CrossMark

Invariant measure for quantum trajectories

T. Benoist[1] · M. Fraas[2] · Y. Pautrat[3] · C. Pellegrini[1]

Abstract

We study a class of Markov chains that model the evolution of a quantum system subject to repeated measurements. Each Markov chain in this class is defined by a measure on the space of matrices, and is then given by a random product of correlated matrices taken from the support of the defining measure. We give natural conditions on this support that imply that the Markov chain admits a unique invariant probability measure. We moreover prove the geometric convergence towards this invariant measure in the Wasserstein metric. Standard techniques from the theory of products of random matrices cannot be applied under our assumptions, and new techniques are developed, such as maximum likelihood-type estimations.

Mathematics Subject Classification 60J05 · 81Q99 (Discrete time Markov chains and Quantum Theory General)

Contents

✉ C. Pellegrini
clement.pellegrini@math.univ-toulouse.fr

T. Benoist
tristan.benoist@math.univ-toulouse.fr

M. Fraas
martin.fraas@gmail.com

Y. Pautrat
yan.pautrat@math.u-psud.fr

[1] Institut de Mathématiques de Toulouse, UMR5219, Université de Toulouse, CNRS, UPS, 31062 Toulouse Cedex 9, France

[2] Instituut voor Theoretische Fysica, KU Leuven, 3001 Leuven, Belgium

[3] Laboratoire de Mathématiques d'Orsay, Univ. Paris-Sud, CNRS, Université Paris-Saclay, 91405 Orsay, France

🖄 Springer

1 Introduction

We consider a complex vector space $\mathbb{C}^k$ and its projective space $\mathrm{P}(\mathbb{C}^k)$ equipped with its Borel σ-algebra $\mathcal{B}$. For a nonzero vector $x \in \mathbb{C}^k$, we denote $\hat{x}$ the corresponding equivalence class of x in $\mathrm{P}(\mathbb{C}^k)$. For a linear map $v \in \mathrm{M}_k(\mathbb{C})$ we denote $v \cdot \hat{x}$ the element of the projective space represented by $v\,x$ whenever $v\,x \neq 0$. We equip $\mathrm{M}_k(\mathbb{C})$ with its Borel σ-algebra and let μ be a measure on $\mathrm{M}_k(\mathbb{C})$ with a finite second moment, $\int_{\mathrm{M}_k(\mathbb{C})} \|v\|^2 \, \mathrm{d}\mu(v) < \infty$, that satisfies the stochasticity condition

$$\int_{\mathrm{M}_k(\mathbb{C})} v^* v \, \mathrm{d}\mu(v) = \mathrm{Id}_{\mathbb{C}^k} \tag{1}$$

(we discuss this condition below).

In this article we are interested in particular Markov chains $(\hat{x}_n)$ on $\mathrm{P}(\mathbb{C}^k)$, defined by

$$\hat{x}_{n+1} = V_n \cdot \hat{x}_n,$$

where V_n is an $\mathrm{M}_k(\mathbb{C})$-valued random variable with a probability density $\|vx_n\|^2/\|x_n\|^2 \, \mathrm{d}\mu(v)$. Condition (1) ensures this is a probability density for any $x_n \neq 0$. More precisely, such a Markov chain is associated with the transition kernel given for a set $S \in \mathcal{B}$ and $\hat{x} \in \mathrm{P}(\mathbb{C}^k)$ by

$$\Pi(\hat{x}, S) = \int_{\mathrm{M}_k(\mathbb{C})} \mathbf{1}_S \left(v \cdot \hat{x} \right) \|vx\|^2 \mathrm{d}\mu(v), \tag{2}$$

where x is an arbitrary normalized vector representative of $\hat{x}$. Moreover, the event $\{vx = 0\}$ always has probability 0, hence the Markov chain is well-defined on $\mathrm{P}(\mathbb{C}^k)$. We recall that for any probability measure ν, $\nu\Pi$ is the probability measure defined by

$$\nu\Pi(S) = \int_{\mathrm{P}(\mathbb{C}^k)} \Pi\left(\hat{x}, S\right) \mathrm{d}\nu(\hat{x})$$

for any $S \in \mathcal{B}$. A measure ν is called invariant if $\nu\Pi = \nu$.

We are interested in the large-time distribution of $(\hat{x}_n)$. Note that $\hat{x}_n$ can be written as

$$\hat{x}_n = V_n \dots V_1 \cdot \hat{x}_0,$$

so that the study of $\hat{x}_n$ can be formulated in terms of random products of matrices. Markov chains associated to random products of matrices have been studied in a more general setting where the weight appearing in the transition kernel (2) is proportional to $\|vx\|^s$ for some $s \geq 0$, instead of $\|vx\|^2$. The classical case of products of independent, identically distributed random matrices pioneered by Kesten, Furstenberg and Guivarc'h corresponds to $s = 0$. In that case, for i.i.d. invertible random matrices $Y_1, Y_2, \ldots$, denoting $S_n = Y_n \dots Y_1$, one is usually interested in the asymptotic properties of

$$\log \|S_n x\|,$$

for any $x \neq 0$. In particular, a law of large numbers, a central limit theorem and a large deviation principle have been obtained for this quantity, under contractivity and strong irreducibility assumptions [8,11,16]. Such results are closely linked to the uniqueness of the invariant measure of the Markov chain

$$\hat{x}_n = S_n \cdot \hat{x}.$$

These results were generalized to the case $s > 0$ in [10]. Our framework corresponds to the case $s = 2$; in this case, and with the additional assumption (1), we provide a new method to study this Markov chain, and use it to derive the above results without assuming invertibility of the matrices, and with an optimal irreducibility assumption. We compare our approach with respect to that of [10] at the end of this section.

The method that we employ is motivated by an interpretation of this process as statistics of a quantum system being repeatedly indirectly measured. Let us expand on this as we introduce more notation and terminology. The set of states of a quantum system described by a finite dimensional Hilbert space $\mathbb{C}^k$ is the set of density matrices $\mathcal{D}_k := \{\rho \in M_k(\mathbb{C}) \mid \rho \geq 0, \ \mathrm{tr}\,\rho = 1\}$. This set is convex and the set of its extreme points is called the set of pure states. This latter set is in one to one correspondence with the projective space $P(\mathbb{C}^k)$ by the bijection $P(\mathbb{C}^k) \ni \hat{x} \mapsto \pi_{\hat{x}} \in \mathcal{D}_k$ with $\pi_{\hat{x}}$ the orthogonal projector on the corresponding ray in $\mathbb{C}^k$. The time evolution of the system conditioned on a measurement outcome is encoded in a matrix v that updates the state of the system. The support of μ is endowed with the meaning of the possible updates, and the system is updated according to v with a probability density $\mathrm{tr}(v\rho v^*)\,\mathrm{d}\mu(v)$. Given v, a state ρ is mapped to a state $v\rho v^*/\mathrm{tr}(v\rho v^*)$. Iterating this procedure defines a random sequence (ρ_n) in $\mathcal{D}_k$ called a quantum trajectory: after n measurements with resulting matrices $v_1, \dots, v_n$ the state of the system becomes

$$\rho_n = \frac{v_n \dots v_1 \rho_0 v_1^* \dots v_n^*}{\mathrm{tr}\left(v_n \dots v_1 \rho_0 v_1^* \dots v_n^*\right)} \tag{3}$$

where $(v_1, \dots, v_n)$ has probability density $\mathrm{tr}(v_n \dots v_1 \rho_0 v_1^* \dots v_n^*)\,\mathrm{d}\mu^{\otimes n}(v_1, \dots, v_n)$. In other words, the process Eq. (3) describes an evolution of a repeatedly measured quantum system.

A key result in the theory of quantum trajectories is the purification theorem obtained by Kümmerer and Maassen [17] showing that quantum trajectories (ρ_n) defined on $\mathcal{D}_k$ almost surely approach the set of pure states (which are the extreme points of $\mathcal{D}_k$) if and only if the following purification condition is satisfied:

(Pur): Any orthogonal projector π such that for any $n \in \mathbb{N}$, $\pi v_1^* \ldots v_n^* v_n \ldots v_1 \pi \propto \pi$ for $\mu^{\otimes n}$-almost all $(v_1, \ldots, v_n)$, is of rank one

(we write $X \propto Y$ for X, Y two operators if there exists $\lambda \in \mathbb{C}$ such that $X = \lambda Y$).

Under this assumption, the long-time behavior of the Markov chain is essentially dictated by its form on the set of pure states, i.e. for $\rho_0 = \pi_{\hat{x}_0}$. It is an immediate observation that

$$\mathrm{tr}\left(v\pi_{\hat{x}_0}v^*\right) = \|vx_0\|^2, \quad \frac{v\pi_{\hat{x}_0}v^*}{\mathrm{tr}\left(v\pi_{\hat{x}_0}v^*\right)} = \pi_{v\cdot\hat{x}_0}, \tag{4}$$

for all $v \in \mathrm{M}_k(\mathbb{C})$. This way our Markov chain $(\hat{x}_n)$ corresponds to the quantum trajectory (ρ_n) described above when ρ_0 is a pure state $\pi_{\hat{x}_0}$.

Although ideas underlying our method are based on the connection of $(\hat{x}_n)$ with this physical problem, we will not explicitly use it in the paper. The notion of quantum trajectory originates in quantum optics [6], and Haroche's Nobel prize winning experiment [9] is arguably the most prominent example of a system described by the above formalism. The reader interested in the involved mathematical structures might consult for example the review book [13] or the pioneering articles [14,15,17].

We will show that under the condition **(Pur)**, the set of all invariant measures of the Markov chain (3) can be completely classified, depending on the operator ϕ on $\mathcal{D}_k$ describing the average evolution:

$$\phi(\rho) = \int_{\mathrm{M}_k(\mathbb{C})} v\rho v^* \, \mathrm{d}\mu(v). \tag{5}$$

The map ϕ on $\mathcal{D}_k$ is completely positive and trace-preserving.[1] Such a map is often called a quantum channel (see e.g. [22]). It has in particular the property of mapping states to states. Brouwer's fixed point theorem shows that there exists an invariant state, i.e. $\rho \in \mathcal{D}_k$ such that $\phi(\rho) = \rho$. A necessary and sufficient algebraic condition for uniqueness of this invariant state is (see e.g. [5,7,22])

(ϕ-Erg): There exists a unique minimal non trivial subspace E of $\mathbb{C}^k$ such that $\forall v \in \mathrm{supp}\,\mu$, $vE \subset E$.

If **(ϕ-Erg)** holds with $E = \mathbb{C}^k$, then ϕ is said to be irreducible. We chose the name **(ϕ-Erg)** to avoid confusion with the notion of irreducibility for Markov chains. We moreover emphasize that we call this assumption **(ϕ-Erg)** because it relies only on ϕ and not on the different operators v in the support of μ: an equivalent statement of **(ϕ-Erg)** is that there exists a unique minimal nonzero orthogonal projector π such that $\phi(\pi) \leq \lambda\pi$ for some $\lambda \geq 0$ (see e.g. [20]).

We now state the main result of the paper:

---

[1] Complete positivity is stronger than positivity; namely by definition ϕ is completely positive iff $\phi \otimes \mathrm{Id}_{M_n(\mathbb{C})}$ is positive for all $n \in \mathbb{N}$.

Theorem 1.1 *Assume that μ satisfies assumptions* **(Pur)** *and* **(ϕ-Erg)**. *Then, the transition kernel Π has a unique invariant probability measure ν_{inv} and there exist $m \in \{1, \ldots, k\}$, $C > 0$ and $0 < \lambda < 1$ such that for any probability measure ν over $(\text{P}(\mathbb{C}^k), \mathcal{B})$,*

$$W_1 \left(\frac{1}{m} \sum_{r=0}^{m-1} \nu \Pi^{mn+r}, \nu_{\text{inv}} \right) \le C\lambda^n, \tag{6}$$

where W_1 is the Wasserstein metric of order 1.

The Wasserstein metric is constructed with respect to a natural metric on the complex projective space. This metric is defined, for $\hat{x}$, $\hat{y}$ in $\text{P}(\mathbb{C}^k)$, by

$$d\left(\hat{x}, \hat{y}\right) = \left(1 - |\langle x, y\rangle|^2\right)^{\frac{1}{2}}, \tag{7}$$

where x, y are unit length representative vectors of $\hat{x}$, $\hat{y}$, and $\langle \cdot, \cdot \rangle$ is the canonical hermitian inner product on $\mathbb{C}^k$.

