This is a repository copy of *Fundamental limits to quantum channel discrimination*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/147001/

Version: Accepted Version

**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

---

**Reuse**
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:
https://creativecommons.org/licenses/

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
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, 30] 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 $E$ 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 $a$ and an output register $b$, which are prepared in an initial state $\rho_0$ by applying a QO $\Lambda_0$ to some fundamental state of $a$ and $b$. Then, a system $a_1$ is picked from the register $a$ and sent through the channel $E$. The corresponding output $b_1$ is merged with the output register $b_1b\rightarrow b$. This is followed by another QO $\Lambda_1$ applied to $a$ and $b$. Then, we send a second system $a_2 \in a$ through $E$ with the output $b_2$ being merged again $b_2b \rightarrow b$ and so on. After $n$ uses, the registers will be in a state $\rho_n$ which depends on $E$ and the sequence of QOs $\{\Lambda_0,\Lambda_1,\ldots,\Lambda_n\}$ defining the adaptive protocol $P_n$.

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 \[|\psi\rangle_{\text{target}}\] with \[\rho_{\text{target}}\] using two entangled states, so that we have the global state \[\Phi_{\text{AB}}^{\otimes M} = \bigotimes_{i=1}^{M} |\Phi_i\rangle_i\langle\Phi_i|, \quad |\Phi_i\rangle := d^{-1/2} \sum_{k} |k\rangle_{A_i} \otimes |k\rangle_{B_i}.\] To teleport the state of a qudit \(C\), Alice performs a joint measurement on \(C\) and her ensemble \(A\). This is a POVM \(\{\Pi_{CA_i}\}_{i=1}^{M}\) 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 \(i\)th 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 \(\delta_e\). Averaging over the outcomes, we define the teleported state \(\rho_B^{\otimes M} = \Gamma_M(\rho_C)\), where \(\Gamma_M\) is the corresponding PBT channel \(52\). In the limit of many ports \(M\), we have that \(\Gamma_M\) approaches an identity channel \(I\) 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 \(\Gamma_M\) and the identity channel \(I\) satisfies

\[
\delta_M := ||\Gamma - \Gamma_M||_\diamondleq 2d(d-1)M^{-1}.
\]

**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 \(\rho_C\) is subject to the approximate channel

\[
\mathcal{E}^M(\rho_C) := \mathcal{E} \circ \Gamma_M(\rho_C).
\]

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 \(M\) qudits before selecting the output port, as shown in Fig. 1c. This leads to the following simulation for the approximate channel

\[
\mathcal{E}^M(\rho_C) = \mathcal{T}^M(\rho_C \otimes \rho_{\mathcal{E}}^{\otimes M}),
\]

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

\[
||\mathcal{E} - \mathcal{E}^M||_\diamondleq \delta_M,
\]

where \(\delta_M\) 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 $\Phi_{AB}^\otimes$. To teleport an input qubit state $\rho_C$, Alice applies a suitable POVM $\{\Pi_i\}$ to the input qubit $C$ and her $A$ qubits. The outcome $i$ is communicated to Bob, who selects the $i$-th among his $B$ qubits (tracing all the others). The performance does not depend on the specific “port” $i$ selected and the average output state is given by $\Gamma_M$ where $\Gamma_M$ is the PBT channel. The latter reduces to the identity channel in the limit of many ports $M \rightarrow \infty$. (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 $\mathcal{E}^M \rightarrow \mathcal{E}$ in diamond norm. (c) Equivalently, Bob can apply $\mathcal{E}^\otimes_M$ to all his qubits $B$ 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 $\mathcal{M}$ of Eq. (7). For large $M$, one has $\mathcal{M} \rightarrow \mathcal{E}$ in diamond norm, 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 $\rho_E^\otimes_M$, iterate the process for all $n$ uses, and collapse all the simulation LOCCs and QOs as shown in Eq. (7). As a result, we may write the multi-copy Choi decomposition

$$\rho_n^M = \Lambda(\rho_E^\otimes_n) \, ,$$

for a trace-preserving QO $\Lambda$. 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 $\rho_n$
of the protocol can be decomposed as follows

\[ \| \rho_n - \tilde{\Lambda}(\rho_E^{\otimes nM}) \| \leq n\delta_M, \quad (11) \]

where \( \tilde{\Lambda} \) is a trace-preserving QO, \( \rho_E \) is the Choi matrix of \( E \), and \( \delta_M \) is the \( M \)-port simulation error in Eq. (9).

