Resource Theories of Quantum Engineering and Quantum Technologies

Gennaro Zanfardino

UNIVERSITY OF SALERNO

DEPARTMENT OF INDUSTRIAL ENGINEERING

Ph.D. Course in Industrial Engineering
Curriculum in Mechanical Engineering – XXXVIII Cycle

Resource Theories of Quantum Engineering and Quantum Technologies

Supervisor
Prof. Fabrizio Illuminati

Ph.D. Student
Gennaro Zanfardino

Scientific Referees
Dr. Victor ASV Bittencourt
Prof. Karol Życzkowski

Ph.D. Course Coordinator
Prof. Massimo De Santo

Publications in international journals:

1) V. A. S. V. Bittencourt, M. Blasone, and G. Zanfardino. *Chiral and flavor oscillations in lepton-antineutrino spin correlations*, Jour. of Phys.: Conf. Ser. **2533**, No. 1 (2023) DOI 10.1088/1742-6596/2533/1/012024.
2) V. A. S. V. Bittencourt, M. Blasone, and G. Zanfardino. *Chiral and flavor oscillations in a hyperentangled neutrino state*, Physica Scripta **99** 125310 (2024) DOI 10.1088/1402-4896/ad9097.
3) V. A. S. V. Bittencourt, M. Blasone, and G. Zanfardino. *Chiral oscillations in quantum field theory*, Phys. Lett. B **864** 139399 (2025) DOI 10.1016/j.physletb.2025.1595.
4) V. A. S. V Bittencourt, M. Blasone, and G. Zanfardino, *Chiral and flavor oscillations in quantum field theory: an overview*, Jour. of Phys.: Conference Series. **3017** (2025) DOI 10.1088/1402-4896/ad9097
5) A. L. Muthali, P. I. Sund, C. F. D. Faurby, C. Ekici, J. C. Adcock, J. Glindemann, G. Zanfardino, L. Leuzzi, R. Santagati, F. Illuminati, G. Ruocco, M. Leonetti, S. Paesani, and Y. Ding, *Photonic simulation of disordered spin systems on a silicon chip*, Adv. Phot. Nex. **5** 2, 025001 (2026) DOI 10.1117/1.APN.5.2.025001
6) G. Zanfardino, S. Paesani, L. Leuzzi, R. Santagati, F. Sciarrino, F. Illuminati, G. Ruocco, and M. Leonetti, *Multiphoton quantum simulation of the generalized Hopfield memory model*, Phys. Rev. Lett. **136** 070602 (2026) DOI 10.1103/945c-11wt.
7) G. Zanfardino, W. Roga, G. Tartaglione, M. Takeoka, F. Illuminati, *Geometric measures of quantum nonlocality: characterization, quantification, and comparison by distances and operations*, J. Phys. A **59** 055301 (2026) DOI 10.1088/1751-8121/ae3bba.
8) M. Blasone, and G. Zanfardino, *Unified field-theoretical description of flavor and chiral oscillations* , Phys. Lett. B **874** 140292 (2026) DOI 10.1016/j.physletb.2026.140292

Pre-print available on arXiv:

1) G. Tartaglione, G. Zanfardino, and F. Illuminati, *Uncertainty equality for $SU(d)$ observables and experimentally friendly detection of k-inseparability by purity measurements*. arXiv pre-print arXiv:2603.17844

Works in progress:

1) G. Tartaglione, G. Zanfardino, and F. Illuminati, *Bell nonlocality of*

multipartite Werner states.
2) G. Tartaglione, G. Zanfardino, and F. Illuminati, *Optimal Discrimination of atomic energy levels using two-mode squeezed states of light*.
3) V. A. S. V. Bittencourt, M. Blasone, F. Illuminati, B. Micciola, G. Tartaglione, and G. Zanfardino, *Quantum informational approach on the optimal discrimination between Dirac or Majorana nature of neutrinos*.
4) F. Illuminati, T. Giordani, M. Leonetti, L. Leuzzi, G. Ruocco, R. Santagati, F. Sciarrino, N. Spagnolo, and G. Zanfardino, *Experimental realization of photonic networks for the simulation of the Hopfield associative memory dynamics*.
5) G. Zanfardino, and F. Illuminati, *Exponential speedup in the quantum simulation of the Hopfield associative memory dynamics*.
6) V. Citro, F. Illuminati, G. Zanfardino, *Quantum simulations for fluid dynamics*.

Contents

List of Figures

Abstract

Quantum resource theory focuses on understanding and quantifying the various types of physical resources required to perform specific tasks, protocols, and operations in quantum engineering and quantum technologies that cannot be achieved by classical means.

The main goal of this PhD thesis is to explore how quantum resources can be exploited and characterized in different contexts of quantum technologies and quantum engineering. We begin by reviewing the resource theories of two fundamental features of quantum systems – coherence and entanglement – and their operational relevance. Building on these concepts, we introduce a geometric framework for defining *bona fide* measures of quantum nonlocality, a necessary step toward establishing a consistent resource theory of Bell nonlocality. The framework provides a unified geometric measure applicable to arbitrary Bell inequalities and measurement scenarios, allowing for explicit evaluation whenever the local set is characterized. Structural results are obtained for several important classes of states: the closest local states to Werner and isotropic states are shown to share the same symmetry as the original state, and, in the two-qubit case, the closest local state to any Bell-diagonal state is itself Bell-diagonal. Leveraging these results, for such states and for the Clauser-Horne-Shimony-Holt (CHSH) and Collins-Gisin-Linden-Massar-Popescu (CGLMP) inequalities, we derive explicit expressions for the geometric measures of nonlocality using various distance metrics.

Building on these concepts, we derive an exact uncertainty–purity relation valid for arbitrary finite-dimensional quantum states. While standard uncertainty relations are formulated for pairs of non-commuting observables, generalized uncertainty relations involving multiple observables have also been developed. In this thesis, the total uncertainty is defined over all local $SU(N)$ observables and it is shown to equal the sum of the purities of the global and reduced states, expressing an uncertainty conservation law. This equality reveals a direct quantitative link between uncertainty, purity, and entanglement, from which we derive a general criterion for k-separability that serves as a necessary–and for pure states, sufficient–condition, as well as experimentally friendly for detecting entanglement.

In the second part, we address the problem of quantum state discrimination, a fundamental task in quantum information and metrology. After reviewing the principles of optimal discrimination and the Helstrom bound – illustrated through the well-known example of quantum illumination – we extend the analysis to novel scenarios. In particular, we study the discrimination between two-level quantum systems using different optical resources, showing that also

in this case entangled two-mode squeezed light achieves the best performance compared to coherent and single-mode squeezed states. Finally, we apply the same theoretical framework in particle physics to address one of the fundamental open problems in the field: the nature of the neutrino, namely whether it is a Dirac or a Majorana particle. We investigate this question within our resource-theoretic approach, analyzing the ultimate discrimination between the two hypotheses based on the neutrinos' internal degrees of freedom, such as spin and chirality, and determining the optimal measurement strategy that maximizes their distinguishability.

The third part of the thesis focuses on the simulation of complex classical systems using quantum architectures. First, we establish a connection between multiphoton quantum interference in linear-optical networks and Hopfield-like Hamiltonians of classical neural networks. In particular, we demonstrate that a system of N_{ph} indistinguishable photons distributed over M modes realizes a p-body Hopfield Hamiltonian with $p = 2N_{\text{ph}}$, enabling the photonic simulation of memory retrieval and spin-glass phases. Building on this idea, we propose a general and platform-independent framework for simulating Hopfield neural networks using generic n-qubit systems. In this approach, the effective coupling terms naturally emerge from the coherence properties of the quantum density matrix, revealing quantum coherence as the essential resource enabling an exponential mapping from n qubits to 2^n classical neurons. Moreover, within the same formalism, we develop a specific protocol that achieves an exponential quantum speedup in simulating the network dynamics, thereby establishing a direct operational manifestation of the underlying resource.

Overall, the thesis establishes deep conceptual and operational connections between quantum resources and their applications in quantum discrimination, quantum metrology, and quantum simulation. These results contribute to the development of a unified framework for quantum resource engineering, providing both theoretical insight and practical tools for the design of future quantum technologies.

Chapter 1

Quantum resource theories

1.1 Introduction to quantum resource theories

Quantum resource theory focuses on understanding and quantifying the various types of physical resources required to perform specific quantum tasks or operations (Chitambar and Gour 2019). For instance, the resource theory of quantum entanglement studies how entanglement can be used to implement protocols that cannot be realized classically, e.g. teleportation above the classical fidelity threshold (Bennett et al. 1993; Popescu 1994), or how the use of entangled resources improves the efficiency of well-established protocols such as, e.g., quantum illumination (Tan et al. 2008) quantum reading (Pirandola 2011). Quantum resource theories have been developed for coherence (Streltsov et al. 2017), steering (Gallego and Aolita 2015), discord (Pirandola 2014; Dakić et al. 2012), but only partially for nonlocality (De Vicente 2014), the difficulty being that the set of local states, i.e. the set of state which admits a local hidden variable model description in any scenario, is not known (Brunner et al. 2014). In this chapter, first of all we briefly review the resource theories of coherence and entanglement, after that we will introduce some geometric measures for quantifying Bell nonlocality in Hilbert space, a necessary ingredient for building a resource theory. Our effort is motivated by the attempt to treat nonlocality and entanglement resource theories on the same footing, in order to highlight and understand analogies and differences (recall that for pure states discord, entanglement, steering and Bell nonlocality all coincide, while for mixed states a stringent hierarchy is established (see Fig. 1.1).

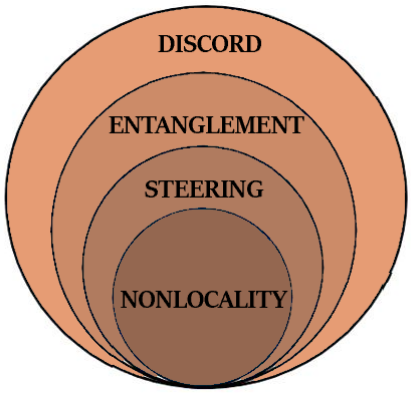


Figure 1.1: *Hierarchy of the features of a quantum system which can be used as quantum resources.*

1.1.1 General features

By identifying the general framework of resource theories we can distinguish their basic elements and features (Chitambar and Gour 2019; Theurer 2021; Kuroiwa and Yamasaki 2020; Gonda and Spekkens 2023). For a given context, a resource theory is specified by the set of resource states and the *free operations* acting on them. These elements induce a structure, often a partial ordering, among the states: one state precedes another if it can be obtained via a free operation. Introducing a meaningful measure of the resource further refines this ordering, allowing one to identify null-resource and maximal-resource states. Consistency requires that free operations cannot generate the resource from null-resource states and cannot increase it as quantified by the chosen measure. When composite systems are considered, the above-mentioned elements allow for distillation of the resource; this process amounts essentially to the localization of high-quality resource states in some subsystem via free operations acting on the larger system, see e.g. for entanglement (Horodecki et al. 1998).

In summary, a central task in resource theories is then to identify the *free states*, that is the states that do not carry the desired resource. The second ingredient of the theory are the transformations that do not create the resource, and thus preserve at least the free states. We already name them free operations, or sometimes resource non-generating transformations. There are a few general properties that free operations should satisfy (Theurer 2021; Chitambar and Gour 2019):

- Free operations preserve the set of free states.
- The trivial extension $\Phi \otimes \mathbb{1}$ of a free operation Φ is also a free operation

(as a consequence, tensor products of free operations are free operations).

- The operation of adding an auxiliary system is free.
- The operation of discarding a subsystem (partial trace) is free.

Next, we need appropriate quantifiers measuring how valuable is a given resource state. Typically, there are a few properties that a well defined quantifier should satisfy. Depending on how many of such properties are satisfied, the quantifier may be known as a *resource monotone* or as a full-fledged resource measure. The fundamental axiomatic properties include (Plenio and Virmani 2007; Christandl and Winter 2004; Eisert and Plenio 1999; Wootters 1998; Plenio 2005; Wei and Goldbart 2003; Blasone et al. 2008; Vedral et al. 1997; Vedral and Plenio 1998; Baumgratz et al. 2014; Bengtsson and Życzkowski 2017):

1. *Faithfulness*: A faithful resource quantifier must vanish if and only if the state is free.
2. *Monotonicity under free operations*: A resource quantifier must be monotonically non-increasing under free operations.
3. *Convexity*: A resource quantifier should be convex (monotonic under mixing).
4. *Monotonicity under subselection*: A resource quantifier must be monotonically non-increasing under subselection (quantum operations for which the measurement outcomes are retained and suitably selected).
5. *Computability*: A resource quantifier should be computable analytically or numerically.

A further central question in resource theories is the definition of the states carrying the maximal amount of a resource. Considering the case of bipartite entanglement for the purpose of illustration, the maximally entangled pure state is determined by showing that it is possible to deterministically obtain any other pure state from the identified maximally entangled one via the application of the free operations. In the case of entanglement, the latter are the *local operations and classical communication* (LOCC) transformations. On the other hand, the notion of maximally resourceful states is not so clear in the multipartite case and for mixed states. For example, considering entanglement again, for mixed states one needs to define the rate of distillation and the entanglement cost. This shows that conditions of convertibility under free operations are significant issues in quantum resource theories.

The current chapter is organized as follows: in the next section, we describe the fundamental ingredients necessary to construct a quantum resource theory. In Sections 1.2 and 1.3, we briefly review the quantum resource theories of coherence and entanglement, respectively. Following this, we discuss the quantum teleportation protocol, which serves as a prime example of utilizing entanglement as a resource. In section 1.3.2 We will address the problem of entanglement characterization, that beside the big effort is still an open question (Gühne and Tóth 2009; Horodecki et al. 2009), specifically finding a sufficient condition for k-separability starting from the uncertainty relations on SU(N) operators. Finally, in Section 1.5, we present our approach to developing a quantum resource theory for Bell nonlocality.

1.2 Quantum resource theory of coherence

Coherence is at the core of quantum mechanics, as it represents the coherent superposition of quantum states, one of the biggest differences compared to classical systems. From a physical perspective, coherence is a manifestation of the wave-like nature of quantum systems and is responsible for genuinely quantum interference effects. It can be experimentally observed in interference phenomena, such as the double-slit experiment (Griffiths and Schroeter 2018), where the superposition of different quantum paths gives rise to observable fringes. A necessary condition for the presence of coherence is the existence of superpositions of states corresponding to different eigenvalues of the observable under consideration. In this sense, coherence is inherently basis-dependent and reflects the ability of a quantum state to exhibit interference in a given measurement context. It is important to stress that the notion of quantum coherence considered here, defined with respect to a fixed reference basis, as in the resource-theoretic framework introduced in (Åberg 2006; Baumgratz et al. 2014; Streltsov et al. 2017), should be distinguished from the concept of coherent states originally introduced by Glauber (Glauber 1963a; Glauber 2007) for the electromagnetic field, as well as from related vector coherent states such as the SU(2) coherent states (Gilmore 1972; Perelomov 1972), whose definition does not depend on the choice of a specific basis.

A quantifier of coherence is called *coherence monotone*.

For a given density operator ρ the therms which quantify the coherence are the off-diagonal ones in its matrix representation.

Therefore, if we fix an orthonormal basis $\{|i\rangle\}$ we can define *incoherent states* those states (pure or not) whose density operator is diagonal in the fixed basis and we indicate with $\mathcal{I}$ the set of all incoherent states. Hence, if d is the

dimension of the Hilbert space which describes the system, $\rho \in \mathcal{I}$ if and only if

$$\rho = \sum_{i=1}^{d} p_i \left|i\right\rangle \left\langle i\right| \tag{1.1}$$

with $p_i \geq 0$ and $\sum_{i=1}^{d} p_i = 1$. As mentioned in the previous section, to build up a resource theory of coherence monotone it is necessary to define the *incoherent operations (IO)*, i.e., operations which do not increase coherence. There are different ways to define incoherent operations (Baumgratz et al. 2014; Streltsov et al. 2017). For example, by using that any quantum operation can be always specified by a set of Kraus operators $\{E_m\}$ such that $\sum_m E_m^\dagger E_m = I$, one defines incoherent operations such that if $\rho \in \mathcal{I}$ then

$$E_m \rho E_m^\dagger \in \mathcal{I} \tag{1.2}$$

for any m. This definition of IO can be used for both trace preserving (e.g., unitary) and non-trace preserving quantum operations, i.e., when the outcomes of the measurement are unknown or known respectively.

The relation (1.2) intuitively ensures that for any observer which measures the outcome m, no coherence arises under an IO starting from an incoherent state ρ.

We define the *maximally coherent state* as

$$\left|\Psi\right\rangle \equiv \frac{1}{\sqrt{d}} \sum_{i=1}^{d} \left|i\right\rangle . \tag{1.3}$$

This state is called maximally coherent state because any state can be obtained performing incoherent operations on it. The proof of this statement is given in the appendix section A.2.

C is a coherence monotone if:

1) $C(\rho) \geq 0$
and equality holds if and only if $\rho \in \mathcal{I}$.

Then we have to impose the monotonicity under incoherent operations and we distinguish the two cases of trace preserving and non-trace preserving operations.

For the trace preserving incoherent operations, we impose that

$2a)$ $C(\epsilon(\rho)) \leq C(\rho)$
for any trace preserving incoherent operation described by ϵ.

While for sub-selective measurements, i.e., measurements for which the

outcome is recorded and an appropriate post-selection scheme is applied (Nielsen and Chuang 2002), we have

2*b*) $C(\rho) \geq \sum_n p_n C(\rho_n)$

for any $\{E_n\}$, with

$$\sum_n E_n^\dagger E_n = I, \quad E_n \mathcal{I} E_n^\dagger \subseteq I, \quad \rho_n = \frac{E_n \rho E_n^\dagger}{\mathrm{Tr}(E_n \rho E_n^\dagger)} = \frac{E_n \rho E_n^\dagger}{p_n} \tag{1.4}$$

for any n.

Finally, since mixing is obtained classically and coherence is a quantum property of the system, it is also intuitive to impose that the coherence does not increase under mixing, that is

3) $\sum_n p_n C(\rho_n) \geq C\Big(\sum_n p_n \rho_n\Big)$

for any set of states $\{\rho_n\}$ and any probability distribution $\{p_n\}$.

At this point, in order to introduce the concept of distance between quantum states, we briefly recall the general definition of a distance. In general, a distance $\mathcal{D}$ is a function defined on a vector space that assigns a non-negative real number to each pair of elements and satisfies the following properties: non-negativity, $\mathcal{D}(\mathbf{x}, \mathbf{y}) \geq 0$; identity of indiscernibles, $\mathcal{D}(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$; symmetry, $\mathcal{D}(\mathbf{x}, \mathbf{y}) = \mathcal{D}(\mathbf{y}, \mathbf{x})$; and the triangle inequality, $\mathcal{D}(\mathbf{x}, \mathbf{z}) \leq \mathcal{D}(\mathbf{x}, \mathbf{y}) + \mathcal{D}(\mathbf{y}, \mathbf{z})$.

In the quantum setting, a quantum distance is a distance defined on the space of quantum states, including both pure states in Hilbert space and mixed states described by density operators, i.e., the convex hull of pure states, and quantifies their distinguishability.

Given a quantum distance $\mathcal{D}$ we can define as coherence measurement the quantity

$$C_{\mathcal{D}}(\rho) \equiv \min_{\sigma \in \mathcal{I}} \{\mathcal{D}(\rho, \sigma)\} \tag{1.5}$$

Monotonicity under trace preserving of $\mathcal{D}$ quantum operations

$$\mathcal{D}(\rho, \sigma) \geq \mathcal{D}(\epsilon(\rho), \epsilon(\sigma)) \tag{1.6}$$

is required for property 2a), indeed, if we call σ_{min} the density operator in $\mathcal{I}$

which minimizes (1.5) we have for an incoherent operation ϵ

$$\begin{aligned} C_{\mathcal{D}}(\rho) &= \mathcal{D}(\rho, \sigma_{min}) \geq \mathcal{D}(\epsilon(\rho), \epsilon(\sigma_{min})) \\ &\geq \min_{\sigma \in \mathcal{I}} \{\mathcal{D}(\epsilon(\rho), \sigma)\} = C_{\mathcal{D}}(\epsilon(\rho)) \end{aligned} \tag{1.7}$$

where in the second inequality we used that $\epsilon(\sigma^*) \in \mathcal{I}$, where σ^* is the state for which the minimum is obtained.

From the definition of distances, property 1) is automatically fulfilled: $\mathcal{D}(\rho, \sigma) = 0$ if and only if $\rho = \sigma$.

Finally if we requite joint convexity for $\mathcal{D}$ (Nielsen and Chuang 2002), it implies property 3) because

$$\begin{aligned} C_{\mathcal{D}}\Big(\sum_n p_n \rho_n\Big) &\leq \mathcal{D}\Big(\sum_n p_n \rho_n, \sum_n p_n \sigma_n^*\Big) \\ &\leq \sum_n p_n \mathcal{D}(\rho_n, \sigma_n^*) = \sum_n p_n C_{\mathcal{D}}(\rho_n) \end{aligned} \tag{1.8}$$

where the $\sigma_n^* \in \mathcal{I}$ are the density operators which minimize $\mathcal{D}(\rho_n, \sigma_n)$.

This considerations suggest that the coherence measure $\mathcal{M}_{\mathrm{Tr}}$, induced by the trace distance

$$D_{\mathrm{Tr}}(\rho_1, \rho_2) = \frac{1}{2}\|\rho_1 - \rho_2\|_1 = \frac{1}{2}\mathrm{Tr}|\rho_1 - \rho_2| = \frac{1}{2}\mathrm{Tr}\left[\sqrt{(\rho_1 - \rho_2)(\rho_1 - \rho_2)}\right], \tag{1.9}$$

is a good candidate to be a coherence quantifier since it is monotonic decreasing under quantum operations and jointly convex (Nielsen and Chuang 2002) which imply 2a) and 3).

In the same way, the geometric coherence measure $\mathcal{M}_{\mathrm{B}}$, induced by the Bures distance

$$D_{\mathrm{B}} \equiv \sqrt{2\left[1 - F(\rho, \sigma)\right]}, \tag{1.10}$$

where F is the fidelity

$$F(\rho, \sigma) = \mathrm{Tr}\left[\sqrt{\rho^{\frac{1}{2}}\sigma\rho^{\frac{1}{2}}}\right] \tag{1.11}$$

is a good candidate for a coherence monotone, but it does not satisfy 2b) (Shao et al. 2015).

Another coherence monotone is given by the relative entropy of coherence

$$C_S(\rho) \equiv \min_{\sigma \in \mathcal{I}} \{[\rho \log \rho - \rho \log \sigma]\} \tag{1.12}$$

where S is the von Neumann entropy defined by

$$S(\rho) = \mathrm{Tr}\big[\rho \log \rho\big] \tag{1.13}$$

Beside that relative entropy is not a distance, it has been shown that it fulfills 1), 2a), 2b) and 3) (Baumgratz et al. 2014). Furthermore, it has an upper limit $C_S(\rho) \leq \log d$ where the equality holds for the maximally coherent state (1.3).

We will better discuss in section 1.5 how these distances works for developing a geometric measures of Bell nonlocality.

1.2.1 Coherence as a resource

In this section we will give an example of how coherence can be used as a resource in the simplest case of generating a single-qubit Hadamard gate. We will prove that, starting with the maximally coherent state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{1.14}$$

we can obtain the unitary transformation H such that $H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $H|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. The transformation H is called Hadamard gate and it is one of the essential building of quantum algorithms (Nielsen and Chuang 2002). Its matrix representation in the standard basis $\{|0\rangle, |1\rangle\}$ is

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{1.15}$$

A scheme of the protocol is shown in figure 1.2.

To prove this statement, consider an ancillary system formed by another qubit and define the two operators

$$E_0 \equiv \frac{1}{\sqrt{2}}[|00\rangle\langle 00| + |01\rangle\langle 01| + |00\rangle|10\rangle - |01\rangle\langle 11|] \tag{1.16}$$

and

$$E_1 \equiv \frac{1}{\sqrt{2}}[|11\rangle\langle 00| + |10\rangle\langle 01| - |11\rangle|10\rangle + |10\rangle\langle 11|] \tag{1.17}$$

acting on the two qubits, where the first one is the ancilla.

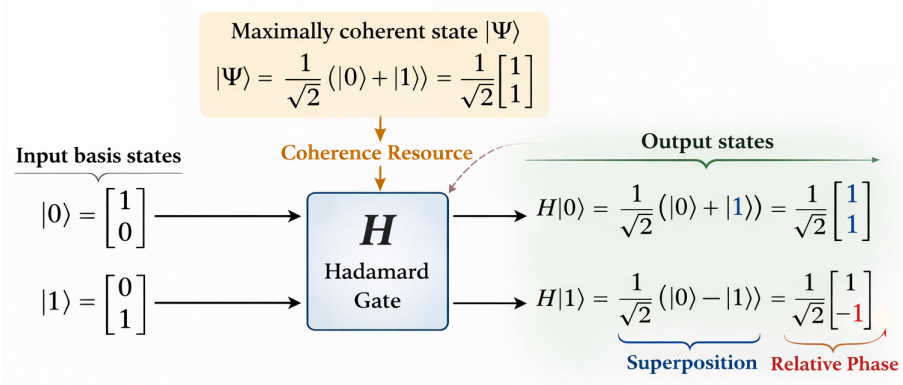


Figure 1.2: *Hadamard gate obtained by the maximally coherent state*

Since $E_0^\dagger E_0 + E_1^\dagger E_1 = I$ and $E_i \mathcal{I} E_i^\dagger \in \mathcal{I}$ for $i = 0, 1$, the $\{E_i\}$ are Kraus operators representing an incoherent operation. The last two statements can be proved by direct calculation, in particular, for the second, the Krauss operators above defined, only leads to a phase-shift, hence they do not create coherence.

Consider $|\Phi\rangle = c_0 |0\rangle + c_1 |1\rangle$, $|\Psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and define

$$|\chi\rangle = |\Phi\rangle \otimes |\Psi\rangle = \frac{1}{\sqrt{2}}(c_0 |00\rangle + c_0 |01\rangle + c_1 |10\rangle + c_1 |11\rangle) \tag{1.18}$$

Hence

$$\begin{aligned} E_0 |\chi\rangle &= \frac{1}{\sqrt{2}}[c_0(|00\rangle + |01\rangle) + c_1(|00\rangle - |01\rangle)] \\ &= \frac{1}{\sqrt{2}} |0\rangle \otimes [c_0 H |0\rangle + c_1 H |1\rangle] = \frac{1}{\sqrt{2}} |0\rangle \otimes H |\Phi\rangle \end{aligned} \tag{1.19}$$

In the same way we obtain

$$E_1 |\chi\rangle = \frac{1}{\sqrt{2}} |1\rangle \otimes H |\Phi\rangle \tag{1.20}$$

and,

$$p_0 = \mathrm{Tr}\left[E_0 |\chi\rangle \langle\chi| E_0^\dagger\right] = \frac{1}{2}, \qquad p_1 = \mathrm{Tr}\left[E_1 |\chi\rangle \langle\chi| E_1^\dagger\right] = \frac{1}{2} \tag{1.21}$$

Therefore, if the quantum operation is sub-selective then we can obtain the two states

$$\rho_0 = \frac{E_0 |\chi\rangle \langle\chi| E_0^\dagger}{p_0} = |0\rangle \langle 0| \otimes H |\Phi\rangle \langle\Phi| H \tag{1.22}$$

or

$$\rho_1 = \frac{E_1 |\chi\rangle \langle\chi| E_1^\dagger}{p_1} = |1\rangle \langle 1| \otimes H |\Phi\rangle \langle\Phi| H \tag{1.23}$$

In both cases, tracing out $\mathcal{S}_1$ we obtain

$$\rho = H |\Phi\rangle \langle\Phi| H \tag{1.24}$$

with probability 1.

Similarly if the quantum operation is not sub-selective, we have

$$\epsilon(\rho) = E_0 |\chi\rangle \langle\chi| E_0^\dagger + E_1 |\chi\rangle \langle\chi| E_1^\dagger = \frac{1}{2} I \otimes H |\Phi\rangle \langle\Phi| H = p_0\rho_0 + p_1\rho_1 \tag{1.25}$$

and tracing out the first qubit, again we obtain (1.24) again.

Note that with the incoherent operation defined by $\{E_i\}$, the coherence in the second subsystem is consumed in order to obtain the Hadamard gate, in this sense we say that coherence is used as a resource.

1.3 Entanglement as a resource

Entanglement is one of the most profound expression of quantum mechanics. It has been studied since the works of Einstein, Podolsky, and Rosen (Einstein et al. 1935) and Schrödinger (Schrödinger 1935), who analyzed the peculiarities of composite quantum systems. The study of entanglement is not only fundamental for deepening our understanding of the foundations of quantum theory but also crucial for numerous practical applications, such as quantum teleportation (Bennett et al. 1993), metrology (Giovannetti et al. 2011), cryptography (Ekert 1991), and more.

In analogy to the section 1.2, here we will search for some criteria for defining a monotone to quantify entanglement. We will focus only on the simpler bipartite case, leaving the multipartite case for the next sections.

We say that a state ρ is (bi)separable, when it can be written in the form

$$\rho = \sum_i p_i \rho^{(A)} \otimes \rho^{(B)} \tag{1.26}$$

where $\{p_i\}$ is a probability distribution and $\rho^{(A)}$ ($\rho^{(B)}$) is a state describing subsystem A (B) alone.

We say that E is an entanglement monotone iff E fulfills the following properties (Plenio and Virmani 2007):

a) $E(\rho) \geq 0$ with $E(\rho) = 0$ iff ρ is separable.

b) Invariance under local unitary operations that is

$$E(\rho) = E(U_A \otimes U_B E(\rho) U_A^\dagger \otimes U_B^\dagger)$$

for any ρ.

c) The entanglement cannot increase under a generic local operation and classical communication (LOCC) ε (free operations), that is

$$E[\varepsilon(\rho)] \leq E(\rho)$$

which is valid for a generic bipartite system. The property a) is quite intuitive since it says that Entanglement occurs iff the state cannot be separated.

The property b) is introduced because local unitary operations are only a change of basis in the Hilbert space of the local subsystem and hence the entanglement cannot be affected by these transformations. The last property c) is related to the intuitive fact that entanglement cannot increase with only local operations and classical communications, since generating entanglement requires an interaction between the two subsystems. Therefore the LOCC are identified as the free operations which do not increase the (entanglement) resource.

If we call $\mathcal{S}$ the set of all separable density operators, in analogy with the coherence described in the previous section, a good entanglement monotone is given by (Vedral et al. 1997)

$$E(\rho) \equiv \min_{\sigma \in \mathcal{S}} D(\rho, \sigma) \tag{1.27}$$

where D is a generic measure of distance. In the same way of previous section, by using the properties of the trace distance (Nielsen and Chuang 2002), we can prove that (1.27) is an entanglement monotone. Or again, we can consider the relative entropy 1.12 also for the entanglement.

For two-qubit systems, one widely used entanglement monotone is the *concurrence* $\mathfrak{C}$ (Hill and Wootters 1997; Wootters 1998), which for a generic two-qubit density operator ρ is defined by

$$\mathfrak{C}(\rho) \equiv \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\} \tag{1.28}$$

where the λ_i are the square roots of the eigenvalues of the operator $\rho\tilde{\rho}$ in decreasing order, with

$$\tilde{\rho} = (\sigma_y \otimes \sigma_y)\rho^*(\sigma_y \otimes \sigma_y). \tag{1.29}$$

Concurrence is closely related to the Entanglement of Formation (EoF). For a generic density operator $\rho = \sum_i p_i \left|\Psi_i\right\rangle \left\langle\Psi_i\right|$, the EoF is defined as

$$\mathrm{EoF}(\rho) = \min \sum_i p_i S(\left|\Psi_i\right\rangle \left\langle\Psi_i\right|), \tag{1.30}$$

where S is the von Neumann entropy, and the minimum is taken over all possible decompositions of ρ as a mixture of pure states (Wootters 1998). In the case ρ is pure, (1.30) coincides with the von Neumann entropy of ρ.

For two-qubit states, concurrence and EoF are related by (Wootters 1998)

$$\mathrm{EoF} = h\left(\frac{1+\sqrt{1-\mathfrak{C}^2}}{2}\right), \tag{1.31}$$

where $h(x)$ is the Shannon entropy

$$h(x) = -x\log_2(x) - (1-x)\log_2(1-x). \tag{1.32}$$

An alternative entanglement monotone is the *negativity* (Peres 1996; Vidal and Werner 2002; Horodecki et al. 2009), given by

$$\mathcal{N}(\rho) \equiv \frac{\|\rho^{T_i}\|_1 - 1}{2}, \quad i = 1, 2. \tag{1.33}$$

where ρ^{T_i} is the partial transpose matrix of ρ respect to the subsystem i. The negativity is a quantifier of entanglement only in the case of 2×2 and 2×3 systems (Horodecki 1997), since property 1) is not always fulfilled for higher dimensions.

As we will better discuss in section 1.3.2, the set of separable states is not fully characterized in general, since there exists no computable criterion for classifying whether a given state ρ is entangled or not.

1.3.1 Quantum teleportation

One of the first applications of entanglement as a resource is *quantum teleportation*, theorized in 1993 by Bennett *et al.* (Bennett et al. 1993) and experimentally demonstrated in 1997 (Boschi et al. 1998; Bouwmeester et al. 1997).

Suppose that Alice possesses a two-single-qubit state $\left|\Psi\right\rangle = a\left|0\right\rangle + b\left|1\right\rangle$, where the coefficients a and b are unknown to her.

Quantum teleportation, schematized in Fig. 1.3 is a procedure that allows this quantum state to be transferred to Bob, who is located in a distant laboratory B.

As we will see, the teleportation protocol must satisfy the following con-

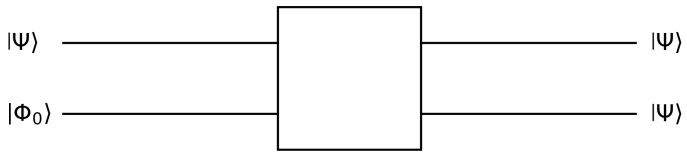


Figure 1.3: *Scheme of quantum teleportation protocol.*

straints:

- No physical matter is transmitted from A to B; only information is exchanged between the two laboratories.
- The teleportation process is not instantaneous, and therefore causality is preserved.
- Once the state has been teleported from A to B, the original state in A is destroyed, in accordance with the no-cloning theorem (Nielsen and Chuang 2002).
- Entanglement is a necessary resource to achieve error-free teleportation.

A first possible approach to the problem might be to determine the two unknown coefficients a and b. However, this is not feasible with only a single copy of the system. If Alice performs a measurement on her qubit, the state will collapse to one of the eigenstates $|0\rangle$ or $|1\rangle$, providing no further information about a and b. To estimate these coefficients, Alice would need a set of N copies of $|\Psi\rangle$, with N large enough to perform repeated measurements and determine a and b statistically. Nevertheless, such copies cannot be produced due to the *no-cloning theorem*, which states that it is impossible to create an independent and identical copy of an unknown quantum state, as shown in the appendix A.3.

Hence, we need a different procedure to teleport the state $|\Psi\rangle$ from A to B. We will show that this goal can be achieved by using the maximally entangled Bell state $|\Phi_0\rangle$, defined as

$$|\Phi_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle) \tag{1.34}$$

Let us suppose that we prepare a two qubit Bell state $|\Phi_0\rangle$, with the first sent to Alice and the second to Bob. The total system then consists of Alice's state $|\Psi\rangle$ (which is to be teleported) and the two entangled states $|\Phi_0\rangle$. The overall

state is therefore represented by a vector in $\mathcal{H}^2 \otimes \mathcal{H}^2 \otimes \mathcal{H}^2$ and can be written as

$$\begin{aligned}|\Psi\rangle \otimes |\Phi_0\rangle &\equiv (a\,|0\rangle_{A_2} + b\,|1\rangle_{A_2}) \otimes \frac{1}{\sqrt{2}}(|0\rangle_{A_1} \otimes |0\rangle_B + |1\rangle_{A_1} \otimes |1\rangle_B) \\ &= \frac{1}{\sqrt{2}}\Big(a\,|0\rangle_{A_2}\,|0\rangle_{A_1}\,|0\rangle_B + a\,|0\rangle_{A_2}\,|1\rangle_{A_1}\,|1\rangle_B \\ &\quad + b\,|1\rangle_{A_2}\,|0\rangle_{A_1}\,|0\rangle_B + b\,|1\rangle_{A_2}\,|1\rangle_{A_1}\,|1\rangle_B\Big).\end{aligned} \tag{1.35}$$

where, for simplicity of notation, the tensor products have been omitted in the last expression. The indices A_1, B, and A_2 have been introduced to label the three different Hilbert spaces of the total system, where A_1 and A_2 correspond to Alice's states and B to Bob's state.

By using the equations from (A.3) to express Alice's state in terms of the Bell states, we obtain

$$\begin{aligned}|\Psi\rangle \otimes |\Phi_0\rangle =& \frac{1}{2}\Big[\,|\Phi_0\rangle_{A_1A_2}\,(a\,|0\rangle_B + b\,|1\rangle_B) + |\Phi_1\rangle_{A_1A_2}\,(a\,|1\rangle_B + b\,|0\rangle_B) \\ &+ |\Phi_2\rangle_{A_1A_2}\,(a\,|1\rangle_B - b\,|0\rangle_B) + |\Phi_3\rangle_{A_1A_2}\,(a\,|0\rangle_B - b\,|1\rangle_B)\Big]\end{aligned} \tag{1.36}$$

Therefore, Alice has a probability of $\frac{1}{4}$ of measuring each of the four Bell states.

Since the Pauli matrices σ_i satisfy the relation $\sigma_i\sigma_i^\dagger = I$, they can be regarded as unitary transformations that Bob can apply to his state.

Now, if Alice performs a measurement on her two-qubit state and obtains the Bell state $|\Phi_0\rangle_{A_1A_2}$, Bob's state automatically collapses into

$$|\Psi\rangle = a\,|0\rangle_B + b\,|1\rangle_B \tag{1.37}$$

and the teleportation has been successful.

If Alice obtains $|\Phi_1\rangle$, then Bob's state will collapse in

$$|\Psi_1\rangle = a\,|1\rangle_B + b\,|0\rangle_B \tag{1.38}$$

but he does not know it.

Nevertheless, if Alice communicates (with classical communication) to Bob the result of her own measurement, he can apply the unitary transformation σ_1

to the state and he obtains

$$\sigma_1 |\Psi_1\rangle = a |0\rangle_B + b |1\rangle_B = |\Psi\rangle \tag{1.39}$$

Therefore, also in this case the teleportation has been accomplished.

In the same way it can be proved that for the other two cases Bob can obtain the state $|\Psi\rangle$ by the application of σ_2 and σ_3.

This is a simple example of how *local operations* (LO), such as Alice's measurements and Bob's unitary transformations, together with *classical communication* (CC), can be used to implement protocols between different laboratories. Moreover, we have demonstrated the usefulness of entanglement as a resource. In fact, to perform teleportation, we started with two entangled states located at A and B, since it is not possible to teleport a state using LOCC alone.

As previously mentioned, there is no transfer of matter from A to B. Causality is also preserved, since Alice must communicate the result of her measurement to Bob, which requires at least the time it takes for light to travel from A to B. Furthermore, once the state $|\Psi\rangle$ is created in Bob's laboratory, Alice state in A is changed, as it collapses after Alice's measurement. Therefore, this process does not contradict the no-cloning theorem.

It is also worth noting that, at the end of the experiment, neither Alice nor Bob know the coefficients a and b of the state $|\Psi\rangle$. Hence, the state is successfully teleported from A to B while remaining completely unknown.

1.3.2 Entanglement qualification and quantification

We have already highlighted the importance of quantum features as resources and the challenges in quantifying these quantities. In the present and in the following sections, focusing on entanglement, we will discuss some of the well-known (computable and non-computable) criteria for its detection and the related measures for its quantification. We will show that, despite the significant efforts made and the remarkable results achieved so far, the problem of determining whether a given quantum state ρ is entangled or not remains unsolved.

After that, building on the criterion introduced in (Hofmann and Takeuchi 2003), we will demonstrate that a necessary condition for separability can be expressed as an inequality involving only the purity of the state ρ and the purities of its subsystems. When this inequality is violated, the state is guaranteed to be entangled. Our approach establishes a duality between the violation of local uncertainty relations for SU(N) operators and the purities of the state and its subsystems, offering a valuable tool for entanglement detection. The relation between entanglement and purity has been investigated also in continuous

variable systems (Adesso et al. 2004). As we will show, besides its conceptual relevance, the derived formulation provides a very useful framework for the experimental detection of entanglement.

We further analyze the effectiveness of this criterion both numerically and analytically, demonstrating that its performance improves as the amount of entanglement increases.

As we will show, this new formulation also provides experimentally friendly tools for entanglement detection, particularly relevant in higher-dimensional Hilbert spaces, where the number of quantities that must be measured to detect entanglement is exponentially reduced.

For the specific case of two-qubit states, we explore the connection between this criterion and the violation of CHSH inequalities. Additionally, we provide a geometric interpretation of the criterion for Bell-diagonal states, highlighting its conceptual and practical implications.

Finally, we relate our result to the uncertainty relations, in particular providing that for the Heisenberg-Robertson relations, the product of the observable inequality is saturated with an additional term related to the linear entropy.

1.3.3 Entanglement criteria

A quantum state $\rho \in \mathcal{D}(\mathcal{H})$, with Hilbert space $\mathcal{H} = \mathcal{H}^{(A_1)} \otimes \cdots \otimes \mathcal{H}^{(A_n)}$, is (fully) separable when there exists a set of states $\{\rho_i^{(A_l)} \in D(\mathcal{H}^{(A_l)}) \mid l \in \{1, \ldots, n\}\ ,\ i = 1, \ldots, P\}$ for some integer P, and a probability distribution $\{p_i\}$, such that

$$\rho = \sum_i p_i \rho_i^{(A_1)} \otimes \cdots \otimes \rho_i^{(A_n)} \,. \tag{1.40}$$

If such a combination does not exist, we say that the state is entangled. It can happen that a state is not fully separable but only k-separable for a certain k-partition, with $k < n$, i.e, when for the fixed k-partition, the state ρ can be factorized as

$$\rho = \sum_i p_i \rho_i^{(A_1)} \otimes \cdots \otimes \rho_i^{(A_k)} \,, \tag{1.41}$$

with $\rho^{(A_l)} \in \mathcal{H}^{(A_l)} \equiv \bigotimes_{s \in A_l} \mathcal{H}^{(s)}$, for $l = 1, \ldots, k$ and $\{A_l\}_{l=1}^k$ is a k-partition of the index set $I = \{I_1, \ldots, I_n\}$, this is $A_l \subset I$, $\bigcup_{l=1}^k A_l = I$, $A_l \cap A_m = \emptyset$ if $l \neq m$. Clearly, if a state is fully separable it is also k-separable for any k-partition.

Pure states

When a multipartite quantum state is pure, determining whether it is entangled or not is straightforward. Since a pure state is a projector, its purity is equal to 1:

$$\mathcal{P} \equiv \mathrm{Tr}\left[|\Psi\rangle\langle\Psi|^2\right] = \mathrm{Tr}\left[|\Psi\rangle\langle\Psi|\right] = 1. \tag{1.42}$$

Therefore, using the definition in Eq. (1.41), a pure quantum state is not k-entangled if and only if it can be written as

$$|\Psi\rangle\langle\Psi| = |\Psi\rangle^{(I_1)}\langle\Psi| \otimes \cdots \otimes |\Psi\rangle^{(I_k)}\langle\Psi|, \tag{1.43}$$

which is called a k-product state. Thus, a pure state ρ is a k-product state if and only if each subsystem i is in a pure state, which means

$$\mathcal{P}^{(I_i)} = \mathrm{Tr}\left[\left(\rho^{(I_i)}\right)^2\right] = 1, \tag{1.44}$$

where $\rho^{(I_i)} = \mathrm{Tr}_{\overline{I}_i}[\rho]$, with $\overline{I}_i$ denoting the complement of I_i, for any $i = 1, \ldots, k$.

The purity of a given state equals 1 only when there is maximum information about the system, meaning there are no probabilistic classical correlations. This corresponds to the minimum von Neumann entropy, $S_{vn} = 0$.

This fact can be interpreted as follows: if one has maximum information about the global system, the only case in which the state is k-separable for a certain partition is when one has maximum information about all its k subsystems. This implies that complete knowledge of the whole system can be obtained by separately analyzing its subsystems. While this is evident for classical systems, it is not true for quantum entangled systems. In the presence of entanglement, examining the subsystems individually results in a loss of information due to quantum correlations, meaning that observing all subsystems separately is not equivalent to observing the system as a whole.

According to these facts, one can define the linear entropy

$$\mathcal{S}_{L_i} = N_{I_i}\mathcal{P}^{(I_i)} - 1 \tag{1.45}$$

as a measure of bipartite entanglement between the subsystem I_i and the rest of the system, where N_{I_i} is the dimension of $\mathcal{H}^{(I_i)}$. We always have $0 \leq \mathcal{S}_{L_i} \leq 1$, where the minimum is reached for $\rho^{(I_i)}$ pure and the maximum for $\rho^{(I_i)}$ maximally mixed in $\mathcal{H}^{(I_i)}$.

Alternatively, one can use the von Neumann entropy of the given subsystem I_i, instead of the linear entropy, as a measure of entanglement between I_i and

its complement.

These quantities are not good quantifiers of entanglement in the case of (global) mixed states, since they take into account all forms of correlations, both classical and quantum.

Positive maps criteria

The characterization of k-separability has been established (Horodecki et al. 2001). Before introducing this result, we define that, for a fixed n-partition of a quantum system, an operator A is positive if and only if

$$\mathrm{Tr}\Big[A\left(P_{I_1}\otimes P_{I_2}\otimes\cdots\otimes P_{I_n}\right)\Big]\geq 0 \tag{1.46}$$

for all product projections $P_{I_1}\otimes P_{I_2}\otimes\cdots\otimes P_{I_n}$, where each P_{I_i} is a projector on $\mathcal{H}^{(I_i)}$.

Theorem 1: A state ρ defined on $\mathcal{H}=\mathcal{H}^{(I_1)}\otimes\mathcal{H}^{(I_2)}\otimes\cdots\otimes\mathcal{H}^{(I_n)}$ is separable if and only if $\mathrm{Tr}\left[A\rho\right]\geq 0$ holds for all positive observables A on $\mathcal{H}$.

Such observables are A are also called entanglement witnesses (Horodecki et al. 1996; Terhal 2000; Bertlmann and Krammer 2008; Bertlmann and Friis 2023).

An equivalent form of the previous result, based on *linear maps positive on product states (LMPP)*, is the following

Theorem 2: A state ρ defined on $\mathcal{H}=\mathcal{H}^{(I_1)}\otimes\mathcal{H}^{(I_2)}\otimes\cdots\otimes\mathcal{H}^{(I_n)}$ is separable if and only if

$$I\otimes\tilde{\Lambda}(\rho)\geq 0 \tag{1.47}$$

for all LMPP $\tilde{\Lambda}$, i.e., all linear maps $\tilde{\Lambda}:B(\mathcal{H}^{(I_2)}\otimes\cdots\otimes\mathcal{H}^{(I_n)})\to B(\mathcal{H}^{(I_1)})$ such that $\tilde{\Lambda}(P_{I_2}\otimes\cdots\otimes P_{I_n})\geq 0$ for all product projections $P_{I_2}\otimes\cdots\otimes P_{I_n}$, where each P_{I_i} is a projector on $\mathcal{H}^{(I_i)}$ and $B(\mathcal{H}^{(I_i)})$ is the set of all bound operators acting on $\mathcal{H}^{(I_i)}$.

Both results are in general non-computable.

Positive partial transpose criterion and negativity

The previous criterion on positive maps is related to another computable one, known as the positive partial transpose (PPT) criterion, which provides a necessary and sufficient condition for k-separability in bipartite systems with Hilbert space dimensions 2×2, 2×3 (and 3×2). In all other cases, it remains only a necessary condition.

The PPT criterion for bipartite systems was introduced by Peres (Peres 1996),

who defined the partially transposed density operator as

$$\rho^{T_B} \equiv (\mathbb{1}_2 \otimes T_B)\, \rho, \tag{1.48}$$

where the operator $\mathbb{1}_2 \otimes T_B$ applies the transposition only to the second subsystem.

In particular, the result found by Peres states that for a given quantum state ρ, if ρ^{T_B} is not positive semi-definite, then ρ is entangled.

The proof follows immediately from Theorem 2 in the previous section, since the partial transpose is a positive map.

Subsequently, it has been established that this condition is also sufficient for detecting entanglement in the case of 2×2 and 2×3 systems (Horodecki 1997). To outline the proof, we first introduce the concept of completely positive (CP) maps: a map $\Lambda : B(\mathcal{H}^A) \to B(\mathcal{H}^B)$ is completely positive if the extended maps

$$\Lambda_n = \Lambda \otimes \mathbb{1}_n : B(\mathcal{H}^A) \otimes \mathcal{M}_n \to B(\mathcal{H}^B) \otimes \mathcal{M}_n \tag{1.49}$$

are positive for any integer n, where $\mathcal{M}_n$ denotes the set of complex $n \times n$ matrices. Clearly, a completely positive map is also a positive map, but the converse is generally not true—this distinction is crucial for the proof.

The proof is completed by noting that any positive map Λ from $B(\mathcal{H}^B)$ to $B(\mathcal{H}^A)$ in 2×2 and 2×3 systems can be expressed as

$$\Lambda = \Lambda_1 + T^2 \Lambda_2, \tag{1.50}$$

where Λ_1 and Λ_2 are completely positive maps and T is the partial transposition operator.

The associated entanglement measure to the PPT criterion is the *negativity*, defined by

$$\mathcal{N}(\rho) = \frac{\|\rho^{T_B}\|_1 - 1}{2} = \left| \sum_i \lambda_i^{(-)} \right| \tag{1.51}$$

where the $\lambda_i^{(-)}$ are the negative eigenvalues of ρ^{T_B} and $\|\cdot\|_1$ is the trace-norm or $1-$norm.

Note that the negativity is the same if we transpose A instead of B, since the trace norm is invariant for (total) transposition:

$$\|\rho^{T_B}\|_1 = \| \left(\rho^{T_B}\right)^T \|_1 = \|\rho^{T_A}\|_1 \,. \tag{1.52}$$

Finally, we mention that the states that are entangled but do not violate the PPT criterion are called *bound entangled states*, for their close relation

with entanglement distillation (Horodecki et al. 1998). In particular, it has been shown that if by considering an arbitrary large number of copies of ρ it is possible to distillate Bell pairs, then ρ must violate the PPT criterion.

Reduction criterion

Another computable criterion was introduced by Horodecki, known as the *reduction criterion* (Horodecki and Horodecki 1999). This criterion provides a sufficient condition for entanglement and is formulated as follows:

Theorem: Given a bipartite state ρ, if

$$\rho - \mathbb{1}_2 \otimes \rho_B > 0 \tag{1.53}$$

or

$$\rho - \rho_A \otimes \mathbb{1}_2 > 0, \tag{1.54}$$

then the state ρ is entangled.

In the same work, it was shown that this criterion is equivalent to the PPT criterion in the case of 2×2 and 2×3 systems, but it becomes weaker in higher dimensions.

The reduction criterion is also related to entanglement distillation (Bennett et al. 1996b; Bennett et al. 1996a). Specifically, if the entanglement of a state is distillable, then the state necessarily violates at least one of the two inequalities of the criterion.

Majorization criterion and entropic criteria

The majorization criterion for bipartite entanglement was introduced by Nielsen and Kempe (Nielsen and Kempe 2001) and is based on majorization theory (Marshall et al. 1979; Horn and Johnson 2012; Bhatia 2013).

In order to formulate the criterion, let us introduce some preliminary concepts. In particular, given two d-dimensional vectors $\mathbf{x}$ and $\mathbf{y}$ whose elements x_i and y_i are arranged in decreasing order, we say that $\mathbf{y}$ majorizes $\mathbf{x}$, denoted as $\mathbf{x} \prec \mathbf{y}$, if

$$\sum_{i=1}^{k} x_i \leq \sum_{i=1}^{k} y_i, \tag{1.55}$$

for any $k = 1, \ldots, d$, with equality holding for $k = d$.

Two important characterizations of majorization are the following:

Theorem 1: The majorization $\mathbf{x} \prec \mathbf{y}$ holds if and only if there exists a

doubly stochastic matrix D such that $\mathbf{x} = D\mathbf{y}$.

Theorem 2: The majorization $\mathbf{x} \prec \mathbf{y}$ holds if and only if, for any convex function $f : \mathbb{R}^d \to \mathbb{R}^d$, the following inequality holds:

$$\mathrm{Tr}[f(\mathbf{x})] \leq \mathrm{Tr}[f(\mathbf{y})]. \tag{1.56}$$

We also denote the spectrum of a density operator ρ by $\lambda(\rho)$.

At this point, we are ready to introduce the main majorization criterion for bipartite entanglement.

Theorem 3: If a state $\rho^{(AB)}$ composed of two subsystems A and B is separable, then the following relations hold:

$$\lambda(\rho^{(AB)}) \prec \lambda(\rho^{(A)}), \qquad \lambda(\rho^{(AB)}) \prec \lambda(\rho^{(B)}), \tag{1.57}$$

where, by convention, the last elements of $\lambda(\rho^{(A)})$ and $\lambda(\rho^{(B)})$ are set to zero to match the dimension of $\lambda(\rho^{(AB)})$.

The criterion is not sufficient for separability, indeed, if we consider the two-qubit state

$$\rho = p\,|00\rangle\langle 00| + (1-p)\,|\Psi\rangle\langle\Psi| \tag{1.58}$$

where $|\Psi\rangle = \frac{|00\rangle+|11\rangle}{2}$ is a Bell state, by using the majorization criterion we obtain that $\lambda(\rho^{(AB)})$ majorizes $\lambda(\rho^{(A)})$ and $\lambda(\rho^{(B)})$ for $p \geq \frac{1}{3}$, while the state is entangled for $p > 0$. Therefore, there are entangled states not detected by the majorization criterion.

It is worth mentioning that the majorization criterion is not weaker than other criteria based on entropic quantities. In particular, there exist two classes of entropic criteria based on conditional Rényi entropies and conditional Tsallis entropies (Vollbrecht and Wolf 2002). The conditional α-Rényi entropies of a given state ρ are defined as

$$S_\alpha(B|A, \rho^{(AB)}) = S_\alpha(\rho^{(AB)}) - S_\alpha(\rho^{(A)}), \tag{1.59}$$

where

$$S_\alpha(\rho) \equiv \frac{\log\{\mathrm{Tr}[\rho^\alpha]\}}{1-\alpha} \tag{1.60}$$

for any $\alpha \in \mathbb{R}$, known as the α-Rényi entropy.

Conversely, the α-Tsallis entropy of a state ρ is defined as

$$T_\alpha(\rho) = \frac{1 - \mathrm{Tr}[\rho^\alpha]}{1-\alpha}, \tag{1.61}$$

and the associated conditional α-Tsallis entropy is given by

$$T_\alpha(B|A,\rho^{(AB)}) = \frac{\mathrm{Tr}\left[\left(\rho^{(A)}\right)^\alpha\right] - \mathrm{Tr}\left[\left(\rho^{(AB)}\right)^\alpha\right]}{(1-\alpha)\mathrm{Tr}\left[\left(\rho^{(A)}\right)^\alpha\right]}. \tag{1.62}$$

Theorem 4: The conditional Rényi entropies and the conditional Tsallis entropies are non-negative if $\rho^{(AB)}$ is a separable state.

It is immediate that the last statement follows as a corollary of Theorems 2 and 3.

Finally, we note that all the criteria discussed in this section are related to the spectrum of the density matrix $\rho^{(AB)}$ and those of its subsystems. In fact, it is possible to show that any criterion for entanglement detection based solely on spectral properties can never provide a necessary and sufficient condition. Indeed, consider the two states

$$\rho_1 = \frac{1}{3}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \rho_2 = \begin{pmatrix} \frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{2}{3} \end{pmatrix} \tag{1.63}$$

which have the same spectra, as well as the same spectra for their reduced density matrices. However, while the first state is entangled, the second one is not.

Cross-norm and computable cross norm

Another characterization of separability was provided by Rudolph (Rudolph 2000). The statement is general, but for simplicity, we will consider only the bipartite case. Let us denote by $B(\mathcal{H}^{(A)})$ and $B(\mathcal{H}^{(B)})$ the sets of bounded operators acting on $\mathcal{H}^{(A)}$ and $\mathcal{H}^{(B)}$, respectively, and by $\|\cdot\|_1$ the trace norm.

Furthermore, let us define the *cross norm* of a generic operator $t \in B(\mathcal{H}^{(A)} \otimes \mathcal{H}^{(B)})$ as

$$\|t\|_\gamma \equiv \inf\left\{ \sum_{i=1}^n \|u_i\|_1 \|v_i\|_1 \;\middle|\; t = \sum_{i=1}^n u_i \otimes v_i \right\}, \tag{1.64}$$

where u_i and v_i are operators that act on $\mathcal{H}^{(A)}$ and $\mathcal{H}^{(B)}$ respectively, and the infimum is taken over all possible tensor product decompositions of t.

The characterization of separability can be stated as follows:

Theorem 1: A density operator ρ is separable if and only if $\|\rho\|_\gamma = 1$.

Such a (non-computable) criterion is called the *cross norm (CN) criterion*.

Related to the CN criterion, there is another computable one, also discovered by Rudolph (Rudolph 2003; Rudolph 2005), called the *computable cross norm (CCN) criterion*.

Let us define two orthonormal (with respect to the Hilbert-Schmidt product) bases $\{|\tilde{\sigma}_i^A\rangle\}_i$ and $\{|\tilde{\sigma}_i^B\rangle\}_i$ of $B(\mathcal{H}^{(A)})$ and $B(\mathcal{H}^{(B)})$, respectively. Then, ρ can be expanded as

$$\rho = \sum_{i,j} \tilde{t}_{ij} \tilde{\sigma}_i^A \otimes \tilde{\sigma}_j^B, \tag{1.65}$$

and we can consider the associated operator

$$\widetilde{T} = \sum_{i,j} \tilde{t}_{ij} \left|\tilde{\sigma}_i^A\right\rangle \left\langle\tilde{\sigma}_j^B\right|. \tag{1.66}$$

An example of such a decomposition is given in Eq. (1.79), with the associated operator $\widetilde{T}(\rho) = T$.

It is possible to show the following result (Rudolph 2003):

$$\left|\left|\widetilde{T}(\rho)\right|\right|_1 = \min\left\{ \sum_i \|u_i\|_2 \|v_i\|_2 \;\middle|\; \rho = \sum_i u_i \otimes v_i \right\}, \tag{1.67}$$

and this result is independent of the particular decomposition (1.65). The CCN criterion reads as follows:

Theorem 2: If $\left|\left|\widetilde{T}(\rho)\right|\right|_1 \geq 1$, then the state ρ is entangled.

