

# Discriminating qubit amplitude damping channels

Milajiguli Rexiti\*

*School of Science and Technology, University of Camerino, 62032 Camerino, Italy*

Stefano Mancini†

*School of Science and Technology, University of Camerino, 62032 Camerino, Italy*

*and INFN-Sezione di Perugia, 06123 Perugia, Italy*

(Dated: February 5, 2021)

We address the issue of the discrimination between two-qubit amplitude damping channels by exploring several strategies. For the single-shot, we show that the excited state does not always give the optimal input, and that side entanglement assistance has limited benefit. On the contrary, feedback assistance from the environment is more beneficial. For the two-shot, we prove the in-utility of entangled inputs. Then focusing on individual (local) measurements, we find the optimal adaptive strategy.

## I. INTRODUCTION

A quantum channel is a linear stochastic (precisely a linear, completely positive and trace-preserving) map on the set of density operators [1]. As such, it can describe any physical process. It is often essential to distinguish between two (or even more) physical processes. Hence the issue of quantum channel discrimination becomes pervasive well beyond the boundary of information theory [2–6]. Quite generally, channel discrimination is a challenging task [7–10]. In fact, although it can be traced back to the (somehow old problem of) states discrimination [11, 12], it involves a double optimization: on the output measurement and the input state. Recently, bounds on the error probability were found for general strategies [13, 14].

Among quantum channels, the amplitude damping plays a prominent role as it describes the energy loss of a system, which is the most common effect occurring in an open system. As a matter of fact, this channel is often invoked as an example when dealing with discrimination (see e.g. [13, 14]). However, a systematic and thorough study of amplitude damping channel discrimination starting from dimension two is still lacking. Here we address this issue and unveil several unexpected results. For one-shot discrimination, the optimal input state turns out to not always be the excited state. Furthermore, side entanglement has a limited benefit because of a limited parameter region where it brings improvement and the smallness of such improvement. In contrast, feedback assistance from the environment results more beneficial. By such feedback, we mean the possibility to access the environment, measure it and then use this (classical) information to adjust the state of (or the measurement process in) the main system according to the desired goal [15]. Additionally, we prove that entangled inputs are not useful for two-shot, although collective measurement can give the minimum error probability. Then, restricting the attention to individual measurements, we find the optimal adaptive strategy (useful for a complete LOCC strategy [16]).

The paper is organized in two main Sections. Sec.II is devoted to one-shot discrimination. In Subsec.II A the optimal input is found. In Subsec.II B side entanglement is considered. Feedback assistance model is presented in Subsec.II C. Then, Sec.III is devoted to two-shot discrimination.

---

\*Electronic address: milajiguli.milajiguli@unicam.it

†Electronic address: stefano.mancini@unicam.it

In Subsec.III A the optimal input is found when also collective measurements are allowed. The optimal adaptive strategy for individual measurements is devised in Subsec.III B. Finally, in Sec.IV conclusions are drawn.

## II. ONE-SHOT DISCRIMINATION

It is customary to consider the amplitude damping channel as coming from the unitary interaction of the system with the environment generated by

$$H = \eta \left( a^\dagger b + ab^\dagger \right), \quad (1)$$

where  $a, a^\dagger$  (res.  $b, b^\dagger$ ) are the ladder operators of the system (resp. environment).

For a qubit system and qubit environment, in the computational basis  $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$  we have

$$H = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \eta & 0 \\ 0 & \eta & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (2)$$

and then the corresponding unitary  $U = e^{-iH}$  reads

$$U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \eta & -i \sin \eta & 0 \\ 0 & -i \sin \eta & \cos \eta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (3)$$

The map on the system's states can be written as

$$\mathcal{N}(\rho) = K_0 \rho K_0^\dagger + K_1 \rho K_1^\dagger, \quad (4)$$

where the Kraus operators are given by

$$K_0 = \langle 0|U|0\rangle = \begin{pmatrix} 1 & 0 \\ 0 & \cos \eta \end{pmatrix}, \quad (5)$$

$$K_1 = \langle 1|U|0\rangle = \begin{pmatrix} 0 & -i \sin \eta \\ 0 & 0 \end{pmatrix}, \quad (6)$$

with the bra-ket taken on the environment. The quantity  $\sin^2 \eta$  represents the decay probability ( $\eta \in [0, \frac{\pi}{2}]$ ).

### A. Optimal input

Let us analyze the distinguishability of two amplitude damping channels characterized by parameters  $\eta_0$  and  $\eta_1$ . By referring to Fig.1, we assume that each one acts with probability  $P_0 = P_1 = 1/2$  on an input state

$$|\psi\rangle = \sqrt{1-x} |0\rangle + e^{-i\varphi} \sqrt{x} |1\rangle, \quad (7)$$

where  $x \in [0, 1]$  and  $\varphi \in [0, 2\pi)$ . Without loss of generality, we take  $\eta_0 > \eta_1$ . We can also set  $\varphi = 0$ , because of the symmetric action of  $\mathcal{N}$  with respect to the  $z$  axis in the Bloch sphere.

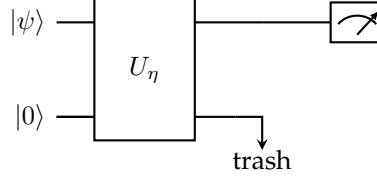


Figure 1: Schematic representation of the channel discrimination through unitary dilation. One has to determine whether  $U_{\eta_0}$  or  $U_{\eta_1}$  acted by controlling only the main system (top line).

According to (3), we obtain

$$U_{\eta_i} |\psi\rangle |0\rangle = \sqrt{1-x} |00\rangle + \sqrt{x} (-i \sin \eta_i |01\rangle + \cos \eta_i |10\rangle), \quad i = 0, 1. \quad (8)$$

Consequently, the output state of the channel reads

$$\begin{aligned} \rho_i = & (1-x + x \sin^2 \eta_i) |0\rangle \langle 0| + \sqrt{x(1-x)} \cos \eta_i |0\rangle \langle 1| \\ & + \sqrt{x(1-x)} \cos \eta_i |1\rangle \langle 0| + x \cos^2 \eta_i |1\rangle \langle 1|, \quad i = 0, 1. \end{aligned} \quad (9)$$

So the problem of channel discrimination is now translated into the discrimination between two mixed states,  $\rho_0$  and  $\rho_1$ , each occurring with probability  $1/2$ . To this end, it is known that the optimal measurement is given by the projection on the positive and negative subspaces of  $\rho_0 - \rho_1$  [1] (also known as Helstrom measurement [11]). Denoting by  $|v_0\rangle$  and  $|v_1\rangle$  its (normalized) eigenvectors corresponding respectively to positive and negative eigenvalues, we can evaluate

$$P_0(0) \equiv \langle v_0 | \rho_0 | v_0 \rangle, \quad (10a)$$

$$P_1(1) \equiv \langle v_1 | \rho_1 | v_1 \rangle. \quad (10b)$$

Then, the probability of success in discriminating between  $\rho_0$  and  $\rho_1$ , each occurring with probability  $1/2$ , is given by

$$P_{succ} = \frac{1}{2} P_0(0) + \frac{1}{2} P_1(1). \quad (11)$$

This turns also out to be [1]

$$P_{succ} = \frac{1}{2} \left( 1 + \frac{1}{2} \|\rho_0 - \rho_1\|_1 \right), \quad (12)$$

where  $\|T\|_1 \equiv \text{Tr} \sqrt{T^\dagger T}$ .