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

\[ \| \rho_n(u) - \tilde{\Lambda}(\rho_E^{\otimes nM}) \| \leq n\delta_M. \quad (12) \]

Ultimate bound for channel discrimination. - We are now ready to show the lower bound for minimum error probability \( p_n(E_0 \neq E_1) \) in Eq. (3). Consider an arbitrary protocol \( \mathcal{P}_n \), for which we may write Eq. (11). Combining Lemma 2 with the triangle inequality leads to

\[ \| \rho_n(0) - \rho_n(1) \| \leq 2n\delta_M + |\tilde{\Lambda}(\rho_E^{\otimes nM}) - \tilde{\Lambda}(\rho_E^{\otimes nM})| \]

\[ \leq 2n\delta_M + |\rho_E^{\otimes nM} - \rho_{E_1}^{\otimes nM}|, \quad (13) \]

where we also use the monotonicity of the trace distance under channels. Because \( \tilde{\Lambda} \) 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. (11) and (9), we get the following.

**Theorem 3** Consider the adaptive discrimination of two channels \( \{ \rho_{n} \} \) in dimension \( d \). After \( n \) probings, the minimum error probability satisfies the bound

\[ p_n(E_0 \neq E_1) \geq B := \frac{1}{2} - n\delta_M - D(\rho_E^{\otimes nM}, \rho_{E_1}^{\otimes nM})/2, \quad (14) \]

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

Let us bound the trace distance in Eq. (14) as \[ D^2 \leq 1 - F^{2nM}, \quad F := \text{Tr} \sqrt{\rho_E \rho_{E_1} \rho_E \rho_{E_1}}, \quad (15) \]

where \( F \) is the fidelity [63] between the Choi matrices of the channels. If we also exploit Eq. (16), we may write

\[ B \geq \frac{1}{2} - \frac{\sqrt{1 - F^{2nM}}}{2} \cdot \frac{d(d-1)n}{M}. \quad (16) \]

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 [64] is \( M = 4d(d-1)n \), so that

\[ B \geq \frac{1}{2} - 1 - 2\sqrt{1 - F^{2nM}}/4. \]

In particular, consider two infinitesimally-close channels, so that \( F \simeq 1 - \epsilon \) where \( \epsilon \simeq 0 \). By expanding in \( \epsilon \), we may write

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

Discriminating between two close quantum channels is a problem in many physical scenarios. For instance, this is typical in quantum optical resolution (discussed below), quantum illumination (discussed below), ideal quantum reading (treated in Ref. [48]), quantum metrology [75–77], quantum 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.

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 \( \epsilon \approx n\eta^2/16 \), where \( \eta \) 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 \( B \simeq 1/4 \exp(-2\eta s/\eta) \).

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 \( \eta \approx 0 \) in a thermal background which has \( b \approx 1 \) 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 \( |i\rangle := |0 \cdots 010 \cdots 0 \rangle \) has one photon in the \( i \)-th 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 \( B \simeq 1/4 \exp(-4n\eta d/s) \).

**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.
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.

Acknowledgements.– 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. Calzamiglia, 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, 060502 (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:1711.09009 (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-
In fact, recall that the diamond distance is defined as
\[ ||E_0 \otimes \epsilon - E_1 \otimes \epsilon||_0 := \sup_{\rho_{ar}} ||E_0 \otimes I(\rho_{ar}) - E_1 \otimes I(\rho_{ar})||, \]
where \( I \) is an identity map acting on a reference system \( r \). The upper bound in Eq. \( 4 \) is achieved by a specific (non-adaptive) protocol, where an (optimal) input state \( \rho_{ar} \) is prepared and its \( a \)-parts transmitted through \( E_a \otimes \epsilon \).

Note that, if we allow the simulator \( S \) to include part of the channel \( E \), then we may simulate any channel \( E \). 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) \( S \) 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.

Explicitly, this channel takes the form \( \Gamma = E_a \otimes \epsilon \), where \( E_a \) denotes the trace over all ports \( B \) but \( B_a \).

Note that this resource state may be optimized in a suitable way to increase the performance of the protocol \[53\]. 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).

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 \( M \)).

