On the asymptotic behaviour of the Aragón Artacho–Campoy algorithm

Salihah Alwadani[*], Heinz H. Bauschke[†], Walaa M. Moursi[‡] and Xianfu Wang[§]

May 28, 2018

Abstract

Aragón Artacho and Campoy recently proposed a new method for computing the projection onto the intersection of two closed convex sets in Hilbert space; moreover, they proposed in 2018 a generalization from normal cone operators to maximally monotone operators. In this paper, we complete this analysis by demonstrating that the underlying curve converges to the nearest zero of the sum of the two operators. We also provide a new interpretation of the underlying operators in terms of the resolvent and the proximal average.

2010 Mathematics Subject Classification: Primary 47H05, 90C25; Secondary 47H09, 49M27, 65K05, 65K10

Keywords: Aragón Artacho–Campoy algorithm, convex function, Douglas–Rachford algorithm, maximally monotone operator, Peaceman–Rachford algorithm, projection, proximal average, proximal operator, resolvent, resolvent average, zero the sum.

1 Introduction

Throughout this note,

$$\boxed{X \text{ is a real Hilbert space}} \tag{1}$$

with inner product $\langle \cdot \mid \cdot \rangle$ and associated norm $\|\cdot\|$. The notation of our paper is standard and mainly follows [6] to which we also refer to basic results on convex analysis and monotone operator theory. A central problem is to find a zero (critical point) of the sum of two maximally monotone operators. The Douglas–Rachford and Peaceman–Rachford algorithms (see Fact 2.1 below) are classical approaches to solve this problem. If the monotone operators are normal cone operators of closed convex nonempty subsets of X, then one obtains a feasibility problem. Suppose, however, that we are interested in finding the nearest point in the intersection. One may then apply several classical best approximation algorithms (see, e.g., [6, Chapter 30]). In the recently published paper [1], Aragón Artacho and Campoy presented a novel algorithm, which we term the *Aaragón Artacho–Campoy Algorithm (AACA)* to solve this best approximation problem. Even more recently, they extended this algorithm in [2] to deal with general maximally monotone operators.

The aim of this paper is to re-derive the AACA from the view point of the proximal and resolvent average. We also complete their analysis by describing the asymptotic behaviour of the underlying curve.

[*]Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. Email: salihahalwadani@hotmail.com.

[†]Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. Email: heinz.bauschke@ubc.ca.

[‡]Electrical Engineering, Stanford University, Stanford, CA 94305, USA and Mansoura University, Faculty of Science, Mathematics Department, Mansoura 35516, Egypt. Email: wmoursi@stanford.edu.

[§]Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. Email: shawn.wang@ubc.ca.

This note is organized as follows. In Section 2, we collect a few facts and results that will make the subsequent analysis more clear. Section 3 contains a new variant of a convergence result for AACA (Theorem 3.2) as well as the announced asymptotic behaviour of the curve (Theorem 3.4).

2 Auxiliary results

Fact 2.1 (Douglas–Rachford and Peaceman–Rachford) Let A and B be maximally monotone on X. Suppose that $\operatorname{zer}(A+B) = (A+B)^{-1}(0) \neq \varnothing$, let $\lambda \in \left]0,1\right]$, and set

$$T = (1-\lambda)\operatorname{Id} + \lambda R_B R_A, \tag{2}$$

where $J_A = (\operatorname{Id} + A)^{-1}$ and $R_A = 2J_A - \operatorname{Id}$. Let $x_0 \in X$ and define

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = Tx_n. \tag{3}$$

Then there exists $\bar{x} \in \operatorname{Fix} T$ such that $\bar{z} = J_A\bar{x} \in \operatorname{zer}(A+B)$ and the following hold:

(i) If A or B is strongly monotone, then $\operatorname{zer}(A+B) = \{\bar{z}\}$.

(ii) If $\lambda < 1$, then $x_n \rightharpoonup \bar{x}$ and $J_A x_n \rightharpoonup \bar{z}$.

(iii) If $\lambda < 1$ and A or B is strongly monotone, then $J_A x_n \to \bar{z}$.

(iv) If $\lambda = 1$ and A is strongly monotone, then $J_A x_n \to \bar{z}$.

Proof. This follows from [6, Theorem 26.11 and Proposition 26.13]. See also [8]. ■

The proof of the following result, which is a slight generalization of [2, Proposition 3.1], is straightforward and hence omitted.

