The Telescope, the Fluxions, the invention of Logarithms and the frenzy of multiplication, often for its own sake, that follow'd have for Emerson all been steps of an unarguable approach to God, a growing clarity,— Gravity, the Pulse of Time, the finite speed of Light present themselves to him as aspects of God's character. It's like becoming friendly with an erratic, powerful, potentially dangerous member of the Aristocracy. He holds no quarrel with the Creator's sovereignty, but is repeatedly appall'd at the lapses in Attention, the flaws in Design, the squand'rings of life and energy, the failures to be reasonable, or to exercise common sense,— first appall'd, then angry. We are taught,— we believe,— that it is love of the Creation which drives the Philosopher in his Studies. Emerson is driven, rather, by a passionate Resentment. (Thomas Pynchon — Mason & Dixon)

Contents

I Tools **5**

1 Quantum theory **7**
1.1 States 7
1.2 Transformations 12
1.3 Measurements 14
1.4 Gell-Mann operators and Bloch vectors 17

2 Mathematical optimization **21**
2.1 Convexity 21
2.2 Linear programming 21
2.3 Semidefinite programming 21
2.3.1 Optimizing the measurement incompatibility robustness . . . 21

II Prepare and measure **23**

3 The prepare-and-measure scenario **25**
3.1 Prepare-and-measure behaviors 26
3.1.1 Classical preparations 26
3.1.2 Quantum preparations 28
3.2 Prepare-and-measure communication protocols 30
3.2.1 Random access coding 30
3.2.2 Dense coding 32

4 Classicality in the prepare and measure scenario **33**

5 Dense coding in the prepare and measure scenario **35**
5.1 Optimizing preparations and measurements for dense coding 35

Appendices **37**

Appendix A Proofs for chap. 4 **39**

Appendix B Proofs for chap. 5 **41**

Bibliography **49**

+.6cm+.6cm

Introduction

Part I

Tools

Chapter 1

Quantum theory

States, transformations and measurements are the basic building blocks in the description of any physical system. In this chapter I review the mathematical structures that quantum theory assigns to each of these building blocks, emphasizing the aspects that most drastically differ from their classical counterparts.

1.1 States

A structure at the core of the most unusual of quantum behaviors is *entanglement* — a property that may or not be present in composite quantum systems. The unusual correlations that entanglement may originate were pointed out as early as 1935 by Einstein, Podolsky and Rosen [1], and also discussed by Schrödinger, who coined the term (originally, "Verschränkung") [2]. This discussion was later rescued by John Bell in 1964 [3], and since then been extensively tested and developed [4], ultimately becoming a central feature of many quantum informational protocols, such as superdense coding [5], quantum key distribution [6] and teleportation [7]. Entanglement theory is a vast and endlessly interesting field of study in itself, some aspects of which I now review with special focus on bipartite systems, and making no attempt of comprehensiveness.

* * *

Density operator

A *quantum state* is described by a *density operator*, commonly denoted by ρ. Any density operator is a linear, unit-trace and positive-semidefinite operator in a Hilbert space $\mathcal{H}$. Conversely, any operator satisfying these properties represents a valid quantum state. Hence,

$$\mathcal{D}(\mathcal{H}^d) = \{\rho \in \mathcal{L}(\mathcal{H}^d) \mid \mathrm{tr}\rho = 1, \rho \succeq 0\}, \tag{1.1}$$

where $\mathcal{L}(\mathcal{H}^d)$ is the set of linear operators in $\mathcal{H}^d$, is the space of density operators in dimension d. Only finite-dimensional Hilbert spaces will be considered.

As $\rho \succeq 0 \implies \rho = \rho^\dagger$, we can use the spectral decomposition to write $\rho = \sum_m m\Pi_m$, where each Π_m is a projection onto the eigenspace of the eigenvector of ρ associated with eigenvalue m. Such eigenvectors are orthogonal. They need not be normalized, but we can always and will take them as being. Orthogonality implies that $\Pi_m\Pi_n = \delta_{mn}\Pi_m$ and $1 \leq \mathrm{rank}(\rho) \leq d$, and normalization that $\mathrm{tr}\,(\Pi_m) = 1$. Furthermore, all $m \geq 0$, and $\mathrm{tr}\,(\rho) = 1 \implies \sum_m m = 1$.

Pure states

You may recall the more usual definition that a quantum state is described by a unit vector in a Hilbert space and, conversely, that such unit vectors describe quantum states. This is only true for a subset of states called *pure quantum states*. Following along the tradition, we will denote pure quantum states as $|\psi\rangle$, where ψ is some label that describes the state. Similar notation is used for the dual vector $(|\psi\rangle)^\dagger \equiv \langle\psi|$, which is useful to write the inner product between any two vectors in the same space as $\langle\psi|\phi\rangle$, and the outer product as $|\psi\rangle\langle\phi|$. An useful geometric intuition on these products is to interpret the inner product as the overlap between $|\psi\rangle$ and $|\phi\rangle$, and an outer product $|\psi\rangle\langle\psi|$ as a projection onto $|\psi\rangle$.

Recalling that all eigenvectors of a density operator ρ are normalized, we may now interpret the $\Pi_m \equiv |m\rangle\langle m|$ as projections onto the pure states labeled by $|m\rangle$, and the spectral decomposition $\rho = \sum_m m\Pi_m$ as a probability distribution, weighted by the eigenvalues m, over those. Any pure state $|\psi\rangle$ can equivalently be described as $\rho = |\psi\rangle\langle\psi|$, and whenever $\operatorname{rank}(\rho) = 1$, we may infer that ρ stands for a pure state. Equivalently, whenever ρ is such a one-dimensional projection, the *purity* $\operatorname{tr}\left(\rho^2\right) = 1$, while in general $1/d \leq \operatorname{tr}\left(\rho^2\right) \leq 1$. This is one of the reasons why density matrices are more general than pure states. All other density operators (i.e., those of non-unit rank) are said to describe *mixed quantum states*. Although the spectral decomposition of ρ suggests that a mixed state can be interpreted as a probability distribution over pure states, the understanding of a mixed state as lack of knowledge on the exact state of the system should not be taken literally. One of several reasons for this assertion is that there may be many pure state ensembles generating the same density operator [8].

Given a basis $\{|e_i\rangle\}_{i=1}^d$ for $\mathcal{H}^d$, any pure state $|\psi\rangle$ in $\mathcal{H}^d$ can be written as $|\psi\rangle = \sum_{i=1}^d c_i |e_i\rangle$, where the $c_i \in \mathbb{C}$ and $\sum_i |c_i|^2 = 1$ due to $\langle\psi|\psi\rangle = 1$. We will frequently be interested in $\mathcal{H}^2$, in which it's common to work with the orthonormal *computational basis* $\{|0\rangle, |1\rangle\}$. The vector representations associated with the computational basis elements are $|0\rangle \equiv (1\ 0)^\intercal$ and $|1\rangle \equiv (0\ 1)^\intercal$. Any $|\psi\rangle \in \mathcal{H}^2$ can thus be identified with $|\psi\rangle = c_1|0\rangle + c_2|1\rangle = (c_1\ c_2)^\intercal$. An extension to a generalized d-dimensional computational basis $\{|i\rangle\}_{i=0}^{d-1}$ is similarly done. Due to its analogy with two-level classical systems (bits), a $|\psi\rangle \in \mathcal{H}^2$ is termed a quantum bit (*qubit*) and, similarly, any $|\psi\rangle \in \mathcal{H}^d$ is a qu*d*it.

Entanglement

Entanglement — and its opposite concept, *separability* —, are properties related to composite quantum systems. If we choose 2 for the number of subsystems, the underlying Hilbert space $\mathcal{H}$ of a state ρ can be correspondingly factored as $\mathcal{H} \equiv \mathcal{H}_A \otimes \mathcal{H}_B$, for choices of $\mathcal{H}_A$ and $\mathcal{H}_B$ respecting $\dim \mathcal{H}_A \dim \mathcal{H}_B = \dim \mathcal{H}$. Using the tensor product for composition is the quantum analogue of using the Cartesian product to build composite phase spaces in classical mechanics.

Letting $\{|\psi_i\rangle\}_{i=1}^{d_A}$ and $\{|\varphi_\alpha\rangle\}_{\alpha=1}^{d_B}$ be orthonormal bases for $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, we can easily build an orthonormal basis for $\mathcal{H}$ as $\{|\psi_i\rangle \otimes |\varphi_\alpha\rangle\}_{i=1,\alpha=1}^{d_A,d_B}$. This means that any vector $|\psi\rangle \in \mathcal{H}$ has a decomposition

$$|\psi\rangle \sum_{i=1}^{d_A} \sum_{\alpha=1}^{d_B} c_{i\alpha} |\psi_i, \varphi_\alpha\rangle . \tag{1.2}$$

Analogously, with $\{|\psi_i\rangle\langle\psi_j|\}_{i,j=1}^{d_A}$ as a basis for $\mathcal{L}(\mathcal{H}_A)$ and $\{|\varphi_\alpha\rangle\langle\varphi_\beta|\}_{\alpha,\beta=1}^{d_B}$ one for

$\mathcal{L}(\mathcal{H}_B)$, any operator $O \in \mathcal{L}(\mathcal{H})$ may be decomposed as

$$O = \sum_{ij\alpha\beta} O_{ij\alpha\beta} \, |\psi_i\rangle\langle\psi_j| \otimes |\varphi_\alpha\rangle\langle\varphi_\beta| \tag{1.3}$$

Suppose A is an operator acting only on the part $O^A \in \mathcal{L}(\mathcal{H}_A)$ of O. The corresponding operator in $\mathcal{H}$ is just $A \otimes \mathbf{1}_B$. To find a description for O^A, we notice that the expectation value $\mathrm{tr}\,(A \otimes \mathbf{1}_B\, O)$ should be equal to $\mathrm{tr}\left(A\, O^A\right)$, and to comply with it we define the *partial trace* over B, $\mathrm{tr}_B : \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B) \mapsto \mathcal{L}(\mathcal{H}_A)$, as

$$\mathrm{tr}_B\,(O) \equiv \sum_{ij\alpha\beta} O_{ij\alpha\beta}\, \mathrm{tr}\,(|\psi_i\rangle\langle\psi_j|) \otimes |\varphi_\alpha\rangle\langle\varphi_\beta| = \sum_{i\alpha\beta} O_{ii\alpha\beta}\, |\varphi_\alpha\rangle\langle\varphi_\beta|$$

and call $O^A = \mathrm{tr}_B\,(O)$ the reduced operator. This definition can be trivially adapted to tracing out $\mathcal{H}_A$ instead, and to dealing with more than two subsystems. Moreover, it can be shown that this is the unique operation satisfying the expectation value equality condition, and that it is also completely positive and trace preserving (the significance of these latter conditions will be discussed later) [][**To-Do:**]. This operation is especially useful when applied to density operators, in which case we call $\mathrm{tr}_B\,(\rho) = \rho^A$ the *reduced state* (of ρ in subsystem A).

