Exact solution for the quantum and private capacities of bosonic dephasing channels
CONTENTS

I. Summary of contents 2

II. Preliminaries, notation, and definitions 2
   A. Quantum states and channels 2
   B. Entropies and relative entropies 3
   C. Continuous-variable systems 4
   D. Unassisted capacities of quantum channels 4
   E. Two-way assisted capacities of quantum channels 6
   F. Teleportation simulation 8
   G. Bosonic dephasing channel 9

III. Capacities of bosonic dephasing channels 11
   A. Infinite Toeplitz matrices and theorems of Szegő and Avram–Parter type 11
   B. Proof of main result 12
   C. Extension to multimode channels 18

IV. Examples 20
   A. Wrapped normal distribution 20
   B. Von Mises distribution 22
   C. Wrapped Cauchy distribution 23

References 23
I. SUMMARY OF CONTENTS

The supplementary material provides detailed arguments for the claims made in the main text. We begin with some background material on quantum states, channels, entropies, continuous-variable systems, capacities, and bosonic dephasing channels. After that, we prove our main claims about the capacities of all bosonic dephasing channels. Finally, we discuss some examples of bosonic dephasing channels of physical interest.

II. PRELIMINARIES, NOTATION, AND DEFINITIONS

A. Quantum states and channels

An arbitrary quantum system is mathematically represented by a separable complex Hilbert space $\mathcal{H}$. We start by reviewing a few basic concepts from the theory of operators acting on a Hilbert space $\mathcal{H}$. An operator $X : \mathcal{H} \to \mathcal{H}$ acting on $\mathcal{H}$ is bounded if its operator norm $\|X\|_\infty := \sup_{\|\psi\| = 1} \|X\psi\|$ is finite, i.e., if $\|X\|_\infty < \infty$. The Banach space of bounded operators acting on $\mathcal{H}$ equipped with the norm $\| \cdot \|_\infty$ will be sometimes denoted by $B(\mathcal{H})$. A bounded operator $X \in B(\mathcal{H})$ is positive semi-definite if $\langle \psi | X | \psi \rangle \geq 0$ for all $|\psi\rangle \in \mathcal{H}$.

A bounded operator $T$ such that the series defining $\text{Tr}[T] = \text{Tr} \sqrt{T^*T} = \|T\|_1 < \infty$ converges is said to be of trace class. Trace class operators acting on $\mathcal{H}$ form another Banach space, denoted by $\mathcal{T}(\mathcal{H})$, once they are equipped with the trace norm $\| \cdot \|_1$. The set of positive semi-definite trace class operators forms a cone, denoted here by $\mathcal{T}_+(\mathcal{H})$. Since trace class operators are compact, the spectral theorem applies [1, Theorem VI.16]. This means that every $T \in \mathcal{T}(\mathcal{H})$ can be decomposed as $T = \sum_{k=0}^\infty t_k |e_k\rangle\langle f_k|$, where $\|T\|_1 = \sum_k |t_k| < \infty$, and $\{|e_k\rangle\}, \{|f_k\rangle\}$ are orthonormal bases of $\mathcal{H}$.

Quantum states of a system $A$ are described by density operators, i.e., positive semi-definite trace class operators with trace 1, acting on $\mathcal{H}_A$. The distance between two density operators $\rho, \sigma$ acting on the same Hilbert space can be measured in two different but compatible ways, either with the trace distance $\frac{1}{2} \|\rho - \sigma\|_1$, endowed with a direct operational interpretation via the Helstrom–Holevo theorem for state discrimination [2, 3] or with the fidelity $F(\rho, \sigma) := \|\sqrt{\rho} \sqrt{\sigma}\|_1^2$ [4]. Two fundamental relations known as the Fuchs–van de Graaf inequalities establish the essential equivalence of these two distance measures. They are as follows [5]:

$$1 - \sqrt{F(\rho, \sigma)} \leq \frac{1}{2} \|\rho - \sigma\|_1 \leq \sqrt{1 - F(\rho, \sigma)}.$$  \hfill (1)\

Physical transformations between states of a system $A$ and states of a system $B$ are represented mathematically as quantum channels, i.e., completely positive trace-preserving maps $\Lambda : \mathcal{T}(\mathcal{H}_A) \to \mathcal{T}(\mathcal{H}_B)$ [6–8]. A linear map $\Lambda : \mathcal{T}(\mathcal{H}_A) \to \mathcal{T}(\mathcal{H}_B)$ is

(i) positive if $\Lambda(\mathcal{T}_+(\mathcal{H}_A)) \subseteq \mathcal{T}_+(\mathcal{H}_B)$;

(ii) completely positive, if $\text{id}_N \otimes \Lambda : \mathcal{T}(\mathbb{C}^N \otimes \mathcal{H}_A) \to \mathcal{T}(\mathbb{C}^N \otimes \mathcal{H}_B)$ is a positive map for all $N \in \mathbb{N}$, where $\text{id}_N$ represents the identity channel acting on the space of $N \times N$ complex matrices;

(iii) trace preserving, if $\text{Tr} \Lambda(X) = \text{Tr} X$ holds for all trace class $X$.  

B. Entropies and relative entropies

Let \( p, q \) be two probability density functions defined on the same measurable space \( X \) with measure \( \mu \). For \( \alpha \in (0, 1) \cup (1, \infty) \), define their \( \alpha \)-Rényi divergence by

\[
D_\alpha(p \| q) := \frac{1}{\alpha - 1} \log_2 \int_X d\mu(x) \ p(x)^\alpha q(x)^{1-\alpha}.
\]

(2)

This definition can be extended to \( \alpha \in \{0, 1, \infty\} \) [9, Definition 3] by taking suitable limits. For our purposes, it suffices to consider the Kullback–Leibler divergence [10] obtained by taking the limit \( \alpha \to 1^- \) in (2). It is defined as

\[
D_1(p \| q) = D(p \| q) := \int_X d\mu(x) \ p(x) \log_2 \frac{p(x)}{q(x)}.
\]

(3)

The following technical result is important for this paper.

**Lemma 1** [9, Theorems 3 and 5]. For all fixed \( p, q \), the \( \alpha \)-Rényi divergence is monotonically non-decreasing in \( \alpha \). Moreover, \( \lim_{\alpha \to 1^-} D_\alpha(p \| q) = D(p \| q) \), and if \( D_{\alpha_0}(p \| q) < \infty \) for some \( \alpha_0 > 1 \) (and therefore \( D_\alpha(p \| q) < \infty \) for all \( \alpha \in (0, \alpha_0) \)) then also

\[
\lim_{\alpha \to 1^+} D_\alpha(p \| q) = D(p \| q).
\]

(4)

As a special case of the above formalism, one can define the differential Rényi entropy of a probability density function \( p \) on \( X \) by setting

\[
h_\alpha(p) := \frac{1}{1-\alpha} \log_2 \int_X d\mu(x) \ p(x)^\alpha,
\]

(5)

for all \( \alpha \in (0, 1) \cup (1, \infty) \). For \( \alpha = 1 \) we obtain the standard differential entropy, given by

\[
h(p) := -\int_X d\mu(x) \ p(x) \log_2 p(x),
\]

(6)

whenever the integral is well defined.

We now consider entropies and relative entropies between quantum states. For the sake of simplicity we assume throughout this subsection that all quantum systems are finite dimensional. Indeed, in this paper we shall not consider entropies and relative entropies of infinite-dimensional states.

The most immediate way to extend (2) to the case of two quantum states \( \rho, \sigma \) is to define the Petz–Rényi relative entropy [11]

\[
D_\alpha(\rho \| \sigma) := \frac{1}{\alpha - 1} \log_2 \text{Tr} \rho^\alpha \sigma^{1-\alpha},
\]

(7)

where as usual \( \alpha \in (0, 1) \cup (1, \infty) \), and it is conventional to set \( D_\alpha(\rho \| \sigma) = +\infty \) whenever \( \alpha > 1 \) and \( \text{supp} \rho \nsubseteq \text{supp} \sigma \), where \( \text{supp} \) denotes the support of \( X \), i.e., the orthogonal complement of its kernel. Although (7) is a sensible definition, it is often helpful to consider an alternative quantity. The sandwiched \( \alpha \)-Rényi relative entropy is defined as [12, 13]

\[
\overline{D}_\alpha(\rho \| \sigma) := \frac{2\alpha}{\alpha - 1} \log_2 \left\| \sigma^{\frac{1-\alpha}{\alpha}} \rho^{\frac{1\cdot\alpha}{\alpha}} \right\|_{2\alpha}.
\]

(8)

Here, for \( \beta > 0 \) we define the corresponding Schatten norm of a matrix \( X \) as

\[
\|X\|_{\beta} := \left( \text{Tr} \left[ |X|^\beta \right] \right)^{1/\beta},
\]

(9)
where $|X| := \sqrt{X^\dagger X}$. As before, it is understood that $\tilde{D}_\alpha(\rho\|\sigma) = +\infty$ when $\alpha > 1$ and $\text{supp} \rho \not\subseteq \text{supp} \sigma$. Importantly, when $[\rho, \sigma] = 0$, i.e., $\rho$ and $\sigma$ commute, (7) and (8) coincide, and are equal to the $\alpha$-Rényi divergence between the spectra of $\rho$ and $\sigma$. Namely,

$$[\rho, \sigma] = 0 \implies \tilde{D}_\alpha(\rho\|\sigma) = D_\alpha(\rho\|\sigma).$$

(10)

Taking the limit as $\alpha \to 1$ of either (7) or (8) yields the (Umegaki) relative entropy, given by [14–16]

$$D(\rho\|\sigma) := \lim_{\alpha\to1} \tilde{D}_\alpha(\rho\|\sigma) = \lim_{\alpha\to1} D_\alpha(\rho\|\sigma) = \Tr \left[ \rho (\log_2 \rho - \log_2 \sigma) \right].$$

(11)

The final quantity we need to define is the simplest of all, namely, the (von Neumann) entropy of a density operator $\rho$:

$$S(\rho) := -\Tr \left[ \rho \log_2 \rho \right].$$

(12)

C. Continuous-variable systems

A single-mode continuous-variable system is mathematically modelled by the Hilbert space $\mathcal{H}_1 := L^2(\mathbb{R})$, which comprises all square-integrable complex-valued functions over $\mathbb{R}$. The operators $x$ and $p := -i \frac{d}{dx}$ satisfy the canonical commutation relation $[x, p] = i \mathbb{1}$, where $\mathbb{1}$ denotes the identity operator (in this case, acting on $\mathcal{H}_1$). Introducing the annihilation and creation operators

$$a := \frac{x + ip}{\sqrt{2}}, \quad a^\dagger := \frac{x - ip}{\sqrt{2}},$$

(13)

this can be recast in the form

$$[a, a^\dagger] = \mathbb{1}.$$ 

(14)

Creation operators map the vacuum state $|0\rangle$ to the Fock states

$$|k\rangle := \frac{(a^\dagger)^k}{\sqrt{k!}} |0\rangle.$$ 

(15)

Fock states are eigenvectors of the photon number operator $a^\dagger a$, which satisfies

$$a^\dagger a \ |k\rangle = k \ |k\rangle.$$ 

(16)

D. Unassisted capacities of quantum channels

In this section, we briefly define the quantum and private capacities of a quantum channel. We begin with the quantum capacity. An $(|M|, \epsilon)$ code for quantum communication over the channel $\mathcal{N}_{A\to B}$ consists of an encoding channel $\mathcal{E}_{M\to A}$ and a decoding channel $\mathcal{D}_{B\to M}$ such that the channel fidelity of the coding scheme and the identity channel $\text{id}_M$ is not smaller than $1 - \epsilon$:

$$F(\text{id}_M, \mathcal{D}_{B\to M} \circ \mathcal{N}_{A\to B} \circ \mathcal{E}_{M\to A}) \geq 1 - \epsilon,$$

(17)

where the channel fidelity of channels $\mathcal{N}_1$ and $\mathcal{N}_2$ is defined as [17]

$$F(\mathcal{N}_1, \mathcal{N}_2) := \inf_\rho F(\text{id} \otimes \mathcal{N}_1)(\rho), (\text{id} \otimes \mathcal{N}_2)(\rho),$$

(18)
FIG. 1: A depiction of a quantum communication protocol that uses the channel $n$ times.

with the optimization over every bipartite state $\rho$ and the reference system allowed to be arbitrarily large. See Figure 1 for a depiction of a quantum communication protocol that uses the channel $n$ times.

