Channel correction via quantum erasure

Francesco Buscemi*
ERATO-SORST Quantum Computation and Information Project,
Japan Science and Technology Agency,
Daini Hongo White Bldg. 201, 5-28-3 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan.

October 31, 2006

Abstract

By exploiting a generalization of recent results on environment-assisted channel correction, we show that, whenever a quantum system undergoes a channel realized as an interaction with a probe, the more efficiently the information about the input state can be erased from the probe, the higher is the corresponding entanglement fidelity of the corrected channel, and viceversa. The present analysis applies also to channels for which perfect quantum erasure is impossible, thus extending the original quantum eraser arrangement, and naturally embodies a general information-disturbance tradeoff.

In a simple double-slit interference experiment with matter beams, two basic aspects of Quantum Theory reveal their deeply counter-intuitive interplay: the first aspect is the complementarity of particle- and wave-like matter’s behaviours, which is strictly related to the non-commutative feature of Quantum Mechanics, and hence to all forms of “uncertainty relations” \[\text{\cite{1}}.\] The second aspect is the active role played by the observer in a quantum measurement process, since the bare presence of accessible information about the particle’s path irreversibly affects the result of the measurement, as interference fringes are destroyed. A great contribution in understanding these two points and their connection came from the so-called quantum eraser; namely,

*Email: buscemi@qci.jst.go.jp
Homepage: http://www.qci.jst.go.jp/~buscemi
a variety—proposed in Ref. [2] and realized in Ref. [3]—of the usual double-slit interference experiment, in which it is possible to mark either particle- or wave-like (exploring also halfway [4]) properties of the beam, long after the beam itself passed through the slits. This delayed choice is done by measuring one observable among a set of non-commuting observables of the probe, which, previous to the measurement, has been made suitably interact with the beam in order to store the which-path information. Such which-path information sits in the correlations established during the interaction between the beam and the probe. Since the experimenter—as it is usually said—could in principle retrieve such information, no fringes appear in the interference pattern. Still, the experimenter can choose to erase the which-path information from the probe, that is, if she measures a probe’s observable whose outcomes are completely independent of—i.e. complementary to—the which-path information, then—i.e. conditional on the probe observable outcome—fringes appear back again [2]. This procedure is called quantum erasure.

Even though in literature the use of the expression “quantum eraser” is limited to very specific double-slit experiment settings, one wonders whether the intuition the quantum eraser suggests, namely, that it may be possible to undo the effect of a quantum channel by erasing some information about its dynamics, is in fact a general feature provided by Quantum Mechanics or not. In the present paper, after briefly reviewing the original quantum eraser arrangement [2] and the theory of environment-assisted correction [5], we will show that it is indeed possible to extend the mentioned approach on a general basis. More explicitly, we will consider general quantum evolutions, mathematically described as channels—i.e. completely positive trace-preserving maps [6]—acting on the input system and physically modelled as unitary interactions of the input system with a probe, the latter playing the role of a controllable environment. We will then derive two inequalities relating the erasure efficiency with the entanglement fidelity of the corresponding corrected channel, showing that perfect erasure is equivalent to perfect correction and, even if perfect erasure is impossible, there exists a tradeoff relation between information erasure and channel correction efficacy. In this sense we can think that a sort of quantum erasure relation holds valid in all conceivable situations, also providing, as a byproduct, a quite general information-disturbance tradeoff relation.
The quantum eraser example.— Let a particle pass through a double-slit. Its state can be described in full generality by a density matrix $\rho$, such that, if the orthogonal states $|1\rangle$ and $|2\rangle$ correspond to the particle passing through the slit number 1 or number 2, respectively, the probability that the particle actually passed through the slit number 1 (2) is $p(1) = \langle 1|\rho|1\rangle$ ($p(2) = \langle 2|\rho|2\rangle$). Notice that $\rho$ can be a pure state, as in the original quantum eraser proposal $\rho = |+\rangle\langle +|$, with $|+\rangle = 2^{-1/2}(|1\rangle + |2\rangle)$. The probe, starting in the idle state $|0\rangle_p$, interacts with the particle by means of a controlled unitary $U(|i\rangle \otimes |0\rangle_p) = |i\rangle \otimes |i\rangle_p$, $i = 1, 2$. In other words, some of the degrees of freedom of the probe get entangled with the which-path information about the particle’s way. Then the experimenter could in principle measure the observable $\{|1\rangle\langle 1|, |2\rangle\langle 2|\}$ on the probe, extracting the which-path information, and actually interference fringes disappear from the overall pattern on the screen, because of the entanglement between the system and the probe. Moreover, no fringes appear even in the conditional (on the probe outcomes) patterns, since, if $p(i|j)$ denotes the conditional probability that the particle passed through the $i$-th slit given that the probe measurement produced the $j$-th outcome, we have $p(1|1) = p(1), p(2|2) = p(2)$, and $p(1|2) = p(2|1) = 0$, that is, cross terms are null.

