



HAL
open science

Projection Methods for Uniformly Convex Expandable Sets

Stéphane Chrétien, Pascal Bondon

► **To cite this version:**

Stéphane Chrétien, Pascal Bondon. Projection Methods for Uniformly Convex Expandable Sets. Mathematics , 2020, 10.3390/math8071108 . hal-02896463

HAL Id: hal-02896463

<https://centralesupelec.hal.science/hal-02896463>

Submitted on 10 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Article

Projection Methods for Uniformly Convex Expandable Sets

Stéphane Chrétien ^{1,2,3,*}  and Pascal Bondon ⁴ ¹ Laboratoire ERIC, Université Lyon 2, 69500 Bron, France² The Alan Turing Institute, London NW1 2DB, UK³ Data Science Division, The National Physical Laboratory, Teddington TW11 0LW, UK⁴ Laboratoire des Signaux et Systèmes, CentraleSupélec, CNRS, Université Paris-Saclay, 91190 Gif-sur-Yvette, France; pascal.bondon@l2s.centralesupelec.fr

* Correspondence: stephane.chretien@univ-lyon2.fr

Received: 21 February 2020; Accepted: 22 June 2020; Published: 6 July 2020



Abstract: Many problems in medical image reconstruction and machine learning can be formulated as nonconvex set theoretic feasibility problems. Among efficient methods that can be put to work in practice, successive projection algorithms have received a lot of attention in the case of convex constraint sets. In the present work, we provide a theoretical study of a general projection method in the case where the constraint sets are nonconvex and satisfy some other structural properties. We apply our algorithm to image recovery in magnetic resonance imaging (MRI) and to a signal denoising in the spirit of Cadzow's method.

Keywords: cyclic projections; nonconvex sets; uniformly convex sets; strong convergence

1. Introduction

1.1. Background and Goal of the Paper

Many problems in applied mathematics, engineering, statistics and machine learning can be reduced to finding a point in the intersection of some subsets of a real separable Hilbert space \mathcal{H} . In mathematical terms, let $(S_i)_{i \in I}$ be a finite family of proximal subsets (i.e., sets S such that any $x \in \mathcal{H}$ admits a closest point in S) [1] of \mathcal{H} with a non-empty intersection, we address the problem of finding a point in the intersection of the sets $(S_i)_{i \in I}$ using a successive projection method.

When the sets are closed and convex, the problem is known as the convex feasibility problem, and is the subject of extensive literature; see [2–5] and references therein. In this paper, we take one step further into the scarcely investigated topic of finding a common point in the intersection of nonconvex sets. Convex feasibility problems have been applied to an extremely wide range of topics in systems analysis and control, signal and image processing. Important examples are: model order reduction [6], controller design [7], tensor analysis [8], image recovery [9], electrical capacitance tomography [10], MRI [11,12], and stabilisation of quantum systems and application to quantum computation [13,14].

Extension of this problem to the nonconvex setting also has many applications, related to sparse estimation and more specifically low rank constraints, such as in control theory [15], signal denoising [16], phase retrieval [17], structured matrix estimation [18,19] and has great potential impact on the design of scalable algorithm in many machine learning problems such as Deep Neural Networks [20]. Studies of projection methods in the nonconvex settings have been quite scarce in the literature [21–23] and a lot of work still remains to be done in order to understand the behavior of such methods for these difficult problems.

In the present paper, our goal is to investigate how the results in [22] can be improved in such a way that strong convergence can be obtained. The results proved in [22] make the assumption that the sets involved in the feasibility problem can be written as a possibly not countable union of closed convex sets, and one of the sets is boundedly compact. Our study will be based on the assumption that the convex sets in the family are uniformly convex, which allows to remove the boundedly compactness assumption. As will be seen in Section 4, uniform convexity can be obtained by simply modifying the algorithm even in the case where the sets S_i are not expandable into a family of uniformly convex sets.

1.2. Preliminary on Projections and Expansion into Convex Sets

The notation Card will denote the cardinality of a set. \mathcal{H} will denote a real separable Hilbert space with scalar product $\langle \cdot | \cdot \rangle$, norm $|\cdot|$, and distance d . Let S be a subset of \mathcal{H} . \bar{S} denotes the closure of S . S is proximal if each point in \mathcal{H} has at least one projection onto S . If S is proximal, then S is closed. When S is proximal, P_S is the projection point-to-set mapping onto S defined by

$$(\forall x \in \mathcal{H}) \quad P_S(x) = \{y \in S \mid |x - y| = \inf_{z \in S} |x - z|\}.$$

$p_S(x)$ denotes an arbitrary element of $P_S(x)$ and $d(x, S) = |x - p_S(x)|$. $B(x, \rho)$ denotes the closed ball of center x and radius ρ . In the case of a projection onto a convex set C , Kolmogorov's criterion is:

$$\langle x' - P_C(x), x - P_C(x) \rangle \leq 0 \tag{1}$$

for all $x' \in C$. Our work is a follow-on project to the work in [22] which introduced the notion of convex expandable sets in order to tackle certain feasibility problems involving rank-constrained matrices as described in [23]. We recall what it means for a nonconvex set to be expandable into a family of convex sets.

Definition 1. A subset S of \mathcal{H} is said to be expandable into a family of convex sets when it is the union of non-trivial, i.e., not reduced to a single point, convex subsets C_j of \mathcal{H} , i.e.,

$$S = \bigcup_{j \in J} C_j \tag{2}$$

where J is a possibly uncountable index set. Any family $(C_j)_{j \in J}$ satisfying (2) is called a convex expansion of S .

Remark 1. Uncountable unions often appear in practice. An important example is the set of matrices of rank r in $\mathbb{R}^{n \times m}$ for $0 < r < \min\{n, m\}$. This set is the union of the rays passing through the null matrix and any matrix of rank r different from zero. This constraint often appears in signal processing problems such as signal denoising [16] and more generally, structured matrix estimation problems [19].

Uniform convexity is an important concept, which allows to prove many strong convergence results in the framework of successive projection methods ([4], Section 5.3).

Definition 2. A convex subset C of \mathcal{H} is uniformly convex with respect to a positive-valued nondecreasing function $b(t)$ if

$$\forall (x, y) \in C^2, \quad B((x + y)/2, b(|x - y|)) \subset C.$$

Uniform convexity will be instrumental in our analysis. We define the following condition which is stronger than the condition of being expandable into a family of convex sets.

Definition 3. A subset S of \mathcal{H} is said to be expandable in uniformly convex sets when the sets C_j in (2) are moreover uniformly convex with the same function $b(\cdot)$. Any family $(C_j)_{j \in J}$ of uniformly convex sets satisfying (2) is called a uniformly convex expansion of S .

1.3. The Projection Algorithm

For any point x in \mathcal{H} , for the sake of simplicity, $p_{S_i}(x)$ will be noted by $p_i(x)$. We will assume throughout that the family $\{S_i\}_{i \in I}$ is finite.

Definition 4. (Method of Projection onto Uniformly Convex Expandable Sets)

Given a point x_0 in \mathcal{H} , two real numbers α and λ in $(0, 1]$, a nonnegative integer M and a sequence (I_n) of subsets of I satisfying the following conditions

$$\begin{aligned}
 C_1. \quad & (\forall n \in \mathbb{N}) \quad I_n \subset I, \quad I = \bigcup_{j=n}^{n+M} I_j \text{ and } \text{Card}(I_n) > 1 \\
 C_2. \quad & (\forall n \in \mathbb{N}) \quad (\forall i \in I_n) \quad \alpha \leq \alpha_{i,n} \text{ and } \sum_{j \in I_n} \alpha_{j,n} = 1
 \end{aligned}$$

the projection-like method considered in this paper is iteratively defined by

$$(\forall n \in \mathbb{N}) \quad \left\{ \begin{array}{l} 1. \quad \left\{ \begin{array}{l} - \text{If } \exists j \in I_n \text{ such that } p_j(x_n) \in \bigcap_{i \in I_n} S_i \\ \text{Then} \quad \left\{ \begin{array}{l} - \text{If } I_n \neq I \text{ set } I_n = I \text{ and go to 1} \\ - \text{If } I_n = I \text{ set } \lambda_n = 1, \alpha_{j,n} = 1 \text{ and go to 2} \end{array} \right. \\ - \text{Else go to 2} \end{array} \right. \\ 2. \quad x_{n+1} = x_n + \lambda_n \sum_{i \in I_n} \alpha_{i,n} (p_i(x_n) - x_n) \end{array} \right. \quad (3)$$

with λ_n satisfying

$$C_3. \quad (\forall n \in \mathbb{N}) \quad \lambda \leq \lambda_n \leq \mu_n \text{ where } \mu_n = \begin{cases} \frac{\sum_{i \in I_n} \alpha_{i,n} |x_n - p_i(x_n)|^2}{|x_n - \sum_{i \in I_n} \alpha_{i,n} p_i(x_n)|^2} & \text{if } x_n \notin \bigcap_{i \in I_n} S_i \\ 1 & \text{if } x_n \in \bigcap_{i \in I_n} S_i \end{cases}$$

Remark 2. Using the assumption $\text{Card}(I_n) > 1$, the coefficients $\alpha_{i,n}$, $i \in I_n$, can be chosen in such a way that the denominator in the definition of μ_n is not equal to zero. Note that one can relax the constraint $\text{Card}(I_n) > 1$ in the case where μ_n is well defined at every iteration, thus allowing to recover the cyclic projection method.

