i,n} d(x_n, S_{i,n})^2 = \sum_{i\in I_n} w_{i,n} \langle x_n - p_{i,n}, q'_{i,n}\rangle \\ &= \sum_{i\in I_n} w_{i,n} \langle x_n - x, q'_{i,n}\rangle + \sum_{i\in I_n} w_{i,n} \langle x - p_{i,n}, q'_{i,n}\rangle \\ &\le \left\langle x_n - x, \sum_{i\in I_n} w_{i,n} q'_{i,n}\right\rangle. \qquad (4.11)\end{aligned}$$

Hence, $\sum_{i\in I_n} w_{i,n} q'_{i,n} = 0 \Rightarrow \max_{i\in I_n} d(x_n, S_{i,n})^2 = 0 \Rightarrow x_n \in \bigcap_{i\in I_n} S_{i,n}$. The three other implications are easily obtained. ☐

PROPOSITION 4.6. $\max_{i\in I_n} d(x_n, S_{i,n}) \xrightarrow{n} 0$.

Proof. Fix $n \in \mathbb{N}$ and suppose $x_n \notin \bigcap_{i\in I_n} S_{i,n}$. Then Proposition 4.5, the convexity of $\|\cdot\|$, (C2), and (4.6) yield $0 \ne \|\sum_{i\in I_n} w_{i,n} q'_{i,n}\| \le \sum_{i\in I_n} w_{i,n} \|q'_{i,n}\| \le \max_{i\in I_n} d(x_n, S_{i,n})$. Consequently, we derive from (4.4), Lemma 4.1, (4.6), and (4.11) that

$$\begin{aligned} d(x_n, H_n) &= \frac{\sum_{i\in I_n} w_{i,n} \langle x_n - p_{i,n}, q'_{i,n}\rangle - \gamma_n}{\|\sum_{i\in I_n} w_{i,n} q'_{i,n}\|} \\ &\ge \varepsilon \frac{\sum_{i\in I_n} w_{i,n} d(x_n, S_{i,n})^2}{\max_{i\in I_n} d(x_n, S_{i,n})} \\ &\ge \delta\varepsilon \max_{i\in I_n} d(x_n, S_{i,n}). \qquad (4.12)\end{aligned}$$

On the other hand, if $x_n \in \bigcap_{i\in I_n} S_{i,n}$, then (4.12) is immediate. Since $d(x_n, H_n) \xrightarrow{n} 0$ by Proposition 3.1(viii), the assertion is proved. ☐

We observe in passing that when Algorithm 1.1 is implemented with surrogate cuts and satisfies the tightness condition then, for every index $i \in I$ and every suborbit $(x_{n_k})_{k\ge0}$ such that $i \in \bigcap_{k\ge0} I_{n_k}$, it follows from Proposition 4.6 that $d(x_{n_k}, S_{i,n_k}) \xrightarrow{k} 0$ and from (A3) that $x_{n_k} \overset{k}{\rightharpoonup} x \Rightarrow x \in S_i$. This is essentially the *focusing* property introduced in [4].

We wind up this section by furnishing convenient criteria for tightness and strong tightness.

PROPOSITION 4.7. *Algorithm* 1.1 *with surrogate cuts is tight if, for every* $i \in I$ *and every suborbit* $(x_{n_k})_{k\ge0}$ *such that* $i \in \bigcap_{k\ge0} I_{n_k}$*, one of the following conditions is fulfilled.*

(i) *It holds that*

$$d\big(x_{n_k}, S_{i,n_k}\big) \xrightarrow{k} 0 \quad \Rightarrow \quad d\big(x_{n_k}, S_i\big) \xrightarrow{k} 0 \qquad (4.13)$$

and any of conditions (i)–(iii) *in Proposition* 4.2 *is satisfied.*

(ii) g_i *is as in Example* 4.2 *with the additional assumption that its subdifferential is bounded on* C*, i.e., that*

$$\big(\exists \zeta_i \in \mathbb{R}_+^*\big)(\forall x \in C) \quad \partial g_i(x) \subset B(0, \zeta_i), \qquad (4.14)$$

and the sets $(S_{i,n_k})_{k\ge0}$ *are as in* (4.8).

(iii) *g_i is as in Example* 4.3 *and the sets* $(S_{i,n_k})_{k\geq 0}$ *are as in* (4.9).

Proof. Take an arbitrary orbit $(x_n)_{n\geq 0}$. Proposition 4.6 asserts that $\max_{i\in I_n} d(x_n, S_{i,n}) \xrightarrow{n} 0$. Hence, given $i \in I$ and an increasing sequence $(n_k)_{k\geq 0} \subset \mathbb{N}$ such that $i \in \bigcap_{k\geq 0} I_{n_k}$, we have $d(x_{n_k}, S_{i,n_k}) \xrightarrow{k} 0$ and must show $\overline{\lim}_k g_i(x_{n_k}) \leq 0$. (i): (4.13) yields $d(x_{n_k}, S_i) \xrightarrow{k} 0$. However, under condition (i) (and in particular condition (ii) or (iii)) of Proposition 4.2, $d(x_{n_k}, S_i) \xrightarrow{k} 0 \Rightarrow \overline{\lim}_k g_i(x_{n_k}) \leq 0$. (ii): $(\forall k \in \mathbb{N}) \max\{0, g_i(x_{n_k})\} = \|t'_{i,n_k}\| \cdot d(x_{n_k}, S_{i,n_k}) \leq \zeta_i d(x_{n_k}, S_{i,n_k})$ by (4.14). (iii): $(\forall k \in \mathbb{N})\ g_i(x_{n_k}) = \|T_i x_{n_k} - x_{n_k}\| = \eta_i^{-1} d(x_{n_k}, S_{i,n_k})$. □

PROPOSITION 4.8. *Algorithm* 1.1 *with surrogate cuts is strongly tight if one of the following conditions holds.*

(i) *For any of its orbits* $(x_n)_{n\geq 0}$, *we have*

$$\max_{i\in I_n} d(x_n, S_{i,n}) \xrightarrow{n} 0 \quad \Rightarrow \quad \max_{i\in I_n} d(x_n, S_i) \xrightarrow{n} 0, \tag{4.15}$$

and any of conditions (i)–(iv) *in Proposition* 4.3 *is satisfied.*

(ii) $(g_i)_{i\in I}$ *is as in Example* 4.2 *with the additional assumption that the subdifferentials are equibounded on* C, *i.e., that*

$$\big(\exists \zeta \in \mathbb{R}_+^*\big)(\forall i \in I)(\forall x \in C) \quad \partial g_i(x) \subset B(0, \zeta), \tag{4.16}$$

and the sets $((S_{i,n})_{i\in I_n})_{n\geq 0}$ *are as in* (4.8).

(iii) $(g_i)_{i\in I}$ *is as in Example* 4.3, *with the additional assumption that* $\eta \triangleq \inf_{i\in I} \eta_i > 0$, *and the sets* $((S_{i,n})_{i\in I_n})_{n\geq 0}$ *are as in* (4.9).

Proof. Take an arbitrary orbit $(x_n)_{n\geq 0}$. Then Proposition 4.6 entails $\max_{i\in I_n} d(x_n, S_{i,n}) \xrightarrow{n} 0$. Let us show $\overline{\lim}_n \max_{i\in I_n} g_i(x_n) \leq 0$. (i): Define a sequence $(\mathrm{i}(n))_{n\geq 0} \subset I$ by $(\forall n \in \mathbb{N})\ g_{\mathrm{i}(n)}(x_n) = \max_{i\in I_n} g_i(x_n)$. Then (4.15) $\Rightarrow d(x_n, S_{\mathrm{i}(n)}) \xrightarrow{n} 0$. However, under condition (i) (and in particular any of conditions (ii)–(iv)) of Proposition 4.3, $d(x_n, S_{\mathrm{i}(n)}) \xrightarrow{n} 0 \Rightarrow \overline{\lim}_n g_{\mathrm{i}(n)}(x_n) \leq 0$, as desired. (ii) and (iii): Fix $n \in \mathbb{N}$. Following the proof of Proposition 4.7(ii) and (iii), we obtain, respectively, $\max_{i\in I_n} g_i(x_n) \leq \zeta \max_{i\in I_n} d(x_n, S_{i,n})$ and $\max_{i\in I_n} g_i(x_n) = \max_{i\in I_n} \|T_i x_n - x_n\| \leq \eta^{-1} \max_{i\in I_n} d(x_n, S_{i,n})$. □

It is readily noted that (4.13) and (4.15) are satisfied in particular when exact projections onto the constraint sets are used instead of projections onto approximating supersets.

