

# Markov operators on cones and non-commutative consensus

Stéphane Gaubert\*  
INRIA and CMAP  
École Polytechnique  
91128 Palaiseau Cédex, France  
Stephane.Gaubert@inria.fr

Zheng Qu\*  
CMAP, INRIA  
École Polytechnique  
91128 Palaiseau Cédex, France  
zheng.qu@polytechnique.edu

**Abstract**—The analysis of classical consensus algorithms relies on contraction properties of Markov matrices with respect to the Hilbert semi-norm (infinitesimal version of Hilbert’s projective metric) and to the total variation norm. We generalize these properties to the case of operators on cones. This is motivated by the study of “non-commutative consensus”, i.e., of the dynamics of linear maps leaving invariant cones of positive semi-definite matrices. Such maps appear in quantum information (Kraus maps), and in the study of matrix means. We give a characterization of the contraction rate of an abstract Markov operator on a cone, which extends classical formulæ obtained by Doeblin and Dobrushin in the case of Markov matrices. In the special case of Kraus maps, we relate the absence of contraction to the positivity of the “zero-error capacity” of a quantum channel. We finally show that a number of decision problems concerning the contraction rate of Kraus maps reduce to finding a rank one matrix in linear spaces satisfying certain conditions and discuss complexity issues.

## I. INTRODUCTION

Given a multi-agent system, the design of distributed algorithms allowing agents to reach consensus by exchanging locally computed results has been studied extensively in the field of communication networks, control theory and parallel computation, [Hir89], [BT89], [BGPS06]. Due to its wide area of applications, the convergence analysis of consensus algorithms has attracted much interest. Classical linear consensus algorithms can be written as time dependent dynamical systems, involving sequences of stochastic matrices. Their convergence has been shown to be controlled by connectivity properties of the adjacency graph of these matrices [Mor05], [VJAJ05], [OT09], [AB09]. These properties are closely related to estimates arising in the study of the coupling time of nonhomogeneous Markov chains. In particular, a basic tool is the characterization, which goes back to Doeblin and Dobrushin, of the Lipschitz constant of a Markov operator with respect to the total variation norm on the space of probability measures (Dobrushin ergodicity coefficient). In [MDA05], the Dobrushin’s ergodicity coefficient was used to determine the contraction rate of the Lyapunov function  $\Delta(x) := \max_i x_i - \min_i x_i$  for a consensus operator. Various convergence results on classical consensus algorithms, obtained by different means, can also be related to Dobrushin’s characterization [Mor05], [VJAJ05], [AB09].

\*The authors were partially supported by the PGM0 programme of Fédération Mathématique Jacques Hadamard and EDF.

In [SSR10], the authors developed a noncommutative generalization of consensus theory, in which the space of the system is represented by a Hermitian matrix, instead of a vector in  $\mathbb{R}^n$ . This is motivated in particular by quantum information theory, in which one is interested to the convergence of “noncommutative Markov chains” to equilibrium. In this setting, generalized consensus algorithms update the system state by a stochastic quantum map, which is the adjoint of a Kraus map, or quantum channel. Other important motivations to study consensus over spaces of matrices arise from the study of non-linear order-preserving positively homogeneous maps on cones, including generalized matrix means. The differentials of these maps are precisely noncommutative consensus operators. It is known that the global convergence to a fixed point, for such non-linear maps, can be inferred from the contraction properties of their differentials [Nus88]. Thus, computing, or giving tight effective bound, of the contraction rate of linear consensus operators, appear to be a key issue.

To study the convergence of noncommutative consensus, the authors of [SSR10] proposed to use as a Lyapunov function the Hilbert’s projective metric between the state matrix and the identity matrix. In particular, they applied Birkhoff’s contraction formula to give a bound on the contraction rate on this Lyapunov function.

One of the main goals of this paper is a generalization of the approach in [MDA05] to the study of noncommutative consensus [SSR10]. To this end, we use the Hilbert’s semi-norm  $\|X\|_H = \lambda_{\max}(X) - \lambda_{\min}(X)$  as a generalization of the Lyapunov function  $\Delta$ . In particular, we clarify the connection between the contraction rate of a consensus operator with respect to the Hilbert’s semi-norm and the Lipschitz constant of the dual density operator with respect to the trace distance (nuclear norm). We shall see that the contraction rate with respect to the Hilbert’s seminorm is always less than the contraction rate with respect to the Lyapunov function proposed in [SSR10]. This is because Hilbert’s projective metric is precisely the weak Finsler structure in which the infinitesimal lengths are given by a family of rescaled Hilbert’s seminorms [Nus94].

The convex duality nature of some of our results will be more apparent by working in a higher generality. Hence, we shall consider consensus dynamics in which the state belongs to an arbitrary normal cone. The Hilbert’s seminorm is still

well-defined and can serve as a Lyapunov function. Then, the consensus algorithms are defined by abstract Markov operators, which generalize both stochastic matrices and stochastic quantum maps. This generality is also convenient in various applications (for instance, generalized means often involve operators defined on Cartesian products of symmetric cones).

We first observe that the contraction rate of an abstract Markov operator with respect to Hilbert's semi-norm can be thought of as an operator norm on a quotient space (Lemma 2.5). Then, the Lipschitz constant of the dual operator with respect to the total variation distance is exactly the dual operator norm. Our first main result, Theorem 3.6, gives a formula for the two (primal and dual) operator norms using a characterization of the extreme points representation of the dual unit ball. When specialized to the noncommutative consensus, this leads to a noncommutative version of the Dobrushin's formula, Corollary 4.1, as well to a characterization of the absence of contraction (i.e., contraction rate equal to 1), Corollary 4.3. It follows that the zero-error capacity of a memoryless quantum channel is positive if and only if the contraction rate of the associated consensus operator is equal to 1. Finally we show that deciding if a time-invariant noncommutative system is globally convergent can be reduced to finding a rank one matrix in a certain matrix subspace, Theorem 4.5.

A Markov matrix is primitive if and only if it is irreducible and aperiodic. A primitive stochastic matrix always leads to a global convergence of the consensus system. The notions of irreducibility and primitivity have been extended to a completely positive map. We apply Theorem 3.6 and prove that an irreducible completely positive map is primitive if and only if the time-invariant consensus system is globally convergent, Proposition 5.6. By applying Burnside's theorem, we get that checking irreducibility reduces to deciding if a certain matrix algebra is the whole matrix algebra, Proposition 5.3. Therefore, irreducibility can be checked in polynomial time. By using a characterization of primitivity [SPGWC10], we deduce that for an irreducible matrix, global convergence to consensus can also be checked in polynomial time, Corollary 5.8.

The paper is organized as follows. In Section II we give the necessary preliminaries to study abstract Markov operators on arbitrary normal cones. We give in particular a representation of the extreme points of the dual unit ball with respect to the Hilbert's semi-norm. In Section III we show a convergence result on the iterates of a Markov operator given that the contraction rate of it is less than 1. We use the representation in Section II to characterize the contraction rate. Section IV is devoted to applications of the general results established in the previous sections to the non-commutative case. Some complexity results, relying on properties of irreducibility and primitivity, are presented in Section V.

