Transmission distance in the space of quantum channels

Diego G. Bussandri,[1,2] Pedro W. Lamberti,[2,3,4] and Karol Życzkowski[4,5]

[1]*Instituto de Física La Plata and Departamento de Física, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, C.C. 67, 1900 La Plata, Argentina*

[2]*Consejo Nacional de Investigaciones Científicas y Técnicas de la República Argentina, Avenida Rivadavia 1917, C1033AAJ Caba, Argentina*

[3]*Facultad de Matemática, Astronomía, Física y Computación, Universidad Nacional de Córdoba, Avenida Medina Allende s/n,Ciudad Universitaria, X5000HUA Córdoba, Argentina*

[4]*Faculty of Physics, Astronomy and Applied Computer Science, Institute of Theoretical Physics, Jagiellonian University, ul. Łojasiewicza 11, 30-348 Kraków, Poland*

[5]*Center for Theoretical Physics, Polish Academy of Sciences, Al. Lotników 32/46, 02-668 Warszawa, Poland*

(Received 24 April 2023; revised 14 June 2023; accepted 16 June 2023; published 6 July 2023)

We analyze two ways to obtain distinguishability measures between quantum maps by employing the square root of the quantum Jensen-Shannon divergence, which forms a true distance in the space of density operators. The arising measures are the transmission distance between quantum channels and the entropic channel divergence. We investigate their mathematical properties and discuss their physical meaning. Additionally, we establish a chain rule for the entropic channel divergence, which implies the amortization collapse, a relevant result with potential applications in the field of discrimination of quantum channels and converse bounds. Finally, we analyze the distinguishability between two given Pauli channels and study exemplary Hamiltonian dynamics under decoherence.

DOI: 10.1103/PhysRevA.108.012604

I. INTRODUCTION

The notion of quantum channel distinguishability is at the core of quantum information theory, and it plays a central role in a variety of contexts. Different works investigate the mathematical and physical conditions for a suitable measure of distance between quantum maps and, correspondingly, various such measures have been introduced, with trace distance and quantum fidelity being the most widely used [1]. Constructing a universal distance measure in the space of quantum maps that fulfils all the suitable requirements is strongly motivated by the recent literature. However, finding such a *gold standard* is rather difficult [2], and one tries to identify distance measures capable to compare theoretically idealized quantum channels with their noisy experimental implementations.

Within the list of relevant requirements for a measure studied, an important property is the triangle inequality, as it allows one to construct a true distance and it serves as a tool to establish other features, including the chaining property. Recently, Virosztek [3] and Sra [4] demonstrated that the square root of the quantum Jensen-Shannon divergence (QJSD) satisfies the triangle inequality for any quantum states of an arbitrary finite dimension. This extensively used entropic distinguishability measure has appealing properties and it has been widely used in quantum information theory [5–8].

The main aim of this paper is to extend the transmission distance, defined as the square root of the quantum Jensen-Shannon divergence [9], to the space of quantum channels. We study two different approaches to carry out this goal: Making use of the Choi-Jamiołkowski isomorphism, we arrive at the *transmission distance between quantum channels*. Furthermore, by optimizing the channel output over all possible inputs, we investigate the *entropic channel divergence*. Additionally, we establish their operational interpretations. In the case of the transmission distance between quantum channels, we point out the connection with the capacity of a dense coding protocol. The entropic channel divergence, on the other hand, corresponds to a particular case of the quantum reading capacity: the maximum amount of information that can be obtained in a readout process (channel decoding) per cell of a digital memory [10].

Going beyond the required properties for having well-behaved measures of distance between quantum operations, we establish a chain rule for the entropic channel divergence. This chain rule was originally proposed in Eq. (4) of Ref. [11] for the quantum relative entropy, motivated by its classical counterpart. However, the extension of the quantum relative entropy to the space of quantum maps through optimization of its inputs does not satisfy this particular chain rule.

We address the issue of the *amortized distinguishability of quantum channels*, relevant to analyze the problem of hypothesis testing for quantum channels [12]. The idea behind amortized distance measures is to consider two quantum states as inputs of two different quantum channels to explore the biggest distance between these channels without considering the original distinguishability that the input states may have. The chain rule leads to another property called *amortization collapse* [12], which occurs if the channel divergence is equal to its amortized version. In such a case, one obtains useful single-letter converse bounds on the capacity of adaptive channel discrimination protocols [13].

Finally, we will examine two specific applications for the entropic distinguishability measures: (a) Pauli channels, with a focus on studying noise in the standard quantum teleportation channel [14], and (b) the distinguishability of Hamiltonians under decoherence, a particular case within the discrimination of superoperators proposed in Refs. [15,16].

This paper is organized as follows. In Sec. II we summarize the main properties of the transmission distance in the space of quantum states. In Sec. III we introduce the transmission distance between quantum channels through the Choi-Jamiołkowski isomorphism and study its properties. The entropic channel divergence is proposed and analyzed in Sec. IV.

The chain rule and the amortization collapse of the entropic channel divergence are presented in Sec. IV and in Sec. V C we consider a set of quantum maps, for which the proposed measures are equal. In Sec. V the physical motivations and operational meanings of the introduced distances are discussed. In Sec. VI, we compute analytically the distances for Pauli channels and for arbitrary Hamiltonians under decoherence. Section VII concludes the paper with a brief review of results obtained.

II. QJSD AND TRANSMISSION DISTANCE IN THE SPACE OF QUANTUM STATES

Let $\mathcal{M}_N$ be the space of density matrices ρ (positive and normalized operators, $\rho \geqslant 0$ and $\mathrm{Tr}\rho = 1$, respectively) defined on a N-dimensional Hilbert space.

The von Neumann entropy, $S(\rho) = -\mathrm{Tr}[\rho \log_2 \rho]$, satisfies the concavity property [17]

$$S(\overline{\rho}) \geqslant \sum_i p_i S(\rho_i), \tag{1}$$

for a given ensemble of quantum states $\{p_i, \rho_i\}_i$, with the weighted average $\overline{\rho} = \sum_i p_i \rho_i$. This property gives rise to a suitable symmetric measure of distinguishability between the states comprising the ensemble (according to the classical probability vector $\boldsymbol{p} = \{p_i\}_i$) called the *Holevo quantity* [18,19] or quantum Jensen-Shannon divergence [3,4,9,20,21]:

$$\mathrm{QJSD}_{\boldsymbol{p}}(\rho_1, \dots, \rho_n) = S(\overline{\rho}) - \sum_i p_i S(\rho_i). \tag{2}$$

Making use of the quantum relative entropy [17] between two states ρ and σ,

$$S_r(\rho||\sigma) = \mathrm{Tr}[\rho(\log_2 \rho - \log_2 \sigma)], \tag{3}$$

the quantum divergence can be recast in the form

$$D_{\mathrm{JS}}^{\boldsymbol{p}}(\rho_1, \dots, \rho_n) = \sum_i p_i S_r(\rho_i||\overline{\rho}). \tag{4}$$

This equality allows us to interpret the quantity $\mathrm{QJSD}_{\boldsymbol{p}}(\rho_1, \dots, \rho_n)$ as *total divergence to the average* (or *information radius*) quantifying how much information is discarded if we describe the system employing just the convex combination $\overline{\rho} = \sum_i p_i \rho_i$. An analogous interpretation can be given in the classical setup [22,23].

In the case of a binary ensemble of states ρ and σ combined with equal weights, we can employ a simplified notation:

$$D_{\mathrm{JS}}(\rho, \sigma) = S\left(\frac{\rho + \sigma}{2}\right) - \frac{1}{2}S(\rho) - \frac{1}{2}S(\sigma). \tag{5}$$

Regarding mathematical properties, the QJSD satisfies the *indiscernibles identity* [21]:

$$\begin{aligned} &0 \leqslant \mathrm{QJSD}(\rho, \sigma) \leqslant 1 \;\text{ with} \\ &\mathrm{QJSD}(\rho, \sigma) = 0 \iff \rho = \sigma, \\ &\mathrm{QJSD}(\rho, \sigma) = 1 \iff \mathrm{supp}(\rho) \perp \mathrm{supp}(\sigma), \end{aligned} \tag{6}$$

where $\mathrm{supp}(\rho) \perp \mathrm{supp}(\sigma)$ denotes ρ and σ with orthogonal supports.

The quantum relative entropy satisfies the monotonicity [17] with respect to any completely positive trace-preserving (CPTP) map Φ. This property, also called the *data processing inequality* [21], is thus inherited by the quantum divergence:

$$\mathrm{QJSD}[\Phi(\rho), \Phi(\sigma)] \leqslant \mathrm{QJSD}(\rho, \sigma). \tag{7}$$

Furthermore, monotonicity implies that QJSD satisfies the restricted additivity,

$$\mathrm{QJSD}(\rho_1 \otimes \sigma, \rho_2 \otimes \sigma) = \mathrm{QJSD}(\rho_1, \rho_2), \tag{8}$$

and the invariance with respect to an arbitrary unitary transformation U acting on both states:

$$\mathrm{QJSD}(U\rho U^\dagger, U\sigma U^\dagger) = \mathrm{QJSD}(\rho, \sigma).$$

In the single-qubit case, $N = 2$, Briët and Harremoës showed [9] that the square root of the QJSD, known as the *transmission distance*,

$$d_t(\rho, \sigma) := \sqrt{\mathrm{QJSD}(\rho, \sigma)}, \tag{9}$$

satisfies the triangle inequality,

$$d_t(\rho, \sigma) \leqslant d_t(\rho, \chi) + d_t(\chi, \sigma), \tag{10}$$

for any $\rho, \sigma, \chi \in \mathcal{M}_2$. Recently, this result has been established for an arbitrary finite dimension N and extended to the cone of positive matrices [3,4].

The transmission distance can be bounded by other known distance measures. For instance, the *trace distance* $T(\rho, \sigma) = \frac{1}{2}\mathrm{Tr}[\sqrt{(\rho - \sigma)^2}]$ allows one to obtain the bounds

$$\frac{T(\rho, \sigma)}{\sqrt{2\ln(2)}} \leqslant d_t(\rho, \sigma) \leqslant \sqrt{T(\rho, \sigma)}, \tag{11}$$

valid for an arbitrary dimension N. The upper bound was derived in Ref. [9], while the lower one follows from inequalities in Ref. [5]:

$$2(1-\alpha)^2 T(\rho, \sigma)^2 \leqslant \mathrm{Tr}[\rho(\ln \rho - \ln \overline{\rho}_\alpha)],$$

with $\overline{\rho}_\alpha = \alpha\rho + (1-\alpha)\sigma$ and $0 < \alpha < 1$. Inserting $\alpha = 1/2$, one arrives at

$$\frac{T(\rho, \sigma)^2}{2\ln(2)} \leqslant S_r\left(\rho \middle\| \frac{\rho + \sigma}{2}\right) \;\text{and}$$
$$\frac{T(\sigma, \rho)^2}{2\ln(2)} \leqslant S_r\left(\sigma \middle\| \frac{\rho + \sigma}{2}\right).$$

The constant $\log_2$ appears above as the quantum relative entropy (3) is defined here with logarithm base 2. Therefore, we obtain

$$\frac{T(\rho,\sigma)^2}{2\ln(2)} \leqslant \frac{1}{2}S_r\left(\rho\middle\|\frac{\rho+\sigma}{2}\right)+\frac{1}{2}S_r\left(\sigma\middle\|\frac{\rho+\sigma}{2}\right),$$

and by taking the square root we arrive at the lower bound in inequality (11).

A complementary upper bound for the transmission distance in terms of the square root of the quantum fidelity, $F(\rho,\sigma) = (\mathrm{Tr}[\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}])^2$,

$$\sqrt{\mathrm{QJSD}(\rho,\sigma)} \leqslant D_E(\rho,\sigma), \tag{12}$$

was established in Ref. [24]. The quantity D_E is called the *entropic distance* [21],

$$D_E(\rho,\sigma) = \sqrt{H_2\left\{\frac{1}{2}[1-\sqrt{F(\rho,\sigma)}]\right\}}, \tag{13}$$

as it is a function of the binary entropy, $H_2(x) = -x\log_2 x - (1-x)\log_2(1-x)$ for $x \in [0,1]$.

III. TRANSMISSION DISTANCE BETWEEN QUANTUM CHANNELS AND JAMIOŁKOWSKI ISOMORPHISM

In the preceding section, we recalled the transmission distance in the space $\mathcal{M}_N$ of quantum states. Let us introduce now a measure of distinguishability between completely positive trace-preserving maps, $\mathcal{E} : \mathcal{M}_N \to \mathcal{M}_N$, by using the Choi-Jamiołkowski isomorphism which establishes a one-to-one correspondence between a quantum operation $\mathcal{E}$ and the corresponding bipartite quantum state $\rho_{\mathcal{E}}$ [25]:

$$\rho_{\mathcal{E}} = (\mathcal{E}\otimes\mathbb{1})(|\Phi\rangle\langle\Phi|). \tag{14}$$

Here

$$|\Phi\rangle = \sum_i \frac{1}{\sqrt{N}}|i\rangle_a|i\rangle_b \tag{15}$$

denotes the maximally entangled, generalized Bell state, represented in some orthonormal basis $\{|i\rangle_x\}_i \in \mathcal{H}_N$ of the N-dimensional Hilbert space. The bipartite state $\rho_{\mathcal{E}}$ is called the *Choi state* of the map $\mathcal{E}$ and represents a mixed state in $\mathcal{M}_{N^2}$. It emerges by applying $\mathcal{E}$ to the principal system, maximally entangled with an ancilla of the same dimension N.

Making use of this isomorphism, we apply Eq. (9) to define the *transmission distance between channels* $\mathcal{E}$ and $\mathcal{F}$:

$$d_t^{\mathrm{iso}}(\mathcal{E},\mathcal{F}) := d_t(\rho_{\mathcal{E}},\rho_{\mathcal{F}}). \tag{16}$$

Instead of QJSD we use its square root d_t to assure that the triangle inequality is satisfied [3] and Eq. (16) can serve as a metric between quantum maps [1].

