

## Quantum reading capacity

**Stefano Pirandola<sup>1,5,6</sup>, Cosmo Lupo<sup>2,5</sup>, Vittorio Giovannetti<sup>3</sup>,  
Stefano Mancini<sup>2,4</sup> and Samuel L Braunstein<sup>1</sup>**

<sup>1</sup> Computer Science, University of York, York YO10 5GH, UK

<sup>2</sup> School of Science and Technology, University of Camerino,  
I-62032 Camerino, Italy

<sup>3</sup> NEST, Scuola Normale Superiore and Istituto Nanoscienze-CNR,  
I-56126 Pisa, Italy

<sup>4</sup> INFN-Sezione di Perugia, I-06123 Perugia, Italy

E-mail: [pirs@cs.york.ac.uk](mailto:pirs@cs.york.ac.uk)

*New Journal of Physics* **13** (2011) 113012 (32pp)

Received 17 August 2011

Published 11 November 2011

Online at <http://www.njp.org/>

doi:10.1088/1367-2630/13/11/113012

**Abstract.** The readout of a classical memory can be modelled as a problem of quantum channel discrimination, where a decoder retrieves information by distinguishing the different quantum channels encoded in each cell of the memory (Pirandola 2011 *Phys. Rev. Lett.* **106** 090504). In the case of optical memories, such as CDs and DVDs, this discrimination involves lossy bosonic channels and can be remarkably boosted by the use of nonclassical light (quantum reading). Here we generalize these concepts by extending the model of memory from single-cell to multi-cell encoding. In general, information is stored in a block of cells by using a channel-codeword, i.e. a sequence of channels chosen according to a classical code. Correspondingly, the readout of data is realized by a process of ‘parallel’ channel discrimination, where the entire block of cells is probed simultaneously and decoded via an optimal collective measurement. In the limit of a large block we define the quantum reading capacity of the memory, quantifying the maximum number of readable bits per cell. This notion of capacity is nontrivial when we suitably constrain the physical resources of the decoder. For optical memories (encoding bosonic channels), such a constraint is energetic and corresponds to fixing the mean total number of photons per cell. In this case, we are able to prove a separation between the quantum reading capacity and the maximum information rate achievable by

<sup>5</sup> These authors are equal first authors.

<sup>6</sup> Author to whom any correspondence should be addressed.

classical transmitters, i.e. arbitrary classical mixtures of coherent states. In fact, we can easily construct nonclassical transmitters that are able to outperform any classical transmitter, thus showing that the advantages of quantum reading persist in the optimal multi-cell scenario.

## Contents

<b>1. Introduction</b>	<b>2</b>
<b>2. Basic model of memory: single-cell encoding</b>	<b>4</b>
2.1. Multi-copy probing and optical memories . . . . .	6
2.2. Assisted readout of optical memories . . . . .	7
<b>3. The simplest case: optical memory with binary cells</b>	<b>8</b>
3.1. Quantum versus classical reading . . . . .	12
<b>4. General model of memory: multi-cell encoding</b>	<b>12</b>
<b>5. Limit of a large block size</b>	<b>14</b>
5.1. Quantum reading capacity . . . . .	16
<b>6. Classical reading capacity</b>	<b>19</b>
<b>7. Nonclassical transmitters</b>	<b>22</b>
<b>8. Conclusion and discussion</b>	<b>28</b>
<b>Acknowledgments</b>	<b>28</b>
<b>Appendix A. Reduction to pure transmitters</b>	<b>29</b>
<b>Appendix B. Triviality of the unconstrained version of the capacity</b>	<b>29</b>
<b>Appendix C. Output entropy of the coherent-state transmitter</b>	<b>30</b>
<b>References</b>	<b>31</b>

## 1. Introduction

One of the central problems in the field of quantum information is the statistical discrimination of quantum states [1–3]. This is a fundamental issue in many protocols, including those of quantum communication [4] and quantum cryptography [5, 6]. A similar problem is the statistical discrimination of quantum channels, also called ‘quantum channel discrimination’ (QCD) [7–11]. In its basic formulation, QCD involves a discrete ensemble of quantum channels which are associated with some *a priori* probabilities. A channel is randomly extracted from the ensemble and given to a party who tries to identify it by using input states and output measurements. The optimal performance is quantified by a minimum error probability, which is generally nonzero in the presence of constraints (e.g. for a fixed number of queries or restricted space of the input states). In general, this is a double-optimization problem whose optimal choices are unknown, a feature that makes its exploration nontrivial. Moreover, QCD may also involve continuous ensembles. A special case is ‘quantum channel estimation’ where the ensemble is indexed by a continuous parameter with flat distribution and the goal is to estimate this parameter with minimal uncertainty (see e.g. [12] and references therein).

Besides its difficult theoretical resolution, QCD is also interesting for its potential practical implementations. For instance, it is at the basis of the decoding procedure of two-way quantum

cryptography [13] where the secret information is encoded in a Gaussian ensemble of phase-space displacements. QCD also appears in the quantum illumination of targets [14–19], where the sensing of a remote low-reflectivity object in a bright thermal environment corresponds to the binary discrimination between a very noisy/lossy channel (presence of target) and a completely depolarizing channel (absence of target).

More recently, QCD has been connected with another fundamental task: the readout of classical digital memories [20]. Thanks to this connection, the work [20] has laid the foundations for treating digital memories, such as optical discs, in the field of quantum information theory (see also the follow-up of [21–24]). The storage of data, i.e. the writing of the memory, corresponds to a process of channel encoding, where information is recorded into a cell by storing a quantum channel picked from some pre-established ensemble. Then readout corresponds to the process of channel decoding, which is equivalent to discriminating between the various channels of the ensemble. This is done by probing the cell using an input state, also called a ‘transmitter’, and measuring the output by a suitable detector or ‘receiver’.

In [20], this model was developed directly in the bosonic setting, in order to apply the results to optical memories, such as CDs and DVDs. The central investigation there regarded the comparison between classical and nonclassical transmitters, where ‘classical transmitters’ correspond to probabilistic mixtures of coherent states and encompass all the sources of light that are used in today’s data storage technology. By contrast, ‘nonclassical transmitters’ are only produced in quantum optics laboratories and are typically based on entangled, squeezed or Fock states [25]. As shown in [20], we can construct nonclassical transmitters that are able to outperform any classical transmitter. In particular, this happens in the regime of low energy, where a few photons are irradiated over each cell of the memory. This regime is particularly relevant for increasing data-transfer rates and storage capacities. Following the terminology of [20], we call ‘quantum reading’ the use of nonclassical transmitters to read data from classical digital memories.

The main results on the quantum reading of memories were based on the single-cell scenario, where each memory cell is written and read independently of all others. However, a supplementary analysis in [20] has also shown that the advantages of quantum reading persist when we extend the encoding of information from a single- to a multi-cell model. Assuming block encoding of data, one can use error correcting codes that make the readout flawless. In this scenario, it is possible to show that the error correction overhead can be made negligible at low energies only when we adopt nonclassical transmitters.

Motivated by these results, the present work provides a full general treatment of the quantum reading of memories in the multi-cell scenario. This is achieved by formalizing the most general kind of classical digital memory. In this model, information is stored in a block of cells by using a channel codeword, i.e. a sequence of channels chosen according to some classical code. Then, the readout of data is realized by a process of parallel channel discrimination. This means that the entire block of cells is probed in parallel and then decoded by an optimal collective measurement. Such a description encompasses all possible encoding and decoding strategies. Since the storage capacity of classical memories is usually very large, a typical memory can be made by many such encoding blocks or, alternatively, by a few of them but large in size. The most general scenario corresponds to the case where the whole memory is represented by a single, very large, encoding block which is read in a parallel fashion. In this limit, we can simplify the description of the memory and resort to the Holevo bound to quantify the amount of readable information. This bound enables us to define the quantum reading

capacity of the classical memory, which corresponds to the maximum *readable* information per cell. If we do not impose constraints, this capacity exactly equals the amount of information stored in each cell of the memory. However, this is no longer the case when we introduce physical constraints on the resources accessible to the reading device, as typically happens in realistic implementations.

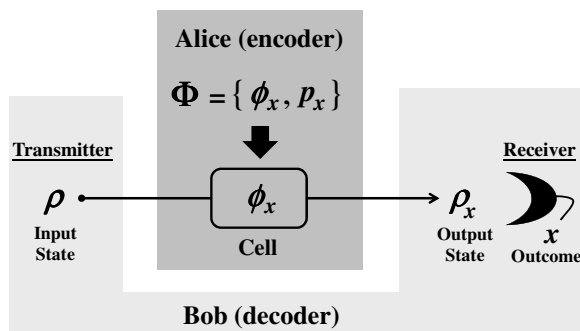
In the case of optical memories, which involve the discrimination of bosonic channels, the energy constraint is the most fundamental [4]. Thus, the quantum reading capacity is properly formulated for fixed input energy. This means that we fix the mean total number of photons irradiated over each cell of the memory. The computation of this capacity would be very important in the low-energy regime, which is the most interesting for its potential implications. Despite its calculation being extremely difficult, we are able to provide lower bounds for the most basic optical memories, i.e. the ones based on the binary encoding of lossy channels. For these memories we are able to derive a simple lower bound which quantifies the maximum information readable by classical transmitters. We call this bound the ‘classical reading capacity’ of the memory and it represents an extension to the multi-cell scenario of the ‘classical discrimination bound’ introduced in [20]. Remarkably, the optimal classical transmitter which irradiates  $n$  mean photons per cell can be realized by using a single coherent state with the same mean number of photons. Thanks to this result, we can easily investigate whether a particular nonclassical transmitter is able to outperform any classical transmitter. This is indeed what we find in the regime of a few photons. Thus, in the low-energy regime, we can prove a separation between the quantum reading capacity and the classical reading capacity, which is equivalent to the statement that the advantages of quantum reading persist into the optimal multi-cell scenario.

This paper is organized as follows. In sections 2 and 3, we review some of the key points of [20] and its supplementary materials, which are preliminary for the new results of sections 4–7. In particular, in section 2, we review the basic notions regarding the memory model with single-cell encoding. Then, in section 3, we discuss the simplest example of optical memory and its quantum reading. Once we have reviewed these notions, we introduce the model with multi-cell encoding in section 4. In section 5, we take the limit of large block size and define the quantum reading capacity of the memory, both unconstrained and constrained. In particular, we specialize the constrained capacity to the case of optical memories (bosonic channels). In section 6, we compute the lower bound relative to classical transmitters, i.e. the classical reading capacity. In section 7, we prove that this bound is separated, by showing simple examples of nonclassical transmitters which outperform classical ones in the regime of a few photons. Finally, section 8 presents our conclusions.

## 2. Basic model of memory: single-cell encoding

In an abstract sense, a classical digital memory can be modelled as a one-dimensional (1D) array of cells (the generalization to two or more dimensions is straightforward). The writing of information by some device or encoder, which we just call ‘Alice’ for simplicity, can be modelled as a process of channel encoding [20]. This means that Alice has a classical random variable  $X = \{x, p_x\}$  with  $k$  values  $x = 0, \dots, k - 1$  distributed according to a probability distribution  $p_x$ . Each value  $x$  is then associated with a quantum channel  $\phi_x$  via the one-to-one correspondence

$$x \leftrightarrow \phi_x, \quad (1)$$



**Figure 1.** Basic process of storage and readout. A memory cell can be characterized by an ensemble of quantum channels  $\Phi = \{\phi_x, p_x\}$ . Alice picks a quantum channel  $\phi_x$  (with probability  $p_x$ ) and stores it in a target cell. In order to read the information, Bob exploits a transmitter and a receiver. In the simplest scenario, this corresponds to inputting a suitable quantum state  $\rho$  and measuring the output  $\rho_x = \phi_x(\rho)$  by a suitable detector. The detector gives the correct answer  $x$  up to some error probability  $P_{\text{err}}$ . *Multi-copy probing.* Since the cell encodes the quantum channel in a stable way, we can probe the cell many times. More generally, this means that Bob can input a multipartite state  $\rho(s) \in \mathcal{D}(\mathcal{H}^{\otimes s})$  which describes  $s$  quantum systems. As a consequence, the output will be  $\rho_x(s) = \phi_x^{\otimes s}[\rho(s)]$ , whose global detection gives  $x$  up to some error probability. *Optical memory.* The encoded channel  $\phi_x$  is a bosonic channel (in particular, a single mode). In this case, Bob uses an input state  $\rho(s, n)$  describing  $s$  bosonic modes that irradiate  $n$  mean photons over the cell.

thus defining an ensemble of quantum channels

$$\Phi = \{\phi_x, p_x\}. \quad (2)$$

Mathematically speaking, each channel of the ensemble is a completely positive trace-preserving (CPT) map acting on the state space  $\mathcal{D}(\mathcal{H})$  of some chosen quantum system (Hilbert space  $\mathcal{H}$ ). Furthermore, the various channels are assumed to be different from each other. This means that, for any pair  $\phi_x$  and  $\phi_{x'}$  with  $x \neq x'$ , there is at least one state  $\rho \in \mathcal{D}(\mathcal{H})$  such that  $F[\phi_x(\rho), \phi_{x'}(\rho)] < 1$ , where  $F(\rho, \sigma) = [\text{Tr}(\sqrt{\rho} \sigma \sqrt{\rho})^{1/2}]^2$  is the quantum fidelity. Thus, in order to write information, Alice randomly picks a quantum channel  $\phi_x$  from the ensemble and stores it in a target cell. This operation is repeated independently and identically for all the cells of the memory, so that we can characterize both the cell and the memory by specifying  $\Phi$  (see figure 1).

