



# An Ergodic Theorem for Quantum Processes with Applications to Matrix Product States

Ramis Movassagh<sup>1</sup> , Jeffrey Schenker<sup>2</sup> 

<sup>1</sup> IBM Quantum, MIT-IBM Watson AI Lab, Cambridge, MA 02142, USA. E-mail: ramis@us.ibm.com

<sup>2</sup> Department of Mathematics, Michigan State University, East Lansing, MI 48824, USA.  
E-mail: jeffrey@math.msu.edu

Received: 27 September 2021 / Accepted: 19 June 2022

Published online: 18 July 2022 – © The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

**Abstract:** Any discrete quantum process is represented by a sequence of quantum channels. We consider ergodic quantum processes obtained by a map that takes the points along the trajectory of a discrete ergodic dynamical system to the space of quantum channels. Under a natural irreducibility condition, we obtain a theorem showing that the state under such a process converges exponentially fast to an ergodic sequence depending on the process, but independent of the initial state. As an application, we describe the thermodynamic limit of ergodic matrix product states and prove that the 2-point correlations of local observables in such states decay exponentially with their distance in the bulk.

## 1. Introduction

The change of a physical system over a discrete unit of time, including the internal dynamics and interaction with the environment, can be represented by a quantum channel. The evolution of the system at discrete times is then obtained by the application of a sequence of quantum channels, which may be termed a *quantum process*. Mathematically, a quantum channel is a completely positive and trace preserving linear transformation of the system's density matrix,  $\rho \mapsto \phi(\rho)$ . In a finite dimensional Hilbert space, any such map can be written in the Kraus form [14]

$$\phi(\rho) = \sum_{i=1}^d B^i \rho B^{i\dagger}, \quad (1.1)$$

where  $\dagger$  denotes the adjoint (conjugate transpose) and the following holds

$$\sum_{i=1}^d B^{i\dagger} B^i = \mathbb{I}. \quad (1.2)$$

The net change in the state resulting from the quantum process is obtained from composition of the channels acting on the initial state

$$\rho_n = \phi_n \circ \cdots \circ \phi_1(\rho_0) = \sum_{i_1, \dots, i_n=1}^d B_n^{i_n} \cdots B_1^{i_1} \rho_0 B_1^{i_1 \dagger} \cdots B_n^{i_n \dagger}. \quad (1.3)$$

In the present work, we study general *ergodic sequences* of channels in the following sense. Consider a map from the points of a discrete, ergodic dynamical system  $\Omega$  to the space quantum channels. Starting from any point on  $\Omega$ , we obtain an ergodic sequence of quantum channels by evaluating the map at the points along the corresponding trajectory.

Here we answer the following questions: What is the action of an ergodic composition given by equation (1.3)? Is there a convergence to a simple and general limit? We obtain a general theorem (Theorem 1) for an ergodic sequence of quantum channels, with an underlying assumption of non-negligible decoherence. This theorem states that the sequence of states  $\rho_n$  converges to a fixed-point sequence that only depends on the sequence of channels and is independent of the initial state. Theorem 2 then shows that the composition of such channels converges exponentially fast to a stochastic sequence of replacement (rank-one) channels. A corollary of this result is the well-known convergence in the translation invariant case to a fixed replacement channel.

Theorems 1 and 2 also apply to sequences of completely positive maps, without imposing the trace preserving condition Eq. (1.2). Such sequences are naturally related to the expectation values of observables in a matrix product state (MPS). We apply our results to an ergodic MPS, wherein the matrices in the MPS form an ergodic sequence. We derive a formula for the expectation values of observables in an MPS. We then prove (Theorem 3) that the correlation functions of local observables decay exponentially with their distance.

**1.1. Background and relation to other works.** The generic aspects of the behavior of quantum systems have long been of interest. However, because of the theoretical challenge of dealing with the general case, in the past ‘ergodic’ quantum channels were considered in various works, each of which, to the best of our knowledge, is a very special subset of possibilities in this work. For example, in [3], a channel was chosen at random from some ensemble and then repeatedly applied, i.e., the sets  $\{B_k^i : i = 1, \dots, d\}$  were all equal. In [2], time dynamics were analyzed for a quantum system with repeated independently chosen random interactions with an environment. Other instances studied include certain independent random channels and their compositions (e.g., from a finite set of random isometries) [4,5]. See [6] for a review. Our work considers a general ergodic sequence and therefore serves as a vast generalization of the past work. In particular, this work allows for long-range correlations among the channels, or even pseudo-randomness generated by quasi-periodic dynamics. This includes the previously considered extreme cases of independently and identically distributed (iid) and (time)-translation invariant channels.

The formalism of quantum channels naturally lends itself to the calculation of expectation values of observables and correlation functions of local observables of physical low-dimensional quantum systems, which are well described by density matrix renormalization group [26] and its natural representation in terms of MPS [24]. Previous works on matrix product states have focused on the translation invariant case [1,8,21]. Theorem 2 allows us to move beyond the translation invariant case to analyze the thermodynamic limit of ergodic (one-dimensional) MPS. The ergodic MPS that we consider

may be translation invariant, quasi-periodic, or formed from random matrices with arbitrary correlations.

In order to reify our theory, in a companion paper [18] we apply our main result (Theorem 2) to the translation invariant case as well as a natural example in which each channel is an independent random Haar isometry:

$$\phi_j(\rho) = \text{tr}_r[U_j \rho \otimes Q_r U_j^\dagger],$$

where  $Q_r$  is a pure state on  $\mathbb{C}^{r \times r}$ ,  $U_j$  is a sequence of independent Haar distributed  $Dr \times Dr$  unitaries, and  $\text{tr}_r$  is the partial trace from  $\mathbb{C}^{Dr \times Dr}$  to  $\mathbb{C}^{D \times D}$ . We analyze the asymptotics with respect to the dimension of the environment ( $r$ ), or of the system ( $D$ ), or both tending to infinity, and prove that the limiting states  $\rho_n$  are given by

$$\rho_n = \frac{1}{D} \mathbb{I}_D + \frac{1}{\sqrt{1+rD^2}} W_n,$$

where  $W_n$  are asymptotically Gaussian with distribution proportional to  $e^{-\frac{D}{2} \text{tr}[W^2]} \delta(\text{tr}[W])$ , where  $\delta$  is the Dirac delta measure. We also present consequences for ergodic MPSs in [18], using Theorem 3 and the theory developed in Sect. 2.3 of the present paper to analytically compute the entanglement spectrum of an ergodic MPS across any cut as well as the one- and two-point correlation functions in an ergodic MPS.

**1.2. Physical implications.** In the companion paper [18], we discuss the physical consequences of the theorems presented here. Theorems 1 and 2, to the best of our knowledge, are the first general theorems proved that apply to correlated quantum processes. Similarly, Theorem 3 for the first time demonstrates a general exponential decay of correlations for ergodic MPS, and therefore, a vast class of ground states of interacting quantum matter.

Physically realistic quantum processes inevitably have temporal correlations, even if the underlying process is Markovian. In the latter case, any two consecutive times are correlated. Similarly, correlated quantum channels arise naturally in the context of MPS for the study of non-trivial systems and states of interacting quantum many-body systems. For example, in any finite system simulation of one-dimensional systems, the matrices that result in the density matrix renormalization group procedure will inevitably be correlated. As such the consideration of iid channels and the MPS formed from them is mostly of theoretical interest.

Three physical corollaries of our theorems are [18]:

(1) Engineered non-equilibrium phases of matter realized by time-periodic driven Hamiltonians (e.g., in Floquet systems) [16, 17, 23], are *only* meta-stable in presence of interactions with an environment at positive temperature.

(2) An ergodic sequence of quantum channels with non-negligible decoherence converges to the same final sequence irrespective of the initial state. These channels are asymptotically replacement channels, which implies that the process cannot even convey classical information with respect to the initial state. This is intuitively seen from a unique fixed point that is reached irrespective of the input quantum state. The channels may be very correlated or even time-translation invariant however. For example, in the near-term quantum computing era when the random quantum circuits have decoherence at each step of computation, the initial memory of the state is exponentially lost with the number of applied gates.

**(3)** It was previously proved that a non-vanishing gap in the thermodynamic limit implies an exponential decay of correlations [11, 19], and that in one-dimension an area law for entanglement entropy holds [10]. Brandão and Horodecki [1] proved that in one-dimension the exponential decay of correlations implies an area law. We prove somewhat of a partial converse, that says finitely correlated states with an ergodic MPS representation have correlation function that decay exponentially with distance.

*1.3. Illustration.* One of the simplest examples of a channel is the amplitude damping channel of a qubit. The state space of a single qubit can be identified with the closed unit ball in  $\mathbb{R}^3$ , the so-called *Bloch sphere*. A point  $\vec{v} = (v_1, v_2, v_3)$  in the Bloch sphere corresponds to the density matrix  $\rho_{\vec{v}} = \frac{1}{2}(I + \vec{v} \cdot \vec{\sigma})$ , where  $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$  are the 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 amplitude damping channel with rate  $\gamma \in [0, 1]$  and axis  $\vec{u} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$  is the completely positive trace preserving map

$$\phi_{\gamma, \vec{u}}(\rho) = E \rho E^\dagger + F \rho F^\dagger,$$

where

$$E = S_{\vec{u}} \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix} S_{\vec{u}}^\dagger \quad \text{and} \quad F = S_{\vec{u}} \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix} S_{\vec{u}}^\dagger,$$

with  $S_{\vec{u}}$  defined by

$$S_{\vec{u}} = \begin{pmatrix} \cos \theta/2 & -e^{-i\phi} \sin \theta/2 \\ e^{i\phi} \sin \theta/2 & \cos \theta/2 \end{pmatrix}.$$

This channel models relaxation, with the rate  $\gamma$ , of a qubit to the pure state  $\frac{1}{2}(I + \vec{u} \cdot \vec{\sigma})$ , i.e., the ground state of the spin Hamiltonian  $H_{\vec{u}} = -\vec{u} \cdot \vec{\sigma}$ . More generally, one may consider relaxation coupled with the Schrödinger dynamics of  $H_{\vec{u}}$  over an interval  $\delta$  to obtain the generalized amplitude damping channel

$$\phi_{\gamma, \vec{u}, \delta}(\rho) = \phi_{\gamma, \vec{u}} \left( e^{-i\delta H_{\vec{u}}} \rho e^{i\delta H_{\vec{u}}} \right).$$

One may obtain a large family of ergodic quantum processes, for instance, by allowing  $\vec{u}$  to evolve according to an ergodic process on the Bloch sphere  $S^2$  (with  $\gamma$  and  $\delta$  fixed). Such processes would model the relaxation of a qubit toward an axis that fluctuates with time, such as might be expected if the qubit Hamiltonian fluctuates while the interaction with the environment remains fixed. The longtime behavior of such processes can depend quite strongly on the nature of the ergodic process on  $S^2$ . More generally one could allow for all three parameters  $(\gamma, \vec{u}, \delta)$  to evolve according to an ergodic process on  $[0, 1] \times S^2 \times [0, \infty)$ .

In Fig. 1 we plot the first 2000 steps  $\rho_j = \phi_j \circ \dots \circ \phi_1(\rho)$ ,  $j = 1, \dots, 2000$  for four distinct ergodic quantum processes of the form  $\phi_j = \phi_{\gamma, \vec{u}_j, \delta}$ ,  $\vec{u}_j = (\cos \alpha_j, \sin \alpha_j, 0)$ . The four processes considered are

- (1) Random channels, with  $\delta = \pi/12$ ,  $\gamma = 0.01$  and  $\alpha_j$  chosen independently and uniformly from  $[0, 2\pi)$ .