Given a factorization of $\mathcal{H}$, a state ρ acting on $\mathcal{H}$ is said to be *separable* if and only if it can be written as

$$\rho = \sum_i p_i \rho_i^A \otimes \rho_i^B \tag{1.4}$$

where the p_i are probabilities, and $\rho_i^A \in \mathcal{H}_A$; correspondingly for ρ_i^B. A state that is not separable is entangled. For pure states, this reduces to $|\psi\rangle_{AB} = |\psi\rangle_A \otimes |\psi\rangle_B$, and when this form is not possible, $|\psi\rangle_{AB}$ is entangled.

Asking whether a state $\rho \in \mathcal{H}$ is entangled is only meaningful when the factorization structure is specified. As a matter of fact, any bipartite pure entangled state $|\psi\rangle \in \mathcal{H}_A^2 \otimes \mathcal{H}_B^2$ can be made separable by a fitting choice of factorization [9]. This observation implies that entanglement is a property of a state *with respect to* a choice of subsystems, and not of the state in itself. Naturally, one may also discuss entanglement in larger number of subsystems [4], but the complexity scales significantly fast, and the discussion would be of little usefulness to our objective.

A rationale for the definition of separability comes from a preparation procedure [10]. Consider two separate laboratories, each equipped with a device that prepares quantum states, and sharing a source of (classical) randomness. Given a random number i, generated by the source with probability p_i, the laboratories locally prepare states ρ_i^A and ρ_i^B. Now suppose the first laboratory measures $\mathcal{M}_A$, and the second $\mathcal{M}_B$, each on their respective preparation. For a pair of measurement effects $E_{m_A} \in \mathcal{M}_A$ and $E_{m_B} \in \mathcal{M}_B$, we then have

$$p(m_A, m_B) = \sum_i p_i \mathrm{tr}\left(E_{m_A} \rho_i^A\right) \mathrm{tr}\left(E_{m_B} \rho_i^B\right) = \mathrm{tr}\left(E_{m_A} \otimes E_{m_B} \rho\right).$$

In the last equality, ρ matches the definition of a separable state.

With this discussion, it is also clearer that separable states are *not* uncorrelated. However, they may only exhibit correlations as strong as the ones possible in classical systems, and for this reason are said to be *classically correlated*. Entangled states, conversely, manifest correlations that are not classically reproducible,

which makes it a intrinsically nonclassical property.

Properly justified, the definition of entanglement is quite amicable. It is not, however, computationally friendly, and determining whether a given state ρ can be decomposed as in eq. (1.4) or not can be a daunting task. Even in bipartite structures, the problem is fully solved only under special circumstances, such as for pure states, dimensionally limited Hilbert spaces, or for some special families of quantum states, including Werner states and isotropic states. These will become important in due time, so we discuss them now.

Schmidt number

For bipartite pure states of any dimension, the problem can be fully solved through the so-called Schmidt decomposition. The Schmidt decomposition theorem states that any $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ can be decomposed as

$$|\psi\rangle = \sum_{i=1}^{d} \eta_i \, |i_A\rangle \otimes |i_B\rangle \,, \tag{1.5}$$

where $\{|i_A\rangle\}_{i=1}^{d_A}$ and $\{|i_B\rangle\}_{i=1}^{d_B}$ are orthonormal bases for $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, $\eta_i \geq 0$, and $d = \min\{d_A, d_B\}$. We call η_i *Schmidt coefficients*, denote by $r_S(\psi)$ the number of non-zero coefficients (*Schmidt rank*), and say $\{\eta_i(\psi)\}_{i=1}^{r_S(\psi)}$ is the *Schmidt spectrum*.

Contrasted to eq. (1.2), this is a remarkable simplification. For one, this representation requires a single sum, but also, d is the *minimum* local dimension, irrespective of how (finitely) large the other dimension may be. It also resembles the definition of separability versus entanglement for pure states. Actually, it can be shown that a pure bipartite state $|\psi\rangle$ is entangled if and only if its Schmidt rank $r_S(\psi)$ is larger than one; equivalently, if the Schmidt decomposition has more than one term, or if any $\eta_i = 1$, because $\sum_i \eta_i^2 = 1$.

This observation urges us to ask: are some states *more entangled* than others, and can a Schmidt "something" be used to measure this? Attempting to adequately discuss *entanglement measures* would be going too far. However, as some introductory basic concepts will turn useful, we discuss them with no intention on reproducing the thoroughness that can be found in [11, 12, 4, 13].

The first important thing is that there is a whole zoo of entanglement measures, such as concurrence, negativity, entanglement of formation, etc. The second is that they do not always agree with each other on the ordering they impose on the set of entangled states. Thus, depending on the intended use, there may be some measure more adequate than another. Nevertheless, there are several desirable properties for an entanglement measure $E : \mathcal{D}(d) \mapsto \mathbb{R}$ to satisfy, such as $E(\rho) = 0$ if ρ is separable, and that it should not increase under local operations with classical communication (LOCC); a condition reminiscent of, but actually weaker than the operational definition of separability given above.

Making the Schmidt spectrum respect the first condition is easy (just subtract 1). Using it to impose a partial ordering on the set of pure states requires some extra caution, but it can be done through majorization. We first order the *Schmidt spectrum* $\{\eta_i(\psi)\}_{i=1}^{r_S(\psi)}$ of some state $|\psi\rangle$ in non-increasing order, then define

$$\psi \prec \varphi \quad \Longleftrightarrow \quad \sum_{i=1}^{r} \eta_i^2(\psi) \leq \sum_{i=1}^{r} \eta_i^2(\varphi), \ \forall r.$$

If $\psi \prec \varphi$, we say that ψ is majorized by φ, or that φ majorizes ψ. In relation to entanglement, ψ would then be *more* entangled than $|\varphi\rangle$, in the sense that we may convert $|\psi\rangle$ to $|\varphi\rangle$ solely by means of LOCC [14, 12]. With this in mind, states for which $\eta_i = 1/\sqrt{d}$ are said to be *maximally entangled*.

Although the Schmidt decomposition only works for pure states, the idea of the Schmidt rank can be nicely generalized to an entanglement measure over mixed states. The so called *Schmidt number* [15] is given by

$$r_S(\rho) = \min_{\{|\psi_i\rangle\}_i} \{\max_i [r_S(\psi_i)]\}, \tag{1.6}$$

where I reuse the notation r_S from the Schmidt rank because the two notions are equivalent for pure states.

Arguably opaque, this definition is better understood through a procedure. Starting from ρ, we find an ensemble of pure states $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ for it, list the $r_S(\psi_i)$ for each state of the ensemble, and take the maximum (this constitutes the inner maximization). However, the decomposition we choose for ρ is not, in general, unique. So we do this procedure for all possible sets of pure states $\{\psi_i\}_i$ that may be used to build ρ, and take the minimum element of the resultant set, which is what the outer minimization means. Denoting as S_k the set of density operator with Schmidt number less than or equal to k, it will be of most importance for us that $S_{k-1} \subset S_k$, that S_1 is the set of separable states, that each S_k is a convex set, and that its extremal points are the pure states.

PPT criterion

Going back to the problem of determining whether a given ρ is entangled, whenever we limit the dimensions as $\mathcal{H} = \mathcal{H}^2 \otimes \mathcal{H}^2$ or $\mathcal{H} = \mathcal{H}^2 \otimes \mathcal{H}^3$, the *positive partial transpose* (PPT, also "Peres-Horodecki") criterion provides a necessary and sufficient condition. In all other cases, the condition is still sufficient, though not necessary. To understand the sufficiency affirmation, we recall that the *partial transpose* is a transposition operation acting only on some subsystems. Reusing the decomposition in eq. (1.3), the partial transpose over B is defined as

$$O^{\intercal_B} = (\mathbf{1}_A \otimes T)\, O = \sum_{ij\alpha\beta} O_{ij\alpha\beta}\, |\psi_i\rangle\langle\psi_j| \otimes |\varphi_\beta\rangle\langle\varphi_\alpha| = \sum_{ij\alpha\beta} O_{ij\beta\alpha}\, |\psi_i\rangle\langle\psi_j| \otimes |\varphi_\alpha\rangle\langle\varphi_\beta|\,,$$

where T stands for the transposition map. Now suppose that we take a separable state ρ and transpose, for instance, its second subsystem (the argument is equivalent for transposing the other). Then $\rho^{\intercal_B} = \sum_i p_i \rho_i^A \otimes \left(\rho_i^B\right)^\intercal$. Building on the fact that ρ_i^B was a valid density operator, and that the transpose preserves its trace and eigenvalues, it follows that $\left(\rho_i^B\right)^\intercal$ is also a density operator. With ρ_i^A left unchanged, this implies that $\rho^{\intercal_B} \succeq 0$. Consequently, all separable states have positive partial transpose, which also implies that if the partial transpose of some ρ has negative eigenvalues, it must be entangled.

Shortly after Peres made this argument [16], Horodecki et al. showed that for $d_A d_B \leq 6$ the PPT criterion is actually necessary *and* sufficient: no entangled states in these factorizations have positive partial transpose [17]. For larger dimensions, though, they also prove this is not always true; except under special circumstances.

Werner states

One such special case is that of *Werner states*. They were central to the first proof that entanglement and Bell nonlocality are not equivalent concepts [10]. More specifically, it was shown that a large set of entangled Werner states are nevertheless

local under projective measurements. Later, the bound for locality in $d = 2$ was improved several times [18, 19, 20], the result was extended to POVMs [21], and they were also made pivotal in the study of quantum steering [22] and as a test bed for the capabilities of many quantum informational protocols, such as (semi)device-independent entanglement witnesses. What makes them especially tractable is that they are highly symmetric, as Werner states are bipartite states in $\mathcal{H}^d \otimes \mathcal{H}^d$ for which $(U \otimes U)\, \rho\, (U^\dagger \otimes U^\dagger) = \rho$. It can be shown that this is a one-parameter family of states, and that they may be written as

$$W_d(\alpha) = \left(\frac{d-1+\alpha}{d-1}\right) \frac{\mathbf{1}}{d^2} - \left(\frac{\alpha}{d-1}\right) \frac{S}{d}. \tag{1.7}$$

Here, $S = \sum_{i,j=0}^{d-1} |ij\rangle\langle ji|$ is the swap operator. When written in this form, $\alpha = 0$ stands for the maximally mixed state, and $\alpha \leq 1$. Werner states are entangled if and only if $\alpha > \frac{1}{d+1}$ but, under projective measurements, they are unsteerable if and only if $\alpha \leq 1 - \frac{1}{d}$ [22].

Isotropic states

A second family of states that will become useful in due time are the *isotropic states*. Bipartite and also highly symmetric, they are defined as states in $\mathcal{H}^d \otimes \mathcal{H}^d$ for which $(U \otimes U^*)\, \rho\, (U^\dagger \otimes U^{*\dagger}) = \rho$. They were originally constructed to aid in proofs of entanglement distillability criteria [23]. Later, together with Werner states, they were used to show that entanglement, EPR steering and Bell nonlocality form a strict hierarchy [22, 24], and they have likewise been useful in a multitude of benchmarks. They can be described through a single real, linear parameter α by

$$\chi(\alpha) = (1-\alpha)\, \frac{\mathbf{1}}{d^2} + \alpha \left|\Phi^+\right\rangle\!\left\langle\Phi^+\right|, \tag{1.8}$$

with $|\Phi^+\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |ii\rangle$, a maximally entangled state. For $d = 2$, they are identical to Werner states up to local unitaries, but this is not true for larger dimensions. They are also nonseparable if and only if $\alpha > \frac{1}{d+1}$, and are unsteerable under projective measurements if and only if $\alpha \leq \frac{H_d - 1}{d-1}$, where $H_d = \sum_{n=1}^{d} 1/n$ [22]. Letting α run in $[0, 1]$, the isotropic state $\chi(0)$ is the maximally mixed state, and for $\chi(1)$ we have a maximally entangled one.

