Article:
Pirandola, Stefano orcid.org/0000-0001-6165-5615, Laurenza, Riccardo, Lupo, Cosmo orcid.org/0000-0002-5227-4009 et al. (1 more author) (2019) Fundamental limits to quantum channel discrimination. npj Quantum Information. 3. ISSN 2056-6387

https://doi.org/10.1038/s41534-019-0162-y

Fundamental limits to quantum channel discrimination

Stefano Pirandola, Riccardo Laurenza, and Cosmo Lupo
Computer Science and York Centre for Quantum Technologies, University of York, York YO10 5GH, United Kingdom

What is the ultimate performance for discriminating two arbitrary quantum channels acting on a finite-dimensional Hilbert space? Here we address this basic question by deriving a general and fundamental lower bound. More precisely, we investigate the symmetric discrimination of two arbitrary qudit channels by means of the most general protocols based on adaptive (feedback-assisted) quantum operations. In this general scenario, we first show how port-based teleportation can be used to completely simplify these adaptive protocols into a much simpler non-adaptive form, designing a new form of teleportation stretching. Then, we prove that the minimum error probability affecting the channel discrimination cannot beat a bound determined by the Choi matrices of the channels, establishing an ultimate and elegant formula for quantum hypothesis testing. As a consequence of this bound, we derive the ultimate limits for adaptive quantum illumination and single-photon quantum optical resolution. Finally, we show that our methodology can also be applied to other tasks, such as quantum metrology, quantum communication and secret key generation.

PACS numbers: 03.67.-a, 42.50.-p, 03.65.-w

Quantum hypothesis testing [1–8] is a central area in quantum information theory [9, 10], with many studies for both discrete variable (DV) [11] and continuous variable (CV) systems [12]. A number of tools [13–17] have been developed for its basic formulation, known as quantum state discrimination. In particular, since the seminal work of Helstrom in the 70s [1], we know how to bound the error probability affecting the discrimination of two arbitrary quantum states [18]. Remarkably, a similar bound is still missing for the discrimination of two arbitrary quantum channels. The main problem in quantum channel discrimination (QCD) [19–23] is that the strategies involve an optimization over the input state and the output measurement, and this process may also be adaptive in the most general case, so that feedback from the output is used to update the input.

The ultimate performance of adaptive QCD is not known because of the extreme difficulty to handle feedback-assistance in quantum protocols. At the same time, it is also known that adaptiveness needs to be considered in QCD. In fact, apart from the cases where two channels are classical [24], jointly programmable or teleportation-covariant [25, 26], feedback may greatly improve the discrimination. For instance, Ref. [27] presented two channels which can be perfectly distinguished by using feedback in just two adaptive uses, while they cannot be perfectly discriminated by any number of uses of a block (non-adaptive) protocol, where the channels are probed in an identical and independent fashion, i.e., using multiple copies of the same input state. This implies that the best discrimination performance is not related to the diamond-norm distance [28], when multiple copies of two quantum channels are considered.

Here we fill this gap by deriving a universal lower bound for the error probability affecting the discrimination of two arbitrary quantum channels. To derive this bound we design a technique which reduces an adaptive protocol over an arbitrary finite-dimensional quantum channel into a block protocol over multiple copies of the channel's Choi matrix. This is obtained by using port-based teleportation (PBT) [29–36] for channel simulation and suitably generalizing the technique of teleportation stretching [37, 38]. This reduction is shown in general for adaptive protocols with any task (not just QCD). When applied to QCD, it allows us to bound the ultimate error probability by using the Choi matrices of the channels. As a direct application of this result, we bound the ultimate adaptive performance of quantum illumination [39–47] and the ultimate adaptive resolution of any single-photon diffraction-limited optical system. Other implications for quantum metrology, quantum and private communications are discussed in Ref. [48].

Adaptive protocols.– Let us formulate the most general adaptive protocol over an arbitrary quantum channel $\mathcal{E}$ defined between Hilbert spaces of dimension d (more generally, this can be taken as the dimension of the input space). We first provide a general description and then we specify the protocol to the task of QCD. A general adaptive protocol involves an unlimited number of quantum systems which may be subject to completely arbitrary quantum operations (QOs). More precisely, we may organize the quantum systems into an input register $\mathbf{a}$ and an output register $\mathbf{b}$, which are prepared in an initial state ρ_0 by applying a QO Λ_0 to some fundamental state of $\mathbf{a}$ and $\mathbf{b}$. Then, a system a_1 is picked from the register $\mathbf{a}$ and sent through the channel $\mathcal{E}$. The corresponding output b_1 is merged with the output register $b_1\mathbf{b} \to \mathbf{b}$. This is followed by another QO Λ_1 applied to $\mathbf{a}$ and $\mathbf{b}$. Then, we send a second system $a_2 \in \mathbf{a}$ through $\mathcal{E}$ with the output b_2 being merged again $b_2\mathbf{b} \to \mathbf{b}$ and so on. After n uses, the registers will be in a state ρ_n which depends on $\mathcal{E}$ and the sequence of QOs $\{\Lambda_0, \Lambda_1, \ldots, \Lambda_n\}$ defining the adaptive protocol $\mathcal{P}_n$.

In a protocol of quantum communication, the registers belong to remote users and, in absence of entanglement-assistance, the QOs are local operations (LOs) assisted

by two-way classical communication (CC), also known as adaptive LOCCs. The output is generated in such a way to approximate some target state [37]. In a protocol of quantum channel estimation, the channel is labelled by a continuous parameter $\mathcal{E} = \mathcal{E}_\theta$ and the QOs include the use of entanglement across the registers. The output state will encode the unknown parameter $\rho_n = \rho_n(\theta)$, which is detected and the outcome processed into an optimal estimator [25]. Here, in a protocol of binary and symmetric QCD, the channel is labelled by a binary digit, i.e., $\mathcal{E} = \mathcal{E}_u$ where $u \in \{0, 1\}$ has equal priors. The QOs are generally entangled and they generate an output state encoding the information bit, i.e., $\rho_n = \rho_n(u)$.

The output state $\rho_n(u)$ of an adaptive discrimination protocol $\mathcal{P}_n$ is finally detected by an optimal positive-operator valued measure (POVM). For binary discrimination, this is the Helstrom POVM, which leads to the conditional error probability

$$p(\mathcal{E}_0 \neq \mathcal{E}_1|\mathcal{P}_n) = \frac{1 - D\left[\rho_n(0), \rho_n(1)\right]}{2}, \tag{1}$$

where $D(\rho, \sigma) := ||\rho - \sigma||/2$ is the trace distance [11]. The optimization over all discrimination protocols $\mathcal{P}_n$ defines the minimum error probability affecting the n-use adaptive discrimination of $\mathcal{E}_0$ and $\mathcal{E}_1$, i.e., we may write

$$p_n(\mathcal{E}_0 \neq \mathcal{E}_1) := \inf_{\mathcal{P}_n} \; p(\mathcal{E}_0 \neq \mathcal{E}_1|\mathcal{P}_n). \tag{2}$$

This is generally less than the n-copy diamond distance between the two channels $\mathcal{E}_0^{\otimes n}$ and $\mathcal{E}_1^{\otimes n}$ [49]

$$p_n(\mathcal{E}_0 \neq \mathcal{E}_1) \leq \frac{1 - \frac{1}{2}||\mathcal{E}_0^{\otimes n} - \mathcal{E}_1^{\otimes n}||_\diamond}{2}. \tag{3}$$

Can we complete Eq. (3) with a lower bound? Up today this has been only proven for jointly-programmable channels, i.e., channels $\mathcal{E}_0$ and $\mathcal{E}_1$ admitting a simulation $\mathcal{E}_u(\rho) = \mathcal{S}(\rho \otimes \pi_u)$ with a trace-preserving QO $\mathcal{S}$ and different program states π_0 and π_1. In this case, we have $p_n \geq [1 - D(\pi_0^{\otimes n}, \pi_1^{\otimes n})]/2$ [25]. In particular, this is true if the channels are jointly teleportation-covariant, so that $\mathcal{S}$ becomes teleportation and the program state is a Choi matrix $\rho_{\mathcal{E}_u}$. For these channels, Ref. [25] found that Eq. (3) holds with an equality and we may write $||\mathcal{E}_0^{\otimes n} - \mathcal{E}_1^{\otimes n}||_\diamond = ||\rho_{\mathcal{E}_0}^{\otimes n} - \rho_{\mathcal{E}_1}^{\otimes n}||$. More precisely, the question to ask is therefore the following: Can we establish a *universal* lower bound for $p_n(\mathcal{E}_0 \neq \mathcal{E}_1)$ which is valid for *arbitrary* channels? As we show here, this is possible by resorting to a more general (multi-program) simulation of the channels, i.e., of the type $\mathcal{S}(\rho \otimes \pi_u^{\otimes M})$ [50].

Port-based teleportation.– Let us describe the protocol of PBT with qudits of arbitrary dimension $d \geq 2$ [30, 31]. The parties exploit two ensembles of $M \geq 2$ qudits, i.e., Alice has $\mathbf{A} := \{A_1, \ldots, A_M\}$ and Bob has $\mathbf{B} := \{B_1, \ldots, B_M\}$ representing the output "ports". The generic ith pair (A_i, B_i) is prepared in a maximally-entangled state, so that we have the global state [53]

$$\Phi_{\mathbf{AB}}^{\otimes M} = \bigotimes_{i=1}^{M} |\Phi\rangle_i\langle\Phi|, \quad |\Phi\rangle_i := d^{-1/2} \sum_k |k\rangle_{A_i} \otimes |k\rangle_{B_i}. \tag{4}$$

To teleport the state of a qudit C, Alice performs a joint measurement on C and her ensemble $\mathbf{A}$. This is a POVM $\{\Pi_{C\mathbf{A}}^i\}_{i=1}^M$ with M possible outcomes (see Refs. [30, 31] for the details). Once Alice communicates the outcome i to Bob, he discards all the ports but the ith one, which contains the teleported state (see Fig. 1a).

The measurement outcomes are equiprobable and independent of the input, and the output state is invariant under permutation of the ports [54]. Averaging over the outcomes, we define the teleported state $\rho_B^M = \Gamma_M(\rho_C)$, where Γ_M is the corresponding PBT channel [55]. In the limit of many ports M, we have that Γ_M approximates an identity channel $\mathcal{I}$ so that Bob's output becomes a perfect replica of Alice's input. More precisely, for any M, we prove [48] the following error in diamond distance.

