# Sandwiched Rényi Convergence for Quantum Evolutions

Alexander Müller-Hermes[1] and Daniel Stilck França[2]

[1]Department of Mathematical Sciences, University of Copenhagen, 2100 Copenhagen, Denmark
[2]Department of Mathematics, Technische Universität München, 85748 Garching, Germany
February 27, 2018

We study the speed of convergence of a primitive quantum time evolution towards its fixed point in the distance of sandwiched Rényi divergences. For each of these distance measures the convergence is typically exponentially fast and the best exponent is given by a constant (similar to a logarithmic Sobolev constant) depending only on the generator of the time evolution. We establish relations between these constants and the logarithmic Sobolev constants as well as the spectral gap. An important consequence of these relations is the derivation of mixing time bounds for time evolutions directly from logarithmic Sobolev inequalities without relying on notions like $l_p$-regularity. We also derive strong converse bounds for the classical capacity of a quantum time evolution and apply these to obtain bounds on the classical capacity of some examples, including stabilizer Hamiltonians under thermal noise.

Alexander Müller-Hermes: muellerh@posteo.net
Daniel Stilck França: dsfranca@mytum.de

arXiv:1607.00041v3 [quant-ph] 26 Feb 2018

# Contents

# 1   Introduction

Consider a quantum system affected by Markovian noise modeled by a quantum dynamical semigroup $T_t$ (with time parameter $t \in \mathbb{R}^+$) driving every initial state towards a unique full rank state $\sigma$. Using the framework of logarithmic Sobolev inequalities as introduced in [1, 2] the speed of the convergence towards the fixed point can be studied. Specifically, the $\alpha_1$-logarithmic Sobolev constant (see [1, 2]) is the optimal exponent $\alpha \in \mathbb{R}^+$ such that the inequality

$$D(T_t(\rho)\|\sigma) \leq e^{-2\alpha t} D(\rho\|\sigma) \tag{1}$$

holds for the quantum Kullback-Leibler divergence, given by $D(\rho\|\sigma) = \operatorname{tr}[\rho(\ln(\rho) - \ln(\sigma))]$, for all $t \in \mathbb{R}^+$ and all states $\rho$.

The framework of logarithmic Sobolev constants is closely linked to properties of noncommutative $l_p$-norms, and specifically to hypercontractivity [1, 2]. Noncommutative $l_p$-norms also appeared recently in the definition of generalized Rényi divergences (so called "sandwiched Rényi divergences" [3, 4]). It is therefore natural to study the relationship between logarithmic Sobolev inequalities and noncommutative $l_p$-norms more closely. The approach used here is to define constants (which we call $\beta_p$ for a parameter $p \in [1, \infty)$), which resemble the logarithmic Sobolev constants, but where the distance measure is a sandwiched Rényi divergence instead of the quantum Kullback-Leibler divergence. More specifically, the constants $\beta_p$ will be the optimal exponents such that inequalities of the form (1) hold for the sandwiched Rényi divergences $D_p$, given by

$$D_p(\rho\|\sigma) = \begin{cases} \frac{1}{p-1} \ln \left( \operatorname{tr} \left[ \left( \sigma^{\frac{1-p}{2p}} \rho \sigma^{\frac{1-p}{2p}} \right)^p \right] \right) & \text{if } \ker(\sigma) \subseteq \ker(\rho) \text{ or } p \in (0, 1) \\ +\infty, & \text{otherwise,} \end{cases} \tag{2}$$

instead of the quantum Kullback-Leibler divergence $D$.

Our main results are two-fold:

- We derive inequalities between the new $\beta_p$ and other quantities such as logarithmic Sobolev constants and the spectral gap of the generator of the time evolution. These inequalities not only reveal basic properties of the $\beta_p$, but can also be used as a technical tool to strengthen results involving logarithmic Sobolev constants.

- We apply our framework to derive bounds on the mixing time of quantum dynamical semigroups. Using the interplay between the $\beta_p$ and the logarithmic Sobolev constants we show how to derive a mixing time bound with the same scaling as that of the one derived in [2] directly from a logarithmic Sobolev constant. Previously, this was only known under the additional assumption of $l_p$-regularity (see [2]) of the generator or for the $\alpha_1$-logarithmic Sobolev constant. It is still an open question whether $l_p$-regularity holds for all primitive generators.

As an additional application of our methods we derive time-dependent strong converse bounds on the classical capacity of a quantum dynamical semigroup. We apply these to some examples of systems under thermal noise. These include stabilizer Hamiltonians, such as the $2D$ toric code, and a truncated harmonic oscillator. To the best of our knowledge, these are the first bounds available on the classical capacity of these channels. We also apply our bound to depolarizing channels, whose classical capacity is known [5], to benchmark our findings.

## 2 Notation and Preliminaries

Throughout this paper $\mathcal{M}_d$ will denote the space of $d \times d$ complex matrices. We will denote by $\mathcal{D}_d$ the set of $d$-dimensional quantum states, i.e. positive semi-definite matrices $\rho \in \mathcal{M}_d$ with trace 1. By $\mathcal{M}_d^+$ we denote the set of positive definite matrices and by $\mathcal{D}_d^+ = \mathcal{M}_d^+ \cap \mathcal{D}_d$ the set of full rank states.

In [3, 4] the following definition of sandwiched quantum Rényi divergences was proposed:

**Definition 2.1** (Sandwiched $p$-Rényi divergence). *Let $\rho, \sigma \in \mathcal{D}_d$. For $p \in (0,1) \cup (1,\infty)$, the **sandwiched $p$-Rényi divergence** is defined as:*

$$
D_p(\rho\|\sigma) = \begin{cases} \frac{1}{p-1} \ln\left( tr\left[ \left( \sigma^{\frac{1-p}{2p}} \rho \sigma^{\frac{1-p}{2p}} \right)^p \right]\right) & \text{if } ker(\sigma) \subseteq ker(\rho) \text{ or } p \in (0,1) \\ +\infty, & \text{otherwise} \end{cases} \tag{3}
$$

*where $ker(\sigma)$ is the kernel of $\sigma$.*

Note that we are using a different normalization than in [3, 4], which is more convenient for our purposes. The logarithm in our definition is in base $e$, while theirs is in base 2. When we write log in later sections we will mean the logarithm in base 2.

Taking the limit $p \to 1$ gives the usual quantum Kullback-Leibler divergence [6]

$$
\lim_{p \to 1} D_p(\rho\|\sigma) = D(\rho\|\sigma) := \begin{cases} \mathrm{tr}[\rho(\ln(\rho) - \ln(\sigma))] & \text{if } \ker(\sigma) \subseteq \ker(\rho) \\ +\infty, & \text{otherwise} \end{cases}.
$$

Similarly by taking the limit $p \to \infty$ we obtain the max-relative entropy [3, Theorem 5]

$$
\lim_{p \to \infty} D_p(\rho\|\sigma) = D_\infty(\rho\|\sigma) = \ln\left( \|\sigma^{-\frac{1}{2}}\rho\sigma^{-\frac{1}{2}}\|_\infty \right).
$$

The sandwiched Rényi divergences increase monotonically in the parameter $p \geq 1$ (see [7, Theorem 7]) and we have

$$
D(\rho\|\sigma) = D_1(\rho\|\sigma) \leq D_p(\rho\|\sigma) \leq D_q(\rho\|\sigma) \leq D_\infty(\rho\|\sigma). \tag{4}
$$

for any $q \geq p \geq 1$ and all $\rho, \sigma \in \mathcal{D}_d$. Next we state two simple consequences of this ordering, which will be useful later.

**Lemma 2.1.** *For $\sigma \in \mathcal{D}_d^+$ and $p \in [1, +\infty)$*

$$
\sup_{\rho \in \mathcal{D}_d} D_p(\rho\|\sigma) = \ln\left( \|\sigma^{-1}\|_\infty \right). \tag{5}
$$

*Proof.* Using (4) for $\rho \in \mathcal{D}_d$ we have

$$
D_p(\rho\|\sigma) \leq D_\infty(\rho\|\sigma) = \ln\left( \|\sigma^{-\frac{1}{2}}\rho\sigma^{-\frac{1}{2}}\|_\infty \right) \leq \ln\left( \|\sigma^{-1}\|_\infty \right).
$$

Here we used that any quantum state $\rho \in \mathcal{D}_d$ fulfills $\rho \leq \mathbb{1}_d$. Clearly, choosing $\rho = |v_{\min}\rangle\langle v_{\min}|$ for an eigenvector $|v_{\min}\rangle \in \mathbb{C}^d$ corresponding to the eigenvalue $\|\sigma^{-1}\|_\infty$ of $\sigma^{-1}$ achieves equality in the previous bound.

$\square$

Using (4) together with the well-known Pinsker inequality [8, Theorem 3.1] for the quantum Kullback-Leibler divergence we have

$$\frac{1}{2}\|\sigma - \rho\|_1^2 \le D(\rho\|\sigma) \le D_p(\rho\|\sigma) \tag{6}$$

for any $p \ge 1$ and all $\rho, \sigma \in \mathcal{D}_d$. The constant $\frac{1}{2}$ has been shown to be optimal in the classical case (see [9]), i.e. restricting to $\rho$ that commute with $\sigma$, and is therefore also optimal here.

## 2.1 Noncommutative $l_p$-spaces

In the following $\sigma \in \mathcal{D}_d^+$ will denote a full rank reference state. For $p \ge 1$ we define the **noncommutative $p$-norm** with respect to $\sigma$ as

$$\|X\|_{p,\sigma} = \left( \mathrm{tr}\left[ \left| \sigma^{\frac{1}{2p}} X \sigma^{\frac{1}{2p}} \right|^p \right] \right)^{\frac{1}{p}} \tag{7}$$

for any $X \in \mathcal{M}_d$. The space $(\mathcal{M}_d, \|\cdot\|_{p,\sigma})$ is called a (weighted) noncommutative $l_p$-space. For a linear map $\Phi : \mathcal{M}_d \to \mathcal{M}_d$ and $p, q \ge 1$ we define the **noncommutative $p \to q$-norm** with respect to $\sigma$ as

$$\|\Phi\|_{p\to q,\sigma} = \sup_{Y \in \mathcal{M}_d} \frac{\|\Phi(Y)\|_{q,\sigma}}{\|Y\|_{p,\sigma}}.$$

We introduce the **weighting operator** $\Gamma_\sigma : \mathcal{M}_d \to \mathcal{M}_d$ as

$$\Gamma_\sigma(X) = \sigma^{\frac{1}{2}} X \sigma^{\frac{1}{2}}.$$

For powers of the weighting operator we set

$$\Gamma_\sigma^p(X) = \sigma^{\frac{p}{2}} X \sigma^{\frac{p}{2}}$$

for $p \in \mathbb{R}$ and $X \in \mathcal{M}_d$. We define the so called **power operator** $I_{p,q} : \mathcal{M}_d \to \mathcal{M}_d$ as

$$I_{p,q}(X) = \Gamma_\sigma^{-\frac{1}{p}}\left( \left| \Gamma_\sigma^{\frac{1}{q}}(X) \right|^{\frac{q}{p}} \right) \tag{8}$$

for $X \in \mathcal{M}_d$. It can be verified that

$$\|I_{p,q}(X)\|_{p,\sigma}^p = \|X\|_{q,\sigma}^q$$

for any $X \in \mathcal{M}_d$. As in the commutative theory, the noncommutative $l_2$-space turns out to be a Hilbert space, where the **weighted scalar product** is given by

$$\langle X, Y \rangle_\sigma = \mathrm{tr}\left[ \Gamma_\sigma\left( X^\dagger \right) Y \right] \tag{9}$$

for $X, Y \in \mathcal{M}_d$. With the above notions we can express the sandwiched $p$-Rényi divergence (3) for $p > 1$ in terms of a noncommutative $l_p$-norm as

$$D_p(\rho\|\sigma) = \frac{1}{p-1} \ln\left( \|\Gamma_\sigma^{-1}(\rho)\|_{p,\sigma}^p \right). \tag{10}$$

For a state $\rho \in \mathcal{D}_d$ the positive matrix $\Gamma_\sigma^{-1}(\rho) \in \mathcal{M}_d$ is called the **relative density** of $\rho$ with respect to $\sigma$. Note that any $X \ge 0$ with $\|X\|_{1,\sigma} = 1$ can be written as $X = \Gamma_\sigma^{-1}(\rho)$ for some state $\rho \in \mathcal{D}_d$. We will simply call operators $X \ge 0$ that satisfy $\|X\|_{1,\sigma} = 1$ relative densities when the reference state is clear.

We refer to [1, 2] and references therein for proofs and more details about the concepts introduced in this section.

## 2.2  Quantum dynamical semigroups

A family of quantum channels, i.e. trace-preserving completely positive maps, $\{T_t\}_{t \in \mathbb{R}_0^+}$, $T_t : \mathcal{M}_d \to \mathcal{M}_d$, parametrized by a non-negative parameter $t \in \mathbb{R}_0^+$ is called a **quantum dynamical semigroup** if $T_0 = \mathrm{id}_d$ (the identity map in $d$ dimensions), $T_{t+s} = T_t \circ T_s$ for any $s, t \in \mathbb{R}_0^+$ and $T_t$ depends continuously on $t$. Any quantum dynamical semigroup can be written as $T_t = e^{t\mathcal{L}}$ (see [10, 11]) for a Liouvillian $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ of the form

$$\mathcal{L}(X) = \mathcal{S}(X) - \kappa X - X \kappa^{\dagger},$$

where $\kappa \in \mathcal{M}_d$ and $\mathcal{S} : \mathcal{M}_d \to \mathcal{M}_d$ is completely positive such that $\mathcal{S}^*(\mathbb{1}_d) = \kappa + \kappa^{\dagger}$, where $\mathcal{S}^*$ is the adjoint of $\mathcal{S}$ with respect to the Hilbert-Schmidt scalar product. We will also deal with tensor powers of semigroups. For a quantum dynamical semigroup $\{T_t\}_{t \in \mathbb{R}^+}$ with Liouvillian $\mathcal{L}$ we denote by $\mathcal{L}^{(n)}$ the Liouvillian of the quantum dynamical semigroup $\{T_t^{\otimes n}\}_{t \in \mathbb{R}^+}$.

In the following we will consider quantum dynamical semigroups having a full rank fixed point $\sigma \in \mathcal{D}_d^+$, i.e. the Liouvillian generating the semigroup fulfills $\mathcal{L}(\sigma) = 0$ (implying that $e^{t\mathcal{L}}(\sigma) = \sigma$ for any time $t \in \mathbb{R}_0^+$). We call a quantum dynamical semigroup (or the Liouvillian generator) **primitive** if it has a unique full rank fixed point $\sigma$. In this case for any initial state $\rho \in \mathcal{D}_d$ we have $\rho_t = e^{t\mathcal{L}}(\rho) \to \sigma$ as $t \to \infty$ (see [12, Theorem 14]).

The notion of primitivity can also be defined for discrete semigroups of quantum channels. For a quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ we will sometimes consider the discrete semigroup $\{T^n\}_{n \in \mathbb{N}}$. Similar to the continuous case we will call this semigroup (or the channel $T$) primitive if there is a unique full rank state $\sigma \in \mathcal{D}_d^+$ with $\lim_{n \to \infty} T^n(\rho) = \sigma$ for any $\rho \in \mathcal{D}_d$. We refer to [12] for other characterizations of primitive channels and sufficient conditions for primitivity.