The numbers α and λ ensure that the steps taken at each iteration are sufficiently large as compared to the distance of the iterations to their projections on each S_i , $i \in I$. The use of the integer M ensures that each S_i is involved at least every M iterations. The ‘‘If’’ condition in Step 1 of the algorithm ensures that x_n will move at each iteration by enforcing that it does not yet belong to all the selected sets S_i , $i \in I_n$. Our method is an extension of Ottavy’s method [4] to the case of nonconvex sets. The idea of using a potentially variable index set I_n at each iteration allows to recover several variants, such as cyclic projections [2,21,22,24] or parallel projections [2,23,24]. Notice that, due to finiteness of the cardinality of I , finding an index j such that $p_j(x_n) \in \bigcap_{i \in I_n} S_i$ is easy if such an index exists.

1.4. Our Contributions

The main contributions of our work are a proof that strong convergence holds in the proposed setting, and a new constructive modification of the projection method in order to accommodate the case of non-necessarily uniformly convex expandable sets. Based on our findings, we will then weaken our assumptions and provide a new algorithm which does not need uniform convexity, while preserving strong convergence to a feasible solution. Applications to an infinite dimensional MRI problem and to low rank matrix estimation for signal denoising are presented in Section 5.

2. Ottavy's Framework

2.1. Successive Projection Point-to-Set Mapping

Let α and λ be two real numbers chosen in $(0, 1)$. Define the point-to-set mapping

$$D : x \in H \rightarrow D(x) = \{\bar{x} \in \mathcal{H}; \bar{x} = x + \lambda(x)(\bar{x} - x)\}$$

where \bar{x} is a weighted average,

$$\bar{x} = \sum_{i \in I(x)} \alpha_i(x) p_i(x)$$

$I(x)$ is a subset of selected constraints such that

$$I(x) \subset I \quad \text{and} \quad I(x) \neq \emptyset$$

$\alpha_i(x), i \in I$ is a set of normalized nonvanishing weightings such that

$$\alpha_i(x) \in [0, 1], \sum_{i \in I(x)} \alpha_i(x) = 1, \quad \text{and} \quad p_i(x) \neq x \Rightarrow \alpha_i(x) \geq \alpha$$

$\lambda(x)$ is the relaxation parameter such that

$$\lambda(x) \in [\lambda, +\infty) \quad \text{and} \quad x \neq \bar{x} \Rightarrow \lambda(x) \in [\lambda, \mu(\bar{x})]$$

with

$$\mu(\bar{x}) = \left[\sum_{\alpha \in I(x)} \alpha_i(x) |x - p_i(x)|^2 \right] / |x - \bar{x}|^2.$$

2.2. Ottavy's Lemma

In [4], Ottavy proved the very useful lemma.

Lemma 1. For any $x \in \mathcal{H}, \bar{x} \in D(x)$, and $z \in \cap_{i \in I} S_i$, the following results hold:

- (i) $|\bar{x} - x|^2 \leq \beta (|x - z|^2 - |\bar{x} - z|^2)$
- (ii) $2\langle x - z | x - \bar{x} \rangle \leq \left(1 + \frac{2}{\lambda}\right) (|x - z|^2 - |\bar{x} - z|^2)$
- (iii) $\max_{i \in I(x)} |p_i(x) - x|^2 \leq \frac{(2 + \lambda)}{2\alpha\lambda^2} (|x - z|^2 - |\bar{x} - z|^2).$

3. A Strong Convergence Result

Let $(x_n)_{n \in \mathbb{N}}$ be the sequence of iterates of the projection method defined by (3). In the present section, we show strong convergence of this sequence to a point in $\cap_{i \in I} S_i$.

For every n in \mathbb{N} and for each $i \in I$, define

$$T_{i,n} = \cup\{C \subset S_i \mid C \text{ is convex and } p_i(x_n) \in C\} \qquad T_n = \cap_{i \in I_n} T_{i,n}.$$

When the (finite) family $(S_i)_{i \in I}$ is expandable into a family of convex sets, $T_{i,n} \neq \emptyset$ for every n in \mathbb{N} and for each $i \in I$.

Lemma 2. If the following conditions are satisfied

- C₄. $(\forall n \in \mathbb{N}) T_n \neq \emptyset$
- C₅. The finite family $(S_i)_{i \in I}$ is expandable in convex sets

then, for every n in \mathbb{N} and for every z_n in T_n , we have

- (i) $|x_{n+1} - x_n|^2 \leq \frac{2}{\lambda} (|x_n - z_n|^2 - |x_{n+1} - z_n|^2)$
- (ii) $2 \langle x_n - z_n | x_n - x_{n+1} \rangle \leq (1 + \frac{2}{\lambda}) (|x_n - z_n|^2 - |x_{n+1} - z_n|^2)$
- (iii) $\sup_{i \in I_n} |p_i(x_n) - x_n|^2 \leq \frac{(2 + \lambda)}{2\alpha\lambda^2} (|x_n - z_n|^2 - |x_{n+1} - z_n|^2)$.

Proof. Fix n in \mathbb{N} and z_n in T_n . According to C₅, each point $p_i(x_n)$ belongs to a set $C_{i,j_n} \subset S_i$, a convex set in the uniformly convex expansion of S_i indexed by j_n . Hence $p_i(x_n)$ is also the projection of x_n onto C_{i,j_n} , and $z_n \in C_{i,j_n}$. Replacing respectively x, \tilde{x} and z by x_n, x_{n+1} and z_n in Ottavý's Lemma (Lemma 1, or [4, Lemma 7.1]), we obtain the desired results. \square

Lemma 3. If the conditions C₄, C₅, and

- C₆. $(\exists (z_n)_{n \in \mathbb{N}})$ such that $z_n \in T_n$ for all n in \mathbb{N} and $\sum_{n \in \mathbb{N}} |z_{n+1} - z_n| < +\infty$

are satisfied, then

- (i) The sequences $(|x_n - z_n|)_{n \in \mathbb{N}}$ and $(|x_{n+1} - z_n|)_{n \in \mathbb{N}}$ are bounded.
- (ii) The series $\sum_{n \in \mathbb{N}} (|x_n - z_n|^2 - |x_{n+1} - z_n|^2)$ is convergent.
- (iii) For each i in I , the sequence $(|p_i(x_n) - x_n|)_{n \in \mathbb{N}}$ converges to zero.
- (iv) The sequence $(|x_n - z_n|)_{n \in \mathbb{N}}$ is convergent.
- (v) The sequence $(z_n)_{n \in \mathbb{N}}$ converges strongly to a point in $\bigcap_{i \in I} S_i$.

Proof. (i) Let $(z_n)_{n \in \mathbb{N}}$ be a sequence satisfying C₆, and let k in \mathbb{N} . We deduce from Lemma 2(i) that $|x_{k+1} - z_k| \leq |x_k - z_k|$. Then we have

$$|x_{k+1} - z_{k+1}| \leq |x_{k+1} - z_k| + |z_{k+1} - z_k| \leq |x_k - z_k| + |z_{k+1} - z_k|.$$

Therefore

$$|x_{n+1} - z_{n+1}| \leq |x_0 - z_0| + \sum_{k=0}^n |z_{k+1} - z_k|$$

and then C₆ ensures that the sequences $(|x_n - z_n|)_{n \in \mathbb{N}}$ and $(|x_{n+1} - z_n|)_{n \in \mathbb{N}}$ are bounded.

(ii) Now, the cosine law, followed by the Cauchy-Schwarz inequality give

$$\begin{aligned} |x_{k+1} - z_{k+1}|^2 &= |x_{k+1} - z_k|^2 + 2 \langle x_{k+1} - z_k | z_k - z_{k+1} \rangle + |z_k - z_{k+1}|^2 \\ &\leq |x_{k+1} - z_k|^2 + 2 |x_{k+1} - z_k| |z_{k+1} - z_k| + |z_{k+1} - z_k|^2. \end{aligned} \tag{4}$$

Using Lemma 2(i), we then obtain that

$$\begin{aligned} 0 \leq |x_k - z_k|^2 - |x_{k+1} - z_k|^2 &\leq |x_k - z_k|^2 - |x_{k+1} - z_{k+1}|^2 + 2 |x_{k+1} - z_k| |z_{k+1} - z_k| \\ &\quad + |z_{k+1} - z_k|^2. \end{aligned}$$

Thus

$$0 \leq \sum_{k=0}^n (|x_k - z_k|^2 - |x_{k+1} - z_k|^2) \leq |x_0 - z_0|^2 + 2 \sum_{k=0}^n |x_{k+1} - z_k| |z_{k+1} - z_k| + \sum_{k=0}^n |z_{k+1} - z_k|^2.$$

Since the sequence $(|x_{n+1} - z_n|)_{n \in \mathbb{N}}$ is bounded and the series $\sum_{n \in \mathbb{N}} |z_{n+1} - z_n|$ is convergent, the result (i) follows at once.

