Convex Set Theoretic Image Recovery by Extrapolated Iterations of Parallel Subgradient Projections

Patrick L. Combettes, *Senior Member, IEEE*

***Abstract*—Solving a convex set theoretic image recovery problem amounts to finding a point in the intersection of closed and convex sets in a Hilbert space. The projection onto convex sets (POCS) algorithm, in which an initial estimate is sequentially projected onto the individual sets according to a periodic schedule, has been the most prevalent tool to solve such problems. Nonetheless, POCS has several shortcomings: It converges slowly, it is ill suited for implementation on parallel processors, and it requires the computation of exact projections at each iteration. In this paper, we propose a general parallel projection method (EMOPSP) that overcomes these shortcomings. At each iteration of EMOPSP, a convex combination of subgradient projections onto some of the sets is formed and the update is obtained via relaxation. The relaxation parameter may vary over an iteration-dependent, extrapolated range that extends beyond the interval]0,2] used in conventional projection methods. EMOPSP not only generalizes existing projection-based schemes, but it also converges very efficiently thanks to its extrapolated relaxations. Theoretical convergence results are presented as well as numerical simulations.**

I. Introduction

THE IMAGE recovery problem is to estimate an image from signals physically or mathematically related to it. For instance, in image restoration the goal is to estimate the original form of a degraded image, whereas in image reconstruction the goal is to estimate an image from partial information pertaining to one (or several) of its transforms, e.g., Fourier, Radon, or wavelet transform. Classical point estimation theory, in which one seeks a solution that is optimal in some sense, offers standard solution techniques that have been employed extensively in image recovery, e.g., [3] and [27]. This framework, however, often provides limited flexibility in the objective and rational incorporation of constraints, especially when they arise from nonprobabilistic *a priori* knowledge. On that score, set theoretic estimation, which revolves around the notion of feasibility, constitutes a solid alternative [13].

In convex set theoretic image recovery the solution space Ξ is a Hilbert space in which the original image is described solely by a family of convex constraints $(\Psi_i)_{i\in I}$ arising from *a priori* knowledge about the problem and from the observed data. A family of closed and convex property sets $(S_i)_{i\in I}$ is constructed as

$$(\forall i \in I)\ \ S_i = \{a \in \Xi \mid a \text{ satisfies } \Psi_i\} \tag{1}$$

so that the recovery problem reduces to the convex feasibility problem

$$\text{Find } a^\star \in S = \bigcap_{i\in I} S_i. \tag{2}$$

Detailed accounts of set theoretic signal and image recovery can be found in [13] and [41]; recent work includes [12], [14], [32], [34], and [37]. Although the range of applications of set theoretic image recovery has expanded tremendously over the past two decades, most studies have relied on a single solution method for solving (2), namely *projections onto convex sets* (POCS). Assuming that the family of sets is finite, say $I = \{1, \cdots, m\}$, POCS generates an image in S as the weak limit of a sequence $(a_n)_{n\geq 0}$ of periodic projections onto the sets, that is

$$(\forall n \in \mathbb{N})\ \ a_{n+1} = P_{n\,(\text{modulo}\, m)+1}(a_n) \tag{3}$$

where P_i designates the operator of projection onto S_i. The popularity of POCS somewhat overshadows the following theoretical and numerical shortcomings:

- POCS converges slowly;
- POCS can process only one set per iteration and it is therefore not well suited for parallel computing;
- POCS requires the computation of an exact projection at each iteration, an often numerically involved subproblem;
- POCS is limited to problems with a finite number of constraints.

The purpose of this paper is to introduce a general parallel projection method that overcomes the above limitations of POCS. In this iterative method, which will be called *extrapolated method of parallel subgradient projections* (EMOPSP), a sequence $(a_n)_{n\geq 0}$ of images is constructed as follows. At iteration n, approximate projections $(P_{i,n}(a_n))_{i\in I_n}$ of the current iterate a_n onto a subfamily of property sets $(S_i)_{i\in I_n\subset I}$ are computed simultaneously and averaged via convex combination to form $d_n = \sum_{i\in I_n} w_{i,n} P_{i,n}(a_n)$. The approximate projections are implemented as subgradient projections, so that all the projection operations actually become linear (affine) ones. An extrapolation parameter $L_n \geq 1$ is then determined and the update is obtained as $a_{n+1} = a_n + \lambda_n(d_n - a_n)$, where the relaxation parameter λ_n lies in the interval $]0, 2L_n[$. As this

Manuscript received February 9, 1995; revised April 17, 1996. This work was supported by the National Science Foundation under Grant MIP-9308609. The associate editor coordinating the review of this paper and approving it for publication was Prof. J. Zhang.

The author is with the Department of Electrical Engineering, City College and Graduate School, City University of New York, New York, NY 10031 USA (e-mail: plc@ee-mail.engr.ccny.cuny.edu).

Publisher Item Identifier S 1057-7149(97)00341-2.

relaxation range extends well beyond the range $]0,2[$ used in conventional methods, EMOPSP converges very efficiently.

The scope of this work is quite general, as we pose the recovery problem in an abstract Hilbert space in which countably many convex constraints are available to define the original image. Although applications in image recovery are emphasized, our results apply to any convex set feasibility problem, such as the set theoretic estimation and design problems described in [13]. The remainder of the paper is organized as follows. Some basic notations, assumptions, and definitions are given in Section II. Section III is a review of the projection methods that have been used in signal recovery. In Section IV, the convex feasibility problem (2) is reexamined in a product space and a first approximate projection method is proposed. EMOPSP is then fully developed in Section V. Section VI is devoted to numerical applications to image recovery problems, in which EMOPSP is compared to existing methods. The conclusion appears in Section VII. Finally, Appendix A contains the proofs of our results and Appendix B a list of acronyms.

II. Preliminaries

A. General Notations

$\mathbb{R}$ is the set of real numbers, $\mathbb{R}_+$ the set of nonnegative real numbers, $\mathbb{R}_+^*$ the set of positive real numbers, $\mathbb{N}$ the set of nonnegative integers, and $\mathbb{N}^*$ the set of positive integers. The cardinality of a set A is denoted by card$A$, its complement by $\complement A$, and its characteristic function by 1_A, i.e., $1_A(a) = 1$ if $a \in A$ and $1_A(a) = 0$ if $a \notin A$. The transpose of a matrix L is denoted by ${}^tL$ and the complex conjugate of z by $\overline{z}$. The underlying image space is a real Hilbert space Ξ with scalar product $\langle \cdot \mid \cdot \rangle$, norm $\|\cdot\|$, and distance d. The zero vector in Ξ is denoted by 0 and the closed ball of center a and radius γ by $B(a,\gamma)$. $\overset{\circ}{A}$ is the interior of a set A.

B. Convex Analysis

Convex analysis plays a prominent role in this paper, and we need to review key results. Complements and details will be found in [4], [20] and, for finite dimensional spaces, [36].

In Ξ, a sequence $(a_n)_{n\geq 0}$ converges to a point a strongly if $(\|a_n - a\|)_{n\geq 0}$ converges to 0, and weakly if, for every b in Ξ, $(\langle a_n - a \mid b\rangle)_{n\geq 0}$ converges to 0. A subset A of Ξ is boundedly compact if its intersection with every closed ball is compact. Now let A be a nonempty closed and convex subset of Ξ. Then A is weakly closed. The distance from a point $a \in \Xi$ to A is $d(a,A) = \inf\{d(a,b) \mid b \in A\}$. There exists a unique point $P_A(a) \in A$ such that $d(a, P_A(a)) = d(a,A)$, which is called the projection of a onto A. The projection operator P_A is characterized by

$$(\forall (a,b) \in \Xi^2) \;\; \langle a - P_A(a) \mid P_A(b) - P_A(a)\rangle \leq 0. \tag{4}$$

Take an affine half-space $Q = \{h \in \Xi \mid \langle h \mid b\rangle \leq \kappa\}$ (where $b \neq 0$ and $\kappa \in \mathbb{R}$) such that $a \notin Q \supset A$. Then the affine hyperplane $H = \{h \in \Xi \mid \langle h \mid b\rangle = \kappa\}$ separates a and A and

$$P_Q(a) = P_H(a) = a + \frac{\kappa - \langle a \mid b\rangle}{\|b\|^2} b. \tag{5}$$

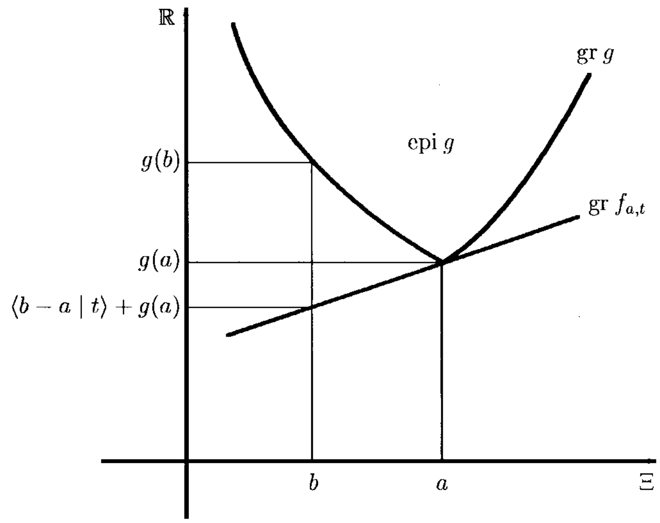

Fig. 1. Geometrical interpretation of subgradients in $\Xi \times \mathbb{R}$.

If in addition $P_A(a) \in H$, then H supports A at $P_A(a)$.

Let $g : \Xi \to \mathbb{R}$ be a functional and let $\Xi \times \mathbb{R}$ be the canonical hilbertian product space. The η-level curve ($\eta \in \mathbb{R}$), η-section ($\eta \in \mathbb{R}$), graph, and epigraph of g are, respectively, defined as

$$\begin{cases} \mathrm{lev}(g,\eta) & = & \{a \in \Xi \mid g(a) = \eta\}, \\ \mathrm{sec}(g,\eta) & = & \{a \in \Xi \mid g(a) \leq \eta\}, \\ \mathrm{gr}\, g & = & \{(a,\eta) \in \Xi \times \mathbb{R} \mid g(a) = \eta\} \\ \mathrm{epi}\, g & = & \{(a,\eta) \in \Xi \times \mathbb{R} \mid g(a) \leq \eta\}. \end{cases} \tag{6}$$

g is lower semicontinuous (l.s.c.) if epi g is closed or, equivalently, if the sets $(\mathrm{sec}(g,\eta))_{\eta\in\mathbb{R}}$ are closed. We shall say that g is lower semiboundedly-compact (l.s.b.co.) if the sets $(\mathrm{sec}(g,\eta))_{\eta\in\mathbb{R}}$ are boundedly compact. From now on, g is convex, i.e., epi g is convex or, equivalently $(\forall(\alpha,a,b) \in [0,1]\times\Xi^2)$ $g(\alpha a + (1-\alpha)b) \leq \alpha g(a) + (1-\alpha)g(b)$. Then the sets $(\mathrm{sec}(g,\eta))_{\eta\in\mathbb{R}}$ are convex and g is continuous if $\dim \Xi < +\infty$; if g is l.s.c., it is continuous. A vector t is called a subgradient of g at a if the continuous affine functional $f_{a,t} : b \mapsto \langle b - a \mid t\rangle + g(a)$, which has "slope" t and takes the same value as g at a, minorizes g on Ξ.[1] In geometrical terms, gr $f_{a,t}$ is an affine hyperplane supporting epi g at $(a, g(a))$ in $\Xi\times\mathbb{R}$ (see Fig. 1). The subdifferential of g at a is the set of its subgradients, i.e.,[2]

$$\partial g(a) = \{t \in \Xi \mid (\forall b \in \Xi) \;\; \langle b - a \mid t\rangle \leq g(b) - g(a)\}. \tag{7}$$

If g is continuous at a, then it is subdifferentiable at a: $\partial g(a) \neq \varnothing$; if g is (Gâteaux) differentiable at a, then there is a unique subgradient, $\nabla g(a)$, called gradient $\partial g(a) = \{\nabla g(a)\}$. We have

$$(\forall a \in \complement A) \;\; \nabla d(a,A) = \frac{a - P_A(a)}{\|a - P_A(a)\|}. \tag{8}$$

[1] The general theory of subdifferentiability is relatively recent [31].

