Error filtration and entanglement purification for quantum communication

N. Gisin
Group of Applied Physics, 20, rue de l'Ecole-de-Médecine, CH-1211 Geneva 4, Switzerland

N. Linden
Department of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, United Kingdom

S. Massar
Laboratoire d'Information Quantique and Centre for Quantum Information and Communication (QUIC), C.P. 165/59, Université Libre de Bruxelles, Avenue F. D. Roosevelt 50, 1050 Bruxelles, Belgium

S. Popescu
H. H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, United Kingdom and Hewlett-Packard Laboratories, Stoke Gifford, Bristol BS12 6QZ, United Kingdom
(Received 12 January 2005; published 28 July 2005)

The key realization that led to the emergence of the new field of quantum information processing is that quantum mechanics, the theory that describes microscopic particles, allows the processing of information in fundamentally new ways. But just as in classical information processing, errors occur in quantum information processing, and these have to be corrected. A fundamental breakthrough was the realization that quantum error correction is in fact possible. However, most work so far has not been concerned with technological feasibility, but rather with proving that quantum error correction is possible in principle. Here we describe a method for *filtering* out errors and *entanglement purification* which is particularly suitable for quantum communication. Our method is conceptually new, and, crucially, it is easy to implement in a wide variety of physical systems with present-day technology and should therefore be of wide applicability.

DOI: 10.1103/PhysRevA.72.012338 PACS number(s): 03.67.Pp, 03.67.Hk, 03.67.Mn

I. INTRODUCTION

When quantum communication was first proposed, it was felt that interactions of the system with the environment, and the consequent loss of information into the environment, would produce errors which would be uncorrectable even in principle. However, it was discovered using two independent approaches, namely error-correction codes for quantum memories [1,2] and entanglement purification [3], that quantum error correction is in fact possible. This transformed the field from an intellectual game into a potentially revolutionary new technology.

In these pioneering papers and their extensions (see, for instance [4–7]), the authors were not concerned with immediate technological feasibility, but rather with proving a point of principle. And the difficulty with all these protocols is that in order to be implemented they require controlled interactions between many particles. This is technically impractical at present and it is likely to remain so for the foreseeable future. For this reason, several authors have proposed methods for error correction which are comparatively simpler to implement [8–12]. However, all these methods require special resources such as squeezed states [8,9] or particular entangled states [10,11] or optical memories [12] that are at the limit of present technology even for proof of principle, let alone for practical schemes. The difficulty of realizing error correction is well illustrated by the complexity of a recent experimental realization [13].

Here we describe a different approach to dealing with errors that work in a conceptually different way from existing error-correction methods. Our method realizes *error filtration* and is particularly useful for quantum communication. It can easily be implemented with present-day technologies and has therefore the potential to make a significant impact on the nascent field of quantum communication. In the longer term, since our methods are applicable to any type of communication, they may find uses inside quantum computers and other quantum devices, if and when these are built.

The difference between error correction and error filtration is the following. In error *correction*, the aim is actively to correct the errors that occur during transmission so that the decoded signal is as close as possible to the emitted signal. In error *filtration*, the aim is to detect with high probability when an error has occurred, and in that case to discard the signal. In effect, what this method does is transform a general error (phase noise, depolarization, etc.) into an erasure, which is far more benign. Our scheme can be thought of as a form of error detection [14–16].

We present two methods: one for improving a channel which is useful for distributing arbitrary states (channel multiplexing), the other for distributing a standard entangled state (source multiplexing). Of course, the first method can also be used for distributing entangled states, but when we *know* what entangled state we want we can also manipulate the source. Thus our methods also provide new ways of *purifying* entangled states. The mapping between these two problems follows the general correspondence between error correction and entanglement purification without two-way classical communication discussed in [4].

The simplicity of implementation of error filtration is underlined by an experiment, reported elsewhere [17], in which a particular scheme is implemented.

In Sec. II we introduce our method in as simple and intuitive way as possible. This should provide the reader with the background to understand the more detailed examples that follow. Then in Sec. III we analyze in detail an example of implementation of error filtration in the case of a simple particle and in Sec. IV we analyze in detail an example of implementation of error filtration in the case of two entangled particles. Finally in the Appendices we extend the protocols presented in the main text in a number of ways: We consider protocols with more general encoding and decoding operations (Appendices A and B), protocols where the particle has internal degrees of freedom (Appendix C), protocols in which error filtration is used in series (Appendix D), error filtration for signals comprising many particles (Appendix E) and for classical wave signals (Appendix F), and general protocols for the transmission of entangled states (Appendix G).

II. BASIC PRINCIPLE

The central idea of the first method is what we will call *channel multiplexing*, i.e., to use more transmission channels than the minimum necessary to send the quantum state.

The main concept of our method is extremely simple and we illustrate it by means of an example. Consider a quantum system that propagates through a communication channel, such as a photon going through an optical fiber. During the propagation, errors can occur; it is these errors that we want to identify and get rid of. Our method consists of replacing the original single communication channel by an interferometer consisting of two communication channels in parallel. Now, instead of sending the photon through the original single channel, we send it in a superposition of two states, one going through each arm of the interferometer, i.e., in a superposition of going through the two channels (see Fig. 1).

The errors in each channel are arbitrary, i.e., we impose no restrictions on the properties of the individual channels. We will, however, arrange things so that the errors on the different channels are independent. This is a simple matter of engineering. (For instance, we can always increase the space separation of the channels.)

When emerging from the two channels, the two wave packets interfere. The output beam splitter is tuned in such a way that when no error occurs, the photon emerges with certainty in one of the output channels; we call this the "useful" output. That is, we arrange the interferometer so that, in the absence of errors, there is complete constructive interference in one output channel and complete destructive interference in the other output channel.

In the presence of errors, the output in the useful channel is better than if we had not multiplexed. The reason for this may be understood as follows.

Let us first consider that all the channels are similar, i.e., that they produce the same amount of errors. The total probability for an error is the same whether we send the photon only through one single channel, as in the original scheme,

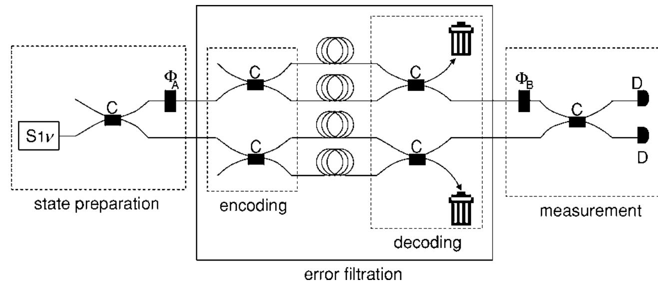

FIG. 1. Implementation of error filtration using multiple optical fibers. A source ($S_{1\nu}$) produces a single photon that is coupled into an optical fiber. The photon is split into two using a fiber coupler (C). Note that an arbitrary state in a two-dimensional space (a qubit) can be prepared in this way by changing the coupling ratio of the coupler and by modifying the phase ϕ_A. In order to protect the state against noise, each basis state is multiplexed into two transmission states using 50/50 couplers. A single qubit is thus encoded into four transmission states, each traveling through a separate fiber. Because the photon cannot jump from one fiber to the other, it will only be affected during transmission by phase noise. Furthermore the noise in the different fibers will be independent. Hence the noise is of the type (independent phase noise on the different transmission channels) studied in the main text. The decoding is the reverse of the encoding procedure: two transmission states are combined into one receiver state using fiber couplers. This is done in such a way that in the absence of noise, constructive interference occurs and the photon always emerges in the receiver states. Due to the noise, the photon may not emerge in the receiver states, in which case it is discarded. But if the photon emerges in the receiver states, then the noise has been filtered out. The receiver can then use the filtered state. For instance he may, as described in the text, test the quality of the filtered state by carrying out the measurement shown (D, single photon detector). The measurement basis is chosen by varying the phase ϕ_B and the coupling ratio of the measurement coupler. Note, however, that the measurement step is included in the figure for illustration only—it is not part of the filtration protocol per se. The receiver can use the filtered signal for other purposes.

or in a superposition of going through the two channels. (Indeed, we do not consider two photons, one going through one channel, the other going through the other channel, each accumulating errors, but a single photon, going either in one channel or in the other.)

There are two complementary ways of describing noise [18,19]. The first is to view the system as undergoing an evolution which depends on random parameters. The second more general approach is to describe the system and its environment as a single combined system. This combined system evolves unitarily, and the noise manifests itself as entanglement between the system and the environment. In the second description, which we adopt here, whenever an error occurs there is a registration in the environment. Due to the fact that the channels are independent (as described above), when an error affects the photon going through the first channel, it is the state of the first channel that it is affected (i.e., that registers the fact that the error occurred). Similarly, when the photon goes through the second channel and an error occurs, it is the state of the second channel that is affected.

Thus, whenever an error occurs, one could, in principle, by looking at the state of the two channels, find out which channel the photon went through. This causes the original superposition (of the photon going through both channels) to collapse. Hence, at the output beam splitter, the two wave packets no longer interfere, and the photon may sometimes end up at the useful output, while other times it can end up at the other output. In this latter case, we *know* that an error occurred—otherwise the photon could not have ended up there—so we discard it. We thus get rid of some errors. And since multiplexing did not increase the total number of errors, we end up with fewer errors.

Once this simple principle is understood, it is clear that many variations are possible. For example, in Fig. 1 we view the channels as being different optical fibers. The same scheme works for free space propagation of photons, or for propagation of electrons through mesoscopic wires. Also two "separate channels" need not be implemented as two different physical objects. For example, one can use a single optical fiber and represent the different channels by different "time bins," see Fig. 2 and [20]. In this case, of course, one has to take care to make the channels independent; making the time bins sufficiently separated in time can easily achieve this goal.

Obviously, we can use a greater degree of multiplexing, i.e., replace one channel not by two but by many. This further decreases the amount of noise in the useful output. We also can cut a channel into pieces and use filtration on each piece separately, in series, one after the other. It is also not critical that the different channels have the same amount of noise (as we assumed for simplicity above).

It is also clear that there is a lot of freedom in the relative phase of the superposition of the photon wave packets going through the different channels. All that is required is that, in the absence of noise, at the output there is constructive interference so that the photon emerges with certainty in one particular output channel (the "useful" output port). This depends, however, on both the input into the interferometer and on the output. We can arrange any input, and then tune the output accordingly. (In some examples below, we arrange the superposition by Hadamard transform; in other examples, we use the Fourier transform; but these are just examples from the large family discussed in Appendices A, B, and C).

Finally it is clear that the method will work independently of how the quantum information is encoded in the particle. For instance, it may be encoded in the relative phase between time bins (as in Fig. 2), or an internal degree of freedom such as the polarization of the photon or spin of an electron. In all cases, the method will improve the output signal.

Error filtration is a general framework, appropriate to a wide variety of physical implementations and types of noise. We will present a number of examples in detail below and present further extensions in the Appendices.

Among other ingredients, our scheme uses multiple paths to encode the quantum state of a single quantum system; that this could be of interest in the context of quantum information processing and communication has been noted in [21].

We note that some implementations of our method have superficial similarities to the symmetrization method described in [22] which also realizes a form of error detection.