Singlet fraction

Maximally entangled states are resources for many informational protocols, and the performance of some protocols can be characterized through the *singlet fraction*, which measures the maximal overlap of a resource ρ with a maximally entangled state. Starting from $|\Phi^+\rangle$, all other maximally entangled states $|\Phi\rangle$ can be reached through local unitaries alone, $|\Phi\rangle = (U_A \otimes U_B)\, |\Phi^+\rangle$, thus the singlet fraction is determined by

$$\zeta(\rho) = \max_{\Phi} \langle\Phi|\rho|\Phi\rangle\,.$$

In particular, the singlet fraction of the isotropic states is

$$\zeta\left[\chi(\alpha)\right] = \alpha + \frac{1-\alpha}{d^2}. \tag{1.9}$$

1.2 Transformations

We now know how to describe static physical systems, which is something that physical systems rarely are. In introductory quantum theory, we learn that the

evolution of closed quantum systems is governed by the Schrödinger equation. Its solution dictates that an initial state ρ that transforms to ρ' does so unitarily following $\rho' = U\rho U^\dagger$. Here, U must be an unitary operator, which means that $UU^\dagger = U^\dagger U = \mathbf{1}$. Given U, we can always find a Hamiltonian H and a suitable interaction time to perform the evolution. Nonetheless, these are not the most general transformations that a quantum state can undergo.

CPTP maps

Taking the operational approach, we will define the most general transformation as any one that takes a density operator into another, i.e., any $\mathcal{N} : \mathcal{D}(\mathcal{H}) \mapsto \mathcal{D}(\mathcal{H}')$, and look to what properties this implies. First of all, $\mathcal{N}$ must be linear, and the reasoning for that goes as usual: if we mix $\rho, \sigma \in \mathcal{H}$ as $w\rho + (1-w)\sigma$ then put it through $\mathcal{N}$, we surely expect the resulting state to be equivalent to independently passing ρ and σ through $\mathcal{N}$ and then mixing. Equation-wise, this means that $\mathcal{N}[w\rho + (1-w)\sigma] = w\mathcal{N}(\rho) + (1-w)\mathcal{N}(\sigma)$. Additionally, it is easy to argue that $\mathcal{N}$ must be such that, for any $\rho \in \mathcal{D}(\mathcal{H})$, it is trace-preserving, $\mathrm{tr}[\mathcal{N}(\rho)] = 1$, and positive, $\mathcal{N}(\rho) \succeq 0$.

Positivity, however, is not a strong enough condition. Suppose that we add an auxiliary system $\mathcal{H}_B$ of arbitrary dimension to $\mathcal{H}$, thus $\mathcal{H} \mapsto \mathcal{H} \otimes \mathcal{H}_B$, and that $\mathcal{N} \equiv \mathcal{N}_{\mathcal{H}\to\mathcal{H}'} \otimes \mathbf{1}_B$. In some situations like this, a positive $\mathcal{N}_{\mathcal{H}\to\mathcal{H}'}$ does not guarantee that $\mathcal{N}$ will map all input states to positive operators. This is precisely the case of the transposition map previously discussed, where the fact the the partial transposition may generate non-positive operators is used as an entanglement criterion.

Amending this requires the stronger condition of *complete positivity* (CP), whereby any CP map $\mathcal{N}$ generates a valid density operator no matter what the dimension of $\mathcal{H}_B$ is.

Together, complete positivity and trace preservation are the conditions that define a CPTP map, or *quantum channel* $\mathcal{N}$, which is the most general type of quantum evolution we will consider.

Kraus representation

These requirements are easily agreeable, but they are not very practical. What we must now do is find a computation-friendly representation for CPTP maps. This can be found in the Kraus representation theorem [25], stating that general map $\mathcal{N} : \mathcal{L}(\mathcal{H}) \mapsto \mathcal{L}(\mathcal{H}')$ is CPTP (i.e., a quantum channel) if and only if it has a decomposition

$$\mathcal{N}(O) = \sum_{i=1}^{D} K_i O K_i^\dagger,$$

with $O \in \mathcal{L}(\mathcal{H})$, all $K_i : \mathcal{L}(\mathcal{H}) \mapsto \mathcal{L}(\mathcal{H}')$, and $\sum_{i=1}^{d} = K_i^\dagger K_i = \mathbf{1}_{\mathcal{H}}$. The limit of the sum, D, will be at least 1 (in this case we recover an unitary evolution), and will never need to be larger than $\dim(\mathcal{H})\dim(\mathcal{H}')$.

Choi representation

The Kraus representation is just one of several other convenient representations for CPTP maps [26], of which we will need the one called *Choi-matrix representation*. It comes from an application of the *Choi-Jamiołkowski isomorphism* [27, 28]. This remarkable result states that any quantum channel $\mathcal{N} : \mathcal{L}(\mathcal{H}) \mapsto \mathcal{L}(\mathcal{H}')$ can be uniquely mapped to a bipartite state

$$\rho_{\mathcal{N}} = (\mathbf{1}_{\mathcal{H}} \otimes \mathcal{N})\, \rho_{\Phi^+} \in \mathcal{D}(\mathcal{H}) \otimes \mathcal{D}(\mathcal{H}), \tag{1.10}$$

where $\rho_{\Phi^+} = \sum_{i,j=1}^{d_{\mathcal{H}}} |ii\rangle\langle jj|$ is the density operator of the previously introduced

maximally entangled state. Taking a closer look, this is actually saying that the space of CPTP maps to the space of bipartite quantum states, and the way to do it is to apply the $\mathcal{N}$ in question to half of that maximally entangled state, while doing nothing to the second part. To finish the isomorphism, we must also know how to go back, which is done through

$$\mathcal{N}_\rho = \mathrm{tr}_{\mathcal{H}} \left[\rho_{\mathcal{N}} \left(\rho^{\intercal} \otimes \mathbf{1}_{\mathcal{H}'} \right) \right]. \tag{1.11}$$

This inverse mapping shows us that we can also take a bipartite state and turn it into a quantum channel. Together, eqs. (1.10) and (1.11) consist in the so-called *channel-state duality*. The representation of quantum channels as bipartite states will become handy when dealing with optimization problems over channels, where a CPTP constraint can be cast as a positivity condition (see secs. 2.3 and 5.1).

1.3 Measurements

To finish our review of quantum theory, we must remember how to measure a state. While in classical mechanics we usually gloss over the concept of measurements, taking for granted that any physical property of a state is trivially accessible, in quantum theory we must not. On a par with entanglement, measurement incompatibility is a concept at the heart of the most interesting quantum phenomena, especially those with no classical counterpart. It is unavoidably linked to some of the most intriguing consequences of quantum theory. Decision problems on the Einstein-Podolsky-Rosen steering scenario [29, 30], for one, can be one-to-one mapped to joint-measurability problems, in the sense that a measurement set is incompatible if and only if it can be used to demonstrate steering. In the also widely studied Bell nonlocality scenario [31], generating nonlocal statistic require incompatible measurements, but not every set of incompatible measurements is sufficient to observe nonlocality [32, 33, 34]. In this section I review the main mathematical structures related to quantum measurements, together with some aspects of measurement incompatibility.

* * *

Any property of a quantum system ρ must be assessed by measuring it. Abstractly, a measurement procedure takes a quantum input (i.e., the state) and returns a classical output (the measurement result). Inside certain constraints, which measurement is chosen for a given application depends on the property to be measured (e.g., the $\mathbf{z}$ component of a spin 1/2 particle), and the number and values of the possible outcomes are associated to the choice of measurement (e.g., either $\pm\hbar/2$ for the spin 1/2 measurement). A quantum measurement procedure is inherently probabilistic: quantum theory goes only so far as telling us how to ascribe probabilities to each possible outcome. Continuing with the spin example, this means that the quantum formalism will only tell us the probability of getting either the $\pm\hbar/2$ result. When the measurement is actually performed, we may end up with any outcome predicted to have non-zero probability of happening. Depending on this outcome (or rather, on our knowledge of it), the very state ρ that was measured is changed in a non-reversible way. This non-reversibility implies

that we cannot fully access ρ with a single copy of the system, a fact that is at the heart of several quantum informational protocols. Although the century-long debate on the nature of quantum measurements is certainly interesting, we will take the pragmatic route and focus on its operational definition.

POVMs

A positive semidefinite (PSD) operator is a Hermitian operator $E : \mathcal{H} \mapsto \mathcal{H}$ such that $\langle\psi|E|\psi\rangle \geq 0,\ \forall\,|\psi\rangle \in \mathcal{H}$. In this case, we denote $E \succeq 0$. We require a quantum measurement to be described by a set $M = \{E_m\}_m$ of PSD operators obeying the completeness relation $\sum_m E_m = \mathbf{1}$. The conditions $E_m \succeq 0$ and $\sum_m E_m$ can be interpreted as enforcing that $p(m) \geq 0,\ \forall m$ and $\sum_m p(m) = 1$, for any possible ρ. Any such set M is called a *positive operator-valued measure* (or POVM; also called *unsharp measurement*), and each of its elements a *measurement effect*. The possible outcomes are labeled by m. Whenever a measurement M is performed on a state ρ, we get result m with probability $p(m) = \mathrm{tr}\,(E_m\rho)$. When $\rho = |\psi\rangle\langle\psi|$, this definition recovers the Born rule in its usual form, $p(m) = \langle\psi|E_m|\psi\rangle$.

Projective measurements

A special case of POVMs arise when every $E_m = \Pi_m$, and $\Pi_m\Pi_n = \Pi_m\delta_{mn}$, where the Π_m are projection operators. It then follows that $1 \leq |M| \leq d$. This is the case of *projective measurements* (or PVMs; commonly also called *ideal*, or *sharp*, or *von Neumann* measurements). Quantum mechanics courses usually introduce projective measurements through the concept of *observables*, which are Hermitian operators. Recalling these can be decomposed as $A = \sum_m m\Pi_m$, we can see they define a PVM where the possible outcomes are the eigenvalues m of observable A, and the projection operators are a set of orthonormal eigenvectors.

A further restriction on measurements is sometimes put on the rank of each effect. A rank-1 projective measurement happens when M is projective and $|M| = d$ or, equivalently, when the associated observable A has no degenerate eigenvalues.

An useful intuition on POVMs is to interpret them as noisy projective measurements. Consider a sharp measurement $M' = \{|0\rangle\langle 0|\,, |1\rangle\langle 1|\}$. Performed on an arbitrary state ρ, we will get the result associated to the the i-th projection with probability $p(i) = \mathrm{tr}\,(|i\rangle\langle i|\,\rho)$, where $i \in \{0, 1\}$. Suppose, though, that our experimental apparatus is such that it states the wrong result with probability $(1 - w)$. Then $p(i) = w\,|i\rangle\langle i| + (1 - w)\,|i \oplus 1\rangle\langle i \oplus 1|$, and a measurement describing this situation is $M = \{w\,|0\rangle\langle 0| + (1 - w)\,|1\rangle\langle 1|\,, w\,|1\rangle\langle 1| + (1 - w)\,|0\rangle\langle 0|\}$. It is easy to see that M's effects are not orthogonal, hence this is not a projective measurement. But it is a valid unsharp measurement.