The one-shot quantum capacity $Q_{\varepsilon}(N_{A\rightarrow B})$ of the channel $N_{A\rightarrow B}$ is defined as

$$Q_{\varepsilon}(N) := \sup_{E,D} \{ \log_2 |M| : \exists (|M|, \varepsilon) \text{ quantum communication protocol for } N_{A\rightarrow B} \},$$

(19)

where the optimization is over every encoding channel $E$ and decoding channel $D$. The (asymptotic) quantum capacity of $N_{A\rightarrow B}$ is then defined as

$$Q(N) := \inf_{\varepsilon \in (0,1)} \liminf_{n \to \infty} \frac{1}{n} Q_{\varepsilon}(N^\otimes n),$$

(20)

where $N^\otimes n$ denotes $n$ copies of the channel $N$ used in parallel. The strong converse quantum capacity of $N_{A\rightarrow B}$ is defined as

$$Q^\dagger(N) := \sup_{\varepsilon \in (0,1)} \limsup_{n \to \infty} \frac{1}{n} Q_{\varepsilon}(N^\otimes n).$$

(21)

The above way of defining quantum capacity is standard, by now, in several references on quantum information theory [18], [19, Section VIII], following the same approach for defining various other capacities in classical and quantum information theory [20, Eqs. (1.6)–(1.7)], [21, Section V-A], [22, Eq. (1)], [23, Eq. (10)]. There are several different ways of defining quantum capacity (see also [24]), but it is known that they lead to the same quantity in the asymptotic limit [25].

It is a classic result of quantum information theory that the quantum capacity is equal to the regularized coherent information of the channel [26–31]:

$$Q(N) = \lim_{n \to \infty} \frac{1}{n} Q^{(1)}(N^\otimes n) = \sup_{n \in \mathbb{N}} \frac{1}{n} Q^{(1)}(N^\otimes n),$$

$$Q^{(1)}(N) := \sup_{\langle \Psi \rangle_{AA'}} I_{\text{coh}}(A)B_{\nu},$$

(22)
where

\[ \gamma_{AB} := (\text{id}_A \otimes \mathcal{N}_{A' \rightarrow B})(\Psi_{AA'}), \]
\[ I_{\text{coh}}(A|B)_\rho := S(\rho_B) - S(\rho_{AB}). \]  

(23)

This gives us a method for evaluating the quantum capacity of particular channels of interest, including the bosonic dephasing channels.

Let us now recall basic definitions related to private capacity. Let \( \mathcal{U}_{A \rightarrow BE}^N \) be an isometric channel extending the channel \( \mathcal{N}_{A \rightarrow B} \) [6]. An \((|M|, \varepsilon)\) code for private communication over the channel \( \mathcal{N}_{A \rightarrow B} \) consists of a set \( \{\rho_{A'}^m\}_m \) of encoding states and a decoder, specified as a positive operator-valued measure (POVM) \( \{\Lambda_{B}^m\}_m \). It achieves an error \( \varepsilon \) if there exists a state \( \sigma_E \) of the environment, such that the following inequality holds for every message \( m \):

\[ F \left( \sum_{m'} |m'\rangle\langle m'| \otimes \text{Tr}_B[\Lambda_{B}^{m'} \mathcal{U}_{A \rightarrow BE}^N(\rho_{A'}^m)], |m\rangle\langle m| \otimes \sigma_E \right) \geq 1 - \varepsilon. \]  

(24)

The one-shot private capacity \( P_\varepsilon(\mathcal{N}_{A \rightarrow B}) \) of the channel \( \mathcal{N}_{A \rightarrow B} \) is defined as

\[ P_\varepsilon(N) := \sup_{\{\rho_{A'}^m\}_m, \{\Lambda_{B}^m\}_m} \{ \log_2 |M| : \exists (|M|, \varepsilon) \text{ private communication protocol for } \mathcal{N}_{A \rightarrow B} \}, \]  

(25)

where the optimization is over every set \( \{\rho_{A'}^m\}_m \) of encoding states and decoding POVM \( \{\Lambda_{B}^m\}_m \). The (asymptotic) private capacity of \( \mathcal{N}_{A \rightarrow B} \) is then defined as

\[ P(N) := \inf_{\varepsilon \in (0, 1)} \liminf_{n \to \infty} \frac{1}{n} P_\varepsilon(\mathcal{N}^\otimes n). \]  

(26)

The strong converse private capacity of \( \mathcal{N}_{A \rightarrow B} \) is defined as

\[ P^\dagger(N) := \sup_{\varepsilon \in (0, 1)} \limsup_{n \to \infty} \frac{1}{n} P_\varepsilon(\mathcal{N}^\otimes n). \]  

(27)

The following inequalities are direct consequences of the definitions:

\[ Q(N) \leq Q^\dagger(N), \]
\[ P(N) \leq P^\dagger(N). \]  

(28)

(29)

E. Two-way assisted capacities of quantum channels

In this section, we define the quantum and private capacities when the channel of interest is assisted by local operations and classical communication (LOCC). We begin with the LOCC-assisted quantum capacity. An \((n, |M|, \varepsilon)\) protocol \( \mathcal{P} \) for LOCC-assisted quantum communication consists of a separable state \( \sigma'_{A_1'B_1'} \) (which is understood to be separable with respect to the bipartition \( A_1'A_1 : B_1'B_1 \)), the set \( \{\mathcal{L}_{A_1'B_1 \rightarrow A_1'B_2'}^{(i-1)}\}_{i=2}^n \) of LOCC channels, and the LOCC channel \( \mathcal{L}_{A_n'B_n \rightarrow M_AM_B}^{(n)} \). (See [32] for the definition of an LOCC channel.) We can also imagine that the state \( \sigma'_{A_1'B_1'} \) is produced by an LOCC
FIG. 2: An LOCC-assisted protocol that involves \( n \) uses of the quantum channel \( N \). Its action is described formally in Eq. (30).

preprocessing channel \( L^{(0)}_{A_0B_0\rightarrow A_1'B_1} \) with the \( A_0'B_0 \) system initialised in a product state. The final state of the protocol is

\[
\eta_{MA'MB} := \left( L^{(n)}_{A_nB_n\rightarrow MA'MB} \circ N_{A_n\rightarrow B_n} \circ L^{(n-1)}_{A_{n-1}B_{n-1}A_{n-1}'B_{n-1}'} \circ \cdots \circ L^{(1)}_{A_1B_1\rightarrow A_1'A_2B_1'} \circ N_{A_1\rightarrow B_1}\right)(\sigma_{A_1'A_1'B_1'}),
\]

satisfying

\[
F(\eta_{MA'MB}, \Phi_{MA'MB}) \geq 1 - \varepsilon,
\]

where \( \Phi_{MA'MB} \) is a maximally entangled state of Schmidt rank \( |M| \). Such a protocol is depicted in Figure 2. We note here that it suffices for such a protocol to generate the maximally entangled state \( \Phi_{MA'MB} \) because entanglement and quantum communication are equivalent communication resources when classical communication is freely available, due to the teleportation protocol [33].

The \( n \)-shot LOCC-assisted quantum capacity of the channel \( N_{A\rightarrow B} \) is defined as

\[
Q_{\rightarrow, n, \varepsilon}(N) := \sup_{\mathcal{P}} \left\{ \frac{1}{n} \log_2 |M| : \exists (n, |M|, \varepsilon) \text{ LOCC-assisted q. comm. protocol } \mathcal{P} \text{ for } N_{A\rightarrow B} \right\},
\]

where the optimization is over every LOCC-assisted quantum communication protocol \( \mathcal{P} \). The (asymptotic) LOCC-assisted quantum capacity of \( N_{A\rightarrow B} \) is then defined as

\[
Q_{\rightarrow}(N) := \inf_{\varepsilon \in (0, 1)} \lim_{n \to \infty} Q_{\rightarrow, n, \varepsilon}(N).
\]

The strong converse LOCC-assisted quantum capacity of \( N_{A\rightarrow B} \) is defined as

\[
Q_{\rightarrow}^*(N) := \sup_{\varepsilon \in (0, 1)} \lim_{n \to \infty} Q_{\rightarrow, n, \varepsilon}(N).
\]

An \((n, |M|, \varepsilon)\) protocol \( \mathcal{K} \) for secret key agreement over a quantum channel is defined essentially the same as an LOCC-assisted protocol for quantum communication, except that the target final state of the protocol is more general. That is, the final step of the protocol is an LOCC channel \( L^{(n)}_{A_nB_n\rightarrow MA'BSA'SB} \), where \( S_A \) and \( S_B \) are extra systems of the sender Alice and the receiver Bob. Let us then denote the final state of the protocol by \( \eta_{MA'MBSA'SB} \). Such a protocol satisfies

\[
F(\eta_{MA'MBSA'SB}, \gamma_{MA'MBSA'SB}) \geq 1 - \varepsilon,
\]

where \( \gamma_{MA'MBSA'SB} \) is a private state of dimension \( |M| \) \([34, 35]\), having the form

\[
\gamma_{MA'MBSA'SB} := U_{MA'MBSA'SB} (\Phi_{MA'MB} \otimes \theta_{S_A'S_B}) U_{MA'MBSA'SB}^\dagger.
\]
In the above, $U_{MAMB}^{SASA}$ is a twisting unitary of the form

$$U_{MAMB}^{SASA} = \sum_{i,j} |i\rangle|i\rangle_{MA} \otimes |j\rangle|j\rangle_{MB} \otimes U_{SASA}^{ij},$$

with each $U_{SASA}^{ij}$ a unitary. Also, $\Phi_{MAMB}^{SASA}$ is a maximally entangled state of Schmidt rank $|M|$ and $\theta_{SASA}$ is an arbitrary state. The fact that such a protocol is equivalent to the more familiar notion of secret key agreement, involving three parties generating a tripartite secret key state of the form

$$\frac{1}{|M|} \sum_{m=0}^{|M|-1} |m\rangle_M |m\rangle_M \otimes \sigma_E,$$

is the main contribution of [34, 35] (see [18] for another presentation).

