Loss of Memory and Convergence of Quantum Markov Processes

Keiji Matumoto
National Institute of Informatics,
Hitotsubashi 2-1-2, Chiyoda-ku, Tokyo 101-8430 Japan

May 5, 2014

1 Abstract

In a quantum (inhomogeneous) Markov process \( \rho_1 := \Gamma_1 (\rho) \), \( \rho_2 := \Gamma_2 (\rho_1) \), \( \cdots \), where \( \Gamma_i \) are CPTP maps and \( \rho \) is the initial state, the state of the system is either oscillatory or convergent to a point or convergent to an oscillatory orbit. Whichever the case it is, "information" about the initial state is always monotone non-increasing and convergent. This fact motivate us to define an equivalence class of families of quantum states, which embodies the bundle of all "information quantities" about the initial state. We show, for any quantum inhomogeneous Markov process over a finite dimensional Hilbert space, the trajectory in the space of the all equivalence classes is "monotone decreasing" and convergent to a point, relative to a reasonably defined topology. Also, a characterization of weak ergodicity in this picture is given.

2 Introduction

A classical (inhomogeneous) Markov process is defined by a sequence \( \{ P_i \}_{i=1}^{\infty} \) of transition probability matrices, and an initial probability distribution \( p \),

\[ p_1 := P_1 p, \quad p_2 := P_2 p_1, \quad \cdots \]

A quantum version of (inhomogeneous) Markov process may be defined by a sequence \( \{ \Gamma_i \}_{i=1}^{\infty} \) of completely positive and trace preserving (CPTP) maps, and an initial density matrix \( \rho \),

\[ \rho_1 := \Gamma_1 (\rho), \quad \rho_2 := \Gamma_2 (\rho_1), \quad \cdots \]

(If the probability space is a discrete set, the former is just a special case of the latter.) A classical or quantum Markov process may converge to a state, may oscillate, or may asymptotically come close to an oscillating orbit, depending on eigenvalues of \( P \) or \( \Gamma \).
Whichever the case it is, "information" about the initial state \((p\) or \(\rho\)) is a non-increasing function of time. This fact may be described mathematically as follows. Let \(\mathcal{E} = \{\rho_\theta; \theta \in \Theta\}\) be a family of initial states, \(D(\rho_{\theta_1}, \rho_{\theta_2}, \cdots \rho_{\theta_k})\) (an "information quantity") be a positive \(k\)-points function which is non-increasing by CPTP maps. \(D(\rho_{\theta_1}, \rho_{\theta_2}) := \|\rho_{\theta_1} - \rho_{\theta_2}\|_1\), e.g.) Also, let

\[
\begin{align*}
\mathcal{E}_0 : \mathcal{E} = \{\rho_\theta; \theta \in \Theta\}, \\
\mathcal{E}_1 := \{\rho_{\theta,1}; \theta \in \Theta\}, \rho_{\theta,1} := \Gamma_1(\rho_\theta), \\
\mathcal{E}_2 := \{\rho_{\theta,2}; \theta \in \Theta\}, \rho_{\theta,2} := \Gamma_2(\rho_{\theta,1}),
\end{align*}
\]

then

\[
D(\rho_{\theta_1}, \rho_{\theta_2}, \cdots \rho_{\theta_k}) \succeq D(\rho_{\theta_1,1}, \rho_{\theta_2,1}, \cdots \rho_{\theta_k,1}) \\
\succeq D(\rho_{\theta_1,2}, \rho_{\theta_2,2}, \cdots \rho_{\theta_k,2}) \\
\cdots \\
\succeq 0.
\]

Obviously, the sequence \(\{D(\rho_{\theta_1,i}, \cdots \rho_{\theta_k,i})\}_{i=1}^\infty\) converges, being monotone decreasing and bounded from below.

So we ask the following question. Is there an object which embodies the totality of information quantities, which is "monotone decreasing", and "converges" to a point as time passes? In this paper, as such an object, we propose an equivalence class of state families over a Hilbert space; \(\mathcal{E} = \{\rho_\theta; \theta \in \Theta\}\) is equivalent to \(\mathcal{F} = \{\sigma_\theta; \theta \in \Theta\}\) if and only if

\[
D(\rho_{\theta_1}, \rho_{\theta_2}, \cdots \rho_{\theta_k}) = D(\sigma_{\theta_1}, \sigma_{\theta_2}, \cdots \sigma_{\theta_k}),
\]

holds for any CPTP monotone decreasing functional \(D\). Also, we introduce order structure in the space of these equivalence classes; \([\mathcal{E}] \succeq [\mathcal{F}]\) if and only if

\[
D(\rho_{\theta_1}, \rho_{\theta_2}, \cdots \rho_{\theta_k}) \succeq D(\sigma_{\theta_1}, \sigma_{\theta_2}, \cdots \sigma_{\theta_k})
\]

holds for any \(k\)-point functional \(D\) which is monotone decreasing by CPTP maps.

Obviously, the sequence \(\{[\mathcal{E}_i]\}_{i=0}^\infty\) is monotone decreasing

\[
[\mathcal{E}] = [\mathcal{E}_0] \succeq [\mathcal{E}_1] \succeq \cdots,
\]

and the value of each \(D\) is convergent. But to make above rough statement rigorous, we have to prove the existence of the family

\[
\mathcal{E}_\infty := \{\rho_{\theta,\infty}; \theta \in \Theta\},
\]

such that

\[
\lim_{i \to \infty} D(\rho_{\theta_1,i}, \rho_{\theta_2,i}, \cdots \rho_{\theta_k,i}) = D(\rho_{\theta_1,\infty}, \rho_{\theta_2,\infty}, \cdots \rho_{\theta_k,\infty})
\]

\(2\)
holds for any well-behaved functional $D$, and that
\[
\lim_{i \to \infty} |E_i| = |E_\infty|
\]
holds with respect to a reasonably defined topology.

The line of arguments in this paper is more or less reminiscent of [5]. However, there are some notable differences. First, [5] is dealing with classical Markov processes (over the finite set), while we are dealing with its quantum counterpart. Second, in [5], $\Theta$ is a finite set; an initial state is concentrated at one of the site. Due to these two, [5] can utilize Blackwell measure [9], for which there is one-to-one correspondence with an equivalence class of families of probability distributions over all measurable spaces. In quantum case, however, the counterpart of Blackwell measure so far proposed is a state over a very huge algebra [2], and thus not handy to deal with. Hence, we prefer to treat the equivalence classes directly, rather than using the quantum version of Blackwell measure.

3 Equivalent classes of finite dimensional state families

Let $\mathcal{B}(\mathcal{H})$ and $\mathcal{S}(\mathcal{H})$ be the set of operators and density operators over $\mathcal{H}$, respectively. Let $\mathcal{C}(\mathcal{H})$ denote CPTP maps from $\mathcal{B}(\mathcal{H})$ to itself. Let $\mathcal{H} := \mathbb{C}^d$, and $\Theta$ be a set. Denote by $\mathcal{S}(\mathcal{H})^\Theta$ the set of all families of states in $\mathcal{H}$ parameterized by $\theta \in \Theta$.