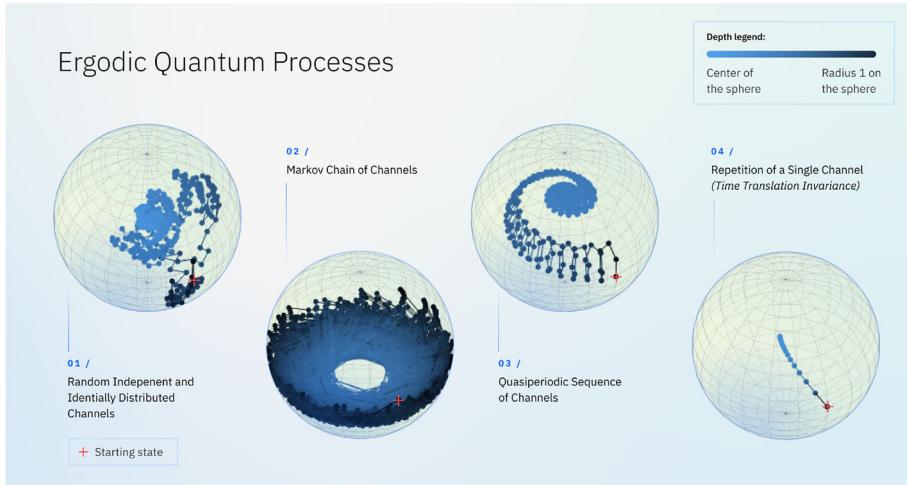


Fig. 1. Evolution of the state of a qubit under four ergodic quantum processes

- (2) A Markov chain of channels with  $\delta = \pi/12$ ,  $\gamma = 0.4$ ,  $\alpha_1 = 0$  and  $\alpha_{j+1} = \alpha_j - x_j$ , where  $x_j$  are independent and  $x_j = 1$  or  $0$  each with probability one half.
- (3) A quasi-periodic family of channels with  $\delta = \pi/12$ ,  $\gamma = 0.01$ ,  $\alpha_1 = 0$  and  $\alpha_{j+1} = \alpha_j + 1$ .
- (4) A periodic family of channels with  $\delta = \pi/12$ ,  $\gamma = 0.1$ ,  $\alpha_1 = 0$  and  $\alpha_{j+1} = \alpha_j + 2\pi/3$ . By sampling the sequence  $\rho_j$  only at  $j \equiv 0 \pmod{3}$ , we get a trajectory obtained by repeating the single channel  $\phi_3 \circ \phi_2 \circ \phi_1$ .

## 2. Ergodic Theory of Quantum Processes and Matrix Product States

2.1. *Notation.* Let  $\mathbb{M}_D = \mathbb{C}^{D \times D}$  denote the space of  $D \times D$  matrices. Recall the trace-norm of  $M \in \mathbb{M}_D$ ,  $\|M\|_1 = \text{tr}[|M|]$  as well as the Hilbert-Schmidt inner product and norm,  $\langle \tilde{M}, M \rangle = \text{tr}[\tilde{M}^\dagger M]$  and  $\|M\|_2^2 = \text{tr}[M^\dagger M]$ . Let  $\mathcal{L}(\mathbb{M}_D)$  denote the set of linear maps from  $\mathbb{M}_D$  to itself. Given  $\phi \in \mathcal{L}(\mathbb{M}_D)$ , we define the adjoint map  $\phi^*$  via the Hilbert-Schmidt inner product:  $\text{tr}[\tilde{M}^\dagger \phi(M)] = \text{tr}[[\phi^*(\tilde{M})]^\dagger M]$ .

Let  $\mathbb{P}_D$  denote the closed cone of positive semi-definite matrices in  $\mathbb{M}_D$ ,

$$\mathbb{P}_D = \left\{ M \in \mathbb{M}_D : z^\dagger M z \geq 0 \text{ for all } z \in \mathbb{C}^D \right\}.$$

The interior of  $\mathbb{P}_D$  is the open cone of positive definite matrices,

$$\mathbb{P}_D^\circ = \left\{ M \in \mathbb{M}_D : z^\dagger M z > 0 \text{ for all } z \in \mathbb{C}^D \text{ with } z \neq 0 \right\}.$$

A map  $\phi \in \mathcal{L}(\mathbb{M}_D)$  is *positive* if  $\phi(\mathbb{P}_D) \subset \mathbb{P}_D$ , i.e.,  $\phi$  maps positive semi-definite matrices to positive semi-definite matrices. The map is *strictly positive* if  $\phi(\mathbb{P}_D \setminus \{0\}) \subset \mathbb{P}_D^\circ$ , i.e.,  $\phi$  maps positive semi-definite matrices to positive definite matrices. A *completely positive* map is one such that  $\phi \otimes \mathbb{I}_r : \mathcal{L}(\mathbb{M}_D) \otimes \mathcal{L}(\mathbb{M}_r)$  is positive for every  $r$ , where  $\mathbb{I}_r$  denotes the identity map on  $\mathbb{M}_r$ ; let  $\mathcal{CP}(\mathbb{M}_D)$  denote the set of completely positive maps over  $\mathbb{M}_D$ ,

$$\mathcal{CP}(\mathbb{M}_D) = \{\phi \in \mathcal{L}(\mathbb{M}_D) : \phi \text{ is completely positive.}\}.$$

By Kraus's theorem [14, 25],  $\phi \in \mathcal{CP}(\mathbb{M}_D)$  if and only if  $\phi$  is of the form equation (1.1). A map  $\phi \in \mathcal{L}(\mathbb{M}_D)$  is *trace preserving* if  $\text{tr}[\phi(M)] = \text{tr}[M]$  for all  $M$ ; equivalently  $\phi^*(\mathbb{I}_D) = \mathbb{I}_D$ . A *quantum channel* is a completely positive trace preserving map.

Let  $(\Omega, \mathcal{F}, \text{Pr})$  be a probability space with

- (1)  $T : \Omega \rightarrow \Omega$  an invertible, ergodic, and measure preserving map, and
- (2)  $\phi_0 : \Omega \rightarrow \mathcal{CP}(\mathbb{M}_D)$  a completely positive map valued random variable (taking the Borel  $\sigma$ -algebra on  $\mathcal{CP}(\mathbb{M}_D)$ ).

Recall that  $T$  is *ergodic* provided  $\text{Pr}[A] = 0$  or  $1$  for any measurable set  $A$  with  $T^{-1}(A) = A$ . We follow the convention in probability theory and suppress the independent variable  $\omega \in \Omega$  in most formulas; when it is needed we will use a subscript to denote the value of a random variable at a particular  $\omega \in \Omega$ , e.g.,  $\phi_{0;\omega}$ . To specify  $\phi_0$  we could introduce matrix-valued random variables  $B_0^i : \Omega \rightarrow \mathbb{M}_D$ , for  $i = 1, \dots, d$ , and take

$$\phi_0(M) = \sum_{i=1}^d B_0^i M B_0^{i\dagger}. \quad (2.1)$$

If we further impose the condition

$$\sum_{i=1}^d B_0^{i\dagger} B_0^i = \mathbb{I}_D \quad \text{almost surely,} \quad (2.2)$$

then  $\phi_0$  is almost surely trace preserving, so  $\phi_0$  is almost surely equal to a quantum channel valued random variable. We note, however, that the matrices  $B_0^i$ ,  $i = 1, \dots, d$ , are not uniquely determined by the channel  $\phi_0$ . For this reason, we formulate our results directly in terms of the channel valued random variable  $\phi_0$  without reference to a specific Kraus representation.

**2.2. Ergodic theorems for quantum processes.** The main focus of this paper is to study the composition of a sequence of maps obtained by evaluating  $\phi_0$  along the trajectories of the ergodic map  $T$ :

$$\phi_{n;\omega} = \phi_{0;T^n\omega}, \quad (2.3)$$

with  $n \in \mathbb{Z}$ . For our general result, we do not require the maps to be quantum channels, i.e., trace preserving. Nonetheless, we take inspiration from the quantum channel case and consider the dynamics  $\rho_n = \phi_n(\rho_{n-1})$  induced by the sequence  $(\phi_n)_{n=0}^\infty$  on (non-normalized) states of a  $D$ -dimensional quantum system with Hilbert space  $\mathcal{H} = \mathbb{C}^D$ .

Consider the process

$$\Phi_N = \phi_N \circ \dots \circ \phi_0 \quad (2.4)$$

for  $N \geq 0$ . The only assumption we need is that

**Assumption 1.** With probability one there exists an  $N_0 > 0$  such that  $\Phi_N$  is strictly positive for all  $N \geq N_0$ .

Physically, this assumption states that no proper subspace of the system is invariant under the dynamics. For more discussion of the physical motivation behind Assumption 1, see [18].

Although Assumption 1 is physically natural, it is not formulated in a way that is easily verifiable. However, it is *equivalent* to two more easily verified assumptions:

**Lemma 2.1.** *Assumption 1 is equivalent to the following two statements taken together:*

- (1) *For some  $n_0 > 0$ ,  $\Pr[\Phi_{n_0} \text{ is strictly positive}] > 0$ .*
- (2) *With probability one,  $(\ker \phi_0) \cap \mathbb{P}_D = (\ker \phi_0^*) \cap \mathbb{P}_D = \{0\}$ .  
That is, if  $\phi_0(M) = 0$  or  $\phi_0^*(M) = 0$  with  $M \in \mathbb{P}_D$ , then  $M = 0$ .*

*Remarks.* (1) Conditions (1) and (2) are manifestly verifiable by a finite computation, while Assumption 1, as stated, is not. The proof of Lemma 2.1 is given below in Sect. 3.3. (2) A map  $\phi$  is strictly positive if and only if  $\phi^*$  is strictly positive.<sup>1</sup> Thus condition (1) is equivalent to  $\phi_0^* \circ \dots \circ \phi_{N_0}^*$  being strictly positive with positive probability. (3) If  $\phi_0$  is trace preserving, i.e., a quantum channel, then  $\text{tr}[\phi_0(M)] = \text{tr}[M]$  for any  $M$ , so  $\ker \phi_0 \cap \mathbb{P}_D = 0$ . However, the other half of condition (2) (that  $\ker \phi_0^* \cap \mathbb{P}_D = \{0\}$ ) does not necessarily hold. For example, if  $D$  is even and  $\phi(M) = PMP + SMS^\dagger$  with  $P$  a projection onto a subspace of dimension  $D/2$  and  $S$  a partial isometry from  $\mathbb{I}_D - P$  to  $P$ , then  $\phi$  is a channel but  $\phi^*(\mathbb{I}_D - P) = 0$ .

The classical Perron–Frobenius theorem [9, 22] has been generalized to linear maps preserving a convex cone, e.g., see [15]. Based on such a generalization, Evans and Høegh-Krohn [7] obtained results for positive maps on  $\mathbb{M}_D$ . It follows from [7, Theorem 2.3] that, if  $\Phi_N$  is strictly positive, then there is a unique (up to scaling) strictly positive matrix  $R_N \in \mathbb{M}_D$  such that

$$\Phi_N(R_N) = \lambda_N R_N, \quad (2.5)$$

where  $\lambda_N$  is the spectral radius of  $\Phi_N$ . Similarly, there is a unique (up to scaling) strictly positive matrix  $L_N$  such that

$$\Phi_N^*(L_N) = \lambda_N L_N. \quad (2.6)$$

We extend the process to  $-N < 0$  by defining

$$\Phi_{-N} = \phi_0 \circ \dots \circ \phi_{-N}. \quad (2.7)$$

By Assumption 1,  $\Phi_N$  is strictly positive for all sufficiently large  $N > 0$ . In Lemma 3.13 below, we show below that, with probability one, we also have  $\Phi_{-N}$  strictly positive for all  $N$  sufficiently large. Thus the left and right eigen-matrices  $R_N$  and  $L_N$  are unique for large  $|N|$ . We normalize  $R_N$  and  $L_N$  so that  $\text{tr}[R_N] = \text{tr}[L_N] = 1$ .