(iii) Lemma 3(i) and Lemma 2(iii), imply that $\lim_{n \rightarrow \infty} \sup_{i \in I_n} |p_i(x_n) - x_n|^2 = 0$. Thus for any $k \in [0, M]$,

$$\lim_{n \rightarrow \infty} \sup_{i \in I_{n+k}} |p_i(x_n) - x_n| = 0.$$

Then we deduce from C_1 that $\lim_{n \rightarrow \infty} \sup_{i \in I} |p_i(x_n) - x_n| = 0$, which completes the proof of (iii).

(iv) Squaring the triangle inequality gives

$$\begin{aligned} |x_{k+1} - z_k|^2 - |x_{k+1} - z_{k+1}|^2 &\leq 2|x_{k+1} - z_{k+1}| |z_{k+1} - z_k| + |z_{k+1} - z_k|^2 \\ &\leq 2|x_{k+1} - z_k| |z_{k+1} - z_k| + 3|z_{k+1} - z_k|^2. \end{aligned}$$

Using (4), we deduce that

$$\left| |x_{k+1} - z_{k+1}|^2 - |x_{k+1} - z_k|^2 \right| \leq 2|x_{k+1} - z_k| |z_{k+1} - z_k| + 3|z_{k+1} - z_k|^2.$$

Since the sequence $(|x_{n+1} - z_n|)_{n \in \mathbb{N}}$ is bounded, and $\sum_{n \in \mathbb{N}} |z_{n+1} - z_n|$ is convergent, we deduce from the last inequality that $\sum_{n \in \mathbb{N}} (|x_{n+1} - z_{n+1}|^2 - |x_{n+1} - z_n|^2)$ is convergent. Then using Lemma 3(i), we obtain that the series $\sum_{n \in \mathbb{N}} (|x_{n+1} - z_{n+1}|^2 - |x_n - z_n|^2)$ is convergent, which is equivalent to the convergence of the sequence $(|x_n - z_n|)_{n \in \mathbb{N}}$.

(v) Since, by C_6 , $(z_n)_{n \in \mathbb{N}}$ is a Cauchy sequence, it is strongly convergent to a point z in the Hilbert space \mathcal{H} . For every n in \mathbb{N} , $z_n \in T_n = \bigcap_{i \in I_n} T_{i,n} \subset \bigcap_{i \in I_n} S_i$. Fix i in I . Due to condition C_1 , there exists a subsequence $(z_{\sigma(n)})_{n \in \mathbb{N}}$ of $(z_n)_{n \in \mathbb{N}}$ satisfying $z_{\sigma(n)} \in S_i$ for all $n \in \mathbb{N}$. Therefore, $z \in \overline{S_i}$. Since this assertion is true for all i in I , and each S_i is closed, we obtain that $z \in \bigcap_{i \in I} S_i$. \square

Remark 3. In the convex case, Lemma 3(iii) is known in the simpler case where $(z_n)_{n \in \mathbb{N}}$ is chosen to be a constant sequence in $\bigcap_{i \in I} S_i$ and is a consequence of the Fejér monotonicity [5].

We now state the main result of this section, which is strong convergence under the assumption that the sets $S_i, i \in I$, are expandable in uniformly convex sets.

Theorem 1. If the conditions C_4, C_6 , and C_7 .

The finite family $(S_i)_{i \in I}$ is expandable in uniformly convex sets with respect to a function b are satisfied, then the sequence $(x_n)_{n \in \mathbb{N}}$ converges strongly to a point in $\bigcap_{i \in I} S_i$.

Proof. Notice that C_7 implies C_5 . According to C_5 , each point $p_i(x_n)$ belongs to a convex set $C_{i,j_n} \subset S_i$ in the uniformly convex expansion of S_i , indexed by j_n . Hence $p_i(x_n)$ is also the projection of x_n onto C_{i,j_n} , and $z_n \in C_{i,j_n}$. Let $(z_n)_{n \in \mathbb{N}}$ be a sequence satisfying C_6 . Fix n in \mathbb{N} , and i in I . Define

$$H(x_n, p_i(x_n)) = \{x \in \mathcal{H} \mid \langle x - p_i(x_n) \mid p_i(x_n) - x_n \rangle \geq 0\}.$$

According to C_7 , $p_i(x_n)$ and z_n belong to a uniformly convex set $C_{i,j_n} \subset S_i$. Thus

$$B((z_n + p_i(x_n))/2, \rho_{i,n}) \subset C_{i,j_n},$$

where $\rho_{i,n} = b(|z_n - p_i(x_n)|)$. Since $p_i(x_n)$ is also the projection of x_n onto C_{i,j_n} , it results from the Kolmogorov criterion (1) that

$$C_{i,j_n} \subset H(x_n, p_i(x_n)).$$

Then

$$B((z_n + p_i(x_n))/2, \rho_{i,n}) \subset C_{i,j_n} \subset H(x_n, p_i(x_n)). \tag{5}$$

Now, consider the translation $T : x \mapsto y = x - (p_i(x_n) - z_n)/2$. Clearly,

$$x \in B((z_n + p_i(x_n))/2, \rho_{i,n}) \text{ if and only if } y \in B(z_n, \rho_{i,n}). \tag{6}$$

On the other hand, using $z_n \in C_{i,j_n}$, we get that $x \in H(x_n, p_i(x_n))$ implies $y \in H(x_n, p_i(x_n))$. Indeed,

$$\begin{aligned} \langle y - p_i(x_n) \mid p_i(x_n) - x_n \rangle &= \langle x - (p_i(x_n) - z_n)/2 - p_i(x_n) \mid p_i(x_n) - x_n \rangle \\ &= \langle x - p_i(x_n) \mid p_i(x_n) - x_n \rangle - \frac{1}{2} \langle p_i(x_n) - z_n \mid p_i(x_n) - x_n \rangle \end{aligned}$$

and since $z_n \in C_{i,j_n}$, Kolmogorov’s criterion gives

$$\frac{1}{2} \langle p_i(x_n) - z_n \mid p_i(x_n) - x_n \rangle \leq 0,$$

which implies

$$\begin{aligned} \langle y - p_i(x_n) \mid p_i(x_n) - x_n \rangle &= \langle x - (p_i(x_n) - z_n)/2 - p_i(x_n) \mid p_i(x_n) - x_n \rangle \\ &\geq \langle x - p_i(x_n) \mid p_i(x_n) - x_n \rangle \end{aligned}$$

and, after recalling that $x \in H(x_n, p_i(x_n))$,

$$\langle y - p_i(x_n) \mid p_i(x_n) - x_n \rangle \geq 0$$

which implies that $y \in H(x_n, p_i(x_n))$ as desired. Hence (5), together with (6) imply

$$B(z_n, \rho_{i,n}) \subset H(x_n, p_i(x_n)).$$

Now, assume $x_n \notin S_i$, and define

$$\begin{aligned} y_n &= z_n + (p_i(x_n) - x_n) / |p_i(x_n) - x_n| \\ u_n &= z_n + \rho_{i,n}(z_n - y_n). \end{aligned}$$

One easily checks that $|u_n - z_n| = \rho_{i,n}$, i.e., $u_n \in B(z_n, \rho_{i,n})$. Thus, using (2.2), we get $u_n \in H(x_n, p_i(x_n))$, i.e., $\langle u_n - p_i(x_n) \mid p_i(x_n) - x_n \rangle \geq 0$. Therefore

$$\langle z_n - p_i(x_n) \mid p_i(x_n) - x_n \rangle \geq \langle z_n - u_n \mid p_i(x_n) - x_n \rangle = \rho_{i,n} |p_i(x_n) - x_n|. \tag{7}$$

Since this inequality is obviously satisfied when $x_n \in S_i$, it holds whether $x_n \notin S_i$ or not. Now, let

$$\rho_n = \min_{i \in I_n} \rho_{i,n}$$

and

$$\rho = \inf_{n \in \mathbb{N}} \rho_n.$$

The case: $\rho = 0$.

In the case where $\rho = 0$, we have

- either there exist n_0 in \mathbb{N} and i_0 in I_{n_0} , such that $\rho_{i_0, n_0} = 0$,
- or $\liminf_{n \rightarrow \infty} \rho_n = 0$.

In the first case, since $b(\cdot)$ vanishes only at zero, we have $z_{n_0} = p_{i_0}(x_{n_0})$ and $p_{i_0}(x_{n_0}) \in T_{n_0} \subset \bigcap_{i \in I_{n_0}} S_i$. According to (3), $p_{i_0}(x_{n_0}) \in \bigcap_{i \in I} S_i$, and for all $n > n_0$, $x_n = p_{i_0}(x_{n_0})$. Therefore, the sequence $(x_n)_{n \in \mathbb{N}}$ has converged to a solution in a finite number of steps. In the second case, there exists a subsequence $(\rho_{\sigma(n)})_{n \in \mathbb{N}}$ which converges to zero. Fix i in I . Since $b(\cdot)$ is nondecreasing,

$$\rho_{\sigma(n)} \geq \min_{j \in I_{\sigma(n)}} b\left(\left| |z_{\sigma(n)} - p_i(x_{\sigma(n)})| - |p_i(x_{\sigma(n)}) - p_j(x_{\sigma(n)})| \right|\right). \tag{8}$$

According to Lemma 3(ii) and (iii), $(|p_i(x_n) - p_j(x_n)|)_{n \in \mathbb{N}}$ converges to zero for all j in I , and $(|z_n - p_i(x_n)|)_{n \in \mathbb{N}}$ converges to a limit c independent of i . Hence, for all $\epsilon > 0$ there exists N in \mathbb{N} such that for all $n \geq N$, $|z_{\sigma(n)} - p_i(x_{\sigma(n)})| \geq c - \epsilon/2$ and $|p_i(x_n) - p_j(x_n)| \leq \epsilon/2$ for all j in I . Thus, for all $n \geq N$, and for all j in I , $|z_{\sigma(n)} - p_i(x_{\sigma(n)})| - |p_i(x_{\sigma(n)}) - p_j(x_{\sigma(n)})| \geq c - \epsilon$.

