J. Math. Anal. Appl. 490 (2020) 124211

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Tight coefficients of averaged operators via scaled relative graph

Xinmeng Huang [a], Ernest K. Ryu [b,*], Wotao Yin [c]

[a] School of Mathematical Sciences, University of Science and Technology of China, China
[b] Department of Mathematical Sciences, Seoul National University, Republic of Korea
[c] Department of Mathematics, University of California, Los Angeles, United States of America

ARTICLE INFO

Article history:
Received 3 December 2019
Available online 12 May 2020
Submitted by P.L. Combettes

Keywords:
Averaged operator
Composition of operators
Nonexpansive operator
Euclidean geometry
Three operators

ABSTRACT

Many iterative methods in optimization are fixed-point iterations with averaged operators. As such methods converge at an $\mathcal{O}(1/k)$ rate with the constant determined by the averagedness coefficient, establishing small averagedness coefficients for operators is of broad interest. In this paper, we show that the averagedness coefficients of the composition of averaged operators by Ogura and Yamada (2002) [21] and the three-operator splitting by Davis and Yin (2017) [9] are tight. The analysis relies on the scaled relative graph, a geometric tool recently proposed by Ryu et al. (2019) [25].

© 2020 Elsevier Inc. All rights reserved.

1. Introduction

Since their introduction in [1], averaged operators have been widely used in the analysis of nonlinear fixed-point iterations. In particular, a wide range of optimization methods can be analyzed as fixed-point iterations with a composition of averaged operators, which are themselves averaged [6]. The smallest (best) averagedness coefficient for this setup was presented by Ogura and Yamada [21] and was introduced to the broader optimization community by Combettes and Yamada [7] and Bauschke and Combettes [3]. More recently, Davis and Yin presented a three-operator splitting method and established its convergence by showing the associated operator is averaged [9]. Whether these averagedness coefficients are tight, loosely defined as being unable to be improved without additional assumptions, was not known.

The *Scaled Relative Graph* (SRG) is a geometric tool for analyzing fixed-point iterations recently proposed by Ryu, Hannah, and Yin [25]. The SRG maps the action of a nonlinear operator to a subset the 2D plane, analogous to how the spectrum maps the action of a linear operator to the complex plane. A strength of the SRG is that it is well-suited for *tight* analysis.

* Corresponding author.
E-mail addresses: hxm8888@mail.ustc.edu.cn (X. Huang), ernestryu@snu.ac.kr (E.K. Ryu), wotaoyin@math.ucla.edu (W. Yin).

https://doi.org/10.1016/j.jmaa.2020.124211
0022-247X/© 2020 Elsevier Inc. All rights reserved.

2D geometric illustrations have been used by Eckstein and Bertsekas [10,11], Giselsson [16,15], Banjac and Goulart [2], and Giselsson and Moursi [17] to qualitatively understand convergence of optimization algorithms. The SRG is a rigorous formulation of such illustrations.

In this paper, we use the SRG to show tightness of the averagedness coefficients of the composition of averaged operators by Ogura and Yamada and the three-operator splitting by Davis and Yin. Section 2 discusses general preliminaries and sets up the notation. Section 3 presents results on the composition of averaged operators. Section 4 presents results on the Davis–Yin splitting.

1.1. Contribution and prior work

The contribution of this paper is in the results Corollaries 1 and 2, which establish tightness of the averagedness coefficients, and the geometric proof technique based on the SRG.

The geometric arguments of Section 4 are entirely new. The geometric arguments of Section 3 overlap with the classical work on "circular arithmetic" initiated by Gargantini and Henrici [14]. In [18,19], Hauenchild introduced the notion of "optimal circular multiplication", which considers the smallest circle enclosing the Minkowski product (defined in Section 2) of two disks on the complex plane. This is *not* the same as what we consider in Section 3, since we find the smallest circle under the additional requirement that it goes through the point $(1, 0)$. These two notions coincide sometimes, but not always. In [22], Polyak, Scherbakov, and Schmulyian perform calculations similar to that of Theorem 1 in the context of control theoretic stability analysis. In fact, Theorem 1 of [22] is, after a change of variables, the same as Theorem 1 of this work. However, the proof in [22] is not rigorous as it omits what we call Step 2 and Step 3 in our proof of Theorem 1. In [12,13], Farouki et al. also perform similar envelope calculations that are, after a change of variables, the same as that of Theorem 1 of this work. However, Farouki et al. also do not prove Steps 2 and 3; they merely state, without providing or outlining a proof, in Section 6.7 of [12] that "one can easily see" this fact. To summarize, in the proof of Theorem 1, Step 1 coincides with existing work, while Steps 2 and 3 are new. Furthermore, the proof of Corollary 1, which connects the geometric analysis to the composition of averaged operators using the SRG, is new.

2. Preliminaries

We follow the standard notation of [3,24]. Write $\mathcal{H}$ for a real Hilbert space equipped with the inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$. Write $A : \mathcal{H} \rightrightarrows \mathcal{H}$ to denote A is a multi-valued operator on $\mathcal{H}$. Write $I : \mathcal{H} \to \mathcal{H}$ for the identity operator. Define the resolvent of A as $J_A = (I + A)^{-1}$. We say $A \colon \mathcal{H} \rightrightarrows \mathcal{H}$ is monotone if

$$\langle u - v, x - y \rangle \geq 0, \qquad \forall x, y \in \mathcal{H},\ u \in Ax,\ v \in Ay.$$

Write $\mathcal{M}$ for the class of monotone operators. For $\beta \in (0, \infty)$, we say a single-valued operator $A \colon \mathcal{H} \to \mathcal{H}$ is β-cocoercive if

$$\langle Ax - Ay, x - y \rangle \geq \beta \|Ax - Ay\|^2,$$

for all $x, y \in \mathcal{H}$ and write $\mathcal{C}_\beta$ for the class of β-cocoercive operators. For $\theta \in (0, 1)$, we say an operator A is θ-averaged if $A = (1-\theta)I + \theta N$ for some nonexpansive operator N and write $\mathcal{N}_\theta$ for the class of θ-averaged operators.

