REVIEWS OF MODERN PHYSICS, VOLUME 86, OCTOBER–DECEMBER 2014

Quantum channels and memory effects

Filippo Caruso*

QSTAR, Largo Enrico Fermi 2, I-50125 Firenze, Italy,
LENS and Università di Firenze, via Carrara 1, I-50019 Sesto Fiorentino, Italy,
Dipartimento di Fisica e Astronomia, Università di Firenze, via Sansone 1,
I-50019 Sesto Fiorentino, Italy,
and Institut für Theoretische Physik, Universität Ulm, Albert-Einstein-Allee 11,
D-89069 Ulm, Germany

Vittorio Giovannetti†

NEST, Scuola Normale Superiore and Istituto Nanoscienze-CNR, Piazza dei Cavalieri 7,
I-56126 Pisa, Italy

Cosmo Lupo‡

MIT, Research Laboratory of Electronics, 77 Massachusetts Avenue,
Cambridge, Massachusetts 02139, USA,
and School of Science and Technology, University of Camerino, I-62032 Camerino, Italy

Stefano Mancini§

School of Science and Technology, University of Camerino, I-62032 Camerino, Italy
and INFN-Sezione di Perugia, Via A. Pascoli, I-06123 Perugia, Italy

(published 10 December 2014)

Any physical process can be represented as a quantum channel mapping an initial state to a final state. Hence it can be characterized from the point of view of communication theory, i.e., in terms of its ability to transfer information. Quantum information provides a theoretical framework and the proper mathematical tools to accomplish this. In this context the notion of codes and communication capacities have been introduced by generalizing them from the classical Shannon theory of information transmission and error correction. The underlying assumption of this approach is to consider the channel not as acting on a single system, but on sequences of systems, which, when properly initialized allow one to overcome the noisy effects induced by the physical process under consideration. While most of the work produced so far has been focused on the case in which a given channel transformation acts identically and independently on the various elements of the sequence (memoryless configuration in jargon), correlated error models appear to be a more realistic way to approach the problem. A slightly different, yet conceptually related, notion of correlated errors applies to a single quantum system which evolves continuously in time under the influence of an external disturbance which acts on it in a non-Markovian fashion. This leads to the study of memory effects in quantum channels: a fertile ground where interesting novel phenomena emerge at the intersection of quantum information theory and other branches of physics. A survey is taken of the field of quantum channels theory while also embracing these specific and complex settings.

DOI: 10.1103/RevModPhys.86.1203 PACS numbers: 03.67.−a, 89.70.−a, 02.50.−r

CONTENTS

*filippo.caruso@lens.unifi.it
†v.giovannetti@sns.it
‡cosmo.lupo@unicam.it
§stefano.mancini@unicam.it

I. INTRODUCTION

In his seminal 1948 work "A Mathematical Theory of Communication," C. E. Shannon established the basis of modern communication technology (Verdú, 1998). Neglecting all the semantic aspects (which are irrelevant at the level of engineering) he stressed that "the fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point." In particular, "the system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design" (Shannon, 1948). At the heart of this view is what one may call the "channel" formalism, where any noisy communication line is depicted as a stochastic map connecting input signals selected by the sender of the message (Alice), who is operating at one end of the line, to their corresponding output counterparts accessible to the receiver of the messages (Bob), who is operating at the other end. In the same article Shannon also proved that the performance of a transmission line can be gauged by a single quantity, the *capacity* of the channel, which measures the maximum rate at which information can be reliably transferred when Alice and Bob, operating on long sequences of transmitted signals, follow a preestablished protocol (error-correcting code procedure) aimed to nullify the detrimental effects of the communication noise. The rationale behind this approach (which is typical to communication theory) is that communication is expensive while local operations are somehow free (unless external constraints are explicitly imposed by the selected implementation).

Rolf Landauer was the first to put on firm ground the fact that information is not just an abstract, mathematical notion but has instead an intrinsic physical nature which poses limits on the possibility of processing and transferring it (Landauer, 1961). That is why quantum mechanics, the most advanced physical theory, comes into play in the study of communication processes (Bennett and Shor, 1998). In this context (*quantum information theory*) it is recognized that any message two parties wish to exchange must be written into the states of some quantum system, say a photonic pulse propagating along an optical fiber (Caves and Drummond, 1994), and that the processing, the transmission, and the reading of such data must be carried out following the rather unconventional prescriptions established by quantum mechanics. As in the classical setting this scenario is properly formalized by introducing the notion of *quantum channels* as those mappings which, generalizing the notion of a channel in Shannon theory, link the initial states of the quantum

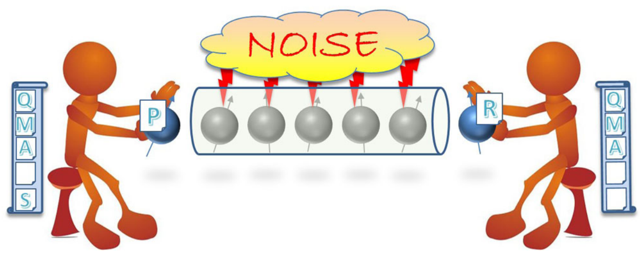


FIG. 1 (color online). Typical communication scenario of two users communicating by encoding (decoding) the letters of a message onto (from) physical systems that are transmitted through a channel. They are subjected to the unavoidable presence of noise in the channel that introduces errors (eventually correlated) in the transmitted message.

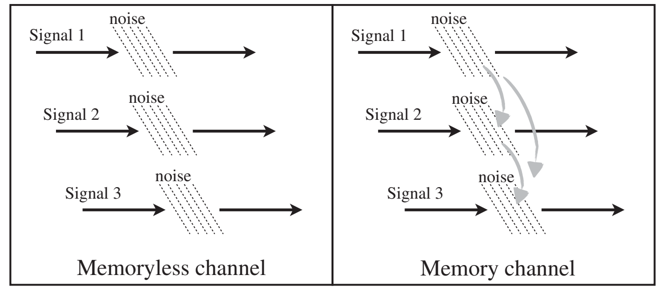


FIG. 2. Pictorial representation of memoryless (left panel) and memory (right panel) channels. In the former case the individual carriers which compose a sequence of transmitted signals experience the same noisy transformation. In the latter case instead cross talking among the various transmission events can happen and, as indicated by the gray arrows in the figure, the noise model which describes the transformation of the nth carrier depends in principle upon the previous communication exchanges.

information carriers (controlled by Alice) to their output states (controlled by Bob); see Fig. 1. Interestingly, to evaluate the quality of these exotic communication lines several non-equivalent notions of coding procedures, as well as of corresponding capacities, must be introduced. Indeed, while Alice and Bob might still be willing to use a quantum channel to exchange purely classical messages, new forms of communication can now be envisioned. For instance, Alice may be interested in transferring to Bob purely quantum messages, e.g., the *unknown quantum state* of a quantum memory that is located in her laboratory, or half of a *maximally entangled* state (Gühne and Tóth, 2009; Horodecki *et al.*, 2009) that she has locally produced. The ability to sustain this special kind of transmission defines what is called the quantum capacity of a quantum communication line. This in general differs from the classical capacity that instead, as in the Shannon setting, measures the ability to transfer purely classical messages. But there is more. Quantum teleportation (Bennett *et al.*, 1993) and the superdense coding protocol (Bennett and Wiesner, 1992) have shown that quantum *entanglement* (Gühne and Tóth, 2009; Horodecki *et al.*, 2009) is a catalytic resource for communication. Indeed, even though entanglement alone does not constitute a communication link between distant parties (Alber *et al.*, 2001), allowing Alice and Bob to use preshared entanglement in the design of their communication protocols can boost the performance of basically any communication line they have access to (even in terms of the quantum capacity). This fact naturally brings in the notion of *entanglement-assisted* capacities of a quantum channel (Bennett, Shor *et al.*, 1999), which is yet a different way of gauging the performance of a communication line.

The vast majority of the work on quantum channels has been concerned with the study of memoryless configurations where sequences of exchanged quantum carriers are supposed to undergo the action of noisy transformations which affect them independently and identically; see the left panel of Fig. 2. In this scenario coding theorems have been derived which allow one to express the various capacities of the communication line in terms of rather compact entropic formulas. For instance, the classical (respectively, quantum) capacity of a memoryless quantum channel is characterized in terms of the Holevo (respectively, coherent) information (Schumacher and Nielsen, 1996; Lloyd, 1997; Schumacher and Westmoreland, 1997; Barnum, Nielsen, and Schumacher, 1998; Barnum, Smolin, and Terhal, 1998; Holevo, 1998a; Shor, 2002a; Devetak, 2005). The memoryless assumption is indeed a useful hypothesis which permits one to simplify the input-output mapping induced by the noise. It also provides a realistic description for those communication schemes where the temporal rates at which signals are fed into the communication line are sufficiently low to allow for a resetting of the channel environment and to prevent signal cross talking. Nonetheless this is not always justified. For instance, with increasing signal feeding rates, successive transmissions happen so rapidly that the environment may retain a “memory” of past events; see the right panel of Fig. 2. Optical fibers are an example in which such effects can occur and have been explored experimentally (Ball, Dragan, and Banaszek, 2004; Banaszek *et al.*, 2004). Similarly, in quantum information processors, especially in solid-state implementations, qubits may be so closely spaced that the same environmental degree of freedom will interact jointly with several of them (even if they are not nearest neighbors) leading to cross talks and correlations in the noise (Duan and Guo, 1998; Hu, Zhou, and Guo, 2007). The consideration of spatial and temporal memory effects is therefore becoming increasingly pressing with the continuing miniaturization of information processing devices and with increasing communication rates through channels. Moreover, from a fundamental point of view, quantum memory channels provide a general framework which encompasses the memoryless ones as a special case.

Apparently, the interest toward information transmission through quantum channels with memory spread after a model introduced by Macchiavello and Palma (2002). Here an example of a qubit channel with Markovian correlated noise was analyzed in which the encoding of information by means of entangled input states may increase the transmission rate of classical information. Subsequently, the study of quantum channels with memory has largely been confined to channels with Markovian correlated noise with the aim of deriving bounds on the classical capacity; see, e.g., Hamada (2002), Bowen and Mancini (2004), and Bowen, Devetak, and Mancini (2005). Then, coding theorems have been devised for a class of quantum memory channels having structural properties that guarantee "regular" asymptotic behavior (Kretschmann and Werner, 2005; Datta and Dorlas, 2007; Bjelaković and Boche, 2008). This approach can be traced back to the work of Dobrushin (1963), who considered a wide class of channels exhibiting stationary or ergodic behavior. A completely different path was taken by Hayashi and Nagaoka (2003), who applied the "information-spectrum" method to obtain a coding theorem for the classical capacity, following the work by Verdú and Han (1994) on classical channels with memory. In the context of continuous-variable (CV) systems (Braunstein and van Loock, 2005; Weedbrook *et al.*, 2012), generalizing results obtained in the classical setting (Shannon, 1949) for the capacity of power-constrained Gaussian channels, quantum "water-filling" formulas have been recently derived (Schäfer, Karpov, and Cerf, 2009; Pilyavets, Lupo, and Mancini, 2012). Exact expressions for classical and quantum capacities have been computed (Lupo, Giovannetti, and Mancini, 2010a) using an "unraveling" technique based on the Toeplitz distribution theorem (Gray, 1972), which allows one to map memory correlations into effective memoryless models. All these results for continuous-variable memory channels have been derived from fundamental results in the memoryless setting (Giovannetti *et al.*, 2004c; Wolf, Pérez-García, and Giedke, 2007) [see also recent developments in Giovannetti, Holevo, and García-Patrón (2013) and Giovannetti, García-Patrón *et al.* (2013)].

Beyond quantum communication, a detailed study of quantum channels, and of the mechanisms responsible for memory correlations, has implications in the broader research field of quantum open system dynamics (Breuer and Petruccione, 2002). As a matter of fact the input-output scheme that underlines the channel formalism reminds us of what in physics is conventionally described as a series of scattering events, the scattered particles playing the role of the messages, while the scattering matrix plays the role of the communication line. More generally quantum channels can be used to mimic all those physical processes (temporal evolution, data processing, etc.) which imply a state change of a system of interest from an initial to a final configuration under the influence of an external agent (the system environment). By exploiting this connection, insight into the system evolution can then be gained by analyzing its quality as a communication line. Along this direction models have been introduced in which memory effects of a communication line are described as arising from the interaction with a multipartite environment initialized in a correlated state (Giovannetti and Mancini, 2005), allowing remarkable links between information theoretical quantities, such as capacities, and statistical properties, such as phase transitions, of the underlining many-body environment (Plenio and Virmani, 2007, 2008).

When studying open system dynamics one shall not only deal with the problem of the input-output evolution of a sequence of otherwise independent carriers. Indeed it is also interesting to address the problem of the evolution of a single carrier in time, to see whether the associated trajectory can be described as a collection of quantum channels which are applied sequentially on that system; see Fig. 3. When this is not the case one can talk of memory effects induced by backaction mechanisms arising from the interaction of the

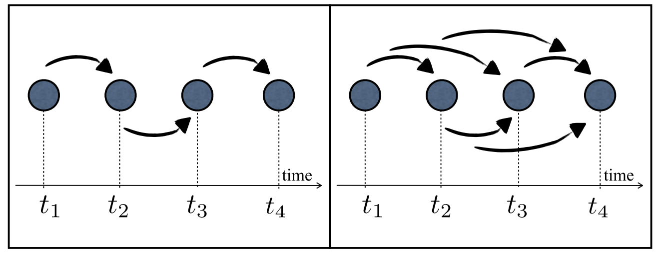


FIG. 3 (color online). Temporal correlations. Left panel: Pictorial representation of a trajectory which describes the evolution of a system under dynamical semigroup approximation (discrete evolution steps have been assumed for simplicity). In this case, as indicated by the arrows, the state of the system of interest (dots) at time t_{j+1} depends only upon the state the system had at time step immediately before, i.e., at t_j. Right panel: Representation of the evolution when the dynamical semigroup approximation fails; in this case the state at time t_{j+1} depends in general on the states the system assumed along the whole trajectory.

system of interest and its own environment. At variance from those described these effects have a clear dynamical character which is absent in the scheme of Fig. 2 where the temporal evolution is fixed. The study of this topic is intimately related to the semigroup structure of the set of quantum channels, hence with dynamical maps and master equations (Alicki and Lendi, 1987). It turns out that dynamical evolutions which can be split into infinitesimal pieces correspond to the set of solutions of (possibly time-dependent) master equations standardly used to describe open systems dynamics (Wolf and Cirac, 2008). A more general approach to open quantum systems uses the Nakajima-Zwanzig projection operator technique (Nakajima, 1958; Zwanzig, 1960) which shows that, under fairly general conditions, the master equation for the reduced density operator takes the form of a nonlocal equation in which memory effects are taken into account through the introduction of a memory kernel. Then, the problem [first put forward by Daffer, Wódkiewicz, and McIver (2003) and Daffer *et al.* (2004)] becomes finding those conditions on the memory kernel ensuring that the time evolution map is a bona fide quantum channel (Chruściński and Kossakowski, 2012). While our review is mostly devoted to analyze the memory effects which arise in the input-output paradigm schematized in Fig. 2, for completeness we also briefly report on the most recent results which have being produced in the study of dynamical maps.

This work aims at providing an overview of the field of quantum channels in a broad framework that also includes memory effects. As such it does not pretend to be omnicomprehensive, but it rather touches quantum communication subjects for which it has been already possible to venture beyond memoryless assumptions. We start considering in Sec. II quantum channel maps and parallel them to physical processes transforming input states into output ones. This presumes basic knowledge on the structure of quantum states, entanglement, and measurement that are not reviewed here; a rather complete report on these topics can indeed be found elsewhere, e.g., in the books (Preskill, 1998; Nielsen and Chuang, 2000; Bengtsson and Zyczkowski, 2006; Petz, 2008; Schumacher and Westmoreland, 2010; Holevo, 2011) or in the review articles (Gühne and Tóth, 2009; Horodecki *et al.*, 2009). Here the focus is on the representation of quantum channel maps and their properties related to the way they act on quantum states, in particular, their composability. For pedagogical reasons we present several examples of quantum channels. In Sec. II we provide tools to study quantum channels such as fidelities, distances, and entropies that are used throughout the paper.

We then move on discussing in Sec. III the transition from the memoryless setting to a more general scenario which allows for correlations in quantum communication. Here much attention is devoted to the structural properties of quantum channels. Various classes of memory channels are then reviewed in Sec. III.D.

In order to use a quantum channel for information transmission one has to cope with the problem of noise altering the transmitted information. For this reason we next present in Sec. IV the subject of quantum error correction and discuss achievable information transmission rates. Actually, this section briefly reviews basic notions of standard quantum error correction (mostly suitable for uncorrelated errors) and decoherence-free subspaces (mostly suitable for completely correlated errors). For more details the interested reader can refer to Lidar and Brun (2013). We then present avenues, not yet fully explored, for correcting partially correlated errors and discuss convolutional codes that work with a structure similar to that of memory channels.

After having introduced the notion of transmission rates, it is natural to ask what are their maximum rates that can be achieved in a quantum channel. Thus we address the issue of quantum channel capacities in Sec. V. A series of papers, at various levels, deal with capacities of memoryless quantum channels. For instance, the article by Bennett and Shor (1998) can be seen as a sort of *manifesto* for quantum information theory. The review by Caves and Drummond (1994) presents instead a rather detailed account of the mathematical and technological issues one faces when dealing with quantum communication with photonic sources (even if some of the open problems discussed there were solved in more recent years, this article remains a useful guide to the field). Galindo and Martín-Delgado (2002) provides a rather compact overview on quantum information theory and discusses in a simple but clear form the basic aspects of the Shannon approach. A more mathematically oriented point of view is presented by Keyl (2002). Holevo and Giovannetti (2012) focus on channel capacities and their entropic characterization, while finally Weedbrook *et al.* (2012) give a detailed introduction to the field of Gaussian bosonic channels. In Sec. V coding theorems that allow one to express capacities in a closed form by means of entropic quantities are succinctly reviewed for memoryless quantum channels. We then indulge in the possibility of using them in the memory setting (revisiting what kind of memory permits it) and on their generalization.

For practical purposes we subsequently present in Sec. VI quantum channel models that are exactly solvable in terms of capacities. Already in the memoryless case these examples are few and they are even less in the memory case. Anyway it is much instructive to see techniques used to solve optimization problems imposed by capacity evaluation.

Finally Sec. VII is devoted to the characterization of the temporal correlations which may arise in the description of the trajectory of quantum system (see Fig. 3). In particular, we review the divisibility property of quantum channels and relate it to properties of dynamical maps and master equations.

A summary of the main results and an outlook on physical realizations are given in Sec. VIII. Appendixes A and B provide elementary material about distance measures for states and quasilocal algebras, respectively. In Appendix C an alternative proof of the structure decomposition theorem for nonanticipatory channels is presented, while in Appendix D an explicit derivation of capacity upper bounds is provided.

II. QUANTUM CHANNELS: BASIC DEFINITIONS AND PROPERTIES

In a typical communication scenario two parties (Alice the sender of the message and Bob the receiver) aim to exchange (classical or quantum) information by encoding it into (possibly arbitrarily long) sequences of signals which propagate through the medium that separates them; see Fig. 1. A train of transmitted signals defines a sequence of independent uses of the communication line (*channel uses*), and their input-output evolution from Alice to Bob is determined by the noise which tampers with the transmission process. In classical information theory (Gallager, 1968) this is schematized by assigning an input alphabet $\mathcal{X}$ and an output alphabet $\mathcal{Y}$ whose elements x and y represent, respectively, the individual signals at the input and output of the transmission line. The noise instead is assigned in terms of a stochastic process characterized by conditional probabilities that, given an input sequence $(x_1, x_2, \ldots)$ of elements of $\mathcal{X}$ transmitted by Alice, Bob will receive the sequence $(y_1, y_2, \ldots)$ of elements of $\mathcal{Y}$.

In quantum information the channel uses are represented by the degrees of freedom (e.g., polarization, spins) of a collection $\{q_1, q_2, \ldots\}$ of identical information carrying objects (e.g., optical pulses, flying atoms, or ions) which are locally produced by Alice and organized in a time-ordered sequence. In this setting the noise can then be described by assigning a proper mapping which acts on the (global) input states of the information carriers to produce the associated (global) output states received by Bob. The formalism is rather general and provides the proper mathematical tools apt to describe all those physical processes that involve the transformation of a quantum system, induced either by the direct temporal evolution of its density matrix on a fixed time interval, or by the transmission through a medium (see Fig. 4). Concrete examples of these mappings can be encountered for instance when studying long-distance quantum communication and cryptography (as quantum key distribution) (Gisin *et al.*, 2002; Scarani *et al.*, 2009). In this case applications are often experimentally realized by identifying the information carriers with single-photon pulses that travel in free space or over optical fibers, where air turbulence and absorption losses effectively limit the covered distance from tens to hundreds of kilometers with the currently available technologies (Hughes *et al.*, 2002; Ursin *et al.*, 2004, 2007; Gisin and Thew, 2007; Schmitt-Manderbach *et al.*, 2007; Villoresi *et al.*, 2008; Ma *et al.*, 2012; Yin *et al.*, 2013). All these effects can be

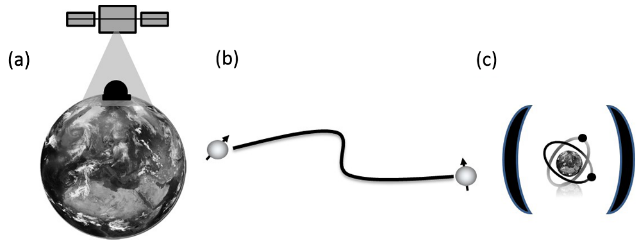


FIG. 4 (color online). Some examples of communication lines which can be described by the quantum channel formalism: (a) quantum communication with Earth- and space-based transmitter terminals, (b) transmission of photons along optical fibers, and (c) storing information on atoms or molecules trapped in optical cavities. The evolution, in space or time, of a quantum state can always be described by the formalism of quantum channels. The potential presence of temporally or spatially correlated noise may lead, then, to dealing with memory quantum channels.

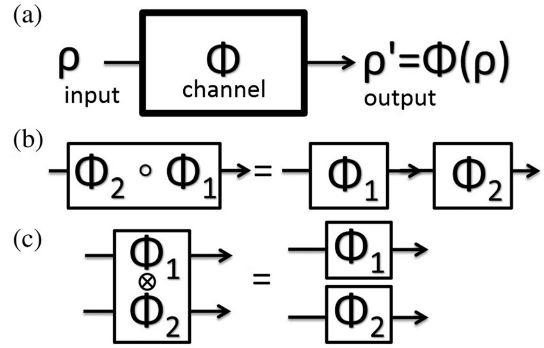


FIG. 5. (a) Quantum channel Φ mapping a state ρ into a new one ρ', (b) concatenation $\Phi_1 \circ \Phi_2$ of two channels Φ_1 and Φ_2, and (c) tensor product $\Phi_1 \otimes \Phi_2$ of two channels Φ_1 and Φ_2.

faithfully described in terms of a combined action of amplitude damping channels (dissipation and absorption) and phase-flip channels (dephasing phenomena) (see Sec. II.I). Similarly, further examples of input-output mapping which admit a proper characterization in terms of quantum channels can be found when analyzing the effectiveness of atomic or molecular systems trapped in an optical cavity (Leibfried *et al.*, 2003) as quantum memory elements useful for information storage. In all these physical implementations, the noise processes may sometimes show temporal or spatial correlations, leading unavoidably to the additional presence of memory effects in the corresponding quantum channel representation. Section III discusses in detail how such effects can be characterized. Before doing so, however, it is useful to recall that quantum mechanics imposes some fundamental structural constraints on the transformations describing the evolution of quantum systems, which must apply independently from the underlying physical mechanisms that govern the process and independently from the composite nature of the input system.

A. Completely positive and trace preserving (CPTP) transformations

Let Φ be a mapping [see Fig. 5(a)] describing the input-output relations of a generic quantum system Q (e.g., the carriers $\{q_1, q_2, \ldots\}$ introduced previously) evolving under the action of some physical process

$$\rho_Q \in \mathfrak{S}(\mathcal{H}_Q) \mapsto \rho_{Q'} = \Phi(\rho_Q) \in \mathfrak{S}(\mathcal{H}_{Q'}). \quad (1)$$

Here $\mathfrak{S}(\mathcal{H}_Q)$ and $\mathfrak{S}(\mathcal{H}_{Q'})$ stand for the sets of density operators (non-negative operators with unit trace) defined on the Hilbert spaces $\mathcal{H}_Q$, $\mathcal{H}_{Q'}$ (the latter may be different in general) associated, respectively, to the input and output systems (unless explicitly stated in what follows it is assumed that these spaces are finite dimensional). Since $\rho_{Q'}$ must be a valid density operator, it results naturally to require the map Φ to be as follows:

(i) linear when extended to the set $\mathcal{T}(\mathcal{H}_Q)$ of trace-class linear operators of $\mathcal{H}_Q$. As a matter of fact Φ must transform mixtures of input density operators into a mixture of the associated outputs, i.e., $\sum_i p_i \rho_Q(i) \mapsto \sum_i p_i \Phi(\rho_Q(i))$, with p_i the probability associated with the input state $\rho_Q(i)$[1];

(ii) trace preserving (i.e., it must preserve the normalization of all input states);

(iii) positive (i.e., when acting on Q it must preserve the positivity of density operators).

Actually the latter condition turns out to not be enough to guarantee the positivity of $\Phi(\rho_Q)$ when considering ρ_Q as coming from a joint state ρ_{QA} of system Q and system A by tracing out the latter. This is due to possible quantum correlations (entanglement) existing between systems Q and A. Hence, condition (iii) is made tighter as follows:

(iii′) completely positive (i.e., when acting on Q the map Φ must preserve the positivity of any density operator, including those describing a joint state ρ_{QA} of Q and an arbitrary ancillary system A).

A violation of any conditions (i), (ii), or (iii′) implies the impossibility of maintaining the statistical interpretation of the theory (Stinespring, 1955; Sudarshan, Mathews, and Rau, 1961; Holevo, 1972; Lindblad, 1975; Kraus, 1983).

Finally any transformation fulfilling all these conditions is said to be a CPTP linear map and is a quantum channel. Special instances of these last maps are provided by the isometric channels

$$\rho_Q \mapsto \mathcal{U}(\rho_Q) \coloneqq U \rho_Q U^\dagger, \quad (2)$$

which are induced by the action of an isometric transformation U connecting $\mathcal{H}_Q$ to $\mathcal{H}_{Q'}$, i.e., $U^\dagger U = \mathbb{1}$ with $U^\dagger$ being the adjoint of U and $\mathbb{1}$ being the identity on $\mathcal{H}_Q$. In particular, when $Q = Q'$ and U corresponds to a unitary operator, Eq. (2) defines a *unitary channel* on Q which admits the channel $\mathcal{U}^{-1}(\cdots) \coloneqq U^\dagger(\cdots)U$ as a CPTP inverse. Furthermore, if U is the identity operator $\mathbb{1}$ on $\mathcal{H}_Q$, the resulting transformation is the identity channel, denoted id, which maps any state into itself, i.e., $\mathrm{id}(\rho_Q) = \rho_Q$ for all ρ_Q.

B. Composition rules and structural properties

While referring the interested reader to Breuer and Petruccione (2002), Keyl (2002), Bengtsson and Zyczkowski

[1]The linearity requirement ensures that the extension of Φ from $\mathfrak{S}(\mathcal{H}_Q)$ to $\mathcal{T}(\mathcal{H}_Q)$ is unique.

(2006), Petz (2008), and Holevo (2011, 2012) for an exhaustive characterization, here the most relevant structural properties of the set $\mathfrak{P}(Q \mapsto Q')$, formed by the CPTP maps connecting system Q to system Q', are reviewed.

1. Convexity

Given $\Phi, \Lambda \in \mathfrak{P}(Q \mapsto Q')$, and $p \in [0,1]$, the transformation

$$\rho_Q \mapsto p\Phi(\rho_Q) + (1-p)\Lambda(\rho_Q) \tag{3}$$

is still an element of $\mathfrak{P}(Q \mapsto Q')$. For instance, convex combinations of unitary channels on Q define the random unitary channel subset of $\mathfrak{P}(Q \mapsto Q)$.

2. Concatenation of channels

Given two CPTP channels, Φ from Q to Q' and Ω from Q' to Q'', one can define [see Fig. 5(b)] their concatenation $\Omega \circ \Phi$ as the following CPTP transformation from Q to Q'':

$$\rho_Q \mapsto (\Omega \circ \Phi)(\rho_Q) := \Omega(\Phi(\rho_Q)). \tag{4}$$

Channels concatenation allows one to introduce a relation of equivalence between CPTP maps. In particular, two maps $\Phi, \Lambda \in \mathfrak{P}(Q \mapsto Q')$ are said to be unitarily equivalent if there exist unitary channels $\mathcal{U} \in \mathfrak{P}(Q \mapsto Q)$ and $\mathcal{U}' \in \mathfrak{P}(Q' \mapsto Q')$ such that

$$\Phi = \mathcal{U}' \circ \Lambda \circ \mathcal{U}, \tag{5}$$

the relation being reversible in $\Lambda = \mathcal{U}^{-1} \circ \Phi \circ \mathcal{U}'^{-1}$. From the above properties it follows also that $\mathfrak{P}(Q \mapsto Q)$ equipped with the concatenation rule (4) possesses a non-Abelian semigroup structure, the channel id being the identity element of the set and the unitary channels $\mathcal{U}$ being the only invertible elements.

3. Tensor product of channels

Given two CPTP channels, Φ from Q to Q' and Ω from R to R', one can define [see Fig. 5(c)] their tensor product $\Phi \otimes \Omega$ as the CPTP transformation from the composite system QR to the composite system $Q'R'$, which given an arbitrary tensor operator $N_Q \otimes M_R \in \mathcal{T}(\mathcal{H}_Q \otimes \mathcal{H}_R)$ transforms it into

$$(\Phi \otimes \Omega)(N_Q \otimes M_R) := \Phi(N_Q) \otimes \Omega(M_R). \tag{6}$$

Special instances are provided by the transformation $\Phi \otimes \text{id}$ obtained by tensoring Φ with the identity channel acting on an external system: such a map is called an extension of Φ and represents the action of such channel when the system Q (where Φ was originally defined) is described as part of an enlarged composite system—notice that this structure was implicitly assumed when stating point (iii′) of Sec. II.A.

Concatenations and tensor products of quantum channels represent two alternative ways of composing CPTP maps which, to some extent, mimic, respectively, the in-series and in-parallel composition rules of electrical circuit elements. In particular, as discussed in Sec. VII, channel concatenation is naturally suited to characterize the temporal correlations of a single quantum system schematized in the left panel of Fig. 3 (the sequential applications of CPTP maps corresponding to different stages of the system evolution). On the contrary the tensor product (6) allows one to describe spatial correlations which might be present in the evolution of composite quantum systems. Also, as discussed in Sec. III, tensor products can be employed to describe the transformations that a sequence of information carriers encounters when transmitted through a communication line; see Fig. 2.

C. Stinespring representation, Kraus representation, and Choi-Jamiolkowski isomorphism

It can be shown (Stinespring, 1955) that a mapping (1) satisfies the CPTP conditions detailed in Sec. II.B if and only if it admits dilations that allow one to represent it in terms of a unitary coupling with an external environment E (which is possibly fictitious and may not correspond to the actual physical environment responsible for the system evolution). For instance, taking for simplicity $Q = Q'$, one can write

$$\Phi(\rho_Q) = \text{Tr}_E[U_{QE}(\rho_Q \otimes \omega_E)U_{QE}^\dagger], \tag{7}$$

where ω_E is a fixed state of E, U_{QE} is the unitary transformation coupling the latter to the input system Q, and Tr_E denotes the partial trace over the environment.[2] Equation (7) is not unique. Nonetheless by enlarging the environment E to describe the environment state as a pure state, $\omega_E = |\omega\rangle_E\langle\omega|$ (see Sec. II.J.1 for a proper definition of this purification mechanism), the choice of U_{QE} can be shown to be unique up to a local isometric transformation on E. Under this condition the dilation (7) provides what is generally known as the Stinespring representation for Φ.

The CPTP conditions are also equivalent to the possibility of expressing Φ in operator sum (or Kraus) representation (Sudarshan, Mathews, and Rau, 1961; Kraus, 1971; Choi, 1975),

$$\Phi(\rho_Q) = \sum_j K_j \rho_Q K_j^\dagger, \tag{8}$$

with $\{K_j\}$ being operators on $\mathcal{H}_Q$ satisfying the normalization condition $\sum_j K_j^\dagger K_j = \mathbb{1}$. The number of nonzero operators in Eq. (8) is called the Kraus rank. As in the case of the unitary dilation (7), their choice is in general not unique. One can, however, guarantee that a Kraus representation exists with no more than d^2 elements (d being the dimension of $\mathcal{H}_Q$). Kraus and Stinespring representations can also be put in mutual correspondence by identifying the operator K_j with the linear operator ${}_E\langle e_j|U_{QE}|\omega\rangle_E$ of $\mathcal{H}_Q$, where $\{|e_j\rangle_E\}$ is an orthonormal basis of E.

It is finally worth recalling that there exists a fundamental relation, known as the Choi-Jamiolkowski (CJ) isomorphism (Jamiolkowski, 1972; Choi, 1975), which permits one to describe any CPTP Φ as a density operator of a composite

[2] Note that complete positivity can be violated if the initial state of Q is entangled with the channel environment, in which case, however, the mapping represented by Eq. (8) is defined only on a proper subset of $\mathfrak{S}(\mathcal{H}_Q)$ (Jordan, Shaji, and Sudarshan, 2004; Shaji and Sudarshan, 2005).

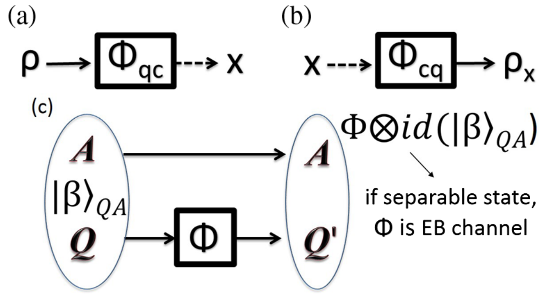


FIG. 6 (color online). (a) qc channels mapping a quantum state ρ into a classical symbol x, while the reversed mapping is represented by cq channels (b); (c) definition of the Choi-Jamiolkowski (CJ) state in terms of a channel Φ and a maximally entangled state $|\beta\rangle_{QA}$ of the system Q and an auxiliary system A subjected to the identity channel. When the CJ state is separable, the corresponding map Φ is called entanglement breaking (EB).

system QA with A being an auxiliary system having the same dimension d as Q; see Fig. 6(c). The explicit connection is obtained by applying the map Φ to half of a maximally entangled state (Gühne and Tóth, 2009; Horodecki et al., 2009) $|\beta\rangle_{QA} = \sum_j |e_j\rangle_Q \otimes |e_j\rangle_A / \sqrt{d}$ of QA to create the so-called CJ state of the channel

$$\rho_{QA}^{(\Phi)} \coloneqq (\Phi \otimes \mathrm{id})(|\beta\rangle_{QA}\langle\beta|), \tag{9}$$

where id stands for the identity map on A and $\{|e_j\rangle_Q\}$ and $\{|e_j\rangle_A\}$ represent orthonormal basis of Q and A, respectively.

D. Heisenberg picture: The dual channel

Equation (1) implicitly assumes the Schrödinger picture in which the states of the system are evolved while the observables are kept fixed. In the Heisenberg picture, in which instead the states are fixed and the observables evolve in time, the CPTP transformation $\Phi \in \mathfrak{P}(Q \mapsto Q')$ is replaced by its dual map

$$O \in \mathcal{B}(\mathcal{H}_{Q'}) \mapsto \Phi^*(O) \in \mathcal{B}(\mathcal{H}_Q), \tag{10}$$

operating on the bounded operator algebra $\mathcal{B}(\mathcal{H}_{Q'})$ of the receiver observable and defined through the identity

$$\mathrm{Tr}[\Phi(\rho_Q)O] = \mathrm{Tr}[\rho_Q \Phi^*(O)], \tag{11}$$

which holds for all $O \in \mathcal{B}(\mathcal{H}_{Q'})$ and for all $\rho_Q \in \mathfrak{S}(\mathcal{H}_Q)$. The Heisenberg-picture transformation Φ^* is linear and completely positive, but in general it is not trace preserving. On the other hand, it is always unital, i.e., it maps the identity operator into itself. Operator sum representations for Φ^* can be easily constructed from those of Φ [Eq. (8)], yielding

$$\Phi^*(O) = \sum_j K_j^\dagger O K_j. \tag{12}$$

Notice also that in the dual picture the concatenation of channels goes in reverse order with respect to the Schrödinger picture, i.e., given Φ and Ω CPTP maps,

$$(\Omega \circ \Phi)^* = \Phi^* \circ \Omega^*. \tag{13}$$

E. cq and qc channels

Besides considering physical transformations which represent the evolution of quantum carriers, in quantum information it is useful to describe processes which map classical inputs into quantum states (cq channels) or, vice versa, quantum states into classical outputs (qc channels); see Figs. 6(a) and 6(b). Specifically the former define state preparation procedures where a symbol x extracted from a classical alphabet $\mathcal{X}$ with probability p_x is encoded into a state $\rho_Q^{(x)}$ of the quantum system Q, thus producing an average density operator $\sum_{x\in\mathcal{X}} p_x \rho_Q^{(x)}$. qc channels instead correspond to measurement procedures which, given $\rho_Q \in \mathfrak{S}(\mathcal{H}_Q)$, produce classical outcomes $x \in \mathcal{X}$ with conditional probabilities

$$p_x(\rho_Q) = \mathrm{Tr}[E_x \rho_Q], \tag{14}$$

with $\{E_x\}_{x\in\mathcal{X}}$ being a set of positive operators on $\mathcal{H}(Q)$, satisfying the normalization condition $\sum_x E_x = \mathbb{1}$, which defines the statistics of the measurement in the positive-operator valued measure (POVM) representation (Breuer and Petruccione, 2002; Petz, 2008; Holevo, 2011). cq and qc channels can both be extended to CPTP maps (1) by introducing an ancillary quantum system Q' of dimension equal to the cardinality of $\mathcal{X}$, and characterized by an orthonormal set $\{|e_x\rangle_{Q'}\}_{x\in\mathcal{X}}$ (Holevo, 1998b). For instance, taking $p_x = {}_{Q'}\langle e_x|\rho_{Q'}|e_x\rangle_{Q'}$ with $\rho_{Q'} \in \mathfrak{S}(\mathcal{H}_{Q'})$, the cq channel defined above induces the following CPTP mapping from Q' to Q:

$$\rho_{Q'} \mapsto \Phi_{\mathrm{cq}}(\rho_{Q'}) \coloneqq \sum_{x\in\mathcal{X}} {}_{Q'}\langle e_x|\rho_{Q'}|e_x\rangle_{Q'} \rho_Q^{(x)}. \tag{15}$$

Analogously the qc channel induces the following CPTP mapping from Q to Q':

$$\rho_Q \mapsto \Phi_{\mathrm{qc}}(\rho_Q) \coloneqq \sum_{x\in\mathcal{X}} |e_x\rangle_{Q'}\langle e_x| \mathrm{Tr}[E_x \rho_Q]. \tag{16}$$

The concatenation $\Phi_{\mathrm{cq}} \circ \Phi \circ \Phi_{\mathrm{qc}}$, with Φ being a generic quantum channel from Q to Q, can also be represented as a CPTP map (from Q' to Q') and describes the typical scenario where a collection of classical messages (represented by elements of the set $\mathcal{X}$) is transferred to Bob via a quantum link (represented by Φ) who "reads" them through the POVM $\{E_x\}_{x\in\mathcal{X}}$. In particular, when applied to elements of the orthonormal set $\{|e_x\rangle_{Q'}\}_{x\in\mathcal{X}}$, $\Phi_{\mathrm{cq}} \circ \Phi \circ \Phi_{\mathrm{qc}}$ induces a classical stochastic process where $x \in \mathcal{X}$ is mapped into $x' \in \mathcal{X}$ with conditional probability

$$p(x'|x) = \mathrm{Tr}[E_{x'} \Phi(|e_x\rangle_{Q'}\langle e_x|)]. \tag{17}$$

F. Entanglement breaking and positive partial transpose (PPT) channels

The cq and qc channels defined in Sec. II.E are particular instances of a larger group of CPTP transformations, called entanglement breaking (EB). As the name suggests, a channel Φ is EB if, when operating on half of a joint input state ρ_{QA} of Q and of an ancillary system A, produces output states

$(\Phi \otimes \mathrm{id})(\rho_{QA})$ that are separable (i.e., not entangled) (Gühne and Tóth, 2009; Horodecki *et al.*, 2009); see Fig. 6(c). These maps are closed under convex combination and channel concatenations, that is, given Φ_1 and Φ_2 EB, then $p\Phi_1 + (1-p)\Phi_2$ and $\Phi_1 \circ \Phi_2$ are also EB for all $p \in [0,1]$ (more generally concatenating an EB map with a generic CPTP map produces an EB channel). Necessary and sufficient conditions for being EB can be found in Horodecki, Shor, and Ruskai (2003) and Ruskai (2003) and, for the special case of infinite-dimensional systems, in Holevo (2008). In particular, Φ is EB if and only if its associated CJ state (9) is separable. Alternatively Φ is EB if and only if it is possible to identify a POVM $\{E_k\}_k$ on Q and collection of states $\{\rho_{Q'}^{(k)}\}_k$ in the output space Q' such that

$$\Phi(\rho_Q) = \sum_k \rho_{Q'}^{(k)} \mathrm{Tr}[E_k \rho_Q], \tag{18}$$

for all inputs ρ_Q [this last condition immediately shows that maps (15) and (16) are indeed EB].

EB channels form a proper subset of PPT channels. The latter are defined as those channels which produce output states $(\Phi \otimes \mathrm{id})(\rho_{QA})$ with PPT (Horodecki, Horodecki, and Horodecki, 1996; Peres, 1996; Rains, 2001). A necessary and sufficient condition for such a property is that the channel's CJ state (9) is PPT. Channels which are PPT but not EB are called entanglement binding maps. Generalizations of EB channels have been presented by De Pasquale and Giovannetti (2012) to describe those CPTP maps that become EB only after a certain number of concatenations, and by Moravčíková and Ziman (2010) and Filippov, Rybár, and Ziman (2012) to describe maps that when acting jointly on a composite system break entanglement among the subsystems that compose it.

G. Complementary channels and degradability

Associated with the Stinespring representation (7) is the notion of the complementary channel of Φ (see Fig. 7). The latter is the CPTP map $\tilde{\Phi} \in \mathfrak{P}(Q \mapsto E)$ which sends the initial states of the system Q into the states of the environment E through the transformation

$$\rho_Q \mapsto \tilde{\Phi}(\rho_Q) \coloneqq \mathrm{Tr}_Q[U_{QE}(\rho_Q \otimes |\omega\rangle_E\langle\omega|)U_{QE}^\dagger], \tag{19}$$

where Tr_Q denotes the partial trace over the system Hilbert space. The purity of the environmental state $\omega_E = |\omega\rangle_E\langle\omega|$

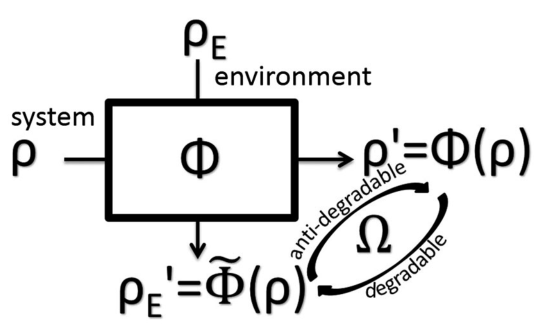


FIG. 7. Graphical representation of a channel Φ, its complementary one $\tilde{\Phi}$, and the degradability properties, in terms of the system and an external environment E.

ensures the uniqueness of $\tilde{\Phi}$ up to an isometric transformation on E. Channels (19) defined in terms of unitary dilations (7) with nonpure states ω_E are called weak complementaries of Φ and in general do not enjoy such symmetry (Caruso and Giovannetti, 2006; Caruso, Giovannetti, and Holevo, 2006).

The definition of $\tilde{\Phi}$ allows us to introduce another property of quantum channels, which is called degradability (Devetak and Shor, 2005). A map Φ is degradable when one can recover the final environment state $\tilde{\Phi}(\rho_Q)$ just by applying a third CPTP map to the output system state. More formally, a degradable map is such that there exists a CPTP map $\Omega \in \mathfrak{P}(Q \mapsto E)$ satisfying

$$\tilde{\Phi} = \Omega \circ \Phi. \tag{20}$$

Similarly, a channel is called *antidegradable* when the opposite relation holds, i.e.,

$$\Phi = \Omega \circ \tilde{\Phi}, \tag{21}$$

for some $\Omega \in \mathfrak{P}(E \mapsto Q)$, as shown in Fig. 7. Special examples of antidegrabable channels are the symmetric channels introduced by Smith, Smolin, and Winter (2008): these are CPTP maps for which Φ and $\tilde{\Phi}$ coincide (hence they are both degradable and antidegradable). Structural properties of degradable and antidegradable channels have been extensively analyzed by Cubitt, Ruskai, and Smith (2008), showing for instance that EB channels are always antidegrabable. Analogous definitions can be obtained for weak-complementary channels: in this case one says that Φ is weakly degradable if Eq. (20) holds—there is no need to define a weakly antidegradability condition as the latter can be shown to be equivalent to the antidegradabilty condition (Caruso, Giovannetti, and Holevo, 2006).