Proposition 2.2 *Let C be maximally monotone on X, let $w \in X$, let $\gamma \in \left]0,1\right]$, and set*

$$C_\gamma \colon X \rightrightarrows X \colon x \mapsto C(\gamma^{-1}(x-(1-\gamma)w)) + (1-\gamma)\gamma^{-1}(x-w). \tag{4}$$

Then C_γ is maximally monotone and its resolvent is given by

$$J_{C_\gamma} \colon X \to X \colon x \mapsto \gamma J_C x + (1-\gamma)w. \tag{5}$$

Remark 2.3 (resolvent and proximal average) Consider the setting of Proposition 2.2. Because J_{C_γ} is a convex combination of the resolvents J_C and $P_{\{w\}}$, we see that C_γ is nothing but a *resolvent average* of C and $N_{\{w\}}$. See [3] for a detailed study of resolvent averages. We note that if C is σ_C-monotone, i.e., $C - \sigma_C \operatorname{Id}$ is monotone, then

$$C_\gamma \text{ is } \gamma^{-1}(\sigma_C + 1 - \gamma)\text{-monotone}. \tag{6}$$

This can be verified directly (as in [2, Proposition 3.1]) or it also follows from [3, Theorem 3.20].

Now suppose that additionally $C = \partial h$ for some proper lower semicontinuous convex function h on X. Then $C_\gamma = \partial h_\gamma$ and $J_{C_\gamma} = \operatorname{Prox}_{h_\gamma}$, where

$$h_\gamma \colon X \to \left]-\infty,+\infty\right] \tag{7a}$$

$$x \mapsto \inf\left\{ \gamma h(y_1) + (1-\gamma)\iota_{\{w\}}(y_2) + \frac{\gamma(1-\gamma)}{2}\|y_1-y_2\|^2 \;\middle|\; \gamma y_1 + (1-\gamma)y_2 = x \right\} \tag{7b}$$

$$= \gamma h\big(\gamma^{-1}(x-(1-\gamma)w)\big) + \frac{\gamma(1-\gamma)}{2}\|\gamma^{-1}(x-(1-\gamma)w) - w\|^2 \tag{7c}$$

$$= \gamma h\big(\gamma^{-1}(x-(1-\gamma)w)\big) + \frac{1-\gamma}{2\gamma}\|x-w\|^2 \tag{7d}$$

is the *proximal average* of h and $\iota_{\{w\}}$. See [7] and the reference therein for more on the proximal average.

3 The Aragón Artacho–Campoy algorithm (AACA)

From now on, we suppose that

$$\boxed{A \text{ and } B \text{ are maximally monotone on } X,\ w \in X, \text{ and } \gamma \in \,]0,1[.} \tag{8}$$

Let $\sigma_A \geqslant 0$ and $\sigma_B \geqslant 0$ be such that

$$\boxed{A - \sigma_A \operatorname{Id} \text{ and } B - \sigma_B \operatorname{Id} \text{ are monotone,}} \tag{9}$$

and we also assume that

$$\boxed{A + B \text{ is maximally monotone}} \tag{10}$$

which will make all results more tidy. (See also Remark 3.3 below.) Next, as in Remark 2.3, we introduce the resolvent averages between A, B and $N_{\{w\}}$:

$$\boxed{A_\gamma \colon X \rightrightarrows X \colon x \mapsto A\big(\gamma^{-1}(x - (1-\gamma)w)\big) + \gamma^{-1}(1-\gamma)(x - w)} \tag{11}$$

and

$$\boxed{B_\gamma \colon X \rightrightarrows X \colon x \mapsto B\big(\gamma^{-1}(x - (1-\gamma)w)\big) + \gamma^{-1}(1-\gamma)(x - w).} \tag{12}$$

Proposition 3.1 *The following hold true:*

(i) *A_γ, B_γ, and $A_\gamma + B_\gamma$ are maximally monotone.*

(ii) *A_γ, B_γ, and $A_\gamma + B_\gamma$ are strongly monotone, with constants $\gamma^{-1}(\sigma_A + 1 - \gamma)$, $\gamma^{-1}(\sigma_B + 1 - \gamma)$, and $\gamma^{-1}(\sigma_A + \sigma_B + 2 - 2\gamma)$, respectively.*

(iii) *$\operatorname{zer}(A_\gamma + B_\gamma)$ is nonempty and a singleton.*

Proof. (i): Clear. (ii): This follows from (6). (iii): Items (i) and (ii) imply that $A_\gamma + B_\gamma$ is maximally monotone and strongly monotone. Now apply [6, Corollary 23.37(ii)]. ■