Lemma 1 *In arbitrary (finite) dimension d, the diamond distance between the M-port PBT channel Γ_M and the identity channel $\mathcal{I}$ satisfies*

$$\delta_M := ||\mathcal{I} - \Gamma_M||_\diamond \leq 2d(d-1)M^{-1}. \tag{5}$$

Channel simulation via PBT.– Let us discuss how PBT can be used for channel simulation. As shown in Fig. 1b, suppose that Bob applies an arbitrary channel $\mathcal{E}$ to the teleported output, so that Alice's input ρ_C is subject to the approximate channel

$$\mathcal{E}^M(\rho_C) := \mathcal{E} \circ \Gamma_M(\rho_C). \tag{6}$$

We note that the port selection commutes with $\mathcal{E}$, because the POVM acts on a different Hilbert space. Therefore, Bob can equivalently apply $\mathcal{E}$ to each port before Alice's CC, i.e., apply $\mathcal{E}^{\otimes M}$ to his $\mathbf{B}$ qudits before selecting the output port, as shown in Fig. 1c. This leads to the following simulation for the approximate channel

$$\mathcal{E}^M(\rho_C) = \mathcal{T}^M(\rho_C \otimes \rho_{\mathcal{E}}^{\otimes M}), \tag{7}$$

where $\mathcal{T}^M$ is a trace-preserving LOCC and $\rho_{\mathcal{E}}$ is the channel's Choi matrix (see Fig. 1d). By construction, the simulation LOCC $\mathcal{T}^M$ is universal, i.e., it does not depend on the channel $\mathcal{E}$. This means that, at fixed M, the channel $\mathcal{E}^M$ is fully determined by the program state $\rho_{\mathcal{E}}$. We can bound the accuracy of the simulation. From Eq. (6) and the monotonicity of the diamond norm, we get

$$||\mathcal{E} - \mathcal{E}^M||_\diamond \leq \delta_M, \tag{8}$$

where δ_M is the simulation error in Eq. (5), with the dimension d being the one of the input Hilbert space.

PBT stretching of an adaptive protocol.– In order to use the PBT simulation to simplify an adaptive protocol, there are two main steps. First of all, we need to replace

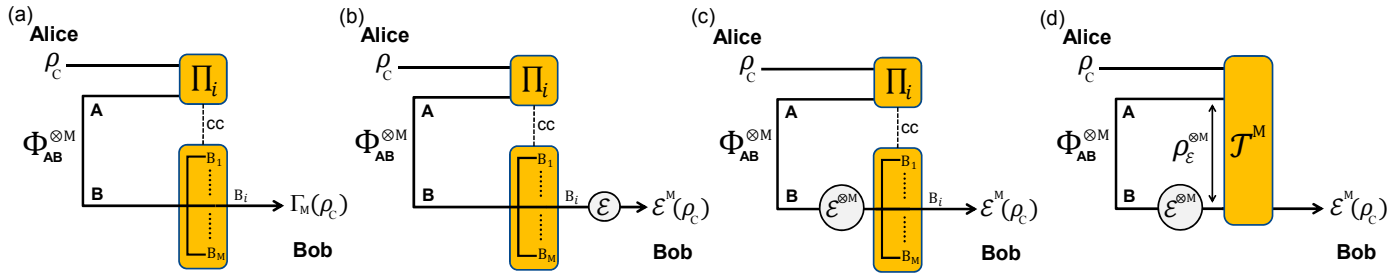

FIG. 1: From port-based teleportation (PBT) to the Choi-simulation of an arbitrary quantum channel. **(a)** Schematic representation of the PBT protocol. Alice and Bob share an $M \times M$ qudit state which is given by M maximally-entangled states $\Phi_{\mathbf{AB}}^{\otimes M}$. To teleport an input qubit state ρ_C, Alice applies a suitable POVM $\{\Pi_i\}$ to the input qubit C and her $\mathbf{A}$ qubits. The outcome i is communicated to Bob, who selects the i-th among his $\mathbf{B}$ qubits (tracing all the others). The performance does not depend on the specific "port" i selected and the average output state is given by $\Gamma_M(\rho_C)$ where Γ_M is the PBT channel. The latter reduces to the identity channel in the limit of many ports $M \to \infty$. **(b)** Suppose that Bob applies a quantum channel $\mathcal{E}$ on his teleported output. This produces the output state $\mathcal{E}^M(\rho_C)$ of Eq. (6). For large M, one has $\mathcal{E}^M \to \mathcal{E}$ in diamond norm. **(c)** Equivalently, Bob can apply $\mathcal{E}^{\otimes M}$ to all his qubits $\mathbf{B}$ in advance to the CC from Alice. After selection of the port, this will result in the same output as before. **(d)** Now note that Alice's LO and Bob's port selection form a global LOCC $\mathcal{T}^M$ (trace-preserving by averaging over the outcomes). This is applied to a tensor-product state $\rho_{\mathcal{E}}^{\otimes M}$ where $\rho_{\mathcal{E}}$ is the Choi matrix of the original channel $\mathcal{E}$. Thus the approximate channel $\mathcal{E}^M$ is simulated by applying $\mathcal{T}^M$ to $\rho_C \otimes \rho_{\mathcal{E}}^{\otimes M}$ as in Eq. (7).

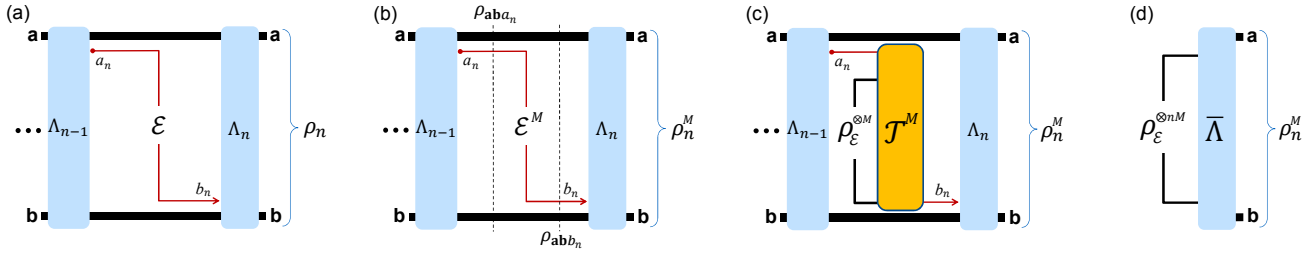

FIG. 2: Port-based teleportation stretching of a generic adaptive protocol over a quantum channel $\mathcal{E}$. This channel is fixed in quantum/private communication, while it is unknown and parametrized in estimation/discrimination problems. **(a)** We show the last transmission $a_n \to b_n$ through $\mathcal{E}$, which occurs between two adaptive QOs Λ_{n-1} and Λ_n. This last step produces the output state ρ_n. **(b)** In each transmission, we replace $\mathcal{E}$ with its M-port simulation $\mathcal{E}^M$ so that the output of the protocol becomes ρ_n^M which approximates ρ_n for large M. Note that, in the last transmission, the register state $\rho_{\mathbf{ab}a_n}$ undergoes the transformation $\rho_{\mathbf{ab}b_n} = \mathcal{I}_{\mathbf{ab}} \otimes \mathcal{E}^M(\rho_{\mathbf{ab}a_n})$. **(c)** Each propagation through $\mathcal{E}^M$ is replaced by its PBT simulation. For the last transmission, this means that $\rho_{\mathbf{ab}b_n} = \mathcal{I}_{\mathbf{ab}} \otimes \mathcal{T}^M(\rho_{\mathbf{ab}a_n} \otimes \rho_{\mathcal{E}}^{\otimes M})$ where $\mathcal{T}^M$ is the LOCC of the PBT and $\rho_{\mathcal{E}}$ is the Choi matrix of the original channel. **(d)** All the adaptive QOs Λ_i and the simulation LOCCs $\mathcal{T}^M$ are collapsed into a single (trace-preserving) QO $\bar{\Lambda}$. Correspondingly, n instances of $\rho_{\mathcal{E}}^{\otimes M}$ are collected. As a result, the approximate output ρ_n^M is given by $\bar{\Lambda}$ applied to the tensor-product state $\rho_{\mathcal{E}}^{\otimes nM}$ as in Eq. (10).

each channel $\mathcal{E}$ with its M-port approximation $\mathcal{E}^M$ while controlling the propagation of the simulation error δ_M from the channel to the output state. This step is crucial also in simulations via standard teleportation [26, 38] (see also Refs. [57–61]). Second, we need to "stretch" the protocol [37] by replacing the approximate channel $\mathcal{E}^M$ with its Choi matrices $\rho_{\mathcal{E}}^{\otimes M}$ and then suitably re-organize all the remaining QOs. Here we describe the technique for a generic task, before specifying it for QCD.

Given an adaptive protocol $\mathcal{P}_n$ over a channel $\mathcal{E}$ with output ρ_n, consider the same protocol over the simulated channel $\mathcal{E}^M$, so that we get the different output ρ_n^M. Using a "peeling" argument [48], we bound the output error in terms of the channel simulation error

$$||\rho_n - \rho_n^M|| \leq n||\mathcal{E} - \mathcal{E}^M||_\diamond \leq n\delta_M. \tag{9}$$

Once understood that the output state can be closely approximated, let us simplify the adaptive protocol over $\mathcal{E}^M$. Using the simulation in Eq. (7), we may replace each channel $\mathcal{E}^M$ with the resource state $\rho_{\mathcal{E}}^{\otimes M}$, iterate the process for all n uses, and collapse all the simulation LOCCs and QOs as shown in Fig. 2. As a result, we may write the multi-copy Choi decomposition

$$\rho_n^M = \bar{\Lambda}(\rho_{\mathcal{E}}^{\otimes nM}) \ , \tag{10}$$

for a trace-preserving QO $\bar{\Lambda}$. Now, we can combine the two ingredients of Eqs. (9) and (10), into the following.

Lemma 2 (PBT stretching) *Consider an adaptive quantum protocol (with arbitrary task) over an arbitrary d-dimensional quantum channel $\mathcal{E}$ (which may be unknown and parametrized). After n uses, the output ρ_n*

of the protocol can be decomposed as follows

$$||\rho_n - \bar{\Lambda}(\rho_{\mathcal{E}}^{\otimes nM})|| \leq n\delta_M, \tag{11}$$

where $\bar{\Lambda}$ is a trace-preserving QO, $\rho_{\mathcal{E}}$ is the Choi matrix of $\mathcal{E}$, and δ_M is the M-port simulation error in Eq. (5).

When we apply the lemma to protocols of quantum or private communication, where the QOs Λ_i are LOCCs, then we may write Eq. (11) with $\bar{\Lambda}$ being a LOCC. In protocols of channel estimation or discrimination, where $\mathcal{E}$ is parametrized, we may write Eq. (11) with $\rho_{\mathcal{E}}$ storing the parameter of the channel. In particular, for QCD we have $\{\mathcal{E}_u\}_{u=0,1}$ and the output $\rho_n(u)$ of the adaptive protocol $\mathcal{P}_n$ can be decomposed as follows

$$||\rho_n(u) - \bar{\Lambda}(\rho_{\mathcal{E}_u}^{\otimes nM})|| \leq n\delta_M. \tag{12}$$

Ultimate bound for channel discrimination.– We are now ready to show the lower bound for minimum error probability $p_n(\mathcal{E}_0 \neq \mathcal{E}_1)$ in Eq. (3). Consider an arbitrary protocol $\mathcal{P}_n$, for which we may write Eq. (1). Combining Lemma 2 with the triangle inequality leads to

$$\begin{aligned} ||\rho_n(0) - \rho_n(1)|| &\leq 2n\delta_M + ||\bar{\Lambda}(\rho_{\mathcal{E}_0}^{\otimes nM}) - \bar{\Lambda}(\rho_{\mathcal{E}_1}^{\otimes nM})|| \\ &\leq 2n\delta_M + ||\rho_{\mathcal{E}_0}^{\otimes nM} - \rho_{\mathcal{E}_1}^{\otimes nM}||, \end{aligned} \tag{13}$$

where we also use the monotonicity of the trace distance under channels. Because $\bar{\Lambda}$ is lost, the bound does no longer depend on the details of the protocol $\mathcal{P}_n$, which means that it applies to all adaptive protocols. Thus, using Eq. (13) in Eqs. (1) and (2), we get the following.