H. Causal, localizable, local operations and classical communication (LOCC), and separable channels

Additional structures arise when a quantum channel Φ acts on a multipartite system, e.g., a bipartite one $Q = Q_1Q_2$. It is useful to imagine that the two subsystems are associated with spatially or temporally separated laboratories where local CPTP maps can be applied and that can exchange classical or quantum information. In particular, bipartite channels can be characterized in terms of (1) how the output of one subsystem changes if a local transformation is applied to the input of the other subsystem; and (2) which resources (e.g., a preshared quantum state, local operations on the subsystems, classical, or quantum communication) are needed to simulate the bipartite channel.

The notions of causal and semicausal channels developed by Eggeling, Schlingemann, and Werner (2002) and Piani *et al.* (2006) provide a means of characterizing how the output of one subsystem depends on the input of the other. In this context a quantum channel $\Phi \in \mathfrak{P}(Q \mapsto Q)$ acting on a bipartite system $Q = Q_1Q_2$ is said to be $Q_1 \nrightarrow Q_2$ semicausal (Beckman *et al.*, 2001) if for any local CPTP map $\Psi \in \mathfrak{P}(Q_1 \mapsto Q_1)$ applied to Q_1 before the action of Φ, there is no detectable effect in the subsystem Q_2, i.e.,

$$\mathrm{Tr}_{Q_1}[\Phi(\rho_Q)] = \mathrm{Tr}_{Q_1}\{\Phi[(\Psi \otimes \mathrm{id})(\rho_Q)]\}, \tag{22}$$

where ρ_Q is a generic (possibly entangled) input state of the two carriers and where Tr_{Q_1} denotes the partial trace with respect to Q_1. In other words, for $Q_1 \nrightarrow Q_2$ semicausal maps cross talking from Q_1 to Q_2 is prevented. Similarly, one introduces the notion of $Q_2 \nrightarrow Q_1$ semicausal map. When both properties are satisfied, the map is called causal or nonsignaling. Special examples of nonsignaling channels are the tensor product channels $\Phi = \Phi_1 \otimes \Phi_2$ with $\Phi_{1,2}$ being CPTP maps operating locally on Q_1 and Q_2, respectively.

Another way of characterizing a bipartite quantum channel is in terms of the physical resources which are needed to simulate it. A bipartite channel is said to be localizable if it can be implemented by applying local CPTP maps on the subsystems with the assistance of a preshared bipartite quantum state (Beckman et al., 2001). Notice that the simulation of localizable channels does not require classical nor quantum communication between the two laboratories. Formally, this is the case when Φ can be represented as

$$\Phi(\rho_{Q_1Q_2}) = \mathrm{Tr}_{A_1A_2}[(\Psi \otimes \Omega)(\rho_{Q_1Q_2} \otimes \omega_{A_1A_2})], \tag{23}$$

where $\omega_{A_1A_2}$ is a shared bipartite state, and Ψ and Ω are quantum channels acting locally on subsystems Q_1A_1 and Q_2A_2, respectively; see Fig. 8. Otherwise, Φ is called $Q_1 \to Q_2$ semilocalizable if also one-way quantum communication from Q_1 to Q_2 is required to simulate the channel. Accordingly in this case Eq. (23) is replaced by

$$\Phi(\rho_{Q_1Q_2}) = \mathrm{Tr}_A[(\Psi \circ \Omega)(\rho_{Q_1Q_2} \otimes \omega_A)], \tag{24}$$

with ω_A being the state of an ancillary system A which is transmitted from one laboratory to the other and acts as the mediator between Q_1 and Q_2, and Ψ and Ω are quantum channels acting on the systems Q_2A and Q_1A, respectively.

By comparison of Eqs. (22) and (24) it follows that all $Q_1 \to Q_2$ semilocalizable maps are $Q_2 \nrightarrow Q_1$ semicausal, which in turn implies that all localizable maps are causal. Moreover, it can be proven that semicausality implies semilocalizability, hence semicausal and semilocalizable maps coincide, although causal and localizable maps do not (Beckman et al., 2001; Eggeling, Schlingemann, and Werner, 2002; Piani et al., 2006).

An important class of bipartite quantum channels is finally represented by the set of transformations which can be

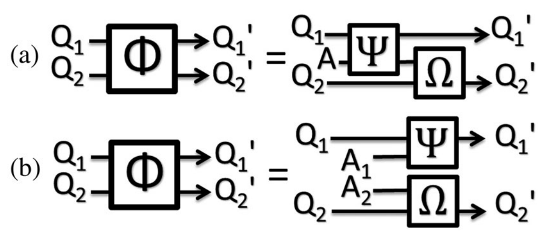


FIG. 8. Graphical structure of (a) semilocalizable and (b) localizable channels on bipartite systems Q_1Q_2, where the two parties act locally with the maps, respectively, Ψ and Ω on their subsystems and some auxiliary resources.

simulated only with LOCC [see Chitambar et al. (2012) for a recent survey]. These channels are hence termed LOCC channels. Classical communication is generally allowed in both directions between the two laboratories; the most general LOCC channel however can always be equivalently obtained by concatenating local CPTP maps and one-way classical communication. A LOCC transformation can hence be simulated by a finite number of iterations of the following sequence of operations: (1) on one of the two subsystems, say Q_1, a local CPTP map Ψ_1 is applied; (2) classical information is sent from Q_1 to Q_2, possibly conditioned on the local output of the map Ψ_1; (3) conditioned on the received classical information, a local CPTP map Ψ_2 is applied on subsystem Q_2; (4) the sequence of operations is repeated with the roles of Q_1 and Q_2 exchanged. A closely related class of bipartite channels is that of separable channels, defined as those channels admitting a Kraus representation in which all the Kraus operators are in the form of a direct product of operators acting on the local subsystems. It is easy to see that all the LOCC channels are separable. Interestingly enough, there exist separable channels which are not LOCC (Bennett, DiVincenzo et al., 1999).

I. Examples

Here some examples of quantum channels are presented.

1. Qubit channels

Qubit channels are the simplest, yet nontrivial, example of quantum channels: they are CPTP transformations $\Phi \in \mathfrak{P}(Q \mapsto Q)$ that map the states of a bidimensional quantum system (qubit) into states of the same system (in this case $\mathcal{H}_Q = \mathbb{C}^2$). A compact characterization of these channels can be obtained by adopting the Bloch ball representation, according to which any density operator ρ of the system is uniquely identified with the corresponding (Bloch) vector $\mathsf{r} = (\mathsf{r}_x, \mathsf{r}_y, \mathsf{r}_z) \in \mathbb{R}^3$ of length $|\mathsf{r}| \leq 1$ via the correspondence

$$\rho = \rho(\mathsf{r}) := \tfrac{1}{2}(\mathbb{1} + \mathsf{r} \cdot \mathsf{s}), \tag{25}$$

where $\mathsf{s} = (\sigma_x, \sigma_y, \sigma_z)^\top$ is a column vector formed by the Pauli matrices. In this framework any qubit channel Φ induces affine transformations of the form

$$\mathsf{r} \mapsto \mathsf{r}' = \mathsf{M}\mathsf{r} + \mathsf{t}, \tag{26}$$

with M and t being, respectively, a fixed 3×3 real matrix and a fixed three-dimensional real vector satisfying certain consistency requirements (King and Ruskai, 2001; Ruskai, Szarek, and Werner, 2002; Ruskai, 2003). In particular, qubit unital channels are obtained for $\mathsf{t} = 0$ and $\mathsf{M}^\top\mathsf{M} \leq \mathsf{I}$, the inequality being saturated if and only if Φ describes a unitary transformation [the latter case corresponds to having $\mathsf{M} \in \mathrm{SO}(3, \mathbb{R})$]. Exploiting this fact and the matrix singular value decomposition (Horn and Johnson, 1990) one can use the unitary equivalence of Eq. (5) to identify a canonical form for the qubit channel Φ, where the matrix M of Eq. (26) is written as $\mathsf{O'DO}$ with D being a real diagonal 3×3 matrix, and with $\mathsf{O'}$ and O being elements of $\mathrm{SO}(3,\mathbb{R})$.

An important class of qubit channels that have been extensively analyzed in the literature are those admitting a representation (8) with only two Kraus operators K_0 and K_1. In the canonical basis formed by the eigenvectors $\{|0\rangle,|1\rangle\}$ of the Pauli operator σ_z they can be parametrized as

$$K_0=\begin{pmatrix}\cos\theta & 0\\ 0 & \cos\phi\end{pmatrix},\qquad K_1=\begin{pmatrix}0 & \sin\phi\\ \sin\theta & 0\end{pmatrix},\tag{27}$$

with $\theta,\phi\in[0,\pi]$, up to unitary rotation. The corresponding affine mapping (26) is obtained with $\mathsf{M}=\mathrm{diag}(\cos(\phi-\theta),\cos(\phi+\theta),[\cos(2\theta)+\cos(2\phi)]/2)$ and $\mathsf{t}=(0,0,[\cos(2\theta)-\cos(2\phi)]/2)$. In the Stinespring representation (7) these maps describe situations in which the qubit system interacts with the smallest nontrivial environment (i.e., another qubit initialized in a pure state) and can be shown to be degradable for $\cos(2\theta)/\cos(2\phi)\geq 0$ and antidegradable otherwise (Giovannetti and Fazio, 2005; Caruso and Giovannetti, 2007; Wolf and Pérez-García, 2007). In particular, setting $\cos(2\theta)=1$ and $\cos(2\phi)=2\eta-1$, Eq. (27) defines the amplitude damping channel with damping rate η. For $\sin\theta=\pm\sin\phi$ instead one gets unital maps. Specifically for $\phi=\theta$, Eq. (27) describes the bit-flip channel that exchanges the states $|0\rangle$ and $|1\rangle$ with probability $p_x=\sin^2\phi$. By applying the unitary matrix (Hadamard transform)

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$$

to K_1,K_2 in Eq. (27) one recovers a unitarily equivalent channel called the phase-flip channel (or phase damping channel) that introduces a π shift between the states $|0\rangle$ and $|1\rangle$ with probability $p_z=\sin^2\phi$; see Fig. 9. Still Eq. (27) for $\phi=-\theta$ describes the bit-phase flip channel which with probability $p_y=\sin^2\phi$ exchanges the states $|0\rangle$ and $|1\rangle$ and also adds a relative π shift to them.

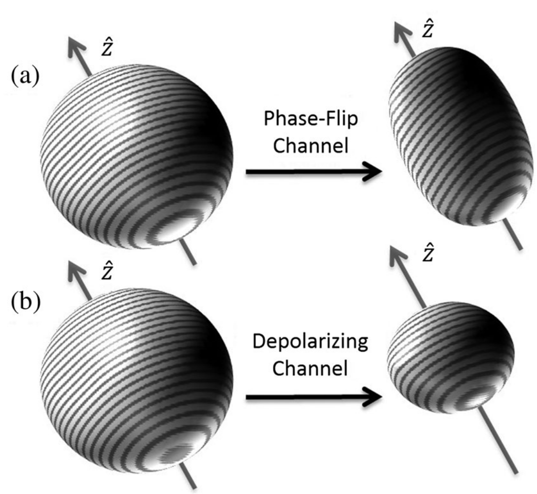


FIG. 9. Pictorial representation of the shrinking effects on the Bloch sphere via (a) phase-flip and (b) depolarizing channels.

Convex combinations of these three maps plus the identity channel define the class of Pauli channels: i.e.,

$$\Phi(\rho)=p_0\rho+p_1\sigma_x\rho\sigma_x+p_2\sigma_y\rho\sigma_y+p_3\sigma_z\rho\sigma_z,\tag{28}$$

with non-negative parameters $p_0+p_1+p_2+p_3=1$. Via the canonical representation detailed previously any other unital qubit map can be obtained from Eq. (28) through the concatenation Eq. (5).

2. Erasure channels

Erasure channels describe those communication scenarios in which errors are somehow heralded (i.e., the receiver Bob can determine whether or not something bad has happened to Alice's original message): accordingly they provide the simplest examples of CPTP maps operating among spaces of different dimensionality. Given a system Q described by the Hilbert space $\mathcal{H}_Q$, an erasure map is a stochastic transformation connecting $\mathfrak{S}(\mathcal{H}_Q)$ with $\mathfrak{S}(\mathcal{H}_{Q'})$, where $\mathcal{H}_{Q'}=\mathcal{H}_Q\oplus|e\rangle$ and $\oplus$ denotes the direct sum of the input Hilbert space with an extra "erasure" state $|e\rangle$ (the error "flag") which is orthogonal to each of the vectors of Q. In particular, as described by Bennett, DiVincenzo, and Smolin (1997) and Grassl, Beth, and Pellizzari (1997), the channel sends the input ρ_Q to itself with probability $1-p$ and to $|e\rangle$ with probability p.

3. Weyl covariant channels

Given a quantum system of finite dimension d (*qudit*) and the canonical basis $\{|e_k\rangle\}_{k=0,\ldots,d-1}$, consider the group $\mathbb{Z}_d\times\mathbb{Z}_d$ as a discrete phase space and take the unitary representation of such a group in the Hilbert space $\mathcal{H}$ of the system as

$$\mathsf{z}=\begin{pmatrix}x\\ y\end{pmatrix}\mapsto W_{\mathsf{z}}=U^xV^y,\tag{29}$$

where $x,y\in\mathbb{Z}_d$ and U, V are unitary operators on $\mathcal{H}$ generalizing the Pauli operators σ_x and σ_z in the following way (Gottesman, 1999):

$$U|e_k\rangle=|e_{k+1(\mathrm{mod}\,d)}\rangle,\qquad V|e_k\rangle=\exp\left(\frac{2\pi ik}{d}\right)|e_k\rangle.\tag{30}$$

The operators W_{z} are the discrete Weyl operators and they satisfy the canonical commutation relations

$$W_{\mathsf{z}}W_{\mathsf{z}'}=e^{(2\pi i/d)\mathsf{z}^\top\Sigma\mathsf{z}'}W_{\mathsf{z}'}W_{\mathsf{z}},\tag{31}$$

where the row vector $\mathsf{z}^\top$ is the transpose of the column vector z and where

$$\Sigma:=\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\tag{32}$$

is the matrix representation of the symplectic form with 0 (respectively, 1) the null (respectively, identity) $d\times d$ matrix.

The algebra $\mathcal{B}(\mathcal{H})$ of operators on $\mathcal{H}$ can be considered as a Hilbert space supplied by the Hilbert-Schmidt inner product.

There the Weyl operators form an orthogonal basis $\mathrm{Tr}(W_{\mathsf{z}}W_{\mathsf{z}'}^{\dagger}) = d\delta_{\mathsf{z}\mathsf{z}'}$, hence for all $O \in \mathcal{B}(\mathcal{H})$ one has

$$O = \sum_{\mathsf{z}} f_O(\mathsf{z}) W_{\mathsf{z}}, \qquad f_O(\mathsf{z}) = \frac{1}{d}\mathrm{Tr}(O W_{\mathsf{z}}^{\dagger}). \tag{33}$$

In this scenario a CPTP map Φ is said to be Weyl covariant (Fukuda and Holevo, 2005) when for all $\mathsf{z} \in \mathbb{Z}_d \times \mathbb{Z}_d$ it fulfills the identity

$$\Phi \circ \mathcal{W}_{\mathsf{z}} = \mathcal{W}_{\mathsf{z}} \circ \Phi, \tag{34}$$

with $\mathcal{W}_{\mathsf{z}}$ representing the quantum channel (2) associated with the unitary W_{z}. Applying Eq. (34) to the operator W_{z}, from Eq. (31) it follows that $\Phi(W_{\mathsf{z}})$ commutes with W_{z}. Hence, by means of Eq. (33), one can write

$$\Phi(W_{\mathsf{z}}) = \phi(\mathsf{z}) W_{\mathsf{z}}, \tag{35}$$

with $\phi(\mathsf{z})$ a complex-valued function, termed the "characteristic function of the channel" [without loss of generality it can be assumed $\phi(0) = 1$].

Special examples of Weyl covariant channels are provided by Weyl channels (Amosov, 2007) defined as those CPTP maps which admit Kraus decomposition in terms of random Weyl operators, i.e., $\Phi = \sum_{\mathsf{z}} p_{\mathsf{z}} \mathcal{W}_{\mathsf{z}}$, with p_{z} a probability distribution over $\mathbb{Z}_d \times \mathbb{Z}_d$. These are unital maps and their associated characteristic function is given by $\phi(\mathsf{z}) = \sum_{\mathsf{z}'} e^{(2\pi i/d)\mathsf{z}'\Sigma\mathsf{z}^T} p_{\mathsf{z}'}$. In particular, any d-depolarizing channel

$$\Phi(\rho) = \lambda\rho + (1-\lambda)\frac{1}{d}\mathbb{1}, \tag{36}$$

where $\lambda \in [0,1]$ is a Weyl channel having $\phi(\mathsf{z}) = \lambda$ for $\mathsf{z} \neq 0$. One notes also that since for $d = 2$ (Fig. 9) the Weyl operators reduce to the standard Pauli operators including identity, any unital qubit ($d = 2$) channel which is unitarily equivalent to the Pauli channels (28) is a Weyl channel. It is also possible to define the transpose d-depolarizing transformation,

$$\Phi(\rho) = \lambda\rho^{\top} + (1-\lambda)\frac{1}{d}\mathbb{1}, \tag{37}$$

which defines a CPTP map for

$$\lambda \in \left[1 - \frac{d}{d+1}, 1 - \frac{d}{d-1}\right]$$

(here $\rho^{\top}$ is the transpose of ρ with respect to a given basis) (Fannes et al., 2004).

4. Continuous variable quantum channels

Up to now mainly finite dimensional Hilbert spaces have been considered. This has been done to avoid technicalities related with the proper definition of the domains of the functionals. It is true however that the most common implementations of quantum communication lines are typically realized with CV systems (Braunstein and van Loock, 2005; Weedbrook et al., 2012) which at the quantum level are associated with an infinite-dimensional space—consider for instance the transferring of classical signals encoded into light pulses propagating along optical fibers or in free space (Caves and Drummond, 1994).

CV systems admit a description in terms of a discrete set of (say n) bosonic oscillators, typically a set of normal modes of the electromagnetic field, defined by ladder operators $a_1, a_1^{\dagger}, a_2, a_2^{\dagger}, \ldots, a_n, a_n^{\dagger}$ obeying canonical commutation relations $[a_k, a_{k'}^{\dagger}] = \delta_{kk'}$. Introducing the generalized "positions" and "momenta" coordinates $\{x_k, y_k\}_{k=1,\ldots,n}$, the density operators ρ of the CV system can be represented in terms of the (symmetrically ordered) characteristic functions $\chi(\mathsf{z}) = \mathrm{Tr}[\rho V(\mathsf{z})]$, where $\mathsf{z} \coloneqq (x_1, y_1, \ldots, x_n, y_n)^{\top}$ defines the vector of phase-space variables and where

$$V(\mathsf{z}) \coloneqq \exp\left[\frac{1}{\sqrt{2}}\sum_k (x_k + i y_k) a_k^{\dagger} - \mathrm{H.c.}\right]$$

denotes n-mode Weyl operators. The function $\chi(\mathsf{z})$ and the operator $V(\mathsf{z})$ represent the infinite-dimensional counterparts of $f_O(\mathsf{z})$ and W_{z} introduced in Sec. II.I.3. In particular, $V(\mathsf{z})$ fulfills commutation relations analogous to Eq. (31), i.e.,

$$V(\mathsf{z})V(\mathsf{z}') = e^{2\pi i \mathsf{z}^{\top}\Sigma^{(n)}\mathsf{z}'} V(\mathsf{z}')V(\mathsf{z}), \tag{38}$$

where now $\Sigma^{(n)}$ is the $2n \times 2n$ block matrix defined by

$$\Sigma^{(n)} \coloneqq \bigoplus_{k=1}^{n} \Sigma,$$

with Σ the single-mode phase-space canonical symplectic form deducible from Eq. (32) for $d = 1$.

Because of their physical relevance and relative simplicity of their mathematical description, a remarkable class of states is the class of Gaussian states (Eisert and Plenio, 2003; Ferraro, Olivares, and Paris, 2005; Holevo, 2011). They correspond to multimode (thermal) Gibbs states of Hamiltonians which are quadratic in the ladder operators of the system and are formally identified by the property of possessing Gaussian characteristic functions, i.e.,

$$\chi(\mathsf{z}) = \exp(i\mathsf{m}^{\top}\mathsf{z} - \tfrac{1}{2}\mathsf{z}^{\top}\mathsf{C}^{(n)}\mathsf{z}). \tag{39}$$

In Eq. (39) m is the vector of first moments

$$\mathsf{m}_k = \mathrm{Tr}(\mathsf{x}_k\rho), \tag{40}$$

where

$$\sqrt{2}\mathsf{x} \coloneqq ((a_1 + a_1^{\dagger}), -i(a_1 - a_1^{\dagger}), (a_2 + a_2^{\dagger}), -i(a_2 - a_2^{\dagger}), \ldots). \tag{41}$$

Furthermore, $\mathsf{C}^{(n)}$ is the covariance matrix (CM)

$$\mathsf{C}_{hk}^{(n)} = \tfrac{1}{2}\mathrm{Tr}(\mathsf{x}_h\mathsf{x}_k\rho) + \tfrac{1}{2}\mathrm{Tr}(\mathsf{x}_k\mathsf{x}_h\rho) - \mathrm{Tr}(\mathsf{x}_h\rho)\mathrm{Tr}(\mathsf{x}_k\rho), \tag{42}$$

obeying the generalized uncertainty relation (Simon, Mukunda, and Dutta, 1994)

$$\mathsf{C}^{(n)} - i\Sigma^{(n)}/2 \geq 0. \tag{43}$$

Concerning quantum channels in CV systems, attention has been mainly devoted to the study of Gaussian channels, i.e., CPTP maps that map Gaussian input states to Gaussian output states (Holevo and Werner, 2001; Weedbrook *et al.*, 2012). A part from attenuation and thermalization events arising from linear interactions with bosonic baths, they also include squeezing and linear amplification processes. When applied to a (not necessarily Gaussian) density operator ρ with characteristic function $\chi(\mathsf{z})$, a Gaussian channel Φ will transform it into an output density operator $\rho' = \Phi(\rho)$ having characteristic function

$$\chi'(\mathsf{z}) = \chi(\mathsf{X}^{(n)\top}\mathsf{z})f(\mathsf{z}), \tag{44}$$

where $\mathsf{X}^{(n)}$ is a matrix inducing a linear transformation on the $2n$-dimensional phase-space vector z, and the function $f(\mathsf{z})$ is Gaussian, i.e., $f(\mathsf{z}) = \exp(i\mathsf{d}^{(n)\top}\mathsf{z} - \frac{1}{2}\mathsf{z}^\top\mathsf{Y}^{(n)}\mathsf{z})$. The linear term proportional to the vector $\mathsf{d}^{(n)}$ accounts for a translation (displacement) of the mean m, while the quadratic term proportional to the matrix $\mathsf{Y}^{(n)}$ adds a term to the CM. CPTP conditions are ensured if and only if

$$\mathsf{Y}^{(n)} - i(\Sigma^{(n)} - \mathsf{X}^{(n)}\Sigma^{(n)}\mathsf{X}^{(n)\top})/2 \geq 0. \tag{45}$$

Gaussian transformations which are also unitary are characterized by the property that $\mathsf{X}^{(n)}$ is a symplectic matrix (i.e., $\mathsf{X}^{(n)}\Sigma^{(n)}\mathsf{X}^{(n)\top} = \Sigma^{(n)}$), and $\mathsf{Y}^{(n)} = 0$. An n-mode Gaussian channel is hence characterized by the triad $(\mathsf{d}^{(n)}, \mathsf{X}^{(n)}, \mathsf{Y}^{(n)})$ satisfying the constraint (45). The concatenation of two Gaussian channels with associated triads $(\mathsf{d}_1^{(n)}, \mathsf{X}_1^{(n)}, \mathsf{Y}_1^{(n)})$ and $(\mathsf{d}_2^{(n)}, \mathsf{X}_2^{(n)}, \mathsf{Y}_2^{(n)})$ is in turn characterized by the triad $(\mathsf{X}_2^{(n)}\mathsf{d}_1^{(n)} + \mathsf{d}_2^{(n)}, \mathsf{X}_2^{(n)}\mathsf{X}_1^{(n)}, \mathsf{X}_2^{(n)}\mathsf{Y}_1^{(n)}\mathsf{X}_2^{(n)\top} + \mathsf{Y}_2^{(n)})$. It follows that, by applying suitable Gaussian unitaries at the input and output of the channel, one can always reduce the channel to a canonical form, in which $\mathsf{d}^{(n)} = 0$, and the matrices $\mathsf{X}^{(n)}$, $\mathsf{Y}^{(n)}$ take a particular symmetric form. For the case of channels acting on one or two modes, the reduction to canonical forms allows one to classify Gaussian quantum channels according to invariance under unitary transformations (Serafini, Eisert, and Wolf, 2005; Caruso and Giovannetti, 2006; Holevo, 2007b; Caruso, Eisert *et al.*, 2008).

The basic processes of linear attenuation and amplification are modeled by single-mode Gaussian channels with $\mathsf{X}^{(1)} = \sqrt{\eta}$, $\mathsf{Y}^{(1)} = |1-\eta|/2$. For $\eta \leq 1$ these channels describe linear losses (with attenuation factor η), while for $\eta > 1$ they model the process of parametric amplification (with gain η). If extra Gaussian noise affects the attenuation or amplification process, one gets the noisy versions of the lossy and amplifier channel. In particular, the lossy and noisy Gaussian channel is defined by $\mathsf{X}^{(1)} = \sqrt{\eta}$ and $\mathsf{Y}^{(1)} = (1-\eta)(N_{\text{th}} + 1/2)$ ($\eta \in [0,1]$ and $N_{\text{th}} \geq 0$), and the additive noise Gaussian channel by $\mathsf{X}^{(1)} = 1$ and $\mathsf{Y}^{(1)} = N_{\text{add}}$ ($N_{\text{add}} \geq 0$). Notice that the additive noise can be obtained from the lossy and noisy channel by taking the limit of $\eta \to 1$ and $N_{\text{th}} \to \infty$ under the condition $(1-\eta)(N_{\text{th}} + 1/2) = N_{\text{add}}$.

J. Transfer fidelities and channel distances

In quantum information distance measures are of fundamental importance: by determining how far apart two states or two transformations are from each other, they are an essential guidance in the optimization of the data processing.

1. Input-output and entanglement fidelity of a quantum channel

A proper way to determine how much a system Q is modified by the action of a channel $\Phi \in \mathfrak{P}(Q \mapsto Q)$ can be obtained by considering the fidelity functional $F(\rho_1, \rho_2)$ (Uhlmann, 1976; Jozsa, 1994) (the definition relevant properties are recalled in Appendix A). Accordingly, for each input ρ_Q one defines the input-ouput (or transfer) fidelity associated with the map Φ as

$$F(\rho_Q; \Phi) \coloneqq F(\rho_Q, \Phi(\rho_Q)), \tag{46}$$

which for a pure state $|\psi\rangle_Q$ is linked to the error probability $P_e(|\psi\rangle_Q; \Phi)$ of not getting the right state at the channel output, via the identity

$$P_e(|\psi\rangle_Q; \Phi) = 1 - F(|\psi\rangle_Q; \Phi). \tag{47}$$

An overall estimate of the disturbance introduced by the channel can then be obtained by looking at how different from unity is the minimum or (alternatively) the average of $F(\rho_Q; \Phi)$ evaluated with respect to all possible pure input states of Q, i.e., the quantities

$$F_{\min}(\Phi) \coloneqq \min_{|\psi\rangle_Q} F(|\psi\rangle_Q; \Phi), \tag{48}$$

$$\bar{F}(\Phi) \coloneqq \int d\mu(\psi) F(|\psi\rangle_Q; \Phi), \tag{49}$$

the rationale being that $F_{\min}(\Phi) = 1$, as well as $\bar{F}(\Phi) = 1$, can occur if and only if Φ coincides with the identity channel id. The average in Eq. (49) is performed with respect to the Haar measure $d\mu(\psi)$ of the group whose action on a vector $|\psi\rangle$ is able to generate the entire space of states (Bengtsson and Zyczkowski, 2006); the minimization in Eq. (48) instead can be generalized to also include mixed states by exploiting the concavity property of the fidelity (Nielsen and Chuang, 2000; Wilde, 2013), i.e., $F_{\min}(\Phi) = \min_{\rho_Q} F(\rho_Q; \Phi)$.

To gauge the disturbance of the channel Φ, one may also consider its entanglement fidelities (Schumacher, 1996), defined as the input-output fidelities of the extended map $\Phi \otimes \text{id}$ when operating on purifications of the density matrices ρ_Q. Recall that a purification of a density matrix $\rho_Q \in \mathfrak{S}(\mathcal{H}_Q)$ is any pure state $|\psi_\rho\rangle_{QR} \in \mathcal{H}_Q \otimes \mathcal{H}_R$ of the enlarged system formed by Q and by an ancillary system R, which fulfills the property $\rho_Q = \text{Tr}_R(|\psi_\rho\rangle_{QR}\langle\psi_\rho|)$ (Gühne and Tóth, 2009; Horodecki *et al.*, 2009). The entanglement fidelity is then written as

$$F_e(\rho_Q; \Phi) \coloneqq F(|\psi_\rho\rangle_{QR}; \Phi \otimes \text{id}), \tag{50}$$

where id is the identity map on R. It is important to stress that $F_e(\rho_Q; \Phi)$ is independent of the way $|\psi_\rho\rangle_{QR}$ is constructed and of the choice of the ancillary system: as a matter of fact, given $\{K_j\}_j$ a set of Kraus operators of Φ, it can be expressed as

$$F_e(\rho_Q;\Phi) = \sum_j |\mathrm{Tr}(\rho_Q K_j)|^2. \tag{51}$$

Operationally the entanglement fidelity functional (50) can be used to detect the detrimental effects on the transmission of half of the entangled state $|\psi\rangle_{QR}$ through the channel Φ. This quantity is related to the input-output fidelity (46) via the inequality

$$F_e(\rho_Q;\Phi) \le F(\rho_Q;\Phi), \tag{52}$$

implying that values of $F_e(\rho_Q;\Phi)$ close to 1 force $F(\rho_Q;\Phi)$ to approach unity too. Slightly weaker versions of the opposite implication can also be proven (Kretschmann and Werner, 2004; Holevo, 2012). In particular, given $\epsilon > 0$, if $F(|\psi\rangle_Q;\Phi) \ge 1-\epsilon$ for all input states $|\psi\rangle_Q$ belonging to the support of the density matrix ρ_Q, then (Barnum, Knill, and Nielsen, 2000)

$$F_e(\rho_Q;\Phi) \ge 1 - 3\epsilon/2. \tag{53}$$

Furthermore, taking ρ_Q to be the completely mixed state of Q, i.e., the density matrix $\mathbb{1}/d$ (d being the dimension of $\mathcal{H}_Q$) whose purification is a maximally entangled state of QR, from Eq. (51) it follows that

$$F_e(\mathbb{1}/d,\Phi) = \frac{1}{d^2}\sum_j |\mathrm{Tr}(K_j)|^2, \tag{54}$$

which can be put in correspondence with the average fidelity (49) through the identity (Horodecki, Horodecki, and Horodecki, 1999; Nielsen, 2002)

$$\bar{F}(\Phi) = \frac{dF_e(\mathbb{1}/d;\Phi)+1}{d+1}. \tag{55}$$

2. Distance measures for channels

Distance measures for quantum channels (and in general for quantum operations) are typically written as $|||\Phi - \Psi|||$, where $|||\Lambda|||$ denotes a proper norm of the superoperator Λ. Suitable choices are

$$|||\Lambda|||_k \coloneqq \sup_{\|O\|_k \le 1} \|\Lambda(O)\|_k, \tag{56}$$

where the index k identifies a norm for the operators of the system. Specifically for $k=1$, $\|O\|_1 = \mathrm{Tr}\sqrt{O^\dagger O}$ is the trace norm; for $k=2$, $\|O\|_2 = \sqrt{\mathrm{Tr} O^\dagger O}$ is the Hilbert-Schmidt norm; and finally for $k=\infty$, $\|O\|_\infty = \sup_{|\psi\rangle} \langle\psi|O|\psi\rangle$ is the standard operator norm—recall that they obey the following ordering $\|O\|_\infty \le \|O\|_2 \le \|O\|_1$ (Horn and Johnson, 1990). While correctly defined, when applied to CPTP maps, the norms (56) become unstable under channel extension. In particular, $|||\Lambda\otimes \mathrm{id}|||_k$ can explicitly depend upon the dimensionality of the ancillary system for which the id channel is defined. In order to amend this, regularizations have been proposed. Of particular relevance is the norm of complete boundedness, or cb norm (Paulsen, 2003), and the diamond norm (Kitaev, 1997). Given a generic (not necessarily CPTP) map $\Lambda\colon \mathfrak{S}(\mathbb{C}^n) \mapsto \mathfrak{S}(\mathbb{C}^k)$ they are defined, respectively, as

$$|||\Lambda|||_{cb} \coloneqq \sup_m |||\Lambda\otimes \mathrm{id}_m|||_\infty, \tag{57}$$

$$|||\Lambda|||_\diamond \coloneqq |||\Lambda\otimes \mathrm{id}_n|||_1, \tag{58}$$

where id_m denotes the identity channel on $\mathfrak{S}(\mathbb{C}^m)$. While not obvious at least in the case of $|||\cdots|||_\diamond$, both these norms are stable under channel extension. Furthermore, they are related through the identity $|||\Lambda|||_{cb} = |||\Lambda^*|||_\diamond$, where Λ^* is the dual of Λ (Johnston, Kribs, and Paulsen, 2009).

In the context of quantum communication, the properties of the cb norm have been extensively reviewed by Holevo and Werner (2001), Keyl (2002), Kretschmann (2003), Kretschmann and Werner (2004, 2005), Belavkin, D'Ariano, and Raginsky (2005), and Johnston, Kribs, and Paulsen (2009). Recall that it is well behaved under tensor product composition rule (6), since it has the property

$$|||\Lambda_1\otimes\Lambda_2|||_{cb} = |||\Lambda_1|||_{cb}|||\Lambda_2|||_{cb}. \tag{59}$$

Furthermore, if Λ is completely positive then $|||\Lambda|||_{cb} = \|\Lambda(\mathbb{1})\|_\infty$. Accordingly if Λ is CPTP and Λ^* its dual channel (10), one has that $|||\Lambda|||_{cb}$ can take any value up to d (the dimension of the channel input space) while $|||\Lambda^*|||_{cb} = 1$ always.

Finally, another useful distance measure for quantum channels is the one introduced by Grace *et al.* (2010) as a distance between unitary operations acting on a bipartite quantum system, where only the effect of the operations on one component (the subsystem of interest) is relevant in the measure, while the effect on the other component (environment) can be arbitrary.

K. Channels and entropies

In the study of quantum communication, entropic quantities play a fundamental role in characterizing quantum channels in terms of their efficiency as communication lines (Barnum, Nielsen, and Schumacher, 1998). A comprehensive characterization of these functionals can be obtained moving into the so-called "church of the larger Hilbert space," a construction based on the Stinespring dilation form (7), where the input state ρ_Q of the system Q is also represented as a reduced density operator of a pure state $|\psi_\rho\rangle_{QR}$ of a larger system QR via a purification; see Fig. 10. We denote by

$$S(\rho) \coloneqq -\mathrm{Tr}(\rho\log_2\rho) \tag{60}$$

the von Neumann entropy of the density operator ρ (Wehrl, 1978; Ohya and Petz, 1993; Petz, 2008) which generalizes to quantum mechanical systems the Shannon entropy of a classical random variable X taking values in the alphabet $\mathcal{X}$, defined as

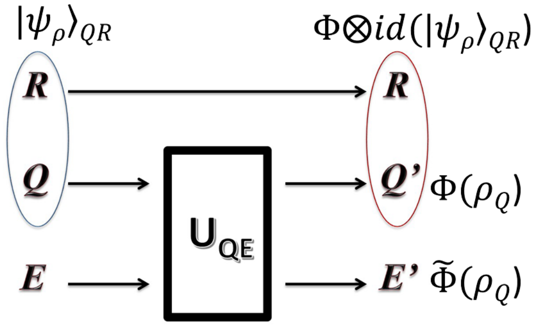


FIG. 10 (color online). Graphical representation of a quantum channel $\Phi \in \mathfrak{P}(Q \mapsto Q')$ as the unitary interaction U_{QE} between the system state ρ_Q and the environmental one ρ_E. The action of $\Phi \otimes \mathrm{id}$ on the purification of $|\psi_\rho\rangle_{QR}$ and the complementary map $\tilde{\Phi}$ are also shown.

$$H(X) := -\sum_{x\in\mathcal{X}} p(x) \log_2 p(x), \tag{61}$$

with $p(x)$ being the probability that X acquires the value x (Gallager, 1968; Cover and Thomas, 1991). Then given a quantum channel $\Phi \in \mathfrak{P}(Q \mapsto Q')$ and an input state ρ_Q, there are three important entropic quantities related to the pair (ρ_Q, Φ). First is the entropy of the input state $S[Q] := S(\rho_Q)$ (input entropy), which exploiting the fact that the purification $|\psi_\rho\rangle_{QR}$ is pure can also be expressed as the entropy of the ancillary system R, i.e., $S[Q] = S[R]$. Second is the entropy of the output state $\Phi(\rho_Q)$, i.e., $S[Q'] := S[\Phi(\rho)]$ (output entropy). Finally there is the entropy of exchange (Schumacher, 1996; Barnum, Nielsen, and Schumacher, 1998) computed as the von Neumann entropy of the environment E after the interaction with Q, i.e., the entropy measured at the output of the complementary channel $\tilde{\Phi}$ defined in Eq. (19),

$$S[E'] = S(\rho_Q; \Phi) := S(\tilde{\Phi}(\rho_Q)) = S[(\Phi \otimes \mathrm{id})(|\psi_\rho\rangle_{QR}\langle\psi_\rho|)] \tag{62}$$

(the last identity follows from the fact that the global state of Q, R, and E is always pure). A complete analysis of the relations between these three quantities was reviewed by Schumacher and Nielsen (1996) and Holevo and Werner (2001). In particular, they satisfy

$$S[\Phi(\rho_Q)] + S(\rho_Q; \Phi) \geq S(\rho_Q), \tag{63}$$

$$|S[\Phi(\rho_Q)] - S(\rho_Q; \Phi)| \leq S(\rho_Q), \tag{64}$$

and the quantum Fano inequality

$$S(\rho_Q; \Phi) \leq h[F_e(\rho_Q; \Phi)] + [1 - F_e(\rho_Q; \Phi)] \log_2(d^2 - 1), \tag{65}$$

with d the dimension of the channel input,

$$h(p) := -p\log_2 p - (1-p)\log_2(1-p), \tag{66}$$

the Shannon binary entropy function (Gallager, 1968; Cover and Thomas, 1991), and $F_e(\rho_Q; \Phi)$ the entanglement fidelity introduced in Eq. (50).

The input, output, and exchange entropies are the building blocks for constructing several information quantities. For instance, one defines the quantum mutual information $I(\rho_Q; \Phi)$ between the system Q at the output of the map Φ and the system R which enters in the purification $|\psi_\rho\rangle_{QR}$, i.e.,

$$I(\rho_Q; \Phi) := S(\rho_Q) + S[\Phi(\rho_Q)] - S(\rho_Q; \Phi). \tag{67}$$

This is a non-negative quantity which is known to be concave and subadditive with respect to density operators on Q (Adami and Cerf, 1997; Bennett *et al.*, 2002).

Subtracting $S(\rho_Q)$ from $I(\rho_Q; \Phi)$ one also defines the channel coherent information (Schumacher and Nielsen, 1996; Barnum, Nielsen, and Schumacher, 1998),

$$J(\rho_Q; \Phi) := S[\Phi(\rho)] - S(\rho_Q; \Phi) = S[\Phi(\rho)] - S[\tilde{\Phi}(\rho_Q)], \tag{68}$$

where in the last expression it is noted that $J(\rho_Q; \Phi)$ can also be expressed as the difference between the output entropy of Φ and of its complementary counterpart $\tilde{\Phi}$. Unlike $I(\rho_Q; \Phi)$, the function $J(\rho_Q; \Phi)$ is in general neither non-negative, nor convex or subadditive (DiVincenzo, Shor, and Smolin, 1998; Smith and Smolin, 2007). However, both the quantum mutual and the coherent information satisfy data-processing inequalities. In particular, given Φ_1 and Φ_2 CPTP channels, one has (Holevo, 2012)

$$I(\rho_Q; \Phi_2 \circ \Phi_1) \leq \min\{I_Q(\rho_Q; \Phi_1), I(\Phi_1(\rho_Q); \Phi_2)\}, \tag{69}$$

while

$$J(\rho_Q; \Phi_2 \circ \Phi_1) \leq J(\rho_Q; \Phi_1). \tag{70}$$

A further entropic quantity useful for characterizing the channel Φ is the channel Holevo information. In contrast to the previous expressions this is a functional of an input ensemble $\mathcal{E} := \{p_j; \rho_Q^{(j)}\}_j$ Alice feeds into the channel (here $\{p_j\}_j$ is a probability distribution while $\{\rho_Q^{(j)}\}_j$ is a collection of input states). Accordingly one has

$$\begin{aligned}\chi(\mathcal{E}; \Phi) &:= S[\Phi(\rho_Q)] - \sum_j p_j S[\Phi(\rho_Q^{(j)})] \\ &= \sum_j p_j S(\Phi(\rho_Q) \| \Phi(\rho_Q^{(j)})),\end{aligned} \tag{71}$$

where $\rho_Q = \sum_j p_j \rho_Q^{(j)}$ is the average state associated with $\mathcal{E}$ and in the last identity the quantum relative entropy $S(\rho_1 \| \rho_2) := \mathrm{Tr}[\rho_1(\log \rho_1 - \log \rho_2)]$ (Lindblad, 1975; Schumacher and Westmoreland, 2000) has been used. Via the Holevo bound (Holevo, 1973a, 1973b) the quantity $\chi(\mathcal{E}; \Phi)$ provides an upper bound on the information one could retrieve on the random variable X associated with index j of the ensemble $\mathcal{E}$ if allowed to measure the corresponding states at the output of the channel Φ. Specifically, indicating with Y the random variables associated with the estimation of j after a POVM has been performed on the density matrix $\Phi(\rho_Q^{(j)})$, one has

$$I(X:Y) \le \chi(\mathcal{E};\Phi), \tag{72}$$

with $I(X:Y) := H(Y) + H(X) - H(X,Y)$ being the (Shannon) mutual information associated with the couple X and Y (Gallager, 1968; Cover and Thomas, 1991). The functional $\chi(\mathcal{E};\Phi)$ obeys the data-processing inequality

$$\chi(\mathcal{E};\Phi_2 \circ \Phi_1) \le \chi(\mathcal{E};\Phi_1), \tag{73}$$

for all Φ_1, Φ_2 CPTP maps and for all ensemble $\mathcal{E}$. Another quantity related to the Holevo information is the minimum output entropy of the channel $S_{\min}(\Phi) = \min_\rho S[\Phi(\rho)]$ that quantifies the minimum disturbance induced by the channel. A connection with the coherent information can be established via the identity (Devetak, 2005; Holevo, 2012)

$$\chi(\mathcal{E};\Phi) - \chi(\mathcal{E};\tilde{\Phi}) = J(\rho_Q;\Phi) - \sum_j p_j J(\rho_Q^{(j)};\Phi), \tag{74}$$

with ρ_Q being the average density matrix of the ensemble $\mathcal{E} = \{p_j;\rho_Q^{(j)}\}_j$.

III. FROM MEMORYLESS TO MEMORY QUANTUM CHANNELS

Having in mind the multiuses communication scenario detailed at the beginning of Sec. II, in which a time-ordered sequence of carriers $Q := \{q_1, q_2, \ldots\}$ propagates from Alice to Bob along a noisy channel, we start this section by discussing the simplest case, where they are affected by uncorrelated identical maps and then move on to consider correlations among uses, i.e., memory effects.

