Quantum channels and their entropic characteristics

A S Holevo\textsuperscript{1} and V Giovannetti\textsuperscript{2}

\textsuperscript{1} Steklov Mathematical Institute, Gubkina 8, 119991 Moscow, Russia
\textsuperscript{2} NEST, Scuola Normale Superiore and Istituto Nanoscienze-CNR, piazza dei Cavalieri 7, I-56126 Pisa, Italy

E-mail: holevo@mi.ras.ru and v.giovannetti@sns.it

Received 25 August 2011, in final form 15 November 2011
Published 7 March 2012
Online at stacks.iop.org/RoPP/75/046001

Abstract

One of the major achievements of the recently emerged quantum information theory is the introduction and thorough investigation of the notion of a quantum channel which is a basic building block of any data-transmitting or data-processing system. This development resulted in an elaborated structural theory and was accompanied by the discovery of a whole spectrum of entropic quantities, notably the channel capacities, characterizing information-processing performance of the channels. This paper gives a survey of the main properties of quantum channels and of their entropic characterization, with a variety of examples for finite-dimensional quantum systems. We also touch upon the ‘continuous-variables’ case, which provides an arena for quantum Gaussian systems. Most of the practical realizations of quantum information processing were implemented in such systems, in particular based on principles of quantum optics. Several important entropic quantities are introduced and used to describe the basic channel capacity formulae. The remarkable role of specific quantum correlations—entanglement—as a novel communication resource is stressed.

(Some figures may appear in colour only in the online journal)

This article was invited by M Lewenstein.

Contents

1. Introduction

2. Channels and open systems
   2.1. Classical and quantum information carriers
   2.2. Classical and quantum channels
   2.3. Dynamics of isolated and open quantum systems
   2.4. Heisenberg picture
   2.5. Positivity, complete positivity and the Choi–Jamiolkowski representation
   2.6. Compositions rules
   2.7. Complementary channels
   2.8. Quantum channels involving a classical stage

3. Bosonic Gaussian channels
   3.1. Example: channel with additive Gaussian quantum noise
   3.2. Multimode bosonic systems
   3.3. Quantum Gaussian channels
   3.4. General properties of the Gaussian channels
   3.5. The case of one mode

4. Entropy, information and channel capacities
   4.1. Information transmission over a classical channel
   4.2. Quantifying information in a quantum world
   4.3. The classical capacity of a quantum channel: part I
   4.4. The classical capacity of a quantum channel: part II
   4.5. Entropy exchange and quantum mutual information
   4.6. Entanglement as an information resource
   4.7. Coherent information and perfect error correction
   4.8. The quantum capacity
   4.9. Quantum wiretap channel
   4.10. Capacities for Gaussian channels
   4.11. The variety of quantum channel capacities
   4.12. Relative entropy

5. Summary and outlook

Acknowledgments

References
1. Introduction

The concept of a communication channel and its capacity is central in information theory [37] where probabilistic models of noise have been used with great success. The importance of considering quantum channels is due to the fact that any physical communication line is after all a quantum one, and it can be treated classically only in so far as the quantum noise associated with the fundamental quantum-mechanical uncertainty [42] is negligible as compared with the classical fluctuations. This is not the case in many modern applications such as optical communication [33] or quantum system engineering [135], calling for a genuinely quantum approach.

Mathematically, the notion of a quantum channel is related to that of a dynamical map, which emerged in the 1960–70s and goes back to the ‘operations’ of Haag–Kastler [70] and Ludwig [128]. The essential property of dynamical maps is positivity, ensuring that states are transformed into states in the Schrödinger picture. These transformations are a quantum analog of the Markov maps (stochastic matrices) in probability theory and a natural further development was to consider semigroups of such processes as a model of quantum Markovian dynamics (see [41, 116]).

However, an approach based on positivity remained not extremely productive until it was observed independently by several researchers (e.g. by Kraus [117], Lindblad [122], Gorini et al [66], Evans and Lewis [50] in statistical mechanics, and by Holevo [82] in the context of quantum communication theory) that the reduced dynamics of open quantum systems has a stronger property of complete positivity, introduced earlier in the purely mathematical context by Stinespring [174] and studied in detail in the finite-dimensional case by Choi [36]. Moreover, when interpreted in physical terms, the basic Stinespring’s dilation theorem implies that complete positivity is not only necessary, but also sufficient for a dynamical map to be extendable to the unitary dynamics of an open quantum system interacting with an environment. In this way the notion of a dynamical map in statistical mechanics (and that of a quantum channel in information theory) was finally identified with that of a completely positive, properly normalized map acting on the relevant operator space (generated either by states or by observables) of the underlying quantum system.

A comprehensive study of the dynamical aspects of open quantum systems was presented by Spohn [158], and a number of important entropic quantities introduced at this early stage were surveyed by Wehrl [184].

Quantum information theory brought a new turn to the whole subject by giving a deeper operational insight into the notion of channels and to their entropic characteristics. Fundamental results of classical information theory are the coding theorems which establish the possibility of transmitting and processing data reliably (i.e. asymptotically error-free) at rates not exceeding certain threshold values (capacities) that characterize the system under consideration. Coding theorems provide explicit formulae for such thresholds in terms of entropic functionals of the channel. From a different perspective, quantum information theory gives a peculiar view onto irreversible evolutions of quantum systems, providing the quantitative answer to the question, to what extent can the effects of irreversibility and noise in the channel be reversed by using intelligent pre-processing of the incoming and post-processing of the outgoing states? The issue of the information capacity of quantum communication channels arose soon after the publication of the pioneering paper by Shannon [160] and goes back to works of Gabor [53] and Gordon [65], asking for fundamental physical limits on the rate and on the quality of information transmission. This laid a physical foundation and raised the question of consistent quantum-theoretical treatment of the problem.

Important steps in this direction were made in the 1970s when quantum detection and estimation theory was developed, making a quantum probabilistic frame for these problems, see the books of Helstrom [80] and Holevo [87]. At that time the concepts of a quantum communication channel and of its capacity for transmitting classical information were established, along with a fundamental upper bound for that quantity. From this point the subject of the present survey begins. In the 1990s a new interest in noisy quantum channels arose in connection with the emerging quantum information science, with a more detailed and deeper insight—e.g. see the books [76, 135] or the review papers [3, 16, 107, 172] and references therein (further references to original works will also be given in this review). The more recent developments are characterized by an emphasis on the new possibilities (rather than on the restrictions) implied by the specifically quantum features of the information-processing agent, notably entanglement as a novel communication resource. Along these lines the achievability of the information bound and a number of further quantum coding theorems were discovered making what can be described as quantum Shannon theory. It was realized that a quantum communication channel is characterized by a whole spectrum of capacities depending on the nature of the information resources and specific protocols used for the transmission.

On the other hand, the question of information capacity turned out important for the theory of quantum computing, particularly in connection with quantum error-correcting codes, communication and algorithmic complexities, and quantum cryptography where the channel environment not only introduces noise but also models an eavesdropper interfering with private communication.

The aim of this paper is to provide a survey of the principal features of quantum communication channels. We will discuss mathematical representations for quantum channels and focus on the question of how their efficiency in transferring signals can be characterized in terms of various entropic quantities. The paper is organized as follows. Section 2 is devoted to the structural theory of quantum channels. It also gives a survey of the main particular classes of channels in the case of a finite level quantum system with a variety of examples. In section 3 we pass to the ‘continuous-variables’ case, which provides an arena for quantum Gaussian systems. Many experimental demonstrations of quantum information processing were realized in such systems, in particular based on principles of quantum optics. Several important entropic quantities are introduced in section 4 and used to describe the

Rep. Prog. Phys. 75 (2012) 046001 A S Holevo and V Giovannetti

A S Holevo and V Giovannetti
2. Channels and open systems

2.1. Classical and quantum information carriers

In any information-processing scheme, data are encoded into the states of some physical systems (e.g. the etched surface of a DVD or the EM microwaves emitted by a cellphone) which play the role of information carriers or fundamental data-processing elements. In its simplest version, the classical theory of information assigns to each of them a finite phase space $X$ (also called alphabet) whose elements represent the possible configurations the carrier can assume (e.g. on/off, dot/dash/space in Morse code, etc). More precisely the completely determined states (pure states) of a carrier are identified with the points $x \in X$, while the mixed states, which represent statistical ensemble of pure states, are described by probability distributions $\{p_x\}$ on $X$. The Shannon entropy of such a distribution $[37]$,

$H(X) = H(\{p_x\}) \equiv -\sum_x p_x \log_2 p_x, \quad (1)$

provides a measure of the ‘uncertainty’ or of the ‘lack of information’ in the corresponding ensemble (also, as will be clarified in section 4.1, $H(X)$ describes the potential information content of a random source producing symbols $x$ with probabilities $p_x$). Choosing the binary logarithm in equation (1) means that we are measuring information in binary digits—bits—which is convenient because of the basic role of two-state processing systems. The minimal value of entropy, equal to 0, is attained on pure states (here and in the rest of the paper we adopt the usual convention $0 \log_2 0 = 0$), while the maximal, equal to $\log_2 |X|$, is obtained on the uniform distribution $p_x \equiv |X|^{-1}$, where $|X|$ denotes the size of the alphabet $X$.

Such a simple phase space description does not hold when the information carrier used to encode the data is a quantum system (e.g. a two-level ion confined in space by using strong electromagnetic fields [134] or a single polarized photon propagating along an optical fiber [63]). Instead the latter can be represented in terms of a Hilbert space $\mathcal{H}$ [42, 181], which again for simplicity we take as finite dimensional. In this context, the pure quantum states are described by projections $|\psi\rangle\langle\psi|$ onto unit vectors $|\psi\rangle$ of $\mathcal{H}$ (as in the classical case they define the completely determined configurations of the carrier, which in the quantum case correspond to specific state preparation procedures). Mixed states are represented by statistical ensembles of pure states $|\psi^n\rangle\langle\psi^n|$ with probabilities $p_\alpha$. Each mixed state is formally described by the corresponding density operator $\rho \equiv \sum_\alpha p_\alpha |\psi^n\rangle\langle\psi^n|$ which contains, in a highly condensed form, the information about the procedure for producing such a state (specifically it describes a preparation procedure of the carrier characterized by a stochastically fluctuating parameter $\alpha$), and which has the following properties [181]:

(i) $\rho$ is a linear Hermitian positive operator in $\mathcal{H}$;
(ii) $\rho$ has unit trace, i.e. $Tr \rho = 1$.

Analogously to their classical counterparts, quantum mixed states form a convex set $\mathcal{S}(\mathcal{H})$ whose extremal points are represented by pure states. Unlike classical mixed states, however, they can be written as a mixture of pure states in many different ways, i.e. one and the same density matrix $\rho$ represents different stochastic preparation procedures of the carrier. A distinguished ensemble is given by the spectral decomposition $\rho = \sum_j \lambda_j |e^j\rangle\langle e^j|$, where $\lambda_j$ are the eigenvalues and $|e^j\rangle$ the orthonormal eigenvectors of the operator $\rho$. The eigenvalues $\lambda_j$ of a density operator form a probability distribution and the Shannon entropy of this distribution is equal to the von Neumann entropy of the density operator [181],

$S(\rho) \equiv -Tr \rho \log_2 \rho = -\sum_j \lambda_j \log_2 \lambda_j = H(\{\lambda_j\}), \quad (2)$

which provides a measure of uncertainty and, as is explained later, of the information content of the quantum state $\rho$. Again, the minimal value of the entropy, equal to 0, is attained on pure states while the maximal, equal to $\log_2 d$,—on the chaotic state $\rho = I/d$ (with $I$ denoting the unit operator in $\mathcal{H}$ and $d = \dim \mathcal{H}$ being the dimensionality of the space).

There is a way to formally embed a finite classical system associated with a classical information carrier into a quantum system by introducing the Hilbert space with the orthonormal basis $\{|x\rangle\}$ indexed by phase space points $x \in X$. Then, to the classical states $\{p_x\}$ of the classical carrier correspond the diagonal density operators $\rho = \sum_x p_x |x\rangle\langle x|$ which commute with each other. Truly quantum systems, however, also admit configurations which are described by density operators that are not simultaneously diagonalizable (i.e. noncommuting). Indeed the full information content of a quantum state cannot be reduced to a classical message and therefore deserves the special name quantum information. This is related to the fact that a quantum state implicitly contains the statistics of all possible quantum measurements, including mutually exclusive (complementary) ones: a distinctive feature which ultimately leads to the impossibility of cloning of quantum information (see section 2.3 for details).

Example. Take a binary alphabet $X = \{0, 1\}$ with associated quantum orthogonal basis $\{|0\rangle, |1\rangle\}$. The density operator $|+\rangle\langle+|$ corresponding to the vector $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ is a proper quantum state of the system which, in general, does not commute with the density matrices $p_0|0\rangle\langle 0| + p_1|1\rangle\langle 1|$ used to embed the statistical distributions $\{p_0, p_1\}$ on $X$.

In information theory, both classical and quantum, messages are typically transmitted using block coding strategies that exploit long sequences of carriers. Therefore, one has to systematically do with composite systems corresponding to repeated or parallel uses of communication channels. Entanglement reflects unusual properties of composite quantum systems which are described by a tensor...
rather than Cartesian (as in the classical case) product of the component systems. According to the superposition principle [42], the Hilbert space $\mathcal{H}_{AB} \equiv \mathcal{H}_A \otimes \mathcal{H}_B$ of the composite system $AB$, along with product vectors $|\psi_A \otimes \psi_B\rangle \equiv |\psi_A\rangle \otimes |\psi_B\rangle$ with $|\psi_A\rangle \in \mathcal{H}_A$ and $|\psi_B\rangle \in \mathcal{H}_B$, contains all possible linear combinations $\sum_j |\psi_A^j \otimes \psi_B^j\rangle$. The pure states given by product vectors are called separable while their superpositions are entangled. A mixed state of $AB$ is called separable if it can be expressed as a mixture of product states, and entangled if it cannot [135, 185].

Entanglement is an intrinsically ‘nonclassical’ sort of correlation between subsystems which typically emerges due to quantum interactions. If a classical composite system $AB$ is in a pure (i.e. completely determined) state then apparently the subsystems $A$, $B$ are also in their uniquely defined pure partial states. Strikingly, this is not so for the entangled states of a quantum composite system $AB$. Consider a generic pure state $|\psi_{AB}\rangle$ of such a system. By Schmidt decomposition it can be expressed as

$$|\psi_{AB}\rangle = \sum_j \lambda_j |e_A^j \otimes e_B^j\rangle,$$

(3)

where $\{\lambda_j\}$ is a probability distribution which is uniquely determined by $|\psi_{AB}\rangle$, while $\{|e_A^j, e_B^j\rangle\}$ are some orthonormal bases in $\mathcal{H}_{A,B}$ [135]. Then the local state of the subsystem $A$ is given by the reduced density operator

$$\rho_A = \text{Tr}_B |\psi_{AB}\rangle \langle \psi_{AB}| = \sum_j \lambda_j |e_A^j\rangle \langle e_A^j|,$$

(4)

obtained by taking the partial trace $\text{Tr}_B[\cdots]$ of the original joint state (3) over the degrees of freedom of $B$ (similarly for the partial state of $B$). Thus, unless the Schmidt form (3) factorizes (i.e. the numbers $\lambda_j$ are all but one equal to zero), the density matrices of $A$ and $B$ represent mixed states. In particular, they have the same non-zero eigenvalues $\lambda_j$ and hence equal entropies,

$$S(\rho_A) = S(\rho_B) = -\sum_j \lambda_j \log_2 \lambda_j,$$

(5)

which are strictly positive if the state $|\psi_{AB}\rangle$ is a genuine superposition of product vectors, i.e. entangled. In fact, the entropy (5) measures how entangled the state (3) is; it is zero if and only if $|\psi_{AB}\rangle$ is separable, and takes its maximal value $\log_2 d$ for the maximally entangled states which have the Schmidt representation (3) with uniform coefficients $\lambda_j = 1/\sqrt{d}$ (measures of entanglement can also be defined for mixed states of $AB$—we refer the reader to [28, 69, 104] for detailed reviews on that subject).

This passage from the pure joint state of a composite system to the mixed partial state of one of its components can be inverted. Indeed, for any density matrix $\rho_A$ of a system $A$ there is a pure state of a composite system $AB$, where $B$ is large enough to contain a copy $A$, given by the vector (3) such that $\rho_A$ is its partial state. This pure state is called purification of $\rho_A$. Nothing like this exists in classical system theory.

For the sake of completeness it is finally worth mentioning that entanglement is not the only way in which quantum-mechanical states exhibit nonclassical features. Another aspect which is currently attracting a growing interest is the possibility of exploiting the nonorthogonality of states, such as $|0\rangle$ and $|+\rangle$ introduced before, to build correlated joint configurations which, even though being separable, cannot be accounted for by a purely classical theory. A measure for such nonclassical, yet unentangled, correlations is provided by so-called quantum discord, see [81, 139] for details.

### 2.2. Classical and quantum channels

A communication channel is any physical transmission medium (e.g. a wire) that allows two parties (say, the sender and the receiver) to exchange messages. In classical information theory [37], a channel is abstractly modeled by specifying an input alphabet $\mathcal{X}$ and an output alphabet $\mathcal{Y}$ (with $\mathcal{Y}$ not necessarily the same as $\mathcal{X}$). The elements $x$ of $\mathcal{X}$ represent the input signals (or letters) the sender wishes to transfer. Alternatively, one can think of $x$s as the pure (classical) states at the input of the information carrier that propagate through the physical medium that describes the communication line. Similarly, the elements $y$ of $\mathcal{Y}$ represent the output counterparts of the input signals which arrive at the receiver after propagation through the medium. The physical properties of the communication process, including the noise that may affect it, are then summarized by a conditional probability $p(y|x)$ of receiving an output letter $y \in \mathcal{Y}$ when an input signal $x \in \mathcal{X}$ is sent. Hence, if an input probability distribution $P = \{p_x\}$ is given, reflecting the frequencies of different input signals, then the input and the output become random variables $X, Y$ with the joint probability distribution $p_{x,y} = p(y|x)p_x$. Accordingly, the input probability distribution $P = \{p_x\}$ is transformed by the channel to the output probability distribution $P' = \{p'_y\}$, where $p'_y = \sum_x p(y|x)p_x$. A noiseless channel is such that $\mathcal{X} = \mathcal{Y}$ and the probabilities $p(y|x)$ are either 0 or 1 (i.e. it amounts to a permutation of the alphabet $\mathcal{X}$). A particular case is the ideal (identity) channel $I_d$, for which $p(y|x) = \delta_{x,y}$ (Kronecker’s delta).

Looking for a generalization to the quantum domain we should first consider those scenarios in which classical data (i.e. data which could have been stored into the state of a classical carrier represented by the alphabet $\mathcal{X}$) are transmitted through a physical communication line which employs quantum carriers to convey information (e.g. an optical fiber operating at very low intensity of the light [33, 63]). In these configurations the signaling process requires an initial encoding stage in which the elements of $\mathcal{X}$ are ‘written’ into the quantum states of the carrier, and a final decoding stage in which the received quantum states of the carrier are mapped back into classical data by some measurement, as schematically shown in figure 1. The encoding stage of such a scheme is characterized by assigning a classical-quantum (c-q) mapping $x \rightarrow \rho_x$ that defines which density matrix $\rho_x$ of the quantum carrier represents the symbol $x \in \mathcal{X}$ of the input classical message. Physically, $x$ can thus be interpreted as a parameter of a state preparation procedure (the correspondence $x \rightarrow \rho_x$ containing, in a condensed form, the description of the physical process generating the state $\rho_x$). Note that in order to preserve the statistical structure of the process, a mixed input state of
Figure 1. Transferring classical data $x \in X$ via a quantum information carrier. The first stage of the process requires the encoding of $x$ into a quantum state $\rho_x$ of the carrier (c–q mapping); then the carrier propagates along the communication line and its state gets transformed into the density matrix $\rho'_x$ (q–q mapping); finally there is the decoding stage where the receiver of the message tries to recover $x$ by performing some measurement on $\rho'_x$ and obtaining the classical outcome $y$ (q–c mapping).