To study the convergence of a primitive semigroup to its fixed point $\sigma$ we introduce the time evolution of the relative density $X_t = \Gamma_\sigma^{-1}(\rho_t)$. For any Liouvillian $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ with full rank fixed point $\sigma \in \mathcal{D}_d^+$ define

$$\hat{\mathcal{L}} = \Gamma_\sigma^{-1} \circ \mathcal{L} \circ \Gamma_\sigma \tag{11}$$

to be the generator of the time evolution of the relative density. Indeed it can be checked that

$$X_t = \Gamma_\sigma^{-1}\left(e^{t\mathcal{L}}(\rho)\right) = e^{t\hat{\mathcal{L}}}(X)$$

for any state $\rho \in \mathcal{D}_d$ and relative density $X = \Gamma_\sigma^{-1}(\rho)$. Note that $\|X_t\|_{1,\sigma} = \|X\|_{1,\sigma} = 1$ for all $t \in \mathbb{R}_0^+$. Clearly the semigroup generated by $\hat{\mathcal{L}}$ is completely positive and unital, but it is not trace-preserving in general. In the special case where

$$\Gamma_\sigma^{-1} \circ \mathcal{L} \circ \Gamma_\sigma = \mathcal{L}^*, \tag{12}$$

the map $\hat{\mathcal{L}}$ generates the adjoint of the initial semigroup, i.e. the corresponding time evolution in Heisenberg picture. A semigroup fulfilling (12) is called **reversible** (or said to fulfill **detailed balance**), and in this case the Liouvillian $\hat{\mathcal{L}}$ is a Hermitian operator w.r.t. the $\sigma$-weighted scalar product. We again refer to [2, 1] for more details on these topics. For discrete semigroups we similarly set $\hat{T} = \Gamma_\sigma^{-1} \circ T \circ \Gamma_\sigma$.

One important class of semigroups are **Davies generators**, which describe a system weakly coupled to a thermal bath under an appropriate approximation [13]. Describing

them in detail goes beyond the scope of this article and here we will only review their most basic properties. We refer to [14, 15, 16] for more details.

Suppose that we have a system of dimension $d$ weakly coupled to a thermal bath of dimension $d_B$ at inverse inverse temperature $\beta > 0$. Consider a Hamiltonian $H_{\text{tot}} \in \mathcal{M}_d \otimes \mathcal{M}_{d_B}$ of the system and the bath of the form

$$H_{\text{tot}} = H \otimes \mathbb{1}_B + \mathbb{1}_S \otimes H_B + H_I,$$

where $H \in \mathcal{M}_d$ is the Hamiltonian of the system, $H_B \in \mathcal{M}_{d_B}$ of the bath and

$$H_I = \sum_\alpha S^\alpha \otimes B^\alpha \in \mathcal{M}_d \otimes \mathcal{M}_{d_B} \tag{13}$$

describes the interaction between the system and the bath. Here the operators $S^\alpha$ and $B^\alpha$ are self-adjoint. Let $\{\lambda_k\}_{k \in [d]}$ be the spectrum of the Hamiltonian $H$. We then define the Bohr-frequencies $\omega_{i,j}$ to be given by the differences of eigenvalues of $H$, that is, $\omega_{i,j} = \lambda_i - \lambda_j$ for different values of $\lambda$. We will drop the indices on $\omega$ from now on to avoid cumbersome notation, as is usually done. Moreover, we introduce operators $S^\alpha(\omega)$ which are the Fourier components of the coupling operators $S^\alpha$ and satisfy

$$e^{iHt} S^\alpha e^{-iHt} = \sum_\omega S^\alpha(\omega) e^{i\omega t}.$$

The canonical form of the Davies generator at inverse temperature $\beta > 0$ in the Heisenberg picture, $\mathcal{L}_\beta^*$, is then given by

$$\mathcal{L}_\beta^*(X) = i[H, X] + \sum_{\omega, \alpha} \mathcal{L}_{\omega, \alpha}^*(X),$$

where

$$\mathcal{L}_{\omega, \alpha}^*(X) = G^\alpha(\omega) \left( S^\alpha(\omega)^\dagger X S^\alpha(\omega) - \frac{1}{2} \{S^\alpha(\omega)^\dagger S^\alpha(\omega), X\} \right).$$

Here $\{X, Y\} = XY + YX$ is the anticommutator and $G^\alpha : \mathbb{R} \to \mathbb{R}$ are the transition rate functions. Their form depends on the choice of the bath model [15]. For our purposes it will be enough to assume that these are functions that satisfy the KMS condition [17], that is, $G^\alpha(-\omega) = G^\alpha(\omega) e^{-\beta\omega}$. Although this presentation of the Davies generators is admittedly very short, for our purposes it will be enough to note that under some assumptions on the operators $S^\alpha(\omega)$ [18, 19] and on the transition rate functions, the semigroup generated by $\mathcal{L}_\beta$ converges to the thermal state $\frac{e^{-\beta H}}{\text{tr}(e^{-\beta H})}$ and is reversible [17]. In the examples considered here this will always be the case.

## 2.3 Logarithmic Sobolev inequalities and the spectral gap

To study hypercontractive properties and convergence times of primitive quantum dynamical semigroups the framework of logarithmic Sobolev inequalities has been developed in [1, 2]. Here we will briefly introduce this theory. For more details and proofs see [1, 2] and the references therein.

We define the **operator valued relative entropy** (for $p > 1$) of $X \in \mathcal{M}_d^+$ as

$$S_p(X) = -p \frac{d}{ds} I_{p+s, p}(X)|_{s=0}. \tag{14}$$

With this we can define the $p$-relative entropy:

**Definition 2.2** (*p*-relative entropy)**.** *For any full rank $\sigma \in \mathcal{M}_d^+$ and $p > 1$ we define the p-**relative entropy** of $X \in \mathcal{M}_d^+$ as*

$$Ent_{p,\sigma}(X) = \langle I_{q,p}(X), S_p(X) \rangle_\sigma - \|X\|_{p,\sigma}^p \ln(\|X\|_{p,\sigma}), \tag{15}$$

*where $\frac{1}{q} + \frac{1}{p} = 1$. For $p = 1$ we can consistently define*

$$Ent_{1,\sigma}(X) = tr[\Gamma_\sigma(X)(\ln(\Gamma_\sigma(X)) - \ln(\sigma))].$$

*by taking the limit $p \to 1$.*

The *p*-relative entropy is not a divergence in the information-theoretic sense (e.g. it is not contractive under quantum channels). It was originally introduced to study hypercontractive properties of semigroups in [1], where they also show it is positive for positive operators. There is however a connection to the quantum relative entropy as

$$\mathrm{Ent}_{p,\sigma}\left(I_{p,1}\left(\Gamma_\sigma^{-1}(\rho)\right)\right) = \frac{1}{p}D(\rho\|\sigma).$$

As a special case of the last equation we have

$$\mathrm{Ent}_{1,\sigma}\left(\Gamma_\sigma^{-1}(\rho)\right) = D(\rho\|\sigma).$$

We may also use it to obtain an expression for $\mathrm{Ent}_{2,\sigma}$:

$$\mathrm{Ent}_{2,\sigma}(X) = tr\left[\left(\Gamma_\sigma^{\frac{1}{2}}(X)\right)^2 \ln\left(\frac{\Gamma_\sigma^{\frac{1}{2}}(X)}{\|X\|_{2,\sigma}}\right)\right] - \frac{1}{2}tr\left[\left(\Gamma_\sigma^{\frac{1}{2}}(X)\right)^2 \ln(\sigma)\right].$$

We also need Dirichlet forms to define logarithmic Sobolev inequalities:

**Definition 2.3** (Dirichlet form)**.** *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a Liouvillian with full rank fixed point $\sigma \in \mathcal{D}_d^+$. For $p > 1$ we define the p-**Dirichlet form** of $X \in \mathcal{M}_d^+$ as*

$$\mathcal{E}_p^{\mathcal{L}}(X) = -\frac{p}{2(p-1)}\left\langle I_{q,p}(X), \hat{\mathcal{L}}(X) \right\rangle_\sigma$$

*where $\frac{1}{p} + \frac{1}{q} = 1$ and $\hat{\mathcal{L}} = \Gamma_\sigma^{-1} \circ \mathcal{L} \circ \Gamma_\sigma$ denotes the generator of the time evolution of the relative density (cf. (11)). For $p = 1$ we may take the limit $p \to 1$ and consistently define the 1-Dirichlet form by*

$$\mathcal{E}_1^{\mathcal{L}}(X) = -\frac{1}{2}tr\left[\Gamma_\sigma\left(\hat{\mathcal{L}}(X)\right)(\ln(\Gamma_\sigma(X)) - \ln(\sigma))\right].$$

Formally, by making this choice we introduce the logarithmic Sobolev framework for $\hat{\mathcal{L}}$ (i.e. the generator of the time-evolution of the relative density) instead of $\mathcal{L}^*$. While this is a slightly different definition compared to [2], where the Heisenberg picture is used, they are the same for reversible Liouvillians.

In [1] the Dirichlet forms were introduced to study hypercontractive properties of semigroups. As we will see in Theorem 3.1, they appear naturally when we compute the entropy production of the Sandwiched Rényi divergences. From Corollary 3.1 we will be able to infer that the Dirichlet form is positive for positive operators, a fact already proved in [1]. Both the $\mathrm{Ent}_{p,\sigma}$ and the Dirichlet form are intimately related to hypercontractive

properties of semigroups, as we have for a relative density $X$, some constant $\alpha > 0$ and $p(t) = 1 + e^{2\alpha t}$ that

$$\frac{d}{dt} \ln \left( \|X_t\|_{p(t),\sigma} \right) = \frac{\alpha e^{\alpha t}}{(1 + e^{\alpha t}) \|X_t\|_{p(t),\sigma}^{p(t)}} \left( \mathrm{Ent}_{p(t),\sigma}(X_t) - \frac{1}{\alpha} \mathcal{E}_{p(t)}(X_t) \right),$$

as shown in [1].

Notice that when working with $\mathcal{E}_2^{\mathcal{L}}$ we may always suppose the Liouvillian is reversible without loss of generality. This follows from the fact that

$$\mathcal{E}_2^{\mathcal{L}}(X) = - \left\langle X, \hat{\mathcal{L}}(X) \right\rangle_\sigma$$

is invariant under the additive symmetrization $\hat{\mathcal{L}} \mapsto \frac{1}{2} \left( \mathcal{L}^* + \Gamma_\sigma^{-1} \circ \mathcal{L} \circ \Gamma_\sigma \right)$ for $X \geq 0$.

We can now introduce the logarithmic Sobolev constants:

**Definition 2.4** (Logarithmic Sobolev constants). *For a Liouvillian $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ with full rank fixed point $\sigma \in \mathcal{D}_d^+$ and $p \geq 1$ the p-**logarithmic Sobolev constant** is defined as*

$$\alpha_p \left( \mathcal{L} \right) = \sup\{\alpha \in \mathbb{R}^+ : \alpha \, Ent_{p,\sigma}(X) \leq \mathcal{E}_p^{\mathcal{L}}(X) \text{ for all } X > 0\} \tag{16}$$

As $\mathrm{Ent}_{2,\sigma}$ does not depend on $\mathcal{L}$ and, as remarked before, $\mathcal{E}_2^{\mathcal{L}}$ is invariant under an additive symmetrization, we may always assume without loss of generality that the Liouvillian is reversible when working with $\alpha_2$.

For any $X \in \mathcal{M}_d^+$ we can define its **variance** with respect to $\sigma \in \mathcal{D}_d^+$ as

$$\mathrm{Var}_\sigma \left( X \right) = \|X\|_{2,\sigma}^2 - \|X\|_{1,\sigma}^2. \tag{17}$$

This defines a distance measure to study the convergence of the semigroup. Given a Liouvillian $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ with fixed point $\sigma \in \mathcal{D}_d^+$ we define its **spectral gap** as

$$\lambda(\mathcal{L}) = \sup \left\{ \lambda \in \mathbb{R}^+ : \lambda \mathrm{Var}_\sigma \left( X \right) \leq \mathcal{E}_2^{\mathcal{L}}(X) \text{ for all } X > 0 \right\} \tag{18}$$

where $\hat{\mathcal{L}} : \mathcal{M}_d \to \mathcal{M}_d$ is given by (11). We can always assume the Liouvillian to be reversible when dealing with the spectral gap, as it again depends on $\mathcal{E}_2^{\mathcal{L}}$.

The spectral gap can be used to bound the convergence in the variance (see [20]), as for any $X \in \mathcal{M}_d^+$ we have

$$\frac{d}{dt} \mathrm{Var}_\sigma(X_t) = 2 \left\langle \hat{\mathcal{L}} \left( X \right), X \right\rangle_\sigma \tag{19}$$

and so

$$\mathrm{Var}_\sigma \left( X_t \right) \leq e^{-2\lambda t} \mathrm{Var}_\sigma \left( X \right). \tag{20}$$

## 3  Convergence rates for sandwiched Rényi divergences

In this section we consider the sandwiched Rényi divergences of a state evolving under a primitive quantum dynamical semigroup and the fixed point of this semigroup. It is clear that these quantities converge to zero as the time-evolved state approaches the fixed point. To study the speed of this convergence we introduce a differential inequality, which can be seen as an analogue of the logarithmic Sobolev inequalities for sandwiched Rényi divergences.

## 3.1 Rényi-entropy production

In [18] the entropy production for the quantum Kullback-Leibler divergence of a Liouvillian was computed. We will now derive a similar expression for the entropy production for the $p$-Rényi divergences for $p > 1$.

