General Projective Splitting Methods for Sums of Maximal Monotone Operators

Jonathan Eckstein*   B. F. Svaiter†

July 23, 2007

Abstract

We describe a general projective framework for finding a zero of the sum of n maximal monotone operators over a real Hilbert space. Unlike prior methods for this problem, we neither assume $n = 2$ nor first reduce the problem to the case $n = 2$. Our analysis defines a closed convex *extended solution set* for which we can construct a separating hyperplane by individually evaluating the resolvent of each operator. At the cost of a single, computationally simple projection step, this framework gives rise to a family of splitting methods of unprecedented flexibility: numerous parameters, including the proximal stepsize, may vary by iteration and by operator. The order of operator evaluation may vary by iteration, and may be either serial or parallel. The analysis essentially generalizes our prior results for the case $n = 2$. We also include a relative error criterion for approximately evaluating resolvents, which was not present in our earlier work.

1 Background and introduction

This paper considers the inclusion

$$0 \in T_1(x) + \cdots + T_n(x), \tag{1}$$

where $n \geq 2$ and $T_1, \cdots, T_n$, are set-valued maximal monotone operators on some real Hilbert space $\mathcal{H}$. Our interest is in *splitting* methods for this problem: iterative algorithms which may evaluate the individual operators T_i or (perhaps approximately) their resolvents $(I + \lambda T_i)^{-1}$, $\lambda > 0$, at various points in $\mathcal{H}$, but never resolvents of sums of the T_i. The idea is that (1) has been formulated so that each individual T_i has some relatively convenient structure, but

*Department of Management Science and Information Systems and RUTCOR, 640 Bartholomew Road, Busch Campus, Rutgers University, Piscataway NJ 08854 USA, jeckstei@rci.rutgers.edu. This research was supported in part by a Faculty Research Grant from Rutgers Business School — Newark and New Brunswick.

†IMPA, Instituto de Matemática Pura e Aplicada, Estrada Dona Castorina, 110. Rio de Janeiro, RJ, CEP 22460-320, Brazil, benar@impa.br. Partially supported by grants 300755/2005-8, 475647/2006-8 and by PRONEX-Optimization.

sums of two or more of the T_i might not. Thus, we seek iterative decomposition algorithms that evaluate only the "easy" resolvents $(I + \lambda T_i)^{-1}$, and not "difficult" compound resolvents such as $(I + \lambda(T_i + T_j))^{-1}$, $i \neq j$.

Algorithms of this form have been studied since the 1970's [14], although their roots in numerical methods for single-valued and in particular linear mappings are much older [5, 17]. In the extensive literature of these methods, the case $n = 2$ predominates. The most attractive convergence theory among $n = 2$ algorithms belongs to the Peaceman-Rachford and Douglas-Rachford class, which form a single related family. References to this class of methods in the context of set-valued monotone operators include [14, 12, 6, 7]. The method of partial inverses [22] is a special case of this approach.

Another family of $n = 2$ methods is the double-backward class; see for example [13, 16, 4, 2]. Such methods have attractive convergence properties, but for a variational inequality different from (1). Only if the proximal parameters λ in the resolvents are driven to zero in a particular way is this approach known to solve (1). It does not appear that double-backward algorithms are used in practice for (1).

Splitting methods of the forward-backward class, generalizing standard gradient projection methods for variational inequalities and optimization problems, are more popular than double-backward methods. References applying such methods to problems in the form (1) with $n = 2$ include [9, 23]. However, such methods must typically impose additional assumptions on at least one of the operators.

Traditionally, splitting algorithms allowing $n > 2$ have either explicitly or implicitly relied on reduction of (1) to the case $n = 2$ in the product space $\mathcal{H}^n$, endowed with with the canonical inner product $\langle (x_1, \ldots, x_n), (y_1, \ldots, y_n) \rangle = \sum_{i=1}^n \langle x_i, y_i \rangle$, as follows: define the closed subspace

$$W \stackrel{\text{def}}{=} \{(w_1, \ldots, w_n) \in \mathcal{H}^n \mid w_1 + w_2 + \cdots + w_n = 0\}, \tag{2}$$

whose orthogonal complement is

$$W^\perp = \{(v_1, \ldots, v_n) \in \mathcal{H}^n \mid v_1 = v_2 = \cdots = v_n\} = \{(v, v, \ldots, v) \mid v \in \mathcal{H}\}.$$

Next, define two operators $A, B : \mathcal{H}^n \rightrightarrows \mathcal{H}^n$ via $A \stackrel{\text{def}}{=} T_1 \otimes T_2 \otimes \cdots \otimes T_n$ and $B \stackrel{\text{def}}{=} N_{W^\perp}$, the normal cone map of $W^\perp$, that is,

$$A(x_1, \ldots, x_n) = T_1(x_1) \times T_2(x_2) \times \cdots \times T_n(x_n) \tag{3}$$

$$B(x_1, \ldots, x_n) = \begin{cases} W, & x_1 = x_2 = \cdots = x_n \\ \emptyset, & \text{otherwise.} \end{cases} \tag{4}$$

Using the maximal monotonicity of $T_1, \ldots, T_n$, it is straightforward to establish that A and B are maximal monotone on $\mathcal{H}^n$, and that

$$\begin{aligned} & 0 \in A(x_1, \ldots, x_n) + B(x_1, \ldots, x_n) \qquad (5) \\ \Leftrightarrow \quad & x_1 = x_2 = \ldots = x_n, \ \exists\, y_i \in T_i(x_i),\ i = 1, \ldots, n : y_1 + y_2 + \cdots + y_n = 0 \\ \Leftrightarrow \quad & x_1 = x_2 = \ldots = x_n \ \text{ solves (1).} \end{aligned}$$

Applying Douglas-Rachford splitting to (5) produces Spingarn's method [22, Section 5], in which one performs independent proximal steps on each of the operators $T_1, \ldots, T_n$, and then computes the next iterate by essentially averaging the results. In this setting, a proximal step on one operator cannot "feed" information into the proximal step for another operator within the same iteration. Applying a different $n = 2$ splitting method to (5) cannot alter this situation: evaluating the resolvent of A as defined in (3) will always yield independent, essentially simultaneous resolvent evaluations for $T_1, \ldots, T_n$.

Here, we propose to take a new, projective approach to splitting algorithms with $n \geq 2$, generalizing our prior work [8] for the case $n = 2$. We make use of a product space, but in a different manner: we define an *extended solution set* corresponding to (1) in $\mathcal{H}^{n+1}$, and use what is an essentially a generic projection algorithm to produce a sequence weakly convergent to a point in that set. The decomposition properties of the algorithm arise from the particular way in which we construct the separating hyperplanes used by this projection method. This approach allows for a generality and flexibility not present in prior splitting methods for (1).

The remainder of this paper is organized as follows: Section 2 defines the extended solution set, and analyzes some of its fundamental properties. To clarify the basic structure of our algorithm, we then introduce it in two stages: Section 3 first describes a generic, abstract family of projection methods for finding a point in the extended solution set, giving general convergence conditions. Section 4 then specializes this abstract family to a concrete family characterized by a large number of parameters, presenting conditions under which it conforms to Section 3's convergence conditions. Section 5 describes some variations and special cases of the algorithm of Section 4, in particular showing that it subsumes Spingarn's method [22]. Section 6 gives some conclusions and topics for future research, while two appendices prove some technical results needed for Sections 3 and 4.

2 The extended solution set and its separators

Consider now the Hilbert space $\mathcal{H} \times \mathcal{H}^n = \mathcal{H}^{n+1}$ under the canonical inner product

$$\langle (v, w_1, \ldots, w_n), (x, y_1, \ldots, y_n) \rangle = \langle v, x \rangle + \textstyle\sum_{i=1}^n \langle w_i, y_i \rangle,$$

and define the closed linear subspace

$$V \stackrel{\text{def}}{=} \mathcal{H} \times W = \left\{ (v, w_1, \ldots, w_n) \in \mathcal{H}^{n+1} \;\middle|\; w_1 + \ldots + w_n = 0 \right\}. \tag{6}$$

We define the *extended solution set* for problem (1) to be

$$S_{\mathrm{e}}(T_1, \ldots, T_n) \stackrel{\text{def}}{=} \{ (z, w_1, \ldots, w_n) \in V \mid w_i \in T_i(z),\ i = 1, \cdots, n \}. \tag{7}$$

For a point $(z, w_1, \ldots, w_n) \in \mathcal{H}^{n+1}$ to be in $S_{\mathrm{e}}(T_1, \ldots, T_n)$, it must satisfy two conditions: (z, w_i) must be in in the graph of T_i for all i, and $w_1 + \cdots + w_n = 0$, so that $(z, w_1, \ldots, w_n) \in V$. We now establish two basic properties of $S_{\mathrm{e}}(T_1, \ldots, T_n)$, the first of which is elementary:

Lemma 1 *Finding a point in $S_{\mathrm{e}}(T_1, \ldots, T_n)$ is equivalent to solving* (1) *in the sense that*

$$0 \in T_1(z) + \cdots + T_n(z) \iff \exists w_1, \ldots, w_n \in \mathcal{H} :\ (z, w_1, \cdots, w_n) \in S_{\mathrm{e}}(T_1, \ldots, T_n). \tag{8}$$

Proof. For any $z \in \mathcal{H}$,

$$\begin{aligned} & 0 \in T_1(z) + \cdots + T_n(z) \\ \iff & \exists w_1, \ldots, w_n \in \mathcal{H} : \textstyle\sum_{i=1}^n w_i = 0, \;\; w_i \in T_i(z), \; i = 1, \cdots, n \\ \iff & \exists w_1, \ldots, w_n \in \mathcal{H} : (z, w_1, \ldots, w_n) \in S_{\mathrm{e}}(T_1, \ldots, T_n). \end{aligned}$$

□

Lemma 2 *If the monotone operators $T_1, \ldots, T_n : \mathcal{H} \rightrightarrows \mathcal{H}$ are maximal, the corresponding extended solution set $S_{\mathrm{e}}(T_1, \ldots, T_n)$ is closed and convex in $\mathcal{H}^{n+1}$.*

Proof. Closedness of $S_{\mathrm{e}}(T_1, \ldots, T_n)$ follows immediately from (7), the closedness of the linear subspace V, and the closedness of the graphs of the maximal monotone operators $T_1, \ldots, T_n$. To prove convexity, take any scalars $p, q \geq 0$, $p + q = 1$, and any

$$(z, w_1, \ldots, w_n), (z', w_1', \ldots, w_n') \in S_{\mathrm{e}}(T_1, \ldots, T_n).$$

If we can establish that

$$p(z, w_1, \ldots, w_n) + q(z', w_1', \ldots, w_n') \in S_{\mathrm{e}}(T_1, \ldots, T_n),$$

the proof will be complete. Since $V \supset S_{\mathrm{e}}(T_1, \ldots, T_n)$ is a linear subspace, it is clear that $p(z, w_1, \ldots, w_n) + q(z', w_1', \ldots, w_n') \in V$. From (7), it thus remains only to show that for all $j = 1, \ldots, n$,

$$pw_j + qw_j' \in T_j(pz + qz'). \tag{9}$$

To this end, fix any $j \in \{1, \ldots, n\}$. By the monotonicity of the T_i, $\langle z - z', w_i - w_i' \rangle \geq 0$ for all $i = 1, \ldots, n$, and so

$$0 \leq \langle z - z', w_j - w_j' \rangle \leq \textstyle\sum_{i=1}^n \langle z - z', w_i - w_i' \rangle = \langle z - z', \sum_{i=1}^n w_i - \sum_{i=1}^n w_i' \rangle.$$