A. Memoryless quantum channels

Memoryless quantum channels describe those scenarios in which the noise acts identically and independently on each element of the sequence Q. Under this assumption the multiuse map associated with the communication line is expressed as a tensor product of a CPTP map $\Phi\colon \mathfrak{S}(\mathcal{H}_q) \mapsto \mathfrak{S}(\mathcal{H}_q)$ that acts on the states of a single carrier q. Therefore, indicating as $\mathcal{H}_Q^{(n)} := \mathcal{H}_{q_1} \otimes \cdots \otimes \mathcal{H}_{q_n}$ the Hilbert space of the first n carriers of the system, its input density operators $\rho_Q^{(n)} \in \mathfrak{S}(\mathcal{H}_Q^{(n)})$ will be mapped into

$$\Phi^{(n)}(\rho_Q^{(n)}) = \Phi^{\otimes n}(\rho_Q^{(n)}), \tag{75}$$

with $\Phi^{\otimes n} := \Phi \otimes \cdots \otimes \Phi$. Equivalently, one can say that the Kraus operators of the memoryless map $\Phi^{(n)}$ can be expressed as a tensor product $K_{i_1} \otimes \cdots \otimes K_{i_n}$ formed by independent and identically distributed sequences extracted from the Kraus set $\{K_i\}_i$ associated with the single carrier channel Φ. A simplified, yet informative, model can be found in Giovannetti (2005). Here the carriers Q are assumed to propagate from Alice to Bob, one by one and at constant speed, while interacting with an external environmental system via a constant coupling described by the unitary operator $U_{qe} \in \mathcal{B}(\mathcal{H}_q \otimes \mathcal{H}_e)$ whose role is to effectively simulate the interaction between the carriers and the medium which separate the two communicating parties. In the model the environment e is assumed to undergo a dissipative process which on a time scale τ tends to reset it into a stable configuration ω_e (see Fig. 11 for a pictorial representation of the scheme). The memoryless regime is achieved in the limit in which the rate ν, at which the carriers propagate from Alice to Bob, is much lower than the inverse of the relaxation time τ, i.e., $\nu \ll 1/\tau$. In this limit in fact each carrier couples with identical and independent environmental states. Defining then $\omega_E^{\otimes n} := \omega_{e_1} \otimes \cdots \otimes \omega_{e_n}$, this allows one to write

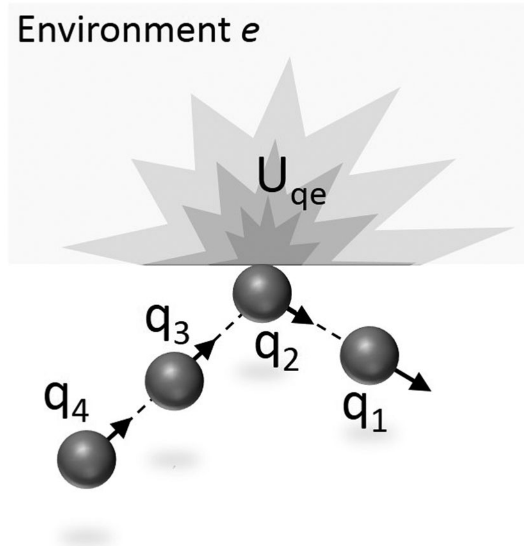


FIG. 11. Pictorial representation of the model of Giovannetti (2005). The different elements represent the sequence of carriers that propagates at a rate ν from Alice to Bob, interacting on the way, with the environment e via the unitary coupling U_{qe}. Among two consecutive interactions the environment tends to relax toward its stable configuration ω_e via a dissipative process characterized by the relaxing time τ. The memoryless channel configuration is achieved when $\nu \ll 1/\tau$.

$$\begin{aligned}\rho_Q^{(n)} \mapsto \mathrm{Tr}_E[&U_{q_n,e_n} \otimes \cdots \otimes U_{q_1,e_1} \\ &(\rho_Q^{(n)} \otimes \omega_E^{\otimes n}) U_{q_1,e_1}^\dagger \otimes \cdots \otimes U_{q_n,e_n}^\dagger],\end{aligned} \tag{76}$$

which reduces to Eq. (75) when identifying $\mathrm{Tr}_e[U_{qe}(\cdots \otimes \omega_e)U_{qe}^\dagger]$ with unitary dilation of the single-use channel Φ.

1. Compound and averaged quantum channels

Before entering into the subject of memory quantum channels, we briefly discuss the situation in which the channel map, although intended as acting like Eq. (75), is not perfectly known to the sender and receiver. Such a situation can be modeled by considering not a single CPTP map, but rather a set $\{\Phi_i\}_i$ of them. Here $\Phi_i\colon \mathfrak{S}(\mathcal{H}_q) \mapsto \mathfrak{S}(\mathcal{H}_q)$ and the set $\{\Phi_i\}_i$ can in principle contain a finite or infinite (countable or not) number of CPTP maps. This leads to the notion of a memoryless compound quantum channel, i.e., the family $\{\Phi_i^{\otimes n}\colon \mathfrak{S}(\mathcal{H}_q^{\otimes n}) \to \mathfrak{S}(\mathcal{H}_q^{\otimes n})\}_{n,i}$. Averaged channels are closely related to compound channels. The difference is that in the former the sender and receiver know an *a priori* probability distribution $\{p_i\}_i$ governing the appearance of the members of a compound channel. It means that for any $n \in \mathbb{N}$ one can write the averaged channel map as

$$\Phi^{(n)}(\rho_Q^{(n)}) = \sum_i p_i \Phi_i^{\otimes n}(\rho_Q^{(n)}). \quad (77)$$

Equation (77) describes a scenario in which, with some probability p_i, all the carriers of the system are operated on by the same identical local transformation Φ_i. The index i can be interpreted as a "switch" selecting different memoryless channels, and Eq. (77) as the average channel over different values of the switch. Classical counterparts of compound and averaged channels were studied a long time ago (Blackwell, Breiman, and Thomasian, 1959; Wolfowitz, 1960; Jacobs, 1962; Ahlswede, 1968). Compound and averaged quantum channels were introduced only recently (Hayashi, 2008; Bjelaković and Boche, 2009). In Sec. III.D.8 one will see that these channels are closely related to a special set of memory channels having long-term memory.

B. Nonanticipatory memory quantum channels

Whenever the tensorial decomposition of Eq. (75) does not apply, one can speak of memory channels or correlated noise channels. Among the plethora of possibilities, the following focuses only on those configurations that have physical relevance and have attracted some interest in the recent literature. In particular, one can treat those models in which the noise respects the time ordering of the carriers Q so that at a given channel use, the output cannot be influenced by successive inputs as pictorially shown in the right panel of Fig. 2. This property generalizes the notion of semicausality discussed in Sec. II.H to the case of multiple (ordered) subsystems. Inspired by the classical theory of communication (Gallager, 1968) one can name the quantum communication lines which fulfills such condition, nonanticipatory quantum channels [note, however, that in the approach of Kretschmann and Werner (2005) these maps are called just causal—see Sec. III.C].

Under the nonanticipatory condition there must exist a family of CPTP maps $\mathcal{F} := \{\Phi^{(n)}; n = 1, 2, \ldots\}$ with $\Phi^{(n)}: \mathfrak{S}(\mathcal{H}_Q^{(n)}) \to \mathfrak{S}(\mathcal{H}_Q^{(n)})$ which allows one to express the output states of the first n carriers in terms of the density matrices of their associated inputs, i.e.,

$$\rho_Q^{(n)} \mapsto \Phi^{(n)}(\rho_Q^{(n)}). \quad (78)$$

Clearly the property (78) requires that the family $\mathcal{F}$ must fulfill the minimal consistency requirement that for all $m < n$ the element $\Phi^{(m)}$ should be obtained as a restriction of $\Phi^{(n)}$ over the degrees of freedom of the first m carriers. That is, given $\rho_Q^{(n)} \in \mathfrak{S}(\mathcal{H}_Q^{(n)})$ and $\rho_Q^{(m)} \in \mathfrak{S}(\mathcal{H}_Q^{(m)})$, one must have

$$\Phi^{(m)}(\rho_Q^{(m)}) = \mathrm{Tr}^{(m)}[\Phi^{(n)}(\rho_Q^{(n)})], \quad (79)$$

whenever $\rho_Q^{(m)} = \mathrm{Tr}^{(m)}[\rho_Q^{(n)}]$, where $\mathrm{Tr}^{(m)}$ stands for the partial trace over all the carriers but the first m.

As noted, in the language introduced in Sec. II.H, nonanticipatory channels can be classified as semicausal with respect to the natural ordering of the channel uses. The representation of semicausal channels given in Eq. (24) can hence be applied, yielding a representation of nonanticipatory quantum channels in which each carrier couples sequentially with a common memory system M. The backaction of M on the message state simulates the memory effects of the transmission. Accordingly, all the nonanticipatory CPTP maps can be expressed as

$$\Phi^{(n)}(\rho_Q^{(n)}) = \mathrm{Tr}_M[U_{q_n M} \ldots U_{q_1 M}(\rho_Q^{(n)} \otimes \omega_M) U_{q_1 M}^\dagger \ldots U_{q_n M}^\dagger], \quad (80)$$

where for all $j = 1, 2, \ldots, n$, $U_{q_j M}$ is a unitary transformation which describes the coupling of the jth carrier with the memory system M, and where ω_M is some given state of M; see Fig. 12(a). The unitary transformations $U_{q_j M}$ may in general depend on the carrier label j. Otherwise, if they are independent of j the memory channel has the additional property of being invariant under translation of the carrier labels. An explicit proof of Eq. (80) was first given by Kretschmann and Werner (2005) in the context of quasilocal algebras (see also Appendix B), under the assumption of translational invariance of the noise (see Sec. III.C). An alternative proof which does not make use of this hypothesis can be found in Appendix C.

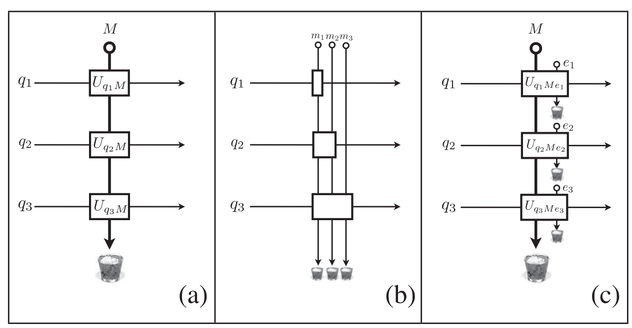


FIG. 12. Unitary dilations for nonanticipatory quantum memory channels. (a) A graphical sketch of the representations of Eq. (80): here the noise correlations among the n channel uses can be described via a series of concatenated unitary interactions with a common reservoir M whose dimension in general depends (exponentially) upon n ($n = 3$ in the example). Notice that while the carrier q_1 might influence the outcome of $q_2 q_3$ via their common interaction with M, $q_2 q_3$ cannot influence the output of the first carrier. (b) The environment M can also be represented as a collection of smaller systems $M_1, M_2, \ldots$ initially prepared into a separable state while, as shown in Eq. (81), the unitary transformation operating on the jth channel use couples it with the first j subsystems only. (c) Unitary dilation (82) where besides M a series of local environment $e_1, e_2, \ldots$ are also present. In all the diagrams the unitary operators (represented by the white boxes) are applied sequentially on the input states of the global system (i.e., the carriers and the environment) starting for the one on the top of the figure. The carriers and the environmental states evolve, respectively, from left to right and from top to bottom while interacting, meeting at a white box. The trash-bin symbol stands for the partial trace operation on the environment.

In Eq. (80) M is in general a large system whose dimension d_M is an explicit function of n (in any case it can always be chosen to be less than or equal to d^{2n} with d being the dimension of a single carrier). In fact, as explained in Appendix C, one can take M to be a composite system of components $m_1, m_2, \ldots, m_n$ whose dimensions can always be chosen to not be larger than d^2. In this configuration then one can assume ω_M to be a pure tensor product state of local terms $|0\rangle_{m_1} \otimes \cdots \otimes |0\rangle_{m_n}$, and write U_{q_jM} as a transformation which couples the jth carrier only with the first j elements of M, i.e.,

$$U_{q_jM} = \mathbb{1}_{m_n} \otimes \cdots \otimes \mathbb{1}_{m_{j+1}} \otimes U_{q_jm_jm_{j-1}\ldots m_1}, \tag{81}$$

with $\mathbb{1}_{m'}$ being the identity operator on the m' components of the environment; see Fig. 12(b).

An alternative, but fully equivalent, representation for nonanticipatory channels is obtained by adding to Eq. (80) a collection of local environments which individually couples with the carriers, i.e.,

$$\Phi^{(n)}(\rho_Q^{(n)}) = \mathrm{Tr}_{ME}[U_{q_nMe_n} \cdots U_{q_1Me_1} \times (\rho_Q^{(n)} \otimes \omega_M \otimes \omega_E^{\otimes n})U_{q_1Me_1}^{\dagger} \cdots U_{q_nMe_n}^{\dagger}], \tag{82}$$

where for all $j = 1, 2, \ldots, n$, $U_{q_jMe_j}$ is now the unitary transformation which describes the coupling of the jth carrier with its own local environment e_j and with the memory system M, where $\omega_E^{\otimes n} \coloneqq \omega_{e_1} \otimes \cdots \otimes \omega_{e_n}$ as in the memoryless case, and ω_M is some given state of M; see Fig. 12(c). In principle one can distinguish different setups in which Alice, Bob, or Eve (third party) has control of the initial or final states of the memory system M (Kretschmann and Werner, 2005). Equation (82) was first introduced by Bowen and Mancini (2004) as a model for representing correlated channels: from Eq. (80) it follows that it provides a general unitary dilation for every nonanticipatory quantum map. It can also be expressed in terms of an n-fold concatenation of a sequence of CPTP maps acting on a single carrier and the memory system M (Bowen and Mancini, 2004; Kretschmann and Werner, 2005). Such concatenation is shown pictorially in Fig. 12(c) and results in the following identity:

$$\Phi^{(n)}(\rho_Q^{(n)}) = \mathrm{Tr}_M[\Phi_{QM}^{(n)}(\rho_Q^{(n)} \otimes \omega_M)], \tag{83}$$

with

$$\Phi_{QM}^{(n)} \coloneqq \Phi_{q_nM} \circ \Phi_{QM}^{(n-1)} = \Phi_{q_nM} \circ \cdots \circ \Phi_{q_1M}, \tag{84}$$

where for $j = 1, 2, \ldots, n$, $\Phi_{q_jM}\colon \mathfrak{S}(\mathcal{H}_{q_j} \otimes \mathcal{H}_M) \to \mathfrak{S}(\mathcal{H}_{q_j} \otimes \mathcal{H}_M)$ is a CPTP map that operates on the jth carrier and on the memory ancilla M and is defined by the unitary dilation

$$\Phi_{q_jM}(\cdots) = \mathrm{Tr}_{e_j}[U_{q_jMe_j}(\cdots \otimes \omega_{e_j})U_{q_jMe_j}^{\dagger}]. \tag{85}$$

In this representation the evolution of M after the interaction with the carriers is provided by the transformation

$$\omega_M \mapsto \Psi^{(n)}(\rho_Q^{(n)}; \omega_M) \coloneqq \mathrm{Tr}_Q[\Phi_{QM}^{(n)}(\rho_Q^{(n)} \otimes \omega_M)], \tag{86}$$

which explicitly depends upon the input state of Q.

Cases of special interest (Kretschmann and Werner, 2005) are those in which, for all j, the Φ_{q_jM} describes the same mapping $\Phi = \Phi_{qM}$ on $\mathfrak{S}(\mathcal{H}_q \otimes \mathcal{H}_M)$ which according to Eq. (84) becomes the *generator* of the n-fold concatenation. That characterizes memory channels which are nonanticipatory and translation invariant (i.e., invariant under translation of the information carriers $q_j \to q_{j+1}$). Memoryless channels can then be included in this class as a limiting case in which the generator Φ can be expressed as a tensor product channel that acts independently on the carrier q and on the memory system M. In terms of the unitary dilation (82) this is equivalent to assuming that the unitaries $U_{q_jMe_j}$ in Eq. (76) factorize in a tensor product $U_{q_je_j} \otimes V_M$, where V_M is a unitary operator on the memory system and $U_{q_je_j}$ acts only on the degree of freedom of the jth carrier and on its local environment e_j.

A special subset of nonanticipatory channels is formed by symbol independent (SI) maps (Bowen, Devetak, and Mancini, 2005). They are communication lines where previous input states do not affect the action of the channel on the current input state. In other words, the symbol independent maps are nonanticipatory (or semicausal) with respect to all possible ordering of the carriers (in this sense they are hence fully nonanticipatory). Accordingly, given a generic subset of the carrier set Q, its output state is uniquely determined by the corresponding input state via a proper CPTP mapping. Following the terminology introduced in Sec. II.H (Beckman et al., 2001; Eggeling, Schlingemann, and Werner, 2002; Piani et al., 2006), they can be said to be nonsignaling (or causal) channels, meaning that the output states of any subset of the carriers cannot be influenced by the input state of the remaining carriers.

Channels which are not SI are said to exhibit intersymbol interference (ISI) (Bowen, Devetak, and Mancini, 2005), that is, the input states of previous carriers affect the action of the channel on the current input. From a physical point of view, in ISI channels there is a non-negligible backaction of the carrier onto the memory during their interaction. So the carrier's state (symbol) influences the subsequent actions of the channel. On the contrary, in SI channels the carrier does not influence the memory during their interaction. Usually this happens because the memory is much larger (in terms of degrees of freedom) of the single carrier. A pedagogical example of ISI channels is the quantum shift channel, where each input state is replaced by the previous input state, i.e., given the jth carrier q_j whose state is ρ_j, then $\Phi_{q_j}(\rho_j) = \rho_{j-1}$.

C. Quasilocal algebras approach

Until now we followed a constructive approach in which memory quantum channels were always thought of as concatenations of smaller units which, starting from an official "first carrier" element, process one quantum signal each. An alternative view where the communication lines are treated as mappings applied on infinitely long message strings was proposed by Kretschmann and Werner (2005) and Bjelaković and Boche (2008). This approach requires some advanced mathematical tools that are briefly reviewed in Appendix B.

To set the stage, suppose we have a quantum channel which transforms input states of an infinitely extended quantum lattice system (representing the infinite message string) into output states on the same system. Kretschmann and Werner (2005) formally assigned this map by working in the Heisenberg picture (see Sec. II.D) via the introduction of a completely positive and unital map $\Phi^*: \mathcal{B}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$ operating on the quasilocal algebras $\mathcal{B}^{\mathbb{Z}}$ and $\mathcal{A}^{\mathbb{Z}}$ (Bratteli and Robinson, 1979) that define the observable quantities on the lattice as described by the receiver Bob and the sender Alice, respectively. In this context one says that the channel is translational invariant or (borrowing from Bjelaković and Boche, 2008) stationary if Φ^* commutes with the shift operator on the lattice, i.e.,

$$\Phi^* \circ T_{\mathcal{B}} = T_{\mathcal{A}} \circ \Phi^* \tag{87}$$

($T_{\mathcal{B}}$ and $T_{\mathcal{A}}$ being the representation of the shift operator on $\mathcal{B}^{\mathbb{Z}}$ and $\mathcal{A}^{\mathbb{Z}}$, respectively). Furthermore, Φ^* is said to be ergodic if it is extremal in the convex set of stationary channels.[3]

Requiring then that future inputs should not affect past measurements, i.e., the nonanticipatory property (79), Kretschmann and Werner (2005) introduced the definition of a causal channel as a completely positive and unital translational invariant map Φ^* that fulfills the constraint

$$\Phi^*(O^{(-\infty,z]} \otimes \mathbb{1}^{[1+z,\infty)}) = \Phi^*(O^{(-\infty,z]}) \otimes \mathbb{1}^{[1+z,\infty)}, \tag{88}$$

for all $z \in \mathbb{Z}$ and for all $O^{(-\infty,z]} \in \mathcal{B}^{(-\infty,z]}$, where $\mathcal{B}^{(-\infty,z]}$ denotes the set of bounded operators defined on lattice elements up to that associated with the label z. In particular, memoryless configurations are obtained when also the condition

$$\Phi^*(\mathbb{1}^{(-\infty,z]} \otimes O^{[1+z,\infty)}) = \mathbb{1}^{(-\infty,z]} \otimes \Phi^*(O^{[1+z,\infty)}) \tag{89}$$

applies for all $O^{(-\infty,z]} \in \mathcal{B}^{[1+z,\infty)}$.

Examples of causal (not necessarily memoryless) maps (88) are provided by concatenated memory channels (Kretschmann and Werner, 2005) which can be easily constructed by adapting the concatenation scheme of Eqs. (83) and (84) to the quantum lattice formalism. Within this context Kretschmann and Werner (2005) proved a structure theorem which shows that any map obeying Eq. (88) can always be represented as concatenated memory channels produced by an assigned generator (see Sec. III.B).

Although cq channels can easily be included in the above formalism by expressing them as CPTP maps via the embedding (15), it is worth reviewing the approach adopted by Bjelaković and Boche (2008) to address this special set of maps. Here a cq channel taking values on the classical alphabet $\mathcal{X}$ is described as a mapping which to each $x \in \mathcal{X}^{\mathbb{Z}}$ (the set of doubly infinite sequences with components from alphabet $\mathcal{X}$) associates a complex value linear functional $W(x,\ldots)$ on $\mathcal{B}^{\mathbb{Z}}$, i.e.,

$$x \mapsto W(x,\ldots). \tag{90}$$

Ultimately, via the Gelfand-Naimark-Segal correspondence (Bratteli and Robinson, 1979), the functional $W(x,\ldots)$ can be identified with a density operator ρ_x defined on the Hilbert space $\mathcal{H}$ carrying a representation π of the quasilocal algebra (see Appendix B), through the identification $W(x,\ldots) = \mathrm{Tr}[\rho_x \pi(\cdots)]$. In this form the stationary condition (87) of the cq channel is that $W(T_{\rm in}x, b) = W(x, T_{\mathcal{B}}b)$ for all $x \in \mathcal{X}^{\mathbb{Z}}$ and all $b \in \mathcal{B}^{\mathbb{Z}}$ (here $T_{\rm in}$ and denote the shift operator on $\mathcal{X}^{\mathbb{Z}}$). The causality condition (88) is instead

$$W(x,b) = W(\tilde{x},b), \tag{91}$$

for $z \in \mathbb{Z}$, $b \in \mathcal{B}^{(-\infty,z]}$, and all $x, \tilde{x} \in \mathcal{X}^{\mathbb{Z}}$ $(x_i = \tilde{x}_i; \forall i \le z)$. Similarly, memoryless configurations (89) are recovered when Eq. (91) applies also for all $b \in \mathcal{B}^{[z,\infty)}$ and all $x, \tilde{x} \in \mathcal{X}^{\mathbb{Z}}$ $(x_i = \tilde{x}_i; \forall i \ge z)$.

D. Taxonomy of nonanticipatory quantum memory channels

Here we review those classes of nonanticipatory quantum channels which have been discussed in the literature.

1. Localizable memory quantum channels

A subset of nonanticipatory quantum channels which represent the natural multipartite generalization of the localizable maps of Beckman *et al.* (2001), Eggeling, Schlingemann, and Werner (2002), and Piani *et al.* (2006), reviewed in Sec. II.H, has been introduced in Giovannetti and Mancini (2005) and Plenio and Virmani (2007, 2008). For such models, the mapping (78) is expressed in terms of (not necessarily identical) local unitary couplings with a correlated many-body environmental system $E \coloneqq \{e_1, e_2, \ldots\}$; see Fig. 13. These transformations are clearly SI: memory effects appear because, unlike the memoryless case (76), the many-body environment is initialized in a state $\omega_E^{(n)}$ which does not factorize, i.e.,

$$\Phi^{(n)}(\rho_Q^{(n)}) = \mathrm{Tr}_E[U_{q_n e_n} \otimes \cdots \otimes U_{q_1 e_1}(\rho_Q^{(n)} \otimes \omega_E^{(n)}) \times U_{q_1 e_1}^\dagger \otimes \cdots \otimes U_{q_n e_n}^\dagger]. \tag{92}$$

It is worth mentioning that a variant of this model (Rossini, Giovannetti, and Montangero, 2008), where the local unitary interaction $U_{q_n e_n} \otimes \cdots \otimes U_{q_1 e_1}$ is replaced by a local Hamiltonian coupling between carriers and environments, is neither SI nor nonanticipatory.

An alternative representation for the localizable mappings described by Eq. (92) has also been provided in Caruso, Giovannetti, and Palma (2010) by generalizing a model presented in Ban, Sasaki, and Takeoka (2002) and Bowen and Bose (2001) for memoryless channels. In this approach the channel noise is effectively described as a quantum

[3]Note that this notion of ergodicity refers to an in-parallel composition of quantum channels and differs from ergodicity of an in-series concatenation discussed by Richter and Werner (1996), Raginsky (2002), Burgarth and Giovannetti (2007), and Burgarth *et al.* (2013), and references therein.

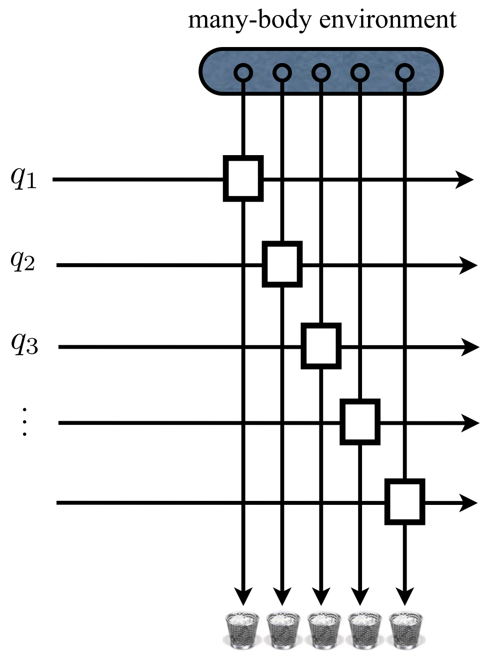


FIG. 13 (color online). Model for a localizable, fully nonanticipatory quantum memory channel. Here the correlations are introduced by allowing the state of the environment (gray element) to be initially entangled. As in the previous figures white boxes represent unitary couplings while the trash bin indicates partial trace over the corresponding degree of freedom. These maps are SI and hence nonanticipatory (therefore they also admit unitary dilations of the form described in Fig. 12).

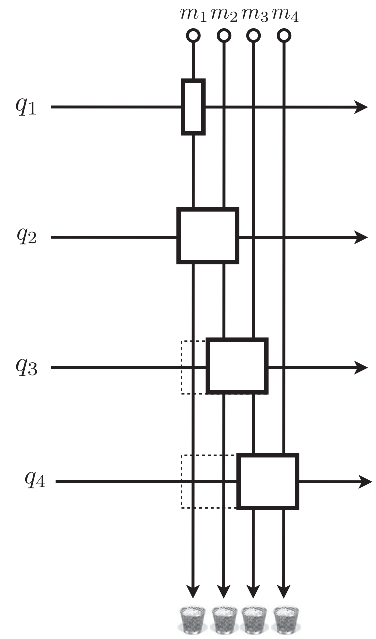


FIG. 14. Unitary dilation (80) for a finite-memory nonanticipatory quantum memory channel with correlation length $k=2$. In the depicted example the total number of channel uses is $n=4$ and each carrier is supposed to interact with only two components of the environment. Consequently, the carrier q_1 can influence the output carrier q_3 only via q_2. For a comparison see the scheme of Fig. 12(b), where instead the first carrier can directly influence q_3 via their common interaction with m_1 (symbolized here by dotted lines).

teleportation protocol (Bennett *et al.*, 1993; Vaidman, 1994; Braunstein and Kimble, 1998) that went wrong because the communicating parties used nonoptimal resources (e.g., the state they shared was not maximally entangled). In the case of Eq. (92) each of the carriers gets teleported independently using the same procedure, the correlations arising from the fact that the communicating parties use a correlated many-body quantum state as a shared resource.

2. Finite-memory channels

The expression finite-memory channels (Bowen and Mancini, 2004) is used to indicate those nonanticipatory channels that admit a representation of the form (82) with M being finite dimensional. The dimension of the memory is determined by the number of Kraus operators in the single channel expansion. Within the representation (80) examples of finite-memory channels are obtained by assuming that the unitary transformations (81) couple the carriers with no more than a fixed number k of environmental subsystems, the parameter k playing the role of the correlation length of the channel. More precisely for all $j \geq k$ one has

$$U_{q_jM} = \mathbb{1}_{m_n} \otimes \cdots \otimes \mathbb{1}_{m_{j+1}} \otimes U_{q_jm_jm_{j-1}\ldots m_{j-k}} \otimes \mathbb{1}_{m_{j-k-1}} \otimes \cdots \otimes \mathbb{1}_{m_1} \qquad (93)$$

(see Fig. 14 for a graphical representation of the case with $k=2$). Note that the case of a memoryless channel can be considered as an extreme example of finite-memory channels, where $k=1$ and each carrier interacts with a devoted component of the multipartite environment M (specifically, for each j, the carrier q_j interacts with m_j only).

3. Perfect memory channels

Memoryless channels have unitary dilations in which the environment has a dimension which is at least exponentially growing in n (i.e., $\log[\dim \mathcal{H}_E^{(n)}] = n \log d_e$) or, equivalently, by possessing a (minimal) operator sum representations whose Kraus sets contain a number of elements which is exponentially growing in n. The same property typically holds also for memory channels with the important exception of the perfect memory channels (Kretschmann and Werner, 2005; Giovannetti, Burgarth, and Mancini, 2009). Perfect memory channels are those admitting a representation as in Eq. (82), where the carriers interact only with the memory system, that is, $U_{q_jMe_j} = U_{q_jM}U_{e_j}$. The simplest example of such communication lines is obtained by assuming that the memory system M in Eq. (80) does not scale with n and it is finite dimensional. Under this hypothesis the maps $\Phi^{(n)}$ explicitly admit a unitary dilation with an environment (the system M) of constant size. A comparison with the dimension of the Hilbert space $\mathcal{H}_Q^{(n)}$ of the information carriers, which grows exponentially with n, shows that information cannot be stuck in the channel environment for a long time. As a consequence, in the asymptotic limit of long carrier sequences, no information is expected to be lost to the environment, yielding optimal communication capacity (see Sec. V). A typical example is provided by the shift channel (see Sec. III.B) which can be described as in Eq. (80) by assuming M to have the same dimension of a single carrier and by taking $U_{q_j M}$ as swapping the states of q_j and M. It is worth noting that Bowen, Devetak, and Mancini (2005) also conjectured that the memory channels that, analogously to the shift channel, display only intersymbol interference, can be represented as perfect memory channels.

More generally the class of perfect memory channels can be extended to include all the CPTP maps (78) that admit unitary dilations (80) in which the dimension d_M of the environmental system M is subexponential in n, i.e.,

$$\lim_{n\to\infty} \frac{1}{n} \log d_M = 0. \tag{94}$$

As explicitly discussed in Sec. V.C.2, in this case the channel is also asymptotically noiseless (Kretschmann and Werner, 2005; Giovannetti, Burgarth, and Mancini, 2009).

4. Markovian channels

An important class of nonanticipatory quantum channels is given by the channels with Markovian correlated noise. They describe noise models in which the carriers are transformed via the applications of strings of local CPTP maps whose elements are randomly generated by a classical Markov process. Explicitly, Markovian channels admit the following representation:

$$\Phi^{(n)}(\rho_Q^{(n)}) = \sum_{i_1,\ldots,i_n} p^{(n)}_{i_n|i_{n-1}} p^{(n-1)}_{i_{n-1}|i_{n-2}} \cdots p^{(2)}_{i_2|i_1} p^{(1)}_{i_1} \times \Phi^{(i_n)}_{q_n} \otimes \Phi^{(i_{n-1})}_{q_{n-1}} \otimes \cdots \otimes \Phi^{(i_1)}_{q_1} (\rho_Q^{(n)}), \tag{95}$$

where $\{\Phi^{(i)}_{q_j}\}_i$ is a set of CPTP maps operating on the jth carrier, $p_i^{(1)}$ is an initial probability distribution, and for $j \geq 2$ the $p^{(j)}_{i|i'}$ are conditional probabilities.

The mapping (95) is SI (see Sec. III.B) since modifying the input state of previous (or subsequent) channel uses does not have any effect on the output states of the carriers that follow (or precede). A unitary dilation of Eq. (82) can be obtained by identifying the initial state ω_M of the memory M with the vector $\sum_i \sqrt{p_i^{(1)}}\, |i\rangle_M$, and by taking the unitary $U_{q_j M e_j}$ in such a way that for all vectors $|\psi\rangle_{q_j}$ of the jth carriers one has

$$U_{q_1 M e_1} |\psi, i', 0\rangle = K^{(i')}_{q_1}(\ell) |\psi, i', \phi^{i'}_{\ell}\rangle \tag{96}$$

and

$$U_{q_j M e_j} |\psi, i', 0\rangle = \sum_i \sqrt{p^{(j)}_{i|i'}}\, K^{(i)}_{q_j}(\ell) |\psi, i, \phi^{i'}_{\ell}\rangle, \tag{97}$$

for $j \geq 2$ (in the above expressions $\{K^{(i)}_{q_j}(\ell)\}_\ell$ is a set of Kraus operators for $\Phi^{(i)}_{q_j}$, $|\psi, i, \phi_{i',\ell}\rangle$ stands for the state $|\psi\rangle_{q_j} |i\rangle_M |\phi_{i',\ell}\rangle_{e_j}$, while $\{|i\rangle_M\}_k$ and $\{|\phi_{i,\ell}\rangle_{e_j}\}_{i,\ell}$ are an orthonormal basis for the memory systems M and e_j, respectively).

Most of the analysis conducted so far focused on the special case of homogeneous Markov processes in which both $p^{(j)}_{i|i'}$ and $\Phi^{(i)}_q$ do not depend upon the carrier label j (i.e., $p^{(j)}_{i|i'} \coloneqq p_{i|i'}$). Under these conditions one also says that the quantum Markov process is regular if the corresponding classical Markov process $p_{i|i'}$ is regular, i.e., if some power of the transition matrix Γ (whose entries are the transition probabilities $p_{i|i'}$) has only strictly positive elements. In this case, for $j \to \infty$ the statistical distribution of the local noise converges to a stationary distribution $p_i^{(\infty)} \coloneqq \lim_{n\to\infty} p_i^{(n)}$, with

$$p_i^{(n)} \coloneqq \sum_{i'} (\Gamma^{n-1})_{i,i'} p_{i'}^{(1)} \tag{98}$$

being the probability of getting $\Phi^{(i)}_{q_n}$ on the nth carrier. The initial probability $p_i^{(1)}$ is said to be stationary if it satisfies the eigenvector equation $\sum_{i'} \Gamma_{i,i'} p_{i'}^{(1)} = p_i^{(1)}$ (when this happens $p_i^{(j)} = p_i^{(1)}$ and the local statistical distribution of $\Phi^{(i)}_{q_j}$ is identical for all the carriers).

The first example of a regular Markov process was analyzed by Macchiavello and Palma (2002). Here the carriers are assumed to be qubits and the CPTP transformations $\Phi^{(i)}_q$ entering in Eq. (95) are unitary rotations $\Phi^{(i)}_q(\cdots) \coloneqq \sigma_{i,q}(\cdots)\sigma_{i,q}$, where $\sigma_{0,q} = \mathbb{1}$ is the identity operator while for $i = x, y, z$, $\sigma_{i,q}$ is the Pauli matrix. The conditional probability $p_{i|i'}$ which describes the associated classical Markov process was finally written as

$$p_{i|i'} = (1-\mu) p_i^{(1)} + \mu \delta_{ii'}, \tag{99}$$

where $\mu \in [0, 1]$ is a correlation parameter (note that for $\mu = 0$ the model describes a memoryless channel while for $\mu = 1$ it describes a long-term memory channel; see Sec. III.D.8). This model of a Markovian correlated Pauli channel shows a remarkable feature when it is used for the transmission of classical information (see Sec. V). That is, when two successive uses of the channel are considered, classical information is optimally encoded in either separable states or maximally entangled states, depending on whether the correlation parameter μ is below or above a certain threshold value. This feature was first conjectured by Macchiavello and Palma (2002), then proven for certain instances of the model by Macchiavello, Palma, and Virmani (2004), and finally proven for general Markovian correlated Pauli channels by Daems (2007). Remarkably, this effect is at the root of the superadditivity property of memoryless quantum channels for transmitting classical information (Hastings, 2009) (see Sec. V.B.5).

An experimental demonstration of the optimality of entangled qubit pairs for encoding classical information through a correlated Pauli channel was provided by Banaszek *et al.* (2004) for mechanically induced correlated birefringence fluctuations, which in turn induce correlated depolarization (Ball, Dragan, and Banaszek, 2004).

A generalized model of the d-dimensional Markovian correlated Pauli channel was considered by Shadman *et al.* (2011) for the problem of sending classical information using a dense-coding protocol. An alternative model of a two-qubit correlated channel was characterized by Caruso, Giovannetti et al. (2008) in terms of the minimum output entropy.

Going beyond the case of two uses of a qubit channel, Markovian correlated depolarization over an arbitrary number of channel uses was studied by Karimipour and Memarzadeh (2006b) and Demkowicz-Dobrzanski, Kolenderski, and Banaszek (2007), and the case of Markovian correlated noise in higher dimensional quantum systems was considered by Karimipour and Memarzadeh (2006a) and Karpov, Daems, and Cerf (2006a, 2006b). Generally speaking, the optimality of entangled states for encoding classical information can be interpreted in terms of a decoherence-free subspace (see Sec. IV) associated with the correlated noise model: this was considered for the Hilbert space defined by multiple uses of a qubit channel (Demkowicz-Dobrzanski, Kolenderski, and Banaszek, 2007) and for the multiphoton Hilbert space associated with the polarization of light (Ball and Banaszek, 2005). In a different context, the same phenomenon was discussed for the problem of quantum communication with polarized light without a shared reference frame (Bartlett, Rudolph, and Spekkens, 2003).

Finally, models of Markovian correlated noise in the framework of quantum systems with continuous variables (see Sec. VI.B.2) were first discussed by Cerf *et al.* (2005, 2006) for the case of two uses of the channel, and then extended to the arbitrary number of uses in Ruggeri and Mancini (2007a), Lupo, Memarzadeh, and Mancini (2009), and Schäfer, Karpov, and Cerf (2009) (see Sec. VI.B.2).

5. Fixed-point channels

Within the representation (83) a channel is said to be a fixed-point memory channel (Bowen, Devetak, and Mancini, 2005) if the initial memory state ω_M of the representation is left invariant after each interaction with the carriers. Specifically, recalling the definition (86) this notion is formalized by the following identity:

$$\Psi^{(n)}(\rho_Q^{(n)};\omega_M) = \omega_M, \qquad \forall\ \rho_Q^{(n)} \in \mathfrak{S}(\mathcal{H}_Q^{(n)}). \tag{100}$$

Fixed-point channels can easily be shown to be symbol independent while the opposite is not necessarily true. Indeed from Eq. (84) one has that the output state of nth carrier $\rho'_{q_n} \coloneqq \mathrm{Tr}_{Q^{(n-1)}}[\Phi^{(n)}(\rho_Q^{(n)})]$ can be expressed as

$$\begin{aligned}\rho'_{q_n} &= \mathrm{Tr}_{Q^{(n-1)}M}[(\Phi_{q_nM}\circ\Phi_{QM}^{(n-1)})(\rho_Q^{(n)}\otimes\omega_M)]\\ &= \mathrm{Tr}_M[\Phi_{qM}(\Psi^{(n-1)}(\rho_Q^{(n)};\omega_M))]\\ &= \mathrm{Tr}_M[\Phi_{q_nM}(\rho_{q_n}\otimes\omega_M)],\end{aligned} \tag{101}$$

which depends only upon the reduced density operator ρ_{q_n} and not on the previous information carriers [in Eq. (101) $\mathrm{Tr}_{Q^{(n-1)}M}$ indicates the partial trace with respect to M and the first $(n-1)$ carriers].

Markovian memory channels are examples of fixed-point memory channels, in which the memory system can be represented by the classical variable of the underlying Markov chain. Being classical, the memory system can be chosen in such a way that it is unaffected by the backaction of the input system. This representation can be made explicitly by choosing a unitary dilation of the form (96). Another example is provided by Plenio and Virmani (2007, 2008), in which the input system interacts with the memory system by a controlled-unitary transformation, where the memory is the control and the system is the target. In this setting, the resulting memory channel is a fixed-point one if the initial state of the memory is diagonal in the control basis.

6. Indecomposable and forgetful channels

An *indecomposable* channel is one where, for each channel input, the long-term behavior of the channel is independent of the initial memory state (Bowen, Devetak, and Mancini, 2005). Such independence can be quantified by evaluating the distance between different trajectories $\Psi^{(n)}(\rho_Q^{(n)};\omega_M)$ and $\Psi^{(n)}(\rho_Q^{(n)};\omega'_M)$ associated through Eq. (86) to two different initial memory configurations ω_M and ω'_M. Specifically a finite-memory quantum channel is said to be indecomposable if for any input state $\rho_Q^{(n)}$ and $\epsilon > 0$ there exists an $N(\epsilon)$ such that for $n \geq N(\epsilon)$,

$$D[\Psi^{(n)}(\rho_Q^{(n)};\omega_M), \Psi^{(n)}(\rho_Q^{(n)};\omega'_M)] \leq \epsilon, \tag{102}$$

for any pair of initial states of the memory ω_M, ω'_M [here D is the trace distance—see Eq. (A1)]. Equivalently Eq. (102) can be stated by saying that for large n, $\Psi^{(n)}(\rho_Q^{(n)};\omega_M)$ converges to a state of M which depends on $\rho_Q^{(n)}$ but not on ω_M [compare this with the behavior (100) of the fixed-point memory channels]. Here one notices that for finite dimensional systems this implies that there exists a family of CPTP channels $\Theta^{(n)}\colon \mathfrak{S}(\mathcal{H}_Q^{(n)}) \to \mathfrak{S}(\mathcal{H}_M)$ which fulfills the identity

$$\mathrm{Tr}_Q[\Phi_{QM}^{(n)}(O_{QM}^{(n)})] \longrightarrow \Theta^{(n)}(O_Q^{(n)}), \tag{103}$$

in the limit $n \to \infty$ for all the operators $O_{QM}^{(n)}$ on Q and M, with $O_Q^{(n)} = \mathrm{Tr}_M[O_{QM}^{(n)}]$. In the Heisenberg picture (see Sec. II.D) this can also be stated as

$$\Phi_{QM}^{(n)\,*}(\mathbb{1}_Q\otimes\cdots) \to \Theta^{(n)*}(\cdots)\otimes\mathbb{1}_M, \tag{104}$$

with $\Phi_{QM}^{(n)\,*}$ and $\Theta^{(n)*}$ being the dual of $\Phi_{QM}^{(n)}$ and $\Theta^{(n)}$, respectively.

The main features of indecomposable channels were revisited through the notion of forgetful channels (Kretschmann and Werner, 2005). The latter was originally introduced in the quasilocal algebra approach detailed in Sec. III.C, where the quantum memory channels are assumed to be translation invariant and nonanticipatory. In the representation (83) this definition coincides with the limiting condition (103), which in Kretschmann and Werner (2005) is written in terms of the cb-norm distance (see Sec. II.J.2). In this context a memory channel is said to be strictly forgetful if there exists a finite integer m such that the right-hand side (rhs) and the left-hand side (lhs) of Eq. (103) exactly coincide for all $n \geq m$. A simple example of forgetful channels can be obtained if $\Phi_{QM}^{(n)}$ is defined by the concatenation (85) of a generator map $\Phi_{qM} \coloneqq p\mathrm{id} + (1-p)\mathrm{SWAP}$, where $p \in [0,1)$, and SWAP denotes the swap channel which exchanges q and M. In this case the only way for $\Phi_{QM}^{(n)}$ not to be forgetful is to choose the ideal channel in every step of the concatenation. However, the probability for this event vanishes in the limit $n \to \infty$ as p^n, implying that Eq. (103) holds.

Several criteria for a quantum memory channel to be forgetful have been proposed (Kretschmann and Werner, 2005). For instance, a sufficient condition is that the cb-norm distance between the rhs and the lhs of Eq. (103) falls below 1 for some finite n. From a physical point of view, one could expect a generic quantum memory channel to be forgetful. Indeed, it can be proven that the subset of forgetful channels is dense and open (according to the topology induced by the cb norm) (Kretschmann and Werner, 2005).

In the case of Markovian channels, the forgetfulness is determined by the asymptotic properties of the underlying Markov chain: in particular, for a discrete-variable memory system, the channel is forgetful if and only if the underlying Markov chain converges to a unique stationary state (Datta and Dorlas, 2009). On the other hand, if the memory system is described by continuous variables, one could have situations in which the Markov chain has a unique stationary state, yet the convergence property (103) in the cb norm is not satisfied. To overcome this limitation, a weaker notion of forgetfulness, named weak forgetfulness, was introduced by Lupo, Memarzadeh, and Mancini (2009) for Markovian channels. Although restricted to this setting, its definition coincides with that of indecomposability [Eq. (102)] and is equivalent to forgetfulness for a discrete-variable Markov chain. Beyond this setting, the model of Gaussian memory channels in Lupo, Giovannetti, and Mancini (2010a) was proven to be indecomposable under restricted conditions on the memory initialization, e.g., if the initial state of the memory is a Gaussian state with finite first and second moments. Finally, the relation between the forgetfulness of the channel and the chaotic quantum evolution of the memory system was studied by Barreto Lemos and Benenti (2010) for a model of a dephasing channel with memory.

7. Decaying input memory cq channels

Forgetful channels represent configurations in which the effect of the far past inputs do not strongly affect present and future outputs. Within the quasialgebra approach a similar notion was developed by Bjelaković and Boche (2008) for the special class of cq channels (see Sec. III.C). Specifically a cq channel defined by Eq. (90) is said to have decaying input memory if for each $\epsilon > 0$ there exists a non-negative integer $m(\epsilon)$ such that

$$|W(x,b) - W(x',b)| \leq \epsilon, \tag{105}$$

for all $b \in \mathcal{B}^{[n,\infty]}$, $n, \in \mathbb{Z}$, whenever $x_i = x_i'$ for $n - m \leq i$ and $m \geq m(\epsilon)$. Note that $\sum_b |W(x,b) - W(x',b)|$ is a distance between quantum states. Then Eq. (105) says that, starting from the nth use, the outputs of two identical channels W are almost (i.e., within a distance ϵ) the same provided that the inputs have started to be identical from the $n - m(\epsilon)$th use. Hence $m(\epsilon)$ gives an estimation of the memory length.

This provides a "continuity" property of the channel which plays a crucial role in establishing coding theorems, an idea that also appears in Sec. IV.D and goes back to the classic paper by McMillan (1953).

8. Long-term memory channels

Long-term quantum memory channels describe those communication lines in which the effect of the memory does not decay with the number of channel uses. These channels are defined as those memory channels which are not forgetful.

Extreme examples are provided by statistical mixtures of memoryless channels (77) of Sec. III.A.1 as pointed out by Datta and Dorlas (2007) and Datta, Suhov, and Dorlas (2008). The memory correlations of this class of channels can be considered to be given by a Markov chain which is aperiodic but not irreducible (Norris, 1997). This can be easily seen by noticing that Eq. (95) reduces to Eq. (77) by setting $p_{i|i'}^{(j)} = \delta_{ii'}$ for all $j = 2, \ldots, n$. Hence, once a particular branch $i = 1, \ldots, M$ has been chosen, the successive inputs are sent through this branch (aperiodicity) and transition between the different branches (which correspond to the different states of the Markov chain) is not permitted (reducibility).