[2] In particular, if $g : \mathbb{R} \to \mathbb{R}$, the subdifferential $\partial g(a)$ is the set of all slopes t of straight lines through $(a, g(a))$ which lie below gr g. Thus, $g : a \mapsto |a|$ is not differentiable at 0 but (7) gives $\partial g(0) = [-1, 1]$.

C. Assumptions

The image to be estimated, h, belongs to Ξ, where it is described by a nonvoid finite or countably infinite family[3] $(S_i)_{i\in I}$ of closed and convex property sets. The solution set for the problem is the feasibility set $S = \bigcap_{i\in I} S_i$. The set theoretic formulation $(\Xi, (S_i)_{i\in I})$ is consistent, i.e., $S \neq \varnothing$. The associated proximity function is the (continuous and convex) functional

$$\begin{aligned} \Phi : \Xi &\to \mathbb{R}_+ \\ a &\mapsto \frac{1}{2}\sum_{i\in I} w_i d(a, S_i)^2 \end{aligned} \tag{9}$$

where

$$\sum_{i\in I} w_i = 1 \text{ and } (\forall i \in I)\ w_i > 0. \tag{10}$$

In words, $\Phi(a)$ measures the degree of infeasibility of an image a. Note that $\Phi(a) = 0 \quad \Leftrightarrow \quad a \in S$. The operators of projection onto the sets $(S_i)_{i\in I}$ are denoted by $(P_i)_{i\in I}$.

III. Projection Methods in Image Recovery

In this section, we give a brief account of the projection methods which have been used in convex set theoretic signal and image recovery (more details can be found in [13]). We assume here that $(S_i)_{i\in I}$ is a finite family of m sets.

A. POCS: Periodic Projections onto Convex Sets

In image recovery, the iterative scheme (3) was first proposed in [23] under the name algebraic reconstruction technique (ART) to find a finite dimensional image in the intersection of affine hyperplanes. In its general form, POCS is described by the algorithm

$$(\forall n \in \mathbb{N})\ \ a_{n+1} = a_n + \lambda_n(P_{n\,(\mathrm{modulo}\,m)+1}(a_n) - a_n) \tag{11}$$

where the relaxation parameters satisfy

$$(\forall n \in \mathbb{N})\ \ \varepsilon \leq \lambda_n \leq 2 - \varepsilon \quad \text{with} \ \ 0 < \varepsilon < 1. \tag{12}$$

The convergence properties of POCS are discussed in the classical paper [24]. While [29] seems to be the first general image recovery application of (11)–(12), the popularity of the method owes much to the expository work [44]. As was mentioned in Section I, a serious drawback of POCS is slow convergence, which is illustrated in Fig. 2: As the angle between the two sets decreases, the progression of the iterates becomes extremely slow. This so-called “angle problem” of POCS had already been pointed out in [24]. POCS has been used mostly in the unrelaxed form (3), i.e., $\lambda_n = 1$ in (11); however, each iteration can be either underrelaxed, i.e., $\lambda_n \leq 1$, or overrelaxed, i.e., $\lambda_n \geq 1$. Unfortunately, this flexibility cannot be exploited to accelerate the iterations. Thus, even in the simple case of affine half-spaces, there is no systematic answer as to whether underrelaxations are faster than overrelaxations or vice-versa [25], [30]. Likewise, in the studies reported in [41], only heuristic rules for specific problems are given.

[3] Meaning $\varnothing \neq I \subset \mathbb{N}$.

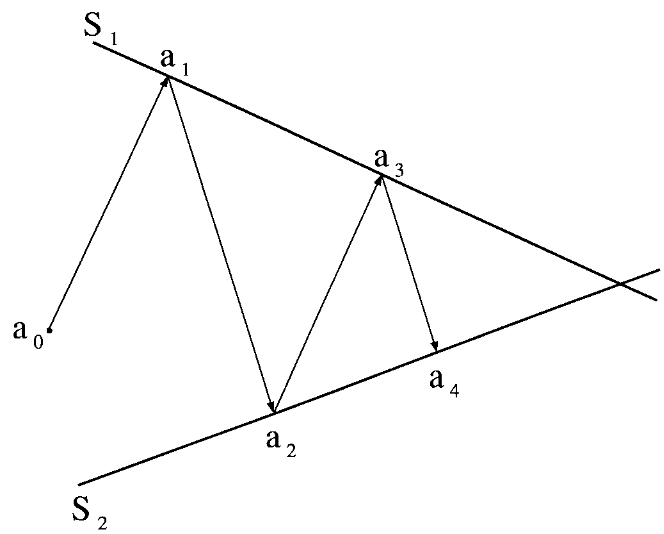

Fig. 2. POCS algorithm.

B. SIRT: Simultaneous Iterative Reconstruction Technique

The simultaneous iterative reconstruction technique (SIRT) was developed for tomographic image reconstruction in [22]. In this method, which can be regarded as the parallel counterpart of ART, the projections of the current iterate onto all the sets (hyperplanes in this case) are averaged to form the update, that is

$$(\forall n \in \mathbb{N})\ \ a_{n+1} = \frac{1}{m}\sum_{i\in I} P_i(a_n). \tag{13}$$

It was soon recognized that, although SIRT gave better results than ART in noisy environments, it did not converge as fast [2], [25]. The fact that SIRT can be slower than POCS is also reported in [43] and illustrated in Fig. 3.

C. PPM: Parallel Projection Method

The parallel projection method (PPM) is a generalization of SIRT governed by the recursion

$$(\forall n \in \mathbb{N})\ \ a_{n+1} = a_n + \lambda_n(\sum_{i\in I} w_i P_i(a_n) - a_n) \tag{14}$$

where (10) and (12) are in force. It was developed for inconsistent feasibility problems in [14]. It was shown there that, when $S = \varnothing$, POCS gives poor solutions while PPM converges to a minimizer of the proximity function Φ of (9), i.e., it yields a weighted least-squares solution.[4] In consistent problems, of course, PPM solves (2) and it converges faster with overrelaxations [14].

D. MOPP: Method of Parallel Projections

Although PPM is quite useful for inconsistent problems, it is not very flexible as a parallel method in that it requires that all the sets be activated at each iteration. As a result, if the number of sets exceeds the number of concurrent processors available,

[4] This, in passing, explains the better behavior of SIRT compared to ART in noisy tomographic reconstruction problems [2], [25], as noisy data often give rise to nonintersecting families of hyperplanes.

the implementation will not be optimal. In order to efficiently spread the computational load of each iteration among the processors and obtain a truly parallel algorithm, it is desirable to have the possibility of activating variable subfamilies of sets. The method of parallel projections (MOPP) meets this requirement. It is described by the algorithm

$$(\forall n \in \mathbb{N}) \;\; a_{n+1} = a_n + \lambda_n \Big(\sum_{i \in I_n} w_{i,n} P_i(a_n) - a_n\Big) \qquad (15)$$

where (12) is in force, where the control sequence $(I_n)_{n \geq 0}$ imposes that every set be activated at least once over any cycle of M consecutive iterations, i.e.,

$$(\forall n \in \mathbb{N}) \;\; \varnothing \neq I_n \subset I \;\text{ and }\; I = \bigcup_{k=n}^{n+M-1} I_k \qquad (16)$$

and where the weights on the projections satisfy

$$(\forall n \in \mathbb{N}) \;\; \begin{cases} \sum_{i \in I_n} w_{i,n} = 1 \\ (\forall i \in I_n) \; w_{i,n} \geq \delta 1_{\complement S_i}(a_n) \end{cases} \qquad (17)$$

for some $\delta \in]0, 1/m]$. MOPP contains as special cases the previous algorithms. Thus, POCS is obtained by letting $(\forall n \in \mathbb{N}) \quad I_n = \{n \,(\mathrm{modulo}\, m) + 1\}$, whereas PPM is obtained by letting $(\forall n \in \mathbb{N}) \;\; I_n = I$ and $(\forall i \in I_n) \;\; w_{i,n} = w_i$, where $(w_i)_{i \in I}$ satisfies (10). MOPP also generalizes the accelerated nonlinear Cimmino algorithm (ANCA) of [26], which is obtained by letting

$$(\forall n \in \mathbb{N}) \begin{cases} \lambda_n = 1; \\ I_n = \{i \in I \mid a_n \notin S_i\}; \\ w_{i,n} = \begin{cases} w_i / \sum_{j \in I_n} w_j & \text{if card}\, I_n \geq 2 \\ w_i & \text{otherwise.} \end{cases} \end{cases} \qquad (18)$$

ANCA was studied for finite dimensional spaces in [26], where it was shown to be faster than SIRT. A general study of the convergence properties of MOPP is presented in [16]. The first study of recursions of type (15) in Hilbert spaces was provided in [33], and a review of the medical imaging applications of parallel algorithms that process blocks of constraints over the iterations is given in [9].

E. Discussion

Intuitively, it would seem that a parallel projection algorithm is numerically more efficient than a serial one such as POCS since the projections can be processed simultaneously as opposed to sequentially. Unfortunately, this is not always the case for unrelaxed methods such as SIRT which are often slower than POCS. On the positive side, an asset of parallel projection methods is that they can be accelerated via overrelaxations, as reported in various experimental and theoretical studies [7], [14], [16], [19], [26], [35]. Since the relaxation parameters in MOPP are confined to the interval $]0, 2[$, this suggests that even greater accelerations could be achieved by pushing the relaxations beyond 2. A key step to prove the convergence of the various methods that have been proposed to solve (2) is to establish the so-called Fejér-monotonicity property

$$(\forall n \in \mathbb{N})(\forall c \in S) \;\; \|a_{n+1} - c\| \leq \|a_n - c\|. \qquad (19)$$

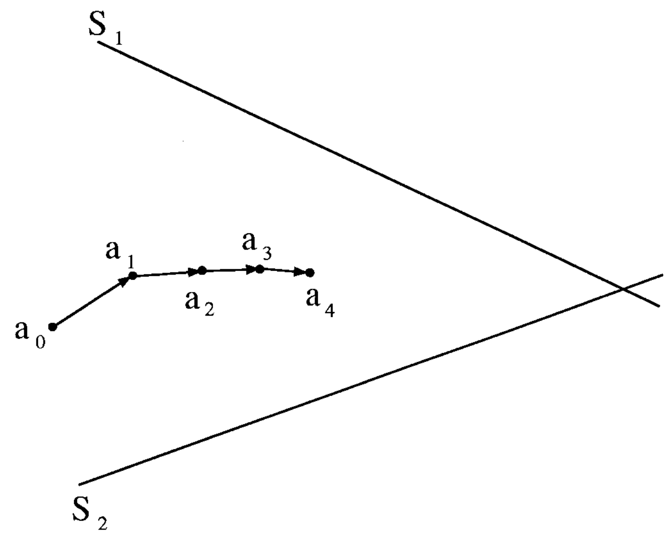

Fig. 3. SIRT algorithm.

When only one set $S_{i(n)}$ is activated at iteration n, as in (11), (19) implies that a_{n+1} cannot lie beyond the reflection $2P_{i(n)}(a_n) - a_n$ of a_n with respect to $S_{i(n)}$, which imposes $\lambda_n \leq 2$. On the other hand, with a parallel scheme such as (15) where several sets are activated simultaneously, one can contemplate the possibility of extrapolating the relaxations beyond 2 and still maintain (19). The question of determining the relaxation range allowable at each iteration is addressed in the next section.

IV. Convex Feasibility in a Product Space

A. Preamble

In this section, card $I = m < +\infty$. $\boldsymbol{\Xi} = \Xi^m$ is the m-fold Cartesian product of the original image space Ξ and is structured as a Hilbert space with the scalar product $\langle\langle \mathbf{a} \mid \mathbf{b} \rangle\rangle = \sum_{i \in I} w_i \langle a_i \mid b_i \rangle$, where $(w_i)_{i \in I}$ is as in (10) and where $\mathbf{a} = (a_i)_{i \in I}$ is a generic m-tuple of images in $\boldsymbol{\Xi}$. The associated norm and distance are denoted by $|||\cdot|||$ and $\mathbf{d}$, respectively. In optimization theory, the product space formalism has been used to decompose minimization problems with multiple constraints into a sequence of elementary problems with a single constraint [8]. The formulation of the feasibility problem (2) as a two-set problem in $\boldsymbol{\Xi}$ is due to Pierra [35] and was used in [14] to solve inconsistent signal feasibility problems. It also provides a convenient framework to develop extrapolated projection methods.

