Decay of quantum conditional mutual information for purely generated finitely correlated states

Pavel Svetlichny and T.A.B. Kennedy
School of Physics, Georgia Institute of Technology, Atlanta, GA, 30332-0430, USA.

(Dated: October 19, 2022)

The connection between quantum state recovery and quantum conditional mutual information (QCMI) is studied for the class of purely generated finitely correlated states (pgFCS) of one-dimensional quantum spin chains. For a tripartition of the chain into two subsystems separated by a buffer region, it is shown that a pgFCS is an approximate quantum Markov chain, and stronger, may be approximated by a quantum Markov chain in trace distance, with an error exponentially small in the buffer size. This implies that, (1) a locally corrupted state can be approximately recovered by action of a quantum channel on the buffer system, and (2) QCMI is exponentially small in the size of the buffer region. Bounds on the exponential decay rate of QCMI and examples of quantum recovery channels are presented.

I. INTRODUCTION

In this paper we investigate properties of quantum conditional mutual information (QCMI) and quantum state recovery for the class of many-particle states known alternatively as purely generated finitely correlated states (pgFCS) or uniform matrix product states.[1–3] These states were introduced in the early 1990’s and used to describe the ground states of translationally invariant one-dimensional spin chains, including the well-known model of Affleck, Kennedy Lieb and Tasaki (AKLT) as a special case.[4] Independently, the matrix product states were introduced by Klümper et al [3] and separately shown by Vidal [5] to be an efficient representation of slightly entangled states, such as the ground states of local gapped Hamiltonians.[6] Proposals for the controlled experimental preparation of such states have been reported in the literature.[7, 8] The larger class of finitely correlated states (FCS) and the closely related matrix product density operators (MPDO),[9] can be used to describe mixed states of systems with local Hamiltonians and finite range interactions, such as finite temperature Gibbs states.[10, 11]

Positivity of QCMI is one of the cornerstones of quantum information theory, being equivalent to the strong subadditivity of quantum entropy.[12, 13] In the theory of quantum state recovery,[14–20] viewed as a generalized quantum error correction problem, the states for which recovery is exact are known as quantum Markov chains. These states are intimately connected with the QCMI as is clear in the following setting. Consider a quantum system (a collection of spatially separated spins) subdivided into three parts, denoted A, B and C, see Figure 1a. The QCMI, denoted I(A : C|B) indicates in broad terms the quantum information mutually shared by subsystems A and C given specific information about the state of subsystem B. Two essential and non-trivial properties of I(A : C|B) are that it is non-negative I(A : C|B) ≥ 0, as a direct consequence of the strong subadditivity of quantum entropy, [13] and that the equality condition I(A : C|B) = 0 defines the class of quantum states with density operator ρABC, which are precisely the quantum Markov chains. In this setting the problem of quantum state recovery, in either exact or approximate form, may be approached by investigating when I(A : C|B) is zero or very small, conditions which require investigation of quantum states described as either exact or approximate quantum Markov chains. This specific quantum information-theoretic task gives QCMI a compelling physical interpretation.

We study quantum state recovery for a one-dimensional quantum system ABC for which the density operator ρABC is the reduced density operator of a pgFCS, respectively, uniform matrix product state. From here on we will principally use the former appellation and refer the reader to the dictionary between these two languages compiled in Appendix A.

As we discuss further below, ρABC is an approximate quantum Markov chain and hence can be approximately recovered.

The pgFCS is a subset of the class of FCS, some of whose properties we review in Section III C. The FCS can efficiently describe systems with limited entanglement and their preparation by using quantum circuits of limited depth.[11, 21] It was recently shown that FCS may be used to approximate Gibbs states of systems described by one-dimensional local Hamiltonians.[10, 11] In Refs. 22 and 23 it was shown that the same class of states can be approximated by matrix product operators (MPO). It is not immediately clear that the resulting MPO are FCS or MPDO. The latter are called locally purifiable in the sense of Ref. 9.) Conversely, it is conjectured that a generic FCS approximates the Gibbs state of some local gapped Hamiltonian.[24] While Gibbs states are of broad physical interest, their sampling has been employed in a quantum algorithm proposed to speed up the computation of semi-definite programs.[25–27]

The QCMI is defined for a system ABC with the Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ by [28]

$$I(A : C|B) := S(\rho_{AB}) + S(\rho_{BC}) - S(\rho_{ABC}) - S(\rho_B),$$

(1)

where $S(\rho) = -\text{tr}(\rho \log \rho)$ is the von Neumann entropy, and omission of a subscript implies taking a partial trace, e.g., $\rho_{AB} = \text{tr}_C(\rho_{ABC})$. As noted earlier, QCMI is non-negative, $I(A : C|B) \geq 0$, as an immediate consequence of the strong subadditivity of von-Neumann entropy. [13] In Refs. 14 and 15 it was established that the condition $I(A : C|B) = 0$ (ρABC is a quantum Markov chain) is equivalent to the existence of a quantum channel, a completely positive trace-preserving map, [28, 29] $\mathcal{R}_{B \to BC}$, which recovers $\rho_{ABC}$ from $\rho_{AB}$ exactly, i.e., $\mathcal{R}_{B \to BC}(\rho_{AB}) = \rho_{ABC}$. From this property follows the particular structure of quantum Markov chains which we summarize in Theorem III.5. [16] The recovery map is not unique, and we
can always explicitly construct the Petz recovery map, which is defined (on the support of $\rho_B$) by [14-16]

$$\mathcal{R}_{B\to BC}(X) := \frac{1}{\rho_{BC}} \left( \rho_B^{\frac{1}{2}} X \rho_B^{\frac{1}{2}} \otimes I_C \rho_{BC} \right),$$

(2)

where $\rho_B^{-\frac{1}{2}}$ is a pseudo-inverse of $\rho_B^\frac{1}{2}$. In the case when a recovery quantum channel does not restore the state exactly, the recovery error is quantified by the trace distance,

$$\varepsilon := \| \rho_{ABC} - \mathcal{R}_{B\to BC}(\rho_{AB}) \|_1,$$

where $\|X\|_1 = \text{tr}(\sqrt{X^\dagger X})$. The recovery error, for a candidate recovery channel $\mathcal{R}_{B\to BC}$, bounds QCMI from above via an Aliki-Fannes type inequality [30, 31] (see Appendix B),

$$I(A : C | B) \leq \frac{1}{2} \varepsilon \ln \dim H_A + \left( 1 + \frac{1}{2\varepsilon} \right) h \left( \frac{\varepsilon}{2 + \varepsilon} \right),$$

(3)

where $h(p) = -p \ln p - (1-p) \ln(1-p)$ is the binary entropy. This bound was used to show that a Gibbs state of a 1D local Hamiltonian is an approximate quantum Markov chain [11]. By constructing a specific recovery channel, for which the recovery error is subexponentially small in the size of the region $B$, it was shown that the QCMI is also subexponentially small. In this paper we similarly construct an approximate recovery channel for pgFCS, and bound QCMI from above using (3).

Theoretical developments regarding the reversibility of quantum channels obtained in a series of works Refs. 17-19, led to the discovery of a recovery channel $\mathcal{R}_{B\to BC}^{\mu}$, called the universal recovery map, which may be used to bound QCMI from below,

$$\frac{1}{4 \ln 2} \left\| \rho_{ABC} - \mathcal{R}_{B\to BC}^{\mu}(\rho_{AB}) \right\|_1^2 \leq I(A : C | B).$$

(4)

The universal recovery map, which is a generalization of the Petz recovery map (2), is given explicitly in Ref. 19 and, as in the case of the Petz recovery map, depends only on the marginal $\rho_{BC}$. The inequality (4) implies that states with small QCMI are guaranteed to have a small recovery error with respect to the universal recovery map. In Ref. 10, it was shown for local Hamiltonian systems of arbitrary dimension, and high enough temperature, that QCMI for Gibbs states decays exponentially with the width of region $B$, separating regions $A$ and $C$. By (4) this implies good recovery by the universal recovery map. (It is conjectured, [11] that in one dimension the exponential decay of QCMI holds for any temperature.) Similar conclusions were found for finite temperature Gibbs states of free fermions, free bosons, conformal field theories, and holographic models [32].

A more physical picture of the recovery map is as an experimental repair process, with an accuracy bounded by how rapidly QCMI decreases as a function of the distance $d(A,C)$ shown in Figure 1.[20] For example, for the lattice of spins shown in Figure 1(a), suppose that quantum state corruption or erasure has happened in the region $C$. It is then natural to partition the lattice by means of a buffer zone $B$ surrounding $C$.[21, 32] Qualitatively, quantum states with rapidly decaying QCMI may be recovered accurately on erased regions with small buffer zones $B$.[11, 21, 32] In one dimension the analysis can be reduced to the partition shown in Figure 1(c),[11] which we focus on in this paper. We briefly consider the setting of Figure 1(b) in Appendix C.

In the language of quantum memories, QCMI is connected to the task of storing quantum information in the presence of quantum error correction,[20] and to state preparation by circuits of local quantum channels,[11, 21, 32] Let us assume that quantum channels can be applied experimentally over regions with size limited by the linear dimension $l$. Conceptually, a quantum state may be prepared on a much larger region by first constructing reduced states on a grid of disconnected regions and then patching them together by the application of a second layer of quantum channels. The deviation of the constructed state from the target state will depend on the value of QCMI for the partitions with $d(A, C) = O(l)$. An example of Gibbs state construction by the application of layers of universal recovery channels is given in Ref. 21. A simpler, one-dimensional example, in which the circuit consists of two layers of quantum channels, is presented in Ref. 11. Note that in the latter reference the quantum channels proposed are not universal recovery channels. In this work we also avoid use of the universal recovery maps, relying instead on the explicit structure of pgFCS.

Our main result, a bound on the recovery error and the related decay of QCMI, is summarized in Theorem II.1 below. The latter can be compared to and contrasted with the behavior of quantum mutual information (QMI), defined by $I(A : C) = S(\rho_A) + S(\rho_C) - S(\rho_{AC})$. The QMI has the property that it bounds the quantum correlations between the regions $A$ and $C$ via the quantum Pinsker’s inequality.[33] Those pgFCS which correspond to injective MPS[6] exhibit exponential decay of both QMI[34] and correlations.[1, 35] A theoretical bound on QMI can be obtained more straightforwardly than a similar bound on QCMI, since in the former case the separating region $B$ is traced out and information about its state is lost. We note that for pgFCS corresponding to non-injective MPS QMI does not necessarily converge to zero with the growth of region $B$ (see Appendix D for examples). By contrast, we show that for any pgFCS the QCMI converges to zero.

The remainder of this paper is organized around the proof of Theorem II.1, which we state in Section II. In Section III we establish our notation and conventions, and provide relevant background and theorems on FCS, pgFCS, quantum channels and quantum Markov chains, which we use throughout the paper. In Section IV we prove Theorem II.1 starting with the simplest case of pgFCS (corresponding to the injective MPS), and proceed to more general cases. While the proof does not require consideration of specific recovery channels, we give examples of the latter in Section V. In Sections III, IV, and V we rely on additional technical results, presented in a sequence of appendices. We summarize our conclusions in Section VI.
II. MAIN RESULT

In this paper we consider a pgFCS density operator, denoted by $\rho_{ABC}$, for the tripartite system $ABC$ depicted in Figure 1(c). We construct a QMC approximating $\rho_{ABC}$ and a recovery map for which both the trace distance error and the recovery error are exponentially small in the size of the separating region $B$. Using inequality (3), we conclude that QCMI decays exponentially. The methods we use are built upon those of Refs. 1, 6, 34, and 35, which we augment with the continuity theorem for Stinespring’s dilation, [36] cited in Section III B as Theorem III.1.

We now state our principle result,

**Theorem II.1.** Let $\rho_{ABC}$ be the reduced density operator for a pgFCS on a finite contiguous one-dimensional region $ABC$, where the subsystem $B$ separates the subsystems $A$ and $C$. Then, provided that the region $B$ is large enough,

1. There exists a QMC, denoted $\rho_{ABC}$, and constants $q > 0$, $K > 0$, such that
   
   $$\|\rho_{ABC} - \tilde{\rho}_{ABC}\|_1 \leq \frac{1}{2}Ke^{-q|B|}.$$  

2. There exists a quantum channel $\mathcal{B}_{B\rightarrow BC} : \mathcal{B}(\mathcal{H}_B) \rightarrow \mathcal{B}(\mathcal{H}_B \otimes \mathcal{H}_C)$, such that

   $$\|\rho_{ABC} - (\text{id}_A \otimes \mathcal{B}_{B\rightarrow BC})(\rho_{AB})\|_1 \leq Ke^{-q|B|}.$$  

3. Furthermore, the QCMI satisfies $I(A : C|B) \leq K' e^{-q'|B|}$, where $K'$ depends linearly on the sizes of the regions $A$ and $B$.

Remarks:

1. We will show that $q = -\frac{1}{2} \log v$, where $v_\text{gap} < v < 1$ with $v_\text{gap} := \max_{|v_i| < 1} \{|v_i|\}$, where $v_i$ are the eigenvalues of the quantum channel $\mathcal{E}$ induced by the pgFCS under consideration (see Section III C). The value of $v$ may be chosen arbitrarily close to $v_\text{gap}$, with compensating increase in $n_0$, defined in the Remark 2(b) immediately below. See Lemma III.3 proved in Appendix E for details.

2. Let $d_s$ be the dimension of the Hilbert space $\mathcal{H}_s$ of a single spin. Our proof requires that size of the region $B$, denoted by $|B|$, is large enough that $\text{dim } \mathcal{H}_s = d_s^B \geq d_s^M$, where $d_s^M$ is the dimension of the memory space, defined for FCS in Section III C. This demand is quite mild, since $\text{dim } \mathcal{H}_s$ grows exponentially with $|B|$.

3. We may write a bound that does not require a choice of the parameter $v$, however this bound will have a logarithmic correction in the exponent,

   $$\|\rho_{ABC} - (\text{id}_A \otimes \mathcal{B}_{B\rightarrow BC})(\rho_{AB})\|_1 \leq K'e^{-q' |B| + u' \log |B|},$$

   where $q' = -\frac{1}{2} \log v_\text{gap}$, and $K'$ and $u'$ are constants. See Lemma E.2 for details. A similar modification may be done to the bound on QCMI.

4. For the setup of Figure 1(b) we have a result similar to Theorem II.1 with an additional factor of 2 in the pre-exponent and $d(A,C) = \min\{|B_1|, |B_2|\}$ replacing $|B|$. In this case we have to demand that both $B_1$ and $B_2$ are sufficiently large, i.e., $\text{dim } \mathcal{H}_{B_1}, \text{dim } \mathcal{H}_{B_2} \geq d_s^M$. We prove this result in Appendix C.

5. In Theorem II.1, the points 2 and 3 are consequences of the stronger statement 1: there exist approximate QMC’s, which cannot be approximated by (exact) QMC’s. [37–39]

III. PRELIMINARIES

In this section we take the opportunity to introduce some relevant background on quantum channels, finitely correlated states and quantum Markov chains, and to compile and restate
several important results from the proof of Theorem II.1 in Section IV.

A. Notation and conventions

All Hilbert spaces considered in this paper are finite-dimensional and are denoted, up to subscripts, by either $\mathcal{H}$ or $\mathcal{K}$. In particular, we denote the Hilbert space of a single spin by $\mathcal{H}_s$, and denote its dimension by $d_s := \dim \mathcal{H}_s$. We define $\mathcal{B}(\mathcal{H})$ to be the space of linear operators acting on $\mathcal{H}$; the space $\mathcal{B}(\mathcal{H})$ is isomorphic to the space of $\dim \mathcal{H} \times \dim \mathcal{H}$ complex matrices. The set of density operators on the Hilbert space $\mathcal{H}$ is denoted by $\mathcal{D}(\mathcal{H})$. The $n$-fold composition of a quantum channel, for instance $\mathcal{E}$, is denoted by $\mathcal{E}^n$.

To lighten the notation we often omit the identity operator/map in tensor products of operators/maps. For example, for a quantum channel $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K}) \otimes \mathcal{B}(\mathcal{H})$, the quantum channel $\Phi^2 = \Phi \circ \Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K}) \otimes \mathcal{B}(\mathcal{K}) \otimes \mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{H})$ denotes the map $X \mapsto (\id \otimes \Phi) (\Phi(X))$, where $\id$ is the identity map on $\mathcal{B}(\mathcal{H})$. Analogously, if $U : \mathcal{H} \to \mathcal{K}_2 \otimes \mathcal{H}$ and $W : \mathcal{H} \to \mathcal{K}_2 \otimes \mathcal{K}_2$ are isometries, then $UW : \mathcal{H} \to \mathcal{K}_1 \otimes \mathcal{K}_2 \otimes \mathcal{H}$ denotes the isometry $(1_{\mathcal{K}_1} \otimes U)W$, which we can also express in terms of an orthonormal basis of $\mathcal{H}$ as

$$UW := \sum_{i,j=1}^{\dim \mathcal{H}} |i\rangle \langle j| \otimes |i\rangle \langle k| \otimes |j\rangle \langle k|.$$

For system $ABC$ the notation $\rho_{AB}$ means the partial trace of $\rho_{ABC}$ over the Hilbert space of system $C$, i.e., $\rho_{AB} := \tr_C(\rho_{ABC})$. We denote the number of spins in the region $A$ by $|A|$. The support of an operator $O : \mathcal{K} \to \mathcal{H}$, i.e., the closure of the set $\{|\psi\rangle \in \mathcal{H} \mid O(|\psi\rangle) \neq 0\}$, is denoted by $\supp(O)$.

B. Quantum channels

A quantum channel $\mathcal{M}$ is a completely positive trace-preserving map $\mathcal{M} : \mathcal{B}(\mathcal{K}_R) \to \mathcal{B}(\mathcal{K}_R \otimes \mathcal{K}_I)$. The complete positivity means that for operators in the space $\mathcal{B}(\mathcal{K}_R \otimes \mathcal{K}_I)$, where $\mathcal{K}_R$ is a reference Hilbert space of arbitrary dimension, the map $\id_R \otimes \mathcal{M}$ is positive, i.e., for any $\dim \mathcal{K}_R \in \mathbb{N}$ and for any $\mathcal{B}(\mathcal{K}_R \otimes \mathcal{K}_I) \ni X \geq 0$, $(\id_R \otimes \mathcal{M})(X) \geq 0$. For finite-dimensional $\mathcal{K}_I$, as considered here, it suffices to demand that $\id_R \otimes \mathcal{M}$ is positive for the case $\dim \mathcal{K}_R = \dim \mathcal{K}_I$. The trace-preservation implies that $\tr_X(X) = \tr_X$ for any $X \in \mathcal{B}(\mathcal{K}_I)$. For $p, q \in [1, \infty)$ we define the $p \to q$ norm $\| \cdot \|_{p \to q} := \sup_{X \in \mathcal{B}(\mathcal{K}_I)} \| \|X\|_q \|_p$, where $\| \cdot \|_p$ is the Schatten $r$-norm. [28]

By Stinespring’s dilation theorem, [29] in the finite-dimensional setting, a quantum channel $\mathcal{M} : \mathcal{B}(\mathcal{K}_I) \to \mathcal{B}(\mathcal{K}_R \otimes \mathcal{K}_I)$ may be represented in the form [28]

$$\mathcal{M}(X) = \tr_{\mathcal{K}_R}(WXW^\dagger),$$

where $W : \mathcal{K}_I \to \mathcal{K}_E \otimes \mathcal{K}_R$ is an isometry, $W^\dagger W = \id_{\mathcal{K}_R}$, called the dilation isometry, and the Hilbert space $\mathcal{K}_E$ is called the dilation space (or environment). The pair $\{\mathcal{K}_E, W\}$ is referred to as a dilation of $\mathcal{M}$. For brevity, we will write that $W$ dilates the quantum channel $\mathcal{M}$. We note that $\dim \mathcal{K}_E \leq \dim \mathcal{K}_I \dim \mathcal{K}_R$. [28] The isometric representation is not unique, [36] and, if $W : \mathcal{K}_I \to \mathcal{K}_E \otimes \mathcal{K}_R$ and $W' : \mathcal{K}_I \to \mathcal{K}_E' \otimes \mathcal{K}_R$ are different dilations with $\dim \mathcal{K}_E' \geq \dim \mathcal{K}_E$, then $W' = UW$ for some isometry $U : \mathcal{K}_E \to \mathcal{K}_E'$, $U^\dagger U = \id_{\mathcal{K}_E}$. In the case of $\dim \mathcal{K}_E = \dim \mathcal{K}_E'$, the isometry $U$ is unitary.