We write complex numbers with the Cartesian and polar coordinate representations $z = x + yi$ and $z = re^{i\varphi} = r\cos(\varphi) + ir\sin(\varphi)$. For notational convenience, we often identify $\mathbb{C}$ with $\mathbb{R}^2$. We use Minkowski-type set notation with sets of complex numbers. In particular, given $\alpha \in \mathbb{C}$ and $Z, W \subseteq \mathbb{C}$, write

$$\alpha Z = \{\alpha z \,|\, z \in Z\}, \quad ZW = \{zw \,|\, z \in Z,\, w \in W\}.$$

The set ZW is called the *Minkowski product* of Z and W. Given a set U, write ∂U to denote its boundary and write $U^\circ = U \backslash \partial U$ to denotes its interior.

2.1. Scaled relative graph

We follow the notation of [25]. The scaled relative graph (SRG) of an operator A is defined as

$$\mathcal{G}(A) = \left\{ \frac{\|u-v\|}{\|x-y\|} \exp\left[\pm i \angle(u-v, x-y)\right] \,\middle|\, u \in Ax,\, v \in Ay,\, x \neq y \right\} \left(\cup \{\infty\} \text{ if } A \text{ is multi-valued} \right),$$

where

$$\angle(a,b) = \begin{cases} \arccos(\frac{\langle a,b\rangle}{\|a\|\|b\|}) & \text{if } a \neq 0,\, b \neq 0 \\ 0 & \text{otherwise,} \end{cases}$$

denotes the angle between $a, b \in \mathcal{H}$. The SRG $\mathcal{G}(A)$ maps the action of the operator A to points onto the extended complex plane. The magnitude of each element of $\mathcal{G}(A)$, $\frac{\|u-v\|}{\|x-y\|}$, represents the size of the change in outputs u, v relative to the size of the change in inputs x, y. The angle, $\angle(u-v, x-y)$, represents how much the change in outputs is aligned with the change in inputs. The SRG of the class of operators $\mathcal{A}$ is defined as

$$\mathcal{G}(\mathcal{A}) = \bigcup_{A \in \mathcal{A}} \mathcal{G}(A).$$

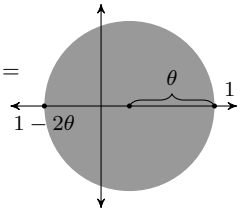
For $\theta \in (0,1)$, define

$$\text{Disk}(\theta) = \{z \in \mathbb{C} \,\big|\, |z - (1-\theta)| \leq \theta\}, \quad \text{Circ}(\theta) = \{z \in \mathbb{C} \,\big|\, |z - (1-\theta)| = \theta\}.$$

The sets have $(1-\theta)$ as their center and include 1 as the right-most point.

Fact 1 *(Proposition 3.3 of [25]). Let $\theta \in (0,1)$. Then*

$\mathcal{G}(\mathcal{N}_\theta) = \text{Disk}(\theta) =$



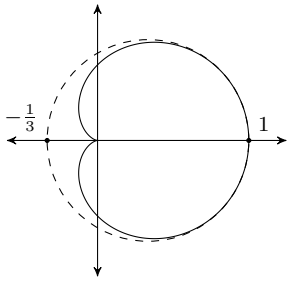
Fact 2 *(Theorem 3.5 of [25]). For the operator class $\mathcal{N}_\theta$, where $\theta \in (0,1)$, inclusion within the operator class is equivalent to the inclusion of the SRG in the 2D plane.*

2.2. Osculating circle, curvature, and envelope

In differential geometry of curves, the *osculating circle* of a sufficiently smooth plane curve C at a point P on the curve is the circle passing through P that approximates C most tightly within infinitesimal neighborhoods of P. The center of the circle lies on the inner normal line, and the reciprocal of its radius is the *curvature* of C at P [8]. For curves defined through polar coordinates as $r(\varphi)$, the curvature $\kappa(\varphi)$ at $r(\varphi)$ is given by [23]:

$$\kappa(\varphi) = \frac{r(\varphi)^2 + 2(\frac{dr}{d\varphi})^2 - r(\varphi)\frac{d^2r}{d\varphi^2}}{(r(\varphi)^2 + (\frac{dr}{d\varphi})^2)^{\frac{3}{2}}} \tag{1}$$

The osculating circle of C at P provides insight on the smallest circle through P enclosing C.



$C = \{re^{i\varphi} \,|\, r = \frac{1}{2}(1 + \cos(\varphi))\}$ and its osculating circle at 1

An *envelope* of a family of curves in the plane is a curve that is tangent to each member of the family at some point. Formally, let each $\{C_t\}_{t\in\mathbb{R}}$ be a parameterized family of curves in $\mathbb{R}^2$ defined by $F(t, \mathbf{x}) = 0$, where $t \in \mathbb{R}$ is the parameter, $\mathbf{x} \in \mathbb{R}^2$, and F is smooth. That is, $C_t = \{\mathbf{x} \,|\, F(t, \mathbf{x}) = 0\}$. The envelope of $\{C_t\}_{t\in\mathbb{R}}$ is defined as the set points satisfying

$$F(t, \mathbf{x}) = 0, \qquad \frac{\partial F}{\partial t}(t, \mathbf{x}) = 0 \tag{2}$$

for some t. The envelope includes the boundary of the region filled by the curves [5, Section 5.17]. See [4,5] for further discussion.

3. Tight characterization of the composition of averaged operators

The composition of two averaged operators is itself an averaged operator, and Ogura and Yamada [21] showed the best known averagedness coefficient for this setup. In this section, we provide an alternate geometric proof of this result and establish its tightness.

