

Exponential Decay of Matrix Φ -Entropies on Markov Semigroups with Applications to Dynamical Evolutions of Quantum Ensembles

HAO-CHUNG CHENG^{1,2}, MIN-HSIU HSIEH², AND MARCO TOMAMICHEL³

¹*Graduate Institute Communication Engineering, National Taiwan University, Taiwan (R.O.C.)*

²*Centre for Quantum Computation and Intelligent Systems,
Faculty of Engineering and Information Technology, University of Technology Sydney, Australia*

³*School of Physics, The University of Sydney, Sydney, Australia*

ABSTRACT. In the study of Markovian processes, one of the principal achievements is the equivalence between the Φ -Sobolev inequalities and an exponential decrease of the Φ -entropies. In this work, we develop a framework of Markov semigroups on matrix-valued functions and generalize the above equivalence to the exponential decay of matrix Φ -entropies. This result also specializes to spectral gap inequalities and modified logarithmic Sobolev inequalities in the random matrix setting. To establish the main result, we define a non-commutative generalization of the carré du champ operator, and prove a de Bruijn's identity for matrix-valued functions.

The proposed Markov semigroups acting on matrix-valued functions have immediate applications in the characterization of the dynamical evolution of quantum ensembles. We consider two special cases of quantum unital channels, namely, the depolarizing channel and the phase-damping channel. In the former, since there exists a unique equilibrium state, we show that the matrix Φ -entropy of the resulting quantum ensemble decays exponentially as time goes on. Consequently, we obtain a stronger notion of monotonicity of the Holevo quantity—the Holevo quantity of the quantum ensemble decays exponentially in time and the convergence rate is determined by the modified log-Sobolev inequalities. However, in the latter, the matrix Φ -entropy of the quantum ensemble that undergoes the phase-damping Markovian evolution generally will not decay exponentially. This is because there are multiple equilibrium states for such a channel.

Finally, we also consider examples of statistical mixing of Markov semigroups on matrix-valued functions. We can explicitly calculate the convergence rate of a Markovian jump process defined on Boolean hypercubes, and provide upper bounds of the mixing time on these types of examples.

1. INTRODUCTION

The core problem when studying dynamical systems is to understand how they evolve as time progresses. For example, we want to understand the equilibrium of a stochastic process. The *Markov semigroup theory* mathematically describes the time evolution of dynamical systems. With a Markov semigroup operator $\{P_t\}_{t \geq 0}$ acting on real-valued functions defined on some Polish space Ω , for example, one can ask: is there an invariant measure μ such that

$$\int_{x \in \Omega} P_t f(x) \mu(dx) = \int_{x \in \Omega} f(x) \mu(dx)$$

holds for all such functions f ? (In the following we use the shorthand $\int f d\mu$ for $\int_{x \in \Omega} f(x) \mu(dx)$.) If there does exist such an invariant measure μ on Ω , then how fast the system evolution $P_t f$ converges to the constant equilibrium $\int f d\mu$ when t goes to infinity? To address these problems, functional inequalities like *spectral gap inequalities* (also called *Poincaré inequalities*) and *logarithmic Sobolev inequalities* (log-Sobolev) play crucial roles [1–6]. More explicitly, the spectral gap inequality with a constant $C > 0$:

$$(1.1) \quad \text{Var}(f) \triangleq \int f^2 d\mu - \left(\int f d\mu \right)^2 \leq C \mathcal{E}(f, f)$$

E-mail address: F99942118@ntu.edu.tw, Min-Hsiu.Hsieh@uts.edu.au, marco.tomamichel@sydney.edu.au.

where $\text{Var}(f)$ denotes the variance of the real-valued function f and $\mathcal{E}(f, f)$ is the “energy” of f (see Section 3 for formal definitions), is equivalent to the so-called \mathbb{L}^2 ergodicity of the semigroup $\{\mathbf{P}_t\}_{t \geq 0}$:

$$(1.2) \quad \text{Var}(\mathbf{P}_t f) \leq e^{-2t/C} \text{Var}(f).$$

On the other hand, the log-Sobolev inequality, which is well-known from the seminal work of Gross [7]:

$$(1.3) \quad \text{Ent}(f^2) \triangleq \int f^2 \log \left(\frac{f^2}{\int f^2 d\mu} \right) d\mu \leq C \mathcal{E}(f, f)$$

is equivalent to an exponential decrease of entropies:

$$(1.4) \quad \text{Ent}(\mathbf{P}_t f) \leq e^{-t/C} \text{Ent}(f).$$

Chafaï [8] generalized the previous results and introduced the classical Φ -entropy functionals to establish the equivalence between exponential decays of the Φ -entropies to the Φ -Sobolev inequalities, which interpolates between spectral gap and log-Sobolev inequalities [9]. Consequently, the optimal constant in those functional inequalities directly determines the convergence rate of the Markov semigroups.

Recently, Chen and Tropp [10] introduced a *matrix Φ -entropy functional* for the matrix-valued function $\mathbf{f} : \Omega \rightarrow \mathbb{C}^{d \times d}$, extending its classical counterpart to include matrix objects, and proved a *subadditive* property. This extension has received great attention, and leads to powerful matrix concentration inequalities [11, 12]. Furthermore, two of the present authors [13] derived a series of matrix Poincaré inequalities and matrix Φ -Sobolev inequalities for the matrix-valued functions. This result partially generalized Chafaï’s work [8, 14].

Equipped with the tools of matrix Φ -entropies [10] and the functional inequalities [13], we are at the position to explore a more general form of dynamical systems; namely those systems consisting of matrix components and their evolution governed by the Markov semigroup:

$$\mathbf{P}_t \mathbf{f}(x) = \int_{y \in \Omega} \mathbf{T}_t(x, dy) \circ \mathbf{f}(y),$$

where $\mathbf{T}_t(x, dy) : \mathbb{C}^{d \times d} \rightarrow \mathbb{C}^{d \times d}$ is a completely positive (CP) map and $\int_{y \in \Omega} \mathbf{T}_t(x, dy)$ is unital. We are able to establish the equivalence conditions for the exponential decay of matrix Φ -entropy functionals.

The contributions of this paper are the following:

- (1) We propose a Markov semigroup acting on matrix-valued functions and define a non-commutative version of the carré du champ operator $\mathbf{\Gamma}$ in Section 3. We obtain the time derivatives of matrix Φ -entropy functionals, a generalization of the de Bruijn’s identity for matrix-valued functions in Proposition 9. The equivalence condition of the exponential decay of matrix Φ -entropy functionals is established in Theorem 12. When Φ is a square function, our result generalizes Eqs. (1.1) and (1.2) to the equivalence condition of matrix spectral gap inequalities (Corollary 13). On the other hand, when $\Phi(u) = u \log u$, we obtain the equivalence between exponential entropy decays and the modified log-Sobolev inequalities (Corollary 14). This is slightly different from Eqs. (1.3) and (1.4).
- (2) We show that the introduced Markov semigroup has a connection with quantum information theory and can be used to characterize the dynamical evolution of quantum ensembles that do not depend on the history. More precisely, when the outputs of the matrix-valued function are restricted to a set of quantum states $\mathbf{f}(x) = \boldsymbol{\rho}_x$ (i.e. positive semi-definite matrices with unit trace), the measure μ together with the function \mathbf{f} yields a *quantum ensemble* $\mathcal{S} \triangleq \{\mu(x), \boldsymbol{\rho}_x\}_{x \in \Omega}$. Its time evolution undergoing the semigroup $\{\mathbf{P}_t\}_{t \geq 0}$ can be described by $\mathcal{S}_t \triangleq \{\mu(x), \int \mathbf{T}_t(x, dy) \circ \boldsymbol{\rho}_y\}_{x \in \Omega}$. Moreover, the matrix Φ -entropy functional coincides the *Holevo quantity* $\chi(\{\mu(x), \boldsymbol{\rho}_x\}_{x \in \Omega}) = \text{Ent}(\mathbf{f})$. Our main theorem hence shows that the Holevo quantity of the ensemble \mathcal{S} exponentially decays through the dynamical process: $\chi(\mathcal{S}_t) \leq e^{t/C} \chi(\mathcal{S})$, where the convergence rate is determined by the modified log-Sobolev inequality¹. This result directly strengthens the celebrated monotonicity of the Holevo quantity [15].

¹Here we assume there exists a unique *invariant measure* (see Section 3 for precise definitions) for the Markov semigroups, which ensures the existence of the average state. We discuss the conditions of uniqueness in Section 5 and 6.

- (3) We study an example of matrix-valued functions defined on a Boolean hypercube $\{0, 1\}^n$ with transition rates p from state 0 to 1 and $(1 - p)$ from 1 to 0 (so-called *Markovian jump process*). In this example, we can explicitly calculate the convergence rate of the Markovian jump process (Theorem 16 and 18) by exploiting the matrix Efron-Stein inequality [13].
- (4) We introduce a random walk of a quantum ensemble, where each vertex of the graph corresponds to a quantum state and the transition rates are determined by the weights of the edges. The time evolution of the ensemble can be described by a statistical mixture of the density operators. By using the Holevo quantity as the entropic measure, our main theorem shows that the states in the ensemble will converge to its equilibrium—the average state of the ensemble. Moreover, we can upper bound the mixing time of the ensemble.

1.1. Related Work. When considering the case of discrete time and finite domain (i.e. Ω is a finite set), our setting reduces to the *discrete quantum Markov* introduced by Gudder [16], and the family $\{\mathsf{T}_t(x, y)\}$ is called the *transition operation matrices* (TOM). We note that this discrete model has been applied in quantum random walks by Attal *et al.* [17–19], and model-checking in quantum protocols by Feng *et al.* [20, 21].

If the state space is a singleton (i.e. $\Omega = \{x\}$) and the trace-preserving property is imposed (i.e. T_t is a quantum channel), our model reduces to the conventional quantum Markov processes (also called the *quantum dynamical semigroups*) [22–26] and the Markov semigroups defined on non-commutative \mathbb{L}_p spaces [27]. This line of research was initiated by Lindblad who studied the time evolution of a quantum state: $\mathsf{T}_t(\rho)$ and was recently extended by Kastoryano *et al.* [24] and Szechr [25, 26] who analyzed its long-term behavior. Subsequently, Olkiewicz and Zegarlinski generalized Gross’ log-Sobolev inequalities [7] to the non-commutative \mathbb{L}_p space [27]. The connections between the quantum Markov processes, hypercontractivity, and the noncommutative log-Sobolev inequalities are hence established [28–34]. The exponential decay properties in non-commutative \mathbb{L}_p spaces are also studied [23, 35–37].

The major differences between our work and the non-commutative setting are the following: (1) the semigroup in the latter is applied on a single quantum state, i.e. $\rho \mapsto \mathsf{T}_t(\rho)$, while the novelty of this work is to propose a Markov semigroup that acts on matrix-valued functions. i.e. $\mathbf{f} \mapsto \mathsf{P}_t \mathbf{f}$. (2) Olkiewicz and Zegarlinski used a \mathbb{L}_p relative entropy:

$$\text{Ent}(\rho) = \frac{1}{d} \text{Tr}[\rho \log \rho] - \frac{1}{d} \text{Tr}[\rho] \log \left(\frac{1}{d} \text{Tr}[\rho] \right)$$

to measure the state. However, every state in the ensemble is endowed with a probability $\mu(x)$ in this work. Thus, we can use the matrix Φ -entropy functionals [10]:

$$H_\Phi(\rho_X) = \sum_{x \in \Omega} \mu(x) \text{Tr}[\rho_x \log \rho_x] - \text{Tr} \left[\left(\sum_x \mu(x) \rho_x \right) \log \left(\sum_x \mu(x) \rho_x \right) \right]$$

as the measure of the ensemble through the dynamical process. In other words, we investigate the time evolution and the long-term behavior of a quantum ensemble instead of a quantum state. The key tools to develop the whole theory require operator algebras (e.g. Fréchet derivatives and operator convex functions), the subadditivity of the matrix Φ -entropy functionals, operator Jensen inequalities [38, 39], and the matrix Φ -Sobolev inequalities [13].

The paper is organized as follows. The notation and basic properties of matrix algebras are presented in Section 2. The Markov semigroups acting on matrix-valued functions are introduced in Section 3. We establish the main results of exponential decays in Section 4. In Section 5 we study the quantum ensemble going through a quantum unital channel and demonstrate the exponential decays of the Holevo quantity. We discuss another example of a statistical mixture of the semigroup in Section 6. We prove an upper bound to the mixing time of a quantum random graph. Finally, we conclude this paper in Section 7.

2. PRELIMINARIES AND NOTATION

2.1. Notation and Definitions. We denote by \mathbb{M}^{sa} the space of all self-adjoint operators on some (separable) Hilbert space. If we restrict to the case of $d \times d$ Hermitian matrices, we refer to the notation

\mathbb{M}_d^{sa} . Denote by Tr the standard trace function. The Schatten p -norm is defined by $\|\mathbf{M}\|_p \triangleq (\text{Tr} |\mathbf{M}|^p)^{1/p}$ for $1 \leq p < \infty$, and $\|\mathbf{M}\|_\infty$ corresponds to the operator norm.

For $\mathbf{A}, \mathbf{B} \in \mathbb{M}^{\text{sa}}$, $\mathbf{A} \succeq \mathbf{B}$ means that $\mathbf{A} - \mathbf{B}$ is positive semi-definite. Similarly, $\mathbf{A} \succ \mathbf{B}$ means $\mathbf{A} - \mathbf{B}$ is positive-definite. Denote by \mathbb{M}^+ (resp. \mathbb{M}_d^+) the positive semi-definite operators (resp. $d \times d$ positive semi-definite matrices). Considering any matrix-valued function $\mathbf{f}, \mathbf{g} : \Omega \rightarrow \mathbb{M}^{\text{sa}}$ defined on some Polish space Ω^2 , we shorthand $\mathbf{f} - \mathbf{g} \succeq \mathbf{0}$ for $\mathbf{f}(x) - \mathbf{g}(x) \succeq \mathbf{0}$, $\forall x \in \Omega$. Throughout the paper, italic boldface letters (e.g. \mathbf{X} or \mathbf{f}) are used to denote matrices or matrix-valued functions.

A linear map $\mathsf{T} : \mathbb{M}^{\text{sa}} \rightarrow \mathbb{M}^{\text{sa}}$ is *positive* if $\mathsf{T}(\mathbf{A}) \succeq \mathbf{0}$ for all $\mathbf{A} \succeq \mathbf{0}$. A linear map $\mathsf{T} : \mathbb{M}^{\text{sa}} \rightarrow \mathbb{M}^{\text{sa}}$ is *completely positive* (CP) if for any \mathbb{M}_d^{sa} , the map $\mathsf{T} \otimes \mathbb{1}$ is positive on $\mathbb{M}^{\text{sa}} \otimes \mathbb{M}^{\text{sa}}$. It is well-known that any CP map $\mathsf{T} : \mathbb{M}^{\text{sa}} \rightarrow \mathbb{M}^{\text{sa}}$ enables a *Kraus decomposition*

$$\mathsf{T}(\mathbf{A}) = \sum_i \mathbf{K}_i \mathbf{A} \mathbf{K}_i^\dagger.$$

The CP map T is *trace-preserving* (TP) if and only if $\sum_i \mathbf{K}_i^\dagger \mathbf{K}_i = \mathbf{I}$ (the identity matrix in \mathbb{M}^{sa}), and is *unital* if and only if $\sum_i \mathbf{K}_i \mathbf{K}_i^\dagger = \mathbf{I}$ (see e.g. [40]). A CPTP map is often called a *quantum channel* or *quantum operation* in quantum information theory [41]. We denote by $|i-1\rangle\langle i-1|$ the zero matrix with the exception that its i -th diagonal entry is 1. The set $\{|0\rangle, |1\rangle, \dots, |d-1\rangle\}$ is the computational basis of the Hilbert space \mathbb{C}^d .