The space of all quantum channels $\mathcal{M} : \mathcal{B}(\mathcal{K}_I) \to \mathcal{B}(\mathcal{K}_R \otimes \mathcal{K}_I)$ with fixed domain and codomain can be endowed with the so-called diamond norm, [40] also called completely bounded norm, $\| \mathcal{M} \|_\diamond := \sup_{\dim X_R} \| \id_R \otimes \mathcal{M} \|_{1 \to 1}$. In the finite-dimensional case the supremum is achieved for $\dim \mathcal{K}_R = \dim \mathcal{K}_I$. [40] The quantum channels are continuous with respect to the diamond norm, a result that is expressed by the Continuity Theorem for Stinespring’s Representation.

**Theorem III.1** (Ref. [36]. Let $\mathcal{K}_I$ and $\mathcal{K}_R$ be finite-dimensional Hilbert spaces, and suppose that $\mathcal{M}_1, \mathcal{M}_2 : \mathcal{B}(\mathcal{K}_I) \to \mathcal{B}(\mathcal{K}_R \otimes \mathcal{K}_I)$ are quantum channels with Stinespring isometries $W_1, W_2 : \mathcal{K}_I \to \mathcal{K}_E \otimes \mathcal{K}_R$ and a common dilation space $\mathcal{K}_E$. We then have

$$\inf_{U_E} \| W_1 - U_E W_2 \|_2^2 \leq \| \mathcal{M}_1 - \mathcal{M}_2 \|_\diamond \leq 2 \inf_{U_E} \| W_1 - U_E W_2 \|,$$

(7)

where the minimization is with respect to all unitary $U_E \in \mathcal{B}(\mathcal{K}_E)$.

We denoted by $\| W \|$ the operator norm of $W$. In the case $\mathcal{M}_1 = \mathcal{M}_2$, we obtain $W_1 = U_E W_2$ for some unitary $U_E$, which implies that the dilating isometry is defined up to a unitary on the dilation space, consistent with (6).

For a quantum channel $\mathcal{M}$, regarded as a linear map on the $\dim \mathcal{K}_I \dim \mathcal{K}_R$-dimensional vector space $\mathcal{B}(\mathcal{K}_I)$, we define the spectrum in the usual way: $X \in \mathcal{B}(\mathcal{K}_I)$ is an eigenvector with the eigenvalue $\nu$, if $X$ is non-zero and $\mathcal{M}(X) = \nu X$. Each eigenvalue $\nu_i (i = 1, \cdots, \dim \mathcal{K}_I)$ of $\mathcal{M}$ satisfies $|\nu_i| \leq 1$, and there always exists at least one fixed point with the eigenvalue equal to 1. [41] Hence the spectral radius of a quantum channel is equal to 1. The set of eigenvalues with $|\nu_i| = 1$ is known as the **peripheral spectrum**. [1] When the peripheral spectrum is a singleton set, consisting of the single eigenvalue 1, it is said to be trivial.

We now employ the spectral decomposition to construct a quantum channel $\mathcal{E}$ from a given quantum channel $\mathcal{E}'$. [42] This will be useful because that the $n$-fold composition $\mathcal{E}^n$ is an excellent approximation to $\mathcal{E}^n$ in the limit $n \to \infty$. [42] A dilating isometry associated to $\mathcal{E}^n$ possesses convenient properties, and we will take advantage of these.

As shown in Appendix F, the space of all maps $\mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is isomorphic to $\mathcal{B}(\mathcal{H} \otimes \mathcal{H})$.

**Definition 1.** Let $\mathcal{H}$ be a finite-dimensional Hilbert space, and $\mathcal{E} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ be a map. Let $E \in \mathcal{B}(\mathcal{H} \otimes \mathcal{H})$ be the operator isomorphic to $\mathcal{E}$, $E := \text{vec}(\mathcal{E})$. Let $E$ have the Jordan decomposition $E = \sum_i (V_i P_i + N_i V_i)$, with the $P_i$ orthogonal projectors onto the subspace of the eigenvalue $V_i$, and the $N_i$ nilpotent operators. Define $F_\mathcal{E} := \sum_{|\nu| = 1} P_i$, the
projector onto the peripheral spectrum of $\mathcal{E}$, and $\tilde{E} := P_\mathcal{E}E = \sum_{|v_i|=1} (v_i P_{v_i} + N_{v_i})$. Then we define the map $\tilde{\mathcal{E}} := \text{vec}^{-1}(\tilde{E}) = \text{vec}^{-1}(P_{\mathcal{E}} \cdot \text{vec}(\mathcal{E}))$.

For $\mathcal{E}$ and $\tilde{\mathcal{E}}$ as given in Definition 1 we have,

**Lemma III.2 (Ref. 42).** If $\mathcal{E}$ is a quantum channel, then $\tilde{\mathcal{E}}$ is a quantum channel.

**Proof.** One may argue (see the proof of Proposition 3.3 in Ref. 1, for example) that the nilpotents $N_{v_i} = 0$ for the peripheral eigenvalues $v_i$ satisfying $|v_i| = 1$. Moreover, the peripheral eigenvalues form a cyclic group under multiplication, hence there exists $m \in \mathbb{N}$, such that $E^m = E$, and $\tilde{E}_m = \tilde{E}$. Then for any $n \in \mathbb{N}$, $\tilde{E}_n^{m(m-1)+1} = \tilde{E}$. As Lemma III.3 below assures, the sequence $\{\tilde{E}_n - \tilde{E}\}_{n \in \mathbb{N}}$ converges to zero in 2-2 norm (and in diamond norm, since all the spaces are finite-dimensional), hence the subsequence $\{E_n^{m(m-1)+1} - \tilde{E}_n^{m(m-1)+1}\}_{n \in \mathbb{N}}$ converges to zero as well. Then $\{E_n^{m(m-1)+1}\}_{n \in \mathbb{N}}$ converges to $\tilde{E}$. For any $X \in \mathcal{B}(\mathcal{H})$, $\text{tr}(E_n^{m(m-1)+1}(X)) = \text{tr}(X)$, hence $\text{tr}(\tilde{E}(X)) = \text{tr}(X)$, thus $\tilde{\mathcal{E}}$ is trace-preserving. To show that $\tilde{\mathcal{E}}$ is completely positive, it suffices to prove that for any $X \geq 0$ in $\mathcal{B}(\mathcal{H} \otimes \mathcal{H})$, the condition $(\text{id} \otimes \mathcal{E})(X) \geq 0$ is satisfied. Select $X \geq 0$ in $\mathcal{B}(\mathcal{H} \otimes \mathcal{H})$, and note that $(\text{id} \otimes E_n^{m(m-1)+1})(X) \geq 0$. Then $\{\text{id} \otimes E_n^{m(m-1)+1}\}_{n \in \mathbb{N}}$ converges to $\text{id} \otimes \tilde{E}$. Since the set of positive semidefinite operators on a finite-dimensional Hilbert space is closed, then $(\text{id} \otimes \tilde{E})(X) = \lim_{n \to \infty}(\text{id} \otimes E_n^{m(m-1)+1})(X) \geq 0$, completing the proof. \qed

The following lemma gathers results on convergence from Refs. 1 and 42. For the sake of completeness, we provide its proof in our notation in Appendix E.

**Lemma III.3 (Theorem III.2 of Ref. 42; Ref. 1).** Let $\mathcal{E} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ be a map with the spectral radius 1, and let $\tilde{\mathcal{E}}$ be the map obtained from $\mathcal{E}$ as described in Definition 1. Then for any $v \in \mathbb{R}$, such that $V_{\text{gap}} < v < 1$, with $V_{\text{gap}}$ defined in Theorem II.1, there exists the constant $c > 0$, depending on $v$, such that

$$
\|E^n - \tilde{E}\|_{2 \to 2} \leq cV^n.
$$

**C. Finitely correlated states**

FCS are a special class of translationally invariant quantum states on a chain of identical finite dimensional quantum systems (spins). The structure of FCS was characterized in Ref. 1, and is summarized here in the form convenient for our purposes. We refer the reader to the original papers Refs. 1 and 2 for the unabridged treatment of FCS.

A FCS can be described by the pair of a full-rank density operator $\sigma \in \mathcal{B}(\mathcal{H}_s)$ and a quantum channel $\Phi : \mathcal{B}(\mathcal{H}_s) \to \mathcal{B}(\mathcal{H}_s)$, satisfying a compatibility condition $\text{tr}_\mathcal{I}(\Phi(\sigma)) = \sigma$. [1] Here $\mathcal{H}_s$ and $\mathcal{H}_s$ are Hilbert spaces with dimensions $d_m$ and $d_s$, respectively. The space $\mathcal{H}_s$ is referred to as the memory space, and $\mathcal{H}_s$ is the Hilbert space of a single spin. The FCS reduced density operator for a continuous region $R$ with $|R|$ spins, is generated by $\Phi$ and $\sigma$ as

$$
\rho_R = \text{tr}_M(\Phi^{(|R|)}(\sigma)).
$$

Here $\text{tr}_M$ is a partial trace over the Hilbert space $\mathcal{H}_M$; it should not be confused with the trace over the spins in the chain outside the region $R$. Colloquially, each $\Phi$ generates a single spin, and a composition of $\Phi$ generates consecutive spins in the chain. Note that the pair of $\Phi$ and $\sigma$ generating $\rho_R$ is not necessarily unique, and using the results of Refs. 1 and 2 we may choose the most convenient representation for our purposes. In this subsection we list all the relevant properties of the representation we choose.

For the FCS $(\Phi, \sigma)$, we define the induced quantum channel $\mathcal{E} : \mathcal{B}(\mathcal{H}_s) \to \mathcal{B}(\mathcal{H}_s)$ by

$$
\mathcal{E}(X) = \text{tr}_s(\Phi(X)),
$$

for which $\sigma$ is a fixed point, $\mathcal{E}(\sigma) = \sigma$.

1. Ergodic FCS

We consider a subcollection of FCS, called ergodic FCS. By definition, an ergodic state is extremal in the convex set of translationally invariant states. However, in the context of this paper, two other equivalent definitions (see Proposition 3.1 of Ref. 1 and Lemma 4.1 of Ref. 41) will be more useful: (i) the state for which the eigenvalue 1 of $\mathcal{E}$ is non-degenerate; (ii) the state for which $\mathcal{E}$ is irreducible in the sense that there does not exist a non-trivial projector $\Pi$, such that $\Pi \mathcal{B}(\mathcal{H}_s) \Pi$ is invariant under $\mathcal{E}$. The importance of ergodic states lies in the fact that any FCS can be decomposed into a convex sum of ergodic FCS, which are also referred to as ergodic components (Corollary 3.2 of Ref. 1). We collect together various observations that lead to this conclusion in the proposition below. The item 1 follows from the Theorem 3.1 of Ref. 41 applied to $\mathcal{E}$ and $\sigma$. The items 2 and 3 are obtained using the same reasoning as in the proof of Proposition 3.3 of Ref. 1.

**Proposition 1.** Without loss of generality, we may assume that a FCS is generated by a pair $(\Phi, \sigma)$, with the induced quantum channel $\mathcal{E}$, for which the following properties hold

1. There exist $J$ orthogonal projectors $\Pi_{\alpha} : \mathcal{H}_M \to \mathcal{H}_M$, $\sum_{\alpha=1}^{J} \Pi_{\alpha} = \mathcal{I}_M$, such that $\mathcal{E}(\Pi_{\alpha} \mathcal{B}(\mathcal{H}_s) \Pi_{\alpha}) \subseteq \Pi_{\alpha} \mathcal{B}(\mathcal{H}_s) \Pi_{\alpha}$. The restriction of $\mathcal{E}$ to $\mathcal{B}(\mathcal{H}_s) \Pi_{\alpha}$ is irreducible.

2. The density operator $\sigma$ is block-diagonal with respect to the decomposition $\mathcal{H}_M = \bigoplus_{\alpha=1}^{J} \Pi_{\alpha} \mathcal{H}_M$, i.e., $\sigma = \bigoplus_{\alpha=1}^{J} \Pi_{\alpha} \sigma \Pi_{\alpha}$.

3. $\Phi(\Pi_{\alpha} \mathcal{B}(\mathcal{H}_s) \Pi_{\alpha}) \subseteq \mathcal{B}(\mathcal{H}_s) \bigoplus \Pi_{\alpha} \mathcal{B}(\mathcal{H}_s) \Pi_{\alpha}$.

4. The restriction of $\mathcal{E}$ to $\Pi_{\alpha} \mathcal{B}(\mathcal{H}_s) \Pi_{\alpha}$ has a non-degenerate eigenvalue equal to 1, corresponding to the fixed point $\Pi_{\alpha} \sigma \Pi_{\alpha}$. 
The decomposition of FCS into ergodic components follows from items 2 and 3 of Proposition 1,

$$
\rho_R = \sum_{\alpha=1}^{J} \lambda_\alpha \rho_R^\alpha,
$$

(10)

where \( \rho_R^\alpha = \text{tr}_M \Phi^\alpha(\sigma_\alpha) \), \( \sigma_\alpha = \Pi_\alpha \sigma \Pi_\alpha / \text{tr}(\Pi_\alpha \sigma \Pi_\alpha) \), and \( \lambda_\alpha = \text{tr}(\Pi_\alpha \sigma \Pi_\alpha) \). We observe that each ergodic component is manifestly a FCS. One can think of ergodic components being, in the sense of items 2 and 3, independent of each other.

The structure of ergodic components can be analyzed further. We summarize their properties in the proposition below, which is a restatement of Proposition 3.3 of Ref. 1 in the language of density operators.

**Proposition 2** (Proposition 3.3, Ref. 1). For an ergodic FCS, generated by \((\Phi, \sigma)\), with the induced quantum channel \(\sigma\), we may assume that the following properties hold

1. The peripheral spectrum of \(\sigma\) consists of \(p \in \mathbb{N}\) non-degenerate eigenvalues \(\{\exp(\pm ik)\} \mid k = 0, \ldots, p-1\}\). Here \(p\) is referred to as the period of the state.

2. There exist \(p\) orthogonal projectors \(\Pi_k : \mathcal{H}_\mathcal{M} \rightarrow \mathcal{H}_\mathcal{M} \), \(\sum_{k=0}^{p-1} \Pi_k = 1_M\), such that \(\sigma(\Pi_k \mathcal{B}(\mathcal{H}_\mathcal{M}) \Pi_k) \subseteq \Pi_{k+1} \mathcal{B}(\mathcal{H}_\mathcal{M}) \Pi_{k+1}\) with the convention \(\Pi_0 := \Pi_0\), which leads to \(\sigma^p(\Pi_k \mathcal{B}(\mathcal{H}_\mathcal{M}) \Pi_k) \subseteq \Pi_k \mathcal{B}(\mathcal{H}_\mathcal{M}) \Pi_k\).

3. The density operator \(\sigma\) is block-diagonal with respect to the decomposition \(\mathcal{H}_\mathcal{M} = \bigoplus_{k=0}^{p-1} \Pi_k \mathcal{H}_\mathcal{M}, \) i.e., \(\sigma = \bigoplus_{k=0}^{p-1} \Pi_k \sigma \Pi_k\). Moreover \(\sigma(\Pi_k \sigma \Pi_k) = \Pi_{k+1} \sigma \Pi_{k+1}\), which leads to \(\text{tr}(\Pi_k \sigma \Pi_k) = 1/p\) for any \(k\).

4. The restriction of \(\sigma^p\) to \(\Pi_k \mathcal{B}(\mathcal{H}_\mathcal{M}) \Pi_k\) has a trivial peripheral spectrum with \(\Pi_k \sigma \Pi_k\) the fixed point.

We observe similarities between the statements of Proposition 1 and Proposition 2. In particular, item 1 of Proposition 1 implies that the algebra of operators on the memory space, \(\mathcal{B}(\mathcal{H}_\mathcal{M})\), contains \(J\) orthogonal subalgebras invariant under \(\sigma\). By item 2 of Proposition 2, each of these subalgebras is supported on the memory space \(\Pi_\alpha \mathcal{H}_\mathcal{M}\) of a corresponding ergodic component, and further contains \(p\) orthogonal subalgebras that are cyclically permuted by \(\sigma\).

2. Purely Generated FCS

A subcollection of FCS, called purely generated FCS, is characterized by a pure quantum channel \(\Phi\), of the form \(\Phi(X) = XV \dagger\), where \(V : \mathcal{H}_\mathcal{M} \rightarrow \mathcal{H}_A \otimes \mathcal{H}_\mathcal{M}\) is an isometry. \([1, 2]\) We will say that pgFCS is induced, or generated, by the pair \((V, \sigma)\). Equivalently these states are characterized by vanishing mean entropy, i.e., for a continuous region \(R\) of \(|R|\) spins,

$$
\lim_{|R| \rightarrow +\infty} S(\rho_R) / |R| = 0.
$$

[2]

In the language of MPS, a pgFCS generated by \((V, \sigma)\) corresponds to the reduced density operator of the properly defined translationally invariant limit of the MPS \(|\Psi\rangle\) as \(n \rightarrow +\infty\) \([6]\) (see Appendix A for the diagrammatic illustration),

$$
|\Psi\rangle = \sum_{s=0}^{d^n} \langle s_{n-1} \cdots s_1 s_0 | R \rangle 
\times \langle s_{n-1} \cdots s_1 s_0 | \rangle,
$$

where \(M^s\) is a \(d_M \times d_M\) complex matrix, defined by its elements,

$$
M^s_n := (\langle s \mid \langle i | V | j \rangle),
$$

and \(|L\rangle, |R\rangle \in \mathbb{C}^{d_M}\) are boundary factors. In the limit of the infinite chain \(\text{tr}_M\) and \(\sigma\) play a similar role as the left and right boundary factors \(|L\rangle\) and \(|R\rangle\), respectively.

For the case of pgFCS the induced channel \(\sigma\) defined in (9) becomes

$$
\sigma(X) = \text{tr}_A (V XV \dagger).
$$

To formulate the study of recoverability and QMI we subdivide a finite region \(R\) into continuous adjacent subregions \(A, B,\) and \(C\), as discussed in the introduction. A pgFCS defined on this region may be expressed as

$$
\rho_{ABC} = \text{tr}_M \left( V_C V_B V_A \sigma V_1^\dagger \sigma V_2^\dagger \right).
$$

(12)

The isometry \(V_A : \mathcal{H}_\mathcal{M} \rightarrow \mathcal{H}_A \otimes \mathcal{H}_\mathcal{M}\), generating all spins in the region \(A\), is the \(|A|\)-fold product of \(V\), \(V_A := \cdots V \cdot V\), in the sense of (5). The isometries \(V_B\) and \(V_C\) generate all spins in the regions \(B\) and \(C\), respectively, and are defined in an analogous manner. The products of these isometries are also defined by (5).

Note that while ergodic FCS and pgFCS form distinct subsets of FCS, their intersection, the ergodic pgFCS, is non-empty and plays an important role in this study. A pgFCS can be decomposed into a convex sum (10) of ergodic pgFCS \(\rho_{ABC}^\alpha\), as shown in Lemma III.4 below, \([1, 2]\]

$$
\rho_{ABC} = \sum_{\alpha=1}^{J} \lambda_\alpha \rho_{ABC}^\alpha.
$$

(13)

**Lemma III.4.** A FCS is a pgFCS if and only if all its ergodic components are pgFCS.

**Proof.** As was shown in Ref. 2, a state is a pgFCS if and only if it has vanishing mean entropy. Note the well-known bounds on the convexity of the quantum entropy,

$$
\sum_{\alpha=1}^{J} \lambda_\alpha S(\rho_{ABC}^\alpha) \leq S(\rho_{ABC}) \leq \sum_{\alpha=1}^{J} \lambda_\alpha S(\rho_{ABC}^\alpha) - \sum_{\alpha=1}^{J} \lambda_\alpha \ln \lambda_\alpha.
$$

Since

$$
\lim_{\|ABC\| \rightarrow +\infty} \sum_{\alpha=1}^{J} \lambda_\alpha \ln \lambda_\alpha / |ABC| = 0,
$$

then

$$
\lim_{\|ABC\| \rightarrow +\infty} S(\rho_{ABC}) / |ABC| = \sum_{\alpha=1}^{J} \lambda_\alpha \lim_{\|ABC\| \rightarrow +\infty} S(\rho_{ABC}^\alpha) / |ABC|.
$$


If there is a decomposition without loss of generality, we may assume, without loss of generality, that every ergodic component is a pgFCS.