Our first result is that  $L_N$  converges as  $N \rightarrow \infty$ , while  $R_N$  converges as  $N \rightarrow -\infty$ .

**Theorem 1.** *There are random matrices  $Z_0, Z'_0 : \Omega \rightarrow \mathbb{M}_D$  such that  $Z_0, Z'_0 \in \mathbb{P}_D^\circ$ ,*

$$\lim_{N \rightarrow -\infty} R_N = Z_0, \quad \text{and} \quad \lim_{N \rightarrow \infty} L_N = Z'_0$$

*almost surely. Furthermore, if we set  $Z_n = Z_{0;T^n\omega}$  and  $Z'_n = Z'_{0;T^n\omega}$ , then*

$$Z_n = \phi_n \cdot Z_{n-1}, \quad \text{and} \quad Z'_n = \phi_n^* \cdot Z'_{n+1},$$

*where  $\cdot$  denotes the projective action of a positive map on the strictly positive  $D \times D$  matrices of trace 1:*

$$\phi_n \cdot M \equiv \frac{1}{\text{tr}[\phi_n(M)]} \phi_n(M).$$

<sup>1</sup> Indeed, if  $\phi$  is strictly positive and  $M \in \mathbb{P}_D$  is non-zero, then we have  $\text{tr}[\phi^*(M)M'] = \text{tr}[M\phi(M')] > 0$  for any non-zero  $M' \in \mathbb{P}_D$ , since  $\phi(M') > 0$ . Thus  $\phi^*(M)$  is strictly positive.

*Remarks.* (1) If the maps  $\phi_n$  are quantum channels, then  $L_N = \frac{1}{D}\mathbb{I}$ , so  $Z'_N = \frac{1}{D}\mathbb{I}$  for all  $n$ . (2) This result is closely related in spirit to Oseledec's Multiplicative Ergodic Theorem [20], a general result on convergence of singular vectors for products of linear transformations. (3) Theorem 1 generalizes a theorem of Hennion on the Perron–Frobenius eigenvectors of products of entry-wise positive matrices [12]. In fact, Hennion's theorem can be seen as a special case of our result applied to the following maps

$$\phi_0(M) = \sum_{\alpha, \beta} A_{0; \alpha, \beta} \mathbf{e}_\alpha \mathbf{e}_\beta^T M \mathbf{e}_\beta \mathbf{e}_\alpha^T, \quad (2.8)$$

with  $A_0$  a random matrix with non-negative entries, and  $\mathbf{e}_\alpha$ ,  $\alpha = 1, \dots, D$ , the standard basis vectors of  $\mathbb{C}^D$ . An equivalent, simpler, expression to equation (2.8) is given by  $\phi_0(M) = \text{diag}(A_0 \text{vec}(M))$ , where  $\text{vec } M$  is the  $D$ -dimensional vector consisting of the diagonal entries of  $M$  and  $\text{diag}(\mathbf{v})$  is a diagonal matrix with the entries of the vector  $\mathbf{v}$  on the diagonal.

Given  $m < n$  in  $\mathbb{Z}$ , let  $P_{n,m}$  denote the rank-one operator

$$P_{n,m}(M) = \text{tr}[Z'_m M] Z_n. \quad (2.9)$$

Our second result states that, for  $n - m$  large, the operator  $\phi_n \circ \dots \circ \phi_m$  is well approximated by  $P_{n,m}$ . To formulate this result precisely, we use the operator norm for a map  $\Phi \in \mathcal{L}(\mathbb{M}_D)$  inherited from the trace norm on  $\mathbb{M}_D$ ,  $\|\Phi\|_1 = \max \{ \text{tr}[|\Phi(M)|] : \text{tr}[|M|] = 1 \}$ .

**Theorem 2.** *Given  $m < n$  in  $\mathbb{Z}$ , let  $\Psi_{n,m} = \phi_n \circ \dots \circ \phi_m$ . There is  $0 < \mu < 1$  so that for each  $x \in \mathbb{Z}$  the following bound holds:*

$$\left\| \frac{1}{\text{tr}[\Psi_{n,m}^*(\mathbb{I})]} \Psi_{n,m} - P_{n,m} \right\|_1 \leq C_{\mu, x} \mu^{n-m} \quad (2.10)$$

for all  $m \leq x$  and  $n \geq x$ , with  $C_{\mu, x}$  finite almost surely.

**2.3. Ergodic matrix product states.** The Kraus matrices associated to a random completely positive map specified as in Eq. (1.1) can be used to define a family of random matrix product states as follows. Let

$$A_n^i = B_{0; T^n \omega}^{i \dagger}, \quad n \in \mathbb{Z} \text{ and } i = 1, \dots, d.$$

Given an interval  $[m, n]$  of  $\mathbb{Z}$ , we define the matrix product state

$$|\psi([m, n])\rangle = \frac{1}{\mathcal{N}} \sum_{i_m, \dots, i_n=1}^d \text{tr}[A_m^{i_m} \dots A_n^{i_n}] |i_m, \dots, i_n\rangle, \quad (2.11)$$

where  $|i_m, \dots, i_n\rangle$  are the elements of the computational basis on  $\bigotimes_{k=m}^n \mathbb{C}^d$ , where  $d = 2$  corresponds to qubits, and the normalization constant is given by

$$\mathcal{N}^2 = \sum_{i_m, \dots, i_n=1}^d \left| \text{tr}[A_m^{i_m} \dots A_n^{i_n}] \right|^2.$$

For simplicity, we restrict our attention here to the periodic boundary condition states, as defined in Eq. (2.11).

There is a close relation between matrix product states and completely positive maps, via which Theorem 2 can be used to characterize the thermodynamic limit ( $m \rightarrow -\infty$  and  $n \rightarrow \infty$ ) of the states defined in Eq. (2.11). A preliminary observation is that the normalization factor  $\mathcal{N}$  can be expressed as

$$\mathcal{N}^2 = \sum_{\alpha, \beta=1}^D \text{tr} \left[ \mathbf{e}_\beta \mathbf{e}_\alpha^T \phi_n \circ \cdots \circ \phi_m \left( \mathbf{e}_\alpha \mathbf{e}_\beta^T \right) \right]. \quad (2.12)$$

Since  $\left\{ \mathbf{e}_\alpha \mathbf{e}_\beta^T : \alpha, \beta = 1, \dots, D \right\}$  is an orthonormal basis for  $\mathbb{M}_D$ , Eq. (2.12) can, in turn, be written as

$$\mathcal{N}^2 = \text{Tr}[\phi_n \circ \cdots \circ \phi_m],$$

where  $\text{Tr}[\phi]$  denotes the trace of a linear map  $\phi \in \mathcal{L}(\mathbb{M}_D)$ . Throughout this discussion, we use  $\text{tr}$  to denote the trace on  $\mathbb{M}_D$  and  $\text{Tr}$  to denote the trace on  $\mathcal{L}(\mathbb{M}_D)$ .

Let us now consider the state  $|\psi(N)\rangle \equiv |\psi([-N, N])\rangle$  defined on  $[-N, N]$  by Eq. (2.11). Given  $-N < m < n < N$  and a local observable  $O$  on the spins in  $[m, n]$ , let

$$\widehat{O}(M) = \sum_{\substack{i_m, \dots, i_n \\ j_m, \dots, j_n}} \langle i_m, \dots, i_n | O | j_m, \dots, j_n \rangle A_n^{i_n \dagger} \cdots A_m^{i_m \dagger} M A_m^{i_m} \cdots A_n^{i_n} \quad (2.13)$$

which is a linear operator on  $\mathbb{M}_D$ . One may easily verify that the (quantum) expectation of  $O$  in  $|\psi(N)\rangle$  is

$$\langle \psi(N) | O | \psi(N) \rangle = \frac{\text{Tr} [\phi_N \circ \cdots \circ \phi_{n+1} \circ \widehat{O} \circ \phi_{m-1} \circ \cdots \circ \phi_{-N}]}{\text{Tr} [\phi_N \circ \cdots \circ \phi_{-N}]} \quad (2.14)$$

We can express the thermodynamic limit of  $\langle \psi(N) | O | \psi(N) \rangle$  in terms of the matrices  $Z_m$  and  $Z'_m$  from Theorem 1 by using Eq. (2.14) and Theorem 2. It is convenient to use Dirac notation for the Hilbert-Schmidt inner product on  $\mathbb{M}_D$ , with which we have  $P_{n,m} = |Z_n\rangle \langle Z'_m|$  (with  $P_{n,m}$  as in Theorem 2). Let

$$W(O) = \lim_{N \rightarrow \infty} \langle \psi(N) | O | \psi(N) \rangle \quad (2.15)$$

denote the thermodynamic limit of  $|\psi(N)\rangle$ , where  $O$  is any local observable. Using Eq. (2.14) and Theorem 2, we compute  $W(O)$  as follows:

$$\begin{aligned} W(O) &= \lim_{N \rightarrow \infty} \frac{\text{Tr} |Z_N\rangle \langle Z'_{n+1}| \widehat{O} |Z_{m-1}\rangle \langle Z'_{-N}|}{\text{Tr} |Z_N\rangle \langle Z'_{n+1}| \phi_n \circ \cdots \circ \phi_m |Z_{m-1}\rangle \langle Z'_{-N}|} \\ &= \frac{\langle Z'_{n+1} | \widehat{O} |Z_{m-1}\rangle}{\langle Z'_{n+1} | \phi_n \circ \cdots \circ \phi_m |Z_{m-1}\rangle}, \end{aligned} \quad (2.16)$$

whenever  $O$  is a local observable on the spins in  $[m, n]$ . Since  $Z_m = \phi_m \cdot Z_{m-1}$ , the normalization in the denominator is given by

$$\langle Z'_{n+1} | \phi_n \circ \cdots \circ \phi_m |Z_{m-1}\rangle = \left[ \prod_{k=m}^n \text{tr}[\phi_k(Z_{k-1})] \right] \langle Z'_{n+1} | Z_n \rangle.$$

As is well known, there is a “gauge-freedom” in the representation of a matrix product state such as  $|\psi(N)\rangle$ : the state itself does not change under the replacement  $A_k^i \mapsto V_{k-1} A_k^i V_k^{-1}$  provided we have  $V_N = V_{-N}$ . See [21] for more discussion on this symmetry. Choosing the matrices  $V_k$  appropriately, one can impose a gauge fixing condition such as  $\sum_{i=1}^d A_i A_i^\dagger = \mathbb{I}$ , which would make the associated channel maps trace preserving. *A priori*, it appears that the matrices required for gauge fixing may depend on  $N$  and it is not clear that they can be chosen consistently with the ergodic structure. However, Theorem 1 allows us to do just that, as we now explain. To begin, let

$$\xi_m = \text{tr}[\phi_m^*(Z'_{m+1})].$$

By Theorem 1,  $(\xi_m)_{m \in \mathbb{Z}}$  is a shift-covariant sequence ( $\xi_{m;\omega} = \xi_{m-1;T\omega}$ ) of positive random variables. Furthermore  $\phi_m^*(Z'_{m+1}) = \xi_m Z'_m$ . Now let

$$\tilde{\phi}_m(M) = \frac{1}{\xi_m} Z'^{1/2}_{m+1} \phi_m \left( Z'^{-1/2}_m M Z'^{-1/2}_m \right) Z'^{1/2}_{m+1}; \quad (2.17)$$

this expression is well defined since the matrices  $Z'_m$  are full rank almost surely by Theorem 1. The maps  $\tilde{\phi}_m$  are an ergodic sequence of completely positive maps, and a short computation shows that they are trace preserving:

$$\begin{aligned} \text{tr}[\tilde{\phi}_m(M)] &= \frac{1}{\xi_m} \text{tr} \left[ Z'_{m+1} \phi_m \left( Z'^{-1/2}_m M Z'^{-1/2}_m \right) \right] = \frac{1}{\xi_m} \text{tr} \left[ \phi_m^*(Z'_{m+1}) Z'^{-1/2}_m M Z'^{-1/2}_m \right] \\ &= \text{tr} \left[ Z'_m Z'^{-1/2}_m M Z'^{-1/2}_m \right] = \text{tr}[M]. \end{aligned}$$

Also,

$$\tilde{\phi}_m(M) = \sum_{i=1}^d \tilde{A}_m^{i\dagger} M \tilde{A}_m^i$$

where  $\tilde{A}_m^i = \frac{1}{\sqrt{\xi_m}} Z'^{1/2}_{m+1} A_m^i Z'^{-1/2}_m$ .