Post-measurement state

After a measurement is performed, we may be interested in what happened to ρ. As already stated, performing a measurement will generally incur in a mapping $\rho \mapsto \rho'$. Lüders rule states that, for a PVM $\{\Pi_m\}_m$ returning result m,

$$\rho' = \frac{\Pi_m\rho\Pi_m}{\mathrm{tr}\,(\rho\Pi_m)}. \tag{1.12}$$

Two interesting facts are that knowing the updated state requires knowledge of the measurement result, and that retaking the same measurement will produce the same outcome with definiteness. Without access to the measurement result, but knowing that $\{\Pi_i\}_i$ was performed, all we can say is that $\rho' = \sum_m \Pi_m\rho\Pi_m$. POVMs results are not always reproducible, and the update rule only exists when all effects can be written as $E_m = K_m^\dagger K_m$, where the K_m are Kraus operators. A

post-measurement state is then given by

$$\rho' = \frac{K_m \rho K_m^\dagger}{\mathrm{tr}\,(\rho E_m)}.$$

Joint-measurability

One of the most intriguing aspects of the quantum measurement formalism is that it implies the existence of quantities that may not be simultaneously measured with arbitrary precision on a single copy of a quantum system. Measurements behaving in this way are called *incompatible measurements*. Throughout this work, we will understand "compatibility" as a synonym of "joint measurability". A set $\mathcal{M} = \{E_m^x\}_{m,x}$ of X quantum measurements indexed by x with $\mathcal{O}(x)$ outcomes each is said to be *jointly measurable* whenever a parent measurement J_ℓ exists. To be a parent measurement, J_ℓ must be a valid quantum measurement from which any $E_m^x \in \mathcal{M}$ may be recovered. Letting $\ell = \ell_1 \ell_2 \ldots \ell_X$, where each $\ell_i \in \{1, \ldots, \mathcal{O}(i)\}$, the latter condition requires that $\sum_\ell J_\ell \delta_{m,\ell_x} = E_m^x$. We interpret this as saying that the measurement statistics that the set $\mathcal{M}$ could generate can be reproduced by applying the single measurement J_ℓ then coarse-graining the results. In the plenty of situations where no parent measurement exist, $\mathcal{M}$ is called incompatible or non-jointly measurable. We will get back to this topic in sec. 2.3.1, where we discuss how to quantify how incompatible a measurement set is.

Several other notions of measurement compatibility exist. Introductory quantum mechanics courses, for instance, usually discuss incompatibility as commutativity: two non-commuting observables A and B can only be simultaneously determined up to a degree of certainty, and a trade-off between the standard deviations on the obtained results is given by Robertson's uncertainty relation $\sigma_A \sigma_B \geq \frac{1}{2}|\langle [A, B] \rangle|$. Notably, in this particular case of sharp measurements, commutativity is equivalent to joint-measurability. Regarding POVMs, though, this is not the case, as with several other inequivalent definitions of incompatibility, such as non-disturbance and coexistence [35].

Simulatability

From now on, let us call $\mathcal{P}(d, n)$ and $\mathbb{P}(d, n)$ the set of POVMs and projective measurements, respectively, with n effects acting on $\mathcal{H}^d$. It is clear that $\mathbb{P}(d, n) \subset \mathcal{P}(d, n)$ for any d and n. Now suppose we want to realize an experiment in which we must perform some $M \in \mathcal{P}(d, n)$ but we only have access to a subset $\mathcal{M} \subset \mathcal{P}(d, n)$ where $M \notin \mathcal{M}$. Could we, somehow, reproduce the results we would obtain with M by only using $\mathcal{M}$ and classical processing? In many cases, the answer is *yes* [36, 37, 38].

In chap. 4 we will be mostly interested in when the some set of POVMs can be simulated using only projective measurements. A good starting point is to notice that the trivial measurement $M = \{\mathbf{1}_d\}_{i=1}^n \subset \mathcal{P}(d, n)$ can be simulated solely by means of classical post-processing. We can simply sample a result from the uniform distribution on $\{1, \ldots, n\}$. Defining the *depolarizing map* as

$$\Phi_t(O) \mapsto tO + (1 - t)\frac{\mathrm{tr}\,(O)}{d}\mathbf{1}, \tag{1.13}$$

and its action on a measurement as $\Phi_t(M) \equiv \{\Phi_t(E_m)\}_m$, we can see that any $M \in \mathcal{P}(d, n)$, when fully depolarized, is simulatable with classical randomness, hence with projective measurements. The question that now arises is whether we need to go all the way through, or if there is some $t > 0$ that suffices. Astonishingly,

a non-trivial, general lower bound of $t = 1/d$ exists, and it turns the whole set $\mathcal{P}(d, n)$, for any number n of effects, simulatable with projective measurements [39]. This bound can be improved for specific cases using optimization techniques, and it is known that, for qubits, $t = \sqrt{2/3} - \epsilon$, for some small ϵ, suffices [36].

1.4 Gell-Mann operators and Bloch vectors

Before ending this chapter, we must discuss an interesting representation for quantum states (and some measurement operators) in terms of vectors instead of operators, which will be central to our discussion in chap. 4.

In sec. 1.1 we argued that any quantum state is a linear, positive semi-definite, unit-trace d-dimensional density operator ρ, and that any such ρ is a quantum state. Thus we defined the set of density operators as

$$\mathcal{D}(d) = \{\rho \in \mathcal{L}(\mathcal{H}^d) \mid \mathrm{tr}\rho = 1, \rho \succeq 0\},$$

where $\mathcal{L}(\mathcal{H}^d)$ is the set of linear operators in $\mathcal{H}^d$. This definition implies that $\rho^\dagger = \rho$ and $\mathrm{tr}(\rho^2) \leq 1$, with equality only holding for pure states.

When describing a quantum state, it is often convenient to choose some basis for $\mathcal{H}^d$ and define ρ through a vector. A widely used choice of basis in $\mathcal{H}^2$ are the Pauli matrices $\{\sigma_x, \sigma_y, \sigma_z\}$. They are Hermitian, traceless operators that, together with the identity, form a basis in which any 2×2 Hermitian matrix can be written. Letting $v = (v_x, v_y, v_z) \in \mathbb{R}^3$ and $\sigma = (\sigma_x, \sigma_y, \sigma_z)$, it's easy to see that

$$\rho = \frac{\mathbf{1}_2}{2} + v \cdot \sigma$$

is any unit-trace Hermitian operator. Restricting to $\rho \succeq 0$ further requires that $|v| \leq 1$. Thus, any element of the ball $\mathcal{B}(2) = \{x \in \mathbb{R}^3 \mid |x| \leq 1\}$ is a 2-dimensional quantum state and, conversely, any 2-dimensional quantum state can be associated to a 3-dimensional real vector v with $|v| \leq 1$. Also importantly, $|v| = 1$ if and only if the state is pure, as can be verified by constraining $\mathrm{tr}(\rho^2) = 1$. We'll call $\mathcal{B}(2)$ the *Bloch ball*, its surface the *Bloch sphere*, and any $v \in \mathcal{B}(2)$ a *Bloch vector*. This geometric interpretation of qubit states will prove to be incredibly convenient.

Generalizing these concepts to $\mathcal{H}^d$ can be done in a similar fashion, but will lead us to an important caveat. Our starting point will be picking a basis. While there are a bunch of useful choices [40], the most adequate for our future endeavors will be the generalized Gell-Mann matrices. The (standard) Gell-Mann matrices naturally extend the Pauli matrices from SU(2) to SU(3), and they are likewise traceless and Hermitian. As these are the most important properties we will want to preserve in our applications, it is sensible to choose the standard SU(d) generators when moving on to $\mathcal{H}^d$. Apart from the identity operator, we will need another $d^2 - 1$ operators to span $\mathcal{H}^d$. These generalized Gell-Mann matrices can be conveniently

written as $\sigma = \{\sigma_{jk}^{(s)}, \sigma_{jk}^{(a)}, \sigma_l^{(d)}\}$, where

$$\begin{aligned}
\sigma_{jk}^{(s)} &= |j\rangle\langle k| + |k\rangle\langle j|, && 1 \leq j < k \leq d, \\
\sigma_{jk}^{(a)} &= -i|j\rangle\langle k| + i|k\rangle\langle j|, && 1 \leq j < k \leq d, \\
\sigma_l^{(d)} &= \sqrt{\frac{2}{l(l+1)}} \left(\sum_{j=1}^{l} |j\rangle\langle j| - l|l+1\rangle\langle l+1| \right), && 1 \leq l \leq d-1.
\end{aligned}$$

Representing an element of $\mathcal{H}^d$ asks for d^2 coefficients. However, because we fix $\mathrm{tr}(\rho) = 1$, there's a redundancy that will leave us with only $d^2 - 1$ free parameters. Preserving the previous notation we then propose that

$$\rho = \frac{\mathbf{1}_d}{d} + v \cdot \sigma = \frac{\mathbf{1}_d}{d} + \sum_{i=1}^{d^2-1} v_i \sigma_i, \tag{1.14}$$

where with the summation we made it explicit that now $v \in \mathbb{R}^{d^2-1}$, and by σ_i we mean the elements of σ defined above. This is, again, clearly a unit-trace and Hermitian operator. The condition $\mathrm{tr}(\rho^2) \leq 1$ further imposes that $|v| \leq \sqrt{\frac{d-1}{2d}}$. Our next step would be to characterize $\mathcal{B}(d)$ and consequently the (generalized) Bloch vectors. Unfortunately, finding a structure on v that guarantees that $\rho \succeq 0$ is not as straightforward as in the 2-dimensional case. Naively trying to associate any vector in a d-dimensional ball with a density operator won't take us far, as we'll soon find out that many choices lead to operators having negative eigenvalues. Albeit the complete characterization of Bloch vectors for arbitrary dimensions has been done [41], it does not lead to a natural parameterization such as in the 2-dimensional case. We will keep calling any $\mathcal{B}(d)$ a Bloch "ball", but it is important to keep in mind that are not balls at all, for there will in general be holes where there are no allowed Bloch vectors.

Even being more complex than for qubits, this geometric interpretation is still fruitful. For one, its inverse map has quite an intuitive interpretation inside of quantum theory. We have so far been interested in defining which Bloch vectors lead to a density operator, but haven't mentioned the converse problem — the one of determining v from ρ — at all. The answer, which can be straightforwardly verified, is that $v_i = \mathrm{tr}(\sigma_i \rho)$. Hence, each component of v is the expectation value of an observable σ_i that can, at least in principle, be experimentally obtained. Luckily, this is the mapping we will usually worry about.