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: \( D \leq nM \sqrt{||\rho_1 - \rho_2||} \) from the subadditivity of the trace distance, and
\[ D \leq \sqrt{nM(\ln 2) \min(S(\rho_1||\rho_2), S(\rho_2||\rho_1))}, \]
from the Pinsker inequality \[61, 65\], where \( S(\rho||\sigma) = \text{Tr}[\rho \log_2 \rho - \log_2 \sigma] \) is the relative entropy \[6, 11\].
Supplementary Information

I. SIMULATION ERROR IN DIAMOND NORM

(Proof of Lemma 1)

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

$$\Gamma_M(U\rho U^\dagger) = U\Gamma_M(\rho)U^\dagger,$$  \hspace{1cm} (S1)

for any input state $\rho$ 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.,

$$\|I - \Gamma_M\|_d = \|\Phi\langle\Phi| - I \otimes \Gamma_M (|\Phi\rangle\langle\Phi|)\|,$$  \hspace{1cm} (S2)

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

$$\|\Phi\langle\Phi| - I \otimes \Gamma_M (|\Phi\rangle\langle\Phi|)\| = 2[1 - f_e(\Gamma_M)].$$  \hspace{1cm} (S3)

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

$$\Lambda_M [(U \otimes U^*)\rho(U \otimes U^*)^\dagger] = (U \otimes U^*)\Lambda_M(\rho)(U \otimes U^*)^\dagger, \hspace{1cm} (S4)$$

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

$$(U \otimes U^*)\Lambda_M(|\Phi\rangle\langle\Phi|)(U \otimes U^*)^\dagger = \Lambda_M(|\Phi\rangle\langle\Phi|).$$  \hspace{1cm} (S5)

This is therefore an isotropic state of the form

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

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

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

where $p^\perp$ is state with support in the orthogonal complement of $|\Phi\rangle$, and $F$ is the singlet fraction

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

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

$$\|\Phi\langle\Phi| - \Lambda_M(|\Phi\rangle\langle\Phi|)\| = \|(1-F)|\Phi\rangle\langle\Phi| - (1-F)p^\perp\| = (1-F)\|\Phi\rangle\langle\Phi\| + (1-F)\|p^\perp\| = 2(1-F) = 2[1 - f_e(\Gamma_M)], \hspace{1cm} (S9)$$

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 $\Gamma_M$ is bounded as

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

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

$$\|I - \Gamma_M\|_d \leq 2d(d - 1)M^{-1}. \hspace{1cm} (S11)$$

II. PROPAGATION OF THE SIMULATION ERROR

This is given for the sake of completeness, we prove the first inequality in Eq. (3) of the main text. Consider the adaptive protocol described in the main text. For the $n$-use output state we may compactly write

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

where $\Lambda$’s are adaptive QOs and $\mathcal{E}$ is the channel applied to the transmitted signal system. Then, $\rho_0$ is the preparation state of the registers, obtained by applying the first QO $\Lambda_0$ to some fundamental state. Similarly, for the $M$-port simulation of the protocol, we may write

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

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 $\mathcal{E}_n$).

Consider now two instances ($n = 2$) of the adaptive protocol. We may bound the trace distance between $\rho_2$ and $\rho_2^M$ using a “peeling” argument [25, 32, 35, 81]

$$\|\rho_2 - \rho_2^M\| = \|\Lambda_2 \circ \mathcal{E} \circ \Lambda_1 \circ \mathcal{E}(\rho_0) - \Lambda_2 \circ \mathcal{E}^M \circ \Lambda_1 \circ \mathcal{E}^M(\rho_0)\| \leq \|\mathcal{E} \circ \Lambda_1 \circ \mathcal{E}(\rho_0) - \mathcal{E}^M \circ \Lambda_1 \circ \mathcal{E}^M(\rho_0)\| \leq \|\mathcal{E} \circ \Lambda_1 \circ \mathcal{E}(\rho_0) - \mathcal{E} \circ \Lambda_1 \circ \mathcal{E}^M(\rho_0)\| + \|\mathcal{E}^M \circ \Lambda_1 \circ \mathcal{E}(\rho_0) - \mathcal{E}^M \circ \Lambda_1 \circ \mathcal{E}^M(\rho_0)\| \leq \|\mathcal{E}(\rho_0) - \mathcal{E}^M(\rho_0)\| + \|\mathcal{E} \circ \Lambda_1 \circ \mathcal{E}^M(\rho_0) - \mathcal{E}^M \circ \Lambda_1 \circ \mathcal{E}^M(\rho_0)\| \leq 2\|\mathcal{E} - \mathcal{E}^M\|_\infty. \hspace{1cm} (S14)$$

