On the convergence of discrete-time linear systems

A linear time-varying Mann iteration converges IFF its operator is strictly pseudocontractive

Belgioioso, G.; Fabiani, Filippo; Blanchini, Franco; Grammatico, Sergio

DOI
10.1109/LCSYS.2018.2840882

Publication date
2018

Document Version
Accepted author manuscript

Published in
IEEE Control Systems Letters

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2018.2840882, IEEE Control Systems Letters

On the convergence of discrete-time linear systems: A linear time-varying Mann iteration converges iff its operator is strictly pseudocontractive

Giuseppe Belgioioso, Filippo Fabiani, Franco Blanchini and Sergio Grammatico

***Abstract*—We adopt an operator-theoretic perspective to study convergence of linear fixed-point iterations and discrete-time linear systems. We mainly focus on the so-called Krasnoselskij–Mann iteration, $x(k+1) = (1-\alpha_k)x(k) + \alpha_k A x(k)$, which is relevant for distributed computation in optimization and game theory, when A is not available in a centralized way. We show that strict pseudocontractiveness of the linear operator $x \mapsto Ax$ is not only sufficient (as known) but also necessary for the convergence to a vector in the kernel of $I - A$. We also characterize some relevant operator-theoretic properties of linear operators via eigenvalue location and linear matrix inequalities. We apply the convergence conditions to multi-agent linear systems with vanishing step sizes, in particular, to linear consensus dynamics and equilibrium seeking in monotone linear-quadratic games.**

***Index Terms*—Linear systems, LMIs, Game theory, Time-varying systems.**

I. Introduction

State convergence is the quintessential problem in multi-agent systems. In fact, multi-agent consensus and cooperation, distributed optimization and multi-player game theory revolve around the convergence of the state variables to an equilibrium, typically unknown a-priori. In distributed consensus problems, agents interact with their neighboring peers to collectively achieve global agreement on some value [1]. In distributed optimization, decision makers cooperate locally to agree on primal-dual variables that solve a global optimization problem [2]. Similarly, in multi-player games, decision makers exchange local or semi-global information to achieve an equilibrium for their inter-dependent optimization problems [3]. Applications of multi-agent systems with guaranteed convergence are indeed vast, e.g. include power systems [4], [5], demand side management [6], social networks [7], [8], robotic and sensor networks [9], [10].

From a general mathematical perspective, the convergence problem is a fixed-point problem [11], or equivalently, a zero-finding problem [12]. For example, consensus in multi-agent systems is equivalent to finding a collective state in the kernel of the Laplacian matrix, i.e., in operator-theoretic terms, to finding a zero of the Laplacian, seen as a linear operator.

Fixed-point theory and monotone operator theory are then key to study convergence to multi-agent equilibria [13]. For instance, Krasnoselskij–Mann fixed-point iterations have been adopted in aggregative game theory [14], [15], monotone operator splitting methods in distributed convex optimization [16] and monotone game theory [3], [17], [18]. Another motivating application is shown in Section VI-A: the Mann fixed-point iteration provides a simple solution to multi-agent consensus in discrete-time, even whenever the agents have no knowledge at all on the graph (connectivity).

The literature on fixed-point iterations is vast. The available results assume *sufficient* conditions, possibly not *necessary* in general, on the problem data to ensure global convergence of fixed-point iterations applied on *nonlinear* mappings.

To the best of our knowledge, we are the first to study *necessary* conditions for the convergence of fixed-point iterations, which is precisely the added value of this paper with respect to the available literature. We focus on the three most popular fixed-point iterations applied on *linear* operators, that essentially are linear time-varying systems with special structure. Our main technical contribution is to show that Krasnoselskij–Mann fixed-point iterations, possibly time-varying, applied on linear operators converge *if and only if* the associated matrix has certain spectral properties (Section III). One motivation for characterizing fixed-point iterations applied on *linear* operators is to provide non-convergence certificates for multi-agent dynamics that arise from distributed convex optimization and monotone game theory (Section VI-B). To achieve our main results, we adopt an operator-theoretic perspective and characterize some regularity properties of linear mappings via eigenvalue location and properties, and linear matrix inequalities (Section IV).

Notation: $\mathbb{R}$, $\mathbb{R}_{\geq 0}$ and $\mathbb{C}$ denote the set of real, non-negative real and complex numbers, respectively. $\mathbb{D}_r := \{z \in \mathbb{C} \mid |z-(1-r)| \leq r\}$ denotes the disk of radius $r > 0$ centered in $(1-r, 0)$, see Fig. 1 for some graphical examples. $\mathrm{bdry}(S)$ denotes the boundary of a set S. $\mathcal{H}(\|\cdot\|)$ denotes a finite-dimensional Euclidean space with inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\| := \sqrt{\langle \cdot, \cdot \rangle}$. $\mathbb{S}^n_{\succ 0}$ is the set of positive definite symmetric matrices and, for $P \in \mathbb{S}^n_{\succ 0}$, $\|x\|_P := \sqrt{x^\top P x}$. Id denotes the identity operator. $R(\cdot) := \begin{bmatrix} \cos(\cdot) & -\sin(\cdot) \\ \sin(\cdot) & \cos(\cdot) \end{bmatrix}$ denotes the rotation operator. Given a mapping $T : \mathbb{R}^n \to \mathbb{R}^n$, $\mathrm{fix}(T) := \{x \in \mathbb{R}^n \mid x = T(x)\}$ denotes the set of fixed points, and

G. Belgioioso is with the Control Systems group, TU Eindhoven, The Netherlands. F. Fabiani is with the Department of Information Engineering, University of Pisa, Italy. F. Blanchini is with the Department of Mathematics and Informatics, University of Udine, Italy. S. Grammatico is with the Delft Center for Systems and Control (DCSC), TU Delft, The Netherlands. E-mail addresses: g.belgioioso@tue.nl, filippo.fabiani@ing.unipi.it, franco.blanchini@uniud.it, s.grammatico@tudelft.nl.
This work was partially supported by NWO under research projects OMEGA (grant n. 613.001.702) and P2P-TALES (grant n. 647.003.003).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2018.2840882, IEEE Control Systems Letters

$\mathrm{zer}(T) := \{x \in \mathbb{R}^n \mid 0 = T(x)\}$ the set of zeros. Given a matrix $A \in \mathbb{R}^{n\times n}$, $\ker(A) := \{x \in \mathbb{R}^n \mid 0 = Ax\} = \mathrm{zer}(A\,\cdot)$ denotes its kernel; $\Lambda(A)$ and $\rho(A)$ denote the spectrum and the spectral radius of A, respectively. $\mathbf{0}_N$ and $\mathbf{1}_N$ denote vectors with N elements all equal to 0 and 1, respectively.