Another useful consequence of this geometric description is that the effect of operations on a state can be interpreted visually. An specific example that will later become important is that of depolarizing channels. These channels are worst-case scenario noise models describing the situation where information on a state ρ is, with some probability p, completely lost. Thus $\rho \mapsto (1-p)\rho + p\frac{\mathbf{1}_d}{d}$, where $\mathbf{1}_d/d$ is the maximally mixed state (the noise). Letting v_ρ be the Bloch vector associated to ρ, and observing that $v = 0$ defines the maximally mixed state, we are led to the conclusion that a depolarizing channel shrinks ρ's Bloch vector towards the origin of the Bloch sphere.

Considering that measurement operators share similarities with density matrices, we may attempt to extend this representation. Our first observation is that not all measurement effects can be written as in eq. 1.14, because they are not required

to have unit-trace. However, as any measurement effect is a positive-semidefinite operator, those of which do have unit-trace can also be described through Bloch vectors. As already pointed out, projective measurements are an important class of measurements where every effect is a projection, which is furthermore orthogonal to all the others. For any projection Π, it is true that $\mathrm{tr}(\Pi) = \mathrm{rank}(\Pi)$. Therefore, all rank-1 projective measurements can be described by a set of Bloch vectors, each associated to one of its effects. As all projections with larger rank can be simulated by rank-1 projections and coarse-graining, we conclude that all projective measurements can be interpreted through Bloch vectors together with post-processing. Because, by definition, $\Pi^2 = \Pi$, then $\mathrm{tr}(\Pi^2) = \mathrm{tr}(\Pi) = 1$, which means that a rank-1 projection operator's Bloch vector is actually on the Bloch sphere, analogously to pure states. Finally, as a measurement's effects must sum to the identity, a nice interpretation of projective measurements on qubits arises: given one of the effect's Bloch vector, the other must be its antipodal.

* * *

Quantum information is the study of the encoding, processing and decoding of information in quantum systems, which are done by preparing a state ρ, transforming, and measuring it. In this chapter we learned that quantum theory tells us we should associate these steps to density operators, CPTP maps, and complete collections of positive-semidefinite operators, respectively. In many cases, we will be interested in quantities that are very nicely described by these structures, but that are nevertheless not easily computable. But luckily, several well-known and efficient optimization techniques are quite well suited to dealing with quantum theory. This is the topic we will review in the next chapter.

Chapter 2

Mathematical optimization

2.1 Convexity

2.2 Linear programming

2.3 Semidefinite programming

2.3.1 Optimizing the measurement incompatibility robustness

Part II

Prepare and measure

Chapter 3

The prepare-and-measure scenario

The prepare-and-measure (PM) scenario is one of the simplest and most fundamental examples of correlation scenarios. In it, a preparation apparatus produces and then sends a physical system, over a communication channel, to a measurement device which reads out information from the received state. Wherefore, it is an adequate setting in which to investigate two of the most fundamental building blocks of physical theories: states and measurements.

Differently from the more widely studied Bell nonlocality and EPR steering scenarios, a quantum prepare-and-measure experiment may behave nonclassically even in the absence of entanglement. Quantum behaviors in PM scenarios must then rely on other strictly quantum features, such as measurement incompatibility [42] and non-orthogonality of states [43], but the exact relations are still unknown.

Other than quantum communication, a second resource that also leads to drastically different behaviors is whether the preparation and measurement devices are independent or not. There are, in general, three possible cases; namely, full independence, shared randomness and entanglement assistance. Together with either classical or quantum communication, this will lead to six inequivalent prepare-and-measure configurations.

Prepare-and-measure scenarios are also the simplest correlation scenarios that presume communication, and, as such, should become an indispensable ingredient in quantum networks [44, 45]. As with other correlation scenarios, quantum behaviors in the PM scenario can be exploited to build informational protocols that show advantage over their classical counterparts (sec. 3.2). On a more fundamental aspect, they are at the core of proposed informational principles to quantum theory [46, 47], and of quantum ontologies [48].

In chapters 4 and 5, novel results regarding some of these settings will be presented. To build towards that end, we now discuss these many different instances of preparation and measurement devices, and show how this scenario can be seen as a physical implementation of two paramountly important communication protocols.

Figure 3.1: [To-Do:]

3.1 Prepare-and-measure behaviors

The simplest prepare-and-measure setup consists of two black-boxes. One is the preparation device P, handled by Alice, and the other the measurement device M, handled by Bob. Nothing about the inner workings of these devices is assumed *a priori* (see), except that P prepares and communicates a physical system to M, which extracts information from the received preparation by measuring it.

Alice is allowed to interact with her device through a classical input $x \in \mathcal{X} \equiv \{1, \ldots, X\} \equiv [X]$. Her choice may weight on the probability $p(m \mid x)$ with which a state labeled by $m \in \mathcal{S}$ is prepared. Here, $\mathcal{S}$ represents the set of possible preparations.

Similarly, Bob can choose a configuration $y \in \mathcal{Y} \equiv [Y]$ for the measurement device, which will thence output $b \in \mathcal{B} \equiv [B]$ with probability $p(b \mid m, y)$. This reflects the fact that the outcome of the experiment may be influenced by both the received message and Bob's choice of measurement (fig. 3.1).

A last, indispensable condition is that neither Alice nor Bob know what happens in each other's labs. In other words, she must have no knowledge of y, as this is a choice made in her causal future, and he cannot know what was her choice x of preparation, as this trivializes the scenario in a sense that will soon be made clearer. Taken together, these constraints amount to saying that all communication between them is mediated by the message m. Given that our intention is to study how this communication influence the observed statistics, this is clearly a natural assumption.

An *round* is a single run of the protocol described above. After many such rounds, Alice and Bob are allowed to share their knowledge with each other, and together they can build the *behavior* $\mathbf{p} = \{p(b \mid x, y)\}_{b,x,y}$ of their devices. This behavior characterizes the prepare-and-measure *experiment*. Naturally, each $p(b \mid x, y) \geq 0$ and $\sum_b p(b \mid x, y) = 1$, $\forall x, y$ or, equivalently, $\mathbf{p}$ is a collection of conditional probability distributions, one for any fixed choice of a pair of settings (x, y).

Behavior

We now arrive at the central question of (semi)device-independent quantum information: if all we have is the behavior $\mathbf{p}$, with no (or restricted) access to the actual workings of the devices, can we still certify some property about the states, measurements, or other quantities of interest? For instance, could we, by only observing $\mathbf{p}$, affirm that P prepares quantum — as opposed to classical —, states? Or that M applies nonprojective measurements? In many cases, the answer, surprisingly, is *yes*.

3.1.1 Classical preparations

Quantum preparations may behave quite distinctly from classical preparations, and knowing how to tell them apart is of the essence for developing quantum communication protocols. An important open question regards the properties that allow some sets of preparations and measurements to behave nonclassically. In chapter 4, I will discuss the problem of classical simulatability of quantum behaviors in

reasonable generality, and prove that measurement incompatibility is not a sufficient condition for nonclassicality in the PM scenario. To make that discussion precise, we must begin by defining what it is that we will call classically-simulatable behaviors or, for shortness, *classical behaviors*.

Starting from the paradigmatic prepare-and-measure scenario, let us further impose that $\mathcal{S}$, the set of possible preparations, contains only *classical states* (dits). Naturally, as this is a communication scenario, the dimension $|\mathcal{S}|$ of the classical system used for encoding the states must be bounded, otherwise communication becomes trivial and all behaviors are possible.

Our first aim is to investigate the set of behaviors that can be achieved when communicating d-dimensional classical systems, or rather, when $\mathcal{S} = \{0, \ldots, d\}$. They will, in general, also depend on X (the size of Alice's input alphabet) and Y (the number of choices for Bob's measurement). Letting this set of behaviors be $\mathcal{C}_{d,X,Y}$, our previous discussion implies that

$$\mathbf{p} \in \mathcal{C}_{d,X,Y} \quad \Longleftrightarrow \quad p\left(b \mid x, y\right) = \sum_{m \in \mathcal{S}} p\left(m \mid x\right) p\left(b \mid m, y\right) \quad \forall b, x, y. \tag{3.1}$$

Essentially, then, $\mathcal{C}_{d,X,Y}$ contains all behaviors that may occur when (i) the devices are *uncorrelated*, (ii) the preparations are classical and (iii) with dimension at most d, (iv) Alice has X preparation choices, and (v) Bob picks one out of Y measurement settings. Independent devices

To simplify the notation, whenever X and Y are arbitrary or clear by context, the subscripts will be omitted.

Briefly detouring, it is interesting to notice that model (3.1) is very much in the spirit of ontological models [48, 49, 50]. To see that, consider $\mathcal{S}$ as a finite, dimension-bounded ontic state space, and x as the preparation procedure. Then, $p\left(m \mid x\right)$ models our epistemic state. Similarly, if we take y as a choice of measurement procedure, we may interpret $p\left(b \mid m, y\right)$ as the indicator function. In this way, any theory that only produces behaviors $\mathbf{p} \in \mathcal{C}_{d,X,Y}$ admits a dimension-bounded ontological model.

Conditions (iii)–(v) above are pretty much natural for any communication scenario, but the case is different for (i) and (ii). While in many situations, such as when we have some trust on our devices, (i) is justifiable, it is not always safe to assume that the devices are uncorrelated. The worst-case scenario for classical variables is when P and M can share an unbounded amount of pre-established classical correlations. Such correlations must reside in the causal past of the experiment (fig. 3.2), but even so, they can be used to achieve better performance in several communication protocols implemented in the PM scenario [51], and lead to quite different geometrical structures [52, 53]. Without knowledge on what information they share, the best we can do is to call it λ and say that π is some probability distribution over this random variable. As both devices can fully access $\lambda \in \Lambda$, Classical correlations

$$\mathbf{p} \in \mathcal{C}^{\lambda}_{d,X,Y} \quad \Longleftrightarrow \quad p\left(b \mid x, y\right) = \int_{\Lambda} \sum_{m \in \mathcal{S}} \pi(\lambda) p\left(m \mid x, \lambda\right) p\left(b \mid m, y, \lambda\right) d\lambda \quad \forall b, x, y. \tag{3.2}$$

Here, $\mathcal{C}^{\lambda}_{d,X,Y}$ is the set of behaviors we get from $\mathcal{C}_{d,X,Y}$ by allowing for shared randomness. As $\pi(\lambda) \geq 0$ and $\sum_{\lambda} \pi(\lambda) = 1$, eq. (3.2) is actually telling us that

Figure 3.2: [To-Do:]

$\mathcal{C}^{\lambda}_{d,X,Y} = \text{conv}\left(\mathcal{C}_{d,X,Y}\right)$. More than that, a slight variation on Fine's theorem [54] (or sec. 2.3 of [55] for a more pedagogical discussion) can show that the set $\mathcal{C}_{d,X,Y}$ has a finite amount of extremal points, called deterministic strategies [56, 51]. They are the points given by eq. (3.1) when the response functions are deterministic, i.e., when $p\left(m \mid x\right) = \delta_{m,f(x)}$ and $p\left(b \mid m, y\right) = \delta_{b,g(m,y)}$, for some functions $f : \mathcal{X} \to [d]$ and $g : [d] \times \mathcal{Y} \to \mathcal{B}$ that are made precise in the aforementioned references. As $\mathcal{C}^{\lambda}_{d,X,Y}$ can now be seen as a convex hull of finitely many points, this proves that $\mathcal{C}^{\lambda}_{d,X,Y}$ is a polytope. Recalling the discussion in sec. 2.1, we emphasize that $\mathcal{C}^{\lambda}_{d,X,Y}$ can thus be described by an intersection of half-spaces, which are given by the linear inequalities defining its facets. This description will turn out to be especially useful during chap. 4, in which we will get back to this topic and work out an example that should clarify this discussion.

