Operational distance and fidelity for quantum channels

Viacheslav P. Belavkin[a]
*Department of Mathematics, University of Nottingham, NG7 2RD Nottingham,
United Kingdom*

Giacomo Mauro D'Ariano[b]
*QUIT Group, INFM Unità di Pavia, Università di Pavia, Dipartimento di Fisica
"A. Volta," via Bassi 6, I-27100 Pavia, Italy and Center for Photonic Communication and
Computing, Department of Electrical and Computer Engineering, Northwestern
University, Evanston, Illinois 60208-3118*

Maxim Raginsky[c]
*Center for Photonic Communication and Computing, Department of Electrical and
Computer Engineering, Northwestern University, Evanston, Illinois 60208-3118*

We define and study a fidelity criterion for quantum channels, which we term the
minimax fidelity, through a noncommutative generalization of maximal Hellinger
distance between two positive kernels in classical probability theory. Like other
known fidelities for quantum channels, the minimax fidelity is well defined for
channels between finite-dimensional algebras, but it also applies to a certain class
of channels between infinite-dimensional algebras (explicitly, those channels that
possess an operator-valued Radon-Nikodym density with respect to the trace in the
sense of Belavkin-Staszewski) and induces a metric on the set of quantum channels
that is topologically equivalent to the CB-norm distance between channels, pre-
cisely in the same way as the Bures metric on the density operators associated with
statistical states of quantum-mechanical systems, derived from the well-known fi-
delity ("generalized transition probability") of Uhlmann, is topologically equivalent
to the trace-norm distance. © *2005 American Institute of Physics.*
[DOI: 10.1063/1.1904510]

I. INTRODUCTION

Many problems in quantum information science,[1,2] both in theory and in experiment, involve
finding a set of quantum-mechanical states or channels that solve some sort of an optimization
problem, typically formulated in terms of a numerical criterion that measures how close a given
pair of states or operations are to each other. (Many such criteria have been proposed to date, each
defined with specific theoretical or experimental considerations in mind; see Ref. 3 for a recent
comprehensive survey.)

Let us first consider the case of quantum states, i.e., density operators. Let $\mathfrak{h}$ be a complex
separable Hilbert space associated to a quantum-mechanical system. Given a pair of density
operators ρ, σ, i.e., positive trace-class operators with unit trace, one can use either the *fidelity*[4–7]

$$F(\rho, \sigma) := \text{Tr}[(\rho^{1/2} \sigma \rho^{1/2})^{1/2}] \tag{1}$$

or the *trace-norm (half-) distance*

---

[a]E-mail address: vpb@maths.nott.ac.uk
[b]Author to whom correspondence should be addressed. E-mail address: dariano@unipv.it
[c]E-mail address: maxim@ece.northwestern.edu

$$D(\rho,\sigma) := \tfrac{1}{2}\|\rho - \sigma\|_{\mathsf{T}}, \tag{2}$$

where $\|\rho\|_{\mathsf{T}} := \mathrm{Tr}|\rho|$ and $|\rho| := (\rho^\dagger \rho)^{1/2}$.[8,9] Loosely speaking, two states ρ and σ are close to each other if $F(\rho,\sigma)$ is large, or if $D(\rho,\sigma)$ is small. In fact, as follows from the key inequality[5,10]

$$1 - F(\rho,\sigma) \leq D(\rho,\sigma) \leq \sqrt{1 - F^2(\rho,\sigma)}, \tag{3}$$

the fidelity and the trace-norm distance are equivalent in the sense that any two density operators that are close to one another in the sense of (1) are also close in the sense of (2), and vice versa.

As for *quantum channels*, i.e., normal completely positive unital mappings from an operator algebra $\mathcal{B}=\mathcal{B}(\mathfrak{h})$ into another algebra $\mathcal{A}=\mathcal{B}(\mathfrak{g})$, where $\mathfrak{g}$ and $\mathfrak{h}$ are complex separable Hilbert spaces, things get somewhat complicated. Consider, for instance, the case when $\mathfrak{g}$ is finite-dimensional, and let $m := \dim \mathfrak{g}$. Fix an orthonormal basis $\{|j\rangle\}_{j=1}^m$ of $\mathfrak{g}$, and let $|\psi\rangle := m^{-1/2}\sum_{j=1}^m |j\rangle \otimes |j\rangle$ be the normalized maximally entangled vector in the product space $\mathfrak{g} \otimes \mathfrak{g}$. Given two quantum channels $\Phi, \Psi : \mathcal{B} \to \mathcal{A}$, one can measure their closeness in terms of the fidelity of the states on $\mathcal{B} \otimes \mathcal{A}$, obtained from the maximally entangled state $\pi = |\psi\rangle\langle\psi|$ by applying the predual channels Φ_{T} and Ψ_{T} (cf. Sec. II for precise definitions) to the first factor in the tensor product:

$$\Phi_{\mathsf{T}} \otimes \mathrm{id}(\pi) = \frac{1}{m}\sum_{i=1}^m \sum_{k=1}^m \Phi_{\mathsf{T}}(|i\rangle\langle k|) \otimes |i\rangle\langle k| \equiv \rho,$$

$$\Psi_{\mathsf{T}} \otimes \mathrm{id}(\pi) = \frac{1}{m}\sum_{i=1}^m \sum_{k=1}^m \Psi_{\mathsf{T}}(|i\rangle\langle k|) \otimes |i\rangle\langle k| \equiv \sigma.$$

The fidelity $F(\rho,\sigma)$, taken as the *channel fidelity*

$$\mathcal{F}(\Phi,\Psi) := F(\Phi_{\mathsf{T}} \otimes \mathrm{id}(\pi), \quad \Psi_{\mathsf{T}} \otimes \mathrm{id}(\pi)), \tag{4}$$

by Raginsky in Ref. 11, enjoys many properties parallel to those of the fidelity (1) for quantum states. Alternatively, one can adopt the (half-) distance[1,12,13]

$$\mathcal{D}(\Phi,\Psi) := \tfrac{1}{2}\|\Phi - \Psi\|_{\mathrm{CB}}, \tag{5}$$

where $\|\cdot\|_{\mathrm{CB}}$ denotes the so-called *norm of complete boundedness* (or CB-norm for short; cf. Sec. II C for details). We note that the CB-norm half-distance (5) can be given in terms of the trace-norm distance (2) between density operators by means of the variational expression[1,12,13]

$$\mathcal{D}(\Phi,\Psi) = \sup_\pi D(\Phi_{\mathsf{T}} \otimes \mathrm{id}(\pi), \quad \Psi_{\mathsf{T}} \otimes \mathrm{id}(\pi)), \tag{6}$$

where the supremum is taken over all density operators π on the tensor product space $\mathfrak{g} \otimes \mathfrak{g}$. By analogy with density operators of the states, we are tempted to say that two quantum channels, Φ and Ψ, are close either if $\mathcal{F}(\Phi,\Psi)$ is large or if $\mathcal{D}(\Phi,\Psi)$ is small. However, in addition to the finite-dimension restriction $\dim \mathfrak{g} < \infty$ [the only case under which the definition (4) of the channel fidelity makes sense], we encounter the following difficulty. It turns out[11] that, as a criterion of closeness, the CB-norm distance (5) is strictly stronger than the fidelity measure (4) in the sense that even when $\mathcal{D}(\Phi,\Psi)$ is large, $\mathcal{F}(\Phi,\Psi)$ may be quite large as well, and may even become equal to one in the limit $\dim \mathfrak{g} \to \infty$. Consider, for instance, the case $\Psi = \mathrm{id}$. Then one can show[11] that

$$1 - \mathcal{D}(\Phi,\mathrm{id}) \leq \mathcal{F}(\Phi,\mathrm{id}) \leq \sqrt{1 - (1/4)\mathcal{D}^2(\Phi,\mathrm{id})}, \tag{7}$$

and we immediately see that when Φ is such that $\mathcal{D}(\Phi,\mathrm{id})$ attains its maximum value of unity, the fidelity $\mathcal{F}(\Phi,\mathrm{id})$ is still bounded between 0 and $\sqrt{3}/2$. To make matters worse, the only bound on (5) in terms of (4) known so far is

$$1 \geqslant \mathcal{D}(\Phi, \Psi) \geqslant 1 - \mathcal{F}(\Phi, \Psi), \tag{8}$$

as follows readily from Eqs. (3) and (6). Furthermore, one can easily find sequences $\{\Phi_m\}$, $\{\Psi_m\}$ of channels $\Phi_m, \Psi_m : \mathcal{B}(\mathbb{C}^m) \to \mathcal{B}(\mathbb{C}^m)$, such that $\mathcal{D}(\Phi_m, \Psi_m) \neq 0$ for all m, while

$$\lim_{m \to \infty} \mathcal{F}(\Phi_m, \Psi_m) = 1.$$

Indeed, consider the unitarily implemented channels

$$\Phi_m(B) = U_m^\dagger B U_m, \quad \Psi(B) = V_m^\dagger B V_m$$

with the unitaries U_m, V_m chosen in such a way that $U_m \neq V_m$ but

$$\lim_{m \to \infty} \frac{1}{m} \mathrm{Tr}(U_m^\dagger V_m) = 1.$$

Thus, the channel fidelity (4), apart from being applicable only in finite-dimensional settings, has the distinct disadvantage of not being equivalent to the CB-norm distance, in contrast to the case of the Uhlmann fidelity (1) and the trace-norm distance (2) on the state space of a quantum-mechanical system.

The goal of this paper is to define and study a new fidelity criterion for quantum channels, which we term the *minimax fidelity* and which is a noncommutative generalization of maximal Hellinger distance between two positive kernels in classical probability theory. Unlike the channel fidelity (4) of Ref. 11, the minimax fidelity is not only well defined for channels between finite-dimensional algebras, but also applies to a certain class of channels between infinite-dimensional algebras (explicitly, those channels that possess an operator-valued Radon-Nikodym density with respect to the trace in the sense of Belavkin-Staszewski[14]) and is equivalent to the CB-norm distance, echoing the way the Uhlmann fidelity (1) for density operators is equivalent to the trace-norm distance (2).

Apart from these technical features, the minimax fidelity $f(\Phi, \Psi)$ between two quantum channels Φ, Ψ has a direct *operational* meaning: intuitively, it is defined as the minimum overlap of output states (density operators) of the predual channels $\Phi_\mathsf{T}, \Psi_\mathsf{T}$, when the operator-sum decompositions[2] of the latter are chosen to be maximally overlapping; this is spelled out in precise terms in Sec. IV E. Our central result (Theorem 1) demonstrates that the minimax fidelity is independent of the order of these two optimizations. Furthermore, the equivalence of our minimax fidelity to the CB-norm distance, which is stated precisely in Sec. V in terms of *dimension-free* bounds, is a promising avenue for the study and characterization of dimension-free bounds (whenever they exist) on other operationally meaningful distance measures for quantum operations[3] in terms of the CB-norm distance. As pointed out in Ref. 15, such bounds are crucial for a successful generalization of the usual quantum capacity of a channel[1,2] (i.e., with respect to the identity channel) to the case of comparing quantum channels to an arbitrary reference channel. We plan to pursue these matters further in a future publication.

The paper is organized as follows. In Sec. II we fix the definitions and notation used throughout the paper. The minimax fidelity is then introduced in Sec. III. Section IV is devoted to the evaluation of the minimax fidelities in the various mathematical settings that arise in quantum information theory. Next, in Sec. V, we list key properties of the minimax fidelity. Finally, in Sec. VI we sketch some example applications of the minimax fidelity to several problems of quantum information theory.

II. PRELIMINARIES, DEFINITIONS, NOTATION

A. Pairings, states, operations

Let $\mathfrak{h}$ be a complex separable Hilbert space; let $\mathcal{B}$ denote the Banach algebra $\mathcal{B}(\mathfrak{h})$ of all bounded linear operators on $\mathfrak{h}$ with the usual operator norm $\|\cdot\|$; and let $\mathcal{B}_\mathsf{T}$ denote the Banach space $\mathcal{B}_\mathsf{T}(\mathfrak{h})$ of trace-class operators on $\mathfrak{h}$ with the trace norm $\|\cdot\|_\mathsf{T}$. The set of normal states on $\mathcal{B}$,

i.e., ultraweakly continuous positive unital linear functionals on $\mathcal{B}$, will be denoted by $\mathcal{S}(\mathcal{B})$ or, whenever we need to exhibit the underlying Hilbert space explicitly, by $\mathcal{S}(\mathfrak{h})$. Generic elements of $\mathcal{S}(\mathcal{B})$ will be denoted by the stylized Greek letters $\varpi, \varrho, \varsigma$. Note that the operator norm on $\mathcal{B}$ can be written as $\|B\| = \sup\{\varrho(|B|) : \varrho \in \mathcal{S}(\mathcal{B})\}$.

We equip $\mathfrak{h}$ (and shall equip all Hilbert spaces introduced in the sequel) with an isometric involution $J = J^{\dagger}$, $J^2 = \mathbb{1}_{\mathfrak{h}}$, having the properties of complex conjugation,

$$J\sum_j \lambda_j \eta_j = \sum_j \overline{\lambda}_j J\eta_j, \quad \forall \, \lambda_j \in \mathbb{C}, \; \eta_j \in \mathfrak{h}.$$

We can thus define the *transpose* of any $B \in \mathcal{B}$ as $\widetilde{B} := JB^{\dagger}J$, as well as introduce the trace pairing[16]

$$(B,\rho) := \mathrm{Tr}(B\widetilde{\rho}) = \mathrm{Tr}(\widetilde{B}\rho), \quad \forall \, B \in \mathcal{B}, \; \rho \in \mathcal{B}_{\mathsf{T}} \tag{9}$$

of $\mathcal{B}$ and $\mathcal{B}_{\mathsf{T}}$. Under this pairing, which differs from the usual one in that $B \in \mathcal{B}$ is paired with the transpose of $\rho \in \mathcal{B}_{\mathsf{T}}$ rather than directly with ρ, normal linear functionals on $\mathcal{B}$ are in a one-to-one correspondence with the elements of $\mathcal{B}_{\mathsf{T}}$. Thus to each normal state ϱ we associate a unique positive trace-class operator with unit trace, denoted by the standard Greek letter ρ and referred to as the *density operator* corresponding to ϱ, via $\varrho(B) = (B,\rho)$ for all $B \in \mathcal{B}$. Similarly, density operators corresponding to states denoted by ϖ and ς will be denoted by π and σ, respectively.

Apart from natural arguments from standard representation theory of operator algebras, one reason why we chose to pair B with the transposed operator $\widetilde{\rho} = J\rho^{\dagger}J$, rather than with ρ, is to be able to keep all notations conveniently parallel to the classical (commutative) case, as will be amply demonstrated throughout the paper. Note also that we can fix a complete orthonormal basis $\{|j\rangle\}$ of $\mathfrak{h}$ and express the pairing (9) in terms of the matrix elements of B and ρ as

$$(B,\rho) = \sum_{j,k} \langle j|B|k\rangle \cdot \langle j|\rho|k\rangle \equiv \sum_{j,k} B_{jk}\rho^{jk},$$

where we have used the covariant indices for the matrix elements of bounded operators in $\mathcal{B}$ and the contravariant indices for the matrix elements of trace-class operators in $\mathcal{B}_{\mathsf{T}}$, when the latter are identified via the pairing (9) with normal linear functionals on $\mathcal{B}$. Yet another reason to opt for the pairing of B with the transposed operator $\widetilde{\rho}$, further elaborated upon in Sec. II B, is that then the density operator ρ of a normal state ϱ will coincide with the operational density of ϱ, understood as a quantum operation from $\mathcal{B}$ into the Abelian algebra $\mathbb{C}$.