Let us now compare our results to those of the article [10] of Guivarc'h and Le Page. They consider a probability distribution μ with support in $\text{GL}_k(\mathbb{C})$, without requiring the normalization condition (1), and study the transition kernel on $\text{P}(\mathbb{C}^k)$ given, for $S \in \mathcal{B}$, by

$$\Pi_s\left(\hat{x}, S\right) \propto \int_{\text{M}_k(\mathbb{C})} \mathbf{1}_S\left(v \cdot \hat{x}\right) \|vx\|^s \text{d}\mu(v).$$

In the case $s = 2$, Theorem A of [10] implies the conclusions of Theorem 1.1 under two assumptions:

- strong irreducibility, in the sense that there is no non-trivial finite union of proper subspaces of $\mathbb{C}^k$ left invariant by all $v \in \text{supp}\,\mu$,
- contractivity, in the sense that there exists a sequence (a_n) in T_μ, the smallest closed sub-semigroup of $\text{GL}_k(\mathbb{C})$ containing $\text{supp}\,\mu$, such that $\lim_{n\to\infty} a_n/\|a_n\|$ exists and is of rank one.

It is, however, immediate that strong irreducibility of μ implies **(ϕ-Erg)** with $E = \mathbb{C}^k$. In addition, if we assume $\text{supp}\,\mu \subset \text{GL}_k(\mathbb{C})$ and $\text{supp}\,\mu$ is strongly irreducible, the equivalence

$$\textbf{(Pur)} \iff T_\mu \text{ is contracting}$$

holds (see "Appendix A"). Our results therefore offer a strong refinement of [10] in the restricted framework of $s = 2$ with $\int v^* v \,\text{d}\mu(v) = \text{Id}_{\mathbb{C}^k}$. This assumption, although mathematically restrictive, is automatically verified in the framework of repeated (indirect) quantum measurements as described earlier in this section.

The article is structured as follows. Section 2 is devoted to the first part of Theorem 1.1, that is the uniqueness of the invariant measure. In Sect. 3 we show the geometric convergence towards the invariant measure with respect to the 1-Wasserstein

metric. In Sect. 4 we discuss the Lyapunov exponents of the process and relate them to the convergence between the Markov chain and an estimate of the chain used in our proofs.

Notation For $x \in \mathbb{C}^k \setminus \{0\}$, $\hat{x}$ is its equivalence class in $\mathrm{P}(\mathbb{C}^k)$ and, for $\hat{x}$ in $\mathrm{P}(\mathbb{C}^k)$, x is an arbitrary norm one vector representative of $\hat{x}$. If e.g. $\mathbb{P}_\nu$ (resp. $\mathbb{P}^\rho$) is a probability measure (depending on some a priori object ν (resp. ρ)) then $\mathbb{E}_\nu$ (resp. $\mathbb{E}^\rho$) is the expectation with respect to $\mathbb{P}_\nu$ (resp. $\mathbb{P}^\rho$). $\mathbb{N}$ represents the set of positive integers $\{1, 2, \ldots\}$.

2 Uniqueness of the invariant measure

This section concerns essentially the first part of Theorem 1.1. More precisely, under **(ϕ-Erg)** and **(Pur)** we show that the Markov chain has a unique invariant measure. Note that an invariant measure always exists since $\mathrm{P}(\mathbb{C}^k)$ is compact. We start by introducing a probability space describing both the state $\hat{x} \in \mathrm{P}(\mathbb{C}^k)$ and the sequence of matrices $(v_1, v_2, \ldots)$ such that $(v_n \ldots v_1 \cdot \hat{x})$ has the same distribution as the Markov chain $(\hat{x}_n)$. Then, in Proposition 2.1, we show that the marginal on the matrix sequence is the same for any Π-invariant probability measure as long as **(ϕ-Erg)** holds. In Proposition 2.2 and Lemma 2.3 we show that $(\hat{x}_n)$ is asymptotically a function of $(v_1, v_2, \ldots)$. We conclude on the uniqueness of the invariant measure in Corollary 2.4.

We now proceed to introduce some additional notation. We consider the space of infinite sequences $\Omega := \mathrm{M}_k(\mathbb{C})^{\mathbb{N}}$, write $\omega = (v_1, v_2, \ldots)$ for any such infinite sequence, and denote by π_n the canonical projection on the first n components, $\pi_n(\omega) = (v_1, \ldots, v_n)$. Let $\mathcal{M}$ be the Borel σ-algebra on $\mathrm{M}_k(\mathbb{C})$. For $n \in \mathbb{N}$, let $\mathcal{O}_n$ be the σ-algebra on Ω generated by the n-cylinder sets, i.e. $\mathcal{O}_n = \pi_n^{-1}(\mathcal{M}^{\otimes n})$. We equip the space Ω with the smallest σ-algebra $\mathcal{O}$ containing $\mathcal{O}_n$ for all $n \in \mathbb{N}$. We let $\mathcal{B}$ be the Borel σ-algebra on $\mathrm{P}(\mathbb{C}^k)$, and denote

$$\mathcal{J}_n = \mathcal{B} \otimes \mathcal{O}_n, \quad \mathcal{J} = \mathcal{B} \otimes \mathcal{O}.$$

This makes $\left(\mathrm{P}(\mathbb{C}^k) \times \Omega, \mathcal{J}\right)$ a measurable space. With a small abuse of notation we denote the sub-σ-algebra $\{\emptyset, \mathrm{P}(\mathbb{C}^k)\} \times \mathcal{O}$ by $\mathcal{O}$, and equivalently identify any $\mathcal{O}$-measurable function f with the $\mathcal{J}$-measurable function f satisfying $f(\hat{x}, \omega) = f(\omega)$.

For $i \in \mathbb{N}$, we consider the random variables $V_i : \Omega \to \mathrm{M}_k(\mathbb{C})$,

$$V_i(\omega) = v_i \quad \text{for} \quad \omega = (v_1, v_2, \ldots), \tag{8}$$

and we introduce $\mathcal{O}_n$-mesurable random variables (W_n) defined for all $n \in \mathbb{N}$ as

$$W_n = V_n \ldots V_1.$$

With a small abuse of notation we identify cylinder sets and their bases, and extend this identification to several associated objects. In particular we identify $O_n \in \mathcal{M}^{\otimes n}$ with $\pi_n^{-1}(O_n)$, a function f on $\mathcal{M}^{\otimes n}$ with $f \circ \pi_n$ and a measure $\mu^{\otimes n}$ with the measure

$\mu^{\otimes n} \circ \pi_n$. Since μ is not necessarily finite, we can not extend $(\mu^{\otimes n})$ into a measure on Ω.

Let ν be a probability measure over $(P(\mathbb{C}^k), \mathcal{B})$. We extend it to a probability measure $\mathbb{P}_\nu$ over $(P(\mathbb{C}^k) \times \Omega, \mathcal{J})$ by letting, for any $S \in \mathcal{B}$ and any cylinder set $O_n \in \mathcal{O}_n$,

$$\mathbb{P}_\nu(S \times O_n) := \int_{S \times O_n} \|W_n(\omega)x\|^2 \mathrm{d}\nu(\hat{x}) \mathrm{d}\mu^{\otimes n}(\omega). \tag{9}$$

From relation (1), it is easy to check that the expression (9) defines a consistent family of probability measures and, by Kolmogorov's theorem, this defines a unique probability measure $\mathbb{P}_\nu$ on $P(\mathbb{C}^k) \times \Omega$. In addition, the restriction of $\mathbb{P}_\nu$ to $\mathcal{B} \otimes \{\emptyset, \Omega\}$ is by construction ν.

We now define the random process $(\hat{x}_n)$. For $(\hat{x}, \omega) \in P(\mathbb{C}^k) \times \Omega$ we define $\hat{x}_0(\hat{x}, \omega) = \hat{x}$. Note that for any n, the definition (9) of $\mathbb{P}_\nu$ imposes

$$\mathbb{P}_\nu(W_n x_0 = 0) = 0.$$

This allows us to define a sequence $(\hat{x}_n)$ of $(\mathcal{J}_n)$-adapted random variables on the probability space $(P(\mathbb{C}^k) \times \Omega, \mathcal{J}, \mathbb{P}_\nu)$ by letting

$$\hat{x}_n := W_n \cdot \hat{x} \tag{10}$$

whenever the expression makes sense, i.e. for any ω such that $W_n(\omega)x \neq 0$, and extending it arbitrarily to the whole of Ω. The process $(\hat{x}_n)$ on $(\Omega \times P(\mathbb{C}^k), \mathcal{J}, \mathbb{P}_\nu)$ has the same distribution as the Markov chain defined by Π and initial probability measure ν.

Let us highlight the relation between $\mathbb{P}_\nu$ and density matrices. To that end, let

$$\rho_\nu := \mathbb{E}_\nu(\pi_{\hat{x}}). \tag{11}$$

By linearity and positivity of the expectation, $\rho_\nu \in \mathcal{D}_k$. Note that, conversely, for a given $\rho \in \mathcal{D}_k$ there exists ν (in general non-unique) such that $\rho_\nu = \rho$. For example, if a spectral decomposition of ρ is $\rho = \sum_j p_j \pi_{x_j}$ then necessarily $\sum_j p_j = 1$, so that $\nu = \sum_j p_j \delta_{\hat{x}_j}$ is a probability measure on $P(\mathbb{C}^k)$, and it satisfies the desired relation (11).

This relation motivates the following definition of probability measures over $(\Omega, \mathcal{O})$. For $\rho \in \mathcal{D}_k$ and any cylinder set $O_n \in \mathcal{O}_n$, let

$$\mathbb{P}^\rho(O_n) := \int_{O_n} \mathrm{tr}(W_n(\omega)\rho W_n^*(\omega)) \mathrm{d}\mu^{\otimes n}(\omega). \tag{12}$$

In particular, for any $S \in \mathcal{B}$ and $A \in \mathcal{O}$,

$$\mathbb{P}_\nu(S \times A) = \int_S \mathbb{P}^{\pi_{\hat{x}}}(A) \, \mathrm{d}\nu(\hat{x}). \tag{13}$$

The following proposition elucidates further the connection between $\mathbb{P}_\nu$ and $\mathbb{P}^{\rho_\nu}$.

Proposition 2.1 *The marginal of $\mathbb{P}_\nu$ on $\mathcal{O}$ is the probability measure $\mathbb{P}^{\rho_\nu}$. Moreover, if* **(ϕ-Erg)** *holds, $\mathbb{P}^{\rho_{\nu_a}} = \mathbb{P}^{\rho_{\nu_b}}$ for any two Π-invariant probability measures ν_a and ν_b.*

Proof By construction it is sufficient to check the equality of the measures on cylinder sets. Let $O_n \in \mathcal{O}_n$; from the definition of $\mathbb{P}_\nu$, and the linearity of the trace and the integral, we have

$$\mathbb{P}_\nu \left(\mathrm{P}(\mathbb{C}^k) \times O_n \right) = \int_{\mathrm{P}(\mathbb{C}^k) \times O_n} \mathrm{tr}\big(W_n^*(\omega) W_n(\omega) \pi_{\hat{x}}\big) \, \mathrm{d}\nu(\hat{x}) \mathrm{d}\mu^{\otimes n}(\omega)$$

$$= \int_{O_n} \mathrm{tr}\left(W_n^*(\omega) W_n(\omega) \int_{\mathrm{P}(\mathbb{C}^k)} \pi_{\hat{x}} \, \mathrm{d}\nu(\hat{x}) \right) \mathrm{d}\mu^{\otimes n}(\omega)$$

$$= \int_{O_n} \mathrm{tr}\big(W_n^*(\omega) W_n(\omega) \rho_\nu\big) \, \mathrm{d}\mu^{\otimes n}(\omega).$$

The equality between the marginal of $\mathbb{P}_\nu$ on $\mathcal{O}$ and $\mathbb{P}^{\rho_\nu}$ follows.