Inserting (10) into (11) (or equivalently using (9) into (12)), yields explicitly

$$P_{succ} = \frac{1}{2} \left\{ 1 + (\cos \eta_1 - \cos \eta_0) \sqrt{x[1-x(1-\gamma^2)]} \right\}, \quad (13)$$

where

$$\gamma \equiv \gamma(\eta_0, \eta_1) \equiv \cos \eta_1 + \cos \eta_0. \quad (14)$$

Maximizing (13) over  $x$  (taking into account that  $0 \leq x \leq 1$ ) we get

$$P_{succ} = \frac{1}{4} \left( 2 + \frac{\cos \eta_1 - \cos \eta_0}{\sqrt{1-\gamma^2}} \right), \quad \gamma < \frac{1}{\sqrt{2}}, \quad (15a)$$

$$P_{succ} = \frac{1}{2} (\sin^2 \eta_0 + \cos^2 \eta_1), \quad \gamma \geq \frac{1}{\sqrt{2}}. \quad (15b)$$

Note that the optimal value of  $x$  results

$$x = \frac{1}{2(1 - \gamma^2)}, \quad \text{for } \gamma < \frac{1}{\sqrt{2}}, \quad (16a)$$

$$x = 1, \quad \text{for } \gamma \geq \frac{1}{\sqrt{2}}. \quad (16b)$$

This means, interestingly, that the optimal input state is not always the excited state  $|1\rangle$  as one would expect (just because it is commonly considered as the most sensitive state to the damping action). The result can be better understood by looking at Fig.2, where it is represented the effect of the amplitude damping in the  $x$ - $z$  sector of the Bloch sphere. There, the radial distance corresponds to the trace distance between the fixed point  $\rho_0 = |0\rangle\langle 0|$  (for which  $\eta_0 = \pi/2$ ) and  $\rho_1$  in Eq.(9), for different values of  $\eta_1$ . One can clearly see that once  $\eta_1$  becomes bigger than  $\pi/4$  (hence  $\gamma$  becomes smaller than  $\frac{1}{\sqrt{2}}$ ), the farthest point from the origin is no longer on the  $z$  axes (see curve corresponding to  $\eta = \pi/3$ ).

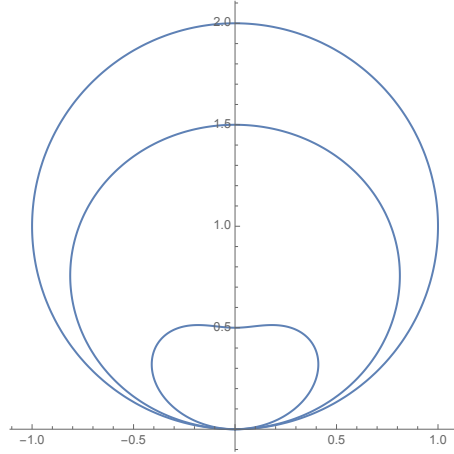


Figure 2: Representation of the amplitude damping effect in the  $x$ - $z$  sector of the Bloch sphere. The radial distance corresponds to the trace distance between the fixed point  $\rho_0 = |0\rangle\langle 0|$  (for which  $\eta_0 = \pi/2$ ) and  $\rho_1$  in Eq.(9), for different values of  $\eta_1$  (from outer to inner curve  $\eta_1 = 0, \pi/6$  and  $\pi/3$ ). The point at the intersection of the  $z$  axes and each curve above the origin corresponds to the state  $\rho_1 = \mathcal{N}(|1\rangle\langle 1|)$ .

## B. Side entanglement

Consider the usage of side entanglement according to the model of Fig.3.

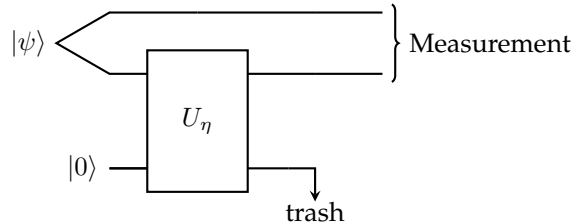


Figure 3: Model for channel discrimination exploiting entanglement between the input system and an accessible reference system.

Let  $|\Psi\rangle = \sqrt{1-y}|01\rangle + \sqrt{y}|10\rangle$  be an entangled state between the reference system and the channel's input. Then, we have to distinguish between the following two states at the measurement stage:

$$\begin{aligned} \text{Tr}_E \left\{ (I \otimes U_{\eta_i}) |\Psi\rangle |0\rangle \langle 0| \langle \Psi| (I \otimes U_{\eta_i})^\dagger \right\} = & \cos^2 \eta_i (1-y) |01\rangle \langle 01| + y |10\rangle \langle 10| \\ & + \sqrt{y(1-y)} \cos \eta_i (|01\rangle \langle 10| + |10\rangle \langle 01|) \\ & + (1-y) \sin^2 \eta_i |00\rangle \langle 00|, \quad i = 0, 1, \end{aligned} \quad (17)$$

where the unitaries are as in (3).

The success probability (12) for the states in (17) leads to

$$P_{\text{succ}}(\eta_0, \eta_1, y) = (\cos \eta_1 - \cos \eta_0) \left\{ (1-y)\gamma + \sqrt{(1-y)[4y + (1-y)\gamma^2]} \right\}. \quad (18)$$

Then, in Fig.4 we show  $P_{\text{succ}}(\eta_0, \eta_1, y^*)$ , where

$$y^*(\eta_0, \eta_1) := \arg\max_y P_{\text{succ}}(\eta_0, \eta_1, y) = \max \left\{ 0, \frac{\gamma-1}{\gamma-2} \right\}. \quad (19)$$

This latter quantity gives the optimal amount of entanglement and is shown in Fig.5.

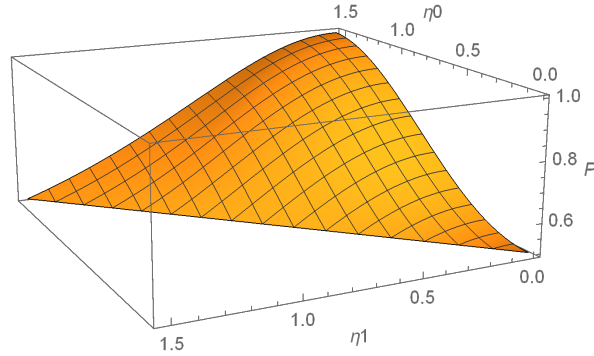


Figure 4: Maximum probability of success  $P(\eta_0, \eta_1, y^*)$  vs  $\eta_0$  and  $\eta_1$ .

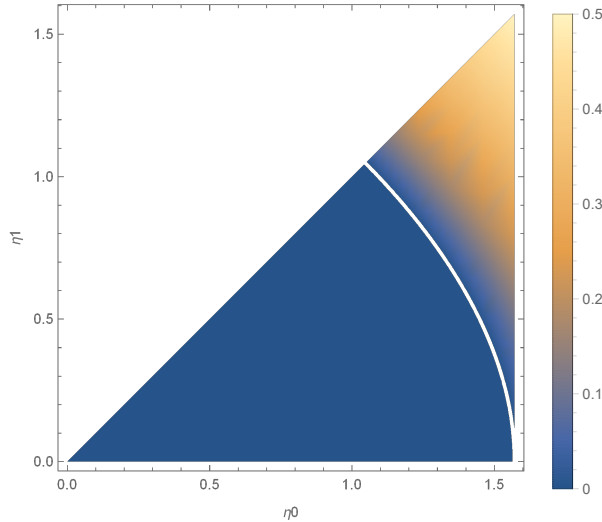


Figure 5: Optimal value  $y^*$  (Eq.(19)) vs  $\eta_0$  and  $\eta_1$ . The white line corresponds to  $\frac{\gamma-1}{\gamma-2} = 0$ .