II. Mathematical definitions

A. Discrete-time linear systems

In this paper, we consider discrete-time linear time-invariant systems,

$$x(k+1) = Ax(k)\,, \tag{1}$$

and linear time-varying systems with special structure, i.e.,

$$x(k+1) = (1-\alpha_k)x(k) + \alpha_k A\, x(k)\,, \tag{2}$$

for some positive sequence $(\alpha_k)_{k\in\mathbb{N}}$. Note that for $\alpha_k = 1$ for all $k \in \mathbb{N}$, the system in (2) reduces to that in (1).

B. System-theoretic definitions

We are interested in the following notion of global convergence, i.e., convergence of the state solution to some vector, which may depend on the initial condition.

Definition 1 (Convergence): The system in (2) is *convergent* if, for all $x(0) \in \mathbb{R}^n$, there exists $\bar{x} \in \mathbb{R}^n$ such that the solution $x(k)$ to (2) satisfies $\lim_{k\to\infty} \|x(k) - \bar{x}\| = 0$. □

Note that in Definition 1, the vector $\bar{x}$ can depend on the initial condition $x(0)$. In the linear time-invariant case, (1), it is known that semi-convergence holds if and only if the eigenvalues of the A matrix are strictly inside the unit disk and the eigenvalue in 1, if present, must be semi-simple, as formalized next.

Definition 2 ((Semi-) Simple eigenvalue): An eigenvalue is *semi-simple* if it has equal algebraic and geometric multiplicity. An eigenvalue is *simple* if it has algebraic and geometric multiplicities both equal to 1. □

Lemma 1: The following statements are equivalent:

i) The system in (1) is *convergent*;
ii) $\rho(A) \leq 1$ and the only eigenvalue on the unit disk is 1, which is semi-simple. □

C. Operator-theoretic definitions

With the aim to study convergence of the dynamics in (1), (2), in this subsection, we introduce some key notions from operator theory in finite-dimensional Hilbert spaces.

Definition 3 (Lipschitz continuity): A mapping $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ is *ℓ-Lipschitz* continuous in $\mathcal{H}(\|\cdot\|)$, with $\ell \geq 0$, if $\forall x, y \in \mathbb{R}^n$, $\|\mathcal{T}(x) - \mathcal{T}(y)\| \leq \ell\,\|x - y\|$. □

Definition 4: In $\mathcal{H}(\|\cdot\|)$, an ℓ-Lipschitz continuous mapping $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ is

- *ℓ-Contractive* (ℓ-CON) if $\ell \in [0,1)$;
- *NonExpansive* (NE) if $\ell \in [0,1]$;
- *η-Averaged* (η-AVG), with $\eta \in (0,1)$, if $\forall x, y \in \mathbb{R}^n$

$$\begin{aligned}\|\mathcal{T}(x) - \mathcal{T}(y)\|^2 &\leq \|x-y\|^2 \\ &\quad - \tfrac{1-\eta}{\eta}\left\|(\mathrm{Id}-\mathcal{T})(x) - (\mathrm{Id}-\mathcal{T})(y)\right\|^2\,, \end{aligned} \tag{3}$$

or, equivalently, if there exists a nonexpansive mapping $\mathcal{B} : \mathbb{R}^n \to \mathbb{R}^n$ and $\eta \in (0,1)$ such that

$$\mathcal{T} = (1-\eta)\mathrm{Id} + \eta\mathcal{B}\,.$$

- *κ-strictly Pseudo-Contractive* (κ-sPC), with $\kappa \in (0,1)$, if $\forall x, y \in \mathbb{R}^n$

$$\begin{aligned}\|\mathcal{T}(x) - \mathcal{T}(y)\|^2 &\leq \|x-y\|^2 \\ &\quad + \kappa\left\|(\mathrm{Id}-\mathcal{T})(x) - (\mathrm{Id}-\mathcal{T})(y)\right\|^2\,. \end{aligned} \tag{4}$$

□

Definition 5: A mapping $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ is:

- *Contractive* (CON) if there exist $\ell \in [0,1)$ and a norm $\|\cdot\|$ such that it is an ℓ-CON in $\mathcal{H}(\|\cdot\|)$;
- *Averaged* (AVG) if there exist $\eta \in (0,1)$ and a norm $\|\cdot\|$ such that it is η-AVG in $\mathcal{H}(\|\cdot\|)$;
- *strict Pseudo-Contractive* (sPC) if there exists $\kappa \in (0,1)$ and a norm $\|\cdot\|$ such that it is κ-sPC in $\mathcal{H}(\|\cdot\|)$.

□

III. Main results: Fixed-point iterations on linear mappings

In this section, we provide necessary and sufficient conditions for the convergence of some well-known fixed-point iterations applied on linear operators, i.e.,

$$\mathcal{A} : x \mapsto Ax, \quad \text{with } A \in \mathbb{R}^{n\times n}. \tag{5}$$

First, we consider the *Banach–Picard* iteration [12, (1.69)] on a generic mapping $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$, i.e., for all $k \in \mathbb{N}$,

$$x(k+1) = \mathcal{T}(x(k))\,, \tag{6}$$

whose convergence is guaranteed if $\mathcal{T}$ is averaged, see [12, Prop. 5.16]. The next statement shows that averagedness is also a necessary condition when the mapping $\mathcal{T}$ is linear.