Definition 1 (Matrix Φ -Entropy Functional [10]). Let $\Phi : [0, \infty) \rightarrow \mathbb{R}$ be a convex function. Given any probability space $(\Omega, \Sigma, \mathbb{P})$, consider a positive semi-definite random matrix \mathbf{Z} that is \mathbb{P} -measurable. Its expectation

$$\mathbb{E}[\mathbf{Z}] \triangleq \int_\Omega \mathbf{Z} \, d\mathbb{P} = \int_{x \in \Omega} \mathbf{Z}(x) \, \mathbb{P}(dx)$$

is a bounded matrix in \mathbb{M}^+ . Assume \mathbf{Z} satisfies the integration conditions: $\text{Tr} [\mathbb{E}[\mathbf{Z}]] < \infty$ and $\text{Tr} [\mathbb{E}[\Phi(\mathbf{Z})]] < \infty$. The matrix Φ -entropy functional H_Φ is defined as³

$$H_\Phi(\mathbf{Z}) \triangleq \text{Tr} [\mathbb{E}[\Phi(\mathbf{Z})] - \Phi(\mathbb{E}[\mathbf{Z}])].$$

Define $\mathcal{F} \subseteq \Sigma$ as a sub-sigma-algebra of Σ such that the expectation $\mathbb{E}[\mathbf{Z}|\mathcal{F}]$ satisfying $\int_E \mathbb{E}[\mathbf{Z}|\mathcal{F}] \, d\mathbb{P} = \int_E \mathbf{Z} \, d\mathbb{P}$ for each measurable set $E \in \mathcal{F}$. The conditional matrix Φ -entropy functional is then

$$H_\Phi(\mathbf{Z}|\mathcal{F}) \triangleq \text{Tr} [\mathbb{E}[\Phi(\mathbf{Z})|\mathcal{F}] - \Phi(\mathbb{E}[\mathbf{Z}|\mathcal{F}])].$$

In particular, we denote by $\text{Ent}(\mathbf{Z}) \triangleq H_\Phi(\mathbf{Z})$ when $\Phi(u) \equiv u \log u$ and call it the *entropy functional*.

Theorem 2 (Subadditivity of Matrix Φ -Entropy Functionals [10, Theorem 2.5]). Let $\mathbf{X} \triangleq (X_1, \dots, X_n)$ be a vector of independent random variables taking values in a Polish space. Consider a positive semi-definite random matrix \mathbf{Z} that can be expressed as a measurable function of the random vector \mathbf{X} . Assume the integrability conditions $\text{Tr} [\mathbb{E}[\mathbf{Z}]] < \infty$ and $\text{Tr} [\mathbb{E}[\Phi(\mathbf{Z})]] < \infty$. If $\Phi(u) = u \log u$ or $\Phi(u) = u^p$ for $1 \leq p \leq 2$, then

$$(2.1) \quad H_\Phi(\mathbf{Z}) \leq \sum_{i=1}^n \mathbb{E} [H_\Phi(\mathbf{Z}|X_{-i})],$$

where the random vector $X_{-i} \triangleq (X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$ is obtained by deleting the i -th entry of \mathbf{X} .

²A Polish space is a separable and complete metric space, e.g. a discrete space, \mathbb{R} , or the set of Hermitian matrices.

³Chen and Tropp [10, Definition 2.4] defined the matrix Φ -entropy functional for the random matrix \mathbf{Z} taking values in \mathbb{M}_d^+ with the normalized trace function:

$$H_\Phi(\mathbf{Z}) \triangleq \overline{\text{tr}} [\mathbb{E}[\Phi(\mathbf{Z})] - \Phi(\mathbb{E}[\mathbf{Z}])],$$

where $\overline{\text{tr}}[\cdot] \triangleq \frac{1}{d} \text{Tr}[\cdot]$. In this paper we adopt the standard trace function; however, the results remain valid for $\overline{\text{tr}}$ as well.

2.2. Matrix Algebra. Let \mathcal{U}, \mathcal{W} be Banach spaces. The *Fréchet derivative* of a function $\mathcal{L} : \mathcal{U} \rightarrow \mathcal{W}$ at a point $\mathbf{X} \in \mathcal{U}$, if it exists⁴, is a unique linear mapping $D\mathcal{L}[\mathbf{X}] : \mathcal{U} \rightarrow \mathcal{W}$ such that

$$\lim_{\|\mathbf{E}\|_{\mathcal{U}} \rightarrow 0} \frac{\|\mathcal{L}(\mathbf{X} + \mathbf{E}) - \mathcal{L}(\mathbf{X}) - D\mathcal{L}[\mathbf{X}](\mathbf{E})\|_{\mathcal{W}}}{\|\mathbf{E}\|_{\mathcal{U}}} = 0,$$

where $\|\cdot\|_{\mathcal{U}(\mathcal{W})}$ is a norm in \mathcal{U} (resp. \mathcal{W}). The notation $D\mathcal{L}[\mathbf{X}](\mathbf{E})$ then is interpreted as “the Fréchet derivative of \mathcal{L} at \mathbf{X} in the direction \mathbf{E} ”.

The Fréchet derivative enjoys several properties of usual derivatives.

Proposition 3 (Properties of Fréchet Derivatives [44, Section 5.3]). *Let \mathcal{U}, \mathcal{V} and \mathcal{W} be real Banach spaces.*

1. (Sum Rule) *If $\mathcal{L}_1 : \mathcal{U} \rightarrow \mathcal{W}$ and $\mathcal{L}_2 : \mathcal{U} \rightarrow \mathcal{W}$ are Fréchet differentiable at $\mathbf{A} \in \mathcal{U}$, then so is $\mathcal{L} = \alpha\mathcal{L}_1 + \beta\mathcal{L}_2$ and $D\mathcal{L}[\mathbf{A}](\mathbf{E}) = \alpha \cdot D\mathcal{L}_1[\mathbf{A}](\mathbf{E}) + \beta \cdot D\mathcal{L}_2[\mathbf{A}](\mathbf{E})$.*
2. (Product Rule) *If $\mathcal{L}_1 : \mathcal{U} \rightarrow \mathcal{W}$ and $\mathcal{L}_2 : \mathcal{U} \rightarrow \mathcal{W}$ are Fréchet differentiable at $\mathbf{A} \in \mathcal{U}$ and assume the multiplication is well-defined in \mathcal{W} , then so is $\mathcal{L} = \mathcal{L}_1 \cdot \mathcal{L}_2$ and $D\mathcal{L}[\mathbf{A}](\mathbf{E}) = D\mathcal{L}_1[\mathbf{A}](\mathbf{E}) \cdot \mathcal{L}_2(\mathbf{A}) + \mathcal{L}_1(\mathbf{A}) \cdot D\mathcal{L}_2[\mathbf{A}](\mathbf{E})$.*
3. (Chain Rule) *Let $\mathcal{L}_1 : \mathcal{U} \rightarrow \mathcal{V}$ and $\mathcal{L}_2 : \mathcal{V} \rightarrow \mathcal{W}$ be Fréchet differentiable at $\mathbf{A} \in \mathcal{U}$ and $\mathcal{L}_1(\mathbf{A})$ respectively, and let $\mathcal{L} = \mathcal{L}_2 \circ \mathcal{L}_1$ (i.e. $\mathcal{L}(\mathbf{A}) = \mathcal{L}_2(\mathcal{L}_1(\mathbf{A}))$). Then \mathcal{L} is Fréchet differentiable at \mathbf{A} and $D\mathcal{L}[\mathbf{A}](\mathbf{E}) = D\mathcal{L}_2[\mathcal{L}_1(\mathbf{A})](D\mathcal{L}_1[\mathbf{A}](\mathbf{E}))$.*

For each self-adjoint and bounded operator $\mathbf{A} \in \mathbb{M}^{\text{sa}}$ with the spectrum $\sigma(\mathbf{A})$ and the spectral measure \mathbf{E} , its *spectral decomposition* can be written as $\mathbf{A} = \int_{\lambda \in \sigma(\mathbf{A})} \lambda d\mathbf{E}(\lambda)$. As a result, each scalar function is extended to a *standard matrix function* as follows:

$$f(\mathbf{A}) \triangleq \int_{\lambda \in \sigma(\mathbf{X})} f(\lambda) d\mathbf{E}(\lambda).$$

A real-valued function f is called *operator convex* if for each $\mathbf{A}, \mathbf{B} \in \mathbb{M}^{\text{sa}}$ and $0 \leq t \leq 1$,

$$f(t\mathbf{A}) + f((1-t)\mathbf{B}) \preceq f(t\mathbf{A} + (1-t)\mathbf{B}).$$

Proposition 4 (Operator Jensen’s Inequality for Matrix-Valued Measures [38], [39, Theorem 4.2]). *Let (Ω, Σ) be a measurable space and suppose that $I \subseteq \mathbb{R}$ is an open interval. Assume for every $x \in \Omega$, $\mathbf{K}(x)$ is a (finite or infinite dimensional) square matrix and satisfies*

$$\int_{x \in \Omega} \mathbf{K}(dx) \mathbf{K}(dx)^\dagger = \mathbf{I}$$

(identity matrix in \mathbb{M}^{sa}). *If $\mathbf{f} : \Omega \rightarrow \mathbb{M}^{\text{sa}}$ is a measurable function for which $\sigma(\mathbf{f}(x)) \subset I$, for every $x \in \Omega$, then*

$$\phi \left(\int_{x \in \Omega} \mathbf{K}(dx) \mathbf{f}(x) \mathbf{K}(dx)^\dagger \right) \preceq \int_{x \in \Omega} \mathbf{K}(dx) \phi(\mathbf{f}(x)) \mathbf{K}(dx)^\dagger \mu(dx)$$

for every operator convex function $\phi : I \rightarrow \mathbb{R}$. Moreover,

$$\text{Tr} \left[\phi \left(\int_{x \in \Omega} \mathbf{K}(dx) \mathbf{f}(x) \mathbf{K}(dx)^\dagger \mu(dx) \right) \right] \leq \text{Tr} \left[\int_{x \in \Omega} \mathbf{K}(dx) \phi(\mathbf{f}(x)) \mathbf{K}(dx)^\dagger \mu(dx) \right]$$

for every convex function $\phi : I \rightarrow \mathbb{R}$.

Proposition 5 ([45, Theorem 3.23]). *Let $\mathbf{A}, \mathbf{X} \in \mathbb{M}^{\text{sa}}$ and $t \in \mathbb{R}$. Assume $f : I \rightarrow \mathbb{R}$ is a continuously differentiable function defined on an interval I and assume that the eigenvalues of $\mathbf{A} + t\mathbf{X} \subset I$. Then*

$$\left. \frac{d}{dt} \text{Tr} f(\mathbf{A} + t\mathbf{X}) \right|_{t=t_0} = \text{Tr}[\mathbf{X} f'(\mathbf{A} + t_0\mathbf{X})].$$

⁴We assume the functions considered in the paper are Fréchet differentiable. The readers is referred to [42, 43] for conditions for when a function is Fréchet differentiable.

3. MARKOV SEMIGROUPS ON MATRIX-VALUED FUNCTIONS

We will introduce the theory of Markov semigroups in this section. We particularly focus on the Markov semigroup acting on matrix-valued functions. The reader may find general references of Markov semigroups on real-valued functions in [2–6].

Throughout this paper, we consider a probability space (Ω, Σ, μ) with Ω being a discrete space or a compact connected smooth manifold (e.g. $\Omega \equiv \mathbb{R}$). We consider the Banach space \mathcal{B} of continuous, bounded and Bochner integrable ([46, 47]) matrix-valued functions $\mathbf{f} : \Omega \rightarrow \mathbb{M}^{\text{sa}}$ equipped with the uniform norm (e.g. $\|\mathbf{f}\| \triangleq \sup_{x \in \Omega} \|\mathbf{f}(x)\|_\infty$). We denote the expectation with respect to the measure μ for any measurable function $\mathbf{f} : \Omega \rightarrow \mathbb{M}^{\text{sa}}$ by

$$\mathbb{E}_\mu[\mathbf{f}] \triangleq \int_\Omega \mathbf{f} \, d\mu = \int_{x \in \Omega} \mathbf{f}(x) \mu(dx).$$

where the integral is the Bochner integral. We also instate the integration condition $\text{Tr}[\mathbb{E}_\mu[\mathbf{f}]] \leq \infty$.

Define a *completely positive kernel* $\mathbb{T}_t(x, dy) : \mathbb{M}^{\text{sa}} \rightarrow \mathbb{M}^{\text{sa}}$ to be a family of CP maps on \mathbb{M}^{sa} depending on the parameter $t \in \mathbb{R}_+ \triangleq [0, \infty)$ such that

$$(3.1) \quad \int_{y \in \Omega} \mathbb{T}_t(x, dy) \text{ is a unital map,}$$

and satisfies the *Chapman-Kolmogorov identity*:

$$(3.2) \quad \int_{y \in \Omega} \mathbb{T}_s(x, dy) \circ \mathbb{T}_t(y, dz) = \mathbb{T}_{s+t}(x, dz) \quad \forall s, t \in \mathbb{R}_+.$$

In particular, we can impose the trace-preserving property: $\int_{y \in \Omega} \mathbb{T}_t(x, dy)$ is a unital quantum channel.

The central object investigated in this work is a family of operators $\{\mathbb{P}_t\}_{t \geq 0}$ acting on matrix-valued functions. These operators are called *Markov semigroups* if they satisfy:

Definition 6 (Markov Semigroups on Matrix-Valued Functions). A family of linear operators $\{\mathbb{P}_t\}_{t \geq 0}$ on the Banach space \mathcal{B} is a *Markov semigroup* if and only if it satisfies the following conditions

- (a) $\mathbb{P}_0 = \mathbb{I}$, the identity map on \mathcal{B} (*intitial condition*).
- (b) The map $t \rightarrow \mathbb{P}_t \mathbf{f}$ is a continuous map from \mathbb{R}_+ to \mathcal{B} (*continuity property*).
- (c) The semigroup properties: $\mathbb{P}_t \circ \mathbb{P}_s = \mathbb{P}_{s+t}$, for any $s, t \in \mathbb{R}_+$ (*semigroup property*).
- (d) $\mathbb{P}_t \mathbf{I} = \mathbf{I}$ for any $t \in \mathbb{R}_+$, where \mathbf{I} is the constant identity matrix in \mathbb{M}^{sa} (*mass conservation*).
- (e) If \mathbf{f} is non-negative (i.e. $\mathbf{f}(x) \succeq \mathbf{0}$ for all $x \in \Omega$), then $\mathbb{P}_t \mathbf{f}$ is non-negative for any $t \in \mathbb{R}_+$ (*positivity preserving*).

Given the CP kernel $\{\mathbb{T}_t(x, dy)\}_{t \geq 0}$, the Markov semigroup acting on the matrix-valued function \mathbf{f} is defined by

$$(3.3) \quad \mathbb{P}_t \mathbf{f}(x) \triangleq \int_{y \in \Omega} \mathbb{T}_t(x, dy) \mathbf{f}(y),$$

where $\mathbb{T}_t(x, dy) \mathbf{f}(y)$ refers to the output of the linear map $\mathbb{T}_t(x, dy)$ acting on $\mathbf{f}(y)$. If $\int_{y \in \Omega} \mathbb{T}_t(x, dy)$ is a CPTP map, then it exhibits the *contraction* property of the semigroup (with respect to the norm $\|\cdot\|$ on the Banach space \mathcal{B}):

$$(3.4) \quad \begin{aligned} \|\mathbb{P}_t \mathbf{f}\| &= \sup_{x \in \Omega} \|\mathbb{P}_t \mathbf{f}(x)\|_\infty \\ &= \sup_{x \in \Omega} \left\| \int_{y \in \Omega} \mathbb{T}_t(x, dy) \mathbf{f}(y) \right\|_\infty \\ &\leq \sup_{y \in \Omega} \|\mathbf{f}(y)\|_\infty \\ &= \|\mathbf{f}\|, \end{aligned}$$

where we use fact that the CPTP map is contractive [48].