Again, $\mathcal{N}_{\theta_1}$ and $\mathcal{N}_{\theta_2}$ are the classes of θ_1- and θ_2-averaged operators. Define $\mathcal{N}_{\theta_1}\mathcal{N}_{\theta_2} = \{N_1N_2 \,|\, N_1 \in \mathcal{N}_{\theta_1},\, N_2 \in \mathcal{N}_{\theta_2}\}$ to be the class of compositions of θ_1- and θ_2-averaged operators.

Theorem 1. *Let $\theta_1, \theta_2 \in (0, 1)$. Then $\mathcal{G}(\mathcal{N}_{\theta_1}\mathcal{N}_{\theta_2})$ is the region enclosed by the outer curve defined by*

$$r(\varphi)^2 - 2r(\varphi)(\cos(\varphi)(1 - \theta_1)(1 - \theta_2) + \theta_1\theta_2) + (1 - 2\theta_1)(1 - 2\theta_2) = 0. \tag{3}$$

Fig. 1 illustrates Theorem 1. To clarify, the equation of Theorem 1 defines at most two non-intersecting closed curves, one enclosing the other, and the SRG is given by the outer curve.

Corollary 1. *Let $\theta_1, \theta_2 \in (0, 1)$. Then $\mathcal{N}_{\theta_1}\mathcal{N}_{\theta_2} \subseteq \mathcal{N}_\theta$ with*

$$\theta = \frac{\theta_1 + \theta_2 - 2\theta_1\theta_2}{1 - \theta_1\theta_2}.$$

The averagedness coefficient θ is tight in the sense that it cannot be reduced without further assumptions.

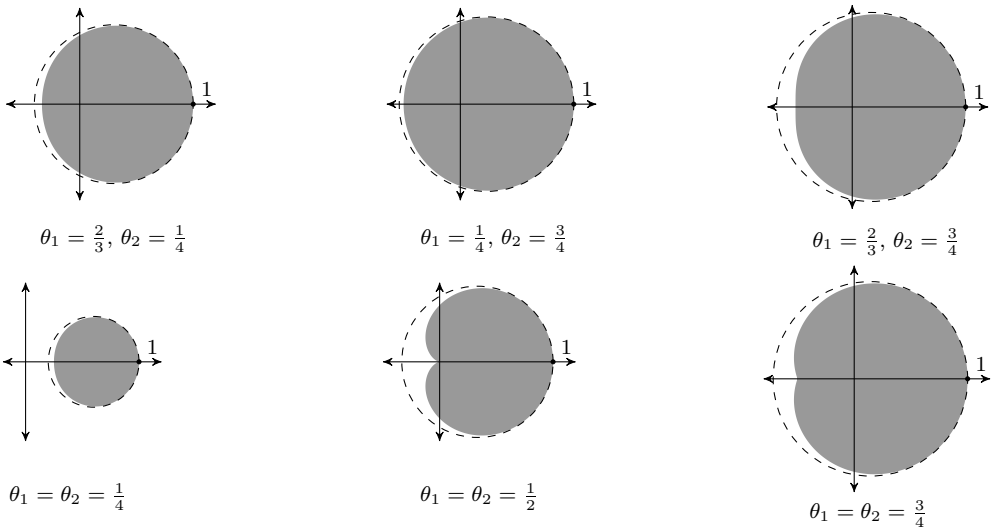


Fig. 1. The shaded regions illustrate $\mathcal{G}(\mathcal{N}_{\theta_1}\mathcal{N}_{\theta_2})$ given by Theorem 1. The circles drawn in dashed lines illustrate $\mathcal{G}(\mathcal{N}_\theta)$ given by Corollary 1.

3.1. Proof of Theorem 1

Define S to be the outer curve defined by (3) and S_{enc} to be the region enclosed by S. By Theorem 4.5 of [25] and the arc property of the averaged operators class, we have

$$\mathcal{G}(\mathcal{N}_{\theta_1}\mathcal{N}_{\theta_2}) = \text{Disk}(\theta_1)\text{Disk}(\theta_2).$$

Therefore, it remains to show $\text{Disk}(\theta_1)\text{Disk}(\theta_2) = S_{\text{enc}}$ on the complex plane. The proof is completed in 3 steps. In Step 1, we show that S is the boundary enclosing $\text{Circ}(\theta_1)\text{Circ}(\theta_2)$ with geometric arguments. In Step 2, we show that S furthermore encloses $\text{Disk}(\theta_1)\text{Disk}(\theta_2)$, i.e. we show $\text{Disk}(\theta_1)\text{Disk}(\theta_2) \subseteq S_{\text{enc}}$. In Step 3, we show $\text{Disk}(\theta_1)\text{Disk}(\theta_2) = S_{\text{enc}}$ with a topological argument.

Step 1. The curve $\text{Circ}(\theta_1)$ is defined by $f_1(z) = 0$ with

$$f_1(z) = (x - (1 - \theta_1))^2 + y^2 - \theta_1^2,$$

where $z = x + yi$. Let

$$z_2(t) = 1 - \theta_2 + \theta_2\cos(t) + \theta_2\sin(t)i$$

be a parameterization of $\text{Circ}(\theta_2)$.