$X$ defined by the probability distribution $P = \{p_x\}$ is mapped by the c–q channel into the density operator $\rho = \sum_x p_x \rho_x$.

On the other hand, the decoding stage of figure 1 is characterized by a quantum–classical (q–c) mapping that establishes the probabilities of the output letters $y \in Y$ corresponding to the quantum state of the carrier emerging from the quantum communication line. Such a mapping is implemented by a quantum measurement and it is characterized by assigning the probability distribution $p_y(\rho)$ that, given a generic state $\rho$ of the carrier, defines the statistics of the possible measurement outcomes. It can be shown [87] that the linear dependence of $p_y(\rho)$ resulting from the preservation of mixtures, along with general properties of probabilities, implies the following functional structure:

$$p_y(\rho) = \text{Tr} \rho M_y,$$  

for all probability distributions $\{p_x\}$ and for arbitrary density operators $\{\rho_x\}$. An exact characterization of such transformations is clearly mandatory as they constitute the proper quantum counterparts of the stochastic mapping $P \rightarrow P'$ of classical communication theory.

Example. The basic example of a quantum system is the qubit—a two-level quantum system (say, the spin of an electron) characterized by Hilbert $\mathcal{H}$ space of dimensionality 2. Via the specification of a canonical basis in $\mathcal{H}$, its linear operators are represented as linear combinations of the Pauli matrices

$$I \equiv \sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix},$$  

$$\sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \quad (9)$$

This allows us to express the density operators $\rho \in \mathcal{S}(\mathcal{H})$ of the system as points $\hat{a} = (a_x, a_y, a_z)$ in the unit ball in the three-dimensional real vector space $\mathbb{R}^3$ (the Bloch ball) by identifying the coordinates $a_x, a_y, a_z$ with the Stokes parameters of the expansion (in which the condition $\text{Tr} \rho = 1$ is taken into account)

$$\rho = \frac{1}{2}(I + a_x \sigma_x + a_y \sigma_y + a_z \sigma_z) = \frac{1}{2}\begin{bmatrix} 1 + a_z & a_x - ia_y \\ a_x + ia_y & 1 - a_z \end{bmatrix}. \quad (10)$$

Equation (8) then implies that a qubit channel is a linear map in $\mathbb{R}^3$ transforming the Bloch ball into a certain ellipsoid inside the ball. Up to rotations at the input and the output, such a transformation has the following canonical form in the basis of Pauli matrices:

$$a_y \rightarrow a'_y = b_y + a_y t_y, \quad y = x, y, z, \quad (11)$$

with $b_y, t_y$ real numbers (satisfying some further constraints to be discussed later). The equation of the output ellipsoid is $\sum_{y=x,y,z} (a'_y - b_y)^2/t_y = 1$. Hence $b_y$ give the coordinates of the center of the output ellipsoid, while $|t_y|$ its half-axes.

2.3. Dynamics of isolated and open quantum systems

The simplest case of q–q mappings (7) is represented by noiseless quantum channels which describe irreversible transformations of the set of an isolated quantum state $\mathcal{S}(\mathcal{H})$ onto itself. The famous Wigner’s theorem implies [41] that any such mapping is implemented by either unitary or anti-unitary conjugation, which amounts to

$$\Phi[\rho] = U \rho U^\dagger, \quad \text{or} \quad \Phi[\rho] = U \rho^\dagger U^\dagger, \quad (12)$$

where $U$ is a unitary operator on the system Hilbert space $\mathcal{H}$, $(\cdot)^\dagger$ denotes the Hermitian adjoint, while $(\cdot)^\dagger$ is the matrix transposition in a fixed basis. For a qubit, the mappings in equation (12) define, respectively, rotations of the Bloch ball (orthogonal maps with determinant $+1$), and reflection with respect to the $xz$-plane (i.e. matrix transposition in the canonical basis) followed by rotations (orthogonal maps with determinant $-1$). It turns out, however, that only the unitary conjugations (the first case in (12)) can be associated with
proper dynamical processes, the anti-unitary being excluded since they cannot be continuously connected with the identity mapping (the transposition $\rho \rightarrow \rho^\dagger$ is indeed typically identified with time reversal).

As was already mentioned in section 2.1, a basic feature of quantum information as distinct from classical is the impossibility of cloning [193]. Clearly, any classical data can be copied exactly in an arbitrary quantity. But a ‘quantum xerox’, i.e. a physical device which would accomplish a similar task for arbitrary quantum states contradicts the principles of quantum mechanics. Indeed, the cloning transformation

$$|\psi\rangle \in \mathcal{H} \rightarrow |\psi\rangle \otimes \cdots \otimes |\psi\rangle \in \mathcal{H}^{\otimes n} \equiv \mathcal{H} \otimes \cdots \otimes \mathcal{H},$$

is nonlinear and cannot be implemented by a unitary operator (even if operating jointly on some external ancillary system that is traced away at the end of the process). Of course, this can be done for each given state $|\psi\rangle$ (and even for each fixed set of orthogonal states) by a corresponding specialized device, but there is no universal cloner for all quantum states. It should also be noted that approximate cloning of arbitrary quantum states is not allowed, as long as the resulting transformation realizes equation (13) to a certain known degree of accuracy, see, e.g., [151]. The quantum xerox is not the only type of machine which is forbidden by the laws of quantum mechanics. For a comprehensive list of such ‘impossible devices’ we refer the reader to [186].

The evolution of an open system, subject to external influences, whether it be the process of establishing equilibrium with an environment or interaction with a measuring apparatus, reveals features of irreversibility. These transformations constitute the class of noisy quantum channels (7) which cause distortion to the transmitted states of a quantum carrier in its propagation through the communication line. A formal characterization of such processes can be obtained by introducing a Hilbert space $\mathcal{H}_E$ to describe the environmental degrees of freedom $E$ which tamper with the communication, and assigning to it an initial state $\rho_E$. Since the carrier and the environment together form an isolated system, their joint (reversible) evolution can be described by a unitary operator $U$ which, acting nontrivially on the composite Hilbert space $\mathcal{H} \otimes \mathcal{H}_E$, defines their interaction—see figure 2. Consequently, the irreversible q–q mappings (7) can be obtained by averaging off $E$ from the resulting output configurations, i.e.

$$\rho \rightarrow \Phi[\rho] = \text{Tr}_E \left[ U (\rho \otimes \rho_E) U^\dagger \right].$$

This expression provides what is generally called a unitary representation for the quantum channel $\Phi$ and generalizes the unitary evolutions of equation (12) to the case of irreversible dynamics. Let us then assume that the initial environment state is pure, i.e. $\rho_E = |\psi_E\rangle\langle\psi_E|$—via purification of $\rho_E$ this is always possible: the resulting expression is often called the Stinespring representation of $\Phi$ [117, 174]. In this case equation (14) can be written as

$$\Phi[\rho] = \text{Tr}_E V \rho V^\dagger,$$

with $V$ being an isometric operator from $\mathcal{H}$ to $\mathcal{H} \otimes \mathcal{H}_E$ defined by

$$V \equiv U|\psi_E\rangle = \sum_{k=1}^{d_E} V_k \otimes |e^k_E\rangle,$$

where $d_E$ is the dimensionality of $\mathcal{H}_E$ while $\{|e^k_E\rangle\}_{k=1}^{d_E}$ is one of its orthonormal bases (recall that a linear operator from one space to another is isometric if it preserves the norms and hence inner products of vectors. In particular, it satisfies the relation $V^\dagger V = I$ with $I$ being the identity operator of the input space).

In this expression $V_k = \langle e^k_E | U |\psi_E\rangle$ are operators acting on $\mathcal{H}$ which are uniquely determined by the relation

$$\langle \psi | V_k |\psi\rangle = \langle \psi \otimes e^k_E | U |\psi \otimes \psi_E\rangle, \quad |\psi\rangle, |\psi\rangle \in \mathcal{H}. \quad (17)$$

They are called Kraus operators and satisfy the completeness relation

$$\sum_{k=1}^{d_E} V_k^\dagger V_k = I. \quad (18)$$

Taking the partial trace in (15) in the basis $\{|e^k_E\rangle\}_{k=1}^{d_E}$ we finally arrive at the so-called operator-sum (or Kraus) decomposition [117] for the evolution (14)

$$\Phi[\rho] = \sum_{k=1}^{d_E} V_k \rho V_k^\dagger. \quad (19)$$

It is important to note that given a quantum channel $\Phi$ expressed as in equation (19) with some set of operators $\{V_k\}$, it is always possible to extend it (in a highly nonunique way) to the open system dynamics (14) with some effective environment $E$ initialized in a state $\rho_E$, and a unitary transformation $U$ that couples it to the system (or equivalently in terms of an isometric transformation $V$ that maps $\mathcal{H}$ into $\mathcal{H} \otimes \mathcal{H}_E$ as in equation (15)). Furthermore, even the Kraus decomposition for a given channel $\Phi$ is not unique. Indeed, given a Kraus set $\{V_j\}$ for $\Phi$, a new Kraus set $\{W_j\}$, which represents the same quantum channel, can be obtained by forming the following linear combinations:

$$W_j = \sum_{k=1}^{d_E} u_{jk} V_k. \quad (20)$$
where \([u_{jk}]\) is any complex matrix which satisfies the isometry constraint \(\sum_j u_{jk}^* u_{jk} = \delta_{kk}\). Let us stress that here, as anywhere in quantum information science, ‘environment’ means those degrees of freedom of the actual physical environment of the open quantum system which are essentially involved in the interaction and in the resulting information exchange with the system (cf. the notion of ‘faked continuum’ in [173]). Therefore, the question of such a minimal environment for a given channel arises quite naturally (clearly the Kraus decomposition with the minimal number of nonzero components \(d_E\) is unique only up to transformations of the form (20) with \(u_{jk}\) being now unitary matrices; in contrast, the size \(d_E\) of the minimal environment is an important invariant of the channel \(\Phi\) reflecting its ‘noisyness’ and irreversibility). In the minimal decomposition, the Kraus operators are linearly independent, so that their number cannot exceed the dimensionality \(d^2\) of the space of linear operators acting on the system—for instance, the reversible channels of equation (12) have a minimal decomposition with only one Kraus element (the unitary \(U\) which defines them).

**Example 1.** The depolarizing channel (with probability of error \(p\)) is given by the formula (in which \(\rho\) is an arbitrary operator, not necessarily of unit trace)

\[
\Phi[\rho] = (1 - p)\rho + p \frac{I}{d} \text{Tr} \rho, 
\]

where \(\text{dim } \mathcal{H} = d\). This relation describes a mixture of the ideal channel \(\text{Id} : \rho \to \rho\) and of the completely depolarizing channel \(\rho \to (1/d) \text{Tr} \rho\) which transforms any state \(\rho\) into the chaotic state \(1/d\). For the depolarizing channel a Kraus decomposition can be obtained by writing the unit operator and the trace in a fixed basis \(\{|e^j\}\)

\[
\Phi[\rho] = (1 - p) V_0 \rho V_0^\dagger + \frac{p}{d} \sum_{i,j=1}^d V_{ij} \rho V_{ij}^\dagger, 
\]

where \(V_0 = I, V_{ij} = |e_i\rangle \langle e_j|\). Such a Kraus set is clearly not minimal since its elements are linearly dependent (e.g. \(V_0 = \sum_{i=1}^d V_{ii}\)).

**Example 2.** As another example consider the qubit channel (11) with \(b_{xy} = 0\), which contracts the Bloch ball along the axes \(y = x, y, z\) with the coefficients \(|t_{xy}|\) (note that the qubit depolarizing channel corresponds to the uniform contraction with \(t_x = t_y = t_z = 1 - p\)). By using the multiplication rules for Pauli matrices one sees that in this case

\[
\Phi[\rho] = \sum_{y=0,\ldots,3} p_y \sigma_y \rho \sigma_y, 
\]

where

\[
p_0 = (1 + t_x + t_y + t_z)/4, \quad p_x = (1 + t_x - t_y - t_z)/4, \\
p_y = (1 - t_x + t_y - t_z)/4, \quad p_z = (1 - t_x - t_y + t_z)/4,
\]

and the nonnegativity of these numbers is the necessary and sufficient condition for the complete positivity of the map \(\Phi\) (see below). In this case (23) gives the minimal Kraus decomposition for the channel (the size of the minimal environment being equal to the number of strictly positive coefficients \(p_y\)).

Coming back to the decomposition (19) we see that it can be written as

\[
\Phi[\rho] = \sum_{k=1}^{d_E} p_k(\rho) \rho_k, 
\]

where \(p_k(\rho) = \text{Tr}_E M_k\) is the probability distribution associated with a POVM of elements \(M_k = V_k^\dagger V_k\), and \(\rho_k = V_k \rho V_k^\dagger / p_k(\rho)\) are density operators. Noting that \(V_k \rho V_k^\dagger = \langle e_k^\dagger | U (\rho \otimes |\psi_E\rangle \langle\psi_E|) U^\dagger | e_k\rangle\), relation (26) can then be interpreted as follows: after the system and the environment have evolved into the state \(U (\rho \otimes |\psi_E\rangle \langle\psi_E|) U^\dagger\) via the unitary coupling \(U\), in the environment a von Neumann measurement in the basis \(\{|e_k\}\) is performed [181]; the outcome \(k\) appears with probability \(p_k(\rho)\) and the posterior state of the system conditioned upon the outcome \(k\) is the density operator \(\rho_k\).

In other words formula (26) gives the decomposition of the system state \(\Phi[\rho]\) after the (nonselective) measurement into an ensemble of posterior states corresponding to different measurement outcomes.

### 2.4. Heisenberg picture

So far we have worked in the Schrödinger picture describing evolutions of states for fixed observables. The passage to the Heisenberg picture, where the observables of the system evolve while the states are kept fixed, is obtained by introducing the dual channel \(\Phi^*\) according to the rule

\[
\text{Tr} \Phi[\rho] X = \text{Tr} \rho \Phi^*[X],
\]

which must hold for all density operators \(\rho \in \mathcal{S}(\mathcal{H})\) and for all operators \(X\) in \(\mathcal{H}\). The map \(\Phi^*\) is also completely positive admitting the operator-sum decomposition

\[
\Phi^*[X] = \sum_k V_k^\dagger X V_k,
\]

where \(\{V_k\}\) is a Kraus set for \(\Phi\). The property of trace preservation for \(\Phi\) is equivalent to the fact that \(\Phi^*\) is unital, i.e., it leaves invariant the unit operator \(I\): \(\Phi^*[I] = I\). A channel \(\Phi\) in the Schrödinger picture may also occasionally be unital, in which case it is called bistochastic; it leaves invariant the chaotic state \(1/d\). This happens always when the channel is self-dual: \(\Phi = \Phi^*\). An example of such a channel is provided by the depolarizing map of equation (21).

### 2.5. Positivity, complete positivity and the Choi–Jamiolkowski representation

In accordance with the fact that the channel \(\Phi\) must transform quantum states into quantum states, equation (19) ensures preservation of positivity (indeed \(\rho \geq 0\) implies \(V_k \rho V_k^\dagger \geq 0\)), while equation (18) guarantees that the mapping \(\Phi\) preserves the trace (i.e. \(\text{Tr} \Phi[\rho] = \text{Tr} \rho\)). Thus, for arbitrary linear operators \(\rho\) in \(\mathcal{H}\), the transformation (19) defines a positive, trace-preserving, linear mapping.
Quite importantly, it satisfies an even stronger property called complete positivity which means that any extension $\Phi \otimes \text{Id}_R$ of the channel $\Phi$ by an ideal channel $\text{Id}_R$ of a ‘parallel’ reference system $R$ is again positive. In other words, given a quantum channel $\Phi$ operating on the system $A$ which admits the Kraus decomposition, for any state $\rho_{AR}$ of the composite system $\mathcal{H}_A \otimes \mathcal{H}_R$ the equation

$$\rho_{AR}' = \rho_{AR}$$

defines again a proper density operator $\rho_{AR}'$ (indeed, $(\Phi \otimes \text{Id}_R)[\rho_{AR}] = \sum_{k=1}^{d_{\rho}} (V_k \otimes I_R)\rho_{AR}(V_k \otimes I_R)^\dagger \geq 0$ while the trace is obviously preserved). The complete positivity together with the trace preservation (CPTP in brief) is necessary and sufficient for a linear map $\Phi$ to have a Kraus decomposition (19) and thus, a unitary representation (14); hence the CPTP conditions are characteristic for the evolution of an open system and can be used to define a quantum channel.

When the initial state $\rho_{AR}$ is nonseparable, the transformation described in equation (29) is called ‘entanglement transmission’. At the physical level it represents those processes by which the sender creates locally some entangled configuration of the systems $A$ and $R$, and then sends half of it (specifically the carrier $A$) to the receiver through a q–q channel. The fact that the mapping $\Phi$ entering in equation (29) is completely positive guarantees that the resulting configuration can still be represented as a density operator, and hence that the overall transformation admits a quantum-mechanical description. Notably not all positive maps which take the set of states $\mathcal{S}(\mathcal{H}_A)$ of a given system $A$ into itself are completely positive.

The canonical example is the transposition $T[\rho_A] = \rho_A^\dagger$ in a fixed basis $\{|e_i^A\rangle\}$, see the second relation in equation (12). Since $\rho_A^\dagger$ and $\rho_A$ share the same spectral properties, it is clear that $T$ maps the states of $A$ into states (hence it is positive). However, applying the map $T \otimes \text{Id}_R$ to a maximally entangled state $\frac{1}{\sqrt{d_A}} \sum |e_i^A \otimes e_j^R\rangle$ of the composite system $AR$ one obtains the operator

$$\frac{1}{d_A} (T \otimes \text{Id}_R) \left[ \sum_{i,j=1}^{d_A} |e_i^A \otimes e_j^R\rangle \langle e_i^A \otimes e_j^R| \right]$$

which is not a proper state of $AR$ since it is not positive (its expectation value on the vector $|e_1^A \otimes e_2^R\rangle - |e_2^A \otimes e_1^R\rangle$ is negative). Such weird behavior of $T$ should not be surprising: as anticipated at the beginning of the section, in physics transposition is related to the time inversion. The operation $T \otimes \text{Id}_R$ is thus something like doing time inversion only in one part of the composite system $AR$. To avoid this unphysical behavior, one should hence work under the explicit assumption that the q–q mappings entering equation (7) satisfy the complete positivity condition.

In the case of CPTP mappings, the right-hand side of equation (29) obtained by substituting the maximally entangled state associated with the vector $\frac{1}{\sqrt{d_A}} \sum |e_i^A \otimes e_j^R\rangle$ for $\rho_{AR}$ provides the Choi–Jamiołkowski state of $\Phi$:

$$\rho_{AR}^{(\Phi)} = \frac{1}{d_A} \sum_{i,j=1}^{d_A} \Phi \left[ |e_i^A \rangle \langle e_i^A| \otimes |e_j^R \rangle \langle e_j^R| \right],$$

which uniquely determines the channel through the inversion formula

$$\Phi[\rho_A] = d_A \text{Tr}_R \left[ \rho_{AR}^{(\Phi)} (I_A \otimes \rho_R^\dagger) \right].$$

Here $\rho_R^\dagger = \sum_{i,j} (e_i^A |\rho_A| e_j^A)i \langle e_j^R | e_j^R\rangle$ is the state of $R$ which is constructed via transposition of the density matrix of $\rho_A$.