From the operator Jensen's inequality (Proposition 4) we have the following two inequalities:

$$(3.5) \quad P_t(\phi \circ \mathbf{f}) = P_t \phi(\mathbf{f}) \succeq \phi(P_t \mathbf{f})$$

for any operator convex function ϕ , and

$$(3.6) \quad \text{Tr}[P_t \phi(\mathbf{f})] \geq \text{Tr}[\phi(P_t \mathbf{f})]$$

for any convex function ϕ .

Since the map $t \rightarrow P_t \mathbf{f}$ is continuous (Definition 6); hence, the derivative of the operator P_t with respect to t , i.e. the convergence rate of P_t , is a main focus of the analysis of Markov semigroups. More precisely, we define the *infinitesimal generator* for any Markov semigroup $\{P_t\}_{t \geq 0}$ by

$$(3.7) \quad \mathbf{L}(\mathbf{f}) \triangleq \lim_{t \rightarrow 0^+} \frac{1}{t} (P_t \mathbf{f} - \mathbf{f}).$$

For convenience, we denote by $\mathcal{D}(\mathbf{L})$ the *Dirichlet domain* of \mathbf{L} , which is the set of matrix-valued functions in \mathcal{B} such that the limit in Eq. (3.7) exists. We provide an equivalent condition of $\mathcal{D}(\mathbf{L})$ in Appendix A.

Combined with the linearity of the operators $\{P_t\}_{t \geq 0}$ and the semigroup property, we deduce that the generator \mathbf{L} is the derivative of P_t at any time $t > 0$. That is, for $t, s > 0$,

$$\frac{1}{s} (P_{t+s} - P_t) = P_t \left(\frac{1}{s} (P_s - \mathbb{1}) \right) = \left(\frac{1}{s} (P_s - \mathbb{1}) \right) P_t.$$

Letting $s \rightarrow 0$ shows that

$$(3.8) \quad \frac{\partial}{\partial t} P_t = \mathbf{L} P_t = P_t \mathbf{L}.$$

The above equation combined with Eq. (3.5) implies the following proposition.

Proposition 7. *Let $\{P_t\}_{t \geq 0}$ be a Markov semigroup with the infinitesimal generator \mathbf{L} . For any operator convex function $\phi : \mathbb{R} \rightarrow \mathbb{R}$ and $\mathbf{f} \in \mathcal{D}(\mathbf{L})$, we have*

$$(3.9) \quad \mathbf{L}(\phi(\mathbf{f})) \succeq D\phi[\mathbf{f}](\mathbf{L}\mathbf{f}).$$

Proof. For any $s > 0$, Eq. (3.5) implies

$$(3.10) \quad \frac{1}{s} (P_s \phi(\mathbf{f}) - \phi(P_0 \mathbf{f})) \succeq \frac{1}{s} (\phi(P_s \mathbf{f}) - \phi(P_0 \mathbf{f})).$$

By letting $s \rightarrow 0^+$ and using the chain rule of the Fréchet derivative (Proposition 3), the right-hand side yields

$$(3.11) \quad \begin{aligned} \lim_{s \rightarrow 0^+} \frac{1}{s} (\phi(P_{0+s} \mathbf{f}) - \phi(P_0 \mathbf{f})) &= D\phi[P_t \mathbf{f}] \left(\frac{\partial}{\partial t} P_t \mathbf{f} \right) \Big|_{t=0} \\ &= D\phi[P_t \mathbf{f}](\mathbf{L} P_t \mathbf{f})|_{t=0} \\ &= D\phi[\mathbf{f}](\mathbf{L}\mathbf{f}), \end{aligned}$$

where the second equality follows from Eq. (3.8). In the last line we apply the property $P_0 \mathbf{f} = \mathbf{f}$ (item (a) in Definition 6).

On the other hand, the left-hand side of Eq. (3.10) can be rephrased as

$$(3.12) \quad \begin{aligned} \lim_{s \rightarrow 0^+} \frac{1}{s} (P_s \phi(\mathbf{f}) - \phi(P_0 \mathbf{f})) &= \lim_{s \rightarrow 0^+} \frac{1}{s} (P_s \phi(\mathbf{f}) - \phi(\mathbf{f})) \\ &= \mathbf{L}(\phi(\mathbf{f})). \end{aligned}$$

Hence, combining Eqs. (3.12), (3.10) and (3.11) arrives at the desired inequality

$$\mathbf{L}(\phi(\mathbf{f})) \succeq D\phi[\mathbf{f}](\mathbf{L}\mathbf{f}).$$

□

In the classical setup (i.e. Markov semigroups acting on real-valued functions), the *carré du champ* operator (see e.g. [4]) is defined by :

$$(3.13) \quad \Gamma(f, g) = \frac{1}{2}(\mathsf{L}(fg) - f\mathsf{L}(g) - g\mathsf{L}(f)).$$

Here we introduce a non-commutative version of the carré du champ operator $\Gamma : \mathcal{D}(\mathsf{L}) \times \mathcal{D}(\mathsf{L}) \rightarrow \mathcal{D}(\mathsf{L})$ of the generator L by

$$(3.14) \quad \Gamma(\mathbf{f}, \mathbf{f}) \triangleq \frac{1}{2}(\mathsf{L}(\mathbf{f}^2) - \mathbf{f}\mathsf{L}(\mathbf{f}) - \mathsf{L}(\mathbf{f})\mathbf{f}),$$

and its symmetric and bilinear extension

$$(3.15) \quad \begin{aligned} \Gamma(\mathbf{f}, \mathbf{g}) &= \Gamma(\mathbf{g}, \mathbf{f}) \triangleq \frac{1}{2}(\Gamma(\mathbf{f} + \mathbf{g}, \mathbf{f} + \mathbf{g}) - \Gamma(\mathbf{f}, \mathbf{f}) - \Gamma(\mathbf{g}, \mathbf{g})) \\ &= \frac{1}{4}(\mathsf{L}(\mathbf{f}\mathbf{g}) + \mathsf{L}(\mathbf{g}\mathbf{f}) - \mathbf{f}\mathsf{L}(\mathbf{g}) - \mathbf{g}\mathsf{L}(\mathbf{f}) - \mathsf{L}(\mathbf{f})\mathbf{g} - \mathsf{L}(\mathbf{g})\mathbf{f}). \end{aligned}$$

We note that when \mathbf{f} commutes⁵ with \mathbf{g} , the carré du champ operator reduces to the conventional expression (cf. Eq. (3.13)):

$$\Gamma(\mathbf{f}, \mathbf{g}) \equiv \frac{1}{2}(\mathsf{L}(\mathbf{f}\mathbf{g}) - \mathbf{f}\mathsf{L}(\mathbf{g}) - \mathbf{g}\mathsf{L}(\mathbf{f})).$$

Recall that the square function $\phi(u) = u^2$ is operator convex. The formula of the Fréchet derivative: $\mathsf{D}\phi[\mathbf{A}](\mathbf{B}) = \mathbf{A}\mathbf{B} + \mathbf{B}\mathbf{A}$ together with Proposition 7 yields

$$\mathsf{L}(\mathbf{f}^2) \succeq \mathbf{f}\mathsf{L}(\mathbf{f}) + \mathsf{L}(\mathbf{f})\mathbf{f}.$$

Hence the carré du champ operator is positive semi-definite: $\Gamma(\mathbf{f}, \mathbf{f}) \succeq \mathbf{0}$. We can also observe that $\Gamma(\mathbf{f}, \mathbf{f}) = \mathbf{0}$ implies that \mathbf{f} is essentially constant, i.e. $\mathsf{P}_t\mathbf{f} = \mathbf{f}$ for $t \geq 0$. Moreover, the non-negativity and the bilinearity of the carré du champ operator directly yield a trace Cauchy-Schwartz inequality:

Proposition 8 (Trace Cauchy-Schwartz Inequality for Carré du Champ Operators). *For all $\mathbf{f}, \mathbf{g} \in \mathcal{D}(\mathsf{L})$,*

$$(\mathrm{Tr} [\Gamma(\mathbf{f}, \mathbf{g})])^2 \leq \mathrm{Tr} [\Gamma(\mathbf{f}, \mathbf{f})] \cdot \mathrm{Tr} [\Gamma(\mathbf{g}, \mathbf{g})].$$

Proof. From Eq. (3.15), for all $s \in \mathbb{R}$ it follows that

$$\Gamma(s\mathbf{f} + \mathbf{g}, s\mathbf{f} + \mathbf{g}) = s^2 \cdot \Gamma(\mathbf{f}, \mathbf{f}) + 2s \cdot \Gamma(\mathbf{f}, \mathbf{g}) + \Gamma(\mathbf{g}, \mathbf{g}) \succeq \mathbf{0}$$

After taking trace, the non-negativity of the above equation ensures the discriminant being non-positive:

$$(2 \cdot \mathrm{Tr} [\Gamma(\mathbf{f}, \mathbf{g})])^2 - 4 \cdot \mathrm{Tr} [\Gamma(\mathbf{f}, \mathbf{f})] \cdot \mathrm{Tr} [\Gamma(\mathbf{g}, \mathbf{g})] \leq 0$$

as desired. □

Note that a bilinear map $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$ on some vector space V is a *scalar inner product* if it satisfies conjugate symmetry and non-negativity ($\langle x, x \rangle \geq 0$ and $x = 0$ when $\langle x, x \rangle = 0$). As a result, the non-commutative carré du champ operator that exhibits the properties of symmetry, linearity, and non-negativity ($\Gamma(\mathbf{f}, \mathbf{f}) \succeq \mathbf{0}$ and $\mathsf{L}\mathbf{f} = \mathbf{0}$ when $\Gamma(\mathbf{f}, \mathbf{f}) = \mathbf{0}$) can be viewed as a matrix-valued inner product (with respect to the generator L) on the space $\mathcal{D}(\mathsf{L})$.

Given the semigroup $\{\mathsf{P}_t\}_{t \geq 0}$, the measure μ is *invariant* for the function $\mathbf{f} \in \mathcal{D}(\mathsf{L})$ if

$$(3.16) \quad \int \mathsf{P}_t\mathbf{f}(x) \mu(\mathrm{d}x) = \int \mathbf{f}(x) \mu(\mathrm{d}x), \quad t \in \mathbb{R}_+.$$

We can observe from Eqs. (3.7) and (3.16) that any invariant measure μ satisfies

$$(3.17) \quad \int \mathsf{L}(\mathbf{f}) \mathrm{d}\mu = \mathbf{0}.$$

⁵Here we means that $[\mathbf{f}(x), \mathbf{g}(x)] = \mathbf{0}$ for all $x \in \Omega$.

We call the measure μ *symmetric* if and only if

$$(3.18) \quad \int \mathbf{f} \mathbf{L}(\mathbf{g}) + \mathbf{L}(\mathbf{g}) \mathbf{f} \, d\mu = \int \mathbf{g} \mathbf{L}(\mathbf{f}) + \mathbf{L}(\mathbf{f}) \mathbf{g} \, d\mu,$$

which implies an integration by parts formula:

$$(3.19) \quad -\frac{1}{2} \int \mathbf{f} \mathbf{L}(\mathbf{g}) + \mathbf{L}(\mathbf{g}) \mathbf{f} \, d\mu = \int \mathbf{\Gamma}(\mathbf{f}, \mathbf{g}) \, d\mu = \int \mathbf{\Gamma}(\mathbf{g}, \mathbf{f}) \, d\mu.$$

The notion of carré du champ operator and the invariant measure μ (i.e. Eq. (3.17)) immediately lead to a symmetric bilinear *Dirichlet form*:

$$(3.20) \quad \mathcal{E}(\mathbf{f}, \mathbf{g}) \triangleq \int \mathbf{\Gamma}(\mathbf{f}, \mathbf{g}) \, d\mu = -\frac{1}{4} \int \mathbf{f} \mathbf{L}(\mathbf{g}) + \mathbf{g} \mathbf{L}(\mathbf{f}) + \mathbf{L}(\mathbf{f}) \mathbf{g} + \mathbf{L}(\mathbf{g}) \mathbf{f} \, d\mu.$$

The non-negativity of the carré du champ operator also yields that $\mathcal{E}(\mathbf{f}, \mathbf{f}) = \int \mathbf{\Gamma}(\mathbf{f}, \mathbf{f}) \, d\mu \succeq \mathbf{0}$, which stands for a kind of second-moment quantity or the energy of the function \mathbf{f} . For convenience, we shorthand the notation $\mathbf{\Gamma}(\mathbf{f}) \equiv \mathbf{\Gamma}(\mathbf{f}, \mathbf{f})$ and $\mathcal{E}(\mathbf{f}) \equiv \mathcal{E}(\mathbf{f}, \mathbf{f})$.

In the following, we often refer to the *Markov Triple* $(\Omega, \mathbf{\Gamma}, \mu)$ with state space Ω , carré du champ operator $\mathbf{\Gamma}$ acting on the Dirichlet domain $\mathcal{D}(\mathbf{L})$ of matrix-valued functions, and invariant measure μ . Additionally, we will apply the Fubini's theorem to freely interchange the order of trace and the expectation with respect to μ .

4. MAIN RESULTS: EXPONENTIAL DECAYS OF MATRIX Φ -ENTROPY FUNCTIONALS

In this section, our goal is to show that the matrix Φ -entropy functional exponentially decays along the Markov semigroup and its relation with the spectral gap inequalities and logarithmic Sobolev inequalities. With the invariant measure μ of the semigroup $\{\mathbf{P}_t\}_{t \geq 0}$ and the Jensen inequality (3.6), we observe that

$$\begin{aligned} H_\Phi(\mathbf{P}_t \mathbf{f}) &= \text{Tr} \left[\mathbb{E}_\mu [\Phi(\mathbf{P}_t \mathbf{f})] - \Phi(\mathbb{E}_\mu [\mathbf{P}_t \mathbf{f}]) \right] \\ &= \text{Tr} \left[\mathbb{E}_\mu [\Phi(\mathbf{P}_t \mathbf{f})] - \Phi(\mathbb{E}_\mu \mathbf{f}) \right] \\ &\leq \text{Tr} \left[\mathbb{E}_\mu [\mathbf{P}_t \Phi(\mathbf{f})] - \Phi(\mathbb{E}_\mu \mathbf{f}) \right] \\ &= H_\Phi(\mathbf{f}), \end{aligned}$$

where in the second and the last lines we use the property of the invariant measure μ , Eq. (3.16). Thus, the matrix Φ -entropy functional is non-increasing along the flow of the semigroup and behaves like the classical Φ -entropy functionals (see e.g. [8]). Moreover, we are able to obtain the time differentiation of the matrix Φ -entropy functional, which can be viewed as the *Boltzmann H-Theorem* for matrix-valued functions.

