Operational applications of the diamond norm and related measures in quantifying the non-physicality of quantum maps

Bartosz Regula,[1, *] Ryuji Takagi,[1, †] and Mile Gu[1, 2, 3, ‡]

[1]*Nanyang Quantum Hub, School of Physical and Mathematical Sciences,*
Nanyang Technological University, 637371, Singapore
[2]*Complexity Institute, Nanyang Technological University, 637371, Singapore*
[3]*Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, 117543, Singapore*

Although quantum channels underlie the dynamics of quantum states, maps which are not physical channels — that is, not completely positive — can often be encountered in settings such as entanglement detection, non-Markovian quantum dynamics, or error mitigation. We introduce an operational approach to the quantitative study of the non-physicality of linear maps based on different ways to approximate a given linear map with quantum channels. Our first measure directly quantifies the cost of simulating a given map using physically implementable quantum channels, shifting the difficulty in simulating unphysical dynamics onto the task of simulating linear combinations of quantum states. Our second measure benchmarks the quantitative advantages that a non-completely-positive map can provide in discrimination-based quantum games. Notably, we show that for any trace-preserving map, the quantities both reduce to a fundamental distance measure: the diamond norm, thus endowing this norm with new operational meanings in the characterisation of linear maps. We discuss applications of our results to structural physical approximations of positive maps, quantification of non-Markovianity, and bounding the cost of error mitigation.

1. Introduction

It is one of the fundamental properties of quantum mechanics that the evolution of quantum states is described by linear maps which are completely positive and trace preserving (CPTP), stemming from the unitary dynamics enforced on a larger Hilbert space [1]. However, in several different settings of practical importance, various applications of quantum dynamics which are not CPTP can be encountered. This motivates the study of such transformations, and in particular a precise understanding of how they can be compared with and approximated by physical quantum channels.

One important application of non-CPTP maps is in entanglement detection, where positive but not completely positive maps can serve as entanglement witnesses [2]. A bipartite state ρ is entangled if and only if there exists a positive map Φ such that $\mathrm{id} \otimes \Phi(\rho)$ is no longer a positive operator, and therefore such a map can reveal the correlations of ρ. This approach has constituted one of the most important ways of detecting entanglement [3, 4], but its experimental implementation encounters an obstacle: how to realise the action of an unphysical linear map in practice? This question prompted the introduction of structural physical approximations (SPA) of non-CPTP maps [5], which aim to enable the physical evaluation of general maps by designing suitable approximations in terms of quantum channels and using them to infer properties of the original map [6–8].

Another setting in which non-CPTP maps are encountered is that of non-Markovian quantum dynamics or, generally, in the reduced dynamics of correlated systems. Specifically, when an open quantum system shares some initial correlations with its environment, the evolution of the composite system-environment state can correspond to a non-CPTP map when looking only at the dynamics of the reduced state of the system [9–12]. Although the physical interpretation of this is a matter of debate and alternative ways to understand such dynamics have been proposed [13–15], it can nevertheless be useful to study such non-CPTP evolutions directly to gain an understanding of reduced dynamics of open quantum systems.

Even broader types of unphysical quantum dynamics can be found in the areas of quantum error correction and error mitigation [16–18]. This is because, in a broad sense, both of these settings are concerned with the following problem: if an unknown system has undergone a noisy evolution as $\rho \mapsto \Theta(\rho)$, how can we reconstruct the original state as closely as possible, that is, how to implement a map Φ such that $\Phi \circ \Theta(\rho) \approx \rho$? Such inverse operations typically cease to be valid quantum channels, and so it becomes necessary to devise approaches to implement them in practice with the use of physical operations.

---

* bartosz.regula@gmail.com
† ryuji.takagi@ntu.edu.sg
‡ mgu@quantumcomplexity.org

In this work, we introduce a general quantitative framework for the characterisation of such unphysical maps by approximating them with quantum channels. We then explicitly give the considered measures operational meaning by connecting them with the performance of practical tasks, including the cost of simulating a given map with quantum channels. Notably, we show that all of the considered measures reduce to the same quantity when the given linear map is trace preserving: they all equal the diamond norm [19, 20], a fundamental computational tool that serves as a measure of quantum channel distance and finds many uses in the practical characterisation of quantum processes [21]. This endows the diamond norm with new meanings in the operational tasks that we consider, and furthermore allows a number of new connections to be established. On the one hand, many known results in the quantification of the diamond norm can be carried over to the setting of our work, and on the other hand, we can use our characterisation to provide new insight into the computation and applications of the diamond norm.

Our approach is based on the notion of robustness measures [22] — inspired by recent applications of such quantities in the study of general resource theories of channels [23–33], we use them to quantify the amount of noise needed to turn a given map into a quantum channel. Such measures allow for several different generalisations to the setting of linear maps, motivating us to study and compare these definitions. The robustness-based approaches can be understood as different ways of designing optimal decompositions of linear maps in terms of quantum channels, and so they generalise the standard structural physical approximations [5]. We express the measures as semidefinite programs and establish various relations and bounds between them.

We apply our first measure in the task of simulating the action of an unphysical map with valid channels, accomplished by allowing the use of ancillary systems which can consists of linear combinations of quantum states. Such an approach allows us to reduce the problem of simulating the dynamics of quantum systems to the much simpler case of simulating the use of a non-positive Hermitian operator. Assessing the difficulty of this procedure then reduces to quantifying how much the given operator deviates from being a valid quantum state, and — employing the trace norm as a natural quantifier of such 'non-quantumness' — we show that the optimal cost of simulating a non-CPTP map in this way is given exactly by the value of the robustness measure.

Furthermore, answering the question of whether any unphysical map can provide measurable operational advantages over quantum channels, we show this to be the case in the setting of discrimination-based quantum games, establishing our second robustness measure as the exact quantifier of this advantage.

Our results also generalise and shed light on the very recent findings of Ref. [33], which considered a similar framework for approximating trace-preserving maps using a robustness- and quasiprobability-based approach. In particular, we show that the measure considered in [33] is actually an alternative expression for the diamond norm of a map, rather than a new quantity.

The paper is structured as follows. In Sec. 2, we introduce the notions of robustness measures and show how they can be applied to non-CPTP linear maps. We establish precise connections with the diamond norm in Sec. 3. We then proceed to show that the robustness measures — and hence the diamond norm — play a crucial role in quantifying the cost of simulating linear maps (Sec. 4) as well as in understanding the advantages a non-CPTP map could provide in input-output quantum games (Sec. 5). We proceed to establish a number of bounds for the measures in Sec. 6. Finally, we discuss the applications of our approach, comparisons with other methods, and explicitly show how the measures can be evaluated for some representative examples in Sec. 7.

2. Robustness of non-CP maps

Let A and B denote two finite-dimensional quantum systems of dimension d_A and d_B, respectively. We will use $\mathbb{L}(A)$ to denote the set of all linear operators, $\mathbb{H}(A)$ to denote the set of all Hermitian operators, and $\mathbb{D}(A)$ to denote all density operators acting on the Hilbert space of system A. We use $\langle X, Y \rangle = \text{Tr}(X^\dagger Y)$ for the Hilbert-Schmidt inner product.

Among all linear maps from $\mathbb{L}(A)$ to $\mathbb{L}(B)$, we will be primarily concerned with Hermiticity-preserving maps $\mathcal{H}(A, B)$, which are defined as maps such that $\Phi(X) \in \mathbb{H}(B) \ \forall X \in \mathbb{H}(A)$. A map is called positive if $\Phi(X) \geq 0 \ \forall X \geq 0$ (w.r.t. the positive semidefinite cone), completely positive (CP) if $\text{id}_A \otimes \Phi$ is positive, trace preserving if $\text{Tr} \, \Phi(X) = \text{Tr} \, X \ \forall X$, and trace non-increasing if $\text{Tr} \, \Phi(X) \leq \text{Tr} \, X \ \forall X$. To each map $\Phi \in \mathcal{H}(A, B)$ we will associate the Choi operator $J_\Phi = (\text{id}_A \otimes \Phi)[|\Omega\rangle\langle\Omega|] \in \mathbb{H}(A \otimes B)$ where $|\Omega\rangle = \sum_i |ii\rangle$. Importantly, a map is Hermiticity-preserving iff $J_\Phi = J_\Phi^\dagger$, CP iff $J_\Phi \geq 0$, and trace preserving iff $\text{Tr}_B \, J_\Phi = \mathbb{1}_A$ (see e.g. [34]). Let

CPTNI(A, B) denote the set of completely positive and trace–non-increasing maps in $\mathcal{H}(A, B)$, and analogously CPTP(A, B) the set of completely positive and trace-preserving maps. For simplicity of notation, we will often simply write CPTP for CPTP(A, B) (and analogously for other sets) when the spaces in consideration are not relevant.

In order to quantify how much a given map deviates from the set of CPTP maps, we will employ the concept of *robustness measures* [22]. It will be insightful to first review how such measures are defined for quantum states. Given a convex set of interest $\mathcal{F} \subseteq \mathbb{D}$, commonly chosen to be the set of free states in a given resource theory, one asks: how much noise from a set $\mathcal{N} \subseteq \mathbb{D}$ has to be added to a state ρ in order to make it a free state? This has the intuitive interpretation of measuring how robust the resources contained in the state ρ are with respect to noise from the set $\mathcal{N}$. Specifically, we write

$$r_{\mathcal{N}}(\rho) := \min \left\{ \lambda \ \middle| \ \frac{\rho + \lambda \omega}{1 + \lambda} =: \sigma \in \mathcal{F}, \ \omega \in \mathcal{N} \right\}. \tag{1}$$

The most common choices of the noise set $\mathcal{N}$ are: $\mathcal{N} = \mathbb{D}$, in which case we obtain the so-called *generalised robustness* equivalently given by

$$r_{\mathbb{D}}(\rho) = \min \left\{ \lambda \ \middle| \ \rho \leq (1 + \lambda)\sigma, \ \sigma \in \mathcal{F} \right\}, \tag{2}$$

and the choice $\mathcal{N} = \mathcal{F}$, which corresponds to the *standard robustness* $r_{\mathcal{F}}$. The latter quantity is directly related to the so-called base norm $\|\rho\|_{\mathcal{F}}$ of the set $\mathcal{F}$, which can be alternatively understood as an optimisation of quasiprobability distributions over the set $\mathcal{F}$:

$$\begin{aligned} 2r_{\mathcal{F}}(\rho) + 1 &= \|\rho\|_{\mathcal{F}} \\ &:= \min \left\{ \lambda_+ + \lambda_- \ \middle| \ \rho + \lambda_- \sigma_- = \lambda_+ \sigma_+, \ \sigma_\pm \in \mathcal{F} \right\} \\ &= \min \left\{ \sum_i |\lambda_i| \ \middle| \ \rho = \sum_i \lambda_i \sigma_i, \ \sigma_i \in \mathcal{F} \right\} \end{aligned} \tag{3}$$

where the third line is a simple consequence of the convexity of $\mathcal{F}$. The definitions straightforwardly extend to unnormalised operators X: defining

$$r_{\mathcal{N}}(X) := \min \left\{ \lambda \ \middle| \ \frac{X + \lambda \omega}{\operatorname{Tr} X + \lambda} =: \sigma \in \mathcal{F}, \ \omega \in \mathcal{N} \right\}, \tag{4}$$

it is important to notice that the trace of X will come into play, and the base norm will equal $\|X\|_{\mathcal{F}} = 2r_{\mathcal{F}}(X) + \operatorname{Tr} X$.

The case of interest to us will be where the set of free states $\mathcal{F}$ contains *all* physical quantum states, $\mathcal{F} = \mathbb{D}$, in which case the different notions of the robustness are equal and one has

$$2r_{\mathbb{D}}(X) + \operatorname{Tr} X = \|X\|_1, \tag{5}$$

that is, the base norm is precisely the trace norm (Schatten 1-norm) $\|\cdot\|_1$.

Robustness of linear maps. A generalisation of these concepts to the case of linear maps can be done in several different ways. Firstly, one has to note that it does not suffice to consider trace-preserving maps in the definitions of this measures. This follows since any linear combination of CPTP maps necessarily satisfies that $\operatorname{Tr}_B J_\Phi \propto \mathbb{1}$, which means that $\operatorname{Tr} \Phi(\rho)$ takes the same value for any input state ρ. Therefore, any Hermiticity-preserving map whose reduced Choi matrix is not proportional to the identity operator cannot be represented as $\lambda_+ \Lambda_+ - \lambda_- \Lambda_-$ for CPTP $\Lambda_\pm$. To circumvent this, we will employ the set of completely positive and trace–non-increasing maps, which can be understood as probabilistic implementations of quantum channels. Importantly, robustness-based definitions which were all equal in the case of states might not be equal any more. We therefore need to explicitly consider three different types of the robustness w.r.t. the sets CPTP or CPTNI:

$$R(\Phi) := \min \left\{ \lambda \ \middle| \ \frac{\Phi + \lambda \Lambda}{1 + \lambda} \in \text{CPTNI}, \ \Lambda \in \text{CPTNI} \right\}, \tag{6}$$

$$\begin{aligned} R'(\Phi) &:= \min \left\{ \lambda \ \middle| \ J_\Phi \leq (1 + \lambda)J_\Lambda, \ \Lambda \in \text{CPTNI} \right\} \\ &= \min \left\{ \lambda \ \middle| \ J_\Phi \leq (1 + \lambda)J_\Lambda, \ \Lambda \in \text{CPTP} \right\}, \end{aligned} \tag{7}$$

$$\begin{aligned} R''(\Phi) &:= \min \left\{ \lambda \ \middle| \ \Phi + \lambda \Lambda \in \text{CP}, \ \Lambda \in \text{CPTNI} \right\} \\ &= \min \left\{ \lambda \ \middle| \ \Phi + \lambda \Lambda \in \text{CP}, \ \Lambda \in \text{CPTP} \right\}, \end{aligned} \tag{8}$$

as well as a generalised notion of a base norm with respect to the set of completely positive and trace–non-increasing maps:

$$\|\Phi\|_\blacklozenge := \min \left\{ \lambda_+ + \lambda_- \ \middle|\ \Phi = \lambda_+ \Lambda_+ - \lambda_- \Lambda_-, \ \Lambda_\pm \in \text{CPTNI} \right\}. \tag{9}$$

In the expressions for R' and R'', we made use of the fact that one can, without loss of generality, restrict the optimisation to CPTP maps; this follows since for any $\Lambda \in \text{CPTNI}$ such that $\text{Tr}_B J_\Lambda \leq \mathbb{1}_A$ we can define the map Λ' by $J_{\Lambda'} = J_\Lambda + \frac{C}{d_B} \otimes \mathbb{1}_B$ where $C = \mathbb{1}_A - \text{Tr}_B J_\Lambda \geq 0$ which satisfies $\Lambda' \in \text{CPTP}$ and achieves the same value of the objective function. We note that closely related definitions were recently also considered in Ref. [33] for the case of trace-preserving maps.