Introducing another Hilbert space $\mathfrak{g}$, the algebra $\mathcal{A} := \mathcal{B}(\mathfrak{g})$, and the trace class $\mathcal{A}_{\mathsf{T}} := \mathcal{B}_{\mathsf{T}}(\mathfrak{g})$, let us consider *quantum operations*, i.e., the completely positive normal linear mappings $\Phi : \mathcal{B} \to \mathcal{A}$ such that $\Phi(\mathbb{1}_{\mathfrak{h}}) \leq \mathbb{1}_{\mathfrak{g}}$; if $\Phi(\mathbb{1}_{\mathfrak{h}}) = \mathbb{1}_{\mathfrak{g}}$, then Φ is referred to as a *quantum channel*. Any quantum operation Φ possesses a unique *predual* $\Phi_{\mathsf{T}} : \mathcal{A}_{\mathsf{T}} \to \mathcal{B}_{\mathsf{T}}$, defined as the transpose of Φ with respect to the trace pairing (9), i.e.,

$$(\Phi(B),\rho) = (B,\Phi_{\mathsf{T}}(\rho)), \quad \forall \, B \in \mathcal{B}, \; \rho \in \mathcal{A}_{\mathsf{T}}. \tag{10}$$

Conversely, given a normal completely positive linear map $\Phi : \mathcal{A}_{\mathsf{T}} \to \mathcal{B}_{\mathsf{T}}$ such that $\mathrm{Tr}_{\mathfrak{h}}\Phi(\rho) \leq \mathrm{Tr}_{\mathfrak{g}}\rho$ for all $\rho \in \mathcal{A}_{\mathsf{T}}$, we define its *dual* with respect to the trace pairing (9) as the unique mapping $\Phi^{\mathsf{T}} : \mathcal{B} \to \mathcal{A}$ for which

$$(B,\Phi(\rho)) = (\Phi^{\mathsf{T}}(B), \rho), \quad \forall \, B \in \mathcal{B}, \rho \in \mathcal{A}_{\mathsf{T}}. \tag{11}$$

Using these definitions, one readily obtains that $\Phi_{\mathsf{T}}^{\mathsf{T}} = \Phi$ for any normal completely positive map $\Phi : \mathcal{B} \to \mathcal{A}$. Alternatively, one may define the predual of a normal completely positive map $\Phi : \mathcal{B} \to \mathcal{A}$ as the unique normal completely positive map $\Phi_{\mathsf{T}} : \mathcal{A}_{\mathsf{T}} \to \mathcal{B}_{\mathsf{T}}$ such that $\Phi_{\mathsf{T}}^{\mathsf{T}} = \Phi$.

If Φ is given in the Kraus form[17] $\Phi(B) = \Sigma F_j^{\dagger} B F_j$, or more generally as an integral

$$\Phi(B) = \int_Z F(z)^\dagger B F(z) \mathrm{d}\mu(z), \tag{12}$$

with respect to a positive measure μ on a measurable space $(Z, \mathcal{B}_Z)$, where the integration is understood in the sense of Bochner,[18] then the predual map Φ_T has the transposed integral form

$$\Phi_\mathsf{T}(\rho) = \int_Z F_\mathsf{T}(z)^\dagger \rho F_\mathsf{T}(z) \mathrm{d}\mu(z),$$

where $\mathfrak{g} \ni \xi \mapsto \langle \xi | F_\mathsf{T}(z)$ are Hilbert-transposed to the operators $\mathfrak{h} \ni \eta \mapsto \langle \eta | F(z)$, that is $F_\mathsf{T}(z) = \widetilde{F(z)}$ for all $z \in Z$.

Any normal state $\varrho \in \mathcal{S}(\mathcal{B})$ is automatically a quantum channel from $\mathcal{B}$ into the Abelian algebra $\mathbb{C}$, and it is readily seen that the density operator ρ of ϱ, understood as acting on $\lambda \in \mathbb{C}$ on the right, $\mathbb{C} \ni \lambda \mapsto \lambda \rho$, is precisely the predual $\varrho_\mathsf{T} : \mathbb{C} \to \mathcal{B}_\mathsf{T}$. Indeed, given $B \in \mathcal{B}$ and $\lambda \in \mathbb{C}$, we have

$$(\varrho(B), \lambda) = (B, \lambda \rho) = (B, \varrho_\mathsf{T}(\lambda)),$$

which proves our claim that $\rho = \varrho_\mathsf{T}$. Thus we also have that $\varrho = \varrho_\mathsf{T}^\mathsf{T} = \rho^\mathsf{T}$.

B. Operational densities

In order to avoid technicalities involving unbounded operators, we shall henceforth assume that all quantum operations we deal with are *completely majorized* by the trace, considered as the map $\tau(\sigma) = \mathbb{1}_\mathfrak{g} \mathrm{Tr}\, \sigma$ of $\mathcal{B}_\mathsf{T}$ into $\mathcal{A} = \mathcal{B}(\mathfrak{g})$, in the sense[14] that there exists a constant $\lambda > 0$ such that the difference $\lambda \tau - \Phi$ is a completely positive map $\mathcal{B}_\mathsf{T} \to \mathcal{A}$. For example, this condition is satisfied by all quantum operations between finite-dimensional algebras.[19] As was proven in Ref. 14, in this case there exists a unique positive operator Φ_τ on the Hilbert space $\mathcal{H} := \mathfrak{g} \otimes \mathfrak{h}$, called the *density* of Φ with respect to the trace τ, such that

$$\Phi(B) = \mathrm{Tr}_\mathfrak{h}[(\mathbb{1}_\mathfrak{g} \otimes \widetilde{B}) \Phi_\tau], \tag{13}$$

where $\mathrm{Tr}_\mathfrak{h}\, Y$, $Y \in \mathcal{B}(\mathcal{H})$, denotes the partial trace of Y with respect to $\mathfrak{h}$,

$$(\mathrm{Tr}_\mathfrak{h} Y, \rho) = (Y, \rho \otimes \mathbb{1}_\mathfrak{h}), \quad \forall\, \rho \in \mathcal{B}_\mathsf{T}(\mathfrak{g}).$$

Moreover, Φ_τ as a linear operator on $\mathcal{H}$ is bounded and majorized by $\lambda : 0 \leqslant \Phi_\tau \leqslant \lambda \mathbb{1}_\mathcal{H}$, and the operation is unital, $\Phi(\mathbb{1}_\mathfrak{h}) = \mathbb{1}_g$ [contractive, $\Phi(\mathbb{1}_\mathfrak{h}) \leqslant \mathbb{1}_\mathfrak{g}$] if and only if $\mathrm{Tr}_\mathfrak{h}\, \Phi_\tau = \mathbb{1}_\mathfrak{g}$ ($\mathrm{Tr}_\mathfrak{h}\, \Phi_\tau \leqslant \mathbb{1}_\mathfrak{g}$). This is equivalent to saying that the predual map $\Phi_\mathsf{T} : \mathcal{A}_\mathsf{T} \to \mathcal{B}_\mathsf{T}$, which, using Eqs. (10) and (13), can be written as

$$\Phi_\mathsf{T}(\rho) = \mathrm{Tr}_\mathfrak{g}[\Phi_\tau(\widetilde{\rho} \otimes \mathbb{1}_\mathfrak{h})], \tag{14}$$

is trace preserving (trace decreasing).

As an example, consider a normal state ϱ on $\mathcal{B}$, which, being a quantum channel into $\mathbb{C}$, satisfies the complete majorization condition with $\lambda = \|\rho\|$, where ρ is the density operator of ϱ. Furthermore, it is easy to see that $\varrho_\tau = \rho$. Indeed, we can write

$$\varrho(B) = (B, \rho) = \mathrm{Tr}(B\widetilde{\rho}) = \mathrm{Tr}(\widetilde{B}\rho) = \mathrm{Tr}_\mathfrak{h}[(\mathbb{1}_\mathbb{C} \otimes \widetilde{B})\rho],$$

and the desired result follows upon comparing this with Eq. (13). This provides additional justification for our definition of the trace pairing in Eq. (9), since we then have that $\varrho_\mathsf{T} = \rho = \varrho_\tau$ for any normal state ϱ.

If the operation $\Phi : \mathcal{B} \to \mathcal{A}$ is given in the generalized Kraus form (12), we can write down its operational density Φ_τ explicitly. To this end, suppose that all operators $F(z)$ are determined by generalized bra-vectors $\Gamma(z) = (F(z)|$, densely defined as the linear functionals

$$\Gamma(z)|\xi \otimes \eta\rangle = \langle\xi|F(z)|\eta\rangle \equiv (F(z)||\xi \otimes \eta\rangle)$$

on the linear span of the ket-vectors $|\xi \otimes \eta\rangle \equiv \widetilde{\xi \otimes \eta}$ in $\mathcal{H} = \mathfrak{g} \otimes \mathfrak{h}$, where $\xi \in \mathfrak{g}$ is also treated as a bra-vector such that $J\xi = \langle\xi|$ and $|\xi\rangle = \tilde{\xi}$. Then the operational density Φ_τ of Φ is given by the corresponding decomposition

$$\Phi_\tau = \int \Gamma(z)^\dagger \Gamma(z) \mathrm{d}\mu(z) \equiv \Gamma^\dagger \Gamma, \tag{15}$$

where the integral is, again, understood in the sense of Bochner.

C. Completely bounded maps

Completely positive linear maps between operator algebras are a special case of *completely bounded* maps.[20] Consider, as before, the algebras $\mathcal{B} = \mathcal{B}(\mathfrak{h})$ and $\mathcal{A} = \mathcal{B}(\mathfrak{g})$. For each $n \in \mathbb{N}$ define the nth *matrix level* $\mathcal{M}_n(\mathcal{B}) \simeq \mathcal{B} \otimes \mathcal{M}_n$, where $\mathcal{M}_n$ denotes the algebra of $n \times n$ matrices with complex entries. That is, $\mathcal{M}_n(\mathcal{B})$ is the space of $n \times n$ matrices with $\mathcal{B}$-valued entries,

$$\mathcal{M}_n(\mathcal{B}) := \{[B_{ij}] : B_{ij} \in \mathcal{B}, 1 \leq i, j \leq n\}.$$

Analogous construction can also be applied to $\mathcal{A}$ to yield the matrix levels $\mathcal{M}_n(\mathcal{A})$. Each matrix level $\mathcal{M}_n(\mathcal{B})$ inherits a *-algebra structure from $\mathcal{B}$ through

$$[B_{ij}][C_{ij}] := \left[\sum_{k=1}^n B_{ik}C_{kj}\right], \quad [B_{ij}]^\dagger := [B_{ji}^\dagger].$$

In fact, by identifying $\mathcal{M}_n(\mathcal{B})$ via a natural *-isomorphism with the algebra $\mathcal{B}(\mathfrak{h}^{(n)})$ of bounded linear operators on $\mathfrak{h}^{(n)}$, the direct sum of n copies of $\mathfrak{h}$, one can make $\mathcal{M}_n(\mathcal{B})$ into a C*-algebra. Thus, each matrix level of $\mathcal{B}$ possesses a unique C*-norm.

Now, for any $n \in \mathbb{N}$ a linear map $\Lambda : \mathcal{B} \to \mathcal{A}$ induces the map $\Lambda^{(n)} := \Lambda \otimes \mathrm{id}_n$ from $\mathcal{M}_n(\mathcal{B})$ into $\mathcal{M}_n(\mathcal{A})$, defined by $\Lambda^{(n)} : [B_{ij}] \mapsto [\Lambda(B_{ij})]$. Let us define the *norm of complete boundedness* (or CB-norm) by $\|\Lambda\|_{\mathrm{CB}} := \sup\{\|\Lambda^{(n)}\| : n \in \mathbb{N}\}$, where

$$\|\Lambda^{(n)}\| := \sup_{B \in \mathcal{M}_n(\mathcal{B}), \|B\| \leq 1} \|\Lambda^{(n)}(B)\|$$

is the usual operator norm of $\Lambda^{(n)}$. A linear map $\Lambda : \mathcal{B} \to \mathcal{A}$ is called *completely bounded* if $\|\Lambda\|_{\mathrm{CB}} < \infty$. Every completely positive map $\Phi : \mathcal{B} \to \mathcal{A}$ is automatically completely bounded, with $\|\Phi\|_{\mathrm{CB}} = \|\Phi(1_\mathfrak{h})\|$. For a general completely bounded map Λ, one has, by definition, $\|\Lambda(1_\mathfrak{h})\| \leq \|\Lambda\| \leq \|\Lambda\|_{\mathrm{CB}}$.

Passing to the predual map $\Lambda_\mathsf{T} : \mathcal{A}_\mathsf{T} \to \mathcal{B}_\mathsf{T}$, we can similarly define induced maps $\Lambda_\mathsf{T}^{(n)} : \mathcal{M}_n(\mathcal{A}_\mathsf{T}) \to \mathcal{M}_n(\mathcal{B}_\mathsf{T})$, $n \in \mathbb{N}$, and the predual CB-norm

$$\|\Lambda\|_{\mathrm{CB}}^\mathsf{T} := \sup_{n \in \mathbb{N}} \|\Lambda^{(n)}\|_\mathsf{T},$$

where

$$\|\Lambda^{(n)}\|_\mathsf{T} := = \sup_{\rho \in \mathcal{M}_n(\mathcal{A}_\mathsf{T}) : \|\rho\|_\mathsf{T} \leq 1} \|\Lambda^{(n)}(\rho)\|_\mathsf{T}.$$

It is easy to see that $\|\Lambda^{(n)}\| = \|\Lambda_\mathsf{T}^{(n)}\|_\mathsf{T}$ for all $n \in \mathbb{N}$, so that $\|\Lambda\|_{\mathrm{CB}} = \|\Lambda_\mathsf{T}\|_{\mathrm{CB}}^\mathsf{T}$. It is also straightforward to see that the "unstabilized" norms $\|\cdot\|$ and $\|\cdot\|_\mathsf{T}$ are tensor supermultiplicative (i.e., $\|\Lambda_1 \otimes \Lambda_2\| \geq \|\Lambda_1\| \|\Lambda_2\|$), whereas the corresponding CB-norms are tensor multiplicative (i.e., $\|\Lambda_1 \otimes \Lambda_2\|_{\mathrm{CB}} = \|\Lambda_1\|_{\mathrm{CB}} \|\Lambda_2\|_{\mathrm{CB}}$).

There is also a useful nonvariational formula for the CB-norm of a map $\Lambda : \mathcal{B} \to \mathcal{A}$. Namely, let ℓ^2 denote the Hilbert space of square-summable infinite sequences of complex numbers, and let $\mathcal{K}(\ell^2)$ denote the space of compact operators on ℓ^2. Then $\|\Lambda\|_{\mathrm{CB}} = \|\Lambda \otimes \mathrm{id}_{\mathcal{K}(\ell^2)}\|$. Since we have

assumed that $\mathcal{B}=\mathcal{B}(\mathfrak{h})$ with $\mathfrak{h}$ a complex separable Hilbert space, and since all complex separable Hilbert spaces are canonically isomorphic to ℓ^2, we may also write $\|\Lambda\|_{\mathrm{CB}}=\|\Lambda\otimes\mathrm{id}_{\mathcal{K}(\mathfrak{h})}\|$.