Since $\sum_1^n w_i = 0$ and $\sum_1^n w_i' = 0$, we conclude that $\langle z - z', w_j - w_j' \rangle = 0$. Now, consider an arbitrary $(\hat{z}, \hat{w}_j) \in \operatorname{Gph}(T_j)$, and observe that we therefore have

$$\begin{aligned} & \langle \hat{z} - (pz + qz'), \hat{w}_j - (pw_j + qw_j') \rangle \\ = \;& p\langle \hat{z} - z, \hat{w}_j - w_j \rangle + q\langle \hat{z} - z', \hat{w}_j - w_j' \rangle + pq\langle z - z', w_j' - w_j \rangle \\ = \;& p\langle \hat{z} - z, \hat{w}_j - w_j \rangle + q\langle \hat{z} - z', \hat{w}_j - w_j' \rangle. \end{aligned}$$

The monotonicity of T_j implies that $\langle \hat{z} - z, \hat{w}_j - w_j \rangle \geq 0$ and $\langle \hat{z} - z', \hat{w}_j - w_j' \rangle \geq 0$, so we conclude that

$$\langle \hat{z} - (pz + qz'), \hat{w}_j - (pw_j + qw_j') \rangle \geq 0.$$

Since $(\hat{z}, \hat{w}_j) \in \operatorname{Gph}(T_j)$ was arbitrary and T_j is maximal, it follows that (9) holds. □

Several variations on the definition of $S_{\mathrm{e}}(T_1, \ldots, T_n)$ are also possible. One possibility is to implicitly define w_n in terms of $w_1, \ldots, w_{n-1}$, obtaining

$$\left\{ (z, w_1, \ldots, w_{n-1}) \;\middle|\; w_i \in T_i(z), i = 1, \ldots, n-1, \; -(w_1 + \cdots + w_{n-1}) \in T_n(z) \right\}. \tag{10}$$

This variation, in the case $n = 2$, is used in our earlier work [8]. Another possible variation is to use the set

$$\left\{ (z_1, \ldots, z_n, w_1, \ldots, w_n) \;\middle|\; \begin{array}{l} z_1 = z_2 = \cdots = z_n \\ w_1 + w_2 + \cdots + w_n = 0 \\ w_i \in T_i(z_i),\ i = 1, \cdots, n \end{array} \right\}, \tag{11}$$

which is the intersection of the sets $\mathrm{Gph}(N_{W^\perp})$ and (after some permutation of indices) $\mathrm{Gph}(T_1) \times \mathrm{Gph}(T_2) \times \cdots \times \mathrm{Gph}(T_n)$ in $\mathcal{H}^{2n}$. Such variations should not lead to material differences in the resulting algorithms.

In view of Lemmas 1 and 2, we attempt to solve (1) by finding a point in $S_{\mathrm{e}}(T_1, \ldots, T_n)$, a problem we in turn approach by using a separator-projection algorithm. The separating hyperplanes used in our algorithm are constructed in a simple manner from points $(x_i, y_i) \in \mathrm{Gph}(T_i)$, $i = 1, \ldots, n$. The following lemma details the construction and properties of these separators:

Lemma 3 *Given $(x_i, y_i) \in \mathrm{Gph}(T_i)$, $i = 1, \cdots, n$, define $\varphi : V \to \mathbb{R}$ via*

$$\varphi(z, w_1, \ldots, w_n) = \textstyle\sum_{i=1}^n \langle z - x_i, y_i - w_i \rangle. \tag{12}$$

Then, for any $(z, w_1, \ldots, w_n) \in S_{\mathrm{e}}(T_1, \ldots, T_n)$, one has $\varphi(z, w_1, \ldots, w_n) \leq 0$, that is,

$$S_{\mathrm{e}}(T_1, \ldots, T_n) \subseteq \{(z, w_1, \ldots, w_n) \in V \mid \varphi(z, w_1, \ldots, w_n) \leq 0\}.$$

Additionally, φ is affine on V, with

$$\nabla\varphi = \left(\textstyle\sum_{i=1}^n y_i, x_1 - \bar{x}, x_2 - \bar{x}, \cdots, x_n - \bar{x}\right), \qquad \text{where} \qquad \bar{x} \stackrel{\mathrm{def}}{=} \tfrac{1}{n}\textstyle\sum_{i=1}^n x_i, \tag{13}$$

and

$$\begin{aligned} \nabla\varphi = 0 \qquad &\Longleftrightarrow \qquad (x_1, y_1, \ldots, y_n) \in S_{\mathrm{e}}(T_1, \ldots, T_n),\ x_1 = \cdots = x_n \\ &\Longleftrightarrow \qquad \varphi(z, w_1, \ldots, w_n) = 0,\ \ \forall (z, w_1, \ldots, w_n) \in V. \end{aligned}$$

Proof. Take any $(z, w_1, \ldots, w_n) \in S_{\mathrm{e}}(T_1, \ldots, T_n)$. For each $i = 1, \cdots, n$, we have $w_i \in T_i(z)$ and $y_i \in T(x_i)$. Since T_i is monotone, we have $\langle z - x_i, w_i - y_i \rangle \geq 0$. Negating and summing these inequalities, we conclude that $\varphi(z, w_1, \ldots, w_n) \leq 0$, proving the first claim.

Next, take $(z, w_1, \ldots, w_n)$ to be an arbitrary element of V. Expanding and regrouping the inner products in (12), we obtain

$$\begin{aligned} \varphi(z, w_1, \ldots, w_n) &= \langle z, \textstyle\sum_{i=1}^n y_i \rangle - \langle z, \textstyle\sum_{i=1}^n w_i \rangle - \textstyle\sum_{i=1}^n \langle x_i, y_i \rangle + \textstyle\sum_{i=1}^n \langle x_i, w_i \rangle & (14) \\ &= \langle z, \textstyle\sum_{i=1}^n y_i \rangle - \textstyle\sum_{i=1}^n \langle x_i, y_i \rangle + \textstyle\sum_{i=1}^n \langle x_i, w_i \rangle & (15) \\ &= \langle z, \textstyle\sum_{i=1}^n y_i \rangle - \textstyle\sum_{i=1}^n \langle x_i, y_i \rangle + \textstyle\sum_{i=1}^n \langle x_i - \bar{x}, w_i \rangle + \langle \bar{x}, \textstyle\sum_{i=1}^n w_i \rangle & (16) \\ &= \langle z, \textstyle\sum_{i=1}^n y_i \rangle + \textstyle\sum_{i=1}^n \langle x_i - \bar{x}, w_i \rangle - \textstyle\sum_{i=1}^n \langle x_i, y_i \rangle & (17) \\ &= \langle (z, w_1, \ldots, w_n), (\textstyle\sum_{i=1}^n y_i, x_1 - \bar{x}, \ldots, x_n - \bar{x}) \rangle - \textstyle\sum_{i=1}^n \langle x_i, y_i \rangle, & (18) \end{aligned}$$

where (15) and (17) follow from $\sum_{i=1}^n w_i = 0$, since $(z, w_1, \ldots, w_n) \in V$. Since $\sum_{i=1}^n (x_i - \bar{x}) = \sum_{i=1}^n x_i - \sum_{i=1}^n x_i = 0$, we have that $(\sum_{i=1}^n y_i, x_1 - \bar{x}, \ldots, x_n - \bar{x}) \in V$. Thus, (18) shows that φ is indeed an affine function on the space V, and $\nabla\varphi = (\sum_{i=1}^n y_i, x_1 - \bar{x}, \ldots, x_n - \bar{x})$.

Finally, we note that $\nabla\varphi = 0$ if and only if $\sum_{i=1}^n y_i = 0$ and $x_1 = \cdots = x_n = \bar{x}$. In that case, since $y_i \in T_i(x_i)$, $i = 1,\ldots,n$, one also has $(x_1, y_1, \ldots, y_n) = (\bar{x}, y_1, \ldots, y_n) \in S_{\mathrm{e}}(T_1,\ldots,T_n)$. In this case, we have $\sum_{i=1}^n \langle x_i, y_i\rangle = \langle x_1, \sum_{i=1}^n y_i\rangle = \langle x_1, 0\rangle = 0$, and we conclude from (18) that φ is the zero function. $\square$

Note that φ is not an affine function on the space $\mathcal{H}^{n+1}$, but only on its subspace V, where the "cross term" $\langle z, \sum_{i=1}^n w_i\rangle$ in (14) must be zero. We will thus implement our algorithm within the subspace V.

Next, it is natural to ask, given a point $(z, w_1, \ldots, w_n)$ in $V \backslash S_{\mathrm{e}}(T_1,\ldots,T_n)$, how to choose the pairs $(x_i, y_i) \in \mathrm{Gph}(T_i)$ so that φ separates $(z, w_1, \ldots, w_n)$ from $S_{\mathrm{e}}(T_1,\ldots,T_n)$, that is, $\varphi(z, w_1, \ldots, w_n) > 0$. In fact, such a choice may be accomplished by a "prox" operation on each of the operators $T_1,\ldots,T_n$. By the maximal monotonicity of the T_i and the classic results of [15], there exists for each $i = 1,\ldots,n$ a unique $(x_i, y_i) \in \mathrm{Gph}(T_i)$ such that $x_i + y_i = z + w_i$. Rearranging this equation, we obtain, $z - x_i = y_i - w_i$, and thus that $\varphi(z, w_1, \ldots, w_n) = \sum_{i=1}^n \|z - x_i\|^2$. Thus, $\varphi(z, w_1, \ldots, w_n) > 0$ unless $x_1 = \cdots = x_n = z$, in which case it is easily deduced that $y_i = w_i$ for all i, and therefore $(z, w_1, \ldots, w_n) \in S_{\mathrm{e}}(T_1,\ldots,T_n)$, contrary to the assumption. Finding the necessary $(x_i, y_i) \in \mathrm{Gph}(T_i)$ is equivalent to evaluating the resolvent $(I + T_i)^{-1}$, which is by assumption tractable for each individual T_i. We will greatly generalize this procedure for determining a separator in Section 4 below.

3 An abstract family of projection algorithms

We now have the necessary ingredients for implementing a projection method: a closed convex set S and at least one tractable procedure for calculating a separator between S and any $p \notin S$. Therefore, we may apply the following algorithmic template:

Algorithm 1 *Suppose $S \neq \emptyset$ is a closed convex set in a real Hilbert space U. Start with an arbitrary $p^0 \in U$. Then, for $k = 0, 1, \ldots,$ repeat:*

1. *Determine a non-constant differentiable affine function $\varphi_k : U \to \mathbb{R}$ such that $\varphi_k(p) \leq 0$ for all $p \in S$.*
2. *Let $\overline{p}^k$ be the projection of p^k onto the halfspace $H_k \overset{\text{def}}{=} \{p \in U \mid \varphi_k(p) \leq 0\}$, that is,*

$$\overline{p}^k = p^k - \frac{\max\{0, \varphi_k(p^k)\}}{\|\nabla\varphi_k\|^2}\nabla\varphi_k. \tag{19}$$

3. *Choose some relaxation parameter $\rho_k \in (0, 2)$, and set*

$$p^{k+1} = p^k + \rho_k(\overline{p}^k - p^k).$$

The last two steps may simply be condensed to

$$p^{k+1} = p^k - \rho_k \frac{\max\{0, \varphi_k(p^k)\}}{\|\nabla\varphi_k\|^2}\nabla\varphi_k. \tag{20}$$

The basic properties of this algorithmic form may be derived by following the analysis of classical projection algorithms, dating back to Cimmino [3] and Kaczmarz [10, 11] in the late 1930's. A comprehensive review of projection algorithms may be found in [1]. As in any (relaxed) projection method, the sequences generated by Algorithm 1 behave as follows: for any $p^* \in S$, then $p^* \in H_k$ and the firm nonexpansiveness property of the projection mapping onto H_k assures that for all $k \geq 0$,

$$\left\|p^* - \overline{p}^k\right\|^2 \leq \left\|p^* - p^k\right\|^2 - \left\|\overline{p}^k - p^k\right\|^2 \tag{21}$$

$$\left\|p^* - p^{k+1}\right\|^2 \leq \left\|p^* - p^k\right\|^2 - \rho_k(2-\rho_k)\left\|p^{k+1} - p^k\right\|^2. \tag{22}$$

The basic behavior of this class of methods is as follows; we omit the proof, which is entirely standard.