Properties of $d_t^{\mathrm{iso}}(\mathcal{E},\mathcal{F})$

A list of required properties for a suitable measure of distinguishability between quantum maps was discussed in Refs. [1,2,16]. In particular, a gold standard should be a true distance and satisfy stability, chaining, the postprocessing inequality, and the unitary invariance. Let us now verify which of these requirements are satisfied by the distance $d_t^{\mathrm{iso}}(\mathcal{E},\mathcal{F})$.

Since the triangle inequality (10) is satisfied for the transmission distance in the state space, the quantity $d_t^{\mathrm{iso}}(\mathcal{E},\mathcal{F})$ is symmetric in its arguments, it satisfies the triangular inequality, it is non-negative, and it vanishes if and only if $\mathcal{E}=\mathcal{F}$). Hence $d_t^{\mathrm{iso}}(\mathcal{E},\mathcal{F})$ forms a true distance in the space of quantum maps.

For this kind of measures one often requires their *stability* with respect to the tensor product:

$$d_t^{\mathrm{iso}}(\mathcal{E}\otimes\mathbb{1},\mathcal{F}\otimes\mathbb{1}) = d_t^{\mathrm{iso}}(\mathcal{E},\mathcal{F}). \tag{17}$$

This fact can be demonstrated employing the restricted additivity (8), and relation $\rho_{\mathcal{E}\otimes\mathbb{1}} = \rho_{\mathcal{E}}\otimes\rho_{\mathbb{1}}$, which yield

$$\begin{aligned} d_t^{\mathrm{iso}}(\mathcal{E}\otimes\mathbb{1},\mathcal{F}\otimes\mathbb{1}) &= \sqrt{\mathrm{QJSD}(\rho_{\mathcal{E}}\otimes\rho_{\mathbb{1}},\rho_{\mathcal{F}}\otimes\rho_{\mathbb{1}})} \\ &= \sqrt{\mathrm{QJSD}(\rho_{\mathcal{E}},\rho_{\mathcal{F}})} = d_t^{\mathrm{iso}}(\mathcal{E},\mathcal{F}). \end{aligned}$$

Another property of *chaining* is relevant to estimate errors in protocols of quantum information processing. It is satisfied by a distance d if for any four maps $\mathcal{E}_1, \mathcal{F}_1, \mathcal{E}_2$, and$\mathcal{F}_2$ the distance between their concatenations can be bounded from above:

$$d(\mathcal{E}_2\circ\mathcal{E}_1,\mathcal{F}_2\circ\mathcal{F}_1) \leqslant d(\mathcal{E}_1,\mathcal{F}_1)+d(\mathcal{E}_2,\mathcal{F}_2). \tag{18}$$

In general, this property is not satisfied by the distance $d_t^{\mathrm{iso}}(\mathcal{E},\mathcal{F})$ defined (16) by the Jamiołkowski isomorphism. To show a counterexample consider the following collection of four selected Choi states analyzed in Ref. [2]:

$$\rho_{\mathcal{E}_1} = \tfrac{1}{2}\mathrm{diag}(1,1,0,0), \tag{19}$$

$$\rho_{\mathcal{E}_2} = \tfrac{1}{2}\mathrm{diag}(1,0,0,1), \tag{20}$$

$$\rho_{\mathcal{F}_1} = \rho_{\mathcal{E}_1}, \tag{21}$$

$$\rho_{\mathcal{F}_2} = \tfrac{1}{2}\mathrm{diag}(0,0,1,1). \tag{22}$$

Hence $\rho_{\mathcal{E}_2\circ\mathcal{E}_1} = \rho_{\mathcal{E}_1}$ and $\rho_{\mathcal{F}_2\circ\mathcal{F}_1} = \rho_{\mathcal{F}_2}$, so the transmission distance between both composed maps reads

$$d_t^{\mathrm{iso}}(\mathcal{E}_2\circ\mathcal{E}_1,\mathcal{F}_2\circ\mathcal{F}_1) = d_t^{\mathrm{iso}}(\mathcal{E}_1,\mathcal{F}_2).$$

As the Choi states $\rho_{\mathcal{E}_1}$ and $\rho_{\mathcal{F}_2}$ have orthogonal supports, the distance $d_t^{\mathrm{iso}}(\mathcal{E}_1,\mathcal{F}_2) = 1$, as it admits the maximal value of the implied identity of indiscernibles (6). Since $\rho_{\mathcal{F}_1} = \rho_{\mathcal{E}_1}$ one has

$$d_t^{\mathrm{iso}}(\mathcal{E}_1,\mathcal{F}_1)+d_t^{\mathrm{iso}}(\mathcal{E}_2,\mathcal{F}_2) = d_t^{\mathrm{iso}}(\mathcal{E}_2,\mathcal{F}_2).$$

Taking into account that $\rho_{\mathcal{E}_2}$ and $\rho_{\mathcal{F}_2}$ do not have orthogonal supports, we obtain the inequality

$$\begin{aligned} d_t^{\mathrm{iso}}(\mathcal{E}_2\circ\mathcal{E}_1,\mathcal{F}_2\circ\mathcal{F}_1) &> d_t^{\mathrm{iso}}(\mathcal{E}_2,\mathcal{F}_2) \\ &= d_t^{\mathrm{iso}}(\mathcal{E}_1,\mathcal{F}_1)+d_t^{\mathrm{iso}}(\mathcal{E}_2,\mathcal{F}_2), \end{aligned}$$

which provides a counterexample of inequality (18).

However, the chaining property holds in a particular case, if one of the maps applied first, $\mathcal{E}_1$ or $\mathcal{F}_1$, is bistochastic: trace preserving and unital. As a consequence of the monotonicity of the transmission distance and the triangle inequality, the chaining property holds for a bistochastic argument, $\mathcal{F}_1 = \mathcal{D}_{\mathrm{bi}}$. To demonstrate the desired inequality,

$$d_t^{\mathrm{iso}}(\mathcal{E}_2\circ\mathcal{E}_1,\mathcal{F}_2\circ\mathcal{D}_{\mathrm{bi}}) \leqslant d_t^{\mathrm{iso}}(\mathcal{E}_1,\mathcal{D}_{\mathrm{bi}})+d_t^{\mathrm{iso}}(\mathcal{E}_2,\mathcal{F}_2), \tag{23}$$

we follow directly the same steps as in Ref. [1]. By applying the triangle inequality, we have

$$d_t^{\text{iso}}(\mathcal{E}_2\circ\mathcal{E}_1,\mathcal{F}_2\circ\mathcal{D}_{\text{bi}}) \leqslant d_t^{\text{iso}}(\mathcal{E}_2\circ\mathcal{E}_1,\mathcal{E}_2\circ\mathcal{D}_{\text{bi}}) + d_t^{\text{iso}}(\mathcal{E}_2\circ\mathcal{D}_{\text{bi}},\mathcal{F}_2\circ\mathcal{D}_{\text{bi}}). \tag{24}$$

The main argument of the proof is to apply contractivity to the right-hand side of the previous equation, reaching Eq. (23).

In the first place, it should be noted that for arbitrary operations $\mathcal{E}$ and $\mathcal{F}$, it holds [1,26] that

$$\rho_{\mathcal{E}\circ\mathcal{F}} = (\mathcal{F}^{\intercal}\otimes\mathcal{E})(|\Phi\rangle\langle\Phi|), \tag{25}$$

where $\mathcal{F}^{\intercal}$ is the quantum operation determined by the transpose of the Kraus operators corresponding to the map $\mathcal{F}$. Specifically, if $\mathcal{F}(\rho)=\sum_i F_i\rho F_i^{\dagger}$, for all quantum state ρ, then $\mathcal{F}^{\intercal}(\rho)=\sum_i F_i^{\intercal}\rho(F_i^{\intercal})^{\dagger}=\sum_i F_i^{\intercal}\rho F_i^{*}$.

If $\mathcal{F}$ is bistochastic ($\sum_i F_iF_i^{\dagger}=\mathbb{1}$), the map $\mathcal{F}^{\intercal}$ is a trace-preserving operation: by taking the transpose in $\sum_i F_iF_i^{\dagger}=\mathbb{1}$ it follows that $\sum_i F_i^{*}F_i^{\intercal}=\mathbb{1}$.

Let us return to Eq. (24). By using Eq. (25), we can write

$$d_t^{\text{iso}}(\mathcal{E}_2\circ\mathcal{D}_{\text{bi}},\mathcal{F}_2\circ\mathcal{D}_{\text{bi}}) = d_t\big[\big(\mathcal{D}_{\text{bi}}^{\intercal}\otimes\mathcal{E}_2\big)(|\Phi\rangle\langle\Phi|),\big(\mathcal{D}_{\text{bi}}^{\intercal}\otimes\mathcal{F}_2\big)(|\Phi\rangle\langle\Phi|)\big],$$

in which the map $\mathcal{D}_{\text{bi}}^{\intercal}$ is a trace-preserving operation. Then, it holds that

$$d_t^{\text{iso}}(\mathcal{E}_2\circ\mathcal{D}_{\text{bi}},\mathcal{F}_2\circ\mathcal{D}_{\text{bi}}) \leqslant d_t^{\text{iso}}(\mathcal{E}_2,\mathcal{F}_2).$$

To see the previous inequality it is necessary to consider that

$$\big(\mathcal{D}_{\text{bi}}^{\intercal}\otimes\mathcal{E}_2\big)(|\Phi\rangle\langle\Phi|) = \big(\mathcal{D}_{\text{bi}}^{\intercal}\otimes\mathbb{1}\big)(\mathbb{1}\otimes\mathcal{E}_2)(|\Phi\rangle\langle\Phi|),$$

and to apply contractivity. By following the same reasoning as before, we can prove that

$$d_t^{\text{iso}}(\mathcal{E}_2\circ\mathcal{E}_1,\mathcal{E}_2\circ\mathcal{D}_{\text{bi}}) \leqslant d_t^{\text{iso}}(\mathcal{E}_1,\mathcal{D}_{\text{bi}}).$$

Finally, we have shown that the right-hand side of Eq. (24) can be bounded by employing contractivity to both terms, leading to the desired result.

The postprocessing inequality [1,16] requires that

$$d_t^{\text{iso}}(\mathcal{R}\circ\mathcal{E},\mathcal{R}\circ\mathcal{F}) \;\leqslant\; d_t^{\text{iso}}(\mathcal{E},\mathcal{F}), \tag{26}$$

for arbitrary quantum maps $\mathcal{R}$, $\mathcal{E}$, and $\mathcal{F}$. The transmission distance d_t^{iso} satisfies this property, as it follows from the monotonicity of this distance.

Inequality (23) and postprocessing inequality (26) allow us to demonstrate the invariance with respect to arbitrary unitary operations $\mathcal{U}$ and $\mathcal{V}$:

$$d_t^{\text{iso}}(\mathcal{U}\circ\mathcal{E}\circ\mathcal{V},\;\mathcal{U}\circ\mathcal{F}\circ\mathcal{V}) = d_t^{\text{iso}}(\mathcal{E},\mathcal{F}). \tag{27}$$

Note that d_t is invariant under a post-transformation of $\mathcal{E}$ with $\mathcal{U}$,

$$d_t^{\text{iso}}(\mathcal{U}\circ\mathcal{E},\mathcal{U}\circ\mathcal{F}) = d_t^{\text{iso}}(\mathcal{E},\mathcal{F}),$$

because of the unitary invariance of the transmission distance in the state space. Thus, it remains to show the identity

$$d_t^{\text{iso}}(\mathcal{E}\circ\mathcal{V},\mathcal{F}\circ\mathcal{V}) = d_t^{\text{iso}}(\mathcal{E},\mathcal{F}). \tag{28}$$

The chaining property in this case states that

$$d_t^{\text{iso}}(\mathcal{E}\circ\mathcal{V},\mathcal{F}\circ\mathcal{V}) \leqslant d_t^{\text{iso}}(\mathcal{E},\mathcal{F}).$$

Simultaneously it holds that

$$d_t^{\text{iso}}(\mathcal{E},\mathcal{F}) = d_t^{\text{iso}}(\mathcal{E}_{\mathcal{V}}\circ\mathcal{V}^{-1},\mathcal{F}_{\mathcal{V}}\circ\mathcal{V}^{-1}) \leqslant d_t^{\text{iso}}(\mathcal{E}_{\mathcal{V}},\mathcal{F}_{\mathcal{V}}),$$

where $\mathcal{E}_{\mathcal{V}}=\mathcal{E}\circ\mathcal{V}$ and $\mathcal{F}_{\mathcal{V}}=\mathcal{F}\circ\mathcal{V}$. Therefore, we conclude that

$$d_t^{\text{iso}}(\mathcal{E},\mathcal{F}) \leqslant d_t^{\text{iso}}(\mathcal{E}_{\mathcal{V}},\mathcal{F}_{\mathcal{V}}) \leqslant d_t^{\text{iso}}(\mathcal{E},\mathcal{F}).$$

This implies Eq. (28) and completes the proof of the unitary invariance (27).