It is worth stressing that the transformation (77) is fully nonanticipatory (i.e., the input state of any subset of carriers cannot influence the output state of the remaining ones): as a consequence, fixing an ordering, it can always be represented as in Fig. 12 with a proper choice of the unitary couplings.

IV. QUANTUM CODES

Coding theory is the branch of information science studying how to use software strategies to counteract the effect of an assigned noise source affecting a communication line or the components (memory elements) of a database. In a sense it can be described as the last resort which can be exploited once no further improvements can be obtained at the level of hardware engineering.

Both in the classical and in the quantum setting, the key idea to prevent the corruption of information is to use redundancy: by properly spreading a given message over many information carriers instead of a single one, one can take advantage of the structural properties of a noise source. Consider for instance the paradigmatic case in which one wishes to store the information contained in (say) k qudits, affected by an assigned error model described by the channel $\Phi^{(k)}$, into a larger set of $n \geq k$ qudits, affected by the noise $\Phi^{(n)}$. Then a coding strategy consists of identifying an encoding CPTP map $\Phi_E^{(k\to n)}: \mathfrak{S}(\mathcal{H}^{\otimes k}) \to \mathfrak{S}(\mathcal{H}^{\otimes n})$, shuffling a state from the smaller space of the k carriers to the larger space of the n carries, and a decoding CPTP map $\Phi_D^{(n\to k)}: \mathfrak{S}(\mathcal{H}^{\otimes n}) \to \mathfrak{S}(\mathcal{H}^{\otimes k})$, moving the information back to the original space, under the requirement that the resulting channel $\Phi_D^{(n\to k)} \circ \Phi^{(n)} \circ \Phi_E^{(k\to n)}$ is somehow "less" noisy than the original transformation $\Phi^{(k)}$ (see Fig. 18). The image $\mathcal{Q}$ of

$\Phi_E^{(k\to n)}$ is called a *quantum error-correcting code* (QECC): it represents the information vault where messages are deposited to prevent the noise from affecting them. The ratio $R = (k/n)\log d$ represents the communication rate (in qubits per channel use) of the code, whose inverse measures how much information spread is involved in the procedure. It is worth noting that if $\Phi_E^{(k\to n)}$ is taken to be an isometry (an option which is often implicitly assumed in QECC) the space $\mathcal{Q}$ becomes a proper vector subspace of $\mathcal{H}^{\otimes n}$ of dimension d^k. When this happens $\Phi^{(n)}\circ\Phi_E^{(k\to n)}$ is just a restriction of $\Phi^{(n)}$ on $\mathcal{Q}$ and the encoding mapping can be fully specified by simply assigning the latter. This also justifies the consideration of a recovery map $\Phi_R^{(n)}$ in place of the decoding map $\Phi_D^{(n\to k)}$ (in the following this simplification is assumed).

Different equivalent ways have been devised to evaluate the quality of a given coding procedure [Cafaro *et al.* (2011) analyzed and compared some of them]. The most commonly used is by means of the input-output fidelities introduced in Sec. II.J.1. In particular, good choices are the minimum or average fidelity functionals, i.e., $F_{\min}(\Phi_D^{(n\to k)}\circ\Phi^{(n)}\circ\Phi_E^{(k\to n)})$ and $\bar{F}(\Phi_D^{(n\to k)}\circ\Phi^{(n)}\circ\Phi_E^{(k\to n)})$. It is important however to distinguish between two different scenarios: the case where the messages to be stored or transmitted are purely classical, and the case where instead they are quantum. In the first scenario the minimization (respectively, average) involved in Eq. (48) [respectively, Eq. (49)] needs not be performed over the entire input space of the k carriers, but only with respect to the orthogonal set of states in $\mathcal{H}^{\otimes k}$ which are encoding the classical messages one wishes to protect. Vice versa in the second scenario, which implies the possibility of producing arbitrary superposition of the input signals, the minimization (respectively, average) is performed over the whole input space. In this last configuration the effectiveness of a correcting code can also be quantified by other distance measures, such as the entanglement fidelity $F_e(\mathbb{1}/d^k; \Phi_D^{(n\to k)}\circ\Phi^{(n)}\circ\Phi_E^{(k\to n)})$ defined in Eq. (54), or the cb-norm distance of $\Phi_D^{(n\to k)}\circ\Phi^{(n)}\circ\Phi_E^{(k\to n)}$ from the identity channel id on the set of the k carriers; see Sec. II.J.2. The simplest example of error-correcting code is obtained for counteracting bit-flip errors by repeating qubit basis states (preferably an odd number of times); see Fig. 15.

When a given code allows for the exact protection of the data stored in the k carriers [e.g., when $F_{\min} = 1$, implying zero-error probability (47) for all possible channel inputs], the code is said to be perfect. However, in realistic situations codes allow only for arbitrary high fidelity at a finite rate in the limit of large code length ($n \to \infty$); see Sec. V. For finite code length, one has to find the optimal compromise between rate, fidelity, and the complexity of the coding and decoding operations. Here the focus is on the general properties of quantum error-correcting codes. The large body of works in this field was mainly concerned with the development of strategies for independent and identically distributed (i.i.d.) errors, i.e., for noise arising from memoryless quantum channels. In this section, after reviewing the basics of such codes, we analyze their effectiveness against correlated errors. Then possible answers to the new challenges posed by memory effects are discussed. Since the extension of the formalism for QECC (mostly relying on group theory) from $\mathcal{H} \simeq \mathbb{C}^2$ to higher dimensional $\mathcal{H}$ is nontrivial, the presentation will be restricted to qubit systems, i.e., to binary quantum codes. For nonbinary codes see Knill (1996), Ashikhmin and Knill (2001), and Ketkar *et al.* (2006).

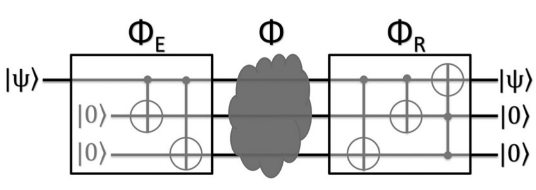


FIG. 15. The encoding and recovery maps for a qubit repetition code, able to counteract uncorrelated bit-flip errors. The circuital implementation of encoding and decoding maps involves controlled NOT operations [for a primer on qubit logical operation, see Mermin (2003)]. Upon encoding the qubit basis states are spreader over three qubits as $|0\rangle \mapsto |000\rangle$ and $|1\rangle \mapsto |111\rangle$. The reverse happens upon recovery provided that no more than one qubit is affected by bit flip.

A. Standard quantum coding theory

Consider a memoryless quantum channel $\Phi^{(n)} = \Phi^{\otimes n}$, characterized by a set of Kraus operators $\{K_{\mathsf{i}}\}$, on the Hilbert space $\mathbb{C}^{2\otimes n}$ of n qubits

$$\rho \mapsto \Phi^{(n)}(\rho) = \sum_{\mathsf{i}} K_{\mathsf{i}}\rho K_{\mathsf{i}}^{\dagger}, \qquad (106)$$

where $K_{\mathsf{i}} = K_{i_1} \otimes \cdots \otimes K_{i_n}$ describes i.i.d. errors on single qubits. An error-correcting code is identified by a 2^k-dimensional subspace $\mathcal{Q} \subset \mathbb{C}^{2\otimes n}$. We denote as $P_{\mathcal{Q}}$ the projector operator associated with $\mathcal{Q}$. It is possible to show that the code is able to correct errors belonging to a subset $\mathfrak{Q}\subseteq\{K_{\mathsf{i}}\}_{\mathsf{i}}$ if and only if there exists a Hermitian matrix S such that

$$P_{\mathcal{Q}} K_{\mathsf{l}}^{\dagger} K_{\mathsf{m}} P_{\mathcal{Q}} = \mathsf{S}_{\mathsf{lm}} P_{\mathcal{Q}}, \qquad (107)$$

for any pair of error operators $K_{\mathsf{l}}, K_{\mathsf{m}} \in \mathfrak{Q}$ (Knill *et al.*, 2002). Because of the unitary freedom of the Kraus representation, Eq. (107) holds true if and only if a Kraus representation of the map $\Phi^{(n)}$ exists, say $\Phi^{(n)}(\rho) = \sum_{\mathsf{i}} K'_{\mathsf{i}}\rho K'^{\dagger}_{\mathsf{i}}$, such that $P_{\mathcal{Q}} K'^{\dagger}_{\mathsf{l}} K'_{\mathsf{m}} P_{\mathcal{Q}} = \delta_{\mathsf{lm}} s_{\mathsf{m}} P_{\mathcal{Q}}$ for some set of non-negative numbers s_{m}. This condition in turn yields the fact that different correctable error operators map the code words in $\mathcal{Q}$ into mutually orthogonal subspaces. This property implies that different errors can be detected by applying a projective measurement able to distinguish the different orthogonal subspaces. Moreover, if the same error applies to two different basis code words, their scalar product will not change. Thus, the geometrical interpretation is that each correctable error maps the code space into an orthogonal subspace without deformations.

Applying again the unitary freedom of the Kraus representation, one can find a basis where any pair of errors acting on a given code word produces either orthogonal states or exactly the same state. This phenomenon will occur if the matrix S results singular. In such a case the quantum code $\mathcal{Q}$ is called degenerate.

Now let $\mathcal{V}^{i_L}$ be the subspace of $\mathbb{C}^{2\otimes n}$ spanned by the corrupted images $\{K'_{\mathsf{l}}|i_L\rangle\}_{\mathsf{l}}$ of the code words $|i_L\rangle$ and let

$\{|v_r^{i_L}\rangle\}$ be an orthonormal basis of $\mathcal{V}^{i_L}$. We define a subspace $\mathcal{V}^{i_L}$ for each code word. Then the recovery map $\Phi_R^{(n)}$ is characterized by Kraus operators $\{R_r\}$ such that

$$R_r = \sum_i |i_L\rangle\langle v_r^{i_L}|. \tag{108}$$

In the qubit context we are considering, an $[\![n,k]\!]$ quantum correcting code $\mathcal{Q}$ is given by a 2^k-dimensional subspace of $\mathbb{C}^{2\otimes n}$ encoding k logical qubits into n physical qubits ($n \geq k$).

One can assume the error operators to be proportional to the Pauli operators acting on the jth qubit and corresponding to no-error, bit-flip error, bit-phase-flip error, and phase-flip error, i.e., $K_{0_j} \propto \mathbb{1}_j$, $K_{1_j} \propto \sigma_{x,j}$, $K_{2_j} \propto \sigma_{y,j}$, and $K_{3_j} \propto \sigma_{z,j}$. This restriction to Pauli errors represents no loss of generality. Indeed, it can be easily shown that if a code corrects a given set of errors it can also correct any linear combination (by complex coefficients) (Nielsen and Chuang, 2000). It is hence sufficient to restrict to Pauli operators since they are a basis on the space of qubit operators.

Given a subset $\mathfrak{Q} \subseteq \{K_\mathrm{i}\}_\mathrm{i}$ of errors that can be corrected one says that the code $\mathcal{Q}$ has $\mathfrak{Q}$-correcting ability. To each K_i one can assign a weight t, an integer $0 \leq t \leq n$ denoting the number of qubits, where operators K_{i_j} $(j = 1, \ldots, n)$ act differently from identity. Then the correction ability of $\mathcal{Q}$ can also be expressed by specifying the value of the distance $d = 2t + 1$ of the code, meaning that $\mathcal{Q}$ corrects all errors affecting at most t qubits.

Suppose that errors are i.i.d. with probability p_e on each qubit, then for any of the

$$\binom{n}{t+1}$$

ways of choosing $t+1$ locations, the probability that errors occur at every one of those locations results in p_e^{t+1}. Therefore one has the following upper bound on the probability that at least $t+1$ errors occur in the block of n qubits:

$$\binom{n}{t+1} p_e^{t+1}.$$

This means that for p_e small the performance of the code $1 - F \approx \mathcal{O}(p_e^{t+1})$ is substantially improved over the unprotected data $1 - F \approx \mathcal{O}(p_e)$.

An upper bound on the rates achievable by nondegenerate quantum codes is given by the quantum version of the (classical) Hamming bound (Ekert and Macchiavello, 1996)

$$2^k \sum_{i=0}^{t} 3^i \binom{n}{i} \leq 2^n, \tag{109}$$

which for large n and d/n fixed yields the approximate bound

$$R \leq 1 - \frac{d}{2n}\log_2 3 - h\left(\frac{d}{2n}\right)$$

with h the binary entropy (66).

There are also upper bounds that apply to all quantum codes, not just nondegenerate ones, such as the quantum Singleton bound (Knill and Laflamme, 1997)

$$n \geq 4t + k. \tag{110}$$

On the other hand, a lower bound on the rates, confirming that good codes indeed exist (Calderbank *et al.*, 1997), comes from the quantum version of the Gilbert-Varshamov theorem, stating that a $[\![n,k]\!]$ quantum code of distance $d = 2t+1$ exists with

$$k \geq \max\left\{k' \middle| 2^{k'} \sum_{i=0}^{2t} 3^i \binom{n}{i} \leq 2^n\right\}. \tag{111}$$

For large n and d/n fixed one gets the approximate bound

$$R \geq 1 - \frac{d}{n}\log_2 3 - h\left(\frac{d}{n}\right).$$

Unfortunately the explicit construction of quantum codes is not an easy task. Historically the first quantum code that appeared was a $[\![9,1]\!]$ code with $d = 3$ (Shor, 1995) whose basis code words read

$$\left[\frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)\right]^{\otimes 3}, \quad \left[\frac{1}{\sqrt{2}}(|000\rangle - |111\rangle)\right]^{\otimes 3}. \tag{112}$$

Its construction relies on a simple argument. A three-qubit code suffices to protect against a single bit flip (see Fig. 15). The reason the three-qubit clusters are repeated 3 times is to protect against phase errors as well.

Then attempts were made following classical linear codes (Hill, 1985). In the classical setting, the state of n bit system is represented by a binary string of length n. An error affecting this string can also be represented by a binary string of the same length, where the 1's indicate the locations of the bits that have been flipped. The action of an error string on a code words string is hence represented by a summation modulo two. The space of binary strings of length n, endowed with the summation modulo two, defines the linear space $\mathbb{F}_2^n$. A classical $[n,k]$ linear code $\mathcal{C}$ is defined as a subspace $\mathcal{C} \subseteq \mathbb{F}_2^n$. The subspace can be characterized by a set of k generators or equivalently by a parity check $(n-k) \times n$ matrix H such that $\mathsf{H}\mathsf{v}^\top = 0$, $\forall \mathsf{v} \in \mathcal{C}$. Errors e_i taking $\mathsf{v} \in \mathcal{C}$ into $\mathsf{v} + \mathsf{e}_i$ can be detected by applying the parity check $\mathsf{H}(\mathsf{v} + \mathsf{e}_i)^\top = \mathsf{H}\mathsf{e}_i^\top = \mathrm{synd}(\mathsf{e}_i)$ (error syndrome) and corrected iff they give rise to distinct syndromes, i.e., $\mathsf{H}(\mathsf{e}_i + \mathsf{e}_j)^\top \neq 0$ for $i \neq j$. If $\mathcal{C}$ is with distance $d = 2t+1$ it means that it is able to correct up to t errors, i.e., bit-flip errors in at most t bits. The set of errors correctable by $\mathcal{C}$ is denoted by $\mathfrak{C}$.

An advantage of linear codes over general error-correcting codes is their compact specification. Using them, quasiclassical [or Calderbank, Shor, and Steane (CSS)] codes were constructed in the following way (Calderbank and Shor, 1996; Steane, 1996). Consider two classical linear codes $\mathcal{C}_1$ and $\mathcal{C}_2$ such that $\mathcal{C}_2^\perp \subseteq \mathcal{C}_1$, where $\mathcal{C}_2^\perp$ is the dual code to $\mathcal{C}_2$, i.e., consisting of those bit strings that are orthogonal to the code words of $\mathcal{C}_2$. If $\mathcal{C}_1$ is a $[n, k_1]$ code with distance d_1 and $\mathcal{C}_2$ is a $[n, k_2]$ code with distance d_2, then the corresponding CSS quantum code is a $[[n, k_1 + k_2 - n]]$ code with distance $\min\{d_1, d_2\}$. Its basis code words are

$$\frac{1}{\sqrt{|\mathcal{C}_2^\perp|}} \sum_{\mathsf{w} \in \mathcal{C}_2^\perp} |\mathsf{u} + \mathsf{w}\rangle, \qquad \mathsf{u} \in \mathcal{C}_1. \tag{113}$$

Performing the Hadamard transform on each qubit of the code one can switch from $\mathcal{C}_1$ to $\mathcal{C}_2$ to account for σ_z errors beside σ_x ones. As a matter of fact it takes Eq. (113) to

$$\frac{1}{\sqrt{2^n |\mathcal{C}_2^\perp|}} \sum_{\mathsf{x} \in \mathbb{F}_2^n} \sum_{\mathsf{w} \in \mathcal{C}_2^\perp} (-1)^{\mathsf{x} \cdot (\mathsf{u} + \mathsf{w})} |\mathsf{x}\rangle.$$

Since $\sum_{\mathsf{w} \in \mathcal{C}_2^\perp} (-1)^{\mathsf{x} \cdot \mathsf{w}} \neq 0$ if $\mathsf{x} \in \mathcal{C}_1^\perp$ and zero otherwise, one is left with a state $\propto \sum_{\mathsf{x} \in \mathcal{C}_1^\perp} |\mathsf{x}\rangle$. This latter, being $\mathcal{C}_1^\perp \subseteq \mathcal{C}_2$, can be considered as an instance of basis code words

$$\frac{1}{\sqrt{|\mathcal{C}_1^\perp|}} \sum_{\mathsf{w} \in \mathcal{C}_1^\perp} |\mathsf{u} + \mathsf{w}\rangle, \qquad \mathsf{u} \in \mathcal{C}_2. \tag{114}$$

Therefore, to correct errors it is enough to implement the parity check of $\mathcal{C}_1$ by measuring in the σ_z basis (for bit-flip errors) and that of $\mathcal{C}_2$ by measuring in the σ_x basis (for phase flip errors). An example along this line is provided by the $[\![7, 1]\!]$ code with $d = 3$ (Steane, 1996).

Another, more general, way to construct quantum codes is to exploit the group structure of the set of errors as is done for *stabilizer codes* (Gottesman, 1997). The method can be summarized as follows. First note that the set of Pauli errors on n qubit can be written as[4]

$$\mathcal{P}_n := \left\{ \bigotimes_{j=1}^{n} \sigma_{x,j}^{X(j)} \bigotimes_{j=1}^{n} \sigma_{z,j}^{Z(j)} | X(j), Z(j) \in \mathbb{F}_2 \right\}, \tag{115}$$

and forms a multiplicative group known as a Pauli group. One can represent the elements of $\mathcal{P}_n$ as $2n$-dimensional binary vectors

$$e \in \mathcal{P}_n \leftrightarrow (\mathsf{e}_X | \mathsf{e}_Z) \equiv \mathsf{e} \in \mathbb{F}_2^n \times \mathbb{F}_2^n, \tag{116}$$

where e_X (respectively, e_Z) is the n-bits vector of components $X^{(j)}$ (respectively, $Z^{(j)}$) specifying on which qubits the σ_x (respectively, σ_z) error occurs and $(\mathsf{e}_X|\mathsf{e}_Z)$ is the joint $\mathsf{e}_X, \mathsf{e}_Z$ vector. Then one considers an Abelian subgroup $\mathcal{G} \subseteq \mathcal{P}_n$:

$$\mathcal{G} = \operatorname{span}\{g_i \in \mathcal{P}_n | 1 \leq i \leq n - k\}, \tag{117}$$

where $g_1, g_2, \ldots, g_{n-k}$ are independent of each other. Note that the operators in $\mathcal{P}_n$ have eigenvalues ± 1. An n-qubit vector $|x\rangle$ is said to be stabilized by the group $\mathcal{G}$ if it is a common eigenvector with eigenvalue $+1$. The set of vectors stabilized by $\mathcal{G}$ forms a 2^k-dimensional subspace

$$\mathcal{Q} = \{|x\rangle \in \mathbb{C}^{2\otimes n} : g|x\rangle = |x\rangle, \forall\ g \in \mathcal{G}\}. \tag{118}$$

The subspace $\mathcal{Q}$ is a $[\![n, k]\!]$ code. Note that all errors belonging to the group $\mathcal{G}$ will leave the code word unaffected. The other errors will in general change the n-qubit states. To detect which error has occurred one measures the set of $n - k$ commuting observables g_i. The results of these measurements are either $+1$ or -1. The corresponding set of measurement results plays the role of the error syndrome. Errors with nontrivial error syndromes are detectable, while errors with different error syndromes are correctable.

One can describe stabilizer codes using the same formalism of classical linear codes. By using the vectors in $\mathbb{F}_2^n \times \mathbb{F}_2^n$ corresponding to the generators $g_1, g_2, \ldots, g_{n-k}$ it is possible to write down the following $(n - k) \times 2n$ parity check matrix:

$$\mathsf{H} := \begin{pmatrix} \mathsf{g}_{1,X} | \mathsf{g}_{1,Z} \\ \vdots \\ \mathsf{g}_{n-k,X} | \mathsf{g}_{n-k,Z} \end{pmatrix} \tag{119}$$

for a $[2n, k]$ classical linear code $\mathcal{C}$ [its jth row is given by the vector $(\mathsf{g}_{j,X}|\mathsf{g}_{j,Z})$] corresponding to $\mathcal{Q}$. Then the analysis of correcting $\mathfrak{Q} \subseteq \mathcal{P}_n$ errors by $\mathcal{Q}$ can be traced back to that of correcting $\mathfrak{C} \subseteq \mathbb{F}_2^n \times \mathbb{F}_2^n$ errors by $\mathcal{C}$.

In this way it results (Gaitan, 2008) in the fact that $\mathcal{Q}$ has $\mathfrak{Q}$-correcting ability iff for every $(\mathsf{e}_{1,X}|\mathsf{e}_{1,Z}), (\mathsf{e}_{2,X}|\mathsf{e}_{2,Z}) \in \mathfrak{C}$ it is

$$\mathsf{H}(\mathsf{e}_{1,X} + \mathsf{e}_{2,X} | \mathsf{e}_{1,Z} + \mathsf{e}_{2,Z})^\top \neq 0. \tag{120}$$

Equation (120) states that $\mathfrak{Q}$ is correctable by $\mathcal{Q}$ iff $\operatorname{synd}(\mathsf{e}_i) \neq \operatorname{synd}(\mathsf{e}_j)$ for all $\mathsf{e}_i, \mathsf{e}_j \in \mathfrak{C}$ [which corresponds to Eq. (107)].

A stabilizer code of distance d has the property that each element of $\mathcal{P}_n$ of weight t less than d either lies in the stabilizer or anticommutes with some element of the stabilizer. An example is provided by the $[\![5, 1]\!]$ code with $d = 3$ introduced by Laflamme *et al.* (1996) and saturating the quantum Hamming bound (109). It is worth remarking that a systematic method to find stabilizer generators exists based on the connection with vectors over Galois field $GF(4)$ (Calderbank *et al.*, 1998).

If the subgroup $\mathcal{S} \subseteq \mathcal{P}_n$ is not Abelian, it can be used as well to construct a QECC provided that entanglement between encoder and decoder is available (Brun, Devetak, and Hsieh, 2006). The trick consists of extending the generators of $\mathcal{S}$ (by attaching extra Pauli operators at their end) in order to generate a new group $\mathcal{S}'$ that is Abelian and for which the above theory can be applied. These are entangled-assisted QECC and the notation $[\![n, k; m]\!]$, with m denoting the number of entangled ancilla qubits ($n - m - k$ giving the number of unentangled ancilla qubits), is used. Entanglement-assisted codes may lead to rates higher than their non-entangled counterparts. The reason is that entanglement allows one to increase the dimension of the decoding Hilbert space to 2^{n+m} compared to 2^n for unentangled ancillas. This also leads to a revision of the Hamming bound (109) with 2^{n+m} on the rhs (Bennett *et al.*, 2002).

Finally note that in constructing codes, besides pursuing the highest possible rate, one should also take into account the complexity of the encoding and decoding procedures. This can be evaluated by means of the number of elementary steps, i.e., the number of elementary gate operations, needed. Efficient encoding and decoding requires polynomial (actually near linear) scaling of complexity versus block code length n. Luckily, stabilizer codes are efficiently encodable or

[4] Actually $\sigma_y = i\sigma_x\sigma_z$; however, the imaginary unit, as well as any global phase factor, is irrelevant in quantum error correction.

decodable (Gottesman, 1997), but usually do not achieve the channel capacity (see Sec. V).

B. Codes concatenation

Unfortunately, using the previous approach it is quite hard to construct good codes with large distances. Families of codes that offer good performance by increasing the distance are the toric codes (Kitaev, 2003) and the quantum version of Reed-Muller codes (Steane, 1999). Aside from them, a particularly simple way to construct codes that can correct multiple errors is to concatenate several single-error correcting codes, i.e., codes with $d = 3$. For simplicity, we illustrate the case of two layers of concatenation and consider single qubit encoding. Assume the inner code (first layer) is a $[\![n_1, k_1]\!]$ stabilizer code $\mathcal{Q}_1$ with distance d_1, and the outer code (second layer) is a $[\![n_2, 1]\!]$ stabilizer code $\mathcal{Q}_2$ with distance d_2. The concatenated code $\mathcal{Q} = \mathcal{Q}_1 \circ \mathcal{Q}_2$ maps k_1 qubits into $n = n_1 n_2$ qubits, with code construction parsing the n qubits into n_2 blocks $B(b)$ ($b = 1, \ldots, n_2$) each containing n_1 qubits. Explicitly, the concatenated code $\mathcal{Q}$ is constructed as follows. For any code word $|c_{\rm out}\rangle$ of the outer code $\mathcal{Q}_2$,

$$|c_{\rm out}\rangle = \sum_{i_1 \cdots i_{n_2}} \alpha_{i_1 \cdots i_{n_2}} |i_1 \cdots i_{n_2}\rangle, \tag{121}$$

with $|i_1 \cdots i_{n_2}\rangle = |i_1\rangle \otimes \cdots \otimes |i_{n_2}\rangle$, replace each basis vector $|i_j\rangle$ by a basis vector $|\phi_{i_j}\rangle$ of the inner code $\mathcal{Q}_1$, so that

$$|c_{\rm conc}\rangle := \sum_{i_1 \cdots i_{n_2}} \alpha_{i_1 \cdots i_{n_2}} |\phi_{i_1}\rangle \otimes \cdots \otimes |\phi_{i_{n_2}}\rangle. \tag{122}$$

Note that the above-mentioned construction produces a $[\![n_1 n_2, k_1]\!]$ code with distance $d \geq d_1 d_2$.

If there are L levels of concatenation of the same single qubit code, and p_e is the error probability on a single qubit, it is possible to show that the code failure probability is bounded by (Gaitan, 2008)

$$p_e^{(L)} \leq p_0 \left(\frac{p_e}{p_0}\right)^{2^L}, \tag{123}$$

where p_0 is an estimate of the threshold error probability that can be tolerated and depends on the chosen single qubit code. Hence, provided that $p_e < p_0$, one can make the code failure probability as small as one wants by adding enough levels to the code.

Finally, it is worth remarking that minimum distance is not everything. It helps in constructing good codes (codes with arbitrarily small error probability). However, very good codes (codes with arbitrarily small error probability and achieving maximum rate) can be constructed even with bad (small) minimum distance. The reason can be understood by means of a metaphor due to Berlekamp [see, e.g., Mac Kay (2003)], and applicable in both classical and quantum frameworks. A blind bat lives in a cave and flies about the center of the cave which corresponds to one code word with its typical distance from the center controlled by the error rate. The boundaries of the cave are made up of stalactites that point in toward the center of the cave. The longest stalactites determine the minimum distance. If there is only a tiny number of such long stalactites, they are relatively unlikely to cause errors when the bat flies beyond the safe distance. It will collide most frequently with more distant (shortest) stalactites, owing to their greater number. So the take-home message is that a given code must be able to correct only "typical" errors.

C. Decoherence-free subspaces

Since the idea of i.i.d. errors was underlying the standard theory of quantum error-correcting codes, it is natural to expect lowered performance when employed on memory channels.

A case study is provided by a regular Markovian channel (see Sec. III.D.4), where the CPTP maps $\Phi_{q_j}^{(i_j)}$ in Eq. (95) are of the form $\Phi_{q_j}^{(i_j)}(\cdots) = \Phi^{(i)}(\cdots) = K_i(\cdots)K_i^\dagger$, with unitary Kraus operators K_i and $p_{i_j|i_{j-1}}^{(j)} = p_{i|i'} = (1-\mu)p_i + \mu\delta_{i,i'}$. A sequence of n uses of the memory channel is hence represented by the map

$$\Phi_\mu(\rho_Q^{(n)}) = \sum_{i_1,\ldots,i_n} p_{i_n|i_{n-1}} p_{i_{n-1}|i_{n-2}} \cdots p_{i_2|i_1} p_{i_1} \times (K_{i_n} \otimes \cdots \otimes K_{i_1}) \rho_Q^{(n)} (K_{i_n} \otimes \cdots \otimes K_{i_1})^\dagger. \tag{124}$$

The correlation parameter μ roughly quantifies the degree of memory of the considered channel. For $\mu = 0$ one obtains the case of i.i.d. (memoryless) noise, while the limit $\mu = 1$ describes completely correlated errors.

Cafaro and Mancini (2010a, 2010b) showed that the performance of stabilizer codes [evaluated by means of entanglement fidelity (54)] is lowered by increasing μ. The same relation between the fidelity and the correlation parameter has been observed for a model of long-term memory channel (see Sec. III.D.8) obtained by a convex combination of uncorrelated and completely correlated quantum channels

$$\Phi(\rho) = (1-\mu)\Phi_{\mu=0}(\rho) + \mu\Phi_{\mu=1}(\rho), \tag{125}$$

where $\Phi_{\mu=0}$ and $\Phi_{\mu=1}$ are given by Eq. (124) in the limiting cases of $\mu = 0$ and $\mu = 1$ (corresponding to uncorrelated and completely correlated errors, respectively). Actually, the effect of the memory is to take the code probability of error back to a linear dependence on the single-error probability [see also (Klesse and Frank (2005)].[5]

[5]It should be noted, however, that the condition of independent errors is not equivalent to the condition of memoryless quantum channel. Indeed, one can say that independent errors are those for which the probability of k errors is of the order of ϵ^k, given that the single error has a (small enough) probability ϵ. On the other hand, in the setting of memoryless channels each qubit independently interacts with its own environment, and different environments do not interact among themselves. While memoryless channels give rise to independent errors, the converse is not necessarily true. In fact, there are situations where although qubits do not interact independently with their environments, the generated errors still satisfy the independence condition, provided that qubits do not directly interact with each other (Hwang, Ahn, and Hwang, 2001; D'Arrigo, Benenti, and Falci, 2009). Thus, in such cases standard QECCs work well enough.

On the other hand, a suitable strategy to deal with completely correlated errors is represented by noiseless codes, also known as decoherence-free subspaces (DFSs). This is a "passive" quantum error correction method where the key idea is that of avoiding decoherence by encoding quantum information into special subspaces that are protected from the interaction with the environment by virtue of some specific dynamical symmetry (Palma, Suominen, and Ekert, 1996; Duan and Guo, 1997; Zanardi and Rasetti, 1997; Lidar, Chuang, and Whaley, 1998; Lidar and Whaley, 2003).

It turns out that a subspace $\mathcal{Q}$ is a DFS if and only if all Kraus operators, when restricted to $\mathcal{Q}$, are equal, up to a multiplicative constant, to a given unitary transformation $U_{\mathcal{Q}}$. In the case of imperfect initialization, i.e., a state not initialized inside a DFS, in a suitable basis the Kraus operators are described by a matrix of the form

$$\mathsf{K}_{\mathsf{k}} = \begin{pmatrix} s_{\mathsf{k}}\mathsf{U}_{\mathcal{Q}} & 0 \\ 0 & \mathsf{M}_{\mathsf{k}} \end{pmatrix}, \tag{126}$$

where M_{k} is an arbitrary matrix that acts on the orthogonal complement $\mathcal{Q}^{\perp}$ and may cause decoherence there (Shabani and Lidar, 2005). Equation (126) implies

$$\mathsf{K}_{\mathsf{k}}^{\dagger}\mathsf{K}_{\mathsf{l}} = \begin{pmatrix} \mathsf{S}_{\mathsf{kl}}\mathsf{I}_{\mathcal{Q}} & 0 \\ 0 & \mathsf{M}_{\mathsf{k}}^{\dagger}\mathsf{M}_{\mathsf{l}} \end{pmatrix}, \tag{127}$$

where $\mathsf{S}_{\mathsf{kl}} = \bar{s}_{\mathsf{k}} s_{\mathsf{l}}$. Applying Eq. (107) to the present setting, it follows that DFSs can be viewed as a special class of QECCs, where upon restriction to the code space $\mathcal{Q}$ all Kraus operators are proportional to the unitary $U_{\mathcal{Q}}$. It is worth noting that in the DFSs case the matrix S has rank 1. Hence, a DFS is an example of a maximally degenerate quantum error-correcting code. In the case of perfect DFS encoding, the necessary and sufficiency conditions are less restrictive than Eq. (126) (Shabani and Lidar, 2005).

An example of the effective application of DFS encoding can be obtained for the case of the "completely correlated" channel (see Sec. III.D.8). By putting $n = 2$, $K_0 = \mathbb{1}$, $K_1 = \sigma_z$, and $p_0 = 1 - p$, $p_1 = p$ in Eq. (124) one obtains a quantum channel with Kraus operators

$$K_{00} = \sqrt{p}\,\mathbb{1} \otimes \mathbb{1}, \qquad K_{01} = 0,$$
$$K_{10} = 0, \qquad K_{11} = \sqrt{(1-p)}\sigma_z \otimes \sigma_z. \tag{128}$$

A DFS is given by $\mathrm{span}\{|01\rangle, |10\rangle\}$ where one can safely encode a qubit

$$|0_L\rangle = |01\rangle, \qquad |1_L\rangle = |10\rangle. \tag{129}$$

By extending this argument, one can say that in the case of completely correlated errors it is possible to exploit the invariance of a subspace to encode information reliably.

Chiribella *et al.* (2011) provided a generalized quantum Hamming bound for nondegenerate codes, which depends on the rank of the CJ state (see Sec. II) associated with the noise process and holds for any kind of (possibly correlated) channel model. The original Hamming bound (109), which was formulated for the case of independent noise on the encoding systems is then recovered as a particular case. On the other hand, for completely correlated noise it was shown how to exploit degeneracy to violate the generalized quantum Hamming bound and achieve perfect quantum error correction with fewer resources than those needed for nondegenerate codes. As an example consider the following channel:

$$\Phi(\rho_Q^{(n)}) = p\rho + \sum_{i=1,\ldots,n,j>i} (p_{X,ij}\sigma_{x,i}\sigma_{x,j}\rho_Q^{(n)}\sigma_{x,i}\sigma_{x,j} + p_{Y,ij}\sigma_{y,i}\sigma_{y,j}\rho\sigma_{y,i}\sigma_{y,j} + p_{Z,ij}\sigma_{z,i}\sigma_{z,j}\rho\sigma_{z,i}\sigma_{z,j}), \tag{130}$$

where the input state is left unchanged with probability $p = 1 - \sum_{i=1,\ldots,n,j>i}(p_{X,ij} + p_{Y,ij} + p_{Z,ij})$, while it undergoes Pauli errors σ_x, σ_y, and σ_z on qubits i and j with probabilities $p_{X,ij}$, $p_{X,ij}$, and $p_{Z,ij}$, respectively. By evaluating the rank of the CJ state associated with the map (130) one obtains the following generalized quantum Hamming bound of Chiribella *et al.* (2011):

$$2^k\left[1 + 3\binom{n}{2}\right] \le 2^n. \tag{131}$$

Then by considering for instance $k = 1$, one gets $n = 7$ as the smallest integer satisfying the bound. However, one can also construct codes with lower values for n. For instance, this is the case of the code

$$|0_L\rangle = |000\rangle, \qquad |1_L\rangle = |111\rangle. \tag{132}$$

Note that these basis code words are not affected by the action of σ_z on any pair of qubits. Consequently, the action of σ_x on a pair of qubits is identical to the action of σ_y on the same pair of qubits. In other words the code is degenerate. Therefore, it is sufficient to consider only the errors due to σ_x operators. This can be realized through a projective measurement onto the subspaces $\mathcal{S}_{00} = \mathrm{span}\{|000\rangle, |111\rangle\}$, $\mathcal{S}_{01} = \mathrm{span}\{|100\rangle, |011\rangle\}$, $\mathcal{S}_{10} = \mathrm{span}\{|010\rangle, |101\rangle\}$, and $\mathcal{S}_{11} = \mathrm{span}\{|001\rangle, |110\rangle\}$. If the measurement outcome is "00," no errors have affected the qubits; on the contrary, if the measurement outcome is "01," errors have affected qubits 2 and 3 and can be corrected by applying there σ_x. Similarly, all other possible errors can be detected and corrected. It is hence clear that this code violates the quantum Hamming bound (131) thanks to the invariance of the coding subspace under the action of the pair of σ_z which allows for perfect error correction.

Having seen that DFSs are suitable to encode information in the presence of completely correlated errors, it is natural to expect that their performance decreases by reducing the degree of errors' correlation (Demkowicz-Dobrzanski, Kolenderski, and Banaszek, 2007). Actually Cafaro and Mancini (2010a, 2011) confirmed this fact for the Markovian model of Eq. (124) and for the model of Eq. (125), respectively. Then one can argue that for the memory channel models (124) and (125), where the memory effects are described by a single parameter μ, there must be a threshold value $\mu^\star$ that allows one to select the best code between the standard and the noiseless ones (D'Arrigo *et al.*, 2008; Cafaro and Mancini, 2010a; Cafaro and Mancini, 2011).

D. Designing quantum codes for correlated errors

The specific features of error models can be used to design new quantum codes that better cope with correlated errors.

The results of Sec. IV.C suggest that it might be convenient to concatenate decoherence-free subspaces with standard quantum error-correcting codes in order to achieve higher entanglement fidelity values in both low and high correlation regimes. This kind of concatenation was first introduced by Lidar, Bacon, and Whaley (1999), and it was investigated in the context of memory channels in Clemens, Siddiqui, and Gea-Banacloche (2004) and subsequently in Cafaro and Mancini (2011).

As an illustrative example, consider encoding one logical qubit into a decoherence-free subspace ($\mathcal{Q}_{\mathrm{DFS}} = \mathcal{Q}_{\mathrm{outer}}$) spanned by the basis code words

$$|0_L\rangle = |+-\rangle, \qquad |1_L\rangle = |-+\rangle, \tag{133}$$

and then encode each qubit of this basis into a three-qubit bit repetition code (132) ($\mathcal{Q}_{\mathrm{bit}} = \mathcal{Q}_{\mathrm{inner}}$). One obtains the fact that the basis code words of the concatenated code $\mathcal{Q} = \mathcal{Q}_{\mathrm{bit}} \circ \mathcal{Q}_{\mathrm{DFS}}$ are given by

$$|0_L\rangle = \tfrac{1}{2}(|000\,000\rangle - |000\,111\rangle + |111\,000\rangle - |111\,111\rangle),$$
$$|1_L\rangle = \tfrac{1}{2}(|000\,000\rangle + |000\,111\rangle - |111\,000\rangle - |111\,111\rangle).$$

The entanglement fidelity for the concatenation of a repetition code and a noiseless code for the models of Eqs. (124) and (125) with $K_0 = \mathbb{1}$, $K_1 = \sigma_x$, $p_0 = 1 - p$, and $p_1 = p$ is reported in Fig. 16. It turns out that in the first case the concatenated code does not work well for partially correlated errors. It is always better to use either the outer or the inner code alone depending on whether one is below or above the threshold value $\mu^\star(p)$. On the contrary, in the second case the concatenated code works optimally almost everywhere. Hence, one can argue that for the model of Eq. (125) the concatenation procedure is particularly advantageous in the presence of partially correlated errors.

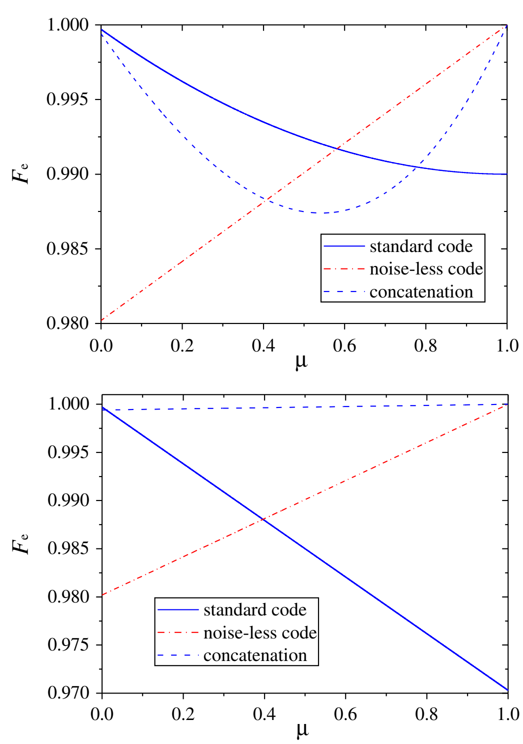


FIG. 16 (color online). Entanglement fidelity for the models (124) (top panel) and (125) (bottom panel) with $p = 10^{-2}$. Solid lines correspond to a standard code. Dot-dashed lines refer to a noiseless code. Dashed lines represent their concatenation.

Another error model often employed is that of burst errors. Such errors can be considered as affecting a sequence of qubits as opposed to random single qubit errors. They are well studied in the classical framework where corresponding error-correcting codes have been developed (Peterson and Weldon, 1972). Vatan, Roychowdhury, and Anantram (1997) considered a quantum analog of burst-error correcting codes. Hamming and Gilbert-Varshamov–type bounds have been derived showing that these codes are more efficient than codes protecting against random errors. In fact, to protect against burst errors of width b (that is, errors occurring on a number b of consecutive qubits with b a fixed constant), it is enough to map $n - \log_2 n - O(b)$ qubits to n qubits, while in the case of t random errors at least $n - t\log_2 n$ qubits should be mapped to n qubits.

A linear code $\mathcal{C}$ has burst-correcting ability b iff, for every burst w_1 and w_2 of width $\leq b$ it is $\mathsf{H}(\mathsf{w}_1 + \mathsf{w}_2)^\top \neq 0$, with H the parity check matrix of $\mathcal{C}$. Vatan, Roychowdhury, and Anantram (1997) presented an explicit construction of quantum codes for correcting burst errors starting from classical binary cyclic codes. A classical binary cyclic code $\mathcal{C}$ is such that if $(c^{(1)}, c^{(2)}, \ldots, c^{(n)})$ is in $\mathcal{C}$, then so is $(c^{(n)}, c^{(1)}, \ldots, c^{(n-1)})$ (Hill, 1985).

The definition of quantum burst-correcting codes straightforwardly follows from Eq. (116). Consider the set $\mathfrak{Q}$ of quantum errors (hence the corresponding set $\mathfrak{C}$ of classical error) such that both

$$\mathfrak{C}_X = \{\mathsf{e}_X \in \mathbb{F}_2^n | \exists \mathsf{e}_Z \in \mathbb{F}_2^n \Rightarrow (\mathsf{e}_X|\mathsf{e}_Z) \in \mathfrak{C}\}, \tag{134}$$

$$\mathfrak{C}_Z = \{\mathsf{e}_Z \in \mathbb{F}_2^n | \exists \mathsf{e}_X \in \mathbb{F}_2^n \Rightarrow (\mathsf{e}_X|\mathsf{e}_Z) \in \mathfrak{C}\} \tag{135}$$

are bursts of width $\leq b$. Then any quantum code $\mathcal{Q}$ having $\mathfrak{Q}$-correcting ability is called a b-burst quantum correcting code.

Now suppose having a $(3b+1)$-burst-correcting binary $[n{\cdot}k]$ cyclic code $\mathcal{C}$ then, to construct the b-burst-correcting $[\![n,k]\!]$ quantum code one can proceed as follows.

Let the $(n-k) \times n$ matrix $\tilde{\mathsf{H}}$ be a parity check matrix for the classical code $\mathcal{C}$. Let $\tilde{\mathsf{H}}_{\to m}$ denote the matrix that is obtained from $\tilde{\mathsf{H}}$ by cyclically shifting the columns m times to the right. Since $\mathcal{C}$ is cyclic, $\tilde{\mathsf{H}}_{\to m}$ is also a parity check matrix of $\mathcal{C}$. Consider the stabilizer quantum code $[\![n,k]\!]$ defined by the parity check matrix [see Eq. (119)]

$$\mathsf{H} = (\tilde{\mathsf{H}} + \tilde{\mathsf{H}}_{\to b} | \tilde{\mathsf{H}} + \tilde{\mathsf{H}}_{\to 2b+1}). \tag{136}$$

Let $\mathsf{e} = (\mathsf{e}_X|\mathsf{e}_Z)$ and $\mathsf{e}' = (\mathsf{e}'_X|\mathsf{e}'_Z)$ be bursts of width $\leq b$, with $\mathsf{e} \neq \mathsf{e}'$ and take

$$\mathsf{w} = \mathsf{e}_X + \mathsf{e}'_X + (\mathsf{e}_X + \mathsf{e}'_X)_{\to b} + \mathsf{e}_Z + \mathsf{e}'_Z + (\mathsf{e}_Z + \mathsf{e}'_Z)_{\to 2b+1}, \tag{137}$$

where $\mathsf{e}_{\to b}$ denotes the vector obtained by cyclically shifting e to the right b times. Then, it is easy to check that $\mathsf{w} \neq 0$ and w is the sum of two bursts of width $\leq 3b+1$. Hence $\mathsf{w} \notin \mathcal{C}$ and

$$\mathsf{H}(\mathsf{e} + \mathsf{e}')^\top = \tilde{\mathsf{H}}\mathsf{w}^\top \neq 0, \tag{138}$$

which guarantees the ability of the quantum code $[\![n,k]\!]$ to correct b-burst errors.