Proposition 4 *Any infinite sequence $\{p^k\}$ generated by Algorithm 1 behaves as follows:*

1. *For any $p^* \in S$, the sequence $\{\|p^k - p^*\|\}$ is nonincreasing — that is, $\{p^k\}$ is Fejér monotone to S.*
2. *If $p^{k_0} \in S$ for some $k_0 \geq 0$ then $p^k = p^{k_0}$ for all $k \geq k_0$.*
3. *If $\{p^k\}$ has a strong accumulation point in S, then the whole sequence converges to that point.*
4. *If S is non-empty, then $\{p^k\}$ is bounded. Moreover, if there exist $\underline{\rho}, \overline{\rho}$ such that $0 < \underline{\rho} \leq \rho_k \leq \overline{\rho} < 2$ for all k, then*

$$\sum_{k=0}^{\infty}\left\|p^k - \overline{p}^k\right\|^2 < \infty \qquad\qquad \sum_{k=0}^{\infty}\left\|p^k - p^{k+1}\right\|^2 < \infty. \tag{23}$$

5. *The sequence $\{p^k\}$ has at most one weak accumulation point in S.*

Note, however, that the basic template of Algorithm 1 is not sufficient to ensure weak convergence of $\{p^k\}$ to a point in S, because the separators φ_k might not be chosen to actually separate p^k from S, or might separate in a pathologically "shallow" way. The analysis of [8] guarantees convergence using the condition $\varphi_k(p^k) \geq \xi\left\|\nabla\varphi_k\right\|^2$ for all $k \geq 0$, where $\xi > 0$ is a fixed constant. We will also use this condition below.

We now restate and specialize Algorithm 1 for the case $U = V$ and $S = S_{\mathrm{e}}(T_1, \ldots, T_n)$, with the separators constructed as in Lemma 3. We do not for the moment assume any particular way of choosing the $(x_i, y_i) \in \mathrm{Gph}(T_i)$ yielding the separator.

Algorithm 2 *Start with an arbitrary $p^0 = (z^0, w_1^0, \ldots, w_n^0) \in V$. Then, for $k = 0, 1, \ldots$, repeat:*

1. *For $i = 1, \ldots, n$, choose some $(x_i^k, y_i^k) \in \mathrm{Gph}(T_i)$.*
2. *If $x_1^k = x_2^k = \cdots = x_n^k$ and $\sum_{i=1}^n y_i^k = 0$, let $w_i^{k+1} = y_i^k$ for $i = 1, \ldots, n$ and $z^{k+1} = x_1^k$. Otherwise, continue:*

3. *Define* $\varphi_k : V \to \mathbb{R}$ *to be the separator derived from* (x_i^k, y_i^k) *via* (12)*, that is,*

$$\varphi_k(z, w_1, \ldots, w_n) \stackrel{\text{def}}{=} \textstyle\sum_{i=1}^n \langle z - x_i^k, y_i^k - w_i \rangle,$$

and let $p^{k+1} = (z^{k+1}, w_1^{k+1}, \ldots, w_n^{k+1})$ *be the projection of* p^k *onto the halfspace* $H_k \stackrel{\text{def}}{=} \{p \in V \mid \varphi_k(p) \leq 0\}$*, with an overrelaxation factor* $\rho_k \in (0, 2)$*, that is,*

$$\bar{x}^k = \tfrac{1}{n} \textstyle\sum_{i=1}^n x_i^k \tag{24}$$

$$\theta_k = \frac{\max\{0, \sum_{i=1}^n \langle z^k - x_i^k, y_i^k - w_i^k \rangle\}}{\left\|\sum_{i=1}^n y_i^k\right\|^2 + \sum_{i=1}^n \left\|x_i^k - \bar{x}^k\right\|^2} \tag{25}$$

$$z^{k+1} = z^k - \rho_k \theta_k \textstyle\sum_{i=1}^n y_i^k \tag{26}$$

$$w_i^{k+1} = w_i^k - \rho_k \theta_k (x_i^k - \bar{x}^k) \qquad i = 1, \ldots, n. \tag{27}$$

Note that the test in step 2 guarantees that the denominator in (25) cannot be zero. Lemma 3's formula for the gradient of φ_k implies that (24)-(27) indeed calculate the overrelaxed projection of p^k onto H_k, and Algorithm 2 is thus Algorithm 1 specialized to $U = V$ and $S = S_{\mathrm{e}}(T_1, \ldots, T_n)$. Note also that $p^k \in V$ and the update (27) ensure $w_1^{k+1} + \cdots + w_n^{k+1} = 0$, so by induction the entire iterate sequence $\{p^k\} = \{(z^k, w_1^k, \ldots, w_n^k)\}$ produced by Algorithm 2 lies in V.

We now perform a preliminary analysis of the convergence properties of Algorithm 2:

Proposition 5 *Suppose that the following conditions are met in Algorithm 2:*

1. $S_{\mathrm{e}}(T_1, \ldots, T_n) \neq \emptyset$.
2. $0 < \underline{\rho} \leq \rho_k \leq \overline{\rho} < 2$ *for all* k.
3. *There exists some scalar* $\xi > 0$ *such that*

$$\varphi(p^k) = \varphi_k(z^k, w_1^k, \ldots, w_n^k) \geq \xi \left\|\nabla \varphi_k\right\|^2 = \xi \left(\left\|\textstyle\sum_{i=1}^n y_i^k\right\|^2 + \textstyle\sum_{i=1}^n \left\|x_i^k - \bar{x}^k\right\|^2 \right) \tag{28}$$

for all $k \geq 0$.

Then $\nabla \varphi_k \to 0$*, that is,* $x_i^k - x_j^k \to 0$ *for all* $i, j = 1, \ldots, n$*, and* $\sum_{i=1}^n y_i^k \to 0$*. Furthermore,* $\varphi_k(p^k) \to 0$*. If it is also true that*

4. *Either* $\mathcal{H}$ *has finite dimension or the operator* $\sum_{i=1}^n T_i$ *is maximal.*
5. $z^k - \bar{x}^k \to 0$.
6. $w_i^k - y_i^k \to 0$*, for* $i = 1, \ldots, n$*,*

then $\{(z^k, w_1^k, \ldots, w_n^k)\}$ *converges weakly to some* $p^\infty = (z^\infty, w_1^\infty, \ldots, w_n^\infty) \in S_{\mathrm{e}}(T_1, \ldots, T_n)$*, which implies that* z^∞ *solves* (1)*. Furthermore,* $x_i^k \xrightarrow{\mathrm{w}} z^\infty$ *and* $y_i^k \xrightarrow{\mathrm{w}} w_i^\infty$ *for* $i = 1, \ldots, n$.

Proof. The hypothesis that $\varphi_k(p^k) \geq \xi \|\nabla\varphi_k\|^2$ implies that $\varphi_k(p^k)$ is always nonnegative, so we obtain from (19) that

$$\left\|p^k - \overline{p}^k\right\| = \frac{\varphi_k(p^k)}{\|\nabla\varphi_k\|} \tag{29}$$

for all k having $\nabla\varphi_k \neq 0$. Substituting $\varphi_k(p^k) \geq \xi \|\nabla\varphi_k\|^2$ into this equation, we obtain

$$\left\|p^k - \overline{p}^k\right\| \geq \xi \|\nabla\varphi_k\|, \tag{30}$$

which clearly also holds for k having $\nabla\varphi_k = 0$. From (23), which holds by hypothesis 2 and Proposition 4(4), we have $\left\|p^k - \overline{p}^k\right\| \to 0$, so (30) implies $\nabla\varphi_k \to 0$. From the expression for $\nabla\varphi_k$ in (13), we immediately conclude $\sum_{i=1}^n y_i^k \to 0$ and $x_i^k - \bar{x}^k \to 0$ for $i = 1, \ldots, n$, and thus $x_i^k - x_j^k \to 0$ for all $i, j = 1, \ldots, n$. Multiplying (29) by $\|\nabla\varphi_k\|$, we obtain that

$$\varphi_k(p^k) = \left\|p^k - \overline{p}^k\right\| \|\nabla\varphi_k\| \tag{31}$$

whenever $\nabla\varphi_k \neq 0$. By Lemma 3, $\varphi_k(p^k) = 0$ whenever $\nabla\varphi_k = 0$, so (31) holds for all $k \geq 0$. Since we have established that $\nabla\varphi_k \to 0$, and we also know that $\left\|p^k - \overline{p}^k\right\| \to 0$, (31) implies that $\varphi_k(p^k) \to 0$. The first set of conclusions are now established; note also that by hypothesis 1 and Fejér monotonicity, the sequence $\{p^k\}$ is bounded.

To prove the remainder of the proposition, we now assume that hyphotheses 4-6 also hold. From hypothesis 5 and $x_i^k - \bar{x}^k \to 0$, we immediately obtain

$$z^k - x_i^k \to 0 \qquad i = 1, \ldots, n. \tag{32}$$

In hypothesis 4, suppose first that $\mathcal{H}$ is finite-dimensional. Let $p^\infty = (z^\infty, w_1^\infty, \ldots, w_n^\infty)$ be any cluster point of the bounded sequence $\{p^k\}$. There then exists a subsequence $\{p^k\}_{k\in\mathcal{K}}$ converging to p^∞. From (32), we must then have $x_i^k \to_{k\in\mathcal{K}} z^\infty$, $i = 1, \ldots, n$. Similarly, hypothesis 6 implies $y_i^k \to_{k\in\mathcal{K}} w_i^\infty$, $i = 1, \ldots, n$. Since $(x_i^k, y_i^k) \in \operatorname{Gph}(T_i)$ for all i and k, and the maximality of the operators T_i imply that the sets $\operatorname{Gph}(T_i)$ are closed, we then obtain $w_i^\infty \in T_i(z^\infty)$ for all $i = 1, \ldots, n$. Furthermore, since $\{p^k\} \subset V$ and V is a closed subspace, we also have $p^\infty \in V$ and thus $p^\infty \in S_{\mathrm{e}}(T_1, \ldots, T_n)$. Finally, we apply Proposition 4(3) to obtain that the entire sequence $\{p^k\}$ converges to $p^\infty \in S_{\mathrm{e}}(T_1, \ldots, T_n)$.

Now assume the other alternative in hypothesis 4, that $\sum_{i=1}^n T_i$ is maximal monotone. Let p^∞ be any weak cluster point of $\{p^k\}$. Then there exists a subsequence $\{p^k\}_{k\in\mathcal{K}}$ weakly convergent to p^∞, and, using hypotheses 5 and 6, we conclude that $(x_i^k, y_i^k) \xrightarrow{\mathsf{w}}_{k\in\mathcal{K}} (z^\infty, w_i^\infty)$, $i = 1, \ldots, n$. Next, we apply Proposition 8 from Appendix A to conclude that $p^\infty = (z^\infty, w_1^\infty, \ldots, w_n^\infty) \in S_{\mathrm{e}}(T_1, \ldots, T_n)$. Since p^∞ was chosen arbitrarily, all weak cluster points of $\{p^k\}$ are in $S_{\mathrm{e}}(T_1, \ldots, T_n)$. Then we may apply Proposition 4(5) to conclude that the entire the sequence $\{p^k\}$ converges weakly to p^∞.

In either case, the remaining conclusions follow from hypothesis 6 and (32). □

4 A general projective splitting scheme

To convert Algorithm 2 into an implementable procedure for solving (1), we must specify a way of choosing the $(x_i^k, y_i^k) \in \operatorname{Gph}(T_i)$ so that the hypotheses of Proposition 5 are satisfied.