To establish bounds on the analyzed transmission distance d_t^{iso} we shall apply the Jamiołkowski isomorphism to extend the standard distance measures defined in the space of states into the space of maps [27]. The trace distance T, fidelity F, Bures distance D_B, and entropic distance D_E between any two maps read, respectively,

$$T(\mathcal{E},\mathcal{F}) = T(\rho_{\mathcal{E}},\rho_{\mathcal{F}}), \tag{29}$$

$$F(\mathcal{E},\mathcal{F}) = F(\rho_{\mathcal{E}},\rho_{\mathcal{F}}), \tag{30}$$

$$D_B(\mathcal{E},\mathcal{F}) = \sqrt{2-2\sqrt{F(\mathcal{E},\mathcal{F})}}, \tag{31}$$

$$D_E(\mathcal{E},\mathcal{F}) = \sqrt{H_2\{D_B^2(\mathcal{E},\mathcal{F})/4\}}. \tag{32}$$

Making use of inequalities (11) and (12) we arrive thus at the bounds relating the transmission distance with other measures:

$$\frac{T(\mathcal{E},\mathcal{F})}{2\sqrt{2}} \leqslant d_t^{\text{iso}}(\mathcal{E},\mathcal{F}) \leqslant \min\{\sqrt{T(\mathcal{E},\mathcal{F})},\;D_E(\mathcal{E},\mathcal{F})\}. \tag{33}$$

Further discussion of the upper bound is provided in Appendix B.

IV. ENTROPIC CHANNEL DIVERGENCE

Let us now explore another approach to introduce a distinguishability measure into the space of maps by using the transmission distance. The quantum Jensen-Shannon divergence plays a key role in quantum information theory as the maximal amount of classical information transmissible by means of quantum ensembles [28]. For a given quantum channel $\mathcal{E}$ one defines its Holevo capacity:

$$C_1(\mathcal{E}) = \max_{\Pi}\,\text{QJSD}_{\boldsymbol{p}}[\mathcal{E}(\rho_1),\ldots,\mathcal{E}(\rho_n)],$$

where the maximum is taken over all ensembles $\Pi=\{p_i,\rho_i\}_{i=1}^{n}$.

Consider now a different setup, in which a fixed state ρ is transformed by channel $\mathcal{E}_i$ with probability p_i. The associated Holevo information [18] reads

$$\mathcal{X}(\rho) = \text{QJSD}_{\boldsymbol{p}}[\mathcal{E}_1(\rho),\ldots,\mathcal{E}_n(\rho)]. \tag{34}$$

Taking two analyzed channels $\mathcal{E}_1$ and $\mathcal{E}_2$ with equal weights, $p_1=p_2=1/2$, we arrive at a *worst-case distance* measure between them:

$$d_t(\mathcal{E},\mathcal{F}) = \sup_{\rho\in\mathcal{M}_N}\sqrt{\text{QJSD}[\mathcal{E}(\rho),\mathcal{F}(\rho)]}. \tag{35}$$

Without loss of generality the supremum can be restricted to pure states [12].

In the above definition one analyzes directly the action of the channels $\mathcal{E}_i$ on the state ρ of size N. A more general approach involves extending the system by a K-dimensional ancilla [1,13] and studying the action of extended channels, $\mathcal{E}_i \otimes \mathbb{1}_K$. The *entropic channel divergence* reads

$$d_t^K(\mathcal{E},\mathcal{F}) = \sup_{\sigma\in\mathcal{M}_{NK}} d_t[(\mathcal{E}\otimes\mathbb{1}_K)(\sigma),(\mathcal{F}\otimes\mathbb{1}_K)(\sigma)], \quad (36)$$

where the state σ acts on an extended space of size NK. Observe that in the special case $K=1$ one has $d_t^{K=1}(\mathcal{E},\mathcal{F}) = d_t(\mathcal{E},\mathcal{F})$, as expected.

Properties of $d_t^K(\mathcal{E},\mathcal{F})$ and the chain rule

Let us discuss some key properties of the entropic channel divergence. By definition, for an arbitrary dimension K of the ancilla, the entropic channel divergence d_t^K is symmetric, is null if and only if the maps are equal, and satisfies the triangle inequality in the space of quantum channels. On the other hand, we have

$$d_t^K[\mathcal{E}\otimes\mathbb{1}_K,\mathcal{F}\otimes\mathbb{1}_K] \geqslant d_t[(\mathcal{E}\otimes\mathbb{1}_K)(\rho^*),(\mathcal{F}\otimes\mathbb{1}_K)(\rho^*)] \quad (37)$$

$$\geqslant d_t[\mathrm{Tr}_K\,(\mathcal{E}\otimes\mathbb{1}_K)(\rho^*),\mathrm{Tr}_K\,(\mathcal{F}\otimes\mathbb{1}_K)(\rho^*)] \quad (38)$$

$$= d_t[\mathcal{E}(\rho_Q^*),\mathcal{F}(\rho_Q^*)] = d_t(\mathcal{E},\mathcal{F}). \quad (39)$$

Inequality (37) holds by definition of d_t^K. The quantum state ρ^* is taken as $\rho^*=\rho_Q^*\otimes\rho_K$ where ρ_Q^* denotes the state which maximizes $d_t(\mathcal{E},\mathcal{F})$ [Eq. (35)] and ρ_K is an arbitrary state in $\mathcal{M}_K$. Equation (38) follows from the fact that the operation Tr_K defines a completely positive trace-preserving map and from applying contractivity. Finally, Eq. (39) holds because $\mathrm{Tr}_K(\mathcal{E}\otimes\mathbb{1}_K)(\rho^*)=\mathcal{E}(\rho_Q^*)$ and $\mathrm{Tr}_K(\mathcal{F}\otimes\mathbb{1}_K)(\rho^*)=\mathcal{F}(\rho_Q^*)$, which lead to Eq. (35) by the definition of ρ_Q^*.

In the same way, for any K' being a multiple of K, it is possible to show the following relation:

$$d_t^K(\mathcal{E},\mathcal{F}) \leqslant d_t^{K'}(\mathcal{E},\mathcal{F}).$$

This inequality suggests that d_t^K is in general not stable under the addition of ancillary systems. Furthermore, it was shown in Ref. [29] that if $K<N$ the channel divergence arising from the trace norm is in general not stable with respect to the tensor product. To ensure stability one supplies the requirement that the sizes of the ancilla and the principal systems are equal, $K=N$. It was demonstrated in Ref. [1] that for $K\geqslant N$ the following equality holds:

$$d_t^K(\mathcal{E},\mathcal{F}) = d_t^N(\mathcal{E},\mathcal{F}).$$

This implies that for $K=N$ the entropic channel divergence is stable under the addition of auxiliary subsystems:

$$d_t^N(\mathcal{E},\mathcal{F}) = d_t^N(\mathcal{E}\otimes\mathbb{1},\mathcal{F}\otimes\mathbb{1}).$$

As a result, it is natural to choose $K=N$ and in this paper the quantity d_t^N will be called *stabilized* entropic channel divergence.

The chaining property, postprocessing inequality and unitary invariance can be straightforwardly demonstrated by using the monotonicity and triangle inequality of the transmission distance in the state space [1].

Once $d_t^K(\mathcal{E},\mathcal{F})$ is defined, we can establish a *chain rule* for the entropic channel divergence, analogously to that obtained for the quantum relative entropy in Ref. [11]—this should not be confused with the chaining property discussed above.

Proposition 1. Let $\mathcal{E}$ and $\mathcal{F}$ denote arbitrary two operations acting over $\mathcal{M}_N$. For arbitrary bipartite quantum states ρ and σ in $\mathcal{M}_{NK}$ the following chain rule holds:

$$d_t[(\mathcal{E}\otimes\mathbb{1}_K)(\rho),(\mathcal{F}\otimes\mathbb{1}_K)(\sigma)] \leqslant d_t(\rho,\sigma) + d_t^K(\mathcal{E},\mathcal{F}). \quad (40)$$

It relates the transmission distance $d_t(\cdot,\cdot)$ between quantum states, defined in (9), and the entropic channel divergence $d_t^K(\cdot,\cdot)$ introduced in Eq. (36).

Proof. It will be convenient to use a simpler notation and write $\mathcal{E}_{NK}(\rho)$ instead of $(\mathcal{E}\otimes\mathbb{1}_K)(\rho)$ for a quantum operation $\mathcal{E}$ acting on $\mathcal{M}_N$. Using this convention, we have

$$\begin{aligned} d_t[\mathcal{E}_{NK}(\rho),\mathcal{F}_{NK}(\sigma)] &\leqslant d_t[\mathcal{E}_{NK}(\rho),\mathcal{E}_{NK}(\sigma)] \\ &\quad + d_t[\mathcal{E}_{NK}(\sigma),\mathcal{F}_{NK}(\sigma)] \\ &\leqslant d_t(\rho,\sigma) + d_t[\mathcal{E}_{NK}(\sigma),\mathcal{F}_{NK}(\sigma)] \\ &\leqslant d_t(\rho,\sigma) + d_t^K(\mathcal{E},\mathcal{F}), \end{aligned} \quad (41)$$

in which we have employed the triangle inequality and the monotonicity of the transmission distance. ∎

Note that the chain rule (40) is valid not only for the stabilized version of the entropic channel divergence but also for the original version (35) and the maps applied directly over the states describing the principal N-dimensional system.

The chain rule (40) has interesting applications in the context of hypothesis testing in quantum channel discrimination [11], due to its connection with the *amortized channel divergence*, introduced in Ref. [12] for an arbitrary generalized divergence $d(\cdot,\cdot)$. By using the transmission distance, we obtain the amortized entropic divergence,

$$d_t^A(\mathcal{E},\mathcal{F}) = \sup_{\rho,\sigma\in\mathcal{M}_{NK}} \{d_t[\mathcal{E}_{NK}(\rho),\mathcal{F}_{NK}(\sigma)] - d_t(\rho,\sigma)\}, \quad (42)$$

which depends on the size K of the ancilla. Note that the chain rule (40) establishes an upper bound for $d_t^A(\mathcal{E},\mathcal{F})$. A lower bound,

$$d_t^A(\mathcal{E},\mathcal{F}) \geqslant d_t^K(\mathcal{E},\mathcal{F}),$$

was shown [12] to hold for an arbitrary distance measure $d(\cdot,\cdot)$. We arrive therefore at the *amortization collapse* of the entropic channel divergence, that is,

$$d_t^A(\mathcal{E},\mathcal{F}) = d_t^K(\mathcal{E},\mathcal{F}). \quad (43)$$

V. PHYSICAL INTERPRETATION

We defined the transmission distance (16) between quantum channels, and the entropic channel divergence (36), and will now discuss their physical meaning.

A. Transmission distance between quantum channels

The transmission distance between quantum channels is *easy* to compute, as its definition does not require any optimization procedure. The calculations are reduced to evaluation of the entropy of a map [27], equal to the von Neumann entropy of the corresponding Choi states. Furthermore, it is possible to estimate experimentally this quantity, since its definition involves the Choi states, which can be obtained by quantum process tomography [1].

Observe that $[d_t^{\mathrm{iso}}(\mathcal{E},\mathcal{F})]^2$ is the Holevo information corresponding to an equiprobable ensemble composed by the states $\rho_\mathcal{E}$ and $\rho_\mathcal{F}$. Additionally, for general discrete ensembles, $[d_t^{\mathrm{iso}}(\mathcal{E},\mathcal{F})]^2$ is connected to the protocol of *dense coding*. Consider a bipartite quantum system in a maximally entangled state, $\rho_r = |\Phi\rangle\langle\Phi|$, usually known as a *resource state*, subjected to local unitary transformations $\mathcal{U}_i$ performed with probability p_i. The output state

$$\rho_i^U = (U_i \otimes \mathbb{1})\rho_r(U_i^\dagger \otimes \mathbb{1}) \tag{44}$$

occurs with probability p_i. This protocol, relying on the initial entanglement between both parties, allows them to transmit classical information encoded in a bipartite system, while conducting operations on a single subsystem only. If the dimension of each subsystem is N, it is possible to send $2\log_2 N$ bits of classical information, even though the classical coding allows one to send only $\log_2 N$ bits.

The capacity of the dense coding protocol with resource ρ_r to transmit classical information for fixed unitary operations U_i, is given [30] by the maximum over $\{p_i\}_i$ of $\mathrm{QJSD}_{\boldsymbol{p}}(\rho_1^U,\ldots,\rho_n^U)$. Since ρ_i^U forms Choi matrices of unitary channels, $\mathcal{U}_i$, the divergence $[d_t^{\mathrm{iso}}(\mathcal{U}_1,\mathcal{U}_2)]^2$ coincides with the capacity of the coding with two unitary operations and equal probabilities, $p_i = 1/2$.

Therefore, $[d_t^{\mathrm{iso}}(\mathcal{E},\mathcal{F})]^2$ is the dense coding capacity connected to maps $\mathcal{E}$ and $\mathcal{F}$, for a noiseless protocol with a maximally entangled resource state ρ_r. Distinguishability of quantum maps using quantum dense coding protocol was advocated by Raginsky [16], who analyzed an analogous measure based on the quantum fidelity instead of the quantum Jensen-Shannon divergence.

B. Entropic channel divergence

One of the fundamental challenges in quantum information theory is the discrimination of quantum channels, belonging to a discrete ensemble of quantum maps with a certain *a priori* probability distribution p_i. To accomplish this task, input states are utilized along with measurements on their corresponding outputs. In Refs. [31,32], quantum channel discrimination was studied specifically in the context of reading digital memories (composed by *cells*), a topic which laid the groundwork for the treatment of digital memories within the field of quantum information.

The storage of data in a digital memory is achieved through a channel encoding protocol, where information is stored in a cell by employing a quantum channel selected from the fixed ensemble [32]. On the other hand, the readout process corresponds to channel decoding, which is equivalent to discriminating between the quantum channels that constitute the ensemble. This decoding operation occurs when a decoder retrieves information from the memory cells. The quantum reading capacity refers to the maximum amount of information that can be read (or decoded) per cell. In the case of a given set of quantum operations denoted as $\mathcal{E}_i$, each associated with a probability p_i, the quantum reading capacity of the most general strategy of readout (parallel discrimination of quantum channels in multicell encoding) is [10,31]

$$\sup_{\rho\in\mathcal{M}_{N\times K}} \mathrm{QJSD}_{\boldsymbol{p}}[(\mathcal{E}_1\otimes\mathbb{1}_K)(\rho),\ldots,(\mathcal{E}_n\otimes\mathbb{1}_K)(\rho)].$$

The entropic channel divergence, Eq. (36), is therefore the square root of the previous quantity in the case of an ensemble of two maps, with $p_i = 1/2$.