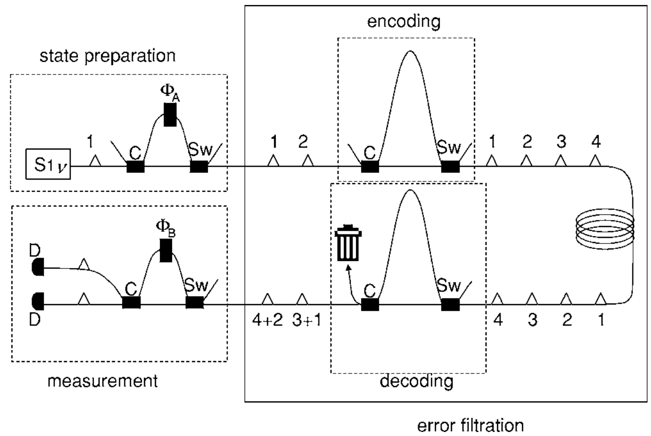

FIG. 2. Implementation of error filtration using time bins propagating in optical fibers. A source produces a single photon in time bin $|1\rangle$. A first Mach-Zender (MZ) interferometer produces a state in a two-dimensional Hilbert space as follows: the fiber coupler C splits the time bin $|1\rangle$ into two pulses which follow the short and long arm of the interferometer. Then the switch (Sw), synchronized with the source, is used to direct the pulses exiting from the first MZ interferometer into the fiber leading to the encoder. In this way, a superposition of two time bins $(|1\rangle + e^{i\phi_A}|2\rangle)/\sqrt{2}$ is produced where the phase ϕ_A encodes the quantum information that must be transmitted. A second MZ interferometer realizes the encoding part of the error filtration protocol: it multiplexes time bin 1 (2) exiting from the state preparation into time bins 1 and 3 (2 and 4). Thus after encoding, the qubit is a superposition of four time bins $(|1\rangle + |3\rangle + e^{i\phi_A}|2\rangle + e^{i\phi_A}|4\rangle)/2$. During transmission, the state is affected by noise. If the time bins are sufficiently separated, then the probability that a photon jumps from one time bin to the other is negligible and the noise affecting each time bin will be essentially independent. Hence one is in the case of independent phase noise considered in the main text. The decoding operation is the reverse of the encoding operation. A first MZ interferometer projects the state onto the $|1\rangle + |3\rangle$, $|2\rangle + |4\rangle$ subspace. The receiver may then use the filtered state. For instance, he can carry out a measurement using a second MZ interferometer. The measurement basis is chosen by varying the phase ϕ_B.

Their are, however, some essential differences due to the fact that the symmetrization method uses multiple copies of the system, whereas we use a single copy. Indeed, the number of errors is proportional to the number of copies and is therefore much larger in the symmetrization method than in our method. Second symmetrizing multiple copies of a quantum system is extremely difficult to implement and does not present practical advantages over implementing error-correction codes.

III. ERROR FILTRATION FOR A SINGLE PARTICLE

Let us consider a source which produces a signal encoded in the quantum state of a single particle: $|\psi\rangle = \sum_{l=1}^{S_{\text{tot}}} c_l |l\rangle_S$, where c_l is the complex amplitude that the particle is in source channel l and $|l\rangle_S$ denotes the state of the particle if it is in source channel l. Let us suppose that there are a certain number T_{tot} of transmission channels with $T_{\text{tot}} = TS_{\text{tot}}$ a multiple of S_{tot}. Denote by $|j\rangle_T (j=1,\ldots,T_{\text{tot}})$ the state of the particle if it is in transmission channel j.

During transmission, the states $|j\rangle_T$ will be affected by noise. For the sake of illustration, we consider the simple, but experimentally relevant (see Fig. 1 and [17]), case where the noise affecting the different states is phase noise. Furthermore, by construction, we can arrange that the noise in the different channels is independent. As mentioned in the preceding section, this does not depend on the intrinsic nature of the channels, but on the way we put the channels together, and it is, in general, rather simple to do.

Then the states $|j\rangle_T$ evolve according to $|j\rangle_T \mapsto e^{i\phi_j}|j\rangle_T$, where ϕ_j are independent random variables. An equivalent description of the noise is to suppose that during transmission the particle interacts with the environment as follows:

$$|j\rangle_T|0\rangle_E \mapsto |j\rangle_T(\alpha_j|0\rangle_E + \beta_j|j\rangle_E), \quad |\alpha_j|^2 + |\beta_j|^2 = 1, \quad (1)$$

where the environment states $|j\rangle_E$ are orthogonal for different j. This represents the physical situation in which the environment starts in the initial state $|0\rangle_E$ and the interaction causes the state of the environment to be disturbed. The amplitudes α_j and β_j describe the amount of disturbance; in general, these parameters will depend on j. For simplicity, we first consider the case in which they are the same on all channels, i.e., α_j and β_j are independent of j.

We suppose that each source channel is encoded and decoded separately. We can therefore focus on a particular one, $|1\rangle_S$, say. The source state $|1\rangle_S$ is encoded into T transmission channels,

$$|1\rangle_S \mapsto U_e|1\rangle_S = \frac{1}{\sqrt{T}}\sum_{j=1}^{T}|j\rangle_T, \quad (2)$$

where we have taken U_e to be the discrete Fourier transform. The noise now occurs, causing the state of system plus environment to change to $(1/\sqrt{T})\Sigma_{j=1}^{T}|j\rangle_T(\alpha|0\rangle_E+\beta|j\rangle_E)$. We decode by performing the inverse Fourier transform, $|j\rangle_T \mapsto U_d|j\rangle_T=(1/\sqrt{T})\Sigma_{k=1}^{T}e^{2i\pi j(k-1)/T}|k\rangle_R$, where $|k\rangle_R$ denotes the state in receiver channel k. Such encoding and decoding transformations can easily be easily realized using linear elements such as beam splitters and phase shifters [23]. One relevant figure of merit is simply the number of such elements which are required to realize the encoding and decoding operations. The Fourier transform which we use here for illustrative purposes is not necessarily optimal in this respect.

One easily computes that the state after decoding is

$$|1\rangle_R\left(\alpha|0\rangle_E + \frac{\beta}{\sqrt{T}}|\tilde{1}\rangle_E\right) + \frac{\beta}{T}\sum_{j=1}^{T}\sum_{k=2}^{T}e^{2i\pi j(k-1)/T}|k\rangle_R|j\rangle_E, \quad (3)$$

where $|\tilde{1}\rangle_E=(1/\sqrt{T})\Sigma_{j=1}^{T}|j\rangle_E$ is a normalized state of the environment. We see that this state has a component in the receiver channel $|1\rangle_R$, which we regard as the "useful" signal, and components in all the other receiver channels, $|k\rangle_R$, $k \neq 1$, which we discard. The useful signal is $|1\rangle_R[\alpha|0\rangle_E + (\beta/\sqrt{T})|\tilde{1}\rangle_E]$. The norm of this state gives the probability that the state appears in the useful channel, $P_{\text{success}}=|\alpha|^2 + (|\beta|^2/T)$.

In summary, the signal does not always reach the useful receiver channel. But when it does, the noise amplitude is reduced by a factor of $\sqrt{T}$. As an illustration, consider the case where the number of source states $S_{\text{tot}}=2$ and the state to be transmitted is $|\psi\rangle=|1\rangle_S+e^{i\Phi}|2\rangle_S$. The ability to preserve the phase Φ is a measure of the quality of the transmission. This is often quantified by the visibility V of the interference fringes seen by the receiver if he measures in the $|1\rangle_R \pm |2\rangle_R$ basis. One finds $V=T|\alpha|^2/(T|\alpha|^2+|\beta|^2)$, which equals $|\alpha|^2$ for $T=1$ (no multiplexing) and increases monotonically to 1 as the amount of multiplexing T tends to infinity.

Above, we considered for simplicity that all the channels were identical, but the method works equally well if the channels are different. This is an essential property since it shows that the method is robust against perturbations. To prove this, suppose that each channel j is characterized by parameters α_j and β_j which obey $|\alpha_j|^2+|\beta_j|^2=1$ and use the same encoding and decoding procedure as above. The amplitude of the quantum state if it ends up in the useful receiver channel [the analogue of the first term of Eq. (3)] is

$$|1\rangle_R\left(\bar{\alpha}|0\rangle_E + \frac{1}{T}\sum_{j=1}^{T}\beta_j|j\rangle_E\right),$$

where $\bar{\alpha}=(1/T)\Sigma_{j=1}^{T}\alpha_j$. The probability that the signal ends up in the useful receiver channel and is unaffected by error is $P_{\text{success\&no error}}=|\bar{\alpha}|^2$, whereas the probability that the signal ends up in the useful channel and is affected by error is $P_{\text{success\&error}}=(1/T^2)\Sigma_{j=1}^{T}|\beta_j|^2 \leq (1/T)(1-|\bar{\alpha}|^2)$. Let us suppose that the α_j all have approximately the same phase (this can easily be arranged by putting a phase shifter in each channel) and that $\bar{\alpha}$ does not tend to zero as the degree of multiplexing increases (this corresponds to supposing that as we add channels they do not become worse). Then the above result shows that the ratio of the probabilities of errors to no errors in the useful output channel decreases as $1/T$, i.e., error filtration works equally well when the transmission channels are not all identical.

An interesting question concerns how our methods scale. There are are a number of different issues to be considered. For example, (a) for fixed length L, how the intensity of the signal changes when we improve fidelity; (b) for a fixed output fidelity, how the signal changes when we increase L; (c) for fixed output fidelity, how the resources required for multiplexing increase as we increase L.

(a) For a fixed length of the communication channel, when we increase the multiplexing in order to increase the fidelity to 1, the total output signal decreases towards a fixed value that depends on the quality of the transmission channel. This value is nothing else than the probability that the signal is not affected by noise in the original, nonmultiplexed channel. In other words, as we increase the filtering power of our method (by increasing the multiplexing factor), we identify the errors better and better and throw away a larger fraction of them. Eventually (in the limit of infinitely many multiplexing channels), a perfect filter will yield a perfectly clean signal (fidelity=1) while throwing away all the errors, but not more than that. [For example, if the noise is phase noise, as in Eq. (1), then the probability of an error not to occur is $|\alpha|^2$. In the limit that the fidelity becomes 1, the probability of receiving a signal tends to $|\alpha|^2$.]

(b) Each unit of length has equal probability of producing an error. Thus the probability for a signal to survive without being affected by noise (and hence to pass our filtration) decreases exponentially with L. This is an inevitable feature of an error filtering method as opposed to one in which errors are corrected. In effect, what our procedure does is to transform a general error (phase noise, depolarization, etc.) into an erasure, which is well known to be a considerable advantage.

(c) The resources required to achieve a fixed fidelity as the length L increases grow polynomially in L. This polynomial scaling is achieved by applying error filtration in "series." By this we mean that the signal is encoded and after a short distance decoded and the noise filtered out. The signal is then re-encoded and re-decoded many times until the end of the communication channel is reached. Suppose that the error rate per unit distance is γ so that the probability that the signal is unaffected by error is $|\alpha|^2 = e^{-\gamma L}$. Suppose that the communication channel has length L, that the signal is encoded and decoded a total of Q times, and that the degree of multiplexing is T. Then using Eq. (3), one can show that the probability that the signal appears at the useful output and is affected by error is (see the Appendix D)

$$\left[e^{-\gamma L/Q} + \frac{1-e^{-\gamma L/Q}}{T}\right]^Q - e^{-\gamma L}. \tag{4}$$

On the other hand, the probability that the signal appears at the useful output and is not affected by error is $e^{-\gamma L}$. Hence one easily deduces from this result that one can maintain a fixed desired high fidelity of the output signal as we increase L by increasing both the multiplexing factor T and the number of filtering units Q linearly with L. In other words, the resources required to maintain a fixed fidelity of output signal scale polynomially with the length of the channel. (The need to repeat the filtration step is very similar to existing error-correction methods where one also needs to perform the correction step many times, each time before the error probability becomes too large. If the correction—or in our case filtering—is not performed a number of times but only once, then the resources needed to obtain a high fidelity increase exponentially with L.)