One simple approach, as mentioned at the end of Section 2, would be to choose the unique $(x_i^k, y_i^k) \in \mathrm{Gph}(T_i)$ satisfying $x_i^k + y_i^k = z^k + w_i^k$. A simple generalization would be to add a proximal parameter $\lambda_i^k > 0$, yielding

$$x_i^k + \lambda_i^k y_i^k = z^k + \lambda_i^k w_i^k. \tag{33}$$

This scheme may in fact be greatly generalized without sacrificing its basic decomposability. Suppose for the moment that in each iteration we perform the proximal calculations for the T_i sequentially, starting with $i = 1$ and finishing with $i = n$. We may then wish to use the "recent" information generated in calculating (x_j^k, y_j^k), where $j < i$, when calculating (x_i^k, y_i^k). Specifically, when calculating (x_i^k, y_i^k), we consider replacing z^k with an affine combination of z^k and the x_j^k, $j < i$. In particular, we first find the unique $(x_1^k, y_1^k) \in \mathrm{Gph}(T_1)$ such that

$$x_1^k + \lambda_1^k y_1^k = z^k + \lambda_1^k w_1^k.$$

We next take some $\alpha_{21}^k \in \mathbb{R}$ and find the unique $(x_2^k, y_2^k) \in \mathrm{Gph}(T_2)$ such that

$$x_2^k + \lambda_2^k y_2^k = (1 - \alpha_{21}^k) z^k + \alpha_{21}^k x_1^k + \lambda_2^k w_2^k.$$

To continue, we choose some $\alpha_{31}^k, \alpha_{32}^k \in \mathbb{R}$ and find the unique $(x_3^k, y_3^k) \in \mathrm{Gph}(T_3)$ such that

$$x_3^k + \lambda_3^k y_3^k = (1 - \alpha_{31}^k - \alpha_{32}^k) z^k + \alpha_{31}^k x_1^k + \alpha_{32}^k x_2^k + \lambda_3^k w_3^k,$$

and so forth. In general, we choose (x_i^k, y_i^k) to satisfy the conditions

$$x_i^k + \lambda_i^k y_i^k = \left(1 - \textstyle\sum_{j=1}^{i-1} \alpha_{ij}^k\right) z^k + \sum_{j=1}^{i-1} \alpha_{ij}^k x_j^k + \lambda_i^k w_i^k \qquad y_i^k \in T_i(x_i^k). \tag{34}$$

In addition to the flexibility afforded by the choice of the α_{ij}^k and λ_i^k, we consider several further generalizations of (34):

- We will allow errors $e_i^k \in \mathcal{H}$ in satisfying (34), so long as they satisfy the approximation criterion (42) below.
- The order of processing the operators may vary from iteration to iteration. At iteration k, we process the operators in the order specified by an arbitrary permutation $\pi^k(\cdot)$ of $\{1, \ldots, n\}$.

Thus, we arrive at the general scheme that for all $i = 1, \ldots, n$ and $k \geq 0$, we have $y_i^k \in T_i(x_i^k)$ and

$$x_{\pi_k(i)}^k + \lambda_i^k y_{\pi_k(i)}^k = \left(1 - \textstyle\sum_{j=1}^{i-1} \alpha_{ij}^k\right) z^k + \sum_{j=1}^{i-1} \alpha_{ij}^k x_{\pi_k(j)}^k + \lambda_i^k w_{\pi_k(i)}^k + e_i^k. \tag{35}$$

Note that the notion of processing the operators in some particular order $\pi_k(\cdot)$ does not necessarily preclude parallelism over i in evaluating (35), depending on how one chooses the α_{ij}^k. For example, if we choose $\alpha_{ij}^k = 0$ for all i, j and set the error terms $e_i^k = 0$, then (35) reduces to (33), which may be calculated independently and in parallel over i.

To analyze this scheme, we will employ some standard matrix analysis: given an $n \times n$ real matrix $\mathbf{L}$, we define $\|\mathbf{L}\|$ to be its operator 2-norm and $\kappa(\mathbf{L})$ to be the smallest eigenvalue of its symmetric part, that is

$$\|\mathbf{L}\| = \max_{\substack{\mathbf{x} \in \mathbb{R}^n \\ \|\mathbf{x}\| = 1}} \|\mathbf{L}\mathbf{x}\| \qquad \operatorname{sym} \mathbf{L} \overset{\text{def}}{=} \tfrac{1}{2}(\mathbf{L} + \mathbf{L}^\top) \qquad \kappa(\mathbf{L}) \overset{\text{def}}{=} \min \operatorname{eig} \operatorname{sym} \mathbf{L}.$$

Note that it is straightforward to show that $\kappa(\mathbf{L}) \leq \|\mathbf{L}\|$, and that for any $\mathbf{x} \in \mathbb{R}^n$, $\langle \mathbf{x}, \mathbf{L}\mathbf{x} \rangle \geq \kappa(\mathbf{L}) \|\mathbf{x}\|^2$. Analogously to the usual linear map $\mathbb{R}^n \to \mathbb{R}^n$ associated with $\mathbf{L}$, we can define a linear mapping $\mathcal{H}^n \to \mathcal{H}^n$ corresponding to $\mathbf{L}$ via

$$\mathbf{L}u = \mathbf{L}(u_1, \ldots, u_n) = (v_1, \ldots, v_n), \quad \text{where} \quad v_i = \sum_{j=1}^{n} \ell_{ij} u_j \in \mathcal{H}, \tag{36}$$

with ℓ_{ij} denoting the elements of $\mathbf{L}$. As one would intuitively expect, this mapping retains key spectral properties that $\mathbf{L}$ exhibits over $\mathbb{R}^n$:

Lemma 6 *Let $\mathbf{L}$ be any $n \times n$ real matrix. For all $u = (u_1, \ldots, u_n) \in \mathcal{H}^n$,*

$$\|\mathbf{L}u\| \leq \|\mathbf{L}\| \, \|u\| \tag{37}$$

$$\langle u, \mathbf{L}u \rangle \geq \kappa(\mathbf{L}) \, \|u\|^2, \tag{38}$$

where $\mathbf{L}u$ is defined by (36), $\langle \cdot, \cdot \rangle$ denotes the canonical inner product for $\mathcal{H}^n$ induced by the inner product for $\mathcal{H}$, and $\|\cdot\|$ applied to elements of $\mathcal{H}^n$ denotes the norm induced by this inner product.

Appendix B proves this result. Of particular interest are the matrices

$$\mathbf{\Lambda}_k \overset{\text{def}}{=} \operatorname{diag}(\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k) \tag{39}$$

$$\mathbf{A}_k \overset{\text{def}}{=} \begin{bmatrix} 1 & & & & \\ -\alpha_{21} & 1 & & & \\ -\alpha_{31} & -\alpha_{32} & 1 & & \\ \vdots & \vdots & & \ddots & \\ -\alpha_{n1} & -\alpha_{n2} & \cdots & -\alpha_{n,n-1} & 1 \end{bmatrix}, \tag{40}$$

that is, $\mathbf{A}_k = [a_{ij}^{(k)}]_{i,j=1,\ldots,n}$, where

$$a_{ij}^{(k)} = \begin{cases} 1, & i = j \\ -\alpha_{ij}^k, & i > j \\ 0, & i < j. \end{cases}$$

We will show that if there exist $\beta, \zeta > 0$ such that

$$\kappa(\mathbf{\Lambda}_k^{-1}\mathbf{A}_k) \geq \zeta \qquad \left\|\mathbf{\Lambda}_k^{-1}\mathbf{A}_k\right\| \leq \beta \qquad \forall\, k \geq 0, \tag{41}$$

then choosing the $(x_i^k, y_i^k) \in \operatorname{Gph}(T_i)$ via (35) will meet all the hypotheses of Proposition 5, and we will obtain weak convergence. Stated in full, including the approximate calculation criterion, the algorithm is as follows:

Algorithm 3 *Choose scalars* $\beta, \zeta > 0$, $0 < \underline{\rho} \leq \overline{\rho} < 2$, *and* $\sigma \in [0, 1)$. *Start with an arbitrary* $p^0 = (z^0, w_1^0, \ldots, w_n^0) \in V$, *that is, any* $z^0, w_1^0, \ldots, w_n^0 \in \mathcal{H}$ *with* $w_1^0 + \cdots + w_n^0 = 0$. *Then, for* $k = 0, 1, \ldots$, *repeat:*

1. *Choose scalars* $\lambda_i^k > 0$, $i = 1, \ldots, n$, *and* α_{ij}^k, $1 \leq j < i \leq n$, *such that* $\kappa(\mathbf{\Lambda}_k^{-1}\mathbf{A}_k) \geq \zeta$ *and* $\left\|\mathbf{\Lambda}_k^{-1}\mathbf{A}_k\right\| \leq \beta$, *where* $\mathbf{\Lambda}_k$ *and* $\mathbf{A}_k$ *are defined by* (39) *and* (40), *respectively. Let* $\pi_k(\cdot)$ *be any permutation of* $\{1, \ldots, n\}$. *For* $i = 1, \cdots, n$, *find* $(x_i^k, y_i^k) \in \mathrm{Gph}(T_i)$ *satisfying* (35),
$$x_{\pi_k(i)}^k + \lambda_i^k y_{\pi_k(i)}^k = \left(1 - \textstyle\sum_{j=1}^{i-1} \alpha_{ij}^k\right) z^k + \sum_{j=1}^{i-1} \alpha_{ij}^k x_{\pi_k(j)}^k + \lambda_i^k w_{\pi_k(i)}^k + e_i^k,$$
where
$$\sum_{i=1}^{n} \left(\lambda_i^k\right)^{-2} \left\|e_i^k\right\|^2 \leq \sigma^2 \kappa(\mathbf{\Lambda}_k^{-1}\mathbf{A}_k)^2 \sum_{i=1}^{n} \left\|x_i^k - z^k\right\|^2. \tag{42}$$

2. *If* $x_1^k = x_2^k = \cdots = x_n^k$ *and* $\sum_{i=1}^n y_i^k = 0$, *let* $w_i^{k+1} = y_i^k$ *for* $i = 1, \ldots, n$, *and* $z^{k+1} = x_1^k$. *Otherwise, continue:*

3. *Choose some* $\rho_k \in [\underline{\rho}, \overline{\rho}]$ *and set*
$$\bar{x}^k = \tfrac{1}{n} \textstyle\sum_{i=1}^n x_i^k \tag{43}$$
$$\theta_k = \frac{\sum_{i=1}^n \langle z^k - x_i^k, y_i^k - w_i^k \rangle}{\left\|\sum_{i=1}^n y_i^k\right\|^2 + \sum_{i=1}^n \left\|x_i^k - \bar{x}^k\right\|^2} \tag{44}$$
$$z^{k+1} = z^k - \rho_k \theta_k \textstyle\sum_{i=1}^n y_i^k \tag{45}$$
$$w_i^{k+1} = w_i^k - \rho_k \theta_k (x_i^k - \bar{x}^k) \qquad i = 1, \ldots, n. \tag{46}$$

The error condition (42) is an n-operator generalization of the relative error tolerance proposed in [19, 18, 21] for modified proximal-extragradient projection methods. Note that β's only role in the statement of the algorithm is to guarantee $\|\mathbf{\Lambda}_k^{-1}\mathbf{A}_k\|$ remains bounded, that is, that $\{\mathbf{\Lambda}_k^{-1}\mathbf{A}_k\}$ is a bounded sequence of matrices. Such boundedness may be assured by any sufficient condition bounding the absolute values $|a_{ij}^{(k)}/\lambda_i^k|$ of all entries of $\{\mathbf{\Lambda}_k^{-1}\mathbf{A}_k\}$. For example, if there exist $\underline{\lambda}, \overline{\alpha} \geq 0$ such that $\lambda_i^k \geq \underline{\lambda}$ and $|\alpha_{ij}^k| \leq \overline{\alpha}$ for all $k \geq 1$, $i = 1, \ldots, n$, and $1 \leq j < i$, then $\{\mathbf{\Lambda}_k^{-1}\mathbf{A}_k\}$ must be bounded, and some β satisfying the condition $\|\mathbf{\Lambda}_k^{-1}\mathbf{A}_k\| \leq \beta$ for all $k \geq 0$ must exist. In practice, we may therefore substitute conditions such as $\lambda_i^k \geq \underline{\lambda}$, $|\alpha_{ij}^k| \leq \overline{\alpha}$ for the condition $\|\mathbf{\Lambda}_k^{-1}\mathbf{A}_k\| \leq \beta$ in step 1 of the algorithm. We now prove convergence of the method:

Proposition 7 *Suppose that either* $\mathcal{H}$ *has finite dimension or the operator* $T_1 + \cdots + T_n$ *is maximal. Suppose also that* (1) *has a solution. Then, in Algorithm 3, the sequences* $\{z^k\}, \{x_1^k\}, \ldots, \{x_n^k\} \subset \mathcal{H}$ *all weakly converge to some* z^∞ *solving* (1). *For each* $i = 1, \ldots, n$, *we also have* $w_i^k, y_i^k \xrightarrow{\mathsf{w}} y_i^\infty$, *where* $y_i^\infty \in T_i(z^\infty)$ *and* $y_1^\infty + \cdots + y_n^\infty = 0$.