D. Miscellany

Any positive operator $B\in\mathcal{B}(\mathfrak{h})$ has a unique positive square root, denoted by $B^{1/2}$ and defined as the positive operator $X\in\mathcal{B}(\mathfrak{h})$ such that $B=X^2$. This definition can be extended to any operator A that is similar to a positive operator $\Delta\in\mathcal{B}(\mathfrak{h})$, in the sense that there exists an operator $S\in\mathcal{B}(\mathfrak{h})$ such that $A=S\Delta S^+$, where S^+ is the pseudoinverse of S, equal to S^{-1} on ran S and to 0 on ker S. In that case, we may *define* $\sqrt{A}:=S\Delta^{1/2}S^+$. From now on, in order to distinguish this extended definition of the square root from the usual one, we shall always use the square root symbol $\sqrt{\cdot}$ for this extended definition, and reserve the exponent notation $\cdot^{1/2}$ for the usual definition.

Consider now two positive operators $A,B\in\mathcal{B}(\mathfrak{h})$. It is easy to see that their product AB is similar to $A^{1/2}BA^{1/2}$ with $S=A^{1/2}$. Note that the operator AB is positive when restricted to the closure of ran A, when the latter is equipped with the weighted inner product $\langle v|\chi\rangle_A:=\langle A^{-1/2}v|A^{-1/2}\chi\rangle$:

$$\langle v|ABv\rangle_A=\langle A^{-1/2}v|A^{-1/2}ABv\rangle=\langle v|Bv\rangle\geqslant 0,\quad\forall\, v\in\overline{\mathrm{ran}A}.$$

Thus we may define $\sqrt{AB}:=S(A^{1/2}BA^{1/2})S^+$ with $S=A^{1/2}$.

This notation, again, allows for a convenient parallelism between the classical (commutative) formalism and the quantum (noncommutative) one. Indeed, consider two mutually commuting positive trace-class operators ρ, σ, let $\{|x\rangle\}$ denote the set of their common eigenvectors, and let $\rho_x\equiv\langle x|\rho|x\rangle$, $\sigma_x\equiv\langle x|\sigma|x\rangle$ denote the corresponding eigenvalues. Then $\sqrt{\rho\sigma}$ is also trace-class, and

$$\mathrm{Tr}\sqrt{\rho\sigma}=\sum_x\sqrt{\rho_x\sigma_x}.$$

If $\mathrm{Tr}\,\rho=1=\mathrm{Tr}\,\sigma$, then $P:=\{\rho_x\}$ and $Q:=\{\sigma_x\}$ are probability distributions, and $\mathrm{Tr}\sqrt{\rho\sigma}$ then gives the classical fidelity (also known as the *Bhattacharyya coefficient*)[10] $F(P,Q)$ between P and Q.

Our main technical tool in this paper is given by the following:

Lemma 1: Let $\mathcal{H}$ be a complex separable Hilbert space, and let R, $S\in\mathcal{B}(\mathcal{H})$ be positive operators such that $R^{1/2}SR^{1/2}$ is trace class. Then the supremum

$$\sup_{X,Y\in\mathcal{B}(\mathcal{H})}\{\mathrm{Tr}(X^\dagger Y+Y^\dagger X):X^\dagger X=R,Y^\dagger Y=S\}=2\,\mathrm{Tr}\sqrt{RS}\qquad(16)$$

is achieved on any $X\in\mathcal{B}(\mathcal{H})$ satisfying the condition $X^\dagger X=R$, say $X=R^{1/2}$, and $Y=Y_o$ satisfying the equation

$$Y_oX^\dagger=(XSX^\dagger)^{1/2}=XY_o^\dagger.\qquad(17)$$

Proof: To prove the lemma one can use either the polar decomposition or the method of Lagrange multipliers. We shall use the latter. Fixing an X satisfying $X^\dagger X=R$, we can write the Lagrange function as

$$\mathsf{L}=\mathrm{Tr}(X^\dagger Y+Y^\dagger X-Y^\dagger YL),$$

where $L=L^\dagger\in\mathcal{B}(\mathcal{H})$ is the operator-valued Lagrange multiplier corresponding to the hermiticity condition $S=Y^\dagger Y=S^\dagger$. At the stationary point

$$\delta\mathsf{L}=\mathrm{Tr}(X^\dagger-LY^\dagger)\delta Y+(X-YL)\delta Y^\dagger=0,$$

so $Y=Y_o$ must satisfy the equation $YL=X$ (the other equation, $LY^\dagger=X^\dagger$, corresponding to $Y^\dagger=Y_o^\dagger$, is obtained by taking the Hermitian adjoint). Thus $Y_o=XL^{-1}$, where L^{-1} should be determined from $L^{-1}X^\dagger XL^{-1}=S$. Multiplying this on the left by X and on the right by $X^\dagger$ yields $(XL^{-1}X^\dagger)^2=XSX^\dagger$, or

$XL^{-1}X^\dagger=(XSX^\dagger)^{1/2}$. Thus, we indeed have that $Y_oX^\dagger=(XSX^\dagger)^{1/2}=XY_o^\dagger$, and therefore that

$$\mathrm{Tr}(Y_oX^\dagger + XY_o^\dagger) = 2\,\mathrm{Tr}((XSX^\dagger)^{1/2}). \tag{18}$$

This extremal value is precisely the maximal value due to convexity of the function being maximized in Eq. (16). Note that, since $(U^\dagger XSX^\dagger U)^{1/2}=U^\dagger(XSX^\dagger)^{1/2}U$ for any unitary U, the value of the supremum in Eq. (16), which coincides with Eq. (18), does not depend on the choice of X satisfying $X^\dagger X=R$. Indeed, by virtue of the polar decomposition $X=UR^{1/2}$,

$$2\,\mathrm{Tr}(XSX^\dagger)^{1/2} = 2\,\mathrm{Tr}(U^\dagger(XSX^\dagger)^{1/2}U) = 2\,\mathrm{Tr}((R^{1/2}SR^{1/2})^{1/2}).$$

Rewriting this trace in the equivalent form $2\,\mathrm{Tr}(X^\dagger Y_o)$ with

$$X^\dagger Y_o = R^{1/2}(R^{1/2}SR^{1/2})^{1/2}R^{-1/2} \equiv \sqrt{RS}$$

corresponding to $X=R^{1/2}$, we obtain the extremal value in Eq. (16). $\qquad\square$

We shall also need the following simple, but useful, result:

Lemma 2: Let S be a compact subset of a complex Banach space V, such that $x \in S$ implies $\lambda x \in S$ for all $\lambda \in \mathbb{C}$ with $|\lambda|=1$. Let $f:V \to \mathbb{C}$ be a continuous function which is homogeneous of order 1, i.e., $f(\lambda x)=\lambda f(x)$ for all $\lambda \in \mathbb{C}$ and all $x \in V$. Then

$$\sup_{x \in S}|f(x)| = \sup_{x \in S}\mathrm{Re}\,f(x). \tag{19}$$

Proof: Let $x^* \in S$ be such that $|f(x^*)|=\sup_{x \in S}|f(x)|$, with $f(x^*)=|f(x^*)|e^{i\,\arg f(x^*)}$. Let $x^{**} := e^{-i\,\arg f(x^*)}x^*$. By the homogeneity of f,

$$f(x^{**}) = e^{-i\,\arg f(x^*)}f(x^*) = |f(x^*)|.$$

But then $|f(x^{**})|=f(x^{**})=\mathrm{Re}\,f(x^{**})$. Since $\mathrm{Re}\,\lambda \leq |\lambda|$ for all $\lambda \in \mathbb{C}$, the lemma is proved. $\qquad\square$

III. OPERATIONAL FIDELITIES AND DISTANCES

A. Classical kernel fidelity

The fidelity distinguishing different quantum operations without the restriction on the Hilbert space dimensionality was suggested by Belavkin in Ref. 21 on the basis of a noncommutative generalization of the maximal Hellinger distance between two positive kernels. Namely, given a locally compact space X and a measure space $(Y, \mathcal{B}_Y, \mu)$, where μ is a positive measure, let us denote by $\mathcal{A}$ the algebra $\mathcal{C}(X)$ of bounded continuous functions on X, and by $\mathcal{B}_\mathsf{T} \equiv \mathcal{C}_\mathsf{T}(Y)$ the space of absolutely μ-integrable complex functions on Y. A *positive kernel* $\mathbf{P}$ is then given in terms of a function $p(\cdot|\cdot):Y \times X \to \mathbb{R}^+$, such that $P_x := p(\cdot|x) \in \mathcal{B}_\mathsf{T}$ for all $x \in X$, while $P := \int_Y p(y|\cdot)\mathrm{d}\mu(y) \in \mathcal{A}$. Given two positive kernels $\mathbf{P}$ and $\mathbf{Q}$, the squared pointwise Hellinger distance

$$d_H^2(P_x,Q_x) := \frac{1}{2}\int (\sqrt{p(y|x)} - \sqrt{q(y|x)})^2\mathrm{d}\mu(y) = \int\left[\frac{1}{2}(p(y|x) + q(y|x)) - \sqrt{p(y|x)q(y|x)}\right]\mathrm{d}\mu(y) \tag{20}$$

is well defined and finite for each $x \in X$, so that we can define

$$d_H^2(\mathbf{P},\mathbf{Q}) := \frac{1}{2}\sup_{x \in X}\int (\sqrt{p(y|x)} - \sqrt{q(y|x)})^2\mathrm{d}\mu(y) \equiv \|d_H^2(P_x,Q_x)\|, \tag{21}$$

the last expression indicating the fact that $d_H^2(\mathbf{P},\mathbf{Q})$ is given by the supremum of the squared pointwise Hellinger distance (20) over all $x \in X$. Note that the squared Hellinger distance $d_H^2(P,Q)$ between two positive distributions $P=p(\cdot)$ and $Q=q(\cdot)$ is the minimal mean quadratic distance

$$d_H^2(P,Q) = \frac{1}{2} \inf_{\chi,\psi \in C(Y)} \left\{ \int |\chi(y) - \psi(y)|^2 d\mu(y) : |\chi(\cdot)|^2 = p(\cdot), |\psi(\cdot)|^2 = q(\cdot) \right\}$$

$$= \left(1, \tfrac{1}{2}(P+Q)\right) - \sup_{\psi : |\psi(\cdot)|^2 = q(\cdot)} \int \sqrt{p(y)} \, \mathrm{Re} \, \psi(y) d\mu(y), \tag{22}$$

where $(f,P) = \int f(y)p(y)d\mu(y)$ denotes the integral pairing of $f \in C(Y)$ with $P \in C_T(Y)$. The *relative fidelity*

$$f(P,Q) = \frac{1}{\sqrt{(1,P)(1,Q)}} \sup_{\psi : |\psi(\cdot)|^2 = q(\cdot)} \int \sqrt{p(y)} \, \mathrm{Re} \, \psi(y) d\mu(y) = \frac{(1,\sqrt{PQ})}{\sqrt{(1,P)(1,Q)}} \tag{23}$$

of the distributions P and Q is obviously related to the distance (22) by

$$d_H^2(P,Q) + \sqrt{(1,P)(1,Q)} f(P,Q) = \left(1, \tfrac{1}{2}(P+Q)\right). \tag{24}$$

If $P_x := p(\cdot|x)$ and $Q_x := q(\cdot|x)$ are conditional distributions with constant integrals $(1,P_x)$ and $(1,Q_x)$, e.g., normalized to unity, this relation also remains valid for the minimal fidelity

$$f(\mathbf{P},\mathbf{Q}) = \inf_{x \in X} f(P_x, Q_x),$$

which can alternatively be defined by the minimax formula

$$f(\mathbf{P},\mathbf{Q}) = \inf_{x \in X} \sup_{\psi : |\psi(\cdot|x)|^2 = Q_x(\cdot)} \frac{(1, \sqrt{P_x} \, \mathrm{Re} \, \psi(\cdot|x))}{\sqrt{(1,P_x)(1,Q_x)}}, \tag{25}$$

where the supremum is achieved on $\psi(\cdot|x) \equiv \psi_\circ(\cdot|x)$ satisfying $\psi(y|x) = \sqrt{q(y|x)}$. In particular, if $\mathbf{P}$ and $\mathbf{Q}$ are probability kernels, $(1,P_x) = 1 = (1,Q_x)$ for all $x \in X$, then

$$d_H^2(\mathbf{P},\mathbf{Q}) = 1 - \inf_{x \in X} \int \sqrt{p(y|x)q(y|x)} d\mu(y) \equiv 1 - f(\mathbf{P},\mathbf{Q}),$$

where

$$f(\mathbf{P},\mathbf{Q}) = \inf_{x \in X} \int \sqrt{p(y|x)q(y|x)} d\mu(y) \equiv \inf_{x \in X} (1, \sqrt{P_x Q_x}) \tag{26}$$

is the minimax fidelity of the classical channels described by these kernels.

B. Quantum operational fidelity

Generalizing Eq. (21), one can define the squared Hellinger distance between quantum operations Φ and Ψ with the respective operational densities $\Phi_\tau, \Psi_\tau \in B(\mathcal{H})$, $\mathcal{H} = \mathfrak{g} \otimes \mathfrak{h}$, as

$$d_H^2(\Phi,\Psi) = \frac{1}{2} \inf_{\Gamma,\Upsilon \in B(\mathcal{H})} \{\|\mathrm{Tr}_\mathfrak{h}(\Gamma - \Upsilon)^\dagger(\Gamma - \Upsilon)\| : \Gamma^\dagger \Gamma = \Phi_\tau, \Upsilon^\dagger \Upsilon = \Psi_\tau\}. \tag{27}$$

The operators $\Gamma, \Upsilon \in B(\mathcal{H})$, such that $\Gamma^\dagger \Gamma = \Phi_\tau$ and $\Upsilon^\dagger \Upsilon = \Psi_\tau$, are naturally thought of as the *purifications* of Φ_τ and Ψ_τ, respectively. This means that we can fix an orthonormal basis $\{|j\rangle\}$ of $\mathcal{H}$, say the product basis $|j\rangle = |i\rangle \otimes |k\rangle \equiv |i,k\rangle$, where $\{|i\rangle\}$ and $\{|k\rangle\}$ are some fixed orthonormal bases of $\mathfrak{g}$ and $\mathfrak{h}$, respectively, and represent any such Γ and Υ as strongly convergent sums

$$\Gamma = \sum_j |j\rangle\langle j|\Gamma \equiv \sum_j |j\rangle(F_j|, \quad \Upsilon = \sum_j |j\rangle\langle j|\Upsilon \equiv \sum_j |j\rangle(V_j|, \tag{28}$$

where the generalized bra-vectors $(F_j|$ define the bounded operators $F_j, V_j : \mathfrak{g} \to \mathfrak{h}$ through

$$\langle k|F_j|i\rangle = (F_j|(|i\rangle \otimes |k\rangle) = \langle j\|\Gamma|i,k\rangle, \quad \langle k|V_j|i\rangle = (V_j|(|i\rangle \otimes |k\rangle) = \langle j\|Y|i,k\rangle.$$

As seen directly from this definition, the mapping $(F|\mapsto F$ is linear: $(aF+bG|\mapsto aF+bG$. Using Eq. (28), we may write

$$\Phi_\tau = \sum_j |F_j)(F_j| \equiv \Gamma^\dagger\Gamma, \quad \Psi_\tau = \sum_j |V_j)(V_j| \equiv Y^\dagger Y, \tag{29}$$

where the sums converge in the strong operator topology. This determines the Kraus decompositions[17] $\Phi(B)=\Sigma_j F_j^\dagger B F_j$, $\Psi(B)=\Sigma_j V_j^\dagger B V_j$ of the maps $\Phi,\Psi:\mathcal{B}\to\mathcal{A}$. Analogously, upon defining the mappings $F,V:\mathfrak{g}\to\mathfrak{h}\otimes\mathcal{H}$ by

$$Fv := \sum_j F_j v \otimes |j\rangle, \quad Vv := \sum_j V_j v \otimes |j\rangle,$$

we can write the maps Φ,Ψ in the Stinespring form[22] as $\Phi(B)=F^\dagger(B\otimes 1_\mathcal{H})F$ and $\Psi(B)=V^\dagger(B\otimes 1_\mathcal{H})V$.