Let introduce preorder $\unlhd$ to $\mathcal{S}(\mathcal{H})^\Theta$ : Given $\mathcal{E} := \{\rho_\theta; \theta \in \Theta\}$, $\mathcal{F} := \{\sigma_\theta; \theta \in \Theta\} \in \mathcal{S}(\mathcal{H})^\Theta$, we write $\mathcal{E} \unlhd \mathcal{F}$ if and only if
\[
\Lambda (\mathcal{E}) = \mathcal{F}, \quad \exists \Lambda \in \mathcal{C}(\mathcal{H}) ,
\]
with
\[
\Lambda (\mathcal{E}) := \{\Lambda (\rho_\theta); \theta \in \Theta\}.
\]
(4) holds if and only if (2) holds for any $k$-point CPTP monotone non-increasing functional $D$ with (24) and for any $k$ [7]. Thus, definition here is the same as the one mentioned in the introduction.

Introduce equivalence relation $\equiv$ in $\mathcal{S}(\mathcal{H})^\Theta$ as follows:
\[
\mathcal{E} \equiv \mathcal{F} \iff \mathcal{E} \unlhd \mathcal{F}, \mathcal{F} \unlhd \mathcal{E} .
\]
(5) We denote by $\Xi (\Theta, \mathcal{H})$ the totality of this equivalence classes. $[\mathcal{E}]$ denotes the equivalence class to which $\mathcal{E}$ belongs.

Introduce pseudo metric $\Delta$ on $\mathcal{S}(\mathcal{H})^\Theta$ as follows:
\[
\Delta (\mathcal{E}, \mathcal{F}) := \max \{\delta (\mathcal{E}, \mathcal{F}), \delta (\mathcal{F}, \mathcal{E})\}, \quad \delta (\mathcal{E}, \mathcal{F}) := \inf_{\Lambda \in \mathcal{C}(\mathcal{H})} \sup_{\theta \in \Theta} \|\Lambda (\rho_\theta) - \sigma_\theta\|_1 .
\]

3
where \( \|A\|_1 = \text{tr} \sqrt{A^* A} \). Observe, by (4),

\[
\Delta (\mathcal{E}, \mathcal{F}) = 0 \iff \mathcal{E} \equiv \mathcal{F},
\]

\[
\mathcal{E} \equiv \mathcal{E}', \mathcal{F} \equiv \mathcal{F}' \implies \delta (\mathcal{E}, \mathcal{F}) = \delta (\mathcal{E}', \mathcal{F}').
\]

(6)

Therefore, each of \( \delta \) and \( \Delta \) naturally defines a two point functional in \( \mathbb{E}(\Theta, \mathcal{H}) \), which is also denoted by \( \delta \) and \( \Delta \):

\[
\delta ([\mathcal{E}], [\mathcal{F}]) := \delta (\mathcal{E}, \mathcal{F}),
\]

\[
\Delta ([\mathcal{E}], [\mathcal{F}]) := \Delta (\mathcal{E}, \mathcal{F}).
\]

The topology over \( \mathbb{E}(\Theta, \mathcal{H}) \) indeed by the metric \( \Delta \) is called strong topology.

By definition, we have

\[
\delta (\mathcal{E}, \mathcal{F}) \leq \delta (\mathcal{E}, \mathcal{E}') + \delta (\mathcal{E}', \mathcal{F}),
\]

\[
\delta (\mathcal{E}, \mathcal{F}) = 0 \iff \mathcal{E} \succeq \mathcal{F}.
\]

(8)

and

\[
\Delta (\mathcal{E}, \mathcal{F}) \leq \Delta (\mathcal{E}, \mathcal{E}') + \Delta (\mathcal{E}', \mathcal{F}),
\]

\[
\Delta (\mathcal{E}, \mathcal{F}) \geq \Delta (\Lambda (\mathcal{E}), \Lambda (\mathcal{F})), \quad \forall \Lambda \in \mathcal{C}(\mathcal{H})
\]

(10)

(11)

Define projection from \( \mathcal{S}(\mathcal{H})^\Theta \) to \( \mathbb{E}(\Theta, \mathcal{H}) \) such that

\[
P : \mathbb{E} \to [\mathcal{E}].
\]

**Lemma 1** Suppose that \( \Theta \) is a finite set. Then, \( \mathbb{E}(\Theta, \mathcal{H}) \) is compact with respect to strong topology.

**Proof.** Define a norm

\[
\|\{X_\theta; \theta \in \Theta\}\|_1 := \max_{\theta \in \Theta} \|X_\theta\|_1,
\]

where \( \{X_\theta; \theta \in \Theta\} \in \mathcal{B}(\mathcal{H})^\Theta \), and equip \( \mathcal{B}(\mathcal{H})^\Theta \) with the topology defined by this norm. Then, since \( \mathcal{B}(\mathcal{H})^\Theta \) is finite dimensional vector space, all the norm are topologically equivalent. Thus, \( \mathcal{S}(\mathcal{H})^\Theta \) is compact with respect to the topology defined above. Also, as is shown below, the projection \( P \) from \( \mathcal{S}(\mathcal{H})^\Theta \) onto \( \mathbb{E}(\Theta, \mathcal{H}) \) is continuous. Therefore, \( \mathbb{E}(\Theta, \mathcal{H}) \) is compact.

Continuity of \( P \) is proved as follows. Denote

\[
\mathcal{E} - \mathcal{F} := \{\rho_\theta - \sigma_\theta; \theta \in \Theta\},
\]

where \( \mathcal{E} := \{\rho_\theta; \theta \in \Theta\} \) and \( \mathcal{F} := \{\sigma_\theta; \theta \in \Theta\} \). Observe

\[
\|\mathcal{E} - \mathcal{F}\|_1 \geq \Delta ([\mathcal{E}], [\mathcal{F}]).
\]

(12)

Also observe, for any point \([\mathcal{E}] \in \mathbb{E}(\Theta, \mathcal{H})\) in an open set \( O \subset \mathbb{E}(\Theta, \mathcal{H}) \), there is \( \varepsilon \) with

\[
\{[\mathcal{F}]; \Delta ([\mathcal{E}], [\mathcal{F}] < \varepsilon \} \subset O.
\]
Therefore, by (12),
\[
P^{-1}(O) \supset P^{-1}\left(\{[\mathcal{F}] : \Delta([\mathcal{E}],[\mathcal{F}]) < \varepsilon\}\right),
\]
\[
= \{\mathcal{F} : \Delta(\mathcal{E},\mathcal{F}) < \varepsilon\},
\]
\[
\supset \{\mathcal{F} : ||\mathcal{E} - \mathcal{F}\|_1 < \varepsilon\},
\]
which means \(P^{-1}(O)\) is open. Therefore, \(P\) is continuous. \(\blacksquare\)

**Remark 2** It is may be worthwhile to mention that the partial order "\(\succeq\)" has a good operational meaning. That is, \(\mathcal{E} \succeq \mathcal{F}\) holds if and only if, for any task defined on the parameter set \(\Theta\), the optimal gain is always larger in \(\mathcal{E}\) than in \(\mathcal{F}\) [7].