B. EPPM: Extrapolated Parallel Projection Method

Following Pierra [35], we first observe that in the product space $\boldsymbol{\Xi}$, the original feasibility problem (2) is equivalent to the simpler two-set problem

$$\text{Find } \mathbf{a}^\star \in \mathbf{S} \bigcap \mathbf{D} \qquad (20)$$

where $\mathbf{S} = \times_{i \in I} S_i = \{\mathbf{a} \in \boldsymbol{\Xi} \mid (\forall i \in I) \quad a_i \in S_i\}$ is the Cartesian product of the property sets and $\mathbf{D} =$

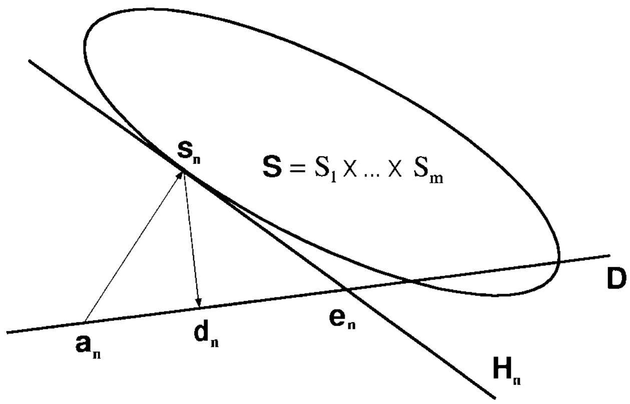

Fig. 4. EPPM algorithm in the product space.

$\{(a, \cdots, a) \in \mathbf{\Xi} \mid a \in \Xi\}$ the diagonal vector subspace. Indeed, $\mathbf{S} \bigcap \mathbf{D} = \{(a, \cdots, a) \in \mathbf{\Xi} \mid (\forall i \in I) \quad a \in S_i\} = \{(a, \cdots, a) \in \mathbf{\Xi} \mid a \in \bigcap_{i \in I} S_i\}$. To solve (20), fix $\mathbf{a}_0 \in \mathbf{D}$ and construct a sequence $(\mathbf{a}_n)_{n \geq 0} \subset \mathbf{D}$ via the alternating projections scheme

$$(\forall n \in \mathbb{N}) \;\; \mathbf{a}_{n+1} = \mathbf{a}_n + \lambda_n (P_{\mathbf{D}} \circ P_{\mathbf{S}}(\mathbf{a}_n) - \mathbf{a}_n). \tag{21}$$

Since this algorithm is a particular instance of POCS, we obtain immediately the following result.

Proposition 1: Every sequence $(\mathbf{a}_n)_{n \geq 0}$ constructed as in (21) with relaxation strategy (12) converges weakly to a point in $\mathbf{S} \bigcap \mathbf{D}$. □

In order to define an alternative relaxation strategy consider Fig. 4, where $\mathbf{s}_n = P_{\mathbf{S}}(\mathbf{a}_n)$ and $\mathbf{d}_n = P_{\mathbf{D}}(\mathbf{s}_n) = P_{\mathbf{D}} \circ P_{\mathbf{S}}(\mathbf{a}_n)$. Let $\mathbf{H}_n$ be the affine hyperplane supporting $\mathbf{S}$ at $\mathbf{s}_n$. Then $\mathbf{H}_n$ separates a_n from $\mathbf{S}$ and intersects $\mathbf{D}$ at $\mathbf{e}_n$. Note that $\mathbf{a}_{n+1} = \mathbf{e}_n$ is attained in (21) by letting λ_n take the value of the extrapolation coefficient

$$L_n = \frac{|||\mathbf{e}_n - \mathbf{a}_n|||}{|||\mathbf{d}_n - \mathbf{a}_n|||} = \frac{|||\mathbf{s}_n - \mathbf{a}_n|||^2}{|||\mathbf{d}_n - \mathbf{a}_n|||^2} = \frac{|||P_{\mathbf{S}}(\mathbf{a}_n) - \mathbf{a}_n|||^2}{|||P_{\mathbf{D}} \circ P_{\mathbf{S}}(\mathbf{a}_n) - \mathbf{a}_n|||^2}. \tag{22}$$

In addition, any update $\mathbf{a}_{n+1}$ on the segment $]\mathbf{a}_n, \mathbf{e}_n]$ is closer to any point in the solution set $\mathbf{S} \bigcap \mathbf{D}$ than $\mathbf{a}_n$, i.e., the Fejér monotonicity property (19) is satisfied. This suggests the relaxation scheme

$$(\forall n \in \mathbb{N}) \;\; \varepsilon \leq \lambda_n \leq L_n \quad \text{with } 0 < \varepsilon < 1. \tag{23}$$

Proposition 2: [35] Every sequence $(\mathbf{a}_n)_{n \geq 0}$ constructed as in (21) with relaxation strategy (23) converges weakly to a point in $\mathbf{S} \bigcap \mathbf{D}$. □

To recast these results in the original image space Ξ, note that [35]

$$\begin{cases} (\forall \mathbf{a} \in \mathbf{D}) \; P_{\mathbf{S}}(\mathbf{a}) = (P_i(a))_{i \in I}, \\ (\forall \mathbf{a} \in \mathbf{S}) \; P_{\mathbf{D}}(\mathbf{a}) = (\sum_{i \in I} w_i a_i, \cdots, \sum_{i \in I} w_i a_i) \end{cases} \tag{24}$$

and that the mapping $\mathbf{a} \mapsto a$ defines an isomorphism from $\mathbf{D}$ into Ξ. Hence, (21) in the product space $\mathbf{\Xi}$ yields (14) in the original space Ξ and Proposition 1 implies the weak convergence of PPM.

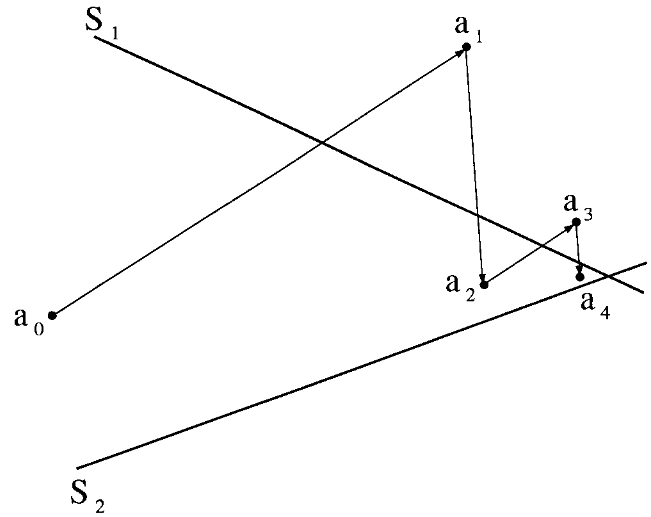

Fig. 5. EPPM algorithm in the original space.

Proposition 3: Every orbit $(a_n)_{n \geq 0}$ of PPM converges weakly to a point in S. □

On the other hand, thanks to (24), (22) and (23) become

$$(\forall n \in \mathbb{N}) \;\; \varepsilon \leq \lambda_n \leq L_n \tag{25}$$

where

$$L_n = \frac{\sum_{i \in I} w_i \|P_i(a_n) - a_n\|^2}{\| \sum_{i \in I} w_i P_i(a_n) - a_n\|^2}. \tag{26}$$

By coupling (14) with (25), we obtain Pierra's extrapolated parallel projection method (EPPM), whose weak convergence follows from Proposition 2.

Proposition 4: [35] Every orbit $(a_n)_{n \geq 0}$ of EPPM converges weakly to a point in S. □

It was observed in [35] that the fast convergence of EPPM was due to the large overrelaxations allowed by (25), as L_n can attain values much larger than 2 and eliminate the "angle problem" of the methods of Section III. For the same problem as in Figs. 2 and 3, Fig. 5 shows the initial portion of an orbit of EPPM obtained with $w_1 = w_2 = 1/2$ and $(\forall n \in \mathbb{N}) \quad \lambda_n = L_n$. Besides fast convergence, this figure also reveals that the orbit has a tendency to "zig-zag," which reduces the effectiveness of the algorithm. To mitigate this phenomenon, it was suggested in [35] to recenter the orbit every three iterations by halving the extrapolations, namely

$$(\forall n \in \mathbb{N}) \;\; \lambda_n = \begin{cases} L_n/2 & \text{if } n = 2 \text{ modulo } 3 \\ L_n & \text{otherwise.} \end{cases} \tag{27}$$

This amounts to taking a smaller step every three iterations, which places the corresponding iterate in a more central position with respect to all the sets than a full extrapolation would.

C. EPPM2: A Generalization of EPPM

In this section we extend EPPM in two directions. We first note that each iteration of EPPM requires the computation of m exact projections. As such constrained quadratic minimization subproblems are often difficult to solve, approximate projections will be employed. Next, we observe that EPPM does not generalize PPM in that the extrapolation parameter L_n is certainly at least equal to 1 thanks to the convexity of $\|\cdot\|^2$ in (26), but not necessarily greater than 2. Hence, to unify and extend both PPM and EPPM, relaxations up to $2L_n$ will be considered. A practical consequence of this extension will be to achieve faster convergence in certain problems through the use of larger overrelaxations than those allowed by PPM and EPPM.

We first generalize (21) by replacing the exact projection of $\mathbf{a}_n$ onto $\mathbf{S}$ by an approximate projection $P_{\mathbf{S}_n}(\mathbf{a}_n)$, namely

$$(\forall n \in \mathbb{N})\ \ \mathbf{a}_{n+1} = \mathbf{a}_n + \lambda_n(P_{\mathbf{D}} \circ P_{\mathbf{S}_n}(\mathbf{a}_n) - \mathbf{a}_n) \tag{28}$$

where $\mathbf{a_0} \in \mathbf{D}$ and where $(\mathbf{S}_n)_{n\geq 0}$ is a sequence of closed and convex subsets of $\mathbf{\Xi}$ such that

$$(\forall n \in \mathbb{N})\ \ \mathbf{S} \subset \mathbf{S}_n \quad \text{and} \quad \mathbf{a}_n \notin \mathbf{S}_n \setminus \mathbf{S}. \tag{29}$$

In words, $\mathbf{S}_n$ is a superset of $\mathbf{S}$, which contains $\mathbf{a}_n$ only when $\mathbf{S}$ does. Next, to double the relaxation ranges, we replace (23) by

$$(\forall n \in \mathbb{N})\ \ \varepsilon \leq \lambda_n \leq (2-\varepsilon)L_n \tag{30}$$

where

$$L_n = \frac{|||P_{\mathbf{S}_n}(\mathbf{a}_n) - \mathbf{a}_n|||^2}{|||P_{\mathbf{D}} \circ P_{\mathbf{S}_n}(\mathbf{a}_n) - \mathbf{a}_n|||^2}. \tag{31}$$

Following [5], we shall say that algorithm (28)–(31) is focusing if for every suborbit $(\mathbf{a}_{n_k})_{k\geq 0}$ it generates we have

$$\begin{cases} \text{weak.}\lim_{k\to+\infty} \mathbf{a}_{n_k} = \mathbf{a} \\ \lim_{k\to+\infty} \mathbf{d}(\mathbf{a}_{n_k}, \mathbf{S}_{n_k}) = 0 \end{cases} \Rightarrow \quad \mathbf{a} \in \mathbf{S}. \tag{32}$$

We observe that, by construction, $(\mathbf{a}_n)_{n\geq 0}$ lies in $\mathbf{D}$. Since $\mathbf{D}$ is weakly closed, $\mathbf{a} \in \mathbf{D}$ in (32).