Taking into account the fact that $\|A^\dagger A\|=\sup_{\varrho\in\mathcal{S}(\mathfrak{g})}\varrho(A^\dagger A)$ and defining the positive function

$$c(\cdot\,;\cdot):\mathcal{B}(\mathcal{H}) \times \mathcal{B}_\mathsf{T}(\mathfrak{g}) \to \mathbb{R},$$

$$c(A;\rho) := \tfrac{1}{2}\mathrm{Tr}(A(\rho\otimes 1_\mathfrak{h})A^\dagger),$$

we can rewrite the fidelity distance (27) in the following minimax form:

$$d_H^2(\Phi,\Psi) = \inf_{\Gamma,Y\in\mathcal{B}(\mathcal{H})}\left\{\sup_{\varrho\in\mathcal{S}(\mathfrak{g})} c(\Gamma-Y;\rho):\Gamma^\dagger\Gamma=\Phi_\tau, Y^\dagger Y=\Psi_\tau\right\}. \tag{30}$$

On the other hand, generalizing Eq. (20) to quantum operations, we can define the squared pointwise distance

$$d_H^2(\Phi,\Psi)(\varrho) := \inf_{\Gamma,Y\in\mathcal{B}(\mathcal{H})}\{c(\Gamma-Y;\rho):\Gamma^\dagger\Gamma=\Phi_\tau, Y^\dagger Y=\Psi_\tau\} \tag{31}$$

between Φ and Ψ on the set $\mathcal{S}(\mathfrak{g})$ of all normal states on $\mathcal{A}=\mathcal{B}(\mathfrak{g})$. Just as with the probability kernels in the commutative setting described in the preceding section, $d_H^2(\Phi,\Psi)$ coincides with the supremum of $d_H^2(\Phi,\Psi)(\varrho)$ over all normal states $\varrho\in\mathcal{S}(\mathfrak{g})$ whenever Φ and Ψ are (proportional to) quantum channels:

Theorem 1: *Let* $\Phi,\Psi:\mathcal{B}\to\mathcal{A}$ *be quantum operations with the respective operational densities* $\Phi_\tau,\Psi_\tau\in\mathcal{B}(\mathcal{H})$. *Suppose that for all* $\varrho\in\mathcal{S}(\mathfrak{g})$ *the pairings*

$$(\Phi_\tau,\rho\otimes 1_\mathfrak{h}) \equiv \varrho[\Phi(1_\mathfrak{h})], \quad (\Psi_\tau,\rho\otimes 1_\mathfrak{h}) \equiv \varrho[\Psi(1_\mathfrak{h})] \tag{32}$$

are constant. Then

$$d_H^2(\Phi,\Psi) = \sup_{\varrho\in\mathcal{S}(\mathfrak{g})} d_H^2(\Phi,\Psi)(\varrho). \tag{33}$$

Furthermore, then we have that

$$d_H^2(\Phi,\Psi) + \sqrt{\|\Phi\|\,\|\Psi\|}f(\Phi,\Psi) = \tfrac{1}{2}(\|\Phi\| + \|\Psi\|), \tag{34}$$

where

$$f(\Phi,\Psi) = \inf_{\varrho\in\mathcal{S}(\mathfrak{g})} \sup_{Y\in\mathcal{B}(\mathcal{H}):Y^\dagger Y=\Psi_\tau} \frac{\mathrm{Re}\,\mathrm{Tr}[\Phi_\tau^{1/2}Y(\rho\otimes 1_\mathfrak{h})]}{\sqrt{\varrho[\Phi(1_\mathfrak{h})]}\sqrt{\varrho[\Psi(1_\mathfrak{h})]}} \tag{35}$$

is the minimax fidelity *between* Φ *and* Ψ.

Proof: Fix an arbitrary $\varrho\in\mathcal{S}(\mathfrak{g})$. From Eq. (32) it follows that

$$\|\Phi\| = \sup_{\varrho \in \mathcal{S}(\mathfrak{g})} \varrho[\Phi(1_\mathfrak{h})] = \varrho[\Phi(1_\mathfrak{h})],$$

and the same goes for Ψ. Therefore, given any pair $\Gamma, Y \in \mathcal{B}(\mathcal{H})$ such that $\Gamma^\dagger\Gamma = \Phi_\tau$ and $Y^\dagger Y = \Psi_\tau$, we can write

$$c(\Gamma - Y; \rho) = \tfrac{1}{2}\mathrm{Tr}((\Gamma - Y)^\dagger(\Gamma - Y)(\rho \otimes 1_\mathfrak{h})) = \tfrac{1}{2}\mathrm{Tr}((\Phi_\tau + \Psi_\tau)(\rho \otimes 1_\mathfrak{h}) - (\Gamma^\dagger Y + \Gamma Y^\dagger)(\rho \otimes 1_\mathfrak{h}))$$

$$= \tfrac{1}{2}(\|\Phi\| + \|\Psi\| - \mathrm{Tr}[(\Gamma^\dagger Y + \Gamma Y^\dagger)(\rho \otimes 1_\mathfrak{h})]),$$

whence it follows that

$$d_H^2(\Phi, \Psi)(\varrho) = \inf_{\substack{\Gamma:\Gamma^\dagger\Gamma = \Phi_\tau \\ Y:Y^\dagger Y = \Psi_\tau}} c(\Gamma - Y; \rho) = \tfrac{1}{2}\left(\|\Phi\| + \|\Psi\| - \sup_{\substack{\Gamma:\Gamma^\dagger\Gamma = \Phi_\tau \\ Y:Y^\dagger Y = \Psi_\tau}} \mathrm{Tr}[(\Gamma^\dagger Y + \Gamma Y^\dagger)(\rho \otimes 1_\mathfrak{h})]\right).$$

Taking the supremum of both sides over all $\varrho \in \mathcal{S}(\mathfrak{g})$, we obtain

$$\sup_{\varrho \in \mathcal{S}(\mathfrak{g})} d_H^2(\Phi, \Psi)(\varrho) = \tfrac{1}{2}\left(\|\Phi\| + \|\Psi\| - \inf_{\varrho \in \mathcal{S}(\mathfrak{g})} \sup_{\substack{\Gamma:\Gamma^\dagger\Gamma = \Phi_\tau \\ Y:Y^\dagger Y = \Psi_\tau}} \mathrm{Tr}[(\Gamma^\dagger Y + \Gamma Y^\dagger)(\rho \otimes 1_\mathfrak{h})]\right). \qquad (36)$$

On the other hand,

$$d_H^2(\Phi, \Psi) = \inf_{\substack{\Gamma:\Gamma^\dagger\Gamma = \Phi_\tau \\ Y:Y^\dagger Y = \Psi_\tau}} \sup_{\varrho \in \mathcal{S}(\mathfrak{g})} c(\Gamma - Y; \rho) = \tfrac{1}{2} \inf_{\substack{\Gamma:\Gamma^\dagger\Gamma = \Phi_\tau \\ Y:Y^\dagger Y = \Psi_\tau}} \sup_{\varrho \in \mathcal{S}(\mathfrak{g})} (\|\Phi\| + \|\Psi\| - \mathrm{Tr}[(\Gamma^\dagger Y + \Gamma Y^\dagger)(\rho \otimes 1_\mathfrak{h})]),$$

which yields

$$d_H^2(\Phi, \Psi) = \tfrac{1}{2}\left(\|\Phi\| + \|\Psi\| - \sup_{\substack{\Gamma:\Gamma^\dagger\Gamma = \Phi_\tau \\ Y:Y^\dagger Y = \Psi_\tau}} \inf_{\varrho \in \mathcal{S}(\mathfrak{g})} \mathrm{Tr}[(\Gamma^\dagger Y + \Gamma Y^\dagger)(\rho \otimes 1_\mathfrak{h})]\right). \qquad (37)$$

Note that the right-hand sides of Eqs. (36) and (37) differ only in the order of the extrema. Thus, establishing the validity of Eq. (33) amounts to justifying the interchange of the extrema.

According to Lemma 1, the supremum over Γ and Y in Eq. (36) can be evaluated by fixing $\Gamma = \Phi_\tau^{1/2}$ first and then varying only over all $Y \in \mathcal{B}(\mathcal{H})$ such that $Y^\dagger Y = \Psi_\tau$. By the polar decomposition, any such Y has the form $U\Psi_\tau^{1/2}$ for some partial isometry U. Thus we have

$$\sup_{\substack{\Gamma:\Gamma^\dagger\Gamma = \Phi_\tau \\ Y:Y^\dagger Y = \Psi_\tau}} \mathrm{Tr}[(\Gamma^\dagger Y + \Gamma Y^\dagger)(\rho \otimes 1_\mathfrak{h})] = 2 \sup_{Y:Y^\dagger Y = \Psi_\tau} \mathrm{Re}\,\mathrm{Tr}[\Phi_\tau^{1/2} Y(\rho \otimes 1_\mathfrak{h})]$$

$$= 2 \sup_{U} \mathrm{Re}\,\mathrm{Tr}[\Phi_\tau^{1/2} U \Psi_\tau^{1/2}(\rho \otimes 1_\mathfrak{h})], \qquad (38)$$

where the supremum in Eq. (38) is taken over all partial isometries U such that

$$\Psi_\tau^{1/2} U^\dagger U \Psi_\tau^{1/2} = \Psi_\tau.$$

Since the expression being minimized is linear in U and the isometries are the extreme points of the unit ball $\mathsf{B}_1(\mathcal{H}) := \{X \in \mathcal{B}(\mathcal{H}) : \|X\| \leq 1\}$ of all bounded operators,[23] we may instead take the supremum over the entire unit ball:

$$\sup_{Y:Y^\dagger Y = \Psi_\tau} \mathrm{Re}\,\mathrm{Tr}[\Phi_\tau^{1/2} Y(\rho \otimes 1_\mathfrak{h})] = \sup_{X \in \mathsf{B}_1(\mathcal{H})} \mathrm{Re}\,\mathrm{Tr}[\Phi_\tau^{1/2} X \Psi_\tau^{1/2}(\rho \otimes 1_\mathfrak{h})]. \qquad (39)$$

Since the expression being maximized in the right-hand side of Eq. (39) is affine in both X and ρ, and since $\mathsf{B}_1(\mathcal{H})$ and $\mathcal{S}(\mathfrak{g})$ are closed convex subsets of $\mathcal{B}(\mathcal{H})$ and $\mathcal{B}_\mathsf{T}(\mathfrak{g})$, respectively, it follows from standard minimax arguments[24] that we can indeed interchange the extrema to obtain $f_-(\Phi, \Psi) = f_+(\Phi, \Psi)$, where

$$f_-(\Phi, \Psi) := \inf_{\varrho \in \mathcal{S}(\mathfrak{g})} \sup_{X \in \mathsf{B}_1(\mathcal{H})} \mathrm{Re} \, \mathrm{Tr}[\Phi_\tau^{1/2} X \Psi_\tau^{1/2}(\rho \otimes \mathbb{1}_\mathfrak{h})]$$

$$f_+(\Phi, \Psi) := \sup_{X \in \mathsf{B}_1(\mathcal{H})} \inf_{\varrho \in \mathcal{S}(\mathfrak{g})} \mathrm{Re} \, \mathrm{Tr}[\Phi_\tau^{1/2} X \Psi_\tau^{1/2}(\rho \otimes \mathbb{1}_\mathfrak{h})],$$

which proves the claim of Eq. (33). The rest is straightforward. $\square$

As seen immediately from Theorem 1, when Φ and Ψ are quantum channels, then

$$d_H^2(\Phi, \Psi) + f(\Phi, \Psi) = 1,$$

with the minimax fidelity given by

$$f(\Phi, \Psi) = \inf_{\varrho \in \mathcal{S}(\mathfrak{g})} \sup_{Y: Y^\dagger Y = \Psi_\tau} \mathrm{Re} \, \mathrm{Tr}[\Phi_\tau^{1/2} Y(\rho \otimes \mathbb{1}_\mathfrak{h})]. \tag{40}$$

IV. EVALUATING THE FIDELITY DISTANCES

A. Fidelities for quantum states and quantum effects

Consider two normal states ϱ, s on $\mathcal{B} = \mathcal{B}(\mathfrak{h})$ as quantum channels from $\mathcal{B}$ into the Abelian algebra $\mathcal{A} = \mathcal{B}(\mathfrak{g})$ with $\mathfrak{g} \simeq \mathbb{C}$. In this case, the operational densities $\varrho_\tau, \mathsf{s}_\tau$ of ϱ, s coincide with the corresponding density operators $\rho, \sigma : \varrho_\tau = \rho$ and $\mathsf{s}_\tau = \sigma$. The predual maps $\varrho_\mathsf{T}, \mathsf{s}_\mathsf{T} : \mathcal{A}_\mathsf{T} \simeq \mathbb{C} \to \mathcal{B}_\mathsf{T}$ can then be thought of as the *state creation operations*, $\varrho_\mathsf{T}(\lambda) = \lambda \rho$ and $\mathsf{s}_\mathsf{T}(\lambda) = \lambda \sigma$ for $\lambda \in \mathbb{C}$.