All of the quantities above are well-defined and take a finite value for any map $\Phi \in \mathcal{H}(A, B)$, as we shall see explicitly by establishing general upper bounds in Sec. 6. The robustness $R(\Phi)$ can be seen to be an upper bound for all other quantities: any feasible decomposition of Φ in Eq. (6) gives feasible solutions for Eq. (7), (8), and for the base norm in Eq. (9). It is a priori unclear whether one can find general conditions under which the inequalities between the different measures are tight. We shall shortly see that equality indeed holds for all trace-preserving linear maps.

All of the introduced quantities can be computed as semidefinite programs, which follows since the constraints for a map to be CPTNI (or CPTP) are linear matrix inequalities. This means that the measures can be evaluated efficiently (in the dimensions of the map) using numerical software. The equivalent dual forms of the problems, which can also provide some insight into the differences between the different definitions of the robustness measures, will be reported shortly in Sec. 6.

3. Relation with the diamond norm

For any Hermiticity-preserving map Φ, the diamond norm (completely bounded trace norm) is defined as [19, 34]

$$\|\Phi\|_\diamond = \max_{\rho \in \mathbb{D}(A \otimes A)} \left\| \text{id}_A \otimes \Phi(\rho) \right\|_1, \tag{10}$$

where, in a slight abuse of notation, we use $\mathbb{D}(A \otimes A)$ to denote the states acting on a bipartite Hilbert space composed of the space A and another space isomorphic thereto.

The diamond norm finds use as a fundamental measure of distance between quantum channels, mirroring the operational role of the trace distance in measuring distances between quantum states [19, 34–36]. It is one of the most widely employed figures of merit in comparing quantum channels and benchmarking channel manipulation protocols. Its quantification and characterisation is therefore crucial to an effective understanding of the properties of quantum processes. Close connections between the diamond norm and the base norm in the space of quantum channels can be inferred already from the operational similarity that the diamond norm bears to the trace norm, the latter being the natural base norm in the space of quantum states. Here we aim to clarify the details of such connections and to explicitly relate the diamond norm with the robustness measures.

We will first introduce the following lemma, which establishes a useful formulation of the diamond norm for Hermiticity-preserving maps. The result is closely related to a more general approach for generalised quantum channels considered previously by Jenčová [37], and can be alternatively deduced from Lem. 4 and Thm. 2 of [37].

Lemma 1. For any Hermiticity-preserving map Φ, it holds that

$$\|\Phi\|_\diamond = \min \left\{ \mu \ \middle|\ J_\Phi = M_+ - M_-, \ M_\pm \geq 0, \ \text{Tr}_B(M_+ + M_-) \leq \mu \mathbb{1}_A \right\} \tag{11}$$

$$= \min \left\{ \mu \ \middle|\ J_\Phi = M_+ - M_-, \ M_\pm \geq 0, \ \text{Tr}_B(M_+ + M_-) = \mu \mathbb{1}_A \right\}. \tag{12}$$

Proof. Let $\|\Phi\|'_\diamond$ denote the quantity in (12). We first notice that the constraint $\text{Tr}_B(M_+ + M_-) = \mu \mathbb{1}_A$ can be relaxed to $\text{Tr}_B(M_+ + M_-) \leq \mu \mathbb{1}_A$ without loss of generality. This follows since for any feasible $M_\pm$ s.t. $\text{Tr}_B(M_+ + M_-) + C = \mu \mathbb{1}$ with $C \geq 0$, one can define feasible solutions $M'_\pm = M_\pm + \frac{C}{2d_B} \otimes \mathbb{1}_B$ which satisfy $\text{Tr}_B(M'_+ + M'_-) = \mu \mathbb{1}_A$ and thus achieve the same optimal value. We thus have

$$\|\Phi\|'_\diamond = \min \left\{ \mu \ \middle|\ J_\Phi = M_+ - M_-, \ M_\pm \geq 0, \ \text{Tr}_B(M_+ + M_-) \leq \mu \mathbb{1}_A \right\}$$
$$= \min \left\{ \|\text{Tr}_B(M_+ + M_-)\|_\infty \ \middle|\ J_\Phi = M_+ - M_-, \ M_\pm \geq 0 \right\}. \tag{13}$$

Taking the Lagrange dual of the above (see Appendix A) gives

$$
\begin{aligned}
\|\Phi\|_\diamond' &= \max\left\{\langle J_\Phi, W\rangle \,\big|\, -\rho\otimes\mathbb{1}_B \le W \le \rho\otimes\mathbb{1}_B,\ \rho\in\mathbb{D}(A)\right\} \\
&= \sup\left\{\langle J_\Phi, W\rangle \,\big|\, -\rho\otimes\mathbb{1}_B \le W \le \rho\otimes\mathbb{1}_B,\ \rho\in\mathbb{D}_{>0}(A)\right\} \\
&= \sup\left\{\left\langle W, \sqrt{\rho\otimes\mathbb{1}_B}\, J_\Phi \sqrt{\rho\otimes\mathbb{1}_B}\right\rangle \,\Big|\, \rho\in\mathbb{D}_{>0}(A),\ -\mathbb{1}_{A\otimes B}\le W\le\mathbb{1}_{A\otimes B}\right\} \\
&= \sup_{\rho\in\mathbb{D}_{>0}(A)}\left\|\sqrt{\rho\otimes\mathbb{1}_B}\, J_\Phi \sqrt{\rho\otimes\mathbb{1}_B}\right\|_1 \\
&= \max_{\rho\in\mathbb{D}(A)}\left\|\sqrt{\rho\otimes\mathbb{1}_B}\, J_\Phi \sqrt{\rho\otimes\mathbb{1}_B}\right\|_1,
\end{aligned}
\tag{14}
$$

where in the second line, by continuity, we restricted our attention to the set of full-rank states $\mathbb{D}_{>0}(A)$ without loss of generality, and in the third line we made the change of variables $W \mapsto \sqrt{\rho^{-1}\otimes\mathbb{1}_B}\, W \sqrt{\rho^{-1}\otimes\mathbb{1}_B}$. The fact that this equals the diamond norm of Φ can be deduced from the results of Ref. [38] already; for completeness, we will show this explicitly. Recalling that $J_\Phi = (\mathrm{id}_A\otimes\Phi)\,|\Omega\rangle\langle\Omega|$ with $|\Omega\rangle$ being the unnormalised maximally entangled state, and using the fact that Φ is only acting on one of the subsystems, we can write

$$
\begin{aligned}
\|\Phi\|_\diamond' &= \max_{\rho\in\mathbb{D}(A)}\left\|\sqrt{\rho\otimes\mathbb{1}_B}\,(\mathrm{id}_A\otimes\Phi)\,[|\Omega\rangle\langle\Omega|]\,\sqrt{\rho\otimes\mathbb{1}_B}\right\|_1 \\
&= \max_{\rho\in\mathbb{D}(A)}\left\|(\mathrm{id}_A\otimes\Phi)\left[\sqrt{\rho}\otimes\mathbb{1}_A\,|\Omega\rangle\langle\Omega|\,\sqrt{\rho}\otimes\mathbb{1}_A\right]\right\|_1 \\
&= \max_{\psi\in\mathbb{D}(A\otimes A)}\left\|\mathrm{id}_A\otimes\Phi(\psi)\right\|_1 \\
&= \max_{\rho\in\mathbb{D}(A\otimes A)}\left\|\mathrm{id}_A\otimes\Phi(\rho)\right\|_1 \\
&= \|\Phi\|_\diamond
\end{aligned}
\tag{15}
$$

where we used that any pure state $\psi\in\mathbb{D}(A\otimes A)$ can be written as $\left(\sqrt{\rho}\otimes\mathbb{1}_A\right)|\Omega\rangle\langle\Omega|\left(\sqrt{\rho}\otimes\mathbb{1}_A\right)$ for a suitable choice of $\rho\in\mathbb{D}(A)$, with $|\psi\rangle$ constituting the canonical purification of ρ. ∎

Compared with the semidefinite programs for the diamond norm of general linear maps originally derived in Refs. [38, 39], the form of the diamond norm presented in Lemma 1 already constitutes a major simplification — both at a conceptual level, allowing for a restatement of the problem in terms of optimising over decompositions of the form $J_\Phi = M_+ - M_-$, and computationally, as the number of optimisation variables is reduced.

As an immediate consequence of the above result, we can use the characterisation of the diamond norm in Eq. (11) to construct valid feasible solutions for the base norm and robustness measures in Eqs. (6)–(9), and vice versa.

Corollary 2. For any Hermiticity-preserving map Φ, it holds that

$$
\begin{aligned}
2\,\|\Phi\|_\diamond &\ge \|\Phi\|_\blacklozenge \ge \|\Phi\|_\diamond, \\
\|\Phi\|_\diamond + 1 &\ge R'(\Phi) \ge \frac{1}{2}\left(\|\Phi\|_\diamond - 2 + \lambda_{\min}(\mathrm{Tr}_B\, J_\Phi)\right) \\
\|\Phi\|_\diamond &\ge R''(\Phi) \ge \frac{1}{2}\left(\|\Phi\|_\diamond - \lambda_{\max}(\mathrm{Tr}_B\, J_\Phi)\right),
\end{aligned}
\tag{16}
$$

where $\lambda_{\min}$ and $\lambda_{\max}$ denote, respectively, the smallest and the largest eigenvalues.

Proof. Any decomposition for the diamond norm of the form $J_\Phi = M_+ - M_-$ with $\mathrm{Tr}_B(M_+ + M_-) \le \mu\mathbb{1}_A$ satisfies $\mathrm{Tr}_B M_\pm \le \mu\mathbb{1}_A$, which provides valid feasible solutions to the norm $\|\cdot\|_\blacklozenge$ and the robustness measures. On the other hand, any decomposition for $\|\cdot\|_\blacklozenge$ of the form $\Phi = \lambda_+\Lambda_+ - \lambda_-\Lambda_-$ with $\Lambda_\pm\in\mathrm{CPTNI}$ gives a feasible decomposition for the diamond norm with $\mathrm{Tr}_B(\lambda_+ J_{\Lambda_+} + \lambda_- J_{\Lambda_-}) \le (\lambda_+ + \lambda_-)\mathbb{1}_A$. Similarly, any decomposition for R' satisfying $J_\Phi \le (1+\lambda)J_\Lambda$ gives a feasible decomposition for $\|\cdot\|_\diamond$ of the form $J_\Phi = (1+\lambda)J_\Lambda - M_-$ where $M_- := (1+\lambda)J_\Lambda - J_\Phi$. Using that

$$
\begin{aligned}
\mathrm{Tr}_B\left[(1+\lambda)J_\Lambda + M_-\right] &= \mathrm{Tr}_B\left[2(1+\lambda)J_\Lambda - J_\Phi\right] \\
&\le \left[2(1+\lambda) - \lambda_{\min}(\mathrm{Tr}_B\, J_\Phi)\right]\mathbb{1}_A,
\end{aligned}
\tag{17}
$$

we get the stated bound. The case of R'' follows analogously. ∎

Equality between the different quantities can be shown for all trace-preserving maps, directly relating the diamond norm with our considered measures.

Theorem 3. For any map $\Phi \in \mathcal{H}(A, B)$ which is trace preserving or, more generally, proportional to a trace-preserving map in the sense that $\text{Tr}_B J_\Phi \propto \mathbb{1}$, it holds that

$$\|\Phi\|_\diamond = \|\Phi\|_\blacklozenge = \min \left\{ \mu_+ + \mu_- \mid J_\Phi = \mu_+ J_{\Lambda_+} - \mu_- J_{\Lambda_-}, \ \Lambda_\pm \in \text{CPTP}, \ \mu_\pm \in \mathbb{R}_+ \right\}. \quad (18)$$

For trace-preserving maps Φ, it additionally holds that

$$R(\Phi) = R'(\Phi) = R''(\Phi) = \frac{\|\Phi\|_\blacklozenge - 1}{2} = \frac{\|\Phi\|_\diamond - 1}{2}. \quad (19)$$

Proof. From the fact that $\text{Tr}_B J_\Phi = t \mathbb{1}_A$ for some $t \in \mathbb{R}$, it is easy to see that every decomposition of the form $J_\Phi = M_+ - M_-$, $M_\pm \geq 0$, $\text{Tr}_B(M_+ + M_-) = \mu \mathbb{1}_A$ as in Lemma 1 has to satisfy

$$\text{Tr}_B M_+ = \frac{\mu + t}{2} \mathbb{1}_A, \quad \text{Tr}_B M_- = \frac{\mu - t}{2} \mathbb{1}_A. \quad (20)$$

This implies that we can equivalently write

$$\|\Phi\|_\diamond = \min \left\{ \mu_+ + \mu_- \mid J_\Phi = \mu_+ M_+ - \mu_- M_-, \ M_\pm \geq 0, \ \text{Tr}_B M_+ = \text{Tr}_B M_- = \mathbb{1}_A \right\} \quad (21)$$

which is precisely Eq. (18). Notice then that any such decomposition gives a valid feasible solution for $\|\Phi\|_\blacklozenge$, together with Cor. 2 yielding equality between the two norms.

When Φ is trace preserving ($t = 1$), we can write

$$\begin{aligned} \|\Phi\|_\diamond &= \min \left\{ \mu_+ + \mu_- \mid J_\Phi = \mu_+ J_{\Lambda_+} - \mu_- J_{\Lambda_-}, \ \Lambda_\pm \in \text{CPTP} \right\} \\ &= \min \left\{ 2\mu_+ - 1 \mid J_\Phi = \mu_+ J_{\Lambda_+} - \mu_- J_{\Lambda_-}, \ \Lambda_\pm \in \text{CPTP} \right\} \\ &= \min \left\{ 2\mu_- + 1 \mid J_\Phi = \mu_+ J_{\Lambda_+} - \mu_- J_{\Lambda_-}, \ \Lambda_\pm \in \text{CPTP} \right\}. \end{aligned} \quad (22)$$

The equality $\|\Phi\|_\diamond = 2R'(\Phi) + 1 = 2R''(\Phi) + 1$ then follows: on the one hand, any decomposition of the form in Eq. (22) gives a feasible decomposition for R' and R'' in Eqs. (7)–(8), and on the other hand, any decomposition for the robustness measures is necessarily of the form in Eq. (22). Equality with the robustness $R(\Phi)$ follows by noting again that any feasible decomposition in Eq. (22) gives a feasible decomposition for $R(\Phi)$, and on the other hand using the relation $R(\Phi) \geq R'(\Phi)$ which holds by definition. ∎

Remark. The expression in Eq. (18) is valid also in the case of trace-annihilating maps ($\text{Tr}_B J_\Phi = 0$), and thus the case of computing the distance $\|\Lambda - \Lambda'\|_\diamond$ between two quantum channels. A simplified expression for this problem appeared previously in [38] and was explicitly expressed as a robustness-type measure in [26].