Proof. Define auxiliary sequences $\{u^k\} = \{(u_1^k, \dots, u_n^k)\} \subset \mathcal{H}^n$ and $\{v^k\} = \{(v_1^k, \dots, v_n^k)\} \subset \mathcal{H}^n$ via

$$u_i^k \overset{\text{def}}{=} x_i^k - z^k \qquad\qquad v_i^k \overset{\text{def}}{=} w_i^k - y_i^k \tag{47}$$

for all $i = 1, \dots, n$ and $k \geq 0$, and note (defining p^k as in Algorithms 1 and 2) that

$$\phi_k(p^k) = \langle u^k, v^k \rangle = \textstyle\sum_{i=1}^n \langle u_i^k, v_i^k \rangle. \tag{48}$$

Further, define $e^k = (e_1^k, \dots, e_n^k) \in \mathcal{H}^n$ for all $k \geq 0$, and observe that by taking square roots and substituting the definitions of e^k and u^k, (42) simplifies via the notation (36) and (39) to

$$\left\| \mathbf{\Lambda}_k^{-1} e^k \right\| \leq \sigma \kappa(\mathbf{\Lambda}_k^{-1} \mathbf{A}_k) \left\| u^k \right\|. \tag{49}$$

Take any $i \in \{1, \dots, n\}$. Subtracting z^k from both sides of (35) and regrouping yields

$$\begin{aligned}
& (x_{\pi_k(i)}^k - z^k) + \lambda_i^k y_{\pi_k(i)}^k = \sum_{j=1}^{i-1} \alpha_{ij}^k (x_{\pi_k(j)}^k - z^k) + \lambda_i^k w_{\pi_k(i)}^k + e_i^k \\
\Leftrightarrow \qquad & (x_{\pi_k(i)}^k - z^k) - \sum_{j=1}^{i-1} \alpha_{ij}^k (x_{\pi_k(j)}^k - z^k) - e_i^k = \lambda_i^k (w_{\pi_k(i)}^k - y_{\pi_k(i)}^k).
\end{aligned}$$

Dividing by λ_i^k and substituting the definitions of u_i^k and v_i^k yields

$$\left(\frac{1}{\lambda_i^k} \right) \left(u_{\pi_k(i)} - \sum_{j=1}^{i-1} \alpha_{ij}^k u_{\pi_k(j)} - e_i^k \right) = v_{\pi_k(i)}. \tag{50}$$

Applying the notation (36) to (50) for $i = 1, \dots, n$ produces

$$v^k = \left(\mathbf{\Pi}_k \mathbf{\Lambda}_k^{-1} \mathbf{A}_k \mathbf{\Pi}_k^\top \right) u^k - \left(\mathbf{\Pi}_k \mathbf{\Lambda}_k^{-1} \right) e^k, \tag{51}$$

where $\mathbf{\Pi}_k$ is the $n \times n$ permutation matrix corresponding to the permutation $\pi_k(\cdot)$. Substituting (51) into (48) and using the Cauchy-Schwarz inequality yields

$$\begin{aligned}
\varphi_k(p^k) &= \langle u^k, \mathbf{\Pi}_k \mathbf{\Lambda}_k^{-1} \mathbf{A}_k \mathbf{\Pi}_k^\top u^k \rangle - \langle u^k, \mathbf{\Pi}_k \mathbf{\Lambda}_k^{-1} e^k \rangle \\
&\geq \kappa\big(\mathbf{\Pi}_k \mathbf{\Lambda}_k^{-1} \mathbf{A}_k \mathbf{\Pi}_k^\top\big) \left\| u^k \right\|^2 - \left\| u^k \right\| \left\| \mathbf{\Pi}_k \mathbf{\Lambda}_k^{-1} e^k \right\| && [\text{using (38)}] \\
&= \kappa\big(\mathbf{\Lambda}_k^{-1} \mathbf{A}_k\big) \left\| u^k \right\|^2 - \left\| u^k \right\| \left\| \mathbf{\Lambda}_k^{-1} e^k \right\| \\
&\geq \kappa\big(\mathbf{\Lambda}_k^{-1} \mathbf{A}_k\big) \left\| u^k \right\|^2 - \sigma \kappa(\mathbf{\Lambda}_k^{-1} \mathbf{A}_k) \left\| u^k \right\|^2 && [\text{using (49)}] \\
&= (1 - \sigma) \kappa(\mathbf{\Lambda}_k^{-1} \mathbf{A}_k) \left\| u^k \right\|^2 \\
&\geq (1 - \sigma) \zeta \left\| u^k \right\|^2. && [\text{using (41)}]
\end{aligned} \tag{52}$$

To meet hypothesis 3 of Proposition 5, we need to convert this lower bound on $\varphi_k(p^k)$, expressed in terms of $\|u^k\|^2$, to one expressed in terms of $\|\nabla \varphi_k\|^2$. To do so, first note that since $\sum_{i=1}^n w_i^k = 0$,

$$\textstyle\sum_{i=1}^n v_i^k = \sum_{i=1}^n (w_i^k - y_i^k) = -\sum_{i=1}^n y_i^k \quad \Rightarrow \quad \left\| \sum_{i=1}^n v_i^k \right\|^2 = \left\| \sum_{i=1}^n y_i^k \right\|^2. \tag{53}$$

Next, define

$$\bar{u}^k \overset{\text{def}}{=} \frac{1}{n}\sum_{i=1}^n u_i^k = \frac{1}{n}\sum_{i=1}^n (x_i^k - z^k) = \bar{x}^k - z^k,$$

and observe that for all $i = 1, \ldots, n$ and $k \geq 0$,

$$u_i^k - \bar{u}^k = x_i^k - z^k - (\bar{x}^k - z^k) = x_i^k - \bar{x}^k. \tag{54}$$

Substituting (53) and (54) into the expression for $\|\nabla\varphi_k\|^2$ arising from Lemma 3, we obtain

$$\begin{aligned} \|\nabla\varphi_k\|^2 &= \left\|\sum_{i=1}^n y_i^k\right\|^2 + \sum_{i=1}^n \left\|x_k^k - \bar{x}^k\right\|^2 \\ &= \left\|\sum_{i=1}^n v_i^k\right\|^2 + \sum_{i=1}^n \left\|u_i^k - \bar{u}^k\right\|^2 \\ &= \langle v^k, \mathbf{E}v^k\rangle + \left\|\mathbf{M}u^k\right\|^2, \end{aligned}$$

where we define $n \times n$ matrices

$$\mathbf{E} \overset{\text{def}}{=} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} \qquad \mathbf{M} \overset{\text{def}}{=} \mathbf{I} - \tfrac{1}{n}\mathbf{E} = \begin{bmatrix} \frac{n-1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & \frac{n-1}{n} & \cdots & -\frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & \cdots & \frac{n-1}{n} \end{bmatrix}.$$

Applying the Cauchy-Schwarz inequality and (37), it then follows that

$$\|\nabla\varphi_k\|^2 \leq \left\|v^k\right\|\left\|\mathbf{E}v^k\right\| + \left\|\mathbf{M}u^k\right\|^2 \leq \|\mathbf{E}\|\left\|v^k\right\|^2 + \|\mathbf{M}\|^2\left\|u^k\right\|^2. \tag{55}$$

Over $\mathbb{R}^n$, the matrix $\mathbf{M}$ represents orthogonal projection onto the nontrivial subspace $T = \{(t_1, \ldots, t_n) \in \mathbb{R}^n \mid t_1 + \cdots + t_n = 0\}$, so we conclude $\|\mathbf{M}\| = 1$. It also follows that $I - M$ represents orthogonal projection onto the nontrivial subspace $T^\perp$, so $\|I - M\| = 1$ and $\|E\| = \|n(I - M)\| = n\|I - M\| = n$. Therefore, (55) reduces to

$$\|\nabla\varphi_k\|^2 \leq n\left\|v^k\right\|^2 + \left\|u^k\right\|^2. \tag{56}$$

Starting with (51), we obtain

$$\begin{aligned} \left\|v^k\right\|^2 &= \left\|\left(\mathbf{\Pi}_k\mathbf{\Lambda}_k^{-1}\mathbf{A}_k\mathbf{\Pi}_k^\top\right)u^k - \mathbf{\Pi}_k\mathbf{\Lambda}_k^{-1}e^k\right\|^2 \\ &\leq \left(\left\|\left(\mathbf{\Pi}_k\mathbf{\Lambda}_k^{-1}\mathbf{A}_k\mathbf{\Pi}_k^\top\right)u^k\right\| + \left\|\mathbf{\Pi}_k\mathbf{\Lambda}_k^{-1}e^k\right\|\right)^2 && [\,\text{triangle inequality}\,] \\ &\leq \left(\left\|\mathbf{\Pi}_k\mathbf{\Lambda}_k^{-1}\mathbf{A}_k\mathbf{\Pi}_k^\top\right\|\left\|u^k\right\| + \left\|\mathbf{\Lambda}_k^{-1}e^k\right\|\right)^2 && [\,\text{using (37)}\,] \\ &\leq \left(\left\|\mathbf{\Lambda}_k^{-1}\mathbf{A}_k\right\|\left\|u^k\right\| + \sigma\kappa(\mathbf{\Lambda}_k^{-1}\mathbf{A}_k)\left\|u^k\right\|\right)^2 && [\,\text{using (49)}\,] \\ &\leq \left(\beta\left\|u^k\right\| + \sigma\beta\left\|u^k\right\|\right)^2 && [\,\kappa(\mathbf{\Lambda}_k^{-1}\mathbf{A}_k) \leq \left\|\mathbf{\Lambda}_k^{-1}\mathbf{A}_k\right\| \leq \beta\,] \\ &= (1+\sigma)^2\beta^2\left\|u^k\right\|^2. \end{aligned} \tag{57}$$

Combining (56) and (57) yields

$$\|\nabla\varphi_k\|^2 \leq \left(n(1+\sigma)^2\beta^2 + 1\right)\left\|u^k\right\|^2.$$

Combining this inequality with (52) yields

$$\varphi_k(p^k) \geq \frac{(1-\sigma)\zeta}{n(1+\sigma)^2\beta^2+1} \left\|\nabla\varphi_k\right\|^2, \tag{58}$$

implying that hypothesis 3 of Proposition 5 is satisfied by setting

$$\xi = \frac{(1-\sigma)\zeta}{n(1+\sigma)^2\beta^2+1} > 0.$$

Note that (58) implies $\varphi_k(p_k)$ is always nonnegative, so that (44) is equivalent to (25), even though the $\max\{0,\cdot\}$ operation is omitted. In view of (58), Proposition 5 guarantees that $\nabla\varphi_k \to 0$ and $\varphi_k(p^k) \to 0$. From (52), we then conclude $u^k \to 0$, from which (57) implies that $v^k \to 0$. Thus, we have $u_i^k = x_i^k - z^k \to 0$ and $v_i^k = w_i^k - y_i^k \to 0$ for all $i = 1, \ldots, n$, fulfilling hypotheses 5 and 6 of Proposition 5. Hypothesis 4 is satisfied by assumption, so all the hypotheses of Proposition 5 hold. The (weak) convergence of the sequences $\{z^k\}$, $\{x_i^k\}$, $\{w_i^k\}$, and $\{y_i^k\}$ then follow from Proposition 5. □

Note that the approximation criterion (42) is implied by the simpler condition

$$\sum_{i=1}^{n}\left(\lambda_i^k\right)^{-2}\left\|e_i^k\right\|^2 \leq \sigma^2\zeta^2\sum_{i=1}^{n}\left\|x_i^k - z^k\right\|^2, \tag{59}$$

which might be more convenient to use in practice. The most convenient way to meet either (42) or (59) will likely depend on the application. One common situation is that only one of the operators, say T_1, has a resolvent difficult enough to warrant approximate computation. In this case, (42) would just simplify to $\|e_{\pi_k^{-1}(1)}\|^2 \leq \sigma^2\kappa(\mathbf{\Lambda}_k^{-1}\mathbf{A}_k)^2\sum_{i=1}^{n}\left\|x_i^k - z^k\right\|^2$. If more than one operator is a candidate for approximate computation, one simple option would be to requre

$$\left(\lambda_i^k\right)^{-2}\left\|e_i^k\right\|^2 \leq \sigma^2\kappa(\mathbf{\Lambda}_k^{-1}\mathbf{A}_k)^2\left\|x_i^k - z^k\right\|^2 \qquad i = 1, \ldots, n,$$

since summing these inequalities yields (42). However, this approach may be more restrictive than necessary. Typically, when an operator T_i is suitable for approximate calculation, the resolvent $(I+\lambda T_i)^{-1}$ is itself evaluated by some kind of iterative method. Thus, a less restrictive option would be to interleave iterations for calculating all of the $(x_i, y_i) \in \operatorname{Gph}(T_i)$, and terminate as soon as (42) itself is satisfied.