Theorem 1: If algorithm (28)–(31) is focusing, then every orbit $(\mathbf{a}_n)_{n\geq 0}$ it generates converges weakly to a point in $\mathbf{S} \bigcap \mathbf{D}$. □

Algorithm (28)–(31) in $\mathbf{\Xi}$ yields a parallel projection method in Ξ that we shall call EPPM2. By virtue of (24), EPPM2 is defined by the recursion

$$(\forall n \in \mathbb{N})\ \ a_{n+1} = a_n + \lambda_n\Big(\sum_{i\in I} w_i P_{i,n}(a_n) - a_n\Big) \tag{33}$$

where

$$(\forall n \in \mathbb{N})\ \ \varepsilon \leq \lambda_n \leq (2-\varepsilon)L_n \tag{34}$$

with

$$L_n = \frac{\sum_{i\in I} w_i \|P_{i,n}(a_n) - a_n\|^2}{\|\sum_{i\in I} w_i P_{i,n}(a_n) - a_n\|^2} \tag{35}$$

and where, for every $i \in I$, $(P_{i,n})_{n\geq 0}$ is a sequence of projection operators onto closed and convex sets $(S_{i,n})_{n\geq 0}$ such that

$$(\forall n \in \mathbb{N})\ \ S_i \subset S_{i,n} \quad \text{and} \quad a_n \notin S_{i,n} \setminus \mathrm{S}_{\mathbf{i}}. \tag{36}$$

Moreover, in view of (32), EPPM2 will be focusing if, for every suborbit $(a_{n_k})_{k\geq 0}$ it generates and every $i \in I$, we have[5]

$$\begin{cases} \text{weak.}\lim_{k\to+\infty} a_{n_k} = a \\ \lim_{k\to+\infty} d(a_{n_k}, S_{i,n_k}) = 0 \end{cases} \Rightarrow \quad a \in S_i. \tag{37}$$

The following statement, which is a direct consequence of Theorem 1, generalizes Propositions 3 and 4.

Theorem 2: If EPPM2 is focusing, then every orbit $(a_n)_{n\geq 0}$ it generates converges weakly to a point in S. □

D. Discussion

The advantage of EPPM2 over MOPP resides in its ability to use approximate projections and larger relaxations, which means that EPPM2 converges in fewer iterations and that the computational cost of each iteration is lower. On the other hand, MOPP is more flexible than EPPM2 in that it can process a variable number of sets at each iteration. In the next section, we introduce a general projection scheme (EMOPSP), which combines the advantageous features of MOPP and EPPM2. Moreover, it employs subgradient projections, which provides a simple way of explicitly computing the approximate projections. The relationships between EMOPSP and the methods discussed so far are shown below.

$$\begin{array}{ccccccc} \text{ART} & & & & \text{SIRT} & & \\ \downarrow & & & & \downarrow & \searrow & \\ \text{POCS} & & \text{ANCA} & & \text{PPM} & & \text{EPPM} \\ & \searrow & \downarrow & \swarrow & & \searrow & \downarrow \\ & & \text{MOPP} & & & & \text{EPPM2} \\ & & & \searrow & & \swarrow & \\ & & & & \text{EMOPSP} & & \end{array} \tag{38}$$

V. Extrapolated Method of Parallel Subgradient Projections (EMOPSP)

A. Subgradient Projections

Each convex constraint Ψ_i is usually formulated through a convex inequality and the associated property set S_i can be written as the 0-section

$$S_i = \text{sec}(g_i, 0) = \{a \in \Xi \mid g_i(a) \leq 0\} \tag{39}$$

of a convex, (lower semi-)continuous functional $g_i : \Xi \to \mathbb{R}$. This representation of a property set is in fact quite general as one can certainly choose $g_i = d(\cdot, S_i)$. The projection $P_i(a_n)$ of an image $a_n \in \complement S_i$ is typically obtained by solving

$$\min_{b\in\Xi} \frac{1}{2}\|a_n - b\|^2 \ \text{ subject to } g_i(b) = 0. \tag{40}$$

In some instances, this program is easily solved and admits a closed-form solution, e.g., [44]. In many cases, however, the exact projection operators are not known, e.g., [11], [40], [42]. In Section IV-C, projections onto approximating supersets were proposed to circumvent the computation of exact projections. A natural choice for the approximating superset $S_{i,n}$ is an affine half-space containing S_i but not a_n. $P_{i,n}(a_n)$ is then simply the projection onto the hyperplane

[5] Note that, in particular, (37) holds if $\lim_{n\to+\infty} d(a_n, S_{i,n}) = 0 \Rightarrow \lim_{n\to+\infty} d(a_n, S_i) = 0$.

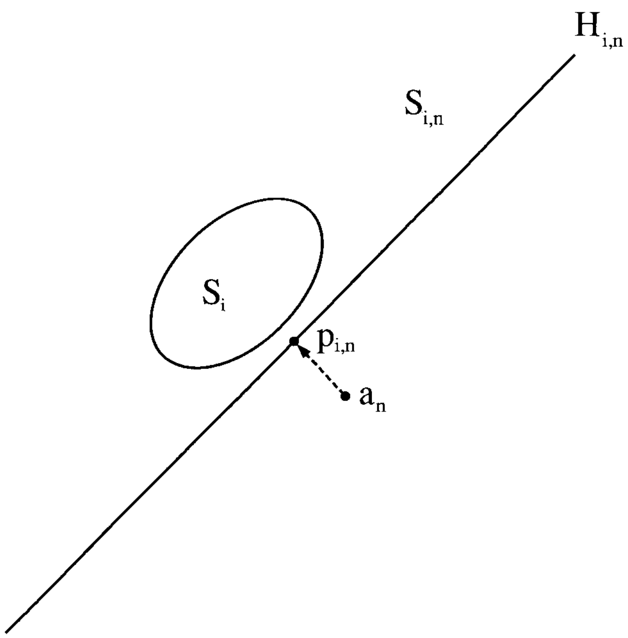

Fig. 6. Projection onto a separating hyperplane.

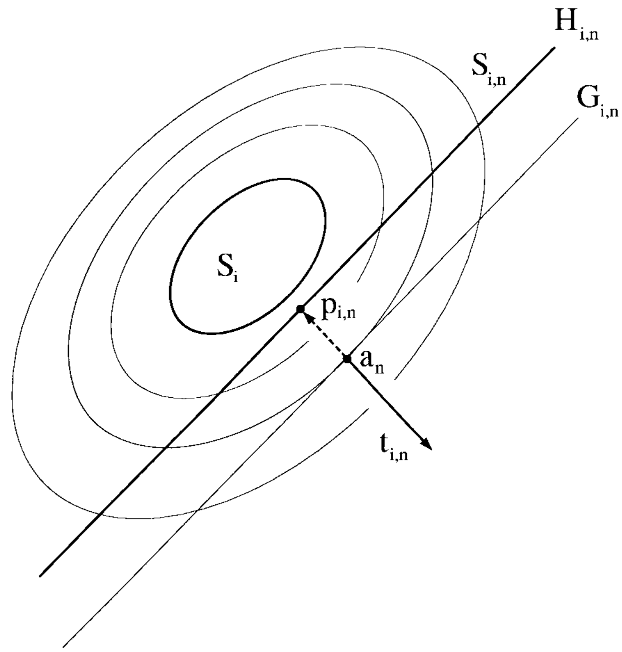

Fig. 7. Subgradient projection. This figure shows various level curves lev(g_i, η) of g_i. The vector $t_{i,n}$ is a subgradient (unique here) of g_i at a_n. Note that $t_{i,n}$ is normal to sec$(g_i, g_i(a_n))$ at a_n: $(\forall a \in \sec(g_i, g_i(a_n)))\ \langle a - a_n \mid t_{i,n}\rangle \leq 0$. Indeed, $a \in \sec(g_i, g_i(a_n)) \Rightarrow g_i(a) - g_i(a_n) \leq 0$ and, therefore, (7) $\Rightarrow \langle a - a_n \mid t_{i,n}\rangle \leq 0$. $G_{i,n} = \{a \in \Xi \mid \langle a_n - a \mid t_{i,n}\rangle = 0\}$ is a hyperplane tangent to sec$(g_i, g_i(a_n))$ at a_n. $H_{i,n} = \{a \in \Xi \mid \langle a_n - a \mid t_{i,n}\rangle = g_i(a_n)\}$ is a hyperplane parallel to $G_{i,n}$ and separating S_i and a_n. It delimits the half-space $S_{i,n}$ of (41), which contains S_i but not a_n. The subgradient projection of a_n onto S_i is the projection $p_{i,n}$ of a_n onto $H_{i,n}$.

$H_{i,n}$ delimiting $S_{i,n}$ and which separates a_n and S_i, as shown in Fig. 6. The nonlinear problem (40) thus becomes an affine one. Methods involving projections onto separating hyperplanes have been proposed previously in connection with less general projection algorithms in [1] and [21]. A practical concern with this conceptually simple approach is to determine efficiently the separating hyperplane $H_{i,n}$. We shall now see that, thanks to (39), the fundamental inequality defining subgradients in (7) can be used to determine $H_{i,n}$ and $S_{i,n}$ explicitly.

Consider the half-space

$$S_{i,n} = \{a \in \Xi \mid \langle a_n - a \mid t_{i,n}\rangle \geq g_i(a_n)\} \tag{41}$$

where $t_{i,n} \in \partial g_i(a_n)$. Notice that $a_n \notin S_i \Rightarrow g_i(a_n) > 0 \Rightarrow a_n \notin S_{i,n}$. Moreover, if we take $a \in S_i$, then $g_i(a) \leq 0$ and, by (7), $t_{i,n} \in \partial g_i(a_n) \Rightarrow \langle a_n - a \mid t_{i,n}\rangle \geq g_i(a_n) - g_i(a) \geq g_i(a_n) \Rightarrow a_n \in S_{i,n}$. Hence, $S_i \subset S_{i,n}$. We conclude that $S_{i,n}$ is a valid approximation of S_i in the sense of (36). From (5) and (41), the projection of $a_n \in \complement S_i$ onto $S_{i,n}$ is then simply given by

$$P_{i,n}(a_n) = a_n - \frac{g_i(a_n)}{\|t_{i,n}\|^2} t_{i,n} \tag{42}$$

and is called a *subgradient projection*. This process is illustrated geometrically in Fig. 7. Thus, only the computation of a subgradient $t_{i,n}$ is required to activate the set S_i instead of the exact projection $P_i(a_n)$. In practice, g_i will often be differentiable, so that $t_{i,n} = \nabla g_i(a_n)$. When $P_i(a_n)$ is tractable, one can take $g_i = d(\cdot, S_i)$ and (42) yields the exact projection thanks to (8). Whence, upon defining the subgradient projection of an arbitrary point $a_n \in \Xi$ by

$$P_{i,n}(a_n) = \begin{cases} a_n - \dfrac{g_i(a_n)}{\|t_{i,n}\|^2} t_{i,n} & \text{if } a_n \in \complement S_i, \\ a_n & \text{otherwise} \end{cases} \tag{43}$$

where $t_{i,n} \in \partial g_i(a_n)$, we obtain a generalization of the notion of projection.

B. Algorithm

The algorithm we propose here has a structure similar to that introduced in [15] to construct common fixed points of firmly nonexpansive operators.