Proposition 1 (Banach–Picard iteration): The following statements are equivalent:

(i) $\mathcal{A}$ in (5) is averaged;
(ii) the solution to the system in (1) converges to some $\overline{x} \in \mathrm{fix}(\mathcal{A}) = \ker(I - A)$. □

If the mapping $\mathcal{T}$ is merely nonexpansive, then the sequence generated by the Banach–Picard iteration in (6) may fail to produce a fixed point of $\mathcal{T}$. For instance, this is the case for $\mathcal{T} = -\mathrm{Id}$. In these cases, a relaxed iteration can be used, e.g. the Krasnoselskij–Mann iteration [12, Equ. (5.15)]. Specifically, let us distinguish the case with time-invariant step sizes, known as Krasnoselskij iteration [11, Chap. 3], and the case with time-varying, vanishing step sizes, known as Mann iteration [11, Chap. 4]. The former is defined by

$$x(k+1) = (1-\alpha)x(k) + \alpha\mathcal{T}(x(k))\,, \tag{7}$$

for all $k \in \mathbb{N}$, where $\alpha \in (0,1)$ is a constant step size.

The convergence of the discrete-time system in (7) to a fixed point of the mapping $\mathcal{T}$ is guaranteed, for any arbitrary $\alpha \in (0,1)$, if $\mathcal{T}$ is nonexpansive [12, Th. 5.15], or if $\mathcal{T}$, defined from a compact, convex set to itself, is strictly pseudo-contractive and $\alpha > 0$ is sufficiently small [11, Theorem

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2018.2840882, IEEE Control Systems Letters

3.5]. In the next statement, we show that if the mapping $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ is linear, and α is chosen small enough, then strict pseudo-contractiveness is necessary and sufficient for convergence.

Theorem 1 (Krasnoselskij iteration): Let $\kappa \in (0,1)$ and $\alpha \in (0, 1-\kappa)$. The following statements are equivalent:

(i) $\mathcal{A}$ in (5) is κ-strictly pseudo-contractive;

(ii) the solution to the system

$$x(k+1) = (1-\alpha)x(k) + \alpha A x(k) \tag{8}$$

converges to some $\overline{x} \in \mathrm{fix}(\mathcal{A}) = \ker(I - A)$. □

Since $\alpha > 0$, (8) is equivalent to $\frac{1}{\alpha}(x(k+1) - x(k)) = (A-I)x(k)$, which represents the Euler approximation of the continuous-time Laplacian-like dynamics $\dot{x} = -(I-A)x$.

In Theorem 1, the admissible step sizes for the Krasnoselskij iteration depend on the parameter κ that quantifies the strict pseudo-contractiveness of the mapping $\mathcal{A} = A\,\cdot$. When the parameter κ is unknown, or hard to quantify, one can adopt time-varying step sizes, e.g. the Mann iteration:

$$x(k+1) = (1-\alpha_k)x(k) + \alpha_k \mathcal{T}(x(k)) \,, \tag{9}$$

for all $k \in \mathbb{N}$, where the step sizes $(\alpha_k)_{k\in\mathbb{N}}$ shall be chosen as follows.

Assumption 1 (Mann sequence): The sequence $(\alpha_k)_{k\in\mathbb{N}}$ is such that $0 < \alpha_k \leq \alpha^{\max} < \infty$ for all $k \in \mathbb{N}$, for some $\alpha^{\max}$, $\lim_{k\to\infty} \alpha_k = 0$ and $\sum_{k=0}^{\infty} \alpha_k = \infty$. □

The convergence of (9) to a fixed point of the mapping $\mathcal{T}$ is guaranteed if $\mathcal{T}$, defined from a compact, convex set to itself, is strictly pseudo-contractive [11, Theorem 3.5]. In the next statement, we show that if the mapping $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ is linear, then strict pseudo-contractiveness is necessary and sufficient for convergence.

Theorem 2 (Mann iteration): Let $(\alpha_k)_{k\in\mathbb{N}}$ be a Mann sequence as in Assumption 1. The following statements are equivalent:

(i) $\mathcal{A}$ in (5) is strictly pseudocontractive;

(ii) the solution to

$$x(k+1) = (1-\alpha_k)x(k) + \alpha_k A x(k) \tag{10}$$

converges to some $\overline{x} \in \mathrm{fix}(\mathcal{A}) = \ker(I - A)$. □

We remark that the sequence in (10) does not necessary converge to a vector in $\ker(I-A)$ if α_k converges too quickly, as it can be shown by simple counterexamples.

IV. Operator-theoretic characterization of linear mappings

In this section, we characterize the operator-theoretic properties of linear mappings via necessary and sufficient linear matrix inequalities and conditions on the spectrum of the corresponding matrices. We exploit these technical results in Section V, to prove convergence of the fixed-point iterations presented in Section III.

Lemma 2 (Lipschitz continuous linear mapping): Let $\ell > 0$ and $P \in \mathbb{S}^n_{\succ 0}$. The following statements are equivalent:

(i) $\mathcal{A}$ in (5) is ℓ-Lipschitz continuous in $\mathcal{H}(\|\cdot\|_P)$;

(ii) $A^\top P A \preccurlyeq \ell^2 P$. □

Proof: It directly follows from Definition 3. ■

Lemma 3 (Linear contractive/nonexpansive mapping): Let $\ell \in (0,1)$. The following statements are equivalent:

(i) $\mathcal{A}$ in (5) is an ℓ-contraction;

(ii) $\exists P \in \mathbb{S}^n_{\succ 0}$ such that $A^\top P A \preccurlyeq \ell^2 P$;

(iii) the spectrum of A is such that

$$\begin{cases} \Lambda(A) \subset \ell\,\mathbb{D}_1 \\ \forall \lambda \in \Lambda(A) \cap \mathrm{bdry}(\ell\,\mathbb{D}_1), \quad \lambda \text{ semi-simple} \end{cases} \tag{11}$$

If $\ell = 1$, the previous equivalent statements hold if and only if $\mathcal{A}$ in (5) is nonexpansive. □

Proof: The equivalence between (i) and (ii) follows from Lemma 2. By the Lyapunov theorem, (iii) holds if and only if the discrete-time linear system $x(k+1) = \frac{1}{\ell} A\, x(k)$ is (at least marginally) stable, i.e., $\Lambda(A) \subset \ell\,\mathbb{D}_1$ and the eigenvalues of A on the boundary of the disk, $\Lambda(A) \cap \mathrm{bdry}(\ell\,\mathbb{D}_1)$, are semi-simple. The last statement follows by noticing that an 1-contractive mapping is nonexpansive. ■