We note that the quantity $\|\cdot\|_\blacklozenge$, applied to trace-preserving maps, was recently considered in Refs. [33] and [40]. It was not noticed in these works that this is simply the diamond norm, and hence many results shown in [33] (e.g. the multiplicativity with respect to tensor product, unitary invariance, bounds with trace norm $\|J_\Phi\|_1$, monotonicity under the action of superchannels, and some explicit expressions) follow directly from known properties of the diamond norm [20, 38, 41, 42].

We will later see that this equivalence does not extend to maps which are not trace preserving (or proportional thereto), and indeed we can have $\|\Phi\|_\blacklozenge = 2 \|\Phi\|_\diamond$ in the extreme case.

4. Quantifying simulation cost

Since the quantum dynamics which can be realised in practice are restricted to completely positive maps, a relevant question then becomes: how can one simulate the action of a non-CPTP map on a quantum state when only CPTP maps are available to us?

A similar question was recently asked in Ref. [33], where the authors applied quasiprobability sampling methods [17, 30, 43] to the desired operation Φ. We take a different approach here and instead allow for the use of an ancillary system X, which can be an affine combination of quantum states, in order to simulate the action of the map Φ as a CPTP map Λ acting jointly on the input quantum state and the ancilla X. The "non-physicality" of the given map Φ is

then pushed into the system X, allowing for the overall transformation Λ to be a valid quantum channel.

The motivation for this approach is that the task of simulating the action of the non-CPTP map Φ is effectively replaced with the simulation of a unit-trace Hermitian operator X, which could be significantly easier to realise in practice, especially since we will see that the dimension of the ancilla can be taken to be arbitrarily small. Standard quasiprobability-based approaches such as the ones employed in [17, 30, 33] aim to estimate the expectation value $\mathrm{Tr}[\Phi(\rho)A]$, where Φ is a non-CPTP map and A an observable, by decomposing the given map as $\Phi = \lambda_i \Lambda_i$ with $\lambda_i \in \mathbb{R}$ and $\Lambda_i \in \mathrm{CPTP}$ (or CPTNI). The expectation value $\mathrm{Tr}[\Phi(\rho)A]$ is then estimated by evaluating $\mathrm{Tr}[\Lambda_i(\rho)A]$ and appropriately sampling from the output distributions with probabilities determined by the coefficients λ_i [17, 43]. In practice, this means that we have to repeatedly realise each operation Λ_i, which requires the implementation of a different quantum circuit for each operation. Consider, on the other hand, a situation in which the dynamics is fixed as some map $\Lambda \in \mathrm{CPTNI}$, and we only need to vary the input states. This can be achieved by writing $\Phi(\cdot) = \Lambda(\cdot \otimes X)$, where we can write any Hermitian operator in a quasiprobability representation as $X = \sum_i \mu_i \rho_i$. The task of sampling from the output distribution is then reduced to feeding in the different states ρ_i into the circuit which realises the fixed operation Λ, thus greatly simplifying the implementation.

As mentioned in Sec. 2, a natural quantifier of how much a given operator $X \in \mathbb{H}(A)$ with $\mathrm{Tr}(X) = 1$ deviates from the set of all quantum states is the trace norm $\|X\|_1$. Indeed, this quantity can be given an explicit interpretation in terms of the optimal cost of a quasiprobability-based estimation of the expectation value of X [43]. We then define the simulation cost of a map as the minimal amount of such "non-physicality" of X needed to simulate the action of the map:

$$S(\Phi) := \min \left\{ \|X\|_1 \ \middle| \ \Lambda(\cdot \otimes X) = \Phi(\cdot), \ \mathrm{Tr}(X) = 1, \ \Lambda \in \mathrm{CPTNI} \right\}. \tag{23}$$

We then have the following.

Theorem 4. For any map $\Phi \in \mathcal{H}(A, B)$, it holds that

$$S(\Phi) = 2R(\Phi) + 1. \tag{24}$$

In the case of a trace-preserving Φ, we have in particular that

$$S(\Phi) = \|\Phi\|_\diamond \tag{25}$$

and an optimal Λ for the simulation can be chosen to satisfy $\Lambda \in \mathrm{CPTP}$.

Proof. Let $\Lambda_\pm \in \mathrm{CPTNI}(A, B)$ be maps that achieve an optimal decomposition for Φ such that $\Phi = (1 + R(\Phi))\Lambda_+ - R(\Phi)\Lambda_-$. Now, consider a non-positive Hermitian operator $X = \mu_+ \omega_+ - \mu_- \omega_- \in \mathbb{H}(A')$ where $\omega_\pm$ are orthogonal quantum states and $\mathrm{Tr}[X] = \mu_+ - \mu_- = 1$. We do not impose any additional conditions on the size of the ancillary system A', meaning that its Hilbert space can be chosen to be an arbitrary space of dimension at least 2. Defining the projector onto the positive part of X as P_+, we then consider the map defined by the action on a basis $|i_1\rangle\langle j_1| \otimes |i_2\rangle\langle j_2| \in \mathbb{L}(A \otimes A')$ as follows:

$$\Lambda(|i_1\rangle\langle j_1| \otimes |i_2\rangle\langle j_2|) := \mathrm{Tr}\left[\frac{P_+}{\mu_+}|i_2\rangle\langle j_2|\right]\Phi(|i_1\rangle\langle j_1|) + \mathrm{Tr}\left[\left(\mathbb{1} - \frac{P_+}{\mu_+}\right)|i_2\rangle\langle j_2|\right]\Lambda_-(|i_1\rangle\langle j_1|)$$

$$= (1 + R(\Phi))\,\mathrm{Tr}\left[\frac{P_+}{\mu_+}|i_2\rangle\langle j_2|\right]\Lambda_+(|i_1\rangle\langle j_1|) \tag{26}$$

$$+ \mathrm{Tr}\left[\left(\mathbb{1} - (1 + R(\Phi))\frac{P_+}{\mu_+}\right)|i_2\rangle\langle j_2|\right]\Lambda_-(|i_1\rangle\langle j_1|)$$

It is easy to check that $\Lambda(\rho \otimes X) = \Phi(\rho)$. Now, we will show that as long as the condition

$$(1 + R(\Phi))\,\mathrm{Tr}\left[\frac{P_+}{\mu_+}\rho\right] \leq 1 \ \forall \rho \tag{27}$$

is satisfied, then Λ is also CPTNI. This can be seen by observing first that (27) gives

$$0 \leq \tilde{P} := (1 + R(\Phi))\frac{P_+}{\mu_+} \leq \mathbb{1}, \tag{28}$$

which implies that $\tilde{P}$ is a valid POVM element. Note that we can rewrite (26) as

$$\Lambda = \Lambda_+ \otimes T_{\tilde{P}} + \Lambda_- \otimes T_{\mathbb{1}-\tilde{P}} \tag{29}$$

$$= (\mathbb{1} \otimes T_{\tilde{P}})(\Lambda_+ \otimes \mathbb{1}) + (\mathbb{1} \otimes T_{\mathbb{1}-\tilde{P}})(\Lambda_- \otimes \mathbb{1}) \tag{30}$$

where $T_P(\cdot) := \mathrm{Tr}\,[P\cdot]$. Since Λ_+, Λ_-, and $T_{\tilde{P}}$, $T_{\mathbb{1}-\tilde{P}}$ are all completely positive, Λ is also completely positive. Since (27) is always satisfied when

$$\mu_+ \geq 1 + R(\Phi), \tag{31}$$

an operator X with $\|X\|_1 = \mu_+ + \mu_- = 1 + 2R(\Phi)$ achieves the desired implementation.

The converse part can be proven by extending an argument in Ref. [23] to our setting. Suppose a non-quantum resource $X = \mu_+ \omega_+ - \mu_- \omega_-$ and CPTNI map Λ realise the simulation of Φ, i.e. $\Phi(\cdot) = \Lambda(\cdot \otimes X)$. Also, define $\Lambda_+(\cdot) := \Lambda(\cdot \otimes \omega_+)$. Then, by linearity of Λ, we get

$$\Lambda_+(\cdot) = \frac{1}{\mu_+}\Lambda(\cdot \otimes X) + \frac{1}{\mu_+}\Lambda(\cdot \otimes \mu_-\omega_-) \tag{32}$$

$$= \frac{1}{\mu_+}\Phi(\cdot) + \frac{\mu_-}{\mu_+}\Lambda(\cdot \otimes \omega_-). \tag{33}$$

Since Λ_+ and $\Lambda_- := \Lambda(\cdot \otimes \omega_-)$ are CPTNI maps, this is a valid linear decomposition of Φ into two CPTNI maps, providing an upper bound for its robustness as $R(\Phi) \leq \mu_- = \mu_+ - 1$. This gives the desired lower bound for the simulation cost as $\mu_+ + \mu_- \geq 1 + 2R(\Phi)$. ∎

An interesting quantitative equivalence emerges between our approach and the method of Ref. [33]. In that work, the authors showed that the minimal overhead required to employ quasiprobability-based simulation techniques [17, 43] to estimate $\mathrm{Tr}[\Phi(\rho)A]$ for a trace-preserving map Φ scales with the norm $\|\Phi\|_\blacklozenge$ (see also the discussion in Sec. 7.2). Since we know from Thm. 3 that

$$2R(\Phi) + 1 = \|\Phi\|_\blacklozenge = \|\Phi\|_\diamond \tag{34}$$

holds for any trace-preserving map, the quantitative cost of the simulation scheme is actually the same as our method, despite the seemingly different approaches employed. In fact, our Thm. 4 shows that it is sufficient to consider decompositions of Φ as

$$\Phi(\cdot) = \mu_+ \Lambda(\cdot \otimes \omega_+) - \mu_- \Lambda(\cdot \otimes \omega_-) \tag{35}$$

where Λ and $X = \mu_+\omega_+ - \mu_-\omega_-$ are as constructed in our protocol. This means that, despite the significant practical simplification obtained by fixing the dynamics of the simulator as Λ and optimising over the quasiprobability representations of X instead, our simulation method does not sacrifice any performance, and the optimal sampling overhead cost of the more direct approach of [33] cannot be any better.

We note that Theorem 4 gives a general way of reducing the task of simulating the action of a linear map Φ to simulating an affine combination of states in the form of the operator X. This could provide methods for the simulation of dynamics even beyond quasiprobability-based approaches like the one discussed above, although the specifics of this will depend on the given simulation method.

State injection and resource simulation. The setting considered here is closely related to state injection methods which generalise quantum teleportation [44] and find use e.g. in the resource theories of entanglement [45–49], stabiliser-state quantum computation [50, 51], and coherence [23, 52]. In such tasks, a resourceful state ϕ (such as a maximally entangled singlet) is used to simulate the action of an arbitrary quantum channel Θ as $\Theta(\cdot) = \Gamma(\cdot \otimes \phi)$, where now Γ is a free operation (such as a protocol consisting of local operations and classical communication only). In this sense, our result can be thought of as the cost of channel simulation in the resource theory of "non-physicality" beyond quantum mechanics, with the operator X acting as a resource. There are many potential ways to interpret such a result: for instance, unit-trace Hermitian operators which are not necessarily positive semidefinite have found use as so-called pseudo-states in [53], where they were used to study correlations beyond quantum mechanics, and as so-called pseudo-density matrices in [54], where they were used to put spatial and temporal correlations on equal footing. Being able to use a Hermitian system X could then be interpreted as having access to such extended sets of correlations. We leave a precise investigation of the connections between the operational setting employed here and resource theories of correlations for future work.

Amortised simulation. A related setting that we can consider is that of *amortised* simulation [23, 55], in which the non-quantum resource X is not consumed completely, but instead we can recover some of it in the form of another resource Y which can be reused. Precisely, we define

$$S_A(\Phi) = \min \left\{ \frac{\|X\|_1}{\|Y\|_1} \, \middle| \, \Lambda(\cdot \otimes X) = \Phi(\cdot) \otimes Y, \ \mathrm{Tr}(X) = \mathrm{Tr}(Y) = 1, \ \Lambda \in \mathrm{CPTNI} \right\}. \tag{36}$$

Clearly, $S_A(\Phi) \le S(\Phi)$ as we can just take X to be optimal for S and Y to be the trivial system 1. One could expect amortisation to lead to a strictly smaller cost of simulating a given map. However, we can show that this is not the case — amortisation cannot improve the simulation cost of any trace-preserving map.

Corollary 5. For any trace-preserving map $\Phi \in \mathcal{H}(A, B)$, it holds that

$$S_A(\Phi) = S(\Phi) = \|\Phi\|_\diamond . \tag{37}$$

Proof. Let Λ be the optimal map such that $\Lambda(\cdot \otimes X) = \Phi(\cdot) \otimes Y$ with $S_A(\Phi) = \|X\|_1 / \|Y\|_1$. Noting that this can be alternatively understood as a simulation protocol for the trace-preserving map $\Phi(\cdot) \otimes Y$, Thm. 4 tells us that any such protocol satisfies

$$\begin{aligned} \|X\|_1 &\ge S\left(\Phi(\cdot) \otimes Y\right) \\ &= \|\Phi(\cdot) \otimes Y\|_\diamond \\ &= \|\Phi\|_\diamond \|Y\|_1 \\ &= S(\Phi) \|Y\|_1 \end{aligned} \tag{38}$$

where we used the multiplicativity of the diamond norm and the fact that $\|Y\|_\diamond = \|Y\|_1$ where we treat Y as a preparation channel with a trivial input space. From this we have that $S_A(\Phi) \ge S(\Phi)$, which concludes the proof. $\blacksquare$

5. Quantifying advantages in quantum games

The study of general linear maps in a resource-theoretic setting motivates the question: is there a well-defined operational task in which having access to *any* non-CPTP map could provide practical advantages over all quantum channels?