5 Variations and special cases

Rewriting (35) as

$$x_{\pi_k(i)}^k + \lambda_i^k y_{\pi_k(i)}^k = z^k + \sum_{j=1}^{i-1}\alpha_{ij}^k(x_{\pi_k(j)} - z^k) + \lambda_i^k w_{\pi_k(i)}^k + e_i^k,$$

it is natural to consider whether the algorithm could be further generalized by treating the y_i^k in a matter symmetric to the x_i^k. That is, for some β_{ij}^k, $1 \le j < i \le n$, one might try to use the $y_{\pi_k(j)}^k$ information generated earlier in the same iteration by replacing (35) with

$$x_{\pi_k(i)}^k + \lambda_i^k y_{\pi_k(i)}^k = z^k + \sum_{j=1}^{i-1} \alpha_{ij}^k (x_{\pi_k(j)} - z^k) + \lambda_i^k \left[w_{\pi_k(i)}^k + \sum_{j=1}^{i-1} \beta_{ij}^k (y_{\pi_k(j)} - w_{\pi_k(j)}^k) \right] + e_i^k.$$

However, if $e_i^k \equiv 0$, it turns out that this modification does not add any generality to the algorithm. The reason is that it is possible to redefine the α_{ij}^k to obtain an equivalent recursion with $\beta_{ij}^k \equiv 0$. We omit the analysis in the interest of brevity; while more complicated, it resembles that for a similar 2-operator result in [8].

5.1 Including a scaling factor

A simple variation of the algorithm may be obtained by multiplying the inclusion (1) through by any scalar $\eta > 0$, arriving at the rescaled formulation

$$0 \in \eta T_1(x) + \cdots + \eta T_n(x),$$

Applying Algorithm 3 to this formulation under the substitutions

$$\begin{aligned} T_i &\leftarrow \eta T_i & \qquad \lambda_i^k &\leftarrow \eta \lambda_i^k \\ y_i^k &\leftarrow \eta y_i^k & \qquad w_i^k &\leftarrow \eta w_i^k \end{aligned}$$

yields, after some minor algebraic manipulation, a procedure identical to Algorithm 3, except that (44)-(46) are modified to incorporate η:

$$y_i^k \in T_i(x_i^k) \qquad i = 1, \ldots, n \qquad (60)$$

$$x_{\pi_k(i)}^k + \lambda_i^k y_{\pi_k(i)}^k = \left(1 - \textstyle\sum_{j=1}^{i-1} \alpha_{ij}^k\right) z^k + \sum_{j=1}^{i-1} \alpha_{ij}^k x_{\pi_k(j)}^k + \lambda_i^k w_{\pi_k(i)}^k + e_i^k \qquad i = 1, \ldots, n \qquad (61)$$

$$\bar{x}^k = \tfrac{1}{n} \textstyle\sum_{i=1}^n x_i^k \qquad (62)$$

$$\theta_k = \frac{\sum_{i=1}^n \langle z^k - x_i^k, y_i^k - w_i^k \rangle}{\eta \left\| \sum_{i=1}^n y_i^k \right\|^2 + \frac{1}{\eta} \sum_{i=1}^n \left\| x_i^k - \bar{x}^k \right\|^2} \qquad (63)$$

$$z^{k+1} = z^k - \rho_k \theta_k \eta \textstyle\sum_{i=1}^n y_i^k \qquad (64)$$

$$w_i^{k+1} = w_i^k - \frac{\rho_k \theta_k}{\eta} (x_i^k - \bar{x}^k) \qquad i = 1, \ldots, n. \qquad (65)$$

This set of recursions produces sequences guaranteed to converge under the same conditions and in the same manner set forth in Proposition 7. Essentially, η sets the relative weight the algorithm ascribes to its two main goals: achieving $\sum_{i=1}^n y_i^k = 0$, and achieving $x_1^k = \cdots = x_n^k$. In practice, η could be adjusted as the algorithm runs if it appears that these goals are not properly balanced; with the theory presented here, however, convergence is only guaranteed for fixed η.

Suppose $n = 2$, $e_1^k = e_2^k = 0$ for all $k \ge 0$, and π_k is the identity map on $\{1, 2\}$ for all $k \ge 0$. Then, letting $\eta = 1/\sqrt{2}$ causes (60)-(65) to reduce precisely, after some changes of notation and minor algebraic manipulations, to the two-operator projective splitting algorithm of [8].

5.2 Spingarn's algorithm

In [22], Spingarn describes a partial inverse method for solving the inclusion $\eta_1 T_1(x) + \eta_2 T_2(x) + \cdots + \eta_n T_n(x) \ni 0$. With $\eta_1 = \eta_2 = \cdots = \eta_n$, this method reduces, in the notation of this paper, to the following set of recursions to solve (1):

$$y_i^k \in T_i(x_i^k) \qquad i = 1, \ldots, n \tag{66}$$

$$x_i^k + y_i^k = z^k + w_i^k \qquad i = 1, \ldots, n \tag{67}$$

$$z^{k+1} = \tfrac{1}{n} \textstyle\sum_{i=1}^n x_i^k \tag{68}$$

$$w_i^{k+1} = y_i^k - \tfrac{1}{n} \textstyle\sum_{j=1}^n y_j^k \qquad i = 1, \ldots, n. \tag{69}$$

The resolvent evaluations entailed in (66)-(67) are in fact the same as suggested for the separator calculation at the end of Section 2 of this paper, and are clearly a special case of our general recursion (35). In fact, we now demonstrate that Spingarn's method (66)-(69) is a special case of the scaled variant (60)-(65) of our algorithm. Consider (60)-(65) with, for all $k \geq 0$,

$$\lambda_i^k = 1 \qquad i = 1, \ldots, n \tag{70}$$

$$\pi_k(i) = i \qquad i = 1, \ldots, n \tag{71}$$

$$\alpha_{ij}^k = 0 \qquad 1 \leq j < i \leq n \tag{72}$$

$$e_i^k = 0 \qquad i = 1, \ldots, n \tag{73}$$

$$\rho_k = 1. \tag{74}$$

Then the main resolvent relation (61) reduces immediately to (67). Rearranging (67) into $z^k - x_i^k = y_i^k - w_i^k$, we deduce that the numerator of (63) is

$$\sum_{i=1}^n \langle z^k - x_i^k, y_i^k - w_i^k \rangle = \sum_{i=1}^n \left\| z^k - x_i^k \right\|^2 . \tag{75}$$

Now consider the denominator of (63). With regard to the first term, we rewrite (67) as $y_i^k = z^k - w_i^k + y_i^k$ and then observe that since $\sum_{i=1}^n w_i^k = 0$,

$$\sum_{i=1}^n y_i^k = \sum_{i=1}^n \left(z^k - w_i^k + x_i^k \right) = n z^k - \sum_{i=1}^n x_i^k = n(z^k - \bar{x}^k). \tag{76}$$

With regard to the second term in the denominator of (63), we calculate

$$\begin{aligned} \sum_{i=1}^n \left\| x_i^k - \bar{x}^k \right\|^2 &= \sum_{i=1}^n \left\| (x_i^k - z^k) - (\bar{x}^k - z^k) \right\|^2 \\ &= \sum_{i=1}^n \left\| x_i^k - z^k \right\|^2 - 2 \left\langle \textstyle\sum_{i=1}^n (x_i^k - z^k), \bar{x}^k - z^k \right\rangle + n \left\| \bar{x}^k - z^k \right\|^2 \\ &= \sum_{i=1}^n \left\| x_i^k - z^k \right\|^2 - 2 \left\langle n(\bar{x}^k - z^k), \bar{x}^k - z^k \right\rangle + n \left\| \bar{x}^k - z^k \right\|^2 \\ &= \sum_{i=1}^n \left\| x_i^k - z^k \right\|^2 - n \left\| \bar{x}^k - z^k \right\|^2 . \end{aligned} \tag{77}$$

Using (76) and (77), we calculate that the denominator of (63) is

$$\eta \left\|\sum_{i=1}^n y_i^k\right\|^2 + \frac{1}{\eta}\sum_{i=1}^n \left\|x_i^k - \bar{x}^k\right\|^2 = n^2\eta \left\|z^k - \bar{x}^k\right\|^2 + \frac{1}{\eta}\left(\sum_{i=1}^n \left\|x_i^k - z^k\right\|^2 - n\left\|\bar{x}^k - z^k\right\|^2\right)$$
$$= \left(n^2\eta - \frac{n}{\eta}\right)\left\|z^k - \bar{x}^k\right\|^2 + \frac{1}{\eta}\sum_{i=1}^n \left\|x_i^k - z^k\right\|^2. \tag{78}$$

Solving the equation $n^2\eta - n/\eta = 0$, we conclude that the first term in (78) will vanish if $\eta = 1/\sqrt{n}$. Combining (63), (75), and (78) with $\eta = 1/\sqrt{n}$, we obtain

$$\theta_k = \frac{\sum_{i=1}^n \left\|x_i^k - z^k\right\|^2}{\frac{1}{\eta}\sum_{i=1}^n \left\|x_i^k - z^k\right\|^2} = \eta = \frac{1}{\sqrt{n}},$$

unless the denominator is zero, in which case $(z^k, w_1^k, \ldots, w_n^k)$ is already a solution to (1). Substituting $\rho_k = 1$, $\theta_k = \eta = 1/\sqrt{n}$, and (76) into (64), we obtain

$$z^{k+1} = z^k + \frac{1}{n}\left(\sum_{i=1}^n y_i^k\right) = z^k + \frac{1}{n}\big(n(z^k - \bar{x}^k)\big) = \bar{x}^k,$$

which is identical to (68). Similarly substituting $\rho_k = 1$ and $\theta_k = \eta = 1/\sqrt{n}$ into (65) yields

$$w_i^{k+1} = w_i^k - \frac{\eta}{\eta}(x_i^k - \bar{x}^k) = w_i^k - x_i^k + \bar{x}^k.$$

From (67), we have $w_i^k - x_i^k = y_i^k - z^k$; using this fact and the definition of $\bar{x}^k$, we then have

$$w_i^{k+1} = y_i^k - z^k + \frac{1}{n}\sum_{i=1}^n x_i^k.$$

Finally, we rearrange (67) into $x_i^k = z^k + w_i^k - y_i^k$ and obtain, using $\sum_{i=1}^n w_i^k = 0$, that

$$\begin{aligned} w_i^{k+1} &= y_i^k - z^k + \frac{1}{n}\sum_{i=1}^n \left(z^k + w_i^k - y_i^k\right) \\ &= y_i^k - z^k + \frac{1}{n}\left(nz^k - \sum_{i=1}^n y_i^k\right) \\ &= y_i^k - \frac{1}{n}\sum_{i=1}^n y_i^k, \end{aligned}$$

which is identical to (69). Thus, we conclude that with the parameter choices (70)-(74) and $\eta = 1/\sqrt{n}$, the scaled projective algorithm (60)-(65) reduces exactly to Spingarn's algorithm (66)-(69).