Furthermore, a study conducted by Laurenza *et al.* in Ref. [30] revealed a relationship between the one-shot capacity of a dense coding protocol and the quantum reading capacity, specifically when employing an arbitrary resource state denoted as ρ_r. The authors demonstrated that the capacity of the dense coding protocol can be expressed in terms of the quantum reading capacity. This finding establishes a conceptual link between the entropic channel divergence, which is associated with the quantum reading capacity, and the transmission distance between quantum channels, which was discussed in the preceding section, and is associated with the dense coding capacity.

C. Relation between the channel divergence d_t^N and the transmission distance d_t^{iso}

Assume that the single-qubit channels we wish to distinguish are covariant with respect to Pauli operators. This means that for each quantum channel $\mathcal{E}$ we can write $\mathcal{E}\circ\mathcal{P} = \mathcal{P}'\circ\mathcal{E}$, where $\mathcal{P}$ and $\mathcal{P}'$ denote Pauli channels. In this case, the channel $\mathcal{E}$ can be simulated via local operations and classical communication (LOCC) operations [30,33], and it is called *Choi stretchable*: Specifically, there exists a LOCC operation $\mathcal{T}$ such that we can write

$$\mathcal{E}(\rho) = \mathcal{T}(\rho\otimes\rho_\mathcal{E}), \tag{45}$$

for all quantum states ρ in the domain of $\mathcal{E}$, and $\rho_\mathcal{E}$ being the corresponding Choi state of the map $\mathcal{E}$. Thus, as $\mathcal{T}$ is a completely positive trace-preserving LOCC operation independent of $\mathcal{E}$ (it can be taken as the standard quantum teleportation protocol) [33], for any two Choi-stretchable quantum operations $\mathcal{E}_1$ and $\mathcal{E}_2$, we have

$$\begin{aligned} d_t^N(\mathcal{E}_1,\mathcal{E}_2) &= \sup_{\rho\in\mathcal{M}_{N\times N}} d_t[(\mathcal{E}_1\otimes\mathbb{1})(\rho),(\mathcal{E}_2\otimes\mathbb{1})(\rho)] \\ &= \sup_{\rho\in\mathcal{M}_{N\times N}} d_t[\mathcal{T}(\rho\otimes\rho_{\mathcal{E}_1}),\mathcal{T}(\rho\otimes\rho_{\mathcal{E}_2})] \\ &\leqslant \sup_{\rho\in\mathcal{M}_{N\times N}} d_t(\rho\otimes\rho_{\mathcal{E}_1},\rho\otimes\rho_{\mathcal{E}_2}) = d_t^{\mathrm{iso}}(\mathcal{E}_1,\mathcal{E}_2). \end{aligned}$$

We applied here the subadditivity of the QJSD in the state space and its monotonicity under completely positive maps. By definition of d_t^N, the inequality $d_t^{\mathrm{iso}}(\mathcal{E}_1,\mathcal{E}_2)\leqslant d_t^N(\mathcal{E}_1,\mathcal{E}_2)$ holds. Thus the equality

$$d_t^{\mathrm{iso}}(\mathcal{E}_1,\mathcal{E}_2) = d_t^N(\mathcal{E}_1,\mathcal{E}_2) \tag{46}$$

is valid for any two Pauli covariant operations $\mathcal{E}_1$ and $\mathcal{E}_2$.

VI. APPLICATIONS

In this section, we explore certain features of the distinguishability measures between quantum operations proposed in Secs. III and IV. We analyze two particular single-qubit problems: distinguishing two unitary Pauli operations and two Hamiltonian evolutions under decoherence.

The three-dimensional Bloch vector $\boldsymbol{r}$ of a single-qubit state allows us to represent the density matrix as

$$\rho = \tfrac{1}{2}(\mathbb{1} + \boldsymbol{r}\cdot\boldsymbol{\sigma}). \tag{47}$$

Here $\boldsymbol{r}\cdot\boldsymbol{\sigma} = \sum_{i=1}^{3} r_i\sigma_i$ with $\{\sigma_i\}_i$ denoting three Pauli matrices. The action of a quantum operation $\mathcal{E}$ over ρ can be described by a distortion matrix $\Lambda_{\mathcal{E}}$ and a translation vector $\boldsymbol{l}_{\mathcal{E}}$:

$$\mathcal{E}(\rho) = \tfrac{1}{2}(\mathbb{1} + \boldsymbol{r}_{\mathcal{E}}\cdot\boldsymbol{\sigma}) \quad \text{with} \quad \boldsymbol{r}_{\mathcal{E}} = \Lambda_{\mathcal{E}}\boldsymbol{r} + \boldsymbol{l}_{\mathcal{E}}. \tag{48}$$

The above form is called the affine decomposition or the Fano representation of the map.

A. Pauli channels

All single-qubit unital operations belong to the class of Pauli channels:

$$\mathcal{P}_p(\rho) = \sum_{\alpha=0}^{3} p_\alpha \sigma_\alpha \rho \sigma_\alpha, \tag{49}$$

where $\{\sigma_\alpha\}_{\alpha=0}^3 = \{\mathbb{1}, \boldsymbol{\sigma}\}$ and $\{p_\alpha\}_{\alpha=0}^3$ is a discrete probability vector. The Fano form of such a map $\mathcal{P}$ reads

$$\boldsymbol{l}_{\mathcal{P}} = \boldsymbol{0},\ \Lambda_{\mathcal{P}} = \mathrm{diag}(c_1, c_2, c_3) = \sum_{\alpha=0}^{3} p_\alpha R_\alpha, \tag{50}$$

with

$$\begin{aligned} R_0 &= \mathrm{diag}(1, 1, 1),\\ R_1 &= \mathrm{diag}(1, -1, -1),\\ R_2 &= \mathrm{diag}(-1, 1, -1),\\ R_3 &= \mathrm{diag}(-1, -1, 1). \end{aligned} \tag{51}$$

Thus, R_α is a diagonal orthogonal matrix defined by the action of the unitary transformations given by the Pauli matrix σ_α and R_0 is connected to the identity map. Additionally, the set $\boldsymbol{c} = (c_1, c_2, c_3)$, in Eq. (50), for which $\mathcal{P}$ is a well-defined CPTP map, specifies a tetrahedron in the three-dimensional space [34], with edges $\{R_\alpha\}_{\alpha=0}^3$ (see Fig. 1). The relation among $\{p_\alpha\}_{\alpha=0}^3$ and the numbers $\{c_i\}_{i=1}^3$ is

$$\begin{aligned} p_0 &= \tfrac{1}{4}(1 + c_1 + c_2 + c_3),\\ p_1 &= \tfrac{1}{4}(1 + c_1 - c_2 - c_3),\\ p_2 &= \tfrac{1}{4}(1 - c_1 + c_2 - c_3),\\ p_3 &= \tfrac{1}{4}(1 - c_1 - c_2 + c_3). \end{aligned} \tag{52}$$

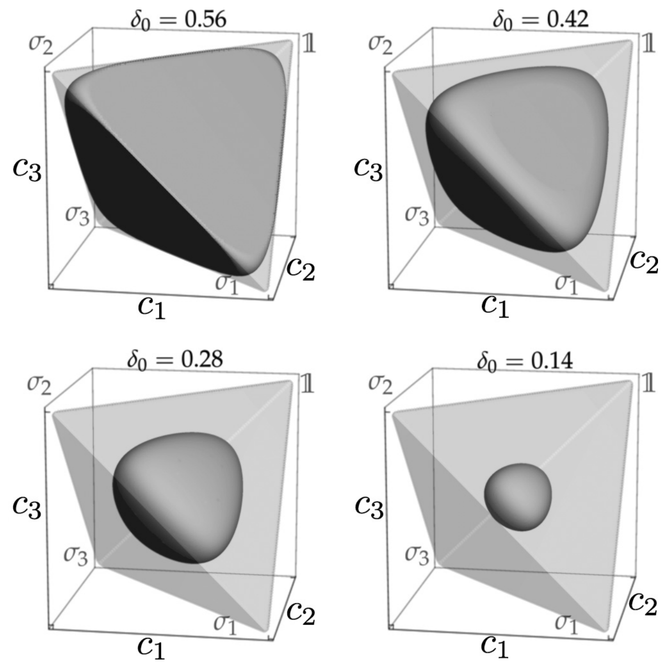

FIG. 1. The tetrahedron of Pauli channels with "spheres" of channels equidistant to the completely depolarizing channel $\mathcal{D}_0$ in the center of the tetrahedron, with respect to the distance $d_t^{\mathrm{iso}}(\mathcal{P}_p, \mathcal{D}_0) = \delta_0$, for radii $\delta_0 \in \{0.56, 0.42, 0.28, 0.14\}$. All depicted quantities are dimensionless.

Particular examples of Pauli maps are the identity, the phase-flip channel $\mathcal{P}_{\mathrm{PF}}$, and the depolarizing map $\mathcal{D}$, corresponding to the distortion matrices

$$\Lambda_{\mathcal{I}} = \mathrm{diag}(1, 1, 1), \tag{53}$$

$$\Lambda_{\mathcal{P}_{\mathrm{PF}}} = \mathrm{diag}(1 - x, 1 - x, 1), \tag{54}$$

$$\Lambda_{\mathcal{D}} = \mathrm{diag}(1 - x, 1 - x, 1 - x), \tag{55}$$

respectively. Completely depolarizing channel $\mathcal{D}_0$ corresponds to Eq. (55) with $x = 1$.

For an arbitrary channel $\mathcal{E}$, the distortion matrix $\Lambda_{\mathcal{E}}$ can be diagonalized by applying local unitary transformations on $\mathcal{E}(\rho)$, reaching the canonical form of the map, which is subsequently given by the translation vector $\boldsymbol{t}_{\mathcal{E}} = (t_1, t_2, t_3)$ and the distortion vector $\boldsymbol{\omega}_{\mathcal{E}} = (\omega_1, \omega_2, \omega_3)$, which results from the diagonalization of $\Lambda_{\mathcal{E}}$ [26,35]. Note that the canonical form of a given unital map, $\boldsymbol{t}_{\mathcal{E}} = \boldsymbol{0}$, gives a Pauli channel (49).

The Choi matrix (14) of any single-qubit channel in its canonical form reads [26]

$$\rho_{\mathcal{E}} = \frac{1}{4}\begin{bmatrix} 1 + \omega_3 + t_3 & 0 & t_1 + i\omega_2 & \omega_1 + \omega_2 \\ 0 & 1 - \omega_3 + t_3 & \omega_1 - \omega_2 & t_1 + i\omega_2 \\ t_1 - i\omega_2 & \omega_1 - \omega_2 & 1 - \omega_3 - t_3 & 0 \\ \omega_1 + \omega_2 & t_1 - i\omega_2 & 0 & 1 + \omega_3 - t_3 \end{bmatrix}.$$

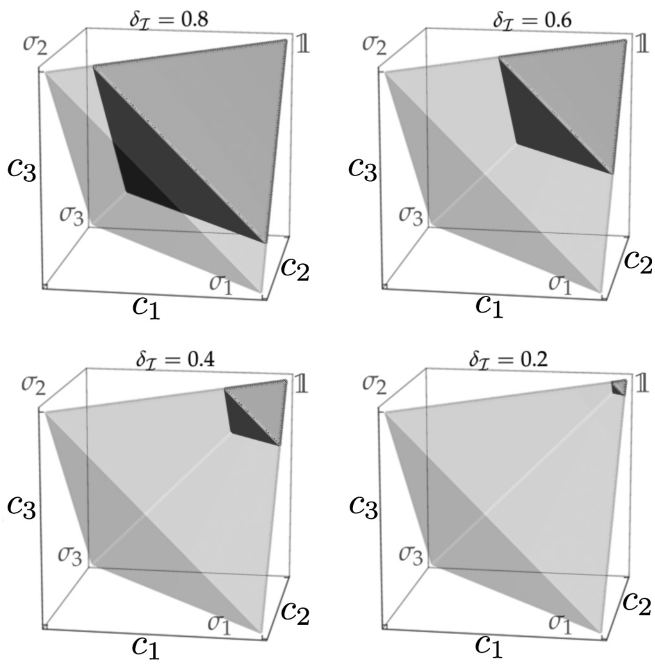

FIG. 2. Surfaces within the Pauli tetrahedron, defined by a constant transmission distance to the identity map $\mathcal{I}$ represented by the corner of the set, $d_t^{\mathrm{iso}}(\mathcal{P}_p, \mathcal{I}) = \delta_{\mathcal{I}}$, for $\delta_{\mathcal{I}} \in \{0.8, 0.6, 0.4, 0.2\}$. All depicted quantities are dimensionless.

If $\boldsymbol{t}_{\mathcal{E}} = \boldsymbol{0}$, $\rho_{\mathcal{E}}$ forms a Bell-diagonal state (i.e., its eigenvectors are the four Bell states) and its eigenvalues are given by the probabilities p_α appearing in (52). Let us analyze the transmission distance between maps, Eq. (16), and the entropic channel divergence, Eq. (36), for $K = 1$ and $K = 2$ (stabilized version).