Since each ergodic component is a pgFCS, we may express $\rho_{ABC}^\alpha$ in the form

$$\rho_{ABC}^\alpha = \text{tr}_{M_\alpha} \left( V_\alpha^a V_B^a V_\alpha^b V_B^b V_\alpha^c V_C^c \right),$$

(14)

where all the objects are defined as in (12), and each $\sigma_\alpha$ is defined on a distinct memory space $\mathcal{H}_{M_\alpha}$, $\alpha = 1, \cdots, J$. Naturally, we denote the elementary isometry inducing the ergodic component by $V_\alpha$, i.e., $V_\alpha^a$ is the $|\alpha|$-fold product $V_\alpha^a := V_\alpha V_\alpha \cdots V_\alpha$ in the sense of (5). Since $\rho_{ABC}^\alpha$ is ergodic, the corresponding quantum channel $\mathcal{E}_\alpha : \mathcal{B}(\mathcal{H}_{M_\alpha}) \to \mathcal{B}(\mathcal{H}_{M_\alpha})$, defined as in (11), $\mathcal{E}_\alpha(X) = \text{tr}_\alpha (V_\alpha XV_\alpha^\dagger)$, and $\sigma_\alpha$ have the properties listed in Proposition 2. Now we rewrite Proposition 1 for the case of pgFCS, making a connection between $(V, \sigma)$ generating the pgFCS $\rho_{ABC}$ and the collection of $(V_\alpha, \sigma_\alpha)$, $\alpha = 1, \cdots, J$, generating its ergodic pgFCS components $\rho_{ABC}^\alpha$.

We add an additional property, item 5, whose interpretation is that ergodic components are not proportional to each other; in Appendix G we show that we may assume, without loss of generality, that this property holds.

**Proposition 3 (Ref. 1).** Without loss of generality, we may assume that a pgFCS is generated by a pair $(V, \sigma)$, with the induced quantum channel $\mathcal{E}$, and the following properties hold.

1. $\mathcal{H}_M = \bigoplus_{\alpha=1}^J \mathcal{H}_{M_\alpha}$, where $J$ is the number of ergodic components in $\rho_{ABC}$, and $\mathcal{H}_{M_\alpha} := \Pi_\alpha \mathcal{H}_M$.

2. $V = \sum_{\alpha=1}^J V_\alpha \Pi_\alpha$, where $\Pi_\alpha$, $\alpha = 1, \cdots, J$, is the orthogonal projector onto $\mathcal{H}_{M_\alpha}$, and $V_\alpha : \mathcal{H}_M \to \mathcal{H}_S \otimes \mathcal{H}_{M_\alpha}$ is an isometry.

3. $\sigma = \bigoplus_{\alpha=1}^J \lambda_\alpha \sigma_\alpha$, where $\sigma_\alpha = \Pi_\alpha \sigma \Pi_\alpha / \text{tr} (\Pi_\alpha \sigma \Pi_\alpha)$, $\lambda_\alpha = \text{tr} (\Pi_\alpha \sigma \Pi_\alpha)$.

4. $(V_\alpha, \sigma_\alpha)$ generates an ergodic pgFCS in the form (14), with the induced quantum channel $\mathcal{E}_\alpha(X) = \mathcal{E}(\Pi_\alpha X \Pi_\alpha) = \text{tr}_\alpha (V_\alpha XV_\alpha^\dagger)$, for which Proposition 2 applies.

5. For $\alpha \neq \beta$ there is no unitary $U : \mathcal{H}_M \to \mathcal{H}_M$ and $\phi \in \mathbb{R}$, such that $V_\alpha = e^{i\phi} UV_\beta U^\dagger$.

**Remarks:**

1. Proposition 3 contains all the pgFCS properties that we use in this paper.

2. The decomposition (13) into ergodic components is not necessarily the finest one that can be performed on a pgFCS – each ergodic component may be further decomposed into a sum of ergodic pgFCS components of period 1. [1] The resulting components may, however, involve adjustment of the memory Hilbert space, and blocking of spins – the procedure under which several consecutive spins are treated as one. We prefer to deal directly with the decomposition of a pgFCS into ergodic pgFCS components of arbitrary period. In Section IV it will be convenient to consider an ergodic pgFCS of period 1 first, in order to establish the arguments in our proof.

**D. Quantum Markov chains**

Quantum Markov chains are defined as those states for which QCMI vanishes, $I(A : C|B) := S(\rho_{AB}) + S(\rho_{BC}) - S(\rho_{ABC}) - S(\rho_B) = 0$, [16] and are fully characterized by the following theorem.

**Theorem III.5 (Ref. 16).** Let $\rho_{ABC} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$. The following three statements are equivalent:

1. $\rho_{ABC}$ is a quantum Markov chain, i.e., $I(A : C|B) = 0$.

2. There exists a quantum channel $\mathcal{B}_{B \to BC} : \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_B \otimes \mathcal{H}_C)$, such that $\rho_{ABC} = \mathcal{B}_{B \to BC}(\rho_{AB})$. On $B(\text{supp}(\rho_B))$ this channel can be taken to be the Petz recovery map $\mathcal{B}_{B \to BC}(X) := \rho_B^\dagger \rho_B^{-1} X \rho_B^{-1} \rho_B^\dagger$, although this choice is not unique.

3. There is a decomposition $\text{supp}(\rho_B) \cong \bigoplus_{k=1}^{\max_k} \mathcal{H}_{B_k} \otimes \mathcal{H}_{C_k}$, i.e., there exists a unitary isomorphism $I : \text{supp}(\rho_B) \cong \bigoplus_{k=1}^{\max_k} \mathcal{H}_{B_k} \otimes \mathcal{H}_{C_k}$, such that $I \rho_{ABC} I^\dagger = \bigoplus_{k=1}^{\max_k} \lambda_k \rho_{ABC}^{B_k} \otimes \rho_{ABC}^{C_k}$, where $\rho_{ABC}^{B_k} \in \mathcal{D}(\mathcal{H}_{B_k} \otimes \mathcal{H}_{B_k})$, $\rho_{ABC}^{C_k} \in \mathcal{D}(\mathcal{H}_{C_k} \otimes \mathcal{H}_{C_k})$, $\lambda_k > 0$ and $\sum_{k=1}^{\max_k} \lambda_k = 1$.

For a pgFCS $\rho_{ABC}$ we will construct an approximating state $\tilde{\rho}_{ABC}$, and show that it is a quantum Markov chain by proving that it satisfies the property 3, whence properties 1 and 2 as well.

The theorem has a useful corollary, whose proof is omitted, which we will employ in the next section.

**Corollary III.5.1.** If $\rho_{ABC} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ and $\rho_{ABC} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ are quantum Markov chains, then $\lambda \rho_{ABC} + (1 - \lambda) \tilde{\rho}_{ABC}$, where $0 \leq \lambda \leq 1$, is a quantum Markov chain on $\mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$.

**IV. PROOF OF THEOREM II.1**

In this section we prove the main result, Theorem II.1. The proof relies on two steps: (i) approximating pgFCS $\rho_{ABC}$ by another state $\tilde{\rho}_{ABC}$ and (ii) the proof that $\tilde{\rho}_{ABC}$ is a quantum Markov chain. Conveniently, the approximation step is the same for any pgFCS, and we outline it immediately below.

The proof that the constructed $\tilde{\rho}_{BC}$ is a quantum Markov chain, proceeds in several steps, which we develop in subsequent subsections. We first consider the simplest case of an ergodic pgFCS of period 1, before generalizing to arbitrary period $p$. Finally, by taking convex sums, we deal with a general case of pgFCS.

We now deal with the approximation step, assuming that $\tilde{\rho}_{ABC}$ is a quantum Markov chain. This is sufficient to prove the Theorem II.1.
Any pgFCS has the form (12),
\[ \rho_{ABC} = \text{tr}_M \left( V_C V_B V_A \sigma V_A^\dagger V_B^\dagger V_C^\dagger \right). \]
We introduce an isometry \( \tilde{V}_B : \mathcal{H}_A \otimes \mathcal{H}_M \to \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_M \) which acts trivially on the space \( \mathcal{H}_A \), and use it to construct the state
\[ \tilde{\rho}_{ABC} = \text{tr}_M \left( V_C \tilde{V}_B V_A \sigma V_A^\dagger \tilde{V}_B^\dagger V_C^\dagger \right). \] (15)

Note that the isometry \( \tilde{V}_B \) generates all the spins in the region \( B \), but unlike \( V_B \), it is not a concatenation of isometric factors generating one spin at a time. It is a key step to choose \( \tilde{V}_B \) to be of a particular form that ensures the state \( \tilde{\rho}_{ABC} \) is simultaneously close (in trace norm distance) to \( \rho_{ABC} \) and is an exact quantum Markov chain.

The trace distance \( \| \rho_{ABC} - \tilde{\rho}_{ABC} \|_1 \) is bounded by twice the error in approximating \( V_B \) by \( \tilde{V}_B \) in operator norm,
\[ \| \rho_{ABC} - \tilde{\rho}_{ABC} \|_1 = \| \text{tr}_M \left( V_C V_B V_A \sigma V_A^\dagger V_B^\dagger V_C^\dagger \right) \|
\leq \| V_B V_A \sigma V_A^\dagger \tilde{V}_B^\dagger V_C^\dagger \|
\leq 2 \| V_B - \tilde{V}_B \|. \] (16)

where in the last line we have used the inequalities \( \| X \|_1 \leq \| X Y \|_1 + \| \tilde{X} Z \|_1 \leq 2 \| X - Y \| \| Z \|_1 \), and the fact that \( \| V_A \sigma V_A^\dagger \|_1 = 1 \).

For \( \tilde{\rho}_{ABC} \) a quantum Markov chain, according to condition 2 of Theorem III.5 there exists a recovery channel \( \mathcal{R}_{B \to BC} \), such that \( \mathcal{R}_{B \to BC}(\tilde{\rho}_{AB}) = \rho_{ABC} \). We may select \( \mathcal{R}_{B \to BC} \) to be the Petz recovery map, defined in statement 2 of Theorem III.5, or one of the alternatives that we discuss in Section V. If we use such an \( \mathcal{R}_{B \to BC} \) to approximately recover \( \rho_{ABC} \) from \( \rho_{AB} \), then the recovery error is given by
\[ \| \rho_{ABC} - \mathcal{R}_{B \to BC}(\rho_{AB}) \|_1 = \| \rho_{ABC} - \mathcal{R}_{B \to BC} \circ \text{tr}_C(\rho_{ABC}) \|_1
\leq \| \rho_{ABC} - \tilde{\rho}_{ABC} \|_1 + \| \mathcal{R}_{B \to BC} \circ \text{tr}_C(\tilde{\rho}_{ABC} - \tilde{\rho}_{ABC}) \|_1
\leq 2 \| \rho_{ABC} - \tilde{\rho}_{ABC} \|_1 \] (17)

where we have used the quantum Markov property, the triangle inequality, and the property that the trace norm is non-increasing under actions of quantum channels.

Combining (16) and (17), we obtain a bound on the recovery error,
\[ \| \rho_{ABC} - \mathcal{R}_{B \to BC}(\rho_{AB}) \|_1 \leq 4 \| V_B - \tilde{V}_B \|. \] (18)

Our goal is to choose the isometry \( V_B \) so that \( \| V_B - \tilde{V}_B \| \) is exponentially small in the size of the region \( B \). Consider the \( |B| \)-fold composition of \( \mathcal{E} \), defined in (11), \( \mathcal{E}^{[B]} \). Since \( \mathcal{E} \) is dilated by the elementary isometry \( V \), then the \( |B| \)-fold composition of \( V \), that is \( V_B \), dilates \( \mathcal{E}^{[B]} \). From \( \mathcal{E} \) we construct the quantum channel \( \mathcal{E} \) via Definition 1, and then approximate \( \mathcal{E}^{[B]} \) by \( \mathcal{E}^{[B]} \), the \( |B| \)-fold composition of \( \mathcal{E} \). As discussed in Section III.B, the maximum dimension of the dilating space of \( \mathcal{E}^{[B]} \) is equal to \( \dim \mathcal{H}_B = d_M^2 \). We will assume that \( \dim \mathcal{H}_B \geq d_M^2 \), so that the dilation space of \( \mathcal{E}^{[B]} \) can be embedded into \( \mathcal{H}_B \). Under this condition, we define \( \tilde{V}_B \) to be an isometry dilating \( \mathcal{E}^{[B]} \). Note that any isometry of the form \( U_B \tilde{V}_B \), where \( U_B : \mathcal{H}_B \to \mathcal{H}_B \) is unitary, will also be a dilating isometry for \( \mathcal{E}^{[B]} \). We will use this freedom of choice to minimize \( \| V_B - \tilde{V}_B \| \).

By the continuity of Stinespring’s dilatation, [36] we can relate the optimum operator norm distance between \( V_B \) and \( \tilde{V}_B \) to the distance between channels \( \mathcal{E}^{[B]} \) and \( \mathcal{E}^{[B]} \) through the inequalities
\[ \inf_{U_B} \| V_B - U_B \tilde{V}_B \|^2 \leq \| \mathcal{E}^{[B]} - \mathcal{E}^{[B]} \|_\diamond \leq 2 \inf_{U_B} \| V_B - U_B \tilde{V}_B \|, \] (19)

where the infimum is taken over the set of all unitaries \( U_B : \mathcal{H}_B \to \mathcal{H}_B \). (Recall that the diamond norm is defined by \( \| \mathcal{E} - \mathcal{E} \|_\diamond := \sup_{\rho \in \mathcal{H}} \| \mathcal{E} \rho - \mathcal{E} \rho \|_{1 \to 1} \). As the dimension of \( \mathcal{H}_B \) is finite, the infimum is attained for some \( U_B \), which we take to be the identity, \( U_B = 1_B \), so that our choice of \( \tilde{V}_B \) is optimal. Using the first inequality in (19), we bound the distance between the isometries as
\[ \| V_B - \tilde{V}_B \| \leq \sqrt{\| \mathcal{E}^{[B]} - \mathcal{E}^{[B]} \|_\diamond}. \] (20)

To bound the diamond norm \( \| \mathcal{E}^{[B]} - \mathcal{E}^{[B]} \|_\diamond \), we take advantage of the 2 to 2 norm, which is easier to estimate. The 2 to 2 norm is defined by \( \| \mathcal{E} \|_{2 \to 2} := \sup_{|X|_2 = 1} \| \mathcal{E}(X) \|_2 \), where the 2-norm, or Hilbert-Schmidt norm, is given by \( \| X \|_2 = \sqrt{\text{tr}(X^* X)} \). Since the Hilbert space \( \mathcal{H}_B \) is finite dimensional, the supremum \( \sup_{|X|_2 = 1} \| \mathcal{E} \rho - \mathcal{E} \rho \|_{1 \to 1} \) is achieved when \( n = d_M, \) hence \( \| \mathcal{E}^{[B]} - \mathcal{E}^{[B]} \|_\diamond = \| \mathcal{E}^{[B]} - \mathcal{E}^{[B]} \|_{2 \to 2} \) (tensoring with the identity map does not change the 2 to 2 norm), and bounding the rank with the space dimension \( \text{rank}(X) \leq d_M \), we obtain the well-known estimate
\[ \| \mathcal{E}^{[B]} - \mathcal{E}^{[B]} \|_{2 \to 2} \leq d_M \| \mathcal{E}^{[B]} - \mathcal{E}^{[B]} \|_{2 \to 2}. \] (21)

From Lemma III.3 applied to the quantum channels \( \mathcal{E} \) and \( \mathcal{E} \) it follows that,
\[ \| \mathcal{E}^{[B]} - \mathcal{E}^{[B]} \|_{2 \to 2} \leq c \| \mathcal{E}^{[B]} \|^2, \] (22)

with \( c > 0 \) and \( v_{\text{gap}} \leq v < 1 \), where \( v_{\text{gap}} := \max_{|V|_2 < 1} \{ |V|_2 \} \), and \( v \) are the eigenvalues of \( \mathcal{E} \). Combining together (20), (21), and (22), we obtain
\[ \| V_B - \tilde{V}_B \| \leq \sqrt{d_M} \| \mathcal{E}^{[B]} \|^2, \] (23)

and, from (18),
\[ \| \rho_{ABC} - \tilde{\rho}_{ABC} \|_1 \leq 2 \sqrt{d_M} \| \mathcal{E}^{[B]} \|^2 \] (24)

Observe that statements 1 and 2 of Theorem II.1 hold with the constants \( K = 4\sqrt{d_M} \) and \( q = \ln (v^{-1}) / 2 \). The bound (24) is meaningful while \( 4\sqrt{d_M} \| \mathcal{E}^{[B]} \|^2 \leq 2 \), since the trace distance between any two states cannot exceed 2. The statement
3 of the theorem follows from the relation (B1) in Appendix B with $K = \sqrt{d_M} \frac{\ln d_b}{2} + 2 \ln(\sqrt{d_M} v_B^{1/2})$. Notice that $\ln(2\sqrt{d_M} v_B^{1/2}) \leq 0$ when the bound (24) is useful. This completes the proof.

It remains to prove that $\hat{\rho}_{ABC}$ given in (15), and defined in terms of the isometry $\check{V}_B$, is a quantum Markov chain.

A. Ergodic pgFCS of period 1

We first prove that $\hat{\rho}_{ABC}$ is a quantum Markov chain for an ergodic pgFCS with $p = 1$. Since $\xi$ has a trivial peripheral spectrum and $\sigma$ is its fixed point, then $\check{\delta}$ constructed according to Definition 1 sends any input into the state proportional to $\sigma$, $\check{\delta}(X) = tr(X)\sigma$. It follows that for any $n \in \mathbb{N}$, $\check{\delta}^n = \check{\delta}$, in particular $\check{\delta}^{|B|} = \check{\delta}$. We observe that the isometry $\check{V}_B = \sum_{i=1}^{d_M} \sqrt{\sigma}|\xi_i \rangle \otimes |i \rangle \langle i|$ is a dilatation of $\check{\delta}^{|B|} = \check{\delta}$, where we have introduced an arbitrary orthonormal set of vectors $\{|\xi_i \rangle \}$ in $\mathcal{H}_B | i = 1, \cdots, d_M \rangle$. The vectors $|i \rangle \in \mathcal{H}_M$ and $\sigma_i > 0$ are the eigenvectors and corresponding eigenvalues of $\sigma$, respectively, i.e., $\sigma = \sum_{i=1}^{d_M} \sigma_i |i \rangle \langle i|$. Indeed,

$$
tr_B(\check{V}_B \check{V}_B^\dagger) = \sum_{i,j,i',j'} \sqrt{\sigma_i \sigma_{i'}} \langle \xi_{i'} \xi_i \rangle \langle j | X | j' \rangle \langle i | i' \rangle
$$

$$
= \sum_{j=1}^{d_M} \langle j | X | j \rangle \sum_{i=1}^{d_M} \sigma_i |i \rangle \langle i| = tr(X)\sigma.
$$

Since $\check{V}_B$ dilates $\check{\delta}$, which is also dilated by $V_B$, there is unitary $U_B : \mathcal{H} \rightarrow \mathcal{H}_B$, such that $\check{V}_B = U_B \check{V}_B$, and we observe that this amounts to a unitary change of the orthonormal basis $\{|\xi_i \rangle \}$. Thus, considering the freedom of choice of the basis vectors $|\xi_i \rangle$, we can assume without loss of generality that $\check{V}_B = V_B$,

$$
\check{V}_B = \sum_{i=1}^{d_M} \sqrt{\sigma_i} |\xi_i \rangle \otimes |i \rangle \langle i|.
$$

Substituting (25) into (15), $\hat{\rho}_{ABC}$ may be written explicitly,

$$
\hat{\rho}_{ABC} = \sum_{i,i',j,j'} \langle j | V_A \sigma V_A^\dagger | j' \rangle \otimes \sigma (V_B) \otimes \sigma (V_C^\dagger).
$$

(26)

We will show by direct computation that $\hat{\rho}_{ABC}$ satisfies the condition 3 of Theorem III.5 and hence is a quantum Markov chain. We denote $\mathcal{H}_b := \text{span}\{|\xi_i \rangle | i = 1, \cdots, d_M \rangle\}$, and note that since their dimensions are equal, $\mathcal{H}_b$ and $\mathcal{H}_M \otimes \mathcal{H}_M$ are isomorphic. We construct the unitary isomorphism $I : \mathcal{H}_b \rightarrow \mathcal{H}_M \otimes \mathcal{H}_M$, defined by $I(\xi_i) = |i \rangle \otimes |i \rangle$. Under this isomorphism, a straightforward calculation gives,

$$
I \hat{\rho}_{ABC} I^\dagger = V_A \sigma V_A^\dagger \otimes tr_M \left(V_C \sigma V_C^\dagger \right),
$$

where $|\rangle \rangle = \sum_{i=1}^{d_M} |i \rangle \otimes |i \rangle \in \mathcal{H}_M \otimes \mathcal{H}_M$ is a non-normalized, maximally entangled state vector, and the partial trace $tr_M$ is over the second factor. If we write $\mathcal{H}_b = \mathcal{H}_b \otimes \mathcal{H}_b$, where $\mathcal{H}_b$ and $\mathcal{H}_b$ are isomorphic to $\mathcal{H}_M$, we observe that $\hat{\rho}_{ABC}$ satisfies condition 3 of Theorem III.5 with $\text{supp} (\rho_b) \equiv \mathcal{H}_b = \mathcal{H}_b \otimes \mathcal{H}_b$.

$$
I \hat{\rho}_{ABC} I^\dagger = \hat{\rho}_{ABC} \otimes \hat{\rho}_{B'},
$$

$$
\hat{\rho}_{ABC} = V_A \sigma V_A^\dagger \otimes tr_M \left(V_C \sigma V_C^\dagger \right),
$$

and is thus a quantum Markov chain.