## II. PRELIMINARIES

### A. Thompson's norm and Hilbert's semi-norm

Throughout the paper,  $(\mathcal{X}, \|\cdot\|)$  is a real Banach space. Let  $\mathcal{C} \subset \mathcal{X}$  be a closed pointed convex cone with non empty interior  $\mathcal{C}_0$ , i.e.,  $\alpha\mathcal{C} \subset \mathcal{C}$  for  $\alpha \in \mathbb{R}^+$ ,  $\mathcal{C} + \mathcal{C} \subset \mathcal{C}$  and  $\mathcal{C} \cap (-\mathcal{C}) = 0$ . We define the partial order  $\leq$  induced by  $\mathcal{C}$  on  $\mathcal{X}$  by  $x \leq y \Leftrightarrow y - x \in \mathcal{C}$ . We assume that the cone is *normal*, meaning that there exists a constant  $K > 0$  such that  $0 \leq x \leq y \Rightarrow \|x\| \leq K\|y\|$ . For  $x \leq y$  we define the order intervals:

$$[x, y] := \{z \in \mathcal{X} | x \leq z \leq y\}.$$

For  $x \in \mathcal{X}$  and  $y \in \mathcal{C}_0$ , following [Nus88], we define

$$\begin{aligned} M(x/y) &:= \inf\{t \in \mathbb{R} : x \leq ty\}, \\ m(x/y) &:= \sup\{t \in \mathbb{R} : x \geq ty\}, \\ \omega(x/y) &:= M(x/y) - m(x/y). \end{aligned} \quad (1)$$

Observe that since  $y \in \mathcal{C}_0$ , and since  $\mathcal{C}$  is closed and pointed, the two sets in (1) are non-empty, closed, and bounded from below and from above, respectively. In particular,  $m$  and  $M$  take finite values. In the sequel,  $\mathbf{e}$  denotes a distinguished element in the interior of  $\mathcal{C}$  (a "unit"). We define

$$\|x\|_T := \max(M(x/\mathbf{e}), -m(x/\mathbf{e}))$$

which we call Thompson's norm, and

$$\|x\|_H := \omega(x/\mathbf{e})$$

which we call Hilbert's semi-norm. It is shown in [Nus94] that the two norms  $\|\cdot\|$  and  $\|\cdot\|_T$  are equivalent.

*Remark 2.1:* These terminologies are motivated by the fact that the Thompson's part metric and the Hilbert's projective metric are Finsler's metrics (see [Nus94]) for which the infinitesimal distance at a point  $\mathbf{e} \in \mathcal{C}_0$  are respectively given by  $\|\cdot\|_T$  and  $\|\cdot\|_H$ . The semi-norm  $\|\cdot\|_H$  is also called Hopf oscillation.

*Example 2.2:* We consider the space  $\mathcal{X} = \mathbb{R}^n$ , the closed convex cone  $\mathcal{C} = \mathbb{R}_+^n$  and the unit element  $\mathbf{e} = \mathbf{1} := (1, \dots, 1)^T$ . It can be checked that the Thompson's norm is nothing but the sup norm

$$\|x\|_T = \max_i |x_i| = \|x\|_\infty,$$

whereas Hilbert's semi-norm is given by:

$$\|x\|_H = \max_{i,j} x_i - x_j = \Delta(x).$$

### B. Simplex in the dual space and dual unit ball

We define the dual space  $\mathcal{X}^*$  of  $\mathcal{X}$  with respect to the norm  $\|\cdot\|_T$  as the space of continuous linear functionals defined on the normed space  $(\mathcal{X}, \|\cdot\|_T)$ . The dual norm  $\|\cdot\|_T^*$  of a continuous linear functional  $z \in \mathcal{X}^*$  is defined by:

$$\|z\|_T^* := \sup_{\|x\|_T=1} \langle z, x \rangle.$$

The cone  $\mathcal{C}$  is still a closed cone in  $(\mathcal{X}, \|\cdot\|_T)$ . We define the dual cone of  $\mathcal{C}$  by:

$$\mathcal{C}^* := \{z \in \mathcal{X}^* | \langle z, x \rangle \geq 0, \forall x \in \mathcal{C}\}.$$

Since  $\mathcal{C}$  is a closed convex cone, we know that  $\mathcal{C} = \mathcal{C}^{**}$ , thus

$$x \in \mathcal{C} \Leftrightarrow \langle z, x \rangle \geq 0, \quad \forall z \in \mathcal{C}^*.$$

Then the Thompson's norm can be calculated by:

$$\|x\|_T = \sup_{z \in \mathcal{C}^*} \frac{|\langle z, x \rangle|}{\langle z, \mathbf{e} \rangle}. \quad (2)$$

We denote by

$$\mathcal{P} = \{\mu \in \mathcal{C}^* \mid \langle \mu, \mathbf{e} \rangle = 1\} \quad (3)$$

the "simplex" of the dual Banach space  $(\mathcal{X}^*, \|\cdot\|_T^*)$ .

*Remark 2.3:* In the case of Example 2.2, the dual space  $\mathcal{X}^*$  is  $\mathcal{X}$  itself and the dual norm  $\|\cdot\|_T^*$  is the  $l_1$  norm:

$$\|x\|_T^* = \sum_i |x_i| = \|x\|_1.$$

The simplex  $\mathcal{P}$  defined in (3) is the simplex in  $\mathbb{R}^n$  in the usual sense:

$$\mathcal{P} = \mathcal{P}_n := \{x \in \mathbb{R}_+^n : \sum_i x_i = 1\}.$$

The next lemma shows the connection between  $\mathcal{P}$  and the unit ball  $B_T^*$  of the space  $(\mathcal{X}^*, \|\cdot\|_T^*)$ . We denote by  $\text{conv}(\cdot)$  the convex hull of a set.

*Lemma 2.4:* The unit ball  $B_T^*$  of the space  $(\mathcal{X}^*, \|\cdot\|_T^*)$  satisfies

$$B_T^* = \text{conv}(\mathcal{P} \cup -\mathcal{P}) \quad (4)$$

*Proof:* It easily follows from (2) that

$$\|x\|_T = \sup_{\mu \in \mathcal{P}} |\langle \mu, x \rangle| = \sup_{\mu \in \mathcal{P} \cup -\mathcal{P}} \langle \mu, x \rangle. \quad (5)$$

Hence  $\|z\|_T^* \leq 1$  if and only if, for all  $x \in \mathcal{X}$ ,

$$\langle z, x \rangle \leq \|x\|_T = \sup_{\mu \in \mathcal{P} \cup -\mathcal{P}} \langle \mu, x \rangle. \quad (6)$$