Scaling and rotating $\text{Circ}(\theta_1)$ by $z_2(t) \in \text{Circ}(\theta_2)$ yields the curve defined by

$$\begin{aligned} 0 =& f_1(z/z_2(t)) \\ =& \left(\frac{x(1 - \theta_2 + \theta_2\cos(t)) + y\theta_2\sin(t)}{|z_2(t)|^2} - (1 - \theta_1)\right)^2 \\ &+ \left(\frac{y(1 - \theta_2 + \theta_2\cos(t)) - x\theta_2\sin(t)}{|z_2(t)|^2}\right)^2 - \theta_1^2. \end{aligned}$$

Multiply both sides of the equation by $|z_2(t)|^4$ and simplify to get

$$(x-(1-\theta_2+\theta_2\cos(t))(1-\theta_1))^2+(y+\theta_2\sin(t)(1-\theta_1))^2$$
$$-(2\theta_2^2-2\theta_2+1+2\theta_2(1-\theta_2)\cos(t))\theta_1^2=0.$$

Apply the envelope formula (2) to eliminate t and obtain the envelope

$$\begin{aligned} f_2(z) &= (x^2+y^2-2x(1-\theta_1)(1-\theta_2)+(1-2\theta_1)(1-2\theta_2))^2-4\theta_1^2\theta_2^2(x^2+y^2)\\ &= 0. \end{aligned} \tag{4}$$

Using polar coordinates with $r=\sqrt{x^2+y^2}\geq 0$ and $x=r\cos(\varphi)$, we can factor (4) as

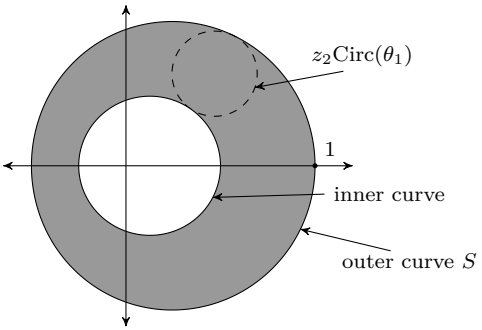
$$\begin{aligned} 0 =&\big(r^2-2r\cos(\varphi)(1-\theta_1)(1-\theta_2)+(1-2\theta_1)(1-2\theta_2)-2\theta_1\theta_2 r\big)\\ &\cdot\big(r^2-2r\cos(\varphi)(1-\theta_1)(1-\theta_2)+(1-2\theta_1)(1-2\theta_2)+2\theta_1\theta_2 r\big), \end{aligned}$$

where $r\geq 0$ and $\varphi\in\mathbb{R}$. By considering the substitution $r\mapsto -r$ and $\varphi\mapsto\varphi+\pi$, we can combine the two factors into one to get (3):

$$r^2-2r\cos(\varphi)(1-\theta_1)(1-\theta_2)+(1-2\theta_1)(1-2\theta_2)-2\theta_1\theta_2 r=0,$$

where $r\in\mathbb{R}$ and $\varphi\in\mathbb{R}$. To clarify, the combined equation allows negative r. The envelope contains the boundary of $\mathrm{Circ}(\theta_1)\mathrm{Circ}(\theta_2)$.

The curve defined by (3) is an instance of the *Cartesian oval*, which contains at most two closed curves one enclosing the other [20]. The following figure illustrates the envelope in solid lines and $\mathrm{Circ}(\theta_1)\mathrm{Circ}(\theta_2)$ as the shaded region. The outer curve S encloses $\mathrm{Circ}(\theta_1)\mathrm{Circ}(\theta_2)$, i.e., $\mathrm{Circ}(\theta_1)\mathrm{Circ}(\theta_2)\subseteq S_{\mathrm{enc}}$.



Step 2. We now show that S encloses not only $\mathrm{Circ}(\theta_1)\mathrm{Circ}(\theta_2)$ but also $\mathrm{Disk}(\theta_1)\mathrm{Disk}(\theta_2)$. Note $\mathrm{Disk}(\theta_1)\mathrm{Disk}(\theta_2)$ is compact as it is the image of a compact set under a continuous map. On the other hand, $\mathrm{Disk}(\theta_1)^\circ\mathrm{Disk}(\theta_2)$ and $\mathrm{Disk}(\theta_1)\mathrm{Disk}(\theta_2)^\circ$ are open as they are unions of open sets. Since

$$\mathrm{Disk}(\theta_1)^\circ\mathrm{Disk}(\theta_2)\cup\mathrm{Disk}(\theta_1)\mathrm{Disk}(\theta_2)^\circ$$

is open, we have

$$\begin{aligned} \partial\left(\mathrm{Disk}(\theta_1)\mathrm{Disk}(\theta_2)\right) &\subseteq \mathrm{Disk}(\theta_1)\mathrm{Disk}(\theta_2)\backslash\left(\mathrm{Disk}(\theta_1)^\circ\mathrm{Disk}(\theta_2)\cup\mathrm{Disk}(\theta_1)\mathrm{Disk}(\theta_2)^\circ\right)\\ &\subseteq \mathrm{Circ}(\theta_1)\mathrm{Circ}(\theta_2). \end{aligned}$$

Since S encloses $\mathrm{Circ}(\theta_1)\mathrm{Circ}(\theta_2)$ which contains the boundary of the compact set $\mathrm{Disk}(\theta_1)\mathrm{Disk}(\theta_2)$, S encloses $\mathrm{Disk}(\theta_1)\mathrm{Disk}(\theta_2)$.

Step 3. We have shown $\text{Disk}(\theta_1)\text{Disk}(\theta_2) \subseteq S_{\text{enc}}$, and it remains to show $\text{Disk}(\theta_1)\text{Disk}(\theta_2) = S_{\text{enc}}$. The question is whether $\text{Disk}(\theta_1)\text{Disk}(\theta_2)$ is simply connected, i.e., whether it contains any "holes". As the previous figure illustrates, $\text{Circ}(\theta_1)\text{Circ}(\theta_2)$ contain holes. We show $\text{Disk}(\theta_1)\text{Disk}(\theta_2)$ does not.