In order to give an instance of such a task, we consider the setting of input-output games, inspired by the work of Ref. [56] and studied in the context of dynamical quantum resources in [27, 28]. The setting is as follows: Alice prepares a state chosen randomly from the ensemble $\{p_i, \sigma_i\}_i$ and sends the state through the map $\Phi \in \mathcal{H}(A, B)$ to Bob, who then measures with a POVM $\{M_j\}_j$. The players are then awarded a score based on a reward function characterised by the coefficients $\{w_{ij}\}_{i,j} \in \mathbb{R}$, and their goal is to maximise the average payoff given by

$$P(\Phi, \{p_i, \sigma_i\}, \{M_j\}, \{w_{ij}\}) = \sum_{i,j} w_{ij} p_i \left\langle M_j, \Phi(\sigma_i) \right\rangle \tag{39}$$

by a suitable choice of the states and measurements. The tuple $\mathcal{G} = (\{p_i, \sigma_i\}, \{M_j\}, \{w_{ij}\})$ then defines the *input-output game* $\mathcal{G}$.

We stress that, although the payoff $P(\Phi, \mathcal{G})$ might lose its physical meaning as a discrimination task when Φ is an arbitrary linear map, already for a *positive* trace-preserving map Φ we have that every output $\Phi(\sigma_i)$ is indeed a valid density matrix and thus the measurement at the output constitutes a well-defined state discrimination task.

We are then interested in quantifying the best possible advantage that a given map Φ could provide over CPTP maps. Such an optimisation is unbounded without any further constraints, so we will consider games for which any completely positive map Γ achieves a non-negative payoff value — this can always be ensured by suitably shifting the payoff function for a given game. We then have the following.

Theorem 6. For any map $\Phi \in \mathcal{H}(A, B)$, it holds that

$$\sup_{\mathcal{G}} \frac{P(\Phi, \mathcal{G})}{\max_{\Lambda \in \mathrm{CPTP}} P(\Lambda, \mathcal{G})} = R'(\Phi) + 1 \tag{40}$$

where the maximisation is over all input-output games $\mathcal{G}$ such that $P(\Gamma, \mathcal{G}) \geq 0 \ \forall \Gamma \in \mathrm{CP}$. In the case of a trace-preserving Φ, we have in particular that

$$\sup_{\mathcal{G}} \frac{P(\Phi, \mathcal{G})}{\max_{\Lambda \in \mathrm{CPTP}} P(\Lambda, \mathcal{G})} = \frac{\|\Phi\|_{\diamond} + 1}{2}, \tag{41}$$

and it suffices to optimise over games such that $P(\Lambda, \mathcal{G}) \geq 0 \ \forall \Lambda \in \mathrm{CPTP}$.

Proof. Any Φ can be written as $\Phi = (1 + R'(\Phi))\Lambda - \Gamma$ where $\Lambda \in \mathrm{CPTP}, \Gamma \in \mathrm{CP}$. On the one hand, we then have for any $\mathcal{G}$ that

$$\begin{aligned} P(\Phi, \mathcal{G}) &= (1 + R'(\Phi)) P(\Lambda, \mathcal{G}) - P(\Gamma, \mathcal{G}) \\ &\leq (1 + R'(\Phi)) P(\Lambda, \mathcal{G}) \\ &\leq (1 + R'(\Phi)) \max_{\Lambda \in \mathrm{CPTP}} P(\Lambda, \mathcal{G}) \end{aligned} \tag{42}$$

where the first inequality follows since $P(\Gamma, \mathcal{G}) \geq 0 \ \forall \Gamma \in \mathrm{CP}$, which shows that the left-hand side of Eq. (40) is upper-bounded by the right-hand side. By Thm. 3, in the case of a trace preserving map Φ we can equivalently write $\Phi = (1 + R'(\Phi))\Lambda_+ - R'(\Phi)\Lambda_-$ where $\Lambda_\pm \in \mathrm{CPTP}$, so one only needs to consider games such that $P(\Lambda, \mathcal{G}) \geq 0 \ \forall \Lambda \in \mathrm{CPTP}$.

On the other hand, by strong Lagrange duality (see App. A) we can write

$$R'(\Phi) + 1 = \max\left\{ \langle W, J_\Phi \rangle \ \middle| \ 0 \leq W \leq \rho \otimes \mathbb{1}_B, \ \rho \in \mathbb{D}(A) \right\}. \tag{43}$$

We can then make the following observations. Firstly, since the set of separable states in $\mathbb{D}(A \otimes B)$ has a non-empty interior [57], any Hermitian operator X can be written as $X = \sum_{i=1}^n x_i \sigma_i \otimes \eta_i$ for some $\sigma_i \in \mathbb{D}(A), \eta_i \in \mathbb{D}(B), x_i \in \mathbb{R}$, and $n \in \mathbb{N}$. Then, choose the optimal W in Eq. (43) and write $W^{T_A} = \sum_{i=1}^n x_i \sigma_i \otimes \eta_i$, where T_A denotes the partial transpose. Defining the set $\{M_i\}_{i=1}^{n+1}$ by $M_i := \eta_i / \left\| \sum_j \eta_j \right\|_\infty$ for $i \leq n$ and $M_{n+1} = \mathbb{1} - \sum_{i=1}^n M_i$, we have that

$$W = \sum_i p_i w_i \, \sigma_i^T \otimes M_i \tag{44}$$

where $p_i = 1/n$ for $i \leq n$ and $p_{n+1} = 0$, and the coefficients w_i are defined by $w_i = x_i n \left\| \sum_j \eta_j \right\|_\infty$. By the Choi-Jamiołkowski isomorphism and the linearity of Φ, we then have for Φ that

$$\langle W, J_\Phi \rangle = \sum_i p_i w_i \langle M_i, \Phi(\sigma_i) \rangle = P(\Phi, \mathcal{G}') \tag{45}$$

with $\mathcal{G}'$ defined by the above choices of $\{p_i, \sigma_i\}, \{M_i\}$, and $\{w_i\}$. Noticing that $W \geq 0 \implies P(\Gamma, \mathcal{G}') \geq 0 \ \forall \Gamma \in \mathrm{CP}$, this finally gives

$$\begin{aligned} \sup_{\mathcal{G}} \frac{P(\Phi, \mathcal{G})}{\max_{\Lambda \in \mathrm{CPTP}} P(\Lambda, \mathcal{G})} &\geq \frac{P(\Phi, \mathcal{G}')}{\max_{\Lambda \in \mathrm{CPTP}} P(\Lambda, \mathcal{G}')} \\ &\geq \frac{R'(\Phi) + 1}{\max_{\Lambda \in \mathrm{CPTP}} R'(\Lambda) + 1} \\ &= R'(\Phi) + 1 \end{aligned} \tag{46}$$

where the second inequality follows since $P(\Phi, \mathcal{G}') = \langle W, J_\Phi \rangle = R'(\Phi) + 1$ holds by assumption while $P(\Lambda, \mathcal{G}') = \langle W, J_\Lambda \rangle \leq R'(\Lambda) + 1$ holds for any map Λ by definition, and the last equality follows since $R'(\Lambda) = 0$ for any $\Lambda \in \mathrm{CPTP}$. ∎

6. General bounds

Useful bounds for the measures can be obtained by relating them with norms or quantities computed at the level of the Choi operator J_Φ, avoiding an optimisation over all CPTNI or CPTP maps. For instance, the following relation with the trace norm generalises known bounds for the diamond norm [42, 58] (see also [33]).

Proposition 7. For any Hermiticity-preserving map $\Phi \in \mathcal{H}(A, B)$, decompose J_Φ into its positive and negative parts as $J_\Phi = J_{\Phi_+} - J_{\Phi_-}$ with $J_{\Phi_\pm} \geq 0$. Then

$$\|J_\Phi\|_1 \geq \|\Phi\|_\blacklozenge \geq \frac{1}{d_A} \|J_\Phi\|_1 \,,$$

$$\max\left\{ \operatorname{Tr} J_{\Phi_+} - 1, \operatorname{Tr} J_{\Phi_-} \right\} \geq R(\Phi) \geq \max\left\{ \frac{1}{d_A} \operatorname{Tr} J_{\Phi_+} - 1, \frac{1}{d_A} \operatorname{Tr} J_{\Phi_-} \right\}, \tag{47}$$

$$\operatorname{Tr} J_{\Phi_+} - 1 \geq R'(\Phi) \geq \frac{1}{d_A} \operatorname{Tr} J_{\Phi_+} - 1,$$

$$\operatorname{Tr} J_{\Phi_-} \geq R''(\Phi) \geq \frac{1}{d_A} \operatorname{Tr} J_{\Phi_-}.$$

Proof. Consider $\|\cdot\|_\blacklozenge$ first. Using the expression

$$\|J_\Phi\|_1 = \min\left\{ \mu_+ + \mu_- \,\big|\, J_\Phi = \mu_+ \omega_+ - \mu_- \omega_-, \ \omega_\pm \in \mathbb{D}(A \otimes B) \right\} \tag{48}$$

we see that any such decomposition provides a feasible solution for $\|\Phi\|_\blacklozenge$, since $\omega_\pm$ constitute valid Choi operators of maps $\Omega_\pm \in \mathrm{CPTNI}(A, B)$. The first inequality thus follows. The second inequality is a consequence of the bound $\|\Phi\|_\blacklozenge \geq \|\Phi\|_\diamond$ from Cor. 2 and the fact that $\frac{1}{d_A}\|J_\Phi\|_1$ is known to lower bound the diamond norm (see e.g. [42, 58]). It can also be explicitly seen by noting that any decomposition of the form $J_\Phi = \lambda_+ J_{\Lambda_+} - \lambda_- J_{\Lambda_-}$ with $\Lambda_\pm \in \mathrm{CPTNI}$ can provide a decomposition for the trace norm by rescaling each $J_{\Lambda_\pm}$ by its trace; specifically,

$$\|J_\Phi\|_1 \leq \lambda_+ \operatorname{Tr} J_{\Lambda_+} + \lambda_- \operatorname{Tr} J_{\Lambda_-}, \tag{49}$$

and using the fact that $\operatorname{Tr} J_{\Lambda_\pm} \leq d_A \|J_{\Lambda_\pm}\|_\infty \leq d_A$ gives the desired bound.

The case of the robustness measures R', R'' follows analogously, where we now use the fact that $\operatorname{Tr} X_+ = \min\left\{ \mu \,\big|\, X \leq \mu\rho, \ \rho \in \mathbb{D} \right\}$ and $\operatorname{Tr} X_- = \min\left\{ \mu \,\big|\, X + \mu\rho \geq 0, \ \rho \in \mathbb{D} \right\}$ for any Hermitian X. For the robustness R, take λ to be the greater of $\operatorname{Tr} J_{\Phi_+} - 1$ and $\operatorname{Tr} J_{\Phi_-}$, and write

$$J_\Phi + \lambda \frac{J_{\Phi_-}}{\lambda} = (1 + \lambda) \frac{J_{\Phi_+}}{1 + \lambda}. \tag{50}$$

Since each $\frac{J_{\Phi_\pm}}{\lambda} \in \mathrm{CPTNI}$, this provides a valid feasible solution for R. On the other hand, $R(\Phi) \geq \max\{R'(\Phi), R''(\Phi)\}$ by definition, from which the lower bound follows. ∎

Both the upper and the lower bounds in Prop. 7 can be tight, as was shown already for the diamond norm [41]. However, better upper bounds can be obtained as follows.

Proposition 8. For any Hermiticity-preserving map $\Phi \in \mathcal{H}(A, B)$, it holds that

$$\begin{aligned}
\|\Phi\|_\diamond &\leq \lambda_{\max}(\operatorname{Tr}_B[J_{\Phi_+} + J_{\Phi_-}]), \\
\|\Phi\|_\blacklozenge &\leq \lambda_{\max}(\operatorname{Tr}_B J_{\Phi_+}) + \lambda_{\max}(\operatorname{Tr}_B J_{\Phi_-}), \\
R(\Phi) &\leq \max\left\{ \lambda_{\max}(\operatorname{Tr}_B J_{\Phi_+}) - 1, \lambda_{\max}(\operatorname{Tr}_B J_{\Phi_-}) \right\}, \\
R'(\Phi) &\leq \lambda_{\max}(\operatorname{Tr}_B J_{\Phi_+}) - 1, \\
R''(\Phi) &\leq \lambda_{\max}(\operatorname{Tr}_B J_{\Phi_-}).
\end{aligned} \tag{51}$$

We note that the bound for the diamond norm, which we stated above for completeness, appeared previously in [42].

Proof. The bounds for $\|\cdot\|_\diamond$, $\|\cdot\|_\blacklozenge$, R' and R'' follow simply by using $J_\Phi = J_{\Phi_+} - J_{\Phi_-}$ as feasible solutions in the definitions.

For the robustness R, take λ to be the greater of $\lambda_{\max}(\operatorname{Tr}_B J_{\Phi_+}) - 1$ and $\lambda_{\max}(\operatorname{Tr}_B J_{\Phi_-})$, and write

$$J_\Phi + \lambda \frac{J_{\Phi_-}}{\lambda} = (1 + \lambda) \frac{J_{\Phi_+}}{1 + \lambda}. \tag{52}$$

Since this is a feasible solution for R, we get $R(\Phi) \leq \lambda$. ∎

As for lower bounds, we will first need to establish dual expressions for the considered measures. The following Proposition is an application of standard convex duality arguments, and we include details in Appendix A for completeness.

Proposition 9. For any $\Phi \in \mathcal{H}(A, B)$, the following dual expressions hold.

$$\|\Phi\|_\diamond = \max \left\{ \langle J_\Phi, W \rangle \,\middle|\, -\rho \otimes \mathbb{1}_B \leq W \leq \rho \otimes \mathbb{1}_B, \ \rho \in \mathbb{D}(A) \right\}$$

$$\|\Phi\|_\blacklozenge = \max \left\{ \langle J_\Phi, W \rangle \,\middle|\, -\rho \otimes \mathbb{1}_B \leq W \leq \sigma \otimes \mathbb{1}_B, \ \rho, \sigma \in \mathbb{D}(A) \right\}$$

$$R(\Phi) = \max \left\{ \langle J_\Phi, W \rangle - \operatorname{Tr} Y \,\middle|\, -X \otimes \mathbb{1}_B \leq W \leq Y \otimes \mathbb{1}_B, \ X, Y \geq 0, \ \operatorname{Tr}(X + Y) = 1 \right\}$$

$$R'(\Phi) = \max \left\{ \langle J_\Phi, W \rangle - 1 \,\middle|\, 0 \leq W \leq \rho \otimes \mathbb{1}_B, \ \rho \in \mathbb{D}(A) \right\}$$

$$R''(\Phi) = \max \left\{ \langle J_\Phi, W - \rho \otimes \mathbb{1}_B \rangle \,\middle|\, 0 \leq W \leq \rho \otimes \mathbb{1}_B, \ \rho \in \mathbb{D}(A) \right\}$$

$$= \max \left\{ \langle J_\Phi, W \rangle - \operatorname{Tr} \Phi(\rho) \,\middle|\, 0 \leq W \leq \rho \otimes \mathbb{1}_B, \ \rho \in \mathbb{D}(A) \right\}. \tag{53}$$

We can then obtain lower bounds by employing the dual optimisation problems. The bound for the diamond norm is well known [20], but we find it is insightful to rederive it using this approach[1].