The $n$-shot secret-key-agreement capacity of the channel $\mathcal{N}_{A\rightarrow B}$ is defined as

$$P_{\leftrightarrow,n,e}(N) := \sup_{\mathcal{K}} \left\{ \frac{1}{n} \log |M| : \exists(n, |M|, e) \text{ secret-key-agreement protocol } \mathcal{K} \text{ for } \mathcal{N}_{A\rightarrow B} \right\},$$

where the optimization is over every secret key agreement protocol $\mathcal{K}$. The (asymptotic) secret key agreement capacity of $\mathcal{N}_{A\rightarrow B}$ is then defined as

$$P_{\leftrightarrow}(N) := \inf_{e \in (0,1)} \liminf_{n \rightarrow \infty} P_{\leftrightarrow,n,e}(N).$$

The strong converse secret key agreement capacity of $\mathcal{N}_{A\rightarrow B}$ is defined as

$$P_{\leftrightarrow}^{\dagger}(N) := \sup_{e \in (0,1)} \limsup_{n \rightarrow \infty} P_{\leftrightarrow,n,e}(N).$$

The following inequalities are direct consequences of the definitions:

$$Q_{\leftrightarrow}(N) \leq Q_{\leftrightarrow}^{\dagger}(N) \leq P_{\leftrightarrow}^{\dagger}(N).$$

Due to the fact that a more general target state is allowed in secret key agreement, the following inequalities hold

$$Q_{\leftrightarrow}(N) \leq P_{\leftrightarrow}(N),$$

$$Q_{\leftrightarrow}^{\dagger}(N) \leq P_{\leftrightarrow}^{\dagger}(N).$$

Finally, due to the fact that classical communication can only enhance capacities, the following inequalities hold:

$$Q(N) \leq Q_{\leftrightarrow}(N),$$

$$Q_{\leftrightarrow}^{\dagger}(N) \leq Q_{\leftrightarrow}^{\dagger}(N) \leq P_{\leftrightarrow}(N),$$

$$P(N) \leq P_{\leftrightarrow}(N),$$

Thus, to establish the collapse of all of the capacities discussed in this section and the previous one, for the case of bosonic dephasing channels, it suffices to prove the lower bound on $Q(N)$ and the upper bound on $P_{\leftrightarrow}^{\dagger}(N)$.

F. Teleportation simulation

The $d$-dimensional quantum teleportation protocol [33] takes as input a $d$-dimensional quantum state $\rho_{A'}$ of a system $A'$, a $d$-dimensional maximally entangled state

$$\Phi_d^{AB} := |\Phi_d\rangle \langle \Phi_d|_AB, \quad |\Phi_d\rangle_{AB} := \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} |k\rangle_A |k\rangle_B,$$

Thus, to establish the collapse of all of the capacities discussed in this section and the previous one, for the case of bosonic dephasing channels, it suffices to prove the lower bound on $Q(N)$ and the upper bound on $P_{\leftrightarrow}^{\dagger}(N)$.
and by using only local operations and one-way classical communication from Alice to Bob reproduces the exact same state $\rho$ on the system $B$. To define it rigorously, for $x, z \in \{0, \ldots, d - 1\}$ let us introduce the unitary matrices

$$
X(x) := \sum_{k=0}^{d-1} |k \oplus x\rangle\langle k|, \quad Z(z) := \sum_{k=0}^{d-1} e^{\frac{2\pi i zk}{d}} |k\rangle\langle k|, \quad U(x, z) := X(x)Z(z),
$$

(45)

where $\oplus$ denotes sum modulo $d$. Then the teleportation channel $T_{A'\rightarrow B}^{(d)}$ is given by

$$
T_{A'\rightarrow B}^{(d)}(X_{A'AB}) := \sum_{x, z=0}^{d-1} U(x, z)_{\tilde{B}} \text{Tr}_{AA'} \left[ X_{A'AB} U(x, z)_{\tilde{A}} \Phi_d^{A'A'} U(x, z)_{\tilde{A}}^\dagger \right] U(x, z)^\dagger_{\tilde{B}}.
$$

(46)

The effectiveness of the standard quantum teleportation protocol is expressed by the identity

$$
T_{A'\rightarrow B}^{(d)}(\rho_{A'} \otimes \Phi_d^{AB}) = \rho_B,
$$

(47)

meaning that the same operator $\rho$ is written in the registers $A'$ and $B$ on the left-hand and on the right-hand side, respectively.

Some channels can be simulated by the action of the standard teleportation protocol on their Choi states [36], in the sense that

$$
T_{A'\rightarrow B}^{(d)}(\rho_{A'} \otimes \Phi_d^{AB}) = N(\rho_{A'}),
$$

(48)

where $\Phi_d^{AB}$ is the Choi state of the channel $N$. For example, this is the case for all Pauli channels. More generally, other channels can be simulated approximately by the action of the teleportation protocol on their Choi states. This concept was introduced in [36] for the explicit purpose of obtaining upper bounds on the LOCC-assisted quantum capacity of a channel in terms of an entanglement measure evaluated on the Choi state. The idea was rediscovered in [37] for the same purpose, and more recently the same idea was used to bound the secret-key-agreement capacity [38] and the strong converse secret-key-agreement capacity [39]. Here we make use of this concept in order to obtain upper bounds on the strong converse secret key agreement capacity of all bosonic dephasing channels. As discussed earlier, it suffices to consider establishing an upper bound on this latter capacity because it is the largest among all the capacities that we consider in this paper.

### G. Bosonic dephasing channel

**Definition 2.** Let $p$ be a probability density function on the interval $[-\pi, \pi]$. The associated **bosonic dephasing channel** is the quantum channel $N_p : T(\mathcal{H}_1) \rightarrow T(\mathcal{H}_1)$ acting on a single-mode system and given by

$$
N_p(\rho) := \int_{-\pi}^{\pi} d\phi \ p(\phi) \ e^{-ia^\dagger a} \phi \rho \ e^{ia^\dagger a} \phi,
$$

(49)

where $a^\dagger a$ is the photon number operator.

The action of the bosonic dephasing channel can be easily described by representing the input operator in the Fock basis. By means of this representation the Hilbert space of a single-mode system, $\mathcal{H}_1$, becomes equivalent to that of square-summable complex-valued sequences, denoted $\ell^2(\mathbb{N})$. Operators on $\mathcal{H}_1$ are
represented by infinite matrices, i.e., operators $S : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$. Given two such operators $S, T$, which we formally write $S = \sum_{h,k} S_{hk} |h\rangle\langle k|$ and $T = \sum_{h,k} T_{hk} |h\rangle\langle k|$, their Hadamard product is defined by

$$S \circ T := \sum_{h,k} S_{hk} T_{hk} |h\rangle\langle k|.$$  \hfill (50)

One of the fundamental facts concerning the Hadamard product is the Schur product theorem [40, Theorem 7.5.3]: it states that if $S \geq 0$ and $T \geq 0$ are positive semi-definite, then also $S \circ T \geq 0$ is such. The theorem is usually stated for matrices, but is immediately generalisable to the operator case as a consequence of the remark below.

**Remark 3.** Let $T : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ be an infinite matrix. Then $T \geq 0$ if and only if $T^{(d)} \geq 0$ for all $d \in \mathbb{N}_+$, where $T^{(d)}$ is the $d \times d$ top left corner of $T$. This follows from the fact that the linear span of the basis vectors $|k\rangle$, $k \in \mathbb{N}$, is dense in $\ell^2(\mathbb{N})$.

Given an infinite matrix $T$ which represents a bounded operator $T : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$, we can define the associated Hadamard channel as

$$\mathcal{L}_T : T(\ell^2(\mathbb{N})) \longrightarrow T(\ell^2(\mathbb{N})) \quad S \quad \mapsto S \circ T.$$ \hfill (51)

When restricted to finite-dimensional systems, Hadamard channels are examples of so-called ‘partially coherent direct sum channels’ [41]. The following is then easily established.

**Lemma 4.** Let $T : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ be a bounded operator represented by an infinite matrix. Then the Hadamard channel $\mathcal{L}_T$ defined by (51) is a completely positive and trace preserving map, i.e., a quantum channel, if and only if

(i) $T \geq 0$ as an operator; and

(ii) $T_{kk} = 1$ for all $k \in \mathbb{N}$.

**Proof.** The two conditions are clearly necessary. In fact, if $T_{kk} \neq 1$ for some $k \in \mathbb{N}_+$, then $\text{Tr}[T \circ |k\rangle\langle k|] = T_{kk} \neq 1 = |k\rangle\langle k|$; i.e., $\mathcal{L}_T$ is not trace preserving. Also, if $T \not\geq 0$ then by Remark 3 there exists $d \in \mathbb{N}_+$ and some $|\psi\rangle \in \mathbb{C}^d$ such that $\langle \psi | T^{(d)} | \psi \rangle < 0$. Rewriting $\langle \psi | T^{(d)} | \psi \rangle = \sum_{h,k=0}^{d-1} \psi_h^* \psi_k T_{hk} = d \langle \psi \rangle |\psi\rangle (T \circ |\psi\rangle |\psi\rangle)$, where $\psi := |\psi\rangle \langle \psi|$ and $|+\rangle := \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} |k\rangle$, shows that in this case $\mathcal{L}_T$ would not even be positive, let alone completely positive.

Vice versa, conditions (i)–(ii) are sufficient. In fact, on the one hand, by (ii), for an arbitrary $X$, we have that $\text{Tr} \, \mathcal{L}_T(X) = \sum_{k} T_{kk} X_{kk} = \text{Tr} X$, i.e., $\mathcal{L}_T$ is trace preserving. On the other hand, if $T \geq 0$ then for all $d \in \mathbb{N}_+$ and for all positive semi-definite bipartite operators $X \geq 0$ acting on $\mathbb{C}^d \otimes \ell^2(\mathbb{N})$ we have that $(I \otimes \mathcal{L}_T) (X) = d \langle + | (I \otimes T) | + \rangle X \geq 0$, where $|+\rangle$ is defined above, and the last inequality follows by the Schur product theorem. Since $d$ is arbitrary, this proves that $\mathcal{L}_T$ is completely positive. \hfill \Box

The theory of Hadamard channels we just sketched out is relevant here due to the following simple observation.