Proposition 9 (de Bruijn's Property for Markov Semigroups). *Fix a probability space (Ω, Σ, μ) . Let $\{\mathbf{P}_t\}_{t \geq 0}$ be a Markov semigroup with infinitesimal generator \mathbf{L} and carré du champ operator $\mathbf{\Gamma}$. Assume that μ is an invariant probability measure for the semigroup. Then, for any suitable matrix-valued function $\mathbf{f} : \Omega \rightarrow \mathbb{M}^{sa}$ in the Dirichlet domain $\mathcal{D}(\mathbf{L})$ with μ being its invariant measure,*

$$(4.1) \quad \frac{\partial}{\partial t} H_\Phi(\mathbf{P}_t \mathbf{f}) = \text{Tr} \mathbb{E}_\mu [\Phi'(\mathbf{P}_t \mathbf{f}) \mathbf{L} \mathbf{P}_t \mathbf{f}] \leq 0, \quad \forall t \in \mathbb{R}_+.$$

When μ is symmetric, one has the following formulation

$$(4.2) \quad \frac{\partial}{\partial t} H_\Phi(\mathbf{P}_t \mathbf{f}) = -\text{Tr} \mathbb{E}_\mu [\mathbf{\Gamma}(\Phi'(\mathbf{P}_t \mathbf{f}), \mathbf{P}_t \mathbf{f})], \quad \forall t \in \mathbb{R}_+.$$

Proof. The proof directly follows from the definition of the matrix Φ -entropy functional and the properties of the Markov semigroup. Namely,

$$\begin{aligned}
\frac{\partial}{\partial t} H_{\Phi}(\mathbf{P}_t \mathbf{f}) &= \frac{\partial}{\partial t} \text{Tr} \left[\mathbb{E}_{\mu} [\Phi(\mathbf{P}_t \mathbf{f})] - \Phi(\mathbb{E}_{\mu} [\mathbf{P}_t \mathbf{f}]) \right] \\
&= \frac{\partial}{\partial t} \text{Tr} \left[\mathbb{E}_{\mu} [\Phi(\mathbf{P}_t \mathbf{f})] - \Phi(\mathbb{E}_{\mu} [\mathbf{f}]) \right] \\
&= \frac{\partial}{\partial t} \text{Tr} \mathbb{E}_{\mu} [\Phi(\mathbf{P}_t \mathbf{f})] \\
&= \text{Tr} \mathbb{E}_{\mu} [\text{D}\Phi[\mathbf{P}_t \mathbf{f}](\mathbf{L}\mathbf{P}_t \mathbf{f})] \\
&= \text{Tr} \mathbb{E}_{\mu} [\Phi'(\mathbf{P}_t \mathbf{f}) \mathbf{L}\mathbf{P}_t \mathbf{f}],
\end{aligned}$$

where the second equality is due to the invariance of μ , Eq. (3.16). The fourth equation is due to the chain rule of Fréchet derivative (see Proposition 3) and Eq. (3.8). We obtain the last identity by Proposition 5.

Proposition 7 yields $\text{D}\Phi[\mathbf{P}_t \mathbf{f}](\mathbf{L}\mathbf{P}_t \mathbf{f}) \preceq \mathbf{L}\Phi(\mathbf{P}_t \mathbf{f})$. By the invariance of μ , we deduce the non-positivity of Eq. (4.1):

$$\mathbb{E}_{\mu} [\text{D}\Phi[\mathbf{P}_t \mathbf{f}](\mathbf{L}\mathbf{P}_t \mathbf{f})] \preceq \mathbb{E}_{\mu} [\mathbf{L}\Phi(\mathbf{P}_t \mathbf{f})] = \mathbf{0}.$$

The symmetric case (4.2) stands by further applying the integration by parts formula, Eq. (3.19), i.e.

$$\begin{aligned}
\text{Tr} \left[\mathbb{E}_{\mu} [\mathbf{\Gamma}(\mathbf{g}, \mathbf{h})] \right] &= -\frac{1}{2} \text{Tr} \left[\mathbb{E}_{\mu} [\mathbf{g} \cdot \mathbf{L}(\mathbf{h}) + \mathbf{L}(\mathbf{h}) \cdot \mathbf{g}] \right] \\
&= -\text{Tr} \left[\mathbb{E}_{\mu} [\mathbf{g} \cdot \mathbf{L}(\mathbf{h})] \right],
\end{aligned}$$

where we apply the cyclic property of the trace function. Hence, Eq. (4.2) follows by taking $\mathbf{g} \equiv \Phi'(\mathbf{P}_t \mathbf{f})$ and $\mathbf{h} \equiv \mathbf{P}_t \mathbf{f}$. \square

In the following, we first give the definitions of the spectral gap inequalities and logarithmic Sobolev inequalities related to Markov semigroups. The main result—the relation between the exponential decays in matrix Φ -entropies and the functional inequalities, is presented in Theorem 12.

Definition 10 (Spectral Gap Inequality for Matrix-Valued Functions). A Markov Triple $(\Omega, \mathbf{\Gamma}, \mu)$ is said to satisfy a spectral gap inequality with a constant $C > 0$, if for all matrix-valued functions $\mathbf{f} : \Omega \rightarrow \mathbb{M}_d^{\text{sa}}$ in the Dirichlet domain $\mathcal{D}(\mathbf{L})$ with μ being its invariant measure,

$$\text{Var}(\mathbf{f}) \leq C \mathcal{E}(\mathbf{f}),$$

where

$$\text{Var}(\mathbf{f}) \triangleq \text{Tr} \mathbb{E}_{\mu} \left[(\mathbf{f} - \mathbb{E}_{\mu}[\mathbf{f}])^2 \right]$$

denotes the variance of the function \mathbf{f} with respect to the measure μ and $\mathcal{E}(\mathbf{f}) \triangleq \text{Tr}[\mathbf{\mathcal{E}}(\mathbf{f})]$. The infimum of the constants among all the spectral gap inequalities is called the *spectral gap constant*.

Definition 11 (Logarithmic Sobolev Inequality for Matrix-Valued Functions). A Markov Triple $(\Omega, \mathbf{\Gamma}, \mu)$ is said to satisfy a logarithmic Sobolev inequality $\text{LS}(C, B)$ with constants $C > 0$, $B \geq 0$, if for all matrix-valued functions $\mathbf{f} : \Omega \rightarrow \mathbb{M}_d^{\text{sa}}$ in the Dirichlet domain $\mathcal{D}(\mathbf{L})$ with μ being its invariant measure,

$$\text{Ent}(\mathbf{f}^2) \leq B \text{Tr} \mathbb{E}_{\mu} [\mathbf{f}^2] + C \mathcal{E}(\mathbf{f}).$$

The logarithmic Sobolev inequality is called *tight* and is denoted by $\text{LS}(C)$ when $B = 0$. When $B > 0$, the logarithmic Sobolev inequality $\text{LS}(C, B)$ is called *defective*.

We also define the *modified logarithmic Sobolev inequality* (MLSI) if there exists a constant C such that

$$\text{Ent}(\mathbf{f}) \leq -C \text{Tr} \mathbb{E}_{\mu} [(\mathbf{I} + \log \mathbf{f}) \mathbf{L}\mathbf{f}].$$

Theorem 12 (Exponential Decay of Matrix Φ -Entropy Functionals of Markov Semigroups). *Given a Markov triple $(\Omega, \mathbf{\Gamma}, \mu)$, the following two statements are equivalent: there exists a Φ -Sobolev constant $C \in (0, \infty]$ such that*

$$(4.3) \quad H_{\Phi}(\mathbf{f}) \leq -C \text{Tr} \mathbb{E}_{\mu} [\Phi'(\mathbf{f}) \mathbf{L}\mathbf{f}],$$

and

$$(4.4) \quad H_\Phi(\mathbf{P}_t \mathbf{f}) \leq e^{-t/C} H_\Phi(\mathbf{f}), \quad \forall t \geq 0$$

for all $\mathbf{f} \in \mathcal{D}(\mathbf{L})$ with μ being its invariant measure.

Proof. The theorem is a consequence of the de Bruijn's property, Proposition 9. More precisely, Eq. (4.1) and the inequality (4.3) imply

$$\frac{\partial}{\partial t} H_\Phi(\mathbf{P}_t \mathbf{f}) = \text{Tr } \mathbb{E}_\mu[\Phi'(\mathbf{P}_t \mathbf{f}) \mathbf{L} \mathbf{P}_t \mathbf{f}].$$

Recall that the function $\mathbf{P}_t \mathbf{f}$ is invariant under the measure μ and therefore satisfies Eq. (4.3). Hence, the above inequality can be rewritten as

$$\begin{aligned} \frac{\partial}{\partial t} H_\Phi(\mathbf{P}_t \mathbf{f}) &= \text{Tr } \mathbb{E}_\mu[\Phi'(\mathbf{P}_t \mathbf{f}) \mathbf{L} \mathbf{P}_t \mathbf{f}] \\ &\leq -\frac{1}{C} H_\Phi(\mathbf{P}_t \mathbf{f}), \end{aligned}$$

from which we obtain

$$H_\Phi(\mathbf{P}_t \mathbf{f}) \leq e^{-t/C} H_\Phi(\mathbf{P}_0 \mathbf{f}) = e^{-t/C} H_\Phi(\mathbf{f}).$$

Conversely, differentiating inequality (4.4) at $t = 0$ gives the desired inequality (4.3). \square

From Theorem 12, we immediately establish the equivalence between the spectral gap inequality and the exponential decay of variance functions.

Corollary 13 (Exponential Decay of Variance and Spectral Gap Inequalities). *A Markov Triple $(\Omega, \mathbf{\Gamma}, \mu)$ satisfies the spectral gap inequality with constant C if and only if, for matrix-valued functions $\mathbf{f} : \Omega \rightarrow \mathbb{M}_d^{sa}$ in the Dirichlet domain $\mathcal{D}(\mathbf{L})$, one has*

$$\text{Var}(\mathbf{P}_t \mathbf{f}) \leq e^{-2t/C} \cdot \text{Var}(\mathbf{f}).$$

Proof. Recall that $H_{u \rightarrow u^2}(\mathbf{f}) \equiv \text{Var}(\mathbf{f})$. Hence, by taking $\Phi(u) = u^2$ in Theorem 12, the corollary follows since the right-hand side of Eq. (4.3) can be rephrased as:

$$\text{Tr } \mathbb{E}_\mu[\Phi'(\mathbf{f}) \mathbf{L} \mathbf{f}] = 2 \text{Tr } \mathbb{E}_\mu[\mathbf{f} \cdot \mathbf{L} \mathbf{f}] = -2 \text{Tr } \mathbb{E}_\mu[\mathbf{\Gamma}(\mathbf{f})] = -2\mathcal{E}(\mathbf{f}).$$

\square

Similarly, by taking $\Phi(u) = u \log u$, we have the equivalence between the modified log-Sobolev inequalities and the exponential decay in entropy functionals.

Corollary 14 (Exponential Decay of Entropy and Modified Log-Sobolev Inequalities). *A Markov Triple $(\Omega, \mathbf{\Gamma}, \mu)$ satisfies the modified log-Sobolev inequality with constant C if and only if, for matrix-valued functions $\mathbf{f} : \Omega \rightarrow \mathbb{M}_d^{sa}$ in the Dirichlet domain $\mathcal{D}(\mathbf{L})$, one has*

$$\text{Ent}(\mathbf{P}_t \mathbf{f}) \leq e^{-t/C} \cdot \text{Ent}(\mathbf{f}).$$

5. TIME EVOLUTIONS OF A QUANTUM ENSEMBLE

In this section, we discuss the applications of the Markov semigroups in quantum information theory and study a special case of the Markov semigroups—quantum unital channel. From the analysis in Section 4, we demonstrate the exponential decays of the matrix Φ -entropy functionals and give a tight bound to the monotonicity of the Holevo quantity: $\chi(\{\mu(x), \mathbf{T}_t(\rho(x))\}_x) \leq e^{-t/C} \chi(\{\mu(x), \rho(x)\}_x)$, when \mathbf{T}_t is a unital quantum dynamical semigroup [22].

To connect the whole machinery to quantum information theory, it is convenient to introduce some basic notation. First of all, we will restrict to the following set of matrix-valued functions in this section: $\mathbf{f}(x) = \rho_x$, where ρ_x is a density operator (i.e. positive semi-definite matrix with unit trace) for all $x \in \Omega$. In other words, the function \mathbf{f} is a classical-quantum encoder that maps the classical message x to a

quantum state ρ_x . Therefore, a quantum ensemble is constructed by the classical-quantum encoder and the given measure μ : $\mathcal{S} \triangleq \{\mu(x), \rho_x\}_{x \in \Omega}$. The Holevo quantity of a quantum ensemble \mathcal{S} is defined as

$$\chi(\mathcal{S}) \triangleq -\text{Tr} [\bar{\rho} \log \bar{\rho}] + \int_{x \in \Omega} \text{Tr} [\rho_x \log \rho_x] \mu(dx),$$

where $\bar{\rho} \triangleq \int \rho_x \mu(dx)$ denotes the average state. It is not hard to verify that the Holevo quantity is a special case of the matrix Φ -entropy functionals [13]: $\chi(\mathcal{S}) = \text{Ent}(\mathbf{f})$.

If we impose the trace-preserving condition on the CP kernel $\int_{y \in \Omega} \mathsf{T}_t(x, dy)$, it becomes a quantum unital channel. Consequently, the Markov semigroup $\{\mathsf{P}_t\}_{t \geq 0}$ acts on the matrix-valued function \mathbf{f} can be interpreted as a time evolution of the quantum ensemble \mathcal{S} , and the results of the matrix-valued functions in Section 4 also work for quantum ensembles.

For each $t \in \mathbb{R}_+$, let the CP kernel be

$$\mathsf{T}_t(x, y) = \begin{cases} \mathsf{T}_t \text{ (a quantum unital channel)}, & \text{if } x = y \\ 0 \text{ (a zero map)}, & \text{if } x \neq y. \end{cases}$$

The set of unital maps $\mathsf{T}_t : \mathbb{M}_d^{\text{sa}} \rightarrow \mathbb{M}_d^{\text{sa}}$ forms a *quantum dynamical semigroup*, which satisfies the semigroup conditions:

- (a) $\mathsf{T}_0(\mathbf{X}) = \mathbf{X}$ for all $\mathbf{X} \in \mathbb{M}_d^{\text{sa}}$.
- (b) The map $t \rightarrow \mathsf{T}_t(\mathbf{X})$ is a continuous map from \mathbb{R}_+ to \mathbb{M}_d^{sa} .
- (c) The semigroup properties: $\mathsf{T}_t \circ \mathsf{T}_s = \mathsf{T}_{s+t}$, for any $s, t \in \mathbb{R}_+$.
- (d) $\mathsf{T}_t(\mathbf{I}) = \mathbf{I}$ for any $t \in \mathbb{R}_+$, where \mathbf{I} is the identity matrix in \mathbb{M}_d^{sa} (*mass conservation*).
- (e) T_t is a positive map for any $t \in \mathbb{R}_+$.

We note that the quantum dynamical semigroup has been studied in the contexts of quantum Markov processes (see Section 1.1). It is shown [22, 49, 50] that any unital quantum dynamical semigroup is generated by a Liouvillian $\mathcal{L} : \mathbb{M}_d^{\text{sa}} \rightarrow \mathbb{M}_d^{\text{sa}}$ of the form

$$\mathcal{L} : \mathbf{X} \mapsto \Psi(\mathbf{X}) - \kappa \mathbf{X} - \mathbf{X} \kappa^\dagger,$$

where $\kappa \in \mathbb{C}^{d \times d}$ and Ψ is a CP map such that $\Psi(\mathbf{I}) = \kappa + \kappa^\dagger$. Therefore, each unital map can be expressed as $\mathsf{T}_t = e^{t\mathcal{L}}$ for all $t \in \mathbb{R}_+$. The Markov semigroup acting on the matrix-valued function \mathbf{f} is hence defined by $\mathsf{P}_t \mathbf{f}(x) = \mathsf{T}_t(\mathbf{f}(x))$ for all $x \in \Omega$. The invariant measure exists if

$$\begin{aligned} \int \mathbf{f}(x) \mu(dx) &= \int \mathsf{P}_t \mathbf{f}(x) \mu(dx) \\ &= \int \mathsf{T}_t(\mathbf{f}(x)) \mu(dx) \\ &= \mathsf{T}_t \left(\int \mathbf{f}(x) \mu(dx) \right), \quad \forall t \in \mathbb{R}_+. \end{aligned}$$

In other words, the expectation $\mathbb{E}_\mu[\mathbf{f}] = \int \mathbf{f}(x) \mu(dx)$ is the fixed point of the unital semigroup $\{\mathsf{T}_t\}_{t \geq 0}$. From our main result—Theorem 12, we can establish the exponential decays of the matrix Φ -entropy functionals through the quantum dynamical semigroup $\{\mathsf{T}_t\}_{t \geq 0}$:

$$H_\Phi(\mathbf{f}) \leq -C \text{Tr} \mathbb{E}_\mu[\Phi'(\mathbf{f}) \mathbf{L}\mathbf{f}] \quad \text{if and only if} \quad H_\Phi(\mathsf{T}_t(\mathbf{f})) \leq e^{-t/C} H_\Phi(\mathbf{f}), \quad \forall t \geq 0,$$

where the infinitesimal generator is given by $\mathbf{L}\mathbf{f}(x) = \mathcal{L}(\mathbf{f}(x))$, for all $x \in \Omega$.