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. (3) 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

\[
\mathcal{F} = \frac{\ell}{x_R},
\]

where \( \ell \) is the size of the object, and

\[
x_R = \frac{\lambda}{N_A}
\]

is the Rayleigh length. Here \( \lambda \) 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 \( \eta \approx \mathcal{F} \) [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 \( a \)

\[
\begin{align*}
(1) : & \quad a \rightarrow \sqrt{\eta} b_1 + \sqrt{1 - \eta} v_1, \\
(2) : & \quad a \rightarrow \sqrt{\eta} b_2 + \sqrt{1 - \eta} v_2,
\end{align*}
\]

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

\[
b_j = \int dx \psi_j(x) a(x),
\]

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

\[
[b_1, b_2^\dagger] = \int dx \psi_1(x) \psi_2^*(x),
\]

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

\[
b_{\pm} := (b_1 \pm b_2)/\sqrt{1 + \delta}.
\]

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

\[
\begin{align*}
(1) : & \quad a \rightarrow \sqrt{\eta} b_+ + \sqrt{1 - \eta} v_1, \\
(2) : & \quad a \rightarrow \sqrt{\eta} b_- + \sqrt{1 - \eta} v_2,
\end{align*}
\]

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

\[
\begin{align*}
(1) : & \quad |1\rangle \rightarrow \eta |\psi_+\rangle \langle \psi_+| + (1 - \eta) |0\rangle \langle 0|, \\
(2) : & \quad |1\rangle \rightarrow \eta |\psi_-\rangle \langle \psi_-| + (1 - \eta) |0\rangle \langle 0|,
\end{align*}
\]

where

\[
|\psi_{\pm}(\alpha, \beta)\rangle = \frac{\sqrt{\eta} |\psi_+(1)\rangle \pm \beta \sqrt{1 - \eta} |\psi_-(1)\rangle}{\sqrt{\eta}},
\]

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

\[
\begin{align*}
(1) : & \quad \alpha |0\rangle + \beta |1\rangle \rightarrow (|\alpha|^2 + \eta |\beta|^2) \times |\psi_+\rangle \langle \psi_+| + (1 - \eta) |\beta|^2 |0\rangle \langle 0|, \\
(2) : & \quad \alpha |0\rangle + \beta |1\rangle \rightarrow (|\alpha|^2 + \eta |\beta|^2) \times |\psi_-\rangle \langle \psi_-| + (1 - \eta) |\beta|^2 |0\rangle \langle 0|,
\end{align*}
\]

where

\[
|\psi_{\pm}(\alpha, \beta)\rangle = \frac{\alpha |0\rangle + \beta \sqrt{\eta} |\psi_+(1)\rangle \pm \beta \sqrt{1 - \eta} |\psi_-(1)\rangle}{\sqrt{|\alpha|^2 + \eta |\beta|^2}}.
\]

As we can see from Eqs. (S27) and (S28), if we apply a Pauli operator \( X \) to the input state \( \alpha |0\rangle + \beta |1\rangle \), we have a swap between \( \alpha \) and \( \beta \). 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 \( \mathcal{H}_2 = \text{span}\{|0\rangle, |1\rangle\} \), 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

$$
\rho_{(1)} = \frac{1 + \eta}{2} |\Psi_+\rangle\langle\Psi_+| + \frac{1 - \eta}{2} |0\rangle\langle0|, \\
\rho_{(2)} = \frac{1 + \eta}{2} |\Psi_-\rangle\langle\Psi_-| + \frac{1 - \eta}{2} |0\rangle\langle0|,
$$

where

$$
|\Psi_\pm\rangle = \frac{|0\rangle|1\rangle + \sqrt{\eta} |1\rangle|0\rangle \mp \sqrt{1 - \eta} |1\rangle|0\rangle}{\sqrt{1 + \eta}}.
$$

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

$$
F(\rho_{(1)}, \rho_{(2)}) = \text{Tr} \sqrt{\rho_{(1)} \rho_{(2)} \rho_{(1)}^{\dagger} \rho_{(2)}^{\dagger}}
$$

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

Assuming that $\delta$ is real, this becomes

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

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.