In order to compute the minimax fidelity $f(\varrho, \mathsf{s})$, we have to consider all $\chi, \psi \in \mathcal{B}$ that give the decompositions $\rho = \chi^\dagger \chi$ and $\sigma = \psi^\dagger \psi$. Note that we can always write these decompositions as purifications

$$\rho = \sum_j |\chi_j\rangle\langle\chi_j|, \quad \sigma = \sum_j |\psi_j\rangle\langle\psi_j|,$$

where $|\chi_j\rangle := \chi|j\rangle, |\psi_j\rangle := \psi|j\rangle$ with respect to a fixed orthonormal basis $\{|j\rangle\}$ of $\mathfrak{h}$. We then have the minimum quadratic distance

$$d_H^2(\varrho, \mathsf{s}) = \frac{1}{2} \inf_{\substack{\chi \in \mathcal{B}: \chi^\dagger \chi = \rho \\ \psi \in \mathcal{B}: \psi^\dagger \psi = \sigma}} \sup_{\varpi \in \mathcal{S}(\mathfrak{g})} \varpi[(\chi - \psi)^\dagger(\chi - \psi)] \equiv \frac{1}{2} \inf_{\substack{\chi \in \mathcal{B}: \chi^\dagger \chi = \rho \\ \psi \in \mathcal{B}: \psi^\dagger \psi = \sigma}} \mathrm{Tr}[(\chi - \psi)^\dagger(\chi - \psi)],$$

where the last equality is due to the fact that $\dim \mathfrak{g} = 1$. Expanding the product under the trace, we can write

$$d_H^2(\varrho, \mathsf{s}) = \frac{1}{2}\left[\mathrm{Tr}(\rho + \sigma) - \sup_{\chi, \psi \in \mathcal{B}} \{\mathrm{Re} \, \mathrm{Tr}(\chi^\dagger \psi) : \chi^\dagger \chi = \rho, \psi^\dagger \psi = \sigma\}\right] \tag{41}$$

$$= 1 - \sup_{\substack{\chi \in \mathcal{B}: \chi^\dagger \chi = \rho \\ \psi \in \mathcal{B}: \psi^\dagger \psi = \sigma}} \mathrm{Re} \, \mathrm{Tr}(\chi^\dagger \psi) \tag{42}$$

$$\equiv 1 - f(\varrho, \mathsf{s}). \tag{43}$$

According to Lemma 1, the supremum in Eq. (42) is attained at any $\chi \in \mathcal{B}$ satisfying the condition $\chi^\dagger \chi = \rho$, say $\chi = \rho^{1/2}$, and $\psi = \psi_\circ$ satisfying the equation $\psi_\circ \chi^\dagger = (\chi \sigma \chi^\dagger)^{1/2} = \chi \psi_\circ^\dagger$:

$$f(\varrho, \mathsf{s}) = \sup_{\substack{\chi \in \mathcal{B}: \chi^\dagger \chi = \rho \\ \psi \in \mathcal{B}: \psi^\dagger \psi = \sigma}} \mathrm{Re} \, \mathrm{Tr}(\chi^\dagger \psi) = \sup_{\psi \in \mathcal{B}(\mathfrak{h})} \{\mathrm{Re} \, \mathrm{Tr}(\rho^{1/2} \psi) : \psi^\dagger \psi = \sigma\} = \mathrm{Tr} \, \sqrt{\rho \sigma}.$$

Observe that the standard Uhlmann fidelity between the density operators ρ and σ, $F(\rho, \sigma)$ in Eq. (1), can be written as $F(\rho, \sigma) = \|\rho^{1/2} \sigma^{1/2}\|_\mathsf{T} = \mathrm{Tr} \, \sqrt{\rho \sigma}$. Thus the minimax fidelity between two normal

states ϱ and ς on $\mathcal{B}$, or, equivalently, between the state creation operations $\varrho_T, \varsigma_T : \mathbb{C} \to \mathcal{B}_T(\mathfrak{h})$, agrees with the Uhlmann fidelity between the respective density operators ρ and σ of ϱ and ς.

Next we turn to the other extreme case, namely that of the *state annihilation operations* Φ, Ψ with the preduals $\Phi_T(\rho) = (\Phi_\tau, \rho), \Psi_T(\rho) = (\Psi_\tau, \rho)$, corresponding to dim $\mathfrak{h} = 1$. They are completely specified by the *effects*, i.e., the positive operators $\Phi_\tau, \Psi_\tau \in \mathcal{B}(\mathfrak{g})$ satisfying $0 \leqslant \Phi_\tau, \Psi_\tau \leqslant \mathbb{1}_\mathfrak{g}$, which can be purified as in (29), where $\Gamma_j = \langle j | \Gamma, Y_j = \langle j | Y$ are the bra-vectors corresponding to an orthonormal basis $\{|j\rangle\}$ in $\mathfrak{g}$. The squared pointwise minimax distance between the state annihilation operations Φ, Ψ, or, equivalently between the effects Φ_τ, Ψ_τ on the set $\mathcal{S}(\mathfrak{g})$ of normal states $\varrho = \rho^T$ on $\mathcal{B}(\mathfrak{g})$ is given by the minimum

$$d_H^2(\Phi, \Psi)(\varrho) = \tfrac{1}{2} \inf_{\Gamma, Y \in \mathcal{B}(\mathfrak{g})} \{ \mathrm{Tr}[(\Gamma - Y)^\dagger (\Gamma - Y)\rho] : \Gamma^\dagger \Gamma = \Phi_\tau, Y^\gamma Y = \Psi_\tau \}$$

of the quadratic distance between their purifications $\Gamma, Y \in \mathcal{B}(\mathfrak{g})$. The solution of this problem is likewise given by Lemma 1 with $R = \Phi_\tau$ and $S = \rho \Psi_\tau \rho$. Thus the optimum

$$d_H^2(\Phi, \Psi)(\varrho) = \tfrac{1}{2} \mathrm{Tr}[(\Phi_\tau + \Psi_\tau)\rho] - \mathrm{Tr}\sqrt{\Phi_\tau(\rho \Psi_\tau \rho)}$$

is attained at any $\Gamma \in \mathcal{B}$ satisfying the condition $\Gamma^\dagger \Gamma = \Phi_\tau$, say $\Gamma = \Phi_\tau^{1/2}$, and the corresponding $Y = Y_\circ$ satisfying the equation $Y_\circ \rho \Gamma^\dagger = \sqrt{\Gamma \rho \Psi_\tau \rho \Gamma^\dagger} = \Gamma \rho Y_\circ^\dagger$. The maximum of this distance over all states,

$$d_H^2(\Phi, \Psi) = \sup_{\varrho \in \mathcal{S}(\mathfrak{g})} d_H^2(\Phi, \Psi)(\varrho)$$

$$\equiv \sup_{\varrho \in \mathcal{S}(\mathfrak{g})} \left(\tfrac{1}{2} \mathrm{Tr}[(\Phi_\tau + \Psi_\tau)\rho] - \mathrm{Tr}\sqrt{\Phi_\tau(\rho \Psi_\tau \rho)} \right)$$

$$= \sup_{\varrho \in \mathcal{S}(\mathfrak{g})} \inf_{\Gamma, Y \in \mathcal{B}(\mathfrak{g})} \{ Tr[(\Gamma - Y)^\dagger (\Gamma - Y)\rho] : \Gamma^\dagger \Gamma = \Phi_\tau, Y^\dagger Y = \Psi_\tau \},$$

is given by the minimax quadratic distance

$$d_H^2(\Phi, \Psi) = \tfrac{1}{2} \inf_{\Gamma, Y \in \mathcal{B}(\mathfrak{g})} \{ \|\Gamma - Y\|^2 : \Gamma^\dagger \Gamma = \Phi_\tau, Y^\dagger Y = \Psi_\tau \},$$

interchange of the extrema following from standard minimax arguments,[24] and the fact that all Γ, Y satisfying, respectively, $\Gamma^\dagger \Gamma = \Phi_\tau$ and $Y^\dagger Y = \Psi_\tau$ are contained in the unit ball of $\mathcal{B}(\mathfrak{g})$.

B. Semiclassical fidelity

It is straightforward to extend the formalism of Sec. III A involving the commutative Hellinger distance between two positive kernels to the case of mappings from a set X into positive trace-class operators on the Hilbert space $\mathfrak{h}$, i.e., $\boldsymbol{\rho} : x \in X \mapsto \rho(x) \in \mathcal{B}_T(\mathfrak{h})$ and $\boldsymbol{\sigma} : x \in X \mapsto \sigma(x) \in \mathcal{B}_T(\mathfrak{h})$ with $\rho(x), \sigma(x) \geqslant 0$ for all $x \in X$. We thus have the pointwise Hellinger distance

$$d_H^2(\rho(x), \sigma(x)) = \left(\mathbb{1}, \tfrac{1}{2}[\rho(x) + \sigma(x)] \right) - \sqrt{(\mathbb{1}, \rho(x))(\mathbb{1}, \sigma(x))} f(\rho(x), \sigma(x))$$

in terms of the trace pairing $(B, \rho) = \mathrm{Tr}(B\tilde{\rho})$ of $B \in \mathcal{B} = \mathcal{B}(\mathfrak{h})$ and $\rho \in \mathcal{B}_T = \mathcal{B}_T(\mathfrak{h})$, where

$$f(\rho(x), \sigma(x)) = \frac{(\mathbb{1}, \sqrt{\rho(x)\sigma(x)})}{\sqrt{(\mathbb{1}, \rho(x))(\mathbb{1}, \sigma(x))}} = \frac{\mathrm{Tr}\sqrt{\rho(x)\sigma(x)}}{\sqrt{\mathrm{Tr}\,\rho(x)\mathrm{Tr}\,\sigma(x)}}.$$

The semi-classical operational distance between $\boldsymbol{\rho} = \rho(\cdot)$ and $\boldsymbol{\sigma} = \sigma(\cdot)$ can then be defined as

$$d_H^2(\boldsymbol{\rho}, \boldsymbol{\sigma}) = \sup_{x \in X} d_H^2(\rho(x), \sigma(x)) \equiv \|d_H^2(\rho(\cdot), \sigma(\cdot))\|. \tag{44}$$

When $\mathrm{Tr}\,\rho(x) = 1 = \mathrm{Tr}\,\sigma(x)$ for all $x \in X$, i.e., when $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$ are *classical-to-quantum*, c-q (or *semiclassical*), channels, Eq. (44) can be written as $d_H^2(\boldsymbol{\rho}, \boldsymbol{\sigma}) = 1 - f(\boldsymbol{\rho}, \boldsymbol{\sigma})$, where

$$f(\boldsymbol{\rho},\boldsymbol{\sigma}) = \inf_{x\in X} \mathrm{Tr}[\rho(x)^{1/2}\sigma(x)\rho(x)^{1/2}]^{1/2} = \inf_{x\in X} \mathrm{Tr}\sqrt{\rho(x)\sigma(x)} \equiv \inf_{x\in X} F(\rho(x),\sigma(x))$$

is the minimax fidelity of $\boldsymbol{\sigma}$ relative to $\boldsymbol{\rho}$.

C. Semiquantum fidelity

Next we consider the opposite of semiclassical operations—namely, the *semiquantum operations* which correspond to quantum measurements as quantum-to-classical (q-c) channels. Such operations are given as

$$\Phi(b) = \int_Y b(y)\Phi_\tau(y)\mathrm{d}\mu(y) \equiv (b,\Phi_\tau)$$

on the algebra $\mathcal{B}=\mathcal{C}(Y)$ of continuous bounded functions $b:Y\to\mathbb{C}$, where $(Y,\mathcal{B}_Y,\mu)$ is a measure space, by specifying the positive operator-valued Bochner μ-integrable functions $\Phi_\tau:Y\to\mathcal{A}$ $=\mathcal{B}(\mathfrak{g})$. If

$$\Phi(1) = (1,\Phi_\tau) = \mathbb{1}_\mathfrak{g},$$

the predual maps $\mathcal{A} \ni \rho \mapsto \Phi_\mathsf{T}(\rho)(\cdot) \in \mathcal{C}_\mathsf{T}(Y)$,

$$\Phi_\mathsf{T}(\rho)(y) := (\Phi_\tau(y),\rho) \equiv \varrho[\Phi_\tau(y)],$$

define for each input quantum state $\varrho \in \mathcal{S}(\mathfrak{g})$ a classical probability density on $(Y,\mathcal{B}_Y,\mu)$, that is, they describe quantum measurements by the positive operator-valued measures (POVMs) $M(\mathrm{d}y)=\Phi_\tau(y)\mathrm{d}\mu(y)$.

In order to avoid technicalities in defining the semi-quantum fidelity distance between two q-c channels $\Phi,\Psi:\mathcal{B}\to\mathcal{A}$, we shall assume that $\Phi_\tau(y),\Psi_\tau(y)$ are weakly continuous bounded functions on Y. Then the squared distance $d_H^2(\Phi,\Psi)$ can be written as

$$d_H^2(\Phi,\Psi) = \inf_{\Gamma,Y:\Gamma^\dagger\Gamma=\Phi_\tau,Y^\dagger Y=\Psi_\tau} \left\| \int (\Gamma(y)-Y(y))^\dagger(\Gamma(y)-Y(y))\mathrm{d}\mu(y) \right\|, \tag{45}$$

where the decompositions $\Gamma^\dagger\Gamma=\Phi_\tau$ and $Y^\dagger Y=\Psi_\tau$ are understood in the pointwise sense as

$$\Phi_\tau(y) = \Gamma(y)^\dagger\Gamma(y), \quad \Psi_\tau(y) = Y(y)^\dagger Y(y), \quad \forall\, y \in Y.$$

The infimum in Eq. (45) is achieved at any $\Gamma \in \mathcal{A}\otimes\mathcal{C}_\mathsf{T}(Y)$ satisfying the condition $\Gamma^\dagger\Gamma=\Phi_\tau$, say $\Gamma(y)=\Phi_\tau(y)^{1/2}$, and the corresponding $Y=Y_\mathrm{o}$ satisfying the equation

$$Y_\mathrm{o}(y)\rho\Gamma(y)^\dagger = [\Gamma(y)\rho\Psi_\tau(y)\rho\Gamma(y)^\dagger]^{1/2} = \Gamma(y)\rho Y_\mathrm{o}(y)^\dagger.$$

The maximum of this minimal distance over all states,

$$d_H^2(\Phi,\Psi) = \sup_{\varrho\in\mathcal{S}(g)} \int \left(\frac{1}{2}\mathrm{Tr}[(\Phi_\tau(y)+\Psi_\tau(y))\rho] - \mathrm{Tr}\sqrt{\Phi_\tau(y)(\rho\Psi_\tau(y)\rho)} \right)\mathrm{d}\mu(y),$$

is equal to $d_H^2(\Phi,\Psi)=1-f(\Phi,\Psi)$ in the measurement operation case $\Phi(1)=\mathbb{1}_\mathfrak{g}=\Psi(1)$, where

$$f(\Phi,\Psi) = \inf_{\varrho\in\mathcal{S}(\mathfrak{g})} \int \mathrm{Tr}\sqrt{\Phi_\tau(y)(\rho\Psi_\tau(y)\rho)}\mathrm{d}\mu(y). \tag{46}$$

D. Operational fidelity formula

Now we can easily evaluate the minimax formula (30) for the fidelity of two general quantum operations $\Phi,\Psi:\mathcal{B}\to\mathcal{A}$, $\mathcal{B}=\mathcal{B}(\mathfrak{h})$, $\mathcal{A}=\mathcal{B}(\mathfrak{g})$. The solution of this problem is also given by Lemma 1 with $R=\Phi_\tau$ and $S=(\rho\otimes\mathbb{1}_\mathfrak{h})\Psi_\tau(\rho\otimes\mathbb{1}_\mathfrak{h})$. For a given $\varrho \in \mathcal{S}(\mathfrak{g})$, the supremum in

$$d_H^2(\Phi,\Psi)(\varrho) = \tfrac{1}{2}(\text{Tr}[(\Phi_\tau + \Psi_\tau)(\rho \otimes 1_\mathfrak{h})] - 2 \sup_{\Gamma,Y \in \mathcal{B}(\mathcal{H})} \{\text{Re Tr}[\Gamma^\dagger Y(\rho \otimes 1_\mathfrak{h})] : \Gamma^\dagger \Gamma = \Phi_\tau, Y^\dagger Y = \Psi_\tau\})$$

is equal to $\text{Tr}\sqrt{\Phi_\tau[(\rho \otimes 1_\mathfrak{h})\Psi_\tau(\rho \otimes 1_\mathfrak{h})]}$, and is achieved at any $\Gamma \in \mathcal{B}(\mathcal{H})$ satisfying the condition $\Gamma^\dagger\Gamma = \Phi_\tau$, say $\Gamma = \Phi_\tau^{1/2}$, and the corresponding $Y = Y_o$ satisfying the equation

$$Y_o(\rho \otimes 1_\mathfrak{h})\Gamma^\dagger = [\Gamma(\rho \otimes 1_\mathfrak{h})\Psi_\tau(\rho \otimes 1_\mathfrak{h})\Gamma^\dagger]^{1/2} = \Gamma(\rho \otimes 1_\mathfrak{h})Y_o^\dagger.$$

When Φ, Ψ are quantum channels, or, equivalently, when the preduals Φ_T, Ψ_T are trace preserving, Theorem 1 says that the maximum of this distance over all states,

$$d_H^2(\Phi,\Psi) = \sup_{\varrho \in \mathcal{S}(\mathfrak{g})} \text{Tr}\left(\tfrac{1}{2}(\Phi_\tau + \Psi_\tau)(\rho \otimes 1_\mathfrak{h}) - \sqrt{\Phi_\tau[(\rho \otimes 1_\mathfrak{h})\Psi_\tau(\rho \otimes 1_\mathfrak{h})]}\right), \tag{47}$$

can be written as $d_H^2(\Phi,\Psi) = 1 - f(\Phi,\Psi)$, where

$$f(\Phi,\Psi) = \inf_{\varrho \in \mathcal{S}(\mathfrak{g})} \text{Tr}\sqrt{\Phi_\tau[(\rho \otimes 1_\mathfrak{h})\Psi_\tau(\rho \otimes 1_\mathfrak{h})]} \tag{48}$$

is the minimax fidelity between Φ and Ψ.