By the strong separation theorem [FHH<sup>+</sup>01, Thm 3.18], if  $z$  did not belong to the closed convex hull  $\overline{\text{conv}}(\mathcal{P} \cup -\mathcal{P})$  (the closure is in the weak star topology of  $\mathcal{X}^*$ ), there would exist a vector  $x \in \mathcal{X}$  and a scalar  $\gamma$  such that  $\langle z, x \rangle > \gamma \geq \langle \mu, x \rangle$ , for all  $\mu \in \mathcal{P} \cup -\mathcal{P}$ , contradicting (6).

$$B_T^* = \overline{\text{conv}}(\mathcal{P} \cup -\mathcal{P}).$$

We claim that the latter closure operation can be dispensed with. Indeed, by the Banach Alaoglu theorem,  $B_T^*$  is weak-star compact. Hence, its subset  $\mathcal{P}$ , which is weak-star closed, is also weak-star compact. If  $\mu \in B_T^*$ , by the characterization of  $B_T^*$  above,  $\mu$  is a limit in the weak star topology, of a net  $\mu_a = s_a \nu_a - t_a \pi_a$  with  $s_a + t_a = 1$ ,  $s_a, t_a \geq 0$  and  $\nu_a, \pi_a \in \mathcal{P}$  for  $a \in \mathcal{A}$ . By passing to a subnet we can assume that  $s_a, t_a : a \in \mathcal{A}$  converge respectively to  $s, t \in [0, 1]$  such that  $s + t = 1$  and  $\nu_a, \pi_a : a \in \mathcal{A}$  converge respectively in the weak-star topology to  $\nu, \pi \in \mathcal{P}$ . It follows that  $\mu = s\nu - t\pi \in \text{conv}(\mathcal{P} \cup -\mathcal{P})$ . ■

### C. Extreme points of the dual unit ball in Hilbert's metric

We first show that the Hilbert semi-norm coincides with the norm on the quotient Banach space of  $(\mathcal{X}, 2\|\cdot\|_T)$  by the closed subspace  $\mathbb{R}\mathbf{e}$ .

*Lemma 2.5:* For all  $x \in \mathcal{X}$ , we have:

$$\|x\|_H = 2 \inf_{\lambda \in \mathbb{R}} \|x + \lambda \mathbf{e}\|_T$$

*Proof:*  $\|x + \lambda \mathbf{e}\|_T = \max((M(x/\mathbf{e}) + \lambda), (-m(x/\mathbf{e}) - \lambda))$  is minimal when  $(M(x/\mathbf{e}) + \lambda) = (-m(x/\mathbf{e}) - \lambda)$ . ■ A standard result of functional analysis shows that if  $\mathcal{W}$  is a closed subspace of a Banach space  $\mathcal{X}$ , then, the dual of the quotient space  $\mathcal{X}/\mathcal{W}$  can be identified isometrically to the space of continuous linear forms on  $\mathcal{X}$  that vanish on  $\mathcal{W}$ , equipped with the dual norm of  $\mathcal{X}^*$ . Specializing this result to  $\mathcal{W} = \mathbb{R}\mathbf{e}$ , we get:

*Lemma 2.6:* The dual space of  $(\mathcal{X}/\mathbb{R}\mathbf{e}, \|\cdot\|_H)$  is  $(\mathcal{M}, \|\cdot\|_H^*)$  where  $\mathcal{M} := \{\mu \in \mathcal{X}^* \mid \langle \mu, \mathbf{e} \rangle = 0\}$  and

$$\|\mu\|_H^* := \frac{1}{2} \|\mu\|_T^*, \quad \forall \mu \in \mathcal{M}. \quad (7)$$

The above lemma implies that the unit ball  $B_H^*$  of the space  $(\mathcal{M}, \|\cdot\|_H^*)$  satisfies:

$$B_H^* = 2B_T^* \cap \mathcal{M}.$$

*Remark 2.7:* In the case of  $\mathbb{R}^n$  equipped with its standard unit vector (Example 2.2), Lemma 2.6 implies that for any two probability measures  $\mu, \nu \in \mathcal{P}_n$ , the dual norm  $\|\mu - \nu\|_H^*$  is the total variation distance between  $\mu$  and  $\nu$ .

$$\|\mu - \nu\|_H^* = \frac{1}{2} \|\mu - \nu\|_1 = \|\mu - \nu\|_{TV}$$

We are now ready to give a representation of the extreme points of  $B_H^*$ . For all  $\nu, \pi \in \mathcal{C}^*$  we write  $\nu \perp \pi$  if  $\|\nu - \pi\|_H^* = 1$ . Let us give more insight on this notion in the finite dimensional case. Then, by compactness of the unit ball of  $(\mathcal{X}, \|\cdot\|_T)$ , we know that  $\|\nu - \pi\|_T^* = 2$  if and only if there is  $-e \leq y \leq e$  such that

$$\langle \pi - \nu, y \rangle = 2.$$

This is equivalent to that there is  $0 \leq x = \frac{y+e}{2} \leq e$  such that:

$$\langle \pi - \nu, x \rangle = 1.$$

Hence  $\mu \perp \nu$  if and only if there is  $0 \leq x \leq \mathbf{e}$  such that:

$$\langle \nu, x \rangle + \langle \pi, \mathbf{e} - x \rangle = 0.$$

We denote by  $\text{extr}(\cdot)$  the set of extreme points of a convex set.

*Proposition 2.8:* Any extreme point of  $B_H^*$  is of the form  $\nu - \pi$ , where  $\nu, \pi \in \text{extr} \mathcal{P}$  and  $\nu \perp \pi$ .

*Proof:* It follows from (4) that every point  $\mu \in B_T^*$  can be written as  $\mu = s\nu - t\pi$  with  $s + t = 1$ ,  $s, t \geq 0$ ,  $\nu, \pi \in \mathcal{P}$ . Moreover, if  $\mu \in \mathcal{M}$ ,  $0 = \langle \mu, \mathbf{e} \rangle = s\langle \nu, \mathbf{e} \rangle - t\langle \pi, \mathbf{e} \rangle = s - t$ , and so  $s = t = \frac{1}{2}$ . Thus every  $\mu \in B_T^* \cap \mathcal{M}$  can be written as

$$\mu = \frac{\nu - \pi}{2}, \quad \nu, \pi \in \mathcal{P}.$$

Assume now by contradiction that  $\nu$  is not extreme in  $\mathcal{P}$  (the case in which  $\pi$  is not extreme can be dealt with similarly).

Then, we can find  $\nu_1, \nu_2 \in \mathcal{P}$ ,  $\nu_1 \neq \nu_2$  and  $s_1, s_2 > 0$ ,  $s_1 + s_2 = 1$ , such that  $\nu = s_1\nu_1 + s_2\nu_2$ . It follows that  $\mu = s_1(\frac{\nu_1 - \pi}{2}) + s_2(\frac{\nu_2 - \pi}{2})$ , where  $\frac{\nu_1 - \pi}{2}, \frac{\nu_2 - \pi}{2}$  are distinct elements of  $B_T^* \cap \mathcal{M}$ , which is a contradiction.  $\blacksquare$

### III. CONTRACTION RATE OF MARKOV OPERATORS

A linear map  $T : \mathcal{X} \rightarrow \mathcal{X}$  is a *Markov operator* with respect to a unit vector  $\mathbf{e}$  in the interior  $\mathcal{C}_0$  of a closed convex pointed cone  $\mathcal{C} \subset \mathcal{X}$  if it satisfies the two following properties:

- (i)  $T$  is positive, i.e.,  $T(\mathcal{C}) \subset \mathcal{C}$ .
- (ii)  $T$  preserves the unit element  $\mathbf{e}$ , i.e.,  $T(\mathbf{e}) = \mathbf{e}$ .