In Fig.6 we report the difference  $P_{succ}(\eta_0, \eta_1, y^*) - P_{succ}(\eta_0, \eta_1, 0)$ . It shows that the improvement due to the side entanglement is relatively tiny. Moreover, it does not occur in all parameters' region. It is worth reminding that in other channels, like depolarizing channels, the effectiveness of side entanglement shows up for values of parameters making the channels entanglement-breaking [9]. Here, we have amplitude damping channel that is never entanglement-breaking, except for  $\eta_i = \pi/2$ . Nevertheless, the region with nonzero values in Fig.6 is close to this border. It is also interesting to note from Fig.5 that a small amount of entanglement is more effective than the maximal ( $y$  never reaches the value  $1/2$ ).

Finally, we note that according to the discussion at the end of SubSec.II A, the component  $|10\rangle$  of  $|\Psi\rangle$  (inputting no excitation to the channel) only plays a role when approaching the region  $\gamma < 1/\sqrt{2}$ . Then, the more we move towards the corner where  $\eta_0 = \eta_1 = \pi/2$ , the more the weight of the two components of  $|\Psi\rangle$  become balanced.

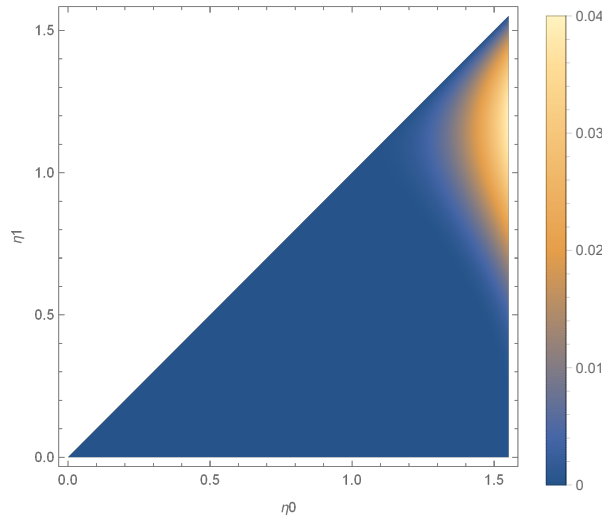


Figure 6: Difference between success probability computed at optimal  $y$  and  $y = 0$  (Eq.(18)) vs  $\eta_0$  and  $\eta_1$ .

### C. The use of feedback

Let us now move on to a discrimination strategy using feedback as illustrated in Fig.7. It presumes the possibility to locally access the environment [15].

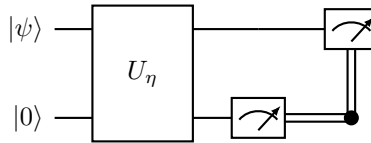


Figure 7: Channel discrimination with feedback. A first measurement is performed on the environment (bottom line), and then according to its outcome, a second measurement is performed on the system (top line). Here and in the following figures, double lines represent classical information.

We perform on the environment a measurement in the basis  $\{|\alpha^+\rangle, |\alpha^-\rangle\}$ , where

$$|\alpha^+\rangle = \cos \alpha |0\rangle + \sin \alpha |1\rangle, \quad (20a)$$

$$|\alpha^-\rangle = -\sin \alpha |0\rangle + \cos \alpha |1\rangle, \quad (20b)$$

are two orthogonal states in the plane  $x - z$  with  $\alpha \in [0, \pi/2]$  to be determined.

Then, depending on the environment's measurement result, we choose a proper observable for the system. In this way, the feedback *actuation* has to be intended as the measurement on the main system performed conditioned to the environment's measurement outcome.

We can hence distinguish between the following possibilities:

- Environment outcome +1 (eigenvalue corresponding to  $|\alpha^+\rangle$ ). This happens with probability:

$$P(+1) = \frac{1}{2} \text{Tr} \left( \langle \alpha^+ | U_{\eta_0} | \psi \rangle |0\rangle \langle 0| \langle \psi | U_{\eta_0}^\dagger | \alpha^+ \rangle \right) + \frac{1}{2} \text{Tr} \left( \langle \alpha^+ | U_{\eta_1} | \psi \rangle |0\rangle \langle 0| \langle \psi | U_{\eta_1}^\dagger | \alpha^+ \rangle \right), \quad (21)$$

and the resulting state on the system is

$$|\varphi_0^+\rangle = \frac{1}{N_0^+} \langle \alpha^+ | \left( U_{\eta_0} | \psi \rangle |0\rangle \right), \quad (22)$$

or

$$|\varphi_1^+\rangle = \frac{1}{N_1^+} \langle \alpha^+ | \left( U_{\eta_1} | \psi \rangle |0\rangle \right), \quad (23)$$

depending on which unitary has acted. Here,  $1/N_i^+, i = 0, 1$  are normalization factors. It is then optimal to measure the observable  $|\varphi_0^+\rangle \langle \varphi_0^+| - |\varphi_1^+\rangle \langle \varphi_1^+|$  on the main system to discriminate between (22) and (23) [1]. This can be done, following (12), with probability [25]  $\frac{1}{2} \left( 1 + \sqrt{1 - |\langle \varphi_0^+ | \varphi_1^+ \rangle|^2} \right)$ . Thus, the probability of success when environment outcome is +1, reads:

$$P_{succ}^{(+)} = P(+1) \times \frac{1}{2} \left( 1 + \sqrt{1 - |\langle \varphi_0^+ | \varphi_1^+ \rangle|^2} \right). \quad (24)$$

- Environment outcome -1 (eigenvalue corresponds to  $|\alpha^-\rangle$ ). This happens with probability:

$$P(-1) = \frac{1}{2} \text{Tr} \left( \langle \alpha^- | U_{\eta_0} | \psi \rangle |0\rangle \langle 0| \langle \psi | U_{\eta_0}^\dagger | \alpha^- \rangle \right) + \frac{1}{2} \text{Tr} \left( \langle \alpha^- | U_{\eta_1} | \psi \rangle |0\rangle \langle 0| \langle \psi | U_{\eta_1}^\dagger | \alpha^- \rangle \right), \quad (25)$$

and the resulting state on the system is

$$|\varphi_0^-\rangle = \frac{1}{N_0^-} \langle \alpha^- | \left( U_{\eta_0} | \psi \rangle |0\rangle \right), \quad (26)$$

or

$$|\varphi_1^-\rangle = \frac{1}{N_1^-} \langle \alpha^- | \left( U_{\eta_1} | \psi \rangle |0\rangle \right), \quad (27)$$

depending on the acted unitary. Here  $1/N_i^-, i = 0, 1$  are normalization factors. It is then optimal to measure the observable  $|\varphi_0^-\rangle \langle \varphi_0^-| - |\varphi_1^-\rangle \langle \varphi_1^-|$  on the main system to discriminate between (26) and (27). Thus, the probability of success when environment outcome is -1, reads:

$$P_{succ}^{(-)} = P(-1) \times \frac{1}{2} \left( 1 + \sqrt{1 - |\langle \varphi_0^- | \varphi_1^- \rangle|^2} \right). \quad (28)$$