B. Ergodic pgFCS of period $p$

For an ergodic pgFCS with $p \geq 1$, Proposition 2 states that the peripheral spectrum of $\xi$ contains $p$ eigenvalues, and thus for $p > 1$ the proof that $\hat{\rho}_{ABC}$ is a quantum Markov chain must be modified.

The properties of $\xi$ listed in Proposition 2 allow us to prove Lemma IV.1 below, stating that (assuming dim $\mathcal{H}_b \geq d_M$) $\check{V}_B$ can be decomposed into $\check{V}_B = \sum_{k=0}^{p-1} \check{V}_k$, where each $\check{V}_k : \mathcal{H}_M \rightarrow \mathcal{H}_b \otimes \mathcal{H}_M$ is a partial isometry with $\text{supp}(\check{V}_k) = \mathcal{K}_K \mathcal{H}_M$, and $\check{V}_k \check{V}_k^\dagger = \Pi_k$. The range of $\check{V}_k$ satisfies

$$
\text{Range}(\check{V}_k) \subseteq \mathcal{K}_K \mathcal{K}_K | \otimes | \text{Range}(\check{V}_k) \subseteq \mathcal{H}_b \otimes \mathcal{H}_M,
$$

where the subspaces $\mathcal{K}_K \subseteq \mathcal{K}_K$ are mutually orthogonal. Below, each $\check{V}_k$ will be seen to have the form (28) similar to (25). The operator $\sigma$ is block diagonal with respect to the set of all projectors $\Pi_k$, $k = 0, \cdots, p - 1$.

Anticipating the validity of Lemma IV.1, below, we may express $\hat{\rho}_{ABC}$ in the form

$$
\hat{\rho}_{ABC} = tr_M \left(V_C \left(\sum_{k=0}^{p-1} \check{V}_k \check{V}_k^\dagger \right) V_A \sigma V_A^\dagger \left(\sum_{k=0}^{p-1} \check{V}_k \check{V}_k^\dagger \right) V_C^\dagger \right),
$$

$$
= \sum_{k=0}^{p-1} tr_M \left(V_C \check{V}_k V_A \sigma V_A^\dagger \check{V}_k^\dagger \right),
$$

(27)
where in the first line we substitute the decomposition of $\tilde{V}_B$ into partial isometries and make the replacement $V_A^\dagger V_A = \sum_{k=1}^p \Pi_k V_A^\dagger \Pi_k$ (see Appendix H for proof), and $\tilde{V}_B^k \Pi_k = \delta_{k,H} \tilde{V}_B^k$ collapses the expression to a single sum in the second line. Since the ranges of $\tilde{V}_B^k$ are mutually orthogonal, the summands corresponding to different $k$ are supported on orthogonal subspaces, so that the sum is direct.

Each term in the direct sum can be expressed as in (26), and by the arguments of Section IV A is shown to be (up to normalization) a quantum Markov chain. Hence by Corollary III.5.1, $\beta_{ABC}$ is also a quantum Markov chain. Thus we can use the bound for the recovery error (24), and the statement of Theorem II.1 holds with the same constants $q, K$ and $K$ as in Section IV A.

We now prove that $\tilde{V}_B$ has the required decomposition.

**Lemma IV.1.** For an ergodic pgFCS of period $p$, the isometry $\tilde{V}_B : \mathcal{H}_M \to \mathcal{H}_B \otimes \mathcal{H}_B$ defined in (15) can be constructed as the sum $\tilde{V}_B = \sum_{k=0}^{p-1} \tilde{V}_B^k$, where

1. $\tilde{V}_B^k$ is a partial isometry, $\tilde{V}_B^k \tilde{V}_B^k = \Pi_k$, and $\tilde{V}_B^k \Pi_k = \delta_{k,H} \tilde{V}_B^k$.

2. The range of $\tilde{V}_B^k$ satisfies $\text{Range}(\tilde{V}_B^k) \subseteq \mathcal{H}_B \otimes \Pi_k \mathcal{H}_M \subseteq \mathcal{H}_B \otimes \mathcal{H}_M$. The subspaces $\mathcal{H}_B^k$ are mutually orthogonal.

3. The density operator $p \text{tr}_M(V_C^\dagger V_A^\dagger V_A^* V_B^k V_C)\sigma_A V_A^\dagger$ is a quantum Markov chain.

Remarks:

1. Explicitly, each $\tilde{V}_B^k$, $k = 0, \cdots, p-1$, has the form

$$\tilde{V}_B^k = \sqrt{p} \sum_{(i,j) \in \mathcal{E}_k} \sqrt{\sigma_i} |\xi_{ij}\rangle \otimes |i\rangle \langle j|,$$

where $\{|\xi_{ij}\rangle \in \mathcal{H}_B \ | \ (i,j) \in \mathcal{E}_k \}$ is a set of orthonormal vectors, $(\xi_{ij}, |\xi_{ij}\rangle) = \delta_{i,j},\delta_{ij}$, the vectors $|i\rangle, |j\rangle \in \mathcal{H}_M$ and the eigenvalues $\sigma_i > 0$ are defined by the spectral decomposition $\sigma = \sum_{i=1}^d \sigma_i |i\rangle \langle i|$. Notice that if $p > 1$, then the set $\cup_{k=0}^{p-1} \mathcal{E}_k$ does not contain all pairs $(i,j)$.

2. The subspaces $\mathcal{H}_B^k$ are defined as $\mathcal{H}_B^k := \text{span}\{|\xi_{ij}\rangle \in \mathcal{H}_B \ | \ (i,j) \in \mathcal{E}_k \}$. It follows that the subspaces $\mathcal{H}_B^k$ are mutually orthogonal, and implies that for $X \in \mathcal{B}(\mathcal{H}_M)$ and $k \neq k'$, the sum $\tilde{V}_B^k \widetilde{X} \tilde{V}_B^{k'} + \tilde{V}_B^{k'} \widetilde{X} \tilde{V}_B^k$ is direct, i.e., $\tilde{V}_B^k \widetilde{X} \tilde{V}_B^{k'} \in \mathcal{B}(\mathcal{H}_B^k) \otimes \mathcal{B}(\mathcal{H}_M)$ and $\tilde{V}_B^{k'} \widetilde{X} \tilde{V}_B^k \in \mathcal{B}(\mathcal{H}_B^{k'}) \otimes \mathcal{B}(\mathcal{H}_M)$ and $\tilde{V}_B^k \widetilde{X} \tilde{V}_B^{k'} + \tilde{V}_B^{k'} \widetilde{X} \tilde{V}_B^k \in \mathcal{B}(\mathcal{H}_B) \otimes \mathcal{B}(\mathcal{H}_M)$.

**Proof.** As we show in Appendix I, the quantum channel $\tilde{\sigma}$, obtained via Definition I from $\tilde{\sigma}$ with the properties listed in Proposition 2, has the form

$$\tilde{\sigma}(X) = p \sum_{k=0}^{p-1} \text{tr}(\Pi_k X \Pi_k) \Pi_{k+1} \sigma \Pi_{k+1}.$$  \hfill (29)

In $\Pi_{k+1}$ the addition in the subscript is interpreted modulo $p$. Then,

$$\tilde{\sigma}^{[B]}(X) = p \sum_{k=0}^{p-1} \text{tr}(\Pi_k X \Pi_k) \Pi_{k+1}[B] \sigma \Pi_{k+1}[B].$$

Notice that for each $k$ and $X \in \mathcal{B}(\mathcal{H}_M)$ the map $X \mapsto p \text{tr}(\Pi_k X \Pi_k) \Pi_{k+1}[B] \sigma \Pi_{k+1}[B]$, while not trace-preserving, is completely positive, and has an associated dilation $\tilde{V}_B^k : \mathcal{H}_M \to \mathcal{H}_B \otimes \mathcal{H}_M$, defined in (28).

It is easy to check that $\tilde{V}_B^k$ is a partial isometry,

$$\tilde{V}_B^k \tilde{V}_B^k = \sum_{(j) \in \mathcal{E}_k} |j\rangle \langle j| = \Pi_k.$$

We define the subspace $\mathcal{H}_B^k := \text{span}\{|\xi_{ij}\rangle \in \mathcal{H}_B \ | \ (i,j) \in \mathcal{E}_k \}$ that has the dimension $\dim \mathcal{H}_B^k = \text{rank}(\Pi_k) \text{rank}(\Pi_k[B])$.

Now we prove that the assumption $\dim \mathcal{H}_B^k \geq d_B^2$ is consistent with mutual orthogonality of the subspaces $\mathcal{H}_B^k$. For $\mathcal{H}_B^k$ to contain $\sum_{k=0}^{p-1} \mathcal{H}_B^k$, we require $\sum_{k=0}^{p-1} \text{rank}(\Pi_k) \text{rank}(\Pi_k[B]) \leq \sum_{k=0}^{p-1} (\text{rank}(\Pi_k))^2 = d_B^2$, this condition is satisfied.

The channel $\tilde{\sigma}^{[B]}$ has the associated dilation isometry $\tilde{V}_B^k := \sum_{k=0}^{p-1} \tilde{V}_B^k$, since

$$\tilde{V}_B^k \tilde{V}_B^k = \sum_{k=0}^{p-1} \tilde{V}_B^k \tilde{V}_B^k = \sum_{k=0}^{p-1} \tilde{V}_B^k \tilde{V}_B^k = \sum_{k=0}^{p-1} \Pi_k = 1_B,$$

and

$$\text{tr}_B(\tilde{V}_B^k X \tilde{V}_B^k) = \sum_{k=0}^{p-1} \text{tr}_B(\tilde{V}_B^k X \tilde{V}_B^k) = \sum_{k=0}^{p-1} \text{tr}_B(\tilde{V}_B^k X \tilde{V}_B^k) = \sum_{k=0}^{p-1} \text{tr}(\Pi_k X \Pi_k) \Pi_{k+1}[B] \sigma \Pi_{k+1}[B] = \tilde{\sigma}^{[B]}(X).$$

We have used mutual orthogonality of $\mathcal{H}_B^k$ in both calculations. Since both $\tilde{V}_B^k, \tilde{V}_B^k : \mathcal{H}_M \to \mathcal{H}_B \otimes \mathcal{H}_M$ dilate $\tilde{\sigma}^{[B]}$, there exists unitary $U_B : \mathcal{H}_B \to \mathcal{H}_B$, such that $\tilde{V}_B^k = U_B \tilde{V}_B^k$. Using the same reasoning as in Section IV A we can, without loss of generality, take $\tilde{V}_B^k = \tilde{V}_B$.

$$\tilde{V}_B = \sum_{k=0}^{p-1} \tilde{V}_B^k = \sqrt{p} \sum_{k=0}^{p-1} \sum_{(i,j) \in \mathcal{E}_k} |\sqrt{\sigma_i} \langle \xi_{ij}| \otimes |i\rangle \langle j|.$$ \hfill (30)

Notice that the case $p = 1$ corresponds to a single projector $\Pi_1 = 1_M$, so that (29) reduces to $\tilde{\sigma}^{[B]}(X) = \text{tr}(X) \sigma$, and (30) reduces to (25).

Comparing (28) to (25), we realize that the we can use the same arguments as in Section IV A to prove that $p \text{tr}_M(V_C^\dagger V_A^\dagger V_A^* V_B^k V_C)$ is a quantum Markov chain. The
for a pgFCS, the isometry
\[ \tilde{\delta} \] separately approximates the corresponding states that the isometry \( \tilde{\rho} \) completes the proof.

### C. The general case of pgFCS

We generalize the proof to the case of a convex sum of ergodic pgFCS considered in the previous subsection. The properties of pgFCS are given in Proposition 3, whose notation and definitions we employ below.

We will prove that \( \tilde{\rho}_{ABC} \), as defined in (15), is a quantum Markov chain by anticipating the Lemma IV.2, which states that the isometry \( \tilde{V}_B \) can be decomposed as a sum of partial isometries corresponding to ergodic components, \( \tilde{V}_B = \sum_{\alpha=1}^{J} \tilde{V}_B^\alpha \), where \( \tilde{V}_B^\alpha : \mathcal{H}_B \to \mathcal{H}_B \otimes \mathcal{H}_M \) is a partial isometry with \( \Pi_\alpha \mathcal{H}_M \) the orthogonal complement of its kernel, and where \( \text{Range}(\tilde{V}_B^\alpha) = \tilde{\mathcal{H}}_\alpha \otimes \Pi_\alpha \mathcal{H}_M \), with subspaces \( \tilde{\mathcal{H}}_\alpha \subseteq \mathcal{H}_B \) being mutually orthogonal. The projectors \( \Pi_\alpha \) are defined in Proposition 3, and correspond to different ergodic components of \( \rho_{ABC} \). Moreover, \( \tilde{V}_B^\alpha \) are such that

\[
\tilde{\rho}_{ABC} := \tilde{\rho}_M(\bigoplus_{\alpha=1}^{J} \tilde{V}_B^\alpha V_B^\dagger A^\sigma A V_B^\dagger C)
\]

separately approximates the corresponding \( \rho_{ABC}^\alpha \) defined in Proposition 3. Then \( \tilde{\rho}_{ABC} \) may be decomposed,

\[
\tilde{\rho}_{ABC} = \tilde{\rho}_M(\bigoplus_{\alpha=1}^{J} \tilde{V}_B^\alpha V_B^\dagger A^\sigma A V_B^\dagger C)
\]

where we have used \( V_C \tilde{V}_B^\alpha = V_C^\dagger \tilde{V}_B^\alpha \) and \( \tilde{V}_B^\alpha V_B^\dagger A^\sigma A V_B^\dagger C = \delta_{\alpha \beta} \tilde{\lambda}_\alpha \tilde{\lambda}_\beta \tilde{V}_B^\alpha V_B^\dagger A^\sigma A V_B^\dagger C \). By the proof in Section IV.2, each \( \tilde{\rho}_{ABC}^\alpha \) is a quantum Markov chain, and by Corollary III.5.1 \( \tilde{\rho}_{ABC} \) is a quantum Markov chain. This is sufficient to complete the proof.

The following lemma shows that \( \tilde{V}_B \) indeed possesses the requisite decomposition. We repeatedly use Proposition 3 in the proof.

### Lemma IV.2

For a pgFCS, the isometry \( \tilde{V}_B : \mathcal{H}_M \to \mathcal{H}_B \otimes \mathcal{H}_M \) defined in (15) can be constructed as the sum \( \tilde{V}_B = \sum_{J=1}^{J} \tilde{V}_B^\alpha \), where \( J \) is the number of ergodic components, and

1. \( \tilde{V}_B^\alpha \) is a partial isometry, \( \tilde{V}_B^\alpha \tilde{V}_B^\dagger = \Pi_\alpha \), and \( \tilde{V}_B^\alpha \Pi_\beta = \delta_{\alpha \beta} \tilde{V}_B^\dagger \).

2. \( \text{Range}(\tilde{V}_B^\alpha) = \tilde{\mathcal{H}}_\alpha \otimes \Pi_\alpha \mathcal{H}_M \subseteq \mathcal{H}_B \otimes \mathcal{H}_M \). The subspaces \( \tilde{\mathcal{H}}_\alpha \) are mutually orthogonal.

3. \( \tilde{\rho}_{ABC}^\alpha = \tilde{\rho}_M(V_C^\alpha \tilde{V}_B^\alpha V_B^\dagger A^\sigma A V_B^\dagger C) \) is a quantum Markov chain.

**Proof.** We express \( \mathcal{E} \) as

\[
\mathcal{E}(X) = \sum_{\alpha, \beta = 1}^{J} \mathcal{E}(\Pi_\alpha X \Pi_\beta) = \sum_{\alpha, \beta = 1}^{J} \text{tr}(V_B \Pi_\alpha X \Pi_\beta \tilde{V}_B^\dagger),
\]

where we have used \( V_B \Pi_\alpha = V_B \Pi_\alpha \). Since the range of \( V_B \) is \( \mathcal{H}_B \otimes \Pi_\alpha \mathcal{H}_M \), then \( \text{tr}(V_B \Pi_\alpha X \Pi_\beta \tilde{V}_B^\dagger) \in \Pi_\alpha \mathcal{D} (\mathcal{H}_M) \Pi_\beta \). We decompose \( \mathcal{E} \) similarly,

\[
\mathcal{E}(X) = \sum_{\alpha, \beta = 1}^{J} \mathcal{E}(\Pi_\alpha X \Pi_\beta).
\]

We now show that the last equation reduces to the direct sum, \( \mathcal{E}(X) = \bigoplus_{\alpha = 1}^{J} \mathcal{E}(\Pi_\alpha X \Pi_\alpha) \). We observe that the map, \( X \to \mathcal{E}(\Pi_\alpha X \Pi_\alpha) \), is obtained via Definition 1 from the map \( X \to \text{tr} (V_B \Pi_\alpha X \Pi_\beta \tilde{V}_B^\dagger) \). Recall that for \( \alpha \neq \beta \) there is no unitary : \( \mathcal{H}_B \to \mathcal{H}_M \) and \( \phi \in \mathbb{R} \), such that \( V_B = e^{i\phi} UV_B \), which implies by Lemma E.1 that the eigenvalues of \( X \to \text{tr}(V_B \Pi_\alpha X \Pi_\beta \tilde{V}_B^\dagger) \) have magnitudes less than 1. Then, following the construction in Definition 1, the map \( X \to \text{tr}(V_B \Pi_\alpha X \Pi_\beta \tilde{V}_B^\dagger) \) is identically zero, \( \mathcal{E}(\Pi_\alpha X \Pi_\beta) = 0 \) for \( \alpha \neq \beta \). Thus, \( \mathcal{E}(X) = \sum_{\alpha = 1}^{J} \mathcal{E}(\Pi_\alpha X \Pi_\alpha) \). Since \( \mathcal{E}(\Pi_\alpha X \Pi_\alpha) \in \Pi_\alpha \mathcal{D} (\mathcal{H}_M) \Pi_\alpha \), from Definition 1 follows that \( \mathcal{E}(\Pi_\alpha X \Pi_\alpha) \in \Pi_\alpha \mathcal{D} (\mathcal{H}_M) \Pi_\alpha \). Since \( \Pi_\alpha \) are orthonormal projectors, the sum is direct, \( \mathcal{E}(X) = \bigoplus_{\alpha = 1}^{J} \mathcal{E}(\Pi_\alpha X \Pi_\alpha) \), and

\[
\mathcal{E}(X) = \bigoplus_{\alpha = 1}^{J} \mathcal{E}(\Pi_\alpha X \Pi_\alpha).
\]

We define \( V_B^\alpha : \mathcal{H}_M \to \tilde{\mathcal{H}}_\alpha \otimes \mathcal{H}_M \) to be an isometric dilation of \( X \to \mathcal{E}(\Pi_\alpha X \Pi_\alpha) \). Here \( \mathcal{H}_M := \Pi_\alpha \mathcal{H}_M \), and \( \tilde{\mathcal{H}}_\alpha \) is the dilation space embedded in \( \mathcal{H}_B \) of dimension \( \dim \tilde{\mathcal{H}}_\alpha \leq (\dim \mathcal{H}_M)^2 = \tilde{d}^2 \). It immediately follows that \( V_B^\alpha \tilde{V}_B^\alpha = \Pi_\alpha \), proving the statement 1. From property 4 of Proposition 3 we know that \( (\tilde{V}_B^\alpha, \sigma_A) \) generates an ergodic pgFCS with the induced quantum channel \( X \to \mathcal{E}(\Pi_\alpha X \Pi_\alpha) \). As mentioned before, \( X \to \mathcal{E}(\Pi_\alpha X \Pi_\alpha) \) is exactly the map obtained via Definition 1 from the map \( X \to \mathcal{E}(\Pi_\alpha X \Pi_\alpha) \). Then, separately for each \( \alpha \), we can repeat the reasoning of Section IV.2, showing that each \( \tilde{V}_B^\alpha \), the dilating isometry of \( X \to \mathcal{E}(\Pi_\alpha X \Pi_\alpha) \), possesses the properties which ensure that \( \tilde{\rho}_{ABC} \) is a quantum Markov chain, proving the statement 3.
Now we show that we can safely choose the subspaces $H_{b_k}$ to be mutually orthogonal. Since $H_M = \bigoplus_{a=1}^d H_{Ma}$, and since we assume $\dim H_B \geq d_B^4$, then $\dim H_B \geq (\sum_{a=1}^d d_{Ma})^2 \geq \sum_{a=1}^d d_{Ma}^2 \geq \sum_{a=1}^d \dim H_{ba_a}$, which implies that the subspaces $H_{ba_a}$ may be chosen to be orthogonal, proving the statement 2. Then we can construct the isometry $\tilde{V}_B = \sum_{a=1}^d \tilde{V}_B^a$. This isometry clearly dilates $\mathcal{E}(B)$, which is also dilated by $\tilde{V}_B$. By the same reasoning as in Section IV A and Section IV B we may take $\tilde{V}_B = \tilde{V}_B$.