If ν is an invariant measure, on the one hand

$$\mathbb{E}_\nu \left(\pi_{\hat{x}_1} \right) = \mathbb{E}_\nu \left(\pi_{\hat{x}_0} \right) = \rho_\nu.$$

On the other hand,

$$\mathbb{E}_\nu \left(\pi_{\hat{x}_1} \right) = \int_{\mathrm{P}(\mathbb{C}^k) \times \mathrm{M}_k(\mathbb{C})} \frac{v \pi_{\hat{x}_0} v^*}{\|vx_0\|^2} \|vx_0\|^2 \mathrm{d}\nu \left(\hat{x}_0 \right) \mathrm{d}\mu(v)$$

$$= \int_{\mathrm{M}_k(\mathbb{C})} v \, \mathbb{E}_\nu \left(\pi_{\hat{x}_0} \right) v^* \mathrm{d}\mu(v)$$

$$= \phi(\rho_\nu),$$

so that ρ_ν is a fixed point of ϕ. Hence if **(ϕ-Erg)** holds, ρ_ν is the unique fixed point of ϕ in $\mathcal{D}_k$. Hence $\rho_{\nu_a} = \rho_{\nu_b}$ and $\mathbb{P}^{\rho_{\nu_a}} = \mathbb{P}^{\rho_{\nu_b}}$ holds. $\qquad\square$

In the following we use the measure $\mathbb{P}^{\mathrm{ch}} = \mathbb{P}^{\frac{1}{k}\mathrm{Id}_{\mathbb{C}^k}}$ associated to the operator $\mathrm{Id}_{\mathbb{C}^k}/k \in \mathcal{D}_k$ as a reference measure. Since for any $\rho \in \mathcal{D}_k$ there exists a constant c such that $\rho \le c \frac{\mathrm{Id}_{\mathbb{C}^k}}{k}$, the measure $\mathbb{P}^\rho$ is absolutely continuous w.r.t. $\mathbb{P}^{\mathrm{ch}}$. We will denote absolute continuity between measures with the symbol $\ll$, so that we have here

$$\mathbb{P}^\rho \ll \mathbb{P}^{\mathrm{ch}},$$

for all $\rho \in \mathcal{D}_k$. The Radon–Nykodim derivative will be made explicit in Proposition 2.2. To that end, we use a particular $(\mathcal{O}_n)$-adapted process. We define a sequence of matrix-valued random variables:

$$M_n := \frac{W_n^* W_n}{\mathrm{tr}\left(W_n^* W_n\right)} \quad \text{if } \mathrm{tr}\left(W_n^* W_n\right) \ne 0$$

and extend the definition arbitrarily whenever $\mathrm{tr}(W_n^* W_n) = 0$. The latter alternative appears with probability 0: indeed, $\mathbb{P}^{\mathrm{ch}}\big(\mathrm{tr}(W_n^* W_n) = 0\big) = 0$ and then by the absolute continuity of $\mathbb{P}^\rho$ with respect to $\mathbb{P}^{\mathrm{ch}}$ we have $\mathbb{P}_\nu\big(\mathrm{tr}(W_n^* W_n) = 0\big) = \mathbb{P}^{\rho_\nu}\big(\mathrm{tr}(W_n^* W_n) = 0\big) = 0$ for any measure ν. The key property of M_n, that we establish in the proof of Proposition 2.2, is that it is an $(\mathcal{O}_n)$-martingale with respect to $\mathbb{P}^{\mathrm{ch}}$.

From the existence of a polar decomposition for W_n, for each n, there exists a unitary matrix-valued random variable U_n such that

$$W_n = U_n \sqrt{\mathrm{tr}\left(W_n^* W_n\right)} M_n^{\frac{1}{2}}. \tag{14}$$

This process (U_n) can be chosen to be $(\mathcal{O}_n)$-adapted.

The key technical results about M_n needed for our proofs are summarized in the following proposition. Recall that any $\mathcal{O}$-measurable function f is extended to a $\mathcal{J}$-measurable function by setting $f(\hat{x}, \omega) = f(\omega)$ for any $(\hat{x}, \omega) \in \mathrm{P}(\mathbb{C}^k) \times \Omega$.

Proposition 2.2 *For any probability measure ν over $(\mathrm{P}(\mathbb{C}^k), \mathcal{B})$, (M_n) converges $\mathbb{P}_\nu$-a.s. and in L^1-norm to an $\mathcal{O}$-measurable random variable M_∞. The change of measure formula*

$$\frac{\mathrm{d}\mathbb{P}^\rho}{\mathrm{d}\mathbb{P}^{\mathrm{ch}}} = k \, \mathrm{tr}(\rho M_\infty) \tag{15}$$

holds true for all $\rho \in \mathcal{D}_k$.

Moreover, the measure μ verifies **(Pur)** *if and only if M_∞ is $\mathbb{P}_\nu$-a.s. a rank one projection for any probability measure ν over $(\mathrm{P}(\mathbb{C}^k), \mathcal{B})$.*

Proof We start the proof by showing that M_n is a $\mathbb{P}^{\mathrm{ch}}$-martingale. Recall that for all $n \in \mathbb{N}$ and all $O_n \in \mathcal{O}_n$,

$$\mathbb{P}^{\mathrm{ch}}(O_n) = \frac{1}{k} \int_{O_n} \mathrm{tr}\big(W_n^*(\omega) W_n(\omega)\big) \, \mathrm{d}\mu^{\otimes n}(\omega).$$

From the definition of W_n, Eq. (8),

$$M_{n+1} = \frac{W_n^* V_{n+1}^* V_{n+1} W_n}{\mathrm{tr}\left(W_n^* W_n\right)} \frac{\mathrm{tr}\left(W_n^* W_n\right)}{\mathrm{tr}\left(W_{n+1}^* W_{n+1}\right)}. \tag{16}$$

This implies that for an arbitrary $\mathcal{O}_n$-measurable random variable Y

$$\begin{aligned}
\mathbb{E}^{\mathrm{ch}}(Y M_{n+1}) &= \frac{1}{k} \int_{\mathrm{M}_k(\mathbb{C})^{n+1}} \frac{W_n^* V_{n+1}^* V_{n+1} W_n}{\mathrm{tr}\left(W_n^* W_n\right)} \, Y \, \mathrm{tr}\big(W_n^* W_n\big) \, \mathrm{d}\mu^{\otimes n+1} \\
&= \frac{1}{k} \int_{\mathrm{M}_k(\mathbb{C})^n} \frac{W_n^* W_n}{\mathrm{tr}\left(W_n^* W_n\right)} \, Y \, \mathrm{tr}\big(W_n^* W_n\big) \, \mathrm{d}\mu^{\otimes n} \\
&= \mathbb{E}^{\mathrm{ch}}(Y M_n),
\end{aligned}$$

where the second equality follows from the stochasticity condition (1), $\int v^* v d\mu(v) =$ $\mathrm{Id}_{\mathbb{C}_k}$. This shows that (M_n) is an $(\mathcal{O}_n)$-martingale w.r.t. $\mathbb{P}^{\mathrm{ch}}$. Since the sequence (M_n) is composed of positive semidefinite matrices of trace one, its coordinates are a.s. uniformly bounded by 1. Therefore, the martingale property implies the L^1 and a.s. convergence of (M_n) to an $\mathcal{O}$-measurable random variable M_∞. Now note that for any $\rho \in \mathcal{D}_k$,

$$\mathrm{tr}\left(W_n^* W_n \rho\right) = \mathrm{tr}(M_n \rho)\, k\, \mathrm{tr}\left(W_n^* W_n \frac{\mathrm{Id}_{\mathbb{C}_k}}{k}\right).$$

This way, the convergence of (M_n) implies the change of measure formula.

We now prove the last part of the proposition. Using the martingale property one can see that for all $n \in \mathbb{N}$, and any fixed $p \in \mathbb{N}$,

$$
\begin{aligned}
V_n^p &:= \sum_{k=0}^{p-1} \mathbb{E}^{\mathrm{ch}}\left(M_{k+n+1}^2 - M_k^2\right) = \sum_{k=0}^{n} \mathbb{E}^{\mathrm{ch}}\left(M_{k+p}^2\right) - \sum_{k=0}^{n} \mathbb{E}^{\mathrm{ch}}\left(M_k^2\right) \\
&= \sum_{k=0}^{n} \mathbb{E}^{\mathrm{ch}}\left(\left(M_{k+p} - M_k\right)^2\right) \\
&= \mathbb{E}^{\mathrm{ch}}\left(\sum_{k=0}^{n} \mathbb{E}^{\mathrm{ch}}\left(\left(M_{k+p} - M_k\right)^2 | \mathcal{O}_k\right)\right).
\end{aligned}
\tag{17}
$$

Since (M_n) is bounded and almost surely convergent, applying Lebesgue's dominated convergence theorem to each $\mathbb{E}^{\mathrm{ch}}(M_{k+n+1}^2)$, $k = 0, \ldots, p-1$ as $n \to \infty$ implies that the term V_n^p is convergent when n goes to infinity. Then, using the monotone convergence theorem in the last line of (17), we get that

$$\lim_{n \to \infty} V_n^p = \mathbb{E}^{\mathrm{ch}}\left(\sum_{k=0}^{\infty} \mathbb{E}^{\mathrm{ch}}\left(\left(M_{k+p} - M_k\right)^2 | \mathcal{O}_k\right)\right).$$

It implies that the series $\sum_{k=0}^{\infty} \mathbb{E}^{\mathrm{ch}}\left((M_{k+p} - M_k)^2 | \mathcal{O}_k\right)$ is almost surely finite. This yields that

$$\lim_{n \to \infty} \mathbb{E}^{\mathrm{ch}}\left(\left(M_{n+p} - M_n\right)^2 | \mathcal{O}_n\right) = 0, \quad \mathbb{P}^{\mathrm{ch}}\text{-a.s.}$$

Since all the norms are equivalent in finite dimension, Jensen's inequality implies

$$\lim_{n \to \infty} \mathbb{E}^{\mathrm{ch}}\left(\left\| M_{n+p} - M_n \right\| | \mathcal{O}_n\right) = 0, \quad \mathbb{P}^{\mathrm{ch}}\text{-a.s.} \tag{18}$$

At this stage we use the polar decomposition of (W_n), Eq. (14), to write

$$M_{n+p} = \frac{M_n^{\frac{1}{2}} U_n^* V_{n+1}^* \cdots V_{n+p}^* V_{n+p} \cdots V_{n+1} U_n M_n^{\frac{1}{2}}}{\mathrm{tr}\left(M_n^{\frac{1}{2}} U_n^* V_{n+1}^* \cdots V_{n+p}^* V_{n+p} \cdots V_{n+1} U_n M_n^{\frac{1}{2}}\right)}.$$

Then we get an expression for the conditional expectation, see the first part of the proof,

$$
\mathbb{E}^{\mathrm{ch}}\big(\,\|M_{n+p} - M_n\|\,|\mathcal{O}_n\big) = \int \left\| \frac{M_n^{\frac{1}{2}} U_n^* v_1^* \ldots v_p^* v_p \ldots v_1 U_n M_n^{\frac{1}{2}}}{\mathrm{tr}\left(M_n^{\frac{1}{2}} U_n^* v_1^* \ldots v_p^* v_p \ldots v_1 U_n M_n^{\frac{1}{2}}\right)} - M_n \right\|
$$
$$
\mathrm{tr}\left(v_1^* \ldots v_p^* v_p \ldots v_1 U_n M_n U_n^*\right)\, \mathrm{d}\mu^{\otimes p}(v_1, \ldots, v_p)
$$
$$
= \int \left\| M_n^{\frac{1}{2}} U_n^* v_1^* \ldots v_p^* v_p \ldots v_1 U_n M_n^{\frac{1}{2}} \right.
$$
$$
\left. - M_n \mathrm{tr}\left(v_1^* \ldots v_p^* v_p \ldots v_1 U_n M_n U_n^*\right) \right\|
$$
$$
\mathrm{d}\mu^{\otimes p}(v_1, \ldots, v_p).
$$

We used non-negativity of $\mathrm{tr}(v_1^* \ldots v_p^* v_p \ldots v_1 U_n M_n U_n^*)$ to get the second equality. The above equation holds for $\mathbb{P}^{\mathrm{ch}}$-almost all realizations $\big(U_n(\omega)\big)$ of (U_n). Since the group of unitary matrices is compact, for any fixed ω there exists a subsequence along which $\big(U_n(\omega)\big)$ converges to a unitary matrix $U_\infty(\omega)$. Taking the limit along this subsequence in the above expression yields (we drop ω for notational simplicity):

$$
\int_{\mathrm{M}_k(\mathbb{C})^p} \left\| M_\infty^{\frac{1}{2}} U_\infty^* v_1^* \ldots v_p^* v_p \ldots v_1 U_\infty M_\infty^{\frac{1}{2}} - M_\infty \mathrm{tr}\left(v_1^* \ldots v_p^* v_p \ldots v_1 U_\infty M_\infty U_\infty^*\right) \right\|
$$
$$
\mathrm{d}\mu^{\otimes p}(v_1, \ldots, v_p) = 0.
$$

This implies that

$$
M_\infty^{\frac{1}{2}} U_\infty^* v_1^* \ldots v_p^* v_p \ldots v_1 U_\infty M_\infty^{\frac{1}{2}} = \mathrm{tr}\left(v_1^* \ldots v_p^* v_p \ldots v_1 U_\infty M_\infty U_\infty^*\right) M_\infty,
$$

for $\mu^{\otimes p}$-almost all $(v_1, \ldots, v_p)$.