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

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

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

where $s$ is the separation in unit of wavelength. Therefore

$$
\epsilon \simeq \eta e^{-s^2/8} \lesssim \eta^2 2^{-16}.
$$

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

$$
B \gtrsim \frac{1}{4} \exp \left(-2n s \sqrt{\eta} \right).
$$

IV. ULTIMATE LIMIT OF ADAPTIVE QUANTUM ILLUMINATION

A. Standard (non-adaptive) protocol

In quantum illumination [39, 43, 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 $\rho_{si}$. 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

$$
d^{|1\rangle := |00\rangle \ldots |01\rangle},
$$

$$
d^{|2\rangle := |00\rangle \ldots |11\rangle},
$$

$$
\vdots
$$

$$
d^{|d\rangle := |01\rangle \ldots |10\rangle},
$$

then the entangled signal-idler state can be written as

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

Let us define the $d$-dimensional identity operator $I^d := \sum_{k=1}^d |k\rangle \langle k|$ which projects onto the subspace spanned by the 1-photon states, and the $(d+1)$-dimensional identity operator $I^{d+1} := \sum_{k=0}^d |k\rangle \langle k|$ which also includes the vacuum state $|0\rangle := |00\ldots0\rangle$. Then, we have the reduced idler state

$$
\rho_i := \text{Tr}_s(\psi_{si}) = d^{-1} I_i^d,
$$

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

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

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

$$
\text{Target absent: } \sigma = \rho^{th}(b) \otimes d^{-1} I_i^d,
$$

$$
\text{Target present: } \rho = (1 - \eta) \sigma + \eta \psi_{si}.
$$

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 $\rho$ and $\sigma$. In the regime of signal-to-noise-ratio $\eta d/b \lesssim 1$, one finds [39]

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

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

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

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

### 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 QO’s $\Lambda$’s to the registers, defining the adaptive protocol $P_n$ (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 $P_n$, 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 $\{|0\}, |1\}, \ldots, |d\}$, where $|i\rangle := |0 \cdots 010 \cdots 0\rangle$ has one photon in the $i$th 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

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

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

Target absent: $\sigma := \rho_{\mathcal{E}_0} = \rho^{ib}(b) \otimes (d + 1)^{-1}\mathbb{I}_{d+1}$,

Target present: $\rho := \rho_{\mathcal{E}_1} = (1 - \eta)\sigma + \eta\Psi_{si}. \tag{S51}$

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 $x = \sqrt{1 - \eta d}$ and note that we may write

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

Then, we may compute

$$\Omega^2 := \sqrt{\sigma} \rho \sqrt{\sigma}$$

$$= \frac{1}{(d + 1)^2} \{ (1 - \eta) [x^4|0\rangle_s|0\rangle + b^2\psi_d^x] \otimes \mathbb{I}_{d+1}$$

$$+ \eta \left[ x^2|00\rangle_s|00\rangle + \sqrt{b}x \sum_{k=1}^d (|00\rangle_s|kk\rangle + |kk\rangle_s|00\rangle)$$

$$+ b \sum_{j,k=1}^d |kk\rangle_s|jj\rangle \right] \}.$$

One can check that $\Omega^2$ 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 $(d + 1)^{-2}M$ where

$$M = \begin{pmatrix}
(1 - \eta)x^4 + \eta x^2 & \eta x \sqrt{b} & \cdots & \eta x \sqrt{b} \\
\eta x \sqrt{b} & b(b + \eta) & \eta b & \cdots & \eta b \\
\eta x \sqrt{b} & \eta b & b(b + \eta) & \cdots & \cdots & \eta b \\
\eta x \sqrt{b} & \cdots & \cdots & \cdots & \cdots & \eta b \\
\eta x \sqrt{b} & \cdots & \cdots & \cdots & \cdots & \eta b \\
\eta x \sqrt{b} & \cdots & \cdots & \cdots & \cdots & \eta b \\
\eta x \sqrt{b} & \cdots & \cdots & \cdots & \cdots & \eta b
\end{pmatrix}. \tag{S53}$$

Once we find the eigenvalues of $\Omega^2$ we take their square root so as to compute those of $\Omega$. 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 $\eta$ and $b$ to get

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

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

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

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

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

$$p_n \geq \frac{1}{4} \exp(-4n \sqrt{\eta}). \tag{S58}$$
V. ADAPTIVE QUANTUM CHANNEL ESTIMATION

A. Adaptive protocols for parameter estimation

Consider here adaptive quantum channel estimation. This corresponds to estimating a continuous parameter \( \theta \) 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 \( \theta \)-dependent output state \( \rho_n(\theta) \) which is generated by the sequence of QOs \( \{A_0, A_1, \ldots, A_n\} \) characterizing the adaptive protocol \( \mathcal{P} \).

Finally, the output state is measured by a POVM \( \mathcal{M} \) providing an optimal unbiased estimator \( \hat{\theta} \) of parameter \( \theta \). The minimum error variance \( \text{Var}(\hat{\theta}) := \langle (\hat{\theta} - \theta)^2 \rangle \) must satisfy the quantum Cramer-Rao bound (QCRB) \[ \text{Var}(\hat{\theta}) \geq (\text{QFI}^a)^{-1}, \]
where \( \text{QFI}^a \) is the quantum Fisher information (QFI) associated with \( n \) adaptive uses.