The *adjoint operator*  $T^* : \mathcal{M} \rightarrow \mathcal{M}$  of  $T : \mathcal{X}/\mathbb{R}\mathbf{e} \rightarrow \mathcal{X}/\mathbb{R}\mathbf{e}$  is by definition:

$$\langle T^*(\mu), x \rangle = \langle \mu, T(x) \rangle, \quad \forall \mu \in \mathcal{M}, x \in \mathcal{X}.$$

*Example 3.1:* When  $\mathcal{X} = \mathbb{R}^n$  and  $\mathcal{C}$  is the standard positive cone (Example 2.2), a linear map  $T(x) = Ax$  is a Markov operator if and only if  $A$  is a row stochastic matrix.

We next extend to the case of Markov operators on cones a number of known properties in the case of Markov operators on the standard positive cone of  $\mathbb{R}^n$  (Markov matrices).

#### A. Duality between Hilbert's semi-norm and the total variation norm

We make the following basic observations for a Markov operator  $T : \mathcal{X} \rightarrow \mathcal{X}$ :

$$M(T(x)/\mathbf{e}) \leq M(x/\mathbf{e}), \quad m(T(x)/\mathbf{e}) \geq m(x/\mathbf{e}), \quad \forall x \in \mathcal{X}.$$

It is easily seen from the above observation that a Markov operator  $T$  is of Lipschitz constant less than 1 with respect to the Thompson's norm. Moreover, it induces a linear continuous application on the normed space  $(\mathcal{X}/\mathbb{R}\mathbf{e}, \|\cdot\|_H)$ , also of Lipschitz constant less than 1.

We are interested in the operator norm of  $T : \mathcal{X}/\mathbb{R}\mathbf{e} \rightarrow \mathcal{X}/\mathbb{R}\mathbf{e}$  with respect to the norm  $\|\cdot\|_H$  given by:

$$\|T\|_H := \sup_{\|x\|_H \leq 1} \|T(x)\|_H$$

The operator norm  $\|\cdot\|_H^*$  of  $T^*$  with respect to the dual norm  $\|\cdot\|_H^*$  is then defined by:

$$\|T^*\|_H^* := \sup_{\|z\|_H^* \leq 1} \|T^*(z)\|_H^*$$

A classical duality result (see [AB99, § 6.8]) shows that an operator and its adjoint have the same operator norm. In particular,

$$\text{Proposition 3.2: } \|T\|_H = \|T^*\|_H^*.$$

*Remark 3.3:* In the case of Example 3.1, the operator norm is contraction rate of the matrix  $A$  with respect to the function  $\Delta$ :

$$\|T\|_H = \tau(A) := \sup_x \frac{\Delta(Ax)}{\Delta(x)},$$

and the dual operator norm is the Lipschitz constant of  $A^T$  on  $\mathcal{P}_n$  with respect to the total variation distance:

$$\|T\|_H^* = \delta(A) := \sup_{\mu, \nu \in \mathcal{P}} \frac{\|A^T\mu - A^T\nu\|_{TV}}{\|\mu - \nu\|_{TV}}$$

#### B. Convergence of iterates to a rank one projector

We have seen that for a Markov operator  $T$ , the operator norm  $\|T\|_H$  satisfies  $\|T\|_H \leq 1$ . The case when  $\|T\|_H < 1$  or equivalently  $\|T^*\|_H^* < 1$  is of special interest, as shown by the following theorem, which shows that the iterates of  $T$  convergence to a rank one projector with a rate bounded by  $\|T\|_H$ .

*Theorem 3.4:* If  $\|T\|_H < 1$  or equivalently  $\|T^*\|_H^* < 1$ , then there is  $\pi \in \mathcal{P}$  such that for all  $x \in \mathcal{X}$

$$\|T^n(x) - \langle \pi, x \rangle \mathbf{e}\|_T \leq (\|T\|_H)^n \|x\|_H,$$

and for all  $\mu \in \mathcal{P}$

$$\|(T^*)^n(\mu) - \pi\|_H^* \leq (\|T\|_H)^n.$$

*Proof:* The intersection

$$\cap_n [m(T^n(x)/\mathbf{e}), M(T^n(x)/\mathbf{e})] \subset \mathbb{R}$$

is nonempty (as a non-increasing intersection of nonempty compact sets), and since  $\|T\|_H < 1$  and

$$\omega(T^n(x)/\mathbf{e}) \leq (\|T\|_H)^n \omega(x/\mathbf{e}),$$

this intersection must be reduced to a real  $\{c(x)\} \subset \mathbb{R}$  depending on  $x$ , i.e.,

$$c(x) = \cap_n [m(T^n(x)/\mathbf{e}), M(T^n(x)/\mathbf{e})].$$

Thus for all  $n \in \mathbb{N}$ ,

$$-\omega(T^n(x)/\mathbf{e}) \leq T^n(x) - c(x)\mathbf{e} \leq \omega(T^n(x)/\mathbf{e}).$$

Therefore by definition:

$$\|T^n(x) - c(x)\mathbf{e}\|_T \leq \omega(T^n(x)/\mathbf{e}).$$

Then we get:

$$\|T^n(x) - c(x)\mathbf{e}\|_T \leq (\|T\|_H)^n \|x\|_H.$$

It is immediate that:

$$c(x)\mathbf{e} = \lim_{n \rightarrow \infty} T^n(x)$$

from which we deduce that  $c : \mathcal{X} \rightarrow \mathbb{R}$  is a continuous linear functional. Thus there is  $\pi \in \mathcal{X}^*$  such that  $c(x) = \langle \pi, x \rangle$ . Besides it is immediate that  $\langle \pi, \mathbf{e} \rangle = 1$  and  $\pi \in \mathcal{C}^*$  because

$$x \in \mathcal{C} \Rightarrow c(x)\mathbf{e} \in \mathcal{C} \Rightarrow c(x) \geq 0 \Rightarrow \langle \pi, x \rangle \geq 0.$$

Therefore  $\pi \in \mathcal{P}$ . Finally for all  $\mu \in \mathcal{P}$  and all  $x \in \mathcal{X}$  we have

$$\begin{aligned} \langle (T^*)^n(\mu) - \pi, x \rangle &= \langle \mu, T^n(x) - \langle \pi, x \rangle \mathbf{e} \rangle \\ &\leq \|\mu\|_T^* \|T^n(x) - \langle \pi, x \rangle \mathbf{e}\|_T \\ &\leq (\|T\|_H)^n \|x\|_H. \end{aligned}$$

Hence

$$\|(T^*)^n(\mu) - \pi\|_H^* \leq (\|T\|_H)^n. \quad \blacksquare$$

*Example 3.5:* Specializing Theorem 3.4 to the case of Example 3.1 we obtain that if  $\tau(A) = \delta(A) < 1$ , then

$$A^n \rightarrow \mathbf{1}\pi^T, \quad n \rightarrow +\infty$$

where  $\pi$  is the unique invariant measure of the stochastic matrix  $A$ . This is a well-known result in the study of ergodicity property and mixing time of Markov chains, see for example [Sen91] and [LPW09].