In the following, we consider the cases of depolarizing and phase-damping channels, and demonstrate the exponential decay phenomenon when all the density operators converge to the same equilibrium. However, as it will be shown in the case of the phase-damping channel, the Φ -Sobolev constant is infinite when the density operators converge to different states.

5.1. Depolarizing Channel. Denote by $\pi \triangleq \mathbf{I}/d$ the maximally mixed state on the Hilbert space \mathbb{C}^d , and let $r > 0$ be a constant. The quantum dynamical semigroup defined by the depolarizing channel is:

$$(5.1) \quad \mathsf{T}_t : \mathbf{f}(x) \mapsto e^{-rt} \mathbf{f}(x) + (1 - e^{-rt}) \operatorname{Tr}[\mathbf{f}(x)] \cdot \pi, \quad \forall x \in \Omega, t \in \mathbb{R}_+.$$

It is not hard to verify that $\{\mathsf{T}_t\}_{t \geq 0}$ forms a Markov semigroup with a unique fixed point (also called the stationary state) $\operatorname{Tr}[\mathbf{f}] \pi$. We assume μ is the invariant measure of \mathbf{f} : $\mathsf{T}_t(\mathbb{E}_\mu[\mathbf{f}]) = \mathbb{E}_\mu[\mathbf{f}] = \operatorname{Tr}[\mathbf{f}] \pi$. The infinitesimal generator and the Dirichlet form can be calculated as

$$\mathsf{L}\mathbf{f} = \lim_{t \rightarrow 0} \frac{1}{t} (\mathsf{T}_t(\mathbf{f}) - \mathbf{f}) = r (\operatorname{Tr}[\mathbf{f}] \pi - \mathbf{f});$$

$$\mathcal{E}(\mathbf{f}) = \frac{1}{2} \mathbb{E}_\mu [\mathsf{L}\mathbf{f}^2 - \mathbf{f} \cdot \mathsf{L}\mathbf{f} - \mathsf{L}\mathbf{f} \cdot \mathbf{f}] = \frac{r}{2} \left(\mathbb{E}_\mu[\mathbf{f}^2] + \operatorname{Tr} \mathbb{E}_\mu[\mathbf{f}^2] \pi - 2 (\operatorname{Tr}[\mathbf{f}] \pi)^2 \right).$$

Now let the matrix-valued function correspond to a set of density operators— $\mathbf{f}(x) := \rho_x$, $\forall x \in \Omega$, with the average state $\bar{\rho} = \pi$. The constant C_2 in the spectral gap inequality is

$$(5.2) \quad C_2 = \sup_{\mathbf{f}: \mathbb{E}_\mu[\mathbf{f}] = \pi} \frac{\operatorname{Var}(\mathbf{f})}{\mathcal{E}(\mathbf{f})} = \sup_{\rho_X: \bar{\rho} = \pi} \frac{2}{r} \cdot \frac{\operatorname{Tr} \mathbb{E}_\mu[\rho_X^2] - \frac{1}{d}}{2 \operatorname{Tr} \mathbb{E}_\mu[\rho_X^2] - \frac{2}{d}} = \frac{1}{r},$$

where we denote by X the random variable such that $\Pr(X = x) = \mu(x)$ for all $x \in \Omega$. Hence, the spectral gap constant is $C_2 = \frac{1}{r}$, and we have the exponential decays of the variance from Corollary 13:

$$(5.3) \quad \operatorname{Var}(\mathsf{T}_t(\rho_X)) \leq e^{-2rt} \cdot \operatorname{Var}(\rho_X).$$

Similarly, the modified log-Sobolev constant C_χ can be calculated by

$$(5.4) \quad \begin{aligned} C_\chi &= \sup_{\mathbf{f}: \mathbb{E}_\mu[\mathbf{f}] = \pi} \frac{\operatorname{Ent}(\mathbf{f})}{-\operatorname{Tr} \mathbb{E}_\mu[(\mathbf{I} + \log \mathbf{f}) \mathsf{L}\mathbf{f}]} \\ &= \sup_{\rho_X: \bar{\rho} = \pi} \frac{1}{r} \cdot \frac{\operatorname{Tr} \mathbb{E}_\mu[\rho_X \log \rho_X] + \log d}{\operatorname{Tr} \mathbb{E}_\mu[\rho_X \log \rho_X] - \frac{\operatorname{Tr} \mathbb{E}_\mu[\log \rho_X]}{d}}. \end{aligned}$$

In the following proposition, we show that $C_\chi = \frac{1}{2r}$ when $d = 2$. Therefore, we are able to establish the exponential decay of the Holevo quantity.

Proposition 15. *Consider a Hilbert space \mathbb{C}^2 . Denote the quantum dynamical semigroup of the depolarizing channel by*

$$\mathsf{T}_t : \rho \mapsto e^{-rt} \rho + (1 - e^{-rt}) \cdot \pi, \quad t \in \mathbb{R}_+.$$

For any quantum ensemble on \mathbb{C}^2 with the average state being π , the modified log-Sobolev constant is $C_\chi = \frac{1}{2r}$. Moreover, we have

$$(5.5) \quad \chi(\{\mu(x), \mathsf{T}_t(\rho_x)\}_{x \in \Omega}) \leq e^{-2rt} \cdot \chi(\{\mu(x), \rho_x\}_{x \in \Omega}).$$

The proof can be found in Appendix B.

We simulate the depolarizing qubit channel with the initial states $\rho_1 = |0\rangle\langle 0|$, $\rho_2 = |1\rangle\langle 1|$ and the uniform distribution in Figure 1. The blue dashed curve and red solid curve show that the upper bounds for the exponential decays in Eqs. (5.5) and (5.3) are quit tight.

We remark that if every state ρ_x in the ensemble goes through different depolarizing channel with rate r_x , i.e.

$$\mathsf{T}_t(x, x) = \mathsf{T}_t^x : \mathbf{f}(x) \mapsto e^{-r_x t} \mathbf{f}(x) + (1 - e^{-r_x t}) \operatorname{Tr}[\mathbf{f}(x)] \cdot \pi,$$

the Sobolev constant C_2 and C_χ will be dominated by the channel with the minimal rate $\inf_{x \in \Omega} r_x := r_{\inf}$. Namely, the spectral gap constant in Eq. (5.2) becomes

$$C_2 = \sup_{\rho_X: \bar{\rho} = \pi} \frac{\operatorname{Tr} \mathbb{E}_\mu[\rho_X^2] - \frac{1}{d}}{\operatorname{Tr} \mathbb{E}_\mu[r_X \rho_X^2] - \frac{\mathbb{E}_\mu[r_X]}{d}} \leq \frac{1}{r_{\inf}},$$

and the modified log-Sobolev constant in Eq. (5.4) is

$$C_\chi \leq \frac{1}{r_{\inf}} \cdot \frac{\text{Tr } \mathbb{E}_\mu[\rho_X \log \rho_X] - \log d}{\text{Tr } \mathbb{E}_\mu[\rho_X \log \rho_X] - \frac{\text{Tr } \mathbb{E}_\mu[\log \rho_X]}{d}}.$$

5.2. Phase-Damping Channel. Fix $d = 2$, and denote the Pauli matrix by

$$\sigma_Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The quantum dynamical semigroup defined by the phase-damping channel is

$$\mathbb{T}_t : \mathbf{f}(x) \mapsto \frac{(1 + e^{-rt})}{2} \mathbf{f}(x) + \frac{(1 - e^{-rt})}{2} \sigma_Z \mathbf{f}(x) \sigma_Z, \quad \forall x \in \Omega, t \in \mathbb{R}_+$$

with the generator:

$$\mathbb{L}\mathbf{f} = \lim_{t \rightarrow 0} \frac{1 - e^{-rt}}{2t} (\sigma_Z \mathbf{f} \sigma_Z - \mathbf{f}) = \frac{r}{2} (\sigma_Z \mathbf{f} \sigma_Z - \mathbf{f}).$$

It is well-known that any diagonal matrix (with respect to the computation basis) is a fixed point of the phase-damping channel \mathbb{T}_t . Now if we assume every matrix $\mathbb{P}_t \mathbf{f}(x)$ converges to different matrices, i.e. $\mathbb{T}_t(\mathbf{f}(x)) \neq \mathbb{T}_t(\mathbf{f}(y))$ for all $x \neq y$ and $t \in \mathbb{R}_+$, then the matrix Φ -entropy functional $H_\Phi(\mathbb{P}_t \mathbf{f})$ is non-zero for all $t \in \mathbb{R}_+$. However, the infinitesimal generator approaches zero as t goes to infinity, i.e.

$$\lim_{t \rightarrow \infty} \mathbb{L} \mathbb{P}_t \mathbf{f} = \lim_{t \rightarrow \infty} \frac{r}{2} (\sigma_Z (\mathbb{P}_t \mathbf{f}) \sigma_Z - \mathbb{P}_t \mathbf{f}) = \mathbf{0},$$

which means that the Φ -Sobolev constant C in Theorem 12 is infinity. In other words, the matrix Φ -entropy $H_\Phi(\mathbb{P}_t \mathbf{f})$ does not decay exponentially in this phase-damping channel.

Remark 5.1. The reason that makes these two examples quite different is the uniqueness of fixed point of the quantum dynamic semigroup \mathbb{T}_t . Since the depolarizing channel has a unique equilibrium state, all the matrices eventually converges. Hence, the Sobolev constants are finite, which leads to the exponential decay phenomenon. On the other hand, the phase-damping channel has multiple fixed points. This ensures the matrix Φ -entropy functionals never vanish.

6. THE STATISTICAL MIXTURE OF THE MARKOV SEMIGROUP

In this section, we study a statistical mixing of Markov semigroups. The interested matrix-valued functions are defined on the Boolean hypercube $\{0, 1\}^n$, which arise in the context of Fourier analysis [51, 52]. Moreover, a matrix hypercontractivity inequality has been established on this particular set of matrix-valued functions [53].

Our first example is the Markovian jump process with transition rates p from state 0 to 1 and $(1 - p)$ from 1 to 0. We will calculate its convergence rate using the matrix Efron-Stein inequality [13]. Second, we consider the statistical mixing of a quantum random graph where each vertex corresponds to a quantum state, and further bound its mixing time.

6.1. Markovian Jump Process on Symmetric Boolean Hypercube. We consider a special case of the Markov semigroup induced by a classical Markov kernel:

$$(6.1) \quad \mathbb{P}_t \mathbf{f}(x) \triangleq \int_{y \in \Omega} p_t(x, dy) \mathbf{f}(y),$$

where $p_t(x, dy)$ is a family of transition probabilities⁶ on Ω , and satisfies the following Chapman-Kolmogorov identity

$$\int_{y \in \Omega} p_s(x, dy) p_t(y, dz) = p_{s+t}(x, dz).$$

⁶For every $t \geq 0$ and $x \in \Omega$, $p_t(x, \cdot)$ is a probability measure on Ω , and $x \mapsto p_t(x, E)$ is measurable for every measurable set $E \in \Sigma$

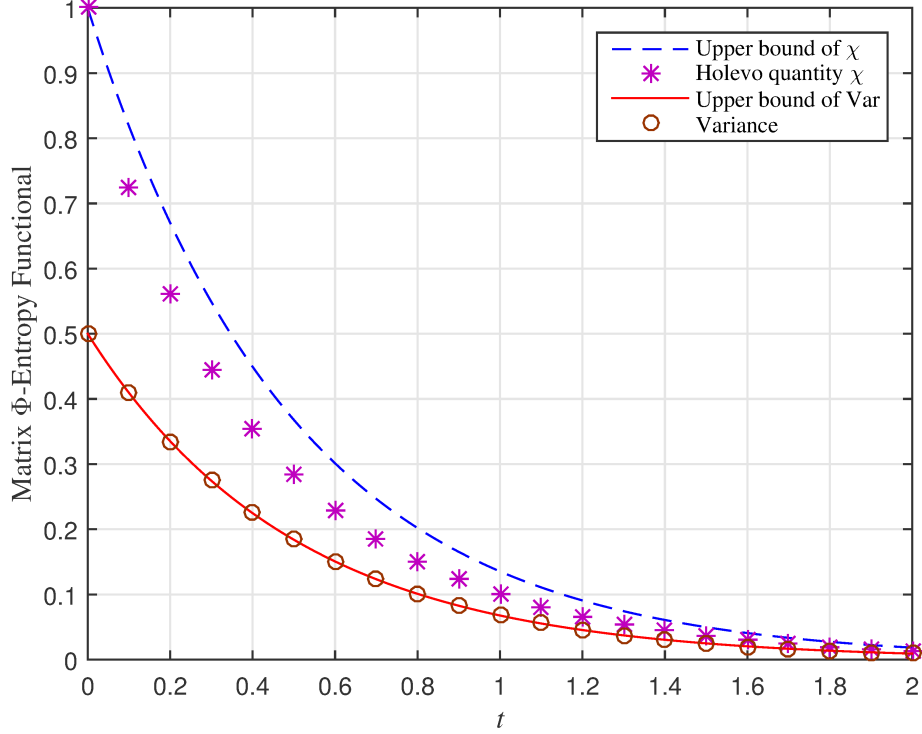


FIGURE 1. This figure illustrates the exponential decay phenomenon of the variance and the Holevo quantity through the depolarizing qubit channel ($d = 2$) (see Eq. (5.1)) with rate $r = 1$. Assume $\rho_1 = |0\rangle\langle 0|$ and $\rho_2 = |1\rangle\langle 1|$ with uniform distribution. The blue dashed curve and red solid curve represent the upper bounds of the Holevo quantity and the variance, respectively, i.e. right-hand sides of Eqs.(5.5) and (5.3). The actual variance and Holevo quantity through the time evolution of the depolarizing qubit channel are plotted by the ‘o’ and ‘*’ lines, , which demonstrates the tightness of the exponential upper bounds.

In other words, the time evolution of the matrix-valued function \mathbf{f} is under a statistical mixture according to Eq. (6.1). Let the state space be a hypercube, i.e. $\Omega \equiv \{0, 1\}^n$ with the measure denoted by

$$\mu_{n,p}(x) = p^{\sum_{i=1}^n x_i} (1-p)^{\sum_{i=1}^n (1-x_i)}, \quad \forall x \in \{0, 1\}^n.$$

We introduce the operator Δ_i that acts on any matrix-valued function $\mathbf{f} : \{0, 1\}^n \rightarrow \mathbb{M}_d^{\text{sa}}$ as follows:

$$\begin{aligned} \Delta_i \mathbf{f} &= \begin{cases} (1-p) \nabla_i \mathbf{f}, & \text{if } x_i = 1 \\ -p \nabla_i \mathbf{f}, & \text{if } x_i = 0 \end{cases} \\ &= \mathbf{f} - \int \mathbf{f} d\mu_{1,p}(x_i), \end{aligned}$$

where

$$\nabla_i \mathbf{f} \triangleq \mathbf{f}(x_1, \dots, x_{i-1}, 1, x_{i+1}, \dots, x_n) - \mathbf{f}(x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_n).$$

The semigroup $\{\mathbf{P}_t\}_{t \geq 0}$ of the Markovian jump process is given by the generator \mathbf{L} with transition rates p from state 0 to 1 and $(1-p)$ from 1 to 0:

$$\mathbf{L} = - \sum_{i=1}^n \Delta_i.$$

Then, we are able to derive the rate of the exponential decay in variance functions.