Assume $c > 0$, and take $\epsilon = c/2$. Then, since $b(\cdot)$ is nondecreasing and vanishes only at zero, we deduce from (8) that for all $n \geq N$, $\rho_{\sigma(n)} \geq b(c/2) > 0$, which contradicts $\lim_{n \rightarrow \infty} \rho_{\sigma(n)} = 0$. Hence, $c = 0$, and $\lim_{n \rightarrow \infty} |z_n - p_i(x_n)| = 0$ for all i in I . Using Lemma 3(ii), we deduce that $\lim_{n \rightarrow \infty} |x_n - z_n| = 0$, and it results from Lemma 3(iv) that $(x_n)_{n \in \mathbb{N}}$ converges strongly to a point in $\bigcap_{i \in I} S_i$.

The case $\rho \neq 0$.

In the case $\rho \neq 0$, we deduce from (7) that

$$\langle z_n - p_i(x_n) \mid p_i(x_n) - x_n \rangle \geq \rho |p_i(x_n) - x_n|$$

and since $\langle p_i(x_n) - x_n \mid p_i(x_n) - x_n \rangle = |p_i(x_n) - x_n|^2 \geq 0$,

$$\langle z_n - x_n \mid p_i(x_n) - x_n \rangle \geq \rho |p_i(x_n) - x_n|. \tag{9}$$

Therefore, combining the definition of x_{n+1} and (9), we have

$$\begin{aligned} \lambda_n \sum_{i \in I_n} \alpha_{i,n} |p_i(x_n) - x_n| &\leq \frac{\lambda_n}{\rho} \sum_{i \in I_n} \alpha_{i,n} \langle z_n - x_n, p_i(x_n) - x_n \rangle \\ &= \frac{1}{\rho} \left\langle z_n - x_n, \lambda_n \sum_{i \in I_n} \alpha_{i,n} (p_i(x_n) - x_n) \right\rangle \\ &= \frac{1}{\rho} \langle z_n - x_n, x_{n+1} - x_n \rangle. \end{aligned}$$

Using Lemma 2(ii) and Lemma 3(i), we deduce that the series $\sum_{n \in \mathbb{N}} |x_{n+1} - x_n|$ is convergent. Then $(x_n)_{n \in \mathbb{N}}$ is a Cauchy sequence, and is therefore strongly convergent to a point x^* in \mathcal{H} . Using Lemma 3(ii), we deduce that $(p_i(x_n))_{n \in \mathbb{N}}$ is also strongly convergent to x^* for any i in I . Since each set S_i is closed, we immediately conclude that $x^* \in \bigcap_{i \in I} S_i$, which completes the proof. \square

4. Projections onto Stepwise Generated Uniformly Convex Sets

In this section, we show how the results of Section 3 may be used to define a strongly convergent cyclic projection algorithm in the case where the sets S_i are only expandable in convex sets, but not into uniformly convex sets. To our knowledge, this method is new, even in the convex case. The underlying idea of the algorithm is as follows. First, note that the condition C_4 is realistic for the type of nonconvex sets considered in this paper, and may be interpreted as a strong consistency condition. On the other hand, in many cases, a trivial sequence $(a_n)_{n \in \mathbb{N}}$ where $a_n \in T_n$ for all n in \mathbb{N} is already known, as in the applications presented in Section 5. Using this sequence, we define a projection method which

converges strongly to a point in $\bigcap_{i \in I} S_i$. For every n in \mathbb{N} , i in I , and a_n in T_n , $B_{i,n}$ will denote the closed ellipsoid with main axis $[a_n, p_i(x_n)]$, with center $1/2(a_n + p_i(x_n))$, which is rotationally invariant around this axis and with maximal radius equal to $|a_n - p_i(x_n)|/2$ and with minimal radius equal to $\gamma|a_n - p_i(x_n)|/2$ for some $\gamma \in (0, 1)$. We denote by $p'_i(x_n)$ the projection of x_n onto $B_{i,n}$; see Figure 1.

Definition 5. (Method of Double Projection onto Convex Expandable Sets)

Assume C_4 is satisfied. Given a point x_0 in \mathcal{H} , two real number α and λ in $(0, 1]$, a nonnegative integer M , a sequence (I_n) of subsets of I and a sequence $(a_n)_{n \in \mathbb{N}}$ where $a_n \in T_n$ for all n in \mathbb{N} , the projection method onto stepwise generated uniformly convex sets is iteratively defined by

$$(\forall n \in \mathbb{N}) \begin{cases} 1. & (\forall i \in I_n) \quad p'_i(x_n) = p_{B_{i,n}}(x_n) \\ & \quad - \text{If } \exists j \in I_n \text{ such that } p'_j(x_n) \in \bigcap_{i \in I_n} S_i \\ 2. & \quad \text{Then} \begin{cases} - \text{If } I_n \neq I \text{ set } I_n = I \text{ and go to 2} \\ - \text{If } I_n = I \text{ set } \lambda_n = 1, \alpha_{j,n} = 1 \text{ and go to 3} \\ - \text{Else go to 3} \end{cases} \\ 3. & x_{n+1} = x_n + \lambda_n \sum_{i \in I_n} \alpha_{i,n} (p'_i(x_n) - x_n) \end{cases}$$

under the conditions C_1, C_2 , and

$$C'_3. \quad (\forall n \in \mathbb{N}) \quad \lambda \leq \lambda_n \leq \mu_n \text{ where } \mu_n = \begin{cases} \frac{\sum_{i \in I_n} \alpha_{i,n} |x_n - p'_i(x_n)|^2}{|x_n - \sum_{i \in I_n} \alpha_{i,n} p'_i(x_n)|^2} & \text{if } x_n \notin \bigcap_{i \in I_n} B_{i,n} \\ 1 & \text{if } x_n \in \bigcap_{i \in I_n} B_{i,n}. \end{cases}$$

In the remainder of this section, the sequence $(x_n)_{n \in \mathbb{N}}$ is defined by the algorithm of Definition 5. The main result is the following.

Theorem 2. If the condition C_5 is satisfied, and the sequence $(a_n)_{n \in \mathbb{N}}$ introduced in Definition 5 satisfies C_6 , then the sequence $(x_n)_{n \in \mathbb{N}}$ converges strongly to a point in $\bigcap_{i \in I} S_i$.

Proof. The method introduced in Definition 5 is a general projection algorithm onto specific uniformly convex sets constructed at each iteration. For every n in \mathbb{N} and each i in I , a_n and $p'_i(x_n)$ belong to $B_{i,n}$. Then Lemma 2 and Lemma 3 where $p_i(x_n)$ and $(z_n)_{n \in \mathbb{N}}$ are respectively replaced with $p'_i(x_n)$ and $(a_n)_{n \in \mathbb{N}}$ hold true. In the same way, taking $(z_n)_{n \in \mathbb{N}} = (a_n)_{n \in \mathbb{N}}$ in the proof of Theorem 1, we deduce that the sequence $(x_n)_{n \in \mathbb{N}}$ converges strongly to some point x^* in \mathcal{H} . Nevertheless, since $p_i(x_n)$ is replaced with $p'_i(x_n)$, it is still not clear why $x^* \in \bigcap_{i \in I} S_i$. Let us now address this specific point. Fix n in \mathbb{N} , i in I , and a_n in T_n . According to C_5 , a_n and $p_i(x_n)$ belong to a convex set $C_{i,j_n} \subset S_i$, and $p_i(x_n)$ is also the projection of x_n onto C_{i,j_n} . Hence the Kolmogorov criterion (1) gives $\langle p_i(x_n) - a_n \mid x_n - p_i(x_n) \rangle \geq 0$. Therefore

$$\langle p_i(x_n) - a_n \mid x_n - p'_i(x_n) \rangle + \langle p_i(x_n) - a_n \mid p'_i(x_n) - p_i(x_n) \rangle \geq 0$$

which implies

$$\langle p_i(x_n) - a_n \mid x_n - p'_i(x_n) \rangle \geq \langle p_i(x_n) - a_n \mid p_i(x_n) - p'_i(x_n) \rangle. \tag{10}$$

Moreover, since the segment $[a_n, p_i(x_n)]$ is the main axis of the ellipsoid $B_{i,n}$ and $p'_i(x_n) \in B_{i,n}$, we have $\langle a_n - p_i(x_n) \mid p'_i(x_n) - p_i(x_n) \rangle \geq 0$. Combining this with (10), we get

$$0 \leq \langle p_i(x_n) - a_n \mid p_i(x_n) - p'_i(x_n) \rangle \leq \langle p_i(x_n) - a_n \mid x_n - p'_i(x_n) \rangle \leq |p_i(x_n) - a_n| |x_n - p'_i(x_n)|.$$

We deduce from Lemma 3(iii) where $p_i(x_n)$ is replaced with $p'_i(x_n)$ that $\lim_{n \rightarrow \infty} |p'_i(x_n) - x_n| = 0$. On the other hand, as $p_i(x_n)$ is the projection of x_n onto C_{i,j_n} , and $a_n \in C_{i,j_n}$, we have $|p_i(x_n) - a_n| \leq |x_n - a_n|$. Moreover, following the same steps as in the proof of Lemma 3(i) gives that $(|x_n - a_n|)_{n \in \mathbb{N}}$

is bounded, from which we deduce that $(|p_i(x_n) - a_n|)_{n \in \mathbb{N}}$ is bounded as well. Therefore we deduce from the last inequality that

$$\lim_{n \rightarrow \infty} \langle a_n - p_i(x_n) \mid p'_i(x_n) - p_i(x_n) \rangle = 0. \tag{11}$$

Let $x = a_n - p_i(x_n)$ and $y = p'_i(x_n) - p_i(x_n)$. Then, we have

$$x - y = a_n - p'_i(x_n)$$

and

$$x - \omega y = \omega \left(\frac{1}{\omega} (a_n + (\omega - 1)p_i(x_n)) - p'_i(x_n) \right) \tag{12}$$

for all $\omega \in \mathbb{R}$.