The existence of classical $(3b+1)$-burst-correcting binary cyclic codes with length $n = 2^m - 1$ and dimension $k = n - m - (3b+1)$ is known with m depending only on b (Peterson and Weldon, 1972). Hence these classical codes lead to almost optimal quantum codes [compared with the bound $n - \log_2 n - O(b)$ given above].

By increasing the length of the bursts, one should increase the length of the burst code as well. Alternatively it might be possible to resort to the *interleaving* technique. By using this method, the code words can be distributed among the qubit stream so that consecutive words are never next to each other. On deinterleaving they are returned to their original positions so that any errors that have occurred become widespread. This ensures that any burst (long) errors now appear as random (short) errors.

Classically the interleaving of m code words $(\mathsf{c}_1, \mathsf{c}_2, \ldots, \mathsf{c}_m)$ of an $[n,k]$ code is achieved by permuting the positions of bits in code words as follows:

$$\begin{aligned}(\mathsf{c}_1, \mathsf{c}_2, \ldots, \mathsf{c}_m) &= ((\mathsf{c}_1^{(1)}, \mathsf{c}_1^{(2)}, \ldots, \mathsf{c}_1^{(n)}), (\mathsf{c}_2^{(1)}, \mathsf{c}_2^{(2)}, \ldots, \mathsf{c}_2^{(n)}), \ldots, (\mathsf{c}_m^{(1)}, \mathsf{c}_m^{(2)}, \ldots, \mathsf{c}_m^{(n)})) \\ &\to ((\mathsf{c}_1^{(1)}, \mathsf{c}_2^{(1)}, \ldots, \mathsf{c}_m^{(1)}), (\mathsf{c}_1^{(2)}, \mathsf{c}_2^{(2)}, \ldots, \mathsf{c}_m^{(2)}), \ldots, (\mathsf{c}_1^{(n)}, \mathsf{c}_2^{(n)}, \ldots, \mathsf{c}_m^{(n)})).\end{aligned} \tag{139}$$

The procedure is equivalent to constructing the code as an $m \times n$ array where every row is a code word of the original $[n,k]$ code $(\mathsf{c}_1, \mathsf{c}_2, \ldots, \mathsf{c}_m)$. Now, a burst of length $\leq bm$ can have at most b symbols in any row of this array. Since each row can correct a burst of length $\leq b$, the code can correct all bursts of length $\leq bm$ (the parameter m is the interleaving degree).

Therefore, given an $[n,k]$ classical code correcting bursts of length $\leq b$, then interleaving this code to the degree m produces an $[nm, km]$ classical code correcting bursts of length $\leq bm$ (Peterson and Weldon, 1972).

Moving to the quantum framework, in order to interleave quantum codes, one needs to exchange the qubits one by one, following Eq. (139). Therefore, the basic step of the quantum interleaving simply is a swapping operation between two qubits. Then the classical result can be extended as follows (Kawataba, 2000): interleaving an $[\![n,k]\!]$ quantum code correcting bursts of length $\leq b$ to the degree m produces an $[\![nm, km]\!]$ quantum code correcting bursts of length $\leq bm$.

E. Convolutional codes

It is possible to extend the notion of stabilizer codes introduced in Sec. IV.A to codes that allow for an overlap between the individual steps of the encoding operation (Chau, 1998, 1999; Ollivier and Tillich, 2003, 2004). From classical coding theory, codes with these properties are called convolutional codes. Although not specifically designed for memory channels, they are intimately related to them.

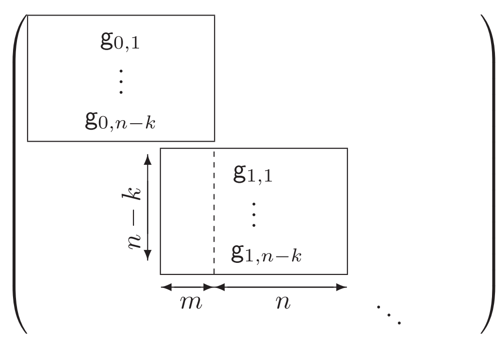
A quantum convolutional stabilizer code is defined by the generators of its stabilizer group just like a block stabilizer code [see Eq. (117)]. Consider a convolutional code encoding k logical qubits per n physical qubits, such that every block has an output of $n+m$ qubits. n of those are output qubits, while m are passed on to the next step. Then, an $[\![n,k]\!]$ m-convolutional stabilizer code is given by the Abelian stabilizer group

$$\mathcal{G} = \mathrm{span}\{g_{j,i} = \mathbb{1}^{\otimes jn} \otimes g_{0,i} | 1 \leq i \leq n-k, 0 \leq j\}, \tag{140}$$

where $g_{0,i} \in \mathcal{P}_{n+m}$ and all $g_{i,j}$ are independent of each other.

Note that the total number of physical qubits (length of the code) is left unspecified. Actually it is useful to set it to infinity by considering a Pauli group $\mathcal{P}_\infty$ with elements defined on a semi-infinite chain of qubits, but acting nontrivially only on a bounded number of them (as it happens with quasilocal algebra; see Appendix B). Then the generators are considered to be padded from the right with identities $\mathbb{1}$.

The structure of the stabilizer group generators can be summarized, following Eq. (119), by a semi-infinite matrix

$\mathsf{H} =$  . (141)

Each line of the matrix represents one of the $g_{j,i}$ and each column a different qubit. Thus, any given entry of H is a Pauli matrix for the corresponding qubit and generator. The rectangles represent which qubits are potentially affected by the action of the generators. Clearly, when $m = 0$ one has each block separately and obtains a block code.

Actually, it is possible to use the invariance by n qubit translation of the generators to find a shorter description. One defines the shift (delay) operator D acting on any element $A \in \mathcal{P}_\infty$ by

$$D[A] = \mathbb{1}^{\otimes n} \otimes A. \tag{142}$$

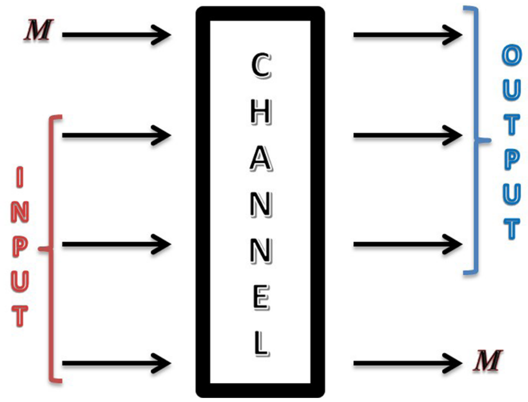


FIG. 17 (color online). The channel encoding the $[\![3,1]\!]$ 1-convolutional stabilizer code. The first input qubit is the memory input, and the last output qubit is the memory output.

Then the generators of the code can be written as

$$g_{j,i} = D^j[g_{0,i}], \qquad 0 \le j, \qquad 1 \le i \le n-k. \tag{143}$$

Using this, one needs only to consider the first $n-k$ generators. All others are obtainable by repeated applications of D.

In addition to applying D^j to an element of the Pauli group A (with bounded support), it is also possible to consider a polynomial $P(D) = \sum_j \alpha_j D^j$ and apply it as $P(D)[A] = \prod_j \alpha_j D^j[A]$. That critically relies on the fact that all copies of A shifted by D^j commute [see, e.g., Ollivier and Tillich (2004)].

It is worth noting that the encoding operations for convolutional codes can be described as quantum memory channels, due to the fact that some of the output qubits of the nth encoding step will be used as inputs in the $(n+1)$th step of the encoding; thus the blocks overlap. These qubits correspond to the memory system of the memory channel describing the encoding map. The encoding operation is the same in every step (neglecting initialization and finalization); thus every step is described by the same channel (see Fig. 17). This can be stated more precisely saying that for every $[\![n,k]\!]$ m-convolutional stabilizer code one can find an encoding operation which is described by a concatenation of a Weyl covariant memory channel with unimodular characteristic function (see Sec. II.I.3). The channel has n input and output qubits and uses m qubits of memory. One use of the channel corresponds to one block in the encoding (Gütschow, 2010).

Unfortunately convolutional codes also carry disadvantages. Because information is transmitted from one block to the next, errors can spread as well. Depending on the encoding algorithm errors that spread without bound on the output side can occur. These are called *catastrophic* errors and have to be avoided by employing noncatastrophic convolutional codes (Grassl and Roetteler, 2006). Necessary and sufficient conditions for an encoder to be noncatastrophic are provided by Poulin, Tillich, and Ollivier (2009) and the minimal amount of resources to satisfy them were determined by Houshmand, Hosseini-Khayat, and Wilde (2013). Furthermore, an attempt to relate such conditions to the property of strict forgetfulness of the memory channel representing the encoder was made by Gütschow (2010). Lacking the boundaries between code blocks, convolutional codes exhibit the same "continuous structure" as channels with memories. As such, they could result particularly suited to protect from correlated noise.

V. CAPACITIES OF QUANTUM CHANNELS

A natural question that arises after having examined correcting codes is what are the maximum communication rates achievable in quantum channels? For classical channels the highest rate (number of bits per channel uses) of reliable information transmission attainable via the application of encoding and decoding error-correcting procedures defines the capacity (Gallager, 1968; Cover and Thomas, 1991). In this context reliability refers to the requirement that the transferred messages have to be received without possibility of misunderstanding, i.e., the communication errors have to be removed by the selected coding strategy. One speaks of zero-error capacity when imposing this constraint for codes of finite length (i.e., codes which operate on a finite number of information carries or channel uses) (Shannon, 1956; Körner and Orlitsky, 1998). However, in many cases of physical and technological interest, it is more reasonable and mathematically more convenient to enforce such a condition only in the asymptotic limit of infinitely long messages. Under this paradigm, in fact, explicit expressions for the channel capacity are available as a function of the noise model which is tampering the communication line. For instance, in the case of a memoryless classical channel characterized by the conditional probability $p(y|x)$ of producing the output symbol y when fed with input x, the associated capacity can be expressed as (Shannon, 1948)

$$C_{\rm SH} = \max_{p(x)} I(X:Y), \tag{144}$$

where the maximization is performed over all probability distributions on x, and where $I(X:Y)$ is the corresponding Shannon mutual information; see Sec. II.K. The proof leading to Eq. (144) relies on the notion of typical sequences (Gallager, 1968; Cover and Thomas, 1991) and it does not provide an explicit recipe for determining the optimal coding and decoding strategies; this is why error-correcting codes and capacities are often treated as distinct subjects with no exception for the quantum realm, except a few notable exceptions, e.g., polar codes (Arikan, 2009). Yet Shannon's result establishes a fundamental benchmark that is useful to test the effectiveness of any coding procedure; an informal and clear introduction to these topics can be found in Preskill (1998) or Galindo and Martín-Delgado (2002). Strong versions of the converse Shannon theorem have been proved (Wolfowitz, 1964; Arimoto, 1973), which establish that if the rate of communication of a (memoryless) classical channel exceeds $C_{\rm SH}$, then the error probability of any coding scheme converges to one in the limit of many channel uses.

The notion of capacity based on asymptotic reliability has also an important operational meaning stated by the reverse Shannon theorem, a result which was only formulated and proved only recently within the context of quantum communication (Bennett, Shor et al., 1999; Bennett et al., 2002; Cuff, 2008). According to it, for any classical noisy channel of capacity $C_{\rm SH}$, if the sender and receiver share an unlimited supply of random bits, an expected $nC_{\rm SH} + o(n)$ uses of a noiseless binary channel are sufficient to exactly simulate n uses of the original channel.

As anticipated in Sec. I, the generalization of the above ideas to the quantum setting leads to the introduction of a plethora of channel capacities, depending on whether classical or quantum information has to be transmitted, and whether additional resources, as preshared entanglement, are exploited. Unification of these quantities under a common formalism based on resource inequalities was presented by Abeyesinghe and Hayden (2003), Devetak, Harrow, and Winter (2004, 2008), Abeyesinghe et al. (2009), and Hsieh and Wilde (2010a, 2010b). Here we do not report this approach; instead we focus on clarifying the operational definitions of these quantities in a framework which does not make explicit reference to the structure of the communication line. Then coding theorems will be reviewed, which allow one to express the capacities in terms of suitable entropic quantities, starting from the case of memoryless channels (the best understood and characterized so far) and then moving on to the more complex scenario of memory channels.

A. Operational definitions

1. Sending bits or qubits on a quantum channel

The classical (respectively, quantum) capacity C (respectively, Q) of a quantum channel defined by the CPTP maps $\Phi^{(n)}$ of Eq. (78) is the maximum rate R at which classical (respectively, quantum) information, encoded on a set of quantum carriers, can be sent reliably from the sender Alice to the receiver Bob (Shor, 1995); see Fig. 18. As in the classical setting (Shannon, 1948) the rate is measured as the ratio $R = k/n$ among the number k of bits (respectively, qubits) transmitted and the number n of carriers employed (the "redundancy" of the code according to Sec. IV). Similarly the reliability condition is introduced by requiring that in the asymptotic limit of $k \to \infty$ the error probability of the procedure can be made arbitrarily small (or, equivalently, the fidelity of the transmission will approach unity), while keeping R constant. In view of these operational definitions, C and Q can be expressed as the following limit (Bennett and Shor, 1998; Bennett et al., 2002):

$$\lim_{\epsilon\to 0}\lim_{k\to\infty}\sup\left\{\frac{k}{n}:\ \exists\Phi_E^{(k\to n)},\exists\Phi_D^{(n\to k)},\min_{m\in\mathcal{M}}F(|m\rangle;\Phi_D^{(n\to k)}\circ\Phi^{(n)}\circ\Phi_E^{(k\to n)})>1-\epsilon\right\}, \tag{145}$$

where, analogously to the notation introduced at the beginning of Sec. IV and as shown in Fig. 18, $\Phi_E^{(k\to n)}$ and $\Phi_D^{(n\to k)}$ are, respectively, encoding and decoding channels mapping elements $|m\rangle$ from a reference input space $\mathcal{M}$ (the messages Alice wishes to send to Bob) to states $\Phi_E^{(k\to n)}(m) \coloneqq \Phi_E^{(k\to n)}(|m\rangle\langle m|)$ of n carriers (the code words of the procedure), and F is the input-output fidelity function introduced in Sec. II.J.1. In particular, to fix the units properly, the expression for C is obtained by taking $\mathcal{M}$ to be a collection of 2^k orthogonal vectors which, without loss of generality can be identified with the elements $\{|0\rangle,|1\rangle\}^{\otimes k}$ of the computational basis of k qubits. On the other hand, for the quantum capacity Q the set $\mathcal{M}$ coincides with the whole $\mathbb{C}^{2\otimes k}$ (as the latter includes the elements of the canonical basis, it trivially follows that for a given communication line one has $Q \le C$). The limits in Eq. (145) are finally computed by first taking a supremum limit in $k \to \infty$ (which always exists) and then sending the error parameter ϵ to zero to enforce the transmission fidelity to approach unity for all input messages. In Eq. (145) this is explicitly enforced by requiring $1-\epsilon$ to lower bound the minimum value achieved on $\mathcal{M}$ by the transmission fidelity. Such a rather strong requirement however can be relaxed by replacing it with a similar constraint that applies only on the average transmission fidelity: this does not affect the limit in Eq. (145) and hence the definitions of either C or Q (Keyl, 2002; Kretschmann and Werner, 2004). Similarly the same value of Q one gets from Eq. (145) can also be obtained by substituting the minimum in F with the entanglement fidelity introduced in Eq. (52) (Barnum, Knill, and Nielsen, 2000; Kretschmann and Werner, 2004).[6]

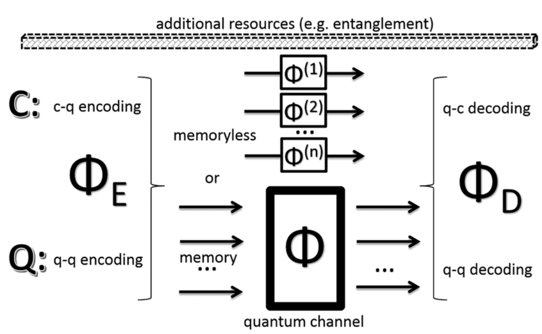


FIG. 18. Classical (C) and quantum (Q) capacities of a (memoryless or memory) quantum channel Φ in terms of all possible classical or quantum encoding and decoding schemes, with the potential use of additional resources as shared entanglement.

[6]All definitions introduced so far assume a notion of capacity in which the error probability is required to nullify only in the asymptotic limit of large enough k. As in classical communication theory (Shannon, 1956; Körner and Orlitsky, 1998), however, a more stringent reliability requirement can be enforced, i.e., imposing that the min-fidelity on $\mathcal{M}$ should equal 1 for a finite number n of channel uses. Under this condition one is led to the definition of zero-error classical and quantum capacities (Medeiros and de Assis, 2005). This corresponds to the maximal communication rate achievable by perfect codes, as introduced in Sec. IV.

Finally, one observes that the definition (145) yields a natural data-processing inequality for the capacities. For instance, given C the classical capacity of a quantum channel whose CPTP mapping $\Phi^{(n)} = \Phi'^{(n)} \circ \Phi''^{(n)}$ is obtained by concatenating the other two CPTP maps, one has

$$C \le \min\{C', C''\}, \tag{146}$$

where C' and C'' are, respectively, the classical capacities of channels described by $\Phi'^{(n)}$ and $\Phi''^{(n)}$ [the same relation also applies for the quantum capacities Q, Q', and Q'' as well as for all the other capacities reviewed in Sec. V.A with the notable exception of those discussed in Sec. V.A.4, where the constraints may introduce spurious effects in the optimization, see, e.g., Giovannetti and Mancini (2005)]. The proof of this rather intuitive fact follows by observing that when passing from C to C' one can interpret $\Phi''^{(n)}$ as part of the decoding procedure of map $\Phi'^{(n)}$. Since the C' is obtained by optimizing the transmission rate with respect to all possible decodings, including those that do not use $\Phi''^{(n)}$ as a preliminary stage, it follows that C' is certainly not smaller than C. Similarly when passing from C to C'', one can interpret $\Phi'^{(n)}$ as part of the encoding stage for $\Phi''^{(n)}$: again since C'' is the optimal rate with respect to all encoding maps one has that it is certainly not smaller than C. An application of the above analysis to two channels $\Phi_1^{(n)}$ and $\Phi_2^{(n)}$ which are unitarily equivalent Eq. (5) shows that they must possess the same capacities: in this case indeed the CPTP concatenation which links the two maps can always be reversed, producing both the inequality $C(\Phi_1^{(n)}) \le C(\Phi_2^{(n)})$ and its counterpart $C(\Phi_2^{(n)}) \le C(\Phi_1^{(n)})$.

2. Capacities assisted by ancillary resources

Entanglement is a fundamental resource in quantum information theory. In the context of quantum communication this fact is testified by the teleportation (Bennett *et al.*, 1993) and superdense coding (Bennett and Wiesner, 1992) protocols. The former provides a nontrivial way of transmitting arbitrary quantum states when quantum carriers are not available but only bits can be exchanged through a classical communication line. The latter instead, in the presence of a noiseless quantum communication line, allows one to send 2 bits of classical information per transferred physical qubit. The necessary additional resource for both procedures is a shared entangled state between the sender Alice and receiver Bob. Application of these ideas to noisy communication lines introduces the notion of entanglement-assisted classical capacity C_{ea} (respectively, quantum capacity Q_{ea}) of a quantum channel Φ (Bennett, Shor *et al.*, 1999; Bennett *et al.*, 2002). Operationally they are defined as the maximum rate of reliable transmission of classical (respectively, quantum) information when the sender and the receiver have at their disposal an unbounded number of preshared maximally entangled states as ancillary side resources. Formal expressions are hence obtained through the same limit given in Eq. (145) with the difference that now the transformations $\Phi_E^{(k\to n)}$ and $\Phi_D^{(n\to k)}$ map elements of the reference input space $\mathcal{M}$ to and from the joint space associated with the n carriers of the channel plus the local quantum memories where Alice and Bob are storing their prior entangled states (specifically $\Phi_E^{(k\to n)}$ acts on Alice's memories, and $\Phi_D^{(n\to k)}$ on Bob's memories). While by definition C_{ea} and Q_{ea} provide natural upper bounds for the unassisted counterparts C and Q, respectively, a direct application of the teleportation and superdense coding protocol shows that for any given quantum channel they are related by the identity $C_{ea} = 2Q_{ea}$ (Bennett, Shor *et al.*, 1999).

Unlimited classical communication between Alice and Bob is another example of an ancillary resource which is known to increase the ability of transferring arbitrary quantum states over a quantum channel via the application of entanglement distillation protocols—admitting only forward classical communication from Alice to Bob is instead of no use in this respect (Bennett *et al.*, 1996). This yields the notion of the (two-way) classical assisted quantum capacity Q_2 (Bennett, Shor *et al.*, 1999; Bennett *et al.*, 2002) which again can be formally expressed as in Eq. (145) by replacing the concatenation $\Phi_D^{(n\to k)} \circ \Phi^{(n)} \circ \Phi_E^{(k\to n)}$ with an arbitrary LOCC process intermediated by the action of the channel (see Sec. II.H). Quantum capacity assisted by providing access to nontrivial zero quantum capacity side channels was analyzed by Smith, Smolin, and Winter (2008).

Both the classical and quantum capacities of a channel can finally also be improved by allowing a feedback communication line (either quantum or classical) which permits the receiver Bob of the messages to signal to the sender Alice. Interestingly enough it has been shown that using feedback in the presence of prior shared entanglement is of no use, i.e., $C_{ea+FB} = C_{ea}$ and $Q_{ea+\mathrm{FB}} = Q_{ea}$ (Bowen, 2004, 2005).

A partial ordering among some of the quantities introduced in this section is provided in Bennett *et al.* (2006).

3. Private classical capacity of a quantum channel

The private classical capacity C_p of a quantum channel is defined as the maximum rate at which classical information can be transmitted privately from the sender to the receiver. Formally this is enforced by requiring that a third party (Eve) who has access to the channel environment, and who is trying to recover Alice messages to Bob, will get it with an error probability that is approaching unity in the asymptotic limit of infinitely long messages (Cai, Winter, and Yeung, 2004; Devetak, 2005). Again C_p can be expressed as the limit in Eq. (145) by further constraining the coding and decoding procedures to satisfy an entropic inequality that implement the privacy requirement. Specifically this is obtained by upper bounding with ϵ the Holevo information (71) of the complementary channel (19) of $\Phi^{(n)}$ and associated with the uniform ensemble $\mathcal{E} = \{p_m = 2^{-k}, \Phi_E^{(k\to n)}(m)\}_{m\in\mathcal{M}}$ generated by the encoding mapping selected by Alice. As the complementary map is the transformation that links the channel inputs to the images they produce on the environment (see Sec. II.G), this choice, via the Holevo bound (72), ensures that the information Eve can recover on Alice's messages vanishes when taking the limit $\epsilon \to 0$. By construction C_p is always smaller or equal to the corresponding C and greater or equal to Q, i.e.,

$$C \ge C_p \ge Q \tag{147}$$

(the last inequality is associated with the fact that the ability of sending all vectors of $\mathbb{C}^{2\otimes k}$ with unit fidelity ensures that no information on the transferred states is passing from Alice to the environment).

4. Constrained capacities

The operational definitions of capacities can be modified in order to account for possible constraints on the input (or output) states on the channel. For instance, three weaker versions of the classical capacity of a quantum channel have been identified (Bennett and Shor, 1998). Specifically one defines the product-state (or classical-quantum or Holevo) classical capacity C_{cq} (which following a rather universal convention hereafter is indicated with the symbol C_1) by requiring that the employed coding maps entering in Eq. (145) produce only separable code words, that is, $\Phi_E^{(k\to n)}(m)$ is a separable state of the n carriers for all messages $|m\rangle$ in $\mathcal{M}$ [one notes incidentally that the analogous of C_1 for the classical private capacity C_p, i.e., the product-state private classical capacity $C_{p,1}$, has been defined by Devetak (2005); see Sec. V.B.2]. Similarly one defines a quantum-classical capacity $C_{\rm qc}$ by leaving the encoding channel unconstrained but imposing the decoding channels $\Phi_D^{(n\to k)}$ to be LOCC with respect to the outputs of different uses of the channel. Finally assuming LOCC operations for $\Phi_D^{(n\to k)}$ and separability for the code words $\Phi_E^{(k\to n)}(m)$ one defines the classical-classical capacity $C_{\rm cc}$. The unconstrained capacity C (often identified in this context also as the quantum-quantum or $C_{\rm qq}$ capacity) is a natural [and in general strict (Hastings, 2009)] upper bound for the others. Similarly, for any assigned quantum channel, $C_{\rm cc}$ is a natural lower bound for $C_{\rm cq}$ and $C_{\rm qc}$, the ordering between the last two being at present unknown.

Of special interest is also a class of physically motivated constrained capacities obtained by introducing a family of observables $\{A^{(n)}\}_{n=1,\ldots,\infty}$ and by imposing that for any n the mean value of $A^{(n)}$ is bounded on the ensemble of states at the input of n uses of the quantum channel. The capacity of a quantum channel under such a constraint can be defined as in Eq. (145) under the additional requirement that for any k and n

$$\mathrm{Tr}(A^{(n)}\rho^{(n)}) \le a, \tag{148}$$

with $\rho^{(n)} = \Phi_E^{k\to n}(\mathbb{1}/2^k)$, $\mathbb{1}/2^k$ being the average state over the set $\mathcal{M}$. In particular, a relevant role is played by additive observables, for which one can put $A^{(n)} = n^{-1}\sum_{k=1}^n A_k$, where $A_k \equiv A$ is the observable for a single input quantum system at the input of the kth use of the channel.

The notion of constrained capacity naturally applies in the context of CV channels (see Sec. II.I.4). Indeed, due to the fact that the carrier Hilbert space is infinite dimensional it turns out that the capacity of a CV channel can be infinite (Holevo and Werner, 2001). In fact, an infinite value for the capacity corresponds to the encoding of information into larger and larger sectors of the Hilbert space. Clearly, that is in contradiction with the finiteness of the resources employed in physical realizations, e.g., the finiteness of the mean energy. It is hence meaningful to introduce a notion of capacity under a physically motivated constraint. Most natural choices are to impose a constraint on the mean value of the energy or the number of bosonic excitation per mode. In the latter case one has $A_n = n^{-1}\sum_{k=1}^n a_k^\dagger a_k$, where $\{a_k, a_k^\dagger\}$ are the canonical ladder operators at the channel input. From a technical point of view, these choices, besides being physically sound, guarantee that the set of states satisfying the constraint form a compact set. This is a crucial feature to ensure that the coding theorems for Gaussian channels under a constrained mean excitation number (or energy) yield expressions formally analogous to the unconstrained case, with the optimization being performed over input ensembles satisfying the constraint (Holevo, 1997, 2004; Holevo and Shirokov, 2006). Since the excitation number and the energy are quadratic in the canonical variables, their mean values can be expressed in terms of the first and second moments of the characteristic function of the input states; see Sec. II.I.4. In particular, the condition of having no more than N mean excitations per mode is expressed in terms of the first and second moments

$$\frac{\mathrm{Tr}(\mathsf{C}^{(n)}) + |\mathsf{m}|^2}{2n} \le N + \frac{1}{2}. \tag{149}$$

5. A superoperator norm approach to quantum capacities

The limit that defines C and Q in Eq. (145) indicates that for sufficiently large k there exists $\Phi_D^{(n\to k)}$ and $\Phi_E^{(k\to n)}$ which makes $\Phi_D^{(n\to k)}\circ\Phi^{(n)}\circ\Phi_E^{(k\to n)}$ close to the identity transformation $\mathrm{id}_\mathcal{M}$ on $\mathcal{M}$. Specifically for the quantum capacity Q, $\mathrm{id}_\mathcal{M}$ is the identity superoperator on $\mathbb{C}^{2\otimes k}$, while for the classical capacity C, the map $\mathrm{id}_\mathcal{M}$ is the fully dephasing channel on $\mathbb{C}^{2\otimes k}$ which leaves the elements of its computational basis $\{|0\rangle, |1\rangle\}^{\otimes k}$ unchanged.

Based on this observation a definition of capacities which is fully equivalent to the approach of Sec. V.A.1 can be given in terms of the cb-norm superoperator distance defined in Sec. II.J.2. In this approach (Keyl, 2002; Kretschmann, 2003), a positive quantity R is said to be an achievable rate for the channel Φ if for all sequences $\{k_i, n_i\}_{i\in\mathbb{N}}$ with $\lim_{i\to\infty} k_i = \infty$ and $\lim\sup_{i\to\infty} k_i/n_i < R$ one has

$$\lim_{i\to\infty}\inf_{\Phi_D,\Phi_E} |||\Phi_E^{*(k_i\to n_i)}\circ\Phi^{*(n_i)}\circ{\Phi_D}^{*(n_i\to k_i)} - \mathrm{id}^*_\mathcal{M}|||_{cb} = 0, \tag{150}$$

where $\Phi_E^{*(k_i\to n_i)}$, $\Phi^{*(n_i)}$, and ${\Phi_D}^{*(n_i\to k_i)}$ are the duals (10) of the maps $\Phi_E^{(k_i\to n_i)}$, $\Phi^{(n_i)}$, and ${\Phi_D}^{(n_i\to k_i)}$ defined in Eq. (145), while $\mathrm{id}^*_\mathcal{M}$ is the dual of the identity map on $\mathcal{M}$. With this prescription the values of Q and C are then identified as the supremum of the corresponding achievable rates. A similar construction was also presented by Holevo and Werner (2001) and Kretschmann and Werner (2004): here, however, the cb-norm distance was used directly in the Schrödinger channel representation.

B. Coding theorems for memoryless channels

Coding theorems provide expressions for the communication capacities of memoryless quantum channels (75) in terms of suitable entropic functions of the input and output states of the channel. While referring the interested reader to Winter (1999a), Holevo (2012), and Wilde (2013) for a detailed review of the subject, here the main results are reported concerning the entanglement-assisted classical capacity, the classical capacity and its product-state version, the private classical capacity, and the quantum capacity. Limitations and applicability of these expressions for the case of memory channels are discussed in Sec. V.C.

1. The Holevo-Schumacher-Westermoreland coding theorem

Preliminary attempts to compute the classical capacity of quantum channels were presented by Hausladen *et al.* (1995, 1996). A closed expression for the product-state classical capacity introduced in Sec. V.A.4 was finally provided by the Holevo-Schumacher-Westermoreland (HSW) coding theorem (Schumacher and Westmoreland, 1997; Holevo, 1998a). It mimics the Shannon formula (144) by establishing the fact that, for a memoryless channel Φ, $C_1(\Phi)$ can be expressed in terms of a maximization of the associated output Holevo information (71) over the set of input state ensembles $\mathcal{E} = \{p_j, \rho_j\}$ (possibly satisfying some additional input constraints), i.e.,

$$C_1(\Phi) = \max_{\mathcal{E}} \chi(\mathcal{E}; \Phi) \tag{151}$$

[owing to the concavity of von Neumann entropy (Wehrl, 1978; Petz, 2008), the maximization can, in fact, be always restricted to the ensemble of pure states]. Besides the original derivations (Schumacher and Westmoreland, 1997; Holevo, 1998a), several independent proofs of Eq. (151) are known (Holevo, 1998b; Ogawa and Nagaoka, 1999, 2002; Winter, 1999b; Hayashi and Nagaoka, 2003; Datta and Dorlas, 2007; Hayashi, 2007, 2009; Lloyd, Giovannetti, and Maccone, 2011; Sen, 2011; Giovannetti, Lloyd, and Maccone, 2012). As is typical with many coding theorems the general argument beyond the HSW result consists of two parts: (i) an inequality which establishes the fact that the rhs of Eq. (151) is an upper bound for the channel capacity (converse part of the theorem), and (ii) a direct part which proves the existence of a coding procedure that saturates such a bound asymptotically in the length of code. Part (i) can be established via the classical Fano inequality (Cover and Thomas, 1991) (relating the average information lost in a classical noisy channel to the error transmission probability) and the Holevo bound inequality (Holevo, 1973a, 1973b) on the achievable information of a quantum source; see Eq. (72). A sketch of this proof can be found in Appendix D. This approach is sufficient to show that any rate exceeding the capacity will necessarily produce an error probability which is nonzero even in the limit of infinitely many channel uses. It is worth noting however that, in contrast to classical information theory, establishing a strong version of the converse part of the theorem (i.e., proving that the transmission error probability will necessarily reach 1 as soon as the rate exceeds the capacity threshold) is particularly demanding in the quantum setting. In fact, strong converse coding theorems have been derived only for limited classes of finite dimensional channels (Ogawa and Nagaoka, 1999; Winter, 1999b; König and Wehner, 2009; Wilde, Winter, and Yang, 2013) while explicit counterexamples have been provided that show that in general they do not apply when considering continuous-variable systems (Wilde and Winter, 2013).

The direct part of the coding theorem which yields to Eq. (151) is based instead on the notion of typical subspace for quantum sources (Ohya and Petz, 1993; Schumacher, 1995). From this it follows that given an ensemble $\mathcal{E} = \{p_j, \rho_j\}$ and an integer number N fulfilling the condition $N \le 2^{n\chi(\mathcal{E};\Phi)}$, one can identify N code words $\rho_1^{(n)}, \ldots, \rho_N^{(n)}$ of the form $\rho_j^{(n)} = \rho_{j_1} \otimes \rho_{j_2} \otimes \cdots \rho_{j_n}$ and an associated POVM that allows Bob to discriminate among the output counterparts of the $\rho_j^{(n)}$ [i.e., the density matrices $\Phi(\rho_{j_1}) \otimes \Phi(\rho_{j_2}) \otimes \cdots \Phi(\rho_{j_n})$] with an error probability that can be bounded below any assigned threshold by sending n to infinity. Accordingly one is thus led to coding-decoding schemes which guarantee faithful transfer of classical messages at rates

$$R = \frac{\log_2 N}{n} \le \chi(\mathcal{E}; \Phi)$$

(the upper limit being attainable for large enough n), which approach Eq. (151) when taking the supremum over $\mathcal{E}$.

It is important to note that the POVM which comes with the proof of the HSW theorem is explicitly a joint one: this is the reason why the rhs of Eq. (151) coincides with $C_1(\Phi)$ and not with the $C_{\rm cc}(\Phi)$ capacity of Sec. V.A.4 (the latter indeed is the maximum rate attainable when allowing only LOCC operations among the various channel outputs). In fact, closed-form expressions for $C_{\rm cc}(\Phi)$ and for $C_{\rm qc}(\Phi)$ are at present still missing. An explicit formula for the unconstrained capacity $C(\Phi)$ defined in Sec. V.A instead can be obtained as a simple generalization of Eq. (151). This is done by adopting a block-coding strategy which, for all n, allows one to represent the density matrices produced by the (possibly nonseparable) encoding maps $\Phi_E^{(k\to n)}$ of Eq. (145) as tensor states over blocks of channels uses on which $\Phi^{\otimes n}$ acts as a single carrier map with an associated product-state capacity $C_1(\Phi^{\otimes n})$. The resulting capacity of Φ can then be obtained by taking the limit over n of the associated rates, i.e.,

$$C(\Phi) = \lim_{n\to\infty} \frac{1}{n} C_1(\Phi^{\otimes n}). \tag{152}$$

2. The private classical capacity theorem

The capacity formula for the private classical capacity $C_p(\Phi)$ introduced in Sec. V.A.3 was given by Cai, Winter, and Yeung (2004) and Devetak (2005). As for the HSW theorem discussed in Sec. V.B.1 it is derived by first providing a closed expression for its product-state version $C_{p,1}(\Phi)$ (i.e., the private classical capacity attainable by using only separable code words), and then using block coding to compute $C_p(\Phi)$ via the identity

$$C_p(\Phi) = \lim_{n\to\infty} \frac{1}{n} C_{p,1}(\Phi^{\otimes n}). \tag{153}$$

Analogously to Eq. (151), $C_{p,1}(\Phi)$ is a functional of the Holevo information (71). In this case, however, one has

$$C_{p,1}(\Phi) = \max_{\mathcal{E}} [\chi(\mathcal{E}; \Phi) - \chi(\mathcal{E}; \tilde{\Phi})], \tag{154}$$

where $\tilde{\Phi}$ is the complementary channel of Φ as defined in Sec. II.G.

The converse part of the proof which leads to Eq. (154) is obtained by joining the classical Fano inequality and the Holevo bound at the output of the map Φ with the privacy requirement imposed on the Holevo information of the complementary channel $\tilde{\Phi}$ of Φ. Vice versa the direct part of the coding theorem is based on the fact that, using a typical subspace argument, for each $\epsilon > 0$ and for each ensemble $\mathcal{E} = \{p_j, \rho_j\}$ satisfying $\chi(\mathcal{E}; \Phi) > \chi(\mathcal{E}; \tilde{\Phi})$, one can identify up to $N \simeq 2^{n\chi(\mathcal{E};\Phi)}$ states of the form $\rho_{j_1} \otimes \rho_{j_2} \otimes \cdots \otimes \rho_{j_N}$ which, when organized in groups of $M \simeq 2^{n\chi(\mathcal{E};\tilde{\Phi})}$ elements each, (i) can be faithfully discriminated on Bob's side, and (ii) fully overlap on Eve's side. Accordingly by assigning to each of such group the same classical message, Alice can now transfer up to $N/M \simeq 2^{n[\chi(\mathcal{E};\Phi) - \chi(\mathcal{E};\tilde{\Phi})]}$ distinct messages to Bob without Eve being able to read them.

3. The quantum capacity theorem

The expression for the quantum capacity $Q(\Phi)$ of a memoryless channel Φ is (Lloyd, 1997; Shor, 2002a; Devetak, 2005)

$$Q(\Phi) = \lim_{n\to\infty} \frac{1}{n} Q_1(\Phi^{\otimes n}), \tag{155}$$

where

$$Q_1(\Psi) = \max_{\rho} J(\rho; \Psi), \tag{156}$$

with $J(\rho; \Psi)$ the associated coherent information (68) and the maximization is over all input states; see also Hayden *et al.* (2008).

The converse part of the coding theorem can be obtained as an application of the quantum Fano inequality (65) which, when imposing a lower limit to the entanglement fidelity, forces the dimensionality of $\mathcal{M}$ to be bounded in terms of the channel coherent information (Barnum, Nielsen, and Schumacher, 1998; Barnum, Knill, and Nielsen, 2000) (in Appendix D.2 a detailed derivation of this relation is presented for the general case of non-necessarily memoryless channels). Following Devetak (2005), a relatively simple proof of the direct part of the theorem instead can be obtained as a modification of the coding theorem for the private classical capacity. The idea here is to extract from the codes which lead to privacy of the classical messages those that allow also the preservation of the coherent superpositions among the various code words. It turns out that this can be enforced by restricting the maximum in Eq. (154) to only those ensemble $\mathcal{E}$ formed by pure state elements: a condition which, thanks to Eq. (74), allows one to identify as achievable rates for quantum communication those obtained from Eq. (154) in which $\chi(\mathcal{E}; \Phi) - \chi(\mathcal{E}; \tilde{\Phi})$ gets replaced by $J(\rho; \Phi)$.

4. The Bennett-Shor-Smolin-Thapliyal theorem and the quantum reverse Shannon theorem

The entanglement-assisted classical capacity of a memoryless channel is given in terms of the quantum mutual information defined in Eq. (67), i.e.,

$$C_{ea}(\Phi) = \max_{\rho} I(\rho; \Phi), \tag{157}$$

where the maximization is over any input ρ [the corresponding quantum capacity version $Q_{ea}(\Phi)$ being half of $C_{ea}(\Phi)$ as already anticipated in Sec. V.A.2]. This result was proven by Bennett, Shor *et al.* (1999) and Bennett *et al.* (2002) by generalizing the dense-coding protocol (Bennett and Wiesner, 1992) to the case of noisy memoryless channel. In dense coding, the sender and the receiver share a maximally entangled state in a Hilbert space of finite dimension, say d^2. The sender encodes classical information by applying d^2 generalized d-dimensional Pauli unitaries to one-half of the maximally entangled states, which is then sent through the channel. These transformations map the given states into d^2 orthogonal states, which the receiver can reliably distinguish. More generally, Alice, the sender, may encode classical information by applying generic CPTP maps on her half of the maximally entangled state and then sending it through the channel. However, Bennett *et al.* (2002) proved that the use of encoding by generalized Pauli unitaries is optimal. This is obtained using the results of Schumacher and Westmoreland (1997) and Holevo (1998a) about encoding classical information into quantum states, which apply also in these settings [see also Hsieh, Devetak, and Winter (2008)]. It is worth noting that the encoding by generalized Pauli unitaries constrains ρ to be the maximally mixed state $\rho = \mathbb{1}/d$. This limitation is circumvented by considering the typical input states in the asymptotic limit of many uses of the channel, in which the average input state is always the maximally mixed one on the typical subspace. Remarkably, due to the subadditivity of the quantum mutual information, the expression for $C_{ea}(\Phi)$ of a memoryless channel does not require the regularization over the channel uses.

An important property of $C_{ea}(\Phi)$ is provided by the quantum reverse Shannon theorem (Bennett *et al.*, 2009; Berta, Christandl, and Renner, 2011; Berta, Renes, and Wilde, 2013; Berta *et al.*, 2013) which generalizes the reverse Shannon theorem discussed in the introductory paragraphs of Sec. V. It establishes the fact that the input-output mapping of n uses of a memoryless quantum channel Φ can be simulated using $nC_{ea}(\Phi) + o(n)$ uses of a noiseless (qubit or bit) channel by providing the sender and the receiver an unlimited supply of prior shared entanglement.

5. Superadditivity and superactivation

The expressions in Eqs. (152), (153), and (155) for the classical, private, and quantum capacity of a memoryless quantum channel require the computation of the regularized limit over the number of uses of the channel $n \to \infty$. Then one has the inequalities $C(\Phi) \geq C_1(\Phi)$, $C_p(\Phi) \geq C_{p,1}(\Phi)$, and $Q(\Phi) \geq Q_1(\Phi)$. For a given memoryless channel Φ, if the first inequality is strict, that is, $C(\Phi) > C_1(\Phi)$, one says that the Holevo information of Φ is superadditive; otherwise it is said to be additive. Similarly, one says that the coherent information is superadditive whenever $Q(\Phi) > Q_1(\Phi)$, and additive otherwise; see, e.g., Smith (2010).

At a higher level of complexity, when two different channels, Φ_1 and Φ_2, are used in parallel the inequalities $C(\Phi_1 \otimes \Phi_2) \geq C(\Phi_1) + C(\Phi_2)$, $C_p(\Phi_1 \otimes \Phi_2) \geq C_p(\Phi_1) + C_p(\Phi_2)$, and $Q(\Phi_1 \otimes \Phi_2) \geq Q(\Phi_1) + Q(\Phi_2)$ follow from

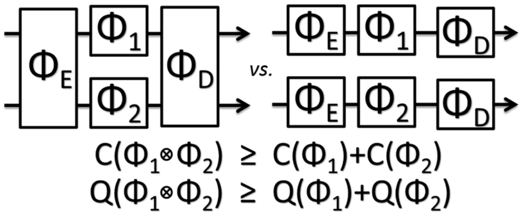


FIG. 19. According to the superadditivity property of channel capacities, the capability of transmitting classical or quantum information over a tensor product of two maps Φ_1 and Φ_2 is, in general, larger than the sum of the individual capacities. This can be intuitively understood by the fact that global encoding or decoding schemes allow one to explore a larger part of the Hilbert space, with respect to the local ones.

simple coding arguments [for instance, the rate $C(\Phi_1) + C(\Phi_2)$ can always be attained by feeding the inputs of Φ_1 and Φ_2 independently with their corresponding optimal codes]; — see Fig. 19. If it happens that one of these inequalities is strict, then one says that the classical (respectively, private, respectively, quantum) capacity is superadditive under the tensor product of the channels Φ_1 and Φ_2 [note that the additivity of (say) the Holevo capacity of Φ_1 and Φ_2 does not necessarily guarantee the additivity of C under the tensor product $\Phi_1 \otimes \Phi_2$].

These (super)additivity issues are instances of a general (super)additivity problem in quantum information theory (Holevo, 2007c). While it was known early on that the coherent information can be superadditive (Shor and Smolin, 1996; DiVincenzo, Shor, and Smolin, 1998), hence the regularization over n in Eq. (155) is in general necessary, the problem of determining whether the Holevo information is additive or superadditive under the tensor product of quantum channels has been for a long time an open problem. The additivity of the Holevo information was shown to be equivalent to the additivity of other quantities in quantum information theory (Shor, 2004), most notably the entanglement of formation (Bennett *et al.*, 1996) and the minimum output entropy (King and Ruskai, 2001), and was put in connection with the behavior of a family of operator norms under composition of quantum channels (Amosov and Holevo, 2000; Hayden and Winter, 2008). Only recently it was established that the Holevo information can indeed be superadditive for certain quantum channels (Hastings, 2009), also implying the superadditivity of the minimum output entropy and the entanglement of formation. Extensions of the results presented by Hastings (2009) can be found in Aubrun, Szarek, and Werner (2011), Brandão and Horodecki (2010), and Fukuda, King, and Moser (2010).