Theorem 16 (Exponential Decay of Variances for Symmetric Bernoulli Random Variables). *Given a Markov Triple $(\{0, 1\}^n, \Gamma, \mu_{n,p})$ of a Markovian jump process, one has*

$$\text{Var}(\mathbf{P}_t \mathbf{f}) \leq e^{-2t} \cdot \text{Var}(\mathbf{f}),$$

for any matrix-valued function $\mathbf{f} : \{0, 1\}^n \rightarrow \mathbb{M}_d^{\text{sa}}$.

Proof. In Corollary 13 we show the equivalence between the exponential decay in variances and the spectral gap inequality (see Definition 10). Therefore, it suffices to establish the spectral gap constant of the Markovian jump process.

Notably, the spectral gap inequality of the Markov jump process is a special case of the *matrix Efron-Stein inequality* in Ref. [13]:

Proposition 17 (Matrix Efron-Stein Inequality [13, Theorem 4.1]). *For any measurable and bounded matrix-valued function $\mathbf{f} : (\mathbb{M}_d^{\text{sa}})^n \rightarrow \mathbb{M}_d^{\text{sa}}$, we have*

$$(6.2) \quad \text{Var}(\mathbf{f}) \leq \frac{1}{2} \text{Tr} \mathbb{E} \left[\sum_{i=1}^n \left(\mathbf{f}(\underline{\mathbf{X}}) - \mathbf{f}(\widetilde{\mathbf{X}}^{(i)}) \right)^2 \right],$$

where $\underline{\mathbf{X}} \triangleq (\mathbf{X}_1, \dots, \mathbf{X}_n) \in (\mathbb{M}_d^{\text{sa}})^n$ denote an n -tuple random vector with independent elements, and $\widetilde{\mathbf{X}}^{(i)} \triangleq (\mathbf{X}_1, \dots, \mathbf{X}_{i-1}, \mathbf{X}'_i, \mathbf{X}_{i+1}, \dots, \mathbf{X}_n)$ is obtained by replacing the i -th component of $\underline{\mathbf{X}}$ by an independent copy of \mathbf{X}'_i .

By taking $\underline{\mathbf{X}}$ to be an n -tuple Bernoulli random vector, and observe that the right-hand side of Eq. (6.2) coincides with $\text{Tr} [\mathcal{E}(\mathbf{f})]$ for the Markov jump process to complete the proof. \square

Similarly, the convergence rate of the exponential decay in entropy functionals can be calculated as follows.

Theorem 18 (Exponential Decay of Matrix Φ -Entropies for Symmetric Bernoulli Random Variables). *Given a Markov Triple $(\{0, 1\}^n, \Gamma, \mu_{n,p})$ of a Markovian jump process, one has*

$$\text{Ent}(\mathbf{P}_t \mathbf{f}) \leq e^{-t} \cdot \text{Ent}(\mathbf{f}),$$

for any matrix-valued function $\mathbf{f} : \{0, 1\}^n \rightarrow \mathbb{M}_d^{\text{sa}}$.

Proof. The theorem is equivalent to proving

$$\text{Ent}(\mathbf{f}) \leq -\text{Tr} \mathbb{E}[(\mathbf{I} + \log \mathbf{f}) \mathbf{L} \mathbf{f}].$$

By virtue of the subadditivity property, we first establish the case $n = 1$, i.e.

$$(6.3) \quad \text{Ent}(\mathbf{f}) \leq -\text{Tr} \mathbb{E}[(\mathbf{I} + \log \mathbf{f}) \mathbf{L} \mathbf{f}], \quad \forall \mathbf{f} : \{0, 1\} \rightarrow \mathbb{M}_d^{\text{sa}}.$$

Taking $\Phi(u) = u \log u$, the first-order convexity property implies that

$$\text{Tr} [\Phi(\mathbf{Y}) - \Phi(\mathbf{X})] \geq \text{Tr} [\mathbf{D}\Phi[\mathbf{X}](\mathbf{Y} - \mathbf{X})], \quad \forall \mathbf{X}, \mathbf{Y} \in \mathbb{M}_d^{\text{sa}}.$$

Let $\mathbf{X} \equiv \mathbf{f}$ and $\mathbf{Y} \equiv \mathbb{E} \mathbf{f}$. Then it follows that

$$\text{Tr} [\Phi(\mathbb{E} \mathbf{f}) - \Phi(\mathbf{f})] \geq \text{Tr} [\mathbf{D}\Phi[\mathbf{f}](\mathbb{E} \mathbf{f} - \mathbf{f})], \quad \forall \mathbf{f} : \{0, 1\} \rightarrow \mathbb{M}_d^{\text{sa}}$$

from which we apply the expectation again to obtain

$$(6.4) \quad \text{Tr} [\mathbb{E} \Phi(\mathbf{f}) - \Phi(\mathbb{E} \mathbf{f})] \leq \text{Tr} [\mathbb{E} [\mathbf{D}\Phi[\mathbf{f}](\mathbf{f} - \mathbb{E} \mathbf{f})]].$$

Then by elementary manipulation, the right-hand side of Eq. (6.4) leads to

$$\begin{aligned} \text{Tr} [\mathbb{E} [\mathbf{D}\Phi[\mathbf{f}](\mathbf{f} - \mathbb{E} \mathbf{f})]] &= \text{Tr} [p(1-p) \mathbf{D}\Phi[\mathbf{f}(1)](\mathbf{f}(1) - \mathbf{f}(0)) + (1-p)p \mathbf{D}\Phi[\mathbf{f}(0)](\mathbf{f}(0) - \mathbf{f}(1))] \\ &= -\text{Tr} \mathbb{E} [\Phi'(\mathbf{f}) \cdot \Delta_1 \mathbf{f}], \end{aligned}$$

and hence arrives at Eq. (6.3).

Then the subadditivity of Φ -entropy in Theorem 2 yields

$$\begin{aligned} \text{Ent}(\mathbf{f}) &\leq \sum_{i=1}^n \mathbb{E} \left[\text{Ent}^{(i)}(\mathbf{f}) \right] \\ &\leq - \sum_{i=1}^n \mathbb{E} \left[\text{Tr} \mathbb{E}_i [\Phi'(\mathbf{f}) \cdot \Delta_i \mathbf{f}] \right] \\ &= - \text{Tr} \mathbb{E} \left[(\mathbf{I} + \log \mathbf{f}) \mathbf{L} \mathbf{f} \right], \end{aligned}$$

which completes the proof. \square

6.2. Mixing Times of Quantum Random Graphs. In the following, we introduce a model of quantum states defined on a random graph and apply the above results to calculate the mixing time. Consider a directed graph Ω with finite vertices. Every arc $e = (x, y)$, $x, y \in \Omega$ of the graph corresponds a non-negative weight $w(x, y)$ (assume $y \neq x$), which represents the *transition rate* starting from node x to y . Here we denote by $(L(x, y))_{x, y \in \Omega}$ the weight matrix that satisfies $L(x, y) \geq 0$ as $x \neq y$. Moreover, a balance condition $\sum_{y \in \Omega} L(x, y) = 0$ for any $x \in \Omega$ is imposed. The Markov transition kernel can be constructed via the exponentiation of the weight matrix L (see e.g. [3, 4, 54]):

$$p_t(x, y) = (e^{tL})(x, y),$$

which stands for the probability from node x to y after time t . Now, each vertex x of the graph is endowed with a density operator $\mathbf{f}(x) = \rho_x$ on some fixed Hilbert space \mathbb{C}^d . The evolution of the quantum states in the graph is characterized by the Markov semigroup acting on the ensembles $\{\rho_x\}_{x \in \Omega}$ according to the rule:

$$(6.5) \quad \rho_{t,x} \triangleq \mathbf{P}_t \mathbf{f}(x) = \sum_{y \in \Omega} \rho_y p_t(x, y).$$

Thus $\rho_{t,x}$ is the quantum state at node x that is mixed from other nodes according to weight $p_t(x, y)$.

It is not hard to observe that the measure μ is invariant for this Markov semigroup $\{\mathbf{P}_t\}_{t \geq 0}$ if

$$\mu(y) = \sum_{x \in \Omega} \mu(x) p_t(x, y), \quad \forall y \in \Omega.$$

We note that there always exists a probability measure satisfying the above equation. However, the probability measure is unique if and only if the Markov kernel $(p_t(x, y))$ is *irreducible*⁷ [6, 54].

As shown in Section 4, all the states $\rho_{t,x} = \mathbf{P}_t \mathbf{f}(x)$ will converge to the average state $\sum_{y \in \Omega} \mu(y) \cdot \rho_y =: \bar{\rho}$ as t goes to infinity, where μ is a unique invariant measure for $\{\mathbf{P}_t\}_{t \geq 0}$. To measure how close it is to the average states $\bar{\rho}$, we exploit the matrix Φ -entropy functionals (with respect to the invariant measure μ) to capture the convergence rate. In particular, we choose $\Phi(u) = u^2$ and $\Phi(u) = u \log u$, which coincide the variance function and the Holevo quantity.

We define the \mathbb{L}^2 and Holevo mixing times as follows:

Definition 19. Let $\rho_t : x \mapsto \rho_{t,x}$ be the ensembles of quantum states after time t . The \mathbb{L}^2 and Holevo mixing times are defined as:

$$\begin{aligned} \tau_2(\epsilon) &\triangleq \inf \{t : \text{Var}(\rho_t) \leq \epsilon\} \\ \tau_\chi(\epsilon) &\triangleq \inf \{t : \chi(\rho_t) \leq \epsilon\}. \end{aligned}$$

By applying our main result (Theorem 12), we upper bound the mixing time of the Markov random graphs.

⁷ A Markov kernel matrix is called irreducible if there exists a finite $t \geq 0$ such that $p_t(x, y) > 0$ for all x and y . In other words, it is possible to get any state from any state. However, the uniqueness of the invariant measure gets more involved when the state space Ω is uncountable. We refer the interested readers to reference [55, Chapter 7] for further discussions. We also remark that it is still unclear whether the classical characterizations of the unique invariant measures can be directly extended to the case of matrix-valued functions. This problem is left as future work.

Corollary 20. *Let $C_2 > 0$ and $C_\chi > 0$ be the spectral gap constant and the modified log-Sobolev constant of the Markov Triple (Ω, Γ, μ) . Then one has*

$$(6.6) \quad \tau_2(\epsilon) \leq \frac{C_2}{2} \left(\log \text{Var}(\boldsymbol{\rho}) + \log \frac{1}{\epsilon} \right)$$

$$(6.7) \quad \tau_\chi(\epsilon) \leq C_\chi \left(\log \chi(\boldsymbol{\rho}) + \log \frac{1}{\epsilon} \right),$$

where $\boldsymbol{\rho} : x \mapsto \boldsymbol{\rho}_x$ denotes the initial ensemble of the graph.

Proof. The corollary follows immediately from Theorem 12. Set $\text{Var}(\boldsymbol{\rho}_t) = \epsilon$. Then we have

$$\epsilon \leq e^{-2\tau_2(\epsilon)/C_2} \text{Var}(\boldsymbol{\rho}),$$

which implies the desired upper bound for the \mathbb{L}^2 mixing time. The upper bound for the mixing time of the Holevo quantity follows in a similar way. \square

Remark 6.1. Note that for every initial ensemble $\boldsymbol{\rho}_0$, it follows that

$$\text{Var}(\boldsymbol{\rho}) = \text{Tr} \left[\sum_{x \in \Omega} \mu(x) \boldsymbol{\rho}_x^2 - \left(\sum_{x \in \Omega} \mu(x) \boldsymbol{\rho}_x \right)^2 \right] \leq \max_{x \in \Omega} \mu(x)/d =: \mu^*/d.$$

Equation (6.6) can be replaced by

$$\tau_2(\epsilon) \leq \frac{C_2}{2} \left(\log \frac{\mu^*}{d} + \log \frac{1}{\epsilon} \right).$$

Moreover, it is well-known [41] that the Holevo quantity $\chi(\{\mu(x), \boldsymbol{\rho}_x\})$ is bounded by the Shannon entropy $H(\mu)$ of the probability distribution μ . Hence, Eq. (6.7) is rewritten as

$$\tau_2(\epsilon) \leq \frac{C_2}{2} \left(\log H(\mu) + \log \frac{1}{\epsilon} \right).$$

\diamond

Consider the random graph generated by the Markovian jump process on a hypercube $\{0, 1\}^n$ with probability $p = 1/2$. Theorems 16 and 18 give the spectral gap constant $C_2 = 1$ and modified log-Sobolev constant $C_\chi = 1$. Hence, for every initial ensemble $\boldsymbol{\rho}$, the mixing times of this quantum random graph is

$$\begin{aligned} \tau_2(\epsilon) &\leq \frac{1}{2} \left(\log \text{Var}(\boldsymbol{\rho}) + \log \frac{1}{\epsilon} \right) \\ \tau_\chi(\epsilon) &\leq \left(\log \chi(\boldsymbol{\rho}) + \log \frac{1}{\epsilon} \right). \end{aligned}$$

7. DISCUSSIONS

Classical spectral gap inequalities and logarithmic Sobolev inequalities have proven to be a fundamental tool in analyzing Markov semigroups on real-valued functions. In this paper, we extend the definition of Markov semigroups to matrix-valued functions and investigate its equilibrium property. Our main result shows that the matrix Φ -entropy functionals exponentially decay along the Markov semigroup, and the convergence rates are determined by the coefficients of the matrix Φ -Sobolev inequality [13]. In particular, we establish the variance and entropy decays of the Markovian jump process using the subadditivity of matrix Φ -entropies [10] and tools from operator algebras.

The Markov semigroup introduced in this paper is not only of independent interest in mathematics, but also has substantial applications in quantum information theory. In this work, we study the dynamical process of a quantum ensemble governed by the Markov semigroups, and analyze how the entropies of the quantum ensemble evolve as time goes on. When the quantum dynamical process is a quantum unital map, our result yields a stronger version of the monotonicity of the Holevo quantity $\chi(\{\mu(x), \mathbb{T}_t(\boldsymbol{\rho}_x)\}_x) \leq e^{-t/C} \cdot \chi(\{\mu(x), \boldsymbol{\rho}_x\}_x)$.

ACKNOWLEDGEMENTS

MH would like to thank Matthias Christandl, Michael Kastoryano, Robert Koenig, Joel Tropp, and Andreas Winter for their useful comments. MH is supported by an ARC Future Fellowship under Grant FT140100574. MT is funded by an University of Sydney Postdoctoral Fellowship and acknowledges support from the ARC Centre of Excellence for Engineered Quantum Systems (EQUS).

APPENDIX A. HILLE-YOSHIDA'S THEOREM FOR MARKOV SEMIGROUPS

Classical Hille-Yoshida's theorem [56, Chapter IX], [4, Appendix A] provides a nice characterization for the Dirichlet domain $\mathcal{D}(\mathbf{L})$ when the underlying Banach space $(\mathcal{B}, \|\cdot\|)$ is the set of real-valued bounded continuous functions on Ω . In the following we show that the Hille-Yosida's theorem can be naturally extended to the Banach space of bounded continuous matrix-valued functions $\mathbf{f} : \Omega \rightarrow \mathbb{M}^{\text{sa}}$ equipped with the uniform norm (e.g. $\|\mathbf{f}\| = \sup_{x \in \Omega} \|\mathbf{f}(x)\|_\infty$). We note that the proof parallels the classical approach; see e.g. [56, Chapter IX] and [5, Theorem 1.7].