**Theorem 3.1** (Derivative of the sandwiched $p$-Rényi divergence)**.** *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a Liouvillian with full rank fixed point $\sigma \in \mathcal{D}_d^+$. For any $\rho \in \mathcal{D}_d$ and $p > 1$ we have*

$$\frac{d}{dt} D_p(e^{t\mathcal{L}}(\rho)\|\sigma)\Big|_{t=0} = \frac{p}{p-1} \frac{tr\left[\left(\sigma^{\frac{1-p}{2p}} \rho \sigma^{\frac{1-p}{2p}}\right)^{p-1} \sigma^{\frac{1-p}{2p}} \mathcal{L}(\rho) \sigma^{\frac{1-p}{2p}}\right]}{tr\left[\left(\sigma^{\frac{1-p}{2p}} \rho \sigma^{\frac{1-p}{2p}}\right)^{p}\right]}. \tag{21}$$

*Using the relative density $X = \Gamma_\sigma^{-1}(\rho)$ and (11) this expression can be written as:*

$$\frac{d}{dt} D_p(e^{t\mathcal{L}}(\rho)\|\sigma)\Big|_{t=0} = \frac{p}{p-1} \|X\|_{p,\sigma}^{-p} \left\langle I_{q,p}(X), \hat{\mathcal{L}}(X)\right\rangle_\sigma \tag{22}$$

*with $\frac{1}{p} + \frac{1}{q} = 1$.*

*Proof.* Rewriting the $p$-Rényi divergence in terms of the relative density $X = \Gamma_\sigma^{-1}(\rho)$ and the corresponding generator $\hat{\mathcal{L}} = \Gamma_\sigma^{-1} \circ \mathcal{L} \circ \Gamma_\sigma$ (see (11)) we have

$$D_p(e^{t\mathcal{L}}(\rho)\|\sigma) = \frac{1}{p-1} \ln\left(\|e^{t\hat{\mathcal{L}}}(X)\|_{p,\sigma}^p\right). \tag{23}$$

By the chain rule

$$\frac{d}{dt} D_p(e^{t\mathcal{L}}\rho\|\sigma)\Big|_{t=0} = \frac{1}{p-1} \|X\|_{p,\sigma}^{-p} \left(\frac{d}{dt}\|e^{t\hat{\mathcal{L}}}(X)\|_{p,\sigma}^p\right)\Big|_{t=0}.$$

Define the curve $\gamma : \mathbb{R}_0^+ \to \mathcal{M}_d$ as $\gamma(t) = \sigma^{\frac{1}{2p}} e^{t\hat{\mathcal{L}}}(X) \sigma^{\frac{1}{2p}}$ and observe that

$$\|e^{t\hat{\mathcal{L}}}(X)\|_{p,\sigma}^p = tr[\gamma(t)^p].$$

As the differential of the function $X \mapsto X^p$ at $A \in \mathcal{M}_d^+$ is given by $pA^{p-1}$, another application of the chain rule yields

$$\frac{d}{dt}\|e^{t\hat{\mathcal{L}}}(X)\|_{p,\sigma}^p\Big|_{t=0} = p\left\langle \gamma(0)^{p-1}, \frac{d\gamma}{dt}(0)\right\rangle.$$

It is easy to check that $\frac{d\gamma}{dt}(0) = \sigma^{\frac{1}{2p}} \hat{\mathcal{L}}(X) \sigma^{\frac{1}{2p}}$. Inserting this in the above equations and writing it in terms of the power operator (8) we finally obtain

$$\frac{d}{dt} D_p(e^{t\mathcal{L}}(\rho)\|\sigma)\Big|_{t=0} = \frac{p}{p-1} \|X\|_{p,\sigma}^{-p} \left\langle I_{q,p}(X), \hat{\mathcal{L}}(X)\right\rangle_\sigma$$

with $\frac{1}{p} + \frac{1}{q} = 1$. Expanding this formula gives (21).

$\square$

By recognizing the $p$-Dirichlet form in the previous theorem we get:

**Corollary 3.1.** *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a Liouvillian with full rank fixed point $\sigma \in \mathcal{D}_d^+$. For any $\rho \in \mathcal{D}_d$ and $p > 1$ we have*

$$\frac{d}{dt} D_p(e^{t\mathcal{L}}(\rho)\|\sigma)\Big|_{t=0} = -2\|X\|_{p,\sigma}^{-p} \mathcal{E}_p^{\mathcal{L}}(X) \leq 0, \tag{24}$$

*where we used the relative density $X = \Gamma_\sigma^{-1}(\rho)$.*

As we remarked before, Corollary 3.1 implies that the Dirichlet form is always positive for relative densities. To see this, recall that the divergences contract under quantum channels [7] and therefore we have that $\frac{d}{dt} D_p(e^{t\mathcal{L}}(\rho)\|\sigma)\Big|_{t=0} \leq 0$. As $\mathcal{E}_p^{\mathcal{L}}(\lambda X) = \lambda^p \mathcal{E}_p^{\mathcal{L}}(X)$ for $\lambda > 0$, this shows that it is positive for all positive operators by properly normalizing $X$.

## 3.2 Sandwiched Rényi convergence rates

For any $p > 1$ we introduce the functional $\kappa_p : \mathcal{M}_d^+ \to \mathbb{R}$ as

$$\kappa_p(X) = \frac{1}{p-1} \|X\|_{p,\sigma}^p \ln\left( \frac{\|X\|_{p,\sigma}^p}{\|X\|_{1,\sigma}^p} \right) \tag{25}$$

for $X \in \mathcal{M}_d^+$. For $p = 1$ we may again take the limit $p \to 1$ and obtain $\kappa_1(X) := \lim_{p\to 1} \kappa_p(X) = \mathrm{Ent}_{1,\sigma}(X)$. Note that $\kappa_p$ is well-defined and non-negative as $\|X\|_{p,\sigma} \geq \|X\|_{1,\sigma}$ for $p \geq 1$. Strictly speaking the definition also depends on a reference state $\sigma \in \mathcal{D}_d^+$, which we usually omit as it is always the fixed point of the primitive Liouvillian under consideration.

Given a Liouvillian $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ with full rank fixed point $\sigma \in \mathcal{D}_d^+$ it is a simple consequence of Corollary 3.1 that for $\rho \neq \sigma$

$$\frac{\frac{d}{dt} D_p(e^{t\mathcal{L}}(\rho)\|\sigma)\Big|_{t=0}}{D_p(\rho\|\sigma)} = -2\frac{\mathcal{E}_p^{\mathcal{L}}(X)}{\kappa_p(X)}, \tag{26}$$

where we used the relative density $X = \Gamma_\sigma^{-1}(\rho)$, which fulfills $\|X\|_{1,\sigma} = 1$. This motivates the following definition.

**Definition 3.1** (Entropic convergence constant for $p$-Rényi divergence)**.** *For any primitive Liouvillian $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ and $p \geq 1$ we define*

$$\beta_p(\mathcal{L}) = \sup\{\beta \in \mathbb{R}^+ : \beta\kappa_p(X) \leq \mathcal{E}_p^{\mathcal{L}}(X) \text{ for all } X > 0\}. \tag{27}$$

Note that as a special case we have $\alpha_1(\mathcal{L}) = \beta_1(\mathcal{L})$. It should be also emphasized that the supremum in the previous definition goes over any positive definite $X \in \mathcal{M}_d^+$ and not only over relative densities. However, it is easy to see that we can equivalently write

$$\beta_p(\mathcal{L}) = \inf\left\{ \frac{\mathcal{E}_p^{\mathcal{L}}(X)}{\kappa_p(X)} : X > 0 \right\} = \inf\left\{ \frac{\mathcal{E}_p^{\mathcal{L}}(X)}{\kappa_p(X)} : X > 0, \|X\|_{1,\sigma} = 1 \right\} \tag{28}$$

as replacing $X \mapsto X/\|X\|_{1,\sigma}$ does not change the value of the quotient $\mathcal{E}_p^{\mathcal{L}}(X)/\kappa_p(X)$. Therefore, to compute $\beta_p$ it is enough to optimize over relative densities (i.e. $X > 0$ fulfilling $\|X\|_{1,\sigma} = 1$). By inserting $\beta_p$ into (26) we have

$$\frac{d}{dt} D_p(e^{t\mathcal{L}}(\rho)\|\sigma) \leq -2\beta_p(\mathcal{L}) D_p\left( e^{t\mathcal{L}}(\rho)\|\sigma \right)$$

for any $\rho \in \mathcal{D}_d$ and Liouvillian $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ with full rank fixed point $\sigma \in \mathcal{D}_d^+$. By integrating this differential inequality we get

**Theorem 3.2.** *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a Liouvillian with full rank fixed point $\sigma \in \mathcal{D}_d^+$. For any $p \geq 1$ and $\rho \in \mathcal{D}_d$ we have*

$$D_p\left(e^{t\mathcal{L}}(\rho)\|\sigma\right) \leq e^{-2\beta_p(\mathcal{L})t} D_p\left(\rho\|\sigma\right) \tag{29}$$

*where $\beta_p(\mathcal{L})$ is the constant defined in (27).*

## 3.3 Computing $\beta_p$ in simple cases

In general it is not clear how to compute $\beta_p$ and it does not depend on spectral data of $\mathcal{L}$ alone. This is not surprising, as the computation of the usual logarithmic Sobolev constants $\alpha_2$ or $\alpha_1$ is also challenging and the exact values are only known for few Liouvillians [21, 2, 22]. In the following we compute $\beta_2$ for the depolarizing semigroups.

**Theorem 3.3** ($\beta_2$ for the depolarizing Liouvillian)**.** *Let $\mathcal{L}_\sigma : \mathcal{M}_d \to \mathcal{M}_d$ denote the depolarizing Liouvillian given by $\mathcal{L}_\sigma(\rho) = tr(\rho)\,\sigma - \rho$ with fixed point $\sigma \in \mathcal{D}_d^+$. Then*

$$\beta_2(\mathcal{L}_\sigma) = \frac{1 - \frac{1}{\|\sigma^{-1}\|_\infty}}{\ln\left(\|\sigma^{-1}\|_\infty\right)}. \tag{30}$$

*Proof.* Without loss of generality we can restrict to $X > 0$ with $\|X\|_{1,\sigma} = 1$ in the minimization (28). Observe that the generator of the time evolution of the relative density (see (11)) for the depolarizing Liouvillian is

$$\hat{\mathcal{L}}_\sigma(X) = tr\left(\sigma^{\frac{1}{2}} X \sigma^{\frac{1}{2}}\right) \mathbb{1} - X.$$

An easy computation yields $\mathcal{E}_2^{\mathcal{L}_\sigma}(X) = \|X\|_{2,\sigma}^2 - 1$ and so

$$\frac{\mathcal{E}_2^{\mathcal{L}_\sigma}(X)}{\kappa_2(X)} = \frac{1 - \frac{1}{\|X\|_{2,\sigma}^2}}{\ln\left(\|X\|_{2,\sigma}^2\right)}.$$

As the function $x \mapsto \frac{1 - \frac{1}{x}}{\ln(x)}$ is monotone decreasing for $x \geq 1$, we have

$$\inf_{X>0} \frac{\mathcal{E}_2^{\mathcal{L}_\sigma}(X)}{\kappa_2(X)} = \frac{1 - \frac{1}{\|\sigma^{-1}\|_\infty}}{\ln\left(\|\sigma^{-1}\|_\infty\right)}, \tag{31}$$

where we used

$$\sup_{X \geq 0, \|X\|_{1,\sigma}=1} \|X\|_{2,\sigma}^2 = \|\sigma^{-1}\|_\infty,$$

which easily follows from Lemma 2.1 by exponentiating both sides of Equation (5) and using the correspondence between relative densities and states. $\qquad\square$

The exact value of $\alpha_2(\mathcal{L}_\sigma)$ is open to the best of our knowledge, but in the case of $\sigma = \frac{1}{d}$ we have $\alpha_2\left(\mathcal{L}_{\frac{1}{d}}\right) = \frac{2(1-2/d)}{\ln(d-1)}$ [2, Theorem 24], which is of the same order of magnitude as $\beta_2$ for these semigroups.

Computing $\beta_p$ for $p \neq 2$ seems not to be straightforward even for depolarizing channels, but for the semigroup depolarizing to the maximally mixed state we can at least provide upper and lower bounds.

**Theorem 3.4** ($\beta_p$ for the Liouvillian depolarizing to the maximally mixed state)**.** *Let* $\mathcal{L}(\rho) = tr(\rho)\frac{1}{d} - \rho.$ *For $p \geq 2$ we have*

$$\frac{p}{2(p-1)}\frac{1}{\ln(d)} \geq \beta_p(\mathcal{L}) \geq \frac{p}{2(p-1)}\frac{d^{\frac{p-1}{p}} - 1}{d^{\frac{p-1}{p}}\ln(d)}.$$

*Proof.* The Dirichlet Form of this Liouvillian for $X > 0$ with $\|X\|_{1,\frac{1}{d}} = 1$ is given by

$$\mathcal{E}_p^{\mathcal{L}}(X) = \frac{p}{2(p-1)}(\|X\|_{p,\frac{1}{d}}^p - \|X\|_{p-1,\frac{1}{d}}^{p-1}).$$

Dividing this expression by $\kappa_p(X)$ we get

$$\frac{\mathcal{E}_p^{\mathcal{L}}(X)}{\kappa_p(X)} = \frac{1 - \frac{\|X\|_{p-1,\frac{1}{d}}^{p-1}}{\|X\|_{p,\frac{1}{d}}^p}}{2\ln\left(\|X\|_{p,\frac{1}{d}}\right)}. \tag{32}$$

By the monotonicity of the weighted norms, we have

$$\frac{\|X\|_{p-1,\frac{1}{d}}^{p-1}}{\|X\|_{p,\frac{1}{d}}^p} \leq \frac{1}{\|X\|_{p,\frac{1}{d}}}$$

and so

$$\frac{\mathcal{E}_p^{\mathcal{L}}(X)}{\kappa_p(X)} \geq \frac{\|X\|_{p,\frac{1}{d}} - 1}{2\|X\|_{p,\frac{1}{d}}\ln\left(\|X\|_{p,\frac{1}{d}}\right)} \tag{33}$$

The expression on the right-hand side of (33) is monotone decreasing in $\|X\|_{p,\frac{1}{d}}$ and so the infimum is attained at

$$\sup_{\|X\|_{1,\frac{1}{d}}=1} \|X\|_{p,\frac{1}{d}} = d^{\frac{p-1}{p}},$$

which again easily follows from Lemma 2.1. The upper bound follows from (32) as

$$\frac{\mathcal{E}_p^{\mathcal{L}}(X)}{\kappa_p(X)} \leq \frac{1}{2\ln\left(\|X\|_{p,\frac{1}{d}}\right)}.$$

which is again monotone decreasing in $\|X\|_{p,\frac{1}{d}}$. $\qquad\square$

From the relations between LS constants [2, Proposition 13], it follows that for the LS constants of the depolarizing channels we have $\alpha_p\left(\mathcal{L}_{\frac{1}{d}}\right) \geq \alpha_2\left(\mathcal{L}_{\frac{1}{d}}\right) = \frac{2(1-2/d)}{\ln(d-1)}$ for $p \geq 1$. The constants $\beta_p$ and $\alpha_p$ are therefore of the same order in this case for small $p \geq 2$.

## 4 Comparison with similar quantities

### 4.1 Comparison with spectral gap

Here we show how $\beta_p$, see (27), compares to the spectral gap (18) of a Liouvillian. This is motivated by similar results for logarithmic Sobolev constants, where it was shown [2, Theorem 16] that $\alpha_1(\mathcal{L}) \leq \lambda(\mathcal{L})$ for reversible semigroups, a result we recover and generalize here.