However, notwithstanding the fact that the Holevo information is generally superadditive, it has been proven to be additive for certain classes of quantum channels (Hiroshima, 2006; Amosov and Mancini, 2009), among which are the qubit unital channels (King, 2002) and the entanglement-breaking channels (Shor, 2002b). In particular, it is worth observing that in the case of qc channels (16) the classical capacity equation (152) reduces to the Shannon capacity (144) for a classical channel with conditional probability $p(y|x) = \langle e_y|E_x|e_y\rangle$. Moreover, as observed by Holevo (2007a) if a quantum channel is (super)additive for the Holevo information, so is its complementary channel. The coherent information has been proven to be additive for degradable channels (see Sec. II), for antidegradable channels [for which the quantum capacity is always zero (Bennett, DiVincenzo, and Smolin, 1997)], and for entanglement-breaking channels (Cubitt, Ruskai, and Smith, 2008) and PPT maps (see Sec. II.G) (Horodecki, Horodecki, and Horodecki, 1996; Peres, 1996). Regarding the private classical capacity $C_p(\Phi)$ it is known that it coincides with its product-state version $C_{p,1}(\Phi)$ for degradable and antidegradable maps (Devetak, 2005): in particular, in both cases one has $C_p(\Phi) = C_{p,1}(\Phi) = Q(\Phi) = Q_1(\Phi)$, which for antidegradable maps implies $C_p(\Phi) = 0$.

A remarkable example of superadditivity for the quantum capacity, called superactivation, has been provided by Smith and Yard (2008) by building up from previous results on the quantum assisted capacities (Smith, Smolin, and Winter, 2008). In particular, it has been shown that it is possible to find channels Φ_1 and Φ_2 with zero-quantum capacity, i.e., $Q(\Phi_1) = Q(\Phi_2) = 0$, for which, by parallel use of the two communication lines Φ_1 and Φ_2 in $\Phi_1 \otimes \Phi_2$, it becomes possible to transmit quantum information, i.e., $Q(\Phi_1 \otimes \Phi_2) > 0$. Specifically, Φ_1 and Φ_2 are given by an antidegradable channel and a PPT channel which possesses a nonzero private classical capacity (Horodecki *et al.*, 2005, 2008). Note that the two channels have zero-quantum capacity for different reasons: the first as a consequence of the no-cloning theorem and the second due to the fact that entanglement cannot be distilled from a PPT state. However, while it cannot be used to distill entanglement, there exist PPT channels that can still be used to establish a secret key between the sender and the receiver. See Brandão, Oppenheim, and Strelchuk (2012) for other examples in terms of *depolarizing* maps and for a more general construction.

It is finally worth noting that in the context of zero-error classical capacity (see footnote 6 in Sec. V.A.1) superactivation effects have been observed by Duan and Shi (2008) and Duan (2009).

C. Coding theorems for memory channels

The operational definitions of channel capacities, introduced in Sec. V.A, apply to both memoryless and memory quantum channels. Indeed they express the optimal classical and quantum information transmission rates between two parties, no matter how complex the internal structure of the communication line is. However, the memory setting is often more complicated than the memoryless one. One notes, in particular, that when dealing with nonanticipatory channels, introduced in Sec. III.B, different notions of coding procedures and capacities can be defined depending on whom, among Alice, Bob or a third party (Eve), controls (or uses for the encoding and decoding procedures) the initial and final states of the memory system M. For instance, for the same communication line one can introduce the classical capacities C_{AB}, C_{AE}, $C_{EB,\mu}$, $C_{EE,\mu}$, with the first (second) index representing the party controlling the initial (final) memory state and μ being Eve's choice for the initial state of the memory M (when considered); the same classification holds for the other forms of capacity, i.e., quantum, private classical, etc. The differences between these various choices have been analyzed by Kretschmann and Werner (2005): here, for simplicity, only the situation in which the third party (Eve) has full control of the memory M will be considered.

1. Entropic bounds

The presence of correlations introduced by noise makes more remote the possibility of formalizing capacities in terms of entropic quantities when residing in the memory setting. A useful strategy is to derive bounds on the capacities, with particular attention to upper bounds, and then show whenever possible their achievability (thus providing coding theorems). The first attempt in this direction was presented by Bowen and Mancini (2004), where for the case of finite-memory channels, i.e., maps with memory of finite dimension (described in Sec. III.D.2), the bounds of Eqs. (158) and (159) discussed later (as well as an analogous inequality for the entanglement-assisted capacity) were derived and shown to be achievable for a class of Markovian channels.

Simple geometric considerations allow one to conclude that both C and Q, independently of the noise model, can never be larger than $\log_2 d$, where d is the dimension of the Hilbert space of an individual carrier (the rationale being that in the space of n carriers one cannot fit more than d^n orthogonal states). This threshold however is not particularly informative as it does not depend upon the CPTP mapping which describes the action of the channel (its value being achieved only by noiseless channels, i.e., by identity or unitary maps). Tighter upper limits on C and Q can be derived from the Holevo bound (72) and the quantum Fano inequality (65), respectively. Specifically as explicitly shown in Appendix D.1 for the classical capacity one gets

$$C \leq \lim_{n\to\infty} \frac{1}{n} \max_{\mathcal{E}} \chi(\mathcal{E}; \Phi^{(n)}), \tag{158}$$

where $\chi(\cdots)$ is the Holevo information defined as in Eq. (71) and the maximization is performed over the ensembles of the first n input carriers. Similarly for the quantum capacity one has

$$Q \leq \lim_{n\to\infty} \frac{1}{n} \max_{\rho} J(\rho, \Phi^{(n)}), \tag{159}$$

where the maximization is now performed over the set of density matrices of the first n carriers, and where $J(\cdots)$ is the coherent information defined in Eq. (68); see Appendix D.2.

By direct comparison with Eqs. (152) and (155), one notes that for memoryless channels (i.e., when $\Phi^{(n)} = \Phi^{\otimes n}$) the bounds given above coincide with the exact values of the corresponding capacities. If the channel has memory correlations however, this feature is typically lost apart from some special configurations that are analyzed in Secs. V.C.2 and V.C.3.

It is worth stressing that the inequalities (158) and (159) refer to the limit of an infinite number of channel uses, i.e., they are asymptotic. Besides them, one could also consider bounds referring to a finite number of channel uses (the so-called one-shot setting). Note that any situation in which a channel is used a finite number of times with arbitrarily correlated noise can be equivalently described as a single use of a larger channel. Bounds on the one-shot classical capacity have been found by Wang and Renner (2012) by using a relative entropy-type measure defined via hypothesis testing, while bounds on the one-shot quantum capacity have been derived by Buscemi and Datta (2010) in terms of a generalization of relative Renyi entropy of zero order.

2. Perfect memory channels

Perfect memory channels admit a Kraus representation with a number of Kraus operators growing subexponentially with the number of channel uses n; see Sec. III.D.3. In other words, the size of the environment is not large enough to "contain" the information sent from Alice to Bob, which is exponentially increasing in n, and so asymptotically the loss of information into the environment is negligible. This intuitively explains that perfect memory channels are asymptotically noiseless and have maximal capacities, i.e., $C = Q = \log_2 d$.

Specifically, one can verify that (Kretschmann and Werner, 2005; Giovannetti, Burgarth, and Mancini, 2009) given a perfect memory channel $\Phi^{(n)}$ [see Eq. (80)], for sufficiently large n there exists a coding procedure which allows zero-error classical communication for reference set $\mathcal{M}$ of size

$$|\mathcal{M}| \geq \frac{d^n}{d_M^2}, \tag{160}$$

with d_M satisfying Eq. (94) and d^n being the size of the n carriers. The corresponding rate is hence

$$R \geq \log_2 d - \frac{2}{n} \log_2 d_M,$$

that for $n \to \infty$ converges to the optimal value $\log_2 d$, implying hence $C = \log_2 d$. Analogously, for sufficiently large n there exists a zero-error quantum communication with a reference set $\mathcal{M}$ of dimension

$$|\mathcal{M}| \geq \frac{d^n}{d_M^4 + d_M^2} \tag{161}$$

with rate

$$R \geq \log_2 d - \frac{1}{n} \log_2 [d_M^4 + d_M^2]$$

which, again, for $n \to \infty$ converges to the optimal value $\log_2 d$.

3. Forgetful channels

Forgetful channels are characterized by the property that the effects of the initial memory state become negligible with time, i.e., memory effects die away exponentially fast, as discussed in Sec. III.D.6. This feature allows one to prove that the upper bounds of Eqs. (158) and (159) can be actually asymptotically achieved (Kretschmann and Werner, 2005). Such important result can be demonstrated by invoking a *double-blocking* encoding procedure which effectively maps forgetful channels into memoryless ones.

Consider a n-fold concatenation of a memory channel $\Phi^{(n)}$. If the channel is strictly forgetful (see Sec. III.D.6), there exists a finite integer m such that for all $n \geq m$ the final state of the memory system does not depend on its initial state. In such a case, it is possible to group the channels $\Phi^{(n)}$ into blocks of length $m + l$, encoding the input in the l channels and ignoring the intermediate m ones. Accordingly the memory channel is reduced to a memoryless one defined on the larger Hilbert spaces $\mathcal{H}_Q{}^{\otimes l+m}$, hence allowing one to extend the coding theorems for memoryless channels. Remarkably, the double-block strategy can be applied even if the channel is forgetful although not strictly forgetful. Therefore, the memoryless expressions for classical and quantum channel capacities in Eqs. (152)–(155) can be applied also in the memory setting for forgetful maps, and the entropic upper bounds in Eqs. (158) and (159) are exactly achieved (the same holds true for entanglement-assisted capacity).

Forgetful channels have been proven to constitute a dense set with the topology induced by the cb-norm distance (Kretschmann and Werner, 2005). That implies that any nonforgetful channel can be approximated by a forgetful one. Notwithstanding, their capacities may be different. An example can be given in the context of Markovian channels. A long-term memory channel, Sec. III.D.8, can be approximated by a forgetful Markovian channel, Sec. III.D.4. However, according to the coding theorem for long-term memory channels discussed in the following section, the capacity of the latter does not approximate the capacity of the former (Datta and Dorlas, 2009).

4. Long-term memory channels

An example of memory channels for which the bounds (158) and (159) are not tight is provided by the long-term quantum memory channels of the form (77); see Sec. III.D.8. In that context it is worth noting that if the set $\{\Phi_i\}_i$ contains a finite number of elements, the determination of capacities of the averaged channel is equivalent to the determination of capacities of the associated compound channel (see Sec. III.A.1), since for finite sums one can always bound the error probability of the individual (memoryless) branches by the error probability of the averaged channel and vice versa. Then, under the condition of $\{\Phi_i\}_{i=1}^{N<+\infty}$, it has been shown that the product-state classical capacity is given by

$$C_1 = \sup_{\mathcal{E}}[\min_i \chi(\mathcal{E}; \Phi_i)], \tag{162}$$

where the supremum is taken over all finite ensembles $\mathcal{E}$ of input states (Datta and Dorlas, 2007). This result was derived by employing a quantum version of Feinstein's fundamental lemma (Feinstein, 1954; Khinchin, 1957) and a generalization of Helstrom's theorem (Helstrom, 1976). The basic idea is to allow Alice and Bob to use the first channel uses to determine which, among the various possible channels Φ_i, happens to be assigned by the statistical process that defines the communication line via Eq. (77). After that, Alice and Bob can use a proper HSW encoding to optimize the communication rate. Accordingly it is clear that the maximum rate for which reliable transmission can be guaranteed is the lowest one among those allowed by the Φ_i. Indeed by operating the channel to the highest rate allowed by the collection of maps $\{\Phi_i\}$ will introduce errors with finite probability.

The product-state capacity can be generalized to give the classical capacity of the channel in the usual manner, that is, by considering inputs that are product states over uses of blocks of n channels, but may be entangled across different uses within the same block. This yields the value

$$C = \lim_{n\to\infty} \frac{1}{n} C_1(\Phi^{(n)}), \tag{163}$$

which, in general, is smaller than the bound (158). Similarly, the entanglement-assisted classical capacity has been proven to be expressed as (Datta, Suhov, and Dorlas, 2008)

$$C_{ea} = \sup_{\rho}[\min_i I(\rho; \Phi_i)], \tag{164}$$

where $I(\rho; \Phi_i)$ is the quantum mutual information (67). Finally, Bjelaković, Boche, and Nötzel (2009) provided the expression for the quantum capacity

$$Q = \lim_{n\to\infty} \frac{1}{n} \max_{\rho}[\inf_i J(\rho; \Phi_i^{\otimes n})]. \tag{165}$$

Actually it was shown, by means of a discretization technique based on τ nets, that this result holds true for the compound channel associated with an arbitrary set $\{\Phi_i\}$ (not only a finite one). Finding the best rate for quantum communication over an arbitrary set of channels can be viewed as a universal coding problem. As such this result looks like a quantum channel counterpart of the universal quantum data compression result discovered by Jozsa *et al.* (1998).

5. Ergodic cq channels with decaying input memory

For cq channels (Sec. II.E) $W: A^{\mathbb{Z}} \times \mathcal{B}^{\mathbb{Z}} \to \mathbb{C}$ which are stationary ergodic (Sec. III.C) and have decaying input memory (Sec. III.D.7), a coding theorem has been derived (Bjelaković and Boche, 2008) such that the classical capacity is given by

$$C(W) = \sup_{p \text{ stationary ergodic}} i(p, W), \tag{166}$$

where

$$i(p, W) := \lim_{n\to\infty} \frac{1}{n}[S(\rho_p^n) + S(\rho_W^n) - S(\rho_{p,W}^n)], \tag{167}$$

with

$$\rho_p^n = \sum_{x^n \in A^n} p^n(x^n)|x^n\rangle\langle x^n|, \tag{168}$$

$$\rho_W^n = \sum_{x^n \in A^n} p^n(x^n)\rho_{x^n}, \tag{169}$$

$$\rho_{p,W}^n = \sum_{x^n \in A^n} p^n(x^n)|x^n\rangle\langle x^n| \otimes \rho_{x^n}. \tag{170}$$

Here ρ_{x^n} denotes the density operator of the output state $W^n(x^n,\cdot)$, $x^n \in A^n$ and $|x^n\rangle = |e_{x_1}\rangle \otimes \cdots \otimes |e_{x_n}\rangle$ for some orthonormal basis $\{|e_i\rangle\}_{i=1}^{|A|}$ of $\mathbb{C}^{|A|}$.

The sup in Eq. (166) is calculated over all stationary ergodic probability measures p on $A^{\mathbb{Z}}$. That is, consider a shift $T\colon A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ of double infinite sequences of A; then p is stationary if $p(Ta) = p(a)$ for all $a \in A^{\mathbb{Z}}$. Moreover, it is ergodic if for all $a \in A^{\mathbb{Z}}$ such that $Ta = a$ it is $p(a) = 0$ or 1.

The above theorem results as an extension of the coding theorem for the input memoryless cq channel whose proof combines Wolfowitz's code construction (Wolfowitz, 1957) and a version of Feinstein's lemma (Blackwell, Breiman, and Thomasian, 1958) based on the notion of the joint input-output probability distribution.

VI. SOLVABLE MODELS

A. Examples of solvable models for memoryless channels

This section collects examples of discrete and continuous memoryless quantum channels for which classical or quantum capacities can be analytically calculated. For most of them the calculation is made feasible by the fact that the Holevo information or the coherent information is additive. Hence the regularization in the limit of infinite n of Eqs. (152) and (155) is not necessary as the capacities equal their product-state version.

1. Discrete-variable memoryless channels

A closed expression for the classical capacity can be obtained for unital qubit channels (mapping the two-dimensional identity operator into itself, see Sec. II.I.1) and for the depolarizing channel acting on a finite dimensional Hilbert space of arbitrary dimension. For these channels the Holevo information has been proven to be additive (King, 2002, 2003).

Since any unital qubit channel is unitary equivalent to a Pauli channel (28) and, as discussed in Sec. V.A, capacities are invariant under unitary tranformations, it is sufficient to consider the latter. As anticipated, the classical capacity for these maps equals its product-state version. The fundamental ingredient to achieve this goal is the inequality Eq. (D9), derived in Appendix D.1 (which for this special channel can be shown to be achievable), and the fact that $S_{\min}(\Phi^{\otimes n})$ happens to be additive, i.e., $S_{\min}(\Phi^{\otimes n}) = nS_{\min}(\Phi)$. The resulting expression for the classical capacity is then computed as

$$C(\Phi) = C_1(\Phi) = 1 - h\left(\frac{1+\xi}{2}\right), \tag{171}$$

where h is the binary Shannon entropy (66) and ξ is the maximum among $|p_0 + p_1 - p_2 - p_3|$, $|p_0 - p_1 + p_2 - p_3|$, and $|p_0 - p_1 - p_2 + p_3|$. One may notice that the capacity reaches its maximum value of 1 if and only if $\xi = 1$, i.e., when at least two of the probabilities p_i are different from zero. Vice versa the capacity nullifies for $\xi = 0$, i.e., when $p_i = 1/4$ which corresponds to a fully depolarizing qubit map sending ρ into the completely mixed state $\mathbb{1}/2$. In a similar way the classical capacity of the qudit depolarizing channel of Eq. (36) can be shown to be

$$C(\Phi) = C_1(\Phi) = \log_2 d + \left[1 - \lambda - \frac{1-\lambda}{d}\right]\log_2\left[\frac{1-\lambda}{d}\right] + \left[\lambda + \frac{1-\lambda}{d}\right]\log_2\left[\lambda + \frac{1-\lambda}{d}\right], \tag{172}$$

and the maximum in Eq. (151) is achieved by a set of d equiprobable orthogonal pure states (King, 2003). Moreover, the entanglement-assisted classical capacity $C_{ea}(\Phi)$ is given by the same expression as for $C(\Phi)$ but replacing d with d^2. A similar expression can be derived for the transpose depolarizing channel of Eq. (37) which has also been proven to have additive Holevo information (Fannes et al., 2004; Datta, Holevo, and Suhov, 2006).

Concerning the large class of qubit maps in Eq. (27), the full quantum capacity, corresponding to the product-state one for the degradable case (since the coherent information is additive) and vanishing for the antidegradable regime, is given by (Giovannetti and Fazio, 2005; Wolf and Pérez-García, 2007)

$$Q(\Phi) = \begin{cases} f(\theta,\phi) & \text{for } \cos(2\theta)/\cos(2\phi) > 0, \\ 0 & \text{for } \cos(2\theta)/\cos(2\phi) \le 0, \end{cases} \tag{173}$$

with

$$f(\theta,\phi) = \max_{q\in[0,1]}\{h[q\cos^2\theta + (1-q)\sin^2\phi] - h[q\sin^2\theta + (1-q)\sin^2\phi]\}. \tag{174}$$

Finally the erasure channel introduced in Sec. II.I.2 is one of the few examples for which one can compute the whole set of capacities. This map is degradable for $p \le 1/2$ and has a quantum capacity $Q(\Phi) = (1-2p)\log_2 d$ with d being the dimension of the input carrier; for $p \ge 1/2$ it is, instead, antidegradable, hence with vanishing Q. Its Holevo information is also additive, yielding a classical capacity equal to $C(\Phi) = (1-p)\log_2 d$ which can also be shown to coincide with the two-way classically assisted quantum capacity $Q_2(\Phi)$. Finally the entanglement-assisted classical capacity is $C_{ea}(\Phi) = 2C(\Phi) = 2(1-p)\log_2 d$ (Bennett, DiVincenzo, and Smolin, 1997; Bennett, Shor et al., 1999); these quantities are shown in Fig. 20.

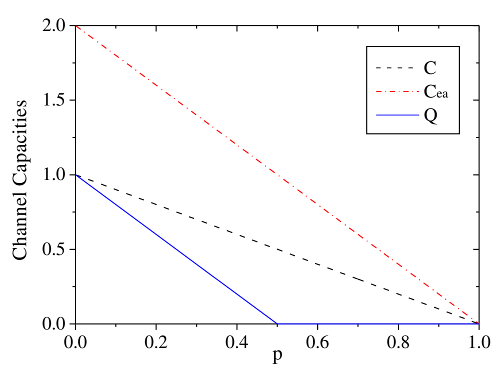


FIG. 20 (color online). Classical capacity (C), entanglement-assisted classical capacity (C_{ea}), and quantum capacity (Q) of the erasure channel mapping the input state into itself with probability p and into an orthogonal state otherwise; see Sec. II.I.2. The local dimension of the carrier is $d = 2$.

2. Continuous-variable memoryless channels

The very first example of a nontrivial CV memoryless channel for which the capacity has been explicitly computed was provided by Holevo, Sohma, and Hirota (1999) who considered a cq channel where classical messages are mapped into Guassian states obtained by continuously displacing an assigned Gibbs reference state. This is a special example of one-mode Guassian channels; see Sec. II.I.4. Under the memoryless condition the latter can be identified by a triad $(\mathsf{d}^{(1)}, \mathsf{X}^{(1)}, \mathsf{Y}^{(1)})$, where $\mathsf{d}^{(1)}$ is a two-component displacement vector, and $\mathsf{X}^{(1)}$, $\mathsf{Y}^{(1)}$ are 2×2 matrices. Therefore, a sequence of n consecutive channel uses is described by a triad $(\mathsf{d}^{(n)}, \mathsf{X}^{(n)}, \mathsf{Y}^{(n)})$, where $\mathsf{d}^{(n)} = \bigoplus_{k=1}^{n} \mathsf{d}^{(1)}$, $\mathsf{X}^{(n)} = \bigoplus_{k=1}^{n} \mathsf{X}^{(1)}$, and $\mathsf{Y}^{(n)} = \bigoplus_{k=1}^{n} \mathsf{Y}^{(1)}$.

The property of (anti)degradability holds for the single-mode channels describing the process of linear attenuation and amplification. These channels are characterized by a single parameter η (see Sec. II.I.4) and are known to be antidegradable (hence having null quantum capacity) for $\eta \leq 1/2$ and degradable (hence having additive coherent information) otherwise (Caruso and Giovannetti, 2006; Caruso, Giovannetti, and Holevo, 2006). For $\eta > 1/2$ their quantum capacity reads (Wolf, Pérez-García, and Giedke, 2007)

$$Q(\Phi) = \log \eta - \log |1-\eta|. \tag{175}$$

Analogously, if the mean number of bosonic excitation at the channel input is constrained to be less than N (see Sec. II.I.4), for $\eta > 1/2$ the constrained quantum capacity reads (Holevo and Werner, 2001; Wolf, Pérez-García, and Giedke, 2007)

$$Q(\Phi, N) = Q_1(\Phi, N) = g(\eta N) - g(|1-\eta|N), \tag{176}$$

where $g(x) := (x+1)\log(x+1) - x\log x$ for $x > 0$ and $g(x) := 0$ for $x \leq 0$.

Under a constraint of N mean input excitations, the Holevo information has been shown to be additive for the lossy bosonic channel ($\eta \in [0,1]$) (Giovannetti *et al.*, 2004c), allowing for a single-letter expression for the classical capacity:

$$C(\Phi; N) = C_1(\Phi; N) = g(\eta N). \tag{177}$$

Similarly, the constrained entanglement-assisted classical capacity (Giovannetti *et al.*, 2003a, 2003b) reads

$$C_{ea}(\Phi; N) = g(N) + g(\eta N) - g[(1-\eta)N]. \tag{178}$$

(Notice that their unconstrained counterparts, unlike the quantum capacity, are unbounded.) Besides being additive, the Holevo information for the lossy bosonic channel is maximized for Gaussian inputs. The same properties have been very recently proven to hold for a broad family of Gaussian channels, which includes the lossy and noisy channel, the linear amplifier, and the additive noise channel (Giovannetti, Holevo, and García-Patrón, 2013; Giovannetti, García-Patrón *et al.*, 2013). For all these channels, this result solves and gives a positive answer to a long-standing conjecture (Holevo and Werner, 2001; Giovannetti *et al.*, 2004a, 2010; Serafini, Eisert, and Wolf, 2005; Hiroshima, 2006; Guha, Erkmen, and Shapiro, 2007; Lloyd *et al.*, 2009; García-Patrón *et al.*, 2012; Giovannetti, Lloyd *et al.*, 2013; König and Smith, 2013a, 2013b) and proves single-letter expressions for their classical capacities [the latter were summarized by Lupo, Pirandola *et al.* (2011)]. For example, the capacity of the lossy and noisy Gaussian channel reads ($\eta \in [0,1]$)

$$C(\Phi; N) = C_1(\Phi; N) = g[\eta N + (1-\eta)N_{\rm th}] - g[(1-\eta)N_{\rm th}], \tag{179}$$

that of the linear amplifier is ($\eta \geq 1$)

$$C(\Phi; N) = C_1(\Phi; N) = g[\eta N + (\eta - 1)] - g[(\eta - 1)], \tag{180}$$

and for the additive noise channel one has

$$C(\Phi; N) = C_1(\Phi; N) = g(N + N_{\rm add}) - g(N_{\rm add}). \tag{181}$$

B. Examples of solvable models for memory channels

The main difficulty in the evaluation of the capacities of quantum channels with memory relies on the requirement of the regularization of the corresponding entropic quantities in the limit of infinite uses of the channel. For the case of forgetful channels this gives the exact expression for the capacities, while in general it provides an upper bound for nonforgetful channels.

Up to now only a few models of memory quantum channels have been fully solved in terms of their capacities. One is the dephasing channel (in the discrete-variable setting), and the other is the lossy bosonic channel (in the continuous-variable setting), with different types of correlations.

1. Discrete memory channels

Referring to the model discussed in Sec. III.A, consider a sequence of qubit carriers propagating at rate ν and interacting each one with a single qubit environment subject in turn to a relaxation process described by amplitude damping with a rate $1/\tau$. Then assume that the carrier-environment interaction is a control unitary, such that when the carrier is in $|0\rangle_{q_j}$ nothing happens to the environment, while when q_j is in $|1\rangle_{q_j}$ the environment undergoes the unitary transformation described by the operator $\gamma\sigma_z + \sqrt{1-\gamma^2}\sigma_x$. One hence has a memory channel whenever the condition $\nu\tau \ll 1$ is not satisfied. However, it is possible (Giovannetti, 2005) to trace this model back to a memoryless phase damping channel $\Phi_{\bar{\gamma}}$ with $p_z = (1-\bar{\gamma})/2$ the probability of σ_z error. Here $\bar{\gamma}$ is a complicated function of several parameters including ν and τ and it reduces to γ for $\nu\tau \ll 1$ (the memoryless limit of Sec. III), while it can be $\bar{\gamma} > \gamma$ for $\nu\tau \geq 1$, thus making $\Phi_{\bar{\gamma}}$ effectively less noisy than Φ_γ.

In the case of the phase damping channels (see Sec. II.I.1) the capacities can be explicitly computed. For instance, since the noise does not affect the populations associated with the computational basis, the classical capacity of the phase damping channel Φ_γ is $C(\Phi_\gamma) = 1$.

On the other hand, the quantum capacity of a phase damping channel Φ_γ is $Q(\Phi_\gamma) = 1 - h(p_z)$ (Devetak and Shor, 2005; Wolf and Pérez-García, 2007), where h is the binary entropy. One hence has $Q(\Phi_{\bar{\gamma}}) \geq Q(\Phi_{\gamma})$ for $\bar{\gamma} > \gamma$, i.e., enhanced quantum capacity by memory effects.

Markovian correlated dephasing was considered by D'Arrigo, Benenti, and Falci (2007), where the degree of correlations is expressed by a correlation parameter $\mu \in [0, 1]$ which characterizes the Markovian transition probabilities as in Eq. (99). Likewise in the above model, the quantum capacity increases when considering a higher degree of memory. In particular, the memoryless dephasing channel capacity is recovered for $\mu = 0$, while for $\mu = 1$ (perfect memory) the channel is asymptotically noiseless, i.e., $Q(\Phi) = 1$ (Bowen and Mancini, 2004). D'Arrigo, Benenti, and Falci (2007) also considered a microscopic model for correlated dephasing defined in terms of a spin-boson model, where quantum information is encoded in a train of qubits and a single bosonic mode represents the memory system. Lower bounds for the quantum capacity of a qubit memory channel with both correlated dephasing and damping have been evaluated numerically starting from a microscopic spin-boson model with Jaynes-Cummings interaction in the presence of strong dephasing noise (Benenti, D'Arrigo, and Falci, 2009, 2012).

Plenio and Virmani (2007, 2008) considered another model of dephasing memory channel for qubits. It can be traced back to the scenario introduced by Giovannetti and Mancini (2005) and schematized in Fig. 13, where each individual information carrier (a qubit in this case) interacts with a corresponding environment particle, the correlations being established by the environment multiparticle state. Specifically they considered the case where the two-particle (two-qubit) interaction is defined by a controlled-phase gate, the environmental particle being the controller qubit that determines which unitary transformation will be applied to the carrier. As a consequence the join state of the carriers gets transformed through mixtures of random sequences of identity and σ_z operators, each sequence being characterized by a (correlated) probability which depends upon the diagonal elements of the environment initial state.

The interesting feature of this model is that it allows one to write explicit formulas for the associated capacities for the channel in terms of properties of the many-body environment that share a close relationship with thermodynamical quantities. In particular, the CJ state of their family of correlated channels is a maximally correlated state (i.e., state of the form $\sum_{i,j} \alpha_{i,j} |ii\rangle\langle jj|$) (Rains, 1999a, 1999b, 2001), and, combining this feature with the forgetfulness of such maps, one can show (Plenio and Virmani, 2007, 2008) that the quantum capacity can be expressed in terms of the regularized diagonal entropy of the system environment, i.e.,

$$Q(\Phi) = 1 - \lim_{n\to\infty} \frac{S[\mathrm{diag}(\rho_{\mathrm{env}})]}{n}, \tag{182}$$

where $\mathrm{diag}(\rho_{\mathrm{env}})$ is the environmental state in the computational basis after eliminating all off-diagonal elements [note that the coding argument used in order to arrive at Eq. (182) has also been independently shown by Hamada (2002)]. For the special case in which the initial state of the environment is described by a classically correlated many-body system (i.e., diagonal in the computational basis), the last term on the right-hand side of Eq. (182) coincides with the thermodynamical entropy of the environment. Hence, the capacity is given by

$$Q(\Phi) = 1 - \left(1 - \beta \frac{\partial}{\partial \beta}\right) \lim_{n\to\infty} \frac{1}{n} \log_2 Z_n, \tag{183}$$

where Z_n is the partition function for n environment spins, and β is the associated inverse temperature. In other words, one can exploit results from classical statistical physics in order to compute the capacity, as shown by Eq. (183).

The calculation of the entropy of the associated many-body system, and hence of the quantum capacity of the memory channel, can be done exactly in certain relevant cases. One of them is the case of many-body systems described by matrix-product states (MPSs) involving only rank-1 matrices. For the sake of simplicity, one focuses on a translationally invariant MPS for a 1D system of two-level particles, with periodic boundary conditions. This environmental state is characterized by two matrices A_0 and A_1 and is given by $|\psi\rangle = \sum_{i_1\cdots i_n} \mathrm{Tr}\{\mathsf{A}_{i_1} \cdots \mathsf{A}_{i_n}\} |i_1 \cdots i_n\rangle$. Then, by dephasing each qubit, the resulting unnormalized state is

$$\rho = \sum_{i_1\cdots i_n} \mathrm{Tr}\left[\prod_{k=1}^{n} (\mathsf{A}_{i_k} \otimes \bar{\mathsf{A}}_{i_k})\right] |i_1 \cdots i_n\rangle\langle i_1 \cdots i_n|, \tag{184}$$

where $\bar{\mathsf{A}}$ is the complex conjugate matrix of A. It is possible to show that, if $|i_1 \cdots i_n\rangle$ has l occurrences of 0 and $n - l$ of 1, and k boundaries between 0's and 1's blocks, then the corresponding diagonal elements of ρ are proportional to $a^l b^{n-l} c^k$, with a (respectively, b) being the eigenvalue of $\mathsf{A}_0 \otimes \bar{\mathsf{A}}_0$ (respectively, $\mathsf{A}_1 \otimes \bar{\mathsf{A}}_1$), and c being the eigenvalue of $(\mathsf{A}_0 \otimes \bar{\mathsf{A}}_0)(\mathsf{A}_1 \otimes \bar{\mathsf{A}}_1)/(ab)$.

Finally it is worth remarking that Wolf *et al.* (2006) showed the existence of Hamiltonians exhibiting quantum phase transitions and with ground states being MPSs involving only matrices of rank 1. Hence, it can be shown that the diagonal elements of such MPSs are equal to the probability $\wp$ of microstates in corresponding classical Ising chains. Therefore, by exploiting this connection, one can easily compute the limit in Eq. (182) by using well-known many-body physics methods. Figure 21 shows the case of the following Hamiltonian:

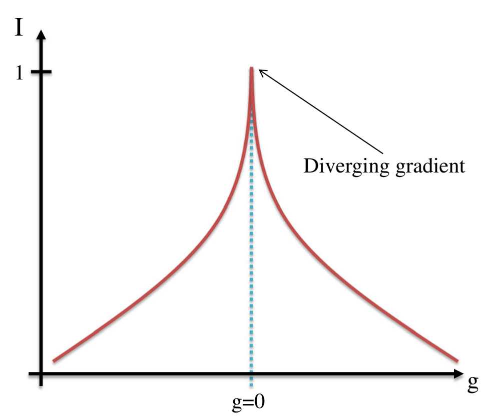


FIG. 21 (color online). Sketch of the capacity behavior in the case of an environment given by the ground state of the Hamiltonian (185). Notice the divergent gradient near the "phase transition", i.e., at $g = 0$. From Plenio and Virmani, 2007.

$$\sum_i 2(g^2-1)\sigma_{z,i}\sigma_{z,i+1} - (1+g)^2\sigma_{x,i} + (g-1)^2\sigma_{z,i}\sigma_{x,i}\sigma_{z,i+1}. \tag{185}$$

In this case, one knows that the ground state is a rank-1 MPS which possesses a nonstandard phase transition at $g = 0$, where indeed some correlation functions are nondifferentiable (though continuous) and the ground state energy is analytic (Wolf *et al.*, 2006).

2. Continuous memory channels

Among Gaussian memory channels, one can identify a subclass of channels for which the memory effects can be *unraveled*. That is, by applying suitable unitary encoding and decoding transformations, n uses of such channels are mapped to n independent single-mode channels used in parallel. By applying known results for the memoryless setting one may then compute the capacities of the memory channel (see Sec. VI.A.2).

Such a unitary mapping from n uses of a Gaussian memory channel to n parallel uses of independent single-mode channels was first considered by Cerf *et al.* (2005, 2006) and Giovannetti and Mancini (2005), and then applied for estimating the communication capacities of Gaussian memory channels in several settings (Lupo, Memarzadeh, and Mancini, 2009; Lupo, Pilyavets, and Mancini, 2009; Schäfer, Karpov, and Cerf, 2009; Lupo, Giovannetti, and Mancini, 2010a, 2010b). A formal definition of the class of memory channels that can be unraveled first appeared in Lupo and Mancini (2010).

If one takes the one-mode channel as a reference point, representing a single use of the channel, n uses of the quantum memory channel are characterized by the triads $(\mathsf{d}^{(n)}, \mathsf{X}^{(n)}, \mathsf{Y}^{(n)})$ such that either $\mathsf{d}^{(n)} \neq \bigoplus_{k=1}^{n} \mathsf{d}^{(1)}$ or $\mathsf{X}^{(n)} \neq \bigoplus_{k=1}^{n} \mathsf{X}^{(1)}$, $\mathsf{Y}^{(n)} \neq \bigoplus_{k=1}^{n} \mathsf{Y}^{(1)}$ (see Sec. II.I.4). A memory channel can be unraveled if there exist unitary transformations $\Phi_E^{(n)}, \Phi_D^{(n)}$, acting on n modes, such that $\Phi_D^{(n)} \phi^{(n)} \Phi_E^{(n)} = \otimes_{k=1}^{n} \phi_k^{(1)}$; that is, n uses of the memory channel are unitary equivalent to the tensor product of n independent, but not necessarily an identical, single-mode Gaussian channel (this mapping is depicted in Fig. 22). Since the application of unitary transformations cannot change the capacities of the channel, they can be equivalently computed for the unraveled channel, in which each input mode is transformed independently (although in general not identically). If one is interested in the calculation of constrained capacities, then one has to take into account how the constraint changes under the action of the encoding and decoding unitaries. A relevant setting is that of encoding transformations preserving the constraint. For the case of the constrained mean input excitation number, the constraint is preserved if

$$\sum_{k=1}^{n} a_k^\dagger a_k = \Phi_E^{*(n)}\left(\sum_{k=1}^{n} a_k^\dagger a_k\right),$$

a condition which is satisfied when the encoding unitary is a linear passive transformation, e.g., in the case of optical realization, when exploiting a network of beam splitters and phase shifters [see, e.g., Ferraro, Olivares, and Paris (2005)].

If a Gaussian memory channel can be unraveled, then its capacities can be computed upon reduction to the case of a memoryless single-mode Gaussian channel. This is the case for the model of a lossy channel with memory introduced by Lupo, Giovannetti, and Mancini (2010a). In this model, the action of the channel upon n uses is defined by the concatenation of n identical unitary transformations coupling the input modes $a_1, a_2, \ldots, a_n$ with a collection of local environmental modes $e_1, e_2, \ldots, e_n$ and the memory mode m. Specifically the evolution of the kth input mode is obtained by a concatenation of two beam-splitter transformations, the first with transmissivity ϵ and the second with transmissivity η; see Fig. 23. This results in a nonanticipatory channel with ISI (see Sec. III.B) having the same structure depicted in Fig. 12(c). By varying the transmissivity parameters, the model is capable of describing different memory schemes, from the memoryless lossy bosonic channel configuration (Giovannetti *et al.*, 2004c) (the input a_k influences only the output b_k), to a channel with perfect memory (all a_k interacts only with the memory mode m_1) (see Sec. III.D.3), to a *quantum shift* channel (Bowen and Mancini, 2004) where each input state is replaced by the previous one (this is obtained by setting $\eta = 0$, $\epsilon = 1$). Extensions of Lupo, Giovannetti, and Mancini (2010a) which encompass memory effects in linear amplification and thermalization processes are presented by Lupo, Giovannetti, and Mancini (2010b) and De Palma, Mari, and Giovannetti (2014), respectively. All these models can be unraveled into the tensor product of one-mode lossy or amplifier channels.

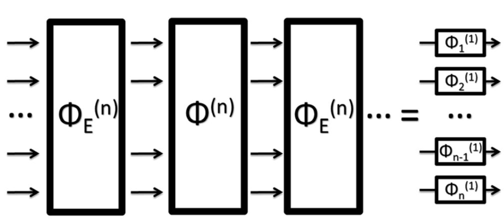


FIG. 22. Unraveling of n uses of a memory channel. Each horizontal line indicates one bosonic mode, propagating from the left to the right. $\phi^{(n)}$ denotes n uses of the memory channel. $E^{(n)}$ and $D^{(n)}$ are preprocessing and postprocessing Gaussian unitaries. $\phi_k^{(1)}$'s are one-mode Gaussian channels.

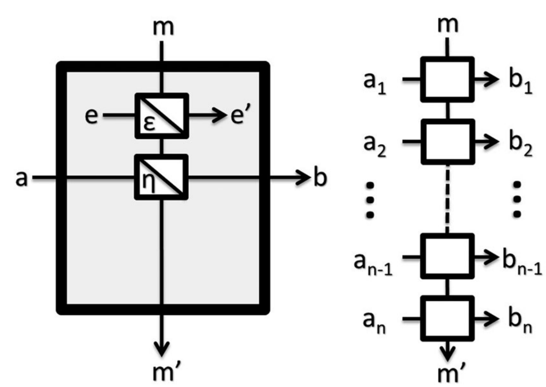


FIG. 23. Left: A single use of the lossy bosonic memory channel. From Lupo, Giovannetti, and Mancini, 2010a. Right: The n-fold concatenation of the memory channel: photons entering in the kth input mode a_k can emerge only in the output ports $b_{k'}$ with $k' \geq k$.

The capacities of these memory channels can hence be computed following four steps: first the memory channel is unraveled into the direct product of the single-mode Gaussian channels; second the optimization of the relevant entropic function is performed modewise under a constrained mean input excitation number; then the distribution of the mean excitation number over the input modes is optimized; finally the asymptotic limit of infinite channel uses is considered. The optimization of the distribution of the mean excitation number leads to a *quantum water filling* solution for the capacity of the memory channel, where the way the mean excitation number is distributed over input modes is analogous to the way water distributes into a vessel (Cover and Thomas, 1991). While algorithms for the optimization were presented by Schäfer, Karpov, and Cerf (2011) and Pilyavets, Lupo, and Mancini (2012), the most delicate point is the consideration of the asymptotic limit (Lupo, Memarzadeh, and Mancini, 2009; Lupo, Giovannetti, and Mancini, 2010a).

Memory channels with additive noise are characterized by having $\mathsf{X}^{(n)} = \mathbb{1}$. They can be realized by means of multimode CV teleportation protocol (Vaidman, 1994; Braunstein and Kimble, 1998; Ban, Sasaki, and Takeoka, 2002), where the teleportation resource is a multimode state (Caruso, Giovannetti, and Palma, 2010). The memory channel considered by Cerf *et al.* (2005, 2006) belongs to this class. The latter was defined for two channel uses, represented by two bosonic modes, which are affected by correlated additive noise. A generalization of this model to the case of more than two channel uses was first introduced by Ruggeri and Mancini (2007a) and subsequently by Lupo, Memarzadeh, and Mancini (2009) and Schäfer, Karpov, and Cerf (2009), where the additive noise, characterized by the matrix $\mathsf{Y}^{(n)}$, constitutes a Markov process.

It is easy to recognize that these models define SI memory channels, which are instances of the general scheme depicted in Fig. 13 and first introduced by Giovannetti and Mancini (2005). Here Gaussian memory effects were introduced by imposing that the n input modes interact modewise with a joint (possibly entangled) Gaussian state of n environmental modes through beam-splitter transformations of transmissivity η [the associated $\mathsf{Y}^{(n)}$ matrix of the channel being $(1-\eta)\mathsf{C}^{(n)}$, where $\mathsf{C}^{(n)}$ is the CM of the environmental state]. These Gaussian memory channels can be unraveled whenever the matrix $\mathsf{Y}^{(n)}$ has a suitable form; furthermore under certain conditions they can be unraveled with the use of energy-preserving unitary preprocessing transformation (Pilyavets, Zborovskii, and Mancini, 2008; Lupo, Pilyavets, and Mancini, 2009) [see also Ruggeri *et al.* (2005)].

It is worth noting that, unlike the case of discrete-variable memory channels (see Sec. III.D.4), there is no transitional behavior in these models of Gaussian memory channels: the optimal input states are either separable or entangled according to the model symmetries (Cerf *et al.*, 2005, 2006; Lupo and Mancini, 2010). As entangled states cannot be prepared locally, it is crucial to identify suboptimal input states that can be prepared efficiently. This issue was considered by Schäfer, Karpov, and Cerf (2012), where it was shown that encoding classical information via Gaussian matrix-product states (Adesso and Ericsson, 2006; Schuch, Cirac, and Wolf, 2008), which can be efficiently prepared, may allow one to achieve a reliable communication rate close to the channel capacity. An analysis of correlated additive Gaussian channels beyond the case of Markovian correlations was presented by Schäfer, Karpov, and Cerf (2011).

Finally, it is worth remarking that the study of Gaussian memory channels has also stimulated and motivated a deep analysis of the communication capacities of the single-mode memoryless Gaussian channel (Schäfer, Karpov, and Cerf, 2010; Lupo, Pirandola *et al.*, 2011; Pilyavets, Lupo, and Mancini, 2012). In particular, Pilyavets, Lupo, and Mancini (2012) and Schäfer, Karpov, and Cerf (2010) provided a complete characterization of one-mode Gaussian channels, respectively, for the case of lossy channels and additive noise.

VII. QUANTUM CHANNELS DIVISIBILITY AND DYNAMICAL MAPS

In this section we leave the input-output scenario, which has characterized all the previous parts of the review, and focus on the memory effects that may arise when studying the dynamical evolution of a system that is evolving in time while interacting with an external environment; see Fig. 3.

As discussed in Sec. II.B the concatenation of CPTP maps defines a new quantum channel. It is also worth considering whether the converse is also true, that is, under which conditions a quantum channel $\Phi \in \mathfrak{P} \coloneqq \mathfrak{P}(Q \mapsto Q)$ acting on a system Q can be expressed as a concatenation of other elements of $\mathfrak{P}$. This is intimately related to the semigroup structure of the set of quantum channels, hence with dynamical maps and master equations.

A. Divisible and indivisible quantum channels

Loosely speaking, by divisibility of a quantum channel $\Lambda \in \mathfrak{P}$ one refers to the possibility of decomposing it in terms of concatenation of other channels, i.e., to the possibility of writing $\Lambda = \Lambda_1 \circ \Lambda_2$, with $\Lambda_i \in \mathfrak{P}$. Obviously, every channel $\Lambda \in \mathfrak{P}$ is divisible in the following way: $\Lambda = (\Lambda \circ \mathcal{U}^{-1}) \circ \mathcal{U}$, with $\mathcal{U}$ any unitary map. A nontrivial definition of (in) divisibility was introduced by Wolf and Cirac (2008). According to that, a quantum channel $\Lambda \in \mathfrak{P}$ is indivisible if every decomposition of the form $\Lambda = \Lambda_1 \circ \Lambda_2$, with $\Lambda_i \in \mathfrak{P}$, implies that either Λ_1 or Λ_2 is a unitary conjugation. Otherwise, Λ is said to be divisible. It happens that quantum channels with maximal Kraus rank (d^2) are divisible (Wolf and Cirac, 2008).

Hereafter the subset of $\mathfrak{P}$ of divisible channels is denoted as $\mathfrak{D}$. The notion of divisibility can then be refined by considering different kinds of divisible quantum channels. First one introduces a notion of Markovianity for quantum channels related to their decomposability, rather than to their composability as done in Sec. III.D. According to Wolf and Cirac (2008) a quantum channel is called Markovian if it is an element of a continuous one-parameter semigroup of CPTP maps.