4.3. Comments. When compared with exact-constraint cuts, surrogate cuts display three advantages. First, they yield versatile block-iterative algorithms that offer great latitude in the selection of the constraints retained at each iteration. Since the pairs $(p_{i,n}, q'_{i,n})_{i\in I_n}$ can be computed simultaneously prior to their aggregation in (4.4), surrogate cuts therefore allow for flexible parallel implementations that can fully take advantage of multiprocessor systems. Second, the processing of a constraint does not require its exact enforcement. Rather, each constraint can be "linearized" by means of a projection onto an outer approximation to the corresponding constraint set. This procedure, illustrated in Examples 4.2 and 4.3, significantly lightens the computational burden of the algorithm when nonaffine constraints are present. Third, surrogate cuts are capable of producing deep cuts, as reported in various theoretical and numerical studies, e.g., [11], [12], [32], [44], [45]. In this connection, the problem of finding optimal weights $(w_{i,n})_{i\in I_n}$ in terms of maximizing $d(x_n, H_n)$ is addressed in [32] and [33].

5. Base construction schemes. Two approaches to the construction of bases for Algorithm 1.1 are described in this section.

5.1. Cumulative bases. Steps 2 and 3 of Algorithm 1.1 suggest an obvious candidate for a base at iteration n, namely, $D_{n+1} = Q_n = E \cap D_n \cap H_n$. This base can be rewritten as

$$D_{n+1} = E \cap \bigcap_{k=0}^{n} H_k. \tag{5.1}$$

In other words, the current base is the intersection of the initial base E with all the previous cuts. This principle is the basis for the cutting plane methods originally proposed in [8] and [31], and reconsidered from a more general viewpoint in [37]. As noted in section 1, a drawback of (5.1) is that the number of cuts accumulated to define the bases grows rapidly as the iterations proceed. In the next proposition, it is pointed out that in the present framework this phenomenon can be mitigated by discarding all the cuts that are inactive at iteration n in the construction of the subsequent bases $(D_{n+k})_{k\geq 1}$.

PROPOSITION 5.1. *Let $\mathbb{K}_{-1} = \varnothing$ and, for every $n \in \mathbb{N}$, $\mathbb{A}_n = \{k \in \mathbb{K}_{n-1} \cup \{n\} \mid x_{n+1} \in \operatorname{bd} H_k\}$. Then the set*

$$D_{n+1} = E \cap \bigcap_{k\in\mathbb{K}_n} H_k, \quad \text{where } \mathbb{A}_n \subset \mathbb{K}_n \subset \mathbb{K}_{n-1} \cup \{n\} \tag{5.2}$$

is a base for Algorithm 1.1 *at iteration n.*

Proof. We need to check that (1.4) holds for D_{n+1} as in (5.2). First, as $E \in \mathfrak{C}(S)$ and, by (1.3), $(H_k)_{k\in\mathbb{K}_n} \subset \mathfrak{C}(S)$, (5.2) implies $D_{n+1} \in \mathfrak{C}(S)$. Next, to show $x_{n+1} = \mathfrak{m}(D_{n+1})$, note that $x_{n+1} = \mathfrak{m}(Q_n)$ and $Q_n = E \cap D_n \cap H_n = E \cap \bigcap_{k\in\mathbb{K}_{n-1}\cup\{n\}} H_k = D_{n+1} \cap B_n$, where $B_n = \bigcap_{k\in(\mathbb{K}_{n-1}\cup\{n\})\smallsetminus\mathbb{K}_n} H_k$. However, it follows from the definition of $\mathbb{A}_n$ and the inclusion $\mathbb{A}_n \subset \mathbb{K}_n$ that $x_{n+1} \in B_n^\circ$ and therefore that B_n is inactive at x_{n+1}. Accordingly, $x_{n+1} = \mathfrak{m}(D_{n+1} \cap B_n) = \mathfrak{m}(D_{n+1})$. The proof is complete. ∎

In particular, if at every iteration $\mathbb{K}_n = \mathbb{K}_{n-1} \cup \{n\}$, then all the cuts are retained and (5.2) relapses to (5.1). At the other end of the spectrum, the simplest bases are obtained by discarding all the inactive cuts, i.e., by taking $\mathbb{K}_n = \mathbb{A}_n$ at every iteration.

5.2. Instantaneous bases. The construction of D_{n+1} described here was first proposed for quadratic forms in Hilbert spaces by Haugazeau in [26] and extended to the present setting in [35].

PROPOSITION 5.2. *Suppose that:*

(A4) *There exists a point $v \in S \cap \operatorname{dom} J$ at which J is continuous.*

Then, given $t'_{n+1} \in \partial J(x_{n+1})$ such that $(\forall x \in Q_n)\langle x_{n+1} - x, t'_{n+1}\rangle \leq 0$, the set

$$D_{n+1} = \big\{x \in \mathcal{X} \mid \langle x_{n+1} - x, t'_{n+1}\rangle \leq 0\big\} \tag{5.3}$$

is a base for Algorithm 1.1 *at iteration n.*

Proof. Since $S \subset Q_n$, (A4) asserts that J is continuous at $v \in Q_n \cap \operatorname{dom} J$, whence, as $x_{n+1} = \mathfrak{m}(Q_n)$, the existence of t'_{n+1} is guaranteed by (2.3). Moreover, the inclusion $Q_n \subset D_{n+1}$ shows that $D_{n+1} \in \mathfrak{C}(S)$. Finally, for $A = D_{n+1}$, (2.3) yields $x_{n+1} = \mathfrak{m}(D_{n+1})$. We have thus established (1.4). ∎

If t'_{n+1} is the zero functional (which may happen only when $x_{n+1} = \mathfrak{m}(\mathcal{X})$), then $D_{n+1} = \mathcal{X}$. On the other hand, if the Gâteaux-derivative, $\nabla J(x_{n+1})$, of J at x_{n+1} exists, then

$$D_{n+1} = \{x \in \mathcal{X} \mid \langle x_{n+1} - x, \nabla J(x_{n+1})\rangle \leq 0\}. \tag{5.4}$$

5.3. Comments. An advantage of cumulative bases is their wide applicability. However, they lead to increasingly complex outer approximations as the algorithm progresses. A partial remedy to this situation is to systematically discard all the inactive cuts. One should, however, beware of its potential side-effect, namely, slower convergence. By contrast, instantaneous bases are very attractive, for they take the form of half-spaces under the relatively mild assumption (A4). Their efficacy can, however, be limited by the search for an acceptable subgradient in (5.3). Of course, this limitation vanishes altogether when J is Gâteaux-differentiable on C, the base being then explicitly given by (5.4).

6. Examples. The analysis of the preceding sections gives rise to four general realizations of Algorithm 1.1 according to whether one selects, on the one hand, exact-constraint or surrogate cuts and, on the other hand, cumulative or instantaneous bases. In this section, these four realizations are presented and the theorems stating their strong convergence to the solution $\overline{x}$ of (P) under the standing assumptions (A1)–(A3) are given. A variety of outer approximation methods are exhibited as special cases and their convergence is deduced from the main theorems. Although we have restricted ourselves to known methods, it is clear that further convergence results can be generated by considering alternative schemes subsumed by Algorithms 6.1–6.4 below.

6.1. Exact-constraint cuts and cumulative bases. If the cuts are generated as in Proposition 4.1 and the bases as in Proposition 5.1, Algorithm 1.1 reads as follows.

ALGORITHM 6.1. *A sequence $(x_n)_{n\geq 0}$ is constructed as follows, where E is supplied by* (A2).