The readout of information corresponds to the inverse process, which is channel decoding or discrimination. The written memory is passed to a decoder, which we call ‘Bob’, who queries the cells of the memory one by one. To retrieve information from a target cell, Bob exploits a transmitter and a receiver. In the simplest case, this means that Bob inputs a suitable quantum state  $\rho$  and measures the corresponding output state  $\rho_x = \phi_x(\rho)$  modified by the specific quantum channel stored in that cell (see figure 1). Note that, given some input state  $\rho$ , the ensemble of possible output states  $\{\phi_x(\rho), p_x\}$  is generally made by non-orthogonal states, which, therefore, cannot be perfectly distinguished by a quantum measurement. In other words, the discrimination cannot be perfect and the quantum detection will output the correct value

$x$  up to some error probability  $P_{\text{err}}$ . It is clear that the main goal for Bob is to optimize the input state and output measurement in order to retrieve the maximal information from the cell.

### 2.1. Multi-copy probing and optical memories

In a classical digital memory information is stored quasi-permanently. This means that the association between a single cell and the channel encoding  $\Phi$  must be stable. As a result, Bob can probe the cell many times by using an input state living in a larger state space. Given some quantum channel

$$\phi_x : \mathcal{D}(\mathcal{H}) \rightarrow \mathcal{D}(\mathcal{H}), \quad (3)$$

Bob can input a multipartite state  $\rho(s) \in \mathcal{D}(\mathcal{H}^{\otimes s})$  with integer  $s \geq 1$ , i.e. describing  $s$  quantum systems. As a result, the output state will be

$$\rho_x(s) = \phi_x^{\otimes s}[\rho(s)]. \quad (4)$$

This state is detected by a quantum measurement applied to the whole set of  $s$  quantum systems (see figure 1). Physically, if we consider the process in the time domain,  $\rho(s)$  describes the global state of  $s$  systems which are *sequentially* transmitted through the cell. In other words, the number  $s$  can also be regarded as a dimensionless readout time [20]. Intuitively, it is expected that the optimal  $P_{\text{err}}$  is a decaying function of  $s$ , so that it is always possible to retrieve all the information in the limit for  $s \rightarrow \infty$ . This suggests that the readout problem is nontrivial only if we impose constraints on the physical resources that are used to probe the memory. In the case of discrete variables (finite-dimensional Hilbert space) the constraint can be stated in terms of a fixed or maximum readout time  $s$ .

More fundamental constraints come into play when we consider an optical memory, which can be defined as a classical memory encoding an ensemble  $\Phi$  of bosonic channels. In particular, these channels can be assumed to be single-mode. Since the underlying Hilbert space is infinite in the bosonic setting, one has unbounded operators such as the energy. Clearly, if we allow the energy to go to infinity, the perfect discrimination of (different) bosonic channels is always possible. As a result, the readout of optical memories has to be modelled as a channel discrimination problem where we fix the input energy. The simplest nontrivial energy constraint corresponds to fixing the mean total number of photons  $n$  irradiated *over* each memory cell [20]. Thus, for fixed  $n$ , Bob's aim is to optimize the input (i.e. the number of bosonic systems  $s$  and their state  $\rho$ ) and the output measurement. In the following we explicitly formalize this constrained problem.

Let us consider an optical memory with the cell  $\Phi = \{\phi_x, p_x\}$  where each element  $\phi_x$  is a single-mode bosonic channel. Then, we denote by  $\rho(s, n)$  a multimode bosonic state  $\rho \in \mathcal{D}(\mathcal{H}^{\otimes s})$  with mean total energy  $\text{Tr}(\rho \hat{n}) = n$ , where  $\hat{n}$  is the total number operator over  $\mathcal{H}^{\otimes s}$ . In other words, this state describes  $s$  bosonic systems which irradiate a total of  $n$  mean photons over the target cell (also see figure 1). We refer to the pair  $(s, n)$  as the signal profile. In the bosonic setting the parameter  $s$  can be interpreted not only as the number of *temporal* modes (therefore readout time) but equivalently as the number of *frequency* modes, thus quantifying the 'bandwidth' of the signal [20]. Now, for a given input  $\rho = \rho(s, n)$  to the cell  $\Phi$ , we have the output state

$$\rho_x(s, n) = \phi_x^{\otimes s}[\rho(s, n)]. \quad (5)$$



This output is subject to a quantum measurement over the  $s$  modes, which is generally described by a positive operator-valued measure (POVM)  $\mathcal{M} = \{\Pi_x\}$  having  $k$  detection operators  $\Pi_x \geq 0$  which sums up to the identity  $\sum_x \Pi_x = I$ . This measurement gives the correct answer  $x$  up to an error probability

$$P_{\text{err}} = 1 - \sum_{x=0}^{k-1} p_x \text{Tr}[\Pi_x \rho_x(s, n)] := P[\Phi|\rho(s, n), \mathcal{M}]. \quad (6)$$

Here we denote by  $P[\Phi|\rho(s, n), \mathcal{M}]$  the error probability in the readout of the cell  $\Phi$  given an input state  $\rho(s, n)$  and an output measurement  $\mathcal{M}$ . We are now interested in minimizing this quantity over both input and output.

As a first step we fix the signal profile  $(s, n)$  and consider the minimization over input states and output measurements. This leads to the quantity

$$P(\Phi|s, n) = \min_{\rho(s, n), \mathcal{M}} P[\Phi|\rho(s, n), \mathcal{M}], \quad (7)$$

which is the minimum error probability achievable for a fixed signal profile  $(s, n)$ . Note that there are some cases where the optimal output POVM is known. For instance, if the output states  $\rho_x(s, n)$  are pure and form a geometrically uniform set [26, 27], then the optimal detection is the square-root measurement [1].

As a final step, we keep the energy  $n$  fixed and we minimize over  $s$ , thus defining the minimum error probability at a fixed energy per cell, i.e.

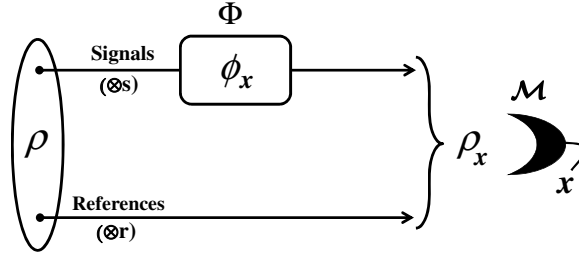
$$P(\Phi|n) = \inf_s P(\Phi|s, n). \quad (8)$$

Thus, given a memory with cell  $\Phi$ , the determination of  $P(\Phi|n)$  provides the ‘optimal’ readout of the cell at a fixed energy  $n$ . It is worth stressing that the minimization over the number of signals  $s$  is not trivial due to the constraint that we impose on the mean total energy (if instead of such a restriction one imposes a bound on the mean energy *per signal*, then the infimum is always achieved in the asymptotic limit as  $s \rightarrow \infty$ ). Also note that we have enclosed the word ‘optimal’ in quotes, since the optimality of equation (8) is still partial, i.e. not including all possible readout strategies. In fact, as we discuss in the following subsection, Bob can also consider the help of ancillary systems while keeping fixed the mean total number of photons  $n$  irradiated over the single cell.

## 2.2. Assisted readout of optical memories

The optimality of equation (8) is true only in the ‘unassisted case’ where all the input modes are sent through the target cell. More generally, Bob can exploit an interferometric-like setup by introducing an ancillary ‘reference’ system which bypasses the cell and assists in the output measurement as depicted in figure 2. In the ‘assisted case’ we consider an input state  $\rho \in \mathcal{D}(\mathcal{H}_S^{\otimes s} \otimes \mathcal{H}_R^{\otimes r})$  which describes  $s$  signal modes (Hilbert space  $\mathcal{H}_S^{\otimes s}$ ) plus a reference bosonic system with  $r$  modes (Hilbert space  $\mathcal{H}_R^{\otimes r}$ ). As before, the minimal energy constraint corresponds to fixing the mean total number of photons irradiated over the target cell, i.e.  $n = \text{Tr}(\rho \hat{n}_S)$  where  $\hat{n}_S$  is the total number operator acting over  $\mathcal{H}_S^{\otimes s}$ .<sup>7</sup> We denote by  $\rho = \rho(s, r, n)$  such a state, where

<sup>7</sup> Here we fix the signal energy effectively irradiated over the cell since this is the weakest constraint we can impose. Clearly this means that we do not impose any constraint on the reference modes. An alternative but more demanding energy constraint corresponds to fixing the mean number of photons of the global input state, i.e. including both signal and reference modes.



**Figure 2.** Assisted readout of an optical memory. Alice stores data in the cell by encoding a single-mode bosonic channel  $\phi_x$  picked from ensemble  $\Phi$ . In general, Bob queries the cell by using a transmitter  $\rho(s, r, n)$  which describes  $s$  signal modes, irradiating  $n$  mean photons over the cell, plus  $r$  reference modes (bypassing the cell). The global output state  $\rho_x(s, r, n)$  is detected by a quantum measurement  $\mathcal{M}$ , which provides the correct answer  $x$  up to error probability  $P[\Phi|\rho(s, r, n), \mathcal{M}]$ .

we make explicit the number of signal modes  $s$ , the number of reference modes  $r$  and the mean total number of photons  $n$  irradiated over the cell. Following the language of [20], we also refer to  $\rho(s, r, n)$  as a transmitter with  $s$  signals,  $r$  references and signalling  $n$  photons.

Now, given a transmitter  $\rho(s, r, n)$  at the input of a target cell  $\Phi = \{\phi_x, p_x\}$ , we have the output state

$$\rho_x(s, r, n) = (\phi_x^{\otimes s} \otimes I^{\otimes r})\rho(s, r, n), \quad (9)$$

where the channel  $\phi_x$  acts on each signal mode, while the identity  $I$  acts on each reference mode. This state is then measured by a POVM  $\mathcal{M} = \{\Pi_x\}$  where  $\Pi_x$  acts on the whole state space  $\mathcal{D}(\mathcal{H}_S^{\otimes s} \otimes \mathcal{H}_R^{\otimes r})$ . The error probability  $P[\Phi|\rho(s, r, n), \mathcal{M}]$  has the form of equation (6) where now both state and measurement are dilated by the reference system. Thus, given a memory with cell  $\Phi$ , the minimum error probability at a fixed signal energy  $n$  is given by

$$P(\Phi|n) = \inf_{s,r} \left\{ \min_{\rho(s,r,n), \mathcal{M}} P[\Phi|\rho(s, r, n), \mathcal{M}] \right\}, \quad (10)$$

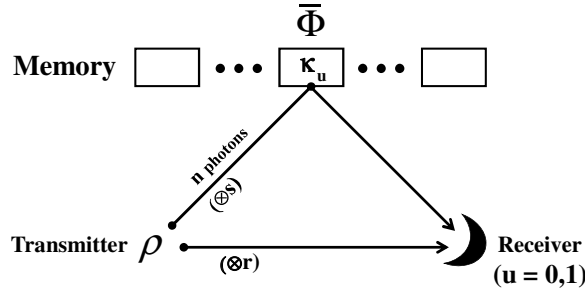
where the minimization includes the reference system too. In general, we always consider the assisted scheme and the corresponding error probability of equation (10). This clearly represents a superior strategy by allowing for the possibility of using entanglement between signal and reference systems. Clearly, the unassisted strategy is recovered by setting  $r = 0$  and  $\rho(s, 0, n) = \rho(s, n)$ .

### 3. The simplest case: optical memory with binary cells

In general, the solution of equation (10) is extremely difficult. In order to investigate the problem, the simplest possible scenario corresponds to an optical memory whose cell encodes two bosonic channels (binary cell) [20]. The situation is particularly advantageous when the channels are pure-loss channels and they are chosen with the same probability. This yields the binary channel ensemble

$$\bar{\Phi} = \{\phi_u, p_u\}_{u=0,1} = \{\phi_0, p_0, \phi_1, p_1\}, \quad (11)$$





**Figure 3.** Optical memory with the binary cell  $\bar{\Phi} = \{\kappa_0, \kappa_1\}$ , which is read in reflection. A bit of information is stored in the reflectivity  $\kappa_u$  of the cell medium ( $u = 0, 1$ ). The encoded bit is read by using a transmitter (with  $s$  signals and  $r$  references), which irradiates  $n$  mean photons over the cell. The output is detected by a dichotomic measurement which provides the value of the bit up to some error probability.

where  $p_0 = p_1 = 1/2$ , and  $\phi_u$  represents a pure-loss channel with transmissivity  $0 \leq \kappa_u \leq 1$ . In the Heisenberg picture, the action of  $\phi_u$  on each signal mode is given by the map

$$\hat{a}_S \rightarrow \sqrt{\kappa_u} \hat{a}_S - \sqrt{1 - \kappa_u} \hat{a}_E, \quad (12)$$

where  $\hat{a}_S$  is the annihilation operator of the signal mode and  $\hat{a}_E$  is that of an environmental vacuum mode. For simplicity we can also denote this ensemble by

$$\bar{\Phi} = \{\kappa_0, \kappa_1\}. \quad (13)$$

When the optical memory is read in reflection (which is usually the case), the two parameters  $\kappa_0$  and  $\kappa_1$  represent the two possible reflectivities of the cell (so that unit reflectivity corresponds to perfect ‘transmission’ of the signal from the transmitter to the receiver). See figure 3 for a schematic representation.