### 4 Convergence of sequences of equivalence classes

Given a sequence of CPTP maps \(\Gamma_i \in \mathcal{C}(\mathcal{H})\) \((i = 1, 2, \cdots)\), define recursively,
\[
\mathcal{E}_0 := \mathcal{E} = \{\rho_0 ; \theta \in \Theta\},
\]
\[
\mathcal{E}_1 := \{\rho_{0,1} ; \theta \in \Theta\}, \quad \rho_{0,1} := \Gamma_1(\rho_0),
\]
\[
\mathcal{E}_2 := \{\rho_{0,2} ; \theta \in \Theta\}, \quad \rho_{0,2} := \Gamma_2(\rho_{0,1}),
\]
and so on, and consider the sequence \([\mathcal{E}_i]_i^{\infty}\).

**Theorem 3** Let \([\mathcal{E}_i]_i^{\infty}\) be defined as above, and \(\Theta\) be any set. Then, there is \(\mathcal{E}_\infty = \{\rho_\infty ; \theta \in \Theta\}\) such that
\[
\lim_{i \to \infty} \Delta([\mathcal{E}_i],[\mathcal{E}_\infty]) = 0,
\]
\[
\mathcal{E} = \mathcal{E}_0 \succeq \mathcal{E}_1 \succeq \cdots \succeq \mathcal{E}_\infty,
\]
\[
\Delta([\mathcal{E}_{i1}],[\mathcal{E}_{i2}]) \geq \Delta([\mathcal{E}_{i1}],[\mathcal{E}_\infty]), \quad i_1 \leq i_2
\]

**Proof.** This proof is much draws upon the one of Lemma 2.1 of [6]. Let \(\Xi\) be a set with \(|\Xi| = d^2\), and \(\{\rho^\xi ; \xi \in \Xi\}\) be a basis of the space of Hermitian operators over \(\mathcal{H}\) viewed as a real vector space. Then, define
\[
\tilde{\mathcal{E}}_0 := \{\rho^\xi ; \xi \in \Xi\},
\]
\[
\tilde{\mathcal{E}}_1 := \{\rho_1^\xi ; \xi \in \Xi\}, \quad \rho_1^\xi := \Gamma_1(\rho^\xi),
\]
\[
\tilde{\mathcal{E}}_2 := \{\rho_2^\xi ; \xi \in \Xi\}, \quad \rho_2^\xi := \Gamma_2(\rho_1^\xi),
\]
and so on, and consider the sequence of equivalence classes of state families \(\{[\tilde{\mathcal{E}}_i]_i^{\infty}\}\), where the equivalence class is defined by [5].

Due to Lemmas [4] and [9] there is an accumulation point of the set \(\{[\tilde{\mathcal{E}}_i]_i^{\infty}\}\). Let that accumulation point be \(\tilde{\mathcal{E}}_\infty\), where
\[
\tilde{\mathcal{E}}_\infty = \{\rho_\infty^\xi ; \xi \in \Xi\}.
\]
Since $\mathbb{E}(\Xi, \mathcal{H})$ is topologized by the topology based on metric $\Delta$, it satisfies the first axiom of countability, due to Lemma 17. Therefore, by Lemma 18, there is a subsequence $\{n_i\}_{i=1}^{\infty}$ such that
\[
\lim_{i \to \infty} \Delta\left(\tilde{E}_{n_i}, \tilde{E}_{\infty}\right) = 0. \tag{16}
\]
Since $\{\rho^\xi; \xi \in \Xi\}$ is a basis of $\mathcal{S}(\mathcal{H})$, there are real valued functions $\alpha_{\theta, \xi}$ with
\[
\rho_\theta = \sum_{\xi \in \Xi} \alpha_{\theta, \xi} \rho^\xi,
\]
and $\{\rho^\xi; \xi \in \Xi\}$ be the dual base,
\[
\alpha_{\theta, \xi} = \text{tr} \rho_{\theta} \rho^\xi.
\]
Then,
\[
|\alpha_{\theta, \xi}| \leq \|\rho_{\theta}\|_1 \|\rho^\xi\| = \|\rho^\xi\|. \tag{17}
\]
Since $\Xi$ is a finite set,
\[
\sup_{\theta \in \Theta, \xi \in \Xi} |\alpha_{\theta, \xi}| < \infty.
\]
Since $\Gamma_i (i = 1, 2, \cdots)$ are linear,
\[
\rho_{\theta,i} = \sum_{\xi \in \Xi} \alpha_{\theta, \xi} \rho^\xi_i.
\]
Define
\[
\rho_{\theta,\infty} := \sum_{\xi \in \Xi} \alpha_{\theta, \xi} \rho^\xi_{\infty}.
\]
Observe, by (11), if $i_1 \leq i_2$,
\[
\delta(E_{i_1}, E_{i_2}) = 0. \tag{18}
\]
Therefore, by (18),
\[
\delta(E_{\infty}, E_{i_1}) \leq \delta(E_{\infty}, E_{i_n}) + \delta(E_{i_1}, E_{i_2}) = \delta(E_{\infty}, E_{i_1}). \tag{19}
\]
Therefore, by choosing $j$ so that $n_j \leq i$, we have
\[
\delta(E_{\infty}, E_{i}) \leq \delta(E_{\infty}, E_{n_j})
\]
\[
= \inf_{\Lambda \in C(\mathcal{H})} \sup_{\theta \in \Theta} \left\| \Lambda \left( \sum_{\xi \in \Xi} \alpha_{\theta, \xi} \rho^\xi_{\infty} \right) - \sum_{\xi \in \Xi} \alpha_{\theta, \xi} \rho^\xi_{n_j} \right\|_1
\]
\[
\leq \inf_{\Lambda \in C(\mathcal{H})} \sup_{\theta \in \Theta, \xi \in \Xi} \left\| \Lambda \left( \rho^\xi_{\infty} \right) - \rho^\xi_{n_j} \right\|_1
\]
\[
\leq d^2 \sup_{\theta \in \Theta, \xi \in \Xi} |\alpha_{\theta, \xi}| \inf_{\Lambda \in C(\mathcal{H})} \sup_{\xi \in \Xi} \left\| \Lambda \left( \rho^\xi_{\infty} \right) - \rho^\xi_{n_j} \right\|_1
\]
\[
= d^2 \sup_{\theta \in \Theta, \xi \in \Xi} |\alpha_{\theta, \xi}| \delta\left(\tilde{E}_{\infty}, \tilde{E}_{n_j}\right),
\]
which, combined with (16) and (17), leads to
\[
\lim_{i \to \infty} \delta (\mathcal{E}_i, \mathcal{E}_i) \leq d^2 \sup_{\theta \in \Theta, \xi \in \Xi} |\alpha_{\theta, \xi}| \lim_{n_j \to \infty} \delta (\tilde{\mathcal{E}}_{n_j}, \tilde{\mathcal{E}}_{n_j}) = 0.
\]
(20)

Similarly, for any \( i \), taking \( j \) large so that \( n_j \geq i \) holds, we have
\[
\delta (\mathcal{E}_i, \mathcal{E}_\infty) \leq \delta (\mathcal{E}_i, \mathcal{E}_{n_j}) + \delta (\mathcal{E}_{n_j}, \mathcal{E}_\infty)
= \inf_{\Lambda \in \mathcal{C}(\mathcal{H})_{\theta, \xi}} \sup_{\rho, \rho' \in \mathcal{S} (\mathcal{H})} \left\| \Lambda \left( \sum_{\xi \in \Xi} \alpha_{\theta, \xi} \rho^\xi_n \right) - \sum_{\xi \in \Xi} \alpha_{\theta, \xi} \rho^\xi_\infty \right\|_1
\leq d^2 \sup_{\theta \in \Theta, \xi \in \Xi} |\alpha_{\theta, \xi}| \delta (\tilde{\mathcal{E}}_{n_j}, \tilde{\mathcal{E}}_{\infty}),
\]
which, with the help of (16) and (17), leads to
\[
\lim_{j \to \infty} \delta (\tilde{\mathcal{E}}_{n_j}, \tilde{\mathcal{E}}_\infty) = 0.
\]
(21)

Combining (20) and (21) leads to (13). (18) and (21) implies (14). (19) and (21) leads to (15).