For an example of the usefulness in understanding these sets, notice that for some fixed X and Y, and some $d' > d$, we have the proper inclusion $\mathcal{C}_d \subset \mathcal{C}_{d'}$, which implies the same for the SR case. Ultimately, this means that larger dimensional communication can carry strictly more information. Now suppose that W_d is a linear functional defining some facet of $\mathcal{C}^{\lambda}_d$ which is not a facet of $\mathcal{C}^{\lambda}_{d'}$ (at least one such W_d must exist, since $\mathcal{C}^{\lambda}_d \subset \mathcal{C}^{\lambda}_{d'}$) and that, for any behavior $\mathbf{p} \in \mathcal{C}^{\lambda}_d$, we have a bound $W_d \cdot \mathbf{p} \leq C_d$. Because W_d does not define a facet of $\mathcal{C}^{\lambda}_{d'}$, there is some $\mathbf{p}' \in \mathcal{C}^{\lambda}_{d'}$ such that $W_d \cdot \mathbf{p}' > C_d$. Hence, if we are given preparation and measurement boxes which are guaranteed to prepare only classical systems, and we observe some behavior that, like $\mathbf{p}'$, violates the bound on W_d, we can certify that our uncharacterized devices is preparing states of dimension at least $d + 1$.

Dimension witnesses

In disguise, I have exemplified what is called a (dimension) witness: any functional and bound that, (i) for some set of behaviors is never violated, but that (ii) can be violated by at least one behavior of some other set, is said to *witness* some property. In our case, we are witnessing dimension in a semi device-independent fashion, as we have assumed the preparations are classical. Lastly, notice that the facets defining any $\mathcal{C}^{\lambda}_d$ polytope are, by definition, dimension witnesses. Furthermore, they are *tight* witnesses, something which not all witnesses must be.

Entanglement assistance

For completeness, I note it is also possible to define a $\mathcal{C}^{\rho}_d$ set where the preparations remain classical, but the devices can be correlated through a shared quantum state ρ [57, 45]. When ρ is an entangled state, it can lead to interesting behaviors associated with advantages in communication protocols [58].

3.1.2 Quantum preparations

To continue generalizing our discussion, let us also remove the assumption of classical preparations.

In that case, $\mathcal{S}$ will be a finite subset of $\mathcal{D}(\mathcal{H})$, and y a choice of quantum measurement. Employing Born's rule, we rewrite the behaviors as $\mathbf{p}_Q = \{\text{tr}\left(\rho_x E_{b|y}\right)\}_{b,x,y}$. Here, all $\rho_x \in \mathcal{S}$, and $\{E_{b|y}\}_b$ is a POVM for each y. For the dimension bound, let $\dim(\mathcal{S}) = \dim \sum_{\rho_x \in \mathcal{S}} \text{supp}\left(\rho_x\right)$ be the smallest Hilbert space dimension needed to represent all density operators in $\mathcal{S}$. Then, in direct analogy to the classical behaviors sets, for any fixed X and Y, we define $\mathcal{Q}_d$ as the set of

Independent devices

behaviors for quantum communication with uncorrelated devices.

Allowing for shared randomness introduces a slight change in the elements of the behaviors. Using the same notation as before, they turn into

Classical correlations

$$p\left(b \mid x, y\right) = \int_{\Lambda} \pi(\lambda) \mathrm{tr}\left(\rho_x^{\lambda} E_{b|y}^{\lambda}\right),$$

making it true that $\mathcal{Q}_d^{\lambda} = \mathrm{conv}\left(\mathcal{Q}_d\right)$.

Lastly, and most importantly for chap. 5, Alice and Bob may share a quantum system $\rho \in \mathcal{H}_A \otimes \mathcal{H}_B$ and use it as a resource to improve their quantum communication. Most generally, Alice can use her share of ρ as an aid in her encoding of x. The way to do it is by applying a local CPTP map Λ_x. As we must bound the communication to d-dimensional systems, $\Lambda_x : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_C)$, where $\mathcal{H}_C \simeq \mathbb{C}^d$. This system in $\mathcal{H}_C$ is then transmitted to Bob, who will afterwards hold $\rho_x = \left(\Lambda_x \otimes \mathbf{1}_B\right)\rho$. His measurements' effects act on $\mathcal{H}_C \otimes \mathcal{H}_B$ (fig. 3.2). Behaviors compatible with experiments implementing this procedure are in the set $\mathcal{Q}_d^{\rho}$.

Entanglement assistance

Sometimes, it is a reasonable assumption to make $\mathcal{H}_A \simeq \mathcal{H}_C$ (see sec. 3.2.2 and chap. 5). Bear in mind, though, that it can lead to some loss of generality: recently, Tavakoli et al. [57] considered a qubit communication protocol where 4-dimensional entanglement shows advantage over having only a 2-dimensional entangled resource. To the best of my knoledge, those, together with Moreno et al.'s ([59] and chap. 5), were the first results considering this most general quantum communication scenario.

As a side note, although I have presented entanglement-assisted quantum communication scenarios as the most general kind of PM scenario, there are still further possible generalizations if we allow Alice's and/or Bob's inputs x and y to also be quantum variables [60]. These are interesting and little explored scenarios, but dealing with them is out of our scope.

Clearly, any of the sets $\mathcal{Q}$, $\mathcal{Q}^{\lambda}$ and $\mathcal{Q}^{\rho}$ is contained in their respective larger-dimensional counterparts. Thence, the same observations made before make it possible to also derive *quantum* dimension witnesses [56]. Whereas witnessing dimension for classical or quantum preparations is already of practical and fundamental interest, we ideally want to also distinguish between classical and strictly quantum behaviors. To that end, it is useful to recognize that a set of pair-wise commuting density operators can only generate classically reproducible statistics [53]. More precisely, if $\mathcal{S}_C$ is a set of X d-dimensional states such that $[\rho_i, \rho_j] = 0$, $\forall \rho_i, \rho_j \in \mathcal{S}_C$, then its behavior under any set of measurements can be mimicked with dits. [**Review!** Not so sure about this justification for "classical states" and whether this is valid for any situation like with SR, without SR etc...]

Nonclassicality witnesses

As not all quantum states commute, it turns out that $\mathcal{C}_d \subset \mathcal{Q}_d$, and similarly for the sets allowing for shared randomness and entanglement assistance. Collecting the aforestated set relationships, we find that

$$\mathcal{C}_d \subset \mathcal{Q}_d \subset \mathcal{Q}_d^{\lambda} \subset \mathcal{Q}_d^{\rho}, \tag{3.3}$$

$$\mathcal{C}_d \subset \mathcal{C}_d^{\lambda} \subset \mathcal{Q}_d^{\lambda} \subset \mathcal{Q}_d^{\rho}, \tag{3.4}$$

$$\mathcal{C}_d \subset \mathcal{C}_d^{\lambda} \subset \mathcal{C}_d^{\rho} \subset \mathcal{Q}_d^{\rho}, \tag{3.5}$$

showing it is thus also possible to build nonclassicality witnesses for the prepare-

and-measure scenario.

* * *

In recent years, a multitude of dimension and nonclassicality witnesses for prepare-and-measure scenarios have been proposed [56, 43, 61, 53, 51, 62, 63] and some of them experimentally tested [64, 65, 66]. More stringent conditions can also be used to perform self-testing of states and measurements, which can be used to certify the use of mutually unbiased bases, nonprojective measurements, or targeted sets of states [67, 68, 69, 70, 71, 72, 59]. All the while, known computational methods can aid in bounding the set of finite dimensional quantum correlations in prepare-and-measure scenarios [73, 74]. The prepare-and-measure scenario and its sequential, or prepare-transform-and-measure variations [75, 76] are also expected to serve as building blocks for quantum networks [44, 77]. Complementing their versatility, they can also be seen as a physical implementation of many paradigmatic communication protocols such as random access coding and dense coding, which I now briefly introduce.

3.2 Prepare-and-measure communication protocols

The preeminent practical interest in prepare-and-measure scenarios is due to the fact that quantum communication can beat classical protocols in many tasks that can me modeled as preparing and measuring states. These include cryptographic key distribution [78], secret sharing [79] and communication complexity scenarios [80]. Simplifiedly, the latter deals with the question of how much is the least amount of communication needed to compute or approximate some function $f(x, y)$ where the inputs x and y are distributed. Prepare-and-measure scenarios clearly resemble this structure. All we have to do is to stipulate a figure of merit that, generally depending on f, measures how well some protocol performs. If we additionally manage to show some separation between the best performance achievable, for instance, in $\mathcal{C}_d^\lambda$ against $\mathcal{Q}_d^\lambda$, we can prove quantum advantage in the task. A special and widely studied choice of $f(x, y) = x_y$ leads us to the so-called *random access coding* (RAC) protocol.

3.2.1 Random access coding

An $(n, d) \mapsto m$ random access code (RAC) is a communication task in which a party — commonly Alice, — is given a n-length $d$itstring $\mathbf{x} = x_1 x_2 \dots x_n$, with each $x_i \in \{0, \dots, d-1\}$, and required to encode it in another ditstring $\mathbf{a}$ of length m, where $m < n$.

The m dits representing $\mathbf{a} = \mathcal{E}(\mathbf{x})$, where $\mathcal{E} : \{0, \dots, d-1\}^n \mapsto \{0, \dots, d-1\}^m$ is an encoder function, are then sent to a second party, whom we'll call Bob. Bob is queried with an $y \in \{1, \dots, n\}$ and must correspondingly guess what was the value of x_y. His guess may be modeled through n decoding functions $\mathcal{D}_y : \{0, \dots, d-1\}^m \mapsto \{0, \dots, d-1\}$ that are chosen depending on the query y. We name *encoding-decoding strategy* the ordered set $(\mathcal{E}, \mathcal{D}_1, \dots, \mathcal{D}_n)$.

When Alice and Bob use the same strategy in all rounds of the protocol, the observed statistics are always deterministic, i.e, the probability of Bob answering

b is given by $p\left(b \mid \mathbf{x}, y\right) = \delta\left[b, \left(\mathcal{D}_y \circ \mathcal{E}\right)(\mathbf{x})\right]$, and in each round of a RAC, Bob's probability of guessing the right value is given by $p(b = x_y \mid \mathbf{x}, y)$.

She and he must cooperate to guess as best as they can. Their performance is typically measured through the *worst-case success probability* p_{worst}, defined as the minimum guess probability $p(b = x_y \mid \mathbf{x}, y)$ occurring for their particular encoding-decoding strategy. When the best possible strategy for an $(n, d) \mapsto m$ scenario is such that $p_{\text{worst}} \leq 0.5$, the RAC is said to not exist, as in that case a better or equivalent performance could be achieved through independently guessing.