We could apply Theorem 1 directly to the sequence  $\tilde{\phi}_m$ , since it is straightforward to see that these maps satisfy conditions (1) and (2) of Lemma 2.1. However, it is easier to simply write down the left and right matrices  $\tilde{Z}'_m$  and  $\tilde{Z}_m$  directly using Eq. (2.17):

$$\tilde{Z}'_m = \frac{1}{D} \mathbb{I} \quad \text{and} \quad \tilde{Z}_m = \frac{1}{\text{tr}[Z'_m Z_m]} Z'^{1/2}_{m+1} Z_m Z'^{1/2}_{m-1}.$$

Since  $\tilde{\phi}_m$  is trace preserving we have  $\tilde{\phi}_m(\tilde{Z}_{m-1}) = \tilde{\phi}_m \cdot \tilde{Z}_{m-1} = \tilde{Z}_m$  for all  $m$  and  $\tilde{\phi}_m^*(\mathbb{I}) = \mathbb{I}$ .

We now return to the expression for the thermodynamic limit  $W(O)$ . Given an interval  $[m, n]$ , one may easily check that

$$\begin{aligned} \langle Z'_{n+1} | \phi_n \circ \cdots \circ \phi_m | Z_{m-1} \rangle \\ = (\xi_n \cdots \xi_m) \text{tr}[Z'_m Z_{m-1}] \text{tr}[\tilde{\phi}_n \circ \cdots \circ \tilde{\phi}_m(\tilde{Z}_{m-1})] = (\xi_n \cdots \xi_m) \text{tr}[Z'_m Z_{m-1}]. \end{aligned}$$

For a local observable  $O$  on the spins in  $[m, n]$ , we define analogous to Eq. (2.13),

$$\tilde{O}(M) \equiv \sum_{\substack{i_m, \dots, i_n \\ j_m, \dots, j_n}} \langle i_m, \dots, i_n | O | j_m, \dots, j_n \rangle \tilde{A}_n^{i_n\dagger} \cdots \tilde{A}_m^{i_m\dagger} M \tilde{A}_m^{i_m} \cdots \tilde{A}_n^{i_n}. \quad (2.18)$$

Inserting these definitions into Eq. (2.16), we find the following remarkably simple formula for the thermodynamic limit  $W$  of the matrix product states:

$$W(O) = \text{tr} [\tilde{O}(\tilde{Z}_{m-1})]. \quad (2.19)$$

Equation (2.19) can be used to obtain a bound on the two-point correlation of two observables  $O_1$  and  $O_2$  located in disjoint intervals  $I_1 = [m_1, n_1]$  and  $I_2 = [m_2, n_2]$  with  $n_1 < m_2$ . For such observables

$$W(O_2 O_1) = \text{tr} [\tilde{O}_2 \circ \tilde{\phi}_{m_2-1} \circ \cdots \circ \tilde{\phi}_{n_1+1} \circ \tilde{O}_1(\tilde{Z}_{m_1-1})].$$

Applying Theorem 2 to  $\tilde{\phi}_m$  allows us to obtain the following

**Theorem 3.** *There is  $0 < \mu < 1$  such that for each  $x \in \mathbb{Z}$  the following correlation inequality holds with  $C_{\mu,x} < \infty$  almost surely:*

$$|W(O_2 O_1) - W(O_2)W(O_1)| \leq C_{\mu,x} \|\tilde{O}_1 - W(O_1)\tilde{\Psi}_1\| \|\tilde{O}_2 - W(O_2)\tilde{\Psi}_2\| \mu^{m_2-n_1}, \quad (2.20)$$

whenever  $\text{supp}[O_j] \in [m_j, n_j]$  and  $\tilde{\Psi}_j = \tilde{\phi}_{n_j} \circ \cdots \circ \tilde{\phi}_{m_j}$  for  $j = 1, 2$  with  $n_1 < x < m_2$ .

**2.4. Overview of the proofs.** Theorems 1, 2, and 3 are proved in Sect. 4 below, using several technical lemmas presented and proved in Sect. 3. The central idea of the proofs is contraction mapping argument for the maps  $\Phi_N$  on  $\mathbb{P}_D$ . This is accomplished in Lemma 3.10, facilitated by the introduction of a non-standard metric on the set of quantum states—see Eq. (3.3) below. The metric and a number of the ideas developed in Sect. 3 are inspired by results in section 10 of Hennion’s paper [12]. Although some of the statements are similar, the proofs in [12] do not directly carry over to the present more general context. Nonetheless, for readers interested in comparing the two papers, we note the following correspondence between lemmas in the present paper and in [12]:

$$\text{Lemma } \left\{ \begin{array}{l} 3.3 \\ 3.5 \\ 3.6 \\ 3.8 \\ 3.9 \\ 3.10 \end{array} \right\} \text{ generalizes Lemma } \left\{ \begin{array}{l} 10.1 \\ 10.2 \\ 10.3 \\ 10.4 \\ 10.5 \\ 10.6 \end{array} \right\} \text{ of [12].}$$

Although we have taken inspiration from [12], we do not use any of the results therein directly and the present paper can be read on its own.

### 3. Technical Results

#### 3.1. Notation.

Let

$$\mathbb{S}_D = \{M \in \mathbb{P}_D : \text{tr}[M] = 1\}, \quad \mathbb{S}_D^\circ = \{M \in \mathbb{P}_D^\circ : \text{tr}[M] = 1\},$$

and

$$\mathcal{P}_D = \{\phi \in \mathcal{L}(\mathbb{M}_D) : \phi \text{ is a positive map, } \ker \phi \cap \mathbb{P}_D = \{0\}, \text{ and } \ker \phi^* \cap \mathbb{P}_D = \{0\}\}.$$

Note that  $\mathcal{P}_D$  is a convex set. Let  $\mathcal{P}_D^\circ$  denote its interior. Since any strictly positive map satisfies the kernel condition in the definition of  $\mathcal{P}_D$ , we have

$$\mathcal{P}_D^\circ = \{\phi \in \mathcal{L}(\mathbb{M}_D) : \phi \text{ is strictly positive.}\}.$$

Assumption 1 ensures that  $\Phi_N = \phi_N \circ \dots \circ \phi_0 \in \mathcal{P}_D^\circ$  for large  $N$ , with probability one. Condition (2) of Lemma 2.1 states that  $\phi_0 \in \mathcal{P}_D$  almost surely, while condition (1) states that  $\Phi_{n_0} \in \mathcal{P}_D^\circ$  with positive probability for some positive integer  $n_0$ . Note that any  $\phi \in \mathcal{P}_D$  maps  $\mathbb{P}_D$  into  $\mathbb{P}_D$ , while  $\phi \in \mathcal{P}_D^\circ$  maps  $\mathbb{P}_D$  into  $\mathbb{P}_D^\circ$ .

Theorem 1 is formulated in terms of the projective action

$$\phi \cdot M \equiv \frac{\phi(M)}{\text{tr}[\phi(M)]},$$

of a positive map on  $\mathbb{S}_D$ . Note that  $\text{tr}[\phi(M)] \neq 0$  for  $\phi \in \mathcal{P}_D$  and  $M \in \mathbb{S}_D$ , so this action is well defined.

**Lemma 3.1.** *Let  $\phi \in \mathcal{P}_D$  then  $\phi$  maps  $\mathbb{P}_D^\circ$  into  $\mathbb{P}_D^\circ$ .*

*Proof.* We first show that  $\phi(\mathbb{I}) \in \mathbb{P}_D^\circ$ . Suppose on the contrary that  $\phi(\mathbb{I}) \in \mathbb{P}_D \setminus \mathbb{P}_D^\circ$ . Let  $P$  denote the orthogonal projection onto the kernel of  $\phi(\mathbb{I})$ . Then  $0 = \text{tr}[P\phi(\mathbb{I})] = \text{tr}[\phi^*(P)]$ , so  $\phi^*(P) = 0$ , contradicting the definition of  $\mathcal{P}_D$ . Thus  $\phi(\mathbb{I}) \in \mathbb{P}_D^\circ$ . Now let  $M$  be any point of  $\mathbb{P}_D^\circ$  and let  $\delta > 0$  such that  $M \geq \delta\mathbb{I}$ . Then  $\phi(M) \geq \delta\phi(\mathbb{I})$ , so  $\phi(M) \in \mathbb{P}_D^\circ$ .  $\square$

The sets  $\mathcal{P}_D$ ,  $\mathcal{P}_D^\circ$  are semi-groups under composition; it follows from Lemma 3.1 that  $\mathcal{P}_D^\circ$  is a two-sided ideal of  $\mathcal{P}_D$ :

**Corollary 3.2.** *Given  $\phi \in \mathcal{P}_D$  and  $\phi' \in \mathcal{P}_D^\circ$ , we have  $\phi \circ \phi' \in \mathcal{P}_D^\circ$  and  $\phi' \circ \phi \in \mathcal{P}_D^\circ$ .*

*Proof.* We have  $\phi' \circ \phi$  in  $\mathcal{P}_D^\circ$ , since  $\phi' \circ \phi(\mathbb{P}_D) \subset \phi'(\mathbb{P}_D) \subset \mathbb{P}_D^\circ$ . On the other hand, for any  $M \in \mathbb{P}_D$  we have  $\phi'(M) \in \mathbb{P}_D^\circ$  and thus  $\phi \circ \phi'(M) \in \mathbb{P}_D^\circ$ , by Lemma 3.1.  $\square$

**3.2. Geometry of  $\mathbb{S}_D$ .** The set  $\mathbb{S}_D$  of density matrices is convex and compact. To implement the contraction argument at the heart of the proof of Theorem 1, it is useful to introduce a special metric on this space based on the following quantity:

$$m(X, Y) = \sup \{\lambda : \lambda Y \leq X\}, \quad (3.1)$$

for  $X, Y \in \mathbb{S}_D$ .

**Lemma 3.3.** *Let  $X, Y \in \mathbb{S}_D$ . Then*

$$\begin{aligned} m(X, Y) &= \min \left\{ \frac{\text{tr}[AX]}{\text{tr}[AY]} : A \in \mathbb{S}_D \text{ and } \text{tr}[AY] \neq 0 \right\} \\ &= \inf \left\{ \frac{\text{tr}[AX]}{\text{tr}[AY]} : A \in \mathbb{S}_D^\circ \right\}. \end{aligned} \quad (3.2)$$

Furthermore, if  $Z \in \mathbb{S}_D$ , Then

- (1)  $0 \leq m(X, Y) \leq 1$
- (2)  $m(X, Z)m(Z, Y) \leq m(X, Y)$
- (3)  $m(X, Y)m(Y, X) = 1$  if and only if  $X = Y$

(4)  $m(X, Y) = 0$  if and only if  $Y\vec{v} \neq 0$  for some  $\vec{v} \in \ker X$ . In particular,  $m(X, Y) > 0$  if  $X \in \mathbb{S}_D^\circ$ .