We say a quantum Markov process is weakly ergodic if and only if the state tends to be independent of the initial state, or
\[
\lim_{i \to \infty} \sup_{\rho, \rho' \in \mathcal{S} (\mathcal{H})} \| \Gamma_i \circ \cdots \circ \Gamma_2 \circ \Gamma_1 (\rho) - \Gamma_i \circ \cdots \circ \Gamma_2 \circ \Gamma_1 (\rho') \|_1 = 0.
\]

Weak ergodicity, by definition, is equivalent to the convergence to one-point family \( \mathcal{E}_\star := \{ \rho_\star ; \theta \in \Theta \} \). This means the information about the initial state is completely lost.

**Theorem 4** A quantum Markov process is weakly ergodic if and only if \( \mathcal{E}_\star = \mathcal{E}_\infty \).

**Proof.** Suppose weaky ergodicity holds. Fix \( \theta_0 \in \Theta \), and let \( \Lambda_i \) be a CPTP map such that \( \Lambda_i (\rho_\star) = \rho_{\theta_0, i} \). Then,
\[
\delta (\mathcal{E}_i, \mathcal{E}_\star) = \inf_{\Lambda} \| \Lambda (\rho_\star) - \rho_{\theta_0, i} \|_1
\leq \| \Lambda_i (\rho_\star) - \rho_{\theta_0, i} \|_1 = \| \rho_{\theta_0, i} - \rho_{\theta_0, i} \|_1 \to 0.
\]

Define a CPTP map \( \Lambda' \) such that
\[
\Lambda' (\rho) = \rho_\star, \forall \rho \in \mathcal{S} (\mathcal{H}).
\]
Then
\[ \delta (E_i, E_\ast) = \inf_\Lambda \| \Lambda (\rho_{\theta,i}) - \rho_\ast \|_1 \leq \| \Lambda' (\rho_{\theta,i}) - \rho_\ast \|_1 = 0. \]

Therefore, we have
\[ \Delta (E_\ast, E_i) \to 0. \]

Since the strong topology is based on the distance \( \Delta \), it is a Hausdorff space. Therefore, any sequence has at most one convergent point. Hence, \( E_\ast = E_\infty \).

Conversely, suppose \( E_\ast = E_\infty \). Then,
\[
\sup_{\theta, \theta' \in \Theta} \| \rho_{\theta,i} - \rho_{\theta',i} \|_1 \leq \sup_{\theta, \theta' \in \Theta} (\| \Lambda (\rho_\ast) - \rho_{\theta,i} \|_1 + \| \Lambda (\rho_\ast) - \rho_{\theta',i} \|_1)
= 2 \sup_{\theta \in \Theta} \| \Lambda (\rho_\ast) - \rho_{\theta,i} \|_1
\]

Since this holds for any \( \Lambda \), we have
\[
\sup_{\theta, \theta' \in \Theta} \| \rho_{\theta,i} - \rho_{\theta',i} \|_1 \leq 2 \inf_{\theta \in \Theta} \| \Lambda (\rho_\ast) - \rho_{\theta,i} \|_1
\leq 2 \Delta (E_\infty, E_i) \to 0.
\]

Letting \( E = \mathcal{S} (\mathcal{H}) \), we have weak ergodicity. Thus the proof is complete. \( \blacksquare \)

Finally, we show \( [E_\infty] \) is a fixed point if the Markov process is homogeneous.

**Theorem 5** Suppose \( \Gamma_i = \Gamma \ (i = 1, 2, \cdots) \). Then,
\[ \Gamma (E_\infty) = E_\infty. \]

**Proof.** Suppose
\[ \Delta (\Gamma (E_\infty), E_\infty) = c \geq 0. \]

Choose \( i \) so that
\[ \Delta (E_i, E_\infty) \leq \frac{c}{3}. \tag{22} \]

By (15),
\[ \Delta (E_{i+1}, E_\infty) \leq \frac{c}{3}. \tag{23} \]

Then,
\[
\Delta (E_i, E_\infty) \geq \Delta (\Gamma (E_i), \Gamma (E_\infty)) = \Delta (E_{i+1}, \Gamma (E_\infty)) \geq \Delta (E_\ast, E_i) - \Delta (E_{i+1}, E_\ast) \geq c - c/3 = 2c/3,
\]

where the inequality in the first line is by (11), the one in the third line is by (10), and the one in the fourth line is by (23).

This, combined with (22), leads to
\[ \Delta (\Gamma (E_\infty), E_\infty) = c = 0. \]

Thus we have the theorem. \( \blacksquare \)
5 Limits of information quantities

Theorem 6 Consider a k-point function
\[ D : \mathcal{S}(\mathcal{H}) \times \mathcal{S}(\mathcal{H}) \times \cdots \times \mathcal{S}(\mathcal{H}) \to \mathbb{R} \cup \{0\} \]
which is monotone decreasing by CPTP maps. Suppose that
\[
|D(X_1, X_2, \cdots, X_k) - D(Y_1, Y_2, \cdots, Y_k)| \\
\leq f(\|X_1 - Y_1\|_1, \|X_2 - Y_2\|_1, \cdots, \|X_k - Y_k\|_1) \tag{24}
\]
holds for any \(X_j, Y_j \in \mathcal{S}(\mathcal{H})\) \((j = 1, 2, \cdots, k)\), with \(f\) being continuous and
\[
f(0, 0, \cdots, 0) = 0. \tag{25}
\]
Then we have
\[
\lim_{i \to \infty} D(\rho_{\theta_1,i}, \cdots, \rho_{\theta_k,i}) = D(\rho_{\theta_1,\infty}, \cdots, \rho_{\theta_k,\infty}). \tag{26}
\]

Proof. Due to monotonicity and positivity of \(D\), the sequence \(\{D(\rho_{\theta_1,i}, \cdots, \rho_{\theta_k,i})\}_{i=1}^{\infty}\)
is monotone decreasing and bounded from below. Therefore, this sequence converges.