**Lemma 5.** When both the input and the output density operators are represented in the Fock basis, the bosonic dephasing channel $\mathcal{N}_p$ acts as the Hadamard channel

$$\mathcal{N}_p(\rho) = \rho \circ T_p,$$ \hfill (52)

$$(T_p)_{hk} := \int_{-\pi}^{\pi} d\phi \, p(\phi) \, e^{-i\phi (h-k)}.$$ \hfill (53)

**Proof.** Due to (16), we have that $\mathcal{N}_p(\rho) = \sum_{h,k} \rho_{hk} \int_{-\pi}^{\pi} d\phi \, p(\phi) \, e^{-i\phi (h-k)} |h\rangle\langle k| = \rho \circ T_p$. \hfill \Box
### III. CAPACITIES OF BOSONIC DEPHASING CHANNELS

#### A. Infinite Toeplitz matrices and theorems of Szegő and Avram–Parter type

Observe that the expression for \((T_p)_{hk}\) only depends on the difference \(h - k\). Matrices with this property are named after the mathematician Otto Toeplitz. Formally, a \textit{Toeplitz matrix} of size \(d \in \mathbb{N}_+\) is a matrix of the form

\[
T = \begin{pmatrix}
a_0 & a_{-1} & a_{-2} & \ldots & a_{-d+1} \\
a_1 & a_0 & a_{-1} & \ldots & a_{-d+2} \\
a_2 & a_1 & a_0 & \ldots & a_{-d+3} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
a_{d-1} & a_{d-2} & a_{d-3} & \ldots & a_0
\end{pmatrix},
\]  

where \(a_0, \ldots, a_{d-1} \in \mathbb{C}\). Alternatively, it can be defined to have entries

\[
T_{hk} = a_{h-k}.
\]  

This definition can be formally extended to the case of \textit{infinite Toeplitz matrices}, simply by letting \(h, k \in \mathbb{N}\) run over all non-negative integers. Note that the top left corners \(T^{(d)}\) of an infinite Toeplitz matrix are Toeplitz matrices themselves.

In applications one often encounters the case in which the numbers \(a_k\) are the Fourier coefficients of an \textit{absolutely integrable function} \(a : [-\pi, \pi] \rightarrow \mathbb{C}\), i.e.

\[
a_k = \int_{-\pi}^{\pi} \frac{d\phi}{2\pi} a(\phi) e^{-ik\phi}. \tag{56}
\]

In this paper, we will consider mainly non-negative functions \(a : [-\pi, \pi] \rightarrow \mathbb{R}_+\).

A result due to Szegő [42, 43] states that the spectrum of the \(d \times d\) top left corners \(T^{(d)}\) of an infinite Toeplitz matrix converges to the generating function \(a : [-\pi, \pi] \rightarrow \mathbb{R}\) (for now assumed to be real-valued), in the sense that

\[
\lim_{d \to \infty} \frac{1}{d} \text{Tr} F(T^{(d)}) = \lim_{d \to \infty} \frac{1}{d} \sum_{j=1}^{d} F(\lambda_j(T^{(d)})) = \int_{-\pi}^{\pi} \frac{d\phi}{2\pi} F(a(\phi)) \tag{57}
\]

whenever \(a\) and \(F : \mathbb{R} \rightarrow \mathbb{R}\) are sufficiently well behaved. Here, \(\lambda_j(T^{(d)})\) denotes the \(j\)th eigenvalue of the matrix \(T^{(d)}\). The scope and extension of Szegő’s result has been expanded over the years by relaxing the conditions to be imposed on \(a\) and \(F\) so that (57) holds. At the same time, an analogous class of results, initially conceived by Parter [44] and Avram [45], has been developed to deal with the case of complex-valued generating functions \(a : [-\pi, \pi] \rightarrow \mathbb{C}\). Results of the Avram–Parter type generalize (57) by stating that

\[
\lim_{d \to \infty} \frac{1}{d} \text{Tr} F(|T^{(d)}|) = \lim_{d \to \infty} \frac{1}{d} \sum_{j=1}^{d} F(s_j(T^{(d)})) = \int_{-\pi}^{\pi} \frac{d\phi}{2\pi} F(|a(\phi)|), \tag{58}
\]

where \(s_j(T^{(d)})\) is now the \(j\)th \textit{singular value} of the matrix \(T^{(d)}\). Both Szegő’s and Avram–Parter’s result have been generalized in successive steps, by Zamarashkin and Tyrtyshnikov [46], Tilli [47], Serra-Capizzano [48], Böttcher, Grudsky, and Maksimenko [49], and others. For a detailed account of these developments, we refer the reader to the textbooks [50–52] and especially to the lecture notes by Grudsky [53]. Here we will just need the following lemma, extracted from the work of Serra-Capizzano.
**Lemma 6** (Serra-Capizzano [48]). If \( \alpha : [-\pi, \pi] \to \mathbb{R}_+ \) is such that
\[
\int_{-\pi}^{+\pi} \frac{d\phi}{2\pi} a(\phi)^{\alpha} < \infty
\]
for some \( \alpha \geq 1 \), and moreover \( F : \mathbb{R}_+ \to \mathbb{R} \) is continuous and satisfies
\[
F(x) = O(x^{\alpha}) \quad (x \to \infty),
\]
then (57) holds.

**Proof.** The original result by Serra-Capizzano [48, Theorem 2] states that the identity (58) involving singular values holds. However, under the stronger hypotheses that we are making here it can be seen that (57) and (58) are actually equivalent. Since \( a \) takes on values in \( \mathbb{R}_+ \), we only need to check that for each \( d \) the singular values and the eigenvalues of \( T^{(d)} \) coincide. To this end, it suffices to note that \( T^{(d)} \) is a positive semi-definite operator, simply because
\[
\sum_{h,k=0}^{d-1} \psi_h^* \psi_k T_{hk} = \sum_{h,k=0}^{d-1} \psi_h^* \psi_k \int_{-\pi}^{+\pi} \frac{d\phi}{2\pi} a(\phi) e^{-i(h-k)\phi}
\]
\[
= \int_{-\pi}^{+\pi} \frac{d\phi}{2\pi} a(\phi) \sum_{h,k=0}^{d-1} \psi_h^* \psi_k e^{-i(h-k)\phi}
\]
\[
= \int_{-\pi}^{+\pi} \frac{d\phi}{2\pi} a(\phi) \left| \sum_{k=0}^{d-1} \psi_k e^{ik\phi} \right|^2
\]
\[
\geq 0
\]
for every \( |\psi\rangle \in \mathbb{C}^d \). This can also be seen as a simple consequence of (the easy direction of) Bochner’s theorem.

\( \Box \)

**B. Proof of main result**

Before stating and proving our main result, let us fix some terminology. For an infinite matrix \( \tau \) that is also a density operator on \( \ell^2(\mathbb{N}) \), the associated **maximally correlated state** \( \Omega[\tau] \) on \( \ell^2(\mathbb{N}) \otimes \ell^2(\mathbb{N}) \) is defined by
\[
\Omega[\tau] := \sum_{h,k=0}^{\infty} \tau_{hk} |h\rangle\langle k| \otimes |h\rangle\langle k|.
\]

Maximally correlated states appear naturally in connecting coherence theory [54–62] (see also the review article [63]) with entanglement theory. When seen in this latter context, they are useful because they represent a particularly simple class of entangled states.

Before we proceed further, let us recall in passing that the Rényi entropy of a probability density defined on the interval \([−\pi, \pi]\) need not be finite; that is, it can be equal to \(-\infty\). It is always bounded from above by \( \log_2(2\pi) \), due to the non-negativity of relative entropy. Indeed, \( h_\alpha(p) \leq \log_2(2\pi) \) for every probability density \( p \) defined on \([−\pi, \pi]\) and for all \( \alpha \in (0, 1) \cup (1, \infty) \), because \( \log_2(2\pi) - h_\alpha(p) = D_\alpha(p||\bar{u}) \geq 0 \), where \( \bar{u} \) is the uniform probability density on \([−\pi, \pi]\). The same is true for the case \( \alpha = 1 \), where we identify \( h_1 \) with the standard differential entropy (6). However, as an example, if we take the probability density to be \( p(x) = c|x|^{-\frac{\alpha}{2}} \) for \( \alpha > 1 \), where \( c \) is a normalization factor, then the Rényi entropy \( h_\alpha(p) \) diverges to \(-\infty\). Note that the condition \( h_\alpha(p) > -\infty \) is equivalent to the condition \( \int_{-\pi}^{+\pi} d\phi \ p(\phi)^\alpha < \infty \). An analogous reasoning can be repeated for the case where \( \alpha = 1 \). Here, \( p(x) \propto |x|^{-1} \left( \log_2 \frac{2\pi}{|x|} \right)^{-2} \) provides an example of a probability distribution for which \( h(p) = -\infty \) (while \( p \) itself is integrable).
Theorem 7. Let \( p : [-\pi, +\pi] \rightarrow \mathbb{R}_+ \) be a probability density function with the property that one of its Rēnyi entropies is finite for some \( \alpha_0 > 1 \), i.e.

\[
\int_{-\pi}^{+\pi} d\phi \ p(\phi)^{\alpha_0} < \infty.
\]

(63)

Then the two-way assisted quantum capacity, the unassisted quantum capacity, the private capacity, the secret-key capacity, and all of the corresponding strong converse capacities of the associated bosonic dephasing channel \( N_p \) coincide, and are given by the expression

\[
\begin{align*}
Q(N_p) &= Q^\uparrow(N_p) = P(N_p) = P^\uparrow(N_p) = Q_{\text{vs}}(N_p) = Q_{\text{vs}}^\uparrow(N_p) = P(N_p) = P^\uparrow_{\text{vs}}(N_p) \\
&= D(p\|u) = \log_2(2\pi) - h(p) \\
&= \log_2(2\pi) - \int_{-\pi}^{\pi} d\phi \ p(\phi) \log_2 \frac{1}{p(\phi)}.
\end{align*}
\]

(64)

Here, \( D(p\|u) \) denotes the Kullback–Leibler divergence between \( p \) and the uniform probability density \( u \) over \([-\pi, \pi]\), and \( h(p) \) is the differential entropy of \( p \).

Remark 8. The condition on the Rēnyi entropy of \( p \) is of a purely technical nature. We expect it to be obeyed in all cases of practical interest. For example, it holds true provided that \( p \) is bounded on \([-\pi, \pi]\).

Remark 9. By using Hölder’s inequality, it can be easily verified that if (63) holds for \( \alpha_0 \geq 1 \) then it holds for all \( \alpha \) such that \( 1 \leq \alpha \leq \alpha_0 \).