Note that since $\|\cdot\|_1 \geq \|\cdot\|_2$, we have $\|\rho\|_\gamma \geq \left|\left|\widetilde{T}(\rho)\right|\right|_1$, meaning the computable cross norm is not stronger than the cross norm criterion. In fact, it is possible to show that it is weaker.

Indeed, consider the state

$$\begin{aligned}\rho = \frac{1}{4}\Big(& \mathbb{1}_2 \otimes \mathbb{1}_2 + s(\mathbb{1}_2 \otimes \sigma_3) + r(\sigma_3 \otimes \mathbb{1}_2) + t(\sigma_1 \otimes \sigma_1) \\ & - t(\sigma_2 \otimes \sigma_2) + (1 + r - s)(\sigma_3 \otimes \sigma_3)\Big).\end{aligned} \tag{1.68}$$

This state is separable for $t = 0$, but it can be verified by direct calculation that

$$T = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & r \\ 0 & t & 0 & 0 \\ 0 & 0 & t & 0 \\ s & 0 & 0 & 1 + r - s \end{pmatrix} \tag{1.69}$$

is such that $\|T\|_1 \leq 1$ even for some $t \neq 0$.

Other properties of the CCN criterion are discussed in (Rudolph 2005). For example, the CCN criterion provides a necessary and sufficient condition for Bell-diagonal states and is also connected to the concept of entanglement of distillation. Furthermore, the CCN criterion is not weaker than the PPT, majorization, and reduction criteria.

Local uncertainty relations and correlation matrix-based criteria

Other approaches to entanglement detection are based on violations of certainty inequalities derived from uncertainty relations (Hofmann and Takeuchi 2003; Gühne 2004) and correlation matrices (Gühne et al. 2007).

Specifically, in (Hofmann and Takeuchi 2003) it was shown that, for a bipartite system, if one considers sets of local non-commuting observables $\{A_i\}$ and $\{B_i\}$ with uncertainty relations

$$\sum_i \Delta^2 A_i \geq L_A, \qquad \sum_i \Delta^2 B_i \geq L_B, \tag{1.70}$$

then for any separable state it must hold that

$$\sum_i \Delta^2 (A_i + B_i) \geq L_A + L_B. \tag{1.71}$$

Therefore, the violation of the latter inequality implies that the state is entangled.

Later, it has been shown that such criteria based on local uncertainty, can be reformulated in a more general picture in terms of the correlation tensor (Gühne et al. 2007), also providing that are stronger than the computable cross norm criterion (Gühne et al. 2006).

Different approaches lead to different criteria based on the correlation tensor. For example, in (Badziag et al. 2008) it has been shown that for n qubits, given the Pauli decomposition of the density operator

$$\rho = \frac{1}{2^n} \sum_{\alpha_1,\ldots,\alpha_n=0}^{3} t_{\alpha_1 \ldots \alpha_n} \sigma_{\alpha_1} \otimes \cdots \otimes \sigma_{\alpha_n} \,, \tag{1.72}$$

where $\sigma_0 = \mathbb{1}_2$ and σ_i are the Pauli matrices, then for any separable state it holds

$$\|t\| \leq t_{\text{max}} \tag{1.73}$$

where, $\|\cdot\|$ is the Hilbert-Schmidt norm and t_{max} is the highest generalized Schmidt coefficient of the tensor t (Peres 2002; Carteret et al. 2000). It is not always possible to compute t_{max}, nevertheless, since $t_{\text{max}} < 1$, a weaker for of

the criterion (1.73) is given by the inequality

$$\|t\| \leq 1\,, \tag{1.74}$$

which must hold for any separable state.

In (De Vicente and Huber 2011) it has been shown that violations of inequalities of the form

$$\|t\| \leq f(N)\,, \tag{1.75}$$

imply genuine multipartite entanglement. In this inequality, $f(N)$ is a function that depends only on the dimension of the system and its subsystems. For instance, in the case of three qudits with the same local dimension N_i, one has

$$\|t\| \leq \sqrt{\frac{8(N_i-1)(N_i^2-1)}{N_i^3}}\,. \tag{1.76}$$

Most of these criteria have been reviewed in various works (Horodecki et al. 2009; Gühne and Tóth 2009).

1.4 Uncertainty relations, purity and the k-separability problem

Let us now consider the generators of the SU(N_{A_l}) algebras, with $l = 1, \ldots, k$. Denoting by $\{\sigma_{\alpha_l}^{(l)}\}_{\alpha_l=0}^{D_{A_l}}$ the matrix representations of the generators of SU(N_{A_l}), i.e., the $N_{A_l} \times N_{A_l}$ matrices acting on $\mathcal{H}^{(A_l)}$, where $\sigma_0^{(l)}$ is the identity matrix and $D_{A_l} = N_{A_l}^2 - 1$, we can introduce the local collective observables $\sigma_{\alpha_1,\ldots,\alpha_k} \equiv \sigma_{\alpha_1}^{(1)} \otimes \cdots \otimes \sigma_{\alpha_k}^{(k)}$, their uncertainties $\Delta^2\sigma_{\alpha_1,\ldots,\alpha_k} \equiv \langle \sigma_{\alpha_1,\ldots,\alpha_k}^2 \rangle_\rho - \langle \sigma_{\alpha_1,\ldots,\alpha_k} \rangle_\rho^2$, and the total uncertainty $\mathcal{U}_T$:

$$\mathcal{U}_T = \sum_{i_1,\ldots,i_k=1}^{D_{A_1},\ldots,D_{A_k}} \Delta^2\sigma_{i_1,\ldots,i_k}\,. \tag{1.77}$$

In this section, we explicitly compute the total uncertainty in Eq. (1.77), showing that this quantity depends linearly only on the purity of the global state and on the purities of its subsystems for a given k-partition. We then discuss its connection with several well-known uncertainty relations, ranging from the original Heisenberg–Robertson relation (Robertson 1929) to its generalized

form derived by Schrödinger (Schrödinger 1930):

$$\Delta A^2 \, \Delta B^2 = \left| \frac{1}{2} \langle [\hat{A}, \hat{B}] \rangle \right|^2 + \left| \frac{1}{2} \langle \{\hat{A}, \hat{B}\} \rangle - \langle \hat{A} \rangle \langle \hat{B} \rangle \right|^2 + \mathcal{R} \, . \tag{1.78}$$

The additional term on the right-hand side of Eq. (1.78), required to saturate the relation, is commonly referred to in the literature as the "remainder" or "rest" term $\mathcal{R}$ (Carruthers and Nieto 1968).

We also discuss the connection with the Maccone–Pati uncertainty relations (Pati and Sahu 2007; Maccone and Pati 2014), which provide strengthened bounds based on the orthogonality of quantum states.

Building upon this expression and on the criterion introduced in (Hofmann and Takeuchi 2003), we derive a necessary condition for k-separability that can be expressed as an inequality involving only the purity of the state ρ and the purities of its subsystems. When this inequality is violated, the state is guaranteed to be entangled. This approach establishes a duality between the violation of local uncertainty relations for $\mathrm{SU}(N)$ operators and the purities of the state and its subsystems, providing a useful tool for entanglement detection. The connection between entanglement and purity has also been investigated in continuous-variable systems (Adesso et al. 2004).

1.4.1 Generalized uncertainty relations for SU(N) observables

In order to compute $\mathcal{U}_T$, we first introduce some definitions, beginning with the simplest case of two qubits to familiarize the reader with the concepts. A two-qubit state can be expanded as

$$\rho = \frac{1}{4} \left[\sum_{\alpha_1, \alpha_2 = 0}^{3} t_{\alpha_1 \alpha_2} \sigma_{\alpha_1 \alpha_2} \right] , \tag{1.79}$$

with $\sigma_{\alpha_1 \alpha_2} = \sigma_{\alpha_1} \otimes \sigma_{\alpha_2}$, $\sigma_0 = \mathbb{1}_2$ and σ_i for $i = 1, 2, 3$ are the Pauli matrices and $t_{\alpha_1 \alpha_2} = \mathrm{Tr}[\rho \sigma_{\alpha_1 \alpha_2}]$. The purity of ρ is then

$$\mathcal{P}^{(A_1, A_2)} = \frac{1}{4} \| T \| \, , \tag{1.80}$$

where T is the (4×4) matrix with elements $t_{\alpha_1 \alpha_2}$ and $\| \cdot \|$ is the squared Hilbert-Schmidt norm (or 2-norm), i.e. the sum of the squared eigenvalues τ_α:

$$\| T \| = \sum_{\alpha_1, \alpha_2 = 0}^{3} t_{\alpha_1 \alpha_2}^2 = \sum_{\alpha = 1}^{4} \tau_\alpha^2 \tag{1.81}$$

If we define the $3-$dimensional vectors $t^{(A_1)}$ and $t^{(A_2)}$ with elements t_{i0} and t_{0i} respectively for $i = 1, 2, 3$, and the (3×3) matrix $t^{(A_1,A_2)}$ with elements t_{ij} for $i, j = 1, 2, 3$, the purities of the subsystems are

$$\mathcal{P}^{A_l} = \frac{1}{2}\left[\|t^{(A_l)}\| + 1\right], \qquad l = 1, 2 \tag{1.82}$$

Since $\|T\| = 1 + \|t^{(A_1)}\| + \|t^{(A_2)}\| + \|t^{(A_1,A_2)}\|$, we obtain

$$\|t^{(A_1,A_2)}\| = 4\mathcal{P}^{(A_1,A_2)} - 2(\mathcal{P}^{(A_1)} + \mathcal{P}^{(A_2)}) + 1\,. \tag{1.83}$$

The generalization of the last equation to a $n-$qudits system, will be useful for finding the desired uncertainty.

In order to obtain such relation, we expand a global k-partite density operator ρ in terms of the generators of the $\mathrm{SU}(N_{A_l})$ algebra, that is

$$\rho = \frac{1}{N}\sum_{\alpha_1,\ldots,\alpha_k=0}^{D_{A_1},\ldots,D_{A_k}} t_{\alpha_1,\ldots,\alpha_k}\sigma_{\alpha_1,\ldots,\alpha_k}\,, \tag{1.84}$$

with $\sigma_{\alpha_1,\ldots,\alpha_k} \equiv \sigma^{(1)}_{\alpha_1}\otimes\cdots\otimes\sigma^{(k)}_{\alpha_k}$, $N = \prod_l N_{A_l}$, and the generators of $\mathrm{SU}(N_{A_l})$ have been normalized such that $\mathrm{Tr}\left[\sigma^{(l)}_{\alpha}\sigma^{(l)}_{\beta}\right] = N_{A_l}\delta_{\alpha\beta}$. In this way, the purity of the state ρ is given by

$$\mathcal{P}^{(A_1,\ldots,A_k)} = \frac{1}{N}\sum_{\alpha_1,\ldots,\alpha_k=0}^{D_{A_1},\ldots,D_{A_k}} t^2_{\alpha_1,\ldots,\alpha_k} = \frac{1}{N}\|T\|\,, \tag{1.85}$$

where T is the tensor with elements $t_{\alpha_1,\ldots,\alpha_n}$. At this point, as in the case of two-qubits, for $g \leq k$, and $i_1 \neq \ldots \neq i_g \in \{1,\ldots,k\}$, let us define the matrices $t^{(A_{i_1},\ldots,A_{i_g})}$ whose elements are given by the $t_{\alpha_1,\ldots,\alpha_k}$ with $\alpha_l = 0$ if $l \notin \{i_1,\ldots,i_g\}$ and $\alpha_l = 1,\ldots,D_{A_l}$ conversely. Hence, by using Eq. (1.85) and

$$\|T\| = 1 + \sum_i \|t^{(A_i)}\| + \sum_{i,j}\|t^{(A_i,A_j)}\| + \cdots + \|t^{(A_1,\ldots,A_k)}\|\,, \tag{1.86}$$

it is possible by induction (see Appendix A.4) to prove the following equality

$$\|t^{(A_1,\ldots,A_k)}\| = \sum_{g=1}^{k}\sum_{\mathcal{A}\in\{A\}_g}(-1)^{k-g}N_{\mathcal{A}}\mathcal{P}^{\mathcal{A}} + (-1)^k \tag{1.87}$$

where, for a fixed $g \leq k$, $\mathcal{A} \equiv A_{i_1}\ldots A_{i_g} \in \{A\}_g$ with $i_1 < \cdots < i_g$ is a

string of g subsystems taken from the original k-partition. Consequently, the sum over $\{A\}_g$ runs over all possible combinations (without repetition) of g subsystems of the k parts, that is

$$\sum_{\mathcal{A}\in\{A\}_g} N_{\mathcal{A}}\mathcal{P}^{\mathcal{A}} \equiv \sum_{i_1<i_2<\cdots<i_g=1}^{k} N_{A_{i_1}}\dots N_{A_{i_g}}\mathcal{P}^{(A_{i_1},\dots,A_{i_g})} \tag{1.88}$$

Equation (1.87) is the generalization to a generic system of the Eq. (1.83) which holds for two-qubits. For example, for a single qudit of dimension N it is

$$\|t\| = N\mathcal{P} - 1\,, \tag{1.89}$$

where $\mathcal{P}$ is the purity of the state. While for a bipartite system AB where the single parts have dimensions N_A and N_B, it is

$$\|t^{(A,B)}\| = N_A N_B \mathcal{P}^{(A,B)} - N_A\mathcal{P}^{(A)} - N_B\mathcal{P}^{(B)} + 1\,. \tag{1.90}$$

Uncertainty relations for SU(N) observables

At this point, we can connect the mean squared values of the SU(N_{A_l}) generators

$$\Delta^2\sigma_{i_1,\dots,i_k} \equiv \langle\sigma^2_{i_1,\dots,i_k}\rangle_\rho - \langle\sigma_{i_1,\dots,i_k}\rangle^2_\rho \tag{1.91}$$

to the relation (1.87). Since $\sigma^2_{\alpha_1,\dots,\alpha_k} = \mathbb{1}_N$, it is

$$\mathcal{U}_T = \sum_{i_1,\dots,i_k=1}^{D_{A_1},\dots,D_{A_k}} \Delta^2\sigma_{i_1,\dots,i_k} = \mathcal{N} - \|t^{(A_1,\dots,A_k)}\|\,, \tag{1.92}$$

where

$$\mathcal{N} = \sum_{i_1,\dots,i_k=1}^{D_{A_1},\dots,D_{A_k}} 1 = \prod_{l=1}^{k} D_{A_l} \tag{1.93}$$

and $\|t^{(A_1,\dots,A_k)}\|$ is expressed in term of the purities by the Eq. (1.87).

The last relation holds for an arbitrary composite quantum state and for any fixed k-partition of the system. It provides an analytical expression for the total uncertainty associated with all tensor products of the SU(N_{A_l}) observables, which is exactly quantified in terms of the purity of the global state and the purities of the k subsystems under consideration.

Once these purities are fixed, the total amount of uncertainty is also fixed. Therefore, states sharing the same set of purities can differ only in the distribution

of uncertainty among the individual terms appearing in Eq. (1.92), while the total sum remains conserved.

Saturation of the Schrödinger uncertainty relation

Consider a composite quantum system partitioned in k subsystems $A_1, \ldots, A_k$, where at least one subsystem, say A_l, is a single qubit. For ease of notation, in the following we drop the subsystem index and denote the qubit simply by $A \equiv A_l$.

Introducing $r_i \equiv t_i^{(A)}$ and recalling that $t_i^{(A)} = \langle \sigma_i^A \rangle$ we can rewrite (1.87) as follows:

$$\sum_i r_i^2 = 2\mathcal{P}_A - 1 \implies \Delta^2\sigma_{i,A} \equiv \langle (\sigma_i^A - \langle \sigma_i^A \rangle)^2 \rangle = 1 - r_i^2 \,, \tag{1.94}$$

where $r_i \equiv t_i^{(A)}$, we have

$$\sum_i \Delta^2\sigma_{i,A} = 2(2 - \mathcal{P}_A) \tag{1.95}$$

and

$$\Delta^2\sigma_{i,A}\Delta^2\sigma_{j,A} = r_k^2 + r_i^2 r_j^2 + 2(1 - \mathcal{P}^{(A)}) \,, \quad \text{for } i \neq j \neq k. \tag{1.96}$$

The last equations are a generalized form of the Heisenberg-Robertson uncertainty relation and its generalized Schrödinger version (1.78)

$$\Delta^2\sigma_{i,A}\Delta^2\sigma_{j,A} = \langle \sigma_k^A \rangle^2 + \langle \sigma_i^A \rangle^2 \langle \sigma_j^A \rangle^2 + 2(1 - \mathcal{P}^{(A)}) \,, \tag{1.97}$$

which holds for a qubit state, pure or mixed, and for $i \neq j \neq k$.

The quantity $\mathcal{R} = 2(1 - \mathcal{P}^{(A)})$ is known as the linear entropy of the system. If the total multipartite system is pure, it quantifies the bipartite entanglement between qubit A and the remainder of the system, since it coincides, up to a factor, with both the negativity and the concurrence. For mixed states, it quantifies the total amount of correlations (both classical and quantum) between the considered qubit and the remainder of the system.

In the case of qudit systems, quantifying the product of two non-compatible observables is not as straightforward as in the case of Pauli operators. Nevertheless, we can still employ Eqs. (1.92) and (1.87), which allow us to quantify the sum of the uncertainties of non-compatible $\mathrm{SU}(N_l)$ observables in terms of the purity of the global state and its subsystems, for any k-partition.

Two-qubit case

In order to find similar relations for a two-qubit state, let us consider the class of "T-diagonal" states, defined as those states for which $t_{ij} = 0$ for $i \neq j$, with $i, j = 1, 2, 3$. This class is particularly relevant, as any two-qubit state can be transformed into a T-diagonal form through local unitary operations (Horodecki et al. 2009), which preserve both purity and entanglement.

We can rewrite Eq. (1.90) as

$$\sum_{i=1}^{3} t_{ii}^2 = 4\mathcal{P}^{(AB)} - 2\mathcal{P}^{(A)} - 2\mathcal{P}^{(B)} + 1\,. \tag{1.98}$$

If we define

$$\Delta\sigma_{ij}^2 = \langle(\sigma_i^A \otimes \sigma_j^B - \langle\sigma_i^A \otimes \sigma_j^B\rangle)^2\rangle = 1 - t_{ij}^2\,, \tag{1.99}$$

we have

$$\sum_{i=1}^{3} \Delta^2\sigma_{ii} = 2(\mathcal{P}^{(A)} + \mathcal{P}^{(B)} - 2\mathcal{P}^{(AB)} + 1) \tag{1.100}$$

and also

$$\Delta^2\sigma_{ii}\Delta^2\sigma_{jj} = \langle\sigma_{kk}\rangle^2 + \langle\sigma_{ii}\rangle^2\langle\sigma_{jj}\rangle^2 + 2(\mathcal{P}^{(A)} + \mathcal{P}^{(B)} - 2\mathcal{P}^{(AB)})\,, \tag{1.101}$$

where $\sigma_{ii} \equiv \sigma_i \otimes \sigma_i$.

If we do not restrict to $T-$diagonal states, we have

$$\sum_{i,j=1}^{3} \Delta^2\sigma_{ij} = 2(\mathcal{P}^{(A)} + \mathcal{P}^{(B)} - 2\mathcal{P}^{(AB)} + 4)\,. \tag{1.102}$$

1.4.2 Sufficient criterion for k-separability

In order to obtain the desired criterion, let us consider the mean squared value $\langle\Delta^2 O\rangle_\rho$ of a Hermitian operator $\hat{O}$ on a state ρ, that is

$$\Delta_\rho^2 O = \left\langle \left(\hat{O} - \langle O\rangle_\rho\right)^2 \right\rangle_\rho\,, \tag{1.103}$$

with $\langle\hat{O}\rangle_\rho \equiv \mathrm{Tr}[\rho\hat{O}]$.

If a state is a convex combination of other states ρ_i, with probabilities p_i, we

can write (Hofmann and Takeuchi 2003)

$$\Delta_{\rho}^{2} O = \sum_{i} p_i \left(\Delta_{\rho_i}^{2} O + \widetilde{\Delta}_{\rho_i}^{2} \right) , \tag{1.104}$$

where $\widetilde{\Delta}_{\rho_i}^{2} = \left(\langle \hat{O} \rangle_{\rho_i} - \langle \hat{O} \rangle_{\rho} \right)^2$.

With this argumentation, it was possible to relate the separability problem to the violation of local uncertainty relations (Hofmann and Takeuchi 2003). The problem was subsequently generalized to any (also nonlocal) observable, using a method that involves the covariance matrix of the chosen observables (Gühne 2004; Gühne et al. 2007).

Now we will show that for a k-separable state ρ, if one considers as observables the product tensors of the Hermitian SU(N_{A_l}) generators, with $l = 1, \ldots, k$, it becomes possible to compute the positive term $\sum_{\text{comb}} \sum_i p_i \tilde{\Delta}_{\rho_i}^2$, even without knowledge of the k-partition of ρ. Here, $\sum_{\text{comb}}$ denotes the sum over all possible tensor products of the SU(N_{A_l}) generators. This term depends (linearly) only on the purity of ρ and the purities of all its subsystems. Consequently, we conclude that when this quantity is negative, ρ cannot be k-separable, thus providing a sufficient and computable condition for k-separability encapsulated into a simple inequality, also suitable for experimental realization.

If a state ρ is k-separable (Eq. (1.41)), since any operator $\rho_i^{(A_l)}$ can be written as convex combination of the projectors corresponding to its (orthogonal) eigenvectors defined by the set $\left\{ |\Psi_{l_i}\rangle^{(A_l)} \right\}$, we can always write ρ as a convex combination of pure k-product states:

$$\rho = \sum_{i} p_i \, |\Psi_i\rangle^{(A_1)} \langle \Psi_i| \otimes \cdots \otimes |\Psi_i\rangle^{(A_k)} \langle \Psi_i| \tag{1.105}$$

for an appropriate probability distribution $\{p_i\}$. For a generic bipartite system, it has been shown that there exists a convex combination of r product states with $r \leq N^2$ (Horodecki 1997). With this last expression, we can choose $\hat{O} = \sigma_{i_1,\ldots,i_n}$ in Eq. (1.104) and then sum over $i_1, \ldots, i_n$ obtaining

$$\sum_{i_1,\ldots,i_k=1}^{D_{A_1},\ldots,D_{A_k}} \Delta_{\rho}^{2} \sigma_{i_1,\ldots,i_k} = \sum_{j} p_j \sum_{i_1,\ldots,i_k} \langle \Delta_{\Psi_j}^{2} \sigma_{i_1,\ldots,i_k} \rangle + \widetilde{\Delta}^2 , \tag{1.106}$$

where

$$\widetilde{\Delta}^2 = \sum_{j} p_j \sum_{i_1,\ldots,i_k} \left(\langle \sigma_{i_1,\ldots,i_k} \rangle_{\Psi_i} - \langle \sigma_{i_1,\ldots,i_k} \rangle_{\rho} \right)^2 \tag{1.107}$$

is a positive quantity and the average $\langle \ldots \rangle_{\Psi_i}$ is performed on the product state

$$|\Psi_i\rangle \equiv |\Psi_i\rangle^{(A_1)} \otimes \cdots \otimes |\Psi_i\rangle^{(A_k)} \ . \tag{1.108}$$

The l.h.s. of Eq. (1.106) is the same as (1.92), while the first term on the right side, by means of the same equation for pure k-product states and Eq. (1.87) is independent from the particular product state $|\Psi_i\rangle$ and can be exactly computed putting all the purities equal to 1. Hence, by inverting Eq. (1.106), the term $\mathcal{N}$ cancels out, and since the probabilities are normalized, we obtain

$$\widetilde{\Delta}^2 = \sum_{g=1}^{k} \sum_{\mathcal{A}\in\{A\}_g} (-1)^{k-g} N_{\mathcal{A}} \left(1 - \mathcal{P}^{\mathcal{A}}\right) \ , \tag{1.109}$$

which must be positive for k-separable states. Therefore we can conclude that for a quantum state ρ and a partition $\{A_l\}_{l=1}^{k}$, if

$$\sum_{g=1}^{k} \sum_{\mathcal{A}\in\{A\}_g} (-1)^{k-g} N_{\mathcal{A}} \left(1 - \mathcal{P}^{\mathcal{A}}\right) < 0 \, , \tag{1.110}$$

then ρ is not k-separable for the considered partition.

The quantities $1 - \mathcal{P}^{(A_{i_1},\ldots,A_{i_g})}$ involved in the last inequality are also called linear entropies, since they are the first order approximation of the von Neumann entropies of the different sub-parts.

The sufficient condition for entanglement given by (1.110) is equivalent to a necessary condition for k-separability, which can also be stated as: if ρ is k-separable, then

$$\|t^{(A_1,\ldots,A_k)}\| < (-1)^k + \sum_{g=1}^{k} \sum_{\mathcal{A}\in\{A\}_g} (-1)^{k-g} N_{\mathcal{A}} \, . \tag{1.111}$$

The set of states satisfying the above condition is a convex set; to prove this we consider two states ρ and σ satisfying (1.111) and we denote its r.h.s. with X. In this way we have

$$\begin{aligned} &\|p t_\rho^{(A_1,\ldots,A_k)} + (1-p) t_\sigma^{(A_1,\ldots,A_k)}\| \\ \leq &[p^2 + (1-p)^2] X + 2p(1-p) \, \langle t_\rho^{(A_1,\ldots,A_k)}, t_\sigma^{(A_1,\ldots,A_k)} \rangle_{\text{HS}} \\ \leq &[p^2 + (1-p)^2] X + 2p(1-p) \sqrt{\|t_\rho^{(A_1,\ldots,A_k)}\| \, \|t_\sigma^{(A_1,\ldots,A_k)}\|} \\ \leq &X \ , \end{aligned} \tag{1.112}$$

where we used the contractivity of the Hilbert-Schmidt norm.

Note that Eq. (1.111) can be further simplified since the r.h.s can be rewritten as

$$(-1)^k + \sum_{g=1}^{k} \sum_{\mathcal{A}\in\{A\}_g} (-1)^{k-g} N_{\mathcal{A}} = \prod_{g=1}^{k} \left(N_{I_g} - 1\right) . \tag{1.113}$$

Relation to other criteria

We already discussed in Section 1.3.3, that the correlation matrix is related to entanglement criteria and entanglement witnesses (De Vicente and Huber 2011; Badziag et al. 2008; Friis et al. 2019). Using the derived formula (1.87), all these criteria can be written in terms of the purities.

Note that in the n-qubit case, $N_{I_g} = 2$ the sufficient condition for separability (1.111), reads as

$$\|t\| > 1 , \tag{1.114}$$

which is exactly inequality (1.74).

Such an inequality is also a necessary condition for nonlocality. In fact, it has been shown that in a scenario involving two measurements per party, if the state is nonlocal, there exists a set of local coordinates such that (Żukowski and Brukner 2002)

$$\sum_{i_1,\ldots,i_n=1}^{2} t_{ij}^2 > 1 . \tag{1.115}$$

Inequality (1.115) is thus a necessary condition for Bell nonlocality of n-qubit states. Observing that the summation runs over only two of the three indices of the correlation matrix t for each qubit, it follows that inequality (1.115) implies inequality (1.114), However, this condition is not sufficient, since there exist local states that also violate the inequality. For this reason, the separability condition we derived is nontrivial, as it is capable of detecting states that admit a local hidden variable description within the framework introduced above.

1.4.3 Exponential advantage of the formulation in terms of purities

In this section we will prove that beyond the conceptual importance of Eq. (1.87) for connecting uncertainty and entanglement, it also provides a direct and more efficient tool for experimentally measuring the Hilbert-Schmidt norm of the correlation matrix t, and thus for all correlation-based criteria for entanglement and nonlocality detection discussed above (Badziag et al. 2008; De Vicente and

Huber 2011; Friis et al. 2019). The key reason is that purities can be efficiently measured, for instance by means of randomized measurement schemes (Van Enk and Beenakker 2012; Elben et al. 2018; Yanay and Tahan 2021). Importantly, the efficiency of such schemes even improves as the number of parties increases, as demonstrated in the n-qubit case (Van Enk and Beenakker 2012).

To illustrate this, let us consider the n-qubit case for a fixed k-partition. If the state ρ for which we aim to detect entanglement or nonlocality is unknown, one needs to measure 3^k terms to determine $\|t\|$, and the situation becomes even more demanding for higher-dimensional qudit systems, where the number of terms increases.

In contrast, for the fixed k-partition, the number of independent parameters required to determine the purities is only $2^k - 1$.

Although the time required to experimentally estimate a single purity element, t_{pur}, may differ from the time needed to measure a single element of the correlation matrix, t_{corr}, it has been shown that t_{pur}, defined as the number of measurements required to estimate a single purity element within a fixed error, does not increase with the size of the system (and thus with the number of elements of the correlation matrix), as proved, e.g., in (Van Enk and Beenakker 2012). Therefore, the total measurement times for the two formulations are, respectively:

$$\begin{aligned} T_{\text{corr}} &= (\#\text{ correlation matrix elements}) \cdot t_{\text{corr}}, \\ T_{\text{pur}} &= (\#\text{ purities}) \cdot t_{\text{pur}}. \end{aligned} \tag{1.116}$$

Since the above quantities obey the scaling behaviors reported in Table 1.1, it follows that for sufficiently large system sizes $T_{\text{pur}} \ll T_{\text{corr}}$, even when $t_{\text{corr}} < t_{\text{pur}}$.

The origin of the scaling is clear: the Hilbert-Schmidt norm of the correlation matrix, as well as the global and local purities, is invariant under local unitary transformations. Therefore, our formula identifies precisely the minimal set of features necessary to determine $\|t\|$.

It is true that, as noted in (De Vicente and Huber 2011), it is not necessary to measure all the correlation terms, since each of them contributes a positive quantity. Hence, once the partial sum exceeds 1, it is already guaranteed that $\|t\| > 1$. Nevertheless, for any state in $\mathcal{D}(\mathcal{H})$ this situation does not typically occur. This observation can be understood using the following argument. A mixed state ρ can be sampled according to the probability distribution $P(\rho) = P_H(\mathbf{F}^{(N)}) \times P(\vec{\lambda})$, where $P_H(\mathbf{F}^{(N)})$ is the distribution of the flag manifold induced by the Haar measure, and $P(\vec{\lambda})$ is a distribution on the eigenvalues simplex, which can be chosen in different ways (Życzkowski and Sommers 2001). Interesting choices as the Hilbert-Schmidt and Bures distributions, lead

to the average purity $\langle \mathcal{P}^{(A_1,\ldots,A_k)} \rangle \sim \frac{\alpha}{N}$ (N large)[1] (Życzkowski and Sommers 2001; Bengtsson and Życzkowski 2017).

Therefore, by using Eq. (1.85), one obtains $\|T\| \sim \alpha$, despite N being large. Consequently, the number of correlation terms required to determine whether $\|t\| > 1$ still scales as reported in Table 1.1.

In order to better understand the regimes in which the use of the purity relation (1.87) provides an advantage, we consider the case of $n = 6$ qubits with $k = n$. In Fig. 1.4a, we plot the average value $\langle \|t\| \rangle$ of $\|t\|$ as a function of the purity, averaged over 50 samples for each fixed purity value. In Fig. 1.4b, we show the average minimal number $\langle N_{\text{meas}} \rangle$ of elements t_{ij}, randomly extracted from t, that are required to ensure $\|t\| > 1$.

From these plots, it is clear that for pure states there is no advantage in using the purity relation to compute $\|t\|$, since the number of terms involved is almost the same. However, as the purity decreases, $\|t\|$ also decreases, causing $\langle N_{\text{meas}} \rangle$ to grow exponentially. This demonstrates that Eq. (1.87) provides an exponential advantage in such cases.

In the case of qudit systems, the number of terms involved in the correlation matrix is even larger, since the number of $\mathrm{SU}(N)$ generators increases quadratically with N. On the other hand, the number of purities involved for a fixed k-partition remains 2^k, as summarized in Table 1.1.

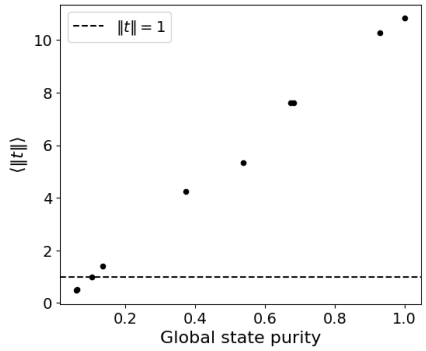
Local dimension	# of matrix elements	# of purities
Qubits ($N_i = 2$)	3^k	$2^k - 1$
Uniform qudits ($N_i = N$)	$(N^2-1)^k$	$2^k - 1$
Qudits (N_i arbitrary)	$\prod_{i=1}^{k} (N_i^2 - 1)$	$2^k - 1$

Table 1.1: *Comparison between the number of correlation-tensor components required to evaluate $\|t\|$ and the number of purities for a fixed k-partition.*

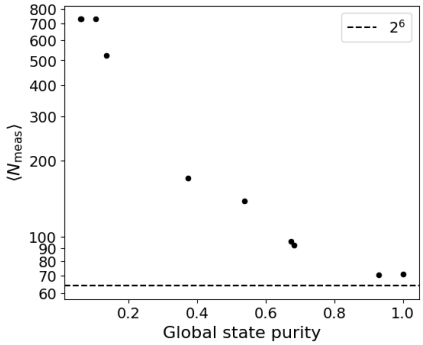
For example, if we consider the multipartite criterion for the case of three qudits, given by inequality (1.76), then for large single-qudit dimensions N_i the number of correlators involved scales as N_i^6, while the number of purities is only $2^3 - 1 = 7$, independent of N_i.

If the state ρ is known, in the bipartite case of two qudits with equal local dimension N_A, the experimental effort required to measure $\|t\|$ can be reduced to N_A^2 terms by exploiting the Schmidt decomposition. In this case, the advantage of our method is therefore lost. Nevertheless, in the multipartite scenario such a simplification is not possible, since a general Schmidt decomposition does not

[1] Specifically, we have $\langle \mathcal{P}^{(A_1,\ldots,A_k)} \rangle_{\text{HS}} = 2N/(N^2+1)$ and $\langle \mathcal{P}^{(A_1,\ldots,A_k)} \rangle_{\text{B}} = (5N^2+1)/(2N(N^2+1))$, corresponding to $\alpha = 2$ and $\alpha = 5/2$, respectively.



(a) Average value of $\|t\|$ over 50 samples for each fixed purity.



(b) Average minimal number N_{meas} of t_{ij} elements required to determine if $\|t\| > 1$, computed over 50 samples per fixed purity. When $\|t\| < 1$, this number reaches $3^6 = 729$, while the number of purities needed to compute $\|t\|$ exactly is $2^6 = 64$. A logarithmic scale has been used for the y axis.

Figure 1.4: *The plots show that as the purity of the global state decreases, relation* (1.87) *offers an exponential advantage.*

exist for k parties, and hence, again, all the terms in $\|t\|$ need to be measured. Consequently, in the multipartite scenario, our approach remains advantageous even when the state is known.

1.4.4 k-separability: examples

In this section we will analyze the inequality (1.110) for the simplest case of two qubits, where it takes the form

$$2\mathcal{P}^{(A,B)} - (\mathcal{P}^{(A)} + \mathcal{P}^{(B)}) > 0\,. \tag{1.117}$$

Pure states

First of all, we note that for pure states $\mathcal{P}^{(AB)} = 1$, the inequality also gives a necessary condition, since $\mathcal{P}^{(A)} = \mathcal{P}^{(B)} \leq 1$ and the equalities hold only when both subsystems A and B are pure, i.e., the state is a product state.

Relation to the violation of the CHSH inequality

For mixed states, it is worth mentioning that the quantity involved in the equality appears also in the violation of the CHSH inequality. Indeed, a necessary

and sufficient condition that a two-qubit state violates the CHSH inequality is (Horodecki et al. 1995)

$$u_1 + u_2 > 1\,, \tag{1.118}$$

where u_1 and u_2 are the two biggest eigenvalues of the matrix $U = (t^{(A,B)})^T t^{(A,B)}$ where T is the transposition. If we call u_3 the minimum eigenvalue of U, we obtain

$$\|t^{(A,B)}\| = \sum_{i=1}^{3} u_i\,. \tag{1.119}$$

Therefore, by means of Eq. (1.83), we can rewrite the inequality (1.118) as

$$2\mathcal{P}^{(A,B)} - (\mathcal{P}^{(A)} + \mathcal{P}^{(B)}) > \frac{u_3}{2}\,. \tag{1.120}$$

Therefore, since $u_3 \geq 0$, if a two-qubit state violates the CHSH inequality, then it violates also the inequality (1.117). Hence the criterion is able to detect the entanglement in all the states that violate the CHSH inequality.

Werner states and Bell-diagonal states - a geometrical interpretation

What happens for two-qubit states that do not violate the CHSH inequality? To answer this question, we analyze the Werner states (Werner 1989), given by the convex combination of a Bell state with the maximally mixed state

$$\rho_\omega = p_\omega \left|\Psi\right\rangle \left\langle\Psi\right| + \frac{1 - p_\omega}{4} \mathbb{1}_4\,, \tag{1.121}$$

where $|\Psi\rangle$ is one of the four Bell states. Werner states are not separable for $p_\omega > \frac{1}{3}$ and violate the CHSH inequality for $p_\omega > \frac{1}{\sqrt{2}}$. The reduced states of a Werner state, are the maximally mixed states, which have $\mathcal{P}^{(A)} = \mathcal{P}^{(B)} = \frac{1}{2}$, therefore the inequality (1.117) is violated for $p_\omega > \frac{1}{\sqrt{3}}$, that is the result already found in (Gühne 2004). This proves that the condition is not necessary, but still able to detect entanglement in states that do not violate the CHSH inequality. The analysis of Werner states is resumed in the figure 1.6.

Werner states are a particular class of Bell-diagonal states, i.e., a convex combination of Bell states. Bell-diagonal has a parametrization (Horodecki et al. 1995):

$$\rho_{BD} = \frac{1}{4}\left(1 + \sum_{i=1}^{3} t_{ii} \sigma_i^A \otimes \sigma_i^B\right)\,, \tag{1.122}$$

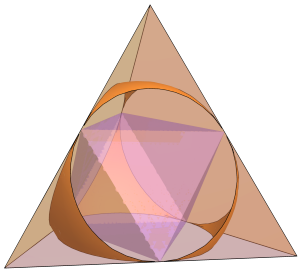


Figure 1.5: *Tetrahedron of Bell-diagonal states together with the octahedron of separable states and the criterion surface. The states in the octahedron lying outside of the criterion surface are entangled states detected by the criterion.*

hence $t_i^{(A)} = t_i^{(B)} = 0$ for $i = 1, 2, 3$ and the matrix $t^{(A,B)}$ is diagonal. Werner states are obtained with $t_{ii} = -\omega$ for $i = 1, 2, 3$. By using Eq. (1.83), for two-qubit states, the inequality (1.117) takes the form

$$\|t^{(A,B)}\| - 1 > 0\,. \tag{1.123}$$

For Bell-diagonal states, $\|t^{(A,B)}\|$ can be interpreted also as the (Hilbert-Schmidt) distance from the maximally mixed state, which indeed has $\|t\| = 1$. Therefore, in the $t-$space the equality $\|t^{(A,B)}\| = 1$ defines a sphere at constant purities $\mathcal{P}^{(A,B)} = \frac{1}{2}$. The set of separable Bell-diagonal states is given by the octahedron $|t_{11}| + |t_{22}| + |t_{33}| = 1$ which is inscribed in the sphere, as shown in the figure 1.5. For Bell-diagonal states, we can estimate the number of entangled states detected by the criterion over the total entangled state by calculating the ratios between the two volumes in the $t-$space obtaining a ratio $R_{detected}$

$$R_{detected} = \frac{2\sqrt{3} - \pi\sqrt{3} + \pi}{2\sqrt{6}} \simeq 0.52\,. \tag{1.124}$$

Randomly extracted two-qubits states

For a general two-qubit state, the inequality in (1.123) remains valid. However, it cannot still be interpreted as the distance from the maximally mixed state. For randomly generated states, the criterion appears to fail when the entanglement (quantified in terms of negativity) is low, while it reliably detects entanglement when the negativity is high, as illustrated in Figures 1.7.

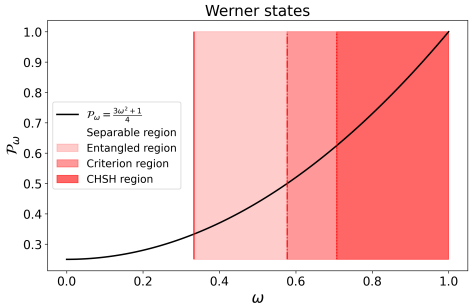
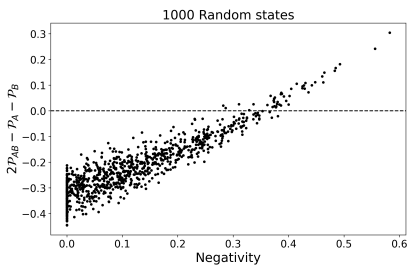
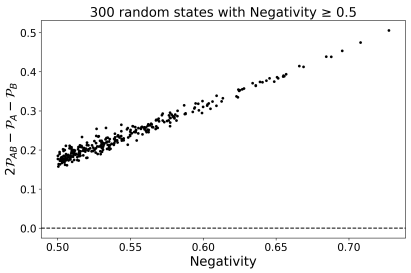


Figure 1.6: *Purity of Werner states and different regions in which the state are separable $\omega \leq \frac{1}{3}$, entangled $\omega > \frac{1}{3}$, in which the criterion is able to detect entanglement $\omega > \frac{1}{\sqrt{3}}$ and where CHSH inequalities are violated $\omega > \frac{1}{\sqrt{2}}$. The corresponding purities are $\mathcal{P}_{sep} = \frac{1}{3}$, $\mathcal{P}_{criterion} = \frac{1}{2}$, $\mathcal{P}_{CHSH} = \frac{5}{8}$.*



(a) Generic random extracted two-qubit states. The criterion fails when the entanglement is too low.



(b) Generic random extracted two-qubit states with negativity ≥ 0.5. The criterion never fails.

Figure 1.7: *For randomly extracted states the criterion seems to work better when the entanglement between the two qubits increases.*

n-qubit systems

In the case of $n-$qubits system, let us suppose that we want to study the fully separability, hence we put $k = n$ in Eq. (1.110). For a fixed g it is

$$\sum_{\{A\}_g} N_{\{A\}_g} \left(1 - \mathcal{P}^{\{A\}_g}\right) = 2^g \sum_{i_1 < \cdots < i_g} \left(1 - \mathcal{P}^{(A_{i_1}, \ldots, A_{i_g})}\right). \tag{1.125}$$

Hence the term

$$q \equiv \sum_{g=1}^{n}(-1)^{n-g}2^g \sum_{i_1<\cdots<i_g} 1 = \sum_{g=0}^{n}(-1)^{n-g}2^g \binom{g}{n} - (-1)^n = 1-(-1)^n \tag{1.126}$$

is null for odd n and 2 for even n.

Therefore, the inequality can be written as

$$\sum_{g=1}^{n}(-1)^{n-g}\sum_{\{A\}_g} 2^g \mathcal{P}^{\{A\}_g} - q > 0\,. \tag{1.127}$$

If all the subsystems are the maximally mixed states we have for fixed $g < n$, $\mathcal{P}^{\{A\}_g} = 2^{-g}$ and hence

$$\begin{aligned} &2^n\mathcal{P}^{(A_1,\ldots,A_n)} + \sum_{g=1}^{n-1}\binom{i}{n}(-1)^{g-n} - q = \\ &2^n\mathcal{P}^{(A_1,\ldots,A_n)} + \sum_{g=0}^{n}\binom{i}{n}(-1)^{g-n} - q - (-1)^n - 1 = \\ &2^n\mathcal{P}^{(A_1,\ldots,A_n)} - 2 > 0\,. \end{aligned} \tag{1.128}$$

Therefore if the purity of the (total) state

$$\mathcal{P}^{(A_1,\ldots,A_n)} > \frac{1}{2^{n-1}}\,, \tag{1.129}$$

the state is fully-separable. The last inequality is the generalization of what we did for the two-qubits Bell-diagonal states. In general, for those states whose reduced states are the maximally mixed states, if the purity of the system is bigger than 2 times the purity of the maximally mixed state, then the state is not fully-separable.

Similar results are obtained if one considers the k-separability of $n-$qubit states.

Higher dimensions

For qudit in higher dimension, let us consider the case of bipartite systems of dimension $N = N_A N_B$. In this case, the separability inequality (1.110) reads as

$$N(1-\mathcal{P}^{(AB)}) - N_A(1-\mathcal{P}^{(A)}) - N_B(1-\mathcal{P}^{(B)}) < 0\,. \tag{1.130}$$

If N increases and the two subsystems have almost the same dimensions $N_A \simeq N_B$ the criterion loses its capacity to detect entangled state. Indeed, if the purities are of the same order of magnitude, the first term goes with $\mathcal{O}(N) = \mathcal{O}(N_A^2)$ and the inequality is never satisfied, except when $\mathcal{O}(1 - \mathcal{P}^{(AB)}) = \mathcal{O}\big(\frac{1}{N_A}\big)$, which implies $\mathcal{P}^{(AB)} \simeq 1$. For example, if we consider two qudits with $N_A = N_B = 10$, the purity of the total state has to be at least $\mathcal{P}^{(AB)} \gtrsim 0.8$. The situation is different when $N_A \gg N_B$ (or vice-versa).

The validity ranges of the criterion improve in the case of multipartite systems, since there are the binomial coefficient terms $\binom{l}{k}$ which can be the same order of magnitude or bigger than N for high number of parties k, as shown also in the case of n qubits.

1.5 Geometric measures of Bell nonlocality

1.5.1 Bell nonlocality

Quantum nonlocality is generally understood as the impossibility of describing experimental correlations using a local hidden variable (LHV) model (Bell 1964; Brunner et al. 2014). A multipartite state is considered nonlocal if there exists at least one experimental configuration, involving non-communicating parties that can share at most classical randomness, which leads to a violation of a Bell-type inequality. For example, in the case of two dichotomic observables per observer ($m = 2$, $\Delta = 2$), the Bell inequality takes the CHSH form:

$$E(x, y) - E(x, y') + E(x', y) + E(x', y') \leq 2, \tag{1.131}$$

where (x, x') and (y, y') are the observables of Alice and Bob, respectively. $E(x, y)$ denotes the statistical expectation value of the measurement outcomes for observables x and y in a given quantum state. For nonlocal states, Eq. (1.133) does not hold, leading to a violation of the CHSH inequality, Eq. (1.131). It is clear that a state that appears local with respect to one Bell inequality may still violate a different Bell inequality (Collins and Gisin 2004; Laskowski et al. 2004; Masanes et al. 2008). It can also occur that a local quantum state violates a Bell inequality after being preprocessed by local filtering operations, thus revealing a *hidden nonlocality* (Popescu 1995; Gisin 1996). Furthermore, there exist quantum states that do not violate any Bell inequality individually, but when multiple copies of the state are considered along with joint measurements, the resulting system can exhibit a violation of some Bell inequality. This phenomenon is known as the *superactivation* of Bell nonlocality (Navascués and Vértesi 2011; Palazuelos 2012).

For a fixed scenario involving n observers, each allowed to perform m_i

measurements labeled by x_i ($i = 1, \ldots, n$), and each measurement having Δ_i possible outcomes a_i, we define the set $\mathcal{L}_{\vec{m},\vec{\Delta}}$ of *local states*, i.e., the set of states that admit a local hidden variable (LHV) description, as follows:

$$p(\vec{a}|\vec{x}) = \sum_{\lambda} p(\lambda) \prod_{i=1}^{n} p(a_i|\lambda, x_i), \tag{1.132}$$

for any choice of $\vec{a}$ and $\vec{x}$, where

$$p(\vec{a}|\vec{x}) = \mathrm{Tr}\left[\rho \bigotimes_{i=1}^{n} M_{a_i|x_i}\right]. \tag{1.133}$$

Here, $M_{a_i|x_i}$ denotes the positive operator-valued measure (POVM) element associated with outcome a_i for the measurement setting x_i. It follows directly that, for separable states, Eq. (1.133) implies Eq. (1.132). Consequently, any nonlocal state must necessarily be entangled. More generally, in the case more scenarios are taken into account, the set of local states $\mathcal{L}$ may coincide with the intersection $\bigcap_{i=1}^{N} \mathcal{L}_{\vec{m},\vec{\Delta}_i}$ of N sets $\{\vec{m}_i, \vec{\Delta}_i\}_{i=1}^{N}$ defined for fixed scenarios.

In the following sections, by analogy with coherence and entanglement, we will define well-behaved geometric monotones of nonlocality (Zanfardino et al. 2026b), thereby laying the foundation for constructing a quantum resource theory of nonlocality.

In what follows, the set of local states will be denoted by $\mathcal{L}$.

The free operations consist of local operations and shared randomness (LOSR) (Chitambar and Gour 2019; Schmid et al. 2020). These operations can be generally defined as follows:

$$\int_{\lambda} d\lambda \Phi_{\lambda}^{A} \otimes \Phi_{\lambda}^{B}, \tag{1.134}$$

where λ denotes the shared classical random variable between Alice and Bob, drawn according to a probability distribution $p(\lambda)$, and Φ_{λ}^{A} and Φ_{λ}^{B} are completely positive and trace-preserving (CPTP) maps. These transformations cannot take a local state outside the set of local states, as they preserve the hidden variable model. Since LOSR operations are also LOCC (although the converse does not hold), they inherit many properties established for the extensively studied class of LOCC transformations and satisfy the conditions listed in Sec. 1.1.1. See 1.2 for a summary table of symbols.

Symbol	Meaning
n	Number of observers (parties) involved in the Bell scenario
m_i	Number of measurement settings available to observer i
x_i	Measurement setting chosen by observer i
Δ_i	Number of possible outcomes for each measurement of observer i
a_i	Measurement outcome obtained by observer i
$\vec{x}$	Vector of measurement settings, $\vec{x} = (x_1, \ldots, x_n)$
$\vec{a}$	Vector of measurement outcomes, $\vec{a} = (a_1, \ldots, a_n)$
$M_{a_i\|x_i}$	POVM element corresponding to outcome a_i for measurement setting x_i
$\mathcal{L}_{\vec{m},\vec{\Delta}}$	Set of local quantum states admitting a local hidden variable model for a fixed Bell scenario

Table 1.2: *Summary of symbols used throughout the manuscript and their meanings.*

1.5.2 Geometric and Entropic Measures of Bell nonlocality

For a given quantum state ρ we take as measure $\mathcal{M}(\rho)$ of Bell nonlocality any contractive distance (with the exception of the Hilbert-Schmidt distance) D between ρ and the closest local state ρ^{loc}:

$$\mathcal{M}(\rho) = \min_{\rho^{loc}\in\mathcal{L}} D(\rho, \rho^{loc})\,. \tag{1.135}$$

Physically meaningful and mathematically sound geometric measures of a quantum resource should satisfy the key properties (1–4) discussed in Sec. 1.1.1. The proof that geometric measures, for suitable distances, satisfy these properties is the same as in Sec. 1.2, since it relies on the fact that the set of free states is convex and that the free operations preserve this set.

Regarding property 5, concerning computability, the set of local states is not known in general. In practice, one can fix a Bell scenario (or Bell inequality) for which the local set—still convex—is known, and compute the measure for this specific case. The resulting quantity not only provides a measure of nonlocality within the chosen scenario, but also represents a lower bound for more general scenarios and for intrinsic nonlocality.

In the following, we introduce geometric and entropic measurements and analyze the properties they satisfy.

Hilbert–Schmidt distance. The Hilbert–Schmidt (HS) distance between two density operators ρ_1 and ρ_2 is defined as

$$D_{\text{HS}}(\rho_1, \rho_2) = \|\rho_1 - \rho_2\|_2 = \sqrt{\text{Tr}[(\rho_1 - \rho_2)^2]}\,. \tag{1.136}$$

Although simple to compute, the HS distance has no direct operational meaning and is not contractive under completely positive trace-preserving (CPTP) maps; for instance, it can increase under trivial extensions of local subsystems (Bengtsson and Życzkowski 2017; Piani 2012). Nevertheless, it provides easily computable bounds for valid contractive distances.

Hellinger distance. The Hellinger distance is defined as

$$D_{\mathrm{He}}(\rho_1, \rho_2) = \|\sqrt{\rho_1} - \sqrt{\rho_2}\|_2 \,. \tag{1.137}$$

For states diagonal in the same basis with eigenvalues $\{e_i^{(1)}\}$ and $\{e_i^{(2)}\}$, it reduces to the classical Hellinger distance (Hellinger 1909):

$$D_{\mathrm{He}}(\rho_1, \rho_2) = \sqrt{2 - 2\sum_i \sqrt{e_i^{(1)} e_i^{(2)}}} \,. \tag{1.138}$$

Bures distance. The Bures distance (Bengtsson and Życzkowski 2017) is defined in terms of the Uhlmann fidelity $F(\rho_1, \rho_2)$ (Uhlmann 1976; Nielsen and Chuang 2002) as

$$D_{\mathrm{B}}(\rho_1, \rho_2) = \sqrt{2\,[1 - F(\rho_1, \rho_2)]}, \quad F(\rho_1, \rho_2) = \mathrm{Tr}\left[\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}\right] \tag{1.139}$$

For commuting states, the Bures and Hellinger distances coincide.

Trace distance. The trace distance is given by (Eq. (1.9))

$$D_{\mathrm{Tr}}(\rho_1, \rho_2) = \frac{1}{2}\|\rho_1 - \rho_2\|_1 = \frac{1}{2}\mathrm{Tr}\left[\sqrt{(\rho_1 - \rho_2)^2}\right] \,. \tag{1.140}$$

It quantifies the optimal error probability of distinguishing two states via the Helstrom measurement (see Sec. 2.2, is jointly convex, and is nonincreasing under CPTP maps (Bengtsson and Życzkowski 2017; Nielsen and Chuang 2002). Consequently, the corresponding measure $\mathcal{M}_{\mathrm{Tr}}(\rho)$ satisfies monotonicity under LOSR and convexity.

Relative entropy. Although not a true distance (as it is asymmetric and does not satisfy the triangle inequality), the quantum relative entropy remains one of the most useful quantities in resource theories (Nielsen and Chuang 2002). It is defined as

$$S(\rho_1\|\rho_2) = \mathrm{Tr}[\rho_1 \log_2 \rho_1 - \rho_1 \log_2 \rho_2] \,. \tag{1.141}$$

For states diagonal in the same basis, it reduces to the classical Kullback–Leibler

divergence (Kullback and Leibler 1951). The associated measure of Bell nonlocality,

$$\mathcal{M}_{\rm Re}(\rho) = \min_{\rho^{\rm loc}\in\mathcal{L}} S(\rho\|\rho^{\rm loc})\,, \tag{1.142}$$

is monotonic under free (LOSR) operations and under subselection, i.e.

$$\mathcal{M}_{\rm Re}(\rho) \geq \sum_i p_i\, \mathcal{M}_{\rm Re}(\rho_i)\,, \tag{1.143}$$

where $\{\Lambda_i\}$ are Kraus operators satisfying $\sum_i \Lambda_i^\dagger\Lambda_i = \mathbb{1}$, $p_i = {\rm Tr}(\Lambda_i\rho\Lambda_i^\dagger)$, and $\Lambda_i\rho^{\rm loc}\Lambda_i^\dagger \in \mathcal{L}$ for all $\rho^{\rm loc}\in\mathcal{L}$. The proof of last inequality is the same as the one obtained for the entropic measure of coherence (Vedral and Plenio 1998).

1.5.3 Application to two-qubit states

An arbitrary two-qubit state ρ can be expanded in terms ofn Pauli matrices σ_i as in Eq. (1.79). These coefficients can be arranged in matrix form as

$$T = \left[\begin{array}{cc} 1 & \vec{t}_{0i} \\ \vec{t}_{i0} & t \end{array}\right], \tag{1.144}$$

where all entries are real, $\vec{t}_{0i}$ is a row vector with components $\{t_{0i}\}$, $\vec{t}_{i0}$ is a column vector with components $\{t_{i0}\}$, and t is a 3×3 real matrix. As first result, we prove the following:

Result 1. *If ρ is a local state with given (t_{0i}, t_{i0}, t) then the state ρ^- given by $(-t_{0i}, -t_{i0}, t)$ is also local.*

Proof. It suffices to consider a local transformation that for $i = 1, 2, 3$ changes σ_i into $-\sigma_i$. □

Result 2. *For any local state ρ with given (t_{0i}, t_{i0}, A) there exists a local state ρ' with the same t but with maximally mixed reduced states, i.e., with $t_{0i} = t_{i0} = 0$*

Proof. Consider the convex combination

$$\rho' = \frac{1}{2}\rho + \frac{1}{2}\rho^-, \quad \text{for which:} \quad T = \left[\begin{array}{cc} 1 & 0 \\ 0 & t \end{array}\right]. \tag{1.145}$$

Now, convex combinations of local states are local and this ends the proof. □

Two particularly relevant families of states are the Bell-diagonal states, i.e.,

convex mixtures of the four Bell states, already discussed in section 1.4.4:

$$\rho_{\mathrm{BD}} = \sum_{i=1}^{4} e_i \left|\Psi_i\right\rangle \left\langle\Psi_i\right| = \left[\mathbb{1}_2 \otimes \mathbb{1}_2 + \sum_{i=1}^{3} t_i\, \sigma_i \otimes \sigma_i\right], \tag{1.146}$$

and the subclass known as Werner states, which are convex combinations of the maximally mixed state and the singlet Bell state.

For this class of states we prove the following two propositions, which hold for every scenario in which a set of Bell inequalities (just "Bell inequality" in the following) defines the convex set $\mathcal{L}$ of local states in that scenario:
Proposition 1. *For any Bell inequality, any jointly convex and unitarily invariant functional D of pairs of two-qubit density matrices*

$$\min_{\rho^{loc}\in\mathcal{L}} D(\rho_w, \rho^{loc}) = \min_{\rho_w^{loc}\in\mathcal{L}} D(\rho_w, \rho_w^{loc}), \tag{1.147}$$

where ρ_w is a Werner state, ρ^{loc} is a local state, and ρ_w^{loc} is a local Werner state.

Proof. The proof of Proposition 1 is shown in section 1.5.6 for isotropic states and generalized Werner states in higher dimensions. Here we just sketch the relevant steps. Starting from Eq. (1.135), we can apply the same $U \otimes U$ to both arguments of D and make use of the unitarily invariance of D; then, we bound this quantity from below by twirling both arguments and using jointly convexity; finally we recognize that twirling projects ρ_w itself and ρ^{loc} into another (local) Werner state ρ_w^{loc}. So $D(\rho_w, \rho_w^{loc}) \leq D(\rho_w, \rho^{loc})$, hence the closest state can be taken to be a Werner. □

Proposition 2. *For any Bell inequality, any jointly convex and unitarily invariant functional D of pairs of two-qubit density matrices, the closest local state to a Bell-diagonal state is Bell-diagonal:*

$$\min_{\rho^{loc}\in\mathcal{L}} D(\rho_{\mathrm{BD}}, \rho^{loc}) = \min_{\rho_{\mathrm{BD}}^{loc}\in\mathcal{L}} D(\rho_{\mathrm{BD}}, \rho_{\mathrm{BD}}^{loc}), \tag{1.148}$$

where ρ_{BD} denotes a Bell-diagonal state, ρ^{loc} a local state, and ρ_{BD}^{loc} a local Bell-diagonal state.