Step 0. *Set $D_0 = E$, $x_0 = \mathfrak{m}(D_0)$, $\mathbb{K}_{-1} = \emptyset$, and $n = 0$.*
Step 1. *Take* $\mathrm{i}(n) \in I$.
Step 2. *Set $x_{n+1} = \mathfrak{m}(D_n \cap S_{\mathrm{i}(n)})$ and $\mathbb{A}_n = \{k \in \mathbb{K}_{n-1} \cup \{n\} \mid x_{n+1} \in \mathrm{bd}\, S_{\mathrm{i}(k)}\}$.*
Step 3. *Take $\mathbb{A}_n \subset \mathbb{K}_n \subset \mathbb{K}_{n-1} \cup \{n\}$ and set $D_{n+1} = E \cap \bigcap_{k \in \mathbb{K}_n} S_{\mathrm{i}(k)}$.*
Step 4. *Set $n = n+1$ and go to Step* 1.

The convergence result below is a direct application of Theorem 3.1.

THEOREM 6.1. *Let $(x_n)_{n\geq 0}$ be an arbitrary orbit of Algorithm* 6.1 *generated under either of the following conditions:* (i) *tightness, I countable, and admissible control; or* (ii) *strong tightness and coercive control. Then $x_n \overset{n}{\to} \overline{x}$.*

EXAMPLE 6.1 (see [5, Thm. 2.4]). *In Algorithm* 6.1, *$\mathcal{X} = \mathbb{R}^N$, E is bounded, J is finite and strictly convex, I is a compact metric space, $(g_i)_{i \in I}$ is a family of finite convex functions such that $(i, x) \mapsto g_i(x)$ is continuous on $I \times \mathcal{X}$, $\mathbb{K}_n = \mathbb{A}_n$ at Step* 3, *and the* most violated constraint *control mode*

$$(\forall n \in \mathbb{N}) \quad g_{\mathrm{i}(n)}(x_n) = \max_{i \in I} g_i(x_n) \tag{6.1}$$

is in force. Then $x_n \overset{n}{\to} \overline{x}$.

Proof. The conditions of Proposition 2.1(ii)(b) are fulfilled, and (A1)–(A2) are therefore satisfied. In addition, it follows from Proposition 2.2(ii) that (A3) is satisfied. For every $x \in \mathcal{X}$, the continuity of $i \mapsto g_i(x)$ on the compact space I yields $\sup_{i\in I} |g_i(x)| < +\infty$. Proposition 4.3(iv) then ensures the strong tightness of the algorithm, while (6.1) is a special instance of the coercive control mode (3.7). The claim is therefore a consequence of Theorem 6.1(ii). ☐

Examples 6.2 and 6.3 below are, respectively, infinite-dimensional formulations of Kelley's basic cutting plane algorithm [31] and of the Kaplan–Veinott supporting hyperplane algorithm [30], [52]. These algorithms were already shown in [35] to be special instances of the framework described in Example 4.1. The problem under consideration is to find the minimizer $\overline{x}$ of a function $J\colon \mathcal{X} \to]-\infty,+\infty]$ over a closed convex set S under assumptions (A1)–(A2). By expressing S as a suitable intersection of half-spaces $(S_i)_{i\in I}$, this problem will be recast in the form of (P).

EXAMPLE 6.2. *Suppose that $S = \operatorname{lev}_{\leq 0} g$, where $g\colon \mathcal{X} \to]-\infty,+\infty]$ is a lower semicontinuous convex function. Let E be a polyhedron and suppose that there exists $\zeta \in \mathbb{R}_+^*$ such that, for every $x \in C$, $\partial g(x) \subset B(0,\zeta)$. Let $x_0 = \mathfrak{m}(E)$ and define $(x_n)_{n\geq 0}$ by the recursion*

(6.2)

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \mathfrak{m}\left(E \cap \bigcap_{k=0}^{n} \{x \in \mathcal{X} \mid \langle x_k - x, t_k'\rangle \geq g(x_k)\}\right), \quad \textit{where } t_n' \in \partial g(x_n).$$

Then $x_n \overset{n}{\to} \overline{x}$.

Proof. Let $I = \{(y,t') \in \mathcal{X} \times \mathcal{X}' \mid t' \in \partial g(y)\}$ be the graph of ∂g. For every $i = (y,t') \in I$, the continuous affine function $g_i\colon x \mapsto \langle x - y, t'\rangle + g(y)$ minorizes g by virtue of (2.2) with $g_i(y) = g(y)$ and it defines a closed affine half-space $S_i = \operatorname{lev}_{\leq 0} g_i$. We can then write $S = \bigcap_{i\in I} S_i$. Since at iteration $n \in \mathbb{N}$ the function $i \mapsto g_i(x_n)$ is maximized for $\mathrm{i}(n) = (x_n, t_n')$ where $t_n' \in \partial g(x_n)$, (6.2) appears as a particular realization of Algorithm 6.1 with $\mathbb{K}_n = \mathbb{K}_{n-1} \cup \{n\}$ at Step 3 and control rule (6.1). The control is therefore coercive since (6.1) $\Rightarrow$ (3.7). Moreover, if $x_n \notin S$, $g_{\mathrm{i}(n)}(x_n) = g(x_n) = \|t_n'\| \cdot d(x_n, S_{\mathrm{i}(n)}) \leq \zeta d(x_n, S_{\mathrm{i}(n)})$. However, Proposition 3.1(viii) states that $d(x_n, S_{\mathrm{i}(n)}) \overset{n}{\to} 0$. Thus, $\overline{\lim}_n g_{\mathrm{i}(n)}(x_n) \leq 0$ and the algorithm is strongly tight. The announced result then follows from Theorem 6.1(ii). ☐

EXAMPLE 6.3. *Suppose that $\mathcal{X}$ is a Hilbert space, that S is bounded with $S^\circ \neq \varnothing$, and that E is a polyhedron. Let $x_0 = \mathfrak{m}(E)$ and $w \in S^\circ$, and define $(x_n)_{n\geq 0}$ by the recursion*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \mathfrak{m}\left(E \cap \bigcap_{k=0}^{n} H_k\right), \tag{6.3}$$

where $H_n \in \mathfrak{C}(S)$ is either the whole space $\mathcal{X}$ or an affine half-space whose boundary supports S at the point $y_n \in \operatorname{bd} S \cap [x_n, w]$, according to whether x_n lies in S or not. Then $x_n \overset{n}{\to} \overline{x}$.

Proof. Let $I = \{i \in \mathcal{X} \mid \|i\| = 1\}$ be the unit sphere in $\mathcal{X}$ and $\sigma\colon i \mapsto \sup_{x\in S}\langle x, i\rangle$ the support function of S. For every $i \in I$, define a closed affine half-space $S_i = \operatorname{lev}_{\leq 0} g_i$, where $g_i\colon x \mapsto \langle x, i\rangle - \sigma(i)$. Then $S = \bigcap_{i\in I} S_i$. By assumption, $B(w,\gamma) \subset S$ for some $\gamma \in \mathbb{R}_+^*$. Now suppose $x_n \notin S$ and let p_n be the projection of x_n onto $\operatorname{bd} H_n \ni y_n$. Then $\operatorname{bd} H_n = \{y \in \mathcal{X} \mid \langle y - y_n, p_n - x_n\rangle = 0\}$ and $d(w, \operatorname{bd} H_n) =$

$\langle w-y_n, p_n-x_n\rangle/\|p_n-x_n\|$. However, $y_n \in [x_n, w]$ and therefore $w-y_n = \alpha_n(y_n-x_n)$, where $\alpha_n = \|w-y_n\|/\|x_n-y_n\|$. Hence, $d(w, \operatorname{bd} H_n) = \alpha_n\langle y_n-x_n, p_n-x_n\rangle/\|p_n-x_n\| = \alpha_n d(x_n, H_n)$. Consequently,

$$(\forall i \in I) \quad d(x_n, H_n) = \frac{d(w, \operatorname{bd} H_n)}{\|w-y_n\|} \cdot \|x_n - y_n\| \geq \eta d(x_n, S_i), \tag{6.4}$$

where $\eta = \gamma/\sup_{y\in \operatorname{bd} S}\|w-y\| > 0$. In addition, for every $i \in I$, $x_n \notin S_i \Rightarrow d(x_n, S_i) = g_i(x_n)$. Hence, $g_{\mathrm{i}(n)}(x_n) \geq \eta \sup_{i\in I} g_i(x_n)$, where $\mathrm{i}(n) \in I$ and $S_{\mathrm{i}(n)} = H_n$, from which it follows that (6.3) is a particular realization of Algorithm 6.1 with $\mathbb{K}_n = \mathbb{K}_{n-1} \cup \{n\}$ at Step 3 and coercive control rule (3.7). Since the affine family $(g_i)_{i\in I}$ is equi-Lipschitzian on $\mathcal{X}$ with constant 1, strong tightness follows from Proposition 4.3(i) and Theorem 6.1(ii) yields $x_n \xrightarrow{n} \overline{x}$. ☐

6.2. Exact-constraint cuts and instantaneous bases. Algorithm 6.2 below is derived from Algorithm 1.1 by coupling the cuts of Proposition 4.1 together with the bases of Proposition 5.2.