Lemma 4 (Linear averaged mapping): Let $\eta \in (0,1)$. The following statements are equivalent:

(i) $\mathcal{A}$ in (5) is η-averaged;

(ii) $\exists P \in \mathbb{S}^n_{\succ 0}$ such that

$$A^\top P A \preccurlyeq (2\eta - 1)\,P + (1-\eta)\left(A^\top P + P A\right);$$

(iii) $\mathcal{A}_\eta := A_\eta \cdot := \left(1 - \frac{1}{\eta}\right) I \cdot + \frac{1}{\eta} A \cdot$ is nonexpansive;

(iv) the spectrum of A is such that

$$\begin{cases} \Lambda(A) \subset \mathbb{D}_\eta \\ \forall \lambda \in \Lambda(A) \cap \mathrm{bdry}(\mathbb{D}_\eta), \quad \lambda \text{ semi-simple.} \end{cases} \tag{12}$$

□

Proof: The equivalence (i) $\Leftrightarrow$ (ii) follows directly by inequality (3) in Definition 4. By [12, Prop. 4.35], $\mathcal{A}$ is η-AVG if and only if the linear mapping $\mathcal{A}_\eta$ is NE, which proves (i) $\Leftrightarrow$ (iii). To conclude, we show that (iii) $\Leftrightarrow$ (iv). By Lemma 3, the linear mapping $\mathcal{A}_\eta$ is NE if and only if

$$\begin{cases} \Lambda(A_\eta) \subset \mathbb{D}_1 \\ \forall \lambda \in \Lambda(A_\eta) \cap \mathrm{bdry}(\mathbb{D}_1), \quad \lambda \text{ semi-simple} \end{cases} \tag{13}$$

$$\Leftrightarrow \begin{cases} \Lambda(A) \subset (1-\eta)\{1\} + \eta\mathbb{D}_1 = \mathbb{D}_\eta \\ \forall \lambda \in \Lambda(A) \cap \mathrm{bdry}(\mathbb{D}_\eta), \quad \lambda \text{ semi-simple} \end{cases} \tag{14}$$

where the equivalence (13) $\Leftrightarrow$ (14) holds because $\Lambda(A_\eta) = (1 - \frac{1}{\eta})\{1\} + \frac{1}{\eta}\Lambda(A)$, and because the linear combination with the identity matrix does not alter the geometric multiplicity of the eigenvalues. ■

Lemma 5 (Linear strict pseudocontractive mapping): Let $\kappa, \eta \in (0,1)$. The following statements are equivalent:

(i) $\mathcal{A}$ in (5) is κ-strictly pseudocontractive;

(ii) $\exists P \in \mathbb{S}^n_{\succ 0}$ such that

$$(1-\kappa) A^\top P A \preccurlyeq (1+\kappa) P - \kappa (A^\top P + P A); \tag{15}$$

(iii) $\mathcal{A}^{\mathrm{s}}_\kappa := A^{\mathrm{s}}_\kappa \cdot := \kappa I \cdot + (1-\kappa) A \cdot$ is nonexpansive;

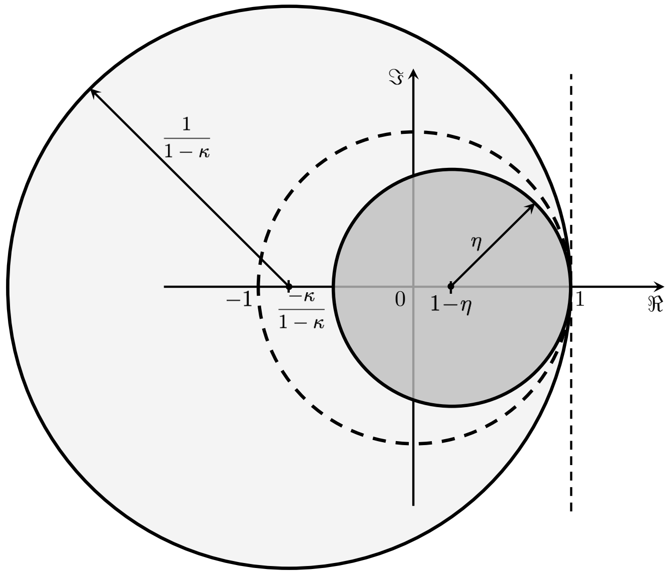

Fig. 1. Spectrum of a linear η-AVG mapping: Disk centered in $1-\eta$ with radius η, $\mathbb{D}_\eta$ (dark-grey disk). Spectrum of a linear κ-sPC mapping: Disk centered in $-\frac{\kappa}{1-\kappa}$ with radius $\frac{1}{1-\kappa}$, $\mathbb{D}_{\frac{1}{1-\kappa}}$ (light-grey disk).

(iv) the spectrum of A is such that

$$\begin{cases}\Lambda(A) \subset \mathbb{D}_{\frac{1}{1-\kappa}} \\ \forall \lambda \in \Lambda(A) \cap \operatorname{bdry}\left(\mathbb{D}_{\frac{1}{1-\kappa}}\right), \quad \lambda \text{ semi-simple}\end{cases} \tag{16}$$

(v) $\mathcal{A}_\alpha := A_\alpha \cdot := (1-\alpha)I \cdot + \alpha A \cdot$ is η-averaged, with $\alpha = \eta(1-\kappa) \in (0,1)$. □