*Proof.* If  $\lambda Y \leq X$ , then  $\lambda \operatorname{tr}[AY] \leq \operatorname{tr}[AX]$  and  $\lambda \leq \frac{\operatorname{tr}[AX]}{\operatorname{tr}[AY]}$  if  $A \in \mathbb{S}_D$  and  $\operatorname{tr}[AY] \neq 0$ . Thus

$$m(X, Y) \leq \inf \left\{ \frac{\operatorname{tr}[AX]}{\operatorname{tr}[AY]} : A \in \mathbb{S}_D \text{ and } \operatorname{tr}[AY] \neq 0 \right\}.$$

To see that the infimum is attained and is equal to  $m(X, Y)$ , note that if  $\lambda = m(X, Y)$  then 0 must be an eigenvalue of  $X - \lambda Y$  with an eigenvector  $\vec{u}$  such that  $Y\vec{u} \neq 0$  (else we could increase  $\lambda$  by a small amount without violating  $\lambda Y \leq X$ ). Let  $A = \vec{u} \langle \vec{u}, \cdot \rangle$ . Then  $\operatorname{tr}[AX] = m(X, Y) \operatorname{tr}[AY]$  and  $\operatorname{tr}[AY] \neq 0$ .

To see that we still obtain  $m(X, Y)$  if we restrict the infimum to range over  $A \in \mathbb{S}_D^\circ$ , note that if  $\lambda Y \not\leq X$ , then we must have  $\langle \vec{u}, X\vec{u} \rangle < \lambda \langle \vec{u}, Y\vec{u} \rangle < 0$  for some  $\vec{u}$ . Since  $X$  is positive, it follows that  $\langle \vec{u}, Y\vec{u} \rangle \neq 0$  and  $\lambda > \frac{\langle \vec{u}, X\vec{u} \rangle}{\langle \vec{u}, Y\vec{u} \rangle}$ . Taking  $M = \vec{u} \langle \vec{u}, \cdot \rangle + \delta \mathbb{I}$  for small enough  $\delta$  we see that  $\lambda > \inf \left\{ \frac{\operatorname{tr}[AX]}{\operatorname{tr}[AY]} : A \in \mathbb{S}_D^\circ \right\}$ . Thus

$$m(X, Y) = \inf \left\{ \frac{\operatorname{tr}[AX]}{\operatorname{tr}[AY]} : A \in \mathbb{S}_D^\circ \right\}.$$

The lower bound in part 1 is clear. To see the upper bound note that  $\frac{\operatorname{tr}[\mathbb{I}X]}{\operatorname{tr}[\mathbb{I}Y]} = 1$ . For part 2, note that if  $\lambda Z \leq X$  and  $\mu Y \leq Z$ , then  $\lambda\mu Y \leq X$ . For part 3, note that if  $m(X, Y)m(Y, X) = 1$  then  $m(X, Y) = m(Y, X) = 1$  so  $X \leq Y$  and  $Y \leq X$ . Finally, for part 4, note that if  $Y\vec{v} \neq 0$  and  $X\vec{v} = 0$  then  $\lambda \langle \vec{v}, Y\vec{v} \rangle > 0 = \langle \vec{v}, X\vec{v} \rangle$  for any  $\lambda > 0$ . Conversely, if  $Y\vec{v} = 0$  for any  $\vec{v} \in \ker X$ , then  $Y$  is reduced by the subspace decomposition  $\ker X \oplus \operatorname{ran} X$ , and with respect to this decomposition

$$X = \begin{pmatrix} 0 & 0 \\ 0 & X' \end{pmatrix} \quad \text{and} \quad Y = \begin{pmatrix} 0 & 0 \\ 0 & Y' \end{pmatrix},$$

where  $X', Y'$  are operators on  $\operatorname{ran} X$ . Furthermore  $\ker X' = \{0\}$ , so  $X' \geq \delta \mathbb{I}$  for some  $\delta > 0$ . It follows that  $\lambda Y' \leq X'$  for small  $\lambda > 0$ . Then  $\lambda Y \leq X$ , so  $m(X, Y) > 0$ .  $\square$

**Corollary 3.4.**  $d_0(X, Y) := -\log m(X, Y) - \log m(Y, X)$  is a metric on  $\mathbb{S}_D$ .

The metric  $d_0(X, Y)$  is slightly unpleasant; it is unbounded and takes the value  $\infty$ . A much nicer metric is given by

$$d(X, Y) = \frac{1 - m(X, Y)m(Y, X)}{1 + m(X, Y)m(Y, X)}. \quad (3.3)$$

**Lemma 3.5.**  $d$  is a metric on  $\mathbb{S}_D$  such that

- (1)  $\sup \{d(X, Y) : X, Y \in \mathbb{S}_D\} = 1$ , and
- (2) if  $X \in \mathbb{S}_D^\circ$  and  $Y \in \mathbb{S}_D$ , then  $d(X, Y) = 1$  if and only if  $Y \in \mathbb{S}_D \setminus \mathbb{S}_D^\circ$ .

*Proof.* Symmetry of  $d$  is clear. Furthermore,  $0 \leq d(X, Y) \leq 1$  and  $d(X, Y) = 0$  if and only if  $m(X, Y)m(Y, X) = 1$ , which holds if and only if  $X = Y$  by Lemma 3.3. To prove the triangle inequality, let

$$f(s) = \frac{1-s}{1+s} = -1 + \frac{2}{1+s}$$

for  $0 \leq s \leq 1$ . Then  $f$  is decreasing and

$$f(s) + f(t) = \frac{2 - 2st}{1 + s + t + st} = 2 \frac{1 + st}{1 + s + t + st} \frac{1 - st}{1 + st} = 2 \frac{1}{1 + \frac{s+t}{1+st}} \frac{1 - st}{1 + st}.$$

The maximum of  $\frac{s+t}{1+st}$  over  $s, t \in [0, 1]$  is 1, from which it follows that  $f(s) + f(t) \geq f(st)$ . The triangle inequality for  $d$  follows from this inequality and part 2 of Lemma 3.3.

To prove that the diameter of  $\mathbb{S}_D$  is 1 as claimed, we simply need to find  $X, Y \in \mathbb{S}_D$  with  $m(X, Y) = 0$ . This holds, for instance, if  $X \in \mathbb{S}_D^\circ$  and  $\ker Y \neq \{0\}$ , which also leads to the result noted in item 2.  $\square$

**Lemma 3.6.** *Let  $X, Y \in \mathbb{S}_D$  with  $X \neq Y$ . Then*

$$d(X, Y) = \frac{|u_1 v_2 - u_2 v_1|}{u_1 v_2 + u_2 v_1}, \quad (3.4)$$

where  $X = u_1 A_- + u_2 A_+$  and  $Y = v_1 A_- + v_2 A_+$  with  $A_\pm$  the endpoints of the intersection of  $\mathbb{S}_D$  with the line through  $X$  and  $Y$ .

*Remark 3.7.* Since  $X$  and  $Y$  lie on the segment connecting  $A_\pm$  we have  $u_1 + u_2 = 1$  and  $v_1 + v_2 = 1$ . Thus we have

$$d(X, Y) = \frac{|u_1 - v_1|}{u_1 + v_1 - 2u_1 v_1} = \frac{|u_2 - v_2|}{u_2 + v_2 - 2u_2 v_2}.$$

*Proof.* Let  $t_+$  and  $t_-$  be the largest and smallest real numbers such that  $tX + (1-t)Y \in \mathbb{S}_D$ . Note that  $t_- \leq 0 \leq 1 \leq t_+$  and  $A_\pm = t_\pm X + (1 - t_\pm)Y$ . Furthermore

$$u_1 = \frac{t_+ - 1}{t_+ - t_-}, \quad u_2 = \frac{1 - t_-}{t_+ - t_-}, \quad v_1 = \frac{t_+}{t_+ - t_-}, \quad \text{and} \quad v_2 = \frac{-t_-}{t_+ - t_-},$$

so Eq. (3.4) is equivalent to

$$d(X, Y) = \frac{t_+ - t_-}{t_- + t_+ - 2t_- t_+}. \quad (3.5)$$

Note that each  $A_\pm$  must have a non-trivial kernel. For example, if  $A_-$  were positive-definite, then  $A_- - \delta(X - Y)$  would be positive definite for small  $\delta$ , contradicting the minimality of  $t_-$ . A similar argument applies to  $A_+$ . Furthermore we must have  $\ker A_- \not\subset \ker A_+$  and  $\ker A_+ \not\subset \ker A_-$ . Indeed, suppose that  $\ker A_+ \subset \ker A_-$ . Then we would have  $A_+ - \delta A_- \geq 0$  for small  $\delta$ , contradicting the maximality of  $t_+$ . The proof that  $\ker A_- \not\subset \ker A_+$  is similar.

Suppose that  $t_+ = 1$ . Then  $X = A_+$  and  $tX + (1-t)Y$  is not positive definite whenever  $t > 1$ , i.e.,  $X - \lambda Y$  is not positive definite for any  $\lambda > 0$ . It follows that  $m(X, Y) = 0$  and thus  $d(X, Y) = 1$ , so Eq. (3.5) holds. Similarly, if  $t_- = 0$  then  $Y = A_-$ ,  $d(X, Y) = 1$ , and Eq. (3.5) holds.

Now suppose that  $t_+ > 1$  and  $t_- < 0$ . Then  $X$  and  $Y$  are in the interior of the interval connecting  $A_-$  and  $A_+$ . Let  $r = \min \{u_i/v_i : i = 1, 2\}$ . Then

$$rY = r(v_1 A_- + v_2 A_+) \leq u_1 A_- + u_2 A_+ = X.$$

Thus  $r \leq m(X, Y)$ . On the other hand

$$m(X, Y)(v_1 A_- + v_2 A_+) = m(X, Y)Y \leq X = u_1 A_- + u_2 A_+.$$

Let  $\vec{w}_+ \in \ker A_+ \setminus \ker A_-$ . Then  $m(X, Y)v_1 \langle \vec{w}_+, A_- \vec{w}_+ \rangle \leq u_1 \langle \vec{w}_+, A_- \vec{w}_+ \rangle$ . Thus  $m(X, Y) \leq u_1/v_1$ . Similarly, working with  $w_- \in \ker A_- \setminus A_+$  we find that  $m(X, Y) \leq u_2/v_2$  so that  $m(X, Y) \leq r$ . Thus  $m(X, Y) = r$ . Similarly,  $m(Y, X) = \min \{v_i/u_i : i = 1, 2\}$ . Thus,

$$m(X, Y)m(Y, X) = \min \left\{ \frac{u_1}{v_1} \frac{v_2}{u_2}, \frac{u_2}{v_2} \frac{v_1}{u_1} \right\},$$

from which Eq. (3.4) follows.  $\square$

**Lemma 3.8.** *Let  $X, Y \in \mathbb{S}_D$ , then  $d(X, Y) \geq \frac{1}{2} \operatorname{tr} |X - Y|$ .*

*Proof.* Based on the remark following Lemma 3.6, we have

$$d(X, Y) = \frac{|u_1 - v_1|}{u_1 + v_1 - 2u_1 v_1} \geq |u_1 - v_1|,$$

where  $X = u_1 A_- + u_2 A_+$  and  $Y = v_1 A_- + v_2 A_+$  with  $A_{\pm}$  as in Lemma 3.6. Since  $u_2 = 1 - u_1$  and  $v_2 = 1 - v_1$ , we have  $X - Y = (u_1 - v_1)(A_- - A_+)$ . Thus

$$-|u_1 - v_1|(A_- + A_+) \leq X - Y \leq |u_1 - v_1|(A_- + A_+),$$

so  $\operatorname{tr} |X - Y| \leq 2|u_1 - v_1|$ .  $\square$

**Lemma 3.9.** *Let  $d_1(X, Y) = \operatorname{tr} |X - Y|$  denote the trace norm metric on  $\mathbb{S}_D$ . Let  $Y \in \mathbb{S}_D^{\circ}$ ,  $X \in \mathbb{S}_D$  and let  $X_n$  be a sequence in  $\mathbb{S}_D$  such that  $\lim_n d_1(X_n, X) = 0$ . Then  $\lim_n d(X_n, Y) = d(X, Y)$ . In particular, the spaces  $(\mathbb{S}_D^{\circ}, d)$  and  $(\mathbb{S}_D^{\circ}, d_1)$  are homeomorphic.*

*Remark.* The spaces  $(\mathbb{S}_D, d)$  and  $(\mathbb{S}_D, d_1)$  are not homeomorphic, and look very different on the boundary  $\mathbb{S}_D \setminus \mathbb{S}_D^{\circ}$ . For instance, if  $P \in \mathbb{S}_D$  is an orthogonal projection onto a proper subspace, then  $Y_t = (1-t)P + t$  converges to  $P$  in  $d_1$  as  $t \rightarrow 0$ , but  $d(P, Y_t) = 1$  for all  $t > 0$  (since  $m(P, Y_t) = 0$ ). The space  $(\mathbb{S}_D, d_1)$  is compact, but  $(\mathbb{S}_D, d)$  has an uncountable number of components.