2.6. Compositions rules

Suppose that we have a family $\{\Phi_a\}$ of quantum channels where the values of the parameter $\alpha$ appear with probabilities $p_a$. Then the mapping defined as

$$\Phi[\rho] = \sum_a p_a \Phi_a[\rho]$$

is still CPTP and hence defines a channel (with a fluctuating parameter $\alpha$). In other words, the set of channels is convex, i.e. closed under convex combinations (mixtures). The extremal (pure) channels are those CPTP transformations which admit a Kraus decomposition (19) with $V_k$ such that $\{V_k V_k^\dagger\}_{k,k'}$ form a collection of linearly independent operators [36] (in particular, the unitary evolutions (12) are extremal).

The set of channels also possesses a semigroup structure under concatenation. Indeed the transformation $\Phi_2 \circ \Phi_1$ obtained by applying $\Phi_2$ at the output of the quantum channel $\Phi_1$, i.e.

$$\langle \Phi_2 \circ \Phi_1 | [\rho] \rangle = \Phi_2 \left[ \Phi_1[\rho] \right],$$

is still CPTP and hence a channel. Note that the product ‘$\circ$’ is in general neither commutative (i.e. $\Phi_2 \circ \Phi_1 \neq \Phi_1 \circ \Phi_2$), nor does it admit inversion (the only quantum channels which admit the CPTP inverse are the reversible transformations defined by the unitary mappings (12)). The possibility of concatenating quantum channels with unitary transformations turns out to be useful in defining equivalence classes which, in several cases, may help in simplifying the study of the noise effects acting on a system. In particular, two maps $\Phi_1$ and $\Phi_2$ are said to be unitary equivalent when there exist unitary transformations $V$ and $U$ such that $\Phi_2 = \Phi_U \circ \Phi_1 \circ \Phi_V$ (here $\Phi_U[\cdot] \equiv U[\cdot]U^\dagger$ is the quantum channel associated with the unitary $U$, similarly for $\Phi_V$). For instance, exploiting this approach, classifications for the quantum channels operating on a qubit [149] or on a single bosonic mode [31, 32, 91] have been realized (see also section 3.5).

Finally, complete positivity allows us to consider tensor products of channels which are most important in information theory, where tensor products describe parallel or block channels. Indeed, one can write the tensor product as a concatenation of completely positive maps

$$\Phi_1 \otimes \Phi_2 = (\Phi_1 \otimes \text{Id}_2) \circ (\text{Id}_1 \otimes \Phi_2),$$

which is again completely positive.
2.7. Complementary channels

For simplicity up to now we have considered channels with the same input and output spaces. From the point of view of transformation theory, however, it is quite natural to have the possibility for them to be different. This is done by considering transformations (19) with operators $V_k$ mapping the input space $\mathcal{H}_A$ into the output space $\mathcal{H}_B$ while still satisfying the completeness relation (18) or, equivalently, by considering in (15) isometric transformation $V$ which connects $\mathcal{H}_A$ to $\mathcal{H}_B \otimes \mathcal{H}_E$. In such cases we shall often use self-explanatory notation $\Phi = \Phi_{A \to B}$.

Example. The quantum erasure channel [13, 68] gives a natural example of a channel with different input and output spaces. It represents a communication line which transmits the input state $\rho$ intact with probability $1 - p$ and ‘erases’ it with probability $p$ by replacing it with an erasure signal $|e\rangle$ that is orthogonal to $\rho$. Formally, it can be defined as the CPTP map

$$\rho \mapsto (1 - p)\rho \oplus p|e\rangle\langle e| \text{Tr}\rho,$$

which operates from the input space $\mathcal{H}_A$ to the output space $\mathcal{H}_B = \mathcal{H}_A \oplus \{|e\rangle\}$ obtained by ‘adding’ an extra orthogonal vector $|e\rangle$ to $\mathcal{H}_A$.

A quite important case where the input and output systems differ arises in the open system description (14) of a channel $\Phi$ that maps a system $A$ into itself. Indeed suppose that, instead of the evolution of $A$, we are interested in the state change of the environment $E$ as a function of the input state $\rho$ of $A$. This is given by the following mapping:

$$\rho \mapsto \text{Tr}_A U (\rho \otimes \rho_E) U^\dagger,$$

where in contrast to equation (14), the partial trace is taken with respect to $\mathcal{H} = \mathcal{H}_A$ instead of $\mathcal{H}_E$. The transformation (36) defines a quantum channel that connects states in $\mathcal{H}_A$ to states of the channel environment. There is a complementarity relationship between the mappings (14) and (36) which reflects various aspects of information–disturbance trade-off in the evolution of an open quantum system. This relation becomes especially strict in the case of pure environmental state $\rho_E$ when it amounts to the notion of complementary channels (if $\rho_E$ is a mixed state such a complementarity is weakened and the two maps are sometimes called weakly complementary [32]).

More generally, complementary channels can also be defined in a three quantum systems setting. Given the spaces $\mathcal{H}_A$, $\mathcal{H}_B$, $\mathcal{H}_E$ that represent the systems $A$, $B$, $E$, and an isometric operator $V : \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_E$ which connects them, we say that the mappings

$$\Phi_\rho(\rho) = \text{Tr}_E V \rho V^\dagger, \quad \tilde{\Phi}_\rho(\rho) = \text{Tr}_B \rho E V^\dagger,$$

with $\rho$ being a density operator in $\mathcal{H}_A$, define two channels which are complementary [47, 90, 112] (note that they reduce to equations (14) and (36), respectively, in the case $B = A$). Expressing the isometry as in equation (16) where again $|\langle e^*_E\rangle\rangle_k$ is an orthonormal basis of $\mathcal{H}_E$ while the $V_k$’s are now operators from $\mathcal{H}_A$ to $\mathcal{H}_B$, the above relations can be rewritten as

$$\Phi[\rho] = \sum_{k=1}^{d_2} V_k \rho V_k^\dagger, \quad \tilde{\Phi}[\rho] = \sum_{k, \ell=1}^{d_E} |\langle e^*_E\rangle\rangle_k \langle e^*_E\rangle\langle \ell| \text{Tr}\left[\rho V_k^\dagger V_\ell\right],$$

the first one being a Kraus decomposition for $\Phi$. It can be shown that the second relation in equation (38) determines all complementary channels of $\Phi$ uniquely up to unitary equivalence, moreover, by using a similar construction for $\tilde{\Phi}$ one can check that its complementary is $\tilde{\Phi} = \Phi$. A similar relation does not hold for the case of weakly complementary channels which are not unique and depend on the selected unitary representation (14) of the original map $\Phi$.

Example. A complementary channel associated with the identity map Id on $\mathcal{H}$ (i.e. Id$[\rho] = \rho$ for all $\rho$) is any map $\Phi$ that transforms all $\rho$ into a fixed pure output state $|\psi_E\rangle$ of the environment, i.e.

$$\Phi(\rho) = |\psi_E\rangle\langle \psi_E| \text{Tr}\rho,$$

(to see this put $\mathcal{H}_A = \mathcal{H}_B$ in equation (37) and take the isometry $V : \mathcal{H}_A \to \mathcal{H}_A \otimes \mathcal{H}_E$ to be the operator $V = I \otimes |\psi_E\rangle$). Further, a channel that transforms the density operators $\rho$ into a fixed non-pure output state $\rho_E$ in $\mathcal{H}_E$ is weakly complementary to the identity channel Id. (Similarly one finds that the complementary to the erasure channel (35) is again an erasure channel where parameter $p$ has been replaced by $1 - p$.)

This example illustrates the fact that perfect transmission of quantum information through a quantum channel is equivalent to the absence of quantum information transfer from the input to the environment $E$, and vice versa. An approximate version of this complementarity expresses the quantum principle of information–disturbance trade-off: the closer the channel $\Phi$ is to an ideal channel, the closer is its complement $\tilde{\Phi}$ to the completely depolarizing one [40].

2.8. Quantum channels involving a classical stage

As shown in figure 1, the process of transmission of classical information through a quantum communication link includes encoding and decoding stages which are special cases of quantum channels with classical input, respectively output. Furthermore, a quantum information-processing system, such as a quantum computer, may include certain sub-routines where the data have essentially classical nature. Here we characterize such mappings by describing them as special instances of CPTP maps.

(i) The c–q channel. In section 2.2 we have seen that this class of channels is uniquely defined by a map $x \to \rho_x$, describing encoding of the classical input $x$ into quantum states $\rho_x$ in the output space $\mathcal{H}_B$. This process can be represented as a quantum channel by introducing the input space $\mathcal{H}_A$ spanned by an orthonormal basis $\{|e^*_E\rangle\rangle_k\rangle_k$, so that

$$\rho \mapsto \Phi^{\text{c-q}}_{A \to B}[\rho] = \sum_{x \in \mathcal{X}} |\langle e^*_E\rangle\rangle_k \rho_x, \quad (40)$$

where $\rho$ is a density operator in $\mathcal{H}_A$ (accordingly output ensembles of the form $\rho = \sum_x \rho_x \rho_x$ are generated by taking the input states $\rho$ such that $|\langle e^*_E\rangle\rangle_k \rho_x = \rho_x$).

An extreme case of a c–q channel is the completely depolarizing channel, describing the ultimate result of an irreversible evolution to a final fixed state as in equation (39).
(ii) The q–c channel. Similarly to the previous case, q–c transformations can be represented as a CPTP map connecting an input quantum system $B'$ to an output quantum system $A'$, by introducing for the latter an orthonormal basis $\{|e_{\alpha}\}_{\alpha < Y}$ to represent the classical outcomes $Y$,

$$\rho \rightarrow \Phi^{\text{q-c}}_{B'\rightarrow A'}[\rho] = \sum_{\alpha < Y} |e_{\alpha}\rangle\langle e_{\alpha}| \text{Tr}_{B'} \rho M_{\alpha}, \quad (41)$$

where $\{M_{\alpha}\}$ is a POVM in the input space $\mathcal{H}_A$.

The transmission of classical information through a quantum communication line can now be modeled as a concatenation of such maps, i.e. $\Phi^{\text{q-c}}_{B'\rightarrow A''} \circ \Phi^{\text{q-c}}_{A''\rightarrow A'} \circ \Phi^{\text{q-c}}_{A\rightarrow B'}$ with $\Phi_{B'\rightarrow B'}$ being a generic CPTP map from $B$ to $B'$. Of particular interest in quantum information theory is also the class of maps that arise when the ordering of such a concatenation is reversed. In particular, identifying $A'$ with $A$ and assuming that the bases $\{|e_{\alpha}\}_{\alpha < Y}$, and $\{|e_{\alpha}\}_{\alpha < Y}'$, coincide, we obtain a q–c–q channel from $B'$ to $B$ of the form

$$\rho \rightarrow \Phi^{\text{q-c}}_{B'\rightarrow B}[\rho] = \left(\Phi^{\text{q-c}}_{A\rightarrow A'} \circ \Phi^{\text{q-c}}_{A'\rightarrow B'}\right)[\rho] = \sum_{x} \rho_x \text{Tr}_{B'} \rho M_x, \quad (42)$$

which describes the quantum measurement on $B'$ followed by a state preparation on $B$ depending on the outcome of the measurement. Such mappings include the c–q and q–c channels as special cases (i.e. c–q channels are obtained by taking $M_x$ to be projections onto the vectors of orthonormal basis, while q–c channels are obtained when $\rho_x$ are orthogonal pure states).

The channels defined above are characterized by the property of entanglement breaking: operating with them on the half of entangled state $\rho_{R'B'}$ of a composite system $B'R'$ produces unentangled state of $BR$ (in particular, their Choi–Jamesiokowski state (30) is separable). This is not surprising, because in the q–c–q channel the quantum information passes through an intermediate stage (represented by system $A$) in which information can only be encoded as a classical state. Less obvious is the fact that any entanglement-breaking channel has the form (42) [101]. Quite non-obviously, the depolarizing channel (21) is entanglement breaking if the error probability $p \geq d/(d+1)$ [17].

In all the cases described above the complete positivity can be seen by explicitly producing a Kraus decomposition for the map. For instance, let $\Phi$ be an entanglement-breaking channel of the form (42). Then by making the spectral decompositions of the operators $\rho_x$, $M_x$ one arrives at a Kraus decomposition with rank 1 operators of the form

$$\Phi[\rho] = \sum_{\alpha} |\psi^{\alpha}\rangle \langle \psi^{\alpha}| \rho |\psi^{\alpha}\rangle \langle \psi^{\alpha}|, \quad (43)$$

where $|\psi^{\alpha}\rangle$ are unit vectors while $|\psi^{\alpha}\rangle \psi^{\alpha}\rangle$ form an overcomplete system for the input space, i.e.

$$\sum_{\alpha} |\psi^{\alpha}\rangle \langle \psi^{\alpha}| = I. \quad (44)$$

According to the general formula (38), the complementary map associated with equation (43) is now obtained as

$$\Phi[\rho] = \sum_{\alpha, \beta} c_{\alpha\beta} |\psi^{\alpha}\rangle \langle \psi^{\alpha}| \rho |\psi^{\beta}\rangle \langle \psi^{\beta}|,$$

where $\{|e^{\alpha}\rangle\}$ is an orthonormal set, and $c_{\alpha\beta} = \langle \psi^{\alpha}| \rho |\psi^{\beta}\rangle$ is a nonnegative semi-definite matrix with units on the diagonal. In the case where the input and output spaces of $\Phi$ coincide and $|\psi^{\alpha}\rangle = |e^{\alpha}\rangle$, such channels amount to elementwise multiplication of the matrices $\{|e^{\alpha}\rangle |e^{\beta}\rangle\}$ and $\{c_{\alpha\beta}\}$, called the Schur (or Hadamard) product. Such channels are called dephasing [47] (or diagonal [112]) since they suppress the off-diagonal elements of the input density matrix. From (43) we see that the dephasing channels are complementary to a particular class of entanglement-breaking channels, namely to c–q channels. For another special subclass of the q–c channels, $\{|\psi^{\alpha}\rangle\}$ of equation (43) constitute an orthonormal basis, so that $c_{\alpha\beta} = \delta_{\alpha\beta}$ and the complementary channel is a q–c channel which is called completely dephasing since it amounts to nullifying the off-diagonal elements of the density matrix $\{|e^{\alpha}\rangle |e^{\beta}\rangle\}$.

3. Bosonic Gaussian channels

Until now we have considered quantum channels from a rather abstract point of view focusing on models in which the quantum carriers are effectively described as finite-dimensional systems (say, collections of qubits). In many real experimental scenarios such a description is valid only approximately. To begin with, the fundamental physical information carrier is the electromagnetic field which is known to be mathematically equivalent to an ensemble of oscillators that in quantum mechanics are described as infinite-dimensional systems [64, 114, 127, 194]. This is a typical example of ‘continuous variables’ bosonic quantum system [27, 35] whose basic observables (oscillator amplitudes) satisfy the canonical commutation relations (CCR) (other examples include vibrational modes of solids, atomic ensembles, nuclear spins in a quantum dot and Bose-Einstein condensates). Many of the current experimental realizations of quantum information processing are carried out in such systems [27, 33, 35, 62, 63]. In particular, the most impressive examples of quantum communication channels have been realized in this way by using optical fibers [129, 175–177] or free space communication [140, 152, 178].

In the context of quantum-optical communication [33] particularly important are the Gaussian states, including coherent and squeezed states realized in lasers and nonlinear quantum-optical devices, and the corresponding class of quantum information processors—the Gaussian channels [29, 49, 96, 182, 183]. The latter provide the proper mathematical description for the most common sources of noise encountered in optical implementations, including attenuation, amplification and thermalization processes [33, 96]. This section is devoted to a review of the basics properties of this special class of channels.
3.1. Example: channel with additive Gaussian quantum noise

As a starter consider the case of a single electromagnetic mode which describes the propagation of photons of a given frequency $\omega$ and fixed polarization through an optical fiber. The physics of the system is fully determined by the annihilation and creation operators

$$a = \frac{1}{\sqrt{2\omega}} (\omega Q + iP), \quad a^\dagger = \frac{1}{\sqrt{2\omega}} (\omega Q - iP),$$

(46)

where $Q$, $P$ are the canonical operators (quadratures) which effectively represent the field and satisfy the Heisenberg CCR

$$[Q, P] = i\hbar,$$

(47)

(in the Schrödinger representation $Q = x$ and $P = i^{-1} d/dx$, so that the physical momentum operator is $p = \hbar P$). Suppose that initially the mode is in the thermal (Gibbs) state represented by the density matrix

$$\rho_0 = \frac{\exp[-\beta H]}{Tr \exp[-\beta H]},$$

(48)

where $H = \frac{1}{2}(\omega^2 Q^2 + P^2) = \omega(a^\dagger a + \frac{1}{2})$ is the Hamiltonian of the system while $\beta > 0$ is the inverse temperature of the state $\rho_0$ related to its average photon number $N = Tr \rho_0 a^\dagger a$ by the identity $\beta = \omega^{-1} \ln(N + 1)/N$. For $\mu \in \mathbb{C}$ (the set of all complex numbers) we define the displaced version of $\rho_0$ as the state

$$\rho_\mu = D(\mu) \rho_0 D(\mu)^\dagger,$$

(49)

which can be obtained, to a good approximation, by pumping the mode with a laser far above the threshold [64, 194]. In the above expressions $\mu$ is the complex amplitude of the mode, while $D(\mu) = \exp(a^\dagger \mu - a \mu^\dagger)$ is the displacement operator which effectively describes the action of the laser.

One can consider the mapping $\mu \to \rho_\mu$ as a c-q channel (40) with continuous alphabet $\mathbb{C}$, which encodes the signal $\mu$ into the quantum state $\rho_\mu$. From a mathematical point of view it is the simplest channel describing transmission of a classical signal on the background of additive quantum Gaussian noise. This model can be naturally generalized to describe many-mode (broadband) c-q channels [33, 86] also in the presence of squeezing [95].

3.2. Multimode bosonic systems

A bosonic system with $s$ degrees of freedom (modes) can be described in terms of canonical observables $Q_1, P_1, \ldots, Q_s, P_s$ which satisfy the Heisenberg CCR

$$[Q_j, P_k] = i\hbar \delta_{jk}, \quad [Q_j, Q_k] = [P_j, P_k] = 0.$$  

(50)

In quantum optics [64, 194] they coincide with the system quadratures, and are related to the bosonic creation–annihilation operators via the identities $a_j^\dagger = (Q_j - i\omega_j P_j)/\sqrt{2\omega_j}$, $a_j = (Q_j + i\omega_j P_j)/\sqrt{2\omega_j}$, where $\omega_j$ are the frequencies of the modes. Introducing the unitary Weyl operators $W(z) \equiv \exp[i Rz]$ where $R \equiv [Q_1, P_1, \ldots, Q_s, P_s]$ and $z$ is the column vector of real parameters $z = [x_1, y_1, \ldots, x_s, y_s]^T$, the relations (50) can be rewritten in the equivalent Weyl–Segal form

$$W(z)W(z') = W(z + z') \exp \left[ i \sum_{j=1}^2 \Delta (z_j, z'_j) \right] = W(z)W(z) \exp[i \Delta (z, z')],$$

(51)

where

$$\Delta (z, z') = \sum_{j=1}^2 (x'_j - x_j, y'_j - y_j) = z^T \Delta z',$$

(52)

is the canonical symplectic form given by the skew-symmetric block matrix

$$\Delta = \text{diag} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

(53)

The $2s$-dimensional real vector space $\mathbb{R}^{2s}$ equipped with (52) is called the symplectic space and describes the classical phase space underlying the quantum bosonic system [87].