Proposition 10. For any $\Phi \in \mathcal{H}(A, B)$ and any input state $\rho \in \mathbb{D}(A)$, let $\Phi(\rho)_\pm$ denote the positive/negative part of the output operator $\Phi(\rho)$. Then

$$\|\Phi\|_\blacklozenge \geq \|\Phi\|_\diamond \geq \max \left\{ \left\| \Phi(\rho) \right\|_1 \,\middle|\, \rho \in \mathbb{D}(A) \right\}$$

$$\geq \|\operatorname{Tr}_B J_\Phi\|_\infty$$

$$\|\Phi\|_\blacklozenge \geq \max \left\{ \operatorname{Tr} \Phi(\rho)_- + \operatorname{Tr} \Phi(\sigma)_+ \,\middle|\, \rho, \sigma \in \mathbb{D}(A) \right\}$$

$$\geq \lambda_{\max}[(\operatorname{Tr}_B J_\Phi)_+] + \lambda_{\max}[(\operatorname{Tr}_B J_\Phi)_-]$$

$$R(\Phi) \geq \max \left\{ \operatorname{Tr} \Phi(X)_- + \operatorname{Tr} \Phi(Y)_+ - \operatorname{Tr} Y \,\middle|\, X, Y \geq 0, \ \operatorname{Tr}(X + Y) = 1 \right\}$$

$$\geq \max \left\{ \lambda_{\max}[(\operatorname{Tr}_B J_\Phi)_+] - 1, \ \lambda_{\max}[(\operatorname{Tr}_B J_\Phi)_-] \right\} \tag{54}$$

$$R'(\Phi) \geq \max \left\{ \operatorname{Tr} \Phi(\rho)_+ - 1 \,\middle|\, \rho \in \mathbb{D}(A) \right\}$$

$$\geq \lambda_{\max}[(\operatorname{Tr}_B J_\Phi)_+] - 1$$

$$R''(\Phi) \geq \max \left\{ \operatorname{Tr} \Phi(\rho)_- \,\middle|\, \rho \in \mathbb{D}(A) \right\}$$

$$\geq \lambda_{\max}[(\operatorname{Tr}_B J_\Phi)_-].$$

Proof. Consider the diamond norm first. The main idea is to restrict the optimisation in the dual expression of $\|\cdot\|_\diamond$ in (53) to operators of the form $W = \rho \otimes Z$ for some operator Z. Then we have

$$\|\Phi\|_\diamond \geq \max \left\{ \langle J_\Phi, \rho \otimes Z \rangle \,\middle|\, -\mathbb{1}_B \leq Z \leq \mathbb{1}_B, \ \rho \in \mathbb{D}(A) \right\}$$

$$= \max \left\{ \langle \Phi(\rho^T), Z \rangle \,\middle|\, -\mathbb{1}_B \leq Z \leq \mathbb{1}_B, \ \rho \in \mathbb{D}(A) \right\} \tag{55}$$

$$= \max \left\{ \left\| \Phi(\rho^T) \right\|_1 \,\middle|\, \rho \in \mathbb{D}(A) \right\},$$

where the second line follows by the Choi-Jamiołkowski isomorphism. Taking $Z \in \{\mathbb{1}, -\mathbb{1}\}$, we get the lower bound

$$\|\Phi\|_\diamond \geq \max \left\{ \pm \langle \operatorname{Tr}_B J_\Phi, \rho \rangle \,\middle|\, \rho \in \mathbb{D}(A) \right\}$$

$$= \|\operatorname{Tr}_B J_\Phi\|_\infty. \tag{56}$$

In the case of $\|\cdot\|_\blacklozenge$, we use feasible solutions of the form $W = \rho \otimes Z + \sigma \otimes V$ with $-\mathbb{1} \leq Z \leq 0$ and $0 \leq V \leq \mathbb{1}$ to obtain the stated bound analogously — the crucial observation being that $\max \left\{ \langle A, B \rangle \,\middle|\, 0 \leq B \leq \mathbb{1} \right\} = \operatorname{Tr} A_+$ for any Hermitian A. The other measures follow in the same way. ∎

Note the similarity between the eigenvalue-based lower bounds of Prop. 10 and the upper bounds of Prop. 8: the upper bounds consider the eigenvalues after decomposing J_Φ as $J_{\Phi_+} - J_{\Phi_-}$, while the lower bounds use the positive and negative parts of $\operatorname{Tr}_B J_\Phi$.

---

[1] We remark the curious fact that, despite the apparent similarity, the bound for the diamond norm in Prop. 10 is not the induced Schatten norm $\|\cdot\|_{1 \to 1}$, as the latter requires an optimisation over non-Hermitian input operators even when the map Φ is Hermiticity-preserving [20].

An immediate consequence is that for any completely positive map Φ, it holds that

$$\|\Phi\|_\blacklozenge = \|\Phi\|_\diamond = \|\mathrm{Tr}_B J_\Phi\|_\infty \tag{57}$$

since the operators J_Φ and $\mathrm{Tr}_B J_\Phi$ are both positive semidefinite. However, the lower bounds allow us to show explicitly that the equality $\|\Phi\|_\blacklozenge = \|\Phi\|_\diamond$ is no longer true for maps which are neither CP nor trace preserving, and in fact the extreme disparity of $\|\Phi\|_\blacklozenge = 2\|\Phi\|_\diamond$ (cf. Cor. 2) can be achieved. Consider for instance the case when

$$\Phi(\cdot) = \langle 0|\cdot|0\rangle\, |0\rangle\langle 0| - \langle 1|\cdot|1\rangle\, |1\rangle\langle 1|. \tag{58}$$

Decomposing $J_\Phi = |0\rangle\langle 0| \otimes |0\rangle\langle 0| - |1\rangle\langle 1| \otimes |1\rangle\langle 1|$ into its positive and negative parts, the bound of Prop. 8 gives $\|\Phi\|_\diamond \leq 1$. However, the best upper bound we get for $\|\Phi\|_\blacklozenge$ is 2, and it is indeed tight: we have $\Phi(|0\rangle\langle 0|) = |0\rangle\langle 0|$ and $\Phi(|1\rangle\langle 1|) = -|1\rangle\langle 1|$, and so Prop. 10 gives $\|\Phi\|_\blacklozenge \geq 2$. A similar argument can be used to show that $R(\Phi) = 1$, which in particular implies that $2R(\Phi) + 1 > \|\Phi\|_\blacklozenge > \|\Phi\|_\diamond$.

All of the bounds that we established in this section can be tight, as we shall demonstrate in what follows.

7. Applications and examples

7.1. Positive maps and structural physical approximation

Positive maps constitute a fundamental way to detect and characterise quantum entanglement [2–4]. One of the most studied approaches to implementing such maps in practice is the structural physical approximation (SPA) [5, 6], which aims to approximate a given positive map Φ with a physical quantum channel by considering decompositions of the form $\Phi + \varsigma\mathcal{D}$, where $\mathcal{D}$ is the completely depolarising channel, $J_\mathcal{D} = \mathbb{1}/d_B$. Such approximations have found use in both understanding the properties of positive maps [7, 59], as well as in realising them in experiments [6, 8, 60].

Intuitively, the robustness measures can then be understood as different approaches to defining an optimised SPA to the map Φ, by allowing channels other than the depolarising map to be used in the decomposition (cf. [33]). We will now discuss the similarities and differences between the approaches by studying two representative examples of positive maps.

Transposition map. Consider first the transposition map $T \in \mathcal{H}(A, A)$. Letting SPA(T) denote the minimal amount ς needed for $(T + \varsigma\mathcal{D})/(1 + \varsigma)$ to be a quantum channel, it can be easily verified that SPA(T) $= d_A$. However, by making a more suitable choice of a channel in the optimisation, our robustness measures construct an approximation as $(T + \lambda\Lambda)/(1 + \lambda)$ where $\lambda = \frac{1}{2}(d_A - 1)$ already suffices to ensure that this is a valid physical channel. From this we see that $R(T) = \frac{1}{2}(d_A - 1)$ and hence $\|T\|_\blacklozenge = d_A$. Quantitatively, the advantage gained by allowing arbitrary channels in such decompositions can therefore be significant.

To understand why a better approximation can be obtained, let us take a closer look at the optimal decomposition for this map. Our generalised approach can take into consideration the fact that the Choi operator of the transposition map, J_T (the swap operator), already has a non-trivial positive part, which means that there is no need to act on that part of the space. More specifically, a better approximation is obtained simply by defining the map $J_\Lambda = \dfrac{\mathbb{1} - J_T}{d_A - 1} \in \mathrm{CPTP}$ and mixing as

$$J_T + \frac{1}{2}(d_A - 1)J_\Lambda \geq 0 \implies \frac{2}{d_A + 1}T + \frac{d_A - 1}{d_A + 1}\Lambda \in \mathrm{CPTP}. \tag{59}$$

Structurally, this is not too different from the SPA — the only maps involved in the combination are the depolarising channel and the transposition map itself, even if the optimal approximation is not simply a convex mixture of the two. Indeed, we could define an optimised structural physical approximation which allows for such decompositions to be used:

$$\begin{aligned}
\mathrm{SPA}'(\Phi) &:= \min\left\{\varsigma' \;\middle|\; J_\Phi + \varsigma'\left[\frac{\lambda_{\max}(J_\Phi)\mathbb{1} - J_\Phi}{\lambda_{\max}(J_\Phi)d_B - 1}\right] \geq 0\right\} \\
&= -\lambda_{\min}(J_\Phi)\frac{d_B\lambda_{\max}(J_\Phi) - 1}{\lambda_{\max}(J_\Phi) - \lambda_{\min}(J_\Phi)},
\end{aligned} \tag{60}$$

with the expression valid for any map such that $\lambda_{\min}(J_\Phi) < \lambda_{\max}(J_\Phi) \neq d_B^{-1}$. This can be used to give a general bound to the robustness measures.

> **Proposition 11.** For any trace-preserving map $\Phi \in \mathcal{H}(A, B)$ such that $\Phi \neq \mathcal{D}$, it holds that
>
> $$R(\Phi) \leq \text{SPA}'(\Phi) \leq \text{SPA}(\Phi). \tag{61}$$

In the case of the transpose, it holds that $\text{SPA}'(T) = R(T) = \frac{1}{2}(d_A - 1)$, so we know that an optimal approximation of the transposition map can be realised with only the depolarising channel, as long as one considers the optimised approach of Eq. (60). However, this is not the case for general maps, and the advantages offered by the generalised robustness approach can provide new insight into optimal approximations of maps, as we shall see in the following.

Choi map. The Choi map $\mathcal{C} \in \mathcal{H}(A, A)$ with $d_A = 3$ is an example of an indecomposable positive map, and is defined by [61]

$$\mathcal{C}(X) := \begin{pmatrix} X_{11} + X_{22} & -X_{12} & -X_{13} \\ -X_{21} & X_{22} + X_{33} & -X_{23} \\ -X_{31} & -X_{32} & X_{33} + X_{11} \end{pmatrix} \tag{62}$$

where X_{ij} denote the matrix elements of X in a chosen basis. A numerical evaluation shows that the optimal decompositions for $\mathcal{C}$ give $\text{SPA}(\mathcal{C}) = \frac{3}{2}$ and $\text{SPA}'(\mathcal{C}) = \frac{2}{3}$. With the robustness, an improved choice can be obtained by choosing $\Lambda = \text{id}$ and mixing as $J_{\mathcal{C}} + \frac{1}{6}J_{\text{id}} \geq 0$, yielding $R(\mathcal{C}) = \frac{1}{6}$. Consequently, mixing with more general maps can not only provide quantitative improvements, but also identify ways of implementing non-CPTP maps which are impossible to find with the standard structural physical approximations.

An interesting difference between the SPA- and robustness-based approaches is that the optimal SPA of the Choi map is a measure-and-prepare (entanglement-breaking) channel [7], while the map obtained in the robustness-based approach is not (as can be verified with the PPT criterion). Since measure-and-prepare channels enjoy an easy implementation in practical settings, it would be an interesting extension of our approach to consider the extent of a quantitative advantage that can be maintained while requiring that the optimal CPTP approximation be entanglement breaking.

We also note that another approach to realising positive maps was studied in Ref. [62] by using multiple copies of the input state, where a related SPA-based approximation was also considered. An extension of the methods of our work to this framework could provide additional insight into the implementability of positive maps.

7.2. Inverse quantum channels

A fundamentally important case of a non-CPTP map encountered in many settings is the inverse linear map of a bijective quantum channel, that is, a map such that $\Lambda^{-1} \circ \Lambda = \Lambda \circ \Lambda^{-1} = \text{id}$.[2] Note that such an inverse is not guaranteed to exist for a general channel, and even when it does, it will not form a valid quantum channel unless Λ is a unitary map. However, many important cases of quantum dynamics are indeed invertible, allowing us to study their inverses in the formalism of our work.

Non-Markovianity. One setting in which channel inverses play a role is the study of non-Markovianity. Among the different ways to define Markovian evolution, a common way is to say that a time-dependent evolution governed by the channel $\Lambda_{t,0}$ is Markovian if it behaves as a physical map over any time interval $[t, t + \delta t]$. Mathematically, any $\Lambda_{t,0}$ satisfying this condition is said to be CP-divisible [63–65], which can be formalised by the statement that for all times t and $s \leq t$ we can write

$$\Lambda_{t,0} = \Xi_{t,s} \circ \Lambda_{s,0}$$

where the propagator $\Xi_{t,s}$ is a CPTP map. For more general channels, the decomposition $\Lambda_{t,0} = \Xi_{t,s} \circ \Lambda_{s,0}$ results in some $\Xi_{t,s}$ that is non-CPTP, indicating that Markovian dynamics break down after some time point s.

Observe that, provided $\Lambda_{t,0}$ is invertible for all t, we can take $\Xi_{t,s} = \Lambda_{t,0} \circ \Lambda_{s,0}^{-1}$. Therefore, the non-physicality of $\Lambda_{t,0} \circ \Lambda_{s,0}^{-1}$ serves as an indicator of non-Markovianity, and — since this map is trace preserving for any trace-preserving Λ — the diamond norm $\left\| \Lambda_{t,0} \circ \Lambda_{s,0}^{-1} \right\|_{\diamond}$ can be

---

[2] We note that in many cases it suffices to consider only left or right inverses, but we assume two-sided invertibility for simplicity.

used as a quantitative measure of non-Markovianity over the time-interval $[t, s]$. This is similar to the original approach of Ref. [63] where a quantifier based on the trace norm of the Choi operator was employed — the advantage of our definition is the ability to interpret this quantity operationally.