Given a transmitter  $\rho(s, r, n)$  at the input of the binary cell  $\bar{\Phi}$ , we have two equiprobable outputs,  $\rho_0(s, r, n)$  and  $\rho_1(s, r, n)$ . In this case the optimal measurement is dichotomic  $\mathcal{M} = \{\Pi_0, I - \Pi_0\}$ , where  $\Pi_0$  is the projection onto the positive part of the Helstrom matrix  $\rho_0(s, r, n) - \rho_1(s, r, n)$  [1]. As a result, the error probability for reading the binary cell  $\bar{\Phi}$  using the transmitter  $\rho(s, r, n)$  is given by

$$P[\bar{\Phi}|\rho(s, r, n)] = \frac{1}{2} \left\{ 1 - \frac{1}{2} D[\rho_0(s, r, n), \rho_1(s, r, n)] \right\}, \quad (14)$$

where  $D$  is the trace distance [1]. This expression has to be optimized on the input only, so that we may write

$$P(\bar{\Phi}|n) = \inf_{\rho} \{ \min_{s,r} P[\bar{\Phi}|\rho(s, r, n)] \}, \quad (15)$$

which is the minimum error probability at fixed signal energy. This quantity allows one to compute the maximum information per cell at fixed signal energy, which is given by

$$I(\bar{\Phi}|n) = 1 - H[P(\bar{\Phi}|n)], \quad (16)$$

where

$$H(x) = -x \log_2 x - (1-x) \log_2 (1-x) \quad (17)$$

is the binary formula for the Shannon entropy. For a binary cell we clearly have  $0 \leq I(\bar{\Phi}|n) \leq 1$ .

Even in this simple binary case, the solution of equation (15) is very difficult. However, we can provide remarkable lower bounds if we restrict the minimization to some suitable class of transmitters. An important class is the one of classical transmitters, since they encompass all the optical resources used for the readout of optical memories in today's storage technology. Furthermore, this class can be easily characterized. Given a transmitter  $\rho(s, r, n)$ , we can write its Glauber–Sudarshan representation [28, 29]

$$\rho(s, r, n) = \int d^{2s} \alpha d^{2r} \beta P(\alpha, \beta) \sigma(\alpha) \otimes \gamma(\beta), \quad (18)$$

where  $\alpha = (\alpha_1, \dots, \alpha_s)^T$  and  $\beta = (\beta_1, \dots, \beta_r)^T$  are vectors of complex amplitudes,

$$\sigma(\alpha) = \bigotimes_{i=1}^s |\alpha_i\rangle_S \langle \alpha_i|, \quad \gamma(\beta) = \bigotimes_{i=1}^r |\beta_i\rangle_R \langle \beta_i| \quad (19)$$

are multi-mode coherent states and the  $P$ -function  $P(\alpha, \beta)$  is a quasi-distribution, i.e. normalized to one but generally nonpositive [28, 29]. In terms of the  $P$ -function, the signal energy constraint reads

$$\int d^{2s} \alpha d^{2r} \beta P(\alpha, \beta) \sum_{i=1}^s |\alpha_i|^2 = n. \quad (20)$$

Now we say that  $\rho(s, r, n)$  is classical (nonclassical) if the  $P$ -function is positive ( $P \geq 0$ ) (nonpositive  $P \not\geq 0$ ). Thus if the transmitter is classical, denoted by  $\rho_c(s, r, n)$ , then it can be represented as a probabilistic mixture of coherent states. The simplest examples of classical transmitters are the coherent state transmitters, which we denote by  $\rho_{\text{coh}}(s, r, n)$ . These are defined by singular  $P$ -functions

$$P(\alpha, \beta) = \delta^{2s}(\alpha - \bar{\alpha}) \delta^{2r}(\beta - \bar{\beta}), \quad (21)$$

so that they have the simple form

$$\rho_{\text{coh}}(s, r, n) = \sigma(\bar{\alpha}) \otimes \gamma(\bar{\beta}).$$

Examples of nonclassical transmitters are constructed using squeezed states, entangled states and number states [25].

As shown in [20], by restricting the optimization to classical transmitters, we can compute the bound

$$P(\bar{\Phi}|n) \leq P_c(\bar{\Phi}|n) := \inf_{s,r} \left\{ \min_{\rho_c(s,r,n)} P[\bar{\Phi}|\rho_c(s,r,n)] \right\}, \quad (22)$$

which is given by

$$P_c(\bar{\Phi}|n) = \frac{1 - \sqrt{1 - \exp[-n(\sqrt{\kappa_0} - \sqrt{\kappa_1})^2]}}{2}. \quad (23)$$

This is the optimal performance of readout by means of classical transmitters. It is important to note that this bound can be reached by a coherent-state transmitter  $\rho_{\text{coh}}(1, 0, n) = |\sqrt{n}\rangle_S \langle \sqrt{n}|$ , i.e. a single-mode coherent state with mean number of photons equal to  $n$ . This achievability

is very easy to show. In fact, given the input state  $|\sqrt{n}\rangle_S$ , we have two possible coherent states  $|\sqrt{\kappa_0 n}\rangle_S$  and  $|\sqrt{\kappa_1 n}\rangle_S$  at the output of the cell. Since these are pure states, it is known that they can be discriminated up to an error probability [1]

$$P_{\text{err}} = \frac{1 - \sqrt{1 - \Omega}}{2}, \quad (24)$$

where  $\Omega$  is their overlap. In the case at hand, we have

$$\Omega = |\langle \sqrt{\kappa_0 n} | \sqrt{\kappa_1 n} \rangle|^2 = \exp[-n(\sqrt{\kappa_0} - \sqrt{\kappa_1})^2], \quad (25)$$

which proves that equation (23) can be reached by a single-mode coherent state. The error probability  $P_c(\bar{\Phi}|n)$  of equation (23) or, equivalently, the mutual information

$$I_c(\bar{\Phi}|n) = 1 - H[P_c(\bar{\Phi}|n)] \quad (26)$$

is called the ‘classical discrimination bound’<sup>8</sup>.

Alternative (and better) bounds can be derived by resorting to nonclassical transmitters. As a prototype of a nonclassical transmitter we consider an Einstein–Podolsky–Rosen (EPR) transmitter [14], which is composed of  $s$  pairs of signals and references, entangled via two-mode squeezing. This transmitter has the form

$$\rho_{\text{epr}}(s, s, n) = |\xi\rangle \langle \xi|^{\otimes s} \quad (27)$$

where  $|\xi\rangle \langle \xi|$  is a two-mode squeezed vacuum (TMSV) state, entangling one signal mode  $S$  with one reference mode  $R$ . In the number-ket representation, we have [25]

$$|\xi\rangle = (\cosh \xi)^{-1} \sum_{m=0}^{\infty} (\tanh \xi)^m |m\rangle_S |m\rangle_R, \quad (28)$$

where the squeezing parameter  $\xi$  quantifies the signal–reference entanglement and gives the energy of a single-mode signal by  $\sinh^2 \xi$ . Since this transmitter involves  $s$  copies of this state, we have to impose the constraint

$$\xi = \text{arcsinh} \sqrt{\frac{n}{s}}, \quad (29)$$

in order to have an average of  $n$  total photons irradiated over the cell. Given an EPR transmitter  $\rho_{\text{epr}}(s, s, n)$  at the input of the binary cell  $\bar{\Phi}$ , we have an error probability  $P[\bar{\Phi}|\rho_{\text{epr}}(s, s, n)]$ . By optimizing over the number of copies  $s$ , we define the upper bound

$$P(\bar{\Phi}|n) \leq P_{\text{epr}}(\bar{\Phi}|n) := \inf_s P[\bar{\Phi}|\rho_{\text{epr}}(s, s, n)]. \quad (30)$$

This bound represents the maximum information which can be read from the binary cell  $\bar{\Phi}$  by using an EPR transmitter which irradiates  $n$  mean photons over the cell. This quantity can be estimated using the quantum Battacharyya bound and its Gaussian formula [30]. After some algebra, we obtain [20]

$$P_{\text{epr}}(\bar{\Phi}|n) \leq B := \frac{\exp(-\mu n)}{2}, \quad (31)$$

<sup>8</sup> Note that the work [20] proved a more general form of this bound where the presence of thermal noise is included. However, when we consider the discrimination of lossy channels with nonzero thermal noise, we do not know if the classical discrimination bound is tight, i.e. achievable by some classical transmitters. As a consequence, in a more general scenario, we do not know if coherent-state transmitters are optimal among all the classical transmitters.

where

$$\mu := \frac{\kappa_0 + \kappa_1 + 2}{2} - 2\sqrt{\kappa_0\kappa_1} - \sqrt{(1 - \kappa_0)(1 - \kappa_1)}. \quad (32)$$

### 3.1. Quantum versus classical reading

Because of the potential implications in information technology, it is important to compare the performances of classical and nonclassical transmitters. The basic question to ask is the following [20]: for fixed signal energy  $n$  irradiated over a binary cell  $\bar{\Phi}$ , can we find some EPR transmitter able to outperform any classical transmitter? In other words, this is equivalent to showing that  $P_{\text{epr}}(\bar{\Phi}|n) < P_c(\bar{\Phi}|n)$  and a sufficient condition corresponds to proving that  $B < P_c(\bar{\Phi}|n)$ . Thus, by using equations (23) and (31), we find that for signal energies

$$n > n_{\text{th}} := \frac{2 \ln 2}{2 - \kappa_0 - \kappa_1 - 2\sqrt{(1 - \kappa_0)(1 - \kappa_1)}}, \quad (33)$$

it is always possible to beat classical transmitters by using an EPR transmitter [20]. For high reflectivity  $\kappa_1 \simeq 1$  and  $\kappa_0 < \kappa_1$  the threshold energy  $n_{\text{th}}$  can be very low. In the case of ‘ideal memories’, defined by  $\kappa_0 < \kappa_1 = 1$ , the bound of equation (31) can be improved. In fact, we can write

$$P_{\text{epr}}(\bar{\Phi}|n) \leq \theta := \frac{\exp[-2n(1 - \sqrt{\kappa_0})]}{2}, \quad (34)$$

and the threshold energy becomes  $n_{\text{th}} = 1/2$  [20]. Thus, for optical memories with high reflectivities and signal energies  $n > 1/2$ , there always exists a nonclassical transmitter able to beat any classical transmitter. In the few-photon regime, roughly given by  $1/2 < n < 10^2$ , the advantages of quantum reading can be numerically remarkable, up to one bit per cell. The implications have been thoroughly discussed in [20] and its supplementary materials. It is important to say that these advantages are also preserved if thermal noise is added to the basic model. This noise can describe the effect of stray photons hitting the memory from the background and other decoherence processes occurring in the reading device. Formally, this corresponds to extending the problem from the discrimination of pure-loss channels to the discrimination of more general Gaussian channels.

A supplementary analysis of quantum reading has also shown that its advantages persist if we consider more advanced designs of memories where information is written on and read from a block of cells (multi-cell/block encoding). Block encoding allows Alice to introduce error correcting codes which make Bob’s readout flawless up to some metadata overhead. By resorting to the Hamming bound and the Gilbert–Varshamov bound, one can show that EPR transmitters enable the low-energy flawless readout of classical memories up to a negligible error correction overhead, contrary to what happens by employing classical transmitters (see the supplementary materials of [20]). In the following section, we extend the block encoding to the most general scenario, i.e. for arbitrary classical memories. Then, by increasing the size of the block (section 5), we will introduce the notion of quantum reading capacity of a classical memory.

## 4. General model of memory: multi-cell encoding

The writing of a memory is based on channel encoding, which generally may involve a block of  $m$  cells. A first simple kind of block encoding is just based on independent and identical

extractions. As usual, Alice encodes a  $k$ -ary variable  $X = \{x, p_x\}$  into an ensemble of quantum channels  $\Phi = \{\phi_x, p_x\}$ . Then, she performs  $m$  independent extractions from  $X$ , generating an  $m$ -letter sequence

$$\mathbf{x} := (x_1, \dots, x_m), \quad (35)$$

with probability  $p_{\mathbf{x}} = p_{x_1}, \dots, p_{x_m}$ . This classical sequence identifies a corresponding ‘channel-sequence’

$$\phi_{\mathbf{x}} := \phi_{x_1} \otimes \dots \otimes \phi_{x_m}, \quad (36)$$

which is stored in the block of  $m$  cells.

In a more general approach, Alice adopts a classical code. This means that Alice disposes a set of  $m$ -letter codewords  $\{\mathbf{x}^0, \dots, \mathbf{x}^i, \dots, \mathbf{x}^{l-1}\}$  with  $l \leq k^m$ . A given codeword

$$\mathbf{x}^i = (x_1^i, \dots, x_m^i) \quad (37)$$

is chosen with some probability  $p_{\mathbf{x}^i}$  and identifies a corresponding ‘channel codeword’

$$\phi_{\mathbf{x}^i} = \phi_{x_1^i} \otimes \dots \otimes \phi_{x_m^i}. \quad (38)$$

Thus, in general, Alice encodes information in a block of  $m$  cells by storing a channel codeword, which is randomly chosen from the ensemble  $\Phi^m = \{\phi_{\mathbf{x}^i}, p_{\mathbf{x}^i}\}$ .