Nevertheless, it is still possible to erase the which-path information stored in the probe by measuring on it the observable $\{|+\rangle\langle +|, |−\rangle\langle −|\}$, where $|−\rangle = 2^{-1/2}(|1\rangle − |2\rangle)$. In fact, the outcomes distribution of this probe observable are independent of the which-way information, that is, the conditional probabilities are $p(1|+) = p(2|+) = p(1−) = p(2−) = 1/2$. In other words, the measure of the observable $\{|+\rangle\langle +|, |−\rangle\langle −|\}$ on the probe perfectly erases all the which-way information. By simple calculations, one can check that the conditional output states of the system are $\rho_+ = \rho$ and $\rho_− = \sigma_z \rho \sigma_z$, namely, a part of an innocuous unitary rotation in the case of $\rho_-$, the system’s state has not been disturbed. That’s the reason why interference fringes appear again in the conditional patterns. Notice that both $\{|1\rangle\langle 1|, |2\rangle\langle 2|\}$ and $\{|+\rangle\langle +|, |−\rangle\langle −|\}$ are observables with rank-one elements: we will show this to be a general property of the erasing measurement.

Environment-assisted channel correction.— As noticed in Ref. [7], the quantum eraser can be well understood within the general theory of environment-assisted channel correction recently introduced by Gregoratti and Werner in Ref. [5]. They consider the most general situation, where an
input system $S$ (described by a Hilbert space $\mathcal{H}_S$ with $d := \dim \mathcal{H}_S < \infty$) in a state $\rho$ unitarily interacts with an environment $E$ initialized in a pure state $|0\rangle$. After the interaction took place, the system output state $\sigma$ is described by

$$\sigma = \text{Tr}_E[U(\rho \otimes |0\rangle\langle 0|)U^\dagger],$$

or, equivalently, by

$$\sigma := \mathcal{E}(\rho) = \sum_k E_k \rho E_k^\dagger,$$

where $\{E_k\}$ is a collection of operators such that $\sum_k E_k^\dagger E_k = I$, in order to preserve the normalization of states, i.e. $\text{Tr}[\mathcal{E}(\rho)] = 1$ for all $\rho$. Eq. (2) is usually referred to as the Kraus form of quantum channels [6], and it is not uniquely defined, as it is not uniquely defined the unitary interaction realizing the channel $\mathcal{E}$ in Eq. (1).

The environment is then assumed to be somehow “controllable” much like as a probe, in the sense that a quantum measurement, described in full generality by a positive operator valued measure (POVM) $M = \{P_j\}$, $P_j \geq 0, \sum_j P_j = I$, can be performed on it. In Ref. [5] it is proved that, given a channel $\mathcal{E}$, for all possible unitary interactions $U$ realizing it as in Eq. (1), and for all possible Kraus decompositions representing it as in Eq. (2), there always exists a POVM $M$ on $E$ with rank-one elements $\{P_j \equiv |\phi_j\rangle\langle \phi_j| := \phi_j\}$, such that

$$E_k \rho E_k^\dagger = \text{Tr}_E[U(\rho \otimes |0\rangle\langle 0|)U^\dagger (I \otimes \phi_k)].$$