ALGORITHM 6.2. *A sequence $(x_n)_{n\geq 0}$ is constructed as follows, where E is supplied by* (A2).

Step 0. *Set $D_0 = E$, $x_0 = \mathfrak{m}(D_0)$, and $n = 0$.*
Step 1. *Take* $\mathrm{i}(n) \in I$.
Step 2. *Set $x_{n+1} = \mathfrak{m}(E \cap D_n \cap S_{\mathrm{i}(n)})$.*
Step 3. *Take $t'_{n+1} \in \partial J(x_{n+1})$ such that $(\forall x \in E \cap D_n \cap S_{\mathrm{i}(n)})\ \langle x_{n+1} - x, t'_{n+1}\rangle \leq 0$ and set $D_{n+1} = \{x \in \mathcal{X} \mid \langle x_{n+1} - x, t'_{n+1}\rangle \leq 0\}$.*
Step 4. *Set $n = n + 1$ and go to Step* 1.

Convergence follows at once from Theorem 3.1.

THEOREM 6.2. *Let $(x_n)_{n\geq 0}$ be an arbitrary orbit of Algorithm* 6.2 *generated under either of the following conditions:* (i) *tightness, I countable, and admissible control; or* (ii) *strong tightness and coercive control. Then $x_n \xrightarrow{n} \overline{x}$.*

EXAMPLE 6.4 (see [26, Thm. 2]). *In Algorithm* 6.2*, $\mathcal{X}$ is a Hilbert space, $E = \mathcal{X}$, J is a coercive quadratic form, $I = \{0, \ldots, M-1\}$ is finite, $(\forall i \in I)$ $g_i \colon x \mapsto d(x, S_i)$, and the* periodic *control mode*

$$(\forall n \in \mathbb{N}) \quad \mathrm{i}(n) = n \ (\mathrm{modulo}\, M), \tag{6.5}$$

is in force. Then $x_n \xrightarrow{n} \overline{x}$.

Proof. The conditions of Proposition 2.1(iii) are satisfied and, consequently, so are (A1) and (A2). In addition, (A3) is secured by Proposition 2.2(ii) since $(g_i)_{i\in I}$ is a family of continuous and—by the convexity of the sets $(S_i)_{i\in I}$—convex functions. In addition, it follows from Proposition 3.1(viii) that $g_{\mathrm{i}(n)}(x_n) = d(x_n, S_{\mathrm{i}(n)}) \xrightarrow{n} 0$ and therefore that the algorithm is tight. Finally, since (6.5) $\Rightarrow$ (3.6), the control is admissible. Hence, the assertion follows from Theorem 6.2(i). ☐

EXAMPLE 6.5 (see [41]). *In Algorithm* 6.2*, $\mathcal{X}$ is a Hilbert space, $E = \mathcal{X}$, J is a coercive quadratic form, I is a compact metric space, $(g_i)_{i\in I}$ is a family of affine functions such that $(i, x) \mapsto g_i(x)$ is continuous on $I \times \mathcal{X}$, and the most violated constraint control mode* (6.1) *is in force. Then $x_n \xrightarrow{n} \overline{x}$.*

Proof. (A1)–(A2) hold by Proposition 2.1(iii). Now fix $i \in I$. Then $g_i \colon x \mapsto \langle x, z_i\rangle + \alpha_i$, where $z_i \in \mathcal{X}$ and $\alpha_i \in \mathbb{R}$, and (A3) holds. For every $x \in \mathcal{X}$, the continuity of $i \mapsto \langle x, z_i\rangle$ on the compact space I implies $\sup_{i\in I}|\langle x, z_i\rangle| < +\infty$. Therefore, the algorithm is strongly tight by Proposition 4.3(iii) and, since (6.1) $\Rightarrow$ (3.7), it operates under coercive control. The desired conclusion is reached by invoking Theorem 6.2(ii). ☐

EXAMPLE 6.6 (see [39]). *In Algorithm* 6.2, $\mathcal{X} = \mathbb{R}^N$, J *is finite and strictly convex,* E *is bounded,* J *is strongly convex (i.e.,* $c : \tau \mapsto \kappa\tau^2$ *with* $\kappa \in \mathbb{R}_+^*$ *in* (A2) [36]*) and differentiable on* E, $I = \{0, \dots, M-1\}$ *is finite,* $(\forall i \in I)$ $g_i \colon x \mapsto d(x, S_i)$, *and either of the following conditions is fulfilled:* (i) *the periodic control mode* (6.5) *is in force; or* (ii) *the most violated constraint control mode* (6.1) *is in force. Then* $x_n \overset{n}{\to} \overline{x}$.

Proof. The conditions of Proposition 2.1(ii)(b)—and therefore (A1)–(A2)—hold. As in Example 6.4, (A3) also holds and the algorithm is (strongly) tight. Hence, (i) and (ii) follow from Theorem 6.2(i) and (ii), respectively. ☐

6.3. Surrogate cuts and cumulative bases. Algorithm 1.1 is implemented with the cuts of Proposition 4.4 and the bases of Proposition 5.1.

ALGORITHM 6.3. *A sequence* $(x_n)_{n\geq 0}$ *is constructed as follows, where* E *is supplied by* (A2).

Step 0. *Take* $(\delta, \varepsilon) \in]0, 1[^2$ *and set* $D_0 = E$, $x_0 = \mathfrak{m}(D_0)$, $\mathbb{K}_{-1} = \varnothing$, *and* $n = 0$.

Step 1. *Take a finite index set* $\varnothing \neq I_n \subset I$ *and set* $H_n = \{x \in \mathcal{X} \mid \sum_{i\in I_n} w_{i,n}\langle x - p_{i,n}, q'_{i,n}\rangle \leq \gamma_n\}$, *where*

(C1) *For every* $i \in I_n$, $p_{i,n}$ *is a projection of* x_n *onto a set* $S_{i,n} \in \mathfrak{C}(S_i)$ *and* $q'_{i,n} \in \Delta(x_n - p_{i,n})$ *is such that* $(\forall x \in S_{i,n})$ $\langle x - p_{i,n}, q'_{i,n}\rangle \leq 0$.

(C2) $(w_{i,n})_{i\in I_n} \subset [0, 1]$, $\sum_{i\in I_n} w_{i,n} = 1$, *and*

$$(\exists j \in I_n) \quad \begin{cases} d(x_n, S_{j,n}) = \max\limits_{i\in I_n} d(x_n, S_{i,n}), \\ w_{j,n} \geq \delta. \end{cases}$$

(C3) $0 \leq \gamma_n \leq (1-\varepsilon)\sum_{i\in I_n} w_{i,n} d(x_n, S_{i,n})^2$.

Step 2. *Set* $x_{n+1} = \mathfrak{m}(D_n \cap H_n)$ *and* $\mathbb{A}_n = \{k \in \mathbb{K}_{n-1} \cup \{n\} \mid x_{n+1} \in \operatorname{bd} H_k\}$.

Step 3. *Take* $\mathbb{A}_n \subset \mathbb{K}_n \subset \mathbb{K}_{n-1} \cup \{n\}$ *and set* $D_{n+1} = E \cap \bigcap_{k\in\mathbb{K}_n} H_k$.

Step 4. *Set* $n = n + 1$ *and go to Step* 1.

Theorem 3.1 now reads as follows.

THEOREM 6.3. *Let* $(x_n)_{n\geq 0}$ *be an arbitrary orbit of Algorithm* 6.3 *generated under either of the following conditions:* (i) *tightness,* I *countable, and admissible control; or* (ii) *strong tightness and coercive control. Then* $x_n \overset{n}{\to} \overline{x}$.