In such a case there exists a (Liouvillian) generator $\mathcal{L}$ such that the quantum channel can be written as $\Lambda(t) = e^{t\mathcal{L}} \in \mathfrak{P}$ for all $t \geq 0$. A standard form for such generators was derived by Gorini, Kossakowski, and Sudarshan (1976) and Lindblad (1976):

$$\mathcal{L}\rho = i[\rho, H] + \sum_{\alpha,\beta} G_{\alpha,\beta}\left(F_\alpha \rho F_\beta^\dagger - \frac{1}{2}\{F_\beta^\dagger F_\alpha, \rho\}\right), \quad (186)$$

where the matrix G is positive semidefinite, $\{,\}$ denotes the anticommutator, and the operators H and F_α, respectively, describe the Hamiltonian and non-Hamiltonian dynamical terms.

Through the (Liouvillian) generator $\mathcal{L}$ one can write down the dynamical (master) equation for the system density operator ρ (Breuer and Petruccione, 2002)

$$\frac{d}{dt}\rho(t) = \mathcal{L}\rho(t). \quad (187)$$

Its solution, for given initial condition $\rho(t_0)$, reads $\rho(t) = \Lambda(t - t_0)\rho(t_0)$ with $\Lambda(t - t_0) = e^{(t-t_0)\mathcal{L}}$ obeying the homogeneous composition law

$$\Lambda(t_1)\circ\Lambda(t_2) = \Lambda(t_1 + t_2), \quad (188)$$

for $t_1, t_2 \geq 0$, hence defining a one-parameter semigroup of CPTP maps. As consequence, Eq. (187) is called a Markovian master equation.

A class of Markovian master equations of this kind can be obtained as the continuous-time limit of a concatenation of identical system-bath interactions. These models, known as collision models (Rau, 1963; Alicki and Lendi, 1987; Terhal and DiVincenzo, 2000; Scarani *et al.*, 2002; Ziman *et al.*, 2002; Ziman and Bužek, 2005; Ziman, Štelmachovič, and Bužek, 2005), are defined by the iterated unitary interactions of the system Q with n identical reservoirs $E = (e_1, \ldots, e_n)$. This cascade process, depicted in Fig. 24, defines a quantum channel of the form

$$\Phi^n(\rho_Q) = \mathrm{Tr}_E[U_{Qe_1}\cdots U_{Qe_n}(\rho_Q \otimes \omega_E^{\otimes n})U_{Qe_1}^\dagger \cdots U_{Qe_n}^\dagger], \quad (189)$$

where U_{Qe_j}'s are n instances of a unitary transformation coupling the system Q with the environmental systems. A comparison with Fig. 12(a) is useful to enlighten the relations between this model and the unitary dilation of memory channels introduced in Eq. (80): basically in passing from the latter to Eq. (189) the environment and the carriers have exchanged their roles transforming the spatial correlations of Eq. (80) into temporal correlations. A hybrid approach which includes both effects was recently introduced by Giovannetti and Palma (2012): as shown in Fig. 25 the scheme has the same structure as Fig. 24 for each row, and the same as Fig. 12(a) for each column. This model provides a link between memory channels and time-continuous dynamical evolutions.

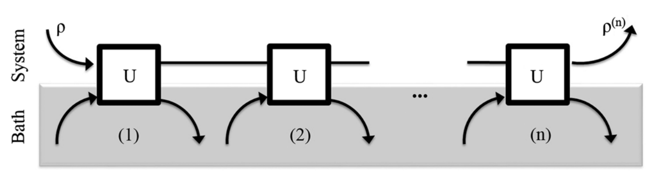


FIG. 24. The cascade structure of a collision model, defined by the concatenation of identical unitaries. From Scarani *et al.*, 2002.

The set of Markovian quantum channels is denoted below by $\mathfrak{M}$. Clearly $\mathfrak{M} \subset \mathfrak{D}$ because any Markovian quantum channel can be divided into a large number of infinitesimal channels being the solution of the (time-independent) master equation (187).

Then one can attempt to single out the class of quantum channels that can be split into infinitesimal pieces, i.e., into channels arbitrarily close to the identity. Clearly it would contain $\mathfrak{M}$. Actually the set $\mathfrak{I}$ of infinitesimal divisible quantum channels can be defined (Wolf and Cirac, 2008) as the closure of the set of all families $\{\Lambda(t_2, t_1) \in \mathfrak{P} | t_1, t_2 \in [0, t]\}$ of quantum channels for which there exists a continuous mapping $[0, t] \times [0, t] \to \mathfrak{P}$ onto $\{\Lambda(t_2, t_1)\}$ such that

(1) $\Lambda(t_3, t_2)\circ\Lambda(t_2, t_1) = \Lambda(t_3, t_1)$, for all $0 \leq t_1 \leq t_2 \leq t_3 \leq t$, and

(2) $\lim_{\epsilon\to 0} |||\Lambda_{\tau+\epsilon,\tau} - \mathrm{id}|||_2 = 0$, for all $\tau \in [0, t)$,

where $|||\cdots|||_2$ is the superoperator norm defined in Eq. (56) of Appendix II.J.2—the closure being intended with respect to the associated distance. For a given family there is a continuous path in $\mathfrak{P}$ (where one can move by concatenating quantum channels) connecting any element of the family with the identity.

Actually one could consider in the above definition a set $\mathfrak{I}_M$ analogous to $\mathfrak{I}$ with the restriction $\Lambda \in \mathfrak{M}$, i.e., of the form $\Lambda(t) = e^{t\mathcal{L}}$. It is obvious that $\mathfrak{I}_M \subseteq \mathfrak{I}$. Intuitively also the converse should be true since any quantum channel close to the identity is "almost Markovian" according to the definition of a Markovian quantum channel. In fact, it was proven by Wolf and Cirac (2008) that any infinitesimal divisible quantum channel can be (arbitrary well) approximated by a product of Markovian quantum channels.

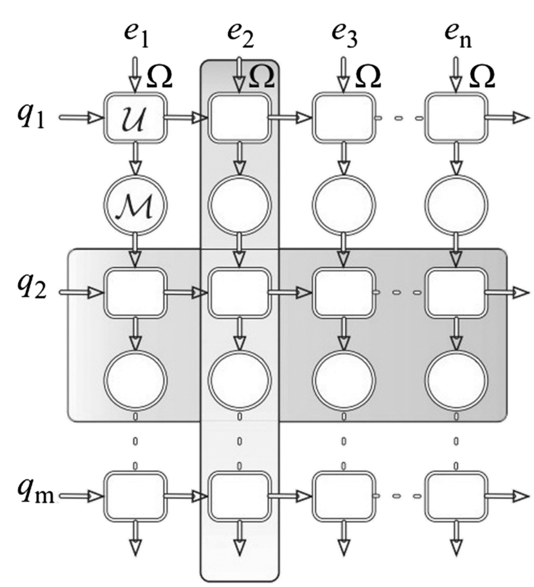


FIG. 25. The cascade structure leading to the master equation for correlated quantum channels discussed by Giovannetti and Palma (2012), and described by Eq. (189). Each row corresponds to a single collision model (see Fig. 24), and each column corresponds to a memory channel [see Fig. 12(a)].

In summary one has the following chain of inclusion $\mathfrak{M} \subset \mathfrak{I} \subset \mathfrak{D} \subset \mathfrak{P}$. The complement of $\mathfrak{D}$ to $\mathfrak{P}$ is given by the indivisible quantum channels.

B. Non-Markovian master equations

The simplest generalization of the dynamical equation (187) is obtained by introducing a time-dependent Liouvillian $\mathcal{L}(t)$ admitting the representation (186), but with time-dependent operators $H(t)$ and $F_\alpha(t)$. Hence, the time-dependent equation for the dynamical map $\Lambda(t, t_0)$

$$\frac{d}{dt}\Lambda_{t,t_0} = \mathcal{L}(t)\circ\Lambda(t, t_0), \qquad \Lambda(t_0, t_0) = \mathrm{id}, \tag{190}$$

has a formal solution

$$\Lambda(t, t_0) = \mathbb{T}\exp\left(\int_{t_0}^{t} \mathcal{L}(\tau)d\tau\right), \tag{191}$$

where $\mathbb{T}$ denotes time ordering. Different from the time-homogeneous case (188), the explicit dependence on time implies that the dynamical map $\Lambda(t, t_0)$ is no more a function of $t - t_0$ only. Notwithstanding, it still satisfies the inhomogeneous composition law

$$\Lambda(t, s)\circ\Lambda(s, t_0) = \Lambda(t, t_0), \tag{192}$$

for any $t \geq s \geq t_0$. The Markovian character is hence preserved by the time-dependent dynamical equation (190) and it implies the infinitesimal divisibility discussed in Sec. VII.A. This is true if one intends the Markovian character simply expressed by an associative binary operation like Eq. (192) (a quantum version of the Chapman-Kolmogorov equation). However, it results that the Chapman-Kolmogorov equation is a necessary but not sufficient condition for having Markov chains (processes) (Vacchini *et al.*, 2011).

On the other hand, from the fact that any infinitesimal divisible quantum channel can be (arbitrary well) approximated by a product of Markovian quantum channels (as discussed in Sec. VII.A), it follows that every infinitesimally divisible quantum channel can be written as a solution of a time-dependent master equation [Wolf and Cirac (2008) proved this fact for $d = 2$ and argued the same for $d > 2$]. Hence, loosely speaking, one can say that the class of infinitesimal divisible channels corresponds to the set of solutions of time-dependent master equations.

A more general dynamical equation comes from the Nakajima-Zwanzig projection operator technique (Nakajima, 1958; Zwanzig, 1960; Breuer and Petruccione, 2002) and reads as follows:

$$\frac{d}{dt}\rho(t) = \int_{t_0}^{t} \mathcal{K}(t - u)\rho(u)du, \qquad \rho(t_0) = \rho_0. \tag{193}$$

Here one has memory effects modeled by the *memory kernel* superoperator $\mathcal{K}(t)$. Hence, the rate of change of the state at time also depends on its history, and the Markovian setting (187) is recovered when $\mathcal{K}(\tau) = 2\delta(\tau)\mathcal{L}$.

The dynamical map $\Lambda(t, t_0)$ associated with the non-Markovian evolution (193) is a solution of

$$\frac{d}{dt}\Lambda(t, t_0) = \int_{t_0}^{t} d\tau\mathcal{K}(t - \tau)\circ\Lambda(\tau, t_0), \qquad \Lambda(t_0, t_0) = \mathrm{id}. \tag{194}$$

It appears to be a function of both t_0 and t. However, one can notice that the dynamics of an open quantum system can be always understood as the reduced dynamics of its unitary dilation (see Sec. II) which includes the environment. Being the unitary dynamics of an isolated system homogeneous in time, it follows that, once the degrees of freedom of the environment are taken into account, the dynamical map will be only a function of the difference $t - t_0$; that is, $\Lambda(t, t_0) \equiv \Lambda(t - t_0)$. This mirrors the fact that any solution of Eq. (194) is also a solution of the time-dependent equation (Chruściński and Kossakowski, 2010)

$$\frac{d}{dt}\Lambda(t - t_0) = \mathcal{L}(t, t_0)\circ\Lambda(t, t_0), \qquad \Lambda(t_0, t_0) = \mathrm{id}, \tag{195}$$

with a time-dependent Liouvillian defined by the logarithmic derivative of the dynamical map

$$\mathcal{L}(t - t_0) := \left[\frac{d}{dt}\Lambda(t - t_0)\right]\circ\Lambda^{-1}(t - t_0).$$

Nevertheless, the explicit dependence of the generator on the initial time t_0 implies that $\mathcal{L}$ is effectively nonlocal in time. Although the formal solution of Eq. (195) is analogous to Eq. (191), it does not satisfy the composition law (192), a fact which represents a signature of memory effects.

Then a fundamental problem is to find those conditions on the memory kernel $\mathcal{K}(t)$ that ensure that the time evolution map $\Lambda(t, t_0)$ is CPTP, i.e., a quantum channel. Contrary to the Markovian case, a full characterization of legitimate memory kernels is still missing.

Chruściński and Kossakowski (2012) provided a class of memory kernels giving rise to legitimate quantum dynamics (quantum channels). The construction is based on a simple idea of normalization: starting from a family of (possibly non-trace-preserving) CPTP maps satisfying a certain additional condition one is able to "normalize" it in order to obtain a legitimate dynamics, i.e., a CPTP map. Non-Markovian master equations have also been described by Rybar *et al.* (2012) and Ciccarello, Palma, and Giovannetti (2013) by generalizing the collision models discussed previously and in Shabani and Lidar (2005) exploiting adaptive strategies that involve the measurements of the system environment followed by local transformations.

C. Markovian vs non-Markovian dynamics

Given a CPTP map, the problem of determining whether or not it admits an infinitesimal generator of the form (186) has been proven to be computationally hard (Wolf *et al.*, 2008; Cubitt, Eisert, and Wolf, 2012).

For CPTP maps that do not belong to $\mathfrak{M}$, a measure of non-Markovianity has been introduced by Wolf *et al.* (2008) in terms of the minimal amount of white noise $\mathcal{L}_\mu$ that has to be added in order to make $\log\Lambda + \mathcal{L}_\mu$ of the form (186).

Besides the Markovianity definition given in Sec. VII.A and the above-mentioned quantifier of (non)Markovianity other proposals have been put forward; see, e.g., Breuer, Laine, and Piilo (2009), Lu, Wang, and Sun (2010), Rivas, Huelga, and Plenio (2010), and Luo, Fu, and Song (2012).

On the one hand, Rivas, Huelga, and Plenio (2010) considered the equivalence between Markovian dynamics and infinitesimal divisibility and introduced a measure of deviation from it. Given a maximally entangled state $|\beta\rangle$ of the system of interest and a suitable ancillary system, due to the Choi-Jamiolkowski isomorphism (9), $\Lambda(t+\epsilon, t)$ is a CPTP map iff $[\Lambda(t+\epsilon,t)\otimes \mathrm{id}]|\beta\rangle\langle\beta| \geq 0$. Then one can consider $\|[\Lambda(t+\epsilon,t)\otimes \mathrm{id}]|\phi\rangle\langle\phi|\|_1$ as a measure of the non-CPTP character of $\Lambda(t+\epsilon,t)$. In fact, due to the trace-preserving property, this quantity equals 1 iff $\Lambda(t+\epsilon,t)$ is CPTP, otherwise it is greater than 1. Actually, the derivative of this quantity has been considered

$$g(t) \coloneqq \lim_{\epsilon\to 0}\frac{\|[\Lambda(t+\epsilon,t)\otimes \mathrm{id}]|\beta\rangle\langle\beta|\|_1 - 1}{\epsilon}. \tag{196}$$

It happens that $g(t) > 0$ iff the map Λ is indivisible.

On the other hand, Breuer, Laine, and Piilo (2009) used a fundamental property of CPTP maps, namely, the fact that they cannot increase the trace distance

$$D\big(\Lambda(t,0)(\rho_1), \Lambda(t,0)(\rho_1)\big) \leq D(\rho_1,\rho_2), \tag{197}$$

for any pair of states ρ_1, ρ_2. If a family of CPTP maps is infinitesimally divisible, the monotonicity of the trace distance holds true locally, that is,

$$\frac{d}{dt}D\big(\Lambda(t,0)(\rho_1), \Lambda(t,0)(\rho_1)\big) \leq 0. \tag{198}$$

According to that the dynamical map $\Lambda(t,0)$ is said to be non-Markovian if there exists a value of t such that Eq. (198) is violated, for some initial states ρ_1, ρ_2. Physically, this implies a temporal increase in the distinguishability of the two quantum states, a consequence of the backflow of information from the surrounding environment.

The criteria relying on Eqs. (196) and (198) allow one to define a computable measure of non-Markovianity. A natural quantifier derived from the criterion of Rivas, Huelga, and Plenio (2010) reads

$$\mathcal{N}_{\mathrm{RHP}}(\Lambda) = \frac{\int_0^\infty g(t)dt}{1+\int_0^\infty g(t)dt}, \tag{199}$$

where $g(t)$ is as in Eq. (196). From the criterion of Breuer, Laine, and Piilo (2009) one defines the non-Markovianity quantifier

$$\mathcal{N}_{\mathrm{BLP}}(\Lambda) = \sup_{\rho_1,\rho_2}\int \frac{d}{dt'}D\big(\Lambda(t',0)(\rho_1), \Lambda(t',0)(\rho_1)\big)|_{t'=t}dt, \tag{200}$$

where the integral is performed only for those t such that Eq. (198) is violated.

It was pointed out that the relation between these two criteria resembles that between separable and PPT states in entanglement theory (Chruściński, Kossakowski, and Rivas, 2011). Indeed, any family of CPTP maps which is Markovian according to the first criterion is as well Markovian according to the second one, that is, $\mathcal{N}_{\mathrm{RHP}}(\Lambda) = 0$ implies $\mathcal{N}_{\mathrm{BLP}}(\Lambda) = 0$, while the converse is in general not true. An example comparing nondivisibility and non-Markovianity, for the case of Gaussian channels, was recently discussed by Benatti, Floreanini, and Olivares (2012) while a test of non-Markovianity for these maps was discussed by Vasile, Maniscalco *et al.* (2011).

One of the few example of non-Markovian dynamics that are exactly solvable for their communication capacities is a single qubit coupled to an environment of noninteracting qubits in a star configuration giving rise to dephasing channel (Arshed, Toor, and Lidar, 2010). Its quantum capacity behavior as a function of time is strongly dependent on the coupling parameters and on the temperature of the bath. For generic values of these parameters, recurrence in the quantum capacity as a function of time is of small amplitude and quickly vanishes. On the contrary, for commensurable values of these parameters the quantum capacity becomes a periodic function of time. This feature indicates the backflow of information from the environment to the central spin: a signature of non-Markovian dynamics. This is also related to the increased distinguishability of states pointed out by the non-Markovianity criterion introduced by Breuer, Laine, and Piilo (2009).

VIII. SUMMARY AND OUTLOOK

In the last decades the subject of quantum channels has become prominent for its usefulness in foundational issues (Kraus, 1983) as well as in technological applications [see the latest striking experiments in quantum communication (Ma *et al.*, 2012; Yin *et al.*, 2013)]. Here this subject has been addressed using a broad approach that embraces memory effects. This is because the consideration of spatial and temporal memory effects is becoming increasingly pressing with the continuing miniaturization of devices and with increasing communication rates. In this scenario defining general properties and determining communication performance become daunting tasks. Hence, we mainly touched topics relevant to and witnessing progress toward these ends.

In the beginning (Sec. II) we reviewed basic features of quantum channel maps and tools for their characterization. Some physical examples of temporal and spatial evolutions of quantum systems, that the general framework of quantum (input-output) channels can describe, were also discussed. Then we focused on multiple channel uses by addressing their structural properties in Sec. III. There several quantum memory channels models were devised and their taxonomy presented. However, it is worth noting that the latter is based on channel representations that consider input and (initial) memory systems mapped onto output and (final) memory systems. In a black-box description accounting only for input to output mapping some of these models could result equivalent. Loosely speaking this could be analogous to the possibility of having different Kraus representations of the same quantum channel. Hence, a property of quantum channels (fixed point, indecomposability, etc.) should be defined in a more general way; that is, the channel has the property if there exists at least one of its memory representations that satisfies it.

Reliable communication through quantum channels can be achieved by employing error-correcting codes as discussed in Sec. IV. Standard quantum codes are designed to counteract independent errors affecting multiple uses of a noisy quantum channel. On the other hand, memory channels produce correlated errors. An extreme case is represented by collective errors affecting a certain number of information carriers at once. In such a case the symmetries of the noise usually allow for the existence of decoherence-free subspaces. In the intermediate situations one has to design new codes to counteract errors that are neither independent nor completely correlated. A relevant strategy to produce such codes is by concatenation of standard codes and decoherence-free subspaces. Another strategy consists of exploiting cyclic property of some codes. Finally a relation between convolutional codes and memory channels was highlighted.

The general definitions of classical and quantum capacities, unassisted as well as assisted by entanglement, were given in Sec. V followed by the definition of constrained capacities suitable for continuous channels. For such definitions we remarked that a superoperator norm approach can be used as well. Then we sketched coding theorems. Actually, concerning the capacities evaluation, the main obstacle is the restricted class of channels for which coding theorems are available. Hopefully this can be enlarged by resorting to stationary or ergodic properties of the quantum channels as outlined by Bjelaković and Boche (2009). Still within such a class, those channels leaving hopes for an exact capacity computation are the forgetful channels (see Sec. V.C.3). For this reason it is of utmost importance to derive general criteria to decide whether or not a given channel is forgetful. Beyond that it would be extremely interesting to establish when memory effects increase the capacity of a quantum channel. It is also worth noting that the effects of correlations among errors are in close connection with the property of superadditivity of the minimum output entropy (Hastings, 2009). The possible memory induced enhancement of the capacity of a quantum channel, looked through the dynamical memory model sketched in Sec. VI.B.1, can be seen as due to a sort of Zeno effect (Misra and Sudarshan, 1977). In fact, by frequently inserting information carriers through the channel one prevents the environment from coming back to its stationary state after the passage of each of them, thus less affecting the carriers themselves. Whereas in the case of quantum channels arising from non-Markovian dynamics like that of Sec. VII, the increment of capacity can be explained by the backflow of information from environment to system.

Known solvable (in terms of capacities) models were discussed in Sec. VI. Among them, dephasing memory channels possess features related to many-body physics and lossy bosonic memory channels show water-filling phenomena similar to fluid mechanics.

Finally, in Sec. VII, we showed the conditions under which a quantum channel can be "divided" into the concatenation of other quantum channels, i.e., its action results as the composition of other quantum channels. This possibility is closely related to quantum channels intended as dynamical maps. Then one can distinguish between Markovian and non-Markovian dynamics, the latter showing memory effects in time. As a consequence we briefly accounted for some measures quantifying deviation from Markovian dynamics, although a general consensus on that subject is not yet reached.

All in all examples of quantum channels showing memory effects are abundant in quantum information processing. An unmodulated spin chain was proposed as a model for short distance quantum communication (Bose, 2003). In such a scheme, the state to be communicated over the channel is placed on one of the spins of the chain, propagates for a specific amount of time, and is then received at a distant spin of the chain. When viewed as a model for quantum communication, it is generally assumed that a reset of the spin chain occurs after each signal, for instance, by applying an external magnetic field, resulting in a memoryless channel. However, a continuous operation without resetting corresponds to a quantum channel with memory (Bayat *et al.*, 2008). Another model of a quantum channel with memory is the so-called one-atom maser or micromaser (Benenti, D'Arrigo, and Falci, 2009). In such a device, excited atoms interact with the photon field inside a high-quality optical cavity. If the photons inside the cavity have sufficiently long lifetime, atoms entering the cavity will feel the effect of the preceding atoms, introducing ISI correlations (see Sec. III.B) among consecutive signal states.

Another source of correlated noise in the propagation of the electromagnetic field is due to atmospheric turbulence, whose effects on the signal propagation can be modeled as random changes of the channel's characteristics (Semenov and Vogel, 2009, 2010). Moreover, the decoherence induced by atmospheric turbulence introduces cross talks (Tyler and Boyd, 2009; Boyd *et al.*, 2011), i.e., ISI correlations (see Sec. III.B), when information is encoded in the transverse degrees of freedom of the electromagnetic field, e.g., the orbital angular momentum. Furthermore, the propagation of the quantum electromagnetic field in linear dispersive media, including the free-space propagation and through linear optical systems, can be described by a quantum channel with memory (Giovannetti *et al.*, 2004b; Shapiro, 2009; Memarzadeh and Mancini, 2010; Lupo, Giovannetti *et al.*, 2011; Lupo *et al.*, 2012), where wave diffraction introduces memory effects.

Memory effects also arise in the context of quantum cryptography. Quite generally, one can categorize the collective attacks within the framework of memoryless channels, while the coherent attacks within the memory channels framework (Gisin *et al.*, 2002; Scarani *et al.*, 2009). However, this link has been subjected to limited attention and probably needs further exploration. Actually, in one-way quantum key distribution memory effects that introduce correlations among transmitted symbols can give an advantage to the eavesdropper (Ruggeri and Mancini, 2007b). Only if the legitimate users have the control of the noise correlations, by properly tuning them, can they reduce eavesdropper information (Vasile, Olivares *et al.*, 2011). Instead, in two-way quantum key distribution checking the presence or absence of noise correlations can help in counteracting eavesdropper attacks (Pirandola *et al.*, 2008). Then an analysis of memory effects in other channel uses configurations, like zero-error channel capacity, channels with feedback, channels with unknown parameters and multiuser channels, should be pursued.

Finally, moving to the framework of time-continuous quantum evolution, non-Markovian effects are relevant in several physical systems characterized by the interaction with a structured environment. Examples are in the framework of solid-state physics, as quantum dots in photonic crystals (Vats, John, and Busch, 2002; Madsen *et al.*, 2011), and in the soft matter framework as the case of exciton dynamics surrounded by their protein environment (Plenio and Huelga, 2008; Caruso *et al.*, 2009; Rebentrost, Chakraborty, and Aspuru-Guzik, 2009; Thorwart *et al.*, 2009; Caruso, Huelga, and Plenio, 2010).

In summary, more efforts are needed to gain a full understanding of quantum channels; however, the presented work constitutes a rather general frame where the still missing pieces of the puzzle could be settled.

ACKNOWLEDGMENTS

The authors would like to thank K. Banaszek, N. J. Cerf, D. Chruscinski, N. Datta, S. F. Huelga, C. Macchiavello, O. Pilyavets, M. B. Plenio, A. Shabani, B. Vacchini, S. Virmani, and M. Wilde for useful discussions and for a careful reading of the manuscript. S. M. acknowledges a fruitful and enjoyable collaboration with Garry Andrew Bowen through whom he became acquainted with the subject of quantum channels. The work of F. C. has been supported by a Marie Curie Intra European Individual Fellowship and a Marie Curie Career Integration Grant within the 7th European Community Framework Programme, under the grant agreements MemoryQuantumICT No. 235086 and QuantumBioTech No. 293449, respectively, and by the Future in Research (FIRB) Programme of the Italian Ministry of Education, University and Research (MIUR), under the FIRB-MIUR grant agreement No. RBFR10M3SB. The work of C. L. and S. M. has been supported by the Future and Emerging Technologies (FET) programme with the Seventh Framework Programme for Research of the European Commission, under the FET-Open grant agreement CORNER, No. FP7-ICT-213681.

APPENDIX A: DISTANCE MEASURES

A proper way to measure the distance between two states $\rho_1, \rho_2 \in \mathfrak{S}(\mathcal{H}_Q)$ of a quantum system Q, is provided by the trace distance defined as

$$D(\rho_1,\rho_2) \coloneqq \tfrac{1}{2}\|\rho_1 - \rho_2\|_1, \tag{A1}$$

with $\|O\|_1 \coloneqq \mathrm{Tr}\sqrt{O^\dagger O}$ being the trace norm of the operator O (Nielsen and Chuang, 2000; Wilde, 2013). While fulfilling all the conditions of a regular distance (i.e., positivity, symmetry, and triangular inequality) the trace distance possesses other interesting properties which makes it operationally well defined. For instance, it is bounded between 0 and 1 (reaching the latter value only when ρ_1 and ρ_2 have orthogonal support). Furthermore, the trace distance is preserved under unitary transformations, i.e., $D(U\rho_1 U^\dagger, U\rho_2 U^\dagger) = D(\rho_1,\rho_2)$ (implying that the distance between physical states does not depend upon the coordinate system used to describe them) but it is contractive under CPTP maps Φ, i.e.,

$$D\big(\Phi(\rho_1),\Phi(\rho_2)\big) \le D(\rho_1,\rho_2) \tag{A2}$$

(implying that the action of noise tends to blur the difference among states). Finally $D(\rho_1,\rho_2)$ can be identified with the maximum distance between the statistical distributions $\{p_x(\rho_1) = \mathrm{Tr}[E_x\rho_1]\}_{x\in X}$ and $\{p_x(\rho_2) = \mathrm{Tr}[E_x\rho_2]\}_{x\in X}$ obtained by performing the same POVM measurement $\{E_x\}_{x\in X}$ on ρ_1 and ρ_2.

Another quantity useful to gauge how close two density matrices ρ_1 and ρ_2 are, is the fidelity (Uhlmann, 1976; Jozsa, 1994)

$$F(\rho_1,\rho_2) \coloneqq \|\rho_1^{1/2}\rho_2^{1/2}\|_1^2, \tag{A3}$$

which for ρ_1 being rank 1, i.e., $\rho_1 = |\psi_1\rangle\langle\psi_1|$, coincides with the probability of finding ρ_2 in the vector $|\psi_1\rangle$, i.e.,

$$F(|\psi_1\rangle,\rho_2) = \langle\psi_1|\rho_2|\psi_1\rangle. \tag{A4}$$

The function $F(\rho_1,\rho_2)$ is symmetric [i.e., $F(\rho_1,\rho_2) = F(\rho_2,\rho_1)$] and always in the range [0, 1] (equal to 1 if and only if $\rho_1 = \rho_2$ and vanishing for density operators with orthogonal supports, e.g., for orthogonal pure states). Furthermore, F is invariant under the action of a unitary evolution $F(U\rho_1 U^\dagger, U\rho_2 U^\dagger) = F(\rho_1,\rho_2)$, and increasing under the CPTP map,

$$F\big(\Phi(\rho_1),\Phi(\rho_2)\big) \ge F(\rho_1,\rho_2). \tag{A5}$$

While not a distance itself, the fidelity is directly linked to the Bures distance (Bures, 1969) via the identity $D_B(\rho_1,\rho_2) \coloneqq [2 - 2\sqrt{F(\rho_1,\rho_2)}]^{1/2}$. Trace distance and fidelity are related by the following inequalities:

$$1 - \sqrt{F(\rho_1,\rho_2)} \le D(\rho_1,\rho_2) \le \sqrt{1 - F(\rho_1,\rho_2)}, \tag{A6}$$

therefore states which have high values of fidelity are also close in trace distance, and vice versa.

APPENDIX B: QUASILOCAL ALGEBRAS

Quasilocal algebras are the proper mathematical tools to describe infinitely extended quantum lattice systems (Bratteli and Robinson, 1979). For the sake of simplicity we consider a chain of infinitely many qubits (spins) placed in a one-dimensional lattice $\mathbb{Z}$. Then, to each lattice site $j \in \mathbb{Z}$ attach a Hilbert space $\mathcal{H}_j \simeq \mathbb{C}^2$ and consider the associated algebra of bounded operators $\mathcal{A}_j = \mathcal{B}(\mathcal{H}_j)$. If one restricts to a finite part of the chain, say a set $\Lambda \subset \mathbb{Z}$, it is possible to define the following tensor products:

$$\mathcal{H}_\Lambda = \otimes_{j\in\Lambda}\mathcal{H}_j, \qquad \mathcal{A}^\Lambda = \mathcal{B}(\mathcal{H}_\Lambda) = \otimes_{j\in\Lambda}\mathcal{A}_j. \tag{B1}$$

Operators $a \in \mathcal{A}^\Lambda$ are called local operators as they are operators "localized" in Λ. Clear examples of local observables are the Pauli operators $\sigma_{x,j}$, $\sigma_{y,j}$, and $\sigma_{z,j}$, attached to the site $j \in \mathbb{Z}$. However, if the region Λ is infinite the situation becomes tricky. In fact, although it is still possible to attach a Hilbert space $\mathcal{H}$ to the whole chain such that local operators σ_j act on $\mathcal{H}$, this cannot be done in a unique way. It is then preferable to proceed by considering the algebraic properties of local operators not requiring a global Hilbert space $\mathcal{H}$. To this end, we first notice that having two finite regions Λ, Λ' such that $\Lambda \subset \Lambda' \subset \mathbb{Z}$, the operators

$$a \in \mathcal{A}^\Lambda \quad \text{and} \quad a \otimes \mathbb{1}_{\Lambda'\setminus\Lambda} \in \mathcal{A}^{\Lambda'} \tag{B2}$$

describe the same physical object. Therefore, one can identify $\mathcal{A}^\Lambda$ with the subalgebra $\mathcal{A}^\Lambda \otimes \mathbb{1}_{\Lambda'\setminus\Lambda}$ of $\mathcal{A}^{\Lambda'}$ through the map

$$\mathcal{A}^\Lambda \ni a \mapsto a \otimes \mathbb{1}_{\Lambda'\setminus\Lambda} \in \mathcal{A}^{\Lambda'}. \tag{B3}$$

Doing so one gets a system of matrix algebras $\mathcal{A}^\Lambda$ which are ordered by inclusion

$$\Lambda \subset \Lambda' \Rightarrow \mathcal{A}^\Lambda \subset \mathcal{A}^{\Lambda'}, \qquad \forall \text{ finite } \Lambda, \Lambda' \subset \mathbb{Z}. \tag{B4}$$

This construction leads to the possibility of defining, for two arbitrary local operators a_1, a_2,

(i) their linear combination $\mu_1 a_1 + \mu_2 a_2$ with $\mu_1, \mu_2 \in \mathbb{C}$;
(ii) their product $a_1 a_2$; and
(iii) their adjoints $a_1^\dagger$ (respectively, $a_2^\dagger$).

To this end one needs only to find a region Λ such that the matrix algebra $\mathcal{A}^\Lambda$ contains both operators a_1, a_2.

More precisely one can introduce the space of all local operators by the union

$$\mathcal{A}^{\rm loc} := \bigcup_{\text{finite } \Lambda\subset\mathbb{Z}} \mathcal{A}^\Lambda, \tag{B5}$$

and then equip this space with a vector space structure, a product (associative and bilinear), a † operation and a unit element $\mathbb{1}$. In such a way $\mathcal{A}^{\rm loc}$ becomes the algebra of local observables.

Going further on, one can associate to each $a \in \mathcal{A}^{\rm loc}$ its norm $\|a\|$ by finding the region Λ such that $a \in \mathcal{A}^\Lambda$ and then using the standard operator norm. A problem is given by the fact that $\mathcal{A}^{\rm loc}$ is not complete with respect to such a norm, i.e., not all Cauchy sequences converge. This problem can be solved by taking the norm closure of $\mathcal{A}^{\rm loc}$ and get the algebra of quasilocal observables

$$\mathcal{A}^\mathbb{Z} = \overline{\mathcal{A}^{\rm loc}}^{\|\cdot\|}. \tag{B6}$$

Here quasilocal stands for the fact that $\mathcal{A}^\mathbb{Z}$ besides all local observables, also contains nonlocal observables which can be approximated in norm by local ones. $\mathcal{A}^\mathbb{Z}$ results in a C^* algebra and its elements can be regarded in many respects as bounded operators (Bratteli and Robinson, 1979).

It is also useful to consider methods to transform abstract elements $a \in \mathcal{A}^\mathbb{Z}$ into operators $\pi(a)$ acting on a Hilbert space $\mathcal{H}$, i.e., representations of quasilocal algebra. A representation π on Hilbert space $\mathcal{H}$ is a homomorphism $\pi\colon \mathcal{A}^\mathbb{Z} \to \mathcal{B}(\mathcal{H})$, i.e., a linear map satisfying $\pi(ab) = \pi(a)\pi(b)$ and $\pi(a^\dagger) = \pi(a)^\dagger$. Unfortunately for spin chains there is not a unique representation that can be used for all purposes, but one has to choose the representation which is most appropriate to the given physical context. This ambiguity also reflects on the definition of states. In fact, one might be tempted to use density operators ρ on the Hilbert space $\mathcal{H}$ which carries the representation π. However, since different representations correspond to different physical contexts one should use all possible representations (in fact, each density operator in any representation can describe a state). Clearly, it would be much better to describe states in a way independent from the representation. Thus a state of $\mathcal{A}^\mathbb{Z}$ is defined as a linear functional $\psi\colon \mathcal{A}^\mathbb{Z} \to \mathbb{C}$ which is positive $[\psi(a^\dagger a) \geq 0, \forall a \in \mathcal{A}]$ and normalized $[\psi(\mathbb{1}) = 1]$. This means that given a representation π and a density operator ρ on $\mathcal{H}$, the corresponding state is the functional $\psi_\rho(a) = \mathrm{tr}\{\pi(a)\rho\}$.

The possibility of finding for each state ψ a Hilbert space $\mathcal{H}$ carrying a representation π and a density operator ρ such that $\psi = \psi_\rho$ is guaranteed by the Gelfand-Naimark-Segal theorem (Bratteli and Robinson, 1979). It states that each state ψ can be represented by a state vector $|v_\psi\rangle$ on a suitable Hilbert space. In other words, we like to say that it is always possible to provide a "purification" of the state ψ.

One can introduce into the quasilocal algebra $\mathcal{A}^\mathbb{Z}$ a shift operation $T\colon \mathcal{A}^\mathbb{Z} \to \mathcal{A}^\mathbb{Z}$ by the following action:

$$\mathcal{A}^\Lambda \ni a \simeq a \otimes \mathbb{1}_\mathcal{A} \mapsto T(a) := \mathbb{1}_\mathcal{A} \otimes a \simeq a \in \mathcal{A}^{\Lambda+1}, \tag{B7}$$

where $A \otimes \mathbb{1}_\mathcal{A}$ (respectively, $\mathbb{1}_\mathcal{A} \otimes A$) stands for the tensor product between A belonging to $\mathcal{A}^\Lambda$ and the identity of $\mathcal{A}$ on the site to the right of Λ (respectively, between the identity of $\mathcal{A}$ on the site to the left of Λ and a belonging to $\mathcal{A}^\Lambda$).

Moving from the action of the shift T, it is possible to introduce the notion of stationary state ψ on $\mathcal{A}^\mathbb{Z}$ when $\psi \circ T = \psi$ holds true. The set of stationary states on $\mathcal{A}^\mathbb{Z}$ turns out to be convex. Then a state ψ on $\mathcal{A}^\mathbb{Z}$ is called ergodic (with respect to the shift) if it is extremal on this set.

APPENDIX C: DECOMPOSITION FOR NONANTICIPATORY QUANTUM CHANNELS

This Appendix provides an explicit derivation of the decomposition of nonanticipatory channels in Eq. (80) based on a generalization of analysis presented in Beckman et al. (2001), Eggeling, Schlingemann, and Werner (2002), Kretschmann and Werner (2005), and Piani et al. (2006).

Let then $\{\Phi^{(n)}; n = 1, 2, \ldots\}$ be a family of CPTP maps describing a nonantipatory quantum channel. Adopting the unitary representation in Eq. (7), for each n one can define a unitary transformation $W^{(n)}_{q_n,q_{n-1},\ldots,q_1,M}$ coupling the first n carriers to a common environment M which allows one to write

$$\Phi^{(n)}(\rho_Q^{(n)}) = \mathrm{Tr}_M[W^{(n)}_{q_n,q_{n-1},\ldots,q_1,M} \times (\rho_Q^{(n)} \otimes \omega_M^{(n)}) W^{(n)\dagger}_{q_n,q_{n-1},\ldots,q_1,M}]. \tag{C1}$$

The environmental state $\omega_M^{(n)}$ is in general a function of n but is independent on the input state $\rho_Q^{(n)}$. Without loss of generality here it is assumed to be a pure state $\omega_M^{(n)} = |\omega^{(n)}\rangle_M\langle\omega^{(n)}|$.

In general the unitary couplings $W^{(n)}_{q_n,q_{n-1},\ldots,q_1,M}$ can have a complicated dependence upon n: however, since the channel is nonanticipatory they must obey the following rule:

$$\mathrm{Tr}_{q_n,M}[W^{(n)}_{q_n,\ldots,q_1,M}(\rho_Q^{(n)}\otimes\omega_M^{(n)})W^{(n)\dagger}_{q_n,\ldots,q_1,M}] = \mathrm{Tr}_M[W^{(n-1)}_{q_{n-1},\ldots,q_1,M}(\rho_Q^{(n-1)}\otimes\omega_M^{(n-1)})W^{(n-1)\dagger}_{q_{n-1},\ldots,q_1,M}],$$

where $\rho_Q^{(n-1)} = \mathrm{Tr}_{q_n}[\rho_Q^{(n)}]$ is the reduced density operator of $\rho_Q^{(n)}$ associated with the first $n-1$ carriers. Applying this relation to a pure input state $\rho_Q^{(n)}$ of the form $|\psi\rangle_{q_n}\otimes|\phi\rangle_{q_{n-1},\ldots,q_1}$ one notes that the vectors $W^{(n)}_{q_n,q_{n-1},\ldots,q_1,M}|\psi\rangle_{q_n}\otimes|\phi\rangle_{q_{n-1},\ldots,q_1}\otimes|\omega^{(n)}\rangle_M$ and $W^{(n-1)}_{q_{n-1},\ldots,q_1,M}|\phi\rangle_{q_{n-1},\ldots,q_1}\otimes|\omega^{(n-1)}\rangle_M$ are both purifications of the state $\Phi^{(n-1)}(\rho_Q^{(n-1)})$. Therefore there must exist a unitary transformation U_{q_nM} acting on the latter which satisfies the identity (Nielsen and Chuang, 2000)

$$W^{(n)}_{q_n,q_{n-1},\ldots,q_1,M}|\omega^{(n)}\rangle_M = U_{q_nM}W^{(n-1)}_{q_{n-1},\ldots,q_1,M}|\omega^{(n-1)}\rangle_M. \quad \text{(C2)}$$

Iterating this n times yields

$$W^{(n)}_{q_n,q_{n-1},\ldots,q_1,M}|\omega^{(n)}\rangle_M = U_{q_nM}U_{q_{n-1}M}\cdots U_{q_1M}|\omega^{(0)}\rangle_M, \quad \text{(C3)}$$

which replaced into Eq. (C1) implies Eq. (80).

APPENDIX D: EXPLICIT DERIVATION OF CAPACITY UPPER BOUNDS

Here we present an explicit derivation of the upper bounds (158) and (159) for the classical and quantum capacities defined in Eq. (145) of a (non-necessarily memoryless) quantum channel $\Phi^{(n)}$.

1. Upper bound for the classical capacity

The derivation of Eq. (158) follows by merging the Holevo bound (Holevo, 1973a, 1973b) with the classical Fano inequality (Cover and Thomas, 1991). For this purpose one is reminded that given two random variables X and Y connected by conditional probability distribution $p(y|x)$ and a correspondence rule which assigns values of X to each of the values of Y, the Fano inequality allows one to lower bound the mutual information $I(X:Y)$ as

$$I(X:Y) \geq H(X) - h(P_e) - P_e\log_2(|X|-1), \quad \text{(D1)}$$

where $|X|$ is the number of elements of the variable X, $h(p)$ is the binary Shannon entropy (66), and finally P_e is the average error probability that the correspondence rule is violated by the conditional probability $p(y|x)$. Specifically, assuming for simplicity that the X and Y span the same alphabet of symbols and that the correspondence rule assigned to Y the same symbol on X, we have $P_e = 1 - \sum_x p(x)p(y=x|x)$. Identify then X with the messages m that Alice is mapping from $\mathcal{M}$ to the n carriers via the coding channel $\Phi_E^{(k\to n)}$, and with Y the elements of $\mathcal{M}$ which Bob is retrieving via his decoding mapping $\Phi_D^{(n\to k)}$. Under the assumption that the probability that Alice is selecting the messages from $\mathcal{M}$ with uniform probability, and reminding one that $\mathcal{M}$ contains 2^k elements, Eq. (D1) yields

$$\begin{aligned} I(X:Y) &\geq k - h(P_e) - P_e\log_2(2^k-1) \\ &\geq k - h(\epsilon) - \epsilon k, \end{aligned} \quad \text{(D2)}$$

where we used the fact that $P_e \leq 1 - F_{\min} < \epsilon$, with $F_{\min}$ being the minimum fidelity achieved by the selected code. By reorganizing the various terms and by dividing by n we then get

$$\frac{k}{n} \leq \frac{I(X:Y)}{(1-\epsilon)n} + \frac{h(\epsilon)}{(1-\epsilon)n}. \quad \text{(D3)}$$

Remember next that $I(X:Y)$ is the mutual information associated with the ensemble of code words $\mathcal{E} = \{p_m = 2^{-k}; \Phi_E^{(k\to n)}(m)\}$ generated by Alice and received by Bob through the channel $\Phi^{(n)}$. The Holevo bound (72) implies then

$$\begin{aligned} \frac{k}{n} &\leq \frac{\chi(\mathcal{E};\Phi^{(n)})}{(1-\epsilon)n} + \frac{h(\epsilon)}{(1-\epsilon)n} \\ &\leq \frac{\max_{\mathcal{E}}\chi(\mathcal{E};\Phi^{(n)})}{(1-\epsilon)n} + \frac{h(\epsilon)}{(1-\epsilon)n}. \end{aligned} \quad \text{(D4)}$$

Since this inequality holds for all encoding and decoding strategies entering in the capacity definition (145), by taking the limits on $k\to\infty$ and $\epsilon\to 0$ we finally get the inequality (158).

The same derivation detailed above can be used to prove the converse part of the HSW theorem for the one-shot classical capacity $C_1(\Phi)$ of a memoryless channel Φ, i.e.,

$$C_1(\Phi) \leq \max_{\mathcal{E}}\chi(\mathcal{E};\Phi). \quad \text{(D5)}$$

Indeed, exploiting the fact that the channel is memoryless and the coding procedure uses only separable code words $\Phi_E^{(k\to n)}(m)$ one can apply the subadditivity of the classical mutual information (Gallager, 1968; Cover and Thomas, 1991) to replace Eq. (D2) as

$$\begin{aligned} \sum_{i=1}^{n} I(X_i:Y_i) &\geq k - h(P_e) - P_e\log_2(2^k-1) \\ &\geq k - h(\epsilon) - \epsilon k, \end{aligned} \quad \text{(D6)}$$

where, for $i=1,\ldots,n$, $I(X_i:Y_i)$ is the mutual information associated with the classical input of the ith channel use. Applying the Holevo bound to all these terms we get

$$\frac{k}{n} \leq \frac{\max_{\mathcal{E}}\chi(\mathcal{E};\Phi)}{(1-\epsilon)} + \frac{h(\epsilon)}{(1-\epsilon)n}, \quad \text{(D7)}$$

that finally leads to Eq. (D5) when taking $\epsilon\to 0$.