Theorem 3 *Consider the adaptive discrimination of two channels $\{\mathcal{E}_u\}_{u=0,1}$ in dimension d. After n probings, the minimum error probability satisfies the bound*

$$p_n(\mathcal{E}_0 \neq \mathcal{E}_1) \geq B := \frac{1 - n\delta_M - D(\rho_{\mathcal{E}_0}^{\otimes nM}, \rho_{\mathcal{E}_1}^{\otimes nM})}{2}, \tag{14}$$

where M may be chosen to maximize the right hand side.

Let us bound the trace distance in Eq. (14) as [62]

$$D^2 \leq 1 - F^{2nM}, \; F := \mathrm{Tr}\sqrt{\sqrt{\rho_{\mathcal{E}_1}}\rho_{\mathcal{E}_2}\sqrt{\rho_{\mathcal{E}_1}}}, \tag{15}$$

where F is the fidelity [63] between the Choi matrices of the channels. If we also exploit Eq. (5), we may write

$$B \geq \frac{1}{2} - \frac{\sqrt{1 - F^{2nM}}}{2} - \frac{d(d-1)n}{M}. \tag{16}$$

In the previous formula there are terms with opposite monotonicity in M. For this reason, the maximum value of B is achieved at some intermediate value of M.

One possible choice [66] is $M = 4d(d-1)n$, so that $B \geq (1 - 2\sqrt{1 - F^{8d(d-1)n^2}})/4$. In particular, consider two infinitesimally-close channels, so that $F \simeq 1-\epsilon$ where $\epsilon \simeq 0$ is the infidelity. By expanding in ϵ, we may write

$$B \geq \frac{1}{4} - n\sqrt{2d(d-1)\epsilon} \simeq \frac{\exp(-4n\sqrt{2d(d-1)\epsilon})}{4}. \tag{17}$$

Discriminating between two close quantum channels is a problem in many physical scenarios. For instance, this is typical in quantum optical resolution [67–69] (discussed below), quantum illumination [39–47] (discussed below), ideal quantum reading [70–74], quantum metrology [75–79] (treated in Ref. [48]), and also tests of quantum field theories in non-inertial frames [80], e.g., for detecting effects such as the Unruh or the Hawking radiation.

Single-photon quantum optical resolution.– Consider a microscope-type problem where we aim at locating a point in two possible positions, either $s/2$ or $-s/2$, where the separation s is very small. Assume we are limited to use probe states with at most one photon and an output finite-aperture optical system (this makes the optical process to be a qubit-to-qutrit channel, so that the input dimension is $d = 2$). Apart from this, we are allowed to use an arbitrary large quantum computer and arbitrary QOs to manipulate its registers. We may apply Eq. (17) with $\epsilon \simeq \eta s^2/16$, where η is a diffraction-related loss parameter [48]. In this way, we find that the error probability affecting the discrimination of the two positions is approximately bounded by $B \gtrsim \frac{1}{4}\exp(-2ns\sqrt{\eta})$ [48].

Adaptive quantum illumination.– Consider the protocol of quantum illumination in the DV setting [39]. Here the problem is to discriminate the presence or not of a target with low reflectivity $\eta \simeq 0$ in a thermal background which has $b \ll 1$ mean thermal photons per optical mode. One assumes that d modes are used in each probing of the target and each of them contains at most one photon. This means that the Hilbert space is $(d+1)$-dimensional with basis $\{|0\rangle, |1\rangle, \ldots, |d\rangle\}$, where $|i\rangle := |0\cdots010\cdots0\rangle$ has one photon in the ith mode. If the target is absent ($u = 0$), the receiver detects thermal noise; if the target is present ($u = 1$), the receiver measures a mixture of signal and thermal noise. In the most general (adaptive) version of the protocol, the receiver belongs to a large quantum computer where the $(d+1)$-dimensional signal qudits are picked from an input register, sent to target, and their reflection stored in an output register, with adaptive QOs performed between each probing. After n probings, the state of the registers $\rho_n(u)$ is optimally detected. Assuming the typical regime of quantum illumination [39], we find that the error probability affecting target detection is approximately bounded by $B \gtrsim \frac{1}{4}\exp(-4nd\sqrt{\eta})$ [48].

Conclusion.– In this work we have established a general and fundamental lower bound for the error probability affecting the adaptive discrimination of two arbitrary quantum channels acting on a finite-dimensional Hilbert space. This bound is conveniently expressed in terms of the Choi matrices of the channels involved, so that it is very easy to compute. It also applies to many scenarios, including adaptive protocols for quantum-enhance optical resolution and quantum illumination. In order to derive our result, we have employed port-based teleportation as a tool for channel simulation, and developed a methodology which simplifies adaptive protocols performed over an arbitrary finite-dimensional channel. This technique can be applied to many other scenarios,

not just quantum hypothesis testing, but also quantum metrology, quantum and private communications. In particular, as we discuss in [48, Sec. V], our technique also allows us to show that adaptive protocols of quantum channel estimation are limited to the Heisenberg scaling, therefore drawing an intimate connection between port-based teleportation and quantum metrology.

Aknowledgements.– This work is supported by the EPSRC via the 'UK Quantum Communications Hub' (EP/M013472/1) and by the Innovation Fund Denmark (Qubiz project). We would like to thank S. Ishizaka, S. L. Braunstein, S. Lloyd, G. Spedalieri, A. Mari, L. Banchi, S. Bäuml, M. Wilde, C. Ottaviani, M. Hayashi, M. Rosati, L. Maccone, A. Harrow, S. Strelchuk, and Z.-W. Wang for comments and insightful questions.

[1] C. W. Helstrom, *Quantum Detection and Estimation Theory* (New York: Academic, 1976).
[2] A. Chefles, Contemp. Phys. **41**, 401 (2000).
[3] S. M. Barnett and S. Croke, Advances in Optics and Photonics **1**, 238-278 (2009).
[4] J. A. Bergou, U. Herzog, and M. Hillery, *Discrimination of Quantum States*, in Quantum state estimation, Lecture Notes in Physics, Springer-Verlag Berlin Heidelberg (2004), pp 417-465.
[5] Y. Sun, J. A. Bergou, and M. Hillery, Phys. Rev. A **66**, 032315 (2002).
[6] J. A. Bergou, and M. Hillery, Phys. Rev. Lett. **94**, 160501 (2005).
[7] J. A. Bergou, V. Bužek, E. Feldman, U. Herzog, and M. Hillery, Phys. Rev. A **73**, 062334 (2006).
[8] J. A. Bergou, E. Feldman, and M. Hillery, Phys. Rev. A **73**, 032107 (2006).
[9] M. Hayashi, *Quantum Information Theory: Mathematical Foundation* (Springer-Verlag Berlin Heidelberg, 2017).
[10] A. Holevo, *Quantum Systems, Channels, Information: A Mathematical Introduction* (De Gruyter, Berlin-Boston, 2012).
[11] M. A. Nielsen, and I. L. Chuang, *Quantum computation and quantum information* (Cambridge University Press, Cambridge, 2000).
[12] C. Weedbrook *et al.,* Rev. Mod. Phys. **84**, 621 (2012).
[13] K. M. R. Audenaert *et al.*, Phys. Rev. Lett. **98**, 160501 (2007).
[14] J. Calsamiglia, R. Munoz-Tapia, L. Masanes, A. Acin, and E. Bagan, Phys. Rev. A **77**, 032311 (2008).
[15] S. Pirandola, and S. Lloyd, Phys. Rev. A **78**, 012331 (2008).
[16] K. M. R. Audenaert, M. Nussbaum, A. Szkola, and F. Verstraete, Commun. Math. Phys. **279**, 251 (2008).
[17] G. Spedalieri, and S. L. Braunstein, Phys. Rev. A **90**, 052307 (2014).
[18] In this paper, we implicitly consider symmetric discrimination, i.e., with identical Bayesian costs.
[19] A. Acin, Phys. Rev. Lett. **87**, 177901 (2001).
[20] M. Sacchi, Phys. Rev. A **72**, 014305 (2005).
[21] G. Wang, and M. Ying, Phys. Rev. A **73**, 042 (2006).
[22] A. Childs, J. Preskill, and J. Renes, J. Mod. Opt. **47**, 155 (2000).
[23] C. Invernizzi, M. G. A. Paris, and S. Pirandola, Phys. Rev. A **84**, 022334 (2011).
[24] M. Hayashi, IEEE Trans. Inf. Theory **55**, 3807 (2009).
[25] S. Pirandola, and C. Lupo, Phys. Rev. Lett. **118**, 100502 (2017).
[26] R. Laurenza, C. Lupo, G. Spedalieri, S. L. Braunstein, and S. Pirandola, *Channel Simulation in Quantum Metrology*, arXiv:1712.06603 (2017).
[27] A. W. Harrow, A. Hassidim, D. W. Leung, and J. Watrous, Phys. Rev. A **81**, 032339 (2010).
[28] V. I. Paulsen, *Completely Bounded Maps and Operator Algebras* (Cambridge University Press, 2002).
[29] S. Ishizaka, and T. Hiroshima, Phys. Rev. Lett. **101**, 240501 (2008).
[30] S. Ishizaka, and T. Hiroshima, Phys. Rev. A **79**, 042306 (2009).
[31] S. Ishizaka, *Some remarks on port-based teleportation*, Preprint arXiv:1506.01555 (2015).
[32] S. Beigi, and R König, New J. Phys. **13**, 093036 (2011).
[33] Z.-W. Wang, and S. L. Braunstein, Sci. Rep. **6**, 33004 (2016).
[34] S. Strelchuk, M. Horodecki, and J. Oppenheim, Phys. Rev. Lett. **110**, 069902 (2013).
[35] M. Studziński, S. Strelchuk, M. Mozrzymas, and M. Horodecki, Sci Rep. **7**, 10871 (2017).
[36] M. Mozrzymas, M. Studziński, S. Strelchuk, and M. Horodecki, *Optimal Port-based Teleportation*, arXiv:1707.08456 (2017).
[37] S. Pirandola, R. Laurenza, C. Ottaviani, and L. Banchi, Nat. Commun. **8**, 15043 (2017). See also arXiv:1510.08863 (2015).
[38] S. Pirandola, S. L. Braunstein, R. Laurenza, C. Ottaviani, T. P. W. Cope, G. Spedalieri, and L. Banchi, *Theory of channel simulation and bounds for private communication*, arXiv:1711.09909 (2017).
[39] S. Lloyd, Science **321**, 1463 (2008).
[40] S.-H. Tan *et al.*, Phys. Rev. Lett. **101**, 253601 (2008).
[41] J. H. Shapiro, and S. Lloyd, New J. Phys. **11**, 063045 (2009).
[42] Z. Zhang, M. Tengner, T. Zhong, F. N. C. Wong, and J. H. Shapiro, Phys. Rev. Lett. **111**, 010501 (2013).
[43] E. D. Lopaeva, I. Ruo Berchera, I. P. Degiovanni, S. Olivares, G. Brida, and M. Genovese, Phys. Rev. Lett. **110**, 153603 (2013).
[44] Z. Zhang, S. Mouradian, F. N. C. Wong, and J. H. Shapiro, Phys. Rev. Lett. **114**, 10506 (2015).
[45] S. Barzanjeh *et al.*, Phys. Rev. Lett. **114**, 080503 (2015).
[46] C. Weedbrook, S. Pirandola, J. Thompson, V. Vedral, and M. Gu, New J. Phys. **18**, 043027 (2016).
[47] G. De Palma, and J. Borregaard, *The ultimate precision of quantum illumination*, arXiv:1802.02158 (2018).
[48] In the Supplementary Information we present the following: (Sec. I) Proof of Lemma 1; (Sec. II) Details of the propagation of the simulation error via the peeling argument; (Sec. III) Calculations related to the ultimate single-photon quantum optical resolution; (Sec. IV) Cal-

culations related to adaptive quantum illumination; (Sec. V) Results related to adaptive quantum channel estimation, including the proof of the Heisenberg limit as a consequence of PBT; (Sec VI) Converse bounds for adaptive quantum and private communications.