Define the map

$$\begin{aligned} \Pi\colon \text{Disk}(\theta_1) \times \text{Disk}(\theta_2) &\to \text{Disk}(\theta_1)\text{Disk}(\theta_2) \\ (z_1, z_2) &\mapsto z_1 z_2. \end{aligned}$$

To clarify, $\text{Disk}(\theta_1) \times \text{Disk}(\theta_2)$ denotes the product set while $\text{Disk}(\theta_1)\text{Disk}(\theta_2)$ denotes the Minkowski product. We have shown that there is a parameterized closed curve

$$\{(z_1(t), z_2(t))\}_{t\in[0,1]} \subseteq \text{Disk}(\theta_1) \times \text{Disk}(\theta_2),$$

such that $\{\Pi(z_1(t), z_2(t))\}_{t\in[0,1]} = S$. Assume for contradiction that $z \in S_{\text{enc}}$ but $z \notin \text{Disk}(\theta_1)\text{Disk}(\theta_2)$. (In other words, we assume for contradiction that z is strictly within the hole of the domain.) Since $\text{Disk}(\theta_1) \times \text{Disk}(\theta_2)$ is simply connected, we can continuously contract $\{(z_1(t), z_2(t))\}_{t\in[0,1]}$ to a point in $\text{Disk}(\theta_1) \times \text{Disk}(\theta_2)$, and the curve under the map Π continuously contracts to a point in $\text{Disk}(\theta_1)\text{Disk}(\theta_2)$. However, this is not possible as $\{\Pi(z_1(t), z_2(t))\}_{t\in[0,1]}$ has a nonzero winding number around z and $z \notin \text{Disk}(\theta_1)\text{Disk}(\theta_2)$. We have a contradiction and we conclude $z \in \text{Disk}(\theta_1)\text{Disk}(\theta_2)$. □

3.2. Proof of Corollary 1

We can visually observe from Fig. 1 that $\text{Circ}(\theta)$, the dashed circle, enclose $\mathcal{G}(\mathcal{N}_{\theta_1}\mathcal{N}_{\theta_2})$. We can also observe that the geometric objects have matching curvature at point 1, and therefore we cannot further reduce the size of the dashed circle while enclosing $\mathcal{G}(\mathcal{N}_{\theta_1}\mathcal{N}_{\theta_2})$. We now make this argument formal with Fact 2 and the following geometric arguments.

Remember, $f_2(x, y) = 0$ defines the boundary $\partial(\text{Disk}(\theta_1)\text{Disk}(\theta_2))$. Define

$$\begin{aligned} g(\varphi) &= f_2(\theta\cos(\varphi) + (1-\theta), \theta\sin(\varphi)) \\ &= \frac{16\theta_1^2\theta_2^2(1-\theta_1)^2(1-\theta_2)^2(\theta_1+\theta_2-2\theta_1\theta_2)^2\sin^4(\varphi/2)}{(1-\theta_1\theta_2)^4}, \end{aligned}$$

i.e., $g(\varphi)$ is f_2 evaluated on the curve $\text{Circ}(\theta)$. We can see that $g(\varphi) > 0$ for all $\varphi \neq 0$ and $g(0) = 0$. This implies $\text{Circ}(\theta)$ and $\partial(\text{Disk}(\theta_1)\text{Disk}(\theta_2))$ intersect at only one point and therefore do not cross. The point $(1-\varepsilon, 0)$ is in $\text{Disk}(\theta_1)\text{Disk}(\theta_2)$ and enclosed by $\text{Circ}(\theta)$ for small enough $\varepsilon > 0$. Since

$$f_2(1-\varepsilon, 0) = -8\theta_1\theta_2 \underbrace{(\theta_1+\theta_2-2\theta_1\theta_2)}_{>0}\varepsilon + \mathcal{O}(\varepsilon^2)$$

for $\varepsilon \to 0$, it is $\text{Circ}(\theta)$ that encloses $\partial(\text{Disk}(\theta_1)\text{Disk}(\theta_2))$. Finally, we conclude $\text{Disk}(\theta)$ contains $\text{Disk}(\theta_1)\text{Disk}(\theta_2)$.

Consider $r(\varphi)$ defined by (3). Through implicit differentiation, we get

$$\frac{dr}{d\varphi}\Big|_{\varphi=0} = 0, \qquad \frac{d^2r}{d\varphi^2}\Big|_{\varphi=0} = \frac{(1-\theta_1)(1-\theta_2)}{\theta_1+\theta_2-2\theta_1\theta_2}.$$

Using (1), the curvature of $r(\varphi)$ at point 1 (given by $\varphi = 0$) is

$$\kappa(\varphi)\big|_{\varphi=0} = \frac{r(\varphi)^2 + 2(\frac{dr}{d\varphi})^2 - r(\varphi)\frac{d^2r}{d\varphi^2}}{(r(\varphi)^2 + (\frac{dr}{d\varphi})^2)^{\frac{3}{2}}}\Big|_{\varphi=0} = \frac{1-\theta_1\theta_2}{\theta_1+\theta_2-2\theta_1\theta_2} = \frac{1}{\theta}.$$

This implies any circle through 1 symmetric about the real axis containing $r(\varphi)$ must have radius at least θ. □

4. Tight characterization of Davis–Yin splitting

Consider the monotone inclusion problem

$$\operatorname*{find}_{x\in\mathcal{H}} \quad 0 \in (A+B+C)x,$$

where A and B are maximal monotone and C is β-cocoercive. Davis and Yin [9] proposed

$$T_\gamma(A,B,C) = I - J_{\gamma B} + J_{\gamma A}(2J_{\gamma B} - I - \gamma C J_{\gamma B}),$$

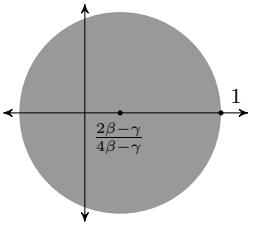
which we call the *Davis–Yin splitting* (DYS), and showed that it is a fixed-point encoding for the monotone inclusion problem in the sense that $\operatorname{Zer}(A+B+C) = J_{\gamma B}(\operatorname{Fix}(T_\gamma(A,B,C)))$, where Zer and Fix respectively denote the set of zeros and fixed points. Define the class of DYS operators as

$$\mathcal{T}_{\beta,\gamma} = \left\{ T_\gamma(A,B,C) \,\middle|\, A,B\in\mathcal{M},\, C\in\mathcal{C}_\beta \right\}.$$

Davis and Yin showed that the DYS operators of $\mathcal{T}_{\beta,\gamma}$ are $\frac{2\beta}{4\beta-\gamma}$-averaged.