1. Transmission distance between Pauli channels

Let $\mathcal{P}_p$ and $\mathcal{P}_q$ be two Pauli channels defined by two probability distributions $\{p_\alpha\}_{\alpha=0}^3$ and $\{q_\beta\}_{\beta=0}^3$, as in Eq. (49). The corresponding Choi matrices of these maps become diagonal in the Bell basis. The quantum Jensen-Shannon divergence between two Pauli channels is therefore equal to the classical Jensen-Shannon divergence evaluated in the classical tetrahedron of four-point probability distributions, determined by the spectra of both Choi states, $p = \{p_\alpha\}_{\alpha=0}^3$ and $q = \{q_\beta\}_{\beta=0}^3$:

$$d_t^{\mathrm{iso}}(\mathcal{P}_p, \mathcal{P}_q) = \sqrt{D_{\mathrm{JS}}^c(\boldsymbol{p}||\boldsymbol{q})}. \quad (56)$$

Using the three-dimensional parametrization in (52), we can plot the surface, within the Pauli tetrahedron, defined by those maps with the same transmission distance to the center of the tetrahedron, which represents the completely depolarizing map $\mathcal{D}_0$:

$$d_t^{\mathrm{iso}}(\mathcal{P}_p, \mathcal{D}_0) = \delta_0. \quad (57)$$

In Fig. 1, such "spheres" with respect to this distance are plotted for four different radii. For a small radius δ_0 such a surface resembles a sphere, while for larger values of δ_0 it becomes deformed by the faces of the tetrahedron.

Analogously, Fig. 2 presents four spheres corresponding to the fixed transmission distance to the identity map, $d_t^{\mathrm{iso}}(\mathcal{P}_p, \mathcal{I}) = \delta_{\mathcal{I}}$, with radii $\delta_{\mathcal{I}}$ listed in the caption.

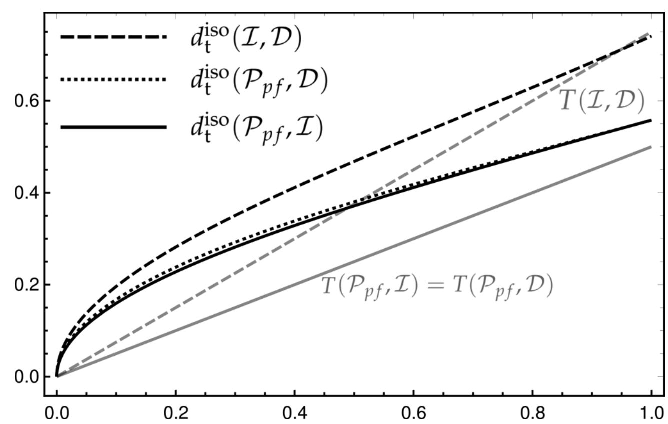

FIG. 3. Phase-flip noise teleportation: Transmission distance d_t^{iso} defined in (16) between the identity map, phase-flip channel, and depolarizing channel [(53)–(55), respectively] as functions of the depolarizing parameter x. For comparison we plot also the trace distance T between the corresponding Choi states [see (29)] and the entropic channel divergence $d_t^{K=1}$ [see (35)]. All quantities shown have no dimensions.

In Fig. 3, we plot the transmission distance between the maps given by (53)–(55), as functions of the depolarizing parameter $x \in [0, 1]$, and the trace distance between the corresponding Choi states. For $x \neq 0$, we observe that

$$d_t^{\mathrm{iso}}(\mathcal{P}_{\mathrm{PF}}, \mathcal{I}) < d_t^{\mathrm{iso}}(\mathcal{P}_{\mathrm{PF}}, \mathcal{D}) < d_t^{\mathrm{iso}}(\mathcal{I}, \mathcal{D}), \quad (58)$$

while for the trace distance (29) the following relations hold:

$$T(\mathcal{P}_{\mathrm{PF}}, \mathcal{D}) = T(\mathcal{P}_{\mathrm{PF}}, \mathcal{I}) < T(\mathcal{I}, \mathcal{D}). \quad (59)$$

2. Entropic channel divergence

Let us calculate the entropic channel divergence (36) for two Pauli channels $\mathcal{P}_p$ and $\mathcal{P}_q$ corresponding to probability distributions p and q, determined by the vectors $\boldsymbol{c}_p = (c_{p1}, c_{p2}, c_{p3})$ and $\boldsymbol{c}_q = (c_{q1}, c_{q2}, c_{q3})$, respectively. For $N = 2$, there are two different entropic divergence measures labeled by the dimension K of the ancilla,

$$d_t^{K=1}(\mathcal{E}, \mathcal{F}) \quad \text{and} \quad d_t^{K=2}(\mathcal{E}, \mathcal{F}),$$

since $d_t^{K'}(\mathcal{E}, \mathcal{F}) = d_t^{K=2}(\mathcal{E}, \mathcal{F})$ for $K' > 2$, as mentioned before. The Pauli channels are Pauli covariant (45), which implies that $d_t^{K=2}(\mathcal{P}_p, \mathcal{P}_q) = d_t^{\mathrm{iso}}(\mathcal{P}_p, \mathcal{P}_q)$ (see Sec. V C).

In the case $K = 1$, one has to optimize the transmission distance between the channels over the initial pure states:

$$d_t^{K=1}(\mathcal{P}_p, \mathcal{P}_q) = \sup_{\rho \in \mathcal{M}_N} \sqrt{\mathrm{QJSD}[\mathcal{P}_p(\rho), \mathcal{P}_q(\rho)]},$$

where

$$\mathrm{QJSD}[\mathcal{P}_p(\rho), \mathcal{P}_q(\rho)] = S[\overline{\mathcal{P}}(\rho)] - \tfrac{1}{2}S[\mathcal{P}_p(\rho)] - \tfrac{1}{2}S[\mathcal{P}_q(\rho)],$$

and $\overline{\mathcal{P}} = (\mathcal{P}_p + \mathcal{P}_q)/2$ is the average channel, which also forms a Pauli map.

Proposition 2. Entropic channel divergence (35), between two Pauli maps $\mathcal{P}_p$ and $\mathcal{P}_q$, given by distortion matrices $\Lambda_p =$

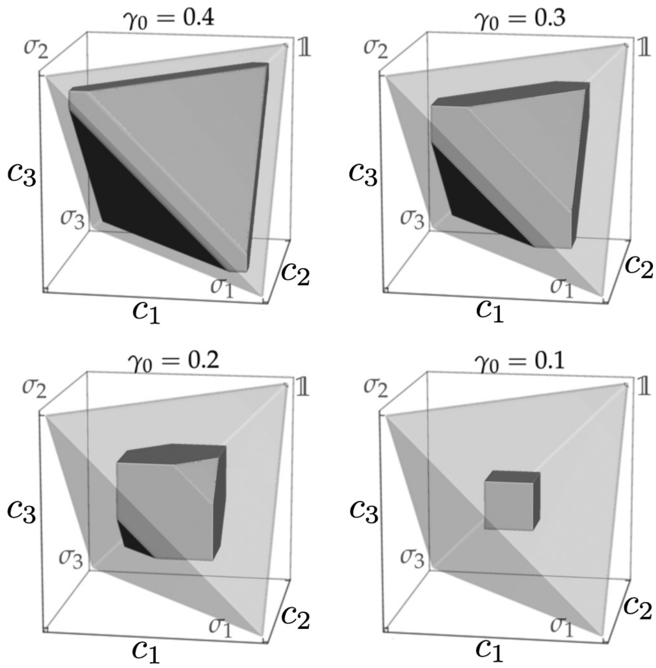

FIG. 4. Spheres with respect to the distance $d_t^{K=1}$ within the tetrahedron of Pauli channels, $d_t^{K=1}(\mathcal{P}_p, \mathcal{D}) = \gamma_0$ for four different radii: $\gamma_0 \in \{0.4, 0.3, 0.2, 0.1\}$. All quantities shown have no dimensions.

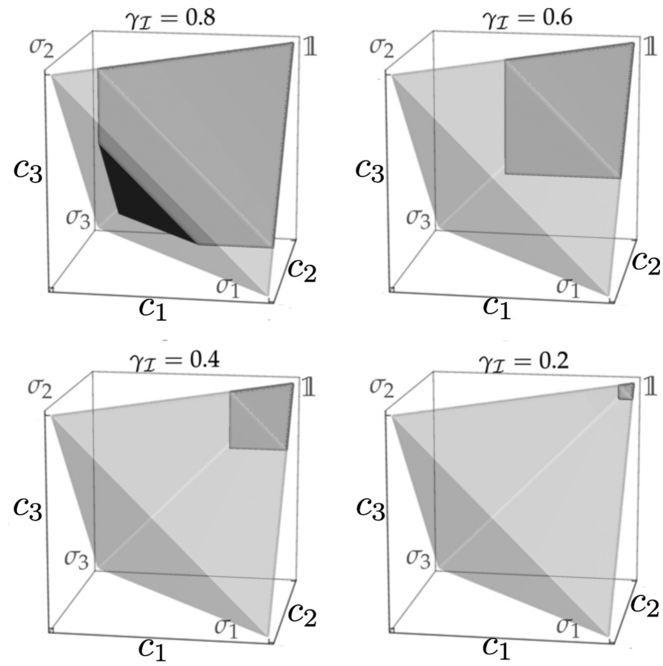

FIG. 5. Surfaces defined by constant entropic channel divergence to the identity map: $d_t^{K=1}(\mathcal{P}_p, \mathcal{I}) = \gamma_\mathcal{I}$ with $\gamma_\mathcal{I} \in \{0.8, 0.6, 0.4, 0.2\}$. All quantities depicted are unitless.

(c_{p1}, c_{p2}, c_{p3}) and $\Lambda_q = (c_{q1}, c_{q2}, c_{q3})$, takes the form

$$d_t^{K=1}(\mathcal{P}_p, \mathcal{P}_q) = \max_i \sqrt{f(\bar{c}_i^2) - \frac{1}{2}\left[f\left(c_{pi}^2\right) + f\left(c_{qi}^2\right)\right]}, \quad (60)$$

where $\bar{c}_i = (c_{pi} + c_{qi})/2$ and

$$f(x) := H_2\left(\frac{1-\sqrt{x}}{2}\right). \quad (61)$$

Here $H_2(x) := -x\log_2 x - (1-x)\log_2(1-x)$ stands for the binary entropy function for $x \in [0, 1]$.

Proof. For an arbitrary Pauli map $S[\mathcal{P}(\rho)]$, we have

$$S[\mathcal{P}(\rho)] = f\left(r_\mathcal{P}^2\right), \quad (62)$$

where $r_\mathcal{P} = |\boldsymbol{r}_\mathcal{P}|$, $\boldsymbol{r}_\mathcal{P} = \Lambda_\mathcal{P}\boldsymbol{r}$ being the Bloch vector of $\mathcal{P}(\rho)$, Eq. (50). Thus, we can write $S[\mathcal{P}(\rho)] = f(r_\mathcal{P}^2) = f(\boldsymbol{r}\cdot\Lambda_\mathcal{P}^2\boldsymbol{r})$ and $\boldsymbol{r}^2 = 1$. Once we have rewritten the entropies of the Pauli channels, the quantum Jensen-Shannon divergence reads

$$\begin{aligned}\mathrm{QJSD}[\mathcal{P}_p(\rho), \mathcal{P}_q(\rho)] = {} & f\left(\boldsymbol{r}\cdot\Lambda_{\bar{\mathcal{P}}}^2\boldsymbol{r}\right) - \tfrac{1}{2}f\left(\boldsymbol{r}\cdot\Lambda_p^2\boldsymbol{r}\right) \\ & - \tfrac{1}{2}f\left(\boldsymbol{r}\cdot\Lambda_q^2\boldsymbol{r}\right). \end{aligned} \quad (63)$$

Let us apply the method of Lagrange multipliers to the Cartesian coordinates of $\boldsymbol{r}$. This leads to the following three equations:

$$\lambda r_i = \left[f'\left(\boldsymbol{r}\cdot\Lambda_{\bar{\mathcal{P}}}^2\boldsymbol{r}\right)\bar{c}_i^2 - \frac{\left(f'\left(\boldsymbol{r}\cdot\Lambda_p^2\boldsymbol{r}\right)c_{pi}^2 + f'\left(\boldsymbol{r}\cdot\Lambda_q^2\boldsymbol{r}\right)c_{qi}^2\right)}{2}\right] r_i,$$

with $i = 1, 2, 3$, which hold simultaneously with the constraint $\boldsymbol{r}^2 = 1$, associated to the Lagrange multiplier λ. Thus, the previous equation defines six possible extreme values of the function (63):

$$\boldsymbol{r}_{\mathrm{opt},1}^{\pm} = \pm(1, 0, 0) = \pm\boldsymbol{r}_1, \quad (64)$$

$$\boldsymbol{r}_{\mathrm{opt},2}^{\pm} = \pm(0, 1, 0) = \pm\boldsymbol{r}_2, \quad (65)$$

$$\boldsymbol{r}_{\mathrm{opt},3}^{\pm} = \pm(0, 0, 1) = \pm\boldsymbol{r}_3. \quad (66)$$

As Eq. (63) is symmetric under reflection, $\boldsymbol{r}' = -\boldsymbol{r}$, we have only three extremes that lead to different values of the QJSD. Correspondingly, the maximum is determined by Eq. (60). ■

In Fig. 4, we plot the three-dimensional spheres within the tetrahedron of Pauli channels such that

$$d_t^{K=1}(\mathcal{P}_p, \mathcal{D}_0) = \gamma_0$$

for four different radii. Analogously, Fig. 5 shows surfaces of maps of the same entropic channel divergence to the identity map, $d_t^{K=1}(\mathcal{P}_p, \mathcal{I}) = \gamma_\mathcal{I}$, for four exemplary values of $\gamma_\mathcal{I}$.

Consider now the distinguishability between the identity map, the phase-flip channel, and the depolarizing channel, specified in (53)–(55), respectively. In this case, for any $x \in [0, 1]$ the following inequalities hold:

$$\begin{aligned} d_t^{K=1}(\mathcal{P}_{\mathrm{PF}}, \mathcal{I}) = d_t^{K=1}(\mathcal{P}_{\mathrm{PF}}, \mathcal{D}) = d_t^{K=1}(\mathcal{I}, \mathcal{D}) \\ = d_t^{\mathrm{iso}}(\mathcal{P}_{\mathrm{PF}}, \mathcal{I}). \end{aligned} \quad (67)$$

Figure 3 shows the dependence of this distance on the depolarizing parameter x.