Assume that $p'_i(x_n) \neq p_i(x_n)$. Using Claim 1, we get that

$$\left\langle a_n - p'_i(x_n) \mid \frac{1}{\omega_{i,n}} (a_n + (\omega_{i,n} - 1)p_i(x_n)) - p'_i(x_n) \right\rangle \leq 0$$

as long as we enforce

$$\omega_{i,n} \geq \frac{\langle p'_i(x_n) - a_n \mid p_i(x_n) - a_n \rangle}{\langle p'_i(x_n) - a_n \mid p_i(x_n) - p'_i(x_n) \rangle}. \tag{13}$$

In particular, we can take

$$\omega_{i,n} = \max \left\{ \frac{\langle p'_i(x_n) - a_n \mid p_i(x_n) - a_n \rangle}{\langle p'_i(x_n) - a_n \mid p_i(x_n) - p'_i(x_n) \rangle}, \epsilon \right\} \tag{14}$$

for any appropriately chosen $\epsilon > 0$. Using Niculescu’s Lemma 4 in Appendix B, we obtain that

$$\left(\sqrt{\frac{1}{\omega_{i,n}}} + \sqrt{\frac{\omega_{i,n}}{1}} \right) \frac{\langle a_n - p_i(x_n), p'_i(x_n) - p_i(x_n) \rangle}{2|a_n - p_i(x_n)|} \geq |p'_i(x_n) - p_i(x_n)|. \tag{15}$$

Note that this inequality also holds without resorting to Claim 1 if $p'_i(x_n) = p_i(x_n)$. According to Theorem 1, $(x_n)_{n \in \mathbb{N}}$ and $(p'_i(x_n))_{n \in \mathbb{N}}$ are strongly convergent to a point x^* in \mathcal{H} . We now split the remainder of the argument into the following complementary cases involving an increasing function $\sigma: \mathbb{N} \mapsto \mathbb{N}$ which parametrises possible subsequences.

- If the sequence $(|a_n - p'_i(x_n)|)_{n \in \mathbb{N}}$ converges to zero, then using Claim 2, $(|p'_i(x_n) - p_i(x_n)|)_{n \in \mathbb{N}}$ also converges to zero. Therefore, $(p_i(x_n))_{n \in \mathbb{N}}$ is also strongly convergent to x^* for each i in I . Since each set S_i is closed, we again conclude that $x^* \in \bigcap_{i \in I} S_i$.
- If the sequence $(|a_n - p'_i(x_n)|)_{n \in \mathbb{N}}$ does not converge to zero, and the sequence $(\langle p'_i(x_n) - a_n \mid p_i(x_n) - p'_i(x_n) \rangle)_{n \in \mathbb{N}}$ converges to zero, then we must have $|a_{\sigma(n)} - p'_i(x_{\sigma(n)})| > \epsilon$ for some $\epsilon > 0$ and some subsequence indexed by an appropriately chosen increasing function σ . This implies in particular that

$$\lim_{n \rightarrow +\infty} |p_i(x_{\sigma(n)}) - p'_i(x_{\sigma(n)})| \left| \cos \left(p'_i(x_{\sigma(n)}) - a_{\sigma(n)}, p_i(x_{\sigma(n)}) - p'_i(x_{\sigma(n)}) \right) \right| = 0.$$

As a result, either

$$\lim_{n \rightarrow +\infty} \cos \left(p'_i(x_{\sigma(n)}) - a_{\sigma(n)}, p_i(x_{\sigma(n)}) - p'_i(x_{\sigma(n)}) \right) = 0 \tag{16}$$

or

$$\lim_{n \rightarrow +\infty} |p_i(x_{\sigma(n)}) - p'_i(x_{\sigma(n)})| = 0.$$

Notice further that (16) holds only if $(|p_i(x_{\sigma(n)}) - p'_i(x_{\sigma(n)})|)_{n \in \mathbb{N}}$ converges to zero. Thus, both cases simplify into the conclusion that $(|p_i(x_{\sigma(n)}) - p'_i(x_{\sigma(n)})|)_{n \in \mathbb{N}}$ converges to zero. Therefore $(p_i(x_{\sigma(n)}))_{n \in \mathbb{N}}$ is also strongly convergent to x^* for each i in I . Since each set S_i is closed, we again conclude that $x^* \in \bigcap_{i \in I} S_i$.

- If both sequences $(|a_n - p'_i(x_n)|)_{n \in \mathbb{N}}$ and $(\langle p'_i(x_n) - a_n \mid p_i(x_n) - p'_i(x_n) \rangle)_{n \in \mathbb{N}}$ do not converge to zero, then $|a_{\sigma(n)} - p'_i(x_{\sigma(n)})| > \epsilon$ and $\langle p'_i(x_{\sigma(n)}) - a_{\sigma(n)} \mid p_i(x_{\sigma(n)}) - p'_i(x_{\sigma(n)}) \rangle > \epsilon$ for some $\epsilon > 0$ and for some subsequence indexed by an appropriately chosen increasing function σ . Since $|a_n - p_i(x_n)|$ is the length of the main axis of $B_{i,n}$, and, as such, bounds from above the distance between any two points in $B_{i,n}$, we deduce that $|a_n - p'_i(x_n)| \leq |a_n - p_i(x_n)|$. Furthermore, since the sequences $(|a_n - p_i(x_n)|)_{n \in \mathbb{N}}$ and $(|a_n - x_n|)_{n \in \mathbb{N}}$ are bounded, we deduce that

$$|a_n - p_i(x_n)| \leq |a_n - x_n| \leq B$$

for some $B > 0$. Then using the Cauchy-Schwarz inequality in the numerator in (14), we get

$$\omega_{i,\sigma(n)} \in [\epsilon, \frac{B^2}{\epsilon}]$$

for some $\epsilon \in (0, B)$. Since $|a_{\sigma(n)} - p_i(x_{\sigma(n)})| > \epsilon$, we deduce from (11) and (15) that $(|p'_i(x_{\sigma(n)}) - p_i(x_{\sigma(n)})|)_{n \in \mathbb{N}}$ converges to zero. Therefore, $(p_i(x_{\sigma(n)}))_{n \in \mathbb{N}}$ is also strongly convergent to x^* for each i in I . Since each set S_i is closed, we again conclude that $x^* \in \bigcap_{i \in I} S_i$.

These previous cases cover all the possible cases and all lead to the conclusion that $x^* \in \bigcap_{i \in I} S_i$. The proof is thus complete. \square

The two following claims were instrumental in the proof of Theorem 2. We now provide their proofs.

Claim 1. Assume that $p'_i(x_n) \neq p_i(x_n)$. Then

$$\frac{\langle p'_i(x_n) - a_n \mid p_i(x_n) - a_n \rangle}{\langle p'_i(x_n) - a_n \mid p_i(x_n) - p'_i(x_n) \rangle} \tag{17}$$

is a well defined nonnegative real number. Moreover, we have

$$\left\langle a_n - p'_i(x_n) \mid \frac{1}{\omega_{i,n}}(a_n + (\omega_{i,n} - 1)p_i(x_n)) - p'_i(x_n) \right\rangle \leq 0 \tag{18}$$

for all

$$\omega_{i,n} \geq \frac{\langle p'_i(x_n) - a_n \mid p_i(x_n) - a_n \rangle}{\langle p'_i(x_n) - a_n \mid p_i(x_n) - p'_i(x_n) \rangle}.$$

Proof. We have $p'_i(x_n) \notin [a_n, p_i(x_n)]$. Indeed, assume for contradiction that $p'_i(x_n) \in [a_n, p_i(x_n)]$. This would imply that $p'_i(x_n)$ lies in the interior of $B_{i,n}$ and therefore $x_n = p'_i(x_n)$. However, since by convexity of C_{i,j_n} , $[a_n, p_i(x_n)] \subset C_{i,j_n}$, this would imply that $x_n \in C_{i,j_n}$, and therefore $p'_i(x_n) = x_n = p_i(x_n)$, the sought-after contradiction.