Note, however, that, important as it is, asymptotic scaling is not always the most relevant issue in practice. In practice, one always deals with a fixed range of distances, and the main question is, what is the advantage that a given method yields for that range? This is determined by the asymptotic formula but also by the precise values of the relevant parameters. Thus, for instance, the BB84 quantum cryptography protocol becomes insecure when the fidelity of the communication channel is below 85%, and a singlet state becomes an unentangled Werner density matrix when its fidelity is below 50%. These figures show that degradation of signal with serious consequences can occur with relatively low levels of noise. With a modest level of multiplexing, our schemes can reduce the noise in communication so as to bring states above these important thresholds.

The protocols presented above are generalized in Appendices A–F which are organized as follows:

Appendix A: General protocol for error filtration. The protocol described in the main body of the article took each source channel and multiplexed it into T transmission channels. The encoding and decoding transformations were the Fourier and inverse Fourier transforms, respectively. In Appendix A we generalize this protocol to other encoding and decoding transformations, and derive a condition for the encoding/decoding to remove as much noise as possible.

Appendix B: Protocol for error filtration with collective encoding. In the previous protocols a given source channel is encoded into a subset of the transmission channels, but the signal carried in a given transmission channel came from a single source channel. Here we show how it is possible to generalize these ideas by allowing each transmission channel to carry signals from more than source channel.

Appendix C: Error filtration for particles with internal degrees of freedom. All the previous protocols can be generalized to the case where the particle has "internal" degrees of freedom. By this we mean that each source and transmission channel has a state space which is a Hilbert space of dimension larger than one.

Appendix D: Using error filtration in series. The previous protocols can loosely be described as using transmission channels in parallel to achieve noise filtration. We may also use the idea of multiple channels in series to filter noise. Here we discuss this question in detail and in particular derive Eq. (4).

Appendix E: Quantum multiexcitation protocol. In the previous protocols the channels contained only a single excitation each. In the case of bosons, we can also consider the situation where each channel contains many quanta. We show that error filtration also works in this case. We first illustrate this by considering the case that the channel states are coherent states; at the end of Appendix E we show how this generalizes to general multiexcitation states.

Appendix F: Error filtration for classical wave signals. A particular case of the multiexcitation protocol considered in Appendix E is the case where the number of quanta is macroscopic and one is dealing with classical wave signals. We show that error filtration can also be used to reduce noise when transmitting classical wave signals. In addition to the types of noise discussed in the previous appendices (phase noise, noise affecting internal degrees of freedom) error filtration can also filter out other types of noise such as amplitude noise (i.e., noise that affects the amplitude of the wave) or even nonlinear noise (when the amount of noise depends on the intensity of the signal). We illustrate how error filtration works for these other types of noise in the case of classical signals.

IV. ERROR FILTRATION FOR TRANSMISSION OF ENTANGLED STATES

The previous protocols illustrated error filtration for one-way communication of quantum signals. We now show that these ideas may be extended to provide protocols when the goal is to send some *known* entangled quantum state from a source to two or more parties. First, it is obvious that this may be achieved, and the errors diminished, by using the previous quantum protocols: one of the parties simply prepares the state locally and transmits it using the one-way communication channels we have described earlier. However, since the task is to distribute *known* entangled states, i.e., those produced by us, not some external process out of our control, it is possible to proceed differently by modifying not only the transmission channels but also the source that produces the state. Our method essentially calls for using a source that produces the same number of entangled particles as the original, but in states with more entanglement than what is ultimately needed. We call this method "source multiplexing." It provides a new method for *entanglement purification*.

Consider the case in which the aim is for two receivers, A and B, to share a quantum state ρ_m, which is as close as possible to the maximally entangled state of dimension m: $|\psi_m^R\rangle=(1/\sqrt{m})\Sigma_{j=1}^m|j\rangle_A\otimes|j\rangle_B$. To improve the fidelity, we use a source that produces two entangled particles in the maximally entangled state of dimension n

$$|\psi_n\rangle=\frac{1}{\sqrt{n}}\sum_{j=1}^{n}|j\rangle_A\otimes|j\rangle_B, \tag{5}$$

where $n\geqslant m$, i.e., we start with more entanglement than we want to end up with.

Again, for simplicity, we consider the case in which the dominant errors are of the form (phase noise)

$$|j\rangle_A|e_0\rangle_A\mapsto|j\rangle_A(\alpha|e_0\rangle_A+\beta|e_j\rangle_A), \tag{6}$$

where $|e_0\rangle$ and $|e_j\rangle$ are states of the environment; α and β are complex numbers satisfying $|\alpha|^2+|\beta|^2=1$. The dominant errors for party B are also of this form. For different states of the system, the disturbed states of the environment are orthogonal $({}_A\langle e_j|e_k\rangle_A={}_B\langle e_j|e_k\rangle_B=0,\quad j,k=1\ldots n,\quad j\neq k;$ ${}_A\langle e_j|e_k\rangle_B=0\ \forall j,\ k)$. α and β describe the amount of disturbance; for simplicity, we have taken them to be independent of j.

One may easily show that after going through the noisy channel, the reduced density matrix of the two particles becomes

$$\rho_n=pP_{\psi_n}+\frac{1-p}{n}\sum_{j=1}^{n}P_{jj}, \tag{7}$$

where $P_{\psi_n}=|\psi_n\rangle\langle\psi_n|$, $P_{jj}=|jj\rangle\langle jj|$, and $p=|\alpha|^4$ is the probability that the state is not affected by phase noise.

Let us suppose that the parties do not carry out error filtration. Then the dimensions n and m of the senders and receivers state are equal. We characterize the quality of the receivers state by the fidelity F_m, that is, the overlap of the receivers state with the state unaffected by noise. One finds that

$$F_m=\mathrm{Tr}(\rho_m P_{\psi_m})=p+\frac{1-p}{m}. \tag{8}$$

Consider now that the source is multiplexed, i.e., $n>m$. The two receivers first carry out a unitary decoding operation,

$$U_d^A|j\rangle_A=\frac{1}{\sqrt{n}}\sum_k e^{-i2\pi jk/n}|k\rangle_A,$$

$$U_d^B|j\rangle_B=\frac{1}{\sqrt{n}}\sum_k e^{+i2\pi jk/n}|k\rangle_A,$$

where we have taken U_d^A, U_d^B to be the Fourier transform and the inverse Fourier transform, respectively. Party A then measures the operators $Q^A=\Sigma_{k=1}^m|k\rangle_{AA}\langle k|$, and party B measures the operator $Q^B=\Sigma_{k=1}^m|k\rangle_{BB}\langle k|$. By measuring $Q^{A(B)}$, we mean that the party keeps the particle if it is in channels 1 to m, and discards the particle otherwise. It is important that this is a measurement that does not affect the state of the particle if the measurement succeeds. Generalizations of this scheme to other decoding operations and measurements are described in Appendix G.

If both measurements succeed, the state becomes

$$\begin{aligned}\rho_f&=\frac{Q^AQ^BU_d^AU_d^B\rho_nU_d^{A\dagger}U_d^{B\dagger}Q^AQ^B}{\mathrm{Tr}(Q^AQ^BU_d^AU_d^B\rho_nU_d^{A\dagger}U_d^{B\dagger})}\\&=\frac{1}{P_{\text{success}}}\Bigg[\frac{mp}{n}P_{\psi_m'}+\frac{1-p}{n^2}\\&\quad\times\sum_{k,k',l,l'=1}^{m}\delta(k-k'-l+l')|k\rangle_A|l\rangle_B\,{}_A\langle k'|_B\langle l'|\Bigg],\end{aligned}$$

where $|\psi_m'\rangle=(1/\sqrt{m})\Sigma_{k=1}^m|k\rangle_A|k\rangle_B$ and $\delta(k-k'-l+l')$ equals 0 except if $k-k'-l+l'=0$ mod n when it is equal to 1; P_{success} is the probability that both measurements succeed and equals

$$P_{\text{success}}=\mathrm{Tr}(Q^AQ^BU_d^AU_d^B\rho_nU_d^{A\dagger}U_d^{B\dagger})=p\frac{m}{n}+(1-p)\frac{m^2}{n^2}. \tag{9}$$

The fidelity of the state ρ_f is

$$F_m'=\mathrm{Tr}(\rho_fP_{\psi_m'})=\frac{(n-1)p+1}{(n-m)p+m}. \tag{10}$$

This fidelity is greater than F_m, the fidelity if an entangled state of dimension m had been transmitted without error filtration, and tends to 1 for large n.

It is important to note that the state will also be purified if party A projects onto the subspace $Q_c^A=\Sigma_{k=c}^{c+m}P_k^A$ and party B simultaneously projects onto the subspace $Q_c^B=\Sigma_{k=c}^{c+m}P_k^B$ for arbitrary c. There are $[n/m]$ such orthogonal projectors (where $[x]$ denotes the largest integer smaller than or equal to x). Thus the total probability that the purification succeeds is $P_{\text{success-total}}=[n/m]P_{\text{success}}$. We note that, as n becomes large, with m fixed, this total success probability tends to p, the probability that no error occurred; in other words, we succeed in filtering out all errors.

In the above procedure, the biggest experimental difficulty is apparently the measurement of Q^A and Q^B. However,

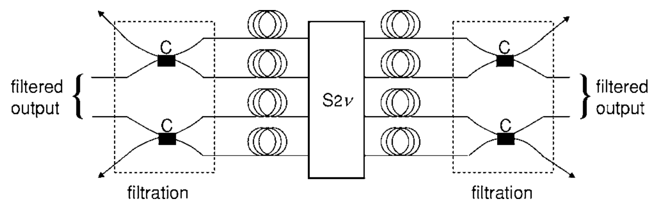

FIG. 3. Implementation of error filtration for entangled particles using multiple optical fibers. A source ($S_{2\nu}$) produces two photons in the entangled state $(|1\rangle_A|1\rangle_B+|2\rangle_A|2\rangle_B+|3\rangle_A|3\rangle_B+|4\rangle_A|4\rangle_B)/2$. States $|i\rangle_{A\ (B)}$ travel to receivers A (B) through different optical fibers. As in Fig. 1, the transmission states will be affected by independent phase noise. The receivers filter out the noise by projecting the state onto the subspaces spanned by $|1\rangle_{A\ (B)}+|2\rangle_{A\ (B)}$, $|3\rangle_{A\ (B)}+|4\rangle_{A\ (B)}$ using 50/50 couplers. When the projections of both parties succeed, the noise has been filtered out. When the projection of either of the parties fails, the state is rejected. Note that this decoding operation is based on the Hadamard transform and is distinct from the decoding operation based on the Fourier transform considered in the main text. However, one can easily show (see Appendix G) that both methods filter out the same amount of noise.

in many applications (for instance quantum cryptography) this measurement is not necessary. Indeed, the parties may proceed as follows: they assume that the filtration has succeeded and carry out the operations they desire as if the particle is present. After these operations, the parties carry out a destructive measurement to check whether the particle is indeed present. It is then that they know whether the filtration succeeded.

The above ideas are illustrated in Figs. 3 and 4 for photons traveling in multiple fibers and photons traveling in multiple time bins in the same fiber. Note that the protocols in Figs. 3 and 4 do not use the Fourier transform as decoding operation, but rather they use a Hadamard transformation. However, it is easy to show that these methods are both effective, see Appendix G.