Finally, putting (24) and (28) together we get the overall probability of success as

$$\begin{aligned}
P_{succ} &= P_{succ}^{(+)} + P_{succ}^{(-)} \\
&= \frac{\chi}{2} \left[ 1 + \sin \alpha \sin \left( \frac{\eta_0 - \eta_1}{2} \right) \sqrt{\frac{x(\mu + \nu)}{c_1 c_2}} \right] \\
&\quad + \frac{1 - \chi}{2} \left[ 1 + \cos \alpha \sin \left( \frac{\eta_0 - \eta_1}{2} \right) \sqrt{\frac{x(\mu - \nu)}{(1 - c_1)(1 - c_2)}} \right],
\end{aligned} \tag{29}$$

where

$$\chi \equiv \frac{1}{2} - \frac{1}{2} \cos(2\alpha) [1 - x(\sin^2 \eta_0 + \sin^2 \eta_1)], \tag{30a}$$

$$\mu \equiv [1 + (2x - 1) \cos \eta_0 \cos \eta_1 + \sin \eta_0 \sin \eta_1], \tag{30b}$$

$$\nu \equiv \cos(2\alpha) [(2x - 1) + \cos \eta_0 \cos \eta_1 + (2x - 1) \sin \eta_0 \sin \eta_1], \tag{30c}$$

$$c_i \equiv \frac{1}{2} - \frac{1}{2} \cos(2\alpha) [1 - 2x \sin^2 \eta_i], \quad i = 0, 1. \tag{30d}$$

Analyzing Eq.(29), we found the maximum success probability as

$$P_{succ} = \frac{1 + \sin(\eta_0 - \eta_1)}{2}, \tag{31}$$

attained when  $x = 1$  and  $\alpha = \frac{\pi}{4}$ . Provided that  $\eta_0 \neq \eta_1$  and  $\eta_0 \neq \frac{\pi}{2} - \eta_1$ , we have the probability of success with feedback greater than without feedback as shown in Fig.8.

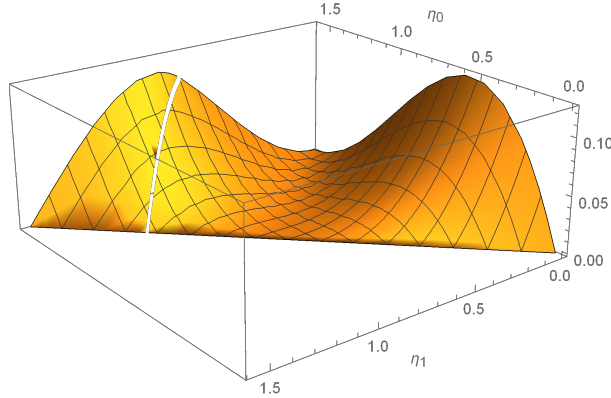


Figure 8: Difference between the maximum probability of success with feedback (31) and without feedback (15) vs  $\eta_0$  and  $\eta_1$ . The white line corresponds to  $\gamma = 1/\sqrt{2}$  (see (14)).

We note that the best measurement on the environment is on the basis  $|\pm\rangle \equiv (|0\rangle \pm |1\rangle)/\sqrt{2}$  and the best input state is always  $|1\rangle$  in contrast to what happened in the absence of feedback. In Fig.8 it is also visible a slight asymmetry of the behavior between the regions  $\eta_1 < \pi/2 - \eta_0$  and  $\eta_1 > \pi/2 - \eta_0$ . We suspect that this is due to the fact that projection on the environment helps in creating states on the main system that are better distinguishable, provided that system and environment are enough entangled prior the measurement. For big  $\eta$ 's they are not, hence the process is less effective.

Finally, it is worth mentioning a similarity of this problem with the discrimination of unitary dilations of the amplitude damping channel using local measurements (analogous similarity was pointed out in Ref.[17] for channel estimation).



### III. TWO-SHOT DISCRIMINATION

In this section, we shall consider the discrimination assuming to have two copies of the channel characterized either by parameter  $\eta_0$  or by parameter  $\eta_1$  (with again equal probability  $1/2$ ).

#### A. Optimal input

Let us first consider the possibility of using entangled inputs. Since for a single shot the optimal input state lies in the  $x - z$  plane, we construct entangled input states as a linear combination of 2-fold tensor product of these states with real coefficients. This amounts to consider the two inputs state as

$$\sqrt{1-x}|01\rangle + \sqrt{x}|10\rangle, \quad (32)$$

or

$$\sqrt{1-x}|00\rangle + \sqrt{x}|11\rangle, \quad (33)$$

with  $x \in [0, 1]$ .

According to the 2-fold action of (4), in case of (32) we will get the output states as

$$\begin{aligned} \rho_i^{(2)} = & \sin^2 \eta_i |00\rangle\langle 00| + (1-x) \cos^2 \eta_i |01\rangle\langle 01| + x \cos^2 \eta_i |10\rangle\langle 10| \\ & + \sqrt{x(1-x)} \cos^2 \eta_i (|01\rangle\langle 10| + |10\rangle\langle 01|), \quad i = 0, 1. \end{aligned} \quad (34)$$

While in case of (33) as

$$\begin{aligned} \rho_i^{(2)} = & ((1-x) + x \sin^4 \eta_i) \eta_i |00\rangle\langle 00| + x \sin^2 \eta_i \cos^2 \eta_i (|01\rangle\langle 01| + |10\rangle\langle 10|) + x \cos^4 \eta_i |11\rangle\langle 11| \\ & + \sqrt{x(1-x)} \cos^2 \eta_i (|00\rangle\langle 11| + |11\rangle\langle 00|), \quad i = 0, 1. \end{aligned} \quad (35)$$

Using (12) with the states (34) gives

$$P_{succ} = \frac{1}{2} (1 + \sin^2 \eta_0 - \sin^2 \eta_1), \quad (36)$$

which is independent of  $x$ . On the other hand, using (12) with the states (35) gives

$$P_{succ} = \frac{1}{2} \left\{ 1 + \frac{x}{4} |\cos^2(2\eta_0) - \cos^2(2\eta_1)| + \frac{\sqrt{x}}{2} |\cos(2\eta_0) - \cos(2\eta_1)| \right\}, \quad (37)$$

which attains its maximum for  $x = 1$ . Therefore, we can conclude that entanglement across the two inputs is useless. Then, we consider as input the 2-fold tensor product of (7) and optimize over  $x$ . In other words, we consider

$$P_{succ} = \max_x \frac{1}{2} \left( 1 + \frac{1}{2} \left\| \rho_0 \otimes \rho_0 - \rho_1 \otimes \rho_1 \right\|_1 \right), \quad (38)$$

where  $\rho_i$ s are given by (9).

The difference between this optimized success probability and (15) is shown in Fig.9.

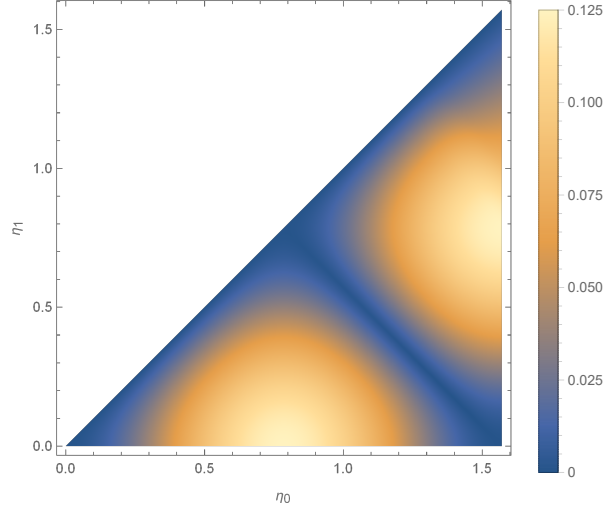


Figure 9: Difference between the maximum probability of success (38) and (15) vs  $\eta_0$  and  $\eta_1$ .