Then, the probability of getting the $k$-th outcome is equal to $p(k) = \text{Tr}[E_k \rho E_k^\dagger]$, the conditional output state is $\sigma_k := E_k \rho E_k^\dagger / p(k)$, and both depend explicitly on the input system state $\rho$ and on the POVM $M = \{\phi_j\}$. In a sense, thanks to Gregoratti and Werner’s theorem, we do not have to worry about the particular form of the interaction $U$ realizing the channel. This picture is very useful, not only because it links two different representation theorems for quantum channels, but also because it makes easy to mathematically describe a feed-forward control from the environment $E$ onto the system $S$. This is obtained by allowing a controlled correction $C_j^\rho$ on the system, conditional on the outcome $j$ obtained from the environment measurement, and generally depending also on the input state $\rho$, as shown in Fig. . This is what we call input-dependent environment-assisted channel correction, actually a slight generalization of the input-independent correction scheme in Ref. [5].

A useful quantity to judge the capability of the channel $\mathcal{E}$ in faithfully and coherently transmitting an input state $\rho$, is given by the entanglement
Figure 1: Input-dependent environment-assisted correction: an input state \( \rho \) unitarily interacts with the pure environment state \(|0\rangle\). Then a rank-one POVM \( \{\phi_j\} \) is measured on the environment side and a conditional correction \( C^\rho_j \) is consequently applied to the conditional output \( \sigma_j \). The final corrected conditional output state is then \( C^\rho_j(\sigma_j) \). In Ref. [5] the correction does not depend on the input state.

Fidelity \( [8] \) \( F_e(\rho) \) defined as

\[
F_e(\rho) = \text{Tr}[|\Omega\rangle\langle\Omega| (E \otimes I)(|\Omega\rangle\langle\Omega|)],
\]

where \( I \) is the identity (ideal) channel and \(|\Omega\rangle\) is a purification of \( \rho \). If \( F_e(\rho) \) is close to one, then the channel \( E \) acts quite like the identity channel on the support of \( \rho \), [8]. With few calculations we find that

\[
F_e(\rho) = \sum_k |\text{Tr} \rho E_k|^2,
\]

and since it does not depend on the particular Kraus decomposition \( \{E_k\} \) chosen in Eq. (2), it is an intrinsic property of the channel. The following simple upper bound then comes from an application of a Cauchy-Schwartz–type inequality

\[
F_e(\rho) \leq \sum_k (\text{Tr} |\rho E_k|)^2 := F_{\text{corr}}(\rho) \leq 1,
\]

(4)

where \( |\rho E_k| \) is the positive part of the polar decomposition \( \rho E_k = U_k^\rho |\rho E_k| \), for unitary \( U_k^\rho \). With an input-dependent environment-assisted correction scheme, it is indeed possible to reach such an upper bound, that we therefore call \( F_{\text{corr}}(\rho) \): we have to choose the conditional correcting channels \( C^\rho_j \) in Fig. to be equal to the unitary channels \( C^\rho_j(\rho) = (U_j^\rho)^\dagger \rho U_j^\rho \), where \( U_j^\rho \) is the unitary part of the polar decomposition \( \rho E_j = U_j^\rho |\rho E_j| \). If \( F_{\text{corr}}(\rho) \) is close to one, it means that the corrected channel \( \sum_k C^\rho_k(\rho E_k) \) acts much like as the ideal channel on the support of \( \rho \). The tricky point now is that \( F_{\text{corr}}(\rho) \) does depend on the particular Kraus decomposition \( \{E_j\} \), that is, on the measurement \( \mathbf{M} = \{\phi_j\} \) performed upon the environment system. We will show that \( F_{\text{corr}}^\mathbf{M}(\rho) \) essentially determines how well the measurement \( \mathbf{M} \) erases from the environment the information about the input state \( \rho \), and viceversa—that is, \( F_{\text{corr}}^\mathbf{M}(\rho) \) and the erasure efficiency are equivalent measures. But before doing
that, we first need to quantitatively describe the information erasure.