[49] In fact, recall that the diamond distance is defined as

$$||\mathcal{E}_0^{\otimes n} - \mathcal{E}_1^{\otimes n}||_\diamond := \sup_{\rho_{ar}} ||\mathcal{E}_0^{\otimes n} \otimes \mathcal{I}(\rho_{ar}) - \mathcal{E}_1^{\otimes n} \otimes \mathcal{I}(\rho_{ar})||,$$

where $\mathcal{I}$ is an identity map acting on a reference system r. The upper bound in Eq. (3) is achieved by a specific (non-adaptive) protocol, where an (optimal) input state ρ_{ar} is prepared and its a-parts transmitted through $\mathcal{E}_u^{\otimes n}$.

[50] Note that, if we allow the simulator $\mathcal{S}$ to include part of the channel $\mathcal{E}$, then we may simulate any channel [37]. This solution can be adopted for the tasks of quantum and private communication. However, in QCD and quantum metrology, we need to consider a universal (channel-independent) $\mathcal{S}$ in order to simplify adaptive protocols. From Ref. [51, 52], we know that single-program universal simulators cannot simulate any quantum channel, so that the use of multi-program versions must necessarily be considered.

[51] M. Christandl, and A. Müller-Hermes, Commun. Math. Phys. **353**, 821(2017).

[52] A. Muller-Hermes, *Transposition in quantum information theory* (Master's thesis, Technical University of Munich, 2012).

[53] Note that this resource state may be optimized in a suitable way to increase the performance of the protocol [31]. This is achieved by adding an operator on Alice's qudits before detection. Such an operator is generally non-trace preserving and, for this reason, it cannot be included in our derivations, where we exploit the monotonicity under trace-preserving QOs (data processing).

[54] The equiprobability of the outcomes and their independence from the input state can be understood by the fact that the scheme is invariant under permutation of the Bell states and, therefore, of the ports. For large M, this can also be understood by the fact that we may always feed a Bell state into the PBT and later use part of that state to teleport an arbitrary input via standard teleportation (whose correction unitaries can be done at the output, i.e., after the identity channel realized by PBT for large M).

[55] Explicitly, this channel takes the form $\Gamma_M(\rho_C) = \sum_{i=1}^M \mathrm{Tr}_{\mathbf{A}\bar{B}_i C}[\Pi^i_{C\mathbf{A}} \left(\rho_C \otimes \Phi^{\otimes M}_{\mathbf{AB}}\right)]$, where $\mathrm{Tr}_{\bar{B}_i}$ denotes the trace over all ports $\mathbf{B}$ but B_i.

[56] C. Majenz, *Entropy in Quantum Information Theory – Communication and Cryptography* (PhD thesis, University of Copenhagen, 2017).

[57] S. Pirandola, S. *Capacities of repeater-assisted quantum communications,* Preprint arXiv:1601.00966 (2016).

[58] R. Laurenza, and S. Pirandola, Phys. Rev. A **96**, 032318 (2017).

[59] R. Laurenza, S. L. Braunstein, and S. Pirandola, *Finite-resource teleportation stretching for continuous-variable systems*, arXiv:1706.06065 (2017).

[60] T. P. W. Cope, L. Hetzel, L. Banchi, and S. Pirandola, Phys. Rev. A **96**, 022323 (2017).

[61] T. P. W. Cope, and S. Pirandola, Quantum Measurements and Quantum Metrology **4**, 44-52 (2017).

[62] C. A. Fuchs, and J. van de Graaf, IEEE Trans. Inf. Theory **45,** 1216 (1999).

[63] This comes from the Fuchs-van de Graaf relations and the multiplicativity of the fidelity over tensor products. Other bounds that can be written are: $D \leq nM \|\rho_{\mathcal{E}_1} - \rho_{\mathcal{E}_2}\|$ from the subadditivity of the trace distance, and $D \leq \sqrt{nM(\ln\sqrt{2})\min\{S(\rho_{\mathcal{E}_1}||\rho_{\mathcal{E}_2}), S(\rho_{\mathcal{E}_2}||\rho_{\mathcal{E}_1})\}}$ from the Pinsker inequality [64, 65], where $S(\rho||\sigma) = \mathrm{Tr}[\rho(\log_2\rho - \log_2\sigma)]$ is the relative entropy [9, 11].

[64] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes* (San Francisco, Holden Day, 1964).

[65] E. A. Carlen and E. H. Lieb, Lett. Math. Phys. **101**, 1-11 (2012).

[66] In general, by setting $M = xnd(d-1)$ for some $x > 2$, we get $B \geq \frac{1}{2} - \frac{1}{x} - \frac{1}{2}\sqrt{1 - F^{2xd(d-1)n^2}}$.

[67] M. Tsang, R. Nair, and X.-M. Lu, Phys. Rev. X **6**, 031033 (2016).

[68] C. Lupo and S. Pirandola, Phys. Rev. Lett. **117**, 190802 (2016).

[69] R. Nair and M. Tsang, Phys. Rev. Lett. **117**, 190801 (2016).

[70] S. Pirandola, Phys. Rev. Lett. **106**, 090504 (2011).

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

[72] M. Dall'Arno, A. Bisio, G. M. D'Ariano, M. Miková, M. Ježek, and M. Dušek, Phys. Rev. A **85**, 012308 (2012).

[73] M. Dall'Arno, A. Bisio, and G. M. D'Ariano, Int. J. Quant. Inf. **10**, 1241010 (2012).

[74] G. Spedalieri, Entropy **17**, 2218-2227 (2015).

[75] S. L. Braunstein, and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).

[76] S. L. Braunstein, C. M. Caves, and G. J. Milburn, Ann. Phys. **247**, 135-173 (1996).

[77] M. G. A. Paris, Int. J. Quant. Inf. **7**, 125-137 (2009).

[78] V. Giovannetti, S. Lloyd, and L. Maccone, Nature Photon. **5**, 222 (2011).

[79] D. Braun *et al.*, Preprint arXiv:1701.05152 (2017).

[80] J. Doukas, G. Adesso, S. Pirandola, and A. Dragan, Class. Quantum Grav. **32**, 035013 (2015).

[81] S. Pirandola, R. Laurenza, and S. L. Braunstein, *Teleportation simulation of bosonic Gaussian channels: Strong and uniform convergence,* arXiv:1712.01615 (2017).

[82] J. H. Shapiro, IEEE J. Sel. Top. Quantum Electron. **15**, 1547 (2009).

[83] C. Lupo, V. Giovannetti, S. Pirandola, S. Mancini, and S. Lloyd, Phys. Rev. A **84**, 010303(R) (2011).

[84] C. Lupo, V. Giovannetti, S. Pirandola, S.Mancini, and S. Lloyd, Phys. Rev. A **85**, 062314 (2012).

[85] M. Hayashi, Commun. Math. Phys. **304**, 689 (2011).

[86] H. M. Wiseman, Australian Optical Society News **16**, 14-19 (2002).

[87] A. A. Berni *et al.*, Nature Photon. **9**, 577 (2015).

[88] K. Kravtsov *et al.*, Phys. Rev. A **87**, 062122 (2013).

[89] D. H. Mahler *et al.*, Phys. Rev. Lett. **111**, 183601 (2013).

[90] Z. Hou, H. Zhu, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, npj Quantum Information **2**, 16001 (2016).

[91] R. Demkowicz-Dobrzański and L. Maccone, Phys. Rev. Lett. **113**, 250801 (2014).

[92] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, Phys. Rev. Lett. **94**, 160502 (2005).

[93] V. Vedral, Rev. Mod. Phys. **74**, 197 (2002).

[94] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, Phys. Rev. Lett. **78**, 2275-2279 (1997).

[95] V. Vedral, and M. B. Plenio, Phys. Rev. A **57**, 1619

(1998).
[96] M. Christandl, *The Structure of Bipartite Quantum States: Insights from Group Theory and Cryptography* (PhD thesis, University of Cambridge, 2006).
[97] E. Kaur, and M. M. Wilde, J. of Phys. A **51**, 035303 (2018).

Supplementary Information

I. SIMULATION ERROR IN DIAMOND NORM (PROOF OF LEMMA 1)

As noted in Ref. [56], the channel Γ_M associated with the qudit PBT protocol of Ref. [29] is covariant under unitary transformations, i.e.,

$$\Gamma_M(U\rho U^\dagger) = U\Gamma_M(\rho)U^\dagger, \tag{S1}$$

for any input state ρ and unitary operator U. Ref. [56] has also shown that, for a channel with such a symmetry, the diamond distance with the identity map is saturated by a maximally entangled state, i.e.,

$$\|\mathcal{I} - \Gamma_M\|_\diamond = \| |\Phi\rangle\langle\Phi| - \mathcal{I}\otimes\Gamma_M\left(|\Phi\rangle\langle\Phi|\right)\|\,, \tag{S2}$$

where $|\Phi\rangle = d^{-1/2}\sum_{k=1}^d |k\rangle|k\rangle$. Here we also show that

$$\| |\Phi\rangle\langle\Phi| - \mathcal{I}\otimes\Gamma_M\left(|\Phi\rangle\langle\Phi|\right)\| = 2[1 - f_e(\Gamma_M)]\ . \tag{S3}$$

In fact, first note that the map $\Lambda_M = \mathcal{I}\otimes\Gamma_M$ is covariant under twirling unitaries of the form $U\otimes U^*$, i.e.,

$$\begin{aligned}&\Lambda_M\left[(U\otimes U^*)\rho(U\otimes U^*)^\dagger\right]\\ &= (U\otimes U^*)\Lambda_M(\rho)(U\otimes U^*)^\dagger,\end{aligned} \tag{S4}$$

for any input state ρ and unitary operator U. This implies that the state $\Lambda_M(|\Phi\rangle\langle\Phi|)$ is invariant under twirling unitaries, i.e.,

$$(U\otimes U^*)\Lambda_M(|\Phi\rangle\langle\Phi|)(U\otimes U^*)^\dagger = \Lambda_M(|\Phi\rangle\langle\Phi|)\ . \tag{S5}$$