Proof of Theorem 7. The smallest of all eight quantities is the unassisted quantum capacity \( Q(N_p) \), and the largest is \( P^\uparrow_{\text{vs}}(N_p) \). Therefore, it suffices to prove that

\[
Q(N_p) \geq D(p\|u), \quad P^\uparrow_{\text{vs}}(N_p) \leq D(p\|u).
\]

(65)

Note that it is elementary to verify that

\[
D(p\|u) = \int_{-\pi}^{+\pi} d\phi \ p(\phi) \log_2 \frac{1}{p(\phi)} = \log_2(2\pi) - \int_{-\pi}^{+\pi} d\phi \ p(\phi) \log_2 \frac{1}{p(\phi)} = \log_2(2\pi) - h(p),
\]

(66)

where the differential entropy \( h(p) \) is defined by (6).

To bound \( Q(N_p) \) from below, we need an ansatz for a state \(|\Psi\rangle_{AA'}\) to plug into (22). Letting \( A \) and \( A' \) be single-mode systems, we can consider the maximally entangled state \(|\Phi_d\rangle_{AA'} := \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} |k\rangle_A |k\rangle_{A'}\), locally supported on the subspace spanned by first \( d \) Fock states \(|k\rangle \) (see (15)), where \( k \in \{0, \ldots, d-1\} \).

Let us also define the truncated matrix

\[
T^{(d)}_p := \Pi_d T_p \Pi_d = \sum_{h,k=0}^{d-1} (T_p)_{hk} |h\rangle\langle k|,
\]

(67)

where

\[
\Pi_d := \sum_{k=0}^{d-1} |k\rangle\langle k|.
\]

(68)

Then note that

\[
\omega_{p,d} := (I \otimes N_p) (\Phi_d)
\]

\[
= \frac{1}{d} \sum_{h,k=0}^{d-1} (I \otimes N_p)(|hh\rangle\langle kk|)
\]

\[
= \frac{1}{d} \sum_{h,k=0}^{d-1} (T_p)_{hk} |hh\rangle\langle kk|
\]

\[
= \frac{1}{d} \Omega[T^{(d)}_p],
\]

(69)
where $\Omega[\tau]$ is defined by (62). Consider that

$$Q(N_p) \geq \limsup_{d \to \infty} I_{\text{coh}}(A)B(I \otimes N_p^{A' \to B})(\Phi_d^{A'})$$

$$= \limsup_{d \to \infty} I_{\text{coh}}(A)B_{\omega_{p,d}}$$

$$= \limsup_{d \to \infty} \left( \log_2 d - S(T_p^{(d)}/d) \right)$$

$$= \limsup_{d \to \infty} \left( \log_2 d + \frac{1}{d} \text{Tr}T_p^{(d)} \left( - \log_2 d + \log_2 T_p^{(d)} \right) \right)$$

$$= \limsup_{d \to \infty} \frac{1}{d} \text{Tr}T_p^{(d)} \log_2 T_p^{(d)}$$

$$= \int_{-\pi}^{\pi} \frac{d\phi}{2\pi} (2\pi p(\phi)) \log_2 (2\pi p(\phi))$$

$$= \log_2(2\pi) - \int_{-\pi}^{\pi} d\phi p(\phi) \log_2 \frac{1}{p(\phi)}. \quad (70)$$

Here: (i) follows from the LSD theorem (22); in (ii) we introduced the state $\omega_{p,d}$ defined by (62); (iii) comes from (23), due to the fact that $S(\Omega[\tau]) = S(\tau)$ on the one hand, and

$$\text{Tr}_A \omega_{p,d}^{AB} = \text{Tr}_A (I \otimes N_p^{A' \to B})(\Phi_d^{A'})$$

$$= \frac{1}{d} \text{Tr}_A \sum_{h,k=0}^{d-1} (T_p)_{hk} |h\rangle \langle k|_A \otimes |h\rangle \langle k|_B$$

$$= \frac{1}{d} \text{Tr} \sum_{h,k=0}^{d-1} (T_p)_{hk} \delta_{hk} |h\rangle \langle k|_B$$

$$= \frac{1}{d} \text{Tr} \sum_{k=0}^{d-1} |k\rangle \langle k|_B$$

$$= \frac{1}{d} I_B$$

and therefore $S(\text{Tr}_A (I \otimes N_p^{A' \to B})(\Phi_d^{A'})) = \log_2 d$ on the other; in (iv) we simply substituted the definition (12) of von Neumann entropy; and finally in (v) we employed Lemma 6 with the choice $a(\phi) = 2\pi p(\phi)$.

This is possible due to our assumption that (63) holds for some $\alpha > 1$. Note that $F(x) = x \log_2 x$ satisfies $|F(x)| < x^\alpha$ for all $\alpha > 1$ and for all sufficiently large $x \in \mathbb{R}^+$. This concludes the proof of the lower bound on $Q(N_p)$ in (65). We remark in passing that the Szegö theorem has been applied before, although with an entirely different scope, in the context of quantum information theory [64].

We now establish the upper bound on $P_{\text{LOCC}}^A(N_p)$ in (65). We claim that there is a sequence of LOCC protocols that can simulate $N_p$ using $\omega_{p,d}$ defined by (69) as a resource state and with error vanishing as $d \to \infty$. To see why this is the case, let $\rho$ be an arbitrary input state, and consider the $d$-dimensional teleportation protocol (46) on $\rho$ that uses $\omega_{p,d}$ as a resource. In formula, let us define

$$N_{p,d}^{A' \to B}(\rho_{A'}) \equiv T_{A'AB \to B}^{(d)} \left( \rho_{A'} \otimes \omega_{p,d}^{AB} \right). \quad (72)$$

We see that

$$N_{p,d}^{A' \to B}(\rho_{A'})$$
We now want to argue that \( x \) where 

\[
\sum_{x,z=0}^{d-1} X(x)_B Z(z)_B \Tr_{A'} \left[ \rho_{A'} \otimes \omega_{p,d} X(x)_A Z(z)_A \Phi_d^{A'} Z(z)_{A'} X(x)_{A'} \right] Z(z)_B X(x)_B
\]

valid for the maximally entangled state (44) in any finite dimension, (ix) and (xi) from (45), (x) from the calculation in (73) shows that 

\[
\sum_{x,z=0}^{d-1} \sum_{h,k=0}^{d-1} \frac{1}{d} (T_p)_{hk} X(x)_B Z(z)_B
\]

\[
\Tr_{A'} \left[ \rho_{A'} \otimes |hhkk|_A X(x)_A Z(z)_A \Phi_d^{A'} X(x)_{A'} Z(z)_{A'} X(x)_{A'} \right] Z(z)_B X(x)_B
\]

\[
\sum_{x,z=0}^{d-1} \sum_{h,k=0}^{d-1} \frac{1}{d} (T_p)_{hk} X(x)_B Z(z)_B
\]

\[
\left( e^{\frac{2\pi i}{d} z(k-h)} \right) \Tr_{A'} \left[ \rho_{A'} \otimes |h+x(k+x)|_A \Phi_d^{A'} \right] Z(z)_B X(x)_B
\]

\[
= \sum_{x=0}^{d-1} \sum_{h,k=0}^{d-1} \frac{1}{d} (T_p)_{hk} \rho_{h\otimes x, k\otimes x} h+x(k+x)_B
\]

\[
= \sum_{h,k=0}^{d-1} \frac{1}{d} \sum_{x=0}^{d-1} (T_p)_{h\otimes x, k\otimes x} \rho_{hk} |h|_B
\]

In the above derivation, (vi) follows from (46), (vii) from (69), (viii) from the formula 

\[
M \otimes 1 \left| \Phi_d \right> = 1 \otimes M^\dagger \left| \Phi_d \right>, \tag{74}
\]

valid for the maximally entangled state (44) in any finite dimension, (ix) and (xi) from (45), (x) from the identity 

\[
\Tr \left[ M_A \otimes N_{A'} \Phi_d^{A'} \right] = \frac{1}{d} \Tr \left[ MN^\dagger \right], \tag{75}
\]

and finally (xii) by a simple change of variable \( h \otimes x \mapsto h, k \otimes x \mapsto k \), once one observes that 

\[
\sum_{x=0}^{d-1} (T_p)_{hk} \mapsto \sum_{x=0}^{d-1} (T_p)_{h\otimes x, k\otimes x} = \sum_{x'=0}^{d-1} (T_p)_{h\otimes (-x'), k\otimes (-x')} = \sum_{x'=0}^{d-1} (T_p)_{h\otimes x', k\otimes x'}, \tag{76}
\]

where \( x' := -x \).

The calculation in (73) shows that 

\[
\langle h | N_{p,d}(\rho) | k \rangle = \left( \frac{1}{d} \sum_{x=0}^{d-1} (T_p)_{h\otimes x, k\otimes x} \right) \rho_{hk}.
\]

We now want to argue that for fixed \( h, k \in \mathbb{N} \) the above quantity converges to \( (T_p)_{hk} \rho_{hk} = \langle h | N_p(\rho) | k \rangle \) as \( d \to \infty \). To this end, note that if \( d \geq h, k \) we have that \( (T_p)_{h\otimes x, k\otimes x} = T_{hk} \) provided that either
x \leq \min \{d-1-h, d-1-k\} = \min \{d-h, d-k\} - 1 \text{ or } x \geq \max \{d-h, d-k\}. \text{ Therefore, } (T_p)_{h@x, k@x} \neq T_{hk} \text{ for at most } |h - k| \text{ values of } x, \text{ out of the } d \text{ possible ones. We can estimate the remainder terms pretty straightforwardly using the inequality } \left| (T_p)_{hk} \right| \leq 1, \text{ valid for all } h, k \in \mathbb{N}. \text{ Doing so yields }

\left| (T_p)_{hk} - \frac{1}{d} \sum_{x=0}^{d-1} (T_p)_{h@x, k@x} \right| \leq (T_p)_{hk} - \frac{d - |h - k|}{d} (T_p)_{hk} + \frac{|h - k|}{d} \leq \frac{2|h - k|}{d} \xrightarrow{d \to \infty} 0.