It is noteworthy that Kelley's basic algorithm, presented in Example 6.2 as a special case of Algorithm 6.1, can also be viewed as a special case of Algorithm 6.3 with a single constraint set and cuts as in (4.8). Along the same lines, we present below a formulation of Kelley's algorithm with a finite number of constraints [31] under assumptions (A1)–(A2).

EXAMPLE 6.7. *In Algorithm* 6.3, I *is finite,* $(g_i)_{i\in I}$ *is a family of finite continuous convex functions satisfying* (4.16)*, the approximations* (4.8) *are used, and the most violated constraint control mode* (6.1) *is in force. Then* $x_n \overset{n}{\to} \overline{x}$.

Proof. Proposition 4.8(ii) asserts that this particular realization of Algorithm 6.3 is strongly tight. Thus, since the control is coercive, the result follows from Theorem 6.3(ii). ☐

6.4. Surrogate cuts and instantaneous bases. The fourth implementation of Algorithm 1.1 is obtained by generating the cuts as in Proposition 4.4 and the bases as in Proposition 5.2.

ALGORITHM 6.4. *A sequence $(x_n)_{n\geq 0}$ is constructed as follows, where E is supplied by* (A2).

Step 0. *Take $(\delta,\varepsilon)\in]0,1[^2$ and set $D_0=E$, $x_0=\mathfrak{m}(D_0)$, and $n=0$.*

Step 1. *Take a finite index set $\varnothing\neq I_n\subset I$ and set $H_n=\{x\in\mathcal{X}\mid\sum_{i\in I_n}w_{i,n}\langle x-p_{i,n},q'_{i,n}\rangle\leq\gamma_n\}$, where*

(C1) *For every $i\in I_n$, $p_{i,n}$ is a projection of x_n onto a set $S_{i,n}\in\mathfrak{C}(S_i)$ and $q'_{i,n}\in\Delta(x_n-p_{i,n})$ is such that $(\forall x\in S_{i,n})\ \langle x-p_{i,n},q'_{i,n}\rangle\leq 0$.*

(C2) *$(w_{i,n})_{i\in I_n}\subset[0,1]$, $\sum_{i\in I_n}w_{i,n}=1$, and*

$$(\exists j\in I_n)\quad\begin{cases}d(x_n,S_{j,n})=\max\limits_{i\in I_n}d(x_n,S_{i,n}),\\ w_{j,n}\geq\delta.\end{cases}$$

(C3) *$0\leq\gamma_n\leq(1-\varepsilon)\sum_{i\in I_n}w_{i,n}d(x_n,S_{i,n})^2$.*

Step 2. *Set $x_{n+1}=\mathfrak{m}(E\cap D_n\cap H_n)$.*

Step 3. *Take $t'_{n+1}\in\partial J(x_{n+1})$ such that $(\forall x\in E\cap D_n\cap H_n)\langle x_{n+1}-x,t'_{n+1}\rangle\leq 0$ and set $D_{n+1}=\{x\in\mathcal{X}\mid\langle x_{n+1}-x,t'_{n+1}\rangle\leq 0\}$.*

Step 4. *Set $n=n+1$ and go to Step* 1.

The convergence conditions below are furnished by Theorem 3.1.

THEOREM 6.4. *Let $(x_n)_{n\geq 0}$ be an arbitrary orbit of Algorithm* 6.4 *generated under either of the following conditions:* (i) *tightness, I countable, and admissible control; or* (ii) *strong tightness and coercive control. Then $x_n\xrightarrow{n}\overline{x}$.*

EXAMPLE 6.8. *In Algorithm* 6.4, *$\mathcal{X}$ is a Hilbert space, I is finite, $J\colon x\mapsto\ \|x-w\|^2/2$ where $w\in\mathcal{X}$, $E=\mathcal{X}$, $(\forall i\in I)\ g_i\colon x\mapsto d(x,S_i)$, $S_{i,n}=S_i$ in* (C1), *$\gamma_n=0$ in* (C3)*, $x_0=w$, and one of the following conditions is fulfilled:*

(i) [27, Thm. 3-2] *The periodic control mode* (6.5) *is in force.*

(ii) [44, Thm. V.1] *The* static *control mode*

$$(\forall n\in\mathbb{N})\quad I_n=I \tag{6.6}$$

is in force and $(\forall i\in I)\ w_{i,n}=1/\mathrm{card}I$ in (C2).

(iii) [14] *The control mode*

$$(\forall n\in\mathbb{N})\quad I_n=\Big\{i\in I\mid d(x_n,S_i)=\max_{j\in I}\,d(x_n,S_j)\Big\} \tag{6.7}$$

is in force and $(w_{i,n})_{i\in I_n}\subset]0,1]$ in (C2).

Then $x_n\xrightarrow{n}\overline{x}$.

Proof. First, the above setting fits into that of Proposition 2.1(iii), and therefore (A1)–(A2) hold. In addition, as in Example 6.4, (A3) holds. Furthermore, Proposition 4.6 yields $\max_{i\in I_n}g_i(x_n)=\max_{i\in I_n}d(x_n,S_{i,n})\xrightarrow{n}0$, which establishes the strong tightness of this implementation of Algorithm 6.4. Accordingly, since in (i) and (ii) the control conforms to the admissibility condition (3.3), the first two assertions follow from Theorem 6.4(i). Finally, since (6.7) is an instance of the coercive control mode (3.5) here, (iii) follows from Theorem 6.4(ii). ∎

EXAMPLE 6.9. *In Algorithm* 6.4, *$\mathcal{X}=\mathbb{R}^N$, J is finite, strictly convex, and differentiable with bounded lower level sets, $E=\mathcal{X}$, I is finite, $(g_i)_{i\in I}$ is a family of finite convex functions, the approximations* (4.8) *are used, and one of the following conditions is fulfilled:*

(i) [29, Thm. 1] *The most violated constraint control mode* (6.1) *is in force.*

(ii) [29, Thm. 2] *The serial admissible control mode* (3.6) *is in force.*

(iii) [29, Thm. 3] *The static control mode* (6.6) *is in force and* $(\forall (n,i) \in \mathbb{N} \times I)$, $x_n \in S_i \Rightarrow w_{i,n} = 0$.

Then $x_n \overset{n}{\to} \overline{x}$.

Proof. Note that the conditions of Proposition 2.1(ii)(a) are satisfied and that (A1)–(A3) hold. Moreover, each g_i is continuous and subdifferentiable on $\mathcal{X}$ and (4.14) holds [46, Thm. 24.7]. Since I is finite, Propositions 3.2 and 4.7(ii) therefore imply that the algorithm is strongly tight. Hence, (i) follows from Theorem 6.4(ii), while (ii) and (iii) follow from Theorem 6.4(i). $\square$

It emerges from the discussions of sections 4.3 and 5.3 that Algorithms 6.3 and 6.4, which employ surrogate cuts, are more advantageous numerically than Algorithms 6.1 and 6.2, which employ exact-constraint cuts. When instantaneous bases are easily generated, as is the case when J is differentiable on C, Algorithm 6.4 stands out as the most attractive implementation of Algorithm 1.1. Its chief asset is to generate at every iteration a simple outer approximation, namely, the intersection of two half-spaces (with the initial base E when $E \neq \mathcal{X}$). An application of Algorithm 6.4 to an important concrete problem is demonstrated next.

6.5. Projection onto an intersection of convex sets. Algorithm 6.4 is applied to the problem of finding the projection $\overline{x}$ of a point w onto the intersection of an arbitrary family of intersecting closed convex sets $(S_i)_{i\in I}$ conforming to (A3) in a real Hilbert space $\mathcal{X}$. As $J\colon x \mapsto \|x-w\|^2/2$ in (P), assumptions (A1) and (A4) are clearly satisfied and, in light of Proposition 2.1(iii), so is (A2) with $E = \mathcal{X}$.

Given $(x,y,z) \in \mathcal{X}^3$, it will be convenient to define

$$H(x,y) = \{h \in \mathcal{X} \mid \langle h-y, x-y\rangle \leq 0\} \tag{6.8}$$

and to denote by $\mathfrak{q}(x,y,z)$ the projection of x onto $H(x,y)\cap H(y,z)$. Thus, $H(x,x) = \mathcal{X}$ and, if $x \neq y$, $H(x,y)$ is a closed affine half-space onto which y is the projection of x.