Given an initial point $a_0 \in \Xi$ and numbers $C \in \mathbb{N}^*$, $\delta \in]0, 1/C[$, and $\varepsilon \in]0, 1[$, EMOPSP is defined by the iterative process

$$(\forall n \in \mathbb{N})\ \ a_{n+1} = a_n + \lambda_n\Big(\sum_{i \in I_n} w_{i,n} P_{i,n}(a_n) - a_n\Big) \tag{44}$$

where at each iteration n

- the family I_n of indices of selected sets satisfies
$$\varnothing \neq I_n \subset I \quad \text{and} \quad \text{card}\,\{i \in I_n \mid a_n \notin S_i\} \leq C \tag{45}$$
- the subgradient projections $(P_{i,n}(a_n))_{i \in I_n}$ are defined by (43);
- the aggregating weights $(w_{i,n})_{i \in I_n}$ conform to (17);
- the relaxation parameter λ_n lies in $[\varepsilon, (2 - \varepsilon)L_n]$, where

$$L_n = \begin{cases} \dfrac{\sum_{i \in I_n} w_{i,n} \|P_{i,n}(a_n) - a_n\|^2}{\|\sum_{i \in I_n} w_{i,n} P_{i,n}(a_n) - a_n\|^2} & \text{if } a_n \notin \bigcap_{i \in I_n} S_i, \\ 1 & \text{otherwise.} \end{cases} \tag{46}$$

At iteration n the image a_n is given and the updating process is performed as follows. First, one selects the subfamily $(S_i)_{i\in I_n}$ of sets to be activated. Then, one takes subgradients $(t_{i,n})_{i\in I_n}$ of $(g_i)_{i\in I_n}$ at a_n and computes simultaneously the subgradient projections $(P_{i,n}(a_n))_{i\in I_n}$. Next, one forms a convex combination $d_n = \sum_{i\in I_n} w_{i,n}P_{i,n}(a_n)$ of these projections and computes the extrapolation parameter L_n. The position of the new iterate a_{n+1} on the ray emanating from a_n and going through d_n is determined by the positive relaxation parameter λ_n, which can take values up to $2L_n$. The set S_i will be said to be violated at iteration n if $a_n \notin S_i$. Since nonviolated sets can be assigned a weight of 0 by (17), they can always be considered as selected. When only one set is violated (*a fortiori* when only one set is selected), then $L_n = 1$ and the relaxation range reverts to the conventional interval $[\varepsilon, 2-\varepsilon]$. Therefore, extrapolations can take place only when $\operatorname{card}\{i \in I_n \mid a_n \notin S_i\} \geq 2$.

Proposition 5: Every orbit $(a_n)_{n\geq 0}$ of EMOPSP satisfies (19). □

Thus, every iteration of EMOPSP brings the update closer to any solution. This is an important property since, in practice, the algorithm will be interrupted after a finite number of steps, when some stopping criterion is satisfied.

C. Control

The control sequence $(I_n)_{n\geq 0}$ determines the subfamilies of sets which are processed at each iteration. Naturally, for the iterates to converge to a solution of (2), suitable conditions must be imposed to ensure that every set is activated repeatedly. We shall say that the control is

- *serial* if

$$(\forall n \in \mathbb{N})\ \ \operatorname{card} I_n = 1 \tag{47}$$

- *static* if

$$(\forall n \in \mathbb{N})\ \ I_n = I \tag{48}$$

- *cyclic* if

$$(\exists M \in \mathbb{N}^*)(\forall n \in \mathbb{N})\ \ I = \bigcup_{k=n}^{n+M-1} I_k \tag{49}$$

- *admissible* if

$$(\exists (M_i)_{i\in I} \subset \mathbb{N}^*)(\forall (i,n) \in I \times \mathbb{N})\ \ i \in \bigcup_{k=n}^{n+M_i-1} I_k \tag{50}$$

- *chaotic* if

$$(\forall n \in \mathbb{N})\ \ I = \bigcup_{k\geq n} I_k. \tag{51}$$

Under static control, all the sets must be processed at each iteration, whereas under cyclic control all the sets must be used at least once within any M consecutive iterations. For instance, we have seen that EPPM2 operates under static control and MOPP under cyclic control. These control modes are restricted to finite families of sets, since all the sets must be activated over a finite number of iterations. On the other hand, under admissible control, a countably infinite number of sets can be handled. It requires that, for every $i \in I$, the set S_i be activated at least once within any M_i consecutive iterations. When $\operatorname{card} I < +\infty$, the admissible control mode coincides with the cyclic mode with $M = \max_{i\in I} M_i$. Following is an example of admissible control sequence with $I = \mathbb{N}^*$ and $\operatorname{card} I_n = 2$, and where the sets with indices $2i+1$ and $2i+2$ are activated every 2^{i+1} iterations.

$$\begin{aligned}(I_n)_{n\geq 0} = (&\{1,2\},\{3,4\},\{1,2\},\{5,6\},\{1,2\},\\ &\{3,4\},\{1,2\},\{7,8\},\{1,2\},\{3,4\},\\ &\{1,2\},\{5,6\},\{1,2\},\{3,4\},\{1,2\},\\ &\{9,10\},\{1,2\},\{3,4\},\{1,2\},\{5,6\},\\ &\{1,2\},\{3,4\},\{1,2\},\{7,8\},\{1,2\},\\ &\{3,4\},\{1,2\},\{5,6\},\{1,2\},\{3,4\},\\ &\{1,2\},\{11,12\},\{1,2\},\{3,4\},\{1,2\},\\ &\{5,6\},\{1,2\},\{3,4\},\{1,2\},\{7,8\},\cdots).\end{aligned}$$

Finally, under chaotic control, every set must be used infinitely often, but in any order. Following is an example of chaotic (but not admissible) control sequence with $I = \mathbb{N}^*$ and $\operatorname{card} I_n = 4$.

$$\begin{aligned}(I_n)_{n\geq 0} = (&\{1,2,3,4\},\{1,2,3,4\},\{5,6,7,8\},\\ &\{1,2,3,4\},\{5,6,7,8\},\{9,10,11,12\},\\ &\{1,2,3,4\},\{5,6,7,8\},\{9,10,11,12\},\\ &\{13,14,15,16\},\{1,2,3,4\},\{5,6,7,8\},\\ &\{9,10,11,12\},\{13,14,15,16\},\\ &\{17,18,19,20\},\cdots).\end{aligned}$$

We have: static $\Rightarrow$ cyclic $\Rightarrow$ admissible $\Rightarrow$ chaotic.

D. Convergence

We now present our main convergence results relevant to the theory and the applications of convex set theoretic image recovery. Recall that the family $(S_i)_{i\in I}$ is finite or countable and that it is defined as in (39), where $(g_i)_{i\in I}$ is a family of real-valued, convex, (lower semi-)continuous functions. These functions are therefore subdifferentable and we shall say that their subdifferentials are locally uniformly bounded if

$$(\forall \gamma \in \mathbb{R}_+^*)(\exists \zeta \in \mathbb{R}_+^*)(\forall i \in I)(\forall a \in B(0,\gamma))\partial g_i(a) \subset B(0,\zeta). \tag{52}$$

As noted in [5], (52) implies that (37) is verified for the half-space (41). Whence, if (52) holds, we obtain at once from Theorem 2 the weak convergence to a point in S of any orbit of EMOPSP with constant weights and static control. Actually, much more is true.

Theorem 3: Suppose that the subdifferentials of $(g_i)_{i\in I}$ are locally uniformly bounded. Then, under admissible control, every orbit $(a_n)_{n\geq 0}$ of EMOPSP converges weakly to a point in S. □

The next theorem pertains to strong convergence under the most flexible type of control, namely chaotic control, at the expense of additional hypotheses.

Fig. 8. Original image.

Fig. 9. Degraded image.

Theorem 4: Suppose that the subdifferentials of $(g_i)_{i\in I}$ are locally uniformly bounded. Then, under chaotic control, every orbit $(a_n)_{n\geq 0}$ of EMOPSP converges strongly to a point in S if either of the following conditions is satisfied:

1) $\overset{\circ}{S} \neq \varnothing$;
2) card $I < +\infty$ and one of the functionals, say g_j, is lower semiboundedly-compact. □

To our knowledge, these results are the most general ones available. Thus, particular cases of Theorem 3 can be found in [5], [6], and [16],[6] while particular cases of Theorem 4(i) can be found in [5] and [33].[7] On the other hand, the following corollary of Theorem 4(ii) generalizes results of [10] and [19].[8]

Corollary 1: Suppose that $\dim \Xi < +\infty$ and card $I < +\infty$. Then, under chaotic control, every orbit $(a_n)_{n\geq 0}$ of EMOPSP converges to a point in S. □

This corollary is of utmost importance for practical digital image recovery applications. Indeed, in such applications, the number of constraints is finite. Furthermore, images are discretized over a bounded domain and therefore represented by a point in the euclidean space. Loosely speaking, Corollary 1 then states that, for any family of convex[9] functionals $(g_i)_{i\in I}$, any sequence generated by EMOPSP converges to a feasible image as long as all the sets are used repeatedly in any order.

[6] [5] considered cyclic control and relaxation range (12); [6] considered exact projections and serial control; [16] considered exact projections, cyclic control, and relaxation range (12).

[7] [5] considered card $I < +\infty$ and relaxation range (12); [33] considered exact projections and relaxation range (23).

[8] [10] considered serial, cyclic control; [19] considered static control. While revising this paper, it came to our attention that Corollary 1 has been established independently in [28].

[9] They are, therefore, continuous since $\dim \Xi < +\infty$.

VI. Numerical Simulations

A. Generalities

In this section, we apply EMOPSP to standard digital image restoration problems in order to provide a numerical illustration of its properties and of its performance compared to conventional methods, especially POCS. We have performed numerical comparisons in a variety of signal and image processing problems and the limited results we present here are quite representative of the performance of EMOPSP.

All images have $N \times N$ pixels ($N = 128$) and will be represented using stacked-vector notations [3]. Ξ is the usual N^2-dimensional euclidean space and the pertinent convergence result is therefore Corollary 1. $\mathfrak{F}$ is the two-dimensional (2-D) discrete Fourier transform (DFT) operator, i.e., $(\forall a \in \Xi)\ \ \mathfrak{F}(a) = \widehat{a}$, where for every (k,l) in $\{0,\cdots,N-1\}^2$

$$\widehat{a}(k,l) = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} a^{(Ni+j)} \exp(-\imath 2\pi(ik+jl)/N). \qquad (53)$$

The original image h of Fig. 8 is degraded by convolutional blur with a uniform 9×9 kernel ℓ and addition of uniform purely white noise u with range $[0,R]$ resulting in a blurred image-to-noise ratio of 35 dB. The degraded image x is shown in Fig. 9. It can be written as $x = Lh + u$, where L is the block-Toeplitz matrix associated with the point spread function ℓ. The problem is to estimate h given x and some *a priori* information about h, ℓ, and u. The first property set $S_1 = (\mathbb{R}_+)^{N^2}$ arises from the nonnegativity of pixel values. Next, it is assumed that the DFT of h is known on one fourth of its support for low frequencies in both directions. The associated property set is $S_2 = \{a \in \Xi \mid \widehat{a}1_K = \widehat{h}1_K\}$, where K contains the set of frequency pairs $\{0,\cdots,N/8-1\}^2$

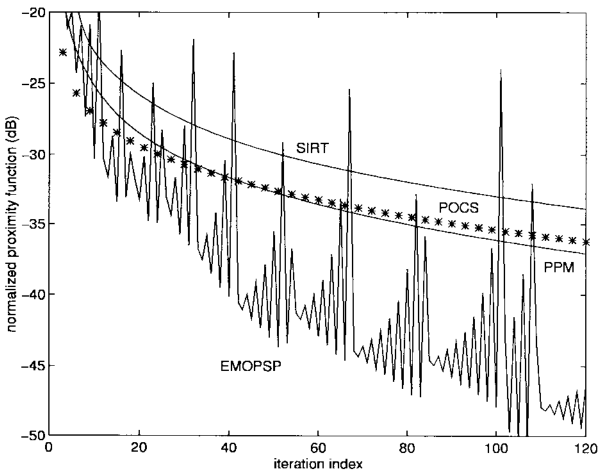

Fig. 10. Scenario 1: EMOPSP versus POCS, SIRT, and PPM.