**Information retrieval and erasure.**— First of all, let us fix some notation. Given a channel $E$ acting on states of the input system $S$, from Eq. (1) we can always construct the so-called complementary channel $\tilde{E}$, defined as $\tilde{E}(\rho) = \text{Tr}_S[U(\rho \otimes |0\rangle \langle 0|)U^\dagger]$. It describes the output state of the environment given that the input state of the system was $\rho$. As a consequence of the Stinespring theorem [9], such a complementary channel is unique up to a partial isometry [10]. Hence, we can consider $\tilde{E}$ as being the canonical complementary channel. Moreover, given a channel $E$ acting on states, there exists a unique dual channel $E'$ acting on observables $O$, defined by the trace relation $\text{Tr}[E(\rho) O] = \text{Tr}[\rho E'(O)]$, for all $\rho$. The trace-preserving condition becomes a unit-preserving condition, i.e. $E'(I) = I$. We then have four channels: the direct one, i.e. $E$; the dual one, i.e. $E'$; the complementary one, i.e. $\tilde{E}$; and the complementary dual one, i.e. $\tilde{E}'$.

If we send through the channel $E$ an ensemble of quantum states $\{\rho_i\}$, such that $\text{Tr}[\rho_i] = p(i)$, $\sum_i \rho_i = \rho$, and $\text{Tr}[\rho] = 1$, at the environment output branch will arrive $\{\tilde{E}(\rho_i)\}$. We then perform a measurement on them by using a rank-one POVM $M = \{\phi_j\}$, thus obtaining a joint probability distribution

$$p(i, j) = \text{Tr}[\tilde{E}(\rho_i) \phi_j] = \text{Tr}[\rho_i \tilde{E}'(\phi_j)].$$

(5)

In the following, we will consistently use the index $i$ for the input ensemble and the index $j$ for the environment outcomes. Notice that, by the already mentioned Gregoratti and Werner’s theorem [5], the choice of a rank-one POVM is not restrictive at all, and moreover it automatically rules out the possibility of a classical post-processing of data [11], which could artificially reduce the information transmission, while here we are interested in a genuinely quantum erasure process. Hence, such a choice is definitely the appropriate one.

In order to quantify the amount of information about the input ensemble $\{\rho_i\}$ that the measure of $M$ retrieves/erases from the environment, it is natural to compute the *mutual information* from $p(i, j)$ as

$$I_{S:E}^M = H(p(i)) + H(p(j)) - H(p(i, j)),$$

(6)

where $H(q(k)) = -\sum_k q(k) \log q(k)$ is the Shannon entropy of the probability distribution $q(k)$. If $I_{S:E}^M$ is close to zero, then $p(i, j) \approx p(i)p(j)$, namely,
the outcomes of the measurement $M$ on the environment are almost independent of the input ensemble $\{\rho_i\}$. It means that the information transmission is poor, and if the same holds for all possible ensemble realizations $\{\rho_i\}$ of $\rho$, we say that the measurement $M$ performs a good erasure with respect to the input state $\rho$. Notice here that the mutual information equals the relative entropy $D(p(i,j)\|p(i)p(j))$ between the joint distribution $p(i,j)$ and the factorized one $p(i)p(j)$, where $D(r(k)\|s(k))$ is defined for two probability distributions $r(k)$ and $s(k)$ as

$$D(r(k)\|s(k)) = \sum_k r(k) \log \frac{r(k)}{s(k)},$$

[12]. The relative entropy quantifies the error we make when encoding a message, drawn according to $r(k)$, as if it were drawn according to $s(k)$. Notice that, if $s(k) = 0$ for some $k$, then $D(r(k)\|s(k))$ diverges. In our case, however, this will never be the case and the following inequalities will play a central role [12, 13, 14]

$$2^{-1} \|r(k) - s(k)\|_1^2 \leq D(r(k)\|s(k)) \leq \beta^{-1} \|r(k) - s(k)\|_1^2,$$  

(7)

where $\beta = \min_k s(k) > 0$, and $\|r(k) - s(k)\|_1 = \sum_k |r(k) - s(k)|$.