Thus, $\tilde{V}_B$ possesses all the required properties. \hfill $\Box$

V. RECOVERY MAPS

In this section we consider the recovery maps that can be used to restore $\tilde{\rho}_{ABC}$, defined in (15), from $\rho_{AB}$. The obvious choice is the Petz recovery map, defined on $\mathcal{B}(\text{supp}(\tilde{\rho}_B))$ in (2), $\mathcal{B}_{B\rightarrow BC} (X) = \frac{1}{\text{tr}_{BC} \tilde{\rho}_B} \tilde{\rho}_B \frac{1}{\tilde{\rho}_B} X \tilde{\rho}_B \frac{1}{\tilde{\rho}_B}$, and which can be extended by a trivial embedding $\text{supp}(\tilde{\rho}_B)^{\perp} \rightarrow \text{supp}(\tilde{\rho}_B)^{\perp} \otimes H_C$ to the map $\mathcal{B}_{B\rightarrow BC} : \mathcal{B}(H_B) \rightarrow \mathcal{B}(H_B) \otimes \mathcal{B}(H_C)$.[14–16]

The Petz recovery map can be defined for any state that is a quantum Markov chain.

The Petz recovery map is, however, not the only map that can exactly reconstruct a quantum Markov chain $\tilde{\rho}_{ABC}$. From Theorem III.5, we note that there is another obvious choice of the recovery channel, which relies on the isomorphism $I : \text{supp}(\tilde{\rho}_B) \rightarrow \bigoplus_{k=1}^{d_B^{10}} H_{b_k} \otimes H_{a_k}$ defined in Theorem III.5, and which provides the decomposition $I \tilde{\rho}_{ABC} I^\dagger = \bigoplus_{k=1}^{d_B^{10}} \lambda_k \tilde{\rho}_{Ab_k} \otimes \tilde{\rho}_{b_k}$. We gave an example of such an isomorphism in Section IV A for the case of an ergodic pgFCS of period 1. Using $I$, for a general quantum Markov chain, we can construct the map $\mathcal{B}_{B\rightarrow BC} : \mathcal{B}(H_B) \rightarrow \mathcal{B}(H_B \otimes H_C)$,

$$\mathcal{B}_{B\rightarrow BC} (X) := I^\dagger \sum_{k=1}^{d_B^{10}} \text{tr}_{b_k} (\mathcal{P}_k (IXI^\dagger) \mathcal{P}_k) \otimes \tilde{\rho}_{b_k} C I,$$

where $\mathcal{P}_k$ is a projector onto $H_{b_k} \otimes H_{a_k}$. We verify that this map is indeed a recovery channel by direct computation,

$$\mathcal{B}_{B\rightarrow BC} (\tilde{\rho}_{AB}) = I^\dagger \sum_{k=1}^{d_B^{10}} \text{tr}_{b_k} (\mathcal{P}_k (I \tilde{\rho}_{AB} I^\dagger) \mathcal{P}_k) \otimes \tilde{\rho}_{b_k} C I$$

$$= I^\dagger \sum_{k=1}^{d_B^{10}} \text{tr}_{b_k} (\mathcal{P}_k (\bigoplus_{k'=1}^{d_B^{10}} \lambda_{k'} \tilde{\rho}_{Ab_{k'}} \otimes \tilde{\rho}_{b_{k'}}) \mathcal{P}_k) \otimes \tilde{\rho}_{b_k} C I$$

$$= I^\dagger \bigoplus_{k'=1}^{d_B^{10}} \lambda_{k'} \text{tr}_{b_{k'}} (\tilde{\rho}_{Ab_{k'}} \otimes \tilde{\rho}_{b_{k'}}) \otimes \tilde{\rho}_{b_k} C I$$

$$= I^\dagger \bigoplus_{k'=1}^{d_B^{10}} \lambda_{k'} \tilde{\rho}_{Ab_{k'}} \otimes \tilde{\rho}_{b_{k'}} C I$$

$$= \tilde{\rho}_{ABC},$$

where we have used $I \tilde{\rho}_{AB} I^\dagger = I \rho_{ABC} I^\dagger = \text{tr}_C (I \tilde{\rho}_{ABC} I^\dagger)$ (recall that $I$ acts non-trivially only on $H_B$), the structure of the quantum Markov chain from Theorem III.5 in the second line, and $\mathcal{P}_k \tilde{\rho}_{Ab_{k'}} \otimes \tilde{\rho}_{b_{k'}} C \mathcal{P}_k = \delta_{k,k'} \tilde{\rho}_{Ab_{k'}} \otimes \tilde{\rho}_{b_{k'}} C$ in the third line.

Unlike the Petz recovery map that is defined in terms of the reduced density operators $\tilde{\rho}_{BC}$ and $\tilde{\rho}_{B}$, the recovery map $\mathcal{B}_{B\rightarrow BC}$ requires knowledge of the isomorphism $I$, the structure of which is not obvious in the case of a general quantum Markov chain. In the case of a quantum Markov chain approximating a pgFCS, and possessing the form (15), we can construct the isomorphism $I$. For simplicity of argument, we illustrate this claim using the case of ergodic pgFCS of period 1. Recall that $\tilde{\rho}_{ABC}$ has the form (15), and $\tilde{V}_B$ has the form (25), $\tilde{V}_B = \sum_{j=1}^{d_B} |\tilde{\xi}_{ij}| \otimes |i\rangle \langle j|$, where $|i\rangle$ and $|j\rangle$ are the eigenvectors of $\sigma$, and $|\tilde{\xi}_{ij}|$ is an orthonormal basis set. This allows us to express $\tilde{\rho}_{ABC}$ as in (26) and define $I$ in terms of the orthonormal vectors $|\tilde{\xi}_{ij}|$, $I = \sum_{i,j=1}^{d_B} (|i\rangle \otimes |j\rangle) (|\tilde{\xi}_{ij}|)$. The required set of orthonormal vectors $\{|\tilde{\xi}_{ij}|\} \subset H_B \ | i, j = 1, \ldots, d_B \}$ is the solution of the minimization problem for the norm $\|\tilde{V}_B - \tilde{V}_B\|$ over unitary change of $d_B$ basis vectors in $H_B$.

Interestingly, we can construct another recovery channel, that has a slightly less optimal bound in terms of the pre-exponential factor, but for which the vectors $|\tilde{\xi}_{ij}|$ can be constructed more explicitly. Again we use the case of ergodic pgFCS of period 1 for illustration.

For $\tilde{V}_B$ in the form (25), any choice of the $|\tilde{\xi}_{ij}|$ induces an exact quantum Markov chain $\tilde{\rho}_{ABC}$ of the form (15). A particular choice of $|\tilde{\xi}_{ij}|$ influences the bound on the recovery error. Expressing

$$\|\tilde{V}_B - \tilde{V}_B\| = \sum_{i,j=1}^{d_B} (|i\rangle \langle V_B|j\rangle - \sqrt{C} |\tilde{\xi}_{ij}|) \otimes |i\rangle \langle j|,\text{,}$$

we observe that any choice of $|\tilde{\xi}_{ij}|$ close to $\sigma_{ij}^{-1/2} |\tilde{\xi}_{ij}|$ is expected to be a good one. Let us construct the positive semidefinite matrix

$$Q_{ij',ij} := |\langle j'| \tilde{V}_B^\dagger (\mathcal{I}_{B} \otimes |i\rangle \langle i|) \mathcal{V}_B|j\rangle|,$$

which is a Gramian matrix for the set of vectors $\{|i\rangle \mathcal{V}_B|j\rangle : |i,j = 1, \ldots, d_B\}$. For large $|B|$, the vectors in the latter set are almost orthogonal to each other, and $\|\mathcal{I}_{B} \tilde{V}_B|j\rangle\|$ is close to $\sigma_{ij}^{1/2}$. To prove this, notice that the map

$$X \mapsto \tilde{V}_B \tilde{V}_B^\dagger (\mathcal{I}_{B} \otimes X) \mathcal{V}_B$$

with $X \in \mathcal{B}(H_M)$ is adjoint (with respect to the Hilbert-Schmidt inner product) to the quantum channel $\mathcal{E}(|B|) = \mathcal{V}_B X \mathcal{V}_B^\dagger$, which converges to $\tilde{\mathcal{E}}(X) = \text{tr}(X) \sigma$ in the limit $|B| \rightarrow \infty$. Hence the map $X \mapsto \tilde{V}_B \tilde{V}_B^\dagger (\mathcal{I}_{B} \otimes X) \mathcal{V}_B$ converges to the adjoint of $\tilde{\mathcal{E}}$, the map $X \mapsto \text{tr}(\mathcal{E}(X) \mathcal{I}_{B})$. Thus, $Q_{ij',ij}$ converges to $|\langle j'| \mathcal{E} (|\sigma| \mathcal{I}_{B}) |i\rangle \mathcal{I}_{B} \mathcal{V}_B|j\rangle| = \sigma_{ij} \delta_{ij}. \text{ }$

Then we define $|\tilde{\xi}_{ij}|$ as

$$|\tilde{\xi}_{ij}| = \sum_{m,n=1}^{d_B} (Q_{ij}^{-1/2})_{mn,ij} |m\rangle \langle n|.$$
(As we show in Appendix J, \( Q \) is guaranteed to be invertible, if the region \( B \) is large enough). We verify that the set of \( |\xi_{ij}\rangle \) is orthonormal,
\[
\langle \xi_{i'j}'|\xi_{ij}\rangle = \sum_{m,n,m',n'} (Q^{-\frac{1}{2}})_{i'j',m'n'} (m'|V_B|^n) (Q^{-\frac{1}{2}})_{nn'ij} = \sum_{m,n,m',n'} (Q^{-\frac{1}{2}})_{i'j',m'n'} Q_{m'n'm} Q^{-\frac{1}{2}}_{nn'ij} = \delta_{i'i} \delta_{j'j}.
\]

In Appendix J we show that for this choice of \( |\xi_{ij}\rangle \),
\[
||V_B - \tilde{V}_B|| \leq \sqrt{2(1 + 2^{-\frac{1}{2}})c d_M v_{\tilde{\nu}}},
\] (34)
with \( c \) and \( \nu \) defined in Lemma III.3, leading to the recovery error
\[
||R_{B\rightarrow BC}(\rho_{AB}) - \rho_{ABC}||_1 \leq 4 \sqrt{2(1 + 2^{-\frac{1}{2}})c d_M v_{\tilde{\nu}}} m^\nu,
\] (35)
which is worse than the recovery error (24) by the factor
\[
\sqrt{2(1 + 2^{-\frac{1}{2}})d_M}.
\]

In Appendix J we present a similar construction for the general case of pgFCS, which is based on the same idea, but is more delicate.

VI. CONCLUSION

Quantum conditional mutual information (QCMI) is nonnegative by virtue of the strong subadditivity of von Neumann entropy. We have studied the QCMI for a homogeneous quantum spin chain described by a purely generated finitely correlated state (pgFCS) \( \rho_{ABC} \). Explicitly, by separating the chain regions A and C by a domain B, and provided the size \(|B|\) of region B is large enough, we show there exists a quantum Markov chain which differs from \( \rho_{ABC} \) in trace distance by an error exponentially small in \(|B|\). It follows that \( \rho_{ABC} \) may be recovered, within the stated error, from the reduced density operator \( \rho_{AB} \) by a quantum channel acting only on the domain B. As a consequence this implies that the QCMI, denoted \( I(A : C|B) \), is also exponentially small in \(|B|\). We have presented quantum channels that, in principle, can perform the state recovery. Besides the obvious choice of the Petz recovery map, we give two other examples derived from the particular structure of the pgFCS. We have carried out numerical experiments on the decay of QCMI, the results of which are consistent with the bound presented in Theorem II.1. The details will be reported elsewhere.

A natural next step would be to prove the exponential decay of QCMI for generic FCS, as conjectured in Ref. 24. Both the results presented here and in Ref. 24 point in this direction. As discussed further in Appendix K, the pgFCS considered here and the class of states considered in Ref. 24 are in general distinct.

Generalization of our result to generic FCS seems to require an approach different from the one we used in this paper. It remains an open question whether a FCS exhibiting an exponential decay of QCMI is necessarily close to a quantum Markov chain in trace distance.

ACKNOWLEDGMENTS

The authors would like to thank Shivan Mittal for helpful discussions and the referee for helpful suggestions.

AUTHOR DECLARATIONS

Conflict of interests

The authors have no conflicts to disclose.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Appendix A

We present in the form of Table I the dictionary for conversion from the language of pgFCS defined by the pair \((V, \sigma)\) (see Section III.C) to the language of matrix product states (in diagrammatic notation).

Appendix B

In this appendix we derive the right-hand side of the inequality (3). In Refs. 17–19, and 43, the bound is based on the Alicki-Fannes inequality, [30] which is the extension of the Fannes inequality. [44] Here we employ an improved version of the Alicki-Fannes inequality from Ref. 31 (we refer the reader to this paper for further references on the topic).

In order to apply the results of Ref. 31, we first express the QCMI in terms of quantum relative entropy, \( D(\rho||\sigma) := \text{tr} \rho (\ln \rho - \ln \sigma) \),
\[
I(A : C|B) = S(\rho_{BC}) - S(\rho_{ABC}) + S(\rho_{AB}) - S(\rho_B) = D(\rho_{ABC}||\rho_{BC}) - D(\rho_{ABC}||\rho_B).
\]

By the monotonicity of quantum relative entropy under the action of quantum channels, \([28, 45]\)
\[
D(\rho_{B\rightarrow BC}(\rho_{AB})||\rho_{B\rightarrow BC}(\rho_{AB})) \leq D(\rho_{ABC}||\rho_{B}) \text{, thus we can estimate}
\]
\[
I(A : C|B) \leq D(\rho_{ABC}||\rho_{BC}) - D(\rho_{B\rightarrow BC}(\rho_{AB})||\rho_{B\rightarrow BC}(\rho_{AB})).
\]

To make the result of Ref. 31 directly applicable, we recast the above inequality in terms of conditional entropy, \( S(R_1|R_2)_\rho := S(\rho_{R_1|R_2}) - S(\rho_{R_2}) \) (we adopt the notation of Ref. 31). We notice that \( D(\rho_{ABC}||\rho_{BC}) = -S(A|BC)_\rho \) and \( D(\rho_{B\rightarrow BC}(\rho_{AB})||\rho_{B\rightarrow BC}(\rho_{AB})) = -S(A|BC)'_\rho', \) where \( \rho_{ABC} = \rho_{B\rightarrow BC}(\rho_{AB}) \). Then
\[
I(A : C|B) \leq S(A|BC)_\rho - S(A|BC)'_\rho'.
\]
We apply Lemma 2 of Ref. 31 to obtain the bound

\[ I(A : C|B) \leq \epsilon' \ln \dim \mathcal{H}_A + (1 + \epsilon')h \left( \frac{\epsilon'}{1 + \epsilon'} \right), \]

where \( h(p) = -p \ln p - (1 - p) \ln (1 - p) \) is the binary entropy and \( \epsilon' = \frac{1}{2}\|\rho_{ABC} - \rho_{B \rightarrow BC}(\rho_{AB})\|_1 \). This bound, unlike the one that follows from the Alicki-Fannes inequality, is applicable for any \( 0 \leq \epsilon' \leq 1 \). Notice that we expressed the inequality (3) in terms of \( \epsilon = \|\rho_{ABC} - \rho_{B \rightarrow BC}(\rho_{AB})\|_1 = 2\epsilon' \). Reverting to our notation,

\[ I(A : C|B) \leq \frac{1}{2}\epsilon|A|\ln d_A + (1 + \frac{1}{2}\epsilon')h \left( \frac{\epsilon'}{2 + \epsilon'} \right), \]

as presented in (3), where we have substituted \( \dim \mathcal{H}_A = d_A |A| \).

We can also derive the simplified bound by estimating

\[ (1 + \epsilon')h \left( \frac{\epsilon'}{1 + \epsilon'} \right) = -\epsilon' \ln \left( \frac{\epsilon'}{1 + \epsilon'} \right) + \ln (1 + \epsilon') \]
\[ = -\epsilon' \ln \epsilon' + (1 + \epsilon') \ln (1 + \epsilon') \]
\[ \leq \epsilon' (2 - \ln \epsilon'), \]

where we have used \( 0 \leq \epsilon' \leq 1 \) and \( \ln (1 + \epsilon') \leq \epsilon' \). This leads to the bound for the QCMI,

\[ I(A : C|B) \leq \epsilon \left( \frac{1}{2}|A|\ln d_A + 1 - \frac{1}{2} \ln \frac{\epsilon}{2} \right), \] (B1)

where \( 0 < \epsilon < 2 \), which we use in the proof of Theorem II.1.

\[ \text{Appendix C} \]

In this appendix we extend Theorem II.1 for the configuration of the regions \( A, B, \) and \( C \) depicted in Figure 1(b). In this case \( \rho_{ABC} \) has the form

\[ \rho_{ABC} = \text{tr}_M \left( V_{A_2} V_{B_2} V_C V_{A_1} \sigma V_{A_1}^\dagger V_{B_1}^\dagger V_{C}^\dagger V_{B_2}^\dagger V_{C}^\dagger \right). \]

We introduce approximating isometries \( \tilde{V}_{B_1} : \mathcal{H}_M \to \mathcal{H}_{B_1} \otimes \mathcal{H}_M \) and \( \tilde{V}_{B_2} : \mathcal{H}_M \to \mathcal{H}_{B_2} \otimes \mathcal{H}_M \) in exactly the same way as in Section IV, which we use to construct the state

\[ \tilde{\rho}_{ABC} = \text{tr}_M \left( V_{A_2} \tilde{V}_{B_2} V_C \tilde{V}_{B_1} V_{A_1} \sigma V_{A_1}^\dagger \tilde{V}_{B_1}^\dagger V_{C}^\dagger \tilde{V}_{B_2}^\dagger V_{C}^\dagger \right). \] (C1)

Then, repeating the calculations in (17) and (16) and using the triangle inequality, we bound the recovery error for the recovery channel \( \mathcal{R}_{B \rightarrow BC} : \mathcal{B}(\mathcal{H}_B) \to \mathcal{B}(\mathcal{H}_B \otimes \mathcal{H}_C) \), for which

\[ \mathcal{R}_{B \rightarrow BC}(\tilde{\rho}_{AB}) = \tilde{\rho}_{ABC}, \]

\[ \|\rho_{ABC} - \mathcal{R}_{B \rightarrow BC}(\rho_{AB})\|_1 \leq 4\|V_{B_1} - \tilde{V}_{B_1}\| + 4\|V_{B_2} - \tilde{V}_{B_2}\|. \]
For both terms we apply the bound (23), which leads to
\[
\|\rho_{ABC} - R_{B\rightarrow BC}(\rho_{AB})\|_1 \leq 4\sqrt{d_{MC}} (\sqrt{\|B_1\|/2} + \sqrt{\|B_2\|/2}) \\
\leq 8\sqrt{d_{MC}v^d(A,C)/2},
\]
where \(d(A,C) = \min\{|B_1|, |B_2|\} \).