Denoting by π_∞ the orthogonal projector onto the range of M_∞, the above condition is equivalent to $\pi_\infty U_\infty^* v_1^* \ldots v_p^* v_p \ldots v_1 U_\infty \pi_\infty = \lambda \pi_\infty$ with $\lambda = \mathrm{tr}(v_1^* \ldots v_p^* v_p \ldots v_1 U_\infty M_\infty U_\infty^*)$. Finally, it follows that

$$
U_\infty \pi_\infty U_\infty^* v_1^* \ldots v_p^* v_p \ldots v_1 U_\infty \pi_\infty U_\infty^* \propto U_\infty \pi_\infty U_\infty^*,
$$

for $\mu^{\otimes p}$-almost all $(v_1, \ldots, v_p)$. Since $U_\infty \pi_\infty U_\infty^*$ is an orthogonal projector, the condition (**Pur**) implies (reintroducing ω) that $\mathrm{rank}(M_\infty(\omega)) = \mathrm{rank}(U_\infty(\omega)\pi_\infty(\omega)U_\infty^*(\omega)) = 1$. Since $M_\infty(\omega)$ is a trace one, positive semidefinite matrix this means that $M_\infty(\omega)$ is a rank one projector. Since this conclusion holds true for $\mathbb{P}^{\mathrm{ch}}$-almost all ω this establishes that the condition (**Pur**) implies that M_∞ is $\mathbb{P}^{\mathrm{ch}}$-almost surely a rank 1 projection.

For the converse implication, assume that M_∞ is $\mathbb{P}^{\mathrm{ch}}$-almost surely a rank one projection but that (**Pur**) does not hold. Then there exists π, a rank two orthogonal projector, such that for all $n \in \mathbb{N}$,

$$\pi W_n^* W_n \pi \propto \pi,$$

$\mu^{\otimes n}$-almost everywhere. Since $\mu^{\otimes n}$-almost everywhere $M_n \propto W_n^* W_n$, we get that

$$\pi M_n \pi \propto \pi,$$

$\mu^{\otimes n}$-almost everywhere. Thus, $\pi M_\infty \pi \propto \pi$ and, under our assumption that rank $M_\infty = 1$ $\mathbb{P}^{\mathrm{ch}}$-a.s. and rank $\pi = 2$, this implies that $\pi M_\infty \pi = 0$, $\mathbb{P}^{\mathrm{ch}}$-almost surely. On the other hand for all $n \in \mathbb{N}$ we have $\mathbb{E}^{\mathrm{ch}}(M_n) = \mathrm{Id}_{\mathbb{C}^k}$, and the L^1 convergence implies that $\mathbb{E}^{\mathrm{ch}}(M_\infty) = \mathrm{Id}_{\mathbb{C}^k}$. Then, $\mathbb{E}^{\mathrm{ch}}(\pi M_\infty \pi) = \pi$ which contradicts $\pi M_\infty \pi = 0$ $\mathbb{P}^{\mathrm{ch}}$-a.s. □

By the polar decomposition, the rank of W_n is equal to the rank of M_n and the proposition thus implies that $W_n \rho_0 W_n^* / \mathrm{tr}(W_n \rho_0 W_n^*)$ approaches the set of pure states for any $\rho_0 \in \mathcal{D}_k$ if and only if **(Pur)** holds. This is the result of Maassen and Kümmerer [17] mentioned in the introduction. Though M_n is not used in [17], the proof relies on similar ideas.

We are now in the position to show that the Markov chain $(\hat{x}_n)$ is asymptotically an $\mathcal{O}$-measurable process. This is expressed in the following lemma. Whenever **(Pur)** holds, we denote by $\hat{z} \in \mathrm{P}(\mathbb{C}^k)$ the $\mathcal{O}$-measurable random variable defined by

$$M_\infty = \pi_{\hat{z}}.$$

Recall that $d(\cdot, \cdot)$, defined by Eq. (7), is our metric on $\mathrm{P}(\mathbb{C}^k)$.

Lemma 2.3 *Assume **(Pur)** holds. Then for any probability measure ν on $(\mathrm{P}(\mathbb{C}^k), \mathcal{B})$,*

$$\lim_{n \to \infty} d\left(\hat{x}_n, U_n \cdot \hat{z}\right) = 0 \quad \mathbb{P}_\nu\text{-a.s.}$$

Proof We start the proof by showing that for any ν

$$\lim_{n \to \infty} M_n^{\frac{1}{2}} \cdot \hat{x} = \hat{z} \quad \mathbb{P}_\nu\text{-a.s.} \tag{19}$$

Let $\hat{x}$ be fixed and recall from Proposition 2.2 that **(Pur)** implies $M_\infty = \pi_{\hat{z}}$. Since $M_\infty x = \langle z, x \rangle z$, in order to show (19), it is enough to show that $\hat{x}$ is $\mathbb{P}_\nu$-almost surely not orthogonal to $\hat{z}$. From Eq. (13) and the change of measure formula in Proposition 2.2,

$$\mathrm{d}\mathbb{P}_\nu\left(\hat{x}, \omega\right) = k \, \mathrm{tr}\left(\pi_{\hat{x}} \pi_{\hat{z}(\omega)}\right) \mathrm{d}\left(\nu(\hat{x}) \otimes \mathbb{P}^{\mathrm{ch}}(\omega)\right).$$

Hence the event $\{\mathrm{tr}(\pi_{\hat{x}} \pi_{\hat{z}}) = |\langle z, x \rangle|^2 = 0\}$ has $\mathbb{P}_\nu$-measure 0. This proves the required claim, and (19) follows from the almost sure convergence of M_n to $\pi_{\hat{z}}$.

Now using the polar decomposition, Eq. (14), and the fact that proportionality of vectors amounts to equality of their equivalence classes in $\mathrm{P}(\mathbb{C}^k)$, we have

$$\hat{x}_n = U_n M_n^{\frac{1}{2}} \cdot \hat{x}_0.$$

The first part of the proof then yields

$$\lim_{n \to \infty} d\left(\hat{x}_n, U_n \cdot \hat{z}\right) = 0, \quad \mathbb{P}_\nu \text{-a.s.}$$

□

The uniqueness of the invariant measure which is the first part of Theorem 1.1 follows as a corollary.

Corollary 2.4 *Assume* (**Pur**) *and* (ϕ-**Erg**). *Then the Markov kernel* Π *admits a unique invariant probability measure.*

Proof For an invariant measure ν, the random variable $\hat{x}_n$ is ν-distributed for all $n \in \mathbb{N}$. In particular, $\mathbb{E}_\nu\left(f(\hat{x}_n)\right)$ is constant for any continuous function f. On the other hand Lemma 2.3 and Lebesgue's dominated convergence theorem imply that

$$\lim_{n \to \infty} \mathbb{E}_\nu\left(f\left(\hat{x}_n\right) - f\left(U_n \cdot \hat{z}\right)\right) = 0.$$

Hence

$$\lim_{n \to \infty} \mathbb{E}_\nu\left(f\left(U_n \cdot \hat{z}\right)\right) = \mathbb{E}_\nu\left(f\left(\hat{x}_0\right)\right). \tag{20}$$

Assume now that there exist two invariant measures ν_a and ν_b. Since $U_n \cdot \hat{z}$ is $\mathcal{O}$-measurable, Proposition 2.1 implies

$$\mathbb{E}_{\nu_a}\left(f\left(U_n \cdot \hat{z}\right)\right) = \mathbb{E}_{\nu_b}\left(f\left(U_n \cdot \hat{z}\right)\right).$$

Then Eq. (20) applied with $\nu = \nu_a$, resp. $\nu = \nu_b$ gives

$$\mathbb{E}_{\nu_a}\left(f\left(\hat{x}_0\right)\right) = \mathbb{E}_{\nu_b}\left(f\left(\hat{x}_0\right)\right)$$

which means that $\nu_a = \nu_b$ and the uniqueness is proved. □

Assuming only (**Pur**) we can actually completely characterize the set of invariant measures.

Proposition 2.5 *Assuming* (**Pur**) *there exists a set* $\{F_j\}_{j=1}^d$ *of mutually orthogonal subspaces of* $\mathbb{C}^k$ *such that for each* $j \in \{1, \ldots, d\}$ *there exists a unique* Π-*invariant probability measure* ν_j *supported on* $\mathrm{P}(F_j)$, *and the set of* Π-*invariant probability measures is the convex hull of* $\{\nu_j\}_{j=1}^d$.

The subspaces F_j are the ranges of the extremal fixed points of ϕ in $\mathcal{D}_k$. This is shown in the proof of Proposition 2.5, which we give in "Appendix B".

Remark 2.6 Assuming (ϕ-**Erg**) only, the chain might or might not have a unique invariant probability measure. Indeed, if $\mathrm{supp}\,\mu \subset \mathrm{SU}(k)$, Assumption (**Pur**) is trivially not verified and, as proved in "Appendix C", the uniqueness of the invariant measure depends on the smallest closed subgroup of $\mathrm{SU}(k)$ containing $\mathrm{supp}\,\mu$. To illustrate this point, in the same appendix, we study two examples with μ supported on and giving equiprobability to two elements of $\mathrm{SU}(2)$ such that (ϕ-**Erg**) holds. In the first example Π has a unique invariant probability measure whereas in the second example Π has uncountably many mutually singular invariant probability measures.

3 Convergence

We now turn to the proof of the second part of Theorem 1.1, namely the geometric convergence in Wasserstein distance of the process $(\hat{x}_n)$ towards the invariant measure. We first recall a definition of this distance for compact metric spaces: for X a compact metric space equipped with its Borel σ-algebra, the Wasserstein distance of order 1 between two probability measures σ and τ on X can be defined using the Kantorovich–Rubinstein duality theorem as

$$
W_1(\sigma, \tau) = \sup_{f \in \mathrm{Lip}_1(X)} \left| \int_X f \, \mathrm{d}\sigma - \int_X f \, \mathrm{d}\tau \right|,
$$

where $\mathrm{Lip}_1(X) = \{f : X \to \mathbb{R} \text{ s.t. } |f(x) - f(y)| \le d(x, y)\}$ is the set of Lipschitz continuous functions with constant one, and d is the metric on X.

The proof of Eq. (6) consists of three parts. In the first part we show a geometric convergence in total variation of $\mathbb{P}^\rho$ to $\mathbb{P}^{\rho_{\mathrm{inv}}}$ under the shift $\theta(v_1, v_2, \ldots) = (v_2, v_3, \ldots)$. In the second one we show a geometric convergence of the chain $(\hat{x}_n)$ towards an $\mathcal{O}$-measurable process $(\hat{y}_n)$. Finally, we combine these results to prove Eq. (6).