The most general strategy of readout can be described as a problem of ‘parallel discrimination of quantum channels’, where Bob probes the entire block in a parallel fashion and detects the output via a collective quantum measurement. In order to query the block, Bob uses  $s$  signal systems per cell besides  $r$  other supplemental reference systems for the benefit of the output measurement. The whole set of  $ms + r$  systems is described by an arbitrary multipartite state  $\rho$  (see figure 4). At the output of the block, Bob has

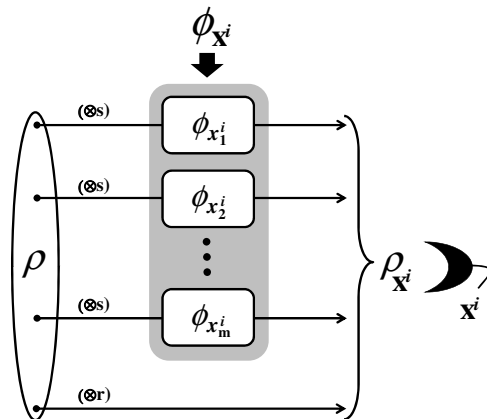
$$\rho_{\mathbf{x}^i} := (\phi_{\mathbf{x}^i}^{\otimes s} \otimes I^{\otimes r})(\rho), \quad (39)$$

where the identity acts on the reference systems, while

$$\phi_{\mathbf{x}^i}^{\otimes s} = \phi_{x_1^i}^{\otimes s} \otimes \dots \otimes \phi_{x_m^i}^{\otimes s} \quad (40)$$

acts on the signal systems. This state is detected by a collective quantum measurement, i.e. a general POVM with  $l$  detection operators with outcome  $i$  corresponding to codeword  $\mathbf{x}^i$ . The correct result is provided up to an error probability  $P_{\text{err}}$ . Clearly, the main goal for Bob is to optimize both input state and output measurements in order to minimize  $P_{\text{err}}$ , thus retrieving the maximal information from the block. One can easily check that, without constraints, Bob is always able to retrieve all the information  $H_{\text{max}}$ .

An important advantage of the block encoding model is that the readout of data can be made flawless even if we consider constraints for the decoder, e.g. we fix the properties of the transmitter. This is possible by increasing the error correction overhead in the memory. In fact, Alice can always use a suitably large block (with a suitable number  $m$  of cells) and an optimal code (encoding  $H_{\text{max}}$  bits) such that the error probability  $P_{\text{err}}$  becomes negligible, i.e. reasonably close to zero. In this case, Bob is able to retrieve all the information from the block, corresponding to an average of  $R = m^{-1}H_{\text{max}}$  bits per cell. In the following section, we study the asymptotic value of the rate  $R$  in the limit of large block size, corresponding to  $m \rightarrow \infty$ . Clearly, the readout process is too difficult to be treated if we consider arbitrary multipartite transmitters, i.e. generally entangled among different cells. Thus, in order to tackle the problem, we restrict the readout to transmitters which are in a tensor-product form. In this case, we can



**Figure 4.** Memory model with block encoding. In order to write data, Alice encodes a channel codeword  $\phi_{x^i}$  in a block of  $m$  cells. To read the data, Bob uses a suitable transmitter and receiver solving a problem of parallel channel discrimination. The transmitter is an arbitrary multipartite state  $\rho$  which probes the entire block by inputting  $s$  systems per cell plus sending an additional  $r$  systems directly to the receiver. The output state  $\rho_{x^i}$  is detected by an optimal collective measurement which provides the correct answer  $x^i$  up to some error probability  $P_{\text{err}}$ . In the case of unconstrained readout,  $P_{\text{err}}$  goes to zero and Bob retrieves all the information  $H_{\text{max}}$  from the block. If the readout is constrained, then  $P_{\text{err}}$  is generally nonzero. However, Bob can still retrieve all the information if Alice increases the size of the block while keeping  $H_{\text{max}}$  constant.

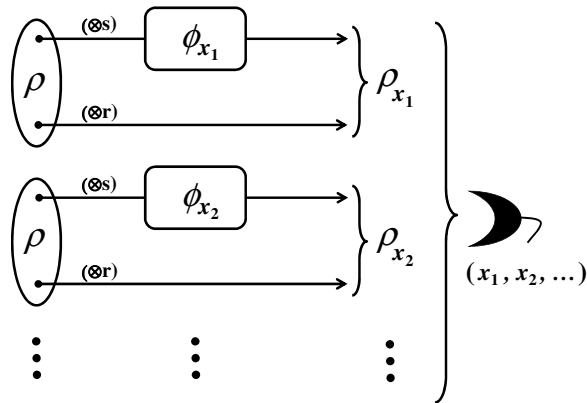
adopt the Holevo bound to quantify the readable information, i.e. the asymptotic rate  $R$ . The maximization of this rate over the transmitters enables us to define the quantum reading capacity of the memory.

## 5. Limit of a large block size

Digital memories typically store a great number of data. This means that an average memory can be made by many encoding blocks or, alternatively, by a small number of them but large in size. The most general scenario corresponds to describing the memory as a single large block of cells where Alice stores data by encoding a very long channel codeword  $\phi_{x^i}$  chosen with some probability  $p_{x^i}$ . Considering the whole memory as a large encoding block allows us to re-introduce the single-cell description. In fact, in the limit of  $m \rightarrow \infty$ , each cell can be described (on average) by a marginal ensemble of quantum channels  $\Phi = \{\phi_x, p_x\}$  encoding a corresponding marginal variable  $X = \{x, p_x\}$ . Thus, independently of the actual classical code used to store information, the description of a large classical memory can always be reduced to its ‘marginal cell’, corresponding to a marginal ensemble of channels  $\Phi$ .

Clearly, this is far from a *characterization* of the memory since the mapping ‘memory  $\rightarrow$  marginal cell’ is highly non-injective. For instance, any marginal ensemble  $\Phi = \{\phi_x, p_x\}$  can be realized by two completely different channel codewords: one is the sequence  $\phi_{x_1} \otimes \phi_{x_2} \otimes \dots$  where each  $\phi_{x_j}$  is independently and identically extracted from  $\Phi$ ; the other is the repetition codeword  $\phi_x \otimes \phi_x \otimes \dots$  which is given by infinite repetitions of the same channel  $\phi_x$  extracted





**Figure 5.** Limit of large block size. A memory can be described as a large (approximately infinite) encoding block, where each cell encodes a marginal ensemble  $\Phi = \{\phi_x, p_x\}$ . In order to read the memory, Bob uses an  $\infty$ -copy transmitter  $\rho(s, r)^{\otimes \infty} = \rho(s, r) \otimes \rho(s, r) \otimes \dots$ , where each copy  $\rho(s, r)$  probes a different cell using  $s$  signals generally coupled with  $r$  other references. All the output systems from the block are collectively detected by an optimal quantum measurement which reconstructs the asymptotic channel codeword.

once from  $\Phi$ . In general, for a fixed marginal cell  $\Phi$ , there are infinitely many block encodings giving that marginal. However, no matter what encoding is used, it is the marginal cell which determines the maximum amount of information that is *readable* from the memory. As we discuss below, this maximal amount is achievable by using an optimal block encoding whose codewords are based on the typical sequences.

Let us restrict the readout of the memory to transmitters which are in tensor product form with respect to different cells. This means that Bob inputs an  $\infty$ -copy state

$$\rho(s, r)^{\otimes \infty} = \rho(s, r) \otimes \rho(s, r) \otimes \dots, \quad (41)$$

where the single-copy  $\rho(s, r) \in \mathcal{D}(\mathcal{H}_S^{\otimes s} \otimes \mathcal{H}_R^{\otimes r})$  describes  $s$  signal systems sent through a target cell plus an additional  $r$  reference systems<sup>9</sup>. Given the  $\infty$ -copy transmitter  $\rho(s, r)^{\otimes \infty}$  at the input of a memory with marginal cell  $\Phi = \{\phi_x, p_x\}$ , the output is still in a tensor product form (see figure 5 for a schematic). The average output of each cell is described by a marginal ensemble of states

$$\mathcal{E} = \{\rho_x(s, r), p_x\}, \quad (42)$$

where

$$\rho_x(s, r) = (\phi_x^{\otimes s} \otimes I^{\otimes r})[\rho(s, r)]. \quad (43)$$

We can also use the notation  $\mathcal{E} = \Phi | \rho(s, r)$  indicating that the output marginal ensemble is generated by applying the marginal cell  $\Phi$  to the input transmitter  $\rho(s, r)$ . In other words, the output  $\mathcal{E}$  is equivalent to  $\Phi$  conditioned to the input  $\rho(s, r)$ . By carrying out an optimal collective

<sup>9</sup> In the remainder of this paper, we refer to ‘transmitter’ both when we consider the  $\infty$ -copy state  $\rho^{\otimes \infty}$  (irradiating the whole memory) or the single-copy state  $\rho$  (irradiating the marginal cell).

measurement on all the output systems, the maximum information per cell that can be retrieved is given by the Holevo bound

$$\chi(\mathcal{E}) = S \left[ \sum_x p_x \rho_x(s, r) \right] - \sum_x p_x S[\rho_x(s, r)], \quad (44)$$

where  $S(\rho) := -\text{Tr}(\rho \log_2 \rho)$  is the von Neumann entropy. This quantity can also be written as *conditional* Holevo information

$$\chi(\mathcal{E}) = \chi[\Phi|\rho(s, r)], \quad (45)$$

underlining the fact that it depends on the marginal cell  $\Phi$  up to fixing the input transmitter.

It is important to note that the achievability of  $\chi$  is assured by the Holevo–Schumacher–Westmoreland (HSW) theorem [32, 33]. According to the HSW theorem, it is possible to construct an asymptotic collective measurement which perfectly discriminates among different sequences of output states  $\rho_{x_1}(s, r) \otimes \rho_{x_2}(s, r) \otimes \dots$  as long as these sequences form a code of  $2^{m\chi}$  typical codewords  $\rho_{\mathbf{x}^i}$  (with  $m$  very large). This optimal code has marginal  $\mathcal{E}$ . Because of equation (43) we can always identify a corresponding set of channel codewords  $\phi_{\mathbf{x}^i}$  generating the typical codewords  $\rho_{\mathbf{x}^i}$  and having marginal  $\Phi$ . As a result, for a given marginal, there always exists an optimal block encoding  $\{\phi_{\mathbf{x}^i}, p_{\mathbf{x}^i}\}$  which enables the receiver to give the correct answer  $\mathbf{x}^i$  with asymptotically zero error while decoding an average of  $\chi$  bits per cell.

Thus, given a memory with marginal cell  $\Phi$  which is read by a transmitter  $\rho(s, r)$ , the conditional quantity  $\chi[\Phi|\rho(s, r)]$  provides the maximum information per cell, which can be reliably retrieved from the memory. This rate can always be achieved by using an optimal block code (with marginal  $\Phi$ ) which is based on the typical sequences. In the following sections, when we define the various reading capacities for a memory with marginal cell  $\Phi$ , it is understood that these capacities are achievable by a suitable choice of the block encoding.

### 5.1. Quantum reading capacity

As we have discussed above, the quantity  $\chi[\Phi|\rho(s, r)]$  represents the maximum information per cell which can be read from a memory with marginal cell  $\Phi$  if we use the transmitter  $\rho(s, r)$ . This readout is asymptotic (i.e. involves a large block) and reliable (i.e. without errors). Now the crucial task is the optimization of  $\chi[\Phi|\rho(s, r)]$  over the transmitters. As a first step, we can consider the readout capacity for a fixed number of input systems  $s$  and  $r$ , i.e.

$$C(\Phi|s, r) = \max_{\rho(s, r)} \chi[\Phi|\rho(s, r)]. \quad (46)$$

By optimizing equation (46) over the number of input systems, we can define the *unconstrained* quantum reading capacity of the memory

$$C(\Phi) = \sup_{s, r} \max_{\rho(s, r)} \chi[\Phi|\rho(s, r)]. \quad (47)$$

This is the maximum information per cell which can be read from a memory with marginal cell  $\Phi$ .<sup>10</sup>

<sup>10</sup> Technically speaking  $C(\Phi)$  is a ‘single-cell’ capacity since there is no entanglement distributed among the different cells. However, the probing of each cell includes the use of a generally entangled reference system, which is shared by Alice and Bob. For this reason, it represents an ‘entanglement-assisted’ capacity. It becomes ‘unassisted’ by setting  $r = 0$  in equation (47).

Since it is unconstrained, this capacity can be greatly simplified and easily computed. First of all, the maximization can be reduced to pure transmitters  $\psi(s, r)$  as a simple consequence of the convexity of the Holevo information [31] (see appendix A for more details). Further, the use of reference systems can be avoided. In other words, it is sufficient to consider the unassisted capacity where we maximize over  $\psi(s, 0) = \psi(s)$ . Finally, the supremum is achieved in the limit for  $s \rightarrow +\infty$ , i.e. we can write

$$C(\Phi) = \lim_{s \rightarrow +\infty} \max_{\psi(s)} \chi[\Phi|\psi(s)]. \quad (48)$$

This quantity is the maximal possible since it equals the maximum amount of information which can be read from the marginal cell of the memory. This is given by the Shannon entropy of the marginal variable  $X = \{x, p_x\}$ . In other words, we have

$$C(\Phi) = H(X) = - \sum_x p_x \log_2 p_x. \quad (49)$$

The proof is very easy (see appendix B for details).