The optimal value  $x^*$  of  $x$  for (38) as function of  $\eta_0$  and  $\eta_1$  is reported in Fig.10. Note that the region where the  $x^*$  value is smaller than one is shrunk with respect to the single-shot case ( $\gamma < 1/2$ ). Although the exact boundary cannot be expressed analytically, we found numerically that

$$\gamma(\eta_0, \eta_1) < \frac{1}{2} \Rightarrow x < 1, \quad (39)$$

where  $\gamma(\eta_0, \eta_1)$  is given by (14). It is worth remarking that for  $x^* = 1$  the optimal observable  $\rho_0 - \rho_1$  constructed with (34) turns out to be local (its normalized eigenvectors are  $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ ), while for  $x^* < 1$  results nonlocal (its normalized eigenvectors are entangled).

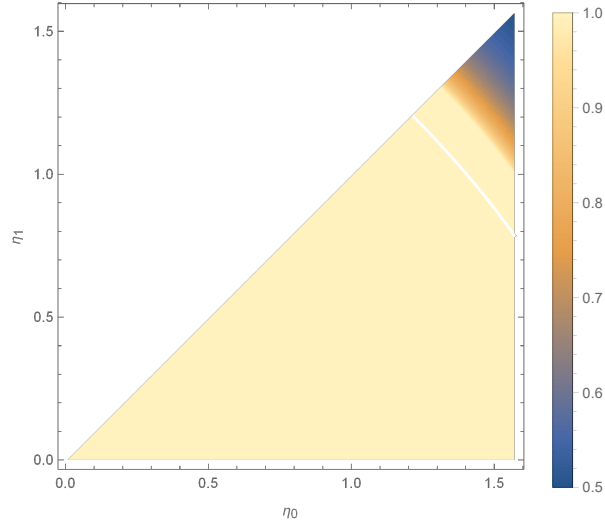


Figure 10: Optimal value  $x^*$  of  $x$  for (38) vs  $\eta_0$  and  $\eta_1$ . The white line represents the boundary  $\gamma = 1/\sqrt{2}$  for the single-shot case.

### B. Adaptive strategy

In the previous subsection, although getting rid of entangled inputs, we saw the necessity of using collective measurement in some parts of the parameters' region. What happens if we restrict to individual measurements to have a completely local strategy (analogously to [16])? We expect an improvement with respect to the one shot-case, but this relies on using an adaptive strategy, which can generally be depicted as in Fig.11.

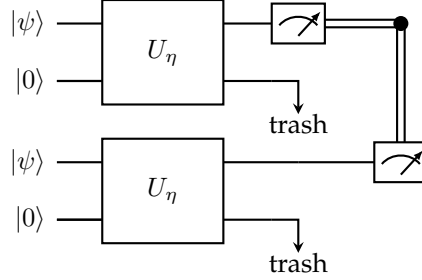


Figure 11: Schematic representation of channel discrimination through a local adaptive strategy. The measurement outcome on the first copy determines the measurement to be performed on the second copy.

At the output of the two channel copies, we have

$$\frac{1}{2}\rho_0^{\otimes 2} + \frac{1}{2}\rho_1^{\otimes 2}, \quad (40)$$

with  $\rho_i$  given by (9).

On the first copy we use the POVM whose elements are  $|v_0\rangle\langle v_0|$  and  $|v_1\rangle\langle v_1|$ , with  $|v_0\rangle$  and  $|v_1\rangle$  the (normalized) eigenvectors of the observable  $\rho_0 - \rho_1$ . On the second copy, we choose a POVM depending on the previous copy's measurement outcome, i.e.,  $\Pi_{x_2}^{x_1}$ ,  $x_i \in \{0, 1\}$ , where the subscript denotes the element of the POVM, while the superscript the dependence from the previous measurement outcome. We define

$$P_j(x_2|x_1) \equiv \text{Tr} [\rho_j \Pi_{x_2}^{x_1}]. \quad (41)$$

Then the success probability will be given by

$$\begin{aligned} P_{succ} = & \frac{1}{2} [P_0(0|0)\langle v_0|\rho_0|v_0\rangle + P_0(0|1)(1 - \langle v_0|\rho_0|v_0\rangle)] \\ & + \frac{1}{2} [P_1(1|1)\langle v_1|\rho_1|v_1\rangle + P_1(1|0)(1 - \langle v_1|\rho_1|v_1\rangle)]. \end{aligned} \quad (42)$$

Eq.(42) should be maximized overall POVMs  $\Pi_{x_2}^{x_1}$ . Actually, the first and fourth terms can be maximized overall  $\Pi_0^0$  being  $\Pi_1^0 = I - \Pi_0^0$ , while the second and third terms can be maximized overall  $\Pi_1^1$  being  $\Pi_0^1 = I - \Pi_1^1$ .

So we can independently perform the following maximizations (for fixed input  $x$ ):

$$\max_{\Pi_0^0: 0 \leq \Pi_0^0 \leq I} \frac{1}{2} \{ \text{Tr} [\rho_0 \Pi_0^0] \langle v_0|\rho_0|v_0\rangle + \text{Tr} [\rho_1 (I - \Pi_0^0)] (1 - \langle v_1|\rho_1|v_1\rangle) \}, \quad (43)$$

and

$$\max_{\Pi_1^1: 0 \leq \Pi_1^1 \leq I} \frac{1}{2} \{ \text{Tr} [\rho_0 (I - \Pi_1^1)] (1 - \langle v_0|\rho_0|v_0\rangle) + \text{Tr} [\rho_1 \Pi_1^1] \langle v_1|\rho_1|v_1\rangle \}. \quad (44)$$

Let us consider the problem of Eq.(43). Defining  $p_0 := \langle v_0 | \rho_0 | v_0 \rangle$  and  $q_0 := \langle v_0 | \rho_1 | v_0 \rangle$  we can recast it into the following form

$$\frac{p_0 + q_0}{2} \max_{\Pi_0^0: 0 \leq \Pi_0^0 \leq I} \left\{ \frac{p_0}{p_0 + q_0} \text{Tr} [\rho_0 \Pi_0^0] + \frac{q_0}{p_0 + q_0} \text{Tr} [\rho_1 (I - \Pi_0^0)] \right\}, \quad (45)$$

which coincides with the optimal discrimination of  $\rho_0$  and  $\rho_1$ , this time appearing with probabilities  $\frac{p_0}{p_0 + q_0}$  and  $\frac{q_0}{p_0 + q_0}$  respectively. Hence, it can be solved by again resorting to Helstrom measurement. In other words,  $\Pi_0^0$  and  $I - \Pi_0^0$  can be constructed as projectors onto positive and negative subspaces of the operator  $\frac{p_0}{p_0 + q_0} \rho_0 - \frac{q_0}{p_0 + q_0} \rho_1$ .