3.1 Convergence for $\mathcal{O}$-measurable random variables

Let us first discuss the origin of the integer m in Eq. (6). Let E be a subspace of $\mathbb{C}^k$ s.t. $vE \subset E$ for any $v \in \mathrm{supp}\,\mu$. Let $(E_1, \ldots, E_\ell)$ be an orthogonal partition of E, i.e. $E = E_1 \oplus \cdots \oplus E_\ell$. We say that $(E_1, \ldots, E_\ell)$ is a ℓ-cycle of ϕ if $vE_j \subset E_{j+1}$ for μ-a.e. v (with the convention $E_{\ell+1} = E_1$).[2] The set of $\ell \in \mathbb{N}$ for which there exists an ℓ-cycle is non-empty (as it contains 1) and bounded (as necessarily $\ell \le k$).

Definition 3.1 The largest $\ell \in \mathbb{N}$ such that there exists a ℓ-cycle of ϕ is called the period of ϕ. We denote this period by m.

Remark 3.2 • The above definition for the period of ϕ is similar to that of the period of a φ-irreducible Markov chain. It is obvious that if $(E_1, \ldots, E_\ell)$ is an ℓ-cycle of ϕ then it is also an ℓ-cycle of Π. However, the Markov chain defined by Π is not φ-irreducible in general. Hence the results of [19] on the period of φ-irreducible Markov chains do not apply and the characterization of the period of Π remains an open problem.
• The above definition shows that the union $\bigcup_{j=1}^m E_j$ is invariant by μ-a.e. v. Hence, the strong irreducibility assumption discussed at the end of the introduction implies that $m = 1$.

The following result is a reformulation of the Perron–Frobenius theorem of Evans and Høegh-Krohn (see [7]). The original formulation in [7] makes the additional assumption that $E = \mathbb{C}^k$ in (ϕ-**Erg**). For the present extension see e.g. [22]. In the following statement, and in the rest of the article, for X an operator on $\mathbb{C}^k$ we denote

---

[2] As suggested by its name, the notion of cycle for ϕ depends only on ϕ and not on the specific measure μ leading to ϕ [20,22].

$\|X\|_1 = \mathrm{tr}|X|$ (all statements are identical with a different norm, but this choice will spare us a few irrelevant constants).

Theorem 3.3 *Assume that (ϕ-Erg) holds. Then there exists a unique ϕ-invariant element ρ_{inv} of $\mathcal{D}_k$ with range equal to the minimal invariant subspace E. In addition, there exist two positive constants c and $\lambda < 1$ such that, with m defined in Definition 3.1, for any $\rho \in \mathcal{D}_k$ and for any $n \in \mathbb{N}$,*

$$\left\| \frac{1}{m} \sum_{r=0}^{m-1} \phi^{mn+r}(\rho) - \rho_{\mathrm{inv}} \right\|_1 \leq c\lambda^n. \tag{21}$$

Proof Theorem 4.2 in [7] implies that ρ_{inv} is the unique ϕ-invariant element of $\mathcal{D}_k$, that the eigenvalues of ϕ of modulus one are exactly the m-th roots of unity, and that they are all simple. The statement follows, with λ any quantity strictly larger than the modulus of the largest non-peripheral eigenvalue.   □

Recall that θ is the left shift operator on Ω, i.e.

$$\theta(v_1, v_2, \ldots) = (v_2, v_3, \ldots).$$

The main result of this section is the following proposition. As announced it concerns the speed of convergence in total variation (expressed in terms of expectation values).

Proposition 3.4 *Assume (ϕ-Erg) holds. Then there exist two positive constants C and $\lambda < 1$ such that for any $\mathcal{O}$-measurable function f with essential bound $\|f\|_\infty$, any $\rho \in \mathcal{D}_k$ and all $n \in \mathbb{N}$,*

$$\left| \mathbb{E}^\rho \left(\frac{1}{m} \sum_{r=0}^{m-1} f \circ \theta^{mn+r} \right) - \mathbb{E}^{\rho_{\mathrm{inv}}}(f) \right| \leq C\|f\|_\infty \lambda^n. \tag{22}$$

Proof We claim that for any bounded $\mathcal{O}$-measurable function f,

$$\mathbb{E}^\rho(f \circ \theta) = \mathbb{E}^{\phi(\rho)}(f). \tag{23}$$

It suffices to prove this relation for all $\mathcal{O}_l$-measurable functions for any integer l. Thus, let l be an integer and f an $\mathcal{O}_l$-measurable function. Then,

$$\mathbb{E}^\rho(f \circ \theta) = \int_{\mathrm{M}_k(\mathbb{C})^{l+1}} f(v_2, \ldots, v_{l+1}) \mathrm{tr}\left(v_{l+1} \ldots v_1 \rho v_1^* \ldots v_{l+1}^* \right) d\mu^{\otimes(l+1)}(v_1, \ldots, v_{l+1})$$

$$= \int_{\mathrm{M}_k(\mathbb{C})^l} f(v_1, \ldots, v_l) \mathrm{tr}\left(v_l \ldots v_1 \phi(\rho) v_1^* \ldots v_l^* \right) d\mu^{\otimes l}(v_1, \ldots, v_l),$$

which is equal to $\mathbb{E}^{\phi(\rho)}(f)$, as claimed.

Applying Eq. (23) multiple times and using the change of measure of Proposition 2.2 we obtain

$$\mathbb{E}^{\rho}\left(\frac{1}{m}\sum_{r=0}^{m-1} f \circ \theta^{mn+r}\right) = \frac{1}{m}\sum_{r=0}^{m-1} \mathbb{E}^{\phi^{nm+r}(\rho)}(f)$$

$$= k\frac{1}{m}\sum_{r=0}^{m-1} \mathbb{E}^{\mathrm{ch}}\left(f \, \mathrm{tr}\left(M_{\infty}\phi^{nm+r}(\rho)\right)\right),$$

for any $\mathcal{O}$-measurable function f. Using $|\mathrm{tr}(M_{\infty}A)| \leq \|A\|_1$ for $A = A^*$ (remark that $M_{\infty} \in \mathcal{D}_k$ by construction) we then obtain

$$\left|\mathbb{E}^{\rho}\left(\frac{1}{m}\sum_{r=0}^{m-1} f \circ \theta^{mn+r}\right) - \mathbb{E}^{\rho_{\mathrm{inv}}}(f)\right| \leq \|f\|_{\infty} k \left\|\frac{1}{m}\sum_{r=0}^{m-1} \phi^{mn+r}(\rho) - \rho_{\mathrm{inv}}\right\|_1$$

and Theorem 3.3 yields the proposition with $C = ck$. $\qquad\square$

3.2 Convergence to an $\mathcal{O}$-measurable process

Let us introduce two relevant processes: for all $n \in \mathbb{N}$, let

$$\hat{z}_n(\omega) = \underset{\hat{x}\in\mathrm{P}(\mathbb{C}^k)}{\mathrm{argmax}} \; \|W_n x\|^2 \tag{24}$$

and

$$\hat{y}_n = W_n \cdot \hat{z}_n. \tag{25}$$

Both random variables $\hat{y}_n$ and $\hat{z}_n$ are $\mathcal{O}_n$-measurable.

The random variable $\hat{z}_n$ corresponds to the maximum likelihood estimator of $\hat{x}_0$. Note that the argmax may not be uniquely defined. We can, however, define it in an $\mathcal{O}_n$-measurable way. The following results will not be affected by such a consideration, and we will not discuss such questions in the sequel. It follows from the definition of $\hat{z}_n$ that

$$\left(W_n^* W_n\right)^{\frac{1}{2}} z_n = \|W_n\| z_n, \quad \mathbb{P}^{\mathrm{ch}} \text{- a.s.} \tag{26}$$

We recall that z_n is a vector representative of the class $\hat{z}_n$.

Concerning $\hat{y}_n$, it can be seen as an estimator of $\hat{x}_n$ given the maximum likelihood estimation of $\hat{x}_0$. The following proposition establishes consistency of this estimator, we show the geometric contraction in the mean of $(\hat{x}_n)$ and $(\hat{y}_n)$. In fact we prove a slightly more general statement that the estimator based on the first n outcomes can be replaced by an estimator based on outcomes between l and $l + n$. We will prove the almost-sure contraction in Proposition 4.4.

Proposition 3.5 *Assume* **(Pur)** *holds. Then there exist two positive constants C and $\lambda < 1$ such that for any probability measure ν over $(P(\mathbb{C}^k), \mathcal{B})$,*

$$\mathbb{E}_\nu \left(d\left(\hat{x}_{n+l}, \hat{y}_n \circ \theta^l \right) \right) \leq C\lambda^n, \tag{27}$$

holds for all non-negative integers l and n.

In order to prove Proposition 3.5 we study the largest two singular values of W_n. As is customary in the study of products of random matrices, we make use of exterior products. We recall briefly the relevant definitions: for $p \in \mathbb{N}$ and p vectors $x_1, \dots, x_p$ in $\mathbb{C}^k$ we denote by $x_1 \wedge \cdots \wedge x_p$ the alternating bilinear form $(y_1, \dots, y_p) \mapsto \det \left(\langle x_i, y_j \rangle \right)_{i,j=1}^p$. Then, the set of all $x_1 \wedge \cdots \wedge x_p$ is a generating family for the set $\wedge^p \mathbb{C}^k$ of alternating bilinear forms on $\mathbb{C}^k$, and we can define a hermitian inner product by

$$\langle x_1 \wedge \cdots \wedge x_p, y_1 \wedge \cdots \wedge y_p \rangle = \det \left(\langle x_i, y_j \rangle \right)_{i,j=1}^p,$$

and denote by $\| x_1 \wedge \cdots \wedge x_p \|$ the associated norm. It is immediate to verify that our metric d, defined by (7), satisfies

$$d(\hat{x}, \hat{y}) = \frac{\| x \wedge y \|}{\| x \| \| y \|}. \tag{28}$$

For an operator A on $\mathbb{C}^k$, we write $\wedge^p A$ for the operator on $\wedge^p \mathbb{C}^k$ defined by

$$\wedge^p A \, (x_1 \wedge \cdots \wedge x_p) = Ax_1 \wedge \cdots \wedge Ax_p. \tag{29}$$

Obviously $\wedge^p(AB) = \wedge^p A \wedge^p B$, so that $\| \wedge^p (AB) \| \leq \| \wedge^p A \| \| \wedge^p B \|$. From e.g. Chapter XVI of [18] or Lemma III.5.3 of [4], we have in addition for $1 \leq p \leq k$

$$\left\| \wedge^p A \right\| = a_1(A) \dots a_p(A), \tag{30}$$

where $a_1(A) \geq \cdots \geq a_k(A)$ are the singular values of A, i.e. the square roots of eigenvalues of A^*A, labelled in decreasing order.

Our strategy to prove Proposition 3.5 is to bound the left hand side of Eq. (27) by a submultiplicative function $f : \mathbb{N} \to \mathbb{R}_+$ and then use Fekete's lemma. We will show that the function

$$f(n) = \int_{M_k(\mathbb{C})^n} \left\| \wedge^2 v_n \dots v_1 \right\| \, d\mu^{\otimes n}(v_1, \dots, v_n) \tag{31}$$

has the desired properties. The following lemma establishes an exponential decay of this function.

Lemma 3.6 *Assume* **(Pur)**. *Then there exist two positive constants C and* $\lambda < 1$ *such that*

$$f(n) \leq C\lambda^n.$$

Proof First, we prove $\lim_{n\to\infty} f(n) = 0$. To prove it, we express the function $f(n)$ using the process W_n as

$$f(n) = \mathbb{E}^{\mathrm{ch}} \left(k \frac{\| \wedge^2 W_n \|}{\mathrm{tr}\left(W_n^* W_n\right)} \right). \tag{32}$$

By definition the eigenvalues of $M_n^{\frac{1}{2}}$ are the singular values of $W_n / \sqrt{\mathrm{tr}(W_n^* W_n)}$. Since by Proposition 2.2, M_n converges $\mathbb{P}^{\mathrm{ch}}$ - a.s. to a rank one projection,

$$\lim_{n\to\infty} a_1 \left(\frac{W_n}{\sqrt{\mathrm{tr}\left(W_n^* W_n\right)}} \right) a_2 \left(\frac{W_n}{\sqrt{\mathrm{tr}\left(W_n^* W_n\right)}} \right) = 0 \quad \mathbb{P}^{\mathrm{ch}} \text{ - a.s.}$$

Using Eq. (30) we then conclude that

$$\lim_{n\mapsto\infty} \frac{\left\| \wedge^2 W_n \right\|}{\mathrm{tr}\left(W_n^* W_n\right)} = 0 \quad \mathbb{P}^{\mathrm{ch}} \text{ - a.s.} \tag{33}$$

Since $\| \wedge^2 W_n \| \leq \| W_n \|^2 \leq \mathrm{tr}(W_n^* W_n)$, the expression (32) and Lebesgue's dominated convergence theorem imply $\lim_{n\to\infty} f(n) = 0$.