Proof: The equivalence (i) $\Leftrightarrow$ (ii) follows directly by inequality (4) in Definition (5). To prove that (ii) $\Leftrightarrow$ (iii), we note that the LMI in (15) can be recast as

$$(\kappa I + (1-\kappa)A)^\top P\,(\kappa I + (1-\kappa)A) \preccurlyeq P, \tag{17}$$

which, by Lemma 3, holds true if and only if $\mathcal{A}^{\mathrm{s}}_\kappa$ is NE. (iii) $\Leftrightarrow$ (iv): By Lemma 3, $\mathcal{A}^{\mathrm{s}}_\kappa$ is NE if and only if

$$\begin{cases}\Lambda(A^{\mathrm{s}}_\kappa) \subset \mathbb{D}_1 \\ \forall \lambda \in \Lambda(A^{\mathrm{s}}_\kappa) \cap \operatorname{bdry}(\mathbb{D}_1), \quad \lambda \text{ semi-simple}\end{cases} \tag{18}$$

$$\Leftrightarrow \begin{cases}\Lambda(A) \subset \left\{-\frac{\kappa}{1-\kappa}\right\} + \frac{1}{1-\kappa}\mathbb{D}_1 = \mathbb{D}_{\frac{1}{1-\kappa}} \\ \forall \lambda \in \Lambda(A) \cap \operatorname{bdry}\left(\mathbb{D}_{\frac{1}{1-\kappa}}\right), \quad \lambda \text{ semi-simple}\end{cases} \tag{19}$$

where the equivalence (18) $\Leftrightarrow$ (19) holds because $\Lambda(A^{\mathrm{s}}_\kappa) = A^{\mathrm{s}}_\kappa := \kappa I + (1-\kappa)A$, and because the linear combination with the identity matrix does not alter the geometric multiplicity of the eigenvalues. (iii) $\Leftrightarrow$ (v): By Definition 4 and [12, Prop. 4.35], $A^{\mathrm{s}}_\kappa \cdot$ is NE if and only if $A_\alpha \cdot = (1-\eta)I \cdot + \eta A^{\mathrm{s}}_\kappa \cdot$ is η-AVG, for all $\eta \in (0,1)$. Since $\alpha = \eta(1-\kappa)$, $A_\alpha = (1-\eta(1-\kappa))\mathrm{Id} + \eta(1-\kappa)A$, which concludes the proof. ■

V. Proofs of the Main Results

Proof of Proposition 1 (Banach–Picard iteration)

We recall that, by Lemma 4, $\mathcal{A}$ is AVG if and only if there exists $\eta \in (0,1)$ such that $\Lambda(A) \subset \mathbb{D}_\eta$ and $\forall \lambda \in \Lambda(A) \cap \operatorname{bdry}(\mathbb{D}_\eta)$, λ is semi-simple and we notice that $\mathbb{D}_\eta \cap \mathbb{D}_1 = \{1\}$ for all $\eta \in (0,1)$. Hence $\mathcal{A}$ is averaged if and only if the eigenvalues of A are strictly contained in the unit circle except for the eigenvalue in $\lambda = 1$ which, if present, is semi-simple. The latter is a necessary and sufficient condition for the convergence of $x(k+1) = A\,x(k)$, by Lemma 1. ■

Proof of Theorem 1 (Krasnoselskij iteration)

(i) $\Leftrightarrow$ (ii): By Lemma 5, $\mathcal{A}$ is $\kappa-$sPC if and only if $(1-\alpha)\mathrm{Id} + \alpha\mathcal{A}$ is η-AVG, with $\alpha = \eta(1-\kappa)$ and $\eta \in (0,1)$; therefore, if and only if $(1-\alpha)\mathrm{Id}+\alpha\mathcal{A}$ is AVG with $\alpha \in (0, 1-\kappa)$. By Proposition (1), the latter is equivalent to the global convergence of the Banach–Picard iteration applied on $(1-\alpha)\mathrm{Id} + \alpha\mathcal{A}$, which corresponds to the Krasnoselskij iteration on $\mathcal{A}$, with $\alpha \in (0, 1-\kappa)$. ■

Proof of Theorem 2 (Mann iteration)

Proof that (i) $\Rightarrow$ (ii): Define the bounded sequence $\beta_k := \frac{1}{\epsilon}\alpha_k > 0$, for some $\epsilon > 0$ to be chosen. Thus, $x(k+1) = (1-\alpha_k)x(k) + \alpha_k A x(k) = (1-\epsilon\beta_k)x(k) + \epsilon\beta_k A x(k) = (1-\beta_k)x(k) + \beta_k\left((1-\epsilon)I + \epsilon A\right)x(k)$. Since $A\cdot$ is sPC, we can choose $\epsilon > 0$ small enough such that $B := (1-\epsilon)\mathrm{Id} + \epsilon A\cdot$ is NE, specifically, we shall choose $\epsilon < \min\left\{\frac{1}{|\lambda|} \mid \lambda \in \Lambda(A) \setminus \{1\}\right\}$. Note that $0 \in \operatorname{fix}(A) = \operatorname{fix}(B) \neq \varnothing$. Since $\beta_k = \frac{1}{\epsilon}\alpha_k \to 0$, we have that $\forall \epsilon, \epsilon' \in (0,1)$, $\exists \bar{k} \in \mathbb{N}$ such that $\beta_k \leq 1-\epsilon'$ for all $k \geq \bar{k}$. Moreover, since $\sum_{k=0}^{\bar{k}} \beta_k < \infty$, for all $x(0) \in \mathbb{R}^n$, we have that the solution $x(\bar{k})$ is finite. Therefore, we can define $h := k - \bar{k} \in \mathbb{N}$ for all $k \geq \bar{k}$, $y(0) := x(\bar{k})$ and $y(h+1) := x(h+\bar{k}+1)$ for all $h \geq 0$. The proof then follows by applying [12, Th. 5.14 (iii)] to the Mann iteration $y(h+1) = (1-\beta_h)y(h) + \beta_h B y(h)$. In fact, as $\beta_h \leq 1-\epsilon'$, $\sum_{h=0}^{\infty} \beta_h(1-\beta_h) \geq \epsilon' \sum_{h=0}^{\infty} \beta_h = \infty$.

Proof that (ii) $\Rightarrow$ (i): For the sake of contradiction, suppose that A is not sPC, i.e., at least one of the following facts must hold: 1) A has an eigenvalue in 1 that is not semi-simple; 2) A has a real eigenvalue greater than 1; 3) A has a pair of complex eigenvalues $\sigma \pm j\omega$, with $\sigma \geq 1$ and $\omega > 0$. We show next that each of these facts implies non-convergence of (9). Without loss of generality (i.e., up to a linear transformation), we can assume that A is in Jordan normal form.