The notion of quantum reading capacity is nontrivial only in the presence of physical constraints. This happens in the bosonic setting, where optical memories are read by fixing the input signal energy. Thus, consider an optical memory with marginal cell  $\Phi = \{\phi_x, p_x\}$ , where  $\phi_x$  represents a single-mode bosonic channel. As the transmitter, consider the  $\infty$ -copy state

$$\rho(s, r, n)^{\otimes \infty} = \rho(s, r, n) \otimes \rho(s, r, n) \otimes \cdots, \quad (50)$$

where  $\rho(s, r, n) \in \mathcal{D}(\mathcal{H}_S^{\otimes s} \otimes \mathcal{H}_R^{\otimes r})$  describes  $s$  signal modes, irradiating  $n$  mean photons on a target cell, plus  $r$  additional reference modes bypassing the cell. At the output we have an infinite tensor product of states of the form

$$\rho_x(s, r, n) = (\phi_x^{\otimes s} \otimes I^{\otimes r})[\rho(s, r, n)], \quad (51)$$

which are detected by an optimal collective measurement. In this way, Bob is able to retrieve an average of  $\chi[\Phi|\rho(s, r, n)]$  bits per cell. Now, we must optimize this quantity over the input transmitters while keeping the signal energy  $n$  fixed. This *constrained* optimization leads to the definition of the quantum reading capacity of the optical memory

$$C(\Phi|n) = \sup_{s,r} \max_{\rho(s,r,n)} \chi[\Phi|\rho(s, r, n)]. \quad (52)$$

This capacity represents the maximum information per cell which can be read from an optical memory with marginal cell  $\Phi$  by irradiating  $n$  mean photons per cell. The computation of equation (52) is not easy. In fact we are only able to provide lower bounds by restricting the class of transmitters involved in the maximization. We do not even know if the optimal transmitters are pure or mixed.

Let us consider a set (or ‘class’)  $\mathcal{P}$  of pure transmitters  $\psi(s, r, n)$  that are characterized by some general property which does not depend on  $s, r$  and  $n$  (for instance, they could be constructed using states of a particular kind, such as coherent states). Then we can always construct a mixed-state transmitter

$$\rho(s, r, n) = \int dy p_y \psi_y(s, r, n), \quad (53)$$

where  $p_y \geq 0$ ,  $\int dy p_y = 1$  and  $\psi_y(s, r, n) \in \mathcal{P}$ . Clearly, the set of mixed-state transmitters identifies a larger class  $\mathcal{A}$  which includes  $\mathcal{P}$ . Now, we can define a lower bound to  $C(\Phi|n)$  by optimizing over the class  $\mathcal{A}$ , i.e.

$$C(\Phi|n) \geq C_{\mathcal{A}}(\Phi|n) = \sup_{s,r} \max_{\rho(s,r,n) \in \mathcal{A}} \chi[\Phi|\rho(s,r,n)]. \quad (54)$$

Similarly, we can consider the further lower bound

$$C_{\mathcal{A}}(\Phi|n) \geq C_{\mathcal{P}}(\Phi|n) = \sup_{s,r} \max_{\psi(s,r,n) \in \mathcal{P}} \chi[\Phi|\psi(s,r,n)]. \quad (55)$$

Here we first ask: is there some class  $\mathcal{P}$  that allows one to saturate equation (55), i.e.  $C_{\mathcal{P}}(\Phi|n) = C_{\mathcal{A}}(\Phi|n)$ ? Then, is it possible to extend this class to all the pure transmitters, so that  $C_{\mathcal{P}}(\Phi|n) = C(\Phi|n)$ ?

Unfortunately, we are not able to answer the second question, so that the issue of the purity of the optimal transmitters remains unsolved. However, we are able to find classes for which  $C_{\mathcal{P}}(\Phi|n) = C_{\mathcal{A}}(\Phi|n)$ . For this task, a sufficient criterion is the concavity of  $C_{\mathcal{P}}(\Phi|n)$ .

**Lemma 1.** *If  $C_{\mathcal{P}}(\Phi|n)$  is concave in  $n$ , then we have  $C_{\mathcal{P}}(\Phi|n) = C_{\mathcal{A}}(\Phi|n)$ .*

**Proof.** Let us consider the transmitter of equation (53), whose signal energy (the mean number of photons) can be written as

$$n = \int dy p_y n_y, \quad n_y = \langle \psi_y | \hat{n} | \psi_y \rangle. \quad (56)$$

Given this transmitter at the input of a marginal cell  $\Phi$ , we can bound the conditional Holevo information

$$\chi[\Phi|\rho(s,r,n)] \leq \int dy p_y \chi[\Phi|\psi_y] \quad (57)$$

$$\leq \int dy p_y C_{\mathcal{P}}(\Phi|n_y) \quad (58)$$

$$\leq C_{\mathcal{P}}\left(\Phi \left| \int dy p_y n_y \right.\right) \quad (59)$$

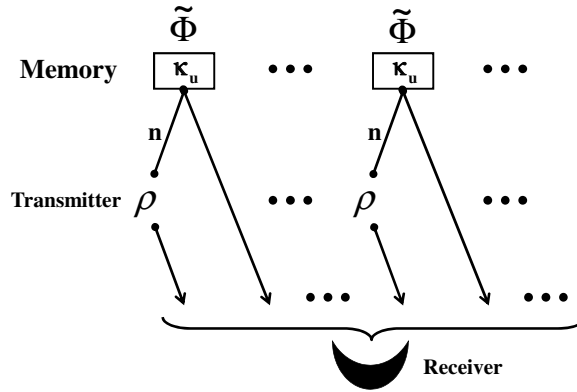
$$= C_{\mathcal{P}}(\Phi|n), \quad (60)$$

where we have used the convexity of  $\chi$  in the first inequality (57), the definition of  $C_{\mathcal{P}}(\Phi|n)$  in the second inequality (58) and its concavity in the last inequality (59). It is clear that equations (57)–(60) hold for every  $\rho(s,r,n) \in \mathcal{A}$  and every  $s$  and  $r$ . As a result, we may write

$$\sup_{s,r} \max_{\rho(s,r,n) \in \mathcal{A}} \chi[\Phi|\rho(s,r,n)] = C_{\mathcal{A}}(\Phi|n) \leq C_{\mathcal{P}}(\Phi|n), \quad (61)$$

which, combined with equation (55), proves the lemma.  $\square$

In the following section, we show that an important class  $\mathcal{P}$  for which  $C_{\mathcal{P}}(\Phi|n)$  is concave is that of coherent-state transmitters. This means that  $C_{\mathcal{P}}(\Phi|n) = C_{\mathcal{A}}(\Phi|n)$ , where  $\mathcal{A}$  is the class of classical transmitters (constructed by convex combination via the  $P$ -function). Thanks to this result we can compute an analytical bound for the readout performance of all the classical



**Figure 6.** Optical memory with binary marginal cell  $\tilde{\Phi}$  which is read in reflection. Here the parameter  $\kappa_u$  represents the unknown reflectivity of the cell medium ( $u = 0, 1$ ). The multi-cell readout is realized by using an  $\infty$ -copy transmitter  $\rho(s, r, n)^{\otimes \infty}$  which irradiates  $n$  mean photons per cell (each cell is probed by  $s$  signals coupled with  $r$  references). The total output is detected by an optimal collective measurement.

transmitters, which we call the ‘classical reading capacity’. This capacity represents the multi-cell generalization of the classical discrimination bound of section 3 and provides a simple lower bound to the quantum reading capacity. In section 6, we compute its analytical form for the most basic optical memories. Then, as we will show in section 7, this classical bound can be easily outperformed by nonclassical transmitters, thus proving its separation from the quantum reading capacity.

## 6. Classical reading capacity

Let us consider a multi-cell generalization of the binary model described in section 3. In the single-cell model of section 3, information was written in each cell in an independent fashion, by encoding one of two possible pure-loss channels,  $\phi_0$  and  $\phi_1$  (binary cell). Here we consider the multi-cell version, where Alice stores a channel codeword in the whole optical memory regarded as a large block. In particular, the block encoding is such that the marginal cell is described by a binary ensemble

$$\tilde{\Phi} = \{\phi_0, p, \phi_1, 1 - p\}, \quad (62)$$

where  $0 \leq p \leq 1$  and  $\phi_u$  is a pure-loss channel with transmissivity  $\kappa_u$  ( $u = 0, 1$ ). Alternatively, we can use the notation

$$\tilde{\Phi} = \{\kappa_0, p, \kappa_1, 1 - p\}. \quad (63)$$

Given this kind of memory, information is read by using an  $\infty$ -copy transmitter  $\rho(s, r, n)^{\otimes \infty}$  (irradiating  $n$  mean photons per cell) and an optimal collective measurement. In particular, when the memory is read in reflection (as is typical of optical discs), the two parameters  $\kappa_0$  and  $\kappa_1$  represent two reflectivities and the scenario is the one depicted in figure 6.

Now let us restrict the readout of this memory to classical transmitters. This means to consider the  $\infty$ -copy input

$$\rho_c(s, r, n)^{\otimes \infty} = \rho_c(s, r, n) \otimes \rho_c(s, r, n) \otimes \cdots, \quad (64)$$

where the single copy  $\rho_c(s, r, n)$  is an arbitrary classical state with  $s$  signal modes,  $r$  reference modes and  $n$  mean photons. The average information which can be read from each cell is provided by the Holevo quantity  $\chi[\tilde{\Phi}|\rho_c(s, r, n)]$ . By optimizing over classical transmitters we get the lower bound

$$C(\tilde{\Phi}|n) \geq C_c(\tilde{\Phi}|n) = \sup_{s,r} \max_{\rho_c(s,r,n)} \chi[\tilde{\Phi}|\rho_c(s, r, n)], \quad (65)$$

which defines the classical reading capacity of the optical memory  $\tilde{\Phi}$ . This capacity represents the multi-cell version of the classical discrimination bound of section 3. As before, we can provide a simple analytical result.

**Theorem 1.** *Let us consider an optical memory with binary marginal cell  $\tilde{\Phi} = \{\kappa_0, p, \kappa_1, 1 - p\}$  which is read by an arbitrary classical transmitter  $\rho_c(s, r, n)$  signalling  $n$  mean photons. Then, the maximum information per cell which can be retrieved is asymptotically equal to*

$$C_c(\tilde{\Phi}|n) = H(\xi), \quad (66)$$

where  $H$  is the binary Shannon entropy and

$$\xi = \frac{1}{2} - \frac{1}{2} \sqrt{1 - 4p(1-p) [1 - e^{-n(\sqrt{\kappa_0} - \sqrt{\kappa_1})^2}]}. \quad (67)$$

In particular, the bound  $C_c(\tilde{\Phi}|n)$  can be reached by using a coherent-state transmitter  $\rho_{\text{coh}}(1, 0, n) = |\sqrt{n}\rangle_S \langle \sqrt{n}|$ , i.e. a single-mode coherent state with  $n$  mean photons.

**Proof.** Consider the class  $\mathcal{P} = \text{coh}$  of coherent-state transmitters  $\rho_{\text{coh}}(s, r, n)$ . By convex combination we may construct the class  $\mathcal{A} = c$  of the classical transmitters  $\rho_c(s, r, n)$ . The first step of the proof is the computation of  $C_{\text{coh}}(\tilde{\Phi}|n)$ , i.e. the readout capacity restricted to coherent state-transmitters. We first prove that  $C_{\text{coh}}(\tilde{\Phi}|n) = \chi[\tilde{\Phi}|\rho_{\text{coh}}(1, 0, n)]$ , i.e. the optimal coherent-state transmitter is the single-mode coherent state  $|\sqrt{n}\rangle_S \langle \sqrt{n}|$ . Then, we analytically compute  $\chi[\tilde{\Phi}|\rho_{\text{coh}}(1, 0, n)]$ . Since this quantity turns out to be concave in  $n$ , we can use lemma 1 to demonstrate that  $C_{\text{coh}}(\tilde{\Phi}|n) = C_c(\tilde{\Phi}|n)$ , thus achieving the result of the theorem.

Given a coherent-state transmitter

$$\rho_{\text{coh}}(s, r, n) = \sigma(\alpha) \otimes \gamma(\beta) = \bigotimes_{i=1}^s |\alpha_i\rangle_S \langle \alpha_i| \otimes \bigotimes_{i=1}^r |\beta_i\rangle_R \langle \beta_i| \quad (68)$$

as the input to the cell  $\tilde{\Phi}$ , the output is

$$\begin{aligned} \rho_u &= \phi_u^{\otimes s}[\sigma(\alpha)] \otimes \gamma(\beta) \\ &= \bigotimes_{i=1}^s \phi_u(|\alpha_i\rangle_S \langle \alpha_i|) \otimes \bigotimes_{i=1}^r |\beta_i\rangle_R \langle \beta_i| \\ &= \bigotimes_{i=1}^s |\sqrt{\kappa_u} \alpha_i\rangle_S \langle \sqrt{\kappa_u} \alpha_i| \otimes \bigotimes_{i=1}^r |\beta_i\rangle_R \langle \beta_i|, \end{aligned} \quad (69)$$



which is still a multimode coherent state. This is a simple consequence of the fact that  $\phi_0$  and  $\phi_1$  are pure-loss channels. Since we are computing the Holevo information of the output ensemble, we have the freedom to apply a unitary transformation over  $\rho_u$ . By using a suitable sequence of beam splitters and phase-shifters we can always transform  $\rho_u$  into the state

$$|\sqrt{\kappa_u n}\rangle_S \langle \sqrt{\kappa_u n}| \otimes |0\rangle\langle 0|^{\otimes r+s-1}. \quad (70)$$

Then, since the Holevo information does not change under the addition of systems, we can trace out the  $r + s - 1$  vacua and just consider the single-mode output state

$$\rho_u = |\sqrt{\kappa_u n}\rangle_S \langle \sqrt{\kappa_u n}|. \quad (71)$$

This is equivalent to considering the single-mode coherent state transmitter

$$\rho_{\text{coh}}(1, 0, n) = |\sqrt{n}\rangle_S \langle \sqrt{n}| \quad (72)$$

at the input of the pure-loss channel  $\phi_u$ . For fixed marginal cell  $\tilde{\Phi}$  and fixed input energy  $n$ , the reduction from the multimode input of equation (68) to the single-mode output of equation (72) is always possible, independently of the actual number of systems,  $s$  and  $r$ , and the specific form of the coherent-state transmitter  $\rho_{\text{coh}}(s, r, n)$  (e.g. how the states are ordered and the energy distributed). Thus, we can write

$$C_{\text{coh}}(\tilde{\Phi}|n) = \sup_{s,r} \max_{\rho_{\text{coh}}(s,r,n)} \chi[\tilde{\Phi}|\rho_{\text{coh}}(s, r, n)] = \chi[\tilde{\Phi}|\rho_{\text{coh}}(1, 0, n)]. \quad (73)$$

In other words, the optimal coherent-state transmitter is the single-mode coherent state  $|\sqrt{n}\rangle_S \langle \sqrt{n}|$ . The next step is the analytical computation of  $\chi[\tilde{\Phi}|\rho_{\text{coh}}(1, 0, n)]$ .