Thus, for all fixed \( h, k \in \mathbb{N} \),

\[ \langle h | N_{p,d}(\rho) | k \rangle \xrightarrow{d \to \infty} \langle h | N_p(\rho) | k \rangle \quad \forall \rho, \quad (79) \]

as claimed. We now argue that this implies the stronger fact that

\[ \lim_{d \to \infty} \left\| (N_{p,d} - N_p) \right\|_{A'\to B} \leq 0, \quad \forall \rho_{A'E}, \quad (80) \]

where it is understood that \( \rho_{A'E} \) is an arbitrary, but fixed state of a bipartite system \( A'E \), with the quantum system \( E \) arbitrary. The above identity is usually expressed in words by saying that \( N_{p,d} \) converges to \( N_p \) in the topology of strong convergence \([65, 66]\). The arguments that allow to deduce (80) from (79) are standard:

(a) The linear span of the Fock states \( \{|k\rangle\}_{k \in \mathbb{N}} \) is dense in \( \mathcal{H}_1 \), and moreover the operators \( N_{p,d}(\rho), N_p(\rho) \) are uniformly bounded in trace norm — since they are states, they all have trace norm 1. Hence, we see that (79) actually holds also when \( |h\rangle, |k\rangle \) are replaced by any two fixed vectors \( |\psi\rangle, |\phi\rangle \in \mathcal{H}_1 \). In formula,

\[ \langle \psi | N_{p,d}(\rho) | \phi \rangle \xrightarrow{d \to \infty} \langle \psi | N_p(\rho) | \phi \rangle \quad \forall \rho, \quad \forall |\psi\rangle, |\phi\rangle \in \mathcal{H}_1. \quad (81) \]

(b) Therefore, by definition \( N_{p,d}(\rho) \) converges to \( N_p(\rho) \) in the weak operator topology. Since the latter object is also a quantum state, an old result due to Davies [67, Lemma 4.3], which can also be seen as an elementary consequence of the ‘gentle measurement lemma’ [68, Lemma 9] (see also [69, Lemmata 9.4.1 and 9.4.2]), states that in fact there is trace norm convergence, i.e.

\[ \lim_{d \to \infty} \left\| (N_{p,d} - N_p) (\rho) \right\|_1 = 0 \quad \forall \rho. \quad (82) \]

(c) The topology of strong convergence is stable under tensor products with the identity channel \([65]\) (see also [66, Lemma 2]). Therefore, (82) and (80) are in fact equivalent. Since we have proved the former, the latter also follows.

We are now ready to prove that \( P_{\omega,w}^f (N_p) \leq D(p || u) \). For a fixed positive integer \( n \in \mathbb{N}_+ \), consider a generic protocol as the one depicted in Figure 2, where the channel \( N \) is now \( N_p \). Since we are dealing with the secret-key-agreement capacity, the final state \( \eta_n \) will approximate a private state \( \gamma_n \) containing \( \lfloor Rn \rfloor \) secret bits \([34, 35]\). Here, \( R \) is an achievable strong converse rate of secret-key agreement, i.e., it satisfies that \( R < P_{\omega,w}^f (N_p) \), where the right-hand side is defined by (40). Call \( \varepsilon_n := \frac{1}{2} \left\| \eta_n - \gamma_n \right\|_1 \) the corresponding trace norm error, so that

\[ \lim_{n \to \infty} \inf \varepsilon_n < 1. \quad (83) \]

Imagine now to replace each instance of \( N_p \) with its simulation \( N_{p,d} \). This will yield at the output a state \( \eta_{n,d} \), in general different from \( \eta_n \); however, because of (80), and since \( n \) here is fixed, we have that the associated error \( \delta_{n,d} \) vanishes as \( d \to \infty \), i.e.

\[ \delta_{n,d} := \frac{1}{2} \left\| \eta_{n,d} - \eta_d \right\|_1 \xrightarrow{d \to \infty} 0. \quad (84) \]
Now, after the above replacement the global protocol can be seen as an LOCC manipulation of \(n\) copies of the state \(\omega_{p,d}\) that is used to simulate \(\mathcal{N}_{p,d}\) as per (72). By the triangle inequality, the trace distance between the final state \(\eta_{n,d}\) and the private state \(\gamma_n\) satisfies

\[
\frac{1}{2} \| \eta_{n,d} - \gamma_n \|_1 \leq \frac{1}{2} \| \eta_{n,d} - \eta_n \|_1 + \frac{1}{2} \| \eta_n - \gamma_n \|_1 \leq \delta_{n,d} + \varepsilon_n.
\]  

(85)

To apply the results of [70], we need to translate the above estimate into one that uses the fidelity instead of the trace distance. Such a translation can be made with the help of the Fuchs–van de Graaf inequalities [5], here reported as (1). We obtain that \(F(\eta_{n,d}, \gamma_n) \geq (1 - \delta_{n,d} - \varepsilon_n)^2\). We can then use [70, Eq. (5.37)] directly to deduce that

\[
(1 - \delta_{n,d} - \varepsilon_n)^2 \leq F(\eta_{n,d}, \gamma_n) \leq 2^{-\frac{n \alpha}{\alpha}} (R - \bar{E}_{R,\alpha}(\omega_{p,d}))
\]

(86)

for all \(1 < \alpha \leq \alpha_0\), where

\[
\bar{E}_{R,\alpha}(\rho_{AB}) \coloneqq \inf_{\sigma \in \mathcal{S}_{AB}} \bar{D}_\alpha(\rho||\sigma)
\]

(87)

is the sandwiched \(\alpha\)-Rényi relative entropy of entanglement, and

\[
\mathcal{S}_{AB} \coloneqq \text{conv}\{ |\psi\rangle\langle\psi|_A \otimes |\phi\rangle\langle\phi|_B : |\psi\rangle_A \in \mathcal{H}_A, |\phi\rangle_B \in \mathcal{H}_B, \langle\psi|\psi\rangle = 1 = \langle\phi|\phi\rangle \}
\]

(88)

is the set of separable states over the bipartite quantum system \(AB\). We can immediately recast (86) as

\[
R \leq \frac{2}{\alpha} \frac{\alpha}{n \alpha - 1} \log_2 \left( \frac{1}{1 - \delta_{n,d} - \varepsilon_n} + \bar{E}_{R,\alpha}(\omega_{p,d}) \right)
\]

(89)

Let us now estimate the quantity \(\bar{E}_{R,\alpha}(\omega_{p,d})\). By taking as an ansatz for a separable state to be plugged into (87) simply \(\Omega[\Pi_d/d]\) (see (62) and (68)), which is manifestly separable because \(\Pi_d\) is diagonal, we conclude that

\[
\bar{E}_{R,\alpha}(\omega_{p,d}) \leq \bar{D}_\alpha(\omega_{p,d} || \Omega \left[ \frac{\Pi_d}{d} \right])
\]

\[
\overset{(\text{xiii})}{=} \frac{1}{\alpha - 1} \log_2 \text{Tr} \omega_{p,d}^\alpha \Omega \left[ \frac{\Pi_d}{d} \right]^{1 - \alpha}
\]

\[
\overset{(\text{xiv})}{=} \frac{1}{\alpha - 1} \log_2 \frac{1}{d} \text{Tr} \left( T_p(d) \right)^\alpha
\]

(90)

Here, (xiii) follows from (10), while in (xiv) we simply recalled (69).

Now, applying Lemma 6 once again with \(a(\phi) = 2\pi p(\phi)\), we surmise that

\[
\lim_{d \to \infty} \frac{1}{d} \text{Tr} \left[ \left( T_p(d) \right)^\alpha \right] = \int_{-\pi}^{+\pi} \frac{d\phi}{2\pi} \left( 2\pi p(\phi) \right)^\alpha = (2\pi)^{\alpha - 1} \int_{-\pi}^{+\pi} d\phi p(\phi)^\alpha.
\]

(91)

Therefore, from (90) we deduce that

\[
\limsup_{d \to \infty} \bar{E}_{R,\alpha}(\omega_{p,d}) \leq \log_2 (2\pi) + \frac{1}{\alpha - 1} \log_2 \int_{-\pi}^{+\pi} d\phi p(\phi)^\alpha = D_\alpha(p||u),
\]

(92)

where \(D_\alpha\) is the \(\alpha\)-Rényi divergence defined by (2). Due to both (92) and (84), taking the limit \(d \to \infty\) in (89) yields

\[
R \leq \frac{2}{\alpha} \frac{\alpha}{n \alpha - 1} \log_2 \left( \frac{1}{1 - \varepsilon_n} + D_\alpha(p||u) \right).
\]

(93)
We are now ready to take the limit \( n \to \infty \). In light of (83), we obtain that
\[
R \leq \liminf_{n \to \infty} \left( \frac{2}{n} \alpha - 1 \log_2 \frac{1}{1 - e_n} + D_\alpha(p\|u) \right) = D_\alpha(p\|u). \tag{94}
\]
The limit as \( \alpha \to 1^+ \) can be computed via Lemma 1 (and in particular (4)), due to the condition (63), which can be rephrased as \( D_{\alpha_0}(p\|u) < \infty \). It gives
\[
R \leq \liminf_{\alpha \to 1^+} D_\alpha(p\|u) = D(p\|u). \tag{95}
\]
Since \( R \) was an arbitrary achievable strong converse rate for secret-key agreement, we deduce that
\[
P_{\to}(N_p) \leq D(p\|u), \tag{96}
\]
completing the proof. \( \square \)

C. Extension to multimode channels

We will now see how to extend our main result, Theorem 7, to the case of a multimode bosonic dephasing channel. An \( m \)-mode quantum system \((m \in \mathbb{N}_+)\) is modelled mathematically by the Hilbert space \( \mathcal{H}_m = \mathcal{H}_1^{\otimes m} = L^2(\mathbb{R})^{\otimes m} = L^2(\mathbb{R}^m) \). The annihilation and creation operators \( a_j, a_j^\dagger \) \((j = 1, \ldots, m)\), defined by
\[
a_1 := a \otimes \mathbb{1} \otimes \cdots \otimes \mathbb{1}, \quad \ldots, \quad a_m := \mathbb{1} \otimes \cdots \otimes a \tag{97}
\]
in terms of the single-mode operators in (13), satisfy the canonical commutation relations
\[
[a_j, a_k] = 0 = [a_j^\dagger, a_k^\dagger], \quad [a_j, a_k^\dagger] = \delta_{jk} \mathbb{1}. \tag{98}
\]
The multimode Fock states \( |k\rangle \), indexed by \( k = (k_1, \ldots, k_m)^\dagger \in \mathbb{N}_+^m \), are given by
\[
|k\rangle := |k_1\rangle \otimes \cdots \otimes |k_m\rangle. \tag{99}
\]

Now, for a probability density function \( p \) on \([-\pi, \pi]^m\), the corresponding multimode bosonic dephasing channel is the quantum channel \( N_p^{(m)} : \mathcal{T}(\mathcal{H}_m) \to \mathcal{T}(\mathcal{H}_m) \) defined by
\[
N_p^{(m)}(\rho) := \int_{[-\pi, \pi]^m} d^m \phi \ p(\phi) \ e^{-i \sum_j a_j^\dagger a_j^\phi_j} \rho \ e^{i \sum_j a_j^\phi_j a_j^\dagger}, \tag{100}
\]
where \( j = 1, \ldots, m \), and \( \phi = (\phi_1, \ldots, \phi_m)^\dagger \).

In perfect analogy with Theorem 7, we can now prove the following.