A similar behavior can be obtained for the distinguishability measures arising from the channel divergence based on the trace distance:

$$d_{\mathrm{Tr}}^{K=1}(\mathcal{E}, \mathcal{F}) = \sup_{\rho\in\mathcal{M}_N} T[\mathcal{E}(\rho), \mathcal{F}(\rho)]. \quad (68)$$

Proposition 3. Let $\mathcal{N}$ and $\mathcal{M}$ be two unital operations for $N = 2$. Then,

$$d_{\mathrm{Tr}}^{K=1}(\mathcal{N}, \mathcal{M}) = \frac{1}{2} \max_i \sqrt{\lambda_i^{\Delta}}, \tag{69}$$

where $\{\lambda_i^{\Delta}\}_i$ is the set of eigenvalues of the matrix

$$\Delta = (\Lambda_{\mathcal{N}} - \Lambda_{\mathcal{M}})^{\intercal}(\Lambda_{\mathcal{N}} - \Lambda_{\mathcal{M}}).$$

A proof of this result is provided in Appendix A. The reasoning presented above implies that

$$\begin{aligned} d_{\mathrm{Tr}}^{K=1}(\mathcal{P}_{\mathrm{PF}}, \mathcal{I}) = d_{\mathrm{Tr}}^{K=1}(\mathcal{P}_{\mathrm{PF}}, \mathcal{D}) = d_{\mathrm{Tr}}^{K=1}(\mathcal{I}, \mathcal{D}) \\ = T(\mathcal{P}_{\mathrm{PF}}, \mathcal{I}). \end{aligned} \tag{70}$$

Dependence of this function on the depolarizing parameter x is also marked in Fig. 3.

3. Noise in the quantum teleportation protocol

Quantum teleportation, one of the most important quantum information protocols, replicates the state of one quantum system into another without having information about the input state. This protocol requires three qubits which are operated by two different entities, usually referred to as Alice and Bob.

The corresponding tasks to teleport the qubit state ρ_a of Alice to Bob, assuming they share a two-qubit state AB in the maximally entangled Bell state $|\Psi\rangle_{AB}$, are the following.

(1) Alice measures a projection onto the Bell basis for the qubits aA and classically communicates its outcome to Bob.

(2) Bob applies suitable unitary operations, according to the shared measurement result, on his qubit B, to replicate the initial input state ρ_a of Alice.

Such a teleportation protocol is called *perfect* and it can be described by the identity channel, $\mathcal{I}_{a\to B}$, with the distortion matrix given by (53), where the subindex $a \to B$ denotes that the channel takes states of qubit a and returns the states of qubit B. However, the maximally entangled state $|\Psi\rangle_{AB}$, preshared by Alice and Bob, can be affected by noise or decoherence. The *standard* teleportation protocol consists of the above steps, but instead of assuming preshared maximally entanglement between the qubits AB one replaces it by a *resource state*:

$$|\Psi\rangle\langle\Psi| \to \rho_{AB}.$$

The quantum channel between qubits a and B defined by this protocol can always be written as a Pauli channel $\mathcal{P}_{a\to B}$, Eq. (49), regardless of which resource state is employed ρ_{AB} [14].

One paradigmatic type of decoherence over $|\Psi\rangle\langle\Psi|$, which is widely studied in the related literature, is given by a Werner state:

$$\rho_{AB} = (1-x)|\Psi\rangle\langle\Psi| + x\frac{\mathbb{1}_A \otimes \mathbb{1}_B}{4},$$

x being its decoherence parameter.

Therefore, this protocol is described by a depolarizing channel $\mathcal{D}_{a\to B}$ with distortion matrix equal to $\Lambda_{\mathcal{D}}$, Eq. (55). Another type of decoherence on $|\Psi\rangle\langle\Psi|$ leads to a teleportation channel described by the phase-flip channel, with distortion matrix given by (54). This protocol will be called *phase-flip noise* teleportation. Hence, Eq. (53) describes the perfect teleportation protocol, while Eqs. (54) and (55) are two different teleportation protocols that consider noise or decoherence affecting their resource state.

Figure 3 shows that for any decoherence parameter x the transmission distance d_t^{iso} between the perfect and the standard teleportation protocols with a Werner state as a resource is greater than the distances to the phase-flip noise teleportation.

An analogous property holds also for the trace distance. In the case of the entropic channel divergence for $K = 1$, the distance between the three different channels is equal [see Eq. (67)], similar to the case of the trace distance, Eq. (70).

The surfaces in Figs. 2 and 5 can be interpreted now as the standard teleportation protocols equally distant to the perfect one, represented by the vertex $c_1 = c_2 = c_3 = 1$. The transmission distance d_t^{iso} between quantum channels is more restrictive regarding the values of the parameters c_i, than the entropic channel divergence and $d_t^{K=1}$, which allows lower values for c_i.

B. Distinguishing operations determined by Hamiltonians

Several applications of quantum information theory involve the problem of distinguishing a particular Hamiltonian from a given set, for instance, to determine errors which occur by a real-life realization of certain information processing tasks. Other examples include identification of a classical static force acting on a given quantum system [15,36]. Consider the distinguishability between two Hamiltonians H_1 and H_2, acting on a two-dimensional Hilbert space.

Since three Pauli matrices, extended by the identity matrix, $\{\mathbb{1}, \boldsymbol{\sigma}\}$, form a Hilbert-Schmidt basis in the space of Hermitian matrices of order 2, any single-qubit Hamiltonian can represented by its Bloch vector:

$$H_m = h_m^0 \mathbb{1} + \boldsymbol{h}_m \cdot \boldsymbol{\sigma}. \tag{71}$$

The noiseless evolution of the state generated by a given Hamiltonian can be described by a unitary transformation, $\mathcal{U}_m(\rho) = U_m \rho U_m^{\intercal}$, with

$$U_m = e^{-itH_m} = e^{-ith_m^0}(\cos t\,\mathbb{1} - i \sin t\,\boldsymbol{h}_m \cdot \boldsymbol{\sigma}) \tag{72}$$

where $\boldsymbol{h}_m \cdot \boldsymbol{h}_m = 1$.

Making use of the Bloch form (48) of the unitary operation $\mathcal{U}_m$ we find the distortion matrix for both channels:

$$\begin{aligned} \Lambda_m &= \cos 2t(\mathbb{1} - \boldsymbol{h}_m \boldsymbol{h}_m^{\intercal}) + \sin 2t[\boldsymbol{h}_m] + \boldsymbol{h}_m \boldsymbol{h}_m^{\intercal} \\ &= e^{2t[\boldsymbol{h}_m]}, \end{aligned} \tag{73}$$

with $m = 1, 2$. The symbol $[\boldsymbol{h}_m]$ denotes the skew-symmetric matrix defined by $[\boldsymbol{h}_m]\boldsymbol{r} = \boldsymbol{h}_m \times \boldsymbol{r}$. This is evidently a unital operation and therefore its translation vector vanishes, $\boldsymbol{l}_m = \boldsymbol{0}$.

To make the model more realistic assume that a single qubit, controlled by a Hamiltonian H_m, suffers decoherence induced by the *depolarizing channel.* The evolution of the

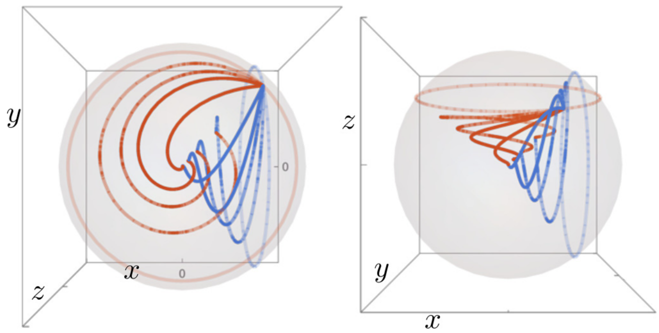

FIG. 6. Visualizations of two unitary channels defined by $\boldsymbol{h}_1 = (0, 0, 1)^\intercal$ and $\boldsymbol{h}_2 = (1, 0, 0)^\intercal$, as a function of time $t \in \{0, \pi\}$ applied over a state with Bloch vector $\boldsymbol{r}_0 = \frac{1}{\sqrt{3}}(1, 1, 1)^\intercal$, under a depolarizing channel with damping rate Γ. Each continues line is the trajectory of $\boldsymbol{r}_1(t) = e^{-\Gamma t} e^{2t[\boldsymbol{h}_1]}\boldsymbol{r}_0$ for different values of $\Gamma \in \{0, 0.2, 0.4, 0.8, 1\}$ (the opacity increase with Γ). The dashed lines correspond to the trajectories $\boldsymbol{r}_2(t) = e^{-\Gamma t} e^{2t[\boldsymbol{h}_2]}\boldsymbol{r}_0$ for the same values Γ. Both figures present the same trajectories from different perspectives. We assume $\hbar = \omega = 1$, so all quantities are dimensionless.

system is governed by the master equation,

$$\frac{d\rho}{dt} = -i[H_m, \rho] - \Gamma\left(\rho - \frac{1}{2}\mathbb{1}\right), \tag{74}$$

with the damping rate Γ. Adopting the convention $\hbar = 1$ we are assured that in these units the frequency is equal to 1.

Any Bloch vector $\boldsymbol{h}_m$ determines, through Eq. (71), the Hamiltonian H_m. Hence the master equation (74) leads to the following dynamics of the Bloch vector $\boldsymbol{r}$:

$$\frac{d\boldsymbol{r}}{dt} = 2(\boldsymbol{h}_m \times \boldsymbol{r}) - \Gamma \boldsymbol{r}, \tag{75}$$

where $\rho = \frac{1}{2}(\mathbb{1} + \boldsymbol{r}\cdot\boldsymbol{\sigma})$.

Solving this equation, we arrive at the time dependence:

$$\boldsymbol{r}(t) = e^{-\Gamma t} e^{2t[\boldsymbol{h}_m]}\boldsymbol{r}_0. \tag{76}$$

The map $\mathcal{E}_m^{\text{dec}}$ can be written as a concatenation of a unitary dynamics and a depolarizing channel, $\mathcal{E}_m^{\text{dec}} = \mathcal{D} \circ \mathcal{U}_m$, with the distortion matrix

$$\Lambda_m^{\text{dec}} = e^{-\Gamma t} e^{2t[\boldsymbol{h}_m]}. \tag{77}$$

In Fig. 6, we show the resulting trajectories from these kinds of channels. We have fixed the initial Bloch vector, $\boldsymbol{r}_0 = \frac{1}{\sqrt{3}}(1, 1, 1)^\intercal$, and evolved it by two Hamiltonians corresponding to $\boldsymbol{h}_1 = (0, 0, 1)^\intercal$ and $\boldsymbol{h}_2 = (1, 0, 0)^\intercal$. Note how the combined channel (unitary transformation and depolarizing channel) becomes less distinguishable as the decoherence parameter Γ increases.

Observe that a rotation of the vector $\boldsymbol{h}_m$ generates a particular transformation on the distortion matrix Λ_m^{dec}. Equation (73) implies that $\tilde{\Lambda}_m^{\text{dec}} = R\Lambda_m^{\text{dec}}R^\intercal$ if $\boldsymbol{h}_m' = R\boldsymbol{h}_m$ with R being an orthogonal matrix and $\tilde{\Lambda}_m^{\text{dec}}$ specified by $\boldsymbol{h}_m'$.

Assume that we need to distinguish between two Hamiltonians, H_1 and H_2, related to vectors $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$, respectively. The evolved state of the system will depend on time and on the damping parameter Γ. A fundamental problem in quantum information is managing the decoherence effects while keeping measurement precision. Our aim is to find the optimal evolution time allowing for the best distinguishability between both Hamiltonians in view of the transmission distance between the channels and the measures proposed in Refs. [15,16].

Comparison of distinguishability measures

We are going to analyze the transmission distance between quantum channels. For $N = 2$, the Choi matrix of an arbitrary quantum channel $\mathcal{M}$ can be written as [14]

$$\rho_{\mathcal{M}} = \frac{1}{4}\left(\mathbb{1}\otimes\mathbb{1} + \mathbb{1}\otimes\boldsymbol{l}\cdot\boldsymbol{\sigma} + \sum_{i,j}\Lambda'_{ij}\sigma_i\otimes\sigma_j\right), \tag{78}$$

where Λ and $\boldsymbol{l}$ denote the distortion matrix and translation vector of the map [see Eq. (48)], while $\Lambda'_{ij} = (C\Lambda^\intercal)_{ij}$, with $C = \text{diag}(1, -1, 1)$.

Following Sec. III, we have to compare the evolved Choi states,

$$\rho_m = \left(\mathcal{E}_m^{\text{dec}}\otimes\mathbb{1}\right)(|\Phi\rangle\langle\Phi|),$$

where $\mathcal{E}_m^{\text{dec}} = \mathcal{D}\circ\mathcal{U}_m$. We use the transmission distance (16), which can be obtained by inserting Eq. (77) into Eq. (78) with $\boldsymbol{l} = 0$. Note that calculation of $d_t^{\text{iso}}(\mathcal{D}\circ\mathcal{U}_1, \mathcal{D}\circ\mathcal{U}_2)$ involves two noncommuting Choi states.

Let us evaluate the entropic channel divergence (35) for unital quantum channels (77), with the distortion matrix proportional to a rotation matrix.