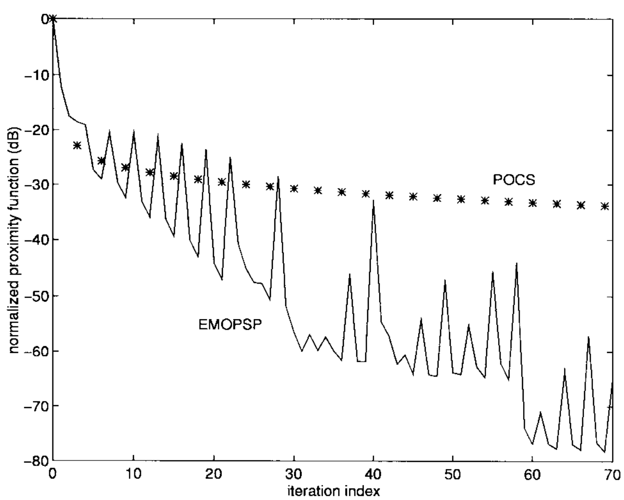

Fig. 11. Scenario 1: EMOPSP with centering versus POCS.

as well as all those resulting from the symmetry properties of the 2-D DFT of real images (a similar set was used in [38]). The projections of an image a_n onto S_1 and S_2 are given by the closed-form expressions

$$\begin{cases} P_1(a_n) = a_n^+ = {}^t[\max\{0, a_n^{(i)}\}]_{0\le i\le N^2-1} \triangleq P_{1,n}(a_n), \\ P_2(a_n) = \mathfrak{F}^{-1}(\widehat{h}1_K + \widehat{a_n}1_{\complement K}) \triangleq P_{2,n}(a_n). \end{cases} \tag{54}$$

To complete the set theoretic formulation, two scenarios will be considered. They both assume knowledge of ℓ but differ in the information available to describe the noise. The first scenario will give rise to a three-set problem in which subgradient projections will be used. The second scenario will give rise to a large scale problem requiring the use of nonstatic control. Every algorithm will be initialized with $a_0 = x$ and the progression of its orbit $(a_n)_{n\ge 0}$ will be tracked by plotting the normalized decibel values $(10\log_{10}(\Phi(a_n)/\Phi(a_0)))_{n\ge 0}$ of the proximity function (9), where

$$(\forall i \in I) \;\; w_i = 1/(\operatorname{card} I). \tag{55}$$

Fig. 12. Scenario 1: restored image.

As a practical stopping rule to compare performance, we shall use the criterion

$$\Phi(a_n) \le 50/\operatorname{card} I. \tag{56}$$

B. Scenario 1: A Three-Set Problem

It is assumed here that the information available about the noise vector u is that its components are independent and all distributed as a random variable U with known second and fourth moments. As shown in [18], with a 95% confidence coefficient, this information leads to the property set

$$S_3 = \{a \in \Xi \mid \|x - La\|^2 \le \rho\} \tag{57}$$

where

$$\rho = N^2\mathsf{E}|U|^2 + 1.96N\sqrt{\mathsf{E}|U|^4 - \mathsf{E}^2|U|^2}. \tag{58}$$

This set has proven quite useful in a number of applications, e.g., [17] and [42], but unfortunately its projection operator must be approximated iteratively via a costly procedure [42]. By contrast, using (43) and the fact that $\nabla g_3(a_n) = \nabla(\|x - La_n\|^2 - \rho) = -2\,{}^tL(x - La_n)$, we simply process the set S_3 at iteration n with the subgradient projection

$$P_{3,n}(a_n) = \begin{cases} a_n + \dfrac{\|r_n\|^2 - \rho}{2\,\|{}^tLr_n\|^2}\,{}^tLr_n & \text{if } \|r_n\|^2 > \rho, \\ a_n & \text{otherwise} \end{cases} \tag{59}$$

where $r_n = x - La_n$. Using standard arguments [3], the upper expression in (59) can be evaluated in the frequency domain efficiently via the 2-D fast Fourier transform (FFT) as

$$P_{3,n}(a_n) = \mathfrak{F}^{-1}\left(\widehat{a_n} + \frac{\|\widehat{r_n}\|^2 - N^2\rho}{2\|\overline{\widehat{\ell}}\,\widehat{r_n}\|^2}\,\overline{\widehat{\ell}}\,\widehat{r_n}\right) \tag{60}$$

where $\widehat{r_n} = \widehat{x} - \widehat{\ell}\,\widehat{a_n}$. The approximate computation of $P_3(a_n)$ proposed in [42] typically requires 10 to 20 iterations of much

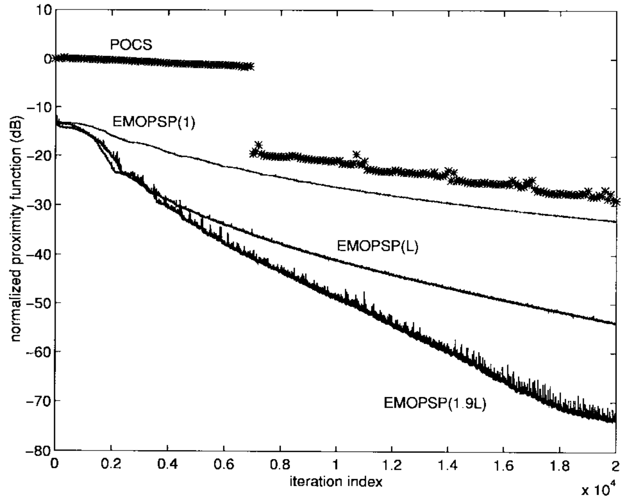

Fig. 13. Scenario 2 with 4 parallel processors.

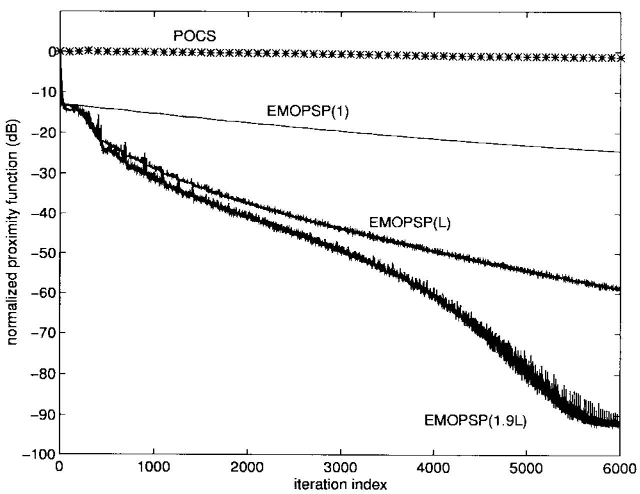

Fig. 14. Scenario 2 with 16 parallel processors.

higher complexity than (60). Consequently, the subgradient projection reduces the cost of processing S_3 by at least an order of magnitude.

The set theoretic formulation for this problem is $(\Xi, (S_i)_{1 \leq i \leq 3})$. In the results shown in Fig. 10, POCS is implemented as in (3), SIRT as in (13), PPM as in (14) with (55) and $(\forall n \in \mathbb{N}) \quad \lambda_n = 1.9$. Furthermore, $P = 3$ parallel processors are available. The -49 dB mark corresponding to (56) was reached by POCS in 1185 iterations. Since only three sets are present, EMOPSP is implemented with static control, fixed weights as in (55), and relaxation strategy $(\forall n \in \mathbb{N}) \quad \lambda_n = L_n$. POCS is faster than SIRT and comparable to PPM, but clearly outperformed by EMOPSP, which uses extrapolated relaxations. Fig. 11 shows that EMOPSP can be further accelerated by using the centering technique (27), resulting in a dramatic improvement over POCS. The restored image obtained in this case appears in Fig 12. Let us observe that these results show performance only in terms of the number of iterations required to reach a given degree of infeasibility. However, to fully appreciate the

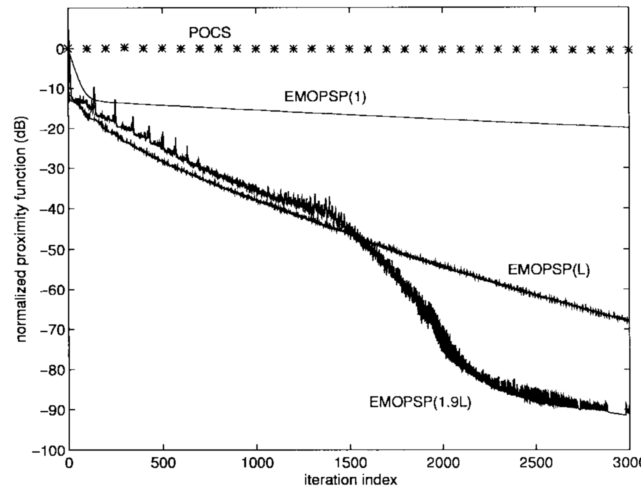

Fig. 15. Scenario 2 with 64 parallel processors.

numerical superiority of EMOPSP, it must be borne in mind that to process the set S_3, POCS, SIRT, and PPM must use the costly projection onto S_3 whereas EMOPSP needs only the approximate projection (59).

C. Scenario 2: A 16 386-Set Problem

We now assume that no probabilistic information is available about the noise vector u and that it is known only that its components lie in $[0, R]$. This information leads to the N^2 property sets [18]

$$\begin{aligned}(\forall i \in \{0, \cdots, N^2-1\})\ & S_{i+3} \\ &= \{a \in \Xi \mid 0 \leq x^{(i)} - \langle a \mid L_i \rangle \leq R\} \end{aligned} \tag{61}$$

where L_i is the ith row of L. According to (5), we have

$$P_{i+3}(a_n) = \begin{cases} a_n + [(x^{(i)} - \langle a_n \mid L_i \rangle)/\|L_i\|^2] L_i \\ \quad \text{if } \langle a_n \mid L_i \rangle > x^{(i)}, \\ a_n + [(x^{(i)} - R - \langle a_n \mid L_i \rangle)/\|L_i\|^2] L_i \\ \quad \text{if } \langle a_n \mid L_i \rangle < x^{(i)} - R, \\ a_n \quad \text{otherwise.} \end{cases} \tag{62}$$

The set theoretic formulation is $(\Xi, (S_i)_{1 \leq i \leq N^2+2})$. Similar problems generating a large number of sets are reported in [34], [39], and [42], where they were solved with POCS. Here, we implement POCS (3) by skipping the nonviolated sets so that each iteration actually produces an update. The -58 dB mark corresponding to (56) was reached by POCS in 76 000 iterations. To implement EMOPSP, computer architectures with $P = 4$, 16, and 64 parallel processors are considered. At each iteration n, the control selects P sets as follows: S_1 and S_2 if they are violated and a block of consecutive violated sets in (61), starting with S_j ($3 \leq j \leq 2 + N^2$ modulo N^2), where S_{j-1} is the last set used at iteration $n-1$. Moreover, three values of λ_n are considered: 1, L_n, and $1.9L_n$. In Figs. 13–15, the corresponding algorithms are labeled as EMOPSP(1), EMOPSP(L), and EMOPSP(1.9L), respectively. POCS starts slowly and approaches the performance of the unrelaxed algorithm EMOPSP(1) after about 7000 iterations. EMOPSP(L) is much faster and EMOPSP(1.9L), which further

Fig. 16. Scenario 2: restored image.

exploits the large relaxation range allowed by our analysis, is even faster. The restoration obtained by EMOPSP(1.9) is shown in Fig. 16.

D. Remarks

EMOPSP is very versatile as all of its parameters can be changed at each iteration (sets selected, approximating supersets, weights on the projections, relaxations). Hence, the above implementations of EMOPSP are somewhat conservative in the sense that they do not fully exploit the flexibility of the method. Although no general conclusion is intended, our intensive simulations with EMOPSP in various problems has revealed the following behavior. When a small number of sets is used, very large extrapolations (say $1.5L_n \leq \lambda_n \leq 1.99L_n$) often create a lot of zig-zagging and are not as effective as the centered extrapolations (27). On the other hand, large extrapolations accelerate the iterations significantly in more sizeable problems. Let us also note that the above results assume that $P > 1$ parallel processors are available. By multiplying the number of iterations needed to obtain a certain level of the proximity function by P, one can easily see that EMOPSP is still faster than POCS in single-processor environments.

We have seen in Section VI-B that the subgradient projection reduced the computational burden associated with the use of the set S_3 in (57) by at least an order of magnitude compared to the projection derived in [42]. Let us add that in [42] the blur was assumed to be space invariant, which made it possible to carry out large matrix inversions efficiently in the frequency domain via circulant approximations. When the blur is space variant, the matrices must be inverted directly which, as noted in [34], makes the use of S_3 practically impossible. On the other hand, (59) does not involve any matrix inversions and can be computed easily regardless of the structure of L. This, therefore, opens the possibility of using S_3 in space-varying blur problems.