*Proof.* We will show that  $m(Y, X) = \lim_n m(Y, X_n)$  and  $m(X, Y) = \lim_n m(X_n, Y)$ . Since  $Y \in \mathbb{S}_D^{\circ}$ , we have  $Y > \delta \mathbb{I}$  for some  $\delta > 0$ . Given  $\epsilon > 0$  we have  $\operatorname{tr} |X_n - X| < \epsilon$  and thus  $X_n \leq X + \epsilon \mathbb{I}$  and  $X \leq X_n + \epsilon \mathbb{I}$  for large enough  $n$ .

We first show that  $m(X, Y) = \lim_n m(X_n, Y)$ . Let  $\epsilon > 0$ . Given  $\lambda \leq m(X, Y)$ , so  $\lambda X \leq Y$ , we have  $\lambda Y \leq X \leq X_n + \epsilon \mathbb{I}$ . Thus  $(\lambda - \frac{\epsilon}{\delta})Y \leq X_n$  for large enough  $n$ . It follows that  $\liminf_n m(X_n, Y) \geq m(X, Y) - \frac{\epsilon}{\delta}$ . On the other hand if  $\lambda \leq \limsup_n m(X_n, Y)$ , then we have  $\lambda Y \leq X_n \leq X + \epsilon \mathbb{I}$  along a subsequence  $n_j \rightarrow \infty$ . Thus  $(\lambda - \frac{\epsilon}{\delta})Y \leq X$ , and so  $\limsup_n m(X_n, Y) \leq m(X, Y) + \frac{\epsilon}{\delta}$ . We have shown

$$m(X, Y) - \frac{\epsilon}{\delta} \leq \liminf_n m(X_n, Y) \leq \limsup_n m(X_n, Y) \leq m(X, Y) + \frac{\epsilon}{\delta}.$$

Taking  $\epsilon \rightarrow 0$ , we see that  $\lim_n m(X_n, Y) = m(X, Y)$  as claimed.

Now we show that  $m(Y, X) = \lim_n m(Y, X_n)$ . Let  $0 < t < 1$  and choose  $\epsilon$  small enough that  $t\epsilon \leq (1-t)\delta$ . Given  $\lambda \leq m(Y, X)$ , we have

$$t\lambda X_n \leq t\lambda(X + \epsilon \mathbb{I}) \leq tY + (1-t)\delta \mathbb{I} \leq Y$$

for large  $n$ . Thus  $\liminf_n m(Y, X_n) \geq tm(X, Y)$ . Similarly, given  $\lambda \leq \limsup_n m(Y, X_n)$ , we have  $t\lambda X \leq t\lambda(X_{n_j} + \epsilon\mathbb{I}) \leq Y$  along a sub-sequence  $n_j \rightarrow \infty$ . Thus

$$tm(Y, X) \leq \liminf_{n \rightarrow \infty} m(Y, X_n) \leq \limsup_{n \rightarrow \infty} m(Y, X_n) \leq \frac{1}{t}m(Y, X).$$

Taking  $t \rightarrow 1$ , we find that  $m(Y, X) = \lim_n m(Y, X_n)$ , completing the proof of (1).

To prove that  $(\mathbb{S}_D^\circ, d_1)$  and  $(\mathbb{S}_D^\circ, d)$  are homeomorphic, note that  $d_1 \leq 2d$  by Lemma 3.8. Thus convergence in  $d$  implies convergence in  $d_1$  on all of  $\mathbb{S}_D$ . On the other hand, if  $X_n \rightarrow X \in \mathbb{S}_D^\circ$  with respect to  $d_1$  then  $\lim_n d(X_n, X) = d(X, X) = 0$ , by the first part of the lemma, so  $X_n$  converge to  $X$  with respect to  $d$ .  $\square$

For any map  $\phi \in \mathcal{P}_D$  we define the *contraction coefficient*

$$c(\phi) = \sup \{d(\phi \cdot X, \phi \cdot Y) : X, Y \in \mathbb{S}_D\}. \quad (3.6)$$

The following Lemma lists various properties of  $c(\phi)$ .

**Lemma 3.10.** *Let  $\phi \in \mathcal{P}_D$ , then*

- (1) *For  $X, Y \in \mathbb{S}_D$ ,  $d(\phi \cdot X, \phi \cdot Y) \leq c(\phi)d(X, Y)$ .*
- (2) *We have  $c(\phi) \leq 1$  and  $c(\phi) < 1$  if and only if  $\phi \in \mathcal{P}_D^\circ$ .*
- (3) *If  $\phi' \in \mathcal{P}_D$ , then  $c(\phi' \circ \phi) \leq c(\phi')c(\phi)$ .*
- (4)  *$c(\phi) = c(\phi^*)$ .*

*Remark.* Thus, if  $\phi \in \mathcal{P}_D^\circ$ , then the projective action of  $\phi$  on  $\mathbb{S}_D$  is *strictly* contractive with respect to the metric  $d$ .

*Proof.* To prove (1), suppose that  $\phi \in \mathcal{P}_D$ . If  $\phi \cdot X = \phi \cdot Y$ , then  $0 = d(\phi \cdot X, \phi \cdot Y) \leq c(\phi)d(X, Y)$ . Now suppose that  $\phi \cdot X \neq \phi \cdot Y$  and let  $t_\pm$  and  $A_\pm$  be as in the proof of Lemma 3.8. Similarly, let  $A'_\pm = s_\pm \phi \cdot X + (1 - s_\pm)\phi \cdot Y$  with  $s_\pm$  the largest and smallest real numbers such that  $s\phi \cdot X + (1 - s)\phi \cdot Y \in \mathbb{S}_D$ .

The linear map  $\phi$  maps the two dimensional space spanned by  $A_-, A_+$  into the two dimensional space spanned by  $A'_-, A'_+$ . Let the matrix of this map (with respect to the bases  $A_-, A_+$  for the domain and  $A'_-, A'_+$  for the range) be

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}.$$

We claim that  $\alpha, \beta, \gamma, \delta \geq 0$ . To see that  $\alpha, \gamma \geq 0$ , note that

$$\phi(A_-) = t_- \phi(X) + (1 - t_-) \phi(Y) = [t_- \operatorname{tr} \phi(X) \phi \cdot X + (1 - t_-) \operatorname{tr} \phi(Y) \phi \cdot Y].$$

Thus

$$\phi \cdot A_- = \frac{t_- \operatorname{tr}[\phi(X)]}{\operatorname{tr}[\phi(A_-)]} \phi \cdot X + \left[ 1 - \frac{t_- \operatorname{tr}[\phi(X)]}{\operatorname{tr}[\phi(A_-)]} \right] \phi \cdot Y.$$

Since  $\phi \cdot A_- \in \mathbb{S}_D$  we must have  $s_- \operatorname{tr}[\phi(A_-)] \leq t_- \operatorname{tr}[\phi(X)] \leq s_+ \operatorname{tr}[\phi(A_-)]$ . Thus  $\phi(A_-) = \alpha A'_- + \gamma A'_+$  with

$$\alpha = \frac{t_- \operatorname{tr}[\phi(X)] - s_- \operatorname{tr}[\phi(A_-)]}{s_+ - s_-} \geq 0, \quad \gamma = \frac{s_+ \operatorname{tr}[\phi(A_-)] - t_- \operatorname{tr}[\phi(X)]}{s_+ - s_-} \geq 0.$$

The verification that  $\beta \geq 0$  and  $\delta \geq 0$  is similar.

We also have  $\alpha\delta + \beta\gamma > 0$ . Indeed if  $\alpha\delta + \beta\gamma$  were zero, then the matrix would have a zero row or a zero column. A zero column would imply that one of  $\phi(A_-)$  or  $\phi(A_+)$  is zero, a contradiction. A zero row would imply that  $\phi(A_+)$  and  $\phi(A_-)$  were both proportional either to  $A'_+$  or  $A'_-$ . Suppose both were proportional to  $A'_-$ . Then both points would lie on the line between 0 and  $A'_-$  and also on the line between  $A'_-$  and  $A'_+$ . Since these lines intersect only in  $A'_-$  we would have  $\phi(A_+) = \phi(A_-) = A'_-$ , contradicting the assumption that  $\phi \cdot X \neq \phi \cdot Y$ .

With these preliminaries, we can now prove (1) by computing  $d(\phi \cdot X, \phi \cdot Y)$ . Let  $X = u_1 A_- + u_2 A_+$ ,  $Y = v_1 A_- + v_2 A_+$ . Then,

$$\begin{aligned} d(\phi \cdot X, \phi \cdot Y) &= \frac{|(\alpha u_1 + \beta u_2)(\gamma v_1 + \delta v_2) - (\gamma u_1 + \delta u_2)(\alpha v_1 + \beta v_2)|}{(\alpha u_1 + \beta u_2)(\gamma v_1 + \delta v_2) + (\gamma u_1 + \delta u_2)(\alpha v_1 + \beta v_2)} \\ &= \frac{|\alpha\delta - \beta\gamma| |u_1 v_2 - u_2 v_1|}{\alpha\gamma 2u_1 v_1 + (\alpha\delta + \beta\gamma)(u_1 v_2 + u_2 v_1) + \beta\delta 2u_2 v_2} \\ &\leq \frac{|\alpha\delta - \beta\gamma|}{\alpha\delta + \beta\gamma} \frac{|u_1 v_2 - u_2 v_1|}{u_1 v_2 + u_2 v_1} \\ &= d(\phi \cdot A_-, \phi \cdot A_+) d(X, Y) \leq c(\phi) d(X, Y). \end{aligned}$$

Turning now to (2), if  $\phi \in \mathcal{P}_D \setminus \mathcal{P}_D^\circ$ , then  $\phi \cdot X \in \mathbb{S}_D \setminus \mathbb{S}_D^\circ$  for some  $X \in \mathbb{S}_D \setminus \mathbb{S}_D^\circ$  (otherwise  $\phi$  would be strictly positive). Thus,  $c(\phi) = 1$ , since by Lemma 3.5,  $d(\phi \cdot X, \phi \cdot Y) = 1$  for  $Y \in \mathbb{S}_D^\circ$ . To see that  $c(\phi) < 1$  for  $\phi \in \mathcal{P}_D^\circ$ , note that  $\phi \cdot$  is a continuous map from  $(\mathbb{S}_D, d_1)$  into  $(\mathbb{S}_D^\circ, d_1)$ . By Lemma 3.9,  $F(X, Y) = d(\phi \cdot X, \phi \cdot Y)$  is a continuous map of  $\mathbb{S}_D \times \mathbb{S}_D$  into  $\mathbb{R}$ , where we take the  $d_1$ -product topology on  $\mathbb{S}_D \times \mathbb{S}_D$ . Since  $\mathbb{S}_D \times \mathbb{S}_D$  is compact we conclude that there are  $X, Y \in \mathbb{S}_D$  such that  $c(\phi) = d(\phi \cdot X, \phi \cdot Y)$ . Since  $\phi \cdot X, \phi \cdot Y \in \mathbb{S}_D^\circ$  we have  $0 < m(\phi \cdot X, \phi \cdot Y) < 1$  and  $0 < m(\phi \cdot Y, \phi \cdot X)$ , so that  $c(\phi) = d(\phi \cdot X, \phi \cdot Y) < 1$ .