Note that the QFI can be computed as
\[
\text{QFI}^a = \frac{4d_B^2[\rho_n(\theta), \rho_n(\theta + d\theta)]}{d\theta^2},
\]
where \( d_B(\rho, \sigma) := \sqrt{2[1 - F(\rho, \sigma)]} \) is the Bures distance, with \( F(\rho, \sigma) \) being the Bures fidelity of \( \rho \) and \( \sigma \). The ultimate precision of adaptive quantum metrology is given by optimizing the QFI over all adaptive protocols, i.e.,
\[
\text{QFI}^a := \sup \text{QFI}^a(\mathcal{P}).
\]

Contrarily to the cases of sequential or parallel strategies, the ultimate performance of adaptive quantum metrology is poorly studied, 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 \( \theta \) 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. \[ 20 \], 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
\[
\mathcal{E}_\theta(\rho) = \mathcal{S}(\rho \otimes \pi_\theta),
\]
where \( \mathcal{S} \) is a (parameter-independent) trace-preserving QO and \( \pi_\theta \) is a program state (depending on the parameter). To beat the SQL, the channel should not admit a simulation as in Eq. \[ \text{(S61)} \] but a multi-copy version
\[
\mathcal{E}_\theta(\rho) = \mathcal{S}(\rho \otimes \pi_\theta^\otimes M),
\]
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 \( \Gamma_M \) is the \( M \)-port PBT channel. Using Lemma 1 of the main text, the simulation error may be bounded as
\[
\|\mathcal{E}_\theta - \mathcal{E}_\theta^M\|_\mathcal{M} \leq \|I - \Gamma_M\|_\mathcal{M} \leq 2\beta M^{-1},
\]
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. \( \text{(7)} \). In other words, for any input state \( \rho_C \), we may write the simulation
\[
\mathcal{E}_\theta^M(\rho_C) = \mathcal{T}^M(\rho_C \otimes \rho_{\mathcal{E}_\theta}^\otimes),
\]
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.,
\[
\|\rho_n(\theta) - \hat{\Lambda}(\rho_{\mathcal{E}_\theta}^\otimes nM)\|_\mathcal{M} \leq n\delta_M.
\]

C. PBT implies the Heisenberg scaling

Using the decomposition in Eq. \[ \text{(S65)} \], we may write a bound for the optimal quantum Fisher information in Eq. \[ \text{(S60)} \]. For large \( n \), we obtain the Heisenberg scaling
\[
\text{QFI}^a \lesssim n^2 \text{QFI}(\rho_{\mathcal{E}_\theta}),
\]
where
\[
\text{QFI}(\rho_{\mathcal{E}_\theta}) = \frac{4d_B^2(\rho_{\mathcal{E}_\theta}, \rho_{\mathcal{E}_\theta^\otimes n})}{d\theta^2}.
\]

In order to show Eq. \[ \text{(S66)} \], consider the function
\[
q_n(\theta, \delta) = \frac{2d_B[\rho_n(\theta), \rho_n(\theta + \delta)]}{\delta}.
\]
We set \( u_\theta := \Lambda(\rho_{\mathcal{E}_\theta}^\otimes_M) \) and apply twice the triangular inequality, so that we may write
\[
d_B[\rho_n(\theta), \rho_n(\theta + \delta)] \leq d_B[\rho_n(\theta), u_\theta] + d_B[u_\theta, u_{\theta + \delta}] + d_B[u_{\theta + \delta}, \rho_n(\theta + \delta)].
\]
Bounding the Bures distance with the trace distance, we get
\[
d_B^2[\rho_n(\theta), u_\theta] \leq \frac{\|\rho_n(\theta) - u_\theta\|}{\delta} \leq \frac{n\delta_M}{\delta} \leq \frac{\beta n}{M}.
\]
Using Eqs. \[ \text{(S69)} \] and \[ \text{(S70)} \], we may write
\[
q_n(\theta, \delta) \leq 2\frac{d_B[u_\theta, u_{\theta + \delta}]}{\delta} + 4\sqrt{\frac{\beta n}{M}}.
\]
We may bound $d_B$ in Eq. (S71) as follows
\[ d_B[u_\theta, u_{\theta+\delta}] \leq d_B[\rho_{\theta}^{\otimes nM}, \rho_{\theta+\delta}^{\otimes nM}] \leq \sqrt{2[1 - F(\rho_{\theta}^{\otimes nM}, \rho_{\theta+\delta}^{\otimes nM})]} \leq \sqrt{2[1 - F^n M(1 - F^n)]} \leq \sqrt{2(1 - F^n M)} \leq \sqrt{2nM(1 - F^n)}. \] (S72)