We can prove that \(\rho_{ABC}\) is a quantum Markov chain using the same reasoning as in Sections IV A, IV B, and IV C. We present the proof only for the case of ergodic pgFCS of period 1. The extensions of the argument to the general case of pgFCS are analogous to the ones of Sections IV B and IV C.

Both \(V_{B_1}\) and \(V_{B_2}\) have the form (25),
\[
\begin{align*}
\tilde{V}_{B_1} &= \sum_{i,j=1}^{d_{B_1}} \sqrt{\sigma_i} |\xi_{ij}\rangle \otimes |i⟩⟨j|, \\
\tilde{V}_{B_2} &= \sum_{m,n=1}^{d_{B_2}} \sqrt{\sigma_m} |\zeta_{mn}\rangle \otimes |m⟩⟨n|.
\end{align*}
\]
\]

We introduce the isomorphism \(I : \mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2} \rightarrow \mathcal{H}_M \otimes \mathcal{H}_M \otimes \mathcal{H}_M \otimes \mathcal{H}_M\), defined by \(I(\xi_{ij}) \otimes |\zeta_{mn}\rangle = |j⟩ \otimes |i⟩ \otimes |n⟩ \otimes |m⟩\), which maps \(\tilde{\rho}_{ABC}\) to
\[
I\tilde{\rho}_{ABC}I^\dagger = V_{A_1} \sigma V_{A_1}^\dagger \otimes V_C \sqrt{\sigma} |+\rangle \langle +| + \sqrt{\sigma} V_C^\dagger \\
\otimes \text{tr}_M(V_{A_2} \sqrt{\sigma} |+\rangle \langle +| \sqrt{\sigma} V_{A_2}^\dagger),
\]
where \(|+\rangle = \sum_{i=1}^{d_{B_1}} |i⟩ \otimes |i⟩ \in \mathcal{H}_M \otimes \mathcal{H}_M\). By Proposition III.5 this state is a quantum Markov chain.

### Appendix D

In this appendix we construct explicit examples of pgFCS, for which QMI converges to a finite limit as the size of the region \(B\) is increased, \(\lim_{|B| \rightarrow +\infty} I(A:C) > 0\). By contrast, Theorem II.1 proves that QCM converges to zero for any pgFCS.

To construct the examples we use two isometries. The first one is \(V_1 : \mathbb{C}^2 \rightarrow \mathbb{C}^3 \otimes \mathbb{C}^2\), defined by
\[
V_1 = \frac{1}{\sqrt{2}} |1⟩ \otimes |−⟩ \langle +| - \frac{1}{\sqrt{2}} |0⟩ \otimes |+⟩ \langle +| + \frac{\sqrt{3}}{2} |0⟩ \otimes |−⟩ \langle −| - \frac{1}{\sqrt{2}} |−⟩ \otimes |+⟩ \langle −|,
\]
where \(\{|−⟩, |0⟩, |1⟩\}\) is an orthonormal basis in \(\mathbb{C}^3\), and \(\{|−⟩, |+⟩\}\) is an orthonormal basis in \(\mathbb{C}^2\). This isometry is of the type defined in Ref. 1, that generates the ground state of the generalized AKLT model. We define \(\delta_1 : \mathcal{B}(\mathbb{C}^2) \rightarrow \mathcal{B}(\mathbb{C}^2)\) by \(\delta_1(X) = \text{tr}_3(V_1 XV_1^\dagger)\), where \(\text{tr}_3\) is the partial trace over \(\mathbb{C}^3\). For the current discussion it is important that the pair \((V_1, \mathbb{I}_2)/2\) generates an ergodic pgFCS of period 1. Hence, \(\delta_1\), obtained from \(\delta_1\) using Definition 1 satisfies \(\delta_1(X) = \text{tr}(X)\mathbb{I}_2/2\). We obtain the second isometry, \(V_2 : \mathbb{C}^2 \rightarrow \mathbb{C}^3 \otimes \mathbb{C}^2\), from \(V_1\) by performing the cyclic permutation \(|−⟩ \rightarrow |0⟩ \rightarrow |1⟩ \rightarrow |−⟩\),
\[
V_2 = \frac{1}{\sqrt{2}} |1⟩ \otimes |−⟩ \langle +| - \frac{1}{\sqrt{2}} |0⟩ \otimes |+⟩ \langle +| + \frac{1}{\sqrt{2}} |0⟩ \otimes |−⟩ \langle −| - \frac{1}{\sqrt{2}} |−⟩ \otimes |+⟩ \langle −|.
\]

We observe \(\text{tr}_3(V_2 XV_2^\dagger) = \text{tr}_3(V_1 \mathbb{I}_2 V_1^\dagger) = \delta_1(X)\) for any \(X \in \mathcal{B}(\mathbb{H}_M)\).

Our first example is a pgFCS with two ergodic components. For a Hilbert space of the spin \(\mathcal{H}_s := \mathbb{C}^3\) and a memory Hilbert space \(\mathcal{H}_M := \mathbb{C}^4 \cong \mathbb{C}^2 \otimes \mathbb{C}^2\), the quantum state is generated by
\[
V = \begin{pmatrix} V_1 & 0 \\ 0 & V_2 \end{pmatrix}, \quad \sigma = \frac{1}{4} \begin{pmatrix} \mathbb{I}_2 & 0 \\ 0 & \mathbb{I}_2 \end{pmatrix}.
\]

It is obvious that this state by construction has two ergodic components (see Proposition 3). We choose the regions \(A\) and \(C\) to each consist of a single spin, i.e., \(\mathcal{H}_A, \mathcal{H}_C = \mathbb{C}^2\). In the limit \(|B| \rightarrow +\infty\),
\[
\rho_{AC} := \lim_{|B| \rightarrow +\infty} \text{tr}_M(V \text{tr}_B(V_3 V_3^\dagger) V^\dagger)
\]
\[
= \text{tr}_M(V \delta(V \sigma V^\dagger) V^\dagger)
\]
\[
= \frac{1}{8} \text{tr}_2(V_1 V_1^\dagger) \otimes \text{tr}_2(V_2 V_2^\dagger) + \frac{1}{8} \text{tr}_2(V_2 V_2^\dagger) \otimes \text{tr}_2(V_2 V_2^\dagger),
\]
where \( \text{tr}_2 \) is the partial trace over \( C^2 \). In the second line \( \tilde{\mathcal{E}} \) is the channel obtained from \( \mathcal{E}: C^4 \to C^4 \), \( \mathcal{E}(X) = \text{tr}_3(VXV^\dagger) \), via Definition 1, which acts as
\[
\begin{pmatrix}
X_{11} & X_{12}\\
X_{21} & X_{22}
\end{pmatrix}

\mapsto
\begin{pmatrix}
\tilde{\mathcal{E}}_1(X_{11}) & 0 \\
0 & \tilde{\mathcal{E}}_1(X_{22})
\end{pmatrix}
\begin{pmatrix}
\frac{1}{2\text{tr}_2(X_{11})}\mathbb{1}_2 & 0 \\
0 & \frac{1}{2\text{tr}_2(X_{22})}\mathbb{1}_2
\end{pmatrix}.
\]
Then
\[
\tilde{\rho}_A \otimes \tilde{\rho}_C = \rho_A \otimes \rho_C = \frac{1}{16} \left( \text{tr}_2(V_1V_1^\dagger) + \text{tr}_2(V_2V_2^\dagger) \right)
\]
\[
\otimes \left( \text{tr}_2(V_1V_1^\dagger) + \text{tr}_2(V_2V_2^\dagger) \right).
\]
Using the explicit expressions (D1) and (D2), we calculate
\[
\tilde{\rho}_A \otimes \tilde{\rho}_C = \frac{1}{64} \left( \begin{pmatrix}
1 & 0 \\
0 & 2
\end{pmatrix} \otimes \begin{pmatrix}
1 & 0 \\
0 & 2
\end{pmatrix} \right)
\]
\[
+ \frac{1}{32} \left( \begin{pmatrix}
1 & 0 \\
0 & 2
\end{pmatrix} \otimes \begin{pmatrix}
1 & 0 \\
0 & 2
\end{pmatrix} \right),
\]
\[
\tilde{\rho}_A \otimes \tilde{\rho}_C = \frac{1}{64} \left( \begin{pmatrix}
2 & 0 \\
0 & 3
\end{pmatrix} \otimes \begin{pmatrix}
2 & 0 \\
0 & 3
\end{pmatrix} \right).
\]
We observe that \( \tilde{\rho}_A \neq \tilde{\rho}_C \), which implies that \( I(A : C) \neq 0 \). We verify this by explicitly calculating \( I(A : C) = 17 \ln 2/16 - 9 \ln 3/8 + 5 \ln 5/16 = 0.0035 \).

Now we construct an ergodic pgFCS of period 2 for which QMI does not converge to zero. We use the same setup as above, changing only the elementary isometry generating the state to
\[
V = \begin{pmatrix}
0 & V_1 \\
V_2 & 0
\end{pmatrix},
\]
where \( V_1 \) and \( V_2 \) are the same isometries (D1) and (D2), respectively. One can check that the only eigenvalues of absolute value 1 of \( \mathcal{E}(X) = \text{tr}_3(VXV^\dagger) \) are 1 and \(-1\), corresponding to \( \frac{1}{2\sqrt{2}} \begin{pmatrix}1 & 0 \\
0 & 1
\end{pmatrix} \) and \( \frac{1}{2\sqrt{2}} \begin{pmatrix}0 & 1 \\
1 & 0
\end{pmatrix} \), respectively. Hence the generated state is indeed an ergodic pgFCS with period 2. To determine the asymptotic behavior of \( \rho_{AC} \) we explicitly express it in terms of \( V_1 \) and \( V_2 \),

\[
\rho_{ABC} = \begin{cases}
\frac{1}{4} \left( \text{tr}_2(V_1V_2 \cdots V_1V_1^\dagger V_2^\dagger \cdots V_2^\dagger V_1^\dagger) \\
+ \frac{1}{4} \left( \text{tr}_2(V_1V_2 \cdots V_1V_1^\dagger V_2^\dagger \cdots V_2^\dagger V_1^\dagger) \right),
\end{cases}
\]
\[
\text{if } |B| \text{ mod } 2 = 1,
\]
\[
\frac{1}{4} \left( \text{tr}_2(V_2V_1 \cdots V_1V_1^\dagger V_2^\dagger \cdots V_2^\dagger V_1^\dagger) \\
+ \frac{1}{4} \left( \text{tr}_2(V_2V_1 \cdots V_1V_1^\dagger V_2^\dagger \cdots V_2^\dagger V_1^\dagger) \right),
\end{cases}
\]
\[
\text{if } |B| \text{ mod } 2 = 0.
\]
Since \( V_2 = U_sV_1 \) for a unitary \( U_s: C^3 \to C^3 \), then
\[
\text{tr}_3(V_1XV_1^\dagger) = \text{tr}_3(V_2XV_2^\dagger)
\]
for any \( X \in \mathcal{B}(\mathcal{H}_M) \). Notice that for any choice of the indices \( \{i_k \in \{1, 2\} | k = 1, 2, \cdots, |B|\} \), we have
\[
\lim_{|B| \to +\infty} \mathcal{E}^{[B]}(V_{i_1} \cdots V_{i_2}XV_{i_1}^\dagger \cdots V_{i_2}^\dagger) = \lim_{|B| \to +\infty} \mathcal{E}^{[B]}(X) = \mathcal{E}^{[B]}(X) \mathbb{1}_2/2.
\]
This implies that in the asymptotic limit \( |B| \to +\infty \), the state \( \rho_{AC} \) oscillates between two states,
\[
\rho_{AC}^{(1)} = \frac{1}{8} \left( \text{tr}_2(V_1V_1^\dagger) + \text{tr}_2(V_2V_2^\dagger) + \text{tr}_2(V_1V_1^\dagger) \right)
\]
\[
\otimes \left( \text{tr}_2(V_1V_1^\dagger) + \text{tr}_2(V_2V_2^\dagger) \right),
\]
For both of these states,
\[
\rho_{AC}^{(1)} \otimes \rho_{AC}^{(2)} = \rho_{AC}^{(2)} \otimes \rho_{AC}^{(2)} = \frac{1}{16} \left( \text{tr}_2(V_1V_1^\dagger) + \text{tr}_2(V_2V_2^\dagger) \right)
\]
\[
\otimes \left( \text{tr}_2(V_1V_1^\dagger) + \text{tr}_2(V_2V_2^\dagger) \right).
\]
Using the explicit expressions (D1) and (D2), we obtain
\[
\rho_{AC}^{(1)} \otimes \rho_{AC}^{(1)} = \rho_{AC}^{(2)} \otimes \rho_{AC}^{(2)} = \frac{1}{64} \left( \begin{pmatrix}
2 & 0 \\
0 & 3
\end{pmatrix} \otimes \begin{pmatrix}
2 & 0 \\
0 & 3
\end{pmatrix} \right).
\]
An explicit calculation shows that \( \tilde{\rho}_{AC}^{(1)} \) and \( \tilde{\rho}_{AC}^{(2)} \) have the same QMI, calculated in the previous example, \( I(A : C) = 17 \ln 2/16 - 9 \ln 3/8 + 5 \ln 5/16 \approx 0.0035 \). Thus, while for this ergodic pgFCS with period 2 QMI actually converges, it still does not converge to zero.

\section*{Appendix E}

In this appendix we present a pair of lemmata for completeness.

The first lemma is a restatement of the Lemma A.2 from Ref. 46 in a slightly different language.

\textbf{Lemma E.1 (Lemma A.2 of Ref. 46).} \textit{Let isometries} \( V_1: \mathcal{H}_M \to \mathcal{H}_s \otimes \mathcal{H}_M \) \textit{and} \( V_2: \mathcal{H}_M \to \mathcal{H}_s \otimes \mathcal{H}_M \) \textit{be such that full-rank density operators} \( \sigma_1, \sigma_2 \in \mathcal{B}(\mathcal{H}_M) \) \textit{are the only fixed points of the quantum channels} \( X \mapsto \text{tr}_1(V_1XV_1^\dagger) \) \textit{and} \( X \mapsto \text{tr}_1(V_2XV_2^\dagger) \), \textit{X} \textit{is in} \( \mathcal{B}(\mathcal{H}_M) \), \textit{respectively, and that these quantum channels have no other eigenvalues of magnitude 1. Then the eigenvalues} \( \nu_i \) \textit{of the map} \( X \mapsto \text{tr}_1(V_1XV_1^\dagger) \) \textit{are such that} \( \nu_i \neq 1 \text{ for all } i \).
that $|\nu| \leq 1$. Furthermore, if there exists the eigenvalue $\nu_i$ of magnitude 1, then $V_i = e^{-i\phi}(1_\times \otimes W^\dagger)V_i W$ for some $\phi \in \mathbb{R}$ and unitary $W: \mathcal{H}_M \to \mathcal{H}_M$.

**Proof.** Let $\nu_i$ be an eigenvalue, so that for some $X$:

$$tr_s(V_iXV_i^\dagger) = \nu_iX.$$  

Using the fact that $\sigma_1$ is invertible, consider:

$$|\nu|tr(X^\dagger \sigma_1^{-1}X)|^2 = |tr(X^\dagger \sigma_1^{-1}tr_s(V_iXV_i^\dagger))|^2$$

$$\leq \nu_i (\lambda_i \otimes \sigma_1^{-1})tr_s(V_iV_i^\dagger) = \nu_i|X||X| = |X|^2.$$  

(E1)

where we have used the Cauchy–Schwarz inequality for the Hilbert–Schmidt inner product $|tr(A^\dagger B)|^2 \leq tr(A^\dagger A)tr(B^\dagger B)$ to get from the second line to the third. Thus $|\nu| \leq 1$. Moreover $|\nu| = 1$ only in the case of equality in the Cauchy–Schwarz inequality, which implies:

$$\sigma_1^{-1/2}XV_2^\dagger = e^{i\phi}(\sigma_1^{-1/2}V_1^\dagger(1_\times \otimes \sigma_1^{-1}X),$$

where $\phi \in \mathbb{R}$. Using that $\sigma_1^{-1/2}$ is invertible and denoting $W = \sigma_1^{-1}X$, we get:

$$WV_2^\dagger = e^{i\phi}V_1^\dagger(1_\times \otimes W).$$

Using that $V_2^\dagger V_2 = 1_M$, we obtain:

$$WW^\dagger = V_1^\dagger(1_\times \otimes WW^\dagger)V_1.$$  

This means that $WW^\dagger$ is the eigenvector with eigenvalue 1 of the map $X \mapsto V_1^\dagger(1_\times \otimes X)V_1$. The latter map is adjoint to the quantum channel $X \mapsto tr_s(V_iXV_i^\dagger)$, with respect to the Hilbert–Schmidt inner product, therefore it has the same spectrum. Hence the adjoint map has only one eigenvector with eigenvalue 1. Since $V_1$ is an isometry, $V_1^\dagger(1_\times \otimes 1_M)V_1 = 1_M$, the identity operator $1_\times \otimes I_\times$ is this eigenvector, and $WW^\dagger = 1_M$. Replacing $tr(X^\dagger \sigma_1^{-1}X)$ with $tr(X^\dagger X\sigma_1^{-1}X)$ in (E1) leads through the same reasoning to $WW^\dagger = 1_M$. Thus $W$ is unitary and

$$V_2 = e^{-i\phi}XV_1^\daggerV_1 W.$$  

Proof. Following the procedure described in Appendix F, we represent the maps $\mathcal{E}$ and $\tilde{\mathcal{E}}$ with the operators $E$ and $\tilde{E}$, respectively, that inherit the spectra of the maps.

As is shown in Appendix F,

$$\|\mathcal{E}^n - \tilde{\mathcal{E}}^n\|_{2 \rightarrow 2} = \|E^n - \tilde{E}^n\|.$$  

(E2)

Then to prove the statement of the lemma we can estimate $\|E^n - \tilde{E}^n\|$.

We make the Jordan decomposition of $E$,

$$E = \sum_{i \in |\nu| = 1}(\nu_iP_{\nu_i} + N_{\nu_i}),$$

where $\nu_i$ is an eigenvalue, $P_{\nu_i}$ is the projector onto the subspace corresponding to the eigenvalue $\nu_i$, and $N_{\nu_i}$ is a nilpotent operator with index $(K_i + 1)$ and $N_{0}^{K_i+1} = 0$. The Jordan decomposition for $\tilde{E}$ can be obtained by setting all $\nu_i$, except for the ones of magnitude 1, to 0,

$$\tilde{E} = \sum_{i \in |\nu| = 1}(\nu_iP_{\nu_i} + N_{\nu_i}).$$

From the properties of the operators $P_{\nu_i}$ and $N_{\nu_i}$ in the Jordan decomposition, $P_{\nu_i}P_{\nu_j} = \delta_{ij}P_{\nu_i}$, $P_{\nu_i}N_{\nu_i} = N_{\nu_i}P_{\nu_i} = \delta_{ij}N_{\nu_i}$, and $N_{\nu_i}N_{\nu_i} = \delta_{ik}N_{\nu_i}^2$, it follows that we can express $E^n$ as

$$E^n = \sum_{i \in |\nu| = 1}(\nu_iP_{\nu_i} + N_{\nu_i})^n.$$  

For all peripheral eigenvalues $\nu_i$, i.e. such that $|\nu| = 1$, $N_{\nu_i} = 0$ (see, for example, the proof of Proposition 3.1 in Ref.1). Thus, we can write the expression for $E^n$ as

$$E^n = \sum_{|\nu| = 1} \nu_i^np_{\nu_i} + \sum_{|\nu| < 1}(\nu_iP_{\nu_i} + N_{\nu_i})^n.$$  

It immediately follows that $E^n = \sum_{|\nu| = 1} \nu_i^np_{\nu_i}$.