To prove (3), note that  $\phi' \circ \phi \cdot X = \phi' \cdot (\phi \cdot X)$ , so that

$$d(\phi' \circ \phi \cdot X, \phi' \circ \phi \cdot Y) \leq c(\phi') d(\phi \cdot X, \phi \cdot Y) \leq c(\phi) c(\phi') d(X, Y),$$

by part (1).

Finally, to prove that  $c(\phi) = c(\phi^*)$ , we use the variational formula (3.2) which implies

$$m(X, Y)m(Y, X) = \inf \left\{ \frac{\text{tr}[AX] \text{tr}[A'Y]}{\text{tr}[A'X] \text{tr}[AY]} : A, A' \in \mathbb{S}_D^\circ \right\}.$$

It follows that

$$\begin{aligned} m(\phi \cdot X, \phi \cdot Y)m(\phi \cdot Y, \phi \cdot X) &= \inf \left\{ \frac{\text{tr}[\phi^*(A)X] \text{tr}[\phi^*(A')Y]}{\text{tr}[\phi^*(A')X] \text{tr}[\phi^*(A)Y]} : A, A' \in \mathbb{S}_D^\circ \right\} \\ &\geq \inf \{m(\phi^* \cdot A, \phi^* \cdot A')m(\phi^* \cdot A', \phi^* \cdot A) : A, A' \in \mathbb{S}_D^\circ\}, \end{aligned}$$

and thus that

$$\begin{aligned} \inf \{m(\phi^* \cdot X, \phi^* \cdot Y)m(\phi^* \cdot Y, \phi^* \cdot X) : X, Y \in \mathbb{S}_D^\circ\} \\ = \inf \{m(\phi \cdot X, \phi \cdot Y)m(\phi \cdot Y, \phi \cdot X) : X, Y \in \mathbb{S}_D^\circ\}, \end{aligned}$$

from which it follows that  $c(\phi) = c(\phi^*)$ .  $\square$

**3.3. Existence of  $Z_0$  and  $Z'_0$ .** We start by proving Lemma 2.1, which states the equivalence of Assumption 1 to two conditions, which we reformulate here in the notation of the Sect. 3.1:

- (1) For some  $n_0 > 0$ ,  $\Pr[\Phi_{n_0} \in \mathcal{P}_D^\circ] > 0$ .
- (2) With probability one,  $\phi_0 \in \mathcal{P}_D$ .

Recall that  $\Phi_N = \phi_N \circ \cdots \circ \phi_0$ , where  $\phi_n = \phi_{T^n \omega}$ . Let

$$\tau = \inf\{N_0 \geq 0 : \Phi_N \in \mathcal{P}_D^\circ \text{ for } N \geq N_0\}.$$

Note that Assumption 1 is equivalent to the statement that  $\tau < \infty$  with probability one.

*Proof of Lemma 2.1.* We first show that Conditions (1) and (2) imply  $\tau < \infty$  with probability one. By ergodicity and condition (1),

$$\Pr\left[\bigcup_{k \geq 0} \{\Phi_{n_0; T^k \omega} \in \mathcal{P}_D^\circ\}\right] = 1.$$

Thus with probability 1 there is  $\sigma < \infty$  such that  $\Phi_{n_0; T^\sigma \omega} = \phi_{\sigma+n_0} \circ \cdots \circ \phi_\sigma \in \mathcal{P}_D^\circ$ . By Condition (1) and the shift invariance of probabilities, we have  $\phi_n \in \mathcal{P}_D$  for all  $n$ , with probability one. By Corollary 3.2, it follows that

$$\Phi_N = \phi_N \circ \cdots \phi_{n_0+\sigma+1} \circ \Phi_{n_0; T^\sigma \omega} \circ \phi_{\sigma-1} \circ \cdots \circ \phi_0$$

is strictly positive for  $N \geq n_0 + \sigma$ , so  $\tau \geq n_0 + \sigma$ .

Conversely, note that Assumption 1 implies Condition (1) directly. To prove Condition (2), note that for  $N > 0$ , we have  $\ker \Phi_N \supset \ker \phi_0$ . It follows that  $\ker \phi_0 \cap \mathbb{P}_D = \{0\}$  if  $\Phi_N \in \mathcal{P}_D^\circ$  for some  $N$ . Thus Assumption 1 implies that  $\ker \phi_0 \cap \mathbb{P}_D = \{0\}$  with probability one. To prove the corresponding statement for  $\phi_0^*$ , first note that if  $\Phi_N = \phi_N \circ \Phi_{N-1}$  is strictly positive, then we have  $\text{tr } \phi_N^*(A) \Phi_{N-1}(B) > 0$  for every  $A, B \in \mathbb{P}_D \setminus \{0\}$ . Thus Assumption 1 implies that, with probability one,  $\ker \phi_N^* \cap \mathbb{P}_D = \{0\}$  for all sufficiently large  $N$ . Let  $A_M$  denotes the event

$$A_M = \bigcap_{N \geq M} \{\omega : \ker \phi_{N; \omega}^* \cap \mathbb{P}_D = \{0\}\}.$$

Then  $(A_M)_{M=0}^\infty$  is an increasing sequence and  $\Pr[\bigcup_M A_M] = 1$ . Thus  $\lim_M \Pr[A_M] = 1$ . However,  $A_M = T^M(A_0)$  so  $\Pr[A_M] = \Pr[A_0]$  for all  $M$ . We conclude that  $\Pr[A_0] = 1$  and thus that  $\ker \phi_0^* \cap \mathbb{P}_D = \{0\}$  with probability one.  $\square$

**Lemma 3.11.** *Let  $c_N = c(\Phi_N)$ , with  $c(\cdot)$  the contraction coefficient in (3.6). Then*

$$\lim_{N \rightarrow \infty} c_N^{1/N} = \inf_N c_N^{1/N} \equiv \kappa$$

*exists almost surely, where  $\kappa \in [0, 1)$  is non-random and*

$$\ln \kappa = \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E} \{\ln c_N\} = \inf_N \frac{1}{N} \mathbb{E} \{\ln c_N\}. \quad (3.7)$$

*Proof.* Note that  $\ln c_{N+M} = \ln c(\Phi_{N+M}) = \ln c(\Phi_{N;T^{M+1}\omega} \circ \Phi_M) \leq \ln c(\Phi_{N;T^{M+1}\omega}) + \ln c(\Phi_M)$  and  $\ln c(\Phi_N) \leq 0$  it follows from the subadditive ergodic theorem [13] that the limit and infimum exist, and that Eq. (3.7) holds. Since  $0 \leq c(\Phi_N) \leq 1$ , we have  $0 \leq \kappa \leq 1$ . By Condition (1) of Lemma 2.1 and Lemma 3.10, we see that  $c_{n_0} < 1$  with positive probability. Thus  $n_0^{-1} \mathbb{E} \{ \ln c_{n_0} \} < 0$  and so  $\ln \kappa < 0$ .  $\square$

We can now prove the existence of the limiting processes:

**Lemma 3.12.** *Let  $\Phi_N^*$  and  $L_N$  be as in Eq. (2.6). As  $N \rightarrow \infty$ ,  $L_N$  converges almost surely to a limit  $Z'_0$  such that:*

- (1)  $Z'_0 \in \mathbb{S}_D^\circ$  almost surely;
- (2)  $\phi_0^* \cdot Z'_{0;T\omega} = Z'_{0;\omega}$ ; and
- (3) for  $Y \in \mathbb{S}_D$  and  $N \geq 0$ , we have  $d(\Phi_N^* \cdot Y, Z'_0) \leq c(\Phi_N)$ .

*Proof.* Let  $B_N = \Phi_N^* \cdot \mathbb{S}_D$ , so  $B_N \subset B_{N-1}$ . It follows from Assumption 1 that  $B_N \subset \mathbb{S}_D^\circ$  for large enough  $N$ . Thus, by Lemma 3.9,  $B_N$  is compact in the  $d$ -topology for large  $N$  (since  $B_N$  is compact in the  $d_1$ -topology for every  $N$ ). Thus  $\cap_N B_N$  is non-empty. On the other hand,

$$\text{diam } B_N = \sup \{ d(\Phi_N^* \cdot X, \Phi_N^* \cdot Y) : X, Y \in \mathbb{S}_D \} \leq c(\Phi_N) \rightarrow 0,$$

by Lemmas 3.10 and 3.11. Thus  $\cap_N B_N = \{Z'_0\}$  for a single point  $Z'_0$ . It is clear that  $Z'_0 \in \mathbb{S}_D^\circ$  almost surely.

We claim that  $L_N \in B_N$ . Indeed, since  $\Phi_N^*(L_N) = \lambda_N L_N$  and  $\text{tr}[L_N] = 1$ , it follows that  $\Phi_N^* \cdot L_N = L_N$ . Thus,  $d(L_N, Z'_0) \leq \text{diam } B_N \rightarrow 0$  almost surely. It follows that  $Z'_{0;T\omega} = \lim_{N \rightarrow \infty} L_{N;T\omega}$ . However  $L_{N;T\omega}$  is a normalized eigenmatrix for  $\Phi_{N;T\omega}^* = \phi_1^* \circ \dots \circ \phi_{N+1}^*$ . Thus  $\phi_0^* \cdot L_{N;T\omega} = \Phi_{N+1}^* \cdot L_{N;T\omega} \in B_{N+1}$ , from which it follows that  $\phi_0^* \cdot L_{N;T\omega} \rightarrow Z'_0$ . Finally, let  $Y \in \mathbb{S}_D$ . Then  $\Phi_N^* \cdot Y \in B_N$ , so  $d(\Phi_N^* \cdot Y, Z'_0) \leq \text{diam } B_N \leq c(\Phi_N)$  as claimed.  $\square$

A similar argument can be applied to  $\Phi_{-N}^*$  to conclude the existence and properties of  $Z_0$ . To this end, let  $\psi_n = \phi_{-n}^*$  and

$$\Psi_N = \psi_N \circ \dots \circ \psi_0 = \phi_{-N}^* \circ \dots \circ \phi_0^* = \Phi_{-N}^*.$$

First we note that the process  $\Psi_N$  satisfies Assumption 1:

**Lemma 3.13.** *With probability one, there is  $N'_0 < \infty$  such that  $\Psi_N \in \mathcal{P}_D^\circ$  for all  $N \geq N'_0$ .*

*Proof.* We will show that Conditions (1) and (2) of Lemma 2.1 hold. Condition (2) for  $\psi_0$  follows from the corresponding statement for  $\phi_0$ , since  $\psi_0 = \phi_0^*$  and  $\psi_0^* = \phi_0$ . To see that Condition (1) holds, note that  $\Psi_{n_0}^* = \Phi_{-n_0;\omega} = \Phi_{n_0;T^{-n_0}\omega}$ . Thus by the shift invariance of probabilities,  $\Pr[\Psi_{n_0}^* \in \mathcal{P}_D^\circ] = \Pr[\Psi_{n_0} \in \mathcal{P}_D^\circ] > 0$ .  $\square$

The existence of  $Z_0$  follows directly from Lemma 3.12 applied to  $\Psi_N = \Phi_{-N}^*$ :

**Lemma 3.14.** *Let  $\Phi_N$  and  $R_N$  be as in Eq. (2.5). As  $N \rightarrow -\infty$ ,  $R_N$  converges almost surely to a limit  $Z_0$  such that:*

- (1)  $Z_0 \in \mathbb{S}_D^\circ$  almost surely;
- (2)  $\phi_0 \cdot Z_{0;T^{-1}\omega} = Z_0$ ; and
- (3) for  $Y \in \mathbb{S}_D$  and  $N \geq 0$ , we have  $d(\Phi_{-N} \cdot Y, Z_0) \leq c(\Phi_{-N})$ .