We should also note that, by scaling the operators $T_1, \ldots, T_n$ by some fixed $\lambda > 0$, and applying an overrelaxed proximal algorithm (see for example [6]) to the partial inverse operator, Spingarn's original algorithm (66)-(69) for (1) is easily generalized to

$$
\begin{aligned}
y_i^k &\in T_i(x_i^k) && i = 1, \ldots, n \\
x_i^k + \lambda y_i^k &= z^k + \lambda w_i^k && i = 1, \ldots, n \\
z^{k+1} &= (1 - \rho_k) z^k + \frac{\rho_k}{n} \textstyle\sum_{i=1}^n x_i^k && \\
w_i^{k+1} &= (1 - \rho_k) w_i^k + \rho_k \left(y_i^k - \tfrac{1}{n} \textstyle\sum_{j=1}^n y_j^k \right) && i = 1, \ldots, n,
\end{aligned}
$$

where $0 < \underline{\rho} \leq \rho_k \leq \overline{\rho} < 2$ for all k. By a similar analysis, but using $\eta = \lambda / \sqrt{n}$, this generalized Spingarn algorithm also turns out to be a special case of the scaled projective method (60)-(65).

6 Conclusions and future research

We have proved convergence of a very general class of projective splitting algorithms, extending the results of [8] by allowing for more than two operators, changing order of operator evaluation, and approximate calculation of resolvents using a "relative" error criterion.

At this point, the key question becomes whether the new flexibility our framework affords in comparison with prior splitting algorithms is of significant practical value. By taking advantage of this new flexibility and larger number of parameters, can one significantly accelerate the convergence of splitting-based algorithms for practical problems?

Answering this question is outside the scope of this paper, but we hope to address it in future research. In particular, one might anticipate dynamically adjusting the parameters $\{\alpha_{ij}^k\}$, $\{\lambda_i^k\}$, $\{\rho_k\}$, and perhaps η, to attempt to optimize some convergence criterion or error bound. However, the details are may very well vary by application. Another interesting topic would be to examine applying the techniques of [20] to force strong convergence, and perhaps to improve practical finite-dimensonal convergence behavior.

References

[1] H. H. Bauschke and J. M. Borwein. On projection algorithms for solving convex feasibility problems. *SIAM Rev.*, 38(3):367–426, 1996.

[2] H. H. Bauschke, P. L. Combettes, and S. Reich. The asymptotic behavior of the composition of two resolvents. *Nonlin. Anal.*, 60(2):283–301, 2005.

[3] G. Cimmino. Calcolo approssimato per le soluzioni dei sistemi di equazioni lineari. *Ric. Sci. Progr. Tecn. Econom. Naz.*, 1:326–333, 1938.

[4] P. L. Combettes. Solving monotone inclusions via compositions of nonexpansive averaged operators. *Optimization*, 53(5-6):475–504, 2004.

[5] J. Douglas, Jr. and H. H. Rachford, Jr. On the numerical solution of heat conduction problems in two and three space variables. *Trans. Amer. Math. Soc.*, 82:421–439, 1956.

[6] J. Eckstein and D. P. Bertsekas. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Program.*, 55(3):293–318, 1992.

[7] J. Eckstein and M. C. Ferris. Operator-splitting methods for monotone affine variational inequalities, with a parallel application to optimal control. *INFORMS J. Comput.*, 10(2):218–235, 1998.

[8] J. Eckstein and B. F. Svaiter. A family of projective splitting methods for the sum of two maximal monotone operators. *Math. Program.*, appeared online, January 2007.

[9] D. Gabay. Applications of the method of multipliers to variational inequalities. In M. Fortin and R. Glowinski, editors, *Augmented Lagrangian Methods: Applications to the Solution of Boundary Value Problems*, chapter IX, pages 299–340. North-Holland, Amsterdam, 1983.

[10] S. Kaczmarz. Angenäherte auflösung von systemen linearer gleichungen. *Bull. Int. Acad. Polon. Sci. A*, 1937:355–357, 1937.

[11] S. Kaczmarz. Approximate solution of systems of linear equations. *Internat. J. Control*, 57(6):1269–1271, 1993. Translated from the German.

[12] J. Lawrence and J. E. Spingarn. On fixed points of nonexpansive piecewise isometric mappings. *Proc. London Math. Soc. (3)*, 55(3):605–624, 1987.

[13] P.-L. Lions. Une méthode itérative de résolution d'une inéquation variationnelle. *Israel J. Math.*, 31(2):204–208, 1978.

[14] P.-L. Lions and B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.*, 16(6):964–979, 1979.

[15] G. J. Minty. Monotone (nonlinear) operators in Hilbert space. *Duke Math. J.*, 29:341–346, 1962.

[16] G. B. Passty. Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. *J. Math. Anal. Appl.*, 72(2):383–390, 1979.

[17] D. W. Peaceman and H. H. Rachford, Jr. The numerical solution of parabolic and elliptic differential equations. *J. SIAM*, 3:28–41, 1955.

[18] M. V. Solodov and B. F. Svaiter. A hybrid approximate extragradient-proximal point algorithm using the enlargement of a maximal monotone operator. *Set-Valued Anal.*, 7(4):323–345, 1999.

[19] M. V. Solodov and B. F. Svaiter. A hybrid projection-proximal point algorithm. *J. Convex Anal.*, 6(1):59–70, 1999.

[20] M. V. Solodov and B. F. Svaiter. Forcing strong convergence of proximal point iterations in a Hilbert space. *Math. Program.*, 87(1):189–202, 2000.

[21] M. V. Solodov and B. F. Svaiter. A unified framework for some inexact proximal point algorithms. *Numer. Funct. Anal. Optim.*, 22(7-8):1013–1035, 2001.

[22] J. E. Spingarn. Partial inverse of a monotone operator. *Appl. Math. Optim.*, 10(3):247–265, 1983.

[23] P. Tseng. A modified forward-backward splitting method for maximal monotone mappings. *SIAM J. Control Optim.*, 38(2):431–446, 2000.

A A technical result for infinite dimension

Proposition 8 *Let $T_1, \ldots, T_n : \mathcal{H} \rightrightarrows \mathcal{H}$ be maximal monotone, and suppose that their sum $T_1 + \cdots + T_n$ is also maximal. Suppose the sequences $\{(x_i^k, y_i^k)\}_{k=1}^{\infty} \subset \mathrm{Gph}(T_i)$, $i = 1, \ldots, n$, and points $z, w_1, \ldots, w_n \in \mathcal{H}$ have the properties*

$$(x_i^k, y_i^k) \xrightarrow{\mathsf{w}} (z, w_i) \qquad i = 1, \ldots, n \tag{79}$$

$$\textstyle\sum_{i=1}^n y_i^k \to 0 \tag{80}$$

$$\|x_i^k - x_j^k\| \to 0 \qquad i, j = 1, \ldots, n. \tag{81}$$

Then $(z, w_1, \ldots, w_n) \in S_{\mathrm{e}}(T_1, \ldots, T_n)$.

Proof. First we claim that

$$0 \in (T_1 + \cdots + T_n)(z). \tag{82}$$

To prove this claim, take an arbitrary $(z', w') \in \mathrm{Gph}(T_1 + \cdots + T_n)$. Then, there exist points $w_i' \in T_i(z')$, $i = 1, \ldots, n$, such that $w' = \sum_{i=1}^n w_i'$. Since all the T_i are monotone,

$$\langle x_i^k - z', y_i^k - w_i' \rangle \geq 0 \qquad i = 1, \ldots, n. \tag{83}$$

Define $\bar{y}^k = \sum_{i=1}^n y_i^k$, and fix any $j \in \{1, \ldots, n\}$. We may rewrite the $i = j$ case of (83) as

$$\Big\langle x_j^k - z', \bar{y}^k - w' + \sum_{\substack{i=1 \\ i \neq j}}^n (w_i' - y_i^k) \Big\rangle \geq 0,$$

and so

$$\langle x_j^k - z', -w' \rangle \geq -\langle x_j^k - z', \bar{y}^k \rangle + \sum_{\substack{i=1 \\ i \neq j}}^n \langle x_j^k - z', y_i^k - w_i' \rangle \tag{84}$$

For any $i \neq j$, we have, courtesy of (83), that

$$\langle x_j^k - z', y_i^k - w_i' \rangle = \langle x_i^k - z', y_i^k - w_i' \rangle + \langle x_j^k - x_i^k, y_i^k - w_i' \rangle \geq \langle x_j^k - x_i^k, y_i^k - w_i' \rangle,$$

and substituting these inequalities into (84) yields

$$\langle x_j^k - z', -w' \rangle \geq -\langle x_j^k - z', \bar{y}^k \rangle + \sum_{\substack{i=1 \\ i \neq j}}^{n} \langle x_i^k - x_j^k, y_i^k - w_i' \rangle \tag{85}$$

We now consider taking $k \to \infty$ in (85). Since $x_j^k \xrightarrow{\mathsf{w}} z$ by (79), the limit of the left-hand side of (85) is $\langle z - z', -w' \rangle$. Since the weakly convergent sequence $\{x_j^k\}$ must be bounded, and $\bar{y}^k \to 0$ by (80), we have $\langle x_j^k - z', \bar{y}^k \rangle \to 0$. Similarly, for each $i \neq j$, we have that $x_i^k - x_j^k \to 0$ by (81), and the weakly convergent sequence $\{y_i^k\}$ must be bounded, so that $\langle x_i^k - x_j^k, y_i^k - w_i' \rangle \to 0$. Thus, taking the limit in (85) yields

$$\langle z - z', -w' \rangle \geq 0.$$

Since $T_1 + \cdots + T_n$ is maximal monotone, and $(z', w') \in \mathrm{Gph}(T_1 + \cdots + T_n)$ was arbitrary, we conclude that (82) holds.