Consider now the 2D plane containing $a_n, p_i(x_n)$ and $p'_i(x_n)$. Set g_n to be at the intersection of the line perpendicular at $p'_i(x_n)$ to $[a_n, p'_i(x_n)]$ with the segment $[a_n, p_i(x_n)]$ (see Figure 1). First, we have

$$\langle p'_i(x_n) - a_n \mid p_i(x_n) - a_n \rangle = \cos(\alpha_{i,n}) |p'_i(x_n) - a_n| |p_i(x_n) - a_n|$$

and

$$|p'_i(x_n) - a_n|^2 = \langle g_n - a_n \mid p'_i(x_n) - a_n \rangle = \cos(\alpha_{i,n}) |g_n - a_n| |p'_i(x_n) - a_n|$$

from which we obtain

$$|g_n - a_n| = \frac{|p'_i(x_n) - a_n|^2 |p_i(x_n) - a_n|}{\langle p'_i(x_n) - a_n \mid p_i(x_n) - a_n \rangle} \tag{19}$$

Since $B_{i,n}$ is an ellipse with maximal axis $[a_n, p_i(x_n)]$, our next step is to express g_n as the convex combination

$$g_n = \theta_{i,n} a_n + (1 - \theta_{i,n}) p_i(x_n) \tag{20}$$

of a_n and $p_i(x_n)$ for some $\theta_{i,n} \in [0, 1]$. Notice that (20) implies

$$|g_n - a_n| = (1 - \theta_{i,n}) |p_i(x_n) - a_n|,$$

which itself gives

$$\theta_{i,n} = 1 - \frac{|g_n - a_n|}{|p_i(x_n) - a_n|}.$$

Using (19), we get

$$\theta_{i,n} = 1 - \frac{|p'_i(x_n) - a_n|^2}{\langle p'_i(x_n) - a_n \mid p_i(x_n) - a_n \rangle} = \frac{\langle p'_i(x_n) - a_n \mid p_i(x_n) - p'_i(x_n) \rangle}{\langle p'_i(x_n) - a_n \mid p_i(x_n) - a_n \rangle}.$$

Notice further that, by construction, since $[a_n, p_i(x_n)]$ is the main axis of $B_{i,n}$ and since, by assumption, $p'_i(x_n) \neq p_i(x_n)$, we have (see Figure 1)

$$\langle p'_i(x_n) - a_n \mid p_i(x_n) - p'_i(x_n) \rangle > 0. \tag{21}$$

Then, setting

$$d_n = \frac{1}{\omega_{i,n}} (a_n + (\omega_{i,n} - 1) p_i(x_n)),$$

we see that we need to take $\omega_{i,n} \geq 1/\theta_{i,n}$, i.e.,

$$\omega_{i,n} \geq \frac{\langle p'_i(x_n) - a_n \mid p_i(x_n) - a_n \rangle}{\langle p'_i(x_n) - a_n \mid p_i(x_n) - p'_i(x_n) \rangle},$$

(which, owing to (21), is well defined), in order to ensure that

$$\langle a_n - p'_i(x_n) \mid d_n - p'_i(x_n) \rangle \leq 0$$

i.e., (18) holds. Finally, (19) and (21) together ensure that (17) is nonnegative. \square

Claim 2. We have

$$|a_n - p'_i(x_n)| \geq |p'_i(x_n) - p_i(x_n)|.$$

Proof. By definition of $p_i(x_n)$, the fact that $a_n \in S_i$ implies that

$$|p_i(x_n) - x_n| \leq |a_n - x_n|.$$

Thus x_n belongs to the half space

$$H = \left\{ x \in \mathcal{H}; \langle x - \frac{1}{2}(a_n + p_i(x_n)) \mid p_i(x_n) - \frac{1}{2}(a_n + p_i(x_n)) \rangle \geq 0 \right\}$$

which is the set of points in \mathcal{H} which are closer to $p_i(x_n)$ than a_n . Clearly, $p'_i(x_n) \in H$ as well, which completes the proof. \square

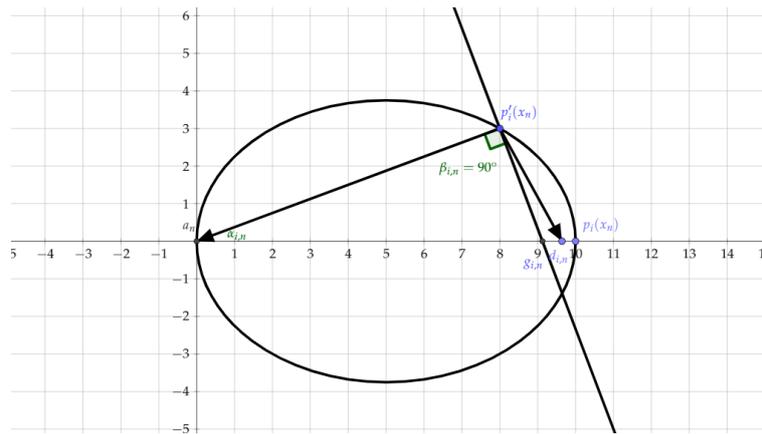


Figure 1. Section of the ellipsoid $B_{i,n}$.

5. Applications

5.1. MRI Image Reconstruction: The Infinite-Dimensional Hilbert Space Setting

The field of inverse problems was extensively investigated in the last five decades, fueled in particular by the many challenged in medical imaging. Projection methods have played an important role in the development of efficient reconstruction techniques [25], a well studied example being the method of projection onto convex sets (POCS) [2]. In recent years, techniques from penalised estimation have gained increasing popularity since the discovery of the compressed sensing paradigm, allowing for fewer measurements to be collected whilst achieving remarkable reconstruction performance. One penalisation approach which has been a center of focus since its introduction is the Rudin Osher Fatemi functional [26], which can be described as follows: the reconstructed image u is the solution of the minimization problem,

$$\arg \min_{u \in BV(\Omega)} \|u\|_{TV(\Omega)} + \frac{\lambda}{2} \int_{\Omega} (y_{obs}(x) - \mathcal{A}[u](x))^2 dx$$

where λ is a positive relaxation parameter and the term $\|u\|_{TV(\Omega)}$ is defined in Appendix A. The term $\int_{\Omega} (y_{obs}(x) - \mathcal{A}[u](x))^2 dx$ is called the fidelity term and the term $\|u\|_{TV(\Omega)}$ is called the regularisation term. This infinite-dimensional minimization problem enforces the solution, which is a priori in $L^2(\Omega)$, to satisfy a finite bounded variation (BV) norm constraint.

One important remark is that the optimisation problem can be turned into a pure feasibility problem by imposing the TV-norm and the fidelity terms to be less than or equal to a certain tolerance. Using this approach, one can define an alternating projection procedure as in Algorithm 1 below, where we will use the forward operator $\mathcal{A} = \mathcal{F}_{partial}$, the partial Fourier Transform which plays a central role in MRI. The usual Fourier transform will be denoted by \mathcal{F} as usual.

In this example, we will use the projections on the sets

$$S_1 = \left\{ u \in L^2(\Omega) \mid \int_{\Omega} (\lambda y_{obs} - \mathcal{A}[u](x))^2 dx \leq \lambda \eta_{fid}, \lambda \in \mathbb{R}_+ \right\},$$

$$S_2 = \left\{ u \in L^2(\Omega) \mid \|u\|_{p,TV(\Omega)} \leq \eta_{TV} \right\},$$

$$S_3 = \left\{ u \in L^2(\Omega) \mid \mathcal{F}_{partial}[u] = \lambda y_{obs}, \lambda \in \mathbb{R}_+ \right\}.$$

The set S_2 is not convex for $p \in (0, 1)$, but, since the epigraph of $|\cdot|_{p,TV(\Omega)}$ is a cone, the set S_2 is expandable into convex sets as discussed in [22,23]. Notice that when the scaling factor λ is dropped in the definition of S_1 and S_3 , the problem is equivalent to the standard TV -penalised reconstruction approach. Introducing this scaling factor allows to recover a solution up to a scaling, which can easily be recovered by rescaling the solution by using the constraint $\mathcal{F}_{partial}[u] = y_{obs}$ as a post-processing step. The introduction of the scaling α in these definition allows the null function to belong to the intersection of the sets S_1, S_2 and S_3 . Thus, we will set $a_n = 0, n \in \mathbb{N}$ in the sequel. In what follows, \tilde{y}_{obs} will be the function will value equal to zero except for the observed frequencies which are taken as the observed values.

Our implementation is described in Algorithm 1. Experiments were made after incorporating the projection onto the TV ball into a freely available code provided in [27]. The fidelity term was set to $\eta_{fid} = 10^{-3}$ and the regularisation term η_{TV} was tuned using cross validation. The projection onto the TV -ball was computed using the method described in [28]. In order to avoid numerical instabilities, the iterates were rescaled every 10 iteration, which does not change the convergence analysis due to conic invariance of our formulation. We chose very thin ellipsoids $B_{i,n}$ by setting $\alpha = 10^{-6}$. A numerical illustration is given in Figure 2 below. Extensive numerical simulations will be presented in a forthcoming paper devoted to a thorough study of projection methods for MRI and X-ray computed tomography (XCT).