This is therefore an isotropic state of the form

$$\Lambda_M(|\Phi\rangle\langle\Phi|) = p|\Phi\rangle\langle\Phi| + \frac{1-p}{d^2}\mathbb{I}, \tag{S6}$$

where $\mathbb{I}$ is the two-qudit identity operator. We may rewrite this state as follows

$$\Lambda_M(|\Phi\rangle\langle\Phi|) = F|\Phi\rangle\langle\Phi| + (1-F)\rho^\perp, \tag{S7}$$

where $\rho^\perp$ is state with support in the orthogonal complement of Φ, and F is the singlet fraction

$$F := \langle\Phi|\Lambda_M(|\Phi\rangle\langle\Phi|)|\Phi\rangle = p + (1-p)d^{-2}. \tag{S8}$$

Thanks to the decomposition in Eq. (S7) and using basic properties of the trace norm, we may then write

$$\begin{aligned}&\| |\Phi\rangle\langle\Phi| - \Lambda_M\left(|\Phi\rangle\langle\Phi|\right)\|\\ &= \|(1-F)|\Phi\rangle\langle\Phi| - (1-F)\rho^\perp\|\\ &= (1-F)\||\Phi\rangle\langle\Phi|\| + (1-F)\|\rho^\perp\|\\ &= 2(1-F)\\ &= 2[1 - f_e(\Gamma_M)],\end{aligned} \tag{S9}$$

where the last step exploits the fact that the singlet fraction F is the channel's entanglement fidelity $f_e(\Gamma_M)$.

Finally, it is known that the entanglement fidelity of Γ_M is bounded as [29]

$$f_e(\Gamma_M) \geq 1 - d(d-1)M^{-1}. \tag{S10}$$

Therefore, combining Eqs. (S2), (S3), and (S10), we derive

$$\|\mathcal{I} - \Gamma_M\|_\diamond \leq 2d(d-1)M^{-1}\ . \tag{S11}$$

II. PROPAGATION OF THE SIMULATION ERROR

This is given for the sake of completeness, we prove the first inequality in Eq. (9) of the main text. Consider the adaptive protocol described in the main text. For the n-use output state we may compactly write

$$\rho_n = \Lambda_n\circ\mathcal{E}\circ\Lambda_{n-1}\circ\cdots\circ\mathcal{E}\circ\Lambda_1\circ\mathcal{E}(\rho_0), \tag{S12}$$

where Λ's are adaptive QOs and $\mathcal{E}$ is the channel applied to the transmitted signal system. Then, ρ_0 is the preparation state of the registers, obtained by applying the first QO Λ_0 to some fundamental state. Similarly, for the M-port simulation of the protocol, we may write

$$\rho_n^M = \Lambda_n\circ\mathcal{E}^M\circ\Lambda_{n-1}\circ\cdots\circ\mathcal{E}^M\circ\Lambda_1\circ\mathcal{E}^M(\rho_0), \tag{S13}$$

where $\mathcal{E}^M$ is in the place of $\mathcal{E}$. (Note that the following reasoning applies to a fixed channel $\mathcal{E}$ or, more generally, to classically-parametrized unknown channel $\mathcal{E}_u$).

Consider now two instances ($n = 2$) of the adaptive protocol. We may bound the trace distance between ρ_2 and ρ_2^M using a "peeling" argument [25, 37, 38, 81]

$$\begin{aligned}\left\|\rho_2 - \rho_2^M\right\| &= \|\Lambda_2\circ\mathcal{E}\circ\Lambda_1\circ\mathcal{E}(\rho_0)\\ &\quad - \Lambda_2\circ\mathcal{E}^M\circ\Lambda_1\circ\mathcal{E}^M(\rho_0)||\\ &\overset{(1)}{\leq} ||\mathcal{E}\circ\Lambda_1\circ\mathcal{E}(\rho_0) - \mathcal{E}^M\circ\Lambda_1\circ\mathcal{E}^M(\rho_0)||\\ &\overset{(2)}{\leq} ||\mathcal{E}\circ\Lambda_1\circ\mathcal{E}(\rho_0) - \mathcal{E}\circ\Lambda_1\circ\mathcal{E}^M(\rho_0)||\\ &\quad + ||\mathcal{E}^M\circ\Lambda_1\circ\mathcal{E}(\rho_0) - \mathcal{E}^M\circ\Lambda_1\circ\mathcal{E}^M(\rho_0)||\\ &\overset{(3)}{\leq} ||\mathcal{E}(\rho_0) - \mathcal{E}^M(\rho_0)||\\ &\quad + ||\mathcal{E}[\Lambda_1\circ\mathcal{E}^M(\rho_0)] - \mathcal{E}^M[\Lambda_1\circ\mathcal{E}^M(\rho_0)]||\\ &\overset{(4)}{\leq} 2||\mathcal{E} - \mathcal{E}^M||_\diamond\ .\end{aligned} \tag{S14}$$

In (1) we use the monotonicity of the trace distance under completely-positive trace-preserving (CPTP) maps (i.e., quantum channels); in (2) we employ the triangle inequality; in (3) we use the monotonicity with respect to the the CPTP map $\mathcal{E}\circ\Lambda_1$ whereas in (4) we exploit the fact that the diamond norm is an upper bound for the trace norm computed on any input state. Generalizing the result of Eq. (S14) to arbitrary n, we achieve the first inequality in Eq. (9) of the main text.

III. ULTIMATE SINGLE-PHOTON QUANTUM OPTICAL RESOLUTION

Consider the problem of discriminating between the following situations:

(1) A point-like source emitting light from position $x = s/2$;

(2) A point-like source emitting light from the shifted position $x = -s/2$.

The discrimination is achieved by measuring the image created by a focusing optical system. More precisely, we consider a linear imaging system in the paraxial approximation that is used to image point-like sources. This is characterized by the Fresnel number

$$\mathcal{F} = \frac{\ell}{x_R}\,, \tag{S15}$$

where ℓ is the size of the object, and

$$x_R = \frac{\lambda}{N_A} \tag{S16}$$

is the Rayleigh length. Here λ is the wavelength and $N_A = R/D$ is the numerical aperture, where R is the radius of the pupil and D is the distance from the object. In the far-field regime, light is attenuated by a loss parameter $\eta \simeq \mathcal{F}$ [82–84]. In particular, because we consider point-like sources, we are in the regime $\eta \ll 1$.

First we need to model the imaging system as a quantum channel acting on the *input* state represented by the light emitted by the source. The two cases are described by the following Heisenberg-picture transformations on the input annihilation operator a

$$(1):\ a \to \sqrt{\eta} b_1 + \sqrt{1-\eta} v_1, \tag{S17}$$
$$(2):\ a \to \sqrt{\eta} b_2 + \sqrt{1-\eta} v_2, \tag{S18}$$

where $b_{1,2}$ are the output operators (encoding the position of the source) and $v_{1,2}$ are associated with a vacuum environment. The modes b_1, b_2 are defined on the image plane and have the form

$$b_j = \int dx\, \psi_j(x)\, a(x)\,, \tag{S19}$$

where $a(x), a(x)^\dagger$ is a continuous family of canonical operators $[a(x), a(y)^\dagger] = \delta(x-y)$ defined on the image plane (for simplicity, we assume unit magnification factor). In general the image modes b_1, b_2 do satisfy the (non-canonical) commutation relations

$$[b_1, b_2^\dagger] = \int dx\, \psi_1(x)\psi_2^*(x)\,, \tag{S20}$$

where ψ_j is the point-spread function associated to the source being in poisiton j. Then, by setting $\delta = \mathrm{Re}\int dx\, \psi_1(x)\psi_2^*(x)$, we can define the effective image operators

$$b_\pm := (b_1 \pm b_2)/\sqrt{2(1\pm\delta)}. \tag{S21}$$

The fact that $\delta \neq 0$ means that the two image fields overlap and the sources cannot be perfectly distinguished. This is a manifestation of diffraction through the finite objective of the optical imaging system.

As a result, we can write the action of the channels as

$$(1):\ a \to \sqrt{\eta_+} b_+ + \sqrt{\eta_-} b_- + \sqrt{1-\eta} v_1, \tag{S22}$$
$$(2):\ a \to \sqrt{\eta_+} b_+ - \sqrt{\eta_-} b_- + \sqrt{1-\eta} v_2, \tag{S23}$$

where $\eta_\pm := (1\pm\delta)\eta/2$. For simplicity, consider a single-photon state at the input. We then have

$$(1):\ |1\rangle \to \eta |\psi_+\rangle\langle\psi_+| + (1-\eta)|0\rangle\langle 0|\,, \tag{S24}$$
$$(2):\ |1\rangle \to \eta |\psi_-\rangle\langle\psi_-| + (1-\eta)|0\rangle\langle 0|\,, \tag{S25}$$

where

$$|\psi_\pm\rangle = \frac{\sqrt{\eta_+}|1\rangle_+ - \sqrt{\eta_-}|1\rangle_-}{\sqrt{\eta}}\,. \tag{S26}$$

More generally, the action of the channels on a generic pure input state is given by

$$(1):\ \alpha|0\rangle + \beta|1\rangle \to \left(|\alpha|^2 + \eta|\beta|^2\right) \times |\psi_+(\alpha,\beta)\rangle\langle\psi_+(\alpha,\beta)| + (1-\eta)|\beta|^2|0\rangle\langle 0|, \tag{S27}$$
$$(2):\ \alpha|0\rangle + \beta|1\rangle \to \left(|\alpha|^2 + \eta|\beta|^2\right) \times |\psi_-(\alpha,\beta)\rangle\langle\psi_-(\alpha,\beta)| + (1-\eta)|\beta|^2|0\rangle\langle 0|, \tag{S28}$$

where

$$|\psi_\pm(\alpha,\beta)\rangle = \frac{\alpha|0\rangle + \beta\sqrt{\eta_+}|1\rangle_+ \pm \beta\sqrt{\eta_-}|1\rangle_-}{\sqrt{|\alpha|^2+\eta|\beta|^2}}\,. \tag{S29}$$

As we can see from Eqs. (S27) and (S28), if we apply a Pauli operator X [11] to the input state $\alpha|0\rangle + \beta|1\rangle$, we have a swap between α and β. This leads to an output state with a different eigenspectrum, so that it cannot be obtained by applying a unitary. This means that the quantum channels are *not* teleportation-covariant.

By limiting ourselves to the space of either no photon or one photon $\mathcal{H}_2 = \mathrm{span}\{|0\rangle, |1\rangle\}$, the the input space of the channels is a qubit, and their output is a qutrit, so that the dimension of the input Hilbert space is $d = 2$. Apart from restricting the input space to qubits, we assume the most general adaptive strategy allowed by quantum mechanics, so that the quantum state of the source may be optimized as a consequence of the output (as generally happens in the adaptive protocol discussed in the main text). In order to compute the ultimate performance, we need to compute the quantum fidelity between the Choi matrices of the two channels in Eqs. (S22) and (S23) suitably truncated to $\mathcal{H}_2$.