Fact 3 *(Proposition 2.1 of [9]). Let $\gamma\in(0,2\beta)$. Then*

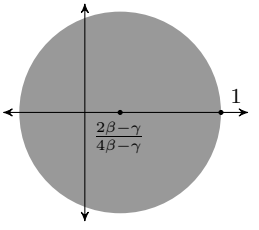
$$\mathcal{G}(\mathcal{T}_{\beta,\gamma}) \subseteq \mathcal{G}\left(\mathcal{N}_{\frac{2\beta}{4\beta-\gamma}}\right) =$$



We show that this characterization is tight in the following sense.

Theorem 2. *Let $\gamma\in(0,2\beta)$. Then*

$$\mathcal{G}(\mathcal{T}_{\beta,\gamma}) = \mathcal{G}\left(\mathcal{N}_{\frac{2\beta}{4\beta-\gamma}}\right) =$$



Corollary 2. *The averagedness parameter of Fact 3 is tight in the sense that it cannot be improved without further assumptions.*

4.1. Proof of Theorem 2

Since $\mathcal{G}(\mathcal{T}_{\beta,\gamma}) \subseteq \operatorname{Disk}\left(\frac{2\beta}{4\beta-\gamma}\right)$ by Fact 3, we show $\operatorname{Disk}\left(\frac{2\beta}{4\beta-\gamma}\right) \subseteq \mathcal{G}(\mathcal{T}_{\beta,\gamma})$. Define the set

$$S_{\beta,\gamma} = \left\{ 1 - z_2 + z_1(2z_2 - 1 - \gamma z_3 z_2) \,\middle|\, z_1,z_2\in\operatorname{Disk}(1/2),\, z_3\in\tfrac{1}{\beta}\operatorname{Disk}(1/2) \right\}.$$

The proof is completed in three steps. In Step 1, we show $S_{\beta,\gamma} \subseteq \mathcal{G}(\mathcal{T}_{\beta,\gamma})$ by appealing to results about the SRG. In Step 2, we show $\text{Circ}\left(\frac{2\beta}{4\beta-\gamma}\right) \subseteq S_{\beta,\gamma}$ with geometric arguments. In Step 3, we strengthen the result of Step 2 to $\text{Disk}\left(\frac{2\beta}{4\beta-\gamma}\right) \subseteq S_{\beta,\gamma}$ using a topological argument.

Step 1. By Lemma 3.2, Proposition 3.3, and Theorems 4.2 and 4.3 of [25], we can identify $z_1, z_2 \in \text{Disk}(1/2)$ with resolvents of maximal monotone operators on $\mathbb{R}^2$ and $z_3 \in \frac{1}{\beta}\text{Disk}(1/2)$ with a β-cocoercive operator on $\mathbb{R}^2$. Therefore, $S_{\beta,\gamma}$ represents the SRGs of operators in $\mathcal{T}_{\beta,\gamma}$, and we conclude $S_{\beta,\gamma} \subseteq \mathcal{G}(\mathcal{T}_{\beta,\gamma})$.

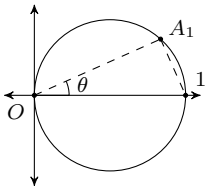
Step 2. Define

$$R_{\beta,\gamma} = \left\{ 1 - z_2 + z_1(2z_2 - 1 - \gamma z_3 z_2) \,\middle|\, z_1 = z_2 \in \text{Disk}(1/2),\ z_3 \in \tfrac{1}{\beta}\text{Disk}(1/2) \right\}.$$

Clearly $R_{\beta,\gamma} \subseteq S_{\beta,\gamma}$. We show

$$\text{Circ}\left(\frac{2\beta}{4\beta-\gamma}\right) \subseteq R_{\beta,\gamma}.$$

Let $A_1 = z_1 = z_2 = \cos(\theta)e^{i\theta} \in \text{Circ}(1/2)$.



With direct calculations, we have

$$A_2 = 2z_1z_2 - z_1 - z_2 + 1 = \cos(2\theta)e^{2\theta i} \in \text{Circ}(1/2)$$

and

$$A_3 = \frac{\gamma}{\beta} z_1 z_2 = \frac{\gamma}{\beta}\cos^2(\theta)e^{2\theta i}.$$

Define $O_1 = \frac{A_3}{2}$. Fig. 2 illustrates the following construction. Define point $P = \frac{2\beta-\gamma}{4\beta-\gamma}$ as the center of the circle $\text{Circ}\left(\frac{2\beta}{4\beta-\gamma}\right)$. Let $O_2 = A_2 - \frac{A_3}{2} = A_2 - O_1$ be the center of the disk $A_2 - A_3\text{Disk}(1/2)$. Let B be the point farthest from P in the disk $A_2 - A_3\text{Disk}(1/2)$. ($B \in R_{\beta,\gamma}$ since $A_2 - A_3\text{Disk}(1/2) \subseteq R_{\beta,\gamma}$.) Then P, O_2, and B are collinear. We have

$$\overline{O_2B} = \frac{\gamma}{2\beta}\cos^2(\theta).$$

Using the cosine rule, we have

$$\begin{aligned}
\overline{PO_2}^2 &= \overline{O_2O}^2 + \overline{OP}^2 - 2\overline{O_2O}\cdot\overline{OP}\cos(2\theta) \\
&= \left(\cos(2\theta) - \frac{\gamma}{2\beta}\cos^2(\theta)\right)^2 + \left(\frac{2\beta-\gamma}{4\beta-\gamma}\right)^2 - 2\left(\cos(2\theta) - \frac{\gamma}{2\beta}\cos^2(\theta)\right)\left(\frac{2\beta-\gamma}{4\beta-\gamma}\right)\cos(2\theta) \\
&= \left(\frac{\gamma}{2\beta}\cos^2(\theta) - \frac{2\beta}{4\beta-\gamma}\right)^2.
\end{aligned}$$