By (13),
\[
D(\rho_{\theta_1,i}, \cdots, \rho_{\theta_k,i}) \geq D(\rho_{\theta_1,\infty}, \cdots, \rho_{\theta_k,\infty}). \tag{27}
\]
Also,
\[
\lim_{i \to \infty} \{D(\rho_{\theta_1,i}, \cdots, \rho_{\theta_k,i}) - D(\rho_{\theta_1,\infty}, \cdots, \rho_{\theta_k,\infty})\} \\
\leq \lim_{i \to \infty} \inf_{\Lambda \in \mathcal{C}(\mathcal{H})} \{D(\rho_{\theta_1,i}, \cdots, \rho_{\theta_k,i}) - D(\Lambda(\rho_{\theta_1,\infty}), \cdots, \Lambda(\rho_{\theta_k,\infty}))\} \\
\leq \lim_{i \to \infty} \inf_{\Lambda \in \mathcal{C}(\mathcal{H})} f(\|\rho_{\theta_1,i} - \Lambda(\rho_{\theta_1,\infty})\|_1, \cdots, \|\rho_{\theta_k,i} - \Lambda(\rho_{\theta_k,\infty})\|_1) \\
\leq \lim_{i \to \infty} f(\inf_{\Lambda \in \mathcal{C}(\mathcal{H})} \|\rho_{\theta_1,i} - \Lambda(\rho_{\theta_1,\infty})\|_1, \cdots, \inf_{\Lambda \in \mathcal{C}(\mathcal{H})} \|\rho_{\theta_k,i} - \Lambda(\rho_{\theta_k,\infty})\|_1) \\
= f(\inf_{\Lambda \in \mathcal{C}(\mathcal{H})} \|\rho_{\theta_1,i} - \Lambda(\rho_{\theta_1,\infty})\|_1, \cdots, \inf_{\Lambda \in \mathcal{C}(\mathcal{H})} \|\rho_{\theta_k,i} - \Lambda(\rho_{\theta_k,\infty})\|_1) \\
= 0,
\]
where the first inequality is by the fact that \(D\) is monotone decreasing by CPTP maps, the one in the third line is by (23), the identity in the fifth line is due to continuity of \(f\), and the inequality in the sixth line is by (13) and (25). Therefore,
\[
\lim_{i \to \infty} D(\rho_{\theta_1,i}, \cdots, \rho_{\theta_k,i}) \leq D(\rho_{\theta_1,\infty}, \cdots, \rho_{\theta_k,\infty}). \tag{28}
\]
Combining (27) and (28), we obtain (26).

Example 7 With
\[
D(\rho_1, \rho_2) := \|\rho_1 - \rho_2\|_1,
\]
the premise of Theorem 6 is obviously satisfied.
Example 8 Let
\[ F(\rho_1, \rho_2) := \frac{1}{\sqrt{\rho_1 \rho_2}} \leq 1. \]

Recall
\[
1 - F(\rho_1, \rho_2) \leq \frac{1}{2} \|\rho_1 - \rho_2\|_1 \leq \sqrt{1 - F(\rho_1, \rho_2)}, \tag{29}
\]
\[
\cos^{-1} F(\rho_1, \rho_2) \leq \cos^{-1} F(\rho'_1, \rho_2) + \cos^{-1} F(\rho_1, \rho'_1). \tag{30}
\]

By (30),
\[
F(\rho_1, \rho_2) \geq \cos(\cos^{-1} F(\rho'_1, \rho_2) + \cos^{-1} F(\rho_1, \rho'_1))
\]
\[= F(\rho'_1, \rho_2)F(\rho_1, \rho'_1) - (\sin \cos^{-1} F(\rho'_1, \rho_2)) (\sin \cos^{-1} (F(\rho_1, \rho'_1))) \]
\[\geq F(\rho'_1, \rho_2)F(\rho_1, \rho'_1) - \sin \cos^{-1} (F(\rho_1, \rho'_1)) \]
\[= F(\rho'_1, \rho_2)F(\rho_1, \rho'_1) - \sqrt{1 - (F(\rho_1, \rho'_1))^2}. \]

Therefore, by (29),
\[
F(\rho'_1, \rho_2) - F(\rho_1, \rho_2) \leq F(\rho'_1, \rho_2) \{1 - F(\rho_1, \rho'_1)\} + \sqrt{1 - (F(\rho_1, \rho'_1))^2}
\]
\[\leq \frac{1}{2} \|\rho_1 - \rho'_1\|_1 + \sqrt{2 \cdot \frac{1}{2} \|\rho_1 - \rho'_1\|_1}. \]

Exchanging \(\rho_1\) and \(\rho'_1\),
\[
F(\rho_1, \rho_2) - F(\rho'_1, \rho_2) \leq \frac{1}{2} \|\rho_1 - \rho'_1\|_1 + \sqrt{\|\rho_1 - \rho'_1\|_1}. \]

Therefore,
\[
|F(\rho_1, \rho_2) - F(\rho'_1, \rho_2)| \leq \frac{1}{2} \|\rho_1 - \rho'_1\|_1 + \sqrt{\|\rho_1 - \rho'_1\|_1}. \]

By the symmetry \(F(\rho_1, \rho_2) = F(\rho_2, \rho_1)\), we have an analogous upper bound to \(|F(\rho'_1, \rho_2) - F(\rho_1, \rho'_2)|\). Therefore,
\[
|F(\rho_1, \rho_2) - F(\rho'_1, \rho'_2)|
\leq |F(\rho_1, \rho_2) - F(\rho'_1, \rho_2)| + |F(\rho'_1, \rho_2) - F(\rho'_1, \rho'_2)|
\leq \frac{1}{2} \|\rho_1 - \rho'_1\|_1 + \sqrt{\|\rho_1 - \rho'_1\|_1} + \frac{1}{2} \|\rho_2 - \rho'_2\|_1 + \sqrt{\|\rho_2 - \rho'_2\|_1}.
\]

Thus,
\[
D_F(\rho_1, \rho_2) := 1 - F(\rho_1, \rho_2)
\]
satisfies the premise of Theorem 6.

Example 9 Let
\[
D_\alpha(\rho_1, \rho_2) := \frac{4}{1 - \alpha^2} \left(1 - \frac{1 - \alpha}{\sqrt{\rho_1 + \rho_2}}\right), \quad (-1 < \alpha < 1)
\]
which is monotone decreasing by CPTP maps [8]. By Lemmas 20-21
\[
\left\| \rho_1^{1+\alpha} - (\rho_1')^{1+\alpha} \right\| \leq \left\| \rho_1 - \rho_1' \right\|^{1+\alpha}
\]
\[
\leq \left\| \rho_1 - \rho_1' \right\|^{1+\alpha}.
\]
Therefore,
\[
|D_\alpha(\rho_1, \rho_2) - D_\alpha(\rho_1', \rho_2)| = \left| \text{tr} \left\{ \rho_1^{1+\alpha} - (\rho_1')^{1+\alpha} \right\} \rho_2 \right|
\]
\[
\leq \left\| \rho_1^{1+\alpha} - (\rho_1')^{1+\alpha} \right\| \left\| \rho_2 \right\|_1^{1+\alpha}
\]
\[
\leq \left\| \rho_1 - \rho_1' \right\|^{1+\alpha} \text{tr} \rho_2^{1+\alpha}.
\]
Similarly,
\[
|D_\alpha(\rho_1', \rho_2) - D_\alpha(\rho_1', \rho_2')| \leq \left\| \rho_2 - \rho_2' \right\|^{1+\alpha} \text{tr} (\rho_1')^{1+\alpha}.
\]
Therefore,
\[
|D_\alpha(\rho_1, \rho_2) - D_\alpha(\rho_1', \rho_2')| \leq \left( \left\| \rho_1 - \rho_1' \right\|^{1+\alpha} + \left\| \rho_2 - \rho_2' \right\|^{1+\alpha} \right),
\]
and the premise of Theorem 6 is satisfied.