---

**Theorem 4.1** (Upper bound spectral gap). *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a primitive and reversible Liouvillian with full rank fixed point $\sigma \in \mathcal{D}_d^+$ and $p \geq 1$. Then*

$$\beta_p(\mathcal{L}) \leq \lambda(\mathcal{L}). \tag{34}$$

*Proof.* Let $(s_i)_{i=1}^d$ denote the spectrum of $\sigma^{1/p}$ and choose a unitary $U$ such that

$$\sigma^{1/p} = U \text{diag}(s_1, s_2, \ldots, s_d) U^\dagger.$$

As $\mathcal{L}$ is reversible, there is a self-adjoint eigenvector $X \in \mathcal{M}_d$ of $\hat{\mathcal{L}}$ corresponding to the spectral gap, i.e. $\hat{\mathcal{L}}(X) = -\lambda(\mathcal{L})X$. Let $\epsilon_0 > 0$ be small enough such that $Y_\epsilon = \mathbb{1}_d + \epsilon X$ is positive for any $|\epsilon| \leq \epsilon_0$. For $|\epsilon| \leq \epsilon_0$ we use Lemma A.1 of the appendix to show

$$\beta_p(\mathcal{L}) \leq \frac{\mathcal{E}_p^{\mathcal{L}}(Y_\epsilon)}{\kappa_p(Y_\epsilon)} = \frac{\lambda(\mathcal{L}) \frac{p}{2(p-1)} \left( 2\epsilon^2 \sum_{1 \leq i \leq j \leq d} f_p(s_i, s_j) b_{ij} b_{ji} + O(\epsilon^3) \right)}{\frac{\epsilon^2}{p-1} \left( p \sum_{1 \leq i \leq j \leq d} f_p(s_i, s_j) b_{ij} b_{ji} \right) + O(\epsilon^3)} \tag{35}$$

where $b_{ij} = (U^\dagger \sigma^{1/2p} X \sigma^{1/2p} U)_{ij}$ and

$$f_p(x, y) = \begin{cases} (p-1)x^{p-2} & \text{if } x = y \\ \frac{x^{p-1} - y^{p-1}}{x-y} & \text{else.} \end{cases} \tag{36}$$

Observe that $f_p(s_i, s_j) > 0$ for $s_i, s_j > 0$. Moreover, as $U^\dagger \sigma^{1/2p} X \sigma^{1/2p} U$ is non-zero and self-adjoint we have $b_{ij} b_{ji} \geq 0$ for all $i, j$ and this inequality is strict for at least one choice of $i, j$. Therefore, the terms of second order in $\epsilon$ in the numerator and denominator of (35) are strictly positive, and we obtain $\lambda(\mathcal{L})$ as the limit of the quotient as $\epsilon \to 0$. $\square$

A similar argument as the one given in the previous proof shows that all real, nonzero elements of the spectrum of $\hat{\mathcal{L}}$ are upper bounds to $\beta_p$ without invoking reversibility.

Note that in the case of $p = 2$ (see the discussion after (16)) we may assume that the Liouvillian is reversible without loss of generality and drop the requirement of reversibility in the previous theorem. Alternatively, we can obtain the same statement directly from a simple functional inequality. In this case we can also give a lower bound on $\beta_2$ in terms of the spectral gap.

**Theorem 4.2** (Upper and lower bound for $\beta_2$). *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a primitive Liouvillian with full rank fixed point $\sigma \in \mathcal{D}_d^+$. Then*

$$\lambda(\mathcal{L}) \frac{1 - \frac{1}{\|\sigma^{-1}\|_\infty}}{\ln(\|\sigma^{-1}\|_\infty)} \leq \beta_2(\mathcal{L}) \leq \lambda(\mathcal{L}). \tag{37}$$

To prove Theorem 4.2 we need the following Lemma.

**Lemma 4.1.** *For any $X \in \mathcal{M}_d$ we have*

$$Var_\sigma(X) \leq \kappa_2(X).$$

*Proof.* For $X > 0$ dividing both sides of the inequality by $\|X\|_{1,\sigma}^2$ yields

$$\frac{\|X\|_{2,\sigma}^2}{\|X\|_{1,\sigma}^2} - 1 \leq \frac{\|X\|_{2,\sigma}^2}{\|X\|_{1,\sigma}^2} \ln \left( \frac{\|X\|_{2,\sigma}^2}{\|X\|_{1,\sigma}^2} \right).$$

This follows from the elementary inequality $x - 1 \leq x \ln(x)$ for $x \geq 1$, where we use the ordering $\|X\|_{2,\sigma} \geq \|X\|_{1,\sigma}$ for any $X \in \mathcal{M}_d$.

$\square$

*Proof of Theorem 4.2.* Using the definition of $\beta_2$ (see (27)) and Lemma 4.1 yields

$$\beta_2 \mathrm{Var}_\sigma(X) \leq \beta_2 \kappa_2(X) \leq \mathcal{E}_2^{\mathcal{L}}(X).$$

Now the variational definition of $\lambda(\mathcal{L})$ (see (18)) implies the second inequality of (37).

To prove the first inequality of (37) consider the depolarizing Liouvillian

$$\mathcal{L}_\sigma(X) = \mathrm{tr}(X)\sigma - X.$$

By Theorem 3.3 we have

$$\frac{1 - \frac{1}{\|\sigma^{-1}\|_\infty}}{\ln\left(\|\sigma^{-1}\|_\infty\right)} \kappa_2(X) \leq \mathcal{E}_2^{\mathcal{L}_\sigma}(X)$$

As $\mathcal{E}_2^{\mathcal{L}_\sigma}(X) = \mathrm{Var}_\sigma(X)$, we have $\mathcal{E}_2^{\mathcal{L}_\sigma}(X) \leq \frac{1}{\lambda(\mathcal{L})} \mathcal{E}_2^{\mathcal{L}}(X)$ by the variational definition of $\lambda(\mathcal{L})$ (see (18)). Inserting this in the above inequality finishes the proof.

$\square$

## 4.2 Comparison with logarithmic Sobolev constants

Here we show how $\beta_p$, see (27), compares to the logarithmic Sobolev constant $\alpha_p$.

**Theorem 4.3.** *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a primitive Liouvillian with full rank fixed point $\sigma \in \mathcal{D}_d^+$. Then for any $p \geq 1$ we have*

$$\beta_p(\mathcal{L}) \geq \frac{\alpha_p(\mathcal{L})}{p}. \tag{38}$$

We will need the following Lemma.

**Lemma 4.2.** *For any full rank state $\sigma \in \mathcal{D}_d^+$, any $p > 1$ and $X \in \mathcal{M}_d^+$ with $\|X\|_{1,\sigma} = 1$ we have*

$$Ent_{p,\sigma}(X) \geq \frac{\kappa_p(X)}{p}. \tag{39}$$

*Proof.* The function $p \mapsto D_p(\rho\|\sigma)$ is monotonically increasing [3, 7] and differentiable (as the noncommutative $l_p$-norm is differentiable in $p$ [1, Theorem 2.7]). Thus, with $f : \mathbb{R}^+ \to \mathbb{R}$ given by $f(t) = t + p$ we have

$$0 \leq \|X\|_{p,\sigma}^p \frac{d}{dt} \left( D_{f(t)}(\rho\|\sigma) \right) \Big|_{t=0} = -\frac{1}{(p-1)^2} \|X\|_{p,\sigma}^p \ln \left( \|X\|_{p,\sigma}^p \right) + \frac{1}{p-1} \frac{d}{dt} \left( \|X\|_{f(t),\sigma}^{f(t)} \right) \Big|_{t=0}.$$

where we used the relative density $X = \Gamma_\sigma^{-1}(\rho)$. The remaining derivative in the above equation has been computed in [1, Theorem 2.7] and we have

$$\frac{d}{dt}\left(\|X\|_{f(t),\sigma}^{f(t)}\right)\Big|_{t=0} = \langle I_{q,p}(X), S_p(X)\rangle_\sigma$$

with the operator valued entropy $S_p$ defined in (14) and $\frac{1}{p} + \frac{1}{q} = 1$. Inserting this expression in the above equation we obtain

$$\frac{1}{p-1}\|X\|_{p,\sigma}^p \ln\left(\|X\|_{p,\sigma}^p\right) \leq \langle I_{q,p}(X), S_p(X)\rangle_\sigma. \tag{40}$$

for any $X \in \mathcal{M}_d^+$ with $\|X\|_{1,\sigma} = 1$, i.e. for any $X = \Gamma_\sigma^{-1}(\rho)$ for some state $\rho \in \mathcal{D}_d$. Now we get

$$\begin{aligned}
p\mathrm{Ent}_{p,\sigma}(X) &= p\langle I_{q,p}(X), S_p(X)\rangle_\sigma - \|X\|_{p,\sigma}^p \ln(\|X\|_{p,\sigma}^p) \\
&\geq \frac{p}{p-1}\|X\|_{p,\sigma}^p \ln(\|X\|_{p,\sigma}^p) - \|X\|_{p,\sigma}^p \ln(\|X\|_{p,\sigma}^p) \\
&= \kappa_p(X)
\end{aligned}$$

where we used (40).

$\square$

*Proof of Theorem 4.3.* There is nothing to show for $p = 1$ as $\alpha_1(\mathcal{L}) = \beta_1(\mathcal{L})$ and we can assume $p > 1$. For $X \in \mathcal{M}_d^+$ with $\|X\|_{1,\sigma} = 1$ we can use Lemma 4.2 and the definition of $\alpha_p(\mathcal{L})$ to compute

$$\frac{\alpha_p(\mathcal{L})}{p}\kappa_p(X) \leq \alpha_p(\mathcal{L})\mathrm{Ent}_{p,\sigma}(X) \leq \mathcal{E}_p^{\mathcal{L}}(X).$$

By the variational definition (27) of $\beta_p$ the claim follows.

$\square$

Theorem 4.3 will be applied in Section 5 to obtain bounds on the mixing time of a Liouvillian with a positive logarithmic Sobolev constant without invoking any form of $l_p$-regularity (see [2]). As usually a logarithmic Sobolev is implied by a hypercontractive inequality [1], we would like to remark that one can also make a similar statement as that of Theorem 4.3 from a hypercontractive inequality. One can easily show that

$$||e^{t\hat{\mathcal{L}}}||_{p(t)\to p,\sigma} \leq 1 \tag{41}$$

for $p(t) = (p-1)e^{-\alpha_p t} + 1$ implies that $\beta_p(\mathcal{L}) \geq \frac{\alpha_p}{p}$.

## 5  Mixing times

In this section we will introduce the quantities of interest and prove the building blocks to prove mixing times from the entropy production inequalities of the last sections, distinguishing between continuous and discrete time semigroups. We will mostly focus on $\beta_2$, as this seems to be the most relevant constant for mixing time applications. This is justified by the fact that the underlying Dirichlet form is a quadratic form and the entropy related to it stems from a Hilbert space norm. Moreover, as the same Dirichlet form is also involved in computations of the spectral gap, it could be easier to adapt existing techniques, such as the ones developed in [23, 19].

**Definition 5.1** (Mixing times)**.** *For either $I = \mathbb{R}^+$ or $I = \mathbb{N}$ let $\{T_t\}_{t\in I}$ be a primitive semigroup of quantum channels with fixed point $\sigma \in \mathcal{D}_d^+$. We define the $l_1$ **mixing time** for $\epsilon > 0$ as*

$$t_1(\epsilon) = \inf\{t \in I \ : \ \|T_t(\rho) - \sigma\|_1 \leq \epsilon \text{ for all } \rho \in \mathcal{D}_d\}.$$

*Similarly we define the $l_2$ **mixing time** for $\epsilon > 0$ as*

$$t_2(\epsilon) = \inf\{t \in I \ : \ Var_\sigma\left(\hat{T}_t\left(X\right)\right) \leq \epsilon \text{ for all } X \in \mathcal{M}_d^+ \text{ with } \|X\|_{1,\sigma} = 1\}.$$

*In the continuous case $I = \mathbb{R}^+$ we will often speak of the mixing times of the Liouvillian generator of a quantum dynamical semigroup which we identify with the mixing times of the semigroup according to the above definition.*

## 5.1 Mixing in Continuous Time

It is now straightforward to get mixing times from the previous results.

**Theorem 5.1** (Mixing time from entropy production)**.** *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a primitive Liouvillian with fixed point $\sigma \in \mathcal{D}_d^+$. Then*

$$t_1(\epsilon) \leq \frac{1}{2\beta_p(\mathcal{L})} \ln\left(\frac{2\ln\left(\|\sigma^{-1}\|_\infty\right)}{\epsilon^2}\right).$$

*Proof.* From (6) and Lemma 2.1 we have

$$\ln\left(\|\sigma^{-1}\|_\infty\right) e^{-2\beta_p(\mathcal{L})t} \geq \frac{1}{2}\|e^{t\mathcal{L}}(\rho) - \sigma\|_1^2. \tag{42}$$

for any $\rho \in \mathcal{D}_d$. The claim follows after rearranging the terms. $\qquad\square$

Using Theorem 4.3 we can lower bound $\beta_p$ in terms of the usual logarithmic Sobolev constant $\alpha_p$. Combining this with Theorem 5.1 shows the following Corollary.

**Corollary 5.1** (Mixing time bound from logarithmic Sobolev inequalities)**.** *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a primitive Liouvillian with fixed point $\sigma \in \mathcal{D}_d^+$. Then*

$$t_1(\epsilon) \leq \frac{p}{2\alpha_p(\mathcal{L})} \ln\left(\frac{2\ln\left(\|\sigma^{-1}\|_\infty\right)}{\epsilon^2}\right). \tag{43}$$

By Corollary 5.1 a nonzero logarithmic Sobolev constant always implies a nontrivial mixing time bound. One should say that the same bound was showed in [2] for $p = 2$, however under additional assumptions (specifically $l_p$-regularity [2]) on the Liouvillian in question. While these assumptions have been shown for certain classes of Liouvillians (including important examples like Davies generators and doubly stochastic Liouvillians [2]) they have not been shown in general. Moreover, the bound in Theorem 5.1 clearly does not depend on $p$ and one could in principle optimize over all $\beta_p$. However, as the computations in subsection 3.3 already indicate, it does not seem to be feasible to compute or bound $\beta_p$ for $p \neq 2$ even in simple cases and one will probably only work with $\beta_2$ in applications.

The bound from Corollary 5.1 also has the right scaling properties needed in recent applications of rapid mixing, such as the results in [24, 25]. In particular, together with the results in [26], the last Corollary shows that the hypothesis of Theorem 4.2 in [24] is

---

always satisfied for product evolutions and not only for the special classes considered in [2].

One may also use these techniques to get mixing times in the $l_2$ norms which are stronger than the ones obtained just by considering that $\beta_2$ is a lower bound to the spectral gap.