### C. Dæbblin theorem for Markov operators on cones

A classical result, which goes back to Dæbblin and Doobrushin, characterizes the Lipschitz constant of a Markov matrix acting on the space of measures, with respect to the total variation norm (see the discussion in Example 3.7 below). We next give a characterization in the case of operators over an arbitrary cone, from which the classical results can be recovered, and new results on non-commutative consensus will follow.

*Theorem 3.6:* Let  $T$  be a Markov operator on  $\mathcal{X}$ . Then,

$$\|T\|_H = \|T^*\|_H^* = \frac{1}{2} \sup_{\substack{\nu, \pi \in \text{extr } \mathcal{P} \\ \nu \perp \pi}} \|T^*(\nu) - T^*(\pi)\|_H^*$$

*Proof:* We already noted that  $\|T\|_H = \|T^*\|_H^*$ . Moreover,

$$\|T^*\|_H^* = \sup_{\mu \in B_H^*} \|T^*(\mu)\|_H^*.$$

By the Banach-Alaoglu theorem,  $B_H^*$  is weak-star compact, and it is obviously convex. The dual space  $\mathcal{M}$  endowed with the weak-star topology is a locally convex topological space. Thus by the Krein-Milman theorem, the unit ball  $B_H^*$ , which is a compact convex set in  $\mathcal{M}$  with respect to the weak-star topology, is the closed convex hull of its extreme points. Moreover, the function

$$\mu \mapsto \|T^*(\mu)\|_H^* = \sup_{\|x\|_H \leq 1} \langle T^*(\mu), x \rangle = \sup_{\|x\|_H \leq 1} \langle \mu, T(x) \rangle$$

which is a sup of weak-star continuous maps is convex and weak-star lower semi-continuous. It follows that

$$\sup_{\mu \in B_H^*} \|T^*(\mu)\|_H^* = \sup_{\mu \in \text{extr}(B_H^*)} \|T^*(\mu)\|_H^*.$$

The announced formula now follows from the characterization of the extreme points of  $B_H^*$  in Proposition 2.8 and of the norm  $\|\cdot\|_H^*$  in Lemma 2.6. ■

*Example 3.7:* In the case of Example 3.1, Theorem 3.6 implies that:

$$\tau(A) = \delta(A) = \frac{1}{2} \sup_{i \neq j} \|A^T e_i - A^T e_j\|_1 \quad (8)$$

since the extreme points of the simplex  $\mathcal{P}_n$  are precisely the standard basis vectors. This is also a known result in the study of Markov chain [Sen91]. The value  $\tau(A)$  is known under the name of the Dobrushin's ergodic coefficient of the stochastic matrix  $A$  [Dob56]. It is explicitly given by:

$$\tau(A) = 1 - \min_{i,j} \sum_{s=1}^n \min(A_{is}, A_{js}). \quad (9)$$

A simple classical situation in which  $\tau(A) < 1$  is when there is a *Dæbblin state*, i.e., an element  $j \in \{1, \dots, n\}$  such that  $A_{ij} > 0$  holds for all  $i \in \{1, \dots, n\}$ .

### D. Hilbert's projective metric

The Hilbert's projective metric between two elements  $x, y \in \mathcal{C}_0$  is defined as:

$$d_H(x, y) = \log(M(x/y)/m(x/y)).$$

Consider a *positive Markov operator*  $T$ , i.e. a Markov operator satisfying

$$T(\mathcal{C}_0) \subset \mathcal{C}_0,$$

and

$$\alpha := \sup_{x \in \mathcal{C}_0} d_H(T(x), \mathbf{e})/d_H(x, \mathbf{e}).$$

The following theorem is a direct corollary of Nussbaum [Nus94].

*Theorem 3.8:* (Corollary of [Nus94, Thm2.3])

$$\|T\|_H = \lim_{\epsilon \rightarrow 0^+} \left( \sup \left\{ \frac{d_H(T(x), \mathbf{e})}{d_H(x, \mathbf{e})} : 0 < d_H(x, \mathbf{e}) \leq \epsilon \right\} \right),$$

By this theorem, it is clear that  $\|T\|_H \leq \alpha$ .

In [SSR10], the authors applied Birkhoff's contraction formula to give an upper bound on  $\alpha$ .

## IV. CONVERGENCE ANALYSIS OF CONSENSUS

Linear consensus in the space  $\mathcal{X}$  gives rise to the system:

$$x(t+1) = T_t(x(t)), \quad x(t) \in \mathcal{X}, \quad t = 1, 2, \dots,$$

where  $T_t$  is a Markov operator depending on  $t$ . The system reaches the consensus if the limit of  $x(t)$  tends to a term co-linear to the unit element  $\mathbf{e}$ .

### A. Consensus in $\mathbb{R}^n$

A linear consensus system in  $\mathbb{R}^n$  is given by a sequence of stochastic matrices  $\{A_t\}_{t \in \mathbb{N}}$ , see [VJAJ05], [Mor05], [SSR10]. There are a lot of results concerning the convergence condition and the convergence speed of the system, see [VJAJ05], [OT09], [Mor05], [AB09]. In particular, most of these papers relate the convergence to the connectivity of the union of graphs of  $\{A_t\}$  or the existence of a sequence of spanning trees. The formula (9) explains why there is a strong connection between the convergence of the consensus system and the connectivity of graphs. Namely, if in the graph corresponding to a stochastic matrix  $A$ , there is a node connected to other nodes, then  $\tau(A) < 1$ . And we have seen in Example 3.5 that  $\tau(A) < 1$  implies a global exponential convergence of the consensus system. For a more detailed description of this point of view, see [MDA05].