Consider then the maximally entangled state $|\Phi_2\rangle = (|0\rangle|1\rangle + |1\rangle|0\rangle)/\sqrt{2}$. The Choi matrices associated with the two truncated channels are equal to

$$\rho_{(1)} = \frac{1+\eta}{2}|\Psi_+\rangle\langle\Psi_+| + \frac{1-\eta}{2}|0\rangle\langle 0|\,, \tag{S30}$$

$$\rho_{(2)} = \frac{1+\eta}{2}|\Psi_-\rangle\langle\Psi_-| + \frac{1-\eta}{2}|0\rangle\langle 0|\,, \tag{S31}$$

where

$$|\Psi_\pm\rangle = \frac{|0\rangle|1\rangle + \sqrt{\eta_+}|1\rangle_+|0\rangle \pm \sqrt{\eta_-}|1\rangle_-|0\rangle}{\sqrt{1+\eta}}. \tag{S32}$$

Notice that $\langle\Psi_+|\Psi_-\rangle = (1+\delta\eta)/(1+\eta)$ where $\delta\eta := \eta_+ - \eta_-$. Therefore we obtain the fidelity

$$F(\rho_{(1)},\rho_{(2)}) = \mathrm{Tr}\sqrt{\sqrt{\rho_{(1)}}\,\rho_{(2)}\,\sqrt{\rho_{(1)}}} \tag{S33}$$

$$= \frac{1+\eta}{2}\left|\frac{1+\delta\eta}{1+\eta}\right| + \frac{1-\eta}{2} = \frac{1-\eta+|1+\delta\eta|}{2}.$$

Assuming that δ is real, this becomes

$$F[\rho_{(1)},\rho_{(2)}] = 1 - \frac{\eta(1-\delta)}{2}, \tag{S34}$$

which allows us to identify $\epsilon = \eta(1-\delta)/2$. A common way to model diffraction is to consider a Gaussian point-spread function, i.e.

$$\psi_j(s) \simeq e^{-(x-x_j)^2/4}, \tag{S35}$$

where x_j is the center of the jth emitter, and the variance of the Gaussian is 1 in units of Rayleigh length. Under this Gaussian model one obtains [67, 68]

$$\delta \simeq e^{-s^2/8}, \tag{S36}$$

where s is the separation in unit of wavelength. Therefore

$$\epsilon \simeq \frac{\eta(1-e^{-s^2/8})}{2} \simeq \frac{\eta s^2}{16}. \tag{S37}$$

By replacing this quantity in Eq. (17) of the main text with $d=2$ we obtain the lower bound

$$B \gtrsim \frac{1}{4}\exp\left(-2ns\sqrt{\eta}\right). \tag{S38}$$

IV. ULTIMATE LIMIT OF ADAPTIVE QUANTUM ILLUMINATION

A. Standard (non-adaptive) protocol

In quantum illumination [39, 40, 45, 46], we aim at determining whether a low-reflectivity object is present or not in a region with thermal noise. We therefore prepare a signal system s and an idler system i in a joint entangled state ρ_{si}. The signal system is sent to probe the target while the idler system is retained for its measurement together with the potential signal reflection from the target. If the object is absent, the "reflected" system is just thermal background noise. If the object is present, then this is composed of the actual reflection of the signal from the target plus thermal background noise. This object can be modelled by a beam splitter, with very small transmissivity $\eta \ll 1$, which combines the each incoming optical mode (signal system) with a thermal mode with b mean number of photons.

In the discrete-variable version of quantum illumination [39], the signal system is prepared in an ensemble of d optical modes, with 1 photon in one of the modes and vacuum in the others. This is the number of modes which are distinguished by the detector in each detection process. If we introduce the following $d-$dimensional computational basis

$$|1\rangle := \overbrace{|00\ldots01\rangle}^{d}, \tag{S39}$$

$$|2\rangle := |00\ldots10\rangle\,, \tag{S40}$$

$$\vdots$$

$$|d-1\rangle := |01\ldots00\rangle\,, \tag{S41}$$

$$|d\rangle := |10\ldots00\rangle\,, \tag{S42}$$

then the entangled signal-idler state can be written as

$$\psi_{si} = |\psi\rangle_{si}\langle\psi|\,, \quad |\psi\rangle_{si} = d^{-1/2}\sum_{k=1}^{d}|kk\rangle_{si}\,. \tag{S43}$$

Let us define the d-dimensional identity operator $\mathbb{I}^d := \sum_{k=1}^{d}|k\rangle\langle k|$ which projects onto the subspace spanned by the 1-photon states, and the $(d+1)$-dimensional identity operator $\mathbb{I}^{d+1} := \sum_{k=0}^{d}|k\rangle\langle k|$ which also includes the vacuum state $|0\rangle := |00\ldots00\rangle$. Then, we have the reduced idler state

$$\psi_i := \mathrm{Tr}_s(\psi_{si}) = d^{-1}\mathbb{I}_i^d, \tag{S44}$$

and we define the thermal state of the environment as [39]

$$\rho^{\mathrm{th}}(b) := (1-db)|0\rangle\langle 0| + b\mathbb{I}^d, \tag{S45}$$

where b is the mean number of thermal photons per mode. Here $b \ll 1$ and $db \ll 1$, where db is the mean number of thermal photons in each detection event.

The output $(d+1)\times d$ state of the reflected signal and retained idler is given by

$$\begin{array}{rl} \text{Target absent:} & \sigma = \rho^{\mathrm{th}}(b)\otimes d^{-1}\mathbb{I}_i^d, \\ \text{Target present:} & \rho = (1-\eta)\sigma + \eta\psi_{si}. \end{array} \tag{S46}$$

If the target is probed n times, then we may use the QCB to bound Q the error probability p_{err} in the discrimination of ρ and σ. In the regime of signal-to-noise-ratio $\eta d/b \lesssim 1$, one finds [39]

$$Q = 1 - \frac{\eta^2 d}{8b} + \mathcal{O}(b^2,\eta b)\,, \tag{S47}$$

which tightens the QCB by a factor d with respect to the unentangled case where $Q \approx 1-\eta^2/(8b)$. From Eq. (S47), we may write the following bound for the error probability of target detection after n probings [39]

$$p_n(\sigma \neq \rho) \leq \frac{1}{2}\exp\left(-\frac{\eta^2 dn}{8b}\right). \tag{S48}$$

In particular, for $\eta d/b \simeq 1$, this can be written as

$$p_n(\sigma \neq \rho) \leq \frac{1}{2}\exp\left(-\frac{\eta n}{8}\right). \tag{S49}$$

B. Adaptive protocol

The adaptive formulation of the discrete variable protocol of quantum illumination assumes an unlimited quantum computer with two register **a** and **b**, prepared in an arbitrary joint quantum state. In each probing, a system a is picked from the input register **a** and sent to the target. Its reflection a' is stored in the output register **b**. A adaptive quantum operation (QO) is applied to both the update registers before the next transmission and so on. Therefore any probing is interleaved by the application of adaptive QOs Λ's to the registers, defining the adaptive protocol $\mathcal{P}_n$ (see also the main text for this description). After n probings, the state of the registers is $\rho_n(u)$ where $u = 0, 1$ is a bit encoding the absence or presence of the target. This state is optimally measured by an Helstrom POVM. By optimizing over all protocol $\mathcal{P}_n$, we define the minimum error probability p_n for adaptive quantum illumination.

Following the constraints and typical regime of DV quantum illumination, we assume that the signal systems are $(d+1)$-dimensional qudits described by a basis $\{|0\rangle, |1\rangle, \ldots, |d\rangle\}$, where $|i\rangle := |0\cdots 010\cdots 0\rangle$ has one photon in the ith mode. For this reason, the two possible quantum illumination channels, $\mathcal{E}_0$ and $\mathcal{E}_1$, are $(d+1)$-dimensional channels. In particular, consider as their input the maximally-entangled state

$$\Psi_{si} = \frac{1}{d+1}\sum_{k,j=0}^{d} |kk\rangle_{si}\langle jj|, \tag{S50}$$

which is similar to ψ_{si} in Eq. (S43) but also includes the vacuum state. Then, we may write the following two $(d+1)\times(d+1)$ dimensional Choi matrices

$$\begin{aligned} &\text{Target absent:} \quad \sigma := \rho_{\mathcal{E}_0} = \rho^{\text{th}}(b)\otimes(d+1)^{-1}\mathbb{I}_i^{d+1}, \\ &\text{Target present:} \quad \rho := \rho_{\mathcal{E}_1} = (1-\eta)\sigma + \eta\Psi_{si}. \end{aligned} \tag{S51}$$

It is clear that $\mathcal{E}_0$ and $\mathcal{E}_1$ are not jointly teleportation-covariant due to the fact that they have different transmissivities ($\eta_0 = 0$ and $\eta_1 = \eta$).

To bound p_n we apply Theorem 2 of the main text and, more specifically, Eq. (17) of the main text, because $\eta \ll 1$ and, therefore, the fidelity between the Choi matrices can be expanded as $F(\sigma,\rho) \simeq 1-\epsilon$. Thus, let us start by computing this fidelity. Let us set $x = \sqrt{1-bd}$ and note that we may write

$$\sqrt{\sigma} = (x|0\rangle_s\langle 0| + \sqrt{b}\mathbb{I}_s^d)\otimes(d+1)^{-1/2}\mathbb{I}_i^{d+1}. \tag{S52}$$

Then, we may compute

$$\begin{aligned} \Omega^2 &:= \sqrt{\sigma}\rho\sqrt{\sigma} \\ &= \frac{1}{(d+1)^2}\left\{(1-\eta)\left[x^4|0\rangle_s\langle 0| + b^2\mathbb{I}_s^d\right]\otimes\mathbb{I}_i^{d+1}\right. \\ &+ \eta\left[x^2|00\rangle_{si}\langle 00| + \sqrt{b}x\sum_{k=1}^{d}\left(|00\rangle_{si}\langle kk| + |kk\rangle_{si}\langle 00|\right)\right. \\ &\left.\left. + b\sum_{j,k=1}^{d}|kk\rangle_{si}\langle jj|\right]\right\}. \end{aligned} \tag{S53}$$

One can check that Ω^2 has d^2 degenerate eigenvalues equal to $b^2(d+1)^{-2}$, d degenerate eigenvalues equal to $(1-\eta)x^4(d+1)^{-2}$, and other $d+1$ eigenvalues $\{\lambda_i\}$ given by the diagonalization of the matrix $(d+1)^{-2}\mathbf{M}$ where

$$\mathbf{M} = \begin{pmatrix} (1-\eta)x^4+\eta x^2 & \eta x\sqrt{b} & \eta x\sqrt{b} & \cdots & \eta x\sqrt{b} \\ \eta x\sqrt{b} & b(b+\eta) & \eta b & \cdots & \eta b \\ \eta x\sqrt{b} & \eta b & b(b+\eta) & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \eta b \\ \eta x\sqrt{b} & \eta b & \cdots & \eta b & b(b+\eta) \end{pmatrix}. \tag{S54}$$