Algorithm 1: Alternating projection method for MRI for $p = 1$.

Result:

- 1 $y = \mathcal{F}_{partial}[u_0]$ observed data;
 - 2 η_{TV} bound on the TV norm;
 - 3 η_{fid} bound on the fidelity term;
 - 4 ϵ convergence tolerance;
 - 5 Compute $y^{(0)}$ from y by filling in unobserved with zeros ;
 - 6 Set $k = 1, u_1 = \mathcal{F}^{-1}(\tilde{y}_{obs})$;
 - 7 **while** Difference between two successive iterates larger than ϵ **do**
 - 8 Compute $p'_1(u_n) = P_{B_{1,n}}(u_n)$;
 - 9 Compute $p'_2(u_n) = P_{B_{2,n}}(u_n)$;
 - 10 Compute $p'_3(u_n) = P_{B_{3,n}}(u_n)$;
 - 11 Set $u_{n+1} = u_n + \lambda_n \sum_{i=1}^3 \alpha_{i,n} (p'_i(u_n) - u_n)$;
 - 12 **end while**
-

Since S_1 and S_3 are convex and S_2 is expandable into convex sets, we deduce from Theorem 2 that Algorithm 1 converges strongly to a point in $\bigcap_{i=1}^3 S_i$.

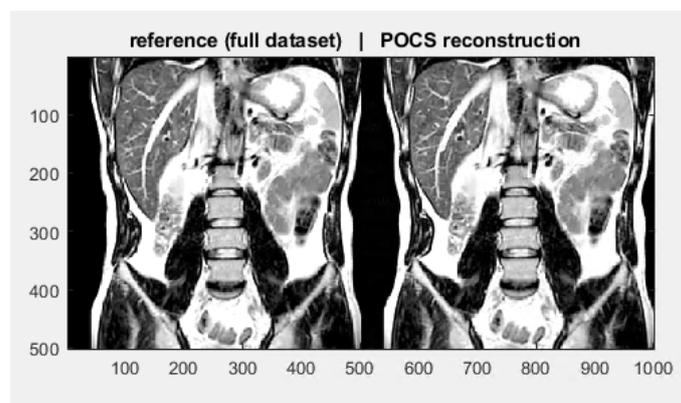


Figure 2. POCS reconstruction with 30 iterations.

5.2. A Uniformly Convex Version of Cadzow’s Method

The method presented in Section 4 is now applied to a denoising problem in the spirit of Cadzow’s method [16].

Let $x = [x_0, \dots, x_{N-1}]^\top$ be a complex-valued vector of length N . Let \mathcal{H} be a linear operator which maps x into the set of Hankel $L \times K$ matrices, with $L \geq K, L + K = N + 1$, defined by setting the (i, j) component to the value x_{i+j-2} i.e.,

$$\mathcal{H}(x) = \begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_{K-1} \\ x_1 & x_2 & x_3 & \cdots & x_K \\ x_2 & x_3 & x_4 & \cdots & x_{K+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{L-1} & x_L & x_{L+1} & \cdots & x_{N-1} \end{bmatrix}.$$

The adjoint operator associated with \mathcal{H} , denoted by \mathcal{H}^* , is a linear map from $L \times K$ matrices to vectors of length N , obtained by averaging each skew-diagonal.

The denoising problem we consider in this section is the one of estimating a signal x corrupted by random noise, as described in the observation model

$$y_t = x_t + \epsilon_t,$$

for $t = 0, \dots, N - 1$, where

- x_t is a sum of damped exponential components, i.e.,

$$x_t = \sum_{k=-\kappa}^{\kappa} c_k \rho_k^t \exp(2i\pi f_k t),$$

with $2\kappa + 1 \leq K, \rho_k$ a real damping factor satisfying $|\rho_k| \leq 1, f_k$ a real number representing a frequency and c_k a complex coefficient,

- ϵ_t is a random noise.

This denoising problem is of crucial importance in many applications and is also known as super-resolution in the literature; see [16,29–33]. The main motivation behind Cadzow’s algorithm is that many signals of interest have low rank Hankelization. In mathematical terms, we can often look for approximations and denoising, for signals which satisfy the constraint that $\text{rank}(\mathcal{H}(x)) = r$ where $r \leq 2\kappa + 1$. Such constraints are important when the observed signal is corrupted with additive noise, which increases the rank of $\mathcal{H}(x)$. Consequently, an intuitive way of estimating x from y is to do *rank reduction*: starting from $x^{(0)} = y = [y_0, \dots, y_{N-1}]^\top$, the Cadzow’s algorithm iteratively updates the estimate via the following rule

$$x^{(n+1)} = \mathcal{H}^*(\mathcal{T}_r(\mathcal{H}(x^{(n)}))), \quad n = 0, 1, \dots$$

Here \mathcal{T}_r computes the truncated singular value decomposition (SVD) of any $L \times K$ matrix X , that is,

$$\mathcal{T}_r(X) = \sum_{j=1}^r \sigma_j u_j v_j^*,$$

with

$$X = \sum_{j=1}^K \sigma_j u_j v_j^*$$

being an SVD of X and with $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_K$ being the singular values. As is now well known, Cadzow’s algorithm is easily interpreted as a method of alternating projections in the matrix domain [23]. Denote by $\mathcal{P}_{\mathcal{M}_r}$ the projection onto the set \mathcal{M}_r of $L \times K$ matrices of rank r , and by $\mathcal{P}_{\mathcal{M}_H}$ the projection onto the space \mathcal{M}_H of $L \times K$ Hankel matrices. Then, Cadzow’s method is easily seen to be equivalent to the following matrix recursions

$$X^{(n+1)} = \mathcal{P}_{\mathcal{M}_H} \left(\mathcal{P}_{\mathcal{M}_r}(X^{(n)}) \right), \quad n = 0, 1, \dots$$

Since \mathcal{M}_H is convex and \mathcal{M}_r is expandable into convex sets, we deduce from the main result of [22] that Cadzow’s method has cluster points in $\mathcal{M}_H \cap \mathcal{M}_r$. On the other hand, applying the method of Section 4 can be advantageous in the application of projection methods to the case of low rank matrix constraints such as in the application of Cadzow’s method. In particular, our method shows that the SVD can be computed only approximately in the first iterations and still belong to the ellipsoids $B_{i,n}$, thus implying possibly important computational savings. Our new method converges to a point in $\mathcal{M}_H \cap \mathcal{M}_r$.

An example showing the efficiency of Cadzow’s method is presented in Figure 3 below. In this experiment, we set the signal-to-noise ratio (SNR) to 3.5 dB and the rank constraint equal to 15 (an appropriate choice for the rank can be made using statistical techniques such as cross-validation). We took $K = \lfloor N/2 - 10 \rfloor$.

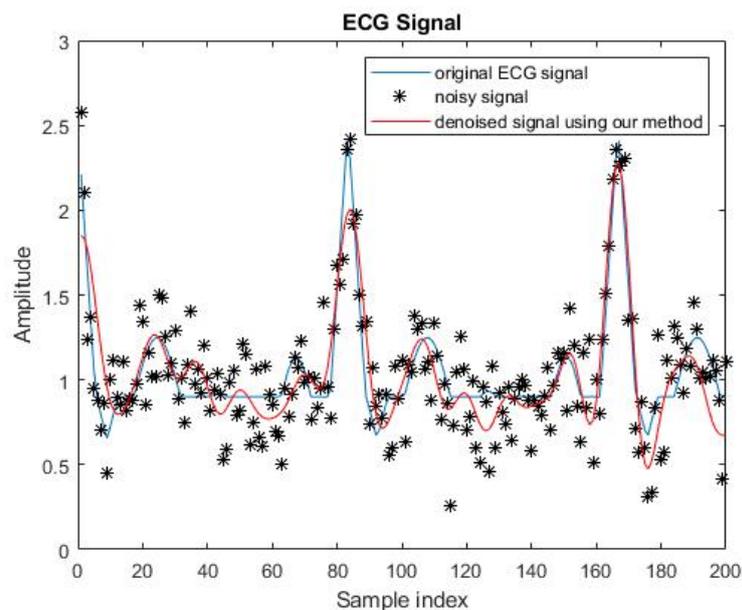


Figure 3. Application of our method to denoising a benchmark Electrocardiogram (ECG) type signal.

6. Conclusions and Future Work

The work presented in this paper is a follow on project to the work of [22]. We obtain strong convergence results in Hilbert spaces in the case of uniformly convex expandable sets, a notion which refines the definition of convex expandable sets introduced in [22,23]. We showed that the proposed methods apply to practical inverse and denoising problems which are instrumental in engineering and signal and data analytics.

Future work is planned in developing new and faster algorithms for nonconvex feasibility problems, using acceleration schemes such as Richardson’s extrapolation [34].

Author Contributions: S.C. and P.B. contributed to all the mathematical findings presented in the paper and to the various stages of the writing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially supported by the iCODE Institute, research project of the IDEX Paris-Saclay, and by the Hadamard Mathematics LabEx (LMH) through the grant number ANR-11-LABX-0056-LMH in the “Programme d’Investissement d’Avenir”.