**Theorem 5.2** ($l_2$-mixing time bound). *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a Liouvillian with fixed point $\sigma \in \mathcal{D}_d^+$. Then*

$$t_2(\epsilon) \leq \frac{1}{2\beta_2(\mathcal{L})} \ln\left(\frac{\ln\left(\|\sigma^{-1}\|_\infty\right)}{\ln(1+\epsilon)}\right). \tag{44}$$

*Proof.* For $X > 0$ with $\|X\|_{1,\sigma} = 1$ we have $\mathrm{Var}_\sigma(X) = \|X - \mathbb{1}\|_{2,\sigma}^2 = \|X\|_{2,\sigma}^2 - 1$ and thus

$$\kappa_2(X) = (1 + \mathrm{Var}_\sigma(X))\ln\left(1 + \mathrm{Var}_\sigma(X)\right).$$

In the following let $X_t = e^{t\hat{\mathcal{L}}}(X)$ denote the time evolution of the relative density $X$. Using (19) and the definition of $\beta_2(\mathcal{L})$ (see (27)) we obtain

$$\frac{d}{dt}\mathrm{Var}_\sigma(X_t) = -2\mathcal{E}_2^{\mathcal{L}}(X_t) \leq -2\beta_2(\mathcal{L})(1 + \mathrm{Var}_\sigma(X_t))\ln(1 + \mathrm{Var}_\sigma(X_t)).$$

Integrating this differential inequality we obtain

$$\ln\left(\frac{\ln(1 + \mathrm{Var}_\sigma(X))}{\ln(1+\epsilon)}\right) \leq \int_0^{t_2(\epsilon)} \frac{1}{(1 + \mathrm{Var}_\sigma(X_t))\ln\left(1 + \mathrm{Var}_\sigma(X_t)\right)}\left[\frac{d}{dt}\mathrm{Var}_\sigma(X_t)\right]dt$$

$$\leq -2\beta_2 t_2(\epsilon).$$

As $1 + \mathrm{Var}_\sigma(X) \leq \|\sigma^{-1}\|_\infty$, the claim follows after rearranging the terms. $\square$

In the remaining part of the section we will discuss a converse to the previous mixing time bounds, i.e. a lower bound on the logarithmic Sobolev constant in terms of a mixing time. This excludes the possibility of a reversible semigroup with both small $\beta_2$ and short mixing time with respect to the $l_2$ distance. For this we generalize [21, Corollary 3.11] to the noncommutative setting.

**Theorem 5.3** (LS inequality from $l_2$ mixing time). *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a primitive, reversible Liouvillian with fixed point $\sigma \in \mathcal{D}_d^+$. Then*

$$\frac{1}{2} \leq \alpha_2(\mathcal{L})\, t_2\left(e^{-1}\right) \leq 2\beta_2(\mathcal{L})\, t_2\left(e^{-1}\right). \tag{45}$$

*Moreover, this inequality is tight.*

*Proof.* We refer to Appendix B for a proof. $\square$

As remarked in [21], even the classical result does not hold anymore if we drop the reversibility assumption. Therefore, this assumption is also needed in the noncommutative setting. By considering a completely depolarizing channel it is also easy to see that no such bound can hold in discrete time.

Theorem 5.3 implies that for reversible Liouvillians $\beta_2$ and $\alpha_2$ cannot differ by a large factor. More specifically we have the following corollary.

**Corollary 5.2.** *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a primitive, reversible Liouvillian with fixed point $\sigma \in \mathcal{D}_d^+$. Then*

$$2\beta_2(\mathcal{L}) \geq \alpha_2(\mathcal{L}) \geq \beta_2(\mathcal{L}) \ln \left( \frac{\ln \left( \|\sigma^{-1}\|_\infty \right)}{\ln(1 + e^{-1})} \right). \tag{46}$$

*Proof.* We showed the first inequality in Theorem 4.3. The second inequality follows by combining (45) and (44). $\qquad\square$

## 5.2 Mixing in Discrete Time

In this section we will obtain mixing time bounds and also entropic inequalities for discrete-time quantum channels $T : \mathcal{M}_d \to \mathcal{M}_d$. We will then use these techniques to derive mixing times for random local channels, which we will define next. These include channels that usually appear in quantum error correction scenarios, such as random Pauli errors on qubits [27, Chapter 10]. They will be based on the following quantity:

**Definition 5.2.** *For a primitive quantum channel $T : \mathcal{M}_d \to \mathcal{M}_d$ with full rank fixed point $\sigma \in \mathcal{D}_d^+$, we define*

$$\beta_D(T) = \beta_2(T^*\hat{T} - id_d). \tag{47}$$

*Here we used $\hat{T} = \Gamma_\sigma^{-1} \circ T \circ \Gamma_\sigma$.*

The definition of $\beta_D(T)$ can be motivated by the following improved data-processing inequality for the 2-sandwiched Rényi divergence.

**Theorem 5.4.** *Let $T : \mathcal{M}_d \to \mathcal{M}_d$ be a primitive quantum channel with full rank fixed point $\sigma \in \mathcal{D}_d^+$. Then for all $\rho \in \mathcal{D}_d$ we have*

$$D_2 \left( T(\rho) \| \sigma \right) \leq (1 - \beta_D(T)) D_2 \left( \rho \| \sigma \right). \tag{48}$$

*Proof.* Let $X = \sigma^{-\frac{1}{2}} \rho \sigma^{-\frac{1}{2}}$ denote the relative density of $\rho$ with respect to $\sigma$. Observe that the 2-Dirichlet form (see Definition 2.3) of the semigroup $\mathcal{L} = T^*\hat{T} - \mathrm{id}_d$ can be written as

$$\mathcal{E}_2^{\mathcal{L}}(X) = \|X\|_{2,\sigma}^2 - \|\hat{T}(X)\|_{2,\sigma}^2.$$

From the definition of $\beta_2(\mathcal{L})$ (see (27)) it follows that

$$\|X\|_{2,\sigma}^2 - \|\hat{T}(X)\|_{2,\sigma}^2 \geq \beta_2 \kappa_2(X),$$

which is equivalent to

$$\ln(\|\hat{T}(X)\|_{2,\sigma}^2) - \ln(\|X\|_{2,\sigma}^2) \leq \ln(1 - \beta_2 \ln(\|X\|_{2,\sigma}^2)).$$

Using the elementary inequality $\ln(1 - \beta_2 \ln(\|X\|_{2,\sigma}^2) \leq -\beta_2 \ln(\|X\|_{2,\sigma}^2)$, that $\ln(\|\hat{T}(X)\|_{2,\sigma}^2) = D_2(T(\rho)\|\sigma)$ and $\ln(\|X\|_{2,\sigma}^2) = D_2(\rho\|\sigma)$ hold, the statement of the theorem follows after rearranging the terms. $\qquad\square$

One should note that, unlike in Theorem 3.2, the constant $\beta_D$ is *not* optimal in (48). As an example take $T(\rho) = \mathrm{tr}[\rho] \frac{\mathbb{1}_d}{d}$ for which $\beta_D(T) = \frac{1 - d^{-1}}{\ln(d)}$, but $D_2 \left( T(\rho) \| \frac{\mathbb{1}_d}{d} \right) = 0$. Also, $\beta_D(T) > 0$ is not a necessary condition for primitivity, as there are primitive quantum

channels that are not strict contractions with respect to $D_2$. To see this, consider the map $T : \mathcal{M}_2 \rightarrow \mathcal{M}_2$ which acts as follows on Pauli operators:

$$T(\mathbb{1}) = \mathbb{1}, \quad T(\sigma_x) = 0, \quad T(\sigma_y) = 0 \text{ and } \quad T(\sigma_z) = \sigma_x.$$

One can check that this is a a primitive quantum channel with $T^2(\rho) = \frac{\mathbb{1}}{2}$ for any state $\rho \in \mathcal{D}_d$. However, $T$ maps the pure state $\frac{1}{2}(\mathbb{1} + \sigma_z)$ to the pure state $\frac{1}{2}(\mathbb{1} + \sigma_x)$, which implies that $D_2$ does not strictly contract under $T$. We can now prove the following bound on the discrete mixing time.

**Theorem 5.5** (Discrete mixing time)**.** *Let $T : \mathcal{M}_d \rightarrow \mathcal{M}_d$ be a primitive quantum channel with full rank fixed point $\sigma \in \mathcal{D}_d^+$ and $\beta_D(T) > 0$. Then*

$$t_1(\epsilon) \leq -\frac{1}{\ln(1 - \beta_D(T))} \ln\left(\frac{2\ln\left(\|\sigma^{-1}\|_\infty\right)}{\epsilon^2}\right).$$

*Proof.* By Theorem 5.4 we have

$$D_2(T^n(\rho)\|\sigma) \leq (1 - \beta_D(T))^n D_2(\rho\|\sigma).$$

for any $\rho \in \mathcal{D}_d$. The claim then follows from (6) and Lemma 2.1. $\qquad\square$

Convergence results for primitive continuous-time semigroups can often be lifted to their tensor powers. In discrete time a similar result holds for the following class of channels:

**Definition 5.3** (Random Local Channels)**.** *For a quantum channel $T : \mathcal{M}_d \rightarrow \mathcal{M}_d$ and probabilities $\mathbf{p} = (p_1, \ldots, p_n)$ with $p_i \geq 0$ and $\sum_i p_i = 1$ we define a* **random local channel** $T_{\mathbf{p}}^{(n)} : \mathcal{M}_d^{\otimes n} \rightarrow \mathcal{M}_d^{\otimes n}$ *by*

$$T_{\mathbf{p}}^{(n)} = \sum_{i=1}^n p_i \ id_d^{\otimes i-1} \otimes T \otimes id_d^{\otimes n-i}. \tag{49}$$

The previous definition can be generalized to the case where not all local channels are identical, i.e. if we have $T_i : \mathcal{M}_d \rightarrow \mathcal{M}_d$ acting on the $i$th system in the expression (49). As long as the local channels are all primitive our results also hold for this more general class of channels. However, for simplicity we will restrict here to the above definition.

**Theorem 5.6.** *Let $T : \mathcal{M}_d \rightarrow \mathcal{M}_d$ be a primitive quantum channel with full rank fixed point $\sigma \in \mathcal{D}_d^+$ such that the Liouvillian $\hat{\mathcal{L}} = T^*\hat{T} - id_d$ fulfills $\beta_2\left(\mathcal{L}^{(n)}\right) \geq q$ for some $q > 0$ and all $n \in \mathbb{N}$. Then for any $n \in \mathbb{N}$ and probabilities $\mathbf{p} = (p_1, \ldots, p_n)$ with $p_i \geq 0$ and $\sum_i p_i = 1$ we have*

$$D_2(T_{\mathbf{p}}^{(n)}(\rho)\|\sigma^{\otimes n}) \leq (1 - qp_{\min}^2)D_2(\rho\|\sigma^{\otimes n}), \tag{50}$$

*for any $\rho \in \mathcal{D}_{d^n}$ and where $p_{\min} = \min p_i$.*

*Proof.* By Theorem 5.4 it is enough to show that $\beta_D(T_{\mathbf{p}}) \geq qp_{\min}^2$.

Observe that the Dirichlet form of $(T_{\mathbf{p}}^{(n)})^*\hat{T}_{\mathbf{p}}^{(n)} - \mathrm{id}_d$ is given by

$$\mathcal{E}_2^{\mathcal{L}}(X) = \sum_{i \neq j} p_i p_j \left\langle X - T_i^*\hat{T}_j(X), X \right\rangle_{\sigma^{\otimes n}} + \sum_i p_i^2 \left\langle X - T_i^*\hat{T}_i(X), X \right\rangle_{\sigma^{\otimes n}}$$

where the map $T_i^* \hat{T}_j$ acts as $T^*$ on the $i$-th system, $\hat{T}$ on the $j$-th and as the identity elsewhere. As $T_i^* \hat{T}_j \leq \mathrm{id}_d$ with respect to $\langle \cdot, \cdot \rangle_{\sigma^{\otimes n}}$ we have

$$\mathcal{E}_2^{\mathcal{L}}(X) \geq \sum_i p_i^2 \left\langle X - T_i^* \hat{T}_i(X), X \right\rangle_{\sigma^{\otimes n}} \geq p_{\min}^2 \mathcal{E}_2^{\mathcal{L}^{(n)}}(X).$$

From the comparison inequality $\mathcal{E}_2^{\mathcal{L}} \geq p_{\min}^2 \mathcal{E}_2^{\mathcal{L}^{(n)}}$ and the assumption $\beta_2\left(\mathcal{L}^{(n)}\right) \geq q$ it then follows that $\beta_D(\Phi) \geq q p_{\min}^2$. $\qquad\square$

As an application we can bound the entropy production and the mixing time in a system of $n$ qubits affected (uniformly) by random Pauli errors. The time evolution of this system is given by the channel $T_n : \mathcal{M}_2^{\otimes n} \to \mathcal{M}_2^{\otimes n}$ given by

$$T_n = \frac{1}{n} \sum_{i=1}^n \mathrm{id}_2^{\otimes i-1} \otimes T \otimes \mathrm{id}_2^{\otimes n-i} \tag{51}$$

with $T(\rho) = \mathrm{tr}(\rho)\frac{\mathbb{1}}{2}$.

**Theorem 5.7.** *For $T_n$ defined as in equation* (51) *we have*

$$D_2\left(T_n(\rho)\|\frac{\mathbb{1}_{2^n}}{2^n}\right) \leq \left(1 - \frac{1}{2n^2}\right) D_2\left(\rho\|\frac{\mathbb{1}_{2^n}}{2^n}\right).$$

*for any $\rho \in \mathcal{D}_{2^n}$.*

*Proof.* From [2] it is known that

$$\alpha_2\left(\mathcal{L}_{\frac{1}{2}}^{(n)}\right) = 1.$$

Now combining Theorem 4.3 and Theorem 5.6 gives

$$\beta_D(T_n) \geq \frac{1}{2n^2} \alpha_2\left(\mathcal{L}_{\frac{1}{2}}^{(n)}\right).$$

$\qquad\square$

**Corollary 5.3.** *Let $T_n$ be defined as in* (51). *Then we have*

$$t_1(\epsilon) \leq -\frac{1}{\ln\left(1 - \frac{1}{2n^2}\right)} \ln\left(\frac{n}{\epsilon^2}\right). \tag{52}$$

*Proof.* This follows directly from the previous theorem and Theorem 5.5. $\qquad\square$

# 6 Strong converse bounds for the classical capacity

When classical information is sent via a quantum channel, the classical capacity is the supremum of transmission rates such that the probability for a decoding error vanishes in the limit of infinite channel uses. In general it is not possible to retrieve the information perfectly when it is sent over a finite number of uses of the channel, and the probability for successful decoding will be smaller than 1. Here we want to derive bounds on this probability for quantum dynamical semigroups. More specifically we are interested in strong converse bounds on the classical capacity. An upper bound on the capacity is

called a strong converse bound if whenever a transmission rate exceeds the bound the probability of successful decoding goes to zero in the limit of infinite channel uses.

We refer to [27, Chapter 12] for the exact definition of the classical capacity and to [28, 29, 30, 4, 31] for more details on strong converses and strong converse bounds.

In [4] the following quantity was used to study strong converses.

**Definition 6.1** (p-information radius)**.** *Let $T : \mathcal{M}_d \to \mathcal{M}_d$ be a quantum channel. The p-**information radius** $T$ is defined as[1]*

$$K_p(T) = \frac{1}{\ln(2)} \min_{\sigma \in \mathcal{D}_d} \max_{\rho \in \mathcal{D}_d} D_p(T(\rho)\|\sigma).$$

We will often refer to a $(m, n, p)$-coding scheme for classical communication using a quantum channel $T$. By this we mean a coding-scheme for the transmission of $m$ classical bits via $n$ uses of the channel $T$ for which the probability of successful decoding is $p$ (see again [27, Chapter 12] for an exact definition). The following theorem shown in [4, Section 6] relates the information radius and the probability of successful decoding.

**Theorem 6.1** (Bound on the success probability in terms of information radius)**.** *Let $T : \mathcal{M}_d \to \mathcal{M}_d$ be a quantum channel, $n \in \mathbb{N}$ and $R \geq 0$. For any $(nR, n, p_{succ})$-coding scheme for classical communication via $T$ we have*

$$p_{succ} \leq 2^{-n\left(\frac{p-1}{p}\right)\left(R - \frac{1}{n} K_p\left(T^{\otimes n}\right)\right)}. \tag{53}$$

We will now apply the methods developed in the last sections to obtain strong converse bounds on the capacity of quantum dynamical semigroups.