## B. Consensus in noncommutative spaces

We now specialize the previous results to the non commutative space  $\mathcal{X} = S_n$ , the set of  $n \times n$  Hermitian matrices. We consider the closed convex cone of positive semi-definite matrices  $\mathcal{C} = S_n^+$ . We chose the unit element  $e = I_n$ , the identity matrix. It can be checked that the Thompson's norm coincides with the operator norm:

$$\|X\|_T = \|X\|_\infty = \max_i \lambda_i(X),$$

and that the Hilbert's semi-norm is:

$$\|X\|_H = \max_{i,j} \lambda_i(X) - \lambda_j(X).$$

The dual space  $\mathcal{X}^*$  is equal to the space  $S_n$ . The dual norm  $\|\cdot\|_T^*$  is the trace norm:

$$\|X\|_T^* = \|X\|_1 = \sum_i |\lambda_i(X)|$$

The simplex defined in (3) is specified by the set of positive semi-definite matrices with trace 1:

$$\mathcal{P} = \mathcal{Q}_n := \{X \in S_n^+ \mid \text{trace}(X) = 1\}.$$

In this case the simplex is the set of all density matrices, the analogue of the set of probability distributions in  $\mathbb{R}^n$ . For any two density matrices  $U, V \in \mathcal{Q}_n$ , the norm  $\|U - V\|_H^*$  is the total variation distance between the two mixed states:

$$\|U - V\|_H^* = \|U - V\|_{TV} := \frac{1}{2} \|U - V\|_1$$

In the sequel we consider a unital completely positive map:

$$\Phi(X) = \sum_{i=1}^m V_i^* X V_i, \quad (10)$$

where the matrices  $V_i$  satisfy the condition

$$\sum_{i=1}^m V_i^* V_i = I_n. \quad (11)$$

This is a generalization of stochastic matrices to the case of operators on the cone of symmetric matrices. The dual operator  $\Psi$  is written as:

$$\Psi(Z) = \sum_{i=1}^m V_i Z V_i^*$$

The dual operator  $\Psi$  is completely positive and trace-preserving.

We apply Theorem 3.6 to this case and obtain the analogue of the formula (8):

*Corollary 4.1:*

$$\begin{aligned} \|\Phi\|_H &= \|\Psi\|_H^* = \frac{1}{2} \sup_{u,v \in S, u \perp v} \|\Psi(uu^*) - \Psi(vv^*)\|_1 \\ &= 1 - \min_{\substack{X: XX^* = I_n \\ u,v \in B_n, u \perp v}} \sum_{i=1}^n \min\{X_i^* \Psi(uu^*) X_i, X_i^* \Psi(vv^*) X_i\} \end{aligned}$$

where  $B_n = \{u \in \mathbb{C}^n \mid \|u\| = 1\}$ ,  $u \perp v$  means  $u^*v = 0$  and  $X_i$  is the  $i$ -th column vector of the unitary matrix  $X$ .

*Remark 4.2:* For the noncommutative case, it is not evident whether more effective characterizations of the contraction rate exists. Note that the contraction ratio of  $\Psi$  over the set of density matrices  $\mathcal{P}_0$  with respect to the trace norm distance, was studied in quantum information theory, see [RKW11] and references therein. They provided a Birkhoff type upper bound (Corollary 9 in [RKW11]):

$$\|\Psi\|_H^* \leq \tanh(\text{diam } \Psi/4)$$

where

$$\text{diam } \Psi := \sup\{d_H(\Psi(X), \Psi(Y)) : X, Y \in S_n^+\}$$

The value  $\text{diam } \Psi$  is not directly computable. Note that, this upper bound is equal to 1 if and only if  $\text{diam } \Psi = \infty$ , which is satisfied if and only if there exists  $u, v \in S$  such that:

$$\text{span}\{V_i u : 1 \leq i \leq m\} \neq \text{span}\{V_i v : i \leq i \leq m\}.$$

We next provide a much tighter, in fact necessary and sufficient, condition for the Lipschitz constant to be 1.

*Corollary 4.3:* The following conditions are equivalent:

1.  $\|\Phi\|_H = \|\Psi\|_H^* = 1$ .
2. There is  $u, v \in B_n$  such that

$$\langle V_i u, V_j v \rangle = 0, \quad \forall i, j \in \{1, \dots, m\}.$$

3. There is a  $n \times n$  rank one matrix  $Y$  such that

$$\text{trace}(V_i^* V_j Y) = 0, \quad \forall i, j \in \{1, \dots, m\}.$$

*Proof:* From Corollary 4.1 we know that  $\|\Phi\|_H = 1$  if and only if there exists  $u, v \in B_n$  such that

$$\left\| \sum_{i=1}^m V_i u u^* V_i^* - \sum_{i=1}^m V_i v v^* V_i^* \right\|_{TV} = 1.$$

It can be checked that this holds if and only if for any  $i, j \in \{1, \dots, m\}$ ,

$$\langle V_i u, V_j v \rangle = 0.$$

Hence the rank one matrix is  $Y = v u^*$ . ■

*Remark 4.4:* The condition that there are two vectors  $u, v \in B_n$  such that  $\|\Psi(uu^*) - \Psi(vv^*)\|_{TV} = 1$ , is equivalent to say, in the quantum channel context, that there exists two pure states which the quantum channel  $\Psi$  takes into two orthogonal subspaces. This distinguishability of two quantum states is a fundamental property of quantum systems [NC00]. In [MA05] the authors extended the definition of zero-error capacity for memoryless quantum channel. In particular they showed that the quantum zero-error capacity is greater than 0 if and only if there exist two pure states distinguishable.

## C. Convergence of noncommutative consensus

We consider a time-invariant noncommutation consensus dynamics [SSR10]:

$$X_{t+1} = \Phi(X_t), \quad X_t \in S_n. \quad (12)$$

where  $\Phi$  is a completely positive map preserving the identity matrix, given in (10). We provide an equivalent condition to the global convergence to consensus of system (12).

Let us consider the sequence of matrix subspaces over  $\mathbb{C}$ :

$$\mathcal{H}_{k+1} = \text{span}\{V_i^* X V_j : X \in \mathcal{H}_k, i, j = 1, \dots, m\}.$$

with  $\mathcal{H}_0 = \text{span}\{I_n\}$ . For all  $k$ , denote the orthogonal space of  $\mathcal{H}_k$  by  $\mathcal{G}_k$ . By (11), we obtain that  $\mathcal{H}_{k+1} \supseteq \mathcal{H}_k$  for all  $k \in \mathbb{N}$ . Therefore the dimension of  $\mathcal{H}_k$  is at least  $1 + k$ . Hence there is  $k_0 \leq n^2 - 1$  such that

$$\mathcal{H}_{k_0+s} = \mathcal{H}_{k_0}, \quad \forall s \in \mathbb{N}.$$

Hence we have:

$$\mathcal{G}_k \supseteq \mathcal{G}_{k+1}, \quad \mathcal{G}_{k_0} = \mathcal{G}_{k_0+s}, \quad \forall k, s \in \mathbb{N} \quad (13)$$

*Theorem 4.5:* The following conditions are equivalent:

- (1) There exists  $k$  such that  $\|\Phi^k\|_H < 1$ .
- (2) Every orbit of the system (12) converges to an equilibrium co-linear to  $I_n$ .
- (3) The subspace  $\bigcap_{k \in \mathbb{N}} \mathcal{G}_k$  does not contain a rank one matrix.
- (4) There exists  $k_0 \leq n^2 - 1$  such that  $\|\Phi^{k_0}\|_H < 1$ .

*Proof:*

(1)  $\Rightarrow$  (2): We apply Theorem 3.4 to the application  $\Phi^k$ .