Second, remark that the function f is submultiplicative. Indeed, for $p, q \in \mathbb{N}$ we have

$$\left\| \wedge^2 \left(v_{p+q} \ldots v_1 \right) \right\| \leq \left\| \wedge^2 \left(v_{p+q} \ldots v_{p+1} \right) \right\| \left\| \wedge^2 \left(v_p \ldots v_1 \right) \right\|$$

and the submultiplicativity follows.

By Fekete's subadditive Lemma, $\frac{\log f(n)}{n}$ converges to $\inf_{n\in\mathbb{N}} \frac{\log f(n)}{n}$, which is (strictly) negative (and possibly equal to $-\infty$) since $f(n) \to 0$. Then there exists $0 < \lambda < 1$ such that $f(n) \leq \lambda^n$ for large enough n, and the conclusion follows. $\square$

We are now in a position to prove Proposition 3.5.

Proof of Proposition 3.5 The Markov property of $(\hat{x}_n)$ implies that

$$\mathbb{E}_\nu \left(d \left(\hat{x}_{n+l}, \hat{y}_n \circ \theta^l \right) \right) = \mathbb{E}_{\nu \Pi^l} \left(d \left(\hat{x}_n, \hat{y}_n \right) \right).$$

Provided inequality (27) is established for $l = 0$, the right hand side of the previous equality can be bounded by $C\lambda^n$. It is hence sufficient to prove the inequality for $l = 0$.

The case $l = 0$ follows from Lemma 3.6 if for any $n \in \mathbb{N}$ and any probability measure ν,

$$\mathbb{E}_\nu \left(d \left(\hat{x}_n, \hat{y}_n \right) \right) \le f(n). \tag{34}$$

To obtain this inequality, note that from the definitions of $\hat{x}_n$, $\hat{y}_n$ and $\hat{z}_n$, we have that

$$\begin{aligned}
d \left(\hat{x}_n, \hat{y}_n \right) &= \frac{\left\| \wedge^2 W_n \left(x_0 \wedge z_n \right) \right\|}{\| W_n x_0 \| \, \| W_n z_n \|} \\
&\le \frac{\left\| \wedge^2 W_n \right\|}{\| W_n x_0 \|^2} \frac{\| W_n x_0 \|}{\| W_n \|} \\
&\le \frac{\left\| \wedge^2 W_n \right\|}{\| W_n x_0 \|^2},
\end{aligned}$$

holds $\mathbb{P}_\nu$-almost surely. To get the first inequality we used $\| W_n z_n \| = \| W_n \|$, and $\| x_0 \wedge z_n \| = d(\hat{x}_0, \hat{z}_n) \le 1$. In addition, by definition of $\mathbb{P}_\nu$,

$$\begin{aligned}
\mathbb{E}_\nu \left(\frac{\left\| \wedge^2 W_n \right\|}{\| W_n x_0 \|^2} \right) &= \int_{P(\mathbb{C}^k) \times M_k(\mathbb{C})^n} \frac{\left\| \wedge^2 W_n \right\|}{\| W_n x_0 \|^2} \| W_n x_0 \|^2 \, d\mu^{\otimes n} \, d\nu(\hat{x}_0) \\
&= \int_{M_k(\mathbb{C})^n} \left\| \wedge^2 W_n \right\| \, d\mu^{\otimes n}(v_1, \ldots, v_n),
\end{aligned}$$

which is $f(n)$. Therefore (34) holds and Lemma 3.6 yields the proof. □

3.3 Convergence in Wasserstein metric

The remainder of Sect. 3 is devoted to the proof of the second part of Theorem 1.1.

Proof of Eq. (6) We are supposed to prove that

$$W_1 \left(\frac{1}{m} \sum_{r=0}^{m-1} \nu \Pi^{mn+r}, \nu_{\text{inv}} \right) = \sup_{f \in \text{Lip}_1(P(\mathbb{C}^k))} \left| \mathbb{E}_\nu \left(\frac{1}{m} \sum_{r=0}^{m-1} f \left(\hat{x}_{mn+r} \right) \right) - \mathbb{E}_{\nu_{\text{inv}}} \left(f \left(\hat{x}_0 \right) \right) \right|$$

is exponentially decaying in n. The expression in the supremum on the right hand side is unchanged by adding an arbitrary constant to f. This freedom allows us to restrict the supremum to functions bounded by 1, i.e. $\| f \|_\infty \le 1$.

Let $f \in \text{Lip}_1(P(\mathbb{C}^k))$ be such a function. Our strategy is to approximate $\hat{x}_{mn+r}$ by $\hat{y}_{mp} \circ \theta^{mq+r}$ with $p = \lfloor \frac{n}{2} \rfloor$ and $q = \lceil \frac{n}{2} \rceil$ so that in particular $p + q = n$. Using telescopic estimates and the invariance of ν_{inv} we then have

$$\left| \mathbb{E}_\nu \left(\frac{1}{m} \sum_{r=0}^{m-1} f\left(\hat{x}_{mn+r} \right) \right) - \mathbb{E}_{\nu_{\text{inv}}} \left(f\left(\hat{x}_0 \right) \right) \right|$$

$$\leq \frac{1}{m} \sum_{r=0}^{m-1} \left| \mathbb{E}_\nu \left(f\left(\hat{x}_{m(p+q)+r} \right) \right) - \mathbb{E}_\nu \left(f\left(\hat{y}_{mp} \circ \theta^{mq+r} \right) \right) \right|$$

$$+ \frac{1}{m} \sum_{r=0}^{m-1} \left| \mathbb{E}_{\nu_{\text{inv}}} \left(f\left(\hat{y}_{mp} \circ \theta^{mq+r} \right) \right) - \mathbb{E}_{\nu_{\text{inv}}} \left(f\left(\hat{x}_{m(p+q)+r} \right) \right) \right|$$

$$+ \left| \frac{1}{m} \sum_{r=0}^{m-1} \mathbb{E}_\nu \left(f\left(\hat{y}_{mp} \circ \theta^{mq+r} \right) \right) - \mathbb{E}_{\nu_{\text{inv}}} \left(f\left(\hat{y}_{mp} \right) \right) \right|.$$

We bound the terms on the right hand side using Proposition 3.5 for the first two terms and Proposition 3.4 for the last term. To this end let C and $\lambda < 1$ be such that the bounds in both these propositions hold true. Since f is 1-Lipschitz continuous we have

$$\left| f\left(\hat{x}_{m(p+q)+r} \right) - f\left(\hat{y}_{mp} \circ \theta^{mq+r} \right) \right| \leq d\left(\hat{x}_{m(p+q)+r}, \hat{y}_{mp} \circ \theta^{mq+r} \right).$$

Proposition 3.5 then implies that

$$\left| \mathbb{E}_\nu \left(f\left(\hat{x}_{m(p+q)+r} \right) \right) - \mathbb{E}_\nu \left(f\left(\hat{y}_{mp} \circ \theta^{mq+r} \right) \right) \right| \leq C\lambda^{mp},$$

and similarly with ν replaced by ν_{inv}. Regarding the last term in the above telescopic estimate we have by Proposition 3.4,

$$\left| \frac{1}{m} \sum_{r=0}^{m-1} \mathbb{E}_\nu \left(f\left(\hat{y}_{mp} \circ \theta^{mq+r} \right) \right) - \mathbb{E}_{\nu_{\text{inv}}} \left(f\left(\hat{y}_{mp} \right) \right) \right| \leq C\lambda^q,$$

where we used the constraint $\| f \|_\infty \leq 1$ discussed at the beginning of the proof.

Putting these estimates together we get

$$\left| \mathbb{E}_\nu \left(\frac{1}{m} \sum_{r=0}^{m-1} f\left(\hat{x}_{mn+r} \right) \right) - \mathbb{E}_{\nu_{\text{inv}}} \left(f\left(\hat{x}_0 \right) \right) \right| \leq 3C\lambda^{\lfloor \frac{n}{2} \rfloor}$$

and this concludes the proof of Eq. (6) and therefore of Theorem 1.1. $\qquad\square$

4 Lyapunov exponents

In this section, we study the almost sure stability exponents. The main results of this section will assume (ϕ-**Erg**) with the additional assumption that $E = \mathbb{C}^k$.

Remark 4.1 Assuming $E = \mathbb{C}^k$ amounts to saying that ϕ has no transient part. Without this assumption, we would have to take into account the almost sure Lyapunov

exponent corresponding to the escape from the transient part. See [3] for a precise account of these ideas.

The relevance of this assumption will stem from the following straightforward inequalities: if ρ is any element of $\mathcal{D}_k$ then one has

$$\frac{d\mathbb{P}^\rho_{|\mathcal{O}_n}}{d\mu^{\otimes n}} \leq \|W_n\|^2,$$

and if ρ is faithful (i.e. definite positive), then

$$\frac{d\mathbb{P}^\rho_{|\mathcal{O}_n}}{d\mu^{\otimes n}} \geq \|\rho^{-1}\|^{-1}\|W_n\|^2.$$

In particular, under the assumption that (ϕ-**Erg**) holds with $E = \mathbb{C}^k$, thus $\rho_{\text{inv}} > 0$ and for any $\rho \in \mathcal{D}_k$, we have

$$\mathbb{P}^\rho \ll \mathbb{P}^{\rho_{\text{inv}}}. \tag{35}$$

Let us start by proving the following lemma, which concerns the ergodicity of θ with respect to the measure $\mathbb{P}^{\rho_{\text{inv}}}$.

Lemma 4.2 *Assume that (ϕ-**Erg**) holds. Then the shift θ on $(\Omega, \mathcal{O})$ is ergodic with respect to the probability measure $\mathbb{P}^{\rho_{\text{inv}}}$.*

Proof Let A, A' in $\mathcal{O}_l$. From the definition of $\mathbb{P}^{\rho_{\text{inv}}}$, for j large enough, $\mathbb{P}^{\rho_{\text{inv}}}\left(A \cap \theta^{-j}(A')\right)$ equals

$$\int_{A \times A'} \text{tr}\left(v'_l \ldots v'_1 \phi^{j-l}\left(v_l \ldots v_1 \rho_{\text{inv}} v_1^* \ldots v_l^*\right) v_1'^* \ldots v_l'^*\right) d\mu^{\otimes l}(v_1, \ldots, v_l)\, d\mu^{\otimes l}\left(v'_1, \ldots, v'_l\right),$$

and the Perron–Frobenius Theorem 3.3 implies

$$\lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} \phi^j\left(v_l \ldots v_1 \rho_{\text{inv}} v_1^* \ldots v_l^*\right) = \text{tr}\left(v_l \ldots v_1 \rho_{\text{inv}} v_1^* \ldots v_l^*\right) \rho_{\text{inv}}$$

for $\mu^{\otimes l}$-almost all $(v_1, \ldots, v_l)$ so that

$$\lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} \mathbb{P}^{\rho_{\text{inv}}}\left(A \cap \theta^{-j}\left(A'\right)\right) = \mathbb{P}^{\rho_{\text{inv}}}(A)\, \mathbb{P}^{\rho_{\text{inv}}}\left(A'\right),$$

which proves the ergodicity. $\square$

Now we can state our result concerning Lyapunov exponents.