**Theorem 10.** Let \( p : [-\pi, +\pi]^m \to \mathbb{R}_+ \) be a probability density function with the property that one of its Rényi entropies is finite for some \( \alpha_0 > 1 \), i.e.
\[
\int_{[-\pi, \pi]^m} d^m \phi \ p(\phi)^{\alpha_0} < \infty. \tag{101}
\]
Then the two-way assisted quantum capacity, the unassisted quantum capacity, the private capacity, the secret-key capacity, and all of the corresponding strong converse rates of the associated multimode bosonic dephasing channel \( N_p^{(m)} \) coincide, and are given by the expression
\[
Q(N_p^{(m)}) = Q_{\to}(N_p^{(m)}) = P(N_p^{(m)}) = P_{\to}(N_p^{(m)}) = Q_{\to}(N_p^{(m)}) = Q_{\to}(N_p^{(m)}) = P_{\to}(N_p^{(m)}) = P_{\to}(N_p^{(m)}) = D(p\|u) = m \log_2(2\pi) - h(p) = m \log_2(2\pi) - \int_{[-\pi, \pi]^m} d^m \phi \ p(\phi) \log_2 \frac{1}{p(\phi)}. \tag{102}
\]
Here, \( D(p\|u) \) denotes the Kullback–Leibler divergence between \( p \) and the uniform probability distribution \( u \) over \([-\pi, \pi]^m \), and \( h(p) \) is the differential entropy of \( p \).
Rather unsurprisingly, one of the key technical tools that we need to prove the above generalization of Theorem 7 is a multi-index version of the Szegö theorem reported here as Lemma 6. In fact, the original paper by Serra-Capizzano [48] deals already with multi-indices, so we can borrow the following result directly from [48, Theorem 2] (cf. also the proof of Lemma 6).

An **multi-index infinite Toeplitz matrix** is an operator $T : ℓ^2(\mathbb{N}^m) \to ℓ^2(\mathbb{N}^m)$ with the property that its matrix entries $T_{h,k}$ (where $|h| = (h_1, \ldots, h_m) \in \mathbb{N}^m$ is a multi-index) depend only on the difference $h - k$, in formula $T_{h,k} = a_{h-k}$. The case of interest is when

$$a_k = \int_{[-\pi,\pi]^m} |m\phi|^2 a(\phi) e^{-ik\cdot\phi}, \quad (103)$$

where $a : [-\pi,\pi]^m \to \mathbb{R}_+$ is a non-negative function, and $k \cdot \phi := \sum_{j=1}^m k_j \phi_j$. As in the setting of Szegö’s theorem, one considers the truncations of $T$ defined for some $d = (d_1, \ldots, d_m) \in \mathbb{N}^m$ by

$$T^{(d)} := \sum_{h_1,k_1=0}^{d_1-1} \cdots \sum_{h_m,k_m=0}^{d_m-1} T_{h,k} |h\rangle\langle k| \quad \text{.} \quad (104)$$

Note that $T^{(d)}$ is an operator on a space of dimension

$$D(d) := \prod_{j=1}^m d_j \quad \text{.} \quad (105)$$

The multi-index Szegö theorem then reads

$$\lim_{d \to \infty} \frac{1}{D(d)} \text{Tr} F(T^{(d)}) = \lim_{d \to \infty} \frac{1}{D(d)} \sum_{j=1}^{D(d)} F(\lambda_j(T^{(d)})) = \int_{[-\pi,\pi]^m} \frac{d\phi}{(2\pi)^m} F(a(\phi)) \quad \text{,} \quad (106)$$

where $F : \mathbb{R} \to \mathbb{R}$, and $d \to \infty$ means that $\min_{j=1,\ldots,m} d_j \to \infty$. Conditions on $a$ and $F$ so that (106) holds are as follows.

**Lemma 11** (Serra-Capizzano [48], multi-index case). If $a : [-\pi,\pi]^m \to \mathbb{R}_+$ is such that

$$\int_{[-\pi,\pi]^m} \frac{|m\phi|^2}{(2\pi)^m} a(\phi)^\alpha < \infty \quad \text{,} \quad (107)$$

for some $\alpha \geq 1$, and moreover $F : \mathbb{R}_+ \to \mathbb{R}$ is continuous and satisfies

$$F(x) = O(x^\alpha) \quad (x \to \infty) \quad \text{,} \quad (108)$$

then (106) holds.

The proof of Theorem 10 follows very closely that of Theorem 7. Let us briefly summarize the main differences.

**Proof of Theorem 10.** As an ansatz in the coherent information (70), we use a multimode maximally entangled state, defined by

$$|\Phi_d\rangle := \frac{1}{D(d)} \sum_{k_1=0}^{d_1-1} \cdots \sum_{k_m=0}^{d_m-1} |k\rangle_A |k\rangle_{A'} \quad \text{,} \quad (109)$$

where $d \in \mathbb{N}^m$ is fixed for now. Since

$$\omega_{p,d} := \left(I \otimes N_p^{(m)}\right)(\Phi_d) = \frac{1}{D(d)} \Omega[T^{(d)}] \quad \text{,} \quad (110)$$
is still a maximally correlated state, the derivation in (70) is unaffected, provided that one employs in (v) Lemma 11.

As for the converse bound on the strong converse rate, one replaces (72) with

\[
(N_{p,d}^{(m)})_{A'\to B} (\rho_{A'}) := \left( \bigotimes_{j=1}^{m} \mathcal{T}_{A'j, A_j B_j \to B_j} \right) (\rho_{A'} \otimes \omega_{p,d}^{AB}),
\]

where \( A_j \) denotes the \( j \)th mode of \( A \), and analogously for \( A' \) and \( B \). Then (73) becomes

\[
(N_{p,d}^{(m)})_{A'\to B} (\rho_{A'}) = \sum_{h_1, k_1=0}^{d_1-1} \cdots \sum_{h_m, k_m=0}^{d_m-1} \left( \frac{1}{D(d)} \right)^{d_1-1} \sum_{j_1=0}^{d_1-1} \cdots \sum_{j_m=0}^{d_m-1} (T_p)_{h_1, k_1} \cdots (T_p)_{h_m, k_m} \rho_{h,k} | h(\mathbf{k})_B .
\]

We can write an inequality analogous to (78) as

\[
\left| (T_p)_{h,k} \Bigg( \frac{1}{D(d)} \right)^{d_1-1} \sum_{j_1=0}^{d_1-1} \cdots \sum_{j_m=0}^{d_m-1} (T_p)_{h_j, k_j} \Bigg| (T_p)_{h,k} \right| D(d) \leq \frac{D(d) - \Pi_{j=1}^{m} (d_j - |h_j - k_j|)}{D(d)} \rightarrow 0.
\]

In the exact same way, one uses the above inequality to prove a generalized version of (80) as

\[
\lim_{d \to \infty} \left\| \left( (N_{p,d}^{(m)} - N_{p}^{(m)})_{A'\to B} \otimes I_E \right) (\rho_{A'E}) \right\|_1 = 0 \quad \forall \rho_{A'E} .
\]

The combination of (90) and (91) now becomes

\[
\bar{E}_{R,\alpha}(\omega_{p,d}) \leq \frac{1}{\alpha - 1} \log_2 \frac{1}{D(d)} \mathbb{E} \left[ \left( T_p^d \right)^ \alpha \right] \rightarrow \frac{1}{\alpha - 1} \log_2 \int_{[-\pi,\pi]^m} d\phi p(\phi)^{\alpha},
\]

so that we find, precisely as in (92), that

\[
\lim_{d \to \infty} \sup \bar{E}_{R,\alpha}(\omega_{p,d}) \leq m \log_2 (2\pi) + \frac{1}{\alpha - 1} \log_2 \int_{[-\pi,\pi]^m} d\phi p(\phi)^{\alpha} = D_\alpha(p||u).
\]

The rest of the proof is formally identical.

\[\square\]

IV. EXAMPLES

A. Wrapped normal distribution

The most commonly studied [71, 72] example of the bosonic dephasing channel is that which yields in (52) a matrix \( T_p \) with entries

\[
(T_p)_{h,k} = e^{-\frac{\gamma}{2} (h-k)^2},
\]

where \( \gamma > 0 \) is a parameter. The probability density function \( p : [-\pi, \pi] \rightarrow \mathbb{R}_+ \) that gives rise to this matrix is a wrapped normal distribution, that is, a normal distribution on \( \mathbb{R} \) with variance \( \gamma \) ‘wrapped’ around the unit circle. In formula, this is given by

\[
p_\gamma(\phi) = \frac{1}{\sqrt{2\pi\gamma}} \sum_{k=-\infty}^{+\infty} e^{-\frac{\gamma}{2} (\phi + 2\pi k)^2}.
\]
FIG. 3: The probability density functions of the wrapped normal (118), von Mises (127), and wrapped Cauchy distributions (133), plotted as a function of $\phi \in [-\pi, \pi]$ for the case where $\gamma = \lambda = \kappa = 0.5$.

Its entropy can be expressed as [73, Chapter 3, § 3.3]

$$h(p_\gamma) = \frac{1}{\ln 2} \left(-\ln \left(\frac{\varphi(e^{-\gamma})}{2\pi}\right) + 2 \sum_{k=1}^{\infty} \frac{(-1)^k e^{-\frac{\gamma}{2}(k^2+k)}}{k \left(1 - e^{-ky}\right)} \right),$$

(119)

where

$$\varphi(q) := \prod_{k=1}^{\infty} \left(1 - q^k\right)$$

(120)

is the Euler function. Therefore, the capacities of the channel $\mathcal{N}_{p_\gamma}$ are given by

$$Q(\mathcal{N}_{p_\gamma}) = P(\mathcal{N}_{p_\gamma}) = P^\dagger(\mathcal{N}_{p_\gamma}) = Q^\dagger(\mathcal{N}_{p_\gamma}) = Q^\rightarrow(\mathcal{N}_{p_\gamma}) = P^\rightarrow(\mathcal{N}_{p_\gamma}) = P^\leftarrow(\mathcal{N}_{p_\gamma}) = D(p_\gamma||u) = \log_2 \varphi(e^{-\gamma}) + \frac{2}{\ln 2} \sum_{k=1}^{\infty} \frac{(-1)^{k-1} e^{-\frac{\gamma}{2}(k^2+k)}}{k \left(1 - e^{-ky}\right)}.$$  

(121)

It is instructive to obtain asymptotic expansions of the above expressions in the limits $\gamma \ll 1$ (small dephasing) and $\gamma \gg 1$ (large dephasing).