Once we find the eigenvalues of Ω^2 we take their square root so as to compute those of Ω. Finally, their sum provides $\text{Tr}\Omega = F(\sigma,\rho)$. We are interested in the regime of low thermal noise $b \ll 1$ and low reflectivity $\eta \ll 1$. There, we may expand at the leading orders in η and b to get

$$F(\sigma,\rho) = 1 - \frac{\eta d + 2b - 2\sqrt{\eta d b}}{2(d+1)} + \mathcal{O}(\eta^2, \eta^{3/2}b^{1/2}, \eta b, b^{3/2}) \tag{S55}$$

$$= 1 - \frac{\eta d}{2(d+1)} + \mathcal{O}(\eta^2, \sqrt{\eta b}, b). \tag{S56}$$

In the typical signal-to-noise-ratio $\eta d/b \simeq 1$ of quantum illumination [39], we may directly re-write Eq. (S55) as $F(\sigma,\rho) \simeq 1-\epsilon$, where

$$\epsilon := \frac{\eta d + 2b - 2\sqrt{d\eta b}}{2(d+1)} \simeq \frac{d\eta}{2(d+1)} < \eta/2. \tag{S57}$$

By replacing the latter in Eq. (17) of the main text (and assuming the correct dimension $d \to d+1$), we get the following lower bound for the minimum error probability p_n of adaptive quantum illumination

$$p_n \geq \frac{1}{4}\exp(-4nd\sqrt{\eta}). \tag{S58}$$

V. ADAPTIVE QUANTUM CHANNEL ESTIMATION

A. Adaptive protocols for parameter estimation

Consider here adaptive quantum channel estimation. This corresponds to estimating a continuous parameter θ encoded in a quantum channel $\mathcal{E}_\theta$ by means of the most general protocols allowed by quantum mechanics, i.e., based on adaptive QOs as described in the main text. After n probings, there is a θ-dependent output state $\rho_n(\theta)$ which is generated by the sequence of QOs $\{\Lambda_0, \Lambda_1, \ldots, \Lambda_n\}$ characterizing the adaptive protocol $\mathcal{P}$. Finally, the output state is measured by a POVM $\mathcal{M}$ providing an optimal unbiased estimator $\tilde{\theta}$ of parameter θ. The minimum error variance $\mathrm{Var}(\tilde{\theta}) := \langle(\tilde{\theta}-\theta)^2\rangle$ must satisfy the quantum Cramer-Rao bound (QCRB) [75] $\mathrm{Var}(\tilde{\theta}) \geq (\mathrm{QFI}_\theta^n)^{-1}$, where QFI_θ^n is the quantum Fisher information (QFI) associated with n adaptive uses.

Note that the QFI can be computed as

$$\mathrm{QFI}_\theta^n = \frac{4d_B^2[\rho_n(\theta), \rho_n(\theta+d\theta)]}{d\theta^2}, \tag{S59}$$

where $d_B(\rho,\sigma) := \sqrt{2[1-F(\rho,\sigma)]}$ is the Bures distance, with $F(\rho,\sigma)$ being the Bures fidelity of ρ and σ. The ultimate precision of adaptive quantum metrology is given by optimizing the QFI over all adaptive protocols, i.e.,

$$\overline{\mathrm{QFI}}_\theta^n := \sup_{\mathcal{P}} \mathrm{QFI}_\theta^n(\mathcal{P}). \tag{S60}$$

Contrarily to the cases of sequential or parallel strategies [85], the ultimate performance of adaptive quantum metrology is poorly studied [86–90], with limited results for DV programmable channels [25, 91], and more recent results for DV and CV teleportation-covariant channels, such as Pauli or Gaussian channels [25].

B. PBT stretching of adaptive quantum metrology

As shown in Ref. [25], the adaptive estimation of a noise parameter θ encoded in a teleportation-covariant channel (i.e., such that the parametrized class of channels $\mathcal{E}_\theta$ is jointly-teleportation covariant) is limited to the standard quantum limit (SQL). More generally, as discussed in Ref. [26], the adaptive estimation of a parameter in a quantum channel cannot beat the SQL if the channel has a single-copy simulation, i.e., of the type

$$\mathcal{E}_\theta(\rho) = \mathcal{S}(\rho \otimes \pi_\theta), \tag{S61}$$

where $\mathcal{S}$ is a (parameter-independent) trace-preserving QO and π_θ is a program state (depending on the parameter). To beat the SQL, the channel should not admit a simulation as in Eq. (S61) but a multi-copy version

$$\mathcal{E}_\theta(\rho) = \mathcal{S}(\rho \otimes \pi_\theta^{\otimes M}), \tag{S62}$$

for some $M > 1$. This is approximately the type of simulation that we can achieve by using PBT.

First of all, we may replace the channel $\mathcal{E}_\theta$ with its M-port approximation $\mathcal{E}_\theta^M := \mathcal{E}_\theta \circ \Gamma_M$, where Γ_M is the M-port PBT channel. Using Lemma 1 of the main text, the simulation error may be bounded as

$$||\mathcal{E}_\theta - \mathcal{E}_\theta^M||_\diamond \leq \delta_M := ||\mathcal{I} - \Gamma_M||_\diamond \leq 2\beta M^{-1}, \tag{S63}$$

where we set $\beta := d(d-1)$. By repeating the steps shown in Fig. 1 of the main text, we may write the metrological equivalent of Eq. (7). In other words, for any input state ρ_C, we may write the simulation

$$\mathcal{E}_\theta^M(\rho_C) = \mathcal{T}^M(\rho_C \otimes \rho_{\mathcal{E}_\theta}^{\otimes M}), \tag{S64}$$

where $\mathcal{T}^M$ is a trace-preserving LOCC and $\rho_{\mathcal{E}_\theta}$ is the Choi matrix of $\mathcal{E}_\theta$. Then, we may also repeat the PBT stretching in Fig. 2 of the main text. In this way, the n-use output state $\rho_n = \rho_n(\theta)$ of an adaptive parameter estimation protocol can be decomposed as in Lemma 2 of the main text, i.e.,

$$||\rho_n(\theta) - \bar{\Lambda}(\rho_{\mathcal{E}_\theta}^{\otimes nM})|| \leq n\delta_M. \tag{S65}$$

C. PBT implies the Heisenberg scaling

Using the decomposition in Eq. (S65), we may write a bound for the optimal quantum Fisher information in Eq. (S60). For large n, we obtain the Heisenberg scaling

$$\overline{\mathrm{QFI}}_\theta^n \lesssim n^2 \mathrm{QFI}(\rho_{\mathcal{E}_\theta}), \tag{S66}$$

where

$$\mathrm{QFI}(\rho_{\mathcal{E}_\theta}) = \frac{4d_B^2(\rho_{\mathcal{E}_\theta}, \rho_{\mathcal{E}_{\theta+d\theta}})}{d\theta^2}. \tag{S67}$$

In order to show Eq. (S66), consider the function

$$q_n(\theta,\delta) = 2\frac{d_B[\rho_n(\theta), \rho_n(\theta+\delta)]}{\delta}. \tag{S68}$$

We set $u_\theta := \bar{\Lambda}(\rho_{\mathcal{E}_\theta}^{\otimes nM})$ and apply twice the triangular inequality, so that we may write

$$\begin{aligned} d_B[\rho_n(\theta), \rho_n(\theta+\delta)] \leq\ & d_B[\rho_n(\theta), u_\theta] + \\ & d_B[u_\theta, u_{\theta+\delta}] + d_B[u_{\theta+\delta}, \rho_n(\theta+\delta)]. \end{aligned} \tag{S69}$$

Bounding the Bures distance with the trace distance, we get

$$d_B^2[\rho_n(\theta), u_\theta] \leq \frac{\|\rho_n(\theta) - u_\theta\|}{2} \leq \frac{n\delta_M}{2} \leq \frac{\beta n}{M}. \tag{S70}$$

Using Eqs. (S69) and (S70), we may write

$$q_n(\theta,\delta) \leq 2\frac{d_B[u_\theta, u_{\theta+\delta}]}{\delta} + \frac{4}{\delta}\sqrt{\frac{\beta n}{M}}. \tag{S71}$$

We may bound d_B in Eq. (S71) as follows

$$\begin{aligned} d_B[u_\theta, u_{\theta+\delta}] &\overset{(1)}{\leq} d_B[\rho_{\mathcal{E}_\theta}^{\otimes nM}, \rho_{\mathcal{E}_{\theta+\delta}}^{\otimes nM}] \\ &\overset{(2)}{=} \sqrt{2[1-F(\rho_{\mathcal{E}_\theta}^{\otimes nM}, \rho_{\mathcal{E}_{\theta+\delta}}^{\otimes nM})]} \\ &\overset{(3)}{=} \sqrt{2(1-F^{nM})} \overset{(4)}{\leq} \sqrt{2nM(1-F)} \\ &\overset{(2)}{=} \sqrt{nM} d_B[\rho_{\mathcal{E}_\theta}, \rho_{\mathcal{E}_{\theta+\delta}}], \end{aligned} \tag{S72}$$

where: (1) we use the monotonicity of the Bures distance under the CPTP map $\bar{\Lambda}$, (2) we use the standard relation between Bures distance and fidelity, (3) we set $F := F(\rho_{\mathcal{E}_\theta}, \rho_{\mathcal{E}_{\theta+\delta}})$ and exploit the multiplicativity of the fidelity over tensor products, and (4) we use the inequality $F^n \geq 1-n+nF$. Therefore, from Eq. (S71), we may derive the inequality

$$q_n(\theta,\delta) \leq 2\sqrt{nM}\frac{d_B[\rho_{\mathcal{E}_\theta}, \rho_{\mathcal{E}_{\theta+\delta}}]}{\delta} + \frac{4}{\delta}\sqrt{\frac{\beta n}{M}}. \tag{S73}$$

Now notice that

$$\lim_{\delta\to 0} 2\frac{d_B[\rho_{\mathcal{E}_\theta}, \rho_{\mathcal{E}_{\theta+\delta}}]}{\delta} = \sqrt{\mathrm{QFI}(\rho_{\mathcal{E}_\theta})}. \tag{S74}$$

This means that for any $\epsilon > 0$, there is $\delta < \delta_\epsilon$ such that

$$q_n(\theta,\delta) \leq \sqrt{nM}\left[\sqrt{\mathrm{QFI}(\rho_{\mathcal{E}_\theta})} + \epsilon\right] + \frac{4}{\delta}\sqrt{\frac{\beta n}{M}}. \tag{S75}$$

Setting $M = n^{1+z}$ (for any $z > 0$) implies

$$\begin{aligned} q_n(\theta,\delta) &\leq \kappa_n(\theta,\delta|\epsilon,z) \\ &:= \sqrt{n^{2+z}}\left[\sqrt{\mathrm{QFI}(\rho_{\mathcal{E}_\theta})} + \epsilon\right] + \frac{4}{\delta}\sqrt{\frac{\beta}{n^z}}. \end{aligned} \tag{S76}$$

Note that, by definition, $\mathrm{QFI}_\theta^n := \lim_{\delta\to 0} q_n(\theta,\delta)^2$. Then, assume that the limit

$$\lim_{n\to\infty}\lim_{\delta\to 0}\frac{q_n(\theta,\delta)^2}{n^{2+z}} \tag{S77}$$

exists for any $z > 0$. Then, using Eq. (S76), which is valid for any n and δ, we may write

$$\begin{aligned} \lim_{n\to\infty}\lim_{\delta\to 0}\frac{q_n(\theta,\delta)}{\sqrt{n^{2+z}}} &\leq \liminf_{n\to\infty,\ \delta\to 0}\frac{\kappa_n(\theta,\delta|\epsilon,z)}{\sqrt{n^{2+z}}} \\ &\leq \sqrt{\mathrm{QFI}(\rho_{\mathcal{E}_\theta})} + \epsilon. \end{aligned} \tag{S78}$$

The previous inequality leads to

$$\lim_{n\to\infty}\frac{\mathrm{QFI}_\theta^n}{n^{2+z}} \leq \left[\sqrt{\mathrm{QFI}(\rho_{\mathcal{E}_\theta})} + \epsilon\right]^2, \tag{S79}$$

for any $\epsilon, z > 0$. Now, sending ϵ and z to zero gives the following scaling for large n

$$\mathrm{QFI}_\theta^n \lesssim n^2 \mathrm{QFI}(\rho_{\mathcal{E}_\theta})\ . \tag{S80}$$

Since this upper bound holds for any protocol P (because $\bar{\Lambda}$ disappears), then the asymptotic scaling in Eq. (S80) may be extended to $\overline{\mathrm{QFI}}_\theta^n$ as in Eq. (S66). In conclusion we have obtained un upper bound for the quantum Fisher information corresponding to the Heisenberg (quadratic) scaling in the number of uses.