ALGORITHM 6.5. *A sequence* $(x_n)_{n\geq 0}$ *is constructed as follows.*

Step 0. *Take* $(\delta,\varepsilon) \in]0,1[^2$ *and set* $x_0 = w$ *and* $n = 0$.

Step 1. *Take a finite index set* $\varnothing \neq I_n \subset I$ *and set* $z_n = x_n + \lambda_n(\sum_{i\in I_n} w_{i,n}p_{i,n} - x_n)$, *where*

(B1) *For every* $i \in I_n$, $p_{i,n}$ *is the projection of* x_n *onto a set* $S_{i,n} \in \mathfrak{C}(S_i)$.

(B2) $(w_{i,n})_{i\in I_n} \subset [0,1]$, $\sum_{i\in I_n} w_{i,n} = 1$, *and*

$$(\exists j \in I_n) \quad \begin{cases} d(x_n, S_{j,n}) = \max\limits_{i\in I_n} d(x_n, S_{i,n}), \\ w_{j,n} \geq \delta. \end{cases}$$

(B3) $$\varepsilon L_n \leq \lambda_n \leq L_n \triangleq \begin{cases} \dfrac{\sum\limits_{i\in I_n} w_{i,n}\|p_{i,n}-x_n\|^2}{\left\|\sum\limits_{i\in I_n} w_{i,n}p_{i,n} - x_n\right\|^2}, & \textit{if } x_n \notin \bigcap\limits_{i\in I_n} S_{i,n}, \\ 1 & \textit{otherwise.} \end{cases}$$

Step 2. *Set* $\pi_n = \langle x_0 - x_n, x_n - z_n\rangle$, $\mu_n = \|x_0 - x_n\|^2$, $\nu_n = \|x_n - z_n\|^2$, $\rho_n =$

$\mu_n\nu_n - \pi_n^2$, *and*

$$x_{n+1} = \mathfrak{q}(x_0, x_n, z_n) \tag{6.9}$$
$$= \begin{cases} z_n & \text{if } \rho_n = 0 \text{ and } \pi_n \geq 0, \\ x_0 + (1 + \pi_n/\nu_n)(z_n - x_n) & \text{if } \rho_n > 0 \text{ and } \pi_n\nu_n \geq \rho_n, \\ x_n + \dfrac{\nu_n}{\rho_n}(\pi_n(x_0 - x_n) + \mu_n(z_n - x_n)) & \text{if } \rho_n > 0 \text{ and } \pi_n\nu_n < \rho_n. \end{cases}$$

Step 3. *Set $n = n + 1$ and go to Step* 1.

In this algorithm, the update x_{n+1} is obtained in (6.9) as the projection of $x_0 = w$ onto the intersection of the half-spaces $H(x_0, x_n)$ and $H(x_n, z_n)$.

PROPOSITION 6.1. *In the present context, Algorithm* 6.4 *reduces to Algorithm* 6.5.

Proof. Since $E = \mathcal{X}$, we obtain $x_0 = \mathfrak{m}(\mathcal{X}) = w$ at Step 0 of Algorithm 6.4. Next, recall that the cut H_n at Step 1 of Algorithm 6.4 is given by (4.7). We shall now show $H_n = H(x_n, z_n)$. Assume $x_n \notin \bigcap_{i\in I_n} S_{i,n}$ and define $y_n = x_n - \sum_{i\in I_n} w_{i,n}p_{i,n}$ ($\neq 0$ by Proposition 4.5), $\sigma_n^2 = \sum_{i\in I_n} w_{i,n}d(x_n, S_{i,n})^2$, and $\lambda_n = (\sigma_n^2 - \gamma_n)/\|y_n\|^2$. Then $z_n = x_n - \lambda_n y_n$, $L_n = \sigma_n^2/\|y_n\|^2$, and (B3) $\Leftrightarrow$ (C3). Moreover, for every $x \in \mathcal{X}$, we have

$$\begin{aligned} x \in H_n &\Leftrightarrow \langle x, y_n\rangle \leq \sum_{i\in I_n} w_{i,n}\langle p_{i,n}, x_n - p_{i,n}\rangle + \gamma_n \\ &\Leftrightarrow \langle x, y_n\rangle \leq \langle x_n, y_n\rangle - \lambda_n\|y_n\|^2 \\ &\Leftrightarrow \langle x - z_n, y_n\rangle \leq 0 \\ &\Leftrightarrow \langle x - z_n, x_n - z_n\rangle \leq 0. \end{aligned} \tag{6.10}$$

Consequently, $H_n = H(x_n, z_n)$. Next, observe that (5.4) yields $D_n = H(x_0, x_n)$. Hence, as $E = \mathcal{X}$, (6.9) coincides with Step 2 of Algorithm 6.4; the expression for $\mathfrak{q}(x, y, z)$ in terms of x, y, and z is drawn from [27, Thm. 3-1]. Note that all the possible cases are exhausted in (6.9) since $\rho_n \geq 0$ and, as also shown in [27, Thm. 3-1], $H(x_0, x_n) \cap H(x_n, z_n) = \varnothing \Leftrightarrow \rho_n = 0$ and $\pi_n < 0$. ☐

Naturally, Algorithm 6.5 contains those described in Example 6.8 as particular instances. Unlike them, however, it can handle an infinite number of constraints, approximate projections, and flexible block-iterative control modes. Strong convergence conditions are given in Theorem 6.4.

For comparison purposes, let us now review alternative iterative schemes that generate sequences converging strongly to the sought projection $\overline{x}$. From an algorithmic standpoint, these schemes are initialized with $x_0 = w$ and operate either in the serial format

$$(\forall n \in \mathbb{N}) \quad \mathrm{i}(n) \in I \quad \text{and} \quad x_{n+1} = R_{\mathrm{i}(n),n}x_n \tag{6.11}$$

or in the static parallel format

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \sum_{i\in I} w_i R_{i,n}x_n \quad \text{with} \quad (w_i)_{i\in I} \subset]0, 1] \quad \text{and} \quad \sum_{i\in I} w_i = 1, \tag{6.12}$$

where I is assumed to be countable and $(R_{i,n})_{(i,n)\in I\times\mathbb{N}}$ is a family of operators from $\mathcal{X}$ into $\mathcal{X}$. Henceforth, $(P_i)_{i\in I}$ designates the family of projectors onto the sets $(S_i)_{i\in I}$.

(1) *Periodic projection method.* Suppose that $(S_i)_{i\in I}$ is a finite family of M closed vector subspaces and set $R_{i,n} = P_i$. Then it was shown in [25] that under the periodic control mode (6.5) the serial projection method (6.11) converges strongly to $\overline{x}$. This result remains valid in the case of closed affine subspaces and it coincides with von Neumann's alternating projection theorem for $M = 2$ (see [18] for further details).

(2) *Dykstra-like methods.* In [6], an extension of the preceding periodic projection method to finite families of closed convex sets was obtained by setting $R_{i,n}x_n = P_i(x_n + y_{i,n})$, where $y_{i,n}$ is the outward normal vector resulting from the previous projection onto S_i. In [21], this serial algorithm was examined from a dual perspective and given an elegant and natural interpretation; moreover, the convergence of its parallel counterpart (6.12) was established. (See also [3] for further analysis.) New developments were reported in [28], where a nonperiodic control mode was used in (6.11) and countably infinite families of sets were considered.

(3) *Anchor point methods.* Suppose that, for every $i \in I$, $S_i = \operatorname{Fix} T_i$ where $T_i\colon \mathcal{X} \to \mathcal{X}$ is firmly nonexpansive, i.e., satisfies (2.15). (Note that if, for some $i \in I$, T_i is merely nonexpansive, it can be replaced by the averaged mapping $T_i^{\mathrm{av}} = (T_i + \mathrm{Id})/2$ which is firmly nonexpansive [24, Thm. 12.1] and satisfies $\operatorname{Fix} T_i^{\mathrm{av}} = \operatorname{Fix} T_i$.) Anchor point methods operate with $R_{i,n}x_n = \alpha_n x_0 + (1-\alpha_n)T_i x_n$, where $(\alpha_n)_{n\geq 0} \subset]0,1]$ converges "slowly" to 0 (e.g., $\alpha_n = 1/(n+1)$ in [2] and [10]). Strong convergence was established in [2] and [38] for the serial version (6.11) under the periodic control rule (6.5) (and I finite) and in [10] for the parallel version (6.12) with I countably infinite.