In view of Proposition 3.1(iii), we denote the unique point in $\operatorname{zer}(A_\gamma + B_\gamma)$ by z_γ:

$$\boxed{\operatorname{zer}(A_\gamma + B_\gamma) = \{z_\gamma\}.} \tag{13}$$

We now essentially re-derive the central convergence result of Aragón–Artacho and Campoy [2, Theorem 3.1]:

Theorem 3.2 (AACA for fixed γ) *Given $x_0 \in X$ and $\lambda \in \,]0,1]$, define the sequence $(x_n)_{n\in\mathbb{N}}$ via*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = (1-\lambda)x_n + \lambda\big(2\gamma J_B + 2(1-\gamma)w - \operatorname{Id}\big) \circ \big(2\gamma J_A + 2(1-\gamma)w - \operatorname{Id}\big)x_n. \tag{14}$$

Then there exists $\bar{x} \in \operatorname{Fix}(R_{B_\gamma} R_{A_\gamma})$ such that $x_n \rightharpoonup \bar{x}$ and $\gamma J_A x_n + (1-\gamma)w \to z_\gamma$.

Proof. On the one hand, by Proposition 2.2,

$$J_{A_\gamma} = \gamma J_A + (1-\gamma)w \quad \text{and} \quad J_{B_\gamma} = \gamma J_B + (1-\gamma)w \tag{15}$$

which implies

$$R_{A_\gamma} = 2\gamma J_A + 2(1-\gamma)w - \operatorname{Id} \quad \text{and} \quad R_{B_\gamma} = 2\gamma J_B + 2(1-\gamma)w - \operatorname{Id} \tag{16}$$

and further

$$R_{B_\gamma} R_{A_\gamma} = \big(2\gamma J_B + 2(1-\gamma)w - \operatorname{Id}\big) \circ \big(2\gamma J_A + 2(1-\gamma)w - \operatorname{Id}\big) \tag{17}$$

On the other hand, both A_γ and B_γ are strongly monotone with constant $\gamma^{-1}(1-\gamma)$. Altogether, the result follows from Fact 2.1 applied to (A_γ, B_γ) instead of (A, B). ■

Remark 3.3 Several comments regarding Theorem 3.2 are in order.

(i) We have opted for a more explicit and thus easier-to-use version of AACA where the effect of w is explicitly recorded.

(ii) While one could make λ depending on n as in [2], we decided instead to stress the new case when $\lambda = 1$, corresponding to the Peaceman–Rachford version and notably absent in [2]. This case deserves interest because it turned out to be the best parameter choice in [5].

(iii) Our assumption of maximal monotonicity makes for a tidy theory. It is used chiefly to guarantee the existence of each z_γ; in [2], this is replaced by some condition regarding the existence of z_γ which seems to be not so easy to check in practice.

(iv) One may apply Theorem 3.2 in a standard product space setting to handle the sum of finitely many maximally monotone operators via AACA, as done in [2].

Of course, the remaining key question is:

What is the behaviour when $\gamma \to 1^-$ for AACA?

While this was answered in some form in [1] when A and B are normal cone operators, no result was offered in [2]. We conclude this paper by providing a complete and satisfying answer, relying on tools by Combettes and Hirstoaga [9] and [10], packed also into [6, Theorem 23.44].

Theorem 3.4 (dichotomy for AACA when $\gamma \to 1^-$) *Let z_γ be as in (13). Then exactly one of the following holds:*

(i) $\operatorname{zer}(A+B) \neq \varnothing$ *and* $z_\gamma \to P_{\operatorname{zer}(A+B)}w$ *as* $\gamma \to 1^-$.

(ii) $\operatorname{zer}(A+B) = \varnothing$ *and* $\|z_\gamma\| \to \infty$ *as* $\gamma \to 1^-$.

Proof. Set $\delta = 2(1-\gamma)$ and note that $\delta \to 0^+ \Leftrightarrow \gamma \to 1^-$. Moreover, set

$$y_\delta = \gamma^{-1}\big(z_\gamma - (1-\gamma)w\big). \tag{18}$$

We have, by definition of z_γ and y_δ,

$$0 \in (A_\gamma + B_\gamma)(z_\gamma) = (A+B)y_\delta + \delta(y_\delta - w). \tag{19}$$

Two cases are now conceivable.