Acknowledgments: The authors would like to thank the reviewers for the high quality of their comments, which greatly helped improve the quality of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Definition of the Total Variation (TV) Norm

Recall that a function u is in $BV(\Omega)$ for a bounded open set $\Omega \subset \mathbb{R}^d$ if it is integrable and there exists a Radon measure Du such that

$$\int_{\Omega} u(x) \operatorname{div} g(x) dx = - \int_{\Omega} \langle g, Du(x) \rangle dx$$

for all $g \in C_c^1(\Omega, \mathbb{R}^d)$. This measure Du is the distributional gradient of u . When u is smooth, $Du(x) = \nabla u(x) dx$, and TV-norm is equivalently the integral of its gradient magnitude,

$$\|u\|_{\text{TV}(\Omega)} = \int_{\Omega} |\nabla u| dx.$$

For $p \in [1, +\infty]$, the L^p -TV norm is defined by

$$\|u\|_{p, \text{TV}(\Omega)} = \left(\int_{\Omega} |\nabla u|^p dx \right)^{\frac{1}{p}}. \tag{A1}$$

For $p \in (0, 1)$, formula (A1) defines a quasi-norm which is nonconvex, called the L^p -TV norm by abuse of language, as is common in the image processing community.

Appendix B. Niculescu’ Lemma

Niculescu’s Lemma [35] (Theorem 2), [36] is a converse of the Cauchy-Schwarz inequality. It can be stated as follows.

Lemma 4. Assume there exist two positive real numbers ω and ω' such that

$$\langle x - \omega y \mid x - \omega' y \rangle \leq 0.$$

Then, we have

$$\frac{\langle x, y \rangle}{\langle x, x \rangle^{\frac{1}{2}} \langle y, y \rangle^{\frac{1}{2}}} \geq \frac{2}{\sqrt{\frac{\omega}{\omega'}} + \sqrt{\frac{\omega'}{\omega}}}.$$

References

1. Bauschke, H.H.; Combettes, P.L. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*; Springer: Berlin/Heidelberg, Germany, 2011; Volume 408.
2. Escalante, R.; Raydan, M. *Alternating Projection Methods*; SIAM: Philadelphia, PA, USA, 2011.
3. Gurin, L.; Polyak, B.T.; Raik, È.V. The method of projections for finding the common point of convex sets. *Zhurnal Vychislitel’noi Matematiki i Matematicheskoi Fiziki* **1967**, *7*, 1211–1228.
4. Ottavy, N. Strong convergence of projection-like methods in Hilbert spaces. *J. Optim. Theory Appl.* **1988**, *56*, 433–461. [[CrossRef](#)]
5. Combettes, P.L.; Puh, H. Iterations of parallel convex projections in Hilbert spaces. *Numer. Funct. Anal. Optim.* **1994**, *15*, 225–243. [[CrossRef](#)]

6. Grigoriadis, K.M. Optimal \mathcal{H}_∞ model reduction via linear matrix inequalities: Continuous-and discrete-time cases. *Syst. Control Lett.* **1995**, *26*, 321–333. [CrossRef]
7. Babazadeh, M.; Nobakhti, A. Direct Synthesis of Fixed-Order \mathcal{H}_∞ Controllers. *IEEE Trans. Autom. Control* **2015**, *60*, 2704–2709. [CrossRef]
8. Li, Z.; Dai, Y.H.; Gao, H. Alternating projection method for a class of tensor equations. *J. Comput. Appl. Math.* **2019**, *346*, 490–504. [CrossRef]
9. Combettes, P. The convex feasibility problem in image recovery. In *Advances in Imaging and Electron Physics*; Elsevier:Amsterdam, Netherlands, 1996; Volume 95, pp. 155–270.
10. Krol, A.; Li, S.; Shen, L.; Xu, Y. Preconditioned alternating projection algorithms for maximum a posteriori ECT reconstruction. *Inverse Probl.* **2012**, *28*, 115005. [CrossRef]
11. Herman, G.T. *Fundamentals of Computerized Tomography: Image Reconstruction From Projections*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2009.
12. McGibney, G.; Smith, M.; Nichols, S.; Crawley, A. Quantitative evaluation of several partial Fourier reconstruction algorithms used in MRI. *Magn. Reson. Med.* **1993**, *30*, 51–59. [CrossRef]
13. Ticozzi, F.; Zuccato, L.; Johnson, P.D.; Viola, L. Alternating projections methods for discrete-time stabilization of quantum states. *IEEE Trans. Autom. Control* **2017**, *63*, 819–826. [CrossRef]
14. Drusvyatskiy, D.; Li, C.K.; Pelejo, D.C.; Voronin, Y.L.; Wolkowicz, H. Projection methods for quantum channel construction. *Quantum Inf. Process.* **2015**, *14*, 3075–3096. [CrossRef]
15. Grigoriadis, K.M.; Beran, E.B. Alternating projection algorithms for linear matrix inequalities problems with rank constraints. In *Advances in Linear Matrix Inequality Methods in Control*; SIAM: Philadelphia, PA, USA, 2000; pp. 251–267.
16. Cadzow, J.A. Signal enhancement—a composite property mapping algorithm. *IEEE Trans. Acoust. Speech Signal Process.* **1988**, *36*, 49–62. [CrossRef]
17. Bauschke, H.H.; Combettes, P.L.; Luke, D.R. Phase retrieval, error reduction algorithm, and Fienup variants: A view from convex optimization. *J. Opt. Soc. Am. A* **2002**, *19*, 1334–1345. [CrossRef] [PubMed]
18. Chu, M.T.; Funderlic, R.E.; Plemmons, R.J. Structured low rank approximation. *Linear Algebra Its Appl.* **2003**, *366*, 157–172. [CrossRef]
19. Markovsky, I.; Usevich, K. Structured low-rank approximation with missing data. *SIAM J. Matrix Anal. Appl.* **2013**, *34*, 814–830. [CrossRef]
20. Elser, V. Learning Without Loss. *arXiv* **2019**, arXiv:1911.00493.
21. Combettes, P.L.; Trussell, H.J. Method of successive projections for finding a common point of sets in metric spaces. *J. Optim. Theory Appl.* **1990**, *67*, 487–507. [CrossRef]
22. Chretien, S.; Bondon, P. Cyclic projection methods on a class of nonconvex sets. *Numer. Funct. Anal. Optim.* **1996**, *17*, 37–56.
23. Chrétien, S. Methodes de projection pour l’optimisation ensembliste non convexe. Ph.D, Thesis, Sciences Po, Paris, France, 1996.
24. Bauschke, H.H.; Borwein, J.M. On projection algorithms for solving convex feasibility problems. *SIAM Rev.* **1996**, *38*, 367–426. [CrossRef]
25. Censor, Y.; Chen, W.; Combettes, P.L.; Davidi, R.; Herman, G.T. On the effectiveness of projection methods for convex feasibility problems with linear inequality constraints. *Comput. Optim. Appl.* **2012**, *51*, 1065–1088. [CrossRef]
26. Rudin, L.I.; Osher, S.; Fatemi, E. Nonlinear total variation based noise removal algorithms. *Phys. D Nonlinear Phenom.* **1992**, *60*, 259–268. [CrossRef]
27. Michael, V. MRI Partial Fourier Reconstruction with POCS. Available online: https://fr.mathworks.com/matlabcentral/fileexchange/39350-mri-partial-fourier-reconstruction-with-pocs?s_tid=prof_contriblnk (accessed on 13 February 2020)
28. Condat, L. Discrete total variation: New definition and minimization. *SIAM J. Imaging Sci.* **2017**, *10*, 1258–1290. [CrossRef]
29. Plonka, G.; Potts, D.; Steidl, G.; Tasche, M. *Numerical Fourier Analysis: Theory and Applications*; Book Manuscript; Springer: Berlin/Heidelberg, Germany, 2018.
30. Al Sarray, B.; Chrétien, S.; Clarkson, P.; Cottez, G. Enhancing Prony’s method by nuclear norm penalization and extension to missing data. *Signal Image Video Process.* **2017**, *11*, 1089–1096. [CrossRef]

31. Barton, E.; Al-Sarray, B.; Chrétien, S.; Jagan, K. Decomposition of Dynamical Signals into Jumps, Oscillatory Patterns, and Possible Outliers. *Mathematics* **2018**, *6*, 124. [[CrossRef](#)]
32. Moitra, A. Super-resolution, extremal functions and the condition number of Vandermonde matrices. In Proceedings of the Forty-Seventh Annual ACM Symposium on Theory Of Computing, Chicago, IL, USA, 22–26 June 2015; pp. 821–830.
33. Chrétien, S.; Tyagi, H. Multi-kernel unmixing and super-resolution using the Modified Matrix Pencil method. *J. Fourier Anal. Appl.* **2020**, *26*, 18. [[CrossRef](#)]
34. Bach, F. On the Unreasonable Effectiveness of Richardson Extrapolation. Available online: <https://francisbach.com/richardson-extrapolation/> (accessed on 13 February 2020).
35. Dragomir, S.S. A generalisation of the Cassels and Greub-Reinboldt inequalities in inner product spaces. *arXiv* **2003**, arXiv:math/0306352.
36. Niculescu, C.P. *Converses of the Cauchy-Schwartz Inequality in the C*-Framework*; RGMIA Research Report Collection; 2001; Volume 4. Available online: <https://rgmia.org/v4n1.php> (accessed on 15 February 2020)



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).