VII. Conclusion

We have presented a general projection method (EMOPSP) for solving convex set theoretic image feasibility problems. It proceeds by extrapolated relaxations of convex combinations of subgradient projections onto variable groups of sets. EMOPSP is superior to the widely used POCS algorithm on four counts: it converges very efficiently, it does not require the computation of exact projections, it can be implemented on concurrent processors in a very flexible fashion, and it can solve problems involving an infinite number of constraints. In view of its overwhelming computational advantages, EMOPSP can be anticipated to become a prominent tool in set theoretic image recovery.

Appendix A
Proofs

Proof of Theorem 1: Since $\mathbf{D}$ is a closed vector subspace of $\boldsymbol{\Xi}$, $P_{\mathbf{D}}$ is linear and (4) yields

$$(\forall \mathbf{a} \in \boldsymbol{\Xi})(\forall \mathbf{b} \in \mathbf{D}) \ \langle\langle \mathbf{b} \mid \mathbf{a}\rangle\rangle = \langle\langle \mathbf{b} \mid P_{\mathbf{D}}(\mathbf{a})\rangle\rangle. \tag{A1}$$

Now, fix $\mathbf{c} \in \mathbf{S} \bigcap \mathbf{D}$, $n \in \mathbb{N}$, and note that $(\mathbf{a}_n, \mathbf{c}) \in \mathbf{D}^{\mathbf{2}}$. Whence, (A1) and (4) yield

$$\begin{aligned}
&\langle\langle \mathbf{a}_n - \mathbf{c} \mid P_{\mathbf{D}} \circ P_{\mathbf{S}_n}(\mathbf{a}_n) - \mathbf{a}_n\rangle\rangle \\
&\quad = \langle\langle \mathbf{a}_n - \mathbf{c} \mid P_{\mathbf{D}}(P_{\mathbf{S}_n}(\mathbf{a}_n) - \mathbf{a}_n)\rangle\rangle \\
&\quad = \langle\langle \mathbf{a}_n - \mathbf{c} \mid P_{\mathbf{S}_n}(\mathbf{a}_n) - \mathbf{a}_n\rangle\rangle \\
&\quad = -|||P_{\mathbf{S}_n}(\mathbf{a}_n) - \mathbf{a}_n|||^2 \\
&\qquad + \langle\langle P_{\mathbf{S}_n}(\mathbf{a}_n) - \mathbf{c} \mid P_{\mathbf{S}_n}(\mathbf{a}_n) - \mathbf{a}_n\rangle\rangle \\
&\quad \leq -|||P_{\mathbf{S}_n}(\mathbf{a}_n) - \mathbf{a}_n|||^2.
\end{aligned} \tag{A2}$$

Using (30), (31), and (A2), we then obtain

$$\begin{aligned}
|||\mathbf{a}_{n+1} - \mathbf{c}|||^2 &= |||\mathbf{a}_n - \mathbf{c}|||^2 + 2\langle\langle \mathbf{a}_n - \mathbf{c} \mid \mathbf{a}_{n+1} - \mathbf{a}_n\rangle\rangle \\
&\quad + |||\mathbf{a}_{n+1} - \mathbf{a}_n|||^2 \\
&= |||\mathbf{a}_n - \mathbf{c}|||^2 \\
&\quad + 2\lambda_n \langle\langle \mathbf{a}_n - \mathbf{c} \mid P_{\mathbf{D}} \circ P_{\mathbf{S}_n}(\mathbf{a}_n) - \mathbf{a}_n\rangle\rangle \\
&\quad + \frac{\lambda_n^2}{L_n} |||P_{\mathbf{S}_n}(\mathbf{a}_n) - \mathbf{a}_n|||^2 \leq |||\mathbf{a}_n - \mathbf{c}|||^2 \\
&\quad - \lambda_n \left(2 - \frac{\lambda_n}{L_n}\right) |||P_{\mathbf{S}_n}(\mathbf{a}_n) - \mathbf{a}_n|||^2 && \text{(A3)} \\
&\leq |||\mathbf{a}_n - \mathbf{c}|||^2 - \varepsilon^2 \mathbf{d}(\mathbf{a}_n, \mathbf{S}_n)^2 && \text{(A4)} \\
&\leq |||\mathbf{a}_n - \mathbf{c}|||^2. && \text{(A5)}
\end{aligned}$$

It follows from (A4) that

$$\mathbf{d}(\mathbf{a}_n, \mathbf{S}_n)^2 \leq (|||\mathbf{a}_n - \mathbf{c}|||^2 - |||\mathbf{a}_{n+1} - \mathbf{c}|||^2)/\varepsilon^2. \tag{A6}$$

However, $(|||\mathbf{a}_n - \mathbf{c}|||^2)_{n\geq 0}$ converges by virtue of (A5) and therefore $(\mathbf{d}(\mathbf{a}_n, \mathbf{S}_n))_{n\geq 0}$ converges to 0. In addition, $(\mathbf{a}_n)_{n\geq 0} \subset \mathbf{D}$ is bounded and it admits a subsequence $(\mathbf{a}_{n_k})_{k\geq 0}$ converging weakly to some point $\mathbf{a} \in \mathbf{D}$. It then follows from (32) that $\mathbf{a} \in \mathbf{S} \bigcap \mathbf{D}$. Finally, since (A5) implies that $(\mathbf{a}_n)_{n\geq 0}$ can have at most one weak cluster point in $\mathbf{S} \bigcap \mathbf{D}$ [6], we conclude that $(\mathbf{a}_n)_{n\geq 0}$ converges weakly to $\mathbf{a}$. ◊

Proof of Proposition 5: At iteration n, if we rederive (A3) in the Hilbertian product space $\boldsymbol{\Xi}_n = \Xi^{\mathrm{card} I_n}$ with norm $|||\mathbf{a}|||_n = (\sum_{i \in I_n} w_{i,n} \|a_i\|^2)^{1/2}$ and bring it back to Ξ, we obtain directly that for every $c_n \in \bigcap_{i \in I_n} S_i$

$$\begin{aligned} &\|a_{n+1} - c_n\|^2 - \|a_n - c_n\|^2 \\ &\leq -\lambda_n \left(2 - \frac{\lambda_n}{L_n}\right) \sum_{i \in I_n} w_{i,n} \|P_{i,n}(a_n) - a_n\|^2 \qquad \text{(A7)} \\ &\leq -\varepsilon^2 \sum_{i \in I_n} w_{i,n} \|P_{i,n}(a_n) - a_n\|^2. \qquad \text{(A8)} \end{aligned}$$

The assertion is then proved by taking $c_n = c \in S$. $\diamond$

Proof of Theorem 3: Fix $c \in S$ and define $(\forall n \in \mathbb{N})\ \beta_n = (\|a_n - c\|^2 - \|a_{n+1} - c\|^2)^{1/2}$. Note that Proposition 5 entails that $(\beta_n)_{n \geq 0}$ converges to 0. Thanks to (52), we can find $\zeta \in \mathbb{R}_+^*$ such that $(\forall (i,n) \in I \times \mathbb{N})\ \|t_{i,n}\| \leq \zeta$. Therefore, for every integer n, (43), (A8), and (17) yield

$$\begin{aligned} \max_{i \in I_n} g_i(a_n) &\leq \zeta \max_{i \in I_n} \|P_{i,n}(a_n) - a_n\| \\ &\leq \zeta \left(\sum_{i \in I_n} w_{i,n} \|P_{i,n}(a_n) - a_n\|^2 / \delta\right)^{1/2} \\ &\leq \zeta \delta^{-1/2} \varepsilon^{-1} \beta_n \triangleq \gamma_n. \qquad \text{(A9)} \end{aligned}$$

Thanks to (A7), we also have

$$\begin{aligned} \|a_{n+1} - a_n\|^2 &= \lambda_n^2 \|\sum_{i \in I_n} w_{i,n} P_{i,n}(a_n) - a_n\|^2 \\ &= \frac{\lambda_n^2}{L_n} \sum_{i \in I_n} w_{i,n} \|P_{i,n}(a_n) - a_n\|^2 \\ &\leq \frac{\lambda_n^2}{L_n} \cdot \frac{\beta_n^2}{\lambda_n(2 - \lambda_n/L_n)} \\ &\leq \frac{\lambda_n}{L_n} \cdot \frac{\beta_n^2}{2 - \lambda_n/L_n} \\ &\leq (2\varepsilon^{-1} - 1)\beta_n^2. \qquad \text{(A10)} \end{aligned}$$

As in the proof of Theorem 1, Proposition 5 implies that $(a_n)_{n \geq 0}$ possesses a subsequence $(a_{n_k})_{k \geq 0}$ converging weakly to some point a and it remains to show $a \in S$. Fix $i \in I$. According to (50) there exists a sequence $(m_k)_{k \geq 0} \subset \mathbb{N}$ such that $(\forall k \in \mathbb{N})\ m_k \in \{n_k, \cdots, n_k + M_i - 1\}$ and $i \in I_{m_k}$. In addition, for every integer k, (A10) yields

$$\begin{aligned} \|a_{m_k} - a_{n_k}\| &\leq \sum_{l=n_k}^{m_k - 1} \|a_{l+1} - a_l\| \\ &\leq M_i (2\varepsilon^{-1} - 1)^{1/2} \max_{n_k \leq l \leq n_k + M_i - 1} \beta_l \\ &\triangleq \alpha_i \beta_{n_l}. \qquad \text{(A11)} \end{aligned}$$

However, since $(\beta_n)_{n \geq 0}$ converges to 0, $(a_{m_k} - a_{n_k})_{k \geq 0}$ converges strongly to 0 and, therefore, $(a_{m_k})_{k \geq 0}$ converges weakly to a. On the other hand, $(\gamma_{m_k})_{k \geq 0}$ converges to 0 in (A9) and it follows that $((a_{m_k}, \gamma_{m_k}))_{k \geq 0}$ converges weakly to $(a, 0)$ in the Hilbert space $\Xi \times \mathbb{R}$. However, thanks to (A9), $((a_{m_k}, \gamma_{m_k}))_{k \geq 0} \subset \operatorname{epi} g_i$ and, since g_i is convex and l.s.c., $\operatorname{epi} g_i$ is closed and convex and, thereby, weakly closed. Consequently, $(a, 0) \in \operatorname{epi} g_i$, i.e., $g_i(a) \leq 0$. We thus obtain $a \in S_i$ and, since i is arbitrary, $a \in S$. $\diamond$

Proof of Theorem 4: (i) Fix $i \in I$. Because of (51), there exists an increasing sequence $(n_k)_{k \geq 0} \subset \mathbb{N}$ such that $i \in \bigcap_{k \geq 0} I_{n_k}$. Since any sequence that satisfies (19) where $\overset{\circ}{S} \neq \varnothing$ converges strongly [5], Proposition 5 implies that $(a_n)_{n \geq 0}$ converges strongly to some point a. It follows from (A9) that $((a_{n_k}, \gamma_{n_k}))_{k \geq 0} \subset \operatorname{epi} g_i$ converges strongly to $(a, 0)$ in $\Xi \times \mathbb{R}$. Since $\operatorname{epi} g_i$ is closed, we get $(a, 0) \in \operatorname{epi} g_i$ and, therefore, $a \in S_i$. As i is arbitrary, we conclude $a \in S$. (ii) Fix $c \in S$ and let $\eta = \zeta \delta^{-1/2} \varepsilon^{-1} \|a_0 - c\|$. As in (i), there exists a suborbit $(a_{n_k})_{k \geq 0}$ such that $j \in \bigcap_{k \geq 0} I_{n_k}$. In view of Proposition 5 and (A9), $(a_{n_k})_{k \geq 0}$ lies in $\sec(g_j, \eta) \bigcap B(c, \|a_0 - c\|)$, which is compact since g_j is l.s.b.co. We can therefore extract a subsequence $(a_{n_{k_l}})_{l \geq 0}$ converging strongly to some point a. It remains to show $a \in S$ for Proposition 5 will then automatically guarantee that the whole sequence $(a_n)_{n \geq 0}$ converges strongly to a. Suppose to the contrary that $a \notin S$ and define $I^+ = \{i \in I \mid a \in S_i\}$, $I^- = I \setminus I^+$, and $\mu = \min_{i \in I^-} g_i(a) > 0$. Take ζ as in the proof of Theorem 3. We derive from (A8) the inequalities

$$\begin{aligned} &(\forall n \in \mathbb{N})(\forall c_n \in \bigcap_{i \in I_n} S_i)\ \|a_{n+1} - c_n\|^2 \\ &\leq \|a_n - c_n\|^2 - \nu \max_{i \in I_n} g_i(a_n)^2. \qquad \text{(A12)} \end{aligned}$$