Case 1: $\operatorname{zer}(A+B) \neq \varnothing$. By [6, Theorem 23.44(i)], we have

$$\lim_{\delta \to 0^+} y_\delta = P_{\operatorname{zer}(A+B)}w; \tag{20}$$

or equivalently, $\lim_{\gamma \to 1^-} z_\gamma = P_{\operatorname{zer}(A+B)}w$.

Case 2: $\operatorname{zer}(A+B) = \varnothing$. By [6, Theorem 23.44(ii)], we have

$$\lim_{\delta \to 0^+} \|y_\delta\| = +\infty; \tag{21}$$

or equivalently, $\lim_{\gamma \to 1^-} \|z_\gamma\| = +\infty$.

Altogether, the proof is complete. ■

Remark 3.5 Here are some comments on Theorem 3.4.

(i) The information presented in Theorem 3.4(ii) is new even when A and B are normal cone operators as in [1].

(ii) Computing $P_{\mathrm{zer}\,(A+B)}w$ via Theorem 3.4 is cumbersome and "doubly iterative": one must first employ an algorithm to find z_γ, and the let γ tend to 1^-. There are, however, some results that allow us to avoid this double iteration and instead solve the problem via a single iteration; see, e.g., the discussion in [4, Section 8].

Let us conclude with a simple example.

Example 3.6 Suppose that $A = P_U$, where U is a closed linear subspace of X, and $B \equiv -v$, where $v \in U^\perp$. Then $\mathrm{zer}\,(A+B) = U^\perp$, if $v = 0$; $\mathrm{zer}\,(A+B) = \varnothing$, if $v \neq 0$. Let $w = 0 \in X$, and let $\gamma \in \,]0,1[$. Then $(\forall x \in X)$ $A_\gamma x = \gamma^{-1}(P_U(x) + (1-\gamma)x)$ and $B_\gamma x = -v + \gamma^{-1}(1-\gamma)x$. Hence $\mathrm{zer}\,(A_\gamma + B_\gamma) = \{z_\gamma\}$, where $z_\gamma = (2(1-\gamma))^{-1}\gamma v$.

Case 1: $v = 0$. Then $z_\gamma \equiv 0 \to 0 = P_{\mathrm{zer}\,(A+B)}(w)$.

Case 2: $v \neq 0$. Then $\|z_\gamma\| = (2(1-\gamma))^{-1}\gamma\|v\| \to +\infty$.

Both cases illustrate Theorem 3.4.

Acknowledgments

SA was partially supported by Saudi Arabian Cultural Bureau. HHB and XW were partially supported by NSERC Discovery Grants. WMM was partially supported by NSERC Postdoctoral Fellowship.

References

[1] F. J. Aragón Artacho and R. Campoy, *A new projection method for finding the closest point in the intersection of convex sets*, Comput. Optim. Appl., 69 (2018), pp. 99–132.

[2] F. J. Aragón Artacho and R. Campoy, *Computing the resolvent of the sum of maximally monotone operators with the averaged alternating modified reflections algorithm*. https://arxiv.org/abs/1805.09720, 2018.

[3] S. Bartz, H. H. Bauschke, S. M. Moffat, and X. Wang, *The resolvent average of monotone operators: dominant and recessive properties*, SIAM J. Optim., 26 (2016), pp. 602–634.

[4] H. H. Bauschke, R. I. Boţ, W. L. Hare, and W. M. Moursi, *Attouch-Théra duality revisited: paramonotonicity and operator splitting*, J. Approx. Th., 164 (2012), pp. 1065–1084.

[5] H. H. Bauschke, R. S. Burachik, and C. Y. Kaya, *Constraint splitting and projection methods for optimal control of double integrator*. https://arxiv.org/abs/1804.03767, 2018.

[6] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, second ed., 2017.

[7] H. H. Bauschke, R. Goebel, Y. Lucet, and X. Wang, *The proximal average: basic theory*, SIAM J. Optim., 19 (2008), pp. 766–785.

[8] P. L. Combettes, *Iterative construction of the resolvent of a sum of maximal monotone operators*, J. Convex Anal., 16 (2009), pp. 727–748.

[9] P. L. Combettes and S. A. Hirstoaga, *Approximating curves for nonexpansive and monotone operators*, J. Convex Anal., 13 (2006), pp. 633–646.

[10] P. L. Combettes and S. A. Hirstoaga, *Visco-penalization of the sum of two monotone operators*, Nonlinear Anal., 69 (2008), pp. 579–591.