With Eq.(44) one can proceed similarly by defining  $p_1 := \langle v_1 | \rho_1 | v_1 \rangle$  and  $q_1 := \langle v_1 | \rho_0 | v_1 \rangle$ . Therefore,  $\Pi_1^1$  and  $I - \Pi_1^1$  will be constructed as projectors onto positive and negative subspaces of the operator  $\frac{q_1}{p_1 + q_1} \rho_0 - \frac{p_1}{p_1 + q_1} \rho_1$ . Ultimately, we can get the success probability as the function  $P_{succ}(\eta_0, \eta_1, x)$  given by

$$P_{succ} = \frac{p_0 + q_0}{2} \frac{1}{2} \left\{ 1 + \left\| \frac{p_0}{p_0 + q_0} \rho_0 - \frac{q_0}{p_0 + q_0} \rho_1 \right\|_1 \right\} + \frac{p_1 + q_1}{2} \frac{1}{2} \left\{ 1 + \left\| \frac{q_1}{p_1 + q_1} \rho_0 - \frac{p_1}{p_1 + q_1} \rho_1 \right\|_1 \right\}. \quad (46)$$

To evaluate the performance of the adaptive strategy, we compare this probability, maximized over  $x$ , with the optimal success probability for two-shot (38). Fig.12 shows the difference between the latter and the former. As we expected, such a difference is nonzero only in the region of Fig.10 where  $x^* < 1$ , however it is very tiny. This shows that the devised local adaptive strategy performs almost like the strategy involving collective measurement.

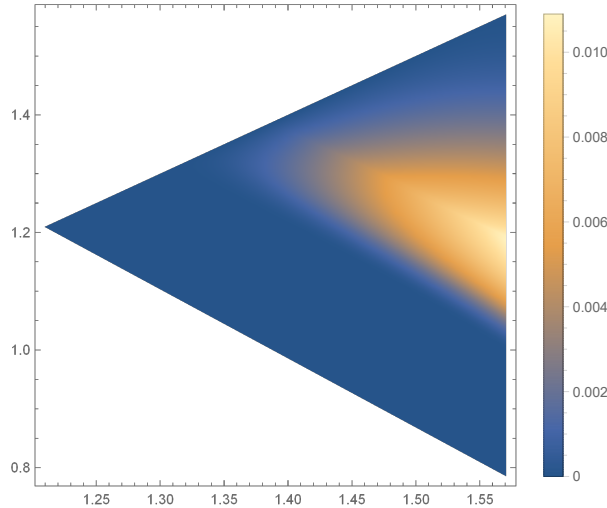


Figure 12: Difference between the maximum probability of success using collective measurement (Eq.(38)) and that using the adaptive strategy (Eq.(46) maximized over  $x$ ) vs  $\eta_0$  and  $\eta_1$  in the region  $\gamma < 1/\sqrt{2}$ .

Furthermore, Fig.13 shows the difference between the maximum probability of success using the adaptive strategy and that of a single-shot. Note that the range of values in Fig.13 is one order of magnitude bigger than Fig.12. As a consequence Fig.13 faithfully follows Fig.9.

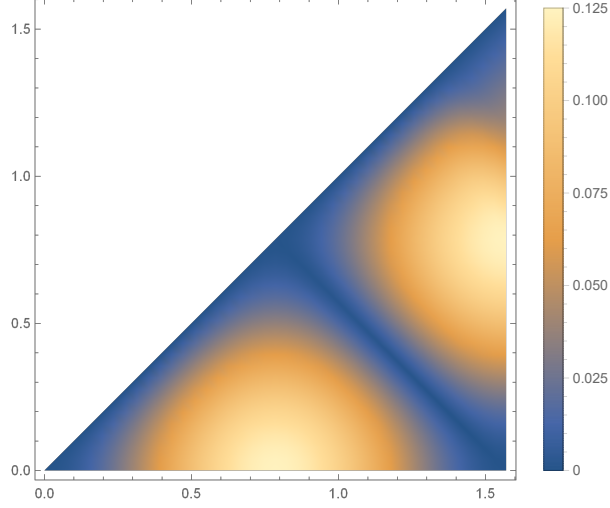


Figure 13: Difference between the maximum probability of success using the adaptive strategy (Eq.(46) maximized over  $x$ ) and that of a single shot (Eq.(15)) vs  $\eta_0$  and  $\eta_1$ .

We close this SubSection saying that the strategy we pursued was based on Bayesian updating using forward optimization, much like Ref.[18]. However in Ref.[19] it was shown that backward optimization can give better performance in case mixed states have to be discriminated. In the present case the latter seems not much effective. Anyway, a comparison of the two methods is provided in Appendix A.

### C. Adaptive strategy with feedback

We now develop an adaptive strategy that includes the feedback from the environment, as illustrated in Fig.14.

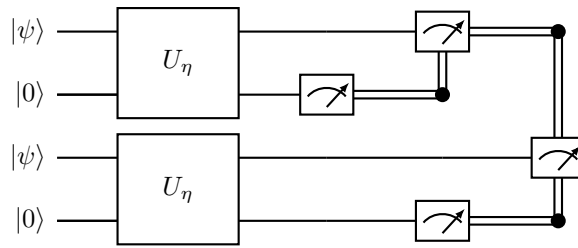


Figure 14: Schematic representation of channel discrimination through a local adaptive strategy including feedback. On each step the measurement to be performed on the main system is determined by the measurement outcome of the environment on that step, together with the measurement outcome of the main system at the previous step.

On the first copy of the unitary, we proceed like in Subsection II C, defining a POVM  $\Pi_{x_1}^{e_1}$  whose elements are

$$\Pi_{x_1=0}^{e_1} = |v_0^{e_1}\rangle\langle v_0^{e_1}|, \quad (47)$$

$$\Pi_{x_1=1}^{e_1} = |v_1^{e_1}\rangle\langle v_1^{e_1}|, \quad (48)$$

where  $|v_0^{e_1}\rangle$  and  $|v_1^{e_1}\rangle$  are (normalized) eigenvectors of  $\varphi_0^{e_1} - \varphi_1^{e_1}$ . Here  $\varphi_j^{e_1} \equiv |\varphi_j^{e_1}\rangle\langle\varphi_j^{e_1}|$  with  $j = 0, 1$  and  $e_1 = \pm$ . Note that following the conclusions of Subsection II C, we are considering the input  $|\psi\rangle$  as  $|1\rangle$  and the measurement on the environment on the basis  $|\pm\rangle$ , so that

$$|\varphi_0^\pm\rangle = \mp i \sin \eta_0 |0\rangle + \cos \eta_0 |1\rangle, \quad (49a)$$

$$|\varphi_1^\pm\rangle = \mp i \sin \eta_1 |0\rangle + \cos \eta_1 |1\rangle. \quad (49b)$$

As a consequence, at the output of the main system, we have to distinguish between pure states (differently to what happened in the previous Subsection). We then define the probability

$$P_j^{e_2}(x_2|x_1, e_2) \equiv \text{Tr} \left[ \varphi_j^{e_2} \Pi_{x_2}^{x_1, e_2} \right], \quad (50)$$

where  $\Pi_{x_2}^{x_1, e_2}$  indicates the POVM elements characterized by  $x_i \in \{0, 1\}$  and  $e_2 = \pm$ .