The above paragraphs illustrate in a simple example how error filtration could be used to filter out noise during transmission of entangled states. In this example we consider a source that produces entangled states of dimension S. The entangled particles are then sent to the two receivers, using $T=S$ transmission channels, who project their state onto a smaller Hilbert space of dimension R. In the first part of Appendix G we generalize this protocol and give conditions on the decoding measurements for the protocol to filter out as much noise as possible. In the second part of Appendix G we consider a source which produces entangled states with the same dimension S as the receiver Hilbert space $R=S$. The number of transmission channels T is taken to be larger than S. Thus one is essentially using the protocols developed for error filtration in the case of a single particle twice, once for each particle. We briefly describe in the second part of Appendix G how such a protocol filters out errors in the case of entangled particles.

V. CONCLUSION

In summary, we have presented a conceptually new way of dealing with errors in quantum communication. This idea,

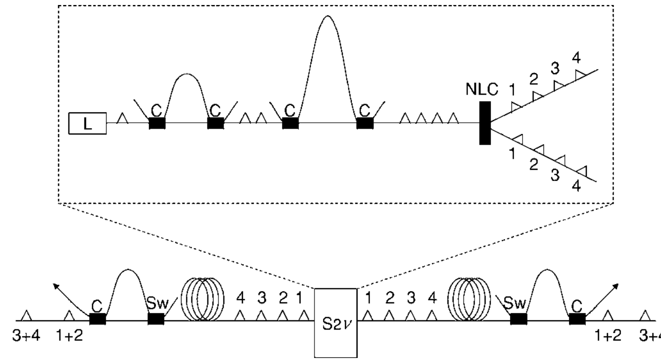

FIG. 4. Implementation of error filtration for entangled particles using time bins. A source ($S_{2\nu}$) produces two photons in the entangled state $(|1\rangle_A|1\rangle_B+|2\rangle_A|2\rangle_B+|3\rangle_A|3\rangle_B+|4\rangle_A|4\rangle_B)/2$ where states $|i\rangle_{A\ (B)}$ correspond to different time bins traveling through the same optical fiber. A possible such source, adapted from [20], is described in the inset: a laser (L) produces intense pulses of light which pass through two unbalanced Mach-Zender (MZ) interferometers so as to produce four coherent equally spaced pulses. The pulses impinge on a nonlinear crystal (NLC) thereby producing the entangled state by parametric down conversion. As in Fig. 2, the transmission states will be affected by independent phase noise. To filter out the noise, the photons are sent through MZ interferometers. A switch (Sw), synchronized with the source, sends time bins 2 and 4 through the long arm and time bins 1 and 3 through the short arm. Time bins 1 and 2 and time bins 3 and 4 then interfere. The state is kept if both photons exit through the lower branch, in which case it has been projected onto the subspace spanned by $|1\rangle_{A\ (B)}+|2\rangle_{A\ (B)}$, $|3\rangle_{A\ (B)}+|4\rangle_{A\ (B)}$. If either of the photons exit through the upper branch, the filtration has failed. As in Fig. 3, this example is based on the Hadamard transform.

error filtration, is a method for reducing errors in quantum communication which can be easily implemented using present-day technology. Indeed, to our knowledge it is the first method that can easily be implemented in practice today. For this reason, we believe it will find a wide range of applications in quantum information processing and communication.

ACKNOWLEDGMENTS

We gratefully acknowledge partial financial support by the European IST project RESQ, by the Swiss NCCR "Quantum Photonics," by the Communauté Française de Belgique under Grant No. ARC 00/05-251, and the IUAP program of the Belgian government under Grant No. V-18. S.M. is supported by the Belgian National Research Foundation (FNRS).

APPENDIX A: GENERAL PROTOCOL FOR ERROR FILTRATION

In Sec. III a specific protocol for error filtration was presented. Here we show how it can be generalized. The generalization consists in allowing more general encoding and decoding operations than the Fourier transform. Nevertheless, here we keep the restriction that each source channel is encoded and decoded separately. Therefore, we can focus on a particular one, $|1\rangle_S$, say.

We start with the state of the source $|1\rangle_S$ but now encode the source state into T transmission channels,

$$|1\rangle_S \mapsto U_e|1\rangle_S, \tag{A1}$$

where the encoding transformation U_e is a unitary map from the source Hilbert space to the transmission Hilbert space. Note that since the transmission Hilbert space is of dimension T, whereas there is a single source state, we have to suppose that there are other input channels on which U_e can act. These additional input channels contain no excitations. Thus $U_e^\dagger U_e=(id)_S$ and $U_e U_e^\dagger=(id)_T$, where $(id)_S$ is the identity operator in the source Hilbert space.

Following encoding, the state of the system plus environment can be written as

$$\sum_{j=1}^{T} {}_T\langle j|U_e|1\rangle_S \quad |j\rangle_T|0\rangle_E, \tag{A2}$$

where we have introduced an orthonormal basis of transmission states $|j\rangle_T$.

The noise now occurs, causing the state to change to

$$\sum_{j=1}^{T} {}_T\langle j|U_e|1\rangle_S\, |j\rangle_T(\alpha|0\rangle_E+\beta|j\rangle_E), \tag{A3}$$

where we have supposed that the noise is phase noise which acts independently on each transmission state.

We now decode by performing a second unitary transformation (the generalization of the inverse Fourier transform),

$$|j\rangle_T \mapsto U_d|j\rangle_T. \tag{A4}$$

Unitarity means that $U_d^\dagger U_d=(id)_T$ and $U_d U_d^\dagger=(id)_R$. Thus the state becomes

$$\sum_{j=1}^{T} {}_T\langle j|U_e|1\rangle_S\; U_d|j\rangle_T(\alpha|0\rangle_E+\beta|j\rangle_E). \tag{A5}$$

Up to this point, the encoding and decoding procedure is very general. We will now restrict ourselves by demanding that if there is no noise, the state should be transmitted exactly into the receiver channel $|1\rangle_R$. Thus we specify

$$U_dU_e|1\rangle_S=|1\rangle_R. \tag{A6}$$

Thus the state of the system and environment becomes

$$|1\rangle_R\left(\alpha|0\rangle_E+\beta\sum_{j=1}^{T} {}_T\langle j|U_e|1\rangle_{SR}\langle 1|U_d|j\rangle_T|j\rangle_E\right) + \beta\sum_{j=1}^{T}\sum_{k=2}^{R} c_{kj}|k\rangle_R|j\rangle_E, \tag{A7}$$

where

$$c_{kj}={}_T\langle j|U_e|1\rangle_S\ {}_R\langle k|U_d|j\rangle_T. \tag{A8}$$

Again, we see that this state has a component in the receiver channel $|1\rangle_R$, which we regard as the "useful" signal and components in all the other receiver channels which we will discard.

The state in the useful receiver channel is

$$|1\rangle_R\left(\alpha|0\rangle_E+\beta\sum_{j=1}^{T} {}_T\langle j|U_e|1\rangle_S\ {}_R\langle 1|U_d|j\rangle_T|j\rangle_E\right). \tag{A9}$$

The probability that the particle ends in the useful channel is the magnitude squared of the state (A9), i.e.,

$$|\alpha|^2+|\beta|^2\sum_{j=1}^{T} |{}_T\langle j|U_e|1\rangle_S\ {}_R\langle 1|U_d|j\rangle_T|^2. \tag{A10}$$

On the other hand, there is a probability of

$$1-\left(|\alpha|^2+|\beta|^2\sum_{j=1}^{T} |{}_T\langle j|U_e|1\rangle_S\ {}_R\langle 1|U_d|j\rangle_T|^2\right) \tag{A11}$$

that the particle appears at one of the "nonuseful" receivers channels, $|k\rangle_R$, $k=2,\ldots,T$. These channels are nonuseful because the particle ending here is always correlated with noise in the environment. Indeed, we see in (A7) that the component containing the receiver channels $|k\rangle_R$, $k=2,\ldots,T$ is

$$\beta\sum_{j=1}^{T}\sum_{k=2}^{T} c_{kj}|k\rangle_R|j\rangle_E, \tag{A12}$$

which has no overlap with the unperturbed state of the environment $|0\rangle_E$.

The probability of noise in the useful receiver channel is

$$|\beta|^2\sum_{j=1}^{T} |{}_T\langle j|U_e|1\rangle_S\ {}_R\langle 1|U_d|j\rangle_T|^2. \tag{A13}$$

Now Schwartz's inequality shows that this magnitude is greater than

$$\frac{|\beta|^2}{T}\left|\sum_{j=1}^{T} {}_T\langle j|U_e|1\rangle_S\ {}_R\langle 1|U_d|j\rangle_T\right|^2=\frac{|\beta|^2}{T}, \tag{A14}$$

with equality when

$$|{}_T\langle j|U_e|1\rangle_S\ {}_R\langle 1|U_d|j\rangle_T|=\frac{1}{T},\quad \text{independent of } j. \tag{A15}$$

Thus for any encoding and decoding scheme satisfying the conditions (A6) and (A15), we find that the noise amplitude is reduced by a factor of $1/\sqrt{T}$.

One example of encoding/decoding schemes satisfying these conditions is the Fourier transform in the previous appendix. A second example is the Hadamard transform; in the case of four transmission channels encoding each source channel, the encoding is

$$|1\rangle_S \mapsto \frac{1}{2}(|1\rangle_T+|2\rangle_T+|3\rangle_T+|4\rangle_T), \tag{A16}$$

$$|2\rangle_S \mapsto \frac{1}{2}(|1\rangle_T + |2\rangle_T - |3\rangle_T - |4\rangle_T),$$

$$|3\rangle_S \mapsto \frac{1}{2}(|1\rangle_T - |2\rangle_T + |3\rangle_T - |4\rangle_T),$$

$$|4\rangle_S \mapsto \frac{1}{2}(|1\rangle_T - |2\rangle_T - |3\rangle_T + |4\rangle_T);$$

the decoding step is

$$|1\rangle_T \mapsto \frac{1}{2}(|1\rangle_R + |2\rangle_R + |3\rangle_R + |4\rangle_R),$$

$$|2\rangle_T \mapsto \frac{1}{2}(|1\rangle_R + |2\rangle_R - |3\rangle_R - |4\rangle_R),$$

$$|3\rangle_T \mapsto \frac{1}{2}(|1\rangle_R - |2\rangle_R + |3\rangle_R - |4\rangle_R),$$

$$|4\rangle_T \mapsto \frac{1}{2}(|1\rangle_R - |2\rangle_R - |3\rangle_R + |4\rangle_R). \quad \text{(A17)}$$

APPENDIX B: PROTOCOL FOR ERROR FILTRATION WITH COLLECTIVE ENCODING

In the previous appendix, a given source channel is encoded into a subset of the transmission channels, but the signal carried in a given transmission channel came from a single source channel. It is possible to generalize these ideas by allowing each transmission channel to carry signals from more than one source channel. This will be of use, for example, when the number of source channels does not divide the number of transmission channels (e.g., two source channels and three transmission channels).

Thus the general situation is that we have $S_{\rm tot}$ source channels which we encode collectively into $T_{\rm tot}$ transmission channels and then decode the transmission channels collectively into $R_{\rm tot}$ receiver channels with $S_{\rm tot}=R_{\rm tot}$. It is clear that the previous protocols, in which a given source channel is encoded into T transmission channels and each source channel is encoded into a different set of transmission channels, are included in this general framework. But other possibilities exist.