Since P, O_2, and B are collinear, we have

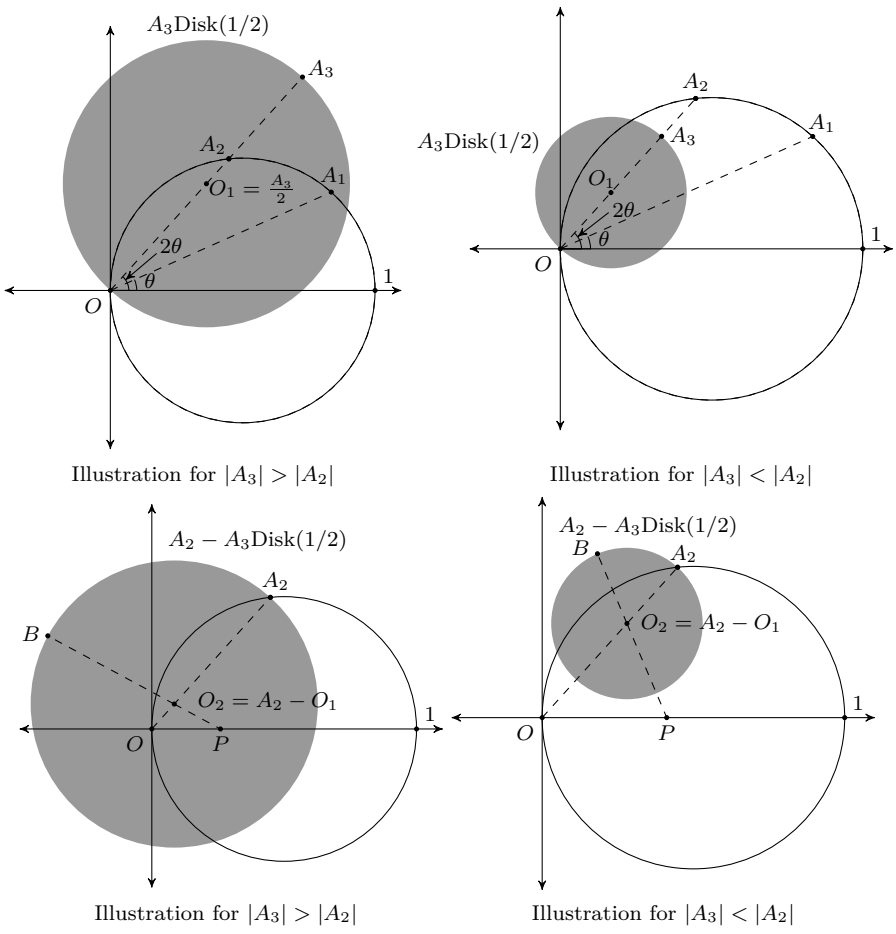


Fig. 2. The first geometric construction of Step 2. The shaded region is a subset of $R_{\beta,\gamma}$ by construction.

$$\overline{PB} = \overline{PO_2} + \overline{O_2B} = \frac{2\beta}{4\beta - \gamma}.$$

Therefore, $B \in \text{Circ}\left(\frac{2\beta}{4\beta-\gamma}\right)$.

Fig. 3 illustrates the following construction. The trajectory of $O_2 = (\cos(2\theta) - \frac{\gamma}{2\beta}\cos^2(\theta))e^{2\theta i}$ as a function of θ is a closed curve within $\text{Circ}\left(\frac{2\beta}{4\beta-\gamma}\right)$. Since $0 < 1 - \frac{2\beta}{4\beta-\gamma} < 1 - \frac{\gamma}{2\beta}$, the curve strictly encloses P. As θ traverses $[-\pi/2, \pi/2)$, O_2 traverses the inner curve and B traverses all of $\text{Circ}\left(\frac{2\beta}{4\beta-\gamma}\right)$. Therefore, we conclude $\text{Circ}\left(\frac{2\beta}{4\beta-\gamma}\right) \subseteq S_{\beta,\gamma}$.

Step 3. Define the map

$$\Pi\colon \text{Disk}(1/2) \times \text{Disk}(1/2) \times \tfrac{1}{\beta}\text{Disk}(1/2) \to S_{\beta,\gamma}$$

$$(z_1, z_2, z_3) \mapsto 1 - z_2 + z_1(2z_2 - 1 - \gamma z_3 z_2).$$

Consider any z strictly enclosed within $\text{Circ}\left(\frac{2\beta}{4\beta-\gamma}\right)$, and assume for contradiction that $z \notin S_{\beta,\gamma}$. We have shown that there is a closed curve

$$\{\eta(t)\}_{t\in[0,1]} \subseteq \text{Disk}(1/2) \times \text{Disk}(1/2) \times \tfrac{1}{\beta}\text{Disk}(1/2)$$