(2)  $\Rightarrow$  (1): The Hilbert's semi-norm defines a norm in the orthogonal space to the identity matrix  $I$ . It follows from Gelfand's formula that

$$\lim_{k \rightarrow +\infty} \|\Phi^k\|_H^{1/k} = \max\{|\lambda| : \Phi(X) = \lambda X, X \in S_n, X \perp I\}$$

Thus if (1) is not true, then there is  $X \in S_n$  and  $\lambda \in \mathbb{C}$  such that  $|\lambda| = 1$ ,  $X \perp I$  and  $\Phi^n(X) = \lambda^n X$ . The system is therefore not globally convergent to an equilibrium co-linear to  $I_n$ .

(3)  $\Leftrightarrow$  (1): Note that for all  $k \in \mathbb{N}$ ,

$$\Phi^k(X) = \sum_{i_1, \dots, i_k} V_{i_k}^* \dots V_{i_1}^* X V_{i_1} \dots V_{i_k}.$$

By Corollary 4.3, we know that  $\|\Phi^k\|_H = 1$  if and only if the subspace  $\mathcal{G}_k$  contains a rank one matrix. Therefore by (13)  $\|\Phi^k\|_H = 1$  for all  $k \in \mathbb{N}$  if and only if the subspace  $\bigcap_{k \in \mathbb{N}} \mathcal{G}_k$  contains a rank one matrix.

(3)  $\Rightarrow$  (4): By (13), there is  $k_0 \leq n^2 - 1$  such that  $\mathcal{G}_{k_0} = \bigcap_{k \in \mathbb{N}} \mathcal{G}_k$ . It follows from Corollary 4.3 that  $\|\Phi^{k_0}\| < 1$  if (3) is true.  $\blacksquare$

*Remark 4.6:* A sufficient condition for the global convergence of the system (12) would be that there is  $k_0 \leq n^2 - 1$  such that

$$\mathcal{H}_{k_0} = \mathcal{M}_n(\mathbb{C}).$$

## V. IRREDUCIBILITY, PRIMITIVITY AND A COMPLEXITY RESULT

In the case of Markov operators on  $\mathbb{R}^n$ , tight conditions for the convergence to consensus can be expressed combinatorially in terms of adjacency graphs. In particular, a well-known result shows that an irreducible positive matrix is primitive (i.e., has a power sending the standard positive cone to its interior) if and only if its adjacency graph is

aperiodic (meaning that the gcd of the lengths of circuits of the adjacency graph is 1). In this section, we present several analogous results in the noncommutative case, and analyze the complexity of associated decision problems.

### A. Irreducibility and Primitivity of a complete positive map

We consider a unital completely positive map written in the form of (10). We denote by  $S_k(\Phi)$  the linear space spanned by all the products of  $k$  elements  $\{V_i\}_{i=1, \dots, m}$ , and by  $D_k(\Phi)$  the linear space spanned by all the products of at most  $k$  elements of  $\{V_i\}_{i=1, \dots, m}$ . We denote by  $\mathcal{A}(\Phi) = \bigcup_{k \geq 1} D_k(\Phi)$  the algebra generated by  $\{V_i : i = 1, \dots, m\}$ .

*Lemma 5.1:* There is  $p \leq n^2$  such that  $\mathcal{A}(\Phi) = D_p(\Phi)$ .

*Proof:* Since  $D_{s+1} \supset D_s$  for all  $s \in \mathbb{N}$ . There is  $p \leq n^2$  such that  $D_s = D_p$  for all  $s \geq p$ .  $\blacksquare$

We next give some definitions analogous to the  $\mathbb{R}^n$  case. Usually, they are given for the Kraus map  $\Psi$ . It is equivalent to define them for the adjoint unital map  $\Phi$ .

*Definition 5.2 (Irreducibility [Far96]):* The map  $\Phi$  is irreducible if there is no face of  $S_n^+$  invariant by  $\Phi$ , where a face  $\mathcal{F}$  of  $S_n^+$  is a closed cone contained in  $S_n^+$  such that if  $P \in \mathcal{F}$  then  $Q \in \mathcal{F}$  for all  $Q \leq P$ .

*Proposition 5.3:* The map  $\Phi$  is irreducible if and only if the algebra  $\mathcal{A}(\Phi)$  is the whole  $n \times n$  matrix algebra.

*Proof:* It was shown [Far96, Theorem 2] that the reducibility is equivalent to the existence of a non-trivial (other than  $\{0\}$  or  $\mathbb{C}^n$ ) common invariant subspace of all  $\{V_i\}$ . By Burnside's theorem, the latter property holds if and only if the algebra  $\mathcal{A}(\Phi)$  is not the whole matrix space.  $\blacksquare$

*Definition 5.4 (Strict positivity):* The map  $\Phi$  is strictly positive if for all  $X \geq 0$ ,  $\Phi(X) > 0$ .

*Definition 5.5 (Primitivity [SPGWC10]):* The map  $\Phi$  is primitive if there is  $p > 0$  such that  $\Phi^p$  is strictly positive.

### B. Characterization of primitivity assuming irreducibility

The following proposition is analogous in the non-commutative case to that a irreducible matrix is primitive if and only if is aperiodic in the  $\mathbb{R}^n$  case.

*Proposition 5.6:* If  $\Phi$  is irreducible, then the system (12) is globally convergent if and only if  $\Phi$  is primitive.

*Proof:* If  $\Phi$  is primitive then there is  $k$  such that  $\|\Phi^k\|_H < 1$ . By Theorem 3.4 the system (12) converges globally. If there is  $k > 0$  such that  $\alpha = \|\Phi^k\|_H < 1$  then by Theorem 3.6 there is unique invariant density matrix  $\Pi \in \mathcal{Q}_n$  of  $\Psi^k$  such that for all  $P \in \mathcal{Q}_n$ ,

$$\|\Psi^{nk}(P) - \Pi\|_H^* \leq \alpha^n, \quad \forall n \geq 0.$$

Therefore  $\Pi$  is also the unique invariant density matrix of  $\Psi$ . We deduce that  $\Pi$  is of full rank by the irreducibility of  $\Psi$ . Again by Theorem 3.6, for all  $x \in \mathbb{C}^n$  with norm equal to 1,

$$\|\Phi^{nk}(xx^*) - (x^* \Pi x) I_n\|_T \leq \alpha^n.$$

That is,

$$-\alpha^n I_n \leq \Phi^{nk}(xx^*) - (x^* \Pi x) I_n \leq \alpha^n I_n$$

Since  $\Pi$  is a density matrix of full rank, there is  $n_0$  such that for all  $x \in \mathbb{C}^n$ ,

$$\Phi^{n_0 k}(xx^*) \geq (\lambda_{\min}(\Pi) - \alpha^{n_0})I_n > 0.$$

Thus  $\Phi$  is primitive. ■

We shall use a characterization of primitivity given in [SPGWC10].