1) A has an eigenvalue in 1 that is not semi-simple. Due to (the bottom part of) the associated Jordan block, the dynamics in (9) contain the two-dimensional dynamics

$$\begin{aligned} y(k+1) &= \left((1-\alpha_k)\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \alpha_k \begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}\right) y(k) \\ &= \begin{bmatrix}1 & \alpha_k\\ 0 & 1\end{bmatrix} y(k). \end{aligned}$$

For $y_2(0) := c > 0$, we have that the solution $y(k)$ is such that $y_2(k) = y_2(0) > 0$ and $y_1(k+1) = y_1(k) + \alpha_k c$, which implies that $y_1(k) = y_1(0) + c\sum_{h=0}^{k-1}\alpha_h$. Thus, $x(k)$ diverges and we have a contradiction.

2) Let A has a real eigenvalue equal to $1+\epsilon > 1$. Again due to the associated Jordan block, the dynamics in (9) must contain the scalar dynamics $s(k+1) = (1-\alpha_k)s(k) + \alpha_k(1+\epsilon)s(k) = (1+\epsilon\,\alpha_k)s(k)$, with solution

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LCSYS.2018.2840882, IEEE Control Systems Letters

$s(k+1) = \left(\prod_{h=0}^{k}(1+\epsilon\,\alpha_h)\right) s(0)$. Since $\epsilon\,\alpha_h > 0$, it holds that $\prod_{h=0}^{\infty}(1+\epsilon\,\alpha_h) \geq \epsilon\sum_{h=0}^{\infty}\alpha_h = \infty$, by Assumption 1. Therefore, $y(k)$ and $x(k)$ diverge and we reach a contradiction.

3) A has a pair of complex eigenvalues $\sigma \pm j\omega$, with $\sigma = 1+\epsilon \geq 1$ and $\omega > 0$. Due to the structure of the associated Jordan block, the vector dynamics in (9) contain the two-dimensional dynamics

$$\begin{aligned} z(k+1) &= \left((1-\alpha_k)\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \alpha_k\begin{bmatrix}\sigma & -\omega\\ \omega & \sigma\end{bmatrix}\right) z(k) \\ &= \begin{bmatrix}1+\epsilon\,\alpha_k & -\omega\alpha_k\\ \omega\alpha_k & 1+\epsilon\,\alpha_k\end{bmatrix} z(k). \end{aligned}$$

Now, we define $\rho_k := \sqrt{(1+\epsilon\alpha_k)^2+\omega^2\alpha_k^2} \geq \sqrt{1+\omega^2\alpha_k^2} > 1$, and the angle $\theta_k > 0$ such that $\cos(\theta_k) = (1+\epsilon\alpha_k)/\rho_k$ and $\sin(\theta_k) = (\omega\alpha_k)/\rho_k$, i.e., $\theta_k = \mathrm{atan}\left(\frac{\omega\alpha_k}{1+\epsilon\alpha_k}\right)$. Then, we have that $z(k+1) = \rho_k R(\theta_k) z(k)$, hence, the solution $z(k)$ reads as

$$z(k+1) = \left(\prod_{h=0}^{k}\rho_h\right) R\left(\sum_{h=0}^{k}\theta_h\right) z(0).$$

Since $\|R(\cdot)\| = 1$, if the product $(\prod_{h=0}^{\infty}\rho_h)$ diverges, then $z(k)$ diverges as well. Thus, let us assume that the product $(\prod_{h=0}^{\infty}\rho_h)$ converges. By the limit comparison test, the series $\sum_{h=0}^{\infty}\theta_h = \sum_{h=0}^{\infty}\mathrm{atan}\left(\frac{\omega\alpha_h}{1+\epsilon\alpha_h}\right)$ converges (diverges) if and only the series $\sum_{h=0}^{\infty}\frac{\omega\alpha_h}{1+\epsilon\alpha_h}$ converges (diverges). The latter diverges since $\sum_{h=0}^{\infty}\frac{\omega\alpha_h}{1+\epsilon\alpha_h} \geq \frac{\omega}{1+\epsilon\alpha^{\max}}\sum_{h=0}^{\infty}\alpha_h = \infty$. It follows that $\sum_{h=0}^{\infty}\theta_h$ diverges, hence $z(k)$ keeps rotating indefinitely, which is a contradiction. ■

VI. Application to Multi-Agent Linear Systems

A. Consensus via discrete-time linear time-varying dynamics

We consider a connected graph of N nodes, associated with N agents seeking consensus via discrete-time, distributed information exchange on a network with fixed Laplacian matrix $L \in \mathbb{R}^{N\times N}$. The aim is to reach consensus assuming that the algebraic connectivity of the graph, i.e., the strictly-positive Fiedler eigenvalue of L, is unknown and no bounds are available.

Thus, in view of [1, Equ. (16)], we study the following linear time-varying dynamics:

$$\begin{aligned} \boldsymbol{x}(k+1) &= \boldsymbol{x}(k) - \alpha_k L\boldsymbol{x}(k) && \text{(20a)}\\ &= (1-\alpha_k)\boldsymbol{x}(k) + \alpha_k(I-L)\boldsymbol{x}(k)\,, && \text{(20b)} \end{aligned}$$

where $\boldsymbol{x}(k) := [x_1(k),\ldots,x_N(k)]^\top \in \mathbb{R}^N$ and, for simplicity, the state of each agent is a scalar variable, $x_i \in \mathbb{R}$.

Since the dynamics in (20) have the structure of a Mann iteration, due to Theorem 2, we have the following result.