The Weyl operators are closely related to displacement operators by inducing the translations $W'(-\Delta z)RW(-\Delta z) = R + z^\dagger$. Moreover, any reasonable function of the canonical observables $\{Q_j, P_j\}$ is or can be approximated by a complex linear combination of the Weyl operators $W(z)$. In particular, a state $\rho$ of the system is uniquely described by its characteristic function $\phi_\rho(z)$ defined by the relations

$$\phi_\rho(z) = Tr \rho W(z), \quad \rho = \int \phi_\rho(z) W(-z) \frac{d^{2s}z}{(2\pi)^s},$$

(54)

where in the second expression the integral is performed on the whole space $\mathbb{R}^{2s}$. A state $\rho$ is called Gaussian if $\phi_\rho(z)$ has the typical Gaussian form

$$\phi_\rho(z) = \exp \left(i m^T z - \frac{1}{2} z^T \alpha z \right),$$

(55)

where $m \in \mathbb{R}^{2s}$ and $\alpha$ is a real symmetric $2s \times 2s$ matrix. Similarly to the case of classical Gaussians, the components of $m^T$ are mean values, i.e. the expectations of the elements of $R = \{Q_1, P_1, \ldots, Q_s, P_s\}$, and $\alpha$ is their (symmetrized) correlation matrix in the state $\rho$. However, in the quantum case $\alpha$ is not only nonnegative definite but satisfies the matrix (Robertson’s) uncertainty relation

$$\alpha \geq \pm i \Delta/2.$$  

(56)

Examples of Gaussian states include vacuum, coherent, squeezed states and also thermal states of oscillator systems, as well as their displacements induced by action of a fixed source. For instance the one-mode state (49) is Gaussian with the characteristic function

$$\phi_{\rho_\mu}(z) = \exp \left[ i (m q x + m p y) - \frac{1}{2} \ominus \frac{1}{2} \left( N + \frac{1}{2} \right) (x^2 + y^2) \right].$$

(57)

where $z = [x, y]^T$, $m = (\omega m_q + i m_p)/\sqrt{2}\omega$, while equation (56) amounts to $N \geq 0$, see, e.g., [87].
3.3. Quantum Gaussian channels

We now analyze those CPTP maps $\rho \rightarrow \rho' = \Phi[\rho]$ which, when operating on Gaussian input states, preserve their structure. It turns out that a convenient way to represent such mappings is to prescribe the action of their duals $\Phi^*$ (defined in section 2.4) on the Weyl operators $W(z)$. One can show that the requirement that the resulting dynamics induces a linear transformation of the canonical observables of an open bosonic system interacting with a Gaussian environment implies

$$\Phi^*[W(z)] = W(Kz) \exp \left( i(\zeta^T z - \frac{1}{2}\zeta^T \zeta) \right),$$  \hspace{1cm} (58)

where $l$ is the vector in $\mathbb{R}^{2s}$ while $K$ and $\beta$ are real $2s \times 2s$-matrices which satisfy the following uncertainty relation:

$$\beta \geq \pm i \left[ \Delta - K^T \Delta K \right]/2.$$ \hspace{1cm} (59)

This inequality is a necessary and sufficient condition for the complete positivity of $\Phi$ [43, 49, 96]. Thus, a Gaussian channel is completely characterized by the three quantities $(K, l, \beta)$, where $\beta$ satisfies (59) (note that there is no restriction on $l$). As any completely positive map, a Gaussian channel admits a Kraus decomposition, explicitly described in [105] in the case of one mode.

Using equations (27) and (54) one can verify that the mean and the correlation matrix $\alpha$ of an input state $\rho$ are transformed by the channel according to the relations

$$m \rightarrow m' = K^T m + l, \quad \alpha \rightarrow \alpha' = K^T \alpha K + \beta,$$ \hspace{1cm} (60)

which, due to equation (59), guarantee that the output counterparts of the Gaussian inputs (54) are again Gaussian states.

A special subclass of the Gaussian channels defined above is obtained for $\beta = 0$. Under this constraint, equation (59) forces the matrix $K$ to be symplectic, i.e. to fulfill the condition

$$K^T \Delta K = \Delta.$$ \hspace{1cm} (61)

Equation (58) implies that these special channels induce the following linear transformation of the canonical observables:

$$R \rightarrow R' = \Phi^*[R] = RK + l^T,$$ \hspace{1cm} (62)

whose components still satisfy the CCR (50) due to equation (61) (for this reason the mapping (62) is called canonical). These transformations represent a unitary evolution of the system which, in the Schrödinger representation, can be expressed as a concatenation of two elementary processes in which one first shifts the coordinates of the system, and then applies a proper symplectic transformation. The first process amounts to the unitary transformation $U_K$ in $\mathcal{H}$ (quantization of the symplectic matrix $K$) such that [87]

$$U_K^* W(z) U_K = W(Kz), \quad \text{hence} \quad U_K^* R U_K = RK.$$ \hspace{1cm} (63)

Any reversible dynamics of bosonic system which in the Heisenberg picture is linear in the canonical observables as in equation (62) can be described in these terms. This is the case for the unitary transformations $\exp[iH]$ induced by Hamiltonians $H$ which are at most quadratic polynomials of the canonical observables, such as the free Hamiltonian of a set of harmonic oscillators, or the Hamiltonian governing a linear amplifier optical process [194].

The presence of irreversible, noisy dynamics for Gaussian channels is signaled by having $\beta \neq 0$ in equation (58). In this case it is possible to produce a unitary representation (14) of $\Phi$ in terms of a unitary coupling with a multimode bosonic environment $E$ characterized by the canonical observables $R_E$ that induces the following linear input–output relation:

$$R \rightarrow R' = RK + R_E K_E, \quad R_E \rightarrow R_E = RL + R_E L_E.$$ \hspace{1cm} (64)

One can show that $\Phi^*[W(z)\rho W(z)^\dagger] = W(Kz)\Phi[\rho]W(K'z)$, \hspace{1cm} (66)

where $K' = \Delta^{-1} K^T \Delta$. Due to this property, the displacement parameter $l$ entering (58) can always be removed as it can be compensated by a proper unitary transformation at the input or at the output of the communication line.

3.4. General properties of the Gaussian channels

Before discussing some specific example of Gaussian channels in the next section we shall list their general properties [96].

(i) Gaussian states are transformed into Gaussian states.

(ii) The dual of a Gaussian channel transforms a polynomial in the canonical variables into a polynomial of the same degree.

(iii) The concatenation of Gaussian channels is again a Gaussian channel. In fact, let $\Phi_2 \circ \Phi_1$ be two Gaussian channels characterized by the parameters $(K_j, l_j, \beta_j)$, then by using (58), the composite map $\Phi_2 \circ \Phi_1$ is Gaussian with parameters

$$K = K_1 K_2, \quad l = K_1^T l_1 + l_2,$$ \hspace{1cm} (65)

$\beta = K_1^T \beta_1 K_2 + \beta_2$.

(iv) Gaussian channels are covariant under the action of Weyl operators. That is for all $\rho$ one has

$$\Phi[W(z)\rho W(z)^\dagger] = W(K'z)\Phi[\rho]W(K'z),$$ \hspace{1cm} (66)

where $K' = \Delta^{-1} K^T \Delta$. Due to this property, the displacement parameter $l$ entering (58) can always be removed as it can be compensated by a proper unitary transformation at the input or at the output of the communication line.

(v) If $K$ is invertible, then $\Phi[I] = |\det K|^{-1} I$, in particular, $\Phi$ is unital if and only if $|\det K| = 1$ [94]. If $\Phi$ is a Gaussian channel with parameters $(K, l, \beta)$, then $|\det K| \Phi^*$ is a Gaussian channel with parameters $(K^{-1}, -(K^{-1})^T l, (K^{-1})^T \beta K^{-1})$ (for channels in one mode this duality was observed in [105]).
Additive classical noise

Let $\Phi$ be a quantum Gaussian channel with parameters $(K, I, \beta)$. It is entanglement breaking if and only if $\beta$ admits the decomposition \cite{93}

$$\beta = \alpha + v,$$

where $\alpha \geq \frac{i}{2} \Delta$, $v \geq \frac{i}{2} K^T \Delta K$. \hfill (67)

In this case $\Phi$ admits a representation of the form

$$\Phi[\rho] = \int W(z) \sigma_B W^\dagger(z) p_\rho(z) \, dz,$$ \hfill (68)

where $p_\rho(z)$ is the probability density in the state $\rho$ of a Gaussian measurement with outcomes $z$, followed by the displacement $W(z)$ of a Gaussian state $\sigma_B$ with covariance matrix $\alpha$. Gaussian measurements are those POVMs which transform quantum Gaussian states into Gaussian probability distributions on $\mathbb{R}^{2s}$, see, e.g., \cite{71} for details. In quantum optics these are implemented by optical homo- and/or hetero-dynes combined with linear multiport interferometers \cite{194}.

3.5. The case of one mode

For ‘continuous-variables’ systems the one-mode channels $(s = 1)$ play a role similar to qubit channels for finite systems. Therefore, it is interesting and important to have their classification under unitary equivalence with respect to canonical transformations \cite{62}, a problem for which a solution is given in \cite{32, 91}. In what follows we consider the most relevant examples of those maps whose efficiency in transferring information will be presented in section 4.10.

(i) Attenuation channels. These are characterized by transformation \cite{58} with $l = 0$

$$K = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \beta = \begin{bmatrix} N_0 + \frac{|1 - k^2|}{2} & 0 \\ 0 & 1 \end{bmatrix},$$ \hfill (69)

where $k \in (0, 1)$ and $N_0$ is nonnegative—the latter being the condition which enforces the inequality \cite{59}. The resulting mapping can be described in terms of a coupling \cite{14}, mediated by a beam splitter of transmissivity $k$ \cite{33, 194}, between the input state of the system and an external bosonic environmental mode initialized in a thermal (Gibbs) state \cite{48} with the mean photon number $N = N_0/[1 - k^2]$.

(ii) Amplification channels. These are described by the same matrices $K$ and $\beta$ of equation \cite{69} where now, however, the parameter $k$ assumes values larger than 1. As in the previous case the resulting mapping can be described in terms of a linear coupling with an environmental state initialized in a thermal state with photon numbers $N_0$. In this case, however, the coupling is provided by a two-mode squeezing Hamiltonian which induces a linear amplification of the impinging field \cite{194}.

(iii) Additive classical noise channels. These maps are characterized by the parameters

$$K = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \beta = N_0 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$ \hfill (70)

with $N_0 \geq 0$. In this case it is possible to write a simple integral representation for the channel

$$\Phi[\rho] = \frac{1}{\pi N} \int D(\zeta) \rho D^\dagger(\zeta) \exp \left(-|\zeta|^2/N\right) \, d^2\zeta,$$ \hfill (71)

with $D(\zeta) = \exp[\zeta a^\dagger - \zeta a]$, which is a continuous analog of the Kraus decomposition expressing the fact that the channel performs random Gaussian displacements of the input state as a result of action of a classical random source, e.g., see, \cite{71}. These maps have also been analyzed in the context of universal cloning machines \cite{34, 124}.

Let us now see which of the above channels are entanglement-breaking based on the decomposability criterion \cite{67}. To do so we rely upon the simple fact that $(N + 1/2)I \geq i\Delta/2$ if and only if $N \geq 0$. Therefore since we have $K^T \Delta K = k^2 \Delta$, it follows that in this case the decomposability condition holds if and only if $\beta \geq \frac{1}{4} (1 + k^2) \Delta$ which is equivalent to $N_0 + |1 - k^2|/2 \geq (1 + k^2)/2$ or

$$N_0 \geq \min \left(1, k^2 \right).$$ \hfill (72)

This gives the condition for entanglement breaking applicable to the channels of classes (i), (ii) and (iii).

4. Entropy, information and channel capacities

The most profound results concerning the essence of information theory—coding theorems—have asymptotic nature and concern the transmission of messages formed by arbitrarily long, ordered sequences of symbols which are emitted by a ‘source’ (say a radio station). In practical applications the index that enumerates the various entries of the message represents either the (discrete) time when they appear sequentially in the communication process, or the number of ‘modes’ in which a given signal can be processed simultaneously in parallel (e.g. the frequency modes in which a radio wave can be decomposed, or the cells that form a memory register). In a statistical description a question inevitably arises regarding what model for correlations between the different symbols of a message should be accepted. The simplest yet most basic is the memoryless model, where both the production of the symbols emitted by the source and their subsequent transformations associated with the propagation through a communication line are statistically identical and independent. The coding theorems of information theory are not at all trivial for this case. At the same time they give a basis for considering more complicated and realistic scenarios taking into account memory effects. Moreover, the effects of entanglement are demonstrated in the memoryless case in the most spectacular way. Therefore, in our presentation we concentrate on the memoryless configurations, providing, where appropriate, references for the more advanced memory models.

In the following we start reviewing some basic results of the classical information theory that allow us to evaluate the information content of a classical message (Shannon’s first coding theorem) and to introduce the notion of capacity.
for a communication channel (Shannon’s second coding theorem). Moving into the quantum domain we will then face the fundamental issue of how to generalize these results in order to have a proper quantum-mechanical measure of information. The task is particularly challenging due to the complex nature of quantum information (see the discussion in section 2.1). In particular, there is a fundamental conceptual distinction between the amount of classical information that it is possible to store in a quantum system, and the amount of quantum information that can be accomplished. At the level of communication theory, this forces the introduction of different alternative generalizations of the Shannon capacity, see, e.g., [16, 166].

4.1. Information transmission over a classical channel

Consider a classical source of information which emits $n$-long sequences $w = (x_1, \ldots, x_n) \in \mathcal{X}^n$ formed by symbols $x_j$ that are randomly extracted from an alphabet $\mathcal{X}$ according to a distribution $P = \{p_x\}$ (the process being repeated identically and independently for each element of the sequence as mentioned previously). In classical information theory, the amount of information that is contained in one of those sequences is measured by the minimal number of binary digits (bits) necessary for its binary representation (coding). The first Shannon coding theorem [160] says that this number is $nH(X)$, where $H(X)$ is the entropy (1) of the distribution $P$ that characterizes the source. This is obtained by showing that for $n \gg 1$ the typical messages $w$ which are most likely to be produced by the source all have approximately equal probabilities $2^{-nH(X)}$, so that their number is $\approx 2^{nH(X)}$. Since the number of bits necessary to enumerate such typical messages is $\approx nH(X)$, it follows that, by tolerating an asymptotically small error probability, one can use this same number of bits to encode all possible messages emitted by the source (e.g. associating with all the non-typical ones a fixed erasure symbol). This result can also be extended to a wide class of stationary correlated source models [37] and provides an operational interpretation of $H(X)$ as the information content per symbol of the average message emitted by the source.

Assume now that the messages produced by the source are transmitted through a classical noisy channel which maps the $n$-long input sequences $w = (x_1, \ldots, x_n)$ into $n$-long output messages $w' = (y_1, \ldots, y_n) \in \mathcal{Y}^n$ whose symbols $y_j \in \mathcal{Y}$ appear with output probability distribution $P' = \{p'_y\}$ where $p'_y = \sum_y p(y|x)p_x$. In this expression $p(y|x)$ is the single-letter conditional probability distribution which fully characterizes the process in the memoryless scenario. Memory channels in which the noise tampering with the transmission process acts on the letters of the message in a correlated fashion require instead the complete joint input–output probabilities for the messages to be assigned, see, e.g., [54].

Due to the stochastic nature of the transformation $w \rightarrow w'$, the distinguishability of the various input messages is not necessarily preserved. Still one can show that by properly selecting the codebook of messages $w$ (encoding), it is possible to convey (reliably) a definite amount of information to the output of the transmission line. Loosely speaking, the second Shannon coding theorem establishes that the amount of information per transmitted symbol that can be recovered from the channel outcomes is given by the Shannon capacity

$$C_{\text{Shan}} = \max_X I(X; Y),$$

(73)

where the maximum is taken over all possible distributions $P$ of the input $X$. In this expression $I(X; Y)$ is the mutual (or Shannon) information,

$$I(X; Y) = H(Y) - H(X|Y) = H(X) + H(Y) - H(X, Y),$$

(74)

with $H(Y)$ being the entropy of the output distribution $P' = \{p'_y\}$ which measures the total information content of the output of the channel, both useful—due to the uncertainty of the signal $X$—and harmful—due to the noise in the channel. Similarly, $H(X|Y) = -\sum_{x,y} p_{x,y} \log_2 p_{x,y}$ is the joint entropy of the pair of random variables $X, Y$ whose joint distribution is $p_{x,y} = p_x p(y|x)$. Finally $H(Y|X)$ is the conditional entropy (also called information loss) defined as

$$H(Y|X) = \sum_x p_x H(Y|X = x)$$

$$= -\sum_x p_x \sum_y p(y|x) \log_2 p(y|x)$$

(75)

$$= H(X, Y) - H(X),$$

(76)

which reflects the effect of the noise in the communication.

More precisely, the second Shannon theorem says that by using special (block) encoding of the messages at the input and decoding at the output of the channel it is possible to transmit $\approx 2^{nC_{\text{Shan}}} (n \rightarrow \infty)$ messages with asymptotically vanishing error, and it is impossible to safely transmit more whatever encoding and decoding are used. Thus, $C_{\text{Shan}}$ is the ultimate asymptotically errorless transmission rate for the channel $p(y|x)$. In practice, the optimal encoding and decoding doing the miraculous job of making the noisy channel partially invertible, i.e. transparent to some selected messages, are extremely difficult to approach, but $C_{\text{Shan}}$ gives the benchmark describing the exponential size of such an ideal codebook. Another insightful interpretation of the quantity $C_{\text{Shan}}$ is given by the ‘reverse Shannon theorem’ (formulated and proved surprisingly rather recently [18], under the influence of ideas from quantum information theory). This theorem implies that the ratio $\log_2 |\mathcal{X}|/C_{\text{Shan}}$ is equal to the number of copies of the noisy channel $p(y|x)$ needed to simulate (asymptotically) the ideal (noiseless) channel with the maximal capacity $\log_2 |\mathcal{X}|$.

The memoryless nature of the classical channel we are considering here is reflected by the additivity property

$$C_{\text{Shan}}^{(n)} = nC_{\text{Shan}},$$

(77)

where $C_{\text{Shan}}^{(n)}$ is the Shannon capacity of the $n$-block channel one obtains by treating $n$ instances of the source as a new source that emits $n$-long strings $w = (x_1, \cdots, x_n)$ as individual supersymbols. Accordingly in the definition of $C_{\text{Shan}}^{(n)}$ the maximum in (73) is taken over all input distributions $P^{(n)}$ on the space $\mathcal{X}^n$, including the correlated ones.
4.2. Quantifying information in a quantum world

Let us start with a quantum source of information that produces \(n\)-long sequences of quantum symbols (pure states) \(|\psi_j\rangle \in \mathcal{H}\), creating factorized states of the form \(|\Psi\rangle = |\psi_j\rangle \otimes \cdots \otimes |\psi_j\rangle \in \mathcal{H}^\otimes n\) (a practical example is a pulsed laser that emits a series of random optical signals characterized by different intensity values). As in the classical case we assume that each symbol constituting the sequence \(|\Psi\rangle\) is produced in a memoryless way, i.e., it is extracted from a given set of possible states independently with certain probability distribution \(\{p_j\}\). We ask what is the minimum number of qubits per symbol necessary to express the generic state produced by such a process. Schumacher and Josza [106] answered this question by showing that in the limit \((n \to \infty)\) this number coincides with the von Neumann entropy \(S(\rho)\) of the density operator \(\rho = \sum_j p_j |\psi_j\rangle \langle \psi_j|\) which represents the average state emitted by the source. More precisely, they proved that the high (asymptotically unit) probability, the sequences \(|\Psi\rangle\) span a subspace of \(\mathcal{H}^\otimes n\) of dimensionality \(\approx 2^{S(\rho)}\). Since the density matrix describing the average sequence \(|\Psi\rangle\) is \(\rho^\otimes n\), this can be rephrased by saying that \(nS(\rho)\) is the logarithmic size of a ‘quantum register’ in which a given quantum message \(\rho^\otimes n\) can be ‘packed’ optimally with negligible loss of information. Thus, in quantum information theory the logarithm of the dimensionality of the space of state vectors carrying information is the measure of information content of the system, and plays a role similar to the logarithm of size of the codebook for classical messages. Analogously to what the first Shannon coding theorem says for the Shannon entropy, the above result provides an operational characterization for the von Neumann entropy \(S(\rho)\), presenting it as a fundamental measure of the amount of quantum information that can be stored in the density matrix \(\rho\).