In terms of it, we can express the success probability as

$$P_{succ} = \sum_{e_1, e_2 = \pm} \frac{1}{8} \left[ P_0^{e_2}(0|0, e_2) \langle v_0^{e_1} | \varphi_0^{e_1} | v_0^{e_1} \rangle + P_0^{e_2}(0|1, e_2) (1 - \langle v_0^{e_1} | \varphi_0^{e_1} | v_0^{e_1} \rangle) \right. \\ \left. + P_1^{e_2}(1|1, e_2) \langle v_1^{e_1} | \varphi_1^{e_1} | v_1^{e_1} \rangle + P_1^{e_2}(1|0, e_2) (1 - \langle v_1^{e_1} | \varphi_1^{e_1} | v_1^{e_1} \rangle) \right], \quad (51)$$

where the factor  $\frac{1}{8}$  in front of the square brackets arises from the probability  $\frac{1}{2}$  of having  $U_{\eta_i} \otimes U_{\eta_i}$ , the probability  $\frac{1}{2}$  for  $e_1$  to take one of the two values, and the probability  $\frac{1}{2}$  for  $e_2$  to take one of the two values. Now for each value of  $e_1$  and  $e_2$  we can separately maximize the terms

$$\max_{\Pi_0^{0, e_2}: 0 \leq \Pi_0^{0, e_2} \leq I} \frac{1}{8} \left\{ \text{Tr} \left[ \varphi_0^{e_2} \Pi_0^{0, e_2} \right] \langle v_0^{e_1} | \varphi_0^{e_1} | v_0^{e_1} \rangle + \text{Tr} \left[ \varphi_0^{e_2} (I - \Pi_0^{0, e_2}) \right] (1 - \langle v_0^{e_1} | \varphi_0^{e_1} | v_0^{e_1} \rangle) \right\}, \quad (52)$$

and

$$\max_{\Pi_1^{1, e_2}: 0 \leq \Pi_1^{1, e_2} \leq I} \frac{1}{8} \left\{ \text{Tr} \left[ \varphi_0^{e_2} (I - \Pi_1^{1, e_2}) \right] \langle v_0^{e_1} | \varphi_0^{e_1} | v_0^{e_1} \rangle + \text{Tr} \left[ \varphi_0^{e_2} \Pi_1^{1, e_2} \right] (1 - \langle v_1^{e_1} | \varphi_1^{e_1} | v_1^{e_1} \rangle) \right\}, \quad (53)$$

following the same reasoning of SubSec.III B.

Ultimately, we arrive at

$$P_{succ} = \frac{1}{2} \left[ 1 + \sin(\eta_0 - \eta_1) \sqrt{1 + \cos^2(\eta_0 - \eta_1)} \right]. \quad (54)$$

The improvement with respect to the single-shot with feedback (Eq.(31)) is shown in Fig.15. We can see that it is positive in all parameters' region (but  $\eta_1 = \eta_0$  and  $\eta_0 = \frac{\pi}{2}, \eta_1 = 0$ ) and is maximum for  $\eta_1 = \eta_0 - \frac{\pi}{4}$ .

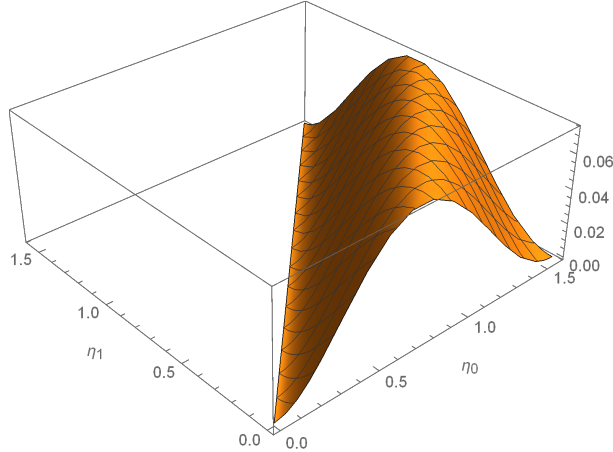


Figure 15: Difference between the probabilities of success (54) and (31) vs  $\eta_0$  and  $\eta_1$ .

#### IV. CONCLUDING REMARKS

We have addressed the issue of discriminating between two-qubit amplitude damping channels by considering single and double-shot. For the one-shot, we showed that the excited state as input is not always optimal, and the side entanglement assistance has a limited benefit. In contrast, feedback assistance from the environment is highly beneficial. About feedback, we suspect that it will be even more effective in case of discrimination of random unitary channels, where the information recovered from the environment neutralizes the channel action (taking it back to identity map).

For the two-shot, we proved the in-utility of entangled inputs. Then, focusing on individual (local) measurements, we found the optimal adaptive strategy. We are confident that this strategy can be extended in a Markovian way to  $n$ -shot, and the asymptotic analysis of its performance would be foreseeable, eventually by employing techniques like dynamical programming or reinforcement learning [21, 22]. Similarly, the adaptive strategy is applied to environment feedback assisted discrimination showing a smaller improvement than the case without feedback (cfr. Figs.13 and 15). This is because feedback has provided a big enhancement already on the first shot.

It is worth remarking that having two copies of the channel one could have used them in sequence, not in parallel as done in Sec.III. In such a case one is led to distinguish between  $\mathcal{N}(\rho_0)$  and  $\mathcal{N}(\rho_1)$  (where  $\rho_i$ s are given in (9)), each occurring with probability  $\frac{1}{2}$ . Simple calculations show that this latter strategy cannot be better than the parallel one. For example, without feedback, its success probability equals (46) in the region  $\eta_1 < \eta_0$  and  $\eta_1 < \frac{\pi}{2} - \eta_0$ , while it results strictly smaller than (46) in the region  $\eta_1 < \eta_0$  and  $\eta_1 > \frac{\pi}{2} - \eta_0$ .

In the future, it is worth extending the performed analysis to  $d$ -dimensional amplitude damping channels, also with restrictions on the set of input states (e.g., energy restriction). After all, investigating amplitude damping channel discrimination in discrete systems can also provide new insights for distinguishing continuous variable lossy channels [20], although this is often traced back to the distinguishability of coherent states (for this latter subject see e.g. [22–24]).

#### Appendix A: Optimal POVMs for local adaptive strategy

Suppose that instead of projecting onto  $|v_0\rangle$  and  $|v_1\rangle$  on the first copy we realize a measurement by a POVM  $\{M, I - M\}$ . Then (42) will become

$$P_{succ} = \frac{1}{2} \left\{ \text{Tr} [\rho_0 \Pi_0^0] \text{Tr} [\rho_0 M] + \text{Tr} [\rho_0 (I - \Pi_1^1)] \text{Tr} [\rho_0 (I - M)] \right. \\ \left. + \text{Tr} [\rho_1 \Pi_1^1] \text{Tr} [\rho_1 (I - M)] + \text{Tr} [\rho_1 (I - \Pi_0^0)] \text{Tr} [\rho_1 M] \right\}. \quad (\text{A1})$$

Defining

$$r_0 \equiv \text{Tr} [\rho_0 M], \quad s_0 \equiv \text{Tr} [\rho_1 M], \quad (\text{A2})$$

$$r_1 \equiv 1 - s_0, \quad s_1 = 1 - r_0, \quad (\text{A3})$$

equation (A1) can be rewritten as

$$P_{succ} = \frac{r_0 + s_0}{2} \left\{ \frac{r_0}{r_0 + s_0} \text{Tr} [\rho_0 \Pi_0^0] + \frac{s_0}{r_0 + s_0} \text{Tr} [\rho_1 (I - \Pi_0^0)] \right\} \\ + \frac{r_1 + s_1}{2} \left\{ \frac{s_1}{r_1 + s_1} \text{Tr} [\rho_0 (I - \Pi_1^1)] + \frac{r_1}{r_1 + s_1} \text{Tr} [\rho_1 \Pi_1^1] \right\}. \quad (\text{A4})$$

Let us now impose that the measurement on the second copy is of the Helstrom kind. Thus  $\Pi_0^0$  and  $(I - \Pi_0^0)$  are projectors onto the positive and negative eigenspaces of

$$\frac{r_0}{r_0 + s_0} \rho_0 - \frac{s_0}{r_0 + s_0} \rho_1. \quad (\text{A5})$$

Analogously  $\Pi_1^1$  and  $(I - \Pi_1^1)$  are projectors onto the positive and negative eigenspaces of

$$\frac{r_1}{r_1 + s_1} \rho_0 - \frac{s_1}{r_1 + s_1} \rho_1. \quad (\text{A6})$$

As such all the quantities appearing in Eq.(A4) will depend upon  $M$  (for a fixed  $x$ ). Then, maximizing  $P_{succ}$  over  $M$ , such that  $0 \leq M \leq I$ , realizes the backward optimization.