**Theorem 6.2.** *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a primitive Liouvillian with full rank fixed point $\sigma \in \mathcal{D}_d^+$ such that for some $p \in (1, \infty)$ there exists $c > 0$ fulfilling $\beta_p(\mathcal{L}^{(n)}) \geq c$ for all $n \in \mathbb{N}$. Then for any $(nR, n, p_{such})$-coding scheme for classical communication via the quantum dynamical semigroup $T_t = e^{t\mathcal{L}}$ we have*

$$p_{succ} \leq 2^{-n\left(\frac{p-1}{p}\right)\left(R - e^{-2ct} \log(\|\sigma^{-1}\|_\infty)\right)}.$$

*Proof.* Using Theorem 3.2 and Lemma 2.1 we have

$$K_p(T_t^{\otimes n}) \leq \frac{1}{\ln(2)} \max_{\rho \in \mathcal{D}_{d^n}} D_p(T_t^{\otimes n}(\rho)\|\sigma) \leq n e^{-2\beta_p\left(\mathcal{L}^{(n)}\right)t} \log(\|\sigma^{-1}\|_\infty).$$

Now Theorem 6.1 together with the assumption $\beta_p(\mathcal{L}^{(n)}) \geq c$ finishes the proof. $\square$

Together with Theorem 4.3 the previous theorem shows that a quantum memory can only reliably store classical information for small times when it is subject to noise described by a quantum dynamical semigroup with "large" logarithmic Sobolev constant, as we will see more explicitly later in Section 7. Moreover, we can use the results from [26] to give a universal lower bound to the decay of the capacity in terms of the spectral gap and the fixed point.

---

[1]The ln(2) factor is due to our different choice of normalization for the divergences.

**Corollary 6.1.** *Let $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ be a primitive Liouvillian with full rank fixed point $\sigma \in \mathcal{D}_d^+$ and spectral gap $\lambda(\mathcal{L})$. Then for any $(nR, n, p_{such})$-coding scheme for classical communication via the quantum dynamical semigroup $T_t = e^{t\mathcal{L}}$ we have*

$$p_{such} \le 2^{-\frac{n}{2}\left(R - e^{-k(\lambda,\sigma)t}\log(\|\sigma^{-1}\|_\infty)\right)}$$

*where $k(\lambda, \sigma) := \frac{\lambda}{2(\ln(d^4\|\sigma^{-1}\|_\infty)+11)}$.*

*Proof.* It was shown in [26, Theorem 9] that $\alpha_2(\mathcal{L}^{(n)}) \ge 2k(\lambda(\mathcal{L}), \sigma)$ for all $n \in \mathbb{N}$. Using Theorem 4.3 we have $\beta_2(\mathcal{L}^{(n)}) \ge k(\lambda(\mathcal{L}), \sigma)$ for all $n \in \mathbb{N}$. Together with Theorem 6.2 this gives the claim. $\qquad\square$

For unital semigroups, i.e. for $\sigma = \frac{\mathbb{1}_d}{d}$, one can improve the bound from the previous theorem slightly using (see [32, Theorem 3.3])

$$k\left(\lambda, \frac{\mathbb{1}_d}{d}\right) = \frac{\lambda(1 - 2d^{-2})}{2(\ln(3)\ln(d^2 - 1) + 2(1 - 2d^{-2})} \tag{54}$$

For $d = 2$ we even have $k(\lambda, \frac{\mathbb{1}_d}{2}) = \frac{\lambda}{2}$ (see [2]).

## 7   Examples of bounds for the classical capacity of Semigroups

We will now apply the estimate on the capacity given by Corollary 6.1 to some examples of semigroups. Here $C(T)$ will denote the classical capacity of a quantum channel $T$.

### 7.1   Depolarizing Channels

In [5] it is shown that for $\mathcal{L}_{\frac{1}{d}}(X) = \operatorname{tr}(X)\frac{\mathbb{1}}{d} - X$ we have

$$C\left(e^{t\mathcal{L}_{\frac{1}{d}}}\right) = \log(d) + \left(e^{-t} + c(t, d)\right)\log\left(e^{-t} + c(t, d)\right) + (d - 1)c(t, d)\log\left(c(t, d)\right) \tag{55}$$

with $c(t, d) = (1 - e^{-t})d^{-1}$. In [30] the strong converse property was established. The semigroup generated by $\mathcal{L}_{\frac{1}{d}}$ is therefore a natural candidate to evaluate the quality of our bounds, as determining its classical capacity can be considered a solved problem. As $\mathcal{L}_{\frac{1}{d}}$ is just the difference of a projection and the identity, it is easy to see that the spectral gap of $\mathcal{L}_{\frac{1}{d}}$ is 1, which gives us the upper bound

$$C\left(e^{t\mathcal{L}_{\frac{1}{d}}}\right) \le \log(d)e^{-\frac{(1-2d^{-2})}{2(\ln(3)\ln(d^2-1)+2(1-2d^{-2})}t} \tag{56}$$

for $d > 2$ and

$$C\left(e^{t\mathcal{L}_{\frac{1}{2}}}\right) \le e^{-\frac{t}{2}} \tag{57}$$
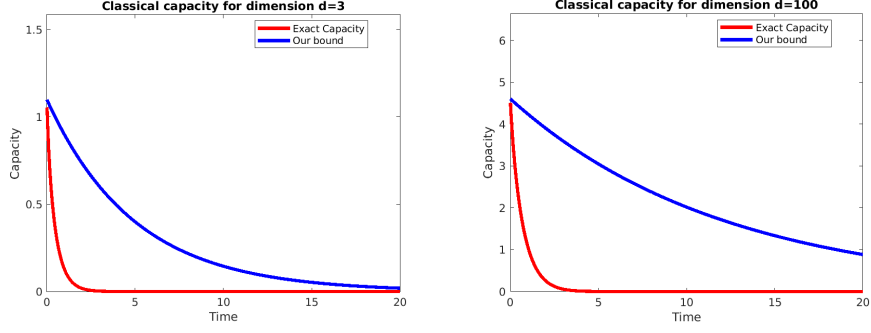
for $d = 2$.

Figure 1: Comparison of the capacity of the depolarizing channel, given in Equation (55), and the bound obtained by our methods, given in Equation (56).

## 7.2 Stabilizer Hamiltonians

Estimates on the spectral gap of Davies generators of stabilizer Hamiltonians were obtained in [19]. In the following we will make the same assumptions as in [19] on the coupling of the system to the bath. That is, we assume that the operators $S^\alpha$ (see (13)) are given by single qubit Pauli operators $\sigma_x, \sigma_y$ or $\sigma_z$. For the transition rates $G^\alpha(\omega)$ we only assume that they satisfy the KMS condition [17], that is, $G^\alpha(-\omega) = G^\alpha(\omega)e^{-\beta\omega}$. This condition implies that the semigroup is reversible. Recall that for Davies generators at inverse temperature $\beta > 0$, which we will denote by $\mathcal{L}_\beta$, the stationary state is always given by the thermal state $\frac{e^{-\beta H}}{\text{tr}(e^{-\beta H})}$.

We will not discuss stabilizer Hamiltonians and groups and their connection to error-correcting codes, but refer to [27, Section 10.5] for more details. Given some stabilizer group $S \subset \mathcal{P}_n$, where $\mathcal{P}_n$ is the group generated by the tensor product of $n$ Pauli matrices, with commutative generators $S = <P_1, \ldots, P_k>$, we define the stabilizer Hamiltonian to be given by

$$H_S = -\sum_{i=1}^{k} P_i.$$

We then have:

**Lemma 7.1.** *Let $H_S$ be the stabilizer Hamiltonian of the stabilizer group $S = <P_1, \ldots, P_k>$ on $n$ qubits. Denote by $\sigma_\beta = \frac{e^{-\beta H_S}}{tr(e^{-\beta H_S})}$ the corresponding thermal state at inverse inverse temperature $\beta > 0$. Then*

$$\|\sigma_\beta^{-1}\|_\infty \leq 2^n e^{2k\beta}.$$

*Proof.* The eigenvalues of each $P_i$ are contained in $\{1, -1\}$, as they are just tensor products of Pauli matrices. From this we have

$$-k\mathbb{1} \leq H_S \leq k\mathbb{1}, \tag{58}$$

as $H_S$ is just the sum of $k$ terms such that $-\mathbb{1} \leq P_i \leq \mathbb{1}$. From (58) it follows that

$$\text{tr}\left(e^{-\beta H_S}\right) \leq 2^n e^{\beta k}, \tag{59}$$

as we have $2^n$ eigenvalues, including multiplicities. Moreover, it also follows that

$$\|e^{\beta H_S}\|_\infty \leq e^{\beta k}. \tag{60}$$

As $\|\sigma_\beta^{-1}\|_\infty = \|e^{\beta H_S}\|_\infty \text{tr}\left(e^{-\beta H_S}\right)$, the claim follows by putting (59) and (60) together. $\quad\square$

In [19, Theorem 15] they show

$$\lambda \geq \frac{h^*}{4\eta^*} e^{-2\beta\bar{\epsilon}}$$

for the spectral gap $\lambda$ of the Davies generators of stabilizer Hamiltonians at inverse temperature $\beta > 0$. Here $\bar{\epsilon}$ is the generalized energy barrier, $h^*$ is the smallest transition rate and $\eta^*$ the longest path in Pauli space. We refer to [19] for the exact definition of these parameters. It is important to stress that in general $\eta^*$ will scale with the number of qubits, so our estimate on the capacity will not be very good as the number of qubits increases.

However, in [19, Theorem 15] they also show the estimate

$$\lambda \geq \frac{h^*}{4} e^{-2\beta\bar{\epsilon}},$$

for the special case in which the generalized energy barrier can be evaluated with canonical paths $\Gamma_1$. We again refer to [19] for the exact definition. For these cases the gap does not scale with the dimension and our estimate is much better. Summing up we obtain:

**Theorem 7.1.** *Let $H_S$ be the stabilizer Hamiltonian of the stabilizer group $S = <P_1, \ldots, P_k>$ on $n$ qubits. Moreover, let $\mathcal{L}_\beta$ be its Davies generator at inverse temperature $\beta > 0$. Then the classical capacity $C(e^{t\mathcal{L}_\beta})$ is bounded by*

$$C(e^{t\mathcal{L}_\beta}) \leq (n + 2\beta k \log(e)) e^{-r(\beta,n,k)t},$$

*with*

$$r(\beta, n, k) = e^{-2\beta\bar{\epsilon}} \frac{h^*}{8\eta^* \left(2k\beta + 5n \ln(2) + 11\right)}$$

*and*

$$r(\beta, n, k) = e^{-2\beta\bar{\epsilon}} \frac{h^*}{8 \left(2k\beta + 5n \ln(2) + 11\right)}$$

*in case the generalized energy barrier can be evaluated with canonical paths $\Gamma_1$. Moreover, this is a bound in the strong converse sense.*

*Proof.* The claim follows immediately after inserting the bounds from Lemma 7.1 and [19, Theorem 15,16] into Corollary 6.1. $\quad\square$

In [19] one can find more explicit bounds for the parameters $\bar{\epsilon}$, $\eta^*$ and $h^*$ for some stabilizer groups. To the best of our knowledge this is the first bound available for the classical capacity of this class of quantum channels. To make the bound in Theorem 7.1 more concrete, we show what we obtain for the $2D$ toric code.

## 7.3 2D Toric Code

Here we consider the 2D toric code as originally introduced in [33], which is a stabilizer code. We consider only square lattices: We take an $N \times N$ lattice with $N^2$ vertical and $(N+1)^2$ horizontal edges; associating a qubit to each edge gives a total of $n = 2N^2 + 2N + 1$ physical qubits. The stabilizer operators are $N(N+1)$ plaquette operators (including the "open" plaquettes along the rough boundary) and $N(N+1)$ vertex operators, all of which are independent. It goes beyond the scope of this article to explain the 2D toric code in detail and we refer to [34, Section 19.4] for a discussion. But from the previous observations we obtain that we have $k = 2N(N+1)$ generators for the stabilizer group of the 2D toric code on $n = 2N^2 + 2N + 1$ qubits. We will make the same assumptions on the the Davies generators at inverse temperature $\beta > 0$ for the toric code as in [19]. These are discussed in the beginning of Subsection 7.2.

In [35] it was proved that the spectral gap for the Davies generators for the 2D toric code at inverse temperature $\beta$ satisfies $\lambda \geq \frac{1}{3}e^{-8\beta}$, a result which was reproved in [19] using different techniques. We therefore obtain:

**Corollary 7.1.** *Let $H$ be the stabilizer Hamiltonian of the 2D toric code on a $N \times N$ lattice and $\mathcal{L}_\beta$ be its Davies generator at inverse temperature $\beta > 0$. Then the classical capacity $C(e^{t\mathcal{L}_\beta})$ is bounded by*

$$C(e^{t\mathcal{L}_\beta}) \leq \left( 2N^2 + 2N + 1 + \log(2)4\beta N(N+1) \right) e^{-r(\beta,L)t}, \tag{61}$$

*with*

$$r(\beta, N) = \frac{e^{-8\beta}}{6\left( (10N^2 + 10N + 5)\ln(2) + 4\beta N(N+1) \right) + 66}.$$

*Moreover, this is a bound in the strong converse sense.*

*Proof.* The claim follows immediately from Lemma 7.1 and the spectral gap estimate of [35] for the toric code. $\square$

From Figure 2 it becomes evident that we cannot retain information in the 2D toric for long times at small inverse temperatures and that we can get nontrivial estimates even for very high dimensions, as the size of the gap does not scale with the size of the lattice. It is conjectured that if the spectral gap of the Davies generators of a Hamiltonian with local, commuting terms satisfies a lower bound which is independent of the size of the lattice, then the logarithmic Sobolev 2 constant also satisfies such a bound [36]. As the Hamiltonian of the 2D toric code is of this form, proving this conjecture would lead to a bound similar to the one in Corollary 7.1, but with a rate $r(\beta, N)$ independent of the size of the lattice. This would of course lead to much better bounds for large lattice sizes.