Theorem 21 (Hille-Yoshida's Theorem for Markov Semigroups). *A linear super-operator \mathbf{L} is the infinitesimal generator of a Markov semigroup $\{\mathbf{P}_t\}_{t \geq 0}$ on \mathcal{B} if and only if*

- *The identity function belongs to the Dirichlet domain: $\mathbf{I} \in \mathcal{D}(\mathbf{L})$ and $\mathbf{L}\mathbf{I} = \mathbf{0}$.*
- *$\mathcal{D}(\mathbf{L})$ is dense in \mathcal{B} .*
- *\mathbf{L} is closed.*
- *For any $\lambda > 0$, $(\lambda \mathbf{1} - \mathbf{L})$ is invertible. The inverse $(\lambda \mathbf{1} - \mathbf{L})^{-1}$ is bounded with*

$$\sup_{\|\mathbf{f}\| \leq 1} \|(\lambda \mathbf{1} - \mathbf{L})^{-1} \mathbf{f}\| \leq \frac{1}{\lambda}$$

and preserves positivity, i.e. $(\lambda \mathbf{1} - \mathbf{L})^{-1} \mathbf{f} \succeq \mathbf{0}$ for all $\mathbf{f} \succeq \mathbf{0}$.

Proof.

Necessary condition.

The first item follows from the fact that $\mathbf{P}_t \mathbf{I} = \mathbf{I}$ and the definition of the infinitesimal generator; see Eq. (3.7). To prove the second item, it suffices to show that

$$(A.1) \quad \mathcal{D}_0 \triangleq \left\{ \frac{1}{t} \int_0^t \mathbf{P}_s \mathbf{f} \, ds : \mathbf{f} \in \mathcal{B}, t > 0 \right\} \subset \mathcal{D}(\mathbf{L}).$$

According to the continuity of the map $t \rightarrow \mathbf{P}_t \mathbf{f}$, \mathcal{D}_0 is dense in \mathcal{B} and hence completes the second item.

To show Eq. (A.1), we invoke the semigroup property in Definition 6 to obtain

$$\begin{aligned} \frac{1}{\tau} (\mathbf{P}_\tau - \mathbf{1}) \int_0^t \mathbf{P}_s \mathbf{f} \, ds &= \frac{1}{\tau} \int_\tau^{\tau+t} \mathbf{P}_s \mathbf{f} \, ds - \frac{1}{\tau} \int_0^t \mathbf{P}_s \mathbf{f} \, ds \\ &= \frac{1}{\tau} \int_t^{\tau+t} \mathbf{P}_s \mathbf{f} \, ds - \frac{1}{\tau} \int_0^\tau \mathbf{P}_s \mathbf{f} \, ds \end{aligned}$$

for any $\tau \in \mathbb{R}_+$. By letting τ tend to zero, we have

$$(A.2) \quad \mathbf{L} \int_0^t \mathbf{P}_s \mathbf{f} \, ds = \int_0^t \mathbf{P}_s \mathbf{L} \mathbf{f} \, ds = \mathbf{P}_t \mathbf{f} - \mathbf{f},$$

for any $\mathbf{f} \in \mathcal{D}(\mathbf{L})$, which shows that $\int_0^t \mathbf{P}_s \mathbf{f} \, ds \in \mathcal{D}(\mathbf{L})$ and hence establishes Eq. (A.1).

To show the closeness of the generator \mathbf{L} , we consider a sequence \mathbf{f}_n in $\mathcal{D}(\mathbf{L})$ converging to a function \mathbf{f} such that

$$\lim_{n \rightarrow \infty} \mathbf{L} \mathbf{f}_n = \mathbf{g}$$

for some $\mathbf{g} \in \mathcal{B}$. Apply Eq. (A.2) on \mathbf{f}_n to obtain

$$\mathbf{P}_t \mathbf{f}_n - \mathbf{f}_n = \mathbf{L} \int_0^t \mathbf{P}_s \mathbf{f}_n \, ds = \int_0^t \mathbf{P}_s \mathbf{L} \mathbf{f}_n \, ds.$$

By taking n go to infinity yields $\mathbf{P}_t \mathbf{f} - \mathbf{f} = \int_0^t \mathbf{P}_s \mathbf{g} \, ds$. Dividing by t and letting $t \downarrow 0$ results in $\mathbf{L} \mathbf{f} = \mathbf{g}$ for every $\mathbf{f} \in \mathcal{D}(\mathbf{L})$.

To show the last point, we define the *resolvent* $R(\lambda, \mathbf{L}) \triangleq (\lambda \mathbf{1} - \mathbf{L})^{-1}$ of the generator \mathbf{L} and show that

$$(A.3) \quad R(\lambda, \mathbf{L})\mathbf{f} = \int_0^\infty e^{-\lambda s} \mathbf{P}_s \mathbf{f} \, ds.$$

By introducing a new semigroup $\tilde{\mathbf{P}}_s = e^{-\lambda s} \mathbf{P}_s$ with associated generator $(\mathbf{L} - \lambda \mathbf{1})$, Eq. (A.2) implies

$$(\mathbf{L} - \lambda \mathbf{1}) \int_0^t e^{-\lambda s} \mathbf{P}_s \mathbf{f} \, ds = e^{-\lambda t} \mathbf{P}_t \mathbf{f} - \mathbf{f}.$$

Taking the limit as $t \rightarrow \infty$ and according to the closeness of \mathbf{L} , we deduce that

$$(\mathbf{L} - \lambda \mathbf{1}) \int_0^\infty e^{-\lambda s} \mathbf{P}_s \mathbf{f} \, ds = \mathbf{f}$$

which establishes Eq. (A.3). Finally, since \mathbf{P}_s is contractive (see Eq. (3.4)),

$$(A.4) \quad \sup_{\|\mathbf{f}\|=1} \|R(\lambda, \mathbf{L})\mathbf{f}\| \leq \int_0^\infty e^{-\lambda s} \, ds = \frac{1}{\lambda}.$$

Hence $R(\lambda, \mathbf{L})$ is a bounded super-operator on $\mathcal{D}(\mathbf{L})$. In particular, we observe the positivity of $R(\lambda, \mathbf{L})$ from Eq. (A.3).

Sufficiency.

For any $\lambda > 0$, we define the Yoshida approximation \mathbf{L}_λ of \mathbf{L} by

$$\mathbf{L}_\lambda \triangleq \mathbf{L} \lambda \mathbf{1} (\lambda \mathbf{1} - \mathbf{L})^{-1} = \lambda^2 (\lambda \mathbf{1} - \mathbf{L})^{-1} - \lambda \mathbf{1}.$$

From Eq. (A.4), \mathbf{L}_λ is a bounded super-operator and hence we define a Markov semigroup

$$\mathbf{P}_t^\lambda \triangleq e^{t\mathbf{L}_\lambda} = \sum_{n=0}^\infty \frac{t^n}{n!} \mathbf{L}_\lambda^n = e^{-t\lambda \mathbf{1}} e^{t\lambda^2 (\lambda \mathbf{1} - \mathbf{L})^{-1}}.$$

for $t \in \mathbb{R}_+$. To finish the proof, it remains to show that \mathbf{P}_t^λ converges to \mathbf{P}_t as $\lambda \rightarrow \infty$ and hence the semigroup properties, contractivity, positivity, and the unit property of the semigroup $\{\mathbf{P}_t^\lambda\}_{t \geq 0}$ can be extended to $\{\mathbf{P}_t\}_{t \geq 0}$. To achieve this, we prove that the family of super-operators $\{\mathbf{L}_\lambda\}_\lambda$ converges to \mathbf{L} :

$$\lim_{\lambda \rightarrow \infty} \mathbf{L}_\lambda \mathbf{f} = \mathbf{L} \mathbf{f},$$

and the convergence of $\{\mathbf{L}_\lambda\}_\lambda$ leads to that of the semigroups $\{\mathbf{P}_t^\lambda\}_\lambda$. In fact, the identity

$$\lambda(\lambda \mathbf{1} - \mathbf{L})^{-1} - (\lambda \mathbf{1} - \mathbf{L})^{-1} \mathbf{L} = \mathbf{1}$$

implies that

$$\lim_{\lambda \rightarrow \infty} \lambda(\lambda \mathbf{1} - \mathbf{L})^{-1} \mathbf{g} = \mathbf{g}$$

for every $\mathbf{g} \in \mathcal{D}(\mathbf{L})$. Hence we conclude that \mathbf{L}_λ converges to \mathbf{L} as λ goes to infinity. Moreover, for any positive real numbers (λ_1, λ_2) and any $\mathbf{f} \in \mathcal{D}(\mathbf{L})$, we have the interpolation formula

$$e^{t\mathbf{L}_{\lambda_1}} \mathbf{f} - e^{t\mathbf{L}_{\lambda_2}} \mathbf{f} = t \int_0^1 e^{t(s\mathbf{L}_{\lambda_1} - (1-s)\mathbf{L}_{\lambda_2})} (\mathbf{L}_{\lambda_1} - \mathbf{L}_{\lambda_2}) \mathbf{f} \, ds.$$

since \mathbf{L}_{λ_1} and \mathbf{L}_{λ_2} are commuting. Combined with the contractivity of the semigroups $e^{t(s\mathbf{L}_{\lambda_1} - (1-s)\mathbf{L}_{\lambda_2})}$ for $s \in [0, 1]$, we have

$$\|e^{t\mathbf{L}_{\lambda_1}} \mathbf{f} - e^{t\mathbf{L}_{\lambda_2}} \mathbf{f}\| \leq t \|(\mathbf{L}_{\lambda_1} - \mathbf{L}_{\lambda_2}) \mathbf{f}\|,$$

which shows that the convergence of $\{\mathbf{L}_\lambda\}_\lambda$ ensures the convergence of the semigroups $\{\mathbf{P}_t^\lambda\}_\lambda$. Finally, we define the super-operators $\mathbf{P}_t = \lim_{\lambda \rightarrow \infty} \mathbf{P}_t^\lambda$ for any $t > 0$. It is clearly that $\{\mathbf{P}_t\}_{t \geq 0}$ is a Markov semigroup with generator \mathbf{L} on $\mathcal{D}(\mathbf{L})$ and completes the proof. \square

APPENDIX B. PROOF OF PROPOSITION 15

For convenience, we set $r = 1$. Our strategy starts from proving the upper bound of $C_\chi \leq \frac{1}{2}$. Then we show that the upper bound is attained when every state in the ensemble approaches the maximally mixed state π .

Firstly, recall the modified log-Sobolev constant C_χ in Eq. (5.4). We assume

$$(B.1) \quad \frac{\text{Tr } \mathbb{E}_\mu[\rho_X \log \rho_X] + \log 2}{\text{Tr } \mathbb{E}_\mu[\rho_X \log \rho_X] - \frac{\text{Tr } \mathbb{E}_\mu[\log \rho_X]}{2}} > \frac{1}{2},$$

which implies

$$(B.2) \quad \text{Tr } \mathbb{E}_\mu[\rho_X \log \rho_X] + 2 \log 2 + \frac{\text{Tr } \mathbb{E}_\mu[\log \rho_X]}{2} > 1.$$

However, since any density operator ρ on \mathbb{C}^2 can be expressed as

$$\rho = U \begin{pmatrix} p & 0 \\ 0 & 1-p \end{pmatrix} U^\dagger$$

for some unitary matrix U and $0 \leq p \leq 1$, we have

$$\text{Tr} \left[\left(\frac{\mathbf{I}}{2} + \rho \right) \log \rho \right] = \left(\frac{1}{2} + p \right) \log p + \left(\frac{1}{2} + 1-p \right) \log(1-p).$$

The above equation is concave in p and maximizes at $p = \frac{1}{2}$ (i.e. ρ is a maximally mixed state). As a result, we have

$$\text{Tr} \left[\left(\frac{\mathbf{I}}{2} + \rho \right) \log \rho \right] \leq -2 \log 2$$

which contradicts Eq. (B.2). Hence, we prove the upper bound $C_\chi \leq \frac{1}{2}$.

Second, we consider the case of $|\Omega| = 2$. Let $\mu(1) = p_1$, $\mu(2) = p_2$, where $p_1 + p_2 = 1$, and denote the Pauli matrix by

$$\sigma_Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Without loss of generality, we set the states in the ensemble to be: $\rho_1 = \pi + \frac{\epsilon}{p_1} \sigma_Z$ and $\rho_2 = \pi - \frac{\epsilon}{p_2} \sigma_Z$, where $0 \leq \epsilon \leq \min\{\frac{p_1}{2}, \frac{p_2}{2}\}$. It is not hard to see the average state is π . Then we can calculate that

$$\begin{aligned} & \frac{\text{Tr } \mathbb{E}_\mu[\rho_X \log \rho_X] + \log 2}{\text{Tr } \mathbb{E}_\mu[\rho_X \log \rho_X] - \frac{\text{Tr } \mathbb{E}_\mu[\log \rho_X]}{2}} \\ &= \frac{-p_1 H_b\left(\frac{1}{2} + \frac{\epsilon}{p_1}\right) - p_2 H_b\left(\frac{1}{2} + \frac{\epsilon}{p_2}\right) + \log 2}{-p_1 H_b\left(\frac{1}{2} + \frac{\epsilon}{p_1}\right) - p_2 H_b\left(\frac{1}{2} + \frac{\epsilon}{p_2}\right) + \log 2 - \frac{p_1 \log\left(\frac{1}{2} + \frac{\epsilon}{p_1}\right)\left(\frac{1}{2} - \frac{\epsilon}{p_1}\right) + p_2 \log\left(\frac{1}{2} + \frac{\epsilon}{p_2}\right)\left(\frac{1}{2} - \frac{\epsilon}{p_2}\right)}{2}} \\ &=: \frac{f(\epsilon)}{g(\epsilon)} =: C(\epsilon), \end{aligned}$$

where $H_b(p) \triangleq -p \log p - (1-p) \log(1-p)$ is the binary entropy function. By taking differentiation with respect to ϵ , it can be shown that $C'''(\epsilon) < 0$ and $C'(\epsilon)|_{\epsilon=0} = 0$. In other words, $C(\epsilon)$ achieves its maximum when ϵ tends to 0.

Finally, by L'Hôpital's rule,

$$\lim_{\epsilon \rightarrow 0} C(\epsilon) = \lim_{\epsilon \rightarrow 0} \frac{f''(\epsilon)}{g''(\epsilon)} = \frac{\frac{4}{p_1} + \frac{4}{p_2}}{\frac{8}{p_1} + \frac{8}{p_2}} = \frac{1}{2}.$$

Similarly, for the cases $|\Omega| > 2$, the modified log-Sobolev constant is attained when every state in the ensemble approaches π . Equation (5.5) then follows by Corollary 14.