Here we will simply illustrate the idea with an example. We consider again the case of phase noise: errors in the transmission channels of the form

$$|k\rangle_T|0\rangle_E \mapsto |k\rangle_T(\alpha|0\rangle_E + \beta|k\rangle_E), \quad k = 1\ldots T_{\rm tot}. \quad \text{(B1)}$$

If each source channel $|j\rangle_S$, $j=1\ldots S_{\rm tot}$ is simply sent along a single transmission channel (trivial encoding), the error amplitude is β. However, an example of the general framework in the previous paragraph is the following encoding/decoding scheme, based on the Fourier transform. The encoding step is

$$|j\rangle_S \mapsto \frac{1}{\sqrt{T}}\sum_{k=1}^{T_{\rm tot}} e^{-2\pi i k(j-1)/T}|k\rangle_T. \quad \text{(B2)}$$

The decoding step is

$$|k\rangle_T \mapsto \frac{1}{\sqrt{T}}\sum_{m=1}^{S_{\rm tot}} e^{2\pi i k(m-1)/T}|m\rangle_R. \quad \text{(B3)}$$

For example, this protocol may be used in the case that $S_{\rm tot}=2=R_{\rm tot}$ and $T_{\rm tot}=3$. Let us write the state emitted by the source as

$$a_1|1\rangle_S + a_2|2\rangle_S, \quad \text{(B4)}$$

where a_1 and a_2 are complex amplitudes obeying $|a_1|^2+|a_2|^2=1$. It may be calculated that the fidelity of the state arriving at the receiver in the channels $|1\rangle_R$ and $|2\rangle_R$ to the incoming state is

$$\frac{|\alpha|^2 + \frac{|\beta|^2}{3}(1 + 2|a_1|^2|a_2|^2)}{|\alpha|^2 + \frac{2|\beta|^2}{3}}. \quad \text{(B5)}$$

The average value of this fidelity over the Bloch sphere of incoming states is

$$\frac{|\alpha|^2 + \frac{4|\beta|^2}{9}}{|\alpha|^2 + \frac{2|\beta|^2}{3}}. \quad \text{(B6)}$$

For any α and β, this is greater than the average fidelity achieved by simply sending each source state through one transmission channel.

APPENDIX C: ERROR FILTRATION FOR PARTICLES WITH INTERNAL DEGREES OF FREEDOM

The fact that the noise corrected by the previous protocols was phase noise, and hence corresponds to random elements of the Abelian group $U(1)$, was not critical. The protocols can be simply extended to the case where each channel can carry a system which has "internal" degrees of freedom (i.e., each channel has a state space which is a Hilbert space of arbitrary dimension I; we shall consider this dimension to be finite here, but this is not essential to the success of the protocol).

Thus we consider an orthonormal set of source states,

$$|i\mu\rangle_S; \quad i = 1\ldots S_{\rm tot}, \quad \mu = 1\ldots I. \quad \text{(C1)}$$

An example is the case where each channel can carry a spin degree of freedom, so that $I=2$. We consider a set of transmission states

$$|j\mu\rangle_T; \quad j = 1\ldots T_{\rm tot}, \quad \mu = 1\ldots I, \quad \text{(C2)}$$

i.e., there are $T_{\rm tot}$ transmission channels each of which carries a state space of dimension I. The transmission states are affected by the following noise:

$$|j\mu\rangle_T|0\rangle_E \mapsto \alpha|j\mu\rangle_T|0\rangle_E + \sum_{\nu=1}^{I}\sum_{\lambda=1}^{L}\beta_\lambda(E_\lambda)_{\mu\nu}|j\nu\rangle_T|j\lambda\rangle_E. \tag{C3}$$

This describes L types of error; each error corresponds to a rotation of the system state. In the case of internal spin degrees of freedom, $I=2$, an example of a set of possible errors is the set of three Pauli matrices,

$$(E_1)_{\mu\nu} = (\sigma_x)_{\mu\nu}, \quad (E_2)_{\mu\nu} = (\sigma_y)_{\mu\nu}, \quad (E_3)_{\mu\nu} = (\sigma_z)_{\mu\nu}. \tag{C4}$$

Thus $|j\lambda\rangle_E$ is the state of the environment if error λ occurred on channel j.

The total probability of error in the channel (C3) is

$$\left|\sum_{\nu=1}^{I}\sum_{\lambda=1}^{L}\beta_\lambda(E_\lambda)_{\mu\nu}|j\nu\rangle_T|j\lambda\rangle_E\right|^2 = \sum_{\lambda=1}^{L}|\beta_\lambda|^2\left|\sum_{\nu=1}^{I}(E_\lambda)_{\mu\nu}\right|^2 = 1-|\alpha|^2. \tag{C5}$$

Thus if the source state $|1\mu\rangle_S$ is simply sent through a single transmission channel with trivial encoding and decoding,

$$|1\mu\rangle_S \mapsto |1\mu\rangle_T \mapsto |1\mu\rangle_R, \tag{C6}$$

the probability of error is

$$1-|\alpha|^2. \tag{C7}$$

Now consider a simple error-filtration protocol in which the source channel is multiplexed to T transmission channels. The encoding step is

$$|1\mu\rangle_S|0\rangle_E \mapsto \frac{1}{\sqrt{T}}\sum_{j=1}^{T}|j\mu\rangle_T|0\rangle_E. \tag{C8}$$

Note that the encoding is independent of μ.

Now the noise occurs during the transmission, and the state becomes

$$\frac{1}{\sqrt{T}}\sum_{j=1}^{T}|j\mu\rangle_T|0\rangle_E \mapsto \frac{1}{\sqrt{T}}\sum_{j=1}^{T}\alpha|j\mu\rangle_T|0\rangle_E + \frac{1}{\sqrt{T}}\sum_{j=1}^{T}\sum_{\nu=1}^{I}\sum_{\lambda=1}^{L}\beta_\lambda(E_\lambda)_{\mu\nu}|j\nu\rangle_T|j\lambda\rangle_E. \tag{C9}$$

The decoding step for this protocol is also the Fourier transform on the j indices,

$$|j\nu\rangle_T \mapsto \frac{1}{\sqrt{T}}\sum_{k=1}^{T}e^{2i\pi j(k-1)/T}|k\nu\rangle_R. \tag{C10}$$

As before we select the term $|1\mu\rangle_R$ at the receiver; thus the state in this receiver channel is

$$\alpha|1\mu\rangle_R|0\rangle_E + \frac{1}{T}\sum_{j=1}^{T}\sum_{\nu=1}^{I}\sum_{\lambda=1}^{L}\beta_\lambda(E_\lambda)_{\mu\nu}|1\nu\rangle_T|j\lambda\rangle_E. \tag{C11}$$

Now the probability that the state was affected by noise is the square of the magnitude of the second term,

$$\frac{1}{T^2}\left|\sum_{j=1}^{T}\sum_{\nu=1}^{I}\sum_{\lambda=1}^{L}\beta_\lambda(E_\lambda)_{\mu\nu}|1\nu\rangle_R|j\lambda\rangle_E\right|^2 = \frac{1}{T}\sum_{\lambda=1}^{L}|\beta_\lambda|^2\left|\sum_{\nu=1}^{I}(E_\lambda)_{\mu\nu}\right|^2, \tag{C12}$$

where we have used the fact that

$${}_E\langle j_1\lambda_1|j_2\lambda_2\rangle_E = \delta_{j_1 j_2}\delta_{\lambda_1\lambda_2} \quad \text{and} \quad {}_R\langle 1\nu_1|1\nu_2\rangle_R = \delta_{\nu_1\nu_2}. \tag{C13}$$

Thus comparing (C12) with (C5), we see that the error filtration protocol has reduced the probability of error by a factor of $1/T$, i.e., each error amplitude has been reduced by a factor of $1/\sqrt{T}$.

This protocol has essentially used the Fourier transform to encode and decode. It is not difficult to extend the protocol to more general encoding/decodings as was done for phase noise in Appendices A and B. Furthermore, the encoding need not be independent of the internal degrees of freedom μ [as was the case in Eq. (C8)].

APPENDIX D: USING ERROR FILTRATION IN SERIES

The previous protocols can loosely be described as using transmission channels in parallel to achieve noise filtration. We may also use the idea of multiple channels in series to filter noise. We illustrate this idea in the case of phase noise.

We compare two situations. Given a source channel we wish to improve, we can use the encoding described in the main text where we multiplex a single source channel into T transmission channels. This has the effect of causing the noise amplitude to be reduced from β to $\beta/\sqrt{T}$, as we showed earlier. If we imagine that the transmission channels have a certain length, l, we could perform the same encoding as in the above protocol, but then use the original decoding procedure at the half-way point (or any other point along the transmission channels), then reperform the encoding, allow the signal to travel for the remaining part of the transmission channel, and finally decode again. As we now show, this protocol gives better error filtration than the protocol without the interior decoding/encoding (assuming that the decoding/encoding module itself does not introduce significant errors). Clearly one could perform the decoding/encoding module at as many interior points as one wishes; we calculate the effect of this below. Thus using error filtration in series is somewhat analogous to the quantum Zeno effect by which evolution is frozen by repeated measurements.

Recall first that if we do not carry out multiplexing, the state of a particle passing through channel 1 is

$$|1\rangle_R(\alpha|0\rangle_E + \beta|1\rangle_E), \tag{D1}$$

whereas if we multiplex into T transmission channels the state of the system plus environment after transmission is

$$|1\rangle_R\left(\alpha|0\rangle_E+\frac{\beta}{\sqrt{T}}|1\rangle_E\right). \tag{D2}$$

Now we imagine decomposing the transmission channel into two halves. We describe the environment Hilbert space as being the tensor product of two Hilbert spaces, one for the first half (E_1) and one for the second half (E_2) of the transmission. After the first half (in the absence of multiplexing), the state is

$$|1\rangle_R(\alpha'|0\rangle_{E1}+\beta'|1\rangle_{E1}), \tag{D3}$$

and after the second half it is

$$|1\rangle_R(\alpha'|0\rangle_{E_1}+\beta'|1\rangle_{E_1})(\alpha'|0\rangle_{E_2}+\beta'|1\rangle_{E_2}). \tag{D4}$$

Comparing with Eq. (D1), we see that $\alpha'^2=\alpha$.

If we carry out multiplexing in series on the two halves, we find that the state after transmission is

$$|1\rangle_R\left(\alpha'|0\rangle_{E_1}+\frac{\beta'}{\sqrt{T}}|1\rangle_{E_1}\right)\left(\alpha'|0\rangle_{E_2}+\frac{\beta'}{\sqrt{T}}|1\rangle_{E_2}\right). \tag{D5}$$

In order to find out the overall probability for an error to have occurred, we write (D5) as

$$|0\rangle_R(\alpha''|0\rangle_E+\beta''|1''\rangle_E), \tag{D6}$$

where $|1''\rangle_E$ is a normalized vector. The probability that the useful receiver state is affected by noise is thus

$$|\beta''|^2=\frac{(1-|\alpha|)(1+2T|\alpha|-|\alpha|)}{T^2}. \tag{D7}$$

It is not difficult to check that this probability is less than the probability of error without the insertion of the decoding/encoding module (this is equal to $|\beta|^2/T$) for any α.

More generally, one can consider what happens if one has a total of Q internal decoding/encoding modules. One gets maximal reduction of error probability when these modules are equally spaced along the transmission channel. In this case the total error probability is found to be

$$\left[|\alpha|^{2/(Q+1)}+\frac{1-|\alpha|^{2/(Q+1)}}{T}\right]^{(Q+1)}-|\alpha|^2. \tag{D8}$$

We note that this probability tends to

$$|\alpha|^{[2(T-1)]/T}-|\alpha|^2 \tag{D9}$$

as the number, Q, of internal decoding/encoding modules tends to infinity.

APPENDIX E: QUANTUM MULTIEXCITATION PROTOCOL

In the previous protocols, the channels contained only a single excitation each. In the case of bosons, we can also consider the situation where each channel contains many quanta. We will illustrate this first by considering the case in which the channel states are coherent states; at the end of this appendix, we show how the protocol may be used for general multiexcitation states.