In such a way we can get the success probability  $P_{succ}^{(\leftarrow)}(\eta_0, \eta_1, x)$ . We also denote by  $P_{succ}^{(\rightarrow)}(\eta_0, \eta_1, x)$  the quantity (46). Then, in Fig.16, we show the difference

$$\max_x P_{succ}^{(\rightarrow)}(\eta_0, \eta_1, x) - \max_x P_{succ}^{(\leftarrow)}(\eta_0, \eta_1, x). \quad (\text{A7})$$

We may note that when the backward optimization performs better, it only gives a tiny improvement (of the order of  $10^{-3}$ ) to the success probability.

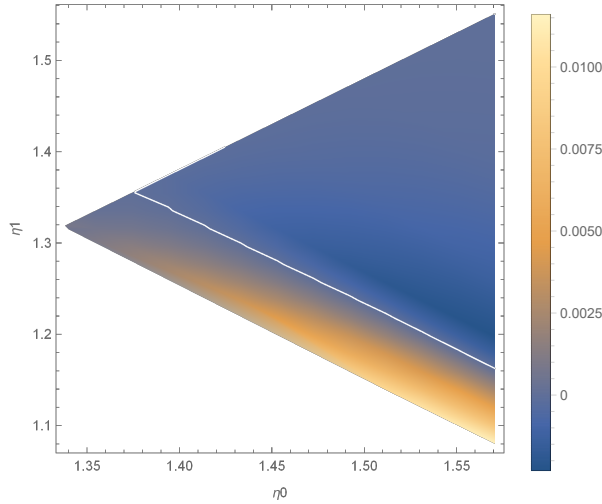


Figure 16: Difference (A7) between forward and backward optimized success probabilities vs  $\eta_0$  and  $\eta_1$ . The white line represents 0. Only the region where  $P_{succ}^{(\rightarrow)}(\eta_0, \eta_1, x)$  differs from (38) is shown.

The authors equally contributed to this work.

The work was financially supported by the Future and Emerging Technologies (FET) programme, within the Horizon-2020 Programme of the European Commission, under the FET-Open grant agreement QUARTET, number 862644.

The authors thank Francisco Revson F. Pereira for useful discussions.



- [2] A. Acin, *Statistical distinguishability between unitary operations*, Physical Review Letters, **87**, 177901 (2001).
- [3] A. Gilchrist, N. K. Langford, and M. A. Nielsen, *Distance measures to compare real and ideal quantum processes*, Physical Review A **71**, 062310 (2005).
- [4] M. Sacchi, *Optimal discrimination of quantum operations*, Physical Review A **71**, 062340 (2005).
- [5] G. Wang, and M. Ying, *Unambiguous discrimination among quantum operations*, Physical Review A **73**, 042301 (2006).
- [6] M. Hayashi, *Discrimination of two channels by adaptive methods and its application to quantum system*, IEEE Transactions on Information Theory **55**, 3807 (2009).
- [7] A. M. Childs, J. Preskill, and J. Renes, *Quantum information and precision measurement*, Journal of Modern Optics, **47**, 155-176 (2000).
- [8] B. Rosgen, and J. Watrous, *On the hardness of distinguishing mixed-state quantum computations*, Proceedings of the 20th IEEE Conference on Computational Complexity, pp. 344-354 (2005).
- [9] M. Sacchi, *Entanglement can enhance the distinguishability of entanglement-breaking channels*, Physical Review A **72**, 014305 (2005).
- [10] M. Sacchi, *Minimum error discrimination of Pauli channels*, Journal of Optics B: Quantum and Semiclassical Optics **7**, S333 (2005).
- [11] C. W. Helstrom, *Quantum detection and estimation theory*, Journal of Statistical Physics, **1**, 231-252 (1969).
- [12] A. S. Holevo, *Statistical decision theory for quantum systems*, Journal of Multivariate Analysis **3**, 337-394 (1973).
- [13] S. Pirandola, R. Laurenza, C. Lupo, and J. L. Pereira, *Fundamental limits to quantum channel discrimination*, arXiv:1803.02834 (2019).
- [14] V. Katariya, and M. M. Wilde, *Evaluating the advantage of adaptive strategies for quantum channel distinguishability*, arXiv:2001.05376 (2020).
- [15] L. Memarzadeh, C. Cafaro, and S. Mancini, *Quantum information reclaiming after amplitude damping*, Journal of Physics A: Mathematical and Theoretical **44**, 045304 (2011).
- [16] P. Tej, S. R. Ahmed, A. R. Usha Devi, and A. K. Rajagopal, *Quantum hypothesis testing and state discrimination*, arXiv:1803.04944 (2018).
- [17] M. Rexit, and S. Mancini, *Adversarial versus cooperative quantum estimation*, Quantum Information Processing **18**, 102 (2019).
- [18] A. Acin, E. Bagan, M. Baig, Ll. Masanes, and R. Muñoz-Tapia, *Multiple-copy two-state discrimination with individual measurements*, Physical Review A **71**, 032338 (2005).
- [19] B. L. Higgins, A. C. Doherty, S. D. Bartlett, G. J. Pryde, and H. M. Wiseman, *Multiple-copy state discrimination: Thinking globally, acting locally*, Physical Review A **83**, 052314 (2011).
- [20] C. Invernizzi, M. G. A. Paris, and S. Pirandola, *Optimal detection of losses by thermal probes*, Physical Review A **84**, 022334 (2011).
- [21] S. Brandsen, M. Lian, K. D. Stubbs, N. Rengaswamy, and H. D. Pfister, *Adaptive procedures for discrimination between arbitrary tensor-product quantum states* 2020 IEEE International Symposium on Information Theory (ISIT), Los Angeles, CA, USA, 2020, pp. 1933-1938.
- [22] M. Bilkis, M. Rosati, R. Morral Yepes, and J. Calsamiglia, *Real-time calibration of coherent-state receivers: Learning by trial and error* Physical Review Research **2**, 033295 (2020).
- [23] S. J. Dolinar, *An optimum receiver for the binary coherent state quantum channel*, MIT Research Laboratory of Electronics Quarterly Progress Report **111**, 115 (1973).
- [24] A. Assalini, N. Dalla Pozza, and G. Pierobon, *Revisiting the Dolinar receiver through multiple-copy state discrimination theory*, Physical Review A **84**, 022342 (2011).
- [25] This expression derives from (12) when employing pure state vectors.