- **Small dephasing.** When $\gamma \to 0^+$, the channel $\mathcal{N}_{p_\gamma}$ approaches the identity over an infinite-dimensional Hilbert space. Therefore, it is intuitive to expect that its capacities will diverge. To determine its asymptotic behavior, it suffices to note that in this limit the entropy of the wrapped normal distribution, which is very concentrated around 0, is well approximated by that of the corresponding normal variable on the whole $\mathbb{R}$, i.e., $\frac{1}{2} \log_2 (2\pi e\gamma)$. Thus

$$Q(\mathcal{N}_{p_\gamma}) = P(\mathcal{N}_{p_\gamma}) = P^\dagger(\mathcal{N}_{p_\gamma}) = Q^\dagger(\mathcal{N}_{p_\gamma}) = Q^\rightarrow(\mathcal{N}_{p_\gamma}) = P^\rightarrow(\mathcal{N}_{p_\gamma}) = P^\leftarrow(\mathcal{N}_{p_\gamma}) \approx \frac{1}{2} \log_2 \frac{2\pi}{e\gamma}. $$

(122)

In practice, already for $\gamma \approx 1$ the above estimate is within about 1% of the actual value.
\* \textbf{Large dephasing.} A straightforward computation using the series representation

\[ -\ln \varphi(q) = \sum_{k=1}^{\infty} \frac{1}{k} \frac{q^k}{1 - q^k} \]  

(123)

yields the expansion

\[
\ln \varphi(q) + 2 \sum_{k=1}^{\infty} \frac{(-1)^{k-1} q^{-\frac{1}{2}(k^2 + k)}}{k (1 - q^k)} = q + \frac{q^2}{2} - \frac{q^3}{3} + \frac{q^4}{4} - \frac{q^5}{5} + \frac{2q^6}{3} + O(q^7)
\]

(124)

This can be plugged into (121) to give

\[
Q(N_{p,\gamma}) = Q^\dagger(N_{p,\gamma}) = P(N_{p,\gamma}) = P^\dagger(N_{p,\gamma}) = Q_{\rightarrow}(N_{p,\gamma}) = Q_{\leftarrow}(N_{p,\gamma}) = P_{\rightarrow}(N_{p,\gamma}) = P_{\leftarrow}(N_{p,\gamma})
\]

\[
= \frac{2}{\ln 2} e^{-\gamma} - \log_2 (1 + e^{-\gamma}) + O\left(e^{-6\gamma}\right)
\]

\[
= \frac{e^{-\gamma}}{\ln 2} + O\left(e^{-2\gamma}\right).
\]

(125)

Incidentally, the combination of these two regimes yields an excellent approximation of the capacities across the whole range of \(\gamma > 0\). Namely,

\[
Q(N_{p,\gamma}) = Q^\dagger(N_{p,\gamma}) = P(N_{p,\gamma}) = P^\dagger(N_{p,\gamma}) = Q_{\rightarrow}(N_{p,\gamma}) = Q_{\leftarrow}(N_{p,\gamma}) = P_{\rightarrow}(N_{p,\gamma}) = P_{\leftarrow}(N_{p,\gamma})
\]

\[
\approx \max \left\{ \frac{1}{2} \log_2 \frac{2\pi}{e\gamma}, \frac{2}{\ln 2} e^{-\gamma} - \log_2 (1 + e^{-\gamma}) \right\}.
\]

(126)

The maximum absolute difference between the left-hand side and the right-hand side for \(\gamma > 0\) is less than \(4 \times 10^{-3}\).

\section*{B. Von Mises distribution}

The \textbf{von Mises distribution} on \([-\pi, +\pi]\) is defined by

\[ p_\lambda(\phi) = \frac{e^{\frac{1}{\lambda} \cos(\phi)}}{2\pi I_0(1/\lambda)}, \]

(127)

where \(I_n\) denotes a modified Bessel function of the first kind. Here, \(\lambda > 0\) is a parameter that plays a role analogous to that \(\gamma > 0\) played in the case of the wrapped normal. The matrix \(T_{p,\lambda}\) obtained in (52) for \(p = p_\lambda\) is given by

\[ (T_{p,\lambda})_{hk} = \frac{I_{|h-k|}(1/\lambda)}{I_0(1/\lambda)}. \]

(128)

The differential entropy of \(p_\lambda\) can be calculated analytically, yielding [73, Chapter 3, Section 3.3]

\[ h(p_\lambda) = \log_2 (2\pi I_0(1/\lambda)) - \frac{1}{\ln 2} \frac{I_1(1/\lambda)}{I_0(1/\lambda)}. \]

(129)

Therefore, the capacities of the corresponding bosonic dephasing channel are given by

\[
Q(N_{p,\lambda}) = Q^\dagger(N_{p,\lambda}) = P(N_{p,\lambda}) = P^\dagger(N_{p,\lambda}) = Q_{\rightarrow}(N_{p,\lambda}) = Q_{\leftarrow}(N_{p,\lambda}) = P_{\rightarrow}(N_{p,\lambda}) = P_{\leftarrow}(N_{p,\lambda})
\]

\[
= \frac{1}{\ln 2} \frac{I_1(1/\lambda)}{I_0(1/\lambda)} - \log_2 I_0(1/\lambda).
\]

(130)
C. Wrapped Cauchy distribution

Our final example of a probability distribution on the circle, and of the bosonic dephasing channel associated to it, is defined similarly to the wrapped normal distribution, but this time starting from the Cauchy probability density function. Namely, for some parameter \( \kappa > 0 \) we set

\[
p_\kappa(\phi) := \sum_{k=-\infty}^{+\infty} \frac{\sqrt{\kappa}}{\pi \left( \kappa + (\phi + 2\pi k)^2 \right)} = \frac{1}{2\pi} \frac{\sinh(\sqrt{\kappa})}{\cosh(\sqrt{\kappa}) - \cos \phi}.
\]

(131)

For a proof of the second identity, see [74, p. 51]. The matrix \( T_{p_\kappa} \) obtained in (52) for \( p = p_\kappa \) is given by

\[
(T_{p_\kappa})_{hk} = e^{-\sqrt{\kappa}|h-k|}.
\]

(132)

The differential entropy of \( p_\kappa \) is equal to \( \log_2\left(2\pi(1-e^{-2\sqrt{\kappa}})\right) \) [73, Chapter 3, § 3.3], implying that the various capacities of the corresponding bosonic dephasing channel \( N_{p_\kappa} \) are equal to

\[
C(N_{p_\kappa}) = \log_2\left(\frac{1}{1-e^{-2\sqrt{\kappa}}}\right).
\]

(133)
[45] F. Avram. On bilinear forms in Gaussian random variables and Toeplitz matrices. *Probability Theory and Related Fields*, 79(1):37–45, 1988. 11
[46] N. L. Zamarashkin and E. E. Tyrtyshnikov. Distribution of the eigenvalues and singular numbers of Toeplitz matrices under weakened requirements on the generating function. *Matematicheskiy Sbornik*, 188(8):83–92, 1997. 11
[47] P. Tilli. Locally Toeplitz sequences: spectral properties and applications. *Linear Algebra and its Applications*, 278(1-3):91–120, 1998. 11
[48] S. Serra-Capizzano. Test functions, growth conditions and Toeplitz matrices. *Rendiconti del Circolo Matematico di Palermo, Serie II, Supplemento*, 68:791–795, 2002. 11, 12, 19
[49] A. Böttcher, S. M. Grudsky, and E. A. Maksimenko. Pushing the envelope of the test functions in the Szegö and Avram-Parter theorems. *Linear Algebra and its Applications*, 429(1):346–366, 2008. 11
[50] A. Böttcher and S. M. Grudsky. *Spectral properties of banded Toeplitz matrices*. Birkhäuser Verlag, Basel, 2000. 11
[51] A. Böttcher and S. M. Grudsky. *Spectral properties of banded Toeplitz matrices*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2005.
[52] C. Garoni and S. Serra-Capizzano. *Generalized locally Toeplitz sequences: theory and applications. Vol. I*. Springer, Cham, 2018. 11
[53] S. M. Grudsky. Eigenvalues of larger Toeplitz matrices: the asymptotic approach. Lecture notes, 2010. 11
[54] J. Åberg. Subspace preservation, subspace locality, and gluing of completely positive maps. *Annals of Physics*, 313(2):326–367, 2004. 12
[55] J. Åberg. Quantifying superposition. 2006. arXiv:quant-ph/0612146.
[56] T. Baumgratz, M. Cramer, and M. B. Plenio. Quantifying coherence. *Physical Review Letters*, 113:140401, 2014.
[57] A. Winter and D. Yang. Operational resource theory of coherence. *Physical Review Letters*, 116:120404, 2016.
[58] E. Chitambar and M.-H. Hsieh. Relating the resource theories of entanglement and quantum coherence. *Physical Review Letters*, 117:020402, 2016.
[59] B. Regula, K. Fang, X. Wang, and G. Adesso. One-shot coherence distillation. *Physical Review Letters*, 121:010401, 2018.
[60] K. Fang, X. Wang, L. Lami, B. Regula, and G. Adesso. Probabilistic distillation of quantum coherence. *Physical Review Letters*, 121:070404, 2018.
[61] L. Lami, B. Regula, and G. Adesso. Generic bound coherence under strictly incoherent operations. *Physical Review Letters*, 122:150402, 2019.
[62] L. Lami. Completing the Grand Tour of asymptotic quantum coherence manipulation. *IEEE Transactions on Information Theory*, 66(4):2165–2183, 2020. 12
[63] A. Streltsov, G. Adesso, and M. B. Plenio. Colloquium: Quantum coherence as a resource. *Reviews of Modern Physics*, 89:041003, 2017. 12
[64] C. Lupo, V. Giovannetti, and S. Mancini. Capacities of lossy bosonic memory channels. *Physical Review Letters*, 104:030501, 2010. 14
[65] M. E. Shirokov and A. S. Holevo. On approximation of infinite-dimensional quantum channels. *Problemy Peredachi Informatsii*, 44(2):3–22, 2008. (English translation: Problems of Information Transmission 44(2):73–90, 2008). 16
[66] M. E. Shirokov. On the energy-constrained diamond norm and its application in quantum information theory. *Problems of Information Transmission*, 54(1):20–33, 2018. 16
[67] E. B. Davies. Quantum stochastic processes. *Communications in Mathematical Physics*, 15(4):277–304, 1969. 16
[68] A. Winter. Coding theorem and strong converse for quantum channels. *IEEE Transactions on Information Theory*, 45(7):2481–2485, 1999. 16
[69] M. M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2nd edition, 2017. 16
[70] M. M. Wilde, M. Tomamichel, and M. Berta. Converse bounds for private communication over quantum channels. *IEEE Transactions on Information Theory*, 63(3):1792–1817, 2017. 17
[71] A. Arqand, L. Memarzadeh, and S. Mancini. Quantum capacity of a bosonic dephasing channel. *Physical Review A*, 102(4):042413, 2020. 20
[72] A. Arqand, L. Memarzadeh, and S. Mancini. Energy-constrained LOCC-assisted quantum capacity of bosonic dephasing channel. 2021. arXiv:2111.04173. 20
[73] F. Papadimitriou. *Spatial Entropy and Landscape Analysis*. RaumFragen: Stadt – Region – Landschaft. Springer Fachmedien Wiesbaden, 2022. 21, 22, 23

[74] K. V. Mardia and P. E. Jupp. *Directional statistics*. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester, 2000. 23