Proof.

The symmetry class that we consider contains the simultaneous local π-rotations of qubits around the three Pauli axes. Rotation R_1 around σ_1 changes σ_2 and σ_3 to $-\sigma_2$ and $-\sigma_3$, respectively. Rotations R_2 and R_3 operate analogously. These transformations are local and unitary, therefore, they cannot change the locality of a quantum state. Moreover, these transformations preserve Bell-diagonal states. Indeed, let us consider the general state representation Eq.

(1.79). Under rotation R_1, we have the following matrix transformation:

$$T \longrightarrow T_1 = \begin{pmatrix} 1 & t_{01} & -t_{02} & -t_{03} \\ t_{10} & a_{11} & -t_{12} & -t_{13} \\ -t_{20} & -t_{21} & t_{22} & t_{23} \\ -t_{30} & -t_{31} & t_{32} & t_{33} \end{pmatrix}. \tag{1.149}$$

The average $\rho' = 1/2(\rho + R_1(\rho))$ has the representation

$$T_1' = \begin{pmatrix} 1 & t_{01} & 0 & 0 \\ t_{10} & t_{11} & 0 & 0 \\ 0 & 0 & t_{22} & t_{23} \\ 0 & 0 & t_{32} & t_{33} \end{pmatrix}. \tag{1.150}$$

Applying rotation R_2 to ρ' and averaging the result with ρ' leads to the state with diagonal representation Eq. (1.146), i.e., a Bell-diagonal state. The total transformation that projects a given state to a Bell-diagonal state takes the compact form

$$\rho \longrightarrow \frac{1}{2}\left(\frac{1}{2}(\rho + R_1(\rho)) + R_2\left(\frac{1}{2}(\rho + R_1(\rho))\right)\right). \tag{1.151}$$

This transformation obviously preserves the Bell-diagonal states, does not change the locality, and is a composition of joint local unitaries and averaging.

The proof of Proposition 2 is thus immediately completed by straightforwardly adapting the reasoning followed in the proof of Proposition 3. □

Note that the previous result is useful for quantifying nonlocality of any two-qubit system with maximally-mixed subsystems, since those states can be transformed in Bell-diagonal states by means of local unitary transformations which do not change the nonlocality of the state.

Indeed, if one considers the singular value decomposition of the matrix A:

$$A = O^1 D O^2, \tag{1.152}$$

where O^1 and O^2 are orthogonal matrices and D is diagonal and non-negative. The orthogonal transformations correspond to local changes of basis:

$$\rho = \frac{1}{4}\Bigg(\mathbb{1}_4 + \sum_{i=1}^{3} t_{0i}\mathbb{1}_2 \otimes \sigma_i + \sum_{i=1}^{3} t_{i0}\sigma_i \otimes \mathbb{1}_2 + \sum_{k=1}^{3} D_k \Bigg(\sum_i O^1_{ik}\sigma_i\Bigg) \otimes \Bigg(\sum_j O^2_{kj}\sigma_j\Bigg)\Bigg). \tag{1.153}$$

1.5.4 CHSH-nonlocality of two-qubit states

As a first application, let us consider Werner and Bell-diagonal states.

1.5.5 Characterization of the CHSH nonlocality

Let (d_1, d_2, d_3) be the singular values in the decomposition Eq. (1.152), and without loss of generality, let $d_1 \geq d_2 \geq d_3$. As proved in Ref. (Horodecki et al. 1995), a state is CHSH-local, i.e., it does not violate the CHSH inequality Eq. (1.131), if and only if

$$d_1^2 + d_2^2 \leq 1. \tag{1.154}$$

Werner states

Hilbert-Schmidt distance Consider the Hilbert-Schmidt (HS) distance between a generic pair $\{\rho_1, \rho_2\}$ of two-qubit states as induced by the HS norm Eq. (1.136). Using Eq. (1.79) and the properties of the Pauli matrices:

$$\|\rho_1 - \rho_2\|_2 = \frac{1}{2} \|\alpha_1 - \alpha_2\|_2 . \tag{1.155}$$

It follows from Proposition 1, Eq. (1.147), that the closest CHSH-local state to a Werner state is also a Werner state with $w \leq 1/\sqrt{2}$. Hence, for CHSH-nonlocal Werner states, i.e., with $w \geq 1/\sqrt{2}$:

$$\mathcal{M}_{\text{HS}}(\rho_w) = \min_{\rho_w^{loc} \in \mathcal{L}} \left\| \rho_w - \rho^{loc} \right\|_2 = \frac{\sqrt{3}}{2} \left(w - \frac{1}{\sqrt{2}} \right) , \tag{1.156}$$

where $w \in \left(1/\sqrt{2}, 1\right]$. The maximum $\mathcal{M}_{\text{HS}}^{max} = \frac{\sqrt{3}}{2} \left(1 - \frac{1}{\sqrt{2}}\right)$ is achieved by the maximally entangled Bell singlet ($w = 1$).

Hellinger distance Recalling again the consequences of Proposition 1, Eq. (1.147), the measure of CHSH nonlocality for Werner states based on the (squared) Hellinger distance, Eq. (1.138), reads

$$\mathcal{M}_{\text{He}}(\rho_w) = 2 - \frac{1}{2} \left(3\sqrt{(1-w)\left(1 - \frac{1}{\sqrt{2}}\right)} + \sqrt{(1+3w)\left(1 + \frac{3}{\sqrt{2}}\right)} \right) , \tag{1.157}$$

where $w \in (1/\sqrt{2}, 1]$. As $|a_1| = |a_2| + |a_3| = w$, the signs of the matrix elements a_i identify the states corresponding to different corners of the state tetrahedron in Fig. 1.5. The Hellinger measure is maximized by the Bell singlet at $w = 1$ with $\mathcal{M}_{\text{He}}^{max} = 2 - \sqrt{1 + \frac{3}{\sqrt{2}}}$:

Bures distance In the case of commuting operators, the Bures distance and the Hellinger distance coincide. This is indeed the case for Werner density matrices. Therefore, the Bures and Hellinger measures of CHSH nonlocality coincide on Werner states ρ_w:

$$\mathcal{M}_{\text{Bu}}(\rho_w) = \mathcal{M}_{\text{He}}(\rho_w)\,. \tag{1.158}$$

Trace distance The geometric measure of CHSH nonlocality of Werner states based on the trace distance Eq. (1.9) reads:

$$\mathcal{M}_{\text{Tr}}(\rho_w) = \min_{\rho_w^{loc} \in \mathcal{L}} D_{\text{Tr}}(\rho_w, \rho_w^{loc}) = \frac{3}{4}\left| w - \frac{1}{\sqrt{2}} \right| = \frac{3}{4}\left(w - \frac{1}{\sqrt{2}} \right)\,, \tag{1.159}$$

where the last equality follows from the fact that, without loss of generality, we can limit the analysis to the interval $w \in (1/\sqrt{2}, 1]$. The maximum value $\mathcal{M}_{\text{Tr}}^{max} = \frac{3}{4}\left(1 - \frac{1}{\sqrt{2}}\right)$ is achieved at $w = 1$.

Relative entropy It is straightforward to verify that for a nonlocal Werner state in the interval $w \geq 1/\sqrt{2}$, its relative entropy Eq. (1.141) with respect to the set $\mathcal{L}$ of local states is minimized by the local state with $w = 1/\sqrt{2}$, so that:

$$\begin{aligned}\mathcal{M}_{\text{Re}}(\rho_w) = {} & \frac{3(1-w)}{4}\log_2\left(\frac{3(1-w)}{4}\right) + \frac{1+3w}{4}\log_2\left(\frac{1+3w}{4}\right) + \\ & - \frac{3(1-w)}{4}\log_2\left(\frac{\sqrt{2}-1}{4\sqrt{2}}\right) - \frac{1+3w}{4}\log_2\left(\frac{\sqrt{2}+3}{4\sqrt{2}}\right)\,.\end{aligned} \tag{1.160}$$

The maximum value achieved by the Bell states at $w = 1$ reads

$$\mathcal{M}_{\text{Re}}^{max} = 2 - \log_2\left(1 + \frac{3}{\sqrt{2}}\right)\,. \tag{1.161}$$

All the discussed (normalized) measure of CHSH nonlocality for Werner states $\widetilde{\mathcal{M}}_{\text{D}} = \mathcal{M}_{\text{D}}(\rho_w)/\mathcal{M}_{\text{D}}^{max}$ as a function of the w parameter are reported in Fig. 1.8.

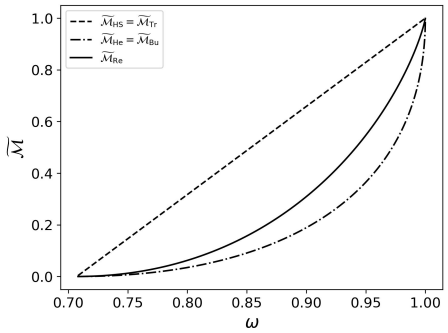


Figure 1.8: *Normalized geometric measures of CHSH nonlocality for Werner states as a function of the Werner parameter w in the interval $\frac{1}{\sqrt{2}} \leq w \leq 1$.*

Bell-diagonal states

As mentioned before, the results found in this section can be applied to any two-qubit state with maximally disordered subsystems, since they can be transformed into Bell-diagonal states by local unitary transformations that do not change their nonlocality.

Since Bell-diagonal states are diagonal in the Bell basis, the evaluation of the previously introduced distances is straightforward for this class of states. For a given state ρ, the closest local state can be identified through numerical analysis by exploiting the Horodecki criterion,

$$\max_{i,j\,,\,i\neq j} \{a_i^2 + a_j^2\} = 1\,. \tag{1.162}$$

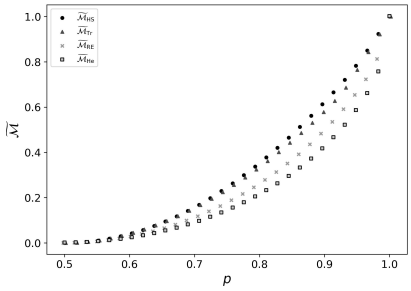
Here, let us consider Bell-diagonal states that are convex combinations of only two of the four Bell states. For instance, let us choose $e_2 = e_4 = 0$, $e_1 = p$, and $e_3 = 1 - p$, with p as a free parameter in the interval $[0, 1]$. Moving to T-space, this instance corresponds to $t_1 = a$, $t_2 = -a$, and $t_3 = 1$, with $a = 2p - 1$.

For this class of states, the locality condition Eq. (1.154) is fulfilled only for $p = \frac{1}{2}$; therefore, such states are nonlocal for $0 \leq p \leq 1$, with the exception of $p = \frac{1}{2}$. The behavior of the four normalized measures of CHSH nonlocality as functions of the convex combination parameter p in the half-interval $p \geq \frac{1}{2}$ is reported in Fig. 1.9a. We see that all measures vanish at $p = \frac{1}{2}$ and reach their maximum at $p = 1$, i.e., on the first Bell state. For all measures, the normalization coefficients turn out to be the same as the ones of the nonlocality measures for Werner states, in agreement with Proposition 1, Eq. (1.147), and Proposition 2, Eq. (1.148). The case $p \leq \frac{1}{2}$ is readily obtained by symmetry.

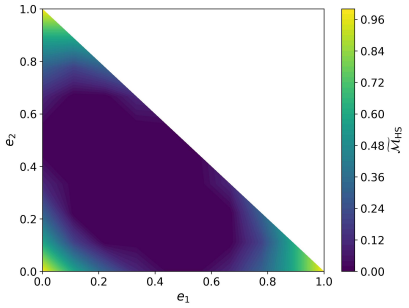
Clearly, the above-described behavior does not depend on the pair of Bell

states chosen as entries of the convex combination.

In the case of the Hilbert-Schmidt norm, Fig. 1.9b displays the corresponding nonlocality measure for a convex combination of three Bell states. For details on the numerical part, see (Zanfardino et al. 2026b) and https://github.com/2zanfardino/geometric_nonlocality.



(a) Normalized measures of CHSH nonlocality for the convex combination of two Bell states as functions of the convex combination parameter p in the interval $p \geq \frac{1}{2}$. The same behavior holds in the symmetric case $p \leq \frac{1}{2}$, with the maximum achieved in $p = 0$.



(b) Normalized Hilbert-Schmidt measure $\widetilde{\mathcal{M}}_{\mathrm{HS}}$ of CHSH nonlocality for Bell-diagonal states as a function of the two free parameters e_1 and e_2, with $e_4 = 0$, $e_3 = 1 - e_1 - e_2$, and normalization coefficient $\mathcal{M}_{\mathrm{HS}}^{max}$. The maximum $\widetilde{\mathcal{M}}_{\mathrm{HS}} = 1$ is reached in the three cases $(e_1, e_2, e_3) = (1, 0, 0)$; $(e_1, e_2, e_3) = (0, 1, 0)$; and $(e_1, e_2, e_3) = (0, 0, 1)$, each corresponding to a different Bell state.

Figure 1.9: *Comparison of CHSH-nonlocality geometric measures for convex combinations of (a) two Bell states and (b) three Bell states.*

1.5.6 Higher dimensions and multipartite systems

The CHSH scenario has been generalized to two-qudit systems, for which some relevant inequalities are the CGLMP inequalities (Masanes 2002; Collins et al. 2002) (number of measurements per party $m = 2$ and number of outcomes $\Delta = d$), and to an arbitrary number of qubits, for $m = 2$, $\Delta = 2$ (Mermin 1990; Belinskii and Klyshko 1993; Żukowski and Brukner 2002) and for arbitrary m (Laskowski et al. 2004). These inequalities are tight, and so they completely characterize the set of local correlations for the selected scenario. For more complex scenarios, while it's often the case that some Bell inequalities are known, they are generally not tight.

Bipartite isotropic states

An isotropic state has the form $\rho_{iso} = \omega \left|\phi^+\right\rangle \left\langle\phi^+\right| + \frac{1-\omega}{d^2} \mathbb{1}_d \otimes \mathbb{1}_d$, where $\left|\phi^+\right\rangle \left\langle\phi^+\right| = \sum_{ij=0}^{d-1} \left|ii\right\rangle \left\langle jj\right|$, and $-\frac{1}{d^2-1} \leq \omega \leq 1$ (Bertlmann and Friis 2023). In the case where we restrict to $0 \leq \omega \leq 1$, this can be interpreted as a mixture of a maximally entangled state affected by white noise. If $\omega > \frac{1}{d+1}$, ρ_{iso} is entangled.

For any local isotropic state, the following result holds:

Proposition 3. *For any Bell inequality and for any jointly convex and unitarily invariant functional D of pairs of two-qudit density matrices,*

$$\min_{\rho^{loc} \in \mathcal{L}} D(\rho_{iso}, \rho^{loc}) = \min_{\rho^{loc}_{iso} \in \mathcal{L}} D(\rho_{iso}, \rho^{loc}_{iso}), \tag{1.163}$$

where ρ_{iso} is an isotropic state, ρ^{loc} is a local state, and ρ^{loc}_{iso} is a local isotropic state.

Proof: The proof is based on the following observations:

1. Obs. 1: Invariance of isotropic states under $U \otimes U^*$ transformations:

$$(U \otimes U^*)\, \rho_{iso}\, (U^\dagger \otimes U^{*\dagger}) = \rho_{iso}, \quad \forall\, U \text{ unitary}. \tag{1.164}$$

2. Obs. 2: Local unitary transformations are resource non-generating (RNG) for nonlocality; in particular,

$$(U \otimes U^*)\, \rho^{loc}\, (U^\dagger \otimes U^{*\dagger}) \in \mathcal{L}, \tag{1.165}$$

   where $\mathcal{L}$ denotes the set of local states.

3. Obs. 3: Any convex combination of local states is itself a local state.

4. Any state ρ can be transformed into an isotropic state by applying the twirling operation:

$$\int dU\, (U \otimes U^*)\, \rho\, (U^\dagger \otimes U^{*\dagger}) = \rho_{iso}. \tag{1.166}$$

From the unitary invariance of D, we have

$$\begin{aligned} D(\rho_{iso}, \rho_{iso}^{loc}) &= D(V\,\rho_{iso}\,V^\dagger, V\,\rho^{loc}\,V^\dagger) \\ &= \int dV\, D(V\,\rho_{iso}\,V^\dagger, V\,\rho^{loc}\,V^\dagger) \\ &\geq D\left(\int dV\, V\,\rho_{iso}\,V^\dagger, \int dV\, V\,\rho^{loc}\,V^\dagger\right) \\ &= D(\rho_{iso}, \rho_{iso}^{loc})\,, \end{aligned} \tag{1.167}$$

where, in going from the third line to the fourth, we defined $V = U \otimes U^*$. Proposition 3 is general, since we do not specify any Bell inequality. In practice, for isotropic states, the set of local states that do not violate the CGLMP inequality is known (Collins et al. 2002; Masanes 2002). This inequality is violated when $\omega > \frac{2}{I_d(QM)}$, with

$$I_d(QM) = 4d \sum_{k=0}^{[d/2]-1} \left(1 - \frac{2k}{d-1}\right)(q_k - q_{-(k+1)}) \tag{1.168}$$

and

$$q_k = \frac{1}{2d^3 \sin^2\left[\frac{\pi(k+\frac{1}{4})}{d}\right]}\,. \tag{1.169}$$

When $m = 2$ and $\Delta = d = 2$, the only facet inequality reduces to CHSH. This inequality is violated when $\omega > \frac{1}{\sqrt{2}}$. For $m = 2$ and $\Delta = d = 3$, the facet inequalities correspond to the CGLMP inequality, which is violated when $\omega > (6\sqrt{3} - 9)/2$.

Here, we present the CGLMP nonlocality measures for isotropic states $\rho_{iso} = \omega\,|\phi^+\rangle\,\langle\phi^+| + \frac{1-\omega}{d^2}\mathbb{1}_d \otimes \mathbb{1}_d$, for the different considered distances and assuming two measurements per party with d outcomes:

$$\mathcal{M}_{\text{HS}}(\rho_{iso}) = \sqrt{1 - \frac{1}{d^2}}\left(\omega - \frac{2}{I_d(QM)}\right) \tag{1.170}$$

$$\mathcal{M}_{\text{Tr}}(\rho_{iso}) = \frac{2(d^2-1)}{d^2}\left(\omega - \frac{2}{I_d(QM)}\right) \tag{1.171}$$

$$\mathcal{M}_{\text{He}} = 2 - \frac{2}{d}\left[(d^2-1)\sqrt{(1-\omega)\left(1-\frac{2}{I_d(QM)}\right)}\right.$$
$$\left. + \sqrt{\left[(d^2-1)\omega+1\right]\left[(d^2-1)\frac{2}{I_d(QM)}+1\right]}\right] \tag{1.172}$$

$$\mathcal{M}_{\text{Re}} = \frac{(d^2-1)\omega+1}{d^2}\log\left(\frac{(d^2-1)\omega+1}{d^2}\right) + \frac{d^2-1}{d^2}\log\left(\frac{1-\omega}{d^2}\right)$$
$$+ \frac{(d^2-1)\frac{2}{I_d(QM)}+1}{d^2}\log\left(\frac{(d^2-1)\frac{2}{I_d(QM)}+1}{d^2}\right)$$
$$+ \frac{d^2-1}{d^2}\log\left(\frac{1-\frac{2}{I_d(QM)}}{d^2}\right) \tag{1.173}$$

Bipartite and multipartite scenarios

Note that the results of the previous propositions also hold for generalized Werner states, defined as those states that are invariant under local unitaries $V \equiv \bigotimes_{i=1}^{n} U$ in arbitrary dimensions and number of parties:

Proposition 4. *For any Bell inequality and for any jointly convex and unitarily invariant functional D, the closest local state to a Werner state is a Werner state.*

The proof follows directly for the same reasons discussed above:

On the other hand, unlike the case of isotropic states, the precise value of the parameter that separates the nonlocal Werner states from the local ones is at present unknown, also in the two-qudit case. As a consequence, considering for instance states that are assured to be either CGLMP-local or CGLMP-nonlocal, our quantifiers provide, respectively, upper and lower bounds on the exact geometric measure of CGLMP nonlocality.

Examples of Bell inequalities are also known in multipartite systems (Mermin 1990; Belinskii and Klyshko 1993). In the case of n qubits and two dichotomic measurements per party ($m = 2$, $\Delta = 2$), a complete characterization of all possible Bell inequalities has been obtained (Żukowski and Brukner 2002).

Nevertheless, for $n > 2$, Werner states are no longer simply given by a convex combination of their symmetric and antisymmetric parts (Eggeling and Werner 2001) and require separate analysis which will be provided in subsequent works.

1.6 Discussion and outlook

In this chapter, after briefly introducing quantum resource theories in Section 1.1, with particular emphasis on the resource theories of coherence and entanglement and some of their applications (Sections 1.2 and 1.3), we have established in Section 1.4 a fundamental link between uncertainty, purity, and entanglement in multipartite quantum systems. By deriving an exact uncertainty relation for a fixed k-partition, we have shown that the total uncertainty of tensor products of $\mathrm{SU}(N_{A_l})$ observables can be fully expressed in terms of the purities of the global and reduced states, revealing a conservation law connecting uncertainty and purity. This leads to a generalized Schrödinger-type uncertainty relation and to a purity-based separability criterion providing a sufficient condition for entanglement.

In parallel, in Section 1.5, we introduced a geometric framework for the quantification of Bell nonlocality in arbitrary scenarios, grounded in the distances between quantum states and the set of local states in Hilbert space. This approach places nonlocality on the same conceptual footing as other quantum resources, such as coherence, entanglement, and discord, allowing for direct comparisons through geometric and entropic measures. We analyzed the resulting structure for different classes of states and inequalities, including Werner and isotropic states, and outlined the general features of a geometric theory of multipartite nonlocality.

Overall, the results presented in this chapter contribute to the broader framework of quantum resource theories by providing unified geometric and information-theoretic perspectives on fundamental quantum resources, particularly nonlocality and entanglement. The concepts and methods developed here form a theoretical foundation that will be instrumental in the following chapters, where we explore their applications to quantum engineering and emerging quantum technologies.

Chapter 2

Applications of quantum resources for quantum engineering and technologies

2.1 Introduction

In this chapter, we analyze, through practical examples, how quantum resource theories can be employed to develop high-performance technologies that, in several cases, surpass their classical counterparts.

In Section 2.2, we introduce the problem of *quantum state discrimination*, discussing the challenges associated with distinguishing non-orthogonal states and highlighting the limitations of existing discrimination protocols.

Section 2.3 focuses on *continuous-variable systems*, with particular emphasis on the protocol of *quantum illumination* – a paradigmatic scenario in which genuinely quantum features, such as entanglement, enable a substantial enhancement in detecting low-reflectivity objects embedded in strong environmental noise, outperforming all classical strategies. Building on established results, we examine how *coherence* (implemented through displacement) and *entanglement* (introduced via squeezing) each contribute to improved detection performance. Our analysis places special emphasis on exploiting both resources simultaneously, using the two-mode coherent-squeezed state as a benchmark that naturally incorporates both coherence and entanglement.

Following this, inspired by the quantum illumination framework, Section 2.4.1 investigates how two-mode squeezed entangled states can enhance the discrimination between two different energy shifts in atomic systems, within the broader context of quantum metrology. We will show that, even in this case, quantum advantage is achieved.

Finally, in Section 2.5, we present an application of quantum discrimination

techniques to particle physics. In particular, we quantify the maximum amount of information that can be extracted when distinguishing Dirac from Majorana neutrinos by probing their intrinsic degrees of freedom during free evolution in time.

For further background on quantum discrimination and detection theory, the reader is referred to Appendix B.1.

2.2 Quantum discrimination

In classical mechanics, it is, in principle, always possible to distinguish between different states of a system by performing suitable measurements. In quantum mechanics, however, this is not generally the case: only orthogonal states can be perfectly distinguished with certainty (Nielsen and Chuang 2002). This limitation arises from the intrinsic probabilistic nature of quantum systems, which implies that non-orthogonal states can be distinguished only with a non-zero probability of error, depending on the choice of measurement observables.

Several criteria exist to determine the optimal observables for distinguishing between quantum states. One of the most widely used is the Helstrom criterion (Helstrom 1967; Helstrom 1969), which selects the measurement that minimizes the probability of error.

To briefly outline the Helstrom approach, suppose that Alice has a set of N quantum states $\{\rho_i\}$, which may also be mixed. She sends one of these states to Bob, who knows the set $\{\rho_i\}$ but not which specific state was sent. Bob must then perform a measurement, described by a POVM $\{M_m\}$, in order to minimize the probability of making an incorrect guess.

A first limitation of the Helstrom approach is that an exact solution exists only for the case of two alternatives, ρ_1 and ρ_2. In this scenario, the optimal measurement that Bob can perform to minimize the error probability is given by the projective measurement defined by $\{\hat{\gamma}^+, \mathbb{I} - \hat{\gamma}^+\}$, where $\hat{\gamma}^+$ denotes the projector onto the positive part of $\rho_1 - \rho_2$. Hence, Bob effectively measures the observable $\rho_1 - \rho_2$ and infers that the state sent by Alice is ρ_1 if the measurement outcome is positive, and ρ_2 if it is negative. According to this strategy, the error probability is given by (Helstrom 1967).

$$p^{err} = \frac{1}{2}\left(1 - \sum_k \lambda_k^+\right) = \frac{1}{2}\left(1 - \frac{1}{2}||\rho_1 - \rho_2||_1\right) \tag{2.1}$$

where the λ_k^+ are the positive eigenvalues of $\rho_1 - \rho_2$, $||\cdot||_1$ the trace norm.

In the case of $N > 2$ states, the Helstrom problem admits exact solutions only in specific instances (Eldar et al. 2004b; Ha and Kwon 2013), and in general, only approximate methods have been developed to estimate both the

optimal error probability and the corresponding measurement strategy (Rosati et al. 2017).

Another limitation of the Helstrom approach is that the optimal POVM is not always feasible to implement experimentally, since it is not always straightforward to construct the projector $\hat{\gamma}^+$

An alternative strategy for quantum state discrimination is known as *unambiguous discrimination* (Ivanovic 1987; Eldar et al. 2004a), which we briefly describe in the case of two pure states, $|\Psi_1\rangle$ and $|\Psi_2\rangle$. In this framework, the POVM $\{M_i\}$ must satisfy a no-error condition: if the measurement outcome is i, then the state sent must be $|\Psi_i\rangle$ with certainty. Formally, this translates to the constraint $p_{ij} = \langle\Psi_i| M_j |\Psi_i\rangle = 0$ for $i \neq j$.

Clearly, this is not achievable with a POVM consisting of only two elements M_1 and M_2, unless the two states are orthogonal (i.e., $\langle\Psi_1|\Psi_2\rangle = 0$). Otherwise, perfect discrimination would be possible, which contradicts the fundamental quantum limitation on distinguishing non-orthogonal states (Bergou 2007).

However, it is possible to introduce a third POVM element, M_3, corresponding to an *inconclusive outcome*, such that the no-error constraint $p_{ij} = 0$ for $i \neq j$ (with $i = 1, 2$) is satisfied. In this case, the price to pay is a non-zero probability of obtaining no information from the measurement about the input state. It has been shown that the optimal POVM – i.e., the one that minimizes the probability of inconclusive outcomes – yields a probability of inconclusive detection given by $p_{\text{inc}} = |\langle\Psi_1|\Psi_2\rangle|$ when the prior probabilities of the two states are equal (Dieks 1988; Peres 1988). This approach has also been generalized to more complex scenarios (Jaeger and Shimony 1995; Eldar et al. 2004a; Chefles and Barnett 1998).

Helstrom approach and unambiguous discrimination approach have been well reviewed in different works (Bergou 2007; Bae and Kwek 2015).

Two-qubits discrimination - Helstrom approach

Consider the case in which Alice has two qubits

$$|\Psi_1\rangle = |0\rangle\,, \quad |\Psi_2\rangle = \cos\theta\,|0\rangle + \sin\theta\,|1\rangle \tag{2.2}$$

The Helstrom error probability formula (2.1) can be easily computed since

$$\rho_1 - \rho_2 = \begin{bmatrix} \sin^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix} \tag{2.3}$$

which has eigenvalues $\lambda_{\pm} = \sin^2\theta \pm \sqrt{\frac{\sin 2\theta}{2}}$ and hence the error probability is

$$p_{\text{err}}^{Hel} = \frac{1}{2}\left[1 - \frac{1}{2}\left(\sin^2\theta + \sqrt{\frac{\sin 2\theta}{2}} + \left|\sin^2\theta - \sqrt{\frac{\sin 2\theta}{2}}\right|\right)\right] \tag{2.4}$$

Hence, if Alice sends m copies of the state, the Helstrom error probability is $(p_{\text{err}}^{Hel})^m$.

2.3 Quantum illumination

2.3.1 Introduction

In this section, we discuss a well-known protocol in which entanglement is used as a resource for distinguishing between non-orthogonal quantum states. It is known that entanglement can be advantageous for such discrimination tasks (Lloyd 2008; Tan et al. 2008), and also for distinguishing between different quantum operations (Sacchi 2005). In the previous sections, we introduced the Helstrom criterion for optimal quantum state discrimination between two alternatives, which is discussed in more detail in Appendix B.1.

Building on these concepts and inspired by the work of Sacchi (Sacchi 2005), it has been shown that the use of entangled photons can significantly reduce the minimum error probability in the task of detecting a low-reflectivity object, i.e., determining whether the object is present or absent (Lloyd 2008; Tan et al. 2008). This advantage persists even in the presence of interaction with a thermal bath.

More recently, it has been demonstrated that the optimal input state for detecting the presence or absence of an object is a two-mode Gaussian squeezed state (De Palma and Borregaard 2018) of the form

$$|\Psi\rangle_{\mathcal{SI}} = \frac{1}{\cosh r}\sum_n (-1)^n e^{in\theta}(\tanh r)^{2n}\,|n\rangle_{\mathcal{S}}\,|n\rangle_{\mathcal{I}} \tag{2.5}$$

where $\xi \equiv re^{i\theta}$ is the squeezing parameter, and the detection efficiency increases with the degree of entanglement. This type of detection scheme is known as *quantum illumination*. In quantum illumination protocols, two entangled modes are employed: the signal mode $\mathcal{S}$, whose photons are sent into the region where the object may be present (hypothesis H^1) or absent (hypothesis H^0), and the idler mode $\mathcal{I}$, whose photons are retained by the observer for a joint measurement performed after the interaction.

Since the exact formula for the error probability in Eq. (2.1) is generally not

computable, upper bounds such as the Bhattacharyya and Chernoff bounds have been adopted (Pirandola and Lloyd 2008; Tan et al. 2008).

An experimental implementation of quantum illumination has also been proposed (Lopaeva et al. 2013).

Eq. (2.1) depends on, the trace distance between the two possible output states appears. As expected, the larger the distance between the states, the lower the probability of error in discriminating them. Since Gaussian states play a fundamental role in quantum illumination and an exact formula exists for computing their fidelity (Banchi et al. 2015), one can also use the Bures distance (see Eqs. (1.10) and (1.11), and (Bures 1969)) as a measure of distinguishability:

$$D_{\mathrm{B}} = \sqrt{2\big[1 - F(\rho_0, \rho_1)\big]}. \tag{2.6}$$

For any pair of density operators ρ and σ, Fuchs-van de Graaf inequality holds (Fuchs and Van De Graaf 2002; Nielsen and Chuang 2002):

$$1 - \sqrt{F\left(\rho, \sigma\right)} \leq D_{\mathrm{Tr}}\left(\rho, \sigma\right) \leq \sqrt{1 - F\left(\rho, \sigma\right)}. \tag{2.7}$$

Based on these considerations, and on the fact that

$$F(\rho_0, \rho_1) = \min_{\{M_m\}} \sum_m \sqrt{p_0(m) p_1(m)}, \tag{2.8}$$

where $p_i(m) = \mathrm{Tr}[M_m \rho_i]$ is the probability of outcome m under the measurement operator M_m for $i = 0, 1$, the possibility of distinguishing single-mode coherent states via homodyne detection has been investigated (Nha and Carmichael 2005). The minimum in Eq. (2.8) is taken over all possible sets of POVMs $\{M_m\}$. Homodyne detection, which maps Gaussian states into Gaussian states, provides an upper bound for the fidelity.

Recently, it has been shown that using homodyne detection with an initially squeezed state does not improve the detection efficiency compared to the optimal scheme with a classical coherent state. Consequently, no quantum advantage is observed in this case (Shapiro 2020).

In this work, we consider as input a two-mode coherent squeezed state. In this case, both resources - coherence and entanglement - are present. By fixing the average number of signal photons and tuning the free parameters that characterize the amount of both resources, this state will serve as a benchmark for studying the contribution of the two resources to the detection task. This approach is also experimentally relevant, as only a limited amount of squeezing can be achieved in practice. Although squeezing is known to be the optimal resource for this task, experimental constraints impose a trade-off. Therefore, once the

maximum achievable squeezing is reached, one may consider adding coherence as a supplementary resource to further improve detection performance.

We will analyze both the optimal discrimination strategy (based on the Helstrom scheme) and homodyne detection, which, although suboptimal, is significantly simpler to implement in experimental setups.

2.3.2 Quantum Chernoff bound for Gaussian states

As already mentioned, in general it is difficult to determine the spectrum of the operator $\rho_1 - \lambda\rho_0$ when ρ_0 and ρ_1 do not commute. In the previous section, it was shown that the optimal strategy for choosing between the two hypotheses is to perform a measurement of the operator $\rho_1 - \lambda\rho_0$ and to choose hypothesis H_1 when the measurement outcome lies in a subspace associated with positive eigenvalues, and hypothesis H_0 otherwise.

In realistic scenarios, the observer is usually able to prepare N independent copies of the system under examination and the error probability decreases exponentially with N as we showed in the previous section.

Calculating this quantity is generally a difficult task. However, there exist useful upper bounds that can be more easily computed in many cases, such as the *quantum Chernoff bound*, which is given by (Chernoff 1952)

$$P^{(N)} \leq P_{QC}^{(N)} = \frac{1}{2}e^{-kN} \tag{2.9}$$

where

$$k \equiv -\log\Big\{ \inf_{0\leq s\leq 1} \mathrm{Tr}[\rho_0^s \rho_1^{1-s}] \Big\} \tag{2.10}$$

which can be rewritten as

$$P_{QC}^{(N)} = \frac{1}{2}\left[\inf_{0\leq s\leq 1} Q_s \right]^N, \qquad \text{where } Q_s \equiv \mathrm{Tr}[\rho_0^s \rho_1^{1-s}], \tag{2.11}$$

and the *"Battacharyya bound"* that is obtained with $s = \frac{1}{2}$.

It is not always possible compute the quantum Chernoff bound. However, in (Pirandola and Lloyd 2008), a closed-form expression for Gaussian states has been derived. This result is particularly useful, as Gaussian states can be efficiently implemented in a wide range of experimental setups.

2.3.3 Gaussian states

In the theory of continuous variable (CV) systems, the state of the system is described by a density operator $\rho \in \mathcal{D}(\mathcal{H}^{\otimes n})$, where $\mathcal{D}(\mathcal{H}^{\otimes n})$ denotes the space of density operators defined on the tensor product of n continuous Hilbert spaces $\mathcal{H}$ (e.g., a system of n particles described by their positions and momenta). In general, such systems can be quite complex, but their analysis can be significantly simplified by studying their properties in the associated symplectic space (Adesso and Illuminati 2007; Ferraro et al. 2005).

To this end, one can define the *characteristic function* as

$$\chi(\vec{\xi}) = \mathrm{Tr}[\rho\, \hat{D}(\vec{\xi})]\,, \qquad \text{where } \hat{D}(\vec{\xi}) \equiv e^{i\vec{\hat{x}}^T\cdot\vec{\xi}}\,. \tag{2.12}$$

$\hat{D}(\vec{\xi})$ is called Weyl operator, with $\vec{\xi} \in \mathbb{R}^{2n}$, and $\vec{\hat{x}}^T = (\hat{q}_1, \hat{p}_1, \ldots, \hat{q}_n, \hat{p}_n)$ is the vector of quadrature operators representing the position $\hat{q}_i$ and momentum $\hat{p}_i$ (each rescaled by a factor $\sqrt{2}$) of the i-th mode in the Hilbert space $\mathcal{H}^i$.

By definition, the components of $\vec{\hat{x}}^T$ satisfy the (non-normalized) canonical commutation relations (with $\hbar = 1$):

$$[\hat{x}_l, \hat{x}_m] = 2i\,\Omega_{lm}\,, \qquad \text{with}\, \Omega \equiv \bigoplus_{k=1}^{n} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{2.13}$$

is the *symplectic form*.

The Fourier transform of the characteristic function is the Wigner function:

$$\mathcal{W}(\vec{x}) = \int_{\mathbb{R}^{2n}} \frac{d^{2n}\xi}{(2\pi)^n}\, e^{-i\vec{x}\cdot\vec{\xi}}\chi(\vec{\xi}). \tag{2.14}$$

The characteristic function fully characterizes the quantum state described by ρ. When the characteristic function (or equivalently the Wigner function) is Gaussian, the state ρ is called a *Gaussian state*. In this case, the characteristic function takes the form:

$$\chi(\vec{\xi}) = \exp\left(-\frac{1}{2}\vec{\xi}^T V \vec{\xi} + i\,\bar{x}\cdot\vec{\xi}\right), \tag{2.15}$$

and the Wigner function becomes:

$$\mathcal{W}(\vec{x}) = \frac{1}{(2\pi)^n\sqrt{\det V}} \exp\left[-\frac{1}{2}(\vec{x}-\bar{x})^T V^{-1}(\vec{x}-\bar{x})\right]. \tag{2.16}$$

Therefore, to specify the Gaussian state ρ, it is sufficient to specify the

displacement vector

$$\bar{x} \equiv \mathrm{Tr}[\rho\vec{\hat{x}}], \tag{2.17}$$

and the covariance matrix (CM) σ, defined by:

$$\sigma_{lm} = \frac{1}{2}\mathrm{Tr}\left[\{\Delta\hat{x}_l, \Delta\hat{x}_m\}\rho\right] = \frac{1}{2}\left(\langle\hat{x}_l\hat{x}_m\rangle + \langle\hat{x}_m\hat{x}_l\rangle\right) - \langle\hat{x}_l\rangle\langle\hat{x}_m\rangle, \tag{2.18}$$

where $\{\cdot,\cdot\}$ denotes the anticommutator and $\langle\hat{O}\rangle \equiv \mathrm{Tr}[\hat{O}\rho]$ denotes the expectation value of operator $\hat{O}$.

The uncertainty principle in Eq. (2.13) can be recast in terms of the covariance matrix as:

$$\sigma + i\Omega \geq 0, \qquad \sigma \geq 0. \tag{2.19}$$

A matrix S is called *symplectic* if it preserves the symplectic form $S\Omega S^T = \Omega$.

Theorem (Williamson): *For every positive-definite covariance matrix V, there exists a symplectic matrix S such that*

$$\sigma = S\left(\bigoplus_{k=1}^{n}\begin{bmatrix}\nu_k & 0\\ 0 & \nu_k\end{bmatrix}\right)S^T, \tag{2.20}$$

where the set $\{\nu_1, \ldots, \nu_n\}$ is called the symplectic spectrum *of V.*

Since $\det S = \pm 1$ and $\sigma \geq 0$, it follows that:

$$\prod_{k=1}^{n}\nu_k = \sqrt{\det\sigma}. \tag{2.21}$$

Choosing σ in diagonal form and applying the uncertainty relation in Eq. (2.19) yields the constraint $\nu_k \geq 1$.

An affine transformation of the quadrature operators,

$$\vec{\hat{x}} \to S\vec{\hat{x}} + \bar{x}, \tag{2.22}$$

induces the following transformation on the covariance matrix:

$$\sigma \to S\sigma S^T, \tag{2.23}$$

and a corresponding unitary transformation in the space of density operators $\mathcal{D}(\mathcal{H}^{\otimes n})$:

$$\rho \to \hat{U}_{\bar{x},S}\,\rho\,\hat{U}^{\dagger}_{\bar{x},S}, \tag{2.24}$$

where the unitary $\hat{U}_{\bar{x},S}$ is fully determined by the affine parameters $(\bar{x}, S)$ and preserves the Gaussian character of the state.

Using Williamson's theorem, one can show that any Gaussian state ρ can be written in the so-called *normal mode decomposition*:

$$\rho = \hat{U}_{\bar{x},S} \left(\bigotimes_{k=1}^{n} \rho_T(\nu_k) \right) \hat{U}_{\bar{x},S}^{\dagger}, \tag{2.25}$$

where

$$\rho_T(\nu_k) \equiv \frac{2}{\nu_k + 1} \sum_{j=0}^{\infty} \left(\frac{\nu_k - 1}{\nu_k + 1} \right)^j |j\rangle_k \langle j| \tag{2.26}$$

is a thermal state with average photon number $\bar{n}_k = \frac{\nu_k - 1}{2}$, and $|j\rangle_k$ are the number states associated with the k-th mode.

Note that the advantage of working in the symplectic space is twofold. First, it allows for the description of infinite-dimensional CV systems using only a finite number of parameters. For instance, a Gaussian state is fully characterized by its displacement vector $\bar{x}$ and its covariance matrix σ. Second, in the symplectic framework, tensor products of subsystems in Hilbert space correspond to direct sums in phase space. This structural simplification means that the properties of individual subsystems can be directly inferred by examining the submatrices of σ associated with the corresponding subspaces. Consequently, the analysis of multipartite systems becomes more transparent and computationally tractable.

Computable formula for Chernoff bound in Gaussian states

In order to obtain a computable bound for the discrimination of two Gaussian states, let us define

$$G_s(x) \equiv \frac{2^s}{(x+1)^s - (x-1)^{1-s}} \tag{2.27}$$

$$\Lambda_s(x) \equiv \frac{(x+1)^s + (x-1)^{1-s}}{(x+1)^s - (x-1)^{1-s}} \tag{2.28}$$

It holds the following result (Pirandola and Lloyd 2008):

Theorem *For two arbitrary bosonic $n-$mode Gaussian states ρ_A and ρ_B with normal mode decomposition defined by $(\overline{x_A}, S_A, \{\alpha_k\})$ and $(\overline{x_B}, S_B, \{\beta_k\})$*

respectively, an upper bound Q_s for the discrimination of the two states is

$$Q_s = \bar{Q}_s e^{-\frac{1}{2}\vec{d}^T[V_A(s)+V_B(1-s)]\vec{d}}, \quad \bar{Q}_s \equiv \frac{2^n \prod_{k=1}^{n} G_s(\alpha_k) G_{1-s}(\beta_k)}{\sqrt{\det[V_A(s)+V_B(1-s)]}} \tag{2.29}$$

with $\vec{d} \equiv \overline{x_A} - \overline{x_B}$, where

$$\begin{aligned} V_A(s) &\equiv S_A \left[\bigoplus_{k=1}^{n} \Lambda_s(\alpha_k)\mathbb{I}^2\right] S_A^T \\ V_B(1-s) &\equiv S_B \left[\bigoplus_{k=1}^{n} \Lambda_{1-s}(\beta_k)\mathbb{I}^2\right] S_B^T \end{aligned} \tag{2.30}$$

2.3.4 Quantum illumination: general framework

The main goal of quantum illumination is to determine whether an object with reflection coefficient $\kappa \ll 1$ is present or not, by sending optical signals into a region inside a cavity where a noisy electromagnetic field is present.

If the object is absent, each signal photon thermalizes, so that only bath photons are detected when measuring the electromagnetic field. Conversely, if the object is present, a small fraction κN of the photons is reflected by the object and therefore does not thermalize.

As a result, the two possible outcomes are described by two different density operators, ρ_1^{out} and ρ_0^{out}, corresponding respectively to the presence or absence of the object. According to the Helstrom criterion, the minimal probability of error is given by Eq. (2.1).

It has been shown that, by using entangled photons prepared in a two-mode squeezed input state (2.5), the probability of error significantly decreases with the number of copies sent (Tan et al. 2008), compared to the case of a coherent input state.

A schematization of quantum illumination protocol is proposed in Figure 2.1.

More recently, it has also been demonstrated that for quantum illumination, the two-mode squeezed vacuum state (2.5) represents the optimal choice for minimizing the error probability (2.1) (De Palma and Borregaard 2018), which performs better than even non-Gaussian resources.

Here, we consider a different input state: the coherent-squeezed state, in which both resources - coherence and entanglement - are present. This state will serve as a benchmark for studying the interplay between the two resources in different regimes of strong and weak noise and in two kind of detection: the

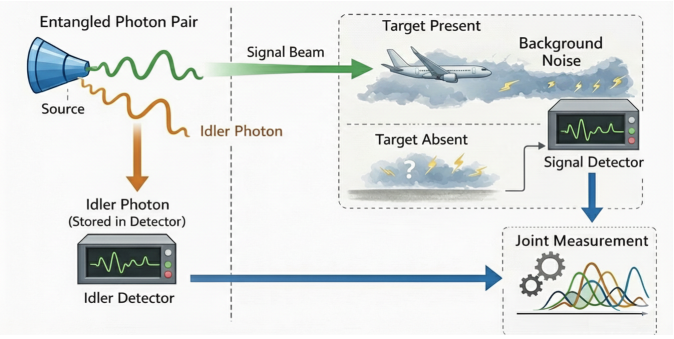


Figure 2.1: *Quantum illumination scheme: a two-mode squeezed state of light is generated. The signal photons are sent into a region where the object may or may not be present, while the idler photons are retained. After the interaction, a joint measurement is performed on the returned light and the retained idler.*

homodyne and the optimal Helstrom scheme-based.

As we have already mentioned, this analysis is not only of theoretical interest, but also relevant for practical applications. In fact, although it is known that the two-mode squeezed state is the optimal resource for quantum illumination, in practice it is not easy to achieve high levels of squeezing (Mizuno et al. 2005; Eberle et al. 2013).

Therefore, once the maximum achievable squeezing is reached, one can consider introducing additional coherence via displacement in order to further enhance the detection scheme.

In the Helstrom formula (2.1), the trace distance between the two possible outcomes appears, and it is clear that the error probability decreases as the distance between the two states increases.

Considering the Bures distance (2.6) instead of the trace distance, one similarly expects that it is easier to discriminate between states that are farther apart. An advantage of using the Bures distance for Gaussian states is that there exists an explicit formula to calculate the fidelity between two Gaussian states (Banchi et al. 2015). Hence, it can be exactly computed, while for Eq. (2.1) only lower and upper bounds are typically available.

By using the inequality

$$F(\rho_0, \rho_1) \leq \int_{-\infty}^{+\infty} \sqrt{p_1(q_\phi) p_2(q_\phi)}\, dq_\phi, \tag{2.31}$$

where $p_i(q_\phi)$ are the outcome probabilities for homodyne detection under the two alternatives, one can compare the fidelity of the two states with the overlap of the outcome probabilities appearing on the right-hand side of (2.31).

The closer this value is to the fidelity, the better we expect the states can be

discriminated.

The thermal bath at fixed temperature $T = \frac{1}{\beta}$ and frequency ω is described by the state

$$\rho_{\text{th}} = \frac{1}{Z}\sum_{n=0}^{\infty} e^{-n\beta\omega} \left|n\right\rangle_{\mathcal{B}} \left\langle n\right|, \quad \text{where } \frac{1}{Z} \equiv 1 - e^{-\beta\omega}, \tag{2.32}$$

and $\left|n\right\rangle_{\mathcal{B}} = (\hat{a}_{\mathcal{B}}^{\dagger})^n \left|0\right\rangle_{\mathcal{B}} / \sqrt{n!}$, with $\hat{a}_{\mathcal{B}}^{\dagger}$ the creation operator for a thermal photon.

Although a physical thermal bath is intrinsically multimode, it is common to model it as an effective single bosonic mode at frequency ω. This approximation is justified because the signal interacts only with bath modes within its detection bandwidth, whose collective effect can be represented by a single thermal mode with mean photon number $N_{\mathcal{B}} = (e^{\beta\omega} - 1)^{-1}$. The approximation breaks down for broadband signals, in the presence of strong dispersion or non-Markovian bath dynamics, or whenever correlations among different bath modes become relevant.

Hence, it has zero displacement and a covariance matrix

$$\sigma_{\text{th}} = \begin{bmatrix} B & 0 \\ 0 & B \end{bmatrix}, \tag{2.33}$$

where $B \equiv 2N_{\mathcal{B}} + 1$ and $N_{\mathcal{B}} = \text{Tr}[\rho_{\text{th}}\, \hat{a}_{\mathcal{B}}^{\dagger}\hat{a}_{\mathcal{B}}]$ is the mean number of photons.

If one starts with an initial state ρ^{in}, the interaction with the thermal bath is modeled by a beam splitter transformation acting on the composite system formed by the signal $\mathcal{S}$ and the thermal bath $\mathcal{B}$. This interaction is described by the unitary operator (Giovannetti et al. 2004)

$$\hat{U}_{\kappa_i} \equiv e^{\tan^{-1}\left(\sqrt{\frac{1-\kappa_i}{\kappa_i}}\right)[\hat{a}_{\mathcal{B}}\hat{a}_{\mathcal{S}}^{\dagger} - \hat{a}_{\mathcal{B}}^{\dagger}\hat{a}_{\mathcal{S}}]} \qquad i = 0, 1, \tag{2.34}$$

where $\hat{a}_{\mathcal{S}}$ and $\hat{a}_{\mathcal{S}}^{\dagger}$ are the annihilation and creation operators for the signal photons, respectively. Here, $\kappa_0 = 0$ when the object is absent and $\kappa_1 = \kappa$ when the object is present.

This leads to the transformations

$$\begin{aligned} \hat{a}_{\mathcal{S}} \quad &\rightarrow \quad \hat{U}_{\kappa_i}^{\dagger}\hat{a}_{\mathcal{S}}\hat{U}_{\kappa_i} = \sqrt{\kappa_i}\hat{a}_{\mathcal{S}} + \sqrt{1-\kappa_i}\hat{a}_{\mathcal{B}} \\ \hat{a}_{\mathcal{B}} \quad &\rightarrow \quad \hat{U}_{\kappa_i}^{\dagger}\hat{a}_{\mathcal{B}}\hat{U}_{\kappa_i} = \sqrt{\kappa_i}\hat{a}_{\mathcal{B}} - \sqrt{1-\kappa_i}\hat{a}_{\mathcal{S}} \end{aligned} \tag{2.35}$$

After the interaction, by tracing out the thermal bath, we obtain two possible output states, ρ_0^{out} and ρ_1^{out}. We can then perform homodyne detection, resulting in two possible probability distributions $p_0(q_\phi)$ and $p_1(q_\phi)$, which, in the case

of ideal single-mode homodyne detection, can be calculated as (Ferraro et al. 2005)

$$p_i(q_\phi) = \int W_i(q,p)\,\delta(q-q_\phi)\,dq\,dp = \int W_i(q_\phi,p)\,dp, \qquad i=0,1, \tag{2.36}$$

where $W_i(q,p)$ are the Wigner functions corresponding to the two possible output states.

In the case of $n-$mode Gaussian state with displacement $\bar{\mathbf{x}}$ and covariance matrix σ the Wigner function is given by

$$W(\mathbf{x}) = \frac{e^{-\frac{1}{2}(\mathbf{x}-\bar{\mathbf{x}})^T\sigma^{-1}(\mathbf{x}-\bar{\mathbf{x}})}}{(2\pi)^n\sqrt{\det[\sigma]}} \tag{2.37}$$

If we consider the possibility of non unitary efficiency of the detector the formula (2.36) turns into

$$p_i(q_\phi) = \int W_i(q,p)W_\eta(q)dqdp\,, \quad \text{with } W_{\delta_\eta}(q) = \frac{1}{\sqrt{2\pi}\delta_\eta^2}e^{-\frac{(q-q_\phi)^2}{2\delta_\eta^2}}\,, \tag{2.38}$$

where δ_η is a parameter related to the efficiency of the detector.

In the following sections, fixing the number N_S of signal photons sent, we will compute the probabilities $p_i(q_\phi)$ starting respectively from a coherent state and a two-mode squeezed state.

We will then compare the overlap of the two spectra with the fidelity according to inequality (2.31), analyzing how they differ by tuning the two parameters associated with squeezing and coherence.

For Gaussian states ρ_0 and ρ_1 with displacements $\bar{x}_0$ and $\bar{x}_1$ and correlations matrices CM σ_0 and σ_1, there is a formula for calculating the fidelity given by (Banchi et al. 2015)

$$F(\rho_0,\rho_1) = F_0(\sigma_0,\sigma_1)e^{-\frac{1}{4}\delta x^T(\sigma_0+\sigma_1)^{-1}\delta x} \tag{2.39}$$

where $\delta x = \bar{x}_1 - \bar{x}_0$.

$$F_0(\sigma_0,\sigma_1) = \frac{F^{tot}}{\sqrt[4]{\det[\sigma_0+\sigma_1]}}\,, \quad F_{tot} = \prod_{k=1}^{N}\left[w_k^{aux} + \sqrt{(w_k^{aux})^2-1}\right]^{\frac{1}{2}} \tag{2.40}$$

with w_k^{aux} eigenvalues of the matrix

$$W^{aux} = -2\sigma^{aux} i\Omega\,, \quad \sigma^{aux} = \Omega^T\left(\sigma_0 + \sigma_1^{-1}\right)\left(\frac{\Omega}{4} + \sigma_2\Omega\sigma_1\right) \tag{2.41}$$

and

$$\Omega = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \otimes \mathbb{I}_N \tag{2.42}$$

is the symplectic form.

2.3.5 Coherent state

A coherent state is obtained by the application of the displacement operator $D(\alpha)$ to the vacuum state:

$$|\alpha\rangle_{\mathcal{S}}^{in} = D(\alpha)\,|0\rangle_{\mathcal{S}} \equiv e^{i\left(\alpha\hat{a}_{\mathcal{S}}^{\dagger} - \alpha^*\hat{a}_{\mathcal{S}}\right)}\,|0\rangle_{\mathcal{S}} = e^{-\frac{|\alpha|^2}{2}}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\,|n\rangle_{\mathcal{S}} \tag{2.43}$$

with $\alpha \in \mathbb{C}$ and $|\alpha|^2 = N_{\mathcal{S}}$ is the mean number of signal photons.

This state is characterized by a displacement and covariance matrix

$$\bar{x}_{coh}^{in} = 2(\mathrm{Re}[\alpha], \mathrm{Im}[\alpha]), \quad \sigma_{coh}^{in} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\,. \tag{2.44}$$

Before the interaction, the total system is in a tensor product

$$\rho_{tot}^{before} = \rho_{\mathcal{S}} \otimes \rho_{th} = |\alpha\rangle_{\mathcal{S}}\langle\alpha| \otimes \rho_{th} \tag{2.45}$$

and hence the total CM is the direct sum of the single parties CM, that is

$$\sigma_{tot}^{before} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & B & 0 \\ 0 & 0 & 0 & B \end{bmatrix} \tag{2.46}$$

At this point, interactions between the signal photon, the thermal bath, and the object occur. The beam splitter transformations (2.34) lead to symplectic

transformations given by (Adesso and Illuminati 2007; Laurat et al. 2005)

$$S_{\kappa_i} = \begin{bmatrix} \sqrt{\kappa_i} & 0 & \sqrt{1-\kappa_i} & 0 \\ 0 & \sqrt{\kappa_i} & 0 & \sqrt{1-\kappa_i} \\ \sqrt{1-\kappa_i} & 0- & \sqrt{\kappa_i} & 0 \\ 0 & \sqrt{1-\kappa_i} & 0 & -\sqrt{\kappa_i} \end{bmatrix} \qquad i = 0, 1 \tag{2.47}$$

Therefore, after the interactions, it is

$$\sigma_{tot}^{out(i)} = S_{\kappa_i} \sigma_{tot}^{in} S_{\kappa_i}^T = \begin{bmatrix} A_i & 0 & C_i & 0 \\ 0 & A_i & 0 & C_i \\ C_i & 0 & B_i & 0 \\ 0 & C_i & 0 & B_i \end{bmatrix} \tag{2.48}$$

where $A_i = \kappa_i + (1-\kappa_i)B$, $B_i = 1 - \kappa_i + \kappa_i B$, and $C_i = \sqrt{\kappa_i}\sqrt{1-\kappa_i}(1-B)$. Hence, for the signal part, we have

$$\sigma_{coh}^{out(0)} = \begin{bmatrix} B & 0 \\ 0 & B \end{bmatrix}, \quad \sigma_{coh}^{out(1)} = \begin{bmatrix} \kappa + (1-\kappa)B & 0 \\ 0 & \kappa + (1-\kappa)B \end{bmatrix} \tag{2.49}$$

For the displacements

$$\begin{aligned} \langle \hat{a}_{\mathcal{S}} \rangle_i &= \mathrm{Tr}\big[\hat{a}_{\mathcal{S}} \rho_{tot}^{out(i)}\big] = \mathrm{Tr}\big[\hat{a}_{\mathcal{S}} \hat{U}_\kappa \rho_{tot}^{in} \hat{U}_{\kappa_i}^\dagger\big] = \mathrm{Tr}\big[\hat{U}_{\kappa_i}^\dagger \hat{a}_{\mathcal{S}} \hat{U}_{\kappa_i} \rho_{tot}^{in}\big] \\ &= \mathrm{Tr}\big[\sqrt{\kappa_i}\hat{a}_{\mathcal{S}} \rho_{tot}^{in}\big] + \mathrm{Tr}\big[\sqrt{1-\kappa_i}\hat{a}_{\mathcal{B}} \rho_{tot}^{in}\big] = \sqrt{\kappa_i}\alpha \end{aligned} \tag{2.50}$$

Therefore, for $\kappa_0 = 0$ and $\kappa_1 = \kappa$, it is

$$\bar{\mathbf{x}}_{coh}^{(0)out} = (0, 0), \quad \bar{\mathbf{x}}_{coh}^{(1)out} = 2\big(\sqrt{\kappa}\mathrm{Re}[\alpha], \sqrt{\kappa}\mathrm{Im}[\alpha]\big) \tag{2.51}$$

Ideal homodyne detection

Since

$$(\mathbf{x} - \bar{\mathbf{x}})^T \big(\sigma_{coh}^{out(i)}\big)^{-1} (\mathbf{x} - \bar{\mathbf{x}}) = \frac{(q_\alpha - \bar{q}^{(i)})^2 + (p_\alpha - \bar{p}^{(i)})^2}{\sqrt{\det\big[\sigma_{coh}^{out(i)}\big]}}, \tag{2.52}$$

by using probability formula (2.36), it is

$$p_i(q_\alpha) = \frac{1}{2\pi\sqrt{\det\left[\sigma^{(i)}\right]}} \int_{-\infty}^{+\infty} \exp\left\{ -\frac{1}{2} \frac{(q_\alpha - \bar{q}^{(i)})^2 + (p_\alpha - \bar{p}^{(i)})^2}{\sqrt{\det\left[\sigma^{(i)}\right]}} \right\} dp_\alpha$$
$$= \frac{1}{\sqrt{2\pi}\left(\det[\sigma^{(i)}]\right)^{\frac{1}{4}}} \exp\left\{ \frac{(q_\alpha - \bar{q}^{(i)})^2}{\sqrt{\det\left[\sigma^{(i)}\right]}} \right\}, \qquad i = 0, 1 \tag{2.53}$$

where, for simplicity, we omitted the indices out and coh and $\bar{q}^{(i)} = 2\sqrt{\kappa_i}\mathrm{Re}[\alpha]$, $\bar{p}^{(i)} = 2\sqrt{\kappa_i}\mathrm{Im}[\alpha]$

Therefore, according to the inequality (2.31), we can calculate the probability of superposition as

$$\widetilde{F} = \int_{-\infty}^{+\infty} \sqrt{p_0(q_\alpha)p_1(q_\alpha)}dq_\alpha = \int_{-\infty}^{+\infty} \frac{\exp\left\{ \frac{(q_\alpha - \bar{q}^{(0)})^2 + (q_\alpha - \bar{q}^{(1)})^2}{\sqrt{\det\left[\sigma^{(i)}\right]}} \right\} dq_\alpha}{\sqrt{2\pi}\left(\det[\sigma^{(0)}]\right)^{\frac{1}{8}}\left(\det[\sigma^{(1)}]\right)^{\frac{1}{8}}}$$
$$= \frac{\sqrt{2}\left(\det[\sigma^{(0)}]\right)^{\frac{1}{8}}\left(\det[\sigma^{(1)}]\right)^{\frac{1}{8}}}{\left(\sqrt{\det[\sigma^{(0)}]} + \sqrt{\det[\sigma^{(1)}]}\right)^{\frac{1}{2}}} \exp\left\{ \frac{(\bar{q}_\alpha^{(0)} - \bar{q}_\alpha^{(1)})^2}{\sqrt{\det\left[\sigma^{(0)}\right]} + \sqrt{\det\left[\sigma^{(1)}\right]}} \right\} \tag{2.54}$$

Since the noise field is much stronger than the signal ($N_{\mathcal{B}} \gg N_{\mathcal{S}}$), and

$$\bar{q}^{(0)} - \bar{q}^{(1)} = 2\sqrt{\kappa_0 - \kappa_1}\,\mathrm{Re}[\alpha], \quad \sqrt{\det\left[\sigma^{(i)}\right]} = \kappa_i + (1 - \kappa_i)B\,, \tag{2.55}$$

it is very difficult to distinguish between the two probability distributions $p_i(q_\alpha)$. Indeed, $2\mathrm{Re}[\alpha] \simeq \sqrt{N_{\mathcal{S}}}$, while $B = 2N_{\mathcal{B}} + 1$, meaning that the spread of the two distributions is much larger compared to the displacement of their peaks.