The latter quantity corresponds to computing the Holevo information of the output ensemble  $\{|\sqrt{\kappa_0 n}\rangle_S, p, |\sqrt{\kappa_1 n}\rangle_S, 1 - p\}$ , which is just the von Neumann entropy of the average output state

$$\bar{\rho} = p|\sqrt{\kappa_0 n}\rangle_S \langle \sqrt{\kappa_0 n}| + (1 - p)|\sqrt{\kappa_1 n}\rangle_S \langle \sqrt{\kappa_1 n}|. \quad (74)$$

The computation of this output entropy is straightforward (see appendix C for details). After simple algebra, we obtain

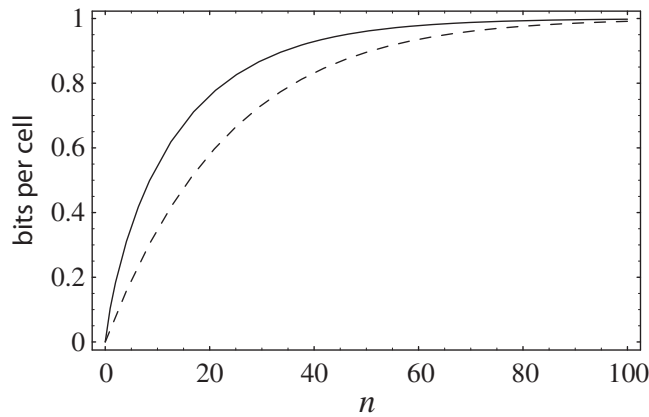
$$\chi[\tilde{\Phi}|\rho_{\text{coh}}(1, 0, n)] = H(\xi), \quad (75)$$

where  $H$  is the binary formula of the Shannon entropy and  $\xi = \xi(\kappa_0, \kappa_1, p, n)$  is given in equation (67). One can easily check that  $H(\xi)$  is a concave function of  $n$ , for any  $\kappa_0, \kappa_1$  and  $p$ . Since  $C_{\text{coh}}(\tilde{\Phi}|n) = H(\xi)$  is concave in the energy  $n$ , we can apply lemma 1 by setting  $\mathcal{P} = \text{coh}$  and  $\mathcal{A} = c$ . Thus we obtain  $C_c(\tilde{\Phi}|n) = C_{\text{coh}}(\tilde{\Phi}|n) = H(\xi)$ , which is the result of equation (66). It is clear that the optimal classical transmitter coincides with the optimal coherent-state transmitter, which is given by  $\rho_{\text{coh}}(1, 0, n)$ .  $\square$

It is interesting to compare the single-cell and multi-cell classical discrimination bounds, in order to estimate the gain which is provided by the parallel readout of cells. For this sake let us consider an optical memory whose binary marginal cell  $\tilde{\Phi}$  stores the maximal data, i.e. one bit of information. This corresponds to the limit situation

$$\tilde{\Phi} \rightarrow \bar{\Phi} = \{\kappa_0, 1/2, \kappa_1, 1/2\} = \{\kappa_0, \kappa_1\}. \quad (76)$$

This case is the direct generalization of the single-cell model analyzed in section 3. Then, we compare the maximum information achievable when using classical transmitters in the multi-cell model, i.e. the classical reading capacity  $C_c(\bar{\Phi}|n)$ , with the maximum information that is



**Figure 7.** Maximum number of bits per cell read by classical transmitters as a function of the signal energy  $n$  (the mean number of photons). We compare the two classical discrimination bounds:  $C_c(\bar{\Phi}|n)$  (multi-cell readout, solid line) and  $I_c(\bar{\Phi}|n)$  (single-cell readout, dashed line). Here we consider  $\kappa_0 = 0.5$  and  $\kappa_1 = 0.9$ .

achievable by classical transmitters in the single-cell model, i.e. the classical discrimination bound  $I_c(\bar{\Phi}|n)$  given in equation (26). In this comparison it is understood that the data retrieved by the multi-cell readout are flawless while those provided by the single-cell readout are generally affected by statistical errors and require post-processing. After post-processing, the  $I_c(\bar{\Phi}|n)$  bits of information are provided error-free. The previous two quantities can be easily expressed in terms of the coherent-state overlap  $\Omega$  defined in equation (25). In fact, we have

$$C_c(\bar{\Phi}|n) = H\left(\frac{1 - \sqrt{\Omega}}{2}\right), \quad I_c(\bar{\Phi}|n) = 1 - H\left(\frac{1 - \sqrt{1 - \Omega}}{2}\right). \quad (77)$$

Now the condition  $C_c(\bar{\Phi}|n) \geq I_c(\bar{\Phi}|n)$  reduces to

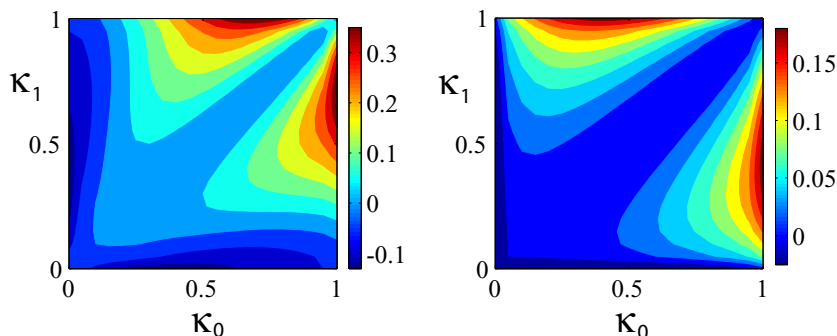
$$H\left(\frac{1 - \sqrt{\Omega}}{2}\right) + H\left(\frac{1 - \sqrt{1 - \Omega}}{2}\right) \geq 1, \quad (78)$$

which is true for any  $\Omega \in [0, 1]$ . To appreciate the advantage of the multi-cell readout we show a numerical example in figure 7.

In the following section, we will construct examples of nonclassical transmitters which are able to outperform the classical reading capacity. This will prove the separation between the quantum reading and the classical reading capacities, thus showing the advantages of quantum reading in the multi-cell scenario.

## 7. Nonclassical transmitters

As before, consider an optical memory with binary marginal cell  $\bar{\Phi} = \{\kappa_0, \kappa_1\}$  (maximal data storage). This time we shall assume that it is read using a nonclassical transmitter  $\rho_{nc}(s, r, n)$ . Since we are in the asymptotic multi-cell scenario, we clearly assume an  $\infty$ -copy input  $\rho_{nc}(s, r, n) \otimes \rho_{nc}(s, r, n) \otimes \dots$  together with an optimal collective measurement of the output.



**Figure 8.** Information gain  $G$  versus reflectivities,  $\kappa_0$  and  $\kappa_1$ , for  $n = 5$  (left panel) and  $n = 1$  (right panel). Here  $G$  provides the number of bits per cell which are gained by the single-copy EPR transmitter  $|\xi\rangle\langle\xi|$  over all the classical transmitters in the readout of an optical memory with marginal cell  $\bar{\Phi} = \{\kappa_0, \kappa_1\}$ . Note that the highest values of  $G$  occur for  $\kappa_0$  or  $\kappa_1$  close to 1 (high reflectivities).

The maximum number of readable bits per cell is given by the conditional Holevo information  $\chi[\bar{\Phi}|\rho_{nc}(s, r, n)]$ . Now we ask: is this quantity larger than the classical reading capacity  $C_c(\bar{\Phi}|n)$ ?

A design of the nonclassical transmitter is the EPR transmitter  $\rho_{\text{epr}}(s, s, n) = |\xi\rangle\langle\xi|^{\otimes s}$  which has been first discussed in section 3. In order to beat classical transmitters, it is sufficient to consider  $\rho_{\text{epr}}(1, 1, n) = |\xi\rangle\langle\xi|$ , i.e. a single TMSV state per cell. This means that we have one signal mode S, irradiating  $n$  mean photons over a target cell, which is entangled with one reference mode R. To quantify the advantage we consider the information gain

$$G = \chi[\bar{\Phi}|\rho_{\text{epr}}(1, 1, n)] - C_c(\bar{\Phi}|n) \quad (79)$$

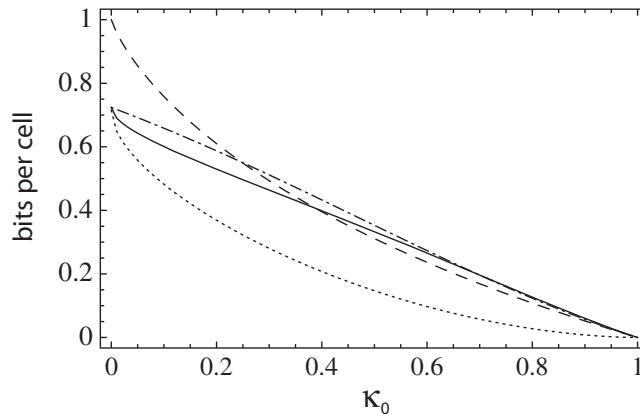
and check its positivity. If  $G > 0$ , then the EPR transmitter  $\rho_{\text{epr}}(1, 1, n)$  beats all the classical transmitters, retrieving  $G$  bits per cell more than any classical strategy. As shown in figure 8, we have  $G > 0$  in the regime of low photons and high reflectivities (i.e.  $\kappa_0$  or  $\kappa_1$  close to 1). This is the typical regime where the quantum reading of optical memories is advantageous, as also investigated in the single-cell scenario [20].

As is evident from figure 8 the best situation corresponds to having one of the two reflectivities equal to 1, i.e. for an ‘ideal memory’  $\bar{\Phi} = \{\kappa_0 < \kappa_1, \kappa_1 = 1\}$ . Given such a memory, we explicitly compare the information read by an EPR transmitter  $\chi_{\text{epr}} = \chi[\bar{\Phi}|\rho_{\text{epr}}(1, 1, n)]$  with the classical reading capacity  $C_c(\bar{\Phi}|n)$ . The comparison is made at low signal energy  $n$ . As shown in figure 9 for the numerical value  $n = 1$ , the EPR transmitter is always able to beat the classical bound.

It is important to note that we can construct other simple examples of nonclassical transmitters that can outperform the classical reading capacity. An alternative example of the nonclassical transmitter can be taken again of the form  $\rho_{nc}(1, 1, n)$  and corresponds to the NOON state [34, 35]

$$|\text{NOON}\rangle = 2^{-1/2}(|2n\rangle_S|0\rangle_R + |0\rangle_S|2n\rangle_R), \quad (80)$$

where the signal and reference are again entangled. A further example of the nonclassical transmitter is of the form  $\rho_{nc}(1, 0, n)$ , i.e. not involving the reference mode. This is the



**Figure 9.** The number of bits per cell as a function of  $\kappa_0$ , for  $\kappa_1 = 1$  (ideal memories) with  $n = 1$  mean photons per cell. We compare the classical reading capacity (dotted line) with the Holevo information retrieved by various nonclassical transmitters: the EPR transmitter (solid line), the NOON state transmitter (dash-dotted line) and the Fock state transmitter (dashed line).

#### Fock state

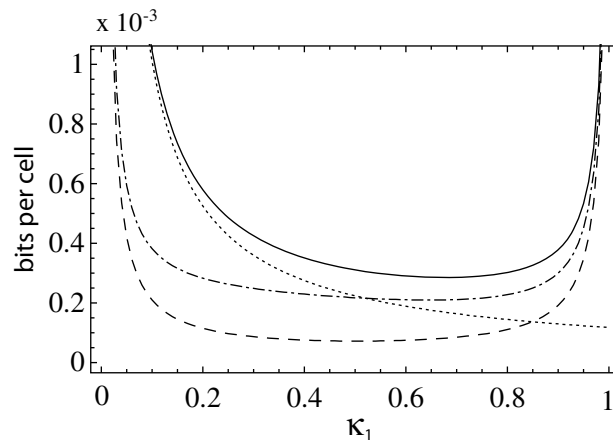
$$|n\rangle_S = (n!)^{-1/2} (a_S^\dagger)^n |0\rangle_S. \quad (81)$$

As shown in figure 9, these transmitters can beat not only the classical reading capacity but also the EPR transmitter  $|\xi\rangle \langle \xi|$  for low values of  $\kappa_0$ . Recently, these kinds of transmitters were also studied in [21] within the context of quantum reading with single-cell readout.