Let us consider as before two input channels. Each channel is now described by an infinite-dimensional Hilbert space and we may describe the states in terms of the creation and annihilation operators,

$$[a_S^1,(a_S^1)^\dagger]=1 \quad \text{and } [\tilde{a}_S^1,(\tilde{a}_S^1)^\dagger]=1, \tag{E1}$$

where a_S^1 refers to the first channel and $\tilde{a}_S^1$ to the second. We will work in the Schrödinger picture of dynamics. Let the initial state of the system be the following coherent state:

$$N(\lambda)\exp\left(\frac{\lambda}{\sqrt{2}}[(a_S^1)^\dagger+e^{i\Phi}(\tilde{a}_S^1)^\dagger]\right)|0\rangle_{\rm sys}, \tag{E2}$$

where $|0\rangle_{\rm sys}$ is the vacuum state for the system and $N(\lambda)$ is a normalization factor. The phase Φ allows us to transmit a signal; it will also be used later to allow us to measure the effect of the noise.

Let us first consider what happens in the absence of filtration, that is, when there is trivial encoding, namely when each source channel evolves into a single transmission channel. The state of the system evolves to

$$N(\lambda)\exp\left(\frac{\lambda}{\sqrt{2}}[(a_T^1)^\dagger+e^{i\Phi}(\tilde{a}_T^1)^\dagger]\right)|0\rangle_{\rm sys}. \tag{E3}$$

The initial state of the environment is a product of states, one for each channel. We denote it $|\xi\rangle_E$. Thus the state of the system plus environment is

$$N(\lambda)\exp\left(\frac{\lambda}{\sqrt{2}}[(a_T^1)^\dagger+e^{i\Phi}(\tilde{a}_T^1)^\dagger]\right)|0\rangle_{\rm sys}|\xi\rangle_E. \tag{E4}$$

The effect of the noise is that there is an interaction between the system and environment. This may be modeled by a unitary transformation of the form

$$U=\exp i[(a_T^1)^\dagger a_T^1 B^1+(\tilde{a}_T^1)^\dagger\tilde{a}_T^1\tilde{B}^1], \tag{E5}$$

where B^1 and $\tilde{B}^1$ are Hermitian operators acting on the environment Hilbert spaces which we do not need to specify further.

Thus after transmission through the noisy channels, the state becomes

$$N(\lambda)\exp\left(\frac{\lambda}{\sqrt{2}}[(a_T^1)^\dagger e^{iB^1}+e^{i\Phi}(\tilde{a}_T^1)^\dagger e^{i\tilde{B}^1}]\right)|0\rangle_{\rm sys}|\xi\rangle_E. \tag{E6}$$

We now decode the signal trivially so that the state at the receiver is given by Eq. (E6). Let us now allow these two receiver channels to interfere. This has the effect of transforming the operators a_R^1 and $\tilde{a}_R^1$ into

$$(a_R^1)\mapsto\frac{1}{\sqrt{2}}(c_R^1+d_R^1) \quad \text{and } (\tilde{a}_R^1)\mapsto\frac{1}{\sqrt{2}}(c_R^1-d_R^1). \tag{E7}$$

We now calculate the current in the channel c_R^1. This is the expected value of the operator

$$(c_R^1)^\dagger c_R^1 \tag{E8}$$

in the final state

$$N(\lambda)\exp\left(\frac{\lambda}{2}[(c_R^1)^\dagger(e^{iB^1}+e^{i\Phi}e^{i\tilde{B}^1})+(d_R^1)^\dagger(e^{iB^1}-e^{i\Phi}e^{i\tilde{B}^1})]\right)$$
$$\times|0\rangle_{\rm sys}|\xi\rangle_E. \tag{E9}$$

This expectation value is

$$\frac{|\lambda|^2}{4}{}_E\langle\xi|(e^{-iB^1}+e^{-i\Phi}e^{-i\tilde{B}^1})(e^{iB^1}+e^{i\Phi}e^{i\tilde{B}^1})|\xi\rangle_E. \tag{E10}$$

Recall that the state of the environment $|\xi\rangle_E$ is a product of states for the individual channels, thus we may write it as

$$|\xi\rangle_E=|\xi^1\rangle_E|\tilde{\xi}^1\rangle_E. \tag{E11}$$

Thus, for example,

$${}_E\langle\xi|e^{-iB^1}e^{i\tilde{B}^1}|\xi\rangle_E={}_E\langle\xi^1|e^{-iB^1}|\xi^1\rangle_{EE}\langle\tilde{\xi}^1|e^{i\tilde{B}^1}|\tilde{\xi}^1\rangle_E. \tag{E12}$$

We assume, as in our discussions of the previous protocols, that the noise on different channels is independent, thus we write

$${}_E\langle\xi^1|e^{-iB^1}|\xi^1\rangle_E=\alpha^*,\quad {}_E\langle\tilde{\xi}^1|e^{i\tilde{B}^1}|\tilde{\xi}^1\rangle_E=\alpha. \tag{E13}$$

Therefore, the expected value of the current in the channel c_R^1 is

$$\frac{|\lambda|^2}{2}(1+|\alpha|^2\cos\Phi). \tag{E14}$$

We now consider what happens when we encode each of the source channels by multiplexing to T transmission channels. We again start with the coherent state (E2). We illustrate the noise filtration in the case when the encoding is the Fourier transform. This encoding has the effect of transforming the creation operators in the coherent state into

$$(a_S^1)^\dagger\mapsto\frac{1}{\sqrt{T}}\sum_{i=1}^{T}(a_T^i)^\dagger,\quad(\tilde{a}_S^1)^\dagger\mapsto\frac{1}{\sqrt{T}}\sum_{i=1}^{T}(\tilde{a}_T^i)^\dagger. \tag{E15}$$

The noise now occurs, causing each creation operator in the coherent state to transform into

$$(a_T^i)^\dagger\mapsto e^{iB^i}(a_T^i)^\dagger,\quad(\tilde{a}_T^i)^\dagger\mapsto e^{i\tilde{B}^i}(\tilde{a}_T^i)^\dagger. \tag{E16}$$

We now decode with the inverse Fourier transform, and consider the signal in the two receiver channels defined by the creation operators $(a_R^1)^\dagger$ and $(\tilde{a}_R^1)^\dagger$. We again allow these to interfere and finally calculate the expected value of the current

$$(c_R^1)^\dagger c_R^1 \tag{E17}$$

in the final state. This is

$$\frac{|\lambda|^2}{4T^2}{}_E\langle\xi|\left(\sum_{i=1}^{T}(e^{-iB^i}+e^{-i\Phi}e^{-i\tilde{B}^i})\right)\left(\sum_{j=1}^{T}(e^{iB^j}+e^{i\Phi}e^{i\tilde{B}^j})\right)|\xi\rangle_E$$
$$=\frac{|\lambda|^2}{2}\left(\frac{1+(T-1)|\alpha|^2}{T}\right)\left(1+\frac{T|\alpha|^2}{1+(T-1)|\alpha|^2}\cos\Phi\right). \tag{E18}$$

Exactly as in the previous protocols, the multiplexing has the effect of reducing the noise.

So far in this appendix we have considered a particularly simple initial state, a coherent state. In this case it is rather straightforward to calculate the effect of our filtration protocol. However, the protocol may be used for any multiexcitation state.

Let us consider that the state of the source is defined by some function of creation operators $(a_S^1)^\dagger$ acting on the vacuum. The effect of the encoding and decoding that we have performed above is to change this state to one in which the operator $(a_S^1)^\dagger$ is transformed to an expression of the form

$$\frac{1}{T}\sum_{j=1}^{T}e^{iB^j}(a_R^1)^\dagger. \tag{E19}$$

Thus any power of the operator $[(a_S^1)^\dagger]^N$ becomes replaced by

$$\frac{1}{T^N}\left(\sum_{j=1}^{T}e^{iB^j}\right)^N[(a_R^1)^\dagger]^N. \tag{E20}$$

We now imagine computing the expectation value of some operator in the state. For T much larger than N we can neglect all terms in the expectation value in which any given operator e^{iB^k}, say, appears to any power greater than 1. Hence when we compute the expection value, we can perform the inner product with the state of the environment, and hence replace (E20) by

$$\frac{1}{T^N}(T\alpha)^N[(a_R^1)^\dagger]^N, \tag{E21}$$

where α is the expected value of e^{iB^k} for channel k. Thus in the limit of large T we see that the effect of the protocol is that the source operator $(a_S^1)^\dagger$ gets transformed to

$$\alpha(a_R^1)^\dagger. \tag{E22}$$

The key point that this protocol achieves (for large T) is that interference between operators is not affected, i.e.,

$$\frac{1}{\sqrt{2}}[(a_S^1)^\dagger+e^{i\Phi}(\tilde{a}_S^1)^\dagger]\mapsto\alpha\frac{1}{\sqrt{2}}[(a_R^1)^\dagger+e^{i\Phi}(\tilde{a}_R^1)^\dagger]. \tag{E23}$$

Destruction of interference is avoided and replaced by overall absorption of quanta. This is the exact analogue of what happens for the single-quanta protocols presented earlier where visibility is improved at the cost of overall reduction in intensity.

We note that while we have focused on the case of large T in the previous paragraph, similar analysis shows that, quite generally, even for finite T, multiplexing has the effect of reducing noise and replacing it by an overall reduction in intensity.

APPENDIX F: ERROR FILTRATION FOR CLASSICAL WAVE SIGNALS

In the previous appendix, we considered the case where the signal contained many excitations. A limiting case is the one where the number of excitations is macroscopic and one is dealing with classical wave signals. Classical waves (as well as the quantum multiexcitation states discussed in the previous appendix) have a number of properties that can all be affected by noise. Scalar classical waves (i.e., described by a single complex amplitude) can be affected by phase noise; waves having "internal degrees of freedom" (such as polarization) can be affected by noise acting on these degrees of freedom. But the wave can also be subjected to noise that affects its amplitude (amplitude noise). Furthermore, it is also possible that the amount of noise in all the above cases depends nonlinearly on the amplitude ("nonlinear noise"). The very same mulplexing scheme that we used for single particles will filter noise in all these situations.

We consider the simple encoding/decoding in which a single source channel is encoded in T transmission channels. Denote by ψ_S a normalized mode of the source channel. Consider a classical signal of amplitude A emitted by the source in mode ψ_S. Denote by $\psi_T^j (j=1,\ldots,T)$ normalized modes of the transmission channels. The encoding operation transforms the signal as

$$A\psi_S \rightarrow \sum_{j=1}^{T} \frac{A}{\sqrt{T}}\psi_T^j = \sum_{j=1}^{T} A_{T_{\text{in}}}^j \psi_T^j, \tag{F1}$$

where $A_{T_{\text{in}}}^j$ is the amplitude at the input of transmission channel j. Suppose that noise acts during the transmission. Then the amplitude in transmission channel j gets transformed as

$$A_{T_{\text{in}}}^j \rightarrow A_{T_{\text{out}}}^j(\xi_j, A_{T_{\text{in}}}^j), \tag{F2}$$

where ξ_j is a random variable (the noise) and $A_{T_{\text{out}}}^j(\xi_j, A_{T_{\text{in}}}^j)$ is the amplitude at the output of the transmission channel if the noise has value ξ_j and the signal at the input of the transmission channel has amplitude $A_{T_{\text{in}}}^j$. It is convenient to rewrite the output amplitude as

$$A_{T_{\text{out}}}^j(\xi_j, A_{T_{\text{in}}}^j) = A_{T_{\text{in}}}^j N^j(\xi_j, A_{T_{\text{in}}}^j), \tag{F3}$$

where N^j is the noise acting on channel j. The noise is linear if N^j depends only on ξ_j but not on $A_{T_{\text{in}}}^j$.