E. Operational fidelity in terms of Kraus and Stinespring decompositions

Consider, as before, two quantum channels $\Phi, \Psi : \mathcal{B} \to \mathcal{A}$, where $\mathcal{B} = \mathcal{B}(\mathfrak{h})$ and $\mathcal{A} = \mathcal{B}(\mathfrak{g})$. Given the minimax fidelity

$$f(\Phi,\Psi) = \inf_{\substack{\varrho \in \mathcal{S}(\mathfrak{g})}} \sup_{\substack{\Gamma:\Gamma^\dagger\Gamma=\Phi_\tau \\ Y:Y^\dagger Y=\Psi_\tau}} \text{Re Tr}[\Gamma^\dagger Y(\rho \otimes 1_\mathfrak{h})] = \inf_{\substack{\varrho \in \mathcal{S}(\mathfrak{g})}} \sup_{\substack{\Gamma:\Gamma^\dagger\Gamma=\Phi_\tau \\ Y:Y^\dagger Y=\Psi_\tau}} |\text{Tr}[\Gamma^\dagger Y(\rho \otimes 1_\mathfrak{h})]|$$

between Φ and Ψ, where the second equality follows from Lemma 2, the supremum over all Γ and Y satisfying, respectively, $\Gamma^\dagger\Gamma = \Phi_\tau$ and $Y^\dagger Y = \Psi_\tau$ can be replaced with the supremum over all Kraus decompositions of Φ and Ψ, i.e., over all collections $\{F_j\}$, $\{V_j\}$ of bounded operators $\mathfrak{g} \to \mathfrak{h}$, determined from Φ_τ, Ψ_τ via Eqs. (29) and (28):

$$f(\Phi,\Psi) = \inf_{\varrho \in \mathcal{S}(\mathfrak{g})} \sup_{\{F_j\},\{V_j\}} \left|\sum_j \varrho(F_j^\dagger V_j)\right|. \tag{49}$$

Just as in the proof of Theorem 1, we may restrict ourselves only to those Γ, Y that can be written as $\Gamma = U\Phi_\tau^{1/2}, Y = V\Psi_\tau^{1/2}$ for some isometries U, V. Thus, if we write $\Phi_\tau^{1/2}$ and $\Psi_\tau^{1/2}$ in the form of Eq. (28) as

$$\Phi_\tau^{1/2} = \sum_j |j\rangle\langle\hat{F}_j|, \quad \Psi_\tau^{1/2} = \sum_j |j\rangle\langle\hat{V}_j|,$$

then it follows that, given isometries U, we can write

$$\Gamma = U\Phi_\tau^{1/2} = \sum_j |j\rangle\left(\sum_\ell U_{j\ell}\hat{F}_\ell\right| \equiv \sum_j |j\rangle\langle\hat{F}_j(U)|,$$

and similarly for $Y = V\Psi_\tau^{1/2}$. Thus

$$f(\Phi,\Psi) = \inf_{\varrho \in \mathcal{S}(\mathfrak{g})} \sup_{U,V} \left|\sum_j \varrho[\hat{F}_j(U)^\dagger \hat{V}_j(V)]\right| = \inf_{\varrho \in \mathcal{S}(\mathfrak{g})} \sup_U \left|\sum_j \varrho[\hat{F}_j(U)^\dagger \hat{V}_j]\right|.$$

Turning now to the infimum over all normal states ϱ on $\mathcal{A} \equiv \mathcal{B}(\mathfrak{g})$, we may equivalently consider all pairs $\{\varphi, \mathcal{K}\}$, where φ is a normal $*$-representation of $\mathcal{A}$ on a Hilbert space $\mathcal{K}$:

$$f(\Phi,\Psi) := \inf_{\{\varphi,\mathcal{K}\}; v \in \mathcal{K}, \|u\|=1} \sup_U \left| \sum_j \langle v | \varphi[\hat{F}_j(U)^\dagger \hat{V}_j] | v \rangle \right|.$$

Since all normal $*$-representations of the full operator algebra $\mathcal{B}(\mathfrak{g})$ are unitarily equivalent to an amplification $B \mapsto B \otimes 1_{\mathfrak{k}}$ for some Hilbert space $\mathfrak{k}$, we can write

$$f(\Phi,\Psi) := \inf_{v \in \mathfrak{g} \otimes \mathfrak{k}; \|u\|=1} \sup_U \left| \sum_j \langle v | \hat{F}_j(U)^\dagger \hat{V}_j \otimes 1_{\mathfrak{k}} | v \rangle \right|. \tag{50}$$

Introducing the vectors $|v,\Phi\rangle, |v,\Psi\rangle \in \mathfrak{g} \otimes \mathfrak{k} \otimes \mathcal{H}$, defined by

$$|v,\Phi\rangle := \sum_j (\hat{F}_j \otimes 1_{\mathfrak{k}}) v \otimes |j\rangle, \quad |v,\Psi\rangle := \sum_j (\hat{V}_j \otimes 1_{\mathfrak{k}}) v \otimes |j\rangle,$$

we obtain yet another form of the minimax fidelity:

$$f(\Phi,\Psi) = \inf_{v \in \mathfrak{g} \otimes \mathfrak{k}} \sup_U |\langle v,\Phi | 1_{\mathfrak{g} \otimes \mathfrak{k}} \otimes U | v,\Psi\rangle|. \tag{51}$$

For a fixed $v \in \mathfrak{g} \otimes \mathfrak{k}$, taking the supremum over U is tantamount to taking the supremum of $|\langle \chi | \xi \rangle|$ over all pairs of unit vectors $\chi, \xi \in \mathfrak{g} \otimes \mathfrak{k} \otimes \mathcal{H}$ such that

$$\mathrm{Tr}_{\mathcal{H}} |\chi\rangle\langle\chi| = \sum_j (\widetilde{F_j} \otimes 1_{\mathfrak{k}}) |v\rangle\langle v| (\widetilde{F_j} \otimes 1_{\mathfrak{k}})^\dagger \equiv \Phi_{\mathsf{T}} \otimes \mathrm{id}(|v\rangle\langle v|),$$

$$\mathrm{Tr}_{\mathcal{H}} |\xi\rangle\langle\xi| = \sum_j (\widetilde{V_j} \otimes 1_{\mathfrak{k}}) |v\rangle\langle v| (\widetilde{V_j} \otimes 1_{\mathfrak{k}})^\dagger \equiv \Psi_{\mathsf{T}} \otimes \mathrm{id}(|v\rangle\langle v|),$$

which, in conjunction with the standard results on the Uhlmann fidelity (1) between density operators,[6,7] finally yields

$$f(\Phi,\Psi) = \inf_{v \in \mathfrak{g} \otimes \mathfrak{k}; \|u\|=1} F(\Phi_{\mathsf{T}} \otimes \mathrm{id}(|v\rangle\langle v|), \Psi_{\mathsf{T}} \otimes \mathrm{id}(|v\rangle\langle v|)) = \inf_{\varrho \in \mathcal{S}(\mathfrak{g} \otimes \mathfrak{k})} F(\Phi_{\mathsf{T}} \otimes \mathrm{id}(\rho), \Psi_{\mathsf{T}} \otimes \mathrm{id}(\rho)).$$

Note that we may always take $\mathfrak{k}$ isomorphic to $\mathfrak{g}$:

$$f(\Phi,\Psi) = \inf_{v \in \mathfrak{g} \otimes \mathfrak{g}, \|u\|=1} F(\Phi_{\mathsf{T}} \otimes \mathrm{id}(|v\rangle\langle v|), \Psi_{\mathsf{T}} \otimes \mathrm{id}(|v\rangle\langle v|)). \tag{52}$$

Given some Kraus decompositions $\{F_j\}, \{V_j\}$ of Φ and Ψ, respectively, we may define the operators

$$F\xi := \sum_j F_j \xi \otimes |j\rangle, \quad V\xi := \sum_j V_j \xi \otimes |j\rangle$$

from $\mathfrak{g}$ into $\mathfrak{h} \otimes \mathcal{H}$ and express Φ and Ψ in the Stinespring form $\Phi(B) = F^\dagger(B \otimes 1_{\mathcal{H}})F$, $\Psi(B) = V^\dagger(B \otimes 1_{\mathcal{H}})V$ (cf. Sec. III B). Then we may rewrite Eq. (49) as

$$f(\Phi,\Psi) = \inf_{\varrho \in \mathcal{S}(\mathfrak{g})} \sup_{F,V} |\mathrm{Tr}(F\rho V^\dagger)|,$$

where the supremum is over all $F, V: \mathfrak{g} \to \mathfrak{h} \otimes \mathcal{H}$ giving the Stinespring decompositions of Φ and Ψ, respectively. We may, as before, fix F and V, say, by considering the "canonical" Kraus decompositions $\{\hat{F}_j\}, \{\hat{V}_j\}$, and instead take the supremum over all unitaries $U \in \mathsf{U}(\mathcal{H})$:

$$f(\Phi,\Psi) = \inf_{\varrho \in \mathcal{S}(\mathfrak{g})} \sup_U |\mathrm{Tr}[(1_{\mathfrak{h}} \otimes U)F\rho V^\dagger]| = \inf_{\varrho \in \mathcal{S}(\mathfrak{g})} \sup_U |\mathrm{Tr}[U \, \mathrm{Tr}_{\mathfrak{h}}(F\rho V^\dagger)]|,$$

which yields another useful formula

$$f(\Phi,\Psi) = \inf_{\varrho \in \mathcal{S}(\mathfrak{g})} \left\| \mathrm{Tr}_{\mathfrak{h}}(F\rho V^{\dagger}) \right\|_{\mathsf{T}} \tag{53}$$

for the minimax fidelity between the channels Φ, Ψ. It is, in fact, not hard to show that the right-hand side of Eq. (53) does not depend on the particular choice of the Stinespring operators F, V, as long as we agree to dilate the input Hilbert space $\mathfrak{h}$ by the "canonical" auxiliary Hilbert space $\mathcal{H} = \mathfrak{g} \otimes \mathfrak{h}$.

We note that the constructions of this section are valid more generally for channels given in terms of the continual Kraus decompositions

$$\Phi(B) = \int_Z F(z)^{\dagger} B F(z) \mathrm{d}\mu(z), \quad \Psi(B) = \int_Z V(z)^{\dagger} B V(z) \mathrm{d}\nu(z),$$

provided that the measures μ and ν are equivalent, i.e., absolutely continuous with respect to each other. Then Eq. (49) is a special instance of the more general expression

$$f(\Phi,\Psi) = \inf_{\varrho \in \mathcal{S}(\mathfrak{g})} \sup_{\{F(z)\},\{V(z)\}} \left| \varrho\left(\int_Z \sqrt{\mathrm{d}\nu/\mathrm{d}\mu} F(z)^{\dagger} V(z) \mathrm{d}\mu(z) \right) \right|,$$

where $\mathrm{d}\nu/\mathrm{d}\mu$ is the Radon-Nikodym derivative of ν with respect to μ, for the case when both μ and ν are counting measures, $\mathrm{d}\mu = \mathrm{d}\nu = 1$, on a finite or countably infinite set.

V. PROPERTIES OF THE OPERATIONAL FIDELITY

In this section we establish several key properties of the minimax fidelity between quantum operations. These properties follow almost immediately from the corresponding properties enjoyed by the fidelity (1) on density operators:

(F.1) F is symmetric, $F(\rho,\sigma) = F(\sigma,\rho)$, bounded between 0 and 1, and $F(\rho,\sigma) = 1$ if and only if $\rho = \sigma$.

(F.2) F is jointly concave over all pairs of density operators.

(F.3) F is unitarily invariant, i.e., $F(\rho,\sigma) = F(U\rho U^{\dagger}, U\sigma U^{\dagger})$ for any unitary U.

(F.4) F is monotone with respect to quantum channels: $F(\Phi_{\mathsf{T}}(\rho), \Phi_{\mathsf{T}}(\sigma)) \geqslant F(\rho,\sigma)$ for every quantum channel Φ.

(F.5) The *Bures distance* $d_B(\cdot,\cdot) := \sqrt{1 - F(\cdot,\cdot)}$ is topologically equivalent to the trace-norm half-distance $D(\cdot,\cdot)$:

$$2^{-1/2} D(\rho,\sigma) \leqslant d_B(\rho,\sigma) \leqslant \sqrt{D(\rho,\sigma)}$$

[cf. Eq. (3)]. Property (F.2), in fact, follows from *strong concavity* of F,[2] i.e.,

$$F\left(\sum_i p_i \rho_i, \sum_i q_i \sigma_i \right) \geqslant \sum_i \sqrt{p_i q_i} F(\rho_i, \sigma_i) \tag{54}$$

for all $0 \leqslant p_i, q_i \leqslant 1$ such that $\Sigma_i p_i = 1 = \Sigma_i q_i$.

Using Eq. (52), we can immediately obtain for the minimax fidelity $f(\cdot,\cdot)$ on pairs of quantum channels the following analogs of properties (F.1)–(F.4) of the fidelity $F(\cdot,\cdot)$ on pairs of density operators:

(f.1) f is symmetric, bounded between 0 and 1, and $f(\Phi,\Psi) = 1$ if and only if $\Phi = \Psi$.

(f.2) f is jointly concave over all pairs of channels.

(f.3) f is invariant under both left and right composition with unitarily implemented channels, i.e.,

$$f(\Theta_U \circ \Phi, \Theta_U \circ \Psi) = f(\Phi,\Psi)$$

and

$$f(\Phi \circ \Theta_V, \Psi \circ \Theta_V) = f(\Phi, \Psi)$$

for any two channels $\Phi, \Psi : \mathcal{B}(\mathfrak{h}) \to \mathcal{B}(\mathfrak{g})$ and any two unitaries $U \in \mathsf{U}(\mathfrak{g})$, $V \in \mathsf{U}(\mathfrak{h})$, where $\Theta_U(B) := U^\dagger B U$, and Θ_V is defined analogously.