This result also admits another operational interpretation. Viewing the quantum source defined above as a special instance of a c–q channel (40) that encodes the classical variable \(x\), the result of [106] indirectly quantifies the amount of quantum resources (qubits) that are necessary to carry out such an encoding. The generalization of this result to arbitrary c–q channels is the subject of the next section.

4.3. The classical capacity of a quantum channel: part I

In section 2.8 we have seen that the simplest quantum model of a communication line is the c–q channel (40) where some classical data \(x \in X\) are mapped into a fixed family of output quantum states \(\{\rho_x\}\) in the receiver space \(\mathcal{H}\). If the letters of the message \(w = (x_1, \ldots, x_n)\) are transmitted independently of each other (no memory) then at the output of the composite channel one has the separable state \(\rho_{x_1} \otimes \cdots \otimes \rho_{x_n}\) in the space \(\mathcal{H}^\otimes n\). Decoding at the output requires a quantum measurement in \(\mathcal{H}^\otimes n\), the outcome of which gives an estimate \(w'\) for \(w\). For each specific choice of such a measurement we can thus define a classical channel which takes the classical random variable \(w\) to its output counterpart \(w'\) (passing through the quantum stage \(w \rightarrow \rho_{x_1} \otimes \cdots \otimes \rho_{x_n}\)).

Its ability to transfer classical information can now be evaluated by the associated Shannon capacity defined as in section 4.1. Define \(C^{(n)}\) to be the maximum of this quantity obtained by optimizing it with respect to all possible measurements on \(\mathcal{H}^\otimes n\). It turns out that in general, due to existence of entangled measurements at the output, differently from (77), one may have \(C^{(n)} > nC^{(1)}\). In other words, for the c–q memoryless channels the transmitted classical information can be strictly superadditive. A proper definition of the capacity hence requires a regularization with respect to the block coding size, i.e.

\[
C_X = \lim_{n \to \infty} \frac{C^{(n)}}{n}. \tag{78}
\]

Remarkably, the quantity \(C_X\), defined rather implicitly by this equation, admits an explicit entropic expression. For a statistical ensemble consisting of the density matrices \(\{\rho_x\}\) with the probabilities \(\{p_x\}\) we define the \(\chi\)-information as

\[
\chi = \chi (\{p_x\}, \{\rho_x\}) = S \left( \sum_x p_x \rho_x \right) - \sum_x p_x S(\rho_x). \tag{79}
\]

This quantity is nonnegative due to the concavity of the von Neumann entropy (see also section 4.12). To a certain extent the \(\chi\)-information can be regarded as a quantum analog of the Shannon information defined in equation (74) (see, however, the discussion in section 4.5). Both these quantities are the differences between the overall output entropy of the channel and a term which can be interpreted as the conditional entropy (loss). Furthermore, \(\chi\) provides a fundamental upper bound, first proved in [83], for the Shannon information \(I(X, Y)\) between the random variable \(X\) having the distribution \(\{p_x\}\) and the random variable \(Y\) describing the outcome of a measurement at the output of the channel aimed at recovering the value \(x\), namely

\[
I(X, Y) \leq \chi (\{p_x\}, \{\rho_x\}). \tag{80}
\]

The coding theorem established by Holevo [84, 85] and by Schumacher–Westmoreland [154] shows that there exist block coding strategies which, in the limit of large \(n\), saturate the bound imposed by equation (80). Thus, the abstractly defined capacity (78) acquires the following compact ‘one-letter’ expression

\[
C_X = \max_{\{p_x\}} \chi (\{p_x\}, \{\rho_x\}), \tag{81}
\]

(the maximization is performed over probabilities \(\{p_x\}\) while keeping \(\{\rho_x\}\) fixed). This relation can be regarded as the ‘classical–quantum’ analog of the second Shannon coding theorem for the noisy channel. For later modifications of its proof, see [77, 126, 136, 137, 188].

An apparent but important consequence of the bound (80) is the inequality

\[
C_X \leq \log_2 \dim \mathcal{H}, \tag{82}
\]

in which equality is attained if the quantum source operates with orthogonal states \(\rho_x\). Thus, the fact that the space \(\mathcal{H}\) contains infinitely many state vectors does not allow one to
increase the classical capacity above the ultimate information resource of the quantum system; increasing the number of signal vectors forces them to become nonorthogonal and hence less and less distinguishable. This is in line with the observation in the previous section where it was shown that the logarithm of dimensionality of the Hilbert space of a quantum system determines the ultimate bound on the amount of quantum information one can store in it.

**Example 1.** Consider the binary input signal $x = \pm 1$, and let $\rho_{\pm}$ be the coherent state of a monochromatic laser beam with the complex amplitude $\pm z$. This defines a c–q channel with two pure nonorthogonal states [90] whose classical capacity (81) can be computed as

$$C^c = h_2 \left(1 + \frac{1}{2}\right),$$

(83)

where $\epsilon = \langle z | z \rangle = \exp(-2|z|^2)$ is the overlap between the two coherent states, while $h_2(\rho) = -\rho \log_2 \rho - (1 - \rho) \log_2(1 - \rho)$ is the Shannon binary entropy. Also the quantity $C^{(1)}$ can be explicitly computed [19] yielding

$$C^{(1)} = 1 - h_2 \left(1 + \frac{1}{2}\right).$$

(84)

For this special case one can then easily verify that $C^c / C^{(1)} > 1$, the ratio tending to $\infty$ in the limit of a weak signal, as $\epsilon \rightarrow 1$, i.e. $z \rightarrow 0$.

**Example 2.** Consider the c–q channel in which the classical alphabet $\mathcal{X}$ is composed of three symbols (say $x = 0, \pm$), which are mapped into three equiangular pure states $|\psi_x\rangle$ of a qubit system,

$$|\psi_0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |\psi_{\pm}\rangle = \begin{bmatrix} -1/2 \\ \pm \sqrt{3}/2 \end{bmatrix}. $$

(85)

In this case of equal probabilities, $p_x = 1/3$, the average density matrix coincides with the chaotic state (i.e. $\rho = \sum_x |\psi_x\rangle \langle \psi_x|/3 = 1/2$). The classical capacity (81) of the channel is hence $C^c = S(1/2) = 1$ which saturates the bound (82) in spite of the fact that the states are not orthogonal. Also for this channel the gap between $C^c$ and $C^{(1)}$ can be shown. One has $C^{(1)} \approx 0.645$, the value attained for the distribution $p_+ = p_- = 1/2$, $p_0 = 0$ and for the measurement corresponding to the orthonormal basis, optimal for discrimination between the two equiprobable states $|\psi_{\pm}\rangle$, $x = \pm$ [150, 164]. However, the optimal encoding and decoding for this case are unknown, as for most of the Shannon theory.

More generally, let $\{|\psi_x\rangle; x \in \mathcal{X}\}$ be an overcomplete system (44) in a $d$-dimensional space (the above example being just a special case with $|\psi_{\pm}\rangle = \sqrt{2}/3 |\psi_x\rangle$). Then the overcompleteness relation can be written as $\sum_x p_x \rho_x = 1/d$, where $p_+ = |\psi_{+}\rangle \langle \psi_{+}|/d$ and $p_- = |\psi_{-}\rangle \langle \psi_{-}|/\langle \psi_{+}\rangle$. This implies that the c–q channel $x \rightarrow \rho_x$ has the capacity $C^c = S(1/d) = \log_2 d$. Furthermore, since the inequality in (80) is strict unless the operators $\rho_x$ commute [83], one can conclude that $C^{(1)} < C^c = \log_2 d$ unless $\{|\psi_x\rangle\}$ is an orthonormal basis.

**Example 3.** Consider the c–q channel model introduced in section 3.1 where a continuous alphabet $\mathcal{C}$ is encoded into quantum states via the mapping $\mu \rightarrow \rho_\mu$ with $\rho_\mu$ given by (49). If one tries to compute the classical capacity of such a channel by a continuous analog of formula (81) where the probabilities $p_x$ are replaced by probability distributions $p(\mu)$ on the complex plane $\mathcal{C}$, i.e.

$$C^c = \max_{p(\mu)} \left\{ S \left( \int p(\mu) \rho_\mu d^2 \mu \right) - \int p(\mu) S(\rho_\mu) d^2 \mu \right\},$$

(86)

one gets an infinite value as is to be expected for a channel with an infinite input alphabet. To obtain a reasonable finite result one should introduce a constraint onto possible input distributions $p(\mu)$ restricting the ‘energy’ that one can actually add in the fiber, namely

$$\int |\mu|^2 p(\mu) d^2 \mu \leq E.$$  

(87)

With this constraint, the maximum (86) can be evaluated by taking into account two facts: first, the states (49) are all unitarily equivalent and have the same entropy $S(\rho_\mu) = S(\rho_0) = g(N) \equiv (N + 1) \log_2(N + 1) - N \log_2 N$, see, e.g., [87]; second, the constraint (87) implies

$$\operatorname{Tr} \tilde{\rho} a^\dagger a \leq N + E,$$

(89)

where $\tilde{\rho} = \int p(\mu) \rho_\mu d^2 \mu$ is the average quantum state transferred in the communication. Consequently, by the maximal entropy principle, the entropy $S(\tilde{\rho})$ is maximized for the Gaussian distribution $p(\mu) = (1/\pi E) \exp(-|\mu|^2/E)$ giving the value $g(N + E)$, whence the channel capacity is

$$C^c = g(N + E) - g(N).$$

(90)

In the classical limit of large energies (i.e. $N \rightarrow \infty$, $E/N \rightarrow \text{const}$) this turns into $C^c = \log_2(1 + E/N)$, which can be regarded as a generalization of the famous Shannon’s formula [160] $C_{\text{Shannon}} = \frac{1}{2}\log_2(1 + E/N)$, for the capacity of memoryless channel with additive Gaussian white noise of power $N$ (the factor 1/2 being absent in the quantum case because one degree of freedom amounts to the two independent identically distributed real amplitudes). The capacity formula can be generalized to many-mode (broadband) c–q Gaussian channels [33, 86] also in the presence of squeezed seed states [95].

In the Shannon theory one considers memory channels with stationary Gaussian noise by making spectral decomposition of the time series in question. It turns out that the modes corresponding to different frequencies can be considered asymptotically independent if the observation time is very large. Then effectively one has a set of parallel independent channels which can be approached as a kind of ‘memoryless’ composite channel [37]. A similar reduction to the memoryless channel in the frequency domain can be elaborated for the c–q channels with additive stationary quantum Gaussian noise [86].
4.4. The classical capacity of a quantum channel: part II

In the previous section we focused on the basic case of quantum channels that can be used to transfer classical messages, the c-q channels. They are formally constructed by fixing a set of possible quantum letters $\rho_s$ in the space $\mathcal{H}$ organized in the $n$-long sequences $\rho_s \otimes \cdots \otimes \rho_s$ that form the codewords in which the classical information can be encoded.

More generally, assuming that the sender of the classical information is allowed to use as input any (possibly entangled [15, 16]) joint density matrix $\rho^{(w)} \in \mathcal{S}(\mathcal{H}^{\otimes n})$ of $n$ quantum information carriers. Here $w$ denotes a classical message encoded into the state $\rho^{(w)}$. We assume that there are $N$ different messages to be transmitted, hence $w$ can be simply the number of messages, $w = 1, \ldots, N$. Assume also that during the transmission stage each component of the codeword $\rho^{(w)}$ is individually affected by the same noisy channel (memoryless regime), producing outputs of the form

$$\rho^{(w)} \rightarrow \Phi^{\otimes n}[\rho^{(w)}],$$

where $\Phi^{\otimes n}$ is the $n$-fold tensor product defined similarly to (34) of the quantum channel $\Phi$ which gives the statistical description (14) of the interaction of a single carrier with the environment (i.e. noise). Note that for each given collection of input density matrices $\{\rho^{(w)}; w = 1, \ldots, N\}$, equation (91) defines an effective c-q channel whose quantum letters are the density matrices $\Phi^{\otimes n}[\rho^{(w)}]$.

The associated classical capacity of such c-q channel is then computed along the lines of section 4.3,

$$C_{\chi}(\Phi^{\otimes n}) = \max_{\{p_w, \{\rho^{(w)}\}\}} \chi \left( \{p_w\}, \{\Phi^{\otimes n}[\rho^{(w)}]\} \right),$$

where now, given the freedom the sender has in selecting the input density matrices of the channel, the maximum is performed over the set of statistical ensembles formed by the probabilities $\{p_w\}$ and the states $\{\rho^{(w)}\}$. Loosely speaking, equation (92) gives the ultimate rate of classical bits that one can transmit when using quantum block-letters of size $n$, hence corresponding to the rate $C_{\chi}(\Phi^{\otimes n})/n$ bits per individual use of the channel $\Phi$. Therefore, the ultimate (asymptotically achievable) rate is given by the expression

$$C(\Phi) = \lim_{n \rightarrow \infty} \frac{1}{n} C_{\chi}(\Phi^{\otimes n}).$$

It should be stressed that the regularizations involved in equations (93) and (78) have different origins. While taking the limit in equation (78) is required by superadditivity due to the possible use of entanglement at the decoding stage of the communication process (see the previous section), the limit in equation (93) is required by the possible superadditivity of the function $C_{\chi}(\Phi^{\otimes n})$ as a consequence of using entangled quantum block-letter $\rho^{(w)}$ at the encoding stage.

In the case where the property of additivity holds, i.e.

$$C_{\chi}(\Phi^{\otimes n}) = n C_{\chi}(\Phi),$$

regularization (93) is not needed and the capacity admits a simple single letter expression $C(\Phi) = C_{\chi}(\Phi)$, where

$$C_{\chi}(\Phi) = \max_{p_s, \rho_s} \left\{ S \left( \sum_s p_s \rho_s \rho_s \right) - \sum_s p_s S(\rho_s) \right\},$$

and the maximization is over all finite ensembles of states $\rho_s$ taken with probabilities $p_s$. The additivity (94) means that using entangled states at the input of the channel $\Phi$ does not increase the quantity of transmitted classical information. The validity of the property (94) was established for a number of channels, including the unital qubit channels [109], the depolarizing channel [110], the erasure channel [13], the purely lossy bosonic channel [58] and the whole class of entanglement-breaking channels [163]. In all of these cases except for the last one the analytical solution of the maximization problem for $C_{\chi}(\Phi)$ is possible.

Example. For the depolarizing channel of equation (21), the additivity (94) combined with the high symmetry of the map $\Phi$ allows one to find its classical capacity [110],

$$C(\Phi) = C_{\chi}(\Phi) = \log_2 d + \left( 1 - p \frac{d - 1}{d} \right) \times \log_2 \left( 1 - p \frac{d - 1}{d} \right) + \frac{d - 1}{d} \log_2 \frac{p}{d},$$

with the maximum attained on the ensemble of $d$ equiprobable orthogonal pure states.

The need of the regularization limit in equation (93) for a generic channel has been debated at length, for a survey see, e.g., [92, 147]. The question as to whether there exist nonadditive quantum channels at all turned out to be extremely difficult and remained open for a rather long time. The additivity of $C_{\chi}$ plays an important role in quantum information and is linked to the additivity of other quantum entropic quantities [130, 167] including the minimal output entropy $\min_{\rho} S(\Phi[\rho])$. A significant step forward was made by Shor [167], who showed in particular that proving the additivity of the minimal output entropy for all pairs of channels, i.e.

$$\min_{\rho} S(\Phi_1 \otimes \Phi_2[\rho]) = \min_{\rho} S(\Phi_1[\rho]) + \min_{\rho} S(\Phi_2[\rho]), \quad \forall \Phi_1, \Phi_2,$$

is equivalent to proving a similar additivity for $C_{\chi}$, i.e.

$$C_{\chi}(\Phi_1 \otimes \Phi_2) = C_{\chi}(\Phi_1) + C_{\chi}(\Phi_2), \quad \forall \Phi_1, \Phi_2,$$

and hence it would imply (94). This in turn stimulated intensive research of several related quantities, which besides the von Neumann entropy [113], include quantum Rényi entropies see, e.g., [5, 6, 46, 61, 108, 111, 159]. The possibility of violating the additivity for these functionals was established in several cases [38, 78, 79, 187, 189], but the violations were not strong enough to imply superadditivity of $C_{\chi}$. The problem has been settled by Hastings [74] who proved that channels which violate the additivity (97) and hence require the regularization in equation (93) do exist at least in very high dimensions (see [52] for the actual dimensionality estimates), among mixtures of unitary channels of the form $\Phi[\rho] = \sum_j \pi_j U_j \rho U_j^\dagger$ (here $\pi_j$ is a probability distribution while $U_j$ are unitary operators), but so far no concrete example has been found.

It is finally worth mentioning that transmission of classical information through quantum channels may display yet another form of superadditivity of a higher complexity level.
as compared with the one implied by the regularizations of equations (78) and (93). Namely, consider two different quantum channels $\Phi_1$ and $\Phi_2$ with the classical capacities $C(\Phi_1)$ and $C(\Phi_2)$ defined as in (93). The problem is whether the capacity $C(\Phi_1 \otimes \Phi_2)$ of the tensor product channel $\Phi_1 \otimes \Phi_2$ can be strictly greater than the sum of the two individual capacities, i.e.

$$C(\Phi_1 \otimes \Phi_2) > C(\Phi_1) + C(\Phi_2).$$

(99)

It should be stressed that due to the regularization present in (93) the superadditivity of the functional $C_{\chi}(\cdot)$ proved in [74] is not sufficient to answer this question. In fact, the strict inequality in (99) would be similar to the superactivation of the quantum capacities [171], see section 4.11. Proving $C(\Phi_1 \otimes \Phi_2) \geq C(\Phi_1) + C(\Phi_2)$ is trivial: this follows simply from the possibility of operating the two channels independently, i.e. using codewords that factorize in the partition $\Phi_1 \otimes \Phi_2$. However, the possibility of having a strict inequality here stems from the freedom of introducing the quantum correlations at the input of such a partition.

4.5. Entropy exchange and quantum mutual information

In the classical case correlation between two random variables $X, Y$ (which, in particular, may describe input and output of a classical channel) can be measured by the mutual information $I(X; Y)$ introduced in equation (74). A quantum analog of this quantity has been identified in section 4.3 as the $\chi$-information of equation (79). The quantity $\chi$, however, does not account for all the correlations which can be established between the input and the output of a quantum channel. To characterize them one needs to look for other quantum generalizations of $I(X; Y)$ (the possibility of multiple quantum generalizations should not be surprising: ultimately this is a consequence of the fact that joint distribution of quantum observables exists only in the very special case when they commute).