Shared randomness is known to improve performance in this task. This is done by allowing the encoding and decoding functions to be correlated by some pre-established random variable λ. For a concrete example, consider the simplest $(2, 2) \mapsto 1$ scenario. When no SR is allowed, $p_{\text{worst}} \leq 0.5$ [81]. With SR, however, Alice and Bob can cooperate in the following way. Before each round, a random variable $\lambda \in \{0, 1\}$ instructs Alice to send $\mathbf{a} = x_\lambda$. Bob knows λ, so if he's queried with $y = \lambda$, he can answer correctly with certainty; otherwise he can just flip a coin. Of course, we still have $p_{\text{worst}} \leq 0.5$ in this situation. However, it is known from [82] that when SR is allowed, the best possible strategy is such that **[Review! Confirmar se é isso mesmo]**

$$p_{\text{worst}} = p_{\text{avg}} \equiv \frac{1}{nd^n} \sum_{\mathbf{x}, y} p\left(b = x_y \mid \mathbf{x}, y\right) \tag{3.6}$$

for an uniform distribution on y, proving that a $(2, 2) \mapsto 1$ SR-RAC with $p_{\text{worst}} = 0.75$ exists.

Something else that can further improve performance is allowing quantum communication. In this case, we interpret the encoders $\mathcal{E}(\mathbf{x}) = \rho_x \in \mathcal{D}(d)$ as preparation procedures and the decoders $\mathcal{D}_y = \{E_{b|y}\}_b$ as quantum measurements.

Quantum random access codes (QRACs) first appeared in [83] and were later rediscovered and linked to quantum automata in [81]. With this popularization, multiple new results and experimental demonstrations regarding the existence and advantage of QRACs over their classical counterparts rapidly ensued [84, 85, 86, 82, 87, 78].

Mapping QRACs to prepare-and-measure scenarios with quantum communication amounts to choosing $|\mathcal{X}| = d^n$ preparations $\rho_x \in \mathcal{D}(d)$, and $|\mathcal{Y}| = n$ choices of POVMs with $|\mathcal{B}| = d$ outcomes each [51]. The set of behaviors is then $\mathcal{Q}_{d,d^n,n}$, and optimizing the protocol means looking for a $\mathbf{p} \in \mathcal{Q}_{d,d^n,n}$ that maximizes either p_{worst} or p_{avg}, depending on which is the chosen figure of merit. In the next chapter we will make use of this mapping to demonstrate an interesting quantum advantage activation phenomenon in QRACs.

Additionally to RACs, SR-RACs and QRACs, one could also investigate SR-QRACs [82], EA-RACs [58] and EA-QRACs, where "EA" stands for "entanglement-assisted". Each of these cases (which I'll herein collectively refer to as simply "RACs") could analogously be mapped to a PM scenario, and the optimal solutions would be searched for inside the PM behavior sets $\mathcal{C}$, $\mathcal{C}^\lambda$, $\mathcal{Q}$, $\mathcal{Q}^\lambda$, $\mathcal{C}^\rho$ and $\mathcal{Q}^\rho$, respectively, all with subscript d, d^n, n.

To end this section I note that while RACs can be cast as an instance of prepare-and-measure, the inverse is not true. Investigating what different kinds of *information retrieval tasks* [88] arise from such other instances could be an interesting

research problem.

3.2.2 Dense coding

Holevo's bound guarantees that n qubits can perfectly encode no more than n bits of information [89]. Calling to mind that n qubits require $2^n - 1$ complex coefficients to be fully described, this result comes to be tremendously surprising. Looking more closely, the setting where this conclusion arises from is a prepare-and-measure scenario with quantum communication but independent devices. Another seminal result proves that when the devices are not independent, but rather, share a maximally entangled state as a resource, it becomes possible to communicate two bits by sending a single qubit.

Dense coding was first proposed by Bennett and Wiesner [5] similarly to the following argument. Let Alice and Bob share a two-qubit maximally entangled state $|\Phi^+\rangle = \frac{1}{\sqrt{2}} \sum_{i=0}^{1} |ii\rangle$ (the same argument is valid for any unitary transformation of it). Her task is to communicate one out of four messages to Bob. We label those messages with choices from the ordered set $\mathcal{X} = (0, 1, 2, 3)$, which could be perfectly encoded with two bits. They previously agree on some special set of transformations, hereby $\Lambda = (\mathbf{1}, \sigma_x, \sigma_y, \sigma_z)$ — where σ_i are the Pauli matrices —, to represent the respective encodings. Alice's encoding of x is a simple matter of applying Λ_x to her share of $|\Phi^+\rangle$. After receiving her qubit, Bob will have the state $(\Lambda_x \otimes \mathbf{1}) |\Phi^+\rangle\langle\Phi^+|$. The four possibilities are mutually orthogonal, thus can be perfectly distinguished by some suitable measurement (e.g., a standard Bell-basis measurement). He will hence be able to recover x from a single communicated qubit.

Purposefully, I have framed the protocol in the previously introduced notation for a prepare-and-measure scenario with quantum communication and entanglement assistance. My intention was to suggest that dense coding can be implemented as a special instance of prepare-and-measure scenarios, as it indeed can. More than that, several generalizations of this protocol are possible. The most straightforward ones are to allow for higher dimensional communication and entanglement [5] or mixed-state entanglement [90]. More recently, discussions have been opened in regard to unbounded entanglement [57] or even dense coding protocols with errors [91, 59]. Chapter 5 will detailedly discuss how we can use insights from prepare-and-measure scenarios to study the latter case. **[Comment:** I plan to do the formal definition on chap. 5 because even if I did it here I'd probably have to repeat it there. That's why I didnt make it already but I can put it here as well if it's better...]

Chapter 4

Classicality in the prepare and measure scenario

Chapter 5

Dense coding in the prepare and measure scenario

5.1 Optimizing preparations and measurements for dense coding

Appendices

Appendix A

Proofs for chap. 4

Appendix B

Proofs for chap. 5

Bibliography

[1] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, May 1935.

[2] E. Schrödinger. Discussion of probability relations between separated systems. *Mathematical Proceedings of the Cambridge Philosophical Society*, 31(4):555–563, 1935.

[3] J. S. Bell. On the einstein podolsky rosen paradox. *Physics Physique Fizika*, 1:195–200, Nov 1964.

[4] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, Jun 2009.

[5] Charles H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on einstein-podolsky-rosen states. *Phys. Rev. Lett.*, 69:2881–2884, Nov 1992.

[6] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. *Theoretical Computer Science*, 560:7–11, 2014. Theoretical Aspects of Quantum Cryptography – celebrating 30 years of BB84.

[7] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels. *Phys. Rev. Lett.*, 70:1895–1899, Mar 1993.

[8] Lane P. Hughston, Richard Jozsa, and William K. Wootters. A complete classification of quantum ensembles having a given density matrix. *Physics Letters A*, 183(1):14–18, 1993.

[9] Marcelo O. Terra Cunha, Jacob A Dunningham, and Vlatko Vedral. Entanglement in single-particle systems. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 463(2085):2277–2286, 2007.

[10] Reinhard F. Werner. Quantum states with einstein-podolsky-rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40:4277–4281, Oct 1989.

[11] Martin B. Plenio and Shashank Virmani. An introduction to entanglement measures. *Quantum Inf. Comput.*, 7(1):1–51, 2007.

[12] Marcelo O. Terra Cunha. Emaranhamento: caracterização, manipulação e consequências. *Universidade Federal de Minas Gerais*, 2005.

[13] Dagmar Bruß. Characterizing entanglement. *Journal of Mathematical Physics*, 43(9):4237–4251, 2002.

[14] M. A. Nielsen. Conditions for a class of entanglement transformations. *Phys. Rev. Lett.*, 83:436–439, Jul 1999.

[15] Barbara M. Terhal and Paweł Horodecki. Schmidt number for density matrices. *Phys. Rev. A*, 61:040301, Mar 2000.

[16] Asher Peres. Separability criterion for density matrices. *Phys. Rev. Lett.*, 77:1413–1415, Aug 1996.

[17] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Separability of mixed states: necessary and sufficient conditions. *Physics Letters A*, 223(1):1–8, 1996.

[18] Antonio Acín, Nicolas Gisin, and Benjamin Toner. Grothendieck's constant and local models for noisy entangled quantum states. *Phys. Rev. A*, 73:062105, Jun 2006.

[19] T. Vértesi. More efficient bell inequalities for werner states. *Phys. Rev. A*, 78:032112, Sep 2008.

[20] Flavien Hirsch, Marco Túlio Quintino, Tamás Vértesi, Miguel Navascués, and Nicolas Brunner. Better local hidden variable models for two-qubit Werner states and an upper bound on the Grothendieck constant $K_G(3)$. *Quantum*, 1:3, April 2017.

[21] Jonathan Barrett. Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a bell inequality. *Phys. Rev. A*, 65:042302, Mar 2002.

[22] H. M. Wiseman, S. J. Jones, and A. C. Doherty. Steering, entanglement, nonlocality, and the einstein-podolsky-rosen paradox. *Phys. Rev. Lett.*, 98:140402, Apr 2007.

[23] Michał Horodecki and Paweł Horodecki. Reduction criterion of separability and limits for a class of distillation protocols. *Phys. Rev. A*, 59:4206–4216, Jun 1999.

[24] Marco Túlio Quintino, Tamás Vértesi, Daniel Cavalcanti, Remigiusz Augusiak, Maciej Demianowicz, Antonio Acín, and Nicolas Brunner. Inequivalence of entanglement, steering, and bell nonlocality for general measurements. *Phys. Rev. A*, 92:032107, Sep 2015.

[25] Mark M. Wilde. *Quantum Information Theory.* Cambridge University Press, 2013.

[26] Christopher J. Wood, Jacob D. Biamonte, and David G. Cory. Tensor networks and graphical calculus for open quantum systems, 2015.

[27] A. Jamiołkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Reports on Mathematical Physics*, 3(4):275–278, 1972.

[28] Min Jiang, Shunlong Luo, and Shuangshuang Fu. Channel-state duality. *Phys. Rev. A*, 87:022310, Feb 2013.

[29] Roope Uola, Ana C. S. Costa, H. Chau Nguyen, and Otfried Gühne. Quantum steering. *Rev. Mod. Phys.*, 92:015001, Mar 2020.

[30] D Cavalcanti and P Skrzypczyk. Quantum steering: a review with focus on semidefinite programming. *Reports on Progress in Physics*, 80(2):024001, dec 2016.

[31] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell nonlocality. *Rev. Mod. Phys.*, 86:419–478, Apr 2014.

[32] Marco Túlio Quintino, Joseph Bowles, Flavien Hirsch, and Nicolas Brunner. Incompatible quantum measurements admitting a local-hidden-variable model. *Phys. Rev. A*, 93:052115, May 2016.

[33] Flavien Hirsch, Marco Túlio Quintino, and Nicolas Brunner. Quantum measurement incompatibility does not imply bell nonlocality. *Phys. Rev. A*, 97:012129, Jan 2018.

[34] Erika Bene and Tamás Vértesi. Measurement incompatibility does not give rise to bell violation in general. *New Journal of Physics*, 20(1):013021, jan 2018.

[35] Teiko Heinosaari, Takayuki Miyadera, and Mário Ziman. An invitation to quantum incompatibility. *Journal of Physics A: Mathematical and Theoretical*, 49(12):123001, feb 2016.

[36] Leonardo Guerini. Simulating quantum measurements and quantum correlations. *Universidade Federal de Minas Gerais*, 2018.