(f.4) f is monotone with respect to both left and right composition with quantum channels, i.e., $f(\Phi \circ \Phi_1, \Psi \circ \Phi_1) \geq f(\Phi, \Psi)$ and $f(\Phi_2 \circ \Phi, \Phi_2 \circ \Psi) \geq f(\Phi, \Psi)$ for any two channels $\Phi, \Psi : \mathcal{B} \to \mathcal{A}$, all channels Φ_1 into $\mathcal{B}$, and all channels Φ_2 on $\mathcal{A}$. Just as in the case of the fidelity between density operators, the minimax fidelity f possesses the strong concavity property

$$f\left(\sum_i p_i \Phi_i, \sum_i q_i \Psi_i\right) \geq \sum_i \sqrt{p_i q_i} f(\Phi_i, \Psi_i). \tag{55}$$

On the other hand, deriving for the minimax fidelity f an analog of property (F.5) of the Uhlmann fidelity F requires a bit more work. To this end, let us consider two channels $\Phi, \Psi : \mathcal{B} \to \mathcal{A}, \mathcal{B} = \mathcal{B}(\mathfrak{h}), \mathcal{A} = \mathcal{B}(\mathfrak{g})$. Suppose first that $\mathfrak{g}$ is infinite dimensional and separable. Then $\mathfrak{g} \simeq \ell^2$, and we can rewrite Eq. (55) as

$$f(\Phi, \Psi) = \inf_{v \in \mathfrak{g} \otimes \ell^2; \|v\|=1} F(\Phi_\mathsf{T} \otimes \mathrm{id}(|v\rangle\langle v|), \Psi_\mathsf{T} \otimes \mathrm{id}(|v\rangle\langle v|)).$$

The space ℓ^2 contains, as a dense subset, the pre-Hilbert space ℓ_0^2 of all infinite sequences of complex numbers with all but finitely many components equal to zero. Using this fact and the continuity property (F.5) of the fidelity F, we obtain

$$f(\Phi, \Psi) = \inf_{v \in \mathfrak{g} \otimes \ell_0^2; \|v\|=1} F(\Phi_\mathsf{T} \otimes \mathrm{id}(|v\rangle\langle v|), \Psi_\mathsf{T} \otimes \mathrm{id}(|v\rangle\langle v|)).$$

Using this expression in conjunction with Eq. (3), we get the bounds

$$f(\Phi, \Psi) \geq 1 - \sup_{v \in \mathfrak{g} \otimes \ell_0^2; \|v\|=1} D(\Phi_\mathsf{T} \otimes \mathrm{id}(|v\rangle\langle v|), \Psi_\mathsf{T} \otimes \mathrm{id}(|v\rangle\langle v|)), \tag{56}$$

$$f^2(\Phi, \Psi) \leq 1 - \sup_{v \in \mathfrak{g} \otimes \ell_0^2; \|v\|=1} D^2(\Phi_\mathsf{T} \otimes \mathrm{id}(|v\rangle\langle v|) - \Psi_\mathsf{T} \otimes \mathrm{id}(|v\rangle\langle v|)). \tag{57}$$

Now, for any completely bounded map $\Lambda : \mathcal{B}(\mathfrak{h}) \to \mathcal{B}(\mathfrak{g})$, the image of the set $\{|v\rangle\langle v| : v \in \mathfrak{g} \otimes \ell_0^2, \|v\|=1\}$ under the predual map $\Lambda_\mathsf{T} \otimes \mathrm{id} : \mathcal{B}_\mathsf{T}(\mathfrak{g} \otimes \ell_0^2) \to \mathcal{B}_\mathsf{T}(\mathfrak{h} \otimes \ell_0^2)$ is contained in the trace-norm closure of the linear span of $\{|\xi\rangle\langle\xi| : \xi \in \mathfrak{h} \otimes \ell_0^2, \|\xi\|=1\}$, which is dual to the tensor product $\mathcal{B}(\mathfrak{h}) \otimes \mathcal{K}(\ell^2)$, where $\mathcal{K}(\ell^2)$ is the space of compact operators on ℓ^2. Thus, by duality we have

$$\sup_{v \in \mathfrak{g} \otimes \ell_0^2; \|v\|=1} D(\Phi_\mathsf{T} \otimes \mathrm{id}(|v\rangle\langle v|), \Psi_\mathsf{T} \otimes \mathrm{id}(|v\rangle\langle v|)) = \tfrac{1}{2}\|(\Phi - \Psi) \otimes \mathrm{id}_{\mathcal{K}(\ell^2)}\| \equiv \mathcal{D}(\Phi, \Psi),$$

where $\mathcal{D}(\Phi, \Psi)$ denotes the CB-norm half-distance $\tfrac{1}{2}\|\Phi - \Psi\|_{\mathrm{CB}}$, and the last equality follows from the formula $\|\Lambda\|_{\mathrm{CB}} = \|\Lambda \otimes \mathrm{id}_{\mathcal{K}(\ell^2)}\|$ for any completely bounded map Λ.

On the other hand, when $\dim \mathfrak{g} = m < \infty$, we can use the fact[20] that, for any completely bounded map Λ into $\mathcal{B}(\mathfrak{g})$,

$$\|\Lambda\|_{\mathrm{CB}} = \|\Lambda \otimes \mathrm{id}_{\mathcal{M}_m}\| = \|\Lambda_\mathsf{T} \otimes \mathrm{id}_{\mathcal{M}_m}\|_\mathsf{T},$$

where $\mathcal{M}_m$ denotes the algebra of $m \times m$ complex matrices, it follows that

$$\sup_{v \in \mathfrak{g} \otimes \mathfrak{g}; \|v\|=1} D(\Phi_\mathsf{T} \otimes \mathbb{1}(|v\rangle\langle v|), \Psi_\mathsf{T} \otimes \mathbb{1}(|v\rangle\langle v|)) = \mathcal{D}(\Phi, \Psi).$$

In either case, we immediately derive the inequality

$$1 - \mathcal{D}(\Phi, \Psi) \leqslant f(\Phi, \Psi) \leqslant \sqrt{1 - \mathcal{D}^2(\Phi, \Psi)}, \tag{58}$$

which, when expressed in terms of the Hellinger distance $d_H(\cdot, \cdot) := \sqrt{1 - f(\cdot, \cdot)}$ as

$$2^{-1/2} \mathcal{D}(\Phi, \Psi) \leqslant d_H(\Phi, \Psi) \leqslant \sqrt{\mathcal{D}(\Phi, \Psi)}, \tag{59}$$

yields the desired property

(f.5) the Hellinger distance $d_H(\cdot, \cdot) := \sqrt{1 - f(\cdot, \cdot)}$ is topologically equivalent to the CB-norm distance [cf. Eq. (59)].

This completes our survey of the basic properties of the minimax fidelity f.

VI. SOME EXAMPLES AND APPLICATIONS

The expressions for the minimax fidelity, derived in Sec. IV for different kinds of quantum operations encountered in quantum information theory, share the common feature of being set up as variational problems, namely, as minimizations of a concave functional over a convex set. This feature of the minimax fidelity renders the problem of computing it amenable to robust numerical methods (see Ref. 3 for detailed discussion of numerical optimization methods for the calculation of fidelitylike measures in quantum information theory). However, there are instances in which the minimax fidelity between two quantum channels can be written down in a more explicit form. In this section we sketch some examples of such instances.

Before we proceed, we would like to remind the reader of the assumption we made in Sec. II B, namely that all the channels we deal with are completely majorized by the trace in the sense of Ref. 14. This assumption, while allowing us to circumvent certain technicalities involving unbounded operators, is somewhat restrictive, as one can easily find examples of quantum channels between infinite-dimensional algebras (e.g., unitarily or isometrically implemented channels; see Ref. 21 for details) that do not satisfy this condition of complete majorization. However, owing to the CB-continuity of the minimax fidelity (cf. Sec. V), we may always regard such channels as CB-limits of sequences of channels with finite-dimensional output algebras. Thus, given a channel $\Phi : \mathcal{B} \to \mathcal{A}, \mathcal{B} = \mathcal{B}(\mathfrak{h}), \mathcal{A} = \mathcal{B}(\mathfrak{g})$ with $\dim \mathfrak{g} = \infty$, we consider a sequence $\{P_n\}$ of finite-dimensional projections such that $P_n \to \mathbb{1}_g$ strongly, and the corresponding sequence $\{\Phi_n\}$ of quantum operations $\Phi_n(B) := P_n \Phi(B) P_n$, so that $\Phi_n(B) \to \Phi(B)$ uniformly as $n \to \infty$ for each $B \in \mathcal{B}$, and each Φ_n is a channel from $\mathcal{B}$ into $P_n \mathcal{A} P_n$, with $\lim_{n \to \infty} \|\Phi - \Phi_n\|_{CB} = 0$.

With this in mind, in the examples below we shall not worry about the issue of bounded versus unbounded operational densities.

A. Unitary maps

In the case of channels Θ_U, Θ_V implemented by the unitaries $U, V : \mathfrak{h} \to \mathfrak{h}$, i.e., $\Theta_U(B) = U^\dagger B U$ and $\Theta_V(B) = V^\dagger B V$, the minimax fidelity $f(\Phi, \Psi)$ is easily evaluated using Eq. (49):

$$f(\Theta_U, \Theta_V) = \inf_{\varrho \in \mathcal{S}(\mathfrak{g})} |\varrho(W)|,$$

where we have defined $W := U^\dagger V$. Let $\mathrm{Sp}(W)$ denote the spectrum of W, which is a closed compact subset of the unit circle $\mathbb{T}$ in the complex plane, and let $E^W(\mathrm{d}z)$ denote the corresponding spectral measure of W. Then we can write

$$f(\Theta_U, \Theta_V) = \inf_{\varrho \in \mathcal{S}(\mathfrak{g})} \left| \int_{\mathrm{Sp}(W)} z M^{W, \varrho}(\mathrm{d}z) \right|,$$

where $M^{W, \varrho}(\mathrm{d}z)$ is the probability measure $\varrho[E^W(\mathrm{d}z)] \equiv (E^W(\mathrm{d}z), \rho)$. Thus

$$f(\Theta_U, \Theta_V) = \mathrm{dist}(0, \overline{\mathrm{co}\,\mathrm{Sp}(W)}), \tag{60}$$

where $\overline{\mathrm{co}\,\mathrm{Sp}(W)}$ denotes the closed convex hull of $\mathrm{Sp}(W)$, and $\underline{\mathrm{dist}(z,S) := \inf\{|z-z'| : z' \in S\}}$ for any $z \in \mathbb{C}$ and $S \subset \mathbb{C}$. Clearly, $f(\Theta_U, \Theta_V) = 1$ if and only if $\overline{\mathrm{co}\,\mathrm{Sp}\,W} \subset \mathbb{T}$, i.e., if and only if $W = \lambda 1_{\mathfrak{h}}$ with $|\lambda| = 1$, which is equivalent to $\Theta_U = \Theta_V$.

When $\dim \mathfrak{h} < \infty$, $\mathrm{Sp}(W)$ is a finite subset of $\mathbb{T}$, so that $\overline{\mathrm{co}\,\mathrm{Sp}(W)}$ is a polygon in the complex plane, and Eq. (60) shows that $f(\Theta_U, \Theta_V)$ is simply the distance d from this polygon to the origin. On the other hand, recalling the formula[12] $\mathcal{D}(\Theta_U, \Theta_V) = \sqrt{1 - d^2}$, we see that the upper bound in Eq. (58) is saturated by the unitarily implemented channels.

B. Random unitary channels

Continuing with the setup from the preceding example, let us consider channels of the form

$$\Phi(B) = \sum_i p_i \Theta_{U_i}(B), \quad \Psi(B) = \sum_i q_i \Theta_{U_i}(B), \tag{61}$$

where Θ_{U_i} are unitarily implemented channels and $\mathbf{p} \equiv \{p_i\}$, $\mathbf{q} \equiv \{q_i\}$ are probability distributions. It then follows immediately from the strong concavity property (55) of the minimax fidelity that

$$f(\Phi, \Psi) \geqslant \sum_i \sqrt{p_i q_i} \equiv F(\mathbf{p}, \mathbf{q}). \tag{62}$$

When $\dim \mathfrak{h} < \infty$, the inequality in (62) becomes equality when the unitaries U_i are orthogonal in the Hilbert-Schmidt sense, $\mathrm{Tr}\,U_i^{\dagger}U_k = \dim \mathfrak{h} \cdot \delta_{ik}$. On the other hand, when $\mathfrak{h}$ is infinite dimensional, this orthogonality condition does not make sense unless we consider channels given in terms of continual Kraus decompositions, so that the sums in Eq. (62) are replaced with integrals with respect to some positive measure μ, and agree to understand orthogonality in the sense of operator-valued Schwartz distributions. As an example, consider the following.

Let $\mathfrak{h} = \mathcal{F}$, the boson Fock space, let a and $a^{\dagger}$ be the field annihilation and creation operators, and let $D(z) := \exp(za^{\dagger} - \bar{z}a)$, $z \in \mathbb{C}$, be the unitary displacement operators obeying the Weyl relation $D(z)D(z') = e^{\mathrm{i}\,\mathrm{Im}\,z\bar{z}'} D(z+z')$. Given a function $f \in L^2(\mathbb{C}, \mathrm{d}z)$, where $\mathrm{d}z := \mathrm{d}(\mathrm{Re}\,z)\mathrm{d}(\mathrm{Im}\,z)$, we define its *Weyl-Fourier transform* as $D(f) := \pi^{-1/2}\int_{\mathbb{C}} f(z)D(z)\mathrm{d}z$. Since f is square integrable, $D(f)$ is a Hilbert-Schmidt operator, and it can be easily shown that

$$\mathrm{Tr}[D(f)^{\dagger}D(g)] = \int_{\mathbb{C}} \overline{f(z)}g(z)\mathrm{d}z \equiv \langle f, g \rangle_{L^2(\mathbb{C})}, \quad \forall f, g \in L^2(\mathbb{C})$$

so that $\mathrm{Tr}[D(z)^{\dagger}D(z')] = \pi \delta^{(2)}(z - z'), z, z' \in \mathbb{C}$, where $\delta^{(2)}(\lambda) := \delta(\mathrm{Re}\,\lambda)\delta(\mathrm{Im}\,\lambda)$ is the Dirac δ-function in the complex plane.

With this in mind, consider the family of channels $\Gamma^{(\mu)} : \mathcal{B}(\mathcal{F}) \to \mathcal{B}(\mathcal{F})$, $\mu \in \mathbb{R}^+$, with the preduals given by

$$\Gamma_{\mathsf{T}}^{(\mu)}(\rho) := \frac{1}{\pi\mu}\int_{\mathbb{C}} D(z)\rho D(z)^{\dagger}\exp(-|z|^2/\mu)\mathrm{d}z$$

(in quantum optics these channels model the so-called *Gaussian displacement noise*[25]). Then the minimax fidelity between $\Gamma^{(\mu)}$ and $\Gamma^{(\nu)}$ is given by

$$f(\Gamma^{(\mu)}, \Gamma^{(\nu)}) = \frac{(\mu\nu)^{1/2}}{\frac{1}{2}(\mu + \nu)}. \tag{63}$$

Owing to the inequality between the geometric and the arithmetic means, the right-hand side of Eq. (63) is always bounded between 0 and 1, and the maximum value of 1 is attained if and only

if $\mu = \nu$, i.e., $\Gamma^{(\mu)} = \Gamma^{(\nu)}$. This, of course, agrees with the properties of the minimax fidelity (cf. Sec. V).