Next, we claim that

$$\lim_{k\to\infty} \langle x_i^k, y_i^k \rangle = \langle z, w_i \rangle \qquad i = 1, \ldots, n. \tag{86}$$

In view of (82), there must exist $u_i \in T_i(z)$, $i = 1, \ldots, n$, such that $\sum_{i=1}^n u_i = 0$. Since the T_i are monotone, we have

$$\langle x_i^k - z, y_i^k - u_i \rangle \geq 0 \qquad i = 1, \ldots, n,$$

which we may rearrange to obtain

$$\langle x_i^k, y_i^k \rangle \geq \langle z, y_i^k - u_i \rangle + \langle x_i^k, u_i \rangle \qquad i = 1, \ldots, n.$$

From (79), it is easily deduced that the right-hand sides of the above inequalities converge respectively to $\langle z, w_i \rangle$. Hence,

$$\liminf_{k\to\infty} \langle x_i^k, y_i^k \rangle \geq \langle z, w_i \rangle \qquad i = 1, \ldots, n. \tag{87}$$

Once again, fix some $j \in \{1, \ldots, n\}$. Then, we observe that

$$\begin{aligned} \langle x_j^k, y_j^k \rangle &= \langle x_j^k, \bar{y}^k \rangle - \sum_{\substack{i=1 \\ i \neq j}}^{n} \langle x_j^k, y_i^k \rangle \\ &= \langle x_j^k, \bar{y}^k \rangle - \sum_{\substack{i=1 \\ i \neq j}}^{n} \big(\langle x_i^k, y_i^k \rangle + \langle x_j^k - x_i^k, y_i^k \rangle \big) . \end{aligned}$$

We now take the lim sup as $k \to \infty$ of the above equation. Using logic resembling that for (85), we observe that $\langle x_j^k, \bar{y}^k \rangle \to 0$ and $\langle x_j^k - x_i^k, y_i^k \rangle \to 0$. Therefore, using (87),

$$\limsup_{k\to\infty} \langle x_j^k, y_j^k \rangle \leq -\sum_{\substack{i=1 \\ i \neq j}}^{n} \langle z, w_i \rangle = -\Big\langle z, \sum_{\substack{i=1 \\ i \neq j}}^{n} w_i \Big\rangle. \tag{88}$$

Since $y_i^k \xrightarrow{\mathrm{w}} w_i$, $i = 1, \ldots, n$, we have $\sum_{i=1}^n y_i^k \xrightarrow{\mathrm{w}} \sum_{i=1}^n w_i$, and therefore, also using that $\sum_{i=1}^n y_i^k \to 0$,

$$\left\langle \sum_{i=1}^n w_i, \sum_{i=1}^n w_i \right\rangle = \lim_{k\to\infty} \left\langle \sum_{i=1}^n y_i^k, \sum_{i=1}^n w_i \right\rangle = \left\langle \lim_{k\to\infty} \sum_{i=1}^n y_i^k, \sum_{i=1}^n w_i \right\rangle = 0,$$

so we must have $\sum_{i=0}^n w_i = 0$. Therefore, (88) may be rewritten

$$\limsup_{k\to\infty} \langle x_j^k, y_j^k \rangle \leq \langle z, w_j \rangle,$$

which, combined with (87), means that $\lim_{k\to\infty} \langle x_j^k, y_j^k \rangle = \langle z, w_j \rangle$. Since $j \in \{1, \ldots, n\}$ was arbitrary, (86) holds.

Finally, we claim that

$$(z, w_i) \in \mathrm{Gph}(T_i), \qquad i = 1, \ldots, n. \tag{89}$$

To prove this inclusion, take any $i \in \{1, \ldots, n\}$ and $(z', w_i') \in \mathrm{Gph}(T_i)$. Then the monotonicity of T_i implies

$$\langle x_i^k - z', y_i^k - w_i' \rangle = \langle x_i^k, y_i^k \rangle - \langle z', y_i^k \rangle - \langle x_i^k, w_i' \rangle + \langle z', w_i' \rangle \geq 0.$$

Applying (79) and (86) while taking the limit as $k \to \infty$ yields

$$\langle z, w_i \rangle - \langle z', w_i \rangle - \langle z, w_i' \rangle + \langle z', w_i' \rangle \geq 0.$$

which is equivalent to $\langle z - z', w_i - w_i' \rangle \geq 0$. Since the T_i are maximal and both $(z', w_i') \in \mathrm{Gph}(T_i)$ and $i \in \{1, \ldots, n\}$ were arbitrary, we conclude that (89) holds.

Finally, since we have already established that $\sum_{i=0}^n w_i = 0$, it follows from (89) that we must have $(z, w_1, \ldots, w_n) \in S_{\mathrm{e}}(T_1, \ldots, T_n)$. □

B Proof of Lemma 6

Proof. If $u = 0$, then $\mathbf{L}u = 0$ and (37)-(38) hold trivially, so it remains to consider the case that at least one u_i is nonzero. Given any such u, let $v = (v_1, \ldots, v_n)$ and ℓ_{ij} be defined as in (36). Define $U \subseteq \mathcal{H}$ to be the finite-dimensional subspace spanned by $u_1, \ldots, u_n$ in $\mathcal{H}$. From (36), we have $v_i \in U$ for $i = 1, \ldots, n$, and thus $u, v \in U^n$. Let $B = (b_1, \ldots, b_{n'})$, $1 \leq n' \leq n$, be some orthonormal basis for U, where n' denotes the dimension of U. From

B, we may create an orthonormal basis $\overline{B} = (\bar{b}_1, \ldots, \bar{b}_{n'n})$ for U^n via

$$\begin{aligned}
\bar{b}_1 &= (b_1, 0, 0, \ldots, 0) \\
\bar{b}_2 &= (0, b_1, 0, \ldots, 0) \\
&\vdots \\
\bar{b}_n &= (0, 0, \ldots, 0, b_1) \\
\bar{b}_{n+1} &= (b_2, 0, 0, \ldots, 0) \\
\bar{b}_{n+2} &= (0, b_2, 0, \ldots, 0) \\
&\vdots \\
\bar{b}_{2n} &= (0, 0, \ldots, 0, b_2) \\
&\vdots \\
\bar{b}_{(n'-1)n+1} &= (b_{n'}, 0, 0, \ldots, 0) \\
\bar{b}_{(n'-1)n+2} &= (0, b_{n'}, 0, \ldots, 0) \\
&\vdots \\
\bar{b}_{n'n} &= (0, 0, \ldots, 0, b_{n'}).
\end{aligned}$$

Let $\overline{\mathbf{u}} \in \mathbb{R}^{n'n}$ be the unique representation of u with respect to this basis, that is, its elements $\overline{u}_m$, $m = 1, \ldots, n'n$, are such that $u = \sum_{m=1}^{n'n} \overline{u}_m \bar{b}_m$. Similarly, let $\overline{\mathbf{v}} \in \mathbb{R}^{n'n}$ be the unique representation of v. By the orthonormality of the basis $\overline{B}$, it follows that $\|u\| = \|\overline{\mathbf{u}}\|$, $\|v\| = \|\overline{\mathbf{v}}\|$, and $\langle u, \mathbf{L}u \rangle = \langle u, v \rangle = \overline{\mathbf{u}}^\top \overline{\mathbf{v}}$. Let us now examine the action of the linear mapping defined by (36) on the basis $\overline{B}$, namely

$$\begin{aligned}
\bar{b}_1 = (b_1, 0, 0, \ldots, 0) &\mapsto (\ell_{11} b_1, \ell_{21} b_1, \ldots, \ell_{n1} b_1) = \ell_{11}\bar{b}_1 + \ell_{21}\bar{b}_2 + \cdots + \ell_{n1}\bar{b}_n \\
\bar{b}_2 = (0, b_1, 0, \ldots, 0) &\mapsto (\ell_{12} b_1, \ell_{22} b_1, \ldots, \ell_{n2} b_1) = \ell_{12}\bar{b}_1 + \ell_{22}\bar{b}_2 + \cdots + \ell_{n2}\bar{b}_n \\
&\vdots \\
\bar{b}_n = (0, 0, \ldots, 0, b_1) &\mapsto (\ell_{1n} b_1, \ell_{2n} b_1, \ldots, \ell_{nn} b_1) = \ell_{1n}\bar{b}_1 + \ell_{2n}\bar{b}_2 + \cdots + \ell_{nn}\bar{b}_n \\
\bar{b}_{n+1} = (b_2, 0, 0, \ldots, 0) &\mapsto (\ell_{11} b_2, \ell_{21} b_2, \ldots, \ell_{n1} b_2) = \ell_{11}\bar{b}_{n+1} + \ell_{21}\bar{b}_{n+2} + \cdots + \ell_{n1}\bar{b}_{2n} \\
\bar{b}_{n+2} = (0, b_2, 0, \ldots, 0) &\mapsto (\ell_{12} b_2, \ell_{22} b_2, \ldots, \ell_{n2} b_2) = \ell_{12}\bar{b}_{n+1} + \ell_{22}\bar{b}_{n+2} + \cdots + \ell_{n2}\bar{b}_{2n} \\
&\vdots \\
\bar{b}_{n'n} = (0, 0, \ldots, 0, b_{n'}) &\mapsto (\ell_{1n} b_{n'}, \ell_{2n} b_{n'}, \ldots, \ell_{nn} b_{n'}) \\
&\qquad = \ell_{1n}\bar{b}_{(n'-1)n+1} + \ell_{2n}\bar{b}_{(n'-1)n+2} + \cdots + \ell_{nn}\bar{b}_{n'n}.
\end{aligned}$$

Thus, in terms of the basis $\overline{B}$, the action of the linear mapping (36) is that of the $n'n \times n'n$ block-diagonal matrix

$$\overline{\mathbf{L}} \stackrel{\text{def}}{=} \underbrace{\begin{bmatrix} \mathbf{L} & & & \\ & \mathbf{L} & & \\ & & \ddots & \\ & & & \mathbf{L} \end{bmatrix}}_{n' \text{ times}},$$

and we must have $\overline{\mathbf{v}} = \overline{\mathbf{L}}\overline{\mathbf{u}}$. We have

$$\operatorname{sym}\overline{\mathbf{L}} = \begin{bmatrix} \operatorname{sym}\mathbf{L} & & & \\ & \operatorname{sym}\mathbf{L} & & \\ & & \ddots & \\ & & & \operatorname{sym}\mathbf{L} \end{bmatrix},$$

so $\operatorname{sym}\overline{\mathbf{L}}$ has the same eigenvalues as $\operatorname{sym}\mathbf{L}$, and thus $\kappa(\overline{\mathbf{L}}) = \kappa(\mathbf{L})$. Using standard eigenvalue analysis in $\mathbb{R}^{n'n}$, we therefore have

$$\overline{\mathbf{u}}^\top\overline{\mathbf{L}}\overline{\mathbf{u}} \geq \kappa(\overline{\mathbf{L}}) \left\|\overline{\mathbf{u}}\right\|^2 = \kappa(\mathbf{L}) \left\|\overline{\mathbf{u}}\right\|^2.$$

Substituting $\|u\| = \|\overline{\mathbf{u}}\|$ and $\langle u, \mathbf{L}u\rangle = \langle u, v\rangle = \overline{\mathbf{u}}^\top\overline{\mathbf{v}} = \overline{\mathbf{u}}^\top\overline{\mathbf{L}}\overline{\mathbf{u}}$ into this relation yields (38). To establish (37), we observe that

$$\begin{aligned} \left\|\overline{\mathbf{L}}\right\|^2 &= \max\left\{ \left\|\overline{\mathbf{L}}\overline{\mathbf{x}}\right\|^2 \;\middle|\; \overline{\mathbf{x}} \in \mathbb{R}^{n'n},\ \|\overline{\mathbf{x}}\| = 1 \right\} \\ &= \max\left\{ \sum_{j=1}^{n'} \left\|\mathbf{L}\mathbf{x}_j\right\|^2 \;\middle|\; \mathbf{x}_1, \ldots, \mathbf{x}_{n'} \in \mathbb{R}^n,\ \sum_{j=1}^{n'} \left\|\mathbf{x}_j\right\|^2 = 1 \right\} \\ &= \max\left\{ \sum_{j=1}^{n'} \max\left\{ \|\mathbf{L}\mathbf{x}\|^2 \;\middle|\; \begin{array}{l} \mathbf{x} \in \mathbb{R}^n \\ \|\mathbf{x}\|^2 = \nu_j \end{array} \right\} \;\middle|\; \begin{array}{l} \nu_1, \ldots, \nu_{n'} \geq 0 \\ \nu_1 + \cdots + \nu_{n'} = 1 \end{array} \right\} \\ &= \max\left\{ \sum_{j=1}^{n'} \nu_j \|\mathbf{L}\|^2 \;\middle|\; \begin{array}{l} \nu_1, \ldots, \nu_{n'} \geq 0 \\ \nu_1 + \cdots + \nu_{n'} = 1 \end{array} \right\} \;=\; \|\mathbf{L}\|^2. \end{aligned}$$

Thus, we may substitute $\left\|\overline{\mathbf{L}}\right\| = \|\mathbf{L}\|$ into the inequality $\left\|\overline{\mathbf{L}}\overline{\mathbf{u}}\right\| \leq \left\|\overline{\mathbf{L}}\right\| \|\overline{\mathbf{u}}\|$, along with $\left\|\overline{\mathbf{L}}\overline{\mathbf{u}}\right\| = \|\overline{\mathbf{v}}\| = \|v\| = \|\mathbf{L}u\|$ and $\|u\| = \|\overline{\mathbf{u}}\|$, to obtain (37). □