Corollary 1: Let $(\alpha_k)_{k\in\mathbb{N}}$ be a Mann sequence. The system in (20) asymptotically reaches consensus, i.e., the solution $\boldsymbol{x}(k)$ to (20) converges to $\overline{x}\,\mathbf{1}_N$, for some $\overline{x} \in \mathbb{R}$. □

Proof: Since the graph is connected, L has one (simple) eigenvalue at 0, and $N-1$ eigenvalues with strictly-positive real part. Therefore, the matrix $I-L$ in (20b) has one simple eigenvalue in 1 and $N-1$ with real part strictly less than 1. By Lemma 5, $(I-L)(\cdot)$ is sPC and by Theorem 2, $\boldsymbol{x}(k)$ globally

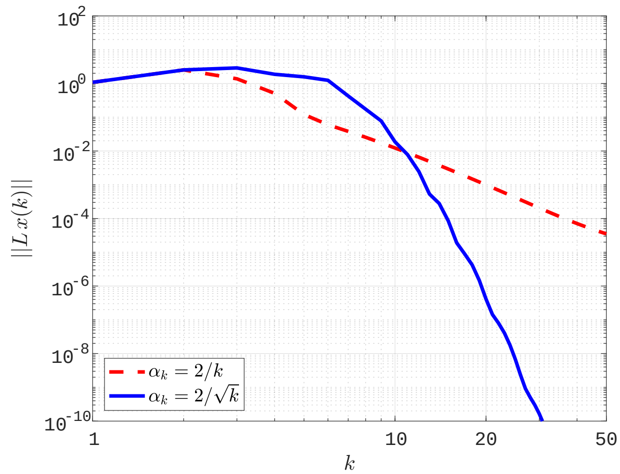

Fig. 2. Disagreement vector norm versus discrete time. Since the mapping $\mathrm{Id} - L\,\cdot$ is strictly pseudocontractive, consensus is asymptotically reached.

converges to some $\overline{\boldsymbol{x}} \in \mathrm{fix}(I-L) = \mathrm{zer}(L)$, i.e., $L\overline{\boldsymbol{x}} = \mathbf{0}_N$. Since L is a Laplacian matrix, $L\overline{\boldsymbol{x}} = \mathbf{0}_N$ implies consensus, i.e., $\overline{\boldsymbol{x}} = \overline{x}\,\mathbf{1}_N$, for some $\overline{x} \in \mathbb{R}$. ■

We have only assumed that the agents agree on a sequence of vanishing, bounded, step sizes, α_k. However, we envision that agent-dependent step sizes can be used as well, e.g. via matricial Mann iterations [11, §4.1]. We stress that the mere convergence of α_k does not imply that consensus is achieved. In fact, if the sequence $\{\alpha_k\}_{k\in\mathbb{N}}$ is summable, then the limit vector $\overline{x}$ may not be in the kernel of the Laplacian matrix.

Let us simulate the consensus dynamics in (20) for a graph with 3 nodes, adjacency matrix $A = [a_{i,j}]$ with $a_{1,2} = a_{1,3} = \frac{1}{2}$, $a_{2,3} = a_{3,1} = 1$, hence with Laplacian matrix

$$L = D_{\text{out}} - A = \begin{bmatrix} 1 & -1/2 & -1/2\\ 0 & 1 & -1\\ -1 & 0 & 1\end{bmatrix}.$$

We note that L has eigenvalues $\Lambda(L) = \left\{0, \frac{3}{2} \pm j\frac{1}{2}\right\}$. Since we do not assume that the agents known about the connectivity of the graph, we simulate with step sizes that are initially larger than the maximum constant-step value for which convergence would hold. In Fig. 2, we compare the norm of the disagreement vectors, $\|Lx(k)\|$, obtained with two different Mann sequences, $\alpha_k = 2/k$ and $\alpha_k = 2/\sqrt{k}$, respectively. We observe that convergence with small tolerances is faster in the latter case with larger step sizes.

B. Two-player zero-sum linear-quadratic games: Non-convergence of projected pseudo-gradient dynamics

We consider two-player zero-sum games with linear-quadratic structure, i.e., we consider $N = 2$ agents, with cost functions $f_1(x_1,x_2) := x_1^\top C x_2$ and $f_2(x_1,x_2) := -x_2^\top C^\top x_1$, respectively, for some square matrix $C = C^\top \neq 0$. In particular, we study discrete-time dynamics for solving the Nash equilibrium problem, that is the problem to find a pair (x_1^*, x_2^*) such that:

$$\begin{cases} x_1^* & \in \underset{x_1\in\mathbb{R}^n}{\mathrm{argmin}}\, f_1(x_1, x_2^*)\\ x_2^* & \in \underset{x_2\in\mathbb{R}^n}{\mathrm{argmin}}\, f_2(x_1^*, x_2). \end{cases}$$

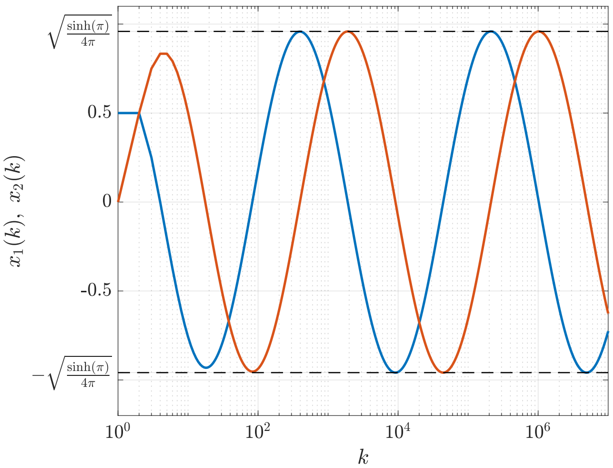

Fig. 3. Solution to the discrete-time system in (21) in semi-log scale. The lack of strict pseudo-contractiveness causes persistent oscillations.

A classic solution approach is the pseudo-gradient method, namely the discrete-time dynamics

$$\begin{aligned} \boldsymbol{x}(k+1) &= \boldsymbol{x}(k) - \alpha_k F \boldsymbol{x}(k) && \text{(21a)} \\ &= (1-\alpha_k)\boldsymbol{x}(k) + \alpha_k (I - F)\boldsymbol{x}(k)\,, && \text{(21b)} \end{aligned}$$

where $F\cdot$ is the so-called pseudo-gradient mapping of the game, which in our case is defined as

$$\mathcal{F}(\boldsymbol{x}) := \begin{bmatrix} \nabla_{x_1} f_1(x_1,x_2) \\ \nabla_{x_2} f_2(x_1,x_2) \end{bmatrix} = \begin{bmatrix} Cx_2 \\ -Cx_1 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \otimes C}_{=:F} \,\boldsymbol{x}\,,$$

and $(\alpha_k)_{k\in\mathbb{N}}$ is a sequence of vanishing step sizes, e.g. a Mann sequence. In our case, (x_1^*, x_2^*) is a Nash equilibrium if and only if $[x_1^*\,; x_2^*] \in \operatorname{fix}(\mathrm{Id} - \mathcal{F}) = \operatorname{zer}(\mathcal{F})$ [17, Th. 1].