Specifically, we observe that quantum mechanics is ultimately a Markovian theory: if we had knowledge of all relevant objects, then all quantum dynamics could be described by Markovian unitary dynamics. That is, any information from the past that is relevant to the future must pass through the present, and hence the optimal prediction of future observational statistics ultimately depends only on the the present state of reality. Non-Markovianity is an artefact of not tracking all relevant information in the present. In our context, this arises as our mathematical characterisation of the candidate channel, $\Lambda_{s,0}$, does not track the state of the environment. The operational relevance of $\left\| \Lambda_{t,0} \circ \Lambda_{s,0}^{-1} \right\|_{\diamond}$ then becomes more evident. Notably, in Sec. 4 we presented a systematic means of simulating any unphysical map $\Xi_{t,s}$ by introducing an ancillary system X. Here, we may think of this as building a Markovian model for $\Xi_{t,s}$ by introducing $X = \sum_i \mu_i \rho_i$ as an "artificial environment". The feeding in of different states ρ_i depending on X then represents a means in which non-Markovian behaviour on the system is realised. While this construction does not immediately look physical (as it allows affine mixtures of quantum states), it can be simulated by a classical computer with sufficient resource overhead. The resource costs of doing so — $\|\Xi_{t,s}\|_{\diamond}$ — thus represents a bound on the information processing capabilities of the environment that enable said non-Markovian behaviour to emerge.

There are multiple approaches for extending this to a time-independent measure of non-Markovianity of Λ. One could, for example, take the supremum of the measure $\left\| \Lambda_{t,0} \circ \Lambda_{s,0}^{-1} \right\|_{\diamond}$ over all t and s. This would then characterise how much extra information processing we need beyond tracking the state of the system at time s to simulate dynamics over the time-interval $[s, t]$. We may also follow an approach based on Ref. [63] and define $\mathcal{I}_{\diamond}(\Lambda) := \int_0^{\infty} g_{\diamond,t}(\Lambda) \, \mathrm{d}t$, where $g_{\diamond,t}$ can be understood as the right-hand derivative of the diamond norm of the dynamics at time t:

$$g_{\diamond,t}(\Lambda) := \lim_{\varepsilon \to 0^+} \frac{\left\| \Lambda_{t+\varepsilon,0} \circ \Lambda_{t,0}^{-1} \right\|_{\diamond} - \left\| \Lambda_{t,0} \circ \Lambda_{t,0}^{-1} \right\|_{\diamond}}{\varepsilon} = \lim_{\varepsilon \to 0^+} \frac{\left\| \Lambda_{t+\varepsilon,0} \circ \Lambda_{t,0}^{-1} \right\|_{\diamond} - 1}{\varepsilon}. \tag{63}$$

$\mathcal{I}_{\diamond}(\Lambda)$ therefore represents the total amount of non-Markovianity in this evolution. A suitable normalisation of this quantity can allow for the comparison of the strength of non-Markovianity in different settings [63, 65]. We leave a careful consideration of these possibilities to future work.

Error mitigation. Another application for the study of channel inverses is error mitigation. This setting considers the scenario where one is tasked with computing expectation values of the type $\mathrm{Tr}[\mathcal{U}(\rho)A]$ for an input state ρ, ideal gate $\mathcal{U}$, and observable A, while operations are followed by a noise channel Θ. A leading approach to this problem, called probabilistic error cancellation [17, 66], is to counteract the noise with the inverse map Θ^{-1}, so that $\mathrm{Tr}[\mathcal{U}(\rho)A] = \mathrm{Tr}[\Theta \circ \Theta^{-1} \circ \mathcal{U}(\rho)A]$. By decomposing Θ^{-1} into a quasiprobability distribution over a convex subset of channels $\mathcal{P} = \{\Lambda_i\}$ such that $\Lambda_i \circ \mathcal{U}$ would be implementable on a (fictitious) noiseless device, standard quasiprobability sampling arguments allows us to construct an unbiased estimator for $\mathrm{Tr}[\mathcal{U}(\rho)A]$ using only operations implementable on a noisy device. The optimal overhead cost of such a procedure scales as $\gamma_{\mathcal{P}}(\Theta)^2$, where [17, 30]

$$\gamma_{\mathcal{P}}(\Theta) = \min \left\{ \sum_i |\lambda_i| \; \middle| \; \Theta^{-1} = \sum_i \lambda_i \Lambda_i, \; \Lambda_i \in \mathcal{P} \right\}$$
$$= \min \left\{ \lambda_+ + \lambda_- \; \middle| \; \Theta^{-1} = \lambda_+ \Lambda_+ - \lambda_- \Lambda_-, \; \Lambda_\pm \in \mathcal{P} \right\}. \tag{64}$$

The specific choice of $\mathcal{P}$ can be made depending on not only the physical setting in consideration, but also on one's precise motivations. On the one hand, a set with a finite number of operations (e.g., Clifford gates) turns Eq. (64) into a linear program [17, 66], making the overhead cost easily computable while sacrificing the expressibility of devices. On the other hand, choosing a larger set with an infinite number of implementable operations takes into account a larger expressibility [30], but makes the computation of Eq. (64) hard in general. Here, to accommodate computability and expressibility at the same time, we take another approach considered in Ref. [33, 67]: we choose $\mathcal{P}$ to be *all* physical quantum channels. We notice that the norm $\|\cdot\|_{\blacklozenge}$ provides the cost of error mitigation in this setting as $\gamma_{\mathrm{CPTP}}(\Theta) = \left\| \Theta^{-1} \right\|_{\blacklozenge} = \left\| \Theta^{-1} \right\|_{\diamond}$, which can be efficiently computed by semidefinite programming. Although this choice of $\mathcal{P}$ might seem too permissive, the lower bound obtained through this approach can actually match known

achievability results (upper bounds) [33], showing new optimality results and even improving on the specialised characterisation of Ref. [30] in some cases. Of note is the fact that, since any inverse map Θ^{-1} of a quantum channel Θ is trace preserving, our Thm. 3 shows a new application of the diamond norm in bounding the cost of error mitigation: it always holds that $\gamma_{\mathcal{P}}(\Theta) \geq \left\|\Theta^{-1}\right\|_{\diamond}$, regardless of the choice of $\mathcal{P}$.

In some cases — such as when experiencing the leakage or loss of some qubits during computation — the noisy evolution can actually correspond to a map which is not trace preserving. Although many previous approaches did not take this into consideration, our methods explicitly extend to such maps, allowing one to understand the simulation of non-trace-preserving linear maps through Thm. 4. Related settings which our methods can characterise include the so-called linear quantum error correction [68], which aims to correct errors of systems undergoing general, non-CPTP dynamics Θ, as well as error mitigation for non-Markovian noise [69], where the mitigation cost can be related to a measure of non-Markovianity. In such cases, our approach can thus help understand the implementation of not only the inverse maps, but also the dynamics themselves.

7.2.1. *Computing the measures*

To showcase the application of our methods and evaluate the measures for some representative examples, we will consider the inverse maps of several fundamental types of noisy quantum evolutions: depolarising, amplitude damping, dephasing, and qubit leakage channels. The expressions for the first two appeared in Ref. [33], which we rederive using the methods and results of this work. We also find for the first three that the optimal decomposition into $\Lambda_{\pm}$ for the norm $\|\Theta^{-1}\|_{\blacklozenge}$ (Eq. (9)), can be taken as convex mixtures of unitaries and state preparations. Thus, $\|\Theta^{-1}\|_{\blacklozenge}$ also serves as the optimal cost $\gamma_{\mathcal{P}}(\Theta^{-1})$ with a smaller set $\mathcal{P}$ as considered in Ref. [30], indicating that the capability to implement all CPTNI maps does not provide any advantage over that of implementing unitaries and state preparations only. Note that the inverses of trace-preserving maps are trace preserving, and so in such cases the equality $\|\Phi\|_{\blacklozenge} = \|\Phi\|_{\diamond} = 2R(\Phi)+1$ holds by Thm. 3, which means that it will suffice to evaluate any one of the measures.

Depolarising noise. The depolarising channel, given by $\mathcal{D}_p(X) := (1 - p)X + p \operatorname{Tr} X \frac{\mathbb{1}}{d_A}$ for some noise parameter $p \in [0, 1)$, has the inverse $\mathcal{D}_p^{-1}(X) = \frac{1}{1-p}X - \frac{p}{1-p}\operatorname{Tr} X \frac{\mathbb{1}}{d_A}$. This gives

$$J_{\mathcal{D}_p^{-1}} = \frac{1}{1-p}|\Omega\rangle\langle\Omega| - \frac{p}{(1-p)d_A}\mathbb{1}_{A\otimes A}. \tag{65}$$

Importantly, one can notice that $\operatorname{Tr}_B J_{\mathcal{D}_p^{-1}{}_+}$ and $\operatorname{Tr}_B J_{\mathcal{D}_p^{-1}{}_-}$ are proportional to identity. As first noticed in [41, 42], this means that the lower bound $\frac{1}{d_A}\left\|J_{\mathcal{D}_p^{-1}}\right\|_1$ of Prop. 7 matches the upper bound $\lambda_{\max}\left(\operatorname{Tr}_B\left[J_{\mathcal{D}_p^{-1}{}_+} + J_{\mathcal{D}_p^{-1}{}_-}\right]\right)$ of Prop. 8[3]. We thus get

$$\left\|\mathcal{D}_p^{-1}\right\|_{\blacklozenge} = \left\|\mathcal{D}_p^{-1}\right\|_{\diamond} = \frac{1}{d_A}\left\|J_{\mathcal{D}_p^{-1}}\right\|_1 = \frac{1 + \left(1 - 2d_A^{-2}\right)p}{1-p}. \tag{66}$$

Dephasing noise. The generalised dephasing channel [70] is defined by $\Delta_{\mathbf{p}}(X) := \sum_{i=0}^{d_A-1} p_i Z_i X Z_i^{\dagger}$, where $\mathbf{p} = (p_0, \ldots, p_{d_A-1})$ is a chosen set of noise parameters $p_i \geq 0$, and Z_i refers to the qudit clock operators

$$Z_i = \sum_{j=0}^{d_A-1} \omega^{ij}|j\rangle\langle j| \tag{67}$$

in some basis $\{|i\rangle\}$, with ω being a primitive d_Ath root of unity. In the case of $d_A = 2$, this recovers the usual qubit dephasing channel $\Delta_p(X) = (1 - p)X + pZXZ^{\dagger}$. One can notice that the action of this channel can be represented by $\Delta_{\mathbf{p}}(X) = X \odot S$ where $\odot$ denotes the element-wise matrix product (Schur/Hadamard product), and

$$(S)_{jk} = \sum_{i=0}^{d_A-1} p_i \omega^{ij}(\omega^{ik})^* = \sum_{i=0}^{d_A-1} p_i \omega^{i(j-k)} \qquad j, k = 0, \ldots d_A - 1 \tag{68}$$

---

[3] In fact, $\|\Phi\|_{\diamond} = \frac{1}{d_A}\|J_\Phi\|_1$ if and only if $\operatorname{Tr}_B\left(J_{\Phi_+} + J_{\Phi_-}\right) \propto \mathbb{1}$ [41, 42].

in the same basis $\{|i\rangle\}$. Provided that the coefficients of S are non-zero (that is, $\Delta_{\mathbf{p}}$ does not act as a completely dephasing channel on any subspace), the map is invertible as $\Delta_{\mathbf{p}}^{-1}(X) = X \odot \overline{S}$ with $\overline{S}$ defined by

$$(\overline{S})_{jk} = \frac{1}{(S)_{jk}} \qquad j, k = 0, \dots d_A - 1. \tag{69}$$

We will now show that $\left\|\Delta_{\mathbf{p}}^{-1}\right\|_{\blacklozenge} = \left\|\Delta_{\mathbf{p}}^{-1}\right\|_{\diamond} = \frac{1}{d_A}\left\|J_{\Delta_{\mathbf{p}}^{-1}}\right\|_1 = \frac{1}{d_A}\left\|\overline{S}\right\|_1$.

The equality $\left\|\Delta_{\mathbf{p}}^{-1}\right\|_{\blacklozenge} = \left\|\Delta_{\mathbf{p}}^{-1}\right\|_{\diamond}$ is a consequence of Thm. 3; note here that we do not actually need to impose that $\Delta_{\mathbf{p}}$ be trace preserving (i.e., that $\sum_i p_i = 1$), since both $\Delta_{\mathbf{p}}$ and $\Delta_{\mathbf{p}}^{-1}$ are always proportional to a trace-preserving map by construction.

To show the equality $\left\|\Delta_{\mathbf{p}}^{-1}\right\|_{\blacklozenge} = \frac{1}{d_A}\left\|\overline{S}\right\|_1$, consider the decomposition of $\overline{S}$ as $\overline{S} = \overline{S}_+ - \overline{S}_-$. Crucially, since S is a circulant matrix, so is $\overline{S}$, and hence it can be diagonalised by the Fourier transform matrix $(F)_{jk} = \frac{1}{\sqrt{d_A}}\omega^{jk}$ [71, 2.2.P10]. Each eigenvector of $\overline{S}$ is therefore of the form

$$|s_m\rangle = \frac{1}{\sqrt{d_A}}\sum_{i=0}^{d_A-1}\omega^{im}|i\rangle, \tag{70}$$

ensuring in particular that all diagonal elements of each density matrix $|s_m\rangle\langle s_m|$ are equal. This entails that $\overline{S}_+$ and $\overline{S}_-$ both have constant diagonals. Define now the maps

$$\Lambda_{\pm}(X) := X \odot \overline{S}_{\pm}. \tag{71}$$

Since $\overline{S}_{\pm} \geq 0$, each such map is completely positive [72, Thm. 3.7], and clearly $\Lambda'_{\pm} := \Lambda_{\pm} d_A / \text{Tr}(\overline{S}_{\pm})$ is trace preserving as we have just seen that $(\overline{S}_{\pm})_{ii} = (\overline{S}_{\pm})_{jj} \, \forall i, j$. Thus we have a decomposition as

$$\Delta_{\mathbf{p}}^{-1} = \frac{\text{Tr}\,\overline{S}_+}{d_A}\Lambda'_+ - \frac{\text{Tr}\,\overline{S}_-}{d_A}\Lambda'_-, \quad \Lambda'_{\pm} \in \text{CPTP}, \tag{72}$$

from which we get the bound $\left\|\Delta_{\mathbf{p}}^{-1}\right\|_{\blacklozenge} \leq \frac{1}{d_A}\left\|\overline{S}\right\|_1$. On the other hand, let $|\psi\rangle = \frac{1}{\sqrt{d_A}}\sum_{i=0}^{d_A-1}|i\rangle$ and use Prop. 10 to get

$$\left\|\Delta_{\mathbf{p}}^{-1}\right\|_{\blacklozenge} \geq \left\|\Delta_{\mathbf{p}}^{-1}(|\psi\rangle\langle\psi|)\right\|_1 = \frac{1}{d_A}\left\|\overline{S}\right\|_1. \tag{73}$$

Finally, the equality $\left\|J_{\Delta_{\mathbf{p}}^{-1}}\right\|_1 = \left\|\overline{S}\right\|_1$ is obtained by noticing that $J_{\Delta_{\mathbf{p}}^{-1}} = \sum_{i,j}(\overline{S})_{ij}|ii\rangle\langle jj|$ which has the same eigenvalues as $\overline{S}$.