One way to proceed is to exploit the following purification trick [153, 154]. Consider a quantum channel $\Phi = \Phi_{A \rightarrow B}$ that maps the input states $\rho = \rho_A \in \mathcal{S}(\mathcal{H}_A)$ into the outputs $\rho_B = \Phi[\rho]$. Let us introduce a reference system $\mathcal{H}_R \simeq \mathcal{H}_A$ and purify the input to $|\psi_{AR}\rangle|\psi_{AR}\rangle$ in the space $\mathcal{H}_A \otimes \mathcal{H}_R$. By equation (5) the state $\rho_R$ of the reference system has the same spectrum as $\rho$ and hence $S(\rho) = S(\rho_R)$. In what follows, for the sake of simplicity, we shall abbreviate the notation for the entropies of partial states by omitting the symbol of the density operator $\rho$, so, for example, the last equality will be written as $S(A) = S(R)$. Now let us focus on the entanglement transmission scenario (29) where the state $|\psi_{AR}\rangle|\psi_{AR}\rangle$ is transmitted via the channel $\Phi_{A \rightarrow B} \otimes \text{Id}_R$, producing the output state

$$\rho_{BR} = (\Phi_{A \rightarrow B} \otimes \text{Id}_R)[|\psi_{AR}\rangle|\psi_{AR}\rangle].$$

(100)

Consider the quantity

$$I(\rho, \Phi) = S(R) + S(B) - S(BR),$$

(101)

which is nonnegative and vanishes if and only if $\rho_{BR} = \rho_B \otimes \rho_R$, see section 4.12. Since the reference system $R$ is a copy of the input $A$ which remains intact over the course of the transmission, one can interpret $I(\rho, \Phi)$ as another substitute for the Shannon mutual information between the input and the output of the quantum channel. It is called the quantum mutual information [2, 123]. It is formed by combining the three entropies: $S(R) = S(\rho)$—the input entropy, $S(BR) = S(\Phi[\rho])$—the output entropy, and the joint entropy $S(BR)$ which deserves a closer look in the isometric representation (37) of the channel $\Phi$.

A useful expression for the quantity (101) is obtained by introducing the channel environment $E$ and the isometry $V : \mathcal{H}_A \rightarrow \mathcal{H}_B \otimes \mathcal{H}_E$ which represent the transformation $\Phi$ as detailed in section 2.3, see figure 3. At the end of the transmission we have the tripartite system $BR E$ characterized by the space $\mathcal{H}_B \otimes \mathcal{H}_E \otimes \mathcal{H}_R$. The total output state $\rho_{BER}$ in $\mathcal{H}_B \otimes \mathcal{H}_E \otimes \mathcal{H}_R$ is pure and is described by the vector $|\psi_{BER}\rangle = (V \otimes I_R)|\psi_{AR}\rangle$. Looking at the split $BR |E$ we have a pure state bipartite system with partial states $\rho_{BR} = \text{Tr}_{E} \rho_{BER}$. $\rho_E = \text{Tr}_{BR} \rho_{BER}$ whose entropies $S(BR)$ and $S(E)$ are equal, see equation (5). This implies

$$S(BR) = S(E) = S(\rho, \Phi),$$

(102)

where the last quantity is called the entropy exchange [123, 153, 154], as it measures the entropy change in the environment (recall that in the isometric representation the initial environment state is assumed pure). Note also that by construction $\rho_E$ is the state at the output of the complementary channel $\Phi$. Therefore by equation (38),

$$S(\rho, \Phi) = S(\Phi[\rho]) = S \left( \left[ \text{Tr}_{R} \rho V_a V_a^\dagger \right]_{a=1,\ldots,d_{C}} \right).$$

(103)

The quantum mutual information can then be expressed as

$$I(\rho, \Phi) = S(\rho) + S(\Phi[\rho]) - S(\rho, \Phi).$$

(104)
**Example.** Consider the depolarizing channel (21) and the chaotic state $I/d$. By using the Kraus decomposition (22) we find that the state $\Phi[I/d]$ has eigenvalues $0$ and $1 - p(d^2 - 1)/d^2$ of multiplicity $1$, and $d^2 - 1$ (the zero appearing because of the linear dependence of the Kraus operators). Hence,

$$S(I/d, \Phi) = - \left(1 - p \frac{d^2 - 1}{d^2}\right) \log_2 \left(1 - p \frac{d^2 - 1}{d^2}\right).$$

Combining with the input and the output entropies $S(I/d) = S(\Phi[I/d]) = \log_2 d$, we obtain

$$I(I/d, \Phi) = \log_2 d^2 + \left(1 - p \frac{d^2 - 1}{d^2}\right) \times \log_2 \left(1 - p \frac{d^2 - 1}{d^2}\right) + p \frac{d^2 - 1}{d^2} \log_2 p.$$  \hspace{1cm} (105)

Up to now we have encountered three entropic quantities: the input entropy $S(\rho)$, the output entropy $S(\Phi[\rho])$ and the entropy exchange $S(\rho, \Phi)$. By making different bipartite splits of the system $BER$, we obtain similarly to (102) the identities

$$S(BE) = S(R) = S(\rho),$$

$$S(RE) = S(B) \equiv S(\Phi[\rho]),$$

and another two information quantities: the loss, i.e. the quantum mutual information between the input and the environment:

$$L(\rho, \Phi) \equiv S(R) + S(E) - S(RE) = S(\rho) + S(\rho, \Phi) - S(\Phi[\rho]),$$

and the noise, i.e. the quantum mutual information between the output and the environment:

$$N(\rho, \Phi) \equiv S(E) + S(B) - S(EB) = S(\rho, \Phi) + S(\Phi[\rho]) - S(\rho).$$

The nonnegativity of the quantities $I(\rho, \Phi), L(\rho, \Phi)$ and $N(\rho, \Phi)$ along with (107), (108) implies that the basic entropies $S(\rho), S(\Phi[\rho])$ and $S(\rho, \Phi)$ satisfy all the triangle inequalities, namely

$$|S(\Phi[\rho]) - S(\rho, \Phi)| \leq S(\rho),$$

and two other inequalities obtained by cyclic permutations.

As in the case of the Shannon mutual information (74), the quantities $I(\rho, \Phi), L(\rho, \Phi)$ and $N(\rho, \Phi)$ can also be expressed in terms of conditional (quantum) entropies. For instance we have

$$I(\rho, \Phi) = S(B) - S(B|R),$$

$$S(B[R]) \equiv S(BR) - S(R) = S(\rho, \Phi) - S(\rho),$$

with $S(B|R)$ being the quantum conditional entropy of $B$ given $R$. Notice however, that while in the classical case the conditional entropy is always nonnegative, in the quantum scenario this is no longer valid as the quantum entropy is not necessarily monotone with respect to enlargement of the system, i.e. $S(R) \nless S(BR)$ (an extreme example of this property is obtained when $BR$ is a pure entangled system: in this case $S(BR) = 0$ while $S(R) > 0$—this cannot happen in the classical statistics where the partial state of a pure state can never be mixed, see also section 2.1). An operational interpretation of the possible negativity of the conditional entropy in terms of ‘quantum state merging’ is given in [100]. Quite remarkably, the monotonicity of the conditional entropy still holds: for any composite system $ABC$ one has

$$S(A|BC) \leq S(A|B),$$

and it is this property which makes the quantum conditional entropy useful. When written in terms of (unconditional) quantum entropy, the inequality (113) amounts to the fundamental property of strong subadditivity:

$$S(ABC) + S(B) \leq S(AB) + S(BC).$$

The proof of this very useful inequality was first given by Lieb and Ruskai [121] and remains rather involved in the quantum case even after subsequent simplifications, see, e.g., [148].

Due to the strong subadditivity (114), the quantum mutual information enjoys important properties similar to that of the Shannon information in the classical case [2, 135].

Specifically, given two quantum channels $\Phi_1$ and $\Phi_2$, one has

(i) subadditivity:

$$I(\rho_{12}, \Phi_1 \otimes \Phi_2) \leq I(\rho_1, \Phi_1) + I(\rho_2, \Phi_2);$$

(ii) data-processing inequalities:

$$I(\rho, \Phi_2 \circ \Phi_1) \leq \min[I(\rho, \Phi_1), I(\Phi_1(\rho), \Phi_2)].$$

4.6. Entanglement as an information resource

In the previous sections we have seen that entangled decodings can enhance transmission of classical information through a quantum channel. An even more impressive gain is achieved when entanglement between the input and the output of the channel is available. The classical capacity of a memoryless communication line defined by a quantum channel $\Phi$ can be substantially increased by using this additional resource in spite of the fact that entanglement alone cannot be used to transmit information. This fundamental observation was first made by Bennett and Wiesner who introduced the notion of superdense coding [19], see also [7, 26, 146] for generalizations to continuous variables and [120, 131, 133, 141] for experimental tests. Here as in some other cases entanglement plays a role of a ‘catalyzer’, disclosing latent information resources of a quantum system.

The scenario of entanglement-assisted communication assumes that prior to information transmission parts of an entangled state $\rho_{AB}$ are distributed between sender $A$ and receiver $B$. Then sender $A$ encodes the classical messages $w$ into different operations $\mathcal{E}_w$ on his part of the entangled state, and the result of these operations is sent to $B$ via the quantum channel $\Phi$. Thus at the end of the transmission the state $(\Phi \circ \mathcal{E}_w \otimes \text{Id}_B)|\rho_{AB}\rangle$ is available to the receiver which extracts the classical information by making quantum
measurement on this state. Optimizing the transmission rate with respect to the entangled states \(\rho_{AB}\), (block) encodings of \(A\) and measurements (decodings) of \(B\) gives the classical entanglement-assisted capacity. The corresponding coding theorem of Bennett, Shor, Smolin and Thapliyal provides a simple formula (the proof of which is far from simple [17, 18, 89]) giving an operational characterization for the quantum mutual information (104):

\[
C_{ea}(\Phi) = \max_{\rho} I(\rho, \Phi),
\]

(117)

where the maximum is taken over all possible input states \(\rho\) of a single channel use. (Equation (117) refers to the case of unlimited shared entanglement. For an analysis of the case in which only limited resources are available see [21, 45, 167].) Remarkably, unlike equation (93), the capacity formula (117) does not contain the limit \(n \to \infty\) because \(C_{ea}(\Phi^{\otimes n}) = n C_{ea}(\Phi)\) due to subadditivity (115) of the quantum mutual information. Also by construction it follows that entanglement-assisted capacity \(C_{ea}(\Phi)\) is always greater than or equal to its unassisted counterpart \(C(\Phi)\).

If \(\Phi = \text{Id}\) is the noiseless channel then this scenario coincides with the original superdense coding protocol [19] for which \(C_{ea}(\text{Id}) = 2 \log_2 d = 2C(\text{Id})\), so the capacity gain is equal to 2. A similar doubling of the capacity holds for the quantum erasure channel. However, in general, the more noisy the channel is, the greater is the gain, and in the limit of very noisy channels the gain can be arbitrarily large. For example, in the case of the depolarizing channel (21), the maximum in (117) is attained on the chaotic state so that \(C_{ea}(\Phi)\) is given by formula (106). Comparing this with the quantity \(C(\Phi)\) given by formula (96), one sees that the gain \(C_{ea}(\Phi)/C(\Phi) \to d+1\) in the limit of large noise \(p \to 1\).

Interestingly, one can have \(C_{ea}(\Phi) > C(\Phi)\) for entanglement-breaking channel, for example, this holds for the depolarizing channel with \(p \geq d/(d+1)\). The explanation given in [17] is that unassisted classical data transmission through an entanglement-breaking channel involves communication cost which is always greater than the difference between \(C_{ea}(\Phi)\) and \(C(\Phi)\).

A protocol in a sense dual to superdense coding is quantum teleportation which was first introduced in [11]. Quantum information theory predicts the possibility of a nontrivial way of transmitting arbitrary quantum state \(\rho\), when the state carrier is not transferred physically but only some classical information is transmitted through a classical communication line, see also [25, 132, 179] for the generalization to continuous variables and [10, 20] for the first experimental test. The necessary additional resource here is a shared entangled state between the sender and the receiver of the classical information (it is impossible to reduce the transmission of an arbitrary quantum state solely to sending the classical information: since the classical information can be copied, it would mean the possibility of cloning the quantum information [193]). The effective maps \(\Phi\) resulting from such protocols are called quantum teleportation channels [23, 99, 143, 192]. They represent a proper subset of the class of quantum communication systems which can be fully specified by assigning a joint initial state \(\rho_{AB}\), characterizing the shared entanglement between the sender and the receiver, and the local operations the two parties are supposed to perform on it.

### 4.7. Coherent information and perfect error correction

An important part of the quantum mutual information \(I(\rho, \Phi)\) is the coherent information [153]

\[
I_c(\rho, \Phi) = S(\rho|\Phi) - S(\rho),
\]

(118)

where \(S(\rho|\Phi)\) is the von Neumann entropy of \(\rho\) given \(\Phi\). This quantity is closely related to the quantum capacity of the channel \(\Phi\) which will be considered in the next section.

The coherent information does not share some ‘natural’ properties of quantum mutual information such as subadditivity and the second data-processing inequality. Moreover, similarly to its classical analog \(H(Y) = H(XY) = -H(X|Y)\) it can be negative. However \(I_c(\rho, \Phi)\) satisfies the first data-processing inequality:

\[
I_c(\rho, \Phi) \leq I_c(\rho, \Phi_1),
\]

(119)

which follows from the relation \(I_c(\rho, \Phi) = I(\rho, \Phi) - S(\rho)\) and from the corresponding property of quantum mutual information \(I(\rho, \Phi)\).

There is a close connection between perfect transmission of quantum information, error correction and certain property of coherent information. The channel \(\Phi\) is perfectly reversible on the state \(\rho = \rho_A\) if there exists a recovery channel \(\mathcal{D}\) from \(B\) to \(A\), such that

\[
(D \otimes \Phi \otimes \text{Id}_R)[\rho_{AB}] = \rho_{AR},
\]

(120)

where \(\rho_{AR}\) is a purification of the state \(\rho_A\) with the reference system \(R\).

One can show [9] that the following statements are equivalent:

(i) the channel \(\Phi\) is perfectly reversible on the state \(\rho\);

(ii) \(L(\rho, \Phi) = 0\), that is \(\rho_{RE} = \rho_R \otimes \rho_E\);

(iii) \(I_c(\rho, \Phi) = S(\rho)\).

Condition (ii) means that information does not leak into the environment, i.e. the channel is ‘secret’ or ‘private’. Thus the perfect reversibility of the channel is equivalent to its privacy. Condition (iii) means that under private transmission of the state \(\rho\) by the channel \(\Phi\), the coherent information \(I_c(\rho, \Phi)\) should attain its maximal value \(S(\rho)\). The equivalence of (ii) and (iii) is obvious since \(S(\rho) - I_c(\rho, \Phi) = S(\rho) + S(\rho; \Phi) - S(\rho|\Phi)\) is always non-negative because \(S(\rho)\) is always non-negative.

In particular, choosing the chaotic state \(\rho = I_A/d_A\), we get the condition \(\log_2 d_A = \log_2 d_A\), which shows that the coherent information should be related to the quantum capacity of the channel \(\Phi\), which characterizes the maximal dimension of the perfectly transmittable states. (In fact, such a relation is valid for asymptotically perfect transmission through the block channel \(\Phi^\otimes n\) when \(n \to \infty\), as we shall see in the following...
section.) In support of this argument we note here that the perfect reversibility property of $\Phi$ on $\rho$ can also be stated by saying that there exists a recovery channel $D$ such that

$$D \circ \Phi[\rho'] = \rho',$$

(121)

for all states $\rho'$ with $\text{supp}[\rho'] \subseteq \mathcal{L} \equiv \text{supp}[\rho]$, with $\text{supp}[\rho]$ being the support of the state $\rho$. We can express the same property by saying that $\Phi$ is perfectly reversible on the subspace $\mathcal{L}$. In other words, the subspace $\mathcal{L}$ is a quantum code correcting the error described by the noisy channel $\Phi$ (and hence, all the related errors $\rho \rightarrow V_j \rho V_j^\dagger$, where $\Phi[\rho] = \sum_j V_j \rho V_j^\dagger$) [115].

On the other hand, note that by formula (102) $S(\rho; \Phi) = S(\Phi[\rho])$, where $\Phi$ is the complementary channel, hence

$$I_c(\rho, \Phi) = I(\rho, \Phi) = S(\Phi[\rho]) - S(\Phi[\rho]) = -I_c(\rho, \Phi).$$

(122)

The next statement which generalizes the observation at the end of section 2.7 gives a characterization in terms of a complementary channel which underlies the coding theorem for the secret classical capacity of the channel in section 4.9.

The following conditions are equivalent:

(i) the channel $\Phi$ is perfectly reversible on the subspace $\mathcal{L}$; (i) the complementary channel (38) is completely depolaring on $\mathcal{L}$, i.e.

$$\Phi[\rho'] = \rho_E,$$

(123)

for any state $\rho'$ with $\text{supp}[\rho'] \subseteq \mathcal{L}$, where $\rho_E$ is a fixed state.

4.8. The quantum capacity

The transformation $\rho \rightarrow \Phi[\rho]$ of quantum states can be regarded as the transfer of quantum information. The discovery of the quantum error-correcting codes [162, 172] is related to the question of asymptotically (as $n \rightarrow \infty$) error-free transmission of quantum information by the channel $\Phi^{\otimes n}$. The quantum capacity $Q(\Phi)$ is defined as the maximum amount of quantum information per one use of the channel which can be transmitted with asymptotically vanishing error [8, 16, 44, 118, 125]. It is related to the dimensionality of the subspace of state vectors in the input space ($\approx 2^n|\Phi(\mathcal{H})|$) that are transmitted asymptotically error-free. For the quantum capacity there is an expression in terms of coherent information (118), namely

$$Q(\Phi) = \lim_{n \rightarrow \infty} \frac{1}{n} \max_{\rho^{(n)}} I_c(\rho^{(n)}, \Phi^{\otimes n}),$$

(124)

the maximum being performed over all input states $\rho^{(n)}$ of $n$ successive channel uses. The relation between $Q(\Phi)$ and the coherent information of the channel was conjectured in [125] and made more precise in [9, 165], while the ultimate proof of equation (124) was given by Devetak [44] exploiting the fact that the quantum capacity of a channel is closely related to its cryptographic characteristics, such as the capacity for the secret transmission of classical information and the rate of the random key distribution. More specifically, as discussed in detail in the following section, a deep analogy with a wiretap channel [190] was used, the role of the eavesdropper in the quantum case played by the environment of the system.

For the ideal channel $\text{Id}$ one easily verifies that $Q(\text{Id}) = \log_2 d$. Analytical expression for the capacity of quantum depolarizing channel (21) is still unknown in the general case, although there are fairly close lower [14, 48, 72] and upper [144, 145, 170, 180] bounds for it. A major difficulty in evaluating $Q(\Phi)$ lies in the nonadditivity of the quantity $I_c(\Phi^{(n)}), \Phi^{\otimes n})$ [48]. However, there is an important class of degradable channels [39, 47] for which the additivity holds, so one can replace equation (124) with the convenient ‘one-letter’ formula

$$Q(\Phi) = Q_1(\Phi) \equiv \max_{\rho} I_c(\rho, \Phi),$$

(125)

where now the maximization is taken over the input states $\rho$ for single channel use. The channel $\Phi$ is called degradable if there exists quantum channel $\Upsilon$ such that $\Phi = \Upsilon \circ \Phi$ [47]. In practical terms, this relation expresses the fact that the complementary channel $\Phi$ is ‘more noisy’ than $\Phi$ (it can be obtained from the latter by ‘adding’ the extra noise $\Upsilon$). Notable examples of degradable channels are the dephasing channels [47, 112] and the amplitude damping channel [56] for which the maximization in (125) can be explicitly performed.