C. Master equation

Consider a strongly continuous semigroup of channels $\{\Phi^{(t)} : \mathcal{B}(\mathfrak{h}) \rightarrow \mathcal{B}(\mathfrak{h})\}_{t \in \mathbb{R}^+}$, with the preduals $\Phi_\mathsf{T}^{(t)}$ satisfying the Lindblad master equation[26]

$$\frac{d\Phi_\mathsf{T}^{(t)}(\rho)}{dt} = X\rho X^\dagger - \frac{1}{2}(X^\dagger X \rho + \rho X^\dagger X) \tag{64}$$

for some $X \in \mathcal{B}(\mathfrak{h})$. Introducing the dilating Hilbert space $\mathcal{H} = \mathfrak{h} \otimes \mathfrak{h}$ with the basis $\{|0\rangle, |1\rangle, \ldots\}$, we can, for an infinitesimal time $t = \varepsilon$, write the predual of the channel $\Phi^{(\varepsilon)}$ in the Stinespring form

$$\Phi_\mathsf{T}^{(\varepsilon)}(\rho) = \mathrm{Tr}_\mathcal{H} A_\varepsilon \rho A_\varepsilon^\dagger, \tag{65}$$

where the map $A_\varepsilon : \mathfrak{h} \rightarrow \mathfrak{h} \otimes \mathcal{H}$ is given by

$$A_\varepsilon v := \left(\mathbb{1}_\mathfrak{h} - \tfrac{1}{2}\varepsilon X^\dagger X\right) v \otimes |0\rangle + \sqrt{\varepsilon} X v \otimes |1\rangle + O(\varepsilon^2), \tag{66}$$

$O(\varepsilon^2)$ indicating terms with norm bounded from above by $M\varepsilon^2$ for some constant $M \geq 0$. Note that $A_0 v = v \otimes |0\rangle$, so that $T^{(0)} = \mathrm{id}$. We can then evaluate the partial trace

$$\mathrm{Tr}_\mathfrak{h}[A_\varepsilon \rho A_0^\dagger] = \left(1 - \tfrac{1}{2}\varepsilon \langle X^\dagger X\rangle_\rho\right)|0\rangle\langle 0| + \sqrt{\varepsilon}\langle X\rangle_\rho |1\rangle\langle 0| + O(\varepsilon^2), \tag{67}$$

where $\langle B\rangle_\rho := \mathrm{Tr}(B\rho)$ for $B \in \mathcal{B}(\mathfrak{h})$. Then, again up to an additive term of operator norm $O(\varepsilon^2)$,

$$\mathrm{Tr}_\mathfrak{h}[A_\varepsilon \rho A_0^\dagger]^\dagger \mathrm{Tr}_\mathfrak{h}[A_\varepsilon \rho A_0^\dagger] \approx \left[\left(1 - \tfrac{1}{2}\varepsilon\langle X^\dagger X\rangle_\rho\right)^2 + \varepsilon|\langle X\rangle_\rho|^2\right]|0\rangle\langle 0|, \tag{68}$$

which allows us to compute, up to $O(\varepsilon^2)$, the minimax fidelity between the channel $T^{(\varepsilon)}$ after an infinitesimal time ε and the identity map. Using Eq. (53), we obtain

$$f(T^{(\epsilon)}, \mathrm{id}) = \inf_{\varrho \in \mathcal{S}(\mathfrak{g})} \left\| \mathrm{Tr}_\mathfrak{h}[A_\varepsilon \rho A_0^\dagger] \right\|_\mathsf{T} \approx \sqrt{1 - \varepsilon C}, \tag{69}$$

where

$$C = \inf_{\varrho \in \mathcal{S}(\mathfrak{g})} \left(\langle X^\dagger X\rangle_\rho - |\langle X\rangle_\rho|^2\right). \tag{70}$$

D. Impossibility of quantum bit commitment

The statement of topological equivalence of the noncommutative Hellinger distance and the CB-norm distance between a pair of quantum channels, i.e., Eq. (59), is essentially the "continuity argument" at the heart of a proof of "impossibility of quantum bit commitment (QBC)."[27] Quantum bit commitment is a cryptographic objective in which one party, Alice, commits a bit to another party, Bob, in such a way that the corresponding protocol is *concealing* (i.e., Bob is not able to retrieve the bit before the opening) and *binding* (i.e., Alice cannot change the bit after the commitment). The impossibility proof asserts that if the protocol is perfectly concealing, then it is necessarily not binding, and invokes a continuity argument for "asymptotically" concealing protocols, stating that Alice's probability of successful cheating approaches unity, while Bob's cheating probability becomes close to the value $\frac{1}{2}$ (pure guessing). (The reader should be aware that the impossibility proof in Ref. 27 is valid for a restricted class of protocols, i.e., those that are nonaborting and have a single commitment step. For wider classes of protocols, it is still a matter of debate whether a secure QBC protocol exists.[28]) In this example we derive the continuity argument from the expression of Alice's and Bob's respective cheating probabilities as a consequence of the topological equivalence between the Hellinger distance and the CB-norm distance in Eq. (59).

From the point of view of Bob, Alice's action of committing the bit is equivalent to a channel $\Phi_{\mathbf{A}^{(b)}}$ on an algebra $\mathcal{B}(\mathfrak{h})$, dim $\mathfrak{h} < \infty$, for each value of the committed bit $b = 0, 1$, where $\mathbf{A}^{(b)} \equiv \{A_j^{(b)}\}_{j=1}^k$ is a collection of operators satisfying the Kraus condition $\Sigma_{j=1}^k A_j^{(b)\dagger} A_j^{(b)} = \mathbb{1}$, and $\Phi_{\mathbf{A}^{(b)}}$ denotes the channel induced by this Kraus decomposition. At the opening, Alice informs Bob about which element of the Kraus decomposition $\mathbf{A}^{(b)}$ she actually used in the commitment. However, prior to unveiling the label j, Alice can perform an *EPR attack* with the purpose of changing the Kraus decomposition to another equivalent decomposition $\mathbf{A}^{(b)}(V) \equiv \{A_j^{(b)}(V)\}$, where $A_j^{(b)}(V) := \Sigma_\ell A_l^{(b)} V_{j\ell}$ for some $V \in \mathsf{U}(\mathbb{C}^k)$. The EPR attack is achieved by Alice via the unitary transformation V on an ancillary k-dimensional space $\mathcal{H}$. The conditional probability that Alice can cheat successfully by convincing Bob that she has committed, say, $b = 1$, while having successfully committed $b = 0$ instead, is given by

$$P_c^A(V, v) = \sum_j \frac{|\langle v|A_j^{(0)\dagger}(V)A_j^{(1)} \otimes \mathbb{1}_{\mathcal{H}}|v\rangle|^2}{\|(A_j^{(1)} \otimes \mathbb{1}_{\mathcal{H}})v\|^2}, \tag{71}$$

where $\|(A_j^{(0)} \otimes \mathbb{1}_{\mathcal{H}})v\|^2$ is the probability that the jth Kraus element is unveiled. Which V should Alice use? Without any knowledge of $|v\rangle$, the best she can do is to adopt a conservative strategy of choosing the V that will maximize her cheating probability in the worst-case scenario, namely for the anonymous state $|v\rangle$ chosen by Bob to minimize $P_c^A(V, v)$. This is the *minimax* choice of V, corresponding to the cheating probability

$$\bar{P}_c^A := \sup_{V \in \mathsf{U}(\mathbb{C}^k)} \inf_{v \in \mathfrak{h} \otimes \mathcal{H}; \|v\|=1} P_c^A(V, v). \tag{72}$$

On the other hand, for equiprobable bit values $b \in \{0, 1\}$ Bob's optimal probability of cheating is given by the probability of error in discriminating between the corresponding output states, more precisely

$$\bar{P}_c^B = \frac{1}{2} + \frac{1}{4} \sup_{v \in \mathfrak{h} \otimes \mathcal{H}; \|v\|=1} \|\rho_{\mathbf{A}^{(0)}}^v - \rho_{\mathbf{A}^{(1)}}^v\|_{\mathsf{T}} = \frac{1}{2}[1 + \mathcal{D}(\Phi_{\mathbf{A}^{(0)}}, \Phi_{\mathbf{A}^{(1)}})], \tag{73}$$

where we have defined $\rho_{\mathbf{A}}^v := \Phi_{\mathbf{A}} \otimes \mathrm{id}(|v\rangle\langle v|)$. Using Jensen's inequality, we can bound Alice's cheating probability $P_c^A(V, v)$ from below as

$$P_c^A(V, v) \geqslant \left|\sum_j |\langle v|A_j^{(0)}(V)^\dagger A_j^{(1)} \otimes \mathbb{1}_{\mathcal{H}}|v\rangle|\right|^2. \tag{74}$$

Note that the value of the max-min in Eq. (72) will not change if we perform the maximization over the closed convex hull of $\mathsf{U}(\mathbb{C}^k)$, i.e., the set $\mathsf{K}(\mathbb{C}^k)$ of all linear contractions on $\mathbb{C}^k$, and the minimization over the closed convex hull of the pure states on $\mathfrak{h} \otimes \mathcal{H}$, i.e., the set $\mathcal{S}(\mathfrak{h} \otimes \mathcal{H})$ of states on $\mathcal{B}(\mathfrak{h} \otimes \mathcal{H})$, thus completing the domain of the max-min to the product $\mathsf{K}(\mathbb{C}^k) \times \mathcal{S}(\mathfrak{h} \otimes \mathcal{H})$ of compact convex sets. Now, the functional

$$F(V, \rho) := \sum_j \mathrm{Re}\, \mathrm{Tr}\{\rho[A_j^{(0)}(V)^\dagger A_j^{(1)} \otimes \mathbb{1}_{\mathcal{H}}]\} \tag{75}$$

is affine in both $V \in \mathsf{K}(\mathbb{C}^k)$ and $\rho \in \mathcal{S}(\mathfrak{h} \otimes \mathcal{H})$, so that we can use standard minimax arguments[24] to justify the interchange of extrema in Eq. (72), and then apply Lemma 2 to obtain

$$\sup_{V \in \mathsf{U}(\mathbb{C}^k)} \inf_{v \in \mathfrak{h} \otimes \mathcal{H}} |F(V, |v\rangle\langle v|)| = \sup_{V \in \mathsf{K}(\mathbb{C}^k)} \inf_{\varrho \in \mathcal{S}(\mathfrak{h} \otimes \mathcal{H})} |F(V, \rho)| \tag{76}$$

$$= \inf_{\varrho \in \mathcal{S}(\mathfrak{h} \otimes \mathcal{H})} \sup_{V \in \mathsf{K}(\mathbb{C}^k)} |F(V, \rho)| \tag{77}$$

$$= \inf_{v \in \mathfrak{h} \otimes \mathcal{H}; \|v\|=1} \sup_{V \in \mathsf{U}(k)} |F(V, |v\rangle\langle v|)|. \tag{78}$$

Now, since a monotone function does not affect the saddle point, we can use Eqs. (50), (72), (74), and (78) to obtain

$$\bar{P}_c^A \geqslant f^2(\Phi_{\mathbf{A}^{(0)}}, \Phi_{\mathbf{A}^{(1)}}).$$

Using Eq. (59) and then Eq. (73), we finally obtain the chain of estimates

$$\bar{P}_c^A \geqslant f^2(\Phi_{\mathbf{A}^{(0)}}, \Phi_{\mathbf{A}^{(1)}}) \geqslant [1 - \mathcal{D}(\Phi_{\mathbf{A}^{(0)}}, \Phi_{\mathbf{A}^{(1)}})]^2 \geqslant [1 - 2(\bar{P}_c^B - 1/2)]^2,$$

where it follows that, for "asymptotically" concealing protocols, i.e., those for which $\bar{P}_c^B \to \frac{1}{2}$, Alice's probability of cheating will approach unity, and the protocol will not be binding.

[1] M. Keyl, Phys. Rep. **369**, 431 (2002).
[2] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge U.P., Cambridge, 2000).
[3] A. Gilchrist, N. K. Langford, and M. A. Nielsen, Distance measures to compare real and ideal quantum processes, arXiv e-print quant-ph/0408063.
[4] P. M. Alberti, Rep. Math. Phys. **51**, 87 (2003).
[5] P. M. Alberti and A. Uhlmann, Acta Appl. Math. **60**, 1 (2000).
[6] R. Jozsa, J. Mod. Opt. **41**, 2315 (1994).
[7] A. Uhlmann, Rep. Math. Phys. **9**, 273 (1976).
[8] R. Schatten, *Norm Ideals of Completely Continuous Operators* (Springer, Berlin, 1960).
[9] B. Simon, *Trace Ideals and Their Applications* (Cambridge U.P., Cambridge, 1979).
[10] C. A. Fuchs and J. van de Graaf, IEEE Trans. Inf. Theory **45**, 1216 (1999).
[11] M. Raginsky, Phys. Lett. A **290**, 11 (2001).
[12] D. Aharonov, A. Kitaev, and N. Nisan, Quantum circuits with mixed states, in *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, 1997, pp. 20–30.
[13] A. Kitaev, Russ. Math. Surveys **52**, 1191 (1997).
[14] V. P. Belavkin and P. Staszewski, Rep. Math. Phys. **24**, 49 (1986).
[15] D. Kretschmann and R. F. Werner, New J. Phys. **6**, 26 (2004).
[16] V. P. Belavkin and M. Ohya, Proc. R. Soc. London, Ser. A **458**, 209 (2002).
[17] K. Kraus, *States, Effects and Operations: Fundamental Notions of Quantum Theory* (Springer, New York, 1983).
[18] K. Yosida, *Functional Analysis* (Springer-Verlag, Berlin, 1965).
[19] M. Raginsky, J. Math. Phys. **44**, 5003 (2003).
[20] V. I. Paulsen, *Completely Bounded Maps and Operator Algebras* (Cambridge U.P., Cambridge, 2003).
[21] V. P. Belavkin, "Contravariant densities, operational distances and quantum channel fidelities," presented at the Sixth International Conference on Quantum Communication, Measurement, and Computing (QCM&C'02), Boston, July 2002; extended version available as arXiv e-print math-ph/0408035 (2004).
[22] W. F. Stinespring, Proc. Am. Math. Soc. **6**, 211 (1955).
[23] S. Sakai, C^*-algebras and W^*-algebras (Springer-Verlag, Berlin, 1971).
[24] D. G. Luenberger, *Optimization by Vector Space Methods* (Wiley, New York, 1969).
[25] M. J. W. Hall, Phys. Rev. A **50**, 3295 (1994).
[26] G. Lindblad, Commun. Math. Phys. **48**, 119 (1976).
[27] H. K. Lo and H. F. Chau, Phys. Rev. Lett. **78** 3410 (1997).
[28] Protocols for quantum bit commitment, claimed to lie outside the framework of the impossibility proof of Ref. 27, were first proposed by H. P. Yuen in arXiv quant-ph/0006109 and quant-ph/0009113; see also H. P. Yuen, in *Quantum Communications and Measurements II*, edited by P. Kumar, G. M. D'Ariano, and O. Hirota (Kluwer Academic/Plenum, New York, 2000), p. 399.