When sending N copies of the signal photons, the overlap becomes

$$\widetilde{F}_N = \int ... \int \sqrt{p_0(q_\alpha^1)p_1(q_\alpha^1)} ... \sqrt{p_0(q_\alpha^N)p_1(q_\alpha^N)}\, dq_\alpha^1 \dots dq_\alpha^N$$
$$= \left(\int_{-\infty}^{+\infty} \sqrt{p_0(q_\alpha)p_1(q_\alpha)}\, dq_\alpha \right)^N = \widetilde{F}^N, \tag{2.56}$$

and similarly, for the fidelity (using formula (2.39)), we find $F_N = F^N$ when N copies of the signal photons are sent.

2.3.6 Two-modes squeezed state and displacement

Now, let us consider a two-mode squeezed state

$$|\xi\rangle = \frac{1}{\cosh r}\sum_{n=0}^{\infty}(-1)^n e^{in\theta}(\tanh r)^n\,|n,n\rangle \tag{2.57}$$

as input, which is obtained by applying the squeezing operator $\hat{U}(\xi)$ to the vacuum state $|00\rangle$, that is

$$\hat{U}(\xi)\,|00\rangle \equiv e^{\xi^*\hat{a}\hat{b}-\xi\hat{a}^\dagger\hat{b}^\dagger}\,|00\rangle \tag{2.58}$$

The squeezing operator leads to the transformations

$$\begin{aligned}\hat{a} \quad &\rightarrow \quad \hat{U}^\dagger(\xi)\hat{a}\hat{U}(\xi) = \hat{a}\cosh r - e^{i\theta}\hat{b}^\dagger\sinh r\\ \hat{b} \quad &\rightarrow \quad \hat{U}^\dagger(\xi)\hat{b}\hat{U}(\xi) = \hat{b}\cosh r - e^{i\theta}\hat{a}^\dagger\sinh r\end{aligned} \tag{2.59}$$

To obtain the CM for a squeezed state one can perform the calculation explicitly or with using the fact that a squeezing transformation generated by $\hat{U}(\xi)$ induces a symplectic transformations represented by (Ferraro et al. 2005)

$$S_\xi = \begin{bmatrix} c-hs & 0 & ks & 0\\ 0 & c+hs & 0 & -ks\\ ks & 0 & c+hs & 0\\ 0 & -ks & 0 & c-hs\end{bmatrix} \tag{2.60}$$

with $c = \cosh r$, $s = \sinh r$, $h = \cos\theta$ and $k = \sin\theta$.

Let us consider $\xi = re^{i\theta}$ with $\theta = \frac{\pi}{2}$ and r such that $\sinh r = \sqrt{N_{\mathcal{S}}}$, $\cosh r = \sqrt{N_{\mathcal{S}}+1}$. Hence the CM for the state (2.57) can be obtained applying the symplectic transformation generated by

$$S_r = \begin{bmatrix} c & 0 & s & 0\\ 0 & c & 0 & -s\\ s & 0 & c & 0\\ 0 & -s & 0 & c\end{bmatrix} \tag{2.61}$$

to the vacuum state CM that is the identity matrix $\mathbb{I}_4$. Therefore one obtain

$$\sigma^{in}_{\mathcal{SI}} = \begin{bmatrix} \cosh 2r & 0 & \sinh 2r & 0\\ 0 & \cosh 2r & 0 & -\sinh 2r\\ \sinh 2r & 0 & \cosh 2r & 0\\ 0 & -\sinh 2r & 0 & \cosh 2r\end{bmatrix} \tag{2.62}$$

or equivalently, in term of $N_{\mathcal{S}}$

$$\sigma_{\mathcal{SI}}^{in} = \begin{bmatrix} S & 0 & C_q & 0 \\ 0 & S & 0 & -C_q \\ C_q & 0 & S & 0 \\ 0 & -C_q & 0 & S \end{bmatrix} \tag{2.63}$$

with $S \equiv 2N_{\mathcal{S}} + 1$ and $C_q \equiv 2\sqrt{N_{\mathcal{S}}(N_{\mathcal{S}} + 1)}$, which is the result in (Tan et al. 2008).

We can also allow the possibility of displacement by considering the two modes coherent-squeezed state of the form

$$D(\vec{\alpha}) \left|\xi\right\rangle \equiv \left|\vec{\alpha}, \xi\right\rangle \tag{2.64}$$

where $\vec{\alpha} = (\alpha_S, \alpha_I)$ is a complex vector such that

$$\begin{aligned} D^{\dagger}(\vec{\alpha}) \hat{a}_{\mathcal{S}} D(\vec{\alpha}) &= \hat{a}_{\mathcal{S}} + \alpha_{\mathcal{S}} \\ D^{\dagger}(\vec{\alpha}) \hat{a}_{\mathcal{I}} D(\vec{\alpha}) &= \hat{a}_{\mathcal{I}} + \alpha_{\mathcal{I}} \end{aligned} \tag{2.65}$$

The signal photon is sent in the place where it could be the object and the noise field, while the idler is retained. Therefore, before the interaction the total system is in the state

$$\rho_{tot}^{in} = \rho_{\mathcal{SI}} \otimes \rho_{th} = \left|\vec{\alpha}, \xi\right\rangle_{\mathcal{SI}} \left\langle\vec{\alpha}, \xi\right| \otimes \rho_{th} \tag{2.66}$$

The interaction is schematized by the same operator (2.34) which couples only the signal photon and the thermal bath, hence the symplectic transformation describing the process is given by

$$S_{\kappa_i} = \begin{bmatrix} \sqrt{\kappa_i} & 0 & 0 & 0 & \sqrt{1-\kappa_i} & 0 \\ 0 & \sqrt{\kappa_i} & 0 & 0 & 0 & \sqrt{1-\kappa_i} \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \sqrt{1-\kappa_i} & 0 & 0 & 0 & \sqrt{\kappa_i} & 0 \\ 0 & \sqrt{1-\kappa_i} & 0 & 0 & 0 & -\sqrt{\kappa_i} \end{bmatrix} \tag{2.67}$$

which acts only on the signal and the bath subspaces.

Acting with the transformation and considering only the firs 4×4 matrix

relative to the signal and the idler part, one obtains

$$\sigma_{\mathcal{SI}}^{(i)out} = \begin{bmatrix} \kappa S + (1-\kappa_i)B & 0 & \sqrt{\kappa_i}C_q & 0 \\ 0 & \kappa_i S + (1-\kappa_i)B & 0 & -\sqrt{\kappa_i}C_q \\ \sqrt{\kappa_i}C_q & 0 & S & 0 \\ 0 & -\sqrt{\kappa_i}C_q & 0 & S \end{bmatrix} \qquad i = 0, 1 \tag{2.68}$$

with $\kappa_0 = 0$ and $\kappa_1 = \kappa$.

The squeezing transformation does not affect the displacement of the state, hence at the output we have

$$\bar{\mathbf{x}}_{squeez}^{(0)out} = (0,0,0,0), \quad \bar{\mathbf{x}}_{squeez}^{(1)out} = 2\big(\sqrt{\kappa}\mathrm{Re}[\alpha], \sqrt{\kappa}\mathrm{Im}[\alpha], 0, 0\big) \tag{2.69}$$

Indeed

$$\begin{aligned} \langle \hat{a}_{\mathcal{S}} \rangle = \mathrm{Tr}\big[\rho^{out}\hat{a}_{\mathcal{S}}\big] &= \mathrm{Tr}\big[\hat{U}_{\kappa_i}\hat{D}(\alpha)\hat{U}_{\xi}\left|0\right\rangle\left\langle 0\right|\hat{U}_{\xi}^{\dagger}\hat{D}^{\dagger}(\alpha)\hat{U}_{\kappa_i}^{\dagger}\hat{a}_{\mathcal{S}}\big] \\ &= \mathrm{Tr}\big[\hat{U}_{\xi}^{\dagger}\hat{D}^{\dagger}(\alpha)(\sqrt{\kappa_i}\hat{a}_{\mathcal{S}})\hat{D}(\alpha)\hat{U}_{\xi}\left|0\right\rangle\left\langle 0\right|\big] \\ &= \sqrt{\kappa_i}\mathrm{Tr}\big[\hat{U}_{\xi}^{\dagger}(\hat{a}_{\mathcal{S}} + \alpha)\hat{U}_{\xi}\big] = \sqrt{\kappa_i}\alpha \end{aligned} \tag{2.70}$$

In the case of the two-mode squeezed-coherent state, the average number of photons is given by

$$N_{\mathcal{S}} = \sinh^2(1 + |\alpha|^2) + |\alpha|^2 \tag{2.71}$$

Ideal homodyne

Since

$$\big(\sigma^{(i)}\big)^{-1} = \frac{1}{A_i S - C_i^2}\begin{bmatrix} S & 0 & -C_i & 0 \\ 0 & S & 0 & C_i \\ -C_i & 0 & A_i & 0 \\ 0 & C_i & 0 & A_i \end{bmatrix} \tag{2.72}$$

with $A_i = S + (1-\kappa_i)B$, $C_i = \sqrt{\kappa_i} C_q$ and $\det[\sigma^{(i)}] = (A_i S - C_i^2)^2$, the Wigner functions for the output states are

$$W_i(\mathbf{x}) = \frac{1}{(2\pi)^2\sqrt{\det[\sigma^{(i)}]}} \exp\Bigg\{ -\frac{1}{2\sqrt{\det[\sigma^{(i)}]}} \Big[S(q_1^{(i)} - \bar{q}_1^{(i)})^2 + A_i (q_2 - \bar{q}_2^{(i)})^2 - 2C_i (q_1 - \bar{q}_1^{(i)})(q_2 - \bar{q}_2^{(i)}) \Big] \Bigg\} \cdot \exp\Bigg\{ -\frac{1}{2\sqrt{\det[\sigma^{(i)}]}} \Big[S(p_1 - \bar{p}_1^{(i)})^2 + A_i (p_2 - \bar{p}_2^{(i)})^2 - 2C_i (p_1 - \bar{p}_1^{(i)})(p_2 - \bar{p}_2^{(i)}) \Big] \Bigg\} \tag{2.73}$$

which can be rewritten as

$$W_i(\mathbf{x}) = \frac{1}{(2\pi)^2\sqrt{\det[\sigma^{(i)}]}} \exp\Big\{ (\mathbf{q} - \bar{\mathbf{q}}^{(i)}) \sigma_q^{(i)} (\mathbf{q} - \bar{\mathbf{q}}^{(i)}) \Big\} \exp\Big\{ (\mathbf{p} - \bar{\mathbf{p}}^{(i)}) \sigma_p^{(i)} (\mathbf{p} - \bar{\mathbf{p}}^{(i)}) \Big\} \tag{2.74}$$

with $\mathbf{q} = (q_1, q_2)$, $\mathbf{p} = (p_1, p_2)$, $\bar{\mathbf{q}}^{(i)} = (\bar{q}_1, \bar{q}_2^{(i)})$, $\bar{\mathbf{p}}^{(i)} = (\bar{p}_1^{(i)}, \bar{p}_2^{(i)})$ and

$$\sigma_q^{(i)} = \frac{1}{\sqrt{\det[\sigma^{(i)}]}} \begin{bmatrix} S & -C_i \\ -C_i & A_i \end{bmatrix}, \qquad \sigma_p^{(i)} = \frac{1}{\sqrt{\det[\sigma^{(i)}]}} \begin{bmatrix} S & C_i \\ C_i & A_i \end{bmatrix} \tag{2.75}$$

Hence

$$p_i(q_1^\alpha, q_2^\alpha) = \int W_i(\mathbf{x}) \delta(q_1 - q_1^\alpha) \delta(q_2 - q_2^\alpha) dq_1 dq_2 dp_1 dp_2 = \frac{\exp\Big\{ -\frac{1}{2} [(\mathbf{q}^\alpha - \bar{\mathbf{q}}^{(i)}) \sigma_q^{(i)} (\mathbf{q}^\alpha - \bar{\mathbf{q}}^{(i)})] \Big\}}{(2\pi)\big(\det[\sigma^{(i)}]\big)^{\frac{1}{4}}} \tag{2.76}$$

Therefore we can calculate the superposition of the output probabilities

$$\widetilde{F} = \int \frac{\exp\left\{ -\frac{1}{2} \left[(\mathbf{q}^\alpha - \bar{\mathbf{q}}^{(0)}) \sigma_m (\mathbf{q}^\alpha - \bar{\mathbf{q}}^{(0)}) \right] \right\}}{2\pi \left(\det[\sigma^{(0)}] \right)^{\frac{1}{8}} \left(\det[\sigma^{(1)}] \right)^{\frac{1}{8}}} \cdot \exp\left\{ -\frac{1}{2} \Delta\mathbf{q}\, \sigma_q^{(1)} (\mathbf{q}^\alpha - \bar{\mathbf{q}}^{(0)}) - \frac{1}{4} \Delta\mathbf{q}\, \sigma_q^{(1)} \Delta\mathbf{q} \right\} dq_1^\alpha dq_2^\alpha \tag{2.77}$$

where $\sigma_m \equiv \frac{\sigma_q^{(0)} + \sigma_q^{(1)}}{2}$ and $\Delta\mathbf{q} = \bar{\mathbf{q}}^{(1)} - \bar{\mathbf{q}}^{(0)}$.

By using the relation for $n-$dimensional Gaussian integrals

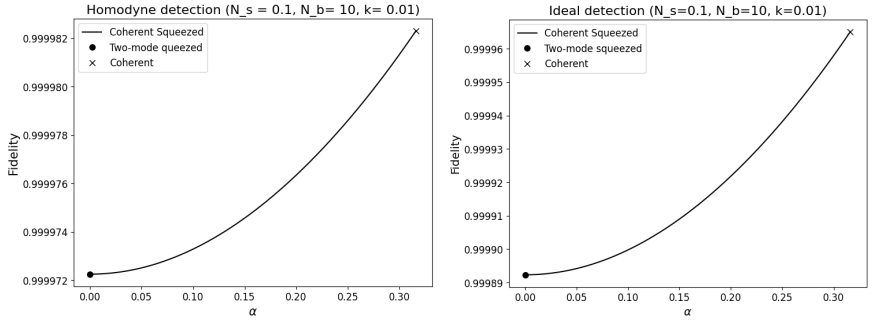
$$\int e^{-\frac{1}{2}\mathbf{x}^T A\mathbf{x}+\mathbf{B}\cdot\mathbf{x}} d\mathbf{x} = \sqrt{\frac{(2\pi)^n}{\det[A]}} e^{-\frac{1}{2}\mathbf{B}^T A^{-1}\mathbf{B}} \tag{2.78}$$

with $A = \sigma_m$ and $\mathbf{B} = \Delta\mathbf{q}\sigma_q^{(1)}$, we obtain

$$\widetilde{F} = \sqrt{\frac{1}{\det[\sigma_m]}} \frac{\exp\left\{-\frac{1}{2}(\Delta\mathbf{q}\sigma_q^{(1)})^T \sigma_m^{-1} \Delta\mathbf{q}\sigma_q^{(1)}\right\}}{\left(\det[\sigma^{(0)}]\right)^{\frac{1}{8}} \left(\det[\sigma^{(1)}]\right)^{\frac{1}{8}}} \tag{2.79}$$

As for the coherent state, if we take N copies of the system, we can compare the fidelity F_N with the probability superposition $\tilde{F}_N$ obtaining the plot in figures from 2.2a to 2.5b. Note that a lower fidelity corresponds to better discrimination performance. As expected, for a fixed number of signal photons $N_{\mathcal{S}}$, the optimal working point is $\alpha = 0$, which corresponds to a two-mode squeezed state.

Since the fidelity is lower than the fidelity of the coherent state, as in (Tan et al. 2008), also in the case of homodyne detection, the entanglement in the squeezing state improves the possibility of discrimination between the two output states.

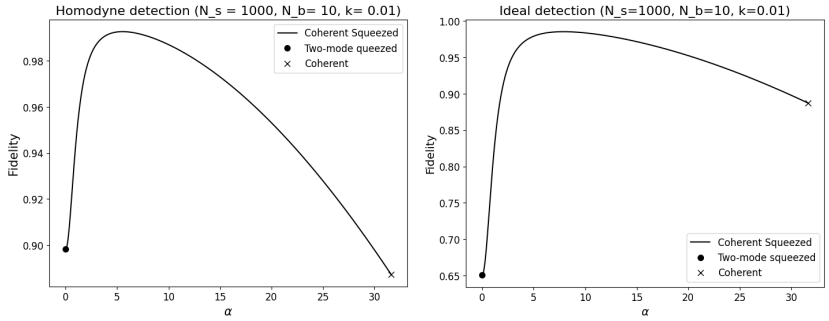


(a) **(b)**

Figure 2.2: *Homodyne and optimal detection for low reflectivity k and low signal photons number $N_{\mathcal{S}}$.*

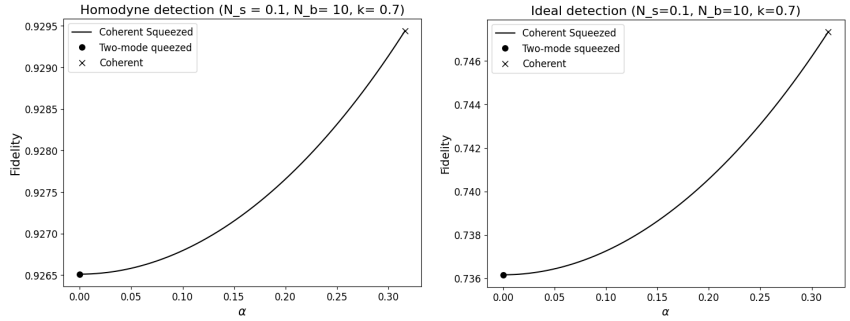
2.4 Discrimination of two level systems with two-mode squeezed states

In this section, motivated by the fact that coherent states of light enable very precise measurements for discriminating atomic energy levels (Gravina et al.



(a) **(b)**

Figure 2.3: *Homodyne and optimal detection for low reflectivity k and large signal photons number N_S.*

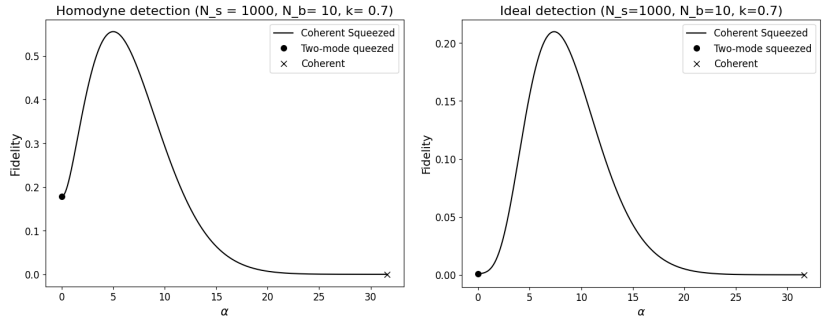


(a) **(b)**

Figure 2.4: *Homodyne and optimal detection for high reflectivity k and low signal photons number N_S.*

2024), and by the promising results of the two-mode squeezed states discussed in the previous section, we consider another discrimination protocol. Here, the two alternatives correspond to two different atoms, each characterized by a distinct energy gap between the ground and excited states. As in the previous section, we use light to discriminate between the two possible alternatives, analyzing the performance under coherent light, single-mode squeezed light, and two-mode squeezed light. We compute the error probability associated with the discrimination task, showing that, once again, the two-mode squeezed state provides the best performance, while coherent and single-mode squeezed states yield similar results.

Specifically, in the next section we introduce the model describing the interaction between radiation and matter under the dipole approximation. We



(a) **(b)**

Figure 2.5: *Homodyne and optimal detection for high reflectivity k and large signal photons number $N_{\mathcal{S}}$.*

also present the Jaynes–Cummings model, which describes the interaction between light and a two-level system (TLS), treating both as quantum systems. We solve the Jaynes–Cummings model for different input states of light, and finally compute the error probability by tracing out the two-level systems.

2.4.1 Radiation-matter interaction

The Hamiltonian of an electron bound to an atom and immersed in an electromagnetic field is given by:

$$\hat{H} = \frac{1}{2m}\left(\hat{\mathbf{p}} + e\mathbf{A}\right)^2 + V(\mathbf{r}) - e\phi(\mathbf{r},t)\,, \tag{2.80}$$

where m and q are the mass and the charge of the electron, $\hat{\mathbf{p}} = -i\nabla$ is its canonical momentum operator, ϕ and $\mathbf{A}$ are the scalar potential and the vector potential of the external electromagnetic field and $V(\mathbf{r})$ represents the potential binding the electron to the atom, such as the Coulomb potential for a hydrogen-like atom:

$$V(\mathbf{r}) = -\frac{e^2}{4\pi\epsilon_0 r} \tag{2.81}$$

The external field is related to the potentials by the relations

$$\mathbf{E}(\mathbf{r},t) = -\nabla\phi(\mathbf{r},t) - \frac{\partial\mathbf{A}(\mathbf{r},t)}{\partial t}\,, \quad \mathbf{B}(\mathbf{r},t) = \nabla\times\mathbf{A}(\mathbf{r},t)\,, \tag{2.82}$$

which are invariant under the gauge transformation:

$$\phi'(\mathbf{r},t) = \phi(\mathbf{r},t) - \frac{\partial\chi(\mathbf{r},t)}{\partial t}\,, \quad \mathbf{A}'(\mathbf{r},t) = \mathbf{A}(\mathbf{r},t) + \nabla\chi(\mathbf{r},t)\,. \tag{2.83}$$

The associated Schrödinger equation is hence

$$i\hbar\frac{\partial\Psi}{\partial t} = \hat{H}\Psi\,. \tag{2.84}$$

By considering a unitary transformation $\hat{U}$ such that $\Psi' = \hat{U}\Psi$, we have

$$i\hbar\frac{\partial\Psi'}{\partial t} = \hat{H}'\Psi'\,, \tag{2.85}$$

with

$$\hat{H}' = \hat{U}\hat{H}\hat{U}^\dagger + i\hbar\frac{\partial U}{\partial t}U^\dagger\,. \tag{2.86}$$

Choosing $\hat{U} = e^{\frac{ie\chi(\mathbf{r},t)}{\hbar}}$, it is

$$\hat{H}' = \frac{1}{2m}[\hat{\mathbf{P}} + e\mathbf{A}']^2 - e\phi' + V(r)\,, \tag{2.87}$$

with $\mathbf{A}'$ and ϕ' given by Eqs. (2.83). If we choose the so-called Coulomb gauge, with $\phi = 0$ and $\nabla\cdot\mathbf{A} = 0$, the potential vector $\mathbf{A}$, in the absence of sources, follows the wave equation

$$\nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial\mathbf{A}}{\partial t} = 0\,, \tag{2.88}$$

with wave plan solution

$$\mathbf{A}(\mathbf{r},t) = \mathbf{A}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + c.c. \tag{2.89}$$

The Hamiltonian simplifies as

$$\hat{H} = \frac{1}{2m}\left(\hat{\mathbf{p}} + e\mathbf{A}\right)^2 + V(\mathbf{r})\,, \tag{2.90}$$

and the transformed one is

$$\hat{H}' = \frac{1}{2m}\left[\hat{\mathbf{p}} + e(\mathbf{A} + \nabla\chi)\right]^2 + e\frac{\partial\chi}{\partial t} + V(\mathbf{r})\,. \tag{2.91}$$

For $|\mathbf{r}|$ on the order of the atomic length ($1\text{Å} = 10^{-10}$ m) and optical light with a wavelength λ in the range of $(400 - 700)$ nm, we have $\lambda \gg |\mathbf{r}|$. Hence, we

can approximate the external field vector potential as being spatially uniform, $\mathbf{A}(\mathbf{r},t) \simeq \mathbf{A}(t)$, which is known as the *dipole approximation*.

By choosing the gauge function $\chi(\mathbf{r},t) = -\mathbf{A}(t)\cdot\mathbf{r}$ (compatible with $\nabla\cdot\mathbf{A} = 0$), the Hamiltonian takes the form:

$$\hat{H}' = \frac{\hat{\mathbf{P}}^2}{2m} + V(r) + e\mathbf{r}\cdot\mathbf{E}(t)\,, \tag{2.92}$$

where the last term represents the interaction of the electron with the external field. The quantity $\hat{\mathbf{d}} = -e\mathbf{r}$ is called the dipole moment.

Thus, we can write:

$$\hat{H} = \hat{H}_0 - \hat{\mathbf{d}}\cdot\mathbf{E}(t) \tag{2.93}$$

where we have omitted the prime notation for simplicity.

2.4.2 Two level systems interacting with light: Jaynes-Cummings model and rotating wave approximation

In order to analyze the quantum features of light, we need to consider a fully quantum model of interaction between light and atoms. For this purpose, we consider a single mode electromagnetic field of the form

$$\hat{\mathbf{E}}(t) = E_0(\hat{a} + \hat{a}^\dagger)\sin(kz)\mathbf{e}\,, \tag{2.94}$$

where E_0 is the amplitude, $k = \frac{2\pi}{\lambda}$ and $\mathbf{e}$ represent the direction of the polarization interacting with a two level system with ground state $|g\rangle$ and excited state $|e\rangle$ with energies E_g and E_e respectively. The interaction term in the Hamiltonian (2.93) reads as

$$\hat{H}_I \equiv g\hat{d}(\hat{a} + \hat{a}^\dagger)\,, \tag{2.95}$$

with $g = -E_0\sin(kz)$ and $\hat{d} = \hat{\mathbf{d}}\cdot\mathbf{e}$.

Introducing the TLS annihilation and creation operators

$$\hat{\sigma}_+ = |e\rangle\langle g|\,, \quad \hat{\sigma}_- = |g\rangle\langle e|\,, \tag{2.96}$$

and considering that, due to parity symmetry, the dipole operator in the standard basis has only nonzero off-diagonal elements (Gerry and Knight 2023), i.e.,

$$\langle g|\hat{d}|g\rangle = \langle e|\hat{d}|e\rangle = 0, \quad \langle e|\hat{d}|g\rangle = \langle g|\hat{d}|e\rangle^* = d\,, \tag{2.97}$$

we can express the dipole operator as

$$\hat{d} = d\hat{\sigma}_- + d^*\hat{\sigma}_+ \,. \tag{2.98}$$

which leads to the interaction Hamiltonian

$$\hat{H}_I = \overline{h}\lambda(\hat{\sigma}_+ + \hat{\sigma}_-)(\hat{a}^\dagger + \hat{a})\,, \tag{2.99}$$

with $\lambda = \frac{dg}{\overline{h}}$. In the free Hamiltonian $\hat{H}_0$, we include the free energies of the TLS and the electromagnetic field, which can be written as

$$\hat{H}_0 = \overline{h}\omega_0\hat{\sigma}_z + \overline{h}\omega\hat{a}^\dagger\hat{a}\,, \tag{2.100}$$

where $\overline{h}\omega_0 = E_e - E_g$. Since the time evolution of the photon creation and annihilation operators is given by

$$\hat{a}(t) = \hat{a}(0)e^{-i\omega t}, \quad \hat{a}^\dagger(t) = \hat{a}^\dagger(0)e^{i\omega t}\,, \tag{2.101}$$

and similarly for the atomic operators

$$\hat{\sigma}_-(t) = \hat{\sigma}_-(0)e^{-i\omega_0 t}, \quad \hat{\sigma}_+(t) = \hat{\sigma}_+(0)e^{i\omega_0 t}\,, \tag{2.102}$$

we obtain

$$\hat{\sigma}_-(t)\hat{a}(t) \sim e^{-i(\omega+\omega_0)t}\,, \quad \hat{\sigma}_+(t)\hat{a}^\dagger(t) \sim e^{i(\omega+\omega_0)t}, \tag{2.103}$$

$$\hat{\sigma}_-(t)\hat{a}^\dagger(t) \sim e^{-i(\omega-\omega_0)t}\,, \quad \hat{\sigma}_+(t)\hat{a}(t) \sim e^{i(\omega-\omega_0)t}\,. \tag{2.104}$$

If $\omega \simeq \omega_0$, the first two terms oscillate much more rapidly compared to the last two. Therefore, it is not surprising that, in a perturbative approach, when integrating over time, they are suppressed relative to the latter.

This fact can also be understood by considering the product $\hat{a}^\dagger\hat{\sigma}_+$, which creates a photon while simultaneously exciting the TLS. This process is energetically unfavorable. A similar argument applies to the term $\hat{\sigma}_-\hat{a}$, which annihilates a photon while de-exciting the TLS. Hence, these processes occur with lower probabilities.

Therefore, the Hamiltonian can be simplified as

$$\hat{H} = \frac{1}{2}\overline{h}\omega_0\hat{\sigma}_z + \overline{h}\omega\hat{a}^\dagger\hat{a} + \overline{h}\lambda(\hat{\sigma}_+\hat{a} + \hat{\sigma}_-\hat{a}^\dagger)\,, \tag{2.105}$$

and the omission of the rapidly oscillating terms is known as the *rotating wave approximation* (RWA). This approximation breaks down for strong light-matter coupling λ, in regimes far from resonance ($\omega \not\simeq \omega_0$), and for sufficiently long evolution times t.

Such a Hamiltonian describes the so-called Jaynes-Cummings model, that is the generalization of the Rabi model in a fully quantum picture (Gerry and Knight 2023).

Defining the operators

$$\hat{P} = |g\rangle\langle g| + |e\rangle\langle e|\,, \quad \hat{N}_e = \hat{a}^\dagger\hat{a} + |e\rangle\langle e|\,, \tag{2.106}$$

we see that they commute with the Hamiltonian, hence they are conserved quantities. Therefore, we can rewrite the Hamiltonian as

$$\hat{H} = \hat{H}_1 + \hat{H}_2\,, \tag{2.107}$$

where

$$\hat{H}_1 = \overline{h}\omega\hat{N}_e + \overline{h}\left(\frac{\omega_0}{2} - \omega\right)\hat{P}\,, \quad \hat{H}_2 = -\overline{h}\Delta + \overline{h}\lambda(\hat{\sigma}_+\hat{a} + \hat{\sigma}_-\hat{a}^\dagger)\,. \tag{2.108}$$

where $\Delta = \omega_0 - \omega$. Only the second term $\hat{H}_2$ governs the time evolution.

Assuming that the atom and the field initially start in a product state of the form

$$|\Psi(0)\rangle = \left(\sum_{n=0}^{\infty} C_n\,|n\rangle\right) \otimes (C_g\,|g\rangle + C_e\,|e\rangle)\,, \tag{2.109}$$

where $\sum_n |C_n|^2 = 1$ and $|C_g|^2 + |C_e|^2 = 1$, we can express the wave function at time $t \geq 0$ as

$$|\Psi(t)\rangle = \sum_{n=0}^{\infty} [C_g C_n C_{n,g}(t)\,|n,g\rangle + C_e C_n C_{e,n}(t)\,|n,e\rangle]\,. \tag{2.110}$$

where the time-dependent coefficients $C_{n,g}(t)$ and $C_{n,e}(t)$ are such that the wave function satisfies the Schrödinger equation

$$i\overline{h}\frac{\partial}{\partial t}\,|\Psi(t)\rangle = \hat{H}_2\,|\Psi(t)\rangle\,, \tag{2.111}$$

with the initial condition $|\Psi(t=0)\rangle = |\Psi(0)\rangle$ given in Eq. (2.109).

Substituting the expression (2.110) into the Schrödinger equation, since

$$i\overline{h}\frac{\partial}{\partial t}\,|\Psi(t)\rangle = i\overline{h}\sum_{n=0}^{\infty}\left(C_g C_n\frac{\partial C_{n,g}(t)}{\partial t}\,|n,g\rangle + C_e C_n\frac{\partial C_{n,e}(t)}{\partial t}\,|n,e\rangle\right)\,, \tag{2.112}$$

and

$$\hat{H}_2 |\Psi(t)\rangle = -\overline{h}\Delta \sum_{n=0}^{\infty} \left(C_n C_g C_{n,g}(t) |n,g\rangle + C_e C_n C_{e,n}(t) |n,e\rangle\right) + \overline{h}\lambda \left(\sum_{n=1}^{\infty} C_g C_n C_{n,g}(t)\sqrt{n} |n-1,e\rangle + \sum_{n=0}^{\infty} C_e C_n C_{n,e}(t)\sqrt{n+1} |n+1,g\rangle \right), \tag{2.113}$$

recalling the indices $n \to n+1$ and $n \to n-1$ respectively in the first and in the second term of the second sum, it is

$$\hat{H}_2 |\Psi(t)\rangle = -\overline{h}\Delta \sum_{n=0}^{\infty} \left(C_n C_g C_{n,g}(t) |n,g\rangle + C_e C_n C_{e,n}(t) |n,e\rangle\right) + \overline{h}\lambda \left(\sum_{n=1}^{\infty} C_g C_{n+1} C_{n+1,g}(t)\sqrt{n+1} |n,e\rangle + \sum_{n=0}^{\infty} C_e C_{n-1} C_{n-1,e}(t)\sqrt{n} |n,g\rangle \right). \tag{2.114}$$

Hence we obtain a set of coupled differential equations for the coefficients:

$$\begin{cases} iC_g C_n \frac{\partial C_{n,g}(t)}{\partial t} = -\Delta C_g C_n C_{n,g}(t) + \lambda\sqrt{n} C_e C_{n-1} C_{n-1,e}(t), \\ iC_e C_n \frac{\partial C_{n,e}(t)}{\partial t} = -\Delta C_e C_n C_{n,e}(t) + \lambda\sqrt{n+1} C_g C_{n+1} C_{n+1,g}(t). \end{cases} \tag{2.115}$$

At this point, we will analyze two different approaches: one for the resonant case $\Delta = 0$ and another for the non-resonant case $\omega \neq \omega_0$.

Resonant case: $\Delta = 0$

If $\Delta = 0$, the above equations simplify to

$$\begin{cases} iC_g C_n \frac{\partial C_{n,g}(t)}{\partial t} = \lambda\sqrt{n} C_e C_{n-1} C_{n-1,e}(t), \\ iC_e C_n \frac{\partial C_{n,e}(t)}{\partial t} = \lambda\sqrt{n+1} C_g C_{n+1} C_{n+1,g}(t). \end{cases} \tag{2.116}$$

Differentiating the first equation with respect to time and substituting the expression for $\frac{\partial C_{n-1,e}(t)}{\partial t}$ from the second equation, we obtain

$$\frac{\partial^2 C_{n,g}(t)}{\partial t^2} = -\lambda^2 n C_{n,g}(t), \tag{2.117}$$

whose solution corresponds to a harmonic oscillator:

$$C_{n,g}(t) = A_n \cos\left(\lambda\sqrt{n}t\right) + B_n \sin\left(\lambda\sqrt{n}t\right). \tag{2.118}$$

Similarly, we find

$$C_{n,e}(t) = \frac{i}{\lambda\sqrt{n+1}}\frac{\partial C_{n+1,g}(t)}{\partial t} = i\left[A_{n+1}\sin\left(\lambda\sqrt{n+1}t\right) + B_{n+1}\cos\left(\lambda\sqrt{n+1}t\right)\right]. \tag{2.119}$$

Imposing the initial conditions at time $t = 0$, we obtain the constraints

$$A_n = C_n C_g, \quad iB_{n+1} = C_n C_e. \tag{2.120}$$

Thus, the complete solution reads

$$\begin{aligned}|\Psi(t)\rangle = &\sum_{n=0}^{\infty}\left[C_n C_g \cos\left(\lambda\sqrt{n}t\right)|n,g\rangle - iC_n C_e \sin\left(\lambda\sqrt{n+1}t\right)|n+1,g\rangle\right] \\ &+ \sum_{n=1}^{\infty}\left[iC_n C_g \sin\left(\lambda\sqrt{n}t\right)|n-1,e\rangle + C_{n+1}C_e \cos\left(\lambda\sqrt{n+1}t\right)|n,e\rangle\right],\end{aligned} \tag{2.121}$$

which can be rewritten in the form

$$\begin{aligned}|\Psi(t)\rangle = &\sum_{n=0}^{\infty}\left[C_n C_g \cos\left(\lambda\sqrt{n}t\right) - i\tilde{\delta}C_{n-1}C_e \sin\left(\lambda\sqrt{n}t\right)\right]|n,g\rangle \\ &+ \sum_{n=1}^{\infty}\left[iC_{n+1}C_g \sin\left(\lambda\sqrt{n+1}t\right) + C_n C_e \cos\left(\lambda\sqrt{n+1}t\right)\right]|n,e\rangle\ .\end{aligned} \tag{2.122}$$

with $\tilde{\delta} \equiv 1 - \delta_{n,0}$.

For example, if the light starts at time $t = 0$ in a state with $N_{ph} > 0$ photons in the same mode $|N_{ph}\rangle$, the solution is

$$\begin{aligned}|\Psi(t)\rangle = &C_g \cos\left(\sqrt{N_{ph}}\lambda t\right)|N_{ph},g\rangle - iC_e \sin\left(\sqrt{N_{ph}+1}\lambda t\right)|N_{ph}+1,g\rangle \\ &+ iC_g \sin\left(\sqrt{N_{ph}}\lambda t\right)|N_{ph}-1,e\rangle + C_e \cos\left(\sqrt{N_{ph}+1}\lambda t\right)|N_{ph},e\rangle\ .\end{aligned} \tag{2.123}$$

Out of resonance case: $\Delta \neq 0$

The approach followed in the previous section does not work in the case $\Delta \neq 0$, since it is not possible to decouple the differential equations (2.115). Nevertheless, it is possible to directly diagonalize the Hamiltonian (2.105) in order to find the solution of the related Schrödinger equation. In fact, noting that in the basis $\{|n,g\rangle, |n,e\rangle\}_{n=0}^{\infty}$, the Hamiltonian has nonzero terms only in the subspaces spanned by $\{|n,e\rangle, |n+1,g\rangle\}$ and in the ground state subspace spanned by $|0,g\rangle$. Since these subspaces are orthogonal to each other, the Hamiltonian (2.105) can be written as a direct sum over these subspaces:

$$\hat{H} = \bigoplus_{n=0}^{\infty} \hat{H}_n \,, \tag{2.124}$$

where $\hat{H}_0 = -\frac{1}{2}\overline{h}\omega_0 \equiv E_{-1}$, and for $n \geq 0$:

$$\hat{H}_n = \begin{pmatrix} n\overline{h}\omega + \frac{1}{2}\overline{h}\omega_0 & \overline{h}\lambda\sqrt{n+1} \\ \overline{h}\lambda\sqrt{n+1} & (n+1)\overline{h}\omega - \frac{1}{2}\overline{h}\omega_0 \end{pmatrix} . \tag{2.125}$$

This matrix can be immediately diagonalized, yielding the eigenvalues

$$E_{\pm}(n) = \left(n + \frac{1}{2}\right)\overline{h}\omega \pm \overline{h}\Omega_n, \quad \Omega_n = \sqrt{\Delta^2 + 4\lambda^2(n+1)}. \tag{2.126}$$

The corresponding eigenvectors are

$$\begin{aligned} |n,+\rangle &= \cos\left(\frac{\Phi_n}{2}\right)|n,e\rangle + \sin\left(\frac{\Phi_n}{2}\right)|n+1,g\rangle \,, \\ |n,-\rangle &= -\sin\left(\frac{\Phi_n}{2}\right)|n,e\rangle + \cos\left(\frac{\Phi_n}{2}\right)|n+1,g\rangle \,, \end{aligned} \tag{2.127}$$

where

$$\Phi_n = \tan^{-1}\left(\frac{2\lambda\sqrt{n+1}}{\Delta}\right). \tag{2.128}$$

This results in

$$\sin\left(\frac{\Phi_n}{2}\right) = \frac{1}{\sqrt{2}}\left[\frac{\Omega_n - \Delta}{\Omega_n}\right]^{\frac{1}{2}}, \quad \cos\left(\frac{\Phi_n}{2}\right) = \frac{1}{\sqrt{2}}\left[\frac{\Omega_n + \Delta}{\Omega_n}\right]^{\frac{1}{2}} . \tag{2.129}$$

The states $|n,\pm\rangle$ are called *dressed states* or *Jaynes-Cummings states*, in contrast to the bare states $|n,e\rangle$ and $|n+1,g\rangle$. While the bare states have an energy difference of $\overline{h}\Delta$, the dressed states exhibit a modified energy splitting given

by $\overline{h}\Omega_n$.

After expanding $|\Psi(0)\rangle$ (Eq. (2.109)) in the energy eigenbasis, applying the unitary evolution and returning to the standard Fock basis (see B.2 for details), the solution of the Schrödinger equation becomes

$$|\Psi(t)\rangle = \sum_{n=0}^{\infty} \Big\{ \left[C_n C_g K_1(n-1,t)\,|n\rangle + C_n C_e K_2(n,t)\,|n+1\rangle\right] \otimes |g\rangle + \left[C_n C_e K_1(n,t) + C_{n+1} C_g K_2(n,t)\right] |n\rangle \otimes |e\rangle \Big\}, \quad (2.130)$$

where

$$K_1(n,t) \equiv \cos^2\left(\frac{\Phi_n}{2}\right) e^{-i\frac{E_-(n)t}{\hbar}} + \sin^2\left(\frac{\Phi_n}{2}\right) e^{-i\frac{E_+(n)t}{\hbar}},$$
$$K_2(n,t) \equiv -i\sin(\Phi_n)\sin\left(\frac{E_2(n)t}{\overline{h}}\right) e^{-i\frac{E_1(n)t}{\hbar}},$$
$$E_1(n) = \left(n+\frac{1}{2}\right)\overline{h}\omega, \quad E_2(n) = \overline{h}\Omega_n, \quad n = 0,\ldots,\infty\,. \quad (2.131)$$

and $K_1(-1,t) \equiv 1$.

In the case there is a second mode non-interacting with the TLS, if we start with the light in the state

$$|\Psi_{light}\rangle = \sum_{n,m=0}^{\infty} C_{n,m}\,|n,m\rangle\,, \quad (2.132)$$

the outcome state will be

$$|\Psi(t)\rangle = \sum_{n,m=0}^{\infty} \Big\{ \left[C_{n,m} C_g K_1(n-1,t)\,|n,m\rangle + C_{n,m} C_e K_2(n,t)\,|n+1,m\rangle\right] \otimes |g\rangle + \left[C_{n,m} C_e K_1(n,t) + C_{n+1,m} C_g K_2(n,t)\right] |n,m\rangle \otimes |e\rangle \Big\}. \quad (2.133)$$

2.4.3 Discrimination of TLSs with light

At this stage, using the results of detection theory reviewed in Section 2.2, we aim to compare different light sources, in particular, coherent, single-mode squeezed, and two-mode squeezed states, in order to discriminate between two atoms, described as two-level systems characterized by two different energy levels $\overline{h}\omega_0^{(i)}$ from the ground state to the excited state.

Clearly, the Helstrom bound (2.1) provides a valid figure of merit even

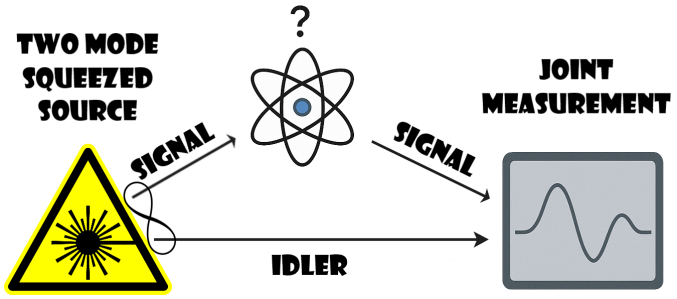


Figure 2.6: *The source generates mode-entangled states of light. Photons in the signal mode are sent to the atom, and after the interaction they are collected and jointly guided with the idler mode to the receiver, where joint measurements are performed.*

in the scenario where both atoms are present in the experimental setup. In this context, the lower the error probability, the more distinguishable the two possible outcome profiles of the chosen observable.

Alternatively, one can consider calculating the fidelity between the two possible outcomes, which is related to the Bures distance between the two alternatives (2.6).

Both methods are connected to the distances between the two alternatives: the greater the distance, the higher the expected resolution.

In analogy with quantum illumination, the signal photons are sent to the atom, while the idler photons are retained. After the interaction, a joint measurement on the entire system (signal + idler) is performed, as sketched in Fig. 2.6.

The results in the optimal Helstrom detection scheme are discussed in the following sections.

Single-mode squeezed vs coherent

Since the light-atom state at time t is given by (2.130) which can be rewritten as

$$\begin{aligned}|\Psi^{(i)}(t)\rangle &= \sum_{n=0}^{\infty}\left[C^{(i)}_{n,g}(t)\,|n,g\rangle + D^{(i)}_{n,g}(t)\,|n+1,g\rangle + (C^{(i)}_{n,e}(t)+D^{(i)}_{n,e}(t))\,|n,e\rangle\right]\\ &= \sum_{n=0}^{\infty}\left[C^{(i)}_{n,g}(t)+D^{(i)}_{n-1,g}(t)\right]|n,g\rangle + (C^{(i)}_{n,e}(t)+D^{(i)}_{n,e}(t))\,|n,e\rangle\Big]\,, \qquad (2.134)\end{aligned}$$

where

$$\begin{aligned}C^{(i)}_{n,g}(t) &= C_n C_g K^{(i)}_1(n-1,t), \quad D^{(i)}_{n,g}(t) = C_n C_e K^{(i)}_2(n,t)\\ C^{(i)}_{n,e}(t) &= C_n C_e K^{(i)}_1(n,t), \quad D^{(i)}_{n,e}(t) = C_{n+1} C_g K^{(i)}_2(n,t)\,, \qquad (2.135)\end{aligned}$$

and $D^{(i)}_{-1,g} \equiv 0$. Hence, the photon state will be

$$\begin{aligned}\rho^{(i)}_{phot} &= \mathrm{Tr}_{tls}\big[\,|\Psi^{(i)}(t)\rangle\,\langle\Psi^{(i)}(t)|\,\big] \\ &= \sum_{n,n'=0}^{\infty}\Big[\big(C^{(i)}_{n,g}+D^{(i)}_{n-1,g}\big)\big(C^{(i)}_{n',g}+D^{(i)}_{n'-1,g}\big)^{*} \\ &\quad+\big(C^{(i)}_{n,e}+D^{(i)}_{n,e}\big)\big(C^{(i)}_{n',e}+D^{(i)}_{n',e}\big)^{*}\Big]\,|n\rangle\,\langle n'|\,, \end{aligned} \tag{2.136}$$

It is worth mentioning that after the interaction, the photonic state is no longer a coherent state. The single-mode squeezed state is characterized by the coefficients

$$C_{2n} \equiv \frac{\sqrt{2n!}}{2^n n!}\frac{(-\tanh r)^n}{\sqrt{\cosh r}}, \quad C_{2n+1}=0, \quad n=0,\ldots,\infty \tag{2.137}$$

where, as in the two-mode squeezed state, the squeezing parameter r is related to the average number of signal photons through the relation $N_{\mathcal{S}} = \sinh^2 r$. The coherent state, instead, is defined by (2.43).

At each fixed interaction time t, the two possible atomic configurations give rise to two distinct photonic output states, $\rho^{(1)}_{\text{phot}}(t)$ and $\rho^{(2)}_{\text{phot}}(t)$. In order to quantify their distinguishability, we assume the implementation of the optimal Helstrom discrimination measurement between these two states, as described in Sec. 2.2 and App. B.1. Specifically, the minimal error probability $p_{\text{err}}(t)$ in Eq. (2.1) is obtained by projecting onto the positive eigenspace γ^{+} of the Helstrom operator $\rho^{(2)}_{\text{phot}}(t) - \rho^{(1)}_{\text{phot}}(t)$. The resulting error probabilities therefore represent the ultimate quantum limit for the discrimination task, independently of any specific detection scheme.

The error probability as a function of time is shown for lower signal photon numbers in Fig. 2.7a and for higher ones in Fig. 2.7b, in comparison with the coherent state. Since for realistic parameters (see Table 2.1) the relevant timescales are quite short, it is convenient to evaluate the time-averaged error probability, $\langle p_{\text{err}}\rangle_t$, to obtain a more accurate estimate. Thus, in Fig. 2.7c, we compare $\langle p_{\text{err}}\rangle_t$ for the coherent and single-mode squeezed states as a function of the same number of signal photons $N_{\mathcal{S}}$. From the plot, we observe that although the single-mode squeezed state exhibits a more stable error probability, the coherent state ultimately outperforms it, achieving a lower average error probability.

It is important to emphasize that the present approach is conceptually different from standard cavity-based spectroscopic techniques, such as transmission or reflection measurements. Rather than estimating the atomic transition frequency from a continuous spectral response, we formulate the problem as a

quantum hypothesis testing task. This distinction is particularly relevant in the low-photon-number regime, where conventional spectroscopic methods become inefficient (Kimble 1998; Maser et al. 2016), and where nonclassical resources can provide a genuine advantage. Moreover, in the case of two-mode squeezed states, the availability of an idler mode enables joint signal–idler measurements, which have no classical analogue in standard cavity transmission schemes and are known to enhance discrimination performance (Lloyd 2008; Tan et al. 2008).
.

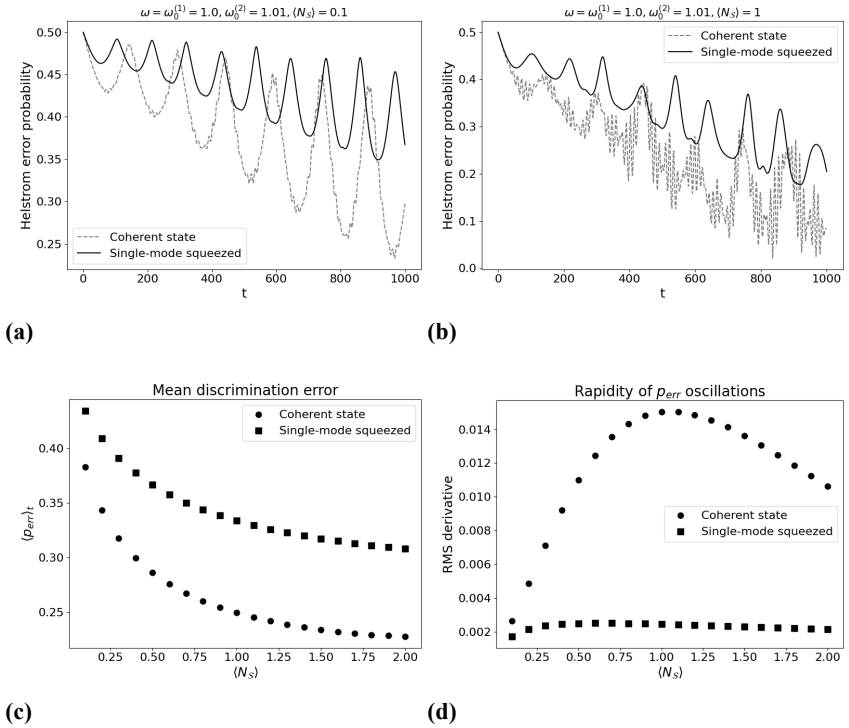


Figure 2.7: *Coherent state vs single-mode squeezed state. Error probability for (a)* $N_S = 0.1$*, (b)* N_S*=1. The values of the interaction strength is* $\lambda = 0.01$*. Error probability averaged over time (c) and the RMS of its derivative in time (d). Time is in natural unit (*$\hbar = 1$*).*

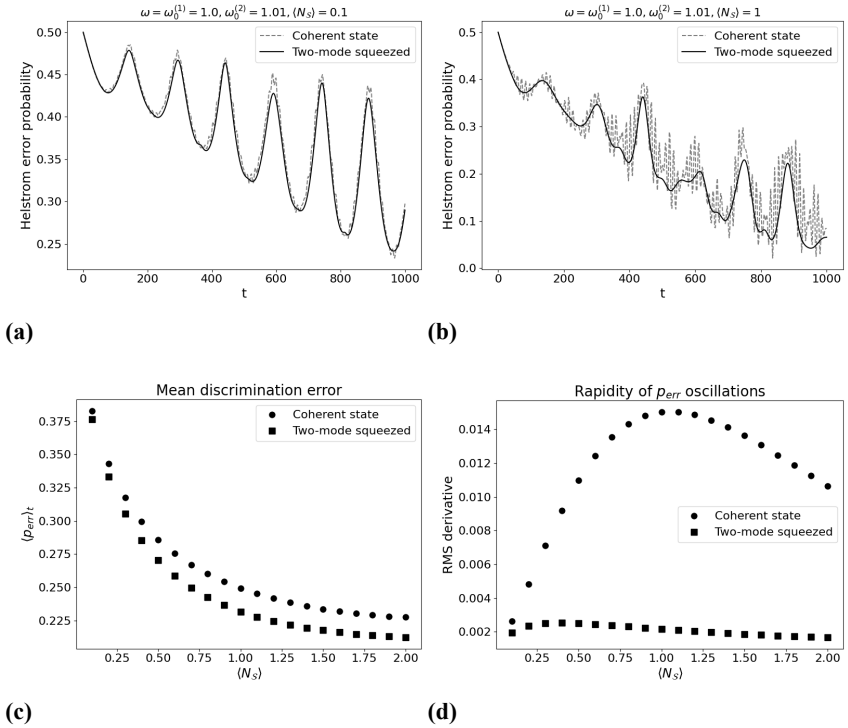


Figure 2.8: *Coherent state vs two-mode squeezed state. Error probability for (a) $N_{\mathcal{S}} = 0.1$, (b) $N_{\mathcal{S}}$=1. The values of the interaction strength is $\lambda = 0.01$. Error probability averaged over time (c) and the RMS of its derivative in time (d). Time is in natural unit ($\hbar = 1$).*

Two-mode squeezed vs coherent

In the case of the two-mode squeezed state, from Eq. (2.133), it is

$$|\Psi(t)\rangle = \sum_{n=0}^{\infty} \Big\{ \Big[C_{g,n}(t)\,|n,n\rangle + D_{g,n}\,|n+1,n\rangle \Big] \otimes |g\rangle + \Big[C_{e,n}(t)\,|n,n\rangle + D_{e,n}\,|n,n+1\rangle \Big] \otimes |e\rangle \Big\}, \tag{2.138}$$

$$\begin{aligned} C_{g,n}(t) &\equiv C_{n,n} C_g K_1(n-1,t), \quad D_{g,n}(t) \equiv C_{n,n} C_e K_2(n,t)\,, \\ C_{e,n}(t) &\equiv C_{n,n} C_e K_1(n,t), \quad D_{e,n}(t) \equiv C_{n+1,n+1} C_g K_2(n,t)\,. \end{aligned} \tag{2.139}$$

Hence, the reduced state of light is

$$\rho_{phot}^{(i)} = \sum_{n,n'=0}^{\infty} \Big[\Big(C_{g,n} C_{g,n'}^{*} + C_{e,n} C_{e,n'}^{*} \Big) |n, n\rangle \langle n', n'| + D_{g,n} D_{g,n'}^{*} |n+1, n\rangle \langle n'+1, n'| + D_{e,n} D_{e,n'}^{*} |n, n+1\rangle \langle n', n'+1| + D_{g,n} C_{g,n'}^{*} |n+1, n\rangle \langle n', n'| + C_{g,n} D_{g,n'}^{*} |n, n\rangle \langle n'+1, n'| + D_{e,n} C_{e,n'}^{*} |n, n+1\rangle \langle n', n'| + C_{e,n} D_{e,n'}^{*} |n, n\rangle \langle n', n'+1| \Big] . \tag{2.140}$$

As before, we compare the performance of the two-mode squeezed state with that of the coherent state, analyzing the same quantities shown in Figs. 2.8a–2.8d. In this case, the signal associated with the two-mode squeezed state is not only more stable, but it also achieves a lower error probability in the discrimination task, thus providing a clear quantum advantage.