Proposition 4. The entropic channel divergence (35) between two unital maps $\mathcal{E}_1$ and $\mathcal{E}_2$, with distortion matrices $\Lambda_1 = \alpha_1 R_1$ and $\Lambda_2 = \alpha_2 R_2$, respectively, reads

$$d_t^{K=1}(\mathcal{E}_1, \mathcal{E}_2) = \sqrt{f(r_{\text{opt}}) - \frac{\left[f\left(\alpha_1^2\right) + f\left(\alpha_2^2\right)\right]}{2}}, \tag{79}$$

with

$$r_{\text{opt}} = \alpha_1^2 + \alpha_2^2 + \alpha_1\alpha_2\left(\text{Tr}\left[\Lambda_1^\intercal\Lambda_2\right] - 1\right), \tag{80}$$

and the function $f(\cdot)$ defined in Eq. (61).

Proof. Employing the same reasoning used to derive Eq. (63), we arrive at

$$\text{QJSD}[\mathcal{E}_1(\rho), \mathcal{E}_2(\rho)] = f\left[\boldsymbol{r}\cdot\left(\Lambda_{\overline{\mathcal{E}}}^\intercal\Lambda_{\overline{\mathcal{E}}}\right)\boldsymbol{r}\right] - \frac{1}{2}f\left(\alpha_1^2\right) - \frac{1}{2}f\left(\alpha_2^2\right), \tag{81}$$

where $\Lambda_{\overline{\mathcal{E}}} = \alpha_1 R_1 + \alpha_2 R_2$. To calculate the entropic channel divergence we need to optimize the function f used in Eq. (61):

$$f\left[\boldsymbol{r}\cdot\left(\Lambda_{\overline{\mathcal{E}}}^\intercal\Lambda_{\overline{\mathcal{E}}}\right)\boldsymbol{r}\right].$$

As $f(x)$ is a decreasing function of x in [0,1], we have to minimize

$$\boldsymbol{r}\cdot\left(\Lambda_{\overline{\mathcal{E}}}^\intercal\Lambda_{\overline{\mathcal{E}}}\right)\boldsymbol{r} = \alpha_1^2 + \alpha_2^2 + 2p_1p_2\boldsymbol{r}\cdot\left(\Lambda_1^\intercal\Lambda_2\right)\boldsymbol{r}, \tag{82}$$

over the sphere $\boldsymbol{r}\cdot\boldsymbol{r} = 1$.

Taking $\Lambda_1^\intercal \Lambda_2 = e^{\phi[\boldsymbol{h}]}$ for some $\boldsymbol{h}$ such that $\boldsymbol{h}^2 = 1$ [see Eq. (73)], we find that

$$\boldsymbol{r} \cdot (\Lambda_1^\intercal \Lambda_2)\boldsymbol{r} = \cos\phi + (1 - \cos\phi)(\cos\gamma)^2,$$

where $\cos\gamma = \boldsymbol{h} \cdot \boldsymbol{r}$. The minimum of the function $\boldsymbol{r} \cdot (\Lambda_1^\intercal \Lambda_2)\boldsymbol{r}$ in the sphere $\boldsymbol{r}^2 = 1$ is correspondingly given by the minimum of the previous function over the parameter γ. It is straightforward to show that $\gamma = \pi/2$ minimizes $\boldsymbol{r} \cdot (\Lambda_1^\intercal \Lambda_2)\boldsymbol{r}$, and therefore

$$\min_{\boldsymbol{r}}\{\boldsymbol{r} \cdot e^{\phi[\boldsymbol{h}]}\boldsymbol{r}\} = \cos\phi.$$

Finally, employing the following equality,

$$\cos\phi = \frac{\mathrm{Tr}[e^{\phi[\boldsymbol{h}]}] - 1}{2},$$

we arrive at

$$\min_{\boldsymbol{r}} \{\boldsymbol{r} \cdot (\Lambda_1^\intercal \Lambda_2)\boldsymbol{r}\} = \frac{\mathrm{Tr}[\Lambda_1^\intercal \Lambda_2] - 1}{2}. \tag{83}$$

By inserting this in Eq. (82), we obtain Eq. (79). ■

On the other hand, if $\boldsymbol{h}_m' = R\boldsymbol{h}_m$, with R denoting an orthogonal matrix of order 3, the corresponding affine matrix Λ_m^{dec} transforms as

$$\tilde{\Lambda}_m^{\text{dec}} = R\Lambda_m^{\text{dec}}R^\intercal. \tag{84}$$

Therefore, the quantum operation $\tilde{\mathcal{E}}_m^{\text{dec}}$ associated with $\tilde{\Lambda}_m^{\text{dec}}$ can be written as $\tilde{\mathcal{E}}_m^{\text{dec}} = \mathcal{R} \circ \mathcal{E}_m^{\text{dec}} \circ \mathcal{R}^{-1}$, where $\mathcal{R}$ is the unitary channel corresponding to the rotation matrix R, while $\mathcal{E}^{\text{dec}}{}_m$ is determined by Λ_m^{dec}.

Since the distance measures between quantum operations satisfy the unitary invariance (27), the distinguishability between operations $\mathcal{E}_1^{\text{dec}}$ and $\mathcal{E}_2^{\text{dec}}$ specified by $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$, respectively, depends only on the angle

$$\theta = \arccos \boldsymbol{h}_1 \cdot \boldsymbol{h}_2,$$

and the damping rate Γ. Thus, without losing generality we can fix the vector $\boldsymbol{h}_1$ in the direction z.

In Fig. 7, we present the transmission distance $d_t^{\text{iso}}(\mathcal{U}_1, \mathcal{U}_2)$ and the Bures distance $D_B(\mathcal{U}_1, \mathcal{U}_2)$ defined in (31). Both quantities are computed in the noiseless case, $\Gamma = 0$, and shown as functions of time t for different values of the angle θ. In this case the entropic channel divergence is equal to the transmission distance (16) between quantum channels. The Bures distance D_B is based on the quantum fidelity between the Choi matrices (see Refs. [15,16]).

The time in which the distinguishability is maximal, according to the measures analyzed, reads

$$t_{\max} = \begin{cases} \pi/2 & \text{if } \cos\theta \geqslant 0 \\ \frac{1}{2}\cos^{-1}\left(\frac{\cos\theta+1}{\cos\theta-1}\right) & \text{if } \cos\theta < 0 \end{cases}. \tag{85}$$

Note that if $\cos\theta > 0$, both unitary operations cannot be distinguished with probability 1 at any time. However, if $\cos\theta \leqslant 0$, there exists a time in which the pure Choi states are orthogonal and can be perfectly distinguished at the selected interaction time $t_{\max}$.

Let us take into account effects of the decoherence. The depolarizing channel (77) transforms the original unitary rotations into channels that send states closer to the maximally mixed state (see Fig. 6) so the problem of distinguishability between the channels becomes more difficult.

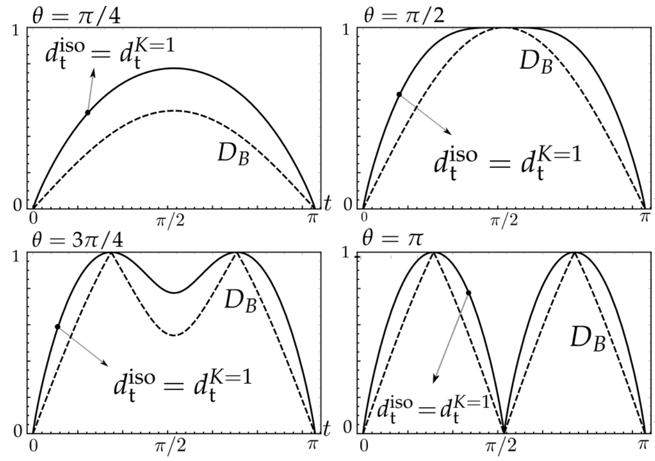

FIG. 7. Transmission distance (16), Bures distance (31), and the entropic channel divergence (35), between two unitary operations (72), whose corresponding vectors $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ form an angle $\theta \in \{\pi/4, \pi/2, 3\pi/4, \pi\}$, as functions of time t. We assume that $\hbar = \omega = 1$, so all quantities are dimensionless.

This problem was already treated in Ref. [15], where it was suggested to select a constant initial state, with the Bloch vector $\boldsymbol{r}_0 = (1, 0, 0)^\intercal$, and to choose the optimal time as the one minimizing the error probability P_{error}. Such an optimal time t_{opt} corresponds to the maximal distinguishability between both evolved states:

$$P_{\text{error}} = \tfrac{1}{2}[1 - \exp(-pt)|\sin t|]. \tag{86}$$

At a time $t_{\text{opt}} = \arctan(1/p)$, P_{err} is minimized and thus the information gained by the measurement is maximized.

Regarding entropic distinguishability measures, Fig. 8 displays behavior of the transmission distance under unitary evolution and decoherence, for angle $\theta = \pi/2$ and exemplary values of the damping rate, $\Gamma \in \{0, 0.3, 0.6, 0.9, 1.2, 1.5, 1.8\}$.

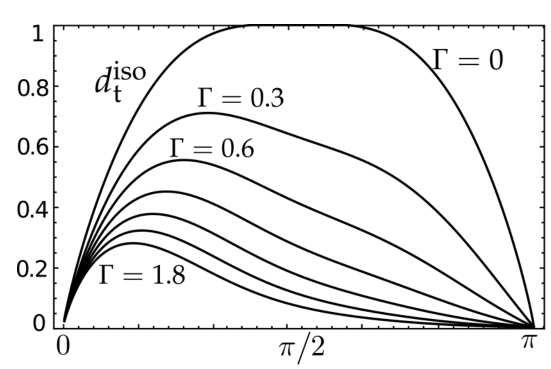

FIG. 8. Transmission distance $d_t^{\text{iso}}(\mathcal{D} \circ \mathcal{U}_1, \mathcal{D} \circ \mathcal{U}_2)$ as a function of time t, where the affine decomposition of the maps $\mathcal{D} \circ \mathcal{U}_i$ is given by (77). The angle between the Bloch vectors defining both Hamiltonians ($\boldsymbol{h}_1$ and $\boldsymbol{h}_2$) is $\theta = \pi/2$. Here $\mathcal{D}$ denotes the depolarizing channel with damping rate Γ, which labels the curves. The larger damping rate, the shorter time $t_{\max}$ of maximal distinguishability. We assume that $\hbar = \omega = 1$, so all quantities depicted are unitless.

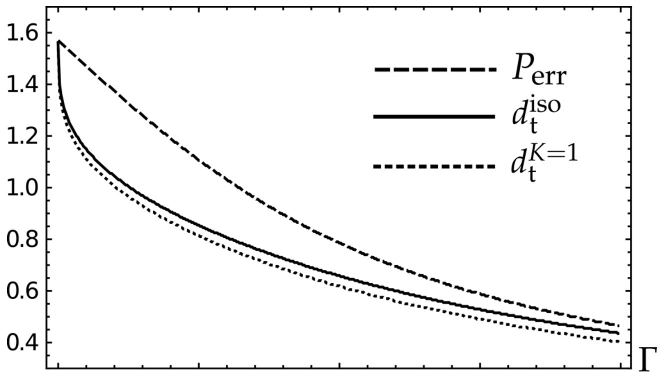

FIG. 9. Optimal times as a function of the noise parameter Γ, in the distinguishability of Hamiltonians (see Sec. VI B). The maps are given by (77). The dashed line corresponds to optimal times for the probability of error, $P_{\rm err}$ [see (86)]. The continuous line represents the transmission distance $d_t^{\rm iso}$ between quantum channels (16), while the dotted line corresponds to the optimal times in the case of the entropic channel divergence, $d_t^{K=1}$ [see (35)]. All quantities are dimensionless ($\hbar = \omega = 1$).

The entropic channel divergence is given by taking $\alpha_i = e^{-\Gamma t}$ and $\Lambda_i = e^{2t[\boldsymbol{h}_i]}$, with $i = 1, 2$, in Eq. (79). In this way one obtains

$$\mathrm{Tr}\big[\Lambda_1^{\intercal}\Lambda_2\big] = 2\cos(2\theta)\sin^4(t) + 2\cos(\theta)\sin^2(2t) + \cos(2t) + \frac{3}{4}\cos(4t) + \frac{5}{4}, \tag{87}$$

where θ denotes the angle between both Bloch vectors, $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$. Inserting (83) into (81), we arrive at the dependence of the entropic channel divergence on the angle θ, the time t, and the damping parameter Γ.

One can pose a natural question: Which interaction time is optimal to distinguish Hamiltonians H_1 and H_2 under decoherence? In the noiseless situation $\Gamma = 0$, the entropic channel divergence results to be equal to the transmission distance between quantum channels, Eq. (16), therefore the interaction time (85) is optimal for this measure as well. In presence of decoherence, each distinguishability measure has its own behavior, leading to different values of optimal interaction times. Figure 9 shows that the best times to measure the distinguishability related to the transmission distance $d_t^{\rm iso}$ are shorter than those arising from minimizing the error probability of distinguishing the two evolved states (86), proposed in Ref. [15]. This is a noteworthy behavior because one might think that the dense coding capacity is completely defined by the distinguishability of the maps $\mathcal{E}$ and $\mathcal{F}$, captured in this case by the entropic channel divergence $d_t^{K=1}$. However, each quantity has its own optimal time suggesting a more complex relation between channel discrimination and dense coding. This topic will be considered in future investigations.

VII. CONCLUDING REMARKS

We have introduced two entropic measures of distinguishability between quantum operations using the square root of the quantum Jensen-Shannon divergence, also called *transmission distance*. We have investigated their properties and physical interpretations.

In the case of the *transmission distance between quantum channels* $d_t^{\rm iso}$, we have shown that this measure satisfies several criteria for a suitable distance measure between maps. Even though this quantity does not satisfy the chaining property, this is the case if one of the maps applied first is bistochastic, which is a key property for estimating errors in quantum information protocols [1]. Furthermore, the transmission distance between quantum channels does not require any optimization procedure and it can be directly obtained by calculating the entropy of a map, defined in Ref. [27]. Regarding the physical interpretation of this measure, $d_t^{\rm iso}$ is the dense coding capacity for a noiseless dense coding protocol. It is therefore fair to expect that the transmission distance between quantum channels is a good candidate for error or diagnostic measures.