The eigenvalues of $\overline{S}$ can be readily obtained due to the fact that it is a circulant matrix [71, 2.2.P10], allowing for a straightforward computation of the trace norm $\left\|\overline{S}\right\|_1$ and altogether giving

$$\left\|\Delta_{\mathbf{p}}^{-1}\right\|_{\diamond} = \frac{1}{d_A}\sum_{m=0}^{d_A-1}\left|\sum_{j=0}^{d_A-1}\left(\sum_{i=0}^{d_A-1}p_i\omega^{j(i-m)}\right)^{-1}\right|. \tag{74}$$

For the qubit dephasing channel with $p \in [0, \frac{1}{2})$, we recover

$$\left\|\Delta_p^{-1}\right\|_{\diamond} = \frac{1}{2}\left\|\begin{pmatrix} 1 & \frac{1}{1-2p} \\ \frac{1}{1-2p} & 1 \end{pmatrix}\right\|_1 = \frac{1}{1-2p}. \tag{75}$$

Since each eigenvector $|s_m\rangle$ for $\overline{S}$ in (70) corresponds to the application of Z_m, $\Lambda'_{\pm}$ in (72) are realised as probabilistic applications of the generalised phase unitaries.

Amplitude damping noise. The qubit amplitude damping channel $\mathcal{A}_{\gamma}(\cdot) = A_0 \cdot A_0^{\dagger} + A_1 \cdot A_1^{\dagger}$ is defined by the Kraus operators $A_0 := |0\rangle\langle 0| + \sqrt{1-\gamma}\,|1\rangle\langle 1|$ and $A_1 := \sqrt{\gamma}\,|0\rangle\langle 1|$. Using the fact that

$$\begin{aligned} |1\rangle\langle 1| &= \frac{1}{1-\gamma}\mathcal{A}_{\gamma}(|1\rangle\langle 1|) - \frac{\gamma}{1-\gamma}|0\rangle\langle 0| \\ &= \frac{1}{1-\gamma}\mathcal{A}_{\gamma}(|1\rangle\langle 1|) - \frac{\gamma}{1-\gamma}\mathcal{A}_{\gamma}(|0\rangle\langle 0|), \end{aligned} \tag{76}$$

we have

$$\mathcal{A}_\gamma^{-1}(|1\rangle\langle 1|) = \frac{1}{1-\gamma} |1\rangle\langle 1| - \frac{\gamma}{1-\gamma} |0\rangle\langle 0| . \tag{77}$$

Proposition 10 thus gives

$$\left\|\mathcal{A}_\gamma^{-1}\right\|_\blacklozenge = \left\|\mathcal{A}_\gamma^{-1}\right\|_\diamond \geq \left\|\mathcal{A}_\gamma^{-1}(|1\rangle\langle 1|)\right\|_1$$
$$= \frac{1+\gamma}{1-\gamma}. \tag{78}$$

A matching upper bound can be obtained by explicitly computing $J_{\mathcal{A}_\gamma^{-1}}$ (see e.g. [17, 30]) and using the upper bound in Prop. 8.

The above shows a rather general method of obtaining lower bounds for linear maps which are inverses of other linear maps, without having to explicitly compute the full inverse map. Indeed, this can be extended to maps which only approximately invert a given channel — useful, for instance, when dealing with non-invertible maps, or when aiming to reduce the cost of implementing a given map by only requiring that it approximately mitigates the error.

Proposition 12. Let $\Phi \in \mathcal{H}(A, B)$ and $\widetilde{\Phi} \in \mathcal{H}(B, A)$ be such that $\left\|\widetilde{\Phi} \circ \Phi(\rho) - \rho\right\|_1 \leq \varepsilon$ for all $\rho \in \mathbb{D}(A)$. Then

$$\left\|\widetilde{\Phi}\right\|_\diamond \geq \max\left\{ \|Z\|_1 (1-\varepsilon) \mid \Phi(Z) \in \mathbb{D}(B) \right\}$$
$$\left\|\widetilde{\Phi}\right\|_\blacklozenge \geq \max\left\{ \operatorname{Tr} Z_- + \operatorname{Tr} Q_+ - \varepsilon(\|Z\|_1 + \|Q\|_1) \mid \Phi(Z), \Phi(Q) \in \mathbb{D}(B) \right\}$$
$$R(\widetilde{\Phi}) \geq \max\left\{ \operatorname{Tr} Z_- + \operatorname{Tr} Q_+ - \operatorname{Tr} \Phi(Q) - \varepsilon(\|Z\|_1 + \|Q\|_1) \mid \Phi(Z), \Phi(Q) \geq 0, \right.$$
$$\left. \operatorname{Tr} \Phi(Z + Q) = 1 \right\} \tag{79}$$
$$R'(\widetilde{\Phi}) \geq \max\left\{ \operatorname{Tr} Z_+ - 1 - \varepsilon \|Z\|_1 \mid \Phi(Z) \in \mathbb{D}(B) \right\}$$
$$R''(\widetilde{\Phi}) \geq \max\left\{ \operatorname{Tr} Z_- - \varepsilon \|Z\|_1 \mid \Phi(Z) \in \mathbb{D}(B) \right\} .$$

Proof. We use Prop. 10 to get that

$$\left\|\widetilde{\Phi}\right\|_\diamond \geq \max\left\{ \left\|\widetilde{\Phi}(\sigma)\right\|_1 \mid \sigma \in \mathbb{D}(B) \cap \operatorname{ran}(\Phi) \right\}$$
$$= \max\left\{ \left\|\widetilde{\Phi} \circ \Phi(Z)\right\|_1 \mid \Phi(Z) \in \mathbb{D}(B) \right\}$$
$$\geq \max\left\{ \|Z\|_1 - \left\|Z - \widetilde{\Phi} \circ \Phi(Z)\right\|_1 \mid \Phi(Z) \in \mathbb{D}(B) \right\} \tag{80}$$
$$\geq \max\left\{ \|Z\|_1 (1-\varepsilon) \mid \Phi(Z) \in \mathbb{D}(B) \right\} .$$

The third line follows by the triangle inequality, and the last line is a consequence of the assumption that $\left\|\widetilde{\Phi} \circ \Phi(\rho) - \rho\right\|_1 \leq \varepsilon$ for all $\rho \in \mathbb{D}(A)$, since we can write any $Z = \mu_+ \rho_+ - \mu_- \rho_-$ for some $\rho_\pm \in \mathbb{D}(A)$ to get $\left\|\widetilde{\Phi} \circ \Phi(Z) - Z\right\|_1 \leq \varepsilon(\mu_+ + \mu_-) \leq \varepsilon \|Z\|_1$. The case of the other measures is analogous: using the variational form of the function $\operatorname{Tr} Z_+$ (and similarly $\operatorname{Tr} Z_-$) we can obtain

$$\operatorname{Tr} \widetilde{\Phi}(\sigma)_+ = \max\left\{ \left\langle \widetilde{\Phi} \circ \Phi(Z), W \right\rangle \mid 0 \leq W \leq \mathbb{1} \right\}$$
$$= \max\left\{ \langle Z, W \rangle - \left\langle Z - \widetilde{\Phi} \circ \Phi(Z), W \right\rangle \mid 0 \leq W \leq \mathbb{1} \right\}$$
$$\geq \max\left\{ \langle Z, W \rangle - \left\|Z - \widetilde{\Phi} \circ \Phi(Z)\right\|_1 \|W\|_\infty \mid 0 \leq W \leq \mathbb{1} \right\} \tag{81}$$
$$\geq \operatorname{Tr} Z_+ - \varepsilon \|Z\|_1$$

where we used the Cauchy-Schwarz inequality. Using these bounds in Prop. 10 yields the stated result. ∎

Leakage error. Consider the qubit leakage error $\mathcal{L}_p(\cdot) = L_p \cdot L_p^\dagger$ where $L_p := |0\rangle\langle 0| + \sqrt{1-p} |1\rangle\langle 1|$. This represents a situation where the excited state is lost with probability $1-p$, and this stochastic

nature is reflected to the fact that $\mathcal{L}_p$ is not trace preserving. The inverse of the leakage error is given by $\mathcal{L}_p^{-1}(\cdot) = L_p^{-1} \cdot L_p^{-1}$. Since this is a completely positive map, Eq. (57) gives

$$\left\| \mathcal{L}_p^{-1} \right\|_{\blacklozenge} = \left\| \mathcal{L}_p^{-1} \right\|_{\diamond} = \left\| \mathrm{Tr}_B J_{\mathcal{L}_p^{-1}} \right\|_{\infty} = \frac{1}{1-p}. \tag{82}$$

Note that the inverse can be realised as

$$\mathcal{L}_p^{-1} = \frac{1}{2}\left(1 + \frac{1}{\sqrt{1-p}} \right) \mathrm{id} - \frac{1}{2}\left(\frac{1}{\sqrt{1-p}} - 1 \right) \mathcal{Z} + \frac{p}{1-p} \Pi_{|1\rangle\langle 1|} \tag{83}$$

where $\mathcal{Z}(\cdot) := Z \cdot Z$ with $Z = |0\rangle\langle 0| - |1\rangle\langle 1|$ being the Pauli Z matrix, and $\Pi_{|1\rangle\langle 1|}(\cdot) := |1\rangle\langle 1| \cdot |1\rangle\langle 1|$ being the projection onto the state $|1\rangle$.

8. Discussion

We introduced a comprehensive quantitative approach to the study of non-completely-positive linear maps, focusing in particular on the task of approximating and simulating them with valid quantum channels. To this end, we considered several quantifiers which generalise measures employed in the study of quantum resources — namely, variants of the robustness and base norm measures. We showed that they satisfy very close relations with the diamond norm, and in particular are exactly equal to it for any trace-preserving linear map. Since such trace-preserving maps are the most commonly encountered examples of dynamics beyond physical quantum channels, this allowed us to establish fruitful interrelations between the quantities, and discover new applications of the fundamentally important quantity that is the diamond norm. We developed in particular two operational connections. Firstly, we introduced a method of simulating general linear maps with quantum channels, shifting the difficulty of realising non-quantum dynamics onto the structurally simpler task of implementing linear combinations of quantum states. We showed that our robustness measure exactly quantifies the cost of realising such schemes in terms of the required state-based resources. Secondly, we showed that another variant of the robustness finds use as an exact quantifier of the performance advantage that a general linear map can enable over quantum channels in a class of state discrimination games. We introduced a number of useful bounds and explicitly employed them to demonstrate the computability of the measures for some representative examples. Finally, we showed how our measures can find use in the quantitative characterisation of several practically relevant settings, namely, structural approximations of positive maps, non-Markovianity quantification, and tightly bounding the cost of probabilistic error mitigation.

Although we focused on the application of our framework to Hermiticity-preserving maps, we note that more general linear maps can be treated in a similar way. The simplest way to approach this is to decompose any linear map Φ into its Hermiticity-preserving and skew-Hermiticity-preserving parts, that is, write $\Phi = \Phi_H + i\Phi_{SH}$ where the constituent maps are defined through $J_{\Phi_H} := \frac{1}{2}(J_\Phi + J_\Phi^\dagger)$ and $J_{\Phi_{SH}} := \frac{1}{2i}(J_\Phi - J_\Phi^\dagger)$. The maps Φ_H and Φ_{SH} are then explicitly Hermiticity-preserving, and our arguments can be applied to them directly. A similar approach was employed in [73] to decompose the two-point quantum correlator $\mathcal{T} : \mathbb{L}(A) \to \mathbb{L}(A \otimes A)$, defined as the map satisfying $\mathrm{Tr}[\mathcal{T}(\rho)(A \otimes B)] = \mathrm{Tr}[A\rho B]$ for all A, B. Indeed, one can show that the decompositions constructed in [73] are also optimal for the robustness-based quantities.

We also note that the diamond norm has been applied as a measure of specific properties of quantum channels, such as their ability to detect coherence [74]. Connections between our methods and such approaches could be fruitful to explore.

A major outstanding issue is to understand how the framework of this work can be extended to non-linear maps, which could allow for the characterisation and more efficient approximation of important unphysical dynamics such as quantum cloners. This question was already asked in the earliest works concerned with approximating non-CPTP maps with quantum channels [5], but it still remains a considerable challenge to devise approaches which could apply to general non-linear transformations.

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, tenth ed. (Cambridge University Press, New York, NY, USA, 2011).

[2] M. Horodecki, P. Horodecki, and R. Horodecki, *Separability of mixed states: Necessary and sufficient conditions*, Physics Letters A **223**, 1 (1996).

[3] O. Gühne and G. Tóth, *Entanglement detection*, Phys. Rep. **474**, 1 (2009).

[4] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Quantum entanglement*, Rev. Mod. Phys. **81**, 865 (2009).

[5] P. Horodecki, *From limits of quantum operations to multicopy entanglement witnesses and state-spectrum estimation*, Phys. Rev. A **68**, 052101 (2003).

[6] P. Horodecki and A. Ekert, *Method for Direct Detection of Quantum Entanglement*, Phys. Rev. Lett. **89**, 127902 (2002).

[7] J. K. Korbicz, M. L. Almeida, J. Bae, M. Lewenstein, and A. Acín, *Structural approximations to positive maps and entanglement-breaking channels*, Phys. Rev. A **78**, 062105 (2008).

[8] J. Bae, *Designing quantum information processing via structural physical approximation*, Rep. Prog. Phys. **80**, 104001 (2017).

[9] P. Pechukas, *Reduced Dynamics Need Not Be Completely Positive*, Phys. Rev. Lett. **73**, 1060 (1994).