REFERENCES

- [1] P. Diaconis and L. Saloff-Coste, “Logarithmic Sobolev inequalities for finite Markov chains,” *The Annals of Applied Probability*, vol. 6, no. 3, pp. 695–750, Aug. 1996. DOI: [10.1214/aoap/1034968224](https://doi.org/10.1214/aoap/1034968224).
- [2] D. Bakry, “L’hypercontractivité et son utilisation en théorie des semigroupes,” in *Lectures on Probability Theory*, Springer, Berlin, 1994, pp. 1–144, ISBN: 978-3540582083. DOI: [10.1007/BFb0073872](https://doi.org/10.1007/BFb0073872).
- [3] —, “Functional inequalities for Markov semigroups,” in *Probability measures on groups*, Tata Institute of Fundamental Research, Mumbai, 2006, pp. 91–147. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-00353724/document>.
- [4] D. Bakry, I. Gentil, and M. Ledoux, *Analysis and Geometry of Markov Diffusion Operators*. Springer International Publishing, 2013, ISBN: 978-3319002262. DOI: [10.1007/978-3-319-00227-9](https://doi.org/10.1007/978-3-319-00227-9).
- [5] A. Guionnet and B. Zegarlinski, “Lectures on logarithmic Sobolev inequalities,” in *Séminaire de Probabilités XXXVI*, Springer Berlin Heidelberg, 2003, pp. 1–134. DOI: [10.1007/978-3-540-36107-7_1](https://doi.org/10.1007/978-3-540-36107-7_1).
- [6] L. Saloff-Coste, “Lectures on finite Markov chains,” in *Lectures on Probability Theory and Statistics*, Springer Berlin Heidelberg, 1997, pp. 301–413. DOI: [10.1007/bfb0092621](https://doi.org/10.1007/bfb0092621).
- [7] L. Gross, “Logarithmic Sobolev inequalities,” *American Journal of Mathematics*, vol. 97, no. 4, p. 1061, 1975. DOI: [10.2307/2373688](https://doi.org/10.2307/2373688).
- [8] D. Chafaï, “Entropies, convexity, and functional inequalities: on Φ -entropies and Φ -sobolev inequalities,” *Journal of Mathematics of Kyoto University*, vol. 44, no. 2, pp. 325–363, 2004. arXiv:[math/0211103](https://arxiv.org/abs/math/0211103) [math.PR].
- [9] R. Latała and K. Oleszkiewicz, “Between Sobolev and Poincaré,” in *Geometric Aspects of Functional Analysis*, Springer Science + Business Media, 2000, pp. 147–168. DOI: [10.1007/bfb0107213](https://doi.org/10.1007/bfb0107213). [Online]. Available: <http://dx.doi.org/10.1007/bfb0107213>.
- [10] R. Y. Chen and J. A. Tropp, “Subadditivity of matrix φ -entropy and concentration of random matrices,” *Electronic Journal of Probability*, vol. 19, no. 0, Jan. 2014. DOI: [10.1214/ejp.v19-2964](https://doi.org/10.1214/ejp.v19-2964). arXiv:[1308.2952](https://arxiv.org/abs/1308.2952) [cs.IT].
- [11] J. A. Tropp, “An introduction to matrix concentration inequalities,” *Foundations and Trends in Machine Learning*, vol. 8, no. 1-2, pp. 1–230, 2015. DOI: [10.1561/22000000048](https://doi.org/10.1561/22000000048). arXiv:[1501.01571](https://arxiv.org/abs/1501.01571) [ma.PR].
- [12] D. Paulin, L. Mackey, and J. A. Tropp, “Efron-Stein inequalities for random matrices,” 2014. arXiv:[1408.3470](https://arxiv.org/abs/1408.3470) [math.PR].
- [13] H.-C. Cheng and M.-H. Hsieh, “New characterizations of matrix Φ -entropies, poincaré and sobolev inequalities and an upper bound to Holevo quantity,” 2015. arXiv:[1506.06801](https://arxiv.org/abs/1506.06801) [quant-ph].
- [14] D. Chafaï, “Binomial-Poisson entropic inequalities and the m/m/ ∞ queue,” *ESAIM: Probability and Statistics*, vol. 10, pp. 317–339, 2006. DOI: [10.1051/ps:2006013](https://doi.org/10.1051/ps:2006013). arXiv:[math/0510488](https://arxiv.org/abs/math/0510488) [math.PR]. [Online]. Available: <http://dx.doi.org/10.1051/ps:2006013>.
- [15] D. Petz, “Monotonicity of quantum relative entropy revisited,” *Reviews in Mathematical Physics*, vol. 15, no. 01, pp. 79–91, 2003. DOI: [10.1142/s0129055x03001576](https://doi.org/10.1142/s0129055x03001576).
- [16] S. Gudder, “Quantum Markov chains,” *Journal of Mathematical Physics*, vol. 49, no. 7, p. 072 105, 2008. DOI: [10.1063/1.2953952](https://doi.org/10.1063/1.2953952).
- [17] S. Attal, F. Petruccione, C. Sabot, and I. Sinayskiy, “Open quantum random walks,” *Journal of Statistical Physics*, vol. 147, no. 4, pp. 832–852, 2012. DOI: [10.1007/s10955-012-0491-0](https://doi.org/10.1007/s10955-012-0491-0).
- [18] S. Attal, F. Petruccione, and I. Sinayskiy, “Open quantum walks on graphs,” *Physics Letters A*, vol. 376, no. 18, pp. 1545–1548, 2012. DOI: [10.1016/j.physleta.2012.03.040](https://doi.org/10.1016/j.physleta.2012.03.040).
- [19] E. Hamza and A. Joye, “Spectral transition for random quantum walks on trees,” *Communications in Mathematical Physics*, vol. 326, no. 2, pp. 415–439, 2014. DOI: [10.1007/s00220-014-1882-7](https://doi.org/10.1007/s00220-014-1882-7).
- [20] Y. Feng, N. Yu, and M. Ying, “Model checking quantum markov chains,” *Journal of Computer and System Sciences*, vol. 79, no. 7, pp. 1181–1198, 2013. DOI: [10.1016/j.jcss.2013.04.002](https://doi.org/10.1016/j.jcss.2013.04.002).
- [21] M. Ying, N. Yu, Y. Feng, and R. Duan, “Verification of quantum programs,” *Science of Computer Programming*, vol. 78, no. 9, pp. 1679–1700, 2013. DOI: [10.1016/j.scico.2013.03.016](https://doi.org/10.1016/j.scico.2013.03.016).
- [22] G. Lindblad, “On the generators of quantum dynamical semigroups,” *Communications in Mathematical Physics*, vol. 48, no. 2, pp. 119–130, 1976. DOI: [10.1007/bf01608499](https://doi.org/10.1007/bf01608499).

- [23] K. Temme, M. J. Kastoryano, M. B. Ruskai, M. M. Wolf, and F. Verstraete, “The χ^2 -divergence and mixing times of quantum Markov processes,” *Journal of Mathematical Physics*, vol. 51, no. 12, p. 122 201, 2010. DOI: [10.1063/1.3511335](https://doi.org/10.1063/1.3511335).
- [24] M. J. Kastoryano, D. Reeb, and M. M. Wolf, “A cutoff phenomenon for quantum Markov chains,” *Journal of Physics A: Mathematical and Theoretical*, vol. 45, no. 7, p. 075 307, 2012. DOI: [10.1088/1751-8113/45/7/075307](https://doi.org/10.1088/1751-8113/45/7/075307).
- [25] O. Szehr and M. M. Wolf, “Perturbation bounds for quantum Markov processes and their fixed points,” *Journal of Mathematical Physics*, vol. 54, no. 3, p. 032 203, 2013. DOI: [10.1063/1.4795112](https://doi.org/10.1063/1.4795112).
- [26] O. Szehr, D. Reeb, and M. M. Wolf, “Spectral convergence bounds for classical and quantum Markov processes,” *Communications in Mathematical Physics*, vol. 333, no. 2, pp. 565–595, 2014. DOI: [10.1007/s00220-014-2188-5](https://doi.org/10.1007/s00220-014-2188-5). [Online]. Available: <http://dx.doi.org/10.1007/s00220-014-2188-5>.
- [27] R. Olkiewicz and B. Zegarlinski, “Hypercontractivity in noncommutative L_p spaces,” *Journal of Functional Analysis*, vol. 161, no. 1, pp. 246–285, Jan. 1999. DOI: [10.1006/jfan.1998.3342](https://doi.org/10.1006/jfan.1998.3342).
- [28] T. Bodineau and B. Zegarlinski, “Hypercontractivity via spectral theory,” *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, vol. 03, no. 01, pp. 15–31, Mar. 2000. DOI: [10.1142/s0219025700000030](https://doi.org/10.1142/s0219025700000030).
- [29] R. Carbone, “Optimal Log-Sobolev inequality and hypercontractivity for positive semigroups on $M_2(\mathbb{C})$,” *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, vol. 07, no. 03, pp. 317–335, Sep. 2004. DOI: [10.1142/s0219025704001633](https://doi.org/10.1142/s0219025704001633).
- [30] K. Temme, F. Pastawski, and M. J. Kastoryano, “Hypercontractivity of quasi-free quantum semigroups,” *Journal of Physics A: Mathematical and Theoretical*, vol. 47, no. 40, p. 405 303, Sep. 2014. DOI: [10.1088/1751-8113/47/40/405303](https://doi.org/10.1088/1751-8113/47/40/405303). arXiv:[1403.5224](https://arxiv.org/abs/1403.5224) [quant-ph].
- [31] C. King, “Hypercontractivity for semigroups of unital qubit channels,” *Communications in Mathematical Physics*, vol. 328, no. 1, pp. 285–301, Mar. 2014. DOI: [10.1007/s00220-014-1982-4](https://doi.org/10.1007/s00220-014-1982-4). arXiv:[1210.8412](https://arxiv.org/abs/1210.8412) [quant-ph].
- [32] T. Cubitt, M. J. Kastoryano, K. Temme, and A. Montanaro, “Quantum reverse hypercontractivity,” 2015. arXiv:[1504.06143](https://arxiv.org/abs/1504.06143) [quant-ph].
- [33] A. Müller-Hermes, D. S. Franca, and M. M. Wolf, “Entropy production of doubly stochastic quantum channels,” 2015. arXiv:[1505.04678](https://arxiv.org/abs/1505.04678) [quant-ph].
- [34] S. Beigi and C. King, “Hypercontractivity and the logarithmic Sobolev inequality for the completely bounded norm,” 2015. arXiv:[1509.02610](https://arxiv.org/abs/1509.02610) [math-ph].
- [35] M. J. Kastoryano and K. Temme, “Quantum logarithmic Sobolev inequalities and rapid mixing,” *Journal of Mathematical Physics*, vol. 54, no. 5, p. 052 202, 2013. DOI: [10.1063/1.4804995](https://doi.org/10.1063/1.4804995). arXiv:[1207.3261](https://arxiv.org/abs/1207.3261) [quant-ph].
- [36] R. Carbone, “Logarithmic Sobolev inequalities and exponential entropy decay in non-commutative algebras,” 2014. arXiv:[1402.6948](https://arxiv.org/abs/1402.6948) [math.OA].
- [37] R. Carbone and A. Martinelli, “Logarithmic Sobolev inequalities in non-commutative algebras,” *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, vol. 18, no. 02, p. 1 550 011, Jun. 2015. DOI: [10.1142/s0219025715500113](https://doi.org/10.1142/s0219025715500113).
- [38] F. Hansen and G. K. Pedersen, “Jensen’s operator inequality,” *London Mathematical Society*, vol. 35, no. 4, pp. 553–564, 2003. DOI: [10.1112/S0024609303002200](https://doi.org/10.1112/S0024609303002200).
- [39] D. R. Farenick and F. Zhou, “Jensen’s inequality relative to matrix-valued measures,” *Journal of Mathematical Analysis and Applications*, vol. 327, pp. 919–929, 2 2007. DOI: [10.1016/j.jmaa.2006.05.008](https://doi.org/10.1016/j.jmaa.2006.05.008).
- [40] C. B. Mendl and M. M. Wolf, “Unital quantum channels - convex structure and revivals of Birkhoff’s theorem,” *Communications in Mathematical Physics*, vol. 289, no. 3, pp. 1057–1086, 2009. DOI: [10.1007/s00220-009-0824-2](https://doi.org/10.1007/s00220-009-0824-2).
- [41] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 2009, ISBN: 978-1107002173. DOI: [10.1017/cbo9780511976667](https://doi.org/10.1017/cbo9780511976667).

- [42] V. V. Peller, “Hankel operators in the perturbation theory of unitary and self-adjoint operators,” *Functional Analysis and Its Applications*, vol. 19, no. 2, pp. 111–123, 1985. DOI: [10.1007/bf01078390](https://doi.org/10.1007/bf01078390). [Online]. Available: <http://dx.doi.org/10.1007/bf01078390>.
- [43] K. Bickel, “Differentiating matrix functions,” *Operators and Matrices*, no. 1, pp. 71–90, 2007. DOI: [10.7153/oam-07-03](https://doi.org/10.7153/oam-07-03). arXiv:1104.0336 [math.FA]. [Online]. Available: <http://dx.doi.org/10.7153/oam-07-03>.
- [44] K. Atkinson and W. Han, *Theoretical Numerical Analysis: A Functional Analysis Framework*. Springer International Publishing, 2009, ISBN: 978-1441904577.
- [45] F. Hiai and D. Petz, *Introduction to Matrix Analysis and Applications*. Springer International Publishing, 2014, ISBN: 978-3319041490. DOI: [10.1007/978-3-319-04150-6](https://doi.org/10.1007/978-3-319-04150-6).
- [46] J. Diestel and J. Uhl, *Vector Measures*. American Mathematical Society, 1977. DOI: [10.1090/surv/015](https://doi.org/10.1090/surv/015).
- [47] J. Mikusiński, *The Bochner Integral*. Springer, 1978. DOI: [10.1007/978-3-0348-5567-9](https://doi.org/10.1007/978-3-0348-5567-9).
- [48] D. Pérez-García, M. M. Wolf, D. Petz, and M. B. Ruskai, “Contractivity of positive and trace-preserving maps under ℓ_p norms,” *Journal of Mathematical Physics*, vol. 47, no. 8, p. 083 506, 2006. DOI: [10.1063/1.2218675](https://doi.org/10.1063/1.2218675).
- [49] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan, “Completely positive dynamical semigroups of N -level systems,” *Journal of Mathematical Physics*, vol. 17, no. 5, pp. 821–825, 1976. DOI: [10.1063/1.522979](https://doi.org/10.1063/1.522979).
- [50] G. Androulakis and M. Ziemke, “Generators of quantum Markov semigroups,” *Journal of Mathematical Physics*, vol. 56, no. 8, p. 083 512, 2015. DOI: [10.1063/1.4928936](https://doi.org/10.1063/1.4928936).
- [51] R. de Wolf, “A brief introduction to Fourier analysis on the Boolean cube,” *Theory of Computing Library*, vol. Graduate Surveys 1, pp. 1–20, 2008. DOI: [10.4086/toc.gs.2008.001](https://doi.org/10.4086/toc.gs.2008.001).
- [52] S. Fehr and C. Schaffner, “Randomness extraction via δ -biased masking in the presence of a quantum attacker,” in *Theory of Cryptography*, Springer, 2008, pp. 465–481. DOI: [10.1007/978-3-540-78524-8_26](https://doi.org/10.1007/978-3-540-78524-8_26).
- [53] A. Ben-Aroya, O. Regev, and R. de Wolf, “A hypercontractive inequality for matrix-valued functions with applications to quantum computing and LDCs,” in *2008 49th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, Institute of Electrical & Electronics Engineers (IEEE), 2008. DOI: [10.1109/focs.2008.45](https://doi.org/10.1109/focs.2008.45). arXiv:0705.3806 [quant-ph]. [Online]. Available: <http://dx.doi.org/10.1109/focs.2008.45>.
- [54] J. R. Norris, *Markov Chains*. Cambridge University Press (CUP), 1997. DOI: [10.1017/cbo9780511810633](https://doi.org/10.1017/cbo9780511810633).
- [55] G. D. Prato, *An Introduction to Infinite-Dimensional Analysis*. Springer, 2006. DOI: [10.1007/3-540-29021-4](https://doi.org/10.1007/3-540-29021-4).
- [56] K. Yosida, *Functional Analysis*, 6th ed. Springer Berlin Heidelberg, 1996, ISBN: 978-3540586548. DOI: [10.1007/978-3-642-61859-8](https://doi.org/10.1007/978-3-642-61859-8).