In conclusion, it is worth remarking that Eq. (158) can be used to prove an inequality which in some cases happens to be useful for deriving an explicit expression for C; see, e.g., Sec. VI.A.1. This is obtained by noting that

$$\max_{\mathcal{E}} \chi(\mathcal{E}; \Phi^{(n)}) \leq \max_{\rho} S[\Phi^{(n)}(\rho)] - \min_{\rho} S[\Phi^{(n)}(\rho)] \leq n \log_2 d - S_{\min}(\Phi^{(n)}), \tag{D8}$$

with $S_{\min}(\Phi^{(n)}) = \min_{\rho} S[\Phi^{(n)}(\rho)]$ being the minimum entropy one can reach at the output of the channel. In the above derivation the first inequality follows directly from the definition of the Holevo information, while the second from the fact that the entropy of a state of $\mathcal{H}_Q^{\otimes n}$ is not larger than $\log_2 d^n$ (d being the dimension of $\mathcal{H}_Q$). Replacing this into Eq. (158) we then get

$$C \leq \log_2 d - \lim_{n\to\infty} \frac{S_{\min}(\Phi^{(n)})}{n}. \tag{D9}$$

2. Upper bound for the quantum capacity

The derivation which follows is an adaptation to the memory channel scenario of the proof of Barnum, Nielsen, and Schumacher (1998). The starting point in this case is the quantum Fano inequality presented in Eq. (65), the data-processing inequality (70) and the bounds (63). From them one can easily verify that given a generic density matrix τ of the reference set $\mathcal{M}$ the following relation holds:

$$\begin{aligned}\frac{S(\tau)}{n} &\leq \frac{J(\tau; \Phi_D^{(n\to k)} \circ \Phi^{(n)} \circ \Phi_E^{(k\to n)})}{n} \\ &\quad + \frac{2}{n} h(F_e(\tau; \Phi_D^{(n\to k)} \circ \Phi^{(n)} \circ \Phi_E^{(k\to n)}) \\ &\quad + \frac{4}{n}[1 - F_e(\tau; \Phi_D^{(n\to k)} \circ \Phi^{(n)} \circ \Phi_E^{(k\to n)})] \log_2(4^k - 1) \\ &\leq \frac{\max_\rho J(\rho; \Phi^{(n)})}{n} + \frac{2}{n} h(1 - 3\epsilon/2) + \frac{6}{n}\epsilon k,\end{aligned} \tag{D10}$$

where the second inequality has been obtained by using again the data-processing inequality (70) and maximizing J over all possible inputs states ρ of the n channels uses, and by using the fact that the minimum fidelity of the code is lower bounded by $1 - \epsilon$ and the fact that this implies that the corresponding entanglement fidelity fulfills $F_e(\tau; \Phi_D^{(n\to k)} \circ \Phi^{(n)} \circ \Phi_E^{(k\to n)}) > 1 - 3\epsilon/2$; see Eq. (53). Specifying Eq. (D10) for the maximally mixed state of $\mathcal{M}$, we then get

$$\frac{k}{n} \leq \frac{\max_\rho J(\rho; \Phi^{(n)})}{n} + \frac{2}{n} h(1 - 3\epsilon/2) + \frac{6}{n}\epsilon k,$$

which in the limit of $k \to \infty$ and $\epsilon \to 0$ yields finally Eq. (159).

REFERENCES

Abeyesinghe, A., I. Devetak, P. Hayden, and A. Winter, 2009, Proc. R. Soc. A **465**, 2537.

Abeyesinghe, A., and P. Hayden, 2003, Phys. Rev. A **68**, 062319.

Adami, C., and N. J. Cerf, 1997, Phys. Rev. A **56**, 3470.

Adesso, G., and M. Ericsson, 2006, Phys. Rev. A **74**, 030305(R).

Ahlswede, R., 1968, Z. Wahrsch. Verw. Geb. **11**, 61.

Alber, G., Th. Beth, M. Horodecki, P. Horodecki, R. Horodecki, M. Rötteler, H. Weinfurter, R. Werner, and A. Zeilinger, 2001, *Quantum Information*, Springer Tracts in Modern Physics (Springer-Verlag, Berlin).

Alicki, R., and K. Lendi, 1987, *Quantum Dynamical Semigroups and Applications*, Lecture Notes in Physics (Springer-Verlag, Berlin).

Amosov, G. G., 2007, J. Math. Phys. (N.Y.) **48**, 012104.

Amosov, G. G., and A. S. Holevo, 2000, arXiv:math-ph/0103015.

Amosov, G. G., and S. Mancini, 2009, Quantum Inf. Comput. **9**, 0594.

Arikan, E., 2009, IEEE Trans. Inf. Theory **55**, 3051.

Arimoto, S., 1973, IEEE Trans. Inf. Theory **19**, 357.

Arshed, N., A. H. Toor, and D. A. Lidar, 2010, Phys. Rev. A **81**, 062353.

Ashikhmin, A., and E. Knill, 2001, IEEE Trans. Inf. Theory **47**, 3065.

Aubrun, G., S. Szarek, and E. Werner, 2011, Commun. Math. Phys. **305**, 85.

Ball, J., and K. Banaszek, 2005, Open Syst. Inf. Dyn. **12**, 121.

Ball, J., A. Dragan, and K. Banaszek, 2004, Phys. Rev. A **69**, 042324.

Ban, M., M. Sasaki, and M. Takeoka, 2002, J. Phys. A **35**, L401.

Banaszek, K., A. Dragan, W. Wasilewski, and C. Radzewicz, 2004, Phys. Rev. Lett. **92**, 257901.

Barnum, H., E. Knill, and M. A. Nielsen, 2000, IEEE Trans. Inf. Theory **46**, 1317.

Barnum, H., M. A. Nielsen, and B. Schumacher, 1998, Phys. Rev. A **57**, 4153.

Barnum, H., J. A. Smolin, and B. M. Terhal, 1998, Phys. Rev. A **58**, 3496.

Barreto Lemos, G., and G. Benenti, 2010, Phys. Rev. A **81**, 062331.

Bartlett, S. D., T. Rudolph, and R. W. Spekkens, 2003, Phys. Rev. Lett. **91**, 027901.

Bayat, A., D. Burgarth, S. Mancini, and S. Bose, 2008, Phys. Rev. A **77**, 050306(R).

Beckman, D., D. Gottesman, M. A. Nielsen, and J. Preskill, 2001, Phys. Rev. A **64**, 052309.

Belavkin, V. P., G. M. D' Ariano, and M. Raginsky, 2005, J. Math. Phys. (N.Y.) **46**, 062106.

Benatti, F., R. Floreanini, and S. Olivares, 2012, Phys. Lett. A **376**, 2951.

Benenti, G., A. D'Arrigo, and G. Falci, 2009, Phys. Rev. Lett. **103**, 020502.

Benenti, G., A. D'Arrigo, and G. Falci, 2012, Eur. Phys. J. D **66**, 147.

Bengtsson, I., and K. Zyczkowski, 2006, *Geometry of Quantum States* (Cambridge University Press, Cambridge, England).

Bennett, C. H., G. Brassard, R. Jozsa, A. Peres, and W. Wootters, 1993, Phys. Rev. Lett. **70**, 1895.

Bennett, C. H., I. Devetak, A. W. Arrow, P. W. Shor, and A. Winter, 2009, arXiv:0912.5537.

Bennett, C. H., I. Devetak, P. W. Shor, and J. A. Smolin, 2006, Phys. Rev. Lett. **96**, 150502.

Bennett, C. H., D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. Wootters, 1999, Phys. Rev. A **59**, 1070.

Bennett, C. H., D. P. DiVincenzo, and J. A. Smolin, 1997, Phys. Rev. Lett. **78**, 3217.

Bennett, C. H., D. P. DiVincenzo, J. A. Smolin, and W. Wootters, 1996, Phys. Rev. A **54**, 3824.

Bennett, C. H., and P. W. Shor, 1998, IEEE Trans. Inf. Theory **44**, 2724.

Bennett, C. H., P. W. Shor, J. A. Smolin, and A. V. Thapliyal, 1999, Phys. Rev. Lett. **83**, 3081.

Bennett, C. H., P. W. Shor, J. A. Smolin, and A. V. Thapliyal, 2002, IEEE Trans. Inf. Theory **48**, 2637.

Bennett, C. H., and S. J. Wiesner, 1992, Phys. Rev. Lett. **69**, 2881.

Berta, M., F. G. S. L. Brandão, M. Christandl, and S. Wehner, 2013, IEEE Trans. Inf. Theory **59**, 6779.

Berta, M., M. Christandl, and R. Renner, 2011, Commun. Math. Phys. **306**, 579.

Berta, M., J. M. Renes, and M. M. Wilde, 2013, arXiv:1301.1594.

Bjelaković, I., and H. Boche, 2008, IEEE Trans. Inf. Theory **54**, 723.

Bjelaković, I., and H. Boche, 2009, IEEE Trans. Inf. Theory **55**, 3360.

Bjelaković, I., H. Boche, and J. Nötzel, 2009, Commun. Math. Phys. **292**, 55.

Blackwell, D., L. Breiman, and A. J. Thomasian, 1958, Ann. Math. Stat. **29**, 1209.

Blackwell, D., L. Breiman, and A. J. Thomasian, 1959, Ann. Math. Stat. **30**, 1229.

Bose, S., 2003, Phys. Rev. Lett. **91**, 207901.

Bowen, G., 2004, IEEE Trans. Inf. Theory **50**, 2429.

Bowen, G., 2005, Int. J. Quantum. Inform. **03**, 123.

Bowen, G., and S. Bose, 2001, Phys. Rev. Lett. **87**, 267901.

Bowen, G., I. Devetak, and S. Mancini, 2005, Phys. Rev. A **71**, 034310.

Bowen, G., and S. Mancini, 2004, Phys. Rev. A **69**, 012306.

Boyd, R. W., B. Rodenburg, M. Mirhosseini, and S. M. Barnett, 2011, Opt. Express **19**, 18 310.

Brandão, F. G. S. L., and M. Horodecki, 2010, Open Syst. Inf. Dyn. **17**, 31.

Brandão, F. G. S. L., J. Oppenheim, and S. Strelchuk, 2012, Phys. Rev. Lett. **108**, 040501.

Bratteli, O., and D. W. Robinson, 1979, *Operators Algebras and Quantum Statistical Mechanics* (Springer, Berlin).

Braunstein, S. L., and H. J. Kimble, 1998, Phys. Rev. Lett. **80**, 869.

Braunstein, S. L., and P. van Loock, 2005, Rev. Mod. Phys. **77**, 513.

Breuer, H. P., E. M. Laine, and J. Piilo, 2009, Phys. Rev. Lett. **103**, 210401.

Breuer, H. P., and F. Petruccione, 2002, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford).

Brun, T., I. Devetak, and M.-H. Hsieh, 2006, Science **314**, 436.

Bures, D., 1969, Trans. Am. Math. Soc. **135**, 199.

Burgarth, D., G. Chiribella, V. Giovannetti, P. Perinotti, and K. Yuasa, 2013, New J. Phys. **15**, 073045.

Burgarth, D., and V. Giovannetti, 2007, New J. Phys. **9**, 150.

Buscemi, F., and N. Datta, 2010, IEEE Trans. Inf. Theory **56**, 1447.

Cafaro, C., L'Innocente, S., C. Lupo, and S. Mancini, 2011, Open Syst. Inf. Dyn. **18**, 1.

Cafaro, C., and S. Mancini, 2010a, Phys. Lett. A **374**, 2688.

Cafaro, C., and S. Mancini, 2010b, Phys. Rev. A **82**, 012306.

Cafaro, C., and S. Mancini, 2011, Int. J. Quantum. Inform. **09**, 309.

Cai, N., A. Winter, and R. W. Yeung, 2004, Probl. Inf. Transm. **40**, 318.

Calderbank, A. R., E. M. Rains, P. Shor, and N. J. A. Sloane, 1998, IEEE Trans. Inf. Theory **44**, 1369.

Calderbank, A. R., E. M. Rains, P. W. Shor, and N. J. A. Sloane, 1997, Phys. Rev. Lett. **78**, 405.

Calderbank, A. R., and P. Shor, 1996, Phys. Rev. A **54**, 1098.

Caruso, F., A. W. Chin, A. Datta, S. F. Huelga, and M. B. Plenio, 2009, J. Chem. Phys. **131**, 105106.

Caruso, F., J. Eisert, V. Giovannetti, and A. S. Holevo, 2008, New J. Phys. **10**, 083030.

Caruso, F., and V. Giovannetti, 2006, Phys. Rev. A **74**, 062307.

Caruso, F., and V. Giovannetti, 2007, Phys. Rev. A **76**, 042331.

Caruso, F., V. Giovannetti, and A. S. Holevo, 2006, New J. Phys. **8**, 310.

Caruso, F., V. Giovannetti, C. Macchiavello, and M. B. Ruskai, 2008, Phys. Rev. A **77**, 052323.

Caruso, F., V. Giovannetti, and M. Palma, 2010, Phys. Rev. Lett. **104**, 020503.

Caruso, F., S. F. Huelga, and M. B. Plenio, 2010, Phys. Rev. Lett. **105**, 190501.

Caves, C. M., and P. D. Drummond, 1994, Rev. Mod. Phys. **66**, 481.

Cerf, N., J. Clavareau, C. Macchiavello, and J. Roland, 2005, Phys. Rev. A **72**, 042330.

Cerf, N., J. Clavareau, J. Roland, and C. Macchiavello, 2006, Int. J. Quantum. Inform. **04**, 439.

Chau, H. F., 1998, Phys. Rev. A **58**, 905.

Chau, H. F., 1999, Phys. Rev. A **60**, 1966.

Chiribella, G., M. Dall'Arno, G. M. D'Ariano, C. Macchiavello, and P. Perinotti, 2011, Phys. Rev. A **83**, 052305.

Chitambar, E., D. Leung, L. Mancinska, M. Ozols, and A. Winter, 2012, arXiv:1210.4583.

Choi, M. D., 1975, Linear Algebra Appl. **10**, 285.

Chruściński, D., and A. Kossakowski, 2010, Phys. Rev. Lett. **104**, 070406.

Chruściński, D., and A. Kossakowski, 2012, Europhys. Lett. **97**, 20005.

Chruściński, D., A. Kossakowski, and A. Rivas, 2011, Phys. Rev. A **83**, 052128.

Ciccarello, F., G. M. Palma, and V. Giovannetti, 2013, Phys. Rev. A **87**, 040103(R).

Clemens, J. P., S. Siddiqui, and J. Gea-Banacloche, 2004, Phys. Rev. A **69**, 062313.

Cover, T. M., and J. A. Thomas, 1991, *Elements of Information Theory* (Wiley, New York).

Cubitt, T. S., J. Eisert, and M. M. Wolf, 2012, Phys. Rev. Lett. **108**, 120503.

Cubitt, T. S., M. B. Ruskai, and G. Smith, 2008, J. Math. Phys. (N.Y.) **49**, 102104.

Cuff, P., 2008, arXiv:0805.0065.

Daems, D., 2007, Phys. Rev. A **76**, 012310.

Daffer, S., K. Wódkiewicz, J. D. Cresser, and J. K. McIver, 2004, Phys. Rev. A **70**, 010304(R).

Daffer, S., K. Wódkiewicz, and J. K. McIver, 2003, Phys. Rev. A **67**, 062312.

D'Arrigo, A., G. Benenti, and G. Falci, 2007, New J. Phys. **9**, 310.

D'Arrigo, A., G. Benenti, and G. Falci, 2009, Phys. Scr. **T135**, 014052.

D'Arrigo, A., E. De Leo, G. Benenti, and G. Falci, 2008, Int. J. Quantum. Inform. **06**, 651.

Datta, N., and T. C. Dorlas, 2007, J. Phys. A **40**, 8147.

Datta, N., and T. C. Dorlas, 2009, J. Stat. Phys. **134**, 1173.

Datta, N., A. S. Holevo, and Y. Suhov, 2006, Int. J. Quantum. Inform. **04**, 85.

Datta, N., Y. Suhov, and T. C. Dorlas, 2008, Quantum Inf. Process. **7**, 251.

Demkowicz-Dobrzanski, R., P. Kolenderski, and K. Banaszek, 2007, Phys. Rev. A **76**, 022302.

De Palma, G., A. Mari, and V. Giovannetti, 2014, Phys. Rev. A **90**, 042312.
De Pasquale, A., and V. Giovannetti, 2012, Phys. Rev. A **86**, 052302.
Devetak, I., 2005, IEEE Trans. Inf. Theory **51**, 44.
Devetak, I., A. Harrow, and A. Winter, 2004, Phys. Rev. Lett. **93**, 230504.
Devetak, I., A. Harrow, and A. Winter, 2008, IEEE Trans. Inf. Theory **54**, 4587.
Devetak, I., and P. W. Shor, 2005, Commun. Math. Phys. **256**, 287.
DiVincenzo, D. P., P. W. Shor, and J. A. Smolin, 1998, Phys. Rev. A **57**, 830.
Dobrushin, R. L., 1963, Am. Math. Soc. Trans. **33**, 323.
Duan, L.-M., and G.-C. Guo, 1997, Phys. Rev. Lett. **79**, 1953.
Duan, L.-M., and G.-C. Guo, 1998, Phys. Rev. A **57**, 737.
Duan, R. Y., 2009, arXiv:0906.2527.
Duan, R. Y., and Y. Y. Shi, 2008, Phys. Rev. Lett. **101**, 020501.
Eggeling, T., D.M. Schlingemann, and R.F. Werner, 2002, Europhys. Lett. **57**, 782.
Eisert, J., and M. B. Plenio, 2003, Int. J. Quantum. Inform. **01**, 479.
Ekert, E., and C. Macchiavello, 1996, Phys. Rev. Lett. **77**, 2585.
Fannes, M., B. Haegeman, M. Mosonyi, and D. Vanpeteghem, 2004, arXiv:quant-ph/0410195.
Feinstein, A., 1954, Trans. IRE Professional Group Inf. Theory **4**, 2.
Ferraro, A., S. Olivares, and M. G. A. Paris, 2005, *Gaussian states in continuous variable quantum information* (Bibliopolis, Napoli).
Filippov, S. N., T. Rybár, and M. Ziman, 2012, Phys. Rev. A **85**, 012303.
Fukuda, M., and A. S. Holevo, 2005, arXiv:quant-ph/0510148.
Fukuda, M., C. King, and D. K. Moser, 2010, Commun. Math. Phys. **296**, 111.
Gaitan, F., 2008, *Quantum Error Correction and Fault Tolerant Quantum Computing* (CRC Press, Boca Raton, FL).
Galindo, A., and M. A. Martín-Delgado, 2002, Rev. Mod. Phys. **74**, 347.
Gallager, R. G., 1968, *Information Theory and Reliable Communication* (Wiley, New York).
García-Patrón, R., C. Navarrete-Benlloch, S. Lloyd, J. H. Shapiro, and N. J. Cerf, 2012, Phys. Rev. Lett. **108**, 110505.
Giovannetti, V., 2005, J. Phys. A **38**, 10 989.
Giovannetti, V., D. Burgarth, and S. Mancini, 2009, Phys. Rev. A **79**, 012311.
Giovannetti, V., and R. Fazio, 2005, Phys. Rev. A **71**, 032314.
Giovannetti, V., R. García-Patrón, N. J. Cerf, and A. S. Holevo, 2013, arXiv:1312.6225.
Giovannetti, V., S. Guha, S. Lloyd, L. Maccone, and J. H. Shapiro, 2004a, Phys. Rev. A **70**, 032315.
Giovannetti, V., S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro, B. J. Yen, and H. P. Yuen, 2004b, Quantum Inf. Comput. **4**, 489.
Giovannetti, V., S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro, and H. P. Yuen, 2004c, Phys. Rev. Lett. **92**, 027902.
Giovannetti, V., A. S. Holevo, and R. García-Patrón, 2013, arXiv:1312.2251.
Giovannetti, V., A. S. Holevo, S. Lloyd, and L. Maccone, 2010, J. Phys. A **43**, 415305.
Giovannetti, V., S. Lloyd, and L. Maccone, 2012, Phys. Rev. A **85**, 012302.
Giovannetti, V., S. Lloyd, L. Maccone, and J. H. Shapiro, 2013, Nat. Photonics **7**, 834.
Giovannetti, V., S. Lloyd, L. Maccone, and P. W. Shor, 2003a, Phys. Rev. A **68**, 062323.
Giovannetti, V., S. Lloyd, L. Maccone, and P. W. Shor, 2003b, Phys. Rev. Lett. **91**, 047901.
Giovannetti, V., and S. Mancini, 2005, Phys. Rev. A **71**, 062304.
Giovannetti, V., and M. Palma, 2012, Phys. Rev. Lett. **108**, 040401.
Gisin, N., G. Ribordy, W. Tittel, and H. Zbinden, 2002, Rev. Mod. Phys. **74**, 145.
Gisin, N., and R. Thew, 2007, Nat. Photonics **1**, 165.
Gorini, V., A. Kossakowski, and E. C. G. Sudarshan, 1976, J. Math. Phys. (N.Y.) **17**, 821.
Gottesman, D., 1997, Stabilizer Codes and Qauntum Error Correction, Ph.D. thesis (California Institute of Technology, Pasadena Ca).
Gottesman, D., 1999, Chaos Solitons Fractals, **10**, 1749.
Grace, M. D., J. Dominy, R. L. Kosut, C. Brif, and H. Rabitz, 2010, New J. Phys. **12**, 015001.
Grassl, M., T. Beth, and T. Pellizzari, 1997, Phys. Rev. A **56**, 33.
Grassl, M., and M. Roetteler, 2006, *Proceedings of the IEEE International Symposium on Information Theory* (IEEE Information Theory Society Editor, New York), p. 1109.
Gray, R. M., 1972, IEEE Trans. Inf. Theory **18**, 725.
Guha, S., B. I. Erkmen, and J. H. Shapiro, 2007, arXiv:0710.5666.
Gühne, O., and G. Tóth, 2009, Phys. Rep. **474**, 1.
Gütschow, J., 2010, Quantum information processing with Clifford quantum cellular automata, Ph.D. thesis (Leibniz University, Hannover), http://www.pik-potsdam.de/members/johannes/dissertation-jg.pdf.
Hamada, M., 2002, J. Math. Phys. (N.Y.) **43**, 4382.
Hastings, M. B., 2009, Nat. Phys. **5**, 255.
Hausladen, P., R. Jozsa, B. Schumacher, M. Westmoreland, and W. K. Wootters, 1996, Phys. Rev. A **54**, 1869.
Hausladen, P., B. Schumacher, M. Westmoreland, and W. K. Wootters, 1995, Ann. N.Y. Acad. Sci. **755**, 698.
Hayashi, M., 2007, Phys. Rev. A **76**, 062301.
Hayashi, M., 2008, arXiv:0805.4092.
Hayashi, M., 2009, Commun. Math. Phys. **289**, 1087.
Hayashi, M., and H. Nagaoka, 2003, IEEE Trans. Inf. Theory **49**, 1753.
Hayden, P., M. Horodecki, A. Winter, and J. Yard, 2008, Open Syst. Inf. Dyn. **15**, 7.
Hayden, P., and A. Winter, 2008, Commun. Math. Phys. **284**, 263.
Helstrom, C. W., 1976, *Quantum Detection and Estimation Theory* (Academic Press, London).
Hill, R., 1985, *A First Course in Coding Theory* (Clarendon Press, Oxford).
Hiroshima, T., 2006, Phys. Rev. A **73**, 012330.
Holevo, A. S., 1972, Probl. Inf. Transm. **8**, 63.
Holevo, A. S., 1973a, Probl. Inf. Transm. **9**, 177.
Holevo, A. S., 1973b, Probl. Inf. Transm. **9**, 110.
Holevo, A. S., 1997, arXiv:quant-ph/9705054.
Holevo, A. S., 1998a, IEEE Trans. Inf. Theory **44**, 269.
Holevo, A. S., 1998b, Russ. Math. Surv. **53**, 1295.
Holevo, A. S., 2004, Theory Probab. Appl. **48**, 243.
Holevo, A. S., 2007a, Theory Probab. Appl. **51**, 92.
Holevo, A. S., 2007b, Probl. Inf. Transm. **43**, 1.
Holevo, A. S., 2007c, *Proceedings of the International Congress of Mathematicians* (European Mathematical Society, Zurich), p. 999.
Holevo, A. S., 2008, Probl. Inf. Transm. **44**, 171.
Holevo, A. S., 2011, *Probabilistic and Statistical Aspects of Quantum Theory* (Edizioni della Normale, Pisa).
Holevo, A. S., 2012, *Quantum Systems, Channels, Information* (De Gruyter, Berlin).
Holevo, A. S., and V. Giovannetti, 2012, Rep. Prog. Phys. **75**, 046001
Holevo, A. S., and M. E. Shirokov, 2006, Theory Probab. Appl. **50**, 86.

Holevo, A. S., M. Sohma, and O. Hirota, 1999, Phys. Rev. A **59**, 1820.

Holevo, A. S., and R. F. Werner, 2001, Phys. Rev. A **63**, 032312.

Horn, R. A., and C. R. Johnson, 1990, *Matrix Analysis* (Cambridge University Press, Cambridge, England).

Horodecki, K., M. Horodecki, P. Horodecki, and J. Oppenheim, 2005, Phys. Rev. Lett. **94**, 160502.

Horodecki, K., L. Pankowski, M. Horodecki, and P. Horodecki, 2008, IEEE Trans. Inf. Theory **54**, 2621.

Horodecki, M., P. Horodecki, and R. Horodecki, 1996, Phys. Lett. A **223**, 1.

Horodecki, M., P. Horodecki, and R. Horodecki, 1999, Phys. Rev. A **60**, 1888.

Horodecki, M., P. Shor, and M. B. Ruskai, 2003, Rev. Math. Phys. **15**, 629.

Horodecki, R., P. Horodecki, M. Horodecki, and K. Horodecki, 2009, Rev. Mod. Phys. **81**, 865.

Houshmand, M., S. Hosseini-Khayat, and M. M. Wilde, 2013, IEEE Trans. Inf. Theory **59**, 1198.

Hsieh, M.-H., I. Devetak, and A. Winter, 2008, IEEE Trans. Inf. Theory **54**, 3078.

Hsieh, M.-H., and M. M. Wilde, 2010a, IEEE Trans. Inf. Theory **56**, 4682.

Hsieh, M.-H., and M. M. Wilde, 2010b, IEEE Trans. Inf. Theory **56**, 4705.

Hu, Y., Z.-W. Zhou, and G.-C. Guo, 2007, New J. Phys. **9**, 27.

Hughes, R. J., J. E. Nordholt, D. Derkacs, and C. G. Petereson, 2002, New J. Phys. **4**, 43.

Hwang, W. Y., D. Ahn, and S. W. Hwang, 2001, Phys. Rev. A **63**, 022303.

Jacobs, K., 1962, *Colloquium on Combinatorial Methods in Probabilistic Theory* (Aarhus, Denmark).

Jamiolkowski, A., 1972, Rep. Math. Phys. **3**, 275.

Johnston, N., D. W. Kribs, and V. I. Paulsen, 2009, Quantum Inf. Comput. **9**, 0016.

Jordan, T. F., A. Shaji, and E. C. G. Sudarshan, 2004, Phys. Rev. A **70**, 052110.

Jozsa, R., 1994, J. Mod. Opt. **41**, 2315.

Jozsa, R., M. Horodecki, P. Horodecki, and R. Horodecki, 1998, Phys. Rev. Lett. **81**, 1714.

Karimipour, V., and L. Memarzadeh, 2006a, Phys. Rev. A **74**, 032332.

Karimipour, V., and L. Memarzadeh, 2006b, Phys. Rev. A **74**, 062311.

Karpov, E., D. Daems, and N. J. Cerf, 2006a, Phys. Rev. A **74**, 032320.

Karpov, E., D. Daems, and N. J. Cerf, 2006b, Open Syst. Inf. Dyn. **13**, 363.

Kawataba, S., 2000, J. Phys. Soc. Jpn. **69**, 3540.

Ketkar, A., A. Klappenecker, S. Kumar, and P. K. Sarvepalli, 2006, IEEE Trans. Inf. Theory **52**, 4892.

Keyl, M., 2002, Phys. Rep. **369**, 431.

Khinchin, A. I., 1957, *Mathematical Foundations of Information Theory* (Dover Publications, New York).

King, C., 2002, J. Math. Phys. (N.Y.) **43**, 4641.

King, C., 2003, IEEE Trans. Inf. Theory **49**, 221.

King, C., and M. B. Ruskai, 2001, IEEE Trans. Inf. Theory **47**, 192.

Kitaev, A. Yu., 1997, Russ. Math. Surv. **52**, 1191.

Kitaev, A. Yu., 2003, Ann. Phys. (N.Y.) **303**, 2.

Klesse, R., and S. Frank, 2005, Phys. Rev. Lett. **95**, 230503.

Knill, E., 1996, arXiv:quant-ph/9608048.

Knill, E., and R. Laflamme, 1997, Phys. Rev. A **55**, 900.

Knill, E., R. Laflamme, A. Ashikhmin, H. N. Barnum, L. Viola, and W. H. Zurek, 2002, *Introduction to Quantum Error Correction* (Los Alamos Science, Los Alamos, NM), Vol. 27.

König, R., and G. Smith, 2013a, Nat. Photonics **7**, 142.

König, R., and G. Smith, 2013b, Phys. Rev. Lett. **110**, 040501.

König, R., and S. Wehner, 2009, Phys. Rev. Lett. **103**, 070504.

Körner, J., and A. Orlitsky, 1998, IEEE Trans. Inf. Theory **44**, 2207.

Kraus, K., 1971, Ann. Phys. (N.Y.) **64**, 311.

Kraus, K., 1983, *States, Effects and Operations*, Vol. 190, Lecture Notes in Physics (Springer-Verlag, Berlin).

Kretschmann, D., 2003, Channel Capacity Quantized, Ph.D. thesis (Institut für Mathematische Physik Technische Universität Braunschweig).

Kretschmann, D., and R. F. Werner, 2004, New J. Phys. **6**, 26.

Kretschmann, D., and R. F. Werner, 2005, Phys. Rev. A **72**, 062323.

Laflamme, R., C. Miquel, J. P. Paz, and W. H. Zurek, 1996, Phys. Rev. Lett. **77**, 198.

Landauer, R., 1961, IBM J. Res. Dev. **5**, 183.

Leibfried, D., R. Blatt, C. Monroe, and D. Wineland, 2003, Rev. Mod. Phys. **75**, 281.

Lidar, D. A., D. Bacon, and K. B. Whaley, 1999, Phys. Rev. Lett. **82**, 4556.

Lidar, D. A., and T. A. Brun, 2013 *Quantum Error Correction* (Cambridge University Press, Cambridge, England).

Lidar, D. A., I. L. Chuang, and K. B. Whaley, 1998, Phys. Rev. Lett. **81**, 2594.

Lidar, D. A., and K. B. Whaley, 2003 "Decoherence-Free Subspaces and Subsystems", *Irreversible Quantum Dynamics*, edited by F. Benatti, and R. Floreanini, Springer Lecture Notes in Physics, Vol. 622 (Springer-Verlag, Berlin), pp. 83–120.

Lindblad, G., 1975, Commun. Math. Phys. **40**, 147.

Lindblad, G., 1976, Commun. Math. Phys. **48**, 119.

Lloyd, S., 1997, Phys. Rev. A **55**, 1613.

Lloyd, S., V. Giovannetti, and L. Maccone, 2011, Phys. Rev. Lett. **106**, 250501.

Lloyd, S., V. Giovannetti, L. Maccone, N. J. Cerf, R. García-Patrón, S. Mitter, S. Pirandola, M. B. Ruskai, J. H. Shapiro, and H. Yuan, 2009, arXiv:0906.2758.

Lu, X.-M., X. Wang, and C. P. Sun, 2010, Phys. Rev. A **82**, 042103.

Luo, S., S. Fu, and H. Song, 2012, Phys. Rev. A **86**, 044101.

Lupo, C., V. Giovannetti, and S. Mancini, 2010a, Phys. Rev. Lett. **104**, 030501.

Lupo, C., V. Giovannetti, and S. Mancini, 2010b, Phys. Rev. A **82**, 032312.

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

Lupo, C., V. Giovannetti, S. Pirandola, S. Mancini, and S. Lloyd, 2012, Phys. Rev. A **85**, 062314.

Lupo, C., and S. Mancini, 2010, Phys. Rev. A **81**, 052314.

Lupo, C., L. Memarzadeh, and S. Mancini, 2009, Phys. Rev. A **80**, 042328.

Lupo, C., O. Pilyavets, and S. Mancini, 2009, New J. Phys. **11**, 063023.

Lupo, C., S. Pirandola, P. Aniello, and S. Mancini, 2011, Phys. Scr. **T143**, 014016.

Ma, X.-S., *et al.*, 2012, Nature (London) **489**, 269.

Macchiavello, C., and G. M. Palma, 2002, Phys. Rev. A **65**, 050301(R).

Macchiavello, C., G. M. Palma, and S. Virmani, 2004, Phys. Rev. A **69**, 010303(R).

Mac Kay, D. J. C., 2003, *Information Theory, Inference, and Learning Algorithms* (Cambridge University Press, Cambridge, England).

Madsen, K. H., S. Ates, T. Lund-Hansen, A. Löffler, S. Reitzenstein, A. Forchel, and P. Lodahl, 2011, Phys. Rev. Lett. **106**, 233601.

McMillan, B., 1953, Ann. Math. Stat. **24**, 196.

Medeiros, R. A. C., and F. M. de Assis, 2005, Int. J. Quantum. Inform. **03**, 135.

Memarzadeh, L., and S. Mancini, 2010, Eur. Phys. J. D **56**, 291.

Mermin, N. D., 2003, Am. J. Phys. **71**, 23.

Misra, B., and E. C. G. Sudarshan, 1977, J. Math. Phys. (N.Y.) **18**, 756.

Moravčíková, L., and M. Ziman, 2010, J. Phys. A **43**, 275306.

Nakajima, S., 1958, Prog. Theor. Phys. **20**, 948.

Nielsen, M. A., 2002, Phys. Lett. A **303**, 249.

Nielsen, M. A., and I. L. Chuang, 2000, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England).

Norris, J. R., 1997, *Markov Chains* (Cambridge University Press, Cambridge, England).

Ogawa, T., and H. Nagaoka, 1999, IEEE Trans. Inf. Theory **45**, 2486.

Ogawa, T., and H. Nagaoka, 2002, *Proceedings of the 2002 IEEE International Symposium Information Theory, Lausanne, Switzerlan* (IEEE, New York), p. 73.

Ohya, M., and D. Petz, 1993, *Quantum Entropy and Its Use* (Springer-Verlag, New York).

Ollivier, H., and J.-P. Tillich, 2003, Phys. Rev. Lett. **91**, 177902.

Ollivier, H., and J.-P. Tillich, 2004, arXiv:quant-ph/0401134.

Palma, M. G., K.-A. Suominen, and A. K. Ekert, 1996, Proc. R. Soc. A **452**, 567.

Paulsen, V. I., 2003, *Completely Bounded Maps and Operator Algebras* (Cambridge University Press, Cambridge, England).

Peres, A., 1996, Phys. Rev. Lett. **77**, 1413.

Peterson, W. W., and E. J. Weldon, 1972, *Error Correcting Codes* (MIT Press, Cambridge, MA).

Petz, D., 2008, *Quantum Information Theory and Quantum Statistics* (Springer, Berlin).

Piani, M., M. Horodecki, P. Horodecki, and R. Horodecki, 2006, Phys. Rev. A **74**, 012305.

Pilyavets, O., C. Lupo, and S. Mancini, 2012, IEEE Trans. Inf. Theory **58**, 6126.

Pilyavets, O., V. Zborovskii, and S. Mancini, 2008, Phys. Rev. A **77**, 052324.

Pirandola, S., S. Mancini, S. L. Braunstein, and S. Lloyd, 2008, Nat. Phys. **4**, 726.

Plenio, M. B., and S. F. Huelga, 2008, New J. Phys. **10**, 113019.

Plenio, M. B., and S. Virmani, 2007, Phys. Rev. Lett. **99**, 120504.

Plenio, M. B., and S. Virmani, 2008, New J. Phys. **10**, 043032.

Poulin, D., J.-P. Tillich, and H. Ollivier, 2009, IEEE Trans. Inf. Theory **55**, 2776.

Preskill, J., 1998, Lecture Notes for Physics 229: Quantum Information and Computation, http://www.theory.caltech.edu/people/preskill/ph229/#lecture.

Raginsky, M., 2002, Phys. Rev. A **65**, 032306.

Rains, E. M., 1999a, Phys. Rev. A **60**, 173.

Rains, E. M., 1999b, Phys. Rev. A **60**, 179.

Rains, E. M., 2001, IEEE Trans. Inf. Theory **47**, 2921.

Rau, J., 1963, Phys. Rev. **129**, 1880.

Rebentrost, P., R. Chakraborty, and A. Aspuru-Guzik, 2009, J. Chem. Phys. **131**, 184102.

Richter, S., and R. F. Werner, 1996, J. Stat. Phys. **82**, 963.

Rivas, A., S. F. Huelga, and M. B. Plenio, 2010, Phys. Rev. Lett. **105**, 050403.

Rossini, D., V. Giovannetti, and S. Montangero, 2008, New J. Phys. **10**, 115009.

Ruggeri, G., and S. Mancini, 2007a, Quantum Inf. Comput. **7**, 265.

Ruggeri, G., and S. Mancini, 2007b, Phys. Lett. A **362**, 340.

Ruggeri, G., G. Soliani, V. Giovannetti, and S. Mancini, 2005, Europhys. Lett. **70**, 719.

Ruskai, M. B., 2003, Rev. Math. Phys. **15**, 643.

Ruskai, M. B., S. Szarek, and E. Werner, 2002, Linear Algebra Appl. **347**, 159.

Rybar, T., S. N. Filippov, M. Ziman, and V. Bužek, 2012, J. Phys. B **45**, 154006.

Scarani, V., H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, 2009, Rev. Mod. Phys. **81**, 1301.

Scarani, V., M. Ziman, P. Stelmachovic, N. Gisin, and V. Bužek, 2002, Phys. Rev. Lett. **88**, 097905.

Schäfer, J., E. Karpov, and N. J. Cerf, 2009, Phys. Rev. A **80**, 062313.

Schäfer, J., E. Karpov, and N. J. Cerf, 2010, Proc. SPIE Int. Soc. Opt. Eng. **7727**, 77270J.

Schäfer, J., E. Karpov, and N. J. Cerf, 2011, Phys. Rev. A **84**, 032318.

Schäfer, J., E. Karpov, and N. J. Cerf, 2012, Phys. Rev. A **85**, 012322.

Schmitt-Manderbach, T., *et al.*, 2007, Phys. Rev. Lett. **98**, 010504.

Schuch, N., J. I. Cirac, and M. M. Wolf, 2008, *Quantum Information and Many Body Quantum Systems*, edited by M. Ericsson and S. Montangero (Edizioni della Normale, Pisa), pp. 129–142.

Schumacher, B., 1995, Phys. Rev. A **51**, 2738.

Schumacher, B., 1996, Phys. Rev. A **54**, 2614.

Schumacher, B., and M. A. Nielsen, 1996, Phys. Rev. A **54**, 2629.

Schumacher, B., and W. D. Westmoreland, 1997, Phys. Rev. A **56**, 131.

Schumacher, B., and W. D. Westmoreland, 2000, arXiv:quant-ph/0004045.

Schumacher, B., and W. D. Westmoreland, 2010, *Quantum Processes, Systems, and Information* (Cambridge University Press, Cambridge, England).

Semenov, A. A., and W. Vogel, 2009, Phys. Rev. A **80**, 021802.

Semenov, A. A., and W. Vogel, 2010, Phys. Rev. A **81**, 023835.

Sen, P., 2011, arXiv:1109.0802.

Serafini, A., J. Eisert, and M. M. Wolf, 2005, Phys. Rev. A **71**, 012320.

Shabani, A., and D. A. Lidar, 2005, Phys. Rev. A **72**, 042303.

Shadman, Z., K. Kampermann, D. Bruss, and C. Macchiavello, 2011, Phys. Rev. A **84**, 042309.

Shaji, A., and E. C. G. Sudarshan, 2005, Phys. Lett. A **341**, 48.

Shannon, C. E., 1948, Bell Syst. Tech. J. **27**, 379.

Shannon, C. E., 1949, Proc. IRE **37**, 10.

Shannon, C. E., 1956, IRE Trans. Inf. Theory **2**, 8.

Shapiro, J. H., 2009, IEEE J. Sel. Top. Quantum Electron. **15**, 1547.

Shor, P. W., 1995, Phys. Rev. A **52**, R2493.

Shor, P. W., 2002a, *The quantum channel capacity and coherent information*, Lecture notes, MSRI Workshop on Quantum Computation (Mathematical Science Research Institute, Berkeley, CA).

Shor, P. W., 2002b, J. Math. Phys. (N.Y.) **43**, 4334.

Shor, P. W., 2004, Commun. Math. Phys. **246**, 453.

Shor, P. W., and J. A. Smolin, 1996, arXiv:quant-ph/9604006.

Simon, R., N. Mukunda, and B. Dutta, 1994, Phys. Rev. A **49**, 1567.

Smith, G., 2010, arXiv:1007.2855.

Smith, G., and J. A. Smolin, 2007, Phys. Rev. Lett. **98**, 030501.

Smith, G., J. A. Smolin, and A. Winter, 2008, IEEE Trans. Inf. Theory **54**, 4208.

Smith, G., and J. Yard, 2008, Science **321**, 1812.

Steane, A. M., 1996, Proc. R. Soc. A **452**, 2551.

Steane, A. M., 1999, IEEE Trans. Inf. Theory **45**, 1701.
Stinespring, W. F., 1955, Proc. Am. Math. Soc. **6**, 211.
Sudarshan, E. C. G., P. M. Mathews, and J. Rau, 1961, Phys. Rev. **121**, 920.
Terhal, B. M., and D. P. DiVincenzo, 2000, Phys. Rev. A **61**, 022301.
Thorwart, M., J. Eckel, J. H. Reina, P. Nalbach, and S. Weiss, 2009, Chem. Phys. Lett. **478**, 234.
Tyler, G. A., and R. W. Boyd, 2009, Opt. Lett. **34**, 142.
Uhlmann, A., 1976, Rep. Math. Phys. **9**, 273.
Ursin, R., T. Jennewein, M. Aspelmeyer, R. Kaltenbaek, M. Lindenthal, P. Walther, and A. Zeilinger, 2004, Nature (London) **430**, 849.
Ursin, R., *et al.*, 2007, Nat. Phys. **3**, 481.
Vacchini, B., A. Smirne, E. M. Laine, J. Piilo, and H.-P Breuer, 2011, New J. Phys. **13**, 093004.
Vaidman, L., 1994, Phys. Rev. A **49**, 1473.
Vasile, R., S. Maniscalco, M. G. A. Paris, H.-P. Breuer, and J. Piilo, 2011, Phys. Rev. A **84**, 052118.
Vasile, R., S. Olivares, M. G. A. Paris, and S. Maniscalco, 2011, Phys. Rev. A **83**, 042321.
Vatan, F., V. P. Roychowdhury, and M. P. Anantram, 1999, IEEE Trans. Inf. Theory **45**, 1703.
Vats, N., S. John, and K. Busch, 2002, Phys. Rev. A **65**, 043808.
Verdú, S., 1998, IEEE Trans. Inf. Theory **44**, 2057.
Verdú, S., and T. S. Han, 1994, IEEE Trans. Inf. Theory **40**, 1147.
Villoresi, P., *et al.*, 2008, New J. Phys. **10**, 033038.
Wang, L., and R. Renner, 2012, Phys. Rev. Lett. **108**, 200501.
Weedbrook, C., S. Pirandola, R. Garcia-Patron, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, 2012, Rev. Mod. Phys. **84**, 621.
Wehrl, A., 1978, Rev. Mod. Phys. **50**, 221.
Wilde, M. M., 2013, *Quantum Information Theory* (Cambridge University Press, Cambridge, England).
Wilde, M. M., and A. Winter, 2013, arXiv:1308.6732.
Wilde, M. M., A. Winter, and D. Yang, 2013, arXiv:1306.1586.
Winter, A., 1999a, Ph.D. thesis (Universität Bielefled).
Winter, A., 1999b, IEEE Trans. Inf. Theory **45**, 2481.
Wolf, M. M., and J. I. Cirac, 2008, Commun. Math. Phys. **279**, 147.
Wolf, M. M., J. Eisert, T. S. Cubitt, and J. I. Cirac, 2008, Phys. Rev. Lett. **101**, 150402.
Wolf, M. M., G. Ortiz, F. Verstraete, and J. I. Cirac, 2006, Phys. Rev. Lett. **97**, 110403.
Wolf, M. M., and D. Pérez-García, 2007, Phys. Rev. A **75**, 012303.
Wolf, M. M., D. Pérez-García, and G. Giedke, 2007, Phys. Rev. Lett. **98**, 130501.
Wolfowitz, J., 1957, Ill. J. Math. **1**, 591.
Wolfowitz, J., 1959/1960, Arch. Ration. Mech. Anal. **4**, 371.
Wolfowitz, J., 1964, *Coding Theorem and Information Theory* (Springer, New York).
Yin, J., *et al.*, 2013, arXiv:1306.0672.
Zanardi, P., and M. Rasetti, 1997, Phys. Rev. Lett. **79**, 3306.
Ziman, M., and V. Bužek, 2005, Phys. Rev. A **72**, 022110.
Ziman, M., P. Štelmachovič, and V. Bužek, 2005, Open Syst. Inf. Dyn. **12**, 81.
Ziman, M., P. Štelmachovič, V. Bužek, V. Scarani, and N. Gisin, 2002, Phys. Rev. A **65**, 042105.
Zwanzig, R., 1960, J. Chem. Phys. **33**, 1338.