Proposition 4.3 *Assume that* (φ**-Erg**) *holds with* $E = \mathbb{C}^k$, *and that* (**Pur**) *holds. Assume* $\int \|v\|^2 \log \|v\| \mathrm{d}\mu(v) < \infty$. *Then there exists numbers*

$$\infty > \gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_k \geq -\infty$$

such that for any probability measure ν *over* $(\mathrm{P}(\mathbb{C}^k), \mathcal{B})$*:*

(1) *for any* $p \in \{1, \ldots, k\}$,

$$\lim_{n\to\infty} \frac{1}{n} \log \left\| \wedge^p W_n \right\| = \sum_{j=1}^{p} \gamma_j, \quad \mathbb{P}_\nu\text{-a.s.,} \tag{36}$$

(2) $\gamma_2 - \gamma_1 < 0$ *with* $\gamma_2 - \gamma_1$ *understood as the limit of* $\frac{1}{n} \log \frac{\|\wedge^2 W_n\|}{\|W_n\|^2}$ *whenever* $\gamma_1 = -\infty$,

(3) *one has the convergence*

$$\lim_{n\to\infty} \frac{1}{n} (\log \|W_n x_0\| - \log \|W_n\|) = 0 \quad \mathbb{P}_\nu\text{-a.s.} \tag{37}$$

Proof Let us start by proving (1). Note that $n \mapsto \log \| \wedge^p W_n \|$ is subadditive by definition. The existence of the $\mathbb{P}^{\rho_{\mathrm{inv}}}$-a.s. limits $\lim_{n\to\infty} \frac{1}{n} \log \| \wedge^p W_n \|$ then follows from $\mathbb{E}^{\rho_{\mathrm{inv}}} (\log \|V\|^2) \leq \int \|v\|^2 \log \|v\|^2 d\mu(v) < \infty$, $\mathbb{P}^{\rho_{\mathrm{inv}}} \circ \theta^{-1} = \mathbb{P}^{\rho_{\mathrm{inv}}}$ and a direct application of Kingman's subadditive ergodic theorem (see e.g. [21]). The fact that these limits are $\mathbb{P}^{\rho_{\mathrm{inv}}}$-a.s. constant comes from the θ-ergodicity of $\mathbb{P}^{\rho_{\mathrm{inv}}}$ proved in Lemma 4.2. Since by Eq. (35) any $\mathbb{P}^\rho$ is absolutely continuous with respect to $\mathbb{P}^{\rho_{\mathrm{inv}}}$, Proposition 2.1 and the $\mathcal{O}$-measurability of $\| \wedge^p W_n \|$ imply the convergence holds $\mathbb{P}_\nu$-a.s. The numbers γ_j are uniquely defined, by defining $\sum_{j=1}^{p} \gamma_j$ as the $\mathbb{P}^{\rho_{\mathrm{inv}}}$-a.s. limit $\lim_{n\to\infty} \frac{1}{n} \log \| \wedge^p W_n \|$ and imposing the rule that $\gamma_{j+1} = -\infty$ if $\gamma_j = -\infty$. This convention and (30) impose that $\gamma_{j+1} \leq \gamma_j$ for $j = 1, \ldots, k-1$.

Concerning (2), recall the quantity $f(n)$ defined in Eq. (31). Then Eq. (32) and the inequality tr $W_n^* W_n \leq k\|W_n\|^2$ give

$$f(n) \geq \mathbb{E}^{\mathrm{ch}} \left(\frac{\|\wedge^2 W_n\|}{\|W_n\|^2} \right).$$

Jensen's inequality implies

$$\frac{1}{n} \log f(n) \geq \mathbb{E}^{\mathrm{ch}} \left(\frac{1}{n} \log \frac{\|\wedge^2 W_n\|}{\|W_n\|^2} \right)$$

so that by Lemma 3.6 and Fatou's lemma, $\log \lambda \geq \gamma_2 - \gamma_1$ with $\lambda \in (0, 1)$.

Finally for (3), from Proposition 2.2, we have

$$\lim_{n \to \infty} \frac{\|W_n x_0\|}{\|W_n\|} = \lim_{n \to \infty} \frac{\left\| M_n^{\frac{1}{2}} x_0 \right\|}{\left\| M_n^{\frac{1}{2}} \right\|} = |\langle x_0, z \rangle| \quad \mathbb{P}_\nu \text{ - a.s.}$$

Since $\mathbb{P}_\nu$ - a.s. $|\langle x_0, z \rangle| > 0$, the proposition holds. $\qquad \square$

From this proposition we deduce the following almost sure convergence rate for the distance between the Markov chain $(\hat{x}_n)$ and the $(\mathcal{O}_n)$-adapted process $(\hat{y}_n)$.

Proposition 4.4 *Assume* **(Pur)** *holds and* **(ϕ-Erg)** *holds with* $E = \mathbb{C}^k$. *Then, for any probability measure ν on* $(P(\mathbb{C}^k), \mathcal{B})$,

$$\limsup_{t \to \infty} \frac{1}{n} \log \left(d\left(\hat{x}_n, \hat{y}_n \right) \right) \le -(\gamma_1 - \gamma_2) < 0, \quad \mathbb{P}_\nu \text{ - a.s.}$$

Proof Identity (28) and the definition of $\hat{z}_n$ imply

$$d(\hat{x}_n, \hat{y}_n) = \frac{\left\| \wedge^2 W_n \, x_0 \wedge z_n \right\|}{\|W_n x_0\| \, \|W_n z_n\|} \le \frac{\left\| \wedge^2 W_n \right\|}{\|W_n x_0\| \, \|W_n\|}.$$

Proposition 4.3 then yields the result. $\qquad \square$

Appendix A: Equivalence of (Pur) and contractivity

We assume $\text{supp}\,\mu \subset \text{GL}_k(\mathbb{C})$. Recall that T_μ is the smallest closed sub-semigroup of $\text{GL}_k(\mathbb{C})$ that contains $\text{supp}\,\mu$. It is said to be contracting if there exists a sequence $(a_n)_{n \in \mathbb{N}} \subset T_\mu$ such that $\lim_{n \to \infty} a_n/\|a_n\|$ exists and is a rank one matrix.

Proposition A.1 *Assume* $\text{supp}\,\mu \subset \text{GL}_k(\mathbb{C})$ *and T_μ is strongly irreducible. Then μ verifies* **(Pur)** *if and only if T_μ is contracting.*

Proof By Proposition 2.2 the implication **(Pur)** $\Rightarrow$ contractivity follows by taking for (a_n) a convergent subsequence of $(W_n(\omega))$ for $\omega \in \text{supp}\,\mathbb{P}^{\text{ch}}$.

We prove the opposite implication by contradiction. Following [12, Lemma 3], under the assumptions of the proposition, T_μ is contracting if and only if, for any two $\hat{x}, \hat{y} \in P(\mathbb{C}^k)$ there exists a sequence of matrices $(a_n) \subset T_\mu$ such that

$$\lim_{n \to \infty} d\left(a_n \cdot \hat{x}, a_n \cdot \hat{y} \right) = 0.$$

Now, assume that contractivity holds but **(Pur)** does not. Namely, that T_μ is contracting but there exists an orthogonal projector π of rank ≥ 2, such that for any $a \in T_\mu$,

$$\pi a^* a \pi \propto \pi.$$

Let x, y in the range of π be orthonormal vectors. Then $\langle ax, ay \rangle = \langle x, y \rangle = 0$, and $\|ax\|$, $\|ay\|$ are nonzero, so that $d(a \cdot \hat{x}, a \cdot \hat{y}) = 1$. As this is true for any a in T_μ, contractivity cannot hold. This contradiction yields the proposition.   □

Appendix B: Set of invariant measures under assumption (Pur)

A quantum channel is a map ϕ on $M_k(\mathbb{C})$ of a form

$$\phi(\rho) = \int_{M_k(\mathcal{C})} v \rho v^* d\mu(v),$$

where μ is a measure satisfying the normalization condition (1). The decomposition of quantum channels to irreducible components was derived in [2,5,22]. The space $\mathbb{C}^k$ is decomposed into orthogonal subspaces, one subspace is transient and in all other subspaces the map has a canonical tensor product structure. We recall these results.

There exists a decomposition

$$\mathbb{C}^k \simeq \mathbb{C}^{n_1} \oplus \cdots \oplus \mathbb{C}^{n_d} \oplus \mathbb{C}^D, \quad k = n_1 + \cdots + n_d + D$$

with the following properties. We denote by $v^{(j)}$ the restriction of v to $\mathbb{C}^{n_j}$.

(e1) All invariant states are supported in the subspace $L = \mathbb{C}^{n_1} \oplus \cdots \oplus \mathbb{C}^{n_d} \oplus 0$,

(e2) The restriction of v to this subspace is block diagonal,

$$v|_L = v^{(1)} \oplus \cdots \oplus v^{(d)} \oplus 0 \quad \mu\text{-a.e.} \tag{38}$$

(e3) For each $j = 1, \ldots, d$ there is a decomposition $\mathbb{C}^{n_j} = \mathbb{C}^{k_j} \otimes \mathbb{C}^{m_j}$, $n_j = k_j m_j$, a unitary matrix U_j on $\mathbb{C}^{n_j}$ and a matrix $\tilde{v}^{(j)}$ on $\mathbb{C}^{k_j}$ such that

$$v^{(j)} = U_j \left(\tilde{v}^{(j)} \otimes \mathrm{Id}_{\mathbb{C}^{m_j}} \right) U_j^* \quad \mu - a.s. \tag{39}$$

(e4) There exists a full rank positive matrix ρ_j on $\mathbb{C}^{k_j}$ such that

$$0 \oplus \cdots \oplus U_j \left(\rho_j \otimes \mathrm{Id}_{\mathbb{C}^{m_j}} \right) U_j^* \oplus \cdots \oplus 0 \tag{40}$$

is a fixed point of ϕ.

It follows from (e3) and (e4) that the set of fixed points for ϕ is

$$U_1 \big(\rho_1 \otimes M_{m_1}(\mathbb{C}) \big) U_1^* \oplus \cdots \oplus U_d \big(\rho_d \otimes M_{m_d}(\mathbb{C}) \big) U_d^* \oplus 0_{M_D(\mathbb{C})}.$$

The decomposition simplifies under the purification assumption.

Proposition B.1 *Assume* **(Pur)** *holds. Then there exists a set* $\{\rho_j\}_{j=1}^d$ *of positive definite matrices and an integer* D *such that the set of* ϕ *fixed points is*

$$\mathbb{C}\rho_1 \oplus \cdots \oplus \mathbb{C}\rho_d \oplus 0_{M_D(\mathbb{C})}.$$

Proof The statement follows from the discussion preceding the proposition if we show that **(Pur)** implies $m_1 = \cdots = m_d = 1$. Assume that one of the m_j, e.g. m_1, is greater than 1. Let x be a norm one vector in $\mathbb{C}^{k_1}$. Then $\pi = U_1\pi_{\hat{x}} \otimes \mathrm{Id}_{\mathbb{C}^{m_1}} U_1^* \oplus 0 \oplus \cdots \oplus 0$ is a projection with rank bigger than 1, and by Eq. (39) we have, in the notation of (38) and (39),

$$\pi v_1^* \ldots v_n^* v_n \ldots v_1 \pi = \left\| \tilde{v}_n^{(1)} \ldots \tilde{v}_1^{(1)} x \right\|^2 \pi$$

for $\mu^{\otimes n}$-almost all $v_1, \ldots, v_n$. This contradicts **(Pur)**. $\qquad\square$

It is clear from Eq. (38) that to each extremal fixed point $0 \oplus \cdots \oplus \rho_j \oplus \cdots \oplus 0$ corresponds a unique invariant measure ν_j supported on its range F_j. The converse is the subject of the next proposition.

Proposition B.2 *Assume* **(Pur)** *holds. Then any* Π*-invariant probability measure is a convex combination of the measures* ν_j, $j = 1, \ldots, d$.