Then we can estimate

$$\|E^n - \tilde{E}^n\| = \|\sum_{|\nu| = 1} (\nu_iP_{\nu_i} + N_{\nu_i})^n - \sum_{|\nu| < 1} (\nu_iP_{\nu_i} + N_{\nu_i})^n\|.$$  

(E3)

$$\leq \sum_{|\nu| < 1} \|(\nu_iP_{\nu_i} + N_{\nu_i})^n\|^2 + \sup_{\|\psi\| = 1} \langle \psi | \sum_{|\nu| < 1} (\nu_iP_{\nu_i} + N_{\nu_i})^nP_{\nu_i}\psi \rangle.$$  

$$\leq \sum_{|\nu| < 1} \|(\nu_iP_{\nu_i} + N_{\nu_i})^n\|^2 + \sup_{\|\psi\| = 1} \langle \psi | \sum_{|\nu| < 1} P_{\nu_i}\psi \rangle.$$  

$$\leq \sum_{|\nu| < 1} \|(\nu_iP_{\nu_i} + N_{\nu_i})^n\|^2 + \sup_{\|\psi\| = 1} \langle \psi | \sum_{|\nu| < 1} P_{\nu_i}\psi \rangle.$$  

We have used the definition of the operator norm and orthogonality of the terms $\nu_iP_{\nu_i} + N_{\nu_i}$ to get the second line. The third
line follows from the identity \( v_i P_{v_i} + N_{v_i} = (v_i P_{v_i} + N_{v_i}) P_{v_i} \), and the bound \( \langle \psi | A^\dagger A | \psi \rangle = ||A||^2 \leq ||A||^2 \langle \psi | \psi \rangle \) is used in the fourth line. The sixth and seventh lines follow from properties of the orthogonal projectors \( \sum_{|\psi_i\rangle \in 1} P_{\psi_i} \leq \sum_{|\psi_i\rangle \neq 1} P_{\psi_i} = \mathbb{1} \).

Now, using the fact that \( N_{v_i} = 0 \), we estimate

\[
\| (v_i P_{v_i} + N_{v_i})^n \| = \| P_{v_i} \sum_{k=0}^{n} \binom{n}{k} v_i^{n-k} N_{v_i}^k \| \\
\leq |v_i|^n \sum_{k=0}^{n} \binom{n}{k} \| N_{v_i} \|^k |v_i|^{n-k} \\
\leq |v_i|^n \sum_{k=0}^{n} \binom{n}{k} \| N_{v_i} \|^k |v_i|^{n-k} \\
\leq |v_i|^n \sum_{k=0}^{n} \binom{n}{k} \left( \frac{n \| N_{v_i} \|}{|v_i|} \right)^k \\
= |v_i|^n \left( \frac{n \| N_{v_i} \|}{|v_i|} + 1 \right)^{K_i} ,
\]

where we have used the triangle inequality and the inequality \( \binom{n}{k} \leq \frac{n^k}{k^n} \) for \( k \leq K_i \). Denoting \( c_1 := \max_{n \in [|v_i|]} \{ \| N_{v_i} \| \} \), and \( K_{\text{max}} := \max_{n \in [|v_i|]} K_i \), we bound

\[
\max_{0 < |v_i| < 1} \| (v_i P_{v_i} + N_{v_i})^n \| \leq v_{\text{gap}}^n (c_1 n + 1)^{K_{\text{max}}} .
\]

Combining (E2), (E3), and (E4) together,

\[
\| E^n - \bar{E}^n \|_{2 \to 2} \leq v_{\text{gap}} (c_1 n + 1)^{K_{\text{max}}},
\]

manifesting an almost exponential decay. Using a slightly weaker bound, we can make the decay to be strictly exponential. We choose some \( v_{\text{gap}} < v < 1 \), expand

\[
v_{\text{gap}} (c_1 n + 1)^{K_{\text{max}}} = \left( \frac{v_{\text{gap}}}{v} \right)^n (c_1 n + 1)^{K_{\text{max}}} v^n ,
\]

and notice that first two factors can be bounded by a constant \( c \), depending on \( v \)

\[
\left( \frac{v_{\text{gap}}}{v} \right)^n (c_1 n + 1)^{K_{\text{max}}} \leq c ,
\]

which leads to the stated estimate,

\[
\| E^n - \bar{E}^n \|_{2 \to 2} \leq c v^n .
\]

\[ \blacksquare \]

Appendix F

In this appendix we review the matrix representation of a quantum channel.[28]

The map \( \mathcal{E} : \mathcal{B}(\mathcal{H}_M) \rightarrow \mathcal{B}(\mathcal{H}_M) \) can be represented by the operator \( \mathcal{E} : \mathcal{H}_M \otimes \mathcal{H}_M \rightarrow \mathcal{H}_M \otimes \mathcal{H}_M \) through the process of vectorization.[28] For the chosen orthonormal basis \( \{ |i\rangle : i = 1 \cdots d_M \} \) of \( \mathcal{H}_M \) the invertible map, vec : \( \mathcal{B}(\mathcal{H}_M) \rightarrow \mathcal{H}_M \otimes \mathcal{H}_M \), is defined through its action on the operators \( |i\rangle \langle j| \), which constitute an orthonormal basis of \( \mathcal{B}(\mathcal{H}_M) \) (with respect to the Hilbert-Schmidt inner product),

\[
\text{vec}(|i\rangle \langle j|) = |i\rangle \otimes |j\rangle , \quad i, j = 1 \cdots d_M .
\]

The inner product on the Hilbert space \( \mathcal{H}_M \otimes \mathcal{H}_M \) may then be related to the Hilbert-Schmidt inner product on the Hilbert space \( \mathcal{B}(\mathcal{H}_M) \), by

\[
\langle \text{vec}(X) | \text{vec}(Y) \rangle_{\mathcal{H}_M \otimes \mathcal{H}_M} = \text{tr}(X^\dagger Y) = \langle X, Y \rangle_{\mathcal{B}(\mathcal{H}_M)} ,
\]

for \( X, Y \in \mathcal{B}(\mathcal{H}_M) \). It follows immediately that

\[
\| X \|_2 = \| \text{vec}(X) \| ,
\]

where \( \| \Psi \| = \langle \Psi | \Psi \rangle^{\frac{1}{2}} \) is the norm of the vector.

Define the operator \( \mathcal{E} : \mathcal{H}_M \otimes \mathcal{H}_M \rightarrow \mathcal{H}_M \otimes \mathcal{H}_M \) by

\[
\mathcal{E} := \text{vec} \circ \mathcal{E} \circ \text{vec}^{-1} ,
\]

leading to the explicit form of the matrix elements of \( \mathcal{E} \),

\[
\langle m | \langle n | \mathcal{E} | i \rangle \otimes | j \rangle = \langle m | \mathcal{E} | i \rangle \langle i | n \rangle , \quad i, j, m, n = 1 \cdots d_M .
\]

The relations (F2) and (F3) imply that the 2 \to 2 norm of the map \( \mathcal{E} \) can be expressed through the operator norm of \( \mathcal{E} \),

\[
\| \mathcal{E} \|_{2 \to 2} := \sup_{\| X \|_2 = 1} \| \mathcal{E}(X) \|_2 \\
= \sup_{\| \text{vec}(X) \| = 1} \| \text{vec} \circ \mathcal{E} \circ \text{vec}^{-1} (\text{vec}(X)) \| \\
= \sup_{\| X \| = 1} \| \mathcal{E}(X) \| \\
= \| \mathcal{E} \| .
\]

The spectra of the map \( \mathcal{E} \) and the operator \( \mathcal{E} \) coincide: if \( \mathcal{E}(X) = \lambda X \) for some \( X \in \mathcal{B}(\mathcal{H}_M) \), then \( \mathcal{E}(\text{vec}(X)) = \lambda \cdot \text{vec}(X) \). Conversely, if \( \mathcal{E}(\text{vec}(\psi)) = \lambda \cdot \text{vec}(\psi) \), for some \( |\psi\rangle \in \mathcal{H}_M \otimes \mathcal{H}_M \), then \( \mathcal{E}(\mathcal{E}^{-1}(|\psi\rangle)) = \lambda \cdot \mathcal{E}^{-1}(|\psi\rangle) \).

Appendix G

In this appendix we show that without loss of generality, we can assume that the decomposition (13) of a pgFCS into the sum of ergodic components does not contain terms that are essentially equal in the sense described below.

Let us suppose there exists \( \phi \in \mathbb{R} \) and a unitary \( W : \mathcal{H}_M \rightarrow \mathcal{H}_M \), such that the isometries \( V_{\alpha} \) and \( V_{\beta} \) satisfy

\[
V_{\beta} = e^{-i\phi} (I_M \otimes W^\dagger) V_{\alpha} W .
\]

This condition defines an equivalence relation and the isometries \( V_{\alpha} \) and \( V_{\beta} \) are then said to be equivalent.

In this appendix we show that the general convex sum representation of the FCS may be written as

\[
\rho_{ABC} = \sum_{\alpha=1}^l \lambda_{\alpha} \rho_{ABC}^\alpha ,
\]

\[ \blacksquare \]
where the $\rho_{ABC}$ define a collection of $J$ ergodic pgFCS constructed as in (12), with a collection of inequivalent isometries $V_A$ and full-rank density operators $\sigma_A$.

Suppose the converse is true, then

$$
\rho_{ABC} = \text{tr}_M \left( V^*_C V^*_B V^*_A \sigma_B V_B W^*_B \otimes V^*_C V^*_A \sigma_A W^*_A \right) \tag{G3}
$$

where we have used the fact that $\sigma_B$ is a fixed point of the map $X \mapsto \text{tr}_r (V_B X V_B^+)$. Thus, $W \sigma_B W^+$ is the fixed point of the quantum channel $X \mapsto \text{tr}_r (V_B X V_B^+)$. Since $\rho_{ABC}$ is ergodic, the density operator $\sigma_A$ is the only fixed point of this map. Then $W \sigma_B W^+ = \sigma_A$. Combining this fact with (G3), we obtain that $\rho_{ABC} = \rho_{ABC}^\sigma$. Thus, $\lambda_{\sigma} \rho_{ABC} = (\lambda_{\sigma} + \lambda_B) \rho_{ABC}$. We can combine in the convex sum (G2) all the terms satisfying the relation (G1) together.

**Appendix II**

In this appendix we present a technical result, required in the proof of Theorem II.1 in Section IV B,

$$
V_A \sigma_A V_A^+ \in \beta(\mathcal{H}_A) \otimes \bigoplus_{k=0}^{p-1} \Pi_k \beta(\mathcal{H}_M) \Pi_k.
$$

We closely follow the reasoning of the proof of Proposition 3.3 of Ref. 1. Here $V_A : \mathcal{H}_M \rightarrow \mathcal{H}_A \otimes \mathcal{H}_M$ is the isometry generating the spins in the region $A$ as in (12), and $\sigma \in \beta(\mathcal{H}_M)$ and the projectors $\{ \Pi_k \in \beta(\mathcal{H}_M) \mid k = 0, \ldots, p - 1 \}$.

Let us express

$$
V_A \sigma_A V_A^+ = \sum_{n,m=1}^{d_A} |\phi_n\rangle \langle \phi_m| \otimes |\phi_n\rangle V_A \sigma_A V_A^+ |\phi_m\rangle,
$$

where $\{ |\phi_n\rangle \in \mathcal{H}_A \mid n = 1, \ldots, d_A \}$ is an orthonormal basis. It is clear that

$$
\langle \phi_n | V_A \sigma_A V_A^+ |\phi_m\rangle = \text{tr}_r (|\phi_n\rangle \langle \phi_m| V_A \sigma_A V_A^+ |\psi\rangle )
$$

and the vector space span$(|\psi\rangle)$ is $\beta(\mathcal{H}_A)$. Notice that $\langle \zeta | V_A \sigma_A V_A^+ |\psi\rangle = \text{tr}_r (|\zeta\rangle \langle \zeta| V_A \sigma_A V_A^+ |\psi\rangle )$ and the vector space span$(|\psi\rangle)$ is $\beta(\mathcal{H}_A)$.

Since any $O_A \geq 0$ satisfies $O_A \leq \langle \sigma_A | O_A | \sigma_A \rangle$, it follows that

$$
\text{tr}(O_A \sigma_A V_A^+) \leq \langle \sigma_A | O_A | \sigma_A \rangle = \langle \sigma_A | O_A | \sigma_A \rangle.
$$

Thus, we have used $V_A \sigma_A V_A^+ = \sigma$. The operators $\text{tr}(O_A \sigma_A V_A^+)$ and $\langle \sigma_A | O_A | \sigma_A \rangle$ can be ordered in this way only if $\text{tr}(O_A \sigma_A V_A^+) \in \bigoplus_{k=0}^{p-1} \Pi_k \beta(\mathcal{H}_M) \Pi_k$, since $\sigma \in \bigoplus_{k=0}^{p-1} \Pi_k \beta(\mathcal{H}_M) \Pi_k$. Then

$$
\text{tr}(O_A \sigma_A V_A^+) \in \bigoplus_{k=0}^{p-1} \Pi_k \beta(\mathcal{H}_M) \Pi_k,
$$

from which it follows that $\langle \zeta | V_A \sigma_A V_A^+ |\psi\rangle \in \bigoplus_{k=0}^{p-1} \Pi_k \beta(\mathcal{H}_M) \Pi_k$, and

$$
V_A \sigma_A V_A^+ \in \beta(\mathcal{H}_A) \otimes \bigoplus_{k=0}^{p-1} \Pi_k \beta(\mathcal{H}_M) \Pi_k.
$$

This leads to the identity $V_A \sigma_A = \sum_{k=0}^{p-1} \Pi_k V_A \sigma_A V_A^+ \Pi_k$.

**Appendix I**

In this appendix we prove that the quantum channel $\tilde{E}$ obtained from $\tilde{E}$ by Definition 1 of Section III B, and used in Section IV B, takes the form

$$
\tilde{E}(X) = \sum_{r=0}^{p-1} \text{tr}(\Pi_r X \Pi_r) \Pi_{r+1} \sigma \Pi_{r+1}.
$$

We use the notation established in Section IV B.

Let $E \in \beta(\mathcal{H}_M) \otimes \beta(\mathcal{H}_M)$ and $\tilde{E} \in \beta(\mathcal{H}_M) \otimes \beta(\mathcal{H}_M)$ be the operators representing $E$ and $\tilde{E}$ as described in Appendix F. The maps adjoint to $E$ and $\tilde{E}$, with respect to Hilbert-Schmidt inner product, are represented by the operators $E'$ and $\tilde{E}'$, respectively. From Proposition 3.3 of Ref. 1 we know that the peripheral spectrum of $E'$ consists of $p$ non-degenerate eigenvalues $\exp(\frac{2\pi i}{p} r)$, where $r = 0, \ldots, p - 1$, which correspond to the eigenvectors $\nu_r$, which are unitary operators. Each of $\nu_r$ can be decomposed as $\nu_r = \sum_{k=0}^{p-1} \exp(\frac{2\pi i}{p} r k) \Pi_k$, where $\{ \Pi_k \}_{k=0, \ldots, p-1}$ is the set of projectors defined in Proposition 2. $\sum_{k=0}^{p-1} \Pi_k = \mathbb{1}_M$.

For brevity of notation in this appendix we denote the vectors $\text{vec}(X) \in \mathcal{H}_M \otimes \mathcal{H}_M$, corresponding to the operators $X \in \beta(\mathcal{H}_M)$ through the isomorphism vec defined in Appendix F, as $|X\rangle := \text{vec}(X)$. Then, in this notation,

$$
|u_r\rangle = \sum_{k=0}^{p-1} \exp\left(\frac{2\pi i}{p} r k\right) |\Pi_k\rangle.
$$
Since the matrix with the elements \( \{ \frac{1}{p} \exp \left( \frac{2\pi i}{p} r k \right) \}_{r k} \) is unitary, we can invert the above relation,
\[
|\Pi_k \rangle = \frac{1}{p} \sum_{r=0}^{p-1} \exp \left( - \frac{2\pi i}{p} r k \right) |u_r \rangle.
\]

Notice that since \( E^\dagger |u_r \rangle = \exp \left( \frac{2\pi i}{p} r \right) |u_r \rangle \), it follows that
\[
E^\dagger |\Pi_k \rangle = |\Pi_{k-1} \rangle,
\]
as was obtained in the proof of Proposition 3.3 of Ref. 1.

As \( \tilde{E} \) constructed using Definition 1 is defined by the peripheral spectrum of \( E \), it follows that \( \tilde{E}^\dagger |u_r \rangle = \exp \left( \frac{2\pi i}{p} r \right) |u_r \rangle \), \( \tilde{E}^\dagger |\Pi_k \rangle = |\Pi_{k-1} \rangle \) and \( \tilde{E}^\dagger = \sum_{r=0}^{p-1} \exp \left( \frac{2\pi i}{p} r \right) P_r \), where \( P_r \) is the projection onto the subspace spanned by \( |u_r \rangle \). We can represent \( P_r = |u_r \rangle \langle v_r | \), where \( \langle v_r | \) satisfy \( \langle v_r | |u_r \rangle = \delta_{r \rho} \). It immediately follows that
\[
\tilde{E} = \sum_{r=0}^{p-1} \exp \left( - \frac{2\pi i}{p} r \right) |v_r \rangle \langle u_r |,
\]
and that \( |v_r \rangle \) are eigenvectors of \( \tilde{E} \) corresponding to the eigenvalues \( \exp \left( - \frac{2\pi i}{p} r \right) \). We introduce the vectors \( |\omega_k \rangle \), defined by
\[
|\omega_k \rangle := \frac{1}{p} \sum_{r=0}^{p-1} \exp \left( - \frac{2\pi i}{p} r k \right) |v_r \rangle.
\]

It is easy to see that \( \tilde{E} |\omega_k \rangle = |\omega_{k+1} \rangle \) and \( |\omega_k \rangle \langle \Pi_k | = \delta_{k \rho} \). Moreover, we can express \( \tilde{E} \) in terms of \( |\omega_k \rangle \) and \( |\Pi_k \rangle \),
\[
\tilde{E} = \sum_{r=0}^{p-1} \exp \left( - \frac{2\pi i}{p} r \right) |v_r \rangle \langle u_r | = \sum_{k, k' = 0}^{p-1} \exp \left( \frac{2\pi i}{p} r (k - k' - 1) \right) |\omega_k \rangle \langle \Pi_{k'} | = \sum_{k=0}^{p-1} |\omega_k \rangle \langle \Pi_{k-1} |.
\]

Since \( \tilde{E} |\omega_{k+1} \rangle = \sum_{k=0}^{p-1} |\omega_{k+1} \rangle = \sum_{k=0}^{p-1} |\omega_k \rangle \), then \( |\omega_k \rangle \) is the fixed point of \( \tilde{E} \). In the case of the ergodic state the only fixed point of \( \tilde{E} \) is \( |\sigma \rangle = \sum_{k=0}^{p-1} |\Pi_k \sigma \Pi_k | \). Then \( |\sigma \rangle = \sum_{k=0}^{p-1} |\Pi_k \sigma \Pi_k | \) with \( |\Pi_k \sigma \Pi_k | = \text{vec}(\Pi_k \sigma \Pi_k) \). Then \( |\sigma \rangle = |\omega \rangle \). Moreover, since \( \langle \Pi_k | \Pi_k \sigma \Pi_k | = \delta_{k k} \), then \( |\omega_k \rangle = |\Pi_k \sigma \Pi_k | \). Then \( \tilde{E} = p \sum_{k=0}^{p-1} |\Pi_k \sigma \Pi_k | \langle \Pi_{k+1} | \langle \Pi_{k+1} | \Pi_k |, \) which corresponds to the quantum channel
\[
\tilde{E}(X) = p \sum_{k=0}^{p-1} \text{tr}(\Pi_k X \Pi_k) |\Pi_k+1 \rangle \sigma \Pi_k+1 |,
\]
as required.