## 7.4 Truncated harmonic oscillator

Consider the Hamiltonian of a truncated harmonic oscillator

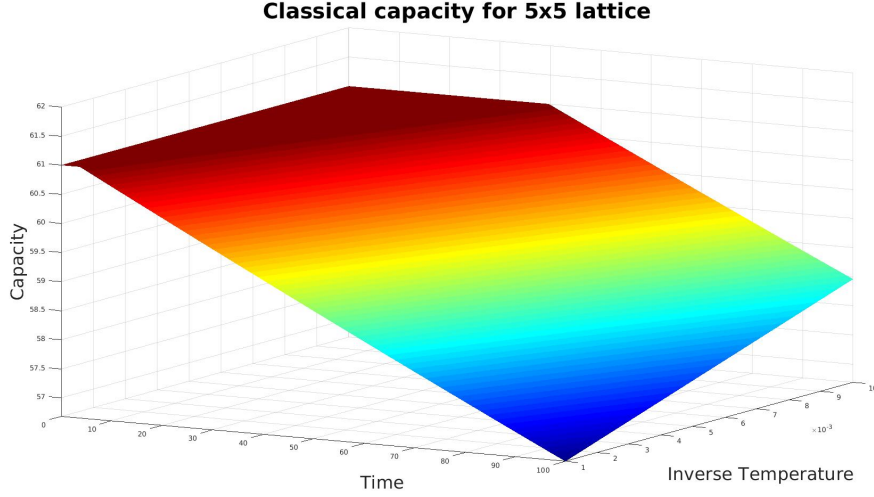$$H = \sum_{n=0}^{d} n|n\rangle\langle n| \in \mathcal{M}_{d+1}.$$

**Classical capacity for 5x5 lattice**

Figure 2: Plot for a $5 \times 5$ lattice of the minimum of the bound in Equation (61) and the trivial bound $C(e^{t\mathcal{L}_\beta}) \leq 2N^2 + 2N + 1 = 61$ as a function of the inverse temperature and time for the Davies generator of the $2D$ toric code.

Suppose that the systems couples to the bath via $S = (a + a^\dagger)$, with

$$a^\dagger = \sum_{n=1}^{d} \sqrt{n} \, |n\rangle \langle n-1| \tag{62}$$

and the transition rate function $G(x) = (1 + e^{-x\beta})^{-1}$. Let $\sigma_\beta = \frac{e^{-\beta H}}{\operatorname{tr}(e^{-\beta H})}$. As the eigenvalues of $e^{-\beta H}$ are just a geometric sequence, we have

$$\|\sigma_\beta^{-1}\|_\infty = \frac{1 - e^{-\beta(d+1)}}{1 - e^{-\beta}} e^{\beta d}. \tag{63}$$

In [23, Section V, Example 1] they show

$$\lambda \geq \frac{1}{2} \min\{((1 + e^{-\beta})d)^{-1}, \left[(G(1)(\sqrt{d-1} - \sqrt{d})^2 + G(-1)(\sqrt{d-2} - \sqrt{d-1})^2)\right]\}, \tag{64}$$

for the spectral gap $\lambda$ of the Davies generator $\mathcal{L}_\beta$ of the truncated harmonic oscillator at inverse temperature $\beta > 0$. We will denote the value of the lower bound in Equation (64) by $\mu(d, \beta)$. As we can compute $\|\sigma_\beta^{-1}\|$ exactly and have a bound on the spectral gap from we can apply Corollary 7.1 to these semigroups.

Note that in this case the bound scales with the dimension. Putting these inequalities together with the bound given in Corollary 6.1 for the capacity, we have for the classical capacity of this semigroup:

$$C(e^{t\mathcal{L}}) \leq \left( \log \left( \frac{1 - e^{-\beta(d+1)}}{(1 - e^{-\beta})} \right) + \beta d \log(e) \right) e^{-r(d,\beta)t}, \tag{65}$$

with

$$r(d, \beta) = \left( 8 \ln (d+1) + 2 \ln \left( \frac{1 - e^{-\beta(d+1)}}{1 - e^{-\beta}} \right) + 2\beta d + 22 \right)^{-1} \mu(d, \beta). \tag{66}$$

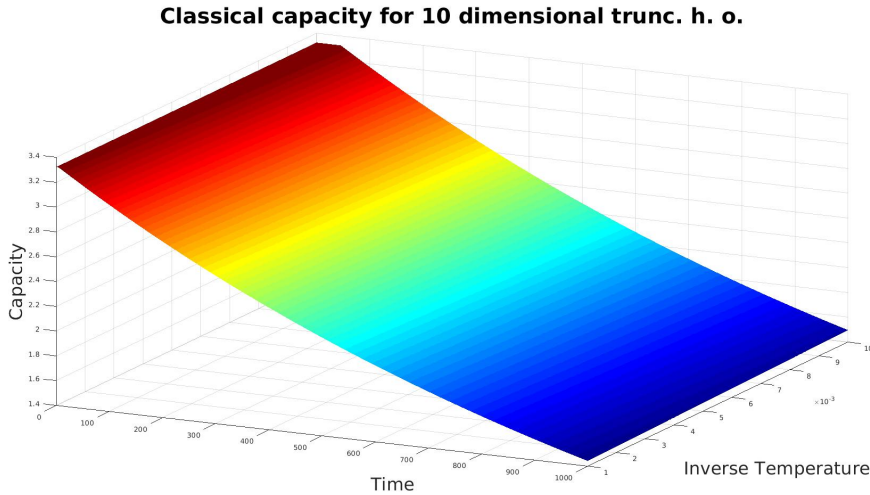**Classical capacity for 10 dimensional trunc. h. o.**

Figure 3: Plot of the minimum of the bound in Equation (65) and the trivial bound $C(e^{t\mathcal{L}_\beta}) \leq \log(10) \simeq$ 3.32 as a function of the inverse temperature and time for the classical capacity of the Davies generator of the truncated harmonic oscillator with $d + 1 = 10$.

In this example we see that, as the estimate available on the gap scales with the dimension, our estimates are not much better than the trivial $\log(d+1)$ for high dimensions unless we are looking at large times.

## 8 Conclusion and open questions

We have introduced a framework similar to logarithmic Sobolev inequalities to study the convergence of a primitive quantum dynamical semigroup towards its fixed point in the distance measure of sandwiched Rényi divergences. These techniques can be used to obtain mixing time bounds and strong converse bounds on the classical capacity of a quantum dynamical semigroup. Moreover, these results show that a logarithmic Sobolev inequality or hypercontractive inequality always implies a mixing time bound without the assumption of $l_p$-regularity (which is still not known to hold for general Liouvillians [2]). Although we have some structural results concerning the constants $\beta_p$, some questions remain open. For logarithmic Sobolev inequalities it is known that $\alpha_2 \leq \alpha_p$ for $p \geq 1$ under the assumption of $l_p$-regularity (see [2]). It would be interesting to investigate if a result of similar flavor also holds for the $\beta_p$. In all examples discussed here, $\beta_2$ and $\alpha_2$ are of the same order and it would be interesting to know if this is always the case. The framework of logarithmic Sobolev inequalities has recently been extended to the nonprimitive case [37]. It should be possible to develop a similar theory for the sandwiched Rényi divergences to get rid of regularity assumptions present in their main results, as we did here for the usual logarithmic Sobolev constants.

We restricted our analysis to the sandwiched Rényi divergences, as they can be expressed in terms of relative densities and noncommutative $l_p$-norms. This allowed us to connect the convergence under the sandwiched divergences to the theory of hypercontractivity and to use tools from interpolation theory which were vital to prove estimates on capacities. There are however other noncommutative generalizations of the Rényi diver-

gences that are known to contract under quantum channels, such as the one discussed in [38, p. 113]. It would be interesting to explore the entropy production and convergence under semigroups for this and other families of divergences in future work.

In a similar vein, it would be interesting to investigate the entropy production or convergence rate for the range $\frac{1}{2} < p < 1$, as the sandwiched Rényi divergences are known to contract under quantum channels for all $p > \frac{1}{2}$ [39]. However, looking closely at the proof of Theorem 3.1, we see that for $p < 1$ the sandwiched Rényi divergence is only differentiable at $t = 0$ if the initial state has full rank. The study of the convergence of these divergences for $p < 1$ therefore requires a different technical approach than that of this work. Finally, it would of course be relevant to obtain bounds on the $\beta_p$ for more examples without relying on the estimate based on the spectral gap, such as Davies generators.

## Acknowledgments

## A  Taylor Expansion of the Dirichlet Form

In order to compute the Taylor expansions of the Dirichlet forms and of the noncommutative $l_p$-norms we define $f_p : \mathbb{R}^2 \to \mathbb{R}$ and $g_p : \mathbb{R}^2 \to \mathbb{R}$ for $p > 1$ as

$$f_p(x,y) = \begin{cases} (p-1)x^{p-2} & \text{if } x = y \\ \frac{x^{p-1}-y^{p-1}}{x-y} & \text{else} \end{cases} \tag{67}$$

and

$$g_p(x,y) = \begin{cases} \frac{p(p-1)}{2}x^{p-2} & \text{if } x = y \\ \frac{(p-1)x^p-px^{p-1}y+y^p}{(x-y)^2} & \text{else.} \end{cases} \tag{68}$$

Note that the following identity holds

$$g_p(x,y) + g_p(y,x) = pf_p(x,y) \tag{69}$$

for any $x, y \in \mathbb{R}$.

**Lemma A.1** (Taylor expansion)**.** *Consider a primitive, reversible Liouvillian $\mathcal{L} : \mathcal{M}_d \to \mathcal{M}_d$ with full rank fixed point $\sigma \in \mathcal{D}_d^+$. Let $X \in \mathcal{M}_d$ be an eigenvector of $\hat{\mathcal{L}} = \Gamma_\sigma^{-1} \circ \mathcal{L} \circ \Gamma_\sigma$ with corresponding eigenvalue $\lambda \in \mathbb{R}$ (i.e. $\hat{\mathcal{L}}(X) = \lambda X$) and $Y_\epsilon = \mathbb{1}_d + \epsilon X$. Then we have*

$$\mathcal{E}_p^{\mathcal{L}}(Y_\epsilon) = \frac{p}{2(p-1)} \left( 2\epsilon^2 \sum_{1 \leq i \leq j \leq d} f_p(s_i, s_j) b_{ij} b_{ji} + O(\epsilon^3) \right). \tag{70}$$

*and*

$$\kappa_p(Y_\epsilon) = \frac{p\epsilon^2}{p-1} \sum_{1 \leq i \leq j \leq d} f_p(s_i, s_j) b_{ij} b_{ji} + O(\epsilon^3). \tag{71}$$

*Where $\sigma^{1/p} = U \, diag\,(s_1, s_2, \ldots, s_d)\, U^\dagger$ and $b_{ij} = (U^\dagger \sigma^{1/2p} X \sigma^{1/2p} U)_{ij}$.*

*Proof.* Using that $X \in \mathcal{M}_d$ is an eigenvector of $\hat{\mathcal{L}}$ a simple computation gives

$$\mathcal{E}_p^{\mathcal{L}}(Y_\epsilon) = \frac{p\epsilon\sigma}{2(p-1)} \mathrm{tr} \left( (A + \epsilon B)^{p-1} B \right)$$

for $A = \sigma^{1/p}$ and $B = \sigma^{1/2p} X \sigma^{1/2p}$. Note that

$$\frac{d^k}{d\epsilon^k} \mathrm{tr} \left( (A + \epsilon B)^{p-1} B \right) \Big|_{\epsilon=0} = \mathrm{tr} \left( D^k F(A)(B, B, \ldots, B) B \right)$$

for the matrix power $F : \mathcal{M}_d \to \mathcal{M}_d$ given by $F(X) = X^{p-1}$. We apply the Daleckii-Krein formula (see [40] and [41, Theorem 2.3.1.] for the version used here) and obtain

$$\frac{d}{d\epsilon} \mathrm{tr} \left( (A + \epsilon B)^{p-1} B \right) \Big|_{\epsilon=0} = \sum_{i=1}^d \sum_{j=1}^d f_p(s_i, s_j) b_{ij} b_{ji}.$$

Using that

$$\mathrm{tr} \left( (A + \epsilon B)^{p-1} B \right) \Big|_{\epsilon=0} = \langle \mathbb{1}_d, X \rangle_\sigma = 0$$

by the orthogonality of eigenvectors, and that $f_p(x, y) = f_p(y, x)$ for any $x, y \in \mathbb{R}$ we obtain (70).

To obtain (71) we write

$$\|Y_\epsilon\|_{p,\sigma}^p = \mathrm{tr} \left( (A + \epsilon B)^p \right)$$

with $A = \sigma^{1/p}$ and $B = \sigma^{1/2p} X \sigma^{1/2p}$ as above. Again it is easy to see that

$$\frac{d^k}{d\epsilon^k} \mathrm{tr} \left( (A + \epsilon B)^p \right) \Big|_{\epsilon=0} = \mathrm{tr} \left( D^k G(A)(B, B, \ldots, B) \right)$$

for the matrix power $G : \mathcal{M}_d \to \mathcal{M}_d$ given by $G(X) = X^p$. Using the Daleckii-Krein formulas we obtain the derivatives

$$\frac{d}{d\epsilon} \mathrm{tr} \left( (A + \epsilon B)^p \right) \Big|_{\epsilon=0} = p \, \langle \mathbb{1}_d, X \rangle_\sigma = 0$$

$$\frac{d^2}{d\epsilon^2} \mathrm{tr} \left( (A + \epsilon B)^p \right) \Big|_{\epsilon=0} = \sum_{i=1}^d \sum_{j=1}^d g_p(\sigma_i, \sigma_j) b_{ij} b_{ji}$$

$$= p \sum_{1 \leq i \leq j \leq d} f_p(\sigma_i, \sigma_j) b_{ij} b_{ji}$$

where we used the identity (69) in the last step. The above shows that

$$\|Y_\epsilon\|_{p,\sigma}^p = 1 + \epsilon^2 p \sum_{1 \le i \le j \le d} f_p(\sigma_i, \sigma_j) b_{ij} b_{ji} + O(\epsilon^3). \tag{72}$$

With the well-known expansion $\ln(1+x) = x - \frac{x^2}{2} + O(x^3)$ we obtain

$$\kappa_p(Y_\epsilon) = \kappa_p(Y_\epsilon) = \frac{p\epsilon^2}{p-1} \sum_{1 \le i \le j \le d} f_p(\lambda_i, \lambda_j) b_{ij} b_{ji} + O(\epsilon^3).$$

which is (71).

$\square$

## B  Interpolation Theorems and Proof of Theorem 5.3

In order to prove Theorem 5.3 we will need the following special case of the Stein-Weiss interpolation theorem [42, Theorem 1.1.1]. This classic result from interpolation spaces has been applied recently to solve problems from quantum information theory, such as in [7, Section III].

**Theorem B.1** (Hadamard Three Line Theorem). *Let $S = \{z \in \mathbb{C} : 0 \le z \le 1\}$ and $F : S \to \mathcal{B}(\mathcal{M}_d)$ be an operator-valued function holomorphic in the interior of $S$ and uniformly bounded and continuous on the boundary. Let $\sigma \in \mathcal{D}_d^+$ and assume $1 \le p \le q \le \infty$. For $0 < \theta < 1$ define $p_0 \le p_\theta \le p_1$ by*

$$\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$$

*Then for $0 \le y \le x \le 1$ we have*

$$\|F(y)\|_{2 \to p_\theta, \sigma} \le \sup_{a,b \in \mathbb{R}} \|F(ia)\|_{2 \to p_0, \sigma}^{1-\theta} \|F(x+ib)\|_{2 \to p_1, \sigma}^{\theta} \tag{73}$$

One important consequence of the Stein-Weiss interpolation theorem is the following interpolation result. We again refer to [42, Theorem 1.1.1] for a proof.