**Example 10** Let \( D(\rho_1, \rho_2) \) be a two point functional satisfying the premise of Theorem 6. Let us define
\[
D_k(\rho_1, \rho_2, \cdots, \rho_k) := \sum_{i,j=1}^k a_{ij} D(\rho_i, \rho_j),
\]
where \( a_{ij} \geq 0 \). Then, \( D_k \) satisfies the premise of Theorem 6.

**Example 11** Let \( D(\rho_1, \rho_2) \) be a two point functional satisfying the premise of Theorem 6. Let us define
\[
\overline{D}_k(\rho_1, \rho_2, \cdots, \rho_k) := \inf_{\overline{\rho} \in \mathcal{S}(\mathcal{H})} \max_{1 \leq j \leq k} D(\rho_j, \overline{\rho}).
\]
Then,
\[
\overline{D}_k(\rho_1, \rho_2, \cdots, \rho_k) - \overline{D}_k(\rho_1', \rho_2', \cdots, \rho_k')
\]
\[
= \inf_{\overline{\rho} \in \mathcal{S}(\mathcal{H})} \sup_{1 \leq j \leq k, 1 \leq j' \leq k} \max \min \left\{ D(\rho_j, \overline{\rho}) - D(\rho'_j, \overline{\rho}') \right\}.
\]
Therefore, letting \( \overline{\rho}_\varepsilon \in \mathcal{S}(\mathcal{H}) \) be a state with
\[
\max_{1 \leq j \leq k} D(\rho_j', \overline{\rho}_\varepsilon) \leq \inf_{\overline{\rho} \in \mathcal{S}(\mathcal{H})} \max_{1 \leq j \leq k} D(\rho'_j, \overline{\rho}) + \varepsilon,
\]
we have
\[ D_k (\rho_1, \rho_2, \cdots, \rho_k) - D_k (\rho'_1, \rho'_2, \cdots, \rho'_k) \leq \max_{1 \leq j \leq k} \min_{1 \leq j' \leq k} \left\{ D (\rho_j, \rho_c) - D (\rho'_j, \rho_c) \right\} + \varepsilon \]
\[ \leq \max_{1 \leq j \leq k} \left\{ D (\rho_j, \rho_c) - D (\rho'_j, \rho_c) \right\} + \varepsilon \]
\[ \leq \sum_{j=1}^{k} \left\{ D (\rho_j, \rho_c) - D (\rho'_j, \rho_c) \right\} + \varepsilon \]
\[ \leq \sum_{j=1}^{k} f \left( \left\| \rho_j - \rho'_j \right\|_1, 0 \right) + \varepsilon. \]

Since \( \varepsilon > 0 \) is arbitrary,
\[ D_k (\rho_1, \rho_2, \cdots, \rho_k) - D_k (\rho'_1, \rho'_2, \cdots, \rho'_k) \leq \sum_{j=1}^{k} f \left( \left\| \rho_j - \rho'_j \right\|_1, 0 \right). \]

Almost analogously, we also have
\[ D_k (\rho_1, \rho_2, \cdots, \rho_k) - D_k (\rho'_1, \rho'_2, \cdots, \rho'_k) \geq - \sum_{j=1}^{k} f \left( \left\| \rho_j - \rho'_j \right\|_1, 0 \right). \]

Therefore, \( D_k \) satisfies the premise of Theorem 6.

6 Classical Markov chains over arbitrary measurable space

We had shown the convergence of quantum Markov chain in case of finite dimensional Hilbert space. Next target maybe the analogous statement for infinite dimensional Hilbert space. Since this is very difficult, instead, we study classical Markov chain, but over the arbitrary measurable space.

Inhomogeneous classical Markov chain with initial probability measure \( P \) is defined by
\[ P_1 := \Gamma_1 (P), P_2 := \Gamma_2 (P_1), \cdots, \]
where \( P_t \) is a probability measure in measurable space \((X, \mathcal{X})\), and \( \Gamma_i \) is a positive linear map from the space \( ca (X, \mathcal{X}) \) of bounded signed measures over \((X, \mathcal{X})\) to \( ca (X, \mathcal{X}) \), such that \( \left\| \Gamma_i (\mu) \right\|_1 = \left\| \mu \right\|_1 \) for any positive element \( \mu \) of \( ca (X, \mathcal{X}) \).

Consider families of probability measures \( \mathcal{E} := \{ P_\theta; \theta \in \Theta \} \) over measurable space \((X, \mathcal{X})\) and \( \mathcal{F} := \{ Q_\theta; \theta \in \Theta \} \) over \((Y, \mathcal{Y})\). Then the relations \( \succeq, \equiv, \) and
two point functions $\Delta, \delta$, are defined in analogy to the ones in Section 3. They also satisfy (6)-(11). The equivalence relation $\equiv$ induces an equivalence class of families of probability measures. We denote by $[E]$ the equivalence class to which $E$ belongs, and $E(\Theta)$ denotes the set of all equivalence classes of probability distribution families parameterized by elements of $\Theta$. (That $E(\Theta)$ is a set and not a proper class is known [4].) For each $\Theta_0 \subset \Theta$, define $E(\Theta_0) := \{ P_{\theta} ; \theta \in \Theta_0 \}$, and denote by $\Pi_{\Theta_0}$ the map which sends $[E] \in E(\Theta)$ to $[E_{\Theta_0}] \in E(\Theta_0)$. Suppose $\Theta_0$ is a finite set, and furnish $E(\Theta_0)$ with the strong topology, i.e., the topology induced by the distance $\Delta$. Then, the weak topology of $E(\Theta)$ is the coarsest topology which makes $\Pi_{\Theta_0}$ continuous for each finite subset $\Theta_0 \subset \Theta$. Put differently, the base of the weak topology is in the form of

$$\bigcap_{\kappa \in K} \Pi_{\Theta_\kappa}^{-1} (U_\kappa) \quad (31)$$

where each $\Theta_\kappa \subset \Theta$ is a finite subset of $\Theta$, $K$ is a set of indeces with $|K| < \infty$, and each $U_\kappa$ is an open set in $E(\Theta_\kappa)$.

**Lemma 12** [4] Let $\Theta$ be a set. Then $E(\Theta)$ is a compact Hausdorff space relative to the weak topology.

**Lemma 13** A sequence $\{[E_i]\}_{i=0}^\infty$ in $E(\Theta)$ converges to $[E_{\infty}]$ if and only if for each finite subset $\Theta_0$, $\{\Pi_{\Theta_0} ([E_i])\}_{i=0}^\infty$ converges to $\Pi_{\Theta_0} ([E_{\infty}])$ relative to the strong topology.