where: (1) we use the monotonicity of the Bures distance under the CPTP map $\Lambda$, (2) we use the standard relation between Bures distance and fidelity, (3) we set $F := F(\rho_{\theta}^{\otimes nM}, \rho_{\theta+\delta}^{\otimes nM})$ 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
\[ q_n(\theta, \delta) \leq 2\sqrt{nM} \left[ \frac{d_B[\rho_{\theta}^{\otimes nM}, \rho_{\theta+\delta}^{\otimes nM}]}{\delta} \right] + \frac{4}{\delta} \sqrt{\frac{3n}{M}}. \] (S73)

Now notice that
\[ \lim_{\delta \to 0} 2 \frac{d_B[\rho_{\theta}^{\otimes nM}, \rho_{\theta+\delta}^{\otimes nM}]}{\delta} = \sqrt{\text{QFI}(\rho_{\theta})}. \] (S74)
This means that for any $\epsilon > 0$, there is $\delta < \delta_\epsilon$ such that
\[ q_n(\theta, \delta) \leq \sqrt{nM} \left[ \sqrt{\text{QFI}(\rho_{\theta})} + \epsilon \right] + \frac{4}{\delta} \sqrt{\frac{3n}{M}}. \] (S75)

Setting $M = n^{1+z}$ (for any $z > 0$) implies
\[ q_n(\theta, \delta) \leq \kappa_n(\theta, \delta, \epsilon, z) \] (S76)
where
\[ \kappa_n(\theta, \delta, \epsilon, z) := \sqrt{n^{2+z}} \left[ \sqrt{\text{QFI}(\rho_{\theta})} + \epsilon \right] + \frac{4}{\delta} \sqrt{\frac{3n}{n^z}}. \]
Note that, by definition, $\text{QFI}_n^{\theta} := \lim_{\delta \to 0} q_n(\theta, \delta)^2$. Then, assume that the limit
\[ \lim_{n \to \infty} \lim_{\delta \to 0} q_n(\theta, \delta)^2 \] (S77)
exists for any $z > 0$. Then, using Eq. (S76), which is valid for any $n$ and $\delta$, we may write
\[ \lim_{n \to \infty} \lim_{\delta \to 0} q_n(\theta, \delta)^2 \leq \lim_{n \to \infty} \inf_{\delta \to 0} \frac{\kappa_n(\theta, \delta, \epsilon, z)}{\sqrt{n^{2+z}}} \leq \sqrt{\text{QFI}(\rho_{\theta})} + \epsilon. \] (S78)

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 $\Lambda_i$ 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 $\rho_n$ of the registers is epsilon-close to a target private state $|\phi\rangle$ with $nR$ private bits, i.e., $\|\rho_n - |\phi\rangle\| \leq \epsilon$. 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 $\mathcal{E}$
\[ K(\mathcal{E}) := \sup_{\mathcal{P}} \lim_{n \to \infty} R_n. \] (S80)

It is known that this capacity is greater than other two-way assisted capacities. In fact, we have [35]
\[ Q_2(\mathcal{E}) = D_2(\mathcal{E}) \leq Q_2(\mathcal{E}) \leq K(\mathcal{E}), \] (S81)
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 $\delta_M$. Correspondingly, we have an $M$-port approximate output state $\rho_{\mathcal{E}}^M$ such that $\|\rho_n - \rho_n^M\| \leq n\delta_M$ as in Eq. (9) of the main text. Then, we may stretch an adaptive protocol $\mathcal{P}$ over $\mathcal{E}^M$ and write $\rho_n^M = \Lambda(\hat{\rho}_{\otimes nM}^M)$ for a trace-preserving LOCC $\Lambda$. Using the triangle inequality, we may write
\[ \|\rho_n^M - |\phi\rangle\| \leq \|\rho_n^M - \rho_n\| + \|\rho_n - |\phi\rangle\| \leq n\delta_M + \epsilon := \gamma. \] (S82)

Now consider an entanglement measure $E$ with the properties listed in Ref. [35, Sec. VIII]. For instance, $E$ may be the relative entropy of entanglement $E_{RE}$ [92, 95] or the squashed entanglement $E_{SE}$
In particular, these measures satisfy a suitable continuity property. For $d$-dimensional states $\rho$ and $\sigma$ such that $\|\rho - \sigma\| \leq \gamma$, we may write the Fannes-type inequality

$$|E(\rho) - E(\sigma)| \leq g(\gamma) \log_2 d + h(\gamma),$$  \hspace{1cm} (S84)

where $g$, $h$ are regular functions going to zero in $\epsilon'$. For the REE and the SE, these functions are $\text{38}$. Therefore, we may write

$$E(\rho) \geq nR_n (\text{normalization})$$

where $H_2$ is the binary Shannon entropy.