Again, by suitable engineering, we can arrange that the noise acts independently on the different channels. Mathematically, this is equivalent to the statement that the ξ_j are independent random variables. If in addition the channels are identical, then we have the further simplification that the functions $N^j(\xi_j, A_{T_{\text{in}}}^j) = N(\xi_j, A_{T_{\text{in}}}^j)$ are independent of j and that the ξ_j are independent identically distributed (i.i.d.) random variables. For simplicity, we assume from now on that the noise in all transmission channels is identical and independent, i.e.,

$$A_{T_{\text{out}}}^j = A_{T_{\text{in}}}^j N(\xi_j, A_{T_{\text{in}}}^j) \quad \text{with } \xi_j \text{ i.i.d. random variables.} \tag{F4}$$

As an illustration, linear phase noise is described by a noise function

$$N_{\text{linear phase}}(\xi, A) = e^{i\varphi(\xi)} \tag{F5}$$

and linear amplitude noise is described by a function

$$N_{\text{linear amplitude}}(\xi, A) = f(\xi) \quad \text{with } f \text{ real} \tag{F6}$$

whereas nonlinear phase noise could, for instance, be described by a noise function

$$N_{\text{nonlinear phase}}(\xi, A) = N_1 e^{i\varphi(\xi)|A|^2} = N_1 N_2, \tag{F7}$$

where N_1 represents some linear phase noise and $N_2 = e^{i\varphi(\xi)|A|^2}$ is independent of N_1 and describes the nonlinear part.

Having described the action of the noise, let us consider the form of the useful signal in the receiver channel. We suppose that the decoding is realized using the Fourier transform. Denote by ψ_R a normalized mode of the receiver channel. The amplitude in the receiver channel is

$$A_R\psi_R = \sum_j \frac{A_{T_{\text{out}}}^j}{\sqrt{T}}\psi_R = A\left(\frac{1}{T}\sum_{j=1}^{T} N(\xi_j, A/\sqrt{T})\right)\psi_R. \tag{F8}$$

In order to interpret the expression Eq. (F8), we first concentrate on the case of linear noise. We will then come back to the case of nonlinear noise. Upon averaging over the random noise variables ξ_j, one finds that the average amplitude in the receiver channel

$$\overline{A_R} = A\bar{N} \tag{F9}$$

is independent of the degree T of multiplexing. On the other hand, the average intensity in the reception channel depends on the degree of multiplexing,

$$\overline{I_R} = \overline{A_R^* A_R} = |\overline{A_R}|^2\left(1 + \frac{1}{T}\frac{\overline{N^* N} - |\bar{N}|^2}{|\bar{N}|^2}\right). \tag{F10}$$

The average intensity thus decreases with the degree of multiplexing. This is because the filtration removes more and more noise as T increases. In the limit of large T, the average intensity is equal to the norm squared of the average amplitude.

This result can be reexpressed in terms of the amplitude fluctuations in the reception channel. These fluctuations are given by

$$\Delta A_R^2 = \overline{A_R^* A_R} - |\overline{A_R}|^2 = |\overline{A_R}|^2 \frac{1}{T}\frac{\overline{N^* N} - |\bar{N}|^2}{|\bar{N}|^2}. \tag{F11}$$

They decrease with T. In the limit of large T, the amplitude in the useful receiver channel A_R no longer fluctuates, i.e., all the noise has been removed.

It is interesting to also look at the amplitudes in the nonuseful receiver channels. These are given by

$$A_R^k \psi_R^k = \sum_j \frac{A^j}{\sqrt{T}} e^{i2\pi jk/T}\psi_R^k, \quad k \neq 0. \tag{F12}$$

One finds that the average amplitude in the nonuseful receiver channels is zero,

$$\overline{A_R^k} = 0, \tag{F13}$$

and that these channels contain nonzero average intensity. This means that these channels contain only noise.

As an illustration of the effect of the noise reduction in the useful receiver channel, we consider the visibility of interference fringes. Suppose the sender prepares two signals,

$$A_{S1}\psi_{S1} + A_{S2}\psi_{S2}, \tag{F14}$$

where $A_{S1}=(1/\sqrt{2})\mathcal{A}$ and $A_{S2}=(1/\sqrt{2})e^{i\phi}\mathcal{A}$. Let us suppose that each signal is independently transmitted by multiplexing it into T channels. The useful receiver signals are

$$A_{R1}\psi_{R1} + A_{R2}\psi_{R2}, \tag{F15}$$

where

$$A_{R1} = \frac{\mathcal{A}}{\sqrt{2}T}\sum_{j=1}^{T} N(\xi_j),$$

$$A_{R2} = \frac{e^{i\phi}\mathcal{A}}{\sqrt{2}T}\sum_{j=T+1}^{2T} N(\xi_j),$$

$$\xi_j \text{ i.i.d. random variables, } j = 1,\ldots,2T. \tag{F16}$$

If the receiver lets the two signals interfere, he will find an intensity

$$\overline{\left|\frac{A_{R1}+A_{R2}}{\sqrt{2}}\right|^2} = \frac{1}{2}|\mathcal{A}|^2|\bar{N}|^2\left(1 + \frac{\overline{|N|^2} - |\bar{N}|^2}{T|\bar{N}|^2} + \cos\phi\right), \tag{F17}$$

hence the visibility of interference fringes he sees is given by

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = \frac{1}{1 + \frac{\overline{|N|^2} - |\bar{N}|^2}{T|\bar{N}|^2}}. \tag{F18}$$

The visibility thus tends to 1 as the degree of multiplexing T increases.

Let us now consider the case of nonlinear noise. The effects of the multiplexing are more complex. Multiplexing reduces noise via two independent processes. First, there is the filtration itself. Second, multiplexing also reduces the intensity in each transmission channel, and so it reduces nonlinearity. Both effects are beneficial, and the exact result depends on the interplay between them, and on the specific form of the nonlinearity.

All the above formulas Eqs. (F9)–(F18) stay valid, but one must replace $N(\xi)$ by $N(\xi, A/\sqrt{T})$, which introduces an additional dependence on T. A first consequence is that the average amplitude $\overline{A_R}$ depends now on the degree of multiplexing. For the noise considered in (F7), $\overline{A_R}=\mathcal{A}\overline{N_1}\,\overline{N_2}$ depends on the degree of multiplexing through $\overline{N_2}$. Let us now compute the noise fluctuations [which appear in Eqs. (F10), (F11), (F17), and (F18)],

$$\frac{1}{T}\frac{\overline{|N|^2} - |\bar{N}|^2}{|\bar{N}|^2}. \tag{F19}$$

As an illustration, consider nonlinear phase noise described in (F7). Let us suppose that the noise is small so that we can expand $N_2(\xi, A) = 1 + i\varphi|A|^2 - \varphi^2|A|^4/2$. Then one easily obtains that

$$\frac{1}{T}\frac{\overline{|N|^2} - |\bar{N}|^2}{|\bar{N}|^2} = \frac{1}{T}\left[\frac{\overline{|N_1|^2} - |\overline{N_1}|^2}{|\overline{N_1}|^2} + \frac{1}{T^2}\frac{|\mathcal{A}|^4}{|\overline{N_1}|^2}(\overline{\varphi^2} - \bar{\varphi}^2) + O\left(\frac{1}{T^4}\right)\right]. \tag{F20}$$

Here we can see the interplay of the two ways in which noise is reduced by multiplexing. The factor $1/T$ that multiplies the whole expression comes directly from filtering, while the factor $1/T^2$ in the square bracket that multiplies the nonlinear noise appears because the reduction in intensity in each transmission channel has reduced the nonlinearity.

Above we considered the situation in which the classical wave was described by a single complex amplitude, i.e., there were no internal degrees of freedom. In the case of quantum systems, we showed in Appendix B that multiplexing works equally well if the particle has internal degrees of freedom. It is not difficult to check that exactly the same is true for classical waves.

All the above remarks which have been made in the case of classical signals are important because they also apply in the case of signals with one or many excitations: error filtration will remove amplitude noise and nonlinear noise in the quantum case as well.

APPENDIX G: PROTOCOLS FOR COMMUNICATION OF ENTANGLED STATES

Here we generalize the error filtration protocol for communication of entangled states given in the main part of the paper.

The protocols we consider in Appendix G have the following general structure. A source emits two S-level systems in an entangled state. System A will be sent to party A whereas system B will be sent to party B. These systems are first encoded as T-level systems. The signals are then transmitted through the noisy channels. Finally they are decoded by parties A and B to R-level systems. The key to the performance of the protocols is to allow the encoding and decoding operations to change the dimension of the space of states.

Denote the states of system A when emitted by the source as

$$|i\rangle_S^A \ (i = 1\ldots S). \tag{G1}$$

These states are encoded by the encoder by a transformation U_e^A; thus the state of the system which emerges is

$$U_e^A|i\rangle_S^A = \sum_{j=1}^{T} {}_T^A\langle j|U_e^A|i\rangle_S^A|j\rangle_T^A. \tag{G2}$$

The initial state of the system plus environment of the channel is thus

$$\sum_{j=1}^{T} {}_T^A\langle j|U_e^A|i\rangle_S^A|j\rangle_T^A|0\rangle_E^A. \tag{G3}$$

After transmission through the channel, the state becomes

$$\sum_{j=1}^{T} {}_T^A\langle j|U_e^A|i\rangle_S^A|j\rangle_T^A(\alpha_{jA}|0\rangle_E^A+\beta_{jA}|j\rangle_E^A). \tag{G4}$$

The decoding acts by a further transformation U_d^A. Thus the final state after encoding and decoding is

$$\sum_{j=1}^{T} {}_T^A\langle j|U_e^A|i\rangle_S^A U_d^A|j\rangle_T^A(\alpha_{jA}|0\rangle_E^A+\beta_{jA}|j\rangle_E^A). \tag{G5}$$

Finally, only some of the receiver states are used. This is accomplished by projecting the final state on a Hilbert subspace of receiver states of dimension R. We will denote this projector by Π^A. There is a certain liberty in choosing this projector as one can modify the projector by a unitary transformation which, on the other hand, can be absorbed into the definition of U_d^A. For concreteness, we fix this arbitrariness by defining

$$\Pi^A=\sum_{l=1}^{R}|k\rangle_R^{A}{}_R^A\langle k|. \tag{G6}$$

Thus when starting with the initial state $|i\rangle_S^A|0\rangle_E^A$, we end up with the (unnormalized) state

$$\sum_{j=1}^{T} {}_T^A\langle j|U_e^A|i\rangle_S^A\Pi^A U_d^A|j\rangle_T^A(\alpha_{jA}|0\rangle_E^A+\beta_{jA}|j\rangle_E^A). \tag{G7}$$

Party B performs similar operations on its state.

A given protocol is a choice of the dimensions of the different Hilbert spaces at each time, and of the encoding and decoding operations. Below we give some specific examples of these choices.

1. Protocol 1: Multiplexing at the source

In this example, the task is to share an R-dimensional maximally entangled state. The protocol works by preparing an S-level system ($S>R$), allowing it to be transmitted through the noisy channel, and then processing it at the end (i.e., the encoding between the source and transmitter channels is trivial and $T=S$). The protocol allows the two parties to end up with a final state of their R-level systems which has higher fidelity than would have been achieved if the R-level system were simply transmitted directly through the channel. We call this method "multiplexing at source" because we use a source which produces more entanglement than the one we wish to produce at the receivers ($S>R$). It is this which enables us to obtain a state at the receivers which is closer to the required state than that we would have obtained had we started with the source simply producing an R-dimensional entangled state.