VI. CONVERSE BOUNDS FOR ADAPTIVE PRIVATE COMMUNICATION

A. Adaptive protocols for quantum/private communication

Let us assume that the adaptive protocol described in the main text has the task of secret key generation, i.e., to establish a secret key between the register **a**, owned by Alice, and the register **b**, owned by Bob. This protocol employs adaptive LOCCs Λ_i interleaved with the transmissions over a d-dimensional quantum channel $\mathcal{E}$. (In this analysis we assume input and output Hilbert spaces with the same dimension d; if the spaces have different dimensions, we may always pad the one with the lower dimension and formally enlarge the channel to include the extra dimensions.) After n adaptive uses of the channel, the output state ρ_n of the registers is epsilon-close to a target private state [92] ϕ_n with nR_n private bits, i.e., $\|\rho_n - \phi_n\| \leq \epsilon$. Optimizing over all the key-generation adaptive protocols $\mathcal{P}$ and taking the limit for large n defines the secret key capacity of the channel $\mathcal{E}$

$$K(\mathcal{E}) := \sup_{\mathcal{P}} \lim_n R_n\ . \tag{S81}$$

It is known that this capacity is greater than other two-way assisted capacities. In fact, we have [38]

$$Q_2(\mathcal{E}) = D_2(\mathcal{E}) \leq P_2(\mathcal{E}) \leq K(\mathcal{E}), \tag{S82}$$

where Q_2 is the two-way assisted quantum capacity (qubits per channel use), D_2 is the two-way assisted entanglement distribution capacity (ebits per channel use), and P_2 is the two-way assisted private capacity (private bits per channel use). We now investigate upper bounds for $K(\mathcal{E})$ which are derived by combining PBT stretching with various entanglement measures, therefore extending one of the main insights of Ref. [37].

B. PBT stretching of private communication and single-letter upper bounds

Consider the M-port approximation $\mathcal{E}^M$ of $\mathcal{E}$, as achieved by the PBT simulation with error δ_M. Correspondingly, we have an M-port approximate output state ρ_n^M such that $\left\|\rho_n - \rho_n^M\right\| \leq n\delta_M$ as in Eq. (9) of the main text. Then, we may stretch an adaptive protocol $\mathcal{P}$ over $\mathcal{E}^M$ and write $\rho_n^M = \bar{\Lambda}(\rho_{\mathcal{E}}^{\otimes nM})$ for a trace-preserving LOCC $\bar{\Lambda}$. Using the triangle inequality, we may write

$$\begin{aligned} \left\|\rho_n^M - \phi_n\right\| &\leq \left\|\rho_n^M - \rho_n\right\| + \|\rho_n - \phi_n\| \\ &\leq n\delta_M + \epsilon := \gamma. \end{aligned} \tag{S83}$$

Now consider an entanglement measure E with the properties listed in Ref. [38, Sec. VIII]. For instance, E may be the relative entropy of entanglement E_{R} (REE) [93–95] or the squashed entanglement E_{SE}

(SE) [96]. In particular, these measures satisfy a suitable continuity property. For d-dimensional states ρ and σ such that $\|\rho-\sigma\| \leq \gamma$, we may write the Fannes-type inequality

$$|E(\rho)-E(\sigma)| \leq g(\gamma)\log_2 d + h(\gamma), \tag{S84}$$

where g, h are regular functions going to zero in ϵ'. For the REE and the SE, these functions are [38]

$$\text{REE:}\;\; g(\gamma)=4\gamma,\;\; h(\epsilon)=2H_2(\gamma), \tag{S85}$$

$$\text{SE:}\;\; g(\gamma)=16\sqrt{\gamma},\;\; h(\gamma)=2H_2(2\sqrt{\gamma}), \tag{S86}$$

where H_2 is the binary Shannon entropy.

By applying Eq. (S84) to Eq. (S83), we get

$$\left|E(\rho_n^M)-E(\phi_n)\right| \leq g(\gamma)\log_2 d + h(\gamma), \tag{S87}$$

where $E(\phi_n) \geq nR_n$ (normalization) and

$$E(\rho_n^M) = E[\bar{\Lambda}(\rho_{\mathcal{E}}^{\otimes nM})] \leq nM\; E(\rho_{\mathcal{E}}), \tag{S88}$$

which exploits the monotonicity of E under trace-preserving LOCCs and the subadditivity over tensor-product states [38]. Therefore, we may write

$$R_n \leq M\; E(\rho_{\mathcal{E}}) + \frac{g(n\delta_M+\epsilon)\log_2 d + h(n\delta_M+\epsilon)}{n}. \tag{S89}$$

Note that for a private state, we may write $\log_2 d \leq cn$ for some constant c [38]. Thus, for any adaptive key generation protocol $\mathcal{P}$ over a d-dimensional quantum channel $\mathcal{E}$, the maximum ϵ-secure key rate that can be generated after n uses is bounded as in Eq. (S89) where E is an entanglement measure (as the REE or the SE), M is the number of ports, and δ_M is the error of the M-port PBT defined in Eq. (5).

We can find alternate bound by extending the definition of channel's REE [37] to a tripartite version. Consider three finite-dimensional systems a', a and b', and a quantum channel $\mathcal{E}=\mathcal{E}_{a\to b}$ from a to the output system b. Consider a generic input state $\rho_{a'ab'}$ transformed into an output state $\omega_{a'bb'} := \mathcal{E}_{a\to b}(\rho_{a'ab'})$ by the action of this channel. Then, one can define a tripartite version of channel's REE as

$$\tilde{E}_{\mathrm{R}}(\mathcal{E}) := \sup_{\rho_{a'ab'}} E_{\mathrm{R}}(a'|bb')_{\omega} - E_{\mathrm{R}}(a'a|b)_{\rho}, \tag{S90}$$

which satisfies $K(\mathcal{E}) \leq \tilde{E}_{\mathrm{R}}(\mathcal{E})$ [97]. Moreover, if two channels are close in diamond norm $\|\mathcal{E}-\mathcal{E}'\|_{\diamond} \leq 2\epsilon$, then one may also write the continuity property [97]

$$|\tilde{E}_{\mathrm{R}}(\mathcal{E})-\tilde{E}_{\mathrm{R}}(\mathcal{E}')| \leq 2\epsilon\log_2 d + f(\epsilon), \tag{S91}$$

$$f(\epsilon) := (1+\epsilon)\log_2(1+\epsilon) - \epsilon\log_2\epsilon, \tag{S92}$$

where d is the dimension of the Hilbert space. Finally, as a straightforward application of one of the tools established in Ref. [37], i.e., the LOCC simulation of a quantum channel $\mathcal{E}$ via a resource state σ [38], one may write the data-processing upper bound $\tilde{E}_{\mathrm{R}}(\mathcal{E}) \leq E_{\mathrm{R}}(\sigma)$.

In our channel simulation via PBT, we have a multi-copy resource state $\sigma=\rho_{\mathcal{E}}^{\otimes M}$ for the M-port approximation $\mathcal{E}^M$ of the d-dimensional channel $\mathcal{E}$. This means that we may write

$$\tilde{E}_{\mathrm{R}}(\mathcal{E}^M) \leq E_{\mathrm{R}}(\rho_{\mathcal{E}}^{\otimes M}) \leq ME_{\mathrm{R}}(\rho_{\mathcal{E}}). \tag{S93}$$

Then, because we have

$$||\mathcal{E}-\mathcal{E}^M||_{\diamond} \leq ||\mathcal{I}-\Gamma_M||_{\diamond} := \delta_M \leq 2d(d-1)M^{-1}, \tag{S94}$$

from Eq. (S91) we may derive

$$\tilde{E}_{\mathrm{R}}(\mathcal{E}) \leq E_{\mathrm{R}}(\rho_{\mathcal{E}}^{\otimes M}) + \delta_M\log_2 d + f(\delta_M/2). \tag{S95}$$

As a result, we may write the upper bound

$$\begin{aligned} K(\mathcal{E}) &\leq E_{\mathrm{R}}(\rho_{\mathcal{E}}^{\otimes M}) + \delta_M\log_2 d + f(\delta_M/2) \\ &\leq ME_{\mathrm{R}}(\rho_{\mathcal{E}}) + \frac{2d(d-1)}{M}\log_2 d + f\left[\frac{d(d-1)}{M}\right] \\ &:= K_{\mathrm{UB}}^M(\mathcal{E}). \end{aligned} \tag{S96}$$

The tightest upper bound is obtained by minimizing $K_{\mathrm{UB}}^M(\mathcal{E})$ over M, which is typically a finite value.

Let us apply the bound to channels that are nearly entanglement-breaking, so that $E_{\mathrm{R}}(\rho_{\mathcal{E}}) \ll 1$. In this case, we expect that the optimal value of M is large. It is easy to see that a sub-optimal choice for M is given by

$$\tilde{M} = \sqrt{\frac{2d(d-1)\log_2 d}{E_{\mathrm{R}}(\rho_{\mathcal{E}})}}, \tag{S97}$$

which provides the upper bound

$$\begin{aligned} K(\mathcal{E}) \leq\; & 2\sqrt{2d(d-1)\log_2 d}\sqrt{E_{\mathrm{R}}(\rho_{\mathcal{E}})} \\ & + f\left[\sqrt{\frac{d(d-1)E_{\mathrm{R}}(\rho_{\mathcal{E}})}{2\log_2 d}}\right]. \end{aligned} \tag{S98}$$

The bound in Eq. (S98) is non-trivial only for almost entanglement-breaking channels, such that $E_{\mathrm{R}}(\rho_{\mathcal{E}}) \lesssim (\log_2 d)/[8d(d-1)]$. In this regime, it is not clear how tight this upper bound may be in comparison with other existing results.