In this appendix we extend the construction of the recovery quantum channel defined in equations (32)-(35) to the general case of pgFCS.

We use the intuition built in Section IV. For a general pgFCS \( \rho_{ABC} \) (12) (which can be decomposed into a convex sum (13)) we build the approximating state \( \rho_{ABC} \) (15) induced by
\[
\tilde{V}_B = \sum_{a_1=0}^{p-1} \sum_{a_0 \in \sigma \Omega} \sqrt{\lambda_{a_0 \sigma} \rho_{a_0 \sigma} \Omega} |\tilde{\xi}_{i a}^a \rangle \langle \tilde{\xi}_{i a}^a |, \quad (J1)
\]
where the index \( \alpha \) distinguishes the components in the convex sum decomposition (13). The terms corresponding to different \( \alpha \) are the isometries inducing ergodic pgFCS and possessing the form (30). The sets \( \sigma \Omega \) are defined in (28). From the Sections IV A, IV B, and IV C we know that \( \tilde{V}_B \) of the form (J1) (with any set of orthonormal vectors \( |\tilde{\xi}_{i a}^a \rangle \) guarantees \( \tilde{\rho}_{ABC} \) is a quantum Markov chain.

At this point it is convenient to restrict ourselves to the case of \( J = 1 \), since the further arguments can be extended to the general case in an obvious manner. Thus in the sequel we deal with an ergodic pgFCS with period \( p \geq 1 \), for which
\[
\tilde{V}_B = \sum_{r=0}^{p-1} \sqrt{p} \sum_{(i, j) \in \sigma \Omega} \sqrt{\sigma} |\tilde{\xi}_{i j}^a \rangle \langle |\tilde{\xi}_{i j}^a | |
\]

In contrast to the main text, we here present an explicit procedure to choose \( |\tilde{\xi}_{i j}^a \rangle \), which leads to a recovery error exponentially small in the size of the region \( B \). These will be derived with the use of the \( d_M^2 \times d_M^2 \) matrix \( Q \).

We define \( \tilde{Q} \) together with the auxiliary matrix \( \tilde{Q} \) of the same dimension by their matrix elements,
\[
Q_{\alpha \beta} = \langle j' | V_B^\dagger j' | i | V_B | j \rangle,
\]

\[
\tilde{Q}_{\alpha \beta} = \langle j' | \tilde{V}_B^\dagger j' | i | \tilde{V}_B | j \rangle,
\]

where \( i, j = 1, \ldots, d_M \). If the region \( B \) is large, these matrices are close to each other in matrix norm. Notice that as Gramian matrices, both \( \tilde{Q} \) and \( \tilde{Q} \) are positive semidefinite. We can explicitly express \( \tilde{Q} \).

\[
\tilde{Q}_{\alpha \beta} = p \sigma_{\alpha \beta} \delta_{i j} \Sigma_{(i, j) \in \sigma \Omega} \tilde{\chi}(i, j),
\]

where \( \chi_{\sigma} \) is a characteristic function of the set \( \sigma \) defined in (28). We observe that \( \tilde{Q} \) is diagonal and positive semidefinite, but not full-rank if \( p > 1 \) (since \( \chi_{\sigma} \) is zero for some pairs \( (i, j) \)). Meanwhile, \( \tilde{Q} \) differs from \( \tilde{Q} \) by a small correction, and in general can have different rank. We want to construct the number of vectors \( |\tilde{\xi}_{i j}^a \rangle \) equal to the cardinality of the set \( \sigma \), which is equal to the rank of \( \tilde{Q} \), hence we want to force the ranks of \( \tilde{Q} \) and \( \tilde{Q} \) to be equal. But first let us rigorously determine how close \( \tilde{Q} \) and \( \tilde{Q} \) are.

Matrices \( \tilde{Q} \) and \( \tilde{Q} \) are partial transposes of \( E \) and \( \tilde{E} \), respectively, defined in Appendix E. The latter are related to the quantum channels \( \tilde{E} \) and \( \tilde{E} \), respectively, defined in Section IV. Then from Ref. 47.

\[
||\tilde{Q} - \tilde{Q}|| \leq d_M ||E - \tilde{E}|| \leq c d_M \nu_{|\nu|}, \quad (J4)
\]
where we have used Lemma III.3 in the second inequality. Now we can represent $Q$ in the form

$$Q = \tilde{Q} + cd_Mv^{[b]}Z,$$  \hspace{1cm} (J5)

where $Z$ is a Hermitian matrix with $\|Z\| \leq 1$.

From (J3) it is clear that the smallest non-zero eigenvalue of $\tilde{Q}$ is equal to $\sigma_{\text{min}}$, the smallest eigenvalue of $\sigma$. The matrices $Q$ and $\tilde{Q}$ are close if the region $B$ is large enough. By dropping from $Q$ the part of the spectrum below $\sigma_{\text{min}}$, we obtain from $Q$ the matrix $Q'$, which approximates to both $Q$ and $\tilde{Q}$ and has the same rank as the latter. We first make the spectral decomposition, $Q = WAW^T$. We now define $\Pi_{\geq \epsilon}$, the projector onto the part of the spectrum with eigenvalues greater than $\epsilon = cd_Mv^{[b]}$, and further define $Q' := WA_{\geq \epsilon}W^T$, where

$$\Lambda_{\geq \epsilon} = \Pi_{\geq \epsilon}\Lambda_{\geq \epsilon}.$$  \hspace{1cm} (J4)

As follows from (J4), if $\sigma_{\text{min}} > 2cd_Mv^{[b]}$, then as $|B| \to \infty$ there is a part of the spectrum of $Q$ with eigenvalues converging to 0 that is well separated from the rest of the spectrum. We assume that this condition is satisfied.

For $Q'$ we have the decomposition analogous to (J5), $Q' = \tilde{Q} + cv^{[b]}Z'$, where $Z'$ is Hermitian and $\|Z'\| \leq 1$.

Since $Q'$ is positive semidefinite, $Q'^{\frac{1}{2}}$ is well-defined. We take the pseudo-inverse of $Q'^{\frac{1}{2}}$, and define $|\xi_{ij}|$,

$$|\xi_{ij}| := \sum_{m,n=1} |Q'^{\frac{1}{2}}|_{mn,ij} \langle m|v_B|n\rangle.$$  \hspace{1cm} (J6)

Now we determine the bound on $\|V_B - \bar{V}_B\|$. By definition of the operator norm,

$$\|V_B - \bar{V}_B\|^2 = \left\| \sum_{i,j=1}^{dt} (\langle i|v_B|j \rangle - \sqrt{p} \sigma_i |\xi_{ij}|) \otimes |i\rangle \langle j| \right\|^2 = \left\| \sum_{i,j=1}^{dt} (\langle i|v_B|j \rangle - \sqrt{p} \sigma_i \sum_{m,n=1} |Q'^{\frac{1}{2}}|_{mn,ij} \langle m|v_B|n\rangle) \otimes |i\rangle \langle j| \right\|^2$$

$$= \max_{\Sigma_i|\eta_i|^2 = \sum_{j,i=1}^{dt}} \sum_{j,i=1}^{dt} \eta_j \sum_{i=1}^{dt} \left( Q_{ij';ij} + \tilde{Q}_{ij';ij} - 2 \sqrt{p} \sigma_i Q'^{\frac{1}{2}}_{ij';ij} \right) \eta_j = \|\text{tr}_1 Q + \text{tr}_1 \tilde{Q} - 2\text{tr}_1((\sqrt{\sigma} \otimes 1)Q'^{\frac{1}{2}})\|,$$

where we have used the definition of $Q$ (J2), the identity $p\sigma_i \langle \xi_{ij}|\xi_{ij}| = Q_{ij';ij}$, which follows from (J2) and (J6), and the identity $\langle j'|v_B^T|i\rangle \langle m|v_B|n\rangle (Q'^{\frac{1}{2}})_{mn,ij} = Q_{ij';mn}(Q'^{\frac{1}{2}})_{mn,ij} = Q_{ij';ij}$. In the last line $\text{tr}_1$ denotes the partial trace over the first pair of indices.

Notice that

$$\sum_{i=1}^{dt} \tilde{Q}_{ij';ij} = \sum_{i=1}^{dt} p\sigma_i \delta_{ij'} \chi_{\sigma_i}(i, j)$$

$$= p \left( \sum_{i=1}^{dt} \chi_{\sigma_i}(i, j) \right) \delta_{ij'}$$

$$= \text{tr}_1(\Pi_{\geq |B|} \sigma \Pi_{\geq |B|}) \delta_{ij'}$$

$$= \frac{1}{p} \delta_{ij'} = \delta_{ij'}.$$

Thus,

$$\text{tr}_1(\tilde{Q}) = 1.$$  \hspace{1cm} (J8)

Now we need to bound $\text{tr}_1 Q$ and $\text{tr}_1((\sqrt{\sigma} \otimes 1)Q'^{\frac{1}{2}})$. Using
The former inequality is based on the previously assumed condition $2c_{dM}V[B] \leq \sigma_{\text{min}}$, which leads to $\|cd_{dM}V[B]\,\hat{Q}^{-\frac{1}{2}}Z'\,\hat{Q}^{-\frac{1}{2}} \leq \frac{\sigma_{\text{min}}}{\sigma_1^2} \leq \frac{1}{2}$. We have used the latter bound to estimate the remainder term in the Taylor expansion. We have also used $\|Z\| \leq \|Z\| \leq 1$. Thus,

$$\hat{Q}^{-\frac{1}{2}} - \frac{1+2^{-\frac{3}{2}}}{2}cd_{dM}V[B]\,\hat{Q}^{-\frac{1}{2}}Z'\,\hat{Q}^{-\frac{1}{2}} \leq \hat{Q}^{-\frac{1}{2}} + \frac{1}{2}cd_{dM}V[B]\,\hat{Q}^{-\frac{1}{2}}Z'\,\hat{Q}^{-\frac{1}{2}},$$

which leads to

$$1 - \frac{1+2^{-\frac{3}{2}}}{2}cd_{dM}V[B]\,\hat{Q}^{-\frac{1}{2}}Z'\,\hat{Q}^{-\frac{1}{2}} \leq \hat{Q}^{-\frac{1}{2}} + \frac{1}{2}cd_{dM}V[B]\,\hat{Q}^{-\frac{1}{2}}Z'\,\hat{Q}^{-\frac{1}{2}},$$

where we have used $\hat{Q}^{-\frac{1}{2}}\hat{Q}^{-\frac{1}{2}} = \frac{1}{2} + 1\,cd_{dM}V[B]\,\hat{Q}^{-\frac{1}{2}}Z'\,\hat{Q}^{-\frac{1}{2}}$. Since $\|\hat{Q}^{-\frac{1}{2}}\| \leq \|\hat{Q}\| \leq d_{M}$, then

$$1 - \frac{1+2^{-\frac{3}{2}}}{2}cd_{dM}V[B]\,\hat{Q}^{-\frac{1}{2}}Z'\,\hat{Q}^{-\frac{1}{2}} \leq \hat{Q}^{-\frac{1}{2}} + \frac{1}{2}cd_{dM}V[B]\,\hat{Q}^{-\frac{1}{2}}Z'\,\hat{Q}^{-\frac{1}{2}}.$$

Combining (J8), (J9), and (J10) together,

$$\|\hat{Q}^{-\frac{1}{2}}\| + 2\|\hat{Q}^{-\frac{1}{2}}\| \leq (1 + 2^{-\frac{3}{2}})cd_{dM}V[B].$$

From (J7) we obtain

$$\|V_B - \hat{V}_B\| \leq \sqrt{2(1 + 2^{-\frac{3}{2}})cd_{dM}V[B]}.$$

This bound is only a factor $\sqrt{2(1 + 2^{-\frac{3}{2}})}d_{M}$ worse than the bound (23), though it requires the condition $\sigma_{\text{min}} > 2cd_{dM}V[B]$ to be satisfied.

Analogous to the bound (23) in Section IV, (J11) remains true for a general pgFCS. In the case $J > 1$ of the convex decomposition (13), the matrices $Q$, $Q'$, and $\hat{Q}$ decompose into direct sums over ergodic components, i.e., $Q = \bigoplus_{i=1}^{J} Q_i$. The generalization to this case is straightforward. For the ergodic pgFCS of period 1, $\hat{Q}$ is full-rank, hence we have the simplification $Q' = Q$, where $Q$ is guaranteed to be invertible provided $\sigma_{\text{min}} > c_{dM}V[B]$. We have discussed this case in Section V.
pgFCS in general does not belong to the class of states considered in Theorem III.2 (the converse is also true), and never belongs to the class considered in Theorem III.1.

The quantum channels considered in Proposition III.3 of Ref. 24 have a forgetful component, i.e., can be decomposed as

$$\Phi(X) = (1 - \eta) \text{tr}(X) \chi + \eta \mathcal{N}(X), \quad \text{(K8)}$$

where $\chi \in \mathcal{D}(\mathcal{H}_S \otimes \mathcal{H}_M)$, $0 < \eta < 1$ and $\mathcal{N} : \mathcal{B}(\mathcal{H}_M) \to \mathcal{B}(\mathcal{H}_S \otimes \mathcal{B}(\mathcal{H}_M))$ is some quantum channel. This class of states is also distinct from pgFCS, since no channel of the form $\Phi(X) = V X V^\dagger$ can be decomposed as (K8). It is intuitively obvious that a channel constructed as a conjugation by isometries does not have a forgetful component. As proof, consider two orthogonal vectors $|\psi\rangle, |\phi\rangle \in \mathcal{H}_M$, $\langle \psi|\phi\rangle = 0$. Then, on the one hand,

$$\text{tr}(\Phi(|\psi\rangle\langle\psi|)\Phi(|\phi\rangle\langle\phi|)) = \text{tr}(V|\psi\rangle\langle\psi|V^\dagger|\phi\rangle\langle\phi|V^\dagger)$$

$$= ||\langle\psi|\phi\rangle|^2$$

$$= 0,$$

and on the other hand, for quantum channels admitting the decomposition (K8),

$$\text{tr}(\Phi(|\psi\rangle\langle\psi|)\Phi(|\phi\rangle\langle\phi|))$$

$$= \text{tr}(((1 - \eta)\chi + \eta \mathcal{N}(|\psi\rangle\langle\psi|))$$

$$\times (((1 - \eta)\chi + \eta \mathcal{N}(|\phi\rangle\langle\phi|)))$$

$$= (1 - \eta)^2 \text{tr}(\chi^2) + \eta (1 - \eta) \text{tr}(\mathcal{N}(|\psi\rangle\langle\psi| + |\phi\rangle\langle\phi|))$$

$$+ \eta^2 \text{tr}(\mathcal{N}(|\psi\rangle\langle\psi|\mathcal{N}(|\phi\rangle\langle\phi|))$$

$$> 0.$$

since all the summands are non-negative, and $(1 - \eta)^2 \text{tr}(\chi^2) > 0$. Hence, the intersection between the class of pgFCS and the states generated by a partially forgetful channel (K8) is empty.

[1] M. Fannes, B. Nachtergaele, and R. F. Werner, Communications in Mathematical Physics 144, 443 (1992).
[2] M. Fannes, B. Nachtergaele, and R. F. Werner, Journal of Functional Analysis 120, 511 (1994).
[3] A. Klümper, A. Schadschneider, and J. Zittartz, Communications in Mathematical Physics, 24, 293 (1993).
[4] I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki, Communications in Mathematical Physics 115, 477 (1988).
[5] G. Vidal, Phys. Rev. Lett. 91, 147902 (2003).
[6] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, Quantum Info. Comput. 7, 401–430 (2007).
[7] C. Schön, K. Hammerer, M. M. Wolf, J. I. Cirac, and E. Solano, Phys. Rev. A 75, 032311 (2007).
[8] L. Zhou, S. Choi, and M. D. Lukin, Phys. Rev. A 104, 032418 (2021).
[9] F. Verstraete, J. I. García-Ripoll, and J. I. Cirac, Phys. Rev. Lett. 93, 207204 (2004).
[10] T. Kuwahara, K. Kato, and F. G. S. L. Brandão, Phys. Rev. Lett. 124, 220601 (2020).
[11] K. Kato and F. G. S. L. Brandão, Communications in Mathematical Physics 370, 117 (2019).
[12] M. M. Wilde, Quantum Information Theory, 2nd ed. (Cambridge University Press, 2017).
[13] E. H. Lieb and M. B. Ruskai, Journal of Mathematical Physics 14, 1938 (1973), https://doi.org/10.1063/1.1666274.
[14] D. Petz, Communications in Mathematical Physics 105, 123 (1986).
[15] D. Petz, The Quarterly Journal of Mathematics 39, 97 (1988).
[16] P. Hayden, R. Jozsa, D. Petz, and A. Winter, Communications in Mathematical Physics 246, 359 (2004).
[17] O. Fawzi and R. Renner, Communications in Mathematical Physics 340, 575 (2015).
[18] D. Sutter, O. Fawzi, and R. Renner, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 472, 20150623 (2016).
[19] M. Junge, R. Renner, D. Sutter, M. M. Wilde, and A. Winter, Annales Henri Poincaré 19, 2955 (2018).
[20] S. T. Flammia, J. Haah, M. J. Kastoryano, and I. H. Kim, Quantum 1, 4 (2017).
[21] F. G. S. L. Brandão and M. J. Kastoryano, Communications in Mathematical Physics 365, 1 (2019).
[22] J. Guth Jarkovský, A. Molnár, N. Schuch, and J. I. Cirac, PRX Quantum 1, 010304 (2020).
[23] T. Kuwahara, A. M. Alhambra, and A. Anshu, Phys. Rev. X 11, 011047 (2021).
[24] C.-F. Chen, K. Kato, and F. G. S. L. Brandão, “Matrix product density operators: when do they have a local parent hamiltonian?” (2021), arXiv:2010.14682 [quant-ph].
[25] F. G. Brandao and K. M. Svore, in 2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS) (2017), pp. 415–426.
[26] F. G. S. L. Brandão, A. Kalev, T. Li, C. Y.-Y. Lin, K. M. Svore, and X. Wu, “Quantum sdp solvers: Large speed-ups, optimality, and applications to quantum learning,” (2019), arXiv:1710.02581 [quant-ph].
[27] J. van Apeldoorn, A. Gilyén, S. Gribling, and R. de Wolf, Quantum 4, 230 (2020).
[28] J. Watrous, The Theory of Quantum Information (Cambridge University Press, 2018).
[29] W. F. Stinespring, Proceedings of the American Mathematical Society 6, 211 (1955).
[30] R. Alicki and M. Fannes, 37, L55 (2004).
[31] A. Winter, Communications in Mathematical Physics 347, 291 (2016).
[32] B. Swingle and J. McGreevy, Phys. Rev. B 94, 155125 (2016).
[33] F. Hiai, M. Ohya, and M. Tsukada, Pacific Journal of Mathematics 96, 99 (1981).
[34] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac, Phys.
Rev. Lett. 100, 070502 (2008).

[35] F. G. S. L. Brandão and M. Horodecki, Communications in Mathematical Physics 333, 761 (2015).

[36] D. Kretschmann, D. Schlingemann, and R. F. Werner, IEEE Transactions on Information Theory 54, 1708 (2008).

[37] D. Sutter, Approximate Quantum Markov Chains (Springer Cham, 2018).

[38] M. Christandl, N. Schuch, and A. Winter, Communications in Mathematical Physics 311, 397 (2012).

[39] B. Ibinson, N. Linden, and A. Winter, Communications in Mathematical Physics 277, 289 (2008).

[40] A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, Classical and Quantum Computation (American Mathematical Society, USA, 2002).

[41] D. E. Evans and R. Høegh-Krohn, Journal of the London Mathematical Society s2-17, 345 (1978).

[42] O. Szehr, D. Reeb, and M. M. Wolf, Communications in Mathematical Physics 333, 565 (2015).

[43] I. Kim, (2017), arXiv:1709.07828 [quant-ph].

[44] M. Fannes, Communications in Mathematical Physics 31, 291 (1973).

[45] G. Lindblad, Communications in Mathematical Physics 40, 147 (1975).

[46] J. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, Annals of Physics 378, 100 (2017).

[47] J. Tomiyama, Proc. Amer. Math. Soc. 88, 635 (1983).