Consider the initial state

$$|\psi_{\mathrm{in}}\rangle=\frac{1}{\sqrt{S}}\sum_{i=1}^{S}|i\rangle_S^A|i\rangle_S^B|0\rangle_E^A|0\rangle_E^B. \tag{G8}$$

The first condition defining this protocol is that the encoding stage is trivial. This means that

$${}_T^A\langle j|U_e^A|i\rangle_S^A={}_T^B\langle j|U_e^B|i\rangle_S^B=\delta_{ij}. \tag{G9}$$

Thus the unnormalized state of the system after decoding and projection is

$$|\psi_{\mathrm{fin}}\rangle=\frac{1}{\sqrt{S}}\sum_{i=1}^{S}\Pi^A\Pi^B U_d^A U_d^B|i\rangle_T^A|i\rangle_T^B(\alpha|0\rangle_E^A+\beta|i\rangle_E^A)\times(\alpha|0\rangle_E^B+\beta|i\rangle_E^B). \tag{G10}$$

We are interested in the maximum fidelity to an R-level $|\psi_R\rangle$ singlet that we can produce,

$$|\psi_R\rangle=\frac{1}{\sqrt{R}}\sum_{i=1}^{R}|i\rangle_R^A|i\rangle_R^B. \tag{G11}$$

The fidelity of the state $|\psi_{\mathrm{fin}}\rangle$ is

$$F=\frac{|\langle\psi_R|\psi_{\mathrm{fin}}\rangle|^2}{\langle\psi_{\mathrm{fin}}|\psi_{\mathrm{fin}}\rangle}. \tag{G12}$$

We now introduce the second condition defining the protocol, namely that U_d^A and U_d^B should be related by being essentially the complex conjugates of each other in the bases we are using. That is, if we write

$$U_d^A|i\rangle_T^A=\sum_j u_{ij}|j\rangle_R^A, \tag{G13}$$

then

$$U_d^B|i\rangle_T^B=\sum_j u_{ij}^*|j\rangle_R^B. \tag{G14}$$

This means in particular that

$$U_d^A U_d^B\sum_i|i\rangle_T^A|i\rangle_T^B=\sum_i|i\rangle_R^A|i\rangle_R^B, \tag{G15}$$

i.e., in the absence of noise the receiver obtains a maximally entangled state.

Now let us compute

$$\begin{aligned}\langle\psi_{\mathrm{fin}}|\psi_{\mathrm{fin}}\rangle&=\frac{1}{S}\sum_{i,i'=1}^{S}{}_T^A\langle i'|{}_T^B\langle i'|(U_d^A)^\dagger(U_d^B)^\dagger\Pi^A\Pi^B U_d^A U_d^B|i\rangle_T^A|i\rangle_T^B\\&\quad\times[|\alpha|^4+(1-|\alpha|^4)\delta_{i,i'}]\\&=\frac{|\alpha|^4}{S}\sum_{i,i'=1}^{S}{}_T^A\langle i'|{}_T^B\langle i'|(U_d^A)^\dagger(U_d^B)^\dagger\Pi^A\Pi^B U_d^A U_d^B|i\rangle_T^A|i\rangle_T^B\\&\quad+\frac{1-|\alpha|^4}{S}\sum_{i=1}^{S}{}_T^A\langle i|{}_T^B\langle i|(U_d^A)^\dagger(U_d^B)^\dagger\\&\quad\times\Pi^A\Pi^B U_d^A U_d^B|i\rangle_T^A|i\rangle_T^B.\end{aligned} \tag{G16}$$

It may be calculated that

$$\langle\psi_{\mathrm{fin}}|\psi_{\mathrm{fin}}\rangle=\frac{|\alpha|^4R}{S}+\frac{1-|\alpha|^4}{S}\sum_{i=1}^{S}[{}_T^A\langle i|(U_d^A)^\dagger\Pi^A U_d^A|i\rangle_T^A]^2, \tag{G17}$$

where we have used the fact that

$${}_T^A\langle i|(U_d^A)^\dagger\Pi^A U_d^A|i\rangle_T^A={}_T^B\langle i|(U_d^B)^\dagger\Pi^B U_d^B|i\rangle_T^B. \tag{G18}$$

Also

$$|\langle \psi_R|\psi_{\text{fin}}\rangle|^2 = \frac{|\alpha|^4 R}{S} + \frac{1-|\alpha|^4}{RS}\sum_{i=1}^{S} [{}_T^A\langle i|(U_d^A)^\dagger \Pi^A U_d^A|i\rangle_T^A]^2. \tag{G19}$$

Thus we may write the fidelity as

$$F = \frac{|\langle \psi_R|\psi_{\text{fin}}\rangle|^2}{\langle \psi_{\text{fin}}|\psi_{\text{fin}}\rangle} = \left(\frac{|\alpha|^4 R}{S} + \frac{1-|\alpha|^4}{S}\frac{Y}{R}\right)\left(\frac{|\alpha|^4 R}{S} + \frac{1-|\alpha|^4}{S}Y\right)^{-1}, \tag{G20}$$

where

$$Y = \sum_{i=1}^{S} [{}_T^A\langle i|(U_d^A)^\dagger \Pi^A U_d^A|i\rangle_T^A]^2. \tag{G21}$$

Y is a positive quantity and by Schwarz's inequality

$$Y = \sum_{i=1}^{S} [{}_T^A\langle i|(U_d^A)^\dagger \Pi^A U_d^A|i\rangle_T^A]^2 \geq \frac{1}{S}\left(\sum_{i=1}^{S} {}_T^A\langle i|(U_d^A)^\dagger \Pi^A U_d^A|i\rangle_T^A\right)^2. \tag{G22}$$

But

$$\sum_{i=1}^{S} {}_T^A\langle i|(U_d^A)^\dagger \Pi^A U_d^A|i\rangle_T^A = R. \tag{G23}$$

Thus

$$Y \geq \frac{R^2}{S} \tag{G24}$$

with equality when

$${}_T^A\langle i|(U_d^A)^\dagger \Pi^A U_d^A|i\rangle_T^A = \frac{R}{S} \quad \text{for all } i. \tag{G25}$$

We will impose (G25) as the third condition defining this protocol. In this case, the fidelity is

$$F = \left(\frac{|\alpha|^4 R}{S} + \frac{1-|\alpha|^4}{S}\frac{R}{S}\right)\left[\frac{|\alpha|^4 R}{S} + (1-|\alpha|^4)\left(\frac{R}{S}\right)^2\right]^{-1}. \tag{G26}$$

We see that by increasing the amount of entanglement produced by the source (i.e., by increasing S), the fidelity is increased and tends to 1 for large S.

2. Protocol 2: Multiplexing of the transmission channels

We may also use the protocols for error filtration presented in the main text and Appendices A and B directly to filter errors when communicating entangled quantum states. We can think of the protocols in the main text and in Appendices A and B as ways of improving a given transmission channel: by multiplexing each source channel to T transmission channels, we can reduce the error amplitude from β to $\beta/\sqrt{T}$.

Consider, then, that a source prepares a state of two S-level systems. This state is preprocessed by multiplexing each source channel into T transmission channels using a general encoding as given in Appendix A. The signal is then decoded and postprocessed to yield a state at the two receivers R_A and R_B. The received state will be of higher fidelity than if the pre- and postprocessing had not been used.

Consider, for example, the following input state:

$$|\psi_{\text{in}}\rangle = \sum_{i=1}^{S} a_i |i\rangle_S^A |i\rangle_S^B, \tag{G27}$$

where a_i are complex amplitudes (this is essentially the most general bipartite state). If each source channel is processed through T transmission channels, in such a way that the original error amplitude is reduced from β to $\beta/\sqrt{T}$, then the final state is

$$|\psi_{\text{fin}}\rangle = \sum_{i=1}^{S} a_i |i\rangle_S^A |i\rangle_S^B \left(\alpha|0\rangle_E^A + \frac{\beta}{\sqrt{T}}|i\rangle_E^A\right)\left(\alpha|0\rangle_E^B + \frac{\beta}{\sqrt{T}}|i\rangle_E^B\right). \tag{G28}$$

The fidelity of the state at the receivers to the state which would have been transmitted if there were no noise (i.e., $\Sigma_{i=1}^S a_i|i\rangle_R^A|i\rangle_R^B$) is

$$F = \frac{|\langle \psi_{\text{in}}|\psi_{\text{fin}}\rangle|^2}{\langle \psi_{\text{fin}}|\psi_{\text{fin}}\rangle} = \frac{|\alpha|^4 + [(|\alpha|^2+|\beta|^2/T)^2 - |\alpha|^4]\sum_{i=1}^{S}|a_i|^4}{(|\alpha|^2+|\beta|^2/T)^2}. \tag{G29}$$

Thus the fidelity increases monotonically with T and tends to 1 as $T\rightarrow\infty$.

[1] P. W. Shor, Phys. Rev. A **52**, R2493 (1995).
[2] A. M. Steane, Phys. Rev. Lett. **77**, 793 (1996).
[3] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. Smolin, and W. Wootters, Phys. Rev. Lett. **76**, 722 (1996).
[4] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).
[5] E. Knill and R. Laflamme, Phys. Rev. A **55**, 900 (1997).
[6] P. Zanardi and M. Rasetti, Phys. Rev. Lett. **79**, 3306 (1997).
[7] D. A. Lidar, I. L. Chuang, and K. B. Whaley, Phys. Rev. Lett. **81**, 2594 (1998).
[8] S. L. Braunstein, Nature (London) **394**, 47 (1998).
[9] S. Lloyd and J.-J. E. Slotine, Phys. Rev. Lett. **80**, 4088 (1998).
[10] D. Bouwmeester, Phys. Rev. A **63**, 040301 (2001).
[11] J.-W. Pan, C. Simon, C. Brukner, and A. Zeilinger, Nature

(London) **410**, 1067 (2001).
[12] L.-M. Duan, M. D. Lukin, J. I. Cirac, and P. Zoller, Nature (London) **414**, 413 (2001).
[13] J.-W. Pan, S. Gasparoni, R. Ursin, G. Weihs, and A. Zeilinger, Nature (London) **423**, 417 (2003).
[14] T. Kløve and V. Korzhik, *Error Detection Codes* (Kluwer, Dordrecht, 1995).
[15] A. E. Ashikhmin, A. M. Barg, E. Knill, and S. N. Litsyn, IEEE Trans. Inf. Theory **46**, 778 (2000).
[16] A. E. Ashikhmin, A. M. Barg, E. Knill, and S. N. Litsyn, IEEE Trans. Inf. Theory **46**, 789 (2000).
[17] L.-Ph. Lamoureux, E. Brainis, N. J. Cerf, Ph. Emplit, M. Haelterman, and S. Massar, Phys. Rev. Lett. **94**, 230501 (2005).
[18] A. Stern, Y. Aharonov, and Y. Imry, Phys. Rev. A **41**, 3436 (1990).
[19] M. Nielsen and I. Chuang, *Quantum Information and Quantum Computation* (Cambridge University Press, Cambridge, 2000).
[20] J. Brendel, W. Tittel, H. Zbinden, and N. Gisin, Phys. Rev. Lett. **82**, 2594 (1999).
[21] D. Kalamidas, Phys. Lett. A **321**, 87 (2004).
[22] A. Barenco *et al.*, SIAM J. Comput. **26**, 1541 (1997).
[23] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Phys. Rev. Lett. **73**, 58 (1994).