Quantity	Value (n.u.)	Value (S.I.)
$\omega \simeq \omega_0^{(i)}$	1.0	$\omega_{\text{phys}} = 2\pi \cdot 10^9 \text{ s}^{-1} \approx 6.28 \times 10^9 \text{ rad/s}$
Δ	0.01	$\Delta_{\text{phys}} \approx 6.28 \times 10^7 \text{ rad/s}$
Λ	0.01	$\Lambda_{\text{phys}} \approx 6.28 \times 10^7 \text{ rad/s}$
t_{max}	1000	$t_{\text{phys}} = \dfrac{500}{\omega_{\text{phys}}} \approx 16.0 \times 10^{-8} \text{ s} = 160 \text{ ns}$

Table 2.1: *Conversion of dimensionless parameters in natural unit (n.u.), with $\hbar = 1$, to the International System of unit (SI), in a realistic scenario with reference frequency $f = 1$ GHz.*

2.5 Application to particle physics: Discrimination between Dirac and Majorana neutrinos

The question of whether neutrinos are Dirac or Majorana particles remains unresolved. Neutrinos possess several degrees of freedom (DoFs), such as flavor, chirality, and spin, which are described by states in different Hilbert spaces. To investigate the Dirac or Majorana nature of neutrinos, one must perform measurements of observables associated with one or more of these DoFs.

The aim of this section is to analyze how much information about the neutrino nature can be extracted through measurements on the different DoFs.

To this end, we describe the time evolution of a neutrino state using the Dirac

bispinor formalism within a quantum information framework, which provides a clear representation of the system and its subsystems (Bittencourt et al. 2021; Bittencourt et al. 2023; Bittencourt et al. 2024). For simplicity, we exclude the flavor degree of freedom and consider only chirality and spin.

By applying the Helstrom approach to derive the optimal error probability, we find that in the ultrarelativistic regime, measurements on chirality allow perfect discrimination between Dirac and Majorana neutrinos, whereas measurements on spin yield only partial information.

It is worth noting that measurements on neutrinos are extremely challenging, particularly those involving chirality. Furthermore, implementing the Helstrom measurement in an experimental setup is not always feasible. Nevertheless, the Helstrom bound remains a valuable figure of merit, as it quantifies the maximum amount of information that can be extracted in a discrimination procedure, meaning that no other measurement on the same observable can outperform it.

Readers who are not familiar with the Dirac equation and the distinction between Dirac and Majorana neutrinos are referred to Appendix B.3

2.5.1 Chiral part

Let us consider a simple example of a neutrino initially prepared in a left-handed state with negative helicity, which, in the Dirac bispinor formalism, can be written as (Bittencourt et al. 2021):

$$|\Psi^D(t)\rangle = N[f_+ e^{-iEt}\,|u_-(\mathbf{p},m)\rangle - f_- e^{-iEt}\,|v_-(-\mathbf{p},m)\rangle]\,, \tag{2.141}$$

with

$$|u_\pm(\mathbf{p},m)\rangle = N\begin{bmatrix} f_\pm\,|\pm\rangle \\ f_\mp\,|\pm\rangle \end{bmatrix}, \qquad |v_\pm(\mathbf{p},m)\rangle = N\begin{bmatrix} f_\pm\,|\pm\rangle \\ -f_\mp\,|\pm\rangle \end{bmatrix}, \tag{2.142}$$

and

$$f_\pm \equiv f_\pm(\mathbf{p},m) = 1 \pm \frac{p}{E+m}, \qquad N \equiv N(\mathbf{p},m) = \sqrt{\frac{E+m}{4E}}. \tag{2.143}$$

The time evolution in the Majorana case is instead given by

$$|\Psi^M(t)\rangle = \frac{1}{\sqrt{2}}\left(|\Psi^D(t)\rangle + |\Psi^D_\perp(t)\rangle\right)$$
$$\frac{N}{\sqrt{2}}\big[f_+ e^{-iEt}\big(\,|u_-(\mathbf{p},m)\rangle - |u_+(\mathbf{p},m)\rangle\,\big)$$
$$+ f_- e^{-iEt}\big(\,|v_-(-\mathbf{p},m)\rangle + |v_+(-\mathbf{p},m)\rangle\,\big)\big]\,, \tag{2.144}$$

with

$$|\Psi^D_\perp(t)\rangle \equiv N\big(f_- e^{-iEt}\big(\,|v_-(-\mathbf{p},m)\rangle + |v_+(-\mathbf{p},m)\rangle\,\big)\,. \tag{2.145}$$

Bispinors $|u\rangle$ and$|v\rangle$ contain both chiral and spin part which lie in different Hilbert spaces. For example we can rewrite

$$|u_-(\mathbf{p},m)\rangle = N(f_-\,|1\rangle_C + f_+\,|-1\rangle_C)\otimes|-\rangle\ , \tag{2.146}$$

and similar expressions for the others. $|\pm1\rangle_C$ are the states with left (+1) and right (-1) chirality and $|\pm\rangle$ are the states for up (+) and down (-) spin.

By using this new notation, we can rewrite

$$|\Psi^D(t)\rangle = (K_1^D\,|1\rangle_C + K_{-1}^D\,|-1\rangle_C)\otimes|-\rangle\ , \tag{2.147}$$

and

$$|\Psi^M(t)\rangle = \frac{1}{\sqrt{2}}\big[(K_1^D\,|1\rangle_C + K_{-1}^D\,|-1\rangle_C)\otimes|-\rangle$$
$$+ (K_1^M\,|1\rangle_C + K_{-1}^M\,|-1\rangle_C)\otimes|+\rangle\,\big]\,, \tag{2.148}$$

where

$$K_1^D = -i\frac{m}{E}\sin Et, \qquad K_{-1}^D = \cos Et - i\frac{p}{E}\sin Et\,,$$

$$K_1^M = -\frac{p}{E}\cos Et - i\sin Et\,, \qquad K_{-1}^M = -\frac{m}{E}\cos Et\,. \tag{2.149}$$

While the Dirac state is separable in spin and chirality, the Majorana state is not. Indeed, if we focus our attention on the chiral part, the reduced state for the Dirac case is:

$$|\Psi^D_C\rangle = K_1^D\,|1\rangle_C + K_{-1}^D\,|-1\rangle_C\ . \tag{2.150}$$

Tracing out the spin for the Majorana case, it is

$$\rho_C^M = |\Psi_C^D\rangle \langle \Psi_C^D| = \frac{1}{2}\begin{bmatrix} |K_1^D|^2 + |K_1^M|^2 & K_1^D {K_{-1}^D}^* + K_1^M {K_{-1}^M}^* \\ K_1^{D*} K_{-1}^D + K_1^{M*} K_{-1}^M & |K_{-1}^D|^2 + |K_{-1}^M|^2 \end{bmatrix} . \tag{2.151}$$

The optimal error probability can be calculated using the Helstrom formula in Eq. (2.1):

$$p_{\text{err}}^C = \frac{1}{2}\left(1 - \sqrt{-\det[\Delta_C]}\right) , \tag{2.152}$$

with $\Delta_C \equiv \rho_C^M - \rho_C^D$. For a 2×2 matrix with zero trace, the eigenvalues are $\lambda_\pm = \pm\sqrt{-\det[\Delta]}$.

Since

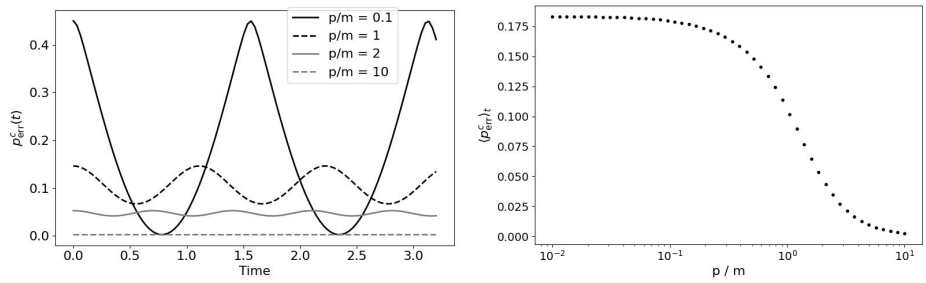
$$\Delta_C = \frac{1}{2}\begin{bmatrix} -\frac{p^2}{E^2} & \frac{pm}{E^2 \cos 2Et - i\frac{m}{E}\cos 2Et} \\ \frac{pm}{E^2 \cos 2Et + i\frac{m}{E}\cos 2Et} & \frac{p^2}{E^2} \end{bmatrix} , \tag{2.153}$$

it is

$$p_{\text{err}}^C = \frac{1}{2}\left[1 - \sqrt{\frac{p^4}{E^4} + \frac{m^2}{E^2}\left(\frac{p^2}{E^2}\cos^2 2Et + \sin^2 2Et\right)}\right] . \tag{2.154}$$

Note that the optimal error probability goes to zero in the ultrarelativistic limit ($p \gg m$), since $\frac{p}{E} \to 1$ and $\frac{m}{E} \to 0$, which means that *a measurement on chirality allows perfect discrimination between Dirac and Majorana neutrinos*.

In Fig. 2.9 it is plotted the optimal error probability for different values of the ratio $\frac{p}{m}$ as a function of time and averaged over time.



(a) **(b)**

Figure 2.9: *Optimal probability for the chirality case for different values of the ratios $\frac{p}{m}$ in 2.9a as function of time and in 2.9b averaged over time.*

2.5.2 Spin part

Now, focusing on the spin part, from Eq. (2.147) it is straightforward to see that the spin state is simply $|\Psi_S^D\rangle = |-1\rangle$. For the Majorana state, it is useful to rewrite Eq. (2.148) as

$$|\Psi^M\rangle = \frac{1}{\sqrt{2}}\Big[\,|1\rangle * C \otimes \big(K * 1^D\,|-\rangle + K_1^M\,|+\rangle\big) + |-1\rangle * C \otimes \big(K * -1^D\,|-\rangle + K_{-1}^M\,|+\rangle\big)\Big]. \tag{2.155}$$

Tracing out the chiral part, the reduced density matrix for the spin reads

$$\rho_S^M = \frac{1}{2}\begin{bmatrix} 1 & K_1^M K_1^{D*} + K_{-1}^M {K_{-1}^D}^* \\ K_1^D K_1^{M*} + K_{-1}^D K_{-1}^{M*} & 1 \end{bmatrix}, \tag{2.156}$$

and

$$\Delta_S \equiv \rho_S^M - \rho_S^S = \begin{bmatrix} -\frac{1}{2} & -\frac{m}{E}\cos 2Et - i\frac{mp}{2E^2}\sin 2Et \\ -\frac{m}{E}\cos 2Et + i\frac{mp}{2E^2}\sin 2Et & \frac{1}{2} \end{bmatrix}, \tag{2.157}$$

which leads to an error probability

$$p_{\text{err}}^S = \frac{1}{2}\left[1 - \frac{1}{2}\sqrt{1 + \frac{m^2}{E^2}\left(\cos^2 2Et + \frac{p^2}{E^2}\sin^2 2Et\right)}\right], \tag{2.158}$$

which in the ultrarelativistic limit ($p >> m$) gives

$$p_{\text{err}}^S \to \frac{1}{4} - \frac{m^2}{8E^2} \simeq \frac{1}{4} \equiv p_{max}^S\,. \tag{2.159}$$

Since in the ultrarelativistic regime the optimal error probability does not vanish, it is impossible to determine with certainty the nature of neutrinos from a single spin measurement.

In the opposite limit of very slow neutrinos ($m \gg p$), we obtain

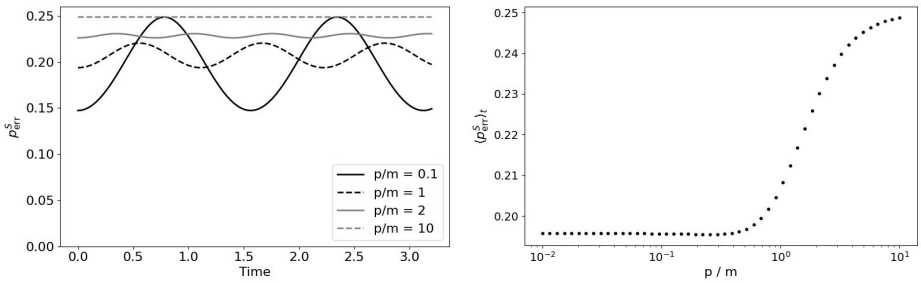
$$p_{\text{err}}^S \to \frac{1}{2}\left[1 - \frac{1}{2}\sqrt{1 + \cos^2 2Et}\right], \tag{2.160}$$

which attains its minimum value

$$p^S_{\text{min}} \to \frac{1}{2}\left[1 - \frac{1}{\sqrt{2}}\right] , \tag{2.161}$$

at specific instants $t = \frac{n\pi}{E}$, with n an integer. The quantities p^S_{max} and p^S_{min} represent the upper and lower bounds, respectively, as shown in Fig. 2.10.

Therefore, even when measuring the spin, there will always remain a nonzero probability of making an error.



(a) **(b)**

Figure 2.10: *Optimal probability for the spin case for different values of the ratios $\frac{p}{m}$. In this case, the error probability has an upper bound $p_{\text{max}} = \frac{1}{4}$, and a lower bound $p_{\text{min}} = \frac{1}{2}\left[1 - \frac{1}{\sqrt{2}}\right]$ obtained respectively in the cases $p \gg m$ and $m \gg p$ for time instants $t = \frac{n\pi}{E}$.*

2.5.3 Entanglement between spin and chirality in Majorana neutrinos

We have already mentioned that, for Dirac neutrinos and for the class of states considered here, spin and chirality remain separable, whereas for Majorana neutrinos they do not. Since the total state is pure, this can be shown by computing the purity of the reduced density operators. For any bipartite system AB, one has $\text{Tr}[\rho_A^2] = \text{Tr}[\rho_B^2]$, so the purity does not depend on which subsystem is traced out. The calculation yields

$$\mathcal{P} = \frac{1}{2} + \frac{m^2}{2E^2}\left(\cos^2 2Et + \frac{p^2}{E^2}\sin^2 2Et\right) . \tag{2.162}$$

We note that this expression coincides (up to an overall factor) with the quantity appearing under the square root in the optimal spin error probability of

Eq. (2.158), which can be rewritten as

$$p_{\text{err}}^{S} = \frac{1}{2}\left[1 - \sqrt{\frac{\mathcal{P}}{2}},\right] . \tag{2.163}$$

2.6 Conclusions and future directions

In this chapter, we briefly introduced several relevant concepts from detection theory, discussing in particular the Helstrom criterion for optimal discrimination between different hypotheses. We then reviewed photonic Gaussian states and the quantum illumination protocol, which combines the Helstrom discrimination strategy with Gaussian resources to detect low-reflectivity objects in a highly noisy environment. Two-mode squeezed Gaussian states provide the optimal light source for this protocol (De Palma and Borregaard 2018), outperforming the best classical strategy based on coherent states (Tan et al. 2008), as well as non-Gaussian resources. Building on these results, we analyzed a different protocol aimed at distinguishing between two atomic energy levels described by a Jaynes–Cummings model. We showed that, also in this case, two-mode squeezed states outperform coherent states, yielding a lower probability of error and greater temporal stability with respect to fast oscillations.

Despite these promising preliminary results, several questions remain open. Do two-mode squeezed states provide the optimal resource for this type of task? Can non-Gaussianity offer additional advantages?

As a future direction, we plan to extend the analysis to non-ideal setups, including the presence of noise, and to study more general light–matter interaction models, such as the fully quantum Rabi model or Dicke models.

In the final section, we analyzed another important problem in high-energy physics from a quantum information perspective: determining whether neutrinos are Dirac or Majorana particles. Since Dirac and Majorana neutrinos evolve with different wave functions, it is possible to exploit their internal degrees of freedom, spin and chirality, to extract information through the optimal Helstrom discrimination scheme.

In the present analysis, flavor oscillations were not included. This simplification is justified for sufficiently short propagation distances, where the time scale of flavor oscillations is much longer than that of chiral oscillations. However, a natural extension of this work is to incorporate flavor oscillations into the framework, enabling the study of the optimal error probability involving all three degrees of freedom: flavor, spin, and chirality (Bernardini and De Leo 2005; Bernardini 2006; Bernardini 2007; Bittencourt et al. 2024).

Furthermore, the same approach to optimal error probability can be applied to scenarios involving a lepton–antineutrino pair (Bittencourt et al. 2023). In

such cases, one may also consider the degrees of freedom of the accompanying charged lepton, such as its spin. If the optimal error probability remains sufficiently low in the ultrarelativistic regime, this could lead to potential experimental setups capable of probing the nature of neutrinos. This prospect is particularly appealing, as the spin of a charged lepton is generally more accessible in laboratory conditions than that of the neutrino itself.

Another possible direction for future research is the extension of the present analysis to the framework of quantum field theory, where neutrinos have also been extensively investigated in recent years (Blasone and Vitiello 1995; Blasone et al. 1999; Bittencourt et al. 2025b; Bittencourt et al. 2025a; Blasone and Zanfardino 2026; Blasone et al. 2025b; Blasone et al. 2025a).

Chapter 3

Quantum simulator for classical neaural networks

3.1 Introduction

The primary goal of this chapter is to investigate how classical complex systems—such as spin glasses and classical neural networks—can be simulated using quantum protocols, and to assess whether such approach offers computational advantages over classical simulation methods. In particular, we focus on the simulation of disordered and frustrated systems, whose dynamics are notoriously difficult to study due to the presence of rugged energy landscapes and a large number of metastable states.

To this end, after introducing the structure of the central nervous system in Section 3.2 as a biological and conceptual foundation, and subsequently presenting the basic mathematical models of neural networks, associative memories, and the Hopfield model in Sections 3.3and 3.4, we show in Section 3.5 that the dynamics of classical disordered systems—such as disordered Ising models or p-spin models, including Hopfield models—can be efficiently mapped onto and simulated by a quantum optical setup. We restrict our analysis to all-to-all interacting systems and study the scaling of the computational time, demonstrating that, at least in principle, this approach can be competitive with the best available classical simulators for disordered and frustrated systems. We also propose a realistic optical architecture for an experimental implementation of the simulator.

Furthermore, in Section 3.6, we introduce a more general and platform-independent simulation scheme based on qubit systems. We show that, using n qubits, it is possible to simulate the dynamics of $N = 2^n$ classical spin glasses defined on arbitrary interaction graphs, identifying quantum coherence as the fundamental resource underlying the protocol. Within this framework, we also

present an algorithm that achieves an exponential speed-up in the update of spin configurations according to the Metropolis rule.

The discussion of neural networks and, in particular, of the central nervous system at the beginning of this chapter is motivated by the close conceptual and mathematical relationship between biological neural systems, associative memory models, and disordered spin systems. Classical neural networks, such as the Hopfield model, were originally inspired by the architecture and functioning of real biological networks, and at the same time admit a natural formulation in terms of interacting spins and energy minimization. For this reason, a brief overview of the structure of the central nervous system provides a useful physical and biological perspective on the origin of the models studied in this chapter.

For additional details on complex systems and spin-glass models, the reader may refer to Appendix C.1.

3.2 The structure of the central nervous system

The central nervous system represents a paradigmatic example of a complex, highly interconnected system composed of a large number of interacting units. Its study has historically played a crucial role in the development of theoretical models of neural computation and associative memory, which later inspired simplified mathematical descriptions in terms of artificial neural networks and spin systems.

All neurons share the same fundamental components, regardless of their size or shape (Müller et al. 2012): a central structure called the *cell body* or *soma*, and several root-like extensions known as *dendrites*. Each neuron also possesses a single tubular fiber called the *axon*, which branches at its terminal end into multiple smaller extensions. The cell body typically measures approximately $10–18,\mu$m in diameter, while both dendrites and axons have diameters of only a few micrometers.

Dendrites act as receptors for signals from neighboring neurons, whereas the axon transmits neural activity to other nerve cells. The connection between the end of an axonal branch and another neuron is called a *synapse*, where the two cells are separated by a 200 nm gap known as the *synaptic cleft*.

Neural signals are transmitted electrically within a neuron and chemically between different neurons.

Electrical transmission is initiated by a chemical discharge originating in the cell body, which travels along the axon and reaches its synaptic terminals.

In the resting state, the interior of the neuron (the *protoplasm*) is negatively charged with respect to the surrounding neural fluid, exhibiting a resting potential of -70 mV. This potential is maintained by the cell membrane, which is impermeable to Na^+ ions.

When signals arrive at the synaptic connections, the resting potential undergoes *depolarization*. If the potential increases above $-60\,\mathrm{mV}$, the membrane suddenly becomes permeable to Na^+ ions, resulting in an electrical discharge that propagates along the axon toward the synapses. The membrane then gradually returns to its initial state, restoring the resting potential over the course of several milliseconds, during which time the neuron cannot be excited again.

Although in principle a single synapse can induce a neuron to "fire," this is rarely the case. The cell body operates as a *summing* device that integrates the depolarizing effects of multiple input signals. These signals decay over a characteristic time of 5–$10\,\mathrm{ms}$. However, if several signals arrive at the same synapse within this time window, their excitatory effects may accumulate.

A high rate of repeated firings in a neuron corresponds to a strong signal intensity. When the total depolarization in the cell body exceeds a critical threshold (about $10\,\mathrm{mV}$), the neuron fires.

The strength of a synaptic connection depends on several factors, and there is substantial evidence that it is not permanently fixed. As originally proposed by Hebb (Hebb 2005), the magnitude of synaptic connections can be modified in response to changes in activity. An active synapse that repeatedly contributes to the activation of its postsynaptic neuron will become stronger, whereas others will gradually weaken.

This mechanism is known as *Hebb's rule* and plays a fundamental role in the complex process of learning.

To convey an order-of-magnitude estimate of the size and complexity of the human brain, it is useful to consider the structure of the human cortex. Each cortical neuron is estimated to receive converging inputs from approximately 10^4 synapses, while transmitting its output to hundreds of other neurons. The neuronal density is about 1.5×10^5 neurons per mm^2, leading to an estimated total of approximately 3×10^{10} neurons and on the order of 10^{15} synaptic connections in the human brain. Notably, the majority of these synaptic connections are formed during the first few months after birth.

3.3 Neural networks - basic concepts

In mathematical terms, a neural network model can be defined as a graph with the following properties (Müller et al. 2012):

1. State variables n_i associated with each node i, where $i = 1, \ldots, N$.
2. Real-valued weights J_{ik} associated with each link between nodes i and k.
3. A real-valued bias θ_i associated with each node i.

4. A transfer function $f_i(n_k, J_{ik}, \theta_i)$ for $k \neq i$, defined for each node i, which determines the state of node i as a function of its bias, the weights of the links, and the states of the nodes connected to it.

Two examples of neural networks are shown in Fig. 3.1.

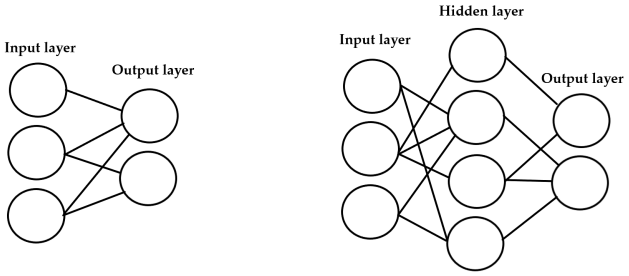


Figure 3.1: *Examples of neural networks with two layers (input and output) and more than two layers.*

The nodes are called *neurons*, the links are referred to as *synapses*, and the bias is known as the *activation threshold*. The transfer function typically takes the form $f\left(\sum_k J_{ik} s_k - \theta_i\right)$, where $f(x)$ is typically a discontinuous step function or a sigmoidal function, as shown in Figure 3.2.

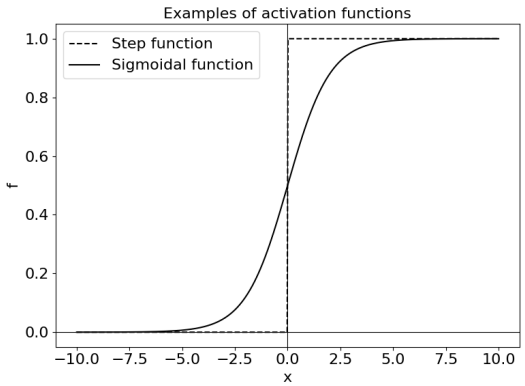


Figure 3.2: *Examples of activation functions.*

Nodes without incoming links are called *input neurons*, while those with no outgoing links are called *output neurons*. A network is called a *feed-forward network* if it contains no closed paths

3.3.1 Network dynamic

In 1943, a general theory of information processing was proposed based on networks of binary switching elements, which we now call neurons (McCulloch and Pitts 1943). They assumed $n_i \in \{0, 1\}$, with

$$0 \rightarrow \text{resting state}, \quad 1 \rightarrow \text{active state}. \tag{3.1}$$

They also assumed that changes occur in discrete steps $t = 0, 1, \ldots$, where t denotes an abstract, discrete time index labeling successive updates of the network state rather than physical time. At each time step, the state of all neurons is updated according to the same deterministic rule, defining the dynamical evolution of the system. This influence is described by the function

$$h_i(t) = \sum_j J_{ij} n_j(t), \tag{3.2}$$

which represents the post-synaptic polarization potential for the neuron.

The properties of the neural network are completely determined by the functional relation between $h_i(t)$ and $n_i(t + 1)$. In the simplest case, there is a certain threshold θ_i for the neuron n_i, with

$$n_i(t + 1) = \theta(h_i(t) - \theta_i), \tag{3.3}$$

where $\theta(x)$ is the step function defined by

$$\theta(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ 1, & \text{if } x > 0. \end{cases} \tag{3.4}$$

This is called the McCulloch-Pitts model. In their work, they analyzed its relation to a Turing machine. The problem is how to choose the correct weights J_{ij} to accomplish a specific task.

The same question was addressed by Caianiello in (Caianiello 1961), who proposed a learning algorithm to determine J_{ij}. This algorithm, called mnemonic equations, incorporates the basic principle of Hebb's learning rule.

Subsequently, further studies (Rosenblatt 1962) led to a particular kind of network called the *perceptron*, due to its similarity with the biological mechanisms of processing sensory information.

A perceptron, in its simplest form, consists of two separate layers of neurons: the input layer and the output layer

The neurons in the output layer receive signals from the input layer but not vice versa, and the neurons within a single layer do not communicate with each other. Thus, the flow of information is unidirectional. Furthermore, a

mechanism was introduced to determine the couplings J_{ij} for a given outcome.

Nevertheless, it was shown in (Minsky and Papert 1969) that certain problems cannot be solved by a perceptron with only two layers, as considered in (Rosenblatt 1962). For example, the exclusive-or (XOR) gate, in which one output depends on two inputs and is active if exactly one of the two inputs is active. It has been demonstrated that there is no solution for the couplings J_1^{out} and J_2^{out} in a single-layer perceptron.

It is now well known that the XOR problem can be solved by a three-layer perceptron, although no algorithm was available at the time to determine the couplings for such systems.

Another important step in the study of neural networks was taken by Little (Little 1974), who discovered a one-to-one correspondence between the McCulloch–Pitts model and Ising-like models of magnetization, in which the spins take values $s_i = \pm 1$. Hopfield later studied similar models (Hopfield 1982), with the main difference being that in Little's model the spins are updated simultaneously, whereas in Hopfield's model they are updated one at a time at each time step (either in a fixed sequence or randomly).

In particular, a strong analogy arises between neural network models and spin-glass models when the deterministic rule

$$s_i(t+1) = \theta\big(h_i(t) - \theta_i\big) \tag{3.5}$$

is replaced by a stochastic update function. The associated probability function includes a parameter T that plays the role of an effective "temperature," analogous to the temperature in magnetization models. Although T is not a physical temperature, it introduces stochasticity and disorder into the network. In the limit of vanishing stochasticity, the McCulloch–Pitts model is recovered.

In some cases, stochastic evolution can make the network less vulnerable to dynamical instabilities, thereby improving its overall performance. While this may seem counterintuitive at first, it is consistent with the fact that biological evolution has adapted the human brain to operate at finite temperatures rather than at absolute zero.

3.4 Associative memories

Associative memory refers to the process of storing and recalling information by associating it with other information. A storage device is called associative memory if it allows information to be recalled based on partial knowledge of its content without requiring knowledge of its storage location (also known as content-addressable memory). This is not the case in conventional computers, where recalling information requires precise knowledge of the relative memory address.

For instance, if we follow the example in (Müller et al. 2012), when trying to recall the name of a certain person, it is completely useless to know that it was the 2728th name we encountered during our lifetime. However, it is useful to know that it starts with a certain letter, even if the choice is not unique.

We can define associative memory as follows: let us suppose that we have $|\Lambda|$ binary patterns containing N bits of information ν_i^μ, $(i = 1, \ldots, N;\ \ \mu = 1, \ldots, |\Lambda|)$ that are stored in the memory. If a new pattern n_i is presented, that stored pattern ν_i^λ has to be recalled which most strongly resembles the presented pattern. This means that ν_i^λ and n_i should differ in as few places as possible, i.e., the square deviation

$$H_\mu = \sum_{i=1}^{N} (n_i - \nu_i^\mu)^2 \tag{3.6}$$

is minimal for $\mu = \lambda$. H_μ is called the Hamming distance.

At this point, we ask whether it is possible to store a set of patterns in a neural network in such a way that each stored pattern corresponds to a stable state of the network dynamics. The desired behavior is that, when the network is initialized in a configuration close to one of the stored patterns, its autonomous time evolution drives it toward the corresponding pattern, thereby enabling pattern retrieval through the intrinsic dynamics of the system.

Let us define for simplicity new variables

$$s_i = 2n_i - 1, \qquad \xi_i = 2\nu_i - 1 \tag{3.7}$$

which assume values ± 1 instead of 0 and 1. The squared deviations are

$$(n_i - \nu_i^\mu)^2 = \frac{1}{4}(s_i - \xi_i^\mu)^2 = \frac{1}{4}[s_i^2 - 2s_i\xi_i^\mu + (\xi_i^\mu)^2] = \frac{1}{2}(1 - s_i\xi_i^\mu) \tag{3.8}$$

Therefore, one has to find the maximal value of

$$A_\mu(s_i) = \sum_{i=1}^{N} \xi_i^\mu s_i, \qquad (\mu = 1, \ldots, |\Lambda|) \tag{3.9}$$

for a given pattern s_i. A_μ is called scalar product or overlap of the $N-$component vectors ξ_i^μ and s_i (see also C.1). According to Hopfield, the dynamic evolution of the network is now defined as follows. To the individuals neurons are assigned new values $s_i(t+1)$ in some randomly chosen sequence. The new solution are computed according to the law

$$s_i(t+1) = \text{sgn}[h_i(t)] \equiv \text{sgn}\left[\sum_{i=1}^{N} J_{ij} s_j(t)\right] \tag{3.10}$$

where we assume $J_{ii} = 0$ for any $i = 1, \ldots, N$.

3.4.1 Hebbian learning

Now the problem is to find the couplings J_{ij} such that the constraints mentioned in the previous section are satisfied. Here and in the following, the term *network dynamics* refers to the discrete-time update rule introduced in Sec. 3.3.1, according to which the state of each neuron at time $t+1$ is determined by the local field $h_i(t)$ through the threshold activation function. First of all, we will show that each stored pattern corresponds to a stable configuration of the network and, after that, we will see that small deviations from it will be automatically corrected by the network dynamics (stability). For simplicity, we will start with the case of a single stored pattern ξ_i and then we will generalize to the case of multiple stored patterns.

The state corresponding to this pattern remains invariant under the dynamics. This happens when $h_i = \sum_j J_{ij}\xi_j$ has the same sign as ξ_i. Such a condition is satisfied by choosing

$$J_{ij} = \frac{1}{N}\xi_i\xi_j \tag{3.11}$$

indeed, assuming $s_i(t) = \xi_i$, since $\xi_i^2 = 1$, we have

$$h_i = \frac{1}{N}\sum_{j=1}^{N}\xi_i\xi_j^2 = \frac{1}{N}\xi_i\sum_{j=1}^{N}1 = \xi_i \tag{3.12}$$

Hence

$$s_i(t+1) = \text{sgn}(h_i) = \text{sgn}(\xi_i) = \xi_i = s_i(t) \tag{3.13}$$

For stability, if we start with the network in another state different from ξ_i, let us assume without loss of generality, that the state is

$$s_i(t=0) = \begin{cases} -\xi_i, & \text{for } i = 1, \ldots, n \\ \xi_i, & \text{for } i = n+1, \ldots, N \end{cases} \tag{3.14}$$

for some $n \leq N$. Then we have

$$h_i(t=0) = \sum_{j=1}^{N} J_{ij} s_j(t=0) = \frac{1}{N}\xi_i \sum_{j=1}^{N} \xi_j s_j(t=0)$$
$$= \frac{1}{N}\xi_i\big[-n + (N-n)\big] = \left(1 - \frac{2n}{N}\right) \tag{3.15}$$

Therefore, if $n \leq \frac{N}{2}$ it is

$$\text{sgn}\left[\left(1 - \frac{2n}{N}\right)\xi_i\right] = \text{sgn}(\xi_i) \tag{3.16}$$

and

$$s_i(t=1) = \text{sgn}\left(1 - \frac{2n}{N}\right)\text{sgn}(\xi_i) = \xi_i \tag{3.17}$$

Hence the transition is corrected in one time step. In this case, we say that the stored pattern acts as an *attractor* fot the network dynamics.

Note that in the case $n \geq \frac{N}{2}$, the state is the opposite $s_i(t=1) = -\xi_i$. In the case of $|\Lambda|$ stored patterns $\xi_i^1, \ldots \xi_i^{|\Lambda|}$, Hebb's rule generalizes as

$$J_{ij} = \frac{1}{N}\sum_{\mu=1}^{|\Lambda|} \xi_i^\mu \xi_j^\mu \tag{3.18}$$

Hence, if $\xi_i^\mu = \xi_j^\mu$ (both active or non-active), J_{ij} is positive for most patterns and the synapse becomes excitatory. If $\xi_i^\mu = -\xi_j^\mu$ the contribution to the synapse is inhibitory. For a selected pattern ν, we can study the stability:

$$h_i^{(\nu)} = \sum_{j=1}^{N} J_{ij}\xi_j^\nu = \frac{1}{N}\sum_{\mu=1}^{|\Lambda|}\xi_i^\mu \sum_{j=1}^{N}\xi_j^\mu \xi_j^\nu$$
$$= \frac{1}{N}\left[\xi_i^\nu \sum_{j=1}^{N}(\xi_j^\nu)^2 + \sum_{\mu\neq\nu=1}^{|\Lambda|}\xi_i^\mu \sum_{j=1}^{N}\xi_j^\mu\xi_j^\nu\right]$$
$$= \xi_i^\nu + \frac{1}{N}\sum_{\mu\neq\nu=1}^{|\Lambda|}\xi_i^\mu\sum_{j=1}^{N}\xi_j^\mu\xi_j^\nu \tag{3.19}$$

The first term is the same as one stored pattern. The second term contains all $N(|\Lambda|-1)$ randomly signed contributions (± 1). For large N and $|\Lambda|$, by the

central limit theorem, its value is of the order of

$$\sim \frac{1}{N}\sqrt{N(|\Lambda|-1)} = \sqrt{\frac{|\Lambda|-1}{N}} \tag{3.20}$$

Therefore, if $|\Lambda| \ll N$ this term goes to 0 and does not influence $h_i^{(\nu)}$.

We define

$$\alpha = \frac{|\Lambda|}{N} \tag{3.21}$$

as a control parameter that, as will be shown below, distinguishes between the *memory retrieval phase*, i.e., the regime in which stored patterns can be correctly recalled, and the *memory black-out phase*, in which memory retrieval fails.

Indeed, for $\alpha \ll 1$, it is

$$s_i(t=1) = \text{sgn}[h_i^{(\nu)}(t=0)] = \text{sgn}[\xi_i^\nu] = \xi_i^\nu \tag{3.22}$$

Furthermore, if $n < N$ neurons start in the "wrong" state, we have

$$\begin{aligned} h_i &= \left(1 - \frac{2n}{N}\right)\xi_i^\nu + \frac{1}{N}\sum_{\mu\neq\nu=1}^{|\Lambda|}\xi_i^\mu \sum_{j=1}^{N}\xi_j^\mu\xi_j^\nu \\ &= \left(1 - \frac{2n}{N}\right)\xi_i^\nu + O\left(\sqrt{\frac{|\Lambda|-1}{N}}\right) \end{aligned} \tag{3.23}$$

Again, under the conditions $n, |\Lambda| \ll N$ we have $\text{sgn}(h_i) = \xi_i^\nu$ and the network will converge to the desired pattern in only one (global) update.

If the number of patterns $|\Lambda|$ is similar to the number of neurons N, the second term in Eq. (3.23) is randomly distributed and it affects the sign of h_i. It is possible to prove that under certain conditions, this occurs when $\alpha \geq 0.14$.

The limit is higher when various of the pattern are orthogonal to each other:

$$\frac{1}{N}\sum_{i=1}^{N}\xi_i^\mu\xi_i^\nu = \delta_{\mu\nu} \tag{3.24}$$

since the second term is zero.

Finally, we mention that there are learning protocols which are superior to Hebb's rule and are able to storage of up to $2N$ patterns also in presence of non-zero overlaps (Müller et al. 2012), as the p-body Hopfield model, briefly discussed in Sec. 3.4.8.

3.4.2 Application to magnetic connection

In the Hopfield network, the energy of a given configuration of neuron states $\{s_i\}$ can be expressed through the Hamiltonian

$$\mathcal{H} = -\frac{1}{2}\sum_{i\neq j=1}^{N} J_{ij}s_i s_j, \tag{3.25}$$

where the couplings J_{ij} are symmetric. This formulation makes explicit the analogy with the classical Ising model in statistical mechanics, where each $s_i = \pm 1$ represents a spin variable and J_{ij} corresponds to the interaction between spins. In this picture, a Hopfield network can be viewed as an Ising system with interactions specifically designed to store certain patterns as energy minima.

The contribution of a given neuron i to the energy is

$$E_i(t) = -s_i(t)\left[\sum_{j\neq i} J_{ij}s_j(t)\right] = -s_i(t)h_i(t), \tag{3.26}$$

where $h_i(t)$ is the local field acting on neuron i.

If the couplings are determined by Hebb's rule (3.18) and the configuration corresponds to a stored pattern ξ_i^ν, the energy can be written as

$$E = -\frac{1}{2N}\sum_{i,j=1}^{N}\sum_{\mu=1}^{|\Lambda|}\xi_i^\mu\xi_j^\mu\xi_i^\nu\xi_j^\nu = -\frac{1}{2N}\left[N^2 + \sum_{\mu\neq\nu=1}^{|\Lambda|}\left(\sum_{i=1}^{N}\xi_i^\mu\xi_i^\nu\right)^2\right]. \tag{3.27}$$

Assuming the patterns are uncorrelated and the values ± 1 are equally distributed, the term $\sum_i \xi_i^\mu\xi_i^\nu$ is of order $\sqrt{N}$. Consequently, we have

$$E \simeq -\frac{1}{2N}\left\{N^2 + O[(|\Lambda|-1)N]\right\} = \frac{N}{2} + O\left(\frac{|\Lambda|-1}{2}\right). \tag{3.28}$$

Therefore, the presence of all other patterns causes a slight shift of the total energy.

3.4.3 Effect of spin errors on the energy landscape

If the spins have the wrong orientation with respect to the stored patterns, the fluctuating term is not significantly affected. Nevertheless, the term with $\mu = \nu$

undergoes a substantial change. Indeed, if the first n spins are incorrectly oriented (i.e., $s_i = -\xi_i$ for $i = 1, \ldots, n$), we have

$$\sum_{i,j=1}^{N} \xi_i^\nu \xi_j^\nu s_i s_j = \left(\sum_{i=1}^{N} \xi_i^\nu s_i \right)^2 = \left(-\sum_{i=1}^{n} \xi_i^\nu \xi_i^\nu + \sum_{i=n+1}^{N} \xi_i^\nu \xi_i^\nu \right)^2$$
$$= (N - 2n)^2. \tag{3.29}$$

Hence, the energy of the configuration rises in proportion to the extent of its deviation from the stored pattern:

$$E \simeq -\frac{1}{2N}(N - 2n)^2 + O\left(\frac{|\Lambda| - 1}{2}\right) = E[\xi^\nu] + 2n - \frac{2n^2}{N}. \tag{3.30}$$

Therefore, the stored patterns are (at least local) minima of the energy.

As we will see in Appendix C.1, the energy functional has, in addition, infinitely many other local minima. However, all these spurious minima are less pronounced than the ones associated with the stored pattern ξ_i^μ, at least for moderate values of $\alpha = \frac{|\Lambda|}{N}$.

The presence of many local minima in the energy landscape, including both the stored patterns and spurious states, implies that simulating the dynamics of such a network can become computationally expensive, especially for large N and moderate to high memory loads α. This combinatorial complexity is precisely the reason why alternative platforms capable of accelerating the exploration of the energy landscape are of interest. In particular, as we discuss in the following sections, quantum systems offer the potential to simulate these dynamics more efficiently, providing a promising route to study large networks and disordered systems that are otherwise difficult to explore classically.

3.4.4 Parallel and sequential dynamics

Starting from the single-neuron energy in Eq. (3.26), if the neuron i is updated with the deterministic law, we have

$$E_i(t+1) = -s_i(t+1)\left[\sum_{j \neq i} J_{ij} s_j(t)\right] = -\text{sgn}[h_i(t)]h_i(t) = -|h_i(t)| \tag{3.31}$$

$$\leq -s_i(t)h_i(t) = E_i(t), \tag{3.32}$$

which clearly decreases the energy.

Such dynamics has been studied by Glauber (Glauber 1963b) and are called

Glauber dynamics. These dynamics ensure that, for an appropriate neural activation function, the network enters thermodynamic equilibrium.

The same argument does not work for the synchronous dynamics (Little model). Since all neurons assume new states in parallel at the same moment, the contribution of an individual neuron to the energy cannot be considered separately. In these cases, a condition for stability is needed considering the following *Liapunov function* (Bruck and Goodman 1988): since

$$E_L(t+1) - E_L(t) = -\sum_{i,j=1}^{N} J_{ij} s_i(t)[s_j(t+1) - s_j(t-1)], \tag{3.33}$$

the network must eventually settle into a steady state with the same network configuration being repeated every two time step:

$$s_i(t+1) = s_i(t-1), \quad i = 1, \ldots, N. \tag{3.34}$$

Hence, the network can reach either a steady state, as in the Hopfield model, or permanently cycle between two different configurations (Bruck and Goodman 1988; Marcus and Westervelt 1989).

The reason why only cycles of length two can occur in synchronous dynamics is a consequence of the form of the Liapunov function and Eq. (3.34). Since $E_L(t)$ is a quadratic form in the neuron states and the couplings are symmetric, the network cannot enter cycles of length greater than two without violating the condition $E_L(t+1) \leq E_L(t)$. Therefore, the system either reaches a fixed point (a stable configuration) or a two-step oscillation, but longer cycles are forbidden by the energy constraints inherent in the network dynamics.

3.4.5 Stochastic Neurons and mean activity

In this section, we briefly analyze what happens when the deterministic update rule (3.5) is replaced by a stochastic one. In the stochastic case, neuron i at time step $t+1$ assumes the value $+1$ or -1 with a certain probability P. We require this probability to be defined by a function such that $P(+1) = f(h_i)$, $P(-1) = f(-h_i)$, and

$$P[s_i(t+1)] = f[h_i(t)], \tag{3.35}$$

where $f \to 0$ as $h \to -\infty$ and $f \to 1$ as $h \to +\infty$. Functions with this behavior are called *sigmoidal functions*. A commonly used example is the Fermi distribution:

$$f(h) = \frac{1}{1 + e^{-2\beta h}}, \tag{3.36}$$

which also satisfies the symmetry property $f(h) + f(-h) = 1$. In this context, the parameter β plays the role of an inverse temperature. Note, however, that β does not correspond to a physical temperature in the usual thermodynamic sense.

In the limit $\beta \to \infty$ (i.e., $T \to 0$), the Fermi function tends to the step function $\theta(h)$.

The state of a single neuron in the stochastic regime is not particularly informative, as it is determined randomly. However, the *mean activity* of a neuron exhibits meaningful properties. For a single neuron, the mean activity is given by

$$\langle s \rangle = (+1) \cdot f(h) + (-1) \cdot f(-h) = \frac{1 - e^{-2\beta h}}{1 + e^{-2\beta h}} = \tanh(\beta h), \tag{3.37}$$

which corresponds to the expected value of the neuron's state. As $\beta \to \infty$, this expression approaches $+1$ for $h > 0$ and -1 for $h < 0$, in agreement with the deterministic case.

In general, calculating the mean activity in the stochastic regime is not straightforward (see Appendix C.1 for more details on disordered and frustrated systems). In particular, due to the nonlinearity of the function f, we typically have

$$\langle f(h) \rangle \neq f(\langle h \rangle), \tag{3.38}$$

It should be noted that the approximation $\langle f(h_i) \rangle \simeq f(\langle h_i \rangle)$ is valid only when the fluctuations of the local field h_i around its mean are small, and correlations between the neurons are negligible. In highly frustrated or disordered networks, or when the number of stored patterns is large, these conditions may be violated, making the approximation less accurate. In such cases, the full distribution of h_i must be considered, and more sophisticated methods–such as the replica technique discussed in Appendices C.1 and C.1.2–are required to compute the mean activity reliably.

Assuming that

$$\langle f(h) \rangle \simeq f(\langle h_i \rangle) = f\left(\sum_{j=1}^{N} J_{ij} \langle s_j \rangle \right), \tag{3.39}$$

we obtain

$$\langle s_i \rangle = \langle f(h_i) \rangle - \langle f(-h_i) \rangle = \tanh\left(\beta \sum_{j=1}^{N} J_{ij} \langle s_j \rangle \right), \tag{3.40}$$

which constitutes a set of N nonlinear equations in N unknowns $\langle s_i \rangle$.

3.4.6 Single-pattern solution

The solution of Eq. (3.40) depends on the couplings J_{ij}. To gain insight into the possible behaviors of the system, let us consider the simplest case of uniform couplings,$J_{ij} = \frac{1}{N}$. Under this assumption, we obtain

$$\langle s_i \rangle = \tanh\left(\frac{\beta}{N}\sum_{j=1}^{N}\langle s_j \rangle\right). \tag{3.41}$$

Since the right-hand side does not depend on i, by symmetry we have

$$\langle s_i \rangle = \langle s \rangle = \tanh(\beta\langle s \rangle). \tag{3.42}$$

The solutions of this self-consistency equation fall into two different cases, illustrated in Figure 3.3.

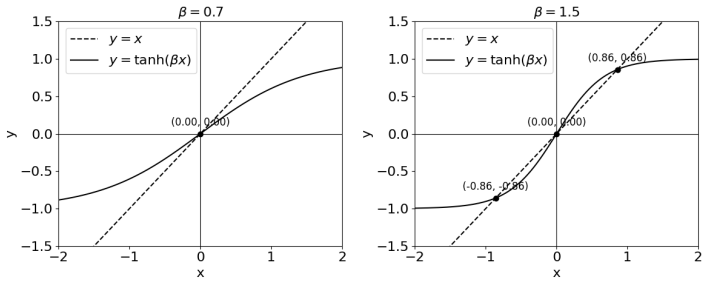


Figure 3.3: *Numerical intersection between* x *and* $\tanh(\beta x)$ *for* $\beta = 0.7$ *and* $\beta = 1.5$.

For sufficiently low values of β (i.e., high temperature T), the only solution is $\langle s \rangle = 0$, indicating that no preferred configuration exists. As β increases beyond a critical value $\beta_c = 1$, two additional non-zero solutions, $\pm S_0$, appear. In this regime, the zero solution becomes unstable: small deviations from $\langle s \rangle = 0$ grow under iteration of the self-consistency equation, driving the system toward one of the stable solutions $\pm S_0$. Consequently, neurons tend to be, on average, either more in the resting state ($-S_0$) or more in the active state ($+S_0$), analogous to spontaneous magnetization in mean-field theory of ferromagnets.

Nevertheless, we emphasize once again that mean-field theory is not exact for stochastic neurons and only provides a qualitative description. This limitation

arises because the approximation in Eq. (3.39) does not hold in general. A more accurate treatment is offered by spin-glass theory, as discussed in Appendix C.1.

3.4.7 Multiple patterns

In the case of multiple stored patterns, it can be shown that the equation becomes

$$\langle s_i \rangle = \tanh\left(\frac{\beta}{N}\sum_{j=1}^{N}\sum_{\mu=1}^{|\Lambda|}\xi_i^\mu \xi_j^\mu \langle s_j \rangle\right), \tag{3.43}$$

which does not admit an immediate analytic solution.

Previously, we showed that in deterministic networks, the stored patterns are stable configurations. Assuming that a similar situation holds in the stochastic case, we consider a solution of the form

$$\langle s_i \rangle = m\xi_i^\nu, \tag{3.44}$$

where m is a proportionality factor. Substituting into the mean-field equation gives

$$\begin{aligned}\langle s_i \rangle &= \tanh\left[\beta m\xi_i^\nu + \frac{\beta m}{N}\sum_{\mu\neq\nu}^{|\Lambda|}\xi_i^\mu\sum_{j=1}^{N}\xi_j^\mu\xi_j^\nu\right] \\ &= \tanh\left[\beta m\xi_i^\nu + \beta m\, O\left(\sqrt{\frac{|\Lambda|-1}{N}}\right)\right].\end{aligned} \tag{3.45}$$

The factor m is then determined by the boundary condition in the limit $N \to \infty$ with $|\Lambda| \ll N$, as in the single-pattern case.

Therefore, in this scenario as well, at high temperature T (i.e., low β), the only solution is the trivial one with zero magnetization, $\langle s_i \rangle = 0$. When the temperature is sufficiently low, the stored pattern is not always recognized perfectly in practice. This imperfection also depends on the number of stored patterns, quantified by the parameter $\alpha = \frac{|\Lambda|}{N}$. If only a few patterns are stored, the system can successfully retrieve them. However, as the number of stored patterns increases, the system eventually fails to recognize any of them.

This behavior is characteristic of a phase transition, where the control parameters are the temperature T and the load parameter α. For high temperatures, the system never settles into a stable configuration, and by analogy with magnetic systems, this regime is called the *paramagnetic phase*. For sufficiently low temperatures and small α, the system reliably evolves toward stable configurations

corresponding to the stored patterns; this is known as the *retrieval phase*. If T is low but α is large, the system exhibits zero magnetization and fails to retrieve any pattern. This regime is called the *glassy phase*. Unlike the paramagnetic phase, in the glassy phase one must consider a different order parameter—the overlap between different replicas of the system—introduced in Appendix C.1.

3.4.8 Generalized p-body Hopfield model

As we have discussed for the classical 2-body Hopfield model, the number of patterns that can be reliably stored scales linearly with the system size N, with a storage capacity of $\alpha \approx 0.14$. A natural generalization of this model is the Hopfield p-body network, also referred to as the *dense associative memory model* (Peretto and Niez 1986; Baldi and Venkatesh 1987; Gardner 1987; Abbott and Arian 1988; Horn and Usher 1988; Arenzon and Almeida 1993; Kosmatopoulos et al. 1995; Krotov and Hopfield 2016; Krotov and Hopfield 2016; Agliari et al. 2020; Ramsauer et al. 2021; Agliari et al. 2023; Albanese et al. 2024; Albanese et al. 2025).

The p-body Hopfield model can be defined through the Hamiltonian

$$\mathcal{H} = -\sum_{i_1<i_2<\cdots<i_p} J_{i_1 i_2 \ldots i_p}\, \xi_{i_1}\xi_{i_2}\ldots\xi_{i_p}, \tag{3.46}$$

where the spins $\sigma_i = \pm 1$ represent the state of neuron i and the p-body couplings are given by the generalization of the Hebbian prescription in Eq. (3.18):

$$J_{i_1 i_2 \ldots i_p} = \frac{1}{N^{p-1}} \sum_{\mu=1}^{|\Lambda|} \xi_{i_1}^{\mu}\xi_{i_2}^{\mu}\ldots\xi_{i_p}^{\mu}. \tag{3.47}$$

Here, $\Lambda \equiv \{\xi_i^{\mu}\}$ is the set of stored patterns.

This construction reduces to the classical 2-body Hopfield model for $p = 2$

In the p-body Hopfield model, interactions involve p neurons simultaneously rather than pairs, leading to a dramatically increased storage capacity. Specifically, the number of storable patterns now scales as N^{p-1}, which makes the p-body formulation particularly relevant for high-capacity associative memory applications. This generalization preserves the associative retrieval properties of the classical Hopfield network while allowing for a significantly denser encoding of information.

3.5 Optical simulator for the p-body Hopfield model

Recent advancements in photonic quantum technologies have spotlighted bosonic quantum interference in linear-optical networks as a cornerstone for applications ranging from fault-tolerant quantum computing (Knill et al. 2001; Gimeno-Segovia et al. 2015; Bartolucci et al. 2023) to quantum advantage demonstrations (Aaronson and Arkhipov 2013) and high-precision metrology (Pirandola et al. 2018). Experiments now manipulate systems exceeding 100 photons, achieving computational power rivaling that of supercomputers (Zhong et al. 2020; Madsen et al. 2022; Aspuru-Guzik and Walther 2012).

In this section, based on our recent work (Zanfardino et al. 2026a), we will show a correspondence between multiphoton interference in linear optics and generalized p-body Hopfield models.

We show that a photonic setup composed of: (i) N_{ph} photons in superposition across M modes, (ii) a layer of M phase shifters, and (iii) a linear-optical interferometer, can implement the dynamics of a $p = 2N_{\text{ph}}$ Hopfield model. This architecture enables analog simulation of high-order HMs, mitigating the computational overhead inherent in classical simulations of such models.

Experimental demonstrations of optical HMs with $p = 2$ have shown promise for efficient spin system simulations (Farhat et al. 1985; Leonetti et al. 2021), while recent photonic systems explore Hopfield-inspired memory functions (Leonetti et al. 2024). By extending these ideas to $p = 4$, we provide photonic evidence of the hallmark phase transition from retrieval to spin-glass behavior.

This photonic-statistical physics correspondence opens new avenues for using quantum photonic systems to simulate and analyze high-order neural networks, offering potentially exponential advantages in speed and scalability.

3.5.1 Mapping

We consider a linear-optical quantum processor composed of single-photon inputs, a universal interferometer (realizable through phase shifters and beam splitters (Reck et al. 1994; Clements et al. 2016)), and photodetectors. This architecture aligns with the principles of scalable photonic quantum computing (Knill et al. 2001; Gimeno-Segovia et al. 2015; Bartolucci et al. 2023) and near-term quantum simulators (Sparrow et al. 2018). Photonic systems offer unique advantages, including ultrafast operations, intrinsic parallelism, and low energy consumption (Leonetti et al. 2024; Leonetti et al. 2021; Fan et al. 2023; Pierangeli et al. 2019), making them promising platforms for optical computing.

We propose a mapping between the dynamics of N_{ph} indistinguishable photons in a linear-optical network with M modes and a classical Hopfield

model (HM) with $p = 2N_{\text{ph}}$-body interactions. A Fock state is specified by a configuration vector $\vec{c} = \{c_1, \ldots, c_{N_{\text{ph}}}\}$, where each photon occupies one of M available modes. The total number of such configurations is $|C_M| = \binom{N_{\text{ph}}+M-1}{N_{\text{ph}}} \sim M^{N_{\text{ph}}}/N_{\text{ph}}!$ in the large-M limit.

The transition amplitude from an input configuration $\vec{c}$ to an output configuration $\vec{k}$ through an interferometer characterized by a unitary matrix U is given by the permanent of a submatrix (Scheel 2004) (see Appendix C.2):

$$\langle \vec{k}| \mathcal{U}(|\vec{c}\rangle) = \frac{1}{\sqrt{\mu(\vec{c})\mu(\vec{k})}} \text{Perm}\left(U_{\vec{k}|\vec{c}}\right), \tag{3.48}$$

where $\mu(\vec{c})$ and $\mu(\vec{k})$ are the mode multiplicities of the configurations $\vec{c}$ and $\vec{k}$, respectively and, for a given matrix A, we define

$$A_{\vec{x}|\vec{k}} = \begin{pmatrix} A_{x_1,k_1} & A_{x_1,k_2} & \ldots & A_{x_1,k_{N_{\text{ph}}}} \\ A_{x_2,k_1} & A_{x_2,k_2} & \ldots & A_{x_2,k_{N_{\text{ph}}}} \\ \vdots & \vdots & & \vdots \\ A_{x_{N_{\text{ph}}},k_1} & A_{x_{N_{\text{ph}}},k_2} & \ldots & A_{x_{N_{\text{ph}}},k_{N_{\text{ph}}}} \end{pmatrix}. \tag{3.49}$$

The corresponding output detection probability from an input state $|\vec{c}\rangle$ and an output state $|\vec{k}\rangle$ is

$$\Pr_{\mathcal{U}}(\vec{k}|\vec{c}) = \frac{1}{\mu(\vec{c})\mu(\vec{k})} \left|\text{Perm}\left(U_{\vec{k}|\vec{c}}\right)\right|^2. \tag{3.50}$$

To establish the mapping to the p-body Hopfield model, we initialize the photonic system in a general superposition state $|\psi\rangle = \sum_{\vec{c}} a_{\vec{c}} |\vec{c}\rangle$. We then apply local phase shifts $\phi_c \in \{0, \pi\}$ to each mode, which define Ising spin variables $\sigma_c = e^{i\phi_c} = \pm 1$. This modifies the amplitudes as:

$$a'_{\vec{c}} = a_{\vec{c}} \prod_{j=1}^{N_{\text{ph}}} \sigma_{c_j}. \tag{3.51}$$

Subsequently, the state is passed through a scattering medium, described by a unitary transformation S. The amplitude to observe an output configuration $\vec{k}$ becomes (Scheel 2004; Zanfardino et al. 2026a)

$$\langle \vec{k}| \mathcal{S}(|\psi'\rangle) = \sum_{\vec{c}\in C_M} \frac{a_{\vec{c}}}{\sqrt{\mu(\vec{c})\mu(\vec{k})}} \text{Perm}(S_{\vec{k}|\vec{c}}) \prod_{j=1}^{N_{\text{ph}}} \sigma_{c_j}. \tag{3.52}$$

The probability of detecting an output configuration $\vec{k} \in \Lambda$, where Λ is a subset of all possible outputs, is

$$\Pr(\Lambda|\boldsymbol{\sigma}) = \sum_{\vec{x},\vec{y}\in C_M} J_\Lambda(\vec{x},\vec{y}) \prod_{j=1}^{N_{\text{ph}}} \sigma_{x_j}\sigma_{y_j}, \tag{3.53}$$

where the coupling tensor is given by

$$J_\Lambda(\vec{x},\vec{y}) = \frac{a_{\vec{x}}a^*_{\vec{y}}}{\sqrt{\mu(\vec{x})\mu(\vec{y})}} \sum_{\vec{k}\in\Lambda} \frac{\text{Perm}(S_{\vec{k}|\vec{x}})\,\text{Perm}^*(S_{\vec{k}|\vec{y}})}{\mu(\vec{k})}. \tag{3.54}$$

This defines an effective $p = 2N_{\text{ph}}$ Hopfield Hamiltonian:

$$\mathcal{H}[\boldsymbol{\sigma}|\Lambda] = -M\ \Pr(\Lambda|\boldsymbol{\sigma}), \tag{3.55}$$

where Λ acts as the set of stored memory patterns.