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

$$|E(\rho^M_n) - E(\phi_n)| \leq g(\gamma) \log_2 d + h(\gamma),$$  \hspace{1cm} (S87)

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

$$E(\rho^M_n) = E[\Lambda(\rho^M_n)] \leq nM \log_2 d + h(\gamma),$$  \hspace{1cm} (S88)

which exploits the monotonicity of $E$ under trace-preserving LOCCs and the subadditivity over tensor-product states $\text{38}$. Therefore, we may write

$$R_n \leq M \log_2 d + h(\gamma).$$  \hspace{1cm} (S89)

Note that for a private state, we may write $\log_2 d \leq cn$ for some constant $c$ $\text{38}$. Thus, for any adaptive key generation protocol $\mathcal{P}$ over a $d$-dimensional quantum channel $\mathcal{E}$, the maximum $\epsilon$-secure key rate that can be generated after $n$ uses is bounded as in Eq. (S89) where $E$ is an entanglement measure (as in the REE or the SE), $M$ is the number of ports, and $\delta_M$ 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 $\text{37}$ to a tripartite version. Consider three finite-dimensional systems $a'$, $a$ and $b'$, and a quantum channel $\mathcal{E} = \mathcal{E}_{a \rightarrow b}$ from $a$ to the output system $b$. Consider a generic input state $\rho_{a'ab'}$ transformed into an output state $\omega_{a'b'} : = \mathcal{E}_{a \rightarrow b}(\rho_{a'ab'})$ by the action of this channel. Then, one can define a tripartite version of channel’s REE as

$$\tilde{E}_R(\mathcal{E}) := \sup_{\rho_{a'ab'}} E_R(a'|bb')_\omega - E_R(a'|a)b_\rho,$$  \hspace{1cm} (S90)

which satisfies $K(\mathcal{E}) \leq \tilde{E}_R(\mathcal{E})$ $\text{97}$. 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 $\text{97}$

$$|\tilde{E}_R(\mathcal{E}) - \tilde{E}_R(\mathcal{E}')| \leq 2\epsilon \log_2 d + f(\epsilon),$$  \hspace{1cm} (S91)

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

where $d$ is the dimension of the Hilbert space. Finally, as a straightforward application of one of the tools established in Ref. $\text{97}$, i.e., the LOCC simulation of a quantum channel $\mathcal{E}$ via a resource state $\sigma$ $\text{38}$, one may write the data-processing upper bound $\tilde{E}_R(\mathcal{E}) \leq E_R(\sigma)$.

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

$$\tilde{E}_R(\mathcal{E}^M) \leq E_R(\rho_{E}^M) \leq ME_R(\rho_{E}).$$  \hspace{1cm} (S93)

Then, because we have

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

from Eq. (S91) we may derive

$$\tilde{E}_R(\mathcal{E}) \leq E_R(\rho_{E}^M) + \delta_M \log_2 d + f(\delta_M/2).$$  \hspace{1cm} (S95)

As a result, we may write the upper bound

$$K(\mathcal{E}) \leq E_R(\rho_{E}^M) + \delta_M \log_2 d + f(\delta_M/2) \leq ME_R(\rho_{E}) + \frac{2d(d-1)}{M} \log_2 d + f \left[ \frac{d(d-1)}{M} \right]$$

$$:= K_{UB}^M(\mathcal{E}).$$  \hspace{1cm} (S96)

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

Let us apply the bound to channels that are nearly entanglement-breaking, so that $E_R(\rho_{E}) \ll 1$. 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

$$M \approx \sqrt{\frac{2d(d-1)\log_2 d}{E_R(\rho_{E})}},$$  \hspace{1cm} (S97)

which provides the upper bound

$$K(\mathcal{E}) \leq 2\sqrt{2d(d-1)\log_2 d}E_R(\rho_{E})$$

$$+ f \left( \frac{d(d-1)E_R(\rho_{E})}{2\log_2 d} \right).$$  \hspace{1cm} (S98)

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