*Theorem 5.7 ([SPGWC10]):* A unital completely positive map  $\Phi$  is primitive if and only if there is  $q \leq (n^2 - m + 1)n^2$  such that the space  $S_q(\Phi)$  is of dimension  $n^2$ .

*Corollary 5.8:* Let  $\Phi$  be a unital completely positive map given in (10), with rational components. Then checking whether  $\Phi$  is irreducible can be done in polynomial time. If  $\Phi$  is irreducible, then checking whether the system (12) is globally convergent can be done in polynomial time.

*Proof:* To decide if  $\Phi$  is irreducible, we shall compute the increasing sequence of matrix subspaces  $D_s(\Phi)$ ,  $s = 1, 2, \dots$ , and look for the first integer  $k$  such that  $D_k(\Phi) = D_{k+1}(\Phi)$ . For a given  $s$ , we shall represent  $D_s(\Phi)$  by a basis,  $D_s(\Phi) = \text{span}\{M_1, \dots, M_l\}$  where  $M_i$  are all linearly independent. Recall that extracting a basis from a family of rational vectors can be done in polynomial time in the bit model. It follows that a basis representation of

$$D_{s+1}(\Phi) = \text{span}\{M_j, V_i M_j : 1 \leq i \leq m, 1 \leq j \leq l\}.$$

can be computed in polynomial time from a basis representation of  $D_s(\Phi)$ . Moreover,  $k < n^2$ . Hence, every basis element arising in the representation of  $D_s(\Phi)$  is obtained by  $O(n^2)$  matrix products. It follows that the number of bits needed to code the basis elements remain polynomially bounded in the length of the input. Hence, a basis representation of the algebra  $\mathcal{A}(\Phi) = \cup_{s \geq 1} D_s(\Phi)$  can be obtained in polynomial time.

Suppose now that  $\Phi$  is irreducible. Then by Proposition 5.6, deciding if the system (12) is globally convergent reduces to checking whether  $\Phi$  is primitive. By Theorem 5.7,  $\Phi$  is primitive if and only if  $S_q(\Phi)$  is of dimension  $n^2$  for some  $q \leq (n^2 - m + 1)n^2$ . Arguing as above, a basis representation of  $S_q(\Phi)$  can be computed in polynomial time. ■

*Remark 5.9:* A natural method to decide whether  $\Phi^k$  is a contraction for  $k$  large enough, would be to check whether 1 is the only eigenvalue of  $\Phi$  on the unit circle and if it is algebraically simple. However, doing so in exact arithmetics appear to be not so tractable, whereas Corollary 5.8 leads to a polynomial algorithm in the bit model, when the map  $\Phi$  is irreducible. When  $\Phi$  is reducible, we do not know the algorithmic complexity of checking the conditions of Theorem 4.5. A referee pointed out that reducible maps do occur in practical applications in which one needs to drive a quantum system to a pure state. The application to noncommutative consensus over infinite dimensional Hilbert spaces also remains to be developed.

*Acknowledgement.* We thank the referees for their insightful remarks and suggestions.

## REFERENCES

- [AB99] C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis. A Hitchiker's Guide*. Springer, 1999.
- [AB09] David Angeli and Pierre-Alexandre Bliman. Convergence speed of unsteady distributed consensus: decay estimate along the settling spanning-trees. *SIAM J. Control Optim.*, 48(1):1–32, 2009.
- [BGPS06] Stephen Boyd, Arpita Ghosh, Balaji Prabhakar, and Devavrat Shah. Randomized gossip algorithms. *IEEE Trans. Inform. Theory*, 52(6):2508–2530, 2006.
- [BT89] Dimitri P. Bertsekas and John N. Tsitsiklis. *Parallel and distributed computation: numerical methods*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1989.
- [Dob56] R. Dobrushin. Central limit theorem for non-stationary Markov chains. I. *Teor. Veroyatnost. i Primenen.*, 1:72–89, 1956.
- [Far96] D. R. Farenick. Irreducible positive linear maps on operator algebras. *Proc. Amer. Math. Soc.*, 124(11):3381–3390, 1996.
- [FHH<sup>+</sup>01] Marián Fabian, Petr Habala, Petr Hájek, Vicente Montesinos Santalucía, Jan Pelant, and Václav Zizler. *Functional analysis and infinite-dimensional geometry*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 8. Springer-Verlag, New York, 2001.
- [Hir89] Morris W. Hirsch. Convergent activation dynamics in continuous time networks. *Neural Networks*, 2(5):331–349, 1989.
- [LPW09] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.
- [MA05] Rex A. C. Medeiros and Francisco M. De Assis. Quantum zero-error capacity. *Int. J. Quantum Inform.*, 03:135, 2005.
- [MDA05] Cao M., Spielman D., A., and Morse A., S. A lower bound on convergence of a distributed network consensus algorithm. In *Proc. of the joint 44th IEEE Conference on Decision and Control and European Control Conference*, pages 2356–2361. IEEE, 2005.
- [Mor05] Luc Moreau. Stability of multiagent systems with time-dependent communication links. *IEEE Trans. Automat. Control*, 50(2):169–182, 2005.
- [NC00] Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, Cambridge, 2000.
- [Nus88] R. D. Nussbaum. Hilbert's projective metric and iterated nonlinear maps. *Mem. Amer. Math. Soc.*, 75(391):iv+137, 1988.
- [Nus94] Roger D. Nussbaum. Finsler structures for the part metric and Hilbert's projective metric and applications to ordinary differential equations. *Differential Integral Equations*, 7(5-6):1649–1707, 1994.
- [OT09] Alex Olshevsky and John N. Tsitsiklis. Convergence speed in distributed consensus and averaging. *SIAM J. Control Optim.*, 48(1):33–55, 2009.
- [RKW11] David Reeb, Michael J. Kastoryano, and Michael M. Wolf. Hilbert's projective metric in quantum information theory. *J. Math. Phys.*, 52(8):082201, 33, 2011.
- [Sen91] E. Seneta. Applications of ergodicity coefficients to homogeneous Markov chains. In *50 Years after Doeblin: Developments in the Theory of Markov Chains, Markov Processes and Sums of Random Variables*, Blaubeuren, Germany, 2-7 November 1991.
- [SPGWC10] Mikel Sanz, David Pérez-García, Michael M. Wolf, and Juan I. Cirac. A quantum version of Wielandt's inequality. *IEEE Trans. Inf. Theory*, 56(9):4668–4673, September 2010.
- [SSR10] Rodolphe Sepulchre, Alain Sarlette, and Pierre Rouchon. Consensus in non-commutative spaces. In *Proceedings of the 49th IEEE Conference on Decision and Control*, pages 6596–6601, Atlanta, USA, Dec 2010.
- [VJAJ05] D. Blondel Vincent, Hendrickx Julien, M., Olshevsky Alex, and Tsitsiklis John, N. Convergence in multiagent coordination, consensus, and flocking. In *Proceedings of the joint 44th IEEE Conference on Decision and Control and European Control Conference*, pages 2996–3000. IEEE, 2005.