In the specific case of $N_{\text{ph}} = 2$, the system implements 4-body interactions, and the coupling tensor becomes

$$\begin{aligned} J_\Lambda(x_1,x_2,y_1,y_2) = \frac{a_{\vec{x}}a^*_{\vec{y}}}{\sqrt{\mu(\vec{x})\mu(\vec{y})}} &\sum_{[k_1,k_2]\in\Lambda} \frac{1}{\mu(\vec{k})} \\ &\cdot [S_{k_1,x_1}S_{k_2,x_2} + S_{k_1,x_2}S_{k_2,x_1}] \\ &\cdot [S_{k_1,y_1}S_{k_2,y_2} + S_{k_1,y_2}S_{k_2,y_1}]^* . \end{aligned} \tag{3.56}$$

3.5.2 DFT for a fully connected model

In order to realize a fully connected Hopfield network, it is advantageous for the input superposition coefficients $a_{\vec{x}}$ to be as uniform as possible. This can be achieved by distributing the N_{ph} photons over the M modes in a symmetric fashion. We begin with the initial Fock state in which all N_{ph} photons occupy mode 0:

$$|\vec{0}\rangle = \frac{\left(\hat{a}_0^\dagger\right)^{N_{\text{ph}}}}{\sqrt{N_{\text{ph}}!}} |\mathbf{0}\rangle = |N_{\text{ph}},0,\ldots,0\rangle . \tag{3.57}$$

We then apply a Discrete Fourier Transform (DFT) unitary transformation, whose matrix elements are defined as

$$U^{\text{DFT}}_{kl} = \frac{1}{\sqrt{M}}\, e^{-2\pi i \frac{kl}{M}}, \tag{3.58}$$

where $k, l \in \{0, 1, \ldots, M-1\}$ and M denotes the total number of modes, corresponding to the size of the DFT matrix.

The mode transformation acts on the annihilation operators as

$$\hat{a}_i \longrightarrow \sum_j U_{ij}^{\mathrm{DFT}}\, \hat{a}_j, \tag{3.59}$$

where the transformation is applied identically to all photons, which are initially localized in mode 0.

From definition in Eq. (3.49), the amplitude of the state $|\vec{x}\rangle$ after applying the DFT to the initial state $|\vec{0}\rangle$ is then given by

$$a_{\vec{x}} = \langle \vec{x} |\, \mathcal{U}_{\mathrm{DFT}}\, |\vec{0}\rangle = \frac{1}{\sqrt{N_{\mathrm{ph}}!\, \mu(\vec{x})}}\, \mathrm{Perm}\left(U_{\vec{x}|\vec{0}}^{\mathrm{DFT}}\right), \tag{3.60}$$

where $U_{\vec{x}|\vec{0}}^{\mathrm{DFT}}$ is a matrix whose N_{ph} columns are all equal to the first column of U^{DFT}, corresponding to mode 0. Therefore,

$$\begin{aligned} a_{\vec{x}} &= \frac{1}{\sqrt{N_{\mathrm{ph}}!\, \mu(\vec{x})}}\, \mathrm{Perm}\left(\begin{bmatrix} U_{\vec{x}|0}^{\mathrm{DFT}} & U_{\vec{x}|0}^{\mathrm{DFT}} \cdots\cdots\cdots U_{\vec{x}|0}^{\mathrm{DFT}} \end{bmatrix}\right) \\ &= \frac{1}{\sqrt{N_{\mathrm{ph}}!\, \mu(\vec{x})}} \sum_{\sigma \in \pi(\vec{x})} \prod_{i \in \vec{x}} U_{i,0}^{\mathrm{DFT}} \\ &= \sqrt{\frac{N_{\mathrm{ph}}!}{\mu(\vec{x})}} \prod_{i \in \vec{x}} \frac{1}{\sqrt{M}} = \sqrt{\frac{N_{\mathrm{ph}}!}{M^{N_{\mathrm{ph}}}\, \mu(\vec{x})}}. \end{aligned} \tag{3.61}$$

Thus, the resulting state after the DFT is a uniform superposition over all configurations:

$$|\Psi\rangle = \sum_{\vec{x}} a_{\vec{x}}\, |\vec{x}\rangle\,, \tag{3.62}$$

which serves as the desired input for the associative memory protocol. At this stage, we apply the controlled phase shifts and the scattering transformation defined by the unitary matrix S, which results in the output probability distribution given in Eq. (3.53).

3.5.3 Order parameters for retrieval and glassy phase

We define the parameter identifying a generalized *memory pattern* embedded in the interferometric scattering system. Such a pattern is labeled by the output configuration $\vec{k} = \{k_1, k_2\}$ of the two-photon current and is given by the

component $\vec{k}, \vec{x}$ of the permanent of the scattering matrix, which we denote using the tensor (Zanfardino et al. 2026a)

$$\mathbb{X}_{\vec{x}}^{(\vec{k})} \equiv \text{Perm}\left(S_{\vec{k}|\vec{x}}\right) = S_{k_1,x_1}S_{k_2,x_2} + S_{k_1,x_2}S_{k_2,x_1}. \tag{3.63}$$

The overlap between the input photonic mode phases $\boldsymbol{\sigma} = \{\sigma_1, \ldots, \sigma_N\}$ and such a matrix pattern is defined as

$$m_{\vec{k}} \equiv \frac{1}{M}\,\boldsymbol{\sigma}^T \mathbb{X}^{(\vec{k})}\,\boldsymbol{\sigma} \tag{3.64}$$

In terms of Eq. (3.64), the cost function (see Eq. (3.53)) can be rewritten as

$$\mathcal{H}[\boldsymbol{\sigma}|\Lambda] = -M\text{Pr}(\Lambda|\boldsymbol{\sigma}) = M\sum_{\vec{k}\in\Lambda}\left|\hat{m}_{\vec{k}}\right|^2, \tag{3.65}$$

where the normalized overlaps

$$\hat{m}_{\vec{k}} \equiv \frac{m_{\vec{k}}}{\sqrt{\sum_{k_1,k_2}^{1,M}|m_{\{k_1,k_2\}}|^2}} \tag{3.66}$$

are introduced so that their real and imaginary parts lie in $[-1, 1]$.

For every target $\vec{k}$ and scattering matrix S there corresponds a permanent $\mathbb{X}_{\vec{x}}^{(\vec{k})}$, represented as a rank-2 tensor (an $M \times M$ matrix indexed by x_1 and x_2), encoding a memory embedded in the neural network.

The Hamiltonian in Eq. (3.65) is constructed such that its minima with respect to $\boldsymbol{\sigma}$ are attracted to the maxima of individual $m_{\vec{k}}$ values, i.e., each $\text{Perm}(S_{\vec{k}|\ldots})$ acts as an attractor in the energy landscape of the $\boldsymbol{\sigma}$ dynamics.

This behavior characterizes the retrieval phase, provided that the number of such patterns is smaller than the maximum capability of the system. However, as the pattern density α increases, new spurious attractors emerge, which do not coincide with the planted ones, as described in the previous sections. When the storage capacity of the network is saturated, these spurious attractors dominate the dynamics and manifest as spin-glass states. Eventually, the original attractors are lost altogether—this marks the onset of the blackout phase.

In the spin-glass phase, the minima $\boldsymbol{\sigma}$ of the Hamiltonian do not correspond to any scattering-matrix-permanent pattern. Spin-glass states have no relation to the scattering elements embedded into the Hopfield neural network via Eqs. (3.54) and (3.56). The memory overlaps (3.66) become vanishingly small, and there is no absolute reference to identify a state.

For this reason, one resorts to measuring the similarity between parallel dynamical histories of the system. That is, one ideally considers the overlap between configurations evolving in exact *replicas* of the same system (with

the same random matrix S). This overlap order parameter does not refer to a specific state (such as an alignment, a vector pattern, or a matrix permanent), but rather quantifies the similarity between two trajectories within the same rugged energy landscape.

If we denote by $\boldsymbol{\sigma}^{(a)}$ the M-component spin configuration of replica a, the overlap is defined as (see also Appendix C.1)

$$q_{ab} \equiv \frac{1}{M} \sum_{j=0}^{M-1} \sigma_j^{(a)} \sigma_j^{(b)}. \tag{3.67}$$

If replica symmetry holds, all states are equivalent, and every thermalized pair of replicas will exhibit the same overlap.

If the energy landscape is more complex, as in structural glasses (Parisi et al. 2020) and in proper spin glasses (Mézard et al. 1987; Mydosh 1993), replica symmetry breaks at low temperature, and Eq. (3.67) can take on different values, with varying probability, reflecting the intricate organization of metastable states. In this case, the entire distribution $P(q)$ becomes the relevant order parameter, also known as the Parisi order parameter.

3.5.4 Metropolis photonic simulation and autocorrelation function

From now on, we focus on the $N_{\text{ph}} = 2$ case, which maps into a 4-body Hopfield model (4HM), as shown in Eq. (3.56). This equivalence can be exploited to extract the energy of a 4HM via the analog measurement of the occurrence probability in Eq. (3.53). The simulations presented here were performed using the fully bunched input configuration, shown in Fig. 3.5d and detailed in 3.5.6.

In a single experimental iteration $n-1 \to n$, one of the mode phases is flipped ($0 \to \pi$ or $\pi \to 0$), and the new two-photon current is then measured. The change is straightforwardly accepted in the simulated dynamics if $\mathcal{H}$ decreases (i.e., if $\Pr(\Lambda|\vec{\sigma})$ increases). However, if we introduce synaptic noise in the form of a statistical mechanical *temperature* T, the change can also be accepted even if the energy cost momentarily *increases*, according to the Markov Chain Monte Carlo Metropolis algorithm (Metropolis et al. 1953), that is, for $\Delta\mathcal{H} \equiv \mathcal{H}[\boldsymbol{\sigma}(n)] - \mathcal{H}[\boldsymbol{\sigma}(n-1)]$:

$$\begin{aligned} &\text{if } \Delta\mathcal{H} < 0 && \text{change accepted} \\ &\text{if } \Delta\mathcal{H} > 0 && \text{change accepted with probability } e^{-\Delta\mathcal{H}/T}. \end{aligned} \tag{3.68}$$

At each Monte Carlo step t (i.e., a sequence of M single phase-shift iterations), the configuration $\boldsymbol{\sigma}(t)$ is recorded. Different time configurations are then used

to compute the self-correlation function $F_{\text{self}}(\tau)$, defied by

$$F(\tau) = \sum_{i=1}^{M} \langle \sigma_i(t)\sigma_i(t+\tau)\rangle_t \equiv \frac{1}{N_{\text{MCs}} \cdot N} \sum_{i=1}^{M} \sum_{t=0}^{N_{\text{MCs}}} \sigma_i(t)\sigma_i(t+\tau), \qquad (3.69)$$

where N_{MCs} denotes the number of Monte Carlo steps, and the averaging is performed over the Monte Carlo steps t and $\tau \in [0, N_{\text{MCs}}]$.

Results for $F_{\text{self}}(\tau)$ at various temperatures and for two different values of the storage size ratio $\alpha \equiv P/M^{N_{\text{ph}}} = P/M^2$ are reported in Fig. 3.4 for $M = 50$. The parameter α is the ratio between the number of output scattering patterns of the two-photon current, $P = |\Lambda|$, and their maximum scaling with the number M of degrees of freedom in the network for this p-neuron interacting model mapping a photonic interferometric quantum system. That is, $P_{\text{all}} \sim M^{p/2} = M^2$.[1]

For small $\alpha = 0.0004$, corresponding to a single two-photon output channel, Fig. 3.4a shows that $F_{\text{self}}(\tau)$ tends to a plateau when T is sufficiently low, indicating relaxation to the unique fixed point associated with memory retrieval. As synaptic noise increases, at a certain critical temperature T, the configuration $\boldsymbol{\sigma}$ decorrelates to zero (faster for higher T), and retrieval is eventually lost.

As α increases (Fig. 3.4b), $F_{\text{self}}(\tau)$ decays to zero at progressively lower temperatures, making memory retrieval increasingly difficult. The emergence of spurious excited states (Amit 1989) further drives the dynamics toward attractors that do not correspond to the scattering patterns embedded in the network via Eq. (3.56). These embedded patterns are global minima of the energy landscape $\mathcal{H}[\boldsymbol{\sigma}|\Lambda]$, while the spurious states correspond to local minima.

The presence of such excited states implies that a plateau in $F_{\text{self}}(\tau)$ does not necessarily indicate successful retrieval if the initial configuration $\boldsymbol{\sigma}(0)$ lies within the *basin of attraction* of a spurious state. Ultimately, when α exceeds a critical value, the proliferation of spurious states leads to the onset of a spin-glass phase, resulting in memory black-out (Amit et al. 1985).

Fig. 3.4c) compares the fluctuations induced by nonzero synaptic noise ($T > 0$) with those arising from measurement noise (see 3.5.7). In the figure, open circles represent the standard deviation σ_T of temperature-induced energy fluctuations as a function of T. Solid lines, on the other hand, denote the standard deviation σ_{exp} due to experimental noise.

The proposed analog simulation remains valid only when the optical measurement allows the estimation of $\Pr(\Lambda|\boldsymbol{\sigma}(t))$ with an error that is negligible compared to the fluctuations induced by the simulation temperature, i.e., $\sigma_{\text{exp}} < \sigma_T$.

[1] This differs from the standard generalized p-spin Hopfield models with P stored memory patterns and M neurons, where, despite variations in model definitions (Gardner 1987; Abbott and Arian 1988; Horn and Usher 1988), the typical scaling is always $\alpha \sim P/M^{p-1}$.

By increasing the experimental "exposure time"—which controls the number of detected photon pairs given a source with finite pair-production rate—σ_{exp} can be reduced, thereby expanding the accessible measurement window.

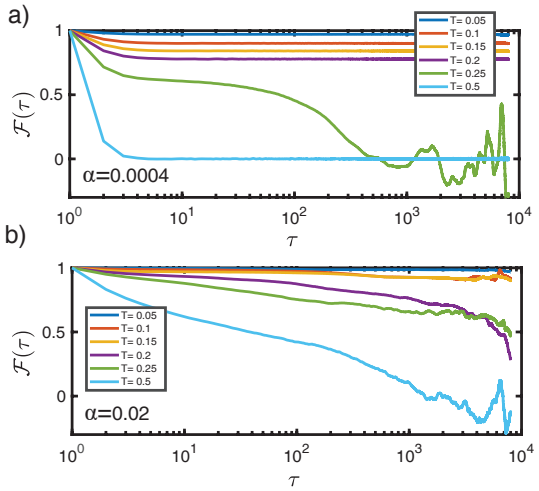


Figure 3.4: *Panels a) and b) report $\mathcal{F}(\tau)$ for various temperatures (see legend), for storage size ratio values $\alpha = 0.0004$ (retrieval regime) and $\alpha = 0.02$ (spin-glass phase), respectively.*

3.5.5 Analysis of the spin-glass phase transitions

An interesting property of the Hopfield model (HM) is the emergence of spin-glass phase transitions as the parameter α increases. In this section, we analyze the behavior of the spin systems mapped through our approach and provide numerical evidence of such a phase transition. To explore the full phase space through numerical simulations on a classical computer, we consider a photonic system configuration that is particularly well-suited for classical simulations, as shown in Fig. 3.5d.

Specifically, we focus on the case in which the largest set $\Lambda = \{(0,0),(1,1),\ldots,(M-1,M-1)\} \equiv \Lambda_{\text{FB}}$ consists of all configurations where the photons are fully bunched in any of the output modes. In this scenario, the output probability for arbitrary values of M and N_{ph}, cf. Eqs. (3.53, 3.54), and for any $\Lambda \subseteq \Lambda_{\text{FB}}$, simplifies to (see 3.5.6):

$$\Pr(\Lambda|\boldsymbol{\sigma}) = \frac{1}{M^{N_{\text{ph}}}} \sum_{k\in[0,M-1]} \left| \sum_{j=0}^{M-1} S_{k,j}\sigma_j \right|^{2N_{\text{ph}}}, \tag{3.70}$$

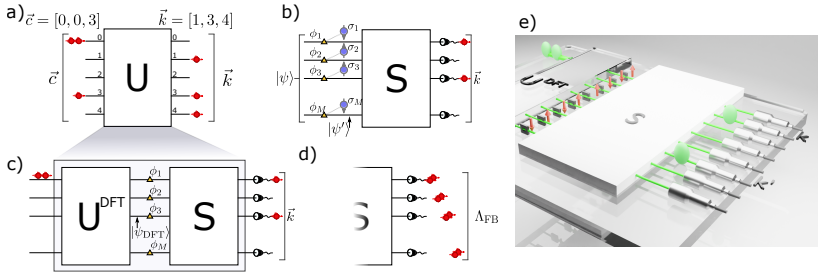


Figure 3.5: *a) Standard description of the linear optical transformation from an input configuration $\vec{c}$ to an output configuration $\vec{k}$ through a linear scattering matrix U. An example is shown for $N_{\text{ph}} = 3$ photons and $M = 5$ modes. b) Schematic of the mapping of a pHM to a photonic system composed of an input state $|\psi\rangle$, a set of M phase-shifters which can take binary values $\phi_i \in \{0, \pi\}$ and correspond to spin states $\sigma_i = exp(i\phi_i) = \pm 1$, and a scattering matrix S. c) Full schematic including the initial Discrete Fourier Transform U^{DFT} for generating the near-uniform input state $|\psi_{DFT}\rangle$. d) Simplified scheme with the set of output states Λ_{FB} given by fully bunched configurations, where all photons exit the same output mode. e) "On-chip" experimental scheme corresponding to panel c).*

The index k runs over a subset of optical modes, while j runs over all modes. Considering this model significantly accelerates classical simulations while preserving the full structure of the mapped Hopfield model (HM). We employ this configuration to study the phase diagram of the 4HM case.

According to Eq. (3.70), we can access a storage coefficient $\alpha \leq 1/M$, e.g., $\alpha \leq 0.02$ for $M = 50$.

Monte Carlo simulations are implemented using the Exchange Monte Carlo (EMC) method (Hukushima and Nemoto 1996), where independent system dynamics at different temperatures are run in parallel across multiple CPUs. All results are obtained for a system of size $M = 50$ spins, with a total of 2×10^5 Monte Carlo steps used for system thermalization. Exchanges between replicas at different temperatures are attempted every 200 Monte Carlo steps.

The scattering matrix S is sampled randomly from the Haar measure, and each EMC simulation is repeated over 20 different disorder realizations, each corresponding to an independently generated S matrix. For every scattering sample, a total of 36 replicas are simulated for each pair of values (α, T) in order to extract the relevant order parameters.

Specifically, we consider two standard order parameters from spin-glass and memory retrieval theory: the overlap between simulated replicas q_{ab}, and the *memory* overlap $\hat{m}_{\vec{k}}$ (see 3.5.3).

The results are presented in Fig. 3.6, where we identify three distinct phases: the retrieval phase, the spin-glass phase, and the paramagnetic phase.

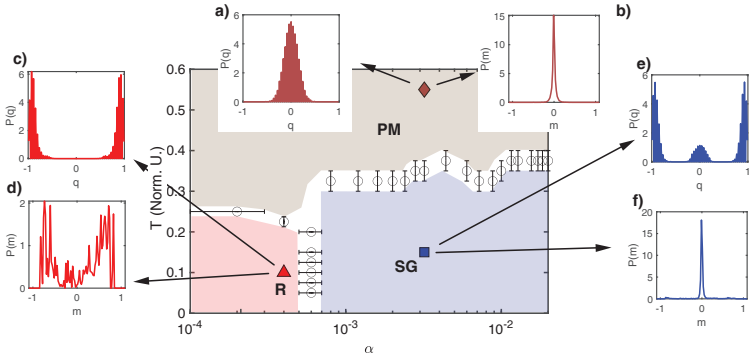


Figure 3.6: *Phase Diagram of the 4-Hopfield model realized with a quantum interferometer of 2-photons on $M = 50$ modes. The insets report $P(q)$ and $P(m)$ for the paramagnetic phase (a), b), $\alpha = 0.0032$; $T = 0.55$), retrieval phase (c), d) $\alpha = 0.0004$; $T = 0.1$), and spin glass phase (e), f) $\alpha = 0.0032$; $T = 0.15$).*

3.5.6 *Efficient probability calculation for fully-bunched model*

In this section we show how the probabilities can be efficiently calculated for the simplified model, shown in Fig. 3.5d, where Λ_{FB} is composed of all on only the configuration where output photons are bunched together (i.e. are in the same output mode).

Because the input contains all photons in the same mode (assumed to be the 0-th for simplicity), the probability for a generic output configuration $\vec{k} = \{k_1, \dots, k_{N_{\text{ph}}}\}$ can be calculated via Eq. (3.50) as:

$$\begin{aligned}\Pr(\vec{k}|\boldsymbol{\sigma}) &= \left|\langle\vec{k}|\mathcal{U}(|\vec{0}\rangle)\right|^2 = \frac{1}{N_{\text{ph}}!\mu(\vec{k})}\left|\text{Perm}\left(U_{\vec{k}|\vec{0}}\right)\right|^2 \\ &= \frac{N_{\text{ph}}!}{\mu(\vec{k})}\prod_{i=1}^{N_{\text{ph}}}\left|S\cdot\text{diag}(\boldsymbol{\sigma})\cdot U^{\text{DFT}}\right|^2_{k_i,0} = \frac{N_{\text{ph}}!}{M^{N_{\text{ph}}}\mu(\vec{k})}\prod_{i=1}^{N_{\text{ph}}}|S\boldsymbol{\sigma}|^2_{k_i} \\ &= \frac{N_{\text{ph}}!}{M^{N_{\text{ph}}}\mu(\vec{k})}\prod_{i=1}^{N_{\text{ph}}}\left|\sum_{j=0}^{M-1}S_{k_i,j}\sigma_j\right|^2_{k_i},\end{aligned} \tag{3.71}$$

where $\text{diag}(\boldsymbol{\sigma})$ is a diagonal matrix having for entries the spin values $\sigma_1, \dots, \sigma_M$ and where we have used that for the Discrete Fourier Transform elements in the

0-th column are simply given by $U_{i,0}^{\text{DFT}} = 1/\sqrt{M}$. The situation is simplified even further when we consider an output configuration where all output photons bunch in the same μ-th mode, i.e., in which $k_i = \mu \in [0, M-1]$, for each photon $i \in [1, N_{\text{ph}}]$.

For this case the output probability is simplified to

$$\Pr(\vec{k_\mu}|\boldsymbol{\sigma}) = \frac{1}{M^{N_{\text{ph}}}} \left| \sum_{j=0}^{M-1} S_{\mu,j}\sigma_j \right|^{2N_{\text{ph}}} . \tag{3.72}$$

Defining the set $\Lambda_{\text{FB}} = \{\vec{k_0}, \ldots, \vec{k_{M-1}}\}$ composed of all configurations with bunched photons on any of the modes, when considering the sum over more output photonic configurations of this kind, composing a subset $\Lambda \subseteq \Lambda_{\text{FB}}$ we obtain:

$$\Pr(\Lambda|\boldsymbol{\sigma}) = \frac{1}{M^{N_{\text{ph}}}} \sum_{\mu\in[0,M-1]} \left| \sum_{j=0}^{M-1} S_{\mu,j}\sigma_j \right|^{2N_{\text{ph}}} , \tag{3.73}$$

where by $\mu \in [0, M-1]$ we mean that the index μ run through a subset of all possible modes in $[0, M-1]$.

3.5.7 Effects of noise

The measured probability $\Pr(k_1, k_2)$, see Eq. (3.53), is affected by two distinct sources of uncertainty. The first originates from measurement noise. The error on the average frequency of observing the photonic output configuration $\vec{k}$ can be estimated by repeating N_{exp} random transmissions that either yield or do not yield a detection of configuration $\vec{k}$. As a binomial process, the experimental average frequency of occurrences of $\vec{k}$ fluctuates around its theoretical value $\Pr(\vec{k})$ with variance $\Pr(\vec{k})(1 - \Pr(\vec{k}))$. The corresponding experimental error is therefore

$$\sigma_{\text{exp}}(\vec{k}) = \sqrt{\Pr(\vec{k})(1 - \Pr(\vec{k}))/N_{\text{exp}}}. \tag{3.74}$$

If $\Pr(\vec{k})$ is small, the process approaches the Poissonian limit, where the error reduces to $\sigma_{\text{exp}}(\vec{k}) \simeq \sqrt{\Pr(\vec{k})/N_{\text{exp}}}$.

The second noise source arises from the dynamical evolution in the presence of synaptic noise, modeled as a temperature T in the statistical–physics description. Temperature enters explicitly through the acceptance probability for overcoming an energy barrier (see Eq. (3.68)). Reformulated in terms of

the variation of the probability of a single output configuration $\vec{k}$ induced by a change in spin phases, and defining $\Delta\text{Pr}(\vec{k}) = \text{Pr}(\vec{k}|\vec{\boldsymbol{\sigma}}(t)) - \text{Pr}(\vec{k}|\vec{\boldsymbol{\sigma}}(t-1))$, the acceptance probability reads $e^{\frac{N\Delta\text{Pr}(\vec{k})}{T}}$. At fixed, nonzero temperature, the long-time values of $\text{Pr}(\vec{k})$ fluctuate around their temperature-dependent mean. These fluctuations are characterized by

$$\sigma_T^2(\vec{k}) = \left\langle \left(\Delta\text{Pr}(\vec{k})\right)^2 \right\rangle. \tag{3.75}$$

The typical probability thus depends on temperature, i.e., on the noise in the scattering channels or, in the neural network mapping, the synaptic noise. A one-to-one correspondence can be established between the average value $\text{Pr}(\vec{k})$ of a given two-photon configuration and the temperature T, interpreted as the long-time plateau where the energy contribution of the scattering channels leading to $\vec{k}$, $\mathcal{H}[\boldsymbol{\sigma}|\vec{k}] = -M\,\text{Pr}(\vec{k}|\boldsymbol{\sigma})$, relaxes. Consequently, a dependence $\sigma_{\text{exp}}(T)$ can be obtained through Eq. (3.74).

In experimental simulations where the two-photon current is used to estimate $\text{Pr}(\vec{k})$, the user must ensure that

$$\sigma_{\text{exp}}(\vec{k}) \ll \sigma_T(\vec{k}). \tag{3.76}$$

Since N_{exp} directly controls $\sigma_{\text{exp}}(\vec{k})$, the experiment must be designed with an appropriate "Experimental Time per Iteration'' (ETI) so that $\sigma_{\text{exp}}(\vec{k})$ remains smaller than $\sigma_T(\vec{k})$.

3.6 Quantum simulator for classical complex systems

In this section, we develop a more general and platform-independent scheme for the simulation of two-body spin Hamiltonians. The main contributions of this study can be summarized as follows:

- The proposed scheme is platform independent and not restricted to optical implementations. Whereas in previous optical setups the number of simulated neurons scales linearly with the number of modes, our approach enables the simulation of the dynamics of 2^n spins using only n qubits in the case of two-body Hamiltonians.
- Unlike previous models limited to fully connected (all-to-all) architectures, the proposed framework allows for the simulation of Hopfield networks on arbitrary interaction graphs, with tunable coupling strengths and the possibility of including external fields.

- The effective coupling terms in the Hopfield model naturally emerge from the coherence terms of the quantum density matrix—an aspect of particular interest in light of the quantum resource theories discussed in the first chapter of this thesis (see Sec. 1.2). Quantum coherence appears to be the essential ingredient enabling the exponential mapping from a system of n interacting qubits to a system of 2^n interacting spins.
- We compare the quantum simulator with classical counterparts and identify the regimes in which a genuine quantum computational advantage can be achieved.

Regarding the last point, quantum computing offers a promising route toward simulation advantages, primarily due to the exponential scaling of the Hilbert space with the number of subsystems. For qubits, for example, the Hilbert space dimension grows as 2^n, where n is the number of qubits. This exponential growth enables quantum systems to represent and manipulate information in ways that are infeasible for classical computers.

However, a fundamental challenge arises from quantum measurement. Any measurement induces the collapse of the quantum state onto one of its eigenstates. Since the number of such eigenstates also scales exponentially with the system size, obtaining reliable statistical information typically requires a very large number of copies of the quantum state. This requirement can offset the potential computational advantage of quantum systems (Nielsen and Chuang 2002).

In contrast, the framework proposed in this work overcomes this obstacle by allowing one to perform measurements on only a subset of the qubits composing the full system. This partial-measurement strategy not only preserves the computational potential associated with the exponentially large Hilbert space, but also carries conceptual and theoretical significance, as it represents a novel approach to quantum computation.

It is well known that quantum coherence is fragile and highly susceptible to noise, posing a major challenge for quantum technologies. However, as we will show, our scheme does not require long coherence times, since the simulation of Hopfield dynamics does not rely on maintaining quantum coherence over extended durations. This feature represents a potential advantage for experimental realizations.

3.6.1 Mapping

We denote a generic state of n qubits by $|i\rangle$, with $i = 0, \ldots, 2^n - 1$, where i is expressed in binary representation.

The general density operator for a n-qubit state can then be written as

$$\rho^{\text{in}} = \sum_{i,j=0}^{2^n-1} \rho_{ij}^{\text{in}} \left|i\right\rangle \left\langle j\right| , \tag{3.77}$$

which is our initial state.

Adding Controlled Phase Shift

The phase shift can be performed using a controlled-phase quantum gate, which at each step of the dynamics adds a -1 phase factor to only one of the states in the superposition, as in the case of photons. Such a gate, CZ_r, has a diagonal matrix form with all elements equal to 1 on the diagonal, except for the element C_{rr}, which is -1 when the phase of the state $|r\rangle$ is shifted. After applying the CZ_r gate, the density operator elements acquire a factor of -1 corresponding to the row and column r, except for the element ρ_{rr}.

By shifting the phase by π only for the state $|r\rangle$ with a controlled-phase gate CZ_r:

$$\text{CZ}_r \left|i\right\rangle = \left|i\right\rangle \quad \text{if } i \neq r, \qquad \text{CZ}_r \left|r\right\rangle = -\left|r\right\rangle \quad \text{else.} \tag{3.78}$$

In terms of the density operator, this transformation is given by:

$$\begin{aligned} &\tilde{\rho}_{i,j} = \rho_{i,j} \quad \text{if } i,j \neq r \quad \text{or if } i = j = r \\ &\tilde{\rho}_{r,i} = \tilde{\rho}_{i,r}^{*} = -\rho_{r,i} \quad \text{if } i \neq r \end{aligned} \tag{3.79}$$

Thus, the phase shift results in a negative sign only for the rth row and column of ρ, excluding the element ρ_{rr}, which acquires two negative signs that cancel each other out.

If this shift is applied t times, with the shifted states $|r\rangle$ chosen randomly at each step, the resulting density operator is

$$\tilde{\rho} \;=\; \sum_{i,j=0}^{2^n-1} \sigma_i(t)\, \sigma_j(t)\, \rho_{ij}^{\text{in}} \;\left|i\right\rangle \left\langle j\right| , \tag{3.80}$$

where $\sigma_i(t) = +1$ if the phase in mode i has been shifted an even number of times, and $\sigma_i(t) = -1$ if it has been shifted an odd number of times (equivalently, $\sigma_i(t) = (-1)^{m_i(t)}$ where $m_i(t)$ is the number of shifts applied to mode i up to time t).

Unitary transformation

After that, if we apply a unitary transformation U, the resulting density operator will be:

$$\rho^{out} = \sum_{i,j,q,q'=0}^{2^n-1} U^*_{iq'} U_{jq} \sigma_i(t)\sigma_j(t)\rho^{\text{in}}_{ij} \left|q'\right\rangle \left\langle q\right| = \sum_{q,q'=0}^{2^n-1} \rho^{out}_{q'q} \left|q'\right\rangle \left\langle q\right| \tag{3.81}$$

where $\rho^{out}_{q'q} \equiv \sum_{i,j} U^*_{iq'} U_{jq} \sigma_i(t)\sigma_j(t)$.

As we will show in the next section, this transformation can be suitably chosen to tune the coupling terms of the spin Hamiltonians. When the goal is to simulate a spin-glass Hamiltonian, the unitary transformation can be selected at random, exactly as in the case of the optical simulator (see Sec. 3.5).

Outcome probabilities and mapping

At the end, the probability that the system collapses into a generic target state $|k\rangle$ is given by

$$p_k = \rho^{\text{out}}_{kk} = \sum_{i \geq j} J^{(k)}_{ij} \, \sigma_i(t)\sigma_j(t) \equiv -\mathcal{H}_k, \tag{3.82}$$

where

$$J^{(k)}_{ij} = 2\,\text{Re}\Big\{ U_{ik} U^*_{jk} \, \rho^{\text{in}}_{ij} \Big\} . \tag{3.83}$$

If the number of qubits is large, this probability becomes quite small. To obtain a larger probability, one can instead consider the sum of the probabilities associated with a set of output configurations $|k\rangle \in \Lambda$, where $\Lambda \subseteq \{|i\rangle\}$ denotes a subset of all possible n-qubit configurations.

In this case, the probability becomes

$$p_\Lambda = \sum_{i \geq j=0}^{2^n-1} J^{\Lambda}_{ij} \, \sigma_i(t)\sigma_j(t), \tag{3.84}$$

where

$$J^{\Lambda}_{ij} \equiv \sum_{k \in \Lambda} J^{(k)}_{ij} = 2 \sum_{k \in \Lambda} \text{Re}\{ U_{ik} U^*_{jk} \rho^{\text{in}}_{ij} \} . \tag{3.85}$$

Adding an external magnetic field

It is also possible to introduce an external magnetic field term in the spin-glass Hamiltonian, which takes the form

$$\mathcal{H}_{ext} = \sum_{i=0}^{2^n-1} H_i \sigma_i(t) \tag{3.86}$$

This can be done just by keeping constant the phases of one of the qubits, e.g. the first one.

After the controlled phase, the state will be

$$\begin{aligned}\tilde{\rho} = \rho_{00}^{in} |0\rangle \langle 0| + \sum_{i,j=1}^{2^n-1} \sigma_i(t)\sigma_j(t)\rho_{ij}^{\text{in}} |i\rangle \langle j| \\ + \sum_{i=1}^{2^n-1} \rho_{0i}^{in} |0\rangle \langle i| \sigma_i(t) + \sum_{i=1}^{2^n-1} |i\rangle \langle 0| \rho_{i0}^{in} \sigma_i(t)\end{aligned} \tag{3.87}$$

After the random unitary, it is

$$\begin{aligned}\rho^{out} = \rho_{00}^{in} \sum_{q,q'=0}^{2^n-1} U_{0q} U_{0q'}^* |q'\rangle \langle q| + \sum_{\substack{i,j=1 \\ q,q'=0}}^{2^n-1} \sigma_i(t)\sigma_j(t) U_{jq} U_{iq'}^* \rho_{ij}^{\text{in}} |q'\rangle \langle q| \\ + \sum_{\substack{i=1 \\ q,q'=0}}^{2^n-1} U_{0q'}^* U_{iq} \rho_{0i}^{in} |q'\rangle \langle q| \sigma_i(t) + \sum_{\substack{i=1 \\ q,q'=0}}^{2^n-1} U_{0q} U_{iq'}^* |q'\rangle \langle q| \rho_{i0}^{in} \sigma_i(t)\end{aligned} \tag{3.88}$$

Hence the outcome probability in the state $|k\rangle$ when all the qubits are measured, is

$$\begin{aligned}p_k = \rho_{kk} = \sum_{i \geq j=1}^{2^n-1} J_{ij}^{(k)} \sigma_i(t)\sigma_j(t) + \sum_{i=1}^{2^n-1} H_i^{(k)} \sigma_i(t) + \rho_{00} |U_{0k}|^2 \\ \equiv -\mathcal{H}_{int} + \mathcal{H}_{ext} + \mathcal{H}_0\end{aligned} \tag{3.89}$$

with

$$J_{ij}^{(k)} = 2\text{Re}\big[U_{jk} U_{ik}^* \rho_{ij}^{\text{in}}\big], \quad i,j = 1, \ldots, 2^{n-1} \tag{3.90}$$

$$H_i^{(k)} = 2\text{Re}\big[U_{0k} U_{ik}^* \rho_{0i}^{in}\big], \quad i = 1, \ldots, 2^{n-1} \tag{3.91}$$

$$\mathcal{H}_0 = \rho_{00}^{in} |U_{0k}|^2 \tag{3.92}$$

where the last term is a constant.

3.6.2 Examples and applications

First example: 2-body Hopfield with all-to-all couplings and a first simulation

As an example, let us consider the case in which we want to simulate the dynamic of an all-to-all interacting 2-body Hopfield model.

In this case, all the elements of input density operator must be different from 0 in the canonical basis. Such a state can be obtained by applying a Quantum Fourier Transform (QFT) to a system of n qubits in the state $|0\rangle$. After the QFT the state is a superposition of all possible states $|i\rangle$, with $i = 0, \ldots, 2^n - 1$ with equal weights $\frac{1}{\sqrt{2^n}}$, that is

$$|\Psi\rangle^{in} = \frac{1}{\sqrt{2^n}} \sum_{i=0}^{2^n-1} |i\rangle \tag{3.93}$$

After t controlled phase shift the state will be

$$|\Psi\rangle_\phi = \frac{1}{\sqrt{2^n}} \sum_{i=0}^{2^n-1} \sigma_i(t) |i\rangle \tag{3.94}$$

with corresponding density matrix

$$\rho_\phi = |\Psi\rangle_\phi \langle\Psi| = \frac{1}{2^n} \sum_{i,j=0}^{2^n-1} \sigma_i(t)\sigma_j(t) |i\rangle \langle j| \tag{3.95}$$

The unitary matrix at the end mix all the terms, giving the output state

$$\rho^{out} = \frac{1}{2^n} \sum_{i,j,l,m=0}^{2^n-1} U^*_{l,i} U_{m,j} \sigma_i(t)\sigma_j(t) |l\rangle \langle m| \tag{3.96}$$

Measuring all the qubits, the outcome probability in the state $|k\rangle$ will be

$$p_k = \rho^{out}_{kk} = \frac{1}{2^n} \sum_{i,j=0}^{2^n-1} U^*_{k,i} U_{k,j} \sigma_i(t)\sigma_j(t) \tag{3.97}$$

which is can be written as a Hopfield-like Hamiltonian

$$\mathcal{H}_k = -p_k = -\sum_{i \geq j=0}^{2^n-1} J_{ij}^k \sigma_i(t)\sigma_j(t) \tag{3.98}$$

with

$$J_{ij}^k = \frac{2}{2^n}\text{Re}(U_{k,i}^* U_{k,j}) \tag{3.99}$$

If at the end we measure the probability of different output states $|k\rangle \in \Lambda$, it is

$$p_\Lambda = -\mathcal{H}^\Lambda = \sum_{i \geq j=0}^{2^n-1} J_{ij}^\Lambda \sigma_i(t)\sigma_j(t) \tag{3.100}$$

where the coupling constants become

$$J_{ij}^\Lambda = \frac{2}{2^n}\sum_{k\in\Lambda} \text{Re}\{U_{i,k}^* U_{j,k}\} \tag{3.101}$$

In the generic case of n qubits, the complete circuit is shown in Fig. 3.7.

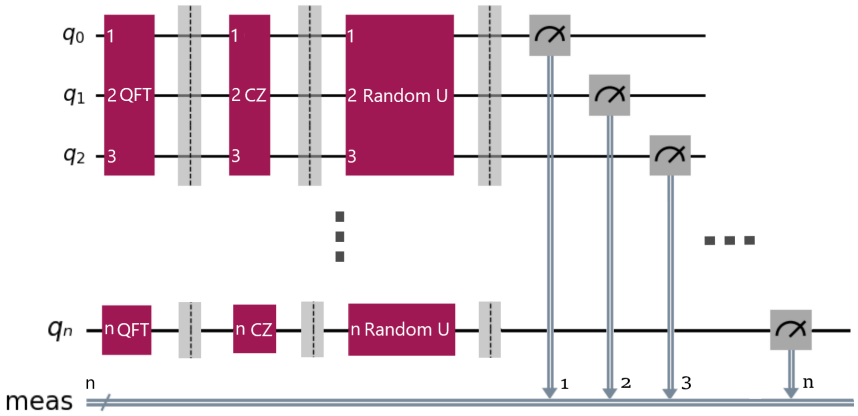


Figure 3.7: *Generic scheme of a fully connected network for n qubits (2^n spins). First, the Quantum Fourier Transform (QFT) is applied, enabling full connectivity through appropriate combinations of Hadamard and controlled-shift gates. This is followed by the application of the controlled-Z (CZ) gate. Finally, a random unitary operation is performed, and measurements are made on any qubit.*

Other examples

1D Ising on a chain

The density operator of a 1D Ising chain for $N = 2^n$ spins with the same diagonal elements $\rho_{ii} = \frac{1}{N}$ (such that the sum is 1), and with real and equal off-diagonal elements is a symmetric, Toepliz tridiagonal matrix

$$\rho_{1D}^{Ising} = \begin{bmatrix} \frac{1}{N} & b & 0 & \cdots & \cdots & 0 \\ b & \frac{1}{N} & b & 0 & \cdots & 0 \\ 0 & b & \frac{1}{N} & b & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b & \frac{1}{N} & b \\ 0 & \cdots & \cdots & 0 & b & \frac{1}{N} \end{bmatrix} \tag{3.102}$$

with eigenvectors

$$\lambda_l = \frac{1}{N} - 2b\cos\left(\frac{l\pi}{N+1}\right) \qquad l = 1, \ldots, N \tag{3.103}$$

The condition to be semidefinite-positive is therefore fulfilled for $b \leq \frac{1}{2N}$. If we consider one mode k as target, the Ising spin-glass Hamiltonian (3.82) reads as

$$\mathcal{H}_k = p_k = \frac{1}{2}\sum_{\langle i \geq j \rangle} J_{ij}^k \sigma_i(t)\sigma_j(t) \tag{3.104}$$

where the summation over $\langle i \geq j \rangle$ is restricted over the nearest-neighbor elements and

$$\begin{aligned} J_{ii}^k &= \frac{2}{N}|U_{ik}|^2, \qquad i = 1, \ldots, N \\ J_{i,j}^k &= 2b\mathrm{Re}\big(U_{ik}^* U_{jk}\big), \qquad j = i+1 \end{aligned} \tag{3.105}$$

If we consider a set Λ of target probabilities, from Eq. (3.85), we have

$$\mathcal{H} = p_\Lambda = \sum_{\langle i \geq j \rangle} J_{ij}^\Lambda \sigma_i(t)\sigma_j(t) \tag{3.106}$$

with

$$J_{ii}^{\Lambda} = \frac{2}{N}\sum_{k\in\Lambda}|U_{i,k}|^2 \qquad i = 1,\ldots,N$$
$$J_{ij}^{\Lambda} = 2b\text{Re}\left\{\sum_{k\in\Lambda}U_{i,k}^{*}U_{j,k}\right\} \qquad j = i+1 \tag{3.107}$$

For $b = \frac{1}{N}$, it is

$$\text{Tr}[(\rho_{1D}^{Ising})^2] = \frac{3N-1}{2N^2} \tag{3.108}$$

To simulate the 1D Ising model with nearest-neighbor interactions only, you just need to choose the initial state that has off-diagonal terms only to the left and right of the diagonal terms:

Ising 1D on a chain with arbitrary coupling strenghts

The initial state should take the form

$$\rho^{\text{in}} = \begin{bmatrix} \rho_{00} & \rho_{01} & 0 & \ldots & 0 & 0 \\ \rho_{01}^{*} & \rho_{11} & \rho_{12} & \ldots & 0 & 0 \\ 0 & \rho_{12}^{*} & \rho_{22} & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \rho_{N-2,N-2} & \rho_{N-2,N-1} \\ 0 & 0 & 0 & \ldots & \rho_{N-2,N-1}^{*} & \rho_{N-1,N-1} \end{bmatrix} \tag{3.109}$$

Ising on a ring

We can add the terms $\rho_{0,N-1}$ e $\rho_{N-1,0} = \rho_{0,N-1}^{*}$ if we want to close the chain:

$$\rho^{\text{in}} = \begin{bmatrix} \rho_{00} & \rho_{01} & 0 & \ldots & 0 & \rho_{0,N-1} \\ \rho_{01}^{*} & \rho_{11} & \rho_{12} & \ldots & 0 & 0 \\ 0 & \rho_{12}^{*} & \rho_{22} & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \rho_{N-2,N-2} & \rho_{N-2,N-1} \\ \rho_{0,N-1}^{*} & 0 & 0 & \ldots & \rho_{N-2,N-1}^{*} & \rho_{N-1,N-1} \end{bmatrix} \tag{3.110}$$

General case

In general, to simulate any lattice symmetric in i and j, you just need to multiply the elements of the lattice adjacency matrix by the elements of the input density operator ρ^{in}, provided that ρ^{in} has all diagonal elements equal to 1 to preserve the trace of the density matrix.

Therefore, given a symmetric adjacency matrix A with $A_{ii} = 1$ for any i, considering an input state of the form

$$\rho_{in} = \sum_{i,j=0}^{2^n-1} A_{ij}\rho_{ij} \left|i\right\rangle \left\langle j\right| \tag{3.111}$$

it is possible to map the error probabilities in the set Λ in the spin-glass Hamiltonian

$$\mathcal{H} = p_\Lambda = \sum_{i\geq j} J_{ij}^{\Lambda}\sigma_i(t)\sigma_j(t) \tag{3.112}$$

on the lattice given by the adjacency matrix A, providing that the input state (3.111) is Hermitian and semi-definite positive. The Hermitianity condition can be simply obtained by choosing a symmetric adjacency matrix $A_{ij} = A_{ji}$.

3.6.3 Pattern retrieval

As shown in Eqs. (3.82) and (3.85), the Hopfield-like coupling terms J_{ij} are embedded in the input density operator matrix elements ρ_{ij} and in the unitary transformation U_{ij}.

To illustrate this with a practical example, let us consider the case of two stored patterns, represented by the letters F and G, as shown in Figures 3.8, with 4 qubits (16 neurons) and 6 qubits (64 neurons) respectively.

In the standard Hopfield network model, the synaptic coupling matrix J_{ij} for a set of P stored patterns $\{\xi_i^\mu\}$, with $\mu = 1, \ldots, P$, is defined as:

$$J_{ij} = \frac{1}{N}\sum_{\mu=1}^{P} \xi_i^\mu\, \xi_j^\mu, \quad J_{ii} = 0. \tag{3.113}$$

Here, N denotes the total number of neurons, and each $\xi_i^\mu \in \{0, 1\}$ represents the state of neuron i in pattern μ. In this case, $\mu = F, G$ correspond to the two stored letters, where black pixels are associated with 1 and white pixels with 0.

Note that the diagonal terms J_{ii} contribute only a constant zero-point energy to the Hamiltonian. Therefore, to embed J_{ij} into a density operator ρ that is

Hermitian, positive semi-definite, and trace-class, we define:

$$\rho_{ij} = J_{ij}, \quad i \neq j, \quad \rho_{ii} = \frac{1}{2^n}, \tag{3.114}$$

which satisfies the above constraints.

As the unitary transformation, we choose the generalized n-qubit Hadamard gate. The transition probability from the initial state $|0\rangle^{\otimes n}$ is then given by:

$$p_0 = \sum_{i,j} J_{ij}\, \sigma_i \sigma_j, \tag{3.115}$$

where J_{ij} is defined according to Eq. (3.114).

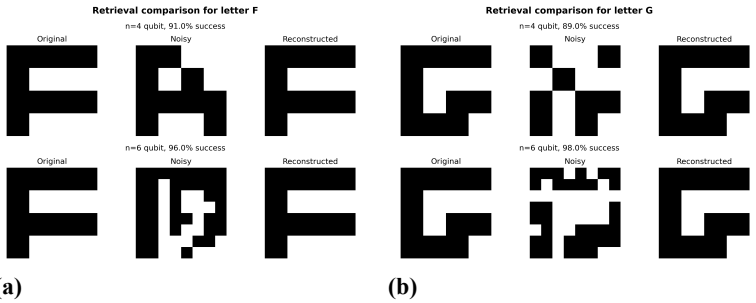


Figure 3.8: *Retrieval of two stored patterns, F and G, with a noise level of* $\eta = 0.2$*, 20 Monte Carlo steps and* 10^3 *samples. Increasing from 4 qubits (16 neurons) to 6 qubits (64 neurons) improves the retrieval efficiency with only two additional qubit.*

3.6.4 Order parameters and simulations

In order to explore the different phases, following (Zanfardino et al. 2026a), we define two order parameters. We perform the Metropolis simulation for n_{rep} replicas, whose equilibrium configurations consist of $N = 2^n$ neurons, denoted by $\boldsymbol{\sigma}^{(\alpha)}$ with $\alpha = 1, \ldots, n_{\text{rep}}$. We consider the (normalized) scalar product between the equilibrium configurations of different replica pairs:

$$q(\alpha, \beta) = \frac{1}{N}\, \boldsymbol{\sigma}^{(\alpha)} \cdot \boldsymbol{\sigma}^{(\beta)}. \tag{3.116}$$

The overlap distribution order parameter is then defined as the probability distribution $p(q)$.

Furthermore, we define the magnetization as

$$M^{(\alpha,\beta)} \equiv \boldsymbol{\sigma}^{(\alpha)} J \boldsymbol{\sigma}^{(\beta)} = \sum_{i,j} \sigma_i^{(\alpha)} J_{ij} \sigma_j^{(\beta)}, \tag{3.117}$$

where J denotes the $N \times N$ coupling tensor, with the associated magnetization distribution $p(M)$.

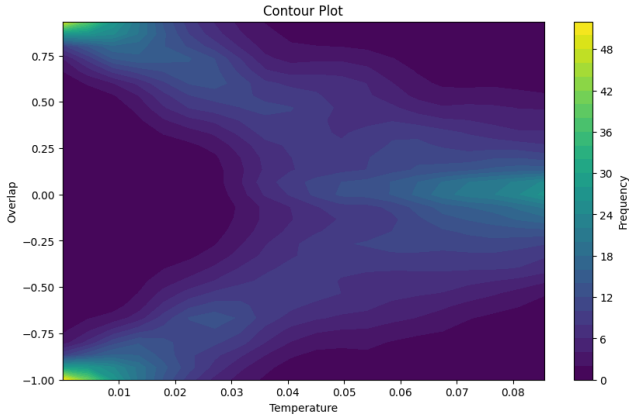


Figure 3.9: *Overlap between $n_{rep} = 50$ replicas of the spin-glass system as function of the temperature for $\alpha = \frac{1}{16}$ obtained by Metropolis simulation in the case of $n = 4$ qubits.*

3.6.5 Metropolis update time complexity: classical vs quantum

Any application of a CZ_r gate results in a new probability distribution $p_k(t+1)$, corresponding to a new Hamiltonian $\mathcal{H}_k(t+1)$, where the spin/neuron r is flipped relative to $\mathcal{H}(t)$. Denoting the energy variation as $\Delta\mathcal{H} \equiv \mathcal{H}(t+1) - \mathcal{H}(t)$, the Metropolis update rule is given by Eq. (3.68) In the current section, we demonstrate that our scheme provides an exponential speedup in performing a single Metropolis update step compared to its classical counterpart.

Classical Scheme

If we consider the fully connected Hopfield model with N neurons, in a classical setup the update of a single neuron, without parallelizing the computation, requires a computational cost of $\sim \mathcal{O}(N)$. Indeed, if at time t the Hamiltonian

is given by

$$\mathcal{H} = -\sum_{i,j=1}^{N} J_{ij}\sigma_i(t)\sigma_j(t), \tag{3.118}$$

and we update neuron r as

$$\sigma_r(t) \to \sigma_r(t+1) = -\sigma_r(t), \tag{3.119}$$

the change in the Hamiltonian is

$$\Delta\mathcal{H} = \mathcal{H}(t+1) - \mathcal{H}(t) = 2\sum_{j=1}^{N} J_{jr}\sigma_r(t)\sigma_j(t), \tag{3.120}$$

and hence the number of operations scales as $\sim \mathcal{O}(N)$. One way to improve this on a classical computer is to parallelize the computation as follows: if we define

$$\Sigma^{(1)}_{r,j} \equiv \sigma_r(t)\sigma_j(t), \qquad T^{(1)}_{r,j} \equiv J_{jr}\Sigma^{(1)}_{r,j}, \tag{3.121}$$

then

$$\Delta\mathcal{H} = 2\sum_{j=1}^{N} T^{(1)}_{r,j}. \tag{3.122}$$

We then define

$$T^{(2)}_{r,j} \equiv T^{(1)}_{r,j} + T^{(1)}_{r,j+1}, \qquad T^{(3)}_{r,j} \equiv T^{(2)}_{r,j} + T^{(2)}_{r,j+1}, \qquad \dots \tag{3.123}$$

Therefore, in the first step, one can compute all $T^{(1)}_{r,j}$ terms in parallel. The subsequent summation can also be performed in parallel in pairs, i.e., computing $T^{(2)}_{r,j}$ in parallel, and so on, until the total sum is obtained. This scheme requires $\sim \mathcal{O}(\log_2(N))$ computational steps to compute the desired sum (3.120).

Clearly, this scale is effective when the number of available parallel working cores, N_{cores}, is equal to or greater than the number of neurons N. If the number of neurons exceeds the number of cores, the actual scaling of the algorithm becomes $\sim \mathcal{O}\left(\frac{N}{N_{\text{cores}}}\log_2(N)\right)$, and thus, for large systems with $N \gg N_{\text{cores}}$, the logarithmic scaling advantage is lost.

3.6.6 Quantum scheme

In the quantum setup, the controlled-phase shifter of an n-qubit state $|r\rangle$, given by Eq. (3.78), is a unitary transformation with a $2^n \times 2^n$ matrix representation:

$$\mathrm{CZ}_r = \mathrm{diag}(1, 1, \ldots, 1, -1, 1, \ldots, 1, 1) \tag{3.124}$$

in the basis $\{|i\rangle\}$, ordered in increasing order, where the -1 corresponds to the k-th element.

It has been shown (Barenco et al. 1995) that for a given single-qubit unitary transformation

$$U = \begin{pmatrix} u_{00} & u_{10} \\ u_{10} & u_{11} \end{pmatrix} \tag{3.125}$$

the corresponding unitary transformation acting on n qubits is given by

$$U_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & u_{00} & u_{01} \\ 0 & 0 & 0 & \cdots & u_{10} & u_{11} \end{pmatrix} \tag{3.126}$$

acting on n qubits, can be simulated by $\sim \mathcal{O}(n^2)$ basic quantum operations. Furthermore, if $U = \sigma_z$, the scaling is $\sim \mathcal{O}(48n)$ basic quantum operations. Here, a basic operation refers to a 2-qubit gate.

Therefore, since the number of neurons that can be simulated with n qubits is $N = 2^n$, we obtain that the neuron update can be simulated logarithmically in the number of neurons N. In this case, the logarithmic scale of simulation is not limited by the number N_{cores} of parallel working cores, but is intrinsic to the quantum nature of the system and hence

$$\frac{Time_{cl}}{Time_{quant}} \sim \mathcal{O}\left(\frac{N \log_2 N}{48 N_{cores} \log_2 N}\right) = \mathcal{O}\left(\frac{N}{48 N_{cores}}\right) \tag{3.127}$$

Note that, although the quantum scheme offers an exponential advantage over its classical counterpart for the single update task, the total simulation time remains exponential in the number of neurons. This is due to the fact that the number of Monte Carlo steps required for the system to reach thermalization grows exponentially. Nevertheless, the method could still prove practically

useful for simulating large neuronal systems.

3.6.7 Quantum advantage

In the previous sections, we showed that, in the quantum scheme, the update of the neurons scales logarithmically, in contrast to the linear scaling observed in the classical scheme. However, in order to evaluate the energy corresponding to a fixed neuron configuration, one needs to measure the probability of the n qubits being in that configuration. In practice, this requires reconstructing many copies of the system and statistically estimating such probability. The challenge lies in the fact that the number of possible qubit configurations grows as 2^n, so the probability of a specific configuration scales as $\sim \frac{1}{2^n}$. Consequently, an exponentially large number of system copies would be required, which is a well-known limitation of many quantum protocols (Nielsen and Chuang 2002).

In our protocol, however, we can circumvent this exponential overhead by measuring only one of the n qubits, say the first one for simplicity. Using Eqs. (3.98) and (3.101), we showed that the resulting outcome probability can still be described by a spin-glass Hamiltonian. Importantly, the probability that the first qubit is found in state $|0\rangle$ or $|1\rangle$ remains on the order of $\sim \frac{1}{2}$, independent of the system size. Hence, with this strategy, we avoid the exponential scaling problem. In this case, the error on $\mathcal{H}_k$ scales as $\sim \frac{1}{2\sqrt{N_{copies}}}$ and is independent on the system size! For example, this can be achieved by considering as unitary transformation in Eq. (3.85) (local) Hadamard acting on each qubit and measuring an odd number $|\Lambda| \simeq \frac{N}{2}$ of output canals, which always gives a non-zero coefficient since the sum of odd numbers ± 1 cannot be null. Therefore, we are able to formulate the following

Theorem *For any n-qubit state ρ, the proposed algorithm can simulate the dynamics of classical spin Hamiltonians for $N = 2^n$ spins:*

$$p_\Lambda = -\mathcal{H}^\Lambda = \sum_{i \geq j=0}^{2^n-1} J_{ij}^\Lambda \sigma_i(t)\sigma_j(t), \tag{3.128}$$

with

$$J_{ij}^\Lambda = 2 \sum_{k \in \Lambda} \mathrm{Re}\{U_{i,k}^* U_{j,k} \rho_{ij}\}. \tag{3.129}$$

Furthermore, if the state ρ can be prepared in polynomial time $\sim$ poly(n) with respect to the number of qubits, whenever $|\Lambda| \simeq \frac{N}{2}$, the update of the spin configuration, according to the rule (3.68)*, can be performed in $\sim \mathcal{O}(\log(N))$ time.*

It is important to note that, in the end, the total simulation time of the full dynamics remains exponential in the number of parties, since the number of Metropolis steps required for relaxation is itself exponential. Nevertheless, our protocol achieves a significant improvement at the level of a single update step: the scaling is reduced from linear to logarithmic. This reduction can have very practical implications regarding the computational resources required.

3.7 Conclusions and future directions

In this chapter, we reviewed the fundamental concepts of associative neural networks, with particular emphasis on the Hopfield model. We then discussed how the dynamics of the p-body Hopfield model can be simulated through multiphoton interference, following the recent work in (Zanfardino et al. 2026a), which leads to a first experimental realization on a silicon chip (Muthali et al. 2026). In this framework, we analyzed the autocorrelation function associated with the resulting dynamics and introduced suitable order parameters to characterize the three main phases of these systems: the retrieval, glassy, and paramagnetic phases.