[10] A. Shaji and E. C. G. Sudarshan, *Who's afraid of not completely positive maps?* Physics Letters A **341**, 48 (2005).

[11] C. A. Rodríguez-Rosario, K. Modi, A. Kuah, A. Shaji, and E. C. G. Sudarshan, *Completely positive maps and classical correlations*, J. Phys. A: Math. Theor. **41**, 205301 (2008).

[12] H. A. Carteret, D. R. Terno, and K. Życzkowski, *Dynamics beyond completely positive maps: Some properties and applications*, Phys. Rev. A **77**, 042113 (2008).

[13] R. Alicki, *Comment on "Reduced Dynamics Need Not Be Completely Positive"*, Phys. Rev. Lett. **75**, 3020 (1995).

[14] K. Modi, *Operational approach to open dynamics and quantifying initial correlations*, Sci. Rep. **2**, 581 (2012).

[15] D. Schmid, K. Ried, and R. W. Spekkens, *Why initial system-environment correlations do not imply the failure of complete positivity: A causal perspective*, Phys. Rev. A **100**, 022112 (2019).

[16] P. Shor, *Fault-tolerant quantum computation*, in *Proceedings of 37th Conference on Foundations of Computer Science* (1996) pp. 56–65.

[17] K. Temme, S. Bravyi, and J. M. Gambetta, *Error Mitigation for Short-Depth Quantum Circuits*, Phys. Rev. Lett. **119**, 180509 (2017).

[18] Y. Li and S. C. Benjamin, *Efficient Variational Quantum Simulator Incorporating Active Error Minimization*, Phys. Rev. X **7**, 021050 (2017).

[19] A. Y. Kitaev, *Quantum computations: Algorithms and error correction*, Russ. Math. Surv. **52**, 1191 (1997).

[20] J. Watrous, *Notes on super-operator norms induced by Schatten norms*, arXiv:quant-ph/0411077 (2004).

[21] J. Watrous, *Theory of Quantum Information* (University of Waterloo, 2011).

[22] G. Vidal and R. Tarrach, *Robustness of entanglement*, Phys. Rev. A **59**, 141 (1999).

[23] M. G. Díaz, K. Fang, X. Wang, M. Rosati, M. Skotiniotis, J. Calsamiglia, and A. Winter, *Using and reusing coherence to realize quantum processes*, Quantum **2**, 100 (2018).

[24] R. Takagi and B. Regula, *General Resource Theories in Quantum Mechanics and Beyond: Operational Characterization via Discrimination Tasks*, Phys. Rev. X **9**, 031053 (2019).

[25] Z.-W. Liu and A. Winter, *Resource theories of quantum channels and the universal role of resource erasure*, arXiv:1904.04201 (2019).

[26] G. Gour and A. Winter, *How to Quantify a Dynamical Quantum Resource*, Phys. Rev. Lett. **123**, 150401 (2019).

[27] R. Uola, T. Kraft, and A. A. Abbott, *Quantification of quantum dynamics with input-output games*, Phys. Rev. A **101**, 052306 (2020).

[28] X. Yuan, Y. Liu, Q. Zhao, B. Regula, J. Thompson, and M. Gu, *Universal and Operational Benchmarking of Quantum Memories*, arXiv:1907.02521 (2020).

[29] R. Takagi, K. Wang, and M. Hayashi, *Application of the Resource Theory of Channels to Communication Scenarios*, Phys. Rev. Lett. **124**, 120502 (2020).

[30] R. Takagi, *Optimal resource cost for error mitigation*, arXiv:2006.12509 (2020).

[31] B. Regula and R. Takagi, *Fundamental limitations on quantum channel manipulation*, arXiv:2010.11942 (2020).

[32] B. Regula and R. Takagi, *One-shot manipulation of dynamical quantum resources*, arXiv:2012.02215 (2020).

[33] J. Jiang, K. Wang, and X. Wang, *Physical Implementability of Quantum Maps and Its Application in Error Mitigation*, arXiv:2012.10959 (2020).

[34] J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, Cambridge, 2018).

[35] M. F. Sacchi, *Optimal discrimination of quantum operations*, Phys. Rev. A **71**, 062340 (2005).

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

[37] A. Jenčová, *Base norms and discrimination of generalized quantum channels*, Journal of Mathematical Physics **55**, 022201 (2014).

[38] J. Watrous, *Semidefinite programs for completely bounded norms*, Theory Comput. **5**, 217 (2009).

[39] J. Watrous, *Simpler semidefinite programs for completely bounded norms*, Chic. J. Th. Comp. Sci. **19**, 1 (2013).

[40] C. Piveteau, D. Sutter, and S. Woerner, *Quasiprobability decompositions with reduced sampling overhead*, arXiv:2101.09290 (2021).

[41] U. Michel, M. Kliesch, R. Kueng, and D. Gross, *Comments on "Improving Compressed Sensing With the Diamond Norm"–Saturation of the Norm Inequalities Between Diamond and Nuclear Norm*, IEEE Trans. Inf. Theory **64**, 7443 (2018).

[42] I. Nechita, Z. Puchała, Ł. Pawela, and K. Życzkowski, *Almost all quantum channels are equidistant*, Journal of Mathematical Physics **59**, 052201 (2018).

[43] H. Pashayan, J. J. Wallman, and S. D. Bartlett, *Estimating Outcome Probabilities of Quantum Circuits Using Quasiprobabilities*, Phys. Rev. Lett. **115**, 070501 (2015).

[44] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, *Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels*, Phys. Rev. Lett. **70**, 1895 (1993).

[45] M. Berta, F. G. S. L. Brandão, M. Christandl, and S. Wehner, *Entanglement Cost of Quantum Channels*, IEEE Trans. Inf. Theory **59**, 6779 (2013).

[46] S. Pirandola, R. Laurenza, C. Ottaviani, and L. Banchi, *Fundamental limits of repeaterless quantum communications*, Nat. Commun. **8**, 15043 (2017).

[47] M. M. Wilde, *Entanglement cost and quantum channel simulation*, Phys. Rev. A **98** (2018), 10.1103/PhysRevA.98.042338.

[48] G. Gour and C. M. Scandolo, *Entanglement of a bipartite channel*, Phys. Rev. A **103**, 062422 (2021).

[49] S. Bäuml, S. Das, X. Wang, and M. M. Wilde, *Resource theory of entanglement for bipartite quantum channels*, arXiv:1907.04181 (2019).

[50] D. Gottesman and I. L. Chuang, *Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations*, Nature **402**, 390 (1999).

[51] J. R. Seddon and E. T. Campbell, *Quantifying magic for multi-qubit operations*, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **475**, 20190251 (2019).

[52] K. Ben Dana, M. García Díaz, M. Mejatty, and A. Winter, *Resource theory of coherence: Beyond states*, Phys. Rev. A **95**, 062327 (2017).

[53] J. Geller and M. Piani, *Quantifying non-classical and beyond-quantum correlations in the unified operator formalism*, J. Phys. A: Math. Theor. **47**, 424030 (2014).

[54] J. F. Fitzsimons, J. A. Jones, and V. Vedral, *Quantum correlations which imply causation*, Sci. Rep. **5**, 18281 (2015).

[55] E. Kaur and M. M. Wilde, *Amortized entanglement of a quantum channel and approximately teleportation-simulable channels*, J. Phys. A: Math. Theor. **51**, 035303 (2017).

[56] D. Rosset, F. Buscemi, and Y.-C. Liang, *Resource Theory of Quantum Memories and Their Faithful Verification with Minimal Assumptions*, Phys. Rev. X **8**, 021033 (2018).

[57] K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein, *Volume of the set of separable states*, Phys. Rev. A **58**, 883 (1998).

[58] M. Kliesch, R. Kueng, J. Eisert, and D. Gross, *Improving Compressed Sensing With the Diamond Norm*, IEEE Trans. Inf. Theory **62**, 7445 (2016).

[59] F. Shultz, *The structural physical approximation conjecture*, Journal of Mathematical Physics **57**, 015218 (2015).

[60] H.-T. Lim, Y.-S. Kim, Y.-S. Ra, J. Bae, and Y.-H. Kim, *Experimental Realization of an Approximate Partial Transpose for Photonic Two-Qubit Systems*, Phys. Rev. Lett. **107**, 160401 (2011).

[61] M.-D. Choi, *Some assorted inequalities for positive linear maps on C*-algebras*, J. Oper. Theory **4**, 271 (1980).

[62] Q. Dong, M. T. Quintino, A. Soeda, and M. Murao, *Implementing positive maps with multiple copies of an input state*, Phys. Rev. A **99**, 052352 (2019).

[63] Á. Rivas, S. F. Huelga, and M. B. Plenio, *Entanglement and Non-Markovianity of Quantum Evolutions*, Phys. Rev. Lett. **105**, 050403 (2010).

[64] D. Chruściński and S. Maniscalco, *Degree of Non-Markovianity of Quantum Evolution*, Phys. Rev. Lett. **112**, 120404 (2014).

[65] Á. Rivas, S. F. Huelga, and M. B. Plenio, *Quantum non-Markovianity: Characterization, quantification and detection*, Rep. Prog. Phys. **77**, 094001 (2014).

[66] S. Endo, S. C. Benjamin, and Y. Li, *Practical Quantum Error Mitigation for Near-Future Applications*, Phys. Rev. X **8**, 031027 (2018).

[67] Y. Xiong, D. Chandra, S. X. Ng, and L. Hanzo, *Sampling overhead analysis of quantum error mitigation: Uncoded vs. coded systems*, IEEE Access **8**, 228967 (2020).

[68] A. Shabani and D. A. Lidar, *Maps for general open quantum systems and a theory of linear quantum error correction*, Phys. Rev. A **80**, 012309 (2009).

[69] H. Hakoshima, Y. Matsuzaki, and S. Endo, *Relationship between costs for quantum error mitigation and non-markovian measures*, Phys. Rev. A **103**, 012611 (2021).

[70] I. Devetak and P. W. Shor, *The Capacity of a Quantum Channel for Simultaneous Transmission of Classical and Quantum Information*, Commun. Math. Phys. **256**, 287 (2005).

[71] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, 2012).

[72] V. Paulsen, *Completely Bounded Maps and Operator Algebras* (Cambridge University Press, 2002).
[73] F. Buscemi, M. Dall'Arno, M. Ozawa, and V. Vedral, *Direct observation of any two-point quantum correlation function*, arXiv:1312.4240 (2013).
[74] T. Theurer, D. Egloff, L. Zhang, and M. B. Plenio, *Quantifying Operations with an Application to Coherence*, Phys. Rev. Lett. **122**, 190405 (2019).
[75] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, New York, 2004).
[76] J. P. Ponstein, *Approaches to the Theory of Optimization* (Cambridge University Press, 2004).

Appendix A: Dual forms

Here we derive the dual expressions of the measures, as stated in Prop. 9. The derivation follows standard arguments in convex optimisation [75, 76] (see also [24, App. B]). Let us explicitly consider the case of the diamond norm. As our starting point, we will take the primal optimisation problem as in Lem. 1:

$$\|\Phi\|_{\diamond,p} = \min \left\{ \mu \mid J_\Phi = M_+ - M_-, \ M_\pm \geq 0, \ \mathrm{Tr}_B(M_+ + M_-) \leq \mu \mathbb{1}_A \right\}. \tag{A1}$$

The Lagrangian of this problem is given by

$$\begin{aligned}
L(\mu, M_+, M_-; W, P, Q, R) &= \mu - \langle M_+ - M_- - J_\Phi, W \rangle - \langle M_+, P \rangle - \langle M_-, Q \rangle \\
&\quad - \langle \mu \mathbb{1}_A - \mathrm{Tr}_B(M_+ + M_-), R \rangle \\
&= \mu(1 - \mathrm{Tr}\, R) + \langle M_+, -W - P + R \otimes \mathbb{1}_B \rangle \\
&\quad + \langle M_-, W - Q + R \otimes \mathbb{1}_B \rangle + \langle J_\Phi, W \rangle
\end{aligned} \tag{A2}$$

where $W, P, Q \in \mathbb{H}(A \otimes B)$, $R \in \mathbb{H}(A)$ are Lagrange multipliers, and we used that $\langle S, \mathrm{Tr}_B T \rangle = \langle S \otimes \mathbb{1}_B, T \rangle$ holds for any $S \in \mathbb{H}(A), T \in \mathbb{H}(A \otimes B)$. The dual problem is then defined as

$$\begin{aligned}
\|\Phi\|_{\diamond,d} &:= \sup_{\substack{W \in \mathbb{H} \\ P,Q,R \geq 0}} \inf_{\substack{\mu \in \mathbb{R} \\ M_+, M_- \in \mathbb{H}}} L(\mu, M_+, M_-; W, P, Q, R) \\
&= \sup_{\substack{W \in \mathbb{H} \\ P,Q,R \geq 0}} \begin{cases} \langle J_\Phi, W \rangle & \text{if } \mathrm{Tr}\, R = 1 \text{ and } W + P = R \otimes \mathbb{1}_B \text{ and } W - Q = -R \otimes \mathbb{1}_B \\ -\infty & \text{otherwise} \end{cases} \\
&= \sup \left\{ \langle J_\Phi, W \rangle \mid W \geq -R \otimes \mathbb{1}, \ W \leq R \otimes \mathbb{1}, \ R \geq 0, \ \mathrm{Tr}\, R = 1 \right\},
\end{aligned} \tag{A3}$$

with the supremum achieved since the feasible set is compact. A strictly feasible solution, that is, a feasible solution for which the inequality constraints are strict, can be constructed by decomposing $J_\Phi = J_{\Phi_+} - J_{\Phi_-}$ and defining $M_\pm := J_{\Phi_\pm} + \varepsilon \mathbb{1}_{A \otimes B}$ with μ suitably large. By Slater's theorem (see e.g. [76]), the existence of a strictly feasible solution ensures that $\|\Phi\|_{\diamond,p} = \|\Phi\|_{\diamond,d}$.

The dual forms of the other measures are obtained in full analogy with the derivation above. The crucial observation is that an optimisation of the form

$$\|\Phi\|_\blacklozenge = \min \left\{ \lambda_+ + \lambda_- \mid \Phi = \lambda_+ \Lambda_+ - \lambda_- \Lambda_-, \ \Lambda_\pm \in \mathrm{CPTNI} \right\} \tag{A4}$$

can be rewritten as

$$\|\Phi\|_\blacklozenge = \min \left\{ \lambda_+ + \lambda_- \mid J_\Phi = M_+ - M_-, \ M_\pm \geq 0, \ \mathrm{Tr}_B M_\pm \leq \lambda_\pm \mathbb{1}_A \right\} \tag{A5}$$

which allows us to follow the same approach.