In Sec. IV, we have introduced the *entropic channel divergence* d_t^K, parametrized by the size K of the ancilla. In addition to the requirements mentioned in Refs. [1,16], we have shown that d_t^K satisfies the chain rule. This property allows one to prove the amortization collapse of the entropic channel divergence, which can be useful to obtain new single-letter converse bounds on the capacity of adaptive protocols in channel discrimination theory [12]. Regarding physical motivation, d_t^K is the square root of the quantum reading capacity in the equiprobable case [31], and it can be identified as the capacity of a dense coding protocol with a resource influenced by decoherence [30].

In Sec. V C, we have considered the case of Choi-stretchable channels. For these kinds of quantum operations, $d_t^{\rm iso}$ and d_t^N are equal, establishing a particular situation, in which the transmission distance between quantum operations is equal to the stabilized entropic channel divergence (36).

To demonstrate the analyzed measures in action, we have investigated the distinguishability of two Pauli channels and provided analytical expressions for the distance $d_t^{\rm iso}$ and the entropic divergence $d_t^{K=1}$. As the standard teleportation protocol can be written as a Pauli map, we have studied the presence of noise in quantum teleportation by calculating both distinguishability measures. The transmission distance $d_t^{\rm iso}$ between quantum channels occurred to be the most sensitive to decoherence, while the trace distance between the corresponding Choi states is more sensitive than the entropic channel divergence.

In the case of a Hamiltonian evolution under decoherence, we have compared the distance $d_t^{\rm iso}$ and the divergence $d_t^{K=1}$ between the quantum operations with the Bures distance between the corresponding Choi states and the *probability of error*, originally studied in Ref. [15]. In the absence of noise, the distance measures defined by employing the transmission distance become equal, $d_t^{\rm iso} = d_t^{K=1}$, showing a smoother behavior than the Bures distance and exhibiting equal times of maximal distinguishability.

To distinguish between dynamics generated by two Hamiltonians subjected to decoherence, we have studied the entropic measures $d_t^{\rm iso}$ and $d_t^{K=1}$ and compared them with the error probability $P_{\rm err}$. For these measures we identified the time window of maximal distinguishability while varying the decoherence rate Γ. The above observations suggest that the measures of the distance between quantum operations based on the square root of the Jensen-Shannon divergence (in this case equivalent to the Holevo quantity) introduced in this paper will find their applications in further theoretical and experimental studies.

ACKNOWLEDGMENTS

D.G.B. and P.W.L. are grateful to the Jagiellonian University for the hospitality during their stay in Cracow. They acknowledge financial support by Consejo Nacional de Investigaciones Científicas y Técnicas, and by Universidad Nacional de Córdoba, Argentina. K.Ż. is supported by Narodowe Centrum Nauki under Quantera Project No. 2021/03/Y/ST2/00193 and by the Foundation for Polish Science under Team-Net Project No. POIR.04.04.00-00-17C1/18-00.

APPENDIX A: CHANNEL DIVERGENCE WITH TRACE DISTANCE BETWEEN UNITAL CHANNELS

Let us calculate

$$d_{\rm Tr}^{K=1}(\mathcal{E},\mathcal{F}) = \sup_{\rho\in\mathcal{M}_N} T[\mathcal{N}(\rho),\mathcal{M}(\rho)],$$

for two arbitrary unital quantum operations $\mathcal{N}$ and $\mathcal{M}$ with $N=2$, $T(\cdot,\cdot)$ being the trace distance.

Performing required calculations we arrive at an expression,

$$d_{\rm Tr}^{K=1}(\mathcal{N},\mathcal{M}) = \tfrac{1}{2}\max_{\boldsymbol{r}} \sqrt{\boldsymbol{r}\cdot\Delta\boldsymbol{r}}, \tag{A1}$$

where $\boldsymbol{r}$ denotes the Bloch vector ρ and $\Delta = (\Lambda_{\mathcal{N}} - \Lambda_{\mathcal{M}})^{\mathsf{T}}(\Lambda_{\mathcal{N}} - \Lambda_{\mathcal{M}})$.

We need now to optimize $\sqrt{\boldsymbol{r}\cdot\Delta\boldsymbol{r}}$ over the sphere $\boldsymbol{r}^2=1$. As Δ is a symmetric positive square matrix, we can take its spectral decomposition:

$$\Delta = \sum_i \lambda_i^{\Delta}\boldsymbol{x}_i\boldsymbol{x}_i^{\mathsf{T}}, \tag{A2}$$

where $\boldsymbol{x}_i$ denotes the eigenvector of Δ corresponding to the eigenvalue λ_i. One obtains, therefore,

$$\boldsymbol{r}\cdot\Delta\boldsymbol{r} = \sum_i \lambda_i^{\Delta}(\boldsymbol{r}\cdot\boldsymbol{x}_i)^2.$$

Having in mind that $\lambda_i^{\Delta}\geqslant 0$ and $(\boldsymbol{r}\cdot\boldsymbol{x}_i)^2\in[0,1]$ for any i, it is clear that the maximum is achieved when $\boldsymbol{r}=\boldsymbol{x}_k$ with k such that $\lambda_k^{\Delta}\geqslant\lambda_i^{\Delta}$ for all i. This implies directly Eq. (69), specifically,

$$d_{\rm Tr}^{K=1}(\mathcal{N},\mathcal{M}) = \tfrac{1}{2}\max_i \sqrt{\lambda_i^{\Delta}},$$

where $\{\lambda_i^{\Delta}\}_i$ is the set of eigenvalues of the matrix

$$\Delta = (\Lambda_{\mathcal{N}} - \Lambda_{\mathcal{M}})^{\mathsf{T}}(\Lambda_{\mathcal{N}} - \Lambda_{\mathcal{M}}).$$

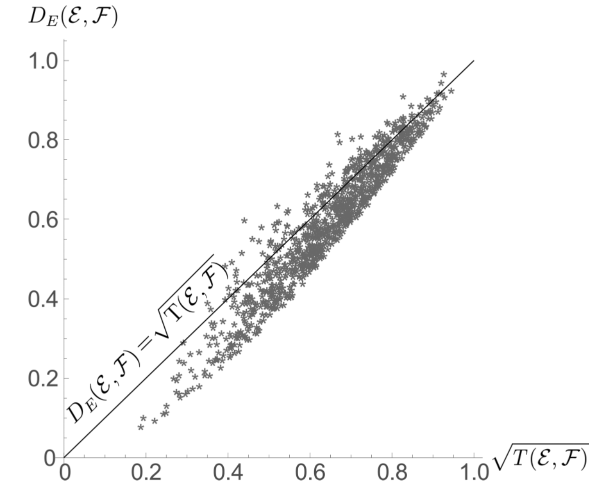

FIG. 10. Square root of the trace distance $\sqrt{T(\mathcal{E},\mathcal{F})}$ between random Choi states, and their entropic distance $D_E(\mathcal{E},\mathcal{F})$, defined in Eqs. (29) and (32), between 1000 pairs of channels taken randomly according to the flat measure in the regular tetrahedron of Pauli channels. As points are scattered on both sides of the diagonal, these results show that the *min* function should be used in the upper bound (B1). All quantities depicted are dimensionless.

APPENDIX B: UPPER BOUND FOR THE TRANSMISSION DISTANCE

Two different upper bounds for the transmission distance d_t between quantum states can be found in the literature: one in terms of the entropic distance D_E defined in Eq. (13) [21] and the other one based on the square root of the trace distance $\sqrt{T}$ [9].

In Eq. (33) we have included the corresponding bound for quantum maps:

$$d_t^{\rm iso}(\mathcal{E},\mathcal{F}) \leqslant \min\{\sqrt{T(\mathcal{E},\mathcal{F})},\ D_E(\mathcal{E},\mathcal{F})\}. \tag{B1}$$

Note that the function *minimum* appears in this bound. In Fig. 10, we analyze an ensemble of random pairs of Choi states of order 4, corresponding to unital Pauli maps, and compare the distances given by $\sqrt{T}$ and D_E between them. Numerical results show that for some pairs of channels it holds that $\sqrt{T} > D_E$ and for others $\sqrt{T} < D_E$. These observations imply that using the function *minimum* in Eq. (B1) is justified as it makes the upper bound stronger.

[1] A. Gilchrist, N. K. Langford, and M. A. Nielsen, Distance measures to compare real and ideal quantum processes, Phys. Rev. A **71**, 062310 (2005).

[2] Z. Puchała, J. A. Miszczak, P. Gawron, and B. Gardas, Experimentally feasible measures of distance between quantum operations, Quantum Inf Process **10**, 1 (2011).

[3] D. Virosztek, The metric property of the quantum Jensen-Shannon divergence, Adv. Math. **380**, 107595 (2021).

[4] S. Sra, Metrics induced by Jensen-Shannon and related divergences on positive definite matrices, Linear Algebra and its Applications **616**, 125 (2021).

[5] K. M. Audenaert, Quantum skew divergence, J. Math. Phys. **55**, 112202 (2014).

[6] C. Radhakrishnan, M. Parthasarathy, S. Jambulingam, and T. Byrnes, Distribution of Quantum Coherence in Multipartite Systems, Phys. Rev. Lett. **116**, 150504 (2016).

[7] N. Megier, A. Smirne, and B. Vacchini, Entropic Bounds on Information Backflow, Phys. Rev. Lett. **127**, 030401 (2021).

[8] F. Settimo, H. P. Breuer, and B. Vacchini, Entropic and trace-distance-based measures of non-Markovianity, Phys. Rev. A **106**, 042212 (2022).

[9] J. Briët and P. Harremoës, Properties of classical and quantum Jensen-Shannon divergence, Phys. Rev. A **79**, 052311 (2009).

[10] S. Pirandola, On quantum reading, quantum illumination, and other notions, IOP SciNotes **2**, 015203 (2021).

[11] K. Fang, O. Fawzi, R. Renner, and D. Sutter, Chain Rule for the Quantum Relative Entropy, Phys. Rev. Lett. **124**, 100501 (2020).

[12] M. M. Wilde, M. Berta, C. Hirche, and E. Kaur, Amortized channel divergence for asymptotic quantum channel discrimination, Lett. Math. Phys. **110**, 2277 (2020).

[13] F. Leditzky, E. Kaur, N. Datta, and M. M. Wilde, Approaches for approximate additivity of the Holevo information of quantum channels, Phys. Rev. A **97**, 012332 (2018).

[14] F. Shahbeigi and S. J. Akhtarshenas, Quantumness of quantum channels, Phys. Rev. A **98**, 042313 (2018).

[15] A. M. Childs, J. Preskill, and J. Renes, Quantum information and precision measurement, J. Mod. Opt. **47**, 155 (2000).

[16] M. Raginsky, A fidelity measure for quantum channels, Phys. Lett. A **290**, 11 (2001).

[17] M. Ohya and D. Petz, *Quantum Entropy and its Use*, (Springer-Verlag, Heidelberg, 2004).

[18] A. S. Holevo, Bounds for the quantity of information transmitted by a quantum communication channel, Probl. Peredachi Inf. **9**, 3 (1973).

[19] A. S. Holevo and V. Giovannetti, Quantum channels and their entropic characteristics, Rep. Prog. Phys. **75**, 046001 (2012).

[20] A. P. Majtey, P. W. Lamberti, and D. P. Prato, Jensen-Shannon divergence as a measure of distinguishability between mixed quantum states, Phys. Rev. A **72**, 052310 (2005).

[21] P. W. Lamberti, A. P. Majtey, A. Borras, M. Casas, and A. Plastino, Metric character of the quantum Jensen-Shannon divergence, Phys. Rev. A **77**, 052311 (2008).

[22] C. Manning and H. Schutze, *Foundations of Statistical Natural Language Processing* (MIT, Cambridge, MA, 1999).

[23] F. Nielsen, On the Jensen-Shannon symmetrization of distances relying on abstract means, Entropy **21**, 485 (2019).

[24] W. Roga, M. Fannes, and K. Życzkowski, Universal Bounds for the Holevo Quantity, Coherent Information, and the Jensen-Shannon Divergence, Phys. Rev. Lett. **105**, 040505 (2010).

[25] J. Watrous, *The Theory of Quantum Information* (Cambridge University, New York, 2018).

[26] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University, New York, 2006).

[27] W. Roga, K. Życzkowski, and M. Fannes, Entropic characterization of quantum operations, Int. J. Quantum Inform. **09**, 1031 (2011).

[28] J. Watrous, Mixing doubly stochastic quantum channels with the completely depolarizing channel, Quantum Inf. Comput. **9**, 406 (2009).

[29] D. Aharonov, A. Kitaev, and N. Nisan, Quantum circuits with mixed states, Proc. Annu. ACM Symp. Theory Comput. **1**, 20 (1998).

[30] R. Laurenza, C. Lupo, S. Lloyd, and S. Pirandola, Dense coding capacity of a quantum channel, Phys. Rev. Res. **2**, 023023 (2020).

[31] S. Pirandola, C. Lupo, V. Giovannetti, S. Mancini, and S. L. Braunstein, Quantum reading capacity, New J. Phys. **13**, 113012 (2011).

[32] S. Pirandola, Quantum Reading of a Classical Digital Memory, Phys. Rev. Lett. **106**, 090504 (2011).

[33] T. P. Cope, L. Hetzel, L. Banchi, and S. Pirandola, Simulation of non-Pauli channels, Phys. Rev. A **96**, 022323 (2017).

[34] M. B. Ruskai, S. Szarek, and E. Werner, An analysis of completely positive trace-preserving maps on M2, Linear Algebra and its Applications **347**, 159 (2002).

[35] S. Luo, Quantum discord for two-qubit systems, Phys. Rev. A **77**, 042303 (2008).

[36] J. Preskill, Quantum information and physics: Some future directions, J. Mod. Opt. **47**, 127 (2000).