**Proof.** Since ‘only if’ is trivial, we show ‘if’. We show that, for any set in the form of (31), there is $N$ such that

$$\{[E_i]\}_{i=N}^\infty \subset \bigcap_{\kappa \in K} \Pi_{\Theta_\kappa}^{-1} (U_\kappa) \quad (32)$$

assuming that $\{\Pi_{\Theta_\kappa} ([E_i])\}_{i=0}^\infty$ converges to $\Pi_{\Theta_\kappa} ([E_{\infty}])$ for each $\kappa \in K$. By assumption, for any open set $U_\kappa$ in $E(\Theta_\kappa)$, there is $N_\kappa$ such that

$$\{\Pi_{\Theta_\kappa} ([E_i])\}_{i=N_\kappa}^\infty \subset U_\kappa.$$

Therefore,

$$\{[E_i]\}_{i=N_\kappa}^\infty \subset \bigcap_{\kappa \in K} \Pi_{\Theta_\kappa}^{-1} (U_\kappa).$$

Therefore, setting $N := \max_{\kappa \in K} N_\kappa$, we have $32$. Thus, the proof is complete.

We consider families of probability measures

$$\mathcal{E}_0 := \mathcal{E} = \{ P_{\theta} ; \theta \in \Theta \},$$

$$\mathcal{E}_1 := \{ P_{\theta,1} ; \theta \in \Theta \}, P_{\theta,1} = \Gamma_1 (P_{\theta,0}),$$

$$\mathcal{E}_2 := \{ P_{\theta,2} ; \theta \in \Theta \}, P_{\theta,2} = \Gamma_2 (P_{\theta,1}), \cdots,$$

and so on.
Theorem 14 Let \( \{ \mathcal{E}_i \}_{i=0}^{\infty} \) be as above and \( \Theta \) be any set. Then, for any finite subset \( \Theta_0 \), we have

\[
\lim_{i \to \infty} \Delta (\Pi_{\Theta_0} ([\mathcal{E}_i]), \Pi_{\Theta_0} ([\mathcal{E}_\infty])) = 0.
\]

(33)

or equivalently, the sequence \( \{\mathcal{E}_i\}_{i=0}^{\infty} \) in \( \mathbb{E}(\Theta) \) converges relative to the weak topology.

Proof. Since \( \mathbb{E}(\Theta) \) is compact by Lemma 12, there is an accumulation point \( [\mathcal{E}_\infty] \) of the sequence \( \{\mathcal{E}_i\}_{i=0}^{\infty} \). Since \( \Pi_{\Theta_0} \) is continuous for each finite set \( \Theta_0 \), the sequence \( \{\Pi_{\Theta_0} ([\mathcal{E}_i])\}_{i=0}^{\infty} \) is an accumulation point of \( \{\Pi_{\Theta_0} ([\mathcal{E}_i])\}_{i=0}^{\infty} \). Since \( \mathbb{E} (\Theta_0) \) satisfies the first axiom countability, there is a subsequence \( \{n_i\}_{i=0}^{\infty} \) such that

\[
\lim_{i \to \infty} \Delta (\Pi_{\Theta_0} ([\mathcal{E}_{n_i}]), \Pi_{\Theta_0} ([\mathcal{E}_\infty])) = 0.
\]

For any \( i_1 \leq i_2 \)

\[
\delta (\Pi_{\Theta_0} ([\mathcal{E}_{i_1}]), \Pi_{\Theta_0} ([\mathcal{E}_{i_2}])) = 0.
\]

Therefore, taking \( j \) with \( n_j \leq i \),

\[
\delta (\Pi_{\Theta_0} ([\mathcal{E}_\infty]), \Pi_{\Theta_0} ([\mathcal{E}_i])) \leq \delta (\Pi_{\Theta_0} ([\mathcal{E}_\infty]), \Pi_{\Theta_0} ([\mathcal{E}_{n_j}])) + \delta (\Pi_{\Theta_0} ([\mathcal{E}_{n_j}]), \Pi_{\Theta_0} ([\mathcal{E}_i]))
\]

\[
= \delta (\Pi_{\Theta_0} ([\mathcal{E}_\infty]), \Pi_{\Theta_0} ([\mathcal{E}_{n_j}]))
\]

\[
\leq \Delta (\Pi_{\Theta_0} ([\mathcal{E}_{n_j}]), \Pi_{\Theta_0} ([\mathcal{E}_\infty])) \to 0, \quad i \to \infty, n_j \to \infty.
\]

Also, taking \( j \) with \( n_j \geq i \),

\[
\delta (\Pi_{\Theta_0} ([\mathcal{E}_i]), \Pi_{\Theta_0} ([\mathcal{E}_\infty])) \leq \delta (\Pi_{\Theta_0} ([\mathcal{E}_i]), \Pi_{\Theta_0} ([\mathcal{E}_{n_j}])) + \delta (\Pi_{\Theta_0} ([\mathcal{E}_{n_j}]), \Pi_{\Theta_0} ([\mathcal{E}_\infty]))
\]

\[
= \delta (\Pi_{\Theta_0} ([\mathcal{E}_{n_j}]), \Pi_{\Theta_0} ([\mathcal{E}_\infty]))
\]

\[
\leq \Delta (\Pi_{\Theta_0} ([\mathcal{E}_{n_j}]), \Pi_{\Theta_0} ([\mathcal{E}_\infty])) \to 0, \quad n_j \to \infty.
\]

Therefore, we have (33). □

We say a Markov process is weakly ergodic if and only if

\[
\lim_{i \to \infty} \sup_{P, P'} \| \Gamma_i \circ \cdots \circ \Gamma_2 \circ \Gamma_1 (P) - \Gamma_i \circ \cdots \circ \Gamma_2 \circ \Gamma_1 (P') \|_1 = 0,
\]

and \( L^1 \)-weakly ergodic if and only if

\[
\forall P, P', \lim_{i \to \infty} \| \Gamma_i \circ \cdots \circ \Gamma_2 \circ \Gamma_1 (P) - \Gamma_i \circ \cdots \circ \Gamma_2 \circ \Gamma_1 (P') \|_1 = 0
\]

Theorem 15 Let \( \mathcal{E}_* \) be a state family such that

\[
\mathcal{E}_* = \{ P_*; \theta \in \Theta \}.
\]

A Markov process is weakly ergodic if and only if, for any \( \mathcal{E} \), \( \{[\mathcal{E}_i]\}_{i=1}^{\infty} \) converges to \( [\mathcal{E}_*] \) relative to the strong topology. Also, a Markov process is \( L^1 \)-weakly ergodic if and only if, for any \( \mathcal{E} \), \( \{[\mathcal{E}_i]\}_{i=1}^{\infty} \) converges to \( [\mathcal{E}_*] \) relative to the weak topology.