Now, fix $l \in I^-$. Note that a belongs to $\complement \sec(g_l, \mu/2)$, which is open since g_l is l.s.c. Hence, we can find $\gamma \in \mathbb{R}_+^*$ such that

$$(\forall b \in B(a, \gamma))\ g_l(b) \geq \mu/2. \qquad \text{(A13)}$$

Let us fix an integer p such that $a_p \in B(a, \gamma)$. Then $a_p \notin S_l$. Let us show that $l \notin I_p$. Indeed, if we had $l \in I_p$, it would follow from (A12) and (A13) that, for γ sufficiently small

$$\begin{aligned} \|a_{p+1} - c\|^2 &\leq \|a_p - c\|^2 - \nu g_l(a_p)^2 \\ &\leq (\gamma + \|a - c\|)^2 - \nu \mu^2/4 \\ &< \|a - c\|^2. \qquad \text{(A14)} \end{aligned}$$

However, this would contradict Proposition 5, which implies that $\|a - c\| \leq \|a_{p+1} - c\|$. Hence, $l \notin I_p$. Since l is arbitrary, $l \notin I_p \Rightarrow I^- \bigcap I_p = \varnothing \Rightarrow I_p \subset I^+ \Rightarrow a \in \bigcap_{i \in I_p} S_i$. (A12) then yields $\|a_{p+1} - a\| \leq \|a_p - a\| \leq \gamma$, i.e., $a_{p+1} \in B(a, \gamma)$. Thus, the above arguments can be replicated for index $p+1$ to give $l \notin I_{p+1}$, $a_{p+2} \in B(a, \gamma)$ and, by induction, $l \notin \bigcup_{k \geq 0} I_{p+k}$. But this is absurd since the control is chaotic. Accordingly, we conclude $a \in S$. $\diamond$

Proof of Corollary 1: If $\dim \Xi < +\infty$, the subdifferential of each g_i is bounded on closed and bounded sets [36]. Since $\operatorname{card} I < +\infty$, it follows that the family $(g_i)_{i \in I}$ satisfies (52). Finally, each g_i is l.s.b.co. since, in finite dimensional spaces, every closed set is boundedly compact. $\diamond$

Appendix B
Acronyms

ANCA	Accelerated nonlinear Cimmino algorithm (15) + (18).
ART	Algebraic reconstruction technique (3).
EMOPSP	Extrapolated method of parallel subgradient projections (17) + (43) + (44) + (45) + (46).

EPPM	Extrapolated parallel projection method (10) + (14) + (25) + (26).
EPPM2	(Generalized) extrapolated parallel projection method (10) + (33) + (34) + (35).
MOPP	Method of parallel projections (12) + (15) + (16) + (17).
POCS	Projection onto convex sets (11) + (12).
PPM	Parallel projection method (10) + (12) + (14).
SIRT	Simultaneous iterative reconstruction technique (13).

REFERENCES

[1] R. Aharoni, A. Berman, and Y. Censor, "An interior points algorithm for the convex feasibility problem," *Adv. in Appl. Math.,* vol. 4, pp. 479–489, Dec. 1983.

[2] A. H. Andersen, "Algebraic reconstruction in CT from limited views," *IEEE Trans. Med. Imaging,* vol. 8, pp. 50–55, Mar. 1989.

[3] H. C. Andrews and B. R. Hunt, *Digital Image Restoration.* Englewood Cliffs, NJ: Prentice-Hall, 1977.

[4] J.-P. Aubin, *L'Analyse Non Linéaire et Ses Motivations Économiques.* Paris: Masson, 1984; *Optima and Equilibria - An Introduction to Nonlinear Analysis.* New York: Springer-Verlag, 1993.

[5] H. H. Bauschke and J. M. Borwein, "On projection algorithms for solving convex feasibility problems," *SIAM Rev.*, vol. 38, pp. 367–426, Sept. 1996.

[6] F. E. Browder, "Convergence theorems for sequences of nonlinear operators in Banach spaces," *Math. Z.,* vol. 100, pp. 201–225, July 1967.

[7] D. Butnariu and Y. Censor, "On the behavior of a block-iterative projection method for solving convex feasibility problems," *Int. J. Comput. Math.,* vol. 34, pp. 79–94, 1990.

[8] J. Céa, *Optimisation - Théorie et Algorithmes.* Paris: Dunod, 1971.

[9] Y. Censor, "Parallel application of block-iterative methods in medical imaging and radiation therapy," *Math. Programming,* vol. 42, pp. 307–325, 1988.

[10] Y. Censor and A. Lent, "Cyclic subgradient projections," *Math. Programming,* vol. 24, pp. 233–235, 1982.

[11] A. E. Çetin, "An iterative algorithm for signal reconstruction from bispectrum," *IEEE Trans. Signal Processing,* vol. 39, pp. 2621–2628, Dec. 1991.

[12] A. E. Çetin and R. Ansari, "Signal recovery from wavelet transform maxima," *IEEE Trans. Signal Processing,* vol. 42, pp. 194–196, Jan. 1994.

[13] P. L. Combettes, "The foundations of set theoretic estimation," *Proc. IEEE,* vol. 81, pp. 182–208, Feb. 1993.

[14] ——, "Inconsistent signal feasibility problems: Least-squares solutions in a product space," *IEEE Trans. Signal Processing,* vol. 42, pp. 2955–2966, Nov. 1994.

[15] ——, "Construction d'un point fixe commun à une famille de contractions fermes," *C. R. Acad. Sci. Paris Sér. I Math.,* vol. 320, pp. 1385–1390, June 1995.

[16] P. L. Combettes and H. Puh, "Iterations of parallel convex projections in Hilbert spaces," *Numer. Funct. Anal. Optim.,* vol. 15, pp. 225–243, 1994.

[17] P. L. Combettes and H. J. Trussell, "Methods for digital restoration of signals degraded by a stochastic impulse response," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. 37, pp. 393–401, Mar. 1989.

[18] ——, "The use of noise properties in set theoretic estimation," *IEEE Trans. Signal Processing,* vol. 39, pp. 1630–1641, July 1991.

[19] L. T. Dos Santos, "A parallel subgradient projections method for the convex feasibility problem," *J. Comput. Appl. Math.,* vol. 18, pp. 307–320, June 1987.

[20] I. Ekeland and R. Temam, *Analyse Convexe et Problèmes Variationnels.* Paris: Dunod, 1974; *Convex Analysis and Variational Problems.* Amsterdam, The Netherlands: North-Holland, 1976.

[21] S. D. Flåm and J. Zowe, "Relaxed outer projections, weighted averages, and convex feasibility," *BIT,* vol. 30, pp. 289–300, 1990.

[22] P. Gilbert, "Iterative methods for the three-dimensional reconstruction of an object from projections," *J. Theoret. Biol.,* vol. 36, pp. 105–117, July 1972.

[23] R. Gordon, R. Bender, and G. T. Herman, "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography," *J. Theoret. Biol.,* vol. 29, pp. 471–481, Dec. 1970.

[24] L. G. Gurin (Gubin), B. T. Polyak, and E. V. Raik, "The method of projections for finding the common point of convex sets," *USSR Comput. Math. Math. Phys.,* vol. 7, pp. 1-24, 1967.

[25] G. T. Herman, *Image Reconstruction from Projections, the Fundamentals of Computerized Tomography.* New York: Academic, 1980.

[26] A. N. Iusem and A. R. De Pierro, "Convergence results for an accelerated nonlinear Cimmino algorithm," *Numer. Math.,* vol. 49, pp. 367–378, Aug. 1986.

[27] A. K. Katsaggelos, Ed., *Digital Image Restoration.* New York: Springer-Verlag, 1991.

[28] K. C. Kiwiel, "Block-iterative surrogate projection methods for convex feasibility problems," *Linear Algebra Appl.,* vol. 215, pp. 225–259, Jan. 1995.

[29] A. Lent and H. Tuy, "An iterative method for the extrapolation of band-limited functions," *J. Math. Anal. Appl.,* vol. 83, pp. 554–565, Oct. 1981.

[30] J. Mandel, "Convergence of the cyclical relaxation method for linear inequalities," *Math. Programming,* vol. 30, pp. 218–228, 1984.

[31] J.-J. Moreau, "Fonctionnelles sous-différentiables," *C. R. Acad. Sci. Paris Sér. A,* vol. 257, pp. 4117–4119, Dec. 1963.

[32] S. Oh, C. Ramon, R. J. Marks II, A. C. Nelson, and M. G. Meyer, "Resolution enhancement of biomagnetic images using the method of alternating projections," *IEEE Trans. Biomed. Eng.,* vol. 40, pp. 323–328, Apr. 1993.

[33] N. Ottavy, "Strong convergence of projection-like methods in Hilbert spaces," *J. Optim. Theory Appl.,* vol. 56, pp. 433–461, Mar. 1988.

[34] M. K. Özkan, A. M. Tekalp, and M. I. Sezan, "POCS-based restoration of space-varying blurred images," *IEEE Trans. Image Processing,* vol. 3, pp. 450–454, July 1994.

[35] G. Pierra, "Méthodes de projections parallèles extrapolées relatives à une intersection de convexes," Res. Rep. INPG, Grenoble, France, Sept. 1975; and "Decomposition through formalization in a product space," *Math. Programming,* vol. 28, pp. 96–115, Jan. 1984.

[36] R. T. Rockafellar, *Convex Analysis.* Princeton, NJ: Princeton Univ. Press, 1970.

[37] C. Sánchez-Avila, "An adaptive regularized method for deconvolution of signals with edges by convex projections," *IEEE Trans. Signal Processing,* vol. 42, pp. 1849–1851, July 1994.

[38] M. I. Sezan and H. Stark, "Image restoration by the method of convex projections: Part 2, applications and numerical results," *IEEE Trans. Med. Imag.,* vol. 1, pp. 95–101, Oct. 1982.

[39] M. I. Sezan and H. J. Trussell, "Prototype image constraints for set-theoretic image restoration," *IEEE Trans. Signal Processing,* vol. 39, pp. 2275–2285, Oct. 1991.

[40] P. Y. Simard and G. E. Mailloux, "A projection operator for the restoration of divergence-free vector fields," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 10, pp. 248–256, Mar. 1988.

[41] H. Stark, Ed., *Image Recovery: Theory and Application.* San Diego, CA: Academic, 1987.

[42] H. J. Trussell and M. R. Civanlar, "The feasible solution in signal restoration," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. 32, pp. 201–212, Apr. 1984.

[43] ——, "The Landweber iteration and projection onto convex sets," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. 33, pp. 1632–1634, Dec. 1985.

[44] D. C. Youla and H. Webb, "Image restoration by the method of convex projections: Part 1, theory," *IEEE Trans. Med. Imaging,* vol. 1, pp. 81–94, Oct. 1982.

Patrick L. Combettes (S'84-M'90-SM'96) was born in Salon de Provence, France, on May 24, 1962. He received the Diplôme d'Ingénieur from l'Institut National des Sciences Appliquées de Lyon, France, the Ph.D. degree from North Carolina State University, Raleigh, and the Habilitation à Diriger les Recherches from the Université de Paris XI, Paris-Sud, Orsay, France, in 1985, 1989, and 1996, respectively.

In August 1990, he joined the Department of Electrical Engineering at the City College and the Graduate School of the City University of New York, NY, where he is now an Associate Professor. He has held several visiting scientist positions at the CNRS/Université de Paris XI, Laboratoire des Signaux et Systèmes, Gif sur Yvette, France. His current research interests are in signal theory.

Dr. Combettes is a member of the Société Mathématique de France and the American Mathematical Society. He received the IEEE Signal Processing Society's 1993 Paper Award.