Subsequently, we generalized these results to the case of 2-body spin Hamiltonians defined on arbitrary lattices, moving beyond the specific optical architecture and into a more general quantum-system setting. We showed that with n qubits it is possible to simulate the dynamics of 2^n spins/neurons. Moreover, we demonstrated the possibility of achieving a quantum advantage in the simulation time of Metropolis-like updates, which are required for dynamical simulation.

Since coherence is the key quantum resource enabling this protocol, a natural extension of our work involves the study of dilute models, known for their enhanced memory-retrieval capacity (Folli et al. 2018; Hwang et al. 2019). Such models can be obtained, for example, by starting from the fully connected case and subsequently applying dephasing maps that reduce the coherence of the quantum state, and consequently the effective number of coupled spins or neurons.

Finally, both for the optical and for the general scheme, we plan to extend this framework towards convolutional neural networks and the study of learning processes, exploring whether similar quantum techniques can offer advantages in more structured neural-network architectures.

Appendix

A.1 Bell states

Bell states are maximally entangled states for two-qubit systems (Cohen-Tannoudji et al. 2019), and they form a basis for the composite Hilbert space $\mathcal{H}^{(2)} \otimes \mathcal{H}^{(2)}$.

Starting with the normalized vector

$$|\Phi_0\rangle = \frac{1}{\sqrt{2}} \left(|00\rangle + |11\rangle\right) \tag{A.1}$$

and performing the transformations

$$|\Phi_i\rangle = \left(\hat{I}_2 \otimes \hat{\sigma}_i\right) |\Phi_0\rangle \tag{A.2}$$

with $i = 1, 2, 3$, where $\hat{\sigma}_i$ are the Pauli matrices, the resulting states, apart from irrelevant global phase factors, are

$$\begin{aligned} |\Phi_1\rangle &= \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \\ |\Phi_2\rangle &= \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \\ |\Phi_3\rangle &= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) \end{aligned} \tag{A.3}$$

Together, the Bell states form a (non-separable) basis of $\mathcal{H}^{(2)} \otimes \mathcal{H}^{(2)}$.

A.2 Maximally coherent state

The statement that any state σ can be obtained by only using incoherent operations on the maximally coherent state (1.3) can be proved as follows.

Define for $n = 1, ..., d$

$$E_n \equiv \sum_{i=1}^{d} c_i \left|i\right\rangle \left\langle m_{i+n-1}\right| \tag{A.4}$$

with $c_i \in \mathbb{C}$, $\sum_{i=1}^{d} |c_i|^2$ and $m_x \equiv mod(x-1, d) + 1$ where $mod(x, y)$ is the rest of the ratio $\frac{x}{y}$. With some algebra can be showed that $m_x = x - \lfloor \frac{x-1}{d} \rfloor d$ where $\lfloor x \rfloor$ is the integer part of x.
Hence if $i + n - 1 \leq d$ then $m_{i+n-1} = i + n - 1$ and if $i + n - 1 > d$ then $m_{i+n-1} = i + n - 1 - d$, and therefore if we fix i (or n) we obtain all the values $1, ..., d$ when n (i) varies. Therefore

$$\begin{aligned}\sum_{n=1}^{d} E_n^\dagger E_n &= \sum_{n=1}^{d} \sum_{i=1,j}^{d} c_i^* c_j \left|m_{i+n-1}\right\rangle \left\langle i|j\right\rangle \left\langle m_{j+n-1}\right| \\ &= \sum_{i=1}^{d} |c_i|^2 \sum_{n=1}^{d} \left|m_{i+n-1}\right\rangle \left\langle m_{i+n-1}\right| = \sum_{n=1}^{d} \left|n\right\rangle \left\langle n\right| = I\end{aligned} \tag{A.5}$$

and hence the $\{E_n\}$ are Kraus operators.
If $\rho \in \mathcal{I}$, then

$$E_n \rho E_n^\dagger = \sum_{i,j,k=1}^{d} s_i c_j c_k^* \delta_{i,m_{j+n-1}} \delta_{i,m_{k+n-1}} \left|j\right\rangle \left\langle k\right| = \sum_{i=1}^{d} s_{m_{i+n-1}} |c_i|^2 \left|i\right\rangle \left\langle i\right| \ , \tag{A.6}$$

which is still in $\mathcal{I}$. Hence $\{E_n\}$ is incoherent.
Now if we take the maximally coherent state $\left|\Psi\right\rangle$ in (1.3) the action of E_n gives

$$E_n \left|\Psi\right\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} c_i \left|i\right\rangle \sum_{j=1}^{d} \delta_{j,m_{i+n-1}} = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} c_i \left|i\right\rangle \tag{A.7}$$

for any n. Hence if we perform a sub-selective measurement we obtain the state

$$\rho_n = \frac{E_n \left|\Psi\right\rangle \left\langle\Psi\right| E_n^\dagger}{p_n} \equiv \left|\Phi\right\rangle \left\langle\Phi\right| \tag{A.8}$$

with $|\Phi\rangle = \sum_{i=1}^{d} c_i |i\rangle$ with probability $p_n = Tr(E_n |\Psi\rangle \langle\Psi| E_n^\dagger) = \frac{1}{d}$ and hence any pure state $|\Phi\rangle$ can be prepared.
If we have a mixture

$$\sigma = \sum_{j=1}^{d} W_j |\Phi_j\rangle \langle\Phi_j| \tag{A.9}$$

with $\Phi_j = \sum_{i=1}^{d} c_i^j |i\rangle$, we define

$$E_n^j \equiv \sqrt{W_j} \sum_{i=1}^{d} c_i^j |i\rangle \langle m_{i+n-1}| \tag{A.10}$$

Also in this case, it can be proven in the same way that $\{E_n^j\}$ is an incoherent operation.
Finally

$$\sum_{n,j} E_n^j |\Psi\rangle \langle\Psi| E_n^{k\dagger} = \sum_j W_j |\Phi_j\rangle \langle\Phi_j| = \sigma \tag{A.11}$$

and hence any state σ can be obtained from the maximally coherent state $|\Psi\rangle$ performing incoherent operations.

A.3 No-cloning theorem

In order to copy a state $|\Psi\rangle$, one would require an experimental apparatus initially prepared in a normalized ancilla state $|\Phi_0\rangle$. This apparatus interacts with the combined state $|\Psi\rangle |\Phi_0\rangle$ such that it is transformed into $|\Psi\rangle |\Psi\rangle$, as illustrated in Figure A.1. Since the total system is closed, the interaction must be described by an *agnostic global unitary transformation*.

$$U(|\Psi\rangle \otimes |\Phi\rangle) = e^{i\alpha} |\Psi\rangle |\Psi\rangle \tag{A.12}$$

where α is a real number, since the state can be specified up to an arbitrary phase factor, and where the term *agnostic* means that the transformation should be independent of the state $|\Psi\rangle$, which is unknown.

To prove the no-cloning theorem, we will show that the unitary transformation U in (A.12) does not exist. Let us therefore consider two different

Figure A.1: *Cloning machine which starts in the state $|\Phi_0\rangle$ and creates a copy of the state $|\Psi\rangle$. The box represents the unitary transformation U which acts on both subsystems.*

normalized states $|\Psi_1\rangle$ and $|\Psi_2\rangle$ that we wish to clone. It is then

$$U(|\Psi_1\rangle \otimes |\Phi_0\rangle) = e^{i\alpha_1} |\Psi_1\rangle \otimes |\Psi_1\rangle \tag{A.13}$$

and similarly for $|\Psi_2\rangle$.

Since U is unitary, it must preserve scalar products, that is

$$\langle\Psi_2|\Psi_1\rangle \langle\Phi_0|\Phi_0\rangle = e^{i(\alpha_1-\alpha_2)} \langle\Psi_2|\Psi_1\rangle^2 , \tag{A.14}$$

which implies

$$|\langle\Psi_2|\Psi_1\rangle| = |\langle\Psi_2|\Psi_1\rangle|^2 \tag{A.15}$$

The last equation is valid if and only if $\langle\Psi_2|\Psi_1\rangle = 0$, i.e., $|\Psi_1\rangle$ and $|\Psi_2\rangle$ are orthogonal or if $\langle\Psi_2|\Psi_1\rangle = 1$, i.e., $|\Psi_1\rangle$ and $|\Psi_2\rangle$ represent the same state, which proves the statement.

A.4 Proof of Equation (1.87)

We will prove Eq. (1.87) by induction. For a fixed $k-$partition, the induction basis for a single-part l is proved tracing out the other $k-1$ parts of ρ in (1.84) which gives

$$\rho^{(I_l)} = \frac{1}{N_l} \left(\mathbb{1}_{N_l} + \sum_{i=1}^{D_l} t_i \sigma_i \right) \tag{A.16}$$

hence, by using $\mathrm{Tr}[\sigma_i \sigma_j] = N_l \delta_{ij}$ we obtain

$$||t^{(I_l)}|| = N_l \mathcal{P}^{(I_l)} - 1 \tag{A.17}$$

for any $l = 1, \ldots, k$.

Now, let us suppose that Eq. (1.87) is valid (for the fixed $k-$partition) for any integer $l < k$, that is

$$||t^{(I_1,\ldots,I_l)}|| = \sum_{g=1}^{l} \sum_{\{I\}_g} (-1)^{l-g} N_{\{I\}_g} \mathcal{P}^{\{I\}_g} + (-1)^l \tag{A.18}$$

for any combination $I_1, \ldots, I_l \in \{I\}_l$. We can invert (1.86)

$$||T|| = 1 + \sum_i ||t^{(I_i)}|| + \sum_{i,j} ||t^{(I_i,I_j)}|| + \cdots + ||t^{(I_1,\ldots,I_k)}|| = 1 + \sum_{g=1}^{k} \sum_{\{I\}_g} ||t^{(\{I\}_g)}|| \tag{A.19}$$

to obtain $||t^{(I_1,\ldots,I_k)}||$. The constant terms (independent by the purities) on the left side of last equation, is

$$\sum_{g=0}^{k-1} (-1)^g \binom{g}{k} = \sum_{g=0}^{k} (-1)^g \binom{g}{k} - (-1)^k = -(-1)^k \tag{A.20}$$

since $\sum_{g=0}^{k} (-1)^g \binom{g}{k} = (1-1)^k = 0$. In the same way, the for a fixed single-party purity $\mathcal{P}^{(I_l)}$, it is

$$N_l \mathcal{P}^{(I_l)} \sum_{g=0}^{k-2} (-1)^g \binom{g}{k-1} = -(-1)^{k-1} N_l \mathcal{P}^{I_l} \tag{A.21}$$

and in general, for the terms

$$N_1 \ldots N_l \mathcal{P}^{(I_1,\ldots,I_l)} \sum_{g=0}^{k-l} (-1)^g \binom{g}{k-l+1} = -(-1)^{k-l} N_1 \ldots N_l \mathcal{P}^{(I_1,\ldots,I_l)} \tag{A.22}$$

Bringing all terms on the left side, we find

$$||t^{(I_1,\ldots,I_k)}|| = \sum_{g=1}^{k} \sum_{\{I\}_g} (-1)^{k-g} N_{\{I\}_g} \mathcal{P}^{\{I\}_g} + (-1)^k \tag{A.23}$$

that is our thesis.

Appendix

B.1 Distinguishability between quantum states

In classical mechanics it is always possible to distinguish between two different states of a system by performing suitable measurements. This is not the case in quantum mechanics, where it is well known that only orthogonal states can be perfectly distinguished.

To formally define distinguishability, we first introduce the concept of quantum measurement.

Def. *Measurement* (Nielsen and Chuang 2002)

A quantum measurement is described by a collection $\{M_m\}$ of operators acting on the Hilbert space of the system, where the index m labels the possible measurement outcomes. These operators must satisfy the following properties:

- If the system is in the state $|\Psi\rangle$ before the measurement, then the probability of obtaining outcome m is

$$p(m) = \langle\Psi|M_m^\dagger M_m|\Psi\rangle\,, \tag{B.1}$$

and the post-measurement state (conditional on outcome m) is

$$\frac{M_m\,|\Psi\rangle}{\sqrt{\langle\Psi|M_m^\dagger M_m|\Psi\rangle}}. \tag{B.2}$$

- The measurement operators satisfy the completeness relation

$$\sum_m M_m^\dagger M_m = \mathbb{1}, \tag{B.3}$$

which ensures that the probabilities sum to one:

$$\sum_m p(m) = \sum_m \langle\Psi|M_m^\dagger M_m|\Psi\rangle = \langle\Psi|\Psi\rangle = 1. \tag{B.4}$$

Using this, we now define the *distinguishability* as:

A set of quantum states $\{|\Psi_i\rangle\}$ is said to be *distinguishable* if there exists a measurement $\{M_m\}$ and an outcome i such that

$$p_i(i) = \langle\Psi_i|M_i^\dagger M_i|\Psi_i\rangle = 1, \tag{B.5}$$

where the subscript i indexes both the state and the corresponding measurement outcome.

From the completeness relation it follows that for $j \neq i$:

$$p_i(j) = \langle\Psi_i|M_j^\dagger M_j|\Psi_i\rangle = 0. \tag{B.6}$$

Indeed,

$$1 = \langle\Psi_i|\Psi_i\rangle = \sum_j \langle\Psi_i|M_j^\dagger M_j|\Psi_i\rangle = \sum_{j\neq i} \langle\Psi_i|M_j^\dagger M_j|\Psi_i\rangle + 1, \tag{B.7}$$

which implies, by the positivity of $M_j^\dagger M_j$, that all terms with $j \neq i$ must vanish. Hence, Equation (B.6) holds.

As already mentioned, perfect distinguishability is only possible when the states are orthogonal. Consider n orthonormal states $|\Psi_1\rangle$, ..., $|\Psi_n\rangle$. The measurement that perfectly distinguishes them is defined by the projectors $M_i \equiv |\Psi_i\rangle\langle\Psi_i|$ for $i = 1, \ldots, n$, along with $M_0 = \mathbb{1} - \sum_i |\Psi_i\rangle\langle\Psi_i|$. For these, we have

$$p_i(i) = \langle\Psi_i|\Psi_i\rangle = 1. \tag{B.8}$$

Now consider two non-orthogonal states $|\Psi_1\rangle$ and $|\Psi_2\rangle$ and suppose they are perfectly distinguishable. That is, there exists a measurement $\{M_m\}$ such that

$$p_1(1) = \langle\Psi_1|M_1^\dagger M_1|\Psi_1\rangle = 1, \quad p_2(2) = \langle\Psi_2|M_2^\dagger M_2|\Psi_2\rangle = 1. \tag{B.9}$$

Then, from Eq. (B.6), it must also hold that

$$p_1(2) = p_2(1) = 0. \tag{B.10}$$

However, since the states are non-orthogonal, $|\Psi_2\rangle$ can be expressed as a

linear combination of $|\Psi_1\rangle$ and a vector $|\Psi_1^{\perp}\rangle$ orthogonal to $|\Psi_1\rangle$:

$$|\Psi_2\rangle = a\,|\Psi_1\rangle + b\,|\Psi_1^{\perp}\rangle\,, \tag{B.11}$$

with $|a|^2 + |b|^2 = 1$. Then,

$$p_2(1) = \langle\Psi_2|M_1^{\dagger}M_1|\Psi_2\rangle = |a|^2 > 0, \tag{B.12}$$

which contradicts the assumption that $p_2(1) = 0$. This shows that non-orthogonal states cannot be perfectly distinguished.

Over the years, various studies have explored quantum distinguishability and its potential applications, including for quantum states (Helstrom 1969; Paris and Rehacek 2004), quantum unitaries (Acín 2001; Sacchi 2005), and completely positive trace-preserving (CPTP) maps (Pirandola et al. 2019). One notable development is the concept of *optimal unambiguous discrimination*, which can be introduced with a simple example (Nielsen and Chuang 2002).

Consider two non-orthogonal quantum states:

$$|\Psi_1\rangle = |0\rangle\,, \qquad |\Psi_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle).$$

As discussed above, non-orthogonality implies that the states cannot be perfectly distinguished. However, consider a measurement $\{M_i\}_{i=1,2,3}$ with the following properties:

$$M_1^{\dagger}M_1 = \frac{\sqrt{2}}{1+\sqrt{2}}\,|1\rangle\,\langle 1|\,, \tag{B.13}$$

$$M_2^{\dagger}M_2 = \frac{\sqrt{2}}{1+\sqrt{2}}\frac{(|0\rangle - |1\rangle)(\langle 0| - \langle 1|)}{2}, \tag{B.14}$$

$$M_3^{\dagger}M_3 = \mathbb{1} - M_1^{\dagger}M_1 - M_2^{\dagger}M_2. \tag{B.15}$$

After performing this measurement, if the outcome is $m = 1$, the state must be $|\Psi_2\rangle$, since

$$p_1(1) = \langle\Psi_1|M_1^{\dagger}M_1|\Psi_1\rangle = 0,$$

and similarly, if $m = 2$, the state is $|\Psi_1\rangle$ because $p_2(2) = 0$. The issue arises when the outcome is $m = 3$, as it provides no information about the initial state. In this way, the ability to unambiguously identify the states (when $m = 1, 2$) is traded off against the possibility of obtaining an inconclusive outcome ($m = 3$), from which nothing can be inferred.

Another approach to studying the distinguishability of non-orthogonal or, more generally, mixed quantum states is through measurements that minimize

the probability of error. The *minimal error probability* is defined as follows:

Def. *Minimal error probability for two states*

Given two quantum states ρ_1 and ρ_2, occurring with prior probabilities p_1 and p_2 respectively, consider the set $\mathcal{M}^2$ of all two-outcome quantum measurements $\{M_m\}_{m=1,2}$. The minimal error probability is given by (Sacchi 2005):

$$p_{err} = \min_{\mathcal{M}} \{p_1 \operatorname{Tr}[\rho_1 M_2] + p_2 \operatorname{Tr}[\rho_2 M_1]\} . \tag{B.16}$$

Since the probability of obtaining outcome i when the system is in state ρ is $\operatorname{Tr}[\rho M_i]$, the quantity in Eq. (B.16) represents the total probability of making an error, i.e., inferring ρ_2 when the system is in ρ_1 and vice versa.

Using the completeness relation $M_1 + M_2 = \mathbb{1}$, the expression to minimize can be rewritten as:

$$\begin{aligned} p_{err} &= p_1 - \operatorname{Tr}[(p_1\rho_1 - p_2\rho_2)M_1] = p_2 + \operatorname{Tr}[(p_1\rho_1 - p_2\rho_2)M_2] \\ &= \frac{1}{2}\left(1 - \operatorname{Tr}[(p_1\rho_1 - p_2\rho_2)(M_1 - M_2)]\right) . \end{aligned} \tag{B.17}$$

Thus, minimizing the error probability in Eq. (B.16) is equivalent to maximizing the following expression:

$$\max_{\mathcal{M}} \{\operatorname{Tr}[(p_1\rho_1 - p_2\rho_2)(M_1 - M_2)]\} . \tag{B.18}$$

Since the operator $p_1\rho_1 - p_2\rho_2$ is Hermitian, it can be decomposed into its positive and negative parts (which correspond to orthogonal projectors). As we will show in the next section, the optimal measurement is obtained by choosing $M_1 = \Pi^+$ and $M_2 = \Pi^-$, where Π^+ and Π^- are the projectors onto the positive and negative eigenspaces of $p_1\rho_1 - p_2\rho_2$, respectively.

Therefore, the minimal error probability is:

$$p_{err} = \frac{1}{2}\left(1 - \|p_1\rho_1 - p_2\rho_2\|_1\right) , \tag{B.19}$$

where $\|\cdot\|_1$ denotes the trace norm of an operator.

B.1.1 Detection theory

Detection theory investigates the reliability of detectors for optical signals, taking into account not only the effects of random noise but also the intrinsic uncertainty arising from the quantum nature of the objects being detected. While this quantum uncertainty is negligible at infrared frequencies, it becomes significant at higher frequencies (Helstrom 1967; Helstrom 1969).

A detector can be modeled as a cavity (or "box") into which an object, whose presence is to be detected, may or may not be placed. This cavity is also subject to an external electromagnetic field that simulates random noise. If, at a given moment, an electromagnetic signal enters the cavity, it interacts both with the ambient electromagnetic field and with the potential object inside. After this interaction, the observer performs a measurement on the electromagnetic field of the returning signal and must decide whether the object is present or not (See for example quantum illumination in section 2.3).

Let ρ_0 and ρ_1 be the density operators describing the return signal in the absence or presence of the object, respectively. Let H^0 and H^1 denote the corresponding hypotheses: H^0 for ρ_0 (object absent) and H^1 for ρ_1 (object present). To distinguish between these hypotheses, the observer performs a quantum measurement described by a Hermitian operator $\hat{X}$, satisfying

$$\hat{X} \left| x_m \right\rangle = x_m \left| x_m \right\rangle , \tag{B.20}$$

where $\{|x_m\rangle\}$ are normalized eigenvectors corresponding to the eigenvalues x_m. It is typically assumed that the observable $\hat{X}$ is complete, meaning that all eigenvalues are distinct and the eigenvectors form an orthonormal basis: $\langle x_m | x_n \rangle = \delta_{mn}$.

The observer is not interested in the specific measurement outcomes x_m, but instead in the probability π_m that hypothesis H^1 is true given that the system collapses into state $|x_m\rangle$. This motivates the definition of a detection operator Π as

$$\Pi \equiv \sum_m \pi_m \left| x_m \right\rangle \left\langle x_m \right| . \tag{B.21}$$

The task is then to determine the optimal operator Π that provides the most accurate decision on the presence or absence of the object.

There exist several criteria for selecting the optimal operator. Helstrom proposed two fundamental approaches:

1. ***Bayes criterion***: This approach seeks to minimize the average cost of decision errors, taking into account prior probabilities and a cost matrix.

2. ***Neyman–Pearson criterion***: This approach aims to maximize the probability of correctly detecting the object (true positive rate), subject to a fixed maximum allowable false alarm probability (i.e., detecting an object when it is not present, corresponding to choosing H^1 when H^0 is true).

The probability that the observer chooses hypothesis H^0 when H^1 is actually

true (i.e., a missed detection) is given by

$$Q_0 = p(H^0 \mid H^1) = \sum_m \pi_m \langle x_m|\rho_0|x_m\rangle = \mathrm{Tr}[\rho_0\Pi], \tag{B.22}$$

while the detection probability, i.e. the probability of correctly choosing H^1 when H^1 is true, is

$$Q_d = p(H^1 \mid H^1) = \mathrm{Tr}[\rho_1\Pi]. \tag{B.23}$$

Let p_0 be the a priori probability that H^0 is true, and let C_{ij} be the cost incurred for choosing hypothesis H^i when H^j is actually true. Then the average cost of decision is

$$\begin{aligned}\overline{C} &= p_0[C_{00}(1-Q_0) + C_{10}Q_0] + (1-p_0)[C_{01}(1-Q_d) + C_{11}Q_d] \\ &= p_0C_{00} + (1-p_0)C_{01} - (1-p_0)(C_{01}-C_{11})(Q_d - \lambda Q_0),\end{aligned} \tag{B.24}$$

where

$$\lambda = \frac{p_0}{1-p_0}\cdot\frac{C_{10}-C_{00}}{C_{01}-C_{11}}. \tag{B.25}$$

Assuming reasonably that $C_{01} > C_{11}$ and $C_{10} > C_{00}$, the Bayes criterion prescribes minimizing the average cost (B.24). This is achieved by maximizing the quantity

$$Q_d - \lambda Q_0 = \mathrm{Tr}[(\rho_1 - \lambda\rho_0)\Pi]. \tag{B.26}$$

Alternatively, under the Neyman–Pearson criterion, one aims to maximize Q_d while fixing Q_0 at a predefined level. Using the method of Lagrange multipliers, the same optimization problem arises, maximizing $Q_d - \lambda Q_0$, with λ serving as the Lagrange multiplier to enforce the constraint on Q_0.

To solve this optimization, consider a basis in which the Hermitian operator $\rho_1 - \lambda\rho_0$ is diagonal, with eigenvectors $\{|\eta_k\rangle\}$ and corresponding eigenvalues $\{\eta_k\}$. Then

$$\rho_1 - \lambda\rho_0 = \sum_k \eta_k \left|\eta_k\right\rangle \left\langle\eta_k\right|, \tag{B.27}$$

and the relevant trace becomes

$$\mathrm{Tr}[(\rho_1 - \lambda\rho_0)\Pi] = \sum_k \eta_k \Pi_{kk}. \tag{B.28}$$

Since $0 \le \Pi_{kk} \le 1$, this expression is maximized by setting $\Pi_{kk} = 1$ for all k

such that $\eta_k \geq 0$, and $\Pi_{kk} = 0$ otherwise.

Moreover, because

$$\sum_k |\Pi_{nk}|^2 \leq 1 \quad \text{and} \quad |\Pi_{mn}| \leq \sqrt{\Pi_{mm}\Pi_{nn}}, \tag{B.29}$$

it follows that the off-diagonal elements vanish unless both Π_{mm} and Π_{nn} are equal to 1. Thus, Π is a diagonal projector onto the subspace spanned by the eigenvectors corresponding to non-negative eigenvalues of $\rho_1 - \lambda\rho_0$:

$$\Pi = \sum_{k:\eta_k \geq 0} |\eta_k\rangle \langle\eta_k| . \tag{B.30}$$

Therefore, the optimal decision strategy is: measure the observable $\rho_1 - \lambda\rho_0$, and decide H^1 if the measurement outcome corresponds to a non-negative eigenvalue; otherwise, decide H^0.

By substituting the optimal operator in (B.17), (B.19) is then obtained.

B.2 Derivation of Jaynes-Cummings solution

If we express the initial state $|\Psi(0)\rangle$, given in (2.109), in terms of the eigenbasis (2.127), the time-evolved state $\exp(-i\hat{H}t)\,|\Psi(0)\rangle$ becomes

$$\begin{aligned} |\Psi(t)\rangle = C_0 C_g\,|0,g\rangle + \sum_{n=0}^{\infty} \Bigg\{ & C_{n+1}C_g \Bigg[\sin\left(\frac{\Phi_n}{2}\right) e^{-i\frac{E_+(n)t}{\hbar}}\,|n,+\rangle \\ & + \cos\left(\frac{\Phi_n}{2}\right) e^{-i\frac{E_-(n)t}{\hbar}}\,|n,-\rangle \Bigg] \\ & + C_n C_e \Bigg[\cos\left(\frac{\Phi_n}{2}\right) e^{-i\frac{E_+(n)t}{\hbar}}\,|n,+\rangle \\ & - \sin\left(\frac{\Phi_n}{2}\right) e^{-i\frac{E_-(n)t}{\hbar}}\,|n,-\rangle \Bigg] \Bigg\} . \end{aligned} \tag{B.31}$$

By substituting the expression of the dressed states in terms of the bared

ones, it is

$$
\begin{aligned}
|\Psi(t)\rangle = & C_0 C_g \,|0,g\rangle \\
& + \sum_{n=0}^{\infty} \Bigg\{ \Bigg[C_{n+1} C_g \left(\cos^2\left(\frac{\Phi_n}{2}\right) e^{-i\frac{E_-(n)t}{\hbar}} + \sin^2\left(\frac{\Phi_n}{2}\right) e^{-i\frac{E_+(n)t}{\hbar}} \right) \\
& \quad + C_n C_e \frac{\sin(\Phi_n)}{2} \left(e^{-i\frac{E_+(n)t}{\hbar}} - e^{-i\frac{E_-(n)t}{\hbar}} \right) \Bigg] |n+1,g\rangle \\
& + \Bigg[C_n C_e \left(\cos^2\left(\frac{\Phi_n}{2}\right) e^{-i\frac{E_-(n)t}{\hbar}} + \sin^2\left(\frac{\Phi_n}{2}\right) e^{-i\frac{E_+(n)t}{\hbar}} \right) \\
& \quad + C_{n+1} C_g \frac{\sin(\Phi_n)}{2} \left(e^{-i\frac{E_+(n)t}{\hbar}} - e^{-i\frac{E_-(n)t}{\hbar}} \right) \Bigg] |n,e\rangle \Bigg\}
\end{aligned}
\tag{B.32}
$$

Since the eigenvalues can be written as

$$
E_\pm = E_1(n) \pm E_2(n) \tag{B.33}
$$

with

$$
E_1(n) = \left(n + \frac{1}{2}\right)\overline{h}\omega, \quad E_2(n) = \overline{h}\Omega_n \tag{B.34}
$$

the wave function can be rewritten in the form

$$
\begin{aligned}
|\Psi(t)\rangle = & C_0 C_g \,|0,g\rangle \\
& + \sum_{n=0}^{\infty} \Bigg\{ \Bigg[C_{n+1} C_g \cos^2\left(\frac{\Phi_n}{2}\right) e^{-i\frac{E_-(n)t}{\hbar}} + C_{n+1} C_g \sin^2\left(\frac{\Phi_n}{2}\right) e^{-i\frac{E_+(n)t}{\hbar}} \\
& \quad - i C_n C_e \sin(\Phi_n) \sin\left(\frac{E_2(n)t}{\overline{h}}\right) e^{-i\frac{E_1(n)t}{\hbar}} \Bigg] |n+1,g\rangle \\
& + \Bigg[C_n C_e \cos^2\left(\frac{\Phi_n}{2}\right) e^{-i\frac{E_-(n)t}{\hbar}} + C_n C_e \sin^2\left(\frac{\Phi_n}{2}\right) e^{-i\frac{E_+(n)t}{\hbar}} \\
& \quad - i C_{n+1} C_g \sin(\Phi_n) \sin\left(\frac{E_2(n)t}{\overline{h}}\right) e^{-i\frac{E_1(n)t}{\hbar}} \Bigg] |n,e\rangle \Bigg\}
\end{aligned}
\tag{B.35}
$$

By defining

$$
\begin{aligned}
K_1(n,t) &\equiv \cos^2\left(\frac{\Phi_n}{2}\right) e^{-i\frac{E_-(n)t}{\hbar}} + \sin^2\left(\frac{\Phi_n}{2}\right) e^{-i\frac{E_+(n)t}{\hbar}}, \quad n = 0,\ldots,\infty \\
K_2(n,t) &\equiv -i\sin(\Phi_n)\sin\left(\frac{E_2(n)t}{\overline{h}}\right) e^{-i\frac{E_1(n)t}{\hbar}}, \quad n = 0,\ldots,\infty
\end{aligned}
\tag{B.36}
$$

The state can be rewritten in a more compact form

$$|\Psi(t)\rangle = \sum_{n=0}^{\infty} \Big\{ [C_n C_g K_1(n-1,t)\,|n\rangle + C_n C_e K_2(n,t)\,|n+1\rangle] \otimes |g\rangle + [C_n C_e K_1(n,t) + C_{n+1} C_g K_2(n,t)]\,|n\rangle \otimes |e\rangle \Big\} \quad \text{(B.37)}$$

where, for convention, we defined $K_1(-1,t) \equiv 1$ in order to include the first term corresponding to $|0,g\rangle$ in the compact expression, which is exactly (2.130). Similar considerations holds in the case of two-mode light states (2.133).

B.3 Dirac equation and Dirac vs Majorana neutrinos

In relativistic quantum mechanics, a free spin-$1/2$ particle of mass m is described by the Dirac equation

$$(i\gamma^\mu \partial_\mu - m)\Psi(x) = 0, \quad \text{(B.38)}$$

which can be written in Hamiltonian form as

$$i\partial_t \Psi(x,t) = \hat{\mathbf{p}} \cdot \boldsymbol{\alpha}\, \Psi(x,t) + m\beta\Psi(x,t), \quad \text{(B.39)}$$

where $\boldsymbol{\alpha} = \gamma^0\boldsymbol{\gamma}$ and $\beta = \gamma^0$.

The Dirac spinor Ψ is a four-component bispinor that can be decomposed into left- and right-handed components,

$$\Psi_L = \frac{1-\gamma^5}{2}\Psi, \qquad \Psi_R = \frac{1+\gamma^5}{2}\Psi, \quad \text{(B.40)}$$

which live in distinct chiral subspaces. In the massless case, the Dirac equation decouples into two independent Weyl equations for Ψ_L and Ψ_R. The mass term couples these components, allowing transitions between opposite chiralities (chiral oscillations).

The distinction between Dirac and Majorana neutrinos originates from the structure of the mass term in the Lagrangian (Pal 2014; Bittencourt et al. 2021).

For a Dirac fermion, the mass term reads

$$\mathcal{L}^{(D)}_{\text{mass}} = -m_D \overline{\Psi}\Psi = -m_D \left(\overline{\Psi}_R \Psi_L + \overline{\Psi}_L \Psi_R\right). \quad \text{(B.41)}$$

This term couples left- and right-handed components and conserves total lepton number. A Dirac neutrino therefore requires independent degrees of freedom for particle and antiparticle.

In contrast, a Majorana mass term couples a field to its charge conjugate.

The charge-conjugated spinor is defined as

$$\Psi^c = C\overline{\Psi}^T, \tag{B.42}$$

where C is the charge conjugation matrix. A purely Majorana mass term can be written as

$$\mathcal{L}^{(M)}_{\text{mass}} = -\frac{m_L}{2}\overline{\Psi^c_L}\,\Psi_L - \frac{m_R}{2}\overline{\Psi^c_R}\,\Psi_R + \text{h.c.} \tag{B.43}$$

Such a term violates lepton number conservation and does not require independent particle and antiparticle degrees of freedom.

A fermion field is called *Majorana* if it satisfies the Majorana condition

$$\Psi = \Psi^c, \tag{B.44}$$

meaning that the particle is identical to its antiparticle.

B.3.1 Majorana bispinors and chirality

In the bispinor formalism, the Majorana condition imposes a constraint between the left- and right-handed components. Writing a generic spinor as

$$\Psi = \begin{pmatrix} \xi_R \\ \xi_L \end{pmatrix}, \tag{B.45}$$

the Majorana condition implies

$$\xi_L = i\sigma_y \xi_R^*, \tag{B.46}$$

so that the upper and lower components are no longer independent.

A key consequence concerns the average chirality,

$$\langle \gamma^5 \rangle = \langle \Psi | \gamma^5 | \Psi \rangle. \tag{B.47}$$

Under charge conjugation one finds

$$\langle \gamma^5 \rangle_c = -\langle \gamma^5 \rangle. \tag{B.48}$$

If the Majorana condition (B.44) holds, then $\Psi = \Psi^c$ and therefore

$$\langle \gamma^5 \rangle = -\langle \gamma^5 \rangle \quad \Rightarrow \quad \langle \gamma^5 \rangle = 0. \tag{B.49}$$

Thus, any Majorana bispinor necessarily has vanishing average chirality at all times, independently of the specific dynamics.

By contrast, a Dirac bispinor does not satisfy this constraint and may exhibit nonzero average chirality, particularly in the ultrarelativistic regime where chirality approximately coincides with helicity.

The structural difference highlighted above underlies the discrimination strategy discussed in Sec. 2.5.

In the ultrarelativistic limit, a Dirac neutrino produced in weak interactions is approximately in a definite chirality state. A Majorana neutrino, instead, must be a superposition of left- and right-handed components due to the constraint (B.44), leading to vanishing average chirality.

Appendix

C.1 Spin-glasses

Disorder and frustration

Spin-glass systems have two fundamental ingredients: disorder and frustration which we will discuss in the next sections.

Disorder

A system exhibits disorder when some of its degrees of freedom (DoFs) are not known exactly, but only through their probability distributions. This scenario is quite common in statistical mechanics. For example, in a system composed of many particles that has reached an equilibrium configuration, the particles continuously change the values of their parameters in phase space. When the number of particles increases significantly, it becomes impossible to know these dynamic values exactly from the outside; instead, we can only describe them using a probability distribution. This kind of dynamical disorder is called annealed, in contrast to quenched disorder. As the term suggests, quenched disorder occurs when the system's degrees of freedom are not known exactly, as in the annealed case, but the parameters are fixed or *quenched* and do not change their values over time.

An example of a system with quenched disorder is dilution. If we take a ferromagnetic model with up or down spins on a lattice, dilution occurs when some sites have defects such as vacancies or non-magnetic materials. In these cases, the number of defects is usually known, but their exact positions are not. Thus, one can only say with a certain probability whether a site contains a spin or a defect. Clearly, in this case, the disorder caused by defects is quenched, in the sense that it does not evolve over time, unlike the spins, which are annealed

DoFs. If the dilution is small enough, it only causes a lowering of the critical temperature of the phase transition. If it is significant enough, it can destroy the ferromagnetic phase entirely.

Quenched disorder is also present in Ising model

$$H = -\sum_{ij} J_{ij} S_i S_j - \sum_i H_i S_i \tag{C.1}$$

when external fields H_i are drawn from a probability distribution $p(H_i)$. In this case, if the probability distribution is symmetric, there is a possibility that when the system is in the ferromagnetic phase, for example with spins up, all the spins may flip to down. Thus, the ferromagnetic phase is preserved but with opposite magnetization. Clearly, this does not happen in the Ising model without quenched disorder.

Spin-glasses, instead, are characterized from the fact that the quenched disorder is present in the coupling constants J_{ij}, which are then extracted from a probability distribution $p(J_{ij})$.

Frustration

Frustration, the second fundamental quantity which characterize spin-glass systems, occurs when the system is not able to find a configuration which minimize the energy. For example, if we consider the ferromagnetic Ising model ($J_{ij} > 0$ on the triangular lattice, it is obvious that the minimum is reached by the system with all the spins pointing in the same direction (up or down). Nevertheless, if we consider the antiferromagnetic (AFM) model ($J_{ij} < 0$) it is clear that once two of three spins of the elementary cell are fixed, the third one "does not know" which configuration has to choose to minimize the energy. It is possible to show that frustration occurs whenever on a closed path Γ it is

$$\sum_{i,j \in \Gamma} J_{ij} < 0 \tag{C.2}$$

The number of such paths that satisfy the previous equation defines the degree of frustration.

If the lattice is squared, frustration does not occur, indeed the minimum configuration can be obtained separating the lattice in two squared sub-lattices, one with all spin up and the other with all spin down (see Fig. C.1).

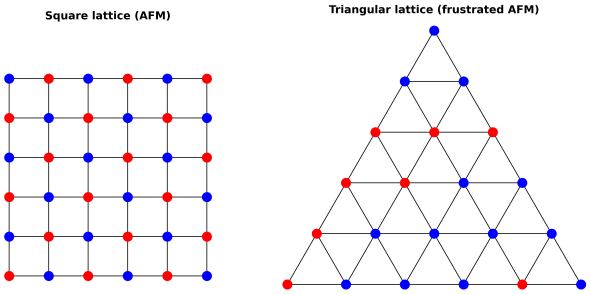


Figure C.1: *The antiferromagnetic model is not frustrated for squared lattice, but it is for triangular lattice.*

C.1.1 Self-averaging

It might seem from the previous discussion that two systems with different realizations of quenched disorder might behave very differently from each other. However, this is not the case for any self-averaging systems, which have general properties that are independent of the specific disorder configuration. In this section we briefly explain what is self-averaging.

When we make a measurement of an observable we obtain a value $\tilde{A}$, which is generally different from the true value, denoted as A. To obtain a more precise measurement, we can perform several measurement and consider the average value $\langle A \rangle$ and we assume that when the number of measurement goes to infinity, supposing an ideal experimental setup without systematic errors, it is $\frac{A-\langle A \rangle}{\langle A \rangle} \to 0$. The concept of self-averaging is quite similar; the only difference is that the system takes this average value on its own. Indeed, suppose we divide the system into N parts, and the observable takes the value A_i in each subsystem, with $i = 1, \ldots, N$. If the observable is an extensive quantity, we can consider the average value

$$\overline{A} = \frac{1}{N} \sum_{i=1}^{N} A_i \tag{C.3}$$

If the subsystems are large enough, we can assume that each of them has its own realization of the disorder, and what we observe is an average over these different realizations. The system is said self-averaging if and only if in the thermodynamic limit it holds that $\frac{A-\overline{A}}{\overline{A}} \to 0$. The concept is the same for intensive quantities, except that we do not need the normalization factor.

C.1.2 Annealed calculation

In statistical mechanics the free energy F is a fundamental quantity from which it is possible to extract a lot of information about the system in exam. To find F it is necessary to find the partition function $\mathcal{Z}$, which is related to the free energy by

$$F = -K_B T \ln \mathcal{Z} \tag{C.4}$$

When quenched disorder is present we have to perform the quenched average on the free energy over all the realization of the disorder. For a fixed realization of the disorder, indicated with J, the relative partition function is $\mathcal{Z}_J$, hence we have to perform the average over J, given by

$$F = \overline{F_J} = -K_B T \overline{\ln \mathcal{Z}_J} \neq -K_B T \ln \overline{\mathcal{Z}_J} \tag{C.5}$$

where the last inequality holds because the logarithm is non-linear and from a mathematical point of view, this fact causes a lot of trouble.

A simple argumentation shows that performing the quenched average over $\mathcal{Z}_{\mathcal{J}}$ and not on its logarithm in general brings to a wrong result: if we define $\beta \equiv \frac{1}{K_B T}$, it is

$$\overline{\mathcal{Z}_J} = \int d\{J\} P(J) \int d\{S_i\} e^{-\beta H_J(\{S_i\})} = \int d\{J\} P(J) e^{-\beta F_J} \tag{C.6}$$

where $\{S_i\}$ indicate all the configuration of the annealed parameters, $H_J(\{S_i\})$ is the Hamiltonian with fixed configuration of the disorder J and $P(J)$ is the probability distribution of the quenched disorder.

We can rewrite

$$\overline{\mathcal{Z}_J} = \int dF \int d\{J\} P(J) e^{-\beta F} \delta(F - F_J) = \int dF P(F) e^{-\beta F} \tag{C.7}$$

where $P(F) \equiv \int d\{J\} P(J) \delta(F - F_J)$. Assuming that the distribution $P(F)$ is a Gaussian with mean value $\bar{F}_J$ and variance ΔF

$$P(F) = \frac{1}{\sqrt{2\pi}\Delta F} e^{-\frac{(\overline{F_J} - F)^2}{\Delta F^2}} \tag{C.8}$$

The integral gives

$$\overline{\mathcal{Z}_J} = e^{-\beta[\overline{F_J} - \beta(\Delta F)^2]} = e^{-\beta F'} \tag{C.9}$$

which gives a free energy

$$F' = \overline{F_J} - \beta(\Delta F)^2 \tag{C.10}$$

different from the correct value $\overline{F_J}$ since the last term is not negligible for low T. For higher T the value is correct, indeed when $\beta \to 0$

$$\overline{\mathcal{Z}_J} = \overline{e^{-\beta F}} \simeq \overline{1 - \beta F_J} = 1 - \beta\overline{F_J} \simeq e^{-\beta\overline{F_J}} \tag{C.11}$$

Here the problem is in the Eq. (C.6), since there is no distinction between the annealed and quenched parameters. Quenched variables are treated exactly as the annealed ones but this is not the case since they are fixed and they do not evolve in the phase space.

C.1.3 The replica trick

To evaluate correctly the quenched average in Eq. (C.5) it is possible to utilize the replica trick, the same adopted by Giorgio Parisi for the resolution of the Sherrington and Kirkpatrick (SK) model (Mézard et al. 1987). The method is based on the limit relation

$$\ln \mathcal{Z}_J = \lim_{n\to 0} \frac{Z_J^n - 1}{n} \tag{C.12}$$

or

$$\ln Z_J = \lim_{n\to 0} \frac{\ln \mathcal{Z}_J^n}{n} \tag{C.13}$$

Hence the average over the replica configurations gives

$$\overline{\ln \mathcal{Z}_J} = \lim_{n\to 0} \frac{\overline{\mathcal{Z}_J^n} - 1}{n} \tag{C.14}$$

or

$$\overline{\ln \mathcal{Z}_J} = \lim_{n\to 0} \frac{\ln \overline{\mathcal{Z}}_J^n}{n} \tag{C.15}$$

The quantity Z_J^n can be viewed as the partition function of n independent replicas of the system, that is

$$\mathcal{Z}_J^n = \sum_{\{S_1\},\ldots,\{S_n\}} e^{-\beta \sum_{\alpha=1}^n H\left(\{S_i^\alpha\}\right)} \tag{C.16}$$

The replica trick is useful since it is easier to calculate the quantity $\overline{Z_J^n}$ instead than the direct calculation of $\overline{\ln Z_J}$. It is immediate to notice that the limit $n \to 0$ has no physical interpretation: which means to replicate a system 0 times? Also from a mathematical point of view, the limit operation in Eq. makes no sense (C.14) since n is an integer. There were several problems in the solution found by Parisi for the SK model, nevertheless, after several years it has been shown that this is the correct solution to the problem (Guerra 2003; Talagrand 2003).

There is also a "more physical" derivation of the replica method formula for performing the quenched average. Suppose that the quenched parameters are also considered annealed with the following two assumptions: they are in contact with a thermal bath at a very high temperature T_J, and their dynamics are slower compared to the real annealed parameters, in the sense that within the time interval in which the quenched parameters change their values, the annealed ones have already reached the equilibrium configuration. The fact that $T_J \to \infty$ implies that the dynamic of the quenched variables is independent by the annealed ones, hence we can assume that the Hamiltonian with a fixed disorder configuration J occurs with a certain probability $P(\{J\})$ and the probability of a given configuration for the annealed parameters (indicated with $\{S_i\}$ is the usual one $e^{-\beta H(\{S_i\},\{J\})}$ and the quantity

$$F(\{J\}) = -K_B T \ln \sum_{\{S_i\}} e^{-\beta H(\{S_i\},\{J\})} \tag{C.17}$$

We can imagine that due to the interaction with spins, the distribution probability is no longer $P(\{J\})$ but the modified version $P(\{J\})e^{-\beta F_J(\{J\})}$. Hence the quantity

$$\tilde{\mathcal{Z}} = \sum_{\{J\}} P(\{J\}) e^{-\beta F_J(\{J\})} \tag{C.18}$$

plays the role of an effective partition function. The free energy of the system is hence

$$\tilde{F} = -\frac{1}{\beta_J} \ln \tilde{Z} \tag{C.19}$$

with $\beta_J = \frac{1}{K_B T_J}$. We can rewrite

$$\tilde{F} = -\frac{1}{\beta} \frac{1}{n} \ln \tilde{Z} \tag{C.20}$$

with $n \equiv \frac{\beta_J}{\beta}$. Hence,

$$\ln \mathcal{Z} = -\beta \tilde{F} = \frac{1}{n} \ln \left[\sum_{\{J\}} P(\{J\}) \left(\sum_{\{S_i\}} e^{-\beta H(\{S_i\},\{J\})} \right) \right] = \frac{1}{n} \ln \overline{Z_J^n} \tag{C.21}$$

in the limit $T_J \to \infty$ ($\beta_J \to 0$), formula (C.14) is obtained.

C.1.4 Spin-glasses phenomenology

Until now we made general consideration about disorder system. In this section we will enter more into the details of spin-glasses. We remark that disorder in spin-glass systems is present in the coupling constants J_{ij} between spins on different sites. What does we observe from an experimental point of view? Spin-glasses have a phase transition from a completely disorder phase to an ordered one as the normal ferromagnetic systems, but with several differences. First of all, in the ordered phase, there is no global magnetization but only a local one. Indeed, it is $\frac{1}{N} \sum_i m_i = 0$, where $m_i \equiv \langle S_i \rangle$ the average $\langle \cdot \rangle$ is the annealed one. Hence, one of the parameter which is useful to study this phase transition is the local susceptibility defined by

$$\chi_{ij} \equiv \frac{\partial \langle S_i \rangle}{\partial H_j} = \frac{\partial m_i}{\partial H_j} \tag{C.22}$$

which is the variation of the averaged spin on the site i modifying the external field H_j on the site j. In fact, only diagonal elements χ_{ii} are different from zero for a symmetric probability distribution of the quenched disorder. For the fluctuation-dissipation theorem, it is

$$\chi_{ii} = \beta(\langle S_i^2 \rangle - \langle S_i \rangle^2) = \beta(1 - m_i^2) \tag{C.23}$$

Therefore, one defines

$$\chi_{loc} = \frac{1}{N} \sum_i \chi_{ii} = \beta \left(1 - \frac{1}{N} \sum_{i=1}^{N} m_i^2 \right) \tag{C.24}$$

Experimentally, there is a critical temperature for which it is observed a cusp of the local susceptibility. This is quite different from the ferromagnets, where the susceptibility diverges at the critical temperature.

Another quantity which gives information on the spin-glass systems is the

spin-glass susceptibility

$$\chi_{SG} = \frac{1}{N}\sum_{i,j}\chi_{ij}^2 = \frac{\beta^2}{N}\sum_{i,j}\overline{(\langle S_i S_j\rangle - \langle S_i\rangle\langle S_j\rangle)^2} \tag{C.25}$$

which is not a real susceptibility, but it is defined by using the fluctuation-dissipation formula. Note that the quenched average is taken only after the square, such that the average over the disorder is not zero.

Another peculiar property of spin-glass systems is the presence of metastable states. In the ferromagnetic case, the free-energy structure is very simple: above the critical temperature T_c one has a parabola and below T_c there is a potential with two double-well potential (see Fig. C.2). When the symmetry is spontaneously broken, the system choose one of the two possible configuration which minimize the energy where all the spins are up or down respectively and the ferromagnetic phase is created. For the spin-glasses the situation is much more complicated due to the presence of the disorder and the frustration. In particular, the free energy spectrum is a complex landscape structure with many valleys (see Fig. C.2). In the thermodynamic limit, some of these valleys become infinitely deeps, while the others becomes deeper but remain finite. Such finite valleys are the metastable states, where the system can remain trapped for a certain amount of time. This fact can cause an increasing of the time needed to reach the equilibrium configuration.

In the thermodynamic limit, the stable (very deep) valleys are called pure states, and each of them is formed by many less deep valleys which are the metastable states. If we take the annealed average quantity of a certain observable A, it is

$$\begin{aligned}\langle A\rangle &= \frac{1}{\mathcal{Z}}\sum_{\{S_i\}}A(\{S_i\})e^{-\beta H(\{S_i\})} = \sum_a \omega_a \sum_{\{S_i\}\in a}\frac{1}{\mathcal{Z}_a}A(\{S_i\})e^{-\beta H(\{S_i\})} \\ &= \sum_a \omega_a \langle A\rangle_a\end{aligned} \tag{C.26}$$

where a indicates the pure states, $\omega_a \equiv \frac{\mathcal{Z}_a}{\mathcal{Z}}$ are the weights of the pure states, and $\langle A\rangle_a$ is the average in the phase a with respective partition function $\mathcal{Z}_a$ given by

$$\mathcal{Z}_a \equiv \sum_{\{S_i\}\in a} e^{-\beta H(\{S_i\})} \tag{C.27}$$

As with ferromagnets, spin glasses can also exhibit broken ergodicity. The situation is straightforward in the case of ferromagnets: the phase space is divided into two equal parts, one corresponding to a magnetization with all

spins up and the other with all spins down. When symmetry is broken, the system falls into one of these two parts of the phase space and there is no possibility of escape in the thermodynamic limit $N \to \infty$. When N is large but finite, there is always a certain probability that the system can occasionally escape from one phase and move to the other. In the case of spin glasses, there are many valleys, so the system falls into one of these valleys, and there is a probability that within a finite time, it escapes from this valley. If this time is too long, ergodicity is clearly broken. If the escape time is finite, ergodicity can be restored. The time it takes to escape from a particular valley defines the degree of broken ergodicity. To understand these properties, it is necessary to study the dynamics of spin glasses. Nevertheless, in the next section, we will define some order parameters that will provide qualitative information without a complete analysis of the dynamics of these systems.

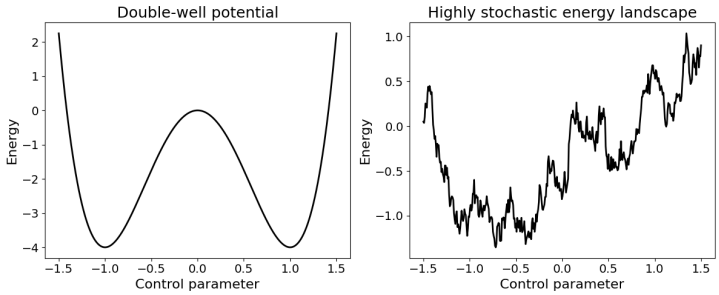


Figure C.2: *In the double-well potential case, there is no disorder, allowing the system to easily reach the global minimum. In contrast, the stochastic case features many local minima, which significantly increase the relaxation time.*

Order parameters for spin-glasses

A quantity that well describes the spin-glasses properties is the order parameter

$$q_{EA} = \lim_{t\to\infty} \lim_{N\to\infty} \overline{\langle S(t)S(t+t_\omega)\rangle} \tag{C.28}$$

where t_ω is a certain reference time called waiting time. If the system is ergodic, the whole phase space is visited and the correlations for long time go to zero, which implies $\overline{\langle S(t)S(t+t_\omega)\rangle} \to \overline{\langle S(t)\rangle \langle S(t+t_\omega)\rangle} = 0$ in the limit of vanishing external field. This is not true when the system is in a pure state and ergodicity is broken. This quantity is nothing else that the averaged squared

magnetization

$$q_{EA} = \frac{1}{N} \sum_i \overline{\sum_a \omega_a (m_i^a)^2} \tag{C.29}$$

which is called Edwards-Anderson order parameter. It is fundamental to take the square before the quenched average such that the quantity does not vanish.

Another interesting order parameter is the statistical mechanics order parameter

$$q_{SM} = \frac{1}{N} \sum_i \overline{\left(\sum_a \omega_a m_i^a \right)} = \frac{1}{N} \sum_i \overline{\sum_{a,b} \omega_a \omega_b m_i^a m_i^b} \tag{C.30}$$

which keeps into account also the contribution intervalley not considered in the Edwards-Anderson version. It holds that $q_{SM} \leq q_{EA}$ and the equality is obtained if and only if the system is in a pure state. Hence, the quantity $\Delta = q_{EA} - q_{SM}$ describes the degree of broken ergodicity.

With using Eq. (C.24) we can rewrite the susceptibility as

$$\chi_{loc}^{BE} = \beta(1 - q_{EA}) \tag{C.31}$$

in the case of broken ergodicity, and

$$\chi_{loc}^{E} = \beta(1 - q_{SM}) \tag{C.32}$$

when the system is ergodic. When the system is not in a pure state, another quantity which takes into account the contribution due to the overlap between different phases is

$$q^{ab} = \frac{1}{N} \sum_i m_i^a m_i^b \tag{C.33}$$

In (Parisi 1983) it was defined its probability distribution, given by

$$P_J(q) \equiv \langle \delta(q - q_J^{ab}) \rangle = \sum_{a,b} \omega_a \omega_b \delta(q - q^{ab}) \tag{C.34}$$

for a fixed configuration of the disorder J, and

$$P(q) = \overline{P_J(q)} = \overline{\sum_{a,b} \omega_a \omega_b \delta(q - q^{ab})} \tag{C.35}$$

for the average over the disorder.

C.1.5 Resolution of the p-spin-glass model

Several spin-glass models have been studied over the years. In this section, we will provide an outline for solving the p-spin-glass model with continuous variables. We will not delve into the detailed calculations; instead, we will offer an overview of how the replica trick is useful for this type of problem. The p-spin model is described by a Hamiltonian of the form

$$\mathcal{H} = -\sum_{i_1,\dots,i_p} J_{i_1\dots i_p}\sigma_{i_1}\dots\sigma_{i_p} \tag{C.36}$$

where there are N spins and the interaction terms are between $p \geq 2$ spins. The spin variables S_i are replaced with σ_i to emphasize that they are now continuous. Conceptually, there are no significant differences between the two models; we use continuous variables for simplicity of calculations. To recover the usual spin-glass properties we add the spherical constraint on the spin variables which reads as

$$\sum_{i=1}^{N}\sigma_i^2 = N \tag{C.37}$$

The goal is to find the configuration of spins that minimizes the Hamiltonian (C.36) under this constraint. The p-spin model has also been studied as a usual ferromagnet without disorder due to the presence of metastable states. Indeed, there are two critical temperatures, T_d and T_c with $T_d > T_c$. In the phase $T_c < T < T_d$, ferromagnetic states develop, but with a free energy bigger than the paramagnetic ones. Below T_c the ferromagnetic states become the global minimum of the Hamiltonian, transitioning from metastable to stable states.

In our case, also disorder is taken into account and we assume that any coupling constant is distributed with a Gaussian distribution of the form

$$P(J) = \sqrt{\frac{N^{p-1}}{\pi p!}}e^{-\frac{N^{p-1}}{p!}J^2} \tag{C.38}$$

Where the normalization factor is obtained by noting that N spins can be p-grouped in $\binom{N}{p} = \frac{N!}{p!(N-p)!}$ that in the thermodynamic limit $N \geq p$ can be approximated with $\frac{N^p}{p!}$. The sum over all the p-uples is $\sim \sqrt{\overline{J^2}}\sqrt{\frac{N^p}{p!}}$ which has to be of the order of $\sqrt{N}$. Hence $\sqrt{\overline{J^2}} \sim \sqrt{\frac{N^{p-1}}{p!}}$.

Annealed calculation

Before solving the p-spin-glass model with quenched disorder, we will perform the calculation treating all variables as annealed to show that the result obtained is wrong. For a fixed disorder configuration it is

$$\mathcal{Z}_J = \int d\{\sigma_i\} \prod_{i_1>\cdots>i_p} e^{\beta J_{i_1\ldots i_p}\sigma_{i_1}\ldots\sigma_{i_p}} \tag{C.39}$$

and averaging over the disorder variables

$$\overline{\mathcal{Z}_J} = \int d\{\sigma_i\} \int \prod_{i_1>\cdots>i_p} d\{J_{i_1,\ldots,i_p}\} \sqrt{\frac{N^{p-1}}{\pi p!}} e^{-\frac{N^{p-1}}{p!}J^2_{i_1\ldots i_p}+\beta J_{i_1\ldots i_p}\sigma_{i_1}\ldots\sigma_{i_p}} \tag{C.40}$$

The last integral can be exactly solved and gives

$$\overline{\mathcal{Z}_J} = \Omega_N e^{N\frac{\beta^2}{4}} \tag{C.41}$$

where Ω_N is the spherical surface with radius 1 in N dimensions. Hence

$$F = -K_B T \ln \overline{\mathcal{Z}_J} = -\frac{\beta}{4} - TS_\infty \tag{C.42}$$

with $S_\infty \equiv \frac{\ln \Omega_N}{N}$. It has been shown, that the free energy obtained with the quenched calculation is lower than one in Eq. (C.42). Another proof that the result is wrong is obtained computing the entropy

$$S = -\frac{\partial F}{\partial T} = -\frac{1}{4K_B T^2} + S_\infty \tag{C.43}$$

that for low T goes to negative values. For the continuous variable system, there is no problem since we have no interpretation of the entropy in terms of the number of configurations in the phase space. Nevertheless, it can be shown that the same result is obtained also considering the discrete model. In this case the entropy cannot be negative, since it is the logarithm of the number of configurations.

While for low T the result is wrong, for higher T the result is correct and it coincides with the one obtained performing the quenched calculation. This is expected for the exact approximation made in the Eq. (C.11).