(4) *Periodic quasi-projection method.* This method, proposed in [26], was described in Example 6.4 as an offspring of Algorithm 6.2. It is equivalent to executing (6.11) under the periodic control mode (6.5) and with $R_{\mathrm{i}(n),n}x_n$ as the "quasi-projection" of x_n onto $S_{\mathrm{i}(n)}$, i.e., the projection of x_n onto $S_{\mathrm{i}(n)} \cap H(x_0, x_n)$.

Overall, Algorithm 6.5 appears to enjoy more flexibility than the above methods in terms of parallel implementation and more versatility in terms of the types of constraints it can handle. Indeed, Dykstra-like and anchor point methods are not well suited for parallel block-processing due to their serial or static parallel structure. The scope of Dykstra-like methods is further limited by the fact that they require the ability to compute projections, which is possible only in special situations. In this regard, anchor point methods are somewhat less restrictive, as any firmly nonexpansive mapping admitting the set under consideration as fixed point set can be used. In addition, Dykstra-like methods require that a normal vector be carried along for each set (except for affine subspaces), which makes their implementation costly in terms of memory allocation and management. Finally, it is noted that the quasi-projection method is a rather conceptual one, the computation of quasi-projections being usually a serious obstacle to its implementability in practice.

7. Further results. In this section, we present convergence results for two variants of (P) in which the original assumptions are altered. $\mathcal{X}$ is assumed to be a Hilbert space.

7.1. Inconsistent constraints. It has been assumed so far that the constraints are consistent, i.e., that $S \neq \varnothing$ in (P). In this section, we place ourselves in the following context: S may be empty and I is finite. As before, $(P_i)_{i\in I}$ are the projectors onto the closed convex sets $(S_i)_{i\in I}$.

As in the convex feasibility problems of [9] and [17], the exact, but possibly empty, feasibility set S can be replaced by the set $\widetilde{S}$ of points which best approximate the constraints in an averaged squared-distance sense. Fix weights $(w_i)_{i\in I} \subset]0,1]$ such that $\sum_{i\in I} w_i = 1$, define a (continuous and convex) *proximity function* $\Phi\colon x \mapsto (1/2)\sum_{i\in I} w_i d(x,S_i)^2$, and let $\widetilde{S}$ be the (closed and convex) set of minimizers of Φ over $\mathcal{X}$. (P) is then replaced by

$$(\widetilde{\mathrm{P}}) \qquad \text{find } \widetilde{x} \in \widetilde{S} \quad \text{such that } J(\widetilde{x}) = \inf_{x\in\tilde{S}} J(x)$$

under assumptions (A1) and

(A0) $\widetilde{S} \neq \varnothing$,

($\widetilde{\mathrm{A2}}$) for some $\widetilde{E} \in \mathfrak{C}(\widetilde{S})$, there exists a point $\widetilde{u} \in \widetilde{S} \cap \operatorname{dom} J$ such that $\widetilde{C} \triangleq \widetilde{E} \cap \operatorname{lev}_{\leq J(\tilde{u})} J$ is bounded and J is uniformly convex on $\widetilde{C}$.

Some remarks are in order. First, if $S \neq \varnothing$, then $\widetilde{S} = S$. Second, if $S = \varnothing$, then assumption (A0) holds when one of the sets in $(S_i)_{i\in I}$ is bounded or when they are all closed affine half-spaces [17]. Third, it follows from (A0), (A1), and ($\widetilde{\mathrm{A2}}$) that ($\widetilde{\mathrm{P}}$) admits a unique solution $\widetilde{x}$.

The next step is to regard ($\widetilde{\mathrm{P}}$) as a program of the general form (P) with a single constraint set, namely $\widetilde{S}$. Consequently, ($\widetilde{\mathrm{P}}$) can be solved via Algorithms 6.3 or 6.4 by constructing suitable surrogate cuts for $\widetilde{S}$.

THEOREM 7.1. *Let $(x_n)_{n\geq 0}$ be an arbitrary orbit of Algorithms* 6.3 *or* 6.4 *in which the cut at Step* 1 *is taken to be*

$$H_n = \left\{ x \in \mathcal{X} \mid \left\langle x - \sum_{i\in I} w_i P_i x_n, x_n - \sum_{i\in I} w_i P_i x_n \right\rangle \leq \gamma_n \right\}, \tag{7.1}$$

where $0 \leq \gamma_n \leq (1-\varepsilon)\|x_n - \sum_{i\in I} w_i P_i x_n\|^2$. Then $x_n \overset{n}{\rightharpoonup} \widetilde{x}$.

Proof. The claim follows from Theorems 6.3(i) and 6.4(i). Indeed, the control is admissible since only one constraint set is present. Next, let us show that (7.1) is a valid cut at iteration n. To this end, let $T = \sum_{i\in I} w_i P_i$. Then T is firmly nonexpansive and $\operatorname{Fix} T = \widetilde{S}$ [9]. Hence (A3) holds by Proposition 2.2(iv)(e) and (7.1) is drawn from (4.10). Finally, tightness follows from Proposition 4.7(iii). ◻

7.2. Convex feasibility problems. If, instead of (A2), it is assumed that J is constant on S, then (P) turns into the convex feasibility problem

$$(\mathrm{CFP}) \qquad \text{find } \overline{x} \in S = \bigcap_{i\in I} S_i.$$

A general strategy for solving (CFP) is to construct a sequence $(x_n)_{n\geq 0}$ in which x_{n+1} is a relaxed projection of x_n onto a cut H_n. An implementation of this outer approximation scheme with surrogate cuts leads to the following block-iterative algorithm.

ALGORITHM 7.1. *In Algorithm* 6.5*, pick x_0 arbitrarily at Step* 0*, extend the relaxation range in* (B3) *to "$\varepsilon \leq \lambda_n \leq (2-\varepsilon)L_n$," and reduce Step* 2 *to "$x_{n+1} = z_n$."*

THEOREM 7.2. *Let $(x_n)_{n\geq 0}$ be an arbitrary orbit of Algorithm* 7.1 *generated under either of the following conditions:* (i) *tightness, I countable, and admissible control or* (ii) *strong tightness and coercive control. Then $(x_n)_{n\geq 0}$ converges weakly to a point in S.*

Proof. A slight modification of the results of [13, section 2] shows that $x_{n+1} - x_n \overset{n}{\to} 0$, $\max_{i \in I_n} d(x_n, S_{i,n}) \overset{n}{\to} 0$, and $(x_n)_{n \geq 0}$ converges weakly to a point in S if $\mathfrak{W}(x_n)_{n \geq 0} \subset S$, from which, by arguing along the same lines as in Theorem 3.1(i) (respectively, Theorem 3.1(ii)), we obtain (i) (respectively, (ii)). ∎

It follows from Propositions 4.7 and 4.8 that Theorem 7.2 covers the weak convergence results of [10], [12], and [13]. A closely related algorithm is proposed in [34, section 11] with similar weak convergence results.

Acknowledgments. The author wishes to express his gratitude to the two anonymous referees for their valuable remarks.

REFERENCES

[1] V. Barbu and Th. Precupanu, *Convexity and Optimization in Banach Spaces*, 2nd ed., D. Reidel, Boston, MA, 1986.

[2] H. H. Bauschke, *The approximation of fixed points of compositions of nonexpansive mappings in Hilbert space*, J. Math. Anal. Appl., 202 (1996), pp. 150–159.

[3] H. H. Bauschke and J. M. Borwein, *Dykstra's alternating projection algorithm for two sets*, J. Approx. Theory, 79 (1994), pp. 418–443.

[4] H. H. Bauschke and J. M. Borwein, *On projection algorithms for solving convex feasibility problems*, SIAM Rev., 38 (1996), pp. 367–426.

[5] J. W. Blankenship and J. E. Falk, *Infinitely constrained optimization problems*, J. Optim. Theory Appl., 19 (1976), pp. 261–281.