**Theorem B.2** (Riesz-Thorin Interpolation Theorem). *Let $L : \mathcal{M}_d \to \mathcal{M}_d$ be a linear map, $1 \le p_0 \le p_1 \le +\infty$ and $1 \le q_0 \le q_1 \le +\infty$. For $\theta \in [0,1]$ define $p_\theta$ to satisfy*

$$\frac{1}{p_\theta} = \frac{\theta}{p_0} + \frac{1-\theta}{p_1}$$

*and $q_\theta$ analogously. Then for $\sigma \in \mathcal{D}_d^+$ we have:*

$$\|L\|_{p_\theta \to q_\theta, \sigma} \le \|L\|_{p_0 \to q_0, \sigma}^{\theta} \|L\|_{p_1 \to q_1, \sigma}^{1-\theta}$$

With these tools at hand we can finally prove Theorem 5.3:

*Proof.* Define $E : \mathcal{M}_d \to \mathcal{M}_d$ by $E(X) = \operatorname{tr}(\sigma X) \mathbb{1}_d$ and set $\tau = t_2(\epsilon)$ for some $\epsilon > 0$. In the following we use $T_z = e^{\tau z \mathcal{L}}$ for $z \in S = \{z \in \mathbb{C} : 0 \le \operatorname{Re} z \le 1\}$. We will show that for $s \in [0,1]$:

$$\|(T_s - E)(X)\|_{\frac{2}{1-s}, \sigma} \le \epsilon^s \|X\|_{2,\sigma}. \tag{74}$$

The family of operators $T_z - E$ clearly satisfies the assumptions of the Stein-Weiss interpolation theorem. We therefore have

$$\|T_s - E\|_{2 \to \frac{2}{1-s}, \sigma} \le \sup_{a,b \in \mathbb{R}} \|T_{ia} - E\|_{2 \to 2, \sigma}^{1-s} \|T_{1+ib} - E\|_{2 \to \infty, \sigma}^{s}. \tag{75}$$

Observe that by reversibility of $\mathcal{L}$ the map $T_{ia}$ is a unitary operator with respect to $\langle \cdot, \cdot \rangle_\sigma$. We also have $T_{ia} \circ E = E$, as $T_{ia}(\mathbb{1}_d) = \mathbb{1}_d$. This gives

$$\|(T_{ia} - E)(X)\|_{2,\sigma} = \|T_{ia}(X - E(X))\|_{2,\sigma} = \|X - E(X)\|_{2,\sigma} \le \|X\|_{2,\sigma},$$

where the last equality follows from $\|X - \operatorname{tr}(\sigma X)\mathbb{1}_d\|_{2,\sigma} = \min_{c \in \mathbb{R}} \|X - c\mathbb{1}_d\|_{2,\sigma}$. We therefore have

$$\|T_{ia} - E\|_{2 \to 2, \sigma}^{1-s} \le 1. \tag{76}$$

Furthermore, by the unitarity of $T_{ib}$ we can compute

$$\|(T_{1+ib} - E)(X)\|_{\infty, \sigma} = \|T_{ib} \circ (T_1 - E)(X)\|_{\infty, \sigma} \le \|T_1 - E\|_{2 \to \infty, \sigma} \|X\|_{2, \sigma}$$

Using duality of the norms and that both $T_1$ and $E$ are self-adjoint we have

$$\|T_{1+ib} - E\|_{2 \to \infty, \sigma} \le \|T_1 - E\|_{2 \to \infty, \sigma} = \|T_1 - E\|_{1 \to 2, \sigma} = \epsilon \tag{77}$$

using the definition of $\tau$ in the last equality. Inserting (76) and (77) into (75) we get

$$\epsilon^{-s} \|(T_s - E)(X)\|_{\frac{2}{1-s}, \sigma} \le \|X - E(X)\|_{2,\sigma}, \tag{78}$$

as $\|(T_s - E)(X)\|_{\frac{2}{1-s}, \sigma} = \|(T_s - E)(X - E(X))\|_{\frac{2}{1-s}, \sigma}$.

Taking the derivative of (78) with respect to $s$ on both sides at $s = 0$ we get

$$\frac{1}{2\|X - E(X)\|_{2,\sigma}} \left( -2\|X - E(X)\|_{2,\sigma}^2 \ln(\epsilon) + \operatorname{Ent}_{2,\sigma}(|X - E(X)|) - 2\tau \mathcal{E}(X) \right) \le 0. \tag{79}$$

Rearranging the terms in (79) we obtain

$$\operatorname{Ent}_{2,\sigma}(|X - E(X)|) \le 2\tau \mathcal{E}(X) + 2\operatorname{Var}(X) \ln(\epsilon). \tag{80}$$

In [1, Theorem 4.2] the following inequality (known as Rothaus' inequality) was shown

$$\operatorname{Ent}_{2,\sigma}(X) \le \operatorname{Ent}_{2,\sigma}(|X - E(X)|) + 2\operatorname{Var}(X). \tag{81}$$

Combining inequalities (81) with (80) and setting $\epsilon = \frac{1}{e}$ we get $t_2\left(\frac{1}{e}\right) \alpha_2(\mathcal{L}) \ge \frac{1}{2}$ by the definition of the LS constant.

To prove that the inequality is tight, consider the depolarizing Liouvillian $\mathcal{L}_\sigma(X) = \operatorname{tr}(X)\sigma - X$ for some full rank $\sigma \in \mathcal{D}_d^+$. It is easy to see that $\operatorname{Var}_\sigma\left(e^{t\hat{\mathcal{L}}_\sigma}X\right) = e^{-t}\operatorname{Var}_\sigma(X)$ and so $t_2(e^{-1}) = 1 + \ln(\|\sigma^{-1}\|_\infty - 1)$. Restricting to operators commuting with $\sigma$, it follows from [21, Theorem A.1] that

$$\alpha_2(\mathcal{L}_\sigma) \le (1 - 2\|\sigma^{-1}\|_\infty^{-1})\frac{1}{2\ln(\|\sigma^{-1}\|_\infty - 1)}.$$

Thus, for a sequence $\sigma_n \in \mathcal{D}_d^+$ converging to a state that is not full rank we have

$$\lim_{n \to \infty} t_2\left(e^{-1}\right) \alpha_2(\mathcal{L}_{\sigma_n}) = \frac{1}{2}$$

. $\qquad \square$

# References

[1] R. Olkiewicz and B. Zegarlinski. Hypercontractivity in noncommutative lp spaces. *J. Funct. Anal.*, 161(1):246 – 285, 1999. ISSN 0022-1236. DOI: 10.1006/jfan.1998.3342.

[2] M. J. Kastoryano and K. Temme. Quantum logarithmic Sobolev inequalities and rapid mixing. *J. Math. Phys.*, 54(5):052202, May 2013. DOI: 10.1063/1.4804995.

[3] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel. On quantum Rényi entropies: A new generalization and some properties. *J. Math. Phys.*, 54(12): 122203, 2013. DOI: 10.1063/1.4838856.

[4] M. M. Wilde, A. Winter, and D. Yang. Strong Converse for the Classical Capacity of Entanglement-Breaking and Hadamard Channels via a Sandwiched Rényi Relative Entropy. *Commun. Math. Phys.*, 331:593–622, October 2014. DOI: 10.1007/s00220-014-2122-x.

[5] C. King. The capacity of the quantum depolarizing channel. *IEEE Trans. Inf. Theory.*, 49(1):221–229, 2003. DOI: 10.1109/TIT.2002.806153.

[6] H. Umegaki. Conditional expectation in an operator algebra. IV. Entropy and information. *Kodai Math. Sem. Rep.*, 14(2):59–85, 1962. DOI: 10.2996/kmj/1138844604.

[7] S. Beigi. Sandwiched Rényi divergence satisfies data processing inequality. *J. Math. Phys.*, 54(12):122202, December 2013. DOI: 10.1063/1.4838855.

[8] F. Hiai, M. Ohya, and M. Tsukada. Sufficiency, KMS condition and relative entropy in von Neumann algebras. *Pacific J. Math.*, 96(1):99–109. DOI: 10.1142/9789812794208˙0030.

[9] G. L. Gilardoni. On Pinsker's and Vajda's type inequalities for Csiszar's f-divergences. *IEEE Trans. Inf. Theory.*, 56(11):5377–5386, Nov 2010. ISSN 0018-9448. DOI: 10.1109/TIT.2010.2068710.

[10] G. Lindblad. On the generators of quantum dynamical semigroups. *Commun. Math. Phys.*, 48(2):119–130. DOI: 10.1007/BF01608499.

[11] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan. Completely positive dynamical semigroups of N-level systems. *J. Math. Phys.*, 17:821–825, May 1976. DOI: 10.1063/1.522979.

[12] D. Burgarth, G. Chiribella, V. Giovannetti, P. Perinotti, and K. Yuasa. Ergodic and mixing quantum channels in finite dimensions. *New J. Phys.*, 15(7):073045, July 2013. DOI: 10.1088/1367-2630/15/7/073045.

[13] H. Spohn and J. L. Lebowitz. Irreversible thermodynamics for quantum systems weakly coupled to thermal reservoirs. *Adv. Chem. Phys*, 38:109–142. DOI: 10.1002/9780470142578.ch2.

[14] E.B. Davies. *Quantum Theory of Open Systems*. Academic Press, 1976. ISBN 9780122061509.

[15] H.P. Breuer and F. Petruccione. *The Theory of Open Quantum Systems*. OUP Oxford, 2007. ISBN 9780199213900. DOI: 10.1093/acprof:oso/9780199213900.001.0001.

[16] E.B. Davies. Generators of dynamical semigroups. *J. Funct. Anal.*, 34(3):421 – 432, 1979. ISSN 0022-1236. DOI: 10.1016/0022-1236(79)90085-5.

[17] A. Kossakowski, A. Frigerio, V. Gorini, and M. Verri. Quantum detailed balance and KMS condition. *Commun. Math. Phys.*, 57(2):97–110, Jun 1977. ISSN 1432-0916. DOI: 10.1007/BF01625769.

[18] H. Spohn. Entropy production for quantum dynamical semigroups. *J. Math. Phys.*, 19(5):1227–1230, 1978. DOI: 10.1063/1.523789.

[19] K. Temme. Thermalization time bounds for Pauli stabilizer hamiltonians. *Commun. Math. Phys.*, 350(2):603–637, Mar 2017. ISSN 1432-0916. DOI: 10.1007/s00220-016-2746-0.

[20] K. Temme, M. J. Kastoryano, M. B. Ruskai, M. M. Wolf, and F. Verstraete. The $\chi^2$-divergence and mixing times of quantum Markov processes. *J. Math. Phys.*, 51 (12):122201, December 2010. DOI: 10.1063/1.3511335.

[21] P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.*, 6(3):695–750, 08 1996. DOI: 10.1214/aoap/1034968224.

[22] A. Müller-Hermes, D. Stilck França, and M. M. Wolf. Relative entropy convergence for depolarizing channels. *J. Math. Phys.*, 57(2), 2016. DOI: 10.1063/1.4939560.

[23] K. Temme. Lower bounds to the spectral gap of Davies generators. *J. Math. Phys.*, 54(12):122110, December 2013. DOI: 10.1063/1.4850896.

[24] T. S. Cubitt, A. Lucia, S. Michalakis, and D. Perez-Garcia. Stability of local quantum dissipative systems. *Commun. Math. Phys.*, 337(3):1275–1315, 2015. ISSN 1432-0916. DOI: 10.1007/s00220-015-2355-3.

[25] F. G. S. L. Brandão, T. S. Cubitt, A. Lucia, S. Michalakis, and D. Perez-Garcia. Area law for fixed points of rapidly mixing dissipative quantum systems. *J. Math. Phys.*, 56(10):102202, October 2015. DOI: 10.1063/1.4932612.

[26] K. Temme, F. Pastawski, and M. J. Kastoryano. Hypercontractivity of quasi-free quantum semigroups. *J. Phys. A.*, 47(40):405303, 2014. DOI: 10.1088/1751-8113/47/40/405303.

[27] M.A. Nielsen and I.L. Chuang. *Quantum Computation and Quantum Information*. Cambridge Series on Information and the Natural Sciences. Cambridge University Press, 2000. ISBN 9780521635035. DOI: 10.1017/cbo9780511976667.

[28] A. Winter. Coding theorem and strong converse for quantum channels. *IEEE Trans. Inf. Theory.*, 45(7):2481–2485, Nov 1999. ISSN 0018-9448. DOI: 10.1109/18.796385.

[29] T. Ogawa and H. Nagaoka. Strong converse to the quantum channel coding theorem. *IEEE Trans. Inf. Theory.*, 45(7):2486–2489, Nov 1999. ISSN 0018-9448. DOI: 10.1109/18.796386.

[30] R. König and S. Wehner. A Strong Converse for Classical Channel Coding Using Entangled Inputs. *Phys. Rev. Lett*, 103(7):070504, August 2009. DOI: 10.1103/PhysRevLett.103.070504.

[31] M. Tomamichel, M. M. Wilde, and A. Winter. Strong converse rates for quantum communication. In *2015 IEEE International Symposium on Information Theory (ISIT)*, pages 2386–2390. IEEE, 2015. DOI: 10.1109/TIT.2016.2615847.

[32] A. Müller-Hermes, D. Stilck França, and M. M. Wolf. Entropy production of doubly stochastic quantum channels. *J. Math. Phys.*, 57(2):022203, 2016. DOI: 10.1063/1.4941136.

[33] S. B. Bravyi and A. Y. Kitaev. Quantum codes on a lattice with boundary. *arXiv preprint quant-ph/9811052*, 1998.

[34] D. A. Lidar and T. A. Brun. *Quantum error correction*. Cambridge University Press, 2013. DOI: 10.1017/CBO9781139034807.

[35] R. Alicki, M. Fannes, and M. Horodecki. On thermalization in Kitaev's 2D model. *J. Phys. A.*, 42(6):065303, 2009. DOI: 10.1088/1751-8113/42/6/065303.

[36] M. J. Kastoryano and F. G. S. L. Brandão. Quantum Gibbs samplers: The commuting case. *Commun. Math. Phys.*, 344(3):915–957, Jun 2016. ISSN 1432-0916. DOI: 10.1007/s00220-016-2641-8.

[37] I. Bardet. Estimating the decoherence time using non-commutative Functional Inequalities. *ArXiv preprint quant-ph/1710.01039*, October 2017.

[38] M. Ohya and D. Petz. *Quantum Entropy and Its Use.* Theoretical and Mathematical Physics. Springer, 2004. ISBN 9783540208068. DOI: 10.1007/978-3-642-57997-4.

[39] R. L. Frank and E. H. Lieb. Monotonicity of a relative Rényi entropy. *J. Math. Phys.*, 54(12):122201, 2013. DOI: 10.1063/1.4838835.

[40] J. L. Daletskii and S. G. Krein. Integration and differentiation of functions of hermitian operators and applications to the theory of perturbations. *AMS Translations (2)*, 47:1–30, 1965. DOI: 10.1090/trans2/047/01.

[41] F. Hiai. Matrix analysis: matrix monotone functions, matrix means, and majorization. *Interdisciplinary Information Sciences*, 16(2):139–246, 2010. DOI: 10.4036/iis.2010.139.

[42] J. Bergh and J. Lofstrom. *Interpolation spaces: an introduction*, volume 223. Springer, 2012. DOI: 10.1007/978-3-642-66451-9.