By Theorem 2, convergence of the system in (21) holds if and only if $I - F$ is strictly pseudocontractive. In the next statement, we show that this is not the case for F in (21).

Corollary 2: Let $(\alpha_k)_{k\in\mathbb{N}}$ be a Mann sequence and $C = C^\top \neq 0$. The system in (21) does not globally converge. □

Proof: It follows by Lemma 5 that the mapping $\mathrm{Id} - F\cdot$ is sPC if and only if the eigenvalues of F either have strictly-positive real part, or are semi-simple and equal to 0. Since $\Lambda\left(\left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]\right) = \{\pm j\}$, we have that the eigenvalues of $F = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right] \otimes C$ are either with both positive and negative real part, or on the imaginary axis and not equal to 0, or equal to 0 and not semi-simple. Therefore, $\mathrm{Id} - F\cdot$ is not sPC and the proof follows by Theorem 2. ■

Let us numerically simulate the discrete-time system in (21), with the following parameters: $n = 1$, $C = 1$, $x_1(0) = 1/2$, $x_2(0) = 0$, and $\alpha_k = 1/(k+1)$ for all $k \in \mathbb{N}$. Figure 3 shows persistent oscillations, due to the lack of strict pseudo-contractiveness of $I - F$. In fact, $\Lambda(I - F) = \{1 \pm j\}$. The example provides a non-convergence result: pseudo-gradient methods do not ensure global convergence in convex games with (non-strictly) monotone pseudo-gradient mapping, not even with vanishing step sizes and linear-quadratic structure.

VII. Conclusion and Outlook

Convergence in discrete-time linear systems can be equivalently characterized via operator-theoretic notions. Remarkably, the time-varying Mann iteration applied on linear mappings converges if and only if the considered linear operator is strictly pseudocontractive. This result implies that Laplacian-driven linear time-varying consensus dynamics with Mann step sizes do converge. It also implies that projected pseudo-gradient dynamics for Nash equilibrium seeking in monotone games do not necessarily converge.

Future research will focus on other, more general, linear fixed-point iterations, e.g. the general Mann iteration [11, §4.1], and of discrete-time linear systems with uncertainty.

References

[1] R. Olfati-Saber, A. Fax, and R. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.
[2] A. Nedić, A. Ozdaglar, and P. Parrillo, "Constrained consensus and optimization in multi-agent networks," *IEEE Trans. on Automatic Control*, vol. 55, no. 4, pp. 922–938, 2010.
[3] P. Yi and L. Pavel, "A distributed primal-dual algorithm for computation of generalized Nash equilibria via operator splitting methods," *IEEE Conf. on Decision and Control*, pp. 3841–3846, 2017.
[4] F. Dörfler, J. Simpson-Porco, and F. Bullo, "Breaking the hierarchy: Distributed control and economic optimality in microgrids," *IEEE Trans. on Control of Network Systems*, vol. 3, pp. 241–253, 2016.
[5] F. Dörfler and S. Grammatico, "Gather-and-broadcast frequency control in power systems," *Automatica*, vol. 79, pp. 296–305, 2017.
[6] A.-H. Mohsenian-Rad, V. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, "Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid," *IEEE Trans. on Smart Grid*, vol. 1, no. 3, pp. 320–331, 2010.
[7] J. Ghaderi and R. Srikant, "Opinion dynamics in social networks with stubborn agents: Equilibrium and convergence rate," *Automatica*, vol. 50, pp. 3209–3215, 2014.
[8] S. R. Etesami and T. Başar, "Game-theoretic analysis of the hegselmann-krause model for opinion dynamics in finite dimensions," *IEEE Trans. on Automatic Control*, vol. 60, pp. 1886–1897, 2015.
[9] S. Martínez, F. Bullo, J. Cortés, and E. Frazzoli, "On synchronous robotic networks – Part i: Models, tasks, and complexity," *IEEE Trans. on Automatic Control*, vol. 52, pp. 2199–2213, 2007.
[10] M. Stanković, K. Johansson, and D. Stipanović, "Distributed seeking of Nash equilibria with applications to mobile sensor networks," *IEEE Trans. on Automatic Control*, vol. 57, no. 4, pp. 904–919, 2012.
[11] V. Berinde, *Iterative Approximation of Fixed Points*. Springer, 2007.
[12] H. H. Bauschke and P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*. Springer, 2010.
[13] E. K. Ryu and S. Boyd, "A primer on monotone operator methods," *Appl. Comput. Math.*, vol. 15, no. 1, pp. 3–43, 2016.
[14] S. Grammatico, F. Parise, M. Colombino, and J. Lygeros, "Decentralized convergence to Nash equilibria in constrained deterministic mean field control," *IEEE Trans. on Automatic Control*, vol. 61, no. 11, pp. 3315–3329, 2016.
[15] S. Grammatico, "Dynamic control of agents playing aggregative games with coupling constraints," *IEEE Trans. on Automatic Control*, vol. 62, no. 9, pp. 4537 – 4548, 2017.
[16] P. Giselsson and S. Boyd, "Linear convergence and metric selection for Douglas–Rachford splitting and ADMM," *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 532–544, 2017.
[17] G. Belgioioso and S. Grammatico, "Semi-decentralized Nash equilibrium seeking in aggregative games with coupling constraints and non-differentiable cost functions," *IEEE Control Systems Letters*, vol. 1, no. 2, pp. 400–405, 2017.
[18] S. Grammatico, "Proximal dynamics in multi-agent network games," *IEEE Trans. on Control of Network Systems*, `https://doi.org/10.1109/TCNS.2017.2754358`, 2018.