14
Proof. The first statement is proved by the argument almost parallel to the proof of Theorem 13. Thus we only prove the second statement. Suppose the Markov chain is $L_1$-weakly ergodic. Fix a finite subset $\Theta_0 \subset \Theta$, and $\theta_0 \in \Theta_0$. Let $\Lambda_i$ be a CPTP map with $\Lambda_i(P_*) = P_{\theta_0,i}$. Then,

$$\sup_{\theta \in \Theta_0} \| P_{\theta,i} - \Lambda_i(P_*) \|_1$$

$$\leq \sum_{\theta \in \Theta_0} \| P_{\theta,i} - \Lambda_i(P_*) \|_1$$

$$= \sum_{\theta \in \Theta_0} \| P_{\theta,i} - P_{\theta_0,i} \|_1 \to 0,$$

or equivalently,

$$\lim_{i \to \infty} \| \Gamma_i \circ \cdots \circ \Gamma_2 \circ \Gamma_1 (P) - \Gamma_i \circ \cdots \circ \Gamma_2 \circ \Gamma_1 (P') \|_1 = 0.$$

On the other hand, denoting by $\Lambda_*$ the CPTP map with $\Lambda_*(\rho) = \rho_*$ for any $\rho_*$,

$$\delta (\Pi_{\Theta_0} (E_i), \Pi_{\Theta_0} (E_*)) \leq \sup_{\theta \in \Theta_0} \| \Lambda_* (\rho_{\theta,i}) - \rho_* \|_1 = 0.$$

Therefore, we have

$$\lim_{i \to \infty} \Delta (\Pi_{\Theta_0} (E_*), \Pi_{\Theta_0} (E_i)) = 0.$$

Since this is true for any finite subset $\Theta_0 \subset \Theta$, by Lemma 13, $\{E_i\}_{i=1}^\infty$ converges to $[E_*]$ relative to the weak topology.

Next, suppose that, for any $E$, $\{[E_i]\}_{i=1}^\infty$ converges to $[E_*]$ relative to the weak topology. Especially, let $\Theta = \{1,2\}$ and $P_1 = P, P_2 = P'$. Then, since $\Theta$ is finite set, we have

$$\lim_{i \to \infty} \Delta (E_* ,E_i) = 0$$

Then, we have

$$\| \Gamma_i \circ \cdots \circ \Gamma_2 \circ \Gamma_1 (P) - \Gamma_i \circ \cdots \circ \Gamma_2 \circ \Gamma_1 (P') \|_1 \leq \| P_{\theta,i} - \Lambda_i(P_*) \|_1 \to 0,$$

$$\leq \sup_{\theta \in \Theta} \| P_{\theta,i} - \Lambda_i(P_*) \|_1 \leq \inf_{\Lambda_i} \| P_{\theta,i} - \Lambda_i(P_*) \|_1 = \delta (E_* ,E_i) \to 0.$$

Therefore, the Markov chain is $L_1$-weakly ergodic.

Theorem 16 Suppose $\Gamma_i = \Gamma$ ($i = 1,2,\cdots$). Then,

$$\Gamma (E_* ) = E_*.$$
Proof. By (??) and using the almost parallel argument as the proof of Theorem 5 leads to

$$\Pi_{\Theta_0} (\Gamma (E_\infty)) = \Gamma (\Pi_{\Theta_0} (E_\infty)) = \Pi_{\Theta_0} (E_\infty).$$

Observe $E (\Theta)$ is a Hausforff space relative to weak topology by Lemma 12 Therefore, if $\Gamma (E_\infty) \neq E_\infty$, $\Gamma (E_\infty) \in \bigcap_{\kappa \in K} \Pi^{-1}_{\Theta_\kappa} (U_\kappa)$, $E_\infty \in \bigcap_{\kappa \in K'} \Pi^{-1}_{\Theta_\kappa} (U'_\kappa)$, (34)

$$\bigcap_{\kappa \in K} \Pi^{-1}_{\Theta_\kappa} (U_\kappa) \cap \bigcap_{\kappa \in K'} \Pi^{-1}_{\Theta_\kappa} (U'_\kappa)$$

(35)

holds where $\Theta_\kappa$ is a finite subset of $\Theta$, and $K, K'$ is a finite set of indeces, and each $U_\kappa$, $U'_\kappa$ is an open set in $E (\Theta_\kappa)$. For (35) to hold, it is necessary that $K$ and $K'$ share at least one element $\kappa_0$. Also, it is necessary that $U_{\kappa_0} \cap U'_{\kappa_0} = \emptyset$.

By (34), we have to have $\Pi_{\Theta_0} (\Gamma (E_\infty)) \in U_{\kappa_0}$ and $\Pi_{\Theta_0} (E_\infty) \in U'_{\kappa_0}$. Since $E (\Theta_\kappa)$ is a Hausforff space, this means

$$\Pi_{\Theta_0} (\Gamma (E_\infty)) \neq \Pi_{\Theta_0} (E_\infty).$$

This contradicts with the assumption. Therefore, we have to have $\Gamma (E_\infty) \equiv E_\infty$.

7 Discussions

We had found out that any quantum Markov chain ”converges”, if you introduce a proper equivalence class. This equivalence class, as had been pointed out, has good decision theoretic meaning [7]. The mode of convergence in case of finite dimensional Hilbert space is ”strong convergence”, that is, convergence with respect to the metric $\Delta$ (13). But, even for the classical Markov chains, such strong statement does not hold in general. Instead, what we could prove was weak convergence (33). Hence, also in quantum case, this is what we can expect at most. The author conjecture weak convergence (33) holds for any quantum Markov chains over arbitrary Hilbert spaces.

Also, we could characterize weak ergodicity and $L^1$-weak ergodicity in view of convergence of the state family, in case of finite dimensional quantum systems and arbitrary classical systems. The author conjectures the similar assertion should holds for arbitrary quantum systems.

References

[1] R. Bhatia, ”Matrix analysis”, Springer (1996)
A General topology

A set in a topological space is a *neighborhood* of a point $x$ if and only if the set contains an open set to which $s$ belongs. A *base* of the topology is a family of open sets such that for each point $s$, every neighborhood of $s$ contains a member of the family. A topological space satisfies the *first axiom of countability* if the topology has a countable base.

**Lemma 17** (11, Chapter 4 of [3]) Every pseudo-metric space satisfies the first axiom of countability.

A point $s$ is an *accumulation point* of a subset $A$ of a topological space if and only if every neighborhood of $s$ contains points of $A$ other than $s$.

**Lemma 18** (8, Chapter 2 of [3]) Suppose the first axiom of countability is satisfied. Then, $s$ is an accumulation point of a sequence $S$ if and only if there is a subsequence converging to $s$.

By Theorem 5.2 of [3], we have:

**Lemma 19** If a topological space $X$ is compact, each sequence $S$ has an accumulation point.
B  Perturbation of matrix functions

A function $f$ is said to be operator monotone if and only if $f(A) \geq f(B)$ holds for any Hermitian matrices $A, B$ with $A \geq B$.

**Lemma 20** (Theorem V.1.9 of [1]) $f(t) = t^\alpha$ ($0 \leq \alpha \leq 1$) is an operator monotone function on $[0, \infty)$

**Lemma 21** (Theorem X.1.1 of [1]) Let $f$ be an operator monotone function on $[0, \infty)$ such that $f(0) = 0$. Then, for all positive operators $A, B$,

$$\|f(A) - f(B)\| \leq f(\|A - B\|).$$

C