**Main result.**— Let us now fix the input state $\rho$ with ensemble realization $\{\rho_i\}$ and the environment POVM $M = \{\phi_j\}$. Then $K_j = \rho^{1/2} \tilde{E}^j(\phi_j) \rho^{1/2}/p(j)$ turn out to be normalized states, for all $j$, with $\sum_j p(j)K_j = \rho$. Moreover, by noticing that $\tilde{E}^j(\phi_j) = E_j^\dagger E_j$, [15], we can rewrite the upper bound in Eq. (4) as

$$F_{\text{corr}}^M(\rho) = \sum_j p(j) F(\rho, K_j)^2,$$  

(8)

where $F(\rho, \sigma) := \text{Tr}[\sqrt{\sqrt{\rho} \sigma \sqrt{\rho}}]$ is the Uhlmann’s fidelity between two mixed states [16]. By exploiting the well-known relation [13] between the fidelity and the trace norm of the difference, defined as $\|\rho - \sigma\|_1 = \text{Tr}[|\rho - \sigma|]$, that is, $F(\rho, \sigma)^2 \leq 1 - 2^{-2}\|\rho - \sigma\|_1^2$, together with Eq. (7), we obtain the following
chain of inequalities

\[ F_M^{\text{corr}}(\rho) \leq 1 - 2^{-2} \sum_j p(j) \| \rho - K_j \|^2 \]

\[ \leq 1 - 2^{-2} \sum_j p(j) \left( \sum_i |p(i) - p(i|j)| \right)^2 \]

\[ \leq 1 - 2^{-2} \beta \sum_j p(j) D(p(i|j)||p(i)) \]

\[ = 1 - 2^{-2} \beta I_{S,E}^M \leq 1, \] (9)

where \( \beta = \min_i p(i) > 0 \). In the second inequality we used the fact that the trace distance between two states is never smaller than the trace distance between the probability distributions obtained by measuring the same POVM \( \{\rho^{1/2}\rho_i\rho^{-1/2}\} \) on both states; notice that \( \{\rho^{1/2}\rho_i\rho^{-1/2}\} \) is a well-defined POVM on the support of \( \rho \), since \( \rho^{1/2}\rho_i\rho^{-1/2} \geq 0 \) and \( \sum_i \rho^{1/2}\rho_i\rho^{-1/2} = I|_{\text{Supp}(\rho)} \). In the last equality we used the trivial identity \( \sum_j p(j) D(p(i|j)||p(i)) = D(p(i, j)||p(i)p(j)) \). Notice moreover that inequality (9) holds for every ensemble realization \( \{\rho_i\} \) of \( \rho \).

Equation (9) informs us that if \( F_M^{\text{corr}}(\rho) \) is sufficiently close to one (but it can be also strictly less than one) for a particular environment POVM \( M \), then the corresponding information transmission from the system to the environment is close to zero for every ensemble realization \( \{\rho_i\} \) of \( \rho \), that is, the measurement \( M = \{\phi_j\} \) is erasing the information about the input state \( \rho \) registered into the environment during the interaction. Equivalently, non-null information extraction always causes disturbance on the input ensemble, even allowing input-dependent environment-assisted correction schemes. Our approach hence embodies a quite general information-disturbance trade-off [18].

Also the converse statement is true. In this case we should check that, for a suitable environment rank-one POVM \( M = \{\phi_j\} \), the information transmission is poor for all possible ensemble realizations \( \{\rho_i\} \) of \( \rho \), that is, the measurement \( M = \{\phi_j\} \) is erasing the information about the input state \( \rho \) registered into the environment during the interaction. Equivalently, non-null information extraction always causes disturbance on the input ensemble, even allowing input-dependent environment-assisted correction schemes. Our approach hence embodies a quite general information-disturbance trade-off [18].
reconstruction formula then holds

\[ O = \sum_i \text{Tr}[O, \rho_i^{-1/2} \rho_i \rho_i^{-1/2}] \rho_i', \tag{10} \]