Another class of interest is the so-called anti-degradable channels: $\Phi$ is called anti-degradable [32, 56], if there exists a channel $\Upsilon'$, such that $\Phi = \Upsilon' \circ \Phi$. Apparently $\Phi$ is degradable if and only if $\Phi$ is anti-degradable. Any anti-degradable channel $\Phi$ has null quantum capacity, $Q(\Phi) = 0$. A formal proof of this fact will be given in section 4.12; there is however a simple heuristic argument based on the no-cloning theorem, see equation (13), which explains why $Q(\Phi)$ cannot be positive. Suppose that $Q(\Phi) > 0$ for the anti-degradable channel $\Phi$. This implies that, via encoding and decoding protocols, the sender will be able to transfer to the receiver at the output of the channel $\Phi$ an arbitrary (unknown) pure quantum state $|\psi\rangle$. Since $\Phi$ is anti-degradable, this implies that the same protocol will also allow one to recover the same quantum message at the output of the complementary channel $\Phi$ (one can reconstruct the associated output of $\Phi$ by applying the channel $\Upsilon'$). Now the contradiction arises by observing that in the isometric representation introduced in section 4.5, $\Phi$ and $\Phi$ describe the two reduced quantum information flows that enter, respectively, to the receiver and to the channel environment. Two independent observers collecting those data will thus be able to get a copy of the same state $|\phi\rangle$, realizing de facto a cloning machine, which is impossible. Notable examples of anti-degradable channels are provided by the entanglement-breaking channel (43) whose complementary counterparts are the dephasing channel (45) (the channel $\Upsilon'$ in this case can be taken as $\Upsilon'|\rho\rangle = \sum_{a} |\phi_a\rangle|e_a\rangle \langle e_a| \langle \phi_a|)$.

Example. The quantum erasure channel $\rho_{\Phi}$ introduced in equation (35) is degradable for $p \in [0, 1/2]$ and anti-degradable for $p \in [1/2, 1]$. Its quantum capacity is computed [13] as

$$Q(\Phi_{p}) = \begin{cases} (1 - 2p) \log_2 d, & p \in [0, 1/2]; \\ 0, & p \in [1/2, 1]. \end{cases}$$

(126)
Another class of channels which, as the anti-degradable ones, possesses null quantum capacity is given by the so-called PPT (partial positive transpose) or binding channels [102]. These maps include as a special case the entanglement-breaking channels of section 2.8, and are characterized by the property that their associated Choi–Jamiołkowski state is PPT, i.e. it remains positive under partial transposition in the reference space. States that possess this property might not be separable but their entanglement (called bound entanglement) is ‘weak’ as it does not allow for distillation [103, 104]. Binding maps (if not entanglement-breaking) allow for a certain amount of entanglement transfer: the latter however is always non-distillable and, even though it admits private communication between the sender and the receiver (see section 4.9), it cannot be used for the faithful transfer of quantum information (not even in the presence of a two-way classical communication side line) [102]. For a recent study of the connections between anti-degradable and PPT channels, as well as on the characterization of the set of maps that have null quantum capacity see [169].

4.9. Quantum wiretap channel

Here we review in brief the argument behind the derivation of equation (124). Consider the situation of classical information transmission in which there are three parties: sender A, receiver B and the eavesdropper E. A mathematical model of the quantum wiretap channel comprises three Hilbert spaces \( \mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_E \) and the isometric map \( V : \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_E \), which transforms the input state \( \rho_A \) into the state \( \rho_{BE} = \Phi_{AB\to BE}[\rho_A] = V \rho_A V^\dagger \) of the system \( BE \), with partial states \( \rho_B = \Phi [\rho_A] = Tr_E V \rho_A V^\dagger, \quad \rho_E = \tilde{\Phi} [\rho_A] = Tr_B V \rho_A V^\dagger. \) (127)

Note that this description is formally identical to that of the complementary channels in section 2.7, where \( E \) denoted the environment (in the wiretap model this is supposed to be completely under control of the eavesdropper).

Assume now that A sends the states \( \{\rho_A^x\} \) with probabilities \( \{p_x\} \); then the parties B and E receive, correspondingly, the states \( \{\rho_B^x\} \) and \( \{\rho_E^x\} \), and upper bounds for Shannon information they receive are the quantities \( \chi (\{p_x\}, \{\rho_B^x\}) \) and \( \chi (\{p_x\}, \{\rho_E^x\}) \), as stated in equation (80). In analogy with the classical wiretap channel [190] the ‘secrecy’ of the transmission can be characterized by the quantity \( \chi (\{p_x\}, \{\rho_B^x\}) - \chi (\{p_x\}, \{\rho_E^x\}) \) (here by secrecy we mean the amount of information which can be shared between A and B without informing E). In fact, the capacity for secret transmission of classical information is shown to be [44]

\[
C_p (\Phi_{A\to BE}) = \lim_{n \to \infty} \max_{\{\rho_B^{(0)} \}, \{\rho_E^{(0)} \}} \left[ \chi (\{p_x^{(n)}\}, \{\rho_B^{(n)}\}) - \chi (\{p_x^{(n)}\}, \{\rho_E^{(n)}\}) \right],
\]

where the maximum is taken over the families of states \( \Sigma^{(n)} = \{\rho_B^{(n)}\} \) in \( \mathcal{H}_B \) and the probability distributions \( \rho^{(n)} = \{\rho_B^{(n)}\} \) (we use the notation \( \rho_B^{(n)} = \Phi^{\otimes n}[\rho_A^{(n)}], \rho_E^{(n)} = \tilde{\Phi}^{\otimes n}[\rho_A^{(n)}] \)).

Assuming that the input states \( \rho_A^x \) are pure, and denoting by \( \rho = \rho_A = \sum_x p_x \rho_A^x \) the average state of the input ensemble, from (129) we obtain the key relation [155]

\[
I_c (\rho, \Phi) = S (\Phi [\rho]) - S (\tilde{\Phi} [\rho])
\]

\[
= \left[ S (\Phi [\rho]) - \sum_x p_x S (\Phi [\rho_A^x]) \right] - \left[ S (\tilde{\Phi} [\rho]) - \sum_x p_x S (\tilde{\Phi} [\rho_A^x]) \right]
\]

\[
= \chi (\{p_x\}, \{\rho_A^x\}) - \chi (\{p_x\}, \{\rho_A^x\}),
\]

(129) where we used equation (122) and the fact that the states \( \rho_B^x = V \rho_A^x V^\dagger \) are pure implying

\[
S (\Phi [\rho_A^x]) = S (\rho_A^x) = S (\tilde{\Phi} [\rho_A^x]).
\]

(130) for all \( x \). Identity (129) provides the fundamental connection between the quantum and the secret classical capacities which underlies the proof of equation (124) given in [44]. Here we only note that since in the computation of \( C_p (\Phi_{A\to BE}) \) one takes into account all ensembles (not just the pure ones for which equation (129) hold), we get the following inequality:

\[
C_p (\Phi_{A\to BE}) \geq Q (\Phi),
\]

(131) which in general can be strict—see, e.g., [97, 98]—with the notable exception of degradable channels for which

\[
C_p (\Phi_{A\to BE}) = Q (\Phi) = Q_1 (\Phi).
\]

(132)

To conclude this discussion of wiretap channels we briefly mention the vast field of quantum cryptography which constitutes a self-consistent portion of quantum information science (for reviews related to this subject we refer to [63, 183]).

4.10. Capacities for Gaussian channels

In this section we briefly discuss the classical and quantum capacities for Gaussian channels: in particular, we focus on the single-mode CPTP maps analyzed in section 3.5.

When speaking of the classical capacities for continuous-variable systems, to obtain reasonable finite quantities, one should introduce an energy constraint onto input ensembles similarly to what was done for the c–q Gaussian channel in example 3 of section 4.3. On the other hand, for the quantum capacity, even though constraining the input may be reasonable from a practical point of view, this is not strictly necessary as \( Q (\Phi) \) remains finite even in the unbounded case. Therefore, in what follows, we define \( C_q (\Phi, E) \) to be the value of (95) where the maximum is taken over ensembles \( \{p_x, \rho_x\} \) with the average state \( \rho = \sum_x p_x \rho_x \) satisfying the energy constraint

\[
\text{Tr} \rho a^\dagger a \leq E
\]

(133) \((a \text{ and } a^\dagger \text{ being the annihilation and creation operator of the mode}). \) Analogously, the entanglement-assisted classical capacity \( C_{eo} (\Phi, E) \) is defined by the expression (117) where the maximum taken over the states \( \rho \) satisfies equation (133).
It is not known in general if $C_\gamma(\Phi, E)$ is additive while taking tensor products of Gaussian channels which prevents from identifying $C_\gamma(\Phi, E)$ with the full constrained capacity $C(\Phi, E)$. However, the additivity property (94) holds for entanglement-breaking channels satisfying (72). Another case where additivity of $C_\gamma(\Phi, E)$ with the energy constraint was established is the case of pure loss channel, i.e. attenuator ($k < 1$) with $N_0 = 0$ (environment in the vacuum state) [58]. The actual computation of $C_\gamma(\Phi, E)$ is in general also an open problem: there is a natural conjecture that the maximum in $C_\gamma(\Phi, E)$ for a quantum Gaussian channel with quadratic energy constraint is attained on a Gaussian ensemble of pure Gaussian states [96, 161, 191], but so far this was only established for c-q channels and the pure loss channel [58]. It is also worth pointing out that this conjecture can be conveniently reformulated in terms of a property of the minimum output entropy [55, 57, 59]. If the conjecture is true, then in cases (i), (ii) and (iii) defined in section 3.5 the optimal ensemble is the continuous ensemble formed by the coherent states distributed with a Gaussian density $p(\xi) = \frac{1}{\pi\sigma_\xi} \exp(-|\xi|^2/\sigma_\xi)$, yielding

$$C(\Phi, E) = g(E') - g(D_\gamma(E)/2),$$

where $E' = k^2E + \max[0, k^2 - 1] + N_0$ is the mean number of quanta at the input when the input number of quanta is $E$, and where $g(x)$ is defined as in equation (88).

The computation of the entanglement-assisted capacity $C_{\text{ea}}(\Phi, E)$ is a relatively simpler problem as, on the one hand, no regularization over multiple uses is required, and on the other hand, the quantity to be maximized (the quantum mutual information) is a concave function [96] which, even in the presence of the linear constraint (133), admits a regular method of solution. In particular, for a Gaussian channel the maximum is always attained on a Gaussian state which can be found as a solution of certain Kuhn–Tucker equations. The expression of $C_{\text{ea}}(\Phi, E)$ for the one-mode channels was derived in [96] and generalized to the multimode case in [60]. Explicitly it is given by

$$C_{\text{ea}}(\Phi, E) = g(D_\gamma(E)/2),$$

where $g(x)$ and $E'$ are as before and where $D_\gamma(E) = D_\gamma(E') = D(E') + N_0(E') = E - 1$, with $D(E) = ((E + N_0(E) + 1)^2 - 4k^2E(E + 1)^2)$.

In general, entanglement-breaking channels have zero quantum capacity $Q(\Phi) = 0$. However, in any case, the domain (72) is superseded by the broader domain $N_0 \geq \min(1, k^2) - 1/2$ where the channel is anti-degradable and hence has zero quantum capacity [32]. On the other hand, in the case $N_0 = 0$, $k^2 > 1/2$ the channel is degradable [31, 32], hence the quantum capacity of the attenuation/amplification channel with $N_0 = 0$ and the coefficient $k$ is equal to the maximized single-letter Gaussian coherent information

$$Q(\Phi) = \sup_\rho I_\rho(\rho, \Phi) = \max\left\{0, \log_2 \frac{k^2}{|k^2 - 1|}\right\},$$

an expression conjectured in [96] and proved in [192]. The case with $N_0 > 0$ remains an open question. Enforcing the energy constraint as in equation (133), the above expression is replaced by

$$Q(\Phi, E) = \max\{0, g(D_\gamma(E)/2) - g(D_\gamma(E)/2)\}.$$

A plot of the above quantities is presented in figure 4 for the special case of the attenuation channel (69) with $N_0 = 0$, for which the expressions simplify as follows:

$$C(\Phi, E) = g(k^2E),$$

$$C_{\text{ea}}(\Phi, E) = g(E) + g(k^2E) - g((1 - k^2)E) = 2Q_{\text{ea}}(\Phi, E),$$

$$Q(\Phi, E) = \max\{0, g(k^2E) - g((1 - k^2)E)\}.$$ (138)

Here $Q_{\text{ea}}(\Phi, E)$ is the entanglement-assisted quantum capacity of the channel $\Phi$ (i.e. the quantum capacity which is achievable when the sender and the receiver are provided with prior shared entanglement): from general results [17] it is known to be always equal to half of the corresponding classical entanglement-assisted capacity. We also stress that for $N_0 = 0$ the reported value (138) for $C(\Phi, E)$ is the exact value of the classical capacity [58].

4.11. The variety of quantum channel capacities

The three capacities defined in equations (93), (117) and (124) are related by $Q(\Phi) \leq C(\Phi) \leq C_{\text{ea}}(\Phi)$ and form a basis for defining and investigating the diversity of various capacities of a quantum communication channel, which arises by the application of additional resources, such as reverse or direct communication, correlation or entanglement. In classical information theory it is well known that feedback does not increase the Shannon capacity which is essentially the unique characteristic of the classical channel. In the quantum case,
a similar property is established [22] for the entanglement-assisted capacity \( C_e(\Phi) \). Regarding the quantum capacity \( Q(\Phi) \), it is known that it cannot be increased with additional unlimited forward classical communication [9, 17]. However, \( Q(\Phi) \) can be increased if there is a possibility of transmitting the classical information in the backward direction. Such a protocol would allow one to create the maximum entanglement between the input and output, which can be used for quantum state teleportation. By this trick, even channels with zero quantum capacity supplemented with a classical feedback can be used for the reliable transmission of quantum information [22, 135] (a notable exception is PPT channels which have null quantum capacity even in the presence of feedback). Furthermore Smith and Yard [171] recently provided an explicit example of an interesting phenomenon named superactivation: there exist cases in which, given two quantum channels \( \Phi_1, \Phi_2 \) with zero quantum capacity (\( Q(\Phi_1) = Q(\Phi_2) = 0 \)), it is possible to have \( Q(\Phi_1 \otimes \Phi_2) > 0 \) (the latter being the capacity of the communication line \( \Phi_1 \otimes \Phi_2 \) obtained by using jointly \( \Phi_1 \) and \( \Phi_2 \)). The example of [171] was built by joining an anti-degradable channel \( \Phi_1 \) with a PPT channel \( \Phi_2 \), and a more general construction was given in [24].

A comparison of the various capacities which arise from employing additional resources is presented in figure 5, where the symbol ‘≤’ should be understood as ‘less than or equal to for all channels and strictly less for some channels’ [12]. In such a scheme \( Q_{\leq} \) and \( C_{\leq} \) denote, respectively, the quantum and classical capacities in the presence of a feedback. The symbol \( Q_{\geq} \) represents the quantum capacity in the presence of two-side classical communication [14]. The corresponding classical capacity \( C_{\geq} \) is computed under the limitation that the side communication is independent of the message transmitted through the main channel [12]. It is also known that \( C_{\leq} = 2Q_{\leq} \) [17] and that for some other pairs both inequalities are possible. Furthermore, one can construct the so-called ‘mother’ protocol which can implement all possible methods of transmission, including those mentioned above when using various additional resources (such as feedback or entanglement) [1]. We finally point out that of all the quantities entering in the scheme of figure 5, the function \( Q_1(\Phi) \) defined by equation (125) is the only one which has no clear operational definition in terms of information rate (apart from the cases in which it coincides with the quantum capacity—e.g. for the degradable channels). Still we have inserted it in the list as it provides a lower bound for \( Q(\Phi) \). Note also that \( C_{\leq}(\Phi) \) corresponds to the classical capacity of the channel \( \Phi \) restricted to separable encodings.

In classical information theory the role of the Shannon capacity is twofold: the converse statement of the coding theorem gives a fundamental upper bound on the channel performance which serves as an important benchmark for a real information-processing system. This is also the case for the present state of the art with quantum channel capacities. On the other hand, the direct statement of the coding theorem asserts that the bound is asymptotically achievable, although the proof does not give a practical recipe. In fact for almost 50 years after the Shannon proof the real performance was well below the theoretical bound and only more recently efficient practical codes appeared with rates approaching the capacity. This is still to be done in the quantum case, and the existing quantum error-correcting codes are the first promising steps in that direction. As an important example, in the papers [67, 73] symplectic codes were successfully applied to demonstrate constructively achievable rates close to the capacity for the additive classical noise channel [71].

4.12. Relative entropy

Before concluding our survey of quantum channel capacities and their entropic expressions, it is useful to consider a fundamental quantity—the quantum relative entropy—which underlies many information characteristics. Given two density operators \( \rho, \sigma \) in \( \mathcal{H} \), one defines the relative entropy of \( \rho \) with respect to \( \sigma \) as

\[
S(\rho\|\sigma) \equiv \text{tr} \rho (\log \rho - \log \sigma).
\]  

(139)

This quantity provides an important (asymmetric) measure of distinguishability of the states \( \rho, \sigma \); it is nonnegative and is equal to zero if and only if \( \rho = \sigma \) [138, 184].

The most important property of relative entropy which underlies several key facts in quantum information theory and nonequilibrium statistical mechanics is monotonicity under the action of quantum channels

\[
S(\rho\|\sigma) \geq S(\Phi[\rho]\|\Phi[\sigma]).
\]  

(140)

This means that states become less distinguishable after an irreversible evolution \( \Phi \). The first proof of the monotonicity property is due to Lindblad [122] who derived it from the strong subadditivity of quantum entropy, equation (114).

To stress the relevance of the relative entropy in the context of quantum information, let us first note that the quantum mutual information (101) can be written as

\[
I(\rho, \Phi) = S(\rho_B\|\rho_B) - I(\rho_B, \Phi_B),
\]  

(141)

implying \( I(\rho, \Phi) > 0 \) unless \( \rho_{RB} = \rho_B \otimes \rho_B \) when \( I(\rho, \Phi) = 0 \) (as already mentioned in section 4.5), and the data-processing inequality (116) as a particular instance of equation (140).

Next, the \( \chi \)-information (79) also admits a representation in terms of \( S(\rho\|\sigma) \), namely [156]

\[
\chi(\{p_x, \{\rho_x\}) = \sum_x p_x S(\rho_x\|\sum_y p_y \rho_y).
\]  

(142)
From this expression one can invoke the monotonicity property (140) to imply the following data-processing inequality
\[ \chi \left( \{ p_1 \}, \{ \Phi(p_1) \} \right) \leq \chi \left( \{ p_1 \}, \{ p_2 \} \right), \]
which in particular yields the bound (80) when \( \Phi \) is taken to be the q-c channel transforming quantum states into probability distributions (in this case the \( \chi \)-function on the left-hand side coincides with the Shannon information of the measurement process associated with the q-c mapping).

Inequality (143) also has profound implications for coherent information. Applying it to identity (129) we obtain that if the channel \( \Phi \) is anti-degradable, then for any state \( I_c(\rho, \Phi) \leq 0, \)

\[ (144) \]
(correspondingly, for degradable channel \( I_c(\rho, \Phi) \geq 0). \]

From the coding theorem (124) it then follows that all anti-degradable channels have zero quantum capacity as anticipated at the end of section 4.8.