Proof Let ν be a Π-invariant probability measure. Let f be a continuous function. From Lemma 2.3,

$$\mathbb{E}_\nu(f) = \lim_{n \to \infty} \mathbb{E}_\nu\big(f\left(U_n \cdot \hat{z}\right)\big).$$

Proposition 2.1 implies

$$\mathbb{E}_\nu(f) = \lim_{n \to \infty} \mathbb{E}^{\rho_\nu}\big(f\left(U_n \cdot \hat{z}\right)\big)$$

with $\rho_\nu \in \mathcal{D}_k$ a fixed point of ϕ. By Proposition B.1, **(Pur)** implies that there exist non negative numbers $t_1, \ldots, t_d$ summing up to one such that $\rho_\nu = t_1\rho_1 \oplus \cdots \oplus t_d\rho_d \oplus 0_{M_D(\mathbb{C})}$. From the definition of $\mathbb{P}^{\rho_\nu}$,

$$\mathbb{P}^{\rho_\nu} = t_1\mathbb{P}^{\rho_1} + \cdots + t_d\mathbb{P}^{\rho_d}$$

where we used the abuse of notation $\rho_j \equiv 0 \oplus \cdots \oplus \rho_j \oplus \cdots \oplus 0$. Using Proposition 2.1, it follows that

$$\mathbb{E}_\nu(f) = \lim_{n \to \infty} t_1 \mathbb{E}_{\nu_1}\big(f\left(U_n \cdot \hat{z}\right)\big) + \cdots + t_d \mathbb{E}_{\nu_d}\big(f\left(U_n \cdot \hat{z}\right)\big).$$

Then Lemma 2.3 and the Π-invariance of each measure ν_j yield the proposition. $\quad\square$

Appendix C: Products of special unitary matrices

Proposition C.1 *Assume* supp $\mu \subset$ SU(k). *Let G be the smallest closed subgroup of* SU(k) *such that* supp $\mu \subset G$. *For any $\hat{x} \in$ P$(\mathbb{C}^k)$, let $[\hat{x}]_G$ be the orbit of $\hat{x}$ with respect to G and the action $G \times$ P$(\mathbb{C}^k) \ni (v, \hat{x}) \mapsto v \cdot \hat{x}$. Namely, $[\hat{x}]_G := \{\hat{y} \in$ P$(\mathbb{C}^k) \mid \exists v \in G$ s.t. $\hat{y} = v \cdot \hat{x}\}$. Then, for any $\hat{x}$, there exists a unique Π-invariant probability measure supported on $[\hat{x}]_G$, and this unique invariant measure is uniform in the sense that for any $v \in G$ it is invariant by the map $\hat{x} \mapsto v \cdot \hat{x}$.*

Corollary C.2 *With the assumption and definitions of the last proposition, if $G = $ SU(k), Π has a unique invariant probability measure and this probability is the uniform one on* P$(\mathbb{C}^k)$.

Proof The corollary being a trivial consequence of $G = $ SU$(k) \Rightarrow [\hat{x}]_G = $ P$(\mathbb{C}^k) \, \forall \hat{x} \in$ P$(\mathbb{C}^k)$, we are left with proving the proposition.

Let P_μ be the Markov kernel on G defined by the left multiplication: $P_\mu f(v) = \int_G f(uv) d\mu(u)$. Since G is compact as a closed subset of SU(k), following [1, Proposition 4.8.1, Theorem 4.8.2], the unique P_μ-invariant probability measure μ_G on G is the normalized Haar measure on G. Since G is compact, Prokhorov's theorem implies that for any $u \in G$,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \delta_u P_\mu^k = \mu_G \quad \text{weakly.} \tag{41}$$

Let $\hat{x} \in$ P$(\mathbb{C}^k)$. Since supp $\mu \subset G$, for any $\hat{y} \in [\hat{x}]_G$, $\Pi(\hat{y}, [\hat{x}]_G) = 1$. Then, $[\hat{x}]_G$ being compact, there exists a Π-invariant measure ν supported on $[\hat{x}]_G$.

Let f be a continuous function on $[\hat{x}]_G$. Then,

$$\nu(f) = \frac{1}{n} \sum_{k=1}^{n} \nu \Pi^k f = \frac{1}{n} \sum_{k=1}^{n} \int_{G^k \times [\hat{x}]_G} f\left(v_k \ldots v_1 \cdot \hat{y}\right) d\mu^{\otimes k}(v_1, \ldots, v_k) d\nu(\hat{y}).$$

For each $\hat{y} \in [\hat{x}]_G$ let $u_y \in G$ be such that $\hat{y} = u_y \cdot \hat{x}$. The map $v \mapsto v u_y \cdot \hat{x}$ being continuous, setting $u = u_y$, the weak convergence (41) and Lebesgue's dominated convergence theorem imply,

$$\nu(f) = \int_G f\left(v \cdot \hat{x}\right) d\mu_G(v).$$

It follows that ν is the image measure of μ_G by the application $v \mapsto v \cdot \hat{x}$. The left multiplication invariance of the Haar measure μ_G yields the invariance of ν by the map $\hat{x} \mapsto v \cdot \hat{x}$ for any $v \in G$. $\square$

Example C.3 Let $\mu = \frac{1}{2}(\delta_{v_1} + \delta_{v_2})$ with,

$$v_1 = \begin{pmatrix} e^i & 0 \\ 0 & e^{-i} \end{pmatrix} \quad \text{and} \quad v_2 = \begin{pmatrix} \cos 1 & i \sin 1 \\ i \sin 1 & \cos 1 \end{pmatrix}.$$

Then $G = \mathrm{SU}(2)$ and the uniform measure on $\mathrm{P}(\mathbb{C}^2)$ is the unique Π-invariant probability measure.

Proof Following Proposition C.1, it is sufficient to prove that any element of $\mathrm{SU}(2)$ is the limit of a sequence of products of v_1 and v_2.

Let $\sigma_1, \sigma_2, \sigma_3$ be the usual Pauli matrices:

$$\sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{and} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The Pauli matrices being generators of $\mathrm{SU}(2)$ in its fundamental representation, for any $u \in \mathrm{SU}(2)$, there exist three reals $\theta_1, \theta_2, \theta_3 \in \mathbb{R}$ s.t.,

$$u = \exp(i(\theta_1\sigma_1 + \theta_2\sigma_2 + \theta_3\sigma_3)).$$

Especially, $v_1 = \exp(i\sigma_3)$ and $v_2 = \exp(i\sigma_1)$. Since for any $j = 1, 2, 3$, $\exp(i\theta_j\sigma_j) = \exp(i(\theta_j + 2\pi)\sigma_j)$, taking limits of sequences of powers of v_1 or v_2, for any $\theta \in \mathbb{R}$, both

$$e^{i\theta\sigma_1} \quad \text{and} \quad e^{i\theta\sigma_3}$$

are elements of G. It remains to show that any $u \in \mathrm{SU}(2)$ is a product of elements equal to $\exp(i\theta\sigma_1)$ or $\exp(i\theta\sigma_3)$ with θ real.

Fix $(\theta_1, \theta_2, \theta_3) \in \mathbb{R}^3$. Then using spherical coordinates in $\mathbb{R}^3$, there exist $r \in \mathbb{R}_+$, $\theta \in [0, \pi]$ and $\varphi \in [0, 2\pi[$ such that $\theta_1 = r\cos\theta$, $\theta_2 = r\sin\theta\cos\varphi$ and $\theta_3 = r\sin\theta\sin\varphi$. Then by direct computation,

$$e^{i(\theta_1\sigma_1 + \theta_2\sigma_2 + \theta_3\sigma_3)} = e^{-i\frac{\varphi}{2}\sigma_1} e^{i\frac{\theta}{2}\sigma_3} e^{ir\sigma_1} e^{-i\frac{\theta}{2}\sigma_3} e^{i\frac{\varphi}{2}\sigma_1}.$$

It follows that as a product of elements of G, $e^{i(\theta_1\sigma_1 + \theta_2\sigma_2 + \theta_3\sigma_3)} \in G$, hence $G = \mathrm{SU}(2)$ and the example holds. $\square$

Example C.4 Let $\mu = \frac{1}{2}(\delta_{v_1} + \delta_{v_2})$ with,

$$v_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \quad \text{and} \quad v_2 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Then $G = \{\pm\mathrm{Id}_{\mathbb{C}^2}, \pm v_1, \pm v_2, \pm v_1 v_2\}$. For $z \in \mathbb{C}$, let $e_z = (1, z)^\mathsf{T}$ and $e_\infty = (0, 1)^\mathsf{T}$. With the conventions $\infty^{-1} = 0$, $0^{-1} = \infty$ and $-\infty = \infty$, for any $z \in \mathbb{C} \cup \{\infty\}$, $[\hat{e}_z]_G = \{\hat{e}_z, \hat{e}_{z^{-1}}, \hat{e}_{-z}, \hat{e}_{-z^{-1}}\}$ and the measure $\frac{1}{4}(\delta_{\hat{e}_z} + \delta_{\hat{e}_{-z}} + \delta_{\hat{e}_{z^{-1}}} + \delta_{\hat{e}_{-z^{-1}}})$ is a Π-invariant probability measure.

The proof of this example is obtained by an explicit computation.

References

1. Applebaum, D.: Probability on Compact Lie Groups, Volume 70 of Probability Theory and Stochastic Modelling. Springer, Berlin (2014)
2. Baumgartner, B., Narnhofer, H.: The structures of state space concerning quantum dynamical semigroups. Rev. Math. Phys. **24**(02), 1250001 (2012)
3. Benoist, T., Pellegrini, C., Ticozzi, F.: Exponential stability of subspaces for quantum stochastic master equations. Ann. Henri Poincaré **18**, 2045–2074 (2017)
4. Bougerol, P., Lacroix, J.: Products of Random Matrices with Applications to Schrödinger Operators, Volume 8 of Progress in Probability and Statistics. Birkhäuser Boston, Inc., Boston (1985)
5. Carbone, R., Pautrat, Y.: Irreducible decompositions and stationary states of quantum channels. Rep. Math. Phys. **77**(3), 293–313 (2016)
6. Carmichael, H.: An Open Systems Approach to Quantum Optics: Lectures Presented at the Université Libre de Bruxelles, October 28 to November 4, 1991. Springer, Berlin (1993)
7. Evans, D.E., Høegh-Krohn, R.: Spectral properties of positive maps on C^*-algebras. J. Lond. Math. Soc. (2) **17**(2), 345–355 (1978)
8. Furstenberg, H., Kesten, H.: Products of random matrices. Ann. Math. Stat. **31**(2), 457–469 (1960)
9. Guerlin, C., Bernu, J., Deleglise, S., Sayrin, C., Gleyzes, S., Kuhr, S., Brune, M., Raimond, J.-M., Haroche, S.: Progressive field-state collapse and quantum non-demolition photon counting. Nature **448**(7156), 889–893 (2007)
10. Guivarc'h, Y., Le Page, É.: Spectral gap properties for linear random walks and Pareto's asymptotics for affine stochastic recursions. Ann. Inst. H. Poincaré Probab. Stat. **52**(2), 503–574 (2016)
11. Guivarc'h, Y., Raugi, A.: Frontière de Furstenberg, propriétés de contraction et théorèmes de convergence. Probab. Theory Relat. Fields **69**(2), 187–242 (1985)
12. Guivarc'h, Y., Raugi, A.: Products of random matrices: convergence theorems. In: Cohen, J.E., Kesten, H., Newman, C.M. (eds.) Random Matrices and Their Applications (Brunswick, Maine, 1984), Volume 50 of Contemporary Mathematics, pp. 31–54. American Mathematical Society, Providence (1986)
13. Holevo, A.: Statistical Structure of Quantum Theory. Springer, Berlin (2001)
14. Kümmerer, B., Maassen, H.: An ergodic theorem for quantum counting processes. J. Phys. A **36**(8), 2155 (2003)
15. Kümmerer, B., Maassen, H.: A pathwise ergodic theorem for quantum trajectories. J. Phys. A **37**(49), 11889–11896 (2004)
16. Le Page, E.: Theoremes limites pour les produits de matrices aleatoires. In: Heyer, H. (ed.) Probability Measures on Groups. Lecture Notes in Mathematics, vol. 928. Springer, Berlin, Heidelberg (1982)
17. Maassen, H., Kümmerer, B.: Purification of quantum trajectories. Lect. Notes Monogr. Ser. **48**, 252–261 (2006)
18. Mac Lane, S., Birkhoff, G.: Algebra, 3rd edn. Chelsea Publishing Co., New York (1988)
19. Meyn, S., Tweedie, R.L.: Markov Chains and Stochastic Stability, 2nd edn. Cambridge University Press, Cambridge (2009)
20. Schrader, R.: Perron-Frobenius theory for positive maps on trace ideals. In: Mathematical Physics in Mathematics and Physics (Siena, 2000), vol. 30, pp. 361–378 (2001)
21. Walters, P.: An Introduction to Ergodic Theory. Graduate Texts in Mathematics, vol. 79. Springer, Berlin (1982)
22. Wolf, M.M.: Quantum channels & operations: guided tour. http://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QChannelLecture.pdf (2012). Lecture notes based on a course given at the Niels–Bohr Institute. Accessed 28 Feb 2017