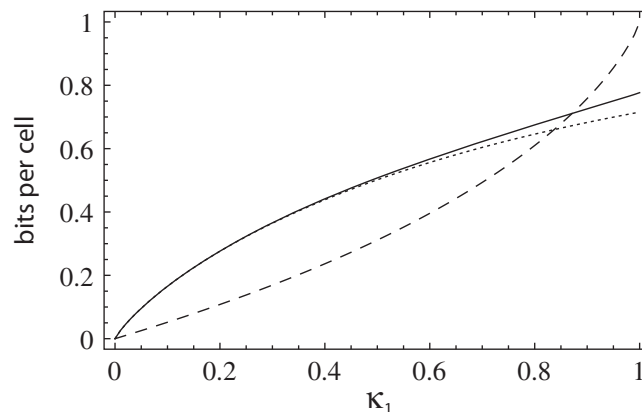
It is interesting to compare the performances of all these transmitters in the low-energy readout of optical memories with very close reflectivities. This is shown in figure 10 for  $\kappa_1 - \kappa_0 = 0.01$  and  $n = 1$ . The EPR transmitter  $|\xi\rangle \langle \xi|$  is optimal almost everywhere, while the classical bound beats the other nonclassical transmitters for low values of  $\kappa_1$  (and  $\kappa_0$ ). This is also compatible with the result of optimality of the TMSV state for the problem of estimating the unknown loss parameter of a bosonic channel [36]. As is evident from figure 10 a larger separation from the classical bound occurs for high reflectivities, i.e.  $\kappa_1$  close to 1.

It is also interesting to see what happens in the regime of low reflectivity by considering a binary marginal cell  $\Phi = \{\kappa_0, \kappa_1\}$  with  $\kappa_0 = 0$ . For this comparison we introduce another nonclassical transmitter of the form  $\rho_{nc}(1, 0, n)$ . This is the squeezed coherent state  $|\alpha, \xi\rangle = D(\alpha)S(\xi)|0\rangle$ , where  $D(\alpha)$  is the displacement operator and  $S(\xi)$  the squeezing operator [25]. The squeezed coherent state is chosen with the squeezing along the same direction of the displacement. Without loss of generality we consider this direction to be the axis of the position quadrature. This means to choose two real parameters,  $\alpha$  and  $\xi$ , which are optimized under the photon number constraint  $\alpha^2 + \sinh^2 \xi = n$ . As shown in figure 11, the presence of squeezing is sufficient to outperform the classical reading capacity in the regime of low reflectivity. However, better performances can be achieved by the Fock state for high values of  $\kappa_1$ .

From the previous analysis, it is evident that, in the regime of low photon number (down to one photon per cell), we can easily find nonclassical transmitters able to beat any classical transmitters, i.e. the classical reading capacity. This is particularly evident for high reflectivities ( $\kappa_1$  or  $\kappa_0$  close to 1). Thus, for the most basic optical memories, classical and quantum reading



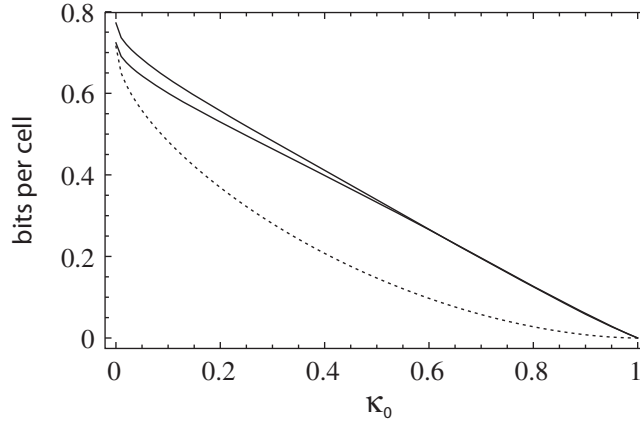
**Figure 10.** The number of bits per cell as a function of  $\kappa_1$ , for  $\kappa_1 - \kappa_0 = 0.01$  (close reflectivities) with  $n = 1$ . We compare the classical reading capacity (dotted line) with different nonclassical transmitters: the EPR transmitter (solid line), the NOON state transmitter (dash-dotted line) and the Fock state transmitter (dashed line).



**Figure 11.** The number of bits per cell as a function of  $\kappa_1$ , for  $\kappa_0 = 0$  with  $n = 1$ . We compare the classical reading capacity (dotted line) with the squeezed coherent state transmitter (solid line) and the Fock state transmitter (dashed line).

capacities are separated at low energies. In other words, the advantages of quantum reading are fully extended from the single- to the optimal multi-cell scenario.

At this point a series of important considerations are in order. First of all, note that we have only considered nonclassical transmitters irradiating one signal per mode (entangled or not with a single reference mode), i.e. transmitters of the kind  $\rho_{nc}(1, 0, n)$  or  $\rho_{nc}(1, 1, n)$ . The reason is because these transmitters are sufficient to beat the classical bound. However, better performances could be reached by optimizing over the number of signals and references. In the case of EPR transmitters, we expect that  $\rho_{\text{epr}}(2, 2, n)$ , which is composed of two TMSV states signalling  $n/2$  mean photons each, would be able to outperform  $\rho_{\text{epr}}(1, 1, n)$ , i.e. a single TMSV state signalling  $n$  mean photons. This is shown in figure 12 for the case of an ideal memory and



**Figure 12.** The number of bits per cell as a function of  $\kappa_0$ , for  $\kappa_1 = 1$  (ideal memory) with  $n = 1$ . We compare the classical reading capacity (dotted line) with two different EPR transmitters:  $\rho_{\text{epr}}(1, 1, n)$  (lower solid line) and  $\rho_{\text{epr}}(2, 2, n)$  (upper solid line).

$n = 1$  mean photons. This advantage could further improve for EPR transmitters  $\rho_{\text{epr}}(s, s, n)$  with higher values of  $s$ . For this reason, in order to reach the quantum reading capacity, it is necessary to optimize over an arbitrary number of signal and reference modes, as foreseen by the general definition of equation (52).

Another important consideration is related to the practical realization of quantum reading. In order to be experimentally feasible, the detection scheme should be as simple as possible. For this reason, it is interesting to compare the classical reading capacity (which refers to the general multi-cell readout) with the performances of EPR transmitters in the single-cell scenario, where each cell is encoded and detected independently of all others. Thus, we consider an ideal memory  $\bar{\Phi} = \{\kappa_0 < \kappa_1, \kappa_1 = 1\}$  which is irradiated by a few mean photons per cell (in particular, we consider the numerical value  $n = 5$ ). Given this memory, we compare the optimal performance  $C_c(\bar{\Phi}|n)$  of classical transmitters assuming the multi-cell readout (optimal block encoding and collective measurement) with the performance of EPR transmitters  $\rho_{\text{epr}}(s, s, n)$  assuming the single-cell readout (single-cell encoding and individual measurements). The latter quantity is given by the mutual information

$$I_{\text{epr}}(\bar{\Phi}|s, n) = 1 - H\{P[\bar{\Phi}|\rho_{\text{epr}}(s, s, n)]\}, \quad (82)$$

where  $H$  is the binary formula for the Shannon entropy and  $P[\bar{\Phi}|\rho_{\text{epr}}(s, s, n)]$  is the error probability of the single-cell readout. One can compute the upper bound

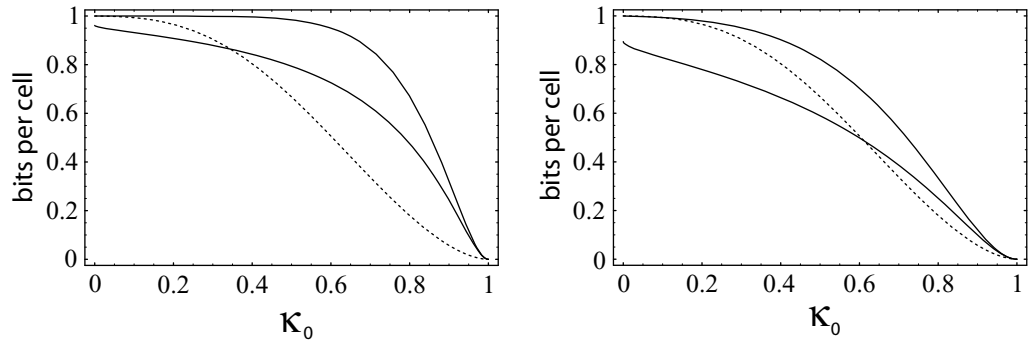
$$P[\bar{\Phi}|\rho_{\text{epr}}(s, s, n)] \leq \Theta := \frac{1}{2} \left[ 1 + \frac{n}{s} (1 - \sqrt{\kappa_0}) \right]^{-2s}, \quad (83)$$

which provides a lower bound for the mutual information

$$I_{\text{epr}}(\bar{\Phi}|s, n) \geq Q(\bar{\Phi}|s, n) := 1 - H(\Theta). \quad (84)$$

Thus,  $Q(\bar{\Phi}|s, n)$  provides the *minimum* number of bits per cell that are read by an EPR transmitter  $\rho_{\text{epr}}(s, s, n)$ . For fixed signal energy  $n$ , it is easy to check that this quantity





**Figure 13.** The number of bits per cell as a function of  $\kappa_0$ , for  $\kappa_1 = 1$  (ideal memory) with  $n = 5$ . Left panel: we compare the classical reading capacity  $C_c(\bar{\Phi}|n)$  (multi-cell readout, dotted line) with EPR transmitters used in the single-cell readout (solid lines). The lower solid line refers to  $Q(\bar{\Phi}|1, n)$ , i.e. a single energetic TMSV state  $\rho_{\text{epr}}(1, 1, n)$ , while the upper solid line refers to  $Q(\bar{\Phi}|\infty, n)$ , i.e. the optimal EPR transmitter  $\rho_{\text{epr}}(\infty, \infty, n)$  corresponding to infinite copies of TMSV states with vanishing signal energy. Right panel: as in the left panel, except that now we compare  $C_c(\bar{\Phi}|n)$  with  $Q(\bar{\Phi}|1, n/2)$  and  $Q(\bar{\Phi}|\infty, n/2)$ . Despite assuming a stronger energy constraint involving the mean total number of photons in both signal and reference modes, the single-cell quantum reading is still able to outperform asymptotic multi-cell classical reading.

is increasing in  $s$ . This means that for any integer  $s$  we have

$$Q(\bar{\Phi}|1, n) \leq Q(\bar{\Phi}|s, n) \leq Q(\bar{\Phi}|\infty, n), \quad (85)$$

where  $Q(\bar{\Phi}|1, n)$  corresponds to a single energetic TMSV state  $\rho_{\text{epr}}(1, 1, n)$  and  $Q(\bar{\Phi}|\infty, n)$  corresponds to  $\rho_{\text{epr}}(\infty, \infty, n)$ , i.e. infinite copies of TMSV states with vanishing energy. The quantity  $Q(\bar{\Phi}|\infty, n)$  is computed by taking the limit  $s \rightarrow \infty$  in equation (83). In this limit, we have  $\Theta \rightarrow \theta$ , with  $\theta$  given in equation (34). In the left panel of figure 13 we explicitly compare the two extremal values  $Q(\bar{\Phi}|1, n)$  and  $Q(\bar{\Phi}|\infty, n)$  with the classical reading capacity  $C_c(\bar{\Phi}|n)$ . As we can see, the single-cell quantum reading is able to beat asymptotic multi-cell classical reading.

Finally, it is interesting to check whether single-cell quantum reading represents a superior readout strategy even if we consider a stronger energy constraint; for instance, if we fix the mean total number of photons in both the signal and reference modes for each copy of the transmitter. Note that this approach has also been considered in [37] for the analysis of loss detection in bosonic channels. While this stronger energy constraint does not make any difference to the classical reading capacity (since the optimal classical transmitter involves signal modes only) it clearly affects the EPR transmitters where the mean total energy of the TMSV states is split exactly in two between signal and reference modes. Imposing this stronger energy constraint corresponds to comparing  $C_c(\bar{\Phi}|n)$  with  $Q(\bar{\Phi}|1, n/2)$  and  $Q(\bar{\Phi}|\infty, n/2)$ . As shown by the right panel of figure 13, we see that the single-cell quantum reading is still able to beat asymptotic multi-cell classical reading.

## 8. Conclusion and discussion

In this paper we have extended the model of quantum reading to the optimal and asymptotic multi-cell scenario. Here the classical memory is modelled as a large block of cells where information is stored by encoding a suitable channel codeword (channel encoding). This information is then retrieved by probing the whole memory in a parallel fashion and detecting the output via an optimal collective measurement (channel discrimination). In this general scenario, we define the quantum reading capacity of the memory.

It is important to note that this notion of capacity also applies to more practical scenarios where the memory is composed of a number  $>1$  of large encoding blocks which are independently read by the decoder. In fact, as long as each block is read in parallel by a tensor product transmitter and a collective measurement, and the marginal cells of different blocks are identical, our derivation can be repeated as before. In the case where the (large) blocks have different marginal cells  $\Phi_1, \Phi_2, \dots$ , the quantum reading capacity of the memory is given by a weighted sum of  $C(\Phi_1), C(\Phi_2), \dots$ . In other words, if the  $i$ th block occurs with probability  $p_i$ , then the memory can be described by the set  $\{p_1, \Phi_1, p_2, \Phi_2, \dots\}$  and its quantum reading capacity is given by  $p_1 C(\Phi_1) + p_2 C(\Phi_2) + \dots$ .