Another useful application of the monotonicity (140) concerns what may be called the generalized H-theorem which states that a bistochastic (unital) evolution does not decrease the entropy. In other words, the H-theorem implies that given a unital channel \( \Phi \) and arbitrary input state \( \rho \), one has

\[ S(\Phi(\rho)) \geq S(\rho). \]

(145)

To see this in the finite dimensional case (dim \( \mathcal{H} = d < \infty \)), we use the identity

\[ S(\rho) = \log d - S(\rho) + I/d, \]

(146)

where \( I/d \) is the chaotic state, and note that

\[ S(\Phi(\rho)) = \log d - S(\Phi(\rho)) + I/d \]
\[ = \log d - S(\Phi(\rho)) + S(I/d) \]
\[ \geq \log d - S(\rho) + I/d = S(\rho), \]

(147)

where equation (140) and the fact that \( I/d = I/d \) for bistochastic channels were used.

For a general channel \( \Phi : \mathcal{S}(\mathcal{H}) \rightarrow \mathcal{S}(\mathcal{H}) \) an interesting characteristic is the minimal entropy gain defined by the quantity

\[ G(\Phi) = \inf_{\rho} \left[ S(\Phi(\rho)) - S(\rho) \right]. \]

(148)

In contrast to nonadditivity of the other similar quantity—the minimal output entropy \( \inf_{\rho} S(\Phi(\rho)) \)—which we introduced at the end of section 4.4, the quantity \( G(\Phi) \) is additive with respect to the tensor product of channels (i.e. \( G(\Phi_1 \otimes \Phi_2) = G(\Phi_1) + G(\Phi_2) \)) as a simple consequence of the strong subadditivity (114), see [4]. For a finite-dimensional system it is easy to see that

\[ -\log_2 d \leq G(\Phi) \leq 0 \]

(149)

(this follows directly from the fact that von Neumann entropies are upperbounded by \( \log_2 d \)). However, as shown in [94], there is a much better lower estimate

\[ -\log_2 \| \Phi[I] \| \leq G(\Phi). \]

(150)

which also holds for infinite-dimensional systems assuming that the channel \( \Phi \) is such that \( \Phi[I] = \sum_{j=1}^\infty V_j V_j^\dagger \) is a bounded operator. This implies that the generalized H-theorem is also valid for infinite-dimensional unital evolutions \( \Phi \) if we restrict ourselves to the input states \( \rho \) with finite entropy. The inequality \( G(\Phi) \leq 0 \) no longer holds; for example, if \( \Phi \) is a bosonic Gaussian channel with parameters \( (K, l, \beta) \), \( \det K \neq 0 \) (see section 3.4) then the minimal entropy gain is equal to

\[ G(\Phi) = \log_2 | \det K |. \]

(151)

and is attained on Gaussian states [94]. The quantity \( | \det K | \) is the coefficient of the change in the classical phase space volume under the linear transformation \( K \) (which can take arbitrary positive values). These results also suggest that for a general irreversible quantum evolution the role of this coefficient is played by the quantity \( \| \Phi[I] \|^{-1} \).

5. Summary and outlook

In this review we considered evolutions of open quantum systems from a quantum information viewpoint. More specifically, we discussed several scenarios where a sender and a receiver establish a communication line by using some physical degrees of freedom (the information carrier) subject to quantum noise from the environment. The resulting transformation of the sender’s input states into the receiver’s output states—the quantum channel in the Schrödinger picture—is described by a completely positive trace preserving (CPTP) map. Several alternative representations of such maps were introduced and their main features were analyzed. In particular, we have reviewed the operator-sum (Kraus) and unitary/isometric representations of a quantum channel, as well as its characterization in terms of the corresponding Choi–Jamiolkowski state. Composition rules along with the notions of dual (Heisenberg picture) and complementary channels have been presented. The general treatment was supplied with discussion of important particular cases of qubit, depolarizing, erasure channels as well as entanglement-breaking and dephasing channels. A whole section was dedicated to bosonic Gaussian channels which constitute a basic class of information processors for continuous-variable systems and provide a representation for some of the most usable quantum communication protocols.

In the second part of the review we surveyed approaches to evaluation of the quality of a given quantum channel, based on the notions of channel capacities and their entropic expressions. For the sake of clarity, we restricted the analysis to the basic case of memoryless communication models, where the noise operates identically and independently on each of the sender’s inputs. We started with recollections of the Shannon entropy of a classical random source and the Shannon capacity of a classical channel (relevant for a communication line where all the stages of information transfer—coding, transmission, decoding—are treated in terms of classical random processes), explaining how these quantities acquire operational meaning in the asymptotic of very long messages. Moving to the quantum domain, we demonstrated how these quantities admit multiple generalizations. Several different notions of channel capacities had to be introduced in view
of the fact that a quantum communication line can be used to transfer either classical or quantum messages, and assisted by various additional resources such as entanglement shared between the communicating parties or classical feedback, which in the Shannon theory either do not exist (entanglement) or do not increase the capacity (feedback). These quantities have operational definitions generalizing that of the Shannon capacity and admit closed expressions in entropic terms given by fundamental quantum coding theorems. The quantum correlations (entanglement) display themselves in the increase in information transmission rates as compared with protocols without entanglement. The notable cases of this phenomenon discussed in our paper include the superadditivity of the classical information transmission rates with respect to entangled decodings and encodings, the gain of the input–output entanglement assistance and superactivation of zero quantum channel capacity.

The information-theoretic view thus opens a completely new perspective of quantum irreversible evolutions—noisy communication channels—and enlightens the fascinating landscape of channel entropic characteristics, in which entanglement plays a crucial role.

Acknowledgments

VG acknowledges support from MIUR through FIRB-IDEAS Project No RBID08B3FM. AH was partially supported by an RFBR grant and the RAS program ‘Dynamical systems and control theory’.

References

[1] Abeyesinghe A, Devetak I, Hayden P and Winter A 2009 The mother of all protocols: restructuring quantum information’s family tree Proc. R. Soc. A 465 2537–63 (arXiv:quant-ph/0606225v1)
[2] Adami C and Cerf N J 1997 Capacity of noisy quantum channels Phys. Rev. A 56 3470–85
[3] Alber G, Beth T, Horodecki M, Horodecki P, Horodecki R, Rötteler M, Weinfurter H and Werner R F 2001 Quantum Information: An Introduction to Basic Theoretical Concepts and Experiments (Berlin: Springer)
[4] Alicki R 2004 Isotropic quantum spin channels and additivity questions arXiv:quant-ph/0402080v1
[5] Alicki R and Fannes M 2005 Note on multiple additivity of minimal Renyi entropy output of the Werner–Holevo channels Open Syst. Inform. Dyn. 11 339–42
[6] Amosov G G, Holevo A S and Werner R F 2000 On some additivity problems of quantum information theory Problems Inform. Transm. 36 25
[7] Ban M 1999 Quantum dense coding via a two-mode squeezed-vacuum state J. Opt. B: Quantum Semiclass. Opt. 1 L9–L11
[8] Barnum H, Knill E and Nielsen M A 2000 On quantum fidelities and channel capacities IEEE Trans. Inform. Theory 46 1317–29
[9] Barnum H, Nielsen M A and Schumacher B 1998 Information transmission through a noisy quantum channel Phys. Rev. A 57 4153–75
[10] Bouwmeester D, Pan J, Mattle K, Eibl M, Weinfurter H and Zeilinger A 1997 Experimental quantum teleportation Nature 390 575–9
[11] Bennett C H, Brassard G, Crépeau C, Jozsa R, Peres A and Wootters W K 1993 Teleporting an unknown quantum state via dual classical and EPR channels Phys. Rev. Lett. 70 1895–9
[12] Bennett C H, Devetak I, Shor P W and Smolin J A 2006 Inequalities and separations among assisted capacities of quantum channels Phys. Rev. Lett. 96 150502 (arXiv:quant-ph/0406086v1)
[13] Bennett C H, DiVincenzo D P and Smolin J A 1997 Capacities of quantum erasure channels Phys. Rev. Lett. 78 3217–20
[14] Bennett C H, DiVincenzo D P, Smolin J A and Wootters W K 1996 Mixed state entanglement and quantum error correction Phys. Rev. A 54 3824–51 (arXiv:quant-ph/9604024v2)
[15] Bennett C H, Fuchs C A and Smolin J A 1997 Entanglement-enhanced classical communication on a noisy quantum channel QCM96: Proc. Quantum Communication, Computing and Measurement ed O Hirota et al (New York: Plenum) pp 79–88
[16] Bennett C H and Shor P W 1998 Quantum information theory IEEE Trans. Inform. Theory 44 2724–42
[17] Bennett C H, Shor P W, Smolin J A and Thapliyal A V 1999 Entanglement-assisted classical capacity of noisy quantum channel Phys. Rev. Lett. 83 3081
[18] Bennett C H, Shor P W, Smolin J A and Thapliyal A V 2002 Entanglement-assisted capacity and the reverse Shannon theorem IEEE Trans. Inform. Theory 48 2637–55 (arXiv:quant-ph/0106052v2)
[19] Bennett C H and Wiesner S J 1992 Communication via one- and two-particle operators on Einstein–Podolsky–Rosen states Phys. Rev. Lett. 69 2881–4
[20] Boschi D, Branca S, De Martini F, Hardy L and Popescu S 1998 Experimental realization of teleporting an unknown pure quantum state via dual classical and Einstein–Podolsky–Rosen channels Phys. Rev. Lett. 80 1121
[21] Bowen G 2002 Entanglement required in achieving entanglement-assisted channel capacities Phys. Rev. A 66 052313
[22] Bowen G 2004 Quantum feedback channels IEEE Trans. Inform. Theory 50 2429–33
[23] Bowen G and Bose S 2001 Teleportation as a depolarizing quantum channel, relative entropy, and classical capacity Phys. Rev. Lett. 87 267901
[24] Brandão F G S L, Oppenheim and Strelchuk S 2011 When does noise increase the quantum capacity? arXiv:1107.4385v1
[25] Braunstein S L and Kimble H J 1998 Teleporting of continuous variable states Phys. Rev. Lett. 80 869–72
[26] Braunstein S L and Kimble H J 2000 Dense coding for continuous variables Phys. Rev. A 61 042302
[27] Braunstein S L and van Loock P 2005 Quantum information with continuous variables Rev. Mod. Phys. 77 513–77
[28] Bruss D 2002 Characterizing entanglement J. Math. Phys. 43 4237–51
[29] Caruso F, Eisert J, Giovannetti V and Holevo A S 2008 Multi-mode bosonic Gaussian channels New J. Phys. 10 083030
[30] Caruso F, Eisert J, Giovannetti V and Holevo A S 2011 The optimal unitary dilation for bosonic Gaussian channels Phys. Rev. A 84 022306 (arXiv:1009.1108)
[31] Caruso F and Giovannetti V 2006 Degradability of bosonic Gaussian channels Phys. Rev. A 74 062307
[32] Caruso F and Giovannetti V and Holevo A S 2006 One-mode bosonic Gaussian channels: a full weak-degradability classification New J. Phys. 8 310
[33] Caves C M and Drummond P B 1994 Quantum limits of bosonic communication rates Rev. Mod. Phys. 66 481–538
[34] Cerf N J, Ipe A and Rottenberg X 2000 Cloning of continuous quantum variables Phys. Rev. Lett. 85 1754–7
[35] Cerf N J, Leuch J and Polzik E S 2007 Quantum Information with Continuous Variables of Atoms and Light (London: Imperial College Press)
[36] Choi M-D 1975 Completely positive maps on complex matrices Linear Algebra Appl. 10 285–90
[37] Cover T M and Thomas J A 1991 Elements of Information Theory (Wiley Series in Telecommunications) (New York: Wiley)
[38] Cubitt T, Harrow A W, Leung D, Montanaro A and Winter A 2008 Counterexamples to additivity of minimum output $p$-Renyi entropy for $p$ close to 0 Commun. Math. Phys. 284 281–90 (arXiv:0712.3628v3)
[39] Cubitt T S, Ruskai M B and Smith G 2008 The structure of degradable quantum channels J. Math. Phys. 49 102104 (arXiv:0802.1360v2)
[40] D’Ariano G M, Kretschmann D, Schlingemann D and Werner R F 2007 Reexamination of quantum bit commitment: the possible and the impossible Phys. Rev. A 76 032328 (arXiv:quant-ph/0605224)
[41] Davies E B 1976 Quantum Theory of Open Systems (London: Academic)
[42] Dirac P A M 1958 The Principles of Quantum Mechanics 4th edn (Oxford: Oxford University Press)
[43] Domen B, Vanheuverzwijn P and Verbeure A 1979 Completely positive quasi-free maps on the CCR algebra Rep. Math. Phys. 15 27–39
[44] Devetak I 2005 The private classical information capacity and quantum information capacity of a quantum channel IEEE Trans. Inform. Theory 51 44–55 (arXiv:quant-ph/0304127v6)
[45] Devetak I, Harrow A W and Winter A 2004 A family of quantum protocols Phys. Rev. Lett. 93 230504
[46] Devetak I, Junge M, King C and Ruskai M B 2006 Multiplicativity of completely bounded $p$-norms implies a new additivity result Commun. Math. Phys. 266 37–63
[47] Devetak I and Shor P 2005 The capacity of a quantum channel for simultaneous transition of classical and quantum information Commun. Math. Phys. 256 287–303 (arXiv:quant-ph/051131v3)
[48] Di Vincenzo D, Shor P W and Smolin J 1998 Quantum channel capacities of very noisy channels Phys. Rev. A 57 830–9
[49] Eisert J and Wolf M M 2007 Gaussian quantum channels Quantum Information with Continuous Variables of Atoms and Light (London: Imperial College Press) pp 23–42 (arXiv:quant-ph/0505151v1)
[50] Evans D F and Lewis J T 1977 Dilations on irreversible quantum channels: the exact solution Commun. Math. Phys. 56 33–8 (arXiv:quant-ph/0506066v2)
[51] Fujiwara A and Hashizume T 2002 Additivity of the capacity of depolarizing channels Phys. Lett. A 299 469–75
[52] Fukuda M, King C and Moser D 2009 Comments on Hastings’ additivity counterexample arXiv:0905.3697v1
[53] Gabor D 1950 Communication theory and physics Progr. Math. Phys. 54 1116–87
[54] Gallager R G 1968 Information Theory and Reliable Communication (New York: Wiley)
[55] Garcia-Patron R, Navarrete-Benlloch C, Lloyd S, Shapiro J H and Cerf N J 2011 Majorization theory approach to the Gaussian channel minimum entropy conjecture arXiv:1111.1986v1
[56] Giovannetti V and Fazio R 2005 Information-capacity description of spin-chain correlations Phys. Rev. A 71 032314
[57] Giovannetti V, Guha S, Lloyd S, Maccone L and Shapiro J H 2004 Minimum output entropy of bosonic channels: a conjecture Phys. Rev. A 70 032315
[58] Giovannetti V, Guha S, Lloyd S, Maccone L, Shapiro J H and Yuen H P 2004 Classical capacity of the lossy bosonic channel: the exact solution Phys. Rev. Lett. 92 027902
[59] Giovannetti V, Holevo A S, Lloyd S and Maccone L 2010 Generalized minimal output entropy conjecture for one-mode Gaussian channels: definitions and some exact results J. Phys. A: Math. Theor. 43 415305
[60] Giovannetti V, Lloyd S, Maccone L and Shor P W 2003 Broadband channel capacities Phys. Rev. A 68 062323
[61] Giovannetti V, Lloyd S and Ruskai M B 2005 Conditions for multiplicativity of maximal $p$-norms of channels for fixed integer $p$ J. Math. Phys. 46 042105
[62] Gisin N and Thew R 2007 Quantum communication Nature Photon. 1 165
[63] Gisin N, Ribordy G, Tittel W and Zbinden H 2002 Quantum cryptography Rev. Mod. Phys. 74 145–95
[64] Glauber R 1963 The quantum theory of optical coherence Phys. Rev. 130 2529–39
[65] Gordon J P 1964 Noise at optical frequencies: information theory Proc. Int. School Physics ‘Enrico Fermi’, Course XXXI Quantum Electronics and Coherent Light ed P A Miles (New York: Academic) pp 156–81
[66] Gorini V, Frigerio A, Verri M, Kossakowski A and Sudarshan E C G 1978 Properties of quantum Markovian master equations Rep. Math. Phys. 13 149–73
[67] Gottesman D, Kitaev A and Preskill J 2001 Encoding a qubit in an oscillator Phys. Rev. A 64 012310 (arXiv:quant-ph/0008040)
[68] Grassl M, Beth T and Pellizzari T 1997 Codes for the quantum erasure channel Phys. Rev. A 56 33–8 (arXiv:quant-ph/9610042v2)
[69] Gühne O and Tóth G 2009 Entanglement detection Phys. Rep. 474 1–75
[70] Haag R and Kastler D 1964 An algebraic approach to quantum field theory J. Math. Phys. 5 848–61
[71] Hall M J H 1994 Gaussian noise and quantum-optical communication Phys. Rev. A 50 3295–303
[72] Hamada M 2002 Lower bounds on the quantum capacity and highest error exponent of general memoryless channels IEEE Trans. Inform. Theory 48 2547–57
[73] Harrington J and Preskill J 2001 Achievable rates for Gaussian quantum channels Phys. Rev. A 64 062301 (arXiv:quant-ph/0505151v1)
[74] Hausladen P, Josza R, Schumacher B, Westmoreland M and Wootters W 1996 Classical information capacity of a quantum channel Phys. Rev. A 54 1869–76
[75] Hayashi M 2006 Quantum Information: An Introduction (Berlin: Springer)
[76] Hayashi M and Nagaoka H 2003 General formulas for capacity of classical–quantum channels IEEE Trans. Inform. Theory 49 1753–68
[77] Hayden P 2007 The maximal $p$-norm multiplicativity conjecture is false arXiv:0707.3291
[78] Hayden P and Winter A 2008 Counterexamples to the maximal $p$-norm multiplicativity conjecture for all $p > 1$ Commun. Math. Phys. 284 263–80
[79] Helstrom C W 1976 Quantum Detection and Estimation Theory (New York: Academic)
[80] Henderson L and Vedral V 2001 Classical, quantum and total correlations J. Phys. A: Math. Gen. 34 6899–905
[81] Holevo A S 1972 On the mathematical theory of quantum communication channels Problemy Peredachi Informatsii 8 62–71
[187] Werner R F and Holevo A S 2002 Counterexample to an additivity conjecture for output purity of quantum channels J. Math. Phys. 43 4353–7
[188] Winter A 1999 Coding theorem and strong converse for quantum channels IEEE Trans. Inform. Theory 45 2481–5
[189] Winter A 2007 The maximum output $p$-norm of quantum channels is not multiplicative for any $p > 2$ arXiv:0707.0402v3
[190] Wyner A D 1975 The wire-tap channel Bell. Syst. Tech. J. 54 1355–87
[191] Wolf M, Giedke G and Cirac J I 2006 Extremality of Gaussian quantum states Phys. Rev. Lett. 96 080502 (arXiv:quant-ph/0509154v1)
[192] Wolf M, Pérez-García D and Giedke G 2007 Quantum capacities of bosonic channels Phys. Rev. Lett. 98 130501
[193] Wootters W K and Zurek W H 1982 A single quantum cannot be cloned Nature 299 802–3
[194] Yamamoto Y and Haus H A 1986 Preparation, measurement and information capacity of optical quantum states Rev. Mod. Phys. 58 1001–20