where the operators \( \rho_i' \)'s are limited but neither positive definite nor semi-definite, in general. Then, by exploiting the well-known inequality \( F(\rho, \sigma)^2 \geq 1 - \| \rho - \sigma \|_1 \), \[13\], we have the following relations

\[ F_{\text{corr}}^M(\rho) \geq 1 - \sum_j p(j) \| \rho - K_j \|_1 \]

\[ = 1 - \sum_{ij} p(j) \| p(i) \rho_i' - p(i|j) \rho_i' \|_1 \]

\[ \geq 1 - |\Gamma| \sum_{ij} p(j) |p(i) - p(i|j)| \]

\[ \geq 1 - \sqrt{2|\Gamma|} \sqrt{I_{S:E}^M}, \tag{11} \]

where \( |\Gamma| = \max_i \| \rho_i' \|_1 < \infty \). In the first equality we used the reconstruction formula \[10\]. Notice that, if \( I_{S:E}^M \) is sufficiently close to zero for one particular informationally complete input ensemble realization \( \{ \rho_i \} \), then \( F_{\text{corr}}^M(\rho) \) is close to one, and, by Eq. \[9\], \( I_{S:E}^M \) is also close to zero for all possible input ensemble realizations of \( \rho \). The implications then turn out to be equivalences.

We can hence conclude, by stating that for every channel realized as an interaction of an input state \( \rho \) with a probe, even if perfect quantum erasure is impossible, the more a given POVM erases from the probe the information about the input state (for all its possible ensemble realizations) stored during the interaction, the closer (on the support of \( \rho \)) the corresponding corrected channel is with respect to the ideal one, and viceversa \[20\]. To find the optimal erasure measurement for a given channel and a given input state remains an open problem. Incidentally, it is worth noticing that if we consider \( \rho = I/d \), then the corresponding input-dependent environment-assisted correction coincides with the one presented in Ref. \[5\], and the quantum capacity of the corresponding corrected channel is maximum over the whole input Hilbert space \( \mathcal{H}_S \). It is possible to achieve perfect erasure, that is \( F_{\text{corr}}^M(I/d) = 1 \)—indeed just an invertible \( \rho \) suffices—if and only if the channel admits a random-unitary decomposition, that is, \( \mathcal{E}(\rho) = \sum_j p(j) U_j \rho U_j^\dagger \), for some probability distribution \( p(j) \) and unitary operators \( U_j \)'s. In this
case, for the corresponding measurement $M$, $I_{S,E}^M$ is rigorously zero for every possible input ensemble $[5, 21]$.

Acknowledgments. — Discussions with M Hayashi, L Maccone, and M F Sacchi are gratefully acknowledged. Japan Science and Technology Agency is acknowledged for support through the ERATO-SORST Quantum Computation and Information Project.

References

[1] Here we are not interested in whether complementarity is “more fundamental” than uncertainty relations or vice versa. However, there was a long controversial discussion about this point. See: M O Scully, B-G Englert, and H Walther, Nature 351, 111 (1991); P Storey, S Tan, M Collett, and D Walls, Nature 367, 626 (1994); B-G Englert, M O Scully, and H Walther, Nature 375, 367 (1995); P Storey, S Tan, M Collett, and D Walls, Nature 375, 368 (1995); H Wiseman and F Harrison, Nature 377, 584 (1995).


[15] This can be simply seen since, from Eq. (3),
\[ \text{Tr}[\rho \tilde{E}(\phi_j)] = \text{Tr}[E_j \rho E_j^\dagger], \]
for all \( \rho \).


[20] Formally speaking, the two quantities \( (1 - F_{\text{corr}}^M(\rho)) \) and
\[ \mathcal{I}(\rho) := \max_{\{\rho_i\}, \sum_i \rho_i = \rho} I_{S:E}^M \]
are equivalent, since, from Eqs. [9] and [11],
\[ a\mathcal{I}(\rho)^2 \leq (1 - F_{\text{corr}}^M(\rho))^2 \leq b\mathcal{I}(\rho), \]
with \( 0 < a < b < \infty \).