In our paper, we have then discussed how the quantum reading capacity is a nontrivial quantity to compute under the assumption of physical constraints for the decoder. In the case of optical memories, where data encoding is realized by bosonic channels, the main constraint of physical interest is energetic. This leads to defining the quantum reading capacity of an optical memory as the maximum number of bits per cell which can be read by irradiating  $n$  mean photons per cell.

Despite the difficulty of a general calculation of this capacity, we are able to provide nontrivial lower bounds in the case of optical memories with binary cells. The first lower bound, which we call the classical reading capacity, represents the maximum number of bits per cell which can be read by classical transmitters. This bound has a simple analytical formula and can be achieved using a single-mode coherent state transmitter. Besides this result, we have also computed other bounds by considering particular kinds of nonclassical transmitters, including the ones constructed with TMSV states (EPR transmitters), NOON states and Fock states. We have shown that, in the regime of a few photons and high reflectivities, these nonclassical transmitters are able to outperform any classical transmitter, thus showing a separation between the classical and the quantum reading capacities. It is remarkable that using a single-mode or two-mode transmitter per cell is already sufficient to beat any classical strategy. Furthermore, we have shown that the classical reading capacity can be outperformed even if we restrict the EPR transmitters to single-cell readout and adopt a stronger energy constraint where the energy of the reference modes is also taken into account.

In conclusion, our study considers the optimal multi-cell encoding for classical memories where we fully extend the advantages of quantum reading, i.e. the readout by nonclassical transmitters. These advantages are particularly evident in the regime of a few photons with nontrivial consequences for the technology of data storage.

## Acknowledgments

The research leading to these results has received funding from EPSRC under grant no. EP/J00796X/1 (HIPERCOM), the European Union (EU) under grant agreement no.

MOIF-CT-2006-039703 and the Italian Ministry of University and Research under the FIRB-IDEAS project RBID08B3FM. The work of CL and SM was supported by the EU under the FET-Open grant agreement HIP, no. FP7-ICT-221889. SP thanks Jeffrey H Shapiro, Saikat Guha and Ranjith Nair for discussions.

## Appendix A. Reduction to pure transmitters

For the sake of completeness, we show here that the maximization in equation (47) can be restricted to pure transmitters. This is a direct consequence of the convexity of the Holevo information.

Let us consider a classical memory with a marginal cell  $\Phi$  which is read by an arbitrary transmitter with  $s$  signals and  $r$  references

$$\rho(s, r) = \int dy p_y \psi_y, \quad \psi_y = |\psi_y\rangle\langle\psi_y|, \quad (\text{A.1})$$

where  $p_y \geq 0$  and  $\int dy p_y = 1$ . Then, the conditional Holevo information obeys the inequality

$$\chi[\Phi|\rho(s, r)] \leq \int dy p_y \chi(\Phi|\psi_y). \quad (\text{A.2})$$

In order to prove equation (A.2), let us consider an auxiliary system associated with the variable  $y$ . We denote by  $\{|y\rangle\}$  an orthonormal basis of this system. Then, we can express the transmitter as

$$\rho(s, r) = \text{Tr}_y \left( \int dy p_y \psi_y \otimes |y\rangle\langle y| \right). \quad (\text{A.3})$$

Since the Holevo information cannot increase under partial trace, we have

$$\chi[\Phi|\rho(s, r)] \leq \chi \left( \Phi \left| \int dy p_y \psi_y \otimes |y\rangle\langle y| \right. \right) = \int dy p_y \chi(\Phi|\psi_y). \quad (\text{A.4})$$

Thus, for any input transmitter  $\rho(s, r)$  we can always choose a pure transmitter  $\psi(s, r) = |\psi\rangle\langle\psi|$  such that  $\chi[\Phi|\rho(s, r)] \leq \chi[\Phi|\psi(s, r)]$ . As a result, the maximization in equation (47) can be restricted to pure transmitters  $\psi(s, r)$ .

## Appendix B. Triviality of the unconstrained version of the capacity

Here we provide a simple sketched proof showing that the unconstrained quantum reading capacity simply equals the whole data stored in the marginal cell of the memory.

Let us consider a pure transmitter in the tensor-product form  $\psi(s) = \psi^{\otimes s}$  at the input of a memory with marginal cell  $\Phi = \{\phi_x, p_x\}$ . At the output of the cell the arbitrary state is given by  $\rho_x(s) = [\phi_x(\psi)]^{\otimes s}$ , where  $\phi_x(\psi)$  is the single-copy output state. Since the quantum channels  $\phi_x$  are different, for any pair  $\phi_x$  and  $\phi_{x'}$  there is at least an input (pure) state  $\psi$  such that

$$F[\phi_x(\psi), \phi_{x'}(\psi)] = \varepsilon < 1. \quad (\text{B.1})$$

For the sake of simplicity, let us assume that this state  $\psi$  is the same for all the channels, i.e. equation (B.1) holds for any  $x \neq x'$ . Then, by exploiting the multiplicativity of the fidelity under tensor product states, we get

$$F[\rho_x(s), \rho_{x'}(s)] = \varepsilon^s, \quad (\text{B.2})$$

for any  $x \neq x'$ . Now, since this quantity goes to zero for  $s \rightarrow +\infty$  we have that the multi-copy output states  $\rho_x(s)$  become asymptotically orthogonal. This implies that

$$\chi[\Phi|\psi(s)] = \chi(\{\rho_x(s), p_x\}) \rightarrow H(X). \quad (\text{B.3})$$

The proof can be easily extended to the weakest case where equation (B.1) holds for different input states  $\psi_i$  where  $i = 0, \dots, k-1$  for a  $k$ -ary variable  $X$ . In this general case, for high values of  $s \gg k$ , we consider input states

$$\psi(s) = \psi_0^{\otimes s_0} \otimes \dots \otimes \psi_{k-1}^{\otimes s_{k-1}}, \quad (\text{B.4})$$

where  $s_0 + \dots + s_{k-1} = s$ . It is easy to check that the output states become asymptotically orthogonal.

### Appendix C. Output entropy of the coherent-state transmitter

Let us compute the von Neumann entropy of the state

$$\bar{\rho} = p|\alpha_0\rangle\langle\alpha_0| + (1-p)|\alpha_1\rangle\langle\alpha_1|, \quad (\text{C.1})$$

where  $|\alpha_0\rangle, |\alpha_1\rangle$  are two arbitrary coherent states and  $0 \leq p \leq 1$ . This state can be written in terms of two orthonormal vectors by using the Gram–Schmidt procedure. In fact, let us introduce the basis

$$|0\rangle := |\alpha_0\rangle, \quad |1\rangle := \frac{|\alpha_1\rangle - \omega|\alpha_0\rangle}{\sqrt{1-|\omega|^2}}, \quad (\text{C.2})$$

where

$$\omega := \langle\alpha_0|\alpha_1\rangle = \exp\left(-\frac{|\alpha_0|^2 + |\alpha_1|^2 - 2\alpha_0^*\alpha_1}{2}\right). \quad (\text{C.3})$$

In terms of this basis, the previous state is expressed by the density matrix

$$\bar{\rho} = \begin{pmatrix} p + (1-p)|\omega|^2 & (1-p)\omega\sqrt{1-|\omega|^2} \\ (1-p)\omega^*\sqrt{1-|\omega|^2} & (1-p)(1-|\omega|^2) \end{pmatrix}. \quad (\text{C.4})$$

Now it is easy to compute the spectral decomposition

$$\bar{\rho} = \xi_0|\xi_0\rangle\langle\xi_0| + \xi_1|\xi_1\rangle\langle\xi_1|, \quad (\text{C.5})$$

where  $|\xi_0\rangle$  and  $|\xi_1\rangle$  are orthonormal, and

$$\xi_0 = \xi \in [0, 1], \quad \xi_1 = 1 - \xi. \quad (\text{C.6})$$

In particular, the two eigenvalues are determined by

$$\xi = \frac{1}{2} - \frac{1}{2}\sqrt{1 - 4p(1-p)(1-|\omega|^2)}. \quad (\text{C.7})$$

The von Neumann entropy of the state  $S(\bar{\rho}) := -\text{Tr}(\bar{\rho} \log_2 \bar{\rho})$  is simply equal to the Shannon entropy of the classical variable whose probabilities correspond to the eigenvalues. Thus, here we have  $S(\bar{\rho}) = H(\xi)$ , where  $H$  is the binary formula of the Shannon entropy.

In the specific case where the two coherent states have amplitudes

$$\alpha_0 = \sqrt{\kappa_0 n}, \quad \alpha_1 = \sqrt{\kappa_1 n}, \quad (\text{C.8})$$

their overlap is given by

$$|\omega|^2 = \exp \left[ -n(\sqrt{\kappa_0} - \sqrt{\kappa_1})^2 \right]. \quad (\text{C.9})$$

By replacing this quantity in equation (C.7), we obtain the expression of equation (67). Thus we have computed the output entropy of a single-mode coherent-state transmitter  $|\sqrt{n}\rangle\langle\sqrt{n}|$  at the input of a binary cell with transmissivities  $\kappa_0$  and  $\kappa_1$ . Since the von Neumann entropy is invariant under unitaries, this is also the output entropy of an arbitrary coherent-state transmitter which irradiates  $n$  mean photons over the binary cell.

## References

- [1] Helstrom C W 1976 *Quantum Detection and Estimation Theory* (New York: Academic)
- [2] Fuchs C A and de Graaf J V 1999 *IEEE Trans. Inf. Theory* **45** 1216
- [3] Chefles A 2000 *Contemp. Phys.* **41** 401
- [4] Caves C M and Drummond P D 1994 *Rev. Mod. Phys.* **66** 481
- [5] Gisin N, Ribordy G, Tittel W and Zbinden H 2002 *Rev. Mod. Phys.* **74** 145
- [6] Scarani V, Bechmann-Pasquinucci H, Cerf N J, Dusek M, Lutkenhaus N and Peev M 2009 *Rev. Mod. Phys.* **81** 1301
- [7] Childs A, Preskill J and Renes J 2000 *J. Mod. Opt.* **47** 155–76
- [8] Acin A 2001 *Phys. Rev. Lett.* **87** 177901
- [9] Sacchi M 2005 *Phys. Rev. A* **72** 014305
- [10] Duan R, Feng Y and Ying M 2009 *Phys. Rev. Lett.* **103** 210501
- [11] Harrow A W, Hassidim A, Leung D W and Watrous J 2010 *Phys. Rev. A* **81** 032339
- [12] Giovannetti V, Lloyd S and Maccone L 2011 *Nat. Photonics* **5** 222
- [13] Pirandola S, Mancini S, Lloyd S and Braunstein S L 2008 *Nat. Phys.* **4** 726
- [14] Tan S-H, Erkmen B I, Giovannetti V, Guha S, Lloyd S, Maccone L, Pirandola S and Shapiro J H 2008 *Phys. Rev. Lett.* **101** 253601
- [15] Lloyd S 2008 *Science* **321** 1463
- [16] Shapiro J H and Lloyd S 2009 *New J. Phys.* **11** 063045
- [17] Yuen H P and Nair R 2009 *Phys. Rev. A* **80** 023816
- [18] Guha S and Erkmen B 2009 *Phys. Rev. A* **80** 052310
- [19] Usha Devi A R and Rajagopal A K 2009 *Phys. Rev. A* **79** 062320
- [20] Pirandola S 2011 *Phys. Rev. Lett.* **106** 090504
- [21] Nair R 2011 *Phys. Rev. A* **84** 032312
- [22] Bisio A, Dall'Arno M and D'Ariano G M 2011 *Phys. Rev. A* **84** 012310
- [23] Nair R and Yen B J 2011 Optimal quantum states for image sensing in loss arXiv:1107.1190
- [24] Hirota O 2011 Error free quantum reading by quasi Bell state of entangled coherent states arXiv:1108.4163
- [25] Walls F D and Milburn G 2008 *Quantum Optics* (Berlin: Springer)
- [26] Ban M, Kurokawa K, Momose R and Hirota O 1997 *Int. J. Theor. Phys.* **36** 1269
- [27] Kato K, Osaki M, Sasaki M and Hirota O 1999 *IEEE Trans. Commun.* **47** 248
- [28] Sudarshan E C G 1963 *Phys. Rev. Lett.* **10** 277
- [29] Glauber R J 1963 *Phys. Rev.* **131** 2766
- [30] Pirandola S and Lloyd S 2008 *Phys. Rev. A* **78** 012331
- [31] Nielsen M A and Chuang I L 2000 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)

- [32] Schumacher B and Westmoreland W D 1997 *Phys. Rev. A* **56** 131
- [33] Holevo A S 1998 *IEEE Trans. Inf. Theory* **44** 269
- [34] Kok P, Lee H and Dowling J P 2002 *Phys. Rev. A* **65** 052104
- [35] Dowling J P 2008 *Contemp. Phys.* **49** 125
- [36] Monras A and Illuminati F 2011 *Phys. Rev. A* **83** 012315
- [37] Invernizzi C, Paris M G A and Pirandola S 2011 *Phys. Rev. A* **84** 022334