[37] Leonardo Guerini, Jessica Bavaresco, Marcelo Terra Cunha, and Antonio Acín. Operational framework for quantum measurement simulability. *Journal of Mathematical Physics*, 58(9):092102, 2017.

[38] Erkka Haapasalo, Teiko Heinosaari, and Juha-Pekka Pellonpää. Quantum measurements on finite dimensional systems: relabeling and mixing. *Quantum Information Processing*, 11(6):1751–1763, 2012.

[39] Michał Oszmaniec, Leonardo Guerini, Peter Wittek, and Antonio Acín. Simulating positive-operator-valued measures with projective measurements. *Phys. Rev. Lett.*, 119:190501, Nov 2017.

[40] Reinhold A Bertlmann and Philipp Krammer. Bloch vectors for qudits. *Journal of Physics A: Mathematical and Theoretical*, 41(23):235303, may 2008.

[41] Gen Kimura. The bloch vector for n-level systems. *Physics Letters A*, 314(5):339–349, 2003.

[42] Claudio Carmeli, Teiko Heinosaari, and Alessandro Toigo. Quantum random access codes and incompatibility of measurements. *EPL (Europhysics Letters)*, 130(5), June 2020.

[43] Nicolas Brunner, Miguel Navascués, and Tamás Vértesi. Dimension Witnesses and Quantum State Discrimination. *Physical Review Letters*, 110(15), April 2013.

[44] Joseph Bowles, Nicolas Brunner, and Marcin Pawłowski. Testing dimension and nonclassicality in communication networks. *Physical Review A*, 92(2), August 2015.

[45] Davide Poderini, Samuraí Brito, Ranieri Nery, Fabio Sciarrino, and Rafael Chaves. Criteria for nonclassicality in the prepare-and-measure scenario. *Physical Review Research*, 2(4), October 2020.

[46] Marcin Pawłowski, Tomasz Paterek, Dagomir Kaszlikowski, Valerio Scarani, Andreas Winter, and Marek Żukowski. Information causality as a physical principle. *Nature*, 461(7267):1101–1104, October 2009.

[47] Marcin Pawłowski and Valerio Scarani. Information Causality. *arXiv:1112.1142 [quant-ph]*, December 2011. arXiv: 1112.1142.

[48] R. W. Spekkens. Contextuality for preparations, transformations, and unsharp measurements. *Phys. Rev. A*, 71:052108, May 2005.

[49] Nicholas Harrigan, Terry Rudolph, and Scott Aaronson. Representing probabilistic data via ontological models, 2008.

[50] Ernesto F. Galvão. Economical ontological models for discrete quantum systems. *Phys. Rev. A*, 80:022106, Aug 2009.

[51] Julio I. de Vicente. A general bound for the dimension of quantum behaviours in the prepare-and-measure scenario. *Journal of Physics A: Mathematical and Theoretical*, 52(9):095304, February 2019. Publisher: IOP Publishing.

[52] Julio I. de Vicente. Shared randomness and device-independent dimension witnessing. *Physical Review A*, 95(1):012340, January 2017.

[53] Michele Dall'Arno, Elsa Passaro, Rodrigo Gallego, and Antonio Acin. Robustness of Device Independent Dimension Witnesses. *Physical Review A*, 86(4):042312, October 2012. arXiv: 1207.2574.

[54] Arthur Fine. Hidden variables, joint probability, and the bell inequalities. *Phys. Rev. Lett.*, 48:291–295, Feb 1982.

[55] Valerio Scarani. *Bell Nonlocality*. Oxford University Press, 1 edition, August 2019.

[56] Rodrigo Gallego, Nicolas Brunner, Christopher Hadley, and Antonio Acín. Device-Independent Tests of Classical and Quantum Dimensions. *Physical Review Letters*, 105(23):230501, November 2010.

[57] Armin Tavakoli, Jef Pauwels, Erik Woodhead, and Stefano Pironio. Correlations in entanglement-assisted prepare-and-measure scenarios. *arXiv:2103.10748 [quant-ph]*, March 2021. arXiv: 2103.10748.

[58] Marcin Pawłowski and Marek Żukowski. Entanglement-assisted random access codes. *Physical Review A*, 81(4):042326, April 2010.

[59] George Moreno, Ranieri Nery, Carlos de Gois, Rafael Rabelo, and Rafael Chaves. Semi-device-independent certification of entanglement in superdense coding. *arXiv:2102.02709 [quant-ph]*, February 2021. arXiv: 2102.02709.

[60] Leonardo Guerini, Marco Túlio Quintino, and Leandro Aolita. Distributed sampling, quantum communication witnesses, and measurement incompatibility. *Physical Review A*, 100(4):042308, 2019. arXiv: 1904.08435.

[61] Jamie Sikora, Antonios Varvitsiotis, and Zhaohui Wei. Device-independent dimension tests in the prepare-and-measure scenario. *Physical Review A*, 94(4), October 2016. Publisher: American Physical Society.

[62] Joseph Bowles, Marco Túlio Quintino, and Nicolas Brunner. Certifying the Dimension of Classical and Quantum Systems in a Prepare-and-Measure Scenario with Independent Devices. *Physical Review Letters*, 112(14), April 2014.

[63] Stephanie Wehner, Matthias Christandl, and Andrew C. Doherty. A lower bound on the dimension of a quantum system given measured data. *Physical Review A*, 78(6), December 2008. arXiv: 0808.3960.

[64] Martin Hendrych, Rodrigo Gallego, Michal Mičuda, Nicolas Brunner, Antonio Acín, and Juan P. Torres. Experimental estimation of the dimension of classical and quantum systems. *Nature Physics*, 8(8), August 2012.

[65] Johan Ahrens, Piotr Badziag, Adán Cabello, and Mohamed Bourennane. Experimental device-independent tests of classical and quantum dimensions. *Nature Physics*, 8(8):592–595, 2012.

[66] Vincenzo D'Ambrosio, Fabrizio Bisesto, Fabio Sciarrino, Johanna F. Barra, Gustavo Lima, and Adán Cabello. Device-Independent Certification of High-Dimensional Quantum Systems. *Physical Review Letters*, 112(14):140503, April 2014.

[67] Armin Tavakoli, Jędrzej Kaniewski, Tamás Vértesi, Denis Rosset, and Nicolas Brunner. Self-testing quantum states and measurements in the prepare-and-measure scenario. *Physical Review A*, 98(6), December 2018.

[68] Máté Farkas and Jędrzej Kaniewski. Self-testing mutually unbiased bases in the prepare-and-measure scenario. *Physical Review A*, 99(3), March 2019.

[69] Shi-Hui Wei, Fen-Zhuo Guo, Xin-Hui Li, and Qiao-Yan Wen. Robustness self-testing of states and measurements in the prepare-and-measure scenario with \textdollar3\textbackslashto 1\textdollar random access code. *Chinese Physics B*, 28(7), July 2019. Publisher: IOP Publishing.

[70] Piotr Mironowicz and Marcin Pawłowski. Experimentally feasible semi-device-independent certification of four-outcome positive-operator-valued measurements. *Phys. Rev. A*, 100:030301, Sep 2019.

[71] Armin Tavakoli, Massimiliano Smania, Tamás Vértesi, Nicolas Brunner, and Mohamed Bourennane. Self-testing nonprojective quantum measurements in prepare-and-measure experiments. *Science Advances*, 6(16), April 2020.

[72] Nikolai Miklin and Michał Oszmaniec. A universal scheme for robust self-testing in the prepare-and-measure scenario. *Quantum*, 5:424, Apr 2021.

[73] Miguel Navascués and Tamás Vértesi. Bounding the Set of Finite Dimensional Quantum Correlations. *Physical Review Letters*, 115(2), July 2015.

[74] Miguel Navascués, Adrien Feix, Mateus Araújo, and Tamás Vértesi. Characterizing finite-dimensional quantum behavior. *Physical Review A*, 92(4), October 2015.

[75] Karthik Mohan, Armin Tavakoli, and Nicolas Brunner. Sequential random access codes and self-testing of quantum measurement instruments. *New Journal of Physics*, 21(8):083034, August 2019.

[76] Nikolai Miklin, Jakub J. Borkała, and Marcin Pawłowski. Semi-device-independent self-testing of unsharp measurements. *Physical Review Research*, 2(3):033014, July 2020.

[77] Yukun Wang, Ignatius William Primaatmaja, Emilien Lavie, Antonios Varvitsiotis, and Charles Ci Wen Lim. Characterising the correlations of prepare-and-measure quantum networks. *npj Quantum Information*, 5(1), February 2019.

[78] Marcin Pawlowski and Nicolas Brunner. Semi-device-independent security of one-way quantum key distribution. *Physical Review A*, 84(1), July 2011. arXiv: 1103.4105.

[79] Christian Schmid, Pavel Trojek, Mohamed Bourennane, Christian Kurtsiefer, Marek Żukowski, and Harald Weinfurter. Experimental single qubit quantum secret sharing. *Phys. Rev. Lett.*, 95:230505, Dec 2005.

[80] Harry Buhrman, Richard Cleve, Serge Massar, and Ronald de Wolf. Nonlocality and communication complexity. *Reviews of Modern Physics*, 82(1):665–698, March 2010.

[81] Andris Ambainis, Ashwin Nayak, Ammon Ta-Shma, and Umesh Vazirani. Dense quantum coding and a lower bound for 1-way quantum automata. In *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing*, STOC '99, page 376–383, New York, NY, USA, 1999. Association for Computing Machinery.

[82] Andris Ambainis, Debbie Leung, Laura Mancinska, and Maris Ozols. Quantum Random Access Codes with Shared Randomness. *arXiv:0810.2937 [quant-ph]*, June 2009. arXiv: 0810.2937.

[83] Stephen Wiesner. Conjugate coding. *ACM Sigact News*, 15(1):78–88, 1983.

[84] Ashwin Nayak. Optimal lower bounds for quantum automata and random access codes. In *40th Annual Symposium on Foundations of Computer Science (Cat. No. 99CB37039)*, pages 369–376. IEEE, 1999.

[85] Andris Ambainis, Ashwin Nayak, Amnon Ta-Shma, and Umesh Vazirani. Dense quantum coding and quantum finite automata. *Journal of the ACM (JACM)*, 49(4):496–511, 2002.

[86] Masahito Hayashi, Kazuo Iwama, Harumichi Nishimura, Rudy Raymond, and Shigeru Yamashita. (4, 1)-quantum random access coding does not exist—one qubit is not enough to recover one of four bits. *New Journal of Physics*, 8(8):129, 2006.

[87] Robert W Spekkens, Daniel H Buzacott, Anthony J Keehn, Ben Toner, and Geoff J Pryde. Preparation contextuality powers parity-oblivious multiplexing. *Physical review letters*, 102(1):010401, 2009.

[88] Pierre-Emmanuel Emeriau, Mark Howard, and Shane Mansfield. Quantum advantage in information retrieval, 2020.

[89] Alexander Semenovich Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problemy Peredachi Informatsii*, 9(3):3–11, 1973.

[90] Adriano Barenco and Artur K. Ekert. Dense coding based on quantum entanglement. *Journal of Modern Optics*, 42(6):1253–1259, 1995.

[91] Ashwin Nayak and Henry Yuen. Rigidity of superdense coding, 2020.