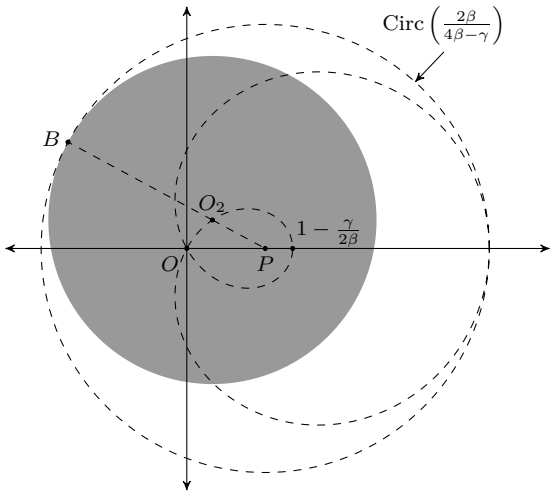


Fig. 3. The second geometric construction of Step 2.

such that $\{\Pi(\eta(t))\}_{t\in[0,1]}$ is Circ $\left(\frac{2\beta}{4\beta-\gamma}\right)$. The closed curve $\{\Pi(\eta(t))\}_{t\in[0,1]}$ strictly encloses z. Since $\mathrm{Disk}(1/2)\times\mathrm{Disk}(1/2)\times\frac{1}{\beta}\mathrm{Disk}(1/2)$ is simply connected, we can continuously contract $\{\eta(t)\}_{t\in[0,1]}$ to a point in $\mathrm{Disk}(1/2)\times\mathrm{Disk}(1/2)\times\frac{1}{\beta}\mathrm{Disk}(1/2)$, and $\{\Pi(\eta(t))\}_{t\in[0,1]}$ continuously contracts to a point in $S_{\beta,\gamma}$. However, this is not possible as $\{\Pi(\eta(t))\}_{t\in[0,1]}$ has a nonzero winding number around z and $z\notin S_{\beta,\gamma}$. We have a contradiction and we conclude $z\in S_{\beta,\gamma}$ and Disk $\left(\frac{2\beta}{4\beta-\gamma}\right)\subseteq S_{\beta,\gamma}$. □

Acknowledgments

This work was supported by NSF DMS-1720237 and ONR N000141712162.

References

[1] J.B. Baillon, R.E. Bruck, S. Reich, On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces, Houst. J. Math. 4 (1) (1978) 1–9.
[2] G. Banjac, P.J. Goulart, Tight global linear convergence rate bounds for operator splitting methods, IEEE Trans. Autom. Control 63 (12) (2018) 4126–4139.
[3] H.H. Bauschke, P.L. Combettes, Convex Analysis and Monotone Operator Theory in Hilbert Spaces, 2nd edition, Springer-Verlag, 2017.
[4] V. Boltyanskii, Envelopes, Macmillan, New York, 1964, p. 1.
[5] J.W. Bruce, P.J. Giblin, Curves and Singularities: A Geometrical Introduction to Singularity Theory, Cambridge University Press, 1992.
[6] P.L. Combettes, Solving monotone inclusions via compositions of nonexpansive averaged operators, Optimization 53 (5–6) (2004) 475–504.
[7] P.L. Combettes, I. Yamada, Compositions and convex combinations of averaged nonexpansive operators, J. Math. Anal. Appl. 425 (1) (2015) 55–70.
[8] J.L. Coolidge, The unsatisfactory story of curvature, Am. Math. Mon. 59 (6) (1952) 375–379.
[9] D. Davis, W. Yin, A three-operator splitting scheme and its optimization applications, Set-Valued Var. Anal. 25 (4) (2017) 829–858.
[10] J. Eckstein, Splitting methods for monotone operators with applications to parallel optimization, Ph.D. thesis, MIT, 1989.
[11] J. Eckstein, D.P. Bertsekas, On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators, Math. Program. 55 (1–3) (1992).
[12] R.T. Farouki, H.P. Moon, B. Ravani, Minkowski geometric algebra of complex sets, Geom. Dedic. 85 (1) (2001) 283–315.
[13] R.T. Farouki, H. Pottmann, Exact Minkowski products of N complex disks, Reliab. Comput. 8 (1) (2002) 43–66.
[14] I. Gargantini, P. Henrici, Circular arithmetic and the determination of polynomial zeros, Numer. Math. 18 (4) (1971) 305–320.
[15] P. Giselsson, Lunds universitet, lecture notes: large-scale convex optimization, http://www.control.lth.se/education/doctorate-program/large-scale-convex-optimization/, 2015. Last visited on 2018/12/01.
[16] P. Giselsson, S. Boyd, Linear convergence and metric selection for Douglas–Rachford splitting and ADMM, IEEE Trans. Autom. Control 62 (2) (2017) 532–544.

[17] P. Giselsson, W.M. Moursi, On compositions of special cases of Lipschitz continuous operators, preprint, arXiv:1912.13165, 2019.
[18] M. Hauenschild, Arithmetiken für komplexe kreise, Computing 13 (3) (1974) 299–312.
[19] M. Hauenschild, Extended circular arithmetic, problems and results, in: Interval Mathematics, Academic Press, New York, 1980, pp. 367–376.
[20] J.D. Lawrence, A Catalog of Special Plane Curves, Dover, New York, 1972.
[21] N. Ogura, I. Yamada, Non-strictly convex minimization over the fixed point set of an asymptotically shrinking nonexpansive mapping, Numer. Funct. Anal. Optim. 23 (1–2) (2002) 113–137.
[22] B.T. Polyak, P.S. Scherbakov, S.B. Shmulyian, Construction of value set for robustness analysis via circular arithmetic, Int. J. Robust Nonlinear Control 4 (3) (1994) 371–385.
[23] A.N. Pressley, Elementary Differential Geometry, Springer Science & Business Media, 2010.
[24] E.K. Ryu, S. Boyd, Primer on monotone operator methods, Appl. Comput. Math. 15 (1) (2016) 3–43.
[25] E.K. Ryu, R. Hannah, W. Yin, Scaled relative graph: nonexpansive operators via 2D Euclidean geometry, preprint, arXiv:1902.09788, 2019.