Replica trick for p-spin-glass

In order to use the replica formulas (C.14) or (C.15) it is necessary to compute

$$\begin{aligned}\overline{\mathcal{Z}_J^n} &= \int d\{\sigma_i^\alpha\} \int \prod_{i_1>\cdots>i_p} d\{J_{i_1\ldots i_p}\} \sqrt{\frac{N^{p-1}}{\pi p!}} e^{-\frac{N^{p-1}}{p!} J_{i_1\ldots i_p}^2 + \beta \sum_{\alpha=1}^n J_{i_1\ldots i_p} \sigma_{i_1}^\alpha \ldots \sigma_{i_p}^\alpha} \\ &= \int d\{\sigma_i^\alpha\} e^{\frac{\beta^2 p!}{4N^{p-1}} \sum_{\alpha,\beta=1}^n \left(\sum_{i=1}^N \sigma_i^\alpha \sigma_i^\beta\right)^p} \end{aligned} \tag{C.44}$$

Note that we started with independent, non-interacting replicas, and the effect of the quenched average is to couple the different replicas. This integral cannot be directly solved since it is not Gaussian.

Automatically appears the term

$$q^{\alpha\beta} = \frac{1}{N} \sum_{i=1}^N \sigma_i^\alpha \sigma_i^\beta \tag{C.45}$$

that is the overlap between the replicas α and β, and the spherical constraint gives $q^{\alpha\alpha} = 1$. If we introduce the following delta function

$$1 = \int dq^{\alpha\beta} \delta \left(N q^{\alpha\beta} - \sum_{i=1}^N \sigma_i^\alpha \sigma_i^\beta \right) \tag{C.46}$$

in the integral above for $\alpha \neq \beta$, we obtain

$$\overline{\mathcal{Z}_J^n} = \int d\{\sigma_i^\alpha\} \prod_{\alpha,\beta} \int dq^{\alpha\beta} \left(N q^{\alpha\beta} - \sum_{i=1}^N \sigma_i^\alpha \sigma_i^\beta \right) e^{\frac{\beta^2 p!}{4N^{p-1}} N^p \sum_{\alpha,\beta=1}^n (q^{\alpha\beta})^p} \tag{C.47}$$

Now, with using the integral representation of the delta function

$$\delta \left(N q^{\alpha\beta} - \sum_{i=1}^N \sigma_i^\alpha \sigma_i^\beta \right) = \int d\lambda^{\alpha\beta} e^{\lambda^{\alpha\beta} \left(N q^{\alpha\beta} - \sum_{i=1}^N \sigma_i^\alpha \sigma_i^\beta \right)} \tag{C.48}$$

the integral in the $\{\sigma_i^\alpha\}$ can be solved giving

$$\overline{\mathcal{Z}_J^n} = \int dq^{\alpha\beta} d\lambda^{\alpha\beta} e^{-N\mathcal{S}(Q,\Lambda)} \tag{C.49}$$

where

$$\mathcal{S}(Q,\Lambda) \equiv -\frac{\beta^2}{4} N \sum_{\alpha,\beta=1}^{n} \left(q^{\alpha\beta}\right)^p - N \sum_{\alpha,\beta=1}^{n} \lambda^{\alpha\beta} q^{\alpha\beta} - \frac{1}{2} \ln \det(2\Lambda) \quad \text{(C.50)}$$

where the formula

$$\int dx_1 \dots dx_n e^{-\frac{1}{2}\sum_{i,j=1}^{n} x_i A_{ij} x_j} = \sqrt{\frac{(2\pi)^n}{\det A}} \quad \text{(C.51)}$$

has been used. The advantage of the integral in Eq. (C.49) is that it can be approximated with the saddle point method which says that in the limit of large N the integral can be approximated with the maximum of the function $\mathcal{S}$ at the exponent.

It has to be noticed that the thermodynamic limit $N \to \infty$ should be taken after the replica limit $n \to 0$. Nevertheless, we are taking before the thermodynamic limit, inverting the order of the two limits and no one ensures that the procedure is correct. Only after years it has been shown that the result is correct.

Hence we have to minimize the function $\mathcal{S}$ respect to its parameters. First of all, we can derive respect to $\lambda^{\alpha\beta}$ with using

$$\frac{\partial}{\partial \lambda^{\alpha\beta}} \ln \det \Lambda = (\Lambda^{-1})^{\alpha\beta} \quad \text{(C.52)}$$

which can be easily obtained writing

$$\det \Gamma = \lambda^{1,1} C^{1,1} + \lambda^{1,2} C^{1,2} + \dots \quad \text{(C.53)}$$

where $C^{1,i}$ is the minor of the element $\lambda^{1,i}$. Hence,

$$\frac{\partial}{\partial \lambda^{\alpha\beta}} \ln \det \Lambda = \frac{C^{\alpha\beta}}{\det \Lambda} = (\Lambda^{-1})^{\beta\alpha} = (\Lambda^{-1})^{\alpha\beta} \quad \text{(C.54)}$$

for the symmetry of Λ. Therefore $\frac{\partial \mathcal{S}}{\partial \lambda^{\alpha\beta}} = 0$ brings to the minimum condition

$$2\Lambda^* = Q^{-1} \quad \text{(C.55)}$$

which leads to

$$\mathcal{S}(\Lambda^*, Q) = -\frac{\beta^2}{4} N \sum_{\alpha,\beta=1}^{n} \left(q^{\alpha\beta}\right)^p - N \sum_{\alpha,\beta=1}^{n} \mathbb{1}^{\alpha\beta} - \frac{1}{2} \ln \det(Q^{-1}) \quad \text{(C.56)}$$

where the second term can be neglected since it gives only a constant contribu-

tion. Hence the free energy reads as

$$F = \lim_{n\to 0} -\frac{1}{\beta n}\overline{\mathcal{Z}^n} = \lim_{n\to 0} -\frac{1}{2\beta n}\left[\frac{\beta^2}{2}\sum_{\alpha,\beta=1}^{n}(q^{\alpha\beta})^p + \ln\det Q\right] \tag{C.57}$$

Minimizing respect to the $q^{\alpha\beta}$ leads to

$$\frac{\partial \mathcal{S}}{\partial q^{\alpha\beta}} = \frac{\beta^2 p}{2}(q^{\alpha\beta})^{p-1} + (Q^{-1})^{\alpha\beta} = 0 \tag{C.58}$$

At this point we will give some ansatz for the form of Q to find a solution for the last equation.

Replica symmetric solution

The simplest ansatz that one can think is that all the overlaps between the replicas are equal to each other. Hence the matrix Q is parameterized with all 1 on the diagonal and the same overlap $q_0 \leq 1$ for different replicas, that is

$$q^{\alpha\beta} = q_0 + (1 - q_0)\delta_{\alpha\beta} \tag{C.59}$$

It can be shown that

$$(Q^{-1})^{\alpha\beta} = \frac{1}{1-q_0}\delta_{\alpha\beta} - \frac{q_0}{(1-q_0)[1+(n-1)q_0]} \tag{C.60}$$

that substituted in the Eq. (C.58) in the limit $n \to 0$ gives

$$h(q_0) \equiv \frac{\beta^2 p}{2}q_0^{p-1} - \frac{q_0}{(1-q_0)^2} = 0 \tag{C.61}$$

which has a solution with $q_0 = 0$ that is the paramagnetic one with a corresponding free energy $F = -\frac{\beta}{4}$ which is the same result of the annealed calculation in Eq. (C.42) except for a constant shift S_∞. Furthermore, there is also a non-paramagnetic solution gives by

$$q_0^{p-2}(1-q_0)^2 = \frac{2}{p}(K_B T)^2 \tag{C.62}$$

From the figure C.3, one can see that the solution is obtained if the temperature T is small enough. There are two solution to the equation, but only the bigger one is physical, since q_0 has to decrease when T increases.

The solution found is nice but it has been shown to be unstable (De Almeida and Thouless 1978; Crisanti and Sommers 1992). The problem is in the ansatz of replica symmetric solution. In the next section we will show that the correct

solution can be found breaking the symmetry of the replicas.

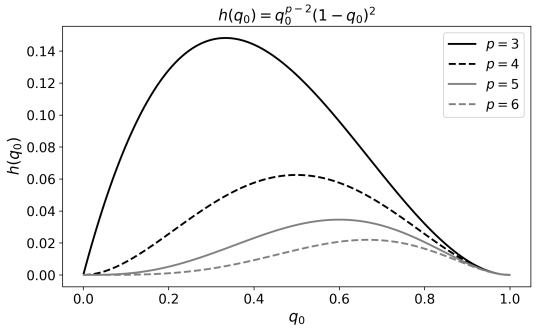


Figure C.3: *Left-hand side of Eq.* (C.61) *for different values of p.*

One-step replica symmetry breaking (solution for $p \geq 3$)

A first step to break the replica symmetry can be done by introducing two parameters $m \leq n$ and $q_1 \geq q_0$ such that the matrix Q has all 1 on the diagonal, it is formed by blocks $m \times m$ along the diagonal with elements q_1 in the off-diagonal elements, and the remaining elements are q_0. For example, if $m = 3$ we have

$$Q = \begin{bmatrix} Q_1 & Q_0 & Q_0 & \cdots \\ Q_0 & Q_1 & Q_0 & \\ Q_0 & Q_0 & Q_1 & \\ \vdots & & & \ddots \end{bmatrix}, \quad Q_1 = \begin{bmatrix} 1 & q_1 & q_1 \\ q_1 & 1 & q_1 \\ q_1 & q_1 & 1 \end{bmatrix}, \quad Q_0 = \begin{bmatrix} q_0 & q_0 & q_0 \\ q_0 & q_0 & q_0 \\ q_0 & q_0 & q_0 \end{bmatrix}. \tag{C.63}$$

The reasons of such a parameterization proposed by Parisi are the following: we can imagine that q_1 describes the overlaps in a fixed valley, which are stronger than q_0 which represents the overlap between different valleys. The strong assumption is that all the interactions intra-valley and inter-valley are equivalents. Since for symmetry it is $\sum_{\alpha,\beta} q^{\alpha\beta} = n \sum_{\alpha} q^{\alpha\beta}$ and there are $n(m-1)$ elements q_1 and $n(n-m)$ elements q_0, we can write the probability distribution in Eq. (C.35) for the $q^{\alpha\beta}$ as

$$P(q) = \frac{m-1}{n-1}\delta(q-q_1) + \frac{n-m}{n-1}\delta(q-q_0) \tag{C.64}$$

Note that in the limit of $n \to 0$, since the probability has to be positive, the order between m and n has to be inverted, such that $1 \leq m \leq n \to 0$, that is another strange thing of the replica trick.

Now, we have to minimize respect to m, q_0 and q_1 according to the saddle point procedure. The first term in Eq. (C.57) gives

$$\lim_{n\to 0}\frac{1}{n}\sum_{\alpha,\beta}(q^{\alpha\beta})^p = \lim_{n\to 0}\sum_{\alpha}(q^{\alpha\beta})^p = \lim_{n\to 0}[1+(m-1)q_1^p+(n-m)q_0^p]$$
$$= 1+(m-1)q_1^p - mq_0^p \tag{C.65}$$

The second term in Eq. (C.57) is more complicated to treat, but it can be shown that the diagonalization of Q gives three eigenvalues λ_i

$$\begin{aligned}
&\lambda_1 = 1-q_1\,, && d_1 = n-\frac{n}{m}\,,\\
&\lambda_2 = m(q_1-q_0)+1-q_1\,, && d_2 = \frac{n}{m}-1\,,\\
&\lambda_3 = nq_0+\lambda_2\,, && d_3 = 1\,,
\end{aligned} \tag{C.66}$$

where the d_i are the respective degeneration. Therefore, the free energy reads as

$$-2\beta F = \frac{\beta^2}{2}[1+(m-1)q_1^p - mq_0^p] + \frac{m-1}{m}\ln(1-q_1)$$
$$+\frac{1}{m}\ln[m(q_1-q_0)+1-q_1] + \frac{q_0}{m(q_1-q_0)+1-q_1} \tag{C.67}$$

Which reduces to the replica symmetric case if $q_1 = q_0$ or $m = 1$:

$$-2\beta F_{sym} = \frac{\beta^2}{2}[1-q_0^p] + \ln(1-q_0) + \frac{q_0}{1-q_0} \tag{C.68}$$

The saddle point equations $\frac{\partial F}{\partial q_0} = 0$, $\frac{\partial F}{\partial q_1} = 0$ and $\frac{\partial F}{\partial m} = 0$ lead respectively to

$$\begin{aligned}
&-\frac{\beta^2}{2}mpq_0^{p-1} + \frac{mq_0}{[m(q_1-q_0)+1-q_1]^2} = 0,\\
&(1-m)\left\{\frac{\beta^2}{2}pq_1^{p-1} - \frac{q_1}{(1-q_1)[(m-1)q_1+1]}\right\},\\
&\frac{\beta^2}{2}q_1^p + \frac{1}{m^2}\ln\left[\frac{1-q_1}{1-(1-m)q_1}\right] + \frac{q_1}{m[1-(1-m)q_1]} = 0
\end{aligned} \tag{C.69}$$

Such equations give a solution with $q_1 = 0$ that is the same of the replica symmetric one, and for $m = 1$ a non-trivial one given by the equation

$$g(q_1) \equiv \frac{\beta^2}{2}q_1^p + \ln(1-q_1) + q_1 = 0 \tag{C.70}$$

As shown in the figure C.4 there is a critical temperature T_s for which a solution $q_1 = q_s \neq 0$ and $m = 1$ is admitted. When $T < T_s$ it is possible to show that m decreases and q_1 increases.

In the case $p \geq 3$, the solution is stable with a free energy lower than the paramagnetic one. The index s for the critical temperature stays for static.

Looking at the probability distribution of the overlaps in Eq. (C.64), we pass from a 0-peaked distribution $P(q) = \delta(q)$ for $T > T_s$ to a probability with two peaks in q_0 and q_1. The second peak is present also when $T = T_s$, which means that it such states were already present also in the paramagnetic phase but with null thermodynamic weights. This fact indicates the presence of metastable states.

As we already mentioned, these considerations holds only for $p \geq 3$. In the case $p = 2$ the one-step symmetry breaking is not enough to found the minimum configuration and we need to continue to break the replica symmetry as we will briefly illustrate in the next section.

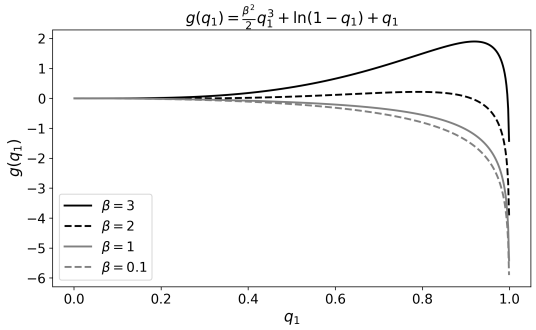


Figure C.4: *Left-hand side of Eq.* (C.70) *for different values of* β *and* $p = 3$.

Infinite-step replica symmetry breaking

We won't make here the explicit calculations but we only talk about the idea of infinite-step replica symmetry breaking for solving the spin-glass problem when $p = 2$, also called Sherrington and Kirkpatrick model.

One-step replica symmetry breaking is not enough to minimize the free energy of the system. Following the idea adopted by Parisi, a second-step for breaking the replica symmetry is to repeat the same procedure on the $\frac{n}{m}$ submatrices along the diagonal. If we call $m_1 \equiv m$, and we define another parameter $m_2 \leq m_1$, in any $m_1 \times m_1$ matrix we can add on the diagonal $\frac{m_1}{m_2}$ matrices with all 1 on the diagonal and another parameter q_2 for the off-diagonal terms.

After that, we repeat the procedure for the squared matrix $m_1 \times m_1$ intro-

ducing two other parameters $m_3 \leq m_2$ and $q_3 \geq q_2$ and so on. Iterating the process M times one obtain a set of parameters m_i and q_i such that

$$1 \leq m_M \leq m_{M-1} \leq \cdots \leq m_1 \leq n \tag{C.71}$$

$$1 \geq q_M \geq q_{M-1} \geq \cdots \geq q_1 \geq q_0 \geq 0 \tag{C.72}$$

Now let us imagine to iterate infinite times this process such that the parameters m_i tend to a continuous variable x and the overlaps become a function $q(x)$ of these parameters.

Note that in the limit $n \to 0$ the order relation in the first Eq. (C.71) is inverted, giving

$$m_M \geq m_{M-1} \geq \cdots \geq m_1 \geq n \to 0 \tag{C.73}$$

In the limit of small $t \equiv \frac{T-T_f}{T_F}$, where T_f is the critical temperature, called temperature of freezing, it can be shown that the minimum of energy has a solution

$$q(x) = \begin{cases} \frac{1}{2}x & \text{if} \quad x < 2|t| \\ |t| & \text{if} \quad x \geq 2|t| \end{cases} \tag{C.74}$$

when the external field H is zero. In the case of external field there is a plateau at $q_{min} = \frac{3}{4}\left(\frac{H}{J}\right)^{\frac{4}{3}}$ as showed in figure C.5. If the field increases, the plateau increases until the minimum reaches the value $q = |t|$ which gives a constant solution and hence the replica become symmetric again. This means that the external field destroys the spin-glass phase and the point in which this happens is called Almeida-Thouless transition line.

In the case T is not close to T_f the general solution for the free energy is given by

$$\beta f = \frac{(\beta J)^2}{4}\left[1 + \int_0^1 q^2(x)dx + -2q(1)\right] - \int_{-\infty}^{+\infty} \frac{dz}{2\pi} e^{-\frac{z^2}{2}} f_0(0, z\sqrt{q(0)})\,, \tag{C.75}$$

where f_0 solves the Parisi equation

$$\frac{\partial f_0(x,H)}{\partial x} = -\frac{J^2}{2}\frac{dq}{dx}\left[\frac{\partial f_0}{\partial H} + x\left(\frac{\partial f_0}{\partial H}\right)^2\right] \tag{C.76}$$

It has also been shown that the solution found by Parisi is stable (De Dominicis and Kondor 1983).

Furthermore, it is possible to compute the susceptibilities defined in the previous section, showing that they present a cusp in correspondence of the transition temperature T_f according to the observations.

The probability distribution for the overlap becomes

$$P(q) = (1 - 4|t|)\delta(q - 2|t|) + \begin{cases} 2x & \text{if} \quad x < 2|t| \\ 0 & \text{if} \quad x \geq 2|t| \end{cases} \tag{C.77}$$

Also in this case we have indications of the presence of metastable states. The only difference is that here there is a continuous spectrum of these states with different thermodynamic weights and only few of these dominates respect to the others.

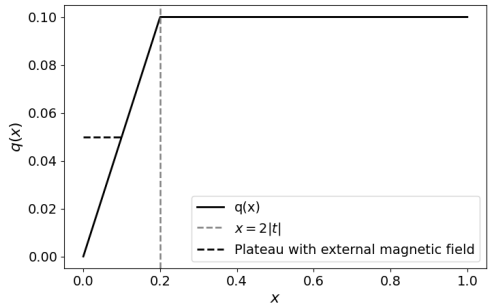


Figure C.5: *Plot of the Eq.* (C.74) *with* $t = 0.1$ *in both cases of zero external magnetic field and non-zero external magnetic field.*

C.2 Derivation of Equation (3.48)

The first derivation of the form of $\Pr_{\mathcal{S}}(\vec{k}|\vec{c})$ in the context of linear optics was obtained in Ref. (Scheel 2004). We describe it here for completeness. The evolution of (indistinguishable) photons through a scattering medium is described by a unitary matrix S, under which the creation operators transform as

$$\hat{a}_i^\dagger \to \sum_{j=1}^{M} S_{ji} \hat{a}_j^\dagger. \tag{C.78}$$

The generic input state of N_{ph} photons distributed over M modes is

$$|\Psi\rangle = \sum_{\vec{c}} a_{\vec{c}} |\vec{c}\rangle = \sum_{\vec{c}} \frac{1}{\sqrt{\mu(\vec{c})}} a_{c_1,\ldots,c_{N_{\text{ph}}}} \hat{a}_{c_1}^\dagger \ldots \hat{a}_{c_{N_{\text{ph}}}}^\dagger |\mathbf{0}\rangle \,, \tag{C.79}$$

where $|\mathbf{0}\rangle = |0\rangle^{\otimes M}$ and $\sum_{\vec{c}}(\ldots) \equiv \prod_{j=1}^{N_{\text{ph}}} \sum_{c_j=0}^{M-1}(\ldots)$. By applying the transformation (C.78), each term $|\vec{c}\rangle$ in the superposition transforms as

$$S(|\vec{c}\rangle) = \frac{1}{\sqrt{\mu\left(\vec{c}\right)}} \sum_{\vec{q}} S_{q_1 c_1} \ldots S_{q_{N_{\text{ph}}} c_{N_{\text{ph}}}} \hat{a}^\dagger_{q_1} \ldots \hat{a}^\dagger_{q_{N_{\text{ph}}}} |\mathbf{0}\rangle \ , \tag{C.80}$$

with $\mu(\vec{c}) = n_1! \cdot \cdots \cdot n_M!$, where n_i is the occupation number of the mode i and $\sum_{i=0}^{M-1} n_i = N_{\text{ph}}$. Therefore, the transition amplitude to the state

$$|\vec{k}\rangle = \frac{1}{\sqrt{\mu(\vec{k})}} \hat{a}^\dagger_{k_1} \ldots \hat{a}^\dagger_{k_{N_{\text{ph}}}} |\mathbf{0}\rangle \tag{C.81}$$

can be explicitly computed as $\langle \vec{k}|S(\vec{c})\rangle$. By using the bosonic canonical commutation relations

$$[\hat{a}_i, \hat{a}^\dagger_j] = \delta_{ij}, \quad [\hat{a}_i, \hat{a}_j] = [\hat{a}^\dagger_i, \hat{a}^\dagger_j] = 0 \tag{C.82}$$

with some algebra we obtain

$$\langle 0|\hat{a}_{k_1} \ldots \hat{a}_{k_{N_{\text{ph}}}} \hat{a}^\dagger_{q_1} \ldots \hat{a}^\dagger_{q_{N_{\text{ph}}}} |0\rangle = \sum_{\{q\}} \delta_{k_1, q_1} \ldots \delta_{k_{N_{\text{ph}}}, q_{N_{\text{ph}}}} \tag{C.83}$$

where the sum runs over all the permutations of the indices $q_1, \ldots q_{N_{\text{ph}}}$. Therefore, substituting the last expression in the Eq. (C.80), the only terms that survive are the one with $\vec{q} = \vec{k}$ and all their permutations, which leads to

$$\langle \vec{k}|S(\vec{c})\rangle = \frac{1}{\sqrt{\mu\left(\vec{c}\right)\mu(\vec{k})}} \text{Perm}\left(S_{\vec{k}|\vec{c}}\right) \tag{C.84}$$

that is, Eq. (3.48).

Bibliography

Aaronson, S. and A. Arkhipov (2013). “Boson sampling is far from uniform”. In: *arXiv preprint arXiv:1309.7460*.

Abbott, L. F. and Y. Arian (1988). “Storage capacity of generalized networks”. In: *Phys. Rev. A* 36, p. 5091.

Åberg, J. (2006). “Quantifying superposition”. In: arXiv:quant-ph/0612146.

Acín, A. (2001). “Statistical distinguishability between unitary operations”. In: *Phys. Rev. Lett.* 87, p. 177901.

Adesso, G. and F. Illuminati (2007). “Entanglement in continuous-variable systems: recent advances and current perspectives”. In: *J. Phys. A: Math. Theor.* 40, p. 7821.

Adesso, G., A. Serafini, and F. Illuminati (2004). “Extremal entanglement and mixedness in continuous variable systems”. In: *Phys. Rev. A* 70, p. 022318.

Agliari, E., L. Albanese, F. Alemanno, A. Alessandrelli, A. Barra, F. Giannotti, D. Lotito, and D. Pedreschi (2023). “Dense Hebbian neural networks: a replica symmetric picture of supervised learning”. In: *Physica A* 626, p. 129076.

Agliari, E., F. Alemanno, A. Barra, M. Centonze, and A. Fachechi (2020). “Neural networks with a redundant representation: Detecting the undetectable”. In: *Phys. Rev. Lett.* 124.2, p. 028301.

Albanese, L., A. Alessandrelli, A. Annibale, and A. Barra (2024). “Replica symmetry breaking in supervised and unsupervised Hebbian networks”. In: *J. Phys. A: Math. Theor.* 57.16, p. 165003.

Albanese, L., A. Alessandrelli, A. Barra, and P. Sollich (2025). “Yet another exponential Hopfield model”. In: *Physica A*, p. 131223.

Amit, D. J. (1989). *Modelling brain functions*. Cambridge University Press.

Amit, D. J., H. Gutfreund, and H. Sompolinsky (1985). “Storing Infinite Numbers of Patterns in a Spin-Glass Model of Neural Networks”. In: *Phys. Rev. Lett.* 55.14, pp. 1530–1533.

Arenzon, J. J. and R. M. C. de Almeida (1993). “Neural networks with high-order connections”. In: *Phys. Rev. E* 48.5, p. 4060.

Aspuru-Guzik, A. and P. Walther (2012). “Photonic quantum simulators”. In: *Nat. Phys.* 8.4, pp. 285–291.

Badziag, P., Č. Brukner, W. Laskowski, T. Paterek, and M. Żukowski (2008). "Experimentally friendly geometrical criteria for entanglement". In: *Phys. Rev. Lett.* 100.14, p. 140403.

Bae, J. and L. C. Kwek (2015). "Quantum state discrimination and its applications". In: *J. Phys. A: Math. Theor.* 48, p. 083001.

Baldi, P. and S. Venkatesh (1987). "Number of stable points for spin-glasses and neural networks of higher order". In: *Phys. Rev. Lett.* 58, p. 913.

Banchi, L., S. L. S. L. Braunstein, and S. Pirandola (2015). "Quantum fidelity for arbitrary Gaussian states". In: *Phys. Rev. Lett.* 115, p. 260501.

Barenco, A., C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter (1995). "Elementary gates for quantum computation". In: *Phys. Rev. A* 52.5, p. 3457.

Bartolucci, S., P. Birchall, H. Bombin, H. Cable, C. Dawson, M. Gimeno-Segovia, E. Johnston, K. Kieling, N. Nickerson, M. Pant, F. Pastawski, T. Rudolph, and C. Sparrow (2023). "Fusion-based quantum computation". In: *Nat. Commun.* 14.1, p. 912.

Baumgratz, T., M. Cramer, and M. B. Plenio (2014). "Quantifying coherence". In: *Phys. Rev. Lett.* 113, p. 140401.

Belinskii, A. V. and D. N. Klyshko (1993). "Interference of light and Bell's theorem". In: *Physics-Uspekhi* 36.8, pp. 653–667.

Bell, J. S. (1964). "On the Einstein Podolsky Rosen paradox". In: *Physics* 1, pp. 195–200.

Bengtsson, I. and K. Życzkowski (2017). *Geometry of quantum states: an introduction to quantum entanglement*. Cambridge, UK: Cambridge University Press.

Bennett, C. H., H. J. Bernstein, S. Popescu, and B. Schumacher (1996a). "Concentrating partial entanglement by local operations". In: *Phys. Rev. A* 53.4, p. 2046.

Bennett, C. H., G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters (1993). "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels". In: *Phys. Rev. Lett.* 70, pp. 1895–1899.

Bennett, C. H., D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters (1996b). "Mixed-state entanglement and quantum error correction". In: *Phys. Rev. A* 54.5, p. 3824.

Bergou, J. A. (2007). "Quantum state discrimination and selected applications". In: *J. Phys.: Conf. Ser.* 84, p. 012001.

Bernardini, A. E. (2006). "Flavor coupled with chiral oscillations in the presence of an external magnetic field". In: *Eur. Phys. J. C* 46.1, pp. 113–122.

— (2007). "Chiral oscillations in terms of the zitterbewegung effect". In: *Eur. Phys. J. C* 50.3, pp. 673–678.

Bernardini, A. E. and S. De Leo (2005). "Flavor and chiral oscillations with Dirac wave packets". In: *Phys. Rev. D* 71.7, p. 076008.

Bertlmann, R. and N. Friis (2023). *Modern Quantum Theory: From Quantum Mechanics to Entanglement and Quantum Information*. Oxford University Press.

Bertlmann, R. A. and P. Krammer (2008). "Geometric entanglement witnesses and bound entanglement". In: *Phys. Rev. A* 77.2, p. 024303.

Bhatia, R. (2013). *Matrix Analysis*. Vol. 169. Springer Science & Business Media.

Bittencourt, V. A. S. V., A. E. Bernardini, and M. Blasone (2021). "Chiral oscillations in the non-relativistic regime". In: *Eur. Phys. J. C* 81.5, p. 411.

Bittencourt, V. A. S. V., M. Blasone, and G. Zanfardino (2023). "Chiral and flavor oscillations in lepton-antineutrino spin correlations". In: *J. Phys.: Conf. Ser.* Vol. 2533. 1. IOP Publishing, p. 012024.

— (2024). "Chiral and flavor oscillations in a hyperentangled neutrino state". In: *Phys. Scr.* 99.12, p. 125310.

— (2025a). "Chiral and flavor oscillations in quantum field theory: an overview". In: *J. of Phys.: Conference Series*. Vol. 3017. 1. IOP Publishing, p. 012026.

— (2025b). "Chiral oscillations in quantum field theory". In: *Phys. Lett. B* 864, p. 139399.

Blasone, M., F. Dell'Anno, S. De Siena, and F. Illuminati (2008). "Hierarchies of geometric entanglement". In: *Phys. Rev. A* 77.6, p. 062304.

Blasone, M., F. Giacosa, L. Smaldone, and G. Torrieri (2025a). "Chiral and flavor oscillations in the interaction picture". In: *J. of Phys.: Conference Series*. Vol. 3017. 1. IOP Publishing, p. 012027.

— (2025b). "Chiral oscillations in finite time quantum field theory". In: *EPJ C* 85.5, p. 523.

Blasone, M., P. A. Henning, and G. Vitiello (1999). "The exact formula for neutrino oscillations". In: *Phys. Lett. B* 451.1-2, pp. 140–145.

Blasone, M. and G. Vitiello (1995). "Quantum field theory of fermion mixing". In: *Ann. Phys.* 244.2, pp. 283–311.

Blasone, M. and G. Zanfardino (2026). "Unified field-theoretical description of flavor and chiral oscillations". In: *Phys. Lett. B*, p. 140292.

Boschi, D., S. Branca, F. De Martini, L. Hardy, and S. Popescu (1998). "Experimental realization of teleporting an unknown pure quantum state via dual classical and Einstein-Podolsky-Rosen channels". In: *Phys. Rev. Lett.* 80, pp. 1121–1125.

Bouwmeester, D., J. W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger (1997). "Experimental quantum teleportation". In: *Nature* 390, pp. 575–579.

Bruck, J. and J. W. Goodman (1988). "A generalized convergence theorem for neural networks". In: *IEEE Trans. Inf. Theory* 34.5, pp. 1089–1092.

Brunner, N., D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner (2014). "Bell nonlocality". In: *Rev. Mod. Phys.* 86.2, pp. 419–478.

Bures, D. (1969). “An extension of Kakutani’s theorem on infinite product measures to the tensor product of semifinite w*-algebras”. In: *Trans. Am. Math. Soc.* 135, pp. 199–212.

Caianiello, E. R. (1961). “Outline of a theory of thought-processes and thinking machines”. In: *J. Theor. Biol.* 1.2, pp. 204–235.

Carruthers, P. and M. M. Nieto (1968). “Phase and angle variables in quantum mechanics”. In: *Rev. Mod. Phys.* 40.2, p. 411.

Carteret, H. A., A. Higuchi, and A. Sudbery (2000). “Multipartite generalisation of the Schmidt decomposition”. In: *arXiv preprint quant-ph/0006125*.

Chefles, A. and S. M. Barnett (1998). “Optimum unambiguous discrimination between linearly independent symmetric states”. In: *Phys. Lett. A* 250, pp. 223–229.

Chernoff, H. (1952). “A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations”. In: *Ann. Math. Stat.*, pp. 493–507.

Chitambar, E. and G. Gour (2019). “Quantum resource theories”. In: *Rev. Mod. Phys.* 91.2, p. 025001.

Christandl, M. and A. Winter (2004). “”Squashed entanglement”: an additive entanglement measure”. In: *J. Math. Phys.* 45.3, pp. 829–840.

Clements, W. R., P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley (2016). “Optimal design for universal multiport interferometers”. In: *Optica* 3.12, pp. 1460–1465.

Cohen-Tannoudji, C., B. Diu, and F. Laloë (2019). *Quantum Mechanics, Volume 3: Fermions, Bosons, Photons, Correlations and Entanglement*. John Wiley and Sons.

Collins, D. and N. Gisin (2004). “A relevant two qubit Bell inequality inequivalent to the CHSH inequality”. In: *J. Phys. A* 37.5, pp. 1775–1787.

Collins, D., N. Gisin, N. Linden, S. Massar, and S. Popescu (2002). “Bell inequalities for arbitrarily high-dimensional systems”. In: *Phys. Rev. Lett.* 88.4, p. 040404.

Crisanti, A. and H. J. Sommers (1992). “The spherical p-spin interaction spin glass model: the statics”. In: *Z. Phys. B: Condens. Matter* 87, pp. 341–354.

Dakić, B., Y. O. Lipp, X. Ma, M. Ringbauer, S. Kropatschek, S. Barz, T. Paterek, V. Vedral, A. Zeilinger, Č. Brukner, and P. Walther (2012). “Quantum discord as a resource for remote state preparation”. In: *Nat. Phys.* 8, pp. 666–670.

De Almeida, J. R. L. and D. J. Thouless (1978). “Stability of the Sherrington-Kirkpatrick solution of a spin glass model”. In: *J. Phys. A* 11, p. 983.

De Dominicis, C. and I. Kondor (1983). “Eigenvalues of the stability matrix for Parisi solution of the long-range spin-glass”. In: *Phys. Rev. B* 27, p. 606.

De Palma, G. and J. Borregaard (2018). “Minimum error probability of quantum illumination”. In: *Phys. Rev. A* 98, p. 012101.

De Vicente, J. I. (2014). “On nonlocality as a resource theory and nonlocality measures”. In: *J. Phys. A* 47, p. 424017.

De Vicente, J. I. and M. Huber (2011). “Multipartite entanglement detection from correlation tensors”. In: *Phys. Rev. A* 84.6, p. 062306.

Dieks, D. (1988). “Overlap and distinguishability of quantum states”. In: *Phys. Lett. A* 126, pp. 303–306.

Eberle, T., V. Händchen, and R. Schnabel (2013). “Stable control of 10 dB two-mode squeezed vacuum states of light”. In: *Opt. Express* 21.9, pp. 11546–11553.

Eggeling, T. and R. F. Werner (2001). “Separability properties of tripartite states with $U \otimes U \otimes U$ symmetry”. In: *Phys. Rev. A* 63.4, p. 042111.

Einstein, A., B. Podolsky, and N. Rosen (1935). “Can quantum-mechanical description of physical reality be considered complete?” In: *Phys. Rev.* 47.10, p. 777.

Eisert, J. and M. B. Plenio (1999). “A comparison of entanglement measures”. In: *J. Mod. Opt.* 46.1, pp. 145–154.

Ekert, A. K. (1991). “Quantum cryptography based on Bell’s theorem”. In: *Phys. Rev. Lett.* 67.6, p. 661.

Elben, A., B. Vermersch, M. Dalmonte, J. I. Cirac, and P. Zoller (2018). “Rényi entropies from random quenches in atomic Hubbard and spin models”. In: *Phys. Rev. Lett.* 120.5, p. 050406.

Eldar, Y. C., A. Megretski, and G. C. Verghese (2004a). “Optimal detection of symmetric mixed quantum states”. In: *IEEE Trans. Inf. Theory* 50, pp. 1198–1207.

Eldar, Y. C., M. Stojnic, and B. Hassibi (2004b). “Optimal quantum detectors for unambiguous detection of mixed states”. In: *Phys. Rev. A* 69, p. 062318.

Fan, Z., J. Lin, J. Dai, T. Zhang, and K. Xu (2023). “Photonic Hopfield neural network for the Ising problem”. In: *Opt. Express* 31.13, pp. 21340–21350.

Farhat, N. H., D. Psaltis, A. Prata, and E. Paek (1985). “Optical implementation of the Hopfield model”. In: *Appl. Opt.* 24.10, pp. 1469–1475.

Ferraro, A., S. Olivares, and M. G. A. Paris (2005). “Gaussian states in continuous variable quantum information”. In: *arXiv preprint* quant-ph/0503237.

Folli, V., G. Gosti, M. Leonetti, and G. Ruocco (2018). “Effect of dilution in asymmetric recurrent neural networks”. In: *Neural Networks* 104, pp. 50–59.

Friis, N., G. Vitagliano, M. Malik, and M. Huber (2019). “Entanglement certification from theory to experiment”. In: *Nat. Rev. Phys.* 1.1, pp. 72–87.

Fuchs, C.A. and J. Van De Graaf (2002). “Cryptographic distinguishability measures for quantum-mechanical states”. In: *IEEE Transactions on Information Theory* 45.4, pp. 1216–1227.

Gallego, R. and L. Aolita (2015). “Resource theory of steering”. In: *Phys. Rev. X* 5.4, p. 041008.

Gardner, E. (1987). “Multiconnected neural network models”. In: *J. Phys. A* 20, p. 3453.

Gerry, C. C. and P. L. Knight (2023). *Introductory Quantum Optics*. Cambridge University Press.

Gilmore, R. (1972). “Geometry of symmetrized states”. In: *Ann. of Phys.* 74.2, pp. 391–463.

Gimeno-Segovia, M., P. Shadbolt, D. E. Browne, and T. Rudolph (2015). “From three-photon Greenberger-Horne-Zeilinger states to ballistic universal quantum computation”. In: *Phys. Rev. Lett.* 115.2, p. 020502.

Giovannetti, V., S. Guha, S. Lloyd, L. Maccone, and J. H. Shapiro (2004). “Minimum output entropy of bosonic channels: a conjecture”. In: *Phys. Rev. A* 70, p. 032315.

Giovannetti, V., S. Lloyd, and L. Maccone (2011). “Advances in quantum metrology”. In: *Nat. Photon.* 5.4, pp. 222–229.

Gisin, N. (1996). “Hidden quantum nonlocality revealed by local filters”. In: *Phys. Lett. A* 210.3, pp. 151–156.

Glauber, R. J. (1963a). “Coherent and incoherent states of the radiation field”. In: *Phys. Rev.* 131, pp. 2766–2788.

— (1963b). “Time-dependent statistics of the Ising model”. In: *J. Math. Phys.* 4.2, pp. 294–307.

— (2007). *Quantum Theory of Optical Coherence*. John Wiley and Sons.

Gonda, T. and R. W. Spekkens (2023). “Monotones in general resource theories”. In: *Compositionality* 5, p. 7.

Gravina, S., N. A. Chishti, S. Di Bernardo, E. Fasci, A. Castrillo, A. Laliotis, and L. Gianfrani (2024). “Comb-Referenced Doppler-Free Spectrometry of the Hg 200 and Hg 202 Intercombination Line at 254 nm”. In: *Phys. Rev. Lett.* 132.21, p. 213001.

Griffiths, D. J. and D. F. Schroeter (2018). *Introduction to quantum mechanics*. Cambridge University Press.

Guerra, F. (2003). “Broken replica symmetry bounds in the mean field spin glass model”. In: *Commun. Math. Phys.* 233, pp. 1–12.

Gühne, O. (2004). “Characterizing entanglement via uncertainty relations”. In: *Phys. Rev. Lett.* 92.11, p. 117903.

Gühne, O., P. Hyllus, O. Gittsovich, and J. Eisert (2007). “Covariance matrices and the separability problem”. In: *Phys. Rev. Lett.* 99.13, p. 130504.

Gühne, O., M. Mechler, G. Tóth, and P. Adam (2006). “Entanglement criteria based on local uncertainty relations are strictly stronger than the computable cross norm criterion”. In: *Phys. Rev. A* 74.1, p. 010301.

Gühne, O. and G. Tóth (2009). “Entanglement detection”. In: *Phys. Rep.* 474.1-6, pp. 1–75.

Ha, D. and Y. Kwon (2013). “Complete analysis for three-qubit mixed-state discrimination”. In: *Phys. Rev. A* 87, p. 062302.

Hebb, D. O. (2005). *The organization of behavior: A neuropsychological theory*. Psychology Press.

Hellinger, E. (1909). "Neue Begründung der Theorie quadratischer Formen von unendlich vielen Veränderlichen". In: *J. Reine Angew. Math.* 136. In German, pp. 210–271.

Helstrom, C. W. (1967). "Detection theory and quantum mechanics". In: *Inf. Control* 10, pp. 254–291.

— (1969). "Quantum detection and estimation theory". In: *J. Stat. Phys.* 1, pp. 231–252.

Hill, S. and W. K. Wootters (1997). "Entanglement of a pair of quantum bits". In: *Phys. Rev. Lett.* 78, pp. 5022–5025.

Hofmann, H. F. and S. Takeuchi (2003). "Violation of local uncertainty relations as a signature of entanglement". In: *Phys. Rev. A* 68.3, p. 032103.

Hopfield, J. J. (1982). "Neural networks and physical systems with emergent collective computational abilities". In: *Proc. Natl. Acad. Sci. U.S.A.* 79.8, pp. 2554–2558.

Horn, D. and M. Usher (1988). "Capacity of multiconnected memory models". In: *J. Phys. France* 49, p. 389.

Horn, R. A. and C. R. Johnson (2012). *Matrix Analysis*. Cambridge University Press.

Horodecki, M. and P. Horodecki (1999). "Reduction criterion of separability and limits for a class of distillation protocols". In: *Phys. Rev. A* 59.6, p. 4206.

Horodecki, M., P. Horodecki, and R. Horodecki (1998). "Mixed-state entanglement and distillation: Is there a "bound" entanglement in nature?" In: *Phys. Rev. Lett.* 80.24, p. 5239.

— (2001). "Separability of n-particle mixed states: necessary and sufficient conditions in terms of linear maps". In: *Phys. Lett. A* 283.1-2, pp. 1–7.

Horodecki, M., P. Horodecki, and Horodecki R. (1996). "Separability of mixed states: necessary and sufficient conditions". In: *Phys. Lett. A* 223.25, pp. 1–8.

Horodecki, P. (1997). "Separability criterion and inseparable mixed states with positive partial transposition". In: *Phys. Lett. A* 232.5, pp. 333–339.

Horodecki, R., P. Horodecki, and M. Horodecki (1995). "Violating Bell inequality by mixed spin-1/2 states: necessary and sufficient condition". In: *Phys. Lett. A* 200, pp. 340–344.

Horodecki, R., P. Horodecki, M. Horodecki, and K. Horodecki (2009). "Quantum entanglement". In: *Rev. Mod. Phys.* 81, pp. 865–942.

Hukushima, K. and K. Nemoto (1996). "Exchange Monte Carlo method and application to spin glass simulations". In: *Journal of the Physical Society of Japan* 65.6, pp. 1604–1608.

Hwang, S., V. Folli, E. Lanza, G. Parisi, G. Ruocco, and F. Zamponi (2019). "On the number of limit cycles in asymmetric neural networks". In: *J. Stat. Mech.* 2019.5, p. 053402.

Ivanovic, I. D. (1987). "How to differentiate between non-orthogonal states". In: *Phys. Lett. A* 123, pp. 257–259.

Jaeger, G. and A. Shimony (1995). "Optimal distinction between two non-orthogonal quantum states". In: *Phys. Lett. A* 197, pp. 83–87.

Kimble, H. J. (1998). "Strong interactions of single atoms and photons in cavity QED". In: *Physica Scripta* T76, p. 127.

Knill, E., G. J. Laflamme, and R. Milburn (2001). "A scheme for efficient quantum computation with linear optics". In: *Nature* 409.6816, pp. 46–52.

Kosmatopoulos, E. B., M. M. Polycarpou, M. A. Christodoulou, and P. A. Ioannou (1995). "High-order neural network structures for identification of dynamical systems". In: *IEEE Trans. Neural Networks* 6.2, pp. 422–431.

Krotov, D. and J. J. Hopfield (2016). "Dense associative memory for pattern recognition". In: *Adv. Neural Inf. Process. Syst.* 29.

Kullback, S. and R. A. Leibler (1951). "On information and sufficiency". In: *Ann. Math. Stat.* 22.1, pp. 79–86.

Kuroiwa, K. and H. Yamasaki (2020). "General quantum resource theories: distillation, formation and consistent resource measures". In: *Quantum* 4, p. 355.

Laskowski, W., T. Paterek, M. Żukowski, and C. Brukner (2004). "Tight multipartite Bell's inequalities involving many measurement settings". In: *Phys. Rev. Lett.* 93, p. 200401.

Laurat, J., G. Keller, J. A. O. Huguenin, C. Fabre, T. Coudreau, A. Serafini, G. Adesso, and F. Illuminati (2005). "Entanglement of two-mode Gaussian states: characterization and experimental production and manipulation". In: *J. Opt. B: Quantum Semiclass. Opt.* 7.12, S577.

Leonetti, M., G. Gosti, and G. Ruocco (2024). "Photonic stochastic emergent storage for deep classification by scattering-intrinsic patterns". In: *Nat. Commun.* 15.1, p. 505.

Leonetti, M., E. Hörmann, L. Leuzzi, G. Parisi, and G. Ruocco (2021). "Optical computation of a spin glass dynamics with tunable complexity". In: *Proc. Natl. Acad. Sci. U.S.A.* 118.21, e2015207118.

Little, W. A. (1974). "The existence of persistent states in the brain". In: *Math. Biosci.* 19.1-2, pp. 101–120.

Lloyd, S. (2008). "Enhanced sensitivity of photodetection via quantum illumination". In: *Science* 321, pp. 1463–1465.

Lopaeva, E. D., I. R. Berchera, I. P. Degiovanni, S. Olivares, G. Brida, and M. Genovese (2013). "Experimental realization of quantum illumination". In: *Phys. Rev. Lett.* 110, p. 153603.

Maccone, L. and A. K. Pati (2014). “Stronger uncertainty relations for all incompatible observables”. In: *Phys. Rev. Lett.* 113.26, p. 260401.

Madsen, L. S., F. Laudenbach, M. F. Askarani, F. Rortais, T. Vincent, J. F. F. Bulmer, F. M. Miatto, L. Neuhaus, L. G. Helt, M. J. Collins, et al. (2022). “Quantum computational advantage with a programmable photonic processor”. In: *Nature* 606.7912, pp. 75–81.

Marcus, C. M. and R. M. Westervelt (1989). “Dynamics of iterated-map neural networks”. In: *Phys. Rev. A* 40.1, p. 501.

Marshall, A. W., I. Olkin, and B. C. Arnold (1979). *Inequalities: Theory of majorization and its applications*. Springer.

Masanes, L. (2002). “Tight Bell inequality for d-outcome measurements correlations”. In: *arXiv:quant-ph/0210073*.

Masanes, L., Y. C. Liang, and A. C. Doherty (2008). “All bipartite entangled states display some hidden nonlocality”. In: *Phys. Rev. Lett.* 100.9, p. 090403.

Maser, A., B. Gmeiner, T. Utikal, S. Götzinger, and V. Sandoghdar (2016). “Few-photon coherent nonlinear optics with a single molecule”. In: *Nat. Photonics* 10.7, pp. 450–453.

McCulloch, W. S. and W. Pitts (1943). “A logical calculus of the ideas immanent in nervous activity”. In: *Bull. Math. Biophys.* 5, pp. 115–133.

Mermin, N. D. (1990). “Extreme quantum entanglement in a superposition of macroscopically distinct states”. In: *Phys. Rev. Lett.* 65.15, pp. 1838–1840.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (June 1953). “Equation of State Calculations by Fast Computing Machines”. In: *The Journal of Chemical Physics* 21.6, pp. 1087–1092.

Mézard, M., G. Parisi, and M. A. Virasoro (1987). *Spin Glass Theory and Beyond: An Introduction to the Replica Method and Its Applications*. World Scientific.

Minsky, M. and S. Papert (1969). “An introduction to computational geometry”. In: *Cambridge, Mass., MIT* 479.480, p. 104.

Mizuno, J., K. Wakui, A. Furusawa, and M. Sasaki (2005). “Experimental demonstration of entanglement-assisted coding using a two-mode squeezed vacuum state”. In: *Phys. Rev. A* 71.1, p. 012304.

Müller, B., J. Reinhardt, and M. T. Strickland (2012). *Neural Networks: An Introduction*. Springer Science & Business Media.

Muthali, A.L., P. I. Sund, C.F.D. Faurby, Çağın. Ekici, J.C. Adcock, J. Glindemann, G. Zanfardino, L. Leuzzi, R. Santagati, F. Illuminati, G. Ruocco, M. Leonetti, S. Paesani, and Y. Ding (2026). “Photonic simulation of disordered spin systems on a silicon chip”. In: *Adv. Phot. Nex.* 5.2, pp. 025001–025001.

Mydosh, J. A. (1993). *Spin Glasses, An Experimental Introduction*. CRC Press.

Navascués, M. and T. Vértesi (2011). “Activation of nonlocal quantum resources”. In: *Phys. Rev. Lett.* 106.6, p. 060403.

Nha, H. and H. J. Carmichael (2005). “Distinguishing two single-mode Gaussian states by homodyne detection: an information-theoretic approach”. In: *Phys. Rev. A* 71, p. 032336.

Nielsen, M. A. and I. L. Chuang (2002). *Quantum Computation and Quantum Information*. Cambridge University Press, pp. 558–559.

Nielsen, M. A. and J. Kempe (2001). “Separable states are more disordered globally than locally”. In: *Phys. Rev. Lett.* 86.22, p. 5184.

Pal, P. B. (2014). *An introductory course of particle physics*. Taylor & Francis.

Palazuelos, C. (2012). “Superactivation of quantum nonlocality”. In: *Phys. Rev. Lett.* 109.19, p. 190401.

Paris, M. and J. Rehacek (2004). *Quantum state estimation*. Vol. 649. Lecture Notes in Physics. Springer Science and Business Media.

Parisi, G. (1983). “Order parameter for spin-glasses”. In: *Phys. Rev. Lett.* 50, p. 1946.

Parisi, G., P. Urbani, and F. Zamponi (2020). *Theory of Simple Glasses, Exact Solutions in Infinite Dimensions*. Cambridge University Press.

Pati, A. K. and P. K. Sahu (2007). “Sum uncertainty relation in quantum theory”. In: *Phys. Lett. A* 367.3, pp. 177–181.

Perelomov, A. M. (1972). “Coherent states for arbitrary Lie group”. In: *Comm. in Math. Phys.* 26.3, pp. 222–236.

Peres, A. (1988). “How to differentiate between non-orthogonal states”. In: *Phys. Lett. A* 128, pp. 19–20.

— (1996). “Separability criterion for density matrices”. In: *Phys. Rev. Lett.* 77.8, p. 1413.

— (2002). *Quantum theory: concepts and methods*. Springer.

Peretto, P. and J. J. Niez (1986). “Long term memory storage capacity of multiconnected neural networks”. In: *Biol. Cybern.* 54, pp. 53–63.

Piani, M. (2012). “Problem with geometric discord”. In: *Phys. Rev. A* 86, p. 034101.

Pierangeli, D., G. Marcucci, and C. Conti (2019). “Large-scale photonic Ising machine by spatial light modulation”. In: *Phys. Rev. Lett.* 122.21, p. 213902.

Pirandola, S. (2011). “Quantum reading of a classical digital memory”. In: *Phys. Rev. Lett.* 106.9, p. 090504.

— (2014). “Quantum discord as a resource for quantum cryptography”. In: *Sci. Rep.* 4, p. 6956.

Pirandola, S., B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd (2018). “Advances in photonic quantum sensing”. In: *Nat. Photonics* 12.12, pp. 724–733.

Pirandola, S., R. Laurenza, C. Lupo, and J. L. Pereira (2019). “Fundamental limits to quantum channel discrimination”. In: *npj Quantum Inf.* 5, p. 50.

Pirandola, S. and S. Lloyd (2008). “Computable bounds for the discrimination of Gaussian states”. In: *Phys. Rev. A* 78, p. 012331.

Plenio, M. B. (2005). “Logarithmic negativity: a full entanglement monotone that is not convex”. In: *Phys. Rev. Lett.* 95.9, p. 090503.
Plenio, M. B. and S. Virmani (2007). “An introduction to entanglement measures”. In: arXiv:quant-ph/0504163.
Popescu, S. (1994). “Bell’s inequalities versus teleportation: What is nonlocality?” In: *Phys. Rev. Lett.* 72.6, pp. 797–800.
— (1995). “Bell’s inequalities and density matrices: revealing “hidden” nonlocality”. In: *Phys. Rev. Lett.* 74.14, pp. 2619–2622.
Ramsauer, H., B. Schäfl, J. Lehner, P. Seidl, M. Widrich, T. Adler, L. Gruber, M. Holzleitner, M. Pavlović, G. K. Sandve, V. Greiff, D. Kreil, M. Kopp, G. Klambauer, J. Brandstetter, and S. Hochreiter (2021). *Hopfield networks are all you need*. arXiv: `2008.02217 [cs.NE]`.
Reck, M., A. Zeilinger, H. J. Bernstein, and P. Bertani (1994). “Experimental realization of any discrete unitary operator”. In: *Phys. Rev. Lett.* 73.1, p. 58.
Robertson, H. P. (1929). “The uncertainty principle”. In: *Phys. Rev.* 34.1, p. 163.
Rosati, M., G. De Palma, A. Mari, and V. Giovannetti (2017). “Optimal quantum state discrimination via nested binary measurements”. In: *Phys. Rev. A* 95, p. 042307.
Rosenblatt, F. (1962). *Principles of Neurodynamics*. Spartan Books, New York.
Rudolph, O. (2000). “A separability criterion for density operators”. In: *J. Phys. A* 33.21, p. 3951.
— (2003). “Some properties of the computable cross-norm criterion for separability”. In: *Phys. Rev. A* 67.3, p. 032312.
— (2005). “Further results on the cross norm criterion for separability”. In: *Quantum Inf. Process.* 4, pp. 219–239.
Sacchi, M. F. (2005). “Optimal discrimination of quantum operations”. In: *Phys. Rev. A* 71, p. 062340.
Scheel, S. (2004). *Permanents in linear optical networks*.
Schmid, D., D. Rosset, and F. Buscemi (2020). “The type-independent resource theory of local operations and shared randomness”. In: *Quantum* 4, p. 262.
Schrödinger, E. (1930). “Sitzungsberichte der Preussischen Akademie der Wissenschaften”. In: *Physikalisch-mathematische Klasse* 14, p. 296.
— (1935). “Discussion of probability relations between separated systems”. In: *Proc. Cambridge Philos. Soc.* 31.4, pp. 555–563.
Shao, L. H., Z. Xi, H. Fan, and Y. Li (2015). “Fidelity and trace-norm distances for quantifying coherence”. In: *Phys. Rev. A* 91, p. 042120.
Shapiro, J. H. (2020). “The quantum illumination story”. In: *IEEE Aerospace and Electron. Syst. Mag.* 35, pp. 8–20.
Sparrow, C., E. Martín-López, N. Maraviglia, A. Neville, C. Harrold, J. Carolan, Y. N. Joglekar, T. Hashimoto, N. Matsuda, J. L. O’Brien, et al. (2018). “Simulating the vibrational quantum dynamics of molecules using photonics”. In: *Nature* 557.7707, pp. 660–667.

Streltsov, A., G. Adesso, and M. B. Plenio (2017). “Colloquium: Quantum coherence as a resource”. In: *Rev. Mod. Phys.* 89.4, p. 041003.

Talagrand, M. (2003). *Spin glasses: a challenge for mathematicians: cavity and mean field models*. Vol. 46. Springer Science & Business Media.

Tan, S. H., B. I. Erkmen, V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, S. Pirandola, and J. H. Shapiro (2008). “Quantum illumination with Gaussian states”. In: *Phys. Rev. Lett.* 101, p. 253601.

Terhal, B. M. (2000). “Bell inequalities and the separability criterion”. In: *Phys. Lett. A* 271, pp. 319–326.

Theurer, T. A. (2021). “Resource theories of states and operations”. Doctoral Dissertation. Universität Ulm.

Uhlmann, A. (1976). “The “transition probability” in the state space of a *-algebra”. In: *Rep. Math. Phys.* 9, pp. 273–279.

Van Enk, S. J. and C. W. J. Beenakker (2012). “Measuring $\mathrm{Tr}\rho^n$ on single copies of ρ using random measurements”. In: *Phys. Rev. Lett.* 108, p. 110503.

Vedral, V. and M. B. Plenio (1998). “Entanglement measures and purification procedures”. In: *Phys. Rev. A* 57, pp. 1619–1633.

Vedral, V., M. B. Plenio, M. A. Rippin, and P. L. Knight (1997). “Quantifying entanglement”. In: *Phys. Rev. Lett.* 78, pp. 2275–2279.

Vidal, G. and R. F. Werner (2002). “Computable measure of entanglement”. In: *Phys. Rev. A* 65, p. 032314.

Vollbrecht, K. G. H. and M. M. Wolf (2002). “Conditional entropies and their relation to entanglement criteria”. In: *J. Math. Phys.* 43.9, pp. 4299–4306.

Wei, T. C. and P. M. Goldbart (2003). “Geometric measure of entanglement and applications to bipartite and multipartite quantum states”. In: *Phys. Rev. A* 68.4, p. 042307.

Werner, R. F. (1989). “Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model”. In: *Phys. Rev. A* 40, pp. 4277–4281.

Wootters, W. K. (1998). “Entanglement of formation of an arbitrary state of two qubits”. In: *Phys. Rev. Lett.* 80, pp. 2245–2248.

Yanay, Y. and C. Tahan (2021). “Quantum state tomography and purity measurement of a multiqubit system with random single-pulse rotations”. In: *Phys. Rev. A* 103.6, p. 062411.

Zanfardino, G., S. Paesani, L. Leuzzi, R. Santagati, F. Illuminati, G. Ruocco, and M. Leonetti (2026a). “Multiphoton quantum simulation of the generalized Hopfield memory model”. In: *Physical Review Letters* 136.7, p. 070602.

Zanfardino, G., W. Roga, G. Tartaglione, M. Takeoka, and F. Illuminati (2026b). “Geometric measures of quantum nonlocality: characterization, quantification, and comparison by distances and operations”. In: *J. Phys. A* 59.5, p. 055301.

Zhong, H.-S., H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, et al. (2020). “Quantum computational advantage using photons”. In: *Science* 370.6523, pp. 1460–1463.

Żukowski, M. and Č. Brukner (2002). “Bell’s theorem for general N-qubit states”. In: *Phys. Rev. Lett.* 88.21, p. 210401.

Życzkowski, K. and H. J. Sommers (2001). “Induced measures in the space of mixed quantum states”. In: *J. Phys. A* 34.35, p. 7111.