## 4. Proofs of the Theorems

**4.1. Proof of Theorem 1.** We have already shown the existence of the limits  $\lim_{N \rightarrow \infty} L_N = Z'_0$  and  $\lim_{N \rightarrow -\infty} R_N = Z_0$ . Let  $Z_n = Z_{0;T^n\omega}$  and  $Z'_n = Z'_{0;T^n\omega}$ . Then  $Z_0 = \phi_0 \cdot Z_{-1}$  and  $Z_0 = \phi_0^* \cdot Z'_1$  by Lemmas 3.12 and 3.14. Thus  $Z_n = \phi_n \cdot Z_{n-1}$  and  $Z'_n = \phi_n^* \cdot Z'_{n+1}$  as claimed.  $\square$

**4.2. Proof of Theorem 2.** Let  $\mu \in (\kappa, 1)$  be as in Lemma 3.11. Given  $m < n$ , let  $\Psi_{n,m} = \phi_n \circ \dots \circ \phi_m$  and  $P_{n,m}(M) = \text{tr}[Z'_m M] Z_n$ . To prove Eq. (2.10), we must show that

$$\text{tr} \left| \frac{1}{\text{tr}[\Psi_{n,m}^*(\mathbb{I})]} \Psi_{n,m}(M) - P_{n,m}(M) \right| \leq C_{\mu,x} \mu^{n-m} \text{tr}|M| \quad (4.1)$$

whenever  $m \leq x \leq n$ . In fact, it suffices to prove Eq. (4.1) for  $M \in \mathbb{S}_D$ . Indeed, any matrix  $M$  can be written as a linear combination

$$M = \sum_{j=1}^4 a_j M_j \quad \text{with } M_j \in \mathbb{S}_D \text{ and } \sum_{j=1}^4 |a_j| \leq 2 \text{tr}(|M|). \quad (4.2)$$

Thus Eq. (4.1) for  $M \in \mathbb{S}_D$  implies the same bound for general  $M$ , with the constant  $C_{\mu,x}$  increased by a factor of 2. (To see that Eq. (4.2) holds, note that for self-adjoint  $M$  we have  $M = \text{tr}[M_+] \rho_+ - \text{tr}[M_-] \rho_-$ , where  $\rho_{\pm} = \frac{1}{\text{tr}[M_{\pm}]} M_{\pm}$  with  $M_{\pm}$  the positive and negative parts of  $M$ . For a general matrix  $M$ , we proceed by applying this decomposition to the real and imaginary parts  $M = M_r + i M_i$ , where  $M_r = \frac{1}{2}(M + M^\dagger)$  and  $M_i = \frac{1}{2i}(M - M^\dagger)$ .)

Now let  $M \in \mathbb{S}_D$  be fixed. Note that  $\Psi_{n,m} = \Phi_{m-n;T^n\omega}$  and  $\Psi_{n,m}^* = \Phi_{n-m;T^m\omega}^*$ . By Lemma 3.8 and Lemma 3.14 we have

$$\text{tr} \left[ \left| \frac{1}{\text{tr} \Psi_{n,m}(M)} \Psi_{n,m}(M) - Z_n \right| \right] \leq 2c(\Psi_{n,m}).$$

By Lemma 3.8 and Lemma 3.12, we have

$$\left| \frac{\text{tr}[\Psi_{n,m}(M)]}{\text{tr}[\Psi_{n,m}^*(\mathbb{I})]} - \text{tr}[Z'_m M] \right| \leq 2c(\Psi_{n,m}),$$

where we have noted that  $\text{tr}[\Psi_{n,n}^*(\mathbb{I})M] = \text{tr}[\Psi_{n,m}(M)]$ . Thus

$$\text{tr} \left[ \left| \frac{1}{\text{tr}[\Psi_{n,m}^*(\mathbb{I})]} \Psi_{n,m}(M) - P_{n,m}(M) \right| \right] \leq 2c(\Psi_{n,m}) \left( 1 + \frac{\text{tr}[\Psi_{n,m}(M)]}{\text{tr}[\Psi_{n,m}^*(\mathbb{I})]} \right) \leq 4c(\Psi_{n,m}), \quad (4.3)$$

since  $\text{tr}[\Psi_{n,m}(M)] = \text{tr}[\Psi_{n,m}^*(\mathbb{I})M] \leq \text{tr}[\Psi_{n,m}^*(\mathbb{I})]$  for  $M \in \mathbb{S}_D^\circ$ .

To prove Eq. (4.1), first suppose that  $m \leq x < n$ . Then  $c(\Psi_{n,m}) \leq c(\Psi_{n,x+1})c(\Psi_{x,m})$ . By Lemma 3.11 we have  $c(\Psi_{n,x+1}) \leq D_{\mu,x} \mu^{n-x}$  and  $c(\Psi_{x,m}) \leq D_{\mu,x} \mu^{x-m}$  for suitable  $D_{\mu,x} < \infty$ . Thus  $c(\Psi_{n,m}) \leq D_{\mu,x}^2 \mu^{n-m}$ , so Eq. (4.1) follows from Eq. (4.3). For  $m \leq x = n$ , we have  $c(\Psi_{n,m}) = c(\Psi_{x,m}) \leq D_{\mu,x} \mu^{x-m}$ , so Eq. (4.1) holds in this case as well.  $\square$

**4.3. Proof of Theorem 3.** We have

$$\begin{aligned} & W(O_2 O_1) - W(O_2) W(O_1) \\ &= \text{tr} [(\tilde{O}_2 - W(O_2) \tilde{\Psi}_2) \circ \tilde{\phi}_{m_2-1} \circ \cdots \circ \tilde{\phi}_{n_1+1} \circ (\tilde{O}_1 (\tilde{Z}_{m_1-1}) - W(O_1) \tilde{\Psi}_1 (\tilde{Z}_{m_1-1}))]. \end{aligned}$$

By Theorem 2, there is  $0 < \mu < 1$  such that

$$\text{tr} [|\tilde{\phi}_{m_2-1} \circ \cdots \circ \tilde{\phi}_{n_1+1}(M) - \tilde{Z}_{n_1} \text{tr } M|] \leq C_{\mu,x} \mu^{m_2-n_1} \text{tr}[|M|],$$

for any  $D \times D$  matrix  $M$ , where we have used the fact that  $\tilde{\phi}_{n_1+1}^* \circ \cdots \circ \tilde{\phi}_{m_2-1}^*(\mathbb{I}) = \mathbb{I}$ . Equation (2.20) then follows.  $\square$

*Acknowledgements.* RM acknowledges the support of the IBM Research Frontiers Institute and funding from the MIT-IBM Watson AI Lab under the project Machine Learning in Hilbert space. JS acknowledges the support of the National Science Foundation under Grant Nos. 1500386 and 1900015, and thanks Lubashan Pathirana for insightful discussions. We thank Natalie Taylor and Jules Murphy at IBM for their help with the graphic design of Fig. 1.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

1. Brandão, F.G.S.L., Horodecki, M.: An area law for entanglement from exponential decay of correlations. *Nat. Phys.* **9**(11), 721–726 (2013)
2. Bruneau, L., Joye, A., Merkli, M.: Repeated interactions in open quantum systems. *J. Math. Phys.* **55**(7), 075204 (2014)
3. Burgarth, D., Chiribella, G., Giovannetti, V., Perinotti, P., Yuasa, K.: Ergodic and mixing quantum channels in finite dimensions. *New J. Phys.* **15**(7), 073045 (2013)
4. Collins, B., Nechita, I.: Random quantum channels I: graphical calculus and the bell state phenomenon. *Commun. Math. Phys.* **297**(2), 345–370 (2010)
5. Collins, B., Nechita, I.: Random quantum channels II: entanglement of random subspaces, Rényi entropy estimates and additivity problems. *Adv. Math.* **226**(2), 1181–1201 (2011)
6. Collins, B., Nechita, I.: Random matrix techniques in quantum information theory. *J. Math. Phys.* **57**(1), 015215 (2016)
7. Evans, D.E., Høegh-Krohn, R.: Spectral properties of positive maps on  $c^*$ -algebras. *J. Lond. Math. Soc.* **2**(2), 345–355 (1978)
8. Fannes, M., Nachtergael, B., Werner, R.F.: Valence bond states on quantum spin chains as ground states with spectral gap. *J. Phys. A Math. Gen.* **24**, L185–L189 (1991)
9. Frobenius, G.: Über matrizen aus nicht negativen elementen. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin* **1912**(1), 456–477 (1912)
10. Hastings, M.B.: An area law for one-dimensional quantum systems. *J. Stat. Mech. Theory Exp.* **2007**(08), P08024 (2007)
11. Hastings, M.B., Koma, T.: Spectral gap and exponential decay of correlations. *Commun. Math. Phys.* **265**(3), 781–804 (2006)
12. Hennion, H.: Limit theorems for products of positive random matrices. *Ann. Probab.* **25**(4), 1545–1587 (1997)
13. Kingman, J.F.C.: Subadditive ergodic theory. *Ann. Probab.* **1**, 883–909 (1973)
14. Kraus, K.: General state changes in quantum theory. *Ann. Phys.* **64**, 311–335 (1971)
15. Krein, M.G., Rutman, M.A.: Linear operators leaving invariant a cone in a Banach space. *Am. Math. Soc. Transl.* **1950**(26), 128 (1950)
16. Li, J., Harter, A.K., Liu, J., de Melo, L., Joglekar, Y.N., Luo, L.: Observation of parity-time symmetry breaking transitions in a dissipative floquet system of ultracold atoms. *Nat. Commun.* **10**(1), 1–7 (2019)
17. Lindner, N.H., Refael, G., Galitski, V.: Floquet topological insulator in semiconductor quantum wells. *Nat. Phys.* **7**(6), 490–495 (2011)
18. Movassagh, R., Schenker, J.: Theory of ergodic quantum processes. *Phys. Rev. X* **11**, 041001 (2021)
19. Nachtergael, B., Sims, R.: Lieb–Robinson bounds and the exponential clustering theorem. *Commun. Math. Phys.* **265**(1), 119–130 (2006)

20. Oseledec, V.: A multiplicative ergodic theorem, characteristic Lyapnov exponents of dynamical systems (transactions of the Moscow Mathematical Society, 19). American Mathematical Society, Providence (1968)
21. Perez-Garcia, D., Verstraete, F., Wolf, M.M., Cirac, J.I.: Matrix product state representations. *Quantum Inf. Comput.* **7**(5), 401–430 (2007)
22. Perron, O.: Zur theorie der matrices. *Math. Ann.* **64**(2), 248–263 (1907)
23. Shtanko, O., Movassagh, R.: Stability of periodically driven topological phases against disorder. *Phys. Rev. Lett.* **121**(12), 126803 (2018)
24. Verstraete, F., Murg, V., Cirac, J.I.: Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Adv. Phys.* **57**(2), 143–224 (2008)
25. Watrous, J.: The Theory of Quantum Information. Cambridge University Press, Cambridge (2018)
26. White, S.R.: Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.* **69**(19), 2863 (1992)

Communicated by A. Giuliani