[6] J. P. Boyle and R. L. Dykstra, *A method for finding projections onto the intersection of convex sets in Hilbert spaces*, in Advances in Order Restricted Statistical Inference, Lecture Notes in Statist. 37, Springer-Verlag, New York, 1986, pp. 28–47.

[7] H. Brézis, *Analyse Fonctionnelle*, 2nd ed., Masson, Paris, 1993.

[8] E. W. Cheney and A. A. Goldstein, *Newton's method for convex programming and Tchebycheff approximation*, Numer. Math., 1 (1959), pp. 253–268.

[9] P. L. Combettes, *Inconsistent signal feasibility problems: Least-squares solutions in a product space*, IEEE Trans. Signal Process., 42 (1994), pp. 2955–2966.

[10] P. L. Combettes, *Construction d'un point fixe commun à une famille de contractions fermes*, C. R. Acad. Sci. Paris Sér. I Math., 320 (1995), pp. 1385–1390.

[11] P. L. Combettes, *The convex feasibility problem in image recovery*, in Advances in Imaging and Electron Physics, P. Hawkes, ed., vol. 95, Academic Press, New York, 1996, pp. 155–270.

[12] P. L. Combettes, *Convex set theoretic image recovery by extrapolated iterations of parallel subgradient projections*, IEEE Trans. Image Process., 6 (1997), pp. 493–506.

[13] P. L. Combettes, *Hilbertian convex feasibility problem: Convergence of projection methods*, Appl. Math. Optim., 35 (1997), pp. 311–330.

[14] G. Crombez, *Finding projections onto the intersection of convex sets in Hilbert spaces*, Numer. Funct. Anal. Optim., 16 (1995), pp. 637–652.

[15] R. Dautray and J.-L. Lions, *Analyse Mathématique et Calcul Numérique pour les Sciences et les Techniques*, Masson, Paris, 1984. English translation: Springer-Verlag, New York, 1988.

[16] M. A. H. Dempster and R. R. Merkovsky, *A practical geometrically convergent cutting plane algorithm*, SIAM J. Numer. Anal., 32 (1995), pp. 631–644.

[17] A. R. De Pierro and A. N. Iusem, *A parallel projection method for finding a common point of a family of convex sets*, Pesqui. Oper., 5 (1985), pp. 1–20.

[18] F. Deutsch, *The method of alternating orthogonal projections*, in Approximation Theory, Spline Functions and Applications, S. P. Singh, ed., Kluwer, The Netherlands, 1992, pp. 105–121.

[19] B. C. Eaves and W. I. Zangwill, *Generalized cutting plane algorithms*, SIAM J. Control, 9 (1971), pp. 529–542.

[20] R. E. Edwards, *Functional Analysis—Theory and Applications*, 2nd ed., Dover, New York, 1995.

[21] N. Gaffke and R. Mathar, *A cyclic projection algorithm via duality*, Metrika, 36 (1989), pp. 29–54.
[22] U. García-Palomares, *Parallel projected aggregation methods for solving the convex feasibility problem*, SIAM J. Optim., 3 (1993), pp. 882–900.
[23] F. Glover, *A multiphase-dual algorithm for the zero-one integer programming problem*, Oper. Res., 13 (1965), pp. 879–919.
[24] K. Goebel and W. A. Kirk, *Topics in Metric Fixed Point Theory*, Cambridge University Press, Cambridge, UK, 1990.
[25] I. Halperin, *The product of projection operators*, Acta Sci. Math. (Szeged), 23 (1962), pp. 96–99.
[26] Y. Haugazeau, *Sur la minimisation de formes quadratiques avec contraintes*, C. R. Acad. Sci. Paris Sér. A Math., 264 (1967), pp. 322–324.
[27] Y. Haugazeau, *Sur les Inéquations Variationnelles et la Minimisation de Fonctionnelles Convexes*, Thèse, Université de Paris, Paris, France, 1968.
[28] H. Hundal and F. Deutsch, *Two generalizations of Dykstra's cyclic projections algorithm*, Math. Programming, 77 (1997), pp. 335–355.
[29] A. N. Iusem and B. F. Svaiter, *A row-action method for convex programming*, Math. Programming, 64 (1994), pp. 149–171.
[30] A. A. Kaplan, *Determination of the extremum of a linear function on a convex set*, Soviet Math. Dokl., 9 (1968), pp. 269–271.
[31] J. E. Kelley, *The cutting-plane method for solving convex programs*, J. SIAM, 8 (1960), pp. 703–712.
[32] K. C. Kiwiel, *Block-iterative surrogate projection methods for convex feasibility problems*, Linear Algebra Appl., 215 (1995), pp. 225–259.
[33] K. C. Kiwiel, *Monotone Gram matrices and deepest surrogate inequalities in accelerated relaxation methods for convex feasibility problems*, Linear Algebra Appl., 252 (1997), pp. 27–33.
[34] K. C. Kiwiel and B. Łopuch, *Surrogate projection methods for finding fixed points of firmly nonexpansive mappings*, SIAM J. Optim., 7 (1997), pp. 1084–1102.
[35] P. J. Laurent and B. Martinet, *Méthodes duales pour le calcul du minimum d'une fonction convexe sur une intersection de convexes*, in Symposium on Optimization, Lecture Notes in Math. 132, Springer-Verlag, New York, 1970, pp. 159–180.
[36] E. S. Levitin, *Perturbation Theory in Mathematical Programming and Its Applications*, Wiley, New York, 1994.
[37] E. S. Levitin and B. T. Polyak, *Constrained minimization methods*, USSR Comput. Math. Math. Phys., 6 (1966), pp. 1–50.
[38] P.-L. Lions, *Approximation de points fixes de contractions*, C. R. Acad. Sci. Paris Sér. A Math., 284 (1977), pp. 1357–1359.
[39] A. I. Lobyrev, *The minimization of a strictly convex function on an intersection of convex sets*, Optimizacija, 5 (1972), pp. 128–132.
[40] Y. I. Merzlyakov, *On a relaxation method of solving systems of linear inequalities*, USSR Comput. Math. Math. Phys., 2 (1963), pp. 504–510.
[41] W. Oettli, *Solving optimization problems with many constraints by a sequence of subproblems containing only two constraints*, Math. Nachr., 71 (1976), pp. 143–145.
[42] N. Ottavy, *Strong convergence of projection-like methods in Hilbert spaces*, J. Optim. Theory Appl., 56 (1988), pp. 433–461.
[43] N. S. Papageorgiou, *Convex integral functionals*, Trans. Amer. Math. Soc., 349 (1997), pp. 1421–1436.
[44] G. Pierra, *Eclatement de contraintes en parallèle pour la minimisation d'une forme quadratique*, in Lecture Notes in Comput. Sci. 41, Springer-Verlag, New York, 1976, pp. 200–218.
[45] G. Pierra, *Decomposition through formalization in a product space*, Math. Programming, 28 (1984), pp. 96–115.
[46] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
[47] R. T. Rockafellar, *Conjugate Duality and Optimization*, SIAM, Philadelphia, PA, 1974.
[48] I. Singer, *Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces*, Springer-Verlag, New York, 1970.
[49] D. M. Topkis, *Cutting-plane methods without nested constraint sets*, Oper. Res., 18 (1970), pp. 404–413.
[50] D. M. Topkis, *A cutting-plane algorithm with linear and geometric rates of convergence*, J. Optim. Theory Appl., 36 (1982), pp. 1–22.
[51] V. V. Vasin, *Iterative methods for solving ill-posed problems with a priori information in Hilbert spaces*, USSR Comput. Math. Math. Phys., 28 (1988), pp. 6–13.

[52] A. F. Veinott, *The supporting hyperplane method for unimodal programming*, Oper. Res., 15 (1967), pp. 147–152.
[53] A. A. Vladimirov, Yu. E. Nesterov, and Yu. N. Čekanov, *Uniformly convex functionals*, Vestnik Moskov. Univ. Ser. XV Vychisl. Mat. Kibernet., 3 (1978), pp. 12–23.
[54] C. Zălinescu, *On uniformly convex functions*, J. Math. Anal. Appl., 95 (1983), pp. 344–374.
[55] W. I. Zangwill, *Nonlinear Programming—A Unified Approach,* Prentice-Hall, Englewood Cliffs, NJ, 1969.